one CMRA embodiment of this invention, the fiber optic array plate is etched completely through the entire width of the plate so that there is a series of open channels running from the top face of the plate through to the bottom face of the plate. A porous filter element is then be contacted to one face of such an etched fiber optic in order to form the CMRA.

In this embodiment, the microreactor array component is formed from a plate comprised of a fused fiber optic bundle. In such a fiber optic plate typically the distance between the top surface or face and the bottom surface or face is no greater than 5 cm, preferably no greater than 2 cm, and most preferably between 1 cm and 1 mm thick.

A series of microchannels extending from the top face to the botton face are created by treating the fiber optic plate, *e.g.*, with acid. Each channel can form a reaction chamber. (see *e.g.*, Walt, *et al.*, 1996. *Anal. Chem.* 70: 1888).

The CMRA array typically contains more than 1,000 reaction chambers, preferably more than 400,000, more preferably between 400,000 and 20,000,000, and most preferably between 1,000,000 and 16,000,000 cavities or reaction chambers. When a fiber optic plate is used as the microreactor element, the shape of each reaction chamber (from a top view) is frequently substantially hexagonal, but the reaction chambers may also be cylindrical. In some embodiments, each reaction chamber has a smooth wall surface, however, we contemplate that each reaction chamber may also have at least one irregular wall surface. The array is typically constructed to have reaction chambers with a center-to-center spacing between 5 to 200 μm, preferably between 10 to 150 μm, most preferably between 50 to 100 μm. In one embodiment, we contemplate that each reaction chamber has a width in at least one dimension of between 0.3 μm and 100 μm, preferably between 0.3 μm and 20 μm, most preferably between 0.3 μm and 10 μm. In a separate embodiment, we contemplate larger reaction chambers, preferably having a width in at least one dimension of between 20 μm and 70 μm.

UMRAs

Yet another technique for creating a membrane-based microreactor array eliminates the need for discrete microchannels or microwells altogether. This membrane reactor array is comprised simply of a porous filter against which molecules are concentrated by concentration polarization (or by which particles are packed by filtration) just as described above. In this instance, however, some of the concentration-polarized molecules may be made to form a continuous 2-D layer atop the rejecting layer of the

IDT Exhibit 1020.6045

porous membrane, while other reaction participants are deposited and/or otherwise immobilized or localized at discrete, independent sites within or atop the structure. Such membrane reactor arrays are referred to hereinafter as "unconfined membrane reactor arrays" or UMRAs.

Discrete reactions can be made to occur in discrete locations by "seeding" the surface of this membrane array with individual molecules or particles that initiate the reaction. For example, a catalyst could be added by pipetting or "spotting" small regions of the surface with a dilute solution thereof. Alternatively, a catalyst could be added in bulk solution at such low concentration such that, upon filtration onto the surface, the density of catalyst deposited in the plane of the UMRA would be reasonably high but discontinuous, *i.e.*, the catalyst might be "dotted" over the surface of the 2-D layer. In yet another embodiment, the catalyst (e.g., an enzyme) might be bound to a particulate or colloidal support (e.g., by covalent immobilization); a dilute suspension thereof would then be filtered through the UMRA, causing deposition of catalyst beads or particles at discrete sites on the surface.

The UMRA array typically contains more than 1,000 reaction chambers, preferably more than 400,000, more preferably between 400,000 and 20,000,000, and most preferably between 1,000,000 and 16,000,000 cavities or reaction chambers. When a fiber optic plate is used as the microreactor element, the shape of each reaction chamber (from a top view) is frequently substantially hexagonal, but the reaction chambers may also be cylindrical. In some embodiments, each reaction chamber has a smooth wall surface, however, we contemplate that each reaction chamber may also have at least one irregular wall surface. The array is typically constructed to have reaction chambers with a center-to-center spacing between 5 to 200 μm, preferably between 10 to 150 μm, most preferably between 50 to 100 μm. In one embodiment, we contemplate that each reaction chamber has a width in at least one dimension of between 0.3 μm and 100 μm, preferably between 0.3 μm and 20 μm, most preferably between 0.3 μm and 10 μm. In a separate embodiment, we contemplate larger reaction chambers, preferably having a width in at least one dimension of between 20 μm and 70 μm.

In still other instances, it may be advantageous to deposit a particular reactant molecule (e.g., an oligonucleotide or construct thereof) at discrete sites on the surface of a UMRA (e.g., for pyrosequencing). Again, this may be accomplished either by pipetting solutions thereof onto the surface of the UMRA, by ultrafiltering extremely dilute solutions of reactants through the UMRA, or, preferably, by immobilizing said reactants

IDT Exhibit 1020.6046

on particulate or colloidal supports and then depositing these onto the UMRA surface by ultrafiltration.

A distinguishing feature of the UMRA relative to the CMRA is that lateral diffusion of molecules in the UMRA is not confined by the walls of a microchannel or microvessel as is the case with the CMRA. To restrict the impact of lateral diffusion of molecules in a UMRA, molecules are swept through the UMRA and into the filtrate before their lateral transport can proceed to an extent such that it becomes problematic (Fig. 7). (Note that for the sake of simplicity, Fig. 7 shows only the rejecting surface or "skin layer" of the ultrafilter; the substrate region -- and other membrane supports, if present -- have been omitted for the sake of clarity.) Again, manipulation of the flow velocity (and species Peclet numbers) permits control of the extent of lateral transport as well as the residence times of molecules within the UMRA. The flow rate can be slowed, thus increasing residence time and the extent of lateral transport; or the flow rate can be increased, with a concomitant decrease in residence time and lateral transport. It should be noted that asymmetric ultrafiltration membranes of the "integral composite" type may be employed in a UMRA in either orientation – that is, with either their rejecting layer "up" (and spongy substrate region "down") or *vice versa*.

In addition to its having a porous ultrafiltration membrane component, a UMRA can also employ more porous and substantially non-rejecting filters (or functionally equivalent matrices) either upstream or downstream of the ultrafilter and either placed against and/or attached to the ultrafiltration membrane itself (Fig. 8). This non-selective secondary filter, characterized by its much larger pore sizes, can be used to provide mechanical support to the permselective ultrafiltration membrane. It can also be used as a mesh or matrix that provides some mechanical support and protection to the concentrated molecules at the membrane surface, thereby stabilizing this layer of molecules. In particular, the provision of such a mesh or matrix can shield the surface layer of concentrated macromolecules and particles from the shearing action of any tangential flow that may be directed along the upper surface of the UMRA. When asymmetric, integrally composite UF membranes of the type mentioned in the preceding paragraph are employed, their spongy and relatively thick substrate region may conveniently serve as a stabilizing matrix if the UF membrane is oriented with the skinned rejecting surface "down".

A UMRA can be assembled and operated substantially in the manner of a CMRA. Species Peclet numbers can be manipulated by controlling the flow rate, and different

IDT Exhibit 1020.6047

macromolecules and/or particles can be packed in sequence to create a stacked microreactor. Molecules can be added at concentrations below their K_d or C_g values, and then concentrated above their K_d or C_g values by concentration polarization at the filter surface. Molecules can be attached to other molecules, to polymers, or to particles. Molecules can also be enmeshed within polymers.

Many different types of reactions can be performed in a CMRA or UMRA. In one embodiment, each cavity or reaction chamber of the array contains reagents for analyzing a nucleic acid or protein. Typically those reaction chambers that contain a nucleic acid (not all reaction chambers in the array are required to) contain only a single species of nucleic acid (i.e., a single sequence that is of interest). There may be a single copy of this species of nucleic acid in any particular reaction chamber, or they may be multiple copies. It is generally preferred that a reaction chamber contain at least 100 copies of a nucleic acid sequence, preferably at least 100,000 copies, and most preferably between 100,000 to 1,000,000 copies of the nucleic acid. In one embodiment the nucleic acid species is amplified to provide the desired number of copies using PCR, RCA, ligase chain reaction, other isothermal amplification, or other conventional means of nucleic acid amplification. In one embodimant, the nucleic acid is single stranded. In other embodiments the single stranded DNA is a concatamer with each copy covalently linked end to end.

The nucleic acid may be immobilized in the reaction chamber, either by attachment to the chamber itself or by attachment to a mobile solid support that is delivered to the chamber. A bioactive agent could be delivered to the array, by dispersing over the array a plurality of mobile solid supports, each mobile solid support having at least one reagent immobilized thereon, wherein the reagent is suitable for use in a nucleic acid sequencing reaction.

The array can also include a population of mobile solid supports disposed in the reaction chambers, each mobile solid support having one or more bioactive agents (such as a nucleic acid or a sequencing enzyme) attached thereto. The diameter of each mobile solid support can vary, we prefer the diameter of the mobile solid support to be between 0.01 to 0.1 times the width of each cavity. Not every reaction chamber need contain one or more mobile solid supports. There are three contemplated embodiments; one where at least 5% to 20% of of the reaction chambers can have a mobile solid support having at least one reagent immobilized thereon; a second embodiment where 20% to 60% of the reaction chambers can have a mobile solid support having at least one reagent immobilized thereon; and a third embodiment where 50% to 100% of the reaction

IDT Exhibit 1020.6048

chambers can have a mobile solid support having at least one reagent immobilized thereon.

The mobile solid support typically has at least one reagent immobilized thereon. For the embodiments relating to pyrosequencing reactions or more generally to ATP detection, the reagent may be a polypeptide with sulfurylase or luciferase activity, or both. Alternatively, enzymes such as hypoxanthine phosphoribosyltransferase, xanthine oxidase, uricase or peroxidase could be utilized (*e.g.*, Jansson and Jansson (2002), incorporated herein by reference). The mobile solid supports can be used in methods for dispersing over the array a plurality of mobile solid supports having one or more nucleic sequences or proteins or enzymes immobilized thereon.

In another aspect, the invention involves an apparatus for simultaneously monitoring the array of reaction chambers for light generation, indicating that a reaction is taking place at a particular site. In this embodiment, the reaction chambers are sensors, adapted to contain analytes and an enzymatic or fluorescent means for generating light in the reaction chambers. In this embodiment of the invention, the sensor is suitable for use in a biochemical or cell-based assay. The apparatus also includes an optically sensitive device arranged so that in use the light from a particular reaction chamber would impinge upon a particular predetermined region of the optically sensitive device, as well as means for determining the light level impinging upon each of the predetermined regions and means to record the variation of the light level with time for each of the reaction chamber.

In one specific embodiment, the instrument includes a light detection means having a light capture means and a second fiber optic bundle for transmitting light to the light detecting means. We contemplate one light capture means to be a CCD camera. The second fiber optic bundle is typically in optical contact with the array, such that light generated in an individual reaction chamber is captured by a separate fiber or groups of separate fibers of the second fiber optic bundle for transmission to the light capture means.

The invention provides an apparatus for simultaneously monitoring an array of reaction chambers for light indicating that a reaction is taking place at a particular site. The reaction event, *e.g.*, photons generated by luciferase, may be detected and quantified using a variety of detection apparatuses, *e.g.*, a photomultiplier tube, a CCD, CMOS, absorbance photometer, a luminometer, charge injection device (CID), or other solid state detector, as well as the apparatuses described herein. In a preferred embodiment, the quantitation of the emitted photons is accomplished by the use of a CCD camera fitted

IDT Exhibit 1020.6049

with a fused fiber optic bundle. In another preferred embodiment, the quantitation of the emitted photons is accomplished by the use of a CCD camera fitted with a microchannel plate intensifier. A back-thinned CCD can be used to increase sensitivity. CCD detectors are described in, *e.g.*, Bronks, *et al.*, 1995. *Anal. Chem.* 65: 2750-2757.

An exemplary CCD system is a Spectral Instruments, Inc. (Tucson, AZ) Series 600 4-port camera with a Lockheed-Martin LM485 CCD chip and a 1-1 fiber optic connector (bundle) with 6-8 μm individual fiber diameters. This system has 4096 x 4096, or greater than 16 million pixels and has a quantum efficiency ranging from 10% to > 40%. Thus, depending on wavelength, as much as 40% of the photons imaged onto the CCD sensor are converted to detectable electrons.

In other embodiments, a fluorescent moiety can be used as a label and the detection of a reaction event can be carried out using a confocal scanning microscope to scan the surface of an array with a laser or other techniques such as scanning near-field optical microscopy (SNOM) are available which are capable of smaller optical resolution, thereby allowing the use of "more dense" arrays. For example, using SNOM, individual polynucleotides may be distinguished when separated by a distance of less than 100 nm, *e.g.*, 10nm x 10nm. Additionally, scanning tunneling microscopy (Binning et al., *Helvetica Physica Acta*, 55:726-735, 1982) and atomic force microscopy (Hanswa et al., *Annu Rev Biophys Biomol Struct*, 23:115-139, 1994) can be used.

Manufacture of CMRAs and UMRAs and Uses Thereof

The invention provides a CMRA and UMRA which are both an array comprising densely packed, independent chemical reactions The reaction site of the CMRA is a microreactor vessel or a microwell. The invention also includes method for making the CMRA and UMRA dense array of discrete reaction sites. The invention also provides a method for charging a microreactor with reaction participants, the method comprising, effecting convective flow of fluid normal to the plane of and through the array of reaction sites. In a preferred embodiment of both the CMRA and UMRA, the reactants may be either attached or unattached to a solid support. The attached reactants are covalently bound to a solid support.

The invention also includes a method for efficiently supplying relatively lower molecular weight reagents and reactants to discrete reaction sites. The CMRA itself comprises a microreactor comprising a microchannel with an ultrafiltration membrane at

IDT Exhibit 1020.6050

one end. Specifically, the CMRA comprises a series of microchannels; a concentration polarization to create a packed column of molecules; a sequential packing of molecules via the concentration polarization to create stacked columns; and a flow of reagents through a packed column. In a preferred embodiment of the CMRA , each array is an independent chemical reactor.

The invention also provides a method of generating a CMRA, the method comprising the steps of: (a) flowing of reagents through a packed column/CMRA for sequential processing of chemicals; (b) adding random fragments of DNA, obtaining about one fragment per microchannel, by filtration of a mixture onto a CMRA; and (c) concentrating molecules into a gel inside the microchannels of a CMRA. Molecules are added to the CMRA below the Kd inside a microchannel, permitting polymerization and assembly of molecules only inside the microchannels. Crosslinkers can be added to decrease the diffusional mobility of molecules in a microchannel. Furthermore, polymers can also be added that enmesh molecules inside a microchannel, as well as smaller molecules attached to larger molecules or to larger particles or to beads to decrease the diffusional mobility of the smaller molecule. Smaller molecules, that would otherwise pass through the filter, attached to larger molecules or to larger particles or to beads to retain the smaller molecules in the microchannel of a CMRA can be added as well.

The CMRA is generated by using Anopore membranes. More specifically, the CMRA is fabricated by bonding an ultrafiltration membrane to microfabricated array of microchannels. Fluid inlets are then radially distributed to equalize pressure across the membrane, reducing variation pressure over the CMRA. An ultrafiltration membrane, without the microchannels, is then used to create a dense 2-D array of chemical reactions ("unconfined membrane reactor array" or UMRA) in which reactions are seeded by filtering a catalyst or reactant or enzyme onto the filter surface and whereby convective flow washes away laterally diffusing molecules before they contaminate adjacent reactions. Concentration polarization is necessary to create the packed columns of molecules for the CMRA and UMRA followed by the sequential packing of molecules via concentration polarization to create stacked columns. Reagents are then flowed through a packed column of the CMRA or UMRA for sequential processing of chemicals. The ultrafiltration membrane is then bonded to a second, more porous membrane to provide mechanical support to the molecules concentrated by concentration polarization. The membrane is a Molecular/Por membrane (Spectrum Labs).

IDT Exhibit 1020.6051

The CMRA and UMRA have a multitude of uses including: PCR, as well as other DNA amplification techniques and DNA sequencing techniques, such as pyrosequencing. Both the CMRA and UMRA can be utilized to achieve highly parallel sequencing without separation of DNA fragments and associated sample prep. The CMRA and UMRA can also be used for combinatorial chemistry. For detection purposes, an array of photodetectors are utilized for monitoring light producing reactions within the CMRA or UMRA. In a preferred embodiment, the array of photodetectors is a CCD camera. Another method of detection of discrete reactions withinn the CMRA and UMRA is to monitor changes in light absorption as an indicator of a chemical reaction in a CMRA using an array of photodetectors.

Two examples are offered below as specifically contemplated uses for CMRAs and UMRAs, but these are meant only to be representative and should not be considered as the only applications or embodiments of the present invention.

Sequencing of DNA via Pyrophosphate Detection

The methods and apparatuses described are generally useful for any application in which the identification of any particular nucleic acid sequence is desired. For example, the methods allow for identification of single nucleotide polymorphisms (SNPs), haplotypes involving multiple SNPs or other polymorphisms on a single chromosome, and transcript profiling. Other uses include sequencing of artificial DNA constructs to confirm or elicit their primary sequence, or to identify specific mutant clones from random mutagenesis screens, as well as to obtain the sequence of cDNA from single cells, whole tissues or organisms from any developmental stage or environmental circumstance in order to determine the gene expression profile from that specimen. In addition, the methods allow for the sequencing of PCR products and/or cloned DNA fragments of any size isolated from any source.

Sequencing of DNA by pyrophosphate detection ("pyrophosphate sequencing") is described in various patents (Hyman, 1990, US Patent Number 4,971,903; Nyren *et al.*, US Patents 6,210,891 and 6,258,568 and PCT Patent application WO98/13523; Hagerlid *et al.*, 1999, WO99/66313; Rothberg, US Patent 6,274,320, WO01/20039) and publications (Hyman, 1988; Nyrén *et al.*, 1993; Ronaghi *et al.*, 1998, Jensen, 2002; Schuller, 2002). The contents of the foregoing patents, patent applications and publications cited here are incorporated herein by reference in their entireties. Pyrophosphate sequencing is a technique in which a complementary sequence is

IDT Exhibit 1020.6052

polymerized using an unknown sequence (the sequence to be determined) as the template. This is, thus, a type of sequencing technique known as "sequencing by synthesis". Each time a new nucleotide is polymerized onto the growing complementary strand, a pyrophosphate (PPi) molecule is released. This release of pyrophosphate is then detected. Iterative addition of the four nucleotides (dATP, dCTP, dGTP, dTTP) or of analogs thereof (e.g., α-thio-dATP), accompanied by monitoring of the time and extent of pyrophosphate release, permits identification of the nucleotide that is incorporated into the growing complementary strand.

Pyrophosphate can be detected via a coupled reaction in which pyrophosphate is used to generate ATP from adenosine 5'-phosphosulfate (APS) through the action of the enzyme ATP sulfurylase (Fig.9). The ATP is then detected photometrically via light released by the enzyme luciferase, for which ATP is a substrate. (It may be noted that dATP is added as one of the four nucleotides for sequencing by synthesis and that luciferase can use dATP as a substrate. To prevent light emission on addition of dATP for sequencing, a dATP analog such as α-thio-dATP is substituted for dATP as the nucleotide for sequencing. The α-thio-dATP molecule is incorporated into the growing DNA strand, but it is not a substrate for luciferase.)

Pyrophosphate sequencing can be performed in a CMRA or UMRA in several different ways. One such protocol follows:

(1) pack luciferase;

(2) pack ATP sulfurylase;

(3) pack the DNA whose sequence is to be determined (preferably, many copies of a single sequence) and DNA polymerase (*e.g.* Klenow fragment); and

(4) flow a mixture of dXTP, APS, and luciferin through the CMRA or UMRA, cycling through the four nucleotides (dCTP, dGTP, dTTP, α-thio-dATP) one at a time. It will be noted that these are all low-molecular weight molecules, so they will pass through the ultrafiltration membrane of the CMRA or UMRA (at least if the ultrafilter's MWCO is appropriately chosen) without said molecules undergoing appreciable concentration polarization.

The upstream-to-downstream flow of fluid into and through the CMRA or UMRA thus causes:

IDT Exhibit 1020.6053

(a) addition of the appropriate dXTP by the polymerase and attendant production of PP_i in the region of the DNA being sequenced (with APS and luciferin flowing through passively);

(b) production of ATP from APS and PP_i when the latter are brought into contact with the sulfurylase enzyme (with luciferin flowing through passively); and

(c) production of light from ATP and luciferin in the vicinity of the luciferase enzyme.

Light production is then monitored by a photodetector. For example, a CCD camera, optically coupled by a lens or other means to the CMRA or UMRA, is capable of monitoring light production simultaneously from many microchannels or discrete reaction sites (Fig. 10). CCD cameras are available with millions of pixels, or photodetectors, arranged in a 2-D array. Light originating from one microchannel, microwell, or discrete reaction site in or on a CMRA or UMRA can be made to strike one or a few pixels on the CCD. Thus, if each microchannel, microwell, or reaction site is arranged to contain and conduct an independent sequencing reaction, each reaction can be monitored by one (or at most a few) CCD elements or photodetectors. By using a CCD camera or other imaging means comprising millions of pixels, the progress of millions of independent sequencing reactions is simultaneously monitored.

Further, each microreactor vessel, well, or reaction site can be made to hold the amplification products from only a single strand of DNA, and if different wells hold the amplification products of different strands of DNA, then the simultaneous sequencing of millions of different strands of DNA is possible. The distribution of DNA to be sequenced can be accomplished in many ways, two of which follow:

(a) The amplification products of a single oligonucleotide strand are attached to a bead, and beads from many independent amplification reactions are combined and placed onto a CMRA or UMRA; or

(b) Many different strands of DNA are added in dilute concentration and applied to the CMRA or URMA such that many if not most microchannels, microvessels, or discrete reaction sites contain only a single strand of DNA. The DNA is then amplified within or upon the CMRA or URMA through one series of reactions, and then it is directly sequenced via addition of the reagents described above. One such technique (polymerase chain reaction or PCR) for amplification of DNA within the microchannels of a CMRA (or UMRA) is described below.

IDT Exhibit 1020.6054

Delivery of the DNA to be sequenced and the enzymes and substrates necessary for pyrophosphate-based sequencing can be accomplished in a number of ways.

In a preferred embodiment, one or more reagents are delivered to the CMRA or UMRA immobilized or attached to a population of mobile solid supports, *e.g.*, a bead or microsphere. The bead or microsphere need not be spherical, irregular shaped beads may be used. They are typically constructed from numerous substances, *e.g.*, plastic, glass or ceramic and bead sizes ranging from nanometers to millimeters depending on the width of the reaction chamber. Preferably, the diameter of each mobile solid support can be between 0.01 and 0.1 times the width of each reaction chamber. Various bead chemistries can be used *e.g.*, methylstyrene, polystyrene, acrylic polymer, latex, paramagnetic, thoria sol, carbon graphite and titanium dioxide. The construction or chemistry of the bead can be chosen to facilitate the attachment of the desired reagent.

In another embodiment, the bioactive agents are synthesized first, and then covalently attached to the beads. As is appreciated by someone skilled in the art, this will be done depending on the composition of the bioactive agents and the beads. The functionalization of solid support surfaces such as certain polymers with chemically reactive groups such as thiols, amines, carboxyls, etc. is generally known in the art. Accordingly, "blank" beads may be used that have surface chemistries that facilitate the attachment of the desired functionality by the user. Additional examples of these surface chemistries for blank beads include, but are not limited to, amino groups including aliphatic and aromatic amines, carboxylic acids, aldehydes, amides, chloromethyl groups, hydrazide, hydroxyl groups, sulfonates and sulfates.

These functional groups can be used to add any number of different candidate agents to the beads, generally using known chemistries. For example, candidate agents containing carbohydrates may be attached to an amino-functionalized support; the aldehyde of the carbohydrate is made using standard techniques, and then the aldehyde is reacted with an amino group on the surface. In an alternative embodiment, a sulfhydryl linker may be used. There are a number of sulfhydryl reactive linkers known in the art such as SPDP, maleimides, α-haloacetyls, and pyridyl disulfides (see for example the 1994 Pierce Chemical Company catalog, technical section on cross-linkers, pages 155-200, incorporated here by reference) which can be used to attach cysteine containing proteinaceous agents to the support. Alternatively, an amino group on the candidate agent may be used for attachment to an amino group on the surface. For example, a large

IDT Exhibit 1020.6055

number of stable bifunctional groups are well known in the art, including homobifunctional and heterobifunctional linkers (see Pierce Catalog and Handbook, pages 155-200).

In an additional embodiment, carboxyl groups (either from the surface or from the candidate agent) may be derivatized using well known linkers (see Pierce catalog). For example, carbodiimides activate carboxyl groups for attack by good nucleophiles such as amines (see Torchilin et al., *Critical Rev. Thereapeutic Drug Carrier Systems*, 7(4):275-308 (1991)). Proteinaceous candidate agents may also be attached using other techniques known in the art, for example for the attachment of antibodies to polymers; see Slinkin et al., *Bioconj. Chem.* 2:342-348 (1991); Torchilin et al., *supra*; Trubetskoy et al., *Bioconj. Chem.* 3:323-327 (1992); King et al., *Cancer Res.* 54:6176-6185 (1994); and Wilbur et al., *Bioconjugate Chem.* 5:220-235 (1994). It should be understood that the candidate agents may be attached in a variety of ways, including those listed above. Preferably, the manner of attachment does not significantly alter the functionality of the candidate agent; that is, the candidate agent should be attached in such a flexible manner as to allow its interaction with a target.

Specific techniques for immobilizing enzymes on beads are known in the prior art. In one case, NH_2 surface chemistry beads are used. Surface activation is achieved with a 2.5% glutaraldehyde in phosphate buffered saline (10 mM) providing a pH of 6.9 (138 mM NaCl, 2.7 mM KCl). This mixture is stirred on a stir bed for approximately 2 hours at room temperature. The beads are then rinsed with ultrapure water plus 0.01% Tween 20 (surfactant) –0.02%, and rinsed again with a pH 7.7 PBS plus 0.01% tween 20. Finally, the enzyme is added to the solution, preferably after being prefiltered using a 0.45 μm amicon micropure filter.

In some embodiments, the reagent immobilized to the mobile solid support can be a polypeptide with sulfurylase activity, a polypeptide with luciferase activity or a chimeric polypeptide having both sulfurylase and luciferase activity. In one embodiment, it can be a ATP sulfurylase and luciferase fusion protein. Since the product of the sulfurylase reaction is consumed by luciferase, proximity between these two enzymes may be achieved by covalently linking the two enzymes in the form of a fusion protein. In other embodiments, the reagent immobilized to the mobile solid support can be the nucleic acid whose sequence is to be determined or analyzed.

IDT Exhibit 1020.6056

Many variations and alternative embodiments of the present invention as applied to DNA sequencing and other applications will be readily apparent and are considered to be within the scope of the present invention.

Polymerase Chain Reaction (PCR)

PCR can also be performed in an CMRA (or a UMRA). Participants in a PCR reaction include template DNA, primers, polymerase, and deoxynucleotides. Diffusion coefficients for some of these molecules follow:

Molecule	D (10^{-5} cm^2/s)
Taq Polymerase	0.06*
DNA (1100mer, duplex)	0.007**

* assumed to be slightly less than ovalbumin, which has a lower molecular weight (45kD for ovalbumin *cf.* 90kD for Taq Polymerase); D for ovalbumin is $0.08 \cdot 10^{-5}$ (Cussler, 1997).

** Liu *et al.* 2000

A DNA sequence to be amplified is packed into the CMRA, with as few as a single copy of a sequence per microchannel or microwell. Polymerase and primer are packed next, and nucleotides are added via flow. Thermal cycling then proceeds with heating of the CMRA by IR or other means of thermal control. Continuous flow of dXTP solution (and in some instances primer) through the CMRA ensures the retention of amplification products within the microchannels or microwells of the CMRA as the solution of dXTP (and perhaps primer) is continually added. The result is an independent array of PCR reactions conducted in independent CMRA microchannels or microwells, with each amplifying different DNA sequences.

It may be noted that other DNA amplification techniques can be performed in a similar manner. Such alternative techniques include bridge amplification onto a bead, rolling circle amplification (RCA) to form linear oligonucleotide concatamers, and hyperbranching amplification.

IDT Exhibit 1020.6057

EXAMPLES

EXAMPLE 1: Pyrophosphate-Based Sequencing in a UMRA

Materials

Reagents. Sepharose beads are 30±10 um and can bind $1x10^9$ biotin molecules per bead (very high binding capacity). Sequences of Oligonucleotides used in PCR on the membrane; Cy3-labelled probe J (5'-[Cy3]ATCTCTGCCTACTAACCATGAAG-3') (SEQ ID NO:1), Biotinyalted probe (5'-RBiot(dT18) GTTTCTCTCCAGCCTCTCACCGA-3') (SEQ ID NO:2), SsDNA template (5'-ATC TCT GCC TAC TAA CCA TGA AGA CAT GGT TGA CAC AGT GGA ATT TTA TTA TCT TAT CAC TCA GGA GAC TGA GAC AGG ATT GTC ATA AGT TTG AGA CTA GGT CGG TGA GAG GCT GGA GAG AAA C-3') (SEQ ID NO:3), and Non-Biotinylated probe (5'-GTTTCTCTCCAGCCTCTCACCGA-3') (SEQ ID NO:4), **Seq1** 5'-ACG TAA AAC CCC CCC CAA AAG CCC AAC CAC GTA CGT AAG CTG CAG CCA TCG TGT GAG GTC-3' (SEQ ID NO:5), PRB1 5'-BS-GAC CTC ACA CGA TGG CTG CAG CTT-3' (SEQ ID NO:6)

Preparation of Beads. Conjugation of biotinylated single stranded DNA probe to the streptavidin-bound Sepharose beads (Pharmacia Biotech, Uppsala, Sweden) was performed using the following protocol: 100 µl of Sepharose beads (1×10^8 beads ml^{-1}) and 300 µl of binding wash buffer (as recommended by the manufacturer) were applied to a Microcon 100 (Amicon, Beverly, MA) membrane. The tube was spun at 5000 rpm for 6 min in a micro centrifuge. The beads were washed twice with 300 µl of binding wash buffer. The tube was inverted into a new tube and the beads were spun out of the membrane (1 min at 6000 rpm). The beads were allowed to settle and the supernatant was removed. 100 µl (1 pmol μl^{-1}) of biotinylated probe solution was added to the beads. The final sample volume was about 100 µl. The tube was placed on a rotator for 1 hr at room temperature to allow the conjugation of the biotinylated probe with the beads. After conjugation, the beads were washed 3 times with TE buffer as described above. The final bead volume was about 100 µl and stored at –20 °C.

IDT Exhibit 1020.6058

Methods

Loading enzymes on to the first membrane. Substrate consists of 300 μm D-Luciferin (Pierce, Rockford, IL), 4 μM of APS (Adenosin 5'-phosphosulfate sodium salt, Sigma), 4 mg/ml PVP (Polyvinyle Pyrolidone, Sigma), and 1 mM of DTT (Dithiothreitol, Sigma). 10 μl of recombinant luciferase (14.7 mg ml^{-1}, from Sigma) and 75 μl of ATP sulfurylase (1.3 mg ml^{-1}, from Sigma) were mixed in 1 ml of substrate. This enzyme mixture was pipetted onto an ultra filtration membrane (MWCO 30,000; Millipore Incorporation, Bedford, MA)) wetted with substrate. Suction was applied to trap the enzyme mixture into the microstructure of the membrane. The enzymes adsorbed onto the membrane were found to be stable for 18 hr at room temperature.

Loading DNA Sepharose beads on to the second membrane. 100 μl of Sepharose beads (1 x 10^8 beads per ml) with bound DNA (3 x 10^6 copies per bead) was diluted in substrate and applied to a second Nylon membrane (Rancho Dominguez, CA). The beads are 30 ± 10 μm and the nylon membrane has a pore diameter 30 μm. Vacuum suction was applied to fix the beads onto the mesh of the membrane. The nylon membrane was then placed on top of the enzyme membrane. 80 μl of *Bst* DNA polymerase enzyme (8000 U μl^{-1}) was pipetted into the nylon membrane and the membrane was incubated for 30 min at room temperature. After incubation a glass window was placed on top of the membrane as shown in Figure 11. The membrane holder was connected to Fluidic 1.1 consisting of a multiposition valve (from Valco Instruments (Houston, TX) and two peristaltic pumps (Instech Laboratories Inc., Plymouth Meeting, PA).
During the substrate flow, the upper pump was running at a flow rate 0.5 ml min^{-1} for 2 min and the lower pump was running constantly at a flow rate of 50 μl min^{-1}. During the flow of pyrophosphate (ppi) or nucleotide, the flow rate of the upper pump was 0.1 ml min^{-1} for 2 min.

Imaging system. The imaging system consists of a CCD camera (Roper Scientific 2k x 2k with a pixel size of 24 μm) and two lenses (50 mm, f 1/1.2). One lens collects the light produced at the result of the interaction of the ppi with the enzyme mixture and the other lens focus the light into the CCD camera. The acquisition time for all of the experiments was 1 sec. A run-off, in which dNTP's (6.5 μM of dGTP, 6.5 μM of dCTP,

IDT Exhibit 1020.6059

6.5 μM of dTTP, and 50 μM of dATP-αs and for 2 min) were added to the DNA beads gave 5000 counts above background. Background was normally 160 counts.

Results

Sensitivity of convective sequencing. Figure 12 shows a pyrogram (a measurement of photons generated as a consequence of pyrophosphate produced) for the oligo seq1 immobilized onto the Sepharose beads. The number of oligo copies per bead was about 1000. The experimental conditions were the same as described in Figure 13. From Figure 14 it is estimated that the signal for the 0.1 uM (ppi) was about 300 counts. The signal for one base (nucleotide A) was ~38 counts at a copy number of 1000 DNAs per bead. This high sensitivity for convective sequencing is in part due to retention of soluble enzyme activity; since the enzyme mixtures are not immobilized to the beads but are physically trapped in the microstructure of the ultra filtration membrane. Also, the enzyme concentration on the membrane may be increased, resulting in enhanced sensitivity.

It is noted that while in this embodiment two membranes were used (in contact with one another), a single membrane would be equally satisfactory, and it would be routine for the ordinarily skilled artisan to trap the nucleic acids to be sequenced and the necessary sequencing enzymes and substrates on a single membrane.

Effect of immobilization of the luciferase and ATP sulfurylase on Sepharose beads. Sepharose beads are 30±10 μm and have binding capacity of 1 x 10^9 biotins per bead. Hence one bead can accommodate the immobilization of luciferase and ATP sulfurylase after the bead becomes loaded with 10^4 - 10^6 copies of nucleic acid template. The effect of immobilization of luciferase and atp sulfurylase on the Sepharose beads with DNA molecules has been studied. 0.5 μl of the oligo seq 1 was added to 100 μl of sepharose beads with prb1 primer. The mixture was placed in a PCR thermocyler and heated to 95 °C and allowed to cool to 4 °C at rate 0.1 °C s^{-1}. The excess oligo was removed by washing the beads twice with annealing buffer. a 100 μl mixture of luciferase and atp sulfurylase was added to 5 μl of beads and incubate with rotation at 4 °C for 1 hr. The pyrogram for the Sepharose beads with the enzyme mixture showed about 3.5 times more sensitivity than the beads without enzymes immobilized to them.

IDT Exhibit 1020.6060

EXAMPLE 2: Production of a UMRA

An ultrafiltration membrane is used to create a dense 2-D array of chemical reactions ("unconfined membrane reactor array" or UMRA) in which reactions are seeded by filtering a catalyst or reactant or enzyme onto the filter surface and whereby convective flow washes away laterally diffusing molecules before they contaminate adjacent reactions. Concentration polarization is necessary to create the packed columns of molecules for the UMRA followed by the sequential packing of molecules via concentration polarization to create stacked columns. Reagents are then flowed through a packed column of the UMRA for sequential processing of chemicals. The ultrafiltration membrane may then be bonded to a second, more porous membrane to provide mechanical support to the molecules concentrated by concentration polarization. The membrane is a Molecular/Por membrane (Spectrum Labs) or Anopore™ and Anodisc™ families of ultrafiltration membranes sold, for example, by Whatman PLC.

The substrate material is preferably made of a material that facilitates detection of the reaction event. For example, in a typical sequencing reaction, binding of a dNTP to a sample nucleic acid to be sequenced can be monitored by detection of photons generated by enzyme action on phosphate liberated in the sequencing reaction. Thus, having the substrate material made of a transparent or light conductive material facilitates detection of the photons. Reagents, such as enzymes and template, are delivered to the reaction site by a mobile solid support such as a bead.

The UMRA has handling properties similar to a nylon membrane. Reaction chambers are formed directly on the membrane, such that each reaction site is formed by the woven fibers of the membrane itself. Alternatively, a fiber optic bundle is utilized for the surface of the UMRA. The surface itself is cavitated by treating the termini of a bundle of fibers, *e.g.*, with acid, to form an indentation in the fiber optic material. Thus, cavities are formed from a fiber optic bundle, preferably cavities are formed by etching one end of the fiber optic bundle. Each cavitated surface can form a reaction chamber, or fiber optic reactor array (FORA). The indentation ranges in depth from approximately one-half the diameter of an individual optical fiber up to two to three times the diameter of the fiber. Cavities are introduced into the termini of the fibers by placing one side of the optical fiber wafer into an acid bath for a variable amount of time. The amount of time varies depending upon the overall depth of the reaction cavity desired (see *e.g.*, Walt, *et al.*, 1996. *Anal. Chem.* 70: 1888). The opposing side of the optical fiber wafer

IDT Exhibit 1020.6061

(*i.e.*, the non-etched side) is typically highly polished so as to allow optical-coupling (*e.g.*, by immersion oil or other optical coupling fluids) to a second, optical fiber bundle. This second optical fiber bundle exactly matches the diameter of the optical wafer containing the reaction chambers, and acts as a conduit for the transmission of light product to the attached detection device, such as a CCD imaging system or camera. The fiber optic wafer is then thoroughly cleaned, *e.g.* by serial washes in 15% H_2O_2/15%NH_4OH volume:volume in aqueous solution, followed by six deionized water rinses, then 0.5M EDTA, six deionized water washes, 15% H_2O_2/15%NH_4OH and six deionized water washes(one-half hour incubations in each wash).

References

Anderson JR, Chiu DT, Jackman RJ, Cherniavskaya O, McDonald JC, Wu H, Whitesides SH, and Whitesides GM. 2000. Fabrication of topologically complex three-dimensional microfluidic systems in PDMS by rapid prototyping. *Anal. Chem.* 72:3158-64.

Aoki A. 1992. Microwell cluster for processing electron microscope sections for immunocytochemistry. *Biotech. Histochem.* 67:98-9.

Asoh, H, Nishio, K, Nakao, M. Tamamura, T and Masuda, H. 2001. Conditions for fabrication of ideally ordered anodic porous alumina using pretextured aluminum. J. Electrochem. Soc. 148:B152-B156.

Bickerstaff GF. 1997. *Immobilization of Enzymes and Cells.* Totowa, New Jersey, USA: Humana Press.

Cherukuri SC, Demers RR, Fan ZHH, Levine AW, McBride SE, Zanzucchi PJ. 1999. Method and system for inhibiting cross-contamination in fluids of combinatorial chemistry device. US Patent Number 5,980,704.

Crank J. 1975. *The Mathematics of Diffusion.* 2nd ed. London: Oxford Univ. Press.

Cussler EL. 1997. *Diffusion: Mass Transfer in Fluid Systems.* 2nd ed. New York: Cambridge Univ. Press.

Dannoux T, Pujol G, and Root D. 2000. High-density test plate and process of making. US Patent number US6030829.

Deng T, Wu H, Brittain ST, and Whitesides GM. 2000. Prototyping of masks, masters, and stamps/molds for soft lithography using an office printer and photographic reduction. *Anal. Chem.* 72:3176-80.

IDT Exhibit 1020.6062

Ehrfeld, W, Hessel, V, and Löwe, H. 2000. *Microreactors: New Technology for Modern Chemistry.* New York: Wiley-VCH.

Eykamp, W. 1995. Microfiltration and ultrafiltration, pp. 1-43 in *Membrane Separation Technology: Principles and Applications*, RD Noble and SA Stern, eds. Amsterdam: Elsevier Science BV.

Ferguson JA, Boles C, Adams CP, and Walt DR. 1996. A fiber-optic DNA biosensor microarray for the analysis of gene expression. *Nature Biotech.* 14:1681-4.

Furneaux, RC, Rigby, WR, and Davidson, AP. 1989. The formation of controlled porosity membranes from anodically oxidized aluminum. *Nature*. 337:147-9.

Gamble RC, Theriault TP, Baldeschwieler JD. 1999. Apparatus for the chemical synthesis of molecular arrays. US Patent Number 5,981,733.Hughes TR, Mao, M, Jones, AR, Burchard J., Marton MJ, Shannon KW, Lefkowitz SM, Ziman M, Schelter JM, Meyer MR, Kobayashi S, Davis C, Dai HY, He YDD, Stephaniants SB, Cavet G, Walker WL, West A, Coffey E, Shoemaker DD, Stoughton R, Blanchard AP, Friend SH, and Linsley PS. 2001. Expression profiling using microarrays fabricated by an ink-jet oligonucleotide synthesizer. *Nat. Biotechmol.* 19(4):342-347.

Hyman ED. 1988. A new method for sequencing DNA. *Anal. Biochem.* 174:423-36.

Jansson V and Jansson K. 2002. Enzymatic chemiluminescence assay for inorganic pyrophosphate. *Analytical Biochemistry*. 304: 135-137.

Jensen N. 2002. Flow-through processing on a microchip for DNA pyrosequencing. *Natl. Nanofabrication Users Network.* 66-67.

Kane RS, Takayama S, Ostuni E, Ingber DE, and Whitesides GM. 1999. Patterning proteins and cells using soft lithography. *Biomaterials*. 20:2363-76.

Kulkarni, SS, Funk, EW, and Li, NN. 1992. Ultrafiltration, pp. 391-453 in *Membrane Handbook,* WSW Ho and KK Sirkar, eds. New York: Van Nostrand Reinhold.

Lehninger, AL. 1975. *Biochemistry*, 2nd ed. New York: Worth Publishers.

Lonsdale, HK. 1982. The growth of membrane technology. *J. Memb. Sci.* 10:81-181.

Madou MJ 1997. Fundamentals of Microfabrication. Boca Raton, FL, USA: CRC Press.

Mardilovich, PP, Govyadinov, AN, Mukhurov, NI, Rzhevskii, AM, and Paterson, R. New and modified anodic alumina membranes. Part I: Treatment of anodic alumina membranes. Part II: Comparison of solubility of amorphous (normal) and polycrystalline anodic alumina membranes. *J. Memb. Sci.* 98:131-155.

Martin CR. 1994. Nanomaterials – A membrane-based synthetic approach. *Science*. 266:1961-1966.

IDT Exhibit 1020.6063

Matsuda T and Chung DJ. 1994. Microfabricated surface designs for cell culture and diagnosis. *ASAIO J.* 1:M594-7.

Michael KL, Taylor LC, Schultz SL, and Walt DR. 1998. Randomly ordered addressable high-density optical sensor arrays. *Anal. Chem.* 70:1242-8.

Mulder, MHV. 1995. Polarization phenomena and membrane fouling, pp. 45-84 in *Membrane Separation Technology: Principles and Applications*, RD Noble and SA Stern, eds. Amsterdam: Elsevier Science BV.

Nyrén P, Pettersson B, and Uhlén M. 1993. Solid phase DNA minisequencing by an enzymatic luminometric inorganic pyrophosphate detection assay. *Anal. Biochem.* 208:171-5.

Rai-Choudhury P. 1997. *Handbook of Microlithography, Micromachining and Microfabrication. Vol. 1: Microlithography* and *Vol. 2: Micromachining and Microfabrication.* SPIE Press.

Ronaghi M, Uhlén M, and Nyrén P. 1998. A sequencing method based on real-time pyrophosphate. *Science*. 281:363-5.

Schuller J. 2002. Microfluidics for DNA pyrosequencing. *Natl. Nanofabrication Users Network.* 76-77.

Tanford, C. 1961. *Physical Chemistry of Macromolecules.* New York: Wiley.

Taylor LC and Walt DR. 2000. Application of high-density optical microwell arrays in a live-cell biosensing system. *Anal. Biochem.* 278:132-42

Zhu H, Klemic J F, Chang S, Bertone P, Casamayor A, Klemic K G, Smith D, Gerstein M, Reed M A, and & Snyder M. 2000. Analysis of yeast protein kinases using protein chips. Nature Genet. 26:283–89.

IDT Exhibit 1020.6064

We Claim:

1. A CMRA comprising:
   (a) a microreactor element comprising an array of open microchannels or open microwells, the longitudinal axes of said microchannels or microwells arranged in a substantially parallel manner; and
   (b) a porous filter element in contact with the microreactor element to form a bottom to the microchannels or microwells, thereby defining a series of reaction chambers, wherein the porous filter element comprises a permselective membrane that blocks the passage of nucleic acids, proteins and beads there across, but permits the passage of low molecular weight solutes, organic solvents and water there across.

2. The CMRA of claim 1, wherein the microreactor element comprises a plate formed from a fused fiber optic bundle, wherein the microchannels extend from the top face of the plate through to the bottom face of the plate.

3. The CMRA of claim 1 further comprising an additional porous support between the microreactor element and the porous filter element.

4. The CMRA of claim 1 wherein the porous filter element comprises an ultrafilter.

5. The CMRA of claim 1 further comprising at least one mobile solid support disposed in each of a plurality of the microchannels of the microreactor element.

6. The CMRA of claim 5 wherein the mobile solid support is a bead.

7. The CMRA of claim 1 or claim 6 wherein the mobile solid support has an enzyme and/or a nucleic acid immobilized thereon.

8. A method of making the CMRA of claim 1 comprising attaching a microreactor element to a porous filter element.

IDT Exhibit 1020.6065

9. A UMRA comprising a porous filter element against which molecules are concentrated by concentration polarization wherein discrete reaction chambers are formed in discrete locations on the surface of or within the porous filter element by depositing reactant molecules at discrete sites on or within the porous filter element.

10. The UMRA of claim 9 wherein the reaction chambers are formed by depositing mobile solid supports having said reactant molecules immobilized thereon, on the surface of, or within, the porous element.

11. The UMRA of claim 9 wherein the porous filter element comprises an ultrafilter.

12. The UMRA of claim 9 wherein the mobile solid support is a bead.

13. The UMRA of claim 10 or claim 12 wherein the mobile solid support has an enzyme and/or a nucleic acid immobilized thereon.

14. A UMRA comprising:

(a) a porous membrane with discrete reaction sites formed by depositing mobile solid supports having said reactant molecules thereon, on the surface of, or within the porous membrane;

(b) a nucleic acid template immobilized to a solid support; and

(c) optionally, at least one immobilized enzyme.

15. The UMRA of claim 14 wherein the mobile solid support is a bead.

16. The UMRA of claim 14 wherein the porous membrane is nylon membrane.

17. The UMRA of claim 14 wherein the porous membrane is made of a woven fiber.

18. The UMRA of claim 14 wherein the porous membrane pore size at least 0.02 μm.

19. The UMRA of claim 1 wherein the solid support is selected from the group consisting of a bead, glass surface, fiber optic or the porous membrane.

IDT Exhibit 1020.6066

20. The UMRA of claim 14 wherein the immobilized enzyme is immobilized to a bead or the porous membrane.

21. The UMRA of claim 14 wherein the immobilized enzyme is selected from the group consisting of ATP sulfurylase, luciferase, hypoxanthine phosphoribosyltransferase, xanthine oxidase, uricase or peroxidase.

22. An array comprising:

(a) a first porous membrane with a plurality of discrete reaction sites disposed thereon, and/or within, wherein each reaction site has immobilized template adhered to the surface; and

(b) a second porous membrane with at least one enzyme located on the surface of, and/or within, the membrane, wherein the second porous membrane is in direct contact with the first porous membrane.

23. The array of claim 22 wherein the first or second membrane is a nylon membrane.

24. The array of claim 22 wherein the porous membrane is made of a woven fiber.

25. The array of claim 22 wherein the each reaction site is defined by the pores of the porous membrane.

26. The array of claim 22 wherein the first and second porous membranes have a pore size of at least 0.2 μm.

27. The array of claim 22 wherein the template is immobilized to a bead or to the porous membrane.

28. The array of claim 22 wherein the enzyme is selected from the group consisting of ATP sulfurylase, luciferase, hypoxanthine phosphoribosyltransferase, xanthine oxidase, uricase or peroxidase.

IDT Exhibit 1020.6067

29. A CMRA comprising an array of open microchannels or microwells attached to a porous filter or membrane.

30. The CMRA of claim 29 further comprising a mechanical support, wherein the mechanical support separates the microchannels from the porous membrane.

31. The CMRA of claim 30 wherein the mechanical support is selected from the group consisting of plastic mesh, wire screening or molded or machined spacers.

32. The CMRA of claim 29 wherein the porous membrane is a nylon membrane.

33. The CMRA of claim 29 wherein the porous membrane is made of a woven fiber.

34. The CMRA of claim 29 wherein the membrane pore size is at least 0.02 μm.

35. The CMRA of claim 29 wherein the microchannels are formed by concentration polarization.

36. An apparatus for determining the nucleic acid sequence in a template nucleic acid polymer, comprising:

(a) a CMRA or UMRA;

(b) nucleic acid delivery means for introducing template nucleic acid polymers to the discrete reaction sites;

(c) nucleic acid delivery means to deliver reagents to the reaction sites to create a polymerization environment in which the nucleic acid polymers will act as template polymers for the synthesis of complementary nucleic acid polymers when nucleotides are added;

(d) convective flow delivery means to immobilize reagents to the porous membrane;

(e) detection means for detecting the formation of inorganic pyrophosphate enzymatically; and

(f) data processing means to determine the identity of each nucleotide in the complementary polymers and thus the sequence of the template polymers.

IDT Exhibit 1020.6068

37. The apparatus of claim 36 wherein the porous membrane is a nylon membrane.

38. The apparatus of claim 36 wherein the nylon membrane is made of a woven fiber.

39. The apparatus of claim 36 wherein the pore size is at least 0.02 μm.

40. The apparatus of claim 36 wherein the discrete reaction sites are formed by concentration polarization.

41. The apparatus of claim 36 wherein the convective flow delivery means is a syringe or a peristaltic pump.

42. The apparatus of claim 36 wherein the template nucleic acid is attached to a solid support.

43. The apparatus of claim 42 wherein the solid support is selected from the group consisting of a bead, glass surface, fiber optic or porous membrane.

44. The apparatus of claim 36 wherein the enzyme detecting inorganic pyrophosphate is selected from the group consisting of ATP sulfurylase, luciferase, hypoxanthine phosphoribosyltransferase, xanthine oxidase, uricase or peroxidase.

45. The apparatus of claim 36 wherein the detection means is a CCD camera.

46. The apparatus of claim 36 wherein the data processing means is a computer.

47. An apparatus for processing a plurality of analytes, the apparatus comprising:
(a) a CMRA or an UMRA;
(b) fluid means for delivering processing reagents from one or more reservoirs to the flow chamber so that the analytes disposed therein are exposed to the reagents; and

IDT Exhibit 1020.6069

(c) detection means for detecting a sequence of optical signals from each of the reaction sites, each optical signal of the sequence being indicative of an interaction between a processing reagent and the analyte disposed in the reaction site, wherein the detection means is in communication with the reaction site.

48. The apparatus of claim 47 wherein the porous membrane is a nylon membrane.

49. The apparatus of claim 47 wherein the pore size is at least 0.02 μm.

50. The apparatus of claim 47 wherein the template nucleic acid is attached to a solid support.

51. The apparatus of claim 50 wherein the solid support is selected from the group consisting of a bead, glass surface, fiber optic or porous membrane.

52. The apparatus of claim 47 wherein the convective flow delivery means is a peristaltic pump.

53. The apparatus of claim 47 wherein the enzyme detecting inorganic pyrophosphate is selected from the group consisting of ATP sulfurylase, luciferase, hypoxanthine phosphoribosyltransferase, xanthine oxidase, uricase or peroxidase.

54. The apparatus of claim 47 wherein the detection means is a CCD camera.

55. The apparatus of claim 47 wherein the data processing means is a computer.

56. An apparatus for determining the base sequence of a plurality of nucleotides on an array, the apparatus comprising:

(a) a CMRA or UMRA;

(b) reagent delivery means for adding an activated nucleotide 5'-triphosphate precursor of one known nitrogenous base to a reaction mixture to each reaction site, each reaction mixture comprising a template-directed nucleotide polymerase and a single-stranded polynucleotide template

IDT Exhibit 1020.6070

hybridized to a complementary oligonucleotide primer strand at least one nucleotide residue shorter than the templates to form at least one unpaired nucleotide residue in each template at the 3'-end of the primer strand, under reaction conditions which allow incorporation of the activated nucleoside 5'-triphosphate precursor onto the 3'-end of the primer strands, provided the nitrogenous base of the activated nucleoside 5'-triphosphate precursor is complementary to the nitrogenous base of the unpaired nucleotide residue of the templates;

(c) detection means for detecting whether or not the nucleoside 5'-triphosphate precursor was incorporated into the primer strands in which incorporation of the nucleoside 5'-triphosphate precursor indicates that the unpaired nucleotide residue of the template has a nitrogenous base composition that is complementary to that of the incorporated nucleoside 5'-triphosphate precursor; and

(d) means for sequentially repeating steps (b) and (c), wherein each sequential repetition adds and detects the incorporation of one type of activated nucleoside 5'-triphosphate precursor of known nitrogenous base composition; and

(e) data processing means for determining the base sequence of the unpaired nucleotide residues of the template in each reaction chamber from the sequence of incorporation of said nucleoside precursors.

57. The apparatus of claim 56 wherein the porous membrane is a nylon membrane.

58. The apparatus of claim 53 wherein the pore size is at least 0.02 μm.

59. The apparatus of claim 53 wherein the template nucleic acid is attached to a solid support.

60. The apparatus of claim 59 wherein the solid support is selected from the group consisting of a bead, glass surface, fiber optic or porous membrane.

IDT Exhibit 1020.6071

61. The apparatus of claim 53 wherein the convective flow delivery means is a peristaltic pump.

62. The apparatus of claim 53 wherein the enzyme detecting inorganic pyrophosphate is selected from the group consisting of ATP sulfurylase, luciferase, hypoxanthine phosphoribosyltransferase, xanthine oxidase, uricase or peroxidase.

63. The apparatus of claim 53 wherein the detection means is a CCD camera.

64. The apparatus of claim 53 wherein the data processing means is a computer.

65. An apparatus for determining the nucleic acid sequence in a template nucleic acid polymer, comprising:

(a) a CMRA or UMRA;

(b) nucleic acid delivery means for introducing a template nucleic acid polymers onto the reaction sites;

(c) nucleic acid delivery means to deliver reagents to the reaction chambers to create polymerization environment in which the nucleic acid polymers will act as a template polymers for the synthesis of complementary nucleic acid polymers when nucleotides are added;

(d) reagent delivery means for successively providing to the polymerization environment a series of feedstocks, each feedstock comprising a nucleotide selected from among the nucleotides from which the complementary nucleic acid polymer will be formed, such that if the nucleotide in the feedstock is complementary to the next nucleotide in the template polymer to be sequenced said nucleotide will be incorporated into the complementary polymer and inorganic pyrophosphate will be released;

(e) detection means for detecting the formation of inorganic pyrophosphate enzymatically; and

(f) data processing means to determine the identity of each nucleotide in the complementary polymers and thus the sequence of the template polymers.

66. The apparatus of claim 65 wherein the porous membrane is a nylon membrane.

IDT Exhibit 1020.6072

67. The apparatus of claim 65 wherein the pore size is at least 0.02 μm.

68. The apparatus of claim 65 wherein the template nucleic acid is attached to a solid support.

69. The apparatus of claim 68 wherein the solid support is selected from the group consisting of a bead, glass surface, fiber optic or porous membrane.

70. The apparatus of claim 65 wherein the convective flow delivery means is a peristaltic pump.

71. The apparatus of claim 65 wherein the enzyme detecting inorganic pyrophosphate is selected from the group consisting of ATP sulfurylase, luciferase, hypoxanthine phosphoribosyltransferase, xanthine oxidase, uricase or peroxidase.

72. The apparatus of claim 65 wherein the detection means is a CCD camera.

73. The apparatus of claim 65 wherein the data processing means is a computer.

74. A system for sequencing a nucleic acid comprising the following components:
(a) a CMRA or UMRA;
(b) at least one enzyme immobilized on a solid support;
(c) means for flowing reagents over said porous membrane;
(d) means for detection; and
(e) means for determining the sequence of the nucleic acid.

75. The system of claim 74 wherein the porous membrane is a nylon membrane.

76. The system of claim 74 wherein the porous membrane has a pore size is at least 0.02 μm.

77. The system of claim 74 wherein the reaction sites are formed by concentration polarization.

IDT Exhibit 1020.6073

78. The system of claim 74 wherein the immobilized enzyme is selected from the group consisting of ATP sulfurylase, luciferase, hypoxanthine phosphoribosyltransferase, xanthine oxidase, uricase or peroxidase.

79. The system of claim 74 wherein the solid support is selected from the group consisting of a bead, glass surface, fiber optic or porous membrane.

80. The system of claim 74 wherein the means for detection is a CCD camera.

81. The system of claim 74 wherein the means for determining a sequence is by pyrophosphate sequencing.

82. A system for sequencing a nucleic acid comprising the following components:
(a) a CMRA or UMRA
(b) at least one enzyme immobilized on a solid support;
(c) means for flowing reagents over said porous membrane;
(d) means for enzymatic detection; and
(e) means for determining the sequence of the nucleic acid.

83. The system of claim 82 wherein the porous membrane is a nylon membrane.

84. The system of claim 82 wherein the porous membrane has a pore size is at least 0.02 μm.

85. The system of claim 82 wherein the reaction sites are formed by concentration polarization.

86. The system of claim 82 wherein the immobilized enzyme is selected from the group consisting of ATP sulfurylase, luciferase, hypoxanthine phosphoribosyltransferase, xanthine oxidase, uricase or peroxidase.

87. The system of claim 82 wherein the solid support is selected from the group consisting of a bead, glass surface, fiber optic or porous membrane.

IDT Exhibit 1020.6074

88. The system of claim 82 wherein the means for detection is a CCD camera.

89. The system of claim 82 wherein the means for determining a sequence is by pyrophosphate sequencing.

90. A method for carrying out separate parallel independent reactions in an aqueous environment, comprising:

(a) delivering a fluid containing at least one reagent to an array, using the CMRA of claim 1 or the UMRA of claim 9, wherein each of the reaction sites immersed in a substance such that when the fluid is delivered onto each reaction site, the fluid does not diffuse onto an adjacent site;

(b) washing the fluid from the array in the time period after the starting material has reacted with the reagent to form a product in each reaction site;

(c) sequentially repeating steps (a) and (b).

91. The method of claim 90 wherein the product formed in any one reaction chamber is independent of the product formed in any other reaction chamber, but is generated using one or more common reagents.

92. The method of claim 90 wherein the starting material is a nucleic acid sequence and at least one reagent in the fluid is a nucleotide or nucleotide analog.

93. The method of claim 90 wherein the fluid additionally comprises a polymerase capable of reacting the nucleic acid sequence and the nucleotide or nucleotide analog.

94. The method of claim 90 additionally comprising repeating steps (a) and (b) sequentially.

95. The method of claim 90 wherein the substance is mineral oil.

96. The method of claim 90 wherein the reaction sites are defined by concentration polarization.

IDT Exhibit 1020.6075

97. A method of determining the base sequence of nucleotides in an array format, the method comprising the steps of:

(a) adding an activated nucleoside 5'-triphopsphate precursor of one known nitrogenous base composition to a plurality of reaction sites localized on a CMRA or UMRA, wherein the reaction site is comprised of a template-directed nucleotide polymerase and a heterogenous population of single stranded templates hybridized to complementary oligonucleotide primer strands at least one nucleotide residue shorter than the templates to form at least one unpaired nucleotide residue in each template at the 3' end of the primer strand under reaction conditions which allow incorporation of the activated nucleoside 5'-triphosphate precursor onto the 3' end of the primer strand under reaction conditions which allow incorporation of the activated nucleoside 5'-triphosphate precursor onto the 3' end of the primer strands, provided the nitrogenous base of the activated nucleoside 5'-triphosphate precursor is complementary to the nitrogenous base of the unpaired nucleotide residue of the templates;

(b) detecting whether or not the nucleoside 5'-triphosphate precursor was incorporated into the primer strands in which incorporation of the nucleoside 5'-triphosphate precursor indicates that the unpaired nucleotide residue of the template has a nitrogenous base composition that is complementary to that of the incorporated nucleoside 5'-triphosphate precursor; and

(c) sequentially repeating steps (a) and (b), wherein each sequential repetition adds and detects the incorporation of one type of activated nucleoside 5'-triphosphate precursor of known nitrogenous base composition;

(d) determining the base sequence of the unpaired nucleotide residues of the template from the sequence of incorporation of said nucleoside precursors.

98. The method of claim 97 wherein the detection of the incorporation of the activated precursor is accomplished enzymatically.

IDT Exhibit 1020.6076

99. The method of claim 98 wherein the enzyme is selected from the group consisting of ATP sulfurylase, luciferase, hypoxanthine phosphoribosyltransferase, xanthine oxidase, uricase or peroxidase.

100. The method of claim 97 wherein the enzyme is immobilized to a solid support.

101. The method of claim 97 wherein the solid support is selected from the group comprising a bead, glass surface, fiber optic or porous membrane.

102. A method of determining the base sequence of a plurality of nucleotides on an array, said method comprising:

(a) providing a plurality of sample DNA's, each disposed within a plurality of reaction sites on a CMRA or UMRA;

(b) detecting the light level emitted from a plurality of reaction sites on respective proportional of an optically sensitive device;

(c) converting the light impinging upon each of said portions of said optically sensitive device into an electrical signal which is distinguishable from the signals from all of said other regions;

(d) determining a light intensity for each of said discrete regions from the corresponding electrical signal;

(e) recording the variations of said electrical signals with time.

103. The method of claim 102 wherein the porous membrane is a nylon membrane.

104. The method of claim 102 wherein the pore size is at least 0.2 μm.

105. The method of claim 102 wherein the detection is performed enzymatically.

106. The method of claim 105 wherein the enzyme is selected from the group consisting of ATP sulfurylase, luciferase, hypoxanthine phosphoribosyltransferase, xanthine oxidase, uricase or peroxidase.

IDT Exhibit 1020.6077

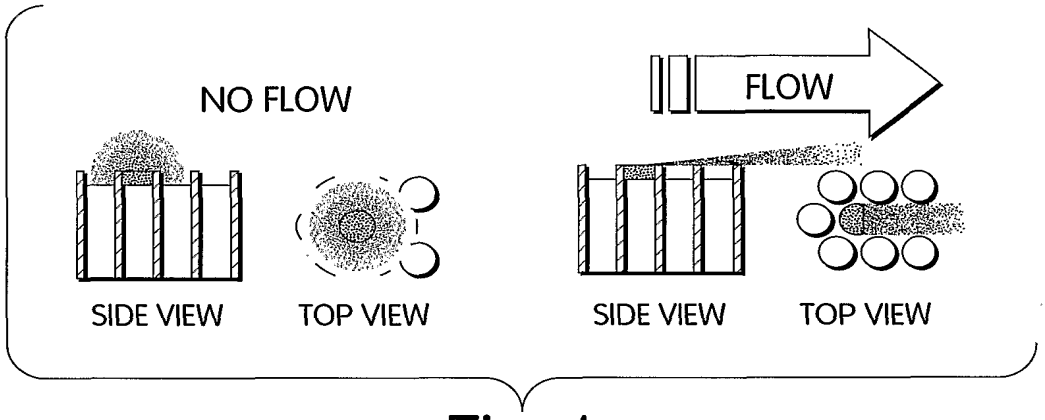


Fig. 1

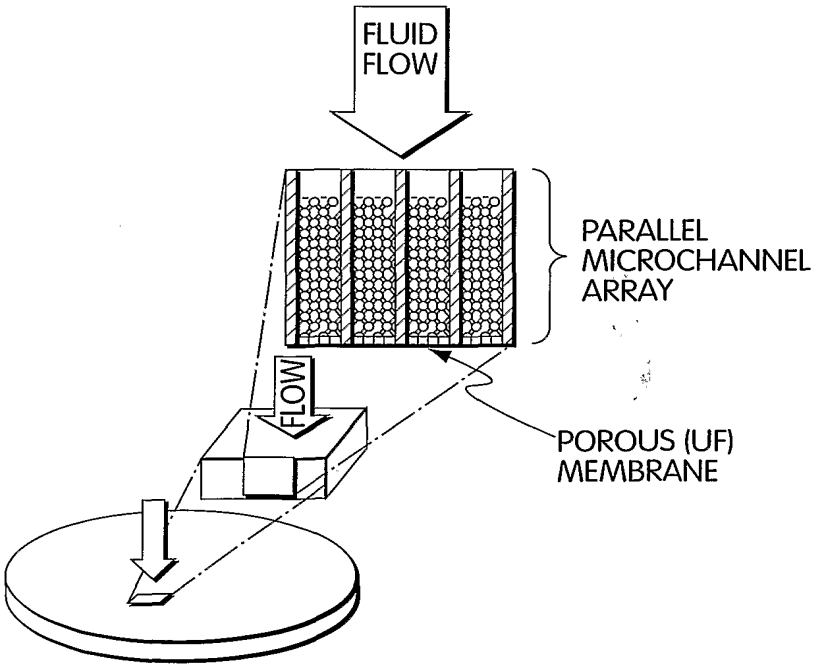


Fig. 2

IDT Exhibit 1020.6078

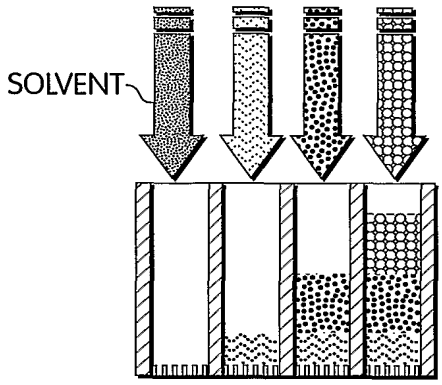


Fig. 3

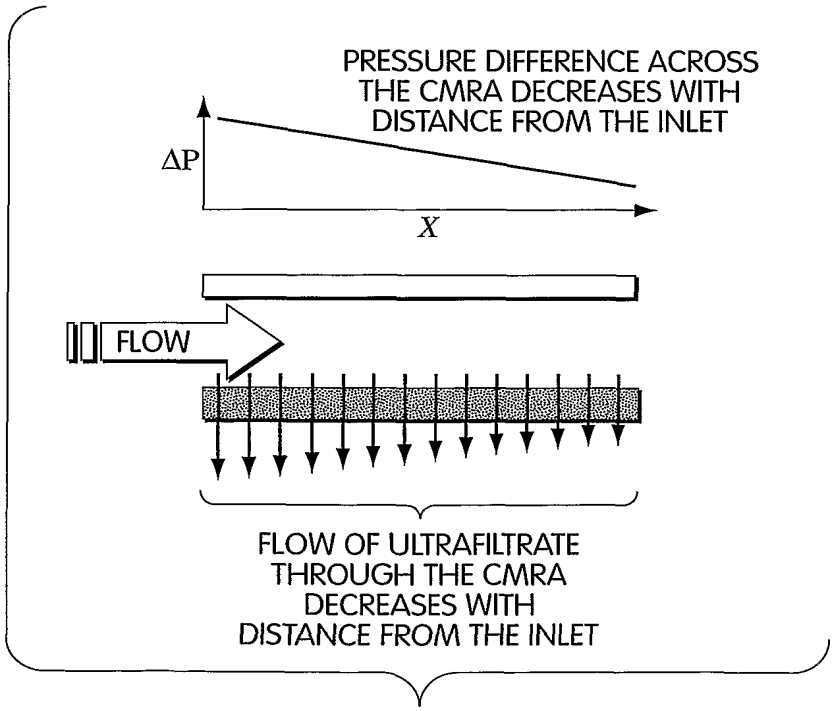


Fig. 4

IDT Exhibit 1020.6079

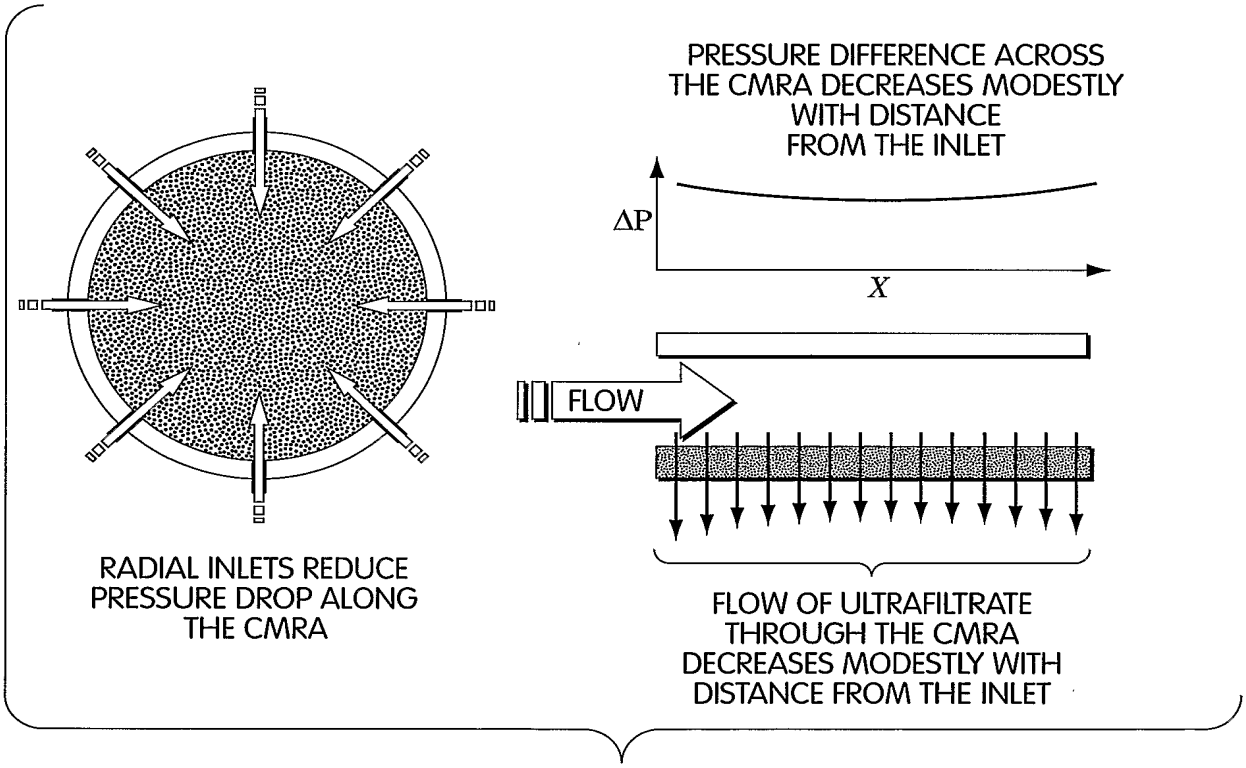


Fig. 5

IDT Exhibit 1020.6080

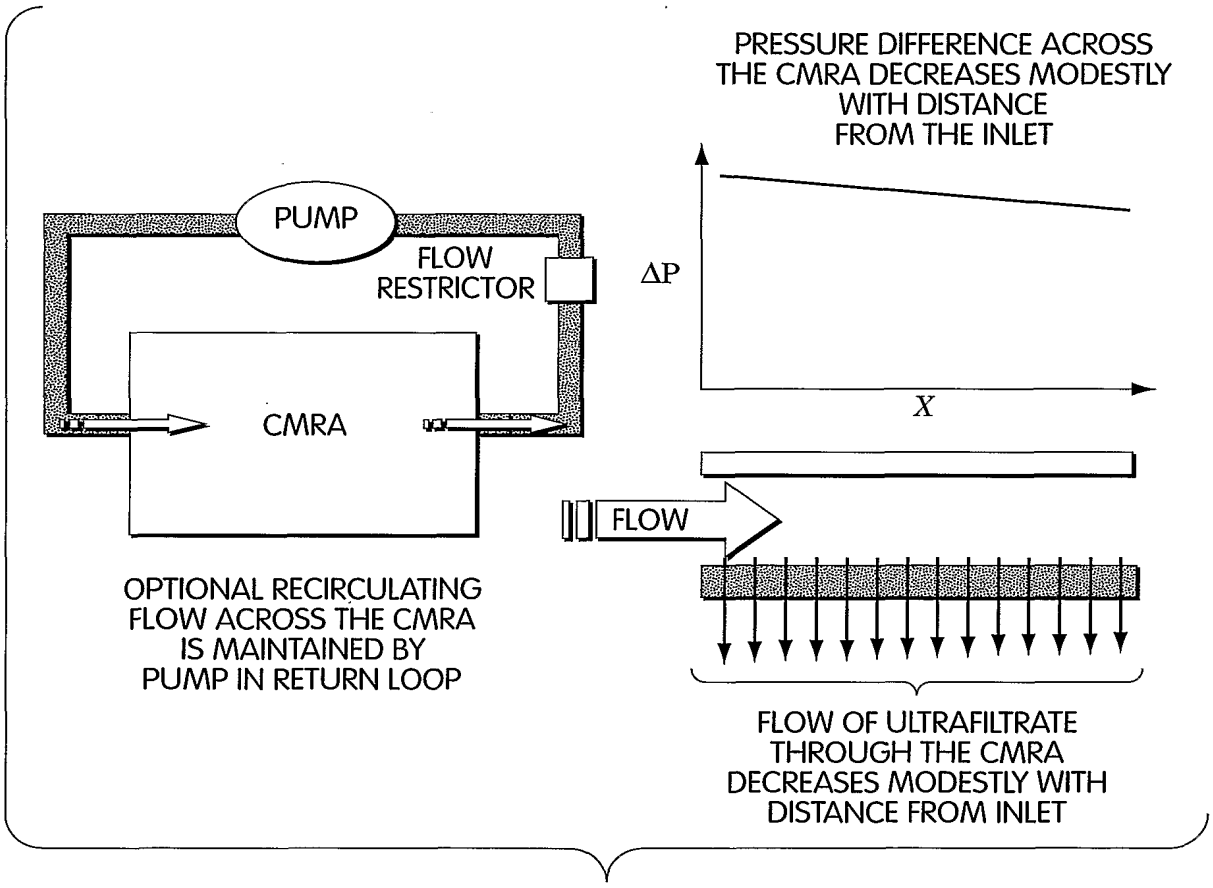


Fig. 6

IDT Exhibit 1020.6081

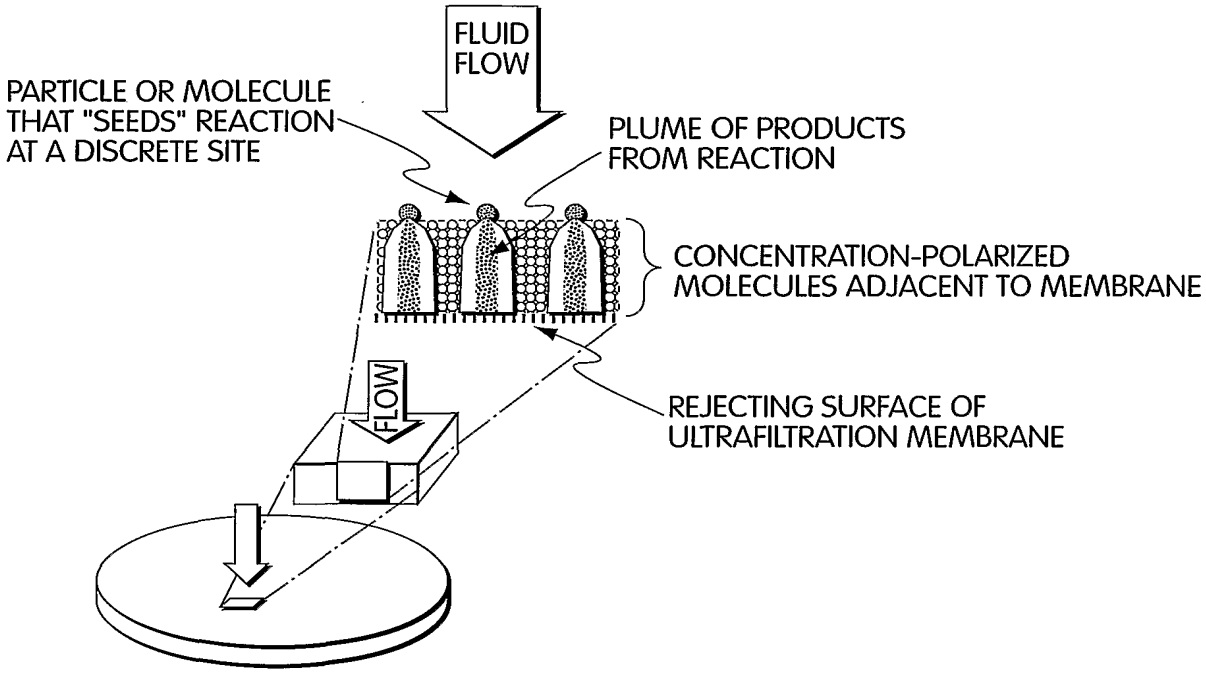


Fig. 7

IDT Exhibit 1020.6082

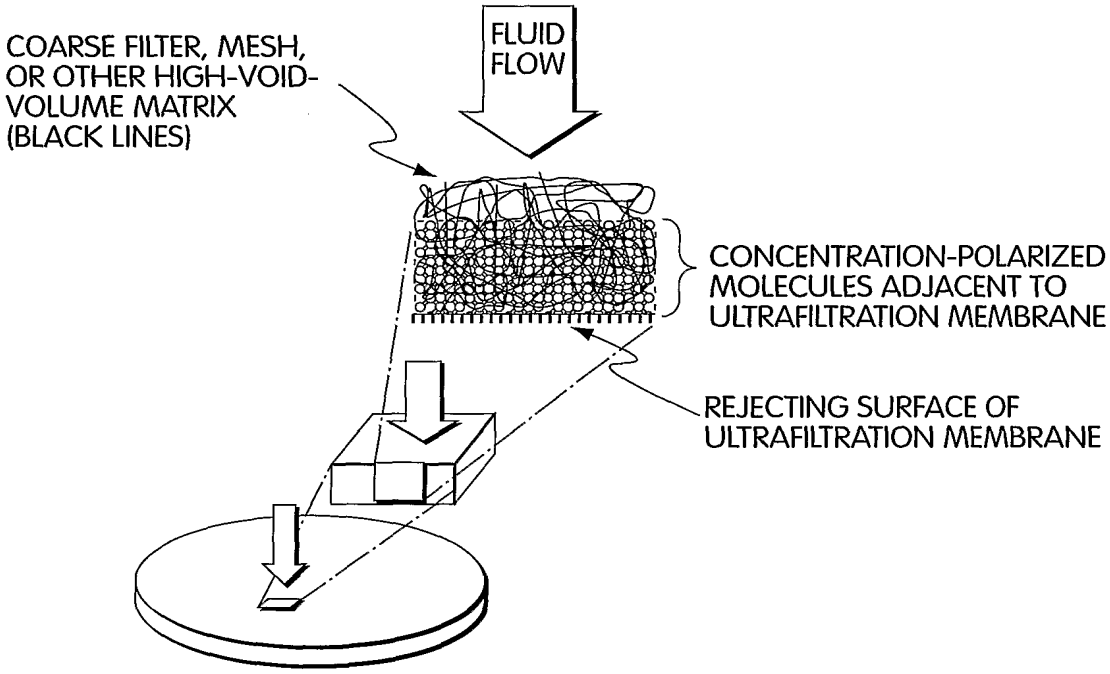


Fig. 8

IDT Exhibit 1020.6083

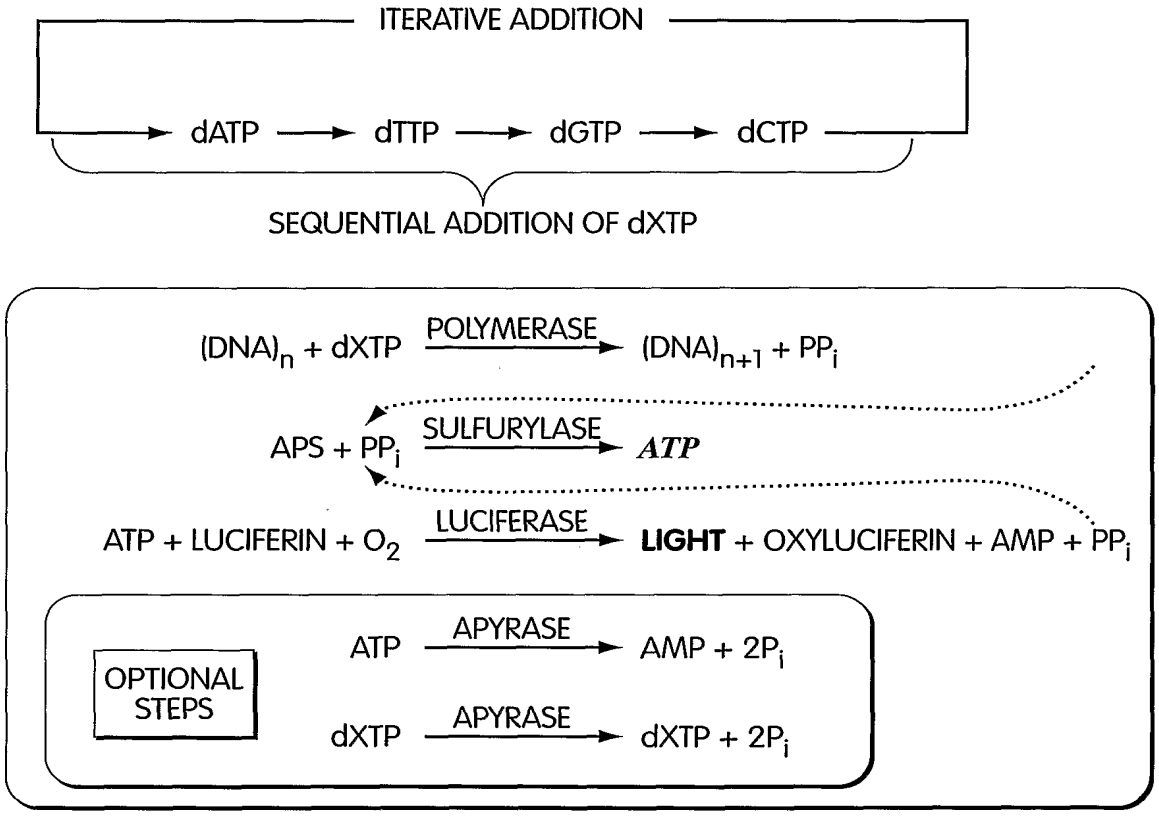


Fig. 9

IDT Exhibit 1020.6084

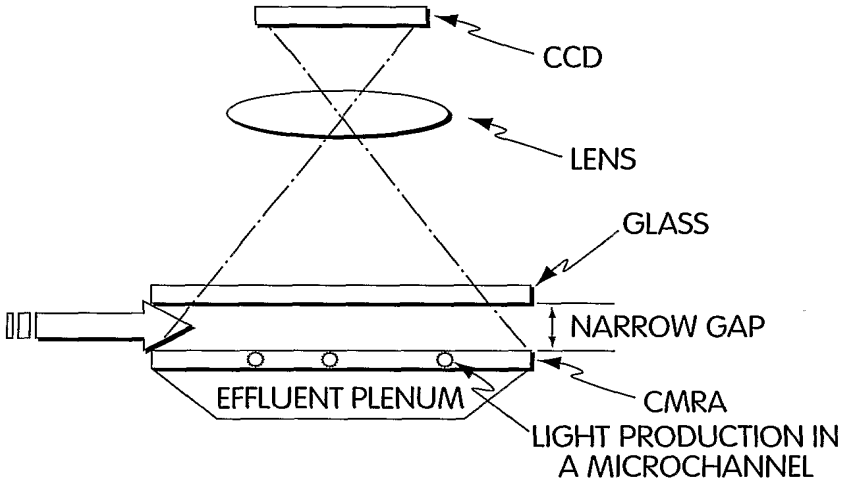


Fig. 10

IDT Exhibit 1020.6085

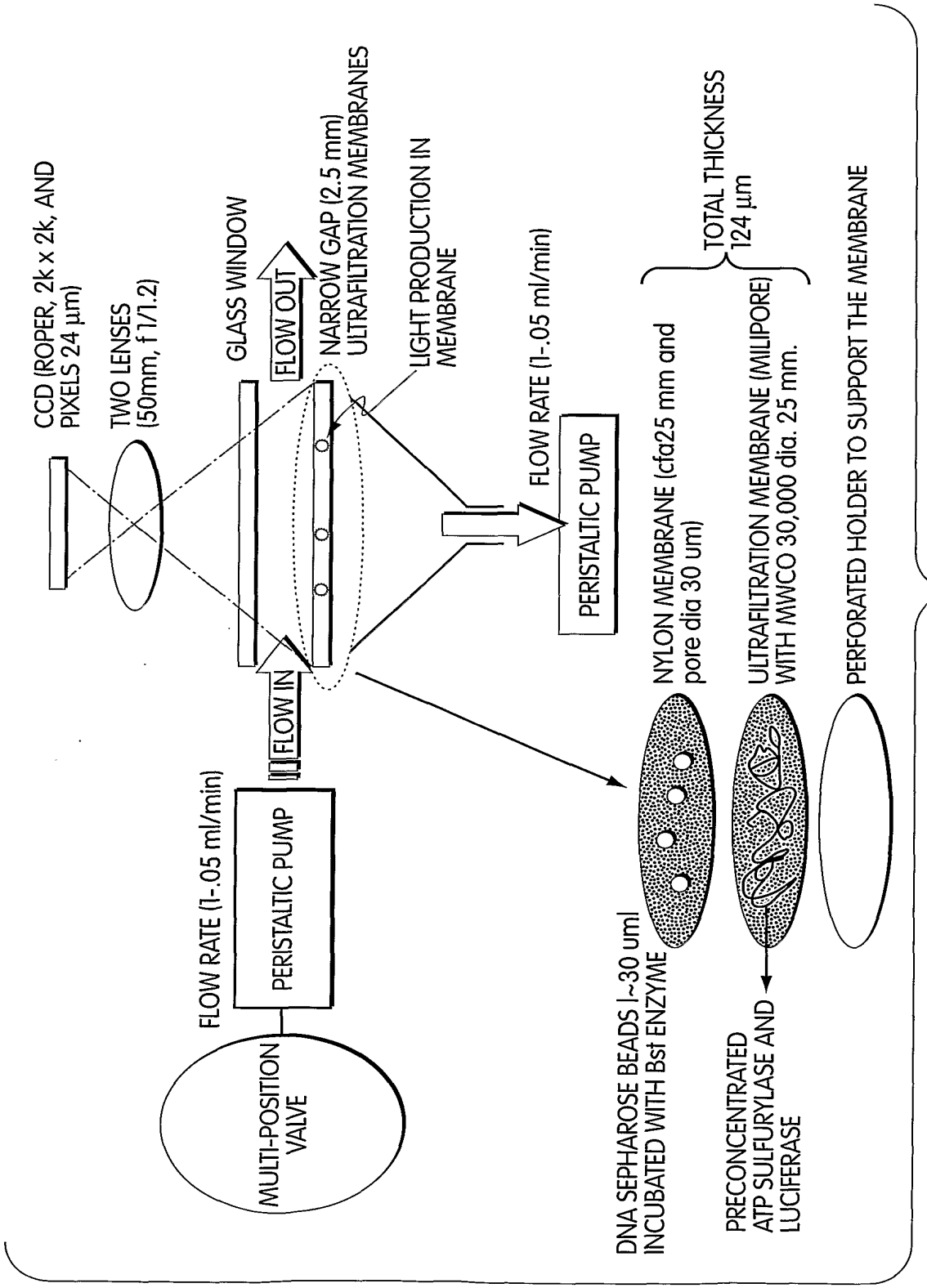


Fig. 11

IDT Exhibit 1020.6086

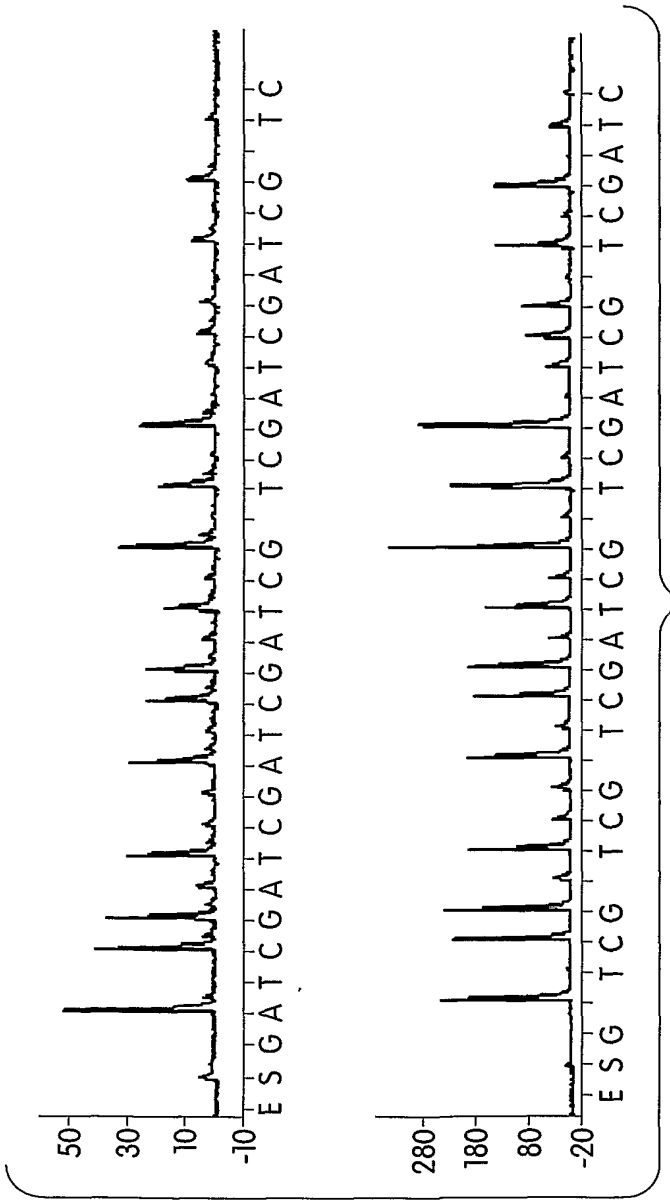


Fig. 12

IDT Exhibit 1020.6087

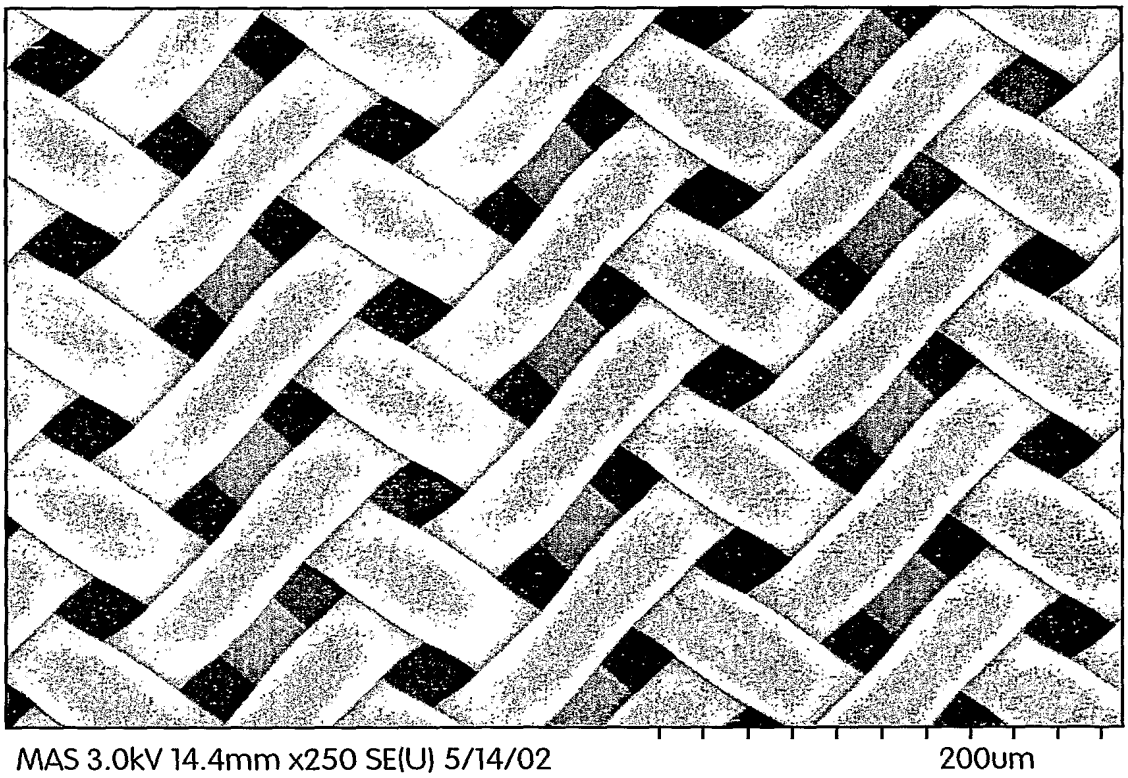


Fig. 13

IDT Exhibit 1020.6088

Espacenet

Bibliographic data: WO0146464 (A1) — 2001-06-28

BRANCHED COMPOUND FOR USE IN NUCLEIC ACID DETECTION AND ANALYSIS REACTIONS

Inventor(s): HUBER MARTIN [AT]; SCHMIDT WOLFGANG [AT]; MUELLER MANFRED [AT]; HILLER REINHOLD [AT] ± (HUBER, MARTIN, ; SCHMIDT, WOLFGANG, ; MUELLER, MANFRED, ; HILLER, REINHOLD)

Applicant(s): LION BIOSCIENCE AG [DE]; HUBER MARTIN [AT]; SCHMIDT WOLFGANG [AT]; MUELLER MANFRED [AT]; HILLER REINHOLD [AT] ± (LION BIOSCIENCE AG, ; HUBER, MARTIN, ; SCHMIDT, WOLFGANG, ; MUELLER, MANFRED, ; HILLER, REINHOLD)

Classification: - **international:** ***C07B61/00; C07F7/18; C07H21/00; C07K1/107; C07K14/00; C07K17/00; C12Q1/68;*** (IPC1-7): C07H21/00; C12Q1/68

- **cooperative:** **C07F7/1836**; **C07H1/00**; **C07H21/00**; **C07K1/1077**; **C07K14/001**; **C07K17/00**; **C12Q1/6834**; **C12Q1/6837**; **C12Q1/686**; **C40B50/14**; C07B2200/11 more

Application number: WO2000EP13098 20001221 Global Dossier

Priority number(s): EP19990125484 19991221 ; EP19990125485 19991221 ; US20000211217P 20000613 ; US20000211209P 20000613

Also published as: AU2170901 (A) AU2843201 (A) AU3163101 (A) EP1244812 (A1) EP1250344 (A2) more

Abstract of WO0146464 (A1)

The present invention relates to a branched compound for use in nucleic acid detection and analysis reactions, as well as a process for the synthesis thereof. The invention also concerns a process for a nucleic acid synthesis reaction making use of the branched compound as well as uses of the novel compound. The invention in addition concerns a kit comprising the branched compound.

IDT Exhibit 1020.6089

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
28 June 2001 (28.06.2001)

PCT

(10) International Publication Number
WO 01/46464 A1

(51) International Patent Classification[7]: C12Q 1/68, C07H 21/00

(21) International Application Number: PCT/EP00/13098

(22) International Filing Date:
21 December 2000 (21.12.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

99125484.8	21 December 1999 (21.12.1999)	EP
99125485.5	21 December 1999 (21.12.1999)	EP
60/211,217	13 June 2000 (13.06.2000)	US
60/211,209	13 June 2000 (13.06.2000)	US

(71) Applicant *(for all designated States except US)*: **LION BIOSCIENCE AG** [DE/DE]; Im Neuenheimer Feld 515, 69120 Heidelberg (DE).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **HUBER, Martin** [AT/AT]; Rennweg 84, A-1030 Vienna (AT). **SCHMIDT, Wolfgang** [AT/AT]; Hetzgasse 12, A-1030 Vienna (AT). **MÜLLER, Manfred** [AT/AT]; Schömergasse 30, A-3400 Klosterneuburg (AT). **HILLER, Reinhold** [AT/AT]; Löhnergasse 8, A-1120 Vienna (AT).

(74) Agents: **SCHOHE, Stefan** et al.; Boehmert & Boehmert, Franz-Joseph-Strasse 38, 80801 München (DE).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

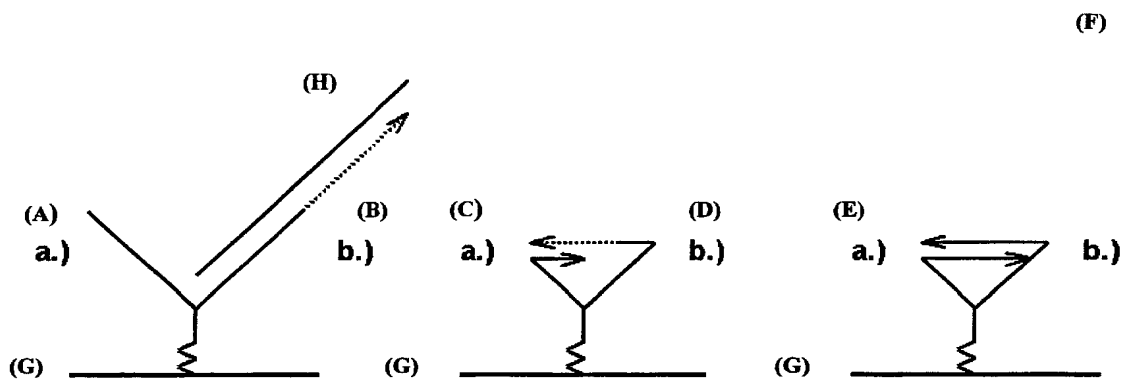
Published:

— *With international search report.*

— *Before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments.*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: BRANCHED COMPOUND FOR USE IN NUCLEIC ACID DETECTION AND ANALYSIS REACTIONS



(57) Abstract: The present invention relates to a branched compound for use in nucleic acid detection and analysis reactions, as well as a process for the synthesis thereof. The invention also concerns a process for a nucleic acid synthesis reaction making use of the branched compound as well as uses of the novel compound. The invention in addition concerns a kit comprising the branched compound.

WO 01/46464 A1

IDT Exhibit 1020.6090

"Branched Compound for Use in Nucleic Acid Detection and Analysis Reactions"

Description

Background of the Invention

Genetic analysis often involves analysis of the nucleic acid sequence, structure or composition of a given organism or sample. frequently such analysis incorporates the step of, or requires nucleic acid amplification. One of the well known methods for nucleic acid amplification is the PCR, or polymerase chain reaction method also disclosed in US 4,683,195 and US 4,683,202. Here, a nucleic acid sample serves as a template for a polymerase dependant *in-vitro* replication starting from for two separate primers. Polymerases are enzymes capable of synthesizing RNA or DNA making use of RNA or DNA as a template. Often times the analysis is performed on RNA (ribo-nucleic acid), here amplification additionally requires an enzymatic reverse transcription into DNA (deoxyribonucleic acid), but equally often on DNA. PCR is becoming powerful tool in diagnostics. PCR kits are becoming available for the detection and analysis of various patho-genic organisms as well *e.g.* mutant alleles of human genes.

PCR is mostly performed *in-vitro*, *i.e.* in a tube whereby the components are mostly supplied in liquid format. Alternatively, one or more of the components, these usually being, a polymerase, a buffer, a template, two or more oligonucleotides, may be bound to some form of a solid-phase.

One very common problem with the PCR being performed in a standard non solid-phase format is the limitation with respect to the number of primer pairs that may be used simultaneously in one reaction. In contrast a solid-phase set-up would theoretically enable the use of tens to thou-sands of primer pairs.

More recently, it has been proposed to use one primer bound to a solid-support together with a free primer in solution in order to simultaneously amplify and bind a PCR product onto a surface (Oroskar, A. A., Rasmussen, S. E., et al., Detection of immobilised amplicons by ElISA-like techniques, Clinical Chemistry 42:1547 (1997)). The drawback of this approach is that the actual multiplexing , *i.e.* the use of a multitude of primer pairs in one reaction is not really facilitated as the free primers in solution may still spuriously anneal.

A solid-phase method is disclosed in US 5,641,658 or WO 96/04404 here, the oligonucleotides used in the reaction are bound to a support. Such a set-up would have a number of advantages

IDT Exhibit 1020.6091

over the standard PCR process, *e.g.* it would be easy to combine a multitude of primer pairs in one reaction without the drawbacks of false amplification products from false primer pairings. Thus, multiplexing would be facilitated.

A problem with such an approach is the efficiency of the reaction, thus also the product yield during amplification is poor and consequently the reliability of the entire process. Often these problems are associated with the fact that the reaction conditions *i.e.* the availability of the primer in the reaction are not optimal. One may envision that the 3' OH ends of the oligonucleotides are not available due to the fact that parts of the oligonucleotide are bound to the support. Such immobilised single-stranded DNAs which have conventionally been used are prepared by binding a single-stranded DNA at the terminal molecule or a suitable functional group introduced into the molecule. However, such conventional method has a drawback in that it is impossible to bind the single-stranded DNA to the carrier only at its terminal molecule by the use of the conventional method because an existing amino or hydroxy group, or other functional group artificially introduced, on the nucleotide molecules other than terminal also participates in the binding with the carrier. The immobilised DNA obtained with conventional methods is one in which the DNA molecules are bonded to the support at various sites of the strand. As will be recognised by those skilled in the art, such a molecule is poorly suited for providing sufficient experimental results in any aspects.

A further problem is that the molecules bound to the support are sterically hindered by the support itself from taking part in *e.g.* enzymatic reactions. EP0787205 discloses the use of linker between the oligonucleotide and the solid-support. However, the primary problem is not addressed here. The primers on a solid-support are not freely available in the reaction. Thus primers remain a limiting factor.

For efficient amplification to occur it would be desired to provide equimolar amount of each primer of a given pair of primers. Due to the fact that primers are bound to a support where an unequal distribution and binding of primers will occur stochastically it is impossible to achieve equimolar amounts of primer in the reaction.

Often reactive groups are used to bind the Primers to the support. Such reactive groups have been *e.g.* amino groups. It is known in the art that such groups are very unstable, consequently when a given primer pair is arrayed on a support wherein the terminal group of the primers are

e.g. amino groups it is to be expected, based on the instability of the groups, that after arraying the primers in each pair will not be present equimolar manner.

One method for the covalent attachment of oligonucleotides on glass supports is to treat the glass with an aminosilane and couple a 5' amino-modified oligonucleotide via covalent bond formation using PDC (Guo et al. 1994 NAR 22:5456-5465). This procedure has two major technical drawbacks. As outlined above, the terminal primary amino group is unstable and therefore also not useful for long-term storage which is one pre-requisite in DNA chip technology. One should mention here that glass slides are often chosen as solid support in molecular DNA analysis and often referred to as DNA chips. Second, the generation of an aminoreactive glass support requires complete inactivation of unreacted groups after spotting in order to avoid a high background which in fact makes multiple solid-phase approaches entirely impossible.

One attempt to solve the problem of poor efficiency, poor yields, poor reproducibility could potentially be to incubate the primers with the template for extended amounts of time to enable improved binding efficiency however, it has been found that this has, depending on the solid support, severe drawbacks with respect to the enzymatic function. Some preferred solid supports such as glass will bind the enzyme, even denature it, and thus disable the reaction. It will be recognised by those skilled in the art that it is not desired to establish conditions in which the enzymatic function is decreased.

It was therefore an object of the present invention to provide for a process, means and substances that lead to a high efficiency in *in-vitro* solid-phase nucleic acid synthesis reactions, thus also to a higher product yield and consequently a higher reliability of such reactions.

It was further an object of the present invention to provide for a process, means and substances to be used in *in-vitro* solid-phase nucleic acid synthesis reactions that are suited for providing good experimental results in many aspects such as, equimolarity of oligonucleotide amount, reduction of background signal, reduction of falsely synthesised products, availability of oligonucleotides in the reaction.

It was a further object within the concept of the invention to provide for novel molecules capable of solving the above problems. Thus it was *e.g.* an object of the present invention to provide for molecules that are readily available in a DNA synthesis reaction and not sterically hindered by *e.g.* a solid-support or the like.

It was a also an object within the concept of the invention to provide for processes for making the molecules according to the invention.

It was a further object within the concept of the invention to provide for a kit comprising one or more of the molecules according to the invention.

Equally it was an object of the present invention to provide for a kit comprising compounds necessary for performing the process according to the invention.

Further objects of the invention are apparent to the skilled person from the specification.

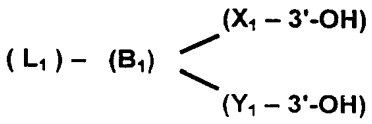
The objects of the present invention were accomplished by providing for novel molecules with advantageous characteristics, processes for making these molecules, novel processes for in vitro nucleic acid synthesis for use with and without the novel molecules as well as kits containing molecules according to the invention for use in processes according to the invention.

The object of the invention was accomplished by providing for a branched compound for use in nucleic acid detection and analysis reactions, comprising one linking moiety for attachment of the branched compound to a support at least one first branching moiety, and at least two nucleic acid moieties which are either directly or indirectly linked to the first branching moiety, whereby each of the nucleic acid moieties exhibits a free 3′-OH. The branched compound according to the invention preferentially has a first branching moiety that is formed by the linking moiety or a part thereof.

Also within the branched compound the nucleic acid moieties preferentially comprise a 5′end and a free 3′-OH so that the at least two nucleic acid moieties are linked to a second branching moiety which in turn is linked to the first branching moiety.

Thus the object of the present invention was accomplished by providing for a branched compound for use in nucleic acid detection and analysis reactions, as represented in formula 1,

formula 1:



Here, L_1 represents a linker moiety of variable length for attachment to a solid support, wherein B_1 represents at least one branching moiety, X_1 is one of at least two nucleic acid moieties, Y_1 represents the second of at least two nucleic acid moieties and the nucleic acid moieties may be extended by a polymerase.

Herein, a solid-support may be any support used to perform solid phase reactions. solid phase reactions and solid-support reactions are used with equal meaning and shall be understood as such reactions in which one or more compounds is attached to a solid matter of any given shape or chemical structure. Thus a solid-phase reaction is a reaction in which at least one compound may not move freely through diffusion in the reaction.

The inventors observe a large enhancement in reaction efficiency when applying the compound according to the invention to e.g. *in-vitro* nucleic acid detection and analysis reactions.

The inventors attribute this unexpected behaviour to the fact that two or more two nucleic acid moieties are in close spatial proximity.

The molecule according to the invention thus solves one of the problems in the prior art namely the availability of oligonucleotides for extension. Since the reaction product of *e.g.* X_1 may, in a subsequent cycle of a number of cycles of an in an *in-vitro* nucleic acid synthesis reaction, serve as a template for a reaction with Y_1 and since these are in close molecular vicinity, this reaction will run very efficiently. Also, the variable linker moiety L_1, makes it possible to overcome the drawbacks of known solid-phase reactions where the nucleic acids were in close vicinity to the solid-phase and thus cross-react with it to remain refrained from taking part in the actual reaction.

The compound according to the invention thus combines for the very first time a number of characteristics in itself which the inventors have found to be important for success in e.g. *in-vitro*

nucleic acid detection and analysis reactions. (i) The compound according to the invention provides for the possibility to define the distance from a given solid-support. This is accomplished through the linker moiety L_1 which also serves to couple the compound to the solid-support. (ii) The compound according to the invention provides for the possibility to bring in close vicinity two or more nucleic acids (e.g. X_1 and Y_1) involved in one reaction, thus solving the problem of a lack of diffusion of *e.g.* nucleic acids in *in-vitro* nucleic acid synthesis reactions. (iii) The compound according to the invention provides for nucleic acids X_1 and Y_1 which are not sterically hindered as the branching moiety B_1 permits them to become available in the reaction.

Figure 9B shall illustrate schematically some of the advantageous the branched compound according to the invention provides for.

In solid-phase, which is one way the molecule according to the invention may be used an efficient *in-vitro* nucleic acid synthesis reaction is dependant on either one of two events occurring, (i) a template molecule, whereby this stems from the sample with which the reaction was started, to reach a primer X_1, through diffusion or, (ii) a reaction product from a previous extension *e.g.* from X_1 to be in spatial vicinity and thus serve as a template for an extension by Y_1 (See also Fig. 9A and 9B).

It is obvious that the molecule according to the invention may comprise further branching moieties B_1 attached to one linker L_1 as well as more than 2 nucleic acids per branching moiety B_1. Herein, a branching molecule is defined as a molecule capable of binding 2 or more nucleic acids as well as a linker L_1.

For a better understanding we shall define the components of nucleic acid molecules as follows:

Deoxyribonucleic acid and ribonucleic acid are thread-like macromolecules, DNA comprising a chain of deoxyribnucleotides, and RNA comprising a chain of ribonucleotides. A "nucleotide" consists of a nucleoside and one or more phosphate groups; a "nucleoside" consists of a nitrogenous base linked to a pentose sugar. Typically, the phosphate group is attached to the fifth carbon ("5") hydroxyl group ("OH") of the pentose sugar; however, it can also be attached to the third-carbon hydroxy group ("3'-OH"). In a molecule of DNA, the pentose sugar is deoxy-ribose, in RNA, the pentose-sugar is ribose. The DNA bases are adenine ("A"), cytosine ("C"), guanine ("G"), and thymine ("T"). These base are the same for RNA except that uracil ("U") replaces thymine in natural RNA, i.e. RNA isolated from a living organism. Accordingly, the major nu-

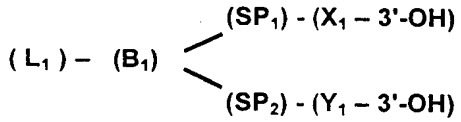
cleosides of DNA, collectively referred to as "deoxynucleosides", are as follows: deoxyadenosine ("dA"), deoxycytosine ("dC"), deoxyguanosine ("dG"), and deoxythymidine ("dT").

Herein, the abbreviations A, C, G and T shall refer to the DNA as well as RNA bases unless specified otherwise.

In this context nucleic acid moieties X_1 and Y_1, may be selected from the group comprising deoxynucleotides, dideoxynucleotides, ribonucleotides or modified nucleotides as specified above. If the term nucleotides is used, any of the above may be meant unless specified otherwise.

The inventors have found that it may, in some cases be advantageous to extend the distance between the branching molecule B and the oligonucleotides or oligonucleotide analogues by additionally adding one or more spacers to the molecule according to the invention. Here, the molecule additionally comprises spacers $\mathbf{SP_1}$ and $\mathbf{SP_2}$ with which the molecular distance between the free 3' OH ends of oligonucleotides or oligonucleotide analogues and the branch $\mathbf{B_1}$ may be defined. A possible structure is shown below in formula 2. Thus in a preferred embodiment of the present invention the branched compound additionally comprises two or more spacers $\mathbf{SP_1}$ and $\mathbf{SP_2}$ according to the following formula 2:

formula 2:



wherein $\mathbf{L_1}$ represents a linker moiety of variable length for attachment to a solid support, wherein $\mathbf{B_1}$ represents at least one branching moiety, wherein $\mathbf{X_1}$ is one of at least two nucleic acid moieties, wherein $\mathbf{Y_1}$ represent the second of at least two nucleic acid moieties, wherein the nucleic acid moieties may be extended by a polymerase and wherein $\mathbf{SP_1}$ and $\mathbf{SP_2}$ respectively represent spacer moieties which serve to define the distance between the branch moiety $\mathbf{B_1}$ and

the nucleic acid moieties X_1 **and** Y_1. The Inventors have found that these spacers enhance the reaction efficiency.

It is obvious that the spacers are optional and one may envision molecules where only one of the two oligonucleotides or oligonucleotide analogues are separated by a spacer from the branching molecule B_1. However, this is one of the essential features of the present invention as the variation of spacer length allows for adjustment to different template lengths and types.

Such spacers SP_1 and SP_2 are polymeric units comprising monomers. Although they may be chosen from the group comprising nucleotides, nucleotide analogues, abasic deoxyribose moieties, abasic ribose moieties, carbon chains which in turn may aliphatic or aromatic or even polyglycoles, they are not limited to these.

Also other molecules than those mentioned here may be used as a spacer as long as they fulfil the functional requirement of defining a distance between the branch and the free 3'-OH and do not hinder the polymerase from extending oligonucleotides or oligonucleotide analogue. The inventors have empirically found that a suitable distance between the 3'-OH group and the branch molecule lies between 3 and 350 angstrom, preferably between 30 and 120 angstrom. These values were established by testing different spacer types and calculating the resulting lengths from the molecule type.

In a preferred embodiment SP_1 and SP_2 are polymeric units comprising monomers chosen from the group comprising nucleotides, nucleotide analogues, abasic deoxyribose moieties, abasic ribose moieties, carbon chains.

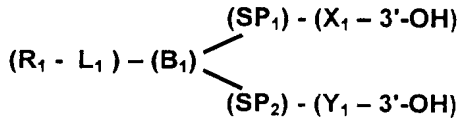
In an even more preferred embodiment of the invention the SP_1 and SP_2 are polymeric units comprising monomers of nucleotides or nucleotide analogues. Herein, nucleotides are to be understood as stretches of the same nucleotide or any given mixture of different nucleotides. The inventors have found between one and a hundred, more preferably between 10 and 30 to be suited, according to the distances described in more detail below.

The inventors have empirically found that a suitable distance between the 3'-OH group and the branch molecule lies between 3 and 350 angstrom, preferably between 30 and 120 angstrom.

These values were established by testing different spacer types and calculating the resulting lengths from the molecule type.

In a preferred embodiment of the invention the branched compound further comprising a reactive group (R_1) as represented by formula 3.

formula 3:



wherein L_1 represents a linker moiety of variable length for attachment to a solid support, wherein B_1 represents at least one branching moiety, wherein X_1 is one of at least two nucleic moieties, wherein Y_1 represent the second of at least two nucleic acid moieties, wherein the nucleic acid moieties may be extended by a Polymerase, wherein SP_1 and SP_2 respectively represent spacer moieties which serve to define the distance between the branch moiety B_1 and the nucleic acid moieties X_1 and Y_1.

The inventors have found such a a reactive group R_1 to greatly facilitate the binding of the branched compound according to the invention to a given solid-support. Such a reactive group may be selected from a group comprising amino group, hydroxyl group, carboxyl group, aldehyde group, thiol group, organo-silane group and phosphate group.

Other groups, not mentioned here, accomplishing the purpose of binding the molecule according to the invention to a solid-phase are equally suited and shall equally be considered to be covered by the invention.

In a preferred embodiment of the branched compound according to the invention characterised in that the linker moiety (L_1) of variable length consists of polymeric units selected from the group comprising, nucleotides, nucleotide analogues, abasic deoxyribose moieties, abasic ribose moie-

ties, carbon chains between about 3 carbons and 22 carbons, preferably between 3 carbons and 18 carbons in length, such chains may be aliphatic or aromatic, and polyglycoles.

Also other molecules than those mentioned here may be used as a linker moiety ($\mathbf{L_1}$) of variable length as long as they fulfil the functional requirement of defining a distance between the solid-support and the branch moiety. The inventors have empirically found that a suitable distance between the branch moiety and the solid-support lies between 3 and 350 angstrom, preferably between 30 and 120 angstrom. These values were established by testing different linker moieties and calculating the resulting lengths from the molecule type.

In an even more preferred embodiment of the invention the linker moiety ($\mathbf{L_1}$) of variable length consists of polymeric units comprising monomers of nucleotides or nucleotide analogues. Herein, nucleotides are to be understood as stretches of the same nucleotide or any given mixture of different nucleotides. The inventors have found between one and a hundred, more preferably between 10 and 30 to be suited, according to the distances described in more detail above.

As described above, two or more nucleic acid moieties ($\mathbf{X_1}$, $\mathbf{Y_1}$, $\mathbf{Z_1}$, etc.) are coupled to the branch moiety $\mathbf{B_1}$. It is an essential feature of the invention that at least to of these nucleic acid moieties carry a free 3' hydroxyl group and may thus be extended by a polymerase.

In a preferred embodiment of the invention X_1 and Y_1 consist of between 2 and 100 nucleotides in a more preferred embodiment of the invention between 10 and 30 nucleotides. Herein, oligonucleotide moiety will be used when the nucleic acid moieties consist of between 2 and 100 nucleotides.

Although it is possible that the two or more oligonucleotide moieties $\mathbf{X_1}$ and $\mathbf{Y_1}$ have the same sequence it is preferred if consist of different nucleotide sequences.

In a preferred embodiment the oligonucleotide moiety is characterised in that $\mathbf{X_1}$ and $\mathbf{Y_1}$ stem from the sense and anti-sense strand of a nucleic acid molecule. That means, if the were to anneal to a mixture of e.g. DNA, they would anneal to the plus and the minus strand respectively of a given DNA double helix.

In an even more preferred embodiment, the oligonucleotide moieties thus point towards each other with their free 3'-OH groups. Here, their positioning upon hybridisation to a DNA double

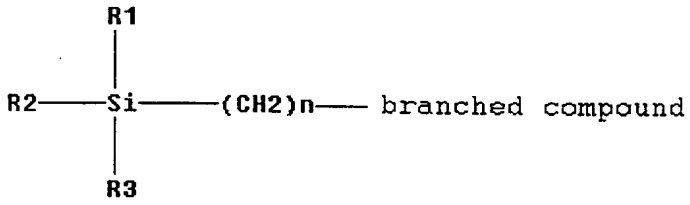
helix is so, that, a stretch of target DNA is located between the two positions of the respective oligonucleotide moieties X and Y.

The inventors have found that this embodiment leads to excellent results when the branched compound according to the invention is applied to solid-phase DNA synthesis reactions.

This embodiment is preferred but shall not limit the possibilities of designing the oligonucleotides which may in principle be designed in endless alternative ways of which the following are examples: The oligonucleotide moieties may (i) have the same sequence, (ii) have different sequences, (iii) have different sequences that hybridise to different targets, (iv) have different sequences that hybridise to the same target at different locations on the target, (v) have different sequences that hybridise to the plus and minus strand of one DNA double helix and many so on. One skilled in the art will readily recognise that there are numerous ways of designing the oligonucleotide moieties for numerous different purposes, all of which shall be covered by the scope of the invention.

As outlined above the reactive group R_1 serves to couple the branched compound to a solid-support. In a preferred embodiment the inventors have found it to be very advantageous if the reactive group is an organo-silane group. Formula four below depicts the entire molecule according to this preferred embodiment.

formula 4:

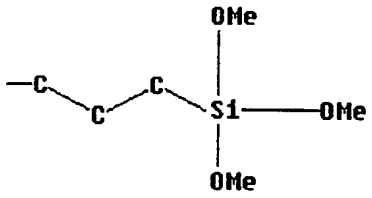


Here, R_1, R_2, and R_3 are identical or different alkoxy groups and "branched compound" refers to the entire compound according to the invention excluding the reactive group R1 and n is an integer from 0 to 18. This novel compound has surprisingly shown unexpected results when compared to similar molecules previously used in solid-phase enzymatic reactions with respect to but not limited to the following effects, its adsorption capacity to a solid-support, it's availability in enzymatic reactions, thus its contribution to the efficiency of *e.g.* solid-phase nucleic acid synthesis reactions.

The alkoxy groups R_1, R_2, and R_3 may, e.g. by methoxy, ethoxy or the like. Within the scope of the invention are organo-silane moieties comprising mixtures of different alkoxy groups. For example, R_1 may be a methoxy, R_2 an ethoxy and R_3 a methoxy. The skilled artisan is credited with the ability to discern alternative combinations which shall be within the scope of the invention.

On form of reactive group has proven to work particularly well here, **(R_1)** is a propyltrimethoxysilyl group according to the following formula 5,

formula 5:



In a preferred embodiment of the present invention the branched compound according to the invention is attached to a solid-phase. Here, the branched compound is connected to a support via the linking moiety L_1 of which the reactive group R_1 has undergone a reaction with a group on the surface of the solid-support to form a covalent or non-covalent bond.

Such solid-phase may be chosen from the group comprising nitrocellulose, nylon, controlled-pore glass beads (CPG), polystyrene, activated dextran, modified polystyrene, styrene-acrylnitril-copolymers, polycarbonate, cellulose, polyamide and glass.

In a particularly preferred embodiment of the invention the solid-phase is glass. Such glass may be a glass slide, as used *e.g.* for microscopy, glass vessels or containers, glass fibres, glass beads or other Si comprising glass entities.

The compound according to the invention may be spotted onto, pipetted onto, sprayed onto or otherwise brought onto such a glass support. Possible methods are, application by means of a

needle, capillary, dispenser and piezo pipette is preferable, *e.g.* an apparatus similar to the kind known for ink jet printers.

The compound according to the invention may be used in various ways some of which shall be mentioned here.

The compound is particularly suited and may thus be used in preferred embodiments of the invention for nucleic acid synthesis reactions, primer extension reactions, reverse transcription reactions of RNA into DNA or nucleic acid hybridisation experiments (see also Figures 1 to 3).

In a preferred embodiment the compound is used in *in-vitro* nucleic acid synthesis reactions and hybridisation experiments. Here, the compound may be bound to a solid support such as glass. The compound represents one or more nucleic acid probes to which a target, *i.e.* the sample is bound.

In a preferred embodiment it may be used in nucleic acid hybridisation or synthesis reactions for identifying the sequence of a nucleic acid.

In a further preferred embodiment the branched compound may be used in nucleic acid hybridisation or synthesis reactions for analysing the expression pattern of genes. Here, oligonucleotides or nucleic acid fragments, *e.g.* PCR products represent the nucleic acid moieties according to the invention which are attached via the rest of the branched compound to a suited solid-support to form an array of various probes, labelled RNA or pre-amplified RNA is hybridised to the array and the positive hybridisations scored as expressed genes. Such methods are described in GB Pat. 2318791.

The compound according to the invention may be used to distinguish single base mismatches. The general principle of such an approach is disclosed in U.S. Pat. 5,700,638 where, such experiments are described in Example 2. US Pat. 5,552,270 also describes such an approach. Here, the compound according to the invention comprises an oligonucleotide of defined sequence. An array is generated comprising numerous different sequences each suited to test a defined sequence.

The branched compound according to the invention may be used to map genomes of organisms. Here, oligonucleotides or nucleic acid fragments, e.g. PCR products represent the nucleic acid

moieties according to the invention which are attached via the rest of the branched compound to a suited solid-support to form an array of various probes, labelled genomic fragments are sequentially hybridised in numerous steps in such a way that the relationship between individual fragments becomes apparent. Such an approach is disclosed in U.S. Pat. 5,219,726.

One skilled in the art will find numerous further uses for the branched compound according to the invention all which shall equally be covered by the invention.

The invention shall also cover a process for a nucleic acid synthesis reaction of one or more selected regions of one or more target nucleic acids comprising the steps of a) combining the sample containing the target region with at least on nucleotide triphosphate, a thermally stable polymerase, a buffer and at least one branched compound according to the invention, b) exposing the reaction mixture to at least one temperature cycle including at least a high temperature denaturation phase and a lower temperature extension phase, and thereby producing at least a partially amplified product.

The inventors have found that the branched compound according to the invention gives unexpectedly good results when used on solid-phase in *in-vitro* DNA synthesis reactions (see also examples).

Thus in a preferred embodiment of the above process at least one branched compound according to the invention is bound to a solid-support.

In this process according to the invention it is preferred that the branched compound according to the invention comprises preferably between 3 and 100 nucleotides more preferably between 10 and 50 nucleotide most preferably between 12 and 25 nucleotides for each nucleic acid moiety. It is essential in this context that the nucleic acid moiety is of sufficient length to bind the desired target nucleic acid molecule with sufficient stringency. PNA-DNA hybrid oligonucleotides (see Finn, P. J. et al., N.A.R. 24, 3357-3363 (1996), Koch, T. et al., Tetrahedron Letters 36, 6933-6936 (1995), Stetsenko, D. A. et al., Tetrahedron letters 37, 3571-3574 (1996), Bergmann, F et al., Tetrahedron Letters 36 6823-6826 (1995) and Will, D. W. et al., Tetrahedron 51, 12069-12082 (1995)) are also considered as primers for the process of the present invention.

As a thermally stable polymerase, a DNA polymerase may be selected from the group comprising Taq DNA polymerase, Tth DNA polymerase or Klentaq (Taq DNA polymerase (-exo5'-3'),

Korolev et al., (1995) Proc. Natl. Acad. Sci. USA 92, 9246-9268. The use of Taq DNA polymerase in the method of the present invention is especially preferred. One skilled in the art will recognise that numerous different thermally stable polymerases may be used.

Alternatively, as a thermally stable polymerase, a DNA polymerase which has a decreased discrimination against the four ddNTPs with respect to wild-type Taq DNA polymerase in the buffer or under the conditions used for thermal cycling is preferred. More preferably, a DNA polymerase Taq polymerase carrying a "Tabor-Richardson" mutation or a functional derivative thereof which also lacks 5'-3' exonuclease activity such as, for example, AmplitaqFS™ (Taq DNA polymerase (-exo5'-3')(F667Y), Tabor and Richardson (1995), loc. cit.), Taquenase™ (Taq DNA polymerase Δ235(-exo5'-3')(F667Y), Tabor and Richardson (1995), loc. cit.) and Thermo-Sequenase™ (Taq DNA polymerase (-exo5'-3')(F667Y), Tabor and Richardson (1995), loc. cit.) as well as mixtures thereof or other DNA polymerases and mixtures thereof which are thermally stable can be used in the process of the present invention. The use of Thermo Sequenase™ or any other DNA polymerase having a high ability to incorperate ddNTPs in the method of the present invention is especially preferred.

Alternatively as a thermally stable polymerase, a DNA polymerase which has a decreased discrimination against labelled nucleotides may be used.

The present invention, *i.e.* the process also provides for the use of two or more polymerases in the process or additional enzymes such as amplification enhancing reagents such as thermally stable pyrophosphatase or enzymes which enhance the processivity of the polymerase such as PCNA (proliferating cell nuclear antigen) homologues. Enzyme mixtures may be equally applied.

The number of thermal cycles may range from about 1 to about 50 depending on the amount of template DNA and its purity. Generally, the inventors have found that very surprisingly extremely short cycles give good results. As the availability of the compound according to the invention is high in the process according to the invention the cycle period may be short, thus disadvantageous denaturing of proteins, *e.g.* the polymerase when in contact with glass occurs at a lower rate and the reaction may run efficiently without loss of function of enzyme.

Routinely, cycling consists of (i) a denaturing cycle, (ii) an annealing cycle and (iii) an extension cycle. Alternatively, only two cycles may be applied, (i) a denaturing cycle and (ii) an annealing and extension cycle.

Preferably the denaturing cycle is performed at between 100°C and 85°C, more preferably at between 98°C and 90°C, most preferably at between 96°C and 92°C.
Preferably the annealing cycle is performed at between 80°C and 45°C, more preferably at between 70°C and 50°C, most preferably at between 60°C and 55°C.

Preferably the extension cycle is performed at between 80°C and 50°C, more preferably at between 75°C and 60°C, most preferably at between 73°C and 68°C.

Preferably the denaturing cycle is performed for 3 minutes, more preferably for 30 seconds, most preferably for under 10 seconds.

Preferably the annealing cycle is performed for 3 minutes, more preferably for 30 seconds, most preferably for under 10 seconds.

Preferably the extension cycle is performed for 3 minutes, more preferably for 30 seconds, most preferably for under 10 seconds, however the extension time vary depending on the length of the template, in particular the extension time may be raised if the template length increases.

Buffers components which can be used can include, but are not limited to, Tris-HCl at a pH of about 7.0 to 10 and concentration of about 2 to 60 mM, ammonium sulfate at a concentration of about 10-20 mM, preferably 15 mM, $MgCl_2$ at a concentration of about 1 to 10 mM, and optionally, about 0.05 mM mercaptoethanol, about 0.28% Tween® 20 and/or about 0.02% Nonidet® 40.

Nucleotide triphosphates are preferably deoxynucleotides and can be selected from, but are not limited to, dGTP, dATP, dTTP and dCTP. In addition, derivatives of deoxynucleotides, which are defined as those deoxynucleotides which are capable of being incorperated by a thermally stable DNA polymerase into nascent DNA molecules synthesized in the thermal cycling reaction, can also be used according to the invention. Such derivatives include, but are not limited to thionucleotides, 7-deaza-2'-dGTP, 7-deaza-2'-dATP as well as deoxyinosine triphosphate which

may also be used as a replacement deoxynucleotide for dATP, dGTP, dTTP or dCTP. The above mentioned deoxynucleotides and derivatives thereof are preferably used at a concentration of about 50 μM to about 4 mM.

Preferably the nucleotides are mixes of all four nucleotides A, C, G, and T and at 200 μM each.

In a preferred embodiment one or more of the nucleotides incorporated are labelled. For example, single and differential labels may consist of the group comprising enzymes such as β-galactosidase, alkaline phosphatase and peroxidase, enzyme substrates, coenzymes, dyes, chromophores, fluorescent, chemiluminescent and bioluminescent labels such as FITC, Cy5, Cy5.5, Cy7, Texas-Red and IRD40(Chen et al. (1993), J. Chromatog. A 652: 355-360 and Kambara et al. (1992), Electrophoresis 13: 542-546), ligands or haptens such as biotin, and radioactive isotopes such as ^{3}H, ^{35}S, ^{32}P ^{125}I and ^{14}C.

In one embodiment of the method of the invention, the nucleic acid molecule to be amplified can be present in the form of total genomic DNA, which is preferably in an uncloned or unpurified form. Preferably, the genomic DNA has a length greater than or equal to 2 kb. Generally all forms of template may be used, *e.g.* purified nucleic acids, *i.e.* nucleic acids where one fraction may be enriched or not, one example being plasmid DNA the other purified genomic DNA. The process may be suited for use with complex mixtures of DNA such being purified but not substantially fractionated genomic DNA or non-complex mixtures such being purified and substantially fractionated DNA e.g. plasmid DNA.

In a further preferred embodiment of the method of the invention, the nucleic acid molecule to be amplified can be present in the form of RNA. One polymerase or a mixture of two polymerases maybe utilised: a first DNA polymerase for example, Taq polymerase, and a second DNA polymerase with the capability to reverse transcribe RNA into DNA preferably Taq DNA polymerase (Jones et al., Nucl. Acids Res. 17: 8387-8388 (1989)) or Tth DNA polymerase (Myers et al., Biochemistry 30: 7666-7672 (1991)).

While the foregoing has been set forth in detail, the Examples are presented for elucidation, and not limitation. Modifications and improvements on the compound and or process according to the invention disclosed above which are within the purview and abilities of those in the art are included within the scope of the claims.

The invention also covers a kit for use in molecular biology or chemistry comprising at least the compound according to the invention. The kit may also comprise other reagents or enzymes such as buffers, nucleotides or the like.

While the foregoing has been set forth in detail, the Examples are presented for elucidation, and not limitation. Modifications and improvements on the compound and or process according to the invention disclosed above which are within the purview and abilities of those in the art are included within the scope of the claims.

Example 1:

DNA Oligonucleotides were synthesised using standard cyanoethyl posphoramidite chemistry on an ABI 392 DNA/RNA Synthesizer (Caruthers MH, Barone AD, Beaucage SL, Dodds DR, Fisher EF, McBride LJ, Matteucci M, Stabinsky Z, Tang JY Chemical synthesis of deoxyoligo-nucleotides by the phosphoramidite method, Methods Enzymol 154, 287-313 (1987)). Primary amino groups with an hexyl carbon atom spacer were introduced using the N-MMT-C6-aminomodifier (Cruachem Cat. No. 22-8401-17). Oligonucleotides were synthesized using the trityl-on mode and purified via HPLC (BioCad, PE Biosystems) using Oligo R3 reversed phase chromatography media (PE Biosystems Cat. No. 1-1339-06) and an on-column detritylation standard protocol. The oligonucleotides contained specific 20 or 19 nucleotides together with a T_9 or T_{15} spacer and were designed for the amplification of a 223 bp fragment of the polylinker sequence of pBlueScript KS + (Stratagene Cat. No. 212207).

Primer sequences:
SEQ ID NO 1:
"BSfor": 5'-NH_2-C6-TTTTTTTTTAAGCGCGCAATTAACCCTCA-3'-OH
SEQ ID NO 2:
"BSrev": 5'-NH_2-C_6-TTTTTTTTTTTTTTTAAAACGACGGCCAGTGAGC-3'-OH

The single branched amplification molecule was synthesised using cyanoethyl posphoramidite chemistry on an ABI 392 DNA/RNA Synthesizer. The 3' terminal amino group was introduced by starting synthesis with an aminolinker column (Cruachem Cat. No. 19-7850-80). The 5'-branch was introduced using either symmetric or assymmetric branching CE phosphoramidites (Cruachem Cat. No. 22-8424-17 or 22-8425-17). Reactive carboxyl groups at the 5'-ends were

introduced using Carboxy-dT CE phosphoramidite (Glen Research Cat.No. 10-1035-02), while primary amino groups were introduced using the N-MMT-C6-aminomodifier (Cruachem Cat. No. 22-8401-17).

Nucleic acid sequences:

SEQ ID NO 3:

Branch1 $(5'\text{-}NH_2\text{-}C_6\text{-}TTT)_2$ – Branch - TTTTTTTTTTTTTTT-3'-NH_2

SEQ ID NO 4:

Branch2 $(5'\text{-dT-}COOH)_2$ – Branch - TTTTTTTTTTTTTTT-3'-NH_2

The aminomodified branch (2 nmol) was incubated with the aminomodified oligonucleotides in 0.2 % (w/v) 1,4-phenylene-diisothiocyanate (PITC) in pyrrimidine/dimethylformamide 1/9 in a 100 µl volume at ambient temperature for 2 hours or longer. Purification was then achieved by precipitation with 0.3 M sodium acetate / ethanol and electrophoretic separation on a 15 % polyacrylamide / 8 M urea gel. (Size fractionation: 21 nt + 29 nt + 34 nt = 84 nt)

The carboxymodified branch (2 nmol) was incubated with the aminomodified oligonucleotides in molar excess. The reaction was accomplished by incubation of the carboxy and amino components in 100 µl aqueous solution containing 9 mg of the peptide condensing agent 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide (EDC) 6 hours at 4 °C. Purification was then achieved by precipitation with 0.3 M sodium acetate / ethanol and electrophoretic separation on a 15 % polyacrylamide / 8 M urea gel. (Size fractionation: 17nt + 29nt + 34nt = 80nt)

Purified single branched amplification molecules containing the two covalently bound primers were attached via their 3'-amino group to an amino-reactive solid support *i.e.* silylated slides (Cell Associates, Cat. No. CSS-100). The single branched amplification molecules were resuspended in 100 mM sodium carbonate buffer at pH 9.0 at a concentration of 100 µM. During the arraying process using a GMS 417 Arrayer (Genetic MicroSystems) which was adjusted to 30 hits per dot (*i.e.* spot) up to 0.75 pmoles of oligonucleotide were immobilised in spots of 150 µm diameter on the glass support. After spotting slides were incubated for 4 hours in a humid chamber to allow rehydration of the DNA, washed once in 0.2% SDS, twice in water, and incubated for 5 minutes in 0.25 % sodium borohydride in phosphate buffered saline for blocking free aminoreactive groups of the glass slide. Arrays are then submerged in water 2 minutes at 95 °C, transferred to 0.2% SDS for 1 minute and rinsed twice in water. After drying, arrays can be stored in the dark at ambient temperature.

Amplification reactions were performed in a 25 µl volume containing 50 mM Tris-HCl at pH 8.5, 30 mM KCl, 3 mM magnesium chloride, 12.5 µg bovine serum albumine, 200µM each dGTP, dATP and dTTP, 100 µM dCTP, 4 µM Cy5-dCTP (Amersham Pharmacia Biotech, Cat. No. PA 55021), 0.1 unit Taq DNA Polymerase and 3 fmole of template DNA (pBlueScript KS + plasmid DNA). The reaction chamber was formed by sealing the reaction droplet on the slide with FrameSeal Slide Chambers (Biozym Diagnostik, Cat. No. 621741). The silde is then attached to a modified frame (originally designed for holding glass capillaries) which fits to the reaction chamber of the thermal-cycler device (Idaho Technology, Thermo-Cycler 1605). Reactions are cycled with the following profile one initial cycle at 94°C for 30 seconds, thirty cycles for 10 seconds at 94 °C, for 15 seconds at 60°C, for 25 seconds at 74 °C and one final cycle for 30 seconds at 74°C.

After the incubation in the thermal-cycler slides were removed and the frame seals lifted off. The slides were transferred to a micro-array washing device (Telechem International, ArrayIt Microarray Wash Station) and rinsed with 0.2 % SDS for 15 minutes, cleaned in water for 15 minutes and dried at ambient temperature.
Scanning of the fluorescent PCR products was accomplished in a fluorescent scanning device (Genetic MicroSystems, Array Scanner).

Example 2:

Oligonucleotides were synthesised using standard cyanoethyl posphoramidite chemistry on an ABI 392 DNA/RNA Synthesizer. Primary amino groups with an hexyl carbon atom spacer were introduced using the N-MMT-C6-aminomodifier (Cruachem Cat. No. 22-8401-17). Oligonucleotides were synthesized using the trityl-on mode and purified via HPLC (BioCad, PE Biosystems) using Oligo R3 reversed phase chromatography media (PE Biosystems Cat. No. 1-1339-06) and an on-column detritylation standard protocol.

The purified oligonucleotide or single branched amplification molecule containing the 5' hexylamino modification (100 pmoles) was incubated in 100 µl of a 100 mM 1,4-phenylene diisothiocyanate (DITC) containing solution of pyrridine : dimethylformamide (ratio 1 : 9) for 2 hours at ambient temperature

After standard precipitation with ethanol/sodium acetate the oligonucleotide was dissolved in 100 µl of a 1 % 3-aminopropyltrimethoxysilane solution in 95 % acetone/water and incubated for 2 hours at ambient temperature (Reaction B). After standard precipitation the oligonucleotide bearing the glass-reactive silane group was re-dissolved in 95 % aqueous acetone and stored at 4 °C. Attachment to various supports, such as glass was accomplished by spotting (Reaction C).

Example 3:

The synthesis of branched oligonucleotides was performed using an automated DNA synthesizer and reagents commercially available from Glen Research (Virginia, USA) following the manufacturer's recommendations for dilution, coupling conditions and deprotection.

1) The synthesis was started using standard 3'-cyanoethyl phosphoramidite chemistry and a 3'-amino C7 modifier column with CPG 1000 Å support (Glen Research Cat. No. 20-2958-xx). 15 T-residues were synthesised that way in 3' -> 5' direction followed by addition of an asymmetric doubler phosphoramidite (Glen Research cat. No. 10-1921-xx) containing two different protecting groups, namely a DMT and a Fmoc group, generating a branch suitable for the consecutive addition of two different sequences (*i.e.* the two different DNA primers in our case).

2) The 5' DMT group of the branch was deprotected under acid conditions (normally with trichloroacetic acid) to generate a free 5'-hydroxyl in order to start synthesis of the first specific DNA oligonucleotide primer.

3) After changing to (non-standard) 5'-cyanoethyl phosphoramidite chemistry (Glen Research, Cat. No. 10-0001-xx, 10-0101-xx, 10-0201-xx, and 10-0301-xx) the first specific oligonucleotide was synthesised in 5' -> 3' direction.

4) The free 3'-hydroxyl was protected by acetylation, which was easily achieved by addition of an extra "capping" step during the synthesis cycle.

5) The Fmoc protecting group of the branch was removed under basic conditions (*e.g.* with 20% piperidine in DMF) in order to start synthesis of the second specific DNA oligonucleotide primer.

6) The second specific oligonucleotide was synthesised, also in 5' -> 3' direction.

7) After deprotection and removal from the solid support with standard ammonia treatment the Y-shaped oligonucleotide contained three 3'-ends, of which two were the functional 3'-hydroxyl groups of the two different DNA primers and one contained an C7-amino modification for coupling to the silane moiety according to the invention.

Figure Captions:

Fig. 1:

Fig. 1 shows a schematic representation of a solid-phase *in-vitro* nucleic acid synthesis reaction. The figure shows a oligonucleotide wherein, the oligonucleotide as an arrow and an organo-silane moiety as zig-zag line, both coupled via the organo-silane moiety to a solid support (G). Here, an *in-vitro* nucleic acid synthesis reaction is schematically shown. A target molecule (H) anneals to the oligonucleotide (B) where it is used as a template and oligonucleotide is extended by a polymerase. Thereafter, the extended oligonucleotide may serve as a template for a further oligonucleotide (C) which must be in sufficiently close vicinity. The oligonucleotide (C) now uses the template (D) and is extended to the full-length product (E). As schematically shown here, efficient synthesis relies on equimolarity and close vicinity of two or more different oligonucleotide required in one reaction.

Fig. 2:

Fig. 2 shows a schematic representation of a solid-phase *in-vitro* nucleic acid synthesis reaction. The figure shows a branched compound according to the invention wherein, the branched compound comprises two free 3'-OH groups and an organo-silane moiety as zig-zag line, both coupled via the organo-silane moiety to a solid support (G). Here, an *in-vitro* nucleic acid synthesis reaction is schematically shown. A target molecule (H) anneals to the nucleic acid moiety of the branched compound (B) where it is used as a template and the nucleic acid moiety is extended by a polymerase. Thereafter, the extended nucleic acid moiety may serve as a template for a further nucleic acid moiety (C) which is by definition in sufficiently close vicinity. The nucleic acid moiety (C) now uses the template (D) and is extended to the full-length product (E).

Fig. 3:

Fig. 3 shows a schematic representation of a solid-phase *in-vitro* nucleic acid synthesis reaction starting from mRNA including the step of prior reverse transcription coupled to an amplification step. The figure shows a branched compound according to the invention wherein, the branched compound comprises two free 3'-OH groups and an organo-silane moiety as zig-zag line, both coupled via the organo-silane moiety to a solid support (G). Here (A to F), an *in-vitro* nucleic acid reverse transcription with coupled amplification reaction is schematically shown. A target mRNA molecule (H) anneals to the nucleic acid moiety (B) where it is used as a template and the nucleic acid moiety is extended by a polymerase with reverse transcriptase activity. Thereafter, the extended nucleic acid moiety may serve as a template for a further nucleic acid moiety (C) which is by definition in sufficiently close vicinity. The nucleic acid moiety (C) now uses the template (D) and is extend to the full-length product (E). As schematically shown here, efficient synthesis relies on equimolarity and close vicinity of two or more different nucleic acid moiety required in one reaction. The compound according to the invention takes place in reverse transcription as well as subsequent amplification.

Fig. 4:

Fig. 4 shows a comparison between binding amino-modified oligonucleotides to silane-treated glass (panels A, B) and binding oligonucleotides carrying a terminal silane group to normal glass (panels C, D). In a preferred embodiment of this invention, the compound according to the invention carries such a organo-silane group as shown here in combination with a standard oligonucleotide. In a preferred embodiment of the invention, the reactive group R_1 is an organo-silane group and thus the compound according to the invention will spot as in this example. The 3' ^{32}P-labeled oligonucleotides (10.000 cpm) were spotted onto the respective glass slides and scanned (panels A, C). As a control also unmodified oligonucleotides were spotted onto both types of slides (not shown, see Table) to determine unspecific background binding. After extensive washing and rinsing steps the slides were dried and scanned again (panels B, D). Herein, cev (counts equivalent value) were determined by quantification of the Phosphor Imager gel scans using the ImageQuant software.

Fig. 5:

In a preferred embodiment of this invention, the branched compound according to the invention carries such a organo-silane group as shown here in combination with a standard oligonucleotide. In a preferred embodiment of the invention, the reactive group R_1 is an organo-silane group and thus the compound according to the invention will spot as in this example. Fig. 5 shows an experiment for the determination of the binding capacity of an organo-silane modified oligonucleotide to normal untreated glass. 1 nmol of an amino-modified oligonucleotide was spiked with 1 fmol of gel-purified 3'-end labelled oligonucleotide (10.000 cpm) and derivatized with an organo-silane according to the invention, spotted four times onto the slides and scanned (panel A). The slides were then washed, dried and scanned again (panel B). 98,7 % of the silane-treated oligonucleotides were covalently bound to the glass support.

Fig. 6:

Fig. 7 shows schematically the reactions leading to the derivatization of an oligonucleotide with an organo-silane moiety according to a preferred embodiment of the invention. An oligonucleotide containing a 5' hexylamino modification is first reacted with 1,4-phenylene diisothiocyanate (reaction A) and further reacted with 1, 3-aminopropyltrimethoxy-silane (reaction B). The compound bearing the glass-reactive silane group can then be attached to *e.g.* glass supports (reaction C).

Fig. 7:

Fig. 7 shows a preferred embodiment of the compound according to the invention. Here, the nucleic acid moieties were attached separately.

Fig. 8:

Fig. 8 shows a preferred embodiment of the compound according to the invention. Here, the nucleic acid moieties were attached separately using two identical bifunctional groups.

Fig. 9 A:

Fig. 9 A shows a prior-art nucleic acid synthesis on solid-phase (E) . One must note, that template (A) is supplied through diffusion only. In the even that, the oligonucleotides of sense (C) and antisense (D) are not in sufficient vicinity or are sterically hindered (not shown here) they may not serve as template upon elongation. Thus, once the oligonucleotide (B) is elongated factors such as concentration, steric hindrance etc. will decide whether it may serve as a template in a subsequent cycle.

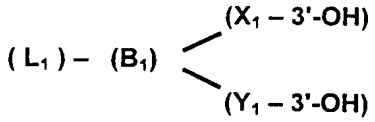
Fig. 9 B:

Fig. 9 B shows a nucleic acid synthesis process on solid-phase (G) according to the invention. Once diffusion template (H) reaches any of the nucleic acid moieties, the may serve themselves as a template (C and D) or neighbouring branched compounds (E and F) may serve as templates. Thus, the process according to the invention overcomes fundamental problems connected to the prior-art.

Claims

1. Branched compound for use in nucleic acid detection and analysis reactions,
as represented in formula 1,

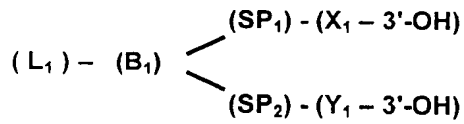
formula 1:



wherein $\mathbf{L_1}$ represents a linker moiety of variable length for attachment to a solid support,
wherein $\mathbf{B_1}$ represents at least one branching moiety,
wherein $\mathbf{X_1}$ is one of at least two nucleic acid moieties,
wherein $\mathbf{Y_1}$ represents the second of at least two nucleic acid moieties,
wherein $\mathbf{X_1}$ and $\mathbf{Y_1}$ are oligonucleotide moieties and stem from the sense and anti-sense strand of a nucleic acid molecule and

2. Branched compound according to claim 1,
as represented in formula 2,

formula 2:

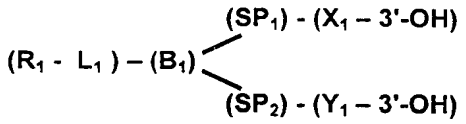


wherein $\mathbf{L_1}$ represents a linker moiety of variable length for attachment to a solid support,
wherein $\mathbf{B_1}$ represents at least one branching moiety,
wherein $\mathbf{X_1}$ is one of at least two nucleic acid moieties,

wherein Y_1 represent the second of at least two nucleic acid moieties,
wherein X_1 and Y_1 are oligonucleotide moieties and stem from the sense and anti-sense strand of a nucleic acid molecule, and
wherein SP_1 and SP_2 respectively represent spacer moieties which serve to define the distance between the branch moiety B_1 and the nucleic acid moieties X_1 and Y_1.

3. Branched compound according to claim 1 or 2,
further comprising a reactive group **(R_1)** as represented by formula 3,

formula 3:



wherein L_1 represents a linker moiety of variable length for attachment to a solid support,
wherein B_1 represents at least one branching moiety,
wherein X_1 is one of at least two nucleic moieties,
wherein Y_1 represent the second of at least two nucleic acid moieties,
wherein X_1 and Y_1 are oligonucleotide moieties and stem from the sense and anti-sense strand of a nucleic acid molecule,
wherein SP_1 and SP_2 respectively represent spacer moieties which serve to define the distance between the branch moiety B_1 and the nucleic acid moieties X_1 and Y_1, and
wherein R_1 is a reactive group selected from a group comprising amino group, hydroxyl group, carboxyl group, aldehyde group, thiol group, organo silane group and phosphate group

4. Branched compound according to any of claims 1 to 3, wherein the at least two nucleic acid moieties are linked to a second branching moiety, which in turn is linked to the first branching moiety B_1.

5. Branched compound according to any of claims 1 to 4, characterised in that SP_1 and SP_2 are polymeric units comprising monomers chosen from the group comprising nucleotides,

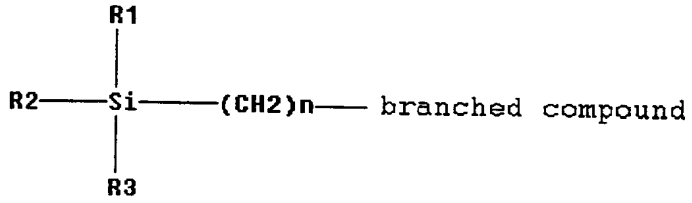
nucleotide analogs, abasic deoxyribose moieties, abasic ribose moieties, carbon chains and polyglycoles.

6. Branched compound according to any of claims 1 to 5, characterised in that the linker moiety of variable length (L_1) consists of polymeric units selected from the group comprising, nucleotides, nucleotide analogs, abasic deoxyribose moieties, abasic ribose moieties, carbon chains and polyglycoles.

7. Branched compound according to any of claim 1 to 6, characterised in that X_1 and Y_1 consist of between 2 and 100 nucleotides, preferably between 10 and 30 nucleotides.

8. Branched compound according to any of claims 3 to 7, characterised in that (R_1) is a terminal organo-silane group and the entire compound is thus structured according to formula 4:

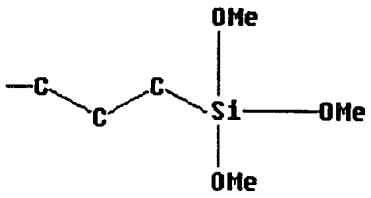
formula 4:



wherein, R_1, R_2, and R_3 are identical or different alkoxy groups, wherein
the term "branched compound" is the compound according to the above claims excluding the reactive group R_1 and
n is an integer from 0 to 18.

9. Branched compound according to any of claims 3 to 8,
characterised in that (R_1) is a propyltrimethoxysilyl group according to the following formula 5,

formula 5:



10. Branched compound according to any of claims 3 to 9, characterised in that it is attached to a solid phase with the reactive group (R_1).

11. Branched compound according to claim 10, characterised in that the solid phase is chosen from the group comprising nitrocellulose, nylon, controlled-pore glass beads (CPG), polystyrene, activated dextran, modified polystyrene, styrene-acrylnitril-copolymers, polycarbonate, cellulose, polyamide and glass.

12. Branched compound according to claim 11, characterised in that the solid phase is glass.

13. Use of a branched compound according to any of claims 1 to 12 in nucleic acid synthesis reactions, primer extension reactions, reverse transcription reactions of RNA into DNA, nucleic acid hybridisation experiments, particularly in nucleic acid hybridisation or synthesis reactions for identifying the sequence of a nucleic acid or in nucleic acid hybridisation or synthesis reactions for analysing the expression pattern of genes.

14. Process for a nucleic acid synthesis reaction of one or more selected regions of one or more target nucleic acids comprising the steps of

a) combining the sample containing the target region with at least on nucleotide triphosphate, a polymerase, a buffer and at least one branched compound according to any of claims 1 to12,

b) exposing the reaction mixture to at least one temperature cycle including at least a high temperature denaturation phase and a lower temperature extension phase, and thereby producing at least a partially amplified product.

15. Process according to claim 14, wherein the compound according to the invention is bound to a solid support.

16. Process according to claims 14 and 15, characterised in that

a) an in-vitro nucleic acid reverse transcription with

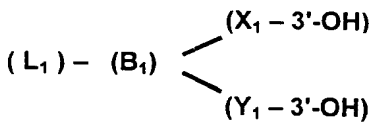
b) a coupled amplification reaction is performed.

17. Process according to any of claims 14 to 16, characterised in that the buffer used contains between Bovine Serum Albumine (BSA) at a concentration of between 0.1 µg per µl and 5 µg per µl, preferably at a concentration of about 0.5 µg per µl.

18. Kit, comprising the branched compound according to any of the claims 1 to 12.

19. Process for the synthesis of a branched compound for use in nucleic acid detection and analysis reactions, as represented in formula 6,

formula 6:

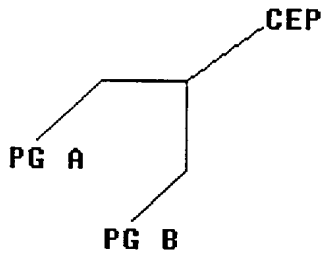


wherein $\mathbf{L_1}$ represents a linker moiety of variable length for attachment to a solid support,
wherein $\mathbf{B_1}$ represents at least one branching moiety,
wherein $\mathbf{X_1}$ is one of at least two nucleic acid moieties,
wherein $\mathbf{Y_1}$ represents the second of at least two nucleic acid moieties, and
wherein $\mathbf{X_1}$ and $\mathbf{Y_1}$ are different oligonucleotide moieties,
characterised in that,
in a first step a linker moiety $\mathbf{L_1}$ is supplied or synthesised,
in a second step at least one branching moiety $\mathbf{B_1}$ is reacted with or synthesised onto the linker reagent,
in a third step the first of at least two nucleic acid moieties $\mathbf{X_1}$ is synthesised separately and subsequently bound to the branching moiety $\mathbf{B_1}$,

in a fourth step the second of at least two nucleic acid moieties Y_1 is synthesised separately and subsequently bound to the branching moiety B_1,

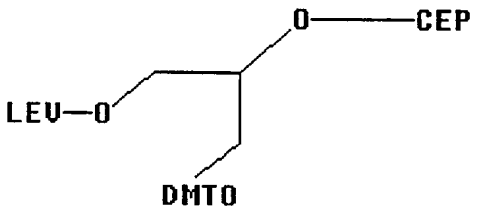
20. Process according to claim 19, characterised in that X_1 and Y_1 are derived from the sense and anti-sense strand of a nucleic acid molecule.

21. Process according to claim 19 or 20, characterised in that the branching moiety **(B_1)** is an asymmetric branching moiety with the following structure,



wherein PG A is a protecting group chosen from the group comprising LEV (levulinyl), DMTO (trityl), MMT (monomethoxytrityl),
wherein PG B is a protecting group chosen from the group comprising LEV (levulinyl), DMTO (trityl), MMT (monomethoxytrityl),
wherein CEP stands for cyanothylphosphoramidite and
wherein PG A and PG B may not be identical.

22. Process according to any of claims 19 to 21, characterised in that the branching moiety **(B_1)** is an asymmetric branch with the following structure,



wherein LEV stands for the protecting group levulinyl,
wherein DMTO stands for the protecting group trityl and
wherein CEP stands for cyanoethylphosphoramidite.

23. Branched compound, produced according to a process according to any of claims 19 to 22.

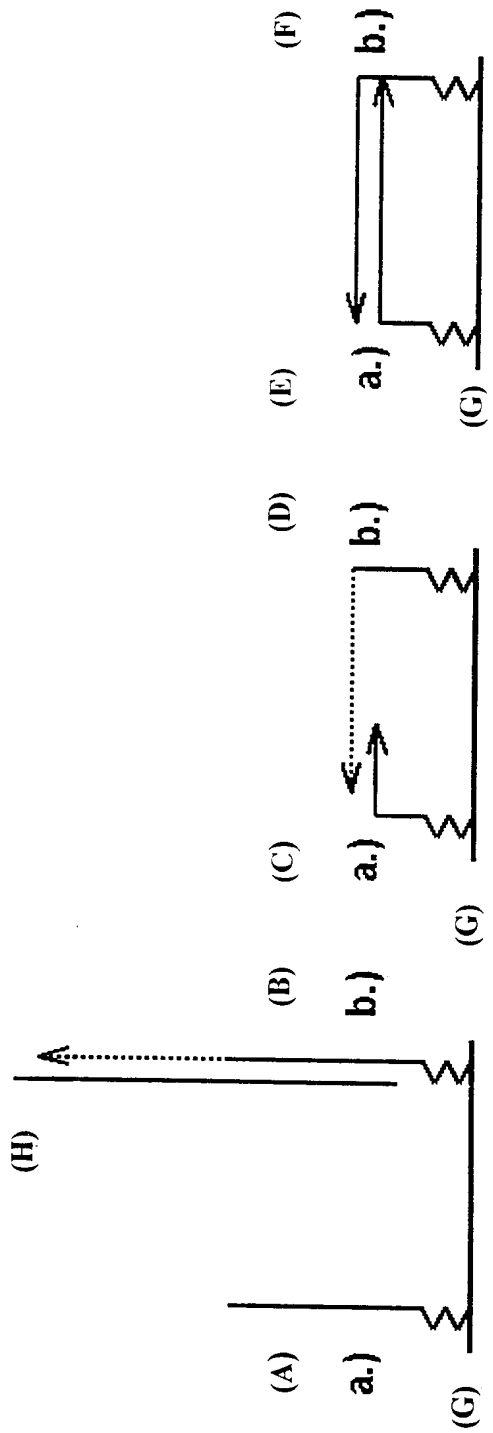


Fig. 1

IDT Exhibit 1020.6123

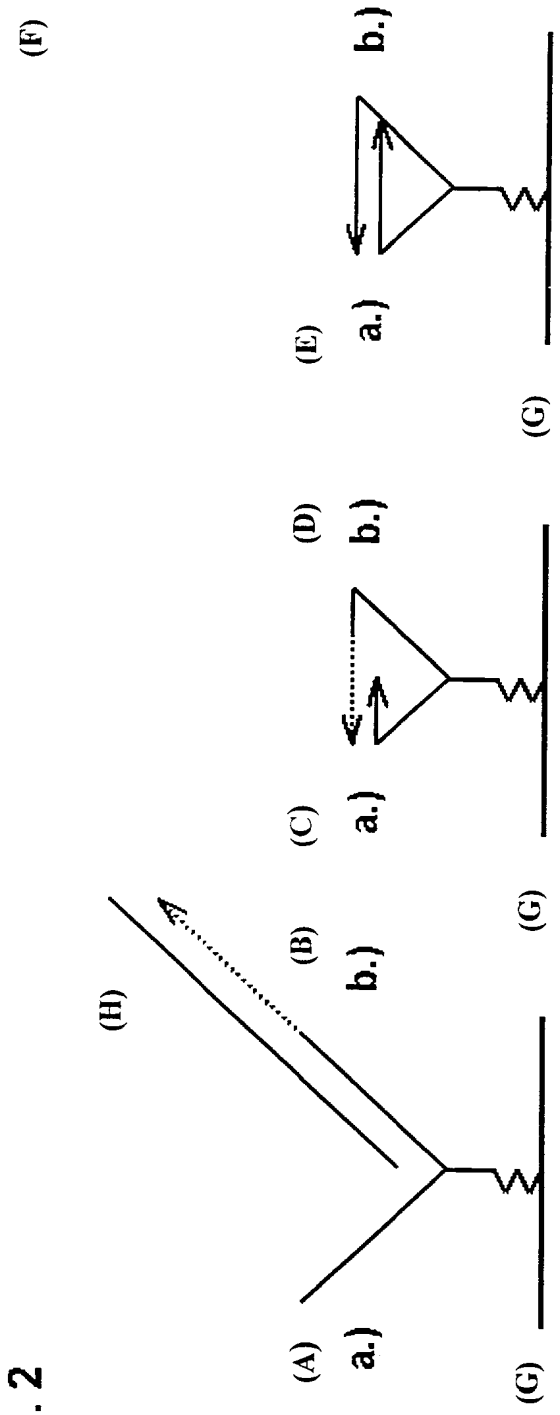


Fig. 2

IDT Exhibit 1020.6124

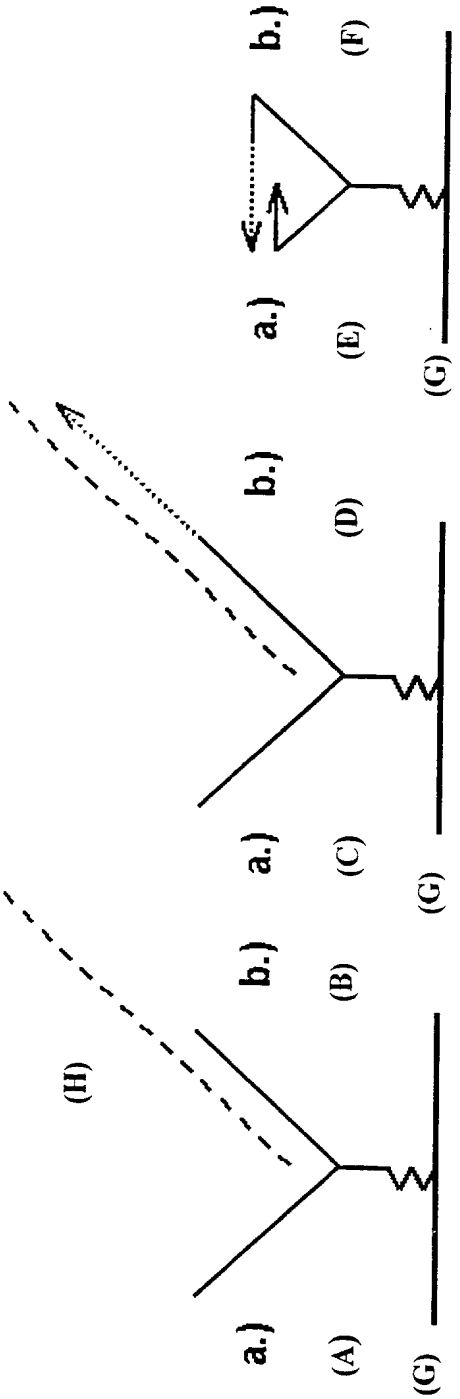
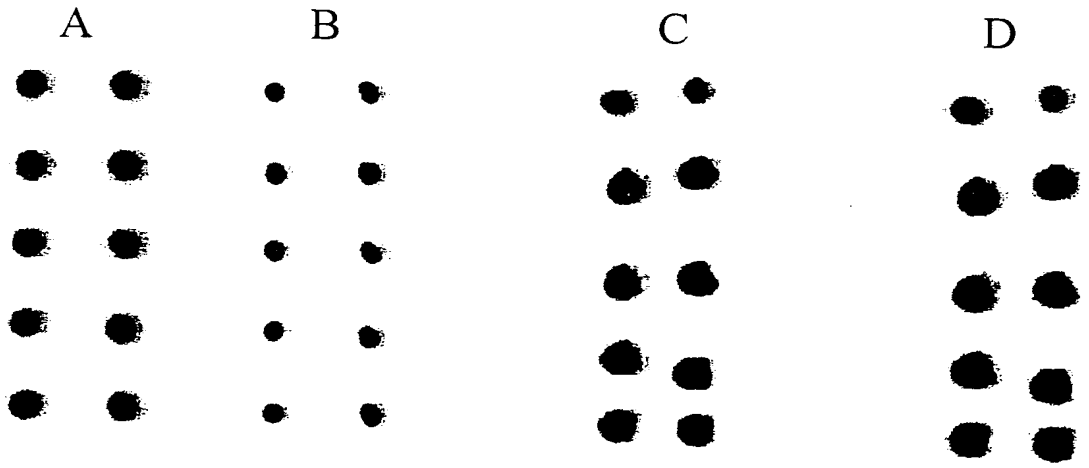


Fig. 3

IDT Exhibit 1020.6125

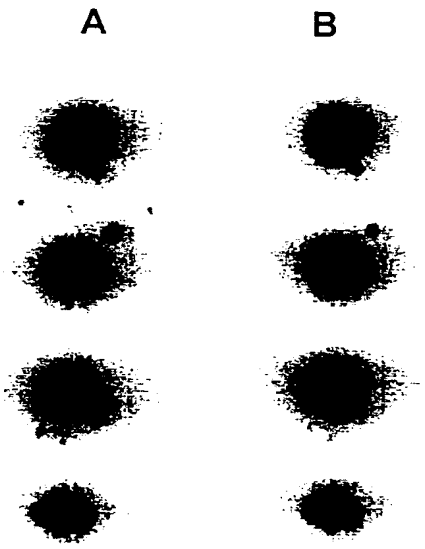
Fig. 4



	w/o wash cev (n=10)	wash cev (n=10)	bound %	background ratio
Aminosilane treated slides				
5'-NH_2 oligo	958753	166152	17.3%	--
Control oligo w/o modification	851570	27591	3.2%	1:5.4
Compound according to the invention, normal glass				
5'-aminosilane oligo	960840	874335	91.0%	-
Control oligo w/o modification	863096	5265	0.6%	1:152

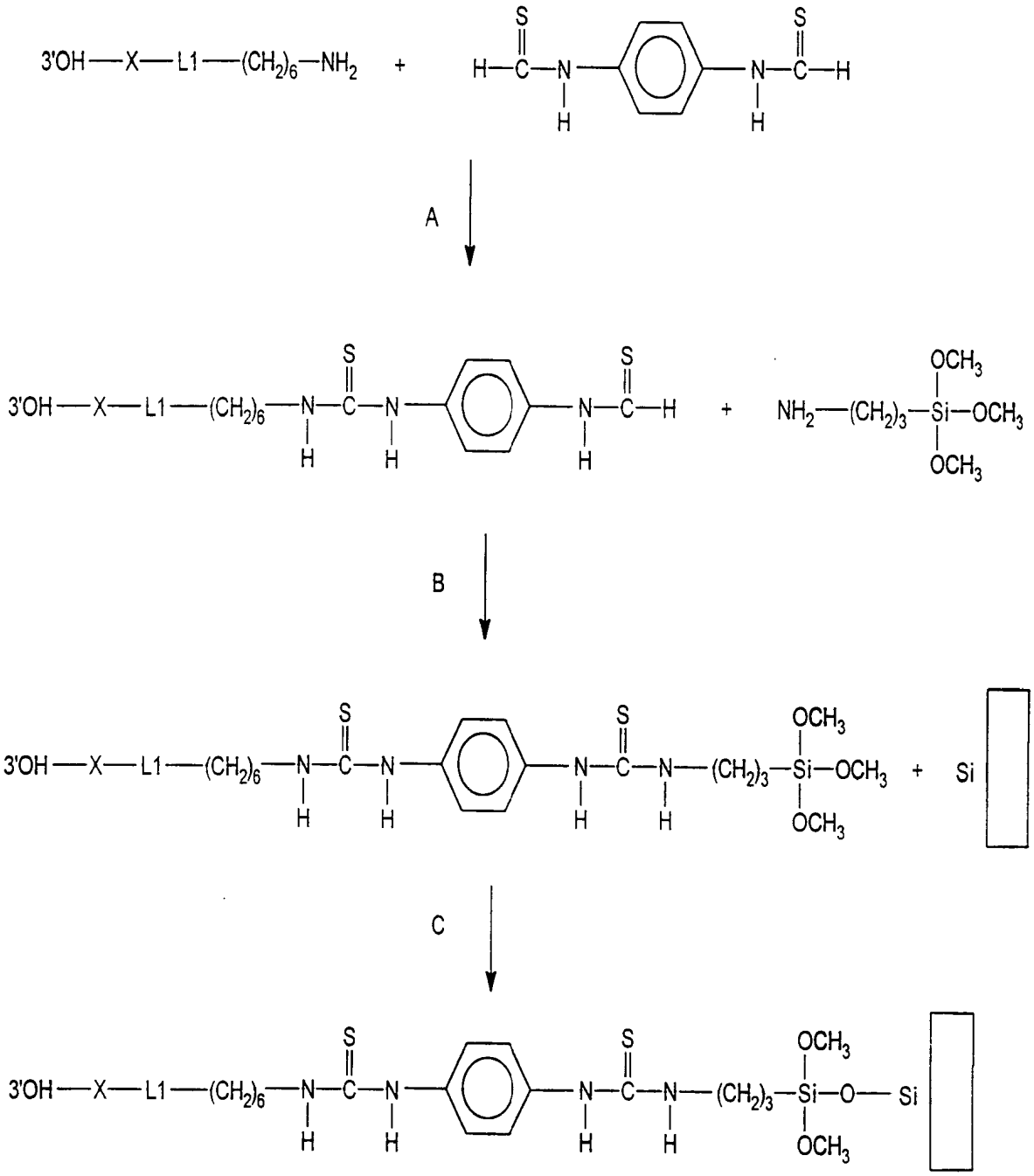
IDT Exhibit 1020.6126

Fig. 5



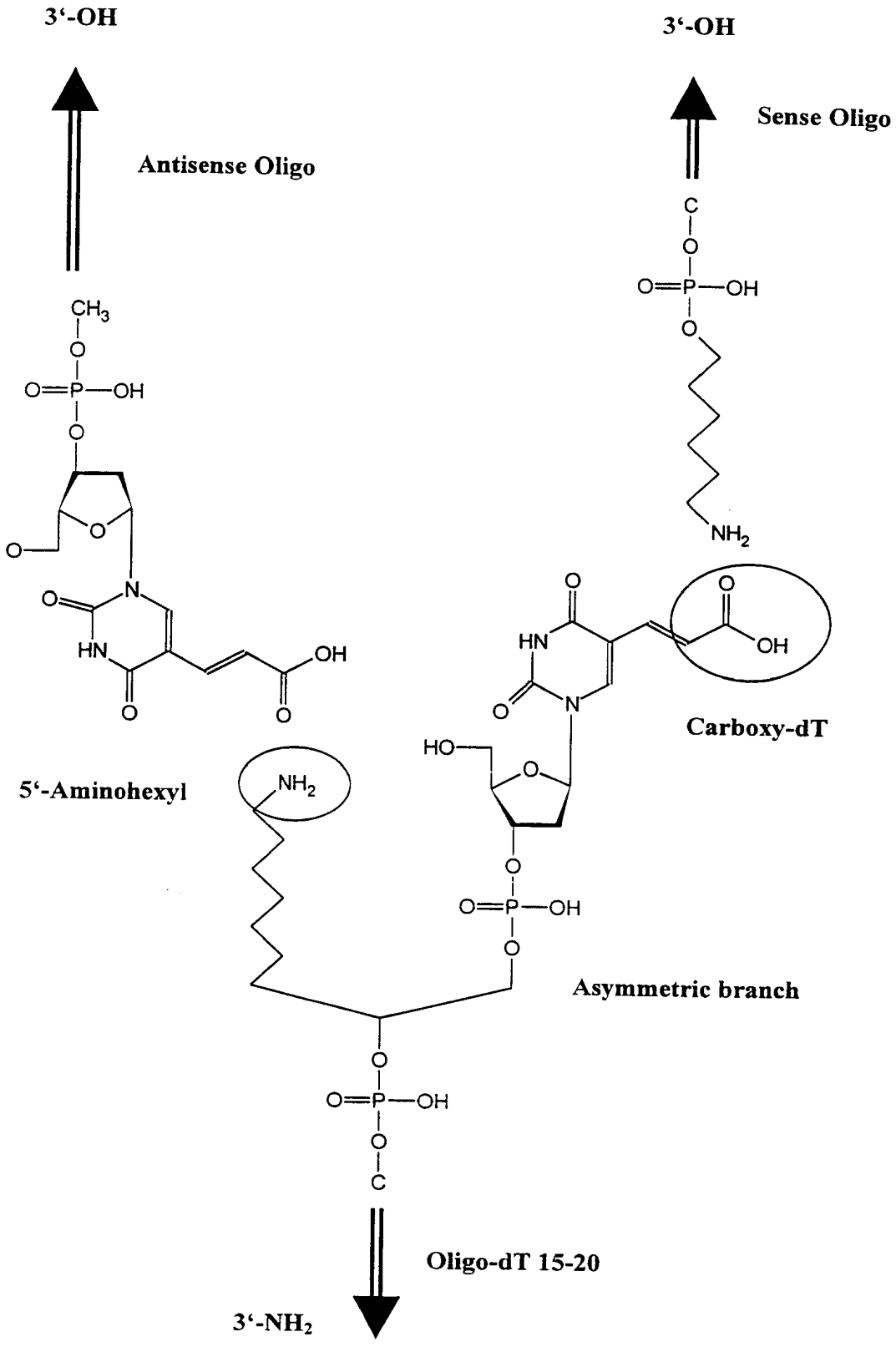
IDT Exhibit 1020.6127

Fig. 6



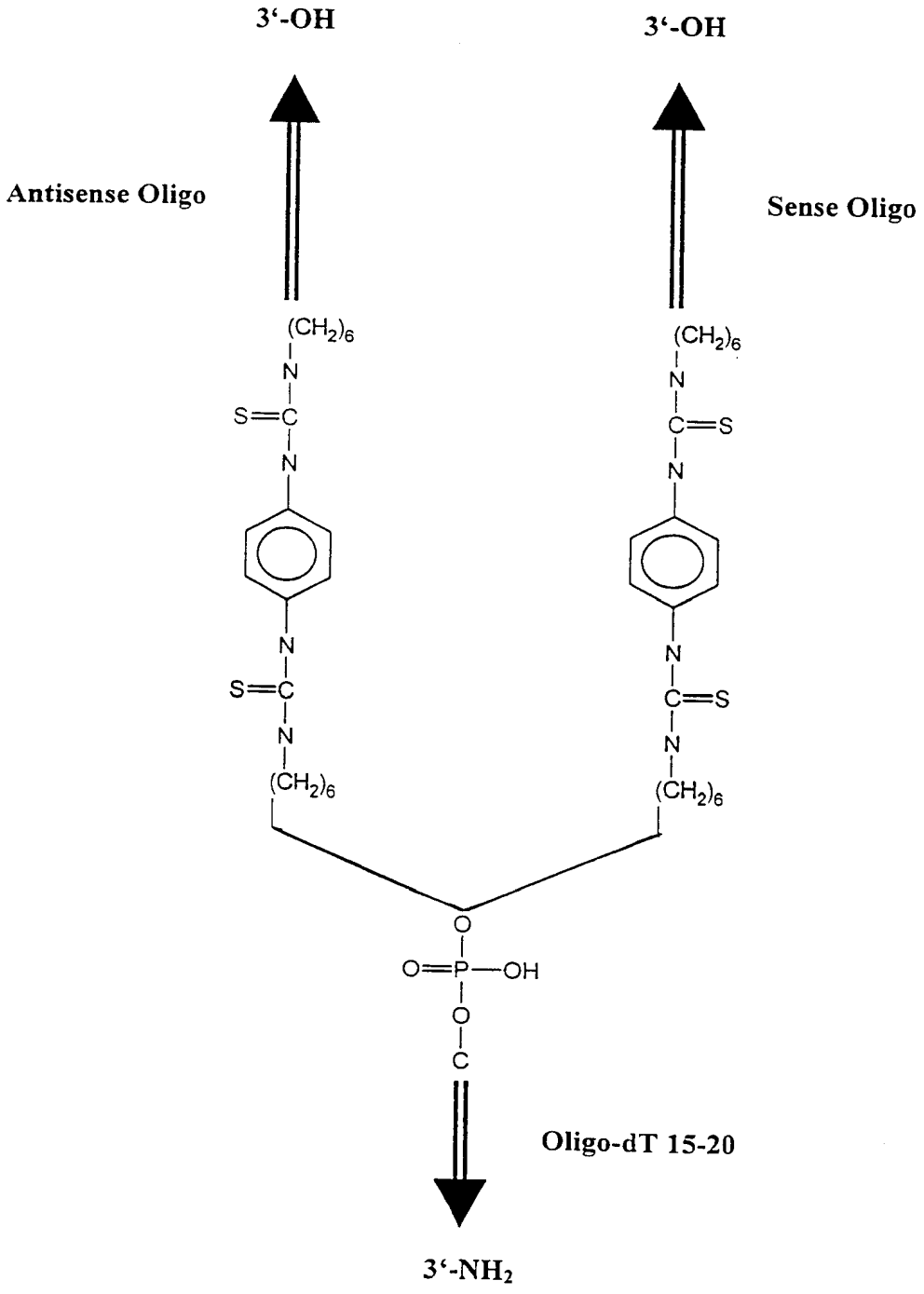
IDT Exhibit 1020.6128

Fig. 7



IDT Exhibit 1020.6129

Fig. 8



IDT Exhibit 1020.6130

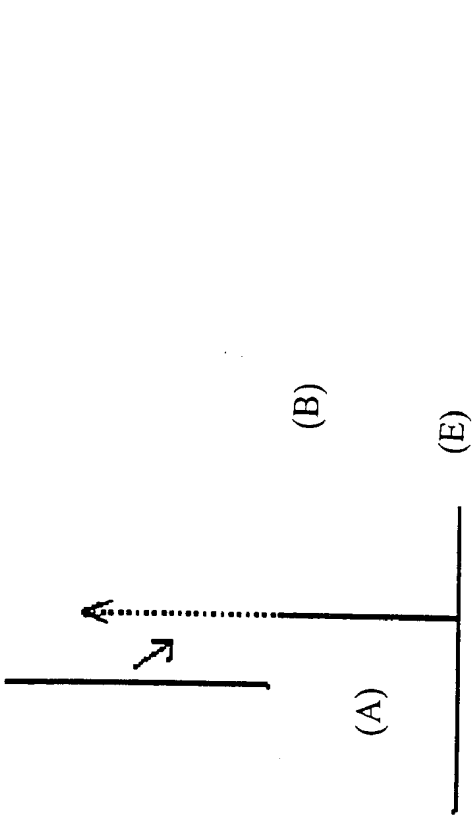
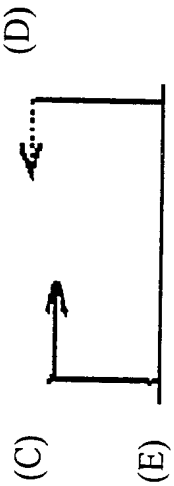


Fig. 9A



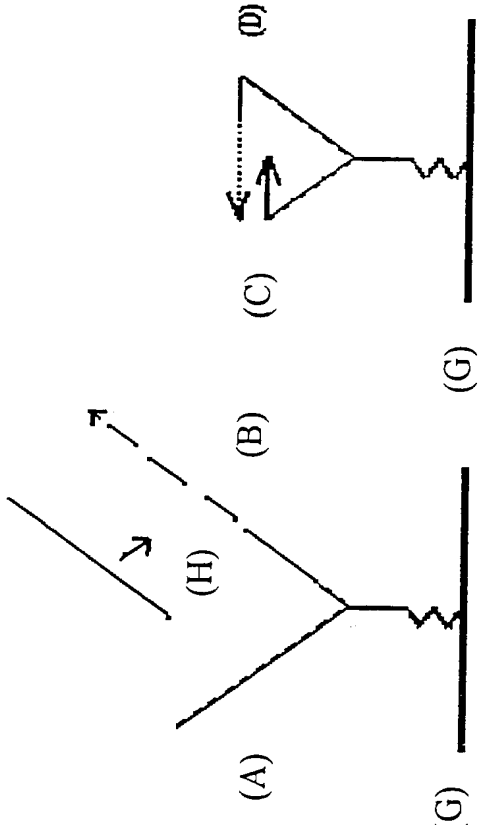


Fig. 9B

IDT Exhibit 1020.6131

INTERNATIONAL SEARCH REPORT

Intern. ıal Application No
PCT/EP 00/13098

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12Q1/68 C07H21/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C12Q C07H

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)
EPO-Internal, WPI Data, PAJ, MEDLINE, BIOSIS, EMBASE, CHEM ABS Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	FR 2 726 286 A (GENSET) 3 May 1996 (1996-05-03) the whole document ---	1-3,5-7, 10-18
X	GROTLI, MORTEN (1) ET AL: "Solid -phase synthesis of branched RNA and branched DNA/RNA chimeras." TETRAHEDRON, (1997) VOL. 53, NO. 33, PP. 11317-11346., XP002140600 the whole document --- -/--	1-3,5-7, 10-12, 18-20,23

[X] Further documents are listed in the continuation of box C. [X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier document but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
11 April 2001	20/04/2001

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL – 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Reuter, U

3

Form PCT/ISA/210 (second sheet) (July 1992)

IDT Exhibit 1020.6132

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
X	SPROAT ET AL.: "Novel solid-phase synthesis of branched oligoribonucleotides, including a substrate for the DNA debranching enzyme" JOURNAL OF THE CHEMICAL SOCIETY. PERKIN TRANSACTIONS I, 1994, pages 419-431, XP002140601 the whole document	1-3,5-7, 10-12, 18-20,23
A	BRAICH R S ET AL: "Regiospecific solid -phase synthesis of branched oligonucleotides. Effect of vicinal 2',5'- (or 2',3'-) and 3',5'-phosphodiester linkages on the formation of hairpin DNA." BIOCONJUGATE CHEMISTRY, (1997 MAY-JUN) 8 (3) 370-7., XP002140602 the whole document	1-23
A	WO 96 39413 A (HYBRIDON INC ;IYER RADHAKRISHNAN P (US); YU DONG (US); AGRAWAL SUD) 12 December 1996 (1996-12-12) the whole document	1-23
A	US 5 641 658 A (ADAMS CHRISTOPHER P ET AL) 24 June 1997 (1997-06-24) the whole document	1-23
A	US 5 916 750 A (INAMDAR ANITA ET AL) 29 June 1999 (1999-06-29) the whole document	1-23
A	US 5 912 332 A (AGRAWAL SUDHIR ET AL) 15 June 1999 (1999-06-15) the whole document	1-23
A	MASKOS U ET AL: "OLIGNUCLEOTIDE HYBRIDISATIONS ON GLASS SUPPORTS: A NOVEL LINKER FOR OLIGONUCLEOTIDE SYNTHESIS AND HYBRIDISATION PROPERTIES OF OLIGNUCLEOTIDES SYNTHESISED IN SITU" NUCLEIC ACIDS RESEARCH,GB,OXFORD UNIVERSITY PRESS, SURREY, vol. 20, no. 7, 1 January 1992 (1992-01-01), pages 1679-1684, XP000651031 ISSN: 0305-1048 the whole document	1,8,9
A	US 5 624 802 A (URDEA MICHAEL S ET AL) 29 April 1997 (1997-04-29) the whole document	1-23

3

IDT Exhibit 1020.6133

INTERNATIONAL SEARCH REPORT

Information on patent family members

Intern. ial Application No
PCT/EP 00/13098

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
FR 2726286 A	03-05-1996	AT 169683 T	15-08-1998
		DE 69504094 D	17-09-1998
		DE 69504094 T	11-03-1999
		DK 787205 T	17-05-1999
		EP 0787205 A	06-08-1997
		ES 2122692 T	16-12-1998
		WO 9613609 A	09-05-1996
		JP 10507923 T	04-08-1998
WO 9639413 A	12-12-1996	US 5750674 A	12-05-1998
		AT 194990 T	15-08-2000
		AU 5869696 A	11-12-1996
		AU 5871196 A	24-12-1996
		AU 5958196 A	24-12-1996
		CA 2222169 A	28-11-1996
		DE 69609511 D	31-08-2000
		DE 69609511 T	29-03-2001
		DK 836612 T	30-10-2000
		EP 0836612 A	22-04-1998
		US 6117993 A	12-09-2000
		WO 9637504 A	28-11-1996
		WO 9639414 A	12-12-1996
		US 5614622 A	25-03-1997
		US 6140482 A	31-10-2000
		US 5955599 A	21-09-1999
		US 5962674 A	05-10-1999
		ES 2151173 T	16-12-2000
US 5641658 A	24-06-1997	CA 2196604 A	15-02-1996
		EP 0784701 A	23-07-1997
		JP 10505492 T	02-06-1998
		WO 9604404 A	15-02-1996
		US 6090592 A	18-07-2000
		US 6060288 A	09-05-2000
US 5916750 A	29-06-1999	AU 5912098 A	03-08-1998
		EP 1019534 A	19-07-2000
		JP 2000511562 T	05-09-2000
		WO 9830719 A	16-07-1998
US 5912332 A	15-06-1999	AU 3672797 A	20-02-1998
		WO 9804571 A	05-02-1998
US 5624802 A	29-04-1997	US 5124246 A	23-06-1992
		US 5594118 A	14-01-1997
		US 5359100 A	25-10-1994
		US 5614362 A	25-03-1997
		AT 138691 T	15-06-1996
		AU 641393 B	23-09-1993
		AU 5563690 A	29-11-1990
		CA 2014855 A	18-10-1990
		DE 69027211 D	04-07-1996
		EP 0469066 A	05-02-1992
		ES 2087153 T	16-07-1996
		KR 156007 B	15-10-1998
		US 5656731 A	12-08-1997
		WO 9013667 A	15-11-1990
		US 5571670 A	05-11-1996

IDT Exhibit 1020.6134

INTERNATIONAL SEARCH REPORT

Information on patent family members

Intern ıal Application No
PCT/EP 00/13098

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 5624802 A		AT 133713 T	15-02-1996
		DE 3854956 D	14-03-1996
		DK 294589 A	15-08-1989
		EP 0317077 A	14-05-1989
		ES 2083363 T	16-04-1996
		GR 3019046 T	31-05-1996
		IE 75211 B	27-08-1997
		JP 2749277 B	13-05-1998
		JP 7303498 A	21-11-1995
		JP 2109999 A	23-04-1990
		JP 2565552 B	18-12-1996
		WO 8903891 A	05-05-1989

IDT Exhibit 1020.6135

Espacenet

Bibliographic data: WO0123613 (A1) — 2001-04-05

HUMAN RNASE H AND OLIGONUCLEOTIDE COMPOSITIONS THEREOF

Inventor(s): CROOKE STANLEY T [US]; LIMA WALTER F [US]; WU HONGJIANG [US]; MANOHARAN MUTHIAH [US] ± (CROOKE, STANLEY, T, ; LIMA, WALTER, F, ; WU, HONGJIANG, ; MANOHARAN, MUTHIAH)

Applicant(s): ISIS PHARMACEUTICALS INC [US]; CROOKE STANLEY T [US]; LIMA WALTER F [US]; WU HONGJIANG [US]; MANOHARAN MUTHIAH [US] ± (ISIS PHARMACEUTICALS, INC, ; CROOKE, STANLEY, T, ; LIMA, WALTER, F, ; WU, HONGJIANG, ; MANOHARAN, MUTHIAH)

Classification: - international: ***C07H21/00; C12N15/113; C12N9/22;*** (IPC1-7): C07H21/00; C07H21/02; C07H21/04; C12P19/34; C12Q1/68

- cooperative: **C07H21/00**; **C12N15/113**; **C12N9/22**; C07K2319/00; C12N2310/31; C12N2310/32; C12N2310/321; C12N2310/322; C12N2310/3341; C12N2310/346 more

Application number: WO2000US26729 20000929 Global Dossier

Priority number(s): US19990409926 19990930

Also published as: AT312202 (T) AU7621600 (A) DE60024660 (T2) DE60024660 (T4) EP1222309 (A1) more

Abstract of WO0123613 (A1)

The present invention provides oligonucleotides that can serve as substrates for human Type 2 RNase H. The present invention is also directed to methods of using these oligonucleotides in enhancing antisense oligonucleotide therapies.

IDT Exhibit 1020.6136

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
5 April 2001 (05.04.2001)

PCT

(10) International Publication Number
WO 01/23613 A1

(51) International Patent Classification[7]: C12Q 1/68, C12P 19/34, C07H 21/02, 21/04, 21/00

(21) International Application Number: PCT/US00/26729

(22) International Filing Date:
29 September 2000 (29.09.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
09/409,926 30 September 1999 (30.09.1999) US

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier application:
US 09/409,926 (CON)
Filed on 30 September 1999 (30.09.1999)

(71) Applicant *(for all designated States except US)*: **ISIS PHARMACEUTICALS, INC.** [US/US]; 2292 Faraday Avenue, Carlsbad, CA 92008 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **CROOKE, Stanley, T.** [US/US]; 3211 Piragua Street, Carlsbad, CA 92009 (US). **LIMA, Walter, F.** [US/US]; 2405 Bancroft, San Diego, CA 92104 (US). **WU, Hongjiang** [CN/US]; 7831 Calle Pinabete, Carlsbad, CA 92009 (US). **MANOHARAN, Muthiah** [US/US]; 7634 Reposado Drive, Carlsbad, CA 92009 (US).

(74) Agents: **CALDWELL, John, W.** et al.; Woodcock Washburn Kurtz Mackiewicz & Norris LLP, 46th floor, One Liberty Place, Philadelphia, PA 19103 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *With international search report.*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 01/23613 A1

(54) Title: HUMAN RNASE H AND OLIGONUCLEOTIDE COMPOSITIONS THEREOF

(57) Abstract: The present invention provides oligonucleotides that can serve as substrates for human Type 2 RNase H. The present invention is also directed to methods of using these oligonucleotides in enhancing antisense oligonucleotide therapies.

IDT Exhibit 1020.6137

HUMAN RNASE H AND OLIGONUCLEOTIDE COMPOSITIONS THEREOF

FIELD OF THE INVENTION

The present invention relates to a human Type 2 RNase H which has now been cloned, expressed and purified to electrophoretic homogeneity. The present invention further relates to oligonucleotide compositions that may serve as substrates for human RNase H1 or human Type 2 RNase H.

BACKGROUND OF THE INVENTION

RNase H hydrolyzes RNA in RNA-DNA hybrids. This enzyme was first identified in calf thymus but has subsequently been described in a variety of organisms (Stein, H. and Hausen, P., *Science*, 1969, 166, 393-395; Hausen, P. and Stein, H.,*Eur. J. Biochem.*, 1970, 14, 278-283). RNase H activity appears to be ubiquitous in eukaryotes and bacteria (Itaya, M. and Kondo K. *Nucleic Acids Res.*, 1991, 19, 4443-4449; Itaya et al., *Mol. Gen. Genet., 1991* 227, 438-445; Kanaya, S., and Itaya, M., *J. Biol. Chem.*, 1992, 267, 10184-10192; Busen, W., *J. Biol. Chem., 1980,* 255, 9434-9443; Rong, Y. W. and Carl, P. L., , 1990, *Biochemistry* 29, 383-389; Eder et al., *Biochimie*, 1993 75, 123-126). Although RNase Hs constitute a family of proteins of varying molecular weight, nucleolytic activity and

substrate requirements appear to be similar for the various isotypes. For example, all RNase Hs studied to date function as endonucleases, exhibiting limited sequence specificity and requiring divalent cations (*e.g.*, Mg^{2+}, Mn^{2+}) to produce cleavage products with 5' phosphate and 3' hydroxyl termini (Crouch, R. J., and Dirksen, M. L., *Nuclease*, Linn, S, M., & Roberts, R. J., Eds., Cold Spring Harbor Laboratory Press, Plainview, NY 1982, 211-241).

In addition to playing a natural role in DNA replication, RNase H has also been shown to be capable of cleaving the RNA component of certain oligonucleotide-RNA duplexes. While many mechanisms have been proposed for oligonucleotide mediated destabilization of target RNAs, the primary mechanism by which antisense oligonucleotides are believed to cause a reduction in target RNA levels is through this RNase H action. Monia et al., *J. Biol. Chem.*, 1993, 266:13, 14514-14522. *In vitro* assays have demonstrated that oligonucleotides that are not substrates for RNase H can inhibit protein translation (Blake et al., *Biochemistry*, 1985, 24, 6139-4145) and that oligonucleotides inhibit protein translation in rabbit reticulocyte extracts that exhibit low RNase H activity. However, more efficient inhibition was found in systems that supported RNase H activity (Walder, R.Y. and Walder, *J.A.*, *Proc. Nat'l Acad. Sci. USA*, 1988, 85, 5011-5015; Gagnor et al., *Nucleic Acid Res.*, 1987, 15, 10419-10436; Cazenave et al., *Nucleic Acid Res.*, 1989, 17, 4255-4273; and Dash et al., *Proc. Nat'l Acad.* Sci. USA, 1987, 84, 7896-7900.

Oligonucleotides commonly described as "antisense oligonucleotides" comprise nucleotide sequences sufficient in identity and number to effect specific hybridization with a particular nucleic acid. This nucleic acid or the protein(s) it encodes is generally referred to as the "target." Oligonucleotides are generally designed to bind either directly to mRNA transcribed from, or to a selected DNA

portion of, a preselected gene target, thereby modulating the amount of protein translated from the mRNA or the amount of mRNA transcribed from the gene, respectively. Antisense oligonucleotides may be used as research tools, diagnostic aids, and therapeutic agents.

"Targeting" an oligonucleotide to the associated nucleic acid, in the context of this invention, also refers to a multistep process which usually begins with the identification of the nucleic acid sequence whose function is to be modulated. This may be, for example, a cellular gene (or mRNA transcribed from the gene) whose expression is associated with a particular disorder or disease state, or a foreign nucleic acid from an infectious agent. The targeting process also includes determination of a site or sites within this gene for the oligonucleotide interaction to occur such that the desired effect, either detection or modulation of expression of the protein, will result.

RNase H1 from E.coli is the best-characterized member of the RNase H family. The 3-dimensional structure of E.coli RNase HI has been determined by x-ray crystallography, and the key amino acids involved in binding and catalysis have been identified by site-directed mutagenesis (Nakamura et al., *Proc. Natl. Acad. Sci. USA*, 1991, 88, 11535-11539; Katayanagi et al., *Nature*, 1990, 347, 306-309; Yang et al., Science, 1990, 249, 1398-1405; Kanaya et al., *J. Biol. Chem.*, 1991, 266, 11621-11627). The enzyme has two distinct structural domains. The major domain consists of four α helices and one large β sheet composed of three antiparallel β strands. The Mg^{2+} binding site is located on the β sheet and consists of three amino acids, Asp-10, Glu-48, and Gly-11 (Katayanagi et al., *Proteins: Struct., Funct., Genet.*, 1993, 17, 337-346). This structural motif of the Mg^{2+} binding site surrounded by β strands is similar to that in DNase I (Suck, D., and Oefner, C., *Nature,* 1986, 321, 620-625). The minor domain is believed to constitute the predominant binding region of the enzyme and

is composed of an α helix terminating with a loop. The loop region is composed of a cluster of positively charged amino acids that are believed to bind electrostatistically to the minor groove of the DNA/RNA heteroduplex substrate. Although the conformation of the RNA/DNA substrate can vary, from A-form to B-form depending on the sequence composition, in general RNA/DNA heteroduplexes adopt an A-like geometry (Pardi et al., *Biochemistry*, 1981, 20, 3986-3996; Hall, K. B., and Mclaughlin, L. W., *Biochemistry*, 1991, 30, 10606-10613; Lane et al., *Eur. J. Biochem.*, 1993, 215, 297-306). The entire binding interaction appears to comprise a single helical turn of the substrate duplex. Recently the binding characteristics, substrate requirements, cleavage products and effects of various chemical modifications of the substrates on the kinetic characteristics of E.coli RNase HI have been studied in more detail (Crooke, S.T. et al., *Biochem. J.*, 1995, 312, 599-608; Lima, W.F. and Crooke, S.T., *Biochemistry*, 1997, 36, 390-398; Lima, W.F. et al., *J. Biol. Chem.*, 1997, 272, 18191-18199; Tidd, D.M. and Worenius, H.M., *Br. J. Cancer*, 1989, 60, 343; Tidd, D.M. et al., *Anti-Cancer Drug Des.*, 1988, 3, 117.

In addition to RNase HI, a second E.coli RNase H, RNase HII has been cloned and characterized (Itaya, M., *Proc. Natl. Acad. Sci. USA*, 1990, 87, 8587-8591). It is comprised of 213 amino acids while RNase HI is 155 amino acids long. E. coli RNase HIM displays only 17% homology with E.coli RNase HI. An RNase H cloned from S. typhimurium differed from E.coli RNase HI in only 11 positions and was 155 amino acids in length (Itaya, M. and Kondo K., *Nucleic Acids Res.*, 1991, 19, 4443-4449; Itaya et al., *Mol. Gen. Genet.*, 1991, 227, 438-445). An enzyme cloned from S. cerevisae was 30% homologous to E.coli RNase HI (Itaya, M. and Kondo K., *Nucleic Acids Res.*, 1991, 19, 4443-4449; Itaya et al., *Mol. Gen. Genet.*, 1991, 227, 438-445). Thus, to date, no enzyme cloned from a

species other than E. coli has displayed substantial homology to E.coli RNase H II.

Proteins that display RNase H activity have also been cloned and purified from a number of viruses, other bacteria and yeast (Wintersberger, *U. Pharmac. Ther.*, 1990, 48, 259-280). In many cases, proteins with RNase H activity appear to be fusion proteins in which RNase H is fused to the amino or carboxy end of another enzyme, often a DNA or RNA polymerase. The RNase H domain has been consistently found to be highly homologous to E.coli RNase HI, but because the other domains vary substantially, the molecular weights and other characteristics of the fusion proteins vary widely.

In higher eukaryotes two classes of RNase H have been defined based on differences in molecular weight, effects of divalent cations, sensitivity to sulfhydryl agents and immunological cross-reactivity (Busen et al., *Eur. J. Biochem.*, 1977, 74, 203-208). RNase H Type 1 enzymes are reported to have molecular weights in the 68-90 kDa range, be activated by either Mn^{2+} or Mg^{2+} and be insensitive to sulfhydryl agents. In contrast, RNase H Type 2 enzymes have been reported to have molecular weights ranging from 31-45 kDa, to require Mg^{2+}, to be highly sensitive to sulfhydryl agents and to be inhibited by Mn^{2+} (Busen, W., and Hausen, P., *Eur. J. Biochem.*, 1975, 52, 179-190; Kane, C. M., *Biochemistry*, 1988, 27, 3187-3196; Busen, W., *J. Biol. Chem.*, 1982, 257, 7106-7108.).

An enzyme with Type 2 RNase H characteristics has been purified to near homogeneity from human placenta (Frank et al., *Nucleic Acids Res.*, 1994, 22, 5247-5254). This protein has a molecular weight of approximately 33 kDa and is active in a pH range of 6.5-10, with a pH optimum of 8.5-9. The enzyme requires Mg^{2+} and is inhibited by Mn^{2+} and n-ethyl maleimide. The products of cleavage reactions have 3' hydroxyl and 5' phosphate termini.

Despite the substantial information about members of the RNase family and the cloning of a number of viral, prokaryotic and yeast genes with RNase H activity, until now, no mammalian RNase H had been cloned. This has hampered efforts to understand the structure of the enzyme(s), their distribution and the functions they may serve.

In the present invention, a cDNA of human RNase H with Type 2 characteristics and the protein expressed thereby are provided.

SUMMARY OF THE INVENTION

The present invention provides oligonucleotides that can serve as substrates for human RNase H1. These oligonucleotides are mixed sequence oligonucleotides comprising at least two portions wherein a first portion is capable of supporting human RNase H1 cleavage of a complementary target RNA and a further portion which is not capable of supporting such human RNase H1 cleavage.

The present invention provides a mixed sequence oligonucleotide comprising at least 12 nucleotides and having a 3' end and a 5' end, said oligonucleotide being divided into a first portion and a further portion,

said first portion being capable of supporting cleavage of a complementary target RNA by human RNase H1 polypeptide,

said further portion being incapable of supporting said RNase H cleavage;

wherein said first portion comprises at least 6 nucleotides and is positioned in said oligonucleotide such that at least one of said 6 nucleotides is 8 to 12 nucleotides from the 3' end of said oligonucleotide.

In a preferred embodiment the oligonucleotide comprises at least one CA nucleotide sequence. In another embodiment the first portion of the mixed sequence oligonucleotide of the present invention comprises nucleotides having a B-form conformational geometry. In a further embodiment each of the

nucleotides of the first portion of the oligonucleotide are 2'-deoxyribonucleotides. In a still further embodiment each of the nucleotides of the first portion of the oligonucleotide is a 2'-F arabinonucleotide or a 2'-OH arabinonucleotide. In yet another embodiment the nucleotides of the first portion are joined together in a continuous sequence by phosphate, phosphorothioate, phosphorodithioate or boranophosphate linkages. In yet a further embodiment all of the nucleotides of the further portion of the oligonucleotide are joined together in a continuous sequence by 3'-5' phosphodiester, 2'-5' phosphodiester, phosphorothioate, Sp phosphorothioate, Rp phosphorothioate, phosphorodithioate, 3'-deoxy-3'-amino phosphoroamidate, 3'-methylenephosphonate, methylene(methylimino), dimethylhydrazino, amide 3, amide 4 or boranophosphate linkages.

Yet another object of the present invention is to provide methods for identifying agents which modulate activity and/or levels of human RNase H1. In accordance with this aspect, the polynucleotides and polypeptides of the present invention are useful for research, biological and clinical purposes. For example, the polynucleotides and polypeptides are useful in defining the interaction of human RNase H1 and antisense oligonucleotides and identifying means for enhancing this interaction so that antisense oligonucleotides are more effective at inhibiting their target mRNA.

Yet another object of the present invention is to provide a method of prognosticating efficacy of antisense therapy of a selected disease which comprises measuring the level or activity of human RNase H in a target cell of the antisense therapy. Similarly, oligonucleotides can be screened to identify those oligonucleotides which are effective antisense agents by measuring binding of the oligonucleotide to the human RNase H1.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows the effects of conditions on human RNase H1 activity.

Figure 2 shows a denaturing polyacrylamide gel analysis of human RNase H1 cleavage of a 17-mer RNA-DNA gapmer duplex.

Figure 3 shows analysis of human Rnase H1 cleavage of a 25-mer Ras RNA hybridized with phosphodiester oligodeoxynucleotides of different lengths.

Figure 4 shows analysis of human RNase H1 cleavage of RNA-DNA duplexes with different sequences, length and 3' or 5' overhangs. The sequence represented by G corresponds to SEQ ID NO: 33.

Figure 5 shows product and processivity analysis of human RNase H1 cleavage on 17-mer Ras RNA-DNA duplexes.

Figure 6 provides the human Type 2 RNase H primary sequence (286 amino acids; SEQ ID NO: 1) and sequence comparisons with chicken (293 amino acids; SEQ ID NO: 2), yeast (348 amino acids; SEQ ID NO: 3) and E. coli RNase H1 (155 amino acids; SEQ ID NO: 4) as well as an EST deduced mouse RNase H homolog (GenBank accession no. AA389926 and AA518920; SEQ ID NO: 5). Boldface type indicates amino acid residues identical to human. "@" indicates the conserved amino acid residues implicated in E. coli RNase H1 Mg^{2+} binding site and catalytic center (Asp-10, Gly-11, Glu-48 and Asp-70). "*" indicates the conserved residues implicated in E. coli RNases H1 for substrate binding.

DETAILED DESCRIPTION OF PREFERRED EMBODIMENTS

A Type 2 human RNase H has now been cloned and expressed. The enzyme encoded by this cDNA is inactive against single-stranded RNA, single-stranded DNA and double-stranded DNA. However, this enzyme cleaves the RNA in an RNA/DNA duplex and cleaves the RNA in a duplex comprised of RNA and a chimeric oligonucleotide with 2' methoxy flanks and a 5-deoxynucleotide center gap. The rate of cleavage of the

RNA duplexed with this so-called "deoxy gapmer" was significantly slower than observed with the full RNA/DNA duplex. These properties are consistent with those reported for E.coli RNase H1 (Crooke et al., *Biochem. J.*, 1995, 312, 599-608; Lima, W. F. and Crooke, S. T., *Biochemistry*, 1997, 36, 390-398). They are also consistent with the properties of a human Type 2 RNase H protein purified from placenta, as the molecular weight (32 kDa) is similar to that reported by Frank et al., *Nucleic Acids Res.*, 1994, 22, 5247-5254) and the enzyme is inhibited by Mn^{2+}.

Thus, in accordance with one aspect of the present invention, there are provided isolated polynucleotides which encode human Type 2 RNase H polypeptides having the deduced amino acid sequence of Figure 6. By "polynucleotides" it is meant to include any form of RNA or DNA such as mRNA or cDNA or genomic DNA, respectively, obtained by cloning or produced synthetically by well known chemical techniques. DNA may be double- or single-stranded. Single-stranded DNA may comprise the coding or sense strand or the non-coding or antisense strand.

Methods of isolating a polynucleotide of the present invention via cloning techniques are well known. For example, to obtain the cDNA contained in ATCC Deposit No. 98536, primers based on a search of the XREF database were used. An approximately 1 Kb cDNA corresponding to the carboxy terminal portion of the protein was cloned by 3' RACE. Seven positive clones were isolated by screening a liver cDNA library with this 1 Kb cDNA. The two longest clones were 1698 and 1168 base pairs. They share the same 5' untranslated region and protein coding sequence but differ in the length of the 3' UTR. A single reading frame encoding a 286 amino acid protein (calculated mass: 32029.04 Da) was identified (Figure 6). The proposed initiation codon is in agreement with the mammalian translation initiation consensus sequence described by Kozak, M., *J. Cell Biol.*, 1989, 108, 229-241, and is preceded by an

in-frame stop codon. Efforts to clone cDNA's with longer 5' UTR's from both human liver and lymphocyte cDNA's by 5' RACE failed, indicating that the 1698-base-pair clone was full length.

In a preferred embodiment, the polynucleotide of the present invention comprises the nucleic acid sequence of the cDNA contained within ATCC Deposit No. 98536. The deposit of E. coli DH5α containing a BLUESCRIPT® plasmid containing a human Type 2 RNase H cDNA was made with the American Type Culture Collection, 12301 Park Lawn Drive, Rockville, Maryland 20852, USA, on September 4, 1997 and assigned ATCC Deposit No. 98536. The deposited material is a culture of E. coli DH5α containing a BLUESCRIPT® plasmid (Stratagene, La Jolla CA) that contains the full-length human Type 2 RNase H cDNA. The deposit has been made under the terms of the Budapest Treaty on the international recognition of the deposit of micro-organisms for the purposes of patent procedure. The culture will be released to the public, irrevocably and without restriction to the public upon issuance of this patent. The sequence of the polynucleotide contained in the deposited material and the amino acid sequence of the polypeptide encoded thereby are controlling in the event of any conflict with the sequences provided herein. However, as will be obvious to those of skill in the art upon this disclosure, due to the degeneracy of the genetic code, polynucleotides of the present invention may comprise other nucleic acid sequences encoding the polypeptide of Figure 6 and derivatives, variants or active fragments thereof.

Another aspect of the present invention relates to the polypeptides encoded by the polynucleotides of the present invention. In a preferred embodiment, a polypeptide of the present invention comprises the deduced amino acid sequence of human Type 2 RNase H provided in Figure 6 as SEQ ID NO: 1. However, by "polypeptide" it is also meant to include fragments, derivatives and analogs of SEQ ID NO: 1 which

retain essentially the same biological activity and/or function as human Type 2 RNase H. Alternatively, polypeptides of the present invention may retain their ability to bind to an antisense-RNA duplex even though they do not function as active RNase H enzymes in other capacities. In another embodiment, polypeptides of the present invention may retain nuclease activity but without specificity for the RNA portion of an RNA/DNA duplex. Polypeptides of the present invention include recombinant polypeptides, isolated natural polypeptides and synthetic polypeptides, and fragments thereof which retain one or more of the activities described above.

In a preferred embodiment, the polypeptide is prepared recombinantly, most preferably from the culture of E. coli of ATCC Deposit No. 98536. Recombinant human RNase H fused to histidine codons (his-tag; in the present embodiment six histidine codons were used) expressed in E.coli can be conveniently purified to electrophoretic homogeneity by chromatography with Ni-NTA followed by C4 reverse phase HPLC. The purified recombinant polypeptide of SEQ ID NO: 1 is highly homologous to E.coli RNase H, displaying nearly 34% amino acid identity with E.coli RNase H1. Figure 6 compares the protein sequences deduced from human RNase H cDNA (SEQ ID NO: 1) with those of chicken (SEQ ID NO: 2), yeast (SEQ ID NO: 3) and E.coli RNase HI (Gene Bank accession no. 1786408; SEQ ID NO: 4), as well as an EST deduced mouse RNase H homolog (Gene Bank accession no. AA389926 and AA518920; SEQ ID NO: 5). The deduced amino acid sequence of human RNase H (SEQ ID NO: 1) displays strong homology with yeast (21.8% amino acid identity), chicken (59%), E.coli RNase HI (33.6%) and the mouse EST homolog (84.3%). They are all small proteins (<40 KDa) and their estimated pIs are all 8.7 and greater. Further, the amino acid residues in E.coli RNase HI thought to be involved in the Mg^{2+} binding site, catalytic center and substrate binding region are completely conserved in the cloned human RNase H sequence (Figure 6).

The human Type 2 RNase H of SEQ ID NO: 1 is expressed ubiquitously. Northern blot analysis demonstrated that the transcript was abundant in all tissues and cell lines except the MCR-5 line. Northern blot analysis of total RNA from human cell lines and Poly A containing RNA from human tissues using the 1.7 kb full length probe or a 332-nucleotide probe that contained the 5' UTR and coding region of human RNase H cDNA revealed two strongly positive bands with approximately 1.2 and 5.5 kb in length and two less intense bands approximately 1.7 and 4.0 kb in length in most cell lines and tissues. Analysis with the 332-nucleotide probe showed that the 5.5 kb band contained the 5' UTR and a portion of the coding region, which suggests that this band represents a pre-processed or partially processed transcript, or possibly an alternatively spliced transcript. Intermediate sized bands may represent processing intermediates. The 1.2 kb band represents the full length transcripts. The longer transcripts may be processing intermediates or alternatively spliced transcripts.

RNase H is expressed in most cell lines tested; only MRC5, a breast cancer cell line, displayed very low levels of RNase H. However, a variety of other malignant cell lines including those of bladder (T24), breast (T-47D, HS578T), lung (A549), prostate (LNCap, DU145), and myeloid lineage (HL-60), as well as normal endothelial cells (HUVEC), expressed RNase H. Further, all normal human tissues tested expressed RNase H. Again, larger transcripts were present as well as the 1.2 kb transcript that appears to be the mature mRNA for RNase H. Normalization based on G3PDH levels showed that expression was relatively consistent in all of the tissues tested.

The Southern blot analysis of EcoRI digested human and various mammalian vertebrate and yeast genomic DNAs probed with the 1.7 kb probe shows that four EcoRI digestion products of human genomic DNA (2.4, 4.6, 6.0, 8.0 Kb) hybridized with the 1.7 kb probe. The blot re-probed with a 430 nucleotide

probe corresponding to the C-terminal portion of the protein showed only one 4.6 kbp EcoRI digestion product hybridized. These data indicate that there is only one gene copy for RNase H and that the size of the gene is more than 10 kb. Both the full length and the shorter probe strongly hybridized to one EcoRI digestion product of yeast genomic DNA (about 5 kb in size), indicating a high degree of conservation. These probes also hybridized to the digestion product from monkey, but none of the other tested mammalian genomic DNAs including the mouse which is highly homologous to the human RNase H sequence.

A recombinant human RNase H (his-tag fusion protein) polypeptide of the present invention was expressed in E.coli and purified by Ni-NTA agarose beads followed by C4 reverse phase column chromatography. A 36 kDa protein copurified with activity measured after renaturation. The presence of the his-tag was confirmed by Western blot analyses with an anti-penta-histidine antibody (Qiagen, Germany).

Renatured recombinant human RNase H displayed RNase H activity. Incubation of 10 ng purified renatured RNase H with RNA/DNA substrate for 2 hours resulted in cleavage of 40% of the substrate. The enzyme also cleaved RNA in an oligonucleotide/RNA duplex in which the oligonucleotide was a gapmer with a 5-deoxynucleotide gap, but at a much slower rate than the full RNA/DNA substrate. This is consistent with observations with E.coli RNase HI (Lima, W. F. and Crooke, S. T., Biochemistry, 1997, 36, 390-398). It was inactive against single-stranded RNA or double-stranded RNA substrates and was inhibited by Mn^{2+}. The molecular weight (~36kDa) and inhibition by Mn^{2+} indicate that the cloned enzyme is highly homologous to E.coli RNase HI and has properties consistent with those assigned to Type 2 human RNase H.

The sites of cleavage in the RNA in the full RNA/DNA substrate and the gapmer/RNA duplexes (in which the oligonucleotide gapmer had a 5-deoxynucleotide gap) resulting from the recombinant enzyme were determined. In the full

RNA/DNA duplex, the principal site of cleavage was near the middle of the substrate, with evidence of less prominent cleavage sites 3' to the primary cleavage site. The primary cleavage site for the gapmer/RNA duplex was located across the 5 nucleotide adjacent to the junction of the 2' methoxy wing and oligodeoxy nucleotide gap nearest the 3' end of the RNA. Thus, the enzyme resulted in a major cleavage site in the center of the RNA/DNA substrate and less prominent cleavages to the 3' side of the major cleavage site. The shift of its 10 major cleavage site to the nucleotide in apposition to the DNA 2' methoxy junction of the 2' methoxy wing at the 5' end of the chimeric oligonucleotide is consistent with the observations for E.coli RNase HI (Crooke et al. (1995) Biochem. J. 312, 599-608; Lima, W. F. and Crooke, S. T. (1997) 15 Biochemistry 36, 390-398). The fact that the enzyme cleaves at a single site in a 5-deoxy gap duplex indicates that the enzyme has a catalytic region of similar dimensions to that of E.coli RNase HI.

Accordingly, expression of large quantities of a 20 purified human RNase H polypeptide of the present invention is useful in characterizing the activities of a mammalian form of this enzyme. In addition, the polynucleotides and polypeptides of the present invention provide a means for identifying agents which enhance the function of antisense 25 oligonucleotides in human cells and tissues.

For example, a host cell can be genetically engineered to incorporate polynucleotides and express polypeptides of the present invention. Polynucleotides can be introduced into a host cell using any number of well known techniques such as 30 infection, transduction, transfection or transformation. The polynucleotide can be introduced alone or in conjunction with a second polynucleotide encoding a selectable marker. In a preferred embodiment, the host comprises a mammalian cell. Such host cells can then be used not only for production of 35 human Type 2 RNase H, but also to identify agents which

increase or decrease levels of expression or activity of human Type 2 RNase H in the cell. In these assays, the host cell would be exposed to an agent suspected of altering levels of expression or activity of human Type 2 RNase in the cells. The level or activity of human Type 2 RNase in the cell would then be determined in the presence and absence of the agent. Assays to determine levels of protein in a cell are well known to those of skill in the art and include, but are not limited to, radioimmunoassays, competitive binding assays, Western blot analysis and enzyme linked immunosorbent assays (ELISAs). Methods of determining increase activity of the enzyme, and in particular increased cleavage of an antisense-mRNA duplex can be performed in accordance with the teachings of Example 5. Agents identified as inducers of the level or activity of this enzyme may be useful in enhancing the efficacy of antisense oligonucleotide therapies.

The present invention also relates to prognostic assays wherein levels of RNase in a cell type can be used in predicting the efficacy of antisense oligonucleotide therapy in specific target cells. High levels of RNase in a selected cell type are expected to correlate with higher efficacy as compared to lower amounts of RNase in a selected cell type which may result in poor cleavage of the mRNA upon binding with the antisense oligonucleotide. For example, the MRC5 breast cancer cell line displayed very low levels of RNase H as compared to other malignant cell types. Accordingly, in this cell type it may be desired to use antisense compounds which do not depend on RNase H activity for their efficacy.

Similarly, oligonucleotides can be screened to identify those which are effective antisense agents by contacting human Type 2 RNase H with an oligonucleotide and measuring binding of the oligonucleotide to the human Type 2 RNase H. Methods of determining binding of two molecules are well known in the art. For example, in one embodiment, the oligonucleotide can be radiolabeled and binding of the oligonucleotide to human

Type 2 RNase H can be determined by autoradiography. Alternatively, fusion proteins of human Type 2 RNase H with glutathione-S-transferase or small peptide tags can be prepared and immobilized to a solid phase such as beads. Labeled or unlabeled oligonucleotides to be screened for binding to this enzyme can then be incubated with the solid phase. Oligonucleotides which bind to the enzyme immobilized to the solid phase can then be identified either by detection of bound label or by eluting specifically the bound oligonucleotide from the solid phase. Another method involves screening of oligonucleotide libraries for binding partners. Recombinant tagged or labeled human Type 2 RNase H is used to select oligonucleotides from the library which interact with the enzyme. Sequencing of the oligonucleotides leads to identification of those oligonucleotides which will be more effective as antisense agents.

The oligonucleotides of the present invention are formed from a plurality of nucleotides that are joined together via internucleotide linkages. While joined together as a unit in the oligonucleotide, the individual nucleotides of oligonucleotides are of several types. Each of these types contribute unique properties to the oligonucleotide. A first type of nucleotides are joined together in a continuous sequence that forms a first portion of the oligonucleotide. The remaining nucleotides are of at least one further type and are located in one or more remaining portions or locations within the oligonucleotide. Thus, the oligonucleotides of the invention include a nucleotide portion that contributes one set of attributes and a further portion (or portions) that contributes another set of attributes.

One attribute that is desirable is eliciting RNase H activity. To elicit RNase H activity, a portion of the oligonucleotides of the invention is selected to have B-form like conformational geometry. The nucleotides for this B-form portion are selected to specifically include ribo-

pentofuranosyl and arabino-pentofuranosyl nucleotides. 2'-Deoxy-*erythro*-pentfuranosyl nucleotides also have B-form geometry and elicit RNase H activity. While not specifically excluded, if 2'-deoxy-*erythro*-pentfuranosyl nucleotides are included in the B-form portion of an oligonucleotide of the invention, such 2'-deoxy-*erythro*-pentfuranosyl nucleotides preferably does not constitute the totality of the nucleotides of that B-form portion of the oligonucleotide, but should be used in conjunction with ribonucleotides or an arabino nucleotides. As used herein, B-form geometry is inclusive of both C2'-endo and O4'-endo pucker, and the ribo and arabino nucleotides selected for inclusion in the oligonucleotide B-form portion are selected to be those nucleotides having C2'-endo conformation or those nucleotides having O4'-endo conformation. This is consistent with Berger, *et. al.*, *Nucleic Acids Research*, **1998**, *26*, 2473-2480, who pointed out that in considering the furanose conformations in which nucleosides and nucleotides reside, B-form consideration should also be given to a O4'-endo pucker contribution.

A-form nucleotides are nucleotides that exhibit C3'-endo pucker, also known as north, or northern, pucker. In addition to the B-form nucleotides noted above, the A-form nucleotides can be C3'-endo pucker nucleotides or can be nucleotides that are located at the 3' terminus, at the 5' terminus, or at both the 3' and the 5' terminus of the oligonucleotide. Alternatively, A-form nucleotides can exist both in a C3'-endo pucker and be located at the ends, or termini, of the oligonucleotide. In selecting nucleotides that have C3'-endo pucker or in selecting nucleotides to reside at the 3' or 5' ends of the oligonucleotide, consideration is given to binding affinity and nuclease resistance properties that such nucleotides need to impart to the resulting the oligonucleotide.

Nucleotides selected to reside at the 3' or 5' termini of oligonucleotides of the invention are selected to impart

nuclease resistance to the oligonucleotide. This nuclease resistance can also be achieved via several mechanisms, including modifications of the sugar portions of the nucleotide units of the oligonucleotides, modification of the internucleotide linkages or both modification of the sugar and the internucleotide linkage.

A particularly useful group of nucleotides for use in increasing nuclease resistance at the termini of oligonucleotides are those having 2'-O-alkylamino groups thereon. The amino groups of such nucleotides can be groups that are protonated at physiological pH. These include amines, monoalkyl substituted amines, dialkyl substituted amines and heterocyclic amines such as imidazole. Particularly useful are the lower alkyl amines including 2'-O-ethylamine and 2'-O-propylamine. Such O-alkylamines can also be included on the 3' position of the 3' terminus nucleotide. Thus the 3' terminus nucleotide could include both a 2' and a 3'-O-alkylamino substituent.

In selecting for nuclease resistance, it is important not to detract from binding affinity. Certain phosphorus based linkage have been shown to increase nuclease resistance. The above described phosphorothioate linkage increase nuclease resistance, however, it also causes loss of binding affinity. Thus, generally for use in this invention, if phosphorothioate internucleotide linkage are used, other modification will be made to nucleotide units that increase binding affinity to compensate for the decreased affinity contribute by the phosphorothioate linkages.

Other phosphorus based linkages having increase nuclease resistance that do not detract from binding affinity include 3'-methylene phosphonates and 3'-deoxy-3'-amino-phosphoroamidate linkages. A further class of linkages that contribute nuclease resistance but do not detract from binding affinity are non-phosphate in nature. Preferred among these are methylene(methylimino) linkages, dimethylhydraxino

linkages, and amine 3 and amide 4 linkages as described (Freier and Altmann, *Nucleic Acid Research*, **1997**, *25*, 4429-4443).

There are a number of potential items to consider when designing oligonucleotides having improved binding affinities. It appears that one effective approach to constructing modified oligonucleotides with very high RNA binding affinity is the combination of two or more different types of modifications, each of which contributes favorably to various factors that might be important for binding affinity.

Freier and Altmann, *Nucleic Acids Research,* (1997) 25:4429-4443, recently published a study on the influence of structural modifications of oligonucleotides on the stability of their duplexes with target RNA. In this study, the authors reviewed a series of oligonucleotides containing more than 200 different modifications that had been synthesized and assessed for their hybridization affinity and T_m. Sugar modifications studied included substitutions on the 2'-position of the sugar, 3'-substitution, replacement of the 4'-oxygen, the use of bicyclic sugars, and four member ring replacements. Several nucleobase modifications were also studied including substitutions at the 5, or 6 position of thymine, modifications of pyrimidine heterocycle and modifications of the purine heterocycle. Numerous backbone modifications were also investigated including backbones bearing phosphorus, backbones that did not bear a phosphorus atom, and backbones that were neutral.

Four general approaches might be used to improve hybridization of oligonucleotides to RNA targets. These include: preorganization of the sugars and phosphates of the oligodeoxynucleotide strand into conformations favorable for hybrid formation, improving stacking of nucleobases by the addition of polarizable groups to the heterocycle bases of the nucleotides of the oligonucleotide, increasing the number of H-bonds available for A-U pairing, and neutralization of

backbone charge to facilitate removing undesirable repulsive interactions. We have found that by utilizing the first of these, preorganization of the sugars and phosphates of the oligodeoxynucleotide strand into conformations favorable for hybrid formation, to be a preferred method to achieve improve binding affinity. It can further be used in combination with the other three approaches.

Sugars in DNA:RNA hybrid duplexes frequently adopt a C3' endo conformation. Thus modifications that shift the conformational equilibrium of the sugar moieties in the single strand toward this conformation should preorganize the antisense strand for binding to RNA. Of the several sugar modifications that have been reported and studied in the literature, the incorporation of electronegative substituents such as 2'-fluoro or 2'-alkoxy shift the sugar conformation towards the 3' *endo* (northern) pucker conformation. This preorganizes an oligonucleotide that incorporates such modifications to have an A-form conformational geometry. This A-form conformation results in increased binding affinity of the oligonucleotide to a target RNA strand.

As used herein, the terms "substituent" and "substituent group" refers to groups that are attached to nucleosides of the invention. Substituent groups are preferably attached to selected sugar moieties but can alternatively be attached to selected heterocyclic base moieties. Selected nucleosides may have substituent groups at both the heterocyclic base and the sugar moiety, however a single substituent group is preferred at a sugar 2', 3' or 5'-positions with the 2'-position being particularly preferred.

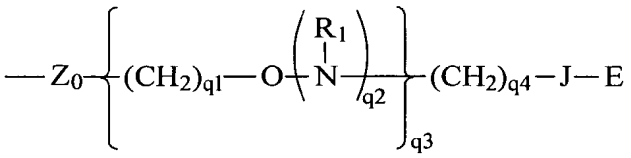
Substituent groups include fluoro, O-alkyl, O-alkylamino, O-alkylalkoxy, O-alkylaminoalkyl, O-alkyl imidazole, and polyethers of the formula $(O\text{-alkyl})_m$, where m is 1 to about 10. Preferred among these polyethers are linear and cyclic polyethylene glycols (PEGs), and (PEG)-containing groups, such as crown ethers and those which are disclosed by

Ouchi *et al.* (*Drug Design and Discovery* **1992**, *9*, 93), Ravasio *et al.* (*J. Org. Chem.* **1991**, *56*, 4329) and Delgardo *et. al.* (*Critical Reviews in Therapeutic Drug Carrier Systems* **1992**, *9*, 249), each of which is herein incorporated by reference in its entirety. Further sugar modifications are disclosed in Cook, P.D., *Anti-Cancer Drug Design*, **1991**, *6*, 585-607. Fluoro, O-alkyl, O-alkylamino, O-alkyl imidazole, O-alkylaminoalkyl, and alkyl amino substitution is described in United States Patent Application serial number 08/398,901, filed March 6, 1995, entitled Oligomeric Compounds having Pyrimidine Nucleotide(s) with 2' and 5' Substitutions, hereby incorporated by reference in its entirety.

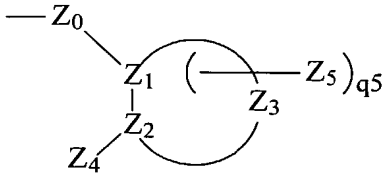
Additional substituent groups amenable to the present invention include -SR and $-NR_2$ groups, wherein each R is, independently, hydrogen, a protecting group or substituted or unsubstituted alkyl, alkenyl, or alkynyl. 2'-SR nucleosides are disclosed in United States Patent No. 5,670,633, issued September 23, 1997, hereby incorporated by reference in its entirety. The incorporation of 2'-SR monomer synthons are disclosed by Hamm *et al.*, *J. Org. Chem.*, **1997**, *62*, 3415-3420. 2'-NR_2 nucleosides are disclosed by Goettingen, M., *J. Org. Chem.*, **1996**, *61*, 6273-6281; and Polushin *et al.*, *Tetrahedron Lett.*, **1996**, *37*, 3227-3230.

Further representative substituent groups include hydrogen, hydroxyl, C_1-C_{20} alkyl, C_2-C_{20} alkenyl, C_2-C_{20} alkynyl, halogen, amino, thiol, keto, carboxyl, nitro, nitroso, nitrile, trifluoromethyl, trifluoromethoxy, O-alkyl, O-alkenyl, O-alkynyl, S-alkyl, S-alkenyl, S-alkynyl, NH-alkyl, NH-alkenyl, NH-alkynyl, N-dialkyl, O-aryl, S-aryl, NH-aryl, O-aralkyl, S-aralkyl, NH-aralkyl, N-phthalimido, imidazole, azido, hydrazino, hydroxylamino, isocyanato, sulfoxide, sulfone, sulfide, disulfide, silyl, aryl, heterocycle, carbocycle, intercalator, reporter molecule, conjugate, polyamine, polyamide, polyalkylene glycol, or polyether;

or each substituent group has one of formula I or II:



I



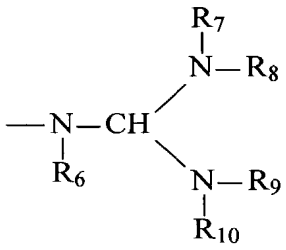
II

wherein:

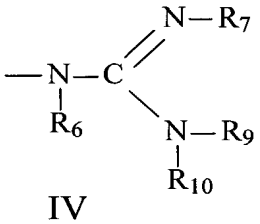
Z_0 is O, S or NH;

J is a single bond, O or C(=O);

E is C_1-C_{10} alkyl, $N(R_1)(R_2)$, $N(R_1)(R_5)$, $N=C(R_1)(R_2)$, $N=C(R_1)(R_5)$ or has one of formula III or IV;



III



IV

each R_6, R_7, R_8, R_9 and R_{10} is, independently, hydrogen, $C(O)R_{11}$, substituted or unsubstituted C_1-C_{10} alkyl, substituted or unsubstituted C_2-C_{10} alkenyl, substituted or unsubstituted C_2-C_{10} alkynyl, alkylsulfonyl, arylsulfonyl, a chemical functional group or a conjugate group, wherein the substituent groups are selected from hydroxyl, amino, alkoxy, carboxy, benzyl, phenyl, nitro, thiol, thioalkoxy, halogen, alkyl, aryl, alkenyl and alkynyl;

or optionally, R_7 and R_8, together form a phthalimido moiety with the nitrogen atom to which they are attached;

or optionally, R_9 and R_{10}, together form a phthalimido moiety with the nitrogen atom to which they are attached;

each R_{11} is, independently, substituted or unsubstituted C_1-C_{10} alkyl, trifluoromethyl, cyanoethyloxy, methoxy, ethoxy, t-butoxy, allyloxy, 9-fluorenylmethoxy, 2-(trimethylsilyl)-

ethoxy, 2,2,2-trichloroethoxy, benzyloxy, butyryl, iso-butyryl, phenyl or aryl;

R_5 is T-L,

T is a bond or a linking moiety;

L is a chemical functional group, a conjugate group or a solid support material;

each R_1 and R_2 is, independently, H, a nitrogen protecting group, substituted or unsubstituted C_1-C_{10} alkyl, substituted or unsubstituted C_2-C_{10} alkenyl, substituted or unsubstituted C_2-C_{10} alkynyl, wherein said substitution is OR_3, SR_3, NH_3^+, $N(R_3)(R_4)$, guanidino or acyl where said acyl is an acid amide or an ester;

or R_1 and R_2, together, are a nitrogen protecting group or are joined in a ring structure that optionally includes an additional heteroatom selected from N and O;

or R_1, T and L, together, are a chemical functional group;

each R_3 and R_4 is, independently, H, C_1-C_{10} alkyl, a nitrogen protecting group, or R_3 and R_4, together, are a nitrogen protecting group;

or R_3 and R_4 are joined in a ring structure that optionally includes an additional heteroatom selected from N and O;

Z_4 is OX, SX, or $N(X)_2$;

each X is, independently, H, C_1-C_8 alkyl, C_1-C_8 haloalkyl, $C(=NH)N(H)R_5$, $C(=O)N(H)R_5$ or $OC(=O)N(H)R_5$;

R_5 is H or C_1-C_8 alkyl;

Z_1, Z_2 and Z_3 comprise a ring system having from about 4 to about 7 carbon atoms or having from about 3 to about 6 carbon atoms and 1 or 2 hetero atoms wherein said hetero atoms are selected from oxygen, nitrogen and sulfur and wherein said ring system is aliphatic, unsaturated aliphatic, aromatic, or saturated or unsaturated heterocyclic;

Z_5 is alkyl or haloalkyl having 1 to about 10 carbon atoms, alkenyl having 2 to about 10 carbon atoms, alkynyl

having 2 to about 10 carbon atoms, aryl having 6 to about 14 carbon atoms, $N(R_1)(R_2)$ OR_1, halo, SR_1 or CN;

each q_1 is, independently, an integer from 1 to 10;

each q_2 is, independently, 0 or 1;

q_3 is 0 or an integer from 1 to 10;

q_4 is an integer from 1 to 10;

q_5 is from 0, 1 or 2; and

provided that when q_3 is 0, q_4 is greater than 1.

Representative substituents groups of Formula I are disclosed in United States Patent Application Serial No. 09/130,973, filed August 7, 1998, entitled "Capped 2'-Oxyethoxy Oligonucleotides," hereby incorporated by reference in its entirety.

Representative cyclic substituent groups of Formula II are disclosed in United States Patent Application Serial No. 09/123,108, filed July 27, 1998, entitled "RNA Targeted 2'-Modified Oligonucleotides that are Conformationally Preorganized," hereby incorporated by reference in its entirety.

Particularly preferred substituent groups include $O[(CH_2)_nO]_mCH_3$, $O(CH_2)_nOCH_3$, $O(CH_2)_nNH_2$, $O(CH_2)_nCH_3$, $O(CH_2)_nONH_2$, and $O(CH_2)_nON[(CH_2)_nCH_3)]_2$, where n and m are from 1 to about 10.

Some preferred oligomeric compounds of the invention contain, at least one nucleoside having one of the following substituent groups: C_1 to C_{10} lower alkyl, substituted lower alkyl, alkaryl, aralkyl, O-alkaryl or O-aralkyl, SH, SCH_3, OCN, Cl, Br, CN, CF_3, OCF_3, $SOCH_3$, SO_2CH_3, ONO_2, NO_2, N_3, NH_2, heterocycloalkyl, heterocycloalkaryl, aminoalkylamino, poly-alkylamino, substituted silyl, an RNA cleaving group, a reporter group, an intercalator, a group for improving the pharmacokinetic properties of an oligomeric compound, or a group for improving the pharmacodynamic properties of an oligomeric compound, and other substituents having similar properties. A preferred modification includes 2'-methoxy-ethoxy [2'-O-$CH_2CH_2OCH_3$, also known as 2'-O-(2-methoxyethyl) or

2'-MOE] (Martin *et al.*, *Helv. Chim. Acta*, 1995, *78*, 486), *i.e.*, an alkoxyalkoxy group. A further preferred modification is 2'-dimethylaminooxyethoxy, *i.e.*, a $O(CH_2)_2ON(CH_3)_2$ group, also known as 2'-DMAOE. Representative aminooxy substituent groups are described in co-owned United States Patent Application serial number 09/344,260, filed June 25, 1999, entitled "Aminooxy-Functionalized Oligomers"; and a United States Patent Application entitled "Aminooxy-Functionalized Oligomers and Methods for Making Same," filed August 9, 1999, presently identified by attorney docket number ISIS-3993, hereby incorporated by reference in their entirety.

Other preferred modifications include 2'-methoxy (2'-O-CH_3), 2'-aminopropoxy (2'-$OCH_2CH_2CH_2NH_2$) and 2'-fluoro (2'-F). Similar modifications may also be made at other positions on nucleosides and oligomers, particularly the 3' position of the sugar on the 3' terminal nucleoside or in 2'-5' linked oligomers and the 5' position of 5' terminal nucleoside. Oligomers may also have sugar mimetics such as cyclobutyl moieties in place of the pentofuranosyl sugar. Representative United States patents that teach the preparation of such modified sugars structures include, but are not limited to, U.S. Patents 4,981,957; 5,118,800; 5,319,080; 5,359,044; 5,393,878; 5,446,137; 5,466,786; 5,514,785; 5,519,134; 5,567,811; 5,576,427; 5,591,722; 5,597,909; 5,610,300; 5,627,0531 5,639,873; 5,646,265; 5,658,873; 5,670,633; and 5,700,920, certain of which are commonly owned, and each of which is herein incorporated by reference, and commonly owned United States patent application 08/468,037, filed on June 5, 1995, also herein incorporated by reference.

Representative guanidino substituent groups that are shown in formula III and IV are disclosed in co-owned United States Patent Application 09/349,040 entitled "Functionalized Oligomers," filed July 7, 1999, hereby incorporated by reference in its entirety.

Representative acetamido substituent groups are disclosed in a United States Patent Application entitled "2'-O-Acetamido Modified Monomers and Oligomers," filed August 19, 1999, presently identified by attorney docket number ISIS-4071, hereby incorporated by reference in its entirety.

Representative dimethylaminoethyloxyethyl substituent groups are disclosed in an International Patent Application entitled "2'-O-Dimethylaminoethyloxyethyl-Modified Oligonucleotides," filed August 6, 1999, presently identified by attorney docket number ISIS-4045, hereby incorporated by reference in its entirety.

Several 2'-substituents confer a 3'-endo pucker to the sugar where they are incorporated. This pucker conformation further assists in increasing the Tm of the oligonucleotide with its target.

The high binding affinity resulting from 2' substitution has been partially attributed to the 2′ substitution causing a C3′ *endo* sugar pucker which in turn may give the oligomer a favorable A-form like geometry. This is a reasonable hypothesis since substitution at the 2′ position by a variety of electronegative groups (such as fluoro and *O*-alkyl chains) has been demonstrated to cause C3′ *endo* sugar puckering (De Mesmaeker *et al.*, *Acc. Chem. Res.*, **1995**, *28*, 366-374; Lesnik *et al.*, *Biochemistry*, **1993**, *32*, 7832-7838).

In addition, for 2'-substituents containing an ethylene glycol motif, a *gauche* interaction between the oxygen atoms around the O-C-C-O torsion of the side chain may have a stabilizing effect on the duplex (Freier *et al.*,Nucleic Acids Research, (1997) 25:4429-4442). Such *gauche* interactions have been observed experimentally for a number of years (Wolfe *et al.*, *Acc. Chem. Res.*, **1972**, *5*, 102; Abe *et al.*, *J. Am. Chem. Soc.*, **1976**, 98, 468). This *gauche* effect may result in a configuration of the side chain that is favorable for duplex formation. The exact nature of this stabilizing configuration

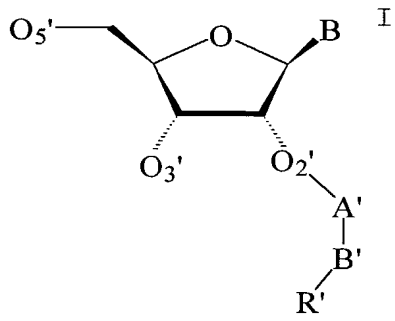
has not yet been explained. While we do not want to be bound by theory, it may be that holding the O-C-C-O torsion in a single *gauche* configuration, rather than a more random distribution seen in an alkyl side chain, provides an entropic advantage for duplex formation.

To better understand the higher RNA affinity of 2'-*O*-methoxyethyl substituted RNA and to examine the conformational properties of the 2'-*O*-methoxyethyl substituent, a self-complementary dodecamer oligonucleotide

2'-*O*-MOE r(CGCGAAUUCGCG) SEQ ID NO: 13

was synthesized, crystallized and its structure at a resolution of 1.7 Ångstrom was determined. The crystallization conditions used were 2 mM oligonucleotide, 50 mM Na Hepes pH 6.2-7.5, 10.50 mM $MgCl_2$, 15% PEG 400. The crystal data showed: space group *C*2, cell constants *a*=41.2 Å, *b*=34.4 Å, *c*=46.6 Å, β=92.4°. The resolution was 1.7 Å at -170°C. The current *R*=factor was 20% (R_{free} 26%).

This crystal structure is believed to be the first crystal structure of a fully modified RNA oligonucleotide analogue. The duplex adopts an overall A-form conformation and all modified sugars display C3'-*endo* pucker. In most of the 2'-*O*-substituents, the torsion angle around the A'-B' bond, as depicted in Structure II below, of the ethylene glycol linker has a *gauche* conformation. For 2'-O-MOE, A' and B' of Structure II below are methylene moieties of the ethyl portion of the MOE and R' is the methoxy portion.



II

In the crystal, the 2'-O-MOE RNA duplex adopts a general orientation such that the crystallographic 2-fold rotation axis does not coincide with the molecular 2-fold rotation axis. The duplex adopts the expected A-type geometry and all of the 24 2'-O-MOE substituents were visible in the electron density maps at full resolution. The electron density maps as well as the temperature factors of substituent atoms indicate flexibility of the 2'-O-MOE substituent in some cases.

Most of the 2'-O-MOE substituents display a *gauche* conformation around the C-C bond of the ethyl linker. However, in two cases, a *trans* conformation around the C-C bond is observed. The lattice interactions in the crystal include packing of duplexes against each other via their minor grooves. Therefore, for some residues, the conformation of the 2'-O-substituent is affected by contacts to an adjacent duplex. In general, variations in the conformation of the substituents (*e.g.* g^+ or g^- around the C-C bonds) create a range of interactions between substituents, both inter-strand, across the minor groove, and intra-strand. At one location, atoms of substituents from two residues are in van der Waals contact across the minor groove. Similarly, a close contact occurs between atoms of substituents from two adjacent intra-strand residues.

Previously determined crystal structures of A-DNA duplexes were for those that incorporated isolated 2'-O-methyl T residues. In the crystal structure noted above for the 2'-O-MOE substituents, a conserved hydration pattern has been observed for the 2'-O-MOE residues. A single water molecule is seen located between O2', O3' and the methoxy oxygen atom of the substituent, forming contacts to all three of between 2.9 and 3.4 Å. In addition, oxygen atoms of substituents are involved in several other hydrogen bonding contacts. For example, the methoxy oxygen atom of a particular 2'-O-

substituent forms a hydrogen bond to N3 of an adenosine from the opposite strand via a bridging water molecule.

In several cases a water molecule is trapped between the oxygen atoms O2′, O3′ and OC′ of modified nucleosides. 2′-*O*-MOE substituents with *trans* conformation around the C-C bond of the ethylene glycol linker are associated with close contacts between OC′ and N2 of a guanosine from the opposite strand, and, water-mediated, between OC′ and N3(G). When combined with the available thermodynamic data for duplexes containing 2′-*O*-MOE modified strands, this crystal structure allows for further detailed structure-stability analysis of other antisense modifications.

In extending the crystallographic structure studies, molecular modeling experiments were performed to study further enhanced binding affinity of oligonucleotides having 2'-O-modifications of the invention. The computer simulations were conducted on compounds of SEQ ID NO: 1, above, having 2'-O-modifications of the invention located at each of the nucleoside of the oligonucleotide. The simulations were performed with the oligonucleotide in aqueous solution using the AMBER force field method (Cornell *et al., J. Am. Chem. Soc.*, **1995**, *117*, 5179-5197)(modeling software package from UCSF, San Francisco, CA). The calculations were performed on an Indigo2 SGI machine (Silicon Graphics, Mountain View, CA).

Further 2'-O-modifications of the inventions include those having a ring structure that incorporates a two atom portion corresponding to the A′ and B′ atoms of Structure II. The ring structure is attached at the 2' position of a sugar moiety of one or more nucleosides that are incorporated into an oligonucleotide. The 2'-oxygen of the nucleoside links to a carbon atom corresponding to the A′ atom of Structure II. These ring structures can be aliphatic, unsaturated aliphatic, aromatic or heterocyclic. A further atom of the ring (corresponding to the B′ atom of Structure II), bears a further oxygen atom, or a sulfur or nitrogen atom. This

oxygen, sulfur or nitrogen atom is bonded to one or more hydrogen atoms, alkyl moieties, or haloalkyl moieties, or is part of a further chemical moiety such as a ureido, carbamate, amide or amidine moiety. The remainder of the ring structure restricts rotation about the bond joining these two ring atoms. This assists in positioning the "further oxygen, sulfur or nitrogen atom" (part of the R position as described above) such that the further atom can be located in close proximity to the 3'-oxygen atom (O3') of the nucleoside.

The ring structure can be further modified with a group useful for modifying the hydrophilic and hydrophobic properties of the ring to which it is attached and thus the properties of an oligonucleotide that includes the 2'-O-modifications of the invention. Further groups can be selected as groups capable of assuming a charged structure, e.g. an amine. This is particularly useful in modifying the overall charge of an oligonucleotide that includes a 2'-O-modifications of the invention. When an oligonucleotide is linked by charged phosphate groups, e.g. phosphorothioate or phosphodiester linkages, location of a counter ion on the 2'-O-modification, e.g. an amine functionality, locally naturalizes the charge in the local environment of the nucleotide bearing the 2'-O-modification. Such neutralization of charge will modulate uptake, cell localization and other pharmacokinetic and pharmacodynamic effects of the oligonucleotide.

Preferred ring structures of the invention for inclusion as a 2'-O modification include cyclohexyl, cyclopentyl and phenyl rings as well as heterocyclic rings having spacial footprints similar to cyclohexyl, cyclopentyl and phenyl rings. Particularly preferred 2'-O-substituent groups of the invention are listed below including an abbreviation for each:

2'-O-(trans 2-methoxy cyclohexyl) -- 2'-O-(TMCHL)

2'-O-(trans 2-methoxy cyclopentyl) -- 2'-O-(TMCPL)

2'-O-(trans 2-ureido cyclohexyl) -- 2'-O-(TUCHL)
2'-O-(trans 2-methoxyphenyl) -- 2'-O-(2MP)

Structural details for duplexes incorporating such 2-O-substituents were analyzed using the described AMBER force field program on the Indigo2 SGI machine. The simulated structure maintained a stable A-form geometry throughout the duration of the simulation. The presence of the 2' substitutions locked the sugars in the C3'-endo conformation.

The simulation for the TMCHL modification revealed that the 2'-O-(TMCHL) side chains have a direct interaction with water molecules solvating the duplex. The oxygen atoms in the 2'-O-(TMCHL) side chain are capable of forming a water-mediated interaction with the 3' oxygen of the phosphate backbone. The presence of the two oxygen atoms in the 2'-O-(TMCHL) side chain gives rise to favorable gauche interactions. The barrier for rotation around the O-C-C-O torsion is made even larger by this novel modification. The preferential preorganization in an A-type geometry increases the binding affinity of the 2'-O-(TMCHL) to the target RNA. The locked side chain conformation in the 2'-O-(TMCHL) group created a more favorable pocket for binding water molecules. The presence of these water molecules played a key role in holding the side chains in the preferable gauche conformation. While not wishing to be bound by theory, the bulk of the substituent, the diequatorial orientation of the substituents in the cyclohexane ring, the water of hydration and the potential for trapping of metal ions in the conformation generated will additionally contribute to improved binding affinity and nuclease resistance of oligonucleotides incorporating nucleosides having this 2'-O-modification.

As described for the TMCHL modification above, identical computer simulations of the 2'-O-(TMCPL), the 2'-O-(2MP) and 2'-O-(TUCHL) modified oligonucleotides in aqueous solution

also illustrate that stable A-form geometry will be maintained throughout the duration of the simulation. The presence of the 2' substitution will lock the sugars in the C3'-endo conformation and the side chains will have direct interaction with water molecules solvating the duplex. The oxygen atoms in the respective side chains are capable of forming a water-mediated interaction with the 3' oxygen of the phosphate backbone. The presence of the two oxygen atoms in the respective side chains give rise to the favorable gauche interactions. The barrier for rotation around the respective O-C-C-O torsions will be made even larger by respective modification. The preferential preorganization in A-type geometry will increase the binding affinity of the respective 2'-O-modified oligonucleotides to the target RNA. The locked side chain conformation in the respective modifications will create a more favorable pocket for binding water molecules. The presence of these water molecules plays a key role in holding the side chains in the preferable gauche conformation. The bulk of the substituent, the diequatorial orientation of the substituents in their respective rings, the water of hydration and the potential trapping of metal ions in the conformation generated will all contribute to improved binding affinity and nuclease resistance of oligonucleotides incorporating nucleosides having these respective 2'-O-modification.

Preferred for use as the B-form nucleotides for eliciting RNase H are ribonucleotides having 2'-deoxy-2'-S-methyl, 2'-deoxy-2'-methyl, 2'-deoxy-2'-amino, 2'-deoxy-2'-mono or dialkyl substituted amino, 2'-deoxy-2'-fluoromethyl, 2'-deoxy-2'-difluoromethyl, 2'-deoxy-2'-trifluoromethyl, 2'-deoxy-2'-methylene, 2'-deoxy-2'-fluoromethylene, 2'-deoxy-2'-difluoromethylene, 2'-deoxy-2'-ethyl, 2'-deoxy-2'-ethylene and 2'-deoxy-2'-acetylene. These nucleotides can alternately be described as 2'-SCH_3 ribonucleotide, 2'-CH_3 ribonucleotide, 2'-NH_2 ribonucleotide 2'-$NH(C_1-C_2$ alkyl) ribonucleotide, 2'-

$N(C_1-C_2$ alkyl$)_2$ ribonucleotide, 2'-CH_2F ribonucleotide, 2'-CHF_2 ribonucleotide, 2'-CF_3 ribonucleotide, 2'=CH_2 ribonucleotide, 2'=CHF ribonucleotide, 2'=CF_2 ribonucleotide, 2'-C_2H_5 ribonucleotide, 2'-CH=CH_2 ribonucleotide, 2'-C≡CH ribonucleotide. A further useful ribonucleotide is one having a ring located on the ribose ring in a cage-like structure including 3',O,4'-C-methyleneribonucleotides. Such cage-like structures will physically fix the ribose ring in the desired conformation.

Additionally, preferred for use as the B-form nucleotides for eliciting RNase H are arabino nucleotides having 2'-deoxy-2'-cyano, 2'-deoxy-2'-fluoro, 2'-deoxy-2'-chloro, 2'-deoxy-2'-bromo, 2'-deoxy-2'-azido, 2'-methoxy and the unmodified arabino nucleotide (that includes a 2'-OH projecting upwards towards the base of the nucleotide). These arabino nucleotides can alternately be described as 2'-CN arabino nucleotide, 2'-F arabino nucleotide, 2'-Cl arabino nucleotide, 2'-Br arabino nucleotide, 2'-N_3 arabino nucleotide, 2'-O-CH_3 arabino nucleotide and arabino nucleotide.

Such nucleotides are linked together via phosphorothioate, phosphorodithioate, boranophosphate or phosphodiester linkages. particularly preferred is the phosphorothioate linkage.

Illustrative of the B-form nucleotides for use in the invention is a 2'-S-methyl (2'-SMe) nucleotide that resides in C2' endo conformation. It can be compared to 2'-O-methyl (2'-OMe)nucleotides that resides in a C3' endo conformation. Particularly suitable for use in comparing these two nucleotides are molecular dynamic investigations using a SGI [Silicon Graphics, Mountain View, CA] computer and the AMBER [UCSF, San Francisco, CA] modeling software package for computer simulations.

Ribose conformations in C2'-modified nucleosides containing S-methyl groups were examined. To understand the

influence of 2'-O-methyl and 2'-S-methyl groups on the conformation of nucleosides, we evaluated the relative energies of the 2'-O- and 2'-S-methylguanosine, along with normal deoxyguanosine and riboguanosine, starting from both C2'-endo and C3'-endo conformations using *ab initio* quantum mechanical calculations. All the structures were fully optimized at HF/6-31G* level and single point energies with electron-correlation were obtained at the MP2/6-31G*//HF/6-31G* level. As shown in Table 1, the C2'-endo conformation of deoxyguanosine is estimated to be 0.6 kcal/mol more stable than the C3'-endo conformation in the gas-phase. The conformational preference of the C2'-endo over the C3'-endo conformation appears to be less dependent upon electron correlation as revealed by the MP2/6-31G*//HF/6-31G* values which also predict the same difference in energy. The opposite trend is noted for riboguanosine. At the HF/6-31G* and MP2/6-31G*//HF/6-31G* levels, the C3'-endo form of riboguanosine is shown to be about 0.65 and 1.41 kcal/mol more stable than the C2'endo form, respectively.

Table 1

Relative energies* of the C3'-endo and C2'-endo conformations of representative nucleosides.

	HF/6-31G	MP2/6-31-G	CONTINUUM MODEL	AMBER
dG	0.60	0.56	0.88	0.65
rG	-0.65	-1.41	-0.28	-2.09
2'-O-MeG	-0.89	-1.79	-0.36	-0.86
2'-S-MeG	2.55	1.41	3.16	2.43

*energies are in kcal/mol relative to the C2'-endo conformation

Table 1 also includes the relative energies of 2'-O-methylguanosine and 2'-S-methylguanosine in C2'-endo and C3'-endo conformation. This data indicates the electronic nature of C2'-substitution has a significant impact on the relative stability of these conformations. Substitution of the 2'-O-methyl group increases the preference for the C3'-endo conformation (when compared to riboguanosine) by about 0.4 kcal/mol at both the HF/6-31G* and MP2/6-31G*//HF/6-31G* levels. In contrast, the 2'-S-methyl group reverses the trend. The C2'-endo conformation is favored by about 2.6 kcal/mol at the HF/6-31G* level, while the same difference is reduced to 1.41 kcal/mol at the MP2/6-31G*//HF/6-31G* level. For comparison, and also to evaluate the accuracy of the molecular mechanical force-field parameters used for the 2'-O-methyl and 2'-S-methyl substituted nucleosides, we have calculated the gas phase energies of the nucleosides. The results reported in Table 1 indicate that the calculated relative energies of these nucleosides compare qualitatively well with the *ab initio* calculations.

Additional calculations were also performed to gauge the effect of solvation on the relative stability of nucleoside conformations. The estimated solvation effect using HF/6-31G* geometries confirms that the relative energetic preference of the four nucleosides in the gas-phase is maintained in the aqueous phase as well (Table 1). Solvation effects were also examined using molecular dynamics simulations of the nucleosides in explicit water. From these trajectories, one can observe the predominance of C2'-endo conformation for deoxyriboguanosine and 2'-S-methylriboguanosine while riboguanosine and 2'-O-methylriboguanosine prefer the C3'-endo conformation. These results are in much accord with the available NMR results on 2'-S-methylribonucleosides. NMR studies of sugar puckering equilibrium using vicinal spin-coupling constants have indicated that the conformation of the sugar ring in 2'-S-methylpyrimidine nucleosides show an

average of >75% S-character, whereas the corresponding purine analogs exhibit an average of >90% S-pucker [Fraser, A., Wheeler, P., Cook, P.D. and Sanghvi, Y.S., *J. Heterocycl. Chem.*, **1993**, *30*, 1277-1287]. It was observed that the 2'-S-methyl substitution in deoxynucleoside confers more conformational rigidity to the sugar conformation when compared with deoxyribonucleosides.

Structural features of DNA:RNA, OMe_DNA:RNA and SMe_DNA:RNA hybrids were also observed. The average RMS deviation of the DNA:RNA structure from the starting hybrid coordinates indicate the structure is stabilized over the length of the simulation with an approximate average RMS deviation of 1.0 Å. This deviation is due, in part, to inherent differences in averaged structures (i.e. the starting conformation) and structures at thermal equilibrium. The changes in sugar pucker conformation for three of the central base pairs of this hybrid are in good agreement with the observations made in previous NMR studies. The sugars in the RNA strand maintain very stable geometries in the C3'-endo conformation with ring pucker values near 0°. In contrast, the sugars of the DNA strand show significant variability.

The average RMS deviation of the OMe_DNA:RNA is approximately 1.2 Å from the starting A-form conformation; while the SMe_DNA:RNA shows a slightly higher deviation (approximately 1.8 Å) from the starting hybrid conformation. The SMe_DNA strand also shows a greater variance in RMS deviation, suggesting the S-methyl group may induce some structural fluctuations. The sugar puckers of the RNA complements maintain C3'-endo puckering throughout the simulation. As expected from the nucleoside calculations, however, significant differences are noted in the puckering of the OMe_DNA and SMe_DNA strands, with the former adopting C3'-endo, and the latter, C1'-exo/C2'-endo conformations.

An analysis of the helicoidal parameters for all three hybrid structures has also been performed to further

characterize the duplex conformation. Three of the more important axis-basepair parameters that distinguish the different forms of the duplexes, X-displacement, propeller twist, and inclination, are reported in Table 2. Usually, an X-displacement near zero represents a B-form duplex; while a negative displacement, which is a direct measure of deviation of the helix from the helical axis, makes the structure appear more A-like in conformation. In A-form duplexes, these values typically vary from -4Å to -5Å. In comparing these values for all three hybrids, the SMe_DNA:RNA hybrid shows the most deviation from the A-form value, the OMe_DNA:RNA shows the least, and the DNA:RNA is intermediate. A similar trend is also evident when comparing the inclination and propeller twist values with ideal A-form parameters. These results are further supported by an analysis of the backbone and glycosidic torsion angles of the hybrid structures. Glycosidic angles (X) of A-form geometries, for example, are typically near -159° while B form values are near -102°. These angles are found to be -162°, -133°, and -108° for the OMe_DNA, DNA, and SMe_DNA strands, respectively. All RNA complements adopt an X angle close to -160°. In addition, "crankshaft" transitions were also noted in the backbone torsions of the central UpU steps of the RNA strand in the SMe_DNA:RNA and DNA;RNA hybrids. Such transitions suggest some local conformational changes may occur to relieve a less favorable global conformation. Taken overall, the results indicate the amount of A-character decreases as OMe_DNA:RNA>DNA:RNA>SMe_DNA:RNA, with the latter two adopting more intermediate conformations when compared to A- and B-form geometries.

Table 2

Average helical parameters derived from the last 500 ps of simulation time.

(canonical A-and B-form values are given for comparison)

Helicoidal Parameter	B-DNA (x-ray)	B-DNA (fibre)	A-DNA (fibre)	DNA:RNA	OMe_DNA: RNA	SMe_DNA: RNA
X-disp	1.2	0.0	-5.3	-4.5	-5.4	-3.5
Inclination	-2.3	1.5	20.7	11.6	15.1	0.7
Propeller	-16.4	-13.3	-7.5	-12.7	-15.8	-10.3

Stability of C2′-modified DNA:RNA hybrids was determined. Although the overall stability of the DNA:RNA hybrids depends on several factors including sequence-dependencies and the purine content in the DNA or RNA strands DNA:RNA hybrids are usually less stable than RNA:RNA duplexes and, in some cases, even less stable than DNA:DNA duplexes. Available experimental data attributes the relatively lowered stability of DNA:RNA hybrids largely to its intermediate conformational nature between DNA:DNA (B-family) and RNA:RNA (A-family) duplexes. The overall thermodynamic stability of nucleic acid duplexes may originate from several factors including the conformation of backbone, base-pairing and stacking interactions. While it is difficult to ascertain the individual thermodynamic contributions to the overall stabilization of the duplex, it is reasonable to argue that the major factors that promote increased stability of hybrid duplexes are better stacking interactions (electrostatic π-π_interactions) and more favorable groove dimensions for hydration. The C2′-S-methyl substitution has been shown to destabilize the hybrid duplex. The notable differences in the rise values among the three hybrids may offer some explanation. While the 2′-S-methyl group has a strong influence on decreasing the base-stacking through high rise values (~3.2 Å), the 2′-O-methyl group makes the overall

structure more compact with a rise value that is equal to that of A-form duplexes (~2.6 Å). Despite its overall A-like structural features, the SMe_DNA:RNA hybrid structure possesses an average rise value of 3.2 Å which is quite close to that of B-family duplexes. In fact, some local base-steps (CG steps) may be observed to have unusually high rise values (as high as 4.5Å). Thus, the greater destabilization of 2'-S-methyl substituted DNA:RNA hybrids may be partly attributed to poor stacking interactions.

It has been postulated that RNase H binds to the minor groove of RNA:DNA hybrid complexes, requiring an intermediate minor groove width between ideal A- and B-form geometries to optimize interactions between the sugar phosphate backbone atoms and RNase H. A close inspection of the averaged structures for the hybrid duplexes using computer simulations reveals significant variation in the minor groove width dimensions as shown in Table 3. Whereas the O-methyl substitution leads to a slight expansion of the minor groove width when compared to the standard DNA:RNA complex, the S-methyl substitution leads to a general contraction (approximately 0.9Å). These changes are most likely due to the preferred sugar puckering noted for the antisense strands which induce either A- or B-like single strand conformations. In addition to minor groove variations, the results also point to potential differences in the steric makeup of the minor groove. The O-methyl group points into the minor groove while the S-methyl is directed away towards the major groove. Essentially, the S-methyl group has flipped through the bases into the major groove as a consequence of C2'-endo puckering.

Table 3

Minor groove widths averaged over the last 500 ps of simulation time

Phosphate Distance	DNA:RNA	OMe_DNA: RNA	SMe_DNA: RNA	DNA:RNA (B-form)	RNA:RNA (A-form)
P5-P20	15.27	16.82	13.73	14.19	17.32
P6-P19	15.52	16.79	15.73	12.66	17.12
P7-P18	15.19	16.40	14.08	11.10	16.60
P8-P17	15.07	16.12	14.00	10.98	16.14
P9-P16	15.29	16.25	14.98	11.65	16.93
P10-P15	15.37	16.57	13.92	14.05	17.69

In addition to the modifications described above, the nucleotides of the oligonucleotides of the invention can have a variety of other modification so long as these other modifications do not significantly detract from the properties described above. Thus, for nucleotides that are incorporated into oligonucleotides of the invention, these nucleotides can have sugar portions that correspond to naturally-occurring sugars or modified sugars. Representative modified sugars include carbocyclic or acyclic sugars, sugars having substituent groups at their 2' position, sugars having substituent groups at their 3' position, and sugars having substituents in place of one or more hydrogen atoms of the sugar. Other altered base moieties and altered sugar moieties are disclosed in United States Patent 3,687,808 and PCT application PCT/US89/02323.

Altered base moieties or altered sugar moieties also include other modifications consistent with the spirit of this invention. Such oligonucleotides are best described as being structurally distinguishable from, yet functionally interchangeable with, naturally occurring or synthetic wild type oligonucleotides. All such oligonucleotides are

comprehended by this invention so long as they function effectively to mimic the structure of a desired RNA or DNA strand. A class of representative base modifications include tricyclic cytosine analog, termed "G clamp" (Lin, *et al., J. Am. Chem. Soc.* **1998**, 120, 8531). This analog makes four hydrogen bonds to a complementary guanine (G) within a helix by simultaneously recognizing the Watson-Crick and Hoogsteen faces of the targeted G. This G clamp modification when incorporated into phosphorothioate oligonucleotides, dramatically enhances antisense potencies in cell culture. The oligonucleotides of the invention also can include phenoxazine-substituted bases of the type disclosed by Flanagan, *et al., Nat. Biotechnol.* **1999**, *17(1)*, 48-52.

Additional modifications may also be made at other positions on the oligonucleotide, particularly the 3' position of the sugar on the 3' terminal nucleotide and the 5' position of 5' terminal nucleotide. For example, one additional modification of the oligonucleotides of the invention involves chemically linking to the oligonucleotide one or more moieties or conjugates which enhance the activity, cellular distribution or cellular uptake of the oligonucleotide. Such moieties include but are not limited to lipid moieties such as a cholesterol moiety (Letsinger *et al., Proc. Natl. Acad. Sci. USA*, 1989, *86*, 6553), cholic acid (Manoharan *et al., Bioorg. Med. Chem. Lett.*, 1994, *4*, 1053), a thioether, e.g., hexyl-*S*-tritylthiol (Manoharan *et al., Ann. N.Y. Acad. Sci.*, 1992, *660*, 306; Manoharan *et al., Bioorg. Med. Chem. Let.*, 1993, *3*, 2765), a thiocholesterol (Oberhauser *et al., Nucl. Acids Res.*, 1992, *20*, 533), an aliphatic chain, e.g., dodecandiol or undecyl residues (Saison-Behmoaras *et al., EMBO J.*, 1991, *10*, 111; Kabanov *et al., FEBS Lett.*, 1990, *259*, 327; Svinarchuk *et al., Biochimie*, 1993, *75*, 49), a phospholipid, e.g., di-hexadecyl-*rac*-glycerol or triethylammonium 1,2-di-O-hexadecyl-*rac*-glycero-3-H-phosphonate (Manoharan *et al.,*

Tetrahedron Lett., 1995, *36*, 3651; Shea *et al.*, *Nucl. Acids Res.*, 1990, *18*, 3777), a polyamine or a polyethylene glycol chain (Manoharan *et al.*, *Nucleosides & Nucleotides*, 1995, *14*, 969), or adamantane acetic acid (Manoharan *et al.*, *Tetrahedron Lett.*, 1995, *36*, 3651), a palmityl moiety (Mishra *et al.*, *Biochim. Biophys. Acta*, 1995, *1264*, 229), or an octadecylamine or hexylamino-carbonyl-oxycholesterol moiety (Crooke *et al.*, *J. Pharmacol. Exp. Ther.*, 1996, *277*, 923).

Human RNase H1 displays a strong positional preference for cleavage, *i.e.*, it cleaves between 8-12 nucleotides from the 5'-RNA:3'-DNA terminus of the duplex. Within the preferred cleavage site, the enzyme displays modest sequence preference with GU being a preferred dinucleotide. The minimum RNA:DNA duplex length that supports cleavage is 6-base pairs and the minimum RNA:DNA "gap size" that supports cleavage is 5-base pairs.

Properties of purified Human RNase H1

The effects of various reaction conditions on the activity of Human RNase H1 were evaluated (Fig. 1). The optimal pH for the enzyme in both Tris HCl and phosphate buffers was 7.0-8.0. At pH's above pH8.0, enzyme activity was reduced. However, this could be due to instability of the substrate or effects on the enzyme, or both. To evaluate the potential contribution of changes in ionic strength to the activities observed at different pHs, two buffers, $NaHPO_4$ and Tris HCl were studied at pH 7.0 and gave the same enzyme activity even though the ionic strengths differed. Enzyme activity was inhibited by increasing ionic strength (Fig. 1B) and N-ethymaleamide (Fig. 1C). Enzyme activity increased as the temperature was raised from 25-42° C (Fig. 1D). Mg^{2+} stimulated enzyme activity with an optimal concentration of 1mM. At higher concentrations, Mg^{2+} was inhibitory (Fig. 1E). In the presence of 1mM Mg^{2+}, Mn^{2+} was inhibitory at all

concentrations tested (Fig. 1F). The purified enzyme was quite stable and easily handled. In fact, the enzyme could be boiled and rapidly or slowly cooled without significant loss of activity (Fig. 1D). The initial rates of cleavage were determined for four duplex substrates studied simultaneously. The initial rate of cleavage for a phosphodiester DNA:RNA duplex was 1050 ± 203 pmol $L^{-1}min^{-1}$ (Table 4A). The initial rate of cleavage of a phosphorothioate oligodeoxynucleotide duplex was approximately four-fold faster than that of the same duplex comprised of a phosphodiester antisense oligodeoxynucleotide (Table 4A). The initial rates for 17-mer and 20-mer substrates of different sequences were equal (Table 4B). However, when a 25-mer heteroduplex, containing the 17-mer sequence in the center of the duplex was digested (RNA 3), the rate was 50% faster. Interestingly, the Km of the enzyme for the 25 mer duplex was 40% lower than that for the 17 mer while the *Vmax's* for both duplexes were the same (see Table 6), suggesting that with the increase in length, a larger number of cleavage sites are available resulting in an increase in the number of productive binding interactions between the enzyme and substrate. As a result, a lower substrate concentration is required for the longer duplex to achieve a cleavage rate equal to that of the shorter duplex.

To better characterize the substrate specificity of Human RNase H1, duplexes in which the antisense oligonucleotide was modified in the 2'-position were studied. As previously reported for *E. coli* RNase H1, Human RNase H1 was unable to cleave substrates with 2'-modifications at the cleavage site of the antisense DNA strand or the sense RNA strand (Table 5). For example, the initial rate of cleavage of a duplex containing a phosphorothioate oligodeoxynucleotide and its complement was 3400 pmol $L^{-1}min^{-1}$ while that of its 2'-propoxy modified analog was undetectable (Table 5). A duplex comprised of a fully modified 2'-methoxy antisense strand also

failed to support any cleavage (Table 5). The placement of 2'-methoxy modifications around a central region of oligodeoxynucleotides reduced the initial rate (Table 5). The smaller the central oligodeoxynucleotide "gap" the lower the initial rate. The smallest "gapmer" for which cleavage could be measured was a 5 deoxynucleotide gap. These data are highly consistent with observations we have previously reported for *E. coli* RNase H1 except that for the bacterial enzyme the minimum gap size was 4 deoxynucleotides.

The Km and *Vmax* of Human RNase H1 for three substrates are shown in Table 6. The Km valves for all three substrates were substantially lower than those of *E. coli* RNase H1 (Table 6). As previously reported for *E. coli* RNase H1, the Km for a phosphorothioate containing duplex was lower than that of a phosphodiester duplex. The *Vmax* of the human enzyme was 30 fold lower than that of the *E. coli* enzyme. The *Vmax* for the phosphorothioate containing substrate was less than the phosphodiester duplex. This is probably due to inhibition of the enzyme at higher concentrations by excess phosphorothioate single strand oligonucleotide as the initial rate of cleavage for a phosphorothioate containing duplex was, in fact, greater than the phosphodiester (Table 4)

Binding affinity and specificity

To evaluate the binding affinity of Human RNase H1, a competitive cleavage assay in which increasing concentrations of noncleavable substrates were added was used. Using this approach, the Ki is formally equivalent to the Kd for the competing substrates. Of the noncleavable substrates studied, Lineweaver-Burk analyses demonstrated that all inhibitors shown in Table 7 were competitive (data not shown). A duplex containing a phosphodiester oligodeoxynucleotide hybridized to a phosphodiester 2'-methoxy oligonucleotide as the noncleavable substrate is considered most like DNA:RNA. Table

7 shows the results of these studies and compares them to previously reported results for the *E. coli* enzyme performed under similar conditions. Clearly, the affinity of the human enzyme for its DNA:RNA like substrate (DNA:2'-O-Me) was substantially greater than that of the *E. coli* enzyme, consistent with the differences observed in Km (Table 6).

E. coli RNase H1 displays approximately equal affinity for RNA:RNA, RNA:2'-O-Me and DNA:2'-O-Me duplexes (Table 7). The human enzyme displays similar binding properties, but is more able to discriminate between various duplexes. For example, the Kd for RNA:RNA was approximately 5 fold lower than the Kd for DNA:2'-O-Me. This is further demonstrated by the Kd for the RNA:2'F duplex. The Kd for the DNA:2'-F duplex was slightly greater than for the RNA:2'-F duplex and the RNA:RNA duplex, but clearly lower than for other duplexes. Thus, both enzymes can be considered double strand RNA binding proteins. However, Human RNase H1 is somewhat less specific for duplexes as compared to single strand oligonucleotides than the *E. coli* enzyme. The enzyme bound to single strand RNA and DNA only 20 fold less well than an RNA:RNA duplex while the *E. coli* enzyme bound to single strand DNA nearly 600 fold less than to an RNA:RNA duplex (Table 7). The affinity of a single strand phosphorothioate oligodeoxynucleotide for both enzymes was significant relative to the affinity for the natural substrate and accounts for the inhibition of the enzymes by members of this class oligonucleotides. Remarkably, Human RNase H1 displayed the highest affinity for a single strand phosphorothioate oligodeoxynucleotide. Thus, this noncleavable substrate is a very effective inhibitor of the enzyme and excess phosphorothioate antisense drug in cells might be highly inhibitory.

Site and sequence preferences for cleavage

Fig. 2 shows the cleavage pattern for RNA duplexed with its phosphorothioate oligodeoxynucleotide and the pattern for

several gapmers. In the parent duplex, RNA cleavage occurred at a single major site with minor cleavage noted at several sites 3′ to this major cleavage site that was 8 nucleotides from 5′-terminus of the RNA. Note that the preferred site occured at a GU dinucleotide. Cleavage of several "gapmers" occurred more slowly and the major cleavage site was at a different position from that of the parent duplex. Further, in contrast to the observations we have made for *E. coli* RNase H1, the major cleavage site in gapmers treated with Human RNase H1 did not occur at the nucleotide apposed to the nucleotide adjacent to the first 2′ methoxy nucleotide in the wing hybridized to the 3′ portion of the RNA.

To further evaluate the site and sequence specificities of Human RNase H1, cleavage of substrates shown in Fig. 3 and Fig. 4 was studied. In Fig. 3, the sequence of the RNA is displayed below the sequencing gels and the length and position of the complementary phosphodiester oligodeoxynucleotide indicated by the solid line below the RNA sequence. This figure demonstrates several important properties of the enzyme. First, the main cleavage site was consistently observed 8-9 nucleotides from the 5′-RNA:3′-DNA terminus of the duplex irrespective of whether there were 5′ or 3′-RNA single strand overhangs. Second, the enzyme, like *E. coli* RNase H1 was capable of cleaving single strand regions of RNA adjacent to the 3′-terminus of an RNA:DNA duplex. Third, the minimum duplex length that supported any cleavage was approximately 6 nucleotides. RNase protection assays were used to confirm that under conditions of the assay the shorter duplexes were fully hybridized, so the differences observed were not due to the failure to hybridize. To assure that the 6-nucleotide duplex was fully hybridized, the reactions were carried out at a 50:1 DNA:RNA ratio (data not shown). Fourth, the figure shows that for duplexes smaller than the nine base pairs, the smaller the duplex, the slower the cleavage rate.

Fifth, the preferred cleavage site was located at a GU dinucleotide.

The site and sequence specificities are further explored in Fig. 4. That the enzyme displays little sequence preference is demonstrated by comparing the rates and sites of cleavage for duplexes A, C and D. In all cases, the preferred site of cleavage was 8-12 nucleotides from the 5'-RNA:3'-DNA terminus of the duplex irrespective of the sequence. Comparison of the cleavage pattern for duplexes A and B shows that cleavage occurred at the 8-12 nucleotide position even when there were RNA overhangs also as shown in Fig. 3. Cleavage of duplex F demonstrated that the site of cleavage was retained even if there were 5'- and 3'-DNA overhangs. In a longer substrate, duplex G, the main site of cleavage was still 8-12 nucleotides from the terminus of the duplex. However, minor cleavage sites were observed throughout the RNA suggesting that this substrate might support binding of more than one enzyme molecule per substrate, but that the preferred site was near the 5'-RNA:3'-DNA terminus. Finally, optimal cleavage seemed to occur when a GU dinucleotide was located 8-12 nucleotides from the 5'-RNA:3'-DNA terminus of the duplex.

To address both the mechanism of cleavage and processivity, the cleavage of 5'-labeled and 3'-labeled substrates was compared (Fig. 5). Lane C shows that CIP treatment prior to and after digestion with Human RNase H1 resulted in a shift in the mobility of the digested fragments suggesting that Human RNase H1 generates cleavage products with 5'-phosphates. Thus, it is similar to *E. coli* RNase H1 in this regard. A second intriguing observation is that the addition of pC to the 3'-end of the RNA caused a shift in the position of the preferred cleavage site (A vs B or C). The four cleavage sites in the center of the duplex observed with a 5'-phosphate labeled RNA were observed in 3' pC-labeled substrates. However, the main cleavage site shifted from base

pair 8 to base pair 12. Interestingly, the sequence at both sites was GU. Thus, it is conceivable that the enzyme selects a position 8-12 nucleotide from the 5'-RNA:3'-DNA terminus, then cleaves at a preferred dinucleotide such as GU. Third, this figure considered along with the cleavage patterns shown in Figs. 3 and 4 demonstrates that this enzyme displays minimal processivity in either the 5' or 3'-direction. In no time course experiment using any substrate have we observed a pattern that would be consistent with processivity. The possibility that the failure to observe processivity in Figs. 3 and 4 was due to processivity in the 3' to 5'-direction is excluded by the results in Fig. 5. Again, this is significantly different from observations we have previously reported for *E. coli* RNase H1.

General properties of Human RNase H1 activity

The present invention also describes the properties of human RNase H1 that have been characterized. As the protein studied is a his-tag fusion and was denatured and refolded, it is possible that the activity of the enzyme in its native state might be greater than we have observed. However, basic properties are certainly likely to reflect the basic properties of the native enzyme. Numerous studies have shown that a his-tag does not interfere with protein folding and crystallization, kinetic and catalytic properties, or nucleic acid binding properties since it is very small (few amino acids) and its pK is near neutral. It is shown in the present invention that the his-tag fusion protein behaves like other RNase H's. It cleaved specifically the RNA strand in RNA:DNA duplexes, resulted in cleavage products with 5'-phosphate termini (Fig. 5) and was affected by divalent cations (Fig. 1). Optimal conditions for Human RNase H1 were similar to, but not identical to, *E. coli* RNase H1. For the human enzyme, the Mg^{2+} optimum was 1mM and 5 mM Mg^{2+} was inhibitory. In the presence of Mg^{2+}, both enzymes were inhibited by Mn^{2+}. The

human enzyme was inhibited by n-ethylmaleimide and was quite stable, easily handled and did not form multimeric structures (Fig. 1). The ease of handling, denaturation, refolding and stability in various conditions suggest that the Human RNase H1 was active as a monomer and has a relatively stable preferred conformation.

Studies on the structure and enzymatic activities of a number of mutants of *E. coli* RNase H1 have recently led to a hypothesis to explain the effects of divalent cations termed an activation/attenuation model. The effects of divalent cations on Human RNase H1 are complex and are consistent with the suggested activation/attenuation model. The amino acids proposed to be involved in both cation binding sites are conserved in Human RNase H1.

Positional and sequence preferences and processivity

The site and sequence specificity of Human RNase H1 differ substantially from *E. coli* RNase H1. Although neither enzyme displays significant sequence specificity and Figs. 2-5 this manuscript, the human enzyme displays remarkable site specificity. Figs. 2-4 show that Human RNase H1 preferentially cleaved 8-12 nucleotides 3′ from the 5′-RNA:3′-DNA terminus of a DNA:RNA duplex irrespective of whether there were 5′ or 3′-RNA or DNA overhangs. The process by which a position is selected and then within that position on the duplex a particular dinucleotide is cleaved preferentially must be relatively complex and influenced by sequence. Clearly, the dinucleotide, GU, is a preferred sequence. In Fig. 3, for example, all the duplexes contained a GU sequence near the optimal position for the enzyme and in all cases the preferential cleavage site was GU. Additionally, in duplexes A and B a second GU was also cleaved, albeit at a very slow rate. The third site in duplexes A and B cleaved was a GG dinucleotide 7 base pairs from the 3′-RNA:5′-DNA terminus. Thus, the data suggest that the enzyme displays strong

positional preference and within the appropriate site, slight preference for GU dinucleotides.

The strong positional preference exhibited by Human RNase H1 suggests that the enzyme fixes its position on the duplex via the 5′-RNA:3′-DNA terminus. Interestingly, the *in-vitro* cleavage pattern observed for the enzyme is compatible with its proposed *in-vivo* role, namely, the removal of RNA primers during DNA replication of the lagging strand. The average length of the RNA primer ranges from 7-14 nucleotides. Consequently, synthesis of the lagging strand results in chimeric sequences consisting of 7-14 ribonucleotides at the 5′-terminus with contiguous stretches of DNA extending in the 3′ direction. The positional preference observed for Human RNase H1, (*i.e.*, 8-12 residues from the 5′-terminus of the RNA), would suggest that cleavage of the chimeric lagging strand by RNase H1 would occur at or near the RNA:DNA junction. The removal of residual ribonucleatides following RNase H digestion has been shown to be performed by the endonuclease FEN1.

Fig. 4 provides additional insight into the positional and sequence preferences of the enzyme. When there was a GU dinucleotide present in the correct position in the duplex, it was cleaved preferentially. When a GU dinucleotide was absent, AU was cleaved as well as other dinucleotides. For duplex G both a GU and a GG dinucleotide were present within the preferred site, and in this case the GG dinucleotide was cleaved slightly more extensively than the GU dinucleotide. Clearly, additional duplexes of different sequences must be studied before definitive conclusions concerning the roles of various sequences within the preferred cleavage sites can be drawn.

In Fig. 5, the 3′-terminus of the RNA was labeled with ^{32}pC. In this case the same four nucleotides were cleaved as when the RNA was 5′ labeled (Fig. 5, panels B & C). However, the GU closer to the 3′-terminus of the RNA was cleaved at

least as rapidly as the 5'-GU. Interestingly in studies on the partially purified enzyme, differences in the cleavage pattern were also observed when 5'-labeled substrates were compared to 3'-labeled substrates. A possible explanation for this observation is that the presence of a 3'-phosphate on an oligonucleotide substrate affects the scanning mechanism the enzyme uses to select preferred positions for cleavage.

In a duplex comprised of RNA annealed to a chimeric oligonucleotide with an oligodeoxynucleotide center flanked by 2'-modified nucleotide wings, the cleavage by Human RNase H1 was directed to the DNA:RNA portion of the duplex as was observed for *E. coli* RNase H1. However, within this region, the preferred sites of cleavage for the human enzyme differed from *E. coli* RNase H1. *E. coli* RNase H1 preferentially cleaved at the ribonucleotide apposed to first 2'-modified nucleotide in the wing of antisense oligonucleotide at the 3'-end of the RNA. In contrast, the human enzyme preferentially cleaved at sites more centered within the gap until the gap was reduced to 5 nucleotides. Further, the minimum gap size for the human enzyme was 5 nucleotides while that of *E. coli* RNase H1 was 4 nucleotides. These differences in behavior suggest differences in the structures of the enzymes and their interactions with substrate that will require additional study.

Although *E. coli* RNase H1 degrades the heteroduplex substrate in a predominantly distributive manner, the enzyme displays modest 5'-3'-processivity. In contrast, Human RNase H1 evidences no 5'-3' or 3'-5'-processivity suggesting that the human enzyme hydrolyzes the substrate in an exclusively distributive manner. The lack of processivity observed with the Human RNase H1 may be a function of the significantly tighter binding affinity (Table 7), thereby reducing the ability of the enzyme to move on the substrate. Alternatively, Human RNase H1 appears to fix its position on the substrate with respect to the 5'-RNA:3'-DNA terminus and

this strong positional preference may preclude cleavage of the substrate in a processive manner. (Fig. 5). Thus, despite the fact that the enzymes are both metal-dependent endonucleases that result in cleavage products with 5'-phosphates (Fig. 5) and both can cleave single-strand 3'-RNA overhangs (Fig. 5), these enzymes display substantial differences.

E. coli RNase H1 has been suggested to exhibit "binding directionality" with respect to the RNA of the substrate such that the primary binding region of the enzyme is positioned several nucleotides 5' to the catalytic center. This results in cleavage sites being restricted from the 5'-RNA:3'-DNA end of a duplex, and cleavage sites occurring at the 3'-RNA:5'-DNA end of the duplex and in 3'-single-strand overhangs. The human enzyme behaves entirely analogously. Thus, we conclude that Human RNase H1 likely has the same binding directionality as the *E. coli* enzyme.

Substrate binding

RNA:RNA duplexes have been shown to adopt an A-form conformation. Many 2'-modifications shift the sugar conformation into a 3'-endo pucker characteristic of RNA. Consequently, when hybridized to RNA, the resulting duplex is "A" form and this is manifested in a more stable duplex. 2'-F Oligonucleotides display duplex forming properties most like RNA, while 2'-methoxy oligonucleotides result in duplexes intermediate information between DNA:RNA and RNA:RNA duplexes.

The results shown in Table 7 demonstrate that like the *E. coli* enzyme, Human RNase H1 is a double strand RNA binding protein. Moreover, it displays some ability to discriminate between various A-form duplexes (Table 7). The observation that the Kd for an RNA:2'-F duplex is equal to that for an RNA:RNA dupex suggests that 2'-hydroxy group is not required for binding to the enzyme. Nevertheless, we cannot exclude the possibility that bulkier 2'-modifications, e.g. 2'-methoxy or 2'-propyl might sterically inhibit the binding of the

enzyme as well as alter the A-form quality of the duplex. The human enzyme displays substantially greater affinity for all oligonucleotides than the *E. coli* enzyme and this is reflected in a lower Km for cleavable substrates (Tables 6 and 7). In addition, the tighter binding affinity observed for Human RNase H1 may be responsible for the 20-fold lower *Vmax* when compared to the *E. coli* enzyme. In this case, assuming that the *E. coli* and human enzymes exhibit similar catalytic rates (Kcat), then an increase in the binding affinity would result in a lower turnover rate and ultimately a lower *Vmax*.

The principal substrate binding site in *E. coli* RNase H1 is thought to be a cluster of lysines that are believed to bind to the phosphates of the substrates. The interaction of the binding surface of the enzyme and substrate is believed to occur within the minor groove. This region is highly conserved in the human enzyme. In addition, eukaryotic enzymes contain an extra N-terminal region of variable length containing an abundance of basic amino acids. This region is homologous with a double strand RNA binding motif and indeed in the S. *cerevasiae* RNase H has been shown to bind to double strand RNA. The N-terminal extension in Human RNase H1 is longer than that in the S. *cerevasiae* enzyme and appears to correspond to a more complete double strand RNA binding motif. Consequently, the enhanced binding of Human RNase H1 to various nucleic acids may be due to the presence of this additional binding site.

As used herein, the term "alkyl" includes but is not limited to straight chain, branch chain, and cyclic unsaturated hydrocarbon groups including but not limited to methyl, ethyl, and isopropyl groups. Alkyl groups of the present invention may be substituted. Representative alkyl substituents are disclosed in United States Patent No. 5,212,295, at column 12, lines 41-50, hereby incorporated by reference in its entirety.

Alkenyl groups according to the invention are to straight chain, branch chain, and cyclic hydrocarbon groups containing at least one carbon-carbon double bond, and alkynyl groups are to straight chain, branch chain, and cyclic hydrocarbon groups containing at least one carbon-carbon triply bond. Alkenyl and alkynyl groups of the present invention can be substituted.

Aryl groups are substituted and unsubstituted aromatic cyclic moieties including but not limited to phenyl, naphthyl, anthracyl, phenanthryl, pyrenyl, and xylyl groups. Alkaryl groups are those in which an aryl moiety links an alkyl moiety to a core structure, and aralkyl groups are those in which an alkyl moiety links an aryl moiety to a core structure.

In general, the term "hetero" denotes an atom other than carbon, preferably but not exclusively N, O, or S. Accordingly, the term "heterocyclic ring" denotes a carbon-based ring system having one or more heteroatoms (*i.e.*, non-carbon atoms). Preferred heterocyclic rings include, for example but not limited to imidazole, pyrrolidine, 1,3-dioxane, piperazine, morpholine rings. As used herein, the term "heterocyclic ring" also denotes a ring system having one or more double bonds, and one or more heteroatoms. Preferred heterocyclic rings include, for example but not limited to the pyrrolidino ring.

Oligonucleotides according to the present invention that are hybridizable to a target nucleic acid preferably comprise from about 5 to about 50 nucleosides. It is more preferred that such compounds comprise from about 8 to about 30 nucleosides, with 15 to 25 nucleosides being particularly preferred. As used herein, a target nucleic acid is any nucleic acid that can hybridize with a complementary nucleic acid-like compound. Further in the context of this invention, "hybridization" shall mean hydrogen bonding, which may be Watson-Crick, Hoogsteen or reversed Hoogsteen hydrogen bonding between complementary nucleobases. "Complementary" as used

herein, refers to the capacity for precise pairing between two nucleobases. For example, adenine and thymine are complementary nucleobases which pair through the formation of hydrogen bonds. "Complementary" and "specifically hybridizable," as used herein, refer to precise pairing or sequence complementarity between a first and a second nucleic acid-like oligomers containing nucleoside subunits. For example, if a nucleobase at a certain position of the first nucleic acid is capable of hydrogen bonding with a nucleobase at the same position of the second nucleic acid, then the first nucleic acid and the second nucleic acid are considered to be complementary to each other at that position. The first and second nucleic acids are complementary to each other when a sufficient number of corresponding positions in each molecule are occupied by nucleobases which can hydrogen bond with each other. Thus, "specifically hybridizable" and "complementary" are terms which are used to indicate a sufficient degree of complementarity such that stable and specific binding occurs between a compound of the invention and a target RNA molecule. It is understood that an oligomeric compound of the invention need not be 100% complementary to its target RNA sequence to be specifically hybridizable. An oligomeric compound is specifically hybridizable when binding of the oligomeric compound to the target RNA molecule interferes with the normal function of the target RNA to cause a loss of utility, and there is a sufficient degree of complementarity to avoid non-specific binding of the oligomeric compound to non-target sequences under conditions in which specific binding is desired, *i.e.* under physiological conditions in the case of *in vivo* assays or therapeutic treatment, or in the case of *in vitro* assays, under conditions in which the assays are performed.

As used herein, "human type 2 RNase H" and "human RNase H1" refer to the same human RNase H enzyme. Accordingly, these terms are meant to be used interchangeably.

The oligonucleotides of the invention can be used in diagnostics, therapeutics and as research reagents and kits. They can be used in pharmaceutical compositions by including a suitable pharmaceutically acceptable diluent or carrier. They further can be used for treating organisms having a disease characterized by the undesired production of a protein. The organism should be contacted with an oligonucleotide having a sequence that is capable of specifically hybridizing with a strand of nucleic acid coding for the undesirable protein. Treatments of this type can be practiced on a variety of organisms ranging from unicellular prokaryotic and eukaryotic organisms to multicellular eukaryotic organisms. Any organism that utilizes DNA-RNA transcription or RNA-protein translation as a fundamental part of its hereditary, metabolic or cellular control is susceptible to therapeutic and/or prophylactic treatment in accordance with the invention. Seemingly diverse organisms such as bacteria, yeast, protozoa, algae, all plants and all higher animal forms, including warm-blooded animals, can be treated. Further, each cell of multicellular eukaryotes can be treated, as they include both DNA-RNA transcription and RNA-protein translation as integral parts of their cellular activity. Furthermore, many of the organelles (*e.g.*, mitochondria and chloroplasts) of eukaryotic cells also include transcription and translation mechanisms. Thus, single cells, cellular populations or organelles can also be included within the definition of organisms that can be treated with therapeutic or diagnostic oligonucleotides.

The following nonlimiting examples are provided to further illustrate the present invention.

EXAMPLES

EXAMPLE 1

5'-*O*-DMT-2'-*O*-(2-methoxyethyl)-5-methyl uridine and 5'-*O*-DMT-3'-*O*-(2-methoxyethyl)5-methyl uridine

2',3'-*O*-dibutylstannylene 5-methyl uridine (345 g) (prepared as per: Wagner *et al.*, *J. Org. Chem.*, **1974**, *39*, 24) was alkylated with 2-methoxyethyl bromide (196 g) in the presence of tetrabutylammonium iodide (235 g) in DMF (3 L) at 70 °C to give a mixture of 2'-*O*- and 3'-*O*-(2-methoxyethyl)-5-methyl uridine (150 g) in nearly 1:1 ratio of isomers. The mixture was treated with DMT chloride (110 g, DMT-Cl) in pyridine (1 L) to give a mixture of the 5'-*O*-DMT-nucleosides. After the standard work-up the isomers were separated by silica gel column chromatography. The 2'-isomer eluted first, followed by the 3'-isomer.

EXAMPLE 2

5'-*O*-DMT-3'-*O*-(2-methoxyethyl)-5-methyl-uridine-2'-*O*-(2-cyanoethyl-N,N-diisopropyl) phosphoramidite

5'-*O*-DMT-3'-*O*-(2-methoxyethyl)-5-methyluridine (5 g, .008 mol) was dissolved in CH_2Cl_2 (30 mL) and to this solution, under argon, diisopropylaminotetrazolide (0.415 g) and 2-cyanoethoxy-N,N-diisopropyl phosphoramidite (3.9 mL) were added. The reaction was stirred overnight. The solvent was evaporated and the residue was applied to silica column and eluted with ethyl acetate to give 3.75 g title compound.

EXAMPLE 3

5'-*O*-DMT-3'-*O*-(2-methoxyethyl)-N^4-benzoyl-5-methyl-cytidine

5'-*O*-DMT-3'-*O*-(2-methoxyethyl)-5-methyl uridine (15 g) was treated with 150 mL anhydrous pyridine and 4.5 mL of acetic anhydride under argon and stirred overnight. Pyridine was evaporated and the residue was partitioned between 200 mL of saturated $NaHCO_3$ solution and 200 mL of ethylacetate. The

organic layer was dried (anhydrous $MgSO_4$) and evaporated to give 16 g of 2'-acetoxy-5'-*O*-(DMT)-3'-*O*-(2-methoxyethyl)-5-methyl uridine.

To an ice-cold solution of triazole (19.9 g) in triethylamine (50 mL) and acetonitrile (150 mL), with mechanical stirring, 9 mL of $POCl_3$ was added dropwise. After the addition, the ice bath was removed and the mixture stirred for 30 min. The 2'-acetoxy-5'-*O*-(DMT)-3'-*O*-(2-methoxyethyl)-5-methyl uridine (16 g in 50 mL CH_3CN) was added dropwise to the above solution with the receiving flask kept at ice bath temperatures. After 2 hrs, TLC indicated a faster moving nucleoside, C-4-triazole-derivative. The reaction flask was evaporated and the nucleoside was partitioned between ethylacetate (500 mL) and $NaHCO_3$ (500 mL). The organic layer was washed with saturated NaCl solution, dried (anhydrous $NgSO_4$) and evaporated to give 15 g of C-4-triazole nucleoside. This compound was then dissolved in 2:1 mixture of NH_4OH/dioxane (100 mL:200 mL) and stirred overnight. TLC indicated disappearance of the starting material. The solution was evaporated and dissolved in methanol to crystallize out 9.6 g of 5'-*O*-(DMT)-3'-*O*-(2-methoxyethyl)5-methyl cytidine.

5'-*O*-DMT-3'-*O*-(2-methoxyethyl)-5-methyl cytidine (9.6 g, 0.015 mol) was dissolved in 50 mL of DMF and treated with 7.37 g of benzoic anhydride. After 24 hrs of stirring, DMF was evaporated and the residue was loaded on silica column and eluted with 1:1 hexane:ethylacetate to give the desired nucleoside.

EXAMPLE 4

5'-*O*-DMT-3'-*O*-(2-methoxyethyl)-N^4-benzoyl-5-methyl-cytidine-2'-*O*-(2-cyanoethyl-N,N-diisopropyl) phosphoramidite

5'-*O*-DMT-3'-*O*-(2-methoxyethyl)-N^4-benzoyl-5-methyl-cytidine-2'-*O*-(2-cyanoethyl-N,N-diisopropyl) phosphoramidite

was obtained from the above nucleoside using the phosphitylation protocol described for the corresponding 5-methyl-uridine derivative.

EXAMPLE 5

N^6-Benzoyl-5'-O-(DMT)- 3'-O-(2-methoxyethyl) adenosine

A solution of adenosine (42.74 g, 0.16 mol) in dry dimethyl formamide (800 mL) at 5 °C was treated with sodium hydride (8.24 g, 60% in oil prewashed thrice with hexanes, 0.21 mol). After stirring for 30 min, 2-methoxyethyl bromide (0.16 mol) was added over 20 min. The reaction was stirred at 5 °C for 8 h, then filtered through Celite. The filtrate was concentrated under reduced pressure followed by coevaporation with toluene (2x100 mL). The residue was adsorbed on silica gel (100 g) and chromatographed (800 g, chloroform-methanol 9:14:1). Selected fractions were concentrated under reduced pressure and the residue was a mixture of 2'-*O*-(2-(methoxyethyl) adenosine and 3'-*O*-(2-methoxyethyl) adenosine in the ratio of 4:1.

The above mixture (0.056 mol) in pyridine (100 mL) was evaporated under reduced pressure to dryness. The residue was redissolved in pyridine (560 mL) and cooled in an ice water bath. Trimethylsilyl chloride (36.4 mL, 0.291 mol) was added and the reaction was stirred at 5 °C for 30 min. Benzoyl chloride (33.6 mL, 0.291 mol) was added and the reaction was allowed to warm to 25 °C for 2 h and then cooled to 5 °C. The reaction was diluted with cold water (112 mL) and after stirring for 15 min, concentrated ammonium hydroxide (112 Ml) was added. After 30 min, the reaction was concentrated under reduced pressure (below 30 °C) followed by coevaporation with toluene (2x100 mL). The residue was dissolved in ethyl acetate-methanol (400 mL, 9:1) and the undesired silyl by-products were removed by filtration. The filtrate was concentrated under reduced pressure and then chromatographed on silica gel (800 g, chloroform-methanol 9:1). Selected

fractions were combined, concentrated under reduced pressure and dried at 25 °C/0.2 mmHg for 2 h to give pure N^6-Benzoyl-2'-*O*-(2-methoxyethyl) adenosine and pure N^6-Benzoyl-3'-*O*-(2-methoxyethyl) adenosine.

A solution of N^6-Benzoyl-3'-*O*-(2-methoxyethyl) adenosine (11.0 g, 0.285 mol) in pyridine (100 mL) was evaporated under reduced pressure to an oil. The residue was redissolved in dry pyridine (300 mL) and DMT-Cl (10.9 g, 95%, 0.31 mol) was added. The mixture was stirred at 25 °C for 16 h and then poured onto a solution of sodium bicarbonate (20 g) in ice water (500 mL). The product was extracted with ethyl acetate (2x150 mL). The organic layer was washed with brine (50 mL), dried over sodium sulfate (powdered) and evaporated under reduced pressure (below 40C). The residue was chromatographed on silica gel (400 g, ethyl acetate-acetonitrile-triethylamine 99:1:195:5:1). Selected fractions were combined, concentrated under reduced pressure and dried at 25 °C/0.2 mmHg to give 16.8 g (73%) of the title compound as a foam. The TLC was homogenous.

EXAMPLE 6

[N^6-Benzoyl-5'-O-(DMT)-3'-O-(2-methoxyethyl) adenosin-2'-O-(2-cyanoethyl-N,N-diisopropyl) phosphoramidite

The title compound was prepared in the same manner as the 5-methyl-cytidine and 5-methyluridine analogs of Examples 2 and 4 by starting with the title compound of Example 5. Purification using ethyl acetate-hexanes-triethylamine 59:40:1 as the chromatography eluent gave 67% yield of the title compound as a solid foam. The TLC was homogenous. ^{31}P-NMR ($CDCl_3$, H_3PO_4 std.) δ 147.89; 148.36 (diastereomers).

EXAMPLE 7

5'-O-(DMT)-N^2-isobutyryl-3'-O-(2-methoxyethyl) guanosine

A. 2,6-Diaminopurine riboside

To a 2 L stainless steel Parr bomb was added guanosine hydrate (100 g, 0.35 mol, Aldrich), hexamethyl) disilazane (320 mL, 1.52 mol, 4.4 eq.), trimethyl) silyl triflouromethanesulfonate (8.2 mL), and toluene (350 mL). The bomb was sealed and partially submerged in an oil bath (170 °C; internal T 150 °C, 150 psi) for 5 days. The bomb was cooled in a dry ice/acetone bath and opened. The contents were transferred with methanol (300 mL) to a flask and the solvent was evaporated under reduced pressure. Aqueous methanol (50%, 1.2 L) was added. The resulting brown suspension was heated to reflux for 5 h. The suspension was concentrated under reduced pressure to one half volume in order to remove most of the methanol. Water (600 mL) was added and the solution was heated to reflux, treated with charcoal (5 g) and hot filtered through Celite. The solution was allowed to cool to 25 °C. The resulting precipitate was collected, washed with water (200 mL) and dried at 90 °C/0.2 mmHg for 5 h to give a constant weight of 87.4 g (89%) of tan, crystalline solid; mp 247 °C (shrinks), 255 °C (dec, lit. (1) mp 250-252 °C); TLC homogenous (Rf 0.50, isopropanol-ammonium hydroxide-water 16:3:1); PMR (DMSO), δ 5.73 (d, 2, 2-NH_2), 5.78 (s, 1, H-1), 6.83 (br s, 2, 6-NH_2).

B. 2'-O-(2-methoxyethyl)-2,6-diaminopurine riboside and 3'-*O*-(2-methoxyethyl)-2,6-diaminopurine riboside

To a solution of 2,6-diaminopurine riboside (10.0 g, 0.035 mol) in dry dimethyl formamide (350 mL) at 0 °C under an argon atmosphere was added sodium hydride (60% in oil, 1.6 g, 0.04 mol). After 30 min., 2-methoxyethyl bromide (0.44 mol) was added in one portion and the reaction was stirred at 25 °C for 16 h. Methanol (10 mL) was added and the mixture was concentrated under reduced pressure to an oil (20 g). The

crude product, containing a ratio of 4:1 of the 2'/3' isomers, was chromatographed on silica gel (500 g, chloroform-methanol 4:1). The appropriate fractions were combined and concentrated under reduced pressure to a semi-solid (12 g). This was triturated with methanol (50 mL) to give a white, hygroscopic solid. The solid was dried at 40 °C/0.2 mmHg for 6 h to give a pure 2' product and the pure 3' isomer, which were confirmed by NMR.

C. 3'-O-2-(methoxyethyl)guanosine

With rapid stirring, 3'-O-(2-methoxyethyl)-2,6-diaminopurine riboside (0.078 mol) was dissolved in monobasic sodium phosphate buffer (0.1 M, 525 mL, pH 7.3-7.4) at 25 °C. Adenosine deaminase (Sigma type II, 1 unit/mg, 350 mg) was added and the reaction was stirred at 25 °C for 60 h. The mixture was cooled to 5 °C and filtered. The solid was washed with water (2x25 mL) and dried at 60 °C/0.2 mmHg for 5 h to give 10.7 g of first crop material. A second crop was obtained by concentrating the mother liquors under reduced pressure to 125 mL, cooling to 5 °C, collecting the solid, washing with cold water (2x20 mL) and drying as above to give 6.7 g of additional material for a total of 15.4 g (31 % from guanosine hydrate) of light tan solid; TLC purity 97%.

D. N^2-Isobutyryl-3'-O-2-(methoxyethyl)guanosine

To a solution of 3'-O-2-(methoxyethyl)guanosine (18.1 g, 0.0613 mol) in pyridine (300 mL) was added trimethyl silyl chloride (50.4 mL, 0.46 mol). The reaction was stirred at 25 °C for 16 h. Isobutyryl chloride (33.2 mL, 0.316 mol) was added and the reaction was stirred for 4 h at 25 °C. The reaction was diluted with water (25 mL). After stirring for 30 min, ammonium hydroxide (concentrated, 45 mL) was added until pH 6 was reached. The mixture was stirred in a water bath for 30 min and then evaporated under reduced pressure to an oil. The oil was suspended in a mixture of ethyl acetate (600 mL) and water (100 mL) until a solution formed. The solution was allowed to stand for 17 h at 25 °C. The

resulting precipitate was collected, washed with ethyl acetate (2x50 mL) and dried at 60 °C/0.2 mmHg for 5 h to give 16.1 g (85%) of tan solid; TLC purity 98%.

E. 5'-O-(DMT)-N^2-isobutyryl-3'-O-(2-methoxyethyl) guanosine

A solution of N^2-Isobutyryl-3'-O-2-(methoxyethyl) guanosine (0.051 mol) in pyridine (150 mL) was evaporated under reduced pressure to dryness. The residue was redissolved in pyridine (300 mL) and cooled to 10-15 °C. DMT-Cl (27.2 g, 95%, 0.080 mol) was added and the reaction was stirred at 25 °C for 16 h. The reaction was evaporated under reduced pressure to an oil, dissolved in a minimum of methylene chloride and applied on a silica gel column (500 g). The product was eluted with a gradient of methylene chloride-triethylamine (99:1) to methylene chloride-methanol-triethylamine (99:1:1). Selected fractions were combined, concentrated under reduced pressure and dried at 40 °C/0.2 mmHg for 2 h to afford 15 g (15.5% from guanosine hydrate) of tan foam; TLC purity 98%.

EXAMPLE 8

[5'-O-(DMT)-N^2-isobutyryl-3'-O-(2-methoxyethyl) guanosin-2'-O-(2-cyanoethyl-N,N-diisopropyl) phosphoramidite

The protected nucleoside from Example 7 (0.0486 mol) was placed in a dry 1 L round bottom flask containing a Teflon stir-bar. The flask was purged with argon. Anhydrous methylene chloride (400 mL) was cannulated into the flask to dissolve the nucleoside. Previously vacuum dried N,N-diisopropylaminohydrotetrazolide (3.0 g, 0.0174 mol) was added under argon. With stirring, bis-N,N-diisopropyl-aminocyanoethylphosphoramidite (18.8 g, 0.0689 mol) was added via syringe over 1 min (no exotherm noted). The reaction was stirred under argon at 25 °C for 16 h. After verifying the completion of the reaction by TLC, the reaction was

transferred to a separatory funnel (1 L). The reaction flask was rinsed with methylene chloride (2x50 mL). The combined organic layer was washed with saturated aq. sodium bicarbonate (200 mL) and then brine (200 mL). The organic layer was dried over sodium sulfate (50 g, powdered) for 2 h. The solution was filtered and concentrated under reduced pressure to a viscous oil. The resulting phosphoramidite was purified by silica gel flash chromatography (800 g, ethyl acetate-triethylamine 99:1). Selected fractions were combined, concentrated under reduced pressure, and dried at 25C/0.2 mmHg for 16 h to give 18.0 g (46%, 3% from guanosine hydrate) of solid foam TLC homogenous. ^{31}P-NMR ($CDCl_3$, H_3PO_4 std.) δ 147.96; 148.20 (diastereomers).

EXAMPLE 9

5'-*O*-DMT-3'-O-(2-methoxyethyl)-5-methyl-uridine-2'-O-succinate

5'-*O*-DMT-3'-*O*-(2-methoxyethyl)-thymidine was first succinylated on the 2'-position. Thymidine nucleoside (4 mmol) was reacted with 10.2 mL dichloroethane, 615 mg (6.14 mmol) succinic anhydride, 570 µL (4.09 mmol) triethylamine, and 251 mg (2.05 mmol) 4-dimethylaminopyridine. The reactants were vortexed until dissolved and placed in heating block at 55 °C for approximately 30 minutes. Completeness of reaction checked by thin layer chromatography (TLC). The reaction mixture was washed three times with cold 10% citric acid followed by three washes with water. The organic phase was removed and dried under sodium sulfate. Succinylated nucleoside was dried under P_2O_5 overnight in vacuum oven.

EXAMPLE 10

5'-*O*-DMT-3'-O-methoxyethyl-5-methyl-uridine-2'-O-succinoyl Linked LCA CPG

5'-*O*-DMT-3'-*O*-(2-methoxyethyl)-2'-*O*-succinyl-thymidine was coupled to controlled pore glass (CPG). 1.09 g (1.52 mmol) of the succinate were dried overnight in a vacuum oven along with 4-dimethylaminopyridine (DMAP), 2,2'-dithiobis (5-nitro-pyridine) (dTNP), triphenylphosphine (TPP), and pre-acid washed CPG (controlled pore glass). After about 24 hours, DMAP (1.52 mmol, 186 mg) and acetonitrile (13.7 mL) were added to the succinate. The mixture was stirred under an atmosphere of argon using a magnetic stirrer. In a separate flask, dTNP (1.52 mmol, 472 mg) was dissolved in acetonitrile (9.6 mL) and dichloromethane (4.1 mL) under argon. This reaction mixture was then added to the succinate. In another separate flask, TPP (1.52 mmol, 399 mg) was dissolved in acetonitrile (37 mL) under argon. This mixture was then added to the succinate/DMAP/dTNP reaction mixture. Finally, 12.23 g pre-acid washed LCA CPG (loading = 115.2 µmol/g) was added to the main reaction mixture, vortexed shortly and placed on shaker for approximately 3 hours. The mixture was removed from the shaker after 3 hours and the loading was checked. A small sample of CPG was washed with copious amounts of acetonitrile, dichloromethane, and then with ether. The initial loading was found to be 63 µmol/g (3.9 mg of CPG was cleaved with trichloroacetic acid, the absorption of released trityl cation was read at 503 nm on a spectrophotometer to determine the loading.) The whole CPG sample was then washed as described above and dried under P_2O_5 overnight in vacuum oven. The following day, the CPG was capped with 25 mL CAP A (tetrahydrofuran/acetic anhydride) and 25 mL CAP B (tetrahydrofuran/pyridine/1-methyl imidazole) for approximately 3 hours on shaker. Filtered and washed with dichloromethane and ether. The CPG was dried under P_2O_5

overnight in vacuum oven. After drying, 12.25 g of CPG was isolated with a final loading of 90 μmol/g.

EXAMPLE 11

3'-O-Methoxyethyl-5-methyl-N-benzoyl-cytidine-2'-O-succinate

5'-*O*-DMT-3'-*O*-(2-methoxy) ethyl-N-benzoyl-cytidine was first succinylated on the 2'-position. Cytidine nucleoside (4 mmol) was reacted with 10.2 mL dichloroethane, 615 mg (6.14 mmol) succinic anhydride, 570 μL (4.09 mmol) triethylamine, and 251 mg (2.05 mmol) 4-dimethylaminopyridine. The reactants were vortexed until dissolved and placed in a heating block at 55 °C for approximately 30 minutes. Completeness of reaction was checked by thin layer chromatography (TLC). The reaction mixture was washed three times with cold 10% citric acid followed by three washes with water. The organic phase was removed and dried under sodium sulfate. The succinylated nucleoside was dried under P_2O_5 overnight in vacuum oven.

EXAMPLE 12

5'-*O*-DMT-3'-O-methoxyethyl-5-methyl-N-benzoyl-cytidine-2'-O-succinoyl linked LCA CPG

5'-*O*-DMT-3'-*O*-(2-methoxyethyl)-2'-*O*-succinyl-N^4-benzoyl cytidine was coupled to controlled pore glass (CPG). 1.05 g (1.30 mmol) of the succinate were dried overnight in a vacuum oven along with 4-dimethylaminopyridine (DMAP), 2,2'-dithiobis (5-nitro-pyridine) (dTNP), triphenylphosphine (TPP), and pre-acid washed CPG (controlled pore glass). The following day, DMAP (1.30 mmol, 159 mg) and acetonitrile (11.7 mL) were added to the succinate. The mixture was "mixed" by a magnetic stirrer under argon. In a separate flask, dTNP (1.30 mmol, 400 mg) was dissolved in acetonitrile (8.2 mL) and dichloromethane (3.5 mL) under argon. This reaction mixture was then added to the succinate. In another separate flask, TPP (1.30 mmol, 338 mg) was dissolved in acetonitrile (11.7

mL) under argon. This mixture was then added to the succinate/DMAP/dTNP reaction mixture. Finally, 10.46 g pre-acid washed LCA CPG (loading = 115.2 µmol/g) were added to the main reaction mixture, vortexed shortly and placed on shaker for approximately 2 hours. A portion was removed from shaker after 2 hours and the loading was checked. A small sample of CPG was washed with copious amounts of acetonitrile, dichloromethane, and then with ether. The initial loading was found to be 46 µmol/g. (3.4 mg of CPG were cleaved with trichloroacetic acid). The absorption of released trityl cation was read at 503 nm on a spectrophotometer to determine the loading. The whole CPG sample was then washed as described above and dried under P_2O_5 overnight in vacuum oven. The following day, the CPG was capped with 25 mL CAP A (tetrahydrofuran/acetic anhydride) and 25 mL CAP B (tetrahydrofuran/pyridine/1-methyl imidazole) for approximately 3 hours on a shaker. The material was filtered and washed with dichloromethane and ether. The CPG was dried under P_2O_5 overnight in vacuum oven. After drying, 10.77 g of CPG was isolated with a final loading of 63 µmol/g.

EXAMPLE 13

5'-*O*-DMT-3'-O-methoxyethyl-N6-benzoyl-adenosine-2'-O-succinate

5'-*O*-DMT-3'-*O*-(2-methoxyethyl)-N^6-benzoyl adenosine was first succinylated on the 2'-position. 3.0 g (4.09 mmol) of the adenosine nucleoside were reacted with 10.2 mL dichloroethane, 615 mg (6.14 mmol) succinic anhydride, 570 µL (4.09 mmol) triethylamine, and 251 mg (2.05 mmol) 4-dimethylaminopyridine. The reactants were vortexed until dissolved and placed in heating block at 55 °C for approximately 30 minutes. Completeness of reaction was checked by thin layer chromatography (TLC). The reaction mixture was washed three times with cold 10% citric acid followed by three washes with water. The organic phase was

removed and dried under sodium sulfate. Succinylated nucleoside was dried under P_2O_5 overnight in vacuum oven.

EXAMPLE 14

5'-*O*-DMT- 3'-O-(2-methoxyethyl)-N6-benzoyl-adenosine-2'-O-succinoyl Linked LCA CPG

Following succinylation, 5'-*O*-DMT-3'-*O*-(2-methoxyethyl)-2'-*O*-succinyl-N^6-benzoyl adenosine was coupled to controlled pore glass (CPG). 3.41 g (4.10 mmol) of the succinate were dried overnight in a vacuum oven along with 4-dimethylaminopyridine (DMAP), 2,2'-dithiobis (5-nitropyridine) (dTNP), triphenylphosphine (TPP), and pre-acid washed CPG (controlled pore glass). The following day, DMAP (4.10 mmol, 501 mg) and acetonitrile (37 mL) were added to the succinate. The mixture was "mixed" by a magnetic stirrer under argon. In a separate flask, dTNP (4.10 mmol, 1.27g) was dissolved in acetonitrile (26 mL) and dichloromethane (11 mL) under argon. This reaction mixture was then added to the succinate. In another separate flask, TPP (4.10 mmol, 1.08 g) was dissolved in acetonitrile (37 mL) under argon. This mixture was then added to the succinate/DMAP/dTNP reaction mixture. Finally, 33 g pre-acid washed LCA CPG (loading = 115.2 µmol/g) were added to the main reaction mixture, vortexed shortly and placed on shaker for approximately 20 hours. Removed from shaker after 20 hours and the loading was checked. A small sample of CPG was washed with copious amounts of acetonitrile, dichloromethane, and then with ether. The initial loading was found to be 49 µmol/g. (2.9 mg of CPG were cleaved with trichloroacetic acid). The absorption of released trityl cation was read at 503 nm on a spectrophotometer to determine the loading. The whole CPG sample was then washed as described above and dried under P_2O_5 overnight in vacuum oven. The following day, the CPG was capped with 50 mL CAP A (tetrahydrofuran/acetic anhydride) and

50 mL CAP B (tetrahydrofuran/pyridine/1-methyl imidazole) for approximately 1 hour on the shaker. The material was filtered and washed with dichloromethane and ether. The CPG was dried under P_2O_5 overnight in vacuum oven. After drying, 33.00 g of CPG was obtained with a final loading of 66 µmol/g.

EXAMPLE 15

5'-*O*-DMT-3'-O-(2-methoxyethyl)-N2-isobutyryl-guanosine-2'-O-succinate

5'-*O*-DMT-3'-*O*-(2-methoxyethy)l-N^2-isobutyryl guanosine was succinylated on the 2'-sugar position. 3.0 g (4.20 mmol) of the guanosine nucleoside were reacted with 10.5 mL dichloroethane, 631 mg (6.30 mmol) succinic anhydride, 585 µL (4.20 mmol) triethylamine, and 257 mg (2.10 mmol) 4-dimethylaminopyridine. The reactants were vortexed until dissolved and placed in heating block at 55 °C for approximately 30 minutes. Completeness of reaction checked by thin layer chromatography (TLC). The reaction mixture was washed three times with cold 10% citric acid followed by three washes with water. The organic phase was removed and dried under sodium sulfate. The succinylated nucleoside was dried under P_2O_5 overnight in vacuum oven.

EXAMPLE 16

5'-*O*-DMT-3'-O-methoxyethyl-N2-isobutyryl-guanosine-2'-O-succinoyl Linked LCA CPG

Following succinylation, 5'-*O*-DMT-3'-*O*-(2-methoxy-ethyl)-2'-*O*-succinyl-N^2-benzoyl guanosine was coupled to controlled pore glass (CPG). 3.42 g (4.20 mmol) of the succinate were dried overnight in a vacuum oven along with 4-dimethylaminopyridine (DMAP), 2,2'-dithiobis (5-nitro-pyridine) (dTNP), triphenylphosphine (TPP), and pre-acid washed CPG (controlled pore glass). The following day, DMAP (4.20 mmol, 513 mg) and acetonitrile (37.5 mL) were added to

the succinate. The mixture was "mixed" by a magnetic stirrer under argon. In a separate flask, dTNP (4.20 mmol, 1.43 g) was dissolved in acetonitrile (26 mL) and dichloromethane (11 mL) under argon. This reaction mixture was then added to the succinate. In another separate flask, TPP (4.20 mmol, 1.10 g) was dissolved in acetonitrile (37. 5 mL) under argon. This mixture was then added to the succinate/DMAP/dTNP reaction mixture. Finally, 33.75 g pre-acid washed LCA CPG (loading = 115.2 µmol/g) were added to the main reaction mixture, vortexed shortly and placed on shaker for approximately 20 hours. Removed from shaker after 20 hours and the loading was checked. A small sample of CPG was washed with copious amounts of acetonitrile, dichloromethane, and then with ether. The initial loading was found to be 64 µmol/g. (3.4 mg of CPG were cleaved with trichloroacetic acid). The absorption of released trityl cation was read at 503 nm on a spectrophotometer to determine the loading. The CPG was then washed as described above and dried under P_2O_5 overnight in vacuum oven. The following day, the CPG was capped with 50 mL CAP A (tetrahydrofuran/acetic anhydride) and 50 mL CAP B (tetrahydrofuran/pyridine/1-methyl imidazole) for approximately 1 hour on a shaker. The material was filtered and washed with dichloromethane and ether. The CPG was dried under P_2O_5 overnight in vacuum oven. After drying, 33.75 g. of CPG was isolated with a final loading of 72 µmol/g.

EXAMPLE 17

5'-*O*-DMT-3'-O-[hexyl-(6-phthalimido)]-uridine

2',3'-*O*-Dibutyl stannylene-uridine was synthesized according to the procedure of Wagner *et. al., J. Org. Chem.*, **1974**, *39*, 24. This compound was dried over P_2O_5 under vacuum for 12 hours. To a solution of this compound (29 g, 42.1 mmol) in 200 mL of anhydrous DMF were added (16.8 g, 55 mmol) of 6-bromohexyl phthalimide and 4.5 g of sodium iodide and the

mixture was heated at 130 °C for 16 hours under argon. The reaction mixture was evaporated, co-evaporated once with toluene and the gummy tar residue was applied on a silica column (500 g). The column was washed with 2 L of EtOAc followed by eluting with 10% methanol (MeOH):90% EtOAc. The product, 2'- and 3'-isomers of O-hexyl-6-N-phthalimido uridine, eluted as an inseparable mixture (R_f=0.64 in 10% MeOH in EtOAc). By ^{13}C NMR, the isomeric ration was about 55% of the 2' isomer and about 45% of the 3' isomer. The combined yield was 9.2 g (46.2%). This mixture was dried under vacuum and re-evaporated twice with pyridine. It was dissolved in 150 mL anhydrous pyridine and treated with 7.5 g of DMT-Cl (22.13 mmol) and 500 mg of dimethylaminopyridine (DMAP). After 2 hours, thin layer chromatography (TLC; 6:4 EtOAc:Hexane) indicated complete disappearance of the starting material and a good separation between 2' and 3' isomers (R_f=0.29 for the 2' isomer and 0.12 for the 3' isomer). The reaction mixture was quenched by the addition of 5 mL of CH_3OH and evaporated under reduced pressure. The residue was dissolved in 300 mL CH_2Cl_2, washed successively with saturated $NaHCO_3$ followed by saturated NaCl solution. It was dried over Mg_2SO_4 and evaporated to give 15 g of a brown foam which was purified on a silica gel (500 g) to give 6.5 g of the 2'-isomer and 3.5 g of the 3' isomer.

EXAMPLE 18

5'-*O*-DMT-3'-O-[hexyl-(6-phthalimido)]-uridine-2'-O-(2-cyanoethyl-N,N,-diisopropyl) phosphoramidite

5'-DMT-3'-*O*-[hexyl-(6-phthalimido)]uridine (2 g, 2.6 mmol) was dissolved in 20 mL anhydrous CH_2Cl_2. To this solution diisopropylaminotetrazolide (0.2 g, 1.16 mmol) and 2.0 mL 2-cyanoethyl-N,N,N',N'-tetraisopropyl phosphoramidite (6.3 mmol) were added with stirred overnight. TLC (1:1 EtOAc/hexane) showed complete disappearance of starting

material. The reaction mixture was transferred with CH_2Cl_2 and washed with saturated $NaHCO_3$ (100 mL), followed by saturated NaCl solution. The organic layer was dried over anhydrous Na_2SO_4 and evaporated to yield 3.8 g of a crude product, which was purified in a silica column (200 g) using 1:1 hexane/EtOAc to give 1.9 g (1.95 mmol, 74% yield) of the desired phosphoramidite.

EXAMPLE 19

Preparation of 5'-*O*-DMT-3'-O-[hexyl-(6-phthalimido)]-uridine-2'-O-succinoyl-aminopropyl CPG

Succinylated and capped aminopropyl controlled pore glass (CPG; 500Å pore diameter, aminopropyl CPG, 1.0 grams prepared according to Damha *et. al., Nucl. Acids Res.* **1990**, *18*, 3813.) was added to 12 mL anhydrous pyridine in a 100 mL round-bottom flask. 1-(3-Dimethylaminopropyl)-3-ethyl-carbodiimide (DEC; 0.38 grams, 2.0 mmol)], triethylamine (TEA; 100 µl, distilled over CaH_2), dimethylaminopyridine (DMAP; 0.012 grams, 0.1 mmol) and nucleoside 5'-O-DMT-3'-O-[hexyl-(6-phthalimido)]uridine (0.6 grams, 0.77 mmol) were added under argon and the mixture shaken mechanically for 2 hours. Additional nucleoside (0.20 grams) was added and the mixture shaken for 24 hours. The CPG was filtered off and washed successively with dichloromethane, triethylamine, and dichloromethane. The CPG was then dried under vacuum, suspended in 10 mL piperidine and shaken 15 minutes. The CPG was filtered off, washed thoroughly with dichloromethane and again dried under vacuum. The extent of loading (determined by spectrophotometric assay of DMT cation in 0.3 M p-toluenesulfonic acid at 498 nm) was approximately 28 µmol/g. The 5'-O-(DMT)-3'-O-[hexyl-(6-phthalimido] uridine-2'-O-succinyl-aminopropyl controlled pore glass was used to synthesize the oligomers in an ABI 380B DNA synthesizer using phosphoramidite chemistry standard conditions. A four base

oligomer 5'-GACU*-3' was used to confirm the structure of 3'-O-hexylamine tether introduced into the oligonucleotide by NMR. As expected a multiplet signal was observed between 1.0-1.8 ppm in ^{1}H NMR.

EXAMPLE 20

5'-*O*-DMT-3'-O-[hexylamino]-uridine

5'-O-(DMT)-3'-O-[hexyl-(6-phthalimido)] uridine (4.5 grams, 5.8 mmol) is dissolved in 200 mL methanol in a 500 mL flask. Hydrazine (1 ml, 31 mmol) is added to the stirring reaction mixture. The mixture is heated to 60-65 °C in an oil bath and refluxed 14 hours. The solvent is evaporated *in vacuo* and the residue is dissolved in dichloromethane (250 mL) and extracted twice with an equal volume NH_4OH. The organic layer is evaporated to yield the crude product which NMR indicates is not completely pure. R_f=0 in 100% ethyl acetate. The product is used in subsequent reactions without further purification.

EXAMPLE 21

3'-O-[Propyl-(3-phthalimido)]-adenosine

To a solution of adenosine (20.0 g, 75 mmol) in dry dimethylformamide (550 ml) at 5 °C was added sodium hydride (60% oil, 4.5 g, 112 mmol). After one hour, N-(3-bromopropyl)phthalimide (23.6 g, 86 mmol) was added and the temperature was raised to 30 °C and held for 16 hours. Ice is added and the solution evaporated *in vacuo* to a gum. The gum was partitioned between water and ethyl acetate (4 x 300 mL). The organic phase was separated, dried, and evaporated *in vacuo* and the resultant gum chromatographed on silica gel (95/5 CH_2Cl_2/MeOH) to give a white solid (5.7 g) of the 2'-*O*-(propylphthalimide)adenosine. Thee fractions containing the 3'-*O*-(propylphthalimide)adenosine were chromatographed a second time on silica gel using the same solvent system.

Crystallization of the 2'-*O*-(propylphthalimide)-adenosine fractions from methanol gave a crystalline solid, m.p. 123-124C. ^{1}H NMR (400 MHZ: DMSO-d_6) δ 1.70(m, 2H, CH_2), 3.4-3.7 (m, 6H, $C_{5'}$, CH_2, OCH_2, Phth CH_2), 3.95 (q, 1H, $C_{4'}$H), 4.30 (q, 1H, $C_{5'}$H), 4.46 (t, 1H, $C_{2'}$H), 5.15 (d, 1H, $C_{3'}$OH), 5.41 (t, 1H, $C_{5'}$OH), 5.95 (d, 1H, $C_{1'}$H) 7.35 (s, 2H, NH_2), 7.8 (brs, 4H, Ar), 8.08 (s, 1H, C_2H) and 8.37 (s, 1H, C_8H). Anal. Calcd. $C_{21}H_{22}N_6O_6$: C, 55.03; H, 4.88; N, 18.49. Found: C, 55.38; H, 4.85; N, 18.46.

Crystallization of the 3'-*O*-(propylphthalimide)-adenosine fractions from H_2O afforded an analytical sample, m.p. 178-179C. ^{1}H NMR (400 MHZ: DMSO-d_6) δ 5.86 (d, 1H, H-1').

EXAMPLE 22

3'-O-[Propyl-(3-phthalimido)]-N6-benzoyl-adenosine

3'-*O*-(3-propylphthalimide)adenosine is treated with benzoyl chloride in a manner similar to the procedure of Gaffney, *et al.*, *Tetrahedron Lett.* **1982**, *23*, 2257. Purification of crude material by chromatography on silica gel (ethyl acetate-methanol) gives the title compound.

EXAMPLE 23

3'-O-[Propyl-(3-phthalimido)]-5'-O-DMT-N6-benzoyl-adenosine

To a solution of 3'-O-(propyl-3-phthalimide)-N^6-benzoyladenosine (4.0 g) in pyridine (250 ml) is added DMT-Cl (3.3 g). The reaction is stirred for 16 hours. The reaction is added to ice/water/ethyl acetate, the organic layer separated, dried, and concentrated *in vacuo* and the resultant gum chromatographed on silica gel (ethyl acetate-methanol triethylamine) to give the title compound.

EXAMPLE 24

3'-O-[Propyl-(3-phthalimido)]-5'-O-DMT-N6-Benzoyl-adenosine-2'-O-(2-cyanoethyl-N,N-diisopropyl) phosphoramidite

3'-*O*-(Propyl-3-phthalimide)-5'-*O*-DMT-N^6-benzoyladenosine is treated with (β-cyanoethoxy)chloro-N,N-diisopropyl)aminophosphane in a manner similar to the procedure of Seela, *et al.*, *Biochemistry* **1987**, *26*, 2233. Chromatography on silica gel (EtOAc/hexane) gives the title compound as a white foam.

EXAMPLE 25

3'-O-(Aminopropyl)-adenosine

A solution of 3'-*O*-(propyl-3-phthalimide)adenosine (8.8 g, 19 mmol), 95% ethanol (400 mL) and hydrazine (10 mL, 32 mmol) is stirred for 16 hrs at room temperature. The reaction mixture is filtered and filtrate concentrated *in vacuo*. Water (150 mL) is added and acidified with acetic acid to pH 5.0. The aqueous solution is extracted with EtOAc (2 x 30 mL) and the aqueous phase is concentrated *in vacuo* to afford the title compound as a HOAc salt.

EXAMPLE 26

3'-O-[3-(N-trifluoroacetamido)propyl]-adenosine

A solution of 3'-*O*-(propylamino)adenosine in methanol (50 mL) and triethylamine (15 mL, 108 mmol) is treated with ethyl trifluoroacetate (18 mL, 151 mmol). The reaction is stirred for 16 hrs and then concentrated *in vacuo* and the resultant gum chromatographed on silica gel (9/1, EtOAc/MeOH) to give the title compound.

EXAMPLE 27

N6-Dibenzoyl-3'-O-[3-(N-trifluoroacetamido)propyl]-adenosine

3'-*O*-[3-(N-trifluoroacetamido)propyl]adenosine is treated as per Example 22 using a Jones modification wherein tetrabutylammonium fluoride is utilized in place of ammonia hydroxide in the work up. The crude product is purified using silica gel chromatography (EtOAc/MeOH 1/1) to give the title compound.

EXAMPLE 28

N6-Dibenzoyl-5'-O-DMT-3'-O-[3-(N-trifluoroacetamido)propyl]-adenosine

DMT-Cl (3.6 g, 10.0 mmol) is added to a solution of N^6-(dibenzoyl)-3'-*O*-[3-(N-trifluoroacetamido)propyl)adenosine in pyridine (100 mL) at room temperature and stirred for 16 hrs. The solution is concentrated *in vacuo* and chromatographed on silica gel (EtOAc/TEA 99/1) to give the title compound.

EXAMPLE 29

N6-Dibenzoyl-5'-O-DMT-3'-O-[3-(N-trifluoroacetamido)propyl]-adenosine-2'-O-(2-cyanoethyl-N,N-diisopropyl) phosphoramidite

A solution of N^6-(dibenzoyl)-5'-*O*-(DMT)-3'-*O*-[3-(N-trifluoroacetamido)propyl]adenosine in dry CH_2Cl_2 is treated with bis-N,N-diisopropylamino cyanoethyl phosphite (1.1 eqiv) and N,N-diisopropylaminotetrazolide (catalytic amount) at room temperature for 16 hrs. The reaction is concentrated *in vacuo* and chromatographed on silica gel (EtOAc/hexane/TEA 6/4/1) to give the title compound.

EXAMPLE 30

3'-O-(butylphthalimido)-adenosine

The title compound is prepared as per Example 21, using N-(4-bromobutyl)phthalimide in place of the 1-bromopropane. Chromatography on silica gel (EtOAC-MeOH) gives

the title compound. ^{1}H NMR (200 MHZ, DMSO-d_6) δ 5.88 (d, 1H, C_1'H).

EXAMPLE 31

N6-Benzoyl-3'-O-(butylphthalimido)-adenosine

Benzoylation of 3'-*O*-(butylphthalimide)adenosine as per Example 22 gives the title compound.

EXAMPLE 32

N6-Benzoyl-5'-O-DMT-3'-O-(butylphthalimido)-adenosine

The title compound is prepared from 3'-*O*-(butylphthalimide)-N^6-benzoyladenosine as per Example 22.

EXAMPLE 33

N6-Benzoyl-5'-O-DMT-3'-O-(butylphthalimido)-Adenosine-2'-O-(2-cyanoethyl-N,N-diisopropyl) phosphoramidite

The title compound is prepared from 3'-*O*-(butylphthalimide)-5'-*O*-DMT-N^6-benzoyladenosine as per Example 24.

EXAMPLE 34

3'-O-(Pentylphthalimido)-adenosine

The title compound is prepared as per Example 21, using N-(5-bromopentyl)phthalimide. The crude material from the extraction is chromatographed on silica gel using $CHCl_3$/MeOH (95/5) to give a mixture of the 2' and 3' isomers. The 2' isomer is recrystallized from EtOH/MeOH 8/2. The mother liquor is rechromatographed on silica gel to afford the 3' isomer.

2'-*O*-(Pentylphthalimido)adenosine: M.P. 159-160 °C. Anal. Calcd. for $C_{23}H_{24}N_6O_5$: C, 57.26; H, 5.43; N, 17.42. Found: C, 57.03; H, 5.46; N, 17.33. 3'-*O*-(Pentylphthalimido)adenosine: ^{1}H NMR (DMSO-d_6) δ 5.87 (d, 1H, H-1').

EXAMPLE 35

N6-Benzoyl-3'-O-(pentylphthalimido)-adenosine

Benzoylation of 3'-*O*-(pentylphthalimido)adenosine is achieved as per the procedure of Example 22 to give the title compound.

EXAMPLE 36

N6-Benzoyl-5'-O-DMT-3'-O-(pentylphthalimido)-adenosine

The title compound is prepared from 3'-*O*-(pentylphthalimide)-N^6-benzoyladenosine as per the procedure of Example 23. Chromatography on silica gel (ethylacetate, hexane, triethylamine), gives the title compound.

EXAMPLE 37

N6-Benzoyl-5'-O-DMT-3'-O-(pentylphthalimido)-adenosine-2'-O-(2-cyanoethyl-N,N-diisopropyl) phosphoramidite

The title compound is prepared from 3'-*O*-(pentylphthalimide)-5'-*O*-(DMT)-N^6-benzoyladenosine as per the procedure of Example 24 to give the title compound.

EXAMPLE 38

3'-O-(Propylphthalimido)uridine

A solution of uridine-tin complex (48.2 g, 115 mmol) in dry DMF (150 ml) and N-(3-bromopropyl)phthalimide (46 g, 172 mmol) was heated at 130 °C for 6 hrs. The crude product was chromatographed directly on silica gel $CHCl_3$/MeOH 95/5. The isomer ratio of the purified mixture was 2'/3' 81/19. The 2' isomer was recovered by crystallization from MeOH. The filtrate was rechromatographed on silica gel using $CHCl_3CHCl_3$/MeOH (95/5) gave the 3' isomer as a foam.

2'-*O*-(Propylphthalimide)uridine: Analytical sample recrystallized from MeOH, m.p. 165.5-166.5C, ^{1}H NMR (200 MHZ, DMSO-d_6) δ 1.87 (m, 2H, CH_2), 3.49-3.65 (m, 4H, C_2'H), 3.80-3.90 (m, 2H, C_3'H_1C_4'H), 4.09(m, 1H, C_2'H), 5.07 (d, 1h, C_3'OH),

5.16 (m, 1H, C_5'OH), 5.64 (d, 1H, CH=), 7.84 (d, 1H, C_1'H), 7.92 (bs, 4H, Ar), 7.95 (d, 1H, CH=) and 11.33 (s, 1H, ArNH). Anal. $C_{20}H_{21}N_3H_8$, Calcd. C, 55.69; H, 4.91;, N, 9.74. Found, C, 55.75; H, 5.12; N, 10.01. 3'-*O*-(Propylphthalimide)uridine: ^{1}H NMR (DMSO-d_6) δ 5.74 (d, 1H, H-1').

EXAMPLE 39

3'-O-(Aminopropyl)-uridine

The title compound is prepared as per the procedure of Example 25.

EXAMPLE 40

3'-O-[3-(N-trifluoroacetamido)propyl]-uridine

3'-*O*-(Propylamino)uridine is treated as per the procedure of Example 26 to give the title compound.

EXAMPLE 41

5'-O-DMT-3'-O-[3-(N-trifluoroacetamido)propyl]-uridine

3'-*O*-[3-(N-trifluoroacetamido)propyl]uridine is treated as per the procedure of Example 28 to give the title compound.

EXAMPLE 42

5'-O-DMT-3'-O-[3-(N-trifluoroacetamido)propyl]-uridine-2'-O-(2-cyanoethyl-N,N-diisopropyl) phosphoramidite

5'-*O*-(DMT)-3'-*O*-[3-(N-trifluoroacetamido)-propyl]uridine is treated as per the procedure of Example 29 to give the title compound.

EXAMPLE 43

3'-O-(Propylphthalimido)-cytidine

The title compounds were prepared as per the procedure of Example 21.

2'-*O*-(propylphthalimide)cytidine: ^{1}H NMR (200 MHZ, DMSO-d_6) δ 5.82 (d, 1H, C_1'H).

3'-*O*-(propylphthalimide)cytidine: ^{1}H NMR (200 MHZ, DMSO-d_6) δ 5.72 (d, 1H, C_1'H).

EXAMPLE 44

3'-O-(Aminopropyl)-cytidine

3'-*O*-(Propylphthalimide)cytidine is treated as per the procedure of Example 25 to give the title compound.

EXAMPLE 45

3'-O-[3-(N-trifluoroacetamido)propyl]-cytidine

3'-*O*-(Propylamino)cytidine is treated as per the procedure of Example 26 to give the title compound.

EXAMPLE 46

N4-Benzoyl-3'-O-[3-(N-trifluoroacetamido)propyl]-cytidine

3'-*O*-[3-(N-trifluoroacetamido)propyl]cytidine is treated as per the procedure of Example 27 to give the title compound.

EXAMPLE 47

N4-Benzoyl-5'-O-DMT-3'-O-[3-(N-trifluoroacetamido)propyl]-cytidine

N^4-(Benzoyl)-3'-*O*-[3-(N-trifluoroacetamido)propyl]-cytidine is treated as per the procedure of Example 28 to give the title compound.

EXAMPLE 48

N4-Benzoyl-5'-O-DMT-3'-O-[3-(N-trifluoroacetamido)propyl]-cytidine-2'-O-(2-cyanoethyl-N,N-diisopropyl) phosphoramidite

N^4-(Benzoyl)-5'-*O*-(DMT)-3'-*O*-[3-(N-trifluoroacetamido)-propyl]cytidine is treated as per the procedure of Example 29 to give the title compound.

EXAMPLE 49

General procedures for oligonucleotide synthesis

Oligonucleotides were synthesized on a Perseptive Biosystems Expedite 8901 Nucleic Acid Synthesis System. Multiple 1-μmol syntheses were performed for each oligonucleotide. Trityl groups were removed with trichloroacetic acid (975 μL over one minute) followed by an acetonitrile wash. All standard amidites (0.1M) were coupled twice per cycle (total coupling time was approximately 4 minutes). All novel amidites were dissolved in dry acetonitrile (100 mg of amidite/1 mL acetonitrile) to give approximately 0.08-0.1 M solutions. Total coupling time was approximately 6 minutes (105 μL of amidite delivered). 1-H-tetrazole in acetonitrile was used as the activating agent. Excess amidite was washed away with acetonitrile. (1S)-(+)-(10-camphorsulfonyl) oxaziridine (CSO, 1.0 g CSO/8.72 mL dry acetonitrile) was used to oxidize (4 minute wait step) phosphodiester linkages while 3H-1,2-benzodithiole-3-one-1,1-dioxide (Beaucage reagent, 3.4 g Beaucage reagent/200 mL acetonitrile) was used to oxidize (one minute wait step) phosphorothioate linkages. Unreacted functionalities were capped with a 50:50 mixture of tetrahyrdofuran/acetic anhydride and tetrahydrofuran/pyridine/1-methyl imidazole. Trityl yields were followed by the trityl monitor during the duration of the synthesis. The final DMT group was left intact. The oligonucleotides were deprotected in 1 mL 28.0-30% ammonium hydroxide (NH_4OH) for approximately 16 hours at 55 °C. Oligonucleotides were also made on a larger scale (20 μmol/synthesis). Trityl groups were removed with just over 8 mL of trichloroacetic acid. All standard amidites (0.1M) were coupled twice per cycle (13 minute coupling step). All novel amidites were also coupled four times per cycle but the coupling time was increased to approximately 20 minutes (delivering 480 μL of amidite). Oxidation times remained the

same but the delivery of oxidizing agent increased to approximately 1.88 mL per cycle. Oligonucleotides were cleaved and deprotected in 5 mL 28.0-30% NH_4OH at 55 °C, for approximately 16 hours.

Table I

3'-*O*-(2-methoxyethyl) containing 2'-5' linked oligonucleotides.

ISIS #	Sequence (5'-3')[1]	Backbone	Chemistry
17176	ATG-CAT-TCT-GCC-CCC-AAG-G**A***	P=S	3'-*O*-MOE
17177	ATG-CAT-TCT-GCC-CCC-AAG-**G*A***	P=S	3'-*O*-MOE
17178	ATG-CAT-TCT-GCC-CCC-AAG$_o$-**G*$_o$A***	P=S/P=O	3'-*O*-MOE
17179	**A***TG-CAT-TCT-GCC-CCC-AAG-G**A***	P=S	3'-*O*-MOE
17180	**A***TG-CAT-TCT-GCC-CCC-AAG-**G*A***	P=S	3'-*O*-MOE
17181	**A*$_o$**TG-CAT-TCT-GCC-AAA-AAG$_o$-**G*$_o$A***	P=S/P=O	3'-*O*-MOE
21415	**A*T***G-CAT-TCT-GCC-AAA-AAG-**G*A***	P=S	3'-*O*-MOE
21416	**A*$_o$T*$_o$**G-CAT-TCT-GCC-AAA-AAG$_o$-**G*$_o$A***	P=S/P=O	3'-*O*-MOE
21945	**A*A*A***	P=O	3'-*O*-MOE
21663	**A*A*A*A***	P=O	3'-*O*-MOE
20389	**A*U*C*G***	P=O	3'-*O*-MOE
20390	**C*G*C*-G*A*A*-T*T*C*-G*C*G***	P=O	3'-*O*-MOE

[1]All nucleosides with an asterisk contain 3'-*O*-(2-methoxyethyl).

EXAMPLE 50

General Procedure for purification of oligonucleotides

Following cleavage and deprotection step, the crude oligonucleotides (such as those synthesized in Example 49) were filtered from CPG using Gelman 0.45 μm nylon acrodisc syringe filters. Excess NH_4OH was evaporated away in a Savant AS160 automatic speed vac. The crude yield was measured on a Hewlett Packard 8452A Diode Array Spectrophotometer at 260 nm. Crude samples were then analyzed by mass spectrometry (MS) on a Hewlett Packard electrospray mass spectrometer and by capillary gel electrophoresis (CGE) on a Beckmann P/ACE system 5000. Trityl-on oligonucleotides were purified by

reverse phase preparative high performance liquid chromatography (HPLC). HPLC conditions were as follows: Waters 600E with 991 detector; Waters Delta Pak C4 column (7.8X300mm); Solvent A: 50 mM triethylammonium acetate (TEA-Ac), pH 7.0; B: 100% acetonitrile; 2.5 mL/min flow rate; Gradient: 5% B for first five minutes with linear increase in B to 60% during the next 55 minutes. Larger oligo yields from the larger 20 µmol syntheses were purified on larger HPLC columns (Waters Bondapak HC18HA) and the flow rate was increased to 5.0 mL/min. Appropriate fractions were collected and solvent was dried down in speed vac. Oligonucleotides were detritylated in 80% acetic acid for approximately 45 minutes and lyophilized again. Free trityl and excess salt were removed by passing detritylated oligonucleotides through Sephadex G-25 (size exclusion chromatography) and collecting appropriate samples through a Pharmacia fraction collector. Solvent again evaporated away in speed vac. Purified oligonucleotides were then analyzed for purity by CGE, HPLC (flow rate: 1.5 mL/min; Waters Delta Pak C4 column, 3.9X300mm), and MS. The final yield was determined by spectrophotometer at 260 nm.

Table II

Physical characteristics of 3'-*O*-(2-methoxyethyl) containing 2'-5' linked oligonucleotides.

	Expected Mass	Observed Mass	HPLC[2] T_R (min.)	#Ods(260nm) Purified
17176	6440.743	6440.300	23.47	3006
17177	6514.814	6513.910	23.67	3330
17178	6482.814	6480.900	23.06	390
17179	6513.798	6513.560	23.20	3240
17180	6588.879	6588.200	23.96	3222
17181	6540.879	6539.930	23.01	
21415	6662.976	6662.700	24.18	4008
21416	6598.969	6597.800	23.01	3060
21945	1099.924	1099.300	19.92	121
21663	1487.324	1486.800	20.16	71
20389	1483.000	1482.000		62
20390	4588.000	4591.000		151

[2]Conditions: Waters 600E with detector 991; Waters C4 column (3.9X300mm); Solvent A: 50 mM TEA-Ac, pH 7.0; B: 100% acetonitrile; 1.5 mL/min. flow rate; Gradient: 5% B for first five minutes with linear increase in B to 60% during the next 55 minutes.

EXAMPLE 51

T_m Studies on modified oligonucleotides

Oligonucleotides synthesized in Examples 49 and 50 were evaluated for their relative ability to bind to their complementary nucleic acids by measurement of their melting temperature (T_m). The melting temperature (T_m), a characteristic physical property of double helices, denotes the temperature (in degrees centigrade) at which 50% helical (hybridized) versus coil (unhybridized) forms are present. T_m is measured by using the UV spectrum to determine the formation and breakdown (melting) of the hybridization complex. Base stacking, which occurs during hybridization, is accompanied by a reduction in UV absorption

(hypochromicity). Consequently, a reduction in UV absorption indicates a higher T_m. The higher the T_m, the greater the strength of the bonds between the strands.

Selected test oligonucleotides and their complementary nucleic acids were incubated at a standard concentration of 4 μM for each oligonucleotide in buffer (100 mM NaCl, 10 mM sodium phosphate, pH 7.0, 0.1 mM EDTA). Samples were heated to 90 °C and the initial absorbance taken using a Guilford Response II Spectrophotometer (Corning). Samples were then slowly cooled to 15 °C and then the change in absorbance at 260 nm was monitored with heating during the heat denaturation procedure. The temperature was increased by 1 degree °C/absorbance reading and the denaturation profile analyzed by taking the 1st derivative of the melting curve. Data was also analyzed using a two-state linear regression analysis to determine the Tm=s. The results of these tests for the some of the oligonucleotides from Examples 49 and 50 are shown in Table III below.

Table III

Tm Analysis of Oligonucleotides

ISIS #	Sequence (5'-3')	Backbone	T_m	# Mods	#2'-5' Linkages
11061	ATG-CAT-TCT-GCC-CCC-AAG-GA (SEQ ID NO: 14)	P=S	61.4	0	0
17176	ATG-CAT-TCT-GCC-CCC-AAG-G**A***	P=S	61.4	1	0
17177	ATG-CAT-TCT-GCC-CCC-AAG-**G*A***	P=S	61.3	2	1
17178	ATG-CAT-TCT-GCC-CCC-AAG$_o$-**G*$_o$A***	P=S/P=O	61.8	2	1
17179	**A***TG-CAT-TCT-GCC-CCC-AAG-G**A***	P=S	61.1	2	1
17180	**A***TG-CAT-TCT-GCC-CCC-AAG-**G*A***	P=S	61.0	3	2
17181	**A*$_o$**TG-CAT-TCT-GCC-AAA-AAG$_o$-**G*$_o$A*** (SEQ ID NO: 15)	P=S/P=O	61.8	3	2
21415	**A*T***G-CAT-TCT-GCC-AAA-AAG-**G*A***	P=S	61.4	4	3
21416	**A*$_o$T*$_o$**G-CAT-TCT-GCC-AAA-AAG$_o$-**G*$_o$A***	P=S/P=O	61.7	4	3

[1]All nucleosides with an asterisk contain 3'-*O*-(2-methoxyethyl).

EXAMPLE 52

NMR experiments on modified oligonucleotides comparison of 3',5' versus 2',5' internucleotide linkages and 2'-substituents versus 3'-substituents by NMR

The 400MHz ^1H spectrum of oligomer d(GAU_2*CT), where U_2*= 2'-*O*-aminohexyluridine showed 8 signals between 7.5 and 9.0 ppm corresponding to the 8 aromatic protons. In addition, the anomeric proton of U* appears as a doublet at 5.9 ppm with $J_{1',2'}$= 7.5Hz, indicative of C2'-endo sugar puckering. The corresponding 2'-5' linked isomer shows a similar structure with $J_{1',2'}$= 3.85 Hz at 5.75 ppm, indicating an RNA type sugar puckering at the novel modification site favorable for hybridization to an mRNA target. The proton spectrum of the oligomer d($GACU_3$*), where U_3*=3'-*O*-hexylamine, showed the expected 7 aromatic proton signals between 7.5 and 9.0 ppm and the anomeric proton doublet at 5.9 ppm with $J_{1',2'}$= 4.4 Hz. This suggests more of a C3'-endo puckering for the 3'-*O*-alkylamino compounds compared to their 2' analogs. ^{31}P NMR of these oligonucleotides showed the expected 4 and 3 signals from the internucleotide phosphate linkages for d(GAU*CT) and d(GACU*), respectively. 3'-5' Linked vs. 2'-5' linked have different retention times in RP HPLC and hence different lipophilicities, implying potentially different extent of interactions with cell membranes.

EXAMPLE 53

T_m Analysis of 2',5'-linked oligonucleotides versus 3',5'-linked oligonucleotides

Thermal melts were done as per standarized literature procedures. Oligonucleotide identity is as follows:
Oligonucleotide A is a normal 3'-5' linked phosphodiester oligodeoxyribonucleotide of the sequence d(GGC TGU* CTG CG)(SEQ ID NO: 16) where the * indicates the attachment site

of a 2'-aminolinker. Oligonucleotide B is a normal 3'-5' linked phosphodiester oligoribonucleotide of the sequence d(GGC TGU* CTG CG) where the * indicates the attachment site of a 2'-aminolinker. Each of the ribonucleotides of the oligonucleotide, except the one bearing the * substituent, are 2'-O-methyl ribonucleotides. Oligonucleotide C has 2'-5' linkage at the * position in addition to a 3'-aminolinker at this site. The remainder of the oligonucleotide is a phosphodiester oligodeoxyribonucleotide of the sequence d(GGC TGU* CTG CG). The base oligonucleotide (no 2'-aminolinker) was not included in the study.

Table IIIa

OLIGONUCLEOTIDE	MODIFICATION	DNA TARGET	RNA TARGET
A	none	52.2	54.1
	2'-aminolinker	50.9	55.5
B	none	51.5	72.3
	2'-aminolinker	49.8	69.3
C	none	NA	
	3'-aminolinker	42.7	51.7

The 2'-5' linkages demonstrated a higher melting temperature against an RNA target compared to a DNA target.

EXAMPLE 54

Snake Venom Phosphodiesterase and Liver Homogenate Experiments on Oligonucleotide Stability

The following oligonucleotides were synthesized following the procedure of Example 49.

Table IV

Modified Oligonucleotides

synthesized to evaluate stability

ISIS #	Sequence (5'-3')	Backbone	Chemistry
22110	TTT-TTT-TTT-TTT-TTT-**T*T*T*-T*** (SEQ ID NO: 17)	P=O	3'-*O*-MOE
22111	TTT-TTT-TTT-TTT-TTT-**T#T#T#-U#** (SEQ ID NO: 18)	P=O	3'-*O*-MOE
22112	TTT-TTT-TTT-TTT-TTT-**T*T*T*-T***	P=S	3'-*O*-MOE
22113	TTT-TTT-TTT-TTT-TTT-**T#T#T#-U#**	P=S	3'-*O*-MOE
22114	TTT-TTT-TTT-TTT-TTT$_o$-**T*$_o$T*$_o$T*$_o$T***	P=S/P=O	3'-*O*-MOE
22115	TTT-TTT-TTT-TTT-TTT$_o$-**T#$_o$T#$_o$T#$_o$-U#**	P=S/P=O	3'-*O*-MOE

[1]All nucleosides with an asterisk contain 3'-*O*-(2-methoxyethyl). All nucleosides with a # contain 2'-*O*-(2-methoxyethyl).

The oligonucleotides were purified following the procedure of Example 50 and analyzed for their structure.

Table V

Properties of Modified Oligonucleotides

ISIS	#Sequence (5'-3')[1]	Expected Mass	Observed Mass	HPLC[2] T_R (min.)	#Ods(260nm) Purified
22110	TTT-TTT-TTT-TTT-TTT-**T*T*T*-T***	6314.189	6315.880	20.39	174
22111	TTT-TTT-TTT-TTT-TTT-**T#T#T#-U#**	6004.777	5997.490	20.89	147
22112	TTT-TTT-TTT-TTT-TTT-**T*T*T*-T***	6298.799	6301.730	25.92	224
22113	TTT-TTT-TTT-TTT-TTT-**T#T#T#-U#**	6288.745	6286.940	24.77	209
22114	TTT-TTT-TTT-TTT-TTT$_o$-**T*$_o$T*$_o$T*$_o$-T***	6234.799	6237.150	24.84	196
22115	TTT-TTT-TTT-TTT-TTT$_o$-**T#$_o$T#$_o$T#$_o$-U#**	6224.745	6223.780	23.30	340

[1]All nucleosides with an asterisk contain 3'-*O*-(2-methoxyethyl). All nucleosides with a # contain 2'-*O*-(2-methoxy) ethyl. [2]Conditions: Waters 600E with detector 991; Waters C4 column (3.9X300mm); Solvent A: 50 mM TEA-Ac, pH 7.0; B: 100% acetonitrile; 1.5 mL/min. flow rate;

Gradient: 5% B for first five minutes with linear increase in B to 60% during the next 55 minutes.

EXAMPLE 55

3'-O-Aminopropyl modified oligonucleotides

Following the procedures illustrated above for the synthesis of oligonucleotides, modified 3'-amidites were used in addition to conventional amidites to prepare the oligonucleotides listed in tables VI and VII. Nucleosides used include: N6-benzoyl-3'-O-propylphthalimido-A-2'-amidite, 2'-O-propylphthaloyl-A-3'-amidite, 2'-O-methoxyethyl-thymidine-3'-amidite (RIC, Inc.), 2'-O-MOE-G-3'-amidite (RI Chemical), 2'-O-methoxyethyl-5-methylcytidine-3'-amidite, 2'-O-methoxyethyl-adenosine-3'-amidite (RI Chemical), and 5-methylcytidine-3'-amidite. 3'-propylphthalimido-A and 2'-propylphthalimido-A were used as the LCA-CPG solid support. The required amounts of the amidites were placed in dried vials, dissolved in acetonitrile (unmodified nucleosides were made into 1M solutions and modified nucleosides were 100 mg/mL), and connected to the appropriate ports on a Millipore Expedite™ Nucleic Acid Synthesis System. Solid support resin (60 mg) was used in each column for 2X1 µmole scale synthesis (2 columns for each oligo were used). The synthesis was run using the IBP-PS(1 µmole) coupling protocol for phosphorothioate backbones and CSO-8 for phosphodiesters. The trityl reports indicated normal coupling results.

After synthesis the oligonucleotides were deprotected with conc. ammonium hydroxide(aq) containing 10% of a solution of 40% methylamine (aq) at 55 °C for approximately 16 hrs. Then they were evaporated, using a Savant AS160 Automatic SpeedVac, (to remove ammonia) and filtered to remove the CPG-resin. The crude samples were analyzed by MS, HPLC, and CE. Then they were purified on a Waters 600E HPLC system with a 991 detector using a Waters

C4 Prep. scale column (Alice C4 Prep.) and the following solvents: A: 50 mM TEA-Ac, pH 7.0 and B: acetonitrile utilizing the AMPREP2@ method. After purification the oligonucleotides were evaporated to dryness and then detritylated with 80% acetic acid at room temp. for approximately 30 min. Then they were evaporated. The oligonucleotides were dissolved in conc. ammonium hydroxide and run through a column containing Sephadex G-25 using water as the solvent and a Pharmacia LKB SuperFrac fraction collector. The resulting purified oligonucleotides were evaporated and analyzed by MS, CE, and HPLC.

Table VI

Oligonucleotides bearing Aminopropyl Substituents

ISIS #	Sequence (5'-3')[1]	Backbone
23185-1	**A***TG-CAT-TCT-GCC-CCC-AAG-G**A***	P=S
23186-1	**A***TG-CAT-TCT-GCC-CCC-AAG-G**A***	P=S
23187-1	**A***$_o$T$_o$G$_o$-C$_o$A$_s$T$_s$-T$_s$C$_s$T$_s$-G$_s$C$_s$C$_s$-C$_s$C$_s$C$_s$-A$_o$A$_o$G$_o$-G$_o$**A***	P=S/P=O
23980-1	**A***$_o$T$_o$G$_o$-C$_o$A$_s$T$_s$-T$_s$C$_s$T$_s$-G$_s$C$_s$C$_s$-C$_s$C$_s$C$_s$-A$_o$A$_o$G$_o$-G$_o$**A***	P=S/P=O
23981-1	**A***TG-CAT-TCT-GCC-CCC-AAG-G**A***	P=S
23982-1	**A***TG-CAT-TCT-GCC-CCC-AAG-G**A***	P=S

[1]All underlined nucleosides bear a 2'-O-methoxyethyl substituent; internucleotide linkages in PS/PO oligonucleotides are indicated by subscript >s= and >o= notations respectively; **A*** = 3'-aminopropyl-A; **A*** = 2'-aminopropyl-A; C = 5-methyl-C

Table VII

Aminopropyl Modified Oligonucleotides

ISIS #	Expected Mass (g/mol)	Observed Mass (g/mol)	HPLC Retention Time(min)	CE Retention Time(min)	Crude Yield (Ods)	Final Yield (Ods)
23185-1	6612.065	6610.5	23.19	5.98	948	478
23186-1	7204.697	7203.1	24.99	6.18	1075	379
23187-1	7076.697	7072.33	23.36	7.56	838	546

23980-1	7076.697	7102.31	23.42	7.16	984	373
23981-1	7204.697	7230.14	25.36	7.18	1170	526
23982-1	6612.065	6635.71	23.47	7.31	1083	463

EXAMPLE 56

***In vivo* stability of modified oligonucleotides**

The *in vivo* stability of selected modified oligonucleotides synthesized in Examples 49 and 55 was determined in BALB/c mice. Following a single i.v. administration of 5 mg/kg of oligonucleotide, blood samples were drawn at various time intervals and analyzed by CGE. For each oligonucleotide tested, 9 male BALB/c mice (Charles River, Wilmington, MA) weighing about 25 g were used. Following a one week acclimatization the mice received a single tail-vein injection of oligonucleotide (5 mg/kg) administered in phosphate buffered saline (PBS), pH 7.0. One retro-orbital bleed (either at 0.25, 0.5, 2 or 4 h post-dose) and a terminal bleed (either 1, 3, 8, or 24 h post-dose) were collected from each group. The terminal bleed (approximately 0.6-0.8 mL) was collected by cardiac puncture following ketamine/xylazine anasthesia. The blood was transferred to an EDTA-coated collection tube and centrifuged to obtain plasma. At termination, the liver and kidneys were collected from each mouse. Plasma and tissue homogenates were used for analysis to determine intact oligonucleotide content by CGE. All samples were immediately frozen on dry ice after collection and stored at -80C until analysis.

The CGE analysis inidcated the relative nuclease resistance of 2',5'-linked oligomers compared to ISIS 11061 (Table III, Example 51) (uniformly 2'-deoxy-phosphorothioate oligonucleotide targeted to mouse *c-raf*). Because of the nuclease resistance of the 2',5'-linkage, coupled with the fact that 3'-methoxyethoxy substituents are present and afford further nuclease protection the

oligonucleotides ISIS 17176, ISIS 17177, ISIS 17178, ISIS 17180, ISIS 17181 and ISIS 21415 were found to be more stable in plasma, while ISIS 11061 (Table III) was not. Similar observations were noted in kidney and liver tissue. This implies that 2',5'-linkages with 3'-methoxyethoxy substituents offer excellent nuclease resistance in plasma, kidney and liver against 5'-exonucleases and 3'-exonucleases. Thus oligonucleotides with longer durations of action can be designed by incorporating both the 2',5'-linkage and 3'-methoxyethoxy motifs into their structure. It was also observed that 2',5'-phosphorothioate linkages are more stable than 2',5'-phosphodiester linkages. A plot of the percentage of full length oligonucleotide remaining intact in plasma one hour following administration of an *i.v.* bolus of 5 mg/kg oligonucleotide is shown in Figure 4.

A plot of the percentage of full length oligonucleotide remaining intact in tissue 24 hours following administration of an *i.v.* bolus of 5 mg/kg oligonucleotide is shown in Figure 5.

CGE traces of test oligonucleotides and a standard phosphorothioate oligonucleotide in both mouse liver samples and mouse kidney samples after 24 hours are shown in Figure 6. As is evident from these traces, there is a greater amount of intact oliogonucleotide for the oligonucleotides of the invention as compared to the standard seen in panel A. The oligonucleotide shown in panel B included one substituent of the invention at each of the 5' and 3' ends of a phosphorothioate oligonucleotide while the phosphorothioate oligonucleotide seen in panel C included one substituent at the 5' end and two at the 3' end. The oligonucleotide of panel D includes a substituent of the invention incorporated in a 2',5' phosphodiester linkage at both its 5' and 3' ends. While less stable than the oligonucleotide seen in panels B and C, it is more

stable than the full phosphorothioate standard oligonucleotide of panel A.

EXAMPLE 57

Control of *c-raf* message in bEND cells using modified oligonucleotides

In order to assess the activity of some of the oligonucleotides, an in vitro cell culture assay was used that measures the cellular levels of c-raf expression in bEND cells.

Cells and Reagents

The bEnd.3 cell line, a brain endothelioma, was obtained from Dr. Werner Risau (Max-Planck Institute). Opti-MEM, trypsin-EDTA and DMEM with high glucose were purchased from Gibco-BRL (Grand Island, NY). Dulbecco's PBS was purchased from Irvine Scientific (Irvine, CA). Sterile, 12 well tissue culture plates and Facsflow solution were purchased from Becton Dickinson (Mansfield, MA). Ultrapure formaldehyde was purchased from Polysciences (Warrington, PA). NAP-5 columns were purchased from Pharmacia (Uppsala, Sweden).

Oligonucleotide Treatment

Cells were grown to approximately 75 % confluency in 12 well plates with DMEM containing 4.5 g/L glucose and 10 % FBS. Cells were washed 3 times with Opti-MEM pre-warmed to 37 °C. Oligonucleotide were premixed with a cationic lipid (Lipofectin reagent, (GIBCO/BRL) and, serially diluted to desired concentrations and transferred on to washed cells for a 4 hour incubation at 37 °C. Media was then removed and replaced with normal growth media for 24 hours for northern blot analysis of mRNA.

Northern Blot Analysis

For determination of mRNA levels by Northern blot analysis, total RNA was prepared from cells by the

guanidinium isothiocyanate procedure (Monia *et al.*, *Proc. Natl. Acad. Sci. USA*, **1996**, *93*, 15481-15484) 24 h after initiation of oligonucleotide treatment. Total RNA was isolated by centrifugation of the cell lysates over a CsCl cushion. Northern blot analysis, RNA quantitation and normalization to G#PDH mRNA levels were done according to a reported procedure (Dean and McKay, *Proc. Natl. Acad. Sci. USA*, **1994**, *91*, 11762-11766). In bEND cells the 2-,5-linked-3'-O-methoxyethyl oligonucleotides showed reduction of *c-raf* message activity as a function of concentration. The fact that these modified oligonucleotides retained activity promises reduced frequency of dosing with these oligonucleotides which also show increased *in vivo* nuclease resistance. All 2',5'-linked oligonucleotides retained the activity of parent 11061 (Table III) oligonucleotide and improved the activity even further. A graph of the effect of the oligonucleotides of the present invention on *c-raf* expression (compared to control) in bEND cells is shown in Figure 7.

EXAMPLE 58

Synthesis of MMI-containing Oligonucleotides

a. Bis-2'-*O*-methyl MMI Building Blocks

The synthesis of MMI (*i.e.*, R = CH_3) dimer building blocks have been previously described (*see, e.g.*, Swayze, *et al.*, *Synlett* **1997**, 859; Sanghvi, *et al.*, *Nucleosides & Nucleotides* **1997**, *16* 907; Swayze, *et al.*, *Nucleosides & Nucleotides* **1997**, *16*, 971; Dimock, *et al.*, *Nucleosides & Nucleotides* **1997**, *16*, 1629). Generally, 5'-*O*-(4,4'-dimethoxytrityl)-2'-*O*-methyl-3'-C-formyl nucleosides were condensed with 5'-*O*-(N-methylhydroxylamino)-2'-*O*-methyl-3'-*O*-TBDPS nucleosides using 1 equivalent of BH_3 pyridine/1 equivalent of

pyridinium para-toluene sulfonate (PPTS) in 3:1 MeOH/THF. The resultant MMI dimer blocks were then deprotected at the lower part of the sugar with 15 equivalents of Et_3N-2HF in THF. Thus the T^*G^{iBu} dimer unit was synthesized and phosphitylated to give T*G(MMI) phosphoramidite. In a similar fashion, A^{Bz}*T(MMI) dimer was synthesized, succinylated and attached to controlled pore glass.

b. Oligonucleotide synthesis

Oligonucleotides were synthesized on a Perseptive Biosystems Expedite 8901 Nucleic Acid Synthesis System. Multiple 1-µmol syntheses were performed for each oligonucleotide. $A^*_{MMI}T$ solid support was loaded into the column. Trityl groups were removed with trichloroacetic acid (975 µL over one minute) followed by an acetonitrile wash. The oligonucleotide was built using a modified thioate protocol. Standard amidites were delivered (210 µL) over a 3 minute period in this protocol. The $T^*_{MMI}G$ amidite was double coupled using 210 µL over a total of 20 minutes. The amount of oxidizer, 3H-1,2-benzodithiole-3-one-1,1-dioxide (Beaucage reagent, 3.4 g Beaucage reagent/200 mL acetonitrile), was 225 µL (one minute wait step). The unreacted nucleoside was capped with a 50:50 mixture of tetrahyrdofuran/acetic anhydride and tetrahydrofuran/pyridine/1-methyl imidazole. Trityl yields were followed by the trityl monitor during the duration of the synthesis. The final DMT group was left intact. After the synthesis, the contents of the synthesis cartridge (1 µmole) were transferred to a Pyrex vial and the oligonucleotide was cleaved from the controlled pore glass (CPG) using 5 mL of 30% ammonium hydroxide (NH_4OH) for approximately 16 hours at 55 °C.

c. Oligonucleotide Purification

After the deprotection step, the samples were filtered from CPG using Gelman 0.45 µm nylon acrodisc

syringe filters. Excess NH_4OH was evaporated away in a Savant AS160 automatic SpeedVac. The crude yield was measured on a Hewlett Packard 8452A Diode Array Spectrophotometer at 260 nm. Crude samples were then analyzed by mass spectrometry (MS) on a Hewlett Packard electrospray mass spectrometer. Trityl-on oligonucleotides were purified by reverse phase preparative high performance liquid chromatography (HPLC). HPLC conditions were as follows: Waters 600E with 991 detector; Waters Delta Pak C4 column (7.8X300mm); Solvent A: 50 mM triethylammonium acetate (TEA-Ac), pH 7.0; B: 100% acetonitrile; 2.5 mL/min flow rate; Gradient: 5% B for first five minutes with linear increase in B to 60% during the next 55 minutes. Fractions containing the desired product (retention time = 41 min. for DMT-ON-16314; retention time = 42.5 min. for DMT-ON-16315) were collected and the solvent was dried off in the SpeedVac. Oligonucleotides were detritylated in 80% acetic acid for approximately 60 minutes and lyophilized again. Free trityl and excess salt were removed by passing detritylated oligonucleotides through Sephadex G-25 (size exclusion chromatography) and collecting appropriate samples through a Pharmacia fraction collector. The solvent was again evaporated away in a SpeedVac. Purified oligonucleotides were then analyzed for purity by CGE, HPLC (flow rate: 1.5 mL/min; Waters Delta Pak C4 column, 3.9X300mm), and MS. The final yield was determined by spectrophotometer at 260 nm.

The synthesized oligonucleotides and their physical characteristics are shown, respectively, in Tables VIII and IX. All nucleosides with an asterisk contain MMI linkage.

Table VIII

ICAM-1 Oligonucleotides Containing MMI Dimers Synthesized for *in Vivo* Nuclease and Pharmacology Studies.

ISIS #	Sequence (5'-3')	Backbone	2'-Chemistry
16314	TGC ATC CCC CAG GCC ACC A*T (SEQ ID NO: 19)	P=S, MMI	Bis-2'-OMe-MMI, 2'-H
16315	T*GC ATC CCC CAG GCC ACC A*T	P=S, MMI	Bis-2'-OMe-MMI, 2'-H
3082	TGC ATC CCC CAG GCG ACC AT	P=S	2'-H, single mismatch
13001	TGC ATC CCC CAG GCC ACC AT	P=S	2'-H

Table IX

Physical Characteristics of MMI Oligomers Synthesized for Pharmacology, and *In Vivo* Nuclease Studies

ISIS #	Sequence (5'-3')	Expected Mass (g)	Observed Mass (g)	HPLC Retn. Time (min)
16314	TGC ATC CCC CAG GCC ACC A*T	6295	6297	23.9
16315	T*G C ATC CCC CAG GCC ACC A*T	6302	6303	24.75

HPLC Conditions: Waters 600E with detector 991; Waters C4 column (3.9X300mm); Solvent A: 50 mM TEA-Ac, pH 7.0; B: 100% acetonitrile; 1.5 mL/min. flow rate; Gradient: 5% B for first five minutes with linear increase in B to 60% during the next 55 minutes.

EXAMPLE 59

Synthesis of Sp Terminal Oligonucleotide

a. 3'-*O*-t-Butyldiphenylsilyl-thymidine (1)

5'-*O*-Dimethoxytritylthymidine is silylated with 1 equivalent of t-butyldiphenylsilyl chloride (TBDPSCl) and 2 equivalents of imidazole in DMF solvent at room temperature. The 5'-protecting group is removed by treating with 3% dichloracetic acid in CH_2Cl_2.

b. 5'-*O*-Dimethoxytrityl-thymidin-3'-*O*-yl-N,N-diisopropylamino (S-pivaloyl-2-mercaptoethoxy) phosphoramidite (2)

5'-*O*-Dimethoxytrityl thymidine is treated with bis-(N,N-diisopropylamino)-S-pivaloyl-2-mercaptoethoxy phosphoramidite and tetrazole in CH_2Cl_2/CH_3CN as described by Guzaev *et al., Bioorganic & Medicinal Chemistry Letters* **1998,** *8,* 1123) to yield the title compound.

c. 5'-*O*-Dimethoxytrityl-2'-deoxy-adenosin-3'-*O*-yl-N,N-diisopropylamino (S-pivaloyl-2-mercapto ethoxy) phosphoramidite (3)

5'-*O*-Dimethoxytrityl-N-6-benzoyl-2'-deoxy-adenosine is phosphitylated as in the previous example to yield the desired amidite.

d. 3'-*O*-t-Butyldiphenylsilyl-2'-deoxy-N_2-isobutyryl-guanosine (4)

5'-*O*-Dimethoxytrityl-2'-deoxy-N_2-isobutyryl-guanisine is silylated with TBDPSCl and imidazole in DMF. The 5'-DMT is then removed with 3% DCA in CH_2Cl_2.

e. $T_{(Sp)}G$ dimers and $T_{(S)}$ Phosphoramidite

Compounds **4** and **2** are condensed (1:1 equivalents) using 1H-tetrazole in CH_3CN solvent followed by sulfurization employing Beaucage reagent (*see, e.g.,* Iyer, *et al., J. Org. Chem.* **1990,** *55,* 4693). The dimers (TG) are separated by column chromatography and the silyl group is

deprotected using t-butyl ammonium fluoride/THF to give Rp and Sp dimers of T_sG. Small amounts of these dimers are completely deprotected and treated with either P1 nuclease or snake venom phosphodiesterase. The R isomer is resistant to P1 nuclease and hydrolyzed by SVPD. The S isomer is resistant to SVPD and hydrolyzed P1 nuclease. The Sp isomer of the fully protected T_sG dimer is phosphitylated to give DMT-T-Sp-G-phosphoramidite.

f. A_sT Dimers and Solid Support Containing $A_{SP}T$ Dimer

Compounds **3** and **1** are condensed using 1H-tetrazole in CH_3CN solvent followed by sulfurization to give AT dimers. The dimers are separated by column chromatography and the silyl group is deprotected with TBAF/THF. The configurational assignments are done generally as in the previous example. The Sp isomer is then attached to controlled pore glass according to standard procedures to give DMT-A_{SP}-T-CPG oligomerization with chirally pure Sp dimer units at the termini.

g. Oligonucleotide Synthesis

The oligonucleotide having the sequence T*GC ATC CCC CAG GCC ACC A*T is synthesized, where T*G and A*T represent chiral Sp dimer blocks described above. DMT-A_{SP}-T-CPG is taken in the synthesis column and the next 16b residues are built using standard phosphorothioate protocols and 3H-1,2-benzodithiol-3-one 1,1 dioxide as the sulfurizing agent. After building this 18 mer unit followed by final detritylation, the chiral Sp dimer phosphoramidite of 5'-DMT-T_{SP}-G amidite is coupled to give the desired antisense oligonucleotide. This compound is then deprotected in 30% NH_4OH over 16 hours and the oligomer purified in HPLC and desalted in Sephader G-25 column. The final oligomer has Sp configuration at the 5'-

terminus and 3'-terminus and the interior has diastereomeric mixture of Rp and Sp configurations.

EXAMPLE 60

Evaluation of in vivo stability of MMI capped oligonucleotides mouse experiment procedures

For each oligonucleotide tested, 9 male BALB/c mice (Charles River, Wilmington, MA), weighing about 25 g was used (Crooke *et al., J. Pharmacol. Exp. Ther.*, 1996, *277*, 923). Following a 1-week acclimation, mice received a single tail vein injection of oligonucleotide (5 mg/kg) administered in phosphate buffered saline (PBS), pH 7.0 One retro-orbital bleed (either 0.25, 0.5, 2 or 4 lv post dose) and a terminal bleed (either 1, 3, 8 or 24 h post dose) were collected from each group. The terminal bleed (approximately 0.6-0.8 mL) was collected by cardiac puncture following ketamine/xylazine anesthesia. The blood was transferred to an EDTA-coated collection tube and centrifuged to obtain plasma. At termination, the liver and kidneys were collected from each mouse. Plasma and tissues homogenates were used for analysis for determination of intact oligonucleotide content by CGE. All samples were immediately frozen on dry ice after collection and stored at -80 °C until analysis.

The capillary gel electrophoretic analysis indicated the relative nuclease resistance of MMI capped oligomers compared to ISIS 3082 (uniform 2'-deoxy phosphorothioate). Because of the resistance of MMI linkage to nucleases, the compound 16314 was found to be stable in plasma while 3082 was not. However, in kidney and liver, the compound 16314 also showed certain amount of degradation. This implied that while 3'-exonuclease is important in plasma, 5'-exonucleases or endonucleases may be active in tissues. To distinguish between these two

possibilities, the data from 16315 was analyzed. In plasma as well as in tissues, (liver and kidney) the compound was stable in various time points. (1, 3 and 24 hrs.). The fact that no degradation was detected proved that 5'-exonucleases and 3'-exonuclease are prevalent in tissues and endonucleases are not active. Furthermore, a single linkage (MMI or Sp thioate linkage) is sufficient as a gatekeeper against nucleases.

Control of ICAM-1 Expression Cells and Reagents

The bEnd.3 cell line, a brain endothelioma, was the kind gift of Dr. Werner Risau (Max-Planck Institute). Opti-MEM, trypsin-EDTA and DMEM with high glucose were purchased from Gibco-BRL (Grand Island, NY). Dulbecco's PBS was purchased from Irvine Scientific (Irvine, CA). Sterile, 12 well tissue culture plates and Facsflow solution were purchased from Becton Dickinson (Mansfield, MA). Ultrapure formaldehyde was purchased from Polysciences (Warrington, PA). Recombinant human TNF-a was purchased from R&D Systems (Minneapolis, MN). Mouse interferon-γ was purchased from Genzyme (Cambridge, MA). Fraction V, BSA was purchased from Sigma (St. Louis, MO). The mouse ICAM-1-PE, VCAM-1-FITC, hamster IgG-FITC and rat IgG_{2a}-PE antibodies were purchased from Pharmingen (San Diego, CA). Zeta-Probe nylon blotting membrane was purchased from Bio-Rad (Richmond, CA). QuickHyb solution was purchased from Stratagene (La Jolla, CA). A cDNA labeling kit, Prime-a-Gene, was purchased from ProMega (Madison, WI). NAP-5 columns were purchased from Pharmacia (Uppsala, Sweden).

Oligonucleotide Treatment

Cells were grown to approximately 75 % confluency in 12 well plates with DMEM containing 4.5 g/L glucose and 10 % FBS. Cells were washed 3 times with Opti-MEM pre-warmed to 37 °C. Oligonucleotide was premixed with Opti-

MEM, serially diluted to desired concentrations and transferred onto washed cells for a 4 hour incubation at 37 °C. Media was removed and replaced with normal growth media with or without 5 ng/mL TNF-α and 200 U/mL interferon-γ, incubated for 2 hours for northern blot analysis of mRNA or overnight for flow cytometric analysis of cell surface protein expression.

Flow Cytometry

After oligonucleotide treatment, cells were detached from the plates with a short treatment of trypsin-EDTA (1-2 min.). Cells were transferred to 12x75 mm polystyrene tubes and washed with 2% BSA, 0.2% sodium azide in D-PBS at 4 °C. Cells were centrifuged at 1000 rpm in a Beckman GPR centrifuge and the supernatant was then decanted. ICAM-1, VCAM-1 and the control antibodies were added at 1 ug/mL in 0.3 mL of the above buffer. Antibodies were incubated with the cells for 30 minutes at 4 °C in the dark, under gentle agitation. Cells were washed again as above and then resuspended in 0.3 mL of FacsFlow buffer with 0.5 % ultrapure formaldehyde. Cells were analyzed on a Becton Dickinson FACScan. Results are expressed as percentage of control expression, which was calculated as follows: [((CAM expression for oligonucleotide-treated cytokine induced cells) - (basal CAM expression))/((cytokine-induced CAM expression) - (basal CAM expression))] X 100. For the experiments involving cationic lipids, both basal and cytokine-treated control cells were pretreated with Lipofectin for 4 hours in the absence of oligonucleotides.

The results reveal the following: 1) Isis 3082 showed an expected dose response (25-200 nM); 2) Isis 13001 lost its ability to inhibit ICAM-1 expression as expected from a mismatch compound, thus proving an antisense mechanism; 3) 3'-MMI capped oligomer 16314 improved the

activity of 3082, and at 200 nM concentration, nearly twice as active as 3082; 4) 5'- and 3'- MMI capped oligomer is the most potent compound and it is nearly 4 to 5 times more efficacious than the parent compound at 100 and 200 nM concentrations. Thus, improved nuclease resistance increased the potency of the antisense oligonucleotides.

EXAMPLE 61

Control of H-ras Expression

Antisense oligonucleotides targeting the H-ras message were tested for their ability to inhibit production of H-ras mRNA in T-24 cells. For these test, T-24 cells were plated in 6-well plates and then treated with various escalating concentrations of oligonucleotide in the presence of cationic lipid (Lipofectin, GIBCO) at the ratio of 2.5 µg/ml Lipofectin per 100 nM oligonucleotide. Oligonucleotide treatment was carried out in serum free media for 4 hours. Eighteen hours after treatment the total RNA was harvested and analyzed by northern blot for H-ras mRNA and control gene G3PDH. The data is presented in figures 8 and 9 in bar graphs as percent control normalized for the G3PDH signal. As can be seen, the oligonucleotide having a single MMI linkage in each of the flank regions showed significant reduction of H-ras mRNA.

EXAMPLE 62

5-Lipoxygenase Analysis and Assays

A. Therapeutics

For therapeutic use, an animal suspected of having a disease characterized by excessive or abnormal supply of 5-lipoxygenase is treated by administering a compound of the invention. Persons of ordinary skill can easily determine optimum dosages, dosing methodologies and repetition rates. Such treatment is generally continued

until either a cure is effected or a diminution in the diseased state is achieved. Long term treatment is likely for some diseases.

B. Research Reagents

The oligonucleotides of the invention will also be useful as research reagents when used to cleave or otherwise modulate 5-lipoxygenase mRNA in crude cell lysates or in partially purified or wholly purified RNA preparations. This application of the invention is accomplished, for example, by lysing cells by standard methods, optimally extracting the RNA and then treating it with a composition at concentrations ranging, for instance, from about 100 to about 500 ng per 10 Mg of total RNA in a buffer consisting, for example, of 50 mm phosphate, pH ranging from about 4-10 at a temperature from about 30 to about 50 °C. The cleaved 5-lipoxygenase RNA can be analyzed by agarose gel electrophoresis and hybridization with radiolabeled DNA probes or by other standard methods.

C. Diagnostics

The oligonucleotides of the invention will also be useful in diagnostic applications, particularly for the determination of the expression of specific mRNA species in various tissues or the expression of abnormal or mutant RNA species. In this example, while the macromolecules target a abnormal mRNA by being designed complementary to the abnormal sequence, they would not hybridize to normal mRNA. Tissue samples can be homogenized, and RNA extracted by standard methods. The crude homogenate or extract can be treated for example to effect cleavage of the target RNA. The product can then be hybridized to a solid support which contains a bound oligonucleotide complementary to a region on the 5' side of the cleavage site. Both the normal and abnormal 5' region of the mRNA would bind to the solid support. The 3' region of the abnormal RNA, which is

cleaved, would not be bound to the support and therefore would be separated from the normal mRNA.

Targeted mRNA species for modulation relates to 5-lipoxygenase; however, persons of ordinary skill in the art will appreciate that the present invention is not so limited and it is generally applicable. The inhibition or modulation of production of the enzyme 5-lipoxygenase is expected to have significant therapeutic benefits in the treatment of disease. In order to assess the effectiveness of the compositions, an assay or series of assays is required.

D. In Vitro Assays

The cellular assays for 5-lipoxygenase preferably use the human promyelocytic leukemia cell line HL-60. These cells can be induced to differentiate into either a monocyte like cell or neutrophil like cell by various known agents. Treatment of the cells with 1.3% dimethyl sulfoxide, DMSO, is known to promote differentiation of the cells into neutrophils. It has now been found that basal HL-60 cells do not synthesize detectable levels of 5-lipoxygenase protein or secrete leukotrienes (a downstream product of 5-lipoxygenase). Differentiation of the cells with DMSO causes an appearance of 5-lipoxygenase protein and leukotriene biosynthesis 48 hours after addition of DMSO. Thus induction of 5-lipoxygenase protein synthesis can be utilized as a test system for analysis of oligonucleotides which interfere with 5-lipoxygenase synthesis in these cells.

A second test system for oligonucleotides makes use of the fact that 5-lipoxygenase is a "suicide" enzyme in that it inactivates itself upon reacting with substrate. Treatment of differentiated HL-60 or other cells expressing 5 lipoxygenase, with 10 μM A23187, a calcium ionophore, promotes translocation of 5-lipoxygenase from the cytosol

to the membrane with subsequent activation of the enzyme. Following activation and several rounds of catalysis, the enzyme becomes catalytically inactive. Thus, treatment of the cells with calcium ionophore inactivates endogenous 5-lipoxygenase. It takes the cells approximately 24 hours to recover from A23187 treatment as measured by their ability to synthesize leukotriene B_4. Macromolecules directed against 5-lipoxygenase can be tested for activity in two HL-60 model systems using the following quantitative assays. The assays are described from the most direct measurement of inhibition of 5-lipoxygenase protein synthesis in intact cells to more downstream events such as measurement of 5-lipoxygenase activity in intact cells.

A direct effect which oligonucleotides can exert on intact cells and which can be easily be quantitated is specific inhibition of 5-lipoxygenase protein synthesis. To perform this technique, cells can be labeled with ^{35}S-methionine (50 µCi/mL) for 2 hours at 37 °C to label newly synthesized protein. Cells are extracted to solubilize total cellular proteins and 5-lipoxygenase is immunoprecipitated with 5-lipoxygenase antibody followed by elution from protein A Sepharose beads. The immunoprecipitated proteins are resolved by SDS-polyacrylamide gel electrophoresis and exposed for autoradiography. The amount of immunoprecipitated 5-lipoxygenase is quantitated by scanning densitometry.

A predicted result from these experiments would be as follows. The amount of 5-lipoxygenase protein immunoprecipitated from control cells would be normalized to 100%. Treatment of the cells with 1 µM, 10 µM, and 30 µM of the macromolecules of the invention for 48 hours would reduce immunoprecipitated 5-lipoxygenase by 5%, 25% and 75% of control, respectively.

Measurement of 5-lipoxygenase enzyme activity in cellular homogenates could also be used to quantitate the amount of enzyme present which is capable of synthesizing leukotrienes. A radiometric assay has now been developed for quantitating 5-lipoxygenase enzyme activity in cell homogenates using reverse phase HPLC. Cells are broken by sonication in a buffer containing protease inhibitors and EDTA. The cell homogenate is centrifuged at 10,000 x g for 30 min and the supernatants analyzed for 5-lipoxygenase activity. Cytosolic proteins are incubated with 10 µM ^{14}C-arachidonic acid, 2mM ATP, 50 µM free calcium, 100 µg/mL phosphatidylcholine, and 50mM bis-Tris buffer, pH 7.0, for 5 min at 37 °C. The reactions are quenched by the addition of an equal volume of acetone and the fatty acids extracted with ethyl acetate. The substrate and reaction products are separated by reverse phase HPLC on a Novapak C18 column (Waters Inc., Millford, MA). Radioactive peaks are detected by a Beckman model 171 radiochromatography detector. The amount of arachidonic acid converted into di-HETE's and mono-HETE's is used as a measure of 5-lipoxygenase activity.

A predicted result for treatment of DMSO differentiated HL-60 cells for 72 hours with effective the macromolecules of the invention at 1 µM, 10 µM, and 30 µM would be as follows. Control cells oxidize 200 pmol arachidonic acid/5 min/10^6 cells. Cells treated with 1 µM, 10 µM, and 30 µM of an effective oligonucleotide would oxidize 195 pmol, 140 pmol, and 60 pmol of arachidonic acid/5 min/10^6 cells respectively.

A quantitative competitive enzyme linked immunosorbant assay (ELISA) for the measurement of total 5-lipoxygenase protein in cells has been developed. Human 5-lipoxygenase expressed in *E. coli* and purified by extraction, Q-Sepharose, hydroxyapatite, and reverse phase HPLC is used

as a standard and as the primary antigen to coat microtiter plates. Purified 5-lipoxygenase (25 ng) is bound to the microtiter plates overnight at 4 °C. The wells are blocked for 90 min with 5% goat serum diluted in 20mM Tris!HCL buffer, pH 7.4, in the presence of 150mM NaCl (TBS). Cell extracts (0.2% Triton X-100, 12,000 x g for 30 min.) or purified 5-lipoxygenase were incubated with a 1:4000 dilution of 5-lipoxygenase polyclonal antibody in a total volume of 100 µL in the microtiter wells for 90 min. The antibodies are prepared by immunizing rabbits with purified human recombinant 5-lipoxygenase. The wells are washed with TBS containing 0.05% tween 20 (TBST), then incubated with 100 µL of a 1:1000 dilution of peroxidase conjugated goat anti-rabbit IgG (Cappel Laboratories, Malvern, PA) for 60 min at 25 °C. The wells are washed with TBST and the amount of peroxidase labeled second antibody determined by development with tetramethylbenzidine.

Predicted results from such an assay using a 30 mer oligonucleotide at 1 µM, 10 µM, and 30 µM would be 30 ng, 18 ng and 5 ng of 5-lipoxygenase per 10^6 cells, respectively with untreated cells containing about 34 ng 5-lipoxygenase.

A net effect of inhibition of 5-lipoxygenase biosynthesis is a diminution in the quantities of leukotrienes released from stimulated cells. DMSO-differentiated HL-60 cells release leukotriene B4 upon stimulation with the calcium ionophore A23187. Leukotriene B4 released into the cell medium can be quantitated by radioimmunoassay using commercially available diagnostic kits (New England Nuclear, Boston, MA). Leukotriene B4 production can be detected in HL-60 cells 48 hours following addition of DMSO to differentiate the cells into a neutrophil-like cell. Cells (2 x 10^5 cells/mL) will be treated with increasing concentrations of the macromolecule for 48-72 hours in the

presence of 1.3% DMSO. The cells are washed and resuspended at a concentration of 2 x 10^6 cell/mL in Dulbecco's phosphate buffered saline containing 1% delipidated bovine serum albumin. Cells are stimulated with 10 μM calcium ionophore A23187 for 15 min and the quantity of LTB4 produced from 5 x 10^5 cell determined by radioimmunoassay as described by the manufacturer.

Using this assay the following results would likely be obtained with an oligonucleotide directed to the 5-LO mRNA. Cells will be treated for 72 hours with either 1 μM, 10 μM or 30 μM of the macromolecule in the presence of 1.3% DMSO. The quantity of LTB_4 produced from 5 x 10^5 cells would be expected to be about 75 pg, 50 pg, and 35 pg, respectively with untreated differentiated cells producing 75 pg LTB_4.

E. In Vivo Assay

Inhibition of the production of 5-lipoxygenase in the mouse can be demonstrated in accordance with the following protocol. Topical application of arachidonic acid results in the rapid production of leukotriene B_4, leukotriene C_4 and prostaglandin E_2 in the skin followed by edema and cellular infiltration. Certain inhibitors of 5-lipoxygenase have been known to exhibit activity in this assay. For the assay, 2 mg of arachidonic acid is applied to a mouse ear with the contralateral ear serving as a control. The polymorphonuclear cell infiltrate is assayed by myeloperoxidase activity in homogenates taken from a biopsy 1 hour following the administration of arachidonic acid. The edematous response is quantitated by measurement of ear thickness and wet weight of a punch biopsy. Measurement of leukotriene B_4 produced in biopsy specimens is performed as a direct measurement of 5-lipoxygenase activity in the tissue. Oligonucleotides will be applied topically to both ears 12 to 24 hours prior to

administration of arachidonic acid to allow optimal activity of the compounds. Both ears are pretreated for 24 hours with either 0.1 µmol, 0.3 µmol, or 1.0 µmol of the macromolecule prior to challenge with arachidonic acid. Values are expressed as the mean for three animals per concentration. Inhibition of polymorphonuclear cell infiltration for 0.1 µmol, 0.3 µmol, and 1 µmol is expected to be about 10%, 75% and 92% of control activity, respectively. Inhibition of edema is expected to be about 3%, 58% and 90%, respectively while inhibition of leukotriene B_4 production would be expected to be about 15%, 79% and 99%, respectively.

EXAMPLE 63

5'-*O*-DMT-2'-deoxy-2'-methylene-5-methyl uridine-3'-(2-cyanoethyl-N,N-diisoproppyl) phosphoramidite

2'-Deoxy-2'-methylene-3',5'-*O*-(tetraisopropyl disiloxane-1,3,diyl)-5-methyl uridine is synthesized following the procedures reported for the corresponding uridine derivative (Hansske, F.; Madej, D.; Robins, M.J. *Tetrahedron* (1984) **40**, 125; Matsuda, A.; Takenusi, K.; Tanaka, S.; Sasaki, T.; Ueda, T. *J. Med. Chem.* (1991) **34**, 812; See also Cory, A. H.; Samano, V.; Robins, M. J.; Cory, J. G. 2'-Deoxy-2'-methylene derivatives of adenosine, guanosine, tubercidin, cytidine and uridine as inhibitors of L1210 cell growth in culture. Biochem. Pharmacol. (1994), 47(2), 365-71.)

It is treated with IM TBAF in THF to give 2'-deoxy-2'-methylene-5-methyl uridine. It is dissolved in pyridine and treated with DMT-Cl and stirred to give the 5'-*O*-DMT-2'-deoxy-2'-methylene-5-methyl uridine. This compound is treated with 2-cyanoethyl-N,N-diisopropyl phosphoramidite and diisopropylaminotetrazolide. In a similar manner the corresponding N-6 benzoyl adenosine, N-

4-benzoyl cytosine, N-2-isobutyryl guanosine phosphoramidite derivatives are synthesized.

EXAMPLE 63

Synthesis of 3′-*O*-4′-C-methyleneribonucleoside

5′-*O*-DMT-3′-*O*-4′-*C*-methylene uridine and 5-methyl uridine are synthesized and phosphitylated according to the procedure of Obika *et al.* (Obika *et al. Bioorg. Med. Chem. Lett.* (1999) **9**, 515-158). The amidites are incorporated into oligonucleotides using the protocols described above.

EXAMPLE 64

Synthesis of 2'-methylene phosphoramidites

5′-*O*-DMT-2′-(methyl)-3′-*O*-(2-cyanoethyl-N,N-diisopropylamine)-5-methyluridine-phosphoramidite, 5′-*O*-DMT-2′-(methyl)-N-6-benzoyl adenosine (3′-*O*-2-cyanoethyl-N,N-diisopropylamino) phosphoramidite, 5′-*O*-DMT-2′-(methyl)-N2-isoburytyl guanosine-3′-*O*-(2-cyanoethyl-N,N-diisopropylamino) phosphoramidite and 5′-*O*-DMT-2′-(methyl)-N-4-benzoyl cytidine-3′-*O*-(2-cyanoethyl-N,N-diisopropylamino) phosphoramidites were obtained by the phosphitylation of the corresponding nucleosides. The nucleosides were synthesized according to the procedure described by Iribarren, Adolfo M.; Cicero, Daniel O.; Neuner, Philippe J. Resistance to degradation by nucleases of (2'S)-2'-deoxy-2'-C-methyloligonucleotides, novel potential antisense probes. Antisense Res. Dev., (1994), 4(2), 95-8; Schmit, Chantal; Bevierre, Marc-Olivier; De Mesmaeker, Alain; Altmann, Karl-Heinz. "The effects of 2'- and 3'-alkyl substituents on oligonucleotide hybridization and stability". Bioorg. Med. Chem. Lett. (1994), 4(16), 1969-74.

The phosphitylation is carried out by using the bisamidite procedure.

EXAMPLE 65

Synthesis of 2'-S-methyl phosphoramidites

5'-*O*-DMT-2'-S-(methyl)-3'-*O*-(2-cyanoethyl-N,N-diisopropylamine)-5-methyl uridine-phosphoramidite, 5'-*O*-DMT-2'-S(methyl)-N-6-benzoyl adenosine (3'-*O*-2-cyanoethyl-N,N-diisopropylamino) phosphoramidite, 5'-*O*-DMT-2'-S-(methyl)-N2-isoburytyl guanosine-3'-*O*-(2-cyanoethyl-N,N-diisopropylamino) phosphoramidite and 5'-*O*-DMT-2'-S-(methyl)-N-4-benzoyl cytidine-3'-*O*-(2-cyanoethyl-N,N-diisopropylamino) phosphoramidites were obtained by the phosphitylation of the corresponding nucleosides. The nucleosides were synthesized according to the procedure described by Fraser *et al.* (Fraser, A.; Wheeler, P.; Cook, P.D.; Sanghvi, Y.S. *J. Heterocycl. Chem.* (1993) **31**, 1277-1287). The phosphitylation is carried out by using the bisamidite procedure.

EXAMPLE 66

Synthesis of 2'-*O*-methyl-β-D-arabinofuranosyl compounds

2'-*O*-Methyl-β-D-arabinofuranosyl-thymidine containing oligonucleotides were synthesized following the procedures of Gotfredson *et. al.* (Gotfredson, C.H. *et. al.* *Tetrahedron Lett.* (1994) **35**, 6941-6944; Gotfredson, C.H. *et. al. Bioorg. Med. Chem.* (1996) **4**, 1217-1225). 5'-*O*-DMT-2'-ara-(O-methyl)-3'-*O*-(2-cyanoethyl-N,N-diisopropylamine)-5-methyl uridine-phosphoramidite, 5'-*O*-DMT-2'-ara-(O-methyl)-N-6-benzoyl adenosine (3'-*O*-2-cyanoethyl-N,N-diisopropylamino) phosphoramidite, 5'-*O*-DMT-2'-ara-(O-methyl)-N2-isoburytyl guanosine-3'-*O*-(2-cyanoethyl-N,N-diisopropylamino) phosphoramidite and 5'-*O*-DMT-2'-ara-(O-

methyl)-N-4-benzoyl cytidine-3'-*O*-(2-cyanoethyl-N,N-diisopropylamino) phosphoramidites are obtained by the phosphitylation of the corresponding nucleosides. The nucleosides are synthesized according to the procedure described by Gotfredson, C.H. *et. al. Tetrahedron Lett.* (1994) **35**, 6941-6944; Gotfredson, C.H. *et. al. Bioorg. Med. Chem.* (1996) **4**, 1217-1225. The phosphitylation is carried out by using the bisamidite procedure.

EXAMPLE 67

Synthesis of 2'-fluoro-β-D-arabinofuranosyl compounds

2'-Fluoro-β-D-arabinofuranosyl oligonucleotides are synthesized following the procedures by Kois,P. et al., Nucleosides Nucleotides 12, 1093,1993 and Damha et al., J. Am. Chem. Soc., 120, 12976,1998 and references sited therin. 5'-*O*-DMT-2'-ara-(fluoro)-3'-*O*-(2-cyanoethyl-N,N-diisopropylamine)-5-methyl uridine-phosphoramidite, 5'-*O*-DMT-2'-ara-(fluoro)-N-6-benzoyl adenosine (3'-*O*-2-cyanoethyl-N,N-diisopropylamino) phosphoramidite, 5'-*O*-DMT-2'-ara-(fluoro)-N2-isoburytyl guanosine-3'-*O*-(2-cyanoethyl-N,N-diisopropylamino) phosphoramidite and 5'-*O*-DMT-2'-ara-(fluoro)-N-4-benzoyl cytidine-3'-*O*-(2-cyanoethyl-N,N-diisopropylamino) phosphoramidites are obtained by the phosphitylation of the corresponding nucleosides. The nucleosides are synthesized according to the procedure described by Kois,P. et al., Nucleosides Nucleotides 12, 1093,1993 and Damha et al., J. Am. Chem. Soc., 120, 12976,1998. The phosphitylation is carried out by using the bisamidite procedure.

EXAMPLE 68

Synthesis of 2'-hydroxyl-β-D-arabinofuranosyl compounds

2'-Hydroxyl-β-D-arabinofuranosyl oligonucleotides are synthesized following the procedures by Resmini and Pfleiderer Helv. Chim. Acta, 76, 158,1993; Schmit et al., *Bioorg. Med. Chem. Lett.* 4, 1969, 1994 Resmini, M.; Pfleiderer, W. Synthesis of arabinonucleic acid (tANA). Bioorg. Med. Chem. Lett. (1994), 16, 1910.; Resmini, Matthias; Pfleiderer, W. Nucleosides. Part LV. Efficient synthesis of arabinoguanosine building blocks (Helv. Chim. Acta, (1994), 77, 429-34; and Damha et al., J. Am. Chem. Soc., 1998, 120, 12976, and references sited therin).

5'-*O*-DMT-2'-ara-(hydroxy)-3'-*O*-(2-cyanoethyl-N,N-diisopropylamine)-5-methyl uridine-phosphoramidite, 5'-*O*-DMT-2'-ara-(hydroxy)-N-6-benzoyl adenosine (3'-*O*-2-cyanoethyl-N,N-diisopropylamino) phosphoramidite, 5'-*O*-DMT-2'-ara-(hydroxy)-N2-isoburytyl guanosine-3'-*O*-(2-cyanoethyl-N,N-diisopropylamino) phosphoramidite and 5'-*O*-DMT-2'-ara-(hydroxy)-N-4-benzoyl cytidine-3'-*O*-(2-cyanoethyl-N,N-diisopropylamino) phosphoramidites are obtained by the phosphitylation of the corresponding nucleosides. The nucleosides are synthesized according to the procedure described by Kois,P. et al., Nucleosides Nucleotides 12, 1093,1993 and Damha et al., J. Am. Chem. Soc., 120, 12976,1998. The phosphitylation is carried out by using the bisamidite procedure.

EXAMPLE 69

Synthesis of difluoromethylene compounds

5'-*O*-DMT-2'-deoxy-2'-difluoromethylene-5-methyluridine-3'-(2-cyanoethyl-N,N-diisopropyl phosphoramidite), 5'-*O*-DMT-2'-deoxy-2'-difluoromethylene-N-4-benzoyl-cytidine, 5'-*O*-DMT-2'-deoxy-2'-diflyoromethylene-N-6-benzoyl adenosine, and 5'-*O*-DMT-2'-deoxy-2'-

difluoroethylene-N_2-isobutyryl guanosine are synthesized following the protocols described by Usman *et. al.* (U.S. Patent 5639649, June 17, 1997).

EXAMPLE 70

Synthesis of 5′-*O*-DMT-2′-deoxy-2'-β-(O-acetyl)-2'-α-methyl-N6- benzoyl-adenosine-3'-(2-cyanoethyl-N,N-diisopropyl phosphoramidite

5′-*O*-DMT-2′-deoxy-2'-β-(OH)-2'-α-methyl-adenosine is synthesized from the compound 5′-3'-protected-2′-keto-adenosine (Rosenthal, Sprinzl and Baker, Tetrahedron Lett. 4233, 1970; see also Nucleic acid related compounds. A convenient procedure for the synthesis of 2'- and 3'-ketonucleosides is shown Hansske *et al.*, Dep. Chem., Univ. Alberta, Edmonton, Can., Tetrahedron Lett. (1983), 24(15), 1589-92.) by Grigand addition of MeMgI in THF solvent, followed by seperation of the isomers. The 2-β-(OH) is protected as acetate. 5′-3'-acetal group is removed, 5'-position dimethoxy, tritylated, N-6 position is benzoylated and then 3'-position is phosphitylated to give 5′-*O*-DMT-2′-deoxy-2'-β-(O-acetyl)-2'-α-methyl-N6-benzoyl-adenosine-3'-(2-cyanoethyl-N,N-diisopropyl)phosphoramidite.

EXAMPLE 71

Synthesis of 5′-*O*-DMT-2′-α-ethynyl-N6- benzoyl-adenosine-3'-(2-cyanoethyl-N,N-diisopropyl phosphoramidite

5′-*O*-DMT-2′-deoxy-2'-β-(OH)-2'-α-ethynyl-adenosine is synthesized from the compound 5′-3'-protected-2′-keto-adenosine (Rosenthal, Sprinzl and Baker, Tetrahedron Lett. 4233, 1970) by Grigand addition of Ethynyl-MgI in THF solvent, followed by seperation of the isomers. The 2'-β-(OH) is removed by Robins' deoxygenation procedure (Robins *et al.*, J. Am. Chem. Soc. (1983), 105, 4059-65. 5′-3'-

acetal group is removed, 5'-position dimethoxytritylated, N-6 position is benzoylated and then 3'-position is phosphitylated to give the title compound.

EXAMPLE 72

2'-O-(guaiacolyl)-5-methyluridine

2-Methoxyphenol (6.2 g, 50 mmol) was slowly added to a solution of borane in tetrahydrofuran (1 M, 10 mL, 10 mmol) with stirring in a 100 mL bomb. Hydrogen gas evolved as the solid dissolved O-2,2'-anhydro-5-methyluridine (1.2g, 5 mmol), and sodium bicarbonate (2.5 mg) were added and the bomb was sealed, placed in an oil bath and heated to 155 °C for 36 hours. The bomb was cooled to room temperature and opened. The crude solution was concentrated and the residue partitioned between water (200 mL) and hexanes (200 mL). The excess phenol was extracted into hexanes. The aqueous layer was extracted with ethyl acetate (3x200 mL) and the combined organic layer was washed once with water, dried over anhydrous sodium sulfate and concentrated. The residue was purified by silica gel flash column chromatography using methanol:methylene chloride (1/10, v/v) as the eluent. Fractions were collected and the target fractions were concentrated to give 490 mg of pure product as a white solid. Rf= 0.545 in CH_2Cl_2/CH_3OH (10:1). MS/ES for $C_{17}H_20N_2O_7$, 364.4; Observed 364.9.

EXAMPLE 73

5'-Dimethoxytrityl-2'-*O*-(2-methoxyphenyl)-5-methyluridine-3'-*O*-(2-cyanoethyl-N,N-diisopropylamino) phosphoramidite

2'-*O*-(guaiacolyl)-5-methyl-uridine is treated with 1.2 equivalents of dimethoxytrityl chloride (DMT-Cl) in pyridine to yield the 5'-*O*-dimethoxy tritylated nucleoside. After evaporation of the pyridine and work up

(CH_2Cl_2/saturated $NaHCO_3$ solution) the compound is purified in a silica gel column. The 5'-protected nucleoside is dissolved in anhydrous methylene chloride and under argon atmosphere, N,N-diisopropylaminohydrotetrazolide (0.5 equivalents) and bis-N,N-diisopropylamino-2-cyanoethyl-phosphoramidite (2 equivalents) are added via syringe over 1 min. The reaction mixture is stirred under argon at room temperature for 16 hours and then applied to a silica column. Elution with hexane:ethylacetate (25:75) gives the title compound.

EXAMPLE 74

5'-Dimethoxytrityl-2'-*O*-(2-methoxyphenyl)-5-methyluridine-3'-*O*-succinate

The 5'-protected nucleoside from Example 73 is treated with 2 equivalents of succinic anhydride and 0.2 equivalents of 4-N,N-dimethylaminopyridine in pyridine. After 2 hours the pyridine is evaporated, the residue is dissolved in CH_2Cl_2 and washed three times with 100 mL of 10% citric acid solution. The organic layer is dried over anhydrous $MgSO_4$ to give the desired succinate. The succinate is then attached to controlled pore glass (CPG) using established procedures (Pon, R.T., Solid phase supports for oligonucleotide synthesis, in *Protocols for Oligonucleotides and Analogs*, S. Agrawal (Ed.), Humana Press: Totawa, NJ, **1993**, 465-496).

EXAMPLE 75

5'-Dimethoxytrityl-2'-*O*-(*trans*-2-methoxycyclohexyl)-5-methyl uridine

2'-3'-*O*-Dibutylstannyl-5-methyl uridine (Wagner *et al.*, *J. Org. Chem.*, **1974**, *39*, 24) is alkylated with *trans*-2-methoxycyclohexyl tosylate at 70 °C in DMF. A 1:1

mixture of 2'-*O*- and 3'-*O*-(*trans*-2-methoxycyclohexyl)-5-methyluridine is obtained in this reaction. After evaporation of the DMF solvent, the crude mixture is dissolved in pyridine and treated with dimethoxytrityl-chloride (DMT-Cl) (1.5 equivalents). The resultant mixture is purified by silica gel flash column chromatography to give the title compound.

EXAMPLE 76

5'-Dimethoxytrityl-2'-*O*-(*trans*-2-methoxycyclohexyl)-5-methyluridine-3'-*O*-(2-cyanoethyl-N,N-diisopropylamino) phosphoramidite

5'-Dimethoxytrityl-2'-*O*-(*trans*-2-methoxycyclohexyl)-5-methyl uridine is phosphitylated according to the procedure described above to give the required phosphoramidite.

EXAMPLE 77

5'-Dimethoxytrityl-2'-*O*-(*trans*-2-methoxycyclohexyl)-5-methyluridine-3'-*O*-(succinyl-amino) CPG

5'-Dimethoxytrityl-2'-*O*-(*trans*-2-methoxycyclohexyl)-5-methyl uridine is succinylated and attached to controlled pore glass to give the solid support bound nucleoside.

EXAMPLE 78

***trans*-2-ureido-cyclohexanol**

Trans-2-amino-cyclohexanol (Aldrich) is treated with triphosgene in methylene chloride (1/3 equivalent). To the resulting solution, excess ammonium hydroxide is added to give after work up the title compound.

EXAMPLE 79

2'-*O*-(*trans*-2-uriedo-cyclohexyl)-5-methyl uridine

Trans-2-uriedo-cyclohexanol (50 mmol) is added to a solution of borane in tetrahydrofuran (1 M, 10 mL, 10 mmol) while stirring in a 10 mL bomb. Hydrogen gas evolves as the reactant dissolves. O2,2'-Anhydro-5-methyluridine (5 mmol) and sodium bicarbonate (2.5 mg) are added to the bomb and sealed. Then it is heated to 140 for 72 hrs. The bomb is cooled to room temperature and opened. The crude material was worked up as illustrated above followed by purification by silica gel flash column chromatography to give the title compound.

EXAMPLE 80

5'-*O*-(Dimethoxytrityl)-2'-*O*-(*trans*-2-uriedo-cyclohexyl) 3'-*O*-(2-cyanoethyl, N,N,-diisopropyl) uridine phosphoramidite

2'-*O*-(*trans*-2-uriedo-cyclohexyl)-5-methyl uridine tritylated at the 5'-OH and phosphitylated at the 3'-OH following the procedures illustrated in example 2 to give the title compound.

EXAMPLE 81

5'-*O*-dimethoxytrityl-2'-*O*-(*trans*-2-uriedo-cyclohexyl)-5-methyl-3'-*O*-(succinyl)-amino CPG uridine

5'-*O*-dimethoxytrityl-2'-*O*-(*trans*-2-uriedo-cyclohexyl)-5-methyl uridine is succinylated and attached to CPG as illustrated above.

EXAMPLE 82

2'-*O*-(*trans*-2-methoxy-cyclohexyl) adenosine

Trans-2-methoxycyclopentanol, *trans*-2-methoxy-cylcohexanol, *trans*-2-methoxy-cyclopentyl tosylate and *trans*-2-methoxy-cyclohexyl tosylate are prepared according

to reported procedures (Roberts, D.D., Hendrickson, W., *J. Org. Chem.*, **1967**, *34*, 2415-2417; *J. Org. Chem.*, **1997**, *62*, 1857-1859). A solution of adenosine (42.74 g, 0.16 mol) in dry dimethylformamide (800 mL) at 5 °C is treated with sodium hydride (8.24 g, 60% in oil prewashed thrice with hexanes, 0.21 mol). After stirring for 30 min, *trans*-2-methoxycyclohexyl tosylate (0.16 mol) is added over 20 minutes at 5 °C. The reaction is stirred at room temperature for 48 hours, then filtered through Celite. The filtrate is concentrated under reduced pressure followed by coevaporation with toluene (2x100 mL) to give the title compound.

EXAMPLE 83

N^6-Benzoyl-2'-*O*-(*trans*-2-methoxycyclohexyl) adenosine

A solution of 2'-*O*-(*trans*-2-methoxy-cyclohexyl) adenosine (0.056 mol) in pyridine (100 mL) is evaporated under reduced pressure to dryness. The residue is redissolved in pyridine (560 mL) and cooled in an ice water bath. Trimethylsilyl chloride (36.4 mL, 0.291 mol) is added and the reaction is stirred at 5 °C for 30 minutes. Benzoyl chloride (33.6 mL, 0.291 mol) is added and the reaction is allowed to warm to 25 °C for 2 hours and then cooled to 5 °C. The reaction is diluted with cold water (112 mL) and after stirring for 15 min, concentrated ammonium hydroxide (112 mL). After 30 min, the reaction is concentrated under reduced pressure (below 30 °C) followed by coevaporation with toluene (2x100 mL). The residue is dissolved in ethyl acetate-methanol (400 mL, 9:1) and the undesired silyl by-products are removed by filtration. The filtrate is concentrated under reduced pressure and purified by silica gel flash column chromatography (800 g, chloroform-methanol 9:1). Selected fractions are combined,

concentrated under reduced pressure and dried at 25 °C/0.2 mmHg for 2 h to give the title compound.

EXAMPLE 84

N^6-Benzoyl-5'-O-(dimethoxytrityl)-2'-*O*-(*trans*-2-methoxycyclohexyl) adenosine

A solution of N^6-benzoyl-2'-*O*-(*trans*-2-methoxycyclohexyl) adenosine (0.285 mol) in pyridine (100 mL) is evaporated under reduced pressure to an oil. The residue is redissolved in dry pyridine (300 mL) and 4,4'-dimethoxytriphenylmethyl chloride (DMT-Cl, 10.9 g, 95%, 0.31 mol) added. The mixture is stirred at 25 °C for 16 h and then poured onto a solution of sodium bicarbonate (20 g) in ice water (500 mL). The product is extracted with ethyl acetate (2x150 mL). The organic layer is washed with brine (50 mL), dried over sodium sulfate (powdered) and evaporated under reduced pressure (below 40 °C). The residue is chromatographed on silica gel (400 g, ethyl acetate-hexane 1:1. Selected fractions were combined, concentrated under reduced pressure and dried at 25 °C/0.2 mmHg to give the title compound.

EXAMPLE 85

[N^6-Benzoyl-5'-*O*-(4,4'-dimethoxytrityl)-2'-*O*-(*trans*-2-methoxycyclohexyl) adenosine-3'-O-yl]-N,N-diisopropylamino-cyanoethoxy phosphoramidite

Phosphitylation of N^6-benzoyl-5'-O-(dimethoxytrityl)-2'-*O*-(*trans*-2-methoxycyclohexyl) adenosine was performed as illustrated above to give the title compound.

EXAMPLE 86

General procedures for chimeric C3'-endo and C2'-endo modified oligonucleotide synthesis

Oligonucleotides are synthesized on a PerSeptive Biosystems Expedite 8901 Nucleic Acid Synthesis System. Multiple 1-mmol syntheses are performed for each oligonucleotide. The 3'-end nucleoside containing solid support is loaded into the column. Trityl groups are removed with trichloroacetic acid (975 mL over one minute) followed by an acetonitrile wash. The oligonucleotide is built using a modified diester (P=O) or thioate (P=S) protocol.

Phosphodiester protocol

All standard amidites (0.1 M) are coupled over a 1.5 minute time frame, delivering 105 µL material. All novel amidites are dissolved in dry acetonitrile (100 mg of amidite/1 mL acetonitrile) to give approximately 0.08-0.1 M solutions. The 2'-modified amidites (both ribo and arabino monomers) are double coupled using 210 µL over a total of 5 minutes. Total coupling time is approximately 5 minutes (210 mL of amidite delivered). 1-H-tetrazole in acetonitrile is used as the activating agent. Excess amidite is washed away with acetonitrile. (1S)-(+)-(10-camphorsulfonyl) oxaziridine (CSO, 1.0 g CSO/8.72 mL dry acetonitrile) is used to oxidize (3 minute wait step) delivering approximately 375 µL of oxidizer. Standard amidites are delivered (210 µL) over a 3-minute period.

Phosphorothioate protocol

The 2'-modified amidite is double coupled using 210 µL over a total of 5 minutes. The amount of oxidizer, 3H-1,2-benzodithiole-3-one-1,1-dioxide (Beaucage reagent, 3.4 g Beaucage reagent/200 mL acetonitrile), is 225 µL (one minute wait step). The unreacted nucleoside is capped with

a 50:50 mixture of tetrahydrofuran/acetic anhydride and tetrahydrofuran/pyridine/1-methyl imidazole. Trityl yields are followed by the trityl monitor during the duration of the synthesis. The final DMT group is left intact. After the synthesis, the contents of the synthesis cartridge (1 mmole) is transferred to a Pyrex vial and the oligonucleotide is cleaved from the controlled pore glass (CPG) using 30% ammonium hydroxide (NH_4OH, 5 mL) for approximately 16 hours at 55 °C.

Oligonucleotide Purification

After the deprotection step, the samples are filtered from CPG using Gelman 0.45 µm nylon acrodisc syringe filters. Excess NH_4OH is evaporated away in a Savant AS160 automatic speed vac. The crude yield is measured on a Hewlett Packard 8452A Diode Array Spectrophotometer at 260 nm. Crude samples are then analyzed by mass spectrometry (MS) on a Hewlett Packard electrospray mass spectrometer. Trityl-on oligonucleotides are purified by reverse phase preparative high performance liquid chromatography (HPLC). HPLC conditions are as follows: Waters 600E with 991 detector; Waters Delta Pak C4 column (7.8X300mm); Solvent A: 50 mM triethylammonium acetate (TEA-Ac), pH 7.0; Solvent B: 100% acetonitrile; 2.5 mL/min flow rate; Gradient: 5% B for first five minutes with linear increase in B to 60% during the next 55 minutes. Fractions containing the desired product/s (retention time = 41 minutes for DMT-ON-16314; retention time = 42.5 minutes for DMT-ON-16315) are collected and the solvent is dried off in the speed vac. Oligonucleotides are detritylated in 80% acetic acid for approximately 60 minutes and lyophilized again. Free trityl and excess salt are removed by passing detritylated oligonucleotides through Sephadex G-25 (size exclusion chromatography) and collecting appropriate samples through a Pharmacia fraction

collector. The solvent is again evaporated away in a speed vac. Purified oligonucleotides are then analyzed for purity by CGE, HPLC (flow rate: 1.5 mL/min; Waters Delta Pak C4 column, 3.9X300mm), and MS. The final yield is determined by spectrophotometer at 260 nm.

Example 87

Rapid amplification of 5'-cDNA end (5' -RACE) and 3'-cDNA end (3'-RACE)

An internet search of the XREF database in the National Center of Biotechnology Information (NCBI) yielded a 361 base pair (bp) human expressed sequenced tag (EST, GenBank accession #H28861), homologous to yeast RNase H (RNH1) protein sequenced tag (EST, GenBank accession #Q04740) and its chicken homologue (accession #D26340). Three sets of oligonucleotide primers encoding the human RNase H EST sequence were synthesized. The sense primers were ACGCTGGCCGGGAGTCGAAATGCTTC (H1: SEQ ID NO: 6), CTGTTCCTGGCCCACAGAGTCGCCTTGG (H3: SEQ ID NO: 7) and GGTCTTTCTGACCTGGAATGAGTGCAGAG (H5: SEQ ID NO: 8). The antisense primers were CTTGCCTGGTTTCGCCCTCCGATTCTTGT (H2: SEQ ID NO: 9), TTGATTTTCATGCCCTTCTGAAACTTCCG (H4; SEQ ID NO: 10) and CCTCATCCTCTATGGCAAACTTCTTAAATCTGGC (H6; SEQ ID NO: 11). The human RNase H 3' and 5' cDNAs derived from the EST sequence were amplified by polymerase chain reaction (PCR), using human liver or leukemia (lymphoblastic Molt-4) cell line Marathon ready cDNA as templates, H1 or H3/AP1 as well as H4 or H6/AP2 as primers (Clontech, Palo Alto, CA). The fragments were subjected to agarose gel electrophoresis and transferred to nitrocellulose membrane (Bio-Rad, Hercules CA) for confirmation by Southern blot, using ^{32}P-labeled H2 and H1 probes (for 3' and 5' RACE products, respectively, in accordance with procedures described by Ausubel et al.,

Current Protocols in Molecular Biology, Wiley and Sons, New York, NY., 1988. The confirmed fragments were excised from the agarose gel and purified by gel extraction (Qiagen, Germany), then subcloned into Zero-blunt vector (Invitrogen, Carlsbad, CA) and subjected to DNA sequencing.

EXAMPLE 88

Screening of the cDNA library, DNA sequencing and sequence analysis

A human liver cDNA lambda phage Uni-ZAP library (Stratagene, La Jolla, CA) was screened using the RACE products as specific probes. The positive cDNA clones were excised into the pBluescript phagemid (Stratagene, La Jolla CA) from lambda phage and subjected to DNA sequencing with an automatic DNA sequencer (Applied Biosystems, Foster City, CA) by Retrogen Inc. (San Diego, CA). The overlapping sequences were aligned and combined by the assembling programs of MacDNASIS v3.0 (Hitachi Software Engineering America, South San Francisco, CA). Protein structure and subsequence analysis were performed by the program of MacVector 6.0 (Oxford Molecular Group Inc., Campbell, CA). A homology search was performed on the NCBI database by internet E-mail.

EXAMPLE 89

Northern blot and Southern blot analysis

Total RNA from different human cell lines (ATCC, Rockville, MD) was prepared and subjected to formaldehyde agarose gel electrophoresis in accordance with procedures described by Ausubel et al., Current Protocols in Molecular Biology, Wiley and Sons, New York, NY, 1988, and transferred to nitrocellulose membrane (Bio-Rad, Hercules CA). Northern blot hybridization was carried out in

QuickHyb buffer (Stratagene, La Jolla, CA) with ^{32}P- labeled probe of full length RNase H cDNA clone or primer H1/H2 PCR-generated 322-base N-terminal RNase H cDNA fragment at 68°C for 2 hours. The membranes were washed twice with 0.1% SSC/0.1% SDS for 30 minutes and subjected to auto-radiography. Southern blot analysis was carried out in 1X pre-hybridization/hybridization buffer (BRL, Gaithersburg, MD) with a ^{32}P-labeled 430 bp C-terminal restriction enzyme PstI/PvuII fragment or 1.7 kb full length cDNA probe at 60°C for 18 hours. The membranes were washed twice with 0.1% SSC/0.1% SDS at 60°C for 30 minutes, and subjected to autoradiography.

EXAMPLE 90

Expression and purification of the cloned RNase protein

The cDNA fragment coding the full RNase H protein sequence was amplified by PCR using 2 primers, one of which contains restriction enzyme NdeI site adapter and six histidine (his-tag) codons and 22 bp protein N terminal coding sequence. The fragment was cloned into expression vector pET17b (Novagen, Madison, WI) and confirmed by DNA sequencing. The plasmid was transfected into E.coli BL21(DE3) (Novagen, Madison, WI). The bacteria were grown in M9ZB medium at 32°C and harvested when the OD_{600} of the culture reached 0.8, in accordance with procedures described by Ausubel et al., *Current Protocols in Molecular Biology*, Wiley and Sons, New York, NY, 1988. Cells were lysed in 8M urea solution and recombinant protein was partially purified with Ni-NTA agarose (Qiagen, Germany). Further purification was performed with C4 reverse phase chromatography (Beckman, System Gold, Fullerton, CA) with 0.1% TFA water and 0.1% TFA acetonitrile gradient of 0% to 80% in 40 minutes as described by Deutscher, M. P., *Guide to Protein*

Purification, Methods in Enzymology 182, Academic Press, New York, NY, 1990. The recombinant proteins and control samples were collected, lyophilized and subjected to 12% SDS-PAGE as described by Ausubel et al. (1988) Current Protocols in Molecular Biology, Wiley and Sons, New York, NY. The purified protein and control samples were resuspended in 6 M urea solution containing 20 mM Tris HCl, pH 7.4, 400 mM NaCl, 20% glycerol, 0.2 mM PMSF, 5 mM DTT, 10 µg/ml aprotinin and leupeptin, and refolded by dialysis with decreasing urea concentration from 6 M to 0.5 M as well as DTT concentration from 5 mM to 0.5 mM as described by Deutscher, M. P., Guide to Protein Purification, Methods in Enzymology 182, Academic Press, New York, NY, 1990. The refolded proteins were concentrated (10 fold) by Centricon (Amicon, Danvers, MA) and subjected to RNase H activity assay.

EXAMPLE 91

RNase H activity assay

^{32}P-end-labeled 17-mer RNA, GGGCGCCGTCGGTGTGG (SEQ ID NO: 12) described by Lima, W. F. and Crooke, S. T., *Biochemistry*, 1997 36, 390-398, was gel-purified as described by Ausubel et al., *Current Protocols in Molecular Biology*, Wiley and Sons, New York, NY, 1988 and annealed with a tenfold excess of its complementary 17-mer oligodeoxynucleotide or a 5-base DNA gapmer, i.e., a 17mer oligonucleotide which has a central portion of 5 deoxynucleotides (the "gap") flanked on both sides by 6 2'-methoxynucleotides. Annealing was done in 10 mM Tris HCl, pH 8.0, 10 mM MgCl, 50 mM KCl and 0.1 mM DTT to form one of three different substrates: (a) single strand (ss) RNA probe, (b) full RNA/DNA duplex and (c) RNA/DNA gapmer duplex. Each of these substrates was incubated with protein samples at 37°C for 5 minutes to 2 hours at the

same conditions used in the annealing procedure and the reactions were terminated by adding EDTA in accordance with procedures described by Lima, W. F. and Crooke, S. T., *Biochemistry*, 1997, 36, 390-398. The reaction mixtures were precipitated with TCA centrifugation and the supernatant was measured by liquid scintillation counting (Beckman LS6000IC, Fullerton, CA). An aliquot of the reaction mixture was also subjected to denaturing (8 M urea) acrylamide gel electrophoresis in accordance with procedures described by Lima, W. F. and Crooke, S. T., *Biochemistry*, 1997, 36, 390-398 and Ausubel et al., *Current Protocols in Molecular Biology*, Wiley and Sons, New York, NY, 1988.

EXAMPLE 92

Effects of phosphorothioate substitution and substrate length on digestion by Human RNase H1 (See Table 4)

Oligoribonucleotides were preannealed with the complementary antisense oligodeoxynucleotide at 10 nM and 20 nM respectively and subjected to digestion by Human RNase H1. The 17mer (RNA no.1) and 25 mer (RNA no.3) RNA sequences are derived from Harvy-RAS oncogen 51) and the 25 mer RNA contains the 17 mer sequence. The 20 mer (RNA no.2) sequence is derived from human hepatitis C virus core protein coding sequence (52). The initial rates were determined as described in Materials and Methods, 1A: Comparison of the initial rates of cleavage of an RNA:phosphodiester (P=O) and an RNA:phosphorothioate (P=S) duplexes, and 1B: Comparison among duplexes of different sequences and lengths.

EXAMPLE 93

Effects of 2′-substitution and deoxy-gap size on digestion rates by human RNase H1 (See Table 5)

Substrate duplexes were hybridized and initial rates were determined as shown in Table 4 and described in Material and Methods. The 17mer RNA is the same used in Table 4, and the 20mer RNA (UGGUGGGCAAUGGGCGUGUU (SEQ ID NO: 20, RNA no.4) is derived from the protein kinase C-zeta (53) sequence. The 17mer and 20mer P=S oligonucleotides were full deoxyphosphorothioate containing no 2′-modifications. The 9, 7, 5, 4 and 3 deoxy gap oligonucleotides were 17mer oligonucleotide with a central portion consisting of nine, seven, five and four deoxynucleotides flanked on both sides by 2′-methoexynucleotides (also see Figure 2). Boxed sequences indicate the position of the 2′-methoxy-modified residues. Dash-boxed sequence indicates the position of the 2′-propoxy-modified residues.

EXAMPLE 94

Kinetic constants for RNase H1 cleavage of RNA:DNA duplexes (See Table 6)

The RNA:DNA duplexes in Table 4 were used to determine *Km* and *Vmax* of Human and *E.coli* RNase H1 as described in the Materials and Methods section.

EXAMPLE 95

Binding constants and specificity of RNase H's (See Table 7)

K_d's were determined as described in Materials and methods. The *K_d's* for *E. coli* RNase H1 was derived from previously reported data (21). The competing substrates (competitive inhibitors) used in the binding study are

divided into two categories: single-strand (ss) oligonucleotides and oligonucleotide duplexes all with the 17mer sequence as in Table 4 (RNA No. 1). The single-strand oligonucleotides included: ssRNA, ssDNA, ss fully modified 2'-methoxy phosphodiester oligonucleotide (ss 2'-O-Me) and ss full phosphorothioate deoxyoligonucleotide (ss DNA, P=S). The duplex substrates include: DNA:DNA duplex, RNA:RNA duplex, DNA:fully modified 2' fluoro or fully modified 2'-methoxy oligonucleotide (DNA:2'-F or 2'-O-Me), RNA:2'-F or 2'-O-Me duplex. Dissociation constants are derived from ≥ 3 slopes of Lineweaver-Burk and /or Augustisson analysis. Estimated errors for the dissociation constants are 2 fold. Specificity is defined by dividing the K_d for a duplex by the K_d for an RNA:RNA duplex.

Table 4

A

	RNA	*Antisense*	*Initial Rate*
1	GGGCGCCGUCGGUGUGG	17mer P=O	1050 ± 203
1	GGGCGCCGUCGGUGUGG	17mer P=S	4034 ± 266

B

RNA No.	*RNA*	*Antisense DNA*	*Initial Rate (pmol L^{-1} min^{-1})*
1	GGGCGCCGUCGGUGUGG	17mer P=O	1050 ± 203
2	ACUCCACCAUAGUACACUCC (SEQ ID NO.: 22)	20mer P=O	1015 ± 264
3	UGGUGGGCGCCGUCGGUGUGGGCAA (SEQ ID NO.: 23)	25mer P=O	1502 ± 182

Table 5

RNA No.	*RNA*	*Antisense DNA*	*Initial Rate (pmol $L^{-1}min^{-1}$)*
1	17mer	CCACACCGACGGCGCCC (SEQ ID NO.: 24)	4034 ± 266
	17mer	CCACACCGACGGCGCCC	1081 ± 168
	17mer	CCACACCGACGGCGCCC	605 ± 81
	17mer	CCACACCGACGGCGCCC	330 ± 56
	17mer	CCACACCGACGGCGCCC	0
	17mer	CCACACCGACGGCGCCC	0
	17mer	CCACACCGACGGCGCCC	0
4*	20mer	AACACGCCCATTGCCCACCA (SEQ ID NO.: 25)	3400 ± 384
	20mer	AACACGCCCATTGCCCACCA	0

* Table legend for sequence

Table 6

Substrates	*Human RNase H*		*E. coli RNase H1*	
	Km (nM)	*(nmol $L^{-1}min^{-1}$)*	*Km (nM)*	*Vmax (nmol $L^{-1}min^{-1}$)*
25mer Ras (RNA no.3):DNA (P=O)	35.4	1.907		
17mer Ras (RNA no.1):DNA (P=0)	56.1	1.961	385	38.8
17mer Ras (RNA no.1):DNA (P=S)	13.9	1.077		

Table 7

Inhibitors	*Human RNase H1*		*E. coli RNase H1*	
	Kd (nM)	*Specificity*	*Kd (nM)*	*Specificity*
DNA:2′-O-Me	458	5.8	3400	2.1
RNA:2′-O-Me	409	5.2	3100	1.9
RNA:RNA	79	1.0	1600	1.0
RNA:2′-F	76	1.0		
DNA:2′-F	99	1.3		
DNA:DNA	3608	45.7	176000	110.0
ssRNA	1400	17.7		
ssDNA	1506	19.6	942000	588.8
ss2′-O-Me	2304	29.2	118000	73.8
ssDNA, P=S	36	0.5	14000	8.8

PROCEDURES

PROCEDURE 1

ICAM-1 Expression

Oligonucleotide Treatment of HUVECs

Cells were washed three times with Opti-MEM (Life Technologies, Inc.) prewarmed to 37 °C. Oligonucleotides were premixed with 10 g/mL Lipofectin (Life Technologies, Inc.) in Opti-MEM, serially diluted to the desired concentrations, and applied to washed cells. Basal and untreated (no oligonucleotide) control cells were also treated with Lipofectin. Cells were incubated for 4 h at 37 °C, at which time the medium was removed and replaced with standard growth medium with or without 5 mg/mL TNF-α 7 & D Systems). Incubation at 37 °C was continued until the indicated times.

Quantitation of ICAM-1 Protein Expression by Fluorescence-activated Cell Sorter

Cells were removed from plate surfaces by brief trypsinization with 0.25% trypsin in PBS. Trypsin activity was quenched with a solution of 2% bovine serum albumin and 0.2% sodium azide in PBS (+Mg/Ca). Cells were pelleted by centrifugation (1000 rpm, Beckman GPR centrifuge), resuspended in PBS, and stained with 3 $1/10^5$ cells of the ICAM-1 specific antibody, CD54-PE (Pharmingin). Antibodies were incubated with the cells for 30 min at 4C in the dark, under gently agitation. Cells were washed by centrifugation procedures and then resuspended in 0.3 mL of FacsFlow buffer (Becton Dickinson) with 0.5% formaldehyde (Polysciences). Expression of cell surface ICAM-1 was then determined by flow cytometry using a Becton Dickinson FACScan. Percentage of the control ICAM-1 expression was calculated as follows: [(oligonucleotide-treated ICAM-1 value) - (basal ICAM-1 value)/(non-treated ICAM-1 value) - (basal ICAM-1 value)]. (Baker, Brenda, *et. al.* 2'-*O*-(2-

Methoxy)ethyl-modified Anti-intercellular Adhesion Molecule 1 (ICAM-1) Oligonucleotides Selectively Increase the ICAM-1 mRNA Level and Inhibit Formation of the ICAM-1 Translation Initiation Complex in Human Umbilical Vein Endothelial Cells, *The Journal of Biological Chemistry*, 272, 11994-12000, **1997**.)

ICAM-1 expression of chimeric C3'-endo and C2'-endo modified oligonucleotides of the invention is measured by the reduction of ICAM-1 levels in treated HUVEC cells. The oligonucleotides are believed to work by RNase H cleavage mechanism. Appropriate scrambled control oligonucleotides are used as controls. They have the same base composition as the test sequence.

Sequences that contain the chimeric C3'-endo (2'-MOE)and C2'-endo (one of the following modifications: 2'-S-Me, 2'-Me, 2'-ara-F,2'-ara-OH, 2'-ara-O-Me) as listed in Table X below are prepared and tested in the above assay. SEQ ID NO: 14, a C-raf targeted oligonucleotide, is used as a control.

Table X

Oligonucleotides Containing chimeric 2'-O-(2-methoxyethyl) and 2'-S-(methyl) modifications.

SEQ ID NO:	Sequence (5'-3')	Target
14	**AsTsGs C^msAsTs** TsC^{m}sTs GsC^{m}s C^{m}s C^{m}sC^{m}s**C^{m}s AsAsGs GsA**	mouse C-raf
26	GsC^{m}sC^{m}s C^{m}sAsAs GsC^{m}sTs GsGsC^{m}s AsTsC^{m}s C^{m}sGsTs C^{m}sA	human ICAM-1

All nucleosides in bold are 2'-O-(methoxyethyl); subscript s indicates a phosphorothioate linkage; underlined nucleosides indicate

2'-S-Me- modification. superscript m on C (C^m)indicates a 5-methyl-C.

Table XI
Oligonucleotides Containing chimeric 2'-O-(2-methoxyethyl) and 2'-O-(methyl) modifications

SEQ ID NO:	Sequence (5'-3')	Target
14	AsTsGs C^msAsTs TsC^{m}sTs GsC^{m}sC^{m}s C^{m}sC^{m}sC^{m}s AsAsGs GsA	mouse C-raf
26	GsC^{m}sC^{m}s C^{m}sAsAs GsC^{m}sTs GsGsC^{m}s AsTsC^{m}s C^{m}sGsTs C^{m}sA	human ICAM-1

All nucleosides in bold are 2'-O-(methoxyethyl); subscript s indicates a phosphorothioate linkage; underlined nucleosides indicate 2'-Methyl modification. Superscript m on C (C^m)indicates a 5-methyl-C.

Table XII
Oligonucleotides Containing chimeric 2'-O-(2-methoxyethyl) and 2'-ara-(fluoro) modifications

SEQ ID NO:	Sequence (5'-3')	Target
14	AsTsGs C^msAsTs TsC^{m}sTs GsC^{m}sC^{m}s C^{m}sC^{m}sC^{m}s AsAsGs GsA	mouse C-raf
26	GsC^{m}sC^{m}s C^{m}sAsAs GsC^{m}sTs GsGsC^{m}s AsTsC^{m}s C^{m}sGsTs C^{m}sA	human ICAM-1

All nucleosides in bold are 2'-O-(methoxyethyl); subscript s indicates a phosphorothioate linkage; underlined nucleosides indicate 2'-ara-(fluoro) modification. superscript m on C (C^m)indicates a 5-methyl-C.

Table XIII
Oligonucleotides Containing chimeric 2'-O-(2-methoxyethyl) and 2'-ara-(OH) modifications

SEQ ID NO:	Sequence (5'-3')	Target
43	AsTsGs C^msAsTs TsC^msTs GsC^msC^ms C^msC^msC^ms AsAsGs GsA	mouse C-raf
44	GsC^msC^ms C^msAsAs GsC^msTs GsGsC^ms AsTsC^ms C^msGsTs C^msA	human ICAM-1

All nucleosides in bold are 2=-O-(methoxyethyl); subscript s indicates a phosphorothioate linkage; underlined nucleosides indicate 2'-ara-(OH) modification. superscript m on C (C^m)indicates a 5-methyl-C.

Table XIV
Oligonucleotides Containing chimeric 2'-O-(2-methoxyethyl) and 2'-ara-(OMe) modifications

SEQ ID NO:	Sequence (5'-3')	Target
14	AsTsGs C^msAsTs TsC^msTs GsC^msC^ms C^msC^msC^ms AsAsGs GsA	mouse C-raf
26	GsC^msC^ms C^msAsAs GsC^msTs GsGsC^ms AsTsC^ms C^msGsTs C^msA-3'	human ICAM-1

All nucleosides in bold are 2=-O-(methoxyethyl); subscript S indicates a phosphorothioate linkage; underlined nucleosides indicate 2'-ara-(OMe) modification. superscript m on C (C^m)indicates a 5-methyl-C.

PROCEDURE 2

Enzymatic Degradation of 2'-O-modified oligonucleotides

Three oligonucleotides are synthesized incorporating the modifications shown in Table 2 below at

the 3'-end. These modified oligonucleotides are subjected to snake venom phosphodiesterase action.

Oligonucleotides (30 nanomoles) are dissolved in 20 mL of buffer containing 50 mM Tris-HCl pH 8.5, 14 mM $MgCl_2$, and 72 mM NaCl. To this solution 0.1 units of snake-venom phosphodiesterase (Pharmacia, Piscataway, NJ), 23 units of nuclease P1 (Gibco LBRL, Gaithersberg, MD), and 24 units of calf intestinal phosphatase (Boehringer Mannheim, Indianapolis, IN) are added and the reaction mixture is incubated at 37C for 100 hours. HPLC analysis is carried out using a Waters model 715 automatic injector, model 600E pump, model 991 detector, and an Alltech (Alltech Associates, Inc., Deerfield, IL) nucleoside/nucleotide column (4.6 x 250 mm). All analyses are performed at room temperature. The solvents used are A: water and B: acetonitrile. Analysis of the nucleoside composition is accomplished with the following gradient: 0-5 min., 2% B (isocratic); 5-20 min., 2% B to 10% B (linear); 20-40 min., 10% B to 50% B. The integrated area per nanomole is determined using nucleoside standards. Relative nucleoside ratios are calculated by converting integrated areas to molar values and comparing all values to thymidine, which is set at its expected value for each oligomer.

Table XV
Relative Nuclease Resistance of 2'-Modified Chimeric Oligonucleotides

5'-TTT TTT TTT TTT TTT T*T*T*T*-3' SEQ ID NO 17
(Uniform phosphodiester)

T* = 2'-modified T
-S-Me
-Me
-2'-ara-(F)
-2'-ara-(OH)
-2'-ara-(OMe)

PROCEDURE 3

General procedure for the evaluation of chimeric C3'-endo and C2'-endo modified oligonucleotides targeted to *Ha-ras*

Different types of human tumors, including sarcomas, neuroblastomas, leukemias and lymphomas, contain active oncogenes of the *ras* gene family. *Ha-ras* is a family of small molecular weight GTPases whose function is to regulate cellular proliferation and differentiation by transmitting signals resulting in constitutive activation of *ras* are associated with a high percentage of diverse human cancers. Thus, *ras* represents an attractive target for anticancer therapeutic strategies.

The parent sequence (SEQ ID NO: 27) of Table XV is a 20-base phosphorothioate oligodeoxynucleotide targeting the initiation of translation region of human *Ha-ras* and it is a potent isotype-specific inhibitor of *Ha-ras* in cell culture based on screening assays (IC_{50}=45 nm). Treatment of cells *in vitro* with this sequence results in a rapid reduction of *Ha-ras* mRNA and protein synthesis and inhibition of proliferation of cells containing an activating *Ha-ras* mutation. When administered at doses of 25 mg/kg or lower by daily intraperitoneal injection (IP), this sequence exhibits potent antitumor activity in a variety of tumor xenograft models, whereas mismatch controls do not display antitumor activity. This sequence has been shown to be active against a variety of tumor types, including lung, breast, bladder, and pancreas in mouse xenograft studies (Cowsert, L.M. *Anti-cancer drug design*, **1997**, *12*, 359-371). A second-generation analog of SEQ ID NO: 27, where the 5' and 3' termini ("wings") of the sequence are modified with 2'-methoxyethyl (MOE) modification and the backbone is kept as phosphorothioate (Table XV, Control gapmer 1 with MOE control), exhibits IC_{50} of 15 nm in cell culture assays. Thus, a 3-fold

improvement in efficacy is observed from this chimeric analog. Because of the improved nuclease resistance of the 2'-MOE phosphorothioate, this sequence increases the duration of antisense effect *in vitro*. This will relate to frequency of administration of this drug to cancer patients. Further, this sequence is currently under evaluation in *ras* dependent tumor models (Cowsert, L.M. *Anti-cancer drug design*, **1997**, *12*, 359-371). The parent compound, SEQ ID NO: 27, is in Phase I clinical trials against solid tumors by systemic infusion.

Antisense oligonucleotides having the 2'-Me modification are prepared and tested in the aforementioned assays in the manner described to determine activity.

Ha-ras Antisense Oligonucleotides With chimeric C3'-endo and C2'-endo modifications and Their Controls.

Table XV

***Ha-ras* Antisense Oligonucleotides With chimeric C3'-endo and C2'-endo modifications and Their Controls.**

SEQ ID NO:	Sequence	Backbone	2'-Modif.	Comments
27	5'-TsCsCs GsTsCs AsTsCs GsCsTs CsCsTs CsAsGs GsG-3'	P=S	2'-H	parent
28	5'-TsCsAs GsTsAs AsTsAs GsGsCs CsCsAs CsAsTs GsG-3'	P=S	2'-H	mismatch control
29	5'-ToToCo GsTsCs AsTsCs GsCsTs CoCoTo CoAoGo GoG-3'	P=O/P=S/ P=O	2'-O-Moe in wings	Parent Gapmer (Mixed Backbone)
	5'-TsCsCs GsTsCs AsTsCs GsCsTs CsCsTs CsAsGs GsG-3'	P=S	2'-O-MOE in wings	Parent Gapmer as uniform thioate
30	5'-ToCoAo GsTsAs AsTsAs GsCsCs GsCsCs GsCoCo	P=O/P=S/ P=O	2'-O-MOE in wings	Parent Gapmer

	CoCoAo CoAoTo GoG-3'			(mixed Backbone)
31	5'-TsCsAs GsTsAs AsTs As GsCsCs GsCsCs CsCsAs CsAsTs GsC-3'	P=S	2'-O-MOE in wings	Control Gapmer as uniform Thioate
	5'-TsCsCs GsTsCs AsTsCs GsCsTs CsCsTs CsAsGs GsG-3'	P=S	2'-O-MOE in wings	Control Gapmer 1 with MOE control
	5'-TsCsAs GsTsAs AsTsAs GsCsCs GsCsCs CsCsAs CsAsTs GsC-3'	P=S	2'-O-MOE in wings	Control Gapmer 2 with MOE control

All underlined portions of sequences are 2'-Me.

PROCEDURE 7

***In vivo* nuclease resistance**

The *in vivo* Nuclease Resistance of chimeric C3'-endo and C2'-endo modified oligonucleotides is studied in mouse plasma and tissues (kidney and liver). For this purpose, the C-raf oligonucleotide series SEQ ID NO: 32 are used and the following five oligonucleotides listed in the Table below will be evaluated for their relative nuclease resistance.

Table XVI

Study of *in vivo* Nuclease Resistance of chimeric C3'-endo (2'-O-MOE) and C2'-endo (2'-S-Me) modified oligonucleotides with and without nuclease resistant caps (2'-5'-phosphate or phosphorothioate linkage with 3'-O-MOE in cap ends).

SEQ ID NO:	Sequence	Backbone	Description
32	5'-ATG CAT TCT GCC CCA AGGA-3'	P=S, 2'-H	(control) rodent C-raf

			antisense oligo
	AoToGo CoAsTs TsCsTs GsCsCs CsCsAo AoGoGo A	P=O/P=S/P=O	2'-MOE/2'-S-Me/ 2'-MOE
	AsTsGs CsAsTs TsCsTs GsCsCs CsCsAs AsGsGs A	P=S	2'-MOE/2'-S-Me/ 2'-MOE
	Ao*ToGo CoAsTs TsCsTs GsCsCs CsCsAo AoGoGo *A	P=O/P=S/P=O	In asterisk, 2'-5' linkage with 3'-O-MOE;2'-O-MOE / 2'-S-Me/2'-O-MOE/2'-5' linkage with 3'-O-MOE in asterisk;
	As*TsGs CsAsTs TsCsTs GsCsCs CsCsAs AsGsGs *A	P=S	In asterisk, 2'-5' linkage with 3'-O-MOE;2'-O-MOE /2'-S-Me/2'-O-MOE/2'-5' linkage with 3'-O-MOE in asterisk.

Table XVII

Study of in vivo Nuclease Resistance of chimeric C3'-endo (2'-O-MOE) and C2'-endo (2'-Me) modified oligonucleotides with and without nuclease resistant caps (2'-5'-phosphate or phosphorothioate linkage with 3'-O-MOE in cap ends).

SEQ ID NO:	Sequence	Backbone	Description
32	5'-ATG CAT TCT GCC CCA AGGA-3'	P=S, 2'-H	(control) rodent C-raf antisense oligo
	AoToGo CoAsTs TsCsTs GsCsCs CsCsAo AoGoGo A'	P=O/P=S/P=O	2'-MOE/2'-Me/ 2'-MOE
	AsTsGs CsAsTs TsCsTs GsCsCs CsCsAs AsGsGs A	P=S	2'-MOE/2'-Me/ 2'-MOE
	Ao*ToGo CoAsTs TsCsTs GsCsCs	P=O/P=S/P=O	In asterisk, 2'-5'

			CsCsAo AoGoGo *A linkage with 3'-O-MOE;2'-O-MOE / 2'-Me/2'-O-MOE/2'-5' linkage with 3'-O-MOE in asterisk;
	As*TsGs CsAsTs TsCsTs GsCsCs CsCsAs AsGsGs *A	P=S	In asterisk, 2'-5' linkage with 3'-O-MOE; 2'-O-MOE/ 2'-Me/2'-O-MOE/2'-5' linkage with 3'-O-MOE in asterisk;

Table XVIII

Study of in vivo Nuclease Resistance of chimeric C3'-endo (2'-O-MOE) and C2'-endo (2'-ara-F) modified oligonucleotides with and without nuclease resistant caps (2'-5'-phosphate or phosphorothioate linkage with 3'-O-MOE in cap ends).

SEQ ID NO:	Sequence	Backbone	Description
32	5'-ATG CAT TCT GCC CCA AGGA-3'	P=S, 2'-H	(control) rodent C-raf antisense oligo
	AoToGo CoAsTs TsCsTs GsCsCs CsCsAo AoGoGo A	P=O/P=S/P=O	2'-MOE/2'-ara-F/ 2'-MOE
	AsTsGs CsAsTs TsCsTs GsCsCs CsCsAs AsGsGs A	P=S	2'-MOE/2'-ara-F/2'-MOE
	Ao*ToGo CoAsTs TsCsTs GsCsCs CsCsAo AoGoGo *A	P=O/P=S/P=O	In asterisk, 2'-5' linkage with 3'-O-MOE;2'-O-MOE / 2'-ara-F/2'-O-MOE/2'-5' linkage with 3'-O-MOE in asterisk;
	As*TsGs CsAsTs TsCsTs	P=S	In asterisk,

	GsCsCs CsCsAs AsGsGs *A		2'-5' linkage with 3'-O-MOE;2'-O-MO 2'-ara-F/2'-O-MOE/2'-5' linkage with 3'-O-MOE in asterisk;

Table XIX

Study of in vivo Nuclease Resistance of chimeric C3'-endo (2'-O-MOE) and C2'-endo (2'-ara-OH) modified oligonucleotides with and without nuclease resistant caps (2'-5'-phosphate or phosphorothioate linkage with 3'-O-MOE in cap ends).

SEQ ID NO:	Sequence	Backbone	Description
32	5'-ATG CAT TCT GCC CCA AGGA-3'	P=S, 2'-H	(control) rodent C-raf antisense oligo
	AoToGo CoAsTs TsCsTs GsCsCs CsCsAo AoGoGo A	P=O/P=S/P=O	2'-MOE/2'-ara-OH/ 2'-MOE
	AsTsGs CsAsTs TsCsTs GsCsCs CsCsAs AsGsGs A	P=S	2'-MOE/2'-ara-OH/ 2'-MOE
	Ao*ToGo CoAsTs TsCsTs GsCsCs CsCsAo AoGoGo *A	P=O/P=S/P=O	In asterisk, 2'-5' linkage with 3'-O-MOE;2'-O-MOE / 2'-ara-OH /2'-O-MOE/2'-5' linkage with 3'-O-MOE in asterisk;
	As*TsGs CsAsTs TsCsTs GsCsCs CsCsAs AsGsGs *A	P=S	In asterisk, 2'-5' linkage with 3'-O-MOE;2'-O-MOE/ 2'-ara-OH/2'-O-		2'-ara-OH/2'-O-MOE/2'-5' linkage with 3'-O-MOE in asterisk;

Table XX

Study of in vivo Nuclease Resistance of chimeric C3'-endo (2'-O-MOE) and C2'-endo (2'-ara-OMe) modified oligonucleotides with and without nuclease resistant caps (2'-5'-phosphate or phosphorothioate linkage with 3'-O-MOE in cap ends).

SEQ ID NO:	Sequence	Backbone	Description
32	5'-ATG CAT TCT GCC CCA AGG A-3'	P=S, 2'-H	(control) rodent C-raf antisense oligo
	AoToGo CoAsTs TsCsTs GsCsCs CsCsAo AoGoGo Aa	P=O/P=S/P=O	2'-MOE/2'-ara-OMe / 2'-MOE
	AsTsGs CsAsTs TsCsTs GsCsCs CsCsAs AsGsGs A	P=S	2'-MOE/2'-ara-OMe /2'-MOE
	Ao*ToGo CoAsTs TsCsTs GsCsCs CsCsAo AoGoGo *A	P=O/P=S/P=O	In asterisk, 2'-5' linkage with 3'-O-MOE;2'-O-MOE / 2'-ara-OMe/2'-O-MOE/2'-5' linkage with 3'-O-MOE in asterisk;
	As*TsGs CsAsTs TsCsTs GsCsCs CsCsAs AsGsGs *A	P=S	In asterisk, 2'-5' linkage with 3'-O-MOE;2'-O-MOE / 2'-ara-OMe/2'-O-MOE/2'-5' linkage with 3'-O-MOE in asterisk.

PROCEDURE 8

Animal studies for in vivo nuclease resistance

For each oligonucleotide to be studied, 9 male BALB/c mice (Charles River, Wilmington, MA), weighing about 25 g

are used (Crooke *et al., J. Pharmacol. Exp. Ther.*, 1996, *277*, 923). Following a 1-week acclimation, the mice receive a single tail vein injection of oligonucleotide (5 mg/kg) administered in phosphate buffered saline (PBS), pH 7.0. The final concentration of oligonucleotide in the dosing solution is (5 mg/kg) for the PBS formulations. One retro-orbital bleed (either 0.25, 9.05, 2 or 4 post dose) and a terminal bleed (either 1, 3, 8 or 24 h post dose) is collected from each group. The terminal bleed (approximately 0.6-0.8 mL) is collected by cardiac puncture following ketamine/xylazine anesthesia. The blood is transferred to an EDTA-coated collection tube and centrifuged to obtain plasma. At termination, the liver and kidneys will be collected from each mouse. Plasma and tissues homogenates will be used for analysis for determination of intact oligonucleotide content by CGE. All samples are immediately frozen on dry ice after collection and stored at
-80C until analysis.

PROCEDURE 9

RNase H studies with chimeric C3'-endo and C2'-endo modified oligonucleotides with and without nuclease resistant caps

^{32}P Labeling of Oligonucleotides

The oligoribonucleotide (sense strand) was 5'-end labeled with ^{32}P using [^{32}P]ATP, T4 polynucleotide kinase, and standard procedures (Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., and Struhl, K., in *Current Protocols in Molecular Biology*, John Wiley, New York (1989)). The labeled RNA was purified by electrophoresis on 12% denaturing PAGE (Sambrook, J., Frisch, E. F., and Maniatis, T., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor

Laboratory Press, Plainview (1989)). The specific activity of the labeled oligonucleotide was approximately 6000 cpm/fmol.

Determination of RNase H Cleavage Patterns

Hybridization reactions were prepared in 120 µL of reaction buffer [20 mM Tris-HC (pH 7.5), 20 mM KCl, 10 mM $MgCl_2$, 0.1 mM DTT] containing 750 nM antisense oligonucleotide, 500 nM sense oligoribonucleotide, and 100,000 cpm ^{32}P-labeled sense oligoribonucleotide. Reactions were heated at 90 °C for 5 min and 1 unit of Inhibit-ACE was added. Samples were incubated overnight at 37 °C degrees. Hybridization reactions were incubated at 37 °C with 1.5 x 10.8^{-8} mg of *E. coli* RNase H enzyme for initial rate determinations and then quenched at specific time points. Samples were analyzed by trichloroacetic acid (TCA) assay or by denaturing polyacrylamide gel electrophoresis as previously described [Crooke, S. T., Lemonidis, K. M., Neilson, L., Griffey, R., Lesnik, E. A., and Monia, B. P., Kinetic characteristics of Escherichia coli RNase H1: cleavage of various antisense oligonucleotide-RNA duplexes, *Biochem J, 312*, 599 (1995); Lima, W.F. and Crooke, S. T., Biochemistry 36, 390-398, 1997].

Those skilled in the art will appreciate that numerous changes and modifications can be made to the preferred embodiments of the invention and that such changes and modifications can be made without departing from the spirit of the invention. It is therefore intended that the appended claims cover all such equivalent variations as fall within the true spirit and scope of the invention.

What is claimed is:

1. A mixed sequence oligonucleotide comprising at least 12 nucleotides in length and having a 3' end and a 5' end and divided into a first portion and a further portion,

said first portion being capable of supporting cleavage of a complementary target RNA by human RNase H1 polypeptide,

said further portion being incapable of supporting said cleavage by said RNase H1;

wherein said first portion comprises at least 6 nucleotides and is positioned in said oligonucleotide such that at least one of said 6 nucleotides is 8 to 12 nucleotides from the 3' end of said oligonucleotide.

2. The oligonucleotide of claim 1 comprising at least one CA nucleotide sequence within said first portion.

3. The oligonucleotide of claim 1 comprising from about 12 to about 50 nucleotides.

4. The oligonucleotide of claim 1 comprising from about 12 to about 25 nucleotides.

5. The oligonucleotide of claim 1 wherein each of said nucleotides of said first portion have B-form conformational geometry and are joined together in a continuous sequence.

6. The oligonucleotide of claim 1 wherein each of said nucleotides of said first portion is, independently, a 2'-deoxyribonucleotide, a 2'-SCH_3 ribonucleotide, a 2'-NH_2 ribonucleotide, a 2'-$NH(C_1$-C_2 alkyl) ribonucleotide, a 2'-$N(C_1$-C_2 alkyl)$_2$ ribonucleotide, a 2'-CF_3 ribonucleotide, a

2'=CH_2 ribonucleotide, a 2'=CHF ribonucleotide, a 2'=CF_2 ribonucleotide, a 2'-CH_3 ribonucleotide, a 2'-C_2H_5 ribonucleotide, a 2'-CH=CH_2 ribonucleotide or a 2'-C≡CH ribonucleotide.

7. The oligonuceotide of claim 1 wherein each of said nucleotides of said first portion is a 2'-deoxyribonucleotide.

8. The oligonucleotide of claim 1 wherein each of said nucleotide of said first portion is, independently, a 2'-CN arabinonucleotide, a 2'-F arabinonucleotide, a 2'-Cl arabinonucleotide, a 2'-Br arabinonucleotide, a 2'-N_3 arabinonucleotide, a 2'-OH arabinonucleotide, a 2'-O-CH_3 arabinonucleotide or a 2'-dehydro-2'-CH_3 arabinonucleotide.

9. The oligonucleotide of claim 1 wherein each of said nucleotides of said first portion is, independently, a 2'-F arabinonucleotide, a 2'-OH arabinonucleotide or a 2'-O-CH_3 arabinonucleotide.

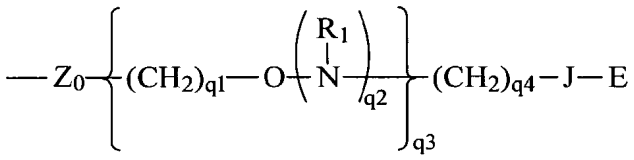
10. The oligonucleotide of claim 1 wherein eachof said nucleotides of said first portion is, independently, a 2'-F arabinonucleotide or a 2'-OH arabinonucleotide.

11. The oligonucleotide of claim 1 wherein said nucleotides of said first portion are joined together in said continuous sequence by phosphate, phosphorothioate, phosphorodithioate or boranophosphate linkages.

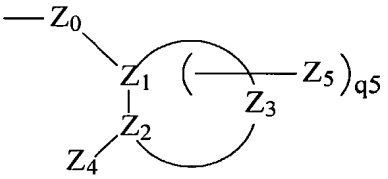
12. The oligonucleotide of claim 1 wherein said further portion includes a plurality of nucleotides, at least some of said nucleotides comprise a 2' substituent group wherein each substituent group is, independently,

hydroxyl, C_1-C_{20} alkyl, C_2-C_{20} alkenyl, C_2-C_{20} alkynyl, halogen, amino, thiol, keto, carboxyl, nitro, nitroso, nitrile, trifluoromethyl, trifluoromethoxy, O-alkyl, O-alkenyl, O-alkynyl, S-alkyl, S-alkenyl, S-alkynyl, NH-alkyl, NH-alkenyl, NH-alkynyl, N-dialkyl, O-aryl, S-aryl, NH-aryl, O-aralkyl, S-aralkyl, NH-aralkyl, N-phthalimido, imidazole, azido, hydrazino, hydroxylamino, isocyanato, sulfoxide, sulfone, sulfide, disulfide, silyl, aryl, heterocycle, carbocycle, intercalator, reporter molecule, conjugate, polyamine, polyamide, polyalkylene glycol, or polyether;

or each substituent group has one of formula I or II:



I



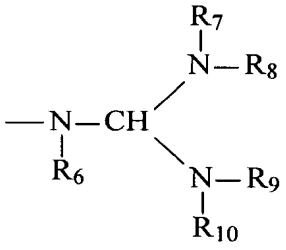
II

wherein:

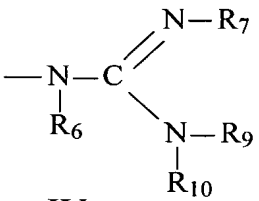
Z_0 is O, S or NH;

J is a single bond, O or C(=O);

E is C_1-C_{10} alkyl, $N(R_1)(R_2)$, $N(R_1)(R_5)$, $N=C(R_1)(R_2)$, $N=C(R_1)(R_5)$ or has one of formula III or IV;



III



IV

each R_6, R_7, R_8, R_9 and R_{10} is, independently, hydrogen, $C(O)R_{11}$, substituted or unsubstituted C_1-C_{10} alkyl, substituted or unsubstituted C_2-C_{10} alkenyl, substituted or unsubstituted C_2-C_{10} alkynyl, alkylsulfonyl, arylsulfonyl, a

chemical functional group or a conjugate group, wherein the substituent groups are selected from hydroxyl, amino, alkoxy, carboxy, benzyl, phenyl, nitro, thiol, thioalkoxy, halogen, alkyl, aryl, alkenyl and alkynyl;

or optionally, R_7 and R_8, together form a phthalimido moiety with the nitrogen atom to which they are attached;

or optionally, R_9 and R_{10}, together form a phthalimido moiety with the nitrogen atom to which they are attached;

each R_{11} is, independently, substituted or unsubstituted C_1-C_{10} alkyl, trifluoromethyl, cyanoethyloxy, methoxy, ethoxy, t-butoxy, allyloxy, 9-fluorenylmethoxy, 2-(trimethylsilyl)-ethoxy, 2,2,2-trichloroethoxy, benzyloxy, butyryl, iso-butyryl, phenyl or aryl;

R_5 is T-L,

T is a bond or a linking moiety;

L is a chemical functional group, a conjugate group or a solid support material;

each R_1 and R_2 is, independently, H, a nitrogen protecting group, substituted or unsubstituted C_1-C_{10} alkyl, substituted or unsubstituted C_2-C_{10} alkenyl, substituted or unsubstituted C_2-C_{10} alkynyl, wherein said substitution is OR_3, SR_3, NH_3^+, $N(R_3)(R_4)$, guanidino or acyl where said acyl is an acid amide or an ester;

or R_1 and R_2, together, are a nitrogen protecting group or are joined in a ring structure that optionally includes an additional heteroatom selected from N and O;

or R_1, T and L, together, are a chemical functional group;

each R_3 and R_4 is, independently, H, C_1-C_{10} alkyl, a nitrogen protecting group, or R_3 and R_4, together, are a nitrogen protecting group;

or R_3 and R_4 are joined in a ring structure that optionally includes an additional heteroatom selected from N and O;

Z_4 is OX, SX, or $N(X)_2$;

each X is, independently, H, C_1-C_8 alkyl, C_1-C_8 haloalkyl, $C(=NH)N(H)R_5$, $C(=O)N(H)R_5$ or $OC(=O)N(H)R_5$;

R_5 is H or C_1-C_8 alkyl;

Z_1, Z_2 and Z_3 comprise a ring system having from about 4 to about 7 carbon atoms or having from about 3 to about 6 carbon atoms and 1 or 2 hetero atoms wherein said hetero atoms are selected from oxygen, nitrogen and sulfur and wherein said ring system is aliphatic, unsaturated aliphatic, aromatic, or saturated or unsaturated heterocyclic;

Z_5 is alkyl or haloalkyl having 1 to about 10 carbon atoms, alkenyl having 2 to about 10 carbon atoms, alkynyl having 2 to about 10 carbon atoms, aryl having 6 to about 14 carbon atoms, $N(R_1)(R_2)$ OR_1, halo, SR_1 or CN;

each q_1 is, independently, an integer from 1 to 10;

each q_2 is, independently, 0 or 1;

q_3 is 0 or an integer from 1 to 10;

q_4 is an integer from 1 to 10; and

q_5 is from 0, 1 or 2;

provided that when q_3 is 0, q_4 is greater than 1.

13. The oligonucleotide of claim 1 wherein each of said nucleotides of said further portion is, independently, a 2'-F ribonucleotide, a 2'-O-(C_1-C_6 alkyl) ribonucleotide, or a 2'-O-(C_1-C_6 substituted alkyl) ribonucleotide wherein the substitution is C_1-C_6 ether, C_1-C_6 thioether, amino, amino(C_1-C_6 alkyl) or amino(C_1-C_6 alkyl)$_2$.

14. The oligonucleotide of claim 1 wherein said nucleotides of said further portion are joined together in a continuous sequence by 3'-5' phosphodiester, 2'-5' phosphodiester, phosphorothioate, Sp phosphorothioate, Rp phosphorothioate, phosphorodithioate, 3'-deoxy-3'-amino

phosphoroamidate, 3'-methylenephosphonate, methylene(methylimino), dimethylhydrazino, amide 3, amide 4 or boranophosphate linkages.

15. The oligonucleotide of claim 1 wherein at least two of said nucleotides of said further portion are joined together in a continuous sequence that is positioned 3' to said first portion.

16. The oligonucleotide of claim 1 wherein at least two of said nucleotides of said further portion are joined together in a continuous sequence that is positioned 5' to said first portion.

17. The oligonucleotide of claim 1 wherein at least two of said nucleotides of said further portion are joined together in a continuous sequence that is positioned 3' to said first portion and at least two of said further portion are joined together in a continuous sequence that is positioned 5' to said first portion.

18. The oligonucleotide of claim 1 wherein at least four of said nucleotides of said further portion are joined together in a continuous sequence that is positioned 3' to said first portion.

19. The oligonucleotide of claim 1 wherein at least four of said nucleotides of said further portion are joined together in a continuous sequence that is positioned 5' to said first portion.

20. The oligonucleotide of claim 1 wherein at least four of said nucleotides of said further portion are joined together in a continuous sequence that is positioned 3' to

said first portion and at least four of said nucleotides of said further portion are joined together in a continuous sequence that is positioned 5' to said first portion.

21. A mixed sequence oligonucleotide comprising at least 8 nucleotides and having a CA nucleotide sequence wherein at least one of the two nucleotides of said CA sequence is positioned 8 to 12 nucleotides from the 3' end of said oligonucleotide.

22. The oligonucleotide of claim 21 wherein said oligonucleotide is capable of supporting cleavage of a complementary target RNA by human RNase H1 polypeptide.

23. A mixed sequence chimeric oligonucleotide comprising at least 8 nucleotides and having a CA nucleotide sequence wherein at least one of the two nucleotides of said CA sequence is positioned 8 to 12 nucleotides from the 3' end of said oligonucleotide.

24. The chimeric oligonucleotide of claim 23 wherein said oligonucleotide is capable of supporting cleavage of a complementary target RNA by human RNase H1 polypeptide.

25. A mixed sequence oligonucleotide comprising 8 to 25 nucleotides and having a CA nucleotide sequence wherein at least one of the nucleotides of said CA sequence is positioned 8 to 12 nucleotides from the 3' end of said oligonucleotide.

26. A mixed sequence chimeric oligonucleotide comprising 8 to 25 nucleotides and having a CA nucleotide sequence wherein at least one of the nucleotides of said CA sequence is positioned 8 to 12 nucleotides from the 3' end

of said oligonucleotide.

27. A chimeric oligonucleotide comprising 8 to 25 nucleotides and having a portion capable of supporting cleavage of a complementary target RNA by human RNase H1 polypeptide wherein said portion supporting said cleavage is at least 6 nucleotides in length and is positioned in said oligonucleotide such that at least one of said 6 nucleotides is positioned 8 to 12 nucleotides from the 3' end of said oligonucleotide.

28. The oligonucleotide of claim 27 wherein said oligonucleotide comprises at least one CA nucleotide sequence within said portion supporting said cleavage.

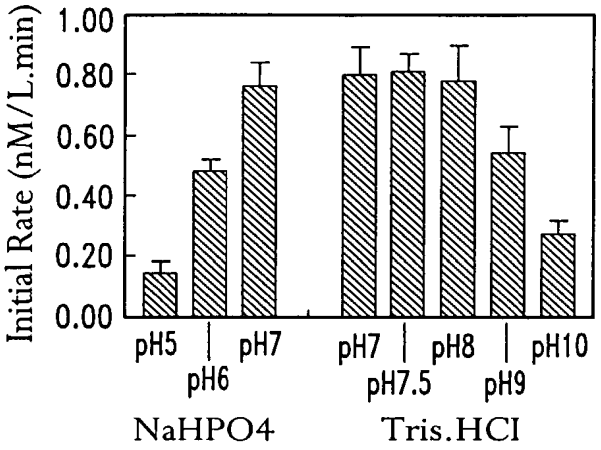


FIG. 1A

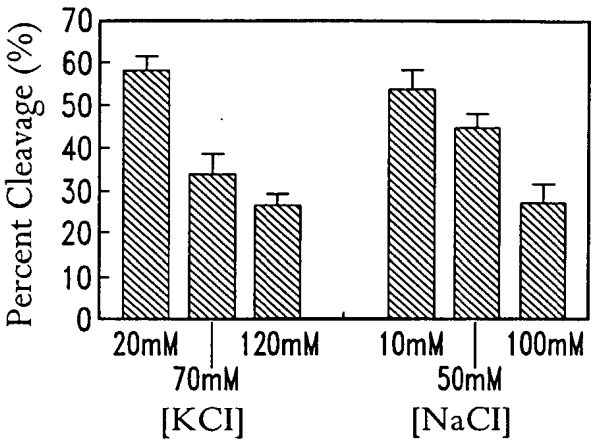


FIG. 1B

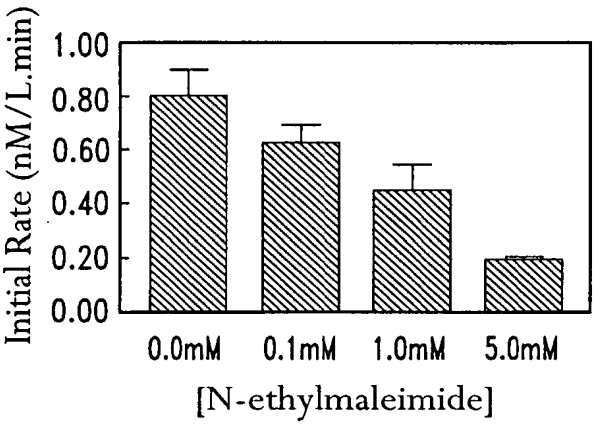


FIG. 1C

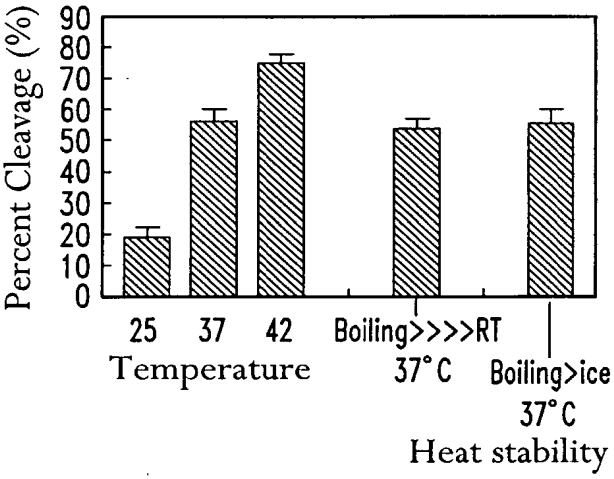


FIG. 1D

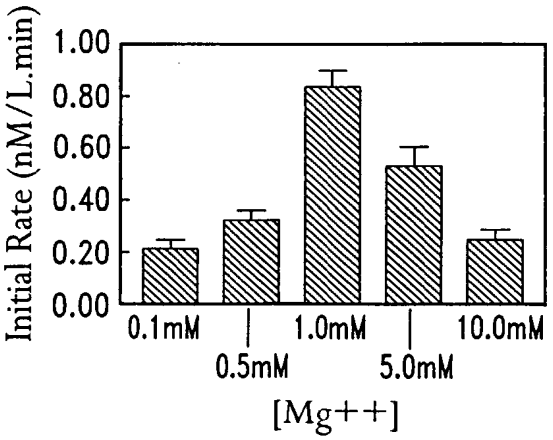


FIG. 1E

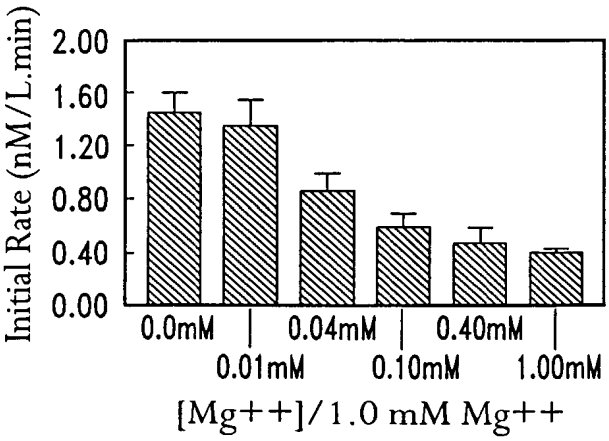


FIG. 1F

IDT Exhibit 1020.6293

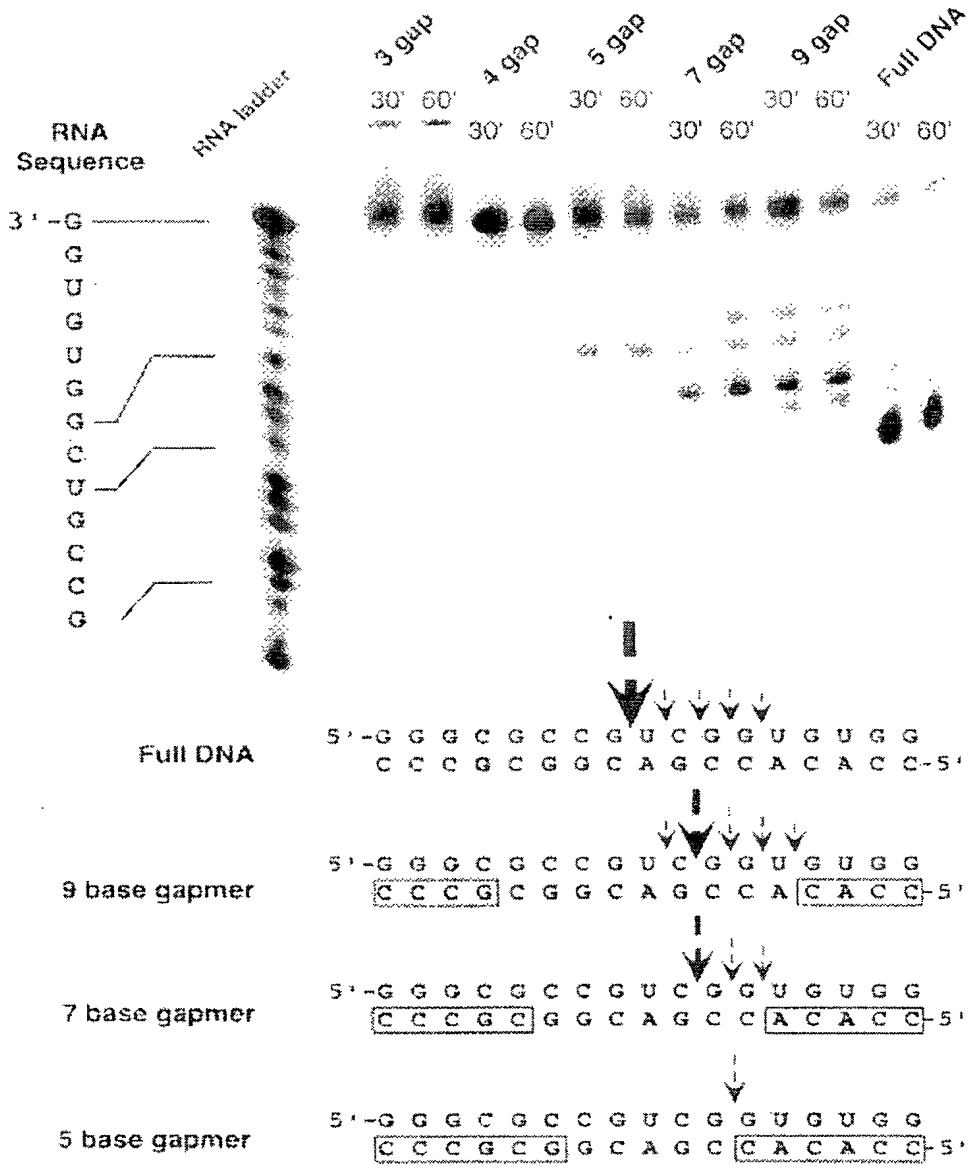


FIG. 2

IDT Exhibit 1020.6294

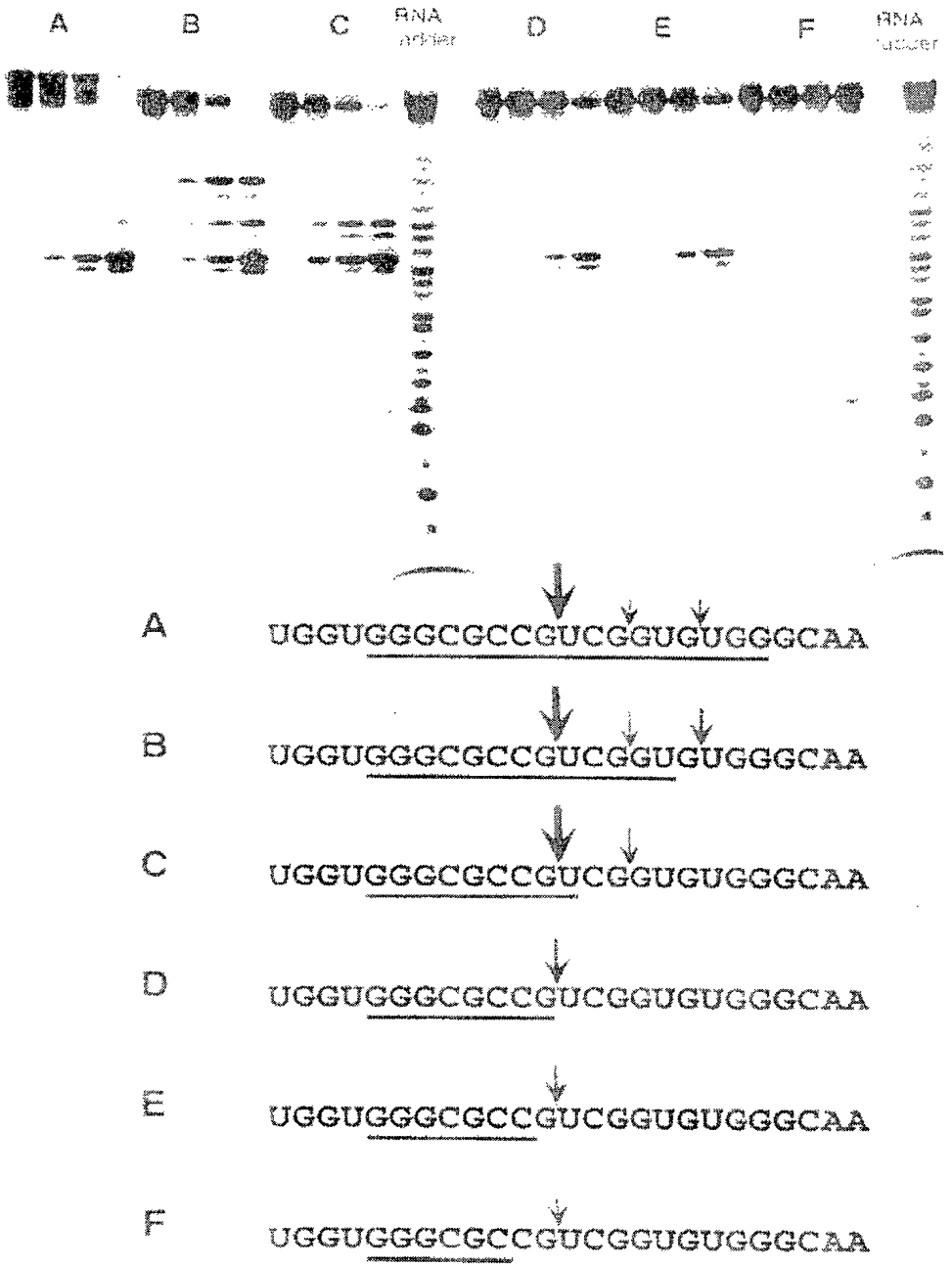


FIG. 3

IDT Exhibit 1020.6295

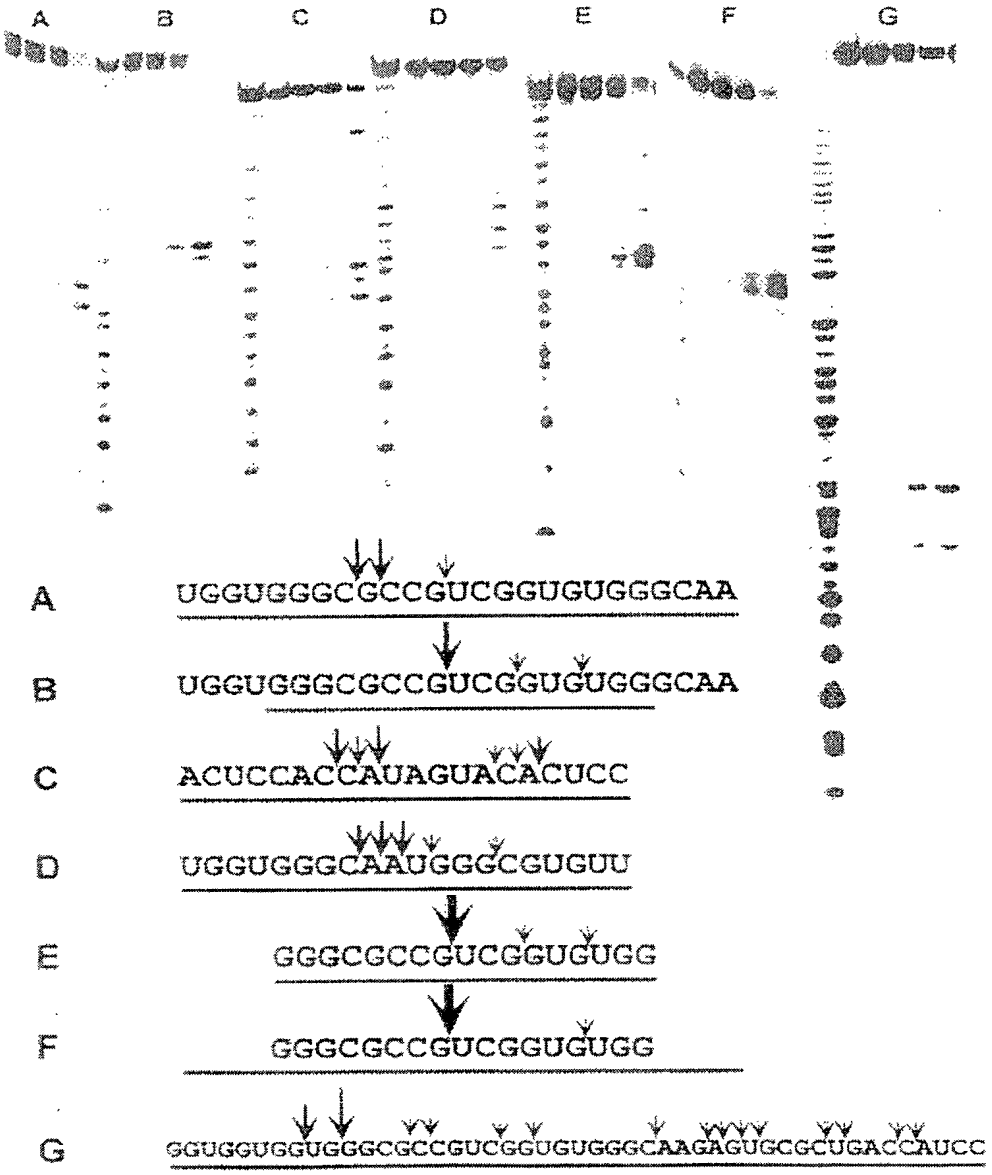


FIG. 4

IDT Exhibit 1020.6296

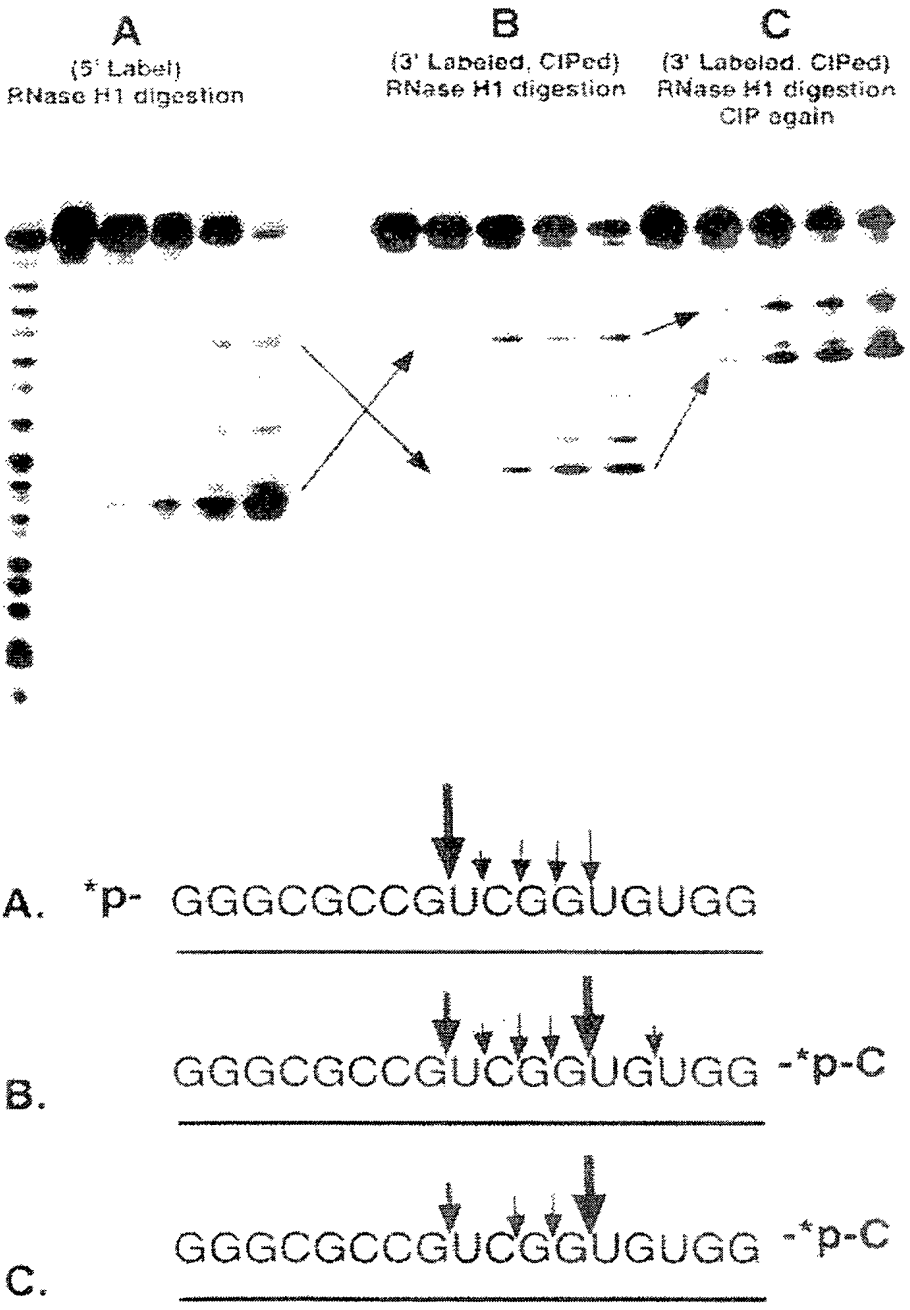


FIG. 5

IDT Exhibit 1020.6297

```
Human       MSWLLFLAHRVALAALPCRRGSRGFGMFYAVRRGRKTGVFLTWNECRAQVDRFPAARFKKFATEDEAWAFVRKSAS
Chicken     MLRWL----VAL-LSHSCFVSKGGGMFYAVRKGROTGVYRTWAECQQQVNRFPSASFKKFATEKEAWAFVGAGPP
Yeast                     MARQGNFYAVRKGRETGIYNTWNECKNQVDGYGGAIYKKFNSYEQAKSFLGQPNT
Mouse EST                 GICGLGMFYAVRRGRRPGVFLSWSECKAQVDRFPAARFKKFATEDEAWAFVRSSSS

Human       PEVSEGHENQHGQESEAKPGKRLREPLDGDG----------------------------------------
Chicken     DGQQSAPAETHGASAVAQENASHREEPETDV----------------------------------------
Yeast       TSNYGSSTHAGGQVSKPHTTQKRVHRRNRPLHYSSLTSSSACSSLSSANTNTFYSVKSNVPNIESKIFNNWKDCQA
Mouse EST   PDGSKGQESAHEQKSQAKTSKRPREPL

                                                                          @@
Human       ----------------------------HESAQPYAKHMKPSVEP-APPVSRDTFSYMGDFVVVYTDGCCSSNGRRKPRA
Chicken     ----------LCCNACKRPYEQSTNEEHTVRRAKH---DEEQSTPVVSEAKFSYMGEFAVVYTDGCCSGNGRNRARA
Yeast       YVKHKRGITFKKFEDQLAAENFISGMSAHDY-KLMNISKESFESKYKLSSNTMYNKSMNVYCDGSSFGNGTSSSRA
E.coli                                                  MLKQVEIFTDGSCLGNPGPGGYG
Mouse EST                                               VVVYTDGCCSSNGRKRARA

                                    @                   @                  * *
Human       GIGVYWGPGHPLNVGI-RLPGRQTNQRAEIHAACKAIEQAKTQNIN-----KLVLYTDSMFTINGITNWVQGWKKN
Chicken     GIGVYWGPGHPLNISE-RLPGRQTNQRAEIHAACKAIEQAKSQNIK-----KLIIYTDSKETINGITSWVENWKTN
Yeast       GYGAYFEGAPEENISEPLLSGAQTNNRAEIEAVSEALKKIWEKLTNEKEKVNYQIKTDSEYVTKLLNDRYMTYDNK
E.coli      AILRYRGREKTFSAGY----TRTTNNRMELMAAIVALEALKEHC---------EVILSTDSQYVRQGITQWIHNWKKR
Mouse EST   GIGVYWGPGHPLNVRI-RLPGRQTNQRAEIHAACKAVMQAKAQNIS-----KLVLYTDSMFTINGITNWVQGWKKN

            *     *
Human       GWKTSAGKEVINKEDFVALERL----------TQGMDIQWMHVPGHSGFIGNEEADRLAREGA-KQSED
Chicken     GWRTSSGGSVINKEDFQKLDSL----------SKGIEIQWMHIPGHAGFQGNEEADRLAREGASKQKL
Yeast       KLEGLPNSDLIVPLVQRFVKVKYYELNKECFKNNGKFQIEWVKGHDGDPGNEMADFLAKKGASRR
E.coli      GWKTADKKPVKNVDLWQRLDAA----------LGQHQIKWEWVKGHAGHPENERCDELARAAAMNPTLEDTGYQVEV
Mouse EST   GWRTSTGKDVINKEDFMELDEL----------TQGMDIQWMHIPGHSGFVGNEE
```

FIG. 6

IDT Exhibit 1020.6298

SEQUENCE LISTING

<110> Crooke, Stanley T.
Lima, Walter F.
Wu, Hongjiang
Manoharan, Muthiah

<120> Human Rnase H and Oligonucleotide Compositions Thereof

<130> ISIS4505

<140> Unassigned
<141> 2000-09-29

<150> USSN 09/409,926
<151> 1999-09-30

<160> 33

<170> PatentIn Ver. 2.1

<210> 1
<211> 286
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Novel Sequence

<400> 1
Met Ser Trp Leu Leu Phe Leu Ala His Arg Val Ala Leu Ala Ala Leu
1 5 10 15

Pro Cys Arg Arg Gly Ser Arg Gly Phe Gly Met Phe Tyr Ala Val Arg
20 25 30

Arg Gly Arg Lys Thr Gly Val Phe Leu Thr Trp Asn Glu Cys Arg Ala
35 40 45

Gln Val Asp Arg Phe Pro Ala Ala Arg Phe Lys Lys Phe Ala Thr Glu
50 55 60

Asp Glu Ala Trp Ala Phe Val Arg Lys Ser Ala Ser Pro Glu Val Ser
65 70 75 80

Glu Gly His Glu Asn Gln His Gly Gln Glu Ser Glu Ala Lys Pro Gly
85 90 95

IDT Exhibit 1020.6299

```
Lys Arg Leu Arg Glu Pro Leu Asp Gly Asp Gly His Glu Ser Ala Gln
            100                 105                 110

Pro Tyr Ala Lys His Met Lys Pro Ser Val Glu Pro Ala Pro Pro Val
        115                 120                 125

Ser Arg Asp Thr Phe Ser Tyr Met Gly Asp Phe Val Val Val Tyr Thr
    130                 135                 140

Asp Gly Cys Cys Ser Ser Asn Gly Arg Arg Lys Pro Arg Ala Gly Ile
145                 150                 155                 160

Gly Val Tyr Trp Gly Pro Gly His Pro Leu Asn Val Gly Ile Arg Leu
                165                 170                 175

Pro Gly Arg Gln Thr Asn Gln Arg Ala Glu Ile His Ala Ala Cys Lys
            180                 185                 190

Ala Ile Glu Gln Ala Lys Thr Gln Asn Ile Asn Lys Leu Val Leu Tyr
        195                 200                 205

Thr Asp Ser Met Phe Thr Ile Asn Gly Ile Thr Asn Trp Val Gln Gly
    210                 215                 220

Trp Lys Lys Asn Gly Trp Lys Thr Ser Ala Gly Lys Glu Val Ile Asn
225                 230                 235                 240

Lys Glu Asp Phe Val Ala Leu Glu Arg Leu Thr Gln Gly Met Asp Ile
                245                 250                 255

Gln Trp Met His Val Pro Gly His Ser Gly Phe Ile Gly Asn Glu Glu
            260                 265                 270

Ala Asp Arg Leu Ala Arg Glu Gly Ala Lys Gln Ser Glu Asp
        275                 280                 285
```

<210> 2
<211> 293
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Novel Sequence

<400> 2

```
Met Leu Arg Trp Leu Val Ala Leu Leu Ser His Ser Cys Phe Val Ser
  1               5                  10                  15
```


IDT Exhibit 1020.6300

```
Lys Gly Gly Gly Met Phe Tyr Ala Val Arg Lys Gly Arg Gln Thr Gly
            20                  25                  30

Val Tyr Arg Thr Trp Ala Glu Cys Gln Gln Gln Val Asn Arg Phe Pro
        35                  40                  45

Ser Ala Ser Phe Lys Lys Phe Ala Thr Glu Lys Glu Ala Trp Ala Phe
    50                  55                  60

Val Gly Ala Gly Pro Pro Asp Gly Gln Gln Ser Ala Pro Ala Glu Thr
65                  70                  75                  80

His Gly Ala Ser Ala Val Ala Gln Glu Asn Ala Ser His Arg Glu Glu
                85                  90                  95

Pro Glu Thr Asp Val Leu Cys Cys Asn Ala Cys Lys Arg Pro Tyr Glu
            100                 105                 110

Gln Ser Thr Asn Glu Glu His Thr Val Arg Arg Ala Lys His Asp Glu
        115                 120                 125

Glu Gln Ser Thr Pro Val Val Ser Glu Ala Lys Phe Ser Tyr Met Gly
    130                 135                 140

Glu Phe Ala Val Val Tyr Thr Asp Gly Cys Cys Ser Gly Asn Gly Arg
145                 150                 155                 160

Asn Arg Ala Arg Ala Gly Ile Gly Val Tyr Trp Gly Pro Gly His Pro
                165                 170                 175

Leu Asn Ile Ser Glu Arg Leu Pro Gly Arg Gln Thr Asn Gln Arg Ala
            180                 185                 190

Glu Ile His Ala Ala Cys Lys Ala Ile Glu Gln Ala Lys Ser Gln Asn
        195                 200                 205

Ile Lys Lys Leu Ile Ile Tyr Thr Asp Ser Lys Phe Thr Ile Asn Gly
    210                 215                 220

Ile Thr Ser Trp Val Glu Asn Trp Lys Thr Asn Gly Trp Arg Thr Ser
225                 230                 235                 240

Ser Gly Gly Ser Val Ile Asn Lys Glu Asp Phe Gln Lys Leu Asp Ser
                245                 250                 255

Leu Ser Lys Gly Ile Glu Ile Gln Trp Met His Ile Pro Gly His Ala
            260                 265                 270
```


IDT Exhibit 1020.6301

```
Gly Phe Gln Gly Asn Glu Glu Ala Asp Arg Leu Ala Arg Glu Gly Ala
        275                 280                 285

Ser Lys Gln Lys Leu
    290
```

<210> 3
<211> 348
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Novel Sequence

<400> 3

```
Met Ala Arg Gln Gly Asn Phe Tyr Ala Val Arg Lys Gly Arg Glu Thr
  1               5                  10                  15

Gly Ile Tyr Asn Thr Trp Asn Glu Cys Lys Asn Gln Val Asp Gly Tyr
             20                  25                  30

Gly Gly Ala Ile Tyr Lys Lys Phe Asn Ser Tyr Glu Gln Ala Lys Ser
         35                  40                  45

Phe Leu Gly Gln Pro Asn Thr Thr Ser Asn Tyr Gly Ser Ser Thr His
     50                  55                  60

Ala Gly Gly Gln Val Ser Lys Pro His Thr Thr Gln Lys Arg Val His
 65                  70                  75                  80

Arg Arg Asn Arg Pro Leu His Tyr Ser Ser Leu Thr Ser Ser Ser Ala
                 85                  90                  95

Cys Ser Ser Leu Ser Ser Ala Asn Thr Asn Thr Phe Tyr Ser Val Lys
            100                 105                 110

Ser Asn Val Pro Asn Ile Glu Ser Lys Ile Phe Asn Asn Trp Lys Asp
        115                 120                 125

Cys Gln Ala Tyr Val Lys His Lys Arg Gly Ile Thr Phe Lys Lys Phe
    130                 135                 140

Glu Asp Gln Leu Ala Ala Glu Asn Phe Ile Ser Gly Met Ser Ala His
145                 150                 155                 160

Asp Tyr Lys Leu Met Asn Ile Ser Lys Glu Ser Phe Glu Ser Lys Tyr
```


IDT Exhibit 1020.6302

```
                165                 170                 175

Lys Leu Ser Ser Asn Thr Met Tyr Asn Lys Ser Met Asn Val Tyr Cys
            180                 185                 190

Asp Gly Ser Ser Phe Gly Asn Gly Thr Ser Ser Ser Arg Ala Gly Tyr
        195                 200                 205

Gly Ala Tyr Phe Glu Gly Ala Pro Glu Glu Asn Ile Ser Glu Pro Leu
    210                 215                 220

Leu Ser Gly Ala Gln Thr Asn Asn Arg Ala Glu Ile Glu Ala Val Ser
225                 230                 235                 240

Glu Ala Leu Lys Lys Ile Trp Glu Lys Leu Thr Asn Glu Lys Glu Lys
                245                 250                 255

Val Asn Tyr Gln Ile Lys Thr Asp Ser Glu Tyr Val Thr Lys Leu Leu
            260                 265                 270

Asn Asp Arg Tyr Met Thr Tyr Asp Asn Lys Lys Leu Glu Gly Leu Pro
        275                 280                 285

Asn Ser Asp Leu Ile Val Pro Leu Val Gln Arg Phe Val Lys Val Lys
    290                 295                 300

Lys Tyr Tyr Glu Leu Asn Lys Glu Cys Phe Lys Asn Asn Gly Lys Phe
305                 310                 315                 320

Gln Ile Glu Trp Val Lys Gly His Asp Gly Asp Pro Gly Asn Glu Met
                325                 330                 335

Ala Asp Phe Leu Ala Lys Lys Gly Ala Ser Arg Arg
            340                 345
```

<210> 4
<211> 216
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Novel Sequence

```
<400> 4
Gly Ile Cys Gly Leu Gly Met Phe Tyr Ala Val Arg Arg Gly Arg Arg
  1               5                  10                  15
```


IDT Exhibit 1020.6303

```
Pro Gly Val Phe Leu Ser Trp Ser Glu Cys Lys Ala Gln Val Asp Arg
             20                  25                  30

Phe Pro Ala Ala Arg Phe Lys Lys Phe Ala Thr Glu Asp Glu Ala Trp
         35                  40                  45

Ala Phe Val Arg Ser Ser Ser Ser Pro Asp Gly Ser Lys Gly Gln Glu
     50                  55                  60

Ser Ala His Glu Gln Lys Ser Gln Ala Lys Thr Ser Lys Arg Pro Arg
 65                  70                  75                  80

Glu Pro Leu Val Val Val Tyr Thr Asp Gly Cys Cys Ser Ser Asn Gly
                 85                  90                  95

Arg Lys Arg Ala Arg Ala Gly Ile Gly Val Tyr Trp Gly Pro Gly His
            100                 105                 110

Pro Leu Asn Val Arg Ile Arg Leu Pro Gly Arg Gln Thr Asn Gln Arg
        115                 120                 125

Ala Glu Ile His Ala Ala Cys Lys Ala Val Met Gln Ala Lys Ala Gln
    130                 135                 140

Asn Ile Ser Lys Leu Val Leu Tyr Thr Asp Ser Met Phe Thr Ile Asn
145                 150                 155                 160

Gly Ile Thr Asn Trp Val Gln Gly Trp Lys Lys Asn Gly Trp Arg Thr
                165                 170                 175

Ser Thr Gly Lys Asp Val Ile Asn Lys Glu Asp Phe Met Glu Leu Asp
            180                 185                 190

Glu Leu Thr Gln Gly Met Asp Ile Gln Trp Met His Ile Pro Gly His
        195                 200                 205

Ser Gly Phe Val Gly Asn Glu Glu
    210                 215
```

<210> 5
<211> 155
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Novel Sequence

IDT Exhibit 1020.6304

<400> 5

```
Met Leu Lys Gln Val Glu Ile Phe Thr Asp Gly Ser Cys Leu Gly Asn
  1               5                  10                  15

Pro Gly Pro Gly Gly Tyr Gly Ala Ile Leu Arg Tyr Arg Gly Arg Glu
             20                  25                  30

Lys Thr Phe Ser Ala Gly Tyr Thr Arg Thr Thr Asn Asn Arg Met Glu
         35                  40                  45

Leu Met Ala Ala Ile Val Ala Leu Glu Ala Leu Lys Glu His Cys Glu
     50                  55                  60

Val Ile Leu Ser Thr Asp Ser Gln Tyr Val Arg Gln Gly Ile Thr Gln
 65                  70                  75                  80

Trp Ile His Asn Trp Lys Lys Arg Gly Trp Lys Thr Ala Asp Lys Lys
                 85                  90                  95

Pro Val Lys Asn Val Asp Leu Trp Gln Arg Leu Asp Ala Ala Leu Gly
            100                 105                 110

Gln His Gln Ile Lys Trp Glu Trp Val Lys Gly His Ala Gly His Pro
        115                 120                 125

Glu Asn Glu Arg Cys Asp Glu Leu Ala Arg Ala Ala Ala Met Asn Pro
    130                 135                 140

Thr Leu Glu Asp Thr Gly Tyr Gln Val Glu Val
145                 150                 155
```

<210> 6
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Novel Sequence

<400> 6
acgctggccg ggagtcgaaa tgcttc 26

<210> 7
<211> 28
<212> DNA
<213> Artificial Sequence

IDT Exhibit 1020.6305

<220>
<223> Description of Artificial Sequence: Novel Sequence

<400> 7
ctgttcctgg cccacagagt cgccttgg 28

<210> 8
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Novel Sequence

<400> 8
ggtctttctg acctggaatg agtgcagag 29

<210> 9
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Novel Sequence

<400> 9
cttgcctggt ttcgccctcc gattcttgt 29

<210> 10
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Novel Sequence

<400> 10
ttgattttca tgcccttctg aaacttccg 29

<210> 11
<211> 34
<212> DNA
<213> Artificial Sequence

IDT Exhibit 1020.6306

<220>
<223> Description of Artificial Sequence: Novel Sequence

<400> 11
cctcatcctc tatggcaaac ttcttaaatc tggc 34

<210> 12
<211> 17
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Novel Sequence

<400> 12
gggcgccgtc ggtgtgg 17

<210> 13
<211> 12
<212> RNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Novel Sequence

<400> 13
cgcgaauucg cg 12

<210> 14
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Novel Sequence

<220>
<221> modified_base
<222> (4)
<223> C= 5-methyl-C

<220>
<221> modified_base
<222> (8)

IDT Exhibit 1020.6307

<223> C= 5-methyl-C

<220>
<221> modified_base
<222> (11)..(15)
<223> C=5-methyl-C

<400> 14
atgcattctg cccccaagga 20

<210> 15
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Novel Sequence

<220>
<221> modified_base
<222> (4)
<223> C= 5-methyl-C

<220>
<221> modified_base
<222> (8)
<223> C=5-methyl-C

<220>
<221> modified_base
<222> (11)..(12)
<223> C= 5-methyl-C

<400> 15
atgcattctg ccaaaaagga 20

<210> 16
<211> 11
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Combined DNA/RNA Molecule:
Oligonucleotide

<220>

IDT Exhibit 1020.6308

<223> Description of Artificial Sequence: Novel Sequence

<400> 16
ggctguctgc g 11

<210> 17
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Novel Sequence

<400> 17
ttttttttt ttttttttt 19

<210> 18
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Combined DNA/RNA Molecule:
Oligonucleotide

<220>
<223> Description of Artificial Sequence: Novel Sequence

<400> 18
ttttttttt ttttttttu 19

<210> 19
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Novel Sequence

<400> 19
tgcatccccc aggccaccat 20

<210> 20
<211> 20

IDT Exhibit 1020.6309

<212> RNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Novel Sequence

<400> 20
uggugggcaa ugggcguguu 20

<210> 21
<211> 17
<212> RNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Novel Sequence

<400> 21
gggcgccguc ggugugg 17

<210> 22
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Novel Sequence

<400> 22
acuccaccau aguacacucc 20

<210> 23
<211> 25
<212> RNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Novel Sequence

<400> 23
uggugggcgc cgucggugug ggcaa 25

<210> 24
<211> 17

IDT Exhibit 1020.6310

<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Novel Sequence

<400> 24
ccacaccgac ggcgccc 17

<210> 25
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Novel Sequence

<400> 25
aacacgccca ttgcccacca 20

<210> 26
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Novel Sequence

<220>
<221> modified_base
<222> (2)..(4)
<223> C=5-methyl-C

<220>
<221> modified_base
<222> (8)
<223> C= 5-methyl-C

<220>
<221> modified_base
<222> (12)
<223> C= 5-methyl-C

<220>
<221> modified_base
<222> (15)..(16)

IDT Exhibit 1020.6311

<223> C= 5-methyl-C

<220>
<221> modified_base
<222> (19)
<223> C=5-methyl-C

<400> 26
gcccaagctg gcatccgtca 20

<210> 27
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Novel Sequence

<400> 27
tccgtcatcg ctcctcaggg 20

<210> 28
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Novel Sequence

<400> 28
tcagtaatag gcccacatgg 20

<210> 29
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Novel Sequence

<400> 29
ttcgtcatcg ctcctcaggg 20

<210> 30

IDT Exhibit 1020.6312

<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Novel Sequence

<400> 30
tcagtaatag ccgccgcccc acatgg 26

<210> 31
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Novel Sequence

<400> 31
tcagtaatag ccgccccaca tgc 23

<210> 32
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Novel Sequence

<400> 32
atgcattctg ccccaagga 19

<210> 33
<211> 47
<212> RNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Novel Sequence

<400> 33
ggugguggug ggcgccgucg gugugggcaa gagugcgcug accaucc 47

IDT Exhibit 1020.6313

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US00/26729

A. CLASSIFICATION OF SUBJECT MATTER

IPC(7) :C12Q 1/68; C12P 19/34; C07H 21/02, 21/04, 21/00
US CL :436/6, 91.1, 91.31, 91.5, 455; 536/23.1, 24.5, 25.3

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 436/6, 91.1, 91.31, 91.5, 455; 536/23.1, 24.5, 25.3

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

WEST, Dialog

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	FRIER et al. The ups and downs of nucleic acid duplex stability: structure-stability studies on chemically-modified DNA:RNA duplexes. Nucleic Acids Res. 1997, Vol. 25, No. 22, pages 4429-4443, see entire text.	1-28

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

*	Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be of particular relevance	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E"	earlier document published on or after the international filing date	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&"	document member of the same patent family
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search 03 NOVEMBER 2000	Date of mailing of the international search report 05 DEC 2000
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer Cynthia Lawrence for JANE ZARA Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet) (July 1998)★

Espacenet

Bibliographic data: WO2006086668 (A2) — 2006-08-17

COMPOSITIONS AND METHODS EMPLOYING 5'-PHOSPHATE-DEPENDENT NUCLEIC ACID EXONUCLEASES

Inventor(s): JENDRISAK JEROME [US]; MEIS JUDITH [US]; MEIS RONALD [US]; RADEK AGNES [US]; DAHL GARY [US] ± (JENDRISAK, JEROME, ; MEIS, JUDITH, ; MEIS, RONALD, ; RADEK, AGNES, ; DAHL, GARY)

Applicant(s): EPICT TECHNOLOGIES [US]; JENDRISAK JEROME [US]; MEIS JUDITH [US]; MEIS RONALD [US]; RADEK AGNES [US]; DAHL GARY [US] ± (EPICENTRE TECHNOLOGIES, ; JENDRISAK, JEROME, ; MEIS, JUDITH, ; MEIS, RONALD, ; RADEK, AGNES, ; DAHL, GARY)

Classification: **- international:** ***C12Q1/68***
- cooperative: **C12N15/1006**; **C12N15/1096**; **C12Q1/6806** more

Application number: WO2006US04798 20060209 Global Dossier

Priority number(s): US20050651409P 20050209 ; US20050685367P 20050527

Also published as: WO2006086668 (A3) US2006240451 (A1)

Abstract of WO2006086668 (A2)

The present invention relates to compositions and methods employing 5'- phosphate-dependent nucleic acid exonucleases. In particular, the present invention provides kits and methods employing 5 '-phosphate-dependent nucleic acid exonucleases for selective enrichment, isolation and amplification of a particular set of desired nucleic acid molecules from samples that also contain undesired nucleic acid molecules for a variety of uses. In preferred embodiments, the desired nucleic acid molecules comprise prokaryotic and/or eukaryotic mRNA.

IDT Exhibit 1020.6315

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
17 August 2006 (17.08.2006)

PCT

(10) International Publication Number
WO 2006/086668 A2

(51) International Patent Classification:
C12Q 1/68 (2006.01)

(21) International Application Number:
PCT/US2006/004798

(22) International Filing Date: 9 February 2006 (09.02.2006)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/651,409 9 February 2005 (09.02.2005) US
60/685,367 27 May 2005 (27.05.2005) US

(71) Applicant *(for all designated States except US)*: **EPICENTRE TECHNOLOGIES** [US/US]; 726 Post Road, Madison, WI 53713 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **JENDRISAK, Jerome** [US/US]; 2721 Mason Street, Madison, WI 53705 (US). **MEIS, Judith** [US/US]; 2728 Tami Trail, Fitchburg, WI 53129 (US). **MEIS, Ronald** [US/US]; 2728 Tami Trail, Fitchburg, WI 53129 (US). **RADEK, Agnes** [HU/US]; 6510 Piedmont Road, Madison, WI 53711 (US). **DAHL, Gary** [US/US]; 2131 Kendall Avenue, Madison, WI 53726 (US).

(74) Agents: **CASIMIR, David, A.** et al.; Medlen & Carroll, LLP, 101 Howard Street, Suite 350, San Francisco, CA 94105 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, LY, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 2006/086668 A2

(54) Title: COMPOSITIONS AND METHODS EMPLOYING 5'-PHOSPHATE-DEPENDENT NUCLEIC ACID EXONUCLEASES

(57) Abstract: The present invention relates to compositions and methods employing 5'- phosphate-dependent nucleic acid exonucleases. In particular, the present invention provides kits and methods employing 5 '-phosphate-dependent nucleic acid exonucleases for selective enrichment, isolation and amplification of a particular set of desired nucleic acid molecules from samples that also contain undesired nucleic acid molecules for a variety of uses. In preferred embodiments, the desired nucleic acid molecules comprise prokaryotic and/or eukaryotic mRNA.

IDT Exhibit 1020.6316

COMPOSITIONS AND METHODS EMPLOYING 5'-PHOSPHATE-DEPENDENT NUCLEIC ACID EXONUCLEASES

The present invention claims priority to U.S. Provisional Patent Application Serial Numbers 60/651,409, filed February 9, 2005 and 60/685,367 filed May 27, 2005, each of which is herein incorporated by reference in its entirety.

FIELD OF THE INVENTION

The present invention relates to compositions and methods employing 5'-phosphate-dependent nucleic acid exonucleases. In particular, the present invention provides kits and methods for employing 5'-phosphate-dependent nucleic acid exonucleases for selective enrichment, isolation and amplification of a particular set of desired nucleic acid molecules from samples that also contain undesired nucleic acid molecules for a variety of uses. In preferred embodiments, the desired nucleic acid molecules comprise prokaryotic and/or eukaryotic mRNA.

BACKGROUND OF THE INVENTION

With the enormous increase in the amount of bacterial genome sequence information over the last several years and the complete sequencing of a number of microbial genomes, the molecular biology field has entered a post-genomic era (Pang et al., Microbiol. Immunol., 48:91, 2004). The investigation of gene expression at the transcriptional level is of basic interest in microbial ecology and pathogenicity studies. This investigation will bridge the gap between structural and functional diversity. Expression analysis is being used to identify gene functions and metabolic pathways in many organisms, including humans, yeast, *Drosophila*, mice, and bacteria.

A major challenge in prokaryotic expression analysis is the preparation and analysis of prokaryotic mRNA. Oligo(dT) selection for poly(A) tails has long been used for isolating mRNAs from eukaryotic sources. However, the lack of relatively stable poly(A) tails, and their short half-lives and low quantity for bacteria make isolation of bacterial mRNA difficult (Adel et al., Nature. Biotechnol. 18:679, 2000) and Coller et al., Proc. Natl. Acad. Sci. U.S.A. 97:3260, 2000). Thus, isolation of mRNA from bacteria has been vitally important, but difficult. There has been an urgent need to develop rapid

IDT Exhibit 1020.6317

and simple methods to purify mRNA from bacteria. Several attempts are described below. However, none have provided sufficient solutions to the problem.

Wendisch et al. (Anal. Biochem. 290:205, 2001); U.S. Pat. No. 6,242,189, herein incorporated by reference in their entireties) developed a method to purify the whole population of cellular mRNAs by polyadenylation with *E. coli* poly(A) polymerase in crude cell extracts obtained by mechanical lysis. However, this method does not select for only mRNA, so other RNA molecules can be polyadenylated in addition to mRNA.

Affymetrix Inc. (Rosenow et al., Nucleic Acids Res. 29(22):e112, 2001) obtained enriched mRNA from *E. coli* with a series of enzymatic steps that specifically eliminate the 16S and 23S rRNA species in the total RNA. Reverse transcriptase and primers specific for 16S and 23S rRNA are used to synthesize complementary DNAs. Then rRNA is removed enzymatically by treatment with RNase H, which specifically digests RNA within an RNA:DNA hybrid. The cDNA molecules are then removed by DNase I digestion and the enriched mRNA is purified on QIAGEN RNeasy columns. The enriched mRNA will contain mRNAs, tRNAs, 5S rRNA, and other small RNAs. This method is complex and difficult to handle, and introduces the risk of mRNA losses due to mispriming of mRNA with rRNA primers.

Ambion Company developed the MICROBE EXRESS Bacterial mRNA Purification Kit that employs a modified capture hybridization approach, to remove abundant 16S and 23S ribosomal RNAs (rRNA) from purified total RNA and enrich bacterial mRNA. Briefly, purified RNA is incubated with the Capture Oligonucleotide Mix in Binding Buffer. Magnetic beads, derivatized with an oligonucleotide that hybridizes to the capture oligonucleotide, are then added to the mixture and allowed to hybridize. The magnetic beads, with 16S and 23S rRNAs attached, are pulled to the side of the tube with a magnet. The enriched RNA in the supernatant is removed and precipitated with ethanol. The enriched mRNA will contain mRNAs, tRNAs, 5S rRNA, and other small RNAs. The process is somewhat tedious and cumbersome, and cannot be applied to all species of bacteria. U.S. Pat. Appln. No. 2003/0175709 likewise describes a method for depleting unwanted RNA species using capture probes that hybridize to bridging oligonucleotides configured to bind to capture beads and target sequences.

IDT Exhibit 1020.6318

Pang et al., *supra*, demonstrate that magnetic capture-hybridization methods may be used for the purification of bacterial mRNAs, using biotin-labeled oligonucleotides as capture probes specific for 5S, 16S and 23S rRNA of bacteria. Ribosomal RNAs hybridize to biotin-labeled oligonucleotide capture probes that are fixed to streptavidin-coated paramagnetic beads. The mRNA remains in the supernatant and is recovered by ethanol precipitation. While this method enriched mRNA, improvements in efficiency are needed.

Thus, the art is in need of improved compositions and methods that provide high efficiency and ease of use for the enrichment and purification of bacterial mRNA for a variety of applications. What is needed are rapid and simple methods for isolation and purification of bacterial mRNA that avoid the tedious and cumbersome use of magnetic beads or purification columns, and/or the synthesis of ribosomal cDNA and subsequent RNase H digestion.

What is further needed in the art are compositions and methods for preparation of bacterial mRNA, whereby said compositions and methods remove 3' mRNA fragments that do not have the 5'-end of primary mRNA transcripts (such as 3' mRNA fragments resulting from exposure of primary mRNA transcripts to a ribonuclease or to physical forces such as shearing).

What is further needed are compositions and methods to prepare cDNA from bacterial mRNA and to amplify bacterial mRNA for gene expression studies such as, but not limited to, for microarray analyses. Moreover, what is needed are compositions and methods for preparation of cDNA from bacterial mRNA, whereby the cDNA is synthesized from full-length mRNA, meaning from mRNA sequences that are not truncated at the 5'-end or the 3'-end.

What is further needed are compositions and methods to isolate bacterial mRNA, to prepare cDNA from bacterial mRNA, and to amplify bacterial mRNA for gene expression studies from samples in which the bacteria are associated with cells of one or more other prokaryotic and/or eukaryotic organisms. By way of example, but not of limitation, what is needed are compositions and methods for purifying bacterial mRNA, making bacterial cDNA, and amplifying bacterial mRNA from Rhizobium nitrogen-fixing bacteria in the roots of legumes, from biofilms such as in the human mouth, or

IDT Exhibit 1020.6319

from pathogenic bacteria or mycoplasma in association with a plant, animal, human, or fungal host. Moreover, what is needed are improved compositions and methods for isolating mRNA, making cDNA, and amplifying mRNA from both pathogen (or symbiont) and host cells at the same time in order to simultaneously study gene expression in both organisms that are associated in pathogen-host (or symbiotic) interactions (or even to simultaneously study gene expression in multiple bacterial species associated in biofilms of any type, or in multiple cells of a single organism or of multiple organisms in any type of association).

What is also needed are compositions and methods that enable synthesis of cDNA and amplification of mRNA from bacteria that are difficult to grow in cell culture.

What is also needed in the art are improved compositions and methods to purify eukaryotic mRNA. What is needed are rapid and simple methods for isolation and purification of eukaryotic mRNA that avoid the tedious and cumbersome use of magnetic beads, purification columns, or membranes that bind polyadenylated nucleic acids (e.g., with oligo(dT)). Moreover, what is needed are compositions and methods for preparation of eukaryotic mRNA, whereby said compositions and methods remove mRNA fragments that do not have the 5'-end of a full-length mRNA transcript (such as 3' mRNA fragments resulting from exposure of the mRNA to a ribonuclease or to physical forces such as shearing and/or heat).

What is further needed are improved compositions and methods to prepare cDNA from eukaryotic mRNA and to amplify eukaryotic mRNA for gene expression studies such as, but not limited to, for microarray analyses. Moreover, what is needed are compositions and methods for preparation of cDNA from eukaryotic mRNA, whereby the cDNA is synthesized from full-length mRNA, meaning from mRNA sequences that are not truncated at the 5'-end or the 3'-end. What is further needed are improved compositions and methods for isolating mRNA, making cDNA, and amplifying mRNA from eukaryotes in a high-throughput manner.

What is also needed are improved kits and methods for analyzing gene expression in biological samples containing eukaryotic mRNA or both eukaryotic and prokaryotic mRNA. For example, it can be very difficult to obtain mRNA, make cDNA, and amplify mRNA in biological samples that contain degraded RNA, such as, but not limited to

IDT Exhibit 1020.6320

slides of formalin-fixed paraffin-embedded ("FFPE") tissue sections. Samples of cells from which it is desired to profile gene expression can be obtained from such sections by various methods known in the art. One method that can be used to obtain sample of cells from such tissue sections is laser capture, such as but not limited to, laser capture microdissection, using methods known in the art. Instruments for laser capture are available commercially from Arcturus, PALM, and other sources. Methods are also known in the art for isolating total RNA from samples from FFPE tissue sections. One such method that can be used is described by Acturus in the product literature for their Paradise™ Kit. Other methods are also known in the art that can be used.

However, once total RNA is obtained, it can be difficult to obtain good quality mRNA for gene expression analysis by methods known in the art, such as but not limited to, analysis by microarrays. In part, this is true because many methods rely on use of oligo(dT) resins or membranes to isolate poly(A) containing RNA, which is a common method for isolating polyadenylated mRNA. The problem is that much of the mRNA may be degraded into fragments that do not have a poly(A) tail and will not be isolated using such oligo(dT) columns or membranes. Also, the mRNA fragments that have a poly(A) tail and therefore are obtained using the oligo(dT) binding method may not contain the sequences that are complementary to many of the oligos on an array or microarray. Therefore, the gene expression analysis can be difficult to interpret. It would be desirable to obtain mRNA samples more easily, particularly for making cDNA, for RNA amplification and other methods of amplification, and especially, for obtaining better and more informative gene expression data than current methods.

SUMMARY OF THE INVENTION

The present invention relates to compositions and methods employing 5'-phosphate-dependent nucleic acid exonucleases. In particular, the present invention provides kits and methods employing 5'-phosphate-dependent nucleic acid exonucleases for selective enrichment, isolation and amplification of a particular set of desired nucleic acid molecules from samples that also contain undesired nucleic acid molecules for a variety of uses. In preferred embodiments, the desired nucleic acid molecules comprise prokaryotic and/or eukaryotic mRNA.

IDT Exhibit 1020.6321

The present invention comprises methods for using a 5' exoribonuclease for a number of applications related to biological research on gene expression, diagnostics, and human and animal therapeutics.

One preferred method of the invention comprises a method for enriching for an RNA having a 5'-triphosphate or a 5'-cap in a biological sample comprising prokaryotic RNA, eukaryotic RNA or both prokaryotic and eukaryotic RNA and at least one undesired nucleic acid, the method comprising treating the sample with purified 5' exoribonuclease under conditions in which the 5' exoribonuclease is active and for sufficient time so that the undesired nucleic acid is digested and the sample is enriched for RNA having a 5'-triphosphate or a 5'-cap.

In different embodiments of this method, the RNA having a 5'-triphosphate or a 5'-cap is selected from the group consisting of: (i) prokaryotic mRNA; (ii) eukaryotic mRNA, including polyadenylated and non-polyadenylated eukaryotic mRNA; (iii) a mixture of both prokaryotic and eukaryotic mRNA; (iv) eukaryotic snRNA; (v) eukaryotic pre-micro RNA; and (vi) prokaryotic or eukaryotic primary RNA transcripts of unknown function.

Another preferred embodiment of this method comprises treating the biological sample with a polynucleotide kinase to phosphorylate the RNA having 5'-hydroxyl groups prior to treating the sample with purified 5' exoribonuclease.

In one embodiment of the method, the RNA having a 5'-triphosphate or a 5'-cap comprises eukaryotic mRNA or both prokaryotic and eukaryotic mRNA, and the method additionally comprises binding the RNA to an oligo(dT) or oligo(dU) resin, membrane or other surface to which oligo(dT) or oligo(dU) is attached and eluting the bound RNA so as to obtain a solution containing polyadenylated eukaryotic mRNA.

In another embodiment of the method, the RNA having a 5'-triphosphate or a 5'-cap comprises prokaryotic mRNA, eukaryotic mRNA or both prokaryotic and eukaryotic mRNA and the undesired nucleic acid comprises prokaryotic rRNA, eukaryotic rRNA or both prokaryotic and eukaryotic rRNA of a size having a svedburg unit greater than about 10S. In specific embodiments of this method the rRNA is selected from the group consisting of prokaryotic rRNA having a svedburg unit of about 16S or 23S and eukaryotic rRNA having a svedburg unit of about 18S, 26S or 28S.

IDT Exhibit 1020.6322

Another embodiment of the invention is a method wherein the RNA having a 5'-triphosphate or a 5'-cap comprises prokaryotic mRNA, eukaryotic mRNA, or both prokaryotic and eukaryotic mRNA, and wherein the method additionally comprises precipitation of said mRNA having a 5'-triphosphate or a 5'-cap with a solution of LiCl and ethanol at concentrations and under conditions in which tRNA and 5S rRNA are not precipitated. In some embodiments of this method, the RNA having a 5'-triphosphate or a 5'-cap comprises prokaryotic mRNA, eukaryotic mRNA, or both prokaryotic and eukaryotic mRNA and the undesired nucleic acid comprises prokaryotic rRNA, eukaryotic rRNA or both prokaryotic and eukaryotic rRNA of a size having a svedburg unit greater than about 10S. In specific embodiments, the rRNA is selected from the group consisting of prokaryotic rRNA having a svedburg unit of about 16S or 23S and eukaryotic rRNA having a svedburg unit of about 18S, 26S or 28S.

Still another method of the invention that uses RNA obtained following 5' exoribonuclease treatment comprises: (i) contacting the sample enriched for RNA having a 5'-triphosphate or a 5'-cap with a poly(A) polymerase under conditions so that polyadenylated RNA is obtained; and (ii) contacting the polyadenylated RNA with an oligo(dT)-containing primer and an RNA-dependent DNA polymerase under conditions that cDNA complementary to the RNA is obtained. In some specific embodiments, the oligo(dT)-containing primer comprises a T7-type RNA polymerase promoter selected from the group consisting of T7 RNA polymerase, T3 RNA polymerase, and SP6 RNA polymerase. In one embodiment of this method, the biological sample is treated with a polynucleotide kinase to phosphorylate the RNA having 5'-hydroxyl groups prior to treating the sample with purified 5' exoribonuclease. In some other specific embodiments of this method, the RNA having a 5'-triphosphate or a 5'-cap comprises prokaryotic mRNA, eukaryotic mRNA, or both prokaryotic and eukaryotic mRNA and the undesired nucleic acid comprises prokaryotic rRNA, eukaryotic rRNA or both prokaryotic and eukaryotic rRNA of a size having a svedburg unit greater than about 10S. In specific embodiments, the rRNA is selected from the group consisting of prokaryotic rRNA having a svedburg unit of about 16S or 23S and eukaryotic rRNA having a svedburg unit of about 18S, 26S or 28S.

IDT Exhibit 1020.6323

In still another method, wherein the biological sample comprises eukaryotic RNA, and wherein the method additionally comprises contacting the sample enriched for RNA having a 5'-triphosphate or a 5'-cap with an oligo(dT)-containing primer and an RNA-dependent DNA polymerase under conditions that cDNA complementary to the RNA is obtained. In some specific embodiments, the oligo(dT)-containing primer comprises a T7-type RNA polymerase promoter selected from the group consisting of T7 RNA polymerase, T3 RNA polymerase, and SP6 RNA polymerase. In some of these embodiments, the method additionally comprises contacting the cDNA complementary to the RNA with one or more DNA or RNA primers and a DNA polymerase and incubating under conditions so as to obtain double-stranded cDNA having a functional promoter for the T7-type RNA polymerase and synthesizing RNA therefrom.

In still other embodiments, the method additionally comprises contacting the cDNA complementary to the RNA with an oligonucleotide sequence tag template and a DNA polymerase and incubating under conditions so as to obtain cDNA having a sequence tag on it the 3'-end. In some embodiments, the method additionally comprises contacting the cDNA having the sequence tag on its 3'-end with a DNA polymerase and an oligonucleotide primer, wherein the either the sequence tag or the oligonucleotide primer, or the combination of both the sequence tag and the oligonucleotide primer encodes the complete sequence of a T7-type RNA polymerase promoter, and incubating under conditions wherein double-stranded cDNA having a functional T7-type RNA polymerase promoter is obtained. In those cases, the method additionally comprises contacting the cDNA with the T7-type RNA polymerase under conditions so as to synthesize RNA using said T7-type RNA polymerase promoter. In some other specific embodiments of these methods, the biological sample is treated with a polynucleotide kinase to phosphorylate the RNA having 5'-hydroxyl groups prior to treating the sample with purified 5' exoribonuclease.

In preferred embodiments, the compositions, kits, and methods of the present invention employ Xrn1p/5' Exoribonuclease I and functional variants, homologues, and equivalents. Xrn1p/5' exoribonuclease I is a magnesium-dependent, 5'-to-3', processive exoribonuclease that acts preferentially on RNA substrates with a 5' phosphate (Stevens, Biochem. Biophys. Res. Commun. 81:656, 1978); Stevens, Biochem. Biophys. Res.

IDT Exhibit 1020.6324

Commun. 86:1126, 1979), Stevens, J. Biol. Chem. 255:3080, 1980); and Stevens and Maupin, Nucleic Acids Res. 15:695, 1987). Nucleic acid molecules with a cap, a triphosphate group, or an hydroxyl group on their 5'-termini (or 5'-ends) are not substantially digested by the enzyme. The exoribonuclease I is not inhibited by proteinaceous RNase inhibitors such as RNASIN (Promega, Madison, WI) or PRIME RNase Inhibitor (Eppendorf, Brinkmann, Westbury, NY). In some embodiments of the present invention, the 5'-phosphate-dependent nucleic acid exonuclease is obtained as described by Stevens (J. Biol. Chem. 255:3080, 1980). In some other embodiments, the 5'-phosphate-dependent nucleic acid exonuclease is obtained by cloning a gene for an exonuclease, such as but not limited to a Saccharomyces Xrn I gene, in a vector, expressing the gene in a host cell, and purifying the expressed protein from cultures of said host cells. For example, but without limitation, in some embodiments, the 5'-phosphate-dependent nucleic acid exonuclease is obtained from recombinant yeast by following the protocol of Johnson and Kolodner (J. Biol. Chem. 266:14046, 1991). In other embodiments, the 5'-phosphate-dependent nucleic acid exonuclease is obtained by cloning the gene for an exonuclease in a vector, expressing the gene in a heterologous host cell, such as but not limited to an *E. coli* host cell, and purifying the expressed protein from cultures of said host cells. Such a composition from recombinant source finds use for more efficient and less expensive production, as well as a higher purity and more consistent quality of 5'-phosphate-dependent nucleic acid exonuclease for use in the methods and kits of the present invention. The expression vector may further express other sequences that assist in the purification, expression, detection, or use of the 5'-phosphate-dependent nucleic acid exonucleases.

The exonucleases of the present invention find use in a number of applications including, but not limited to: analysis of the 5' ends of RNA molecules; removal of ribosomal RNA from RNA preparations; enrichment of prokaryotic and/or eukaryotic mRNA; enrichment of mixtures of prokaryotic and/or eukaryotic mRNA; preparation of prokaryotic and/or eukaryotic mRNA for cDNA synthesis; synthesis of amplified antisense or sense RNA by preparing prokaryotic mRNA using exoribonuclease, then polyadenylating the RNA with poly(A) polymerase, then making cDNA joined to an RNA polymerase promoter, and transcribing the cDNA from the promoter *in vitro*;

IDT Exhibit 1020.6325

enrichment of 5'-capped or 5'-triphosphorylated small nuclear RNAs (snRNA); removal of 5'-phosphorylated splice template oligos (or template switch oligos) and other oligos (e.g., ligation splints) from *in vitro* transcription reactions, including antisense or sense RNA amplification reactions, in order to reduce nonspecific background transcription in RNA amplification and other reactions; removal of linear RNA from circular RNA preparations; removal of linear first-strand cDNA molecules from circular first-strand cDNA molecules, diagnostic applications; and the like.

For example, in some embodiments, the present invention provides a method for isolating a nucleic acid molecule of interest in a mixture of nucleic acids, comprising the steps of: a) providing a 5'-to-3' exoribonuclease and a biological sample comprising a mixture of nucleic acids, said mixture of nucleic acids comprising a nucleic acid molecule of interest (i.e., a desired nucleic acid) and an undesired nucleic acid having a 5'-phosphate; and b) exposing the 5'-to-3' exoribonuclease to the biological sample under conditions such that the exoribonuclease digests the undesired nucleic acid so as to enrich the sample for the nucleic acid molecule of interest. The present invention is not limited by the nature of the 5'-phosphate-dependent nucleic acid exonuclease. In preferred embodiments, the 5'-phosphate-dependent nucleic acid exonuclease includes, but is not limited to, Xrn1p/5' exoribonuclease 1, fragments of Xrn1p/5' exoribonuclease 1 (e.g., that have exoribonuclease activity), and Xrn1p/5' exoribonuclease 1 variants (e.g., containing conserved amino acid changes, truncations, chimeras, etc.). A 5' exonuclease of the present invention comprises any 5'-exonuclease that has greater than 20-fold more 5'-to-3' exonuclease activity for a single-stranded RNA substrate that has a 5'-monophosphorylated terminus than for the same RNA substrate that has a 5'-triphosphorylated or 5'-capped terminus. That is, a 5' exonuclease of the present invention comprises any 5'-to'3' exonuclease that has a relative activity in digesting a particular defined-sequence single-stranded RNA substrate with a 5'-triphosphate or a 5'-cap that is less than 5% of the activity of an RNA substrate having the same sequence but with a 5'-monophosphate. Enzyme activity of a 5' exonuclease of the invention can be measured using a number of different methods. Without limitation, suitable methods that can be used for assaying activity and determining relative activity using RNA substrates with a

IDT Exhibit 1020.6326

5'-triphosphate, a 5'-cap, or a 5'-monophosphate are described by Stevens and Poole (J. Biol. Chem. 270: 16063, 1995).

The present invention is also not limited by the nature of the sample. Samples include, but are not limited to, cell lysates, provided that nucleases that can affect nucleic acids of interest are removed or inhibited in said lysates, mixtures of nucleic acids (e.g., unpurified, partially purified, etc.), environmental samples, etc. Samples can comprise nucleic acids from a 1-cell organism, or if an organism comprises multiple cells, from one or more cells. Samples comprising nucleic acids from multiple cells can comprise cells from one or more types of cells, including cells from different organisms, and/or different cells from the same organism. The present invention finds use with prokaryotic nucleic acid, with eukaryotic nucleic acid, or with a mixture of both eukaryotic and prokaryotic nucleic acid.

In some preferred embodiments of the present invention, the sample comprises (or contains) RNA. A sample of the invention that comprises or contains RNA can contain total RNA from any source, which RNA can be obtained using, for example, a MasterPure RNA Purification Kit, an ArrayPure Nano-Scale RNA Purification Kit, or a MasterPure Yeast RNA Purification Kit (all of which are from EPICENTRE), or using a kit from another commercial source, or using a "home-brew" method known in the art. Alternatively, in some embodiments a sample of the invention that comprises or contains RNA can contain a subfraction of total RNA obtained by any method known in the art, such as, but without limitation, a subfraction based on size (e.g., by purification on an agarose or polyacrylamide gel, or by column purification, including by HPLC), or a subfraction obtained by salt precipitation (e.g., using precipitation with 0.5-2.5 M LiCl (Barlow, JJ et al., Biochem. Biophys. Res. Comm. 13: 61, 1963); Cathala, G et al, DNA 2: 329, 1983) or 2.5 M ammonium acetate). In some embodiments, a sample comprising RNA can also contain DNA. In some preferred embodiments, the nucleic acid molecules of interest comprise mRNA having a 5'-triphosphate or a 5'-cap. However, the nucleic acid molecule of interest (or desired nucleic acid) can comprise any nucleic acid molecule that is resistant to digestion by the 5'-phosphate-dependent nucleic acid exonuclease used, including, but not limited to nucleic acid molecules having a 5'-triphosphate or a 5'-cap, a 5'-hydroxyl group, or circularized nucleic acid molecules.

IDT Exhibit 1020.6327

The undesired nucleic acid can comprise any nucleic acid molecule that is digested by the 5'-phosphate-dependent nucleic acid exoribonuclease used, including, but not limited to nucleic acid molecules having a 5'-monophosphate or a 5'-diphosphate group, which are substrates for the enzyme. In some preferred embodiments, the undesired nucleic acid comprises rRNA that is greater than approximately 200 nucleotides (e.g., without limitation, containing each of 16S and 23S prokaryotic rRNA or 18S and 28S human, mouse, drosophila, or Xenopus rRNA or 18S and 26S yeast rRNA).

In some preferred embodiments, little or no detectable level of the desired nucleic acid molecule of interest is digested by the 5'-phosphate-dependent nucleic acid exonuclease. However, the invention is not limited by the amount of the desired nucleic acid molecule of interest that is digested so long as a substantial amount of the undesired nucleic acid is digested and sufficient amount of the desired nucleic acid remains for the intended purpose, which can vary for different purposes. For example, less of the desired RNA may be needed from some samples and/or applications (e.g., for RNA amplification of mRNA from total RNA in large samples) than are needed from other samples and/or applications (e.g., preparing a cDNA library from only a small number of cells). In preferred embodiments, a substantial amount of the undesired nucleic acid is digested by the 5' exonuclease of the invention, which means that at least 50% (e.g., 60%, 70%, 80%, 90%, 95%, 98%, 99%) of the starting amount (i.e., the amount present prior to treatment with the exonuclease) of the undesired nucleic acid is digested.

In some preferred embodiments, the present invention provides a method for isolating prokaryotic mRNA (e.g., to conduct gene expression analysis) comprising the steps of: a) providing a sample comprising RNA from a prokaryotic organism, said sample containing desired mRNA having a 5'-triphosphate group and at least one undesired nucleic acid; and a 5'-phosphate-dependent nucleic acid exonuclease; b) treating the sample with the 5'-phosphate-dependent nucleic acid exonuclease under conditions such that said undesired nucleic acid is substantially digested and said desired mRNA is not digested; and c) recovering the mRNA (e.g., for additional analysis or use). In some embodiments, the 5'-phosphate-dependent nucleic acid exonuclease used in the method is provided in a kit with buffers, instructions, and optionally, with appropriate control samples, and the like.

IDT Exhibit 1020.6328

In some preferred embodiments, the present invention provides a method for isolating prokaryotic mRNA molecules that have the 5'-ends of primary mRNA transcripts (e.g., to conduct gene expression analysis) comprising the steps of: a) providing a sample comprising RNA from a prokaryotic organism, said sample containing desired mRNA having a 5'-triphosphate group and at least one undesired nucleic acid; a polynucleotide kinase; and a 5'-phosphate-dependent nucleic acid exonuclease; b) treating the sample with the polynucleotide kinase in the presence of ATP and under conditions such that nucleic acid molecules having an hydroxyl group at their 5'-termini are phosphorylated; c) treating the sample with the 5'-phosphate-dependent nucleic acid exonuclease under conditions such that RNA molecules having a 5'-monophosphate group are substantially digested; and d) recovering the mRNA (e.g., for additional analysis or use). In some embodiments, the 5'-phosphate-dependent nucleic acid exonuclease used in the method is provided in a kit with polynucleotide kinase (e.g., T4 polynucleotide kinase), ATP, buffers, instructions, and optionally, with appropriate control samples, and the like.

In some embodiments, the present invention provides methods and kits for producing cDNA copies of prokaryotic mRNA (e.g., for making cDNA libraries). For example, the present invention provides a method for preparing prokaryotic cDNA, comprising the steps of a) providing a sample comprising RNA from a prokaryotic organism, said sample containing desired mRNA having a 5'-triphosphate group and at least one undesired nucleic acid; and a 5'-phosphate-dependent nucleic acid exonuclease; b) contacting the sample with the 5'-phosphate-dependent nucleic acid exonuclease under conditions such that undesired nucleic acids in the sample are substantially digested but mRNA molecules in the sample having a 5'-triphosphate are not digested; and c) producing cDNA copies of the mRNA. The present invention is not limited by the manner in which the cDNA copies are produced.

In some preferred embodiments, the method of making cDNA copies comprises: i) contacting the mRNA molecules with a poly(A) polymerase to produce mRNA molecules having a 3'-poly(A) tail; and ii) contacting the mRNA molecules having a 3'-poly(A) tail with a primer that anneals to said tail (e.g., an oligo(dT)-containing primer) and extending the primer with an RNA-dependent DNA polymerase under conditions

IDT Exhibit 1020.6329

whereby cDNA is obtained. Kits of the invention for conducting such methods may have any one or more reagents (e.g., primers), control samples, buffers, etc. useful in the method. For example, preferred kits comprise: a) a 5'-phosphate-dependent nucleic acid exonuclease; b) a poly(A) polymerase (e.g., without limitation, an *E. coli* poly(A) polymerase encoded by the *pcn*B gene); and c) an RNA-dependent DNA polymerase (e.g., an AMV reverse transcriptase; an MMLV reverse transcriptase; SuperScript I, SuperScript II, SuperScript III, or AMV ThermoScript reverse transcriptase (INVITROGEN); or MonsterScript reverse transcriptase (EPICENTRE).

In some other embodiments, the method of making cDNA copies comprises adding a sequence tag template to the 3'-end of the mRNA molecules using a method as described in U.S. Patent Application No. 2005/0153333, incorporated herein by reference. Thus, one embodiment for making cDNA copies of mRNA molecules comprises: i) contacting the mRNA molecules with a mixture of tagging oligonucleotides, each tagging oligonucleotide having a 5'-portion comprising the same sequence tag template, a 3'-portion comprising a random sequence and a blocked 3' terminus, under conditions in which one of the tagging oligonucleotides anneals with each mRNA molecule and the mRNA molecules are extended with a nucleic acid polymerase to form mRNA molecules having a sequence complementary to the sequence tag template on their 3'-termini; and ii) contacting the mRNA molecules having a sequence complementary to the sequence tag template on their 3'-termini with a primer that anneals to said sequence that is complementary to the sequence tag template and extending the primer with an RNA-dependent DNA polymerase under conditions whereby cDNA is obtained. In some aspects of this embodiment, the nucleic acid polymerase used to extend the mRNA molecules using the sequence tag template as a template in step (i) is the same enzyme as the RNA-dependent DNA polymerase used to obtain cDNA in step (ii), wherein, in other aspects of this embodiment, the nucleic acid polymerase used to extend the mRNA molecules using the sequence tag template as a template in step (i) is different than the RNA-dependent DNA polymerase used to obtain cDNA in step (ii). Kits for synthesizing cDNA according to the embodiment of the invention comprise: a) tagging oligonucleotides; b) a primer that anneals to the sequence that is complementary to the sequence tag template; and c) an RNA-dependent DNA

IDT Exhibit 1020.6330

polymerase (e.g., an AMV reverse transcriptase; an MMLV reverse transcriptase; SuperScript I, SuperScript II, SuperScript III, or AMV ThermoScript reverse transcriptase (INVITROGEN); or MonsterScript reverse transcriptase (EPICENTRE)); and in aspects of this embodiment in which the nucleic acid polymerase that extends the 3'-termini of the mRNA molecules using the sequence that is complementary to the sequence tag template of the tagging oligonucleotide as a template is different from the RNA-dependent DNA polymerase used to obtain cDNA, the kit additionally comprises said nucleic acid polymerase for extending the 3'-termini of the mRNA molecules.

In some preferred embodiments, the present invention provides methods and kits for producing cDNA copies of full-length prokaryotic mRNA (e.g., for making full-length cDNA libraries). For example, the present invention provides a method for preparing prokaryotic cDNA, comprising the steps of a) providing a sample comprising RNA from a prokaryotic organism; a polynucleotide kinase; and a 5'-phosphate-dependent nucleic acid exonuclease; b) contacting the sample with the polynucleotide kinase in the presence of ATP and under conditions such that nucleic acid molecules having an hydroxyl group at their 5'-termini are phosphorylated; c) contacting the sample treated with the polynucleotide kinase with the 5'-phosphate-dependent nucleic acid exonuclease under conditions such that undesired nucleic acids in the sample are digested but mRNA molecules in the sample having a 5'-triphosphate are not digested; and d) producing cDNA copies of the mRNA. The present invention is not limited by the manner in which the cDNA copies are produced.

In some preferred embodiments, the method of making cDNA copies comprises: i) contacting the mRNA molecules with a poly(A) polymerase to produce mRNA molecules having a 3'-poly(A) tail; and ii) contacting the mRNA molecules having a 3'-poly(A) tail with a primer that anneals to said tail and extending the primer with an RNA-dependent DNA polymerase under conditions whereby cDNA is obtained. Kits for conducting such methods may have any one or more reagents (e.g., oligo(dT) primer and/or other primers), control samples, ATP, buffers, etc. useful in the method. For example, preferred kits comprise: a) a 5'-phosphate-dependent nucleic acid exonuclease; b) a polynucleotide kinase (e.g., T4 polynucleotide kinase); c) a poly(A) polymerase; and d) an RNA-dependent DNA polymerase. In some preferred embodiments, the desired

IDT Exhibit 1020.6331

nucleic acid molecules comprise both prokaryotic and eukaryotic mRNA. In some preferred kits in which the RNA in the sample is of sufficient quality for the intended purpose, meaning that a sufficient amount of the RNA comprises the desired mRNA with a 5'-triphosphate or 5'-cap for the intended purpose (which amount can vary for different purposes), the step comprising contacting the sample with the polynucleotide kinase in the presence of ATP and under conditions such that nucleic acid molecules having an hydroxyl group at their 5'-termini are phosphorylated can be omitted and a polynucleotide kinase is not needed in a preferred kit for this method, or if it is included in the kit, the polynucleotide kinase step can be omitted. The invention also envisions that this step can be omitted when a sample is rare or limited in amount and it is desirable to obtain the maximum amount of mRNA and/or cDNA, including degraded portions of mRNA and/or cDNA molecules, for the intended purposes.

Similar methods and kits are provided that are useful for amplifying prokaryotic mRNA. For example, the present invention provides a method for amplifying prokaryotic mRNA, comprising the steps of a) providing a sample comprising RNA from a prokaryotic organism; and a 5'-phosphate-dependent nucleic acid exonuclease; b) contacting the sample with the 5'-phosphate-dependent nucleic acid exonuclease under conditions such that undesired nucleic acids in the sample are substantially digested but mRNA molecules in the sample having a 5'-triphosphate are not digested; and c) amplifying the mRNA. Similar methods and kits are provided that are useful for amplifying prokaryotic mRNA, whereby the methods and kits result in enrichment of sequences corresponding to the 5' portions of mRNA. For example, the present invention provides a method for amplifying prokaryotic mRNA, comprising the steps of a) providing a sample comprising RNA from a prokaryotic organism; a polynucleotide kinase; and a 5'-phosphate-dependent nucleic acid exonuclease; b) contacting the sample with the polynucleotide kinase in the presence of ATP and under conditions such that nucleic acid molecules having an hydroxyl group at their 5'-termini are phosphorylated; c) contacting the sample with the 5'-phosphate-dependent nucleic acid exonuclease under conditions such that undesired nucleic acids in the sample are substantially digested but mRNA molecules in the sample having a 5'-triphosphate are not digested; and d) amplifying the mRNA. The present invention is not limited by the manner in which the

IDT Exhibit 1020.6332

mRNA is amplified. In some embodiments, the mRNA amplified by a method comprising: i) contacting said mRNA molecules with a poly(A) polymerase to produce mRNA molecules having a 3'-poly(A) tail; ii) contacting said mRNA molecules having a 3'-poly(A) tail with a primer that anneals to said tail and extending said primer with an RNA-dependent DNA polymerase under conditions whereby cDNA that is joined to an RNA polymerase promoter sequence is obtained (using any number of methods known in the art; methods are known for joining the promoter to either the 5'-end of the cDNA or the 3'-end of the cDNA, resulting in synthesis of amplified antisense or sense RNA, respectively, by subsequent *in vitro* transcription); iii) producing double stranded cDNA; and iv) contacting said double-stranded cDNA comprising said promoter sequence with an RNA polymerase that binds to said promoter and synthesizes RNA therefrom under conditions whereby RNA corresponding to mRNA in the biological sample is synthesized. Some other embodiments of the present invention use a single-stranded first-strand cDNA template that is functionally joined at either the 3'-end or the 5'-end to a single-stranded promoter that is recognized by an RNA polymerase, such as MiniV RNA polymerase, that can synthesize RNA *in vitro* therefrom; in these embodiments, the method does not comprise a step for synthesis of double-stranded cDNA. Kits for conducting such methods may have any one or more reagents (e.g., primers), control samples, buffers, etc. useful in the method. For example, preferred kits comprise: a) a 5'-phosphate-dependent nucleic acid exonuclease; c) an RNA-dependent DNA polymerase and c) an RNA polymerase (e.g., T7, T3, and SP6 RNA polymerases). Other kits that use single-stranded templates and promoters use an RNA polymerase that synthesize RNA *in vitro* therefrom, such as MiniV RNA polymerase. Other preferred kits further comprise: a) a polynucleotide kinase, such as but not limited to T4 polynucleotide kinase, and ATP; and b) a poly(A) polymerase, such as but not limited to a poly(A) polymerase encoded by an *E. coli pcn*B gene.

In other preferred embodiments, the present invention provides a method for enriching for or isolating mRNA from eukaryotes, comprising the steps of a) providing a 5'-phosphate-dependent nucleic acid exonuclease; and a biological sample comprising a mixture of nucleic acids, said mixture of nucleic acids comprising mRNA of interest and an undesired nucleic acid having a 5'-phosphate; and b) contacting the 5'-phosphate-

IDT Exhibit 1020.6333

dependent nucleic acid exonuclease to the biological sample under conditions such that the exonuclease substantially digests the undesired nucleic acid so as to enrich said sample for the mRNA of interest. Thus, the present invention provides alternative kits and methods for isolating eukaryotic mRNA that do not require binding to oligo(dT) cellulose or other oligo(dT) or oligo(dU) resins or membranes. The method is particularly suitable to high-throughput methods—much more so than methods that employ binding to a resin. However, it will also be understood, based on the description of the present invention, that a 5'-phosphate-dependent exonuclease can be used to obtain a mixture of both prokaryotic and eukaryotic mRNA, and that the eukaryotic mRNA can be separated from prokaryotic mRNA in a mixture by binding polyadenylated eukaryotic mRNA to an oligo(dT) resin, as is known in the art. Still further, although most eukaryotic mRNA is believe to be polyadenylated, some eukaryotic mRNA transcripts may not have a poly(A) tail. The methods disclosed herein can be used to study eukaryotic transcripts that lack a poly(A) tail, which transcripts would not have been detected in many studies that have used oligo(dT) resins and membranes for isolation of the eukaryotic mRNA.

In some preferred embodiments, the present invention provides a method for isolating eukaryotic mRNA comprising the steps of: a) providing a sample from a eukaryotic organism comprising a mixture of nucleic acids, said mixture of nucleic acids comprising desired mRNA of interest and at least one undesired nucleic acid; and a 5'-phosphate-dependent nucleic acid exonuclease; b) treating the sample with the 5'-phosphate-dependent nucleic acid exonuclease under conditions such that the undesired nucleic acid is substantially digested and mRNA having a 5'-cap or a 5'-triphosphate group is not digested; and c) recovering the mRNA (e.g., for additional analysis or use). In some embodiments, the 5'-phosphate-dependent nucleic acid exonuclease used in the method is provided in a kit with buffers, instructions, and optionally, with appropriate control samples, and the like.

One problem in the art that complicates analyses involving mRNA, including analyses involving eukaryotic mRNA, relates to the fact that samples containing the mRNA can include mRNA molecules that are not full-length because they have been degraded: 1) by ribonucleases in the sample, which ribonucleases can be derived either

IDT Exhibit 1020.6334

from the cells from which the mRNA is derived or from contamination of the sample, such as by contamination with a ribonuclease A-type nuclease from human skin or other sources; or 2) by physical forces, such as by shearing forces or by forces due to contact of the mRNA with heat and/or metal ions, such as but not limited to Mn 2+ cations. Many common ribonucleases, such as but not limited to RNase A-type ribonucleases (e.g., including those on human skin) and *E. coli* RNase I, are endoribonucleases that digest RNA to yield products having 5'-hydroxyl and 3'-phosphate groups. Without being bound by theory, it is believed that physical forces can yield a mixture of products having either a 5'-hydroxyl or a 5'-phosphate group. In any case, it is likely that many, if not most, of the degraded mRNA molecules in a sample will have a 5'-hydroxyl group, which means that they will be resistant to digestion with a 5'-exoribonuclease of the present invention. However, since most 5'-exoribonuclease-resistant eukaryotic mRNA molecules will have a 3'-poly(A) tail, these degraded molecules will be substrates for synthesis of oligo(dT)-primed cDNA synthesis and will also be substrates for many antisense RNA amplification methods or sense RNA amplification methods. Thus, these degraded, but 5'-exonuclease-resistant, mRNA molecules can result in a high representation of mRNA sequences corresponding to the 3'-end of the mRNA relative to mRNA sequences corresponding to the 5'-end of the mRNA, which can complicate interpretations of gene expression analyses, and the like. One embodiment of the present invention is a method for assessing the degree of RNA degradation, the method comprising: a) providing a first sample comprising a known quantity of RNA of unknown quality (i.e., of unknown degree of fragmentation or degradation) from any organism or mixture of organisms, the sample containing mRNA; a second control sample comprising a known quantity of substantially undegraded mRNA, preferably, but not necessarily, from the same organism or mixture of organisms; a polynucleotide kinase; and a 5'-phosphate-dependent exonuclease; b) treating one portion of each sample with the polynucleotide kinase in the presence of ATP and under conditions such that nucleic acid molecules having an hydroxyl group at their 5'-termini are phosphorylated; c) contacting the portion of each sample treated with the polynucleotide kinase and a second portion (of the same starting RNA quantity) of each sample that was not treated with the polynucleotide kinase with the 5'-phosphate-dependent nucleic acid exonuclease under conditions such that

IDT Exhibit 1020.6335

undesired nucleic acids in the sample having a 5'-phosphate are substantially digested but mRNA molecules in the sample having a 5'-triphosphate are not digested; and d) producing labeled cDNA copies or labeled amplified RNA of the mRNA in each portion of each sample using an oligo(dT)-containing primer; e), obtaining a ratio for each sample, the ratio having a numerator comprising the amount of labeled cDNA (or amplified RNA) obtained from the polynucleotide kinase-untreated portion and a denominator comprising the amount of labeled cDNA (or amplified RNA) obtained from the polynucleotide kinase-treated portion; and f) comparing the ratio obtained for the first sample with the ratio obtained for the second sample, whereby a higher ratio for the first sample than the ratio for the second sample indicates a higher relative degree of degradation of the RNA in the first sample compared to the second sample. This embodiment of the invention finds utility as an independent method of the invention, or more preferably, it is used as part of a method of the invention for obtaining cDNA or amplified sense or antisense RNA.

As indicated from the above discussion, it is usually desirable to increase the proportion of cDNA molecules and/or amplified sense or antisense RNA molecules that contain sequences corresponding to the 5'-end of the mRNA molecules in the sample (i.e., by removing degraded 3' mRNA fragments) so that a higher percentage of the cDNA and/or the amplified sense or antisense RNA is made from full-length mRNA molecules. Thus, the present invention also provides a method to obtain mRNA from which the degraded mRNA molecules with a 5'-hydroxyl group have been removed, which can improve various analyses involving mRNA, including analyses that involve eukarytic mRNA, prokaryotic mRNA, or both eukaryotic and prokaryotic mRNA. One method of the present invention uses a polynucleotide kinase, such as but not limited to T4 polynucleotide kinase, in order to phosphorylate the 5'-hydroxyl groups of degraded mRNA molecules, thereby making them substrates for digestion by a 5'-exonuclease. Following digestion of the 5'-phosphorylated mRNA with a 5'-exonuclease of the invention, the remaining undigested mRNA molecules will be enriched for full-length mRNA, which will in turn result in cDNA molecules (from cDNA synthesis reactions using an oligo(dT)-containing primer) and amplified sense or antisense RNA molecules (from RNA amplification reactions) that have better representations of sequences

IDT Exhibit 1020.6336

corresponding to the 5'- and 3'-ends of mRNA in the sample. Thus, in some preferred embodiments, the present invention provides a method for isolating desired mRNA of interest comprising the steps of: a) providing a sample containing RNA from an organism, said sample comprising mRNA and at least one undesired nucleic acid; a polynucleotide kinase; and a 5'-phosphate-dependent nucleic acid exonuclease; b) treating the sample with the polynucleotide kinase in the presence of ATP and under conditions such that nucleic acid molecules having an hydroxyl group at their 5'-termini are phosphorylated; c) treating the sample with the 5'-phosphate-dependent nucleic acid exonuclease under conditions such that RNA molecules having a 5'-monophosphate group are substantially digested; and d) recovering the mRNA (e.g., for additional analysis or use). In some embodiments, the 5'-phosphate-dependent nucleic acid exonuclease used in the method is provided in a kit with polynucleotide kinase (e.g., T4 polynucleotide kinase), ATP, buffers, instructions, and optionally, with appropriate control samples, and the like.

In some embodiments, the present invention provides methods and kits for producing cDNA copies of eukaryotic mRNA (e.g., for making cDNA libraries). For example, the present invention provides a method for preparing eukaryotic cDNA, comprising the steps of a) providing a sample comprising RNA from a eukaryotic organism; and a 5'-phosphate-dependent nucleic acid exonuclease; b) contacting the sample with the 5'-phosphate-dependent nucleic acid exonuclease under conditions such that undesired nucleic acids in the sample are substantially digested but mRNA molecules in the sample having a 5'-triphosphate or a 5'-cap are not digested; and c) producing cDNA copies of the mRNA. In some preferred embodiments, the present invention provides methods and kits for producing cDNA copies of full-length eukaryotic mRNA (e.g., for making full-length cDNA libraries). For example, the present invention provides a method for preparing eukaryotic cDNA copies of full-length eukaryotic mRNA, comprising the steps of a) providing a sample comprising RNA from a eukaryotic organism; a polynucleotide kinase; and a 5'-phosphate-dependent nucleic acid exonuclease; b) contacting the sample with the polynucleotide kinase in the presence of ATP and under conditions such that nucleic acid molecules having an hydroxyl group at their 5'-termini are phosphorylated; c) contacting the sample treated with the

IDT Exhibit 1020.6337

polynucleotide kinase with the 5'-phosphate-dependent nucleic acid exonuclease under conditions such that undesired nucleic acids in the sample are substantially digested but desired mRNA molecules in the sample having a 5'-cap or a 5'-triphosphate are not digested; and d) producing cDNA copies of the mRNA. The present invention is not limited by the manner in which the cDNA copies are produced. In some preferred embodiments, the method of making cDNA copies comprises contacting the mRNA molecules with a primer having an oligo(dT) sequence in its 3'-end portion under conditions such that the primer anneals with the mRNA, and extending the primer with an RNA-dependent DNA polymerase under conditions whereby cDNA is obtained. In preferred embodiments, an RNA-dependent DNA polymerase enzyme and suitable reaction conditions are used so that the amount of full-length cDNA is maximized. Suitable enzymes and reaction conditions for this purpose are known in the art. For example, but without limitation, U.S. Patent Nos. 5,962,271; and 5,962,272 describe one method to obtain increased amounts of full-length cDNA, which can be used together with the methods of the present invention. However, the invention is not limited with respect to enzymes or reaction conditions, and any RNA-dependent DNA polymerase and reaction conditions that accomplish the intended purpose can be used in a method or kit of the present invention. Kits for conducting such methods of the present invention may have any one or more reagents (e.g., an oligo(dT)-containing primer and/or other primers), control samples, ATP, buffers, etc. useful in the method. For example, preferred kits comprise: a) a 5'-phosphate-dependent nucleic acid exonuclease; b) a polynucleotide kinase (e.g., T4 polynucleotide kinase); and c) an RNA-dependent DNA polymerase. In some preferred kits in which the mRNA in the sample is known to be of sufficient quality for the intended purpose, meaning that approximately <90%, <75%, <50%, <40%, <30%, <20%, <15%, <10%, <5%, <4%, <3%, <2%, or <1% of the mRNA transcripts are degraded into fragments that lack a 5'-cap or a 5'-triphosphate group, the step comprising contacting the sample with the polynucleotide kinase in the presence of ATP and under conditions such that nucleic acid molecules having an hydroxyl group at their 5'-termini are phosphorylated can be omitted and a polynucleotide kinase is not needed in a preferred kit, or if included, its use can be omitted.

IDT Exhibit 1020.6338

Similar methods and kits are provided that are useful for amplifying eukaryotic mRNA. For example, the present invention provides a method for amplifying eukaryotic mRNA, comprising the steps of a) providing a sample comprising RNA from a eukaryotic organism; and a 5'-phosphate-dependent nucleic acid exonuclease; b) contacting the sample with the 5'-phosphate-dependent nucleic acid exonuclease under conditions such that undesired nucleic acids in the sample are substantially digested but mRNA molecules in the sample having a 5'-triphosphate or a 5'-cap are not digested; and c) amplifying the mRNA. In preferred embodiments, the present invention provides a method for amplifying eukaryotic mRNA, comprising the steps of a) providing a sample comprising RNA from a eukaryotic organism; a polynucleotide kinase; and a 5'-phosphate-dependent nucleic acid exonuclease; b) contacting the sample with the polynucleotide kinase in the presence of ATP and under conditions such that nucleic acid molecules having an hydroxyl group at their 5'-termini are phosphorylated; c) contacting the sample with the 5'-phosphate-dependent nucleic acid exonuclease under conditions such that undesired nucleic acids in the sample are substantially digested but mRNA molecules in the sample having a 5'-triphosphate or a 5'-cap are not digested; and d) amplifying the mRNA. The present invention is not limited by the manner in which the mRNA is amplified. In some embodiments, the mRNA is amplified by a method comprising: i) contacting said mRNA molecules with an oligo(dT)-containing primer that anneals to a poly(A) tail of said mRNA and extending said primer with an RNA-dependent DNA polymerase under conditions whereby cDNA that is joined to an RNA polymerase promoter sequence is obtained (using any number of methods known in the art; methods are known for joining the promoter to either the 5'-end of the cDNA or the 3'-end of the cDNA, resulting in synthesis of amplified antisense or sense RNA, respectively, by subsequent *in vitro* transcription); ii) producing double stranded cDNA; and iii) contacting said double-stranded cDNA comprising said promoter sequence with an RNA polymerase that binds to said promoter and synthesizes RNA therefrom under conditions whereby RNA corresponding to mRNA in the biological sample is synthesized. Some other embodiments of the present invention use a single-stranded first-strand cDNA template that is functionally joined at either the 3'-end or the 5'-end to a single-stranded promoter that is recognized by an RNA polymerase, such as MiniV

IDT Exhibit 1020.6339

RNA polymerase, that can synthesize RNA *in vitro* therefrom; in these embodiments, the method does not comprise a step for synthesis of double-stranded cDNA. Still further, the invention also comprises methods for amplifying mRNA enriched or isolated using a 5'-phosphate-dependent exonuclease of the invention using methods, such as but not limited to, methods described in U.S. Pat. Appln. No. 2004/0180361 of Dahl et al.; PCT Patent Publication Nos. WO 02/16639; WO 00/56877; and AU 00/29742 of Takara Shuzo Company; and in U.S. Pat. No. 6,251,639 and U.S. Pat. Appln. Nos. 2001/0034048; 2003/0017591; 2003/0087251; and 2003/0186234 of Kurn (all incorporated herein by reference), even if said methods comprise synthesis of DNA products rather than RNA products corresponding to mRNA in the sample. Thus, another embodiment of the invention comprises a method for amplifying mRNA enriched or isolated using a 5'-phosphate-dependent exonuclease, the method comprising: i) contacting said mRNA molecules with a first primer that anneals to said mRNA and extending said first primer with an RNA-dependent DNA polymerase under conditions whereby first-strand cDNA is obtained; ii) contacting said first-strand cDNA with a second primer and extending said second primer with a strand-displacing DNA-dependent DNA polymerase under conditions whereby double-stranded cDNA that is joined to a ribonucleotide-containing tail sequence at the 3'-end of at least one cDNA strand is obtained; and iii) contacting said double-stranded cDNA with a ribonuclease H and a ribonucleotide-containing primer that anneals to the tail sequence and extending said ribonucleotide-containing primer with the strand-displacing DNA-dependent DNA polymerase under conditions whereby at least one strand of said cDNA is amplified. Kits for conducting such methods may have any one or more reagents (e.g., primers), control samples, buffers, etc. useful in the method. For example, preferred kits comprise: a) a 5'-phosphate-dependent nucleic acid exonuclease; b); c) an RNA-dependent DNA polymerase; and c) an RNA polymerase (e.g., T7, T3, or SP6 RNA polymerases). Other kits that use single-stranded templates and promoters use an RNA polymerase that synthesize RNA *in vitro* therefrom, such as MiniV RNA polymerase. Other preferred kits further comprise: a) a polynucleotide kinase, such as but not limited to T4 polynucleotide kinase, and ATP. Still other kits also include a ribonucleotide-containing primer, a ribonuclease H, and a strand-displacing DNA-dependent DNA polymerase, such as but not limited to phi29 DNA

IDT Exhibit 1020.6340

polymerase or rBST DNA polymerase large fragment (both available from EPICENTRE).

In some cases in which a sample cannot be replaced (e.g, a forensics sample from a crime scene), it may not be desirable to remove mRNA fragments that do not have a 5'-triphosphate or a 5'-cap prior to synthesis of cDNA and/or RNA amplification. Thus, some embodiments of the present invention comprise a method for preparing prokaryotic mRNA, eukaryotic mRNA or a mixture of both prokaryotic and eukaryotic mRNA that is degraded for synthesis of cDNA and/or RNA amplification, the method comprising: a) providing a sample comprising degraded RNA from a prokaryotic and/or eukaryotic organism; a poly(A) polymerase; and a 5'-phosphate-dependent nucleic acid exonuclease; b) contacting the sample with the 5' exonuclease under conditions such that undesired nucleic acids in the sample are substantially digested but desired RNA molecules are not digested; and c) contacting the sample with a poly(A) polymerase in the presence of ATP and under conditions in which a 3'-poly(A) tail is added to the desired RNA molecules. The invention further comprises a method for preparing cDNA from said 3'-polyadenylated RNA, the method comprising: a) providing a sample comprising 3'-polyadenylated RNA; an RNA-dependent DNA polymerase; and an oligo(dT)-containing primer that anneals to said tail; and b) contacting the 3'-polyadenylated RNA with the oligo(dT)-containing primer and extending the primer with the RNA-dependent DNA polymerase under conditions whereby cDNA is obtained. Other embodiments further comprise methods for amplifying the polyadenylated RNA, the method comprising a) contacting the polyadenylated RNA with a primer that anneals to said tail and extending said primer with an RNA-dependent DNA polymerase under conditions whereby cDNA that is joined to an RNA polymerase promoter sequence is obtained (using any number of methods known in the art; methods are known for joining the promoter to either the 5'-end of the cDNA or the 3'-end of the cDNA, resulting in synthesis of amplified antisense or sense RNA, respectively, by subsequent *in vitro* transcription); b) producing double stranded cDNA; and c) contacting said double-stranded cDNA comprising said promoter sequence with an RNA polymerase that binds to said promoter and synthesizes RNA therefrom under conditions whereby amplified RNA corresponding to polyadenylated RNA in the sample is synthesized. Kits for conducting such methods may have any one

IDT Exhibit 1020.6341

or more reagents (e.g., primers, such as an oligo(dT) primer, or other oligonucleotides), control samples, buffers, etc. useful in the method. For example, preferred kits comprise: a) a 5'-phosphate-dependent nucleic acid exonuclease; and b) a poly(A) polymerase. Other preferred kits for preparing cDNA additionally comprise an RNA-dependent DNA polymerase. Still other preferred kits for methods comprising sense or antisense RNA amplification comprise, in addition to an RNA-dependent DNA polymerase, an RNA polymerase (e.g., without limitation, T7, T3, or SP6 RNA polymerase).

The present invention also provides a method for using a 5'-phosphate-dependent exonuclease of the invention to obtain 5'-ends of degraded mRNA and other primary transcripts that have a 5'-cap or a 5'-triphosphate group.

For example, one method that can be used to isolate the 5'-ends of degraded mRNA and other primary transcripts that have a 5'-cap or a 5'-triphosphate group comprises the steps of:

a) obtaining a sample of total RNA comprising degraded RNA;

b) treating the degraded RNA with a polynucleotide kinase (e.g., T4 polynucleotide kinase) under conditions that result in phosphorylation of 5'-hydroxyl groups of said degraded RNA;

c) treating the RNA from step (b) with a 5'-phosphate-dependent exonuclease of the invention under conditions that result in digestion of RNA having a 5'-phosphate group; and

d) isolating the RNA that is not digested in step (c).

Degraded RNA, such as but not limited to RNA degraded by RNase A-type nucleases, can have 5'-hydroxylated termini. The polynucleotide kinase phosphorylates the 5'-ends of degrade RNA molecules with a hydroxyl group, making these molecules subtrates which can be digested by the 5'-to-3' exoribonuclease. The RNA obtained using this procedure will comprise 5'-end fragments of RNA that have a 5'-cap or a 5'-triphosphate. Without limiting the invention, if sufficient quantities of these fragments are available, the RNA fragments can be directly labeled [e.g., using ULS™ (Kreatech) or other chemical dye labeling compounds] and used for gene expression studies with arrays or microarrays of oligonucleotide sequences that are complementary to the 5'-ends of all known sequences or desired known sequences corresponding to mRNA transcripts

IDT Exhibit 1020.6342

for the organism from which the total RNA was obtained. If sufficient quantities of the RNA fragments are not available from this method (e.g., if the RNA fragments are obtained from only a single cell or a small number of cells), then the 5'-end fragments of RNA having a 5'-cap or a 5'-triphosphate obtained can be amplified by one of several methods. By way of example but not of limitation, the 5'-end fragments can be polyadenylated using poly(A) polymerase using methods known in the art, and then amplified using an Eberwine-type RNA amplification method (e.g., see Van Gelder et al., Proc. Nat. Acad. Sci. USA 87: 1663, 1990), which uses an oligo(dT) promoter primer and generates anti-sense RNA. A 1-round or 2-round TargetAmp™ aRNA or aminoallyl-aRNA Amplification Kit (EPICENTRE Technologies, Madison, WI, USA) or another kit can also be used for this purpose. If the amplified and labeled aRNA is analyzed using a microarray, the oligos on the microarray must be complementary to the aRNA corresponding to the 5'-ends of the mRNA fragments rather than to the sense mRNA sequence. Alternatively, a sense mRNA amplification method, such as available in kits from BD-Clontech and from Genisphere can be use to amplify sense RNA corresponding to the 5'-ends of the mRNA fragments. A method comprising an oligonucleotide sequence tag template can be used to add a sequence tag to the 3'-end of a nucleic acid, including an RNA or DNA molecule, which can then be amplified following addition of a promoter, as described in U.S. Patent Application No. 2005/015333, which is incorporated herein by reference. After labeling, the amplified sense RNA can be annealed to a microarray of antisense oligos complementary to all known sequences or desired known sequences corresponding to the 5'-ends of mRNA transcripts for the organism from which the total RNA was obtained. Still further, the 5'-ends of the mRNA fragments can be amplified without polyadenylation by using a promoter primer with a random sequence for priming first-strand cDNA synthesis (e.g., see Ziman and Davis, WO 2002/044399). Still other methods for amplifying the mRNA 5'-end fragments are described herein which can be used as part of the present invention.

In other embodiments of the invention, a 5'-phosphate-dependent exonuclease is used to obtain 5'-end fragments from intact, undegraded mRNA or other primary transcripts having a 5'-cap or a 5'-triphosphate in a sample which contains other undesired nucleic acids. One way that the total RNA can be analyzed to determine that it is not

IDT Exhibit 1020.6343

significantly degraded is to determine the profile of large rRNA in the sample using an Agilent bioanalyzer and to verify that the scan of the rRNA indicates that the large rRNA is largely intact. If desired, small RNA, such as 5S rRNA and tRNA, can be substantially removed from the total RNA by precipitation of other RNA in the sample using LiCl and ethanol, leaving 5S rRNA and tRNA in solution, prior to performing additional steps to isolate 5'-end fragments of mRNA. Thus, one embodiment is a method for isolating the 5'-ends of mRNA and other primary transcripts that have a 5'-cap or a 5'-triphosphate group (i.e., the nucleic acid of interest) from a sample which also contains undesired nucleic acid (e.g., large rRNA in a sample of total RNA), wherein the nucleic acid of interest is not substantially degraded, the method comprising:

a) obtaining a sample of total RNA that contains mRNA;

b) optionally, obtaining an RNA fraction by that is reduced in 5S rRNA and tRNA LiCl and ethanol precipitation;

c) treating the RNA so as to bring about controlled degradation of the RNA into smaller fragments by physical or chemical means or by degradation with a nuclease;

d) treating the degraded RNA from step (c) with a polynucleotide kinase (e.g., T4 polynucleotide kinase) under conditions that result in phosphorylation of 5'-hydroxyl groups of said degraded RNA;

e) treating the RNA from step (d) with a 5'-to-3' exoribonuclease under conditions that result in digestion of RNA having a 5'-phosphate group; and

f) isolating the RNA that is not digested in step (e).

This procedure will give a solution of fragments of 5'-capped or 5'-triphosphorylated mRNA of the desired size range. These mRNA 5'-end fragments can then be used for any desired application, including for the applications and methods described with respect to the 5'-end fragments of mRNA obtained from FFPE degraded RNA discussed herein. For example, but without limitation, the fragments can be polyadenylated using poly(A) polymerase and amplified using an RNA amplification method that yields labeled aRNA or sense RNA as described herein. In other embodiments of the invention, the 5'-end fragments of mRNA can be amplified and analyzed using rolling circle replication or rolling circle transcription as also described herein.

IDT Exhibit 1020.6344

Some methods that can be used for treating the RNA so as to bring about controlled degradation of the RNA into smaller fragments include chemical treatments, such exposing the RNA to cations, including, but not limited to Mg^{2+} under elevated temperatures, and controlled degradation with a ribonuclease (after optimizing the enzyme, enzyme amount, time and other reactions conditions that yield RNA fragments in the desired size range). One RNase that can be used to fragment the RNA is RNase I, which cleaves the RNA after every base and which can be inactivated by heat after the treatment. Still another embodiment of a method of the invention for controlled degradation of RNA with a nuclease comprises treating the RNA with an RNase H in the presence of one or more DNA oligonucleotides complementary to specific sequences of the RNA of interest, under conditions wherein the DNA oligonucleotides can hybridize to the complementary mRNA and the RNase H is active.

Still another embodiment of the invention is a method for method for amplifying circularized cDNA, which cDNA is prepared from 5'-end fragments of 5'-capped or 5'-triphosphorylated RNA, by rolling circle replication. Thus, one embodiment of the invention is a method for amplifying the 5'-end fragments of mRNA and other primary transcripts, the method comprising the steps of:

a) providing a sample comprising 5'-end fragments of mRNA or other primary transcripts of a desired size range (e.g., without limitation, $\leq$ 100 bases);

b) contacting the 5'-end fragments with poly(A) polymerase under conditions that result in addition of a poly(A) tail of sufficient length to permit annealing of a primer for reverse transcription;

c) contacting the poly(A)-tailed 5'-end fragments with a 5'-phosphorylated primer that anneals to the poly(A) tail and an RNA-dependent DNA polymerase or reverse transcriptase under conditions that result in extension of the primer and synthesis of first-strand cDNA that is complementary to the poly(A)-tailed 5'-end fragments;

d) treating the complex between the first-strand cDNA and the poly(A)-tailed 5'-end fragments of mRNA from step (c) so as to remove the RNA from the first-strand cDNA;

e) contacting the first-strand cDNA from step (d) with a ligase that can ligate single-stranded DNA (ssDNA) in the absence of a ligation splint oligo under conditions that

IDT Exhibit 1020.6345

result in intramolecular ligation (i.e., circularization) of the linear first-strand cDNA so as to obtain circular first-strand cDNA; and

f) contacting the circular first-strand cDNA obtained from step (e) with a strand-displacing DNA polymerase (e.g., without limitation, phi29 DNA polymerase or rBst DNA polymerase large fragment) and a primer that anneals to the circular first-strand cDNA under conditions that result in synthesis of concatemeric second-strand cDNA that is complementary to the circular first-strand cDNA (e.g., by rolling circle replication).

The primer used to prime first-strand cDNA synthesis in step (c) can be an oligo$(dT)_n$ or an oligo$(dT)_nX$ anchored primer, wherein X is a mixture of dAMP, dCMP and dGMP. In step (c) above, the primer has a 5'-phosphate. In other embodiments, the primer has a 5'-hydroxyl group and is phosphorylated using a polynucleotide kinase (e.g., T4 polynucleotide kinase) under kinasing conditions after step (d) and prior to step (e). An RNA-dependent DNA polymerase or reverse transcriptase that does not have RNase H activity, such as but not limited to, an enzyme is selected from among an RNase H-minus mutant of MMLV or AMV reverse transcriptase (e.g., SuperScript™ I, II, or III from Invitrogen) can be used and is preferred in step (c), but an RNA-dependent DNA polymerase that has RNase H activity can be used in some embodiments. Any ligase that ligates ssDNA in the absence of a ligation splint oligo can be used, such as but not limited to, a ligase selected from the group consisting of a ligase derived from phage TS2126 that infect Thermus scotoductus, CircLigase™ ssDNA ligase (EPICENTRE), and ThermoPhage™ ssDNA Ligase (PROKARIA, Reykjavik, Iceland). In some embodiments, the primer used for rolling circle replication in step (f) is an oligo(dA) primer that anneals to the oligo(dT) sequence from the primer that was used to prime synthesis of first-strand cDNA. In other embodiments, a multiplicity of primers is used for rolling circle replication. For example, in one embodiment, a solution comprising all possible combinations of short primers such as hexamers (i.e., the hexamer is synthesized to contain any of the four canonical nucleotide bases at each of the six nucleotide positions) is used. In some aspects of this embodiment, the concatemeric rolling circle replication product is labeled by incorporation of a labeled nucleotide, but any other suitable method known in the art can also be used to detect the rolling circle replication product. In another embodiment in which a multiplicity of primers is used for rolling

IDT Exhibit 1020.6346

circle replication, the primers comprise oligonucleotides that have a sequence that is the same as at least one sequence within a certain defined distance of the 5'-end of the known mRNA transcripts for the organism for which transcription is analyzed. The defined distance from the 5'-end of the mRNA transcripts within which the complementary sequence for the priming oligos is chosen is similar to the average length of the 5'-end fragments provided in step (a). In a preferred aspect of this embodiment, the multiplicity of primers having the same sequence as at least one sequence within a certain defined distance of the 5'-end of the known mRNA transcripts for the organism for which transcription is analyzed are attached to a surface (e.g., in an array or microarray) so that their 3'-ends have an hydroxyl group and the sequences are at sufficient distance from the surface so that the sequences can prime rolling circle replication by primer extension of the tethered primer using a circular first-strand cDNA from step (e), to which the primer is complementary, as a template. In this way, the primer that is attached to the surface is primer extended by rolling circle replication using a strand-displacing DNA polymerase under polymerization conditions. In some aspects of this embodiment, the concatemeric rolling circle replication product is labeled by incorporation of a labeled nucleotide, but any other suitable method known in the art can also be used to detect the rolling circle replication product.

Yet another method that can be used to amplify 5'-end fragments of mRNA and other primary transcripts comprises the steps of:

a) providing a sample comprising 5'-end fragments of mRNA and other primary transcripts of a desired size range (e.g., without limitation, ≤ 100 bases);

b) contacting the 5'-end fragments with a 5'-phosphorylated random hexamer primer and an RNA-dependent DNA polymerase or reverse transcriptase under conditions that result in extension of the primer and synthesis of first-strand cDNA that is complementary to the 5'-end fragments;

c) treating the complex between the first-strand cDNA and the 5'-end fragments of mRNA from step (b) so as to remove the RNA from the first-strand cDNA;

d) contacting the first-strand cDNA from step (c) with a ligase that can ligate single-stranded DNA (ssDNA) in the absence of a ligation splint oligo under conditions that

IDT Exhibit 1020.6347

result in intramolecular ligation (i.e., circularization) of the linear first-strand cDNA so as to obtain circular first-strand cDNA; and

e) contacting the circular first-strand cDNA obtained from step (d) with a strand-displacing DNA polymerase (e.g., without limitation, phi29 DNA polymerase or rBst DNA polymerase large fragment) and a primer that anneals to the circular first-strand cDNA under conditions that result in synthesis of concatemeric second-strand cDNA that is complementary to the circular first-strand cDNA (e.g., by rolling circle replication).

In preferred embodiments this method, the RNA-dependent DNA polymerase or reverse transcriptase is an enzyme that does not have RNase H activity. By way of example, but not of limitation, the enzyme is selected from among an RNase H-minus mutant of MMLV or AMV reverse transcriptase (e.g., SuperScript™ I, II, or III from Invitrogen). A ligase that ligates ssDNA in the absence of a ligation splint oligo includes, but is not limited to a ligase selected from the group consisting of a ligase derived from phage TS2126 that infect Thermus scotoductus, CircLigase™ ssDNA ligase (EPICENTRE), and ThermoPhage™ ssDNA Ligase (PROKARIA, Reykjavik, Iceland). In some embodiments, a multiplicity of primers is used for step (e). For example, in one embodiment, a solution comprising all possible hexamer primers (i.e., the hexamer is synthesized to contain any of the four canonical nucleotide bases at each of the six positions). In some aspects of this embodiment, the concatemeric rolling circle replication product is labeled by incorporation of a labeled nucleotide, but any other suitable method known in the art can also be used to detect the rolling circle replication product. In other embodiments in which a multiplicity of primers is used for rolling circle replication, the primers comprise oligonucleotides that have a sequence that is the same as at least one sequence within a certain defined distance of the 5'-end of the known mRNA transcripts for the organism for which transcription is analyzed. The defined distance from the 5'-ends of the mRNA transcripts is chosen to be similar to the average length of the 5'-end fragments provided in step (a). In a preferred aspect of this embodiment, the multiplicity of primers having the same sequence as at least one sequence within a certain defined distance of the 5'-end of the known mRNA transcripts for the organism for which transcription is analyzed are attached to a surface (e.g., in an array or microarray) so that their 3'-ends have an hydroxyl group and the sequences are at

IDT Exhibit 1020.6348

sufficient distance from the surface so that the sequences can prime rolling circle replication by primer extension of the tethered primer using a circular first-strand cDNA from step (d), to which the primer is complementary, as a template. In this way, the primer that is attached to the surface is primer extended by rolling circle replication using a strand-displacing DNA polymerase under polymerization conditions. In some aspects of this embodiment, the concatemeric rolling circle replication product is labeled by incorporation of a labeled nucleotide, but any other suitable method known in the art can also be used to detect the rolling circle replication product.

Another method for amplifying circularized cDNA prepared from 5'-end fragments of 5'-capped or 5'-triphosphorylated RNA is by rolling circle transcription. Thus, one method that can be used to amplify the 5'-end fragments of mRNA and other primary transcripts comprises the steps of:

a) providing a sample comprising 5'-end fragments of mRNA and other primary transcripts of a desired size range (e.g., without limitation, $\leq$ 100 bases);

b) contacting the 5'-end fragments with poly(A) polymerase under conditions that result in addition of a poly(A) tail of sufficient length to permit annealing of a reverse transcription primer;

c) contacting the poly(A)-tailed 5'-end fragments with a 5'-phosphorylated primer that anneals to the poly(A) tail and an RNA-dependent DNA polymerase or reverse transcriptase under conditions that result in extension of the primer and synthesis of first-strand cDNA that is complementary to the poly(A)-tailed 5'-end fragments;

d) treating the complex between the first-strand cDNA and the poly(A)-tailed 5'-end fragments of mRNA from step (c) so as to remove the RNA from the first-strand cDNA;

e) contacting the first-strand cDNA from step (d) with a ligase that can ligate single-stranded DNA (ssDNA) in the absence of a ligation splint oligo under conditions that result in intramolecular ligation (i.e., circularization) of the linear first-strand cDNA so as to obtain circular first-strand cDNA; and

f) contacting the circular first-strand cDNA obtained from step (e) with an RNA polymerase under conditions that result in synthesis of concatemeric RNA complementary to the first-strand cDNA by rolling circle transcription.

IDT Exhibit 1020.6349

An oligo$(dT)_n$ or an oligo$(dT)_nX$ anchored primer, wherein X=a mixture of dAMP, dCMP and dGMP, can be used to prime first-strand cDNA synthesis in step (c). In step (c) above, the primer has a 5'-phosphate. In other embodiments, the primer has a 5'-hydroxyl group and is phosphorylated using a polynucleotide kinase (e.g., T4 polynucleotide kinase) under kinasing conditions after step (d) and prior to step (e). RNA-dependent DNA polymerases or reverse transcriptases that can be used include, but are not limited to enzymes that lack RNase H activity. By way of example, but not of limitation, the enzyme can be selected from among an RNase H-minus mutant of MMLV or AMV reverse transcriptase (e.g., SuperScript™ I, II, or III from Invitrogen). However, a reverse transcriptase with RNase H activity can also be used. Ligases that can be used to ligate ssDNA include, but are not limited to, ligases selected from the group consisting of a ligase derived from phage TS2126 that infect Thermus scotoductus, CircLigase™ ssDNA ligase (EPICENTRE), and ThermoPhage™ ssDNA Ligase (PROKARIA, Reykjavik, Iceland). The RNA polymerase can be any RNA polymerase that transcribes a single-stranded circular DNA in the absence of a primer and in the absence of a promoter sequence, although a promoter sequence can be present in some embodiments of the invention. RNA polymerases that can be used include, but are not limited to, *E. coli* RNA polymerase (including core enzyme); a T7-type RNA polymerase, including T7 RNAP, T3 RNAP, and SP6 RNAP; or phage N4 mini-vRNAP.

Still another method that can be used to amplify the 5'-end fragments of mRNA and other primary transcripts comprises the steps of:

a) providing a sample comprising 5'-end fragments of mRNA and other primary transcripts of a desired size range (e.g., without limitation, ≤ 100 bases);

b) contacting the 5'-end fragments with a 5'-phosphorylated random hexamer primer and an RNA-dependent DNA polymerase or reverse transcriptase under conditions that result in extension of the primer and synthesis of first-strand cDNA that is complementary to the 5'-end fragments;

c) treating the complex between the first-strand cDNA and the 5'-end fragments of mRNA from step (b) so as to remove the RNA from the first-strand cDNA;

d) contacting the first-strand cDNA from step (c) with a ligase that can ligate single-stranded DNA (ssDNA) in the absence of a ligation splint oligo under conditions that

IDT Exhibit 1020.6350

result in intramolecular ligation (i.e., circularization) of the linear first-strand cDNA so as to obtain circular first-strand cDNA; and

e) contacting the circular first-strand cDNA obtained from step (d) with an RNA polymerase under conditions that result in synthesis of concatemeric RNA complementary to the first-strand cDNA by rolling circle transcription.

RNA-dependent DNA polymerases or reverse transcriptases that can be used include, but are not limited to enzymes that lack RNase H activity. By way of example, but not of limitation, the enzyme can be selected from among an RNase H-minus mutant of MMLV or AMV reverse transcriptase (e.g., SuperScript™ I, II, or III from Invitrogen). However, a reverse transcriptase with RNase H activity can also be used. Ligases that can be used to ligate ssDNA include, but are not limited to, ligases selected from the group consisting of a ligase derived from phage TS2126 that infect Thermus scotoductus, CircLigase™ ssDNA ligase (EPICENTRE), and ThermoPhage™ ssDNA Ligase (PROKARIA, Reykjavik, Iceland). The RNA polymerase can be any RNA polymerase that transcribes a single-stranded circular DNA in the absence of a primer and in the absence of a promoter sequence, although a promoter sequence can be present in some embodiments of the invention. RNA polymerases that can be used include, but are not limited to, *E. coli* RNA polymerase (including core enzyme); a T7-type RNA polymerase, including T7 RNAP, T3 RNAP, and SP6 RNAP; or phage N4 mini-vRNAP.

The present invention also comprises embodiments comprising amplification of circular first-strand cDNA by both rolling circle replication and rolling circle transcription using any combination of the methods described herein.

The present invention further comprises methods methods, compositions, and kits for amplification of linear first-strand cDNA prepared from 5'-capped or 5'-triphosphorylated mRNA using a random hexamer primer and a strand-displacing DNA polymerase. Thus, one method that can be used to amplify the 5'-end fragments of mRNA and other primary transcripts comprises the steps of:

a) providing a sample comprising 5'-end fragments of mRNA and other primary transcripts of a desired size range (e.g., without limitation, approximately 100 - 500 bases);

IDT Exhibit 1020.6351

b) contacting the 5'-end fragments with poly(A) polymerase under conditions that result in addition of a poly(A) tail of sufficient length to permit annealing of a reverse transcription primer;

c) contacting the poly(A)-tailed 5'-end fragments with a primer that anneals to the poly(A) tail and an RNA-dependent DNA polymerase or reverse transcriptase under conditions that result in extension of the primer and synthesis of first-strand cDNA that is complementary to the poly(A)-tailed 5'-end fragments;

d) treating the complex between the first-strand cDNA and the poly(A)-tailed 5'-end fragments of mRNA from step (c) so as to remove the RNA from the first-strand cDNA; and

e) contacting the first-strand cDNA obtained from step (d) with a strand-displacing DNA polymerase (e.g., phi29 DNA polymerase) and a random hexamer primer that anneals to the first-strand cDNA under conditions that result in amplification of the cDNA.

An oligo$(dT)_n$ or an oligo$(dT)_nX$ anchored primer, wherein X=a mixture of dAMP, dCMP and dGMP, can be used to prime first-strand cDNA synthesis in step 4.A.(c). In preferred embodiments of step 4.A. (c), the RNA-dependent DNA polymerase or reverse transcriptase is an enzyme that does not have RNase H activity. By way of example, but not of limitation, the enzyme is selected from among an RNase H-minus mutant of MMLV or AMV reverse transcriptase (e.g., SuperScript™ I, II, or III from Invitrogen). However, reverse transcriptases with RNase H activity can be used in some embodiments. In some embodiments of step 4.A (e), phi29 DNA polymerase is used as the strand-displacing DNA polymerase and the random hexamer primer used comprises either a random hexamer DNA primer having alpha-thiophosphate internucleoside linkages, which are resistant to many exonucleases, or a random hexamer RNA primer.

Still another embodiment of the invention comprises a method for amplifying the 5'-end fragments of mRNA and other primary transcripts, the method comprising the steps of:

a) providing a sample comprising 5'-end fragments of mRNA and other primary transcripts of a desired size range (e.g., without limitation, approximately 100 - 500 bases);

IDT Exhibit 1020.6352

b) contacting the 5'-end fragments with a random primer that anneals to 5'-end fragments of mRNA and other primary transcripts and an RNA-dependent DNA polymerase or reverse transcriptase under conditions that result in extension of the primer and synthesis of first-strand cDNA that is complementary to the 5'-end fragments;

c) treating the complex between the first-strand cDNA and the 5'-end fragments of mRNA from step (b) so as to remove the RNA from the first-strand cDNA; and

d) contacting the first-strand cDNA obtained from step (c) with a strand-displacing DNA polymerase (e.g., phi29 DNA polymerase) and a random hexamer primer that anneals to the first-strand cDNA under conditions that result in amplification of the cDNA.

The random primer used in step 4.B. (b) can comprise a random hexamer DNA primer having alpha-thiophosphate internucleoside linkages. In preferred embodiments of step 4.A. (b), the RNA-dependent DNA polymerase or reverse transcriptase is an enzyme that does not have RNase H activity. By way of example, but not of limitation, the enzyme is selected from among an RNase H-minus mutant of MMLV or AMV reverse transcriptase (e.g., SuperScript™ I, II, or III from Invitrogen). However, reverse transcriptases with RNase H activity can be used in some embodiments. The primer in step 4.B.(d) can comprise either a random hexamer DNA primer having alpha-thiophosphate internucleoside linkages or a random hexamer RNA primer. One DNA polymerase that can be used in this step is phi29 DNA polymerase.

The present invention further provides kits and methods for selectively modifying nucleic acids that are not substrates for a 5' exonuclease of the present invention so that said nucleic acids will be substrates for said 5' exonuclease and thereby, can be removed from desired nucleic acid molecules. For example, but without limitation, an RNA molecule that is not a substrate for a 5'-phosphate-dependent exoribonuclease because it has a triphosphate group or an hydroxyl group on its 5'-terminus can be hybridized to a DNA oligo that anneals near to its 5'-end and the RNA:oligo complex can be contacted with a ribonuclease H, such as but not limited to *E. coli* RNase H or a *Thermus* RNase H (e.g., Hybridase RNase H from EPICENTRE). The RNase H digests the RNA within the portion that is annealed to the oligo, yielding an RNA molecule that has a 5'-phosphorylated RNA that is a substrate for the 5'-phosphate-dependent exoribonuclease that can be substantially digested in the presence of desired nucleic acids. Thus, in some

IDT Exhibit 1020.6353

preferred embodiments, the present invention provides a method for enriching for a desired RNA in the presence of an undesired RNA comprising the steps of: a) providing a 5'-phosphate-dependent nucleic acid exoribonuclease; a sample comprising RNA, said sample containing desired RNA and an undesired RNA that is not a substrate for the exoribonuclease; a DNA oligo that is complementary to a sequence near the 5'-terminus of the undesired RNA; and an RNase H; b) contacting the sample with the DNA oligo and the RNase H under conditions under which the oligo anneals to the undesired RNA and the RNase H is active; c) treating the sample with the exoribonuclease under conditions such that the undesired RNA is substantially digested and the desired RNA is not digested; and d) recovering the desired RNA (e.g., for additional analysis or use). In some embodiments, the 5'-phosphate-dependent nucleic acid exoribonuclease used in the method is provided in a kit that contains the DNA oligo, RNase H, buffers, instructions, and optionally, appropriate control samples, and the like.

The present invention is not limited with respect to kits and methods for modifying nucleic acids to be either desired or undesired nucleic acids that are, respectively, either resistant to digestion or susceptible to digestion by a 5' exonuclease of the present invention. Any appropriate modification method can be used to achieve a particular purpose in a method of the invention. For example, but without limitation, a 5'-exonuclease-resistant nucleic acid with a 5'-hydroxyl can be modified to a 5'-phosphorylated nucleic acid that is a substrate for digestion by treatment with a polynucleotide kinase, such as T4 polynucleotide kinase, and ATP under kinase reaction conditions; or a 5'-phosphorylated nucleic acid that is a substrate for digestion by a 5'-exonuclease can be dephosphorylated to a 5'-exonuclease-resistant nucleic acid with a 5'-hydroxyl by treatment with a DNA phosphatase, such as, but not limited to APex Alkaline Phosphatase (EPICENTRE) or shrimp alkaline phosphatase; or a 5'-exonuclease-resistant 5'-triphosphorylated or 5'-capped mRNA can be modified to a 5'-exonuclease substrate that has a 5'-phosphate by treatment with a pyrophosphatase, such as but not limited to, tobacco acid pyrophosphatase. Those with skill in the art will know or know how to identify other methods for modifying undesired nucleic acids so that they will be substrates for a 5' exonuclease of the invention or for modifying desired nucleic

IDT Exhibit 1020.6354

acids so they will not be substrates for the 5' exonuclease, all of which modification methods are included within the scope of the present invention.

The present invention further provides methods and kits for removing undesired 5'-phosphorylated oligonucleotides from nucleic acid reactions (e.g., without limitation, aRNA and sense RNA amplification reactions), comprising the steps of: a) providing a sample that contains desired nucleic acid molecules that have undergone or are undergoing enzymatic manipulation and undesired 5'-phosphorylated oligonucleotides that are involved in said manipulation of said desired nucleic acid molecules; and b) exposing the sample to a 5'-phosphate-dependent nucleic acid exonuclease under conditions such that the undesired 5'-phosphorylated oligonucleotides are substantially digested and the desired nucleic acid molecules are not digested. The method finds use in a wide variety of applications. By way of example, but not of limitation, U.S. Patent Appln. Nos. 2004/0171041 and 2004/0197802, which are incorporated herein by reference in their entireties, disclose methods for which this aspect of the present invention can be used. For example, the present method can be used for methods disclosed therein that use a splice template oligo (or template switch oligo), such as a promoter splice template oligo to synthesize first-strand cDNA comprising a target sequence in order to reduce subsequent background transcription due to the presence of the splice template oligo if it is not removed. The splice template oligo can have a 5'-phosphate group or it can be phosphorylated during the method using a polynucleotide kinase. The present invention is not limited to removal of splice template oligos. The invention also includes removal of other oligos that are undesired nucleic acids. (e.g., removal of linear promoter primers from circular first-strand cDNA synthesized by primer extension of the promoter primer and subsequent ligation using a ligase in the presence of a ligation splint oligo or using a ligase that ligates ssDNA in the absence of a ligation splint oligo (e.g., CircLigase ssDNA Ligase (EPICENTRE Biotechnologies)), and the like, and from RNA amplification reactions). In some embodiments, the undesired 5'-phosphorylated oligonucleotide includes, but is not limited to, primers, ligation splint oligos, etc. In some embodiments, the enzymatic manipulation comprises a ligation reaction, a nucleic acid synthesis reaction (e.g., *in vitro* transcription reaction, cDNA synthesis reaction, Eberwine-type aRNA amplification reaction, sense RNA

IDT Exhibit 1020.6355

amplification reaction, etc.). Kits for conducting such methods comprise a 5'-phosphate-dependent exonuclease and may have any one or more other reagents (e.g., primers), control samples (e.g., 5'-phosphorylated oligonucleotides), buffers, etc. useful in the method.

The compositions and methods of the present invention provided herein can be used in diagnostic methods. For example, the present invention provides a method for detecting the presence of a target nucleic acid in a sample, comprising the steps of: a) providing a sample comprising a mixture of nucleic acids, said mixture of nucleic acids suspected of containing the target nucleic acid; b) exposing the sample to a 5'-phosphate-dependent nucleic acid exonuclease; and c) detecting the presence of undigested nucleic acid molecules that are not substrates for the exonuclease or detecting the products and/or process of digestion of nucleic acid molecules that are substrates for the exonuclease in order to detect the presence of the target nucleic acid in the sample. The method is not limited by the nature of the target nucleic acid. In some embodiments, the target nucleic acid, if present in the sample, is modified as discussed herein above to make it susceptible or resistant to cleavage by the 5'-phosphate-dependent nucleic acid exonuclease in a target-specific manner. The target nucleic acid may also be differentiated from non-target sequence by its different behavior in response to particular modification treatments (e.g., susceptibility to RNase H digestion in the presence of a DNA oligo or restriction enzyme digestion due to the presence of a particular sequence). The present invention also provides kits, mixtures, and compositions for carrying out such methods. For example, the present invention provides a kit for isolating a nucleic acid molecule of interest in a mixture of nucleic acids comprising one or more of: a) a 5'-phosphate-dependent nucleic acid exonuclease; b) a positive control sample comprising nucleic acid having a 5' phosphate (e.g., rRNA); and c) a negative control sample comprising nucleic acid that is resistant to degradation by the 5'-phosphate-dependent nucleic acid exonuclease (e.g., mRNA).

The present invention further provides a method for isolating 5'-capped or 5'-triphosphorylated or 5'-hydroxylated small RNA molecules in samples containing eukaryotic RNA (e.g., small nuclear RNA (snRNA) or pre-microRNA (pre-miRNA) to study the their functions in gene processing and control of gene expression, etc. in health

IDT Exhibit 1020.6356

and disease) or 5'-capped or 5'-triphosphorylated transcripts of other small RNA molecules, such as but not limited to, primary, the method comprising the steps of: a) providing a sample comprising nuclear RNA from a eukaryotic organism (e.g., obtained from isolated nuclei or from total RNA preparations using methods known in the art), said sample containing desired small RNA having a 5'-cap or 5'-triphosphate or 5'-hydroxyl group and at least one undesired nucleic acid; and a 5'-to-3' exoribonuclease; b) treating the sample with the 5'- to-3' exoribonuclease under conditions such that said undesired nucleic acid is substantially digested and the desired 5'-capped or 5'-triphosphorylated or 5'-hydroxylated small RNA is not digested; and c) recovering the capped or triphosphorylated or hydroxylated small RNA. In some embodiments, the 5'-phosphate-dependent nucleic acid exonuclease used in the method is provided in a kit for isolation of capped or triphosphorylated or hydroxylated small RNA, with buffers, instructions, and optionally, with appropriate control samples, and the like. Pre-miRNA, which has a 5'-triphosphate, is further processed in the cell to micro RNA (miRNA) having a 5'-monophosphate, which is implicated in control of gene expression of many. In some embodiments of methods of the invention, a 5'-group is modified in order to achieve a purpose. For example, without limitation, in one embodiment of the method, small RNA with a 5'-phosphate group can be removed using a nucleic acid phosphatase, such as APex Phosphatase (EPICENTRE), to form a 5'-hydoxyl group, which is resistant to digestion with the 5'-to-3 exoribonuclease, thereby permitting enrichment for the small RNA (such as miRNA). In another embodiment, the small RNA with a 5'-triphosphate or 5'-cap can be treated with tobacco acid pyrophosphatase (EPICENTRE) to form a 5'-phosphate group, thereby permitting digestion of the small RNA, such as pre-miRNA, with the 5'-to-3 exoribonuclease.

The present invention further comprises methods to prepare cDNA from said isolated capped or triphosphorylated small RNA, the method comprising a) providing a sample comprising capped and/or triphosphorylated small RNA; a poly(A) polymerase; and an RNA-dependent DNA polymerase; b) contacting the sample with the poly(A) polymerase in the presence of ATP and under conditions in which a 3'-poly(A) tail is added to the desired small RNA molecules; and c) contacting the 3'-polyadenylated RNA with the oligo(dT)-containing primer and extending the primer with the RNA-dependent

IDT Exhibit 1020.6357

DNA polymerase under conditions whereby cDNA is obtained. Other embodiments comprise methods for amplifying the polyadenylated small RNA, the methods comprising a) contacting the polyadenylated small RNA with a primer that anneals to said tail and extending said primer with an RNA-dependent DNA polymerase under conditions whereby cDNA that is joined to an RNA polymerase promoter sequence is obtained (using any number of methods known in the art; methods are known for joining the promoter to either the 5'-end of the cDNA or the 3'-end of the cDNA, resulting in synthesis of amplified antisense or sense RNA, respectively, by subsequent *in vitro* transcription); b) producing double stranded cDNA; and c) contacting said double-stranded cDNA comprising said promoter sequence with an RNA polymerase that binds to said promoter and synthesizes RNA therefrom under conditions whereby amplified RNA corresponding to polyadenylated RNA in the sample is synthesized. Kits for conducting such methods may have any one or more reagents (e.g., primers), control samples, buffers, etc. useful in the method. For example, preferred kits comprise: a) a 5'-phosphate-dependent nucleic acid exonuclease; b) a poly(A) polymerase; c) an RNA-dependent DNA polymerase; and, for methods comprising RNA amplification, d) an RNA polymerase (e.g., T7, T3, and SP6 RNA polymerases).

One preferred embodiment of the invention is a kit for enrichment of mRNA from a sample comprising mRNA and at least one undesired nucleic acid, the kit comprising a 5'-to-3' exoribonuclease that has an activity in digesting an RNA having a 5'-triphosphate or a 5'-cap that is less than 5% of its activity in digesting said RNA having a 5'-monophosphate.

Another preferred embodiment is a kit for enrichment of mRNA from a sample comprising mRNA and at least one undesired nucleic acid, the kit comprising: a) a 5'-to-3' exoribonuclease, wherein the activity in digesting an RNA having a 5'-triphosphate or a 5'-cap is less than 5% of its activity in digesting said RNA having a 5'-monophosphate; and b) a negative control comprising an RNA having a 5'-triphosphate or a 5'-cap.

Another preferred embodiment is a kit for enrichment of mRNA from a sample comprising mRNA and at least one undesired nucleic acid, the kit comprising: a) a 5'-to-3' exoribonuclease, wherein the activity in digesting an RNA having a 5'-triphosphate or

IDT Exhibit 1020.6358

a 5'-cap is less than 5% of its activity in digesting said RNA having a 5'-monophosphate; and b) a positive control comprising an RNA having a 5'-monophosphate.

Still another preferred embodiment is a kit for enrichment of mRNA from a sample comprising mRNA and at least one undesired nucleic acid, the kit comprising: a) a 5'-to-3' exoribonuclease, wherein the activity in digesting an RNA having a 5'-triphosphate or a 5'-cap is less than 5% of its activity in digesting said RNA having a 5'-monophosphate; b) a negative control comprising an RNA having a 5'-triphosphate or a 5'-cap; and c) a positive control comprising an RNA having a 5'-monophosphate.

Another preferred embodiment of the invention is a kit for isolation of mRNA from a sample comprising mRNA and at least one undesired nucleic acid, the kit comprising: a) a 5'-to-3' exoribonuclease, wherein the activity in digesting an RNA having a 5'-triphosphate or a 5'-cap is less than 5% of its activity in digesting said RNA having a 5'-monophosphate; and b) a solution of LiCl for precipitation of low molecular RNA.

Another preferred embodiment is a kit for isolation of mRNA from a sample comprising mRNA and at least one undesired nucleic acid, the kit comprising: a) a 5'-to-3' exoribonuclease, wherein the activity in digesting an RNA having a 5'-triphosphate or a 5'-cap is less than 5% of its activity in digesting said RNA having a 5'-monophosphate; b) a solution of LiCl for precipitation of low molecular RNA; and c) a negative control comprising an RNA having a 5'-triphosphate or a 5'-cap.

Another preferred embodiment is a kit for isolation of mRNA from a sample comprising mRNA and at least one undesired nucleic acid, the kit comprising: a) a 5'-to-3' exoribonuclease, wherein the activity in digesting an RNA having a 5'-triphosphate or a 5'-cap is less than 5% of its activity in digesting said RNA having a 5'-monophosphate; b) a solution of LiCl for precipitation of low molecular RNA; and c) a positive control comprising an RNA having a 5'-monophosphate.

Still another preferred embodiment is a kit for isolation of mRNA from a sample comprising mRNA and at least one undesired nucleic acid, the kit comprising: a) a 5'-to-3' exoribonuclease, wherein the activity in digesting an RNA having a 5'-triphosphate or a 5'-cap is less than 5% of its activity in digesting said RNA having a 5'-monophosphate; b) a solution of LiCl for precipitation of low molecular RNA; c) a negative control

IDT Exhibit 1020.6359

comprising an RNA having a 5'-triphosphate or a 5'-cap; and d) a positive control comprising an RNA having a 5'-monophosphate.

The present invention further comprises a kit for enrichment of mRNA having a 5'-triphosphate or a 5'-cap from a sample comprising mRNA and at least one undesired nucleic acid, the kit comprising: a) a 5'-to-3' exoribonuclease, wherein the activity in digesting an RNA having a 5'-triphosphate or a 5'-cap is less than 5% of its activity in digesting said RNA having a 5'-monophosphate; and b) a polynucleotide kinase.

Another embodiment of the invention is a kit for isolation of mRNA having a 5'-triphosphate or a 5'-cap from a sample comprising mRNA and at least one undesired nucleic acid, the kit comprising: a) a 5'-to-3' exoribonuclease, wherein the an activity in digesting an RNA having a 5'-triphosphate or a 5'-cap is less than 5% of its activity in digesting said RNA having a 5'-monophosphate; b) a polynucleotide kinase; and c) a solution of LiCl for precipitation of low molecular RNA.

The present invention further comprises a kit for synthesis of first-strand cDNA complementary to mRNA from a sample comprising mRNA and at least one undesired nucleic acid, the kit comprising: a) a 5'-to-3' exoribonuclease, wherein the activity in digesting an RNA having a 5'-triphosphate or a 5'-cap is less than 5% of its activity in digesting said RNA having a 5'-monophosphate; b) at least one primer that anneals to said mRNA in said sample; and c) an RNA-dependent DNA polymerase. One kit according to this embodiment is a kit for synthesis of first-strand cDNA complementary to mRNA from a prokaryotic organism. Another kit according to this embodiment is a kit for synthesis of first-strand cDNA complementary to mRNA from a eukaryotic organism. Still another kit according to this embodiment is a kit for synthesis of first-strand cDNA complementary to mRNA from both a prokaryotic organism and a eukaryotic organism.

Still further, another preferred embodiment of the invention is a kit for synthesis of first-strand cDNA complementary to mRNA from a prokaryotic organism, the kit comprising: a) a 5'-to-3' exoribonuclease, wherein the activity in digesting an RNA having a 5'-triphosphate or a 5'-cap is less than 5% of its activity in digesting said RNA having a 5'-monophosphate; b) a poly(A) polymerase; c) an oligo(dT)-containing primer that is capable of annealing to a poly(A) tail; and d) an RNA-dependent DNA polymerase. Another kit according to this embodiment is a kit for synthesis of first-strand

IDT Exhibit 1020.6360

cDNA complementary to mRNA from both a prokaryotic organism and a eukaryotic organism in a sample comprising mRNA from both the prokaryotic and the eukaryotic organisms.

Still another preferred embodiment of the invention is a kit for synthesis of first-strand cDNA complementary to mRNA from a prokaryotic organism, the kit comprising: a) a 5'-to-3' exoribonuclease, wherein the activity in digesting an RNA having a 5'-triphosphate or a 5'-cap is less than 5% of its activity in digesting said RNA having a 5'-monophosphate; b) a poly(A) polymerase; c) a polynucleotide kinase; d) an oligo(dT)-containing primer that is capable of annealing to a poly(A) tail; and e) an RNA-dependent DNA polymerase. Another kit according to this embodiment is a kit for synthesis of first-strand cDNA complementary both to mRNA from a prokaryotic organism and to full-length mRNA from a eukaryotic organism in a sample comprising mRNA from both the prokaryotic and the eukaryotic organisms.

Yet another preferred embodiment of the present invention is a kit for synthesis of first-strand cDNA complementary to full-length mRNA from a eukaryotic organism, the kit comprising: a) a 5'-to-3' exoribonuclease, wherein the activity in digesting an RNA having a 5'-triphosphate or a 5'-cap is less than 5% of its activity in digesting said RNA having a 5'-monophosphate; b) a polynucleotide kinase; c) an oligo(dT)-containing primer that is capable of annealing to a poly(A) tail; and d) an RNA-dependent DNA polymerase.

Still further, the invention also comprises a kit for RNA amplification of mRNA from a sample comprising mRNA and at least one undesired nucleic acid, the kit comprising: a) a 5'-to-3' exoribonuclease, wherein the activity in digesting an RNA having a 5'-triphosphate or a 5'-cap is less than 5% of its activity in digesting said RNA having a 5'-monophosphate; b) at least one primer that anneals to said mRNA in said sample; c) an oligonucleotide that encodes an RNA polymerase promoter; d) an RNA-dependent DNA polymerase; and e) an RNA polymerase that binds to the RNA polymerase promoter encoded by the oligonucleotide. One kit according to this embodiment is a kit for RNA amplification of mRNA from a prokaryotic organism. Another kit according to this embodiment is a kit for RNA amplification of mRNA from

IDT Exhibit 1020.6361

a eukaryotic organism. Still another kit according to this embodiment is a kit for RNA amplification of mRNA from both a prokaryotic organism and a eukaryotic organism.

One preferred embodiment of the invention is a kit for RNA amplification of mRNA from a prokaryotic organism, the kit comprising: a) a 5'-to-3' exoribonuclease, wherein the activity in digesting an RNA having a 5'-triphosphate or a 5'-cap is less than 5% of its activity in digesting said RNA having a 5'-monophosphate; b) a poly(A) polymerase; c) an oligo(dT)-containing primer that is capable of annealing to a poly(A) tail; d) an oligonucleotide that encodes an RNA polymerase promoter; e) an RNA-dependent DNA polymerase; and f) an RNA polymerase that binds to the RNA polymerase promoter encoded by the oligonucleotide. One kit according to this embodiment of the invention is a kit for RNA amplification of mRNA from both a prokaryotic organism and a eukaryotic organism in a sample comprising mRNA from both the prokaryotic and the eukaryotic organisms.

Still another preferred embodiment of the invention is a kit for RNA amplification of mRNA from a prokaryotic organism, the kit comprising: a) a 5'-to-3' exoribonuclease, wherein the activity in digesting an RNA having a 5'-triphosphate or a 5'-cap is less than 5% of its activity in digesting said RNA having a 5'-monophosphate; b) a poly(A) polymerase; c) a polynucleotide kinase; d) an oligo(dT)-containing primer that is capable of annealing to a poly(A) tail; e) an oligonucleotide that encodes an RNA polymerase promoter; e) an RNA-dependent DNA polymerase; and f) an RNA polymerase that binds to the RNA polymerase promoter encoded by the oligonucleotide. One kit according to this embodiment is a kit for RNA amplification of both mRNA from a prokaryotic organism and full-length mRNA from a eukaryotic organism in a sample comprising mRNA from both the prokaryotic and the eukaryotic organisms.

The present invention further comprises a kit for RNA amplification of mRNA from a eukaryotic organism, the kit comprising: a) a 5'-to-3' exoribonuclease, wherein the activity in digesting an RNA having a 5'-triphosphate or a 5'-cap is less than 5% of its activity in digesting said RNA having a 5'-monophosphate; b) an oligo(dT)-containing primer that is capable of annealing to a poly(A) tail; c) an oligonucleotide that encodes an RNA polymerase promoter; d) an RNA-dependent DNA polymerase; and e) an RNA polymerase that binds to the RNA polymerase promoter encoded by the oligonucleotide.

IDT Exhibit 1020.6362

A preferred embodiment of the invention comprises a kit for RNA amplification of full-length mRNA from a eukaryotic organism, the kit comprising: a) a 5'-to-3' exoribonuclease, wherein the activity in digesting an RNA having a 5'-triphosphate or a 5'-cap is less than 5% of its activity in digesting said RNA having a 5'-monophosphate; b) a polynucleotide kinase; c) an oligo(dT)-containing primer that is capable of annealing to a poly(A) tail; d) an oligonucleotide that encodes an RNA polymerase promoter; e) an RNA-dependent DNA polymerase; and f) an RNA polymerase that binds to the RNA polymerase promoter encoded by the oligonucleotide.

The present invention also comprises a kit for removing ribosomal RNA from a sample, the kit comprising: a) a 5'-to-3' exoribonuclease, wherein the activity in digesting an RNA having a 5'-triphosphate or a 5'-cap is less than 5% of its activity in digesting said RNA having a 5'-monophosphate; and b) a positive control comprising rRNA selected from among a) 16S and 23S prokaryotic rRNA; and b) 18S and 26S or 28S eukaryotic rRNA. A preferred kit of this embodiment comprises a kit for removing ribosomal RNA >300 nucleotides from a sample. Other kits according to these embodiments comprise kits that additionally comprise a negative control comprising an RNA having a 5'-triphosphate, a 5'-cap, or a 5'-hydroxyl.

One preferred embodiment of the invention is a kit for enriching for mRNA having a 5'-triphosphate or a 5'-cap in a biological sample comprising prokaryotic mRNA, eukaryotic mRNA or both prokaryotic and eukaryotic mRNA and at least one undesired nucleic acid, the kit comprising: (i) a solution of purified and stabilized 5'-to-3' exoribonuclease; and (ii) a concentrated solution of a reaction buffer for providing a 1X reaction buffer in which the exoribonuclease is active. One suitable 1X concentration of a reaction buffer that can be used for the invention comprises: 50 mM Tris-HCl (pH 8.0), 2.0 mM $MgCl_2$, and 100 mM NaCl.

In another embodiment, a kit for enriching for mRNA additionally comprises one or more components selected from the group consisting of: (i) a negative control comprising an RNA having a 5'-triphosphate or a 5'-cap; (ii) a positive control comprising an RNA having a 5'-monophosphate; (iii) a solution of LiCl; (iv) a polynucleotide kinase; and (v) a poly(A) polymerase.

IDT Exhibit 1020.6363

In still another embodiment, the kit additionally comprises an RNA-dependent DNA polymerase (reverse transcriptase).

In still another embodiment, the kit additionally comprises one or more components selected from the group consisting of: (i) a poly(A) polymerase; (ii) an oligo(dT)-containing primer; and (iii) an RNA-dependent DNA polymerase.

In yet another embodiment, the kit additionally comprises one or more components selected from the group consisting of: (i) a polynucleotide kinase; (ii) an oligo(dT)-containing primer; and (iii) an RNA-dependent DNA polymerase for obtaining cDNA complementary to full-length mRNA from a eukaryotic organism.

In yet another embodiment, the kit additionally comprises one or more primers selected from the group consisting of (i) an oligo(dT)-containing primer; and (ii) a primer that is complementary to a specific nucleic acid sequence. In some aspects of these embodiments, the primer additionally comprises a sense or an antisense sequence of a double-stranded promoter for a T7-type RNA polymerase selected from the group consisting of T7 RNA polymerase, T3 RNA polymerase, and SP6 RNA polymerase, and wherein the kit additionally comprises the RNA polymerase that can transcribe RNA using said promoter.

In some embodiments the kit additionally comprises: (i) a oligonucleotide sequence tag template; and (ii) a DNA polymerase selected from the group consisting of a DNA-dependent DNA polymerase and an RNA-dependent DNA polymerase that can extend the 3'-end of a nucleic acid that is annealed to the oligonucleotide sequence tag template. In some of these embodiments, the kit additionally comprises: (i) an oligonucleotide that is complementary to the oligonucleotide sequence tag template; and (ii) a T7-type RNA polymerase selected from the group consisting of T7 RNA polymerase, T3 RNA polymerase, and SP6 RNA polymerase; wherein the oligonucleotide sequence tag template and the oligonucleotide that is complementary to the oligonucleotide sequence tag template together encode the complete sequence of a promoter that can be used for transcription by said T7-type RNA polymerase.

In many cases, a kit of the present invention will contain one or more components selected from the group consisting of: (i) a negative control comprising an RNA having

IDT Exhibit 1020.6364

a 5'-triphosphate or a 5'-cap; (ii) a negative control comprising an RNA or a DNA having a 5'-hydroxyl group; (iii) a positive control comprising an RNA having a 5'-monophosphate; (iv) a solution of LiCl; (v) a ribonuclease H enzyme; and (vi) a polynucleotide kinase. The ribonuclease H (RNase H)can be used, together with a DNA oligonucleotide that anneals to a specific sequence of an mRNA molecule, to cleave the mRNA in a particular region prior to adding sequence tag by extension of the 3'-end of the RNase H-cleaved mRNA having a 5'-triphosphate or a 5'cap using an oligonucleotide sequence tag template as a template. The polynucleotide kinase can be used to phosphorylate the 5'-end of degraded RNA molecules having a 5'-hydroxyl group in order to generate 5'-monophosphorylated RNA substrates that can digested using a 5'-to-3' exoribonuclease of the invention.

Some kits of the invention will also comprise one or more components selected from the group consisting of: (i) a poly(A) polymerase; (ii) an oligo(dT)-containing primer; and (iii) an RNA-dependent DNA polymerase, such as, but not limited to, kits for enriching for prokaryotic mRNA. Some of such kits may also contain an RNA-dependent DNA polymerase and one or more primers selected from the group consisting of (i) a random primer, such as a random hexamer primer; (ii) an oligo(dT)-containing primer; and (iii) a primer that is complementary to a specific nucleic acid sequence. Some of such kits may also comprise: (i) a oligonucleotide sequence tag template for adding a tag sequence to the 3'-end of mRNA having a 5'-triphosphate or a 5'-cap; and (ii) a DNA polymerase selected from the group consisting of a DNA-dependent DNA polymerase and an RNA-dependent DNA polymerase that can extend the 3'-end of the mRNA using the oligonucleotide sequence tag template annealed to the mRNA as a template.

The kit of the invention can also comprise: (i) an oligonucleotide sequence tag template for adding a tag sequence to the 3'-end of mRNA having a 5'-triphosphate or a 5'-cap; and (ii) a DNA polymerase selected from the group consisting of a DNA-dependent DNA polymerase and an RNA-dependent DNA polymerase that can extend the 3'-end of the mRNA using the oligonucleotide sequence tag template annealed to the mRNA as a template. The DNA polymerase can be an RNA-dependent DNA polymerase, or, if a DNA-dependent DNA polymerase is used for adding the tag sequence to the 3'-end of the mRNA, the kit additionally contains an RNA-dependent

IDT Exhibit 1020.6365

DNA polymerase that can synthesize cDNA from a primer using the mRNA as a template.

In some embodiments of the invention, the oligonucleotide sequence tag template can additionally encode either a sense or an antisense strand of a double-stranded promoter for a T7-type RNA polymerase selected from the group consisting of T7 RNA polymerase, T3 RNA polymerase, and SP6 RNA polymerase, and in those embodiments, the kit can additionally comprise the RNA polymerase that can transcribe RNA using said promoter.

Still further, a kit of the invention can also comprise a poly(A) polymerase for tailing of prokaryotic mRNA, and an oligo(dT)-containing primer and an RNA-dependent DNA polymerase for synthesis of cDNA. The kit can also contain an oligonucleotide sequence tag template (for adding a sequence tag to the 3'-end of first-strand cDNA); an oligonucleotide that is complementary to the oligonucleotide sequence tag template; wherein the oligonucleotide sequence tag template and the oligonucleotide complementary to the oligonucleotide sequence tag template together can encode the complete sequence of a promoter for a T7-type RNA polymerase. Thus, in some such embodiments, the kit can also contain a T7-type RNA polymerase selected from the group consisting of a T7 RNA polymerase, a T3 RNA polymerase, and an SP6 RNA polymerase, wherein the T7-type RNA polymerase can use the promoter encoded by the combination of the oligonucleotide sequence tag template and the oligonucleotide complementary to the oligonucleotide sequence tag template.

Still other kits of the invention are useful for removing intact prokaryotic or eukaryotic ribosomal RNA (rRNA) of a size having a svedburg unit greater than about 10S from a biological sample. These kits can comprise: (i) a solution of purified and stabilized 5'-to-3' exoribonuclease that is free of contaminating *S. cerevisiae* protein; (ii) a concentrated solution of a reaction buffer for providing a 1X reaction buffer in which the exoribonuclease is active; and (iii) a positive control rRNA, selected from the group consisting of a 16S and a 23S prokaryotic rRNA, and an 18S, a 26S and a 28S eukaryotic rRNA. In a preferred embodiment of this aspect of the invention, the 5'-to-3' exoribonuclease is a wild-type or recombinant *Saccharomyces cerevisiae* Xrn1p/5' exoribonuclease 1.

IDT Exhibit 1020.6366

In addition to the kits described above, it should be understood that the present invention provides a wide variety of kits for use in an equally wide variety of applications. In some preferred embodiments, the kits comprise a 5'-phosphate-dependent nucleic acid exonuclease (e.g., in a form, concentration, etc. suitable for use in the methods described herein). In other preferred embodiments, the kits comprise or consist of a 5'-phosphate-dependent nucleic acid exonuclease and one or more other components with which the 5'-phosphate-dependent nucleic acid exonuclease might be used directly or indirectly, either alone together or with yet other components. The one or more components may be any component involved in any of the methods described herein or in any of the methods described in the references or patents cited herein. The one or more components include, but are not limited to, buffers, salts, enzymes, proteins, antibodies, nucleic acid molecules (e.g., primers, probes, antisense oligonucleotides, RNA, DNA, etc.), nucleotides, control reagents, instructions, solid surfaces (e.g., beads, microarray components, etc.), detection reagents (e.g., fluorescent, luminescent, colorimetric, etc.), general laboratory equipment (tubes, multiwell plates, etc.), software, vectors or vector components, cell culture reagents and materials, and the like.

DESCRIPTION OF THE FIGURE

Figure 1 shows the amino acid (SEQ ID NO:1) and nucleic acid (SEQ ID NO:2) sequences of Saccharomyces cerevisiae Xrn1p.

DEFINITIONS:

The present invention will be understood and interpreted based on the definitions of terms as defined below.

A "5'-to-3' exoribonuclease" or a "5' exoribonuclease" or a "5' Xrn1p exoribonuclease" or a "5' exonuclease" or a "5'-phosphate-dependent exonuclease" or a "5'-phosphate-dependent nucleic acid exonuclease" or a "5 nuclease" of the present invention comprises an 5'-exonuclease that has greater than 20-fold more 5'-to-3' exonuclease activity for a single-stranded RNA substrate that has a 5'-monophosphorylated terminus than for the same RNA substrate that has a 5'-triphosphorylated or 5'-capped terminus, or, any 5'-to-3' exonuclease that has a relative

IDT Exhibit 1020.6367

activity in digesting a particular defined-sequence single-stranded RNA substrate with a 5'-triphosphate or a 5'-cap that is less than 5% of the activity for an RNA substrate having the same sequence but with a 5'-monophosphate. Enzyme activity of a 5' exonuclease of the invention can be measured using a number of different methods. Without limitation, suitable methods that can be used for assaying activity and determining relative activity using RNA substrates with a 5'-triphosphate, a 5'-cap, or a 5'-monophosphate are described by Stevens and Poole (J. Biol. Chem., 270: 16063, 1995).

Preferred embodiments of compositions, kits and methods of the invention employ Xrn1p/5' exoribonuclease 1, including *Saccharomyces cerevisiae* Xrn1p/5' exoribonuclease 1, and/or functional variants and homologues thereof. Xrn1p was first identified and characterized in *Saccharomyces cerevisiae*, but the activity has been reported in numerous organisms, including humans. In preferred embodiments of the present invention, the 5'-to-3' exoribonuclease comprises a solution of the enzyme that has been purified so as to be free of contaminating enzymes with activity on nucleic acids and *Saccharomyces cerevisiae* proteins. In some preferred embodiments, the 5'-to-3' exoribonuclease is obtained by expression of the *Saccharomyces cerevisiae* XRN1 gene that has been cloned in a plasmid, and then replicated and expressed in *Esherichia coli* cells, since the Xrn1p/5' exoribonuclease I obtained from such a recombinant source is of a higher purity, free from contaminating enzymatic activities, and generally at a higher enzyme concentration than is obtained from non-recombinant sources. In preferred embodiments of the invention, the 5'-to-3' exoribonuclease is "stabilized", by which we mean that the 5'-to-3' exoribonuclease is sufficiently pure of proteases and other contaminants which contribute to degradation and loss of enzyme activity and that the 5'-to-3' exoribonuclease is provided in a formulation of enzyme storage buffer in which there is no significant loss of activity during storage at -20 degrees C for six months. One suitable enzyme storage buffer for providing a stabilized 5'-to-3' exoribonuclease comprises a 50% glycerol solution containing 50 mM Tris-HCL (pH 7.5), 100 mM NaCl, 100 mM EDTA, 1 mM DTT and 0.1% of the non-ionic detergent Triton X-100. The *Saccharomyces cerevisiae* purified yeast enzyme was initially found to be a 160-kDa protein but, based on its gene sequence, the actual size is closer to 175 kDa. The amino acid (SEQ ID NO:1) and nucleic acid (SEQ ID NO:2) sequences of Saccharomyces

IDT Exhibit 1020.6368

cerevisiae Xrn1p are shown in Figure 1. "Xrn1p" generally refers to the protein, whereas "XRN1" generally refers to the gene; however, the terms "Xrn1p" and "Xrn1p/5' exoribonuclease I" are used interchangeably herein and the term "Xrn1p", as used herein, refers to both the protein and gene unless indicated otherwise. The enzyme is a processive exoribonuclease, requiring a divalent cation, and generally is stimulated by monovalent cations. The enzyme acts on a variety of substrates, including homopolymers, rRNA, and oligoadenylates. The enzyme also has some DNase activity. Single-stranded molecules are the preferred substrate. The XRN1 gene of Saccharomyces cerevisiae is identical to a gene (DST2, SEP1) encoding a DNA transferase and to genes involved in nuclease fusion, KEM1, and plasmid stability, RAR5 (see e.g., Larimer et al., Gene 120:51, 1992). Another name for this gene is Stpβ (see e.g., Heyer et al., Mol. Cell Biol. 15:2728, 1995).

Without being bound by theory, sequence comparison of several Mg2+-dependent 5'-3' exonucleases from bacteriophages, prokaryotes, and eukaryotes (Solinger et al., Mol. Cell. Biol. 19:5930, 1995) revealed some highly conserved amino acid residues. In the crystal structure of phage T4 RNase H, these residues are clustered in the proposed active site. In particular, some conserved aspartic (D) and glutamic (E) acid residues coordinate two Mg2+ ions in the reactive center of the exonuclease and are believed to play a crucial role in catalysis. Mutation of D206 and D208 of S. cerevisiae Xrn1p abolished exonuclease activity in the mutants.

Moreover, variant forms of Xrn1p are also contemplated as being equivalent to those peptides and DNA molecules that are set forth in more detail herein. For example, it is contemplated that isolated replacement of a leucine with an isoleucine or valine, an aspartate with a glutamate, a threonine with a serine, or a similar replacement of an amino acid with a structurally related amino acid (*i.e.*, conservative mutations) will not have a major effect on the biological activity of the resulting molecule. Accordingly, some embodiments of the present invention provide variants of Xrn1p disclosed herein containing conservative replacements. Conservative replacements are those that take place within a family of amino acids that are related in their side chains. Genetically encoded amino acids can be divided into four families: (1) acidic (aspartate, glutamate); (2) basic (lysine, arginine, histidine); (3) nonpolar (alanine, valine, leucine, isoleucine,

IDT Exhibit 1020.6369

proline, phenylalanine, methionine, tryptophan); and (4) uncharged polar (glycine, asparagine, glutamine, cysteine, serine, threonine, tyrosine). Phenylalanine, tryptophan, and tyrosine are sometimes classified jointly as aromatic amino acids. In similar fashion, the amino acid repertoire can be grouped as (1) acidic (aspartate, glutamate); (2) basic (lysine, arginine, histidine), (3) aliphatic (glycine, alanine, valine, leucine, isoleucine, serine, threonine), with serine and threonine optionally be grouped separately as aliphatic-hydroxyl; (4) aromatic (phenylalanine, tyrosine, tryptophan); (5) amide (asparagine, glutamine); and (6) sulfur -containing (cysteine and methionine) (*e.g.*, Stryer ed., *Biochemistry*, pg. 17-21, 2nd ed, WH Freeman and Co., 1981). It can be readily determined whether a change in the amino acid sequence of a peptide results in a functional polypeptide by assessing the ability of the variant peptide to function in a fashion similar to the wild-type protein. Peptides having more than one replacement can readily be tested in the same manner.

More rarely, a variant includes "nonconservative" changes (*e.g.*, replacement of a glycine with a tryptophan). Analogous minor variations can also include amino acid deletions or insertions, or both. Guidance in determining which amino acid residues can be substituted, inserted, or deleted without abolishing biological activity can be found using computer programs (*e.g.*, LASERGENE software, DNASTAR Inc., Madison, Wis.).

Variants may be produced by methods such as directed evolution or other techniques for producing combinatorial libraries of variants, described in more detail below. In still other embodiments of the present invention, the nucleotide sequences of the present invention may be engineered in order to alter an Xrn1p coding sequence including, but not limited to, alterations that modify the cloning, processing, localization, secretion, and/or expression of the gene product. For example, mutations may be introduced using techniques that are well known in the art (*e.g.*, site-directed mutagenesis to insert new restriction sites, alter glycosylation patterns, or change codon preference, etc.).

Still other embodiments of the present invention provide mutant or variant forms of Xrn1p. It is possible to modify the structure of a peptide having an activity of Xrn1p for such purposes as enhancing activity, or stability (*e.g.*, *ex vivo* shelf life, and/or

IDT Exhibit 1020.6370

resistance to proteolytic degradation *in vivo*). Such modified peptides are considered functional equivalents of peptides having an activity of the subject Xrn1p proteins as defined herein. A modified peptide can be produced in which the amino acid sequence has been altered, such as by amino acid substitution, deletion, or addition.

Moreover, as described above, variant forms (*e.g.*, mutants) of the subject Xrn1p proteins are also contemplated as being equivalent to those peptides and DNA molecules that are set forth in more detail. For example, as described above, the present invention encompasses mutant and variant proteins that contain conservative or non-conservative amino acid substitutions.

This invention further contemplates a method of generating sets of combinatorial mutants of the present Xrn1p proteins, as well as truncation mutants, and is especially useful for identifying potential variant sequences (*i.e.*, mutants) that are functional in exoribonuclease activity. The purpose of screening such combinatorial libraries is to generate, for example, novel Xrn1p variants that have improved or altered exoribonuclease activity.

Therefore, in some embodiments of the present invention, Xrn1p variants are engineered by the present method to provide altered (*e.g.*, increased or decreased) exoribonuclease activity. In other embodiments, Xrn1p variants are engineered to provide heat-stable (i.e., "thermostable") or heat-labile exoribonuclease activity for particular applications. In other embodiments of the present invention, combinatorially-derived variants are generated which have substrate variability different than that of a naturally occurring Xrn1p. Such proteins, when expressed from recombinant DNA constructs, find use in the methods described herein.

Still other embodiments of the present invention provide Xrn1p variants that have intracellular half-lives dramatically different than the corresponding wild-type protein. For example, the altered protein can be rendered either more stable or less stable to proteolytic degradation or other cellular process that result in destruction of, or otherwise inactivate Xrn1p. Such variants, and the genes which encode them, can be utilized to alter the location of Xrn1p expression by modulating the half-life of the protein. For instance, a short half-life can give rise to more transient Xrn1p biological effects and,

IDT Exhibit 1020.6371

when part of an inducible expression system, can allow tighter control of Xrn1p levels within the cell.

In still other embodiments of the present invention, Xrn1p variants are generated by the combinatorial approach to act as antagonists, in that they are able to interfere with the ability of the corresponding wild-type protein to regulate cell function.

In some embodiments of the combinatorial mutagenesis approach of the present invention, the amino acid sequences for a population of Xrn1p homologs, variants or other related proteins are aligned, preferably to promote the highest homology possible. Such a population of variants can include, for example, Xrn1p homologs from one or more species, or Xrn1p variants from the same species but which differ due to mutation or polymorphisms. Amino acids that appear at each position of the aligned sequences are selected to create a degenerate set of combinatorial sequences.

In a preferred embodiment of the present invention, the combinatorial Xrn1p library is produced by way of a degenerate library of genes encoding a library of polypeptides which each include at least a portion of potential Xrn1p protein sequences. For example, a mixture of synthetic oligonucleotides can be enzymatically ligated into gene sequences such that the degenerate set of potential Xrn1p sequences are expressible as individual polypeptides, or alternatively, as a set of larger fusion proteins (*e.g.*, for phage display) containing the set of Xrn1p sequences therein.

There are many ways by which the library of potential Xrn1p homologs and variants can be generated from a degenerate oligonucleotide sequence. In some embodiments, chemical synthesis of a degenerate gene sequence is carried out in an automatic DNA synthesizer, and the synthetic genes are ligated into an appropriate gene for expression. The purpose of a degenerate set of genes is to provide, in one mixture, all of the sequences encoding the desired set of potential Xrn1p sequences. The synthesis of degenerate oligonucleotides is well known in the art (*See e.g.*, Narang, Tetrahedron Lett., 39:39, 1983; Itakura *et al.*, Recombinant DNA, *in* Walton (ed.), *Proceedings of the 3rd Cleveland Symposium on Macromolecules,* Elsevier, Amsterdam, pp 273-289, 1981; Itakura *et al.*, Annu. Rev. Biochem., 53:323, 1984; Itakura *et al.*, Science 198:1056, 1984; Ike *et al.*, Nucl. Acid Res., 11:477, 1983). Such techniques have been employed in the directed evolution of other proteins (*See e.g.*, Scott *et al.*, Science 249:386 [1980];

IDT Exhibit 1020.6372

Roberts *et al.*, Proc. Natl. Acad. Sci. USA 89:2429 [1992]; Devlin *et al.*, Science 249: 404 [1990]; Cwirla *et al.*, Proc. Natl. Acad. Sci. USA 87: 6378 [1990]; as well as U.S. Pat. Nos. 5,223,409, 5,198,346, and 5,096,815; each of which is incorporated herein by reference).

It is contemplated that the Xrn1p nucleic acids (*e.g.*, SEQ ID NO:2, and fragments and variants thereof) can be utilized as starting nucleic acids for directed evolution. These techniques can be utilized to develop Xrn1p variants having desirable properties such as increased, decreased, or altered exoribonuclease activity.

In some embodiments, artificial evolution is performed by random mutagenesis (*e.g.*, by utilizing error-prone PCR to introduce random mutations into a given coding sequence). This method requires that the frequency of mutation be finely tuned. As a general rule, beneficial mutations are rare, while deleterious mutations are common. This is because the combination of a deleterious mutation and a beneficial mutation often results in an inactive enzyme. The ideal number of base substitutions for targeted gene is usually between 1.5 and 5 (Moore and Arnold, Nat. Biotech., 14, 458, 1996; Eckert and Kunkel, PCR Methods Appl., 1:17-24, 1991; Caldwell and Joyce, PCR Methods Appl., 2:28, 1992; and Zhao and Arnold, Nuc. Acids Res. 25:1307, 1997). After mutagenesis, the resulting clones are selected for desirable activity (*e.g.*, screened for Xrn1p activity). Successive rounds of mutagenesis and selection are often necessary to develop enzymes with desirable properties. It should be noted that only the useful mutations are carried over to the next round of mutagenesis.

In other embodiments of the present invention, the polynucleotides of the present invention are used in gene shuffling or sexual PCR procedures (*e.g.*, Smith, Nature, 370:324, 1994; U.S. Pat. Nos. 5,837,458; 5,830,721; 5,811,238; 5,733,731; all of which are herein incorporated by reference). Gene shuffling involves random fragmentation of several mutant DNAs followed by their reassembly by PCR into full length molecules. Examples of various gene shuffling procedures include, but are not limited to, assembly following DNase treatment, the staggered extension process (STEP), and random priming in vitro recombination. In the DNase mediated method, DNA segments isolated from a pool of positive mutants are cleaved into random fragments with DNaseI and subjected to multiple rounds of PCR with no added primer. The lengths of random fragments

IDT Exhibit 1020.6373

approach that of the uncleaved segment as the PCR cycles proceed, resulting in mutations present in different clones becoming mixed and accumulating in some of the resulting sequences. Multiple cycles of selection and shuffling have led to the functional enhancement of several enzymes (Stemmer, Nature, 370:398, 1994; Stemmer, Proc. Natl. Acad. Sci. USA, 91:10747, 1994; Crameri *et al.*, Nat. Biotech., 14:315, 1996; Zhang *et al.*, Proc. Natl. Acad. Sci. USA, 94:4504, 1997; and Crameri *et al.*, Nat. Biotech., 15:436, 1997).

A wide range of techniques are known in the art for screening gene products of combinatorial libraries made by point mutations, and for screening cDNA libraries for gene products having a certain property. Such techniques will be generally adaptable for rapid screening of the gene libraries generated by the combinatorial mutagenesis or recombination of Xrn1p homologs or variants. The most widely used techniques for screening large gene libraries typically comprise cloning the gene library into replicable expression vectors, transforming appropriate cells with the resulting library of vectors, and expressing the combinatorial genes under conditions in which detection of a desired activity facilitates relatively easy isolation of the vector encoding the gene whose product was detected.

Homologues and similar proteins that find use in the compositions and methods of the present invention include proteins and extracts containing proteins from a variety of organisms. For example, Rat1p/5' exonuclease 2 and homologues DHp1p and Dhm1p in mouse and humans may be used in the compositions and methods of the present invention. Work with crude extracts shows a homologue in Xenopus (see e.g., Furuichi et al., Nature 266:235, 1977) and wheat germ (see e.g., Shimotohno et al., Proc. Natl. Acad. Sci. USA 74:2734, 1977). Stevens et al. showed purification of a 5' exoribonuclease from human placental nuclei (Nucl. Acids Res., 15:695, 1987). Heyer et al., show that antibodies against the yeast Xrn1 protein also interact with the homologue protein in Drosophila, Xenopus, and mouse (Mol. Cell. Biol., 15:2728, 1995).

Fragments of the nucleic acids and proteins of the present invention may also be used, so long as the fragments encode or possess the desired enzymatic activity. Page et al. (Nucl. Acids Res., 26:3707, 1998), herein incorporated by reference in its entirety,

IDT Exhibit 1020.6374

provide deletion analysis demonstrating that the C-terminus is dispensable for exonuclease function.

A "unit" of exoribonuclease is the amount of enzyme required to convert 1 μg of *E. coli* ribosomal RNA (i.e., 16S and/or 23S rRNA) to acid-soluble form in sixty minutes at 30 °C under standard reaction assay conditions (e.g., 50 mM Tris-HCl, pH 8.0, 2 mM $MgCl_2$, and 0.1 M NaCl).

The symbol "S" with respect to a ribosomal RNA, such as "16S" or "23S" rRNA, or another macromolecule, designates a "svedberg unit" (named for Theodor Svedberg), which is a unit used to measure the sedimentation rate of a molecule in a colloidal suspension and which indicates the molecular weight of the molecule, with a svedberg unit being equal to 10^{-13} second.

The term "gene" as used herein, refers to a DNA sequence that comprises control and coding sequences necessary for the production of a polypeptide or protein precursor or other encoded molecule (e.g., rRNA, tRNA). The polypeptide can be encoded by a full-length coding sequence or by any portion of the coding sequence, as long as the desired protein activity is retained.

"Nucleoside", as used herein, refers to a compound consisting of a purine (guanine (G) or adenine (A)) or pyrimidine (thymine (T), uridine (U), or cytidine (C)) base covalently linked to a pentose sugar, whereas "nucleotide" refers to a nucleoside phosphorylated at one of the hydroxyl groups of the pentose sugar.

A "cap" is a modified guanine nucleotide that is joined to the 5'-end of the 5' nucleotide of a primary eukaryotic mRNA molecule. A common cap (sometimes referred to as "standard cap") has a structure designated as m^7G[5']ppp[5']N, in which "p" represents a phosphate group, "G" represents a guanosine nucleoside, "m^7" represents a methyl group on the 7-position of guanine (in this example, on the guanine base of the cap nucleotide), and "[5']" indicates the position at which the "p" is joined to the ribose of each respective nucleoside (in this example, indicating that the 5'-carbon of the ribose of the cap is joined via a triphosphate group to the 5'-carbon of the ribose of the first nucleoside ("N") of the mRNA. In addition to this "standard" cap, a variety of other naturally-occurring and synthetic cap analogs are known in the art.

IDT Exhibit 1020.6375

A "nucleic acid" or a "polynucleotide", as used herein, is a covalently linked sequence of nucleotides in which the 3' position of the sugar moiety of one nucleotide is joined by a phosphodiester group to the 5' position of the sugar moiety of the next nucleotide, and in which the nucleotide residues (bases) are linked in specific sequence; i.e., a linear order of nucleotides. An "oligonucleotide", as used herein, is a short polynucleotide or a portion of a polynucleotide. An oligonucleotide typically contains a sequence of about two to about one hundred bases, although longer molecules may also be encompassed by this term. The word "oligo" is sometimes used in place of the word "oligonucleotide".

An "oligonucleotide sequence tag template" or a "splice template oligo" or a "template switch oligo" is an oligonucleotide that complexes with a single-stranded target nucleic acid and is used as a template to extend the 3'-terminus of one or more target sequences in order to add a specific sequence or a "sequence tag". The 3'-portion of an oligonucleotide sequence tag template or a splice template is sufficiently complementary to the 3'-terminus of the target sequence that is to be extended to anneal thereto. A DNA- or RNA-dependent DNA polymerase is then used to extend the target nucleic acid molecule using the sequence in the 5'-portion of the oligonucleotide sequence tag template or splice template oligo as a template. The extension product of the primer-extended molecule has the specific sequence at its 3'-terminus that is complementary to the sequence in the 5'-portion of the oligonucleotide sequence tag template or splice template oligo. With respect to the present invention, an oligonucleotide sequence tag template can be used to add a sequence tag to the 3'-end of a nucleic acid, including an RNA or DNA molecule, as described in U.S. Patent Application No. 2005/015333 of Roy R. Sooknanan, which is incorporated herein by reference.

One embodiment of an oligonucleotide sequence tag template or splice template oligo of the present invention is a "promoter sequence tag template" or "promoter splice template oligo." A promoter sequence tag template or promoter splice template oligo comprises a sequence in its 3'-portion that is sufficiently complementary to the 3'-end of the target sequence to anneal thereto and a sequence in it 5'-portion that is complementary to a sequence comprising a single-stranded transcription promoter of the invention. Thus, a promoter sequence tag template or promoter splice template oligo can provide a

IDT Exhibit 1020.6376

template for synthesis of a sequence comprising a transcription promoter at the 3'-end of first-strand cDNA obtained either by RNA-dependent DNA polymerase (or "reverse transcriptase") primer extension of a target sequence comprising an RNA target nucleic acid, such as, but not limited to, an mRNA target, or by DNA polymerase primer extension of a target sequence comprising DNA or RNA.

A "ligation splint oligo" or "ligation splint" is an oligo that is used to provide an annealing site or a "ligation template" for joining two ends of one nucleic acid (i.e., "intramolecular joining") or two ends of two nucleic acids (i.e., "intermolecular joining") using a ligase or another enzyme with ligase activity. The ligation splint holds the ends adjacent to each other and "creates a ligation junction" between the 5'-phosphorylated and a 3'-hydroxylated ends that are to be ligated.

Nucleic acid molecules are said to have a "5'-terminus" (5' end) and a "3'-terminus" (3' end) because nucleic acid phosphodiester linkages occur at the 5' carbon and 3' carbon of the sugar moieties of the substituent mononucleotides. The end of a polynucleotide at which a new linkage would be to a 5' carbon is its 5' terminal nucleotide. The end of a polynucleotide at which a new linkage would be to a 3' carbon is its 3' terminal nucleotide. A terminal nucleotide, as used herein, is the nucleotide at the end position of the 3'- or 5'-terminus.

Nucleic acid molecules are said to have "5' ends" and "3' ends" because mononucleotides are joined in one direction via a phosphodiester linkage to make oligonucleotides, in a manner such that a phosphate on the 5'-carbon of one mononucleotide sugar moiety is joined to an oxygen on the 3'-carbon of the sugar moiety of its neighboring mononucleotide. Therefore, an end of an oligonucleotide referred to as the "5' end" if its 5' phosphate is not linked to the oxygen of the 3'-carbon of a mononucleotide sugar moiety and as the "3' end" if its 3' oxygen is not linked to a 5' phosphate of the sugar moiety of a subsequent mononucleotide.

Polypeptide molecules are said to have an "amino terminus" (N-terminus) and a "carboxy terminus" (C-terminus) because peptide linkages occur between the backbone amino group of a first amino acid residue and the backbone carboxyl group of a second amino acid residue.

IDT Exhibit 1020.6377

As used herein, the terms "complementary" or "complementarity" are used in reference to a sequence of nucleotides related by the base-pairing rules. For example, the sequence 5'-A-G-T-3', is complementary to the sequence 3'-T-C-A-5'. Complementarity may be "partial," in which only some of the nucleic acids' bases are matched according to the base pairing rules. Or, there may be "complete" or "total" complementarity between the nucleic acids. The degree of complementarity between nucleic acid strands has significant effects on the efficiency and strength of hybridization between nucleic acid strands. This is of particular importance in amplification reactions, as well as detection methods that depend upon hybridization of nucleic acids.

As used herein, the term "probe" refers to an oligonucleotide (i.e., a sequence of nucleotides), whether occurring naturally (e.g., as in a purified restriction digest) or produced synthetically, recombinantly or by PCR amplification, which is capable of hybridizing to another oligonucleotide of interest. A probe may be single-stranded or double-stranded (e.g., and rendered single-stranded or partially single-stranded in use). Probes are useful in the detection, identification and isolation of particular gene sequences. It is contemplated that a probe used in the present invention can be labeled with any "reporter molecule," so that it is detectable in a detection system, including, but not limited to enzyme (i.e., ELISA, as well as enzyme-based histochemical assays), visible, fluorescent, radioactive, and luminescent systems. It is not intended that the present invention be limited to any particular detection system or label. The terms "reporter molecule" and "label" are used herein interchangeably. In addition to probes, primers and deoxyribonucleoside triphosphates may contain labels; these labels may comprise, but are not limited to, ^{32}P, ^{33}P, ^{35}S, enzymes, or visible, luminescent, or fluorescent molecules (e.g., fluorescent dyes).

The term "recombinant protein" or "recombinant polypeptide" as used herein refers to a protein molecule expressed from a recombinant DNA molecule. In contrast, the term "native protein" is used herein to indicate a protein isolated from a naturally occurring (i.e., a nonrecombinant) source. Molecular biological techniques may be used to produce a recombinant form of a protein with identical or similar properties as compared to the native form of the protein. Variants of the native sequence may also be

IDT Exhibit 1020.6378

made to, for example, improve expression, purification, or other desired properties of the polypeptide.

As used herein in reference to an amino acid sequence or a protein, the term "portion" (as in "a portion of an amino acid sequence") refers to fragments of that protein. The fragments may range in size from four amino acid residues to the entire amino acid sequence minus one amino acid. Particularly preferred fragments retain one or more of the enzymatic activities associated with a whole protein (e.g., 5' exoribonuclease activity).

As used herein, the term "fusion protein" refers to a chimeric protein containing the protein of interest (e.g., 5' exoribonuclease 1 and fragments thereof) joined to an exogenous protein fragment (e.g., the fusion partner which contains a non-xrn-1 protein). The fusion partner may enhance the solubility of xrn-1 protein as expressed in a host cell, may provide an affinity tag to allow purification of the recombinant fusion protein from the host cell or culture supernatant, or both. If desired, the fusion protein may be removed from the protein of interest (e.g., xrn-1 or fragments thereof) by a variety of enzymatic or chemical means know to the art.

The term "5' to 3' exonuclease activity" refers to the presence of an activity in a protein that is capable of removing nucleotides from the 5' end of an oligonucleotide. 5' to 3' exonuclease activity may be measured using any of the assays provided herein or known in the art. Some enzymes contain 5' to 3' "exoribonuclease" activity, the ability to remove nucleotides from the 5' end of ribonucleic acid molecules. Some enzymes that are traditionally called "exoribonucleases" may also have 5'-exonuclease activity against DNA substrates. For example, enzymes of the present invention, while historically named "exoribonucleases", have 5'-exonuclease activity on DNA substrates that have a 5'-phosphate, meaning a phosphate group on the 5'-position of the deoxyribose sugar of the nucleotide at the 5'-end of the DNA. Thus, it is understood that such exoribonucleases (e.g., Xrn1 5' exoribonuclease 1), unless specified otherwise, have exonuclease activity for both RNA and DNA substrates.

The terms "cell," "cell line," "host cell," as used herein, are used interchangeably, and all such designations include progeny or potential progeny of these designations. The words "transformants" or "transformed cells" include the primary transformed cells

IDT Exhibit 1020.6379

derived from that cell without regard to the number of transfers. All progeny may not be precisely identical in DNA content, due to deliberate or inadvertent mutations. Nonetheless, mutant progeny that have the same functionality as screened for in the originally transformed cell are included in the definition of transformants.

Nucleic acids are known to contain different types of mutations. A "point" mutation refers to an alteration of a base nucleotide at a single nucleotide position in the sequence compared to the wild type sequence. Mutations may also refer to insertion or deletion of one or more bases, so that the nucleic acid sequence differs from the wild-type sequence.

The term "homology" refers to a degree of complementarity of one nucleic acid sequence with another nucleic acid sequence. There may be partial homology or complete homology (i.e., complementarity). A partially complementary sequence is one that at least partially inhibits a completely complementary sequence from hybridizing to a target nucleic acid and is referred to using the functional term "substantially homologous." The inhibition of hybridization of the completely complementary sequence to the target sequence may be examined using a hybridization assay (Southern or Northern blot, solution hybridization and the like) under conditions of low stringency. A substantially homologous sequence or probe will compete for and inhibit the binding (i.e., the hybridization) of a completely homologous sequence to a target under conditions of low stringency. This is not to say that conditions of low stringency are such that non-specific binding is permitted; low stringency conditions require that the binding of two sequences to one another be a specific (i.e., selective) interaction. The absence of non-specific binding may be tested by the use of a second target that lacks complementarity or that has only a low degree of complementarity (e.g., less than about 30% complementarity). In the case in which specific binding is low or non-existent, the probe will not hybridize to a nucleic acid target.

When used in reference to a double-stranded nucleic acid sequence such as a cDNA or a genomic clone, the term "substantially homologous" refers to any probe which can hybridize to either or both strands of the double-stranded nucleic acid sequence under conditions of low stringency as described herein.

IDT Exhibit 1020.6380

As used herein, the terms "hybridization" or "annealing" are used in reference to the pairing of complementary nucleic acid strands. Hybridization and the strength of hybridization (i.e., the strength of the association between nucleic acid strands) is impacted by many factors well known in the art including the degree of complementarity between the nucleic acids, stringency of the conditions involved affected by such conditions as the concentration of salts, the T_m (melting temperature) of the formed hybrid, the presence of other components (e.g., the presence or absence of polyethylene glycol or betaine), the molarity of the hybridizing strands and the G:C content of the nucleic acid strands.

The term "isolated" when used in relation to a nucleic acid, as in "isolated polynucleotide " or "isolated oligonucleotide " refers to a nucleic acid sequence that is identified and separated from at least one contaminant with which it is ordinarily associated in its source. Thus, an isolated nucleic acid is present in a form or setting that is different from that in which it is found in nature. In contrast, non-isolated nucleic acids (e.g., DNA and RNA) are found in the state they exist in nature. For example, a given DNA sequence (e.g., a gene) is found on the host cell chromosome in proximity to neighboring genes; and a specific RNA sequences (e.g., a specific mRNA sequence encoding a specific protein), is found in the cell as a mixture with numerous other mRNAs that encode a multitude of proteins. The isolated polynucleotide or nucleic acid or oligonucleotide may be present in single-stranded or double-stranded form. When an isolated polynucleotide or nucleic acid is to be utilized to express a protein, the polynucleotide contains at a minimum, the sense or coding strand (i.e., the polynucleotide may be single-stranded), but may contain both the sense and anti-sense strands (i.e., the polynucleotide may be double-stranded).

As used herein, the term "purified" or "to purify" means the result of any process that removes some of a contaminant from the component of interest, such as a protein or nucleic acid. The percent of a purified component is thereby increased in the sample.

As used herein, the term "enzyme" refers to molecules or molecule aggregates that are responsible for catalyzing chemical and biological reactions. Such molecules are typically proteins, but can also comprise short peptides, RNAs, ribozymes, antibodies,

IDT Exhibit 1020.6381

and other molecules. A molecule that catalyzes chemical and biological reactions is referred to as "having enzyme activity" or "having catalytic activity."

In addition to 5'-phosphate-dependent exonucleases, a variety of other enzymes are used in the methods of the present invention. In general, the invention comprises use of any enzyme from any source that has an activity that is equivalent to an enzyme referred to in a method so long as the enzyme functions in the same way in the particular method. By way of example, but not of limitation, an RNA-dependent DNA polymerase can comprise an AMV reverse transcriptase; an MMLV reverse transcriptase; SuperScript I, SuperScript II, SuperScript III, or AMV ThermoScript reverse transcriptase (INVITROGEN); or MonsterScript reverse transcriptase (EPICENTRE), or it can comprise another enzyme that has similar activity in the method; a polynucleotide kinase can comprise T4 polynucleotide kinase or it can comprise another enzyme that has similar activity in the method; a poly(A) polymerase can comprise an *E. coli* poly(A) polymerase encoded by the *pcn*B gene or it can comprise another enzyme that has similar activity in the method; a ribonuclease H can comprise *E. coli* RNase H or a *Thermus* RNase H (e.g., Hybridase RNase H from EPICENTRE) or it can comprise another enzyme that has similar activity in the method; a pyrophosphatase can be tobacco acid pyrophosphatase or it can comprise another enzyme that has similar activity in the method; and a DNA phosphatase can comprise APex Alkaline Phosphatase (EPICENTRE) or shrimp alkaline phosphatase or it can comprise another enzyme that has similar activity in the method. Methods and reaction conditions for use of these and other enzymes that can be used are well known in the art.

An "RNA amplification method" according to the present invention is a method that synthesizes an RNA product and that results in an increase in the number of copies of an RNA sequence or of its complementary sequence compared to the number of copies of the sequence present in a sample. A number of RNA amplification methods are known in the art that comprise steps for joining an RNA polymerase promoter to either the 5'-end of the cDNA or the 3'-end of cDNA that is made by reverse transcription of mRNA using an RNA-dependent DNA polymerase to extend a primer, such as, but not limited to an oligo(dT)-containing primer, which methods result in synthesis of amplified antisense RNA or amplified sense RNA, respectively, by subsequent *in vitro* transcription (e.g.,

IDT Exhibit 1020.6382

using T7, T3, or SP6 RNA polymerase). Kits and methods of the present invention that comprise RNA amplification can comprise any compositions or method, including without limitation, any compositions or methods known in the art. By way of example, but not of limitation, a method that uses an oligo(dT) promoter primer to synthesize an antisense RNA (aRNA), such as methods described by Van Gelder, R.N., et al. (Proc. Natl. Acad. Sci. USA 87: 1663, 1990) can be used; such methods are sometimes referred to as "Eberwine-type" amplification methods herein. Kits for this purpose are commercially available and can be used, including 1-round and 2-round amplification kits such as various 1-round and 2-round TargetAmp™ Aminoallyl-aRNA Amplification Kits or TargetAmp™ aRNA Amplification Kits available from EPICENTRE, or MessageAmp kits available from Ambion (Austin, TX). Still further, a sense RNA amplification method that attaches a promoter to the 3'-end of first-strand cDNA for synthesis of amplified sense RNA, such as methods described by Che, S. and Ginsberg, S.D (Laboratory Investigation 84: 131, 2004) and by Ginsberg and Che (Neurochemical Res. 27: 981, 2002; and PCT Internatl. Publication No. WO 02/065093), can be used in a method or kit of the present invention. The BD SMART mRNA amplification kit (BD Biosciences Clontech, Palo Alto, CA) is an example of a commercial kit of a type that can be used for sense RNA amplification in methods and kits of the present invention. Still another RNA amplification method that can be used in the methods of the present invention uses a oligonucleotide sequence tag template to add a tag sequence to the 3'-end of a target sequence comprising mRNA or first-strand cDNA, which tag sequence can be further used to add a promoter for synthesis of RNA corresponding to the target sequence, as described in U.S. Patent Application No. 2005/015333 of Roy R. Sooknanan, which is incorporated herein by reference. Examples of other methods for RNA amplification that can be used are described in: Murakawa et al., DNA 7:287-295, 1988; Phillips and Eberwine, Methods in Enzymol. Suppl. 10:283-288, 1996; Ginsberg et al., Ann. Neurol. 45:174-181, 1999; Ginsberg et al., Ann. Neurol. 48:77-87, 2000; VanGelder et al., Proc. Natl. Acad. Sci. USA 87:1663-1667, 1990; Eberwine et al., Proc. Natl. Acad. Sci. USA 89:3010-3014, 1992; U.S. Patent Nos. 5,021,335; 5,168,038; 5,545,522; 5,514,545; 5,716,785; 5,891,636; 5,958,688; 6,291,170; and PCT Patent

IDT Exhibit 1020.6383

Applications WO 00/75356 and WO 02/065093, which is incorporated herein by reference.

The present invention is also not limited to RNA amplification methods that require synthesis of double-stranded cDNA. By way of example, but without limitation, the present invention also comprises RNA amplification methods and compositions as described in U.S. Patent Appln. No. 2004/0171041 that use an RNA polymerase that can synthesize RNA using single-stranded templates that are functionally joined to a single-stranded promoter, such as, but not limited to methods that use MiniV RNA polymerase (available from EPICENTRE in the MiniV™ *In Vitro* Transcription Kit); in these embodiments, a single-stranded promoter is joined to either the 5'-end of the cDNA or the 3'-end of cDNA that is made by reverse transcription of mRNA using an RNA-dependent DNA polymerase to extend a primer, resulting in synthesis of amplified antisense RNA or amplified sense RNA, respectively, by subsequent *in vitro* transcription of single-stranded DNA templates (e.g., using MiniV RNA polymerase).

A "T7-type RNA polymerase" as defined herein is a wild-type or mutant form of an RNA polymerase derived from a T7-type bacteriophage, including both phage-encoded enzymes and enzymes obtained by cloning the RNA polymerase gene in a DNA vector and expressing it in a bacterial or other cell. This is based on the fact that the genetic organization of all T7-type bacteriophage that have been examined has been found to be essentially the same as that of T7. Examples of T7-type bacteriophages according to the invention include, but are not limited to Escherichia coli phages T3, phi I, phi II, W31, H, Y, A1, 122, cro, C21, C22, and C23; Pseudomonas putida phage gh-1; Salmonella typhimurium phage SP6; Serratia marcescens phages IV; Citrobacter phage ViIII; and Klebsiella phage No. 11 (Hausmann, Current Topics in Microbiology and Immunology 75:77-109, 1976; Korsten et al., J. Gen. Virol. 43:57-73, 1975; Dunn, et al., Nature New Biology 230:94-96, 1971; Towle, et al., J. Biol. Chem. 250:1723-1733, 1975; Butler and Chamberlin, J. Biol. Chem. 257:5772-5778, 1982). Mutant RNAPs (Sousa et al., U.S. Patent No. 5,849,546; Padilla, R and Sousa, R, Nucleic Acids Res., 15: e138, 2002; Sousa, R and Mukherjee, S, Prog Nucleic Acid Res Mol Biol., 73: 1-41, 2003), such as, but not limited to, T7 RNAP Y639F mutant enzyme, T3 RNAP Y573F mutant enzyme, SP6 RNAP Y631F mutant enzyme, T7 RNAP having altered amino

IDT Exhibit 1020.6384

acids at both positions 639 and 784, T3 RNAP having altered amino acids at both positions 573 and 785, or SP6 RNAP having altered amino acids at both positions 631 and 779 can also be used in some embodiments of methods or assays of the invention. In particular, such mutant enzymes can corporate dNTPs and 2'-F-dNTPs, in addition to ddNTPs and certain other substrates, which are advantageous for synthesis of RNA molecules with specific properties and uses. Phage N4 mini-vRNAP, which has certain domains in common with T7-type RNAPs, and which is also an RNA polymerase of the present invention, is a transcriptionally active 1,106-amino acid domain of the N4 vRNAP, which corresponds to amino acids 998-2103 of N4 vRNAP (Kazmierczak, K.M., et al., EMBO J., 21: 5815-5823, 2002; U.S. Patent Application No. 20030096349, incorporated herein by reference) can be used. Alternatively, an N4 mini-vRNAP Y678F mutant enzyme (U.S. Patent Application No. 20030096349) can incorporate non-canonical nucleotides such as 2'-F-dNTPs. In order to carry out transcription, a RNA polymerase recognizes and binds to a DNA sequence of approximately 25 nucleotides in length called an "RNA polymerase promoter," a "transcription promoter" or simply a "promoter," and initiates transcription therefrom. In most cases, the promoter sequence is double-stranded. As used herein, the strand of a double-stranded promoter that is covalently joined to the template strand for synthesis of RNA is defined as the "sense strand" or "sense promoter sequence" and its complement is defined as the "anti-sense strand" or the "anti-sense promoter sequence."

Still further, the term "RNA amplification method" according to the present invention also includes other methods, such as but not limited to methods described in U.S. Pat. Appln. No. 2004/0180361 of Dahl et al.; PCT Patent Publication Nos. WO 02/16639; WO 00/56877; and AU 00/29742 of Takara Shuzo Company; and in U.S. Pat. No. 6,251,639 and U.S. Pat. Appln. Nos. 2001/0034048; 2003/0017591; 2003/0087251; and 2003/0186234 of Kurn (all incorporated herein by reference), even if said methods comprise synthesis DNA products rather than RNA products corresponding to mRNA in the sample. Thus, the present invention also comprises embodiments in which a nucleic acid of interest (or a desired nucleic acid) obtained following treatment with a 5'-phosphate-dependent exonuclease of the present invention is amplified using any of said

IDT Exhibit 1020.6385

RNA amplification methods, including those that comprise synthesis of DNA amplification products.

A nucleic acid of interest (or a desired nucleic acid) obtained following treatment with a 5'-phosphate-dependent exonuclease of the present invention can also be amplified by PCR or reverse transcription-PCR (i.e., RT-PCR), including without limitation real-time PCR or real-time RT-PCR. Amplification of an mRNA of interest or a cDNA of interest by RT-PCR or PCR, respectively, is useful for validation of results of gene expression analysis using other methods, such as but not limited to, gene expression analysis results obtained by hybridization of labeled RNA or DNA corresponding to mRNA from a sample to nucleic acid arrays or microarrays.

"Rolling Circle Replication" and "Rolling Circle Transcription" are two other methods that can be used to amplify a nucleic acid of interest in a method of the present invention, which methods are defined and described in a number of patents and publications below, which are incorporated herein by reference. In PCT Patent Application No. WO 92/01813, Ruth and Driver disclosed a process for synthesizing circular single-stranded nucleic acids by hybridizing a linear polynucleotide to a complementary oligonucleotide and then ligating the linear polynucleotide. They further disclosed a process for generating multiple linear complements of the circular single-stranded nucleic acid template by extending a primer more than once around the circular template using a DNA polymerase. Japanese Patent Nos. JP4304900 and JP4262799 of Aono Toshiya et al. disclose detection of a target sequence by ligation of a linear single-stranded probe having target-complementary 3'- and 5'-end sequences which are adjacent when the linear probe is annealed to a target sequence in the sample, followed by either rolling circle replication or *in vitro* transcription of the circular single-stranded template. The inventors disclose that in vitro transcription is performed by first annealing to the circular single-stranded template a complementary nucleotide primer having an anti-promoter sequence in order to form a double-stranded promoter, and then transcribing the circular single-stranded template having the annealed anti-promoter primer with an RNA polymerase that has helicase-like activity, such as T7, T3 or SP6 RNA polymerase. In U.S. Patent Nos. 6,344,329; 6,210,884; 6,183,960; 5,854,033; 6,329,150; 6,143,495; 6,316,229; and 6,287,824, Paul M. Lizardi also used rolling circle replication to amplify

IDT Exhibit 1020.6386

and detect nucleic acid sequences. In a series of articles and patents, Eric Kool and co-workers disclosed synthesis of DNA or RNA multimers, meaning multiple copies of an oligomer or oligonucleotide joined end to end (i.e., in tandem) by rolling circle replication or rolling circle transcription, respectively, of a circular DNA template molecule. Rolling circle replication uses a primer and a strand-displacing DNA polymerase, such as phi29 DNA polymerase. With respect to rolling circle transcription, it was shown that circular single-stranded DNA (ssDNA) molecules can be efficiently transcribed by phage and bacterial RNA polymerases (Prakash, G. and Kool, E., J. Am. Chem. Soc. 114: 3523-3527, 1992; Daubendiek, S.L. et al., J. Am. Chem. Soc. 117: 7818-7819, 1995; Liu, D. et al., J. Am. Chem. Soc. 118: 1587-1594, 1996; Daubendiek, S.L. and Kool, E.T., Nature Biotechnol., 15: 273-277, 1997; Diegelman, A.M. and Kool, E.T., Nucleic Acids Res., 26: 3235-3241, 1998; Diegelman, A.M. and Kool, E.T., Chem. Biol., 6: 569-576, 1999; Diegelman, A.M. et al., BioTechniques 25: 754-758, 1998; Frieden, M. et al., Angew. Chem. Int. Ed. Engl. 38: 3654-3657, 1999; Kool, E.T., Acc. Chem. Res., 31: 502-510, 1998; U.S. Patent Nos. 5,426,180; 5,674,683; 5,714,320; 5,683,874; 5,872,105; 6,077,668; 6,096,880; and 6,368,802). Rolling circle transcription of these circular ssDNAs occurs in the absence of primers, in the absence of a canonical promoter sequence, and in the absence of any duplex DNA structure, and results in synthesis of linear multimeric complementary copies of the circle sequence up to thousands of nucleotides in length. Transcription of the linear precursor of the circular ssDNA template yielded only a small amount of RNA transcript product that was shorter than the template. Fire and Xu (U.S. Patent No. 5,648,245; Fire, A. and Xu, S-Q, Proc. Natl. Acad. Sci. USA, 92: 4641-4645, 1995) also disclose methods for using rolling circle replication of small DNA circles to construct oligomer concatamers. Other researchers, including, but not limited to, Mahtani (U.S. Patent No. 6,221,603), Rothberg et al. (U.S. Patent No. 6,274,320), Dean et al. (Genome Res., 11: 1095-1099, 2001), Lasken et al. (U.S. Patent No. 6,323,009), and Nilsson et al. (Nucleic Acids Res., 30(14): e66, 2002) disclose other methods and applications of rolling circle amplification. Pickering et al. (Nucleic Acids Res., 30(12): e60, 2002) discloses a ligation and rolling circle amplification method for homogeneous end-point detection of single nucleotide polymorphisms (SNPs).

IDT Exhibit 1020.6387

As used herein, the terms "buffer" or "buffering agents" refer to materials that when added to a solution, cause the solution to resist changes in pH.

As used herein, the terms "reducing agent" and "electron donor" refer to a material that donates electrons to a second material to reduce the oxidation state of one or more of the second material's atoms.

The term "monovalent salt" refers to any salt in which the metal (e.g., Na, K, or Li) has a net 1+ charge in solution (i.e., one more proton than electron).

As used herein, the term "divalent salt" refers to any salt in which a metal (e.g., Mg, Mn, Ca, or Sr) has a net 2+ charge in solution.

As used herein, the terms "chelator" or "chelating agent" refer to any materials having more than one atom with a lone pair of electrons that are available to bond to a metal ion.

As used herein, the term "solution" refers to an aqueous or non-aqueous mixture.

As used herein, the term "buffering solution" refers to a solution containing a buffering agent.

As used herein, the term "reaction buffer" refers to a buffering solution in which an enzymatic reaction is performed.

As used herein, the term "storage buffer" refers to a buffering solution in which an enzyme is stored.

IDT Exhibit 1020.6388

In keeping with standard polypeptide nomenclature, J. Biol. Chem., 243:3557-3559, 1969, abbreviations for amino acid residues are as shown in the following Table of Correspondence.

TABLE OF CORRESPONDENCE

1-Letter	3-Letter	AMINO ACID
Y	Tyr	tyrosine
G	Gly	glycine
F	Phe	phenylalanine
M	Met	methionine
A	Ala	alanine
S	Ser	serine
I	Ile	isoleucine
L	Leu	leucine
T	Thr	threonine
V	Val	valine
P	Pro	proline
K	Lys	lysine
H	His	histidine
Q	Gln	glutamine
E	Glu	glutamic acid
W	Trp	tryptophan
R	Arg	arginine
D	Asp	aspartic acid
N	Asn	asparagine
C	Cys	cysteine

The phrase "mixed population" or "complex population" refers to any sample containing both desired and undesired nucleic acids. As a non-limiting example, a complex population of nucleic acids may be total genomic DNA, total cellular RNA or a combination thereof. Moreover, a complex population of nucleic acids may have been enriched for a given population, but include other undesirable populations. For example, a complex population of nucleic acids may be a sample that has been enriched for desired messenger RNA (mRNA) sequences but still includes some undesired ribosomal RNA sequences (rRNA).

The term "undetectable levels" of exonuclease activity with respect to a particular nucleic acid substrate refers to levels of digestion of a nucleic acid substrate that are not detectable within the assay or method in which an exonuclease is employed. Undetectable levels include the complete absence of digestion. Although the level of digestion of a nucleic acid of interest (i.e., a desired nucleic acid) can be undetectable in some embodiments, the invention is not limited to embodiments in which digestion of the

IDT Exhibit 1020.6389

nucleic acid of interest by a 5'-phosphate-dependent exonuclease of the invention is undetectable so long as a substantial amount of the undesired nucleic acid is digested and sufficient amount of the desired nucleic acid remains for the intended purpose, which can vary for different purposes. In preferred embodiments of the invention, a substantial amount of the undesired nucleic acid is digested by the 5' exonuclease of the invention, which means that at least 50% (e.g., 60%, 70%, 80%, 90%, 95%, 98%, 99%) of the starting amount (i.e., the amount present prior to treatment with the exonuclease) of the undesired nucleic acid is digested.

The terms "sample" and "specimen" are used in their broadest sense and encompass samples or specimens obtained from any source including biological and environmental sources. As used herein, the term "sample" when used to refer to biological samples obtained from organisms, includes, but it not limited to fluids, solids, tissues, and gases. In preferred embodiments of this invention, biological samples include bodily fluids, isolated cells, fixed cells, cell lysates and the like. However, these examples are not to be construed as limiting the types of samples that find use with the present invention.

A "target nucleic acid" comprises at least one nucleic acid molecule or portion of at least one nucleic acid molecule, whether the molecule or molecules is or are DNA, RNA, or both DNA and RNA, and wherein each the molecule has, at least in part, a defined nucleotide sequence, which is referred to as the "target sequence." A goal of an assay or method of the invention is to detect whether or not the target nucleic acid is present in a sample by means of methods that can detect the target sequence. This is frequently, but without limitation, achieved using a "nucleic acid probe" that is complementary to the target sequence (e.g., a probe that is present on an array or microarray together with other probes that are complementary to other nucleic acid sequences). The target nucleic acid may also have at least partial complementarity with other molecules used in an assay, such as, but not limited to, primers, splice template oligos, ligation splint oligos, capture probes or detection probes. The target nucleic acid may be single- or double-stranded and may be of any length. However, it must comprise a polynucleotide sequence of sufficient sequence specificity and length so as to be useful for its intended purpose. By way of example, but not of limitation, a target nucleic acid

IDT Exhibit 1020.6390

that is to be detected using a sequence-complementary detection probe must have a sequence of sufficient sequence specificity and length so as remain hybridized by the detection probe under assay hybridization conditions wherein sequences that are not target nucleic acids are not hybridized. A target nucleic acid having sufficient sequence specificity and length for an assay of the present invention may be identified, using methods known to those skilled in the art, by comparison and analysis of nucleic acid sequences known for a target and for other sequences which may be present in the sample. For example, sequences for nucleic acids of many viruses, bacteria, humans (e.g., for genes and messenger RNA), and many other biological organisms can be searched using public or private databases, and sequence comparisons, folded structures, and hybridization melting temperatures (i.e., T_m's) may be obtained using computer software known to those knowledgeable in the art. The term "source of target nucleic acid" refers to any sample that contains a naturally occurring target nucleic acid, RNA or DNA.

DETAILED DESCRIPTION OF THE INVENTION

The present invention relates to compositions and methods employing 5'-phosphate-dependent nucleic acid exonucleases. In particular, the present invention provides kits and methods employing 5'-phosphate-dependent nucleic acid exonucleases for a variety of uses where it is desired to digest nucleic acid molecules having 5'-phosphate groups.

As defined herein, any exoribonuclease that selectively digests nucleic acid molecules having 5'-phosphate groups may be used in the methods and kits of the present invention. A 5'-to-3' exoribonuclease of the present invention comprises any 5'-to-3' exonuclease that has greater than 20-fold more 5'-to-3' exonuclease activity for a single-stranded RNA substrate that has a 5'-monophosphorylated terminus than for the same RNA substrate that has a 5'-triphosphorylated or 5'-capped terminus. That is, a 5' exonuclease of the present invention comprises any 5'-to'3' exonuclease that has a relative activity in digesting a particular defined-sequence single-stranded RNA substrate with a 5'-triphosphate or a 5'-cap that is less than 5% of the activity of an RNA substrate having the same sequence but with a 5'-monophosphate. Without limitation, suitable methods that can be used for assaying activity and determining relative activity using RNA

IDT Exhibit 1020.6391

substrates with a 5'-triphosphate, a 5'-cap, or a 5'-monophosphate are described by Stevens and Poole (J. Biol. Chem., 270: 16063, 1995). Preferred embodiments employ Saccharomyces cerevisiae Xrn1p/5' exoribonuclease 1 and/or functional variants and homologues thereof.

In some embodiments, the present invention provides a method of preparing a nucleic acid sample for analysis. It is often desirable to isolate, enrich, or increase the relative percentage of a particular population of sequences within a much larger population of sequences in order to limit analysis to sequences of interest and to reduce interference and unnecessary work that may be caused by the presence of undesirable sequences. The present invention provides a novel method wherein a complex sample or a mixed population of nucleic acids (e.g., containing both nucleic acids of interest and nucleic acids not of interest) is substantially depleted of undesired sequences (e.g., nucleic acids not of interest) and is thus enriched for a population of interest. For example, in preferred embodiments, the present invention provides a method of enriching a nucleic acid population of interest present in a complex sample by treating the sample with a 5'-phosphate-dependent nucleic acid exonuclease, thereby increasing the relative percentage of the nucleic acid population of interest in a given sample for further analysis. In preferred embodiments, the 5'-phosphate-dependent nucleic acid exonuclease digests greater than 50% (e.g., 60%, 70%, 80%, 90%, 95%, 97%, 99%, 99.5%, 99.9%) of the undesired sequence or sequences. Also, in preferred embodiments, the 5'-phosphate-dependent nucleic acid exonuclease digests less than 50%, (e.g., 40%, 30%, 20%, 15%, 10%, 5%, 4%, 3%, 2%, 1%) of the desired sequence or sequences, and most preferably, undetectable amounts of the desired sequence or sequences, but the invention also comprises embodiments in which more of the desired sequence or sequences is/are digested so long as the amount remaining can be used for the intended purpose and a substantial amount of the undesired sequence or sequences is/are digested.

For example, in some embodiments, methods of the present invention enrich for a nucleic acid population of interest (e.g., mRNA sequences) within a mixed population of nucleic acid sequences by targeting undesired sequences (e.g., 16S and 23S prokaryotic rRNA or 18S and 26S or 28S eukaryotic rRNA sequences comprising a 5' phosphate) for digestion by a 5'-phosphate-dependent exonuclease and eliminating them from the mixed

IDT Exhibit 1020.6392

population. In some embodiments, the undesired sequences are degraded by the enzymatic activity of xrn-1 exoribonuclease (See, e.g., Example 1, Table 1).

The mixed population of nucleic acids may be any nucleic acid sample comprising both desired and undesired sequences. The population may include different DNA and/or RNA molecules. In some preferred embodiments, the mixed population is an RNA sample. In further preferred embodiments, the nucleic acid sample is RNA derived from a prokaryotic organism. In other preferred embodiments, the nucleic acid sample comprises RNA from a eukaryotic organism. In still other embodiments, the sample comprises RNA from multiple prokaryotic organisms, or from both prokaryotic and eukaryotic organisms, or from multiple eukaryotic organisms. The mixed population may be derived from a wide variety of sources including for example, tissue samples, blood, isolated cells or environmental samples such as water or soil. The mixed population may be derived from any organism including both eukaryotes and prokaryotes such as human, rat, mouse, Escherichia coli (E. coli), Bacillus subtilis (B. subtilis), Pseudomonas aeruginosa, etc. Methods of deriving nucleic acid samples from eukaryotic and prokaryotic organisms are well known to those of skill in the art (See, e.g., Chapter 4, "Current Protocols in Molecular Biology," Ausubel et al., eds (1997 supplement) Johan Wilen & Sons, Inc. and Chapter 7, Sambrook, Fritsch, Maniatis "Molecular Cloning: A Laboratory Manual" (1989) Cold Spring Harbor Press, etc.).

The population of interest may comprise a subset of the mixed population. The population of interest may include RNA and/or DNA. The population of interest may, for example, be a particular type of RNA. In a preferred embodiment the population of interest is mRNA. The population of interest may comprise any sequence or mixture of sequences and the sequence or mixture of sequences need not be known. The population of interest may be chosen on a variety of bases, including by sequence, function (e.g. messenger RNA (mRNA), ribosomal RNA (rRNA), transfer RNA (tRNA), etc.), or a combination thereof. In some embodiments, the population of interest is provided in a sample that has been depleted of 5S rRNA and/or tRNA. Such a mixture may be generated by salt precipitation using LiCl. The larger rRNAs and mRNAs are precipitated, but the tRNAs and 5S rRNAs remain in solution and can be removed. In some embodiments, the tRNAs and 5S rRNAs are removed by LiCl precipitation after a

IDT Exhibit 1020.6393

sample is treated with exoribonuclease. Kits for conducting such reactions are provided by the present invention (e.g., kits containing LiCl and exoribonuclease). Other methods for isolating subsets of molecules (e.g., PEG precipitation, etc.) may also be employed in the kits and methods of the present invention.

The sequences that are targeted for digestion may comprise undesired sequences in the mixed population. The sequences that are targeted for digestion may comprise any sequence so long as they are distinguishable from the population of interest (e.g., the sequences that are targeted for digestion possesses an accessible 5' phosphate whereas the population of interest does not). Targeted double-stranded DNA (dsDNA) or RNA (dsRNA) or single-stranded DNA (ssDNA) or RNA (ssRNA) that self-anneals to form double-stranded regions can be denatured by various treatments known in the art, such as, but not limited to, treatment with heat or addition of compositions that affect the melting temperature of the targeted polynucleotide, such as, but not limited to betaine. In preferred embodiments, the sequences that are targeted for digestion are RNAs including rRNA (e.g., 16S and 23S prokaryotic rRNA, or 18S and 26S or 28S eukaryotic rRNA). In some embodiments, it may not be necessary to remove substantially all the undesired sequences from the mixed population. In these embodiments it is acceptable to remove only enough of the undesired sequences such that the undesired sequences do not interfere with analysis of the population of interest.

In preferred embodiments, any non-targeted undesirable sequences represent only a small proportion of the mixed population or are readily removed using other methods known in the art. These non-targeted undesirable sequences may include a variety of other nucleic acids such as DNAs, rRNAs, mRNAs or tRNAs. For the sake of simplicity, the presence of non-targeted RNAs will not be discussed throughout the remainder of the application; however, the possibility of their presence is contemplated by the scope of the presently claimed invention, unless described otherwise.

A particular example of the presently claimed invention provides a method of isolating or enriching for mRNAs within a mixed population of RNAs by specifically removing rRNAs >300 nucleotides that possess an accessible 5' phosphate. A mixed population of RNAs may include, for example, mRNAs, tRNAs and rRNAs.

IDT Exhibit 1020.6394

In some embodiments, methods of the present invention are used in combination with methods that expose a 5' phosphate on an undesired nucleic acid sequence (e.g., a method that removes a 5'-triphosphate or a 5' cap on an undesired nucleic acid sequence using an agent, such as, but not limited to, tobacco acid pyrophosphatase).

In other embodiments, methods of the present invention are used in combination with methods that result in addition of a 5'-phosphate to a nucleic acid molecule of interest (e.g., a method for 5'-phosphorylation using a polynucleotide kinase, such as, but not limited to, T4 polynucleotide kinase).

In other embodiments, methods of the present invention are used in combination with methods that result in addition of a cap to the 5'-terminus of a nucleic acid sequence (e.g., a method for 5'-capping of an mRNA molecule having a 5'-triphosphate using a guanyltransferase, such as, but not limited to, *Vaccinia* virus capping enzyme).

In some preferred embodiments, the sample containing the nucleic acids of interest (e.g., mRNA sequences) and the undesired nucleic acids (e.g., rRNA sequences) is partially purified prior to treatment using a composition or method of the present invention. By way of example, but not of limitation, total RNA can be purified from a sample using methods known in the art, including, for example, commercially available purification kits such as the MasterPure™ RNA Purification Kit (EPICENTRE Technologies, Madison, WI) or the RNeasy Kit (Qiagen, Valencia, CA) prior to treatment with an exoribonuclease according to the present invention.

In preferred embodiments, once the undesired nucleic acids (e.g., rRNA sequences) are eliminated, the nucleic acids of interest (e.g., mRNA sequences) are further purified using methods known in the art, including, for example, commercially available purification kits such as the MasterPure™ Complete DNA and RNA Purification Kit (EPICENTRE Technologies, Madison, WI) or the RNeasy Kit (Qiagen, Valencia, CA).

In some embodiments, the presently claimed invention provides a method of differentiating between nucleic acid sequences. For example, the present invention provides a method of distinguishing nucleic acid sequences that comprise an accessible 5' phosphate (e.g., that are targeted for digestion by a 5'-phosphate-dependent exonuclease) from nucleic acid sequences that do not possess an accessible 5' phosphate (e.g., that are

IDT Exhibit 1020.6395

not targeted for digestion by xrn-1). Thus, the present invention provides diagnostic methods. A wide variety of diagnostic methods are provided by the invention. Any method that utilizes an exoribonuclease to distinguish between nucleic acid molecules having and lacking 5' phosphate groups is encompassed by the present invention. For example, the exoribonucleases may be used to identify nucleic acid molecules that have been digested by restriction enzyme—generating a 5' phosphate—as compared to undigested nucleic acid molecules lacking a 5' phosphate.

In some embodiments, the method of differentiating between nucleic acids is used in combination with known diagnostic methods. For example, in some embodiments, methods of using exoribunucleases of the present invention are combined with restriction enzyme digestion used in restriction fragment length polymorphism (RFLP) analysis or other diagnostic methods.

In some embodiments, the enzymatic activity of the exoribonuclease of the present invention is controlled by altering the concentration of a divalent cation (e.g., Mg^{2+}) present in a reaction solution. In some embodiments, the enzymatic activity of exoribonuclease of the present invention is controlled by addition of a chelating agent.

EXAMPLES

The following examples serve to illustrate certain preferred embodiments and aspects of the present invention and are not to be construed as limiting the scope thereof.

Example 1

Activity of Exoribonuclease I

This Example describes the use of Exoribonuclease I in enriching for specific RNA substrates in the presence of other undesired nucleic acids.

A. Materials and Methods

Enzyme

Exoribonuclease I (xrn-1) (Stevens, Biochem. Biophys. Res. Commun. 81:656, 1978); Stevens, Biochem. Biophys. Res. Commun. 86:1126, 1979); Stevens, J. Biol.

IDT Exhibit 1020.6396

Chem 255:3080, 1980); and Stevens and Maupin, Nucleic Acids Res. 15:695, 1987) was purified from a recombinant source using methods similar to those described in the art. The enzyme was stored in 50% (v/v) glycerol containing, 0.05 M Tris-HCl (pH 7.5), 100 mM NaCl, 0.1 mM EDTA, 1 mM DTT, and 0.1% Triton X-100.

One unit of exoribonuclease I is the amount of enzyme that converts 1 μg of *E. coli* ribosomal RNA (comprising 16S and 23S rRNA) to acid-soluble form in 60 minutes at 30°C under standard reaction assay conditions. Exoribonuclease I used in the experiments described herein is free of other detectable contaminating ribonucleases, as well as free from detectable contaminating exo- and endodeoxyribonucleases (although the enzyme itself has DNase activity).

Reaction Conditions

All experiments were performed in a reaction buffer containing 50 mM Tris-HCl (pH 8.0), 2.0 mM $MgCl_2$, and 100 mM NaCl.

B. Results

Capped RNA is resistant to xrn-1 degradation

A 75-nucleotide, capped transcript was made with AmpliCap™ SP6 High Yield Message Maker Kit from EPICENTRE (Madison, WI). Samples containing this capped RNA were treated with tobacco acid pyrophosphatase (TAP; EPICENTRE) to remove the cap (or "de-cap") the RNA and/or with xrn-I, as indicated in Table 1, and analyzed by polyacrylamide gel electrophoresis. The results, shown in Table 1, demonstrated that the 5'-capped RNA was resistant to digestion (or "degradation") by xrn-1. De-capping the RNA with TAP resulted in degradation of the RNA by xrn-1 exonuclease.

IDT Exhibit 1020.6397

Table 1		
TAP Treated	Xrn-1 Treated	Degradation
-	-	-
-	+	-
+	-	-
+	+	+

Degradation of ribosomal RNA

A preparation of increasing amounts of tobacco RNA isolated from tobacco plants was treated with xrn-1 exonuclease and analyzed by agarose gel electrophoresis. Minimal RNA was visible in the lanes containing RNA treated with xrn-1. These results showed that the xrn-1 exonuclease was able to digest abundant ribosomal RNA.

The experiment was next repeated with a mixture of the tobacco RNA and a 1.4-kB capped mRNA transcript prepared using an AmpliCap™ T7 High Yield Message Maker Kit (EPICENTRE), which was added to the preparation of tobacco RNA, and with the 1.4-kB transcript alone. Results of the reactions were analyzed by agarose gel electrophoresis. The capped mRNA was not degraded by xrn-1. In the lanes containing both the tobacco RNA and the capped mRNA treated with xrn-1, the ribosomal RNA bands corresponding to the tobacco RNA were degraded, but those corresponding to the capped mRNA were not degraded. The present invention is not limited to a particular mechanism. Indeed, an understanding of the mechanism is not necessary to practice the present invention. Nonetheless, it is contemplated that the tobacco mRNA remains undegraded (see above) but is not visible due to its low concentration.

The xrn-1-treated tobacco RNA was concentrated approximately 50X by ethanol precipitation and analyzed on an agarose gel. The results indicated that the bulk of the RNA was degraded, but that a small amount of xrn-1-resistant RNA was present that appeared as a heterodisperse size range that could be observed on an agarose gel following ethanol concentration.

IDT Exhibit 1020.6398

5' pppG RNAs are resistant to degradation

A T7 control DNA template was used to synthesize a sample comprising a 1.4-kB RNA transcript using either: a) standard rNTPS without a cap; b) a capping rNTP mixture containing rNTPs and m^7GpppG; or c) a non-capping non-pppG 5' end mix containing rNTPs and an rGMP (which can be incorporated at the 5'-end of a transcript during *in vitro* transcription instead of a 5'-triphosphorylated G). Each transcript preparation was treated with xrn-1, and then xrn-1-treated and xrn-1-untreated samples were analyzed by agarose gel electrophoresis. The 5'pppG RNA was resistant to degradation by xrn-1. The RNA transcript that lacked a 5'-triphosphate or a cap structure was degraded by xrn-1. The RNA with rGMP in the transcription reaction was partially sensitive to degradation by xrn-1, presumably because some of the transcripts had a 5'-monophosphate due to incorporation of the GMP during *in vitro* transcription. These results show that xrn-1 can be used to enrich for uncapped but 5'-triphosphorylated RNA (e.g., bacterial mRNA).

xrn-1 degradation of E. coli RNA

E. coli total RNA was treated with xrn-1 exonuclease. Following xrn-1 treatment, a portion of the RNA preparation was concentrated approximately 20X by ethanol precipitation and analyzed by agarose gel electrophoresis. The majority of the RNA, including the *E. coli* 16S and 23S rRNA, was degraded. Following ethanol precipitation to concentrate the RNA remaining in the xrn-1-treated solution, a heterodisperse size range of undegraded RNA was observed on the agarose gel. Specific mRNAs for different genes were detected by RT-PCR of the xrn-1-treated material.

IDT Exhibit 1020.6399

Example 2

Enrichment of Eukaryotic mRNA from Eukaryotic Total RNA for Use in cDNA Synthesis, RNA Amplification and Preparation of Labeled Target RNA for Gene Expression Analysis Using DNA Microarrays.

Sample source and nucleic acid samples

Cells from a mouse stem cell line that had been transformed with a gene linked to an inducible promoter were treated with the inducing substance and then grown in culture for either 6 days or 12 days following induction. Total RNA was then isolated from each 6-day and 12-day culture using methods known in the art, resulting in samples that were designated as the "6-day sample" and the "12-day sample," respectively. Each RNA sample had prominent 18S and 28S rRNA peaks when they were analyzed using an Agilent 2100 bioanalyzer.

xrn-1 treatment of 6-day and 12-day RNA samples

The 6-day and 12-day RNA samples were each divided into two aliquots. The first aliquot was treated with xrn-1 exonuclease and the second aliquot was not treated with xrn-1 exonuclease. No 18S or 28S rRNA peaks are detectable using the Agilent 2100 bioanalyzer after the total RNA is treated with the xrn-1 exonuclease using the same standard protocol.

cDNA synthesis and RNA amplification of xrn-1-treated and xrn-1-untreated 6-day and 12-day RNA samples

Two hundred picograms of total RNA from the xrn-1-untreated aliquots of 6-day and 12-day RNA samples, or a volume of the xrn-1-treated aliquots equivalent to the volume of 200 pg of total RNA in the xrn-1-untreated aliquots, were each used for cDNA and aminoallyl-antisense RNA ("AA-aRNA") synthesis using the TargetAmp™ 2-Round Aminoallyl-aRNA Amplification Kit 1.0 (EPICENTRE) and SuperScript™_ III and SuperScript™ II reverse trancriptases (INVITROGEN) for cDNA synthesis in the first and second rounds of amplification, respectively, according to the instructions provided in the TargetAmp 1.0 kit; the TargetAmp 1.0 kit makes AA-aRNA corresponding to

IDT Exhibit 1020.6400

mRNA in the sample using an Eberwine-type aRNA amplification process (Van Gelder, R.N., et al., Proc. Natl. Acad. Sci. USA, 87: 1663, 1990). Sufficient amounts of AA-aRNA were obtained from amplification reactions for all four types of samples for use in preparing labeled target RNA for hybridization to Affymetrix mouse whole genome GeneChip®s.

Microarray analysis of amplified aRNA from mRNA in xrn-1-treated and xrn-1-untreated 6-day and 12-day RNA samples

Prior to hybridization to an Affymetrix GeneChip, each AA-aRNA was biotinylated using Biotin-X-X-NHS (EPICENTRE) according to the protocol of the supplier. The biotinylated a-RNA was then fragmented, and hybridized to mouse genome GeneChip arrays (MG 430 2.0, Affymetrix) in triplicates. Hybridization, staining, and washing were conducted as specified by Affymetrix. Data were analyzed using GeneSpring software (Agilent Technologies).

Microarray results using labeled target aRNA prepared from xrn-1-treated and xrn-1-untreated 6-day and 12-day RNA samples

The "% present calls" were 49-51% from all GeneChips hybridized to labeled target aRNA prepared from xrn-1-treated or xrn-1-untreated 12-day RNA samples, and ~43-48% for both 6-day samples (excluding one xrn-1-untreated replicate that had a "% present call" of only 29.2%). The concordance of present and absent calls between the xrn-1-treated and xrn-1-untreated 12-day samples was 73.6%. This concordance was close to the average concordance (82.6%) obtained from pairwise comparisons of three replicate slides hybridized with biotinylated aRNA samples obtained from untreated RNA. Thus, the high concordance indicates good preservation of messenger RNA in the xrn-1 treated samples.

IDT Exhibit 1020.6401

Example 3

Preservation of Relative mRNA Abundance Levels after xrn-1 exonuclease Treatment.

Two micrograms of human reference RNA and human skeletal muscle RNA were treated with 1U of xrn-1 exonuclease for 1 hour at 30°C. The RNA was extracted and concentrated by ethanol precipitation and the entire sample was used in a 40μl reverse transcription reaction using MMLV reverse transcriptase (EPICENTRE) with random nonamer primers at 37°C for 1 hour. qPCR was performed using TAQURATE GREEN Real-time PCR MasterMix (EPICENTRE) and optimized concentrations of target-specific primers. The abundance of beta-2-microglobulin (B2M) in the samples was used for expression level normalization. The expression levels of six different target messages were analyzed. Duplicate cDNA and qPCR reactions were preformed, averaged, and normalized for each comparison. Simultaneous analysis was performed with normalization and test primers, and non-template controls were included. The difference in threshold cycles between a target message in one RNA sample (in relation to the normalizer) and that of the other RNA sample (in relation to the normalizer) was determined. The results showed that xrn-1 exonuclease maintained mRNA abundance levels.

All publications and patents mentioned in the above specification are herein incorporated by reference. Various modifications and variations of the described method and system of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with specific preferred embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention that are obvious to those skilled in the relevant fields are intended to be within the scope of the following claims.

IDT Exhibit 1020.6402

CLAIMS

We claim:

1. A method for enriching for an RNA having a 5'-triphosphate or a 5'-cap in a biological sample comprising prokaryotic RNA, eukaryotic RNA or both prokaryotic and eukaryotic RNA and at least one undesired nucleic acid, the method comprising treating the sample with purified 5' exoribonuclease under conditions in which the 5' exoribonuclease is active and for sufficient time so that the undesired nucleic acid is digested and the sample is enriched for RNA having a 5'-triphosphate or a 5'-cap.

2. The method of claim 1, wherein the RNA having a 5'-triphosphate or a 5'-cap is selected from the group consisting of: (i) prokaryotic mRNA; (ii) eukaryotic mRNA, including polyadenylated and non-polyadenylated eukaryotic mRNA; (iii) a mixture of both prokaryotic and eukaryotic mRNA; (iv) eukaryotic snRNA; (v) eukaryotic pre-micro RNA; and (vi) prokaryotic or eukaryotic primary RNA transcripts.

3. The method of claim 1, wherein the RNA having a 5'-triphosphate or a 5'-cap comprises eukaryotic mRNA or both prokaryotic and eukaryotic mRNA, and wherein the method additionally comprises binding the RNA to an oligo(dT) or oligo(dU) resin, membrane or other surface to which oligo(dT) or oligo(dU) is attached and eluting the bound RNA so as to obtain a solution containing polyadenylated eukaryotic mRNA.

4. The method of claim 1 wherein the RNA having a 5'-triphosphate or a 5'-cap comprises prokaryotic mRNA, eukaryotic mRNA or both prokaryotic and eukaryotic mRNA and the undesired nucleic acid comprises prokaryotic rRNA, eukaryotic rRNA or both prokaryotic and eukaryotic rRNA of a size having a svedburg unit greater than about 10S.

IDT Exhibit 1020.6403

5. The method of claim 4 wherein the rRNA is selected from the group consisting of prokaryotic rRNA having a svedburg unit of about 16S or 23S and eukaryotic rRNA having a svedburg unit of about 18S, 26S or 28S.

6. The method of claim 1 wherein the RNA having a 5'-triphosphate or a 5'-cap comprises prokaryotic mRNA, eukaryotic mRNA, or both prokaryotic and eukaryotic mRNA, and wherein the method additionally comprises precipitation of said mRNA having a 5'-triphosphate or a 5'-cap with a solution of LiCl and ethanol at concentrations and under conditions in which tRNA and 5S rRNA are not precipitated.

7. The method of claim 6 wherein the RNA having a 5'-triphosphate or a 5'-cap comprises prokaryotic mRNA, eukaryotic mRNA, or both prokaryotic and eukaryotic mRNA and the undesired nucleic acid comprises prokaryotic rRNA, eukaryotic rRNA or both prokaryotic and eukaryotic rRNA of a size having a svedburg unit greater than about 10S.

8. The method of claim 7 wherein the rRNA is selected from the group consisting of prokaryotic rRNA having a svedburg unit of about 16S or 23S and eukaryotic rRNA having a svedburg unit of about 18S, 26S or 28S.

9. The method of claim 1, wherein the biological sample is treated with a polynucleotide kinase to phosphorylate the RNA having 5'-hydroxyl groups prior to treating the sample with purified 5' exoribonuclease.

10. The method of claim 1, wherein the method additionally comprises: (i) contacting the sample enriched for RNA having a 5'-triphosphate or a 5'-cap with a poly(A) polymerase under conditions so that polyadenylated RNA is obtained; and (ii) contacting the polyadenylated RNA with an oligo(dT)-containing primer and an RNA-dependent DNA polymerase under conditions that cDNA complementary to the RNA is obtained.

IDT Exhibit 1020.6404

11. The method of claim 9, wherein the biological sample is treated with a polynucleotide kinase to phosphorylate the RNA having 5'-hydroxyl groups prior to treating the sample with purified 5' exoribonuclease.

12. The method of claim 10, wherein the oligo(dT)-containing primer comprises a T7-type RNA polymerase promoter selected from the group consisting of T7 RNA polymerase, T3 RNA polymerase, and SP6 RNA polymerase.

13. The method of claim 1, wherein the biological sample comprises eukaryotic RNA, and wherein the method additionally comprises contacting the sample enriched for RNA having a 5'-triphosphate or a 5'-cap with an oligo(dT)-containing primer and an RNA-dependent DNA polymerase under conditions that cDNA complementary to the RNA is obtained.

14. The method of claim 13, wherein the biological sample is treated with a polynucleotide kinase to phosphorylate the RNA having 5'-hydroxyl groups prior to treating the sample with purified 5' exoribonuclease.

15. The method of claim 13, wherein the oligo(dT)-containing primer comprises a T7-type RNA polymerase promoter selected from the group consisting of T7 RNA polymerase, T3 RNA polymerase, and SP6 RNA polymerase.

16. The method of claim 15, wherein the method additionally comprises contacting the cDNA complementary to the RNA with one or more DNA or RNA primers and a DNA polymerase and incubating under conditions so as to obtain double-stranded cDNA having a functional promoter for the T7-type RNA polymerase and synthesizing RNA therefrom.

17. The method of claim 13, wherein the method additionally comprises contacting the cDNA complementary to the RNA with an oligonucleotide sequence tag

IDT Exhibit 1020.6405

template and a DNA polymerase and incubating under conditions so as to obtain cDNA having a sequence tag on it the 3'-end.

18. The method of claim 17, additionally comprising contacting the cDNA having the sequence tag on its 3'-end with a DNA polymerase and an oligonucleotide primer, wherein the either the sequence tag or the oligonucleotide primer, or the combination of both the sequence tag and the oligonucleotide primer encodes the complete sequence of a T7-type RNA polymerase promoter, and incubating under conditions wherein double-stranded cDNA having a functional T7-type RNA polymerase promoter is obtained.

19. The method of claim 18, wherein the method additionally comprises contacting the cDNA with the T7-type RNA polymerase under conditions so as to synthesize RNA using said T7-type RNA polymerase promoter.

20. The method of claim 1, wherein the 5' exoribonuclease is a wild-type or recombinant *Saccharomyces cerevisiae* Xrn1p/5' exoribonuclease 1.

21. A kit for enriching for mRNA having a 5'-triphosphate or a 5'-cap in a biological sample comprising prokaryotic mRNA, eukaryotic mRNA or both prokaryotic and eukaryotic mRNA and at least one undesired nucleic acid, the kit comprising: (i) a solution of purified and stabilized 5' exoribonuclease; and (ii) a concentrated solution of a reaction buffer for providing a 1X reaction buffer in which the 5' exoribonuclease is active.

22. The kit of claim 21, additionally comprising one or more components selected from the group consisting of: (i) a negative control comprising an RNA having a 5'-triphosphate or a 5'-cap; (ii) a positive control comprising an RNA having a 5'-monophosphate; (iii) a solution of LiCl; (iv) a polynucleotide kinase; and (v) a poly(A) polymerase.

IDT Exhibit 1020.6406

23. The kit of claim 21, additionally comprising an RNA-dependent DNA polymerase (reverse transcriptase).

24. The kit of claim 21, additionally comprising one or more components selected from the group consisting of: (i) a poly(A) polymerase; (ii) an oligo(dT)-containing primer; and (iii) an RNA-dependent DNA polymerase.

25. The kit of claim 21, additionally comprising one or more components selected from the group consisting of: (i) a polynucleotide kinase; (ii) an oligo(dT)-containing primer; and (iii) an RNA-dependent DNA polymerase.

26. The kit of claim 21, additionally comprising one or more primers selected from the group consisting of (i) an oligo(dT)-containing primer; and (ii) a primer that is complementary to a specific nucleic acid sequence.

27. The kit of claim 26, wherein said primer additionally comprises a sense or an antisense sequence of a double-stranded promoter for a T7-type RNA polymerase selected from the group consisting of T7 RNA polymerase, T3 RNA polymerase, and SP6 RNA polymerase, and wherein the kit additionally comprises the RNA polymerase that can transcribe RNA using said promoter.

28. The kit of claim 26, additionally comprising: (i) a oligonucleotide sequence tag template; and (ii) a DNA polymerase selected from the group consisting of a DNA-dependent DNA polymerase and an RNA-dependent DNA polymerase that can extend the 3'-end of a nucleic acid that is annealed to the oligonucleotide sequence tag template.

29. The kit of claim 28, additionally comprising: (i) an oligonucleotide that is complementary to the oligonucleotide sequence tag template; and (ii) a T7-type RNA polymerase selected from the group consisting of T7 RNA polymerase, T3 RNA polymerase, and SP6 RNA polymerase; wherein the oligonucleotide sequence tag

IDT Exhibit 1020.6407

template and the oligonucleotide that is complementary to the oligonucleotide sequence tag template together encode the complete sequence of a promoter that can be used for transcription by said T7-type RNA polymerase.

30. The kit of claim 21, additionally comprising one or more components selected from the group consisting of: (i) a negative control comprising an RNA having a 5'-triphosphate or a 5'-cap; (ii) a negative control comprising an RNA or a DNA having a 5'-hydroxyl group; (iii) a positive control comprising an RNA having a 5'-monophosphate; (iv) a solution of LiCl; (v) a ribonuclease H enzyme; and (vi) a polynucleotide kinase.

31. The kit of claim 30, additionally comprising one or more components selected from the group consisting of: (i) a poly(A) polymerase; (ii) an oligo(dT)-containing primer; and (iii) an RNA-dependent DNA polymerase.
(for prokaryotic mRNA)

32. The kit of claim 30, additionally comprising an RNA-dependent DNA polymerase and one or more primers selected from the group consisting of (i) a random primer, such as a random hexamer primer; (ii) an oligo(dT)-containing primer; and (iii) a primer that is complementary to a specific nucleic acid sequence.

33. The kit of claim 30, additionally comprising: (i) an oligonucleotide sequence tag template for adding a tag sequence to the 3'-end of mRNA having a 5'-triphosphate or a 5'-cap; and (ii) a DNA polymerase selected from the group consisting of a DNA-dependent DNA polymerase and an RNA-dependent DNA polymerase that can extend the 3'-end of the mRNA using the oligonucleotide sequence tag template annealed to the mRNA as a template.

34. The kit of claim 21, additionally comprising: (i) an oligonucleotide sequence tag template for adding a tag sequence to the 3'-end of mRNA having a 5'-triphosphate or a 5'-cap; and (ii) a DNA polymerase selected from the group consisting of

IDT Exhibit 1020.6408

a DNA-dependent DNA polymerase and an RNA-dependent DNA polymerase that can extend the 3'-end of the mRNA using the oligonucleotide sequence tag template annealed to the mRNA as a template.

35. The kit of claim 34, wherein the DNA polymerase is an RNA-dependent DNA polymerase, or if a DNA-dependent DNA polymerase is used for adding the tag sequence to the 3'-end of the mRNA, wherein the kit additionally contains an RNA-dependent DNA polymerase that can synthesize cDNA from a primer using the mRNA as a template.

36. The kit of claim 34, wherein said oligonucleotide sequence tag template additionally encodes either a sense or an antisense strand of a double-stranded promoter for a T7-type RNA polymerase selected from the group consisting of T7 RNA polymerase, T3 RNA polymerase, and SP6 RNA polymerase, and wherein the kit additionally comprises the RNA polymerase that can transcribe RNA using said promoter.

37. The kit of claim 21, additionally comprising: (i) a poly(A) polymerase; (ii) an oligo(dT)-containing primer; and (iii) an RNA-dependent DNA polymerase.

38. The kit of claim 37, additionally comprising: (i) an oligonucleotide sequence tag template; (ii) an oligonucleotide that is complementary to the oligonucleotide sequence tag template; wherein the oligonucleotide sequence tag template and the oligonucleotide complementary to the oligonucleotide sequence tag template together encode the complete sequence of a promoter for a T7-type RNA polymerase.

39. The kit of claim 37, additionally comprising a T7-type RNA polymerase selected from the group consisting of a T7 RNA polymerase, a T3 RNA polymerase, and an SP6 RNA polymerase, wherein the T7-type RNA polymerase can use the promoter

IDT Exhibit 1020.6409

encoded by the oligonucleotide sequence tag template and the oligonucleotide complementary to the oligonucleotide sequence tag template.

40. A kit for removing intact prokaryotic or eukaryotic ribosomal RNA (rRNA) of a size having a svedburg unit greater than about 10S from a biological sample, the kit comprising: (i) a solution of purified and stabilized 5' exoribonuclease; (ii) a concentrated solution of a reaction buffer for providing a 1X reaction buffer in which the exoribonuclease is active; and (iii) a positive control rRNA, selected from the group consisting of a 16S and a 23S prokaryotic rRNA, and an 18S, a 26S and a 28S eukaryotic rRNA.

41. The kit of claim 40, wherein the 5' exoribonuclease is a wild-type or recombinant *Saccharomyces cerevisiae* Xrn1p/5' exoribonuclease 1.

IDT Exhibit 1020.6410

Figure 1

SEQ ID NO:1

MGIPKFFRYISERWPMILQLIEGTQIPEFDNLYLDMNSILHNCT
HGNDDDVTKRLTEEEVFAKICTYIDHLFQTIKPKKIFYMAIDGVAPRAKMNQQRARRF
RTAMDAEKALKKAIENGDEIPKGEPFDSNSITPGTEFMAKLTKNLQYFIHDKISNDSK
WREVQIIFSGHEVPGEGEHKIMNFIRHLKSQKDFNQNTRHCIYGLDADLIMLGLSTHG
PHFALLREEVTFGRRNSEKKSLEHQNFYLLHLSLLREYMELEFKEIADEMQFEYNFER
ILDDFILVMFVIGNDFLPNLPDLHLNKGAFPVLLQTFKEALLHTDGYINEHGKINLKR
LGVWLNYLSQFELLNFEKDDIDVEWFNKQLENISLEGERKRQRVGKKLLVKQQKKLIG
SIKPWLMEQLQEKLSPDLPDEEIPTLELPKDLDMKDHLEFLKEFAFDLGLFITHSKSK
GSYSLKMDLDSINPDETEEEFQNRVNSIRKTIKKYQNAIIVEDKEELETEKTIYNERF
ERWKHEYYHDKLKFTTDSEEKVRDLAKDYVEGLQWVLYYYYRGCPSWSWYYPHHYAPR
ISDLAKGLDQDIEFDLSKPFTPFQQLMAVLPERSKNLIPPAFRPLMYDEQSPIHDFYP
AEVQLDKNGKTADWEAVVLISFVDEKRLIEAMQPYLRKLSPEEKTRNQFGKDLIYSFN
PQVDNLYKSPLGGIFSDIEHNHCVEKEYITIPLDSSEIRYGLLPNAKLGAEMLAGFPT
LLSLPFTSSLEYNETMVFQQPSKQQSMVLQITDIYKTNNVTLEDFSKRHLNKVIYTRW
PYLRESKLVSLTDGKTIYEYQESNDKKKFGFITKPAETQDKKLFNSLKNSMLRMYAKQ
KAVKIGPMEAIATVFPVTGLVRDSDGGYIKTFSPTPDYYPLQLVVESVVNEDERYKER
GPIPIEEEFPLNSKVIFLGDYAYGGETTIDGYSSDRRLKITVEKKFLDSEPTIGKERL
QMDHQAVKYYPSYIVSKNMHLHPLFLSKITSKFMITDATGKHINVGIPVKFEARHQKV
LGYARRNPRGWEYSNLTLNLLKEYRQTFPDFFFRLSKVGNDIPVLEDLFPDTSTKDAM
NLLDGIKQWLKYVSSKFIAVSLESDSLTKTSIAAVEDHIMKYAANIEGHERKQLAKVP
REAVLNPRSSFALLRSQKFDLGDRVVYIQDSGKVPIFSKGTVVGYTTLSSSLSIQVLF
DHEIVAGNNFGGRLRTNRGLGLDASFLLNITNRQFIYHSKASKKALEKKKQSNNRNNN
TKTAHKTPSKQQSEEKLRKERAHDLLNFIKKDTNEKNSESVDNKSMGSQKDSKPAKKV
LLKRPAQKSSENVQVDLANFEKAPLDNPTVAGSIFNAVANQYSDGIGSNLNIPTPPHP
MNVVGGPIPGANDVADVGLPYNIPPGFMTHPNGLHPLHPHQMPYPNMNGMSIPPPAPH
GFGQPISFPPPPPMTNVSDQGSRIVVNEKESQDLKKFINGKQHSNGSTIGGETKNSRK
GEIKPSSGTNSTECQSPKSQSNAADRDNKKDEST

SEQ ID NO:2

atgggtattc caaaattttt caggtacatc tcagaaagat ggcccatgat tttacagctt
attgagggaa cacagattcc tgagtttgat aacttatacc tggatatgaa ttcgatttta
cataattgta cgcatggtaa cgacgatgat gtaaccaagc gattaactga agaagaggtt
tttgcaaaaa tctgtacgta tatcgatcac ctttttcaaa caatcaagcc caagaagatt
ttctacatgg ctattgatgg tgtggcccct cgtgccaaga tgaatcaaca aagagctcgt
agattcagaa ccgctatgga tgcagaaaaa gccttgaaga aggctattga gaatggtgac
gagattccta agggtgagcc atttgattcg aattctatta ctccaggtac ggagtttatg
gccaaattga ccaaaaactt acaatatttt attcacgaca agatttctaa cgattccaaa
tggagggaag tgcaaatcat attttctggc catgaagttc caggtgaagg tgaacacaag
atcatgaact ttataaggca tttaaaatcc caaaaggatt tcaaccagaa tacgagacat
tgtatttacg gtcttgacgc agatttgatt atgctgggtt tgtctactca tgggccacat
tttgcgttat tgagagaaga agtgacattt ggtagaagaa atagtgaaaa aaaatcgctt
gaacatcaaa atttctactt attacatctt tctttattaa gagaatacat ggagttggaa
ttcaaagaaa ttgccgatga aatgcaattt gaatacaatt ttgaacgtat tttggatgat
tttattcttg tcatgttcgt cattggtaat gatttcttgc ccaatttgcc agatttgcac
cttaacaaag gagcatttcc cgttttgtta caaacgttca aagaagctct tttacatact
gatggctaca ttaatgaaca tggtaaaata aatttaaaga gattaggtgt ctggttaaat
tatctgtctc aatttgagtt attaaatttc gaaaaggatg atatagacgt tgagtggttc
aacaagcaat tagagaatat ttctttggag ggtgaaagga aaagacagag ggttggtaaa
aaattactgg taaaacaaca gaagaaatta attggaagta taaaaccatg gttgatggag
caattacagg aaaaattatc gcctgattta ccagatgaag aaattccaac tttagagtta
cctaaggact tagacatgaa agatcattta gaatttttaa aagaattcgc ttttgatttg

IDT Exhibit 1020.6411

ggtcttttta taacgcattc caaatccaaa ggtagttatt cgctaaaaat ggatcttgat
tctattaatc ctgatgaaac agaagaagaa tttcaaaatc gtgttaattc tatcaggaaa
acaataaaaa aatatcaaaa tgctatcatc gtggaggaca aagaagaatt ggaaactgaa
aaaacgattt ataatgaaag gtttgaacgt tggaagcatg agtattatca cgacaagtta
aaatttacga cagacagtga agaaaaagtg agagatcttg ctaaagacta cgttgaaggt
ttacaatggg ttctatatta ttattataga ggatgtccat cttggtcgtg gtactatccg
caccattatg caccaagaat ctccgactta gccaagggtt tagatcaaga cattgaattt
gatttgagca aaccatttac tccattccag caactaatgg cagttttacc ggaaaggtcc
aaaaatctga tacctcccgc ctttaggcca ttaatgtacg atgaacagtc gccaatccac
gatttctatc ccgctgaggt tcaacttgat aaaaacggca agacagctga ttgggaagct
gtggttttga tatcgtttgt agatgaaaaa aggttgattg aggctatgca accttatttg
cgcaagttat cacctgaaga aaaaacgaga aatcaatttg gcaaggactt gatatattcc
tttaatcctc aagttgataa cctttataag agtccgttgg gcggcatttt ttctgatatt
gaacacaatc attgtgtcga aaaagagtac atcaccatcc cattggacag ctccgagatt
cggtatggtt tattacctaa tgctaaactc ggtgccgaaa tgctggcggg tttccccacg
ttattgtctt taccatttac tagttcactg gagtacaatg agacaatggt tttccaacaa
ccttctaaac aacaatcaat ggtcttacaa ataactgaca tatacaaaac gaataatgtt
actttggagg acttttccaa gaggcattta aacaaagtga tttatacaag atggccatat
ttaagagaat ccaaattggt ctctttaacg gatggtaaga ctatctatga atatcaggag
tccaatgata agaaaaagtt cggattcata acgaagcctg cggaaaccca ggacaaaaaa
cttttcaata gtttgaagaa ttcaatgcta aggatgtatg ctaaacagaa agctgttaaa
ataggaccta tggaagccat tgctaccgtc tttccagtga ctggtttggt gagagactct
gatggtggtt atattaagac ctttagccct accccagatt actatccatt gcaactggtt
gttgaatctg ttgtcaacga ggatgaaaga tataaagaaa gaggacccat tcctattgaa
gaggaatttc cattgaattc aaaagttatt ttcttaggtg attatgccta tggtggtgaa
actactattg acggttacag cagtgaccgc agactaaaaa ttactgtaga aaagaagttt
ttggatagtg agcccaccat cggcaaagaa aggttacaaa tggatcatca agccgttaaa
tattatccgt cttatattgt gtccaagaac atgcacttac accccttgtt tttgtctaag
attacttcca agttcatgat tactgacgct actgggaagc atatcaatgt tggtatcccg
gttaagttcg aagctagaca ccaaaaggtt ttaggttacg cgaggaggaa ccctaggggc
tgggaatact caaatttaac tctaaattta ctgaaagagt atagacaaac tttcccagat
ttttttttca ggttgtccaa ggtaggtaat gatatcccag ttttggaaga tcttttcccc
gatacttcca ctaaggatgc catgaattta ttagatggta tcaaacaatg gctaaagtat
gtctcatcga agtttatcgc ggtatctttg gagtctgact ccttaactaa gacatcgatt
gctgccgtgg aagatcatat catgaaatac gcagctaaca tcgaaggtca tgaaagaaaa
cagttagcca aagttcctcg tgaggctgtt ttgaatccaa gatcatcatt tgcactctta
cgtagtcaaa agttcgattt gggtgaccgt gttgtttata tccaagattc tggtaaggta
ccaatttttct caaagggtac agttgttggc tatactactc tcagttcatc attatcaatt
caggtcttat ttgatcatga aatcgtggcc ggtaataatt ttggcggcag gttgcgcacg
aatagaggct tagggcttga tgcctctttc ttattgaata ttactaacag gcagttcatt
tatcactcca aggcttccaa aaaggctttg gaaaagaaaa agcaatctaa caataggaac
aataatacca aaactgctca caagactcct tcaaagcaac aatctgaaga aaaactgaga
aaagaaaggg cacatgattt attgaatttt atcaaaaagg ataccaatga aaagaattct
gaaagtgtag acaacaagag catgggatcg caaaaagatt ccaaacccgc aaagaaagtt
ttgttaaaaa gaccagctca gaaaagcagt gaaaacgtgc aagttgattt ggccaatttt
gaaaaagcac cgcttgataa tccaactgtt gctggatcta ttttcaatgc cgttgcaaat
caatattctg atggtatagg cagtaatttg aatatcccaa ctccacctca cccaatgaat
gtggttgggg gtcctattcc tggagcgaat gatgttgcag atgttggttt gccgtacaat
attcccccag gttttatgac gcatcctaat ggtcttcacc cattacaccc tcaccagatg
ccttacccta atatgaatgg aatgtctatt ccgccaccag caccacatgg gtttggacaa
ccgatttcct tcccacctcc acctcctatg acaaatgttt cagatcaagg aagtcgtatt
gttgtcaatg aaaaggaaag ccaagatttg aaaaaattca ttaatggtaa acagcacagc
aatggttcaa ctattggggg agaaacaaag aacagtagga aaggcgagat taaaccttct
tctggcacaa actctactga atgtcaatcg ccaaagtcac aaagcaatgc tgctgaccgt
gataataaaa aagacgaatc

IDT Exhibit 1020.6412

(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 1 929 039 B2

(12) **NEW EUROPEAN PATENT SPECIFICATION**
After opposition procedure

(45) Date of publication and mention of the opposition decision:
20.11.2013 Bulletin 2013/47

(45) Mention of the grant of the patent:
30.12.2009 Bulletin 2009/53

(21) Application number: **06783930.8**

(22) Date of filing: **21.09.2006**

(51) Int Cl.:
C12Q 1/68 *(2006.01)*

(86) International application number:
PCT/NL2006/000467

(87) International publication number:
WO 2007/037678 (05.04.2007 Gazette 2007/14)

(54) **HIGH THROUGHPUT SCREENING OF MUTAGENIZED POPULATIONS**

SCREENING MUTAGENISIERTER POPULATIONEN MIT HOHEM DURCHSATZ

CRIBLAGE A HAUT DEBIT DE POPULATIONS MUTAGENISEES

(84) Designated Contracting States:
AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HU IE IS IT LI LT LU LV MC NL PL PT RO SE SI SK TR

(30) Priority: **29.09.2005 US 721528 P**

(43) Date of publication of application:
11.06.2008 Bulletin 2008/24

(73) Proprietor: **KEYGENE N.V.**
6708 PW Wageningen (NL)

(72) Inventors:
- **VAN EIJK, Michael, Josephus, Theresia**
  NL-5373 EJ Herpen (NL)
- **VAN TUNEN, Adrianus, Johannes**
  NL-6708 NA Wageningen (NL)

(74) Representative: **de Lang, Robbert-Jan**
Exter Polak & Charlouis B.V. (EP&C)
J.J. Viottastraat 31
1071 JP Amsterdam (NL)

(56) References cited:
US-A1- 2005 130 173

- **WIENHOLDS ERNO ET AL: "Target-selected gene inactivation in zebrafish." METHODS IN CELL BIOLOGY 2004, vol. 77, 2004, pages 69-90, XP002420246 ISSN: 0091-679X**
- **WIENHOLDS ERNO ET AL: "Efficient target-selected mutagenesis in zebrafish." GENOME RESEARCH DEC 2003, vol. 13, no. 12, December 2003 (2003-12), pages 2700-2707, XP002420247 ISSN: 1088-9051**
- **ALTSHULER D ET AL: "An SNP map of the human genome generated by reduced representation shotgun sequencing" NATURE, NATURE PUBLISHING GROUP, LONDON, GB, vol. 407, 28 September 2000 (2000-09-28), pages 513-516, XP002955779 ISSN: 0028-0836**
- **VANDENBUSSCHE MICHIEL ET AL: "Toward the analysis of the petunia MADS box gene family by reverse and forward transposon insertion mutagenesis approaches: B, C, and D floral organ identity functions require SEPALLATA-like MADS box genes in petunia." THE PLANT CELL NOV 2003, vol. 15, no. 11, November 2003 (2003-11), pages 2680-2693, XP002420664 ISSN: 1040-4651**
- **TILL BRADLEY J ET AL: "Large-scale discovery of induced point mutations with high-throughput TILLING." GENOME RESEARCH MAR 2003, vol. 13, no. 3, March 2003 (2003-03), pages 524-530, XP002420665 ISSN: 1088-9051 cited in the application**
- **GUPTA P K ET AL: "Single nucleotide polymorphisms: A new paradigm for molecular marker technology and DNA polymorphism detection with emphasis on their use in plants" CURRENT SCIENCE, vol. 80, no. 4, 25 February 2001 (2001-02-25), pages 524-535, XP002972388 ISSN: 0011-3891**
- **WOLFORD ET AL: 'High-throughput SNP detection by using DNA pooling and denaturing high performance liquid chromatography (DHPLC)' HUM GENET vol. 107, 18 October 2000, SPRINGER-VERLAG, pages 483 - 487**

EP 1 929 039 B2

Printed by Jouve, 75001 PARIS (FR)

(Cont. next page)

IDT Exhibit 1020.6413

- QIU ET AL: 'DNA Sequence-Based "Bar Codes" for Tracking the Origins of Expressed Sequence Tags from a Maize cDNA Library Constructed Using Multiple mRNA Sources' PLANT PHYSIOLOGY, OCT. 2003 vol. 133, AMERICAN SOCIETY OF PLANT BIOLOGISTS, pages 475 - 481
- COLBERT ET AL: 'High-Throughput Screening for Induced Point Mutations' PLANT PHYSIOLOGY, JUNE 2001 vol. 126, AMERICAN SOCIETY OF PLANT PHYSIOLOGISTS, pages 480 - 484

IDT Exhibit 1020.6414

Description

BACKGROUND OF THE INVENTION

Field of the Invention

[0001] The present invention, in the fields of molecular biology and genetics relates to improved strategies for identifying mutations in populations, based on the use of high throughput sequencing technologies. The invention further provides for kits that can be applied in the methods.

Description of the Background Art

[0002] Populations carrying mutations, either induced or naturally occurring are used in modem genomics research to identify genes affecting traits of importance by reverse genetics approaches. This is in particular applicable for plants and crops of agronomic importance, but such populations are also useful, for other organisms such as yeast, bacteria etc. Other organisms, such as animals, birds, mammals etc can also be used, but these populations are typically more cumbersome to obtain or to control. Nevertheless, it is observed that the invention described herein is of a very general nature, and can be applied also to such organisms. Altschuler et al. (Nature 2000, 407, 513-516) discloses a method for analyzing SNPs in which different samples are mixed and analyzed.

[0003] Mutagenized populations represent complementary tools for gene discovery, as such populations are commonly used to screen known genes for loss-of-function mutations or assessing phenotype changes in organisms with the mutated gene. The rate-limiting step is the screening work associated with identification of, respectively, organisms carrying a mutation in the gene of interest, Below, the principles of such populations and the screening methods are described in more detail and more efficient screening methods are presented which increase the value of these tools for gene-discovery.

[0004] A technology that uses mutagenized populations is known as TILLING (Targeted Induced Local Lesions In Genomes) (McCallum et al., Nat. Biotechnol 2000, 18, 455-457, McCallum et al., Plant Physiology, 2000,123, 439-442: Till et al. Genome Research 2003, 13, 524-530 Wienholds et al, Methods in Cell Biology 2004, 70, 69-90) relies on random introduction of large numbers of mutations (mostly nucleotide substitutions) into the genome by treatment with ethyl methane sulfonate (EMS) or by ionizing radiation (fast neutron bombardment,) (Li et al. The Plant Journal, 2001, 27, 235-42). Every plant in the population carries several hundred (or thousand) mutations, some of which affect normal development, morphology or otherwise confer a phenotype due to loss-of-function (knock-out, knock-down) of one or multiple genes or their regulatory sequences. A TILLING population generally contains a sufficient number of plants to cover all genes with multiple independent mutations (5-20 per gene). A mutagenized plant population used in TILLING therefore usually consist of 3000-10,000 plants and can be used in two ways:

Reverse Genetics

[0005] "Reverse Genetics" is the most common way of using TILLING populations. A gene of interest is identified, *e.g.*, by transcript profiling or a candidate gene approach, and the question to be answered is whether this gene affects a particular phenotypic trait of interest. The challenge therefore is to identify one (or several) plants with loss-of-function mutations in this gene. This is commonly performed in a multi-step screening process, typically comprising the following steps:

1. Genomic DNA of a large number of (pooled) M2 plants (*e.g.*, 3072) of the TILLING population is isolated.
2. Pools of equal amounts of DNA from 8 to 32 plants per pool are assembled, with the pooling level depending on the sensitivity of the CEL I screening system (see below). This results in a total of 96- to 384 pooled DNA samples in case of 3072 plants.
3. Labeled PCR primers are used to amplify parts of the gene from all pooled DNAs. Overlapping PCR fragments are used to cover the entire gene (*e.g.*, 3 * 600 bp PCR fragments are amplified from a 1500 bp gene).
4. Heteroduplexes of the PCR products obtained from the pooled DNA samples are prepared and incubated with CEL I or another enzyme which recognizes and cuts single nucleotide sequence mismatches (*e.g.*, mung bean nuclease, S1 nuclease, Surveyor *etc.*) and the treated samples are resolved on a denaturing (sequencing) gel or by capillary electrophoresis.
5. Pools containing a plant carrying a mutation in the gene are identified by observing bands of digestion products resulting from CEL I treatment.

[0006] To identify the plant carrying the mutation, PCRs are repeated on individual DNAs of the plants in the positive

IDT Exhibit 1020.6415

pools, followed by bi-directional Sanger sequencing.

[0007] Plants harboring a mutation are grown and out-crossed to wild-type to establish causal relationship between the mutation and the observed phenotype change.

[0008] The advantage of CEL I screening (steps 3-5 above) is that pre-screening the pooled samples saves costs over sequencing all plants individually by Sanger sequencing.

[0009] However, a limitation of CEL I screening is that not all identified mutations affect gene function (*e.g.*, silent substitutions) and this is not known until the PCR products of individual plants in a positive pool are sequenced. Nevertheless, the CEL I mediated screening method is cost-saving compared to sequencing PCR products of all plants separately.

[0010] Another limitation is that CEL I screening involves running gels and scoring, a relatively cumbersome process that requires confirmation of mutations from the second strand as gel-patterns are not always clear-cut.

[0011] A third disadvantage is that CEL I screening is relatively insensitive to mutation detection at the termini of the PCR product which may lead to some mutations going undetected. Further disadvantages of CEL I are that it has been found that the enzyme is extremely sensitive to reaction conditions such as salt concentrations. This makes that the enzyme can only be used in a limited number of buffers, thereby hampering the broad use of CEL I. Another practical disadvantage associated with the application of CEL I is that the enzyme is not reliable in cutting all mismatched heteroduplexes.

[0012] Finally, CEL I screening is incapable of distinguishing missense mutations (which are the most prevalent) from non-sense mutations, causing a great deal of screening work carried out on positive pools without yielding interesting mutations.

Forward Genetics

[0013] Plants of the mutagenized population are grown and phenotyped for traits of interest. Plants with an interesting phenotype are then crossed to a wild-type plant to out-cross mutations that are not linked to the phenotype of interest. Finally, the mutated gene responsible for the phenotype of interest is identified by positional cloning (using genetic markers), analogous to mapping QTL in conventional genetic mapping populations (F2, RIL etc). Although theoretically possible, mutagenized populations are not commonly used this way.

[0014] The present invention was made in part improve the existing strategies for screening of mutagenized populations. It is an object of the invention to provide efficient methods for screening large populations for the presence of mutations and to improve efficient assessment of the mutations for impact on gene function, *i.e.*, to reduce the amount of effort expended on screening mutations that do not lead to altered gene functions. The present methods were designed to avoid the use of the CEL I enzyme or its equivalents.

SUMMARY OF THE INVENTION

[0015] The present inventors found that using high throughput sequencing strategies, the above-mentioned goals were achieved and mutagenized populations, such as TILLING populations, populations wherein mutations have been introduced using (synthetic) mutagenic or DNA damaging oligonucleotides or, i.e. by Targeted Nucleotide Exchange (TNE) or by Region Targeted Mutagenesis (RTM), or populations that contain naturally occurring mutations such as Single nucleotide polymorphisms (SNPs), small insertions and deletions, and variations in microsatellite repeat number could be efficiently screened for the presence of mutations of interest.

Definitions

[0016] In the following description and examples, a number of terms are used. To provide a clear and consistent understanding of the specification and claims, including the scope to be given such terms, the following definitions are provided. Unless otherwise defined herein, all technical and scientific terms used have the same meaning as commonly understood by those of ordinary skill in the art to which this invention belongs. The disclosures of all publications, patent applications, patents and other references are incorporated by reference herein in their entirety.

[0017] "TILLING" or "Targeting induced local lesions in genomes" is a general reverse genetic strategy providing an allelic series of induced (point) mutations by random chemical or physical mutagenesis in combination with PCR-based screening to identify point mutations in a region of interest. In TILLING screening, regions of interest are amplified by PCR. Heteroduplexes between wild-type fragments and fragments harboring an induced mutation are formed by denaturing and reannealing PCR products. These heteroduplexes are cleaved by CEL I and cleaved products are resolved. Throughput can be increased by pooling. Following discovery of PCR products harboring sequence differences in a pool, PCR products included in the pool are commonly screened again by Sanger sequencing of individual PCR products, thereby identifying the mutant plant and the exact sequence difference in the mutated gene.

IDT Exhibit 1020.6416

[0018] "Mutagenized Population" refers to a population of organisms (usually plants, but other organisms, including animals such as Drosophila and mice may be used to create a mutagenized populations; Schimenti et al., 1998, Genome Research 8:698-710) that have been subjected to mutagenesis (chemical or physical) to yield a library of mutants. TILLING populations may vary widely in size, and for certain purposes, partial TILLING populations can be used that contain 90, 80 70, 60, 50, 40 30 or even only 20% of the original population. As an alternative to mutagenized populations, populations can be used wherein the population is not mutagenized but comprises sub-populations that contain naturally occurring mutations such as Single nucleotide polymorphisms (SNPs), small insertions and deletions, and variations in microsatellite repeat number. These populations are particularly advantageous when mutagenized populations are not readily accessible (humans) or where already large germplasms are available. See for instance Comai et al., The Plant Journal, 2004, 37, 778-786. Such a population can be used in combination with a 'reference DNA'.

[0019] "Targeted Nucleotide Exchange" or "TNE". Targeted nucleotide exchange (TNE) is a process by which a synthetic oligonucleotide, partially complementary to a site in a chromosomal or an episomal gene directs the reversal of a single nucleotide at a specific site. TNE has been described using a wide variety of oligonucleotides and targets. Some of the reported oligonucleotides are RNA/DNA chimeras, contain terminal modifications to impart nuclease resistance.

[0020] "Region targeted mutagenesis" or "RTM". Region targeted mutagenesis is a process by which double-strand breaks at a predefined target site in the genomic DNA are artificially created, resulting in repair of the break by one of various available cellular repair mechanisms, mostly leading to mutations at the site of the break. Double-strand breaks may be created by introduction into the cell nucleus of zinc-finger nucleases (e.g. see Lloyd *et al.*, 2005), meganucleases such as I-Sce1 (Epinat *et al.,* 2003), or triplex-forming oligonucleotides coupled to mutagcnic chemical groups (Havre *et al.*, 1993).

[0021] "Nucleic acid": A nucleic acid, as used herein, may include any polymer or oligomer of nucleotides with pyrimidine and purine bases, preferably cytosine, thymine (or uracil), adenine and guanine, respectively (*See* Lehninger, Principles of Biochemistry, at 793-800 (Worth Pub. 1982)). Any deoxyribonucleotide, ribonucleotide or peptide nucleic acid component, and any chemical variant thereof, such as those with methylated, hydroxymethylated or glycosylated forms of these bases, and the like, are included. The polymers or oligomers may be heterogeneous or homogeneous in composition, and may be isolated from naturally occurring sources or may be artificially or synthetically produced. A nucleic acid may be DNA or RNA, or a mixture thereof, and may exist permanently or transiently in single-stranded or double-stranded form, including homoduplexes, heteroduplexes, and hybrid states.

[0022] "Tagging" refers to the addition of a tag or label to a nucleic acid in order to be able to distinguish it from a second or further nucleic acid. Tagging can be performed, for example, by the addition of a sequence identifier during amplification by using tagged primers or by any other means known in the art. Such a sequence identifier can be a unique base sequence of varying but defined length uniquely used for identifying a specific nucleic acid sample. Typical example are ZIP sequences. Using such a tag, the origin of a sample can be determined upon further processing. In the case of combining processed products originating from different nucleic acid samples, the different nucleic acid samples are generally identified using different tags.

[0023] "Tagged library" refers to a library of tagged nucleic acids.

[0024] "Sequencing" refers to determining the order of nucleotides (base sequences) in a nucleic acid sample, *e.g.*, DNA or RNA.

[0025] "Aligning and alignment" mean the comparison of two or more nucleotide sequences based on the presence of short or long stretches of identical or similar nucleotides. Several methods for alignment of nucleotide sequences are known in the art, as will be further explained below. Sometimes the terms "assembly" or "clustering" are used as synonyms.

[0026] "High-throughput screening" (HTS) is a method of scientific experimentation especially relevant to the fields of biology and chemistry. Through a combination of modem robotics and other specialized laboratory hardware, HTS allows an investigator to effectively screen large numbers of samples simultaneously (or virtually simultaneously).

[0027] "Primers" in general refers to DNA strands which can prime the synthesis of DNA. DNA polymerase cannot synthesize DNA *de novo* without primers: it can only extend an existing DNA strand in a reaction in which the complementary strand is used as a template to direct the order of nucleotides to be assembled. The synthetic oligonucleotide molecules which are used in a polymerase chain reaction (PCR) are referred to herein as primers.

[0028] "Primers with increased affinity" are primers with modified nucleotides such as PNA or LNA, which increases their thermal stability and allows for allele-specific amplification based on single nucleotide sequence differences. In order to achieve this, one or several modified nucleotides are often included, preferably at the 3'-end of the primer.

[0029] "DNA amplification" is typically used to denote the *in vitro* synthesis of double-stranded DNA molecules using PCR. It is noted that other amplification methods exist and they may also be used in the present invention.

BRIEF DESCRIPTION OF THE DRAWINGS

[0030]

IDT Exhibit 1020.6417

Fig 1: Schematic representation of clustered sequences resulting from shotgun sequencing a gene to identify EMS-induced mutations. Mutations are lighter, sequence errors darker colored. Sequence errors are expected to be observed randomly and most often just once.

Fig 2: Schematic representation of clustered tagged sequencing resulting from a 100 bp gene region amplified with 4 bp -tagged PCR primers from a 3-D pooled library. Mutations are lighter, sequence errors darker colored. Plant IDs are known for mutations identified by 3 tags (1,2,3) and (4,5,6) but not for those identified by less than 2 tag (7,8). Sequence errors are expected to be observed randomly and just once.

Fig 3: Illustration of the system of long and short PCR primers to use in tagging the sequences.

[0031] **Fig. 4.** Agarose gel estimation of the PCR amplification yield of eIF4E exon 1 amplification for each of the 28 3D pools.

DETAILED DESCRIPTION OF THE INVENTION

[0032] In one aspect the invention is directed to a method for the detection of a mutation in a target sequence in a member of a mutagenized population comprising the steps of:

(a) Isolating genomic DNA of each member of the mutagenized population to provide for DNA samples of each member in the population;
(b) pooling the DNA obtained in step (a);
(c) amplifying the target sequence with a pair of (optionally labeled) primers from the DNA pools;
(d) pooling the amplification products of step (c) to create a library of amplification products;
(e) optionally, fragmenting the amplification products in the library;
(f) determining the nucleotide sequence of the products and/or fragments using high throughput sequencing;
(g) identifying mutations by clustering (aligning) the sequences of the fragments;
(h) screening the identified mutations for a modified function of the target sequence;
(i) designing a primer directed to hybridize to the identified mutation;
(j) amplifying the DNA pools at step (b) with the primer of step (i) and one of the primers of step (c);
(k) identifying the member(s) carrying the mutation;
(l) optionally, confirming the mutation by amplifying the target sequence from the member(s) of step (k) using the primers of step (c) and determining the sequence of the amplified product.

[0033] The isolation of DNA is generally achieved using common methods in the art such as the collection of tissue from a member of the population, DNA extraction (for instance using the Q-Biogene fast DNA kit), quantification and normalization to obtain equal amounts of DNA per sample. As an example, the present invention is illustrated based on a TILLING population of 3072 plants and a gene of 1500bp.

[0034] The pooling of the isolated DNA can for instance be achieved using a 3-dimensional pooling scheme (Vandenbussche et al., 2003, The Plant Cell, 15: 2680-93). The pooling is achieved preferably using equal amounts of DNA. The 3D-pooling scheme may comprise 15x15x14, resulting in 44 pools (15+15+14) containing 3072/14=219 or 3072/15 =205 different DNA samples per pool. Other pooling schemes can be used.

[0035] The pooling step typically serves to identify the plant containing an observed mutation after one round of PCR screening. Pooling of the DNA further serves to normalize the DNAs prior to PCR amplification to provide for a more equal representation in the libraries for sequencing. The additional advantage of the pooling of the DNA is that not all sequences have to be determined separately, but that the pools allow for rapid identification of the sequences of interest, in particular when tagged libraries are used. This facilitates the screening of large or complex populations in particular.

[0036] The amplification of the target sequence with a pair of optionally labeled primers from the pools can be achieved by using a set of primers that have been designed to amplify the gene of interest. As stated, the primers may be labeled to visualize the amplification product of the gene of interest.

[0037] The amplification products are pooled, preferably in equal or normalized amounts to thereby create a library of amplification products. Exemplary, the complexity of the library will be 3072 plants X 1500 bp gene sequence = 4.6 Mb sequence.

[0038] The amplification products in the library may be randomly fragmented prior to sequencing of the fragments in case the PCR product length exceeds the average length of the sequence traces. Fragmentation can be achieved by physical techniques, i.e., shearing, sonication or other random fragmentation methods. In step (f), at least part, but preferably the entire, nucleotides sequence of at least part of, but preferably of all the fragments contained in the libraries is determined. In certain embodiments, the fragmentation step is optional. For instance, when the read length of the sequencing technique and the PCR fragments length are about the same, there is no need for fragmentation. Also in the case of larger PCR products this may not be necessary if it is acceptable that only part of the PCR product is

IDT Exhibit 1020.6418

sequenced for instance in case of 1500 bp PCR product and read length of 400 (from each side) 700 bp remain unsequenced.

[0039] The sequencing may in principle be conducted by any means known in the art, such as the dideoxy chain termination method (Sanger sequencing), but this is less preferred given the large number of sequences that have to be determined. It is however preferred and more advantageous that the sequencing is performed using high-throughput sequencing methods, such as the methods disclosed in WO 03/004690, WO 03/054142, WO 2004/069849, WO 2004/070005, WO 2004/070007, and WO 2005/003375 (all in the name of 454 Life Sciences), by Seo et al. (2004) Proc. Natl. Acad. Sci. USA 101:5488-93, and technologies of Helios, Solexa, US Genomics, etcetera, which are herein incorporated by reference. It is most preferred that sequencing is performed using the apparatus and/or method disclosed in WO 03/004690, WO 03/054142, WO 2004/069849, WO 2004/070005, WO 2004/070007, and WO 2005/003375 (all in the name of 454 Life Sciences). The technology described allows sequencing of 40 million bases in a single run and is 100 times faster and cheaper than competing technology. The sequencing technology roughly consists of 5 steps: 1) fragmentation of DNA and ligation of specific adaptor to create a library of single-stranded DNA (ssDNA); 2) annealing of ssDNA to beads, emulsification of the beads in water-in-oil microreactors and performing emulsion PCR to amplify the individual ssDNA molecules on beads; 3) selection of /enrichment for beads containing amplified ssDNA molecules on their surface 4) deposition of DNA carrying beads in a PicoTiterPlate®; and 5) simultaneous sequencing in at least 100,000 wells by generation of a pyrophosphate light signal. The method will be explained in more detail below.

[0040] In a preferred embodiment, the sequencing comprises the steps of:

(a) annealing adapted fragments to beads, with a single adapted fragment being annealed to each bead;
(b) emulsifying the beads in water-in-oil microreactors, each water-in-oil microreactor comprising a single bead;
(c) loading the beads in wells, each well comprising a single bead; and generating a pyrophosphate signal.

[0041] In the first step (a), sequencing adaptors are ligated to fragments within the library. The sequencing adaptor includes at least a "key" region for annealing to a bead, a sequencing primer region and a PCR primer region. Thus, adapted fragments are obtained.

[0042] In a second step, adapted fragments are annealed to beads, each bead annealing with a single adapted fragment. To the pool of adapted fragments, beads are added in excess as to ensure annealing of one single adapted fragment per bead for the majority of the beads (Poisson distribution).

[0043] In a next step, the beads are emulsified in water-in-oil microreactors, each water-in-oil microreactor comprising a single bead. PCR reagents are present in the water-in-oil microreactors allowing a PCR reaction to take place within the microreactors. Subsequently, the microreactors are broken, and the beads comprising DNA (DNA positive beads) are enriched.

[0044] In a following step, the beads are loaded in wells, each well comprising a single bead. The wells are preferably part of a PicoTiter™Plate allowing for simultaneous sequencing of a large amount of fragments.

[0045] After addition of enzyme-carrying beads, the sequence of the fragments is determined using pyrosequencing. In successive steps, the PicoTiter™Plate and the beads as well as the enzyme beads therein are subj ected to different deoxyribonucleotides in the presence of conventional sequencing reagents, and upon incorporation of a deoxyribonucleotide a light signal is generated which is recorded. Incorporation of the correct nucleotide will generate a pyrosequencing signal which can be detected.

[0046] Pyrosequencing itself is known in the art and described inter alia on www.biotagebio.com; www.pyrosequencing.com / section technology. The technology is further applied in e.g., WO 03/004690, WO 03/054142, WO 2004/069849, WO 2004/070005, WO 2004/070007, and WO 2005/003375 (all in the name of 454 Life Sciences).

[0047] The mutations are identified by clustering of the sequenced fragments in the amplified library. Identification of the mutations is achieved by aligning the determined sequences of the fragments of the libraries. The majority of the sequences are wild-type (not mutated) but the induced mutations and occasional sequencing errors are also observed. As the amplification libraries are sequenced with multifold redundancy (typically about 4- to 5-fold redundant), multiple observations of the same sequence change is indicative of a mutation rather than a sequencing error. See Fig. 1.

[0048] The clustering provides alignments of the fragments in the amplified library. In this way for each PCR product in the library, a cluster is generated from sequenced fragments, *i.e.*, a contig of the fragments, is build up from the alignment of the sequence of the various fragments obtained from the fragmenting in step (e).

[0049] Methods of alignment of sequences for comparison purposes are well known in the art. Various programs and alignment algorithms are described in: Smith and Waterman (1981) Adv. Appl. Math. 2:482; Needleman and Wunsch (1970) J. Mol. Biol. 48:443; Pearson and Lipman (1988) Proc. Natl. Acad. Sci. USA 85:2444; Higgins and Sharp (1988) Gene 73:237-244; Higgins and Sharp (1989) CABIOS 5:151-153; Corpet et al. (1988) Nucl. Acids Res. 16:10881-90; Huang et al. (1992) Computer Appl. in the Biosci. 8:155-65; and Pearson et al. (1994) Meth. Mol. Biol. 24:307-31. Altschul et al. (1994) Nature Genet. 6:119-29 present a detailed consideration of sequence alignment methods and homology calculations.

IDT Exhibit 1020.6419

[0050] The NCBI Basic Local Alignment Search Tool (BLAST) (Altschul *et al.,* 1990) is available from several sources, including the National Center for Biological Information (NCBI, Bethesda, Md.) and on the Internet, for use in connection with the sequence analysis programs blastp, blastn, blastx, tblastn and tblastx. It can be accessed at <http://www.ncbi.nlm.nih.govBLAST/>. A description of how to determine sequence identity using this program is available at <http://www.ncbi.n1m.nih.gov/BLASTlblast_help.html>.

[0051] In the analysis of mutagenized populations, after the mutations have been identified, the identified mutations are assessed for a modified function of the associated gene, for instance the introduction of a stop codon. This assessment is performed on the sequence itself, for example by six-frame translation. Once the interesting mutations have been identified, the mutations are further investigated to identify the associated member of the population.

[0052] For each mutation that has been classified as an interesting mutation, an allele specific primer is designed that targets the mutation of interest. Thus, the allele specific primer is then used in combination with one of the primers used in the amplification of the pooled DNA samples (either the reverse or the forward primer). One or both of the primers may be labeled. The set of primers is used to amplify the pools of DNA. The positive pools are identified and the mutant plant is identified. In the above-mentioned 3D pooling scheme, the allele specific PCR with the set of primers to screen the 3D pooled DNA sample plates results in the identification of 3 positive pools (one in each dimension), which specifies the library address of the mutant plant.

[0053] In certain embodiments; the allele-specific primers comprise alternative nucleotides such as Locked Nucleic Acids (LNA) or Peptide Nucleic Acids (PNA) to increase their specificity. Such nucleic acids are widely known in the art and are commercially available from a choice of suppliers.

[0054] Confirmation of the mutation is achieved by amplification of the target sequence from the identified mutant plant. This amplification is performed with the primers from step (c). The nucleotide sequence of the amplified product is determined and by comparison with the consensus sequence, the mutation is identified. The sequencing is preferably performed Sanger sequencing.

[0055] In one aspect the invention pertains to a method for the detection of a mutation in a target sequence in a member of a mutagenized population comprising the steps of:

(a) isolating genomic DNA of each member of the mutagenized population to provide DNA samples of each member in the population;
(b) pooling the DNA obtained in step (a);
(c) amplifying a part or segment of the target sequence with a pair of tagged (optionally labeled) primers from the DNA pools, preferably wherein at least one of the primers comprise a gene-specific section, a tag and a sequence primer binding site;
(d) pooling the amplification products of step (c) to create a library of amplification products;
(d) determining the nucleotide sequence of the amplification products using high throughput sequencing;
(f) identifying mutations by clustering (aligning) the sequences of the fragments;
(g) identifying the member(s) having the mutation using the tags;
(h) optionally, confirming the mutation by amplifying the target sequence from the member(s) of step (g) using the primers of step (c) and determining the sequence of the amplified product.

The isolation of genomic DNA of the members of the mutagenized population and the pooling of the isolated DNA can be carried out essentially as described above.

A part or segment of the target sequence is amplified using a pair of tagged primers that may be labeled. Preferably, for each pool of each dimension, a different primer is used. In the above illustration this means that 44 forward and 44 reverse primers are preferred. Preferably, each of the forward and reverse primers comprises

(i) a sequence primer binding site that can be used in the following sequencing step,
(ii) a tag that serves to link the primer (and the resulting amplification product) to the original member of the population, and
(iii) a gene specific sequence that is capable of annealing to the target sequence of interest (*i.e.*, the gene).

[0056] In a typical embodiment the primer has the following order:

5'-Sequence primer binding site---Tag---Gene specific PCR primer sequence-3'

The length of the sequence primer binding site and the gene specific PCR primer sequence are those that are conventional in common PCR use, *i.e.*, independently from about 10 to about 30bp with a preference for from 15 to 25bp. Preferably the part or segment of the sequence that is amplified corresponds to a length that can be sequenced in one run using the high throughput sequencing technologies described below. In certain embodiments the part or segment has a length of between about 50 bp to about 500 bp, preferably from about 75 bp to about 300 bp and more preferably between about 90bp and about 250 bp. As stated above, this length may vary with the sequencing technology employed including those yet to be developed.

IDT Exhibit 1020.6420

[0057] By using primers (forward and/or reverse) containing a tag sequence that is unique for each of the primers representing all pool dimensions, the specific plant origin of each tag sequence is known as the sequence primer anneals upstream of the tag and as a consequence, the tag sequence is present in each amplification product. In certain embodiments, both forward and reverse primers are tagged. In other embodiments, only on of the forward or reverse primers is tagged. The choice between one or two tags depends on the circumstances and depends on the read length of the high throughput sequencing reaction and/ or the necessity of independent validation. In the case of, *e.g.*, a 100bp PCR product that is sequenced unidirectionally, only one tag is needed. In the case of a 200bp PCR product and a 100bp read-length, double tagging is useful in combination with bi-directional sequencing as it improves efficiency 2-fold. It further provides the possibility of independent validation in the same step. When a 100bp PCR product is sequenced bi-directionally with two tagged primers, all traces, regardless of orientation, will provide information about the mutation. Hence both primers provide "address information" about which plant contains which mutation.

[0058] The tag can be any number of nucleotides, but preferably contains 2, 3, 4 or 5 nucleotides. With 4 nucleotides permuted, 256 tags are possible, whereas 3 nucleotides permuted provide 64 different tags. In the illustration used, the tags preferably differ by >1 base, so preferred tags are 4 bp in length. Amplification using these primers results in a library of tagged amplification products.

[0059] In certain embodiments, a system of tags can be used wherein the amplification process includes

(1) a long PCR primer comprising (a) a 5'-constant section linked to (b) a degenerate tag section (NNNN) linked to (c) a gene specific section- 3' and

(2) a short PCR primer in subsequent amplifications that consists of (a) the 5'-contract section linked to (b) non-degenerate tag section-3' (*i.e.*, a selection amongst NNNN).

The non-degenerate tag section can be unique for each sample, for example, ACTG for sample 1, AATC for sample 2, *etc*. The short primer anneals to a subset of the long primer. The constant section of the primer can be used as a sequence primer. See Fig. 3.

[0060] The library preferably comprises equal, amounts of PCR products from all amplified pools. In the illustrative example, the library contains 3072 plants x 100bp= 307kb sequence to be determined.

[0061] The PCR products in the library are subjected to a sequencing process as disclosed above. In particular, the PCR products are attached to beads using the sequence primer binding site that corresponds to the sequence linked to the bead. Thus the present embodiment does not require fragmentation and adapter ligation. Rather, in this embodiment, the adapters have been introduced earlier via the PCR primer design. This improves the reliability of the method. Following the annealing to the beads, sequencing is performed as described above, *i.e.*, (1) emulsification of the beads in water-in-oil microreactors, (2) emulsion PCR to amplify the individual ssDNA molecules on beads; (3) selection of /enrichment for beads containing amplified ssDNA molecules on their surface, (4) transfer of the DNA carrying beads to a PicoTiterPlate®; and (5) simultaneous sequencing in 100,000 wells by a method that generates a pyrophosphate light signal. Typical output is about 200.000x100-200 bp sequences, representing a 66 fold coverage of all PCR products in the library.

[0062] Clustering and alignment is performed essentially as described above. The individual plant containing the mutation can be identified using the tags. In the examples, the combination of the 3 tags denotes the positive pools and the consequently the coordinates of the individual plant in the pools.

[0063] Confirmation of the mutation by re-sequencing of the PCR product of the identified mutant sample is as described above.

[0064] Various pooling strategies can be used with the present invention, examples of which are multidimensional pooling (including 3D pooling) or column-, row- or plate pooling.

[0065] High throughput sequencing methods that can be used here are described, for example, in Shendure et al., Science 309:1728-32. Examples include microelectrophoretic sequencing, hybridization sequencing/sequencing by hybridization (SBH), cyclic-array sequencing on amplified molecules, cyclic-array sequencing on single molecules, non-cyclical, single-molecule, real-time methods, such as , polymerase sequencing, exonuclease sequencing, or nanopore sequencing.

[0066] For optimal results, fragments or amplified products should be sequenced with sufficient redundancy. Redundancy permits distinction between a sequencing error and a genuine possible mutation. In certain embodiments, the redundancy of the sequencing is preferable at least 4, more preferably at least 5, but, as can be seen from the Examples, redundancies of more than 10, preferably more than 25 or even more than 50 are considered advantageous, although not essential for this invention..

[0067] Advantages of the methods of the present invention reside *inter alia* in the fact that mutations can be assessed *in silico* for their impact on gene function, meaning that a selection is made for the active mutations. Mutations conferring only silent substitutions can be selected against, thereby making the overall process more economical and efficient. This is a particular advantage with regard to the known CEL I based TILLING technology because the majority of CEL

IDT Exhibit 1020.6421

l mutations are C/G to T/A transitions, of which only 5% commonly create stop codons (Colbert *et al.* 2001). The vast majority are missense mutations of reduced interest. Efficient recognition of members in a population with stop codon mutations economizes the process and obviates the need for additional screening of individual members of positive pools.

[0068] All mutations can be found with equal probability, irrespective of their position in the PCR product, in particular when the whole target sequence is screened.

[0069] The method further avoids the use of CEL I digestion, heteroduplex formation and cumbersome gel scoring. The invention is therefore insensitive to pooling limitations associated with CEL I technology.

[0070] The invention further relates to kits that may contain one or more compounds selected form the group consisting of: one or more (labeled) primers for a particular gene or trait, mutation- or allele-specific primers. The kits may further contain beads, sequencing primers, software, descriptions for pooling strategies and other components that are known for kits per se. In certain embodiments, kits are provided that are dedicated to find specific mutations, for instance disease-related mutations.

[0071] The invention is now illustrated here in below.

EXAMPLES

[0072] Screening a TILLING population can be advanced by using novel high-throughput sequencing methods, such as that of 454 Life Sciences (Margulies *et al.*, 2005) or Polony Sequencing (Shendure *et al.*, 2005). With the current state-of-the-art, 454 Life Sciences technology produces approximately 20 Mb sequence in a single sequencing run. Read lengths are approximately 100 bp per read. Assuming the screening of a population consisting of 3072 plants for mutations in a 1500 bp gene (as described in the above-cited reference in Chapter 2), two approaches are envisaged and described in more detail below.

(1) an approach where the entire 1500 bp gene is investigated for the presence of EMS-induced mutations; and
(1) an approach where one or several 100 bp stretches are investigated for the presence of EMS-induced mutations.

EXAMPLE I

Screening the Entire 1500 bp Region:

[0073] Genomic DNA of 3072 plants of the TILLING population is isolated. A 3-D pooling scheme of equal amounts of DNA per plant is set up (e.g., 15 x 15 x 14), resulting in 44 pools (15+15+14 = 44) containing 3072/14= 219 or 3072/15 = 205 different DNA samples (Vandenbussche *et al.*, *supra*).

[0074] This pooling step serves to permit identification of a plant containing an observed mutation after one round of PCR screening (step 8). Pooling of genomic DNAs further serves to normalize DNAs prior to PCR amplification to increase the probability that all DNAs are represented equally in the sequence library.

[0075] The 1500 bp gene is amplified from the pooled DNA samples using 1 pair of unlabelled PCR primers.

[0076] Equal amounts of PCR products from all pools wells are pooled to create a pooled PCR products library (complexity 3072 plants x 1500bp = 4.6 Mb sequence).

[0077] The pooled PCR product library is subjected to shotgun sequencing using conventional technologies (such as those provided by 454 Life Sciences) wherein PCR products are randomly fragmented, amplified on individual beads and sequenced on the bead. Output is approximately 200,000 100bp sequences, representing 4- to 5-fold coverage of all PCR products in the library).

[0078] All sequences are clustered. The majority of sequences are wild-type but EMS-induced mutations (and sequence errors) are observed as well. Since PCR products are sequenced with 4-5 fold redundancy, multiple observations of the same sequence change is indicative of a mutation rather than a sequencing error (Figure 1).

[0079] Mutations are assessed for their impact on gene function such as introduction of a stopcodon.

[0080] An allele-specific primer targeting a mutation of interest (with 3' Locked Nucleic Acid; LNA; or Peptide Nucleic Acid; PNA) is designed to be used in combination with either the forward or reverse primer used in step 3 to screen the 3-D pooled DNA sample plate. Allele-specific PCR will result in three positive pools (one of each dimension), which specifies the library address of the mutant plant.

[0081] The mutation is confirmed by amplifying the 1500 bp gene using the primers of step 3, followed by (bi-directional) Sanger sequencing.

IDT Exhibit 1020.6422

EXAMPLE II

Screening 100 bp Stretches

(100 bp is the read length of one 454 sequence run)

[0082] Genomic DNA of 3072 plants of the TILLING population is isolated. A 3-D pooling scheme of equal amounts of DNA per plant is set up (*e.g.*, 15 x 15 x 14), resulting in 44 pools (15+15+14 = 44) containing 3072/14= 219 or 3072/15 = 205 different DNA samples (Vandenbussche *et al.*, *supra*).

[0083] This pooling step serves to permit identification of the plant containing an observed mutation directly from the sequence data. Pooling of genomic DNAs further serves to normalize DNAs prior to PCR amplification to increase the probability that all DNAs are represented equally in the sequence library.

[0084] A 100bp (or 200bp) region of the gene is amplified from a the pools by PCR using tagged unlabelled PCR primers. This requires 44 forward and 44 reverse primers (one for each pool of each dimension) with the following configuration:

[0085] 5'-Sequence primer binding site---4 bp Tag---Gene specific primer sequence-3'.

[0086] By using tailed forward and reverse primers containing a 4 bp sequence tag that is different for each of the 44 primers representing all pool dimensions, the specific plant origin of each sequence is known as the sequence primer anneals upstream of the tag. Hence the tag sequence in present in each sequence trace. A 4 bp tag allows $4^4 = 256$ different tags. A 3bp tag allows 64 different tag sequences - sufficient to distinguish 44 tags- but tag sequences differing by more than 1 base are preferred.

[0087] Equal amounts of PCR products from all pools wells are pooled to create a pooled PCR products library (complexity 3072 plants x 100 bp = 307 kb sequence).

[0088] The pooled PCR product library is provided to 454 for sequencing, *i.e.*, PCR products are amplified and sequenced on the beads. Output is approximately 200,000 100bp sequences, representing 66-fold coverage of all PCR products in the library.

[0089] All sequences (from either direction) are clustered; the majority of sequences are wild-type sequences but EMS-induced mutations (and sequence errors) are observed as well. Since PCR products are sequenced with 66 fold redundancy, multiple observations of the same sequence change are indicative of a mutation rather than a sequencing error (Figure 1).

[0090] The coordinates of the individual plant containing the mutation will be known immediately based on the unique combination of 3 tags sequences that occur in the sequence traces harboring the mutation (Figure 2).

[0091] The mutation is confirmed by amplifying the 1500 bp gene using the primers of step 3, followed by (bi-directional) Sanger sequencing.

Example III. Identifying specific mutations in a mutant library of tomato

Mutant library of tomato

[0092] This example describes the screening of a mutant library of tomato by massive parallel sequencing in order to identify point mutations in a specific locus (target gene). The mutant library used is an isogenic library of inbred determinate tomato cultivar M82 consisting of 5075 M2 families derived from EMS mutagenesis treatments. Seeds of each of the 5075 M2 families were stored at 10% RH and 7°C. The origin and characteristics of the library are described in Menda et al. (Plant J. 38: 861-872, 2004) and in a database on http://zamir.sgn.cornell.edu/mutants/index.html.

DNA isolation

[0093] Leaf material was harvested from 5 individual greenhouse-grown plants of each of 3072 M2 families randomly chosen from the library. As any mutation occurring in the library will segregate in a Mendelian fashion in the M2 offspring, the pooling of the leaf material of 5 individual M2 plants reduced the likelihood of overlooking any mutation as a consequence of segregation to less than 0.1%. Genomic DNA was isolated from the pooled leaf material using a modified CTAB procedure described by Stuart and Via (Biotechniques, 14: 748-750, 1993). DNA samples were diluted to a concentration of 100 ng/µl in TE (10 mM Tris-HCl pH 8.0, 1 mM EDTA) and stored at -20 °C in 96-well microtitre plates.

3D pooling of the DNA samples

[0094] The isolated DNA samples were normalized to a concentration of 20 ng/µl and subsequently pooled 4-fold resulting in 768 samples comprised in eight 96-well microtitre plates. Subsequently, these eight microtitre plates were

IDT Exhibit 1020.6423

subjected to a 3D pooling strategy, resulting in 28 pools of DNA. The 3D pooling strategy consisted of pooling together all DNAs in three different manners, thus ensuring that each single 4-fold pool occurs only once in an X-coordinate pool, only once in a Y-coordinate pool and only once in a Z-coordinate pool.

X-pools were assembled by pooling all DNA samples together per column of eight wells (e.g. A1-H1) from all eight microtitre plates, resulting in 12 X-pools. Each X-pool therefore held 8 (wells in a column) × 8 (plates) = 64 samples of 4-fold pools, representing 256 M2 families.

Y-pools were assembled by pooling all DNA samples together per row of twelve wells (e.g. A1-A12) from all eight microtitre plates, resulting in 8 Y-pools. Each Y-pool therefore held 12 (wells in a row) × 8 (plates) = 96 samples of 4-fold pools, representing 384 M2 families.

Z-pools were assembled by pooling all DNA samples together from an entire microtitre plate, resulting in 8 Z-pools. Each Z-pool therefore held 12 × 8 = 96 samples of 4-fold pools, representing 384 M2 families.

Target locus

[0095] The target locus in this example was part of the tomato gene for eucaryotic initiation factor 4E (eIF4E). This gene has been shown to be involved in susceptibility to infection of potyviruses in Arabidopsis (Duprat et al., Plant J. 32: 927-934, 2002), lettuce (Nicaise et al. Plant Physiol. 132: 1272-1282, 2003) and Solanaceae (Ruffel et al., Plant J. 32: 1067-1075, 2002; Mol. Gen. Genomics 274: 346-353, 2005), and specific mutations in this gene are associated with recessive potyvirus resistance. The mutation screening described in this example was aimed to identify additional mutations in the tomato eIF4E gene as possible sources of new potyvirus resistance. For the tomato eIF4E, only the cDNA sequence was known (NCBI accession numbers AY723733 and AY723734). Using a PCR approach using primers designed on the basis of the cDNA sequence, fragments of the genomic sequence of the eIF4E locus of tomato cultivar Moneyberg were amplified and sequenced. This resulted in a sequence of most of the genomic locus of tomato eIF4E. The locus consists of 4 exons and 3 introns. For the mutation screening, exon 1 of the gene was chosen as the target sequence (SEQ ID 57).

SEQ ID 57: Sequence of exon 1 of tomato Moneyberg eIF4E :

```
ATGGCAGCAGCTGAAATGGAGAGAACGATGTCGTTTGATGCAGCTGAGAAGTTGAAGGCCGCCG
ATGGAGGAGGAGGAGAGGTAGACGATGAACTTGAAGAAGGTGAAATTGTTGAAGAATCAAATGA
TACGGCATCGTATTTAGGGAAAGAAATCACAGTGAAGCATCCATTGGAGCATTCATGGACTTTT
TGGTTTGATAACCCTACCACTAAATCTCGACAAACTGCTTGGGGAAGCTCACTTCGAAATGTCT
ACACTTTCTCCACTGTTGAAAATTTTTGGGG
```

Primer design for target locus amplification

[0096] Primers were designed for the PCR amplification of exon 1 of tomato eIF4E. The forward primers were designed to correspond to the ATG start codon of the Open Reading Frame of exon 1, with 5' of the ATG a tag sequence of four bases, providing a unique identifier for each of the 28 pools. At the far 5' end of the forward PCR primers, a 5'-C was added. All primers were phosphorylated at their 5'end to facilitate subsequent ligation of adaptors. The sequence and names of the 28 forward primers are listed in Table 1. The tag sequences are underlined.

Table 1. Forward primers, sequences and pool identification for exon 1 amplification.

name	sequence	3D pool	SEQ ID #
06I009	CACACATGGCAGCAGCTGAAATGG	X1	1
06I010	CACAGATGGCAGCAGCTGAAATGG	X2	2
06I011	CACGAATGGCAGCAGCTGAAATGG	X3	3
06I012	CACGTATGGCAGCAGCTGAAATGG	X4	4
06I013	CACTCATGGCAGCAGCTGAAATGG	X5	5
06I014	CACTGATGGCAGCAGCTGAAATGG	X6	6
06I015	CAGACATGGCAGCAGCTGAAATGG	X7	7
06I016	CAGAGATGGCAGCAGCTGAAATGG	X8	8

IDT Exhibit 1020.6424

(continued)

name	sequence	3D pool	SEQ ID #
06I017	CAGCAATGGCAGCAGCTGAAATGG	X9	9
06I018	CAGCTATGGCAGCAGCTGAAATGG	X10	10
06I019	CAGTCATGGCAGCAGCTGAAATGG	X11	11
06I020	CAGTGATGGCAGCAGCTGAAATGG	X12	12
06I021	CATCGATGGCAGCAGCTGAAATGG	Y1	13
06I022	CATGCATGGGAGCAGCTGAAATGG	Y2	14
06I023	CTACGATGGCAGCAGCTGAAATGG	Y3	15
061024	CTAGCATGGCAGCAGCTGAAATGG	Y4	16
061025	CTCACATGGCAGCAGCTGAAATGG	Y5	17
06I026	CTCAGATGGCAGCAGCTGAAATGG	Y6	18
06I027	CTCGAATGGCAGCAGCTGAAATGG	Y7	19
06I028	CTCGTATGGCAGCAGCTGAAATGG	Y8	20
06I029	CTCTCATGGCAGCAGCTGAAATGG	Z1	21
06I030	CTCTGATGGCAGCAGCTGAAATGG	Z2	22
06I031	CTGACATGGCAGCAGCTGAAATGG	Z3	23
06I032	CTGAGATGGCAGCAGCTGAAATGG	Z4	24
06I033	CTGCAATGGCAGCAGCTGAAATGG	Z5	25
06I034	CTGCTATGGCAGCAGCTGAAATGG	Z6	26
06I035	CTGTCATGGCAGCAGCTGAAATGG	Z7	27
06I036	CTGTGATGGCAGCAGCTGAAATGG	Z8	28

[0097] The reverse primers were designed to correspond to basepair position 267 to 287 of exon 1 in the non-coding strand. Again, 5' of the priming part the same series of tag sequences of four bases were included, providing a identifier for each of the 28 pools. At the far 5' end of the reverse PCR primers, a 5'-C was added. All primers were phosphorylated at their 5'end to facilitate subsequent ligation of adaptors. The sequence and names of the 28 reverse primers are listed in Table 2. The tags are underlined.

Table 2. Reverse primers, sequences and pool identification for exon 1 amplification.

name	sequence	3D pool	SEQ ID #
06I037	CACACCCCCAAAAATTTTCAACAGTG	X1	29
06I038	CACAGCCCCAAAAATTTTCAACAGTG	X2	30
06I039	CACGACCCCAAAAATTTTCAACAGTG	X3	31
06I040	CACGTCCCCAAAAATTTTCAACAGTG	X4	32
06I041	CACTCCCCCAAAAATTTTCAACAGTG	X5	33
06I042	CACTGCCCCAAAAATTTTCAACAGTG	X6	34
06I043	CAGACCCCCAAAAATTTTCAACAGTG	X7	35
06I044	CAGAGCCCCAAAAATTTTCAACAGTG	X8	36
06I045	CAGCACCCCAAAAATTTTCAACAGTG	X9	37
06I046	CAGCTCCCCAAAAATTTTCAACAGTG	X10	38
06I047	CAGTCCCCCAAAAATTTTCAACAGTG	X11	39

IDT Exhibit 1020.6425

(continued)

name	sequence	3D pool	SEQ ID #
06I048	CAGTGCCCCAAAAATTTTCAACAGTG	X12	40
06I049	CATCGCCCCAAAAATTTTCAACAGTG	Y1	41
06I050	CATGCCCCCAAAAATTTTCAACAGTG	Y2	42
06I051	CTACGCCCCAAAAATTTTCAACAGTG	Y3	43
06I052	CTAGCCCCCAAAAATTTTCAACAGTG	Y4	44
06I053	CTCACCCCCAAAAATTTTCAACAGTG	Y5	45
061054	CTCAGCCCCAAAAATTTTCAACAGTG	Y6	46
06I055	CTCGACCCCAAAAATTTTCAACAGTG	Y7	47
06I056	CTCGTCCCCAAAAATTTTCAACAGTG	Y8	48
06I057	CTCTCCCCCAAAAATTTTCAACAGTG	Z1	49
06I058	CTCTGCCCCAAAAATTTTCAACAGTG	Z2	50
06I059	CTGACCCCCAAAAATTTTCAACAGTG	Z3	51
06I060	CTGAGCCCCAAAAATTTTCAACAGTG	Z4	52
06I061	CTGCACCCCAAAAATTTTCAACAGTG	Z5	53
06I062	CTGCTCCCCAAAAATTTTCAACAGTG	Z6	54
06I063	CTGTCCCCCAAAAATTTTCAACAGTG	Z7	55
06I064	CTGTGCCCCAAAAATTTTCAACAGTG	Z8	56

Target locus amplification

[0098] The exon 1 of the target locus was amplified from the 3D pooled DNAs using the forward and reverse primers described above. For each PCR reaction, a forward and a reverse primer were used with identical tags. For the amplification of exon 1 from each of the 28 3D pools, a different set of forward and reverse primers was used.

[0099] The PCR amplification reaction conditions for each sample were as follows:

25 μl DNA (= 50 ng); 5 μl RNase-mix ; 10 μl 5x Herculase PCR-buffer; 0.6 μl of the four dNTPs (20mM); 1.25 μl forward primer (50 ng/μl); 1.25 μl reverse primer (50 ng/μl); 0.5 μl Herculase DNA polymerase; 28.9 μl milliQ-purified water. The RNase-mix consisted of 157,5 μl milliQ-purified water + 17,5 μl RNase.

[0100] PCR amplifications were performed in a PE9600 thermocycler with a gold or silver block using the following conditions: 2 minutes hot-start of 94°C, followed by 35 cycles of 30 sec at 94°C, 30 sec at 53°C, 1 min at 72°C, and a final stationary temperature of 4°C. The PCR amplification efficiency was checked by analysis of 10 μl of PCR products on a 1% agarose gel. Fig. 4 shows the efficient amplification of exon 1 PCR products from each of the 28 3D pools in comparison to a concentration range of lambda DNA on the same gel.

[0101] Following amplification, equal amounts of PCR products were mixed and purified using the QIAquick PCR Purification Kit (QIAGEN), according to the QIAquick® Spin handbook (page 18). On each column a maximum of 100 μl of product was loaded. Products were eluted in 10 mM Tris-EDTA.

Sequence library preparation and high-throughput sequencing

[0102] Mixed amplification products from the 3D pools were subjected to high-throughput sequencing on a GS20 sequencer using 454 Life Sciences sequencing technology as described by Margulies et al. (Nature 437: 376-380, 2005, and Online Supplements). Specifically, the PCR products were ligated to adaptors to facilitate emulsion-PCR amplification and subsequent fragment sequencing as described by Margulies *et al.*, The 454 adaptor sequences, emulsion PCR primers, sequence primers and sequence run conditions were all as described by Margulies *et al.*, The linear order of functional elements in an emulsion-PCR fragment amplified on Sepharose beads in the 454 sequencing process was as follows:

IDT Exhibit 1020.6426

[0103] 454 PCR adaptor - 454 sequence adaptor - C-nucleotide - 4 bp tag - target amplification primer sequence 1 - target fragment internal sequence - target amplification primer sequence 2 - 4 bp tag - G-nucleotide - 454 sequence adaptor - 454 PCR adaptor - Sepharose bead.

454 sequence run data-processing.

[0104] After base calling with 454 software for each region of the microtiter plate a file with FASTA formatted sequences was produced. These were concatenated into one file. Within this file a search was conducted with a regular expression to a 100% match of the forward primer preceded with 5 nucleotides (C plus four bp tag sequence). The same was done with the reverse primer extended with 5 nucleotides (C plus tag sequence). All sequences were then grouped by their tag sequence (pool indentifiers) in separate files. Each file was analysed with the ssahaSNP tool and the known exon 1 nucleotide sequence as a reference. The ssahaSNP tool reported about all single nucleotide sequence differences and "indels" (single base insertions or deletions as a result of either mutagenesis or erroneous base-calling) of the 454 sequences versus the reference genome. These single nucleotide sequence difference and indel statistics were saved in a database and used for error rate analysis and point mutation identification.

454 sequencing error rate

[0105] The total number of correct sequences obtained from the data processing for all 28 pools combined was 247,052. The sequences were divided in two groups, those that aligned with the forward primer and coding strand (5' end) of the exon 1 PCR product (128,594 = 52%), and those that aligned with the reverse primer and the complementary strand of the PCR product (118,458 = 48%). The number of sequences obtained from each of the different pools and alignment groups ranged from 69 to 7269. On average, each of the 3072 M2 families should be represented 80 times in the total collection of sequences, and each allele 40 times.

[0106] Within the alignment group corresponding to the forward primer, 1338 sequences out of 128,594 (1.2%) showed one or more single nucleotide sequence differences in relation to the eIF4E reference sequence along a stretch of 63 bases of aligned target sequence. For the reverse primer group, 743 sequences out of 118,458 (0.6%) showed one or more single nucleotide sequence differences in relation to the eIF4E reference sequence along a stretch of 102 bases of aligned target sequence. Therefore, the single base substitution error rate for both sequence groups combined equals 0.84% for a 165 base stretch, or 0.0051% per base position (0.5 errors per 10,000 bases). This error rate is similar to the one reported by Margulies *et al.* of 0.004% for individual read substitution errors in test sequences, but much lower than for whole-genome resequencing (0.68%).

[0107] A similar analysis of the occurence of indels in both alignment groups revealed an indel incidence of 3883 (forward primer group) and 3829 (reverse primer group) in a total of 247,052 sequences (is 3.1 % in a 165 bp stretch). The indel occurence rate therefore equals 0.01891 % per base position (1.89 indels per 10,000 bases). The indel rate is significant higher than the base substitution error rate. Both types of sequencing error combined occur on average at a frequency of 2.39 per 10,000 bases, or 0.024 per base position. This error rate is much lower than reported by Margulies *et al.*, and may be explained by the absence of long homopolymer stretches in the eIF4e exon 1 sequence.

Detection of a mutation in the target locus

[0108] Because the objective of this screen is the identification of (EMS)-induced point mutations (preferentially C→T and G→A mutations), all sequences representing indels in comparison to the reference sequence were discarded for the sake of the analysis in this example. Most of the single base substitutions occurred only once in any given 3D pool, some occurred 2 or 3 times, or rarely more often. Since these single base substitutions occur more or less uniformly at every position of the aligned sequence, and at a more or less uniform frequency of 0.005% per base, they were assumed to represent sequencing errors, and not specific mutations that exist in the mutant library. However, at a few specific base positions in the scanned sequence, a much higher incidence of a specific single base sequence difference occurs. Such single base sequence differences reveal mutations in the library, when the following criteria are fulfilled:

1. the single base sequence difference represents an C→T or G→A mutation;
2. the incidence is higher than 20 per 10,000 sequence reads per 3D pool;
3. the single base sequence difference occurs in precisely one and not more than one X-pool, Y-pool and Z-pool.

[0109] In this example, one such mutation was found in the alignment group corresponding to the reverse primer, at base position 221 of the eIF4E exon 1 sequence. This mutation, a G→A mutation (corresponding to C→T in the complementary strand) occured in pool X12 at a frequency of 70 per 10,000 sequences, in pool Y3 at a frequency of 33 per 10,000 and in pool Z6 at 62 per 10,000 sequences. This same mutation at the same position did not occur in any of the

IDT Exhibit 1020.6427

other pools, not even at background error rates.

[0110] The unique occurrence of this G221A mutation in only the three pools allowed the identification of the original 4-fold pool of DNA, representing four M2 families. DNA of each of these four M2 families was amplified individually with the primers 06F598 and 06F599 that are identical to the forward and reverse primers of Tables 1 and 2, but without the 5' five base sequence tags. The amplified PCR products were subjected to conventional Sanger sequencing. The sequence of the eIF4E gene in one of the four families (coded "24") revealed a dual peak at position 221, corresponding to an overlapping G and A. This is indicative of an M2 family pool, in which half the alleles are wild-type, and the other half carry the G221A point mutation (Fig. 2). The sequences of the other M2 families around base position 221 were according to the reference (wild-type).

[0111] The mutation causes an arginine to glutamine substitution. Seeds of this particular M2 family were planted in the greenhouse in order to select for homozygous mutant individuals, that will be used for phenotyping.

[0112] In a similar manner, two other point mutations were identified in the 454 sequence reads. An estimation of the mutation density of the M82 tomato mutant library therefore equals 3 mutations per 165 bp scanned sequence, or 18 mutations per 1000 bases in 3072 M2 families. This corresponds to mutation densities reported for *Arabidopsis* (Greene et al., Genetics 164: 731-740,2003).

References

[0113]

Colbert et al..2001. High-throughput screening for induced point mutations. Plant Physiology 126: 480-484.

Duprat et al., 2002. The Arabidopsis eukaryotic initiation factor (iso)4E is dispensable for plant growth but required for susceptibility to potyviruses. Plant J. 32: 927-934.

Epinat et al., 2003. A novel engineered meganuclease induces homologous recombination in yeast and mammalian cells. Nucleic Acids Research, 31(11): 2952-2962.

Havre et al., 1993. Targeted mutagenesis of DNA using triple helix-forming oligonucleotides linked to psoralen. Proc.Natl.Acad.Sci. USA 90: 7879-7883.

McCallum et al., 2000. Targeted screening for induced mutations. Nature Biotechnology 18: 455-457.

Greene et al., 2003. Spectrum of chemically induced mutations from a large-scale reverse-genetic screen in Arabidopsis. Genetics 164: 731-740.

Lloyd et al., 2005. Targeted mutagenesis using zinc-finger nucleases in Arabidopsis. Proc.Natl.Acad.Sci. USA 102: 2232-2237.

Margulies et al., 2005. Genome sequencing in microfabricated high-density picolitre reactions. Mature 437: 376-380.

Menda et al., 2004. In silico screening of a saturated mutation library of tomato. Plant J. 38: 861-872.

Nicaise et al., 2003. The eukaryotic translation initiation factor 4E controls lettuce susceptibility to the p otyvirus lettuce mosaic virus1. Plant Physiol. 132: 1272-1282.

Ruffel et al., 2002. A natural recessive resistance gene against potato virus Y in pepper corresponds to the eukaryotic initiation factor 4E (eIF4E). Plant J. 32: 1067-1075.

Ruffel et al., 2005. The recessive potyvirus resistance gene pot-1 is the tomato orthologue of the pepper pvr2-eIF4E gene. Mol. Gen. Genomics 274: 346-353.

Shendure et al., 2005. Accurate multiplex polony sequencing of an evolved bacterial genome. Scienceexpress Report, August 4.

Stuart and Via, 1993. A rapid CTAB DNA isolation technique useful for RAPD fingerprinting and other PCR applications. Biotechniques, 14: 748-750.

IDT Exhibit 1020.6428

Vandenbussche et al., 2003. Toward the analysis of the petunia MADS box gene family by reverse and forward transposon insertion mutagenesis approaches: B, C, and D floral organ identity functions require SEPALLATA-like MADS box genes in petunia. The Plant Cell 15:2680-2693.

SEQUENCE LISTING

[0114]

<110> Keygene NV

<120> HIGH THROUGHPUT SCREENING OF MUTAGENIZED POPULATIONS

<130> P27926PC00

<150> 60/721,528
<151> 2005-09-29

<160> 57

<170> PatentIn version 3.3

<210> 1
<211> 24
<212> DNA
<213> artificial;

<400> 1
cacacatggc agcagctgaa atgg 24

<210> 2
<211> 24
<212> DNA
<213> artificial;

<400> 2
cacagatggc agcagctgaa atgg 24

<210> 3
<211> 24
<212> DNA
<213> artificial;

<400> 3
cacgaatggc agcagctgaa atgg 24

<210> 4
<211> 24
<212> DNA
<213> artificial;

<400> 4
cacgtatggc agcagctgaa atgg 24

<210> 5
<211> 24
<212> DNA
<213> artificial;

IDT Exhibit 1020.6429

<400> 5
cactcatggc agcagctgaa atgg 24

<210> 6
<211> 24
<212> DNA
<213> artificial;

<400> 6
cactgatggc agcagctgaa atgg 24

<210> 7
<211> 24
<212> DNA
<213> artificial;

<400> 7
cagacatggc agcagctgaa atgg 24

<210> 8
<211> 24
<212> DNA
<213> artificial;

<400> 8
cagagatggc agcagctgaa atgg 24

<210> 9
<211> 24
<212> DNA
<213> artificial;

<400> 9
cagcaatggc agcagctgaa atgg 24

<210> 10
<211> 24
<212> DNA
<213> artificial;

<400> 10
cagctatggc agcagctgaa atgg 24

<210> 11
<211> 24
<212> DNA
<213> artificial;

<400> 11
cagtcatggc agcagctgaa atgg 24

<210> 12
<211> 24
<212> DNA
<213> artificial;

<400> 12
cagtgatggc agcagctgaa atgg 24

IDT Exhibit 1020.6430

<210> 13
<211> 24
<212> DNA
<213> artificial;

<400> 13
catcgatggc agcagctgaa atgg 24

<210> 14
<211> 24
<212> DNA
<213> artificial;

<400> 14
catgcatggc agcagctgaa atgg 24

<210> 15
<211> 24
<212> DNA
<213> artificial;

<400> 15
ctacgatggc agcagctgaa atgg 24

<210> 16
<211> 24
<212> DNA
<213> artificial;

<400> 16
ctagcatggc agcagctgaa atgg 24

<210> 17
<211> 24
<212> DNA
<213> artificial;

<400> 17
ctcacatggc agcagctgaa atgg 24

<210> 18
<211> 24
<212> DNA
<213> artificial;

<400> 18
ctcagatggc agcagctgaa atgg 24

<210> 19
<211> 24
<212> DNA
<213> artificial;

<400> 19
ctcgaatggc agcagctgaa atgg 24

<210> 20
<211> 24

IDT Exhibit 1020.6431

<212> DNA
<213> artificial;

<400> 20
ctcgtatggc agcagctgaa atgg 24

<210> 21
<211> 24
<212> DNA
<213> artificial;

<400> 21
ctctcatggc agcagctgaa atgg 24

<210> 22
<211> 24
<212> DNA
<213> artificial;

<400> 22
ctctgatggc agcagctgaa atgg 24

<210> 23
<211> 24
<212> DNA
<213> artificial;

<400> 23
ctgacatggc agcagctgaa atgg 24

<210> 24
<211> 24
<212> DNA
<213> artificial;

<400> 24
ctgagatggc agcagctgaa atgg 24

<210> 25
<211> 24
<212> DNA
<213> artificial;

<400> 25
ctgcaatggc agcagctgaa atgg 24

<210> 26
<211> 24
<212> DNA
<213> artificial;

<400> 26
ctgctatggc agcagctgaa atgg 24

<210> 27
<211> 24
<212> DNA
<213> artificial;

IDT Exhibit 1020.6432

<400> 27
ctgtcatggc agcagctgaa atgg 24

<210> 28
<211> 24
<212> DNA
<213> artificial;

<400> 28
ctgtgatggc agcagctgaa atgg 24

<210> 29
<211> 26
<212> DNA
<213> artificial;

<400> 29
cacaccccca aaaattttca acagtg 26

<210> 30
<211> 26
<212> DNA
<213> artificial;

<400> 30
cacagcccca aaaattttca acagtg 26

<210> 31
<211> 26
<212> DNA
<213> artificial;

<400> 31
cacgacccca aaaattttca acagtg 26

<210> 32
<211> 26
<212> DNA
<213> artificial;

<400> 32
cacgtcccca aaaattttca acagtg 26

<210> 33
<211> 26
<212> DNA
<213> artificial;

<400> 33
cactccccca aaaattttca acagtg 26

<210> 34
<211> 26
<212> DNA
<213> artificial;

<400> 34
cactgcccca aaaattttca acagtg 26

IDT Exhibit 1020.6433

<210> 35
<211> 26
<212> DNA
<213> artificial;

<400> 35
cagaccccca aaaattttca acagtg 26

<210> 36
<211> 26
<212> DNA
<213> artificial;

<400> 36
cagagcccca aaaattttca acagtg 26

<210> 37
<211> 26
<212> DNA
<213> artificial;

<400> 37
cagcacccca aaaattttca acagtg 26

<210> 38
<211> 26
<212> DNA
<213> artificial;

<400> 38
cagctcccca aaaattttca acagtg 26

<210> 39
<211> 26
<212> DNA
<213> artificial;

<400> 39
cagtccccca aaaattttca acagtg 26

<210> 40
<211> 26
<212> DNA
<213> artificial;

<400> 40
cagtgcccca aaaattttca acagtg 26

<210> 41
<211> 26
<212> DNA
<213> artificial;

<400> 41
catcgcccca aaaattttca acagtg 26

<210> 42
<211> 26

IDT Exhibit 1020.6434

<212> DNA
<213> artificial;

<400> 42
catgccccca aaaattttca acagtg 26

<210> 43
<211> 26
<212> DNA
<213> artificial;

<400> 43
ctacgccccа aaaattttca acagtg 26

<210> 44
<211> 26
<212> DNA
<213> artificial;

<400> 44
ctagccccca aaaattttca acagtg 26

<210> 45
<211> 26
<212> DNA
<213> artificial;

<400> 45
ctcaccccca aaaattttca acagtg 26

<210> 46
<211> 26
<212> DNA
<213> artificial;

<400> 46
ctcagcccca aaaattttca acagtg 26

<210> 47
<211> 26
<212> DNA
<213> artificial;

<400> 47
ctcgacccca aaaattttca acagtg 26

<210> 48
<211> 26
<212> DNA
<213> artificial;

<400> 48
ctcgtcccca aaaattttca acagtg 26

<210> 49
<211> 26
<212> DNA
<213> artificial;

IDT Exhibit 1020.6435

<400> 49
ctctccccca aaaattttca acagtg 26

<210> 50
<211> 26
<212> DNA
<213> artificial;

<400> 50
ctctgcccca aaaattttca acagtg 26

<210> 51
<211> 26
<212> DNA
<213> artificial;

<400> 51
ctgaccccca aaaattttca acagtg 26

<210> 52
<211> 26
<212> DNA
<213> artificial;

<400> 52
ctgagcccca aaaattttca acagtg 26

<210> 53
<211> 26
<212> DNA
<213> artificial;

<400> 53
ctgcacccca aaaattttca acagtg 26

<210> 54
<211> 26
<212> DNA
<213> artificial;

<400> 54
ctgctcccca aaaattttca acagtg 26

<210> 55
<211> 26
<212> DNA
<213> artificial;

<400> 55
ctgtccccca aaaattttca acagtg 26

<210> 56
<211> 26
<212> DNA
<213> artificial;

<400> 56
ctgtgcccca aaaattttca acagtg 26

IDT Exhibit 1020.6436

<210> 57
<211> 287
<212> DNA
<213> Lycopersicon esculentum

<400> 57

atggcagcag ctgaaatgga gagaacgatg tcgtttgatg cagctgagaa gttgaaggcc 60

gccgatggag gaggaggaga ggtagacgat gaacttgaag aaggtgaaat tgttgaagaa 120

tcaaatgata cggcatcgta tttagggaaa gaaatcacag tgaagcatcc attggagcat 180

tcatggactt tttggtttga taaccctacc actaaatctc gacaaactgc ttggggaagc 240

tcacttcgaa atgtctacac tttctccact gttgaaaatt tttgggg 287

SEQUENCE LISTING

[0115]

<110> Keygene NV

<120> HIGH THROUGHPUT SCREENING OF MUTAGENIZED POPULATIONS

<130> P27926PC00

<150> 60/721,528
<151> 2005-09-29

<160> 57

<170> PatentIn version 3.3

<210> 1
<211> 24
<212> DNA
<213> Artificial

<220>
<223> Primers

<400> 1
cacacatggc agcagctgaa atgg 24

<210> 2
<211> 24
<212> DNA
<213> Artificial

<220>
<223> primers

<400> 2

IDT Exhibit 1020.6437

cacagatggc agcagctgaa atgg 24

<210> 3
<211> 24
<212> DNA
<213> Artificial

<220>
<223> primers

<400> 3
cacgaatggc agcagctgaa atgg 24

<210> 4
<211> 24
<212> DNA
<213> Artificial

<220>
<223> Primer

<400> 4
cacgtatggc agcagctgaa atgg 24

<210> 5
<211> 24
<212> DNA
<213> Artificial

<220>
<223> Primer

<400> 5
cactcatggc agcagctgaa atgg 24

<210> 6
<211> 24
<212> DNA
<213> Artificial

<220>
<223> Primer

<400> 6
cactgatggc agcagctgaa atgg 24

<210> 7
<211> 24
<212> DNA
<213> Artificial

<220>
<223> Primer

<400> 7
cagacatggc agcagctgaa atgg 24

<210> 8

IDT Exhibit 1020.6438

<211> 24
<212> DNA
<213> Artificial

<220>
<223> Primer

<400> 8
cagagatggc agcagctgaa atgg 24

<210> 9
<211> 24
<212> DNA
<213> Artificial

<220>
<223> Primer

<400> 9
cagcaatggc agcagctgaa atgg 24

<210> 10
<211> 24
<212> DNA
<213> Artificial

<220>
<223> Primer

<400> 10
cagctatggc agcagctgaa atgg 24

<210> 11
<211> 24
<212> DNA
<213> Artificial

<220>
<223> Primer

<400> 11
cagtcatggc agcagctgaa atgg 24

<210> 12
<211> 24
<212> DNA
<213> Artificial

<220>
<223> Primer

<400> 12
cagtgatggc agcagctgaa atgg 24

<210> 13
<211> 24
<212> DNA
<213> Artificial

IDT Exhibit 1020.6439

<220>
<223> Primer

<400> 13
catcgatggc agcagctgaa atgg 24

<210> 14
<211> 24
<212> DNA
<213> Artificial

<220>
<223> Primer

<400> 14
catgcatggc agcagctgaa atgg 24

<210> 15
<211> 24
<212> DNA
<213> Artificial

<220>
<223> Primer

<400> 15
ctacgatggc agcagctgaa atgg 24

<210> 16
<211> 24
<212> DNA
<213> Artificial

<220>
<223> Primer

<400> 16
ctagcatggc agcagctgaa atgg 24

<210> 17
<211> 24
<212> DNA
<213> Artificial

<220>
<223> Primer

<400> 17
ctcacatggc agcagctgaa atgg 24

<210> 18
<211> 24
<212> DNA
<213> Artificial

<220>
<223> Primer

IDT Exhibit 1020.6440

<400> 18
ctcagatggc agcagctgaa atgg 24

<210> 19
<211> 24
<212> DNA
<213> Artificial

<220>
<223> Primer

<400> 19
ctcgaatggc agcagctgaa atgg 24

<210> 20
<211> 24
<212> DNA
<213> Artificial

<220>
<223> Primer

<400> 20
ctcgtatggc agcagctgaa atgg 24

<210> 21
<211> 24
<212> DNA
<213> Artificial

<220>
<223> Primer

<400> 21
ctctcatggc agcagctgaa atgg 24

<210> 22
<211> 24
<212> DNA
<213> Artificial

<220>
<223> Primer

<400> 22
ctctgatggc agcagctgaa atgg 24

<210> 23
<211> 24
<212> DNA
<213> Artificial

<220>
<223> Primer

<400> 23
ctgacatggc agcagctgaa atgg 24

IDT Exhibit 1020.6441

<210> 24
<211> 24
<212> DNA
<213> Artificial

<220>
<223> Primer

<400> 24
ctgagatggc agcagctgaa atgg 24

<210> 25
<211> 24
<212> DNA
<213> Artificial

<220>
<223> Primer

<400> 25
ctgcaatggc agcagctgaa atgg 24

<210> 26
<211> 24
<212> DNA
<213> Artificial

<220>
<223> Primer

<400> 26
ctgctatggc agcagctgaa atgg 24

<210> 27
<211> 24
<212> DNA
<213> Artificial

<220>
<223> Primer

<400> 27
ctgtcatggc agcagctgaa atgg 24

<210> 28
<211> 24
<212> DNA
<213> Artificial

<220>
<223> Primer

<400> 28
ctgtgatggc agcagctgaa atgg 24

<210> 29
<211> 26
<212> DNA

IDT Exhibit 1020.6442

<213> Artificial

<220>
<223> Primer

<400> 29
cacaccccca aaaattttca acagtg 26

<210> 30
<211> 26
<212> DNA
<213> Artificial

<220>
<223> Primer

<400> 30
cacagcccca aaaattttca acagtg 26

<210> 31
<211> 26
<212> DNA
<213> Artificial

<220>
<223> Primer

<400> 31
cacgacccca aaaattttca acagtg 26

<210> 32
<211> 26
<212> DNA
<213> Artificial

<220>
<223> Primer

<400> 32
cacgtcccca aaaattttca acagtg 26

<210> 33
<211> 26
<212> DNA
<213> Artificial

<220>
<223> Primer

<400> 33
cactccccca aaaattttca acagtg 26

<210> 34
<211> 26
<212> DNA
<213> Artificial

<220>

IDT Exhibit 1020.6443

<223> Primer

<400> 34
cactgcccca aaaattttca acagtg 26

<210> 35
<211> 26
<212> DNA
<213> Artificial

<220>
<223> Primer

<400> 35
cagaccccca aaaattttca acagtg 26

<210> 36
<211> 26
<212> DNA
<213> Artificial

<220>
<223> Primer

<400> 36
cagagcccca aaaattttca acagtg 26

<210> 37
<211> 26
<212> DNA
<213> Artificial

<220>
<223> Primer

<400> 37
cagcacccca aaaattttca acagtg 26

<210> 38
<211> 26
<212> DNA
<213> Artificial

<220>
<223> Primer

<400> 38
cagctcccca aaaattttca acagtg 26

<210> 39
<211> 26
<212> DNA
<213> Artificial

<220>
<223> Primer

<400> 39

IDT Exhibit 1020.6444

cagtccccca aaaattttca acagtg 26

<210> 40
<211> 26
<212> DNA
<213> Artificial

<220>
<223> Primer

<400> 40 26
cagtgcccca aaaattttca acagtg 26

<210> 41
<211> 26
<212> DNA
<213> Artificial

<220>
<223> Primer

<400> 41 26
catcgcccca aaaattttca acagtg 26

<210> 42
<211> 26
<212> DNA
<213> Artificial

<220>
<223> Primer

<400> 42 26
catgccccca aaaattttca acagtg 26

<210> 43
<211> 26
<212> DNA
<213> Artificial

<220>
<223> Primer

<400> 43 26
ctacgcccca aaaattttca acagtg 26

<210> 44
<211> 26
<212> DNA
<213> Artificial

<220>
<223> Primer

<400> 44 26
ctagccccca aaaattttca acagtg 26

<210> 45

IDT Exhibit 1020.6445

<211> 26
<212> DNA
<213> Artificial

<220>
<223> Primer

<400> 45
ctcaccccca aaaattttca acagtg 26

<210> 46
<211> 26
<212> DNA
<213> Artificial

<220>
<223> Primer

<400> 46
ctcagcccca aaaattttca acagtg 26

<210> 47
<211> 26
<212> DNA
<213> Artificial

<220>
<223> Primer

<400> 47
ctcgacccca aaaattttca acagtg 26

<210> 48
<211> 26
<212> DNA
<213> Artificial

<220>
<223> Primer

<400> 48
ctcgtcccca aaaattttca acagtg 26

<210> 49
<211> 26
<212> DNA
<213> Artificial

<220>
<223> Primer

<400> 49
ctctccccca aaaattttca acagtg 26

<210> 50
<211> 26
<212> DNA
<213> Artificial

IDT Exhibit 1020.6446

<220>
<223> Primer

<400> 50
ctctgcccca aaaattttca acagtg 26

<210> 51
<211> 26
<212> DNA
<213> Artificial

<220>
<223> Primer

<400> 51
ctgaccccca aaaattttca acagtg 26

<210> 52
<211> 26
<212> DNA
<213> Artificial

<220>
<223> Primer

<400> 52
ctgagcccca aaaattttca acagtg 26

<210> 53
<211> 26
<212> DNA
<213> Artificial

<220>
<223> Primer

<400> 53
ctgcacccca aaaattttca acagtg 26

<210> 54
<211> 26
<212> DNA
<213> Artificial

<220>
<223> Primer

<400> 54
ctgctcccca aaaattttca acagtg 26

<210> 55
<211> 26
<212> DNA
<213> Artificial

<220>
<223> Primer

IDT Exhibit 1020.6447

<400> 55
ctgtccccca aaaattttca acagtg 26

<210> 56
<211> 26
<212> DNA
<213> Artificial

<220>
<223> primers

<400> 56
ctgtgcccca aaaattttca acagtg 26

<210> 57
<211> 287
<212> DNA
<213> Lycopersicon esculentum

<220>
<221> misc_feature
<223> Exon 1 of tomato moneyberg eIF4E

<400> 57

atggcagcag ctgaaatgga gagaacgatg tcgtttgatg cagctgagaa gttgaaggcc 60
gccgatggag gaggaggaga ggtagacgat gaacttgaag aaggtgaaat tgttgaagaa 120
tcaaatgata cggcatcgta tttagggaaa gaaatcacag tgaagcatcc attggagcat 180
tcatggactt tttggtttga taaccctacc actaaatctc gacaaactgc ttggggaagc 240
tcacttcgaa atgtctacac tttctccact gttgaaaatt tttgggg 287

Claims

1. A method for the detection of a mutation in a target sequence in a member of a mutagenized population, comprising the steps of:

   (a) isolating genomic DNA of each member of the mutagenized population to provide DNA samples of each population member;
   (b) pooling the DNA obtained in step (a);
   (c) amplifying the target sequence with a pair of (optionally labeled) primers from the DNA pools;
   (d) pooling the amplification products of step (c) to create a library of amplification products;
   (e) optionally, fragmenting the amplification products in the library;
   (f) determining the nucleotide sequence of the amplified products or fragments using high throughput sequencing;
   (g) identifying mutations by clustering/aligning the sequences of the fragments;
   (h) screening the identified mutations fora modified function of the target sequence;
   (i) designing a primer directed to hybridize to the identified mutation;
   (j) amplifying the DNA pools of step (b) with the primer of step (i) and one of the primers of step (c);
   (k) identify the population member(s) carrying the mutation;
   (l) optionally, confirming the mutation by amplifying the target sequence from the member(s) of step (k) using the primers of step (c) and determining the sequence of the amplified product.

2. A method for the detection of a mutation in a target sequence in a member of a mutagenized population, comprising the steps of:

IDT Exhibit 1020.6448

(a) isolating genomic DNA of each member of the mutagenized population to provide for DNA samples of each population member;
(b) pooling the DNA obtained in step (a); (c) amplifying part of the target sequence with a pair of tagged (optionally labeled) primers from the DNA pools, wherein at least one of the primers preferably comprises a gene-specific section, a tag and a sequence primer binding site;
(d) pooling the amplification products of step (c) to create a library of amplification products;
(e) determining the nucleotide sequence of the amplification products using high throughput sequenang;
(f) identifying mutations by clustering/aligning the sequences of the fragments; (g) identifying the member(s) having the mutation using the tags;
(h) optionally, confirming the mutation by amplifying the target sequence from the member(s) of step (g) using the primers of step (c) and determining the sequence of the amplified product.

3. The method according to claim 1 or 2, wherein the mutagenized population is obtained by treating the genome with one or more selected from the group consisting of mutation-inducing chemicals, ionizing radiation, targeted nucleotide exchange or region targeted mutagenesis.

4. A method for the detection of a mutation in a target sequence in a member of a population that comprises sub-populations that contain naturally occurring mutations, comprising the steps of:

(a) isolating genomic DNA of each member of the population to provide DNA samples of each population member,
(b) pooling the DNA obtained in step (a);
(c) amplifying the target sequence with a pair of (optionally labeled) primers from the DNA pools;
(d) pooling the amplification products of step (c) to create a library of amplification products;
(e) optionally, fragmenting the amplification products in the library;
(f) determining the nucleotide sequence of the amplified products or fragments using high throughput sequencing;
(g) identifying mutations by clustering/aligning the sequences of the fragments;
(h) screening the identified mutations for a modified function of the target sequence;
(i) designing a primer directed to hybridize to the identified mutation;
(j) amplifying the DNA pools of step (b) with the primer of step (i) and one of the primers of step (c);
(k) identify the population member(s) currying the mutation;
(l) optionally, confirming the mutation by amplifying the target sequence from the member(s) of step (k) using the primers of step (c) and determining the sequence of the amplified product.

5. A method for the detection of a mutation in a target sequence in a member of a population that comprises sub-populations that contain naturally occurring mutations, comprising the steps of:

(a) isolating genomic DNA of each member of the mutagenized population to provide for DNA samples of each population member;
(b) pooling the DNA obtained in step (a);
(c) amplifying part of the target sequence with a pair of tagged (optionally labeled) primers from the DNA pools, wherein at least one of the primers preferably comprises a gene-specific section, a tag and a sequence primer binding site;
(d) pooling the amplification products of step (c) to create a library of amplification products;
(e) determining the nucleotide sequence of the amplification products using high throughput sequencing;
(f) identifying mutations by clustering/aligning the sequences of the fragments;
(g) identifying the member(s) having the mutation using the tags;
(h) optionally, confirming the mutation by amplifying the target sequence from the member(s) of step (g) using the primers of step (c) and determining the sequence of the amplified product.

6. The method according to claims 1-5, wherein the pooling is a 3D-pooling strategy.

7. The method according to claims 1-6, wherein the high throughput sequencing is performed by synthesis, preferably pyrosequencing.

8. The method according to any of claims 1-7, wherein sequencing is performed on a solid support such as a bead.

9. The method according to any of claims 1-8, wherein the sequencing is based on High Throughput Sequencing, preferably Sequencing-by-Synthesis.

IDT Exhibit 1020.6449

10. The method according to any of claims 1-9, wherein the sequencing is performed by synthesis, preferably pyrosequencing.

11. The method according to claim 7, wherein sequencing comprises the steps of: (i) ligating sequencing adaptors to the fragments; (ii) annealing said adaptor-ligated fragments to beads, each bead annealing with a single fragment; (iii) emulsifying the beads in water-in-oil microreactors such that each microreactor contains a single bead; (iv) performing emulsion PCR to amplify adaptor-ligated fragments on the surface of the beads (v) selecting/enriching beads to which are attached said amplified adaptor-ligated fragments (vi) placing the beads in wells such that each well comprises a single bead; and (vii) generating a pyrophosphate signal.

12. The method according to any of claims 1-11, wherein the primers in steps (c) and/or step (i) contain nucleotides with improved binding affinity.

13. The method according to any of claims 2-12, wherein the part of the target sequence that is amplified in step (c) is from about 80 to about 400 bp.

14. The method according to claim 13, wherein the part of the target sequence that is amplified in step (c) is from about 90 to about 300 bp.

15. The method according to claim 13, wherein the part of the target sequence that is amplfied in step (c) is from 100 to about 200 bp.

Patentansprüche

1. Verfahren zur Feststellung einer Mutation in einer Zielsequenz in einem Teil einer mutagenisierten Population, das die folgenden Schritte umfasst:

   (a) Isolieren von genomischer DNA in jedem Teil der mutagenisierten Population zur Bereitstellung von DNA-Proben eines jeden Populationsteils,
   (b) Poolen der in Schritt (a) erhaltenen DNA,
   (c) Verstärken der Zielsequenz mit einem Paar (wahlweise etikettierter) Primer aus den DNA-Pools,
   (d) Poolen der Verstärkungsprodukte von Schritt (c) zur Schaffung einer Bibliothek von Verstärkungsprodukten,
   (e) wahlweise Fragmentieren der Verstärkungsprodukte in der Bibliothek,
   (f) Bestimmen der Nukleotidsequenz der verstärkten Produkte oder Fragmente unter Verwendung von Sequenzierung mit hohem Durchsatz,
   (g) Identifizieren von Mutationen durch Bündeln/Abgleichen der Sequenzen der Fragmente,
   (h) Screenen der identifizierten Mutationen auf eine modifizierte Funktion der Zielsequenz,
   (i) Ausbilden eines auf Hybridisation zu der identifizierten Mutation gerichteten Primers,
   (j) Verstärken der DNA-Pools von Schritt (b) mit dem Primer von Schritt (i) und einem der Primer von Schritt (c),
   (k) Identifizieren des/der die Mutation tragenden Populationsteils/-teile,
   (l) wahlweise Bestätigen der Mutation durch Verstärken der Zielsequenz aus dem Teil/den Teilen von Schritt (k) unter Verwendung der Primer von Schritt (c) und Bestimmen der Sequenz des verstärkten Produkts.

2. Verfahren zur Feststellung einer Mutation in einer Zielsequenz in einem Teil einer mutagenisierten Population, das die folgenden Schritte umfasst:

   (a) Isolieren von genomischer DNA eines jeden Teils der mutagenisierten Population zur Bereitstellung von DNA-Proben jedes Populationsteils,
   (b) Poolen der im Schritt (a) erhaltenen DNA,
   (c) Verstärken eines Teils der Zielsequenz mit einem Paar markierter (wahlweise etikettierter) Primer aus den DNA-Pools, wobei zumindest einer der Primer vorzugsweise einen genspezifischen Abschnitt, eine Markierung und einen Sequenzprimer-Bindeort umfasst,
   (d) Poolen der Verstärkungsprodukte von Schritt (c) zur Schaffung einer Bibliothek von Verstärkungsprodukten,
   (e) Bestimmen der Nukleotidsequenz der Verstärkungsprodukte unter Verwendung von Sequenzieren mit hohem Durchsatz,
   (f) Identifizieren von Mutationen durch Bündeln/Abgleichen der Sequenzen der Fragmente,
   (g) Identifizieren des Teils/der Teile mit der Mutation unter Verwendung der Markierungen,

IDT Exhibit 1020.6450

(h) wahlweise Bestätigen der Mutation durch Verstärken der Zielsequenz von dem Teil/den Teilen von Schritt (g) unter Verwendung der Primer von Schritt (c)und Bestimmen der Sequenz des verstärkten Produkts.

3. Verfahren nach Anspruch 1 oder 2, bei dem die mutagenisierte Population durch Behandeln des Genoms mit einer oder mehreren Auswahlen aus der Gruppe erreicht wird, die aus mutationsinduzierenden Chemikalien, ionisierender Strahlung, gezieltem Nukleotidaustausch oder regiongezielter Mutagenese besteht.

4. Verfahren zur Feststellung einer Mutation in einer Zielsequenz in einem Teil einer Population, die Sub-Populationen umfasst, die natürlicherweise vorkommende Mutationen enthalten, das die folgenden Schritte umfasst:

(a) Isolieren von genomischer DNA eines jeden Teils der Population zur Bereitstellung von DNA-Proben jedes Populationsteils,
(b) Poolen der in Schritt (a) erhaltenen DNA,
(c) Verstärken der Zielsequenz mit einem Paar (wahlweise etikettierter) Primer aus den DNA-Pools,
(d) Poolen der Verstärkungsprodukte von Schritt (c) zur Schaffung einer Bibliothek von Verstärkungsprodukten,
(e) wahlweise Fragmentieren der Verstärkungsprodukte in der Bibliothek,
(f) Bestimmen der Nukleotidsequenz der verstärkten Produkte oder Fragmente unter Verwendung von Sequenzieren mit hohem Durchsatz,
(g) Identifizieren von Mutationen durch Bündeln/Abgleichen der Sequenzen der Fragmente,
(h) Screenen der identifizierten Mutationen auf eine modifizierte Funktion der Zielsequenz,
(i) Ausbilden eines auf Hybridisation zu der identifizierten Mutation gerichteten Primers,
(j) Verstärken der DNA-Pools von Schritt (b) mit dem Primer von Schritt (i) und einem der Primer von Schritt (c),
(k) Identifizieren des/der die Mutation tragenden Teils/Teile,
(l) wahlweise Bestätigen der Mutation durch Verstärken der Zielsequenz von dem Teil/den Teilen von Schritt (k) unter Verwendung der Primer von Schritt (c) und Bestimmen der Sequenz des verstärkten Produkts.

5. Verfahren zur Feststellung einer Mutation in einer Zielsequenz in einem Teil einer Population, die Sub-Populationen umfasst, die natürlicherweise vorkommende Mutationen enthalten, das die folgenden Schritte umfasst:

(a) Isolieren von genomischer DNA jedes Teils der mutagenisierten Population zur Bereitstellung von DNA-Proben jedes Populationsteils,
(b) Poolen der in Schritt (a) erhaltenen DNA,
(c) Verstärken eines Teils der Zielsequenz mit einem Paar markierter (wahlweise etikettierter) Primer aus den DNA-Pools, wobei zumindest einer der Primer vorzugsweise einen genspezifischen Abschnitt, eine Markierung und einen Sequenzprimer-Bindeort umfasst,
(d) Poolen der Verstärkungsprodukte von Schritt (c) zur Schaffung einer Bibliothek von Verstärkungsprodukten,
(e) Bestimmen der Nukleotidsequenz der Verstärkungsprodukte unter Verwendung von Sequenzieren mit hohem Durchsatz,
(f) Identifizieren von Mutationen durch Bündeln/Abgleichen der Sequenzen der Fragmente,
(g) Identifizieren des Teils/der Teile mit der Mutation unter Verwendung der Markierungen,
(h) wahlweise Bestätigen der Mutation durch Verstärken der Zielsequenz von dem Teil/den Teilen von Schritt (g) unter Verwendung der Primer von Schritt (c) und Bestimmen der Sequenz des verstärkten Produkts.

6. Verfahren nach den Ansprüchen 1 bis 5, bei dem das Poolen eine 3D-Pooling-Strategie ist.

7. Verfahren nach den Ansprüchen 1 bis 6, bei dem das Sequenzieren mit hohem Durchsatz durch Synthese, vorzugsweise Pyrosequenzierung, durchgeführt wird.

8. Verfahren nach einem der Ansprüche 1 bis 7, bei dem das Sequenzieren auf einem festen Träger wie einer Perle durchgeführt wird.

9. Verfahren nach einem der Ansprüche 1 bis 8, bei dem das Sequenzieren auf Sequenzieren mit hohem Durchsatz, vorzugsweise Sequenzieren-durch-Synthese, basiert.

10. Verfahren nach einem der Ansprüche 1 bis 9, bei dem das Sequenzieren durch Synthese, vorzugsweise Pyrosequenzieren, durchgeführt wird.

11. Verfahren nach Anspruch 7, bei dem das Sequenzieren die folgenden Schritte umfasst:

IDT Exhibit 1020.6451

(i) Ligieren von Sequenzierungsadaptoren an die Fragmente, (ii) Absetzen der adapterligierten Fragmente auf Perlen, wobei sich ein einziges Fragment auf jeder Perle absetzt, (iii) Emulgieren der Perlen in Wasser-in-Öl-Mikroreaktionsräumen, derart, dass jeder Mikroreaktionsraum eine einzige Perle enthält, (iv) Durchführen einer PCR-Emulsion zum Verstärken adapterligierter Fragmente auf der Oberfläche der Perlen, (v) Selektieren/Anreichern von Perlen, an denen die verstärkten adapterligierten Fragmente anhaften, (vi) Einbringen der Perlen in Mulden, derart, dass jede Mulde eine einzelne Perle aufweist, und (vii) Erzeugen eines Pyrophosphatsignals.

12. Verfahren nach einem der Ansprüche 1 bis 11, bei dem die Primer in den Schritten (c) und/oder in Schritt (i) Nukleotide mit verbesserter Bindeaffinität enthalten.

13. Verfahren nach einem der Ansprüche 2 bis 12, bei dem der Teil der Zielsequenz, der in Schritt (c) verstärkt wird, etwa 80 bis etwa 400 bp beträgt.

14. Verfahren nach Anspruch 13, bei dem der Teil der Zielsequenz, der in Schritt (c) verstärkt wird, etwa 90 bis etwa 300 bp beträgt.

15. Verfahren nach Anspruch 13, bei dem der Teil der Zielsequenz, der in Schritt (c) verstärkt wird, 100 bis etwa 200 bp beträgt.

Revendications

1. Procédé de détection d'une mutation dans une séquence cible d'un élément d'une population mutagénisée, ledit procédé comprenant les étapes :

(a) d'isolation de l'ADN génomique de chaque élément de la population mutagénisée pour obtenir des échantillons d'ADN de chaque élément de population ;
(b) de collecte de l'ADN obtenu à l'étape (a) ;
(c) d'amplification de la séquence cible avec une paire d'amorces (optionnellement marquées) provenant des collectes d'ADN ;
(d) de collecte des produits de l'amplification de l'étape (c) pour créer une bibliothèque de produits d'amplification ;
(e) en option, de fragmentation des produits d'amplification de la bibliothèque ;
(f) de détermination de la séquence de nucléotides des produits ou fragments amplifiés en utilisant un séquençage à haut débit ;
(g) d'identification des mutations par groupement/alignement des séquences des fragments ;
(h) de criblage des mutations identifiées pour une fonction modifiée de la séquence cible ;
(i) d'élaboration d'une amorce destinée à s'hybrider à la mutation identifiée ;
(j) d'amplification des collectes d'ADN de l'étape (b) avec l'amorce de l'étape (i) et l'une des amorces de l'étape (c) ;
(k) d'identification de l'élément ou des éléments de population véhiculant la mutation ;
(l) en option, de confirmation de la mutation par amplification de la séquence cible provenant de l'élément ou des éléments de l'étape (k) en utilisant les amorces de l'étape (c) et en déterminant la séquence du produit amplifié.

2. Procédé de détection d'une mutation dans une séquence cible d'un élément d'une population mutagénisée, ledit procédé comprenant les étapes :

(a) d'isolation de l'ADN génomique de chaque élément de la population mutagénisée pour obtenir des échantillons d'ADN de chaque élément de population ;
(b) de collecte de l'ADN obtenu à l'étape (a) ; (c) d'amplification de la partie de la séquence cible contenant une paire d'amorces étiquetées (optionnellement marquées) provenant des collectes d'ADN, au moins une des amorces comprenant de préférence une section spécifique à un gène, une étiquette et un site de liaison d'une amorce de séquence ;
(d) de collecte des produits d'amplification de l'étape (c) pour créer une bibliothèque de produits d'amplification ;
(e) de détermination de la séquence de nucléotides des produits d'amplification en utilisant un séquençage à haut débit ;
(f) d'identification des mutations en groupant/alignant les séquences des fragments ;

IDT Exhibit 1020.6452

(g) d'identification de l'élément ou des éléments possédant la mutation en utilisant les étiquettes ;
(h) en option, de confirmation de la mutation par amplification de la séquence cible provenant de l'élément ou des éléments de l'étape (g) en utilisant les amorces de l'étape (c) et en déterminant la séquence du produit amplifié.

3. Procédé selon l'une des revendications 1 ou 2, dans lequel la population mutagénisée est obtenue en traitant le génome avec un ou plusieurs éléments du groupe constitué des produits chimiques induisant une mutation, d'un rayonnement ionisant, d'un échange ciblé de nucléotides ou d'une mutagenèse à région ciblée.

4. Procédé de détection d'une mutation dans une séquence cible d'un élément d'une population qui comprend des sous-populations qui contiennent des mutations survenant naturellement, ledit procédé comprenant les étapes :

(a) d'isolation de l'ADN génomique de chaque élément de la population pour obtenir des échantillons d'ADN de chaque élément population ;
(b) de collecte de l'ADN obtenu à l'étape (a) ;
(c) d'amplification de la séquence cible possédant une paire d'amorces (optionnellement marquées) provenant des collectes d'ADN ;
(d) de collecte des produits d'amplification de l'étape (c) pour créer une bibliothèque de produits d'amplification ;
(e) en option, de fragmentation des produits d'amplification de la bibliothèque ;
(f) de détermination de la séquence de nucléotides des fragments ou produits amplifiés en utilisant un séquençage à haut débit ;
(g) d'identification des mutations par groupement/alignement des séquences des fragments ;
(h) de criblage des mutations identifiées pour une fonction modifiée de la séquence cible ;
(i) d'élaboration d'une amorce destinée à s'hybrider à la mutation identifiée ;
(j) d'amplification des collectes d'ADN de l'étape (b) avec l'amorce de l'étape (i) et l'une des amorces de l'étape (c) ;
(k) d'identification de l'élément ou des éléments de population véhiculant la mutation ;
(l) en option, de confirmation de la mutation par amplification de la séquence cible provenant de l'élément ou des éléments de l'étape (k) en utilisant les amorces de l'étape (c) et en déterminant la séquence du produit amplifié.

5. Procédé de détection d'une mutation dans une séquence cible d'un élément d'une population qui comprend des sous-populations qui contiennent des mutations survenant naturellement, ledit procédé comprenant les étapes :

(a) d'isolation de l'ADN génomique de chaque élément de la population mutagénisée pour obtenir des échantillons d'ADN pour chaque élément de population ;
(b) de collecte de l'ADN obtenu à l'étape (a) ;
(c) d'amplification d'une partie de la séquence cible avec une paire d'amorces étiquetées (optionnellement marquées) provenant des collectes d'ADN, au moins une des amorces comprenant de préférence une section spécifique à un gène, une étiquette et un site de liaison d'une amorce de séquence ;
(d) de collecte des produits d'amplification de l'étape (c) pour créer une bibliothèque de produits d'amplification ;
(e) de détermination de la séquence de nucléotides des produits d'amplification en utilisant un séquençage à haut débit ;
(f) d'identification des mutations par groupement/alignement des séquences des fragments ;
(g) d'identification de l'élément ou des éléments possédant la mutation en utilisant les étiquettes ;
(h) en option, de confirmation de la mutation par amplification de la séquence cible provenant de l'élément ou des éléments de l'étape (g) en utilisant les amorces de l'étape (c) et en déterminant la séquence du produit amplifié.

6. Procédé selon les revendications 1 à 5, dans lequel la collecte est une stratégie de collecte 3D.

7. Procédé selon les revendications 1 à 6, dans lequel le séquençage à haut débit est réalisé par synthèse, de préférence par pyroséquençage.

8. Procédé selon l'une quelconque des revendications 1 à 7, dans lequel le séquençage est réalisé sur un support solide tel qu'une bille.

9. Procédé selon l'une quelconque des revendications 1 à 8, dans lequel le séquençage repose sur un séquençage

IDT Exhibit 1020.6453

à haut débit, de préférence un séquençage par synthèse.

10. Procédé selon l'une quelconque des revendications 1 à 9, dans lequel le séquençage est réalisé par synthèse, de préférence par pyroséquençage.

11. Procédé selon la revendication 7, dans lequel le séquençage comprend les étapes : (i) de ligature d'adaptateurs de séquençage aux fragments ; (ii) d'annelage des fragments ligaturés à des adaptateurs à des billes, chaque bille s'annelant à un seul fragment ; (iii) d'émulsion des billes dans des microréacteurs « eau dans huile » de sorte que chaque microréacteur contient une seule bille ; (iv) de réalisation d'une PCR par émulsion pour amplifier les fragments ligaturés à des adaptateurs sur la surface des billes ; (v) de sélection/d'enrichissement des billes auxquelles sont attachés lesdits fragments amplifiés ligaturés à des adaptateurs ; (vi) de placement des billes dans des puits de sorte que chaque puits comprend une seule bille ; et (vii) de génération d'un signal pyrophosphate.

12. Procédé selon l'une quelconque des revendications 1 à 11, dans lequel les amorces des étapes (c) et/ou (i) contiennent des nucléotides présentant une affinité de liaison améliorée.

13. Procédé selon l'une quelconque des revendications 2 à 12, dans lequel la partie de la séquence cible qui est amplifiée à l'étape (c) contient environ 80 à environ 400 paires de bases.

14. Procédé selon la revendication 13, dans lequel la partie de la séquence cible qui est amplifiée à l'étape (c) contient environ 90 à environ 300 paires de bases.

15. Procédé selon la revendication 13, dans lequel la partie de la séquence cible qui est amplifiée à l'étape (c) contient 100 à environ 200 paires de bases.

IDT Exhibit 1020.6454