[0062] Ligating embodiments may be further defined as comprising: generating ligatable ends on the double stranded nucleic acid molecule; generating a ligatable end on the stem-loop oligonucleotide; and ligating one strand of the ligatable end of the stem-loop oligonucleotide to one strand of an end of the nucleic acid molecule, thereby generating a non-covalent junction, such as a nick, a gap, or a 5' flap structure, in the oligonucleotide-attached nucleic acid molecule. In further aspects, the methods comprise generating blunt ends on the nucleic acid molecule; generating a blunt end on the stem-loop oligonucleotide; and ligating one strand of the blunt end of the stem-loop oligonucleotide to one strand of a blunt end of the nucleic acid molecule, thereby generating a nick in the oligonucleotide-ligated nucleic acid molecule.

[0063] In specific aspect of the invention, the stem-loop oligonucleotide comprises a known sequence, for example a regulatory sequence, including a RNA polymerase promoter sequence. A regulatory sequence may reside at least in part within the stem of the stem-loop oligonucleotide, within the loop of the stem-loop oligonucleotide, or both.

[0064] Methods of the invention may further comprise digesting the DNA molecule with one or more endonucleases to produce DNA fragments; producing blunt ends on the DNA fragments; producing a blunt end on the stem-loop oligonucleotide; and ligating one strand of the blunt end of a stem-loop oligonucleotide to one strand of a blunt end of a DNA fragment, thereby generating a nick in an oligonucleotide-ligated DNA fragment. In a specific embodiment, the endonuclease is a restriction endonuclease, DNAse I, or an apoptotic endonuclease, or a mixture thereof. In particular embodiments, the restriction endonuclease is methylation-specific or methylation-sensitive.

[0065] In additional embodiments, the oligonucleotide-attached nucleic acid molecule comprises a nick having a 3' hydroxy group, wherein there is polymerization from the 3' hydroxy group of at least part of the oligonucleotide-attached nucleic acid molecule.

[0066] Strand displacement or nick translation polymerization may be further defined as polymerization that ceases at a non-replicable base or region in the loop or in a region of the stem adjacent to the loop.

[0067] In a specific aspect of the invention, the method further comprises the step of digesting the double stranded DNA molecule with an endonuclease to generate DNA fragments, wherein the oligonucleotide becomes ligated to one strand of the DNA fragment and

IDT Exhibit 1020.7554

wherein polymerization of an oligonucleotide-ligated DNA fragment excludes at least part of the inverted repeat by subjecting the oligonucleotide-ligated DNA fragment to strand displacement or nick translation polymerization that halts at a base or sequence in the loop or in a region of the stem adjacent to the loop.

[0068] In some embodiments, the stem-loop oligonucleotide is further defined as comprising a non-replicable base or sequence. In particular, in some cases at least part of the non-replicable base or sequence is present in the loop of the oligonucleotide or in a sequence of the stem adjacent to the loop. The non-replicable base or sequence may comprise an abasic site or sequence, hexaethylene glycol, and/or a bulky chemical moiety attached to the sugar-phosphate backbone or the base. In specific embodiments, the abasic site or sequence is introduced by one or more enzymes in the single solution. In a further specific embodiment, the loop of the stem-loop oligonucleotide comprises at least one deoxy-uridine.

[0069] In one particular embodiment, strand displacement or nick translation polymerization of the oligonucleotide-attached nucleic acid molecule generates an endonuclease site. In particular aspects, the methods further comprise the step of digesting the double stranded DNA molecule with an endonuclease to generate DNA fragments, wherein the oligonucleotide becomes blunt end ligated to one strand of the DNA fragment to produce an oligonucleotide-ligated DNA fragment, and wherein strand displacement or nick translation polymerization of the oligonucleotide-ligated DNA fragment generates an endonuclease site.

[0070] In particular aspects, the endonuclease site is a site-specific restriction endonuclease site and at least part of the inverted repeat is removed by cleavage with said restriction endonuclease. The endonuclease site may reside within the stem and/or loop regions. Examples of endonucleases include Eco NI or PacI. The endonuclease site may be a homing endonuclease site, such as I-Ceu I, I-Sce I, PI-Psp I, or PI-Sce I. In specific embodiments, RNA transcription is initiated from the regulatory sequence, thereby producing at least one transcribed polynucleotide.

[0071] In specific embodiments, a regulatory sequence resides at least in part within the stem of the stem-loop oligonucleotide, or at least in part within the loop of the stem-loop oligonucleotide. In a specific aspect, the regulatory sequence comprises a bacteriophage regulatory sequence, such as a T7 regulatory sequence, a T3 regulatory sequence, or a Sp6 regulatory sequence, for example. In specific aspects, a transcribed polynucleotide is replicated

IDT Exhibit 1020.7555

by a reverse transcriptase, which may be initiated by hybridization of the oligonucleotide complementary to the 3' end of the transcribed polynucleotide.

[0072] The stem-loop oligonucleotide may further comprise a recognition sequence for a nicking endonuclease and the oligonucleotide-attached nucleic acid molecule is amplified by strand displacement synthesis and second strand synthesis, in particular aspects of the invention. In a specific aspect, the recognition sequence for a nicking endonuclease resides at least in part within the stem of the stem-loop oligonucleotide. The recognition sequence for a nicking endonuclease may reside at least in part within the loop of the stem-loop oligonucleotide. In specific embodiments, the nicking endonuclease is N.Alw I, N.Bbv CIA, N.Bbv CIB, Nb.Bpu10I, Nb.BsmI, N.BstNBI, or N.Bst 9I.

[0073] In specific aspects, a 5' end of the stem-loop oligonucleotide lacks a phosphate.

[0074] In particular embodiments, the oligonucleotide-attached nucleic acid molecule is further modified, such as by cloning, including by incorporation of the modified molecule into a vector, said incorporation occuring at ends in the modified molecule generated by endonuclease cleavage within the inverted repeat.

[0075] In additional embodiments of the invention, methods of the invention occur in a single suitable solution, wherein the process occurs in the absence of exogenous manipulation. The method may occur at one temperature, such as from about 10°C to about 75°C. . In particular embodiments, the solution comprises one or more of the following: ligase; DNA polymerase; one or more endonucleases; RNA polymerase; reverse transcriptase; RNase H; deoxy-uridine glycosylase (which is also referred to as Uracyl-DNA Glycosylase or UDG); nickase; thermophilic DNA polymerase; ATP; rNTPs; and dNTPs.

[0076] In one aspect of the invention, there is a kit, housed in a suitable container, comprising: a stem-loop oligonucleotide; and a solution suitable for ligation of the oligonucleotide onto a double stranded molecule; and, optionally, one or more of the following: ligase; DNA polymerase one or more restriction enzymes; RNA polymerase; reverse transcriptase; RNase H; nickase; thermophilic DNA polymerase; ATP; rNTPs; and dNTPs. In specific embodiments, the oligonucleotide of the kit comprises a RNA polymerase promoter sequence, such as a T7 RNA polymerase promoter sequence, for example. The one or more

IDT Exhibit 1020.7556

endonucleases may be methylation-specific, methylation-sensitive, or may be a homing endonuclease.

[0077] In specific embodiments, methods of the invention occur in a single suitable solution, wherein the process of preparing and amplifying of a oligonucleotide-attached nucleic acid molecule occurs in the absence of exogenous manipulation. The method may be further defined as occurring at one temperature, such as between about 10°C to about 85°C or between about 10°C to about 85°C. In specific embodiments, the method is further defined as occurring at least two different temperatures, such as wherein at least one of the temperatures is from about 10°C to about 100°C.

[0078] In specific embodiments, the oligonucleotide-attached nucleic acid molecule is immobilized to a solid support, such as non-covalently or covalently.

[0079] Some methods of the invention further comprise the step of digesting the DNA molecule with an endonuclease to generate DNA fragments, wherein the oligonucleotide becomes attached to one strand of the DNA fragment, wherein strand displacement polymerization of the oligonucleotide-ligated DNA fragment and its arrest at a base or sequence in the loop or in a region of the stem adjacent to the loop generates a 5' overhang, and wherein the single stranded 5' overhang hybridizes to a complementary oligonucleotide covalently immobilized to a solid support.

[0080] Additional embodiments of the invention include a library of DNA molecules prepared by the methods of the invention.

[0081] In some embodiments of the invention, there is a method of preparing a DNA molecule, comprising: providing an oligonucleotide; and mixing the oligonucleotide with a double stranded DNA molecule, such that upon attachment of the oligonucleotide to the double stranded DNA molecule, the oligonucleotide attached to the DNA molecule is capable of demonstrating a function that it was incapable of before the attachment to the double stranded DNA molecule, thereby producing an oligonucleotide-attached DNA molecule suitable for modification. In specific aspects, the method occurs in a single suitable solution, wherein the process occurs in the absence of exogenous manipulation. Exemplary modifications include site-specific nicking, site-specific double-strand cleavage, transcription, recombination, amplification, polymerization, nick translation, strand displacement, immobilization or a combination thereof.

IDT Exhibit 1020.7557

[0082] In an additional embodiment of the invention, there is a composition comprising a stem-loop oligonucleotide, said oligonucleotide comprising an inverted repeat, a loop and at least one thermo-sensitive site or site capable of becoming a thermo-sensitive site, wherein the oligonucleotide is capable of producing a suitable primer upon exposure to heat. In additional embodiments, the composition further comprises a repair enzyme, such as one capable of converting damaged bases into an abasic site. Examples of DNA repair enzymes include a DNA glycosylase, such as dU-glycosylase, for example. In a specific aspect, the thermo-sensitive site is an abasic site, such as an apyrimidine site or an apurine site, for example. In an additional specific embodiment, the site capable of becoming a thermo-sensitive site is further defined as a base that can be enzymatically converted to a thermo-sensitive site. The base that can be enzymatically converted to a thermo-sensitive site may be deoxy-uridine and the enzyme may be dU-glycosylase (which is also referred to as Uracyl-DNA Glycosylase or UDG). In a specific aspect, the thermo-sensitive site is introduced into the oligonucleotide during synthesis, such as chemical synthesis. The composition may be capable of comprising a breakage of a strand of the oligonucleotide at one or more deoxy-uridine bases upon exposure to heat. The thermosensitive site or site capable of becoming a thermo-sensitive site may be located in the stem, in the loop, or both. The oligonucleotide may be blunt ended or have a 5' overhang.

[0083] In an additional embodiment, there is a method of producing a primer, comprising providing a composition of the invention and subjecting the composition to heat such that the primer is produced. In a specific aspect, the composition further comprises a DNA repair enzyme and the subjecting to heat is further defined as subjecting to two or more exposures of heat. One exposure to heat may comprise about 37°C and another exposure to heat may comprise about 95°C. One exposure to heat may comprise about 10°C to about 100°C whereas another exposure to heat may comprise about 80°C to about 100°C. The method may further comprise amplification of a DNA molecule utilizing said composition.

[0084] The foregoing has outlined rather broadly the features and technical advantages of the present invention in order that the detailed description of the invention that follows may be better understood. Additional features and advantages of the invention will be described hereinafter which form the subject of the claims of the invention. It should be appreciated by those skilled in the art that the conception and specific embodiment disclosed may be readily utilized as a basis for modifying or designing other structures for carrying out the same purposes of the present invention. It should also be realized by those skilled in the art that

IDT Exhibit 1020.7558

such equivalent constructions do not depart from the spirit and scope of the invention as set forth in the appended claims. The novel features that are believed to be characteristic of the invention, both as to its organization and method of operation, together with further objects and advantages will be better understood from the following description when considered in connection with the accompanying figures. It is to be expressly understood, however, that each of the figures is provided for the purpose of illustration and description only and is not intended as a definition of the limits of the present invention.

BRIEF DESCRIPTION OF THE DRAWINGS

[0085] For a more complete understanding of the present invention, reference is now made to the following descriptions taken in conjunction with the accompanying drawings.

[0086] FIG. 1A shows a standard multi-step DNA adaptor attachment process.

[0087] FIG. 1B compares preparation processes for WGA and WMA libraries from degraded serum and urine DNA described in the present invention and in previously filed patent applications (U.S. Patent Application Serial No. 11/071,864, filed March 3, 2005).

[0088] FIG. 1C is a general representation of the invention that can be described as a reduction of several important multi-step enzymatic DNA processes to a single-step multiplexed enzymatic reaction that is performed in one reaction volume (Enz-O-Mix Method).

[0089] FIG. 2 shows that the Enz-O-Mix approach can be viewed as a simple and cost-effective alternative to the "lab-on-a chip" microfluidic approach that is currently attempting to solve the same problem by reduction of reaction volumes and integration of multiple reactions into a small format.

[0090] FIG. 3 is a schematic description and composition of specific components of the Enz-O-Mix method and reagents involved in the attachment of stem-loop oligonucleotides to DNA ends. Four enzymatic reactions are taking place nearly simultaneously: "polishing" of the DNA ends and the hairpin double-stranded stem-region; ligation of the stem-loop oligonucleotide 3' end to the 5' phosphate of the DNA, leaving a nick between the 3' end of DNA and the 5' end of the hairpin double-stranded stem-region; polymerase extension of the 3' DNA end that propagates toward the end of stem-loop oligonucleotide; and by strand-displacement reaction within the oligonucleotide stem region. The process results in the library of DNA fragments with inverted repeat sequences at their ends.

IDT Exhibit 1020.7559

[0091] FIG. 4 shows three original secondary structures of the stem-loop oligonucleotide with the 3' or 5' protruding, or blunt end, and its final (blunt end) structure within the Enz-O-Mix.

[0092] FIGS. 5A-5B are schematic descriptions of the amplification of a library of DNA fragments with inverted repeat sequences at their ends by RNA transcription. Promoter sequence can be added to a DNA end either as a part of the oligonucleotide stem-region (A), or as a part of the oligonucleotide loop region (B). The products of the amplification are RNA molecules in this embodiment.

[0093] FIG. 6A illustrates the components and enzymatic reactions involved in the one-step DNA amplification by transcription using Enz-O-Mix reagent (Enz-O-Mix 5 in FIG. 19B, for example). The process linearly amplifies DNA and produces single stranded RNA molecules.

[0094] FIG. 6B illustrates additional components and enzymatic reactions involved in the one-step DNA amplification by transcription accompanied by the synthesis of complementary cDNA strand. The transcription (replication) block (see Sections E-H and FIGS. 8-10) is introduced into the stem-loop oligonucleotide upstream of the T7 promoter region to prevent formation of an inverted repeat at the 3' end of RNA molecules and thus increase the efficiency of the oligo T7 priming (the transcription block is shown as a black square). The process can linearly amplify DNA and produce double stranded DNA/RNA hybrids.

[0095] FIG. 7 is an illustration of the inhibitory effect on PCR of inverted repeats attached to both ends of DNA fragments. Heating and replication generates DNA molecules refractory to melting, priming, and PCR amplification either using the universal primer A or any internal site-specific primer pair.

[0096] FIG. 8 is a schematic description of the one-step Enz-O-Mix attachment process for a stem-loop oligonucleotide with non-replicable linker. The following enzymatic reactions are taking place nearly simultaneously: "polishing" of the DNA ends and the stem-loop oligonucleotide double-stranded stem-region; ligation of the oligonucleotide 3' end to the 5' phosphate of the DNA, leaving a nick between the 3' end of DNA and the 5' end of the hairpin double-stranded stem-region; polymerase extension of the 3' DNA end that propagates toward the end of oligonucleotide and stops at the replication block (non-replicable linker) within the loop or outside the loop but no more than about six bases away from the loop. The process

IDT Exhibit 1020.7560

results in the library of DNA fragments with universal sequence A at the ends and an inverted repeat attached only to the 5' end of DNA.

[0097] FIG. 9 is an illustration of the absence of an inhibitory effect on PCR of inverted repeats attached to only 5' ends of DNA fragments. Heating and replication generates DNA molecules with the universal sequence A and no inverted repeats at the ends (shown by dash lines) that can be successfully amplified by PCR.

[0098] FIG. 10A shows the structure of a stem-loop oligonucleotide with non-replicable linker introduced chemically during oligonucleotide synthesis and detailed events occurring at a DNA end during the multi-enzyme attachment process. The process occurs as described in FIG. 8.

[0099] FIG. 10B shows the structure of a stem-loop oligonucleotide with non-replicable linker introduced enzymatically during the attachment reaction and detailed events occurring at a DNA end during the multi-enzyme attachment process. Specifically, dU-glycosylase generates abasic sites within the loop and upper strand of the stem region of the stem-loop oligonucleotide. The abasic sites within a loop generate a non-replicable region, while the sites in a stem destabilize the duplex and facilitate strand displacement reaction. The process results in a library of DNA fragments with universal sequence at the ends and an inverted repeat attached only to the 5' end of DNA. Heating at 95°C during PCR generates breaks at abasic sites and completely eliminates the 5' portion of a stem-loop oligonucleotide.

[0100] FIG. 11A shows a standard adaptor formed by two different oligonucleotides, wherein one of them has a 5' phosphate group, and possible products formed when such adaptor is used in the multi-enzyme one-step ligation process.

[0101] FIG. 11B shows a standard adaptor formed by two different oligonucleotides without phosphate group and two products formed when such an adaptor is used in the multi-enzyme one-step ligation process. Only 50% of formed molecules are desirable products with correct orientation of the adaptor, whereas the other 50% of molecules have inverse adaptor orientation.

[0102] FIG. 11C shows a standard adaptor formed by two different oligonucleotides without a phosphate group, a protective group at the 3' end of one oligonucleotide, and products formed when such an adaptor is used in the multi-enzyme one-step

IDT Exhibit 1020.7561

ligation process. As in the case described on FIG. 11B, only 50% of formed molecules are desirable products with correct orientation of the adaptor, whereas the other 50% of molecules have inversed adaptor orientation.

[0103] FIG. 12 is a schematic description of the one-step Enz-O-Mix attachment process for a stem-loop oligonucleotide with an enzymatically cleavable site generated during the multi-enzyme reaction. Specifically, the endonuclease recognizes and cuts the oligonucleotide specific oligo-sequence R when it adopts a double-stranded conformation as a result of the attachment process. The reaction results in a library of DNA fragments with universal sequence A at the ends and no inverted repeat attached to the ends of DNA fragments.

[0104] FIG. 13A shows the stem-loop oligonucleotide with the recognition sequence for the restriction endonuclease Eco NI located at the loop region. The enzyme can not cut the CCTNNNNNAGG (SEQ ID NO: 20) region within the stem-loop oligonucleotide (not a recognizable structure) but it can cut it efficiently when the oligonucleotide is attached to a DNA end and adopts a canonical (double-stranded) Watson-Crick conformation.

[0105] FIG. 13B shows the stem-loop oligonucleotide with the recognition sequence for the restriction endonuclease PacI located within the loop region. The enzyme can not cut the single stranded TTAATTAA region within the stem-loop oligonucleotide, but it cuts it efficiently when the oligonucleotide is attached to a DNA end and adopts a canonical (double-strand) Watson-Crick conformation.

[0106] FIG. 13C shows the stem-loop oligonucleotide with the recognition sequence for the exemplary homing endonuclease I-Ceu I located within the loop region. The enzyme can not cut the 26-base single-stranded region within the hairpin loop, but it cuts efficiently when the oligonucleotide is attached to a DNA end and adopts a canonical (double-stranded) Watson-Crick conformation.

[0107] FIG. 14 is a schematic description of the one-step Enz-O-Mix attachment process for a stem-loop oligonucleotide with a very short stem. The process results in the library of DNA fragments with universal sequence A at the ends and very short inverted repeat attached to the ends of DNA fragments. Amplification by PCR is possible, because the melting temperature for terminal hairpins is low enough to prevent folding, self-priming and formation of long hairpin molecules, as shown on FIG. 7.

IDT Exhibit 1020.7562

[0108] FIG. 15 illustrates the components and enzymatic reactions involved in exemplary one-step or two-step (after library synthesis) DNA amplification by strand displacement DNA synthesis using Enz-O-Mix reagents (Enz-O-Mix 3 and Enz-O-Mix 6 in FIGS. 19A and 19B, for example). The amplification is initiated and further maintained by the generation of a nick within the attached oligonucleotide sequence. The process linearly amplifies DNA and produces double stranded DNA molecules.

[0109] FIG. 16A shows the stem-loop oligonucleotide with the recognition sequence for the nicking endonuclease N. Bbv CIA located at the loop region. The enzyme can not cut the single stranded GCTGAGG region within the stem-loop oligonucleotide (not a recognizable structure) but it can nick it efficiently when the oligonucleotide is attached to a DNA end and adopts a canonical (double-stranded) Watson-Crick conformation.

[0110] FIG. 16B shows a stem-loop oligonucleotide with the recognition sequence for the nicking endonuclease N. Bbv CIA located at the end of a stem region. The enzyme can not nick efficiently the GCTGAGG region within the stem-loop oligonucleotide (it is too close to the end) but it can nick it efficiently when the oligonucleotide is attached to a DNA end; the second site generated by attachment is not a good substrate for nicking (it is too close to the end).

[0111] FIG. 16C shows two exemplary multifunctional stem-loop oligonucleotides with several regulatory elements within a single (1) or multiple (2) loop regions: T7 promoter sequence (that includes all regions necessary for transcription), recognition sequence N for a nicking endonuclease, recognition sequence for a restriction endonuclease, and a replication/transcription block region (a chemical modification or a string of 2-4 dU bases). A multifunctional stem loop oligonucleotide can support different types of DNA amplification, including PCR, isothermal amplification by transcription, isothermal amplification by strand displacement synthesis, or even a combination thereof and it can also facilitate other modifications, including DNA immobilization, DNA recombination, DNA cloning, and a combination thereof, for example.

[0112] FIG. 17 shows generalization of the Enz-O-Mix method and demonstrates specific but exemplary applications.

[0113] FIG. 18 shows composition of five exemplary different Stem-Loop Oligo Attachment Master Mixes of the invention that are used for the one-step nucleic acid library synthesis.

IDT Exhibit 1020.7563

[0114] FIG. 19A schematically shows compositions of several Enz-O-Mix reagents, specifically, Enz-O-Mix 1, designed to convert fragmented DNA into a Whole Genome Library that can be amplified by PCR in an open, or close-tube format.. Enz-O-Mix 2 and Enz-O-Mix 3 are designed to convert fragmented DNA into a Whole Genome Library that can be amplified after or during the Whole Genome Library synthesis by transcription and strand displacement, respectively.

[0115] FIG. 19B shows schematically compositions of several Enz-O-Mix reagents, specifically, Enz-O-Mix 4, designed to convert HMW DNA into a Whole Genome Library that can be amplified by PCR in an open, or close-tube format.. Enz-O-Mix 5 and Enz-O-Mix 6 are designed to convert fragmented DNA into a Whole Genome Library that can be amplified after or during the Whole Genome Library synthesis by transcription and strand displacement, respectively.

[0116] FIG. 19C shows schematically compositions of several Enz-O-Mix reagents, specifically, Enz-O-Mix 7, designed to convert fragmented DNA into a Whole Methylome Library that can be amplified by PCR.. Enz-O-Mix 8 and Enz-O-Mix 9 are designed to convert fragmented DNA into a Whole Methylome Library that can be amplified after or during the Whole Methylome Library synthesis by transcription and strand displacement, respectively.

[0117] FIG. 19D shows schematically compositions of several Enz-O-Mix reagents, specifically, Enz-O-Mix 10, designed to convert HMW DNA into a Whole Methylome Library that can be amplified by PCR. Enz-O-Mix 11 and Enz-O-Mix 12 are designed to convert fragmented DNA into a Whole Methylome Library that can be amplified after or during the Whole Methylome Library synthesis by transcription and strand displacement, respectively.

[0118] FIG. 19E shows schematically compositions of Enz-O-Mix reagents 13 and 14 designed to convert fragmented and HMW DNA, respectively, into a Whole Genome Library that can be directly cloned into an appropriate vector.

[0119] FIG. 20 illustrates assembly, storage and use of Enz-O-Mix reagent(s) as a kit in high throughput applications and clinical diagnostics, for example.

[0120] FIG. 21A demonstrates the optimization process for Enz-O-Mix Universal Buffer using real-time PCR to monitor amplification of the generated library.

IDT Exhibit 1020.7564

[0121] FIG. 21B demonstrates the optimization process for Enz-O-Mix incubation temperature using real-time PCR to monitor amplification of the generated library.

[0122] FIG. 21C demonstrates the optimization process for Enz-O-Mix reagent (N) concentration using real-time PCR to monitor amplification of the generated library.

[0123] FIG. 22 shows real-time amplification curves of whole genome libraries prepared in a single step by ligation of a stem-loop oligonucleotide with EcoNI resrtiction site generated after ligation and strand-displacement extension of blunt-end AluI restriction fragments in the presence or in the absence of EcoNI. Excision of the terminal inverted repeat by EcoNI results in high amplification efficiency that exceeds amplification efficiency of the library with inverted repeats (in the absence of the Eco NI cleavage) by two orders of magnitude.

[0124] FIG. 23 shows real-time amplification curves of libraries prepared in a single step by ligation of blunt-end AluI restriction fragments to the stem-loop oligonucleotide comprising hexa-ethylene glycol replication stop (HEG Stem-Loop Oligo), a stretch of 5 ribonucleosides (Ribo Stem-Loop Oligo), or just deoxyribonucleosides (Stem-Loop Oligo) in their loop regions. Unlike the libraries made with HEG-containing oligonucleotide, the libraries with ribonucleosides in the loop region are not amplified efficiently, suggesting that ribonucleosides in the template strand do not arrest DNA replication carried by the Klenow fragment of DNA polymerase I.

[0125] FIGS. 24A, 24B, and 24C show the real-time PCR amplification curves of 3 exemplary human STS markers used for evaluation of genomic sequence representation in WGA libraries prepared in a single step by ligation of blunt-end AluI restriction fragments to a stem-loop oligonucleotide containing within a loop either the hexa-ethylene glycol linker (HEG) as a replication stop, or the EcoNI recognition sequence that becomes functional (cleavable) after ligation and strand-displacement extension. Twenty nanograms of purified material of each library is compared to 1 ng of human genomic DNA randomly fragmented to an average size of 1.5 Kb using Hydro Shear device (gDNA). In FIG. 24A, there is PCR using STS # 4 primers. In FIG. 24B, there is PCR using STS # 19 primers. In FIG. 24C, there is PCR using STS # 35 primers (Table II). Results show little or no representational bias for all 3 STS sequences.

[0126] FIG. 25 shows the real-time amplification curves of libraries prepared in a single step by ligation of blunt-end AluI restriction fragments to the stem-loop oligonucleotide containing deoxy-uridine in the 5' stem region and in the loop with T4 DNA ligase and Klenow

IDT Exhibit 1020.7565

fragment of DNA polymerase I, in the presence or in the absence Uracil-DNA glycosylase (which is also referred to as UDG or dU glycosylase). Amplification is conducted by either adding universal primer M_U or using the primer released as a result of the UDG activity, followed by heat-induced degradation of the stem-loop oligonucleotide.

[0127] FIG. 26 illustrates a one-step genomic DNA restriction digestion and whole genome amplification using degradable stem-loop oligonucleotide containing deoxy-uridine in the 5' stem region and in the loop. Human genomic DNA in amounts ranging from 20 ng to 10 pg or a blank sample (no DNA control) are incubated with the stem-loop oligonucleotide in the presence of 4 exemplary enzymatic activities: AluI restriction endonuclease, T4 DNA ligase, Klenow fragment of DNA polymerase I, and Uracil-DNA glycosylase (UDG) for 1 hour at 37°C and amplified by PCR using universal primer M_U.

[0128] FIG. 27 shows PCR amplification curves of specific promoter sites from amplified libraries prepared from non-methylated (N) and artificially methylated (M) cell-free urine DNA from a healthy donor. Libraries were prepared by ligation of a degradable stem-loop oligonucleotide containing deoxy-uridine with (N+, M+) or without (N-, M-) simultaneous cleavage with methylation-sensitive restriction enzymes. Promoter sites from non-methylated cleaved DNA (N+) amplify with significant (at least 10 cycles) delay as compared to uncut DNA (N-, M-) unlike methylated DNA which is refractory to cleavage (M+ versus M-).

[0129] FIG. 28 shows the products of simultaneous whole bacteriophage DNA library synthesis and isothermal amplification by T7 polymerase transcription of lambda DNA digested with restriction endonuclease Bst EII on an agarose gel in the presence (+T4 ligase) or absence of DNA ligase (-T4 ligase).

[0130] FIG. 29 shows the products of simultaneous library synthesis and isothermal amplification of cell-free urine DNA using the nicking activity of the nuclease N.BbvC IB and strand displacement DNA synthesis.

[0131] FIGS. 30A-30B compare two Enz-O-Mix library synthesis / PCR amplification processes (Enz-O-Mix 1 and Enz-O-Mix 4 from FIGS. 19A and 19B). FIG. 30A shows a 2-step, opened-tube protocol. In this case, Enz-O-Mix library synthesis / amplification is performed in two operational steps: step 1- a tube contatining DNA is supplemented with the library (WGA or WMA) synthesis reagents and incubated at 37 oC for 1 h, and step 2 – the tube is opened, supplemented with PCR amplification buffer / reagents and subjected to temperature

IDT Exhibit 1020.7566

cycling. FIG. 30B shows a 1-step, closed-tube protocol. In this case, Enz-O-Mix library synthesis / amplification is performed in one operational step when a tube containing DNA is supplemented with the library synthesis / amplification (PCR) reagents and universal buffer, and subjected to programmed temperature conditions within a thermocycler.

[0132] FIG. 31 shows an expected accumulation of DNA in a tube (a) and a typical temperature profile (b) for a one-step, closed tube Enz-O-Mix PCR-based WGA or WMA.

[0133] FIG. 32 illustrates heat-induced conversion of a stem-loop oligonucleotide adaptor containing dU-bases within the loop and the 5' stem region into a functional universal PCR primer.

[0134] FIG. 33 shows effect of DMSO and magnesium concentration on the whole genome amplification of HMW human DNA isolated from blood using integrated one-step, closed-tube Enz-O-Mix protocol illustrated in FIG. 30, 31, and 32, and real-time PCR detection.

[0135] FIG. 34 illustrate application of the one-step, closed tube Enz-O-Mix WGA process (DNA fragmentation, amplification, and labeling) for micro-array analysis.

[0136] FIGS. 35A and 35B show the one-step process for simultaneous GenomePlex or MethylPlex library synthesis and immobilization. In FIG. 35A, there is non-covalent library immobilization by hybridization of the exposed 5' stem region to an oligo S covalently attached to a tube/well surface. In FIG. 35B, there is a covalent library immobilization by hybridization and ligation of the exposed 5' stem region to an oligo S covalently attached to a tube/well surface.

[0137] FIG. 36 shows tubes, plates, micro-slides, and micro-beads manufactured for one-step GenomePlex or MethylPlex library synthesis and immobilization.

[0138] FIG. 37 shows a hypothetical fluidic device that utilizes the one-step GenomePlex or MethylPlex library synthesis and immobilization process.

[0139] FIG. 38 illustrates hot start locus-specific PCR amplification of DNA using degradable stem-loop primers.

IDT Exhibit 1020.7567

[0140] FIG. 39 shows the specific events that lead to a conversion of the inactive stem-loop oligonucleotides into the active PCR primers.

[0141] FIG. 40 shows accumulation of a specific PCR product (A) and a typical temperature profile (B) for a hot start PCR using degradable stem-loop primers

DETAILED DESCRIPTION OF THE INVENTION

[0142] The present application incorporates by reference herein in its entirety U.S. Patent Application Serial No. 11/071,864, filed March 3, 2005. Also incorporated by reference herein in its entirety is U.S. Patent Application Serial No. unknown, filed March 2, 2006, entitled "Isolation of CpG Islands by Thermal Segregation and Enzymatic Selection-Amplification Method," which itself claims priority to U.S. Provisional Patent Application Serial No. 60/704,541, filed August 2, 2005.

I. Definitions

[0143] In keeping with long-standing patent law convention, the words "a" and "an" when used in the present specification in concert with the word comprising, including the claims, denote "one or more." Some embodiments of the invention may consist of or consist essentially of one or more elements, method steps, and/or methods of the invention. It is contemplated that any method or composition described herein can be implemented with respect to any other method or composition described herein.

[0144] The practice of the present invention will employ, unless otherwise indicated, conventional techniques of molecular biology, microbiology, recombinant DNA, and so forth which are within the skill of the art. Such techniques are explained fully in the literature. See e.g., Sambrook, Fritsch, and Maniatis, MOLECULAR CLONING: A LABORATORY MANUAL, Second Edition (1989), OLIGONUCLEOTIDE SYNTHESIS (M. J. Gait Ed., 1984), ANIMAL CELL CULTURE (R. I. Freshney, Ed., 1987), the series METHODS IN ENZYMOLOGY (Academic Press, Inc.); GENE TRANSFER VECTORS FOR MAMMALIAN CELLS (J. M. Miller and M. P. Calos eds. 1987), HANDBOOK OF EXPERIMENTAL IMMUNOLOGY, (D. M. Weir and C. C. Blackwell, Eds.), CURRENT PROTOCOLS IN MOLECULAR BIOLOGY (F. M. Ausubel, R. Brent, R. E. Kingston, D. D. Moore, J. G. Siedman, J. A. Smith, and K. Struhl, eds., 1987), CURRENT PROTOCOLS IN IMMUNOLOGY (J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach and W. Strober, eds., 1991); ANNUAL REVIEW OF IMMUNOLOGY; as well as monographs in

IDT Exhibit 1020.7568

journals such as ADVANCES IN IMMUNOLOGY. All patents, patent applications, and publications mentioned herein, both supra and infra, are hereby incorporated herein by reference.

[0145] A skilled artisan recognizes that there is a conventional single letter code in the art to represent a selection of nucleotides for a particular nucleotide site. For example, R refers to A or G; Y refers to C or T; M refers to A or C; K refers to G or T; S refers to C or G; W refers to A or T; H refers to A or C or T; B refers to C or G or T; V refers to A or C or G; D refers to A or G or T; and N refers to A or C or G or T.

[0146] The term "blunt end" as used herein refers to the end of a dsDNA molecule having 5' and 3' ends, wherein the 5' and 3' ends terminate at the same nucleotide position. Thus, the blunt end comprises no 5' or 3' overhang.

[0147] The term "double stranded molecule" as used herein refers to a molecule that is double stranded at least in part.

[0148] The term "homing endonuclease" as used herein refers to proteins that are encoded by polynucleotides having mobile, self-splicing introns. Homing endonucleases promote the movement of at least part of the DNA sequences that encode them from one polynucleotide location to another by generating a site-specific double-stranded break at a target site in an allele that lacks the corresponding mobile intron. Examples include but are not limited to at least the following: I-Ceu I, I-Sce I, PI-Psp I, PI-Sce I, or a mixture thereof.

[0149] The terms "hairpin" and "stem-loop oligonucleotide" as used herein refer to a structure formed by an oligonucleotide comprised of 5' and 3' terminal regions that are inverted repeats and a non-self-complementary central region, wherein the self-complementary inverted repeats form a double-stranded stem and the non-self-complementary central region forms a single-stranded loop.

[0150] The term "in the absence of exogenous manipulation" as used herein refers to there being modification of a DNA molecule without changing the solution in which the DNA molecule is being modified. In specific embodiments, it occurs in the absence of the hand of man or in the absence of a machine that changes solution conditions, which may also be referred to as buffer conditions. In further specific embodiments, changes in temperature occur during the modification.

IDT Exhibit 1020.7569

[0151] The term "nonidentical function" as used herein refers to two or more enzymes that do not comprise the same activity. For example, two restriction endonucleases would be considered to have identical function, although a restriction endonuclease and a ligase would be considered to have nonidentical function. Further, endonucleases are defined as enzymes that cut double stranded DNA and therefore that have identical function, although this may be performed in a variety of ways. For example, some restriction endonucleases cut frequently enough that they may be considered to have a DNA fragmentation function (such as to reduce an average DNA size to the size appropriate for efficient WGA, for example), and other restriction endonucleases have a function of cleaving unmethylated restriction sites, for example to generate a Methylome library (such as cleavage that would not reduce an average DNA size), for example.

[0152] The term "polished" as used herein refers to the repair of dsDNA fragment termini that may be enzymatically repaired, wherein the repair constitutes the fill in of recessed 3' ends or the exonuclease activity trimming back of 5' ends to form a "blunt end" compatible with adaptor ligation.

II. Specific Embodiments of the Invention

[0153] The present invention may employ particular compositions and methods as described in the following exemplary embodiments.

A. One-step attachment of double-stranded inverted repeat DNA sequences to DNA fragments using stem-loop oligonucleotides

[0154] In this embodiment of the present invention, as illustrated in FIG. 3, DNA is incubated with an exemplary mixture comprising a stem-loop oligonucleotide with 3' recessed, 3' protruding, or blunt end (FIG. 4); a 3'proofreading DNA polymerase (Klenow fragment of the DNA polymerase I, T4 DNA polymerase, *etc.*); T4 DNA ligase; Enz-O-Mix Universal Buffer; ATP; and dNTPs. Four exemplary enzymatic reactions are taking place simultaneously: "polishing" of the DNA ends and the oligonucleotide double-stranded stem-region; ligation of the oligonucleotide 3' end to the 5' phosphate of the DNA leaving a nick between the 3' end of DNA and the 5' end of the oligonucleotide double-stranded stem-region; polymerase extension of the 3' DNA end that propagates toward the end of the stem-loop oligonucleotide; and a strand-displacement reaction within the oligonucleotide stem region. The process results in a library of DNA fragments with inverted repeat adaptors at their ends.

IDT Exhibit 1020.7570

B. Transcription-mediated amplification of DNA library with attached promoter sequence

[0155] In this embodiment of the present invention, as illustrated in FIG. 5, a DNA library produced by incubation with an exemplary mixture comprising a stem-loop oligonucleotide with 3' recessed, 3' protruding, or blunt end (FIG. 4) and T7 promoter sequence within the stem or loop region; a 3'proofreading DNA polymerase (Klenow fragment of the DNA polymerase I, T4 DNA polymerase, *etc.*); T4 DNA ligase; Enz-O-Mix Universal Buffer; ATP; and dNTPs (as shown in FIG. 3), is used as a template for the transcription-mediated amplification. In FIG. 5A, promoter sequence resides within the stem region of a stem-loop oligonucleotide, and there are two inversely oriented promoters at the ends of all DNA fragments: internally located promoter that initiates synthesis of long RNA molecules, and terminal promoter that initiates synthesis of very short (<10 bases) RNA initiation products (not shown in FIG. 5A) (similar products will be produced from the non-attached stem-loop oligonucleotides present in the mixture). In FIG. 5B, the promoter sequence resides within the loop region of a stem-loop oligonucleotide, and there is only one active promoter element that supports synthesis of long RNA products. The transcription-mediated amplification is achieved by incubation of the synthesized DNA library (see above) with T7 RNA polymerase in the presence of ribonucleotides (rNTPs). The products of this reaction are RNA molecules.

[0156] Supplementation of a transcription-mediated amplification reaction with additional ingredients, such as a reverse transcriptase polymerase (M-MuLV reverse transcriptase, AMV reverse transcriptase, *etc.*) and a primer complementary to the T7 promoter or the internal stem region at the 3' end of the RNA molecule (see FIG. 6B), would allow the synthesis of a complementary cDNA. Efficient priming at the 3' ends of RNA is achieved by blocking the transcription-mediated RNA synthesis within the adaptor region to prevent formation of a folded structure at the 3' ends of RNA. Transcriptional arrest can be achieved by using the same chemical or enzymatic modifications of a stem-loop oligonucleotide as described herein in Sections E, F, G, and H and shown in FIGS. 8, 10, and 16C. The products of such a reaction are double stranded DNA/RNA hybrid molecules.

C. Enz-O-Mix reagents and procedures for one-step library synthesis and linear amplification by transcription

[0157] This embodiment is shown in FIG. 6A and illustrates exemplary components and enzymatic reactions involved in the process catalyzed by Enz-O-Mix reagent (Enz-O-Mix 5, FIG. 19B). This reagent simultaneously synthesizes a DNA library with T7

IDT Exhibit 1020.7571

promoter at the ends and amplifies it linearly by transcription. The Enz-O-Mix reagent 5 comprises a stem-loop oligonucleotide with T7 promoter sequence in the loop region; a DNA fragmentation endonuclease (a restriction enzyme, DNase I, methylation-specific nuclease McrBC, benzonase, apoptotic endonuclease, *etc.*); a 3' proofreading DNA polymerase; T4 DNA ligase; T7 RNA polymerase; Enz-O-Mix Universal Buffer; rNTP; and dNTPs. First, HMW DNA is fragmented by an endonuclease to produce 100-3,000 bp DNA fragments. The T4 DNA polymerase "polishes" the DNA ends and the stem-loop oligonucleotide double-stranded stem-region, and T4 ligase ligates the 3' end of the oligonucleotide to the 5' phosphate of the DNA leaving a nick between the 3' end of DNA and the 5' end of the oligonucleotide double-stranded stem-region. DNA polymerase extends the available 3' DNA end toward the end of a stem-loop oligonucleotide and generates a library of DNA fragments with active conformation of T7 promoter at their ends. RNA polymerase transcribes the library, linearly amplifying DNA and producing single stranded RNA molecules.

[0158] In the case of fragmented DNA (for example, cell-free DNA from blood and/or urine) the reaction does not require a fragmentation endonuclease (Enz-O-Mix 2, FIG. 19A); all other components may be the same as described above.

[0159] In another embodiment, as shown in FIG. 6B, supplementation of a transcription-mediated amplification reaction with the additional ingredients, such as a reverse transcriptase polymerase (M-MuLV reverse transcriptase, AMV reverse transcriptase, *etc.*) and T7 primer, allows the synthesis of a complementary cDNA. Efficient priming at the 3' ends of RNA is achieved by blocking the transcription-mediated RNA synthesis within the adaptor region to prevent formation of a folded structure at the 3' ends of RNA. Transcriptional arrest can be achieved by using the same chemical or enzymatic modifications of a stem-loop oligonucleotide as described herein in Sections E, F, G, and H and shown in FIGS 8, 10, and 16C. The products of such a reaction are double stranded DNA/RNA hybrid molecules.

D. Inhibitory effect of inverted repeats at DNA ends on whole genome amplification and locus-specific PCR

[0160] As shown in FIG. 7, inverted repeats attached to both ends of DNA fragments have an inhibitory effect on PCR process. Heating to 95°C during a first PCR cycle denatures DNA strands and generates hairpin-like structures at both 3' and 5' ends of DNA fragments due to the folding of terminal inverted repeat sequences (the melting temperature for these structures is usually more than 100°C). Subsequent cooling activates a thermostable DNA

IDT Exhibit 1020.7572

polymerase and results in extension of the self-primed 3' ends and creation of long hairpin DNA molecules refractory to melting, priming, and PCR amplification. Neither the universal primer A nor any internal, site-specific primer pair would produce any detectable PCR product.

E. Library created by Enz-O-Mix attachment process with a stem-loop oligonucleotide with a non-replicable linker

[0161] This embodiment is illustrated in FIG. 8 and describes the one-step Enz-O-Mix attachment process for a stem-loop oligonucleotide with a non-replicable linker. The exemplary reaction mix comprises fragmented DNA; a stem-loop oligonucleotide with 3' recessed, 3' protruding or blunt end (FIG. 4), and a non-replicable linker (such as about in the central part of the oligonucleotide); a 3' proofreading DNA polymerase (Klenow fragment of the DNA polymerase I, T4 DNA polymerase, *etc.*); T4 DNA ligase; Enz-O-Mix Universal Buffer; ATP; and dNTPs. At least the following enzymatic reactions are taking place: "polishing" of the DNA ends and the stem-loop oligonucleotide double-stranded stem-region; ligation of the oligonucleotide 3' end to the 5' phosphate of the DNA, leaving a nick between the 3' end of DNA and the 5' end of the oligonucleotide double-stranded stem-region; polymerase extension of the 3' DNA end that propagates toward the end of stem-loop oligonucleotide and stops somewhere within the loop or close to the loop region at the replication block, for example. The process results in the library of DNA fragments with universal sequence A at the ends and an inverted repeat attached only to the 5' end of DNA.

F. Libraries produced by the Enz-O-Mix attachment process with a stem-loop oligonucleotide comprising a non-replicable linker have no inhibitory effect on PCR

[0162] As shown in FIG. 9, a DNA library produced by the Enz-O-Mix (Enz-O-Mix 1 in the case of fragmented DNA, FIG. 19A, or Enz-O-Mix 4 in the case of HMW DNA, FIG. 19B) attachment process with a stem-loop oligonucleotide with a non-replicable linker has no inhibitory effect on WGA and PCR. Heating to 95°C during first PCR cycle denatures DNA strands and generates hairpin-like structures only at the 5' ends of DNA fragments. At a lower PCR temperature, the universal primer A anneals to the A' region at the 3' ends of DNA fragments, and a thermostable DNA polymerase replicates DNA until it reaches a non-replicable linker in the loop or in the region adjacent to the loop stem region of the stem-loop oligonucleotide attached to the 5' DNA end. Replication of the hairpin region is accompanied by a strand-displacement reaction within the oligonucleotide stem region. Synthesized DNA

IDT Exhibit 1020.7573

molecules with the universal sequence A and no inverted repeats at the ends (shown by dashed lines) can then successfully be amplified by PCR.

G. Stem-loop oligonucleotide with a non-replicable linker introduced during chemical synthesis

[0163] FIG. 10A shows the structure of a stem-loop oligonucleotide with a non-replicable linker introduced chemically during oligonucleotide synthesis and shows detailed events occurring at DNA end during the exemplary multi-enzyme attachment process. Many structural modifications within the oligonucleotide, such as the exemplary hexaethelene glycol linker, the abasic site, or a cluster of abasic sites (a gap), a bulky group within the base or backbone, *etc.* can block propagation of the DNA replication. The exemplary process occurs as described in FIG. 8.

H. Stem-loop oligonucleotide with a non-replicable abasic gap generated enzymatically by dU-glycosylase during the attachment process

[0164] This embodiment is illustrated in FIG. 10B and shows the structure of a stem-loop oligonucleotide with non-replicable linker introduced enzymatically during the attachment reaction and shows detailed events occurring at a DNA end during the exemplary multi-enzyme attachment process. The exemplary reaction mix comprises fragmented DNA; a stem-loop oligonucleotide with 3' recessed, 3' protruding or blunt end (FIG. 4), and dU bases within the loop and stem regions (optional); a dU-glycosylase; a 3' proofreading DNA polymerase; T4 DNA ligase; Enz-O-Mix Universal Buffer; ATP; and dNTPs. dU-glycosylase creates abasic sites within the loop and upper strand of the stem region of the oligonucleotide; DNA polymerase "polishes" the ends of DNA and the stem-loop oligonucleotide and generates blunt ends; T4 DNA ligase attaches the 3' end of the oligonucleotide and the 5' end of DNA fragment, leaving a nick between the 3' end of DNA and the 5' end of the hairpin double-stranded stem-region; a DNA polymerase extends the 3' end of DNA toward the end of the stem-loop oligonucleotide and stops somewhere within the loop or close to the loop region of the oligonucleotide, such as at the replication block generated by a contiguous abasic site region. The role of abasic sites in the about central stem region is to destabilize base pairing and facilitate the strand displacement reaction. The process results in a library of DNA fragments with universal sequence at both ends and an inverted repeat attached only to the 5' end of DNA, thereby suitable for PCR amplification. Heating at 95°C during PCR generates breaks at abasic sites and completely eliminates the 5' portion of a stem-loop oligonucleotide.

IDT Exhibit 1020.7574

I. Stem-loop oligonucleotides are advantageous for the Enz-O-Mix attachment process

[0165] This embodiment illustrates the low efficiency and problems associated with use of standard adaptors in the Enz-O-Mix attachment process, although in alternative embodiments standard adaptors may be employed.

[0166] FIG. 11A shows a standard adaptor formed by two different oligonucleotides, wherein one of them has the 5' phosphate group, and possible products formed when such an adaptor is used in the multi-enzyme one-step ligation process. The adaptor should be present at high concentration to provide efficient ligation to DNA ends. Because both ends of a standard adaptor can be ligated to DNA and form adaptor-adaptor conjugates (a dominating product of the reaction), a number of undesirable products may be high, and the yield of desirable molecules may be low, in certain aspects.

[0167] FIG. 11B shows a standard adaptor formed by two different oligonucleotides without phosphate group and two products formed when such an adaptor is used in the multi-enzyme one-step ligation process. Only 50% of formed molecules are desirable products with correct orientation of the adaptor, while the other 50% of molecules have inversed adaptor orientation.

[0168] FIG. 11C shows a standard adaptor formed by two different oligonucleotides without phosphate group but with a protective group at the 3' end of one oligonucleotide and the products formed when such adaptor is used in the multi-enzyme one-step ligation process. A blocked 3' end or recessive 3' end or both are well-known in the art and have been used to prevent formation of products with inversed adaptor orientation during ligation reaction. Presence of a 3' proofreading DNA polymerase in the mix will replace the terminal 3' blocked nucleotide with normal nucleotide (a), and repair and extend the recessed, blocked 3' end (b). As in the case described for FIG. 11B, only 50% of formed molecules are desirable products with correct orientation of the adaptor, while the other 50% of molecules have inversed adaptor orientation.

J. Library generated by the Enz-O-Mix attachment process using a stem-loop oligonucleotide with a cleavable site generated during the multi-enzyme reaction

[0169] In this embodiment of the present invention, as illustrated in FIG. 12, there is a schematic description of the one-step Enz-O-Mix attachment process for a stem-loop

IDT Exhibit 1020.7575

oligonucleotide with an enzymatically cleavable site generated during the multi-enzyme reaction. The reaction mix contains fragmented DNA, a stem-loop oligonucleotide with 3' recessed, 3' protruding or blunt end (FIG. 4), and a specific DNA sequence R within the loop or in the loop and adjacent stem region; a 3'proofreading DNA polymerase (Klenow fragment of the DNA polymerase I, T4 DNA polymerase, *etc.*); T4 DNA ligase; an endonuclease that recognizes and cuts the oligonucleotide specific oligo-sequence R when it adopts a double-stranded conformation; Enz-O-Mix Universal Buffer; ATP; and dNTPs. Four enzymatic reactions are taking place nearly simultaneously: "polishing" of the DNA ends and the stem-loop oligonucleotide double-stranded stem-region; ligation of the oligonucleotide 3' end to the 5' phosphate of the DNA leaving a nick between the 3' end of DNA and the 5' end of the hairpin double-stranded stem-region, polymerase extension of the 3' DNA end that propagates toward the end of stem-loop oligonucleotide and generates inverted repeats at the ends of DNA fragments; and, finally, cleavage of inverted repeats by an endonuclease at the R sites. The process results in a library of DNA fragments with universal sequence A at the ends, and no inverted repeat attached to the ends of DNA fragments, which is thereby competent for PCR amplification.

K. Examples of stem-loop adaptors with a cleavable restriction site generated by DNA synthesis

[0170] FIG. 13A shows an exemplary stem-loop oligonucleotide with recognition sequence for the restriction endonuclease Eco NI located at the loop region. The enzyme can not cut the CCTNNNNNAGG (SEQ ID NO: 20) region within the stem-loop oligonucleotide (an unrecognizable structure) but it will cut it efficiently when the oligonucleotide is attached to a DNA end and adopts a canonical (double-stranded) Watson-Crick conformation. Cleavage with Eco NI eliminates inverted repeats that inhibit PCR applications.

[0171] FIG. 13B shows an exemplary stem-loop oligonucleotide with recognition sequence for the restriction endonuclease Pac I located within the loop region. The enzyme can not cut the TTAATTAAregion within the stem-loop oligonucleotide, but it will cut it efficiently when the oligonucleotide is attached to a DNA end and adopts a canonical (double-stranded) Watson-Crick conformation. Cleavage eliminates inverted repeats that inhibit PCR applications. Use of a rare-cutting restriction enzyme such as Pac I reduces the number of DNA fragments affected by the enzyme cleavage.

IDT Exhibit 1020.7576

[0172] FIG. 13C shows the stem-loop oligonucleotide with recognition sequence for the homing endonuclease I-Ceu I located within the loop region. The enzyme can not cut the 26-base single-stranded region within the loop, but it cuts efficiently when the stem-loop oligonucleotide is attached to a DNA end and adopts a canonical (double-stranded) Watson-Crick conformation. Cleavage eliminates inverted repeats that inhibit PCR applications. Use of a homing endonuclease dramatically reduces and practically eliminates cleavage of genomic DNA.

L. Library generated by the Enz-O-Mix attachment process with a stem-loop oligonucleotide with a short stem

[0173] Amplifiable libraries can be synthesized in the presence of but not limited to a stem-loop oligonucleotide with a non-replicable or cleavable linker. FIG. 14 shows a schematic description of an exemplary one-step Enz-O-Mix attachment process using a stem-loop oligonucleotide with a very short stem. The exemplary reaction mix comprises fragmented DNA, a stem-loop oligonucleotide with 3' recessed, 3' protruding or blunt end (FIG. 4), and a short stem region (5-8 bases); a 3'proofreading DNA polymerase (Klenow fragment of the DNA polymerase I, T4 DNA polymerase, *etc.*); T4 DNA ligase; Enz-O-Mix Universal Buffer; ATP; and dNTPs. Three enzymatic reactions are taking place: "polishing" of the DNA ends and the stem-loop oligonucleotide double-stranded stem-region; ligation of the oligonucleotide 3' end to the 5' phosphate of the DNA, leaving a nick between the 3' end of DNA and the 5' end of the hairpin double-stranded stem-region; and polymerase extension of the 3' DNA end that propagates toward the end of the stem-loop oligonucleotide and generates short inverted repeats at the ends of DNA fragments. The process results in a library of DNA fragments with universal sequence A at the ends and a very short inverted repeat attached to the ends of DNA fragments (FIG. 14, shown in bold). PCR is possible because the melting temperature for terminal hairpins is low enough to prevent folding, self-priming and formation of long hairpin molecules, as shown on FIG. 7. Primer A (representing the 3' portion of a stem-loop oligonucleotide without the complementary 5' region) or even the whole stem-loop oligonucleotide can be utilized as a primer for PCR amplification.

M. Enz-O-Mix reagents and procedures for one-step library synthesis and one-step library synthesis/amplification utilizing strand displacement DNA amplification

[0174] This embodiment is shown in FIG. 15 and illustrates exemplary components and enzymatic reactions involved in the process catalyzed by Enz-O-Mix reagents (Enz-O-Mix 3 and Enz-O-Mix 6, FIGS. 19A and 19B). The reagent Enz-O-Mix 3 can use fragmented (cell-free serum DNA or urine DNA, for example), while the reagent Enz-O-Mix 6 can utilize HMW

IDT Exhibit 1020.7577

DNA and synthesize a library of DNA fragments with the recognition sites for a nicking endonuclease (N.AlwI, N.BbvCIA, N.BbvCIB, Nb.Bpu10I, Nb.BsmI, N.Bst9I, N.BstNBI, *etc.*) at their ends to allow their subsequent amplification by strand displacement DNA synthesis. The presence of nicking endonuclease and strand displacing polymerase in the reagents Enz-O-Mix 3 and Enx-O-Mix 6 allows one to simultaneously synthesize and amplify DNA library by strand displacement DNA synthesis.

[0175] For example, the Enz-O-Mix reagent 3 (suitable for fragmented DNA) comprises a stem-loop oligonucleotide with a recognition sequence for a nicking nuclease; a primer P complementary at least partially to the stem region of a stem-loop oligonucleotide; a 3' proofreading DNA polymerase; a strand displacing DNA polymerase; T4 DNA ligase; a nicking endonuclease; Enz-O-Mix Universal Buffer; and dNTPs. T4 DNA polymerase "polishes" the DNA ends and the stem-loop oligonucleotide double-stranded stem-region. T4 ligase ligates the 3' end of the oligonucleotide to the 5' phosphate of the DNA, leaving a nick between the 3' end of DNA and the 5' end of the oligonucleotide double-stranded stem-region. DNA polymerase extends the available 3' DNA end toward the end of a stem-loop oligonucleotide and generates a library of DNA fragments with functional recognition site for a nicking endonuclease (N.AlwI, N.BbvCIA, N.BbvCIB, Nb.Bpu10I, Nb.BsmI, N.Bst9I, N.BstNBI, *etc.*) at their ends. DNA polymerase initiates multiple rounds of strand displacement DNA synthesis at nicks generated by a nicking nuclease, followed by synthesis of a second DNA strand (primed by primer P), producing double stranded DNA molecules and linearly amplifying DNA.

[0176] In the case of HMW DNA (Enz-O-Mix 6, FIG. 19B), the reaction requires a DNA fragmentation endonuclease (a restriction enzyme, DNase I, methylation-specific nuclease McrBC, benzonase, apoptotic endonuclease *etc.*), and all other components may be the same as described above for Enz-O-Mix 3.

N. Examples of stem-loop adaptors with a nicking endonuclease recognition site generated by DNA synthesis

[0177] FIG. 16A shows the stem-loop oligonucleotide with recognition sequence for the exemplary nicking endonuclease N.Bbv CIA located at the loop region. The enzyme can not nick the single stranded GCTGAGG region within the stem-loop oligonucleotide (an unrecognizable structure), but it will nick it efficiently when the oligonucleotide is attached to a DNA end and adopts a canonical (double-stranded) Watson-Crick conformation.

IDT Exhibit 1020.7578

[0178] FIG. 16B shows an exemplary stem-loop oligonucleotide with recognition sequence for the nicking endonuclease N.Bbv CIA located at the end of the stem region. The enzyme can not nick efficiently the terminally located TTAATTAA region within the stem-loop oligonucleotide, but it will nick it efficiently when the oligonucleotide is attached to a DNA end and the site acquires an internal location. A second nicking site, generated by the oligonucleotide attachment process, would have a terminal location and low nicking efficiency.

[0179] FIG. 16C shows an exemplary stem-loop oligonucleotides 1 and 2 that contains several functional elements within the loop region including the recognition sequence for a nicking endonuclease. Nultifunctional stem-loop oligonucleotide have several regulatory elements within the loop region: T7 promoter sequence (that includes all regions necessary for transcription); recognition sequence N for a nicking endonuclease (N.AlwI, N.BbvCIA, N.BbvCIB, Nb.Bpu10I, Nb.BsmI, N.Bst9I, N.BstNBI, *etc.*); recognition sequence for a restriction endonuclease; and a replication/transcription block region (a chemical modification or a string of 2-4 dU bases). The regulatory elements can reside within a single loop region (L1 in the stem-loop oligonucleotide 1), or be distributed across several loop regions (L1, L2, and L3 in the stem-loop oligonucleotide 2). A multifunctional stem-loop oligonucleotide can support different types of DNA amplification, including PCR, isothermal amplification by transcription, isothermal amplification by strand displacement synthesis, or even a combination thereof and can also facilitate modification such as DNA immobilization, DNA recombination, DNA cloning, or a combination thereof. Activation of the RNA transcription, DNA nicking, DNA double-stranded cleavage, and generation of the 5' overhang functions of the stem-loop adaptor occurs only upon its enzymatic processing and attachment to DNA ends.

O. Enz-O-Mix is a general method for synthesis and amplification of DNA libraries

[0180] This embodiment of the present invention, as illustrated in FIG. 17, shows generalization of the Enz-O-Mix method and demonstrates its specific applications. When Enz-O-Mix reagent is combined with DNA and incubated at optimal temperature (for example, 37°C) for a time necessary to complete the reaction (for example, 1 hour), although it can be more or less, it results in either a DNA library that can be amplified or cloned in a separate reaction, or a DNA library that is amplified isothermally during the incubation with the Enz-O-Mix reagent. For example, the exemplary products of reaction could be as follows: 1) a library of DNA fragments with universal sequence or sequences attached to the ends of DNA fragments so that it can be amplified by PCR (PCR WGA Library); 2) a library of DNA fragments with a promoter

IDT Exhibit 1020.7579

sequence attached at the ends of DNA fragments so that it can be amplified by transcription (Transcription WGA Library); 3) a Methylome Library, *i.e.* the library of adapted DNA fragments where all non-methylated CpG islands are eliminated (such as are described in U.S. Patent Application Serial No. 11/071,864, filed March 3, 2005, which is incorporated by reference herein in its entirety); 4) a library of DNA fragments with ready-to-clone sticky-end adaptors at the ends; 5) an amplified DNA library where the product is DNA or RNA; and/or 6) an amplified Methylome library where the product is DNA or RNA.

P. Stem-loop Oligo Attachment Master Mix is a key component of all Enz-O-Mix reactions

[0181] This embodiment is shown in FIG. 18 and illustrates the components of four different Stem-Loop Oligo Attachment Master Mixes described in the invention and utilized in different Enz-O-Mix reactions involved in the DNA library synthesis. The Stem-Loop Oligo Attachment Master Mix is a key component of the Enz-O-Mix reaction, and it is present in all Enz-O-Mix reagents.

[0182] The exemplary Stem-Loop Oligo Attachment Master Mix I comprises a stem-loop oligonucleotide; a 3' proofreading DNA polymerase; a DNA ligase; Enz-O-Mix Universal Buffer; ATP; and dNTPs (see also FIG. 3 and FIG. 14).

[0183] The exemplary Stem-Loop Oligo Attachment Master Mix II comprises a stem-loop oligonucleotide with a non-replicable chemically introduced linker; a 3' proofreading DNA polymerase; a DNA ligase; Enz-O-Mix Universal Buffer; ATP; and dNTPs (see also FIG. 10A).

[0184] An exemplary Stem-Loop Oligo Attachment Master Mix III comprises a stem-loop oligonucleotide with a non-replicable enzymatically-introduced abasic linker; dU-glycosylase; a 3' proofreading DNA polymerase; a DNA ligase; Enz-O-Mix Universal Buffer; ATP; and dNTPs (see also FIG. 10B).

[0185] An exemplary Stem-Loop Oligo Attachment Master Mix IV comprises a stem-loop oligonucleotide with a cleavable site introduced by DNA replication; a site-specific endonuclease; a 3' proofreading DNA polymerase; a DNA ligase; Enz-O-Mix Universal Buffer; ATP; and dNTPs (see also FIG. 12 and FIGS. 14A, 14B, and 14C).

IDT Exhibit 1020.7580

Q. Enz-O-Mix reagents and procedures for the one-step Whole Genome Libray synthesis and combined Whole Genome Library synthesis-amplification

[0186] Whole Genome Amplification (WGA) is a process that is used to amplify total genomic DNA for many important research and diagnostic applications when the amount of DNA is limited. WGA is frequently preceded by a multi-step WGA library synthesis process. These embodiments describe the Enz-O-Mix reagents and procedures for the one-step Whole Genome Library synthesis and combined one-step, close-tube Whole Genome Library synthesis-amplification processes.

[0187] FIG. 19A shows schematically compositions of three Enz-O-Mix reagents, specifically Enz-O-Mix 1, Enz-O-Mix 2, and Enz-O-Mix 3 designed to convert fragmented DNA into a Whole Genome Library competent for amplification by PCR, transcription, or strand displacement amplification, respectively.

[0188] Reagent Enz-O-Mix 1 may comprise Stem-Loop Oligo Attachment Master Mixes shown in FIG. 18, and it could be also supplemented with a thermophilic DNA polymerase. The library prepared by the Enz-O-Mix 1 process can be used for PCR-mediated WGA in a open-tube (two-step process) or a close-tube (one-step process) formats (as shown in FIG. 30).

[0189] Reagent Enz-O-Mix 2 may comprise Stem-Loop Oligo Attachment Master Mixes shown in FIG. 18 (in this case a stem-loop oligonucleotide has the T7 promoter sequence in the loop or stem region in addition to other necessary base/backbone modifications), and it could be also supplemented with the T7 RNA polymerase and rNTPs. The Enz-O-Mix 2 can be used for a two-step transcription-mediated Whole Genome Amplification (when the library is prepared first, and then amplified) or for a one-step ithothermal amplification process when the Whole Genome library synthesis and its amplification by transcription occurs simultaneously (as shown in FIG. 6).

[0190] Reagent Enz-O-Mix 3 may comprise Stem-Loop Oligo Attachment Master Mixes shown in FIG. 18 (in this case, a stem-loop oligonucleotide has a recognition sequence for a nicking endonuclease in the loop or stem region in addition to other necessary base/backbone modifications). It could be also supplemented with a nicking endonuclease (N.AlwI, N.BbvCIA, N.BbvCIB, Nb.Bpu10I, Nb.BsmI, N.Bst9I, N.BstNBI, *etc.*), a strand displacing DNA polymerase (Klenow exo-, Phi 29, Bst I, *etc.*), and primer P (FIG. 15). The Enz-O-Mix 3 can be used for a two-step strand displacement-mediated Whole Genome Amplification (when the

IDT Exhibit 1020.7581

library is prepared first, and then amplified) or for a one-step ithothermal amplification process when the Whole Genome library synthesis and its amplification by a strand displacement synthesis occurs simultaneously (as shown in FIG. 15).

[0191] FIG. 19B shows schematically compositions of three Enz-O-Mix reagents, specifically Enz-O-Mix 4, Enz-O-Mix 5, and Enz-O-Mix 6 designed to convert high molecular DNA into a Whole Genome Library competent for amplification by PCR, transcription, or strand displacement amplification, respectively.

[0192] Reagent Enz-O-Mix 4 may comprise Stem-Loop Oligo Attachment Master Mixes shown in FIG. 18; it is also supplemented with a DNA fragmentation endonuclease such as restriction enzyme DNase I, Benzonase, methylation-specific nuclease McrBC, apoptotic endonuclease, etc.; and it could be also supplemented with a thermophilic DNA polymerase. The library prepared by the Enz-O-Mix 4 process can be used for PCR-mediated WGA in a open-tube (two-step process) or a close-tube (one-step process) formats (as shown in FIG. 30).

[0193] Reagent Enz-O-Mix 5 may comprise Stem-Loop Oligo Attachment Master Mixes shown in FIG. 18 (in this case, a stem-loop oligonucleotide has the T7 promoter sequence in the loop or stem region in addition to other necessary base/backbone modifications), and it is also supplemented with a DNA fragmentation endonuclease such as restriction enzyme DNase I, benzonase, methylation-specific nuclease McrBC, apoptotic endonuclease, *etc.*; and it could be also supplemented with RNA polymerase and rNTPs. The Enz-O-Mix 5 can be used for a two-step transcription-mediated Whole Genome Amplification of HMW DNA (when the library is prepared first, and then amplified) or for a one-step ithothermal amplification process when the Whole Genome library synthesis and its amplification by transcription occurs simultaneously (as shown in FIG. 6).

[0194] Reagent Enz-O-Mix 6 may comprise Stem-Loop Oligo Attachment Master Mixes shown in FIG. 18 (in this case, a stem-loop oligonucleotide has a recognition sequence for a nicking endonuclease in the loop or stem region in addition to other necessary base/backbone modifications), and it is supplemented with a DNA fragmentation endonuclease such as restriction enzyme, DNase I, benzonase, methylation-specific nuclease McrBC, apoptotic endonuclease. It could also be supplemented with a nicking endonuclease (N.AlwI, N.BbvCIA, N.BbvCIB, Nb.Bpu10I, Nb.BsmI, N.Bst9I, N.BstNBI, *etc.*); a strand displacing DNA polymerase (Klenow exo-, Phi 29, Bst I, *etc.*); and primer P (FIG. 15). The Enz-O-Mix 6 can be

IDT Exhibit 1020.7582

used for a two-step strand displacement-mediated Whole Genome Amplification of HMW DNA (when the library is prepared first, and then amplified) or for a one-step ithothermal amplification process when the Whole Genome library synthesis and its amplification by a strand displacement synthesis occurs simultaneously (as shown in FIG. 15).

R. Enz-O-Mix reagents and procedures for one-step Whole Methylome Library synthesis and combined Whole Methylome Library synthesis-amplification

[0195] Whole Methylome Amplification (WMA) is a process that was recently developed to amplify methylated DNA and suppress amplification of non-methylated genomic (promoter) regions for many important research and diagnostic applications, especially when the amount of DNA is limited (see, for example, U.S. Patent Application Serial No. 11/071,864, filed March 3, 2005, which is incorporated by reference herein in its entirety). In a standard protocol, WMA is preceded by a multi-step WMA library synthesis process. This embodiment describes the exemplary Enz-O-Mix reagents and procedures for the one-step Whole Methylome Library synthesis and combined Whole Methylome Library synthesis-amplification processes.

[0196] FIG. 19C shows schematically compositions of three Enz-O-Mix reagents, Enx-O-Mix 7,Enx-O-Mix 8, and Enx-O-Mix 9, designed to convert fragmented DNA into a Whole Methylome Library competent for amplification by PCR, transcription, or strand displacement, for example.

[0197] Reagent Enz-O-Mix 7 may comprise one of Stem-Loop Oligo Attachment Master Mixes shown in FIG. 18, and it is also supplemented with a mixture of several (about 5-about 12) methylation-sensitive restriction enzymes, such as Aci I, Acc II, Asp LE I, Ava I, Bce AI, Bsa HI, Bsh 1236 I, Bsi E1, Bsi SI, Bst FN I, Bst HH I, Bst UI, Cfo I, Hap II, Hga I, Hha I, HinP1 I, Hin 6I, Hpa II, Hpy 99I, Hpy CH4 IV, Hsp AI, Mvn I, and Ssi I,, to produce a substantial fragmentation of all non-methylated CpG islands (see, for example, U.S. Patent Application Serial No. 11/071,864, filed March 3, 2005, which is incorporated by reference herein in its entirety). Reagent Enz-O-Mix 7 could be also supplemented with a thermophilic DNA polymerase. The Methylome library prepared by the Enz-O-Mix 7 process can be used for PCR-mediated WMA in a open-tube (two-step process) or a close-tube (one-step process) formats (as shown in FIG. 30).

[0198] Reagent Enz-O-Mix 8 may comprise Stem-Loop Oligo Attachment Master Mixes shown in FIG. 18 (in this case a stem-loop oligonucleotide has the T7 promoter sequence

IDT Exhibit 1020.7583

in the loop or stem region in addition to other necessary base/backbone modifications), and it is also supplemented with a mixture of several (about 5-about 12) methylation-sensitive restriction enzymes, such as Aci I, Acc II, Asp LE I, Ava I, Bce AI, Bsa HI, Bsh 1236 I, Bsi E1, Bsi SI, Bst FN I, Bst HH I, Bst UI, Cfo I, Hap II, Hga I, Hha I, HinP1 I, Hin 6I, Hpa II, Hpy 99I, Hpy CH4 IV, Hsp AI, Mvn I, and Ssi I,, to produce a substantial fragmentation of all non-methylated CpG islands (U.S. Patent Application Serial No. 11/071,864, filed March 3, 2005, which is incorporated by reference herein in its entirety). Reagent Enz-O-Mix 8 could be also supplemented with T7 RNA polymerase and rNTPs. The Enz-O-Mix 8 can be used for a two-step transcription-mediated Whole Methylome Amplification (when the Whole Methylome library is prepared first, and then amplified) or for a one-step ithothermal amplification process when the Methylome library synthesis and its amplification by transcription occurs simultaneously.

[0199] Reagent Enz-O-Mix 9 may comprise Stem-Loop Oligo Attachment Master Mixes shown in FIG. 18 (in this case, a stem-loop oligonucleotide has a recognition sequence for a nicking endonuclease in the loop or stem region in addition to other necessary base/backbone modifications), and it is also supplemented with a mixture of several (about 5-12) methylation-sensitive restriction enzymes, such as Aci I, Acc II, Asp LE I, Ava I, Bce AI, Bsa HI, Bsh 1236 I, Bsi E1, Bsi SI, Bst FN I, Bst HH I, Bst UI, Cfo I, Hap II, Hga I, Hha I, HinP1 I, Hin 6I, Hpa II, Hpy 99I, Hpy CH4 IV, Hsp AI, Mvn I, and Ssi I,, to produce a substantial fragmentation of all non-methylated CpG islands (see, for example, U.S. Patent Application Serial No. 11/071,864, filed March 3, 2005, which is incorporated by reference herein in its entirety). Reagent Enz-O-Mix 9 could be also supplemented with nicking endonuclease (N.AlwI, N.BbvCIA, N.BbvCIB, Nb.Bpu10I, Nb.BsmI, N.Bst9I, N.BstNBI, etc.), a strand displacing DNA polymerase (Klenow exo-, Phi 29, Bst I, etc.), and primer P (FIG. 15). The Enz-O-Mix 9 can be used for a two-step strand displacement-mediated Whole Methylome Amplification (when the library is prepared first, and then amplified) or for a one-step ithothermal amplification process when the Methylome library synthesis and its amplification by a strand displacement synthesis occurs simultaneously (as shown in FIG. 15).

[0200] FIG. 19D show schematically compositions of another three Enz-O-Mix reagents, Enx-O-Mix 10, Enx-O-Mix 11, and Enx-O-Mix 12, designed to convert HMW DNA into a Whole Methylome Library competent for amplification by PCR, transcription, or strand displacement, for example.

IDT Exhibit 1020.7584

[0201] Reagent Enz-O-Mix 10 may comprise one of Stem-Loop Oligo Attachment Master Mixes shown in FIG. 18, and it is also supplemented with a DNA fragmentation endonuclease such as restriction enzyme, DNase I, benzonase, methylation-specific nuclease McrBC, apoptotic endonuclease, and a mixture of several (about 5-12) methylation-sensitive restriction enzymes, such as Aci I, Acc II, Asp LE I, Ava I, Bce AI, Bsa HI, Bsh 1236 I, Bsi E1, Bsi SI, Bst FN I, Bst HH I, Bst UI, Cfo I, Hap II, Hga I, Hha I, HinP1 I, Hin 6I, Hpa II, Hpy 99I, Hpy CH4 IV, Hsp AI, Mvn I, and Ssi I,, to produce a substantial fragmentation of all non-methylated CpG islands (see, for example, U.S. Patent Application Serial No. 11/071,864, filed March 3, 2005, which is incorporated by reference herein in its entirety). It could be also supplemented with a thermophilic DNA polymerase. The Methylome library prepared by the Enz-O-Mix 10 process can be used for PCR-mediated WMA in a open-tube (two-step process) or a close-tube (one-step process) formats (as shown in FIG. 30).

[0202] Reagent Enz-O-Mix 11 may comprise Stem-Loop Oligo Attachment Master Mixes shown in FIG. 18 (in this case a stem-loop oligonucleotide has the T7 promoter sequence in the loop or stem region in addition to other necessary base/backbone modifications), and it is also supplemented with a DNA fragmentation endonuclease, such as restriction enzyme, DNase I, benzonase, methylation-specific nuclease McrBC, apoptotic endonuclease, *etc.*;; and a mixture of several (about 5-about 12) methylation-sensitive restriction enzymes, such as Aci I, Acc II, Asp LE I, Ava I, Bce AI, Bsa HI, Bsh 1236 I, Bsi E1, Bsi SI, Bst FN I, Bst HH I, Bst UI, Cfo I, Hap II, Hga I, Hha I, HinP1 I, Hin 6I, Hpa II, Hpy 99I, Hpy CH4 IV, Hsp AI, Mvn I, and Ssi I,, to produce a substantial fragmentation at all non-methylated CpG islands (see, for example, U.S. Patent Application Serial No. 11/071,864, filed March 3, 2005, which is incorporated by reference herein in its entirety). Reagent Enz-O-Mix 11 could be also supplemented with T7 RNA polymerase and rNTPs. The Enz-O-Mix 11 can be used for a two-step transcription-mediated Whole Methylome Amplification (when the Whole Methylome library is prepared first, and then amplified) or for a one-step ithothermal amplification process when the Methylome library synthesis and its amplification by transcription occurs simultaneously.

[0203] Reagent Enz-O-Mix 12 may comprise Stem-Loop Oligo Attachment Master Mixes shown in FIG. 18 (in this case, a stem-loop oligonucleotide has a recognition sequence for a nicking endonuclease in the loop or stem region in addition to other necessary base/backbone modifications), and it is also supplemented with a DNA fragmentation endonuclease, such as restriction enzyme, DNase I, benzonase, methylation-specific nuclease McrBC, apoptotic

IDT Exhibit 1020.7585

endonuclease; and a mixture of several (about 5-12) methylation-sensitive restriction enzymes, such as Aci I, Acc II, Asp LE I, Ava I, Bce AI, Bsa HI, Bsh 1236 I, Bsi E1, Bsi SI, Bst FN I, Bst HH I, Bst UI, Cfo I, Hap II, Hga I, Hha I, HinP1 I, Hin 6I, Hpa II, Hpy 99I, Hpy CH4 IV, Hsp AI, Mvn I, and Ssi I,, to produce a substantial fragmentation of all non-methylated CpG islands (see, for example, U.S. Patent Application Serial No. 11/071,864, filed March 3, 2005, which is incorporated by reference herein in its entirety). Reagent Enz-O-Mix 12 could be also supplemented with nicking endonuclease (N.AlwI, N.BbvCIA, N.BbvCIB, Nb.Bpu10I, Nb.BsmI, N.Bst9I, N.BstNBI, etc.), a strand displacing DNA polymerase (Klenow exo-, Phi 29, Bst I, etc.), and primer P (FIG. 15). The Enz-O-Mix 12 can be used for a two-step strand displacement-mediated Whole Methylome Amplification (when the library is prepared first, and then amplified) or for a one-step ithothermal amplification process when the Methylome library synthesis and its amplification by a strand displacement synthesis occurs simultaneously (as shown in FIG. 15).

S. Enz-O-Mix reagents and procedures for the one-step Whole Genome Libray synthesis for cloning

[0204] DNA cloning is a powerful tool for whole genome DNA sequencing and analysis of cDNA libraries, for example. In a standard process, DNA is randomly fragmented, repaired, attached to linkers, and then integrated into a vector. This embodiment describes the Enz-O-Mix reagents for the one-step preparation of DNA fragments with sticky ends that is easy to clone.

[0205] FIG. 19E shows schematically exemplary compositions of two Enz-O-Mix reagents designed to convert fragmented (Enz-O-Mix 13) or HMW DNA (Enz-O-Mix 14) into a Whole Genome Library for cloning.

[0206] Reagent Enz-O-Mix 13 may comprise a Stem-Loop Oligo Attachment Master Mix IV shown in FIG. 18, and it is the only component of the Enz-O-Mix 13 reagent. The library generated by the Enz-O-Mix 13 process can only utilize a fragmented DNA (for example, cell-free DNA from blood and/or urine, or DNA fragmented enzymatically or mechanically, for example) and be used for cloning.

[0207] Reagent Enz-O-Mix 14 may comprise a Stem-Loop Oligo Attachment Master Mix IV shown in FIG. 18 that is supplemented with a DNA fragmentation endonuclease, such as restriction enzyme, DNase I, benzonase, methylation-specific nuclease McrBC, apoptotic

IDT Exhibit 1020.7586

endonuclease, *etc.* The library generated by the Enz-O-Mix 14 process can utilize HMW DNA from any type of cells and tissues and be used for cloning.

T. Enz-O-Mix reagent kits and their use in high-throughput applications and clinical diagnostics

[0208] In this embodiment of the present invention, as illustrated in FIG. 20, pre-assembled Enz-O-Mix reagents can be stored at -20°C and used as ready-to-use kits.

[0209] Exemplary pre-assembled Enz-O-Mix 1 comprises all major components of the reaction, such as a stem-loop oligonucleotide; DNA polymerase; a DNA ligase; dU-glycosylase (when the hairpin non-replicable region is generated by abasic sites); a hairpin-specific endonuclease (when the hairpin has a cleavable site generated during oligonucleotide attachment process); thermophilic DNA polymerase (close-tube WGA library synthesis / PCR-mediated amplification reaction); ATP; and dNTPs.

[0210] Exemplary pre-assembled Enz-O-Mix 2 comprises a stem-loop oligonucleotide with a promoter sequence in a stem or loop region; DNA polymerase; a DNA ligase; dU-glycosylase (when the non-replicable region is formed by abasic sites); a site-specific endonuclease (when the stem-loop oligonucleotide has a cleavable site created during oligonucleotide attachment process); ATP; and dNTPs; In the case of close-tube, isothermal library synthesis / transcription-mediated amplification reaction it also comprises RNA polymerase and rNTPs.

[0211] Exemplary pre-assembled Enz-O-Mix 3 may comprise a stem-loop oligonucleotide with a nicking endonuclease sequence in a stem or a loop region; DNA polymerase; a DNA ligase; dU-glycosylase (when the non-replicable region is formed by abasic sites); a site-specific endonuclease (when the stem-loop oligonucleotide has a cleavable site generated during oligonucleotide attachment process); a DNA fragmentation endonuclease, such as restriction enzyme(s), DNase I, Benzonase, methylation-specific nuclease McrBC, apoptotic endonuclease *etc.*; ATP; and dNTP, In the case of close-tube, isothermal library synthesis / strand displacement-mediated amplification reactiuon it also comprises a nicking endonuclease (N.AlwI, N.BbvCIA, N.BbvCIB, Nb.Bpu10I, Nb.BsmI, N.Bst9I, N.BstNBI, *etc.*); a strand displacing DNA polymerase (Klenow exo-, Phi 29, Bst I, *etc.*); and a primer.

[0212] Exemplary pre-assembled Enz-O-Mix 4 may comprise all major components of the reaction, such as a stem-loop oligonucleotide; DNA polymerase; a DNA

IDT Exhibit 1020.7587

ligase; dU-glycosylase (when the hairpin non-replicable region is created by abasic sites); a hairpin-specific endonuclease (when the hairpin has a cleavable site created during oligonucleotide attachment process); a DNA fragmentation endonuclease such as restriction enzyme, DNase I, Benzonase, methylation-specific nuclease McrBC, apoptotic endonuclease ; thermophilic DNA polymerase (close-tube WGA library synthesis / PCR-mediated amplification reaction); ATP; and dNTPs.

[0213] Exemplary pre-assembled Enz-O-Mix 5 may comprise a stem-loop oligonucleotide with a promoter sequence in a stem or loop region; DNA polymerase; DNA ligase; dU-glycosylase (when the non-replicable region is formed by abasic sites); a site-specific endonuclease (when the stem-loop oligonucleotide has a cleavable site created during the oligonucleotide attachment process); a DNA fragmentation endonuclease, such as restriction enzyme, DNase I, benzonase, methylation-specific nuclease McrBC, apoptotic endonuclease, *etc.*; ATP; and dNTPs. In the case of close-tube, isothermal library synthesis / transcription-mediated amplification reaction it also comprises RNA polymerase and rNTPs..

[0214] Exemplary pre-assembled Enz-O-Mix 6 comprises a stem-loop oligonucleotide with a nicking endonuclease sequence in a stem or a loop region, DNA polymerase, a DNA ligase, dU-glycosylase (when the non-replicable region is formed by abasic sites), a site-specific endonuclease (when the stem-loop oligonucleotide has a cleavable site created during oligonucleotide attachment process), a DNA fragmentation endonuclease such as restriction enzyme(s), DNase I, Benzonase, methylation-specific nuclease McrBC, apoptotic endonuclease; ATP; and dNTPs. In the case of close-tube, isothermal library synthesis / strand displacement-mediated amplification reactiuon it also comprises a nicking endonuclease (N.AlwI, N.BbvCIA, N.BbvCIB, Nb.Bpu10I, Nb.BsmI, N.Bst9I, N.BstNBI, *etc.*); a strand displacing DNA polymerase (Klenow exo-, Phi 29, Bst I, *etc.*); and a primer.

[0215] Exemplary pre-assembled Enz-O-Mix 7 may comprise all major components of the reaction, such as a stem-loop oligonucleotide ; DNA polymerase; T4 DNA ligase; dU-glycosylase (when the hairpin non-replicable region is created by abasic sites); a hairpin-specific endonuclease (when the hairpin has a cleavable site created during oligonucleotide attachment process); a mixture of several (about 5-about 12) methylation-sensitive restriction enzymes, such as Aci I, Acc II, Asp LE I, Ava I, Bce AI, Bsa HI, Bsh 1236 I, Bsi E1, Bsi SI, Bst FN I, Bst HH I, Bst UI, Cfo I, Hap II, Hga I, Hha I, HinP1 I, Hin 6I, Hpa II, Hpy 99I, Hpy CH4 IV, Hsp AI, Mvn I, and Ssi I,, to produce a substantial fragmentation at all

IDT Exhibit 1020.7588

non-methylated CpG islands (see, for example, U.S. Patent Application Serial No. U.S. Patent Application Serial No. 11/071,864, filed March 3, 2005, which is incorporated by reference herein in its entirety); thermophilic DNA polymerase (close-tube WMA library synthesis / PCR-mediated amplification reaction); ATP; and dNTPs.

[0216] Exemplary pre-assembled Enz-O-Mix 8 may comprise a stem-loop oligonucleotide with a promoter sequence in a stem or loop region; DNA polymerase; a DNA ligase; dU-glycosylase (when the non-replicable region is formed by abasic sites); a site-specific endonuclease (when the stem-loop oligonucleotide has a cleavable site created during the oligonucleotide attachment process); a mixture of several (about 5-about 12) methylation-sensitive restriction enzymes, such as Aci I, Acc II, Asp LE I, Ava I, Bce AI, Bsa HI, Bsh 1236 I, Bsi E1, Bsi SI, Bst FN I, Bst HH I, Bst UI, Cfo I, Hap II, Hga I, Hha I, HinP1 I, Hin 6I, Hpa II, Hpy 99I, Hpy CH4 IV, Hsp AI, Mvn I, and Ssi I,, to produce a substantial fragmentation at all non-methylated CpG islands (see, for example, U.S. Patent Application Serial No. 11/071,864, filed March 3, 2005, which is incorporated by reference herein in its entirety); ATP; and dNTPs. In the case of close-tube, isothermal library synthesis / transcription-mediated amplification reaction it also comprises RNA polymeraseand rNTPs.

[0217] Exemplary pre-assembled Enz-O-Mix 9 may comprise a stem-loop oligonucleotide with a nicking endonuclease sequence in a stem or a loop region; DNA polymerase; a DNA ligase; dU-glycosylase (when the non-replicable region is formed by abasic sites); a site-specific endonuclease (when the stem-loop oligonucleotide has a cleavable site created during oligonucleotide attachment process); a DNA fragmentation endonuclease, such as restriction enzyme(s), DNase I, Benzonase, methylation-specific nuclease McrBC, apoptotic endonuclease, *etc.*; a mixture of several (about 5-about 12) methylation-sensitive restriction enzymes, such as Aci I, Acc II, Asp LE I, Ava I, Bce AI, Bsa HI, Bsh 1236 I, Bsi E1, Bsi SI, Bst FN I, Bst HH I, Bst UI, Cfo I, Hap II, Hga I, Hha I, HinP1 I, Hin 6I, Hpa II, Hpy 99I, Hpy CH4 IV, Hsp AI, Mvn I, and Ssi I,, to produce a substantial fragmentation at all non-methylated CpG islands (see, for example, U.S. Patent Application Serial No. 11/071,864, filed March 3, 2005, which is incorporated by reference herein in its entirety); ATP; and dNTPs. In the case of close-tube, isothermal library synthesis / strand displacement-mediated amplification reactiuon it also comprises a nicking endonuclease (N.AlwI, N.BbvCIA, N.BbvCIB, Nb.Bpu10I, Nb.BsmI, N.Bst9I, N.BstNBI, *etc.*); a strand displacing DNA polymerase (Klenow exo-, Phi 29, Bst I, *etc.*); and a primer.

IDT Exhibit 1020.7589

[0218] Exemplary pre-assembled Enz-O-Mix 10 may comprise all major components of the reaction, such as a stem-loop oligonucleotide; DNA polymerase; a DNA ligase; dU-glycosylase (when the hairpin non-replicable region is created by abasic sites); a hairpin-specific endonuclease (when the hairpin has a cleavable site created during oligonucleotide attachment process); a DNA fragmentation endonuclease, such as restriction enzyme, DNase I, Benzonase, methylation-specific nuclease McrBC, apoptotic endonuclease, *etc.*; a mixture of several (about 5-about 12) methylation-sensitive restriction enzymes, such as Aci I, Acc II, Asp LE I, Ava I, Bce AI, Bsa HI, Bsh 1236 I, Bsi E1, Bsi SI, Bst FN I, Bst HH I, Bst UI, Cfo I, Hap II, Hga I, Hha I, HinP1 I, Hin 6I, Hpa II, Hpy 99I, Hpy CH4 IV, Hsp AI, Mvn I, and Ssi I,, to produce a substantial fragmentation at all non-methylated CpG islands (see, for example, U.S. Patent Application Serial No. No. 11/071,864, filed March 3, 2005, which is incorporated by reference herein in its entirety); thermophilic DNA polymerase (close-tube WGA library synthesis / PCR-mediated amplification reaction); ATP; and dNTPs.

[0219] Exemplary pre-assembled Enz-O-Mix 11 may comprise a stem-loop oligonucleotide with a promoter sequence in a stem or loop region; DNA polymerase; a DNA ligase; dU-glycosylase (when the non-replicable region is formed by abasic sites); a site-specific endonuclease (when the stem-loop oligonucleotide has a cleavable site created during oligonucleotide attachment process); a DNA fragmentation endonuclease, such as restriction enzyme, DNase I, benzonase, methylation-specific nuclease McrBC, apoptotic endonuclease; a mixture of several (about 5-12) methylation-sensitive restriction enzymes, such as Aci I, Acc II, Asp LE I, Ava I, Bce AI, Bsa HI, Bsh 1236 I, Bsi E1, Bsi SI, Bst FN I, Bst HH I, Bst UI, Cfo I, Hap II, Hga I, Hha I, HinP1 I, Hin 6I, Hpa II, Hpy 99I, Hpy CH4 IV, Hsp AI, Mvn I, and Ssi I,, to produce a substantial fragmentation at all non-methylated CpG islands (see, for example, U.S. Patent Application Serial No. 11/071,864, filed March 3, 2005, which is incorporated by reference herein in its entirety); ATP; and dNTPs. In the case of close-tube, isothermal library synthesis / transcription-mediated amplification reaction it also comprises RNA polymerase and rNTPs..

[0220] Exemplary pre-assembled Enz-O-Mix 12 may comprise a stem-loop oligonucleotide with a nicking endonuclease sequence in a stem or a loop region; DNA polymerase; a DNA ligase; dU-glycosylase (when the non-replicable region is formed by abasic sites); a site-specific endonuclease (when the stem-loop oligonucleotide has a cleavable site created during oligonucleotide attachment process); a DNA fragmentation endonuclease, such as

IDT Exhibit 1020.7590

restriction enzyme(s), DNase I, benzonase, methylation-specific nuclease McrBC, apoptotic endonuclease; a mixture of several (about 5-about 12) methylation-sensitive restriction enzymes, such as Aci I, Acc II, Asp LE I, Ava I, Bce AI, Bsa HI, Bsh 1236 I, Bsi E1, Bsi SI, Bst FN I, Bst HH I, Bst UI, Cfo I, Hap II, Hga I, Hha I, HinP1 I, Hin 6I, Hpa II, Hpy 99I, Hpy CH4 IV, Hsp AI, Mvn I, and Ssi I,, to produce a substantial fragmentation at all non-methylated CpG islands (see, for example, U.S. Patent Application Serial No. 11/071,864, filed March 3, 2005, which is incorporated by reference herein in its entirety); ATP; and dNTPs. In the case of close-tube, isothermal library synthesis / strand displacement-mediated amplification reactiuon it also comprises a nicking endonuclease (N.AlwI, N.BbvCIA, N.BbvCIB, Nb.Bpu10I, Nb.BsmI, N.Bst9I, N.BstNBI, *etc.*) a strand displacing DNA polymerase (Klenow exo-, Phi 29, Bst I, *etc.*); and a primer.

[0221] Kits comprising the exemplary Enz-O-Mixes described above may be supplied with the Enz-O-Mix Universal Buffer. When used in high-throughput applications and clinical diagnostics, the Enz-O-Mix reagent is first diluted by the Enz-O-Mix Universal Buffer, aliquoted into microplate wells, pre-incubated at the reaction temperature (for example, about 37°C) for a few minutes, supplemented with DNA, and incubated at the reaction temperature for about 1-about 16 h.

[0222] FIG. 20 demonstrates an exemplary application of the Enz-O-Mix kits for high-throughput library preparation / amplification in the 96- or 384-well format.

U. Enz-O-Mix reaction optimization procedures

[0223] In this embodiment, as shown in FIG. 21, there are exemplary procedures for identifying the most optimal parameters for any Enz-O-Mix process using real-time PCR to monitor amplification of the generated library, or a direct fluorescent detection of the amplified product in the case of combined isothermal library synthesis-amplification process.

[0224] FIG. 21A illustrates the procedure for optimization of the Enz-O-Mix Universal Buffer. Three specific but exemplary buffers are utilized, including Universal Buffers 1, 2 and 3. Buffer 2 results in the early WGA amplification curve, suggesting that the multi-enzymatic reaction was more efficient in this buffer, and the Universal Buffer 2 is particularly suitable for the Enz-O-Mix reaction.

IDT Exhibit 1020.7591

[0225] FIG. 21B illustrates the procedure for concentration optimization of the Enz-O-Mix incubation temperature. Three temperatures are employed; specifically, T_1, T_2 and T_3. Temperature T_3 results in the early WGA amplification curve, suggesting that the multi-enzymatic reaction was more efficient at this temperature and that the parameter T_3 should be selected as a more suitable condition for the Enz-O-Mix reaction.

[0226] FIG. 21C illustrates the procedure for optimization of the Enz-O-Mix component N. Three concentrations of the component N; specifically, C^N_1, C^N_2 and C^N_3, were utilized. Concentration C^N_1 results in the early WGA amplification curve, suggesting that the multi-enzymatic reaction was more efficient at this component N concentration, and that the parameter C^N_1 is a particularly suitable condition for the Enz-O-Mix reaction.

V. High potential for the Enz-O-Mix approach in future biotechnology and diagnostic medicine

[0227] The applications of the Enz-O-Mix method are numerous. The Enz-O-Mix method is easy to automate and use in clinical diagnostic and point-of-care applications. Different enzyme combinations could lead to many novel kits and assays, and applications in such areas as biotechnology, the pharmaceutical industry, molecular diagnostics, forensics, pathology, bio-defense, and bio-computing are contemplated The Enz-O-Mix approach can be viewed as a simple and cost-effective alternative to the "lab-on-a chip" micro-fluidic approach that is currently attempting to solve the same problem (multi-step DNA processing) by reduction of reaction volumes and integration of multiple reactions into a small format.

[0228] Exemplary applications for the invention include but are not limited to the following:

[0229] a) A one-step, closed tube preparation and amplification of genomic libraries (WGA) from highly degraded serum, plasma, and/or urine (such as the supernatant fraction) DNA. DNA amplification and re-amplification can be used as an *in vitro* "immortalization" process to maintain and generate necessary quantities of valuable but limited DNA samples for gene association studies, mutation and microsatellite instability detection in cancer diagnostics, etc.

[0230] b) A one-step, closed tube preparation and amplification of genomic libraries (WGA) from formalin fixed, paraffin embedded tissues for cancer diagnostics and research applications.

IDT Exhibit 1020.7592

[0231] c) A one-step, closed tube preparation and amplification of Methylome libraries;

[0232] d) A one-step preparation and simultaneous immobilization of prepared DNA libraries on a solid support.

EXAMPLES

[0233] The following examples are included to demonstrate preferred embodiments of the invention. It should be appreciated by those of skill in the art that the techniques disclosed in the examples that follow represent techniques discovered by the inventors to function well in the practice of the invention, and thus can be considered to constitute preferred modes for its practice. However, those of skill in the art should, in light of the present disclosure, appreciate that many changes can be made in the specific embodiments that are disclosed and still obtain a like or similar result without departing from the spirit and scope of the invention.

EXAMPLE 1: ONE STEP WHOLE GENOME LIBRARY AMPLIFICATION USING HAIRPIN ADAPTOR WITH EcoNI RESRTICTION SITE GENERATED AFTER LIGATION AND STRAND-DISPLACEMENT EXTENSION

[0234] In this example, a single step WGA process is described wherein a partial hairpin oligonucleotide adaptor containing EcoNI recognition half sites in its stem is extended by a DNA polymerase to form a blunt end hairpin molecule and is ligated *via* its free 3' end to the 5' phosphate of blunt-ended DNA restriction fragments by T4 DNA ligase. Due to the strand-displacement activity of the DNA polymerase, the free 3' end of the restriction fragments are extended using as template the ligated 3' end of the adaptor, thus opening up the hairpin structure and generating functional EcoNI cleavage site (FIG. 12 and 13A). Following cleavage with EcoNI present in the same reaction mix, the DNA library molecules have their terminal inverted repeats excised and become amplifiable by PCR using a primer complementary to the 5' stem portion of the adaptor sequence. The whole process takes place in a single tube and in one step in just about 1 hour, in specific embodiments.

[0235] Human genomic DNA isolated from the peripheral blood of a healthy donor by standard procedures was digested with 10 units of AluI restriction enzyme (NEB) for 1 hour following the manufacturer's protocol.

IDT Exhibit 1020.7593

[0236] For library preparation, 5 nanograms of AluI-digested DNA were incubated in a reaction mixture comprising 1x NEBuffer 4 (20 mM Tris-acetate, 10 mM Mg acetate, 50 mM potassium acetate, 1 mM dithiothreitol, pH 7.9), 1 μM EcoNI Ext Adaptor oligonucleotide (Table I, SEQ ID NO:1), 1 mM ATP, 120 μM dNTPs, 2 Units of Klenow fragment of DNA polymerase I, 400 units of T4 DNA ligase (New England Biolabs, NEB; Beverly, MA), and 10 units of EcoNI restriction enzyme (NEB) in a final volume of 15 μl for 1 hour at 37°C. Enzymes were inactivated for 10 min at 75 °C. Control reaction that contained no EcoNI restriction enzyme was also run in parallel.

[0237] The resulting library was amplified by real-time PCR in a reaction mixture containing: 1 x Titanium *Taq* reaction buffer (Clontech; Mountain View, CA); 200 μM each dNTP; 1:10,000 dilutions of fluorescein and SybrGreen I (Molecular Probes; Carlsbad, CA); 1 μM of K_U primer (Table I, SEQ ID NO:2); and 5 units of Titanium *Taq* polymerase (Clontech; Mountain View, CA) in a reaction volume of 75 μl. PCR amplification was carried out for 15 cycles at 95°C for 15 sec and 65°C for 2 min on i-Cycler real-time PCR instrument (Bio-Rad; Hercules, CA). FIG. 22 shows the amplification curves of the libraries prepared in the presence or in the absence of EcoNI. Cleavage of the terminal inverted repeat by EcoNI improved the efficiency of amplification by 7 cycles, *i.e.* over two orders of magnitude.

[0238] Preliminary representation analysis of the library prepared in the presence of EcoNI was conducted using three exemplary human genomic STS markers (Table II). The material amplified by PCR with the universal K_U primer was purified with Qiaquick filters (Qiagen; Valencia, CA), and 20 ng aliquots were analyzed in real-time PCR. Reactions were carried out for 45 cycles at 94°C for 15 sec and 68°C for 1 min on i-Cycler (Bio-Rad; Hercules, CA). The PCR reaction mixture comprised the following: 1 x Titanium *Taq* reaction buffer (Clontech; Mountain View, CA); 200 μM each dNTP; 1:10, 000 dilutions of fluorescein and SybrGreen I (Molecular Probes); 0.2 μM each forward and reverse STS primer (Table II); 5 units of Titanium *Taq* polymerase (Clontech; Mountain View, CA); and 20 ng library DNA. A control reaction comprising 1 ng of human genomic DNA randomly fragmented to an average size of 1.5 Kb using Hydro Shear device (Gene Machines; Palo Alto, CA) was performed. FIGS. 24A-24C show unbiased representation of the analyzed human STS sequences in the EcoNI - treated library.

IDT Exhibit 1020.7594

EXAMPLE 2: ONE STEP WHOLE GENOME AMPLIFICATION USING HAIRPIN ADAPTOR WITH NON-REPLICABLE 18-ATOM HEXA-ETHYLENEGLYCOL SPACER

[0239] In this example, a single step WGA process is described wherein a hairpin oligonucleotide adaptor containing non-replicable 18-atom hexa-ethyleneglycol (HEG) spacer in its loop is ligated *via* its free 3' end to the 5' phosphate of DNA restriction fragments in the presence of T4 DNA ligase and a DNA polymerase. Due to the strand-displacement activity of the polymerase, the free 3' ends of the restriction fragments are extended until the replication stop is reached (see FIGS. 8 and 10A). This process results in truncated 3' termini of the resulting WGA library such that they do not contain a terminal inverted repeat. The library molecules are then amplified by PCR using a primer complementary to the universal sequence of the extended 3' termini. The whole process takes place in a single tube in one step and is completed in just about 1 hour, in specific embodiments.

[0240] In control experiments, hairpin adaptor molecules that either do not comprise a replication stop or modified bases or that comprise ribo-nucleosides in their loop region are compared for their ability to participate in WGA along with the hexa-ethyleneglycol adaptor described above.

[0241] Human genomic DNA isolated from the peripheral blood of a healthy donor by standard procedures was digested with 10 units of AluI restriction enzyme (NEB) for 1 hour following the manufacturer's protocol.

[0242] Five nanograms of AluI digested DNA were incubated in a reaction mixture comprising 1x NEBuffer 4; 1 μM HEG Adaptor oligonucleotide (Table I, SEQ ID NO:3) or control hairpin adaptors (see below); 1 mM ATP; 120 μM dNTPs; 2 Units of Klenow fragment of DNA polymerase I; and 400 units of T4 DNA ligase (NEB) in a final volume of 15 μl for 1 hour at 37°C. Enzymes were inactivated for 10 min at 75 °C.

[0243] To test the effect of the hexa-ethyleneglycol replication stop, control reactions containing partial hairpin oligonucleotides extended by Klenow to form blunt end hairpin adaptors that either do not contain replication stop (Table I, Hairpin Adaptor, SEQ ID NO:5), or containing ribonucleosides in their loop region (Table I, Hairpin Ribo Adaptor, SEQ ID NO:6) were substituted for the HEG Adaptor (Table I, SEQ ID NO:3) and run in parallel as described above.

IDT Exhibit 1020.7595

[0244] The resulting libraries were amplified by real-time PCR in a reaction mixture comprising the following: 1 x Titanium *Taq* reaction buffer (Clontech; Mountain View, CA); 200 μM each dNTP; 1:10,000 dilutions of fluorescein and SybrGreen I (Molecular Probes; Carlsbad, CA); 1 μM of universal primer (Y_U primer, Table I SEQ ID NO:4 in the case of the HEG adaptor, or K_U primer, Table I, SEQ ID NO:2, in the case of control hairpin adaptors); and 5 units of Titanium *Taq* polymerase (Clontech; Mountain View, CA) in a reaction volume of 75 μl. PCR amplification was carried out at 95°C for 15 sec and 65°C for 2 min on i-Cycler real-time PCR instrument (Bio-Rad; Hercules, CA). As shown on FIG. 23, amplification of the library prepared by ligation of hexa-ethyleneglycol adaptor was 9 cycles more efficient than a library made with adaptor not containing a replication stop in its loop structure. The results shown on FIG. 23 also indicate that an array of 5 consecutive ribonucleosides in the adaptor's loop is not a preferable replication stop for Klenow fragment of DNA polymerase I, since amplification of a library prepared by ligation of such an adaptor amplified only 2 cycles earlier than a library made with adaptor containing only deoxyribonucleosides.

[0245] Preliminary representation analysis of the library prepared by ligation of HEG adaptor was done using three exemplary human genomic STS markers (Table II). The material amplified by PCR with the universal Y_U primer was purified with Qiaquick filters (Qiagen; Valencia, CA), and 20 ng aliquots were analyzed in real-time PCR as described in Example 1. FIGS. 24A-24C show unbiased representation of the analyzed human STS sequences in this exemplary library.

EXAMPLE 3: ONE STEP WHOLE GENOME LIBRARY PREPARATION USING DEGRADABLE HAIRPIN ADAPTOR COMPRISING DEOXY-URIDINE

[0246] In this example, a single step WGA process is described wherein a hairpin oligonucleotide adaptor comprising deoxy-uridine in both its 5' stem region and in its loop (FIG. 10B) is ligated *via* its free 3' end to the 5' phosphates of DNA restriction fragments in the presence of 3 enzymatic activities: T4 DNA ligase, DNA polymerase, and Uracil-DNA glycosylase (which is also referred to as UDG or dU glycosylase). UDG catalyzes the release of free uracil and generates abasic sites in the adaptor's loop region and the 5' half of the hairpin. The strand-displacement activity of the DNA polymerase extends the free 3' end of the restriction fragments until an abasic site is reached, serving as a replication stop (FIG. 10B). This process results in truncated 3' ends of the resulting WGA library fragments such that they do not have terminal inverted repeats. Prior to amplification, samples are heated in order to break

IDT Exhibit 1020.7596

the phosphodiester bonds of the generated abasic sites, and the adaptor molecules are degraded into 2 small fragments and one intact 18 base oligonucleotide complementary to the universal sequence of the extended 3' termini (FIG. 10B). The released 18 base oligonucleotide then serves as a primer in subsequent PCR amplification, or alternatively it can be exogenously supplied. The entire process takes place in a single tube in one step and is completed in just 1 hour.

[0247] Human genomic DNA isolated from the peripheral blood of a healthy donor by standard procedures was digested with 10 units of AluI restriction enzyme (New England Biolabs; Beverly, MA) for 1 hour following the manufacturer's protocol.

[0248] For library preparation, five nanograms of AluI-digested genomic DNA were incubated in an exemplary reaction mixture comprising 1x NEBuffer 4; 2 μM dU Hairpin Adaptor oligonucleotide (Table I, SEQ ID NO:7); 1 mM ATP; 120 μM dNTPs; 2 Units of Klenow fragment of DNA polymerase I; 400 units of T4 DNA ligase; and 2 Units of UDG (New England Biolabs; Beverly, MA) in a final volume of 15 μl for 1 hour at 37°C. Enzymes were inactivated at 75 °C for 10 min. A control reaction containing no UDG was run in parallel.

[0249] The resulting library was amplified by real-time PCR in a reaction mixture comprising the following: 1 x Titanium *Taq* reaction buffer (Clontech, Mountain View, CA); 200 μM each dNTP; 1:10,000 dilutions of fluorescein and SybrGreen I (Molecular Probes; Carlsbad, CA); 1 μM of universal M_U primer (Table I, SEQ ID NO:8); and 5 units of Titanium *Taq* polymerase (Clontech; in a reaction volume of 75 μl). PCR amplifications were carried out at 95°C for 15 sec and 65°C for 2 min on i-Cycler real-time PCR instrument (Bio-Rad; Hercules, CA). In a separate amplification reaction, no external universal M_U primer was supplied to test if endogenously-generated primer is sufficient for library amplification. FIG. 25 shows that virtually identical amplification curves were generated whether or not the universal primer was added exogenously or was generated as a result of the UDG activity followed by heat-induced degradation of the adaptor. FIG. 25 also shows that if UDG was not present during library preparation, this caused an 8 cycle-delay in amplification. This latter result once again indicates the preference of removal of the terminal inverted repeat of a hairpin adaptor for efficient library priming and amplification.

IDT Exhibit 1020.7597

EXAMPLE 4: ONE STEP GENOMIC DNA RESTRICTION DIGESTION AND WHOLE GENOME LIBRARY PREPARATION USING DEGRADABLE HAIRPIN ADAPTOR COMPRISING DEOXY-URIDINE

[0250] In this example, a single step WGA process is described wherein a hairpin oligonucleotide adaptor comprising deoxy-uridine described in Example 3 is ligated *via* its free 3' end to the 5' phosphates of DNA restriction fragments generated from intact genomic DNA in a single exemplary reaction mix comprising 4 exemplary enzymatic activities: AluI restriction endonuclease, T4 DNA ligase, DNA polymerase, and Uracil-DNA glycosylase (UDG) (FIG. 19A, Enz-O-Mix 3). UDG catalyses the release of free uracil and generates abasic sites in the adaptor's loop region and the 5' half of the hairpin. AluI digests the target DNA into restriction fragments. T4 DNA ligase generates a phosphodiester bond between the 3' ends of the adaptor molecules and the 5' phosphates of the restriction fragments. Finally, the strand-displacement activity of the DNA polymerase extends the free 3' end of the restriction fragments using as template the ligated 3' end of the hairpin stem until an abasic site is reached that serves as a replication stop (FIG. 10B). Prior to amplification, samples are heated to degrade the abasic sites of the adaptor, and the resulting library is amplified using the universal primer M_U. The entire process takes place in a single tube in one step and is completed in just 1 hour.

[0251] For library preparation, 20, 10, 3, 1, 0.3, 0.1, 0.03, 0.01, or 0 nanograms of human genomic DNA isolated from the peripheral blood of a healthy donor by standard procedures were incubated in an exemplary reaction mix comprising: 1x NEBuffer4; 2 μM dU Hairpin Adaptor oligonucleotide (Table I, SEQ ID NO:7); 1 mM ATP; 120 μM dNTPs; 2 Units of Klenow fragment of DNA polymerase I; 400 units of T4 DNA ligase; 2 Units of UDG; and 10 units of AluI restriction endonuclease (New England Biolabs; Beverly, MA) in a final volume of 15 μl for 1 hour at 37°C. Enzymes were inactivated for 10 min at 75 °C.

[0252] The resulting libraries were amplified by real-time PCR in the exemplary reaction mixture comprising the following: 1 x Titanium *Taq* reaction buffer (Clontech; Mountain View, CA); 200 μM each dNTP; 1:10,000 dilutions of fluorescein and SybrGreen I (Molecular Probes; Carlsbad, CA); 1 μM of universal M_U primer (Table I, SEQ ID NO:8); and 5 units of Titanium *Taq* polymerase (Clontech; Mountain View, CA) in a reaction volume of 75 μl. PCR amplifications were carried out at 95°C for 15 sec and 65°C for 2 min on i-Cycler real-time PCR instrument (Bio-Rad; Hercules, CA). FIG. 26 shows the amplification curves of the resulting libraries. As shown, a series of titration curves was generated in a concentration-dependent manner. The smallest amount of DNA corresponding to less than two genome

IDT Exhibit 1020.7598

equivalents was still discernible from the control reaction with no DNA, indicating that the described single closed tube process is an efficient and sensitive way of whole genome amplification.

EXAMPLE 5: SIMPLIFIED PROTOCOL COMBINING PREPARATION OF METHYLOME LIBRARIES FROM CELL-FREE URINE DNA AND CLEAVAGE WITH METHYLATION-SENSITIVE RESTRICTION ENZYMES IN ONE STEP

[0253] In this example, the preparation of methylome libraries from cell-free urine DNA by ligation of hairpin oligonucleotide adaptor comprising deoxy-uridine as described in Example 3 is combined with the simultaneous cleavage with a mix of methylation-sensitive restriction enzymes in a single step.

[0254] Cell-free DNA was isolated from urine of healthy donors collected in 50 ml Falcon tubes and stabilized for storage by adding 0.1 volume of 0.5 M EDTA. Urine samples were centrifuged at 1,800 x g for 10 min at ambient temperature to sediment cells, and supernatant was transferred carefully to a fresh tube. Equal volume of 6 M solution of guanidine thiocyanate was added to each sample followed by 1/6 vol of Wizard Miniprep resin (Promega catalog # A7141; Madison, WI). DNA was bound to the resin by rotation for 1 hour at ambient temperature. The resin was then sedimented by brief centrifugation at 500 x g and loaded on Wizard minicolumns (Promega catalog #A7211; Madison, WI) using syringe barrel extensions after carefully decanting out the supernatant. Resin was washed with 5 ml of wash buffer (Promega catalog # A8102; Madison, WI) using Qiagen QIAvac-24 vacuum manifold. Minicolumns were then centrifuged for 2 min at 10,000 x g to remove residual wash buffer and bound DNA was eluted with 50 µl of DNAse-free water at 10,000 x g for 1 min. Eluted DNA was buffered by adding 0.1 vol of 10x TE-L buffer and quantified on a fluorescent spectrophotometer using Pico Green (Molecular Probes; Carlsbad, CA) and l phage D.

[0255] Aliquots of 200 ng of cell-free DNA were artificially methylated by incubation in 20 µl of NEBuffer 2 (New England Biolabs; Beverly, MA) with 4 units of M.SssI CpG methylase (New England Biolabs; Beverly, MA) in the presence of 160 µM SAM for 1 hour at 37°C.

[0256] Twenty five nanograms of methylated or non-methylated DNA were incubated in 1x NEBuffer 4 (New England Biolabs; Beverly, MA) comprising 0.35 units of T4 DNA polymerase (New England Biolabs; Beverly, MA); 1.5 µM of dU-Hairpin Adaptor (Table

IDT Exhibit 1020.7599

I, SEQ ID NO:7); 0.5 units of UDG (New England Biolabs; Beverly, MA); 400 units of T4 DNA ligase (New England Biolabs; Beverly, MA); 30 μM dNTPs, 0.75 mM ATP; 75 μg/ml BSA; 16.7 units of AciI; 16.7 units of HhaI; 8.3 units each of BstUI, HpaII, and Hinp1I (New England Biolabs; Beverly, MA) in a final volume of 20 μl for 1 hour at 37°C. A second aliquot of 25 ng of methylated or non-methylated DNA was incubated in parallel with all the ingredients described above but without the restriction enzymes ("uncut" control).

[0257] One half of each sample (12.5 ng) was then amplified by quantitative PCR in an exemplary reaction mix comprising 1 x Titanium *Taq* reaction buffer (Clontech; Mountain View, CA); 200 μM of each dNTP; fluorescein calibration dye (1:100,000) and SYBR Green I (1:100,000); 1 μM universal primer M_U-1 (Table I, SEQ ID NO:8); 4% DMSO; 200 μM 7-deaza-dGTP (Sigma); and 5 units of Titanium *Taq* polymerase (Clontech) in a final volume of 75 μl. Samples were pre-heated at 72°C for 15 min followed by 95°C for 5 min and 12 cycles at 94°C for 15 sec and 65°C for 2 min on an I-Cycler real-time PCR instrument (Bio-Rad; Hercules, CA). Amplified libraries were purified using MultiScreen PCR cleanup system (Millipore) and quantified by optical density reading.

[0258] Methylation analysis of promoter sites was performed by real-time PCR using aliquots of 160 ng of each digested or non-digested amplified library DNA incubated in exemplary reaction mixtures comprising the following: 1x Titanium *Taq* reaction buffer (Clontech; Mountain View, CA); 200 μM of each dNTP; 4% DMSO; 0.5 M betaine; FCD (1:100,000) and SYBR Green I (1:100,000); 200 nM each forward and reverse primer (Table II, SEQ ID NO:11 and SEQ ID NO:12 for GSTP-1 promoter, SEQ ID NO:13 and SEQ ID NO:14 for MDR-1 promoter, SEQ ID NO:9 and SEQ ID NO:10 for EDNRB promoter, and SEQ ID NO:15 and SEQ ID NO:16 for PTGS-2 promoter); and approximately 1.5 units of Titanium *Taq* polymerase (Clontech; Mountain View, CA) in a final volume of 15μl at 95°C for 3 min followed by 50 cycles at 94°C for 15 sec and 68°C for 1 min.

[0259] FIG. 27 shows PCR amplification curves of specific promoter sites in amplified libraries prepared from methylated (M) or non-methylated (N) urine DNA in the presence (+) or in the absence (-) of methylation-sensitive restriction enzymes. As expected, promoter sites from non-methylated cleaved DNA (N+) amplified with significant (at least 10 cycles) delay as compared to uncut DNA (N-, M-) for all four promoter sites tested. On the other hand, methylated DNA was completely refractory to cleavage (M+). These results demonstrate

IDT Exhibit 1020.7600

that the method disclosed in the present invention can be applied as a simple one-step non-invasive high-throughput diagnostic procedure for detection of aberrant methylation in cancer.

IDT Exhibit 1020.7601

TABLE I OLIGONUCLEOTIDE SEQUENCES

No	Code	Sequence 5' – 3'
1.	EcoNI Ext Adaptor	CCAAACACACCCAACACA<u>CCT</u>AAAAA<u>AGG</u>TGT (SEQ ID NO:1)*
2.	K_U	TGTGTTGGGTGTGTTTGG (SEQ ID NO:2)
3.	HEG Adaptor	AAGAGAGAGGGAAGGAAGAA/HEG/AACTTCCTTCCCTCTC TCTT (SEQ ID NO:3)**
4.	Y_U	CTTCCTTCCCTCTCTCTT (SEQ ID NO:4)
5.	Hairpin Adaptor	CCAAAC ACACCCAACACAAAAAGTGTTG (SEQ ID NO:5)
6.	Hairpin Ribo Adaptor	CCAAAC ACACCCAACACrArArArArAGTGTTG (SEQ ID NO:6)***
7.	dU Hairpin Adaptor	TGTGTTGGGdUGdUGTGTGGdUdUdUdUdUCCACACACAC CCAACACA (SEQ ID NO:7)****
8.	M_U	CCACACACACCCAACACA (SEQ ID NO:8)
9.	EDNRB	GGAGGAGTCTTTCGAGTTCAA (SEQ ID NO:9) Forward CGGGAGGAATACAGACACGTCTT (SEQ ID NO:10) Reverse
10.	GSTP-1	GGAAAGAGGGAAAGGCTTC (SEQ ID NO:11) Forward CCCCAGTGCTGAGTCACGG (SEQ ID NO:12) Reverse
11.	MDR-1	GGGTGGGAGGAAGCATCGTC (SEQ ID NO:13) Forward GGTCTCCAGCATCTCCACGAA (SEQ ID NO:14) Reverse
12.	PTGS-2	AGAACTGGCTCTCGGAAGCG (SEQ ID NO:15) Forward GGGAGCAGAGGGGGTAGTC (SEQ ID NO:16) Reverse
13.	dUT7 Hairpin Adaptor	TGTGTTGGGdUGdUGTGTGGdUdUdUdUdUATTTAATACG ACTCACTATAGGGAGACCACACACACCCAACACA (SEQ ID NO:17)****
14.	dU_N.BbvC IB_Nick Hairpin Adaptor	TGTGTTGGGdUGdUGTGTGGdUdUdUdUdUATTTAATACG AC<u>CCTCAGC</u>ACCACACACACCCAACACA (SEQ ID NO:18)*****

IDT Exhibit 1020.7602

15. dU_N.BbvC IB_Nick ATTTAATACGACCCTCAGCACCAC (SEQ ID NO:19)***** Primer

* Underlined sequences indicate EcoNI palindrome

**HEG = hexa-ethyleneglycol

*** rA = Adenosine

**** dU = deoxy-Uridine

***** Underlined sequence indicates N.BbvC IB nicking endonuclease recognition site

TABLE II. HUMAN STS MARKERS USED FOR REPRESENTATION ANALYSIS BY QUANTITATIVE REAL-TIME PCR

STS #	UniSTS Database Name*
4	SHGC-149956
8	csnpmnat1-pcr1-1
35	SHGC-146602

* Unique names of STS marker sequences from the National Center for Biotechnology Information UniSTS database. Sequences of the STS regions as well as the forward and backward primers used in quantitative real-time PCR can be found in the UniSTS database at the National Center for Biotechnology Information website.

EXAMPLE 6: ENZ-O-MIX DNA LIBRARYPREPARATION AND SIMULTANEOUS AMPLIFICATION BY TRANSCRIPTION

[0260] This example demonstrates an exemplary preparation of a library from lambda phage BstEII restriction digest by ligation of hairpin oligonucleotide adaptor comprising deoxy-uridine and T7 promoter sequence combined with the simultaneous isothermal amplification by *in vitro* transcription with T7 RNA polymerase (see FIG. 6A and FIG. 19A, Enz-O-Mix 2). The adaptor oligonucleotide used in this example (Table I, SEQ ID NO:9) is based on the adaptor sequence described in Example 3 (Table I, SEQ ID NO:8), except that it has the consensus T7 phage promoter sequence incorporated in the loop region. Four enzymatic activities are present in the mix: T4 DNA ligase, T4 DNA polymerase, Uracil-DNA glycosylase (UDG), and T7 RNA polymerase. UDG creates abasic sites by destabilizing the 5' end of the adaptor stem. T4 DNA polymerase generates blunt ends at the restriction fragments leaving intact 5' phosphate groups that are necessary for ligation to the free 3' end of the adaptor by the T4 DNA ligase. T4 DNA polymerase then extends the 3' ends of the newly ligated restriction fragments into the adaptor, thereby displacing the 5' end of the adaptor's stem until an abasic site created by UDG is reached that stops further extension. By doing so, the T4 DNA polymerase

IDT Exhibit 1020.7603

generates a fully functional double-stranded T7 promoter, and *in vitro* transcription is initiated by T7 RNA polymerase.

[0261] One hundred nanograms of a lambda phage BstEII restriction DNA fragments (NEB Cat. # M0208S; New England Biolabs; Beverly, MA) were incubated in 1x NEBuffer 4 (NEB) with 2 μM adaptor oligonucleotide comprising deoxy-uridine and T7 promoter sequence (Table I, SEQ ID NO:9); 1 mM ATP; 40 μM each dNTP; 500 μM each rNTP; 200 μg/ml BSA; 5.6 mM DTT; 1 unit of uracil-DNA glycosylase (New England Biolabs; Beverly, MA); 0.18 unit of T4 DNA polymerase (New England Biolabs; Beverly, MA); 600 units of T4 DNA ligase (New England Biolabs; Beverly, MA); and 37.5 units of T7 RNA polymerase (Epicentre; Madison, WI) for 2 hours at 37°C in a final volume of 15 μl. A negative control containing no T4 DNA ligase was also run in parallel. Products of the library amplification were analyzed by gel electrophoresis on 1.5% agarose gel after staining with Sybr Gold (Molecular Probes; Carlsbad, CA).

[0262] FIG. 28 shows the products of an exemplary *in vitro* transcription amplification. As shown, in the presence of all four enzymatic activities the lambda restriction fragments are converted into a library amplified as RNA products, whereas in the absence of T4 DNA ligase no accumulation of RNA occurs. This demonstrates that library preparation and isothermal amplification can be combined into a single-step process.

EXAMPLE 7: ENZ-O-MIX DNA LIBRARY PREPARATION AND SIMULTANEOUS IMMOBILIZATION ON SOLID SUPPORT

[0263] This embodiment is illustrated in FIG. 29 and describes the one-step Enz-O-Mix attachment process for a stem-loop oligonucleotide with a non-replicable linker that is accompanied by immobilization of the synthesized library to the surface of a vessel where the reaction occurs (for example, a tube, micro-plate well, glass slide, micro-beads, *etc.*). The exemplary reaction mix comprises HMW DNA; a stem-loop oligonucleotide with 3' recessed, 3' protruding or blunt end (FIG. 4), and a non-replicable linker somewhere in the central part of the oligonucleotide; a DNA fragmentation endonuclease, such as restriction enzyme, DNase I, benzonase, methylation-specific nuclease McrBC, apoptotic endonuclease, *etc.*; a 3'proofreading DNA polymerase (Klenow fragment of the DNA polymerase I, T4 DNA polymerase, *etc.*); T4 DNA ligase; Enz-O-Mix Universal Buffer; ATP; and dNTPs. In specific embodiments, the reaction occurs inside a tube or a micro-plate well containing a hybridization-capture oligonucleotide (HCO) covalently attached to the inner surface of the reaction vessel.

IDT Exhibit 1020.7604

Oligonucleotide HCO has a sequence that is complementary to a portion of the stem-loop oligonucleotide located between the 5' end and the non-replicable linker. Four enzymatic reactions and one hybridization reaction are taking place as follows: (1) DNA fragmentation by a nuclease(s); (2) "polishing" of the DNA ends and the stem-loop oligonucleotide double-stranded stem-region; (3) ligation of the oligonucleotide 3' end to the 5' phosphate of the DNA, leaving a nick between the 3' end of DNA and the 5' end of the oligonucleotide double-stranded stem-region; (4) polymerase extension of the 3' DNA end that propagates toward the end of stem-loop oligonucleotide, displaces the 5' portion of the stem-loop oligonucleotide, stops somewhere within the loop or close to the loop region, at the replication block, and generates single stranded overhangs at the 5' ends of DNA molecules; and (5) immobilization of DNA fragments through a hybridization of the generated 5' overhangs to the surface-attached capture oligonucleotide HCO.

[0264] Library immobilization can be non-covalent (as shown on FIG. 29A), or covalent (as shown on FIG. 29B). The first case (non-covalent immobilization) relies solely on the hybridization between the 5' overhang (produced from the stem-loop oligonucleotide during its ligation to DNA ends and subsequent replication) and the capture oligo HCO. The second case (covalent immobilization) involves hybridization as an initial step, but it also involves additional enzymatic steps, such as site-specific cleavage within the single stranded 5' overhang and ligation of the 5' end of DNA to the ligation-capture oligonucleotide construct LCO. The purpose of the cleavage reaction is to generate a 5' phosphate group for the ligation reaction. Such cleavage can be achieved, for example, by incubation with the USER enzyme (a mixture of Uracil DNA glycosylase (UDG) and the DNA glycosylase-lyase Endonuclease VIII; New England Biolabs, Beverly, MA). In that case, UDG catalyses the excision of a uracil base (located somewhere at the 5' portion of the stem-loop oligonucleotide), forming an abasic (apyrimidinic) site while leaving the phosphodiester backbone intact. The lyase activity of Endonuclease VIII breaks the phosphodiester backbone at the 3′ and 5′ sides of the abasic site leaving a 5′ phosphate. A generated phosphate group at the end of the 5' overhang is then ligated by a DNA ligase to the 3' end of the ligation-capture oligonucleotide construct LCO (see FIG. 29B).

[0265] Library synthesis and immobilization can occur within the specifically designed tubes, micro-well plates, on the surface of micro-slides, or micro-beads (FIG. 30). In all of these cases, the hybridization capture oligonucleotide HCO, or the covalent ligation capture

IDT Exhibit 1020.7605

oligo construct LCO, is covalently attached to the inner surface of reaction tubes and plates, spotted on the glass (plastic) surface to create a micro-array, or cover the surface of micro-beads.

[0266] The synthesized and surface-immobilized Whole Genome Library (as described above), or Whole Methylome Library (if the methylation-sensitive restriction nucleases, such as Aci I, Acc II, Asp LE I, Ava I, Bce AI, Bsa HI, Bsh 1236 I, Bsi E1, Bsi SI, Bst FN I, Bst HH I, Bst UI, Cfo I, Hap II, Hga I, Hha I, HinP1 I, Hin 6I, Hpa II, Hpy 99I, Hpy CH4 IV, Hsp AI, Mvn I, and Ssi I,, or the methylation-specific nucleases, such as McrBC, are included into the Enx-O-Mix or used after the library synthesis) can be further processed enzymatically, washed, and amplified or stored in an immobilized format. FIG. 31 illustrates a specific but exemplary use of the one-step library synthesis and immobilization in a hypothetical fluidic device. The process starts by a library synthesis (I), is continued by a library immobilization at the bottom of a reaction vessel (II), and is completed by removal of the first reagent mix, washing, and introduction of a new reagent mix (for example, PCR components).

EXAMPLE 8: ENZ-O-MIX DNA LIBRARY PREPARATION AND SIMULTANEOUS AMPLIFICATION BY STRAND DISPLACEMENT SYNTHESIS

[0267] This example demonstrates an exemplary preparation of a library from cell-free urine DNA by ligation of hairpin oligonucleotide adaptor comprising deoxy-uridine and N.BbvC IB nicking endonuclease recognition site located at the loop region as illustrated in FIG. 16A followed by isothermal amplification with a strand-dispalcing DNA polymerase (see FIG. 6A and FIG. 19A, Enz-O-Mix 3). The adaptor oligonucleotide used in this example (Table I, SEQ ID NO:18) is based on the adaptor sequence described in Example 3 (Table I, SEQ ID NO:8), except that it has recognition sequence 5'-CCTCAGC-3' for N.BbvC IB nicking endonuclease in the loop region. In the first step, a stem-loop adaptor attachment is accomplished by using a mix (FIG. 18, Master Mix III) with 3 enzymatic activities: T4 DNA ligase, T4 DNA polymerase, and Uracil-DNA glycosylase. UDG creates abasic sites by destabilizing the 5' end of the adaptor stem. T4 DNA polymerase generates blunt ends at the restriction fragments leaving intact 5' phosphate groups that are necessary for ligation to the free 3' end of the adaptor by the T4 DNA ligase. T4 DNA polymerase then extends the 3' ends of the newly ligated restriction fragments into the adaptor, thereby displacing the 5' end of the adaptor's stem until an abasic site generated by UDG is reached that stops further extension. By doing so, the T4 DNA polymerase generates a functional double-stranded N.BbvC IB nicking endonuclease recognition site. In the second step, following thermal inactivation of T4 DNA polymerase and T4 DNA ligas, the mix is supplemented with two additional enzymatic activities,

IDT Exhibit 1020.7606

N.BbvC IB nicking endonuclease generating single-starnded nicks at the adaptor (and internal) sites, and the strand-displacing Klenow fragment of DNA polymerase I, initiating strand displacement from the nicks. The newly synthesized strand is extended until a nick or abasic site at the template strand is encountered. In order to fill-in a second strand and to recreate intact N.BbvC IB recognition sites, an oligonucleotide primer comprising the adaptor's loop sequence at its 5'-end plus five bases complementary to the adaptor's stem at its 3'-end (Table I, SEQ ID NO:19) is also added to the mix. As a result, a self-sustained isothermal DNA amplification process is accomplished.

[0268] Cell-free DNA was isolated from urine of healthy donors as described in Example 5. Aliquots of 50 nanograms of DNA were incubated in 1x NEBuffer 4 (New England Biolabs; Beverly, MA) with 2 μM adaptor oligonucleotide comprising deoxy-uridine and N.BbvC IB nicking endonuclease recognition site (Table I, SEQ ID NO:18); 1 mM ATP; 40 μM each dNTP; 200 μg/ml BSA; 1 unit of uracil-DNA glycosylase (New England Biolabs; Beverly, MA); 0.18 unit of T4 DNA polymerase (New England Biolabs; Beverly, MA); and 600 units of T4 DNA ligase (New England Biolabs; Beverly, MA) for 1 hour at 37°C in a final volume of 15 μl. A negative control containing no T4 DNA ligase was run in parallel. Samples were heated at 65°C for 15 min, cooled to 37°C and supplemented with 10 units of N.BbvC IB nicking endonuclease (New England Biolabs; Beverly, MA); 27 ng/μl of *E. coli* Single Stranded Binding Protein (SSB protein, USB Corporation, Clevelang, OH); 5 units of Klenow Exo- fragment of DNA polymerase I (USB Corporation, Clevelang, OH); 0.5 μM of dU_N.BbvC IB_Nick Primer (Table I, SEQ ID NO:19); and 200 μM of each dNTP in a final volume of 30 μl of 1x NEBuffer 4 (New England Biolabs; Beverly, MA). To test the effect of additives on the representation of GC-rich sequences such as promoter sites, reactions containing 4% of dimethyl sulfoxide (DMSO, Sigma-Aldrich, St. Louis, MO) and 7deaza-dGTP (Roche Diagnostics, Indianapolis, IN) were run in parallel. A control reaction containing no N.BbvC IB nicking endonuclease was also included. Samples were incubated for 1 to 3 hours at 37 °C and reactions were stopped by adding 50 mM of EDTA. Products of the library amplification were analyzed by gel electrophoresis on 20 % acrylamide TBE gel (Invitrogen Corporation, Carlsbad, CA), after staining with Sybr Gold (Molecular Probes; Carlsbad, CA).

[0269] FIG. 29 shows the products of an exemplary isothermal DNA amplification after 1 hour incubation. As shown, in the presence of all five enzymatic activities the lambda

IDT Exhibit 1020.7607

restriction fragments are converted into a library amplified as DNA products, whereas in the absence of T4 DNA ligase or N.BbvC IB nicking endonuclease no accumulation of DNA occurs.

EXAMPLE 9: INTEGRATED WHOLE GENOME AMPLIFICATION IN A CLOSED-TUBE REACTION USING DEGRADABLE HAIRPIN ADAPTOR COMPRISING DEOXY-URIDINE

[0270] Traditionally, preparation of DNA libraries and their subsequent amplification by PCR involves multiple separate processes and different buffer systems. Previous embodiments and examples of this invention have demonstrated that Enz-O-Mx approach can substantially simplify these processes and reduce number of steps necessary for WGA and WMA library preparation to one step. However, in previous Examples the Enzo-O-Mix DNA libraries were first, prepared in small volume (15µl) of buffer A (NEB 4), and then amplified by PCR in a larger volume (75 ml) of buffer B (Titanium Taq reaction buffer), FIG. 30A. In all of these cases, Enz-O-Mix library synthesis / amplification is performed in two operational steps: step 1- a tube contatining DNA is supplemented with the library (WGA or WMA) synthesis reagents and incubated at 37 °C for 1 h, and step 2 – a tube is opened, supplemented with PCR amplification buffer / reagents and subjected to temperature cycling. Such an approach is referred to herein as a two-step, opened-tube protocol.

[0271] This example introduces further simplification in preparation and PCR-mediated amplification of WGA / WMA libraries when all the necessary synthesis / amplification reagents are introduced into the tube prior reaction, and library synthesis and subsequent amplification occur in the same volume and the same buffer without opening the tube within a pre-programmed thermocycler (FIG. 30B). Such an approach is referred to herein as an integrated, one-step, closed-tube protocol. Due to its simplicity and lack of any human intervention, the integrated, one-step, closed-tube Enz-O-Mix DNA amplification can be easily automated and used for high-throughput research applications and clinical diagnostics.

[0272] FIG. 31 shows a one possible temperature profile (b) and an envisioned DNA accumulation during amplification (a), where incubation at 37 °C for 1 h (DNA library synthesis) is followed by heating at 95 °C for 10 min (Taq DNA polymerase activation and stem-loop adaptor → universal PCR primer conversion), and then by thermocycling between 65°C and 95 °C for 1 h (PCR-mediated whole genome or whole methylome amplification).

[0273] FIG. 32 shows biochemical and physicochemical reactions involved in the transformation of the stem-loop adaptor oligonucleotide into a functional universal PCR primer

IDT Exhibit 1020.7608

adequate for efficient amplification of synthesized WGA / WMA libraries. Dual utilization of the stem-loop oligonucleotide as adaptor and PCR primer eliminate a necessity to introduce into reaction an additional single stranded oligo-primer (due to the 3' profreading activity of a DNA polymerase involved in the library synthesis process such primer should contain at least several nuclease-resistant bases at its 3' terminus).

[0274] In this example, a WGA process is described wherein a hairpin oligonucleotide adaptor comprising deoxy-uridine described in Example 3 is ligated *via* its free 3' end to the 5' phosphates of DNA restriction fragments generated from intact genomic DNA in a single enclosed container exemplary reaction mix comprising 6 exemplary enzymatic activities: AluI restriction endonuclease, RsaI restriction endonuclease, T4 DNA ligase, T4 DNA polymerase, a hot-start Taq DNA polymerase, and Uracil-DNA glycosylase (UDG). In an initial isothermal incubation UDG catalyses the release of free uracil and generates abasic sites in the adaptor's loop region and the 5' half of the hairpin. AluI and RsaI digest the target DNA into restriction fragments. T4 DNA polymerase generates blunt ends of the restriction fragments. T4 DNA ligase generates a phosphodiester bond between the 3' ends of the adaptor molecules and the 5' phosphates of the restriction fragments. T4 DNA polymerase extends the free 3' end of the restriction fragments using as template the ligated 3' end of the hairpin stem until an abasic site is reached that serves as a replication stop (FIG. 10B). Samples are heated to 72°C to inactivate all thermo-labile enzymes and to activate Taq polymerase, then to 95°C to degrade the abasic sites of the adaptor and to generate active primer from the remaining intact strand of the adaptor that is free of abasic sites. The resulting library is finally amplified by thermal cycling. The entire process takes place in a single enclosed reaction container under programmed temeparture control algorithm and without any intermediate liquid handling.

[0275] Ten nanogram aliquots of human genomic DNA isolated from the peripheral blood of a healthy donor by standard procedures were incubated in an exemplary reaction mix comprising: 1x Titanium Taq buffer (Clontech, Mountain View, CA); 0.6 µl of Titanium Taq (Clontech, Mountain View, CA); 2 µM dU Hairpin Adaptor oligonucleotide (Table I, SEQ ID NO:7); 1 mM ATP; 200 µM dNTPs; 0.36 Units of T4 DNA polymerase; 1200 units of T4 DNA ligase; 2 Units of UDG; 10 units each of AluI and RsaI restriction endonucleases (New England Biolabs; Beverly, MA), and 1:10,000 dilutions of fluorescein and SybrGreen I (Molecular Probes; Carlsbad, CA) in a final volume of 30 µl. To study the effect of Mg^{++} ions and DMSO, reactions supplemented with $MgCl_2$ at final concentrations of 5 mM and

IDT Exhibit 1020.7609

7.5 mM and with DMSO at a final concentration of 4% were also included (It should be noted that the 1x Titanium Taq buffer contains 3.5 mM $MgCl_2$). Reactions were incubated at 37°C for 1 hr, followed by 72°C for 10 min, 95 °C for 10 min, and 11 cycles of 94 °C for 20 sec and 65 °C for 2 min. on i-Cycler real-time PCR instrument (Bio-Rad; Hercules, CA). FIG. 33 shows the amplification curves of the resulting libraries. As shown, Mg^{++} concentrations of 5 mM and 7.5 mM in the reaction buffer supported the enzymatic activities present in the mix better than the basic Titanium Taq buffer containing 3.5 mM $MgCl_2$. The presence of 4% DMSO did not have significant effect on the WGA amplification when higher Mg^{++} concentrations were applied.

EXAMPLE 10: HOT START PCR USING DEGRADABLE STEM-LOOP PRIMERS

[0276] Specificity and the ability to amplify a single DNA or RNA target is one of the most important requirements for application of PCR in molecular diagnostics. Hot start PCR protocol was introduced to reduce the non-specific primer/template and primer/primer annealing events that occur at lower temperatures and subsequently result in non-specific amplification products. Several methods and corresponding commercial products for performing hot start PCR rely upon the physical separation of PCR reagents until the high temperature of the reaction has been reached. Those products include the following:

[0277] 1. Wax beads (Ampliwax PCR Gems, Perkin Elmer) that create a temporary barrier between dNTPs, buffer, and $MgCl_2$ on one side of the wax layer, and DNA template and DNA polymerase on another side.

[0278] 2. Small beads of wax with encapsulated Taq DNA polymerse (Taq Bead Hot Start Polymerase, Promega). In this case Taq DNA polymerase is released when the reaction reaches 60°C.

[0279] 3. Small beads of wax with encapsulated magnesium (StartaSpere, Stratagene).

[0280] 4. Taq DNA polymerase inactivated by antibody (JumStart Taq DNA polymerase, Sigma; TaqStart and TthStart, Clontech; AmpliTaq Gold, PE Biosystems; HotStarTaq, Qiage; *etc.*). Antibody binds to the polymerase and inactivates it at low temperature but denatures and releases the active polymerase at high temperature.

[0281] Example 9 introduced the idea of using a degradable, dU-containing stem-loop oligonucleotide-adaptor as a universal PCR primer (see FIG. 32) for whole genome

IDT Exhibit 1020.7610

amplification. This embodiment extends the use of degradable, dU-containing stem-loop oligonucleotides as "hot start" PCR primers for locus-specific DNA amplification.

[0282] As shown in FIG. 38 and FIG 39, both PCR primers are synthesized in a form of stem-loop ologonucleotides A and B. The 3' portion of the stem-loop oligonucleotides represents the primer sequence (no dU bases); the 5' portion represents the sequence complementary to the primer sequence and comprises one or more dU bases substituting dT bases; and the loop comprises several dT and dU bases.

[0283] PCR reaction is assembled by mixing together DNA, dNTPs, magnesium, PCR buffer, Taq DNA polymerase (or any other thermostable DNA polymerase), and dU-glycosylase (FIG. 38 and FIG. 40). A thermocycler is programmed to have the following conditions:

[0284] a) Incubation step at 37°C for 15 – 30 min that is necessary to convert all dU bases within the stem-loop oligonucleotides A and B into the abasic sites;

[0285] b) Incubation step at 95°C for 10min that is necessary to denature DNA, introduce breaks at the abasic sites within the stem-loop oligonucleotides A and B, and thus release the active primers A and B;

[0286] c) A regular PCR cycling mode that amplifies the DNA region defined by the primer A and B.

[0287] The conversion of the stem-loop oligonucleotides A and B into PCR primers A and B (see FIG. 39) results from several breaks introduced by heating the loop and the 5' portion of the oligonucleotides. At high temperature, the small oligonucleotides originated from the 5' stem regions can not form stable interaction with the remaining intact 3' primer regions and do not affect the PCR process. In general, the position of dU bases within the 5' stem region is dictated by the location of thymines in a DNA sequence, although they can be also introduced at another nucleotide position (thus generating a stem with one ore several mismatched bases). The number of dU bases within the 5' stem and the loop regions can vary from 1 to about 6, and an optimal number can be determined empirically.

[0288] The proposed hot start primer method can be used in PCR and other DNA/RNA amplification methods that utilize thermostable polymerases. It has several advantages over the hot start PCR methods that involve either non-degradable hairpin primers

IDT Exhibit 1020.7611

with short stem [Kaboev, O.K., et al., 2000; Ailenberg, M., and Silverman, M., 2000], or duplex primers [Kong, D., et al., 2004]: a) the hybridization kinetics are not compromised by the presence of a stem region (as it might be in the case of non-degradable stem-loop primer method); b) no inverted repeat is formed at the ends of PCR amplicons that could substantially reduce the efficiency of the PCR amplification process (as it happens in the case of non-degradable stem-loop primer method); and c) oligonucleotides complementary to the priming oligonucleotide (products of the stem-loop oligonucleotide fragmentation) are short, do not form stable interaction with the remaining intact 3' primer regions and, as a result, do not affect the PCR process (as it possible in the duplex primer method). Advantages over other existing hot-start methods include but are not limited to the following: a) no phase separation is necessary and all reaction components are originally present in the same reaction mix simplifying storage and reducing production cost; b) no expensive blocking antibodies are involved; c) no mutagenesis or chemical modification of amino acid reidues in the thermostabile DNA polymerase are required whatsoever.

REFERENCES

[0289] All patents and publications mentioned in the specification are indicative of the level of those skilled in the art to which the invention pertains. All patents and publications are herein incorporated by reference to the same extent as if each individual publication was specifically and individually indicated to be incorporated by reference.

PATENTS AND PATENT APPLICATIONS

[0290] U.S. Patent No. 6,777,187

[0291] U.S. Patent Application No. 10/797,333

[0292] U.S. Patent Application No. 10/795,667

[0293] U.S. Patent Application No. 11/071,864

PUBLICATIONS

[0294] Barns, W.M. PCR amplification of up to 35-kb DNA with high fidelity and high yield from lambda bacteriophage templates. Proc. Natl. Acad. Sci. USA, 91, 2216-2220, (1994).

IDT Exhibit 1020.7612

[0295] Broude, N.E., Stem-loop oligonucleotides: a robust tool for molecular biology and biotechnology. Trends in Biotechnology, 20, 249-256, (2002).

[0296] de Baar, M.P., *et al.* One-tube real-time isothermal amplification assay to identify and distinguish human immunodeficiency using type I subtypes A, B, and C and circulating recombinant forms AE and AG. J. Clin. Microbiol. 39, 1895-1902, (2001).

[0297] Deiman, B., *et al.* Characteristics and applications of nucleic acid sequence-based amplification (NASBA). Mol. Biotechnology, 20, 163-179, (2002); Hill, C.S., Molecular diagnostic testing for infectious diseases using TMA technology. Expert Rev Mol Diagn, 1, 445-455, (2001).

[0298] Goddard, N.L., *et al.* Sequence dependent rigidity of single stranded DNA. Phys. Rev. Lett. 85, 2400-2403, (2000).

[0299] Hamad-Schifferli, K., *et al.* Remote electronic control of DNA hybridization through inductive coupling to an attached metal nanocrystal antenna. Nature, 415, 152-155, (2002).

[0300] Hellyer, T.J., and Naolean, J.G., Strand displacement amplification: a versatile tool for molecular diagnostics. Expert Rev Mol Diagn, 4, 251-261, (2004).

[0301] Kaboev, O.K., *et al.* PCR hot start using primers with the structure of molecular beacons (stem-loop-like structure). Nucleic Acids Res., 28, e94, (2000); Ailenberg, M., and Silverman, M. Controlled hot start and improved specificity in carrying out PCR utilizing touch-up and loop incorporated primers (TULIPS). Bio-Techniques, 29, 1018-1024, (2000).

[0302] Kong, D., et al. PCR hot-start using duplex primers. Biotechnol. Lett. 26, 277-280, (2004).

[0303] Liu, *et al.* Molecular beacons for DNA biosensors with micrometer to submicrometer dimensions. Anal. Biochem. 283, 56-63, (2000).

[0304] Mackay, I.M., *et al.* Real-time PCR in virology. Nucleic Acids Res., 30, 1292-1305, (2002); Elnifro, E.M., *et al.* Multiplex PCR: optimization and application in diagnostic virology. Clin. Microbiol. Rev. 13, 559-570, (2000).

IDT Exhibit 1020.7613

[0305] Riccelli, P.V., *et al.* Hybridization of single-stranded DNA targets to immobilized complementary DNA probes: comparison of stem-loop versus linear capture probes. Nucleic Acids Res., 29, 996-1004, (2001).

[0306] Sokol, D.L., *et al.* Real-time detection of DNA-RNA hybridization in living cells. Proc. Natl. Acad. Sci. USA, 95, 11538-11543, (1998).

[0307] Summerer, D., and Marx, A., A molecular beacon for quantitative monitoring of the DNA polymerase reaction in real-time. Angew. Chem. Int. 41, 3620-3622, (2002).

[0308] Tyagi, S., and Kramer, F.R., Molecular beacons: probes that fluoresce upon hybridization. Nat Biotechnol. 14, 303-308, (1996).

[0309] Tyagi, S., *et al.* Multicolor molecular beacons for allele discrimination. Nat. Biotechnol. 16, 49-53, (1998).

[0310] van Deursen, P.B.H., *et al.* , A novel quantitative multiplex NASBA method: application to measuring tissue factor and CD14 mRNA levels in human monocytes. Nucleic Acids Res , 27, e15, (1999).

[0311] Whitcombe, D., *et al.* Detection of PCR products using self-probing amplicons and fluorescence. Nat. Biotechnol. 17, 804-807, (1999)

[0312] Although the present invention and its advantages have been described in detail, it should be understood that various changes, substitutions and alterations can be made herein without departing from the spirit and scope of the invention as defined by the appended claims. Moreover, the scope of the present application is not intended to be limited to the particular embodiments of the process, machine, manufacture, composition of matter, means, methods and steps described in the specification. As one of ordinary skill in the art will readily appreciate from the disclosure of the present invention, processes, machines, manufacture, compositions of matter, means, methods, or steps, presently existing or later to be developed that perform substantially the same function or achieve substantially the same result as the corresponding embodiments described herein may be utilized according to the present invention. Accordingly, the appended claims are intended to include within their scope such processes, machines, manufacture, compositions of matter, means, methods, or steps.

IDT Exhibit 1020.7614

CLAIMS

What is claimed is:

1. A method of preparing a nucleic acid molecule, comprising:

providing a double stranded nucleic acid molecule; and

attaching one strand of a stem-loop oligonucleotide comprising an inverted repeat and a loop to the double stranded nucleic acid molecule to produce an oligonucleotide-attached nucleic acid molecule.

2. The method of claim 1, wherein the double stranded nucleic acid molecule is a double stranded DNA molecule.

3. The method of claim 1, wherein the attaching is further defined as attaching the oligonucleotide to the double stranded nucleic acid molecule under conditions to produce a non-covalent junction in the oligonucleotide-attached nucleic acid molecule.

4. The method of claim 3, wherein the non-covalent junction comprises a nick, a gap, or a 5' flap structure.

5. The method of claim 1, wherein the attaching is further defined as ligating.

6. The method of claim 1, further comprising displacing one strand of the oligonucleotide from the oligonucleotide-attached nucleic acid molecule by strand displacement or by nick translation polymerization.

7. The method of claim 1, wherein at least part of the oligonucleotide-attached nucleic acid molecule is amplified.

8. The method of claim 7, wherein the amplification comprises polymerase chain reaction.

IDT Exhibit 1020.7615

9. The method of claim 7, wherein the amplification comprises RNA transcription.

10. The method of claim 7, wherein the amplification comprises strand displacement.

11. The method of claim 1, further comprising amplifying the oligonucleotide-attached nucleic acid molecule, wherein at least part of the inverted repeat is excluded from the amplified oligonucleotide-attached nucleic acid molecule.

12. The method of claim 5, wherein the ligating is further defined as comprising:

generating ligatable ends on the double stranded nucleic acid molecule;

generating a ligatable end on the stem-loop oligonucleotide; and

ligating one strand of the ligatable end of the stem-loop oligonucleotide to one strand of an end of the nucleic acid molecule, thereby generating a non-covalent junction in the oligonucleotide-attached nucleic acid molecule.

13. The method of claim 1, further defined as comprising:

generating blunt ends on the nucleic acid molecule;

generating a blunt end on the stem-loop oligonucleotide; and

ligating one strand of the blunt end of the stem-loop oligonucleotide to one strand of a blunt end of the nucleic acid molecule, thereby generating a nick in the oligonucleotide-ligated nucleic acid molecule.

14. The method of claim 1, wherein said stem-loop oligonucleotide comprises a known sequence.

15. The method of claim 1, wherein said stem-loop oligonucleotide comprises a regulatory sequence.

IDT Exhibit 1020.7616

16. The method of claim 2, further comprising:

digesting the DNA molecule with one or more endonucleases to produce DNA fragments;

producing blunt ends on the DNA fragments;

producing a blunt end on the stem-loop oligonucleotide; and

ligating one strand of the blunt end of a stem-loop oligonucleotide to one strand of a blunt end of a DNA fragment, thereby generating a nick in a oligonucleotide-ligated DNA fragment.

17. The method of claim 16, wherein the endonuclease is a restriction endonuclease, DNAse I, or an apoptotic endonuclease.

18. The method of claim 17, wherein the restriction endonuclease is methylation-specific or methylation-sensitive.

19. The method of claim 1, wherein the oligonucleotide-attached nucleic acid molecule comprises a nick having a 3' hydroxy group, wherein there is polymerization from the 3' hydroxy group of at least part of the oligonucleotide-attached nucleic acid molecule.

20. The method of claim 5, wherein the strand displacement or nick translation polymerization is further defined as polymerization that ceases at a non-replicable base or region in the loop or in a region of the stem adjacent to the loop.

21. The method of claim 2, further comprising the step of digesting the double stranded DNA molecule with an endonuclease to generate DNA fragments, wherein the oligonucleotide becomes ligated to one strand of the DNA fragment and wherein polymerization of an oligonucleotide-ligated DNA fragment excludes at least part of the inverted repeat by subjecting the oligonucleotide-ligated DNA fragment to strand displacement or nick translation polymerization that halts at a base or sequence in the loop or in a region of the stem adjacent to the loop.

IDT Exhibit 1020.7617

22. The method of claim 1, wherein the stem-loop oligonucleotide is further defined as comprising a non-replicable base or sequence.

23. The method of claim 22, wherein at least part of the non-replicable base or sequence is present in the loop of the oligonucleotide or in a sequence of the stem adjacent to the loop.

24. The method of claim 22, wherein the non-replicable base or sequence comprises an abasic site or sequence, hexaethylene glycol, or a bulky chemical moiety attached to the sugar-phosphate backbone or the base.

25. The method of claim 24, wherein the abasic site or sequence is introduced by one or more enzymes in the single solution.

26. The method of claim 1, wherein the loop of the stem-loop oligonucleotide comprises at least one deoxy-uridine.

27. The method of claim 6, wherein the strand displacement or nick translation polymerization of the oligonucleotide-attached nucleic acid molecule generates an endonuclease site.

28. The method of claim 2, further comprising the step of digesting the double stranded DNA molecule with an endonuclease to generate DNA fragments, wherein the oligonucleotide becomes ligated to one strand of the DNA fragment to produce an oligonucleotide-ligated DNA fragment, and wherein strand displacement or nick translation polymerization of the oligonucleotide-ligated DNA fragment generates an endonuclease site.

29. The method of claim 27, wherein the endonuclease site is a site-specific restriction endonuclease site and wherein at least part of the inverted repeat is removed by cleavage with said restriction endonuclease.

30. The method of claim 27, wherein the endonuclease site resides within the stem and loop regions.

IDT Exhibit 1020.7618

31. The method of claim 27, wherein the endonuclease site resides within the loop region.

32. The method of claim 27, wherein the endonuclease site resides within the stem region.

33. The method of claim 27, wherein the endonuclease site is a homing endonuclease site.

34. The method of claim 15, wherein RNA transcription is initiated from the regulatory sequence, thereby producing at least one transcribed polynucleotide.

35. The method of claim 34, wherein the transcribed polynucleotide is replicated by a reverse transcriptase.

36. The method of claim 35, wherein reverse transcription is initiated by hybridization of the oligonucleotide complementary to the 3' end of the transcribed polynucleotide.

37. The method of claim 1, wherein the stem-loop oligonucleotide further comprises a recognition sequence for a nicking endonuclease and wherein the oligonucleotide-attached nucleic acid molecule is amplified by strand displacement synthesis and second strand synthesis.

38. The method of claim 37, wherein the recognition sequence for a nicking endonuclease resides at least in part within the stem of the stem-loop oligonucleotide.

39. The method of claim 37, wherein the recognition sequence for a nicking endonuclease resides at least in part within the loop of the stem-loop oligonucleotide.

40. The method of claim 1, wherein a 5' end of the stem-loop oligonucleotide lacks a phosphate.

IDT Exhibit 1020.7619

41. The method of claim 1, wherein the oligonucleotide-attached nucleic acid molecule is further modified.

42. The method of claim 41, wherein the further modification comprises cloning.

43. The method of claim 42, further defined as comprising incorporation of the modified molecule into a vector, said incorporation occuring at ends in the modified molecule generated by endonuclease cleavage within the inverted repeat.

44. The method of claim 1, further defined as occuring in a single suitable solution, wherein the process occurs in the absence of exogenous manipulation.

45. The method of claim 44, further defined as occurring at one temperature.

46. The method of claim 44, wherein the solution comprises one or more of the following:

ligase;
DNA polymerase;
one or more endonucleases;
RNA polymerase;
reverse transcriptase;
RNase H;
deoxy-uridine glycosylase;
nickase;
thermophilic DNA polymerase;
ATP;
rNTPs; and
dNTPs.

47. The method of claim 7, further defined as occuring in a single suitable solution, wherein the process of preparing and amplifying

IDT Exhibit 1020.7620

of a oligonucleotide-attached nucleic acid molecule occurs in the absence of exogenous manipulation.

48. The method of claim 1, wherein the oligonucleotide-attached nucleic acid molecule is immobilized to a solid support.

49. The method of claim 48, wherein the molecule is immobilized non-covalently.

50. The method of claim 48, wherein the molecule is immobilized covalently.

51. The method of claim 2, further comprising the step of digesting the DNA molecule with an endonuclease to generate DNA fragments, wherein the oligonucleotide becomes attached to one strand of the DNA fragment, wherein strand displacement polymerization of the oligonucleotide-ligated DNA fragment and its arrest at a base or sequence in the loop or in a region of the stem adjacent to the loop generates a 5' overhang, and wherein the single stranded 5' overhang hybridizes to a complementary oligonucleotide covalently immobilized to a solid support.

52. A method of preparing a DNA molecule, comprising:

providing an oligonucleotide; and

mixing the oligonucleotide with a double stranded DNA molecule, such that upon attachment of the oligonucleotide to the double stranded DNA molecule, the oligonucleotide attached to the DNA molecule is capable of demonstrating a function that it was incapable of before the attachment to the double stranded DNA molecule, thereby producing an oligonucleotide-attached DNA molecule suitable for modification.

53. The method of claim 52, further defined as occurring in a single suitable solution, wherein the process occurs in the absence of exogenous manipulation.

IDT Exhibit 1020.7621

54. The method of claim 52, wherein the modification comprises site-specific nicking, site-specific double-strand cleavage, transcription, recombination, amplification, polymerization, nick translation, strand displacement, immobilization or a combination thereof.

IDT Exhibit 1020.7622

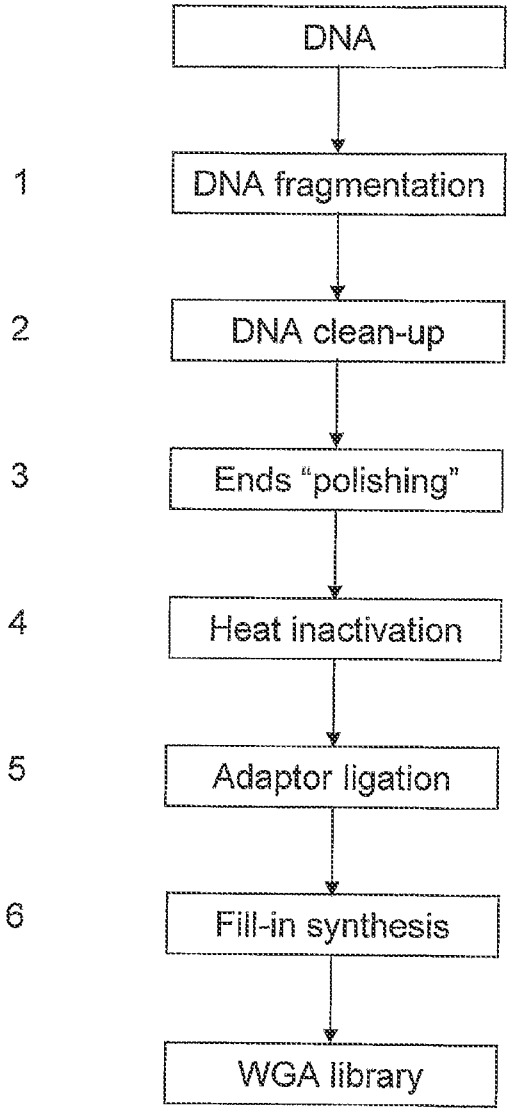

DNA
1
DNA fragmentation
2
DNA clean-up
3
Ends "polishing"
4
Heat inactivation
5
Adaptor ligation
6
Fill-in synthesis
WGA library

FIG. 1A

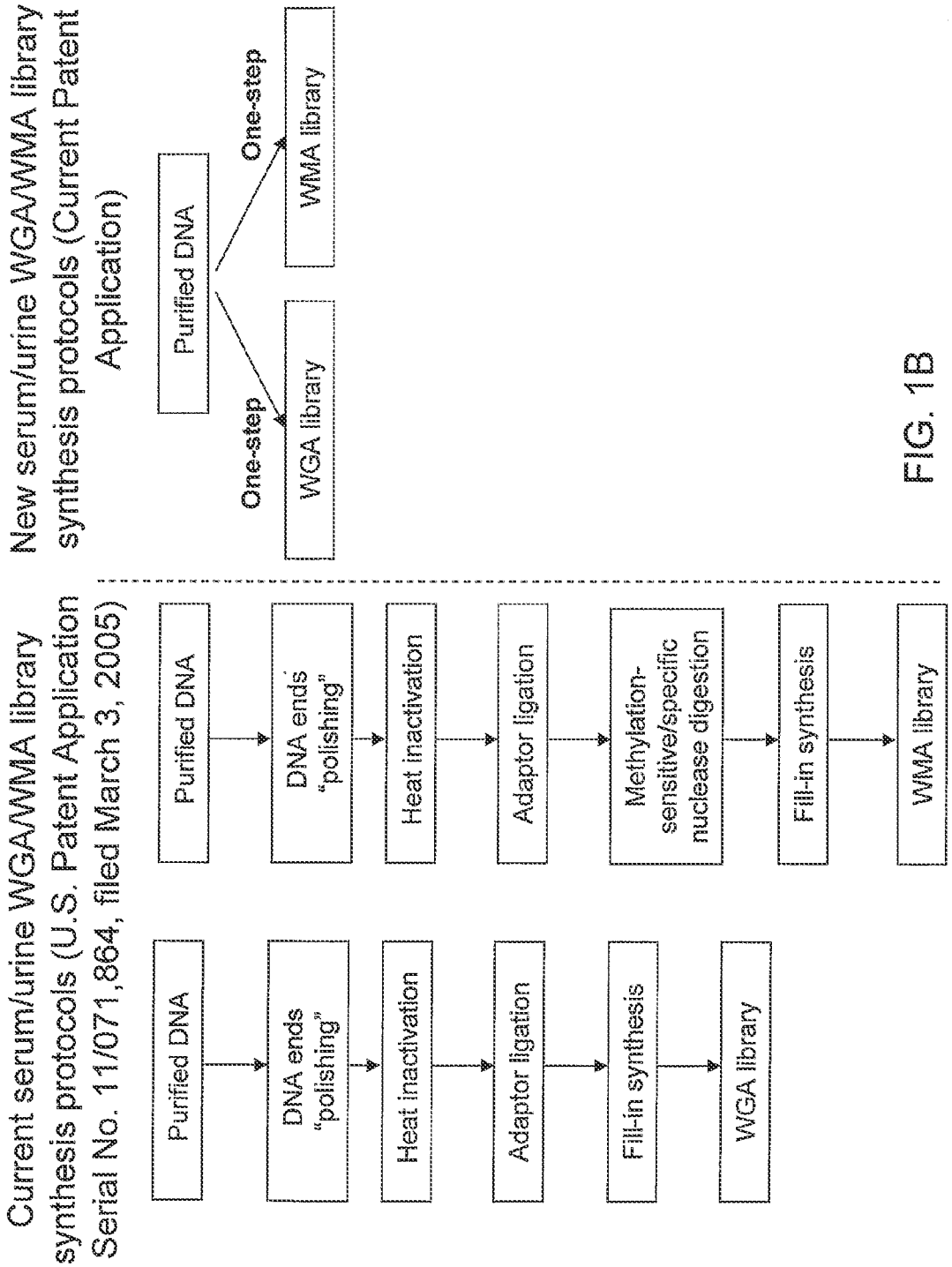

Current serum/urine WGA/WMA library synthesis protocols (U.S. Patent Application Serial No. 11/071,864, filed March 3, 2005)
Purified DNA
DNA ends "polishing"
Heat inactivation
Adaptor ligation
Fill-in synthesis
WGA library
New serum/urine WGA/WMA library synthesis protocols (Current Patent Application)
Purified DNA
One-step
WGA library
One-step
WMA library
Purified DNA
DNA ends "polishing"
Heat inactivation
Adaptor ligation
Methylation-sensitive/specific nuclease digestion
Fill-in synthesis
WMA library

FIG. 1B

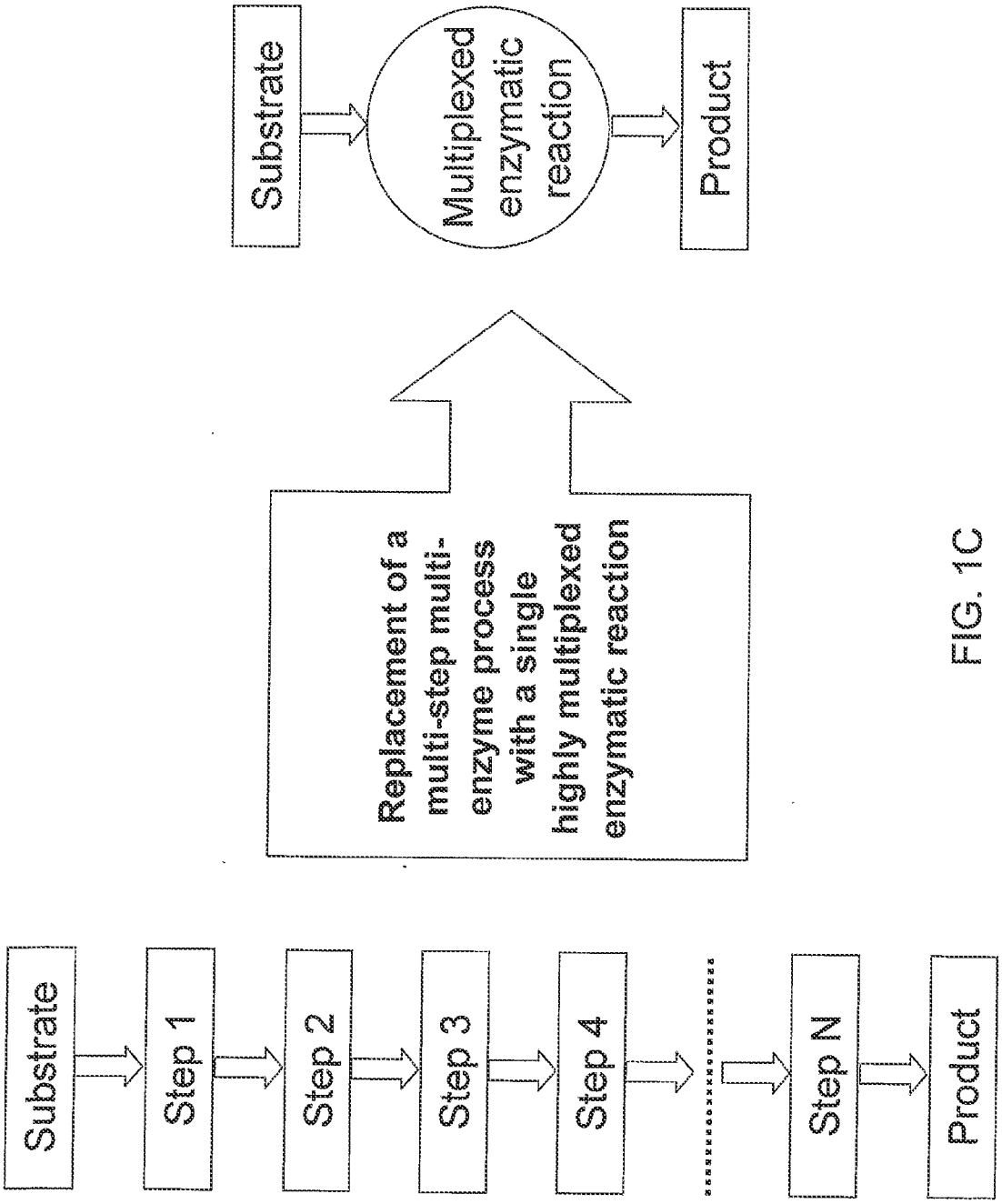

Substrate
Multiplexed enzymatic reaction
Product
Replacement of a multi-step multi-enzyme process with a single highly multiplexed enzymatic reaction
Substrate
Step 1
Step 2
Step 3
Step 4
Step N
Product

FIG. 1C

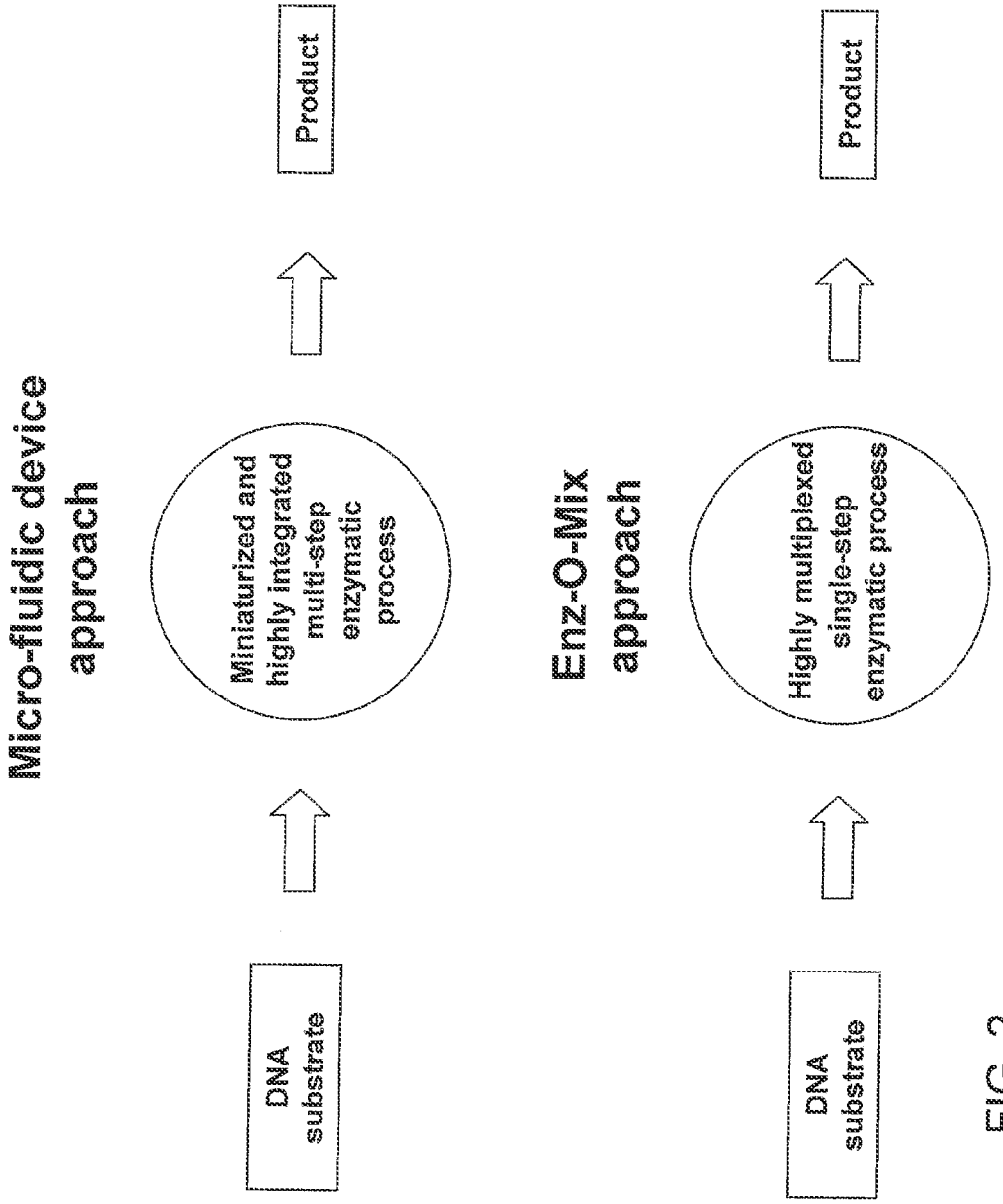

Micro-fluidic device approach
DNA substrate
Miniaturized and highly integrated multi-step enzymatic process
Product
Enz-O-Mix approach
DNA substrate
Highly multiplexed single-step enzymatic process
Product

FIG. 2

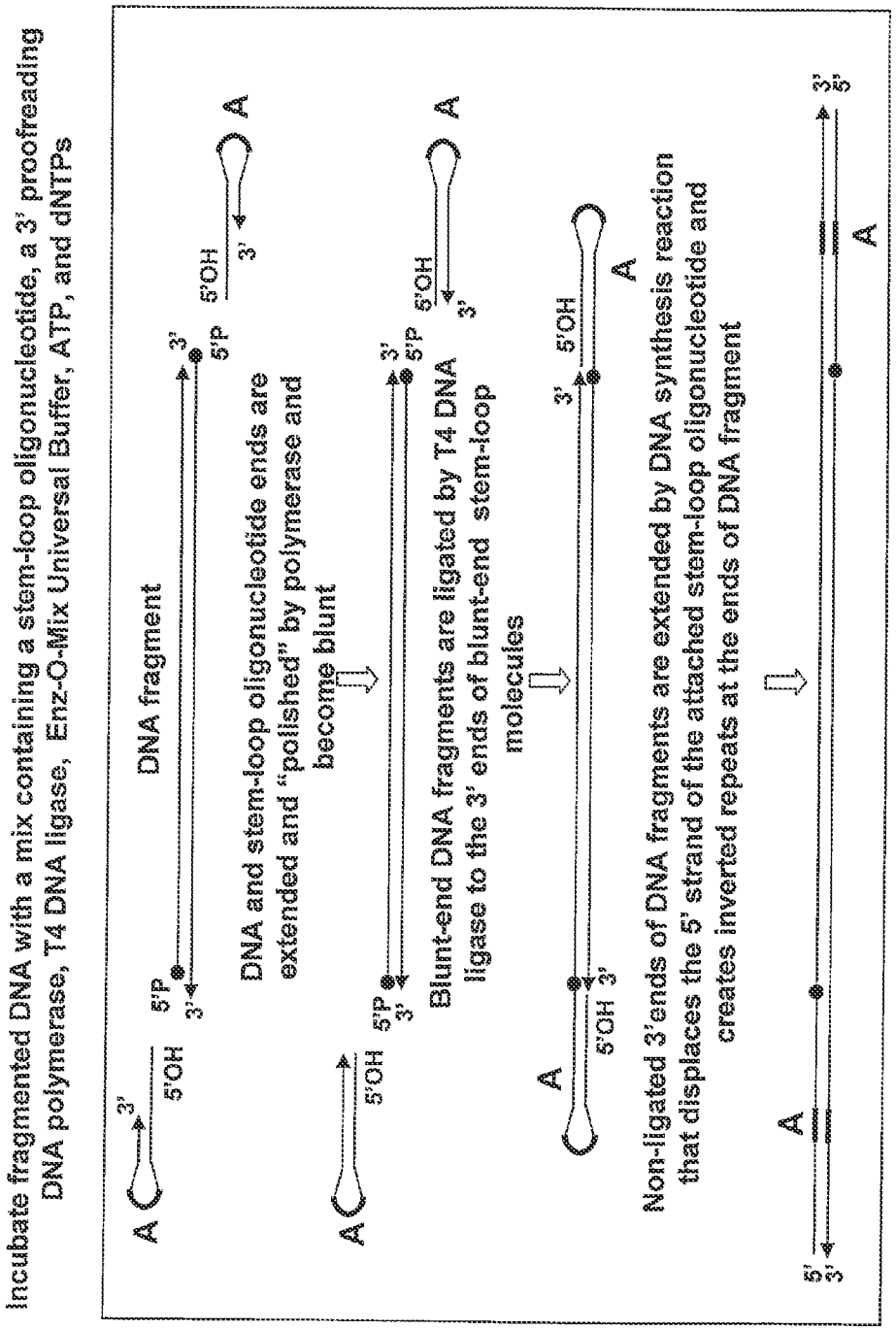

Incubate fragmented DNA with a mix containing a stem-loop oligonucleotide, a 3' proofreading DNA polymerase, T4 DNA ligase, Enz-O-Mix Universal Buffer, ATP, and dNTPs
DNA fragment
DNA and stem-loop oligonucleotide ends are extended and "polished" by polymerase and become blunt
Blunt-end DNA fragments are ligated by T4 DNA ligase to the 3' ends of blunt-end stem-loop molecules
Non-ligated 3'ends of DNA fragments are extended by DNA synthesis reaction that displaces the 5' strand of the attached stem-loop oligonucleotide and creates inverted repeats at the ends of DNA fragment
A
5'OH
3'
5'P

FIG. 3

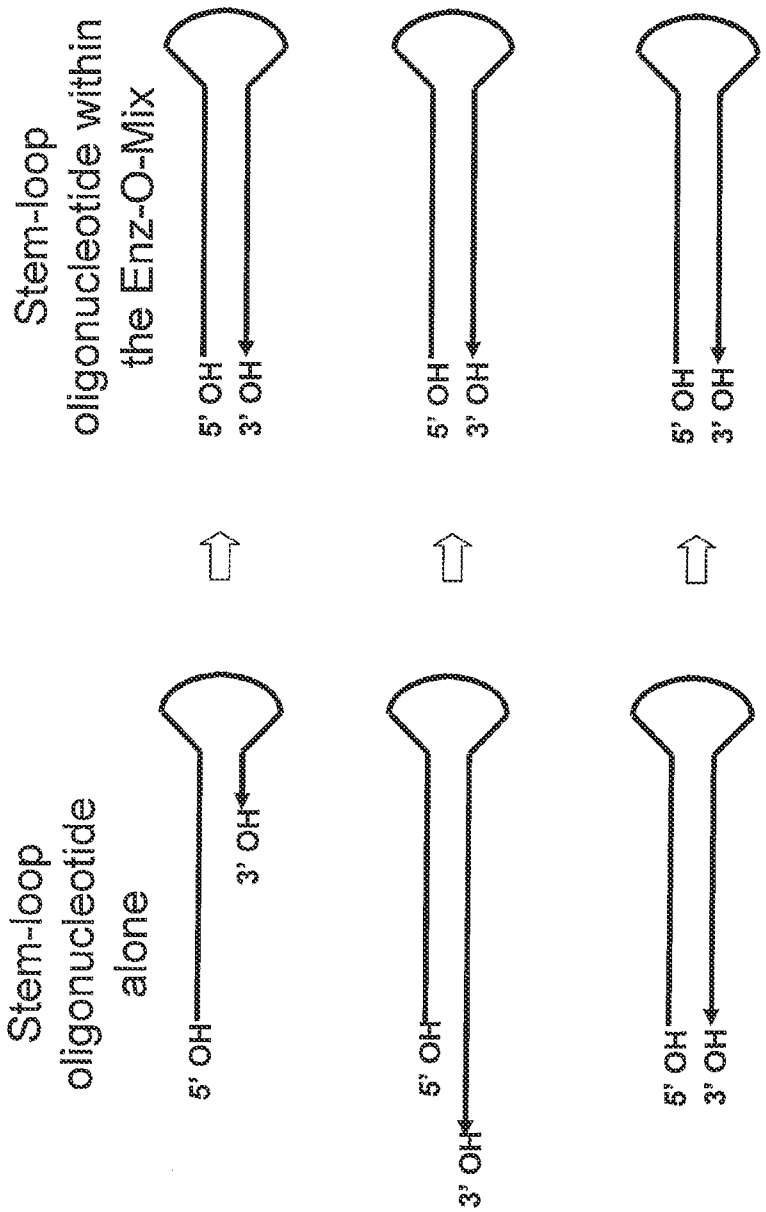

Stem-loop oligonucleotide alone
Stem-loop oligonucleotide within the Enz-O-Mix
5' OH
3' OH
DNA end "polishing" activity

FIG. 4

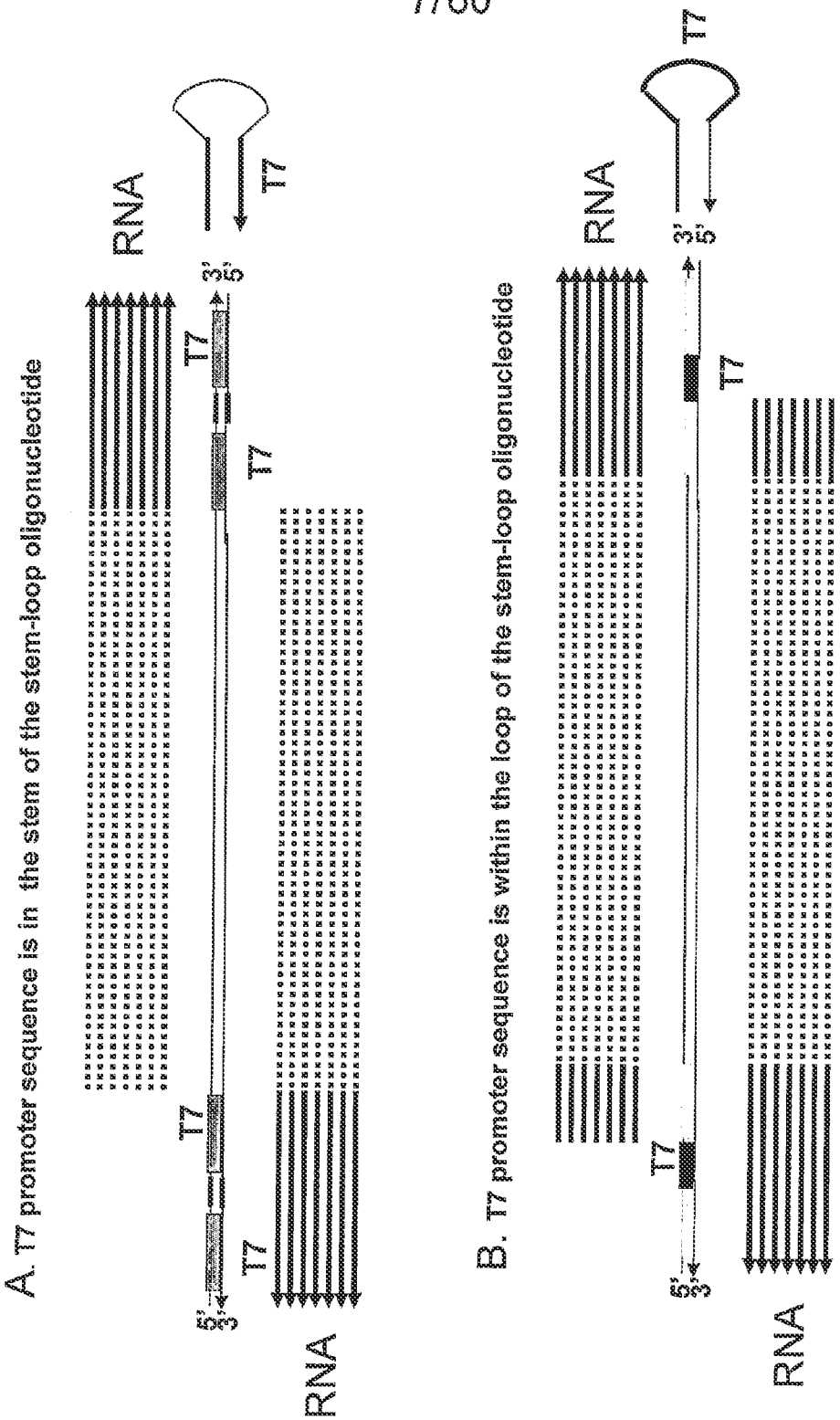

A. T7 promoter sequence is in the stem of the stem-loop oligonucleotide
RNA
T7
3'
5'
T7
T7
T7
5'
3'
RNA
B. T7 promoter sequence is within the loop of the stem-loop oligonucleotide
RNA
T7
3'
5'
T7
T7
5'
3'
RNA

FIG. 5

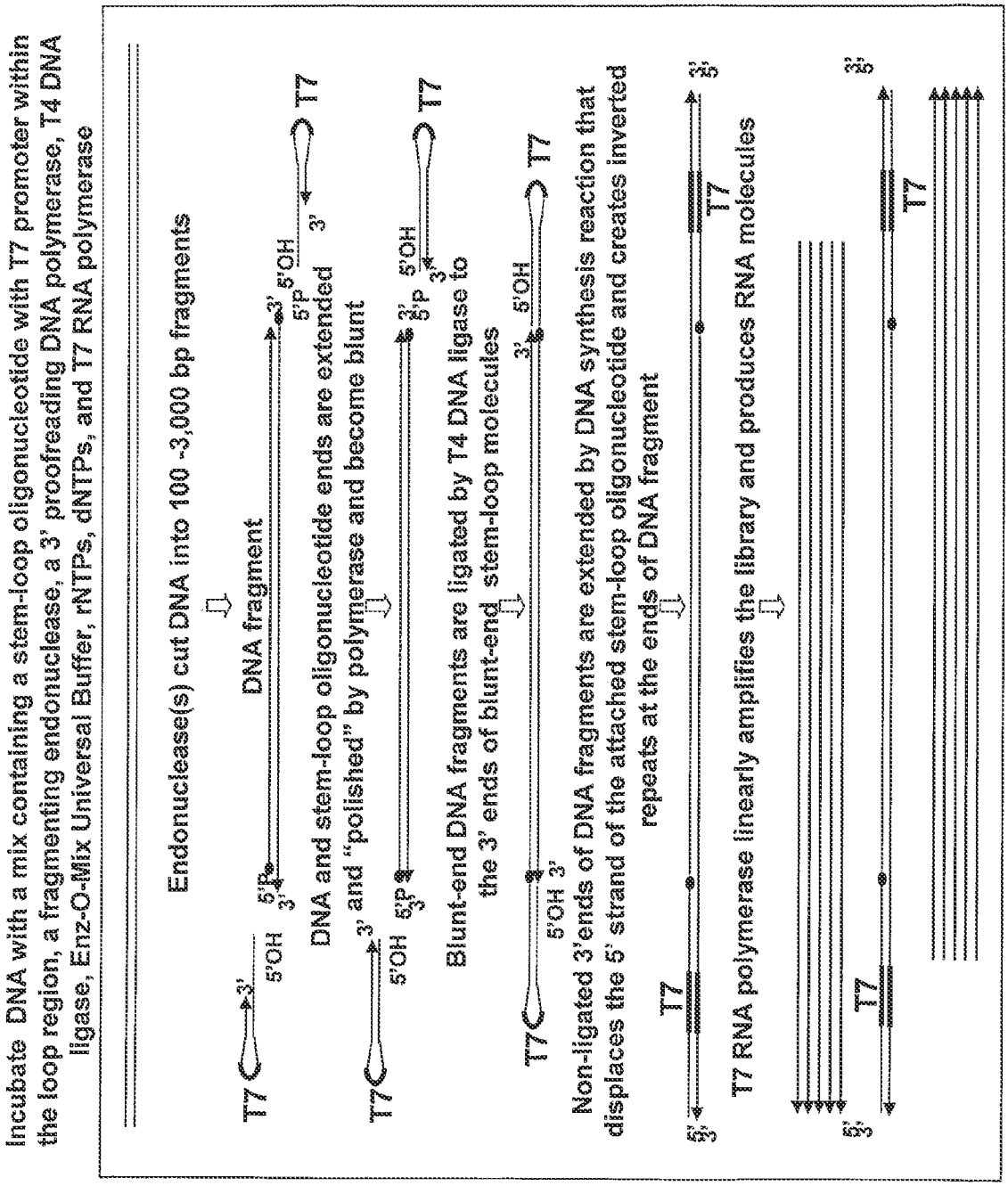

Incubate DNA with a mix containing a stem-loop oligonucleotide with T7 promoter within the loop region, a fragmenting endonuclease, a 3' proofreading DNA polymerase, T4 DNA ligase, Enz-O-Mix Universal Buffer, rNTPs, dNTPs, and T7 RNA polymerase
Endonuclease(s) cut DNA into 100 -3,000 bp fragments
DNA fragment
T7
3'
5'OH
5'P
3'
5'OH
DNA and stem-loop oligonucleotide ends are extended and "polished" by polymerase and become blunt
Blunt-end DNA fragments are ligated by T4 DNA ligase to the 3' ends of blunt-end stem-loop molecules
Non-ligated 3'ends of DNA fragments are extended by DNA synthesis reaction that displaces the 5' strand of the attached stem-loop oligonucleotide and creates inverted repeats at the ends of DNA fragment
T7 RNA polymerase linearly amplifies the library and produces RNA molecules

FIG. 6A

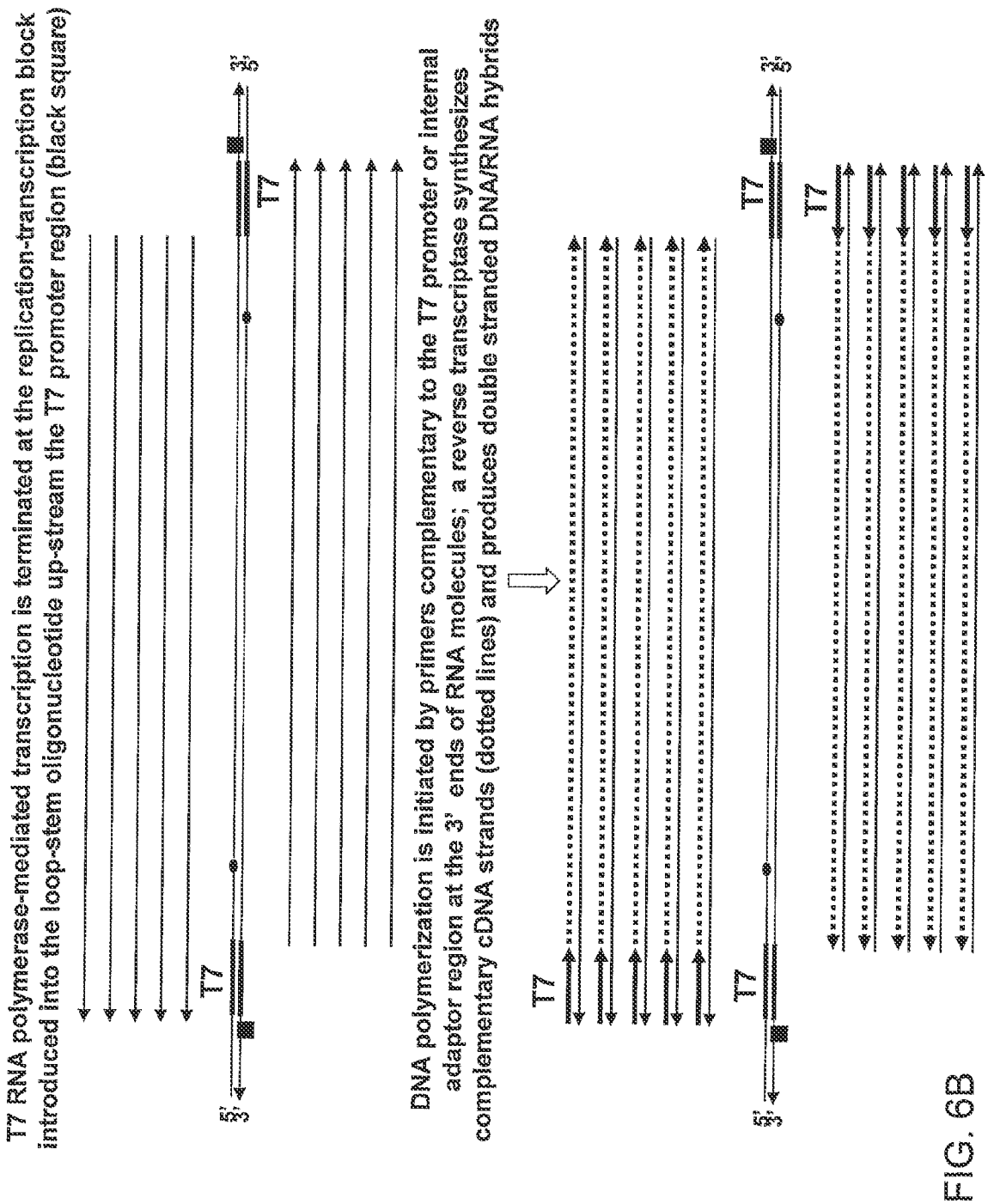

T7 RNA polymerase-mediated transcription is terminated at the replication-transcription block introduced into the loop-stem oligonucleotide up-stream the T7 promoter region (black square)
T7
5'
3'
3'
5'
T7
DNA polymerization is initiated by primers complementary to the T7 promoter or internal adaptor region at the 3' ends of RNA molecules; a reverse transcriptase synthesizes complementary cDNA strands (dotted lines) and produces double stranded DNA/RNA hybrids
T7
T7
5'
3'
3'
5'
T7
T7

FIG. 6B

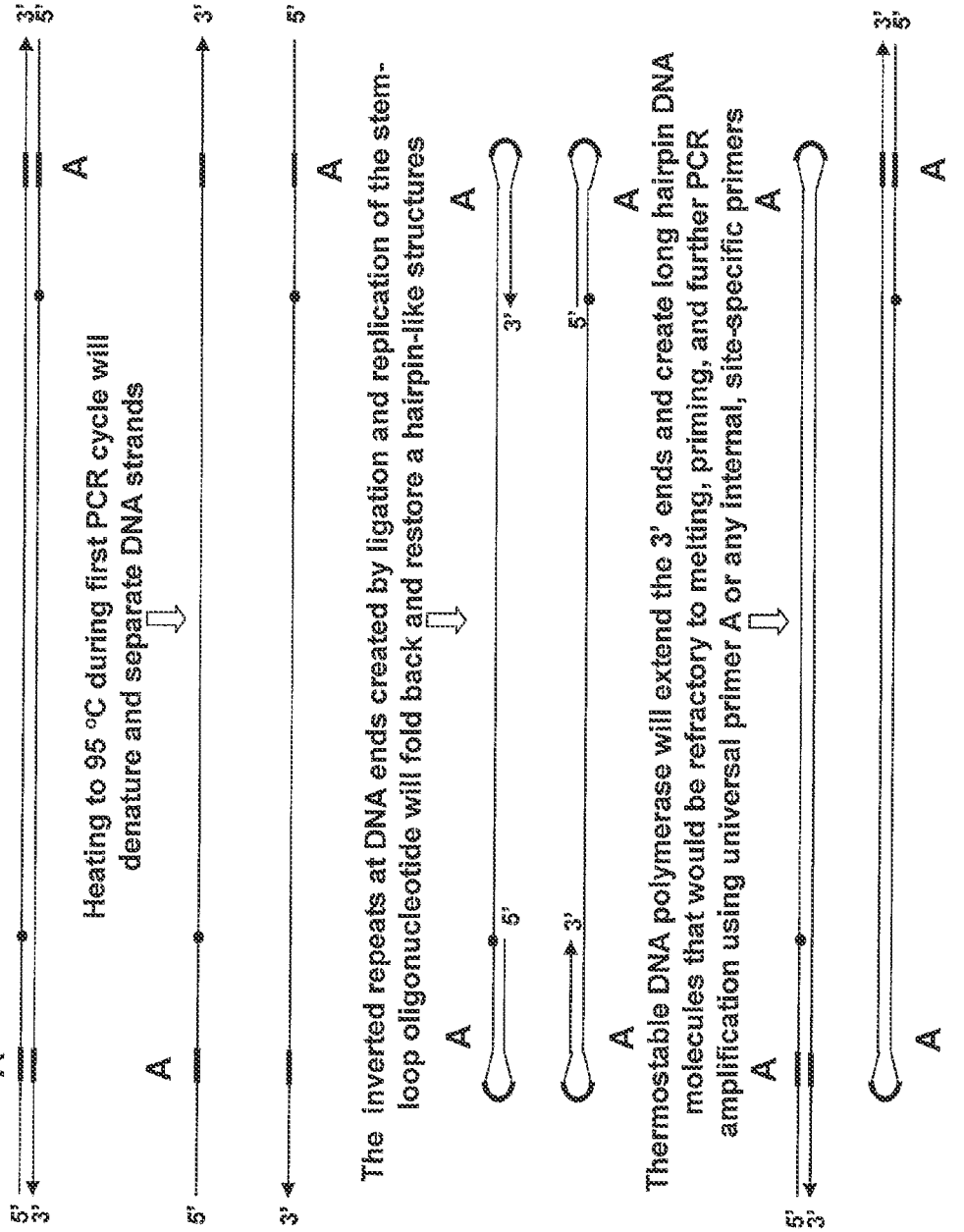

A
3'
5'
5'
3'
Heating to 95 °C during first PCR cycle will denature and separate DNA strands
A
3'
5'
5'
3'
A
The inverted repeats at DNA ends created by ligation and replication of the stem-loop oligonucleotide will fold back and restore a hairpin-like structures
A
5'
3'
A
3'
5'
A
Thermostable DNA polymerase will extend the 3' ends and create long hairpin DNA molecules that would be refractory to melting, priming, and further PCR amplification using universal primer A or any internal, site-specific primers
A
5'
3'
A
3'
5'
A

FIG. 7

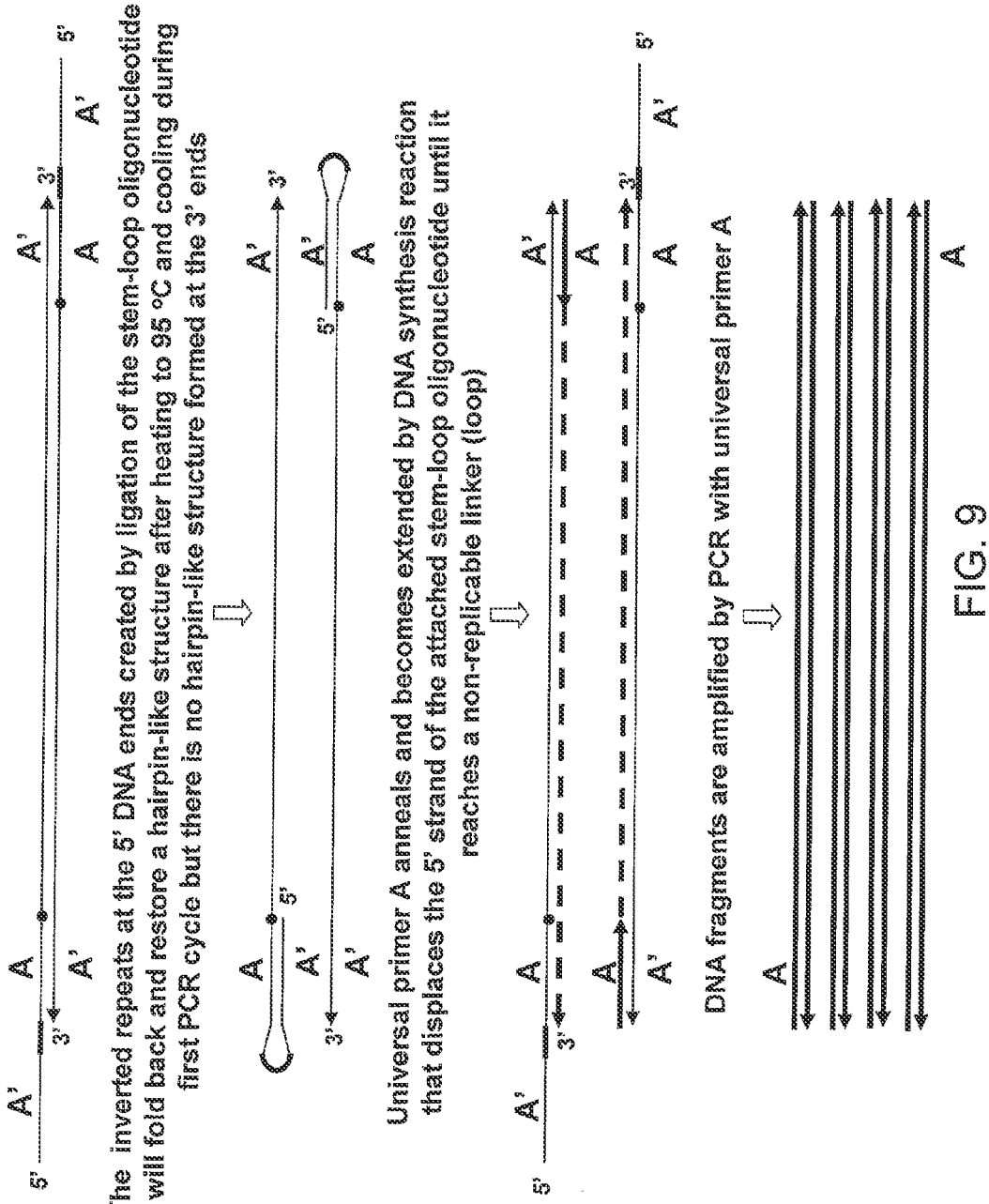

A'
A
A'
5'
3'
A'
A
A'
5'
The inverted repeats at the 5' DNA ends created by ligation of the stem-loop oligonucleotide will fold back and restore a hairpin-like structure after heating to 95 °C and cooling during first PCR cycle but there is no hairpin-like structure formed at the 3' ends
A
A'
5'
3'
A'
A
5'
A'
3'
A'
Universal primer A anneals and becomes extended by DNA synthesis reaction that displaces the 5' strand of the attached stem-loop oligonucleotide until it reaches a non-replicable linker (loop)
A'
A
A'
5'
3'
A
A'
A
3'
5'
A'
A
DNA fragments are amplified by PCR with universal primer A
A
A

FIG. 9

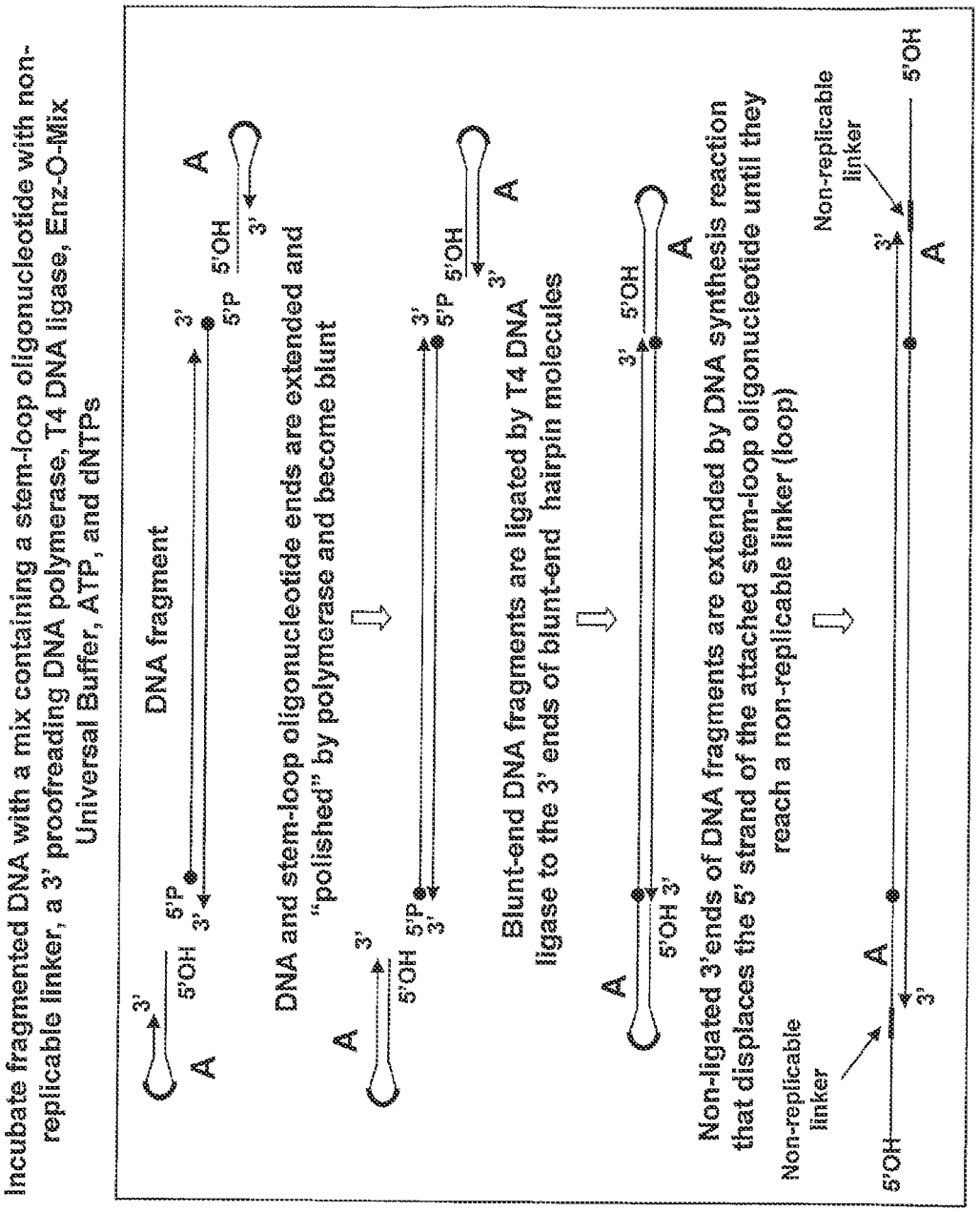

Incubate fragmented DNA with a mix containing a stem-loop oligonucleotide with non-replicable linker, a 3' proofreading DNA polymerase, T4 DNA ligase, Enz-O-Mix Universal Buffer, ATP, and dNTPs
DNA fragment
3'
5'OH
5'P
3'
A
DNA and stem-loop oligonucleotide ends are extended and "polished" by polymerase and become blunt
Blunt-end DNA fragments are ligated by T4 DNA ligase to the 3' ends of blunt-end hairpin molecules
Non-ligated 3'ends of DNA fragments are extended by DNA synthesis reaction that displaces the 5' strand of the attached stem-loop oligonucleotide until they reach a non-replicable linker (loop)
Non-replicable linker

FIG. 8

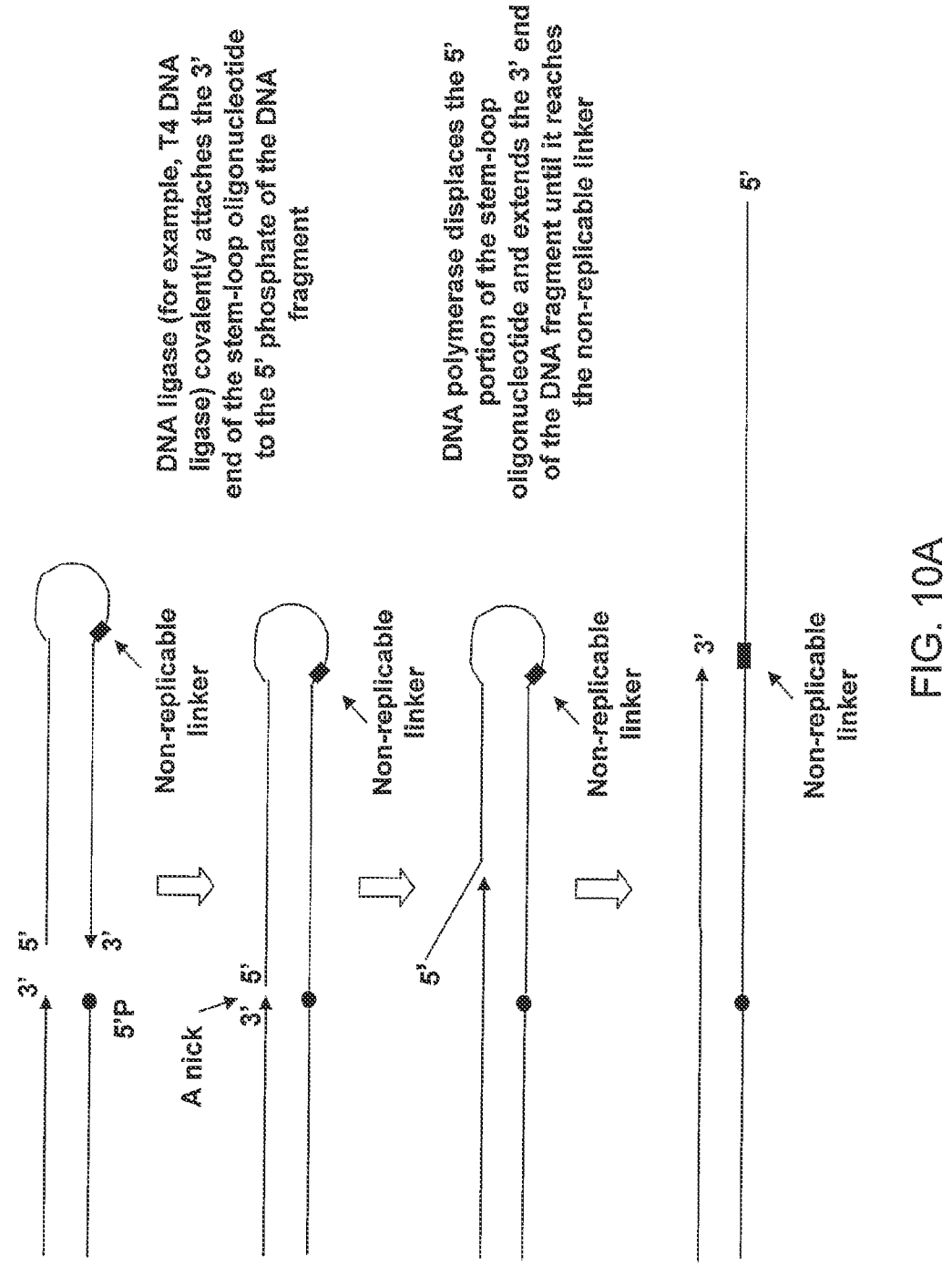

DNA ligase (for example, T4 DNA ligase) covalently attaches the 3' end of the stem-loop oligonucleotide to the 5' phosphate of the DNA fragment
DNA polymerase displaces the 5' portion of the stem-loop oligonucleotide and extends the 3' end of the DNA fragment until it reaches the non-replicable linker
Non-replicable linker
Non-replicable linker
Non-replicable linker
Non-replicable linker
5'
3'
5'
3'
5'P
A nick
3'
5'
5'
3'
5'

FIG. 10A

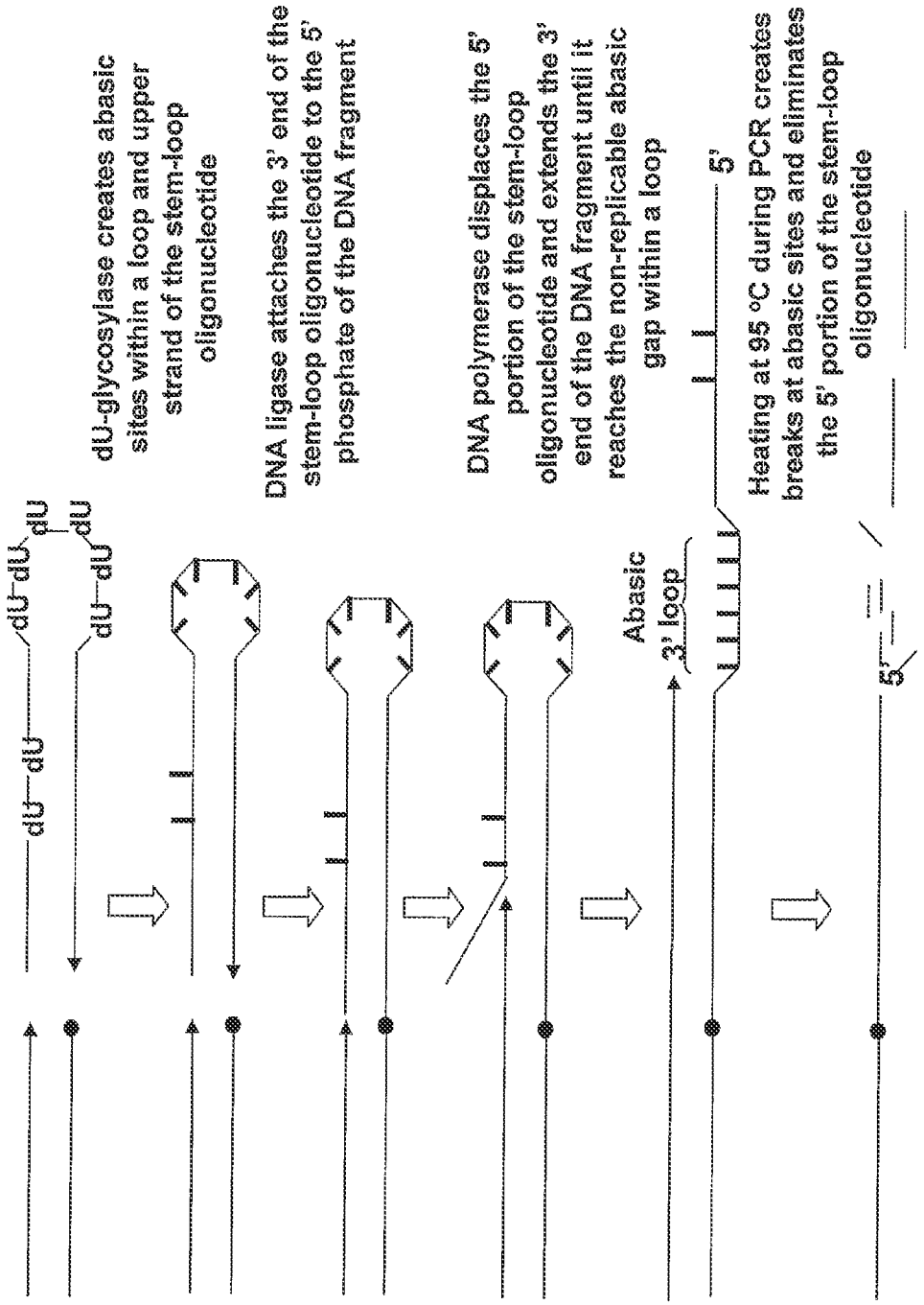

dU-glycosylase creates abasic sites within a loop and upper strand of the stem-loop oligonucleotide
DNA ligase attaches the 3' end of the stem-loop oligonucleotide to the 5' phosphate of the DNA fragment
DNA polymerase displaces the 5' portion of the stem-loop oligonucleotide and extends the 3' end of the DNA fragment until it reaches the non-replicable abasic gap within a loop
Heating at 95 °C during PCR creates breaks at abasic sites and eliminates the 5' portion of the stem-loop oligonucleotide
dU
Abasic
3' loop
5'
5'

FIG. 10B

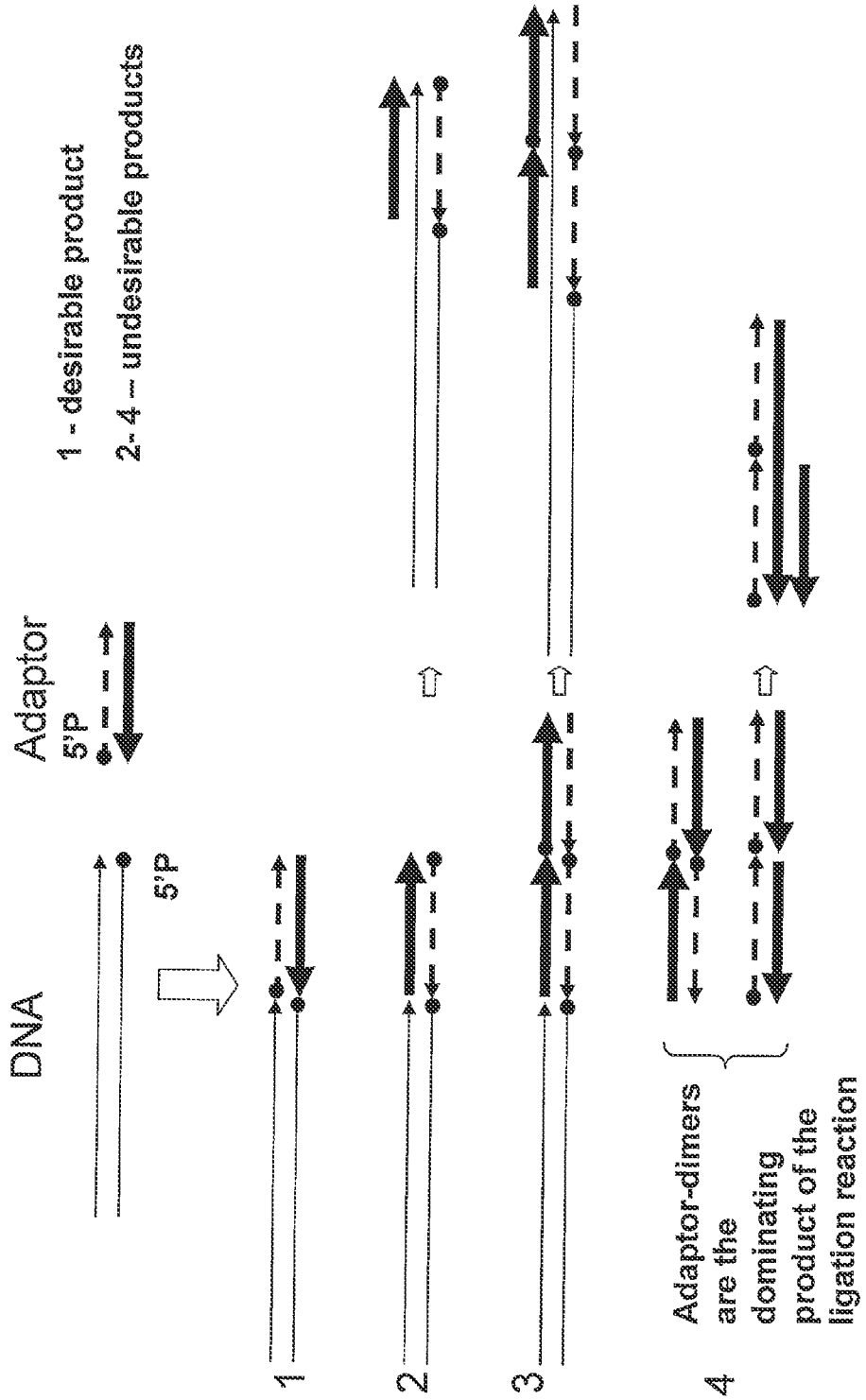

1 - desirable product
2-4 – undesirable products
Adaptor
5'P
DNA
5'P
1
2
3
4
Adaptor-dimers are the dominating product of the ligation reaction

FIG. 11A

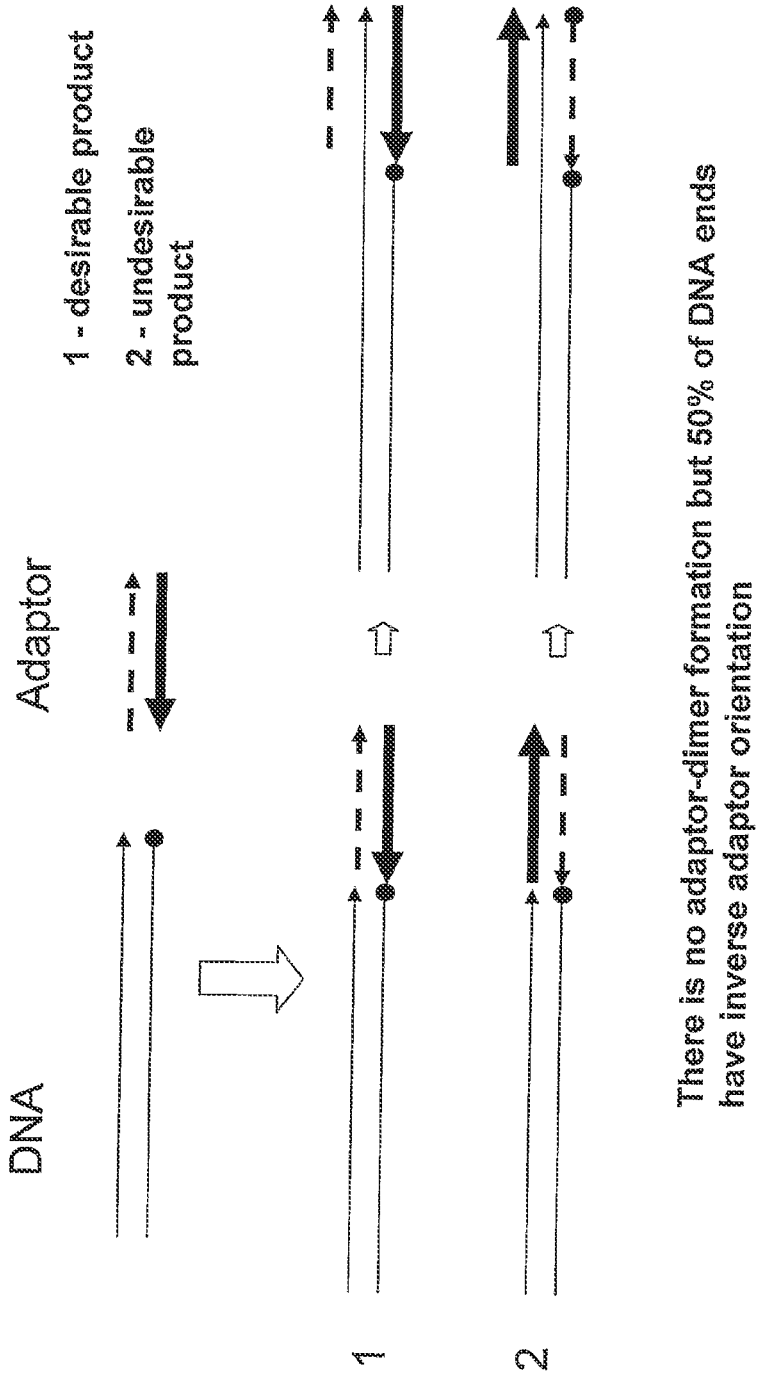
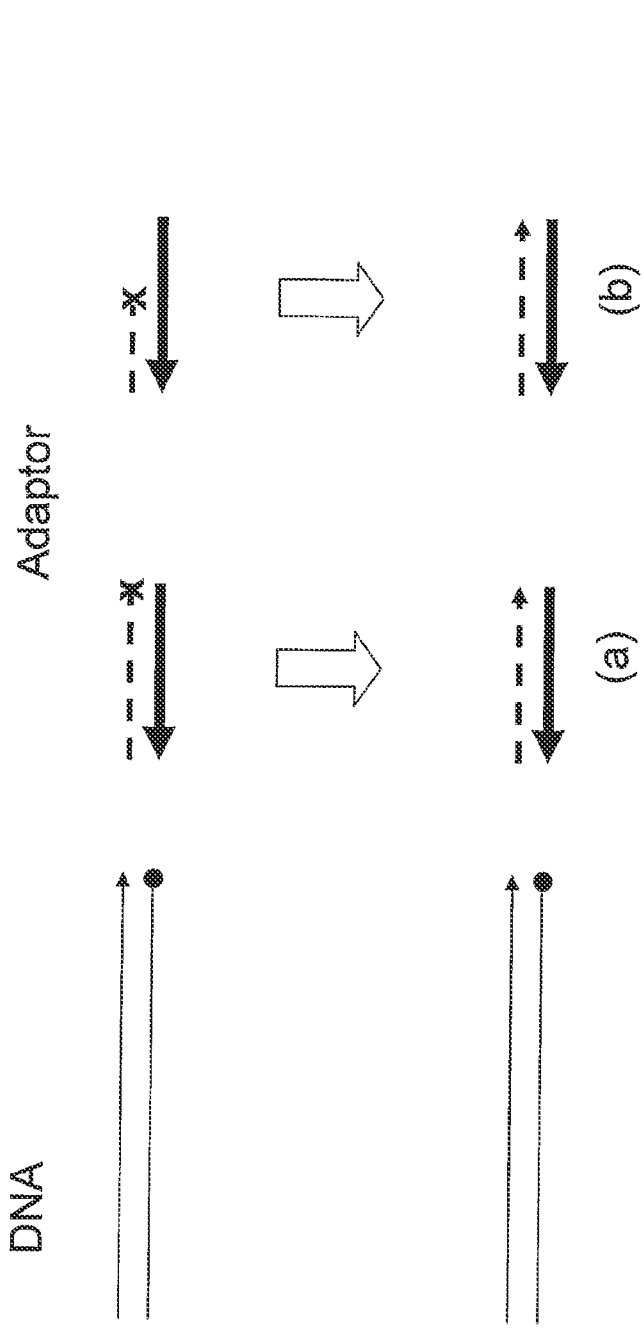

DNA
Adaptor
1 - desirable product
2 - undesirable product
1
2
There is no adaptor-dimer formation but 50% of DNA ends have inverse adaptor orientation

FIG. 11B



3'-blocked oligonucleotides within the standard 2-oligo adaptor will be repaired and extended by 3' proofreading polymerase present in the mix and ligated to DNA ends in two orientations

FIG. 11C

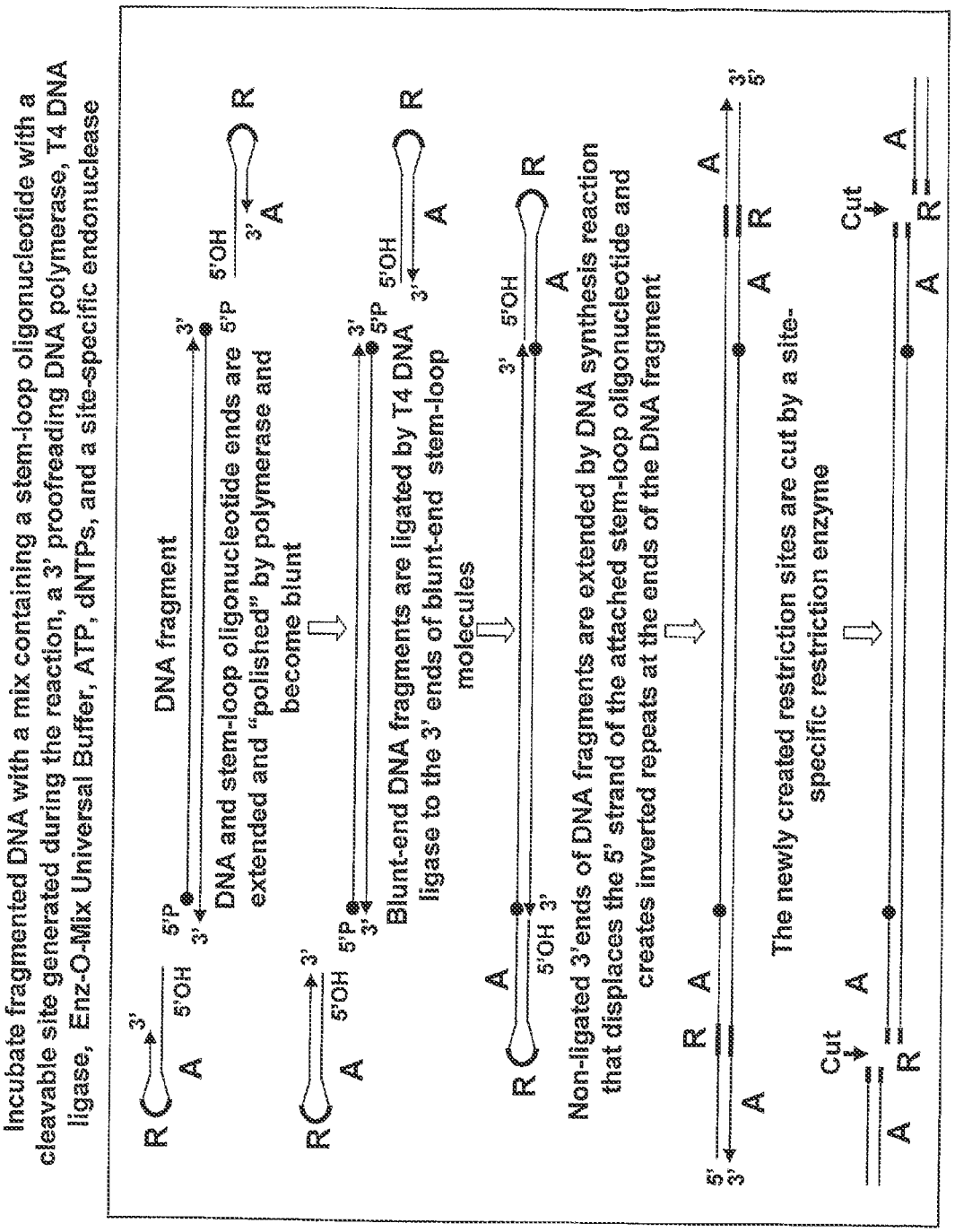

Incubate fragmented DNA with a mix containing a stem-loop oligonucleotide with a cleavable site generated during the reaction, a 3' proofreading DNA polymerase, T4 DNA ligase, Enz-O-Mix Universal Buffer, ATP, dNTPs, and a site-specific endonuclease
DNA fragment
DNA and stem-loop oligonucleotide ends are extended and "polished" by polymerase and become blunt
Blunt-end DNA fragments are ligated by T4 DNA ligase to the 3' ends of blunt-end stem-loop molecules
Non-ligated 3'ends of DNA fragments are extended by DNA synthesis reaction that displaces the 5' strand of the attached stem-loop oligonucleotide and creates inverted repeats at the ends of the DNA fragment
The newly created restriction sites are cut by a site-specific restriction enzyme
Cut
Cut

FIG. 12

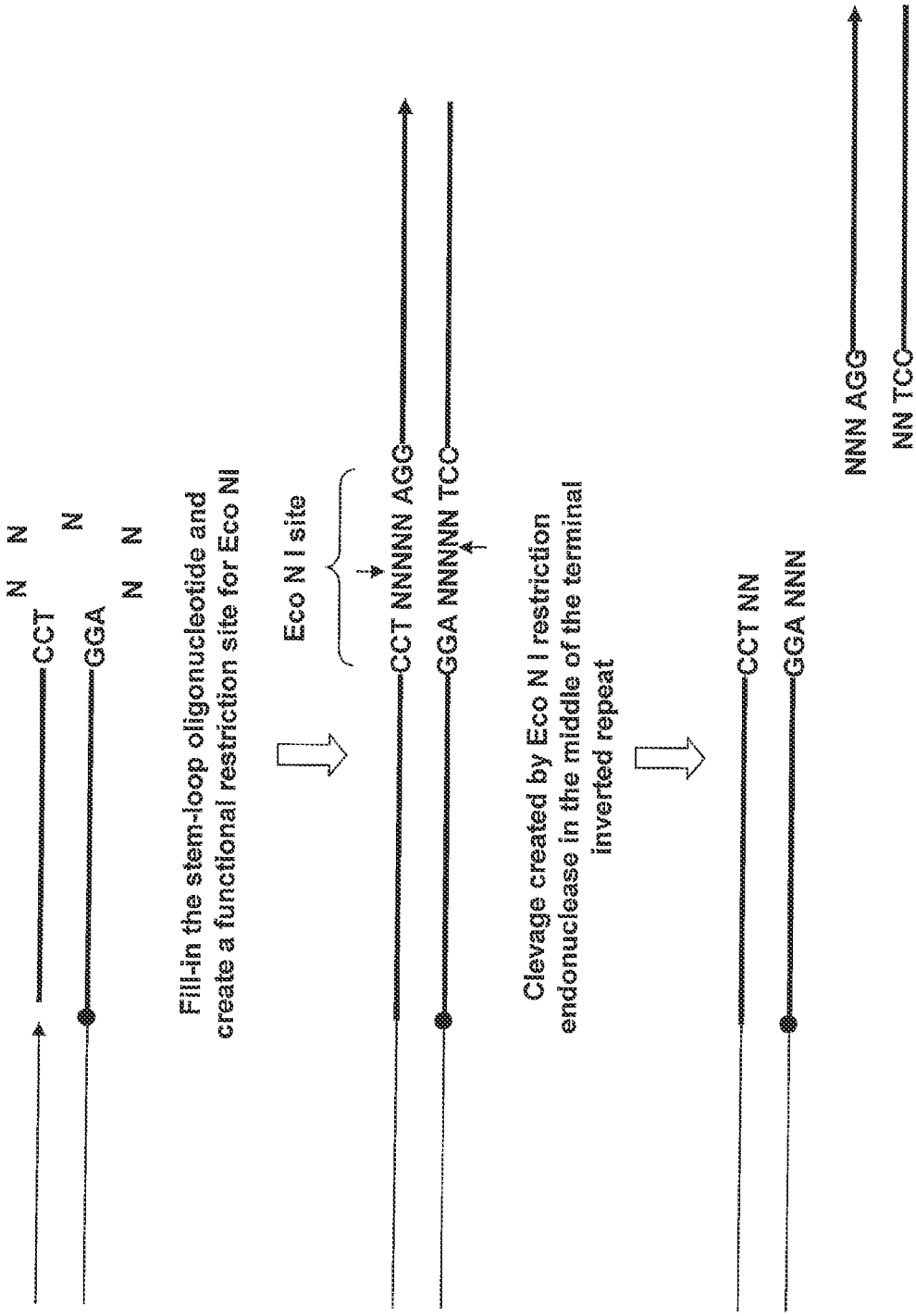

N N
CCT
N
GGA
N N
Fill-in the stem-loop oligonucleotide and create a functional restriction site for Eco NI
Eco N I site
CCT NNNNN AGG
GGA NNNNN TCC
Clevage created by Eco N I restriction endonuclease in the middle of the terminal inverted repeat
CCT NN
GGA NNN
NNN AGG
NN TCC

FIG. 13A

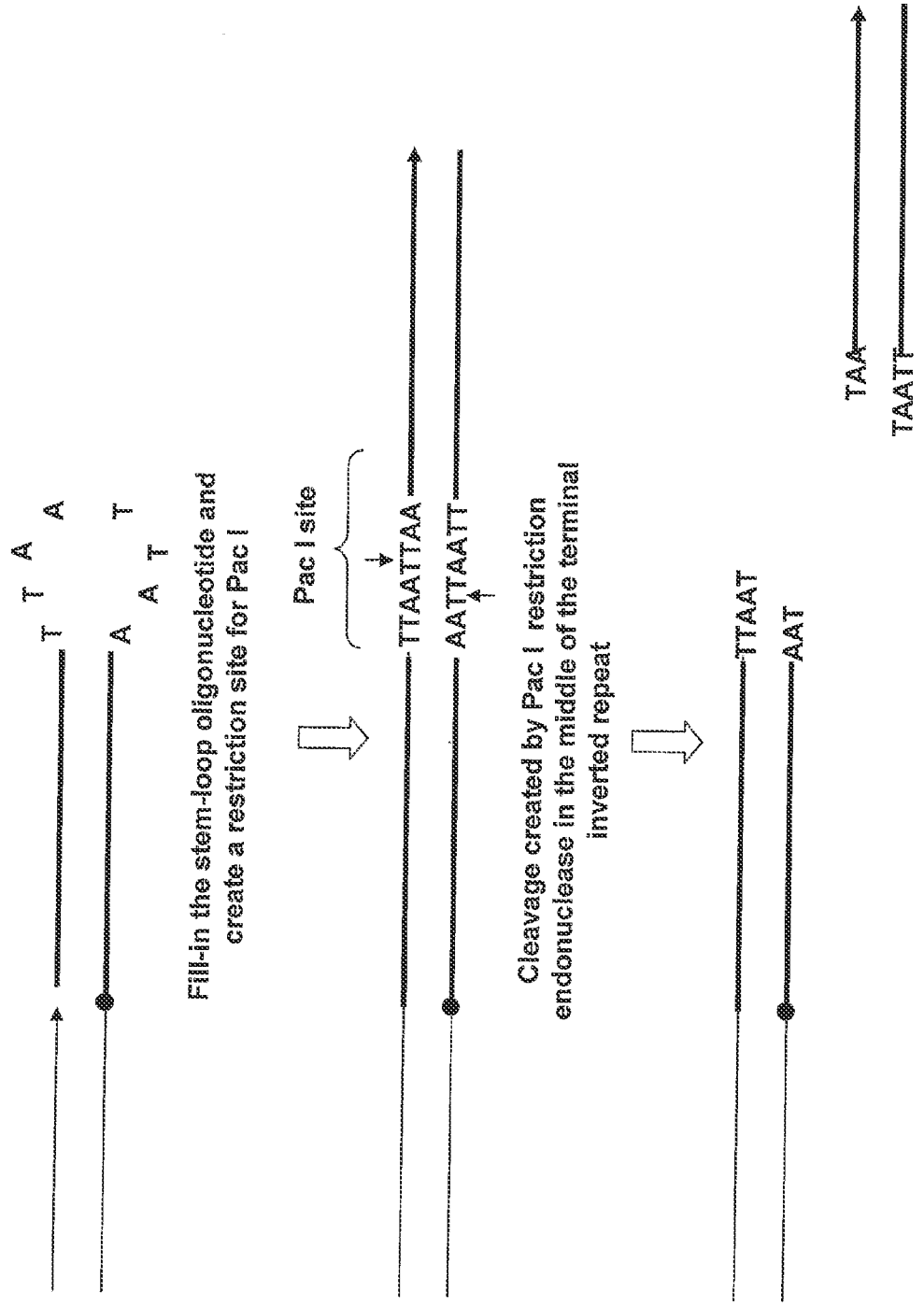

T A A
T T
A T
A T
Fill-in the stem-loop oligonucleotide and create a restriction site for Pac I
Pac I site
TTAATTAA
AATTAATT
Cleavage created by Pac I restriction endonuclease in the middle of the terminal inverted repeat
TTAAT
AAT
TAA
TAATT

FIG. 13B

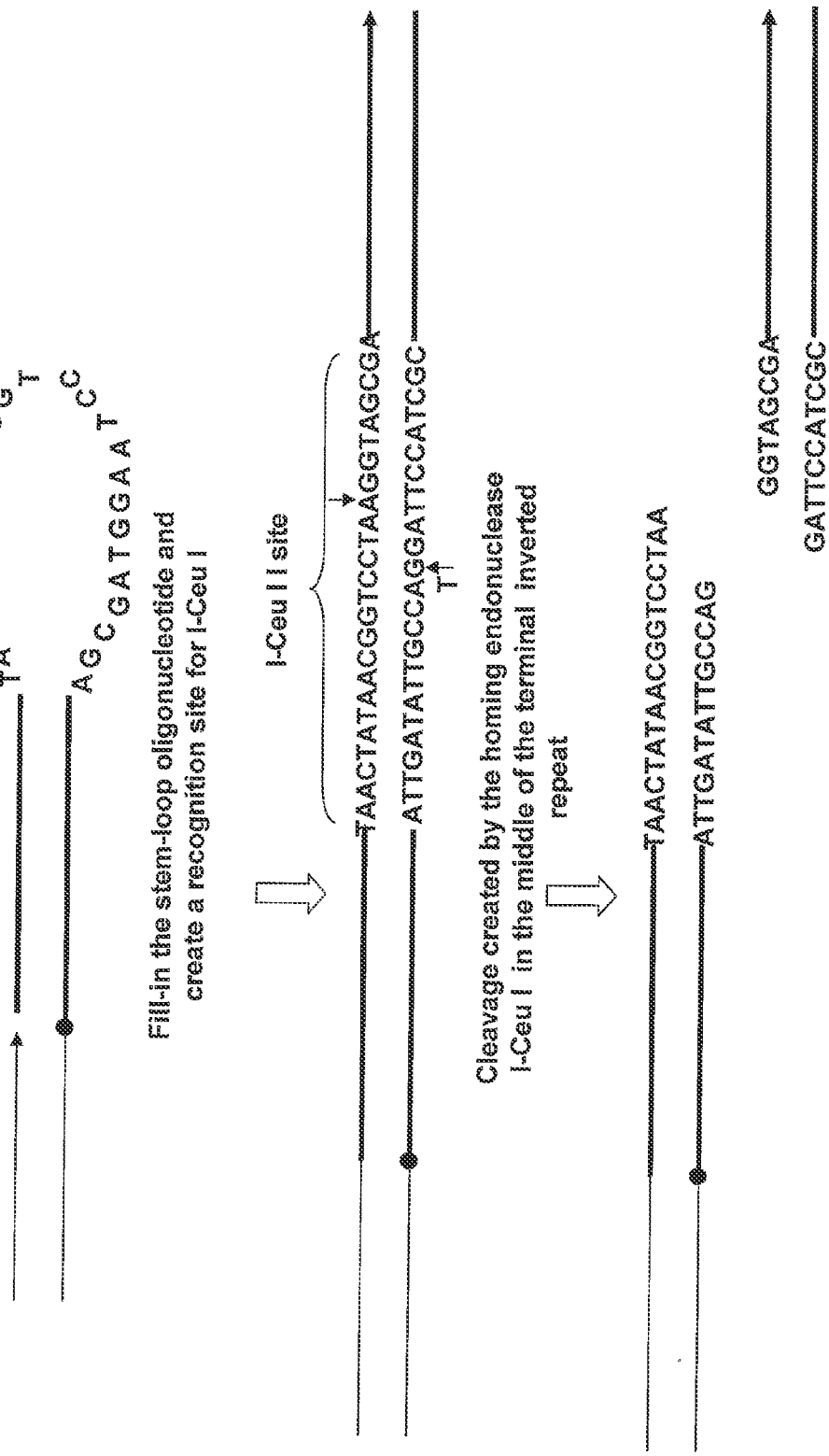

Fill-in the stem-loop oligonucleotide and create a recognition site for I-Ceu I
I-Ceu I I site
Cleavage created by the homing endonuclease I-Ceu I in the middle of the terminal inverted repeat

FIG. 13C

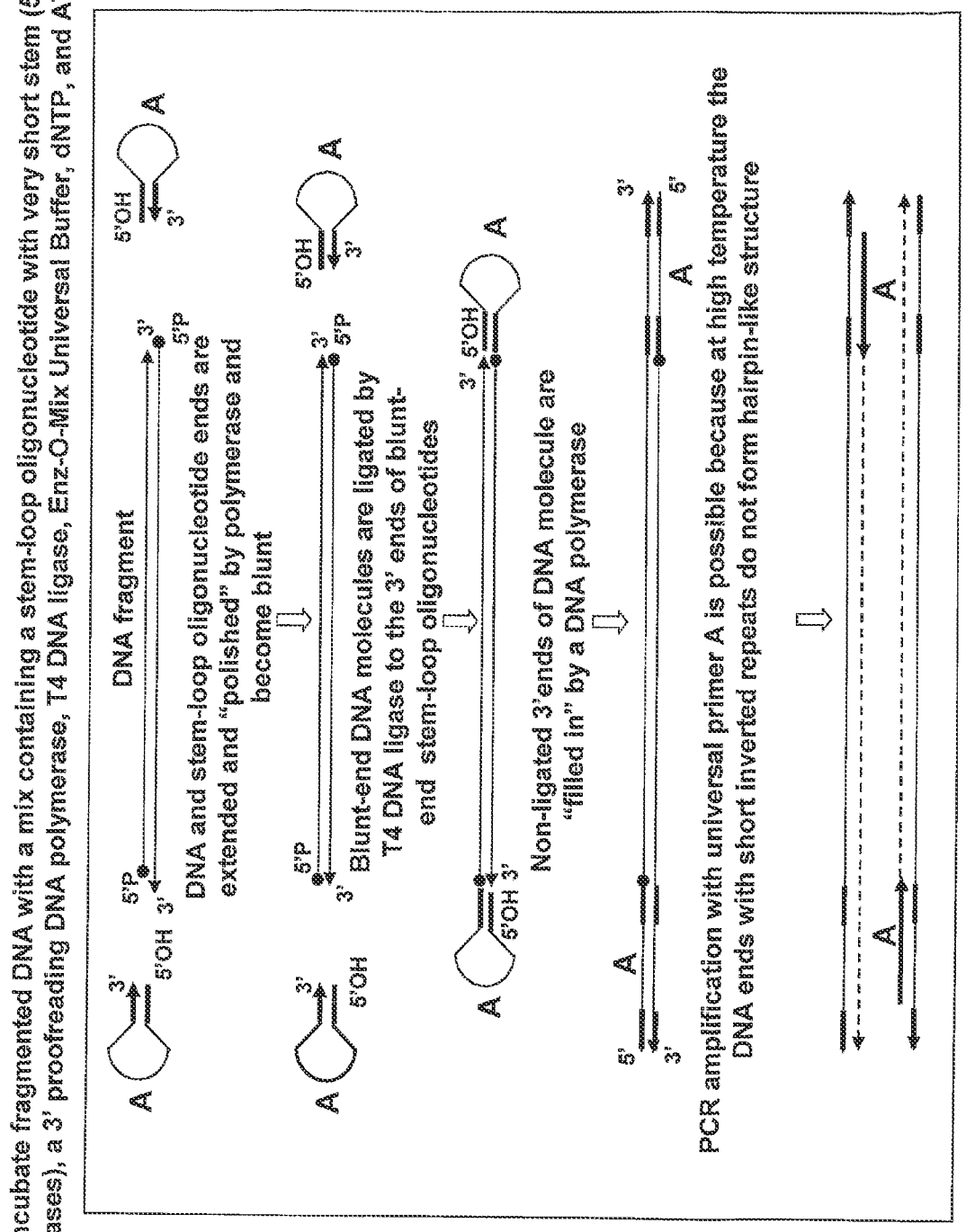

Incubate fragmented DNA with a mix containing a stem-loop oligonucleotide with very short stem (5-8 bases), a 3' proofreading DNA polymerase, T4 DNA ligase, Enz-O-Mix Universal Buffer, dNTP, and ATP
DNA fragment
DNA and stem-loop oligonucleotide ends are extended and "polished" by polymerase and become blunt
Blunt-end DNA molecules are ligated by T4 DNA ligase to the 3' ends of blunt-end stem-loop oligonucleotides
Non-ligated 3'ends of DNA molecule are "filled in" by a DNA polymerase
PCR amplification with universal primer A is possible because at high temperature the DNA ends with short inverted repeats do not form hairpin-like structure

FIG. 14

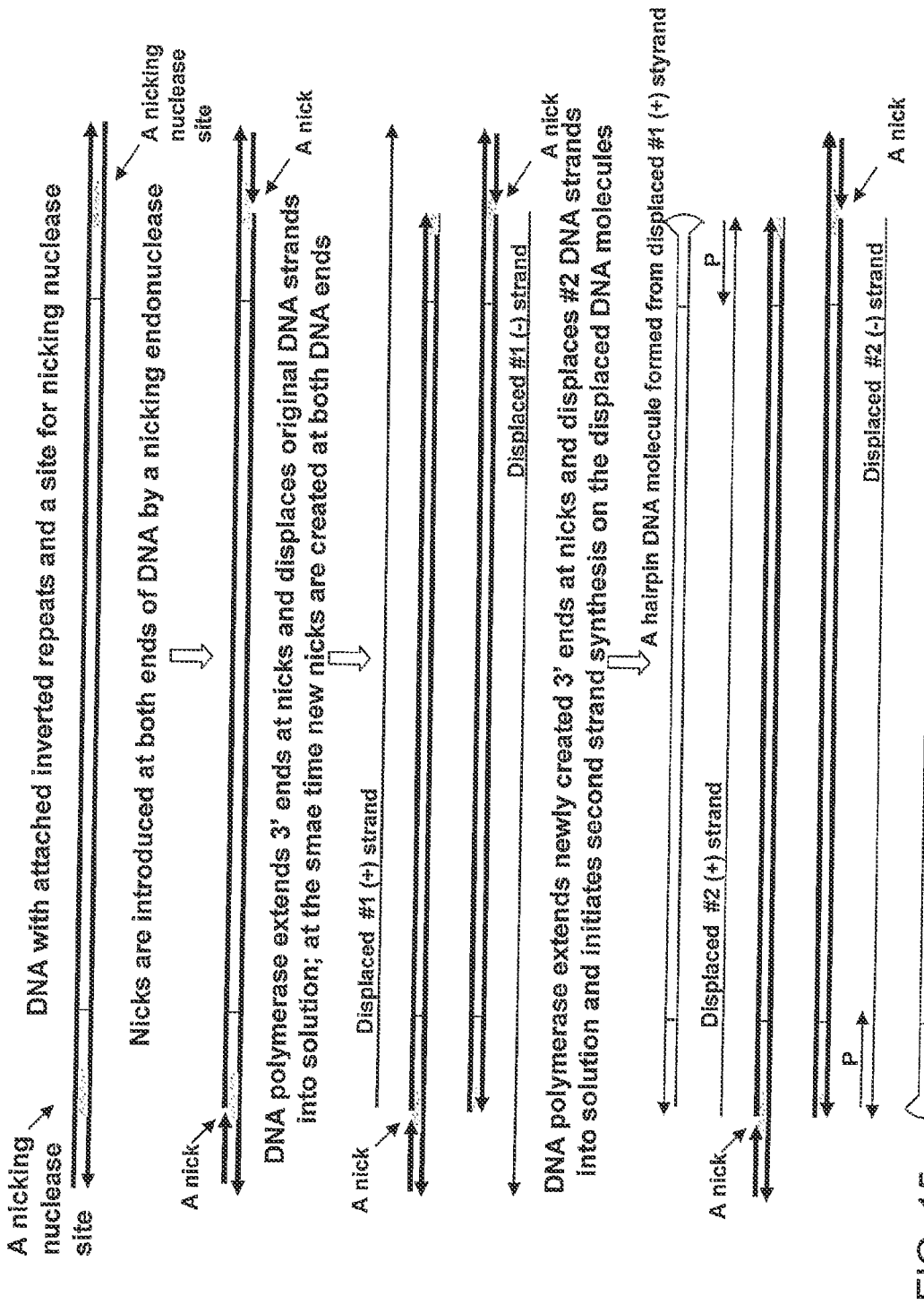

A nicking nuclease site
DNA with attached inverted repeats and a site for nicking nuclease
Nicks are introduced at both ends of DNA by a nicking endonuclease
A nicking nuclease site
A nick
DNA polymerase extends 3' ends at nicks and displaces original DNA strands into solution; at the smae time new nicks are created at both DNA ends
A nick
Displaced #1 (+) strand
Displaced #1 (-) strand
A nick
DNA polymerase extends newly created 3' ends at nicks and displaces #2 DNA strands into solution and initiates second strand synthesis on the displaced DNA molecules
A hairpin DNA molecule formed from displaced #1 (+) styrand
P
Displaced #2 (+) strand
A nick
P
Displaced #2 (-) strand
A nick
A hairpin DNA molecule formed from displaced #1 (-) styrand

FIG. 15

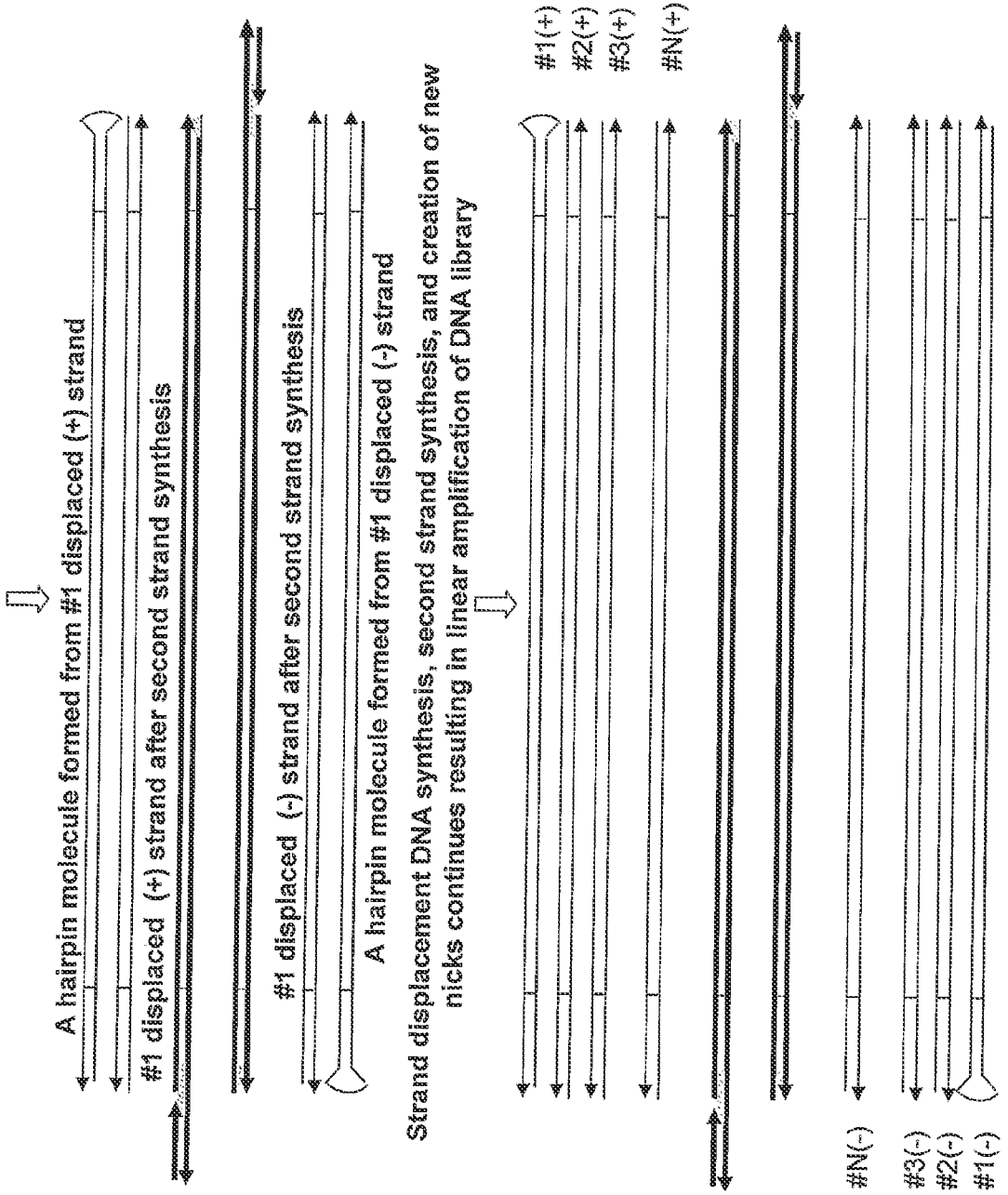


FIG. 15 (CONTINUED)

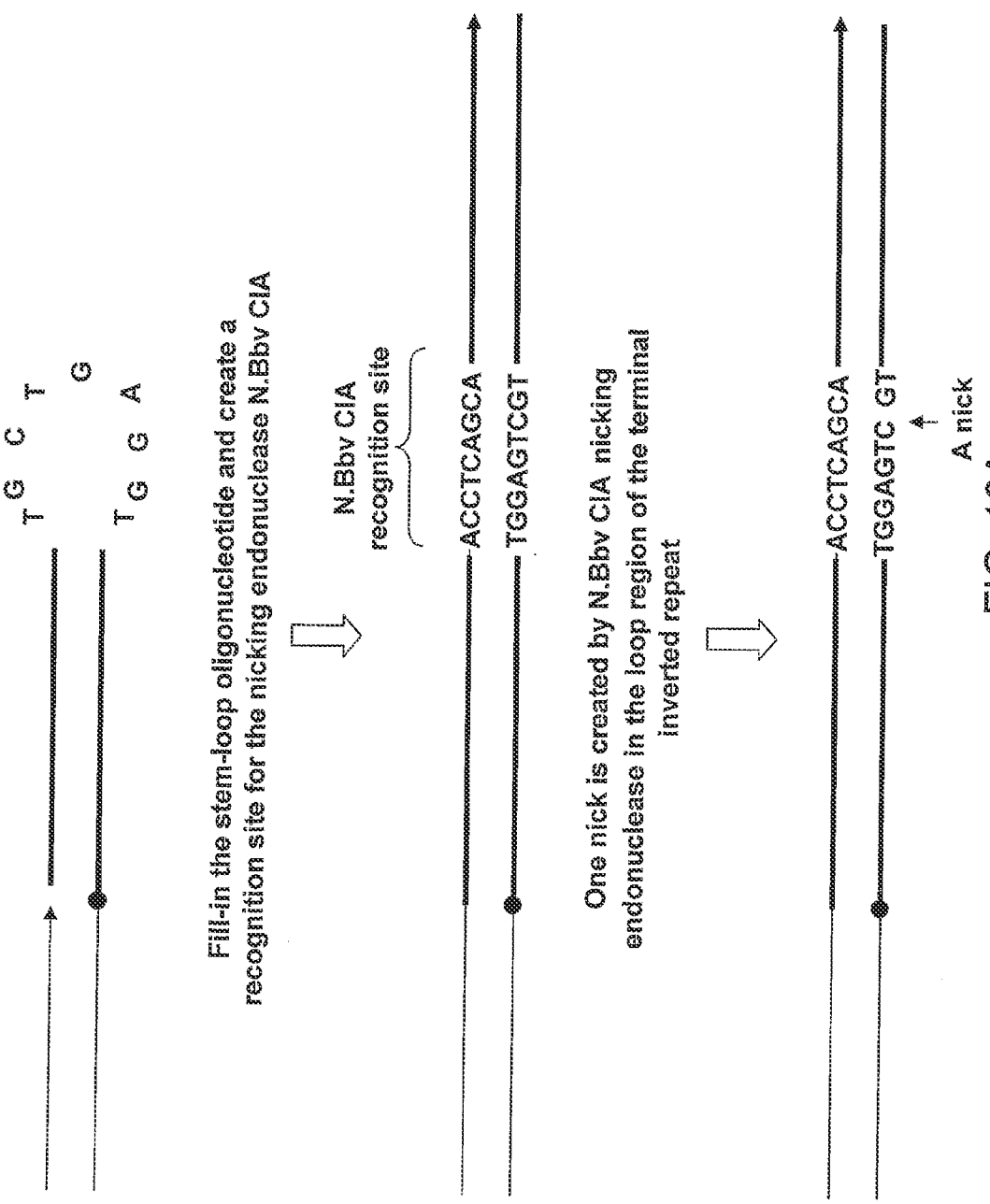

Fill-in the stem-loop oligonucleotide and create a recognition site for the nicking endonuclease N.Bbv CIA
N.Bbv CIA recognition site
ACCTCAGCA
TGGAGTCGT
One nick is created by N.Bbv CIA nicking endonuclease in the loop region of the terminal inverted repeat
ACCTCAGCA
TGGAGTC GT
A nick

FIG. 16A

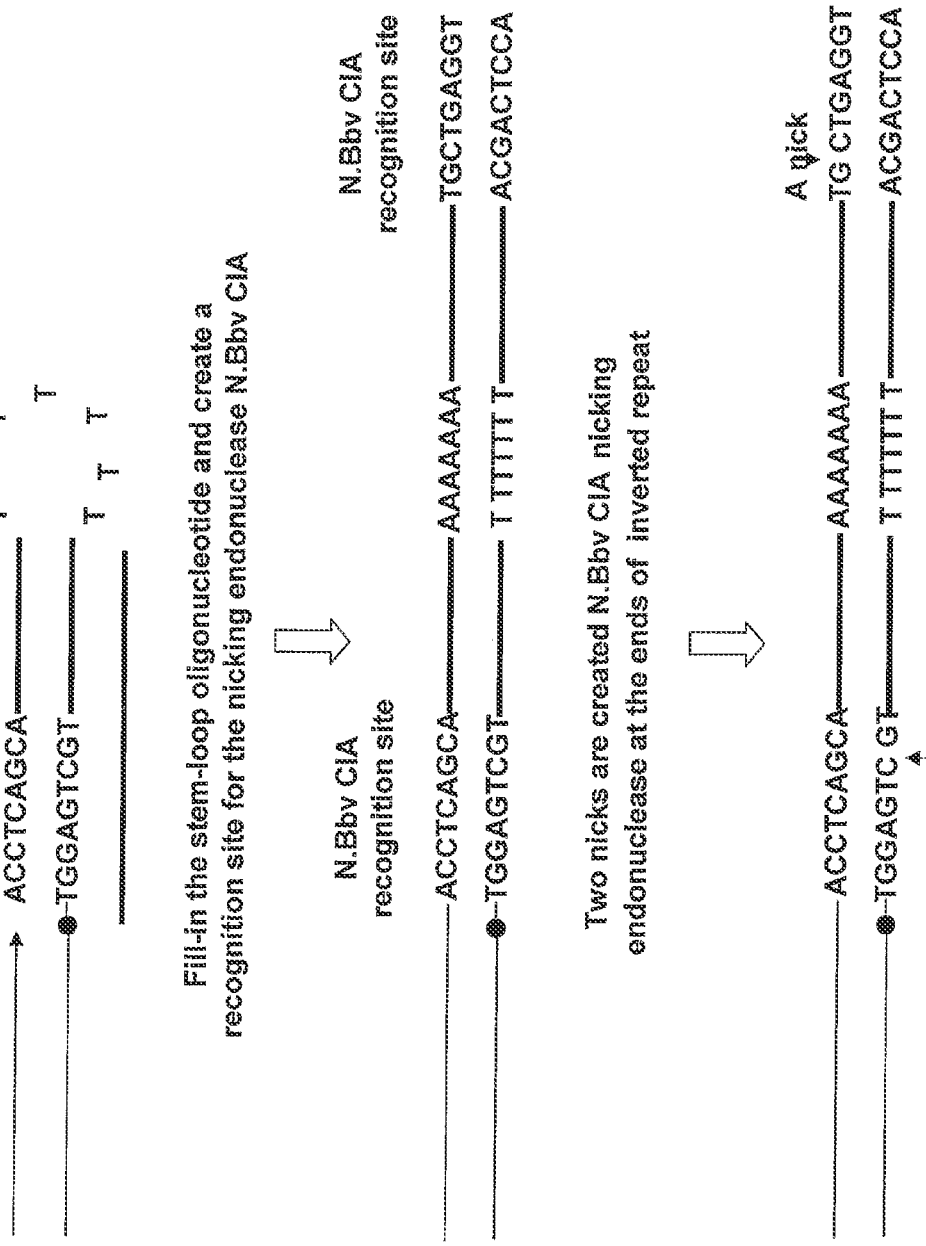
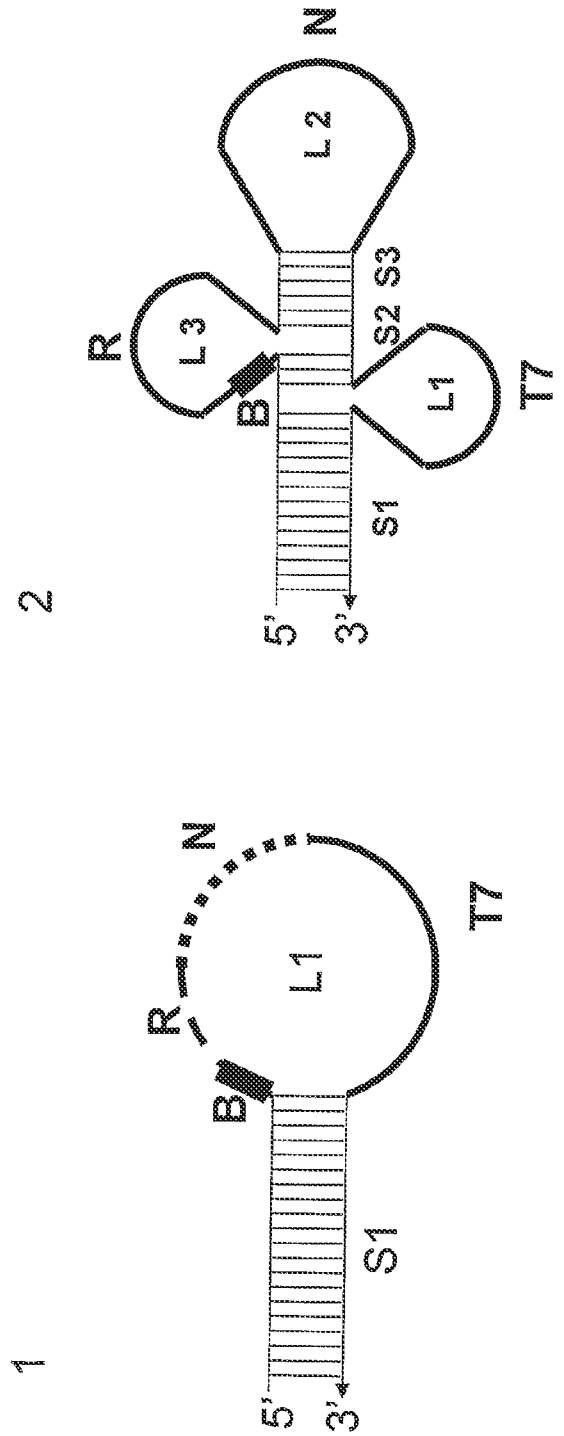

ACCTCAGCA
TGGAGTCGT
Fill-in the stem-loop oligonucleotide and create a recognition site for the nicking endonuclease N.Bbv CIA
N.Bbv CIA recognition site
N.Bbv CIA recognition site
ACCTCAGCA
AAAAAAA
TGCTGAGGT
TGGAGTCGT
T TTTTT T
ACGACTCCA
Two nicks are created N.Bbv CIA nicking endonuclease at the ends of inverted repeat
A nick
A nick
ACCTCAGCA
AAAAAAA
TG CTGAGGT
TGGAGTC GT
T TTTTT T
ACGACTCCA

FIG. 16B



T7 - T7 promoter sequence (including all necessary for efficient transcription downstream and upstream regions;

N - recognition sequence for a nicking endonuclease;

R - recognition sequence for a restriction endonuclease;

B - replication or transcription block (chemical group, abasic site or gap, etc.);

S1, S2, S3 – stem regions 1, 2, and 3;

L1, L2, L3 – loop regions 1, 2, and 3.

FIG. 16C

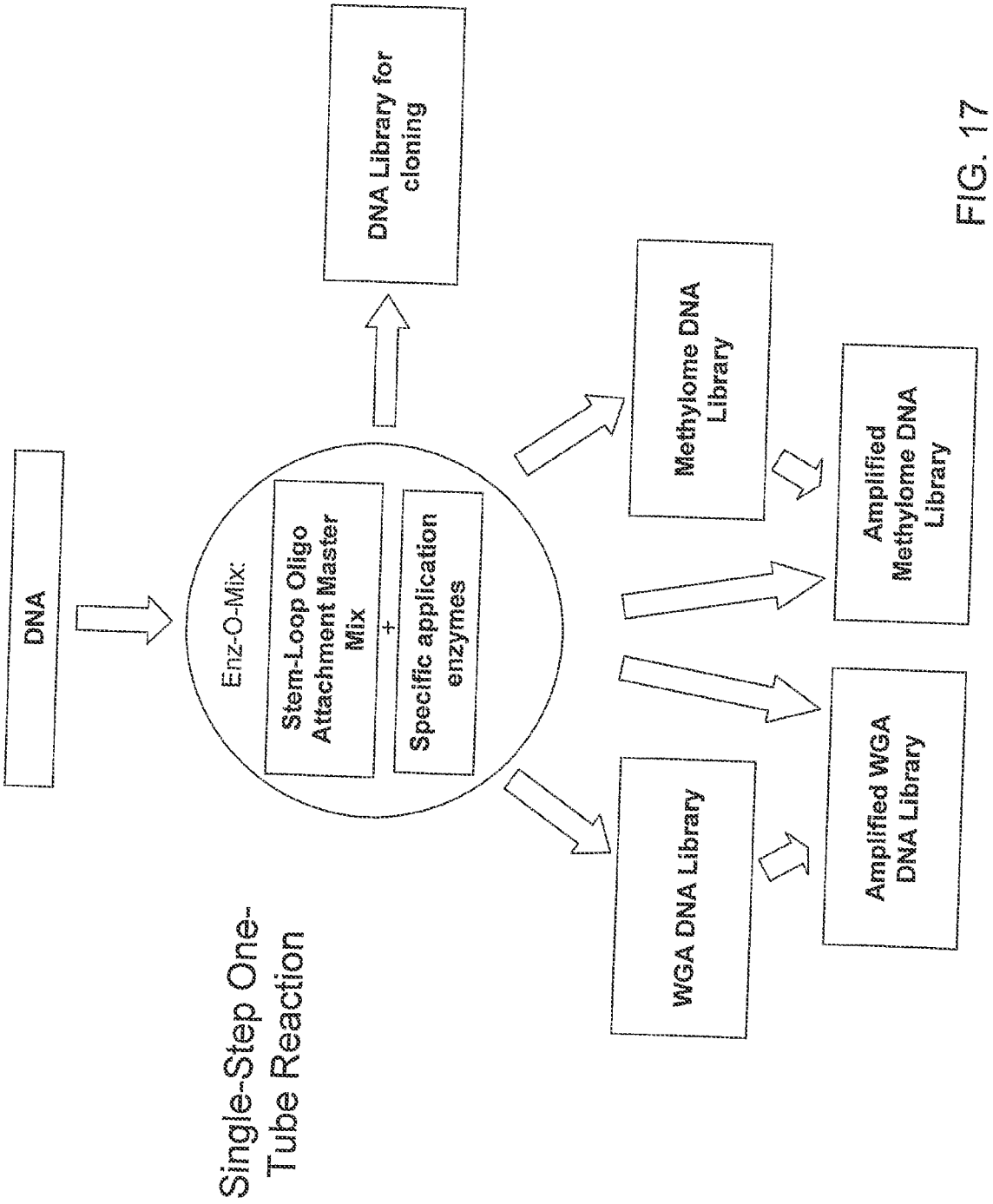

DNA
Enz-O-Mix:
Stem-Loop Oligo Attachment Master Mix
+
Specific application enzymes
DNA Library for cloning
Methylome DNA Library
Amplified Methylome DNA Library
WGA DNA Library
Amplified WGA DNA Library
Single-Step One-Tube Reaction

FIG. 17

Stem-Loop Oligo Attachment Master Mix I	Stem-Loop Oligo Attachment Master Mix II	Stem-Loop Oligo Attachment Master Mix III	Stem-Loop Oligo Attachment Master Mix IV
Stem-loop oligonucleotide 3' proofreading DNA polymerase T4 DNA ligase Universal Buffer ATP dNTPs	Stem-loop oligonucleotide with non-replicable linker (chemical) 3' proofreading DNA polymerase T4 DNA ligase Universal Buffer ATP dNTPs	Stem-loop oligonucleotide with non-replicable linker (abasic) dU-glycosylase 3' proofreading DNA polymerase T4 DNA ligase Universal Buffer ATP dNTPs	Stem-loop oligonucleotide with cleavable site created by replication Site-specific endonuclease 3' proofreading DNA polymerase T4 DNA ligase Universal Buffer ATP dNTPs

FIG. 18

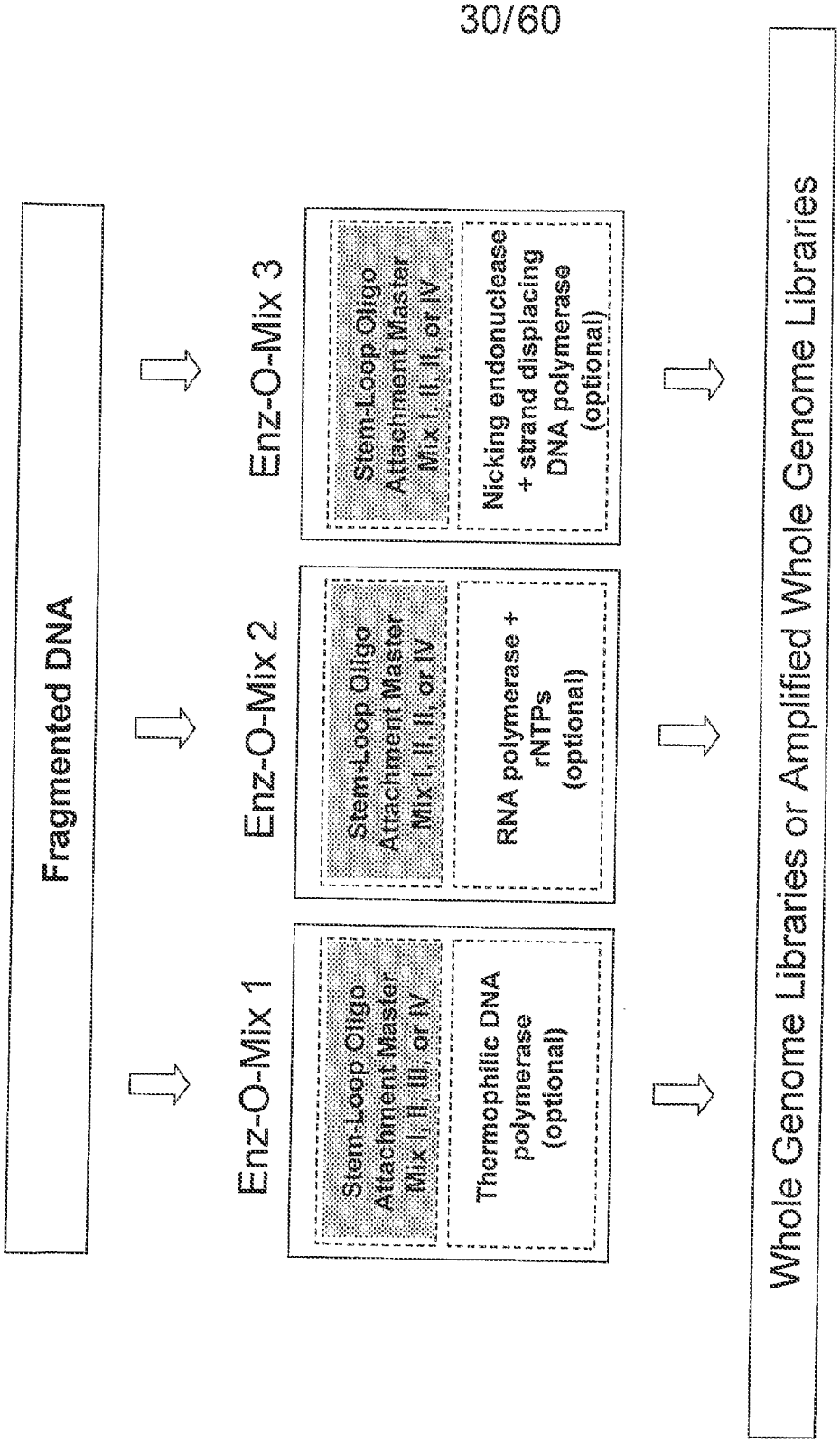

Fragmented DNA
Enz-O-Mix 1
Stem-Loop Oligo Attachment Master Mix I, II, III, or IV
Thermophilic DNA polymerase (optional)
Enz-O-Mix 2
Stem-Loop Oligo Attachment Master Mix I, II, II, or IV
RNA polymerase + rNTPs (optional)
Enz-O-Mix 3
Stem-Loop Oligo Attachment Master Mix I, II, II, or IV
Nicking endonuclease + strand displacing DNA polymerase (optional)
Whole Genome Libraries or Amplified Whole Genome Libraries

FIG. 19A

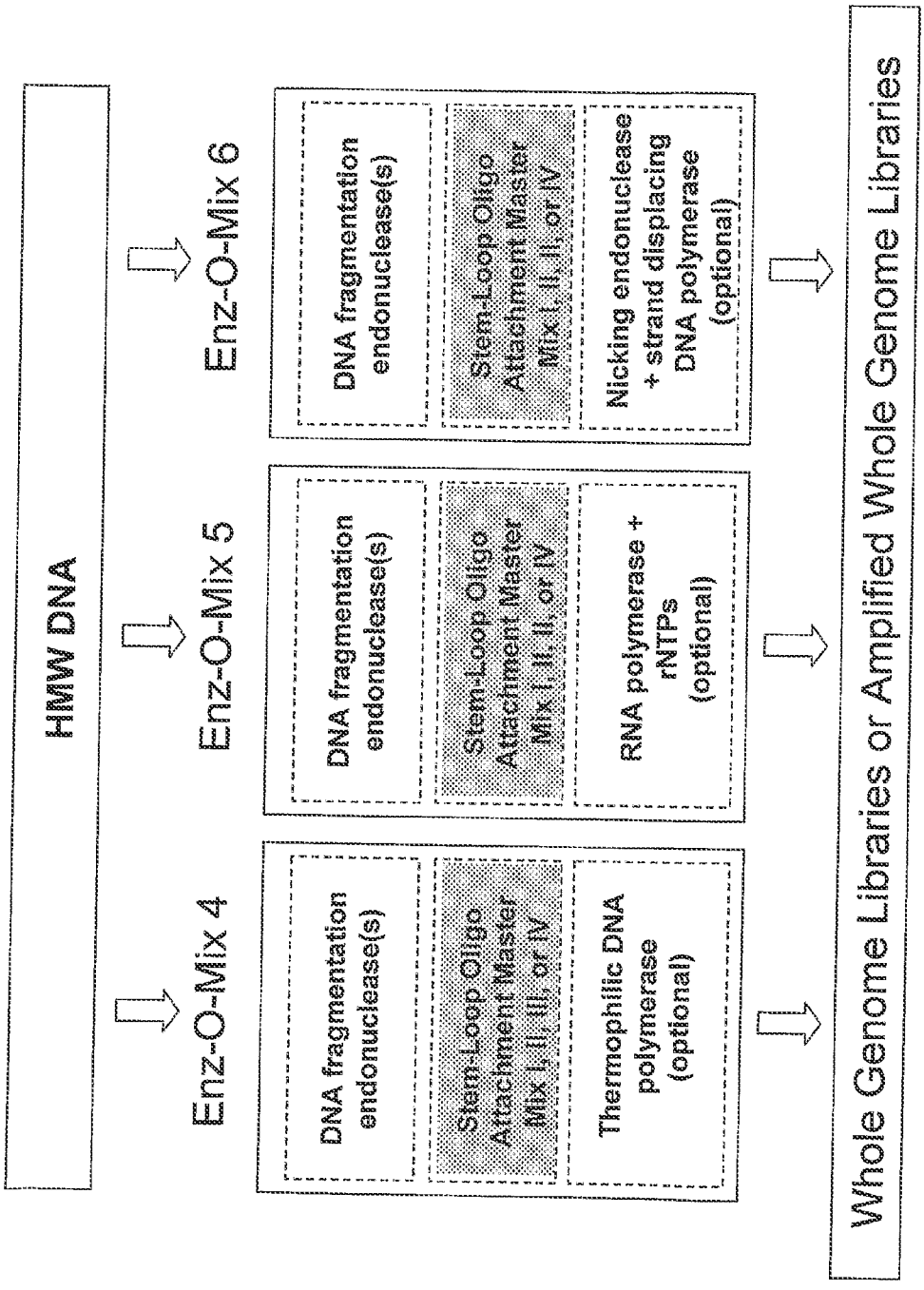

HMW DNA
Enz-O-Mix 4
DNA fragmentation endonuclease(s)
Stem-Loop Oligo Attachment Master Mix I, II, III, or IV
Thermophilic DNA polymerase (optional)
Enz-O-Mix 5
DNA fragmentation endonuclease(s)
Stem-Loop Oligo Attachment Master Mix I, II, II, or IV
RNA polymerase + rNTPs (optional)
Enz-O-Mix 6
DNA fragmentation endonuclease(s)
Stem-Loop Oligo Attachment Master Mix I, II, II, or IV
Nicking endonuclease + strand displacing DNA polymerase (optional)
Whole Genome Libraries or Amplified Whole Genome Libraries

FIG. 19B

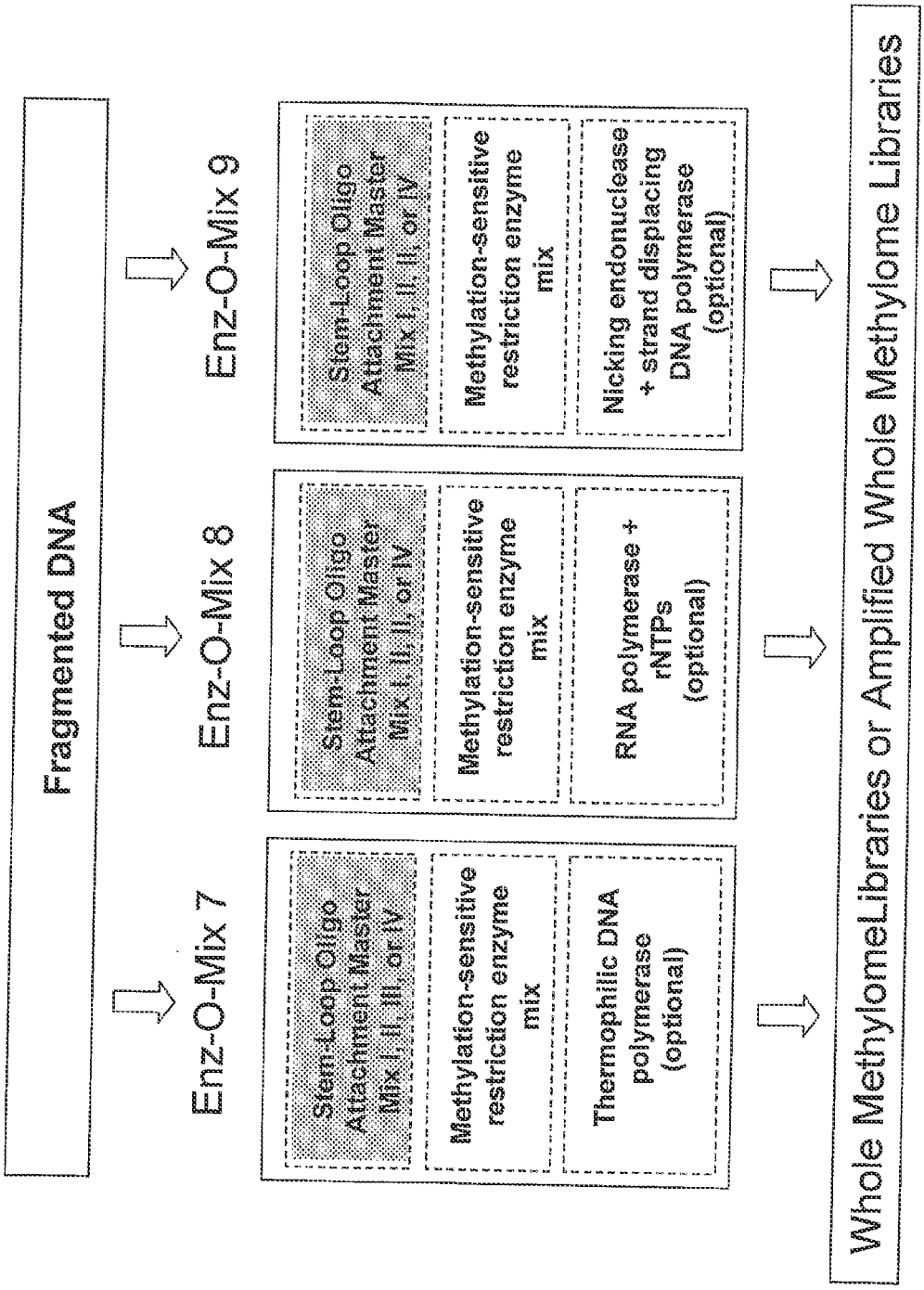

Fragmented DNA
Enz-O-Mix 7
Stem-Loop Oligo Attachment Master Mix I, II, III, or IV
Methylation-sensitive restriction enzyme mix
Thermophilic DNA polymerase (optional)
Enz-O-Mix 8
Stem-Loop Oligo Attachment Master Mix I, II, II, or IV
Methylation-sensitive restriction enzyme mix
RNA polymerase + rNTPs (optional)
Enz-O-Mix 9
Stem-Loop Oligo Attachment Master Mix I, II, II, or IV
Methylation-sensitive restriction enzyme mix
Nicking endonuclease + strand displacing DNA polymerase (optional)
Whole MethylomeLibraries or Amplified Whole Methylome Libraries

FIG. 19C

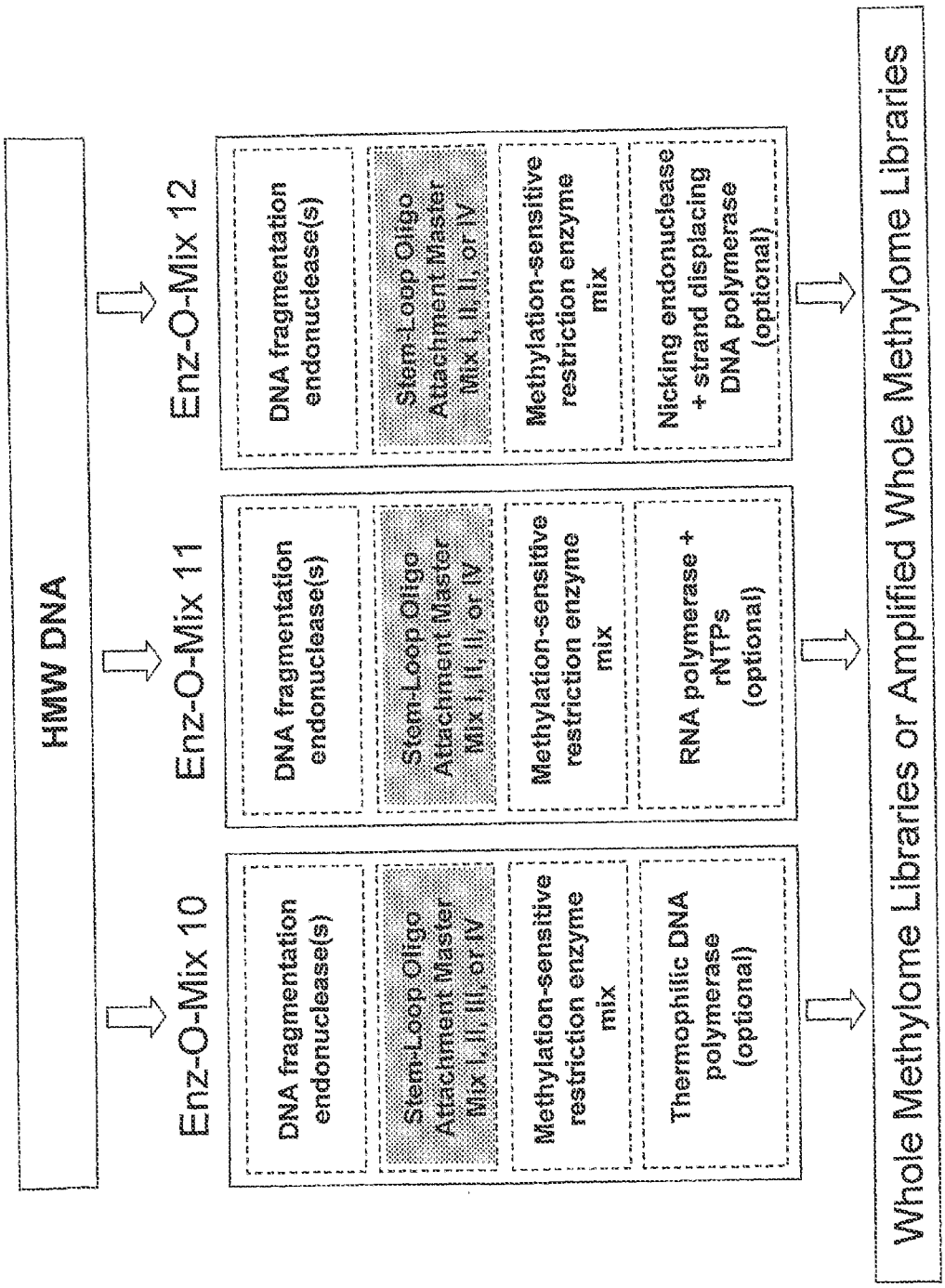

HMW DNA
Enz-O-Mix 10
DNA fragmentation endonuclease(s)
Stem-Loop Oligo Attachment Master Mix I, II, III, or IV
Methylation-sensitive restriction enzyme mix
Thermophilic DNA polymerase (optional)
Enz-O-Mix 11
DNA fragmentation endonuclease(s)
Stem-Loop Oligo Attachment Master Mix I, II, II, or IV
Methylation-sensitive restriction enzyme mix
RNA polymerase + rNTPs (optional)
Enz-O-Mix 12
DNA fragmentation endonuclease(s)
Stem-Loop Oligo Attachment Master Mix I, II, II, or IV
Methylation-sensitive restriction enzyme mix
Nicking endonuclease + strand displacing DNA polymerase (optional)
Whole Methylome Libraries or Amplified Whole Methylome Libraries

FIG. 19D

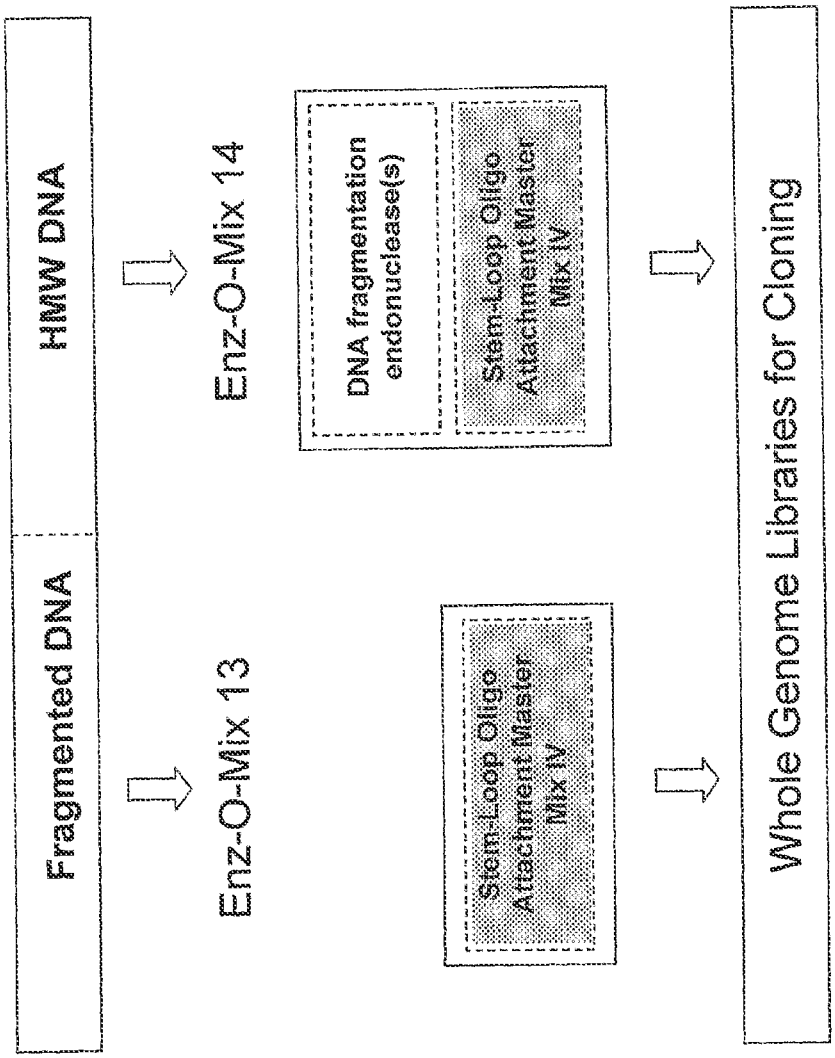

HMW DNA
Fragmented DNA
Enz-O-Mix 14
Enz-O-Mix 13
DNA fragmentation endonuclease(s)
Stem-Loop Oligo Attachment Master Mix IV
Stem-Loop Oligo Attachment Master Mix IV
Whole Genome Libraries for Cloning

FIG. 19E

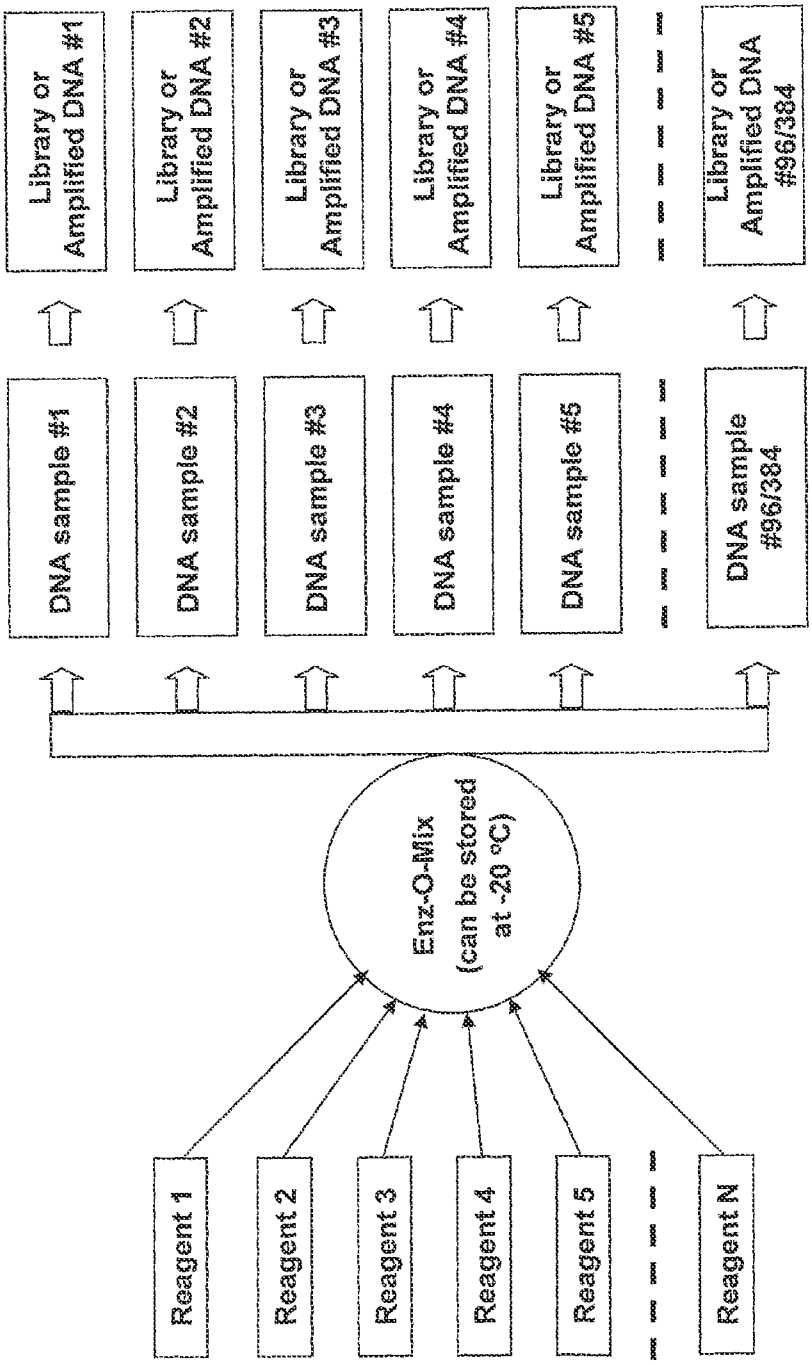

Reagent 1
Reagent 2
Reagent 3
Reagent 4
Reagent 5
Reagent N
Enz-O-Mix
(can be stored
at -20 °C)
DNA sample #1
DNA sample #2
DNA sample #3
DNA sample #4
DNA sample #5
DNA sample
#96/384
Library or
Amplified DNA #1
Library or
Amplified DNA #2
Library or
Amplified DNA #3
Library or
Amplified DNA #4
Library or
Amplified DNA #5
Library or
Amplified DNA
#96/384

FIG. 20

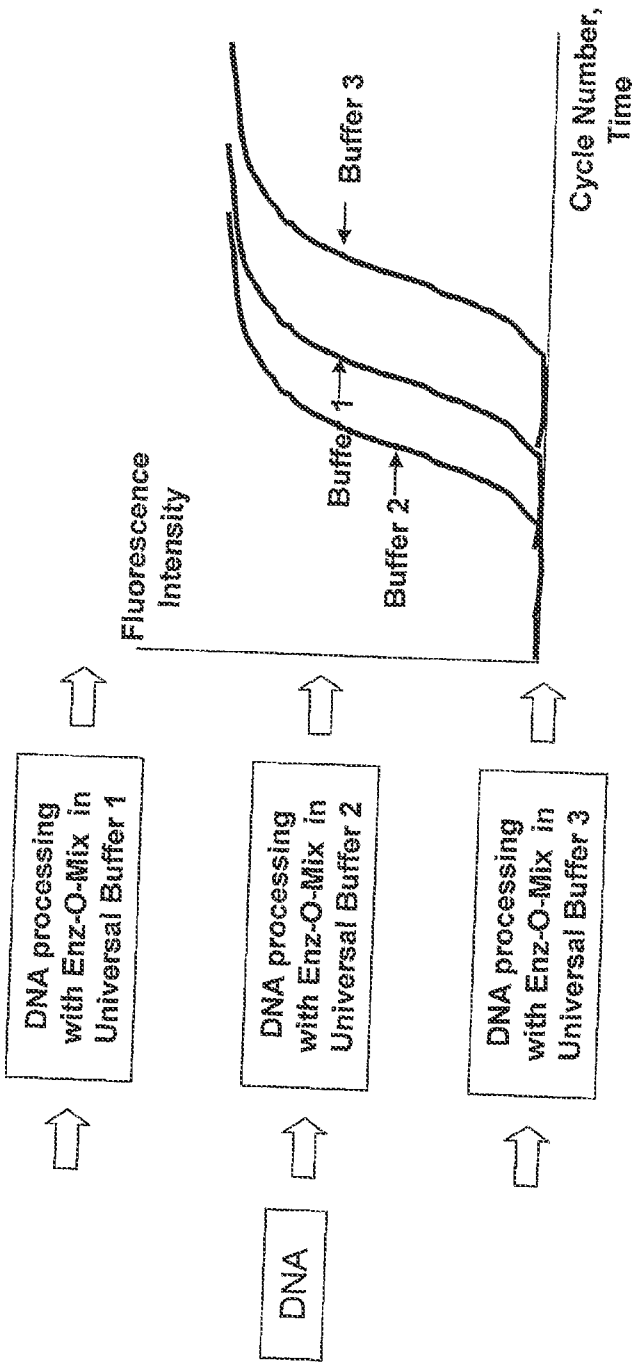

DNA
DNA processing with Enz-O-Mix in Universal Buffer 1
DNA processing with Enz-O-Mix in Universal Buffer 2
DNA processing with Enz-O-Mix in Universal Buffer 3
Fluorescence Intensity
Buffer 1
Buffer 2
Buffer 3
Cycle Number, Time

FIG. 21A

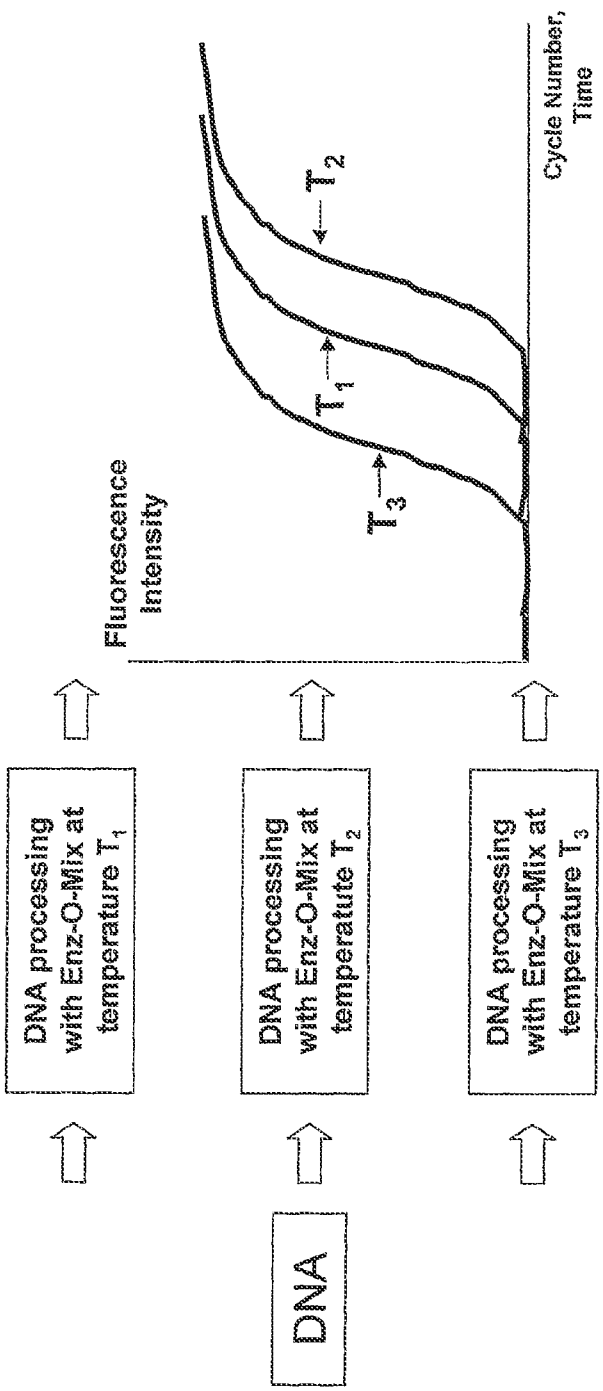

DNA
DNA processing with Enz-O-Mix at temperature T_1
DNA processing with Enz-O-Mix at temperatute T_2
DNA processing with Enz-O-Mix at temperature T_3
Fluorescence Intensity
T_3
T_1
T_2
Cycle Number, Time

FIG. 21B

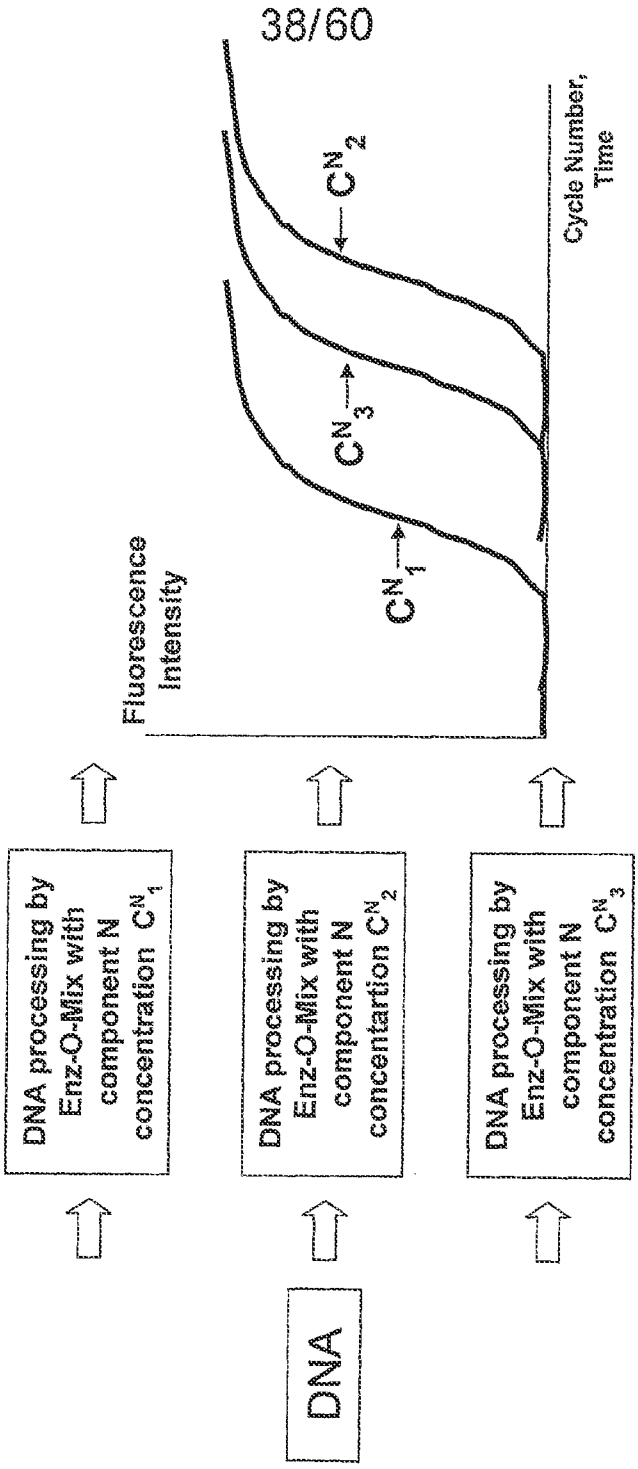

DNA
DNA processing by Enz-O-Mix with component N concentration C^N_1
DNA processing by Enz-O-Mix with component N concentartion C^N_2
DNA processing by Enz-O-Mix with component N concentration C^N_3
Fluorescence Intensity
C^N_1
C^N_3
C^N_2
Cycle Number, Time

FIG. 21C

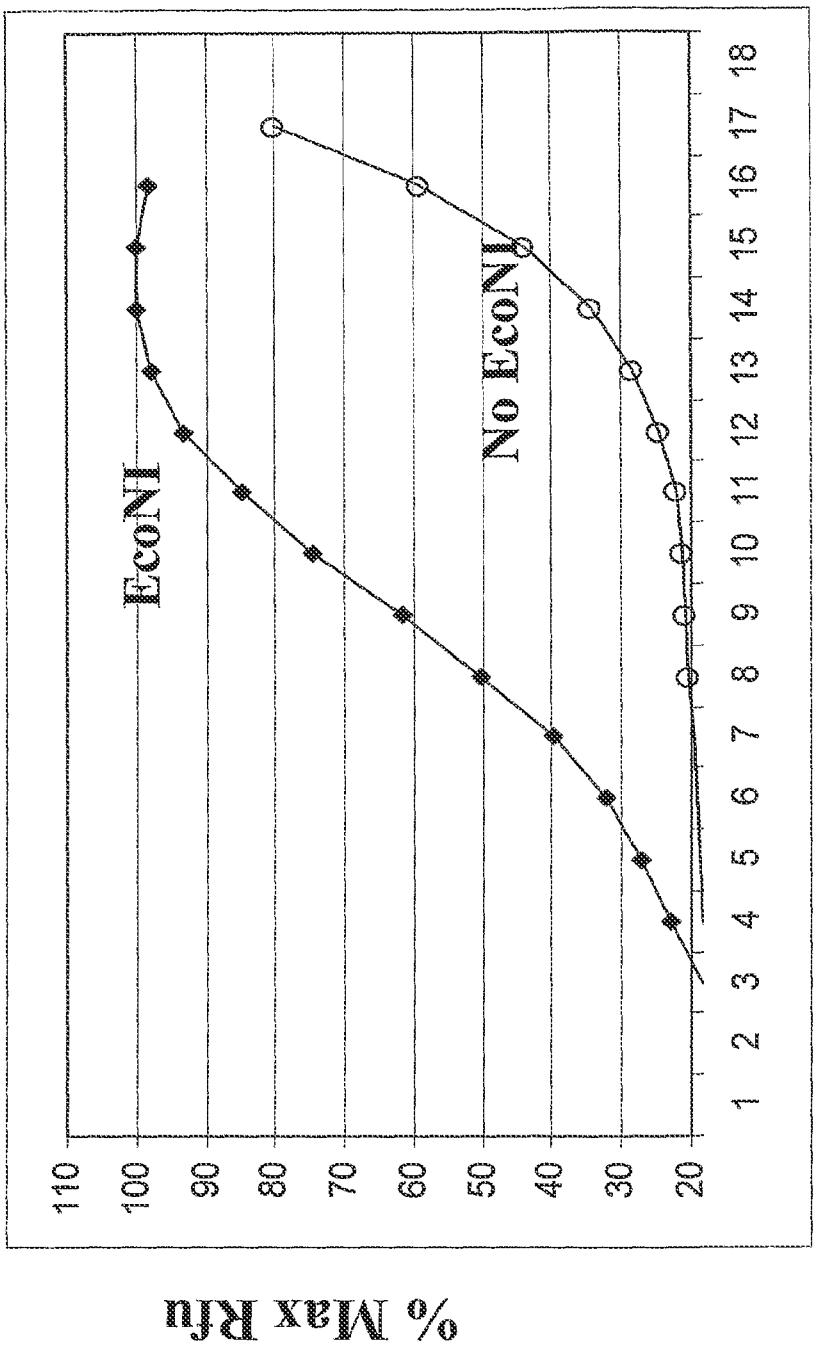

EcoNI
No EcoNI
% Max Rfu
Cycle Number
110
100
90
80
70
60
50
40
30
20
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18

FIG. 22

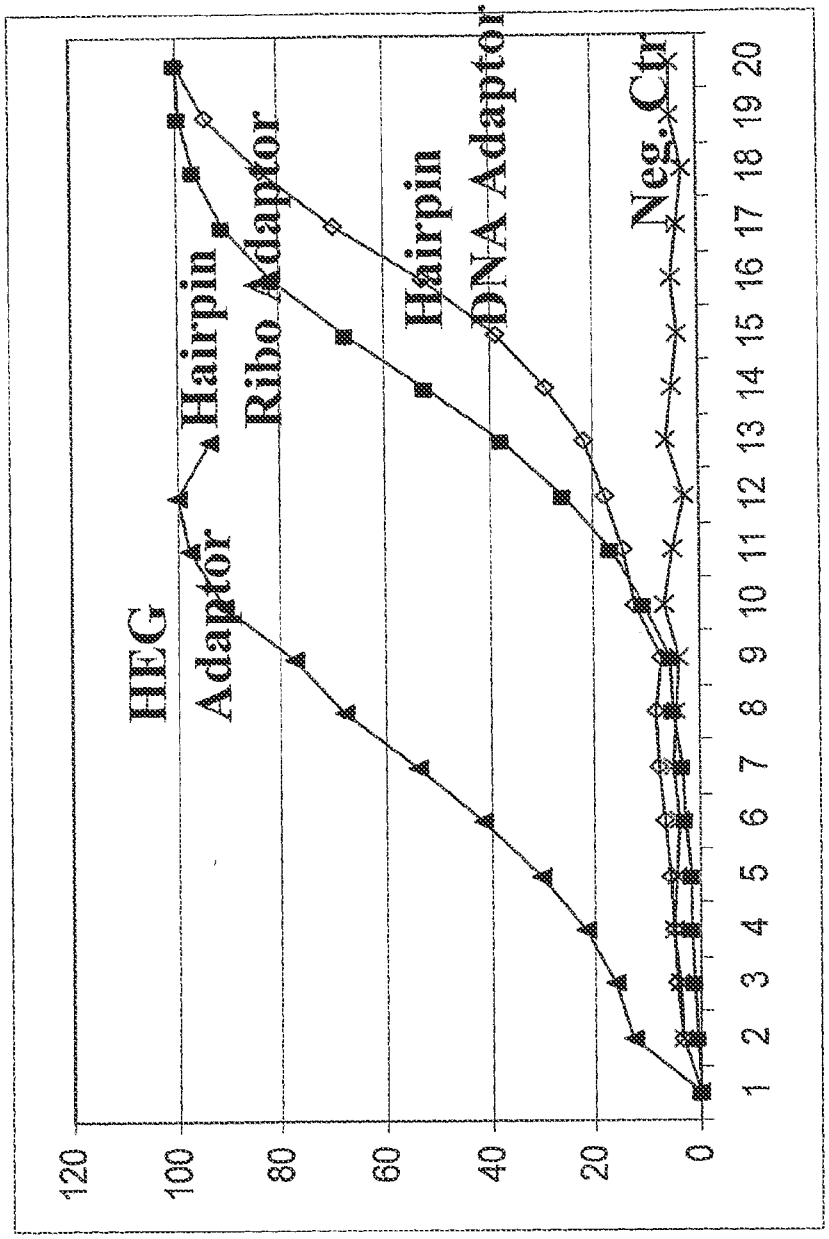

% Max Rfu
120
100
80
60
40
20
0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
Cycle Number
HEG
Adaptor
Hairpin
Ribo Adaptor
Hairpin
DNA Adaptor
Neg. Ctr

FIG. 23

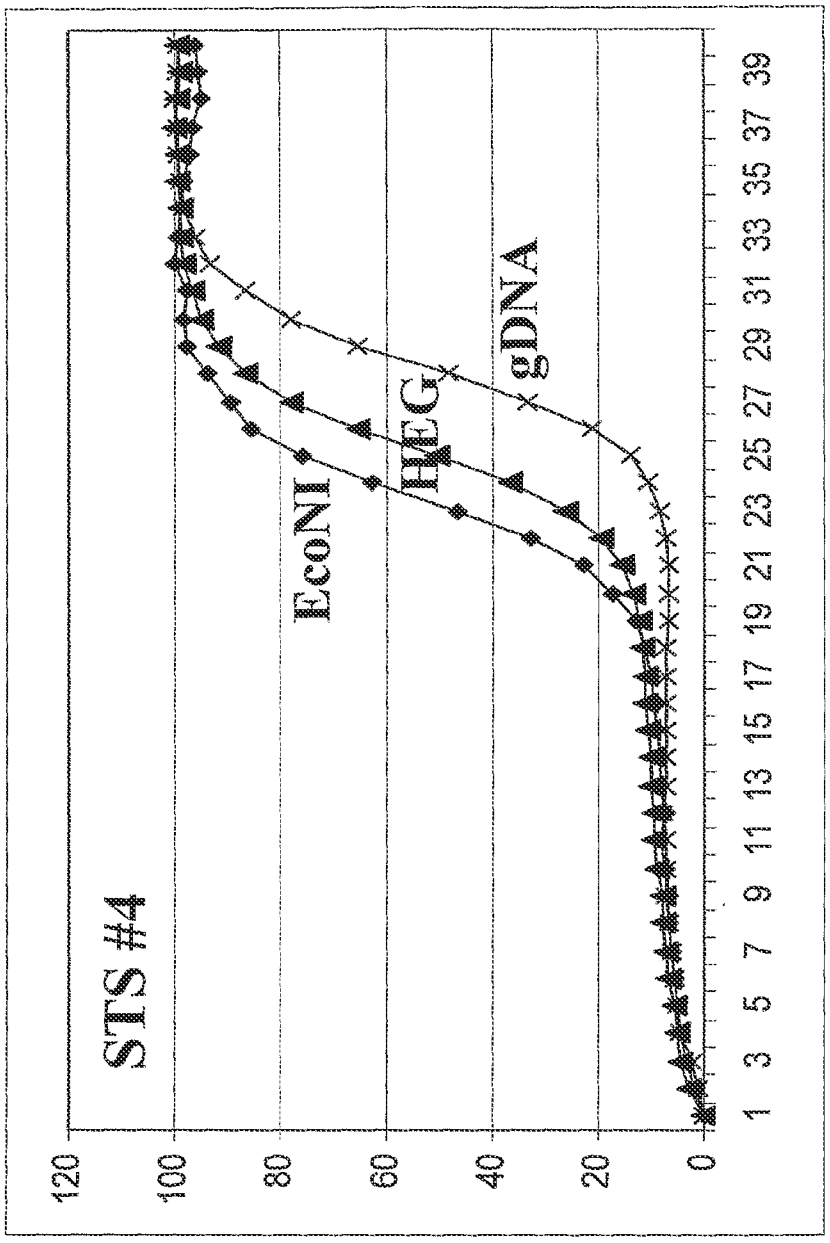

STS #4
EcoNI
HEG
gDNA
% Max Rfu
120
100
80
60
40
20
0
1 3 5 7 9 11 13 15 17 19 21 23 25 27 29 31 33 35 37 39
Cycle Number

FIG. 24A

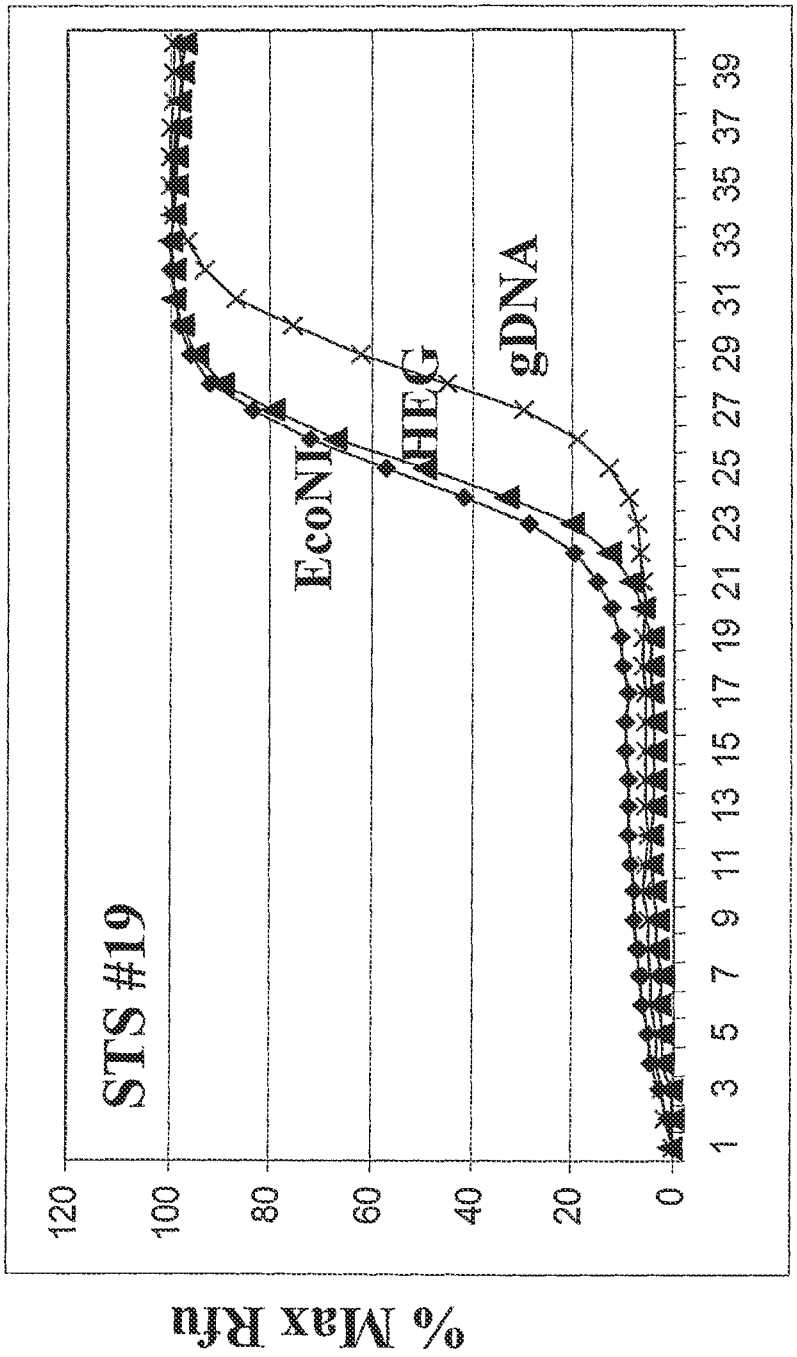

STS #19
EcoNI
AHEG
gDNA
% Max Rfu
120
100
80
60
40
20
0
1 3 5 7 9 11 13 15 17 19 21 23 25 27 29 31 33 35 37 39
Cycle Number

FIG. 24B

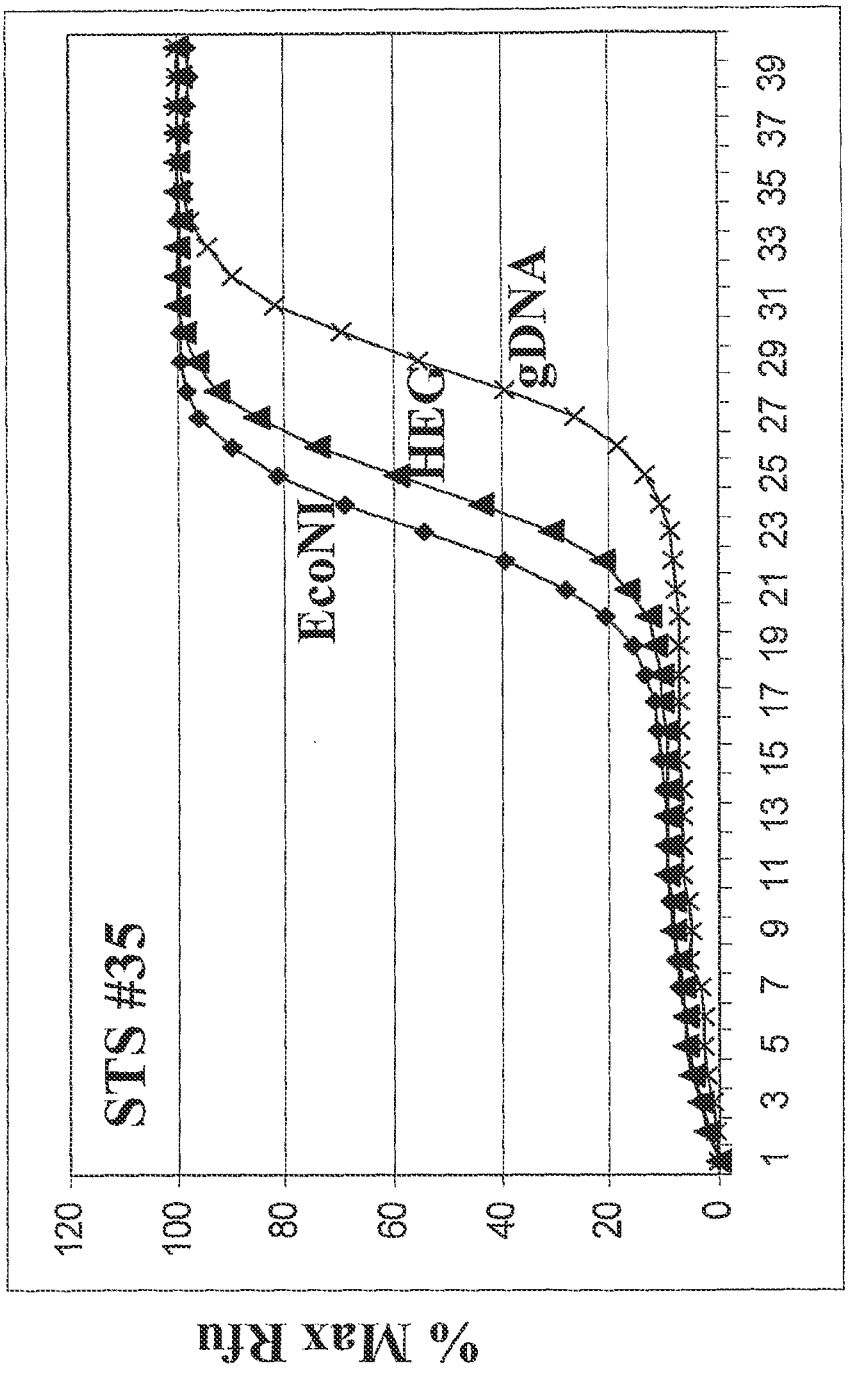

STS #35
EcoNI
HEG
gDNA
% Max Rfu
Cycle Number
0
20
40
60
80
100
120
1 3 5 7 9 11 13 15 17 19 21 23 25 27 29 31 33 35 37 39

FIG. 24C

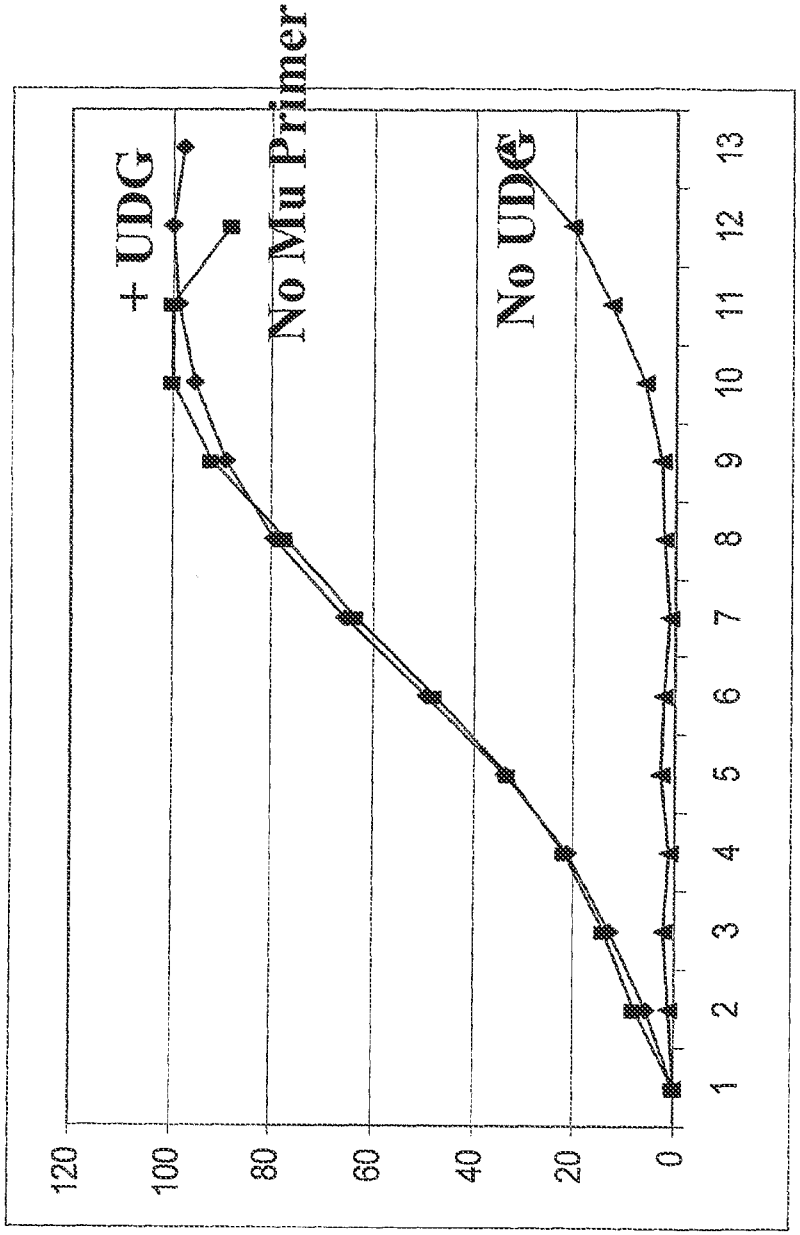

+ UDG
No Mu Primer
No UDG
% Max Rfn
120
100
80
60
40
20
0
1
2
3
4
5
6
7
8
9
10
11
12
13
Cycle Number

FIG. 25

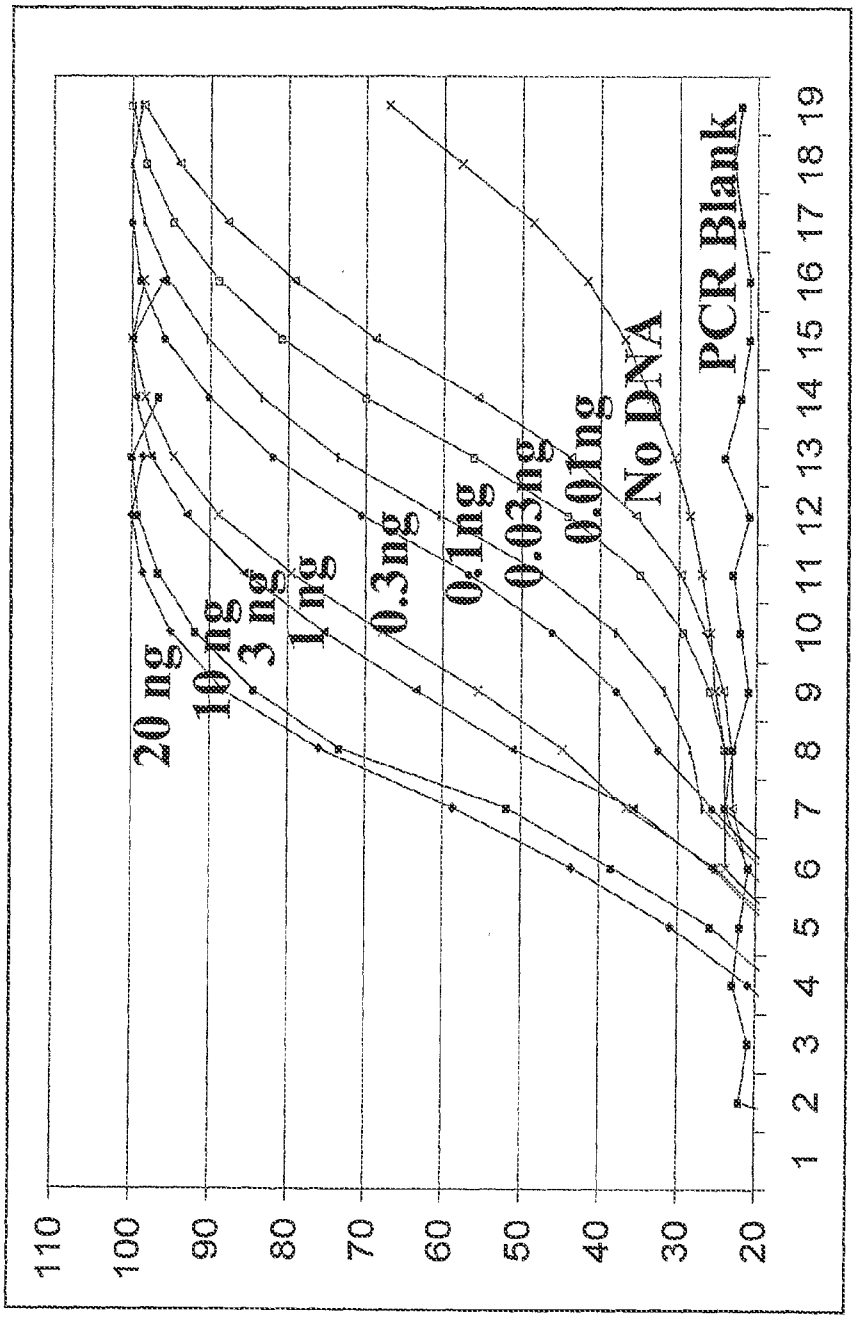

% Max Rfu
110
100
90
80
70
60
50
40
30
20
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19
Cycle Number
20 ng
10 ng
3 ng
1 ng
0.3ng
0.1ng
0.03ng
0.01ng
No DNA
PCR Blank

FIG. 26

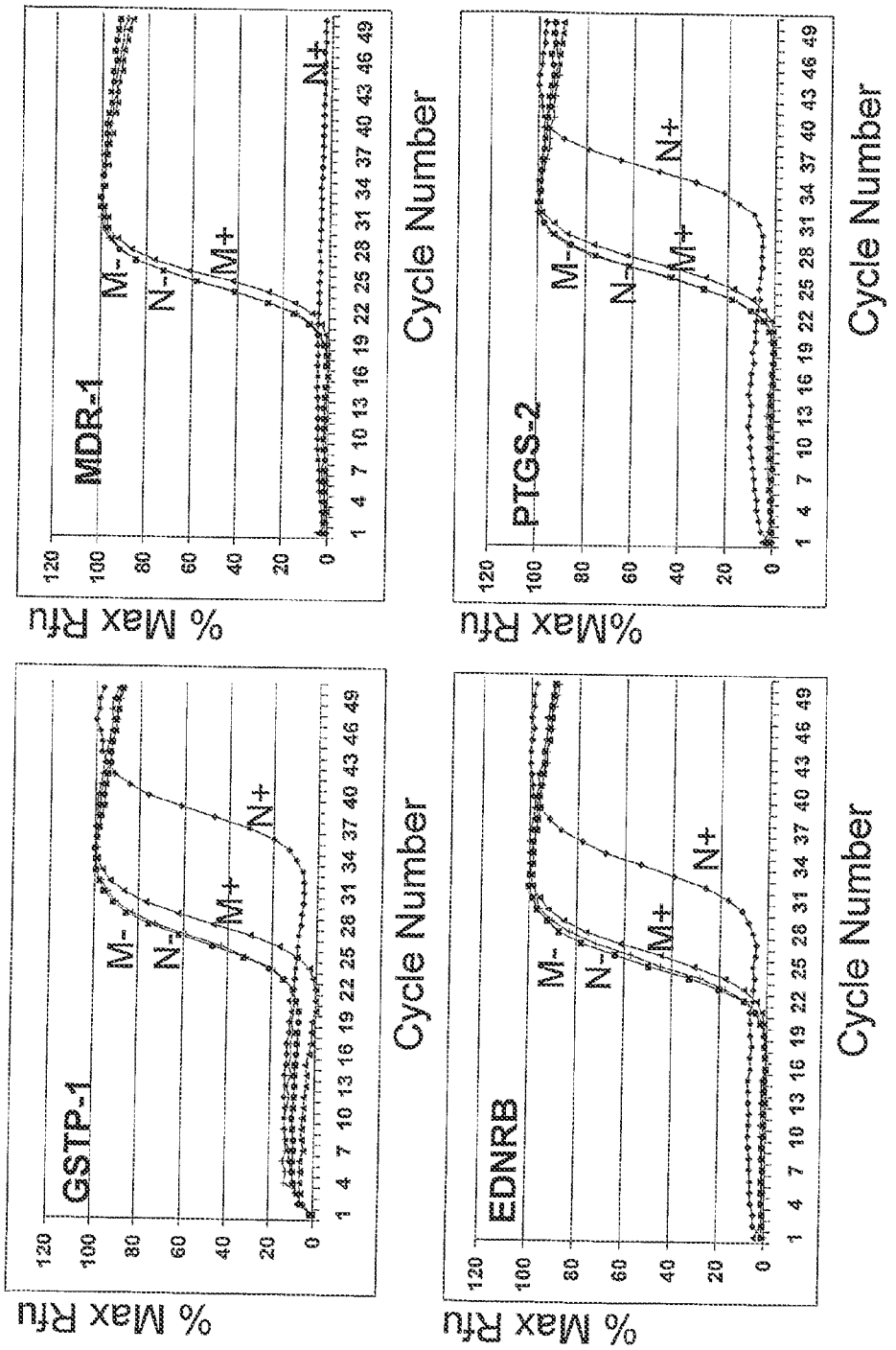

MDR-1
PTGS-2
GSTP-1
EDNRB
M-
N-
M+
N+
% Max Rfu
Cycle Number
N+ = Non-methylated Cut, N- = Non-methylated Uncut
M+ = Methylated Cut M- = Methylated Uncut

FIG. 27

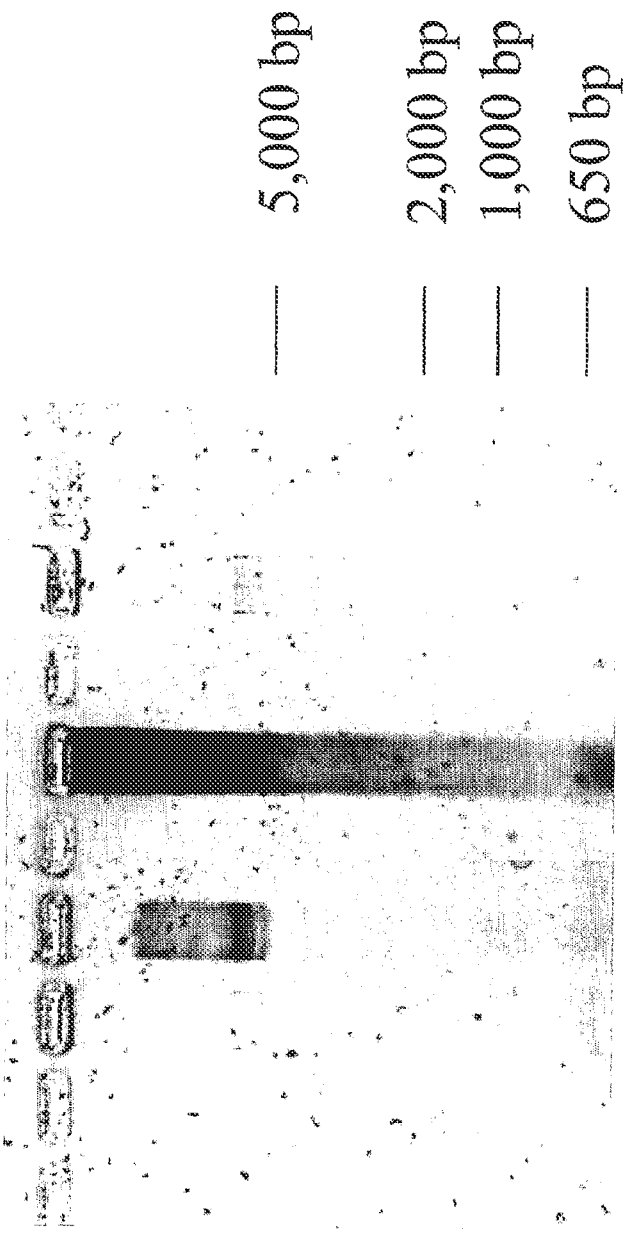

5,000 bp
2,000 bp
1,000 bp
650 bp
1Kb+ ladder
-T4 Ligase
+T4 Ligase
1Kb+ ladder

FIG. 28

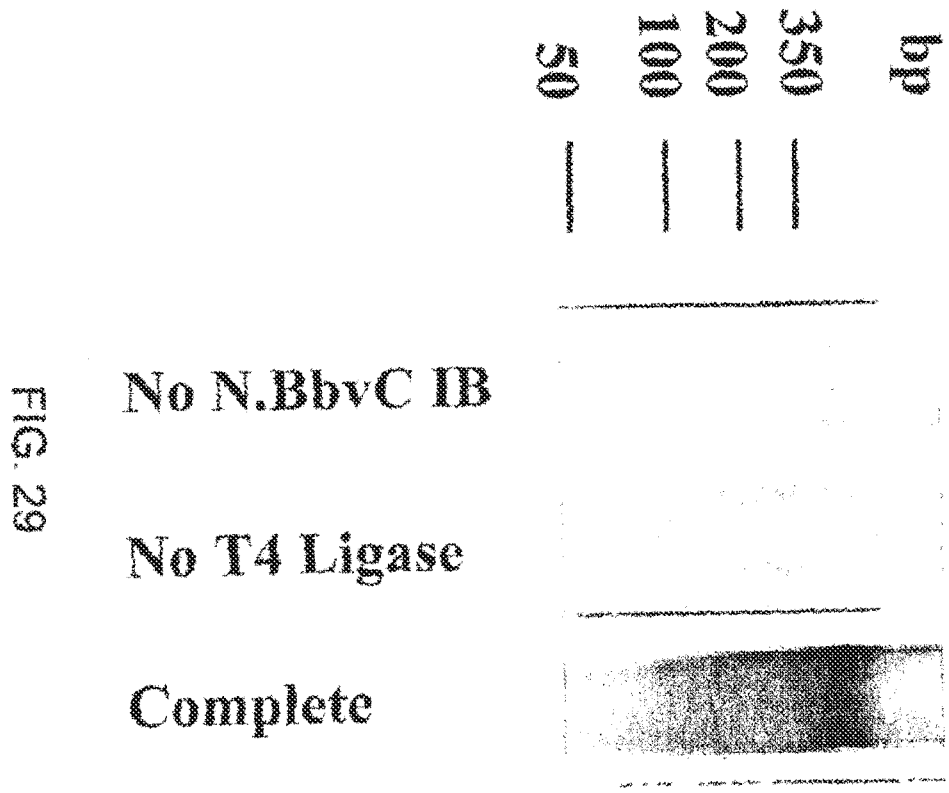


FIG. 29

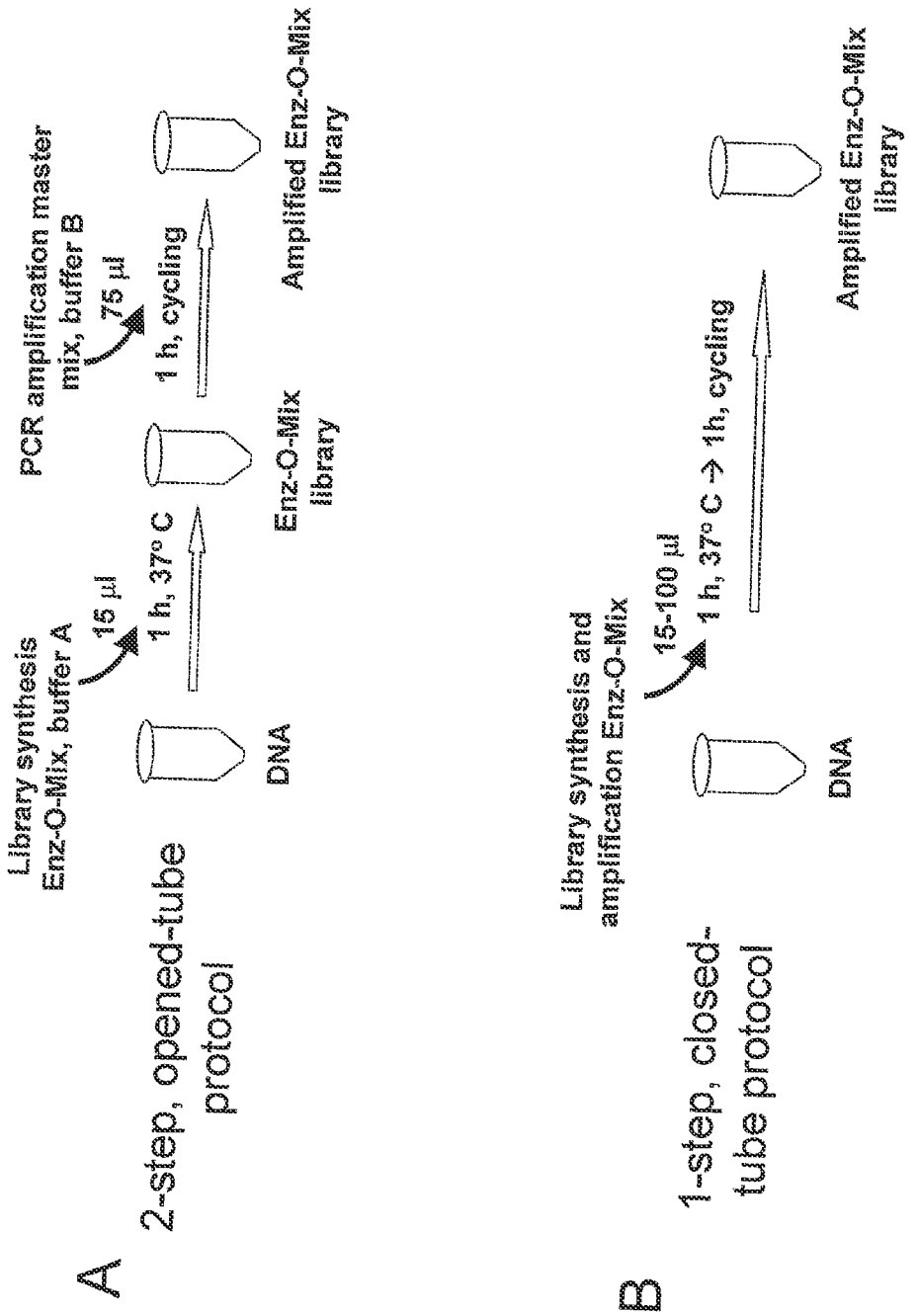

A
2-step, opened-tube protocol
Library synthesis
Enz-O-Mix, buffer A
15 µl
1 h, 37° c
DNA
Enz-O-Mix library
PCR amplification master mix, buffer B
75 µl
1 h, cycling
Amplified Enz-O-Mix library
B
1-step, closed-tube protocol
Library synthesis and amplification Enz-O-Mix
15-100 µl
1 h, 37° c → 1h, cycling
DNA
Amplified Enz-O-Mix library

FIG. 30

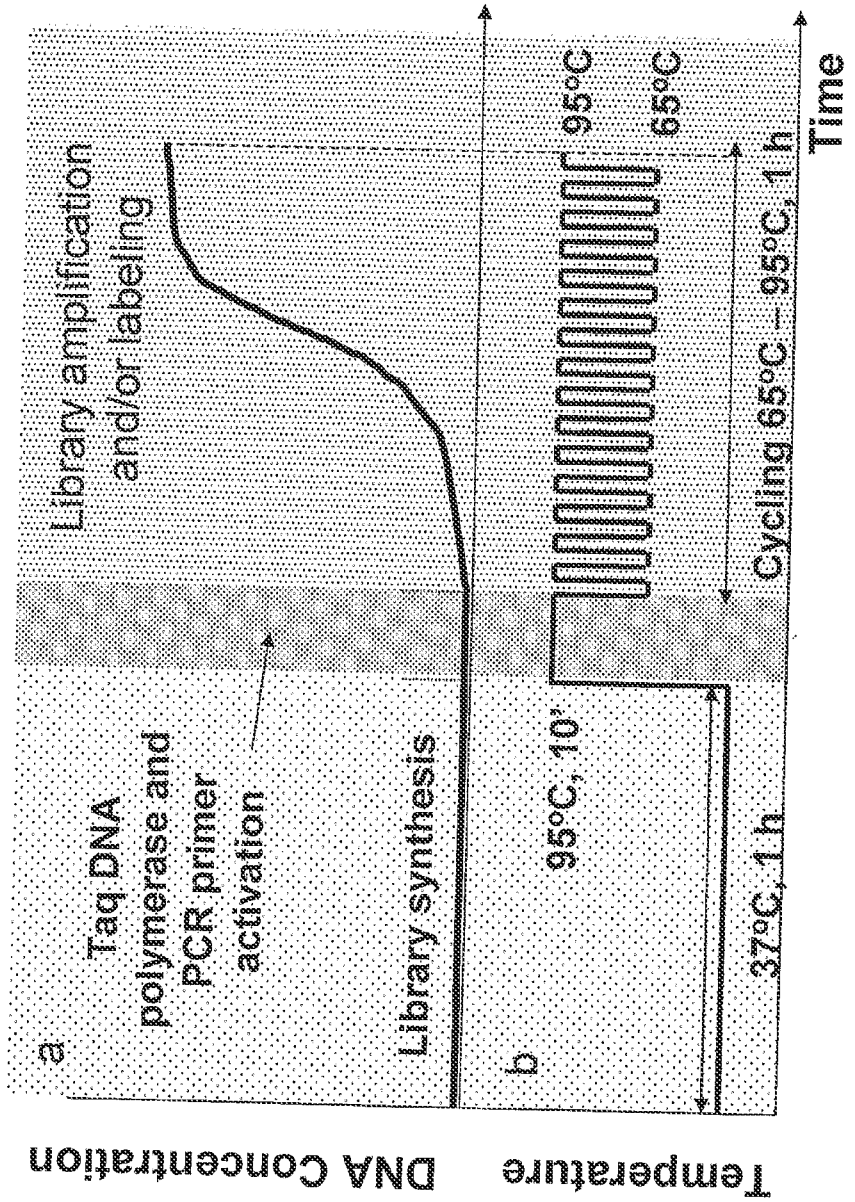

DNA Concentration
Temperature
a
Taq DNA polymerase and PCR primer activation
Library synthesis
Library amplification and/or labeling
b
95°C, 10'
37°C, 1 h
95°C
65°C
Cycling 65°C – 95°C, 1 h
Time

FIG. 31

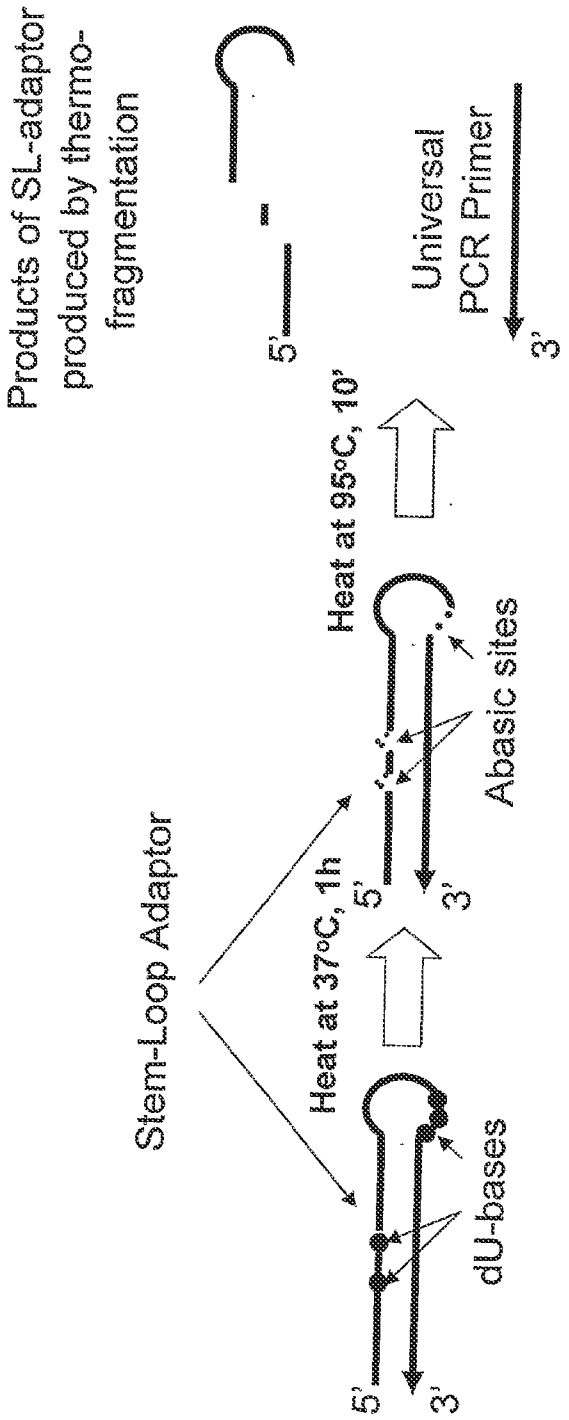

Products of SL-adaptor produced by thermo-fragmentation
5'
3'
Universal PCR Primer
Heat at 95°C, 10'
Abasic sites
Heat at 37°C, 1h
Stem-Loop Adaptor
dU-bases
5'
3'
5'
3'

FIG. 32

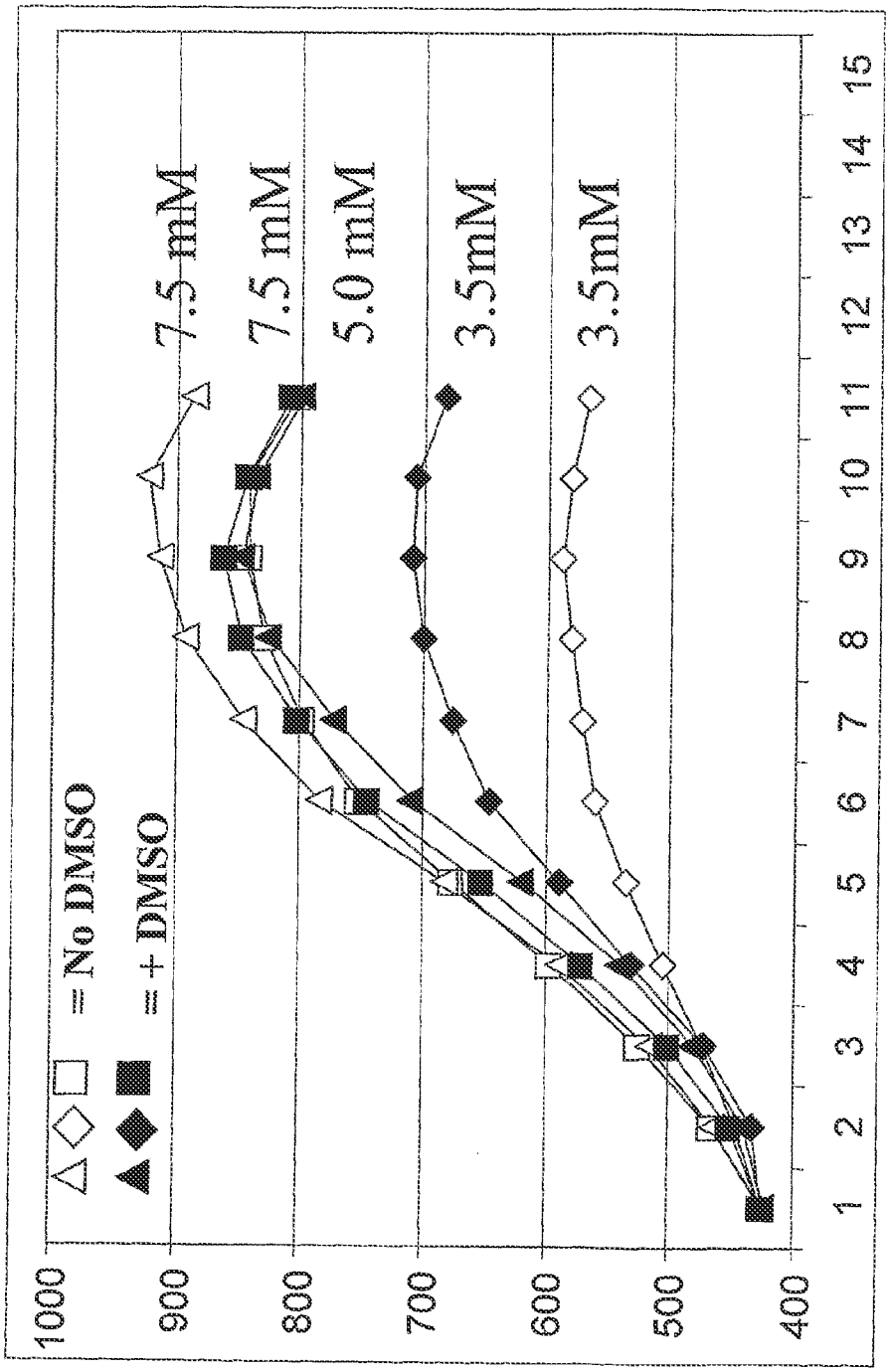

= No DMSO
= + DMSO
7.5 mM
7.5 mM
5.0 mM
3.5mM
3.5mM
1000
900
800
700
600
500
400
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15

FIG. 33

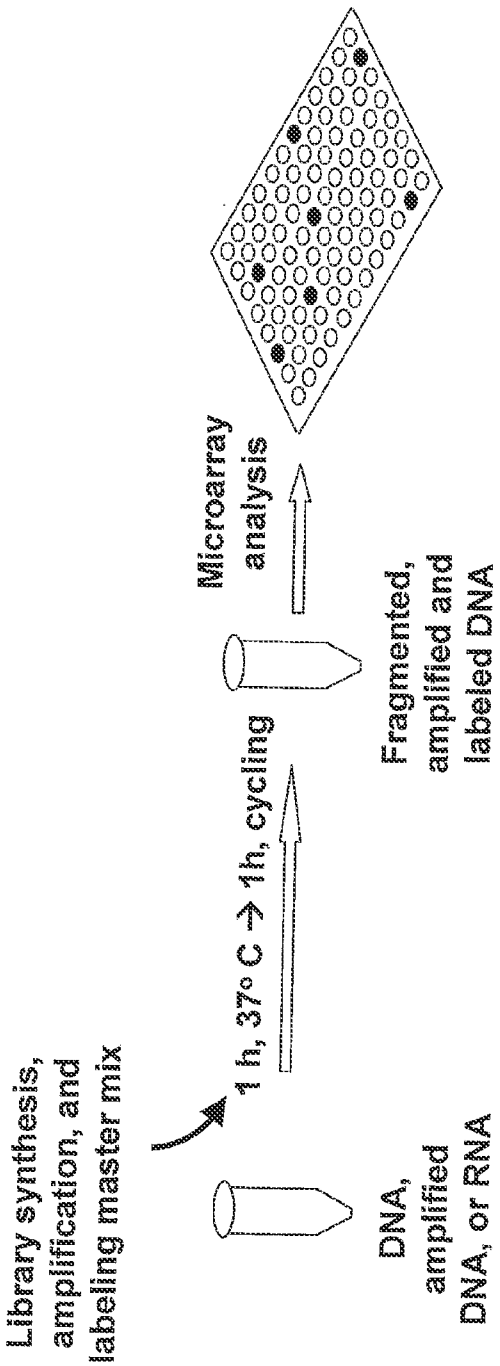

Library synthesis, amplification, and labeling master mix
1 h, 37° C → 1h, cycling
DNA, amplified DNA, or RNA
Fragmented, amplified and labeled DNA
Microarray analysis

FIG. 34

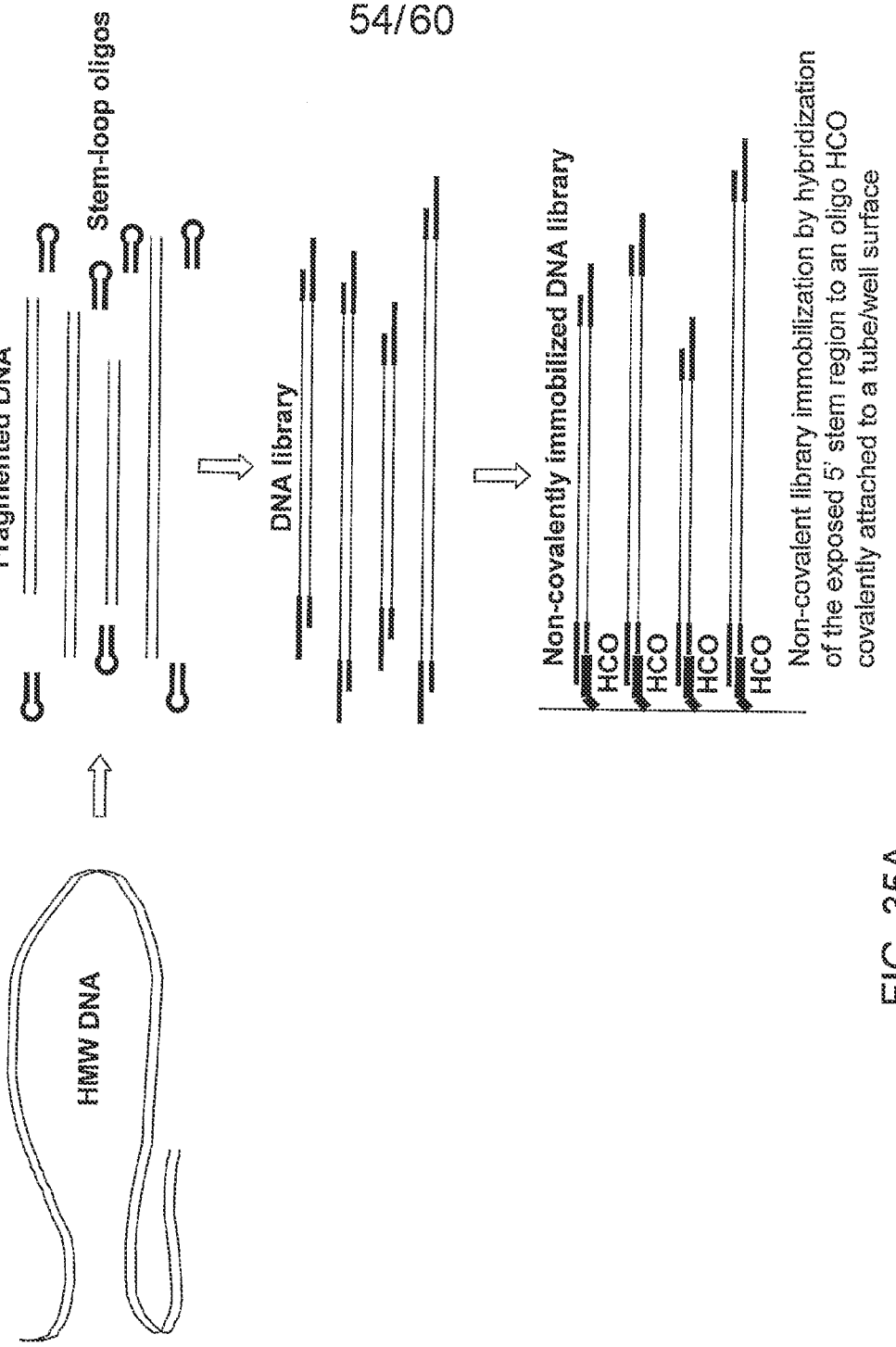

HMW DNA
Fragmented DNA
Stem-loop oligos
DNA library
Non-covalently immobilized DNA library
HCO
HCO
HCO
HCO
Non-covalent library immobilization by hybridization of the exposed 5' stem region to an oligo HCO covalently attached to a tube/well surface

FIG. 35A

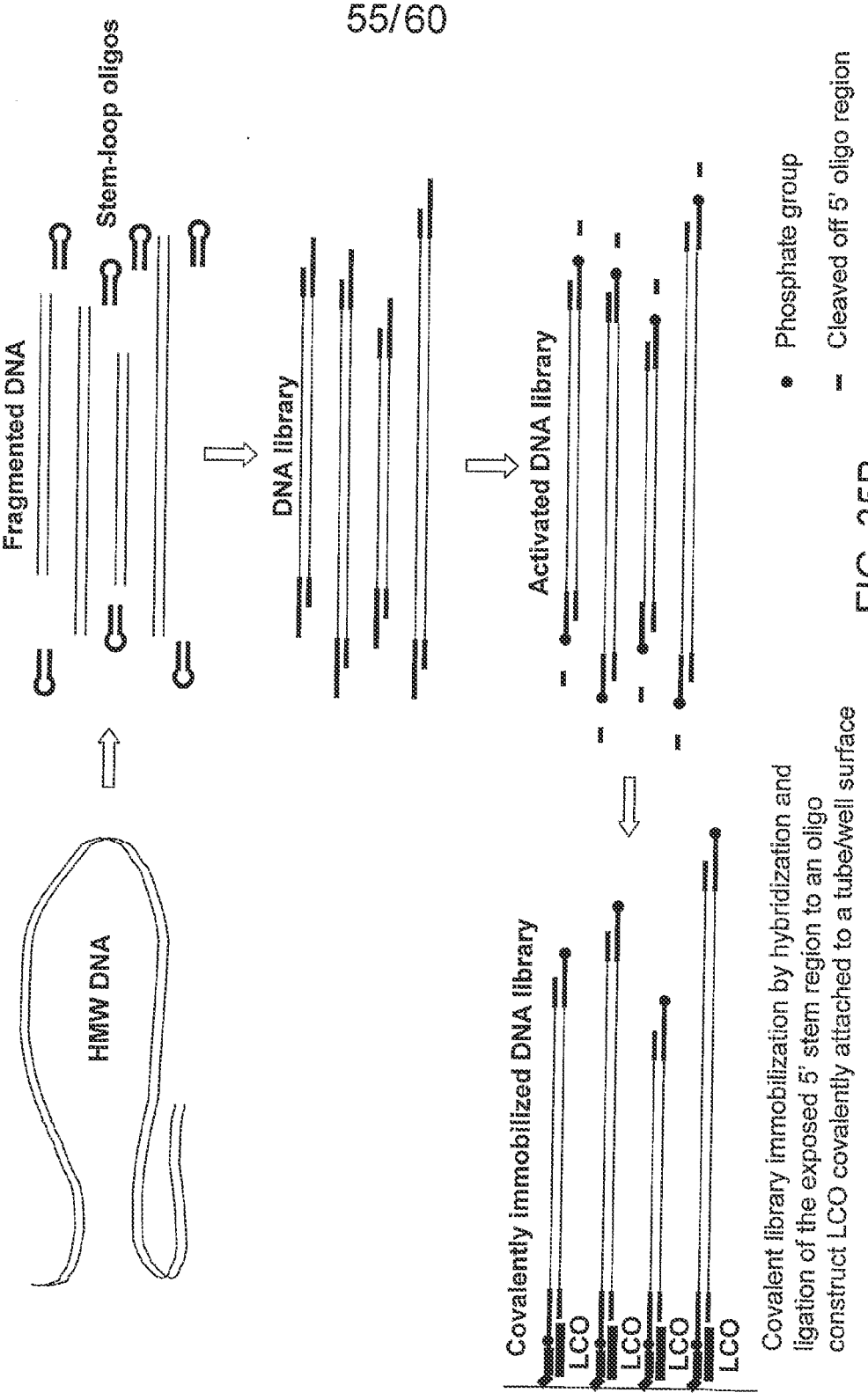

HMW DNA
Fragmented DNA
Stem-loop oligos
DNA library
Activated DNA library
Covalently immobilized DNA library
LCO
LCO
LCO
LCO
Covalent library immobilization by hybridization and ligation of the exposed 5' stem region to an oligo construct LCO covalently attached to a tube/well surface
Phosphate group
Cleaved off 5' oligo region

FIG. 35B

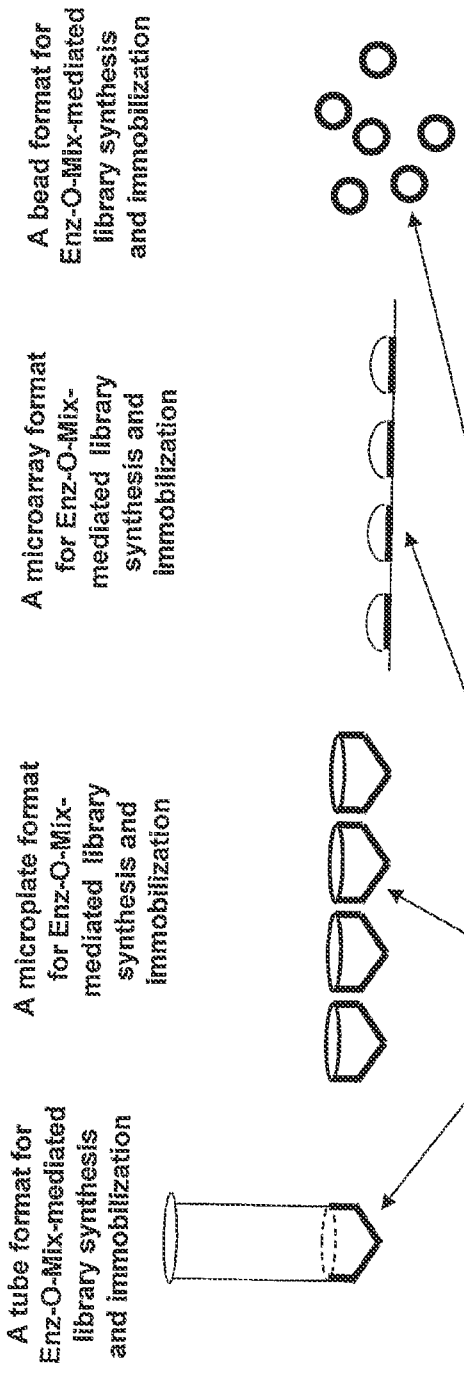

A tube format for Enz-O-Mix-mediated library synthesis and immobilization
A microplate format for Enz-O-Mix-mediated library synthesis and immobilization
A microarray format for Enz-O-Mix-mediated library synthesis and immobilization
A bead format for Enz-O-Mix-mediated library synthesis and immobilization
The solid black lines show the area of a capture oligonucleotide attachment

FIG. 36

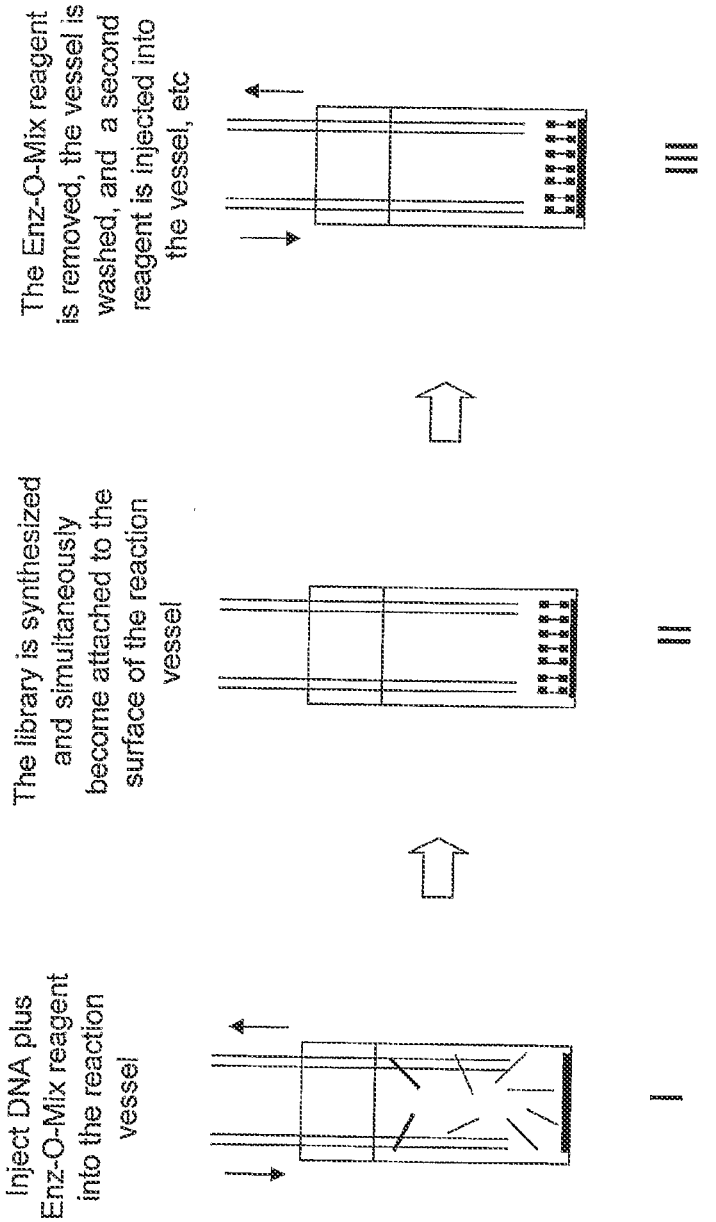

Inject DNA plus Enz-O-Mix reagent into the reaction vessel
The library is synthesized and simultaneously become attached to the surface of the reaction vessel
The Enz-O-Mix reagent is removed, the vessel is washed, and a second reagent is injected into the vessel, etc
I
II
III

FIG. 37

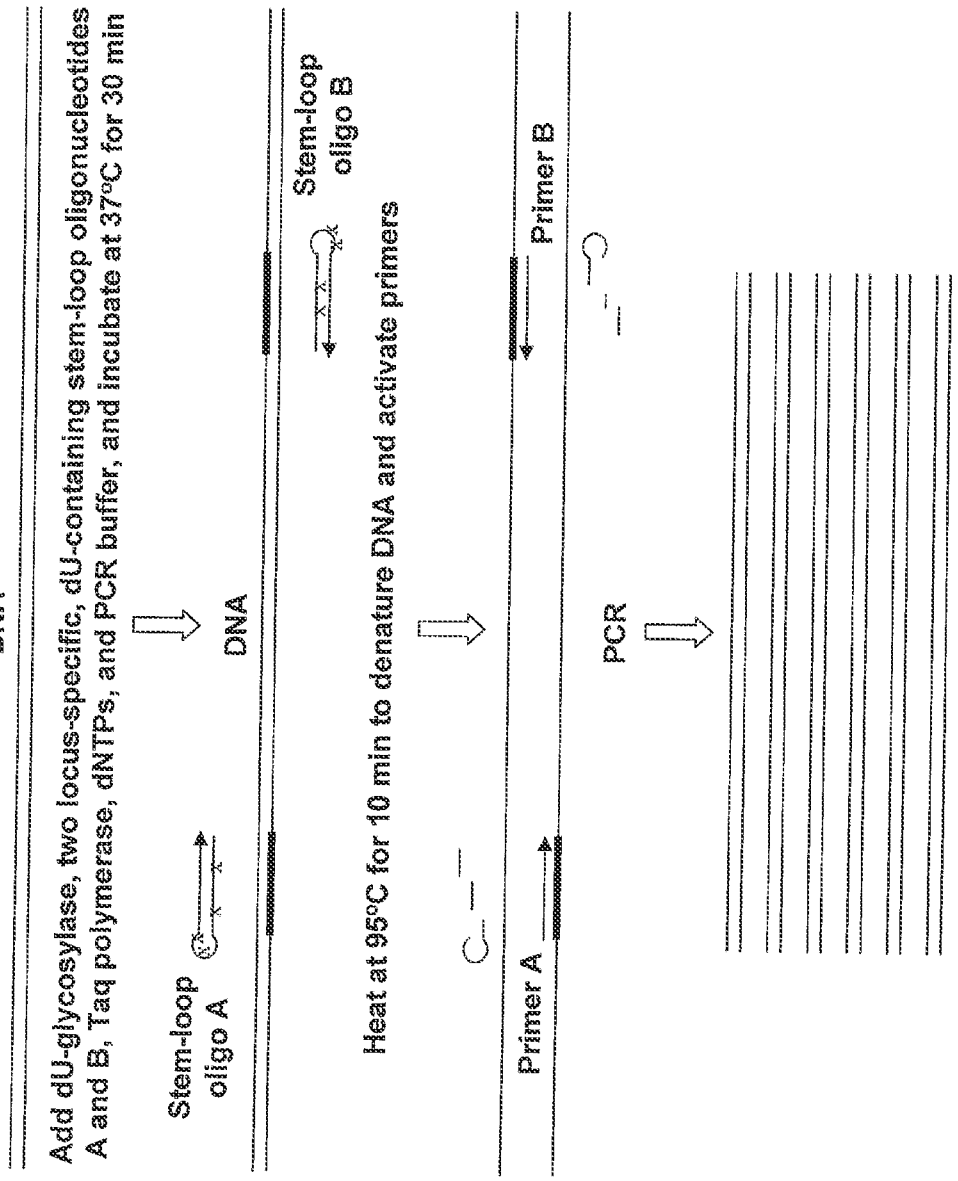

DNA
Add dU-glycosylase, two locus-specific, dU-containing stem-loop oligonucleotides A and B, Taq polymerase, dNTPs, and PCR buffer, and incubate at 37°C for 30 min
Stem-loop oligo A
DNA
Stem-loop oligo B
Heat at 95°C for 10 min to denature DNA and activate primers
Primer A
Primer B
PCR

FIG. 38

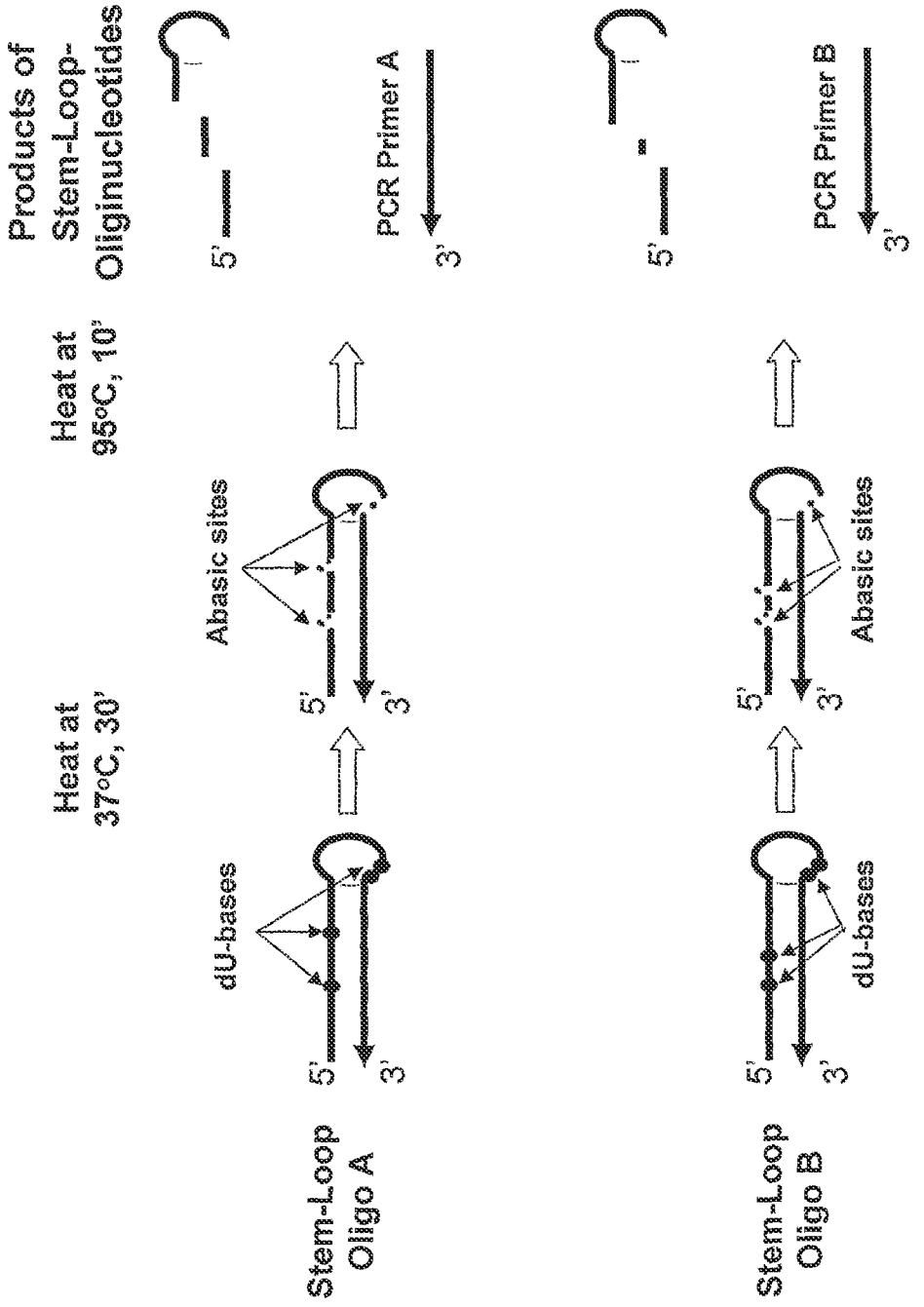

Products of Stem-Loop-Oligonucleotides
Heat at 95°C, 10'
Heat at 37°C, 30'
Stem-Loop Oligo A
dU-bases
Abasic sites
PCR Primer A
Stem-Loop Oligo B
dU-bases
Abasic sites
PCR Primer B
5'
3'

FIG. 39

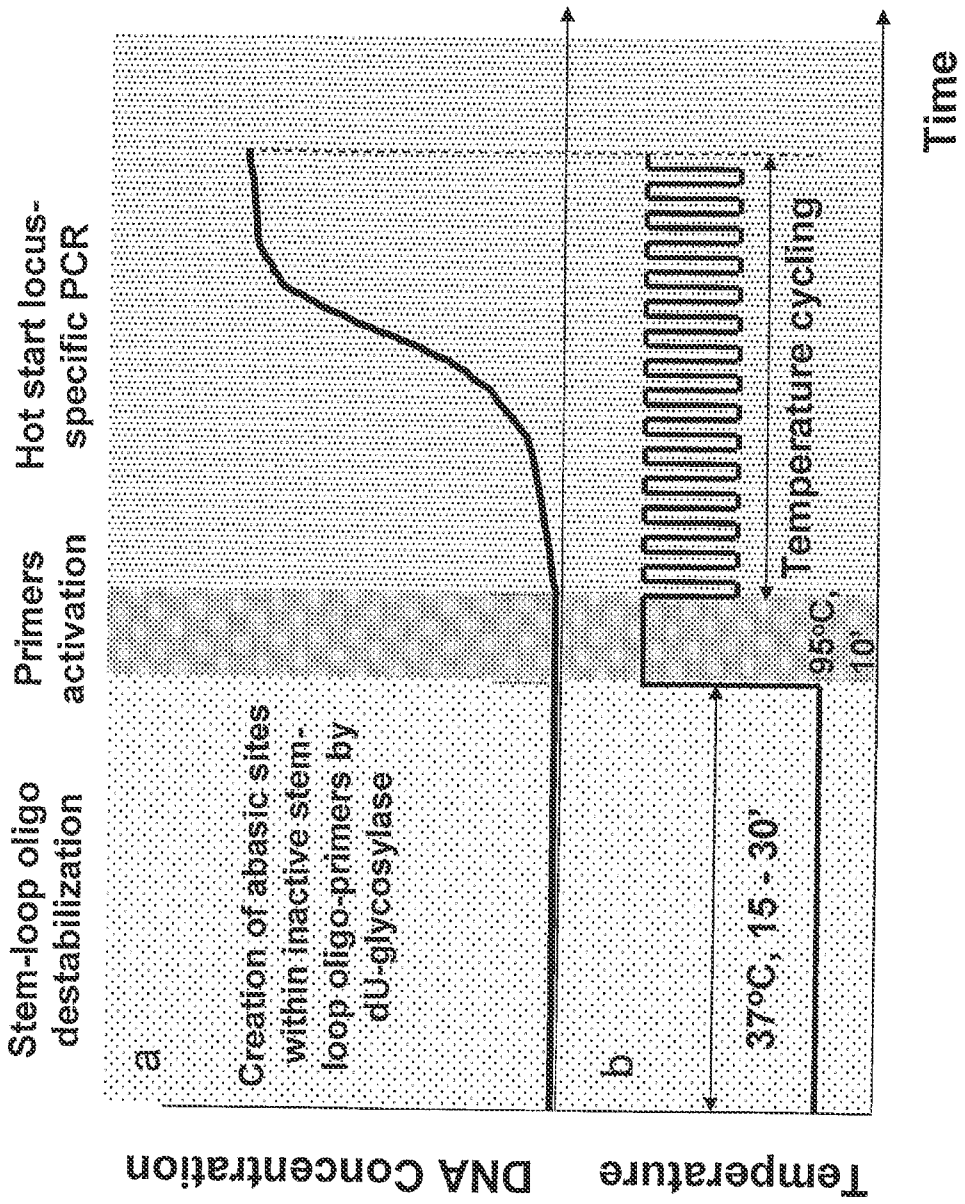

DNA Concentration
Temperature
Time
a
Stem-loop oligo destabilization
Creation of abasic sites within inactive stem-loop oligo-primers by dU-glycosylase
Primers activation
Hot start locus-specific PCR
b
95°C, 10'
37°C, 15 - 30'
Temperature cycling

FIG. 40

RUBC026WO.APP.txt
SEQUENCE LISTING

110> MAKAROV, VLADIMIR L.
KAMBEROV, EMMANUEL
TARRIER, BRENDAN

120> COMPOSITIONS AND METHODS FOR PROCESSING AND AMPLIFICATION OF DNA, INCLUDING USING MULTIPLE ENZYMES IN A SINGLE REACTION

130> RUBC:026WO

140> UNKNOWN
141> 2006-03-02

150> 60/704,932
151> 2005-08-02

160> 20

170> PatentIn Ver. 2.1

210> 1
211> 32
212> DNA
213> Artificial Sequence

220>
223> Description of Artificial Sequence: Synthetic Primer

400> 1
aaacacac ccaacacacc taaaaaaggt gt 32

210> 2
211> 18
212> DNA
213> Artificial Sequence

220>
223> Description of Artificial Sequence: Synthetic Primer

400> 2
tgttgggt gtgtttgg 18

210> 3
211> 40
212> DNA
213> Artificial Sequence

220>
223> Description of Artificial Sequence: Synthetic Primer

300> 3
gagagagg gaaggaagaa aacttccttc cctctctctt 40

210> 4
211> 18

IDT Exhibit 1020.7683

<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic Primer

<400> 4
cttccttccc tctctctt 18

<210> 5
<211> 28
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic Primer

<400> 5
ccaaacacac ccaacacaaa aagtgttg 28

<210> 6
<211> 28
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic Primer

<400> 6
ccaaacacac ccaacacaaa aagtgttg 28

<210> 7
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<221> modified_base
<222> (10)..(24)
<223> U = a, c, g or t

<220>
<223> Description of Artificial Sequence: Synthetic Primer

<400> 7
tgtgttgggu gugtgtgguu uuuuccacac acacccaaca 40

<210> 8
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic Primer

IDT Exhibit 1020.7684

<400> 8
ccacacacac ccaacaca 18

<210> 9
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic Primer

<400> 9
ggaggagtct ttcgagttca a 21

<210> 10
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic Primer

<400> 10
cgggaggaat acagacacgt ctt 23

<210> 11
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic Primer

<400> 11
ggaaagaggg aaaggcttc 19

<210> 12
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic Primer

<400> 12
ccccagtgct gagtcacgg 19

<210> 13
<211> 20
<212> DNA
<213> Artificial Sequence

<220>

IDT Exhibit 1020.7685

<223> Description of Artificial Sequence: Synthetic Primer

<400> 13
gggtgggagg aagcatcgtc 20

<210> 14
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic Primer

<400> 14
ggtctccagc atctccacga a 21

<210> 15
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic Primer

<400> 15
agaactggct ctcggaagcg 20

<210> 16
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic Primer

<400> 16
gggagcagag ggggtagtc 19

<210> 17
<211> 60
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic Primer

<400> 17
tgtgttgggg gtgtggattt aatacgactc actataggga gaccacacac acccaacaca 60

<210> 18
<211> 62
<212> DNA
<213> Artificial Sequence

IDT Exhibit 1020.7686

<220>
<223> Description of Combined DNA/RNA Molecule:xx

<220>
<221> modified_base
<222> (10)..(24)
<223> U = a, c, g or t

<220>
<223> Description of Artificial Sequence: Synthetic Primer

<400> 18
tgtgttgggu gugtgtgguu uuuuatttaa tacgaccctc agcaccacac acacccaaca ca 62

<210> 19
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Combined DNA/RNA Molecule:xx

<220>
<223> Description of Artificial Sequence: Synthetic Primer

<400> 19
atttaatacg accctcagca ccac 24

<210> 20
<211> 11
<212> DNA
<213> Artificial Sequence

<220>
<221> modified_base
<222> (4)..(8)
<223> N = a, c, g, and/or t/u

<220>
<223> Description of Artificial Sequence: Synthetic Primer

<400> 20
cctnnnnnag g 11

IDT Exhibit 1020.7687

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2006/007485

A. CLASSIFICATION OF SUBJECT MATTER
INV. C12Q1/68

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)
C12Q

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, BIOSIS, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 01/57248 A (SOLEXA LTD) 9 August 2001 (2001-08-09) page 6 - page 7	1-5,7, 12-14, 16,17, 21,52-54
X	KABOEV O K ET AL: "PCR hot start using primers with the structure of molecular beacons (hairpin-like structure)" NUCLEIC ACIDS RESEARCH, OXFORD UNIVERSITY PRESS, SURREY, GB, vol. 28, no. 21, 1 November 2000 (2000-11-01), page E94, XP002244959 ISSN: 0305-1048 cited in the application the whole document	1-3, 52-54

-/--

[X] Further documents are listed in the continuation of Box C. [X] See patent family 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
22 September 2006	12/10/2006

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	Aslund, Fredrik

Form PCT/ISA/210 (second sheet) (April 2005)

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2006/007485

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 6 692 915 B1 (NALLUR GIRISH N) 17 February 2004 (2004-02-17) page 23, line 10 - line 14	1-3
A	US 6 261 774 B1 (PAGRATIS NIKOS ET AL) 17 July 2001 (2001-07-17) page 18, line 44 - line 49; figures 41A,B	1,4-11, 14,15,20
A	LIU CHIH LONG ET AL: "Development and validation of a T7 based linear amplification for genomic DNA." BMC GENOMICS, vol. 4, no. 19 Cited June 13, 2003, 9 May 2003 (2003-05-09), XP002394174 ISSN: 1471-2164	15,34
A	KAUR MANJIT ET AL: "Novel amplification of DNA in a hairpin structure: towards a radical elimination of PCR errors from amplified DNA." NUCLEIC ACIDS RESEARCH. 15 MAR 2003, vol. 31, no. 6, 15 March 2003 (2003-03-15), page e26, XP002394810 ISSN: 1362-4962 figures 1,3	
A	WO 98/39485 A (THE REGENTS OF THE UNIVERSITY OF MICHIGAN; MAKAROV, VLADIMIR L; LANGMO) 11 September 1998 (1998-09-11) page 49 - page 50	
A	PIRKER C ET AL: "WHOLE GENOMIC AMPLIFICATION FOR CGH ANALYSIS: LINKER-ADAPTER PCR AS THE METHOD OF CHOICE FOR DIFFICULT AND LIMITED SAMPLES" CYTOMETRY, ALAN LISS, NEW YORK, US, vol. 61A, no. 1, September 2004 (2004-09), pages 26-34, XP008068045 ISSN: 0196-4763	

IDT Exhibit 1020.7689

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2006/007485

Box No. I Nucleotide and/or amino acid sequence(s) (Continuation of item 1.b of the first sheet)

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application and necessary to the claimed invention, the international search was carried out on the basis of:

   a. type of material

   [X] a sequence listing

   [X] table(s) related to the sequence listing

   b. format of material

   [X] on paper

   [X] in electronic form

   c. time of filing/furnishing

   [X] contained in the international application as filed

   [ ] filed together with the international application in electronic form

   [X] furnished subsequently to this Authority for the purpose of search

2. [X] In addition, in the case that more than one version or copy of a sequence listing and/or table relating thereto has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that in the application as filed or does not go beyond the application as filed, as appropriate, were furnished.

3. Additional comments:

IDT Exhibit 1020.7690

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/US2006/007485

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 0157248 A	09-08-2001	AU 3038401 A	14-08-2001
		CA 2398168 A1	09-08-2001
		EP 1252332 A2	30-10-2002
		IS 6489 A	31-07-2002
		JP 2003521693 T	15-07-2003
US 6692915 B1	17-02-2004	NONE	
US 6261774 B1	17-07-2001	US 5637459 A	10-06-1997
		US 5723289 A	03-03-1998
		US 5696249 A	09-12-1997
		US 5670637 A	23-09-1997
		US 5843653 A	01-12-1998
		US 5496938 A	05-03-1996
		US 5476766 A	19-12-1995
WO 9839485 A	11-09-1998	AU 6544598 A	22-09-1998
		US 6537757 B1	25-03-2003
		US 6117634 A	12-09-2000

IDT Exhibit 1020.7691

Espacenet

Bibliographic data: EP2322612 (A1) — 2011-05-18

METHOD FOR PRODUCTION OF cDNA LIBRARY HAVING REDUCED CONTENT OF cDNA CLONE DERIVED FROM HIGHLY EXPRESSED GENE

Inventor(s): OHTOKO KUNIYO [JP]; SUGIYAMA MASAHIDE [JP]; KATO SEISHI [JP] ± (OHTOKO, KUNIYO, ; SUGIYAMA MASAHIDE, ; KATO SEISHI)

Applicant(s): HITACHI HIGH TECH CORP [JP]; JP NAT REHABILITATION CT FOR PERSONS WITH DISABILITIES [JP] ± (HITACHI HIGH-TECHNOLOGIES CORPORATION, ; JAPAN AS REPRESENTED BY DIRECTOR GENERAL OF NATIONAL REHABILITATION CENTER FOR PERSONS WITH DISABILITIES)

Classification: - international: *C12N15/10; C40B40/06; C40B50/06*
- cooperative: **C12N15/1093**; **C12N15/1096**; **C40B40/06**; **C40B50/06** more

Application number: EP20090809882 20090825 Global Dossier

Priority number(s): WO2009JP64755 20090825 ; JP20080217245 20080826

Also published as: EP2322612 (A4) EP2322612 (B1) CN102124106 (A) CN102124106 (B) JP2010051195 (A) more

Abstract of EP2322612 (A1)

A method for efficiently constructing a cDNA library having a reduced content of cDNA clones derived from a highly expressed gene is provided. According to the method for constructing a cDNA library using a double-stranded DNA primer having oligo dT and mRNA as a template, the proportion of cDNA clones derived from a highly expressed gene in a cDNA library was decreased through coexistence with a probe having a property of binding to the mRNA of the highly expressed target gene so as to inhibit a cDNA extension reaction resulting from reverse transcriptase.

IDT Exhibit 1020.7692

(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 2 322 612 A1

(12) **EUROPEAN PATENT APPLICATION**
published in accordance with Art. 153(4) EPC

(43) Date of publication:
18.05.2011 Bulletin 2011/20

(21) Application number: **09809882.5**

(22) Date of filing: **25.08.2009**

(51) Int Cl.:
C12N 15/09 *(2006.01)* ***C40B 40/06*** *(2006.01)*

(86) International application number:
PCT/JP2009/064755

(87) International publication number:
WO 2010/024231 (04.03.2010 Gazette 2010/09)

(84) Designated Contracting States:
AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO PL PT RO SE SI SK SM TR
Designated Extension States:
AL BA RS

(30) Priority: **26.08.2008 JP 2008217245**

(71) Applicants:
- **Hitachi High-Technologies Corporation**
  Tokyo 105-8717 (JP)
- **Japan as represented by Director General of National Rehabilitation Center for Persons with Disabilities**
  Tokorozawa-shi
  Saitama 359-8555 (JP)

(72) Inventors:
- **OHTOKO, Kuniyo**
  Hitachinaka-shi,
  Ibaraki 312-0033 (JP)
- **SUGIYAMA Masahide**
  Tokyo 105-8717 (JP)
- **KATO Seishi**
  Tokorozawa-shi,
  Saitama 359-8555 (JP)

(74) Representative: **Strehl Schübel-Hopf & Partner**
Maximilianstrasse 54
80538 München (DE)

(54) METHOD FOR PRODUCTION OF cDNA LIBRARY HAVING REDUCED CONTENT OF cDNA CLONE DERIVED FROM HIGHLY EXPRESSED GENE

(57) A method for efficiently constructing a cDNA library having a reduced content of cDNA clones derived from a highly expressed gene is provided. According to the method for constructing a cDNA library using a double-stranded DNA primer having oligo dT and mRNA as a template, the proportion of cDNA clones derived from a highly expressed gene in a cDNA library was decreased through coexistence with a probe having a property of binding to the mRNA of the highly expressed target gene so as to inhibit a cDNA extension reaction resulting from reverse transcriptase.

EP 2 322 612 A1

Printed by Jouve, 75001 PARIS (FR)

IDT Exhibit 1020.7693

Description

TECHNICAL FIELD

[0001] The present invention relates to a method for constructing a cDNA library having a reduced content of cDNA clones derived from a gene that is expressed at high frequency.

BACKGROUND ART

[0002] The full-length sequences of genomic DNAs of various organisms such as humans, mice, rice, nematodes, and yeast have been almost completely determined by genome projects. Now in the post-genome era, research is conducted to examine the kinetics of gene groups in order to macroscopically understand the role of each gene in biological phenomena. Thus, technology for comprehensive analysis of all genes is required.

[0003] It is considered that the cells composing a human body have a genome in which about 22,000 types of gene are encoded and several tens of thousands of types of genes are expressed depending on cell type (see International Human Genome Sequencing Consortium, 2004, Nature 432, 931-945). Comprehensive examination of the expression of such genes becomes increasingly important not only for elucidation of biological phenomena, but also for developing pharmaceutical products based on gene information. In particular, comprehensive and detailed analysis of information concerning the protein primary structures encoded by genes requires that DNAs complementary to mRNAs, that is, cDNAs (complementary DNAs), be obtained.

[0004] Genes are roughly classified into 3 classes based on expression levels: a group of highly expressed genes expressed at 12,000 or more copies per cell, a group of moderately expressed genes expressed at about 300 copies per cell, and a group of lowly expressed genes expressed at only about 15 copies per cell (see David J. Bertioli., et al., Nucleic Acids Research, 1995, Vol .23, No. 21, 4520-4523). When a cDNA library is constructed using a mixture of mRNAs with different expression levels as a template, the content of the lowly expressed gene clones contained in the cDNA library is lower than that of the highly expressed gene by three or more orders of magnitude. Accordingly, comprehensive analysis of cDNA clones contained in a cDNA library by analyzing partial nucleotide sequences requires unnecessary procedures such as redundant sequencing of clones derived from the highly expressed genes. Many intracellularly expressed genes are expressed at low levels. Hence, methods for reducing the proportion of clones derived from a highly expressed gene have been developed.

[0005] One such method uses self-association of the cDNAs derived from highly expressed genes. First, a double-stranded cDNA is prepared from mRNA and then denatured to single strands. Subsequently, an annealing reaction is performed and then single-stranded cDNAs are caused to adsorb hydroxyapatite. They are then collected from a mixture of double-stranded cDNA obtained by re-association and the cDNAs that have remained as single strands. This procedure is repeated several times, and single-stranded cDNAs are collected and then cloned. Specifically, many single-stranded cDNAs are present in the case of highly expressed genes, so that they highly likely form double stranded cDNAs via annealing reaction. However, the content of single-stranded cDNAs among lowly expressed genes is low and the association frequency is low, and thus these single-stranded cDNAs remain as single strands. This method involves collecting single-stranded cDNAs derived from lowly expressed genes that have remained unassociated (see JP Patent Publication (Kokai) No. 4-108385 A (1992)).

[0006] Another method is based on hybridization with mRNA. Specifically, mRNA is transcribed from the cDNA of an existing highly expressed gene, biotinylated, and then caused to hybridize to a subject single-stranded cDNA library. cDNA not binding to the magnetic particles is recovered, and cDNA that has reacted with avidin-immobilized magnetic particles so as to bind to the biotinylated mRNA is excluded. Then, the unbound cDNA is cloned and thus a clone of the lowly expressed gene is obtained (see JP Patent Publication (Kokai) No. 2000-325080 A).

[0007] A cDNA library having an increased content of lowly expressed genes can be constructed by these methods. However, these methods require an additional step for removing the cDNA of a highly expressed gene after construction of a cDNA library from mRNA, resulting in an increased number of steps compared with those in a general cDNA library construction method. Also, the cDNA of a lowly expressed gene is nonspecifically lost, causing the problem of a reduced yield thereof.

[0008] In contrast, a method is proposed by which synthesis of 1st strand cDNA of a highly expressed gene is suppressed in a step of synthesizing 1st strand cDNA using mRNA as a template. Specifically, in the step of synthesizing 1 st strand cDNA from mRNA using oligo dT primers and reverse transcriptase, a probe specifically binding to a position near the 3' end of the mRNA of a highly expressed gene is caused to also coexist to stop 1 st strand cDNA synthesis and then a probe-bound portion of the mRNA is degraded by RNaseH treatment (see JP Patent Publication (Kokai) No. 2000-37193 A). However, reverse transcriptase has activity of eliminating a probe that has bound upon transcription. Hence, it is thought that when a probe binding to an arbitrarily site of mRNA is used, 1 st strand cDNA synthesis cannot be inhibited. In the above document, a probe binding to a position near the 3' end of mRNA is used.

IDT Exhibit 1020.7694

[0009] The use of any such method requires a step of annealing single-stranded cDNAs to form double-stranded cDNA and then integrating the resultant into an appropriate vector. Loss of the cDNA clone of a lowly expressed gene occurs during this process. Hence, it has been difficult to efficiently obtain full-length cDNA clones of a lowly expressed gene. Therefore, a method for efficiently constructing a cDNA library having a low content of a highly expressed gene with fewer steps has been desired.

DISCLOSURE OF THE INVENTION

Object to be Achieved by the Invention

[0010] An object of the present invention is to provide a method for efficiently constructing a cDNA library having a reduced content of cDNA derived from a single or a plurality of highly expressed genes in mRNA derived from cells or the like.

Means for Achieving the Object

[0011] The present inventors have discovered that, when cDNA synthesis is performed using a double-stranded DNA primer as described in International Patent Publication WO2004/087916 (Pamphlet), the content of cDNA clones of a target gene can be lowered by causing a probe (for a highly expressed target gene in a nucleic acid sample) to coexist upon an extension reaction of 1 st strand cDNA, so as to inhibit the extension reaction of 1 st strand cDNA and to inhibit a ring-closing reaction of the conjugate of an mRNA/cDNA heteroduplex derived from the gene and a double-stranded DNA primer. Thus, the present inventors have completed the present invention. Specifically, the present invention provides a step for constructing a cDNA library having a reduced content of cDNA derived from the mRNA of a target gene that is a highly expressed gene.

[0012] The present invention relates to a method for constructing a cDNA library having a reduced content of cDNA clones derived from a target gene, for example, comprising the steps of:

(i) annealing a double-stranded DNA primer to an RNA mixture containing mRNA having a cap structure at the 5' end and then further annealing at least one probe that binds to mRNA of the target gene so as to inhibit a reaction with reverse transcriptase;
(ii) synthesizing 1 st strand cDNA from the double-stranded DNA primer using reverse transcriptase and thus preparing a conjugate of a mRNA/cDNA heteroduplex and the double-stranded DNA primer; and
(iii) ligating the 3' end to the 5' end of the DNA strand containing the cDNA of the conjugate of the mRNA/cDNA heteroduplex and the double-stranded DNA primer using ligase for circularization.

A highly expressed gene can be used as a target gene and a target gene can be selected based on a gene database.

[0013] The mRNA having the cap structure in step (i) is contained in a cell extract, for example. The primer sequence of a double-stranded DNA primer to be used herein contains a sequence complementary to the poly (A) sequence of mRNA having a cap structure, for example. Also, as a ligase, T4RNA ligase can be used, for example.

[0014] The above method may further comprise, between step (ii) and step (iii), step (ii') of cleaving the conjugate of the mRNA/cDNA heteroduplex and the double-stranded DNA primer with a restriction enzyme, so as to generate a 5' protruding end or blunt end of the double-stranded DNA primer.

[0015] Also, the above method may further comprise, following step (iii), step (iv) of substituting the RNA strand of the conjugate of the mRNA/cDNA heteroduplex and the double-stranded DNA primer with a DNA strand.

[0016] Also, in the above method, the double-stranded DNA primer may contain a replication origin, or a replication origin and a cDNA expression promoter. The probe that binds to the mRNA of a target gene so as to inhibit a reaction with reverse transcriptase is, for example, an oligonucleotide containing non-natural nucleic acids with a melting temperature (Tm value) higher by 2°C or more than that of a complementary sequence of a partial sequence of the target gene, in which 2 or more consecutive nucleotides from the 5' end side are composed of non-natural nucleic acids. As non-natural nucleic acids, locked nucleic acids, polyamide nucleic acids, or bridged nucleic acids (BNAs) can be used. Probes that bind to the mRNAs of target genes so as to inhibit a reaction with reverse transcriptase may target the mRNAs of different types of target gene.

[0017] Furthermore, the present invention provides, for example, a probe having a sequence complementary to an mRNA sequence. Specifically, the probe comprises an oligonucleotide that is designed so that the 3' end has a structure modified so that it does not serve as the initiation point for an extension reaction with reverse transcriptase, 2 or more consecutive nucleotides from the 5' end are non-natural nucleic acids, and the Tm value of the nucleotide sequence of the entire probe is higher by 2°C or more than that of a nucleotide sequence composed of only natural nucleic acids.

[0018] An example of the probe is used for constructing a cDNA library having a reduced content of cDNA clones

IDT Exhibit 1020.7695

derived from a target gene and binds to the mRNA of the target gene so as to inhibit a reaction with reverse transcriptase.

[0019] Furthermore, the present invention provides, for example, a reagent kit for producing a cDNA library containing the above probe, a double-stranded DNA primer, reverse transcriptase, and T4 RNA ligase.

[0020] In addition, in the present invention, the term "double-stranded DNA primer" refers to a primer in which the 3' end of one DNA strand of the double-stranded DNA protrudes and the nucleotide sequence (primer sequence) of the protruding portion has a nucleotide sequence complementary to the template mRNA sequence. This protruding portion hybridizes to template mRNA and functions as a primer when 1 st strand cDNA is synthesized with reverse transcriptase. For cDNA synthesis, a double-stranded DNA primer wherein the sequence of such a protruding portion comprises 30-70 oligo dTs is used.

Effects of the Invention

[0021] According to the present invention, a cDNA library having a reduced content of cDNA derived from the mRNA of a target gene can be provided using fewer steps than conventional methods. When a highly expressed gene is used as a target gene, the content of a lowly expressed gene becomes relatively high among the prepared cDNA clones. Thus, the complete gene information of lowly expressed genes can be efficiently obtained. Moreover, the present invention enables the cloning of the cDNA of a gene that is expressed in trace amounts and thus is cloned with difficulty by a conventional method. These genes that are expressed at low expression levels can be used as target genes for gene diagnosis for diseases or development of pharmaceutical products.

BRIEF DESCRIPTION OF THE DRAWINGS

[0022]

Fig. 1 shows basic procedures of the method for cDNA synthesis using a method for inhibiting the extension of 1 st strand cDNA according to the present invention.
Fig. 2 shows the contents of genes existing in cDNA libraries constructed from RNA samples used in the examples.
Fig. 3 shows the sequences of probes used for examination.
Fig. 4 is a table showing the contents of Albumin 1 and Major urinary protein 2 in each library, as determined by hybridization.
Fig. 5 is a table showing the contents of Albumin 1 and Major urinary protein 2 in each library, as determined by sequencing of the 5'-terminal nucleotide sequence.

BEST MODE OF CARRYING OUT THE INVENTION

[0023] This description includes part or all of the contents as disclosed in the description and/or drawings of Japanese Patent Application No. 2008-217245, which is a priority document of the present application.

[0024] Regarding the cDNA library having a reduced content of target gene-derived cDNA clones of the present invention, the cDNA library is constructed by a cDNA synthesis method using a double-stranded DNA primer as described in International Patent Publication WO2004/087916 (Pamphlet) resulting in a reduced content of a specific target gene-derived cDNA. The cDNA synthesis method is a method for synthesizing cDNA having a continuous sequence from a nucleotide adjacent to the cap structure of mRNA, for which total RNA of several micrograms can be used as a starting material. Also, cDNA can be synthesized with fewer steps without using PCR. The cDNA synthesis method is further characterized in that a full-length cDNA having the nucleotide sequence corresponding to the full-length sequence ranging from the 5'-terminal cap structure addition site to the 3' terminal poly (A) tail of a complete mRNA guaranteed to have a continuous sequence from the nucleotide at the transcription initiation point can be synthesized with a high yield of 90% or more.

[0025] As mRNA samples, total RNA isolated from eukaryotic cells can be used. Also, mRNA having a polyA structure, which has been purified from the total RNA using an oligo dT-bound carrier may be used. Samples containing these mRNAs desirably merely undergo degradation. Specifically, such RNA sample desirably contains complete mRNA having a cap structure at the 5' end and polyA at the 3' end. The term "RNA mixture containing mRNA having a cap structure" in the present invention may refer to an RNA mixture substantially comprising only mRNA having a cap structure and the RNA mixture may contain mRNA and the like lacking a cap structure in addition to mRNA having a cap structure, for example.

[0026] Any gene can be selected as a target gene for reduction of the content of cDNA clones thereof in a cDNA library. In view of construction of cDNA libraries containing many lowly expressed genes, which has been performed with difficulty using conventional methods, highly expressed genes are preferred. Here, the term "highly expressed gene" refers to a gene the mRNA of which is contained at a high level in a mRNA sample because of its high expression level.

IDT Exhibit 1020.7696

The term "highly expressed gene" also refers to a gene that exists at about 12,000 copies or more per cell, for example. A gene that exists at about 300 copies or more per cell is also included herein. Many highly expressed genes are known as namely housekeeping genes or householding genes. Housekeeping genes are constantly expressed in all cells and are mainly composed of genes encoding proteins required for the survival of cells (e.g., structural proteins, enzymes of energy metabolic systems, and genes involved in mechanisms for protein synthesis, DNA replication, cell division, and the like).

[0027] Specifically, genes of such as glyceraldehyde-3-phosphate dehydrogenase (GAPDH), β-actin, a ribosomal protein, histone, phospholipase, a transferrin receptor, HPRT1, and RNA synthetase are known. These housekeeping genes can be selected as target genes. When a specific cell is used, the number or the type of a highly expressed gene differs depending on cells since a specific gene is specifically expressed at a high level in the relevant cell. Examples of such a gene that is specifically expressed at a high level in such a specific cell include albumin in hepatic cells and genes of major urinary proteins. These highly expressed genes can be selected with the use of public databases of existing genes. For example, through the use of the Bodymap (http://bodymap.ims,u-tokyo.ac.jp) database for genes expressed in each human or mouse tissue, information of highly expressed genes can be obtained for every tissue. Also, if no information exists in the existing databases, a gene expression level in an mRNA sample can be found by a method for obtaining the information of gene expression, such as an EST analysis method, a SAGE method, and a DNA array method. Hence, the gene expression level in a mRNA sample used herein can be found. A list is prepared by numbering genes based on their expression levels and then highly expressed gene candidates to be eliminated are selected. A plurality (e.g., 2 or more types) of target genes are applied in most cases. For example, based on information based on the above database, 2 to 10, 2 to 50, 2 to 100, 2 to several hundred, 2 to 1000, 2 to several thousand, or an even higher number of target genes are selected, in descending order of expression levels.

[0028] A cDNA library having a reduced content of a highly expressed gene can be produced by the method of the present invention. Specifically, a cDNA library having an improved content of a gene, number of copies of which is about 12,000 per cell or less, or is preferably about 300 per cell or less, can be produced.

[0029] A probe that binds to the mRNA of a target gene has a property of blocking the extension of 1 st strand cDNA to be synthesized by reverse transcriptase in the 5' end direction on the mRNA strand.

[0030] The probe has the following characteristics.

(1) Basic structure

[0031] The basic structure of the probe is an oligonucleotide having a sequence complementary to a partial sequence of the mRNA sequence of a target gene, wherein the 3' end is modified so that it does not serve as an initiation point for an extension reaction with reverse transcriptase. For example, the probe has a structure of dideoxyribonucleic acid (ddNTP) lacking a hydroxyl group or has a structure in which a hydroxyl group of a nucleotide at the 3' end is bound to a compound such as biotin.

(2) Binding position

[0032] The nucleotide sequence of mRNA of a target gene to which a probe complementarily binds is designed based on nucleotide sequence information according to actual data or public database of subject gene sequences. At this time, a binding site for a probe is preferably located in a common region among variants since it is known that a plurality of transcript variants are generated from a single gene locus due to differences in promoter, splicing, poly (A) addition site, and the like.

(3) Nucleotide length of nucleotide sequence of complementary strand

[0033] The nucleotide length of a nucleotide sequence of a complementary strand is not particularly limited, as long as it is sufficient for stable binding to the mRNA of a target gene with a length of 20 to 50 nucleotides, preferably of 20 to 40 nucleotides, and further preferably of 20 to 35 nucleotides.

(4) Introduction of non-natural nucleic acid

[0034] The term "non-natural nucleic acid" refers to a nucleic acid having a structure lacking natural nucleic acids, but having nucleotides with structures that are absent in the nature. In the present invention, introduction of one or a plurality of nucleotides that are absent in natural nucleic acids into a probe refers to introduction of a non-natural nucleic acid(s). The probe of the present invention is an oligonucleotide containing non-natural nucleic acids.

[0035] When non-natural nucleic acids are introduced into a probe, Tm value for the probe is preferably determined to be higher by 2°C or more than the Tm value for a nucleotide sequence that is complementary to a partial sequence

IDT Exhibit 1020.7697

of the target sequence and is composed of natural nucleic acids. Also, even higher Tm value is more preferable. Also, Tm value for a probe must be higher than the temperature at which 1 st strand cDNA is synthesized in step (ii). The Tm value for a nucleotide sequence composed of natural nucleic acids can be calculated by a known method. The probe is designed in such a manner so that after the probe has once annealed to the mRNA of a target gene, the probe is not easily separated, resulting in a greater effect of suppressing the extension reaction. Any nucleic acid may be used as a non-natural nucleic acid, as long as it complementarily binds to mRNA. Polyamide nucleic acids or locked nucleic acids, bridged nucleic acids (BNAs), or the like are preferably used, for example. Here, the term "polyamide nucleic acid" refers to a DNA analog containing neutral-amide main-chain bonding (e.g., Nielsen et al., Science 254: 1497, 1991). The term "locked nucleic acid" refers to a bicyclic nucleic acid analog in which an O-methylene (oxy-LNA), an S-methylene (thio-LNA), or an NH_2-methylene component (amino-LNA) binds to a furanose ring at positions 2' and 4'. The term "bridged nucleic acid (BNA)" refers to an artificial synthetic nucleic acid prepared via cross-linking of natural nucleic acid molecules (e.g., JP Patent Publication (Kokai) No. 10-195098 A (1998)). When a locked nucleic acid is used, the Tm value resulting from the mixing of a natural type and a non-natural type can be predicted by referring to Niels Tolstrup's report (Niels Tolstrup et al., Nucleic Acid Research, 2003, 31, 3758-3762), for example. When a non-natural nucleic acid is introduced, preferably 2 or more nucleotides and further preferably 3 or more nucleotides from the 5' end of the probe are non-natural nucleic acids. Also, non-natural nucleic acids to be introduced are preferably continuous. The upper limit of the number of non-natural nucleic acids to be introduced is not limited, but is preferably 10 or less and further preferably 5 or less.

[0036] A probe composed of such natural and non-natural nucleic acids can be chemically synthesized by a conventional method. Such probe to be used herein with a general degree of purification can be used.

[0037] When a plurality of target genes are present, a plurality of probes are used so that they can correspond to all of these target genes. The total number of probes to be used herein is at least the same as that of target gene types. A plurality of probes differing in positions of partial sequences may be used for one target gene.

[0038] Steps of the present invention will be explained according to Fig. 1.

[0039] In step (i), a probe (C) binding to target gene mRNA (A) and a double-stranded DNA primer (D) having oligo dT is annealed to a sample RNA containing the target gene mRNA (A) and non-target gene mRNA (B). Therefore, a target gene mRNA (E) to which the probe (C) and the double-stranded DNA primer (D) are bound is generated, and a non-target gene mRNA (F) to which the double-stranded DNA primer (D) is bound is generated. cDNA can be synthesized with even less than 1 μg of such a sample RNA, but preferably 1 μg or more of a sample RNA is used. A single or 2 or more types of target gene can be used. Hence, a required number of types of the probe (C) having the above characteristics is prepared in accordance with the number of types of target genes. The amount of the probe (C) to be used herein is preferably 50-100 times (in terms of molar ratio) the amount of mRNA (A) or (B), but is not particularly limited thereto. Also, the number of consecutive dTs composing the sequence of the above double-stranded DNA primer (D) preferably ranges from 30 to 70. The double-stranded DNA primer (D) may be such primer containing a replication origin, or a replication origin and a cDNA expression promoter, but the examples thereof are not particularly limited thereto. A restriction enzyme site may be introduced in advance into the double-stranded DNA primer in accordance with the intended use of the cDNA. Annealing temperature may be determined based on the Tm value of the probe (C), so that it is lower than the Tm value of the probe (C).

[0040] Probe design and annealing conditions can be adequately determined depending on the types of target gene.

[0041] In step (ii), reverse transcriptase is caused to act on the target gene mRNA (E) to which the probe (C) and the double-stranded DNA primer (D) are bound, and the non-target gene mRNA (F) to which the double-stranded DNA primer (D) is bound, so that 1 st strand cDNA complementary to the mRNA in the direction from the 3' end to the 5' end is synthesized. Preferable reverse transcriptase lacks endogenous RNaseH activity. In the case of the target gene mRNA (E) to which the probe (C) and the double-stranded DNA primer (D) are bound, 1 st strand cDNA extension stops before the probe, resulting in an incomplete mRNA/DNA heteroduplex (G) that does not reach the 5' end of the mRNA. Meanwhile, in the case of the non-target gene mRNA (F) to which the double-stranded DNA primer (D) is bound, a mRNA/DNA heteroduplex (H) is formed in which a complementary strand DNA is completely synthesized so as to reach the 5' end. Therefore, 1st strand cDNA is synthesized from only the non-target gene mRNA. In step (ii), a mRNA/cDNA heteroduplex-to-double-stranded DNA primer conjugate is formed. In the conjugate, one end of a cDNA strand of the mRNA/cDNA heteroduplex and one end of a single strand of the double-stranded DNA primer are linked. Also, the cDNA of the mRNA/cDNA heteroduplex generated in step (ii) is complementary to mRNA having a cap structure and is a cap-consecutive cDNA containing dC or 5'-dC(dA)n-3' added to the 3' end or a mRNA-derived cap non-consecutive cDNA having no cap structure.

[0042] In step (ii'), other end of the double-stranded DNA primer ligated to the mRNA/DNA heteroduplex is blunt-ended or is preferably treated to have the 5' protruding end through cleavage with a restriction enzyme. In this case, it is required to cause the terminal region of a double-stranded DNA primer to have a restriction enzyme site in advance. In addition, this step can be omitted. Implementation of this step can lead to a significant decrease in the background composed only of a vector contained in a cDNA library.

IDT Exhibit 1020.7698

[0043] In step (iii), the mRNA/DNA heteroduplex-to-double-stranded DNA primer conjugate wherein other end of the double-stranded DNA primer has been blunt-ended or treated to have a 5' protruding end is circularized with ligase. As ligase, various types of DNA ligase or RNA ligase can be used and T4RNA ligase is preferably used. A complete non-target gene mRNA/DNA heteroduplex (H), in which the 1 st strand cDNA extends to the 5' end of mRNA, is circularized with ligase. However, an incomplete target gene mRNA/DNA heteroduplex (G), in which 1st strand cDNA extension is not completed, is not circularized.

[0044] In step (iv), in the thus circularized non-target gene mRNA/DNA heteroduplex (H), the mRNA strand is degraded with RNaseH and the RNA strand is further substituted with a DNA strand with DNA polymerase and DNA ligase. Accordingly, non-target gene cDNA (J) is synthesized.

[0045] If the double-stranded DNA primer (D) contains a replication origin and a drug resistance gene, step (v) is carried out so that transformation into *Escherichia coli* or the like is performed and then the RNA strand may be substituted with a DNA strand in an *in vivo* reaction.

[0046] The thus obtained non-target gene double-stranded cDNA may be cloned into a vector. For example, after cleavage at a restriction enzyme site provided in a double-stranded DNA portion, the resultant may be inserted into a plasmid vector, a phage vector, or the like. At this time, through the use of a vector primer as a double-stranded DNA primer, a step of insertion into another vector can be omitted. A vector primer can be prepared by restrictively cleaving a circular vector DNA at an appropriate cloning site and then ligating a primer sequence complementary to a partial sequence of mRNA to the 3' end, so as to form a 3' protruding end.

[0047] A cDNA library obtained by the above method for constructing a cDNA library has a reduced content of target gene-derived cDNA clones. The content of individual target gene-derived cDNA clones accounts for 2% or less and preferably 1 % or less of the total number of cDNA clones. Also, the percentage by which the number of target gene-derived clones is reduced when a probe has been added to the target gene is 60% or more, preferably 70% or more, and further preferably 80% or more. When a target gene is a highly expressed gene, the above cDNA library mainly contains lowly expressed gene-derived cDNAs. In addition, the cDNA library includes clones containing cap-consecutive cDNA with an extremely high probability of 60% or more, preferably 75% or more, further preferably 90% or more, and most preferably 95% or more.

[0048] A kit for constructing a cDNA library comprises the probe of the present invention to be annealed to a target gene and reagents for synthesis of other cDNAs. Examples of reagents for cDNA synthesis include a double-stranded DNA primer, reverse transcriptase, and T4 RNA ligase.

Examples

[0049] Next, the invention will be described more specifically by referring to examples. The present invention is not limited to the examples. In addition, basic procedures concerning DNA recombination and enzyme reactions were carried out according to the literature (Sambrook and Maniatis, in Molecular Cloning - A Laboratory Manual, Cold spring Harbor Laboratory Press, New York, 1989). Restriction enzymes and various modification enzymes used herein were those produced by Takara Shuzo Co., Ltd., unless otherwise specified. Buffer compositions and reaction conditions for each enzyme reaction were employed according to relevant instructions.

[Example 1]

Suppression of cDNA extension reaction via probe binding

(1) Determination of target gene

[0050] The following experiment was conducted to determine target genes with which experimental effects can be easily confirmed because of their high expression levels. cDNA synthesis was carried out using total RNA (10 μg each) obtained from a male or female mouse liver and 300 ng of pGCAP10 vector primer (International Patent Publication W02004/087916 (Pamphlet)), so that cDNA clones were obtained. The preparation method was carried out according to description in International Patent Publication W02004/087916 (Pamphlet). The 5'-end nucleotide sequences of the thus obtained cDNA clones were analyzed. The thus sequenced 3,217 clones derived from a male mouse liver and the thus sequenced 2,325 clones derived from a female mouse liver were subjected to BLAST search using a GenBank nucleic acid database. As a result, a Major Urinary Protein 2 (MUP2) gene was observed to exhibit the highest degree of overlapping (213 clones (6.6%)) among the cDNA clones obtained from a male mouse liver and the Albumin 1 (A1b1) gene was observed to exhibit the highest degree of overlapping (174 clones (7.5%)) among the cDNA clones obtained from a female mouse liver (Fig. 2). Based on these results, the A1b1 gene and the MUP 2 gene were determined as target genes.

IDT Exhibit 1020.7699

(2) mRNA preparation

[0051] T7 RNA polymerase promoter is located upstream of the cDNA cloning site of the pGCAP10 vector and a *Not* I site is located downstream of the same. Hence, the thus obtained cDNA clones of the A1b1 gene and the MUP2 gene were linearized via *Not* I digestion. With the use of the resultants as templates and T7RNA polymerase, mRNA was prepared using an *in vitro* transcription kit (Ambion).

(3) Probe design

[0052] Probes for stopping a reverse transcription reaction, which is cDNA synthesis reaction, were designed.

[0053] The sequences of translation regions of target gene mRNAs were employed as sequences of the probes. Also, as a non-natural nucleic acid, bridged nucleic acid (BNA) was arranged in the sequence of each probe. There are no examples or detailed information concerning the use of special nucleic acids such as BNA or PNA for stopping reverse transcription or other extension reactions. Examination was carried out by varying the number of BNAs, ranging from 1 to 5. Also, the 3' end was modified with biotin, so as to prevent the probe from functioning as a primer for an extension reaction. The thus designed probe sequences are as shown in Fig. 3. In the sequences in Fig. 3, lower-case letters indicate bridged nucleic acids. The sequences represented by Alb1-0 and Mup-0 are natural nucleic acids. SEQ ID NOS: 1, 2, and 3 in the Sequence Listing represent the sequences of Alb1-0, Mup2-0, and Mup2-1, respectively. Also, we consigned probe synthesis to Gene Design Inc.

(4) Verification of effect of suppressing reaction

[0054] The effect of suppressing the reverse transcription reaction was examined using the probes prepared in (3) above. Probes having "A1b1" in their names were added when A1b1 mRNA was used as a template and probes having "MUP2" in their names were added when MUP2 mRNA was used as a template. cDNA synthesis was carried out for each case. As each reaction solution, a total of 13 μl of a reaction solution was prepared by mixing mRNA (500 ng), poly dT20 primer (100 pmol), dNTP (10 nmol), and a probe (100 pmol) of (3) above. A control reaction solution containing no probe of (3) above was prepared. These reaction solutions were each maintained at 65°C for 5 minutes and then rapidly cooled on ice for 2 minutes. SuperScript II RNase H-reverse transcriptase (Invitrogen Corporation) (200 U), an attached reaction buffer (4 μl), DTT (100 nmol), and ribonuclease inhibitor (40 U) were added so that the volume of each reaction solution was 20 μl. After 1 hour of reaction at 42°C, 1st strand cDNA was synthesized. The reaction solutions were subjected to phenol extraction and then an mRNA/cDNA heteroduplex was collected by ethanol precipitation. The resultant was dissolved in water (10 μl) containing RNase A to degrade mRNA and then the length of cDNA strand in each reaction solution was confirmed by agarose gel electrophoresis. As a result, compared with the control, no stopping of reverse transcription reaction was observed for the A1b1-1 probe, A1b1-2 probe, or Mup 2-1 probe. However, cDNA synthesis was decreased by about 88% in the case of the A1b1-5 probe, by about 66% in the case of the A1b1-4 probe, by about 25% in the case of the A1b1-3 probe, and by about 50% in the cases of the Mup 2-3 probe and the Mup2-2 probe. Regarding conditions of the arrangement and number of BNAs contained in each probe required for suppression of a reverse transcription reaction, it was suggested by the experimental results that 3 or more consecutive nucleotides are effective, in the case of GC-rich sequences. Also, the effect of the MUP2 probe containing AT-rich BNA is weaker than that of the A1b1 probe. Hence, it was considered that conditions under which similar effects could be obtained with any sequence should be examined.

(5) Optimization of effect of suppressing reaction

[0055] Conditions for suppression of cDNA synthesis by A1b1 mRNA were examined using the A1b1-5 probe. Examination was carried out with regard to the amounts of probes to be used, annealing temperature, and annealing time. Reaction was carried out in the above mentioned reaction system by varying only the amount of the probe to be used. Also, regarding annealing temperature, an attached reaction buffer (4 μl), DTT (100 nmol), and ribonuclease inhibitor (40 U) were added so that the volume of the reaction solution was 19 μl, and then the reaction solution was maintained at 42°C or 57°C for 30 minutes to 2 hours. After rapid cooling, SuperScript II RNase H-reverse transcriptase (Invitrogen Corporation) (200 U) was added and then reaction was carried out at 42°C for 1 hour, so that cDNA was synthesized. The detection method was carried out in a manner similar to that in (4) above. As a result, the effect of suppressing reaction by 95% or more was the highest and exhibited under the following conditions: the amount of the probe used was 200 pmol, the annealing temperature was 57°C, and the annealing time was 2 hours. However, the following 3 conditions were determined as optimum conditions under which the effect of suppressing reaction by 90% or more could be obtained in view of the stability of RNA and the amounts of unbound probes, for example: (i) the amount of a probe to be used ranges from 50 to 100 pmol, (ii) the annealing temperature is 57°C, and (iii) the annealing time is 30 minutes.

IDT Exhibit 1020.7700

[0056] Next, the effect of suppressing reaction was examined under the above optimum conditions using the Mup2-3 probe or the Mup 2-2 probe with which an effect equivalent to that in the case of the A1b1-5 probe had not been obtained. The effect of suppressing reaction by 90% or more was exerted in the case of the Mup2-3 probe. However, in the case of the Mup2-2 probe, the effect of suppressing reaction by about 50% was exerted, which was similar to that in the previous case. These results suggested that regarding the effect of suppressing reaction under the optimum conditions employed in this experiment, suppression by 90% or more was observed regardless of sequence as long as a probe containing 5 consecutive BNAs had been used. Also, it was suggested that: a probe that is preferably used for the effect of suppressing reaction has a Tm value (Δ Tm) higher by 2°C (or more) than that of a probe prepared using natural nucleic acids and contains 2 or more consecutive non-natural nucleic acids from the 5' end; and the higher the Δ Tm, the higher the effect of suppressing reaction.

[Example 2]

Construction of cDNA library having reduced content of highly expressed gene

(1) 1 st strand cDNA synthesis

[0057] cDNA synthesis was carried out from total RNA using the above probes. The probes used herein were Albl-5 and Mup2-3. Total RNA (10 μg each) obtained from the liver of a male mouse, a pGCAP10 vector primer (JP Patent Publication (Kokai) No. 4-108385 A (1992)) (150 ng), dNTP (25 nmol), and 100 pmol probe were mixed, so that a total of 8 μl of each mixture was prepared. The mixed solution was heated at 65°C for 5 minutes and then ice-cooled for 2 minutes. An attached buffer (4 μl) of SuperScript II RNase H-reverse transcriptase (Invitrogen Corporation) was added and then the solution was maintained at 57°C for 30 minutes. Subsequently, the solution was ice-cooled for 2minutes, mixed with DTT (100 nmol), ribonuclease inhibitor (40 U), and SuperScript II RNase H-reverse transcriptase (Invitrogen Corporation) (200 U), so as to adjust the solution to a total volume of 20 μl. The mixed solution was subjected to 1 hour of reaction at 42°C, so that 1st strand cDNA was synthesized. After phenol extraction for the reaction solution, the conjugate of mRNA/cDNA heteroduplex and the vector primer was collected by ethanol precipitation and then dissolved in 80 μl of water.

(2) Self ligation

[0058] 10 x H buffer (TAKARA SHUZO Co., Ltd.) and 50 U of restriction enzyme *EcoR* I (TAKARA SHUZO Co., Ltd.) were added to 80 μl of the solution of the conjugate of a mRNA/cDNA heteroduplex and a vector primer to a total volume of 100 μl, followed by 1 hour of reaction at 37°C. The solution was subjected to phenol extraction, the conjugate of the mRNA/cDNA heteroduplex and the vector primer treated with a restriction enzyme was collected by ethanol precipitation, and then the resultant was dissolved in 24 μl of water. The solution was mixed with a reaction solution (50 mM Tris-HC 1 (pH 7.5), 5 mM $MgCl_2$, 10 mM 2-mercaptoethanol, 0.5 mM ATP, and 2 mM DTT), 120 U of T4 RNA ligase (TAKARA SHUZO Co., Ltd.) was added, followed by 16 hours of reaction at 20°C. The mRNA/cDNA heteroduplex was ligated to *Eco*R I end of the vector primer for circularization (self ligation reaction). The reaction solution was subjected to phenol extraction, a self ligation product was collected by ethanol precipitation, and then the self ligation product was dissolved in 50 μg of water, so that a cDNA vector solution was prepared.

(3) Transformation of *Escherichia coli*

[0059] The thus prepared cDNA vector solution (1 μl) was mixed with 20 μl of DH10B Electro-cells (Invitrogen Corporation), so that transformation was carried out by an electroporation method. Electroporation was carried out using MicroPulser (Bio-Rad Laboratories Inc.). The thus obtained transformant was suspended in SOC medium, seeded on LB agar medium containing 50 μg/ml ampicillin, and then cultured at 35°C for 16 to 18 hours, so that cDNA clone-transformed *Escherichia coli* was obtained.

(4) Calculation of the content by hybridization

[0060] The following 4 types of library were constructed.

C: Library (C) to which no probe was added.
Alb: Library (Alb) to which A1b1-5 probe was added.
MUP: Library (MUP) to which Mup2-3 probe was added.
W: Library (W) to which A1b1-5 and Mup2-3 probes were added

IDT Exhibit 1020.7701

[0061] To find the content of the A1b1 gene (GenBank Accession No. NP_033784) and the content of the MUP 2 gene (GenBank Accession No. NP_032673) contained in the constructed cDNA libraries, about 4000 clones were sampled from each library and then subjected to colony hybridization.

[0062] DIG DNA-labeled probe for hybridization was prepared as follows. The A1b1 gene was digested with *Bst* X I and *Not* I restriction enzymes, so that a 1.1-kbp fragment was generated. Also the MUP2 gene was digested with *Avr* II and *Ava* II restriction enzymes, so that a 0.7-kbp fragment was generated. These fragments were subjected to agarose gel electrophoresis, excised, and then increased in amounts by PCR. Subsequently, the DIG DNA-labeled probe was prepared using a DIG High Prime DNA Labeling and Detection Starter Kit I (Roche Applied Science). Next, about 4000 colonies were transferred from each library to a Hybond N + nylon membrane (Amersham). After alkali fixation, hybridization was carried out. Hybridization and detection were carried out according to protocols using DIG High Prime DNA Labeling and Detection Starter Kit I (Roche Applied Science). Hybridization was carried out at 42°C.

[0063] The number and content of clones detected are shown in Fig. 4. The content of A1b1 gene clones decreased from 5.6% to 1.0% in the case of Alb library and also decreased to 1.0% in the case of W library. Therefore, the rate of decrease was 82% in all cases. Also, the content of MUP2 gene clones decreased from 9.9% to 3.1 % in the case of MUP library and decreased to 3.2% in the case of W library. Therefore, the rate of decrease was about 70% in all cases. These results suggest that a decrease by 70% or more can be expected with the addition of the probe. Also, in the case of the W library, the reduction by the same percentage was obtained for each gene even when 2 types of probe had been added.

(5) Calculation of content via analysis of 5'-end nucleotide sequence

[0064] The structural gene region of the nucleotide sequence of the Mup2 gene has 80% or more homology with that of the other Mup genes (Mup1 and Mup3). Therefore, not only the Mup 2 gene, but also the entire Mup group was detected by hybridization. Hence, to find the content more precisely, the nucleotide sequences of about 1000 clones in C and W libraries were analyzed and then numbers of clones, contents, and rates of decreases were found (Fig. 5). As a result, the rate of decrease for A1b1 or for Mup2 was found to be 87%.

[0065] As a result of calculation of contents by hybridization and nucleotide sequence analysis above, it was demonstrated that addition of a probe(s) satisfying requirements that should be satisfied by the probes of the present invention and construction of a cDNA library using a double-stranded DNA primer decrease the content of clones corresponding to the probe(s) by 80% or more and such decreasing effect can be obtained by addition of a plurality of probes.

[0066] All publications, patents, and patent applications cited in this description are herein incorporated by reference in their entirety.

Sequence Listing Free Text

SEQ ID NOS: 1-3 Synthesis

[0067]

IDT Exhibit 1020.7702

SEQUENCE LISTING

<110> HITACHI HIGH-TECHNOLOGIES CORPORATION; Japan as Represented by Director General of National Rehabilitation Center for Persons with Disabilities

<120> A method for producing cDNA library wherein the content of cDNA clones derived from highly expressed genes have been reduced

<130> EPA-46698

<150> JP 2008-217245
<151> 2008-08-26

<160> 3

<170> PatentIn version 3.4

<210> 1
<211> 31
<212> DNA
<213> Artificial

<220>
<223> Synthetic probe

<400> 1
gggatacaga gtaatcaggg tgtcttcttg a 31

<210> 2
<211> 25
<212> DNA
<213> Artificial

<220>
<223> Synthetic probe

<400> 2
taaggattcc atgctcctca catag 25

IDT Exhibit 1020.7703

<210> 3
<211> 31
<212> DNA
<213> Artificial

<220>
<223> Synthetic probe

<400> 3
tttctctaag gattccatgc tcctcacata g 31

Claims

1. A method for constructing a cDNA library having a reduced content of cDNA clones derived from a target gene, comprising the steps of:

   (i) annealing a double-stranded DNA primer to an RNA mixture containing mRNA having a cap structure at the 5' end and then further annealing at least one probe that binds to mRNA of the target gene so as to inhibit the reaction of reverse transcriptase;
   (ii) synthesizing 1st strand cDNA from the double-stranded DNA primer using reverse transcriptase and thus preparing a conjugate of a mRNA/cDNA heteroduplex and the double-stranded DNA primer; and
   (iii) ligating the 3' end to the 5' end of the DNA strand containing the cDNA of the conjugate of the mRNA/cDNA heteroduplex and the double-stranded DNA primer using ligase for circularization.

2. The method according to claim 1, wherein the target gene is a highly expressed gene.

3. The method according to claim 1, comprising a step of selecting the target gene based on a gene database.

4. The method according to claim 1, wherein the mRNA having the cap structure is contained in a cell extract.

5. The method according to claim 1, wherein the primer sequence of a double-stranded DNA primer contains a sequence complementary to the poly (A) sequence of mRNA having a cap structure.

6. The method according to claim 1, wherein the ligase is T4RNA ligase.

7. The method according to claim 1, further comprising, between step (ii) and step (iii), step (ii') of cleaving the conjugate of the mRNA/cDNA heteroduplex and the double-stranded DNA primer with a restriction enzyme, so as to generate a 5' protruding end or blunt end of the double-stranded DNA primer.

8. The method according to claim 1, further comprising, following step (iii), step (iv) of substituting the RNA strand of the conjugate of the mRNA/cDNA heteroduplex and the double-stranded DNA primer with a DNA strand.

9. The method according to claim 1, wherein the double-stranded DNA primer contains a replication origin, or a replication origin and a cDNA expression promoter.

10. The method according to claim 1, wherein the probe that binds to the mRNA of a target gene so as to inhibit a reaction with reverse transcriptase is an oligonucleotide containing non-natural nucleic acids with a melting temperature (Tm value) higher by 2°C or more than that of a complementary sequence of a partial sequence of the target gene.

IDT Exhibit 1020.7704

11. The method according to claim 1, wherein the probe that binds to the mRNA of a target gene so as to inhibit a reaction with reverse transcriptase is an oligonucleotide in which 2 or more consecutive nucleotides from the 5' end side are composed of non-natural nucleic acids.

12. The method according to claim 10, wherein the non-natural nucleic acids are locked nucleic acids, polyamide nucleic acids, or bridged nucleic acids.

13. The method according to claim 1, wherein the probes that bind to the mRNAs of target genes so as to inhibit a reaction with reverse transcriptase target the mRNAs of different types of target gene.

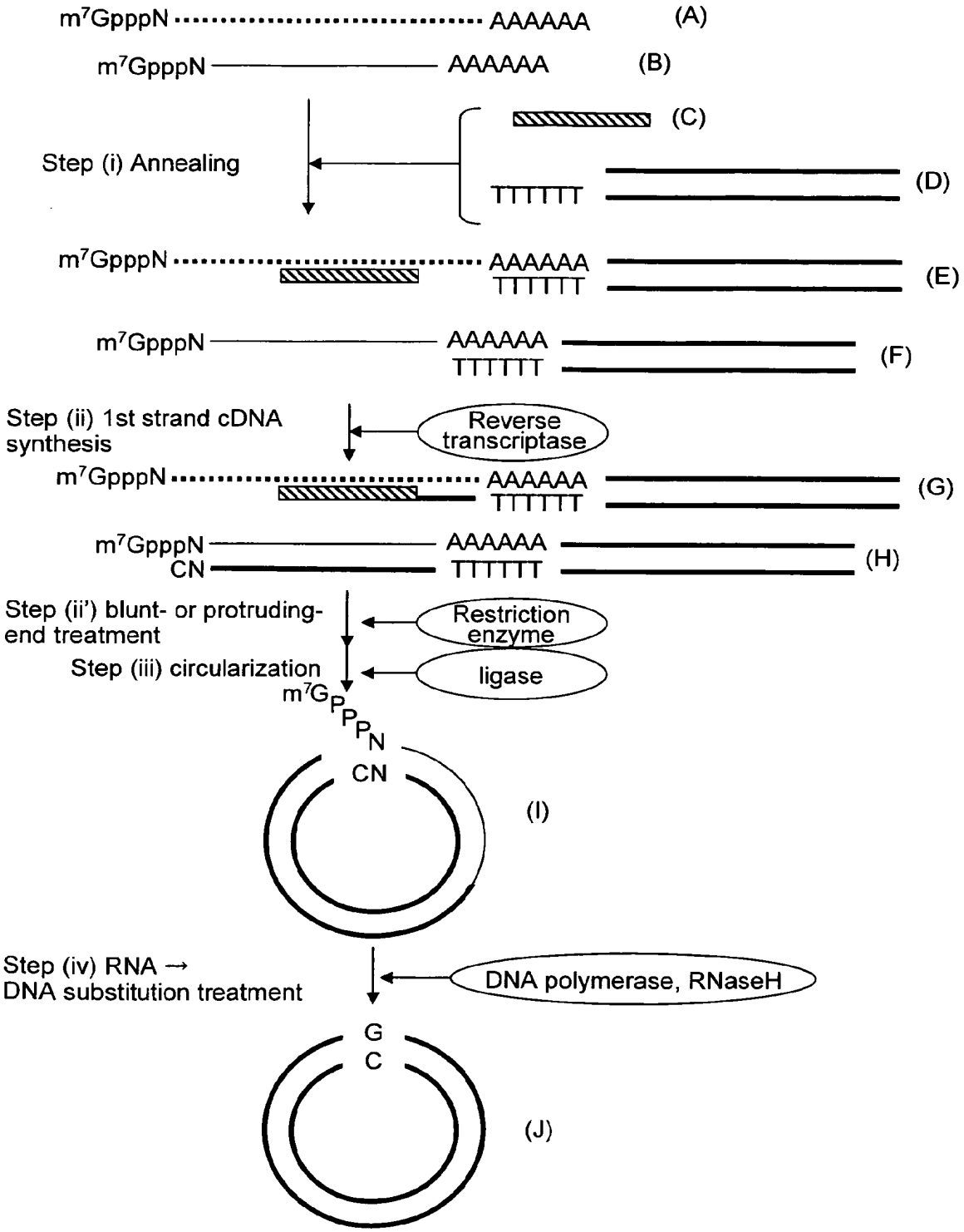
14. A probe having a sequence complementary to an mRNA sequence, which comprises an oligonucleotide that is designed so that the 3' end has a structure modified so that it does not serve as the initiation point for an extension reaction with reverse transcriptase, 2 or more consecutive nucleotides from the 5' end are non-natural nucleic acids, and the Tm value of the nucleotide sequence of the entire probe is higher by 2°C or more than that of a nucleotide sequence composed of only natural nucleic acids.

15. The probe according to claim 14, which is used for constructing a cDNA library having a reduced content of cDNA clones derived from a target gene and binds to the mRNA of the target gene so as to inhibit a reaction with reverse transcriptase.

16. A reagent kit for producing a cDNA library containing the probe according to claim 14, a double-stranded DNA primer, reverse transcriptase, and T4 RNA ligase.

IDT Exhibit 1020.7705

Fig. 1



IDT Exhibit 1020.7706

Fig. 2

cDNA clone content					
	Derived from mouse liver	Female		Male	
		In a total of 2324 clones		In a total of 3216 clones	
Gene name	Accession No.	Number of clones	Content	Number of clones	Content
Albumin1	NP 033784	174	7.5%	148	4.5%
major urinary protein(MUP)2	NP 032673	116	5.0%	213	6.6%
serine (or cysteine) protease inhibitor	NP 033269	72	3.1%	79	2.5%

IDT Exhibit 1020.7707

Fig. 3

Probe name	5' - Sequence - 3'	Modification of 3' end	Number of nucleotides	Tm (°C)	ΔT m
Alb1-0	GGGATACAGAGTAATCAGGGTGTCTTCTTGA	Biotin	31	68	-
Alb1-1	gGGATACAGAGTAATCAGGGTGTCTTCTTGA	Biotin	31	69	1
Alb1-2	GGgATACAGAGTAATCAGGGTGTCTTCTtGA	Biotin	31	70	2
Alb1-3	ggGATACAGAGTAATCAGGGTGTCTTCTTGA	Biotin	31	71	3
Alb1-4	gggATACAGAGTAATCAGGGTGTCTTCTTGA	Biotin	31	72	4
Alb1-5	gggatACAGAGTAATCAGGGTGTCTTCTTGA	Biotin	31	74	6
Mup2-0	TAAGGATTCCATGCTCCTCACATAG	Biotin	25	64	-
Mup2-1	TTTCTCTAAGGATTCCATGCTCCTCACATAG	Biotin	31	66	-
Mup2-2	taaGGATTCCATGCTCCTCACATAG	Biotin	25	66	2
Mup2-3	taaggATTCCATGCTCCTCACATAG	Biotin	25	71	7

Lower-case letter = non-natural nucleic acid

EP 2 322 612 A1

IDT Exhibit 1020.7708

Fig. 4

Library	C	Alb	Rate of decrease (%)	Mup	Rate of decrease (%)	W	Rate of decrease (%)
Number of clones	4000	4000	-	4000	-	4000	-
Albmin 1 (Number of clones)	223	41	-	228	-	41	-
Content (%)	5.6	1.0	82	5.7	-	1.0	82
Major Urinary Protein (Number of clones)	395	437	-	122	-	126	-
Content (%)	9.9	10.9	-	3.1	70	3.2	70

IDT Exhibit 1020.7709

Fig. 5

Library	C	W	Rate of decrease (%)
Number of clones analyzed	1102	1071	-
Albmin 1 (Number of clones)	68	9	-
Content (%)	6.2	0.8	87
Major Urinary Protein 2 (Number of clones)	107	14	-
Content (%)	9.7	1.3	87

IDT Exhibit 1020.7710

INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP2009/064755

A. CLASSIFICATION OF SUBJECT MATTER
C12N15/09(2006.01)i, C40B40/06(2006.01)i

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)
C12N15/09, C40B40/06

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
Jitsuyo Shinan Koho 1922-1996 Jitsuyo Shinan Toroku Koho 1996-2009
Kokai Jitsuyo Shinan Koho 1971-2009 Toroku Jitsuyo Shinan Koho 1994-2009

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
PubMed, WPI, JSTPlus(JDreamII), JMEDPlus(JDreamII), JST7580(JDreamII)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	BENDER, M. et al., "Use of a PNA probe to block DNA-mediated PCR product formation in prokaryotic RT-PCR.", BIOTECHNIQUES, 2007.05, Vol.42, No.5, P.609-614	1-16
A	OLDENBURG, R.P. et al., "Selective amplification of rare mutations using locked nucleic acid oligonucleotides that competitively inhibit primer binding to wild-type DNA.", J. INVEST. DERMATOL., 2008.02, Vol.128, No.2, P.398-402	1-16

☒ Further documents are listed in the continuation of Box C. ☐ See patent family 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 application or patent 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
18 September, 2009 (18.09.09)	06 October, 2009 (06.10.09)
Name and mailing address of the ISA/ Japanese Patent Office Facsimile No.	Authorized officer Telephone No.

Form PCT/ISA/210 (second sheet) (April 2007)

IDT Exhibit 1020.7711

INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP2009/064755

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 2004/087916 A1 (Japan As Represented By Director General Of National Rehabilitation Center For Persons With Disabilities), 14 October, 2004 (14.10.04), Full text & US 2006/0246453 A1 & EP 1616947 A1 & CA 2520811 A	1-16
A	JP 2000-37193 A (Hitachi, Ltd.), 08 February, 2000 (08.02.00), Full text & US 2002/0009731 A1 & US 2002/0102561 A1 & EP 959141 A2	1-16
A	WO 1998/039352 A1 (Takeshi IMANISHI), 11 September, 1998 (11.09.98), Full text & JP 10-304889 A & US 2003/0105309 A1 & EP 1013661 A1	12

Form PCT/ISA/210 (continuation of second sheet) (April 2007)

IDT Exhibit 1020.7712

REFERENCES CITED IN THE DESCRIPTION

This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.

Patent documents cited in the description

- JP 4108385 A **[0005] [0057]**
- JP 2000325080 A **[0006]**
- JP 2000037193 A **[0008]**
- WO 2004087916 A **[0011] [0024] [0050]**
- JP 2008217245 A **[0023]**
- JP 10195098 A **[0035]**

Non-patent literature cited in the description

- International Human Genome Sequencing Consortium. *Nature,* 2004, vol. 432, 931-945 **[0003]**
- **David J. Bertioli. et al.** *Nucleic Acids Research,* 1995, vol. 23 (21), 4520-4523 **[0004]**
- **Nielsen et al.** *Science,* 1991, vol. 254, 1497 **[0035]**
- **Niels Tolstrup et al.** *Nucleic Acid Research,* 2003, vol. 31, 3758-3762 **[0035]**
- **Sambrook ; Maniatis.** Molecular Cloning - A Laboratory Manual. Cold spring Harbor Laboratory Press, 1989 **[0049]**

IDT Exhibit 1020.7713

Espacenet

Bibliographic data: WO02090584 (A2) — 2002-11-14

METHOD FOR LABELLING AND FRAGMENTING DNA

Inventor(s): BOURGET CECILE [FR]; KOTERA MITSUHARU [FR]; LHOMME JEAN [FR]; TREVISIOL EMMANUELLE [FR]; LAAYOUN ALI [FR]; TORA CHRISTELLE [FR]; SOTHIER ISABELLE [FR] ± (BOURGET, CECILE, ; KOTERA, MITSUHARU, ; LHOMME, JEAN, ; TREVISIOL, EMMANUELLE, ; LAAYOUN, ALI, ; TORA, CHRISTELLE, ; SOTHIER, ISABELLE)

Applicant(s): BIO MERIEUX [FR]; UNIV GRENOBLE 1 [FR]; CENTRE NAT RECH SCIENT [FR]; BOURGET CECILE [FR]; KOTERA MITSUHARU [FR]; LHOMME JEAN [FR]; TREVISIOL EMMANUELLE [FR]; LAAYOUN ALI [FR]; TORA CHRISTELLE [FR]; SOTHIER ISABELLE [FR] ± (BIO MERIEUX, ; UNIVERSITE JOSEPH FOURIER (GRENOBLE 1), ; CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE, ; BOURGET, CECILE, ; KOTERA, MITSUHARU, ; LHOMME, JEAN, ; TREVISIOL, EMMANUELLE, ; LAAYOUN, ALI, ; TORA, CHRISTELLE, ; SOTHIER, ISABELLE)

Classification: **- international:** ***C12N15/09; C12Q1/68;*** (IPC1-7): C12Q1/68
- cooperative: **C12Q1/68**

Application number: WO2002FR01542 20020503 Global Dossier

Priority number(s): FR20010006039 20010504

Also published as: WO02090584 (A3) AT340269 (T) AU2002313026 (B2) CA2444926 (A1) CA2444926 (C) more

Abstract of WO02090584 (A2)

The invention concerns a method for labelling and fragmenting a single-stranded or double-stranded deoxyribonucleic acid (DNA) comprising the following steps: chemically fragmenting the DNA by generating an abasic site on said DNA, attaching a marker on at least one of the fragments by means of a labelling reagent, said reagent being covalently and majority bonded on at least one phosphate of said fragment. The invention is applicable in the field of diagnosis.

IDT Exhibit 1020.7714

IDT Exhibit 1020.7715

(12) DEMANDE INTERNATIONALE PUBLIÉE EN VERTU DU TRAITÉ DE COOPÉRATION EN MATIÈRE DE BREVETS (PCT)

(19) Organisation Mondiale de la Propriété Intellectuelle
Bureau international

(43) Date de la publication internationale
14 novembre 2002 (14.11.2002)

PCT

(10) Numéro de publication internationale
WO 02/090584 A2

(51) Classification internationale des brevets[7] : C12Q 1/68

(21) Numéro de la demande internationale :
PCT/FR02/01542

(22) Date de dépôt international : 3 mai 2002 (03.05.2002)

(25) Langue de dépôt : français

(26) Langue de publication : français

(30) Données relatives à la priorité :
01/06039 4 mai 2001 (04.05.2001) FR

(71) Déposants *(pour tous les États désignés sauf US)* : **BIO MERIEUX** [FR/FR]; Chemin de l'Orme, F-69280 MARCY L'ETOILE (FR). **UNIVERSITE JOSEPH FOURIER (GRENOBLE 1)** [FR/FR]; Domaine Universitaire St Martin d'Hères, F-38041 GRENOBLE (FR). **CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE** [FR/FR]; 3 rue Michel Ange, F-75794 PARIS CEDEX 16 (FR).

(72) Inventeurs; et

(75) Inventeurs/Déposants *(pour US seulement)* : **BOURGET, Cécile** [FR/FR]; 3 chemin Villebois, F-38100 GRENOBLE (FR). **KOTERA, Mitsuharu** [FR/FR]; 20 rue des Peupliers, F-38240 MEYLAN (FR). **LHOMME, Jean** [FR/FR]; 13 allée des Brandons, F-38240 MEYLAN (FR). **TREVISIOL, Emmanuelle** [FR/FR]; 16 vieux chemin de Blagnac, F-31700 CORNEBARRIEU (FR). **LAAYOUN, Ali** [FR/FR]; 70 rue Challemel Lacour, F-69007 LYON (FR). **TORA, Christelle** [FR/FR]; La Scierie, F-38780 Oytier Saint Oblas (FR). **SOTHIER, Isabelle** [FR/FR]; 351 chemin du Creuzet, F-69730 GENAY (FR).

(74) Mandataire : **CABINET GERMAIN & MAUREAU**; 12 rue Boileau, F-69006 LYON (FR).

(81) États désignés *(national)* : AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, 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, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

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

Publiée :
— *sans rapport de recherche internationale, sera republiée dès réception de ce rapport*

En ce qui concerne les codes à deux lettres et autres abréviations, se référer aux "Notes explicatives relatives aux codes et abréviations" figurant au début de chaque numéro ordinaire de la Gazette du PCT.

WO 02/090584 A2

(54) **Title:** METHOD FOR LABELLING AND FRAGMENTING DNA

(54) **Titre :** PROCÉDÉ DE MARQUAGE ET DE FRAGMENTATION D'ADN

(57) **Abstract:** The invention concerns a method for labelling and fragmenting a single-stranded or double-stranded deoxyribonucleic acid (DNA) comprising the following steps: chemically fragmenting the DNA by generating an abasic site on said DNA, attaching a marker on at least one of the fragments by means of a labelling reagent, said reagent being covalently and majority bonded on at least one phosphate of said fragment. The invention is applicable in the field of diagnosis.

(57) **Abrégé :** La présente invention concerne un procédé de marquage et de fragmentation d'un acide désoxyribonucléique (ADN) simple ou double brin comprenant les étapes suivantes : fragmenter chimiquement l'ADN en créant au moins un site abasique sur ledit ADN, attacher un marqueur sur au moins un des fragments par l'intermédiaire d'un réactif de marquage, ledit réactif se couplant de manière covalente et majoritaire sur au moins un phosphate dudit fragment. L'invention trouve une application préférentielle dans le domaine du diagnostic.

IDT Exhibit 1020.7716

PROCEDE DE MARQUAGE ET DE FRAGMENTATION D'ADN

La présente invention concerne un procédé de fragmentation et de marquage d'ADN ainsi que les applications de ce procédé en particulier dans le domaine du diagnostic.

L'état de la technique montre que de nombreuses méthodes existent pour marquer des acides nucléiques.

Une première méthode consiste à fixer le marqueur sur la base, que celle-ci soit naturelle ou modifiée. Une deuxième méthode propose de fixer le marqueur sur le sucre, là encore qu'il soit naturel ou modifié. Une troisième méthode a pour objet la fixation du marqueur sur le phosphate.

Le marquage sur la base a été notamment utilisé dans l'approche de marquage des acides nucléiques par incorporation de nucléotides directement marqués.

Le marquage sur le sucre est souvent utilisé dans le cas des oligonucléotides préparées par synthèse chimique.

Le marquage sur le phosphate a été aussi utilisé pour introduire des bras fonctionnalisés et des marqueurs lors de la synthèse chimique des oligonucléotides.

En fait l'homme du métier, qui doit effectuer le marquage d'un nucléotide, ou d'un analogue de nucléotide ou d'un acide nucléique, est enclin à effectuer cette fixation sur la base ou sur le sucre qui lui offre plus de commodité et d'alternatives. C'est d'ailleurs ce qui ressort de l'étude de nombreux documents, tels que EP-A-0.329.198, EP-A-0.302.175, EP-A-0.097.373, EP-A-0.063.879, US-A-5,449,767, US-A-5,328,824, WO-A-93/16O94, DE-A-39 10 151, EP-A-0.567.841 pour la base ou EP-A-0.286.898 pour le sucre.

La fixation du marqueur sur le phosphate est une technique plus complexe que la technique consistant à fonctionnaliser la base ou le sucre et a été bien moins utilisée notamment à cause de la faible réactivité du phosphate (voir par exemple Jencks W.P. et al J. Amer. Chem. Soc., 82, 1778-1785, 1960). De même dans la revue de O'Donnel et Mc Laughlin (« Reporter groups for the analysis of nucleic acid structure », p 216-243, dans « Bioorganic Chemistry : Nucleic Acids », Ed Hecht S.M., Oxford University Press, 1996) portant sur les méthodes d'introduction de sondes dans les fragments d'oligonucléotides,

l'alkylation efficace du phosphodiester internucléotidique est considérée comme étant impossible.

Un deuxième problème se pose pour le marquage d'acides nucléiques spécialement pour les acides nucléiques de grande taille, c'est-à-dire de plus de cent (100) nucléotides, qui doivent s'hybrider avec des sondes nucléiques. Ce problème est lié à la gêne stérique ou au manque de spécificité entre l'acide nucléique et la sonde nucléique. Cela se traduit pas une perte de sensibilité en détection.

La gêne stérique peut être le fait, non seulement, de la longueur de l'acide nucléique, mais également de l'existence ou de la conservation de structures secondaires. La fragmentation permet de détruire ces structures et ainsi d'optimiser l'hybridation. Cette gêne stérique joue un rôle particulièrement important dans le cas de l'hybridation sur des surfaces contenant des sondes de capture à forte densité, par exemple les puces à ADN mises au point par la société Affymetrix ("Accessing Genetic Information with High-Density DNA arrays", M. Shee et al., Science, 274, 610-614. "Light-generated oligonucleotide arrays for rapide DNA sequence analysis", A. Caviani Pease et al., Proc. Natl. Acad. Sci. USA, 1994, 91, 5022-5026).

En ce qui concerne la fragmentation des acides nucléiques, de nombreuses méthodes sont décrites dans l'état de la technique.

Premièrement, la fragmentation peut être enzymatique, c'est-à-dire que la fragmentation des acides nucléiques peut être réalisée par des nucléases (DNases).

Deuxièmement, la fragmentation peut être chimique. Par exemple dans le cas de l'ADN, on peut effectuer la dépurination ou la dépyrimidination de l'ADN, qui est alors fragmenté en présence d'une base par un mécanisme dit de « β-élimination ». La fragmentation de l'ADN peut être réalisé par des mécanismes d'oxydation, d'alkylation, d'addition de radicaux libres entre autres.

Enfin la fragmentation peut être physique, par exemple par sonication ou par voie photochimique.

La difficulté consiste en fait à associer ces deux étapes de fragmentation et de marquage.

La demande de brevet WO-A-99/65926 décrit un procédé de marquage d'un acide ribonucléique (ARN) de synthèse ou naturel qui consiste à fragmenter l'ARN et à marquer au niveau du phosphate terminal. Ce document décrit un certain nombre de fonctions réactives pouvant être utilisées pour le marquage en liaison avec la fragmentation. Ces fonctions permettent de marquer les ARN mais il faut associer une étape de fragmentation pour avoir un marquage efficace car ce marquage se produit sur le phosphate libéré lors de la fragmentation. De plus, il faut ajouter un excès important de réactif de marquage par rapport à l'ARN pour obtenir un marquage efficace, ce qui induit des problèmes de bruit de fond généré par le marqueur en excès. Enfin cette méthode n'est pas applicable à l'ADN double brin.

Il existe donc un besoin pour une technique de marquage de l'ADN simple et efficace qui permette un bon rendement de marquage et donc une bonne sensibilité, qui soit spécifique au niveau de la position de marquage, et en particulier qui n'affecte pas les propriétés d'hybridation des bases impliquées dans la formation de la double hélice par l'intermédiaire des liaisons hydrogènes, et enfin qui permette la fragmentation de l'ADN dans le but d'hybrider ces fragments d'ADN marqués sur des sondes nucléiques, et en particulier sur des sondes nucléiques fixés sur un support solide.

La présente invention décrit un procédé de marquage et de fragmentation d'un acide désoxyribonucléique (ADN) simple ou double brin comprenant les étapes suivantes :
- fragmenter l'ADN en créant au moins un site abasique sur ledit ADN,
- attacher un marqueur sur au moins un des fragments par l'intermédiaire d'un réactif de marquage, ledit réactif se couplant de manière covalente et majoritaire sur au moins un phosphate dudit fragment.

La fragmentation et le marquage s'effectue en une étape ou en deux étapes et le marquage peut s'effectuer indifféremment avant, après ou simultanément avec la fragmentation.

De préférence, la marquage et/ou la fragmentation s'effectue(nt) en solution homogène sensiblement aqueuse. De préférence, le marquage et la fragmentation s'effectuent simultanément, c'est-à-dire que les réactifs nécessaires à ces deux étapes sont mis ensemble dans une solution sensiblement aqueuse avec l'acide nucléique par exemple. C'est notamment le cas pour la fragmentation chimique ou enzymatique. Dans le cas de la fragmentation mécanique par un moyen physique, marquage et fragmentation s'effectuant simultanément signifie que le moyen physique est appliqué à une solution contenant au moins les acides nucléiques et le réactif de marquage.

Par solution sensiblement aqueuse, on entend une solution contenant au moins 50% d'eau. Cette solution contient de préférence des sels comme une solution tampon. Par solution homogène, on entend une solution monophasique telle qu'une solution eau/DMSO, par opposition à une solution biphasique telle qu'une solution eau/chloroforme.

La fragmentation de l'ADN par l'intermédiaire de la création d'un site abasique s'effectue par voie enzymatique, chimique ou physique.

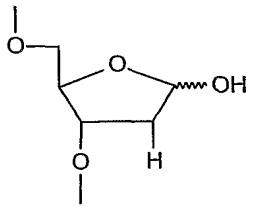
La fragmentation par voie chimique de l'ADN est réalisée en mettant en présence l'acide nucléique avec un moyen chimique de création de site abasique.

Des exemples de conditions de fragmentation chimique par l'intermédiaire d'un site abasique de l'ADN sont donnés dans G. Pratviel et al., Angew. Chem. Int. Ed. Engl., 34, p746-769, 1995; G. Pratviel et al., Adv. Inorg. Chem.,45, p251-312, 1998; D. S. Sigman et al. Chem. Rev., 93, p2295-2316, 1993 ; J. Lhomme et al., Biopolymers (Nucleic Acid Sciences), 52, p65-83, 1999.

Un site abasique est créé quand la liaison N-glycosidique reliant la base modifiée au désoxyribose est hydrolysée, laissant la liaison phosphodiester intacte. Ce phénomène est appelé dépurination (dans le cas des purines) ou dépyrimidination (dans le cas des pyrimidines). La dépurination est beaucoup plus fréquente que la dépyrimidination.

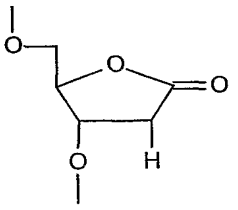
Des agents chimiques, tels que le pH acide, les agents oxydants ou les agents alkylants, peuvent induire le phénomène de dépurination et donc la formation de sites abasiques. Ces agents alkylants sont constitués d'espèces électrophiles qui réagissent avec les sites nucléophiles d'un fragment d'ADN. L'atome d'azote en position 7 de la guanine est le

site principal d'alkylation de l'ADN. En plus de la protonation des purines, l'alkylation est l'une des modifications chimiques susceptibles de rendre la liaison N-glycosidique plus labile.



Site abasique

L'oxydation de l'ADN peut également générer des lésions abasiques dites « alcalilabiles », qui conduisent à la fragmentation de l'ADN en présence d'une base. Par exemple, le résidu ribonolactone peut être généré par réaction du radical hydroxyle (OH) avec le carbone C1'.



Ribonolactone

Le site abasique est très instable. Sous sa forme aldéhydique, il peut subir facilement une β-élimination du phosphodiester fixé en position 3', ce qui conduit à la fragmentation du brin de l'ADN.

La dépurination est spontanée dans des conditions physiologiques (pH 7,4 à 37°C) mais la vitesse de la réaction est très faible de l'ordre de 3.10^{-11} dépurination par seconde, c'est-à-dire inutilisable pour une fragmentation efficace. Pour augmenter la vitesse de réaction, on utilise des agents alkylants qui fragilisent la liaison N-glycosidique ou des enzymes, comme des ADN glycosylases.

Un mode préféré de réalisation de la fragmentation est obtenu par l'utilisation d'un pH acide, c'est-à-dire un pH inférieur à 5, de préférence inférieur à 4. Avantageusement le pH est d'environ 3.

Un tampon formiate de sodium à pH 3 permet de fragmenter de manière efficace les acides nucléiques selon la présente invention. Ce tampon est compatible avec les conditions de marquage en une étape comme cela sera démontré dans les exemples. Encore plus avantageusement, un milieu acide (HCl, carbonate, H_2SO_4) est utilisé.

Dans un mode particulier de la présente invention et dans le but d'augmenter encore la fragmentation, l'acide désoxyribonucléique contient au moins une base modifiée susceptible de générer un site abasique plus facilement.

Diverses bases modifiées sont utilisables comme les N7-alkylpurines, les N3-alkylpurines, les O6-alkylpurines, les 8-bromopurines, les 8-thiopurines, les 8-alkylthiopurines, les 8 azidopurines ou les 8-alkylsulfonylpurines.

Dans le cas où l'acide nucléique à marquer est généré par une technique d'amplification enzymatique comme la PCR, l'utilisation d'une 8-bromopurine permet d'avoir une incorporation efficace dudit nucléotide pendant l'amplification ce qui facilite d'autant le procédé de fragmentation et de marquage selon l'invention, tout en conservant une sensibilité excellente pour l'étape d'amplification enzymatique. Une 8-méthylthiopurine est également incorporée sans gêner l'efficacité d'amplification et d'incorporation par PCR. Une fois incorporée, la base thioéther est oxydée par un peracide ou un dérivé de monopersulfate comme le peroxymonosulfate de potassium par exemple, en sulphonyle correspondant qui est très labile. Il a été démontré que le temps de demi-vie correspondant à l'hydrolyse d'une 8-méthylsulfonylguanosine est inférieur à 2 minutes alors que le temps de demi-vie pour son homologue naturel est de 1000 heures.

Par marqueur, on entend au moins un marqueur capable de générer directement ou indirectement un signal détectable. Une liste non limitative de ces marqueurs suit :

- les enzymes qui produisent un signal détectable par exemple par colorimétrie, fluorescence, luminescence, comme la péroxydase de raifort, la phosphatase alcaline, la β-galactosidase, la glucose-6-phosphate déshydrogénase,

- les chromophores comme les composés fluorescents, luminescents, colorants,
- les groupements à densité électronique détectable par microscopie électronique ou par leur propriété électrique comme la conductivité, l'ampérométrie, la voltamétrie, l'impédance,
- les groupements détectables, par exemple dont les molécules sont de tailles suffisantes pour induire des modifications détectables de leurs caractéristiques physiques et/ou chimiques, cette détection peut être réalisée par des méthodes optiques comme la diffraction, la résonance plasmon de surface, la variation de surface, la variation d'angle de contact ou des méthodes physiques comme la spectroscopie de force atomique, l'effet tunnel, et
- les molécules radioactives comme le ^{32}P, le ^{35}S ou le ^{125}I.

De préférence, le marqueur n'est pas un marqueur radioactif pour éviter les problèmes de sécurité liés à ces marqueurs.

Dans un mode de réalisation particulier de la présente invention, le marqueur est détectable électrochimiquement et en particulier le marqueur est un dérivé d'un complexe de Fer comme un ferrocène.

Des systèmes indirects peuvent être aussi utilisés, comme par exemple des ligands capables de réagir avec un anti-ligand. Les couples ligand/anti-ligand sont bien connus de l'homme du métier, ce qui est le cas par exemple des couples suivants : biotine/streptavidine, haptène/anticorps, antigène/anticorps, peptide/anticorps, sucre/lectine, polynucléotide/ complémentaire du polynucléotide. Dans ce cas, c'est le ligand qui porte la fonction réactive. L'anti-ligand peut être détectable directement par les marqueurs décrits au paragraphe précédent ou être lui-même détectable par un autre couple ligand/anti-ligand.

Un autre exemple de systèmes indirects utilise une liaison covalente spécifique entre le ligand et l'anti-ligand, par exemple méthylcétone et alcoxyamine. Des exemples de ce système sont décrits dans les demandes de brevet WO-A-00/40590 et WO-A-98/05766.

Dans le procédé de la présente invention, la liaison entre le réactif de marquage et l'acide nucléique est covalente, mais il est décrit ci-dessus que des interactions non covalentes peuvent être utilisées notamment dans les systèmes d'empilement ou dans le cas où le marqueur est détectable indirectement. Le terme « attacher » couvre donc ces différentes possibilités.

Ces systèmes de détection indirects peuvent conduire, dans certaines conditions, à une amplification du signal et l'on pourra se reporter aux demandes de brevet antérieures WO-A-00/07982, WO-A-01/92361 et WO-A-95/08000 de la demanderesse pour des exemples d'amplification chimique en utilisant des polymères ou à la demande WO-A-01/44506, toujours de la demanderesse, pour les systèmes d'amplification chimique par empilement.

Dans un mode particulier d'amplification de signal, au moins deux marqueurs sont présents sur le réactif de marquage.

Dans un mode préféré de l'invention, le traceur est un composé fluorescent de faible encombrement stérique comme la fluorescéine, le dansyl, les chromophores du type IR (Li-COR Inc, Lincoln NE, USA), les dérivés cyanines comme le Cy5 et Cy3 (Randolph J.B. and al, Nucleic Acids Res., 25(14), p2923-2929, 1997) et en particulier les dérivés du Cy5 ou bien le traceur est un haptène de faible encombrement stérique comme la biotine ou un dérivé de l'abiétane (voir la demande WO-A-00/07982). Par faible encombrement stérique, on entend un poids moléculaire inférieur à 1000 g/mole.

Dans le cas d'un fluorophore, il est préférable de travailler avec des fluorophores dont la longueur d'onde d'excitation est supérieure à 450 nm, de préférence supérieure à 600 nm.

Dans le cas où le traceur est un haptène qui ne produit pas de signal par lui-même, comme par exemple la biotine, la détection est réalisée par la réaction d'un anti-ligand comme décrit plus haut. Dans le cas de la biotine, on utilise de préférence de la streptavidine ou un anticorps anti-biotine couplé à un composé fluorescent comme la fluorescéine, Cy5 ou la Phycoérythrine. Dans le cas de l'abiétane, on utilise un anticorps monoclonal comme décrit dans WO-A-00/07982.

Le terme « acide désoxyribonucléique» ou ADN signifie un enchaînement d'au moins deux désoxyribonucléotides comprenant éventuellement au moins un nucléotide modifié, par exemple au moins un nucléotide comportant une base modifiée tel que l'inosine, la méthyl-5-désoxycytidine, la diméthylamino-5-désoxyuridine, la désoxyuridine, la diamino-2,6-purine, la bromo-5-désoxyuridine ou toute autre base modifiée permettant l'hybridation.

Cet ADN peut aussi être modifié au niveau de la liaison internucléotidique comme par exemple les phosphorothioates, les H-phosphonates, les alkyl-phosphonates, au niveau du squelette comme par exemple les alpha-oligonucléotides (FR-A-2.607.507) ou les PNA (M. Egholm et al., J. Am. Chem. Soc., 114, 1895-1897, 1992). L'ADN est naturel ou synthétique, et/ ou sous la forme de fragment. L'ADN est simple ou double brin.

En particulier l'ADN est obtenu par une technique d'amplification enzymatique telle que :

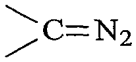
- la PCR (Polymerase Chain Reaction), décrite dans les brevets US-A-4,683,195, US-A-4,683,202 et US-A-4,800,159, et sa dérivée RT-PCR (Reverse Transcription-PCR), notamment dans un format en une étape tel que décrit dans la demande de brevet EP-A-0.569.272,
- la LCR (Ligase Chain Reaction), exposée par exemple dans la demande de brevet EP-A-0.201.184,
- la RCR (Repair Chain Reaction), décrite dans la demande de brevet WO-A-90/01069.

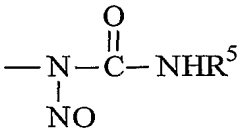
On parle alors d'amplicons pour désigner l'ADN généré par une technique d'amplification enzymatique. L'ADN peut aussi comprendre des ribonucléotides en faible proportion, par exemple inférieur à 10%.

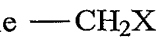
Chacune de ces modifications peut être prise en combinaison pour peu qu'au moins un phosphate soit présent dans l'ADN.

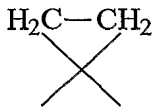
Le réactif de marquage comprend, comme fonction réactive, un motif choisi parmi les composés suivants : diazométhyle ; halogénure d'alkyle ; nitrosourée ; spirocyclopropane ; aziridine ; époxyde ; trifluorosulfonates.

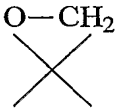
Les fonctions réactives sont décrites ci-dessous :

diazométhyle 

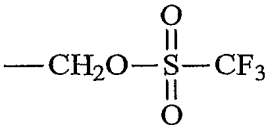
nitrosourée 

halogénure d'alkyle 

spirocyclopropane 

époxyde 

aziridine 

trifluorosulfonate 

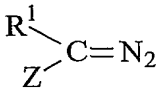
Pour la fonction réactive halogénure d'alkyle, X signifie Br, Cl ou I. Avantageusement le réactif de marquage est la 5-(bromométhyl) fluorescéine.

Pour la fonction réactive nitrosourée, R^5 est un alkyle ou H.

La fonction diazométhyle a déjà été utilisée pour l'alkylation des groupements phosphates, mais un certain nombre de problèmes se pose. D'une part, les dérivés diazo en général sont instables eux-mêmes, ce qui génère des problèmes pour l'utilisation de ces réactifs de marquage dans un kit de marquage, et d'autre part, le produit de couplage est instable ce qui est rédhibitoire si le produit marqué a pour fonction de mettre en évidence la présence d'une molécule cible biologique dans un échantillon quelconque, ou si c'est la cible marquée que l'on souhaite détecter.

Enfin les dérivés portant la fonction diazométhyle sont insolubles dans l'eau ce qui conduit à utiliser des conditions biphasiques pour le couplage avec des acides nucléiques, qui ne sont solubles et stables que dans l'eau ou des tampons aqueux, mais ces conditions ralentissent la vitesse de réaction et donc nuisent à l'efficacité du couplage.

Dans un mode préféré du procédé selon l'invention, le réactif de marquage est choisi parmi les composés de formule (1) :

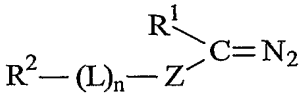


dans laquelle :

- R^1 représente H ou un groupe alkyle, alkyle substitué, aryle ou aryle substitué,
- Z comprend un marqueur détectable.

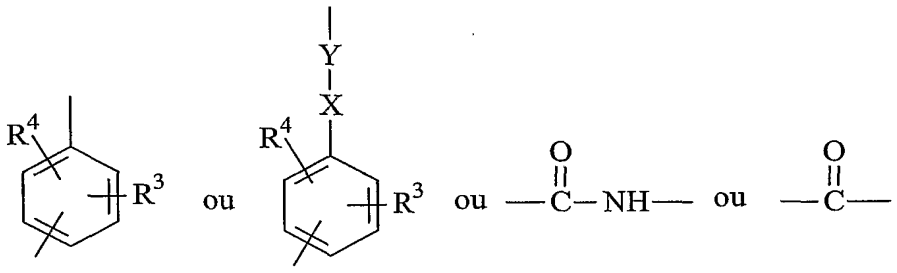
Z et/ou R^1 sont choisis pour stabiliser la fonction diazométhyle, c'est-à-dire au moins un des deux groupes Z ou R^1 a un noyau phényle.

De préférence, le réactif de marquage est choisi parmi les composés de formule (2) :



dans laquelle :

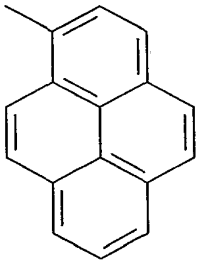
- R^1 représente H ou un groupe alkyle, aryle ou aryle substitué,
- R^2 est un marqueur détectable,
- L est un bras de liaison comportant un enchaînement linéaire d'au moins deux liaisons covalentes et n est égal à 0 ou 1,et
- Z est choisi parmi :



dans lequel :

- R^3 et R^4 représentent indépendamment l'un de l'autre : H, NO_2, Cl, Br, F, I, OR, SR, NR_2, R, NHCOR, CONHR, COOR avec R = alkyle ou aryle,
- -Y-X- représente -CONH-, -NHCO-, $-CH_2O-$, $-CH_2S-$.

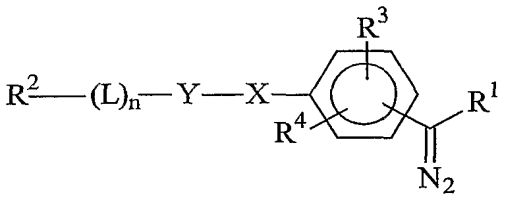
Dans un mode particulier selon la formule (1), Z a la structure suivante :



Dans ce cas et si R^1 est égal à H, le réactif de marquage est le 1-Pyrényldiazométhane (PDAM).

Bien que ce marqueur soit fluorescent, la longueur d'onde d'excitation est trop proche de celles des acides nucléiques. Un détection indirecte par l'utilisation d'un anticorps monoclonal dirigé contre le motif pyrène est préférée. Le mode d'obtention de cet anticorps est bien connu de l'homme du métier (voir par exemple la demande de brevet WO-A-00/07982).

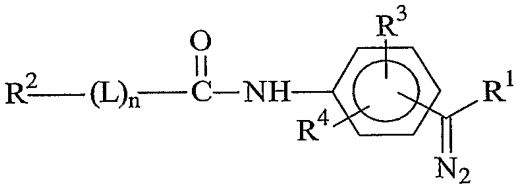
Dans un mode préféré de réalisation du procédé, le réactif de marquage est de formule (3) :



dans laquelle :

- R^1 représente H ou un groupe alkyle, aryle ou aryle substitué,
- R^2 représente un marqueur détectable,
- L est un bras de liaison comportant un enchaînement linéaire d'au moins deux liaisons covalentes et n un nombre entier égal à 0 ou 1,
- R^3 et R^4 représentent indépendamment l'un de l'autre : H, NO_2, Cl, Br, F, I, OR, SR, NR_2, R, NHCOR, CONHR, COOR avec R = alkyle ou aryle,
- -Y-X- représente -CONH-, -NHCO-, $-CH_2O-$, $-CH_2S-$.

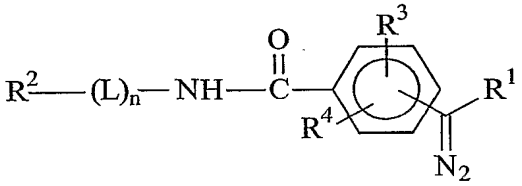
Avantageusement le réactif est de formule (4) :



dans laquelle :

- R^1 représente H ou un groupe alkyle, aryle ou aryle substitué,
- R^2 représente un marqueur détectable,
- L est un bras de liaison comportant un enchaînement linéaire d'au moins deux liaisons covalentes et n un nombre entier égal à 0 ou 1,
- R^3 et R^4 représentent indépendamment l'un de l'autre : H, NO_2, Cl, Br, F, I, OR, SR, NR_2, R, NHCOR, CONHR, COOR avec R = alkyle ou aryle.

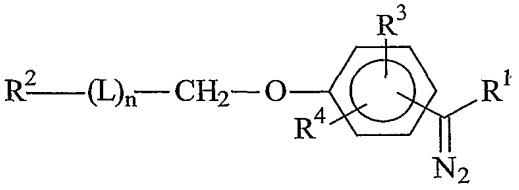
Avantageusement le réactif est de formule (5) :



dans laquelle :

- R^1 représente H ou un groupe alkyle, aryle ou aryle substitué,
- R^2 représente un marqueur détectable,
- L est un bras de liaison comportant un enchaînement linéaire d'au moins deux liaisons covalentes et n un nombre entier égal à 0 ou 1,
- R^3 et R^4 représentent indépendamment l'un de l'autre : H, NO_2, Cl, Br, F, I, OR, SR, NR_2, R, NHCOR, CONHR, COOR avec R = alkyle ou aryle.

Avantageusement le réactif est de formule (6) :

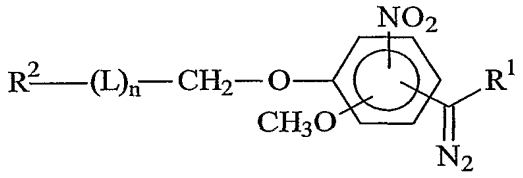


dans laquelle :

- R^1 représente H ou un groupe alkyle, aryle ou aryle substitué,
- R^2 représente un marqueur détectable,
- L est un bras de liaison comportant un enchaînement linéaire d'au moins deux liaisons covalentes et n un nombre entier égal à 0 ou 1,
- R^3 et R^4 représentent indépendamment l'un de l'autre : H, NO_2, Cl, Br, F, I, OR, SR, NR_2, R, NHCOR, CONHR, COOR avec R = alkyle ou aryle.

Dans les formules ci-dessus (3) à (6), avantageusement R^3 et R^4 représentent indépendamment l'un de l'autre : H, NO_2, OCH_3.

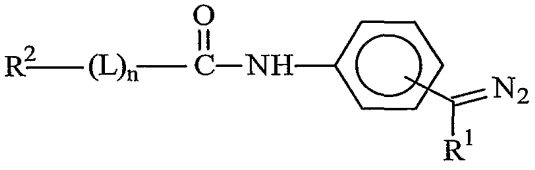
Ainsi, un composé préféré selon la formule (6) est de formule (6') :



dans laquelle :

- R^1 représente H ou un groupe alkyle, aryle ou aryle substitué,
- R^2 représente un marqueur détectable,
- L est un bras de liaison comportant un enchaînement linéaire d'au moins deux liaisons covalentes et n un nombre entier égal à 0 ou 1.

Un composé préféré selon la formule (4) est de formule (4') :



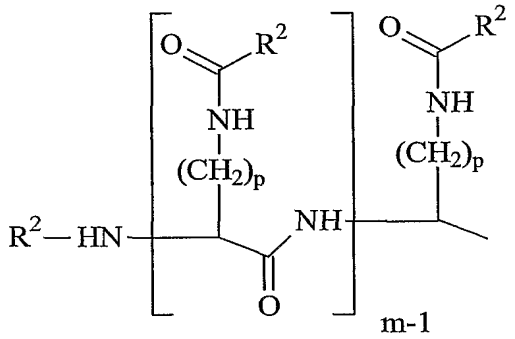
dans laquelle :

- R^1 représente H ou un groupe alkyle, aryle ou aryle substitué,
- R^2 représente un marqueur détectable,

IDT Exhibit 1020.7730

- L est un bras de liaison comportant un enchaînement linéaire d'au moins deux liaisons covalentes et n un nombre entier égal à 0 ou 1.

Dans un mode particulier du procédé où l'on veut amplifier le signal, au moins deux marqueurs sont présents sur le réactif de marquage. En particulier un réactif qui permet de mettre en oeuvre l'amplification de signal selon la présente invention possède une structure R^2-$(L)_n$- de formule (7) ci dessous :



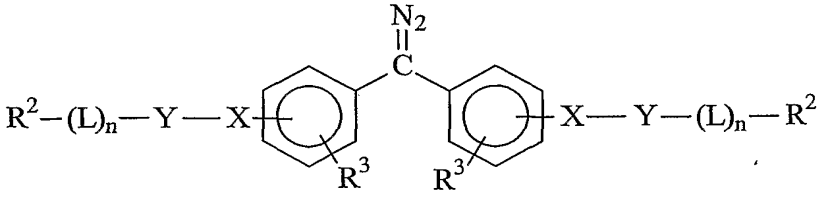
m-1

dans laquelle :

- R^2 représente un marqueur détectable,
- m est un nombre entier compris entre 1 et 100, de préférence compris entre 1 et 20,
- p est un nombre entier compris entre 1 et 10, avantageusement 2 à 6, de préférence 4.

Cette structure de R^2-$(L)_n$ s'applique indifféremment aux formules (2) à (6) précédentes.

Un autre réactif de marquage préféré pour l'amplification de signal est le réactif de formule (8) :

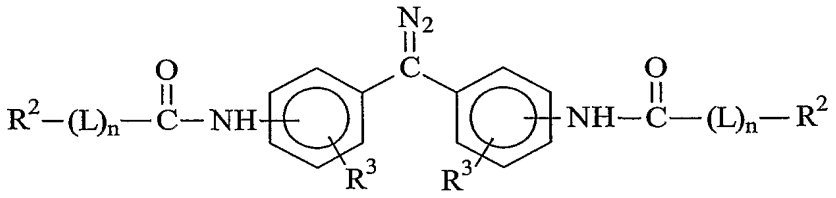


dans laquelle :

- R^2 représente un marqueur détectable,

- R^3 représente H, NO_2, Cl, Br, F, I, OR, SR, NR_2, R, NHCOR, CONHR, COOR avec R = alkyle ou aryle,
- L est un bras de liaison comportant un enchaînement linéaire d'au moins deux liaisons covalentes et n un nombre entier égal à 0 ou 1,
- -Y-X- représente -CONH-, -NHCO-, $-CH_2O-$, $-CH_2S-$.

Avantageusement le réactif pour l'amplification de signal a la formule (9)

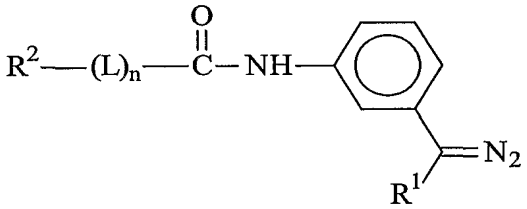


dans laquelle :

- R^2 représente un marqueur détectable,
- R^3 représente H, NO_2, Cl, Br, F, I, OR, SR, NR_2, R, NHCOR, CONHR, COOR avec R = alkyle ou aryle, de préférence R^3 représente H, NO_2 ou OCH_3.,
- L est un bras de liaison comportant un enchaînement linéaire d'au moins deux liaisons covalentes et n un nombre entier égal à 0 ou 1.

Certains réactifs avantageux de l'invention sont :

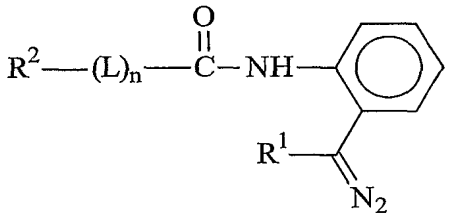
a) de formule (10) :



dans laquelle :

- R^1 représente H ou un groupe alkyle ou aryle ou aryle substitué,
- R^2 représente un marqueur détectable,
- L est un bras de liaison comportant un enchaînement linéaire d'au moins deux liaisons covalentes et n un nombre entier égal à 0 ou 1,

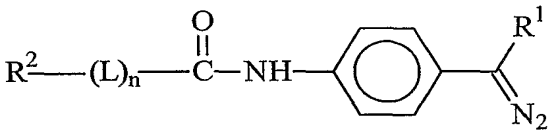
b) de formule (11) :



dans laquelle :

- R^1 représente H ou un groupe alkyle ou aryle ou aryle substitué,
- R^2 représente un marqueur détectable,
- L est un bras de liaison comportant un enchaînement linéaire d'au moins deux liaisons covalentes et n un nombre entier égal à 0 ou 1,

c) de formule (12) :

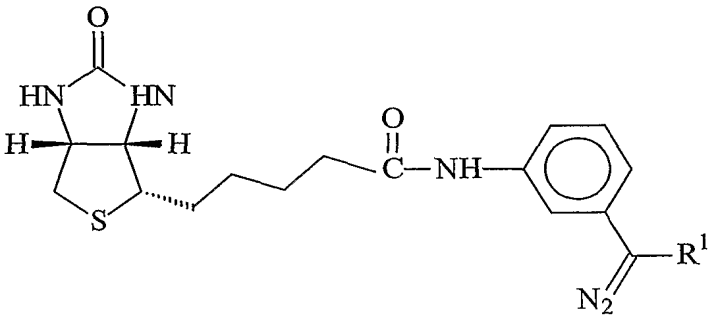


dans laquelle :

- R^1 représente H ou un groupe alkyle ou aryle ou aryle substitué,
- R^2 représente un marqueur détectable,
- L est un bras de liaison comportant un enchaînement linéaire d'au moins deux liaisons covalentes et n un nombre entier égal à 0 ou 1.

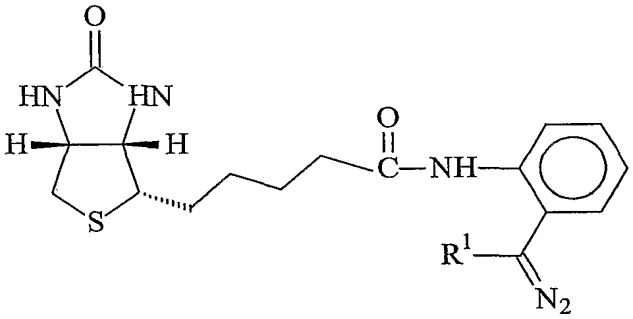
De préférence, le réactif de marquage a :

a) la structure



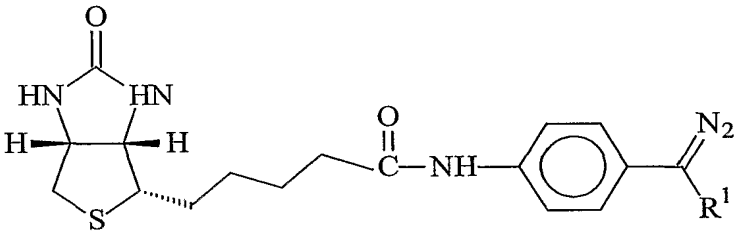
dans laquelle R^1 représente un groupe méthyle ou un phényle, ou

b) la structure :



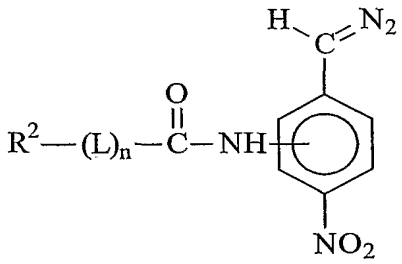
dans laquelle R^1 représente un groupe méthyle ou phényle, ou

c) la structure :



dans laquelle R^1 représente un groupe méthyle ou phényle.

D'autres réactifs préférés selon l'invention ont la formule (13) :



dans laquelle :

- R^2 représente un marqueur détectable,
- L est un bras de liaison comportant un enchaînement linéaire d'au moins deux liaisons covalentes et n un nombre entier égal à 0 ou 1.

En particulier les réactifs de marquage selon le procédé de l'invention sont solubles dans des solvants polaires miscibles à l'eau comme le DMF, le DMSO, CH_3CN, THF, DMA (diméthylacétamide), NMP (N-méthylpyrrolidone), DME (diméthoxyéthane).

De préférence, les réactifs de marquage sont solubles dans le DMSO ou l'eau.

Par solvant miscible à l'eau, on entend un solvant qui est miscible dans une proportion d'au moins 5% en volume avec de l'eau ou un tampon aqueux contenant des sels.

Avantageusement dans les formules précédentes, le bras L comprend un motif éthylène glycol ou polyéthylène glycol pour augmenter la solubilité du réactif dans l'eau.

Ces réactifs peuvent ainsi se fixer en phase homogène sur les acides nucléiques, la phase homogène étant constituée d'une solution sensiblement aqueuse, c'est-à-dire contenant au moins 50% d'eau.

Un objet de la présente invention est de décrire un procédé de détection d'un acide désoxyribonucléique cible (ADN) simple ou double brin comportant les étapes suivantes :

a) fragmenter et marquer lesdits ADN selon l'un quelconque des procédés décrits précédemment,

b) hybrider les fragments marqués sur au moins une sonde nucléique suffisamment spécifique de l'acide nucléique cible,

c) détecter l'hybride formé par l'intermédiaire du marqueur.

Comme cela sera démontré dans les exemples, la dénaturation de l'ADN avant hybridation permet d'augmenter la sensibilité. Cette étape de dénaturation a lieu après la fragmentation et le marquage.

Dans un mode particulier du procédé, une étape d'amplification enzymatique a lieu avant la fragmentation et le marquage. L'étape d'amplification enzymatique génère des amplicons ADN à partir d'un acide nucléique cible ADN, mais aussi à partir d'un acide nucléique cible ARN comme une ARN messager, un ARN de transfert ou un ARN ribosomal. La technique de RT-PCR par exemple est un moyen connu pour amplifier l'ARN.

De préférence, les étapes de fragmentation, marquage et dénaturation ont lieu en même temps, aussi bien dans le cas d'un acide nucléique cible naturel que dans le cas d'un acide nucléique cible obtenu par une technique d'amplification enzymatique.

La fragmentation par création d'un site abasique implique la perte d'une base. Dans les procédés de détection d'acide nucléique cible et en particulier dans les procédés de génotypage c'est-à-dire dans les procédés où l'acide nucléique cible présente un

polymorphisme réparti sur sa séquence, il est nécessaire d'identifier une pluralité de modifications de séquences dont certaines représentent la modification d'une seule base entre la cible et les sondes nucléiques dont la fonction est d'identifier cette modification (sondes de capture). La spécificité est donc essentielle dans ce contexte. De manière surprenante, la présente invention démontre que cette perte d'une base n'affecte en rien la spécificité de l'hybridation pour un acide nucléique marqué et fragmenté selon l'invention.

La présente invention a aussi pour objet un procédé de mise en évidence d'un polymorphisme, réparti en des positions prédéterminées ou non d'un acide nucléique cible par la présence d'une pluralité de délétions et/ou insertions et/ou mutations dans la séquence dudit acide nucléique cible par rapport à une séquence, dite de référence, comportant les étapes suivantes :

a) disposer d'un ADN cible comportant l'ensemble du polymorphisme à étudier, ledit ADN étant éventuellement généré par une technique d'amplification enzymatique,
b) fragmenter et marquer ledit ADN par un des procédés décrits précédemment,
c) hybrider lesdits fragments sur une pluralité de sondes nucléiques dites sondes de capture, la pluralité de sondes de capture étant fixé sur un support solide et la pluralité de sondes de capture couvrant dans son ensemble au moins le polymorphisme à étudier,
d) détecter les hybrides formés entre les fragments marqués et au moins une partie des sondes nucléiques par l'intermédiaire du marqueur et en déduire le polymorphisme de l'ADN cible.

Comme indiqué précédemment une étape de dénaturation après l'étape de fragmentation et marquage permet d'améliorer la sensibilité du procédé, et cette étape est réalisée de préférence simultanément avec l'étape de marquage et de fragmentation.

Le procédé de fragmentation et de marquage selon l'invention est particulièrement utile dans le cas où l'ADN marqué et fragmenté doit s'hybrider avec une multitude d'acides nucléiques, notamment oligonucléotides, fixés sur le support solide à une position prédéterminée pour former une puce à ADN. Par "puce à ADN", on entend un support solide de dimension réduite où sont fixés une multitude de sondes de capture à des positions

prédéterminées. Par multitude, on entend un support solide comportant au moins dix (10) sondes nucléiques de séquences différentes, avantageusement au moins quatre cents (400), de préférence au moins mille (1000).

En effet, la densité des acides nucléiques fixés sur le support solide impose des contraintes stériques importantes lors de l'hybridation et la fragmentation permet d'améliorer cette étape d'hybridation. Des exemples de ces puces à ADN sont donnés par exemple dans les publications de G. Ramsay, Nature Biotechnology, 16, p40-44, 1998; F. Ginot, Human Mutation, 10, p1-10, 1997 ; J. Cheng et al, Molecular diagnosis, 1(3), p183-200, 1996 ; T. Livache et al, Nucleic Acids Research, 22(15), p2915-2921, 1994 ; J. Cheng et al, Nature Biotechnology, 16, p541-546, 1998. Le terme "support solide" tel qu'utilisé ici inclut tous les matériaux sur lesquels peut être fixé un acide nucléique. Le support solide peut être sous la forme d'une plaque de microtitration, d'une membrane, d'une particule ou d'une plaque sensiblement plane.

L'invention concerne l'utilisation d'un ADN marqué, tel que défini ci-dessus, comme sonde de détection d'une acide nucléique cible.

Pour permettre la détection et/ou la quantification et/ou la purification de l'acide nucléique cible, l'ADN marqué est capable de former un complexe d'hybridation avec une sonde nucléique. A titre d'exemple, l'acide nucléique marqué est suffisamment complémentaire de la cible pour s'hybrider spécifiquement en fonction des conditions de réaction et notamment de la température ou de la salinité du milieu réactionnel.

Le procédé de détection est applicable pour le séquençage, le profil d'expression des ARN messagers ou le criblage de mutations à des fins de recherche ainsi que le criblage de drogues dans l'industrie pharmaceutique, le diagnostic de maladies infectieuses (aussi bien en bactériologie, virologie, parasitologie) ou génétiques, le contrôle alimentaire ou industriel.

De nombreuses publications décrivent ce type d'applications : Antoine de Saizieu et al, Nature Biotechnology, 16, p44-48, 1998 ;Thomas R. Gingeras et al Genome Research, 8, p435-448, 1998 ; David J. Lockhart et al, Nature Biotechnology, 14, p1675-1680, 1996 ; Daniel D. Shoemaker et al, Nature Genetics, vol 14, p450-456, 1996 ; R.J. Lipshutz et al;

BioTechniques, 19(3), p442-447, 1995; David G. Wang et al, Science, 280, p1077-1082, 1998 ; Kevin L. Gunderson et al; Genome Research, 8, p1142-1152, 1998 ; Joseph G. Hacia et al; Nature Genetics, vol 14, p441-447, 1996.

La tendance en matière de diagnostic et notamment pour les maladies infectieuses (SIDA ou Tuberculose par exemple) est de baisser le niveau de sensibilité jusqu'à la détection d'une molécule unique dans un échantillon, qui peut représenter plusieurs millilitres dans le cas d'un prélèvement liquide type sang ou urine ou liquide céphalo-rachidien. Ce niveau de sensibilité ne peut être obtenu que si toutes les étapes depuis le prélèvement de l'échantillon jusqu'au rendu de résultat sont optimisées. En particulier dans le cas où une étape d'amplification enzymatique est nécessaire pour obtenir la sensibilité nécessaire (infection virale ou bactérienne comme VIH, VHC ou Tuberculose), un procédé de marquage et/ou fragmentation comme décrit dans la présente invention permet de ne pas affecter la sensibilité de la technique d'amplification, soit parce qu'il n'est pas nécessaire de remplacer les désoxyribonucléotides utilisés dans la technique d'amplification enzymatique, soit parce ce que les désoxyribonucléotides incorporés n'altèrent pas la sensibilité.

Des informations complémentaires peuvent être trouvées dans une autre demande de brevet de la Demanderesse, déposée le 4 mai 2001 sous le numéro d'enregistrement FR01/06040, ainsi que de son extension internationale déposée le même jour que la présente invention.

Les figures et exemples ci-joints représentent des modes particuliers de réalisation et ne peuvent pas être considérés comme limitant la portée de la présente invention.

La figure 1 représente les formules développées de différents réactifs utilisés dans la présente invention ainsi que l'abréviation les désignant (o- signifie *ortho*, m- *méta* et p- *para*).

Les figures 2A à 2I représentent les profils, en fonction du temps, analysés par électrophorèse capillaire du couplage covalent de différents réactifs portant une fonction

diazométhyle sur l'uridine 3' monophosphate (3'-UMP) selon l'exemple 6.1. Les molécules sont les suivantes :

- PDAM en figure 2A,
- DPDAM à 2 mM (millimole(s) par litre) en figure 2B,
- DPDAM à 20 mM en figure 2C,
- PMDAM en figure 2D,
- NPDAM en figure 2E,
- BioDPDAM en figure 2F,
- *méta*-BioPMDAM en figure 2G,
- *para*-BioPMDAM en figure 2H, et
- *ortho*-BioPMDAM en figure 2I.

Les figure 3A à 3D représentent les profils en fonction du temps analysée par électrophorèse capillaire de la réaction du *méta*-BioPMDAM sur quatre (4) nucléotides 3'-monophosphate selon l'exemple 6.2. Les molécules sont les suivantes :

- 3'-CMP en série ribonucléotide selon la figure 3A,
- 3'-AMP en série ribonucléotide selon la figure 3B,
- 3'-GMP en série ribonucléotide selon la figure 3C, et
- 3'-TMP en série désoxyribonucléotide selon la figure 3D.

La figure 4 représente les profils en fonction du temps analysée par électrophorèse capillaire de la réaction du *méta*-BioPMDAM sur un dinucléotide 5'-ApUp selon l'exemple 6.3.

Les figures 5A à 5D représentent le spectre RMN du proton dans D_2O des différents conjugués entre le réactif *méta*-BioPMDAM et quatre (4) ribonucléotides 3'-monophosphate selon l'exemple 6.4. Les molécules sont les suivantes :

- 3'-GMP en figure 5A,
- 3'-AMP en figure 5B,
- 3' CMP en figure 5C, et

- 3' UMP en figure 5D.

La figure 6 représente un schéma de synthèse d'un réactif de marquage portant deux biotines pour l'amplification chimique du signal.

La figure 7 représente le mécanisme de fragmentation en milieu acide par la formation de site abasique.

La figure 8 montre, selon l'exemple 8.1, la cinétique de dégradation à pH acide pour différents nucléosides modifiés (8-bromo-2'-désoxyadénosine (8-BrdA) et la 5-bromo-2'-désoxycytidine (5-BrdC)) ainsi que les quatre nucléosides naturels (dA, dC, dG et dT). Les résultats sont représentés sous forme de pourcentage d'hydrolyse du nucléoside de départ (en ordonnée) par rapport au temps de réaction en minutes (en abscisse).

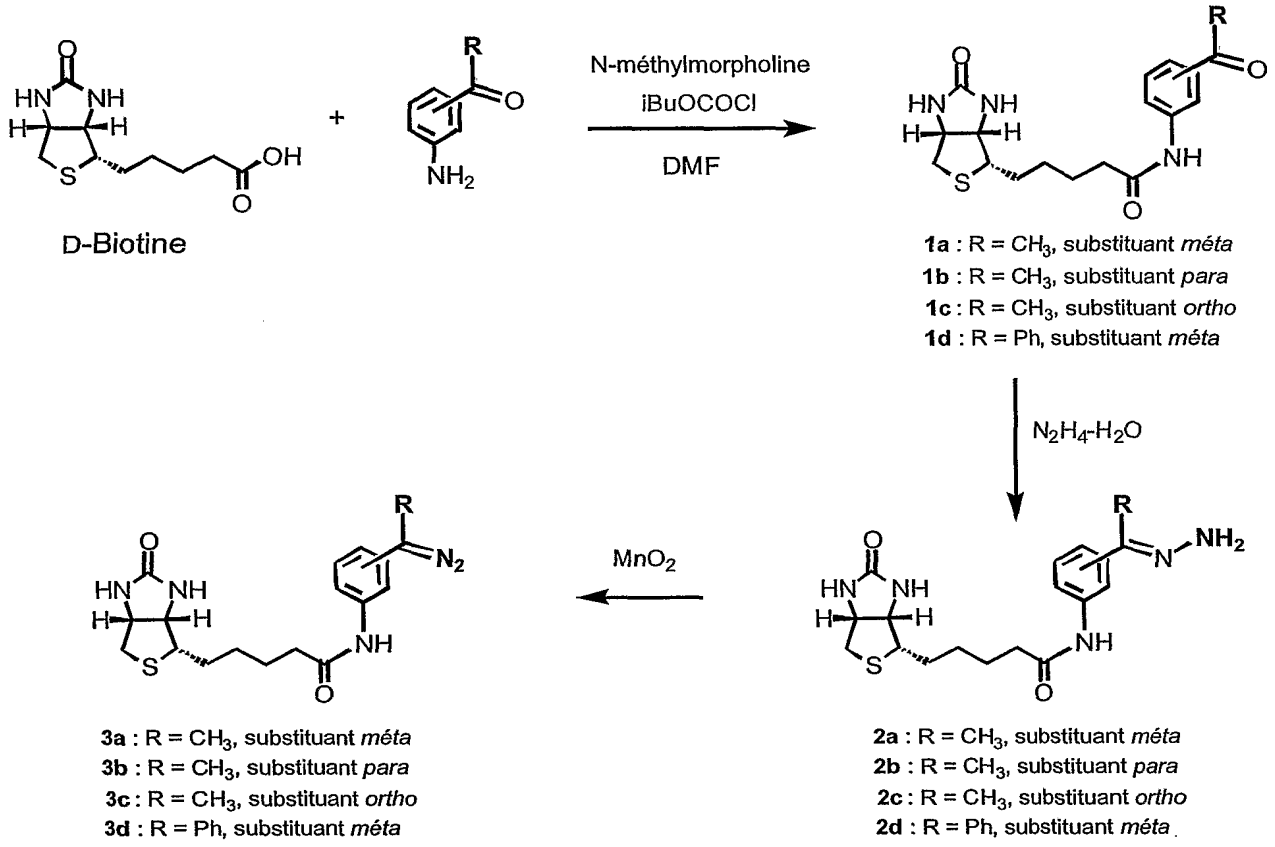
La figure 9 représente, selon l'exemple 11.2, la cinétique de marquage en fonction du temps à une température de 60°C avec le réactif PDAM sur un ODN synthétique 5'-phosphate. Les résultats sont représentés sous forme de pourcentage de marquage par rapport au temps de réaction exprimé en minutes (en abscisse).

La figure 10 représente, selon l'exemple 11.3, le pourcentage de marquage en fonction de la température de réaction. Les résultats sont présentés sur la figure 10 avec en ordonnée le pourcentage de marquage et en abscisse la température de réaction en °C.

La figure 11 représente une voie de synthèse pour un réactif selon la formule 4', en utilisant le réactif commercial 5-Nitrovanilline. La fonction aldéhyde est le précurseur de la fonction diazométhyle.

Exemple 1 : Synthèse de réactifs avec la biotine :

Schéma général de synthèse :



Exemple 1.1 : Synthèse du *méta*-BioPMDAM :

• Composé biotine *méta*-acétophénone **1a** :

On solubilise la D-biotine (1,0 gramme (g), 4,1 millimoles (mmol)) dans 45 millilitres (mL) de DMF anhydre à chaud. On refroidit à 0°C sous argon, puis on ajoute successivement la *N*-méthylmorpholine (590 microlitres (µL), 5,33 mmol) et le chloroformiate d'isobutyle (840 µL, 6,60 mmol). On laisse sous agitation pendant 30 minutes (mn), puis on ajoute la 3-aminoacétophénone (824 mg, 6,10 mmol) et la *N*-méthylmorpholine (480 µL, 4,35 mmol) dans 10 mL de DMF. La solution est maintenue sous agitation à 0°C pendant 2 heures (h), puis on évapore à sec. On reprend le résidu dans 3 mL de MeOH, puis on ajoute 50 mL d'eau. Le précipité obtenu est filtré, lavé avec de l'eau, du CH_2Cl_2 et de l'éther pour donner 1,2 g (80 %) de produit **1a** brut. Une recristallisation dans le couple MeOH-H_2O donne **1a** (1,01 g, 70 %) sous forme d'une poudre blanche.

F 145°C. - IR (KBr) : 3280, 2931, 2857, 1691, 1590,1540, 1487, 1434, 1298, 1266 cm^{-1}. - RMN ^1H (300 MHz, DMSO-d_6) δ = 1,3-1,7 (m, 6 H) ; 2,33 (t, J = 8 Hz, 2 H) ; 2,55 (s, 3 H) ; 2,58 ; (d, J = 12 Hz, 1 H) ; 2,83 (dd, J = 12 et 5 Hz, 1 H) ; 3,13 (m, 1 H) ; 4,15 (m, 1 H) ; 4,31 (m, 1 H) ; 6,34 (s, 1 H) ; 6,41 (s, 1 H) ; 7,44 (t, J = 8 Hz, 1 H) ; 7,64 (d, J = 8 Hz, 1 H) ; 7,85 (d, J = 8 Hz, 1 H) ; 8,17 (s, 1 H) ; 10,05 (s, 1 H). - MS (FAB/glycérol), *m/z*: 362 $[M+H]^+$.

• Composé *méta*-hydrazone **2a :**

Une solution de **1a** (500 mg, 1,38 mmol) et d'hydrazine monohydrate (200 µL, 4,15 mmol) dans de l'éthanol absolu (8 mL) est chauffée à reflux pendant 2 h. Après refroidissement à température ambiante, le précipité blanc est filtré, lavé avec de l'eau, puis avec de l'éther et séché. On obtient ainsi 385 mg (74 %) de produit **2a** sous forme d'une poudre blanche.

F 185°C. - IR (KBr) : 3298, 2931, 2857, 1698, 1665, 1626, 1541, 1494, 1470, 1446, 1330, 1265 cm^{-1}. - RMN ^1H (300 MHz, DMSO-d_6) δ = 1,3-1,7 (m, 6 H) ; 1,98 (s, 3 H) ; 2,26 (t, J = 8 Hz, 2 H) ; 2,56 ; (d, J = 12 Hz, 1 H) ; 2,81 (dd, J = 12 et 5 Hz, 1 H) ; 3,11 (m, 1 H) ; 4,13 (m, 1 H) ; 4,29 (m, 1 H) ; 6,39 (s, 3 H) ; 6,42 (s, 1 H) ; 7,22 (m, 2 H) ; 7,50 (d, J = 8 Hz, 1 H) ; 7,84 (s, 1 H) ; 9,82 (s, 1 H). - MS (FAB/glycérol), *m/z*: 376 $[M+H]^+$.

• Composé *méta*-diazométhane **3a** :

On solubilise **2a** (180 mg, 0,48 mmol) dans 2 mL de DMF. On ajoute alors MnO_2 (340 mg, 3,9 mmol). Après 30 minutes d'agitation à température ambiante, le mélange est filtré à travers un entonnoir fritté contenant de la célite (épaisseur : 0,5 cm) et des tamis moléculaires en poudre 3 Å (0,5cm). Le mélange réactionnel est concentré jusqu'à un volume d'environ 0,5 mL, puis 5 mL d'éther sont ajoutés. Le précipité résultant est filtré, lavé à l'éther puis séché. Le composé **3a** (170 mg, 95 %) est obtenu sous forme d'une poudre rose.

F 160°C. - IR (KBr) : 3278, 2935, 2859, 2038, 1704, 1666, 1605, 1577, 1536, 1458, 1430, 1263 cm^{-1}. - RMN ^1H (300 MHz) δ = 1,3-1,7 (m, 6 H) ; 2,11 (s, 3 H); 2,28 (t, J = 8 Hz, 2 H); 2,57 ; (d, J = 12 Hz, 1 H) ; 2,81 (dd, J = 12 et 5 Hz, 1 H) ; 3,11 (m, 1 H) ; 4,13 (m, 1 H) ; 4,29 (m, 1 H) ; 6,33 (s, 1 H) ; 6,41 (s, 1 H) ; 6,60 (m, 1 H) ; 7,25 (m, 3 H) ; 9,84 (s, 1 H).).

Exemple 1.2 : Synthèse du *para*-BioPMDAM :

• Composé biotine *para*-acétophénone **1b** :

On solubilise la D-biotine (1 g, 4,1 mmol) dans 45 mL de DMF anhydre à chaud. On refroidit à 0°C sous argon puis on ajoute successivement la *N*-méthylmorpholine (590 µL, 5,33 mmol) et chloroformiate d'isobutyle (840 µL, 6,60 mmol). On laisse sous agitation pendant 30 min, puis on ajoute la 4-aminoacétophénone (824 mg, 6,10 mmol). La solution est maintenue sous agitation à 0°C pendant 2 h puis on évapore à sec. On reprend le résidu dans 50 mL d'eau. Le précipité obtenu est filtré, lavé avec de l'eau puis avec 50mL de MeOH à chaud. Le précipité blanc est dissous dans du DMF en chauffant puis la solution obtenue est filtrée et lavée au MeOH. Le filtrat est récupéré et évaporé à sec pour donner 888 mg de **1b** (2,46 mmol, 60 %).

F 260°C. - IR (KBr) : 3260, 2930, 2358, 1706, 1673, 1610, 1526, 1401, 1380, 1322, 1257, 1150 cm^{-1}. - RMN ^1H (200 MHz, DMSO-d_6) δ = 8,82 (s, 1H, NH-CO) ; 7,57 (d, 2H, J = 9 Hz, Ar-H) ; 6,83 (d, 2H, J = 9 Hz, Ar-H) ; 6,40 (s large, 1H, NH-CO-NH) ; 6,32 (s large, 1H, NH-CO-NH) ; 4,28 (m, 1H, CH_2-CH-NH) ; 4,12 (m, 1H, CH-CH-NH) ; 3,11 (m, 1H, CH-S) ; 2,80 et 2,55 (système ABX, 2H, $^2J_{AB}$= 5 Hz, $^3J_{AX}$ = 12 Hz, $^3J_{BX}$ = 0 Hz, CH_2-S) ; 2,35 (t, 2H, J = 8 Hz, CH_2-CO) ; 2,10 (s, 3H, CH_3) ; 1,60-1,34 (m, 6H, $(CH_2)_3$).

• Composé *para*-hydrazone **2b** :

Le composé **1b** (870 mg, 2,4 mmol) est dissous à chaud dans l'éthanol (99 %, 8 mL) puis l'hydrazine monohydrate (995 µL, 19,5 mmol) est ajoutée. La solution est chauffée à reflux pendant 3 h. Le précipité blanc obtenu est filtré, lavé avec de l'eau glacée. On obtient ainsi 820 mg (90 %) de produit **2b** sous forme d'une poudre blanche.

F 305°C. – IR (KBr) : 3281, 3183, 2930, 2857, 1698, 1658, 1593, 1521, 1459, 1401, 1325, 1263, 1187 cm^{-1}. - RMN ^1H (200 MHz, DMSO-d_6) δ = 9,68 (s, 1H, NH-CO) ; 7,52 (s, 4H, J = 9 Hz, Ar-H) ; 6,43 (s large, 1H, NH-CO-NH) ; 6,35 (s large, 1H, NH-CO-NH) ; 6,21 (s, 2H, NH_2) 4,29 (m, 1H, CH_2-CH-NH) ; 4,12 (m, 1H, CH-CH-NH) ; 3,12 (m, 1H, CH-S) ; 2,81 et 2,56 (système ABX, 2H, $^2J_{AB}$= 5 Hz, $^3J_{AX}$ = 12 Hz, $^3J_{BX}$ = 0 Hz, CH_2-S) ; 2,32 (t, 2H, J = 8 Hz, CH_2-CO) ; 1,97 (s, 3H, CH_3) ; 1,63-1,36 (m, 6H, $(CH_2)_3$).

• Composé *para*-diazométhane **3b** :

On solubilise **2b** (200 mg, 0,53 mmol) dans 10 mL de DMF. On ajoute alors 800 mg de MnO_2. Après 10 minutes d'agitation, le mélange est filtré à travers une couche mixte Célite (0,5 cm) - tamis moléculaire (0,5 cm en poudre). Le mélange réactionnel est évaporé à sec puis lavé à l'éther et séché. Le composé **3b** (190 mg, 96 %) est obtenu sous forme d'une poudre rose.

F 180°C (dec). - IR (KBr) : 3257, 2930, 2857, 2032, 1698, 1597, 1524, 1510, 1455, 1404, 1307, 1259, 1180 cm^{-1}. - RMN ^1H (200 MHz, DMSO-d_6) δ = 10,18 (s, 1H, NH-CO) ; 7,88 (d, 2H, J = 6 Hz, Ar-H) ; 7,7 (d, 2H, J = 6 Hz, Ar-H) ; 6,41 (s large, 1H, NH-CO-NH) ; 6,34 (s large, 1H, NH-CO-NH) ; 4,28 (m, 1H, CH_2-CH-NH) ; 4,12 (m, 1H, CH-CH-NH) ; 3,11 (m, 1H, CH-S) ; 2,80 et 2,55 (système ABX, 2H, $^2J_{AB}$= 5 Hz, $^3J_{AX}$ = 12 Hz, $^3J_{BX}$ = 0 Hz, CH_2-S) ; 2,35 (t, 2H, J = 8 Hz, CH_2-CO) ; 2,10 (s, 3H, CH_3) ; 1,60-1,34 (m, 6H, $(CH_2)_3$).

Exemple 1.3 : Synthèse de l'*ortho*-BioPMDAM :

• Composé biotine *ortho*-acétophénone **1c** :

On solubilise la D-biotine (1 g, 4,1 mmol) dans 45 mL de DMF anhydre à chaud. On refroidit à 0°C sous argon puis on ajoute successivement la *N*-méthylmorpholine (590 µL, 5,33 mmol) et chloroformiate d'isobutyle (840 µL, 6,60 mmol). On laisse sous agitation pendant 30 min, puis on ajoute la 2-aminoacétophénone (824 mg, 6,10 mmol). La solution est maintenue sous

agitation à température ambiante pendant 3 h 30 puis on évapore à sec. On reprend le résidu dans 50 mL d'eau. Le précipité obtenu est filtré, lavé avec de l'eau puis avec 50mL de MeOH à chaud. Le précipité obtenu est filtré, lavé avec de l'eau. Une recristallisation est faite en dissolvant le produit dans du MeOH à chaud et en reprécipitant par addition d'eau. Le précipité est filtré, lavé à l'eau, puis à l'éther pour donner 1,1 g (2,95 mmol, 72 %) de produit **1c** brut.

F 150°C. - IR (KBr) : 3248, 2930, 2857, 2359, 1691, 1669, 1651, 1582, 1528, 1448, 1354, 1310, 1245, 1161 cm^{-1}. - RMN ^1H (200 MHz, DMSO-d_6) δ = 11,24 (s, 1H, NH-CO) ; 8,33 (d, 1H, J = 8,5 Hz, Ar-H) ; 7,97 (d, 2H, J = 8 Hz, Ar-H) ; 7,57 (t, 1H, J = 7 Hz, Ar-H) ; 7,18 (t, 1H, J = 7 Hz, Ar-H) ; 6,44 (s large, 1H, NH-CO-NH) ; 6,35 (s large, 1H, NH-CO-NH) ; 4,30 (m, 1H, CH_2-CH-NH) ; 4,14 (m, 1H, CH-CH-NH) ; 3,12 (m, 1H, CH-S) ; 2,80 et 2,55 (système ABX, 2H, $^2J_{AB}$= 5 Hz, $^3J_{AX}$ = 12 Hz, $^3J_{BX}$ = 0 Hz, CH_2-S) ; 2,61 (s, 3H, CH_3) ; 2,37 (t, 2H, J = 8 Hz, CH_2-CO) ; 1,62-1,38 (m, 6H, $(CH_2)_3$).

• Composé *ortho*-hydrazone **2c** :

Le composé **1c** (500 mg, 1,38 mmol) est dissous à chaud dans l'éthanol (99 %, 8 mL) puis l'hydrazine monohydrate (572 µL, 11,1 mmol) est ajoutée. La solution est chauffée à reflux pendant 50 minutes. La solution est évaporée à sec. Le précipité blanc obtenu est filtré, lavé avec de l'eau puis séché à l'éther. On obtient ainsi 416 mg (11,1 mmol, 80 %) de produit **2c** sous forme d'une poudre blanche.

F 161°C. – IR (KBr) : 3412, 3240, 2930, 2857, 2351, 1706, 1677, 1604, 1590, 1531, 1463, 1444, 1372, 1303, 1270, 1169 cm^{-1}.- RMN ^1H (200 MHz, DMSO-d_6) δ =11,97 (s, 1H, NH-CO) ; 8,35 (d, 1H, J = 8 Hz, Ar-H) ; 7,45 (d, 1H, J = 7 Hz, Ar-H) ; 7,19 (t, 1H, J = 7,5 Hz, Ar-H) ; 7,04 (t, 1H, J = 7 Hz, Ar-H) ; 6,61 (s, 2H, NH_2) ; 6,42 (s large, 1H, NH-CO-NH) ; 6,35 (s large, 1H, NH-CO-NH) ; 4,32 (m, 1H, CH_2-CH-NH) ; 4,14 (m, 1H, CH-CH-NH) ; 3,12 (m, 1H, CH-S) ; 2,81 et 2,56 (système ABX, 2H, $^2J_{AB}$= 5 Hz, $^3J_{AX}$ = 12 Hz, $^3J_{BX}$ = 0 Hz, CH_2-S) ; 2,31 (t, 2H, J = 8 Hz, CH_2-CO) ; 2,09 (s, 3H, CH_3) ; 1,63-1,36 (m, 6H, $(CH_2)_3$).

• Composé *ortho*-diazométhane **3c** :

On solubilise **2c** (200 mg, 0,53 mmol) dans 10 mL de DMF. On ajoute alors 800 mg de MnO_2. Après 15 minutes d'agitation, le mélange est filtré à travers une couche mixte Célite(0,5cm)-tamis moléculaire(0,5cm en poudre). Le mélange réactionnel est évaporé à sec puis lavé à l'éther et séché. Le composé **3c** (130 mg, 65 %) est obtenu sous forme d'une poudre rose.

F 110°C. - IR (KBr) : 3248, 2930, 2857, 2367, 2342, 2038, 1699, 1521,1456 cm^{-1}. - RMN ^1H (200 MHz, DMSO-d_6) δ = 9,37 (s, 1H, NH-CO) ; 7,26-7,00 (m, 4H, Ar-H) ; 6,43 (s large, 1H, NH-CO-NH) ; 6,35 (s large, 1H, NH-CO-NH) ; 4,30 (m, 1H, CH_2-CH-NH) ; 4,15 (m, 1H, CH-CH-NH) ; 3,12 (m, 1H, CH-S) ; 2,82 et 2,54 (système ABX, 2H, $^2J_{AB}$= 5 Hz, $^3J_{AX}$ = 12 Hz, $^3J_{BX}$ = 0 Hz, CH_2-S) ; 2,24 (t, 2H, J = 8 Hz, CH_2-CO) ; 2,12 (s, 3H, CH_3) ; 1,63-1,37 (m, 6H, $(CH_2)_3$).

Exemple 1.4 : Synthèse du *méta*-BioDPDAM :

• Composé *méta*-benzophénone **1d** :

On solubilise la D-biotine (500 mg, 2,05 mmol) dans 23 mL de DMF anhydre à chaud. On refroidit à 0°C sous argon, puis on ajoute successivement la *N*-méthylmorpholine (295 µL, 2,67 mmol) et le chloroformiate d'isobutyle (420 µL, 3,28 mmol). On laisse sous agitation pendant 30 min, puis on ajoute la 3-aminobenzophénone (605 mg, 3,07 mmol) et la *N*-méthylmorpholine (240 µL, 2,17 mmol) dans 7 mL de DMF. La solution est maintenue sous agitation à 0°C pendant 2 h, puis on évapore à sec. On reprend le résidu dans 1 mL de MeOH, puis on ajoute 25 mL d'eau. Le précipité obtenu est filtré, lavé avec de l'eau, puis avec de l'éther pour donner 810 mg (93 %) de produit **1d** brut. Une recristallisation dans le couple MeOH-H_2O donne **1d** (630 mg, 72 %) sous forme d'une poudre blanche.

RMN ^{1}H (200MHz, DMSO-d_6) δ =10,10 (s, 1H, NH-CO) ; 8-7,39 (m, 9H, Ar-H) ; 6,43 (s large, 1H, NH-CO-NH) ; 6,35 (s large, 1H, NH-CO-NH) ; 4,27 (m, 1H, CH_2-CH-NH) ; 4,13 (m, 1H, CH-CH-NH) ; 3,12 (m, 1H, CH-S) ; 2,84 et 2,55 (système ABX, 2H, $^2J_{AB}$= 5 Hz, $^3J_{AX}$ = 12 Hz, $^3J_{BX}$ = 0 Hz, CH_2-S) ; 2, 31 (t, 2H, J = 8 Hz, CH_2-CO) ; 1,59-1,36 (m, 6H, $(CH_2)_3$).

• Composé *méta*-hydrazone **2d** :

On solubilise **1d** (350 mg, 0,83 mmol) dans 5,5 mL d'éthanol absolu puis on ajoute l'hydrazine monohydrate (140 µL, 2,48 mmol). La solution est chauffée à reflux pendant une nuit. Après évaporation, le produit est repris dans 1 mL d'éthanol et de l'eau. Le précipité blanc est recristallisé : on le dissout dans un minimum d'éthanol à chaud et l'on ajoute de l'eau jusqu'à apparition d'un léger trouble. Après refroidissement, le précipité obtenu est lavé à l'eau puis séché à l'éther. On obtient ainsi 264 mg (73 %) de produit **2d** sous forme d'une poudre blanche.

RMN ^{1}H (200 MHz, DMSO-d_6) δ = 9,99 (s, 1H, NH-CO) ; 9,80 (s, 2H, NH_2) ; 7,54-6,88 (m, 9H, Ar-H) ; 6,26 (s large, 1H, NH-CO-NH) ; 6,21 (s large, 1H, NH-CO-NH) ; 4,28 (m, 1H, CH_2-CH-NH) ; 4,13 (m, 1H, CH-CH-NH) ; 3,12 (m, 1H, CH-S) ; 2,78 et 2,59 (système ABX, 2H, $^2J_{AB}$= 5 Hz, $^3J_{AX}$ = 12 Hz, $^3J_{BX}$ = 0 Hz, CH_2-S) ; 2,27 (t, 2H, J = 8 Hz, CH_2-CO) ; 1,57-1,36 (m, 6H, $(CH_2)_3$).

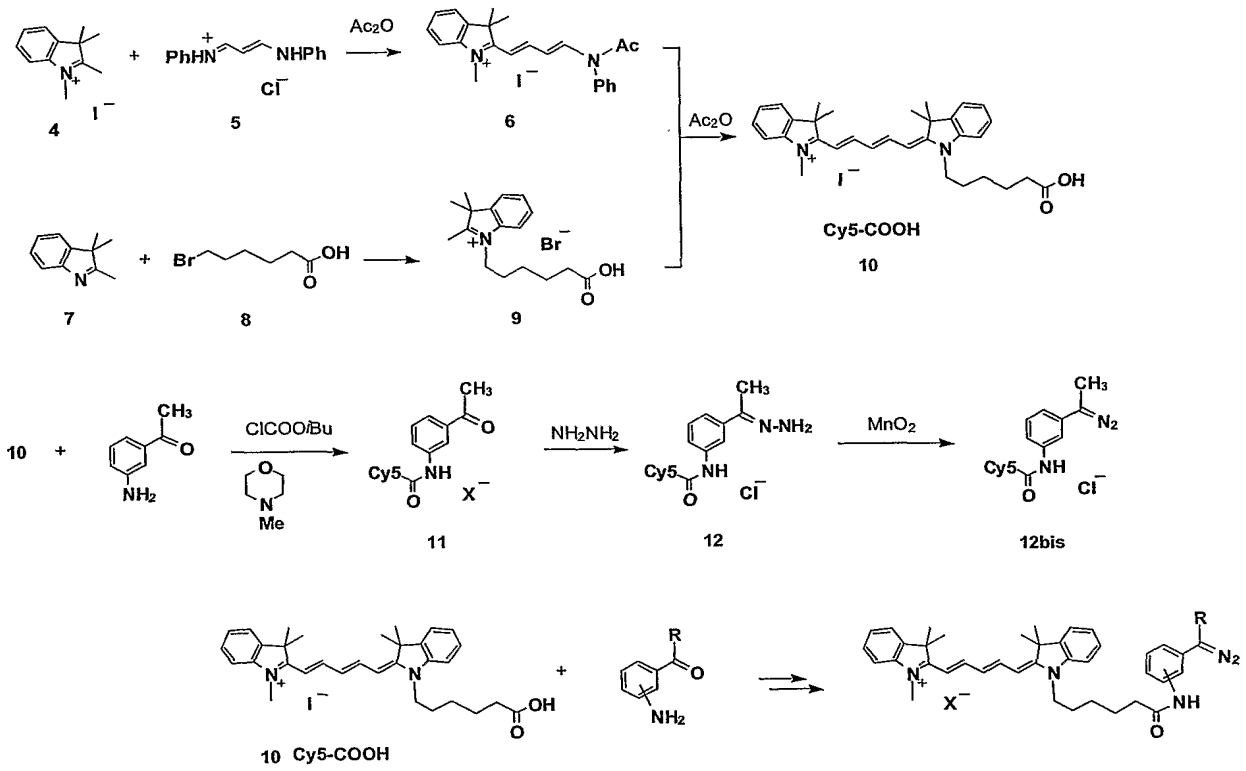
• Composé *méta*-diazodiphényl **3d** :

On solubilise **3d** (500 mg, 0,53 mmol) dans 1 mL de THF. On ajoute alors 80 mg de MnO_2 activé. Après 5 minutes d'agitation à température ambiante, le mélange est filtré à travers une couche mixte Célite (0,5 cm) -tamis moléculaire (0,5 cm en poudre). Le mélange réactionnel est évaporé à sec. Le composé **3d** (47 mg, 100 %) est obtenu sous forme d'une huile violette.

RMN ^{1}H (200 MHz, DMSO-d_6) δ = 9,95 (s, 1H, NH-CO) ; 7,60-6,9 (m, 9H, Ar-H) ; 6,42 (s large, 1H, NH-CO-NH) ; 6,35 (s large, 1H, NH-CO-NH) ; 4,28 (m, 1H, CH_2-CH-NH) ; 4,14

(m, 1H, CH-C<u>H</u>-NH) ; 3,12 (m, 1H, CH-S) ; 2,83 et 2,59 (système ABX, 2H, $^2J_{AB}$= 5 Hz, $^3J_{AX}$ = 12 Hz, $^3J_{BX}$ = 0 Hz, CH_2-S) ; 2,27 (t, 2H, J = 8 Hz, CH_2-CO) ; 1,58-1,35 (m, 6H, $(CH_2)_3$).

<u>Exemple 2 : Réactif de marquage avec le Cy5 : Cy5-PMDAM :</u>



• <u>Composé Iodure de 2-[4-(*N*-acétyl-*N*-phénylamino)buta-1,3-diényl]-1,2,3,3-tétraméthyl[3H] indolium **6** :</u>

Le mélange du monochlorhydrate de malonaldehyde-bis(phénylimine) **5** (18,3 g ; 70,0 mmole), de NaOAc (9,0 g ; 110 mmole) et de l'iodure de 1,2,3,3- tétraméthyl[3H]indolium **4** (4,25 g ; 14,1 mmole) dans de l'anhydride acétique (75 mL) est chauffé à 110°C pendant 20 min. Après refroidissement, on ajoute de l'éther (350 mL) et le solide brun précipité est filtré, lavé avec de l'éther (3 × 100 mL). Le solide est redissout dans 150 mL de CH_2Cl_2, filtré (élimination des sels minéraux) puis évaporé pour donner un solide brun (6,0 g, 90 %).

RMN ^1H ($CDCl_3$) : δ = 8,64 (d ; 1H ; J = 12 Hz ; 1-H) ; 8,14 (t ; 1H ; J = 16 ; 12 Hz ; 3-H) ; 7,63-7,19 (m ; 9H) ; 6,90 (d ; 1H ; J = 15 Hz ; 4-H) ; 5,82 (t ; 1H ; J = 12 ; 13 Hz ; 2-H) ; 4,06 (s ; 3H ; NCH_3) ; (2,16 (s ; 3H ; -$COCH_3$) ; 1,74 (s ; 6H ; CH_3).

• Composé Bromure de 1-(5-carboxypentyl)-2,3,3-triméthyl[3H]indolium **9** :

On mélange le 2,3,3-triméthylindole **7** (10,0 g ; 62,8 mmole) et l'acide 6-bromohexanoïque **8** (12,3 g ; 62,8 mmole) sans solvant et chauffe à 110°C pendant 12 h sous argon. Le mélange réactionnel pâteux rouge-violet est lavé avec de l'acétate d'éthyle (2 × 60 mL, on triture la pâte avec la spatule et on décante le surnageant), puis avec de l'acétone (50 mL, la pâte solidifie). Le solide rose est filtré puis séché sous vide (16,0 g ; 73 %).

• Composé Cy5COOH **10** :

Le mélange de l'iodure **6** (6,0 g ; 12,7 mmole), du bromure **9** (4,5 g ; 12,7 mmole) et de NaOAc (2,6 g ; 32 mmole) dans de l'anhydride acétique (35 mL) est chauffé à 110 °C pendant 20 min. Après refroidissement, on ajoute de l'éther (150 mL) et le précipité est filtré et lavé avec de l'éther (3 × 50 mL). Le solide est dissous dans 100 mL de CH_2Cl_2, filtré et purifié par chromatographie sur colonne de SiO_2 (éluant : MeOH 5-10 %/CH_2Cl_2). On obtient 3,4 g (44 %).

RMN ^1H ($CDCl_3$) : δ = 8,03 (t ; 2H ; J = 10 ; 11 Hz, 2-H, 4-H) ; 7,38-6,91 (m ; 9H ; Ar-H, 3-H) ; 6,41 (d ; 1H ; J = 14 Hz ; 1-H) ; 6,31 (d ; 1H ; J = 13 Hz ; 5-H) ; 4,07 (t ; 2H, J = 7 ; 7 Hz ; α-CH_2) ; 3,68 (s ; 3H ; NCH_3) ; 2,47 (t ; 2H, J = 7 ; 7 Hz ; ε-CH_2) ; 1,71 (m ; 18H ; CH_3, β,γ et δ-CH_2).

• Composé de couplage de la 3-aminoacétophénone avec le Cy5COOH **10** (produit **11**) :

À une solution de Cy5COOH **10** (1,19 g ; 1,9 mmole) dans 12 mL de CH_2Cl_2, on ajoute *N*-méthylmorpholine (360 µL ; 3,2 mmole). La solution est refroidie avec un bain de glace, puis on ajoute le chloroformiate d'isobutyle (480 µL ; 3,7 mmole). Après 5 minutes d'agitation, on additionne la 3-aminoacétophénone (488 mg ; 3,6 mmole). On agite le mélange à température

ambiante pendant 3 h. En ajoutant 250 mL d'éther, on obtient un solide pâteux. Après agitation, on laisse reposer le solide au fond de ballon et on décante le surnageant. On ajoute de nouveau 50 mL d'éther et on triture avec une spatule pour obtenir un solide. Celui-ci est filtré, lavé avec de l'eau, de l'éther, puis séché sous vide. Le produit (iodure) est ensuite dissous dans de l'éthanol, passé sur colonne d'amberlite IRA900 (Cl^- ; 15 g). La solution éthanolique recueillie est évaporée à sec puis passée sur colonne de SiO_2. On obtient 0,93 g (77 %) de solide bleu.

• Composé Cy5-hydrazone **12** :

On ajoute à une solution de l'acétophénone **11** (0,93 g ; 1,46 mmole) dans 5 mL d'éthanol absolu l'hydrazine monohydraté (180 µL ; 3,1 mmole) que l'on agite à température ambiante pendant 7 h. On ajoute 50 mL d'éther et le précipité est filtré et lavé avec de l'éther. On dissout le produit brut dans 50 mL de CH_2Cl_2, la solution est filtrée puis concentrée à 10 mL. Le produit est précipité par ajout de 100 mL d'éther, filtré, lavé à l'éther et séché sous vide. On obtient 540 mg de produit **12** (57 %).

• Composé Cy5PMDAM **12bis** :

A une solution de 100 mg d'hydrazone **12** dans 2 mL de DMF, on ajoute 300 mg de MnO_2 et l'on agite vigoureusement pendant 10 min. On filtre la suspension à travers une couche de célite et lavée avec DMF (3×500 µL). On ajoute 50 mL d'éther et l'huile déposée est triturée avec une spatule et le surnageant est décanté. On répète l'opération de lavage trois fois avec 25 mL d'éther et le solide ainsi obtenu est filtré, séché. On obtient 65 mg (65 %) de produit **12bis**. La pureté du produit est d'environ 80-85 % (RMN ^1H).

Exemple 3 : Autres réactifs synthétisés :

Exemple 3.1 : Synthèse de *para*-nitrophényldiazométhane (NPDAM) :

Le 4-nitrobenzaldéhyde hydrazone est commercialement disponible (référence 28,118-2, Aldrich, France).

On travaille donc sur 600 mg (3,64 mmole) de ce produit que l'on dissout dans 9 mL de THF. La solution est laissée sous agitation pendant 5 minutes puis 1,26 g (4 équivalent, 14,56 mmole) de MnO_2 sont additionnés avec précaution. Le mélange est laissé sous agitation pendant 10 minutes, puis filtré. Le filtrat récupéré est évaporé à sec. Après lavage au pentane, le composé *para*-nitrophényldiazométhane est obtenu sous forme d'une poudre orange vif avec un rendement de 79 % (468 mg, 2,87 mmole).

F 80-82°C. - RMN ^1H (300 MHz, DMSO-d_6) δ = 8,11 (d, 2H, J = 9 Hz, Ar-$\underline{H_3}$) ; 7,18 (d, 2H, J = 9 Hz, Ar-$\underline{H_2}$); 6,06 (s, 1H, C$\underline{H_1}$-N_2).

Exemple 3.2 : Synthèse du phénylméthyldiazométhane (PMDAM) :

Hydrazone de l'acétophénone : On dilue l'acétophénone (2,0 g, 16 mmole) dans 16 mL d'éthanol absolu puis on ajoute l'hydrazine (2,3 mL, 48 mmole). On porte au reflux. Après 2 h, on évapore le solvant et on reprend le résidu dans de l'éther (150 mL). On lave avec de l'eau (100mL). Après séchage sur Na_2SO_4, on évapore l'éther. On obtient l'huile jaune pâle (1,5 g, 11 mmole, 69 %).

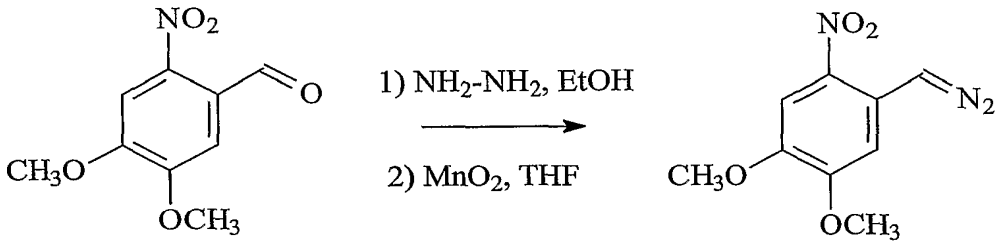
Phénylméthyldiazométhane (PMDAM) : On dissout l'hydrazone (150 mg, 1,1 mmole) dans 3 mL de THF. On ajoute le MnO_2 (480 mg, 5,5 mmole). On agite pendant 30 min à température ambiante. Le milieu se colore en rouge. On filtre et on évapore le solvant. On obtient l'huile rouge (145 mg, 100 %). Ce réactif est utilisé sans purification.

Exemple 3.3 : Synthèse du diphényldiazométhane (DPDAM) :

La Benzophénone hydrazone est commerciale (référence B 960-2 Aldrich, France).

On travaille donc sur 196 mg (1,0 mmole) que l'on dissout dans 5 mL de THF. 435 mg (5 éq, 5,0 mmole) de MnO_2 sont additionnés, le mélange est laissé sous agitation pendant 10 minutes, puis filtré. Le filtrat récupéré est évaporé à sec. On obtient 193 mg (0,99 mmole). Ce réactif est utilisé sans purification.

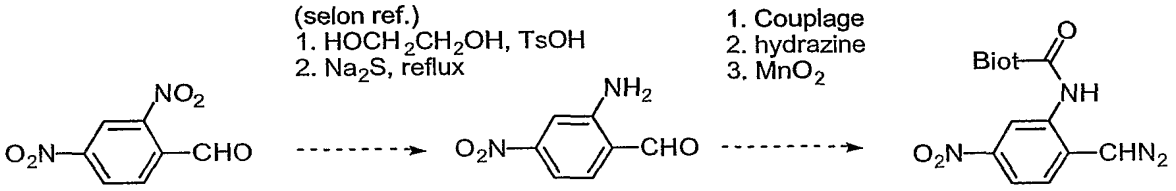
Exemple 3.4 : Synthèse du NVDAM :



NVDAM

La synthèse est réalisée selon le protocole décrit précédemment à partir de 6-nitrovératraldéhyde (Aldrich, référence 27,960-9).

Exemple 4 : Synthèse du dérivé biotinylé à partir du NPDAM :



M. E. Wall et al., *J. Med. Chem.*, **1993**, 36, 2689.

Le dérivé 2-amino-4-nitro benzaldéhyde est préparé selon la méthode de ME Wall et al référencé ci-dessus.

La préparation du diazométhane NPDAM est identique à celle décrite dans l'exemple 3.1 ci-dessus.

Exemple 5 : Préparation des acides nucléiques ADN :

Les amplicons ADN sont générés par PCR à partir de cibles d'ADN génomique *Mycobacterium tuberculosis* 16S (10^{+4} copies comme cibles de départ) en utilisant le kit Fast

Start de Roche, 0,2 mM de chaque désoxyribonucléotide (d-ATP, d-CTP, d-GTP, d-TTP), 0.3 µM d'amorces et 0,4 µL d'enzyme.

Les paramètres de la PCR sont les suivants :

- 95°C : 4 min puis 35 cycles (95°C : 30 sec ; 55°C : 30 sec ; 72°C : 30 sec) puis 4°C .

Les amplicons sont analysés qualitativement par électrophorèse sur gel d'agarose (1,5%, TBE 0,5X). Le volume déposé est de 5 µL et la migration s'effectue durant 20 min à 100 Volts (V). La visualisation des produits PCR est réalisée sous lampe UV après coloration au bromure d'éthidium.

Les conditions pour la culture, l'extraction des Mycobactéries ainsi que les amorces d'amplification sont données dans la demande de brevet WO-A-99/65926.

Exemple 6 : Réactivité des réactifs de marquage sur des nucléotides modèles

La synthèse des réactifs de marquage est décrite dans les exemples 1 à 4 ci-dessus. Le PDAM décrit dans la présente invention est commercialement disponible. (référence P1405, Molecular Probes, Eugene, OR)

Exemple 6.1 : Marquage des monomères UMP 3'-phosphate :

La réactivité des réactifs de marquage portant une fonction diazométhyle a été étudiée pour contrôler la spécificité de la réaction.

Un protocole a été mis au point qui consiste à étudier cette réactivité par électrophorèse capillaire sur un composé modèle le 3'-UMP (Uridine 3'monophosphate, référence U1126 Sigma) dans les conditions standards suivantes :

3'-UMP 0,04 mM ; H_3BO_3 2,0 mM ; Marqueur 2,0 mM [ajouté avec un solvant organique approprié (THF, AcOEt ou DMSO)] ; Solvants H_2O - CH_3CN -Solvant organique (rapport : 1/3/1).

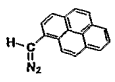
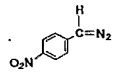
On répartit cette solution en dix fractions de 250 µL que l'on chauffe à 60 °C. Après une durée définie, chaque fraction est traitée avec 250 µL de dichlorométhane. Après agitation, on élimine la phase organique (phase inférieure). On répète cette opération encore deux fois. Après centrifugation (5 min, 5000 tours par minutes (rpm)), la phase aqueuse est analysée par électrophorèse capillaire. Les conditions de l'électrophorèse capillaire (EC) sont les suivantes : l'analyse EC a été effectuée par l'appareil Beckman P/ACE 5000. Un capillaire de silice fondue non traité (75 µm × 50 cm) a été utilisé. Le voltage appliqué est de 30 kV (polarité normale) et la température du capillaire a été maintenue à 23°C. Les électrophorégrammes ont été enregistrés à 254 nm. La solution tampon borate (0,1 M, pH 8,3) a été préparée à partir de l'acide borique en ajustant le pH avec une solution NaOH et filtrée à travers un filtre 0,2 µm. Les échantillons ont été injectés sous pression (0,5 psi, 5 secondes). Chaque analyse est précédée de la régénération du capillaire par passages successifs d'une solution de NaOH (0,1 M, 2 min), eau (2 min) et de tampon borate (2 min) sous pression (20 psi).

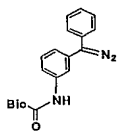
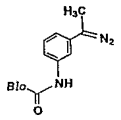
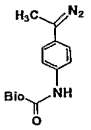
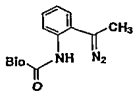
La réaction est effectuée en faisant varier le temps de réaction entre 0 et 4 heures, comme indiqué sur chaque figure, et les résultats sont présentés pour chaque réactif testé sur les figures 2A à 2I avec la concentration en réactif utilisé.

Le temps de réaction est indiqué sur chaque électrophorégramme.

Avec tous les réactifs testés, on obtient la formation exclusive d'un produit monoalkylé, ce qui prouve la spécificité de la réaction.

La réactivité, c'est-à-dire le temps de demi réaction du réactif sur le 3'-UMP, peut être ainsi calculée, par comparaison de la hauteur des pics, et les résultats sont présentés dans le tableau 1 ci-dessous (conditions standards décrites ci-dessus) :

Formules chimiques et noms	 PDAM	 DPDAM	 PMDAM	 NPDAM
Réactivité	5 min	20 h	30 min	4 h

Formules chimiques et noms	 BioDPDAM	 *m*-BioPMDAM	 *p*-BioPMDAM	 *o*-BioPMDAM *
Réactivité	10 h	15 min	5 min	1 min

* Avec le *o*-BioPMDAM, la réactivité est estimée car dans les conditions de concentration de 2 mM en réactif de marquage, la réaction est complète en moins de 5 min. La figure 2I utilise donc une concentration de 0,2 mM.

Tableau 1 : Réactivité (temps de demi réaction) de réactifs

On peut néanmoins noter que, la réaction étant très spécifique et ne conduisant pas à des sous produits pour tous les réactifs testés, il est possible d'augmenter la concentration du réactif de marquage sans conséquence du point de vue de la sélectivité sur le marquage.
Ainsi pour le réactif DPDAM, si l'on augmente la concentration à 20 mM (voir figure 2C), la réactivité (temps de demi réaction) est de 2 heures. Le même résultat est obtenu avec le BioDPDAM (figure 2F) où la réactivité est de 45 min à une concentration de 30 mM.

Exemple 6.2 : Test de différents nucléotides 3'-monophosphate :

De manière à éviter toute erreur d'interprétation, une étude complémentaire sur le marqueur *méta*-BioPMDAM, pris à titre d'exemple significatif, a été effectuée avec les autres nucléotides 3'-monophosphate.

Les nucléotides testés sont les suivants : 3'-AMP (référence 85,194-9, Aldrich), 3'-GMP (référence 151214, ICN), 3'-CMP (référence C1133, Sigma), 3'-TMP (série désoxyribo) (référence T1008, Sigma). Les électrophorégrammes obtenus avec les différents nucléotides sont représentés sur les figures 3A à 3D. Les temps de réaction indiqués sur la figure 3A sont identiques pour les figures 3B à 3D.

Quel que soit le nucléotide de départ (série ribo ou désoxyribo), nous observons la formation exclusive et complète du produit alkylé en 130 min à 60°C. Il est important de noter que dans

le cas de la guanine (base la plus réactive avec les réactifs alkylants usuels), seul le produit alkylé au phosphate est observé prouvant la très grande sélectivité de la réaction.
Cette étude permet également de vérifier que la vitesse de la réaction ne dépend pas de la nature de nucléotide en tant que substrat.

Exemple 6.3 : Etude d'un dinucléotide :

L'alkylation du dinucléotide ApUp (référence A4298, Sigma) a été effectuée avec le *méta*-BioPMDAM afin de vérifier la sélectivité de la réaction vis-à-vis du phosphate terminal par rapport au phosphate internucléotidique. Le suivi de la réaction par électrophorèse est représenté sur la figure 4. Les conditions sont les conditions standards de l'exemple 6.1 déjà décrit.

La formation exclusive d'un seul produit est observée montrant une bonne sélectivité du réactif *méta*-BioPMDAM vis-à-vis du phosphate terminal par rapport au phosphate internucléotidique.

Exmple 6.4 : Caractérisation des adduits avec les quatre (4) nucléotides 3'-monophosphates :

Pour s'assurer que les produits obtenus provenaient bien d'une alkylation au phosphate, la synthèse des adduits des monophosphates 3'-UMP, 3'-CMP, 3'-GMP et 3'-AMP avec le réactif *méta*-BioPMDAM a été effectuée. La réaction d'alkylation est réalisée à l'échelle préparative comme indiqué ci-dessous. Les adduits, obtenus avec des rendements de l'ordre de 70 % sont purifiés puis étudiés par RMN du proton et du phosphore.

Protocole de préparation :

On dissout le 3'-UMP (sous forme de sel disodique ; 9,3 mg ; 21,1 µmol) dans 2 mL d'une solution aqueuse 0,1 M de H_3BO_3, puis on ajoute successivement 2 mL de CH_3CN, 6 mL de MeOH puis le réactif *méta*-BioPMDAM (75 mg ; 0,20 mmol). La réaction est réalisée pendant

2,5 h à température ambiante. Elle est suivie par électrophorèse capillaire. On ajoute 3 mL d'eau, puis l'excès du réactif est éliminé par extraction avec CH_2Cl_2. La phase aqueuse est évaporée. Le résidu est dissous dans une petite quantité d'eau et purifié par passage sur une colonne de gel de silice en phase inverse (Lichroprep RP-18, Merck ; élution MeOH/H_2O (20/80)). On obtient 10 mg (69 %) de l'adduit du 3'-UMP.

Les spectres de RMN du proton obtenus pour les adduits du 3'-NMP (N= G, U, C, A) sont présentés sur les figures 5A à 5D. L'identification des adduits a été réalisé par des expériences de RMN à deux dimensions $^1H/^1H$ (COSY). Deux diastéréoisomères pour chacun de ces adduits dans un rapport 1/1 sont présents.

Un seul pic est présent pour la RMN du phosphore vers 0 ppm (300 MHz, D_2O).

Ces expériences démontrent que la réaction est bien spécifique, qu'un seul adduit est observé et que le marquage a bien lieu sur le phosphore. Il n'y a pas de réaction secondaire d'alkylation sur les bases. Les produits de marquage sont donc particulièrement adaptés pour une étape d'hybridation.

Exemple 7 : Etude de stabilité à la température :

Tous les dérivés du diazométhane, décrit dans le tableau 1 de l'exemple 6.1 ci-dessus, sont conservé à l'état solide dans le congélateur à -20 °C pendant au moins trois (3) mois et aucune perte de réactivité n'est observée.

La stabilité à température ambiante sur la paillasse a été déterminée par RMN ^1H pour les deux réactifs NPDAM et *méta*-BioPMDAM. Nous n'avons observé aucune décomposition en laissant le NPDAM pendant un mois sur la paillasse sans précaution particulière. Nous observons environ 50% de décomposition en laissant le *méta*-BioPMDAM sur la paillasse pendant vingt-cinq (25) jours.

La stabilité à la température du réactif de marquage est une caractéristique essentielle. En effet, la destination finale d'un tel réactif pour une application industrielle est un kit de marquage. Un réactif qui n'est pas stable au moins pendant quinze (15) jours à -20°C, de préférence un (1) mois est invendable. Même s'il existe des moyens de stockage et d'envoi

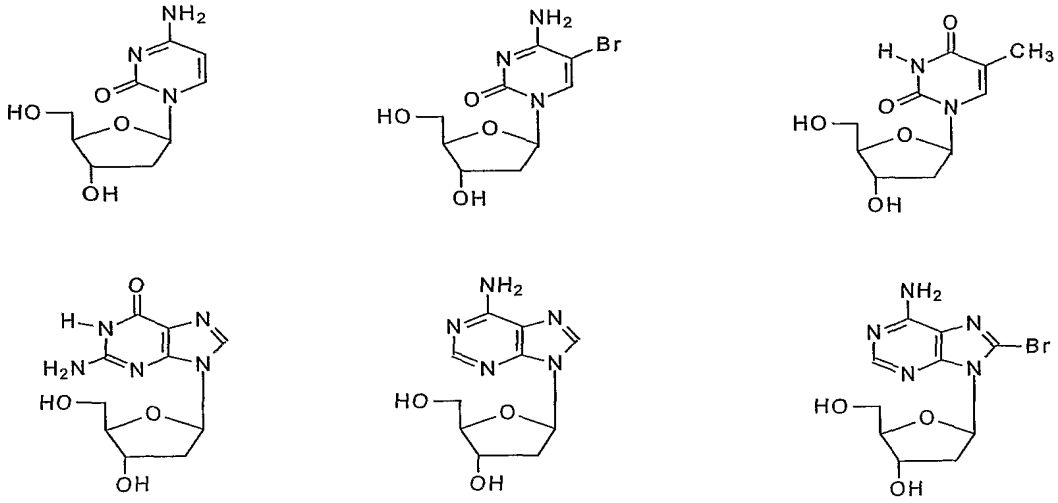
jusqu'à -110°C, il existe une relation entre la stabilité à -110°C et à -20°C et c'est pourquoi la valeur de quinze (15) jours à -20°C, de préférence un (1) mois à -20°C est un minimum industriel. Au delà de -110°C, les laboratoires ne disposent pas des équipements nécessaires (congélateur) pour stocker ces réactifs du point de vue de l'utilisateur et du fabricant, et il n'y a pas de moyen simple type carboglace pour les expédier du point de vue du fabricant.

En ce qui concerne la stabilité à température ambiante, une stabilité de quelques heures, de préférence de un (1) jour, est suffisante pour permettre à l'utilisateur d'effectuer le marquage.

Exemple 8 : Etude de la fragmentation de l'ADN :

Exemple 8.1 : Hydrolyse de différents nucléosides en milieu acide :

Le but de cette étude est de démontrer la différence en terme de stabilité à pH acide entre les nucléosides naturels, les nucléosides modifiés ainsi que les nucléosides de type puriniques et pyrimidiniques. Cette étude permet également de mieux contrôler la fragmentation de l'ADN en tenant compte de sa composition en bases.



Deux nucléosides modifiés, la 8-bromo-2'-désoxyadénosine (8-BrdA) et la 5-bromo-2'-désoxycytidine (5-BrdC) ainsi que les quatre (4) nucléosides naturels (dA, dC, dG et dT) ont été utilisés dans cette étude.

50 nanomoles (nmol) de chaque nucléoside sont incubées dans 50 mM de formiate de sodium pH 3 à 95°C. Les temps d'incubation varient de 0 à 45 min. Après séchage sous vide et reprise par 20 µl d'eau pure, les échantillons (10 nmol) sont alors analysés par HPLC en phase inverse. Les résultats sont donnés sous forme de pourcentage d'hydrolyse du nucléoside de départ (en ordonnée) par rapport au temps d'incubation en minutes (en abscisse) voir figure 8.

Les courbes de la figure 8 montrent que la modification de l'adénine en position 8 par un atome de brome rend ce nucléoside moins stable que le nucléoside naturel. D'autre part, les résultats montrent que, dans les conditions utilisées, la dépurination est beaucoup plus importante que la dépyrimidination.

Cette étude montre que la fragmentation de l'ADN par la dépurination ou la dépyrimidination peut être contrôlée en optimisant les conditions d'hydrolyse, ou en incorporant soit des bases modifiées moins stables que les bases naturelles soit des bases pouvant être modifiées et hydrolysées après leur incorporation.

Exemple 8.2 : Fragmentation de l'ADN double brin incorporant ou non un nucléotide modifié :

Trois amplifications PCR ont été réalisées en parallèle à partir de cible ADN génomique *Mycobacterium tuberculosis* 16S (10^{+4} copies de départ) en utilisant le kit *Fast Start* de Roche, 0,2 mM de chaque désoxyribonucléotide (d-ATP, d-CTP, d-GTP, d-TTP), 0,3 µM d'amorces et 0,4 µL d'enzyme.

Les paramètres de la PCR sont ceux de l'exemple 5.

Dans le premier cas, le protocole est utilisé tel quel : cas de la PCR dite naturelle.

Dans le deuxième cas, le protocole est modifié pour obtenir une PCR à 30 % de 8Br-dATP. Ceci est réalisé en introduisant 0,2 mM de d-CTP, d-GTP et d-TTP. On introduit également 0,14 mM de d-ATP et 0,06 mM de 8-BrdATP. (La 8-BrdATP est commerciale (référence N-2005-1, TriLink Biotechnologies, San Diego CA).

Dans le troisième cas, le protocole est modifié pour obtenir une PCR à 50 % de 8-BrdATP. Ceci est réalisé en introduisant 0,2 mM de d-CTP, d-GTP et d-TTP. On introduit également 0,1 mM de d-ATP et 0,1 mM de 8-BrdATP.

L'étude de la fragmentation seule de ces amplicons a été réalisée dans les conditions décrites ci-dessus : 50 mM de formiate de sodium pH 3 à 95°C.
L'analyse a été effectuée sur gel dénaturant de polyacrylamide (8 % polyacrylamide, 7 M urée, 1 X TBE) utilisant coloration au bromure d'éthidium.

Après 15 min d'incubation à 95 °C dans 50 mM de formiate de sodium pH 3, aucune différence n'est visible entre les trois (3) cibles. Dans tous les cas, nous avons observé une fragmentation totale des amplicons PCR.

La dépurination des amplicons PCR a été également réalisée à différents pH et à différentes températures et temps d'incubation. L'analyse sur gel, dans les conditions ci-dessus, montre qu'à pH 3, la fragmentation est totale après seulement 10 min d'incubation à 95°C. A ce pH, la fragmentation est aussi totale à 60°C après 30 min d'incubation.
A pH 4, 30 min d'incubation sont nécessaires pour l'obtention d'une fragmentation totale des amplicons ADN même à 95°C. Ce résultat est très intéressant et montre que le site abasique généré à pH acide est instable et donc conduit à la fragmentation de la chaîne ADN sans aucun autre traitement particulier.

Exemple 9 : Marquage et fragmentation de l'ADN par le dérivé *méta*-bioPMDAM (3a) en deux étapes :

Le dérivé *méta*-bioPMDAM (3a) a été obtenu selon le schéma réactionnel décrit dans l'exemple 1.1.
Les amplicons ADN ont été préparés par amplification PCR selon le protocole décrit dans l'exemple 5.

Marquage

A 10 µl de PCR, 38 µL d'eau pure (sigma), 50 µL de formiate de sodium à pH 3 (100 mM dans l'eau pure) sont ajoutés et le mélange est incubé pendant 30 minutes à 60°C. Puis 2 µL de *méta*-bioPMDAM (100 mM dans DMSO) sont ajoutés ensuite. La solution a été mélangée par vortex, puis incubée 15 minutes supplémentaires à 60°C.

Les essais sont réalisés en duplicata afin de pouvoir analyser la fragmentation de l'ADN sur gel et l'efficacité de marquage par hybridation et lecture de la puce à ADN.

Purification

La purification est réalisée sur colonnes QIAQUICK™ (Nucleotide Removal kit, Qiagen, Hilden, Allemagne). Le protocole de purification utilisé est celui recommandé par le fournisseur.

Après purification, l'éluat est transféré dans un tube propre contenant 400 µL de tampon d'hybridation (1,75 mL 20X SSPE ; 2,9 mL Bétaine 5M ; 290 µL DTAB 0,1M ; 10 µL Anti-mousse 30%) sont ajoutés. Les références de ces substances sont, pour :

- la bétaine référence B-2754 Sigma, et
- la DTAB référence D-5047 Sigma.

La solution est mélangée par vortex et incubée 10 min à 95°C, afin de séparer les brins d'ADN qui ne sont pas séparés lors de l'étape de marquage pendant la fragmentation (étape de dénaturation). Le tube est ensuite plongé dans un mélange eau-glace à 0°C avant hybridation sur puce à ADN.

Hybridation sur puce à ADN

Après l'étape de marquage pendant la fragmentation, les fragments obtenus sont hybridés sur les puces à ADN conçues pour l'analyse de la région 213-415 de la séquence M20940 « Genbank » de l'ARN 16S de *Mycobacterium tuberculosis*. Cette puce à ADN est décrite dans A. Troesch et al, J. Clin Microbiol., 37(1), p49-55, 1999.

Les étapes d'hybridation ont été réalisées sur les stations fluidiques (Affymetrix, Santa Clara, CA) en utilisant le protocole d'hybridation et les tampons décrits dans A. Troesch et al, J. Clin

IDT Exhibit 1020.7761

Microbiol., 37(1), p49-55, 1999. Une étape supplémentaire est nécessaire pour révéler la biotine (détection indirecte).

L'hybridation est révélée par le couplage de la streptavidine marquée à la phycoérythrine (PE) qui interagit avec la biotine de la *méta*-BioPMDAM dans les conditions suivantes : 300 µL d'eau pure ; 300 µL de tampon Tris 100 mM pH 7 / NaCl 1M / tween 0,05% / Antimousse 0,005% ; 6 µL de BSA (50 mg/mL) ; 6 µL de streptavidine-PE (300 µg / mL). Les références de ces substances sont, pour :

- la Streptavidine-Phycoérythrine : référence R0438, Dako, Danemark,
- la Streptavidine-CY5 : référence C0050, Dako, Danemark,
- l'Antimousse référence M5-575, Ultra Additives Inc., et
- le Tween référence P-7949, Sigma.

Lecture de la puce à ADN :

La lecture de la fluorescence émise à la surface de la puce à ADN après marquage et hybridation ainsi que la génération des données en termes d'intensité du signal et du pourcentage d'homologie sont réalisés par les systèmes de lecture et le logiciel fournis par Affymetrix (GeneChip® Instrument System et GeneChip® Information System, Santa Clara CA).

Le système de lecture fournit des intensités en signal et bruit de fond exprimées en rfu (« relative fluorescence unit »). Le pourcentage d'homologie est donné par rapport à une séquence référence qui dans ce cas est la séquence de *Mycobacterium tuberculosis*.

Les résultats en termes d'intensité moyenne du signal (I), du bruit de fond (B) et du pourcentage d'homologie (%) sont donnés dans le tableau 2 ci-dessous :

De manière générale, on considère qu'un pourcentage d'homologie supérieur à 90% est un résultat satisfaisant bien que l'on cherche en général un résultat supérieur à 95%. Au dessus de 95%, les valeurs ne sont plus indiquées car elles ne sont pas significatives dans le cas de la puce à ADN Mycobactéries. Une intensité élevée avec un bruit de fond faible est le deuxième résultat recherché dans les exemples qui suivent. Dans tous les résultats, le bruit de fond B est déduit de l'intensité moyenne I.

Analyse sur gel de polyacrylamide

Les échantillons destinés à être analysés sur gel sont séchés sous vide, repris par 10 µL d'eau pure et 10 µL de bleu formamide 2X.

La migration est effectuée sur gel d'acrylamide 8%, dans du TBE 1X, une (1) heure à 150 V.

Le pH acide a été utilisé pour la fragmentation de l'ADN. En effet à ce pH, le phénomène de dépurination engendre des sites abasiques très instables conduisant à une fragmentation presque immédiate des séquences ADN à température élevée. Ce type de fragmentation produit des fragments ADN-5' phosphate.

L'analyse sur gel montre que l'incubation des amplicons PCR à 60°C pendant 30 min en solution dans le tampons formiate (50 mM, pH 3) conduit à une fragmentation totale de ces amplicons. Ceci nous a permis d'évaluer le marquage pendant la fragmentation des amplicons ADN en présence du marqueur *méta*-bioPMDAM.

Les résultats de marquage pendant la fragmentation des amplicons ADN en termes de pourcentage d'homologie, d'intensité des signaux et du bruit de fond sont donnés dans le tableau 2 ci-dessous.

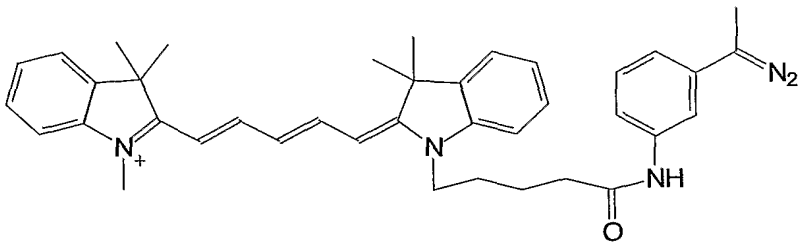
Conditions de marquage des amplicons PCR	**Homologie (%)**	**I (rfu)**	**B (rfu)**	**I/B**
Tampon : formiate pH3, 50 mM Marqueur : 2 mM *méta*-bioPMDAM Incubation : 30 min à 60°C	> 95	4456	593	7,5

Tableau 2 : Marquage et fragmentation des amplicons ADN en termes d'homologie, d'intensité des signaux (I) et du bruit de fond (B)

Cet exemple montre que les dérivés de l'invention peuvent être utilisés pour le marquage des fragments ADN produits par amplification enzymatique dans un protocole en deux étapes. Ils peuvent être également utilisés pour le marquage de l'ADN naturel non amplifié.

Exemple 10 : Marquage de l'ADN par le dérivé Cy5-PMDAM (12bis) :

Le marquage de l'ADN par ce nouveau marqueur portant le fonction diazométhyle a été évalué en utilisant un fragment ADN synthétique.



Cy5-PMDAM (12bis)

Le réactif de marquage Cy5-PMDAM (12bis) est préparé selon le protocole décrit dans l'exemple 2.

Un oligodésoxyribonucléotide (ODN) vingt (20) mers est préparé selon la méthode dite au phosphoramidite. On introduit un phosphate à l'extrémité 5' par un réactif de phosphorylation standard compatible avec la chimie au phosphoramidite. La séquence de cet ODN est constituée de toutes les bases naturelles de l'ADN (séquence de l'ODN : 5'-CTGAACGGTAGCATCTTGAC-3'). Cette séquence est complémentaire de séquences de capture d'une puce à ADN dite « modèle », synthétisée selon la technologie Affymetrix. Cette puce à ADN contient des sondes de capture identiques en séquence et qui sont réparties en damier sur sa surface. La lecture de cette puce à ADN donne des informations quant à la performance du marquage en terme d'intensité mais pas de résultat d'homologie.

Marquage : A 50 picomoles (pmoles) de cet ODN, 10 µL de Cy5-PMDAM (100 mM dans le DMSO) sont ajoutés. Le volume final est de 100 µL. Après homogénéisation, l'incubation est réalisée à 60°C pendant 30 minutes.

Purification et lecture : La purification pour éliminer l'excès de réactif de marquage est réalisée selon l'exemple 9. La lecture sur puce à ADN est effectuée selon l'exemple 17.

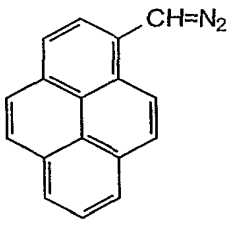
Résultats :

La moyenne des intensités de marquage (I) lue sur la puce à ADN est de 16 644 rfu pour un bruit de fond (B) de 450 rfu.

Ce niveau d'intensité est très élevé et montre que le réactif de marquage Cy5-PMDAM (12bis) est tout à fait compatible avec le marquage des fragments d'ADN sur le groupement phosphate.

Exemple 11 : Marquage et fragmentation de l'ADN par le réactif PDAM :

Exemple 11.1 : Marquage par le 1-Pyrényldiazométhane (PDAM) :



1-Pyrényldiazométhane

Le PDAM est obtenu chez Molecular Probes (Eugene, OR) et solubilisé dans du DMSO anhydre.

Deux ODN de vingt (20) mers ont été utilisés comme modèles ADN : un ODN de 20mer 5'-hydroxyle et le même ODN de 20mer portant un phosphate à l'extrémité 5'. La séquence de l'ODN est décrite dans l'exemple 10. La réaction de marquage est réalisée dans un mélange à 50 % en DMSO et 1,5 mM en 1-Pyrényldiazométhane (PDAM) à 60°C pendant 30 minutes ou une heure.

L'efficacité du marquage a été évaluée par chromatographie sur couche mince (en phase normale) dans un éluant isopropanol/ammoniaque/eau 60/10/30. Après 30 minutes, le couplage est total sur l'ODN 5'-phosphate. Il faut une heure pour obtenir un couplage partiel sur l'ODN 5'-hydroxyle c'est à dire environ 50%.

Les résultats de l'exemple 6 sont confirmés sur une séquence modèle de 20 bases en ce qui concerne le marquage préférentiel des réactifs portant une fonction diazométhyle sur le phosphate terminal. Le marquage sur un phosphate intranucléotidique n'est pas un inconvénient rédhibitoire puisque cela peut conduire à une augmentation de la sensibilité par introduction de plus d'un marqueur sur le fragment d'acide nucléique. Cela permet à l'acide nucléique de s'hybrider avec une bonne sensibilité sur la cible complémentaire tout en conservant une bonne spécificité d'hybridation.

L'homme du métier par des réactions d'optimisation peut ainsi contrôler la spécificité du marquage, en jouant par exemple sur le réactif de marquage, le temps de réaction et la température, pour avoir un marquage exclusif sur le phosphate terminal.

Exemple 11.2 : Etude cinétique de la réaction de marquage par le PDAM :

Cette étude a été réalisée en utilisant l'ODN 20mer 5'-phosphate dans les conditions précédentes en faisant varier le temps de réaction. Les rendements de marquage ont été évalués par analyse chromatographie liquide haute pression (CLHP) en phase inverse dans les conditions suivantes :

Phase inverse colonne Spheri-5 RP-18 5 µm, 220x 4,6 mm (Perkin Elmer). Les tampons et le gradient utilisés sont :

- Tampon A : 0,1 M TEAA ; Tampon B = 50% Tampon A + 50 % CH_3CN, et
- Gradient de 10 à 50 % de B en 30 min à 1 ml/min à température ambiante.

Les résultats sont représentés sur la figure 9 avec en abscisse le temps de réaction, exprimé en minutes, et en ordonnée le pourcentage de marquage.

Le rendement est proche de 90% après seulement 10 min d'incubation à 60°C.

Exemple 11.3 : Effet de la température sur le marquage au PDAM :

Le marquage a été réalisée en utilisant l'ODN 20mer 5'-phosphate dans les conditions précédentes en faisant varier la température d'incubation et avec un temps d'incubation de 10 minutes dans chaque cas.

Les rendements de marquage ont été évalués par analyse CLHP en phase inverse.
Les résultats sont présentés sur la figure 10 avec en ordonnée le pourcentage de marquage et en abscisse la température de réaction en °C
Il est très important de noter que même à température ambiante (25°C), nous observons un marquage de l'ODN. Après 10 minutes d'incubation à 25 °C, le rendement de marquage est environ de 25%. A des températures supérieures à 50°C, des rendements supérieurs à 80% sont obtenus.
Ceci montre l'efficacité et la flexibilité de cette chimie de marquage de l'ADN par les réactifs portant une fonction diazométhyle.

Exemple 12 : Marquage et fragmentation des amplicons ADN obtenus par amplification PCR avec le réactif de marquage *méta*-bioPMDAM (3a) :

Le dérivé *méta*-bioPMDAM (3a) a été obtenu selon le schéma réactionnel décrit dans l'exemple 1.1.
Les amplicons ADN ont été préparés par amplification PCR selon le protocole décrit dans l'exemple 5.

Exemple 12.1 : Comparaison du marquage avec et sans fragmentation :

a. Marquage dans les conditions de fragmentation :
A 10 µL de PCR sont ajoutés 50 µL de formiate de sodium pH 3 (50 mM) et 2 µL de *méta*-BioPMDAM 100 mM dans du DMSO). Le volume est ajusté à 100 µL. La solution est incubée 30 min à 60°C.

b. Marquage sans fragmentation :
A 10 µL de PCR sont ajoutés 2 µL de *méta*-BioPMDAM (100 mM dans du DMSO). Le volume est ajusté à 100 µL. La solution est incubée 30 min à 60°C.

Le reste du protocole est identique à celui de l'exemple 9.

Résultats

Protocole utilisé	Homologie (%)	I (rfu)	B (rfu)	I/B
a. Marquage dans les conditions de fragmentation	>95	3995	569	7,0
b. Marquage sans fragmentation	94,1	500	542	0,9

Tableau 3 : Comparaison du marquage avec et sans fragmentation

Les résultats dans le tableau 3 ci-dessus montre que sans fragmentation, la moyenne des intensités obtenues est au même niveau que le bruit de fond (500 rfu). Le marquage pendant la fragmentation donne un niveau d'intensité bien supérieur (environ 4000 rfu) et un très bon pourcentage d'homologie. La combinaison des deux étapes représente donc bien une amélioration significative pour la détection d'un acide nucléique de longueur supérieure à cent nucléotides.

Exemple 12.2 : Effet de la dénaturation avant hybridation sur puce à ADN :

Deux réactions de marquage ont été réalisées en parallèle dans deux tubes séparés selon le protocole suivant : A 10 µL de PCR sont ajoutés 50 µL de tampon formiate de sodium pH 3 (50 mM) et 2 µL de *méta-bio*PMDAM (100mM dans du DMSO). Le volume total est ajusté à 100 µL et incubé 30 min à 60°C.

Après purification sur colonne (exemple 9), la solution issue du premier tube est incubée 10 min à 95°C (afin de désapparier le double brin d'ADN), puis le tube est plongé dans un mélange eau-glace à 0°C jusqu'à l'hybridation sur la puce ADN.

La solution issue du deuxième tube est hybridée sur la puce à ADN sans dénaturation préalable.

Les fragments biotinylés hybridés sur les sondes de capture à la surface de la puce à ADN sont révélés par l'introduction d'une streptavidine marquée par la phycoérythrine en utilisant les conditions décrites dans l'exemple 9.

Résultats

Conditions utilisées	Homologie (%)	I (rfu)	B (rfu)	I/B
Avec dénaturation	>95	22812	570	40,1
Sans dénaturation	93,5	4795	681	7,0

Tableau 4 : Effet de la dénaturation avant hybridation sur puce à ADN

Les résultats obtenus, présentés dans le tableau 4, avec la dénaturation pré-hybridation sont supérieurs à ceux obtenus sans l'étape de dénaturation. Ceci montre que la dénaturation de l'ADN est nécessaire pour obtenir un bon niveau d'intensité. La fragmentation via les sites abasiques est un moyen pour faciliter la dénaturation d'un ADN double brin et renforcer l'hybridation sur les sondes de captures.

Pour tester d'autres réactifs de marquage et compte tenu des résultats obtenus avec les différentes conditions ci-dessus, nous avons défini un protocole de référence utilisant la fragmentation au tampon formiate de sodium (50 mM, pH 3) et un étape de dénaturation pré-hybridation.

Exemple 13 : Marquage et fragmentation des amplicons PCR par les réactifs biotinylés dans un protocole en une étape :

Les dérivés *méta-*, *ortho-*, et *para*-bioPMDAM ont été préparés selon le protocole décrit dans les exemples 1.1, 1.2 et 1.3. Ils ont été solubilisés dans du DMSO anhydre à une concentration de 100 mM.

Le protocole est identique à celui de l'exemple 12.2 ci-dessus (marquage et fragmentation en une étape puis étape de dénaturation pré-hybridation).

Résultats

Marqueur	Homologie (%)	I (rfu)	B (rfu)	I/B
ortho-bioPMDAM	>95	25951	820	31,6
méta-bioPMDAM	>95	22960	581	39,5
para-bioPMDAM	94,1	43785	1205	36,3

Tableau 5 : Marquage et fragmentation des amplicons PCR par les réactifs biotinylés dans un protocole en une étape

Le protocole optimisé avec fragmentation et marquage en une étape donne d'excellents résultats avec différents réactifs de marquage avec la fonction réactive diazométhyle, comme le montrent les résultats exposés dans le tableau 5.

Exemple 14 : Marquage et fragmentation de l'ADN par le BioDPDAM :

Les amplicons ADN ont été préparés par amplification PCR selon le protocole décrit dans l'exemple 5.

La synthèse du réactif de marquage est décrite dans l'exemple 1.

Le protocole est identique à celui de l'exemple 12.2, y compris la dénaturation à 95 °C avant l'étape d'hybridation.

Résultats

Marqueur utilisé	Homologie (%)	I (rfu)	B (rfu)	I/B
BioDPDAM	93,0	32359	3610	9,1

Tableau 6 : Marquage et fragmentation de l'ADN par le BioDPDAM

Ce résultat, décrit dans le tableau 6, montre que des substitutions aussi importantes que le groupement phényle peuvent être utilisées pour optimiser la réactivité des réactifs de marquage portant une fonction diazométhyle.

Exemple 15 : Marquage et fragmentation de l'ADN par la 5-(bromométhyl)fluorescéine :

Les amplicons ADN ont été préparés par amplification PCR selon le protocole décrit dans l'exemple 5.

A 10 µL de PCR sont ajoutés 50 µL de formiate de sodium pH3 (100 mM) et 2 µL de 5-(bromométhyl)fluorescéine (Molecular probes, Eugen, OR) (100 mM dans du DMSO). Le volume est ajusté à 100 µL. La solution est incubée 30 min à 60°C.

Les conditions de purification sont conformes à celles de l'exemple 9. Une étape de dénaturation est effectuée comme décrit dans l'exemple 12.2.

Les autres conditions d'hybridation et de lecture sont identiques à celles décrites dans l'article de A. Troesch et al. J. Clin. Microbiol., 37(1), p 49-55 1999. La fluorescéine est directement détectable sur le lecteur.

Protocole utilisé	Homologie (%)	I (rfu)	B (rfu)	I/B
Marquage et fragmentation des amplicons PCR avec la 5-(bromo-méthyl)fluorescéine	>95	855	183	4,7

Tableau 7 : Marquage et fragmentation de l'ADN par la 5-(bromométhyl)fluorescéine

Ce résultat du tableau 7 montre que la fragmentation de l'ADN par la création de sites abasiques est tout à fait compatible avec un réactif de marquage portant une fonction réactive halogénure d'alkyle. Ce protocole est réalisé en une étape (fragmentation et marquage) mais avec des intensités plus faibles qu'avec les marqueurs portant une fonction diazométhyle.

Exemple 16 : Marquage et fragmentation des amplicons ADN en présence d'un autre agent chimique de fragmentation dérivé de la phénanthroline :

Des amplicons ADN ont été préparés par amplification PCR selon le protocole décrit dans l'exemple 5. Deux types de conditions sont utilisées :

Conditions a :

A 10 µL de PCR sont ajoutés 20 µL de phénanthroline-$FeSO_4$ (25 mM) et 2 µL de *méta*-BioPMDAM (100 mM dans DMSO). Le volume total est ajusté à 100 µL. Le mélange est incubé 60 min à 95°C.

Conditions b :

A 10 µL de PCR sont ajoutés 50 µL de tampon formiate de sodium pH3 (100 mM dans eau pure) et 2 µL de de *méta*-BioPMDAM (100 mM dans DMSO). Le volume total est ajusté à 100 µL. Le mélange est incubé 60 min à 95°C.

Les autres conditions du protocole sont identiques à celles de l'exemple 9.

Conditions	Homologie (%)	I (rfu)	B (rfu)	I/B
Phénanthroline FeSO4(5 mM) *méta*-bioPMDAM (2 mM) 60 min à 95°C	>95	2236	500	4,5
Tampon formiate (50mM) pH3 *méta*-bioPMDAM (2 mM) 60 min à 95°C	>95	6786	565	12,0

Tableau 8 : Marquage et fragmentation des amplicons ADN en présence de la phénanthroline

Les deux conditions de fragmentation permettent un résultat satisfaisant comme démontré dans le tableau 8.

Le meilleur résultat est obtenu avec les conditions (b) utilisant la fragmentation à pH acide.

Exemple 17 : Fragmentation des amplicons PCR marqués par incorporation de d-UTP-fluorescéine :

Incorporation du nucléotide marqué

Une amplification PCR a été réalisée selon le protocole suivant afin générer des amplicons PCR marqués par la fluorescéine (marquage sur les bases).

A partir de cible ADN génomique *Mycobacterium tuberculosis* 16S (10^{+4} copies) en utilisant le kit *Fast Start* de Roche, 0,2 mM des désoxyribonucléotides d-ATP, d-CTP et d-GTP ainsi que 0,14 mM de d-TTP et 0,06 mM de d-UTP-12-fluoresceine, 0,3 µM d'amorces et 0,4 µL d'enzyme. Le pourcentage du nucléotide marqué par rapport à son homologue naturel d-UTP est de 30%. C'est ce rapport qui est généralement utilisé dans les réactions de marquage des amplicons par incorporation de nucléotides marqués.

La d-UTP-12-fluoresceine est disponible commercialement chez Roche Diagnostics référence 1373242, Mannheim, Allemagne).

Les paramètres de la PCR sont ceux de l'exemple 5.

a. Fragmentation des amplicons PCR marqués à 30% par d-UTP-fluorescéine :

A 10 µL de PCR sont ajoutés 50 µL de tampon formiate de sodium pH 3 (50 mM). Le volume est ajusté à 100 µL. La solution est ensuite incubée 30 min à 60°C.

b. Marquage pendant la fragmentation des amplicons PCR contenant 30% par d-UTP-fluorescéine :

A 10 µL de PCR sont ajoutés 50 µL de tampon formiate de sodium pH 3 (50 mM) et 2 µL de *méta*-BioPMDAM (100 mM dans du DMSO). Le volume est ajusté à 100 µL. La solution est ensuite incubée 30 min à 60°C. Cet essai correspond à un protocole de référence qui permet de comparer les différentes stratégies de marquage en s'affranchissant de la variabilité due à l'étape d'amplification.

Une étape de purification sur colonne et une étape de dénaturation à 95°C sont effectuées dans tous les cas comme dans l'exemple 9.

Protocole (a1) :

Les acides nucléiques obtenus par fragmentation des amplicons marqués à la d-UTP-fluorescéine (conditions *a*)) sont hybridés sur puce à ADN et détectés dans un premier temps par lecture directe des signaux fluorescents émis par la fluorescéine comme décrit dans l'exemple 15.

Protocole (a2) :

Une étape d'amplification du signal a été utilisée pour améliorer la sensibilité de marquage. L'amplification du signal a été réalisée par l'introduction, lors de l'étape d'hybridation, d'un anticorps anti-fluorescéine biotinylé (référence 216-065-084, Jackson ImmunoResearch) puis d'une streptavidine marquée par la phycoérythrine en utilisant les conditions successives suivantes :

- 300 µL d'eau pure,

- 300 µL de tampon Tris 100 mM pH 7 / NaCl 1M / tween 0,05% / Antimousse 0,005% , 2,4 µL de BSA (50 mg/mL), 1,2 µL d'anticorps anti-fluorescéine biotinylé. (1 mg / mL),
- 300 µL d'eau pure, et
- 300 µL de tampon Tris 100 mM pH 7 / NaCl 1M / tween 0,05% / Antimousse 0,005% ; 6 µL de BSA (50 mg/mL) ; 6 µL de streptavidine-PE. (300 µg / mL).

Dans ce protocole, la fluorescéine agit comme un haptène (traceur détectable indirectement par un anticorps marqué) et non pas comme un fluorophore.

Protocole (b) :

Les fragments biotinylés (condition *b)* hybridés sur puce à ADN sont révélés par l'introduction d'une streptavidine marquée par la phycoérythrine en utilisant les conditions suivantes :

- 300 µL d'eau pure, et
- 300 µL de tampon Tris 100 mM pH 7 / NaCl 1M / tween 0,05% / Antimousse 0,005% ; 6 µL de BSA (50 mg/mL) ; 6 µL de streptavidine marquée. (300 µg / mL).

La lecture de la fluorescence émise à la surface de la puce à ADN après marquage et hybridation ainsi que la génération des données en termes d'intensité du signal et du pourcentage d'homologie sont réalisés par les systèmes de lecture et le logiciel fournis par Affymetrix. A ce propos, il est important de noter que le système de lecture utilisé contient deux filtres permettant de détecter directement :

- la fluorescéine dans le cas où les amplicons sont marqués au d-UTP-fluorescéine uniquement, selon le protocole *a1*, ou bien
- la phycoérythrine dans le cas où les amplicons sont marqués :
  - au d-UTP-fluorescéine avec amplification du signal, selon le protocole *a2*, ou
  - par le *méta*-bioPMDAM pendant leur fragmentation, selon le protocole *b*.

Dans les deux cas où la révélation s'effectue par la phycoérythrine, l'utilisation d'un filtre permet de s'affranchir du signal généré par la fluorescéine et c'est bien le signal de la phycoérythrine qui est détecté.

IDT Exhibit 1020.7775

Résultats

Protocole utilisé	Marqueur détecté	Homologie (%)	I (rfu)	B (rfu)	I/B
a1. Fragmentation des amplicons PCR marqués par la d-UTP-fluorescéine	Flu*	81,6	595	342	1,7
a2. Fragmentation des amplicons PCR marqués par la d-UTP-fluorescéine avec amplification du signal	PE*	>95	22107	3461	6,4
b. Fragmentation et marquage sur les mêmes amplicons PCR modifiés par la d-UTP-fluorescéine et marqués par la méta-BioPMDAM	PE*	>95	21700	1503	14,4

* Flu =Fluorescéine et PE = Phycoérythrine

Tableau 9 : Fragmentation des amplicons PCR marqués par incorporation de d-UTP-fluorescéine

Les résultats, du tableau 9 ci-dessus, montrent que la fragmentation chimique utilisant la création de site abasique est compatible avec le marquage enzymatique des amplicons ADN et que le marquage peut avoir lieu avant la fragmentation.

Les niveaux d'intensité ainsi que le pourcentage d'homologie obtenus avec ce protocole d'incorporation enzymatique du fluorophore sont faibles en comparaison de ceux obtenus avec le marquage pendant la fragmentation utilisant le réactif de marquage avec une fonction diazométhyle comme le *méta*-bioPMDAM (conditions (b)).

Pour atteindre le niveau d'intensité obtenue avec le dérivé *méta*-bioPMDAM, une étape d'amplification du signal est nécessaire (conditions a2). Ceci montre bien l'efficacité de la fonction réactive diazométhyle par rapport à l'incorporation traditionnelle de base modifiée comme le d-UTP fluorescéine (protocole de référence (b)).

Exemple 18 : Fragmentation de l'ADN double brin par sonication :

Des amplicons ADN ont été obtenus en utilisant le protocole décrit dans l'exemple 5. Ces amplicons ont été fragmentés par sonication en présence et en absence du marqueur.

a. Marquage des amplicons PCR pendant la sonication :

A 10 µL de réaction PCR sont ajoutés 2 µL de *méta*-BioPMDAM (100 mM dans DMSO). Le volume est ajusté à 100 µL par de l'eau pure le pH est ajusté à 6,5. Le mélange est incubé 30 min à 60°C dans un bain d'une cuve à ultrasons (fréquence 35kHz, modèle T460-H, Bioblock, France).

b. Marquage pendant la fragmentation chimique des amplicons PCR (protocole de référence en une étape) :

A 10 µL de réaction PCR sont ajoutés 50 µL de formiate de sodium pH 3 (50 mM) et 2 µL de *méta*-BioPMDAM (100 mM dans du DMSO). Le volume est ajusté à 100 µL. La solution est incubée 30 min à 60°C.

Les essais sont réalisés en duplicata afin de pouvoir analyser la fragmentation de l'ADN sur gel et l'efficacité de marquage par hybridation et lecture de la puce à ADN comme décrit précédemment (exemple 9 sur la détection de la phycoérythrine).

Analyse sur gel

L'analyse a été effectuée sur gel dénaturant de polyacrylamide (8 % polyacrylamide, 7 M urée, 1 X TBE) utilisant coloration au bromure d'éthidium.

L'analyse sur gel montre que les amplicons ADN sont fragmentés par sonication à 60°C.

Résultats

Conditions	Homologie (%)	I (rfu)	B (rfu)	I/B
a. Marquage pendant la fragmentation par sonication	93,8	2271	631	3,6
b. Marquage pendant la fragmentation chimique (conditions de référence)	>95	19639	1459	13,5

Tableau 10 : Fragmentation de l'ADN double brin par sonication

Les résultats de marquage du tableau 10 pendant la sonication (conditions *a*) sont satisfaisants. Ceci montre que la fragmentation physique par sonication des cibles ADN est compatible avec la chimie de marquage par des réactifs de marquage portant une fonction diazométhyle.

Les faibles résultats de marquage dans ce cas sont certainement dus au fait que le marqueur se dégrade sous l'effet des ultrasons. Les résultats avec la fragmentation par création de site abasique par action du pH acide sont meilleurs.

Exemple 19 : Marquage, fragmentation et dénaturation de l'ADN en une étape :

Les amplicons ADN ont été préparés par amplification PCR selon le protocole décrit dans l'exemple 5.

Deux réactions de marquage ont été réalisées :

a. Marquage, fragmentation et dénaturation à 95°C en une seule étape :

A 10 µL de PCR sont ajoutés 50 µL de tampon formiate de sodium pH3 (50 mM) et 2 µL du marqueur *méta*-BioPMDAM (100mM dans du DMSO anhydre). Le volume final est ajusté à 100 µL. La solution est ensuite incubée 30 min à 95°C. Dans ce cas, le mélange réactionnel a été hybridé sur puce à ADN sans aucune purification préalable.

b. Marquage et fragmentation à 60°C :

A 10 µL de PCR sont ajoutés 50 µL de tampon formiate de sodium pH3 (50 mM) et 2 µL du marqueur *méta*-BioDPDAM (100mM dans du DMSO anhydre). Le volume final est ajusté à 100 µL. La solution est ensuite incubée 30 min à 60°C. Le mélange réactionnel a été ensuite purifié selon le protocole décrit auparavant. Dans ce protocole et avant hybridation sur puce à ADN, les fragments ont été dénaturés selon le protocole décrit dans l'exemple 12.2.

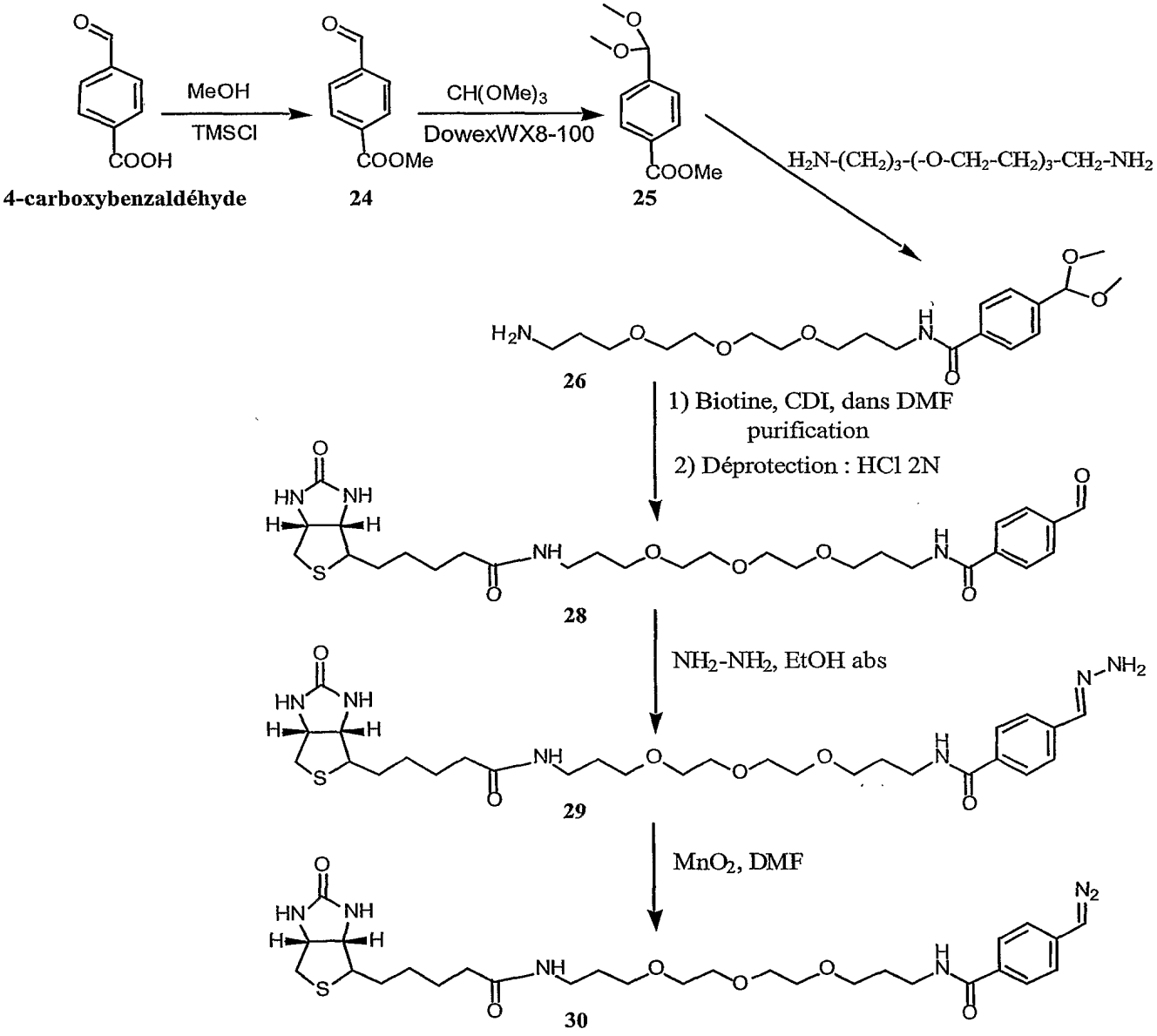
Résultats

Protocole utilisé	Homologie (%)	I (rfu)	B (rfu)	I/B
a. Marquage, fragmentation et dénaturation à 95°C	>95	5426	506	10,7
b. Marquage et fragmentation à 60°C puis dénaturation à 95°C	>95	7015	818	6,8

Tableau 11 : Marquage, fragmentation et dénaturation de l'ADN en une étape

L'exemple 12.2 a démontré l'importance de la dénaturation de l'ADN double brin pour la sensibilité de détection. Ces résultats du tableau 11 montrent qu'avec l'approche fragmentation par création de site abasique, le marquage, la fragmentation et la dénaturation de l'ADN peuvent être réalisées en une seule étape, ce qui représente une amélioration notable du point de vue de la simplicité et du temps pour l'utilisateur et cela sans affecter la sensibilité de détection.

Exemple 20 : Synthèse du *para*-Bio-EG3-PDAM :



- Protection du 4-carboxylbenzaldéhyde :

Le 4-carboxybenzaldéhyde est commercial. On le dissout (3 g ; 20 mmol) dans une solution de chlorure de triméthylsilyle (10 g ; 92 mmol) dans 100 mL de MeOH. Le mélange est laissé sous agitation pendant 40 h à température ambiante. Après évaporation, un solide blanc

correspondant au 4-méthoxycarbonylbenzaldéhyde **24** est isolé, caractérisé par RMN et utilisé tel quel pour la réaction suivante.

RMN ^1H (200MHz, $CDCl_3$) δ =10,07 (s, 1H, -CHO) ; 8,17 (d, 2H, J = 8 Hz, Ar-$H_{2,6}$) ; 7,92 (d, 2H, J = 8 Hz, Ar-$H_{3,5}$) ; 3,93 (s, 3H, CO-O-CH_3).

- Protection du 4'-méthoxycarbonylbenzaldéhyde :

Le 4-méthoxycarbonylbenzaldéhyde (3,35 g ; 20 mmol) est mis en solution avec de l'orthoformiate de triméthyle (4,8 g ; 40 mmol) en présence de Dowex 50WX8-100 (1 g). Le mélange est chauffé à reflux pendant 2 h, puis filtré et évaporé. Après un essai de recristallisation, une analyse par RMN montre que la réaction n'est pas complète et celle-ci est relancée dans 30 mL de MeOH, 30 mL de $CH(OMe)_3$ et 1 g de Dowex 50WX8-100 à température ambiante. On réalise une filtration puis une évaporation, afin d'obtenir 3,55 g (16,89 mmol, 84 %) de produit **25**.

RMN ^1H (200MHz, $CDCl_3$) δ = 8,01 (d, 2H, J = 8 Hz, Ar-$H_{2,6}$) ; 7,50 (d, 2H, J = 8 Hz, Ar-$H_{3,5}$) ; 5,41 (s, 1H, CH) ; 3,93 (s, 3H, -CO-O-CH_3) ; 3,29 (s, 6H, -O-CH_3).

- Composé **26** :

On solubilise le composé **25** (3,1 g ; 14,8 mmol) dans 16 mL (73 mmol) de 4,7,10-Trioxa-1,13-tridécanediamine. La solution obtenue est chauffée à 140-150°C pendant 2 h. Le mélange est ensuite dissous dans 100 mL de DCM (dichlorométhane ou CH_2Cl_2) et lavé 6 fois avec 10 mL d'eau. La phase organique est séchée avec Mg SO_4 puis évaporée jusqu'à l'obtention d'une huile. Cette huile est lavée avec du pentane 3 fois de suite par décantation, puis une nouvelle extraction avec DCM et H_2O est réalisée. Après séchage sur $MgSO_4$ et évaporation, le produit **26** est isolé avec un rendement de 63 % (9,27 mmol).

RMN ^1H (200MHz, $CDCl_3$) δ = 7,78 (d, 2H, J = 8 Hz, Ar-$H_{2,6}$) ; 7,46 (d, 2H, J = 8 Hz, Ar-$H_{3,5}$) ; 5,39 (s, 1H, CH) ; 3,62-3,47 (m, 14H, $H_{7',8',10',11'}$ et $H_{5',13'}$ et $H_{3'}$) ; 3,29 (s, 6H, -O-CH_3) ; 2,72 (m, 2H, $H_{15'}$) 1,87 (m, 2H, $H_{4'}$) ; 1,64 (m, 2H, $H_{14'}$) ; 1,30 (s large, 2H, NH_2).

● Composé biotinylé **27** :

La biotine (500 mg ; 2,05 mmol) est mise en suspension dans 10 mL de DMF puis 365 mg (2,25 mmol) de CDI sont ajoutés. Cette solution est laissée sous agitation pendant 30 min à température ambiante. Le composé **26** (900 mg ; 2,26 mmol) est dissous dans 1 mL de DMF, puis ajouté petit à petit à la solution précédente. Le mélange ainsi obtenu est laissé sous agitation pendant 1 h à température ambiante. Après évaporation, une purification sur colonne par chromatographie-flash (colonne de 20 mm de diamètre) est réalisée avec 250 mL de MeOH-DCM 6 %, puis avec 200 mL de MeOH-DCM 7 % et enfin 200 mL de MeOH-DCM 8 %. Les fractions correspondant au produit **27** sont rassemblées puis évaporées à sec pour donner 1,00 g d'huile avec un rendement estimé à 50 %.

RMN ^{1}H (200MHz, $CDCl_3$) δ = 9,50 (s large, 1H, $NH_{imidazole}$) ; 7,80 (d, 2H, J = 8 Hz, Ar-$H_{2,6}$) ; 7,64 (s, 1H, $H_{imidazole}$) ; 7,46 (d, 2H, J = 8 Hz, Ar-$H_{3,5}$ et 1H, $NH_{2'}$) ; 7,05 (s, 2H, $H_{imidazole}$) ; 6,76 (t, 1H, $NH_{16'}$) , 6,20 (s large, 1H, N$\underline{H}_{B1}$) ; 5,44 (s large, 1H, N$\underline{H}_{B3}$) ; 5,37 (s, 1H, CH) ; 4,42 (m, 1H, H_{B6a}) ; 4,24 (m, 1H, H_{B3a}) ; 3,59-3,44 (m, 14H, $H_{7',8',10',11'}$ et $H_{5',13'}$ et $H_{3'}$) ; 3,29 (m, 8H, $H_{15'}$ et 2-O-CH_3) ;3,07 (m, 1H, H_{B4}) ; 2,84 et 2,66 (système ABX, 2H, $^2J_{AB}$= 5 Hz, $^3J_{AX}$ = 12 Hz, $^3J_{BX}$ = 0 Hz, H_{B6}) ; 2,13 (t, 2H, J = 8 Hz, H_{B10}) ;1,85 (m, 2H, $H_{4'}$) ; 1,66 (m, 2H, $H_{14'}$) ; 1,40-1,37 (m, 6H, $H_{B7, B8, B9}$).

● Composé aldéhydique **28** :

L'acétal **27** est dissous dans 50 mL de chloroforme, puis 20 mL de HCl 2N sont ajoutés . Le mélange diphasique est vigoureusement agité pendant 15 min. On récupère la phase organique qu'on sèche sur $NaHCO_3$ anhydre. On filtre, on évapore et le composé **28** est obtenu sous forme d'une pâte (495 mg ; 0,855 mmol) avec un rendement global de 42 % à partir de la biotine.

RMN ^{1}H (300MHz, $CDCl_3$) δ = 10,05 (s, 1H, CHO) ; 7,98 (d, 2H, J = 8 Hz, Ar-$H_{2,6}$) ; 7,92 (d, 2H, J = 8 Hz, Ar-$H_{3,5}$) ; 7,58 (t, 1H, $NH_{2'}$) ; 6,46 (t, 1H, $NH_{16'}$) , 6,02 (s large, 1H, N$\underline{H}_{B1}$) ;

5,19 (s large, 1H, N$\underline{H}_{B3}$) ; 4,46 (m, 1H, H_{B6a}) ; 4,27 (m, 1H, H_{B3a}) ; 3,66-3,56 (m, 10H, $H_{7',8',10',11'}$ et $H_{5'}$) ; 3,50-3,29 (m, 4H, $H_{3',13'}$) ; 3,28 (m, 2H, $H_{15'}$) ; 2,95 (m, 1H, H_{B4}) ; 2,84 et 2,71 (système ABX, 2H, $^2J_{AB}$= 5 Hz, $^3J_{AX}$ = 12 Hz, $^3J_{BX}$ = 0 Hz, H_{B6}) ; 2,15 (t, 2H, J = 8 Hz, H_{B10}) ;1,89 (m, 2H, $H_{4'}$) ; 1,72-1,63 (m, 6H, $H_{14'}$, $H_{B7, B9}$) ; 1, 23 (m, 2H, H_{B8}).

- Composé hydrazone **29** :

L'aldéhyde **28** (495 mg ; 0,855 mmol) est dissous dans 10 mL d'éthanol absolu. L'hydrazine (350 µL ; 7,20 mmol) est ajoutée, puis le mélange réactionnel est chauffé à reflux pendant 1 h. L'huile obtenue après évaporation est dissoute dans EtOH abs, pour être de nouveau évaporée. On obtient alors une mousse qu'on triture au pentane. La pâte correspondant au produit **29** (511 mg ; 0,862 mmol) est obtenue avec un rendement de 100 %.

RMN ^{1}H (300MHz, $CDCl_3$) δ = 7,76 (d, 2H, J = 8 Hz, Ar-$H_{2,6}$) ; 7,72 (s, 1H, CH) ; 7,56 (d, 2H, J = 8 Hz, Ar-$H_{3,5}$) ; 7,34 (t, 1H, $NH_{2'}$) ; 6,45 (t, 1H, $NH_{16'}$) , 5,98 (s large, 1H, N$\underline{H}_{B1}$) ; 5,78 (s large, 2H, NH_2) ; 5,18 (s large, 1H, N$\underline{H}_{B3}$) ; 4,44 (m, 1H, H_{B6a}) ; 4,26 (m, 1H, H_{B3a}) ; 3,62-3,56 (m, 10H, $H_{7', 8',10',11'}$ et $H_{5'}$) ; 3,48-3,45 (m, 4H, $H_{3',13'}$) ; 3,27 (m, 2H, $H_{15'}$) ; 3,07 (m, 1H, H_{B4}) ; 2,84 et 2,68 (système ABX, 2H, $^2J_{AB}$= 5 Hz, $^3J_{AX}$ = 12 Hz, $^3J_{BX}$ = 0 Hz, H_{B6}) ; 2,11 (t, 2H, J = 8 Hz, H_{B10}) ;1,86 (m, 2H, $H_{4'}$) ; 1,72-1,59 (m, 6H, $H_{14'}$, $H_{B7, B9}$) ; 1, 21 (m, 2H, H_{B8}).

- Composé diazo **30** :

On solubilise l'hydrazone **29** (357 mg ; 0,602 mmol) dans 17,5 mL de DMF. On ajoute alors MnO_2 (700 mg ; 7,7 mmol). Après 12 min d'agitation à température ambiante, le mélange est filtré sur millipore contenant de la célite (épaisseur : 2 cm) et du tamis moléculaire en poudre 3 Å (0,5 cm). Le mélange réactionnel est évaporé à sec. L'huile résiduelle obtenue est lavée à l'éther, trois fois de suite. Le composé **30** (290 mg, 0,491 mmol) est obtenu sous forme d'un solide légèrement rose avec un rendement de 82 %.

RMN ^{1}H (300MHz, DMSO-d_6) δ = 8,28 (t, 1H, $NH_{2'}$) ; 7,77 (d, 2H, J = 8 Hz, Ar-$H_{2,6}$) ; 7,74 (t, 1H, $NH_{16'}$) ; 7,00 (d, 2H, J = 8 Hz, Ar-$H_{3,5}$) ; 6,38 (s large, 1H, N$\underline{H}_{B1}$) ; 6,32 (s large, 1H, N$\underline{H}_{B3}$) ; 5,80 (s, 1H, C$\underline{H}$-N_2) ; 4,27 (m, 1H, H_{B6a}) ; 4,11 (m, 1H, H_{B3a}) ; 3,51-3,44 (m, 10H, $H_{7', 8', 10', 11'}$ et $H_{5'}$) ; 3,37 (m, 2H, $H_{15'}$) ; 3,32 (m, 4H, $H_{3',13'}$) ; 3,05 (m, 1H, H_{B4}) ; 2,79 et 2,58 (système ABX, 2H, $^2J_{AB}$= 5 Hz, $^3J_{AX}$ = 12 Hz, $^3J_{BX}$ = 0 Hz, H_{B6}) ; 2,02 (t, 2H, J = 8 Hz, H_{B10}) ;1,69 (m, 2H, $H_{4'}$) ; 1,59-1,48 (m, 6H, $H_{14'}$, $H_{B7, B9}$) ; 1, 25 (m, 2H, H_{B8}).

La réactivité du composé **30** a été testée sur 3'-Uridine-monophosphate et suivie par électrophorèse capillaire. Les conditions d'analyse sont celles de l'exemple 6.1. Les résultats montrent un temps de demi vie de 45 minutes.

Le réactif est stable à -20°C pendant au moins 1 mois.

Exemple 21 : Marquage et fragmentation des amplicons d'ADN avec le réactif de marquage *para*-Bio-EG3-PDAM :

Les intérêts principaux de ce genre de molécules, c'est-à-dire de dérivés de PDAM portant un bras de liaison à base de polyéthylène glycol, sont de permettre l'éloignement de la fonction diazo par rapport à la biotine, et d'augmenter la solubilité et, au final, la réactivité de ces molécules.

Le dérivé *para*-Bio-EG3-PDAM **30** a été obtenu selon le schéma réactionnel décrit dans l'exemple 20. Les amplicons ADN ont été préparés par amplification PCR selon le protocole décrit dans l'exemple 5. Deux réactions de marquage ont été réalisées.

a. Marquage par le réactif *para*-Bio-EG3-PDAM :

A 10 µL de PCR sont ajoutés 10 µL de *para*-Bio-EG3-PDAM (100 mM dans du DMSO) et 77 µL d'eau Dnase/Rnase free. La solution est homogène et ne présente pas de précipités. Cette solution est incubée 10 min à 95°C, puis 3 µL d'HCl à 0,1 M sont ajoutés et la solution est incubée 10 min à 95°C.

Le reste du protocole est identique à celui de l'exemple 8.

b. Marquage par le réactif *méta*-BioPMDAM :

A 10 µL de PCR sont ajoutés 10 µL de *méta*-BioPMDAM (100 mM dans du DMSO) et 77 µL d'eau Dnase/Rnase free. La synthèse de ce produit est évoqué dans l'exemple 1.1. La solution présente un léger précipité. Cette solution est incubée 10 min à 95°C. Puis 3 µL d'HCl 0,1 M sont ajoutés et la solution est incubée 10 min à 95°C.
Le reste du protocole est identique à celui de l'exemple 8.

Résultats :

Protocole utilisé	Homologie (%)	I (rfu)	B (rfu)	I/B
a. Marquage par le réactif *para*-Bio-EG3-PDAM	> 95%	15151	621	24,4
b. Marquage par le réactif *méta*-BioPMDAM	> 95%	11226	515	21,8

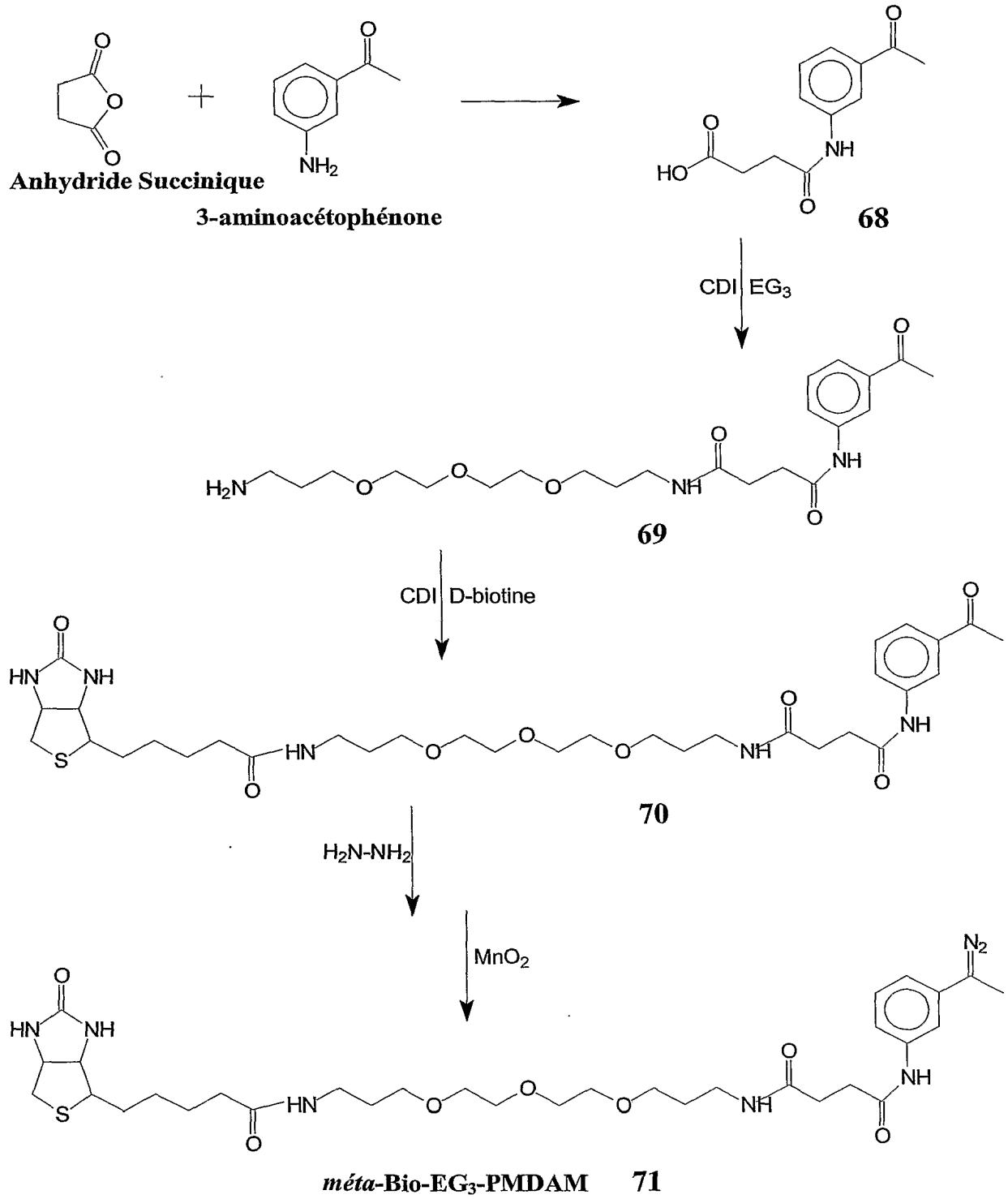
Tableau 12 : Etude comparative du marquage et de la fragmentation d'amplicons d'ADN avec le *para*-Bio-EG3-PDAM et le *méta*-BioPMDAM

Les intensités de signal obtenues dans ce tableau 12 sont très satisfaisantes et le pourcentage d'homologie est élevé. Ce résultat montre que l'introduction d'un bras polyéthylène glycol sur la molécule de marquage diazo permet d'augmenter la solubilité en phase aqueuse du réactif. Le test est donc homogène. De plus, l'augmentation de la solubilité permet d'augmenter la réactivité du marqueur.

Exemple 22 : Synthèse d'autres dérivés PDAM comportant un bras de liaison à base de polyéthylène glycol :

Exemple 22.1 : Synthèse du *méta*-Bio-EG3-PMDAM :

Schéma de synthèse



Composé **68** :

On solubilise la 3-aminoacetophenone (14,5 g, 107 mmol) dans 50 ml de DMF anhydre. On ajoute l'anhydride succinique (10,7 g, 107 mmol) et on laisse sous agitation, sous argon et a température ambiante. Après 6 h, la solution est concentrée sous vide et 50 ml de méthanol sont ajoutés. Le précipité obtenu est filtré et lavé avec méthanol et éther. On obtient ainsi 19,4 g (81%) de produit **68** sous forme d'une poudre de couleur blanc cassé.
RMN ^1H (200 MHz, DMSO-d_6): d = 2,5-2,6 (m, 7H); 7,45 (t, 1H); 7,64 (d, 1H); 7,83 (d, 1H); 8,19 (s, 1H); 10,16 (s, 1H); 12,12 (s, 1H).

Composé **69** :

On solubilise 5,07 g (22 mmol) du composé **68** dans 10 ml de DMF anhydre, sous argon. On met sur glace et on ajoute 5,00 g (32 mmol) de carbonyldiimidazole. Apres 20 min, on ajoute lentement 20 ml (94,6 mmol) du 4, 7, 10-trioxatridecanediamine (EG_3). Après 3h de réaction à température ambiante, on évapore le DMF et reprend le résidu dans 100 ml de CH_2CL_2. On fait des extractions avec du $NaHCO_3$ saturé et H_2O, après quoi la phase organique est séchée avec du Na_2SO_4 anhydre et le solvant évaporé. On obtient ainsi 4,34 g (46%) du produit **69**.
RMN ^1H (200 MHz, DMSO-d_6): d = 1,59 (m, 2H); 1,87 (m, 2H); 2,16 (s, 3H); 2,40 (m, 2H); 2,55 (m, 2H); 3,08 (m, 2H); 3,45 (m, 16H); 7,30 (t, 1H); 7,42 (d, 1H); 7,70 (d, 1H); 7,83 (t, 1H); 7,97 (s, 1H); 10,00 (s, 1H).

Composé biotinylé **70** :

On solubilise la D-biotine (1,0 g, 4,1 mmol) dans 10 ml de DMF anhydre, sous argon. On refroidit sur glace et on ajoute le carbonyldiimidazole (CDI) (0,665 g, 4,1 mmol) dans 10 ml de DMF anhydre. Après 15 min, on ajoute le composé **69** (1,8 g, 4,1 mmol) dans 2 ml de DMF anhydre. On laisse réagir 3 h à 35°C, puis on évapore le DMF et on reprend dans 100ml de CH_2CL_2. On fait des extractions avec du $NaHCO_3$ saturé et H_2O, après quoi la phase organique est séchée avec du Na_2SO_4 anhydre et le solvant évaporé. La caractérisation par

RMN du produit ainsi obtenu montre qu'on obtient un mélange du produit **70** et du EG_3 libre. Une autre étape de purification ets effectuée avant de continuer la synthèse.

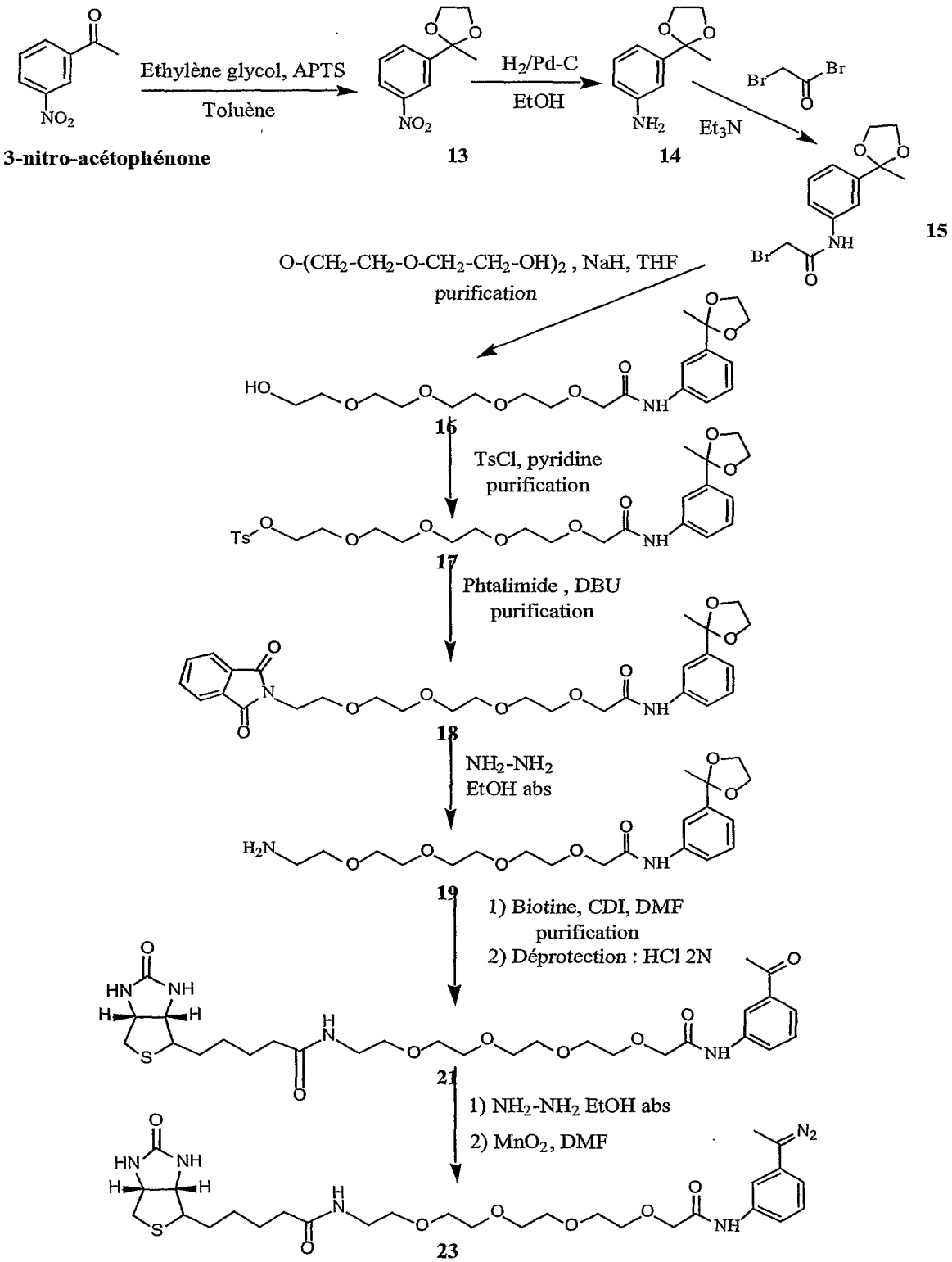
Le composé final, *méta*-Bio-EG_3-PMDAM, est obtenu après deux étapes de synthèse selon le schéma décrit dans l'exemple 1.

L'intérêt de cette synthèse est double. D'une part, on obtient le produit **69** en seulement deux étapes; ce produit peut être utilisé comme précurseur du diazo avec la possibilité d'y accrocher des molécules détectables de nature différente, grâce au groupement amine terminal. Ce groupement permet aussi de greffer le composé **69** sur des supports solides, avec l'objectif d'immobiliser des acides nucléiques. D'une autre part, le composé **71** possède le même centre réactif que la *méta*-Bio-PMDAM (notre molécule de référence), ce qui facilite l'analyse des avantages liés à l'inclusion du bras éthylèneglycol (EG_3).

Le réactif est stable à -20°C pendant au moins 1 mois.

Exemple 22.2 : Synthèse du *méta*-Bio-EG4-PMDAM :

Schéma de synthèse :


Ethylène glycol, APTS
Toluène
H2/Pd-C
EtOH
Et3N
3-nitro-acétophénone
13
14
15
O-(CH2-CH2-O-CH2-CH2-OH)2 , NaH, THF
purification
16
TsCl, pyridine
purification
17
Phtalimide , DBU
purification
18
NH2-NH2
EtOH abs
19
1) Biotine, CDI, DMF
purification
2) Déprotection : HCl 2N
21
1) NH2-NH2 EtOH abs
2) MnO2, DMF
23

● Protection de la 3-nitro-acétophenone **13** :

On dissout 33 g (0,20 mol) de 3-nitro-acétophénone dans 400 mL de toluène, puis on ajoute 40 mL (0,717 mol) d'éthylène glycol et 600 mg (3,15 mmol) d'acide *para*-toluène sulfonique (APTS). On monte un système Dean Stark. On chauffe la solution pendant 3 h à 130°C. Après avoir laissé la solution revenir à température ambiante, on ajoute 400 mL d'acétate d'éthyle, puis on lave celle-ci avec 8 mL d'une solution de $NaHCO_3$ saturée. On sèche la phase organique sur $MgSO_4$. Après évaporation, on obtient un solide jaune pâle **13** (39,72 g ; 0,190 mol) avec un rendement de 95 %.

RMN ^1H (200MHz, $CDCl_3$) δ =7,11 (t, 1H, J = 8 Hz, Ar-H) ; 6,87-6,78 (m, 2H, Ar-H) ; 6,59 (dd, 1H, J = 6,5 Hz, Ar-H) ; 4,00 (m, 2H, $\underline{H_2C}_{acétale}$) ; 3,79 (m, 2H, $\underline{H_2C}_{acétale}$) ; 1,61 (s, 3H, CH_3).

● Préparation de l'amine **14** :

Le composé **13** (39,7 g ; 0,190 mol) est dissous dans 500 mL d'EtOH, puis 1 g de palladium sur charbon 10 % est ajouté. On chauffe pour tout dissoudre, puis on laisse la solution revenir à température ambiante. Après avoir fait le vide et mis la solution sous H_2, on laisse sous forte agitation pendant 5 h. La solution est ensuite filtrée à chaud puis évaporée. Le produit **14** est lavé au pentane, et isolé sous forme d'un solide (34 g ; 0,189 mol) avec un rendement de 99 %.

RMN ^1H (200MHz, $CDCl_3$) δ =7,14 (t, 1H, J = 8 Hz, Ar-H) ; 6,85 (m, 2H, J = 7,5 Hz, Ar-H) ; 6,79 (s, 1H, Ar-H) ; 6,59 (dd, 1H, J = 6,5 Hz, Ar-H) ; 4,00 (m, 2H, $\underline{H_2C}_{acétale}$) ; 3,77 (m, 2H, $\underline{H_2C}_{acétale}$) ; 1,61 (s, 3H, CH_3).

● Composé bromé **15** :

L'amine **14** (12,3 g ; 68,7 mmol) et la triéthylamine (7 g ; 69 mmol) sont mises en solution dans 150 mL de DCM sous argon. On ajoute à –5 °C au goutte-à-goutte une solution de 13,8 g (60 mmol) de bromure de bromoacétyle dissous dans 150 mL de DCM. À la fin de l'addition,

on ajoute 100 mL de $NaHCO_3$ aqueux 1 N. On lave la phase organique deux fois de suite avec $NaHCO_3$ aqueux et on sèche sur $MgSO_4$. Après évaporation à sec, 22,6 g d'huile marron est obtenue correspondant au composé **15** et utilisée telle quelle pour la réaction suivante.

RMN ^1H (200MHz, $CDCl_3$) δ = 8,29 (s large, 1H, NH) ; 7,62 (dt, $1H_4$, J = 5 Hz, Ar-H) ; 7,47 (s, 1H, Ar-H_2) ; 7,38-7,19 (m, 2H, Ar-$H_{5,6}$) ; 4,00 (m, 2H, Br-CH_2) ; 3,75 (m, 4H, H_2C-$H_2C_{acétale}$) ; 1,61 (s, 3H, CH_3).

- Composé alcool **16** :

L'hydrure de sodium (3,3 g ; 82,5 mmol) est lavé trois fois avec du pentane puis mis en suspension dans 150 mL de THF, le tétra éthylène glycol (50 mL ; 0,29 mol) est alors ajouté à température ambiante. On laisse la réaction sous agitation pendant 15 min, puis on refroidit la solution à –5°C.

L'ajout du composé **15** préalablement dilué dans 25 mL de THF se fait au goutte-à-goutte. On laisse sous agitation pendant 30 min pour laisser revenir à température ambiante. On concentre la solution à 100 mL, puis on dilue celle-ci avec 500 mL de $CHCl_3$. On lave trois fois de suite cette phase organique avec 250 mL de $NaHCO_3$ aqueux 1 N, puis on la sèche sur $MgSO_4$ avant de l'évaporer. On purifie le produit sur colonne de silice par chromatographie-flash (colonne de 65 mm de diamètre) avec 1,5 L de MeOH-DCM 5 %, puis avec 500 mL de MeOH-DCM 7 %, et enfin 500 mL de MeOH-DCM 10 %. Les fractions correspondant au composé **16** sont rassemblées, puis évaporées à sec pour donner 17,4 g (42,1 mmol) de produit, avec un rendement de 61 %.

RMN ^1H (200MHz, $CDCl_3$) δ = 8,86 (s large, 1H, NH) ; 7,71 (d, 1H, J = 7,5 Hz, Ar-H_4) ; 7,51 (s, 1H, Ar-H_2) ; 7,29-7,24 (m, 2H, Ar-$H_{5,6}$) ; 4,09 (m, 2H, CO-CH_2-O) ; 3,99 (m, 2H, $H_2C_{acétale}$) ; 3,72-3,53 (m, 20H, O-CH_2-CH_2-O, $H_2C_{acétale}$ et HO-CH_2) ; 1,61 (s, 3H, CH_3).

- Composé tosylate **17** :

L'alcool **16** (4,13 g ;10,0 mmol) est mis en solution dans 5 mL de pyridine. On ajoute ensuite 2,0 g (10,5 mmol) de chlorure de tosyle à température ambiante. On met sous agitation sous

argon pendant 10 h. On dilue avec 100 mL de DCM, on lave trois fois la phase organique avec 20 mL de $NaHCO_3$ aqueux 1 N, puis on la sèche sur $MgSO_4$ avant de la co-évaporer avec du toluène. Une purification sur colonne par chromatographie-flash (colonne de 50 mm de diamètre) est réalisée avec 500 mL de MeOH-DCM 2 %, puis 500 mL de MeOH-DCM 3 %, et enfin 500 mL de MeOH-DCM 4 %. Les fractions correspondant au produit **17** sont rassemblées puis évaporées à sec pour donner 3,68 g (6,48 mmol) d'huile avec un rendement estimé à 65 %.

RMN ^1H (300MHz, $CDCl_3$) δ = 8,86 (s large, 1H, NH) ; 7,76 (d, 4H, J = 5,5 Hz, Ar-H_{tosyl}) ; 7,60 (d, 1H, J = 7,5 Hz, Ar-H_4) ; 7,50 (s, 1H, Ar-H_2) ; 7,32-7,22 (m, 2H, Ar-$H_{5,6}$) ; 4,10 (m, 2H, CO-C$\underline{H_2}$-O) ; 4,00 (m, 2H, $\underline{H_2}C_{acétale}$) ; 3,73-3,54 (m, 20H, O-C$\underline{H_2}$-C$\underline{H_2}$-O, $\underline{H_2}C_{acétale}$ et HO-C$\underline{H_2}$) ; 2,42 (s, 3H, Ar-CH_3) ; 1,61 (s, 3H, CH_3).

- Composé phtalimide **18** :

On met le tosylate **17** (3,68 g ; 6,48 mmol) en solution avec 1,52 g (10,0 mmol) de DBU (1,8-diazobicyclo-[5.4.0]-undecène) puis on ajoute le phtalimide (1,47 g ; 10 mmol). La solution ainsi obtenue est chauffée à 85-90°C pendant 17 h puis évaporée. On purifie le produit sur colonne de silice par chromatographie-flash (colonne de 50 mm de diamètre) avec 1 L d'acétone-DCM 15 %, puis avec 1 L d'acétone-DCM 20 %. Les fractions correspondant au composé **18** sont rassemblées, puis évaporées à sec pour donner 3,15 g (5,8 mmol) de produit, avec un rendement de 90 %.

RMN ^1H (200MHz, $CDCl_3$) δ = 8,73 (s large, 1H, NH) ; 7,79 (m, 2H, Ar-H_{phta}) ; 7,99 (m, 2H, Ar-H_{phta} et Ar-H_4) ; 7,49 (s, 1H, Ar-H_2) ; 7,27-7,18 (m, 2H, Ar-$H_{5,6}$) ; 4,10 (m, 2H, CO-C$\underline{H_2}$-O) ; 4,00 (m, 2H, $\underline{H_2}C_{acétale}$) ; 3,69-3,56 (m, 20H, O-C$\underline{H_2}$-C$\underline{H_2}$-O, $\underline{H_2}C_{acétale}$ et N_{phta}-C$\underline{H_2}$) ; 1,61 (s, 3H, CH_3).

● Composé amine **19** :

Le produit **18** est dissous dans 20 mL d'EtOH absolu en chauffant à reflux à 75-80°C. L'hydrazine (1,07 mL ; 22,1 mmol) est ensuite ajoutée, et la réaction est laissée sous agitation pendant 1 h 15. Le précipité obtenu est filtré sur fritté et la phase éthanolique évaporée. Le précipité blanc est ensuite lavé au DCM, et la phase DCM est évaporée. L'huile jaune obtenue (2,3 g ; 5,57 mmol) est directement utilisée pour la réaction suivante, même si elle contient de l'imidazole qui pourra être éliminé par la suite, lors de l'étape de déprotection de l'acétal.

RMN ^{1}H (300MHz, $CDCl_3$) δ = 8,83 (s large, 1H, NH) ; 7,69 (d, 1H, J = 7,5 Hz, Ar-H_4) ; 7,51 (s, 1H, Ar-H_2) ; 7,30-7,19 (m, 2H, Ar-$H_{5,6}$) ; 4,10 (m, 2H, CO-C$\underline{H_2}$-O) ; 4,00 (m, 2H, $\underline{H_2}C_{acétale}$) ; 3,69-3,56 (m, 20H, O-C$\underline{H_2}$-C$\underline{H_2}$-O, $\underline{H_2}C_{acétale}$ et H_2N-C$\underline{H_2}$) ; 1,61 (s, 3H, CH_3).

● Composé biotinylé **20** :

On solubilise la D-Biotine (1,05 g ; 4,32 mmol) dans 10 mL de DMF anhydre. On ajoute sous argon 790 mg (4,87 mmol) de carbonyle diimidazole (CDI). Après 10 min d'agitation, on ajoute l'amine **19** diluée dans 5 mL de DMF. La solution est laissée 40 min sous agitation, puis évaporée avant d'être purifiée sur colonne par chromatographie-flash.
Pour ceci on utilise une colonne de 50 mm de diamètre avec comme éluant 500 mL de MeOH-DCM 5 %, puis 500 mL de MeOH-DCM 10 %, et enfin 500 mL de MeOH-DCM 15 %. Les fractions correspondant au produit **20** sont rassemblées puis évaporées à sec pour donner une huile jaune (1,66 g ; 2,6 mmol).
L'huile jaune obtenue (2,4 g) contenant d'après le spectre RMN environ 30 % en poids d'imidazole. On en déduit donc que le rendement de la réaction aboutissant au produit **20** est d'environ 60 % par rapport à la biotine de départ.

RMN ^{1}H (300MHz, $CDCl_3$) δ = 8,80 (s large, 1H, NH) ; 7,66 (m, 3H, Ar-H_4 et $H_{imidazole}$) ; 7,54 (s, 1H, Ar-H_2) ; 7,28-7,24 (m, 2H, Ar-$H_{5,6}$) ; 7,07 (s, 2H, $H_{imidazole}$), 6,59 (t, 1H, $NH_{15'}$) ; 6,06 (s large, 1H, N$\underline{H}_{B1}$) ; 5,19 (s large, 1H, N$\underline{H}_{B3}$) ; 4,45 (m, 1H, H_{B6a}) ; 4,27 (m, 1H, H_{B3a}) ; 4,10 (s, 2H, $H_{3'}$) ; 4,00 (m, 2H, $\underline{H_2}C_{acétale}$) ; 3,75-3,49 (m, 18H, O-C$\underline{H_2}$-C$\underline{H_2}$-O et $\underline{H_2}C_{acétale}$,) ;

3,36 (m, 2H, $H_{14'}$) ; 3,09 (m, 1H, H_{B4}) ; 2,85 et 2,66 (système ABX, 2H, $^2J_{AB}$= 5 Hz, $^3J_{AX}$ = 12 Hz, $^3J_{BX}$ = 0 Hz, H_{B6}) ; 2,16 (t, 2H, J = 8 Hz, H_{B10}) ; 1,61 (s, 3H, CH_3) ; 1,59-1,3 (m, 6H, $H_{B9, B8, B7}$).

● Composé cétone **21** :

L'acétal **20** est dissous dans 80 mL de chloroforme, puis 30 mL de HCl 2 N sont ajoutés. On laisse sous forte agitation pendant 45 min. La phase organique est récupérée puis séchée sur $NaHCO_3$ anhydre. Après filtration, la solution est évaporée et l'huile obtenue est lavée au pentane pour donner le produit **21** (1,48 g ; 2,48 mmol) avec un rendement de 99 %.

RMN ^{1}H (300MHz, $CDCl_3$) δ = 8,99 (s large, 1H, NH) ; 8,07 (s, 1H, Ar-H_2) ; 7,98 (d, 2H, J = 8 Hz, Ar-H_4) ; 7,66 (d, 2H, J = 8 Hz, Ar-H_6) ; 7,42 (t, 2H, J = 8 Hz, Ar-H_5) ; 6,38 (t, 1H, $NH_{15'}$) ; 5,78 (s large, 1H, N$\underline{H}_{B1}$) ; 4,96 (s large, 1H, N$\underline{H}_{B3}$) ; 4,47 (m, 1H, H_{B6a}) ; 4,29 (m, 1H, H_{B3a}) ; 4,13 (s, 2H, $H_{3'}$) ; 3,76-3,37 (m, 16H, O-C$\underline{H}_2$-C$\underline{H}_2$-O) ; 3,32 (m, 2H, $H_{14'}$) ; 3,11 (m, 1H, H_{B4}) ; 2,89 et 2,75 (système ABX, 2H, $^2J_{AB}$= 5 Hz, $^3J_{AX}$ = 12 Hz, $^3J_{BX}$ = 0 Hz, H_{B6}) ; 2,59 (s, 3H, CH_3) ; 2,16 (t, 2H, J = 8 Hz, H_{B10}) ; 1,64-1,40 (m, 6H, $H_{B9, B8, B7}$).

● Composé hydrazone **22** :

La cétone **21** est dissoute dans 20 mL d'EtOH absolu. On chauffe à reflux à 75 –80°C. On ajoute ensuite l'hydrazine (816 µL ; 16,81 mmol) et on laisse sous agitation pendant 3 h. Après filtration, on évapore à sec, on redissout dans l'éthanol jusqu'à l'obtention d'une mousse blanche collante. Dans un deuxième temps, on dissout cette mousse dans 50 mL de chloroforme puis on ajoute 20 mL d'une solution de $NaHCO_3$ saturée. On lave bien puis on récupère la phase organique. On la sèche sur Na_2CO_3 anhydre et après filtration, on évapore à sec afin d'obtenir une nouvelle mousse collante. Celle-ci correspond au produit **22** (842 mg g ; 1,38 mmol) et est obtenue avec un rendement de 66 %.

RMN ^{1}H (300MHz, $CDCl_3$) δ = 8,81 (s large, 1H, NH) ; 8,82 (s, 1H, Ar-H_2) ; 7,64 (d, 2H, J = 8 Hz, Ar-H_4) ; 7,32 (m, 4H, Ar-$H_{5,6}$) ; 6,43 (t, 1H, $NH_{15'}$) ; 5,89 (s large, 1H, N$\underline{H}_{B1}$) ; 5,46 (s

large, 2H, NH_2) ; 4,99 (s large, 1H, $N\underline{H}_{B3}$) ; 4,44 (m, 1H, H_{B6a}) ; 4,27 (m, 1H, H_{B3a}) ; 4,11 (s, 2H, $H_{3'}$) ; 3,70-3,37 (m, 16H, O-C$\underline{H}_2$-C$\underline{H}_2$-O) ; 3,32 (m, 2H, $H_{14'}$) ; 3,08 (m, 1H, H_{B4}) ; 2,87 et 2,67 (système ABX, 2H, $^2J_{AB}$= 5 Hz, $^3J_{AX}$ = 12 Hz, $^3J_{BX}$ = 0 Hz, H_{B6}) ; 2,11 (m, 5H, CH_3 et H_{B10}) ; 1,64-1,40 (m, 6H, $H_{B9, B8, B7}$).

- Composé diazo **23** :

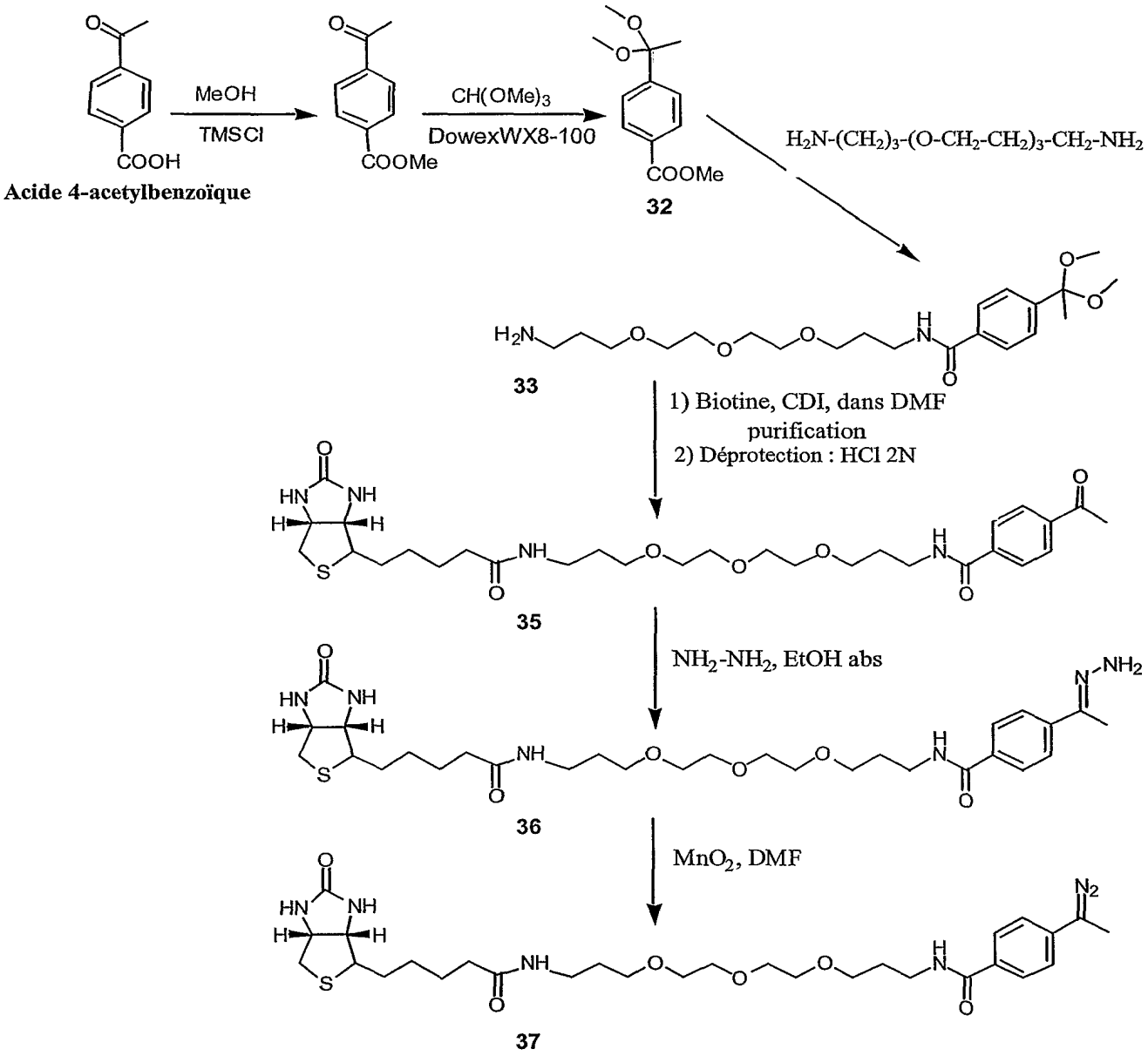
L'hydrazone **22** (100mg ; 0,164 mmol) est dissoute dans 1 mL de DMF sous argon. On ajoute 80 mg de MnO_2 activé et on laisse sous forte agitation pendant 30 min. Le mélange est filtré à travers une couche mixte Célite (3 cm) - tamis moléculaire en poudre (1 cm). La solution est ensuite évaporée à sec. L'huile obtenue en fin d'évaporation est triturée jusqu'à l'obtention d'un poudre rose correspondant au composé **23** (78 mg ; 0,128 mmol ; 78 %).

RMN ^{1}H (300MHz, DMSO-d_6) δ = 9,60 (s large, 1H, NH) ; 7,89 (s, 1H, Ar-H_2) ; 7,76 (t, 1H, $NH_{15'}$) ; 7,35-7,25 (m, 4H, Ar-$H_{5,6}$) ; 6,64 (d, 2H, J = 8 Hz, Ar-H_4) ; 6,36 (s large, 1H, $N\underline{H}_{B1}$) ; 6,32 (s large, 1H, $N\underline{H}_{B3}$) ; 4,28 (m, 1H, H_{B6a}) ; 4,08 (m, 1H, H_{B3a}) ; 4,06 (s, 2H, $H_{3'}$) ; 3,55-3,31 (m, 16H, O-C$\underline{H}_2$-C$\underline{H}_2$-O) ; 3,17 (m, 2H, $H_{14'}$) ; 3,08 (m, 1H, H_{B4}) ; 2,80 et 2,59 (système ABX, 2H, $^2J_{AB}$= 5 Hz, $^3J_{AX}$ = 12 Hz, $^3J_{BX}$ = 0 Hz, H_{B6}) ; 2,13 (m, 5H, CH_3) ; 2,13 (t, 2H, J = 8 Hz, H_{B10}) ; 1,99-1,30 (m, 6H, $H_{B9, B8, B7}$).

La réactivité du composé **23** a été testée sur 3'-Uridine-monophosphate et suivie par électrophorèse capillaire. Les conditions d'analyse sont celles de l'exemple 6.1. Les résultats montrent un temps de demi vie de 30 minutes.
La stabilité du réactif est supérieure à 1 mois à -20°C.

Exemple 22.3 : Synthèse du *para*-Bio-EG3-PMDAM :

Schéma de synthèse :



Acide 4-acetylbenzoïque

32

33

35

36

37

● Protection de l'acide 4-acétylbenzoïque :

On dissout l'acide 4-acétylbenzoïque (1 g ; 6,1 mmol) dans une solution de chlorure de triméthylsilyle (TMSCl, 10 g ; 92 mmol) dans 5 mL de MeOH. Le mélange est chauffé à 90°C pendant une nuit. Après évaporation, un solide blanc correspondant au composé **31** (1,21 g ; 5,75 mmol) est isolé, caractérisé par RMN et utilisé tel quel pour la réaction suivante.

RMN ^1H (200MHz, $CDCl_3$) δ = 8,08 (d, 2H, J = 8 Hz, Ar-$H_{2,6}$) ; 7,59 (d, 2H, J = 8 Hz, Ar-$H_{3,5}$) ; 3,18 (s, 6H, -O-CH_3) ; 1,53 (s, 3H, CH_3).

- Composé **32** :

On dissout le composé **31** (1,21 g ; 5,75 mmol) dans 5 ml d'orthoformiate de triméthyle en présence de Dowex 50WX8-100 (0,3 g). Le mélange est chauffé à 60°C pendant une nuit, puis filtré et évaporé pour donner le composé **32** (1,19 g ; 5,3 mmol) avec un rendement de 87 %.
RMN ^1H (200MHz, $CDCl_3$) δ = 8,00 (d, 2H, J = 8 Hz, Ar-$H_{2,6}$) ; 7,54 (d, 2H, J = 8 Hz, Ar-$H_{3,5}$) ; 3,89 (s, 1H, CO-O- CH_3) ; 3,16 (s, 6H, -O-CH_3) ; 1,51 (s, 3H, CH_3).

- Composé **33** :

On solubilise le composé **32** (1,17 g ; 5,22 mmol) dans 5 mL (22,7 mmol) de 4,7,10-trioxa-1,13-tridécanediamine. La solution obtenue est chauffée à 140°C pendant 4 h. Le mélange est ensuite dissous dans 30 mL de DCM et lavé 3 fois avec 10 mL d'eau. La phase organique est séchée avec Mg SO_4, puis évaporée jusqu'à l'obtention d'une huile correspondant au produit **33** (1,44 g ; 3,49 mmol) avec un rendement de 67 %.

RMN ^1H (200MHz, $CDCl_3$) δ = 7,76 (d, 2H, J = 8 Hz, Ar-$H_{2,6}$) ; 7,51 (d, 2H, J = 8 Hz, Ar-$H_{3,5}$) ; 3,62-3,47 (m, 14H, $H_{7',8',10',11'}$ et $H_{5',13'}$ et $H_{3'}$) ; 3,15 (s, 6H, -O-CH_3) ; 2,73 (m, 2H, $H_{15'}$) 1,88 (m, 2H, $H_{4'}$) ; 1,65 (m, 2H, $H_{14'}$) ; 1,38 (s large, 2H, NH_2).

- Composé biotinylé **34** :

La biotine (780 mg ; 3,19 mmol) est mise en suspension dans 13 mL de DMF. On ajoute ensuite 590 mg (3,60 mmol) de CDI. Cette solution est laissée sous agitation pendant 30 min à température ambiante. Le composé **33** est dissous dans 1 mL de DMF, puis ajouté petit à petit à la solution précédente. Le mélange ainsi obtenu est laissé sous agitation pendant 1 h à température ambiante. Après évaporation du DMF, une purification sur colonne par chromatographie-flash (colonne de 35 mm de diamètre) est réalisée avec 500 mL de MeOH-

DCM 6 %, puis avec 250 mL de MeOH-DCM 8 %, et enfin 250 mL de MeOH-DCM 8 %. Les fractions correspondant au produit **34** sont rassemblées, puis évaporées à sec pour donner 1,05 g d'huile avec un rendement estimé à 30 %.

RMN ^{1}H (300MHz, $CDCl_3$) δ = 8,49 (s large, 1H, $NH_{imidazole}$) ; 7,79 (d, 2H, J = 8 Hz, Ar-$H_{2,6}$) ; 7,66 (s, 1H, $H_{imidazole}$) ; 7,50 (d, 2H, J = 8 Hz, Ar-$H_{3,5}$) ; 7,38 (t, 1H, $NH_{2'}$) ; 7,11 (s, 2H, $H_{imidazole}$) ; 6,67 (t, 1H, $NH_{16'}$) , 5,99 (s large, 1H, N$\underline{H}_{B1}$) ; 5,15 (s large, 1H, N$\underline{H}_{B3}$) ; 4,46 (m, 1H, H_{B6a}) ; 4,27 (m, 1H, H_{B3a}) ; 3,61-3,45 (m, 14H, $H_{7',8',10',11'}$ et $H_{5',13'}$ et $H_{3'}$) ; 3,28 (m, 2H, $H_{15'}$) ; 3,15 (s, 6H, -OCH_3) ; 2,85 (m, 1H, H_{B4}) ; 2,85 et 2,69 (système ABX, 2H, $^2J_{AB}$= 5 Hz, $^3J_{AX}$ = 12 Hz, $^3J_{BX}$ = 0 Hz, H_{B6}) ; 2,14 (t, 2H, J = 8 Hz, H_{B10}) ;1,86 (m, 2H, $H_{4'}$) ; 1,69 (m, 2H, $H_{14'}$) ; 1,49 (s, 3H, CH_3) ; 1,42-1,39 (m, 6H, $H_{B7, B8, B9}$).

- Composé **35** :

L'acétal **34** est dissous dans 45 mL de chloroforme, puis 10 mL de HCl 2N sont ajoutés . Le mélange diphasique est vigoureusement agité pendant 5 min. On récupère la phase organique qu'on sèche sur $NaHCO_3$ anhydre. On filtre, on évapore, et le composé **35** est obtenu sous forme d'un solide jaune clair (504 mg ; 0,87 mmol) avec un rendement global de 27 % à partir de la biotine.

RMN ^{1}H (300MHz, $CDCl_3$) δ = 7,97 (d, 2H, J = 8 Hz, Ar-$H_{2,6}$) ; 7,91 (d, 2H, J = 8 Hz, Ar-$H_{3,5}$) ; 7,51 (t, 1H, $NH_{2'}$) ; 6,50 (t, 1H, $NH_{16'}$) , 6,05 (s large, 1H, N$\underline{H}_{B1}$) ; 5,23 (s large, 1H, N$\underline{H}_{B3}$) ; 4,45 (m, 1H, H_{B6a}) ; 4,27 (m, 1H, H_{B3a}) ; 3,62-3,56 (m, 10H, $H_{7',8',10',11'}$ et $H_{5'}$) ; 3,48-3,46 (m, 4H, $H_{3',13'}$) ; 3,27 (m, 2H, $H_{15'}$) ; 3,10 (m, 1H, H_{B4}) ; 2,85 et 2,71 (système ABX, 2H, $^2J_{AB}$= 5 Hz, $^3J_{AX}$ = 12 Hz, $^3J_{BX}$ = 0 Hz, H_{B6}) ; 2,60 (s, 3H, CH_3) ; 2,14 (t, 2H, J = 8 Hz, H_{B10}) ;1,89 (m, 2H, $H_{4'}$) ; 1,72-1,61 (m, 6H, $H_{14'}$, $H_{B7\ B9}$) ; 1, 40 (m, 2H, H_{B8}).

- Composé hydrazone **36** :

La cétone **35** (500 mg ; 0,864 mmol) est dissoute dans 11 mL de EtOH absolu. L'hydrazine (335 µL ; 6,911 mmol) est ajoutée, puis le mélange réactionnel est chauffé à reflux pendant 1

h. L'huile obtenue après évaporation, est dissoute dans EtOH abs, pour être de nouveau évaporée. On obtient alors une mousse collante correspondant au produit **36** (488 mg ; 0,823 mmol) avec un rendement de 95 %.

RMN ^{1}H (300MHz, $CDCl_3$) δ = 7,76 (d, 2H, J = 8 Hz, Ar-$H_{2,6}$) ; 7,67 (d, 2H, J = 8 Hz, Ar-$H_{3,5}$) ; 7,29 (t, 1H, $NH_{2'}$) ; 6,46 (t, 1H, $NH_{16'}$) , 5,98 (s large, 1H, N$\underline{H}_{B1}$) ; 5,55 (s large, 2H, NH_2) ; 5,14 (s large, 1H, N$\underline{H}_{B3}$) ; 4,45 (m, 1H, H_{B6a}) ; 4,24 (m, 1H, H_{B3a}) ; 3,62-3,51 (m, 10H, $H_{7', 8',10',11'}$ et $H_{5'}$) ; 3,47-3,45 (m, 4H, $H_{3',13'}$) ; 3,27 (m, 2H, $H_{15'}$) ; 3,07 (m, 1H, H_{B4}) ; 2,84 et 2,69 (système ABX, 2H, $^2J_{AB}$= 5 Hz, $^3J_{AX}$ = 12 Hz, $^3J_{BX}$ = 0 Hz, H_{B6}) ; 2,11 (t, 2H, J = 8 Hz, H_{B10} et s, 3H, CH_3) ; 1,86 (m, 2H, $H_{4'}$) ; 1,72-1,59 (m, 6H, $H_{14'}$, $H_{B7, B9}$) ; 1, 21 (m, 2H, H_{B8}).

- <u>Composé diazo **37** :</u>

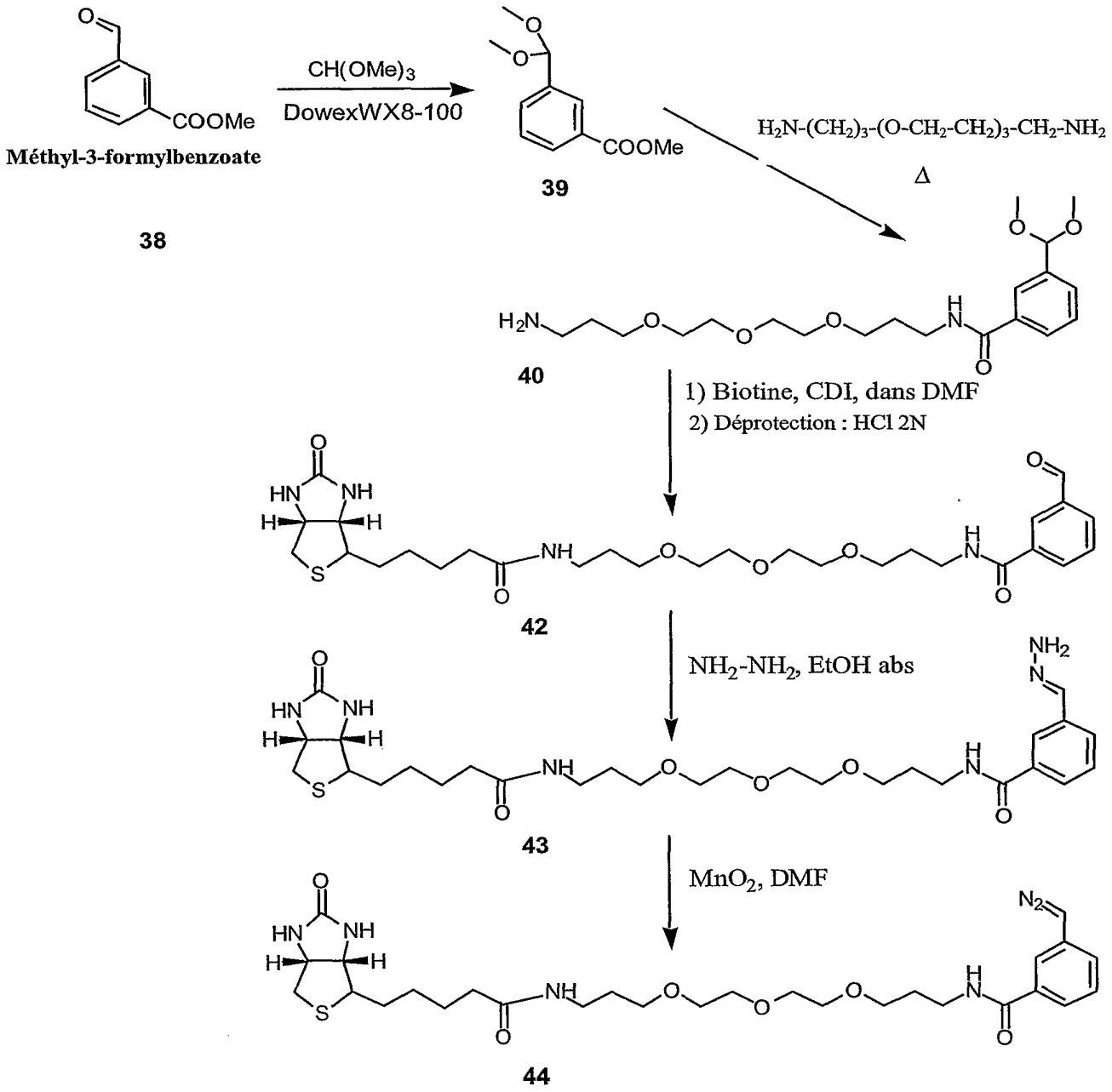
On solubilise l'hydrazone **36** (200 mg ; 0,337 mmol) dans 5 mL de DMF. On ajoute alors MnO_2 (450 mg ; 5,17 mmol). Après 15 min d'agitation à température ordinaire, le mélange est filtré sur millipore contenant de la célite (épaisseur : 2 cm) et du tamis moléculaire en poudre 3 Å (0,5 cm). Le mélange réactionnel est évaporé à sec. L'huile résiduelle obtenue est lavée à l'éther trois fois de suite, jusqu'à l'obtention d'une poudre. Le composé **37** (290 mg, 0,491 mmol) est obtenu sous forme d'un solide rose avec un rendement de 93 %.

RMN ^{1}H (300MHz, DMSO-d_6) δ = 8,33 (t, 1H, $NH_{2'}$) ; 7,83 (d, 2H, J = 8 Hz, Ar-$H_{2,6}$) ; 7,73 (t, 1H, $NH_{16'}$) ; 6,98 (d, 2H, J = 8 Hz, Ar-$H_{3,5}$) ; 6,39 (s large, 1H, N$\underline{H}_{B1}$) ; 6,33 (s large, 1H, N$\underline{H}_{B3}$) ; 4,30 (m, 1H, H_{B6a}) ; 4,12 (m, 1H, H_{B3a}) ; 3,51-3,45 (m, 16H, $H_{7', 8', 10', 11'}$ et $H_{5'}$ et $H_{15'}$ et $H_{3',13'}$) ; 3,07 (m, 1H, H_{B4}) ; 2,79 et 2,58 (système ABX, 2H, $^2J_{AB}$= 5 Hz, $^3J_{AX}$ = 12 Hz, $^3J_{BX}$ = 0 Hz, H_{B6}) ; 2,14 (s, 3H, CH_3) ; 2,04 (t, 2H, J = 8 Hz, H_{B10}) ;1,77 (m, 2H, $H_{4'}$) ; 1,62-1,48 (m, 6H, $H_{14'}$, $H_{B7, B9}$) ; 1, 31 (m, 2H, H_{B8}).

La réactivité du composé **37** a été testée sur 3'-Uridine-monophosphate et suivie par électrophorèse capillaire. Les conditions d'analyse sont celles de l'exemple 6.1. Les résultats montrent un temps de demi vie de 60 minutes.
Le réactif est stable à -20°C pendant au moins 1 mois.

Exemple 22.4 : Synthèse du *méta*-Bio-EG3-PDAM :

Schéma de synthèse :



- Protection du méthyl-3-formylbenzoate **38** :

La résine Dowex 50WX8-100 (2 g) est mise en solution dans 25 mL de MeOH et 25 mL d'orthoformiate de triméthyle puis laissé sous agitation pendant 15 min. Après décantation, on lave la résine deux fois de suite avec 20 mL de MeOH. On met ensuite cette résine dans 100 mL de MeOH, 50 mL de $CH(OMe)_3$ et on ajoute 7,12 g (43,4 mmol) de méthyl-3-formylbenzoate. La solution est laissée 15 min sous agitation, puis filtrée sur papier plissé avant évaporation. Le produit **39** (9 g ; 43,1 mmol) est isolé sous forme d'un liquide jaune clair avec un rendement de 99 %.

RMN ^1H (200MHz, $CDCl_3$) δ = 8,10 (s, H, Ar-H_2) ;7,9 (d, H, J = 8 Hz, Ar-H_4) ; 7,63 (d, H, J = 8 Hz, Ar-H_6) ; 7,42 (t, H, J = 8 Hz, Ar-H_5) ; 5,40 (s, 1H, CH) ; 3,90 (s, 3H, -CO-O-CH_3) ; 3,31 (s, 6H, -O-CH_3).

- Composé **40** :

On solubilise le composé **39** (2 g ; 9,5 mmol) dans 10,4 mL (47,6 mmol) de 4,7,10-Trioxa-1,13-tridécanediamine. La solution obtenue est chauffée à 165°C pendant 2 h. Le mélange est ensuite dissous dans 80 mL de DCM et lavé 4 fois avec 20 mL d'eau. Après séchage sur $MgSO_4$ et évaporation, le produit **40** est isolé avec un rendement de 60 % (2,27 g ; 5,69 mmol).

RMN ^1H (200MHz, $CDCl_3$) δ = 7,84 (s, H, Ar-H_2) ;7,75 (d, H, J = 8 Hz, Ar-H_4) ; 7,53 (d, H, J = 8 Hz, Ar-H_6) ; 7,39 (t, H, J = 8 Hz, Ar-H_5) ; 5,38 (s, 1H, CH) ; 3,64-3,43 (m, 14H, $H_{7',8',10',11'}$ et $H_{5',13'}$ et $H_{3'}$) ; 3,29 (s, 6H, -O-CH_3) ; 2,72 (m, 2H, $H_{15'}$) 1,87 (m, 2H, $H_{4'}$) ; 1,64 (m, 2H, $H_{14'}$) ; 1,30 (s large, 2H, NH_2).

- Composé biotinylé **41** :

La D-biotine (344 mg ; 1,40 mmol) est mise en suspension dans 4 mL de DMF puis 250 mg (1,54 mmol) de CDI sont ajoutés. Cette solution est laissée sous agitation pendant 30 min à température ambiante. Le composé **40** (616 mg ; 1,54 mmol) est dissous dans 2 mL de DMF, puis ajouté petit à petit à la solution précédente. Le mélange ainsi obtenu est laissé sous

agitation pendant 50 min à température ambiante. Après évaporation, une purification sur colonne par chromatographie-flash (colonne de 30 mm de diamètre) est réalisée avec 750 mL de MeOH-DCM 10 %, puis avec 250 mL de MeOH-DCM 15 %. Les fractions correspondant au produit **41** sont rassemblées puis évaporées à sec pour donner 740 mg d'huile avec un rendement estimé à 50 %.

RMN ^{1}H (200MHz, $CDCl_3$) δ = 7,87 (s, H, Ar-H_2) ;7,78 (d, H, J = 8 Hz, Ar-H_4) ; 7,65 (s, 1H, $H_{imidazole}$) ; 7,53 (d, H, J = 8 Hz, Ar-H_6) ; 7,39 (t, H, J = 8 Hz, Ar-H_5) ; 7,07 (s, 2H, $H_{imidazole}$) ; 6,65 (t, 1H, $NH_{16'}$) , 5,95 (s large, 1H, N$\underline{H}_{B1}$) ; 5,38 (s, 1H, CH) ; 5,15 (s large, 1H, N$\underline{H}_{B3}$) ; 4,43 (m, 1H, H_{B6a}) ; 4,27 (m, 1H, H_{B3a}) ; 3,59-3,44 (m, 14H, $H_{7',8',10',11'}$ et $H_{5',13'}$ et $H_{3'}$) ; 3,29 (m, 8H, $H_{15'}$ et 2-O-CH_3) ;3,07 (m, 1H, H_{B4}) ; 2,84 et 2,66 (système ABX, 2H, $^2J_{AB}$= 5 Hz, $^3J_{AX}$ = 12 Hz, $^3J_{BX}$ = 0 Hz, H_{B6}) ; 2,13 (t, 2H, J = 8 Hz, H_{B10}) ;1,85 (m, 2H, $H_{4'}$) ; 1,66 (m, 2H, $H_{14'}$) ; 1,40-1,37 (m, 6H, $H_{B7, B8, B9}$).

- Composé aldéhydique **42** :

L'acétal **41** est dissous dans 20 mL de chloroforme, puis 5 mL de HCl 2N sont ajoutés . Le mélange diphasique est vigoureusement agité pendant 15 min. On récupère la phase organique qu'on sèche sur $NaHCO_3$ anhydre. On filtre, on évapore et le composé **42** est obtenu sous forme d'une huile jaune (593 mg ; 1,02 mmol) avec un rendement global de 87 % à partir de la biotine.

RMN ^{1}H (300MHz, $CDCl_3$) δ = 10,04 (s, 1H, CHO) ;8,34 (s, H, Ar-H_2) ; 8,16 (d, H, J = 8 Hz, Ar-H_4) ; 7,96 (d, H, J = 8 Hz, Ar-H_6) ; 7,72 (t, 1H, $NH_{2'}$) ; 7,39 (t, H, J = 8 Hz, Ar-H_5) ; 6,51 (t, 1H, $NH_{16'}$) ; 6,00 (s large, 1H, N$\underline{H}_{B1}$) ; 5,30 (s large, 1H, N$\underline{H}_{B3}$) ; 4,46 (m, 1H, H_{B6a}) ; 4,27 (m, 1H, H_{B3a}) ; 3,66-3,56 (m, 10H, $H_{7', 8',10',11'}$ et $H_{5'}$) ; 3,50-3,29 (m, 4H, $H_{3',13'}$) ; 3,28 (m, 2H, $H_{15'}$) ; 2,95 (m, 1H, H_{B4}) ; 2,84 et 2,71 (système ABX, 2H, $^2J_{AB}$= 5 Hz, $^3J_{AX}$ = 12 Hz, $^3J_{BX}$ = 0 Hz, H_{B6}) ; 2,15 (t, 2H, J = 8 Hz, H_{B10}) ;1,89 (m, 2H, $H_{4'}$) ; 1,72-1,63 (m, 6H, $H_{14'}$, $H_{B7, B9}$) ; 1, 23 (m, 2H, H_{B8}).

- Composé hydrazone **43** :

L'aldéhyde **42** (593 mg ; 1,02 mmol) est dissous dans 10 mL d'éthanol absolu. L'hydrazine (400 µL ; 8,19 mmol) est ajoutée, puis le mélange réactionnel est chauffé à reflux pendant 50 min. L'huile jaune obtenue après évaporation, est triturée avec de l'éther jusqu'à l'obtention d'une poudre beige correspondant au produit **43** (404 mg ; 0,68 mmol) avec un rendement de 66 %. Une purification sur colonne par chromatographie-flash (colonne de 15 mm de diamètre) est réalisée ensuite sur un échantillon de 150 mg (0,253 mmol) avec 200 mL de MeOH-DCM 20% . Les fractions sont rassemblées puis évaporées à sec pour donner 144 mg de produit **43** avec un rendement de 76 %.

RMN ^{1}H (300MHz, $CDCl_3$) δ = 7,95 (s, H, Ar-H_2) ; 8,16 (d, H, J = 8 Hz, Ar-H_4) ; 7,76 (s, 1H, CH) ; 7,96 (d, H, J = 8 Hz, Ar-H_6) ; 7,38 (t, H, J = 8 Hz, Ar-H_5) ; 6,45 (t, 1H, $NH_{16'}$) ; 5,98 (s large, 1H, N$\underline{H}_{B1}$) ; 5,72 (s large, 2H, NH_2) ; 5,18 (s large, 1H, N$\underline{H}_{B3}$) ; 4,44 (m, 1H, H_{B6a}) ; 4,26 (m, 1H, H_{B3a}) ; 3,62-3,56 (m, 10H, $H_{7', 8', 10', 11'}$ et $H_{5'}$) ; 3,48-3,45 (m, 4H, $H_{3',13'}$) ; 3,27 (m, 2H, $H_{15'}$) ; 3,07 (m, 1H, H_{B4}) ; 2,84 et 2,68 (système ABX, 2H, $^2J_{AB}$= 5 Hz, $^3J_{AX}$ = 12 Hz, $^3J_{BX}$ = 0 Hz, H_{B6}) ; 2,11 (t, 2H, J = 8 Hz, H_{B10}) ;1,86 (m, 2H, $H_{4'}$) ; 1,72-1,59 (m, 6H, $H_{14'}$, $H_{B7, B9}$) ; 1, 21 (m, 2H, H_{B8}).

- <u>Composé diazo **44** :</u>

On solubilise l'hydrazone **43** (100 mg ; 0,187 mmol) dans 4 mL de DMF. On ajoute alors MnO_2 (200 mg ; 2,3 mmol). Après 13 min d'agitation à température ordinaire, le mélange est filtré sur millipore contenant de la célite (épaisseur : 2 cm) et du tamis moléculaire en poudre 3 Å (0,5 cm). Le mélange réactionnel est évaporé à sec. L'huile résiduelle obtenue est lavée à l'éther, trois fois de suite. Le composé **44** (290 mg, 0,491 mmol) est obtenu sous forme d'un solide orange avec un rendement de 83 %.

RMN ^{1}H (300MHz, DMSO-d_6) δ = 8,39 (t, 1H, $NH_{2'}$) ; 7,78 (t, 1H, $NH_{16'}$) ; 7,39-7,34 (m, Ar-H) ; 7,09 (d, Ar-H) ; 6,38 (s large, 1H, N$\underline{H}_{B1}$) ; 6,32 (s large, 1H, N$\underline{H}_{B3}$) ; 5,78 (s, 1H, C$\underline{H}$-N_2) ; 4,27 (m, 1H, H_{B6a}) ; 4,11 (m, 1H, H_{B3a}) ; 3,51-3,44 (m, 10H, $H_{7', 8', 10', 11'}$ et $H_{5'}$) ; 3,37 (m, 2H, $H_{15'}$) ; 3,32 (m, 4H, $H_{3',13'}$) ; 3,05 (m, 1H, H_{B4}) ; 2,79 et 2,58 (système ABX, 2H,

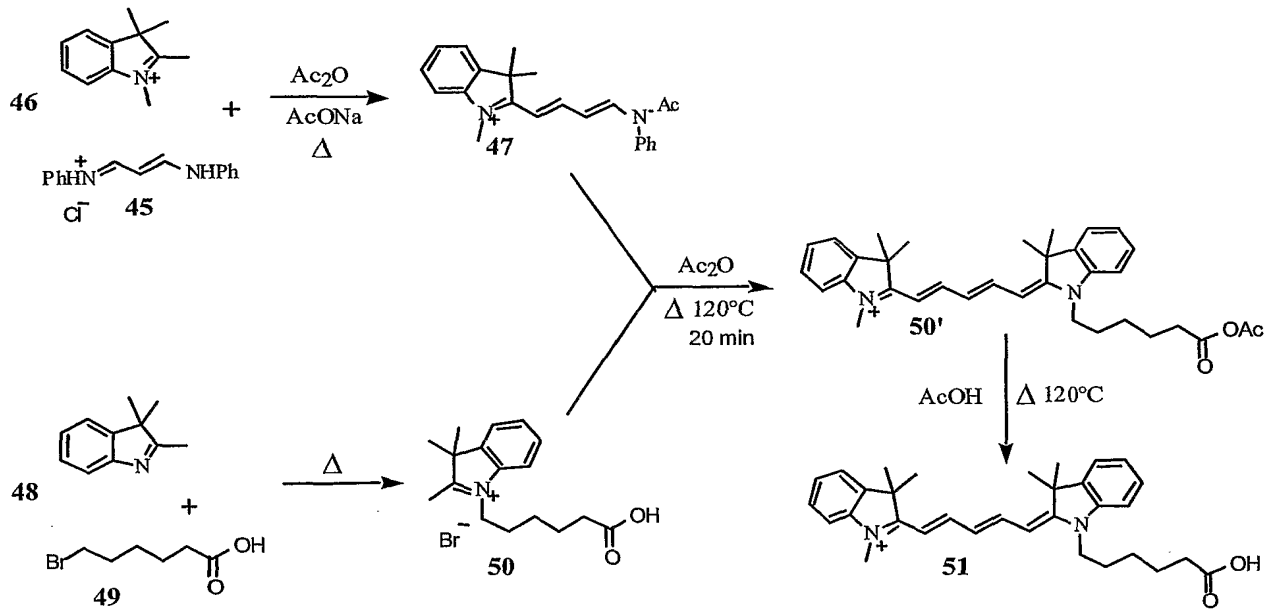
$^2J_{AB}$= 5 Hz, $^3J_{AX}$ = 12 Hz, $^3J_{BX}$ = 0 Hz, H_{B6}) ; 2,02 (t, 2H, J = 8 Hz, H_{B10}) ;1,69 (m, 2H, $H_{4'}$) ; 1,59-1,48 (m, 6H, $H_{14'}$, $H_{B7, B9}$) ; 1, 25 (m, 2H, H_{B8}).

La stabilité du produit est supérieure à 1 mois à -20°C.

Exemple 23 : Synthèse du *para*-Cy5-EG3-PDAM :

Comme cela a déjà été évoqué dans l'exemple 2, la biotine peut être remplacée par un autre marqueur tel que le Cy5. Cet exemple montre que l'on peut également relier la fonction diazo, portée par le PDAM, à ce marqueur Cy5 par l'intermédiaire d'un bras de liaison polyéthylène glycol.

Schéma de synthèse : Le contre ion I^- n'est pas représenté sur les formules 46, 47, 50' et 51.



- Iodure de 2-[4-(*N*-acétyl-*N*-phénylamino)buta-1,3-diényl]-1,2,3,3-tétraméthyl[3H]indolium **47 :**

Le mélange du monochlorhydrate de malonaldehyde-bis(phénylimine) **45** (13 g ; 50,2 mmol), de NaOAc (6,0 g ; 69,7 mmol) et de l'iodure de 1,2,3,3- tétraméthyl[3H]indolium **46** (3,01 g ; 10 mmol) dans de l'anhydride acétique (50 mL) est chauffé à 100 °C pendant 20 min précisément. Après refroidissement, on ajoute de l'éther (350 mL) et le solide brun précipité est filtré, lavé avec de l'éther (3 × 100 mL). Le solide est redissout dans 150 mL de CH_2Cl_2, filtré (élimination des sels minéraux) puis précipité avec 350 mL d'éther pour donner un solide brun (3,54 g, 54 %).

RMN ^1H ($CDCl_3$) : δ = 8,64 (d ; 1H ; J = 12 Hz ; 1-H) ; 8,14 (t ; 1H ; J = 16 ; 12 Hz ; 3-H) ; 7,63-7,19 (m ; 9H) ; 6,90 (d ; 1H ; J = 15 Hz ; 4-H) ; 5,82 (t ; 1H ; J = 12 ; 13 Hz ; 2-H) ; 4,06 (s ; 3H ; NCH_3) ; (2,16 (s ; 3H ; -$COCH_3$) ; 1,74 (s ; 6H ; CH_3).

- Bromure de 1-(5-carboxypentyl)-2,3,3-triméthyl[3H]indolium **50** :

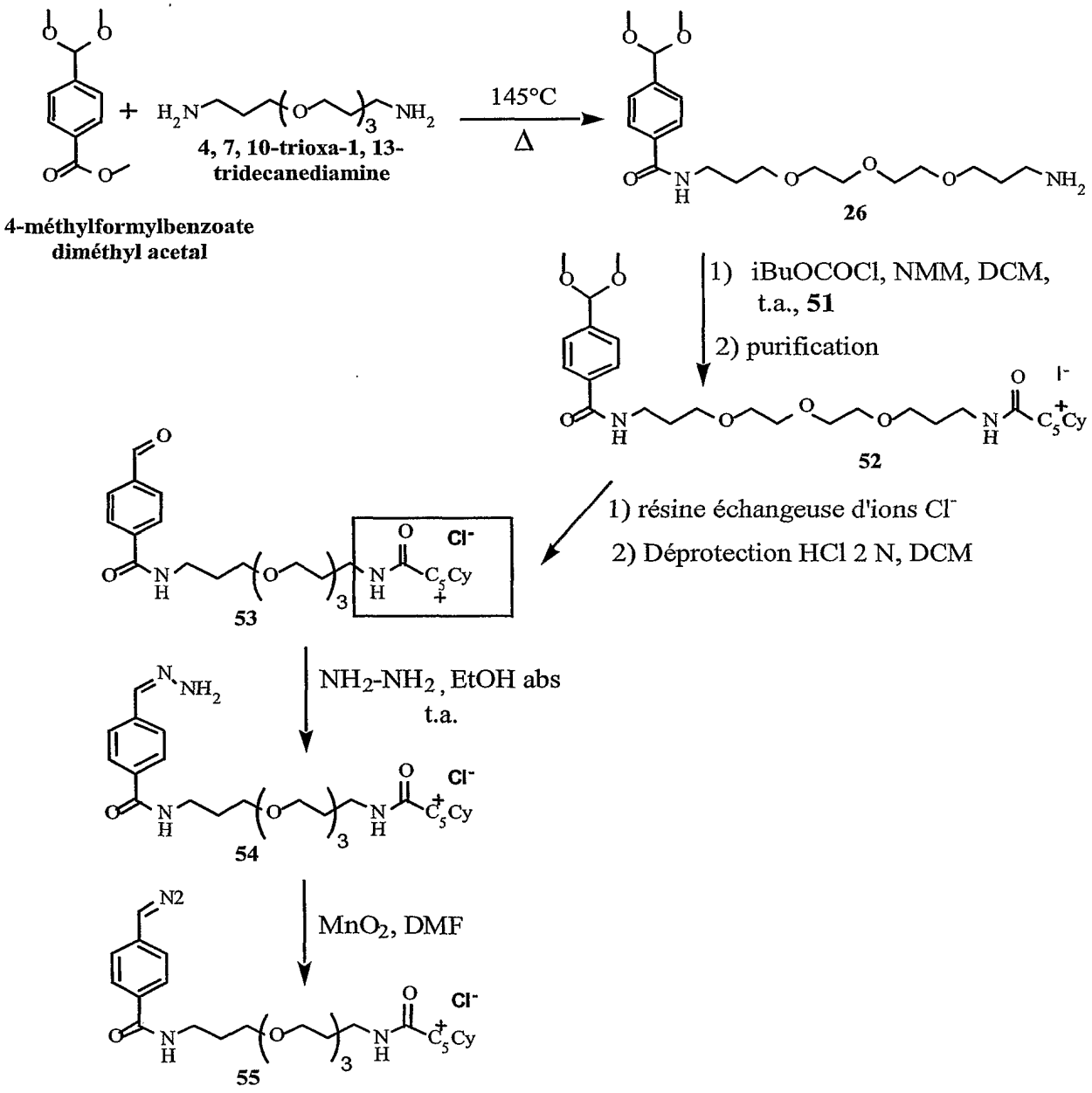
On mélange le 2,3,3-triméthylindole **48** (10,0 g ; 62,8 mmole) et l'acide 6-bromohexanoïque **49** (12,3 g ; 62,8 mmole) sans solvant et chauffe à 110 °C pendant 12 h sous argon. Le mélange réactionnel pâteux rouge violet est lavé avec de l'acétate d'éthyle (2 × 60 mL, on triture la pâte avec la spatule et on décante le surnageant), puis avec de l'acétone (50 mL, la pâte solidifie). Le solide rose est filtré puis séché sous vide (16,0 g ; 73 %).

- Composé Cy5COOH **51** :

Le mélange de l'iodure **47** (2,5 g ; 5,3 mmole), du bromure **50** (1,87 g ; 5,3 mmole) et de NaOAc (1,08 g ; 12,1 mmole) dans de l'anhydride acétique (11 mL) est chauffé à 120 °C pendant 25 min. Après refroidissement, on ajoute de l'éther (200 mL) et le précipité est filtré et lavé avec de l'éther (3 × 50 mL). Le solide correspondant au produit **50'** est dissous dans 100 mL de CH_2Cl_2 puis évaporé. On le dissout ensuite dans 15 mL d'acide acétique et l'on agite pendant 30 min à 120 °C. Le précipité correspondant au produit **51** est ainsi obtenu après ajout de 200 mL d'éther et filtration sur fritté, avec un rendement de 84 % (2,71 g ; 4,44 mmol).

RMN ^1H ($CDCl_3$) : δ = 8,03 (t ; 2H ; J = 10 ; 11 Hz, 2-H, 4-H) ; 7,38-6,91 (m ; 9H ; Ar-H, 3-H) ; 6,41 (d ; 1H ; J = 14 Hz ; 1-H) ; 6,31 (d ; 1H ; J = 13 Hz ; 5-H) ; 4,07 (t ; 2H, J = 7 ; 7 Hz ; α-CH_2) ; 3,68 (s ; 3H ; NCH_3) ; 2,47 (t ; 2H, J = 7 ; 7 Hz ; ε-CH_2) ; 1,71 (m ; 18H ; CH_3, β,γ et δ-CH_2).

5



- Couplage du composé **26** avec le Cy5COOH **51** (produit **52**) :

À une solution de Cy5COOH **51** (1,5 g ; 2,46 mmole) dans 15 mL de CH_2Cl_2, on ajoute *N*-méthylmorpholine (NMM, 405 µL ; 3,68 mmole). La solution est refroidie avec un bain de glace et mise sous argon, puis on ajoute le chloroformiate d'isobutyle (494 µL ; 3,81 mmole). Après 10 min d'agitation, on additionne l'amine **26** (1,86 mg ; 4,67 mmole) diluée dans 8 mL de CH_2Cl_2. On laisse le mélange sous agitation à température ambiante pendant 1 h 30. On ajoute 20 mL de CH_2Cl_2 et on lave avec 25 mL de $NaHCO_3$ (1 N) trois fois de suite. Après séchage sur Na_2CO_3, on filtre la solution pour récupérer la phase dichlorométhane que l'on évapore.

Une purification sur colonne par chromatographie-flash (colonne de 45 mm de diamètre, fractions de 20 mL)) est réalisée avec comme éluant du MeOH-DCM 10% . Les fractions correspondant au produit **52** sont rassemblées puis évaporées à sec pour donner un solide bleu qui est dissous dans CH_2Cl_2. Le produit **52** est alors précipité et lavé à l'éther pour donner un produit bleu avec un rendement de 72 % (1,45 g ; 1,77 mmol).

Le produit **52** (iodure) est ensuite dissous dans 54 mL de méthanol puis passé sur colonne d'amberlite IRA900 (Cl^- ; 15 g). La solution méthanolique recueillie est évaporée à sec pour donner une huile collante que l'on redissout dans CH_2Cl_2 . L'évaporation permet l'obtention du produit **52'** avec un rendement de 87 %.

● Aldéhyde **53** :

L'acétal **52'** est dissous dans 10 mL de DCM, puis 10 mL de HCL 2N sont ajoutés. La solution est laissée sous forte agitation pendant 3h 30 min. Après ajout de 20 mL de DCM, la phase dichlorométhane est récupérée puis séchée sur $NaHCO_3$. Le produit obtenu après évaporation est lavé avec de l'éther pour donner l'aldéhyde **53** avec un rendement de 90 % (1,18g ; 1,46 mmol).

● Hydrazone **54** :

On dissout l'aldéhyde **53** (200 mg ; 0,247 mmole) dans 1 mL d'éthanol absolu et on ajoute l'hydrazine monohydratée (15,6 µL ; 0,321 mmole). La solution est agitée à température

ambiante pendant 30 min. On ajoute 8 mL d'éther ; on lave à l'éther par décantation, trois fois de suite puis on sèche sous vide. On obtient 172 mg d'hydrazine **54** (0,209 mmol ; 85 % de rendement) que l'on conserve au congélateur.

● Diazo **55** :

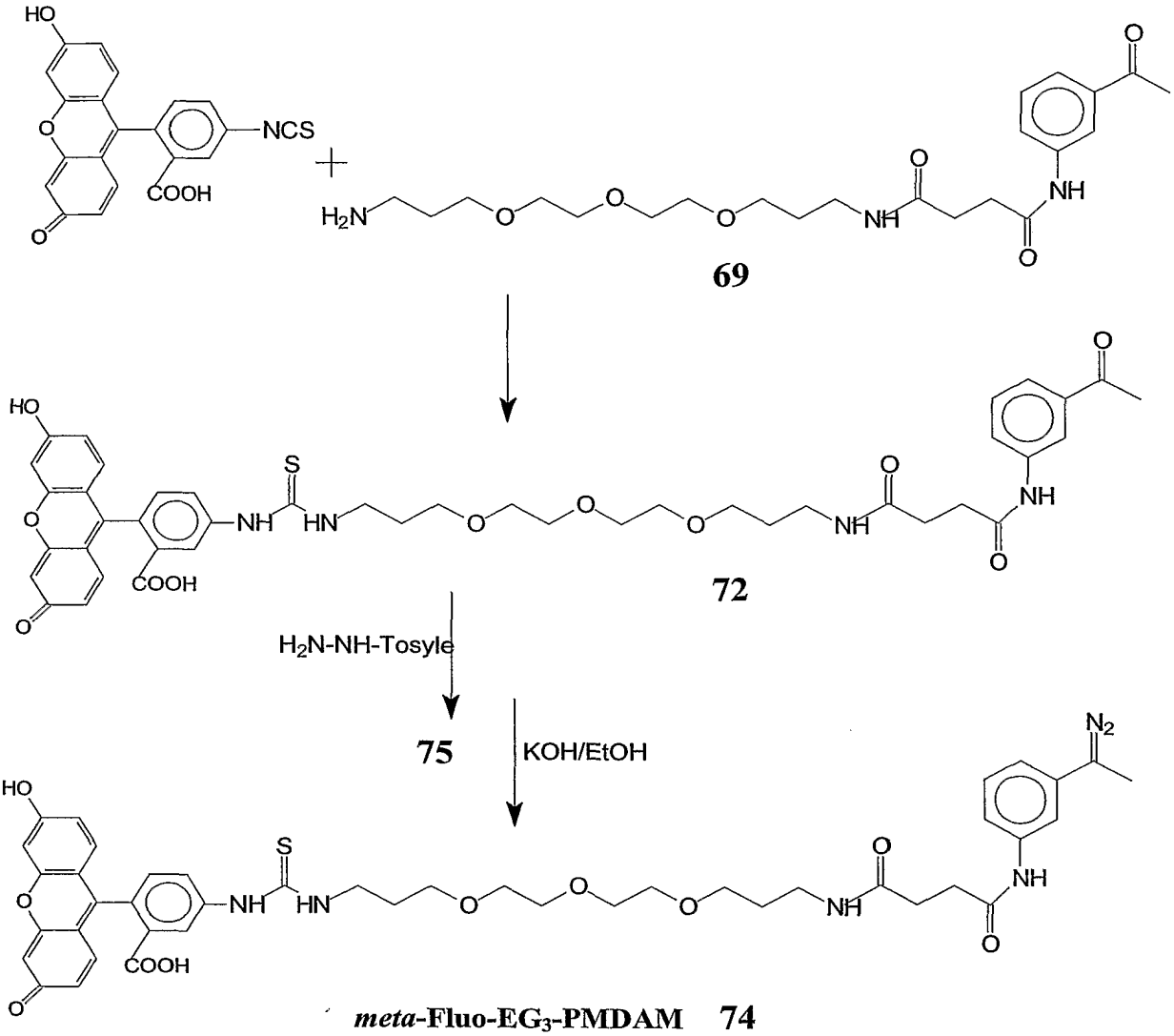
À une solution de 20 mg (0,243 mmol) d'hydrazone **54** dans 2 mL de DMF, on ajoute 100 mg de MnO_2 et agite vigoureusement pendant 5 min sous argon à température ambiante. On filtre la suspension à travers une couche du célite (épaisseur : 2 cm) et du tamis moléculaire en poudre 3 Å (0,5 cm) et lave avec DMF. On évapore la solution puis on triture à l'éther. Le solide ainsi obtenu est séché. On obtient 18 mg (0,185 mmol ; 76%) de diazo **55**.

La stabilité du réactif est supérieure à 1 mois à -20°C.

Exemple 24 : Synthèse du *méta*-Fluo-EG3-PMDAM :

Comme cela a déjà été évoqué dans l'exemple 23, la biotine peut être remplacée par un autre marqueur. Cet exemple montre que l'on peut également relier la fonction diazo, portée par le PMDAM, à ce marqueur fluorescéine par l'intermédiaire d'un bras de liaison polyéthylène glycol.

Schéma de synthèse :



***meta*-Fluo-EG_3-PMDAM 74**

Composé **72** :

On solubilise l'isothiocyanate de fluorescéine- (250 mg, 0,64 mmol) dans 1,6 ml de DMF anhydre, avec 2% de pyridine, sous argon. On ajoute le produit **69** (0,356 g, 0,81 mmol) dissout dans 1,6 ml de DMF anhydre. On laisse réagir 3,5 h à température ambiante, puis on évapore le DMF et on reprend dans 25 ml de H_2O. On fait trois extractions avec 50 ml de CH_2Cl_2, et la phase aqueuse est évaporée. On obtient 255 mg (48%) du produit **72**.

Composé *méta*-Fluo-EG_3-Hydrazone-Tosyle **75** :

On dissout le composé **72** (255 mg, 0,31 mmol) dans 1,5 ml d'éthanol à reflux. On ajoute la *p*-toluen-sulfonyl-hydrazine (69,2 mg, 0,37 mmol) dans 1,5 ml d'éthanol et on laisse réagir pendant 6 h. On évapore a sec, et on lave le solide avec CH_2Cl_2, H_2O et éther. On obtient 18,5 mg (74%) du produit **75** sous forme d'une poudre orange.
RMN ^1H (200 MHz, DMSO-d_6): d = 1,6-1,8 (m, 4H); 2,13 (s, 1H); 2,28 (s, 1H); 2,36 (s, 1H); 2,80 (m, 1H); 3,07 (m, 2H); 3,46 (m, 12H); 6,5-6,7 (m, 6H); 7,1-8,3 (m, 9H).

Composé *méta*-Fluo-EG_3-PMDAM **74** :

On solubilise l'hydrazone **75** (176 mg, 0,18 mmol) dans 720 µl d'une solution de KOH à 10% dans du méthanol anhydre. La solution est laissée a reflux pendant 3 h. On laisse refroidir la solution et un précipité apparaît. La solution est filtrée et évaporée a sec. Le résidu est lavé a l'éther et séché.
L'analyse par RMN montre la disparition des signaux a 2,36 et 2,13 ppm (correspondant aux Méthyles du tosyle et de l'hydrazone) et l'apparition d'un pic à 1,96 ppm (correspondant au méthyle du diazo).

Exemple 25 : Intermédiaire diazométhyle permettant un marquage ultérieur :

Il peut être intéressant d'avoir non pas un marquage direct par le réactif de marquage diazométhyle portant aussi le marqueur R^2 mais de procéder en deux étapes avec un marquage indirect. Dans ce cas le réactif de marquage comprenant la fonction diazométhyle est dit préfonctionnalisé c'est à dire qu'il comprend aussi une fonction chimique apte à réagir ultérieurement avec des marqueurs directs ou indirects. La pré-fonctionnalisation peut intervenir par introduction d'une fonction réactive de covalence au sein du réactif de marquage qui peut réagir avec une fonction anti-réactive de covalence du marqueur direct ou indirect. Ces fonctions peuvent être constituées par une fonction chimique organique électrophile et une fonction chimique organique nucléophile, ou inversement.

Un exemple d'une telle stratégie de marquage est illustrée par le schéma ci-dessous



Dans lequel le réactif de marquage comprend, outre une fonction diazométhyle, une fonction W^1 électrophile ou nucléophile susceptible de réagir avec un marqueur R^2 comprenant une fonction W^2 complémentaire de W^1.

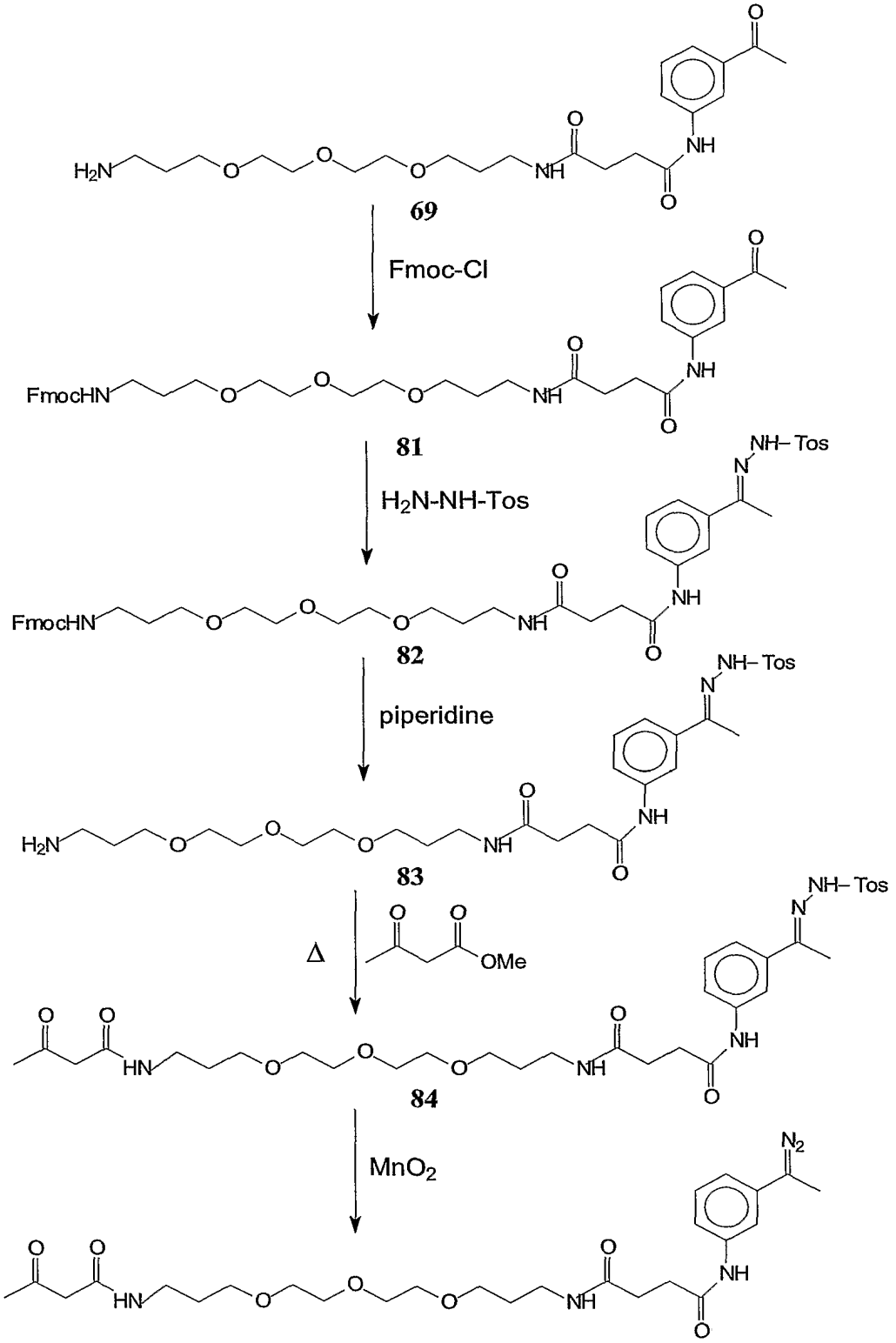
Par exemple si W^1 est une fonction méthylcétone ou aldéhyde, W^2 est une fonction alcoxyamine.

Dans un procédé de marquage d'une molécule biologique comme un acide nucléique, l'acide nucléique est mis en contact avec le réactif de marquage coprenant la fonction diazométhyle et dans une étape ultérieure, le marqueur W^2- R^2 réagit avec l'acide nucléique par l'intermédiaire de la fonction W^1.

L'une des utilisations est par exemple constituée par un procédé d'amplification d'une séquence d'un acide nucléique ou par un procédé d'amplification du signal. Des informations complémentaires sur ce type de marquage peuvent être trouvées dans la demande de brevet WO-A-98/05766, sous priorité du 2 août 1996. et dans la demande de brevet WO-A-00/40590 sous priorité du 05 Janvier 1999.

Exemple 25.1 : Synthèse du MeCO-PMDAM :

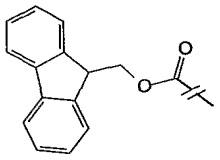
Schéma de synthèse :



MeCO-PMDAM **85**

Le produit **85**, dont la synthèse est décrite dans cet exemple, permet de réaliser le marquage des acides nucléiques naturels, grâce à la réactivité de la fonction diazométhyl avec les groupements phosphate, et d'introduire ainsi une fonction méthylcétone, qui peut être utilisée par la suite pour introduire une molécule détectable (fluorescente, biotine) possédant un groupement alcoxyamine.

Cette synthèse est basée sur des méthodes connues et utilisées en routine en chimie. Le produit de départ est le même que pour la synthèse des marqueurs **71** et **74**. La première étape consiste en la protection de l'amine terminale par le Fluorenylmethyl-formate (Fmoc, **99**). Le choix de ce groupement protecteur est basé sur sa stabilité et conditions de clivage.

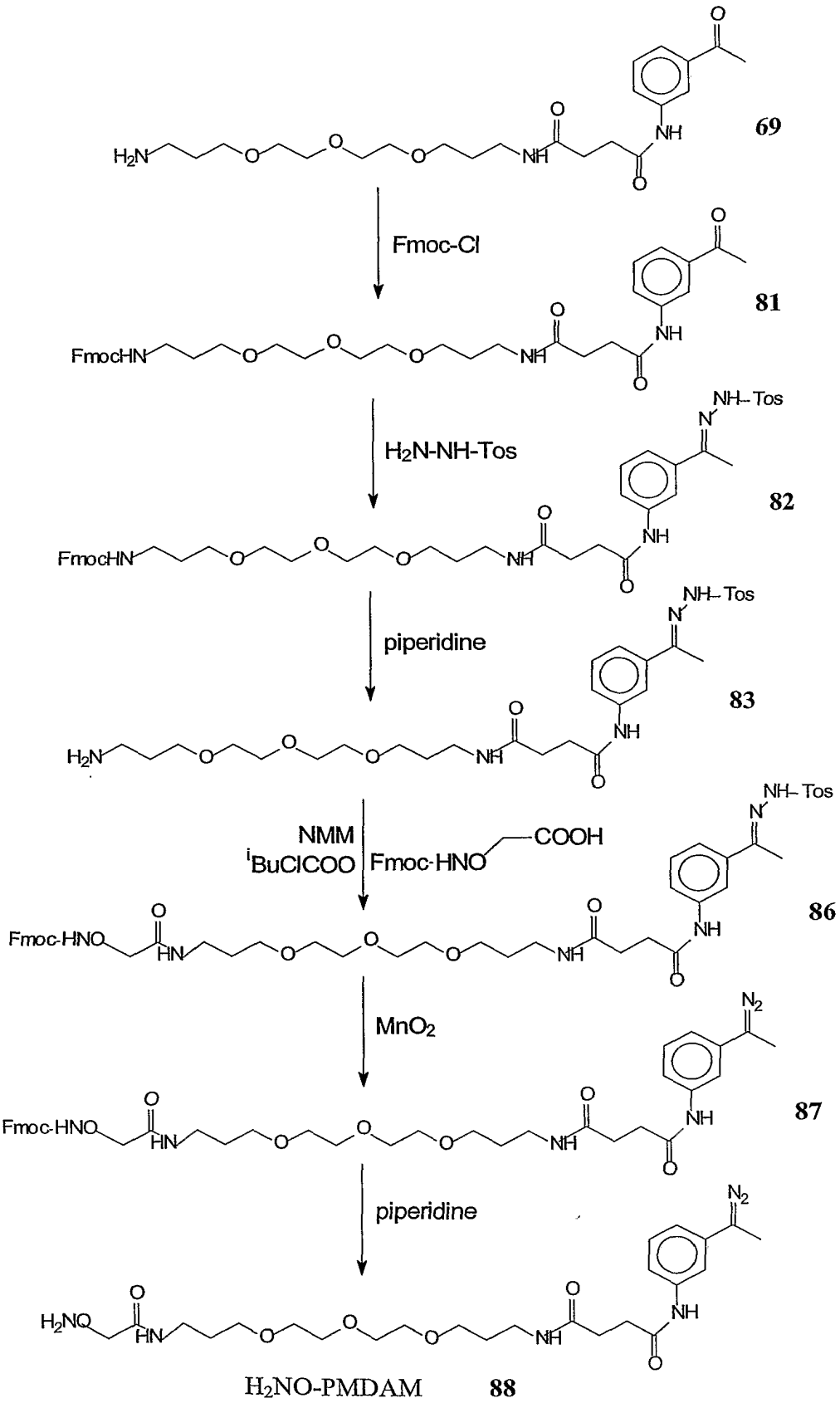


99

Après formation de l'hydrazone protégée **82** par la méthode utilisée précédemment (exemple *méta*-Fluo-EG_3-PMDAM,), l'amine terminale est déprotégée dans des conditions basiques douces qui garantissent la stabilité de l'hydrazone. L'acétoacétate de méthyle est utilisé pour créer la fonction méthylcetone par une réaction d'acylation de l'amine terminale (voir formation des composés **26** et **36**). La formation du diazométhyl est ensuite réalisée par l'une des méthodes décrites précédemment.

Exemple 25.2 : Synthèse du H_2NO-PMDAM :

Schéma de synthèse :


69
Fmoc-Cl
81
H2N-NH-Tos
82
piperidine
83
NMM
iBuClCOO
Fmoc-HNO
COOH
86
MnO2
87
piperidine
H2NO-PMDAM
88

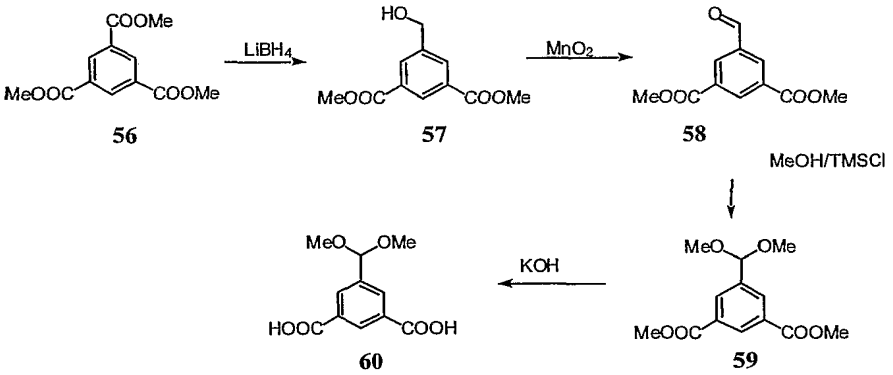
Le produit **88**, dont la synthèse est décrite dans cet exemple, permet de réaliser le marquage des acides nucléiques naturels, grâce a la réactivité de la fonction diazométhyl avec les groupements phosphate, et d'introduire ainsi une fonction alcoxyamine, qui peut être utilisée par la suite pour introduire une molécule détectable (fluorescente, biotine) possédant un groupement méthylcétone.

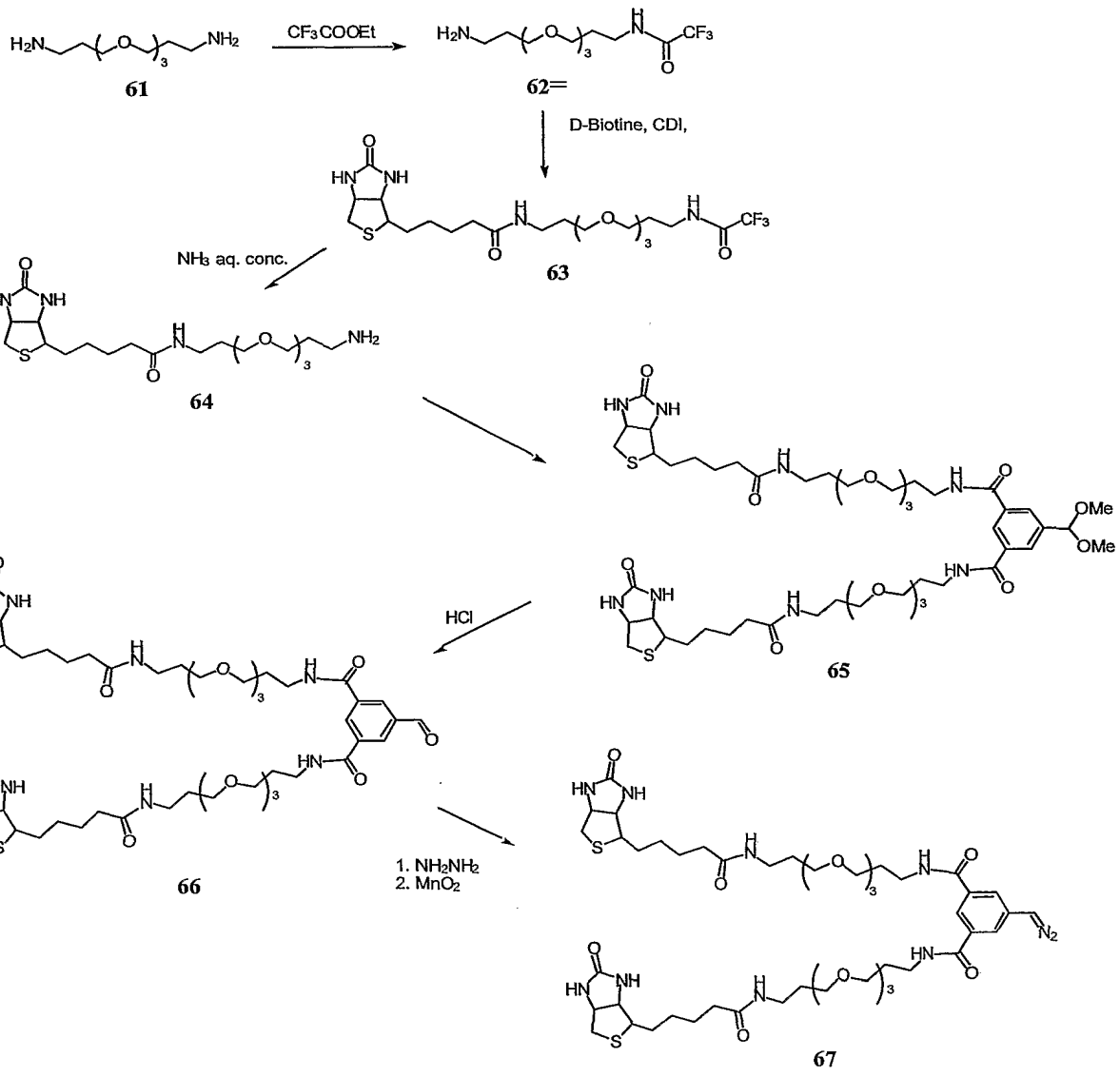
Cette synthèse est basée sur le modèle de la précédente, c'est-à-dire l'utilisation du précurseur **69**, du Fmoc pour la protection de l'amine et du tosyle pour la protection de l'hydrazone. L'introduction de la fonction alcoxyamine (composé **86**) se fait par l'usage du carboxyméthoxylamine (commercial) protégé par la fonction Fmoc (Thèse E. Trévisiol, LEDSS Grenoble, 1999). Etant donné les conditions douces de la déprotection finale (composé **88**), celle ci est réalisée de suite après la formation du diazométhyl.

Exemple 26 : Préparation de dérivés PDAM permettant l'amplification du signal :

Exemple 26.1 : Synthèse de marqueurs bis-biotinylés tels que le [Bio-EG3]2-PDAM :

Schéma de synthèse :


COOMe
MeOOC
COOMe
LiBH4
HO
MnO2
O
56
57
58
MeOH/TMSCl
MeO
OMe
KOH
HOOC
COOH
60
59


H2N
NH2
CF3COOEt
CF3
61
62=
D-Biotine, CDI,
HN
NH
S
O
63
NH3 aq. conc.
64
65
OMe
HCl
66
1. NH2NH2
2. MnO2
N2
67

● Réduction du benzène-1,3,5-tricarboxylate de triméthyle **56** en alcool **57** :

On dissout le triester **56** (12,6 g ; 50,0 mmol) dans 100 mL de THF puis ajoute 1,1 g (50,5 mmol) de $LiBH_4$ à température ambiante. La solution rouge est chauffée à 40-45°C sous agitation sous argon pendant 1 h. Après refroidissement (glace), on détruit l'hydrure en excès avec précaution (dégagement de H_2) par l'addition d'eau (200 mL) puis de HCl 2N (30 mL). On observe le virage de couleur en jaune claire. On extrait cette solution avec CH_2Cl_2 (100mL puis 3 fois 50 mL), la phase organique est lavée avec $NaHCO_3$ anhydre, séchée sur $MgSO_4$ puis le solvant est évaporé jusqu'à l'obtention d'une huile (11,1 g). Par chromatographie flash sur colonne de silice (diamètre = 40 mm, éluant : acétate d'éthyle/cyclohéxane = 1/1), on obtient l'alcool **57** (6,38 g, 57 %).

RMN ^1H (200 MHz, $CDCl_3$) : δ = 8,53(t, 1H, J = 2 Hz) ; 8,18 (d, 2H, J = 2 Hz) ; 4,76 (s, 2H) ; 3,91 (s, 6H) ; 2,30 (s, 1H).

● Oxydation de l'alcool **57** en aldéhyde **58** :

On dissout l'alcool **57** (5,86 g ; 26,1 mmol) dans 100 mL de THF puis ajoute 40,0 g de MnO_2 petit à petit en 5 min à température ambiante. La solution est laissée une nuit sous agitation sous argon. On filtre la solution à travers un entonnoir Büchner muni d'une couche de Célite 545, on lave avec CH_2Cl_2 puis évapore les solvants. Le solide brut (4,4 g) est purifié par chromatographie flash sur une colonne de silice (diamètre = 50 mm, éluant : acétate d'éthyle/cyclohéxane 3/7). On obtient 3,44 g (59 %) de l'aldéhyde **58**.

RMN ^1H (200 MHz, $CDCl_3$) : δ = 10,11 (s, 1H) ; 8,89(t, 1H, J = 1 Hz) ; 8,69 (d, 2H, J = 1 Hz) ; 3,98 (s, 6H).

● Formation de l'acétal **59** :

On dissout l'aldéhyde **58** (3,21 g ; 14,4 mmol) dans 30 mL de méthanol puis ajoute 6,0 mL de TMSCl. La solution est laissée à température ambiante sous agitation sous argon pendant 1 h.

On dilue la solution avec 200 mL de CH_2Cl_2, et agite avec 1 M $NaHCO_3$ (100 mL) (attention dégagement de CO_2). On sépare les deux phases, extrait la phase aqueuse trois fois avec CH_2Cl_2 (25 mL), on rassemble les phases organiques, on sèche sur $MgSO_4$ puis évapore le solvant. On obtient 3,55 g (92 %) de l'acétal **59**.

RMN ^1H (200 MHz, $CDCl_3$) : δ = 8,63(t, 1H, J = 2 Hz) ; 8,29 (d, 2H, J = 2 Hz) ; 5,45 (s, 2H) ; 3,93 (s, 6H) ; 3,32 (s, 6H).

- Hydrolyse du diester **59** en diacide **60** :

On dissout le diester **59** (3,18 g ; 11,9 mmol) dans 10 mL de THF puis ajoute une solution de KOH (2,0 g, pastille 85 %) dans 10 mL de méthanol. Après 15 min à température ambiante, on évapore les solvants. Le résidu est dissous dans H_2O (50 mL). On ajoute H_3PO_4 (environ 2,5 mL, 85 %) jusqu'à pH 3 et le précipité blanc est filtré sur le fritté (#3), lavé avec de l'eau et séché sous vide. On obtient 2, 59 g (91 %) du diacide **60**.

RMN ^1H (200 MHz, DMSO-d_6) : δ = 8,43 (t, 1H, J = 1 Hz) ; 8,15 (d, 2H, J = 1 Hz) ; 5,53 (s, 1H) ; 3,27 (s, 6H).

- Trifluoracétamide **62** :

On dissout la diamine **61** (66 g ; 0,30 mol) dans 250 mL de CH_2Cl_2 puis ajoute le trifluoracétate d'éthyle (11,8 mL, 0,10 mol) goutte à goutte en 5 min à 10 °C sous agitation sous argon. Après 15 min à température ambiante, on transfère la solution dans une ampoule à décanter, on lave avec H_2O (3 × 100 mL), on sèche sur $MgSO_4$ et évapore le solvant. On obtient 22,4 g (71 %) du monoamide **62** de pureté d'environ 85 % (déterminée par F^{19} RMN). Ce composé est conservé à –20 °C et utilisé sans purification.

RMN ^1H (200 MHz, $CDCl_3$) : δ = 3,5-3,6 (m, 12H) ; 3,42 (t, 2H, J = 6 Hz) ; 2,75 (t, 2H, J = 6 Hz) ; 1,81 (quantiplet, 2H, J = 6 Hz) ; 1,67 (quantiplet, 2H, J = 6 Hz) ; 1,30 (s. large, 2H).

RMN ^{19}F (190 MHz, $CDCl_3$) : δ = –76,3.

● Composé **63** :

A une suspension de D-biotine (6,39 g ; 26,2 mmol) dans 50 mL de DMF, on ajoute le carbonyldiimidazole (CDI, 6,32 g , 80 %, 31,2 mmol). Le mélange est chauffé à 55-60 °C sous agitation sous argon pendant 30 min. On observe initialement une dissolution complète de matériel puis une prise en masse avec une précipitation de solide blanc (CO_2 évolution). On ajoute l'amine (huile) à l'aide de 5 mL de CH_2Cl_2 pour rincer et le mélange est chauffé à 55-60 °C pendant 3 h. On évapore le DMF sous vide (< 1 mmHg) et le résidu est agité avec CH_2Cl_2 (700 mL) et 2N HCl (100 mL). Après une filtration des deux phases à travers une couche de Célite 545, on sépare les phases, la phase aqueuse est extraite avec CH_2Cl_2 (15 × 100 mL), les phases organiques sont rassemblées, séchées sur $NaHCO_3$ anhydre et $MgSO_4$, puis le solvant est évaporé. Le résidu huileux est trituré avec 150 mL d'éther pour obtenir une suspension. Le solide pâteux est difficile à filtrer. On décante le surnageant et on répète le lavage à l'éther. Après séchage sous vide, on obtient 9,13 g (64 %) de composé **63**.
RMN ^1H (200 MHz, $CDCl_3$) : δ = 3,5-3,6 (m, 12H) ; 3,42 (t, 2H, J = 6 Hz) ; 2,75 (t, 2H, J = 6 Hz) ; 1,81 (quantiplet, 2H, J = 6 Hz) ; 1,67 (quantiplet, 2H, J = 6 Hz) ; 1,30 (s. large, 2H).

● Composé **64** :

On chauffe une solution du composé **63** dans l'ammoniaque (100 mL, 22 % aqueuse) dans un ballon de 250 mL bouché avec un septum à 55-60 °C pendant 2 h. Après refroidissement, on évapore le solvant. Le résidu est dissous dans le méthanol (20 mL) et passé sur une colonne résine échangeuse d'anion Dowex 21K [hauteur 12 cm × diamètre 35 mm, forme OH^- obtenu par lavage préalable avec NaOH 1N (1,5 L) puis H_2O (1,5 L) puis méthanol (1,5 L)]. Le composé **64** libre de l'ion trifluoracétate passe dans les premières fractions avec 200 mL du méthanol. Après évaporation, on triture le résidu avec 50 mL d'éther puis on le décante. Le lavage à l'éther est répété cinq fois de suite. Après séchage, on obtient le composé **64** (6,43 g, 86 %).

RMN ^1H (300 MHz, $CDCl_3$) : δ = 6,77 (t, 1H, J = 4 Hz) ; 6,32 (s, 1H) ; 5,52 (s, 1H) : 4,45 (m, 1H) ; 4,28 (m, 1H), 3,50-3,68 (m, 12H) ; 3,30 (m, 2H), 3,11 (m, 1H) ; 2,86 (dd, 1H, J = 13 et 5 Hz), 2,75 (t, 2H, J = 13 Hz), 2,68 (d, 1H, J = 13 Hz) ; 2,16 (t, 2H, J = 7 Hz) ; 1,60-1,85 (m, 8H) ; 1,41 (m, 2H).

- [Bio-EG_3]$_2$- acétal **65** :

A une suspension du diacide **60** (120 mg ; 0,500 mmol) dans le dichloroéthane (5 mL), on ajoute le carbonyldiimidazole (225 mg, 90 %, 1,25 mmol) et on chauffe à 55-60 °C pendant 30 min sous agitation sous argon. On ajoute l'amine **64** (550mg ; 1,23 mmol) et la solution est chauffée à 55-60 °C pendant 6 h. Après évaporation, on passe sur une colonne de silice (diamètre : 25 mm, éluant : méthanol 15-30 % dans CH_2Cl_2). On obtient 413 mg (75 %) de composé **65**.

RMN ^1H (300 MHz, $CDCl_3$) : δ = 8,34 (s, 1H) ; 8,06 (s, 2H) ; 7,87 (m, 2H) ; 6,85 (m, 2H) ; 6,60 (s, 2H) ; 5,93 (s, 2H) : 5,40 (s, 1H) ; 4,45 (m, 2H) ; 4,27 (m, 2H), 3,43-3,68 (m, 24H) ; 3,31 (s, 6H) ; 3,25 (m, 4H), 3,08 (m, 2H) ; 2,83 (dd, 2H, J = 13 et 5 Hz), 2,70 (t, 2H, J = 13 Hz) ; 2,13 (t, 4H, J = 7 Hz) ; 1,89 (quintuplet, 4H, J = 7 Hz) ; 1,55-1,70 (m, 12H) ; 1,37 (m, 4H).

- [Bio-EG_3]$_2$- aldéhyde **66** :

L'acétal **65** (413 mg ; 0,376 mmol) en solution dans le méthanol est traité avec HCl 2N (0,5 mL). Après évaporation et lavage avec l'éther, on obtient l'aldéhyde **66** (0,37 g, 90 %).

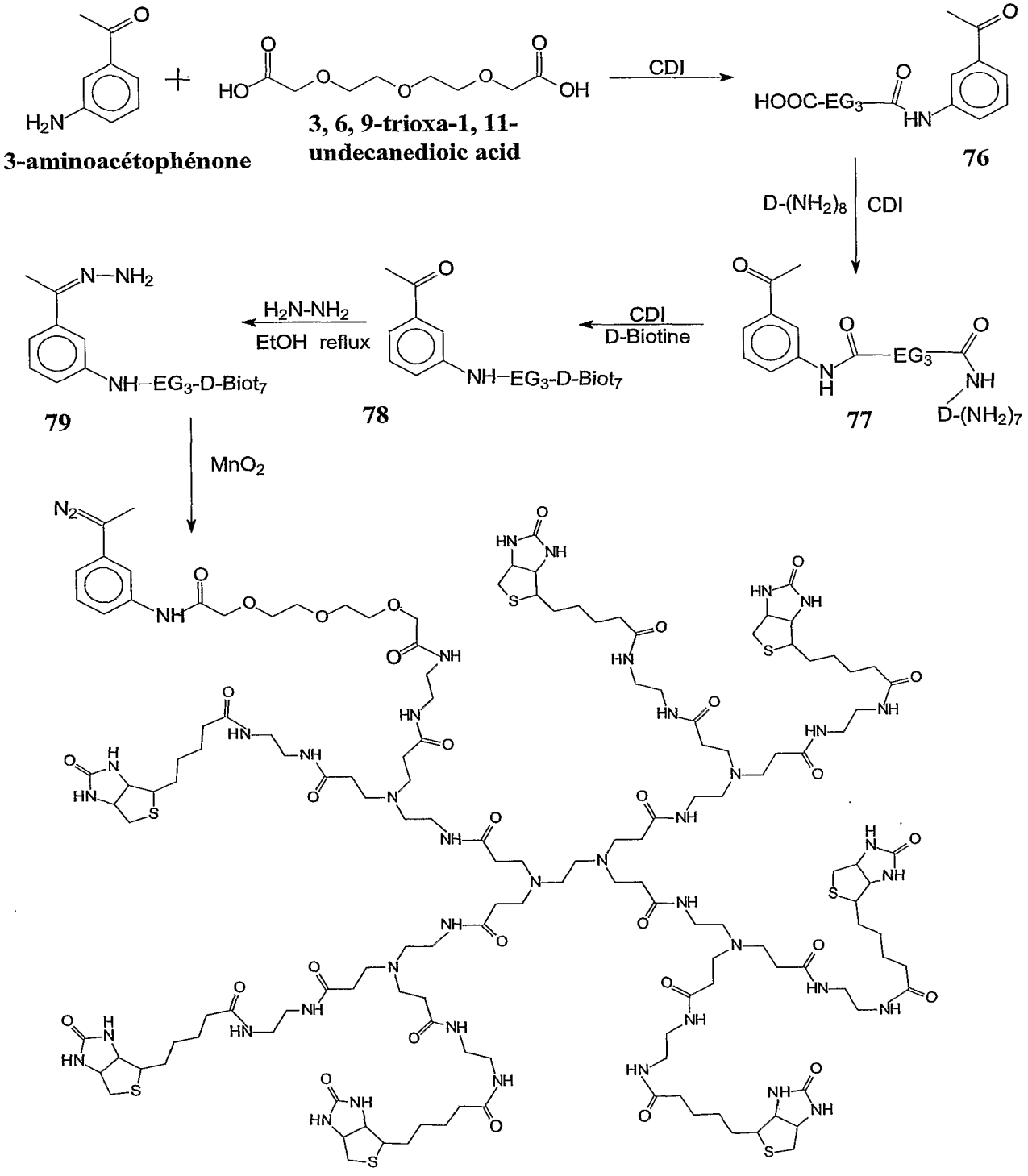
RMN ^1H (300 MHz, DMSO-d_6) : δ = 10,11 (s,1H) ; 8,82 (t, 2H, J = 6 Hz) ; 8,62 (s, 1H) ; 8,47 (s, 2H) ; 7,73 (t, 2H, J = 5 Hz) ; 4,30 (m, 2H) ; 4,11 (m, 2H), 3,30-3,60 (m, 24H) ; 3,06 (m, 6H) ; 2,80 (dd, 2H, J = 12 et 5 Hz), 2,56 (t, 2H, J = 12 Hz) ; 2,03 (t, 4H, J = 7 Hz) ; 1,78 (quintuplet, 4H, J = 7 Hz) ; 1,35-1,60 (m, 12H) ; 1,28 (m, 4H).

- [Bio-EG_3]$_2$- PDAM **67** :

L'aldéhyde **66** est transformé en diazométhane **67** selon la méthode utilisée pour la préparation des diazométhanes (Exemple 1).
La stabilité du réactif est supérieure à 1 mois à -20°C.

Exemple 26.2 : Synthèse du *méta*-Bio7-EG3-PMDAM :

Schéma de synthèse :


3-aminoacétophénone
3, 6, 9-trioxa-1, 11-undecanedioic acid
CDI
76
D-(NH2)8
CDI
77
CDI
D-Biotine
78
H2N-NH2
EtOH reflux
79
MnO2
méta-Bio7-EG3-PMDAM 80

Composé EG_3-acétophénone **76** :

Le 3, 6, 9-trioxa-1, 11-undecanedioic acid (EG_3, 12,64 ml, 74 mmol) est dissout dans 80 ml de DMF anhydre sous argon et refroidit dans un bain de glace. Le dicyclohexylcarbodiimide (DCC, 11,45 g, 55,5 mmol)) est ensuite dissout dans 20 ml de DMF anhydre et ajouté lentement. Après 30 min, on ajoute le 3-aminoacétophénone (5,0 g, 37 mmol) et on laisse réagir 1 h à température ambiante sous argon. Le DMF est ensuite évaporé sous vide et 70 ml de CH_2CL_2 sont ajoutés. La solution est filtrée et extraite avec 3x25 ml d'acide acétique 1%. Les phases aqueuses sont réunies et lavées avec 25 ml de CH_2CL_2. Les phases organiques sont mélangées, séchées avec du sulfate de sodium anhydre et évaporées à sec. Le produit est recristallisé dans le couple MeOH:H_2O. On obtient ainsi 8,74 g (70%) du produit **76**.

RMN ^1H (200 MHz, DMSO-d_6): d = 2,55 (s, 3H); 3,5-3,7 (m, 8H); 4,0 (s, 2H); 4,1 (s, 2H); 7,45 (t, 1H); 7,65 (d, 1H); 7,90 (d, 2H); 8,2 (s, 1H); 9,8 (s, 1H).

Composé $(NH_2)_7$-EG_3-acétophénone **77** :

Le produit **76** (120mg, 0,35 mmol) est dissout dans 15 ml de DMF anhydre sous argon, refroidit sur glace, et le DCC (110 mg, 0,53 mmol) est ensuite ajouté. Après 30 min, on ajoute cette solution sur une solution du dendrimère commercial « Starburst PAMAM Dendrimer, Generation 1 » (Aldrich, St Quentin Fallavier) (1 g, 0,71 mmol, dans 5 ml de méthanol), lentement et sous forte agitation. On laisse réagir 1 h à température ambiante et on évapore. Le résidu est repris dans 10 ml de CH_2CL_2 et extrait deux fois avec 30 ml d'acide acétique 1%.

Composé $Biot_7$-EG_3- acétophenone **78** :

La D-Biotine (1,73 g, 7,08 mmol) est solubilisée dans 80 ml de DMF anhydre sous argon, et la solution est refroidie sur glace. On ajoute successivement la N-méthylmorpholine (NMM, 856 µl, 7,7 mmol) et le chloroformiate d'isobutyle (1022 µl, 7,7 mmol). Après 30 min, on ajoute le produit **77** (1,13 g, 0,7 mmol, dans 5 ml de méthanol) et on laisse réagir 3 h sur glace et sous argon. On concentre sous vide jusqu'à 50 ml et on ajoute 100 ml de CH_2Cl_2 . Un précipité se

IDT Exhibit 1020.7823

forme, qui est filtré, lavé à l'éther, et séché sous vide. On obtient 1,3 g de **78** sous forme d'une poudre blanche.

Composé $Biot_7$-EG_3-Hydrazone **79** :

Le composé **78** (300 mg, 0,09 mmol) est dissout dans 10 ml d'éthanol absolu à reflux. On ajoute l'hydrazine monohydrate (20 ml, 0,40 mmol) et on laisse réagir 3 h à reflux. Après refroidissement, un précipité se forme, qui est filtré, lavé a l'éther et séché sous vide. On obtient ainsi 109 mg (36%) du produit **79**, sous forme d'une poudre blanche.

Composé $Biot_7$-EG_3-PMDAM **80** :

On solubilise l'hydrazone **79** (100 mg, 0,03 mmol) dans 5 ml de DMF anhydre à 70°C. On laisse revenir à température ambiante et on ajoute le MnO_2 (31 mg, 0, 36 mmol). On laisse réagir 10 min et on élimine l'oxyde de manganèse par filtration sur un fritté avec de la célite (0,5 cm et du tamis moléculaire en poudre (0,5 cm). Le filtrat est évaporé a sec, lavé a l'éther e seché sous vide. On obtient ainsi 78 mg (78%) du produit **80**.

Les dendrimères sont des molécules arborescentes possédant aux extrémités plusieurs groupements réactifs tels que des amines, carboxyles, hydroxyles ou autres (pour revue, voir Newcome et al, (1996) *Dendritic Molecules: Concept, Syntheses, Perspectives*. VCH Ed., Weinheim, Germany). La synthèse de ces molécules est aujourd'hui bien maîtrisée, et des nombreux dendrimères sont commercialisés par l'industrie chimique. Le choix du PAMAM (Sigma-Aldrich) a été fait sur la base de sa stabilité, solubilité et souplesse, puisque plusieurs versions de cette molécule, avec un nombre et type de terminaisons différentes, sont disponibles. Le "PAMAM Generation 1" permet d'ajouter sept molécules du marqueur (en une seule étape de synthèse) pour chaque groupement diazométhyl.

IDT Exhibit 1020.7824

Exemple 27 : Marquage et fragmentation en deux étapes d'amplicons ADN avec le *méta*-BioPMDAM :

Les amplicons ADN ont été préparés par amplification PCR selon le protocole décrit dans l'exemple 5. Deux réactions de marquage ont été réalisées.

a. Marquage et fragmentation en deux étapes :

A 10 µL de PCR sont ajoutés 10 µL de *méta*-BioPMDAM (100 mM dans du DMSO) et 77 µL d'eau Dnase/Rnase free. La solution est incubée 10 min à 95°C. Puis 3 µL d'HCl à 0,1M sont ajoutés et la solution est incubée 10 min à 95°C.

b. Marquage et fragmentation en une étape :

A 10 µL de PCR sont ajoutés 10 µL de *méta*-BioPMDAM (100 mM dans du DMSO), 5 µL d'HCl à 0,1 M et 75 µL d'eau Dnase/Rnase free. La solution est incubée 30 min à 60°C.
Le reste du protocole est identique à celui de l'exemple 9.

Résultats :

Protocole utilisé	Homologie (%)	I (rfu)	B (rfu)	I/B
a. Marquage et fragmentation en deux étapes	99,5	14129	624	22,7
b. Marquage et fragmentation en une étape	98,9	4431	667	6,6

Tableau 13 : Etude comparative du marquage et de la fragmentation en deux étapes distinctes et en une seule étape

Comme le démontre le tableau 13, les résultats obtenus avec le protocole en une étape sont satisfaisants. Ceux obtenus avec un marquage et une fragmentation en deux étapes sont encore meilleurs. Cet exemple montre que les étapes de marquage et clivage peuvent être dissociées pour améliorer le marquage en fonction de la cible utilisée.

Exemple 28 : Marquage et fragmentation d'amplicons ADN dans différents formats de réaction :

Les amplicons d'ADN ont été préparés par amplification PCR selon le protocole décrit dans l'exemple 5. Trois réactions de marquage ont été réalisées.

a. Marquage et fragmentation dans un format de 250 µL :

A 50 µL de PCR sont ajoutés 75 µL de *méta*-BioPMDAM (100 mM dans du DMSO) et 102,5 µL d'eau Dnase/Rnase free. La solution est incubée 25 min à 95°C. Puis 22,5 µL d'HCl à 0,1 M sont ajoutés et la solution est incubée 5 min à 95°C.

b. Marquage et fragmentation dans un format de 200 µL :

A 50 µL de PCR sont ajoutés 75 µL de *méta*-BioPMDAM (100 mM dans du DMSO) et 52,5 µL d'eau Dnase/Rnase free. La solution est incubée 25 min à 95°C. Puis 22,5 µL d'HCl à 0,1 M sont ajoutés et la solution est incubée 5 min à 95°C.

c. Marquage et fragmentation dans un format de 150 µL :

A 50 µL de PCR sont ajoutés 75 µL de *méta*-BioPMDAM (100 mM dans du DMSO) et 2,5 µL d'eau Dnase/Rnase free. La solution est incubée 25 min à 95°C.
Puis 22,5 µL d'HCl 0,1M sont ajoutés et la solution est incubée 5 min à 95°C.

Le reste du protocole est identique à celui de l'exemple 9.

Résultats :

Protocole utilisé	Homologie (%)	I (rfu)	B (rfu)	I/B
a. Format de 250 µL	100,0	5606	549	10,2
b. Format de 200 µL	99,4	5886	557	10,6
c. Format de 150 µL	99,4	6800	537	12,7

Tableau 14 : Marquage et fragmentation selon différents formats

Les résultats obtenus en terme de signal et de pourcentage d'homologie sont très satisfaisants dans tous les cas de figures. De plus, bien que le format de la réaction varie de 150 à 250 µL, les résultats ont des valeurs similaires.

Cet exemple montre une flexibilité du format réactionnel du protocole de marquage pouvant accepter différents volumes et notamment différents volumes de produits d'amplification.

Exemple 29 : Comparaison entre un protocole utilisant une étape de purification avant fragmentation et un protocole utilisant une étape de purification après fragmentation :

Les amplicons d'ADN ont été préparés par amplification PCR selon le protocole décrit dans l'exemple 5. Deux réactions de marquage ont été réalisées.

a. Marquage purification puis fragmentation des amplicons ADN :

A 10 µL de PCR sont ajoutés 10 µL de *méta*-BioPMDAM (100 mM dans du DMSO) et 80 µL d'eau Dnase/Rnase free. La solution est incubée 10 min à 95°C. Puis la purification est réalisée selon le protocole décrit dans l'exemple 9. A la solution d'amplicons marqués purifiés 6 µL d'HCl à 0,1 M sont ajoutés. La solution est incubée 10 min à 95°C. Quatre cents (400) µL de tampon d'hybridation préchauffé à 95°C durant 10 min sont ajoutés.

La composition du tampon d'hybridation ainsi que le reste du protocole sont identiques à celui de l'exemple 9.

b. Marquage fragmentation puis purification des amplicons ADN :

A 10 µL de PCR sont ajoutés 10 µL de *méta*-BioPMDAM (100 mM dans du DMSO) et 77 µL d'eau Dnase/Rnase free. La solution est incubée 10 min à 95°C.
Puis 3 µL d'HCl à 0,1 M sont ajoutés et la solution est incubée de nouveau 10 min à 95°C. Le reste du protocole est identique à celui de l'exemple 9.

Résultats :

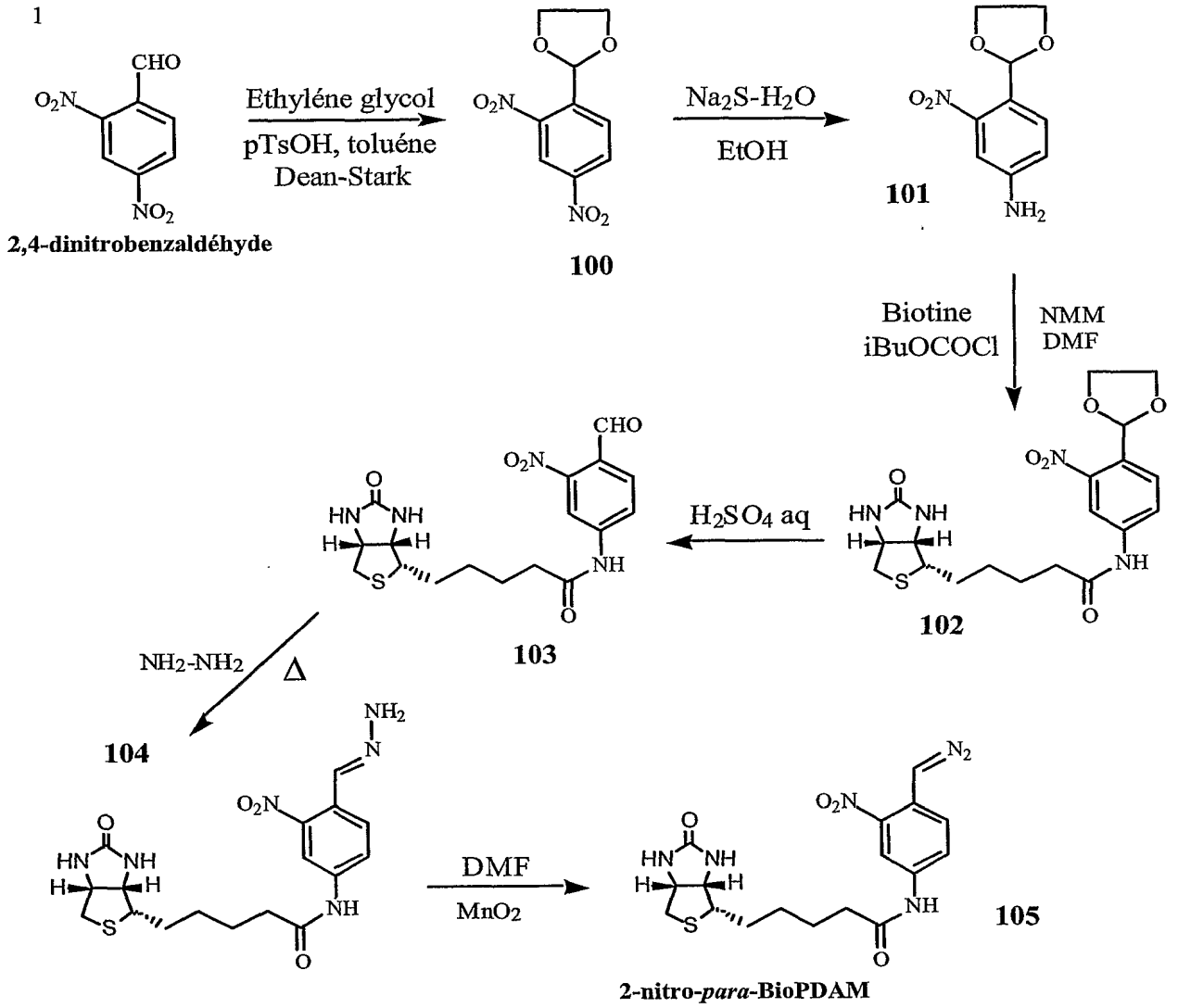
Protocole utilisé	Homologie (%)	I (rfu)	B (rfu)	I/B
a. Purification avant fragmentation	98,9	6256	473	13
b. Fragmentation avant purification	96,1	6066	556	11

Tableau 15 : Comparaison entre protocole utilisant une étape de purification avant fragmentation et protocole utilisant une étape de purification après fragmentation

Ce résultat, présenté dans le tableau 15, montre que l'étape de purification peut être introduite entre les étapes de marquage et de fragmentation. De plus l'introduction de la purification entre les étapes de marquage et de fragmentation permet de réaliser la dénaturation pendant le clivage et d'hybrider sur la puce la totalité des fragments d'amplicons marqués.

Exemple 30 : Synthèse du 2-nitro-*para*-BioPDAM :

Schéma de synthèse :



● Protection de l'aldéhyde :

On dissout 5 g (25,5 mmol) de 2,4-dinitrobenzaldéhyde dans 250 mL de toluène, et l'on ajoute 20 mL d'éthylène glycol et 150 mg d'acide *para*-toluènesulfonique. On chauffe à reflux en récupérant l'eau dans un système Dean-Stark pendant 6 h. On traite avec 150 mL d'EtOAc et 100 mL d'H_2O. On extrait deux fois la solution à l'acétate d'éthyle, on sèche la phase organique avec $MgSO_4$ puis on l'évapore. L'huile obtenue correspondant au produit **100**, est utilisée pour la réaction suivante.

RMN ^{1}H (200 MHz, $CDCl_3$) : δ = 8,70 (d, $1H_{aro}$, J = 2 Hz, H_3) ; 8,44 (dd, $1H_{aro}$, J = 2 Hz, J = 6 Hz, H_5) ; 8,02 (d, $1H_{aro}$, J = 8 Hz, H_6) ; 6,49 (s, 1H, CH) ; 4,12-4,06 (m, 4H, CH_2-CH_2).

● Réduction du dérivé dinitro **100** :

Le 2-4-dinitrobenzaldéhyde protégé (6,4 g ; 25,5 mmol) est dissous dans un mélange éthanol-eau (6/1), puis 2 équivalents de Na_2S nonahydraté (12,3 g ; 51,1 mmol) sont ajoutés. Le mélange réactionnel est ensuite chauffé pendant 30 min. Une évaporation puis une extraction au dichlorométhane sont réalisées. Après séchage et filtration, le milieu réactionnel est évaporé afin d'obtenir une huile, qui est directement purifiée sur colonne de silice (cyclohexane-acétate d'éthyle 60/40). Le composé **101** est isolé avec un rendement de 45 %.

Composé **101** : F 58-60°C. - RMN ^{1}H (200 MHz, $CDCl_3$) : 7,49 (d, $1H_{aro}$, J = 2 Hz, H_3) ; 7,09 (d, $1H_{aro}$, J = 2 Hz, H_6) ; 6,80 (dd, $1H_{aro}$, J = 2 Hz, J = 6 Hz, H_5) ; 6,27 (s, 1H, CH) ; 3,99-3,97 (m, 4H, CH_2-CH_2).

● Couplage avec la biotine :

On solubilise la D-biotine (1,0 g ; 4,1 mmol) dans 20 mL de DMF anhydre et 600 µL de N-méthylmorpholine. On ajoute sous argon le chloroformiate d'isobutyle (700 µL ; 5,5 mmol) en refroidissant dans un bain glacé. On laisse sous agitation pendant 5 min, puis on ajoute 1 g (4,75 mmol) du composé **101** et 500 µL de N-méthylmorpholine. La solution est maintenue sous agitation à température ambiante pendant 4 h, puis évaporée à sec. L'huile obtenue est passée directement sur colonne de silice avec comme solvant d'élution MeOH-DCM 7 % puis 10 %. Le produit **102** (1,1 g ; 2,52 mmol) est obtenu avec un rendement de 62 %.
RMN ^{1}H (200MHz, DMSO-d_6) δ =10,40 (s, 1H, NH-CO) ; 8,31 (d, $1H_{aro}$, J = 2 Hz, H_3) ; 7,77 (dd, $1H_{aro}$, J = 2 Hz, J = 6 Hz, H_5) ; 7,68 (d, $1H_{aro}$, J = 2 Hz, H_6) ; 6,43 (s large, 1H, NH-CO-NH) ; 6,36 (s large, 1H, NH-CO-NH) ; 6,23 (s, 1H, CH) ; 4,28 (m, 1H, CH_2-CH-NH) ; 4,14 (m, 1H, CH-CH-NH) ; 3,92 (s, 4H, CH_2-CH_2) ; 3,12 (m, 1H, CH-S) ; 2,85 et 2,76 (système ABX, 2H, $^2J_{AB}$= 5 Hz, $^3J_{AX}$ = 12 Hz, $^3J_{BX}$ = 0 Hz, CH_2-S) ; 2,29 (t, 2H, J = 8 Hz, CH_2-CO) ; 1,61-1,39 (m, 6H, $(CH_2)_3$).

● Déprotection de l'acétal :

On met le produit **102** (768 mg ; 1,76 mmol) en suspension dans 25 mL de THF. Tout se dissout après ajout de 4 mL de H_2SO_4 2N. On laisse sous agitation pendant 2 h. On évapore puis on rince et lave à l'eau sur fritté. Le composé **103** (694 mg) est obtenu sous forme d'une poudre jaune avec un rendement de 90 %.

F 165°C. - RMN ^1H (200MHz, DMSO-d_6) δ =10,69 (s, 1H, NH-CO) ;10,09 (s, H, CHO) ; 8,43 (d, $1H_{aro}$, J = 2 Hz, H_3) ; 7,91 (s, $2H_{aro}$, H_5 et H_6) ; 6,42 (s large, 1H, NH-CO-NH) ; 6,35 (s large, 1H, NH-CO-NH) ; δ = 6,23 (s, 1H, CH) ; 4,29 (m, 1H, CH_2-CH-NH) ; 4,13 (m, 1H, CH-CH-NH) ; 3,12 (m, 1H, CH-S) ; 2,84 et 2,78 (système ABX, 2H, $^2J_{AB}$= 5 Hz, $^3J_{AX}$ = 12 Hz, $^3J_{BX}$ = 0 Hz, CH_2-S) ; 2,29 (t, 2H, J = 8 Hz, CH_2-CO) ; 1,61-1,39 (m, 6H, $(CH_2)_3$).

● Formation de l'hydrazone **104** :

On met l'aldéhyde **103** en suspension dans l'éthanol et l'on chauffe à 80°C. Quand on ajoute l'hydrazine, tout se dissout et la solution se colore immédiatement en orange. Un précipité se forme après 5 min. On chauffe sous agitation pendant 1 h. On filtre sur fritté puis on sèche le précipité. Le produit **104** (700 mg ; 690 mmol) est obtenu avec un rendement de 98 %.

F 169°C. - RMN ^1H (200MHz, DMSO-d_6) δ =10,31 (s, 1H, NH-CO) ; 8,31 (d, $1H_{aro}$, J = 2 Hz, H_3) ; 7,96 (s, H, CHO) ; 7,87 (d, $1H_{aro}$, J = 2 Hz, H_6) ; 7,68 (dd, $1H_{aro}$, J = 2 Hz, J = 6 Hz, H_5) ; 7,31 (s, 2H, NH_2) ; 6,42 (s large, 1H, NH-CO-NH) ; 6,34 (s large, 1H, NH-CO-NH) ; 4,29 (m, 1H, CH_2-CH-NH) ; 4,13 (m, 1H, CH-CH-NH) ; 3,12 (m, 1H, CH-S) ; 2,84 et 2,78 (système ABX, 2H, $^2J_{AB}$= 5 Hz, $^3J_{AX}$ = 12 Hz, $^3J_{BX}$ = 0 Hz, CH_2-S) ; 2,29 (t, 2H, J = 8 Hz, CH_2-CO) ; 1,61-1,39 (m, 6H, $(CH_2)_3$).

● Formation du diazo **105** :

On dissout le composé **104** (200 mg ; 0,492 mmol) dans 8 mL de DMF. On ajoute 400 mg de MnO_2. On agite vigoureusement pendant 10 min. On filtre sur millipore contenant de la célite (épaisseur : 2 cm) et du tamis moléculaire en poudre 3 Å (0,5 cm). On évapore à sec puis on lave à l'éther. On filtre de nouveau sur millipore. Le composé **105** (180 mg ; 0,445 mmol) est obtenu sous forme d'une poudre orange avec un rendement de 98 %.

F 155°C. - RMN ^1H (200MHz, DMSO-d_6) δ =10,21 (s, 1H, NH-CO) ; 8,60 (d, $1H_{aro}$, J = 2 Hz, H_3) ; 7,77 (d, $1H_{aro}$, J = 6 Hz, H_5) ; 7,22 (d, $1H_{aro}$, J = 6 Hz, H_6) ; 6,60 (s, H, CH-N) ; 6,41 (s large, 1H, NH-CO-NH) ; 6,33 (s large, 1H, NH-CO-NH) ; 4,29 (m, 1H, CH_2-CH-NH) ; 4,13 (m, 1H, CH-CH-NH) ; 3,12 (m, 1H, CH-S) ; 2,84 et 2,78 (système ABX, 2H, $^2J_{AB}$= 5 Hz, $^3J_{AX}$ = 12 Hz, $^3J_{BX}$ = 0 Hz, CH_2-S) ; 2,29 (t, 2H, J = 8 Hz, CH_2-CO) ; 1,61-1,39 (m, 6H, $(CH_2)_3$).

La réactivité du composé **105** a été testée sur 3'-Uridine-monophosphate suivie par électrophorèse capillaire. Les conditions d'analyse sont celles de l'exemple 6.1. Les résultats montrent un temps de demi réaction de 45 minutes.
La stabilité du réactif est supérieure à 1 mois à -20°C

Exemple 31 : Marquage et fragmentation des amplicons d'ADN avec le réactif de marquage 2-nitro-*para*-BioPDAM :

Le dérivé 2-nitro-*para*-BioPDAM a été obtenu selon le schéma réactionnel décrit dans l'exemple 30. Les amplicons ADN ont été préparés par amplification PCR selon le protocole décrit dans l'exemple 5. Deux réactions de marquage ont été réalisées.

a. Marquage par le réactif 2-nitro-*para*-BioPDAM :

A 10 µL de PCR sont ajoutés 2 µL de 2-nitro-*para*-BioPDAM (100 mM dans du DMSO), 5µL d'HCl à 0,1 M et 83 µL d'eau Dnase/Rnase free. Cette solution est incubée 30 min à 60°C.

b. Marquage par le réactif *méta*-bioPMDAM :

A 10 µL de PCR sont ajoutés 2 µL de *méta*-BioPMDAM (100 mM dans du DMSO), 5µL d'HCl à 0,1 M et 83 µL d'eau Dnase/Rnase free. Cette solution est incubée 30 min à 60°C.

Le reste du protocole est identique à celui de l'exemple 9.

Résultat :

Protocole utilisé	Homologie (%)	I (rfu)	B (rfu)	I/B
a. Marquage par le réactif 2-nitro-*para*-BioPDAM	100,0	24392	899	27,1
b. Marquage par le réactif méta-BioPMDAM	98,9	21883	774	28,3

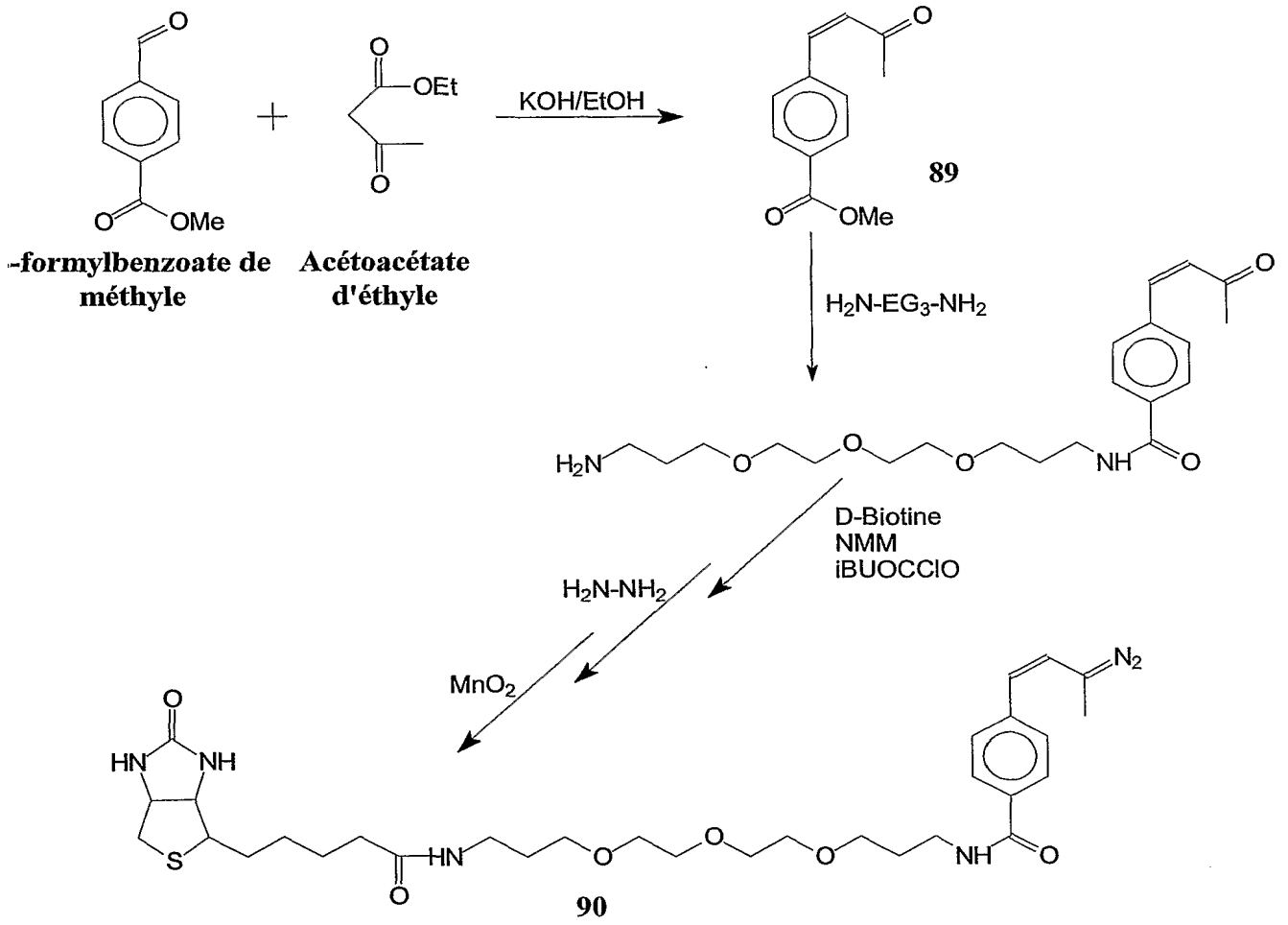
Tableau 16 : Etude comparative du marquage d'ADN par le dérivé 2-nitro-*para*-BioPDAM vis-à-vis du dérivé *méta*-BioPMDAM

Le réactif 2-nitro-*para*-BioPDAM utilisé pour le marquage de l'ADN donne des résultats intéressants en terme d'intensité de marquage et de pourcentage d'homologie.

Exemple 32 : Insertion d'une double liaison entre la fonction diazométhyle et le noyau phenyle, éloignement du DAM et synthèse d'une molécule particulièrement adaptée à cet éloignement :

L'objectif visé est l'éloignement de la fonction diazométhyle (DAM) de la structure aromatique pour minimiser l'effet de l'encombrement stérique lors de l'alkyaltion des phosphates et aussi lors de l'hybridation de l'acide nucléique marqué avec sa séquence complémentaire.

Schéma de synthèse :



-formylbenzoate de méthyle

Acétoacétate d'éthyle

Pour la réaction aldolique de formation du (*para*-méthoxycarbonyl)-styrylméthylcétone **89**, on utilise l'acétoacétate d'éthyle pour la forte acidité des protons du méthylène, ce qui facilite l'attaque du groupement formyl, avec élimination postérieure de H_2O (favorisée par la conjugaison de la double liaison avec l'anneau aromatique) et décarboxylation par hydrolyse due au milieu basique. Le reste de la synthèse est similaire a ce qui a été montré dans les autres exemples.

Le produit final *para*-Bio-EG_3-SMDAM **90**, possède deux carbones de plus entre le diazométhyle et l'anneau aromatique, ce qui limite les possibles problèmes stériques, tout en conservant la stabilisation du diazométhyle par le système aromatique par conjugaison.

Exemple 33 : Capture et détection d'un acide nucléique sur un support solide portant des groupements diazométhyles :

La réactivité d'une résine portant des groupements diazométhyles a été étudiée pour déterminer sa capacité à lier des acides nucléiques.

La 4-(Diazomethyl)phenoxymethyl-polystyrene (référence 17338, Fluka) est une résine décrite pour son pouvoir à lier les groupements carboxyliques, en particuliers ceux présents dans les protéines (G. Bhalay, A.R. Dunstan, Tetrahedron Lett. *39*, 7803 – 1998), mais elle n'est pas décrite pour sa capacité à lier les molécules d'ADN. Nous avons testé la possibilité de capturer des acides nucléiques avec ce réactif, et de les révéler par un test colorimétrique.

L'expérience est réalisée avec une partie des réactifs présents dans le kit HLA-DR oligo-détection (référence 33 202, bioMérieux, France principe de base décrit dans le brevet EP 549 776-B1), permettant la détection d'acides nucléiques amplifiés par PCR dans des microplaques, par lecture colorimétrique. Dans le cadre de l'expérience décrite, on fait réagir simultanément des acides nucléiques produits par PCR avec la résine testée, ainsi qu'avec une molécule de *para*-Bio-EG3-PDAM, dont la synthèse est évoquée dans l'exemple 20. Si l'ADN réagit avec les fonctions diazométhyle présentes sur les deux composés, il sera possible de le révéler, après lavage et élimination des molécules fixées de manière non covalente, grâce à une réaction colorimétrique faisant intervenir une enzyme couplée à la streptavidine. La streptavidine est associée à une péroxydase de raifort, cet enzyme pouvant décomposer une molécule d'OrthoPhénylène Diamine (réactif Color 1 du kit) en un composé détectable à une longueur d'onde de 492 nm.

Exemple 33.1 : Capture et détection de l'ADN :

On incube 10 mg de résine pendant 30 minutes à 60°C avec 50 µL de PCR réalisées comme décrit dans l'exemple 5, dans 400 µL d'eau pure (Sigma) additionnée de 5 µL de *para*-Bio-EG3-PDAM. On lave ensuite cette résine par 500 µL de tampon PBS tween (Réactif Color 0

HLA du kit, PBS pH 7,0 ; Tween 1 %, BND 0,2 g/L ; Ciproflaxacine 0,01 g/L). On remet en suspension alors la résine dans 100 µL de PBS Tween, et 250 µL de tampon d'hybridation streptavidine (BPS pH 7,0 TWEEN 0,5 %) additionné de streptavidine HRP (S-911, MOLECULAR PROBES, EUGENE, OR, USA) dilué au 1/10000. On incube le mélange réactionnel 30 min à température ambiante. On lave ensuite la résine trois (3) fois par 500 µL de tampon PBS tween, et on l'incube à température ambiante en présence de réactif chromogénique (1 Comprimé Color 1, Chlorhydrate d'orthophénylène diamine, dilué dans 5 mL de tampon Color 2, Phosphate de sodium 100 mM, acide citrique 50 mM, H_2O_2 0,03%). Après une incubation de 20 min à l'obscurité, la réaction est ensuite bloquée par 50 µL de H_2SO_4 (1, 8 N réactif Color 3). Le surnageant est alors pipeté et placé dans une microplaque pour lire l'absorbance du milieu réactionnel à 492 nm.

Exemple 33.2 : Témoin sans acide nucléique :

On incube 10 mg de résine pendant 30 minutes à 60°C dans 425 µL d'eau pure (Sigma) additionnée de 5 µL de *para*-Bio-EG3-PDAM. On lave ensuite cette résine par 500 µL de tampon PBS. On traite alors l'échantillon de façon identique à processus décrit à l'exemple 33.1.

Exemple 33.3 : Témoin avec PCR réalisée sans cibles :

On incube 10 mg de résine dans 400 µL d'eau pure additionnée de 5 µL de *para*-Bio-EG3-PDAM pendant 30 minutes à 60°C, avec 50 µL de PCR réalisée avec un volume de 25 µL d'eau pure à la place du volume l'ADN génomique décrit. On lave ensuite cette résine par 500 µL de tampon PBS. On traite alors l'échantillon de façon identique au processus décrit à l'exemple 33.1.

Exemple 33.4 : Témoin avec PCR sans molécule de révélation :

On incube 10 mg de résine avec 50 µL de PCR dans 400 µL d'eau pure pendant 30 minutes à 60°C. On lave ensuite cette résine par 500 µL de tampon PBS. On traite alors l'échantillon de façon identique au processus décrit à l'exemple 33.1.

Exemple 33.5 : Témoin avec acide nucléique non capturé :

On incube 10 mg de résine pendant 30 minutes à 60°C avec 400 µL d'eau pure additionnée de 5 µL de *para*-Bio-EG3-PDAM. On lave ensuite cette résine par 500 µL de tampon PBS tween (Réactif Color 0 HLA du kit, PBS pH 7,0 ; Tween 1 %, BND 0,2 g/L ; Ciproflaxacine 0,01 g/L). On resuspend alors la résine dans 100 µL de PBS Tween, et 250 µL de tampon d'hybridation streptavidine additionné de streptavidine HRP dilué au 1/10000 ème. A cette préparation est rajoutée 50µL d'une préparation d'ADN préparée comme suit :

On rajoute 5 µL de *para*-Bio-EG3-PDAM et 70 µL d'eau pure à 25 µL d'ADN provenant d'une PCR préparé comme décrit dans l'exemple 5. Ce mélange est incubé 30 min à 60°C, puis l'excès de marqueur est éliminé en soumettant la préparation à une purification sur colonne QIAquick (Nucleotide Removal Kit, Qiagen, Hilden, Allemagne) suivant le protocole recommandé par le fournisseur, en effectuant une élution finale dans un volume de 50 µL.

Le mélange réactionnel est incubé 30 min à température ambiante, puis il est traité suivant le processus décrit à l'exemple 33.1.

Résultats

Conditions	Absorbance à 492 nm
Ex. 33.1 : ADN capturé sur Résine	527
Ex. 33.2 : Sans acide nucléique	249
Ex. 33.3 : PCR sans cible	261
Ex. 33.4 : Témoin sans marqueur	264
Ex. 33.5 : Acide nucléique non capturé	249

Tableau 17 : Etude de la réactivité d'une résine portant des groupements diazométhyle

Dans le tableau 17, une forte valeur colorimétrique indique une forte concentration d'enzymes dans le milieu réactionnel, correspondant à une présence importante d'acide nucléique portant des dérivés biotines. Les témoins indiquent que le signal n'est pas dû à une adsorption non spécifique de l'ADN sur la bille, à une réaction du *para*-Bio-EG3-PDAM sur la résine, ou à une adsorption de la streptavidine HRP sur la résine, mais bien à la présence d'ADN capturé de façon covalente et marqué par le *para*-Bio-EG3-PDAM.

Exemple 34 : Marquage d'un produit PCR permettant sa capture et sa détection dans une microplaque :

Il est montré dans cet exemple la possibilité de marquer une molécule d'ADN avec un seul type de molécule portant une fonction diazométhyle, de façon à capturer et à détecter cet acide nucléique en une seule étape, sur une microplaque.

L'expérience est réalisée avec une partie des réactifs présent dans le kit HLA-DR oligo-détection (référence 33 202, bioMérieux), permettant la détection d'acides nucléiques amplifiés par PCR dans des microplaques, par lecture colorimétrique. Dans le cadre de l'expérience décrite, on fait réagir du *para*-Bio-EG3-PDAM, dont la synthèse est décrite dans l'exemple 20, sur des acides nucléiques produits par PCR. L'ADN réagit avec les groupements diazométhyle de la molécule, et se trouve ainsi doté de biotines greffé sur ses phosphate. Il sera alors possible de capturer l'acide nucléique par incubation sur une microplaque où des molécule de streptavidine sont adsorbées, et de le révéler, grâce à une réaction colorimétrique.On utilise un réactif de détection qui est aussi une molécule de streptavidine, associée à une peroxydase de raifort (Horse Radish Peroxidase, HRP). Dans les condition utilisées, la péroxydase peut décomposer une molécule d'OrthoPhénylène Diamine (réactif Color 1 du kit) en un composé détectable à une longueur d'onde de 492 nm.

Exemple 34.1 : Capture et détection de l'ADN issu de PCR, sur microplaque :

On marque 10 µL d'ADN obtenu par amplification PCR comme décrit dans l'exemple 5 en l'incubant 30 min à 60°C dans 80 µL d'eau pure (Sigma) additionnée de 10 µL de *para*-Bio-EG3-PDAM. Après marquage, l'ADN est purifié sur colonne QIAquick (Nucleotide Removal Kit, Qiagen, Hilden, Allemagne) suivant le protocole recommandé par le fournisseur, et l'éluat final est recueilli dans 50 µL de tampon EB (Tris EDTA 10 mM, pH 8,5). Vingt (20) µL de cet éluat est dilué dans 180 µL de tampon PEG (0,1 M Na PO3 ; 0,5 M NaCl ; 0,65% Tween 20 ; 0,14 mg/mL ADN de sperme de Saumon (Gibco) ; 2 % PEG 4000) additionné de streptavidine HRP (S-911, Molecular Probes, Eugene, OR, USA) diluée au 1/10 000. Cent (100) µL de cette préparation sont ensuite incubés 1 heure à 37°C, soit dans un puits de Combiplate 8, streptavidine coated, (référence 95029263, Labsystem, Helsinki, Finlande) soit dans un puits témoin, provenant d'une barrette Maxisorb. (Nunc, Danemark).

Exemple 34.2 : Réalisation de témoins :

Des témoins sont simultanément réalisés de la manière suivante :

A - Témoin de marquage sans ADN :
Quatre-vingt-dix (90) µL d'eau pure additionnés de 10 µL de *para*-Bio-EG3-PDAM sont incubés 30 min à 60°C. Le mélange réactionnel est ensuite traité de façon similaire au processus précédemment décrit à l'exemple 34.1.

B - Témoin de marquage sans marqueurs :
Dix (10) µL d'ADN obtenus par amplification PCR, comme décrit dans l'exemple 5, sont incubés 30 min à 60°C dans 90 µL d'eau pure. Le mélange réactionnel est ensuite traité de façon similaire au processus précédemment décrit à l'exemple 34.1.

Toutes les barrettes sont ensuite lavées par trois fois 100 µl de tampon PBS Tween (Réactif Color 0 HLA du kit, PBS pH 7,0 ; Tween 1 %, BND 0,2 g/L ; Ciproflaxacine 0,01

g/L) puis la présence de streptavidine HRP est révélée par addition de 100 µL du réactif chromogénique (1 Comprimé Color 1, Chlorydrate d'orthophénylène diamine, dilué dans 5 mL de tampon Color 2, Phosphate de sodium 100 mM, acide citrique 50 mM, H_2O_2 0,03%), incubé 20 min dans l'obscurité, la réaction étant ensuite bloquée par 50 µL de H_2SO_4 (1,8 N Réactif Color3). L'absorbance du milieu réactionnel est alors mesurée à 492 nm.

Résultats :

Conditions	Combiplate	Témoin maxisorp
A - ADN marqué	1382	152
B - ADN non marqué	178	136
C - Sans ADN	140	192

Tableau 18 : Détection d'ADN capturé et marqué par du *para*-Bio-EG3-PDAM

L'expérience, exposée dans le tableau 18, montre donc que l'ADN marqué par du *para*-Bio-EG3-PDAM peut être capturé et détecté en une seule étape dans un puits de microplaque. Comme l'indique les témoins de réaction, le signal généré est uniquement dû à l'ADN et ne résulte pas d'une adsorption non spécifique de l'acide nucléique sur la paroi de la microplaque, ou sur la streptavidine, ou bien à une réaction non spécifique de la streptavidine HRP sur l'ADN non marqué, ou sur le plastique de la microplaque.

Exemple 35 : Double marquage d'un produit PCR permettant sa capture et sa détection sur un support solide de type microplaque :

Il est montré dans cet exemple la possibilité de marquer avec deux molécules, porteuses de fonctions diazométhyles, et en une seule étape, une molécule d'ADN, de façon la capturer et à la détecter sur une microplaque.

L'expérience est réalisée avec une partie des réactifs présent dans le kit HLA-DR oligo-détection, permettant la détection d'acides nucléiques amplifié par PCR dans des microplaques, par simple lecture colorimétrique. Dans le cadre de l'expérience décrite, on fait réagir simultanément sur des acides nucléiques produits par PCR du :

- 1-Pyrényldiazométhane (PDAM), et
- *para*-Bio-EG3-PDAM.

Si l'ADN réagit avec les fonctions diazométhyle présentes sur les deux composés, il pourra se fixer sur un support portant des anticorps anti-pyrène, et il sera possible de le révéler avec une molécule de streptavidine associée à une péroxydase de raifort. Cette enzyme peut décomposer une molécule d'OrthoPhénylène Diamine, faisant office de réactif de révélation, en un composé détectable à une longueur d'onde de 492 nm.

Exemple 35.1 : Double marquage de l'ADN et détection sur microplaque :

On a adsorbé sur des barrettes Maxisorp huit (8) puits, des anticorps anti-pyrène, en incubant une nuit à température ambiante 100 µL d'une solution de 1,1 µL d'anticorps anti-pyrène dilués dans 100 µL de tampon bicarbonate (0,05M pH 9,6). De tels anticorps, appelés Rabbit anti-pyrène (Réf. : YSRT-AHP236) sont disponibles chez Accurate Chemical & Scientific (Westbury, New York, Etats-Unis d'Amérique). Bien entendu, ceci aurait pu être réalisé avec d'autres anticorps commercialement disponibles, sans qu'il y ait de résultats divergents par rapport à ceux que nous avons obtenus dans cet exemple.

On a ensuite marqué 10 µL d'ADN obtenu par amplification PCR, comme décrit dans l'exemple 5, en l'incubant 30 min à 60°C dans 40 µL d'eau pure (Sigma), 10 µL de *para*-Bio-EG3-PDAM, 2 µL de PDAM (P-1405, 1-pyrenyldiazomethane, Molecular Probes, Eugene, OR, USA) et 38 µL de DMSO.

Après marquage, l'ADN a été purifié par utilisation d'un kit QIAquick (QIAGEN) et l'éluat final à été recueilli dans 50 µL de tampon EB (Tris EDTA 10 mM, pH 8,5). Vingt (20) µL de cet éluat a été dilué dans 180 µL de tampon PEG (0,1 M Na PO3 ; 0,5 M NaCl ; 0,65 % Tween 20 ; 0,14 mg/mL ADN de sperme de Saumon (Gibco) ; 2 % PEG 4000) additionné de streptavidine HRP (S-911, MOLECULAR PROBES, EUGENE, OR,

USA) diluée au 1/10 000. Cent (100) µL de cette préparation ont été ensuite incubés 1 heure à 37°C, soit dans un puits de barrette Maxisorp adsorbé, soit dans un puits témoin, non adsorbé.

Exemple 35.2 : Réalisation de témoins :

Des témoins ont été simultanément réalisés de la manière suivante :

A - Témoin de marquage avec *para*-Bio-EG3-PDAM seul :

Dix (10) µL d'ADN obtenus par amplification PCR, comme décrit dans l'exemple 5, sont marqués par incubation de 30 min à 60°C dans 90 µL d'eau pure additionnée de 10 µL de *para*-Bio-EG3-PDAM. Après marquage, l'ADN est purifié par utilisation d'un kit QIAquick et l'éluat final est recueilli dans 50 µL de tampon EB. Vingt (20) µL de cet éluat est dilué dans 180 µL de tampon PEG additionné de streptavidine HRP diluée au 1/10 000. Cent (100) µL de cette préparation sont ensuite incubés 1 heure à 37°C, soit dans un puits de barrette Maxisorp adsorbé, soit dans un puits non adsorbé.

B - Témoin de marquage avec PDAM seul :

Dix (10) µL d'ADN obtenus par amplification PCR, comme décrit dans l'exemple 5, sont marqués par incubation de 30 min à 60°C dans 90 µL d'eau pure additionnée de 2 µL de PDAM et 38 µL de DMSO. Après marquage, l'ADN est purifié par utilisation d'un kit QIAquick et l'éluat final est recueilli dans 50 µL de tampon EB. Vingt (20) µL de cet éluat est dilué dans 180 µL de tampon PEG additionné de streptavidine HRP diluée au 1/10 000. Cent (100) µL de cette préparation sont ensuite incubés 1 heure à 37°C, soit dans un puits de barrette Maxisorp adsorbé, soit dans un puits non adsorbé.

C - Témoin sans marquage :

Dix (10) µL d'ADN obtenus par amplification PCR, comme décrit dans l'exemple 5, sont incubés 30 min à 60°C dans 100 µL d'eau pure. Après marquage, l'ADN est purifié par utilisation d'un kit QIAquick et l'éluat final est recueilli dans 50 µL de tampon EB. Vingt (20) µL de cet éluat est dilué dans 180 µL de tampon PEG additionné de streptavidine

HRP diluée au 1/10 000. Cent (100) µL de cette préparation sont ensuite incubés 1 heure à 37°C, soit dans un puits de barrette Maxisorp adsorbé, soit dans un puits non adsorbé.

Les barrettes sont ensuite lavées par trois fois 100 µL de tampon PBS Tween (color 0) puis la présence de streptavidine HRP est révélée par addition de 100 µL du réactif chromogénique (Color 2), incubé 20 min dans l'obscurité, la réaction étant ensuite bloquée par 50 µL de H_2SO_4 (Color3). L'absorbance du milieu réactionnel est alors mesurée à 492 nm.

Résultats :

Conditions	**Anticorps anti-pyrène adsorbés sur plaque (rfu)**	**Témoin plaque non adsorbée (rfu)**
Ex. 35.1 : ADN marqué PDAM + *para*-Bio-EG3-PDAM	348	16
Ex. 35.2.A : ADN marqué *para*-Bio-EG3-PDAM	44	19
Ex. 35.2.B : ADN marqué PDAM	68	12
Ex. 35.2.C : ADN non marqué	75	19

Tableau 19 : Double marquage de l'ADN par le PDAM et le *para*-Bio-EG3-PDAM

Le résultat du tableau 19 montre clairement un signal important résultant de la capture de l'ADN dans les puits par les anticorps anti-pyrène, ainsi que le marquage simultané de celui-ci par la streptavidine HRP, qui s'est fixée. Comme le montre l'absence de signal au niveau des témoins, cette détection est spécifique de l'ADN marqué, et n'est pas dû à l'adsorption non spécifique de l'ADN ou de la streptavidine HRP sur le plastique, ou à une liaison non spécifique de l'enzyme sur l'ADN capturé. Cet exemple montre donc qu'il est possible de

réaliser un double marquage de l'ADN en une seule étape, ce double marquage pouvant être utilisé pour le capturer et le détecter simultanément.

Exemple 36 : Marquage d'un produit PCR permettant simultanément sa capture et sa détection par sondes nucléiques complémentaires : :

Cette expérience permet de démontrer qu'il est possible de détecter spécifiquement un ADN, ledit ADN étant capturé sur une surface solide en utilisant la réactivité de la fonction diazométhyle vis-à-vis d'un groupement phosphate de l'ADN.

L'expérience est réalisée avec une partie des réactifs présents dans le kit HLA-DR oligo-détection (référence 33 202, bioMérieux, France), permettant la détection d'acides nucléiques amplifié par PCR dans des microplaques, par lecture colorimétrique. Dans le cadre de l'expérience décrite, on fait réagir du *para*-Bio-EG3-PDAM sur des acides nucléiques produit par PCR. L'ADN réagit avec les fonctions diazométhyle de la molécule, et se trouve ainsi doté de biotines greffées sur ses phosphates. Il sera alors possible de capturer l'acide nucléique par incubation sur une microplaque où des molécule de streptavidine sont adsorbées, et de le révéler, grâce à une sonde constitué par un oligonucléotide complémentaire de la séquence capturée, associée à une péroxydase de raifort, cette enzyme pouvant décomposer des molécules incolores d'OrthoPhénylène Diamine (réactif Color 1 du kit) en un composé détectable à une longueur d'onde de 492 nm.

Exemple 36.1 : Capture et détection spécifique de l'ADN sur microplaque :

On réalise en double le marquage de 10 µL d'ADN obtenu par amplification PCR, en l'incubant 30 min à 60°C avec 20 µL de *para*-Bio-EG3-PDAM. Après marquage, l'ADN est purifié sur colonne QIAquick (Nucleotide Removal Kit, Qiagen, Hilden, Allemagne) suivant le protocole recommandé par le fournisseur et l'éluat final est recueuilli dans 50 µL de tampon EB (Tris-HCl 10 mM, pH 8,5). Quatre-vingt-cinq (85) µL du mélange de ces

éluats sont dénaturés par 8,5 µL de réactif R4 (NaOH 2N) durant 5 min à température ambiante, puis la solution est ensuite neutralisée par 8,5 µL de réactif R5 (acide acétique 2 N). On rajoute au mélange 850 µL de tampon d'hybridation (R6 – Tris-HCL 10 mM, pH 7,0 BND 0,2 g/L, Ciproflaxacine 0,01 g/L.) et 85 µL d'oligonucléotide de détection (R7 – Phosphate de sodium 4 mM, Phosphate de potassium 1 mM, pH 7,0, Sérum albumine bovine 0,1 %, Phénol 0,5%). Cent (100) µL de cette préparation sont déposés soit sur le contrôle positif d'une barrette R1 fournie avec le kit, (capture par une séquence consensus du gène amplifié), soit sur une plaque Combiplate 8 streptavidine coated (référence 95029263, Labsystem, Helsinki, Finlande), soit sur une plaque témoin Maxisorp (Nunc, Danemark).

Parallèlement, la même réaction d'hybridation est réalisée sur une dilution, dans du tampon EB, au dixième et au centième, de la préparation d'ADN, afin de tester la sensibilité de la technique.

Exemple 36.2 : Réalisation des témoins :

Des témoins ont été simultanément réalisés de la manière suivante :

A - Comparaison au kit HLA-DR :

Dix (10) µL d'ADN obtenus par amplification PCR, comme décrit à l'exemple 5, sont incubés 30 min à 60°C avec 20 µL de *para*-Bio-EG3-PDAM. Après marquage, l'ADN est purifié sur colonne QIAquick, et l'éluat final est recueilli dans un volume de 50 µL de tampon EB. Les 45 µL d'éluat sont dénaturés par 4,5 µL de réactif R4 durant 5 min à température ambiante, puis la solution est neutralisée par 4,5 µL de réactif R5. On rajoute au mélange 450 µL de tampon d'hybridation R6 et 45 µL d'oligonucléotide de détection. Cent (100) µL de cette préparation sont déposés soit sur le contrôle positif d'une barrette R1 fournie avec le kit (hybridation avec une séquence consensus du gène amplifié), soit sur une plaque Combiplate 8 streptavidine, soit sur une plaque témoin Maxisorp.

B - Hybridation réalisée sur un ADN ne s'hybridant pas à la sonde spécifique :

Dix (10) µL d'ADN obtenu par amplification PCR, comme décrit dans l'exemple 5, sont incubés 30 min à 60°C avec 20 µL de *para*-Bio-EG3-PDAM. On traite alors l'échantillon de façon identique à processus décrit précédemment dans l'exemple A.

C - Témoin sans ADN :

Dix (10) µL de réactif R6 (tampon d'hybridation) et 100 µL de réactif R7 (oligonucléotide de détection) sont déposés soit sur le contrôle positif d'une barrette R1 fournie avec le kit, sur une plaque streptavidine, soit sur une plaque témoin de type Maxisorp.

Toutes les barrettes des protocoles ci-dessus sont incubées une heure et demie à 37°C, puis lavées par trois fois 100 µL de tampon PBS Tween (réactif color 0 HLA) puis la présence de la sonde de détection spécifique est révélée par addition de 100 µL du réactif chromogénique (Réactif Color 2, , PBS pH 7,0 ; Tween 1 %, BND 0,2 g/L ; Ciproflaxacine 0,01 g/L), incubé 20 min dans l'obscurité, la réaction étant ensuite bloquée par 50 µL de H_2SO_4 (1,8 N réactif Color 3). L'absorbance du milieu réactionnel est alors mesurée à 492 nm.

Résultats :

Conditions	Barette R1 Kit HLA DR (Capture spécifique)	Combiplate (Streptavidine)	Maxisorp (Témoin)
A - ADN marqué - Ex 36.1	**1215**	**2160**	**16**
A - ADN marqué (dilué au 1/10ème)	NA	**900**	NA
A - ADN marqué (dilué au 1/100ème)	NA	**53**	NA
B - Témoin ADN non marqué - Ex 36.2 A	**1153**	**40**	**17**
C- Témoin ADN bactérien - Ex 36.2 B	**24**	**15**	NA
D - Témoin sans ADN - Ex 36.2 C	**13**	**12**	**17**

Tableau 20 : Détection spécifique sur microplaque d'un ADN issus de PCR

Les résultats du tableau 20 indiquent une excellente amplification de la cible permettant d'envisager une utilisation dans un cadre diagnostic. Cet exemple montre que le marquage au niveau des groupements phosphate permet la capture de l'ADN et n'empêche par l'hybridation spécifique sur celui-ci.

Exemple 37 : Capture et amplification d'ADN issu d'un lysat bactérien et marqué par *para*-Bio-EG3-PDAM :

Cet exemple montre qu'il est possible de capturer et d'amplifier un ADN bactérien en utilisant une capture basée sur la réactivité de la fonction diazométhyle sur le phosphate de l'acide nucléique.

Dans le cas présent, les acides nucléiques contenus dans un lysat bactérien et marqués par du *para*-Bio-EG3-PDAM, sont capturés sur des billes magnétiques recouvertes de streptavidine. L'utilisation de billes magnétiques permet d'immobiliser celles-ci par aimantation lors des lavages successifs visant à éliminer les résidus cellulaires, présents dans le milieu réactionnel, ces résidus devant être éliminés car ils peuvent inhiber l'amplification par PCR qui est pratiqué par la suite.
Les produits d'amplification ont pu être analysés par passage sur puces à ADN.

L'ADN bactérien est obtenu par lyse des cellules contenues dans une culture de *Mycobacterium tuberculosis*. La lyse s'effectue par lyse mécanique. Plus précisément, elle est réalisée par sonication, l'échantillon liquide traité contenant des billes de verre. Un tel procédé est déjà bien décrit par la Demanderesse dans sa demande de brevet WO-A-99/15621, en ce qui concerne les billes, et dans sa demande de brevet WO-A-00/60049, en ce qui concerne la sonication. La sonication peut également être réalisée à l'aide d'un bain liquide.
Toutefois d'autres techniques, connues de l'homme du métier, peuvent être utilisées comme celles décrites dans le brevet US-A-5,902,746 et les demandes de brevet WO-A-98/54306 et WO-A-00/05338. Tous ces titres de Propriété Industrielle appartiennent à la Demanderesse.

L'ADN bactérien a été quantifié par Picogreen (P-7589 ; Molecular Probes, Eugene, OR, USA)suivant le protocole décrit par le fournisseur, à une concentration de 10^7 copies par µL.

Dix (10) µL de lysat sont incubés en présence de 20 µL de *para*-Bio-EG3-PDAM durant 30 minutes à 60°C. Parallèlement, 10 µL de lysat sont incubés dans 20 µL d'eau pure (Sigma) dans les mêmes conditions.

Le milieu réactionnel est ensuite purifié sur colonne QIAquick (Nucleotide Removal Kit, Qiagen, Hilden, Allemagne). Le protocole de purification utilisé est celui recommandé par le fournisseur. Le volume final d'élution est de 50 µL.

Les fragments d'ADN marqués sont ensuite capturés sur des billes magnétiques Dynal (Dynabeads M-280 streptavidin ; référence 112.05 ; Dynal Biotech ASA, Oslo, Norway), qui sont préparées selon le protocole suivant :
Quatre-vingt-dix (90) µL de billes Dynal sont lavés deux fois par 200 µL d'eau pure Free (Sigma), puis sont repris par 200 µL de tampon PEG (0,1 M Na PO_4, pH 7, 0,5 M NaCl ; 0,65 % Tween 20 ; 0,14 mL ADN de sperme de Hareng (Référence 15634-017, GibcoBRL) ; 2 % PEG 4000) et incubés 30 min à 37°C. Elles sont ensuite lavées deux fois par 200 µL de tampon PBS 1X Tween 20 0,5%, puis finalement reprises par 90 µL du même tampon.

Dix (10) µL des éluats d'ADN marqués ou non marqués sont incubés 5 min à température ambiante avec 40 µL de tampon PEG et 2,5 µL de la préparation de billes magnétiques décrite précédemment.

Les billes sont ensuite lavées trois fois par 200 µL de tampon PBS 1 X tween 0,5%, reprises dans 200 µL d'eau et incubées 20 min à 60°C, puis lavées de nouveau quatre fois par 200 µL de PBS tween. Les billes sont finalement reprises par 25 µL d'eau et une PCR est réalisée en suivant le protocole décrit à l'exemple 5. Deux contrôles réactionnels sont effectués l'un avec 25 µL d'eau pure et l'autre avec 2,5 µL de billes préparées et lavées dans les mêmes conditions que les échantillons biologiques, et reprises dans 25 µL d'eau.

Résultat :

Les produits de PCR sont ensuite quantifiés par Picogreen (P-7589 ; Molecular Probes, Eugene, OR, USA) suivant le protocole décrit par le fournisseur. Cette méthode repose sur l'utilisation d'une molécule (Picogreen) ayant la particularité de devenir fluorescente uniquement lorsqu'elle se positionne à l'intérieure d'une molécule d'ADN (en s'intercalant entre les bases). En raison du caractère très spécifique de cette intercalation, et du fait que le signal fluorescent produit est directement proportionnel à la quantité d'ADN présent dans le milieu, il est possible de doser de cette manière, de façon très précise, la concentration d'acide nucléique présente dans un échantillon. Le signal est alors exprimé en rfu (relative fluorescent unit).

L'analyse des résultats de PCR sur gel montre la présence d'une seule bande spécifique à la taille attendue dans les échantillons réalisé à partir d'ADN génomique marqué par *para*-Bio-EG3-PDAM. Les bandes ne sont pas détectées lorsque la PCR a été effectuée à partir d'ADN génomique non marqué. Une quantification de l'ADN par Picogreen permet de confirmer la production d'ADN à partir d'ADN génomique capturé sur billes.

Conditions	rfu
Témoins non Marqués	51
ADN marqués	170
Bruit de fond	20

Tableau 21 : Quantification de l'ADN produit par PCR, à partir d'un lysat bactérien, capturé et purifié par le *para*-Bio-EG3-PDAM

Une analyse sur puce à ADN, suivant le protocole décrit à l'exemple 8 permet de confirmer la spécificité de l'amplification comme le montre le taleau 34 ci-dessous.

	Homologie %	I (rfu)	B(rfu)	I/B
Echantillon	98	11531	723	16

Tableau 22 : Détection spécifique de la cible capturée et purifiée par le *para*-Bio-EG3-PDAM

Cet exemple montre qu'il est possible de préparer un échantillon biologique de façon à amplifier l'acide nucléique qu'il contient, en utilisant une technique de capture basée sur la réactivité de la fonction diazométhyle sur les groupements phosphate de celui-ci.

Exemple 38 : Amplification successive de deux gènes à partir d'un ADN bactérien capturé sur un support solide :

Cet exemple montre qu'il est possible d'amplifier à plusieurs reprises et sur des cibles différentes un ADN capturé grâce à la réactivité de la fonction diazométhyle sur les groupements phosphates de celui-ci .

Dans le cas présent, les acides nucléiques, contenus dans un lysat bactérien et marqués par du *para*-Bio-EG3-PDAM, sont capturés sur des billes magnétiques recouvertes de streptavidine. L'utilisation de billes magnétiques permet de conserver celles-ci par aimantation lors des lavages successifs visant à éliminer les résidus cellulaires présents dans le milieu réactionnel, ces résidus devant être éliminés car ils peuvent inhiber les amplifications par PCR, qui seront pratiquées par la suite. Ces amplifications vont avoir lieu sur deux gènes différents présents dans l'ADN génomiques, respectivement désignés sous les noms de 16S et rpoB. Ces deux gênes sont ensuite analysés par utilisation de puces à ADN.

L'ADN bactérien est obtenu par lyse des cellules contenues dans une culture de *Mycobacterium tuberculosis*, suivant le protocole déjà décrit dans l'exemple 37. L'ADN bactérien a été quantifié par Picogreen suivant le protocole décrit par le fournisseur, à une concentration de 10^7 copies par µL.

Dix (10) µL de lysat sont incubés en présence de 20 µL de *para*-Bio-EG3-PDAM durant 30 minutes à 60°C. Parallèlement, 10 µL de lysat sont incubés dans 20 µL d'eau pure (Sigma) dans les mêmes conditions.

Le milieu réactionnel est ensuite purifié sur colonne QIAquick (Nucleotide Removal Kit, Qiagen, Hilden, Allemagne). Le protocole de purification utilisé est celui recommandé par le fournisseur. Le volume final d'élution est de 50 µL.

Les fragments d'ADN marqués sont ensuite capturés sur des billes magnétiques Dynal, qui sont préparées suivant le protocole suivant :
Quatre-vingt-dix (90) µL de billes Dynal sont lavées deux fois par 200 µL d'eau pure Free (Sigma), puis sont reprises par 200 µL de tampon PEG (0,1 M Na PO_4, pH 7, 0,5 M NaCl ; 0,65 % Tween 20 ; 0,14 mL ADN de sperme de saumon (Gibco) ; 2 % PEG 4000) et incubées 30 min à 37°C. Elles sont ensuite lavées deux fois par 200 µL de tampon PBS 1X Tween 20 0.5%, puis finalement reprises par 90 µL du même tampon.

Dix (10) µL des éluats d'ADN marqués ou non marqués sont incubés 5 min à température ambiante avec 40 µL de tampon PEG et 2,5 µL de la préparation de billes magnétiques décrite précédemment.

Les billes sont ensuite lavées trois fois par 200 µL de tampon PBS 1 X tween 0,5 %, reprise dans 200 µL d'eau et incubées 20 min à 60°C, puis lavées de nouveau quatre fois par 200 µL de PBS tween. Les billes sont finalement reprises par 25 µL d'eau et une PCR est réalisée en suivant le protocole décrit dans l'exemple 5. Deux contrôles réactionnels sont effectués l'un avec 25 µL d'eau pure (Sigma) et l'autre avec 2,5 µL de billes préparées et lavées dans les même, conditions que les échantillons biologiques, et reprise, dans 25 µL d'eau.

Après amplification, le milieu réactionnel est recueilli, et les billes sont séparées et lavées par 150 µL de PBS 1 X Tween 0,5%, puis resuspendues dans 25 µL d'eau pure (Sigma). Une nouvelle amplification est réalisée sur les billes, mais en présence d'amorces destinées à amplifier le gène rpoB.

Des amplifications témoins réalisés à partir d'ADN génomique non capturé sont réalisées en parrallèle sur les deux systèmes d'amplification (rpoB et 16S).

Résultat :

Les produits de PCR obtenus sont ensuite analysés par puces à ADN suivant le protocole décrit à l'exemple 8.

Conditions		Homologie (%)	I (rfu)	B (rfu)	I/B
PCR témoin 16S	16S	96	4662	387	12
Résultats sur séquence rpoB	RpoB	23	237	358	1
PCR témoin rpoB	RpoB	98	5438	397	14
Résultats sur séquence 16S	16S	17	183	391	1
PCR 16S sur billes	16S	98	2726	534	5
Résultats sur séquence rpoB	RpoB	8	161	480	<1
PCR rpoB sur billes lavées	RpoB	97	3205	349	9
Résultats sur séquence 16S	16S	14	84	358	<1

Tableau 23 : Analyse sur puces à ADN des amplicons ADN issus des PCR amplifiées successivement ou non

L'analyse des résultats de PCR sur gel montre la présence d'une seule bande spécifique à la taille attendue dans les échantillons réalisés à partir d'ADN génomique marqué par *para*-Bio-EG3-PDAM. Les bandes ne sont pas détectées lorsque la PCR a été effectuée à partir d'ADN génomique non marqué.

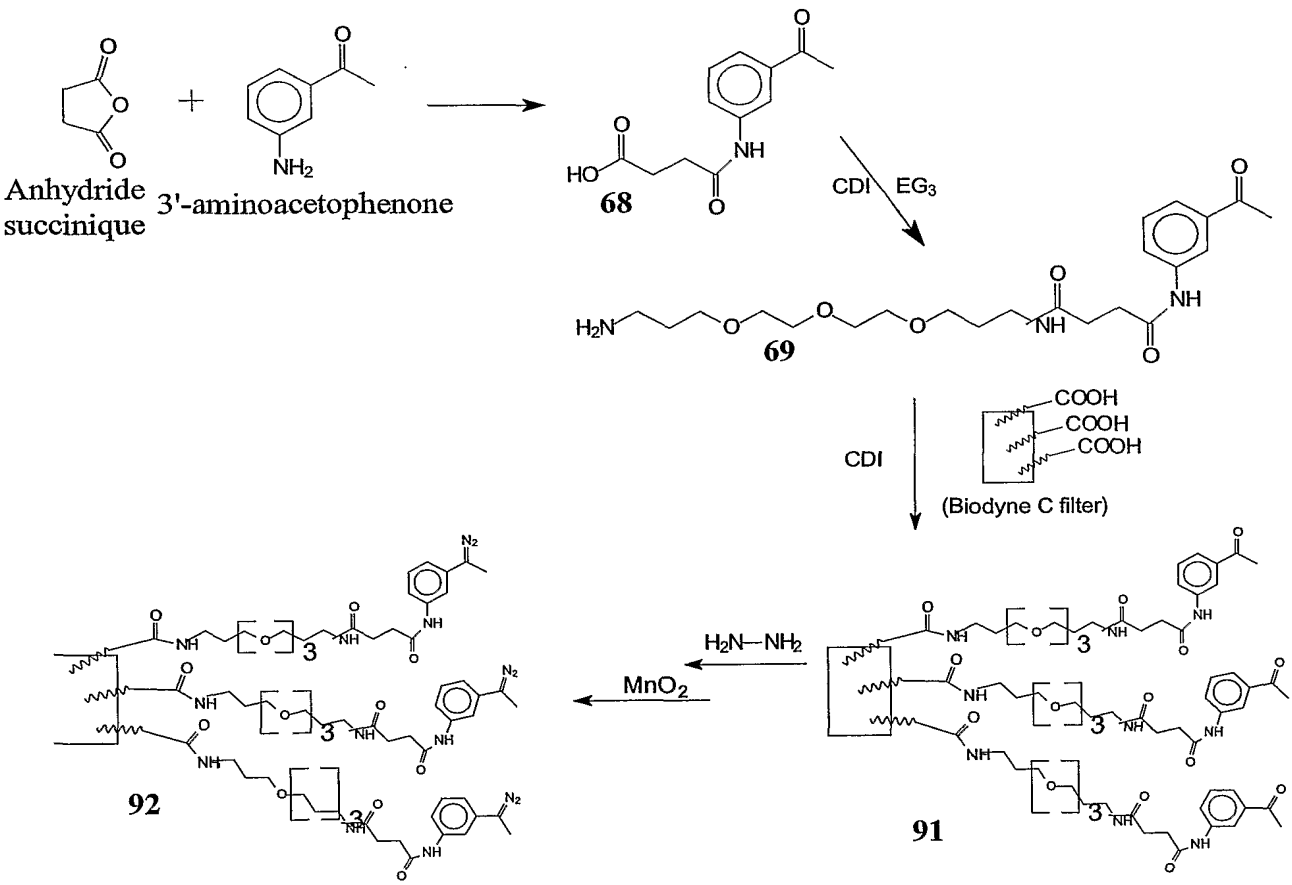
Une analyse par puce à ADN, telle que présentée sur le tableau 23 ci-dessus, permet de confirmer la spécificité des deux amplifications successives, et donc la capacité de réaliser une amplification successive de plusieurs gènes à partir d'ADN immobilisé sur un support solide, ceci permettant d'éviter la mise au point de systèmes multiplex, qui réduisent souvent de façon importante la sensibilité et l'efficacité des amplifications d'acides nucléiques.

Exemple 39 : Capture et Amplification d'ADN sur une membrane de Nylon portant des groupements diazométhyles :

Une membrane de nylon activée de façon à porter des groupements diazométhyles a été utilisée afin de capturer de l'ADN bactérien, dans le but de l'amplifier par PCR.

Exemple 39.1 : Modification du filtre Biodyne C :

Schéma de synthèse :



Composé **68** :

On solubilise la 3'-aminoacetophenone (14,5 g, 107 mmol) dans 50 ml de DMF anhydre. On ajoute l'anhydride succinique (10,7 g, 107 mmol) et on laisse sous agitation, sous argon et à température ambiante. Après 6 h, la solution est concentrée sous vide et 50 ml de méthanol sont ajoutés. Le précipité obtenu est filtré et lavé avec méthanol et éther. On obtient ainsi 19,4 g (81 %) de produit **68** sous forme d'une poudre de couleur blanc cassé.

RMN ^{1}H (200 MHz, DMSO-d_6): d= 2,5-2,6 (m, 7H); 7,45 (t, 1H); 7,64 (d, 1H); 7,83 (d, 1H); 8,19 (s, 1H); 10,16 (s, 1H); 12,12 (s, 1H).

Composé **69** :

On solubilise 5,07 g (22 mmol) du composé **68** dans 10 ml de DMF anhydre, sous argon. On met sur glace et on ajoute 5,00 g (32 mmol) de carbonyldiimidazole. Apres 20 min, on ajoute lentement 20 ml (94,6 mmol) du 4,7,10-trioxatridecanediamine (EG_3). Après 3h de réaction à température ambiante, on évapore le DMF et on reprend le résidu dans 100 ml de CH_2Cl_2. On fait des extractions avec du $NaHCO_3$ saturé et H_2O, après quoi la phase organique est séchée avec du Na_2SO_4 anhydre et le solvant évaporé. On obtient ainsi 4,34 g (46 %) du produit **69**.

RMN ^{1}H (200 MHz, DMSO-d_6): d= 1,59 (m, 2H); 1,87 (m, 2H); 2,16 (s, 3H); 2,40 (m, 2H); 2,55 (m, 2H); 3,08 (m, 2H); 3,45 (m, 16H); 7,30 (t, 1H); 7,42 (d, 1H); 7,70 (d, 1H); 7,83 (t, 1H); 7,97 (s, 1H); 10,00 (s, 1H).

Composé **91** :

Un rectangle de 4 cm^2 est découpé sur une feuille de filtre Biodyne C (référence 60314 ; Pall Gelman Laboratory ; Ann Arbor ; Michigan ; USA), introduit dans un flacon et mis en contact avec 0,97 g (6 mmol) de carbonyldiimidazole (CDI) dans 3 ml de DMF anhydre, sur glace, sous argon et sous forte agitation. Après 20 min, la solution est éliminée et le filtre lavé avec du DMF. Une quantité de 0,53 g du produit **68** (1 mmol) dans 3 ml de DMF anhydre est ensuite ajoutée, et la réaction se réalise pendant la nuit à température ambiante. La solution est ensuite enlevée et le filtre est rincé avec de l'éthanol, séché sous vide et gardé sous argon.

Composé **92** :

Le filtre modifié **91** est mis dans une solution de 97 ml d'hydrazine hydrate (2 mmol) dans 4 ml d'éthanol absolu. La solution est mise à reflux pendant 5 h. Après avoir laissé refroidir, le filtre est lavé avec H_2O, éthanol et éther, séché sous vide et mis sous argon. Ensuite, on ajoute 4 ml de DMF anhydre et 86 mg de MnO_2 (1 mmol), et on laisse réagir sous forte agitation. Après 20 min, la solution est rejetée, et le filtre est rincé avec du DMF et de l'éther. Le filtre modifié-diazométhyl **92** est conservé sous argon, à une température de -19 à -31°C.

Exemple 39.2 : Essais biologiques :

La membrane activée est découpée en petits fragments de 2 mm^2 qui vont être incubés pendant 30 minutes à température ambiante dans 25 µl de lysat bactérien de *Mycobacterium tuberculosis*, préparé par lyse mécanique, suivant la même technique et la même concentration finale que dans l'exemple 37, et 375 µl d'eau pure (Sigma).

La membrane est ensuite placée à 65°C pendant 60 min dans 100 mL de tampon de lavage (5 % Formamide (Sigma), 1X SSPE (Perbio), 0.01 % triton X-100) afin d'éliminer les acides nucléiques non spécifiques adsorbés sur la membrane, puis celle-ci est stockée dans 1 mL d'eau pure avant amplification.

La PCR est pratiquée comme décrite dans le paragraphe 5.1, le volume réactionnel étant complété par une quantité suffisante d'eau pure.

Parallèlement, des contrôles sont effectués suivant le même processus avec des membranes ne pouvant pas lier de façon covalente les acides nucléiques :

- Membrane Biodyne C non modifiée (membrane A),
- Membrane Biodyne modifiée chimiquement suivant le processus décrit, mais non traitée par du DMF anhydre et du MnO_2 ; ce contrôle permet de vérifier le comportement de la membrane en l'absence de groupements diazométhyles (membrane B), et

- Membrane Biodyne C non modifiée, mais traitée 20 min sous forte agitation par du DMF anhydre et du MnO_2, ce contrôle permettant de vérifier que cette dernière étape ne modifie pas l'adsorption de l'ADN sur la membrane (membrane C).

Afin de contrôler l'absence d'inhibition de la PCR provoquée par le traitement des membranes, un autre fragment des membranes A, B, C est amplifié simultanément avec 25 µl de lysat bactériens.(tubes A', B', C')

Les produits de PCR sont ensuite quantifiés par Picogreen suivant le protocole décrit par le fournisseur.

Résultat :

Tests effectués	Signal (rfu)
Membrane modifiée	111
Membrane non modifiée (A)	18
Membrane non modifiée, co-amplifiée avec 25 µL de lysat bactérien (A')	260
Membrane modifiée non activée (B)	26
Membrane modifiée non activée, co-amplifiée avec 25 µL de lysat bactérien (B')	264
Membrane non modifiée ayant subit une activation (C)	21
Membrane non modifiée ayant subit une activation, co-amplifiée avec 25 µL de lysat bactérien (C')	268

Tableau 24 : Quantification de l'ADN obtenu par PCR à partir d'ADN de lysat bactérien capturé sur support solide

Ces résultats du tableau 24 indiquent qu'il est possible de capturer de façon covalente sur un support solide, des acides nucléiques, issus de lysat, grâce à la chimie diazométhyle. L'amplification observée n'est pas due à une adsorption non spécifique de l'ADN sur la membrane. D'autre part, on observe avec les contrôles effectués avec les PCR que les membranes ne provoquent pas d'inhibition de la réaction d'amplification.

Afin de contrôler la nature du produit amplifié sur la membrane, le produit d'amplification a été analysé par passage sur puce à ADN, suivant le protocole décrit auparavant.

5

IDT Exhibit 1020.7857

REVENDICATIONS

1. Procédé de marquage et de fragmentation d'un acide désoxyribonucléique (ADN) simple ou double brin comprenant les étapes suivantes :

- fragmenter l'ADN en créant au moins un site abasique sur ledit ADN,
- attacher un marqueur sur au moins un des fragments par l'intermédiaire d'un réactif de marquage, ledit réactif se couplant de manière covalente et majoritaire sur au moins un phosphate dudit fragment.

2. Procédé, selon la revendication 1, caractérisé par le fait que la fragmentation et le marquage sont effectués en deux étapes.

3. Procédé, selon la revendication 1, caractérisé par le fait que la fragmentation et le marquage sont réalisés en une étape.

4. Procédé, selon l'une quelconque des revendications 1 à 3, caractérisé par le fait que le site abasique est généré par action d'un milieu aqueux acide.

5. Procédé, selon la revendication 4, caractérisé par le fait que le pH du milieu acide est inférieur à 5, de préférence inférieur à 4.

6. Procédé, selon la revendication 5, caractérisé par le fait que le milieu aqueux acide est un tampon formiate de sodium à pH d'environ 3.

7. Procédé, selon l'une quelconque des revendications 1 à 6, caractérisé par le fait que l'ADN contient au moins une base modifiée susceptible de générer un site abasique.

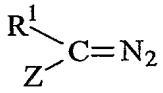
8. Procédé, selon la revendication 7, caractérisé par le fait que la base modifiée susceptible de générer un site abasique est choisi parmi les dérivés de la 8-bromo purine.

9. Procédé, selon l'une quelconque des revendications 1 à 3, caractérisé par le fait que le site abasique est généré par un agent alkylant ou un agent oxydant.

10. Procédé de marquage et de fragmentation, selon l'une quelconque des revendications 1 à 9, caractérisé par le fait que le réactif de marquage comprend, comme fonction réactive, un motif choisi parmi les composés suivants : diazométhyle ; halogénure d'alkyle ; nitrosourée ; spirocyclopropane ; aziridine ; époxyde ; trifuorosulfonates.

11. Procédé, selon la revendication 10, caractérisé par le fait que ledit réactif de marquage est la 5-(bromométhyl)fluorescéine.

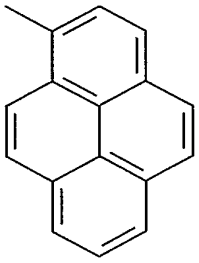
12. Procédé, selon la revendication 10, caractérisé par le fait que le réactif de marquage est choisi parmi les composés de formule (1):



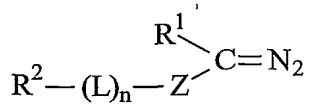
dans laquelle :

- R^1 représente H ou un groupe alkyle, alkyle substitué, aryle ou aryle substitué, et
- Z comprend un marqueur détectable.

13. Procédé, selon la revendication 12, caractérisé par le fait que Z est :

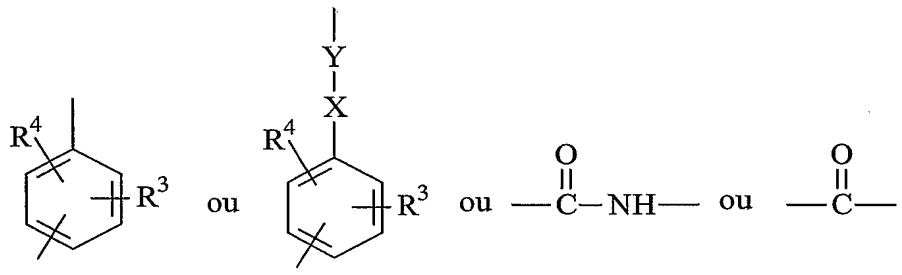


14. Procédé, selon la revendication 12, caractérisé par le fait que le réactif de marquage est choisi parmi les composés de formule (2) :



dans laquelle

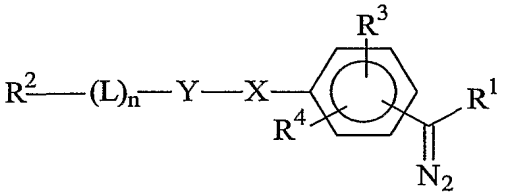
- R^1 représente H ou un groupe alkyle, aryle ou aryle substitué,
- R^2 est un marqueur détectable,
- L est un bras de liaison comportant un enchaînement linéaire d'au moins deux liaisons covalentes et n est égal à 0 ou 1, et
- Z est choisi parmi :



dans lequel :

- R^3 et R^4 représentent indépendamment l'un de l'autre : H, NO_2, Cl, Br, F, I, OR, SR, NR_2, R, NHCOR, CONHR, COOR avec R = alkyle ou aryle, et
- -Y-X- représente -CONH-, -NHCO-, $-CH_2O-$, $-CH_2S-$.

15. Procédé, selon la revendication 14, caractérisé par le fait que le réactif de marquage a la formule (3) :



dans laquelle :

- R^1 représente H ou un groupe alkyle, aryle ou aryle substitué,
- R^2 représente un marqueur détectable,
- L est un bras de liaison comportant un enchaînement linéaire d'au moins 2 liaisons covalentes et n un nombre entier égal à 0 ou 1,
- R^3 et R^4 représentent indépendamment l'un de l'autre : H, NO_2, Cl, Br, F, I, OR, SR, NR_2, R, NHCOR, CONHR, COOR avec R = alkyle ou aryle, et
- -Y-X- représente -CONH-, -NHCO-, $-CH_2O-$, $-CH_2S-$.

16. Procédé, selon l'une quelconque des revendications 14 ou 15, caractérisé par le fait que R^2 est un composé fluorescent ou un haptène.

17. Procédé, selon l'une quelconque des revendications 10 à 16, caractérisé par le fait que ledit réactif de marquage est soluble dans un solvant miscible à l'eau.

18. Procédé de détection d'un acide désoxyribonucléique cible (ADN) simple ou double brin comportant les étapes suivantes :

- fragmenter et marquer lesdits ADN, selon l'un quelconque des procédés selon les revendications 1 à 17,
- hybrider les fragments marqués sur au moins une sonde nucléique suffisamment spécifique de l'acide nucléique cible, et
- détecter l'hybride formé par l'intermédiaire du marqueur.

19. Procédé de détection d'un acide désoxyribonucléique cible (ADN) double brin, selon la revendication 18, comportant en outre une étape de dénaturation après la fragmentation et le marquage.

20. Procédé, selon la revendication 19, caractérisé par le fait que la fragmentation, le marquage et la dénaturation sont réalisés en une seule étape.

21. Procédé de détection d'un acide nucléique cible comportant les étapes suivantes :

- amplifier enzymatiquement l'acide nucléique cible pour générer une multitude d'amplicons ADN double brin,
- fragmenter et marquer lesdits amplicons ADN, selon l'un quelconque des procédés selon les revendications 1 à 17,
- hybrider les fragments marqués sur au moins une sonde nucléique suffisamment spécifique de l'acide nucléique cible, et
- détecter l'hybride formé par l'intermédiaire du marqueur.

22. Procédé, selon la revendication 21, comportant en outre une étape de dénaturation après l'étape de fragmentation et marquage.

23. Procédé, selon la revendication 22, caractérisé par le fait que la fragmentation, le marquage et la dénaturation sont réalisés en une seule étape.

24. Procédé de mise en évidence d'un polymorphisme réparti en des positions prédéterminées ou non d'un acide nucléique cible par la présence d'une pluralité de délétions et/ou insertions et/ou mutations dans la séquence dudit acide nucléique cible par rapport à une séquence dite de référence comportant les étapes suivantes :

- disposer d'un ADN cible comportant l'ensemble du polymorphisme à étudier, ledit ADN étant éventuellement généré par une technique d'amplification enzymatique.
- fragmenter et marquer ledit ADN par un procédé, selon l'une quelconque des revendications 1 à 17.
- hybrider lesdits fragments sur une pluralité de sondes nucléiques dites sondes de capture, la pluralité de sondes de capture étant fixé sur un support solide et la pluralité de sondes de capture couvrant dans son ensemble au moins le polymorphisme à étudier.
- détecter les hybrides formés entre les fragments marqués et au moins une partie des sondes nucléiques par l'intermédiaire du marqueur et en déduire le polymorphisme de l'ADN cible.

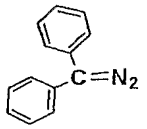
25. Procédé, selon la revendication 24, comportant en outre une étape de dénaturation après l'étape de fragmentation et marquage.

26. Procédé, selon la revendication 25, caractérisé par le fait que la fragmentation, le marquage et la dénaturation sont réalisés en une seule étape.

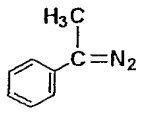
27. Procédé, selon l'une quelconque des revendications 24 à 26, caractérisé par le fait que le support solide comporte au moins dix (10) sondes nucléiques de séquences différentes, avantageusement au moins quatre cents (400), de préférence au moins mille (1000).

28. Utilisation d'un fragment d'acide nucléique marqué, selon le procédé de l'une quelconque des revendication 1 à 17, comme sonde de détection d'un acide nucléique cible.

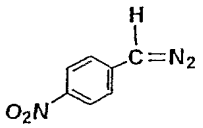
Fig. 1



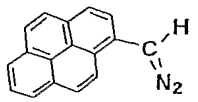
DPDAM
(DiPhénylDAM)



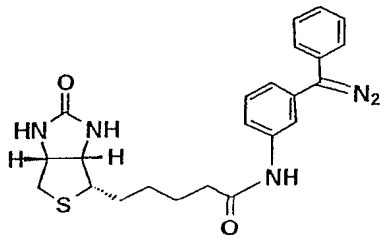
PMDAM
(PhénylMéthylDAM)



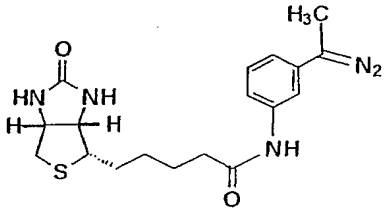
NPDAM
(NitroPhénylDAM)



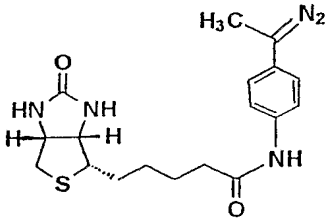
PDAM
(PyrénylDiAzoMéthane)



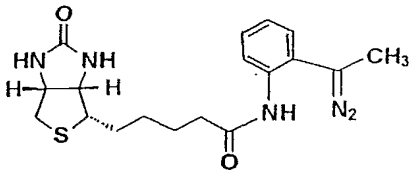
BioDPDAM
(BiotinylDPDAM)



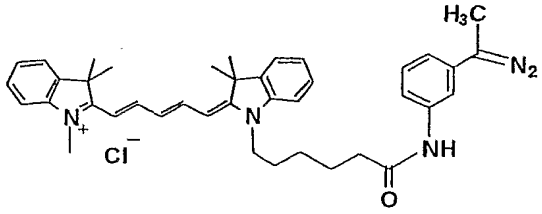
m-BioPMDAM
(*meta*-BiotinylPMDAM)



p-BioPMDAM
(*para*-BiotinylPMDAM)



o-BioPMDAM
(*ortho*-BiotinylPMDAM)



Cy5PMDAM

Fig. 2A

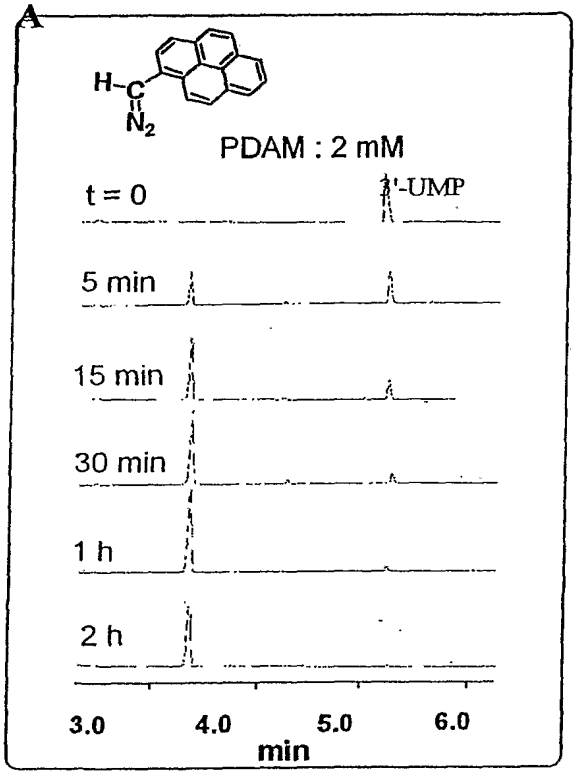


Fig. 2B

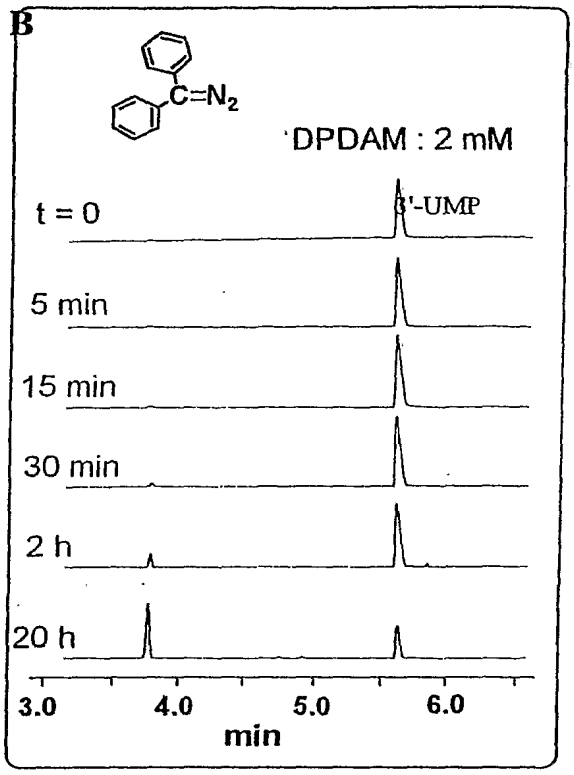


Fig. 2C

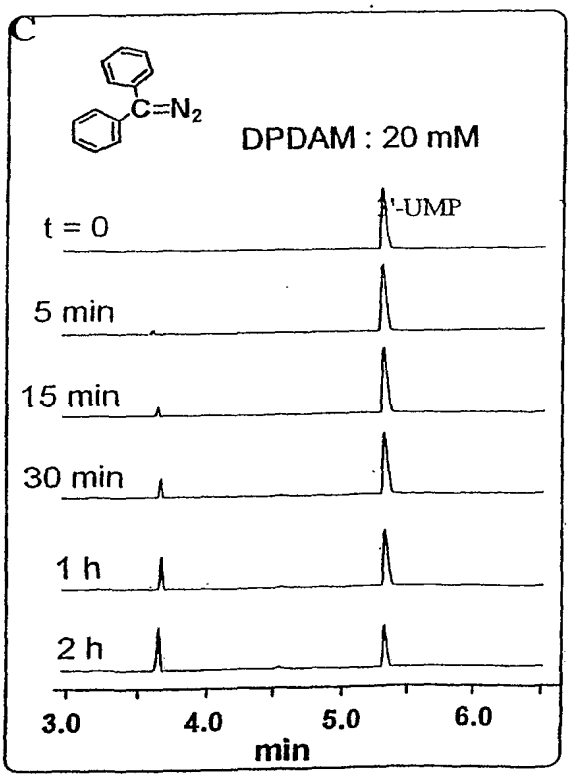


Fig. 2D

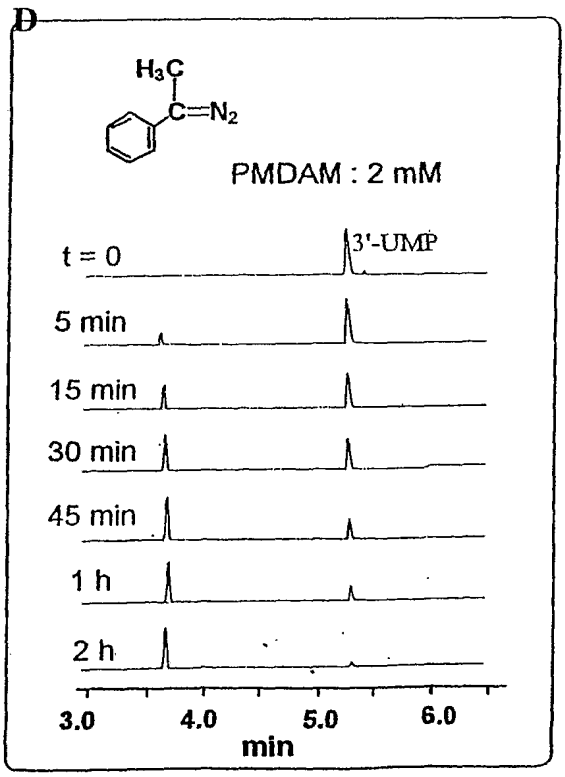


Fig. 2E

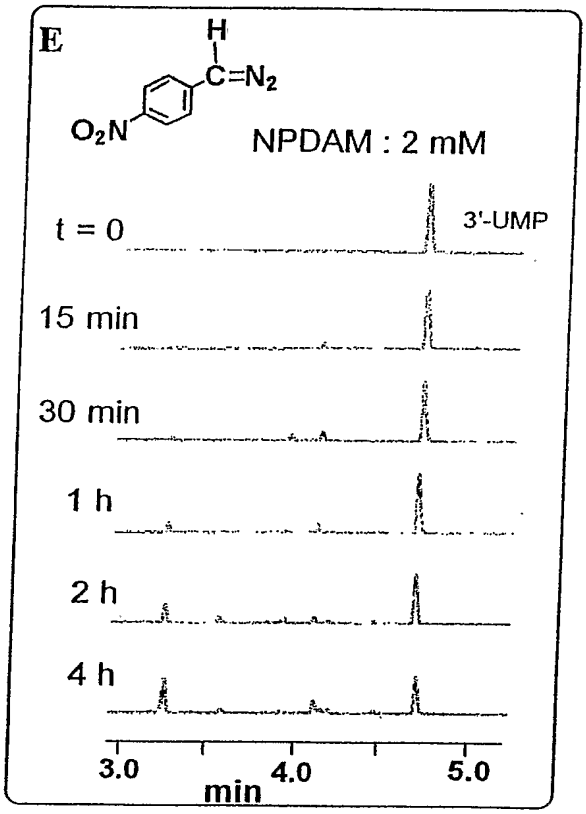


Fig. 2F

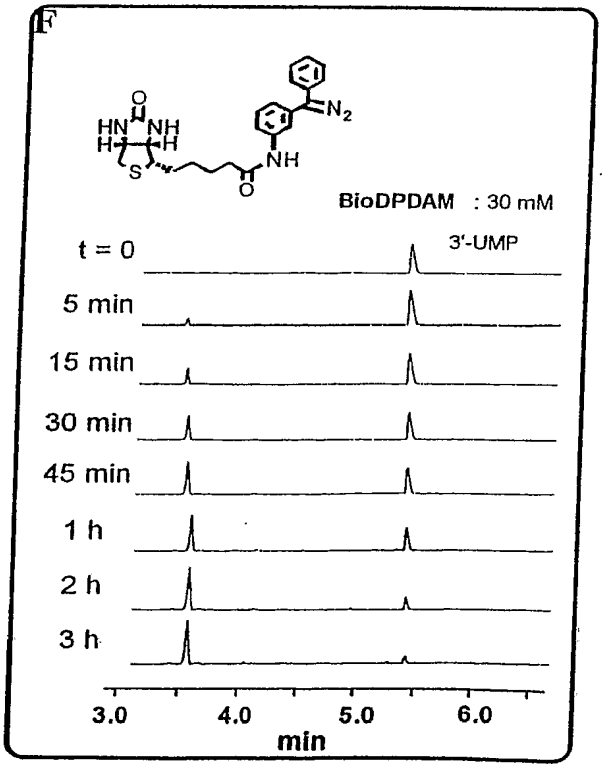


Fig. 2G

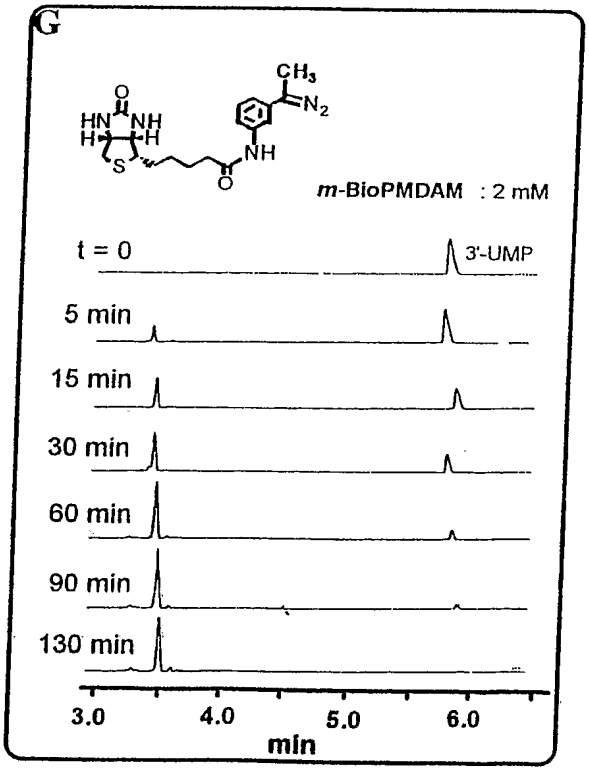


Fig. 2H

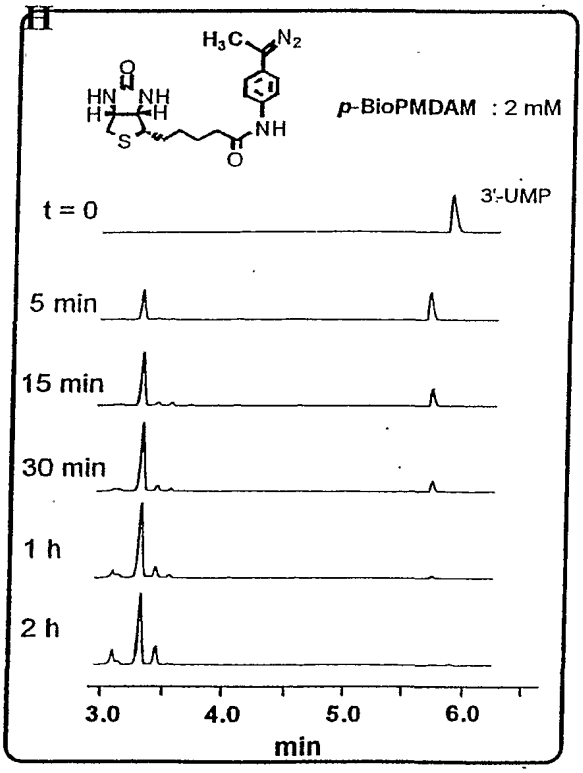

H
p-BioPMDAM : 2 mM
3'-UMP
t = 0
5 min
15 min
30 min
1 h
2 h
3.0
4.0
5.0
6.0
min

Fig. 2I

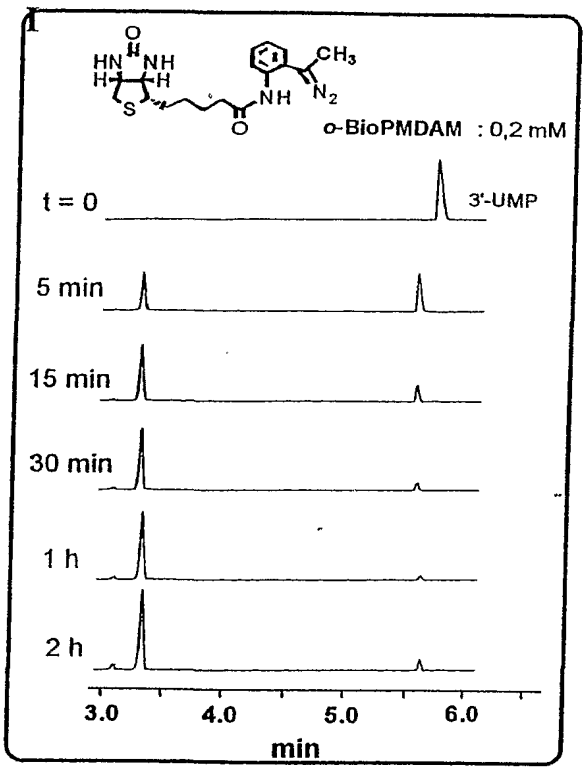

I
o-BioPMDAM : 0,2 mM
3'-UMP
t = 0
5 min
15 min
30 min
1 h
2 h
3.0
4.0
5.0
6.0
min

Fig. 3A

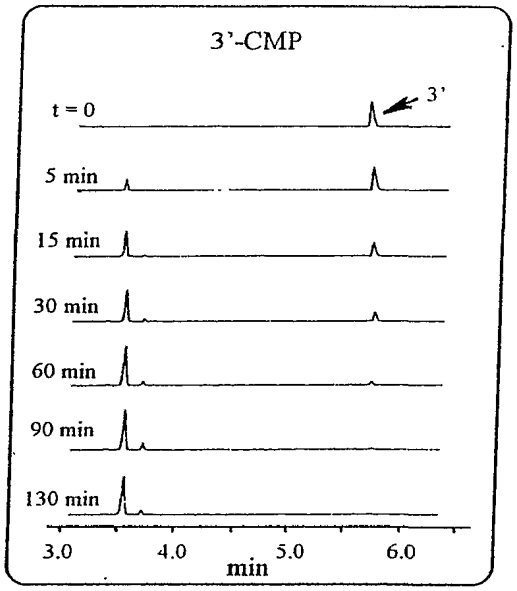


Fig. 3B

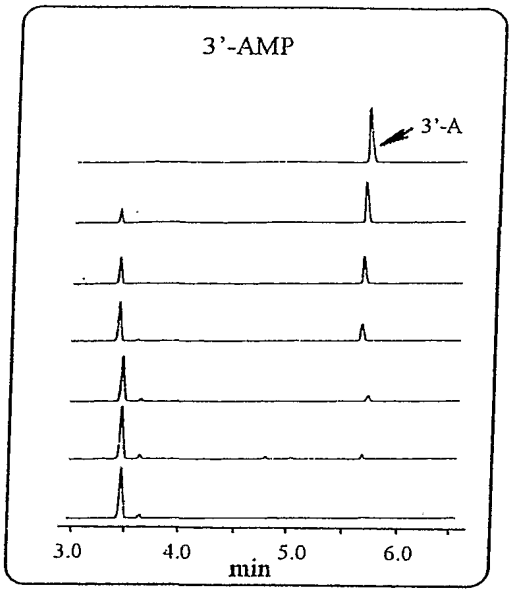


Fig. 3C

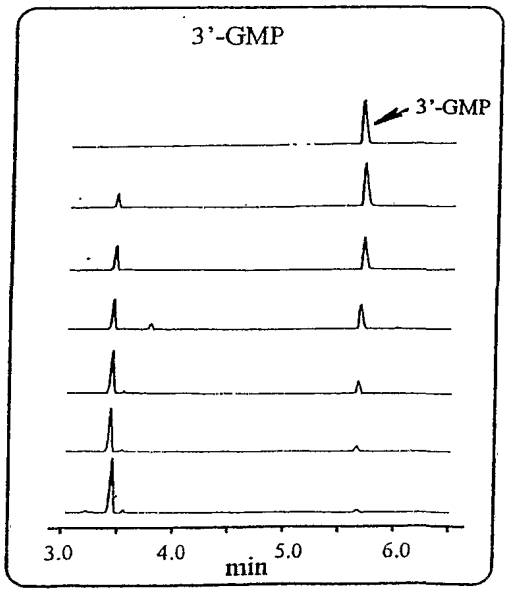


Fig. 3D

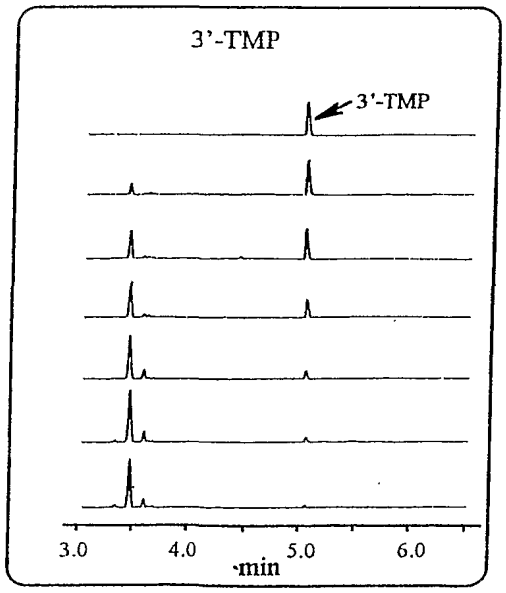
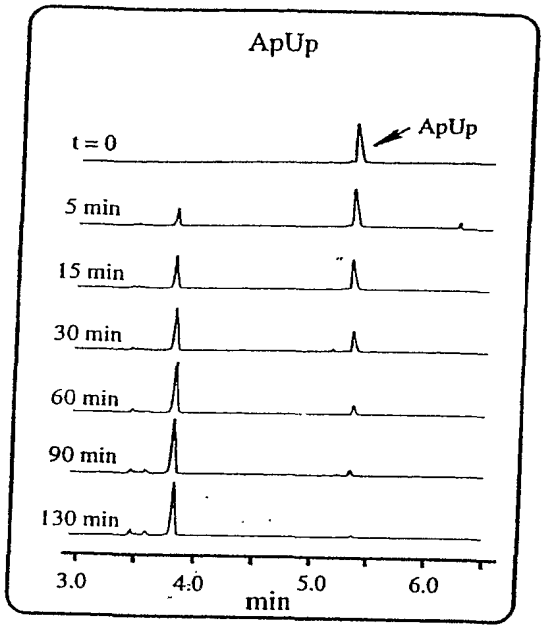


Fig. 4


ApUp
t = 0
ApUp
5 min
15 min
30 min
60 min
90 min
130 min
3.0
4.0
5.0
6.0
min

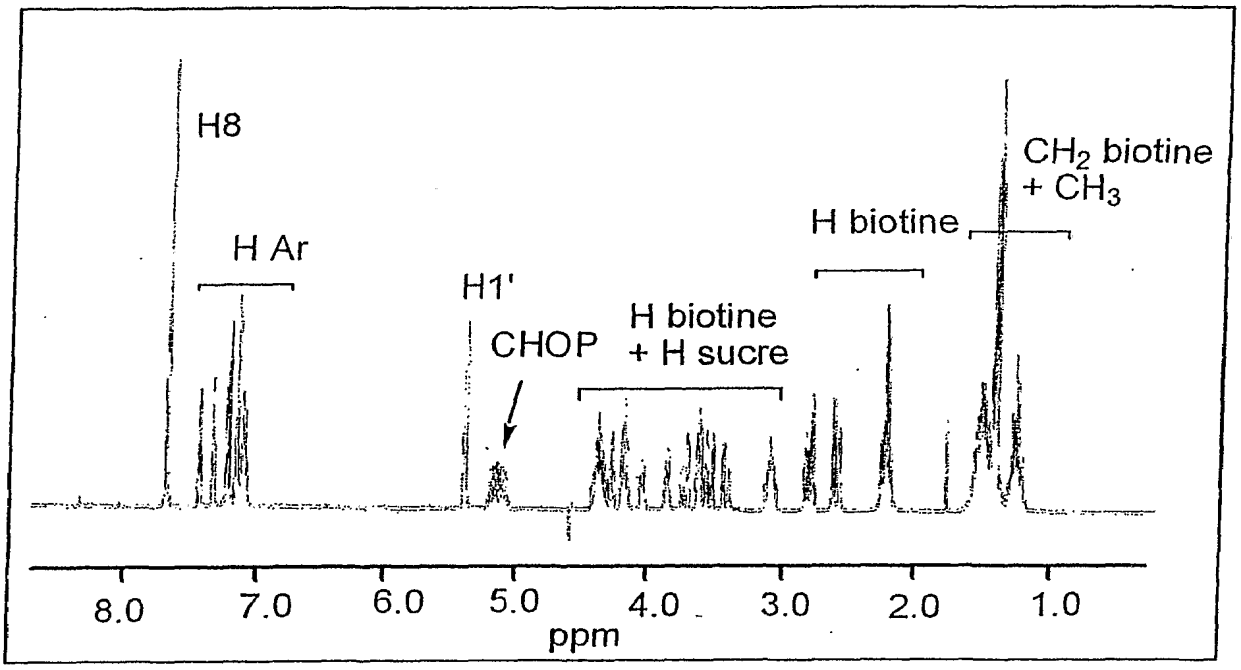

H8
H Ar
H1'
CHOP
H biotine
+ H sucre
H biotine
CH2 biotine
+ CH3
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0
ppm

Fig. 5B

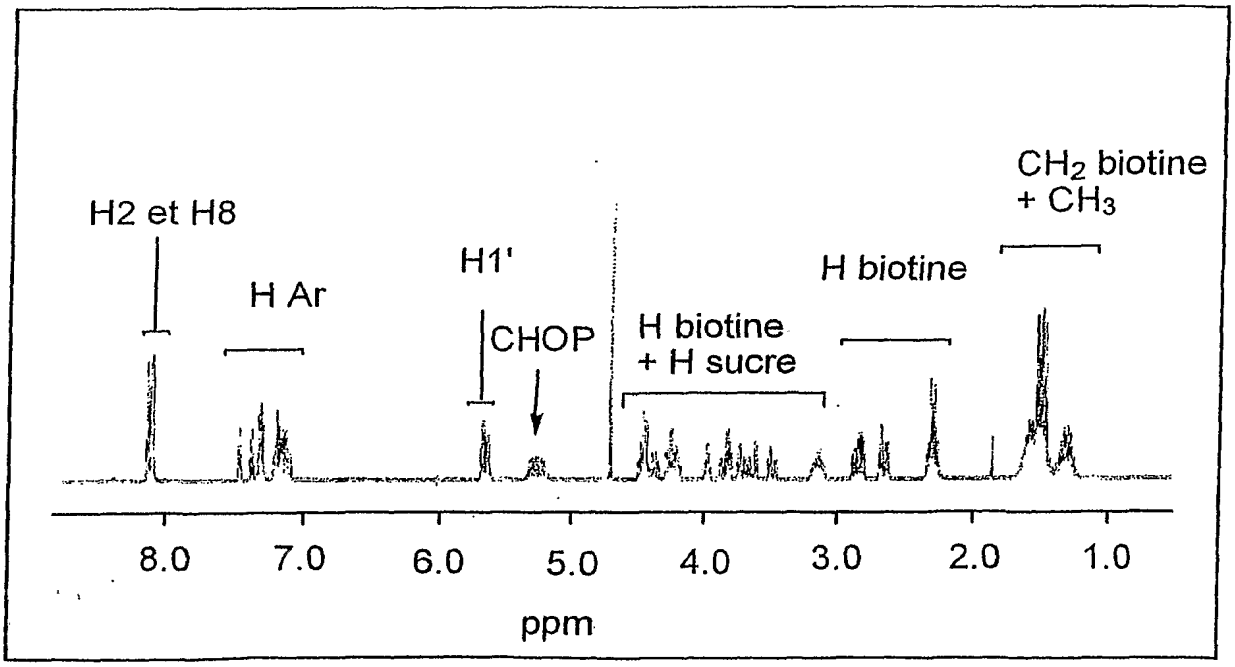

H2 et H8
H Ar
H1'
CHOP
H biotine
+ H sucre
H biotine
CH2 biotine
+ CH3
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0
ppm

Fig. 5C

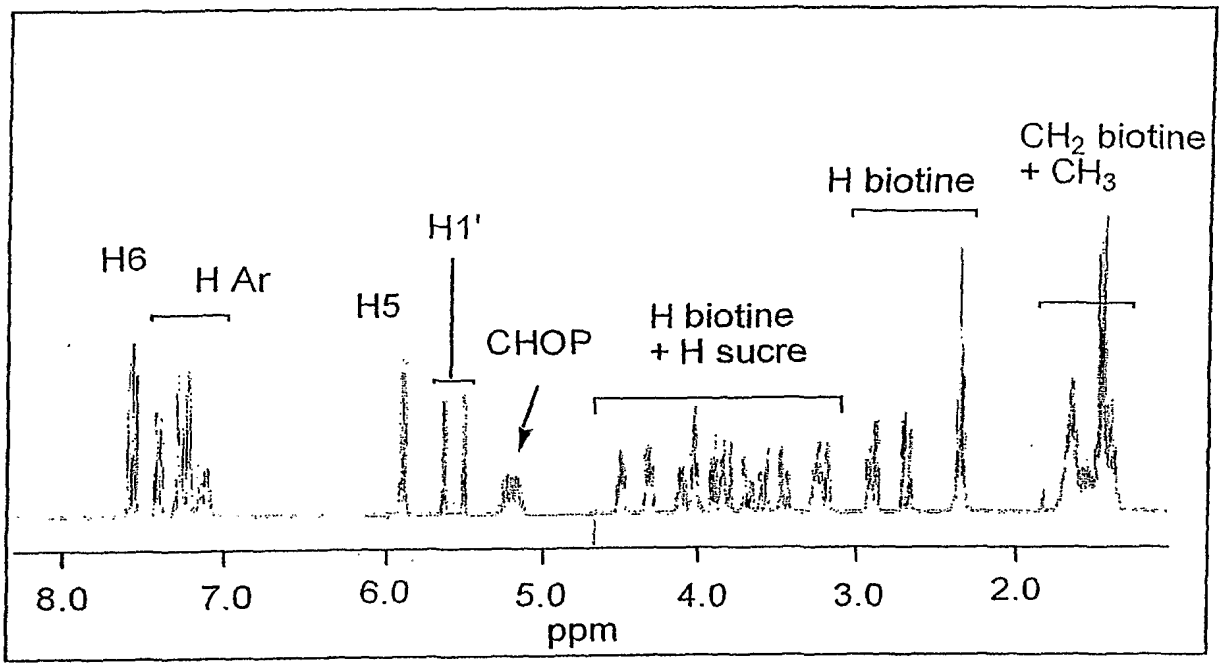

H6
H Ar
H5
H1'
CHOP
H biotine
+ H sucre
H biotine
CH2 biotine
+ CH3
8.0
7.0
6.0
5.0
4.0
3.0
2.0
ppm

Fig. 5D

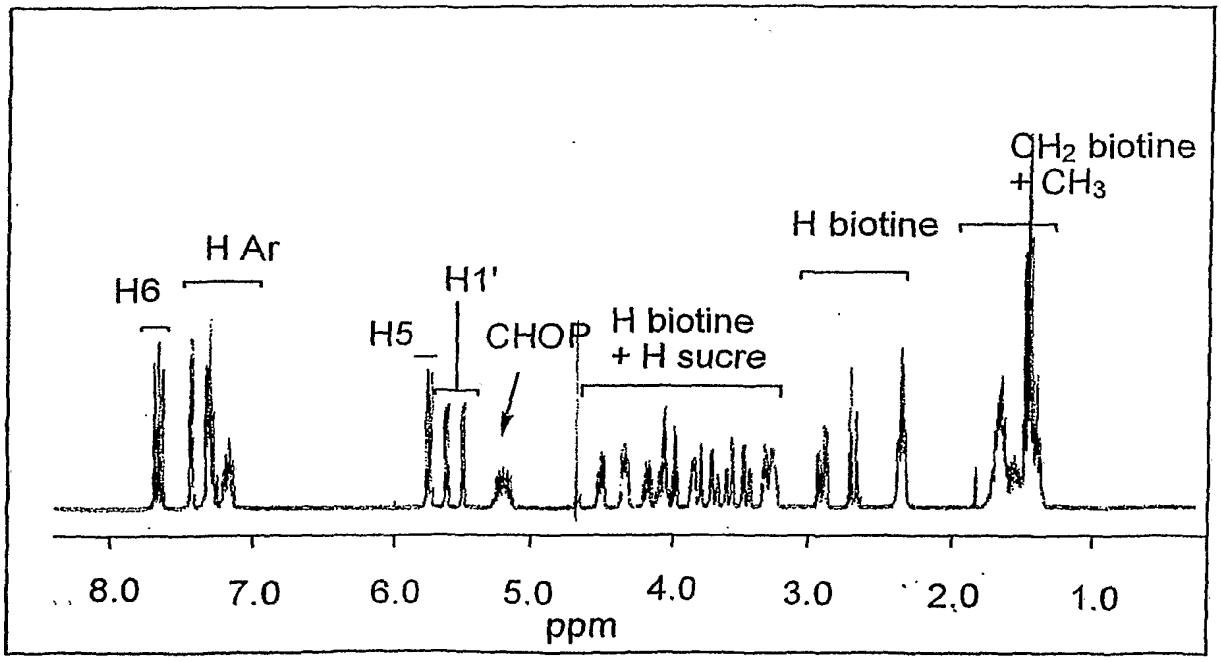

H6
H Ar
H5
H1'
CHOP
H biotine
+ H sucre
H biotine
CH2 biotine
+ CH3
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0
ppm

Fig. 6

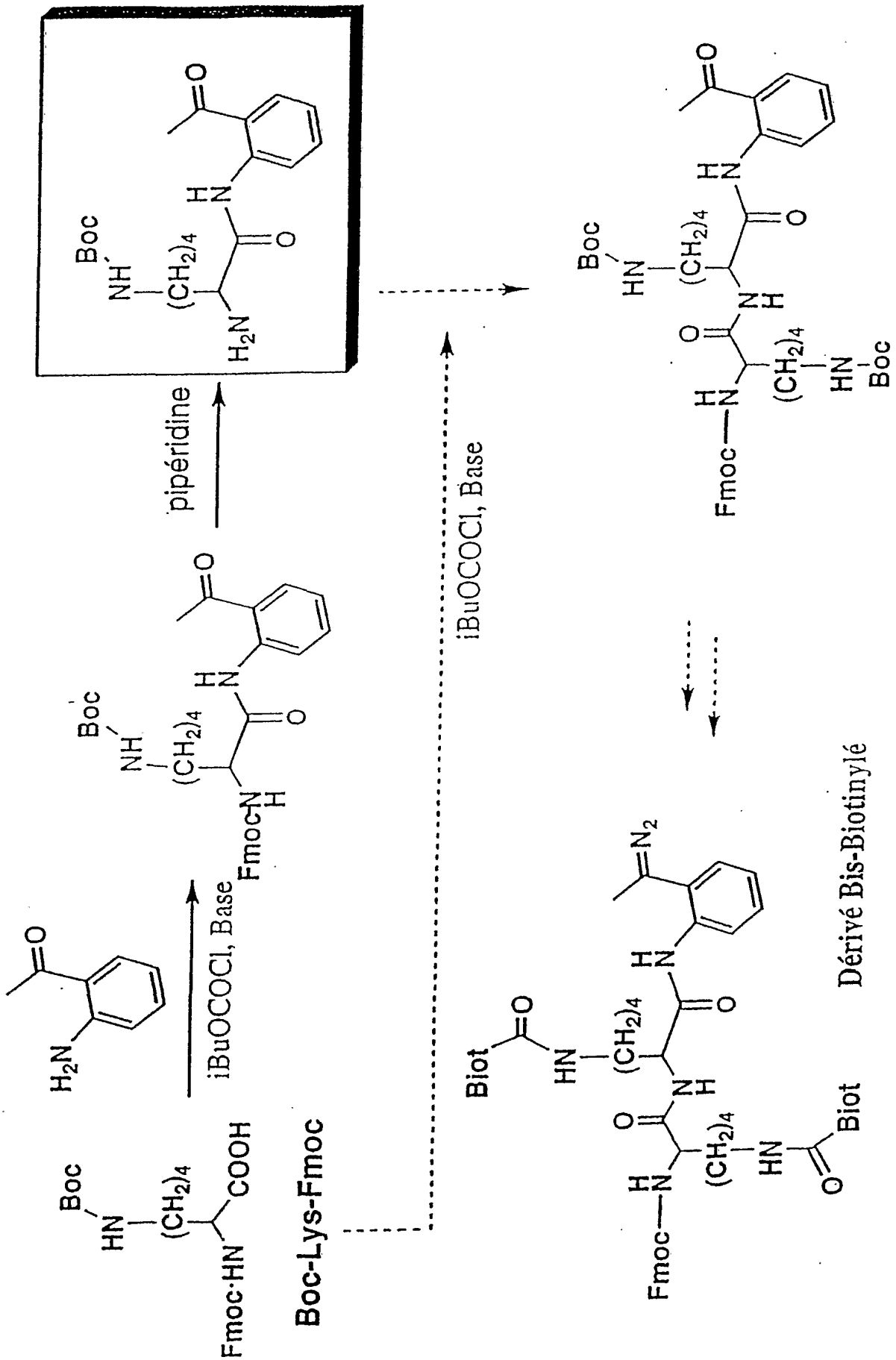

pipéridine
iBuOCOCl, Base
iBuOCOCl, Base
Boc-Lys-Fmoc
Dérivé Bis-Biotinylé

Fig 7

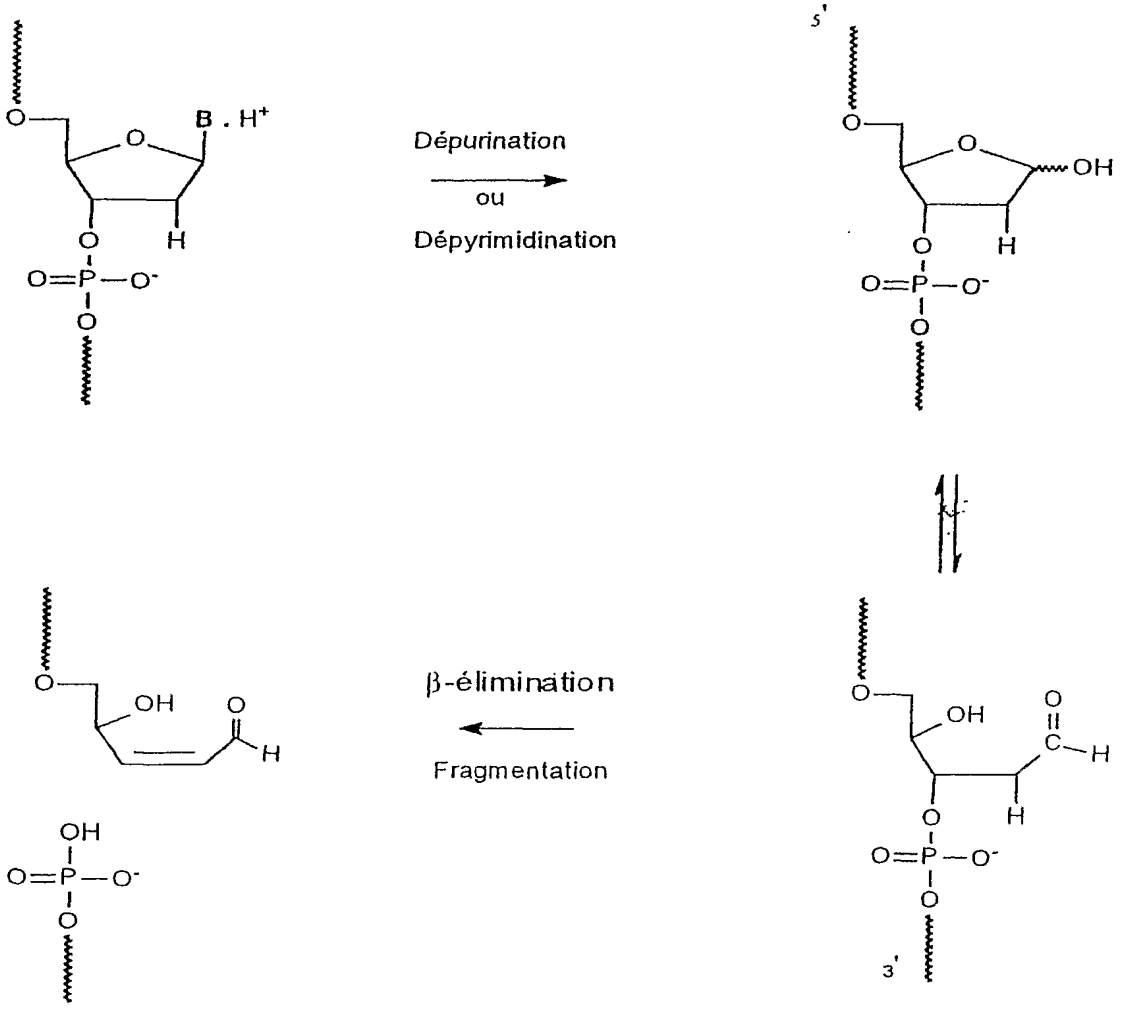

Dépurination
ou
Dépyrimidination
β-élimination
Fragmentation
5'
3'

Fig 8

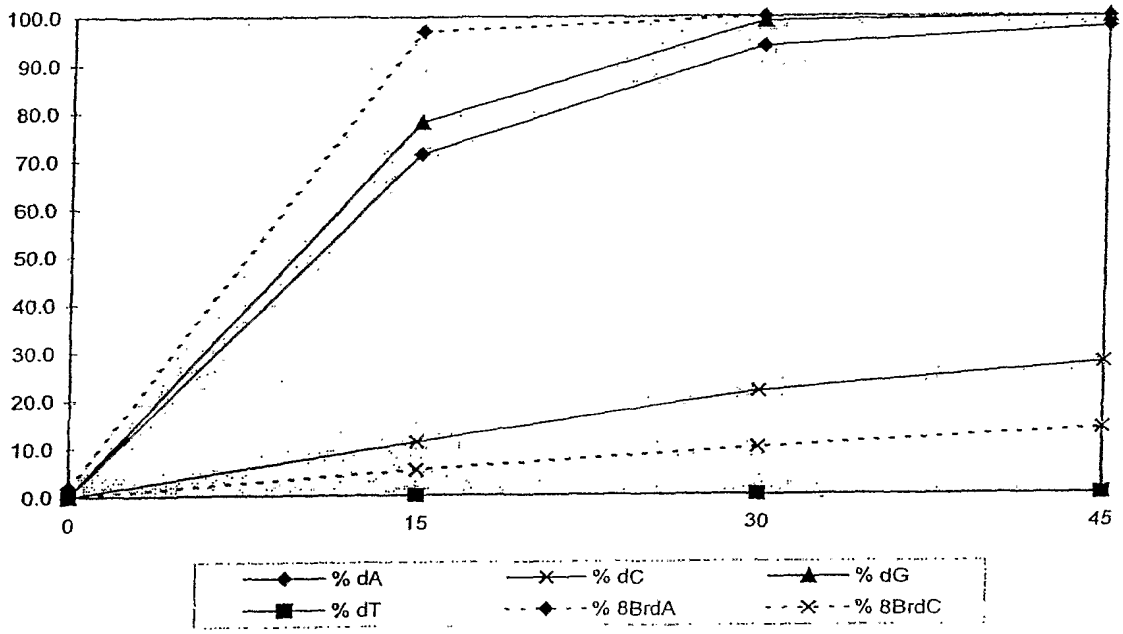

100.0
90.0
80.0
70.0
60.0
50.0
40.0
30.0
20.0
10.0
0.0
0
15
30
45
% dA
% dC
% dG
% dT
% 8BrdA
% 8BrdC

Fig 9

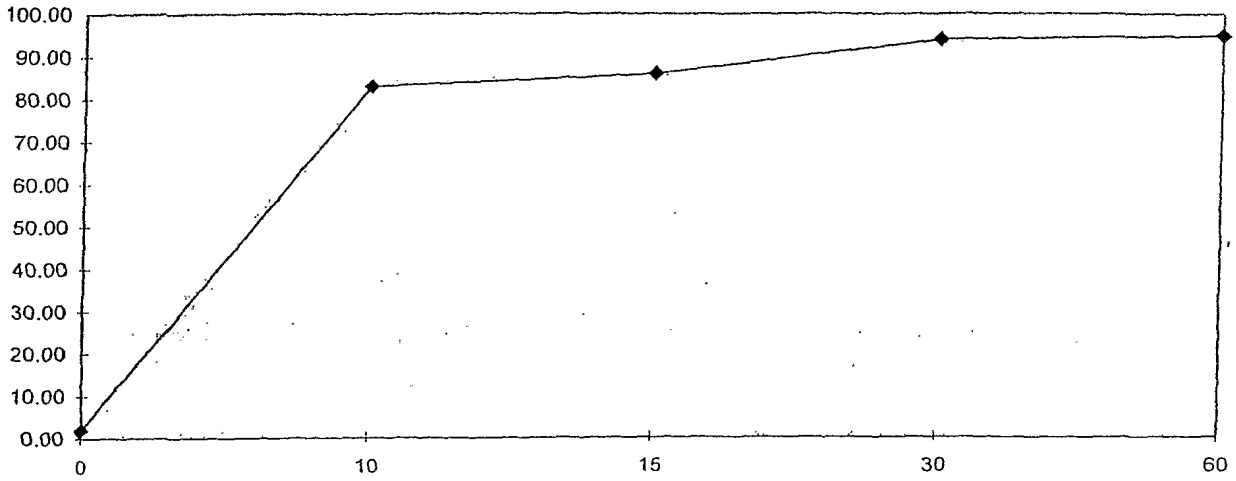
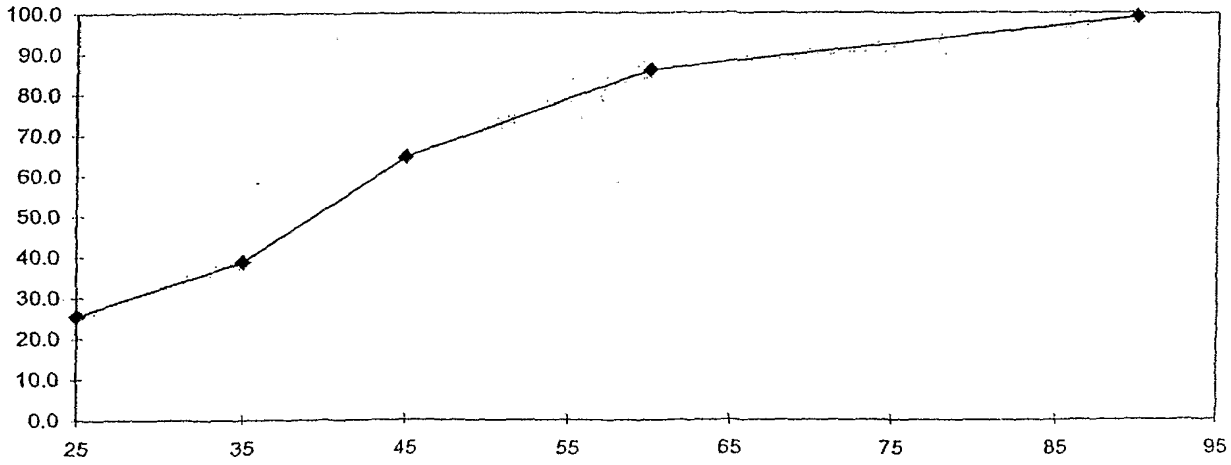

100.00
90.00
80.00
70.00
60.00
50.00
40.00
30.00
20.00
10.00
0.00
0
10
15
30
60

Fig 10


100.0
90.0
80.0
70.0
60.0
50.0
40.0
30.0
20.0
10.0
0.0
25
35
45
55
65
75
85
95

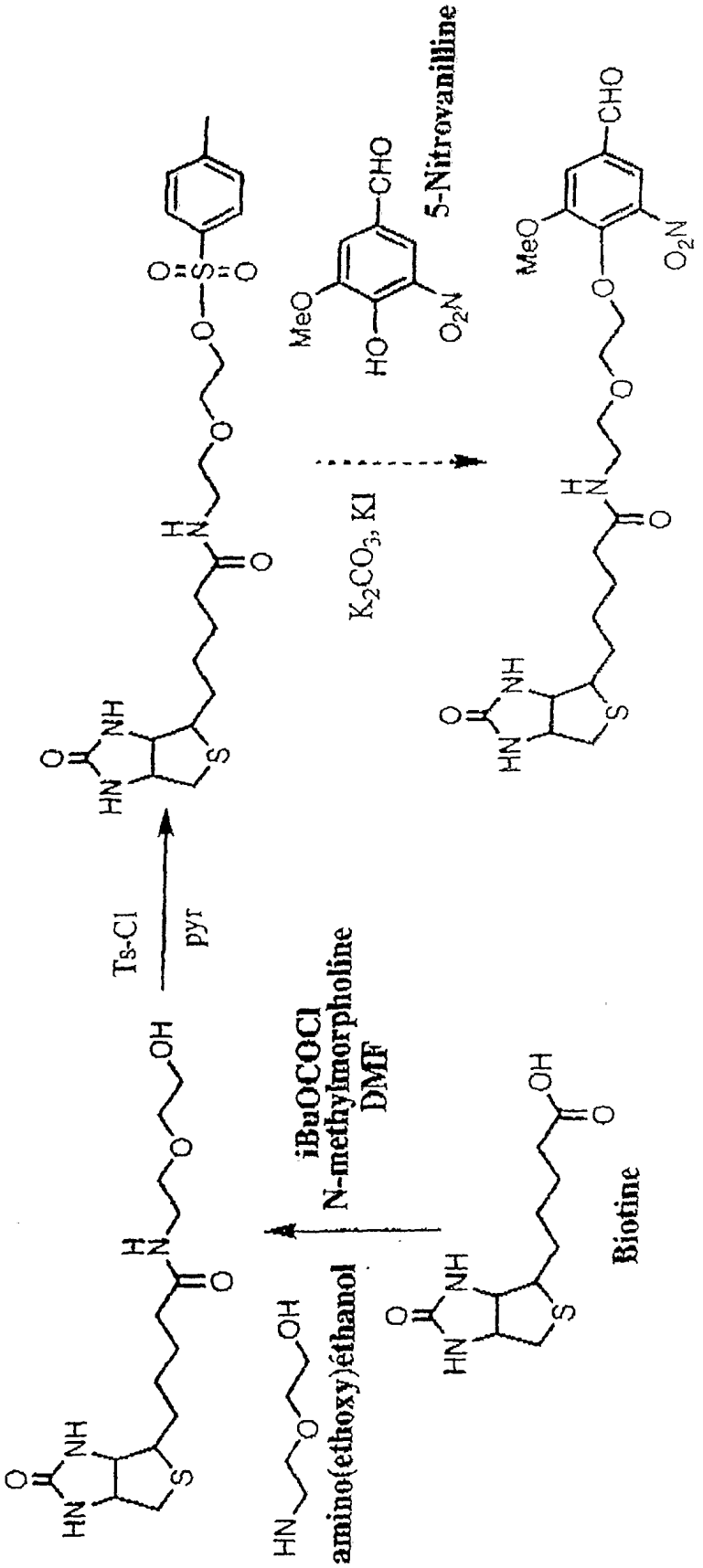

Ts-Cl
pyr
iBuOCOCl
N-methylmorpholine
DMF
amino(ethoxy)éthanol
Biotine
K_2CO_3, KI
5-Nitrovanilline

Fig 11

Espacenet

Bibliographic data: WO0039345 (A1) — 2000-07-06

MUTATION SCANNING ARRAY, AND METHODS OF USE THEREOF

Inventor(s): MAKRIGIORGOS G MIKE [US] ± (MAKRIGIORGOS, G., MIKE)

Applicant(s): DANA FARBER CANCER INST INC [US]; MAKRIGIORGOS G MIKE [US] ± (DANA-FARBER CANCER INSTITUTE, INC, ; MAKRIGIORGOS, G., MIKE)

Classification: - international: ***C12N15/09; C12Q1/68; G01N33/53; G01N33/566; G01N37/00;*** (IPC1-7): C12P19/34; C12Q1/68
- cooperative: **C12Q1/6827**; **C12Q1/6837**; **C12Q1/6858** more

Application number: WO1999US31177 19991229 Global Dossier

Priority number (s): US19980114196P 19981230

Also published as: AU2397400 (A) CA2357601 (A1) EP1141414 (A1) EP1141414 (A4) JP2002538415 (A) more

Abstract of WO0039345 (A1)

The present method is directed to using a mutation scanning array to identify mismatches or polymorphisms in multiple genes or the same gene in multiple individuals. The array can be a chip or a microsphere. Preferably, the array has elements containing immobilized oligonucleotides that collectively span at least 10 different whole genes.

IDT Exhibit 1020.7877

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 : C12Q 1/68, C12P 19/34	A1	(11) International Publication Number: WO 00/39345
		(43) International Publication Date: 6 July 2000 (06.07.00)

(21) International Application Number: PCT/US99/31177

(22) International Filing Date: 29 December 1999 (29.12.99)

(30) Priority Data:
60/114,196 30 December 1998 (30.12.98) US

(71) Applicant *(for all designated States except US)*: DANA–FARBER CANCER INSTITUTE, INC. [US/US]; 44 Binney Street, Boston, MA 02115 (US).

(72) Inventor; and
(75) Inventor/Applicant *(for US only)*: MAKRIGIORGOS, G., Mike [GR/US]; 36 Goodrich Road, Jamaica Plain, MA 02130 (US).

(74) Agents: EISENSTEIN, Ronald, I. et al.; Nixon Peabody LLP, 101 Federal Street, Boston, MA 02110 (US).

(81) Designated States: AU, CA, JP, US, European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.

(54) Title: MUTATION SCANNING ARRAY, AND METHODS OF USE THEREOF

(57) Abstract

The present method is directed to using a mutation scanning array to identify mismatches or polymorphisms in multiple genes or the same gene in multiple individuals. The array can be a chip or a microsphere. Preferably, the array has elements containing immobilized oligonucleotides that collectively span at least 10 different whole genes.

IDT Exhibit 1020.7878

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

IDT Exhibit 1020.7879

MUTATION SCANNING ARRAY, AND METHODS OF USE THEREOF

STATEMENT AS TO FEDERALLY SPONSORED RESEARCH

This invention was made with Government support under grants R29 CA63334, K04 CA69296, and RO1 CA72046, awarded by the USPHS. The Government may have certain rights in the invention.

FIELD OF THE INVENTION

The present invention is directed to mutation scanning arrays, and methods of use thereof. In a preferred embodiment the method can be used to rapidly identify mutations and/or polymorphisms on a nucleic acid segment, or in an arbitrary mixture of nucleic acid segments or genes.

BACKGROUND OF THE INVENTION

The detection of mutations has been an area of great interest in recent years. For example, mutations in certain genes have been associated with a variety of disorders - ranging from blood disorders to cancers. Genetic tests are thus becoming an increasingly important facet of medical care. Consequently, there has been an emphasis on the ability to rapidly and efficiently detect mutations and polymorphisms.

Many electrophoretic techniques have been developed to rapidly screen DNAs for sequence differences by which such mutations can be detected. Denaturing Gradient Get Electrophoresis (DGGE) [Myers, R.M., Maniaris, T. and Lerman, L., *Methods in Enzymology*, 155, 501-527 (1987)], Constant Denaturant Gel Electrophoresis (CDGE) [Borresen, A.L., et al., Proc. Nat. Acad. Sci. USA, 88, 8405 (1991)], Single Strand Conformation Polymorphism (SSCP) [Orrita, M., et al., *Proc. Nat. Acad. Sci. USA*, 86, 2766-2770 (1989)], Heteroduplex Analysis (HA) [Nagamine, C.M., et al., *Am. J. Hum, Genet.*, 45,377-399 (19?9)] and Protein Truncation Test (PTT) [Roest, P.A.M., et al., Hum. Molec. Genet., 2,1719-1721 (1993)] are frequently used methods. Many labs use combinations of these methods to maximize mutation detection efficiency. All these methods require gel electrophoresis. Methods that do not require gel electrophoresis also exist. For example, selective hybridization on immobilized target sequences allows screening for rare known mutations [Zafiropoulos, A., et al., *Biotechniques* 223, 1104-1109

(1997)], while mass-spectrometry has been used to detect mutations by analyzing molecular weight of proteins [Lewis, J.K., et al., *Biotechniques* 24, 102-110 (1998)].

A fundamental problem with currently existing mutation and polymorphism detection methods is that they only screen for mutations on a single gene at a time (i.e. the method is directed to looking at a 'gene of interest', that is suspected of having a mutation). Given that the human genome has 50,000 - 100,000 genes, this is a severe limitation. It is likely that unknown mutations and polymorphisms in several other genes both known and unknown, exist simultaneously with mutations/polymorphisms in the 'gene of interest'. However, mutations in those other genes would likely not be identified. Therefore a method that can perform 'mutation/polymorphism scanning' in for a wide array of genes simultaneously, without the initial need for identifying the gene one is screening would be useful. Gel-electrophoresis - based methods are essentially restricted to examining mutations in a single gene at a time. Attempts have been made to devise non-gel electrophoretic methods to identify mutations, that would not be restricted to a single gene [Cotton et al., *Proc. Natl. Acad. Sci. USA* vol. 85, pp 4397-4401, (1988)] [Nelson, S.F. et al., *Nature Genetics*, 4, 11-8, (1993 May)] [Modrich, P., et al., Methods for Mapping Genetic Mutations. US PATENT 5459039, (1995)]. These methods, however, have had limited success (Nollau P. and Wagener C., *Clinical Chemistry* 43:1114-1128 (1997)) since they are complicated, typically requiring several enzymatic steps and they result in a large number of false positives, i.e. they frequently score mutations and polymorphisms in normal DNA. It would be desirable to have a method that allows highly sensitive and specific identification and rapid purification of sites that contain mutation/polymorphism over large spans of the genome.

Although DNA arrays and methodologies that can simultaneously scan a large set of DNA fragments for gene expression (e.g. the 'repertoire' and amount of genes expressed in normal vs. cancer cells) are known [Wodicka L, *Nature Biotechnology* 15: 1359-1367 (1997); Lockhart, DJ, *Nature Biotechnology* 14: 1675-1680 (1996); Schena, M., *Trends Biotecnnol* 16: 301-306, (1998); Yang, T.T., *Biotechniques* 18: 498-503, (1995)], the ability to scan a large set of random DNA fragments for unknown mutations is a much more demanding process on which the technology is lagging [Ginot F., *Human Mutation* 10: 1-10 (1997)]. Thus far DNA array - based methods to scan for polymorphisms (SNPs) and mutations has been restricted to specific genes [Lipshutz, R.J., *Biotechniques* 19: 442-447 (1995); Wang, D.G., *Science* 280: 1077-1082 (1998)]. Whereas detection of unknown

mutations over several genes requires a selectivity and sensitivity not currently achievable by present arrays [Ginot F., *Human Mutation* 10: 1-10 (1997)]. For example, when it comes to unknown mutation detection, even a single gene with a coding sequence of the size of APC (8.5 kb) is difficult to screen in a single experiment, especially when an excess normal alleles is simultaneously present [Sidransky D., *Science* 278: 1054-1058 (1997)]. Thus, current arrays do not scan whole genes from the 5' to the 3' end, but selectively sample the gene. For example, expression arrays are biased to the 3' end. It would be desirable to have a methodology that could sample genes over larger portions of a gene and multiple genes. A method that permits identification of mismatches over large spans of the genome would also be desirable.

SUMMARY OF INVENTION

We have found a method that permits one to overcome resolutions and other limitations existing in current DNA chip technology and utilize the existing DNA chip technology for mutations scanning hundreds or thousands of genes simultaneously and identifying short (50-300 bases) DNA segments that contain mutations or polymorphisms. This method comprises first identifying a DNA segment containing mismatches. Those mismatches can either be single nucleotide polymorphisms (SNP) or mutations. Thereafter, one selects a DNA segment of from about 50-300 nucleotides containing a mismatch. Those DNA segments can be amplified by PCR and then screened on the DNA chip. Accordingly, instead of selecting a single gene at a time and examining whether it contains mutations, the present methodology first scans DNA to identify and isolate mismatch-containing and thereby mutation-containing DNA fragments (genotypic selection), and then determines which genes these DNA fragments belong to, by using available DNA arrays. Thus, the search for mutations is transformed to the easier task of searching for genes on a DNA array to identify where the mismatch occurs. Accordingly, DNA arrays currently used for multiplexed gene expression scanning [Wodicka L, *Nature Biotechnology* 15: 1359-1367 (1997); Lockhart, DJ, *Nature Biotechnology* 14: 1675-1680 (1996).; Schena, M., *Trends Biotecnnol* 16: 301-306, (1998); Yang, T.T., *Biotechniques* 18: 498-503, (1995)] can be used directly or with minor modifications known to the artisan based upon this disclosure to scan for mutation.

Any method of identifying a DNA segment that contains a polymorphism and, thus, mutations and/or polymorphisms can be used. For example, one

preferred method of identifying DNA fragments containing such mismatches involved the following steps; (a) isolating the nucleic acid, e.g., DNA, to be screened for mutations (referred to as the target DNA), adding PCR primers, and hybridizing it with control DNA, to create mismatches. These mismatches occur at the exact positions of mutations or polymorphisms; (b) removing any pre-existing, spontaneous aldehydes by, for example, treating the DNA with hydroxylamine; (c) using repair glycosylase enzymes (Mut Y and TDG) to convert the mismatches to reactive sites, namely, aldehyde-containing abasic sites (these enzymes recognize mismatches and will 'cut' the nucleic acid base, e.g., adenine at that site to create a reactive site); (d) using compounds (e.g. ligands) with functional groups that at one site can covalently bind to the reactive sites on the DNA, and that at a second site contain unique moieties that can be detected; (e) binding antibodies or avidin to the detectable second sites of the DNA-bound ligands. These antibodies or avidin may carry chemiluminescent or other indicators, so that the total reactive sites on the nucleic acid, e.g., DNA segment(s) tested is quantified, e.g. by chemiluminescence; (f) purifying the segments where a reactive site is present (e.g. by immunoprecipitation, or by ELISA-microplate-based techniques, or by microsphere-based techniques). The rest of the nucleic acid, e.g., DNA that does not contain mismatches can then be discarded; and (g) amplifying the remaining, mismatch containing nucleic acid, e.g., DNA, by PCR using the primers added in the first step. That mismatch-containing DNA can then be used in the above-described method with our mutation scanning arrays. (These arrays are sometimes also referred to as DNA chips.) Using the arrays of the present invention one can determine which gene each identified mismatch belongs to. Thereafter, known techniques can be used to determine whether that mismatch is a mutation that either causes the disorder or is associated with the disorder or simply an allelic variation, i.e. a polymorphism.

More specifically, the present invention permits biochemical approaches for chemically modifying mutations in a target nucleic acid sequence. The mutations are converted to mismatches following hybridization with control nucleic acid sequence. The mismatches in the hybrid nucleic acid, e.g. DNA can then be covalently bound by a ligand molecule, and then identified by a detectable moiety. Subsequently the mismatch-containing DNA can be purified by known means such as immunoprecipitation and the mutation-containing genes detected.

The target nucleic acid can be cDNA or genomic DNA. For example, the DNA can be any mixture containing one or various sizes of DNA, such as cDNA

synthesized from the whole mRNA collected from cells that need to be screened for mutation/polymorphism; or fractions thereof; or the whole genomic DNA collected from cells that need to be screened for mutation/polymorphism; or fractions thereof; or any combination of the above digested into smaller pieces by enzymes. The use of cDNA is preferable.

This method can also be used to detect a variety of other labeled DNA sites such as DNA lesions that are converted to reactive sites by glycosylase enzymes or by chemical means (e.g. clustered DNA-damaged sites; abasic sites; carcinogen-DNA adducts; damaged DNA bases). In these embodiments, mixing of the target DNA with wild-type DNA to create mismatches is not needed. Enzymes will recognize damage and will generate reactive sites directly in the target DNA. Such enzymes include all known glycosylases, such as endonuclease III, uracil glycosylase, T4 endonuclease V, 3-methyladenine DNA glycosylase, 3- or 7-methylguanine DNA glycosylase, hydroxymethyluracile DNA glycosylase, FaPy-DNA glycosylase, M. Luteus UV-DNA glycosylase. Also, chemical agents such as bleomycin, alkylation agents or simple acid hydrolysis can generate reactive sites automatically in target DNA without any enzyme. The crucial step however is again the same, i.e. addition of some compound to the reactive site of a DNA lesion, which allows identification of the DNA-containing the lesion.

BRIEF DESCRIPTION OF THE FIGURES

Figure 1 shows a schematic of how the present technology is applied for identification of mutations in a complex mixture of genes, e.g. screening for C to A transversions over hundreds or thousands of genes simultaneously.

Figure 2 shows the sensitivity of chemiluminescence detection of alkaline phosphatase with a cooled ICCD camera. The inset shows time-dependent buildup of chemiluminescence following addition of chemiluminescent substrate plus enhancer.

Figure 3 shows chemiluminescence detection of aldehyde-containing apurinic/apyrimidinic (AP) sites generated in plasmid DNA following depurination in sodium citrate, pH 3.5 at 38° for up to 60 seconds. The inset depicts fluorescence detection when extensive depurination under identical conditions is applied. Data in the inset (from us (Makrigiorgos GM, Chakrabarti S and Mahmood S. Int J Radiat Biol, 74: 99-109, 1998) were used to convert chemiluminescence units to AP sites (right axis, see text).

Figures 4A and 4B shows sensitive detection of AP sites using FARP. Figure 4A shows detection of AP sites in genomic calf thymus DNA depurinated for 15 seconds, without treatment (bar1) or following treatment (bar 2) with hyrdoxylamine. Figure 4B shows detection of spontaneously generated AP sites in hydroxylamine - treated genomic calf thymus DNA at pH=7.0, at a temperature of 37°C (curve1) or 4°C (curve 2).

Figures 5A-5C shows gel electrophoresis of Mut Y-treated DNA, examined on denaturing gels. Figure 5A shows 49-mer double stranded oligonucleotides that are Mut Y-treated and visualized on polyacrylamide gels following SYBR GOLD staining. Lane 1, no mismatch, no Mut Y. Lane 2, no mismatch, plus Mut Y. Lane 3, A/G mismatch, no Mut Y. Lane 4, A/G mismatch, plus Mut Y. Figure 5B shows double standard homoduplex mixtures (DNA ladder, 27-500 base pairs) are Mut Y-treated and visualized on polyacrylamide gels following SYBR GOLD staining. Lane 1, no Mut Y. Lane 2, plus Mut Y. Figure 5C shows single stranded M13 DNA (7,249 bases) are enzymatically - treated and visualized on agarose gels following ethidium staining. Lane 1, M13 DNA, no Mut Y. Lane 2, M13 DNA, plus Mut Y. Lanes 3-6, molecular weight markers.

Figure 6 shows FARP-based chemiluminescence detection of Mut Y-treated DNA of a single length: 49-mer oligonucleotides are enzymatically - treated, FARP-labeled and captured on microplates. Bar 1, A/G mismatch, no Mut Y. Bar 2, A/G mismatch, plus Mut Y. Bar 3, No mismatch, no Mut Y. Bar 4, no mismatch, plus Mut Y.

Figure 7 shows FARP-based chemiluminescence detection of Mut Y-treated DNA fragments of varying length: Single stranded M13 DNA (7249 bases) and double stranded homoduplex mixtures (DNA ladder, 27-500 base pairs) are enzymatically - treated, FARP-labeled and captured on microplates. Bar 1, M13 DNA, no Mut Y. Bar 2, M13 DNA, plus Mut Y. Bar 3, ladder DNA, no Mut Y. Bar 4, ladder DNA, plus Mut Y.

Figures 8A and 8B show BARP-based chemiluminescence detection of Mut Y-treated DNA fragments of varying length: Figure 8A shows chemiluminescence from single stranded M13 DNA (that forms ~3 mismatches over 7249 bases) and double stranded homoduplex M13 DNA (no mismatches) enzymatically – treated by Mut Y, BARP-labeled and captured on microplates. Bar 1, s.s. M13 DNA, no Mut Y. Bar 2, s.s. M13 DNA, plus Mut Y. Bar 3, d.s. M13 DNA, no Mut Y. Bar 4, d.s. M13 DNA, plus Mut Y. Figure 8B shows gel electrophoresis of the same DNA, and demonstrates that, in agreement with the chemiluminescence results in Figure 8A,

only single stranded M13 plus Mut Y demonstrate DNA digestion (see bands in Lane 2).

Figures 9A and 9B show detection of a mutation. Figure 9A shows chemiluminescence detection of a single mutation (A-to-C transversion) engineered in a p53 gene which is incorporated in a 7091 base pair plasmid. Plasmids containing the mutation were first digested into smaller DNA fragments (400-2,00 bp) by exposure to RSAI enzyme. These were then melted and hybridized with normal plasmids to form mismatches at the position of the mutation. The DNA was then enzymatically – treated with Mut Y to convert mismatches to aldehydes, BARP-labeled and captured on microplates. Bar 1, plasmid with mismatch, no Mut Y. Bar 2, plasmid with mismatch plus Mut Y. Bar 3, normal plasmid, no Mut Y. Bar 4, normal plasmid, plus Mut Y. Figure 9B shows the variation of the chemiluminescence signal obtained when different amounts of mismatch-containing plasmid treated by Mut Y and BARP are applied on microplates.

Figures 10A and 10B compare DNA binding by different compounds. Figure 10A demonstrates the binding of the compound, AED, (2-(aminoacetylamino) ethylenediamine) to reactive sites generated at position of mismatches in DNA by the enzyme Mut Y. The figure shows samples of M13 DNA containing mismatches, treated with enzyme and various compounds, stained with ethidium bromide and examined via gel electrophoresis. A sample of M13 DNA without enzymatic treatment shows a single bright band in lane A. The sample of plasmid DNA treated with the enzyme Mut Y shows multiple bands, demonstrating the expected recognition and cutting of mismatched bases by Mut Y in lane B. Lane C In Lanes C, D and E, the Mut Y treatments are carried out in the presence of 5 mM methoxyamine (C) or in presence of the novel compound AED (D, 5 mM and E, 10 mM AED respectively). The disappearance of the bands in lanes C, D and E is an indication of covalent high labeling of DNA by methoxyamine or by AED, at the positions of reactive sites generated by Mut Y. In Lane F, the treatment of DNA was as in Lanes E and D, but another aldehyde reactive compound (BARP) was used instead of AED. Lane F still shows the same multiple bands as those generated in the absence of compound (see Lane B), indicating an inefficient labeling of aldehyde sites by BARP.

Figure 10B demonstrates the superior DNA binding of AED over BARP or FARP when reactive sites are generated at position of mismatches in DNA by the enzyme TDG. Lanes 1 and 2, G/T mismatch-containing oligonucleotide, no enzyme. Lane 3, G/T oligonucleotide with TDG enzyme. Lane 4, G/T

oligonucleotide with TDG enzyme in the presence of 5 mM methoxyamine. Lane 5, G/T oligonucleotide with TDG enzyme in the presence of 5 mM BARP. Lane 6, G/T oligonucleotide with TDG enzyme in the presence of 5 mM AED. Lane 7, G/T oligonucleotide with TDG enzyme in the presence of 0.5 mM FARP.

Figure 11. AED-based chemiluminescence detection of mismatches obtained when mismatch-containing s.s. M13 DNA is Mut Y treated in the presence of 5 mM AED. Bar 1, M13 DNA without Mut Y enzyme. Bar 2, M13 DNA with Mut Y enzyme.

Figure 12 is a schematic showing how the method of mismatch identification can be used with a DNA chip to detect inherited and acquired predisposition to cancer (see text).

Figure 13 is a procedure to detect polymorphisms and mutations.

Figure 14 demonstrates the results of four different hybridization experiments.

Figure 15 shows how a microbead (miscosphere)-based mutation scanning array is constructed.

DETAILED DESCRIPTION OF THE INVENTION

The present method permits biochemical approaches for chemically identifying the mismatch sites in, for example, the target DNA sequence. The target DNA can be identified by a detectable moiety and subsequently detected and purified by immunoprecipitation, microplates or microsphere technologies. Subsequently, the purified mutation - containing DNA fragments can be used in single-step screening of these mismatches by a wide variety of hybridization techniques (DNA chips, large-scale hybridization arrays, etc.)

For example, in trying to detect unknown mutations it has thus far proven difficult to screen for a single gene of about 8.5 kb (such as APC) in a single experiment, especially when an excess of normal alleles is simultaneously present [Sidransky, D. *Science* 278: 1054-1058 (1997)]. By contrast, the present method can screen multiple genes at once and/or multiple individual, and selects and isolates only those fragments containing a mutation/polymorphism. These mismatch-containing segments can be amplified by PCR and used, for example, in a DNA array to simply search for the matching gene in the array to identify which genes these mutation-containing fragments belong to. Consequently, existing arrays for multiplexed gene expression scanning such as known in the art can be used. For example, Affymetrix Hu6800 DNA Chip, or known arrays [Wodicka, L. et

al, *Nature Biotechnology:* 15: 1359-1367 (1997); Lockart, D.J. et al, *Nature Biotechnology* 14: 1675-1680 (1996); Schena, M. *Trends Biotechical* 16:301-306 (1998); Yang, T.T. et al. *Biotechniques* 18: 498-503 (1995); Ginot F. *Human Mutation* 10: 1-10 (1997)].

In order to increase resolution (i.e. definition of the gene segment containing the mutation/polymorphism) the fragment should be smaller. However, in order to effectively prepare large amounts of mismatch-containing fragments by standard techniques such as PCR, the fragments should be at least about 50 bases. In some instances for ease of operation, a loss in resolution can be tolerated and larger fragments used.

Preferably the mismatch-containing fragment is 50 - 300 bases, more preferably 50 - 200 bases, still more preferably 50 - 100 bases and most preferably about 50 bases.

The nucleotides on the array (gene elements) should be between 8-300 bases preferably no larger than the size of the DNA of the mismatch-containing fragments. For improved resolution, smaller sizes should be used. For example, 50 bases or less, more preferably 8-25 bases. Many arrays presently available use nucleotide fragments of about 25 bases. Typically, these nucleotide segments are selected to be close to the 3' portion of the transcript.

However, other DNA arrays as discussed, infra, can also be used. Such arrays, which contain fragments that span the whole length of the gene (i.e. from both the 5' end of the gene as well as the 3' end) are preferred.

These mutation scanning arrays can be used with any method that permits one to tag a DNA or mRNA mismatch, thereby permitting the rapid identification of mutations in nucleic acids, e.g. DNA, mRNA or DNA segment(s). One preferred method of identifying DNA-containing mismatches is disclosed in copending application U.S.S.N. (09/224,227). That method comprises (a) isolating the nucleic acid, e.g., DNA, to be screened for mutations (referred to as the target DNA), and hybridizing it with control DNA, to create mismatches. These mismatches occur at the exact positions of mutations or polymorphisms; (b) removing any pre-existing, spontaneous aldehydes by, for example, treating the DNA with hydroxylamine; (c) using repair glycosylase enzymes to convert the mismatches to reactive sites, namely, aldehyde-containing abasic sites (these enzymes recognize mismatches and will 'cut' the nucleic acid base, e.g., adenine at that site to create a reactive site); (d) using compounds (e.g. ligands) with functional groups that at one site can covalently bind to the reactive sites on the DNA, and that at a second site

contain unique moieties that can be detected; (e) binding antibodies or avidin to the detectable second sites of the DNA-bound ligands. These antibodies or avidin may carry chemiluminescent or other indicators, so that the total reactive sites on the nucleic acid, e.g., DNA segment(s) tested is quantified, e.g. by chemiluminescence; (f) purifying the segments where a reactive site is present (e.g. by immunoprecipitation, or by ELISA-microplate-based techniques, or by microsphere-based techniques). The rest of the nucleic acid, e.g., DNA that does not contain mutations can then be discarded; and (g) amplifying the remaining, mutation-containing nucleic acid, e.g., DNA, by PCR. That DNA can then be used with our DNA arrays in order to find which gene each identified mismatch belongs to. Thereafter, by known techniques determining whether that mismatch is a mutation that either causes the disorder or is associated with the disorder or simply an allelic variation, i.e. a polymorphism.

The present method will recognize mismatches formed upon hybridization of the target DNA and the control (wild-type) DNA. Those skilled in the art are aware that mismatches may appear as a result of inherited or acquired genetic alterations. Also, that not every mismatch is the result of mutation but that some mismatches simply represent polymorphisms that occur naturally in populations. Both the inherited and the acquired genetic alterations in DNA will cause a mismatch.

Furthermore, those skilled in the art are aware that because all eukariotic cells contain two copies of each chromosome, one paternal and one maternal, differences between the two alleles of each gene may also cause mismatches. In this case one gene copy (e.g. the paternal) will act as a control DNA and the second gene copy (the maternal) will act as the target DNA, and the mismatches will form upon hybridization of maternal and paternal DNA (i.e. simply by self-hybridization of DNA present in cells). These inherited differences can represent either polymorphisms or mutations.

There are a number of ways known in the art to distinguish whether a particular mismatch is an inherited polymorphism or mutation, or an acquired mutation.

One method that can be used to identify acquired mutations is to have the control DNA come from the same individual. For example, when screening a malignant cell the control DNA can be obtained from the corresponding non-malignant cell. By screening first the non-malignant cell alone and then the malignant cell (or a mixture of malignant and non-malignant cells) a comparison of

detected mismatches in the two cases can be made. Differences that appear solely on the malignant cell and not on the normal cell comprise acquired mutations which may have lead to the malignancy.

When inherited (genetic) mutations/polymorphisms (i.e. where the alteration from the wild-type is present at birth and in every cell of the body) need to be identified, only normal cells need to be examined. As explained, inherited differences between the two alleles will cause mismatches upon self-hybridization. Detection of these mismatches will indicate the positions of inherited polymorphisms or mutations. Thereafter, one standard method to discriminate inherited polymorphisms from inherited mutations is to screen kindred and to determine whether or not the mismatch is present in normal kindred (i.e. a benign polymorphism) or only present in kindred showing a particular abnormality (i.e. a debilitating mutation).

The use of databases categorizing mutations and polymorphisms has also been increasingly popular. Thus, comparison of an identified genetic variation with those contained in a database can in many instances be used to determine whether the detected mismatch in DNA is due to a mutation or due to a polymorphism. One can also look at whether the mismatch causes truncation in the expressed protein.

Finally, another method that can be used to discriminate among mutations and polymorphisms is by the use of in-vivo assays. Thus, one can substitute a gene with at least one engineered base substitution mutation for the wild-type gene in an assay to determine whether or not the gene with the mutation can functionally replace a wild-type normal gene. If a gene can replace a wild-type normal gene in an assay and exhibit almost normal function that gene is *not* considered a mutation, but an allelic variation (i.e. polymorphism). If it cannot that gene will be considered a mutation.

One of the advantages of the present approach as opposed to mutation-detection methods presently being used is the ability to identify numerous mutations at diverse places in the genome. This permits one to determine if certain genes not presently associated with a particular abnormality may also have some relationship with that abnormality. For example, with hereditary non-polyposis colorectal cancer (HNPCC), mutations in the MSH2 and MLH1 genes are believed to be responsible for approximately 90% of the cases. A number of other genes have been identified as being responsible for the other 10% of the cases. However, in view of the cost of screening one typically looks primarily at MSH2 and MLH1. It

may turn out when an array of genes are looked at the same time, that mutations in other genes also play a major role, in an individual with a particular condition. These other mutations may be associated with severity of the condition. By monitoring these additional genes and looking at disease state and recovery, one can develop a better idea of prognosis and treatment regimes than is currently available.

When using genomic DNA the skilled artisan is aware that numerous mismatches can and will occur in non-coding genetic regions. Looking at non-coding regions can permit the identification of mutations that affect expression and levels of expression. On the other hand when one is interested in looking for mutations in the expressed proteins it is preferable to use the mRNA to generate cDNA, and then form mismatches that can be detected by the present approach.

The preferred target nucleic acid is DNA, but mRNA can also be used. The DNA can be any mixture containing one or various sizes of DNA, such as cDNA synthesized from the whole mRNA collected from cells that need to be screened for mutation/polymorphism; or fractions thereof; or the whole genomic DNA collected from cells that need to be screened for mutation/polymorphism; or fractions thereof; or any combination of the above digested into smaller pieces by enzymes.

Thereafter, the DNA, mRNA, whole or smaller pieces thereof can be screened to identify those pieces containing a mismatch. The mismatch-containing pieces are then isolated from the remaining nucleic acid segments. As a result of this selection for mismatch-containing segments, it does not matter whether normal alleles outnumber aberrant segments. Moreover, one can use techniques such as PCR to amplify the mismatch-containing segments.

The control will be a wild-type fraction similar to the target. This wild-type likely will have no mutations. The control nucleic acid can be selected depending upon the intent of the test. For example, where acquired mutations in cancer cells are being screened, the control nucleic acid can come from a "normal" cell from the same individual. In other instances, for example, where an inherited (genetic) component may be involved the control DNA would come from a different subject than providing the nucleic acid; or simply differences among the paternal and maternal alleles can be examined by a self-hybridization of the DNA of the examined individual.

Following DNA isolation, the DNA is fragmented to reduce its size to the desired 50-300 base pairs, and generic PCR primers are added to the nucleic acids, in order to amplify the preparation at a later stage.

Thereafter in one embodiment, the target DNA is mixed and hybridized with wild-type DNA to create mismatches at the positions of differences, which are expected to be mutations/polymorphisms. One labels the mismatch and isolates the mismatch-containing DNA from the remaining DNA. In this manner one removes potentially interfering background signal. Thus, the only DNA being used with the DNA array is DNA that contains at least one mismatch.

Any method of labeling the mismatches in bulk can be used. One preferred method involves using repair glycosylases and specific aldehyde compounds to identify mismatches revealed by the glycosylases.

For example, creating a mixture of wild-type and target DNA and creating a mixture of segments containing the wild-type: wild-type, target DNA: target DNA and target DNA: wild-type pairs. The mixture is preferably treated with a compound such as hydroxylamine to remove any spontaneous aldehydes. Thereafter, the mismatches that occurred, which would be in the wild-type: target DNA hybrids, are recognized and converted to reactive sites (aldehydes) by enzymes such as a glycosylase repair enzyme such as Mut Y and thymine DNA glycosylases (TDG) (e.g., from Hela cells or *E. coli*). A unique feature of these enzymes is that they are highly specific, i.e. they act only on mismatches while they leave non-mismatch containing DNA completely intact.

These reactive sites are identified by using a compound containing an aldehyde - binding moiety such as -O-NH_2 (-hydroxylamine), or -$NHNH_2$ (-hydrazine) or -NH_2 (-amine) and also having a second moiety that reacts with a detectable entity (e.g. fluorescein, biotin, digoxigenin, which respectively react with antifluorescein antibody, avidin, and antidigoxigenin antibody. The antibodies may have chemiluminescence tags on them and thereby are detected). A unique feature of the present approach is that the aldehyde - binding moiety binds covalently to the enzyme-generated reactive sites. Combined with the specificity of the mismatch - repair enzymes, the use of covalently bound ligands to the position of mutations results in a sensitivity and specificity which is unparalleled by other methods for detection of mutations and polymorphisms.

The compounds have the general formula:

X-Z-Y,

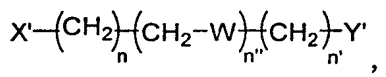
wherein X is a detectable moiety, preferably X is NH_2, SH, $NHNH_2$, a fluorescein derivative, a hydroxycoumarin derivative, a rhodamine derivative, a BODIPY derivative a digoxigenin derivative or a biotin derivative;

Y is $NHNH_2$, O-NH_2 or NH_2, preferably Y is NH_2,

Z is a hydrocarbon, alkylhydroxyl, alkylethoxy, alkylester, alkylether, alkylamide or alkylamine. Z may contain a cleavable group such as S-S. Z may be substituted or unsubstituted.

These reactive sites are identified by using a compound containing an aldehyde - binding moiety (Y) such as -O-NH_2 (-hydroxylamine), or -$NHNH_2$ (-hydrazine) or -NH_2 (-amine) and also having a second moiety (X) that reacts with a detectable entity (e.g. fluorescein, biotin, digoxigenin, which respectively react with antifluorescein antibody, avidin, and antidigoxigenin antibody. The antibodies may have chemiluminescence tags on them and thereby are detected). The aldehyde - binding moiety binds covalently to the enzyme-generated reactive sites. Combined with the specificity of the mismatch - repair enzymes, the use of covalently bound ligands to the position of mutations results in a high sensitivity and specificity.

One preferred embodiment of the invention has a general formula;



wherein X' is $NHNH_2$ or NH_2, preferably NH_2;

Y' is O-NH_2 or NH_2, preferably O-NH_2;

W is -NHC(O)-, -NHC(OH)-, -C(OH)-, -NH-, C-O-, -O-, -S-, -S-S-, -OC(O)-, or C(O)O-;

n is and integer from 0 to 12, preferably 4-7 and more preferably 6;

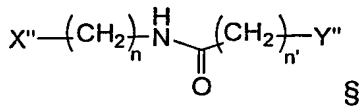
n' is an integer from 0 to 12, preferably 4-7, and more preferable 6, and

n'' is an integer from 1 to 4, preferably 1-2, and more preferably 1.

Preferably, the compound has a molecular weight between 100 - 500, more preferably 100 – 300, still more preferably 150 – 200.

Z and W can be substituted with groups that enhance the solubility of the resultant compound. Preferably the compounds of the formula $X\text{-}(CH2)_n\text{-}(CH2\text{-}W)_{n''}\text{-}(CH2)_{n'}\text{-}Y$ are overall soluble in the solvent used.

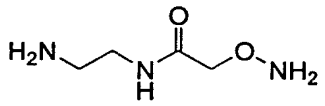
A preferred embodiment has the formula;



§

wherein X", Y", n, and n' are as described as above.

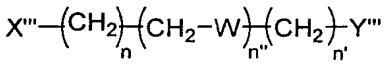
A more preferred compound is 2-(aminoacetylamino) ethylenediamine (AED), ($NH_2CH_2CH_2NHC(O)CH_2ONH_2$).



2-(aminoacetylamino) ethylenediamine (AED)

In another embodiment, the DNA reactive site recognized by enzymes such as glycosylases are identified by using compounds that contain a hydroxylamine reactive group. Examples of hydroxylamine compounds include FARP and FARPhc, both of which are fluorescent. FARP is a novel hydroxylamine containing derivative of fluorescein and FARPhc is a novel hydroxylamine containing derivative of hydroxy-coumarin.

These compounds have the general formula;

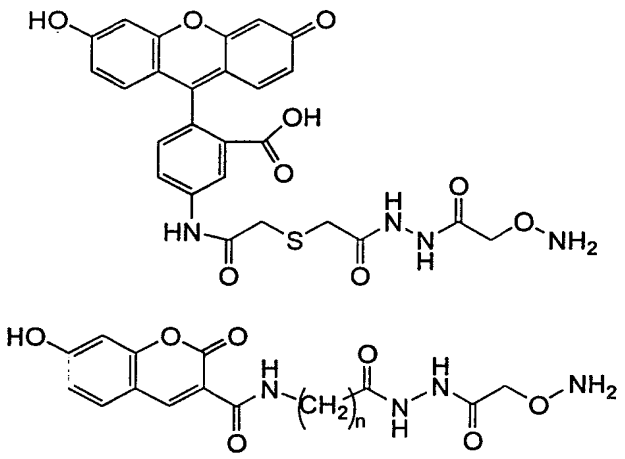


wherein Y'" is O-NH_2;

X'" is a fluorescent molecule, a fluorescein derivative or a hydroxy-coumarin derivative.

W, n, n', n" and n'" are defined as above.

More preferred compounds includes fluorescein aldehyde reactive probe, FARP, and fluorescent reactive probe hydroxycoumarin, FARPhc.

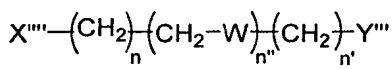


FARP $FARP_{hc}$

DNA samples containing mismatches that are prepared and treated with DNA-glycosylase enzymes as described above, will form covalent oxime bonds to FARP and FARPhc.

In an alternative embodiment, the DNA reactive sites recognized by enzymes such as glycosylases are identified by using compounds that contain a hydrazine reactive group. An example of this class of compounds includes biotin hydrazine. The present invention allows using hydrazine compounds to label reactive sites generated by the DNA-glycosylase enzymes. In yet still another alternative embodiment, the compound is a biotin aldehyde reactive probe, such as BARP, a biotinylated derivative of hydroxylamine [BARP, Kubo, K, et al., *Biochemistry* 31:3703-3708 (1992)].

These biotinylated hydroxylamine or hydrazine compounds have the general formula:



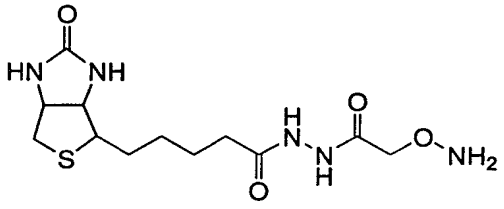
wherein Y''' is O-NH_2 or $NHNH_2$;

X'''' is a detectable molecule, biotin or biotin derivative.

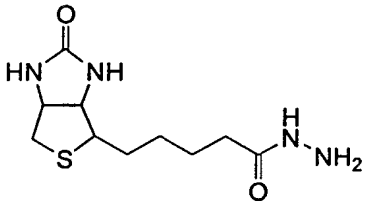
W, n, n' and n'' are defined as above.

For example, a Y moiety such as an amine should react with the aldehyde on for example the DNA, while the X group remains free for further modification and detection.

More preferred compounds includes biotin aldehyde reactive probe, BARP [Kubo, K, et al., *Biochemistry* 31:3703-3708 (1992)] and biotin hydrazide:



or



BARP

It was discovered that, following the recognition of mismatches by glycosylases such as Mut Y or TDG, and the resulting conversion to aldehyde-containing reactive sites, the enzyme has to be kept inactive, otherwise it interferes with the subsequent covalent binding of the ligand compounds. As a result, the conditions for reaction of hydroxylamines, hydrazines or amines with the enzymatically - generated aldehyde-containing reactive sites are at temperature of 4°C-15°C and at pH 6-7. (In the specific case of Y=NH_2 (amines), the presence of a reducing agent such as borohydride, 4°C-15°C for 1-3 hours is also required during binding to reactive sites). Following covalent attachment of the ligand compounds to reactive sites, the enzyme is then inactivated via heating at 70° C, for 10 minutes. Alternatively, to remove the enzyme a standard phenol-chloroform extract ion, or treatment with protein use K can be adopted.

When X=NH_2 (amine), in order for the covalently-bound ligand to be recognizable by an antibody, the free -NH_2 group is first covalently linked to an amine-binding compound with a recognizable group (e.g. a succinimidylester compound such as biotin-LC-succinimidyl ester; biotin-LC-SS-succinimidyl ester [Pierce]; fluorescein-succinimidyl ester; etc.). The reaction and purification conditions of such succinimidyl esters with -NH_2 containing compounds are well known.

When X=SH (sulfhydryl), in order for the covalently-bound ligand to be recognizable by an antibody, the free -SH group is first covalently linked to a sulfhydryl-binding compound with a recognizable group (e.g. a maleimide compound such as biotin-LC- maleimide; biotin-LC-SS- maleimide [Pierce];

fluorescein - aleimide; etc.). The reaction and purification conditions of such maleimides with -SH containing compounds are well known.

It was discovered that binding to reactive sites becomes much more efficient when small hydroxylamines (such as AED) are used. Therefore, the use of small compounds of the formula X'-(CH_2)n-(CH_2-W)n''-(CH_2)n'-Y , and of molecular weight less than 200 is preferred. These compounds are water soluble, can be incubated with DNA at a high molarity (e.g. 10 mM), and are able to diffuse fast enough to bind to reactive sites at a much higher level of efficiency than the other compounds (e.g. FARP, BARP) that have higher molecular weights and are less water soluble.

Purification of the mismatch-containing DNA preferably relies on the utilization of aldehydes as the recognition sites for mismatches resulting in covalent bonding of the marker molecule to these aldehydes. Therefore, the presence of contaminating nucleases that cleave DNA and create 3' hydroxyl groups - containing strand breaks (-a common problem in similar assays-) do not generate binding sites for the marker molecules. This method does not require the use of gel electrophoresis which compares DNA strand by their length or size. Thus, the generation of false positives from strand breaks generated by contaminating nucleases is thereby avoided. The method of the invention only detects labeled DNA following covalent binding of such aldehydes with ligand compounds and subsequent immobilization to a solid support, e.g., microplates. In addition, the length and diversity of DNA fragments are irrelevant to the assay, which is another advantage over gel-electrophoretic methods.

If borohydrite (sodium- or cyano- borohydrite) is present during the enzymatic lyase step, then a covalent cross-linking between enzyme and DNA occurs. In an alternative embodiment, one can introduce covalently-bound marker molecules (biotin, fluorescein, digoxigenin, or other fluorescent or chemiluminescent indicators) at the positions recognized by the enzymes.

To achieve this, the glycosylase enzymes are covalently labeled with an indicator molecule, either prior to DNA binding or after DNA binding and crosslinkage have taken place. For example using biotin as the indicator molecule, and Mut Y (trevigen) as the glycosylase enzyme the following procedure can be used.

To pre-label Mut Y with biotin, (i.e. Prior to DNA binding), the enzyme is incubated with a reactive analogue of biotin (e.g. Succinimidyl-ester of biotin, or biotin-lc-succinimidyl ester (Pierce), or biotin-maleimide, etc). Glycosylases are positively charged enzymes by virtue of the –NH_2-groups on their surface. These –

NH_2- groups can be utilized for covalent crosslinking of the biotin indicator. The reaction of Mut Y with biotin-lc-succinimidyl ester takes place at ph 7.5-8.0, 1h, 4°C at a molar ratio of biotin:Mut Y of 5:1, or 2:1, or 1:1 (higher excess of biotin must be avoided if Mut Y is to retain its biological activity). Following that the unreacted biotin is removed by standard chromatography or by G25 filtration. The biotinylated Mut Y can now by used for enzymatic activity on a test DNA, to recognize, crosslink and label mismatches.

The assay is carried out as follows: first, DNA to be tested for mutations (test DNA) is mixed and cross-hybridized to DNA that contains no mutations (control DNA), in order to generate mismatches at the positions of mutations. Second, biotinylated Mut Y is added at a Mut Y:DNA molar ratio of 100:1, in the presence of 100 mM sodium borohydrite. Other reaction conditions and reaction buffers are standard, as recommended by the Mut Y supplier. Following crosslinkage of biotinylated Mut Y to DNA mismatches (a/g or a/c mismatches), the DNA is contacted by a streptavidin-coated solid support (e.g. Streptavidin-coated magnetic microspheres) in order to isolate DNA molecules containing biotinylated sites (i.e. Mutations). The isolated DNA can then be PCR-amplified and/or screened on appropriate mutation scanning arrays in order to identify the mutated genes. Alternatively, unmodified Mut Y can first be cross-linked to the DNA via borohydrite and then labeled with biotin through any of its available $-NH_2$- groups. Unreacted biotin is then removed from the DNA-enzyme mixture by known means such as by chromatography or G125 filtration. This alternate approach allows the enzyme to fully preserve its activity during the DNA binding process.

Other glycosylases that have lyase activity and can be crosslinked to DNA via borohydrite are known in the art and include: endonuclease III; endonuclease VIII; and hmu-DNA glycosylase.

Glycosylases that do not have lyase activity (e.g., Ap glycosylases: uracil glycosylase; thymine mismatch glycosylase; 3-ma-DNA glycosylase I or II; pd-DNA glycosylase; m-luteus uv glycosylase; etc) but only contain the base excision activity may also be adapted to function in the present invention. This is done by the subsequent use of endonuclease III glycosylase, following the initial base excision by the ap glycosylase. Enconuclease III recognizes the abasic site generated by the AP glycosylase and exersizes it lyase activity at that point. If borohydrite is present, endoIII is crosslinked at that position in DNA.

By following the above described approach, a biotin (or other indicator) can be introduced at the glycosylase-reactive sites in DNA (mismatches; mutations; polymorphisms; damaged bases; methylated bases; etc).

Once any of these compounds are covalently bound to the reactive sites, their reaction with a detectable group such as antibodies (e.g. avidin, antifluorescein etc.) and their subsequent detection (e.g. by chemiluminescence) and isolation (e.g. immunoprecipitation, avidin-coated microplates or microspheres, are well known in the art. For example, when X=NH_2, direct immobilization and purification of the mismatch - containing DNA is possible on microplates coated with activated succinimidyl ester [Costar] or maleic anhydrite [Pierce] which covalently bind the NH_2 group on the DNA-bound linker. When X=fluorescein, direct immobilization and isolation is achieved via antifluorescein - coated microplates [Boehringer]. And when X=biotin, direct immobilization and isolation is achieved via streptavidin - coated microplates (Pierce). In all cases, the immobilized DNA can be detected via alkaline-phosphatase or peroxidase - based chemiluminescence assays.

Those of ordinary skill in the art will recognize that a large variety of other possible detectable moieties can also be coupled to antibodies used to bind the DNA-coupled linkers at the positions of mismatches in this invention. Thereby providing additional methods to detect the antibody-bound mismatches on DNA. See; for example, "Conjugate Vaccines", *Contributions to Microbiology and Immunology*, J.M. Cruse and R.E. Lewis, Jr. (eds), Carger Press, New York, (1989).

The term "substituted," as used herein refers to single or multiple substitutions of a molecule with a moiety or moieties distinct from the core molecule. Substituents include, without limitation, halogens, hetero atoms, (i.e. 0, S and N), nitro moieties, alkyl (preferably C_1 - C_6), amine moieties, nitrile moieties, hydroxy moieties, alkoxy moieties, phenoxy moieties, other aliphatic or aromatic moieties. Preferably the aliphatic or aromatic moieties are lower aliphatic or aromatic moieties, i.e. 12 or less carbons, more preferably 6 or less carbon atoms. Substituted compounds may be referred to as derivatives of the core structure.

Antibodies of the present invention can be detected by appropriate assays, such as the direct binding assay and by other conventional types of immunoassays. For example, a sandwich assay can be performed in which the receptor or fragment thereof is affixed to a solid phase. Incubation is maintained for a sufficient period of time to allow the antibody in the sample to bind to the immobilized labeled DNA on the solid phase. After this first incubation, the solid phase is separated from the

sample. The solid phase is washed to remove unbound materials and interfering substances such as non-specific proteins which may also be present in the sample. The solid phase containing the antibody of interest bound to the immobilized labeled DNA of the present invention is subsequently incubated with labeled antibody or antibody bound to a coupling agent such as biotin or avidin. Labels for antibodies are well-known in the art and include radionucleotides, enzymes (e.g. maleate dehydrogenase, horseradish peroxidase, glucose oxidase, catalase), fluorophores (fluorescein isothiocyanate, rhodamine, phycocyanin, fluorescamine), biotin, and the like. The labeled antibodies are incubated with the solid and the label bound to the solid phase is measured, the amount of the label detected serving as a measure of, for example, the amount of anti-FARP antibody present in the sample. These and other immunoassays can be easily performed by those of ordinary skill in the art.

The present method allows for extremely sensitive mismatch - scanning in diverse DNA fragments, thereby resulting in sensitive and high throughput mutation screening over several hundreds or thousands of genes at once. For example, it becomes possible for the screening and discovery of novel mutations in tumor samples which is instrumental to establish the pathogenesis of cancer and to establish new relations between mutations and cancer or other diseases. The new compounds and methods described above are also useful in analysis of the genetic background (polymorphisms, mutations) of any individual. High throughput genotyping and genotypic selection can be carried out by the present method.

In one embodiment for detecting SNPs or mutations is illustrated in the accompanying Figure 13. In this embodiment, the following protocol can be used:

1. DNA to be screened (e.g. cDNA from an individual's lymphocytes) is digested to smaller fragments (50-300 base pairs) as detailed herein.

2. The digested fragments are self-hybridized to generate mismatches at positions of SNPs and mutations. Self hybridization can be carried out by heating (e.g., 2 minutes, 96°C), then cooling to 65°C for 1 hour. Alternatively, to avoid heating which may generate unwanted abasic sites, addition of formamide plus moderate heating (40-70°C) can be used to perform the self hybridization.

3. Preferably, treatment with hydroxylamine follows, to remove unwanted abasic sites.

4. Treatment with mismatch repair glycosylases (Mut Y and TDG, separately or together) follows to convert mismatches to aldehydes.

5. Treatment with AED, FARP, BARP follows to label, e.g., fluoresceinate the generated aldehydes (corresponding to SNPs/mutations).

6. The sample is denatured, and directly applied to a DNA array, for example, DNA chips or beads where hybridization takes place.

7. Extensive washing follows, to remove unhybridized DNA or unbound FARP.

8. The label, e.g., fluorescence, from all elements on the DNA chip is read via appropriate devices (e.g. a scanning laser). These elements that display fluorescence correspond to gene fragments containing SNPs and mutations.

In the present method, one can use DNA arrays to identify the gene where the mismatch is present. For example, the Affymetrix, Inc. (San Diego, CA) HU6800 DNA chip, the Clontech Atlas™ DNA array (Palo Alto, CA); the Telechem International Array (San Jose, CA); the Genetix Ltd. array (Dorset, UK); and the BioRobotics Ltd. array (Cambridge, UK). The chip such as the Affymetrix DNA chip contains densely-packed DNA or RNA elements. For highest resolution the oligomers on the chip should be small. Preferably 8-50 nucleotides, more preferably 8-25 nucleotides. This will provide the highest resolution. However, the DNA or mRNA on the chips can be as large as the mismatch-containing DNA fragments, e.g. 50-300 nucleotides.

For example, using a conventional array, (e.g., the Affymetrix chip for detecting gene expression) the array will have multiple DNA or RNA elements densely packed, each comprising 25-mer oligonucleotides immobilized on a solid support. For each of the 6,800 genes which are represented on the chip, there are 20 elements each containing 25-mer oligonucleotides with a distinct portion of the mRNA sequence. Thereby the 20 elements 'sample' the mRNA sequence of the gene. In the current version, the immobilized probes are biased towards the 3' end of the mRNA, thus sequences towards the 5' end are not well represented. To use the array for detecting gene expression, users generate cDNA from the genes to be screened in the test sample (typically 1:g) and then perform in-vitro transcription to collect cRNA and biotinylate it(~50:g), 12 :g of which are hybridized on the chip (alternatively, cDNA can directly be applied on the chip without in-vitro transcription). If a gene is present in the test sample, then it hybridizes to an appropriate array element. Because the array is constructed to contain known gene sequences at known positions, all the transcribed genes are detected in a single step. The detection process utilizes addition of a marker-identifier such as a

fluorescent scanner. The magnitude of the signal from each element signifies the degree of gene expression for the specific gene.

It should be noted that the present invention utilizes the same arrays but not anymore to detect gene expression (i.e. difference in signal among array elements), but mutations, which requires only detection of presence or absence of signal (indicating polymorphism/mutation in the specific gene fragment which was captured), thereby making the detection task much simpler.

Inherited single nucleotide polymorphisms (SNPs) and mutations can define a genetic predisposition towards several diseases, including cancer, cardiovascular, neurodegenerative and others. Indeed, acquired SNPs, mutations and loss of heterozygocity are particularly pertinent to cancer development, and early cancer detection. All of the above can be simultaneously detected in a single step by the above-described methodology.

For example, cDNA for tumor and normal issue of a single individual is prepared. (See Figure 11) Because inherited polymorphisms is a frequent event (average 1 SNP per 1000 bases), several genes will have more than one SNP. Also, the tumor genes will contain one or more inherited SNPs as well as occasional acquired SNPs/mutations. Next, the cDNA is digested by enzymes down to small fragments (~100-200 bp), thereby generating fragments that are likely to contain only one -or none- genetic alternations. Then, each sample is melted and self-hybridized, to generate mismatches at positions of SNPs/mutations. The above-described methodology using an X-Z-Y compound is applied as described above, to isolate only the mismatch-containing cDNA.

The mismatch-containing cDNA is PCR-amplified, labeled, e.g. biotinylated, and applied on a an array such as the Affymetrix chip: Each mismatch-containing fragment will hybridize to its complementary oligonucleotide on the array, thereby revealing which gene and which gene region (to within 100-200 base pairs) the SNP/mutation belongs to. By comparing arrays A and B, both the inherited and the acquired SNPs/mutations can be derived. Loss of heterozygocity may occur when an acquired SNP/mutation occurs in the same gene with an inherited SNP/mutation. Such genes can readily be identified by comparing A and B.

Current chip arrays, which are intended for gene expression detection, including the Affymetrix chips, utilize immobilized oligonucleotides which are biased toward the 3' end of mRNA. Accordingly, the 5' end of the gene is underrepresented. Moreover, present mutation detecting technology cannot

adequately utilize existing chips. By contrast, our methodology makes it possible to identify mismatches over diverse sections of the genome using chip technology.

A preferred Mutation Scanning Array should contain immobilized oligonucleotides, preferably 8-25 bases long, which span the whole mRNA sequence of each gene represented on the array, and not biased toward one or the other mRNA end. Rather they should cover the whole genes being studied. In some preferred instances, one uses genomic DNA to make the array and would use both coding, and non-coding portions. As mentioned, the oligonucleotides can be larger, but by increasing size, resolution is lost. The oligonucleotides should sample the mRNA in intervals not bigger than the DNA fragments isolated by present method preferably 50-100 bases but capable of ranging from 50 - 300 bases. In this manner the mismatch-containing fragment will be assured of finding a complementary sequence on the array. When immobilized oligonucleotides on the array are arranged to sample the mRNA at small intervals (e.g. 20 bases) there will be redundancy of information upon hybridization of the mutant fragments to the DNA chip, as each fragment may simultaneously hybridize to two or more immobilized oligonucleotides. In this case, by using the combined information from all array elements, a better resolution of the position of the mutation will be achieved.

This Mutation Scanning Array can be constructed using the same technologies as for the current arrays. The above-described modification will allow SNPs/mutation detection over the whole length of the immobilized genes to be identified. The immobilized genes can be either the whole genomic cDNA library, or an arbitrary fraction of that, or a specific collection of genes that are known to be related to a specific disease (i.e. disease specific arrays).

A major advantage of the present mutation scanning chip technology is that it can detect SNPs/mutations in the presence of an excess of normal alleles in the initial sample because the methodology first isolates the mutants, and the array subsequently identified the gene. This is currently impossible to do with existing technology.

A preferred kit will comprise reagents to isolate mRNA from tissues, synthesize cDNA, fragment DNA to 100 - 200-mers and add PCR, form heteroduplexes, use Mut Y and TDG enzymes to cut the mismatches, remove spontaneous aldehydes, apply the X-Z-Y compounds e.g., FARP/BARP/AED, to detect mismatches, isolate mutant/polymorphic fragments by immobilization on

microplates, recover and PCR mutants/polymorphisms, and finally apply on an array to detect SNPs/mutations at specific genomic positions.

The kit can be used to screen an individual for inherited susceptibility to cancer, cardiovascular disease, neurodegenerative disorders, etc. by mapping positions of heterozygocities and SNPs in the whole genome or in selected fractions of the genes.

The present methodology also permits one to detect early onset of cancer (acquired SNPs/mutations) from tissue biopsies or excretions. The present technology also permits research labs to detect new mutations and correlate them to other diseases.

The ligand compounds described demonstrated excellent detection of DNA mismatch - repair recognition sites. In addition, based on our discovery that small (MW<200-250) compound allow high binding efficiency (>50%) to DNA reactive sites, new compounds (like AED) were designed, synthesized and tested. These were shown to bind reactive sites generated by Mut Y much more efficient than compounds of higher (>250) molecular weight. Preferred compounds are small, water soluble, do not encounter significant steric interactions with DNA and can diffuse fast to the enzymatically - generated reactive sites on DNA. This class of new bifunctional compounds is also uniquely designed to retain their water solubility as the chain length is extended. The simultaneous addition of internal polar functional groups along with methylene groups maintains the water solubility of these compounds in spite of the increased length of the molecule. Care must be taken however to retain a low overall molecular weight for the final compound. Useful polar functional groups include; alcohols, esters, ethers, thioethers, amines and amides. This allows users of this method the flexibility to tailor the chain length of the compounds to suit their specific needs with out the loss of water solubility, which is essential.

In one method for obtaining DNA containing mismatches from a tumor sample, mRNA is isolated from a malignant cell. The corresponding mRNA from a healthy or normal tissue sample is also isolated. The mRNA from the normal tissue will serve as the wild-type control. A cDNA library can be made for each mRNA sample, the cancerous and wild-type. The two cDNA libraries are added together, for example in a 1:1 ratio and hybridized. (See Figure 1) The hybridization produces a mixture of double stranded DNA. The double strands of DNA that consist of cDNA from the malignant cell hybridized with a strand of wild-type DNA

will now typically contain some mismatches that are associated with the malignancy.

The mixture of hybridized cDNA is then treated with hydroxylamine to remove any spontaneous aldehydes, and then the hydroxylamine is removed via G25 filter centrifugation of the samples. The double stranded cDNA which is now void of pre-existing aldehydes, is then treated with a mismatch - repair glycosylase, such as Mut Y or TDG. Mut Y is a DNA-repair enzyme that recognizes mismatched adenosine nucleotides, and TDG recognizes mismatched thymines. Upon recognition, Mut Y or TDG remove the base by cleavage at the point of attachment to the deoxyribose sugar. Removal of the base by this method of cleavage results in the opening of the deoxyribose ring with formation of an aldehyde. Since pre-existing aldehydes were removed by hydroxylamine treatment, the only aldehydes are those generated at positions of mutations.

The resulting strands of cDNA now contain an aldehyde located at each point of mismatch. These resulting aldehydes are then treated with one of the compounds, e.g. the 2-(aminoacetylamino)ethylenediamine (AED) or one of its analogues, at low temperature so that further activity of the Mut Y/TDG enzymes is suppressed. The DNA labeled with AED is then selectively immobilized on microplates as described earlier in this text. The unlabeled DNA is then washed away leaving behind only AED labeled DNA attached to the microplate. The DNA with the labeled mutations, while immobilized on the microplates is then biotinylated and the mutations can be detected, for example, via chemiluminescence. Mutation-containing DNA can then be recovered from microplates for identification of the genes involved via PCR and large-scale hybridization techniques using the present mutation detecting arrays. Consequently, all mismatch containing genes are captured at once and the number of genes that can be simultaneously be screened is only limited by the total genes the DNA array created has to verify and identify the exact position of the mismatche(s) on each particular gene identified by the present invention, conventional procedures such as sequencing can be used.

This method can also be used to detect a variety of other DNA lesions that are converted to reactive sites by glycosylase enzymes or by chemical means (e.g. clustered DNA-damaged sites); abasic sites; carcinogen-DNA adducts; damaged DNA bases). In these embodiments, mixing of for example the target DNA with wild-type DNA to create mismatches is not needed. Enzymes will recognize damage and will generate reactive sites directly in the target DNA. Such enzymes include

all known glycosylases, such as endonuclease III, T4 endonuclease V, 3-methyladenine DNA glycosylase, 3- or 7- methylguanine DNA glycosylase, hydroxymethyluracile DNA glycosylase, FaPy-DNA glycosylase, M. Luteus UV-DNA glycosylase. Also, chemical agents such as bleomycin, alkylation agents or simple acid hydrolysis can generate reactive sites automatically in target DNA without any enzyme. The crucial step however is again the same, i.e. covalent addition of compound to the reactive site of the DNA lesion, which allows subsequent sensitive detection.

The described technology can be used for mutation screening and for research. For example, the use of solid supports at every stage of the assay will substantially shorten the time required to screen tumor samples, improve its cost-effectiveness in terms of man-power as well as its reliability and reproducibility.

An alternative mutation scanning array to the chip array is the use of beads, sometimes referred to as microbeads or microspheres. For instance, magnetic microsphere technology can be utilized to immobilize heteroduplexes at an early stage of the assay. Following mRNA extraction from e.g., a host cell such as cancerous and normal samples, cDNA for e.g. 588 genes can be generated. Thereafter PCR primers that contain a cleavable (S-S) biotin are added. Hybridization of the cancerous cDNA with wild-type alleles generates heteroduplexes at the positions of base substitution mutations, and the DNA sample is immobilized on, for example, the streptavidin - coated magnetic microspheres (available from Dynal Inc.). From this point onwards, all subsequent steps of the ALBUMS assay can be conducted on the solid support.

The microspheres allow chemical/enzymatic treatment of the immobilized DNA and efficient, rapid separation of chemicals from DNA via magnetic immobilization of the microspheres during washing. For example, in one embodiment the assay uses hydroxylamine treatment to remove traces of aldehydes and subsequent complete removal of hydroxylamine via repeated (x3) ultracentrifugation through G25 filters. This can be time-consuming and result in an inevitable loss of sample, which can be important when tissue samples are limited. In contrast, by immobilizing the DNA magnetic microspheres, all subsequent steps become faster, easier and without DNA loss: Hydroxylamine treatment and removal, enzymatic treatment and washing, X-Z-Y treatment and washing, binding antifluorescein-AP to e.g. AED-trapped mismatches and washing,

and finally chemiluminescent detection of mismatches are performed on the magnetic microsphere format.

Alternatively, to recover the DNA from magnetic microspheres and isolate the X-Z-Y, e.g. FARP, containing DNA, instead of adding antifluorescein-AP the immobilized DNA can be recovered by cleaving the disulfide (S-S) bond on the biotin by mild exposure to a reducing reagent (DTT, 50 mM, ~10 min, 25° C).

To construct primers end-labeled with a cleavable moiety such as biotin, oligonucleotides containing a terminal aliphatic amine are ordered, and reacted with e.g. a biotin -S-S- succinimidyl ester (available from Pierce). Reactions of succinimidyl ester with amino-oligonucleotides and subsequent purification by reverse C18 column chromatography are standard procedures on which our group has had prior experience.

Following removal of DNA samples from the magnetic microspheres, the samples will be applied on e.g. antifluorescein-microplates to isolate e.g., FARP-containing heteroduplexes which subsequently will be recovered, PCR amplified and screened on the Clontech DNA hybridization array. Using the above procedures, base substitution mutations can be isolated via ALBUMS, amplified by PCR and screened on the DNA array in less than 24 hours. Thus, this technique results in a standardized procedure with easy access to researchers and clinicians for cost - effective, large - scale mutation screening of a target sample, such as cancer samples.

See also U.S. Patent Nos. 5,736,330 and 5,981,180 and the products of Luminex Corporation (Austin, Texas). Flow cytometry can be used for diverse applications in hematology, oncology, cell biology, etc. Apart from cultured cells, beads (also known as 'microbeads', 'microspheres') tagged with fluorescent probes, or with biomolecules carrying fluorescent probes, are commonly used. During flow cytometry, such fluorescent microbeads are forced to flow down a thin tube and are individually excited by one or more laser beams. Light emitted from each microbead is then individually filtered and measured by an attached light detector. Depending on the signals obtained, individual microbeads can be separated (sorted) from the rest of the microbead population. Common flow cytometers can collect individual light signals and sort 10,000-30,000 microbeads per second. As a result, 10^8 microbeads can be sorted in less than one hour. Specialized flow cytometers can count individual microbeads at much higher rates.

Multiparametric flow cytometry allows each individual microbead to be excited by several lasers at once, and illumination by each laser produces optical emissions at

several discrete wavelengths and intensities, depending also on the type and amount of fluorescently-labeled biomolecules bound to the microbeads. As a result, passage of each individual microbead through the flow cytometer can result to emission of a large set (5-7) of signals, which are individually detected and stored on a computer. The capability of rapidly sorting individual microbeads depending on the collected parallel optical signals makes flow cytometry a powerful tool for analyzing numerous genes in a very short period.

When a microbead tagged with a specific gene, is also tagged with a combination of fluorescent probes it can be used in flow cytometry. The fluorescent probes with which the beads are tagged can have a variety of different fluorescent intensities. Consequently, for each bead there is a unique combination of intensities/fluorescent probes such that, passage of an individual bead through the flow cytometer uniquely identifies the immobilized gene. Further, if the gene immobilized on the bead has hybridized to a 'target DNA', the hybridization produces a unique fluorescent signal which may also be detected by the flow cytometer as the bead passes through. Therefore hybridization in hundreds or thousands of diverse genes can be rapidly detected, quantified and analyzed by this procedure.

Microbeads for flow cytometry are commercially available by several manufacturers (e.g. Polysciences; Molecular Probes). A typical microbead consists of an approximately spherical polystyrene 'core' with a diameter of 0.1-20 μm. The microbead can be tagged with indicator molecules such as fluorescent probes of appropriate wavelength, which are either directly bound to the microbead surface, or bound to a nucleic acid that coats the microbead surface, or fill the 'interior volume' of the microbead. When the microbead passes through the laser beam in a flow cytometer, an intense fluorescent signal is emitted, which is filtered and counted with a photomultiplier. Depending on constraints set by the user (e.g. certain intensity in the observed signal; or a certain combination of emitted fluorescent wavelengths) microbeads can be sorted into separate containers after passage through the cytometer laser beam.

Microbeads can readily be tagged with specific DNA fragments. A standard method to achieve this is to manufacture microbeads with 'functionalized surfaces', e.g. coated with carboxyl- or amino-groups, or with avidin, etc. Such microbeads are widely available. Binding of a nucleic acid on the functionalized surface is achieved via end-labeling the nucleic acid, which is then attached to the microbead surface. For example, if a primary amine is attached to the nucleic acid, then a

carbodiimide-mediated reaction can attach the nucleic acid to the carboxyl-coated microbead. If a biotin is attached to the nucleic acid, then this will bind to the avidin-coated microbead, etc. The single stranded DNA attached to the microbeads will be called hereafter 'control DNA'.

When microbeads coated with control DNA are mixed and allowed to hybridize with single stranded 'target DNA' which is to be analyzed for gene expression, polymorphisms or mutations, hybridization will take place if a sequence complementary to the control DNA sequence exists. Hybridization should preferably be under conditions of at least moderate stringency, more preferably high stringency. These conditions can be obtained by varying salts, temperature, etc. and are well known in the art. See e.g., *Sambrook, et al, Molecular Cloning Second Edition.* The unhybridized DNA can then be removed from the solution, e.g. via centrifugation, if desired. (This is not an absolute requirement since flow cytometry will only count fluorescnce bound to the microbeads and not in solution). Such hybridization of DNA in microbead-format is well known to those experienced in the art, (e.g. to isolate mRNA, poly-dT-coated microbeads are used). To detect signals from microbeads containing hybridized double stranded sequences the target DNA can be pre-labeled with a fluorescent probe. Accordingly, fluorescent signals from microbeads with control DNA hybridized to target DNA can be readily detected via flow cytometry.

In one embodiment, beads can also be labeled with a 'cocktail' of fluorescent probes, each fluorescent probe having a specific emission intensity (Such probes are often used for calibration purposes). Multiparametric flow cytometry, which is a common application, can simultaneously acquire signals from all fluorescent probes bound to the beads and measure their intensity.

For example, a flow cytometer can monitor simultaneously 6 fluorescent emission wavelengths (e.g. by using 3 excitation lasers with two different filters each). The bead can be labeled with, for example, fluorescent probes 1-5 chosen so that they have either distinct excitation or distinct emission wavelengths, with minimal overlap regions. Different amounts of each fluorescent probe can be incorporated on the bead, so that at least 20 distinct fluorescent intensities can result per probe (see, for example Figure 14, where beads labeled with 5 different fluorescent intensities of fluorescein are depicted.

Accordingly, in this example there are up to 3,200,000 combinations of beads that can be constructed, each one different from the others, which upon passage through a flow cytometer would give a unique optical signature. (In

practice one will not need to use so many combinations, as there are only about 70,000 human genes most of which may not need to be counted in a particular application. A typical application may utilize 1,000 human genes, represented by 1,000 – 20,000 different beads, each bead containing a different section of a gene).

If each bead is tagged with a different gene, or gene fragment, any desired number of genes can be analyzed in a single experiment, via flow cytometry. Assume also, that the 'target DNA' which is hybridized to the beads of known optical signature tagged with 'control DNA' is tagged with a 6th fluorescent probe, which gives a unique signal, clearly distinguishable from signals by probes 1-5. Upon passage through a multiparametric flow cytometer, each bead would yield (a) its optical signature by monitoring the intensity of emissions from probes 1-5, and thereby the identity of the control gene bound to the bead; and (b) the intensity of emission from probe 6 which would determine if, and to what extent the target DNA has hybridized to the control gene.

The flow cytometer used preferably has the ability to detect 'rare events'. Each B.U.S. will only rarely emit its unique optical signal since other beads will be counted most of the time. If 10,000 distinct B.U.S. are included in an aliquot, representing 1,000 human genes, the flow cytometer must be able to detect and discriminate 1 in 104 beads, without 'noise'. This is well within the capabilities of common flow cytometers, which can typically discriminate 1 in 105 events.

In some embodiments, enhanced software and computer storage space may be needed to record each event separately and store it in an appropriate 'bar histogram' (each bar representing a unique gene). Such software currently exists for applications that deal with a limited number of uniquely labeled beads/cells. The current invention requires that the software handles thousands of different parameters, following the optical detection of unique signatures from a total of 107 to 108 beads. Extension of the current software and disk storage space to satisfy the demands of the current application can readily be accomplished.

Several control genes and beads can be included in the flow-cytometric determination, in order to ensure that the method performs optimally.

The above method can be adapted for the detection of SNPs and mutations. In this approach, an intermediate step is used to isolate and purify only those DNA fragments that contain SNPs or mutations. This process can be carried out with the compounds and methodologies detailed herein.

Following isolation of those cDNA fragments (typically 50-200 base pair long fragments) that contain SNPs or mutations, the target DNA is hybridized to control