Electronic Acknowledgement Receipt

EFS ID:	42434474
Application Number:	17215923
International Application Number:	
Confirmation Number:	6706
Title of Invention:	SYSTEM AND SENSOR ARRAY
First Named Inventor/Applicant Name:	Omid Farokhzad
Customer Number:	26710
Filer:	Angeline R. Babel/Debra Szumowski
Filer Authorized By:	Angeline R. Babel
Attorney Docket Number:	129319.00770.BWH24258
Receipt Date:	19-APR-2021
Filing Date:	29-MAR-2021
Time Stamp:	16:52:04
Application Type:	Utility under 35 USC 111(a)

Payment information:

Submitted with Payment	no

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1	Information Disclosure Statement (IDS) Form (SB08)	129319_00770_IDS-1_form_6_of_6.PDF	1035019 ef0718fbe04de5eaf7f454acc63c4365948e829d	no	5

Warnings:

Information:					
A U.S. Patent Number Citation or a U.S. Publication Number Citation is required in the Information Disclosure Statement (IDS) form for autoloading of data into USPTO systems. You may remove the form to add the required data in order to correct the Informational Message if you are citing U.S. References. If you chose not to include U.S. References, the image of the form will be processed and be made available within the Image File Wrapper (IFW) system. However, no data will be extracted from this form. Any additional data such as Foreign Patent Documents or Non Patent Literature will be manually reviewed and keyed into USPTO systems.					
2	Non Patent Literature	Zhang_2007.PDF	590349 757a70e9f5e318ba73c5df1d8c53f9665653b300	no	6
Warnings:					
Information:					
3	Non Patent Literature	Zakynthinos_2009.PDF	156135 498c007b5fa352682a5d1457fc9904cfc2c8a735	no	17
Warnings:					
Information:					
4	Non Patent Literature	Zhang_2011.PDF	1922273 6ef360e6228f9cf5dc38736b328533ef6d3d0e2c	no	17
Warnings:					
Information:					
5	Non Patent Literature	Zhang_2014.PDF	1473390 3ae4444ab8a0484ac324fd2ef8e5f77a7c452dd3	no	14
Warnings:					
Information:					
6	Non Patent Literature	Zheng_2005.PDF	10320941 e3553ee66c47e0c33c4f68a8a013ea4752633be7	no	8
Warnings:					
Information:					
7	Non Patent Literature	Zheng_2015.PDF	1283218 a31881113e743a7a83fe8b24aa26a4e7dc9188db	no	9
Warnings:					
Information:					

8	Non Patent Literature	Zheng_2017.PDF	1497383 b035653993a1b295bdeba961288201ae9a6009ea	no	9
Warnings:					
Information:					
9	Non Patent Literature	Zupancic_2014.PDF	472574 9503f22341d6593a58d3d593e50984ef232c02a5	no	10
Warnings:					
Information:					
10	Non Patent Literature	Zwicker_2007.PDF	103525 4524476d388041c15169613511c6d140f783e9f4	no	11
Warnings:					
Information:					
			Total Files Size (in bytes):	18854807	

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

New Applications Under 35 U.S.C. 111

If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

National Stage of an International Application under 35 U.S.C. 371

If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

New International Application Filed with the USPTO as a Receiving Office

If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

PTO/SB/08a (02-18)
Approved for use through 11/30/2020. OMB 0651-0031
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

INFORMATION DISCLOSURE STATEMENT BY APPLICANT
(Not for submission under 37 CFR 1.99)

Application Number	17215923
Filing Date	2021-03-29
First Named Inventor	Farokhzad, Omid C.
Art Unit	1618
Examiner Name	–
Attorney Docket Number	129319.00770

U.S.PATENTS Remove

Examiner Initial*	Cite No	Patent Number	Kind Code[1]	Issue Date	Name of Patentee or Applicant of cited Document	Pages,Columns,Lines where Relevant Passages or Relevant Figures Appear
	1	6291183		2001-09-18	Pirrung	
	2	6361944		2002-03-26	Mirkin	
	3	6759010		2004-07-06	Lewis	
	4	6969615		2005-11-29	Knezevic	
	5	7442921		2008-10-28	Franzen	
	6	8021891		2011-09-20	Bird	
	7	9005994		2015-04-14	Huo	
	8	9234895		2016-01-12	Hood	

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)		
	Application Number	17215923
	Filing Date	2021-03-29
	First Named Inventor	Farokhzad, Omid C.
	Art Unit	1618
	Examiner Name	--
	Attorney Docket Number	129319.00770

	9	9689039		2017-06-27	Wong	
	10	9758811		2017-09-12	Caris Life Sciences Switzerland Holdings GmbH	

If you wish to add additional U.S. Patent citation information please click the Add button. Add

U.S.PATENT APPLICATION PUBLICATIONS Remove

Examiner Initial*	Cite No	Publication Number	Kind Code[1]	Publication Date	Name of Patentee or Applicant of cited Document	Pages,Columns,Lines where Relevant Passages or Relevant Figures Appear
	1	20120046184		2012-02-23	Dawson	
	2	20130058923		2013-03-07	Huo	
	3	20170131276		2017-05-11	Johnston	
	4	20180136231		2018-05-17	Immunovia Ab	
	5	20180172694		2018-06-21	Farokhzad	
	6	20180356414		2018-12-13	Strano	
	7	20180361000		2018-12-20	Weissleder	

If you wish to add additional U.S. Published Application citation information please click the Add button. Add

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)		
	Application Number	17215923
	Filing Date	2021-03-29
	First Named Inventor	Farokhzad, Omid C.
	Art Unit	1618
	Examiner Name	--
	Attorney Docket Number	129319.00770

FOREIGN PATENT DOCUMENTS								Remove
Examiner Initial*	Cite No	Foreign Document Number[3]	Country Code[2] i	Kind Code[4]	Publication Date	Name of Patentee or Applicant of cited Document	Pages,Columns,Lines where Relevant Passages or Relevant Figures Appear	T[5]
	1	1308520	EP		2003-07-05	MicroBioSytems Limited Partnership		
	2	2209893	EP		2010-07-28	Brown		
	3	2010097785	WO		2010-09-02	University of College of Dublin, National Univer.		
	4	2010148365	WO		2011-05-19	Johnston		
	5	2011088128	WO		2011-07-21	University of Central Florida Research Foundation		
	6	2012068226	WO		2012-08-16	Liu		
	7	2018046542	WO		2018-03-15	The University of Manchester		
	8	2018112460	WO		2018-06-21	The Brigham and Women's Hospital, Inc		

If you wish to add additional Foreign Patent Document citation information please click the Add button Add

NON-PATENT LITERATURE DOCUMENTS			Remove
Examiner Initials*	Cite No	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc), date, pages(s), volume-issue number(s), publisher, city and/or country where published.	T[5]

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)		
	Application Number	17215923
	Filing Date	2021-03-29
	First Named Inventor	Farokhzad, Omid C.
	Art Unit	1618
	Examiner Name	--
	Attorney Docket Number	129319.00770

	1	AGASTI et al. Adv Drug Deliv Rev. 2010 March 8; 62(3): 316-328. 2010.
	2	AGGARWAL, P., Hall, J.B., McLeland, C.B., Dobrovolskaia, M.A. & McNeil, S.E. Nanoparticle interaction with plasma proteins as it relates to particle biodistribution, biocompatibility and therapeutic efficacy. Advanced drug delivery reviews 61, 428-437 (2009).
	3	AHN, J.-M. & Cho, J.-Y. Current serum lung cancer biomarkers. Journal of Molecular Biomarkers & Diagnosis 2013 (2013).
	4	ALEXOPOULOS, C., Blatsios, B. & Avgerinos, A. Serum lipids and lipoprotein disorders in cancer patients. Cancer 60, 3065-3070 (1987).
	5	ALEXOPOULOS, C., Pournaras, S., Vaslamatzis, M., Avgerinos, A. & Raptis, S. Changes in serum lipids and lipoproteins in cancer patients during chemotherapy. Cancer chemotherapy and pharmacology 30, 412-416 (1992).
	6	ALI, et al. "Erlotinib-Conjugated Iron Oxide Nanoparticles as a Smart Cancer-Targeted Theranostic Probe for MRI." Scientific reports vol. 6 36650. 11 Nov. 2016, doi:10.1038/srep36650
	7	AMICI, A. et al. In vivo protein corona patterns of lipid nanoparticles. RSC Advances 7, 1137-1145 (2017).
	8	ANDERSEN, J.D. et al. Identification of candidate biomarkers in ovarian cancer serum by depletion of highly abundant proteins and differential in-gel electrophoresis. Electrophoresis 31, 599-610 (2010).
	9	ANDERSON, L. Candidate-Based Proteomics in the Search for Biomarkers of Cardiovascular Disease. The Journal of physiology 2005, 563, 23-60.
	10	ANGEL, T.E. et al. Mass spectrometry-based proteomics: existing capabilities and future directions. Chemical Society Reviews 41, 3912-3928 (2012).
	11	ASKIM, J. R., Mahmoudi, M. & Suslick, K. S. Optical sensor arrays for chemical sensing: the optoelectronic nose. Chemical Society Reviews 42, 8649-8682 (2013).

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)		
	Application Number	17215923
	Filing Date	2021-03-29
	First Named Inventor	Farokhzad, Omid C.
	Art Unit	1618
	Examiner Name	--
	Attorney Docket Number	129319.00770

	12	BAGALKOT, V. et al. Quantum dot-aptamer conjugates for synchronous cancer imaging, therapy, and sensing of drug delivery based on bi-fluorescence resonance energy transfer. Nano letters 7, 3065-3070 (2007).	
	13	BAKHTIARY, Z. et al. Targeted superparamagnetic iron oxide nanoparticles for early detection of cancer: Possibilities and challenges. Nanomedicine: Nanotechnology, Biology and Medicine 12, 287-307 (2016)	
	14	BALLY, M., et al. "Liposome and lipid bilayer arrays towards biosensing applications." Small 6.22 (2010): 2481-2497.	
	15	BECK, H.C., Overgaard, M. & Rasmussen, L.M. Plasma proteomics to identify biomarkers–application to cardiovascular diseases. Translational Proteomics 7, 40-48 (2015).	
	16	BENJAMIN, E. J., Blaha, M. J., Chiuve, S. E., et al. Heart Disease and Stroke Statistics—2017 Update: A Report from the American Heart Association. Circulation 2017, 135, e146-e603.	
	17	BERI, J., Rosenblatt, M.M., Strauss, E., Urh, M. & Bereman, M.S. Reagent for Evaluating Liquid Chromatography–Tandem Mass Spectrometry (LC-MS/MS) Performance in Bottom-Up Proteomic Experiments. Analytical chemistry 87, 11635-11640 (2015).	
	18	BERTRAND, N. et al. Mechanistic understanding of in vivo protein corona formation on polymeric nanoparticles and impact on pharmacokinetics. Nature Communications 8, 777 (2017).	
	19	BIGDELI, A. et al. Exploring Cellular Interactions of Liposomes Using Protein Corona Fingerprints and Physicochemical Properties. ACS nano 10, 3723-3737 (2016).	
	20	BIO-RAD, ProteoMiner Protein Enrichment Technology	
	21	BISKER, et al. Protein-targeted corona phase molecular recognition. Nat Commun 7, 10241 (2016). https://doi.org/10.1038/ncomms10241	
	22	BLOOM, D., Cafiero, E., Jané-Llopis, E., et al. The Global Economic Burden of Noncommunicable Diseases. Program on the Global Demography of Aging;2012.	

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)		
	Application Number	17215923
	Filing Date	2021-03-29
	First Named Inventor	Farokhzad, Omid C.
	Art Unit	1618
	Examiner Name	--
	Attorney Docket Number	129319.00770

	23	BLOOMSTON, M. et al. Fibrinogen ? overexpression in pancreatic cancer identified by large-scale proteomic analysis of serum samples. Cancer research 66, 2592-2599 (2006).	
	24	BODANSKY, O. & McInnes, G. F. Thermal coagulation of serum proteins in cancer, in the postoperative phase of surgery, and in the administration of adrenocorticotropic hormone. Cancer 3, 1-14 (1950).	
	25	BREDE et al. Advances in Chronic Kidney Disease, Volume 20, Issue 6, November 2013, pages 454-465. 2013.	
	26	BREIMAN, L. Random forests. Machine learning 45, 5-32 (2001).	
	27	BRENNER et al., Gene expression analysis by massively parallel signature sequencing (MPSS) on microbead arrays. Nature Biotechnology 18:630–634, 2000	
	28	BYRNE, J.C. et al. 2D-DIGE as a strategy to identify serum markers for the progression of prostate cancer. Journal of proteome research 8, 942-957 (2008).	
	29	CAO, Z., Tang, H.-Y., Wang, H., Liu, Q. & Speicher, D.W. Systematic comparison of fractionation methods for in-depth analysis of plasma proteomes. Journal of proteome research 11, 3090-3100 (2012).	
	30	CAPRIOTTI, A.L. et al. Label-free quantitative analysis for studying the interactions between nanoparticles and plasma proteins. Analytical and bioanalytical chemistry 405, 635-645 (2013).	
	31	CAPRIOTTI, A.L. et al. Shotgun proteomic analytical approach for studying proteins adsorbed onto liposome surface. Analytical and bioanalytical chemistry 401, 1195-1202 (2011).	
	32	CAPUTO, D. et al. A protein corona-enabled blood test for early cancer detection. Nanoscale 9, 349-354 (2017).	
	33	CARACCIOLO, G. et al. Lipid composition: a "key factor" for the rational manipulation of the liposome–protein corona by liposome design. RSC Advances 5, 5967-5975 (2015).	

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)		
	Application Number	17215923
	Filing Date	2021-03-29
	First Named Inventor	Farokhzad, Omid C.
	Art Unit	1618
	Examiner Name	--
	Attorney Docket Number	129319.00770

	34	CARACCIOLO, G., Caputo, D., Pozzi, D., Colapicchioni, V. & Coppola, R. Size and charge of nanoparticles following incubation with human plasma of healthy and pancreatic cancer patients. Colloids and Surfaces B: Biointerfaces 123, 673-678 (2014).
	35	CARACCIOLO, G., et al. "Disease-specific protein corona sensor arrays may have disease detection capacity." Nanoscale Horizons 4.5 (2019): 1063-1076.
	36	CARACCIOLO, G., Farokhzad, O.C. & Mahmoudi, M. Biological Identity of Nanoparticles In Vivo: Clinical Implications of the Protein Corona. Trends in Biotechnology 35, 257–264 (2017).
	37	CAREY, J. R. et al. Rapid identification of bacteria with a disposable colorimetric sensing array. Journal of the American Chemical Society 133, 7571-7576 (2011).
	38	CARTER, A. M. Complement Activation: An Emerging Player in the Pathogenesis of Cardiovascular Disease. Scientifica 2012, 2012.
	39	CARTER, H. B. et al. Early detection of prostate cancer: AUA Guideline. The Journal of urology 190, 419-426 (2013)
	40	CEDERVALL, T. et al. Understanding the nanoparticle–protein corona using methods to quantify exchange rates and affinities of proteins for nanoparticles. Proceedings of the National Academy of Sciences 104, 2050-2055 (2007).
	41	CERASOLI, E. et al. MiS-MALDI: microgel-selected detection of protein biomarkers by MALDI-ToF mass spectrometry. Molecular BioSystems 6, 2214-2217 (2010).
	42	CHOI, Y.-E., Kwak, J.-W. & Park, J. W. Nanotechnology for early cancer detection. Sensors 10, 428-455 (2010)
	43	COLAPICCHIONI ET AL. "Personalized lipsoome-protein corona in the blood of breast, gastric and pancreatic cancer patients", The International Journal of Biochemistry & Cell Biology. S1357-2725. 30016-9.. 10.1016/j. biocel.2015.09.002.
	44	CONSORTIUM, E.P. Europe PMC: a full-text literature database for the life sciences and platform for innovation. Nucleic acids research, gku1061 (2014).

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)		
	Application Number	17215923
	Filing Date	2021-03-29
	First Named Inventor	Farokhzad, Omid C.
	Art Unit	1618
	Examiner Name	--
	Attorney Docket Number	129319.00770

	45	CONSORTIUM, U. UniProt: a hub for protein information. Nucleic acids research, gku989 (2014).
	46	CORBO, C. et al. Unveiling the in Vivo Protein Corona of Circulating Leukocyte-like Carriers. ACS Nano (2017).
	47	CORBO, C., Molinaro, R., Parodi, A., Furman, N. E. T., Salvatore, F., Tasciotti, E. The Impact of Nanoparticle Protein Corona on Cytotoxicity, Immunotoxicity and Target Drug Delivery. Nanomedicine 2016, 11, 81-100.
	48	CORBO, C., Molinaro, R., Tabatabaei, M., Farokhzad, O.C. & Mahmoudi, M. Personalized protein corona on nanoparticles and its clinical implications. Biomaterials Science 5, 378-387 (2017).
	49	CORBO, C., Molinaro, R., Taraballi, F., et al. Effects of the Protein Corona on Liposome–Liposome and Liposome–Cell Interactions. International Journal of Nanomedicine 2016, 11, 3049.
	50	CROFT, D. et al. The Reactome pathway knowledgebase. Nucleic acids research 42, D472-D477 (2013).

If you wish to add additional non-patent literature document citation information please click the Add button Add

EXAMINER SIGNATURE

Examiner Signature		Date Considered	

*EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through a citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

[1] See Kind Codes of USPTO Patent Documents at www.USPTO.GOV or MPEP 901.04. [2] Enter office that issued the document, by the two-letter code (WIPO Standard ST.3). [3] For Japanese patent documents, the indication of the year of the reign of the Emperor must precede the serial number of the patent document. [4] Kind of document by the appropriate symbols as indicated on the document under WIPO Standard ST.16 if possible. [5] Applicant is to place a check mark here if English language translation is attached.

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)		
	Application Number	17215923
	Filing Date	2021-03-29
	First Named Inventor	Farokhzad, Omid C.
	Art Unit	1618
	Examiner Name	--
	Attorney Docket Number	129319.00770

CERTIFICATION STATEMENT

Please see 37 CFR 1.97 and 1.98 to make the appropriate selection(s):

That each item of information contained in the information disclosure statement was first cited in any communication from a foreign patent office in a counterpart foreign application not more than three months prior to the filing of the information disclosure statement. See 37 CFR 1.97(e)(1).

OR

☐ That no item of information contained in the information disclosure statement was cited in a communication from a foreign patent office in a counterpart foreign application, and, to the knowledge of the person signing the certification after making reasonable inquiry, no item of information contained in the information disclosure statement was known to any individual designated in 37 CFR 1.56(c) more than three months prior to the filing of the information disclosure statement. See 37 CFR 1.97(e)(2).

See attached certification statement.

The fee set forth in 37 CFR 1.17 (p) has been submitted herewith.

☒ A certification statement is not submitted herewith.

SIGNATURE

A signature of the applicant or representative is required in accordance with CFR 1.33, 10.18. Please see CFR 1.4(d) for the form of the signature.

Signature	/Angeline R. Babel/	Date (YYYY-MM-DD)	2021-04-19
Name/Print	Angeline R. Babel	Registration Number	60,886

This collection of information is required by 37 CFR 1.97 and 1.98. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 1 hour to complete, including gathering, preparing and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

Privacy Act Statement

The Privacy Act of 1974 (P.L. 93-579) requires that you be given certain information in connection with your submission of the attached form related to a patent application or patent. Accordingly, pursuant to the requirements of the Act, please be advised that: (1) the general authority for the collection of this information is 35 U.S.C. 2(b)(2); (2) furnishing of the information solicited is voluntary; and (3) the principal purpose for which the information is used by the U.S. Patent and Trademark Office is to process and/or examine your submission related to a patent application or patent. If you do not furnish the requested information, the U.S. Patent and Trademark Office may not be able to process and/or examine your submission, which may result in termination of proceedings or abandonment of the application or expiration of the patent.

The information provided by you in this form will be subject to the following routine uses:

1. The information on this form will be treated confidentially to the extent allowed under the Freedom of Information Act (5 U.S.C. 552) and the Privacy Act (5 U.S.C. 552a). Records from this system of records may be disclosed to the Department of Justice to determine whether the Freedom of Information Act requires disclosure of these record s.

2. A record from this system of records may be disclosed, as a routine use, in the course of presenting evidence to a court, magistrate, or administrative tribunal, including disclosures to opposing counsel in the course of settlement negotiations.

3. A record in this system of records may be disclosed, as a routine use, to a Member of Congress submitting a request involving an individual, to whom the record pertains, when the individual has requested assistance from the Member with respect to the subject matter of the record.

4. A record in this system of records may be disclosed, as a routine use, to a contractor of the Agency having need for the information in order to perform a contract. Recipients of information shall be required to comply with the requirements of the Privacy Act of 1974, as amended, pursuant to 5 U.S.C. 552a(m).

5. A record related to an International Application filed under the Patent Cooperation Treaty in this system of records may be disclosed, as a routine use, to the International Bureau of the World Intellectual Property Organization, pursuant to the Patent Cooperation Treaty.

6. A record in this system of records may be disclosed, as a routine use, to another federal agency for purposes of National Security review (35 U.S.C. 181) and for review pursuant to the Atomic Energy Act (42 U.S.C. 218(c)).

7. A record from this system of records may be disclosed, as a routine use, to the Administrator, General Services, or his/her designee, during an inspection of records conducted by GSA as part of that agency's responsibility to recommend improvements in records management practices and programs, under authority of 44 U.S.C. 2904 and 2906. Such disclosure shall be made in accordance with the GSA regulations governing inspection of records for this purpose, and any other relevant (i.e., GSA or Commerce) directive. Such disclosure shall not be used to make determinations about individuals.

8. A record from this system of records may be disclosed, as a routine use, to the public after either publication of the application pursuant to 35 U.S.C. 122(b) or issuance of a patent pursuant to 35 U.S.C. 151. Further, a record may be disclosed, subject to the limitations of 37 CFR 1.14, as a routine use, to the public if the record was filed in an application which became abandoned or in which the proceedings were terminated and which application is referenced by either a published application, an application open to public inspections or an issued patent.

9. A record from this system of records may be disclosed, as a routine use, to a Federal, State, or local law enforcement agency, if the USPTO becomes aware of a violation or potential violation of law or regulation.

to another). The central class of the Reactome data model is a *Reaction*, and subclasses of *Reaction* model these core biological events. Reactions are grouped into pathways, which in turn are assembled into a hierarchy of biological processes. Wherever appropriate, Reactome entities are linked to external reference databases such as UniProt (7), Ensembl (8), ChEBI (9) and Rhea (10). The full Reactome database schema is available at http://www.reactome.org/cgi-bin/classbrowser?DB=gk_current.

A straightforward definition of disease allows us to annotate a broad range of major disease processes at the molecular level. Diseases that can now be annotated in Reactome arise in one of three ways: a mutation, somatic or germ-line, leads to a non-functional gene product so processes that normally depend on that gene product do not take place; a mutation leads to a gene product with a novel function, enabling novel reactions whose products perturb normal human processes; or an infectious agent such as a virus introduces novel gene products whose novel reactions perturb normal human processes.

To identify disease-associated entities and events, a new 'disease' attribute is added, taking its value terms from a disease ontology [currently http://disease-ontology.org/ (11)]. This attribute is multivalued, so an entity or event that has roles in multiple disease processes can be annotated to capture all of those roles.

With this addition, diseases due to infection can be annotated within our data structure. Our existing 'species' attribute allows pathogen-derived proteins, DNA and RNA to be distinguished from molecules encoded in the human genome. The 'species' attribute, also associated with complexes, reactions and pathways, can be multivalued, allowing complexes containing both host and viral proteins or reactions involving host and viral components to be properly identified. Finally, addition of Gene Ontology (GO) (12) host_cell terms to our cell compartment vocabulary allows us to localize pathogen-derived entities accurately, in compliance with GO annotation practice.

To denote molecular properties of a protein modified by a somatic or germline mutation in the gene that encodes it, as opposed to post translational modification, we created a new class. This class of genetic modifications has subclasses to accommodate substitution of a canonical residue by a different one, the insertion or deletion of multiple contiguous residues into a canonical sequence and the generation of a fusion protein containing fragments of two canonical ones (13). The underlying chemical similarity between a protein that differs from its canonical form due to a mutation and one that differs due to co- or post-translational modification allows us to maintain compliance with the PSI-MOD standard for annotation of protein modifications (14).

Reactions involving mutated proteins as catalysts, inputs, outputs and regulators are annotated exactly as wild-type reactions are. To link reactions involving a mutated protein to those involving its normal counterpart, an optional 'normal reaction' attribute is used. A reaction in which a receptor constitutively activated by a mutation transmits a signal is thereby paired with the wild-type reaction in which a normal receptor is activated by ligand binding. A reaction catalyzed by the normal form of an enzyme is paired with the different one catalyzed by its gain-of-function mutant counterpart or with a dead-end reaction (normal inputs, no outputs) associated with its loss-of-function mutant counterpart. A limitation of this strategy at present is that, as Reactome does not capture quantitative data such as reaction rates or binding affinities, quantitative effects of mutations are not readily annotated.

At the level of our event hierarchy, these normal-disease pairings support a disease event hierarchy that parallels our normal event hierarchy, a useful and generalizable organization. These pairings also support a visualization scheme that highlights the relationship between the normal and disease processes. As described previously (15), the physical entities and their interactions that

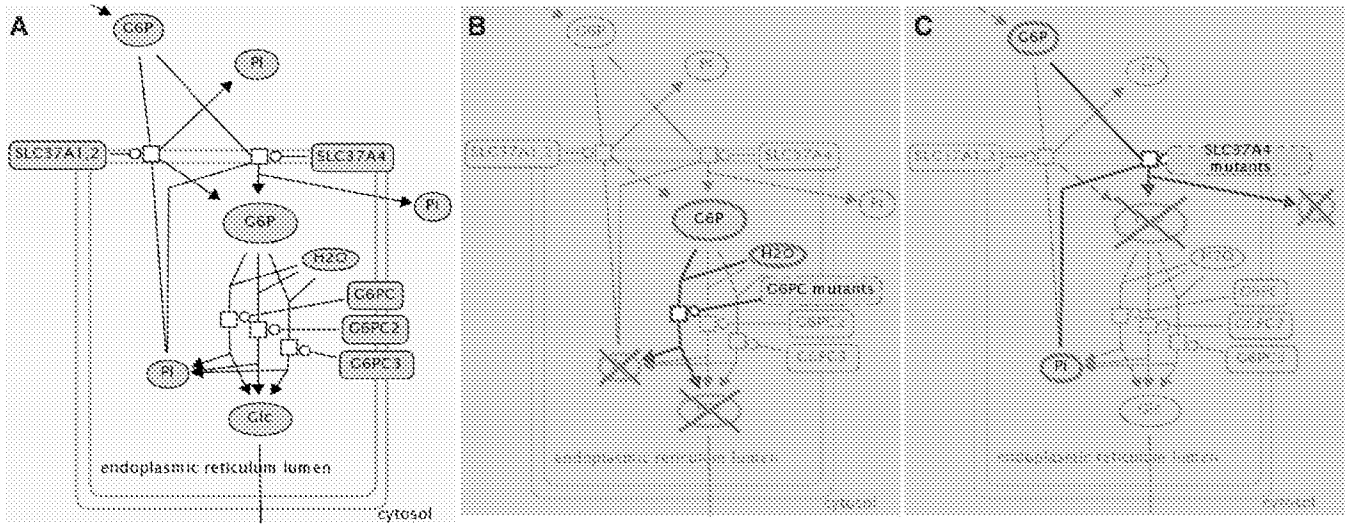


Figure 1. Example of disease curation and visualization in Reactome. The normal process of glucose export from the liver under fasting conditions (**A**) is disrupted by mutations that block glucose-6-phosphate hydrolysis within the endoplasmic reticulum (**B**) or the transport of glucose 6-phosphate and orthophosphate (Pi) between the endoplasmic reticulum and the cytosol (**C**).

Downloaded from https://academic.oup.com/nar/article-abstract/42/D1/D472/1050109
by guest
on 19 December 2017

PREOMICS EXHIBIT 1002
Page 316 of 1861

comprise a pathway are laid out in a pathway diagram that follows the SBGN process description language (http://www.sbgn.org/Documents/Specifications) and that is displayed on our Web site. The software that generates these displays has been extended to create disease displays in which the variant forms of the events responsible for a disease process are superimposed and highlighted on the normal process diagram (Figure 1).

We have used these extensions of the Reactome data model and pathway visualization process to create a new disease pathway classification in our event hierarchy that incorporates existing material such as the HIV and influenza life cycles, amyloid formation and botulinus toxin neurotoxicity, together with new material that includes malignant transformation due to mutations in the EGFR and FGFR signaling pathways and mucopolysaccharidoses. Our current release includes annotations for 420 mutant forms of 68 proteins.

Updated Reactome Web site

Our home page has been redesigned completely to support intuitive access to our pathway browsing and data analysis tools. The new Web site retains a top menu bar to provide easy access to all of our tools and resources, accompanied by a central panel of links to our most widely used tools and a footer that displays all tools and resources. This organization is consistent with the general model for resources associated with EBI. News is now available as an interactive Twitter display, which also provides open real-time feedback, both by the Reactome group and our user community.

Dynamic pathway portal

We have improved the flexibility and performance of our pathway browser by creating a new pathway diagram visualization tool using the canvas element introduced in HTML 5. The canvas element is used to render whole pathway diagrams in a Google map-like way with XML-encoded pathway diagram data retrieved from the server using a RESTful API (Figure 2A). The new diagram visualization tool offers quicker performance and better data overlaying technologies (see later in text) and bypasses the slow step to generate static images during database release.

The event hierarchy panel to the left has been redesigned to provide interactivity and access to the entire listing of all the Reactome pathways. Icons now indicate whether a pathway is new (N) or updated (U), and identify ones that are parts of disease processes (+). Navigation controls in the upper left corner enable zooming and panning across the pathway panel. A diagram thumbnail in the lower left shows the part of the pathway currently displayed in the visualization panel. A widget icon in the upper right corner of the panel links to tools for searching within the displayed pathway, overlaying the pathway with functional interactors (see later in text) and for downloading the diagram as a snapshot or PNG file.

A new tabbed 'Details' panel, below the pathway diagram, provides additional graphical and textual information. The 'Overview' tab provides summary

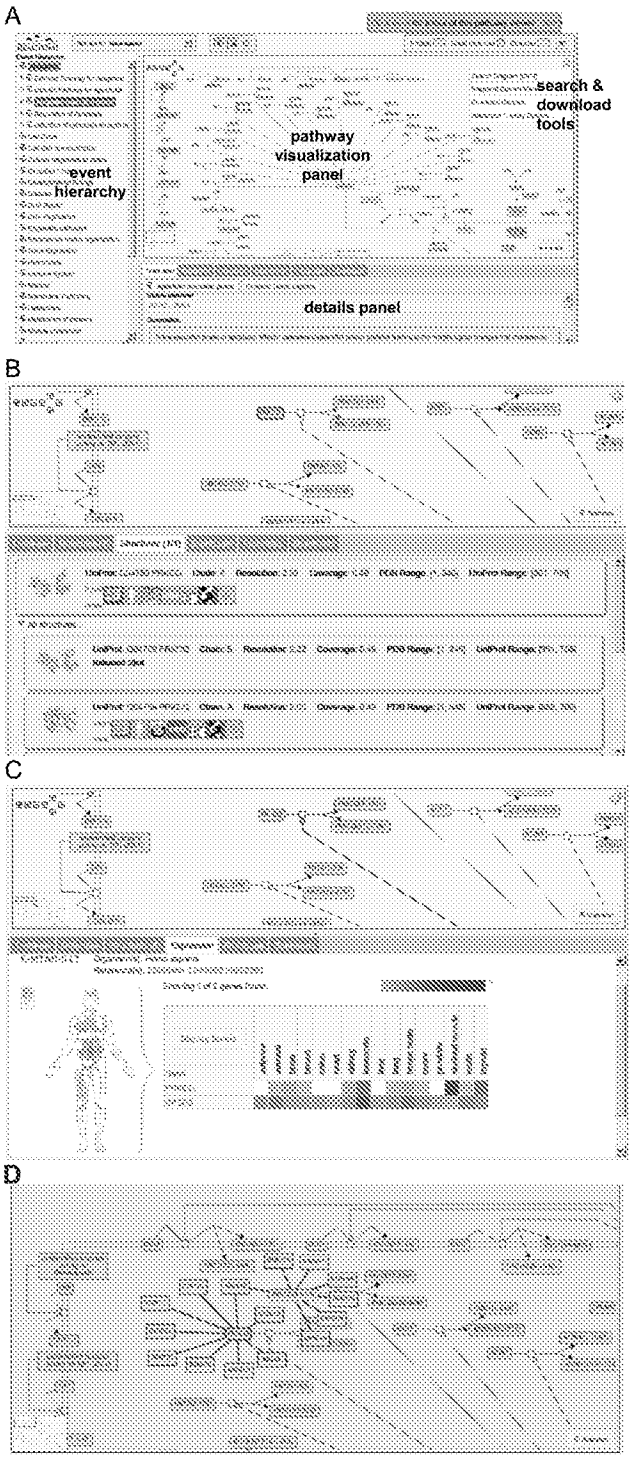


Figure 2. Visualization of the execution phase of apoptosis. (**A**) Choosing the entry for this event from the event hierarchy in the left panel of the web page causes the pathway to be displayed in a large panel to the right, laid out in SBGN process description format. Buttons at the top of the panel open a diagram key and give access to a brief tour of features of the web page. The panel at the bottom of the page contains text descriptions of various features of the pathway and participating molecules, linked to external databases. (**B**) The user has selected the reaction 'Caspase 3-mediated cleavage of PKC delta'; the 'Structures' detail tab displays PDB 3D structural data and citations for proteins in that reaction and the 'Expression' detail tab (**C**) displays condition-specific gene expression data from the Gene Expression Atlas. (**D**) The molecular interaction (MI) overlay displays proteins that interact with user-specified proteins PRKCQ and PRKCD.

Downloaded from https://academic.oup.com/nar/article-abstract/42/D1/D472/1050109
by guest
on 19 December 2017

PREOMICS EXHIBIT 1002

information relating to the pathway, reaction or entity selected. The 'Molecule' tab displays external annotation and linking-out to other bioinformatic resources describing the selected molecule. The 'Structures' tab (Figure 2B) displays 3D structural data and citations for proteins from PDB (16), for small molecules from ChEBI and for the stoichiometry of metabolic reactions from Rhea. The 'Expression' tab (Figure 2C) serves condition-specific gene expression data from the Gene Expression Atlas (17). The 'Processes' tab displays Reactome pathways and reactions associated with the selected pathway node or event. Finally, the 'Download' tab allows users to download the description of the pathway in a variety of formats compatible with third-party tools.

Based on usability testing and feedback from our user community, we modified the context sensitive menus. In this new version, the mouse cursor can be positioned on the pathway background and right-clicked to display a new set of pathway diagram options.

The molecular interaction overlay that provides the display of proteins or small molecules that interact with Reactome pathway proteins (Figure 2D) is revised to include additional interaction databases through the PSIQUIC registry (18) including BindingDB (19), DrugBank (20) and GeneMANIA (21).

Unified pathway analysis portal

We have merged pathway identifier mapping, over-representation and expression analysis tools into a single tabbed data analysis portal with integrated visualization and summary features (Figure 3). The pathway data analysis module accepts gene lists or expression data with numerical values (e.g. expression, abundance, fold expression change and quality scores). It will automatically distinguish between tab-delimited, comma-delimited or Microsoft Excel files and can cope with files that have been ZIP compressed.

The results of an analysis are displayed in three tabs that provide three different views of the same data. Pathway-oriented expression analysis provides a table of top-level pathways plus information relevant to the type of data submitted by the user. If presented with a list of identifiers, an enrichment analysis will be shown; numerical data will be summarized in a graphical way, using color to indicate average expression values. The identifier mapping view lists each of the identifiers supplied by the user with the corresponding UniProt identifier and a list of pathways in Reactome in which it functions. The overview is a completely new feature. It provides a canvas-based aggregated summary of the entire set of Reactome pathways, using visual cues to inform the user about the

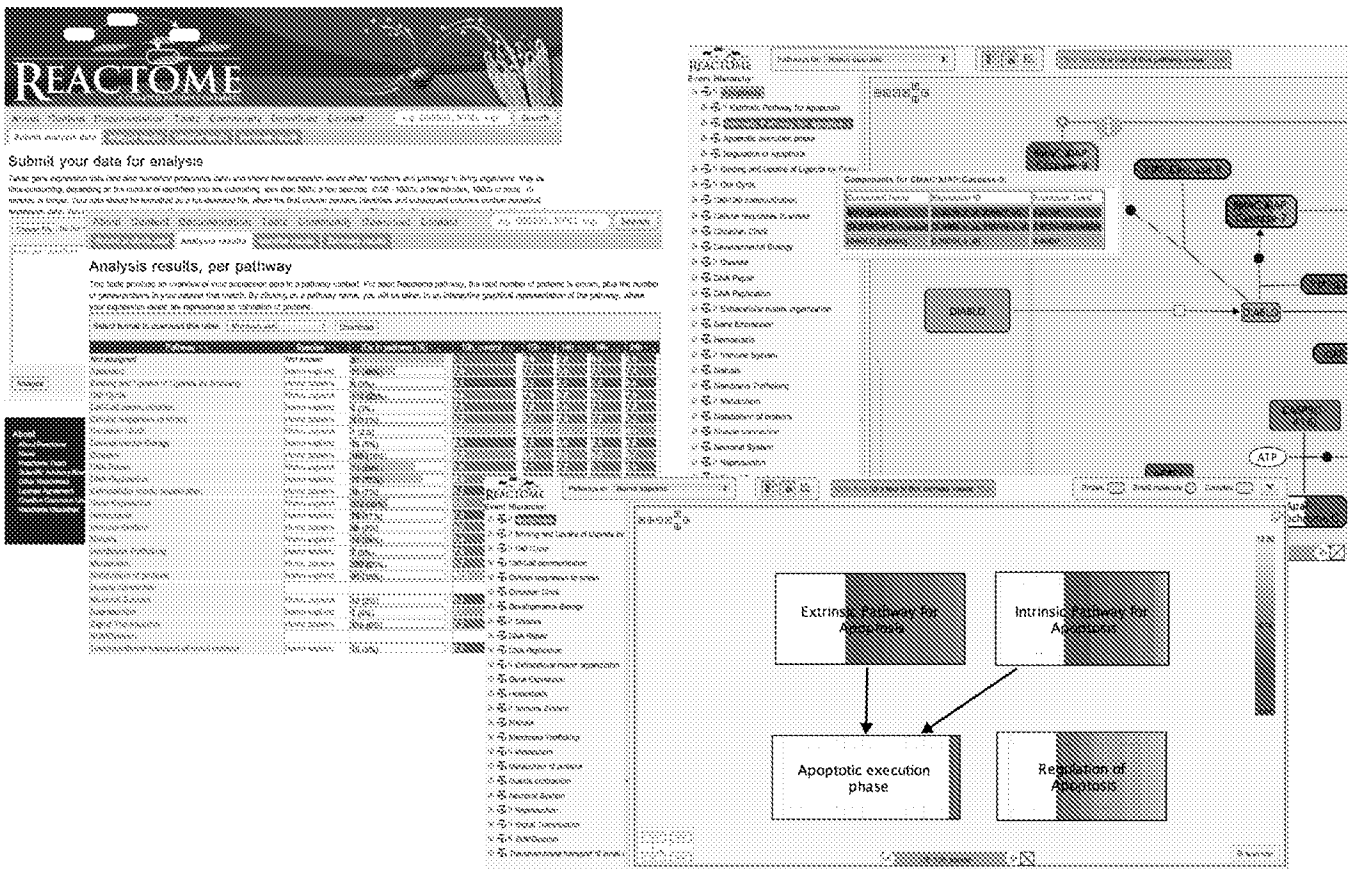


Figure 3. Pathway analysis. The data analysis workflow proceeds in three steps: data submission, a tabular display of results and visual display of results as an overlay in a pathway browser window. The fourth pane shows a detail of the pathway browser window, to highlight the display of expression levels.

Downloaded from https://academic.oup.com/nar/article-abstract/42/D1/D472/1050109
by guest
on 19 December 2017

PREOMICS EXHIBIT 1002
Page 318 of 1861

number of participating entities, relations between the pathways, average expression levels and other details.

The pathway names in the pathway-oriented view are clickable links, which take the user to the relevant pathway diagram. In it, the sub-pathway nodes, complexes and entities are colored according to expression level. Entities that lack gene expression information are colored white. Pathway and complex nodes are colored in vertical segments with each segment representing a protein in the pathway or a component of the complex. The segments are stacked in order from lowest to highest expression level. A new icon at the base of the diagram allows the user to single step through the individual experiments or time points.

Community relationships and data exchange

Reactome continues to collaborate with other data resources, such as GO, NCBI, EBI and WikiPathways (22). Reactome provides a series of link-outs to many online bioinformatics resources from its protein pages; we have added links to GeneCards annotations (23). Reactome is open-source and open-data, and we have continuously supported the major open-data standards in the domain, including BioPAX levels 2 and 3 (24), PSI MITAB (18), Protégé (http://protege.stanford.edu), SBML-ML (25) and SBGN export format. Reactome now provides an SBGN file format generated using libSBGN for individual pathways. The Reactome SBML export has been upgraded to Level 2, Version 4 and is enriched with a wide variety of additional annotations, including Systems Biology Ontology terms (26). Reactome also supports the Protein Ontology in developing an ontology for protein modifications and protein complexes (27). Our new RESTful API provides outside users with direct access to pathway data in Reactome.

The Reactome data model has been adopted by the Gramene group for manual annotation of plant pathways, especially metabolic processes specific to plants. The current release of Plant Reactome (http://plants.reactome.org) includes 131 rice pathways.

Since 2011, Reactome has participated in the Google Summer of Code program as part of the Genome Informatics group, helping to create the software components for the new pathway browser, RESTful API and the pathway overview. The Reactome data and source code continues to be publically accessible under the terms of a Creative Commons Attribution 3.0 Unported License.

ACKNOWLEDGEMENTS

The authors are grateful to the many expert scientists who have collaborated with us as external authors and reviewers of Reactome content. They thank three anonymous reviewers for their comments on this article.

FUNDING

National Human Genome Research Institute at the National Institutes of Health [U41 HG003751]; Ontario Research (GL2) fund; European Bioinformatics Institute; European Commission (PSIMEx); Google Summer of Code Program (2011–2013). Funding for open access charge: National Institutes of Health [U41 HG003751].

Conflict of interest statement. None declared.

REFERENCES

1. Biankin,A.V., Waddell,N., Kassahn,K.S., Gingras,M.C., Muthuswamy,L.B., Johns,A.L., Miller,D.K., Wilson,P.J., Patch,A.M., Wu,J. *et al.* (2012) Pancreatic cancer genomes reveal aberrations in axon guidance pathway genes. *Nature*, **491**, 399–405.
2. Gieger,C., Radhakrishnan,A., Cvejic,A., Tang,W., Porcu,E., Pistis,G., Serbanovic-Canic,J., Elling,U., Goodall,A.H., Labrune,Y. *et al.* (2011) New gene functions in megakaryopoiesis and platelet formation. *Nature*, **480**, 201–208.
3. Wu,G., Feng,X. and Stein,L. (2010) A human functional protein interaction network and its application to cancer data analysis. *Genome Biol*, **11**, R53.
4. Vastrik,I., D'Eustachio,P., Schmidt,E., Gopinath,G., Croft,D., de Bono,B., Gillespie,M., Jassal,B., Lewis,S., Matthews,L. *et al.* (2007) Reactome: a knowledge base of biologic pathways and processes. *Genome Biol*, **8**, R39.
5. Kanchisa,M., Goto,S., Sato,Y., Furumichi,M. and Tanabe,M. (2012) KEGG for integration and interpretation of large-scale molecular datasets. *Nucleic Acids Res*, **40**, D109–D114.
6. Karp,P.D., Krummenacker,M., Paley,S. and Wagg,J. (1999) Integrated pathway-genome databases and their role in drug discovery. *Trends Biotechnol*, **17**, 275–281.
7. UniProt Consortium. (2012) Reorganizing the protein space at the universal protein resource (UniProt). *Nucleic Acids Res*, **40**, D71–D75.
8. Flicek,P., Amode,M.R., Barrell,D., Beal,K., Brent,S., Carvalho-Silva,D., Clapham,P., Coates,G., Fairley,S., Fitzgerald,S. *et al.* (2012) Ensembl 2012. *Nucleic Acids Res*, **40**, D84–D90.
9. Hastings,J., de Matos,P., Dekker,A., Ennis,M., Harsha,B., Kale,N., Muthukrishnan,V., Owen,G., Turner,S., Williams,M. *et al.* (2013) The ChEBI reference database and ontology for biologically relevant chemistry: enhancements for 2013. *Nucleic Acids Res*, **41**, D456–D463.
10. Alcántara,R., Axelsen,K.B., Morgat,A., Belda,E., Coudert,E., Bridge,A., Cao,H., de Matos,P., Ennis,M., Turner,S. *et al.* (2012) Rhea—a manually curated resource of biochemical reactions. *Nucleic Acids Res*, **40**, D754–D760.
11. Schriml,L.M., Arze,C., Nadendla,S., Chang,Y.W., Mazaitis,M., Felix,V., Feng,G. and Kibbe,W.A. (2012) Disease ontology: a backbone for disease semantic integration. *Nucleic Acids Res*, **40**, D940–D946.
12. Gene Ontology Consortium. (2010) The gene ontology in 2010: extensions and refinements. *Nucleic Acids Res*, **38**, D331–D335.
13. Milacic,M., Haw,R., Rothfels,K., Wu,G., Croft,D., Hermjakob,H., D'Eustachio,P. and Stein,L. (2012) Annotating cancer variants and anti-cancer therapeutics in reactome. *Cancers*, **4**, 1180–1211.
14. Montecchi-Palazzi,L., Beavis,R., Binz,P.A., Chalkley,R.J., Cottrell,J., Creasy,D., Shofstahl,J., Seymour,S.L. and Garavelli,J.S. (2008) The PSI-MOD community standard for representation of protein modification data. *Nat. Biotechnol*, **26**, 864–866.
15. Croft,D., O'Kelly,G., Wu,G., Haw,R., Gillespie,M., Matthews,L., Caudy,M., Garapati,P., Gopinath,G., Jassal,B. *et al.* (2011) Reactome: a database of reactions, pathways and biological processes. *Nucleic Acids Res*, **39**, D691–D697.
16. Berman,H.M., Westbrook,J., Feng,Z., Gilliland,G., Bhat,T.N., Weissig,H., Shindyalov,I.N. and Bourne,P.E. (2000) The protein data bank. *Nucleic Acids Res*, **28**, 235–242.
17. Kapushesky,M., Adamusiak,T., Burdett,T., Culhane,A., Farne,A., Filippov,A., Holloway,E., Klebanov,A., Kryvych,N., Kurbatova,N. *et al.* (2012) Gene expression atlas update - a value-added database of microarray and sequencing-based

Downloaded from https://academic.oup.com/nar/article-abstract/42/D1/D472/1050109
by guest
on 19 December 2017

functional genomics experiments. *Nucleic Acids Res*, **40**, D1077–D1081.
18. Aranda,B., Blankenburg,H., Kerrien,S., Brinkman,F.S., Ceol,A., Chautard,E., Dana,J.M., De Las Rivas,J., Dumousseau,M., Galeota,E. *et al.* (2011) PSICQUIC and PSISCORE: accessing and scoring molecular interactions. *Nat. Methods*, **8**, 528–529.
19. Liu,T., Lin,Y., Wen,X., Jorrisen,R.N. and Gilson,M.K. (2007) BindingDB: a web-accessible database of experimentally determined protein-ligand binding affinities. *Nucleic Acids Res*, **35**, D198–D201.
20. Knox,C., Law,V., Jewison,T., Liu,P., Ly,S., Frolkis,A., Pon,A., Banco,K., Mak,C., Neveu,V. *et al.* (2011) DrugBank 3.0: a comprehensive resource for 'omics' research on drugs. *Nucleic Acids Res*, **39**, D1035–D1041.
21. Warde-Farley,D., Donaldson,S.L., Comes,O., Zuberi,K., Badrawi,R., Chao,P., Franz,M., Grouios,C., Kazi,F., Lopes,C.T. *et al.* (2010) The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. *Nucleic Acids Res*, **38**, W214–W220.
22. Kelder,T., van Iersel,M.P., Hanspers,K., Kutmon,M., Conklin,B.R., Evelo,C.T. and Pico,A.R. (2012) WikiPathways: building research communities on biological pathways. *Nucleic Acids Res*, **40**, D1301–D1307.
23. Safran,M., Dalah,I., Alexander,J., Rosen,N., Iny Stein,T., Shmoish,M., Nativ,N., Bahir,I., Doniger,T., Krug,H. *et al.* (2010) GeneCards Version 3: the human gene integrator. *Database*, **2010**, baq020.
24. Demir,E., Cary,M.P., Paley,S., Fukuda,K., Lemer,C., Vastrik,I., Wu,G., D'Eustachio,P., Schaefer,C., Luciano,J. *et al.* (2010) The BioPAX community standard for pathway data sharing. *Nat. Biotechnol*, **28**, 935–942.
25. Bornstein,B.J., Keating,S.M., Jouraku,A. and Hucka,M. (2008) LibSBML: an API Library for SBML. *Bioinformatics*, **24**, 880–881.
26. Courtot,M., Juty,N., Knüpfer,C., Waltemath,D., Zhukova,A., Dräger,A., Dumontier,M., Finney,A., Golebiewski,M., Hastings,J. *et al.* (2011) Model storage, exchange and integration. *Mol. Syst. Biol*, **7**, 543.
27. Natale,D.A., Arighi,C.N., Barker,W.C., Blake,J.A., Bult,C.J., Caudy,M., Drabkin,H.J., D'Eustachio,P., Evsikov,A.V., Huang,H. *et al.* (2011) The protein ontology: a structured representation of protein forms and complexes. *Nucleic Acids Res*, **39**, D539–D545.

Downloaded from https://academic.oup.com/nar/article-abstract/42/D1/D472/1050109
by guest
on 19 December 2017

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

(11) EP 1 308 520 A2

(12) EUROPEAN PATENT APPLICATION

(43) Date of publication:
07.05.2003 Bulletin 2003/19

(21) Application number: **02021593.5**

(22) Date of filing: **27.09.2002**

(51) Int Cl.[7]: **C12Q 1/04**, G01N 33/68, G01N 33/542, G01N 33/543, G01N 33/50

(84) Designated Contracting States:
AT BE BG CH CY CZ DE DK EE ES FI FR GB GR IE IT LI LU MC NL PT SE SK TR
Designated Extension States:
AL LT LV MK RO SI

(30) Priority: **01.11.2001 US 999159**

(71) Applicant: **MicroBioSystems Limited Partnership**
Cheyenne, Wyoming 82001 (US)

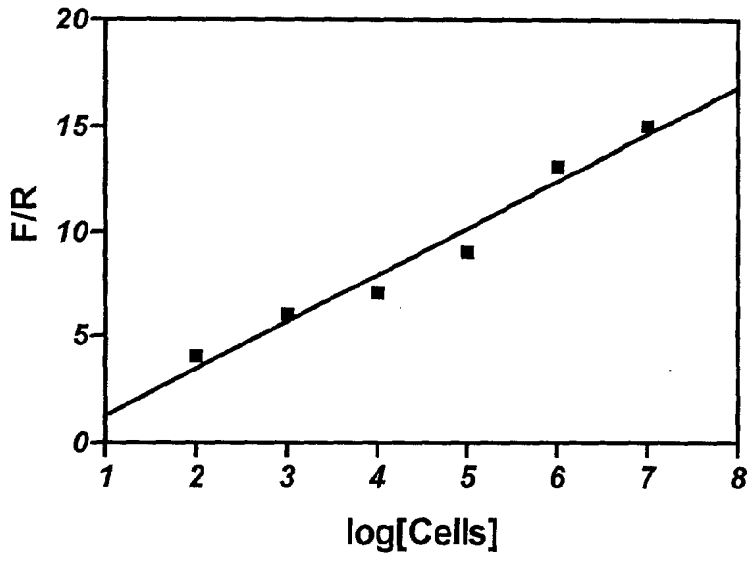
(72) Inventors:
- **Powers, Linda S.**
  Logan, Utah 84321 (US)
- **Ellis, Walther R., Jr.**
  Logan, Utah 84321 (US)
- **Lloyd, Christopher R.**
  Logan, Utah 84341 (US)

(74) Representative: **Bauer, Wulf, Dr.**
Bayenthalgürtel 15
50968 Köln (Marienburg) (DE)

(54) Taxonomic identification of pathogenic micro-organisms and their toxic proteins

(57) The present invention describes a method for the binding of pathogenic microorganisms and their toxic proteins with ligands that have been covalently tethered at some distance from the surface of a substrate: distances of at least fifteen Å are required for microorganism binding ligand tethers and at least six Å are required for protein binding ligand tethers. The ligands described herein include heme compounds, siderophores, polysaccharides, and peptides specific for toxic proteins, outer membrane proteins and conjugated lipids. Non-binding components of the solution to be analyzed are separated from the bound fraction and binding is confirmed by detection of the analyte via microscopy, fluorescence, epifluorescence, luminescence, phosphorescence, radioactivity, or optical absorbance. By patterning numerous ligands in an array on a substrate surface it is possible to taxonomically identify the microorganism by analysis of the binding pattern of the sample to the array.

Figure 1



EP 1 308 520 A2

PREOMICS EXHIBIT 1002
Page 321 of 1861

Description

BACKGROUND OF THE INVENTION

[0001] The present invention relates to a method for the taxonomic identification of pathogenic microorganisms and the detection of their proteinaceous toxins.

[0002] Pathogenic microorganisms, particularly pathogenic bacteria which either occur naturally or which have acquired virulence factors, are responsible for many diseases which plague mankind. Many of these bacteria have been proposed as biowarfare agents. In addition, there is also the risk and likelihood that nonpathogenic microbes could also be used as pathogens after genetic manipulation (e.g., *Escherichia coli* harboring the cholera toxin).

[0003] Typical pathogenic bacteria include those responsible for botulism, bubonic plague, cholera, diphtheria, dysentery, leprosy, meningitis, scarlet fever, syphilis and tuberculosis, to mention a few. During the last several decades, the public perception has been one of near indifference in industrialized nations, principally because of successes that have been achieved in combating these diseases using antibiotic therapy. However, bacteria are becoming alarmingly resistant to antibiotics. In addition, there have been recent revelations of new roles that bacteria perform in human diseases such as *Helicobacter pylori* as the causative agent of peptic ulcers, *Burkholderia cepacia* as a new pulmonary pathogen and *Chlamydia pneumoniae* as a possible trigger of coronary heart disease. Apart from those pathogens, various socioeconomic changes are similarly contributing to the worldwide rise in food-borne infections by bacteria such as *Escherichia coli*, *Salmonella* spp., *Vibrio* spp., and *Campylobacter jejuni.*

[0004] Potential infections are also important considerations in battlefield medicine. A number of bacterial pathogens, including *Bacillis anthracis* and *Yersinia pestis* and their exotoxins, have been used as weapons. And there is always the risk that nonpathogenic microbes can be engineered to be pathogenic and employed as biowarfare agents.

[0005] Pathogenic microorganisms are also of concern to the livestock and poultry industries as well as in wildlife management. For example, *Brucella abortus* causes the spontaneous abortion of calves in cattle. Water supplies contaminated with exotoxin-producing microorganisms have been implicated in the deaths of bird, fish and mammal populations. More recently, mad cow disease has been traced to the oral transmission of a proteinaceous particle not retained by filters. Thus, there is clearly a need for rapid and inexpensive techniques to conduct field assays for toxic proteins and pathogenic microorganisms that plague animals as well as humans.

[0006] As a general proposition, bacterial contamination can be detected by ordinary light microscopy. This technique, however, is only of limited taxonomic value. The investigation and quantitation of areas greater than microns in size are difficult and time consuming. Many commercially available systems rely on the growth of cultures of bacteria to obtain sufficiently large samples (outgrowth) for the subsequent application of differential metabolic tests for species (genus) identification. However, techniques requiring bacterial outgrowth may fail to detect viable but nonculturable cells. To the contrary, the growth media employed may favor the growth of bacteria with specific phenotypes.

[0007] More sensitive and more rapid typing schemes are described in "Strategies to Accelerate the Applicability of Gene Amplification Protocols for Pathogen Detection in Meat and Meat Products" by S. Pillai and S. C. Ricke (Crit. Rev. Microbiol. 21(4), 239-261 (1995)) and "Molecular Approaches for Environmental Monitoring of Microorganisms" by R. M. Atlas, G. Sayler, R. S. Burlage and A. K. Bej (Biotechniques 12(5), 706-717 (1992)). Those techniques employ the polymerase chain reaction (PCR) for amplification of bacterial DNA or RNA, followed by nucleic acid sequencing to detect the presence of a particular bacterial species. Such general amplification and sequencing techniques require technical expertise and are not easily adaptable outside of specialized laboratory conditions. PCR-based techniques utilize the inference of microbial presence since these techniques provide only a positive analysis whenever an intact target nucleic acid sequence, not necessarily a microbe, is detected. PCR is also unable to detect the presence of toxic microbial proteins. Moreover, the detection of specific microorganisms in environmental samples is made difficult by the presence of materials that interfere with the effectual amplification of target DNA in 'dirty' samples.

[0008] Mass spectral analysis of volatile cell components (e.g., fatty acids) after sample lysis or pyrolysis has been used for the detection of bacteria and viruses. One description of the methods used to detect microorganisms with this method can be found in "Characterization of Microorganisms and Biomarker Development from Global ESI-MS/MS Analyses of Cell Lysates" by F. Xiang, G. A. Anderson, T. D. Veenstra, M. S. Lipton and R. D. Smith (Anal. Chem. 72 (11), 2475-2481 (2000)). Unfortunately, identification of the analyte is unreliable as the compositions of a microbe's volatile components change depending upon different environmental growth conditions.

[0009] Another approach utilizes immunochemical capture as described in "The Use of Immunological Methods to Detect and Identify Bacteria in the Environment" by M. Schlotter, B. Assmus and A. Hartmann (Biotech. Adv. 13, 75-80 (1995)), followed by optical detection of the captured cells. The most popular immunoassay method, enzyme-linked immunosorbent assay (ELISA), has a detection limit of several hundred cells. This is well below the ID_{50} of extremely infectious bacteria such as *Shigella flexneri.* Piezoelectric detection techniques, such as those described by "Development of a Piezoelectric Immunosensor for the Detection of *Salmonella typhimurium"* by E. Prusak-Sochaczewski and J. H. T. Luong (Enzyme Microb. Technol. 12: 173-177 (1990)) are even less sensitive having a detection limitation

PREOMICS EXHIBIT 1002
Page 322 of 1861

of about 5 x 10^5 cells. A recent report entitled "Biosensor Based on Force Microscope Technology" by D. R. Baselt, G. U. Lee and R. J. Colton (Biosens. & Bioelectron. 13, 731-739 (1998)) describes the use of an atomic force microscope (AFM) to detect immunocaptured cells; this method has little utility outside a laboratory setting and when the sample volumes are large. Immunoassays are also presently used in the trace analysis of peptides and proteins.

[0010] Moreover, the prior art has made extensive use of immobilized antibodies in peptide/protein/microorganism capture. Those techniques likewise involve significant problems because the antibodies employed are very sensitive to variations in pH, ionic strength and temperature. Antibodies are susceptible to degradation by a host of proteolytic enzymes in "dirty" samples. In addition, the density of antibody molecules supported on surfaces (e.g., microwell plates or magnetic beads) is not as high as is frequently necessary. A good summary of the state of the art, still up-to-date, is "Microbial Detection" by N. Hobson, I. Tothill and A. Turner (Biosens. & Bioelectron. 11, 455-477 (1996)).

[0011] Medical and military considerations call for better toxin and pathogen detection technologies. Real-time assessment of battlefield contamination by a remote sensing unit is necessary to permit and facilitate rapid diagnosis for administration of appropriate countermeasures. A microbe/toxic protein sensor useful in such situation requires the ability to globally discriminate between pathogens and nonpathogens. In addition, such techniques require high sensitivity when less than 100 cells are present and analysis that can be completed in the field in less than 15 minutes. Such techniques should be able to recognize pathogens and provide some assessment of strain virulence or toxigenicity.

[0012] To date, common approaches used for the identification of pathogenic microorganisms and their proteinaceous toxins have employed immunological methodologies. Immunological methods suffer from the sensitivity of antibodies toward pH, ionic strength, and temperature; the antibodies themselves are subject to proteolysis and require careful storage conditions. To overcome these problems the present invention describes the capture of microorganisms and their proteinaceous toxins using non-antibody based ligands. It is accordingly an object of the present invention to provide a method for taxonomically evaluating microbes and proteins that overcome the foregoing disadvantages of technologies that depend upon antibodies.

[0013] It is a more specific object of the invention to provide a method for taxonomically evaluating microbes and proteins that has the capability of discriminating between specific microbial species, pathogens and nonpathogens, and can be likewise used to identify microbial proteins of diagnostic utility.

SUMMARY OF THE INVENTION

[0014] The present invention demonstrates the ability of heme compounds, siderophores, polysaccharides and peptides to bind to pathogenic microorganisms and their proteinaceous toxins; taxonomic identification of a microorganism is attained thorough analysis of the number and kind of ligands to which it binds. The development of this method was done to overcome the aforementioned limitations of antibody-based technologies. The concept of the present invention resides in a method for the taxonomic identification of microorganisms in which microbes are captured through the binding of microbial receptors to specific ligands. A microorganism-containing sample is contacted by the ligand, with the ligand being either tethered to a surface or conjugated to a marker. The target microbe (bacteria, virus, fungi, protozoa, rickettsiae, or other cell) or proteinaceous material (toxin) is then separated from the non-binding sample components and unbound ligand as by washing, magnetic separation or chromatography. Finally, the sample is interrogated by an appropriate method to determine if the ligand has been bound to the target by detecting signals endogenous to the target or marker.

[0015] Electromagnetic radiation is one method used to detect the presence of metabolites characteristic of living microbes, e.g., reduced pyridine nucleotides or other fluorescent metabolites, other biomolecules, e.g., notably tryptophan or tyrosine in proteins, or incorporated dyes for the detection of the presence of the captured microorganisms and/or toxins in accordance with the practice of the invention. For example, if the ligand contains a fluorescent dye, the sample will fluoresce after washing, since the ligand is bound to the cells and the excess is washed away. Other markers, including luminescent, phosphorescent, radioactive and/or colorometric compounds, can be conjugated to the ligand and used to identify a microbe and/or proteinaceous toxin in a similar manner.

[0016] One specific method to detect capture of microorganisms or toxic proteins is described in U.S. Patents 5,760,406 and 5,968,766, where electromagnetic radiation is directed, for example, onto the surface of a ligand-conjugated substrate that has been treated with an analyte-containing solution as outlined above. This detection method could be used to determine if binding of an analyte has occurred. Other detection methods, appropriate for the specific kind of marker conjugated to the ligand, can be also be employed to determine if the ligand has been specifically bound to a microorganism or
toxic protein. An example mentioned previously uses a fluorescent dye conjugated to a ligand coupled to detection of a microbe via fluorescence characteristic of the dye after (1) contact between the microbe and ligand and (2) washing away excess dye-conjugated ligand. It is important to note that if optical methods are used to detect the captured microbe or protein the tether should not be photocleavable.

PREOMICS EXHIBIT 1002
Page 323 of 1861

[0017] Thus, the method of the present invention does not depend on classical antigen-antibody recognition. On the contrary, the concepts of the present invention make use of relatively inexpensive reagents in the capture of microorganisms and microbial proteins contained in the sample.

[0018] In one embodiment of the invention, sensor chips (or beads) are employed. These chips should be formed from a suitable support material such as glass or plastic substrates (e.g., poly(propylene) or poly(vinyl acetate)) that will be compatible with both the chemistries used to conjugate the linker and ligand to the surface and the detection method employed. The sensor chip is formed of a patterned array defining a plurality of sections on the surface of the sensor chip, and each section has bonded thereto a different ligand capable of molecularly recognizing a specific microbial protein or microbial receptor, and hence the microbe itself. Microbial receptors would include, for example, proteins residing in the outer membrane of the microbial cell, pilus or flagellum, which is exposed to the aqueous environment surrounding the cell. The ligand for pathogen/protein capture bonded to the surface of the sensor chip can and should be varied. In general, such ligands may be characterized as heme compounds, siderophores, polysaccharides and anti-adhesion peptides capable of capturing a wide variety of microorganisms and toxic proteins. These ligands can thus be immobilized or bonded to the surface of the sensor chip through an appropriately sized cross-linker also having the capability of reacting with the ligands, whereby the coupling agent establishes a chemical tether between the surface of the sensor chip and the ligand capable of reaction with a variety of different microorganisms and proteins. The sensor chips and arrays (1) are exposed to a solution containing microorganisms or toxic proteins, (2) the non-binding constituents of the solution are removed, (3) followed by interrogation of the ligand-tethered surfaces to detect analyte binding. Analysis of the type or pattern of ligand-tethered surfaces found to have captured the microorganism(s), or microbial proteins not contained within intact microbial cells, can be used to taxonomically identify a microorganism or its toxic protein.

[0019] Thus, the present invention can be used rapidly to identify microorganisms without the need for growing a culture of the microorganism and then microscopically examining the culture thus produced. Likewise, low levels of toxic microbial proteins can similarly be identified. It is also unnecessary to employ enzymes or antibodies in the capture of microbial metabolites as is often used in the prior art. These, and other objects, features and advantages of the present invention will become apparent upon review of the following detailed descriptions of the disclosed embodiments and the appended claims.

BRIEF DESCRIPTION OF THE FIGURES

[0020]

Figure 1 shows the capture of *Salmonella typhimurium* on a glass microscope slide coated with tethered heme. Detection of this pathogenic bacterium was accomplished according to the method and apparatus outlined in U. S. Patent 5,968,766 to Powers.

Figure 2 shows the capture of *Enterobactor aerogenes* diluted in various concentrations of *Bacillus globigii* on a glass microscope slide coated with tethered heme. Detection of this pathogenic bacterium was accomplished according to the method and apparatus outlined in U. S. Patent 5,968,766 to Powers.

DETAILED DESCRIPTION OF THE INVENTION

[0021] The capture of a pathogenic bacterium (*Salmonella typhimurium*) with tethered heme, as outlined in the present invention, is shown in Figure 1. (The method and apparatus outlined in U. S. Patent 5,968,766 was employed for the detection of the captured bacteria. Although numerous compatible bacterial detection methods could have been employed, this method was used due to its ability to detect such small numbers of bacteria on the slide.) Inspection of the figure shows that the detection limit (<100 cells) of the captured microorganism using a tethered heme ligand is lower than that observed using immunological methods (*ca.* 400 cells under optimal conditions). Binding between the microorganism and the heme ligand is not as sensitive to pH, ionic strength and temperature as is binding to an antibody. The heme ligand is also less expensive, requires less careful storage and is not susceptible to proteolysis as are antibodies.

[0022] Figure 2 shows the tethered heme capture of a pathogenic bacterium (*Enterobactor aerogenes*) that has been diluted to the same concentration in solutions of a nonpathogen (*Bacillus globigii*). This figure shows that the tethered heme-coated slide is able to effectively capture the pathogenic bacteria from a solution even when the nonpathogen to pathogen ratio is 10^7:1. Detection of the captured bacteria was accomplished with the apparatus outlined in U. S. Patent 5, 968,766 to Powers.

[0023] In one embodiment of the present invention, a sample containing an unknown analyte microorganism or protein toxin is first contacted by the ligand. The ligand can be tethered to a surface of either a chip or bead. Binding efficiency is dependent upon the length of the tether. Microbes are found to bind most efficiently to ligands that are

PREOMICS EXHIBIT 1002
Page 324 of 1861

around forty Å long. Ligands directed to microbes are covalently attached to the substrate surface by tethers that are at least fifteen Å in length; ligands directed to proteinaceous toxins are at least six Å long. The analyte is then physically separated from the non-binding sample. Analytes captured by ligands tethered to a surface can be separated from non-binding components of the sample by simply washing the surface of the chip or bead. The surface of the substrate is then interrogated to determine if binding of the analyte to the ligand has occurred. The detection of bound microbes on the substrate surface can be made with: microscopy, intrinsic fluorescence, conjugate dye fluorescence, radioactivity, luminescence, phosphorescence, and/or optical absorbance. Identification of the microbe or protein is determined by the identity of the ligand. It is important to note that the tether should not be photocleavable or otherwise chemically labile in the solution used to wash the ligand-tethered surface.

[0024] In one embodiment of the invention, a sample containing an unknown analyte microorganism or protein is first contacted with the sensor chip. The sensor chip is formed of a substrate, such as glass, having a series of sections on the surface thereof. Each section has a different ligand bonded thereto, so as to be capable of binding to specific analytes. The ligands are capable of binding to the analyte for capture, and the presence of the captured analyte is detected using a fluorescence detection system, for example, disclosed and claimed in U.S. Patents 5,760,406 and 5,968,766 and via the intrinsic fluorescence of the proteinaceous toxins. Thus, the ligand of each of the sections of the sensor chip has the capability of capturing a specific microbial cell or microbial protein. The used chip can be saved and used to grow out the captured microorganisms as well.

[0025] In an alternate embodiment of the invention, a sample containing an unknown analyte (microorganism, proteinaceous toxin or other protein) is first contacted with a ligand conjugated to a marker, including, but not limited to, a fluorescent dye. The non-binding sample components and excess ligand are separated from the ligand-bound analyte; this separation can be accomplished by centrifugation (for cells), magnetic sedimentation or chromatography (for proteins). The detection of binding between the analyte and ligand, and thus taxonomic identification of the analyte, is accomplished by detection of the marker (e.g., fluorescence of the dye-conjugate in the example above).

[0026] In another embodiment of the invention, a sample containing an unknown analyte (microorganism or protein) is first contacted with a ligand tethered to a substrate surface with a linker of appropriate length, as noted above. Physical separation and washing remove non-binding components of the solution. As will be appreciated by those skilled in the art, the captured microorganism or protein can be treated with a reactive marker, provided the marker does not react with either the substrate surface or the ligands. Detection of the marker on the area of the surface associated with the ligand(s) that have been exposed to the analyte indicates the presence of a specific analyte.

[0027] In a preferred embodiment of the invention, the ligands used in the present invention may be taken from the group comprised of heme compounds, siderophores, polysaccharides (including oligosaccharides) and peptides.

[0028] As is also well known to those skilled in the art, animal pathogens generally possess heme uptake capability, and thus heme compounds can be used to capture a number of pathogenic species. In addition to heme compounds, other ligands in the form of high-affinity iron chelators, generally referred to as siderophores, can also be used to capture many strains of pathogenic bacteria. Included among such siderophores are alcaligin, mycobactins, pyochelin, staphyloferrin, vibriobactins and yersiniabactins.

[0029] As is also well known to those skilled in the art and as mentioned above, discrimination of animal pathogens by binding to heme compounds and siderophores that have been labeled with markers is also possible. An example would include the incubation of bacteria-containing solutions with a siderophore or heme compound that has been conjugated with a fluorescent, luminescent, phosphorescent, chemiluminescent, or radioactive compound. After washing the cells, detection of animal pathogens can be made by standard fluorescence, colorimetric or radiation detection techniques. The binding of animal pathogens to heme compounds and siderophores that are tethered to a support can also be exploited to separate these microbes from environmental samples, e.g., water, for the purpose of concentration and/or purification.

[0030] In addition to heme compounds and siderophores, eukaryotic surface epitopes (peptides or carbohydrates), which are recognized by microbial cell receptors, can likewise be used as ligands in the practice of the present invention. These ligands include naturally occurring oligosaccharides and polysaccharides as well as those available by chemical synthesis. Other oligosaccharides and their affinity to pathogens from various microorganisms are described by K. A. Karlsson "Microbial Recognition of Target Cell Glycoconjugates" (Structural Biology 5:622-635 (1995)).

[0031] The characteristics of a number of pathogenic bacterial organisms, including the disease caused by each species and their binding characteristics with siderophores, oligosaccharides and heme compounds are set forth in Table I. These characteristics can be used in the capture and identification of such species.

[0032] Peptide ligands can typically be identified by affinity panning of libraries of oligopeptides and then synthesized chemically. Siderophore ligands can be produced by chemical synthesis or isolation from spent microbial culture media. Oligosaccharide ligands can be produced by chemical synthesis or isolated from eukaryotic tissue. Heme compounds can be produced typically by chemical synthesis using protoporphyrin IX as a starting reagent.

PREOMICS EXHIBIT 1002
Page 325 of 1861

Table I

Bacterial Characteristics for Siderophore, Oligosaccharide and Hemin Binding					
Bacterial Species	**Disease Caused**	**Siderophore Binding?**	**Oligosaccharide Binding?**	**Hemin Binding?**	**Exotoxin Produced ?**
Bacillus anthracis	Anthrax	unknown	a pulmonary oligosaccharide	unknown	anthrax toxin
Bordetella pertusis	Whooping cough	Alcaligin, others	N-acetyl-glucosamine	Yes	pertussis toxin
Clostridium botulinum	Botulism	unknown	unknown	Yes	botulinum toxin A
Clostridium perfringens	Gas gangrene	unknown	unknown	unknown	α-toxin
Clostridium tetani	Tetnus	unknown	unknown	unknown	tetanus toxin
Corynebacterium diphthariae	Diphtheria	Aerobactin	unknown	unknown	diphtheria toxin
Escherichia coli 0157:H7	Numerous infections	many	Globobiose, others	Yes	Shiga-like toxin, others
Haemophilus influenzae	Meningitis	Enterobactin	GalNAcβ(1-4) Gal, others	Yes	unknown
Helicobacter pylori	Gastric ulcers	unknown	a mucosal oligosaccharide	unknown	vacuolating cytotoxin A
Klebsiella pneumoniae	Numerous infections	many	GalNAcβ(1-4) Gal, others	Yes	unknown
Mycobacterium tuberculosis	Tubercul osis	Mycobactin T	unknown	unknown	unknown
Neisseria meningitidis	Meningitis	many	unknown	Yes	unknown
Pseudomonas aeruginosa	Numerous infections	Pyochelin, others	Asialo G_{M1}, others	Yes	exotoxin A, elastase, others
Salmonella typhi	typhoid fever	many	unknown	Yes	Yes
Serratia marescens	numerous infections	Aerobactin , Ferrioxamine B	Yes	Yes	serralysin
Shigella dysenteriae	dysentery	Enterobactin, Aerobactin	Yes	Yes	Shiga toxin
Staphylococcus aureus	numerous infections	Staphylofe rrin, others	GalNAcβ(1-4) Gal	Yes	several superantigens
Streptococcus pneumoniae	pneumonia, meningitis	unknown	GlcNAcetyl(1-3) Gal, others	Yes	streptolysin O
Vibrio cholerae	cholera	Vibriobactin, others	Yes	Yes	cholera toxin
Yersinia pestis	bubonic plague	Yersiniabactin, others	unknown	Yes	YopE, others

PREOMICS EXHIBIT 1002
Page 326 of 1861

[0033] Toxins that contain at least one tryptophan or a few tyrosines per molecule can be detected by tryptophan/tyrosine fluorescence after capture using a tethered peptide. A variety of microbes, including algae, fungi, and bacteria, export exotoxins that are amenable to detection using this technology.

[0034] Table II contains examples of toxic, bacterial proteins that can be (1) captured using the technology described herein, and (2) ultimately detected via means of their intrinsic fluoresence. It is important to note that, for *Staphylococcus aureus* enterotoxin B, which represents the most unfavorable case in Table II (due to the presence of just one Trp and 22 Tyr), the following fluorescence study of the sole Trp residue has appeared: B. R. Singh, M. L. Evenson and M. S. Bergdahl "Structural Analysis of Staphylococcal Enterotoxins B and C1 Using Circular Dichroism and Fluorescence Spectroscopy" (Biochemistry 27: 8735-8741 (1988)). As is well known to those skilled in the art, detection of tryptophan/tyrosine fluorescence (normalized to the scattered excitation signal) is sufficient to indicate that spores, nonviable cells, viable vegetative bacterial or fungal cells, viruses, or a microbial toxin are present (i.e., bound to a ligand) on the surface of a sector of the sensor chip.

PREOMICS EXHIBIT 1002
Page 327 of 1861

Table II. Amino Acid Counts for Selected Bacterial Toxins

Bacterium	Toxin	No. Amino Acids	No. Trp	No. Tyr
B. anthracis	protective antigen	753	7	27
B. anthracis	lethal factor	770	5	35
B. cereus	phospholipase C	245	9	15
B. pertussis	pertussis toxin	952	11	50
C. botulinum	toxin A	1296	15	67
C. difficile	toxin A	2366	25	166
C. perfringens	iota-toxin	346	4	18
C. tetani	tetanus toxin	1421	13	78
C. diphtheriae	diphtheria toxin	534	5	16
E. coli	alpha-hemolysin	1023	3	38
H. pylori	vacuolating cytotoxin A	808	8	14
L. monocytogenes	listeriolysin O	523	7	23
P. aeruginosa	elastase	301	4	22

PREOMICS EXHIBIT 1002
Page 328 of 1861

S. marescens 19	serralysin	470	7
S. dysenteriae 17	Shiga toxin	638	7
S. aureus 22	enterotoxin B	239	1
S. aureus 9	toxic-shock toxin-1	194	3

[0035] Thus, as described above, a different ligand is tethered to each of the sections of the sensor chip. The sensor chip is then contacted with a sample containing unknown organisms or proteins, whereby specific ligands on the surface of the chip bind to specific analytes, selectively capturing them. The unbound analytes are then washed away with an appropriate solution (such as a phosphate-buffered saline); and the sensor chip is then subjected to an appropriate detection technique. One possible technique used to detect the presence of bacteria on the sections of the sensor chip is disclosed in U. S. Patents 5,706,406 and 5,968,766, wherein the described apparatus utilizes electromagnetic radiation of appropriate wavelengths to excite fluorescence characteristic of the presence of bound analytes.

[0036] As is well known to those skilled in the art, if a tethered ligand used to capture an analyte is itself fluorescent then this fluorescence may change upon binding to the analyte. (This change in fluorescence could be manifest as either a change in intensity or a shift of the characteristic fluorescence energy.) This change in the fluorescence of the tethered ligand can be used to confirm detection of the analyte.

[0037] In the presence of the present invention, a sample containing unknown microbes can be contacted with the sensor chip, whereby one or more receptors of the bacteria react with various different ligands tethered to the various sections of the chip. Then, the fluorescence of the chip can be measured with a probe for the purpose of detecting which of the sections of the sensor chip have analytes bonded thereto. As examples, mycobacterial siderophores can be used to capture mycobacteria such as *Mycobacterium tuberculosis. Helicobacter pylori* can be captured using tethered N-acetylneuroaminyl-alpha-2,3-galactose. The peptide:

GADRSYLSFIHLYPELAGAGGGC

can be tethered, by means of the terminal cysteine group to expressly capture free *Staphylococcus aureus* toxic-shock toxin-1. The peptide:

GHHKHHHGGGC

can be tethered also by means of the terminal cysteine group, to specifically capture the surface-exposed protein A of *Staphylococcus aureus,* and hence this organism itself. The *Staphylococcus aureus* toxic-shock syndrome toxin-1-binding peptide was described by A. Sato, *et al.* in "Identification from a Phage Display Library of Peptides that Bind to Toxic Shock Syndrome Toxin-1 and that Inhibit Its Binding to Major Histocompatibility Complex (MHC) Class II Molecules" (Biochemistry 35, 10441-10447 (1996)).

[0038] As indicated above, determining the presence of a single captured microorganism or discrete microbial protein can identify some of the analytes of interest. In other cases, however, a series of two or more captured analytes of interest is indicative of the identity of a particular analyte. As an example, consider a sensor chip having an area of three sections along the horizontal axis and three sections along the vertical axis as illustrated below:

A1	**A2**	**A3**
B1	**B2**	**B3**

PREOMICS EXHIBIT 1002
Page 329 of 1861

(continued)

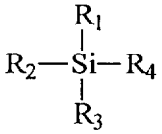
C1	C2	C3

[0039] In this example, the sections identified can be provided with the following ligands tethered to a specific section as set forth in the following table:

Section Location	3 x 3 Array Ligand
A1	asialo G_{M1}
A2	hemin
A3	pyochelin
B1	GalNAcβGal
B2	alcaligin
B3	fibronectin (peptide fragment)
C1	anti-*S. aureus* protein A peptide
C2	staphyloferrin
C3	ferrioxamine B

[0040] It has been found that *Pseudomonas aeruginosa* can be identified as the microorganism when analytes are detected in sections A1, A2, A3, B1 and C3. Similarly, *Staphylococcus aureus* can be identified when sections A2, B1, B3, C1, C2 and C3 contain analyte captured thereon. In this case, capture of an analyte in section C1 is sufficient for taxonomic identification. Capture of cells in sections A2, B1, B3, C2 and C3 reinforces the result. The incorporation of multiple ligands targeting a given analyte onto a sensor chip, in effect, permits multiple, independent analyses to be carried out using a single sample. This increases the statistical reliability of the analytical outcome.

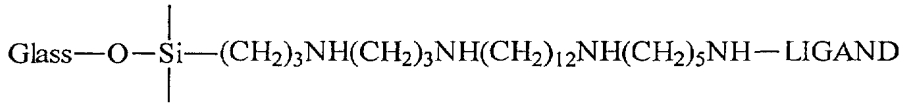
[0041] The various ligands are preferably tethered to a substrate by means of organic coupling agents which are themselves well known to those skilled in the art. When using a glass substrate, or a plastic whose surface that has been chemically oxidized to produce exposed hydroxyl groups, for the sensor chip, it is frequently preferred to employ, in the practice of the invention, organosilane compounds have the following general structure:



wherein R_1 through R_4 are each selected from the group consisting of hydrogen, alkyl groups containing 1 to 6 carbon atoms, aryl groups containing 6-12 carbons and alkoxy groups containing 1 to 4 carbon atoms, with at least one of R_1, R_2 and R_3 being an alkoxy group. R_4 is an organic group containing at least three carbon atoms and also containing a long linker bearing a functional group capable of reaction with the ligand. Without limiting the invention, suitable organic groups comprising the linker include polyamines, polyethers and poly(glycine). Also suitable for use in the practice of the invention are coupling agents containing other functional groups such as epoxy groups, amino groups and unsaturated functional groups, hydroxyl groups, thiol groups and the like, which are capable of reaction with the various ligands. Without limiting the invention as to theory, it is believed that the ligand reacts with the functional group, preferably a terminal functional group on the organosilane compound while the readily hydrolyzable alkoxy group attached directly to the silicon atom has the capability of reacting directly with the surface of the glass or plastic substrate of the sensor chips. This coupling agent (extended silane) may be alternatively constructed in situ by first reacting the parent silane with the surface of the sensor chip, followed by the chemical reactions needed to attach the linker to the immobilized silane. The ligand is then tethered to the surface of the glass or plastic through the coupling agent (i.e., silane bearing the organic linker). Further, the linker should be of sufficient length to present the ligand at the optimal distance (40 Ångstroms) from the surface of the chip. This observation is based on our determination that shorter distances results in decreased bacterial cell capture efficiency.

[0042] Thus the ligand tethered to the glass surface may be illustrated by the following:

PREOMICS EXHIBIT 1002
Page 330 of 1861

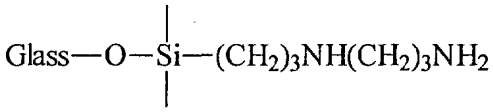


The ligand tethered to an oxidized plastic surface may be illustrated as above by replacing the "Glass-O-Si" moiety with C (carbon from the plastic polymer). The chemical reactions used in tethering ligands to the surface of the sensor chip are known to those skilled in the art and are described in the literature. Such reactions may be found in G. T. Hermanson Bioconjugate Techniques (San Diego: Academic Press, 1966); Hansson et al., "Carbohydrate-Specific Adhesion of Bacteria to Thin Layer Chromatograms: A Rationalized Approach to the Study of Host Cell Glycolipid Receptors" (Analytical Biochemistry 146: 158-163 (1985)); and, Nilsson et al., "A Carbohydrate Biosensor Surface for the Detection of Uropathogenic Bacteria" (Bio/Technology 12: 1376-1378 (December 1994)).

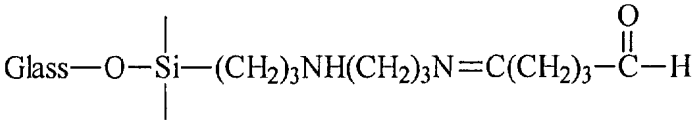
[0043] Illustrative of such reactions are those used to tether ferroxamine as a ligand to the surface of a glass sensor chip. In the first stage, a glass surface containing free hydroxyl groups is first reacted with a 2% solution of gamma-N-(aminopropyl)-gamma-aminopropyltrimethoxysilane to attach the silane to the glass surface:



(I)

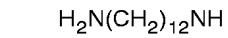


The product of that reaction can then be reacted with glutaraldehyde at a pH of about 8 to form the corresponding aldehyde:

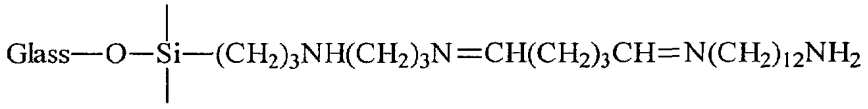


(II)

The aldehyde, in turn, can be reacted with a diamine (III) to yield IV:



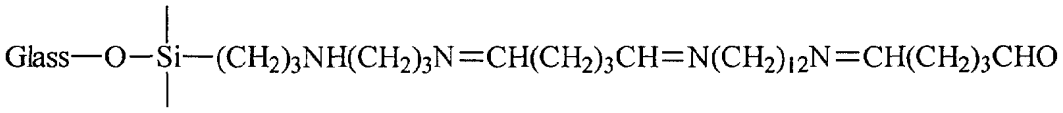
(III)



(IV)

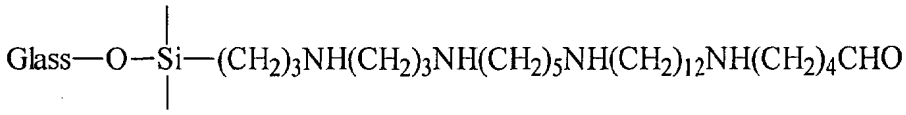
[0044] Next, the product of the preceding reaction is reacted with glutaraldehyde to introduce a (terminal) aldehyde group:

PREOMICS EXHIBIT 1002
Page 331 of 1861



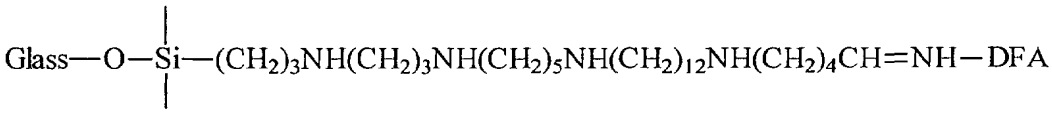
(V)

which can then be reduced using $NaCNBH_3$ to yield:



(VI)

[0045] The foregoing silane coupling agent bonded to the surface can then be derivatized by reaction with deferrioxamine B (or DFA) at an alkaline pH to yield:



(VII)

The DFA can then be complexed with Fe by reaction with a ferrous salt in aqueous medium to form the ligand.

[0046] It will be understood that various changes and modifications can be made in the determination, procedure, and formulation without departing from the spirit of the invention, especially as defined in the following claims:

Claims

1. A method for taxonomic identification of a biological analyte comprising:

(a) exposing a solution containing the analyte to a ligand specific for the analyte of interest that has been covalently tethered to a substrate surface with a photostable linker at a distance of at least 15Å for the capture of microorganisms;
(b) separating the bound analyte from the non-binding components of the solution containing the analyte by physical separation, washing or both; and
(c) interrogation of the ligand-tethered substrate surface for analyte binding.

2. The method of claim 1, wherein the biological analyte is selected from the group comprised of:

(a) bacteria;
(b) viruses;
(c) rickettsiae;
(d) protozoa; and
(e) fungi.

3. The method of claim 1, wherein the ligand is at least one of a heme compound,a siderophore,a polysaccharide, a peptide specific for an outer membrane protein, and/or a peptide specific for a conjugated lipid.
The method of claim 1, wherein the detection of the captured analyte is accomplished trough at least one of microscopy the intrinsic fluorescence of the target,
the fluorescence of a reactive dye conjugate exposed to the sample before capture of the analyte by the tethered ligand surface,
the fluorescence of a reactive dye conjugate exposed to the sample after capture of the analyte by the tethered

PREOMICS EXHIBIT 1002
Page 332 of 1861

ligand surface,
the radioactivity of a reactive compound exposed to the sample before capture of the analyte by the tethered ligand surface,
the radioactivity of a reactive compound exposed to the sample after capture by the tethered ligand surface,
the luminescence of a reactive dye conjugate exposed to the sample before capture of the analyte by the tethered ligand surface,
the luminescence of a reactive dye conjugate exposed to the sample after capture of the analyte by the tethered ligand surface,
the phosphorescence of a reactive dye conjugate exposed to the sample before capture of the analyte by the tethered ligand surface,
the phosphorescence of a reactive dye conjugate exposed to the sample after capture of the analyte by the tethered ligand surface,
the optical absorbance of a reactive dye conjugate exposed to the sample before capture of the analyte by the tethered ligand surface,
the optical absorbance of a reactive dye conjugate exposed to the sample after capture of the analyte by the tethered ligand surface, and/or
the fluorescent quenching of the fluorescent tethered ligand surface upon binding of the analyte.

5. A method for taxonomic identification of a biological analyte comprising:

(a) exposing a solution containing the analyte to a ligand specific for the analyte of interest that has been covalently tethered to a substrate surface with a photostable linker at a distance of at least six Å for the capture of proteins;
(b) separating the bound analyte from the non-binding components of the solution containing the analyte by physical separation, washing or both; and
(c) interrogation of the ligand-tethered substrate surface for analyte binding.

6. The method of claim 5, wherein the biological analyte is selected fromthe group comprised of:

(a) proteinaceous toxins; and
(b) cytosolic proteins.

7. The method of claim 5, wherein the ligand is at least one of the group consisting of a peptide, usually three to twenty amino acids long, specific for a proteinaceous toxin,a peptide, usually three to twenty amino acids long, specific for a proteinaceous hormone, and/or a peptide, usually three to twenty amino acids long, specific for a cytosolic protein.

8. The method of claim 5, wherein the detection of the captured analyte is accomplished through at least one of the intrinsic fluorescence of the protein, the fluorescence of a reactive dye conjugate exposed to the protein before capture of the analyte by the tethered ligand surface,
the fluorescence of a reactive dye conjugate exposed to the protein after capture of the analyte by the tethered ligand surface,
the radioactivity of a reactive compound exposed to the protein before capture of the analyte by the tethered ligand surface,
the radioactivity of a reactive compound exposed to the protein after capture by the tethered ligand surface,
the luminescence of a reactive dye conjugate exposed to the protein before capture of the analyte by the tethered ligand surface,
the luminescence of a reactive dye conjugate exposed to the protein after capture of the analyte by the tethered ligand surface,
the phosphorescence of a reactive dye conjugate exposed to the protein before capture of the analyte by the tethered ligand surface,
the phosphorescence of a reactive dye conjugate exposed to the protein after capture of the analyte by the tethered ligand surface,
the optical absorbance of a reactive dye conjugate exposed to the protein before capture of the analyte by the tethered ligand surface,
the optical absorbance of a reactive dye conjugate exposed to the sample after capture of the analyte by the tethered ligand surface, and/or
the fluorescent quenching of the fluorescent tethered ligand surface upon binding of the protein.

PREOMICS EXHIBIT 1002
Page 333 of 1861

9. A method for taxonomic identification of a biological analyte comprising:

(a) exposing the solution containing the analyte to a ligand specific for the analyte of interest that has been conjugated to a marker;
(b) separating the bound analyte from the excess marker-conjugated ligands;
(c) interrogation of the analyte for ligand binding via detection of the conjugated marker.

10. The method of claim 9, wherein the biological analyte is selected from the group comprised of:

(a) bacteria;
(b) viruses;
(c) proteinaceous toxin;
(d) rickettsiae;
(e) protozoa;
(f) fungi; and
(g) cytosolic protein.

11. The method of claim 9, wherein the separation of the bound analyte from the excess conjugated ligand is accomplished by chromatography. 12.The method of claim 9, wherein the ligand is conjugated to a magnetic particle and the separation of the bound analyte from the non-binding components of the analyte solution is accomplished by magnetic separation with the ligand being tethered to the magnetic particle by at least fifteen Å for capture of microorganisms.

13. The method of claim 9, wherein the ligand is at least one of a heme compound, a siderophore, a polysaccharide, a peptide specific for an outer membrane protein, and/or a peptide specific for a conjugated lipid. 14.The method of claim 9, wherein the marker is at least one of the following:

the marker is fluorescent and the detection is via fluorescence,
the marker is luminescent and the detection is via luminescence,
the marker is radioactive and the detection is via radioactivity, and/or the marker is phosphorescent and the detection is via phosphorescence; 15.A method for taxonomic identification of a biological analyte comprising:

(a) exposing a solution containing the analyte to an array of different ligands that have been covalently tethered to a substrate surface;
(b) separating the bound analyte on the ligand array from the non-binding components of the solution by physical separation, washing or both; and
(c) interrogation of the ligand-tethered substrate surface for analyte binding.

16. The method of claim 15, wherein the ligands utilized in the array are tethered with a photostable linker at a distance of at least fifteen Å from the substrate surface for the capture of microorganisms, and/or for the capture of proteinaceous toxins.

17. The method of claim 15, wherein the biological analyte is selected from the group comprising:

(a) bacteria;
(b) viruses;

(a) proteinaceous toxin;
(b) rickettsiae;
(c) protozoa;
(d) fungi; and
(e) cytosolic protein.

18. The method of claim 15, wherein a ligand is a heme compound.

19. The method of claim 15, wherein one or more of the ligands is at least one of a siderophore, a polysaccharide, a peptide specific for an outer membrane protein, and/or a peptide specific for a conjugated lipid.

PREOMICS EXHIBIT 1002
Page 334 of 1861

20. The method of claim 15, wherein the detection of a captured microorganism is accomplished through microscopy.

21. The method of claim 15, wherein the detection of the captured analyte is accomplished through at least one of the following
the intrinsic fluorescence of the target,
the fluorescence of a reactive dye conjugate exposed to the sample before capture of the analyte by the tethered ligand surface,
the fluorescence of a reactive dye conjugate exposed to the sample after capture of the analyte by the tethered ligand surface,
the radioactivity of a reactive compound exposed to the sample before capture of the analyte by the tethered ligand surface,
the radioactivity of a reactive compound exposed to the sample after capture by the tethered ligand surface,
the luminescence of a reactive dye conjugate exposed to the sample before capture of the analyte by the tethered ligand surface,
the luminescence of a reactive dye conjugate exposed to the sample after capture of the analyte by the tethered ligand surface.
the phosphorescence of a reactive dye conjugate exposed to the sample before capture of the analyte by the tethered ligand surface,
the phosphorescence of a reactive dye conjugate exposed to the sample after capture of the analyte by the tethered ligand surface,
the optical absorbance of a reactive dye conjugate exposed to the sample before capture of the analyte by the tethered ligand surface,
the optical absorbance of a reactive dye conjugate exposed to the sample after capture of the analyte by the tethered ligand surface, and/or
the fluorescent tethered ligand surface upon binding of the analyte.

22. A method for identification of a protein comprising:

(a) exposing the solution containing microorganisms to conditions resulting in the rupture of the microorganism and spilling of the contents into the solution selected from the group comprising chemical treatment, holins, enzymatic treatment, freeze-thaw cycling, bacteriophage infection, and physical treatment.
(b) exposing a solution containing the protein analyte to a ligand specific for the analyte of interest that has been covalently tethered with a photostable linker to a substrate surface;
(c) separating the bound analyte from the non-binding components of the solution containing the analyte by physical separation; and
(d) interrogating the ligand-tethered substrate surface for analyte binding wherein the ligand is tethered at a distance of at least six Å from the substrate surface for the capture of proteins.

23. The method of claim 22, wherein the ligand is a peptide specific for the protein of interest.

24. The method of claim 22, wherein the detection is via the intrinsic fluorescence of the captured protein.

25. A method for capture of a biological analyte onto a substrate wherein:

(a) the ligands used are selected from the group consisting of: heme compounds, siderophores, polysaccharides, and peptides specific for outer membrane proteins, conjugated lipids, and microbial protein targets;
(b) the ligand is tethered at a distance of at least fifteen Å from the substrate surface for the capture of microorganisms;
(d) the substrate is suitable for a column packing;
(e) the biological analytes are selected from the groups of: bacteria, viruses, rickettsiae, protozoa, and fungi.
(f) the capture of the biological analytes on the substrate is used for the concentration of the biological analyte from water samples; and
(g) the capture of the biological analytes on the substrate is used for the concentration of the biological analyte from medical samples.

26. A method for tethering a ligand to a substrate surface via a linker comprising:

PREOMICS EXHIBIT 1002
Page 335 of 1861

(a) reacting of a silane compound [$Si(R_1R_2R_3R_4)$] with an oxidized (hydroxylated) substrate surface wherein one of the substituents contains a linker bearing a functional group capable of reacting with the ligand; and
(b) reacting the ligand with the functional group of the covalently attached silane linker; wherein the linker is chosen so that the ligand is tethered preferably at least 15Å away from the substrate surface.

27. The method in claim 26, wherein R1 through R4 are each selected from the group consisting of hydrogen, alkyl groups containing 1 to 6 carbons, aryl groups containing 6-12 carbons and alkoxy groups containing 1 to 4 carbons, with at least one of R_1, R_2 and/or R_3 being an alkoxy group.

28. The method in claim 26, wherein R_4 is an organic group containing at least three carbon atoms and also containing a long linker bearing a functional group capable of reacting with the ligand.

29. The method in claim 26, wherein the organic groups comprising the linker include polyamines, polyethers, and poly(glycine).

30. The method in claim 26, wherein the coupling agent functional group of the linker (R_4) includes epoxy groups, amino groups, unsaturated functional groups, hydroxyl groups, and thiol groups.

31. The method in claim 26, wherein the ligand reacts with the functional group, preferably a terminal functional group on the organosilane compound while the readily hydrolyzable alkoxy group attached directly to the silicon atom has the capablilty of reacting directly with the surface of the surface substrate.

32. The method in claim 26, wherein the ligand tether is constructed in situ by first reacting the parent silane [$Si(R_1R_2R_3R_4)$] with the ligand followed by reaction of the alkoxysilane with the substrate surface.

PREOMICS EXHIBIT 1002
Page 336 of 1861

FIGURE 1

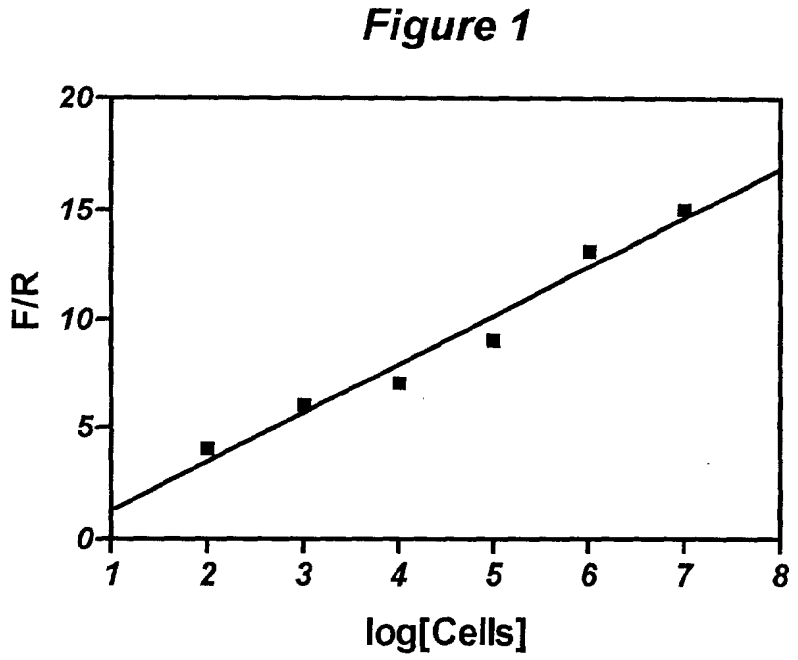
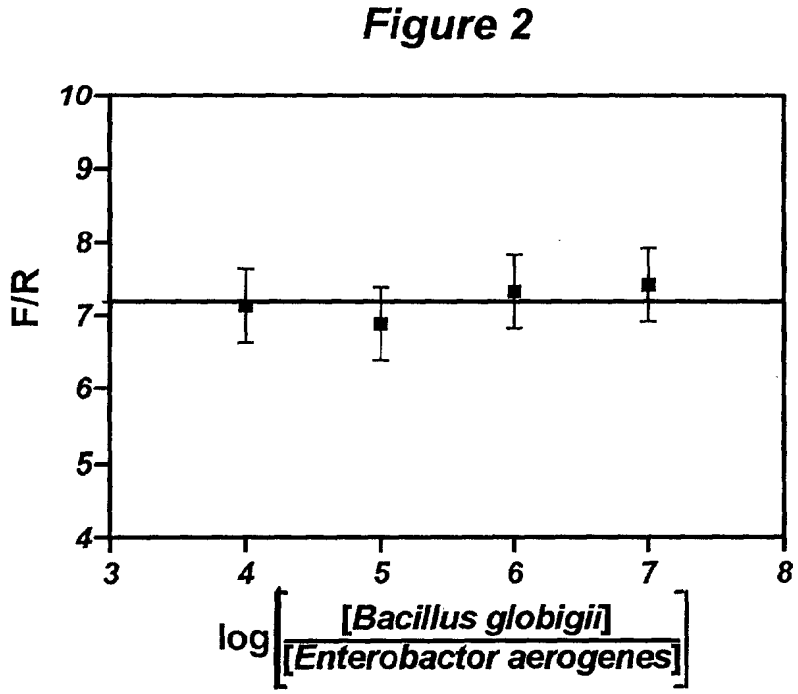


FIGURE 2



PREOMICS EXHIBIT 1002
Page 337 of 1861

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

(11) EP 2 209 893 B1

(12) **EUROPEAN PATENT SPECIFICATION**

(45) Date of publication and mention of the grant of the patent:
20.11.2013 Bulletin 2013/47

(21) Application number: **08837175.2**

(22) Date of filing: **10.10.2008**

(51) Int Cl.:
C12N 15/11 *(2006.01)* ***G01N 33/68*** *(2006.01)*

(86) International application number:
PCT/GB2008/003447

(87) International publication number:
WO 2009/047526 (16.04.2009 Gazette 2009/16)

(54) **USE OF APTAMERS IN PROTEOMICS**

VERWENDUNG VON APTAMEREN IN DER PROTEOMIK

UTILISATION D'APTAMÈRES EN PROTÉOMIQUE

(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 MT NL NO PL PT RO SE SI SK TR**

(30) Priority: **12.10.2007 EP 07020049**

(43) Date of publication of application:
28.07.2010 Bulletin 2010/30

(73) Proprietor: **Pronota NV**
9052 Zwijnaarde Ghent (BE)

(72) Inventors:
- **BROWN, Clive, Gavin**
  Cambridge CB3 6DS (GB)
- **KAS, Koen**
  2970 Schilde (BE)
- **EYCKERMAN, Sven, Agnes, Jan**
  9840 De Pinte (BE)

(74) Representative: **Boxall, Sarah Jane**
Boxall IPM Ltd
9 King Street
Sandwich Kent CT13 9BT (GB)

(56) References cited:
WO-A-00/06770 WO-A-2005/003291

- **CHO ET AL: "Optimization of aptamer microarray technology for multiple protein targets" ANALYTICA CHIMICA ACTA, vol. 564, no. 1, 30 March 2006 (2006-03-30), pages 82-90, XP005351637 ISSN: 0003-2670**
- **ELRICK MOLLISA M ET AL: "Proteomics: recent applications and new technologies." BASIC & CLINICAL PHARMACOLOGY & TOXICOLOGY, vol. 98, no. 5, May 2006 (2006-05), pages 432-441, XP002514963 ISSN: 1742-7835**
- **LEE JEAN W ET AL: "Biomarker assay translation from discovery to clinical studies in cancer drug development: quantification of emerging protein biomarkers" ADVANCES IN CANCER RESEARCH, vol. 96, 1 January 2007 (2007-01-01), pages 269-298, XP009106017 ISSN: 0065-230X**
- **DANIEL BRANTON ET AL: "The potential and challenges of nanopore sequencing" NATURE BIOTECHNOLOGY, vol. 26, no. 10, 1 January 2008 (2008-01-01), pages 1146-1153, XP002507807**

Note: Within nine months of the publication of the mention of the grant of the European patent in the European Patent Bulletin, any person may give notice to the European Patent Office of opposition to that patent, in accordance with the Implementing Regulations. Notice of opposition shall not be deemed to have been filed until the opposition fee has been paid. (Art. 99(1) European Patent Convention).

Printed by Jouve, 75001 PARIS (FR)

PREOMICS EXHIBIT 1002
Page 338 of 1861

Description

[0001] The present invention relates to the identification and quantification of proteins in a complex mixture. In particular, the invention relates to the use of aptamers as tags or proxies for proteins and protein quantities in a complex mixture. In another aspect the invention relates to the use of next generation polynucleotide sequencing methods in proteomics.

[0002] The proteome is usually described as the entire complement of proteins found in a biological system, such as, *e.g.*, a cell, tissue, body fluid, organ or organism. The study of naturally occurring proteins is generally termed 'proteomics' and encompasses study of the proteome expressed at particular times and/or under internal or external conditions of interest. Proteomics approaches frequently aim at global analysis of the proteome and require that large numbers of proteins, *e.g.*, hundreds or thousands, can be routinely resolved, identified and quantified from a single or multiple sample(s).

[0003] Among the promises of proteomics is its ability to recognise new biomarkers, *i.e.*, proteins as biological indicators that signal a changed physiological state, for example due to a disease or a therapeutic intervention. Biomarker discovery usually involves comparing proteomes expressed in distinct physiological states and identifying proteins whose occurrence or expression levels consistently differ between the physiological states (Schrattenholz A, Groebe K. Electrophoresis. (2007) Jun 28(12) 1970-9).

[0004] Proteins in blood are a particular target for the identification of markers of disease states and drug treatments. It is widely assumed that the amounts and/or conformation of proteins in the blood should be statistically related to such states in a manner that outweighs intrinsic natural variability. Blood and other body fluids are a particular target because they bathe affected tissues, transport vital proteins and can be obtained for testing using relatively cheap and straightforward procedures during a medical consultation.

[0005] However, proteins in blood have a very large range of concentrations, with a small number of proteins accounting for over 99.9% of all proteins and the rest occupying a distribution from picogram to milligram per millilitre (Qian W.J. et al., Mol. Cell. Prot. (2006) 5(10) 1727-1744). Due to the limitations of existing proteomics techniques, this abundance range remains hypothetical. Proteomics scientists have employed a variety of methods to reach into this range, with the aim of also minimising disruption to the relative abundances of the proteins. This often requires the exclusion of the high-abundance proteins by selective purification. Attempts have also been made to reduce the complexity of the peptides obtained by selectively focusing in on subsets of all peptides in the sample. These procedures are lengthy and the reproducibility between samples, replicates, machines and laboratories has yet to be demonstrated in a manner that would be a pre-requisite for the statistical discovery of biomarker proteins.

[0006] Conventional Mass Spectrometry (MS)-based proteomics, commonly used in biomarker discovery, proceeds by separating biological samples to isolate single proteins from the mixture under investigation. More recently, this has advanced from 2D-gels to multi-dimensional column-based high performance liquid chromatography (HPLC). Proteins can be broken down into shorter subunits or peptides. Isolated peptides are then fed into a mass spectrometer that ionises the peptides and breaks them up further, yielding a ladder of mass/charge measurements. These measurements and their abundances can also be quantified under a variety of schemes, usually relative to some control. The resulting ladder of spectra is then interpreted against known peptide sequences or blindly from raw data and the obtained mass and sequence information is used to search sequence databases to identify the proteins from which the respective peptides originated.

[0007] However, proteolysis of complex biological samples usually produces hundreds of thousands of peptides which may overwhelm the resolution capacity of known chromatographic and MS systems, causing incomplete resolution and impaired identification of the constituent peptides. Typically in MS-based proteomics, as many as 80% of the spectra derived from the sample cannot be accurately or consistently re-interpreted into discriminatory peptides or thence proteins. Their fragmentation behaviour and abundance in the MS process can also be context dependent, further complicating reproducibility (Liu H, Sadygov RG, Yates JR 3rd., Anal. Chem. (2004) Jul 15 76(14) 4193-201).

[0008] One method to enable proteomic analysis of biological samples is to reduce the complexity of peptide mixtures generated by separation of such samples, before subjecting said peptide mixtures to downstream resolving and identification steps, such as chromatographic separation and/or Mass Spectrometry (MS). Ideally, reducing the complexity of protein peptide mixtures will decrease the average number of distinct peptides present per individual protein of the sample, yet will maximise the fraction of proteins of the sample actually represented in the peptide mixture.

[0009] The use of blood (serum or plasma) is further obscured by biological processing of proteins in a variety of ways that confound MS-based ascertainment (Qian W.J. et al., Mol. Cell. Prot. (2006) 5(10) 1727-1744). Recent studies have shown wildly contradictory attempts to identify and count proteins from clinical samples. The very act of isolating, fragmenting and measuring proteins in conventional MS-based proteomics alters their relative abundance and chemical makeup.

[0010] The output of proteomics-based analysis is a "hit list" of proteins that are significantly correlated in the samples under study. Typically, this list is a selection of proteins of variable statistical significance. A biologically oriented study of the list is usually made and a rational choice is made about the potential biological significance

PREOMICS EXHIBIT 1002
Page 339 of 1861

of each member. The process of reaching a list of hypothetical biomarkers, or proteins of putative statistical significance, is generally called 'Discovery'.

[0011] Efforts are then focused on verifying and validating a small number of chosen "hits". This involves confirmation of the measured protein abundances in a broader population of clinical samples with the objective being to show that the discovered and chosen proteins are genuine and not false positives and that they are specific to the disease or drug state. This process of confirming the quantitative significance of a proteomics measurement and putative biomarker in a more generalised population is usually referred to as 'Validation' (Zolg, Mol. Cell. Prot. (2006) 5(10), 1720-1726). Often this requires the use of an alternative technology to the discovery phase.

[0012] Typically, this is achieved using antibodies raised to the purified putative proteins. These antibodies can then be used in an ELISA-based assay against hundreds of samples, the relevant statistics reapplied and the validity of the protein as a biomarker established. The ultimate intention of this process is also for the antibody to become-part of a clinical assay. However, antibodies are time-consuming and costly to produce and it is not always possible to raise an antibody specific to a given protein. A further complication is the presence of alternative variants of the protein or isoforms that may have crusts of attached sugar molecules and other modifications on their surface. Isoforms may not be evident in the Discovery phase MS based identification and thus complicate the statistical significance of the measurements as well as the raising of specific antibodies (Zolg, Mol. Cell. Prot. (2006) 5(10), 1720-1726; Rifai et al., Nature Biotech (2006) 24(8) 971-83).

[0013] Both Discovery and Validation have high failure rates and are time consuming and costly exercises. Success and failure can usually only be judged after the Validation stage after much time and financial expenditure. Therefore, recent efforts have focused on the refinement of MS-based approaches for discovery and on the bulk generation of antibodies in an effort both to accelerate Discovery, Validation and to derive a clinical product (Anderson and Hunter, Mol. Cell Prot. (2006) 5(4) 573-588; Zangar et al., Expert Review of Proteomics (2006) 3(1) 37-44).

[0014] Accordingly, there is a clear need for new methods to interrogate proteins, particularly complex mixtures of proteins, derived from body fluids. Specifically, there is a need for a technology to identify and measure the abundance of proteins, especially abundance of multiple proteins in a complex mixture, that overcomes the technical drawbacks and scientific and cost constraints that hampers current technology. In particular, tools are needed to replace the costly and sometimes unreliable generation and use of antibodies to identify and validate target proteins. Such tools can also be used to supplant existing Discovery based technologies such as MS, provided a sufficient diversity of antibody alternatives can be generated quickly and cheaply enough, to probe the full range of native proteins in a biological sample.

[0015] A clear desiderata for such a tool is minimal manipulation of the mixture so that a true representation of the proteins in their native configuration, together with any natural modification and variations, in the sample may be obtained. Such a tool should also be highly and accurately reproducible.

[0016] Aptamers are short polymers, usually nucleic acids (DNA, RNA, PNA), that form well-defined three dimensional shapes, allowing them to bind target molecules in a manner that is conceptually similar to antibodies. Aptamers combine the optimal characteristics of small molecules and antibodies, including high specificity and affinity, chemical stability, low immunogenicity, and the ability to target protein-protein interactions. In addition to high specificity, aptamers have very high affinities to their targets. Typically, aptamers generated against proteins have affinities in the picomolar to low nanomolar range. In contrast to monoclonal antibodies, aptamers are chemically synthesised, rather than biologically expressed, offering a significant cost advantage.

[0017] While aptamers provide a useful and effective alternative to antibodies, there still remains a problem on how to quantify proteins, via aptamers. Complex mixtures of proteins and the large dynamic range of proteins present in biological samples provides a particular problem. Typically, aptamers are quantified using radiolabels. While quantification of radiolabels is acceptable, and has been for a number of years, there remains a need to improve and expedite quantification, as well as improve accuracy and reproducibility.

[0018] Thus, in its broadest sense, the present invention relates to a method for measuring the amount of at least one molecule in a biological sample, the method comprising:

a) combining the sample with one or more aptamers and allowing one or more molecules in the sample to bind to the aptamer(s);
b) separating bound from unbound molecules; and
c) quantifying the molecule(s) bound to the or each aptamer,

wherein quantification of the bound molecule(s) is carried out by sequencing at least part of the or each aptamer.

[0019] Expressed in another way, the invention resides in the use of an aptamer as a tag to quantify a protein, wherein the sequence of the aptamer is the tag. It has been appreciated that an aptamer has a unique sequence that may be read, for example, like a barcode. According to the present invention, it is this unique barcode that is the tag for each protein to which the barcode (aptamer) binds. By reading the barcode (sequencing the aptamer), the need for additional tags and/or labels is obviated.

[0020] In contrast with existing methods, use of the aptamer sequence as a tag allows identification and quantification of a protein using only a single molecular

PREOMICS EXHIBIT 1002
Page 340 of 1861

entity. Indeed, sequencing of an aptamer allows counting of the instances or occurrences of each distinct sequence and so allows direct quantification of the protein to which each aptamer has bound. In other words, each aptamer is a direct representation of a protein within a sample or protein mixture when an individual aptamer binds and individual protein. It will be appreciated that, by way of the method of the present invention, manipulation of the original protein or sample is minimised, thereby allowing interrogation of a protein or a mixture of proteins in as near to a natural state as is possible. In this way, a true picture of a protein or population of proteins is obtainable.

[0021] In a preferred embodiment, sequencing of the aptamers is achieved without amplification so an absolute or digital count of the aptamers by counting the occurrence of each species on a sequencing device that separates out each individual molecule prior to sequencing, for example a 'next-generation DNA sequencer. A direct count of the occurrence of each sequence is obtained which, in turn, equates to an absolute count of the proteins to which the aptamers have bound. A subset of the sequences may be represented on the counting device which should be a representative sample of the occurrence of the given sequences in the population of molecules under study. This, again, is in contrast to current aptamer quantification methods which typically require amplification of bound tag or amplification of the aptamer molecule yielding an analogue estimate of its relative abundance (as described in Patent Publication No. US 2007-0166741). This is due to the limitations of using labels and additional tags by which the aptamers that have bound proteins are identified. As a result, quantification is proportional to the actual abundance, *i.e.* an analogue signal, and absolute or individual occurrences cannot be resolved.

[0022] A further advantage of using the aptamer sequence as a tag is that the sequences provide a huge range of possible tags. For example, a sequence of four bases offers 256 different combinations and thus 256 possible unique identifiers or tags. Thus, while the entire aptamer sequence may be obtained, the number of nucleotides sequenced needs only to be sufficient to identify a particular aptamer. In other words, either the whole aptamer sequence is the tag or the tag is part of the aptamer sequence.

[0023] Although the method of the present invention may be used to quantify a single protein in a sample with one aptamer sequence, a further advantage of the present invention is that the range of possible tags provided by unique sequences allows many proteins to be quantified, at the same time, in the same sample. Thus, a panel or library of known aptamer sequences may be used to quantify multiple proteins in a complex mixture and, in this way, the invention allows interrogation of complex biological samples. Furthermore, a panel or library of aptamers may be targeted to specific proteins, in particular, proteins that are known to be present in a sample in low quantities. Since it is rarely possible to identify and quantify low abundance proteins in a sample using known methods, the present invention provides a solution for interrogating these small and, so far, elusive populations. The library or panel should ideally contain an excess of aptamer species for probing samples.

[0024] The invention may analyse a single protein (*e.g.*, gel-excised protein) and is particularly suitable for analysing mixtures of proteins, including complex protein mixtures. The terms "mixture of proteins" or "protein mixture" generally refer to a mixture of two or more different proteins, *e.g.*, a composition comprising two or more different proteins.

[0025] In preferred embodiments, a mixture of proteins to be analysed may include more than about 10, preferably more than about 50, even more preferably more than about 100, yet more preferably more than about 500 different proteins, such as, *e.g.*, more than about 1000 or more than about 5000 different proteins. An exemplary complex protein mixture may involve, without limitation, all or a fraction of proteins present in a biological sample or part thereof.

[0026] The terms "biological sample" or "sample" as used herein generally refer to material, in a non-purified or purified form, obtained from a biological source. By means of example and not limitation, samples may be obtained from: viruses, *e.g.*, viruses of prokaryotic or eukaryotic hosts; prokaryotic cells, *e.g.*, bacteria or archeae, *e.g.*, free-living or planktonic prokaryotes or colonies or bio-films comprising prokaryotes; eukaryotic cells or organelles thereof, including eukaryotic cells obtained from *in vivo* or *in situ* or cultured *in vitro*; eukaryotic tissues or organisms, *e.g.*, cell-containing or cell-free samples from eukaryotic tissues or organisms; eukaryotes may comprise protists, *e.g.*, protozoa or algae, fungi, *e.g.*, yeasts or molds, plants and animals, *e.g.*, mammals, humans or non-human mammals. "Biological sample" may thus encompass, for instance, a cell, tissue, organism, or extracts thereof. A biological sample may be preferably removed from its biological source, *e.g.*, from an animal such as mammal, human or non-human mammal, by suitable methods, such as, without limitation, collection or drawing of urine, saliva, sputum, semen, milk, mucus, sweat, faeces, *etc.*, drawing of blood, cerebrospinal fluid, interstitial fluid, optic fluid (vitreous) or synovial fluid, or by tissue biopsy, resection, *etc.* A biological sample may be further subdivided to isolate or enrich for parts thereof to be used for obtaining proteins for analysing in the invention. By means of example and not limitation, diverse tissue types may be separated from each other; specific cell types or cell phenotypes may be isolated from a sample, *e.g.*, using FACS sorting, antibody panning, laser-capture dissection, *etc.*; cells may be separated from interstitial fluid, *e.g.*, blood cells may be separated from blood plasma or serum; or the like. The sample may be applied to the method of the invention directly or may be processed, extracted or purified to varying degrees before being used.

[0027] The sample may be derived from a healthy sub-

PREOMICS EXHIBIT 1002
Page 341 of 1861

ject or a subject suffering from a condition, disorder, disease or infection. For example, without limitation, the subject may be a healthy animal, *e.g.*, human or non-human mammal, or an animal, *e.g.*, human or non-human mammal, that has cancer, an inflammatory disease, autoimmune disease, metabolic disease, CNS disease, ocular disease, cardiac disease, pulmonary disease, hepatic disease, gastrointestinal disease, neurodegenerative disease, genetic disease, infectious disease or viral infection, or other ailment(s).

[0028] Preferably, protein mixtures derived from biological samples may be treated to deplete highly abundant proteins there from, in order to increase the sensitivity and performance of proteome analyses. By means of example, mammalian samples such as human serum or plasma samples may include abundant proteins, *inter alia* albumin, IgG, antitrypsin, IgA, transferrin, haptoglobin and fibrinogen, which may preferably be so-depleted from the samples. Methods and systems for removal of abundant proteins are known, such as, *e.g.*, immunoaffinity depletion, and frequently commercially available, *e.g.*, Multiple Affinity Removal System (MARS-7, MARS-14) from Agilent Technologies (Santa Clara, California).

[0029] While the present invention has particular application to the quantification of proteins, it is to be appreciated that the invention also has an application for the quantification of other molecules, such as metabolites and the identification of potential small molecule and biological therapeutics.

[0030] In one embodiment, the present invention uses next generation polynucleotide sequencing techniques to identify and quantify aptamer tags. In this way, the present invention provides a more sensitive and efficient method of quantification, compared to currently available methods, and harnesses the large dynamic range, specificity and sensitivity offered by this genomic methodology in the field of proteomics.

[0031] Next generation polynucleotide sequencing techniques require the separation of individual molecules of DNA or RNA or other polynucleotide sequences onto a surface, such as a bead or chip, thereby creating a single molecule array of sequences. The array has a surface density that allows each molecule to be individually resolved, for example by optical microscopy. Sequencing of the polynucleotide molecules on the array allows 'digital', i.e. absolute, counting of the sequences and, thus, direct quantification of the sequences present on the array. In some technologies, sequences may be clonally amplified once arrayed to strengthen and/or clarify the signal from each sequence. Nonetheless, quantification is obtained by counting the occurrence of sequences on the array and not the signal produced from each amplicon. Examples of suitable sequencing and quantification techniques may be found in publications such as WO 00/006770 and Branton et al (Nature Biotechnology (2008) 26 1146-1153).

[0032] Accordingly, in a preferred embodiment, quantification of the bound aptamer sequences is carried out using a single molecule array of the bound sequences. Alternatively, the aptamer sequences may be quantified on a clonal array, in which each sequence is amplified after being arrayed on a surface.

[0033] Because of the high specificity of an aptamer for its target protein, an aptamer can be used as a proxy for a protein within a mixture. After binding to its protein and subsequent elution, quantification of an aptamer via counting the frequency of its sequence on a parallel sequencing device will enable quantification of a particular protein within the sample. This will, by implication or proxy, be in a manner similar to the use of antibodies as proxies for proteins. Indeed, the distribution of aptamer sequences or tags will be representative of the distribution of the corresponding proteins in the biological population or sample from which the protein was obtained.

[0034] Furthermore, counting of clonal molecules on arrays provides an absolute number for the quantity or proportion of each aptamer tag. This is termed "digital counting" and differs from existing DNA/RNA quantification methods which rely on indirect analogue measurement of parameters such as fluorescence of an attached dye. In this way, no manipulation of the proteins or the protein mixture is required, except exposure of the aptamers under suitable conditions. In addition, digital counting is believed to be superior as it avoids many of the ambiguities and signal noise problems of analogue quantification (Smyth GK, Bioinformatics (2007) Sep 19).

[0035] In a specific embodiment and as shown in Figure 1, a library of aptamers is screened and selected against *in vitro* sets of folded proteins derived from biological material e.g. serum. It will be appreciated that the library of aptamer tags may be generated by any method, including SELEX and methods described in co-pending application number EP 07020629.7.

[0036] Aptamers that do not bind proteins are removed or eluted. To achieve this, various known options are available. For example, a single protein or a complex protein sample may be immobilised on a solid support. Stringent washing may be performed to remove unbound and weak binding aptamers. Another option is the use of reversible cross-linking of the aptamer on the protein targets (photoaptamers). Non-SELEX approaches based on equilibrium capillary electrophoresis of equilibrium mixtures (ECEEM) or the non-equilibrium variant (NECEEM) may also be employed rapidly to select 'smart' aptamers with defined binding properties (Drabovich et al, Anal. Chem. (2006) 78, 3171-3178).

[0037] The remaining bound aptamers are removed from their protein hosts and run through a next generation sequencer where they are sequenced and counted.

[0038] As a specific example of next generation sequencing, the bound aptamers are randomly arrayed on a surface, such as a bead or a chip. The aptamers are optionally amplified cyclically yielding groups of clonal single stranded molecules at discrete x and y coordinates on the surface or on an individual bead. After the sequences are bound to the array, a primer is added to

PREOMICS EXHIBIT 1002
Page 342 of 1861

each sequence. A DNA sequencer then executes a stepwise chemistry consisting of reagents that allow the determination of one base on each complementary sequence per cycle. In this way, a complementary sequence mirroring each bound aptamer sequence is built up. An illumination and imaging system allows this process to be photographed and, by repeating the procedure several times, the sequences of the original aptamers can be obtained. Typically such technologies are able to sequence in excess of 40-100 million fragments of DNA at up to 50 base pairs in length and these numbers are increasing rapidly. This process currently takes less than 1 to 3 days from sample preparation to sequencing output and these timescales are shortening.

[0039] It will be appreciated that other methods that fall into the category of "next generation polynucleotide sequencing" may be used and the present invention is not limited to the specific example provided. "Next generation polynucleotide sequencing" is a term generally coined to described DNA/RNA sequencing platforms that have appeared since 2004. Their common characteristic is that they use a different sequencing chemistry from the old technology that was based on "Sanger" sequencing. The new platforms use new chemistries and generally have very high throughput and much lower costs. This has been achieved through the ability to parallelise the sequencing reactions. The new platforms typically proceed by synthesis or strand extension (building), unlike the Sanger method which works by snipping off bases (degrades). In addition, the new platforms operate on tiny numbers of clonal molecules, unlike the Sanger method which has a very large DNA sequencer molecule to base measurement ratio.

[0040] It will also be appreciated that while reference is made to DNA sequencing, the DNA sequencing technology may also be used with RNA or PNA (peptide nucleic acid) sequences. Thus, reference to next generation sequencers or sequencing encompasses DNA, RNA and PNA, as well as all other chemical variants of nucleic acid based polymers and their analogues that are suitable for use in the method of the present invention.

[0041] While a preferred embodiment of the invention uses a two dimensional array of the aptamer sequences, the invention encompasses next generation sequencing in solution and in three dimensions. For the latter, the tags may be concatenated to increase the dynamic range of the array.

[0042] A significant advantage of the present invention is that multiple aptamers directed to different proteins may be used to interrogate a complex protein mixture in one assay or experiment, thereby enabling quantification of many proteins in one sample in parallel. Such multiplexing is clearly amenable to high throughput assays such as next generation sequencing. It is believed that the power of multiple proteins to act as specific and sensitive biomarkers, when used in conjunction, far outweighs the use of single proteins since a number of biomarkers typically act in concert in a diseased state, rather than individually.

[0043] Ideally, the library aptamer sequences comprises a pool of aptamer sequences that are specific to a single protein. Alternatively, the pool comprises aptamer sequences specific to more than one protein. For example, the library may contain aptamer sequences al, a2 and a3 that are known to be specific to protein A. Alternatively, the library may contain aptamer sequences al, a2 and a3, as well as b1 and b2 specific for protein B, cl etc. While aptamer sequences a1, a2 and a3 may be specific for protein A, the invention encompasses the scenario when, for example, aptamer sequence a2 is also specific for protein B. In other words, the library of aptamers may contain sequences that are discretely specific for a protein, as well as sequences that bind to more than one protein. The latter scenario may be useful when searching for related families and groups of proteins within a mixture.

[0044] While it is preferred that the identity of the protein or molecule to which the or each aptamer binds is known, such a feature is not essential. For example, the class of protein to which an aptamer binds may be known but specific members of the class may be altered in a diseased state. The identity of proteins may be elucidated once quantification has been completed. For example, one or more specific aptamers may be assayed against a protein mixture. The aptamer(s) bind specifically to the protein of interest which may then be isolated and purified from the mixture. A significant advantage of the present invention is that actual identification of the protein biomarker need not be ascertained until the end of the interrogation process. Thus, manipulation of the protein sample is not required during interrogation.

[0045] It will be appreciated that polypeptide sequencing may be carried out by any known method. For example, an aptamer of interest may be immobilised on a column, a protein mixture flowed through the column and protein bound to the aptamer on the column analysed using conventional MS-based proteomics.

[0046] In another embodiment, the aptamers carry unique sequence tags. In this way, once aptamers have been identified as binding to one or more proteins of interest, the tags may be removed from the aptamers and the tags are then sequenced and counted. In this way, the tags are used as an address for an aptamer and provide an alternative method of providing quantification.

[0047] It will be appreciated that the method may be used to provide a variety of solutions and answers. In particular, the method may be used to interrogate a biological sample to ascertain whether certain proteins or molecules are up- or down-regulated in a diseased state when compared to a healthy state. Because panels of aptamers may be used at the same time, the invention allows the interrogation of a raft of potential and/or known biomarkers at the same time. The use of aptamer sequences as tags, rather than indirect tags and markers such as radiolabels and antibodies, allows identification of molecules that are up- and down-regulated, as well as

PREOMICS EXHIBIT 1002
Page 343 of 1861

alterations in post-translational modifications, conformational changes, emergence of variants and the like. In this way, the canvass of possible biomarkers that can be identified is far broader than that which is possible if using currently available techniques.

[0048] Indeed, particular areas of proteome may be selectively interrogated using a focussed set of aptamers. Furthermore, because aptamer sequences offer such a large range of possible tags, a number of discrete areas of the proteome may be interrogated at the same time.

[0049] In a further embodiment, an aptamer panel or pool may be challenged with a complex sample. The bound fraction may then be eluted and re-probed against a second, different sample. Free, unbound aptamers will then reflect the difference in protein abundance. In this way, changes in protein abundance can be measured, as well as the emergence of any proteins appearing in the sample as a result of a diseased or posttreatment state, for example. In these embodiments, the complex protein mixture is typically derived from a bodily fluid, such as blood, plasma or serum.

[0050] It will be appreciated that the invention may also be used to validate biomarkers. Specifically, a panel or collection of distinct aptamers may be screened against a protein sample from one or more healthy individuals and compared with protein samples taken from one or more individuals having a disease of interest.

[0051] Because many proteins may be quantified at the same time, the invention is highly amenable to high throughput analysis, thereby significantly reducing the cost and time typically required for biomarker validation.

[0052] An advantage of the present invention with respect to validation of a protein biomarker is that aptamers can be used instead of antibodies. The generation of specific antibodies is expensive and not always successful. On the other hand, it is usually possible to generate a selective aptamer to a given process and the process is far quicker and cheaper. In use, one or more aptamers specific for a particular protein of interest are assayed against a complex protein mixture. Absence of aptamer binding in a healthy sample compared to binding in a diseased sample (or vice versa) can be elucidated via counting of the aptamers. Again, no manipulation of the protein sample is necessary and no identification of proteins *per se* is required. In addition, multiple aptamers may be assayed against a raft of proteins of interest in one experiment.

[0053] By enabling the full complement of proteins to be interrogated, the method of the present invention may also be used to build up a picture of the conformation and abundancy of proteins in individuals. That is, to provide a profile, or map, of a patient's protein make-up. For example, a picture of Prostate Specific Antigen (PSA) may be used to establish prognosis and diagnosis of prostate cancer, as well as providing an indication of disease progression.

[0054] In another example, if a patient is found to have a certain protein variant, information about a patient's protein profile can be used for disease prognosis or susceptibility and can be used to assist with the development of individualised treatment regimes for the patient. For example, people carrying a specific mutation in their CCR5 gene produce a variant CCR5 receptor. These variant receptors do not bind HIV and so a person carrying the variant receptor has a lower chance of becoming HIV^+ and progressing to AIDS as a result.

[0055] The invention may also be used to determine specificity of an aptamer for a single molecule, such as a protein, by ascertaining the variance between replicates of the same sample, as well as variance between samples obtained from different individuals.

[0056] Use of the invention also resides in diagnostics. Specifically, one or more aptamers known to identify validated biomarkers may simply be used to interrogate a bodily fluid, such as blood or serum, taken from an individual. Because the binding of an aptamer indicates the presence of a protein, a panel of aptamers may be used to identify a single or a suite of biomarker proteins that are known to be altered in an abnormal state. For example, profiling may be used to assess quantities of specific proteins against a known baseline. Examples of such uses would be use as a pregnancy test, or a test for PSA levels. In another example, sepsis is an abnormal state in which many factors are known to be up- or down-regulated according to the type of sepsis and the stage of infection.

[0057] Such a diagnostic may be in the form of a kit that includes a panel of aptamers, the binding of which indicates the presence of biomarkers known to be specific to a particular abnormal state in a biological sample derived from an individual.

[0058] In another embodiment, the invention may be used to determine disease progression. For example, the biomarker pattern in sepsis changes over the course of the infection. The identification of biomarkers as indicators of sepsis progression would greatly assist in the drawing up of appropriate treatment regimes.

[0059] The invention also allows patient profiling which, in turn, may be used to direct treatment regimes and possible prophylactic treatment.

[0060] Thus, the invention provides a single platform on which a variety of outputs may be achieved.

[0061] In a yet further aspect, the present invention encompasses the use of clonal or single molecule arrays for the sequencing and counting of aptamers.

[0062] The inventors have realised that aptamers may ideally be adapted to enhance their compatibility with Next-Generation DNA sequencing platforms. Therefore, the aptamer may be enhanced to enable sequencing and identification of the aptamer using such technology platforms. Thus, from another aspect, the present invention encompasses a sequencing-enabled aptamer comprising an aptamer sequence and an adapter sequence. The adapter sequence comprises a sequence which enables attachment of the aptamer to a surface to enable se-

PREOMICS EXHIBIT 1002
Page 344 of 1861

quencing of the aptamer.

[0063] The sequencing-enabled aptamer may additionally comprise a sequencing primer for sequencing of the aptamer.

[0064] Such constructs may include more than one sequencing and attachment adapters. For example, the aptamer sequence may be sandwiched between pairs of sequencing and attachment adapters annealed to each end of the aptamer sequence.

[0065] The sequencing-enabled aptamer may also comprise a label, to enable visualisation of the aptamer. It will be appreciated that any suitable label may be used, including fluorescent labels and radio-labels.

[0066] Sequencing-enabled aptamers of the present invention may be used in the methods, uses and diagnostic kits described above.

[0067] The invention will now be described by way of non-limiting examples, in which:

Figure 1 illustrates the method of the present invention, in which a library of aptamer sequences is mixed with serum proteome. The bound aptamer/protein fraction is eluted before being sequenced and counted on a second generation sequencer. The output from the sequencer is the number of each sequence present in the bound fraction.

Figure 2 illustrates use of the method for biomarker discovery, in which proteomes from different patients are compared with a baseline or control.

Figure 3 illustrates an alternative embodiment of the present invention, in which a library of aptamers is mixed with a first serum proteome sample. After quantification of the bound fraction, the population of aptamers binding to proteins of interest within the serum sample is amplified and mixed with a second serum proteome sample. A comparison of the first and second respective quantification outputs is then made.

Figure 4 is a chromatogram showing the different constituents of the mixture obtained after a 30' incubation of equal volumes of 150nM IgE and 100nM ProNuclFR.

Figure 5 shows CE-chromatograms of 0/20/40... 2000nM IgE incubated with a constant (100 nM) ProNuclF concentration (from bottom to top)

Figure 6 is a graph showing the fraction of bound ProNuclF plotted against changing protein concentration.

Figure 7 is a schematic representation of the adapted sample preparation protocol for next generation sequencing of aptamers.

Figure 8 is a plot showing absolute counts of IgE aptamer sequences retrieved using two different analysis methods.

Figure 9 is a plot showing the fraction of IgE aptamer sequences present in a mixture of sequences counted by Next Generation sequencing plotted against the number of sequences spiked in a mixture of irrelevant (PhiX) sequences.

Example 1

[0068] This example is illustrated in Figure 1 which shows the quantification of proteins in a serum sample using a library of aptamer tags. A diverse library of known aptamers is screened against a protein mixture derived from serum. The protein mixture is immobilised on a solid support and stringent washing is performed to remove unbound and weakly binding aptamers.

[0069] The remaining bound aptamers are then removed from their protein hosts and sequenced and counted on a next generation sequencer. After being arrayed, the bound aptamer sequences may be amplified by standard methods, such as PCR, to increase the clarity of the signal over background noise on the sequencer.

[0070] The output from the sequencer will be an absolute quantification of proteins present in the biological sample, as represented by the library of aptamers.

Example 2

[0071] For use in biomarker discovery and as shown in Figure 2, the library of aptamers is screened against samples derived from different individuals. A suitable number of known control, or healthy samples will also be required to establish a baseline or healthy condition. A number of samples known to represent a diseased state will also be screened. A comparison of the two populations, healthy versus diseased, will allow the identification of aptamers that show significant alternations between the two populations. Once the sequence of the or each aptamer has been elucidated, the protein to which the aptamers bind may then be identified. Identification of the protein may be by next generation sequencing techniques, or proteomic-based methods, such as chromatography and MS.

Example 3

[0072] This Example describes an alternative method of the present invention and is illustrated in Figure 3.

[0073] An aptamer library is screened against a first biological sample. As with either Example 1 or Example 2, the non-bound aptamers are removed and the remaining bound aptamers are released from their host proteins and sequenced and counted on a next generation sequencer.

[0074] The population of sequences that bind to proteins in the first sample is then screened against a second biological sample. As before, the non-bound aptamers are discarded and the remaining aptamers are sequenced and counted. If the two samples are derived from healthy subjects, any variation between the two sample may be attributed to variance within the normal population. Similarly, any variance between the output from two samples derived from patients known to have

PREOMICS EXHIBIT 1002
Page 345 of 1861

a particular disease may also be attributed to variance. However, any significant variance between healthy and diseased subject may be attributed to a change in the protein population as a result of the disease of interest. The aptamer sequences in which significant changes are found may then be translated to identify the protein(s) to which the aptamer(s) bind.

Example 4

[0075] In this example, the aptamer library comprises at least one pool of aptamers specific for one protein. Equally, the panel may contain a number of pools of aptamers, where each pool is specific for a different protein.

[0076] The protein mixture may simply be a biological sample, such as blood or serum. Alternatively, the protein mixture may be obtained by enriching the biological sample for the protein fraction. In this example, care must be taken to minimise disruption to and denaturing of the protein population.

[0077] The library of known, selected aptamer sequences is screened against the protein mixture. In one example, the protein mixture is bound to a column and the panel of aptamers is allowed to flow through the column. Aptamers that do not bind are removed. Aptamers that do bind are removed from their protein hosts and sequenced and counted on a next generation sequencer.

[0078] If the protein of interest is indeed a biomarker for a diseased state, aptamers will be found in the bound fraction from the diseased sample and not in a sample derived from a healthy individual. Alternatively, the role of a protein as a biomarker may be manifest by up- or down-regulation of the protein in a diseased individual when compared to a healthy subject.

Example 5

[0079] In this example, the library of aptamers contains one or more sequences known to be specific for one or more proteins. The identity of the or each protein of interest is also known. The library may comprise a single pool specific to a single protein, or it may comprise many pools specific to an array of proteins.

[0080] The library of aptamers is screened against a biological sample derived from an individual, such as blood. The aptamer library may be held on a support and the biological sample passed over the aptamers. Any unbound proteins are removed and the aptamer:protein complexes retained. The retained aptamers are released from their protein hosts and sequenced and counted on a next generation sequencer.

[0081] The presence and quantity, or absence, of aptamers on the next generation sequencer may be used to diagnose a particular disease or predicament.

Example 6

[0082] A novel oligonucleotide aptamer was designed so that the aptamer was compatible with the Next-Generation DNA sequencing technology platform that was used.

[0083] The aptamer had an oligonucleotide sequence having three distinct regions: a) a functional aptamer region, b) an adapter region, and c) a label.

[0084] The aptamer region in this study was based on a well-studied aptamer (Wiegand et al Journal of Immunology (1996) 157 221-230) having a high affinity for human Immunoglobulin-E.

[0085] The adapter region made the construct compatible with the sample preparation procedures of an Illumina GA1 Next Generation sequencer and was designed according to the manufacturer's protocols.

[0086] The label was fluorescent 6-carboxyfluorescein (FAM), the inclusion of which was to allow visualization of the construct.

[0087] The constructs were obtained by standard solid phase DNA synthesis and purified by polyacrylamide gel electrophoresis to ensure low error rates.

[0088] The construct with the FAM label will be referred to hereinafter as ProNuc1F and without the label as ProNuc1.

[0089] The constructs were used to demonstrate that:

i) such aptamer-derived single strand oligonucleotide sequences are compatible with sample preparation procedures used for Next Generation DNA sequence applications,
ii) the sequences can be sequenced using the Illumina GA1 sequencing technology, i.e. the identity of the ProNuc1 can be retrieved ("read"); and
iii) the count of whole ProNuc1 sequences, following a standard Illumina GA1 sequencing experiment, directly correlates with the number of prepared ProNuc1 sequences spiked into a sample, thereby confirming that Next Generation sequencing can be applied to proteomics in accordance with the present invention.

Example 7

[0090] Capillary electrophoresis (CE) was used to confirm that the oligonucleotide sequence exhibits a distinct protein affinity.

Preparation of reagents and solutions

[0091] A stock 100 nM stock solution of ProNucF1 was prepared from dry aptamer source by diluting the DNA material in TGK buffer (tris(hydroxyamino)methane-glycine-potassium) buffer, pH 8.4. Following dilution, the aptamer stock solution was incubated at 75 °C for 10 minutes and subsequently cooled and stored in ice to ensure the absence of multimers. The heating procedure

PREOMICS EXHIBIT 1002
Page 346 of 1861

also ensures formation of the secondary structure of the aptamer.

[0092] Starting from a concentrated, 5500 nM stock solution of protein a dilution series of IgE was prepared in TGK buffer (Table 1):

Table 1: IgE dilution series

Solution	Concentration (nM)
1	20
2	40
3	60
4	100
5	150
6	200
7	300
8	1000
9	2000

[0093] The samples for CE demonstrating aptamer-protein complex formation were prepared by incubating 5 μl of the ProNucF1 stock solution with 5 μl of protein solution. A total of six replicates were prepared in this way for each protein concentration level. All samples were incubated at room temperature for a minimum of 30 minutes and no longer than 40 minutes.

Apparatus

[0094] Separations were achieved using a Beckman P/ACE 2200 CE-LIF system (Beckman-Coulter, Fullerton, CA, USA). Separations were performed using fused-silica capillaries (50 μm ID, 360 μm OD) with 40.2 cm total length, detector at 30 cm. Capillaries were pretreated by pumping 1 M NaOH, deionized H_2O and buffer through the capillary for 10 min each. Between each electrophoretic separation the capillary was rinsed with base, water, and buffer again to remove any residual sample from the capillary walls. Samples were injected using pressure for 4 s at 1 psi. The LIF detector employed the 488 nm line of a 3 mW Ar-ion laser (Beckman-Coulter) for excitation and emission was collected through a 520 ±10 nm filter. Data were recorded and analyzed with P/ACE software (Beckman-Coulter).

Protein-aptamer binding

[0095] Samples at each protein concentration level were run on the column in normal mode using the conditions described in (Table 2), firstly to ensure that a protein-aptamer complex was indeed formed after incubation of each sample.

Table 2: Normal mode separation conditions

Action	Duration (s)
Rinse 20 psi - 0.1M NaOH	120
Rinse 20 psi - Buffer	120
Inject 1 psi 4 - Sample	4
Inject 0.1 psi - Water	1
Separate 20 kV	760
Rinse 10 psi - Buffer	60

Results

[0096] As an example, a chromatogram for normal mode injection for a sample containing 150 nM of IgE and 100 nM of ProNucIF is shown in Figure 4.

[0097] In order to apply the aptamer-protein complex formation for quantification purposes, the bound-aptamer fraction has to reflect the protein dilution series. To prove, this a constant amount of ProNuc1F was incubated with different protein concentrations (see above).

[0098] From Figure 5 it is clear that, with increasing protein concentration (bottom trace "0"; upper trace 5000 nM), the peak area of the aptamer-protein complex increases. Simultaneously, the signal of the free ProNuc1F decreases to the point where no free aptamer is detected (upper trace). This shows that adding an adapter sequence specific for the Next Generation Sequencing procedure to a known aptamer sequence against IgE does not impair the aptamer capabilities of ProNuc1F - ProNuc1F retains its high and selective affinity towards IgE.

[0099] Figure 6 shows the information in Figure 5 translated into a graph where the "Bound ProNuc1F fraction" is plotted against the protein concentration. From this graph it can be seen that the ProNuc1F aptamer can be used to "read out" quantitative protein information. From this (and other) graphs it is also possible to deduce equilibrium dissociation constants Kd, which give insight in the affinity of ProNucIF for IgE, i.e. between 76 and 92 nM.

Example 8

[0100] Experiments were carried out to demonstrated that the binding characteristics illustrated in the previous Example are applicable in a quantitative context, i.e. differences in protein concentration correlate with the fraction of bound aptamer (aptamer-protein complex). Selection for these aptamer-protein complexes was achieved by means of nonequilibrium capillary electrophoresis of equilibrium mixtures (NECEEM) as developed by Krylov et al (Journal of the American Chemical Society (2002) 124 13674-13675; Analytical Chemistry (2003) 75 1382-1386).

Method

[0101] An adapted sample preparation protocol was

PREOMICS EXHIBIT 1002
Page 347 of 1861

developed to allow sequencing of the functional IgE aptamer containing the added sequencing sequence (ProNuc1).

[0102] Figure 7 outlines the different steps in the sample preparation protocol. After the annealing of primer 1 on Sequencing primer site 1, a double stranded molecule was formed by extension with Taq polymerase. An adaptor was ligated to the double stranded molecule generating the full molecule.

[0103] The IgE aptamer sequence used in this study was a single stranded oligonucleotide with Sequencing Primer Site 1 at its 3' end. In order to make the complementary strand, a tailed primer was annealed with primer 1 and extension was performed using Taq polymerase.

[0104] The recipe for this reaction was:

10 μM Pronucl	2μl
10 μM Primer 1	5μl
2.5 mM dNTPs	4μl
10x Taq polymerase buffer	5μl
50mM $MgCl_2$	2μl
Taq polymerase	0.5μl
Water	31.5μl

[0105] The reaction was mixed and heated to 94°C for 2 minutes, followed by controlled cooling to 60°C and incubation at 72°C for 10 minutes. Reaction products were cleaned using a Qiagen MinElute column, following the manufacturer's instructions, and eluting in 10μl EB buffer.

[0106] 20μl each of a 100μM solution of ProAd1 and ProAd2 adapters were annealed together in a total volume of 50μl 10mM Tris-HCl pH 8.3 by heating to 94°C for 2 minutes and controlled cooling to room temperature. ProAd1 and ProAd2 are synthesized oligonucleotides that are derived from the Illumina GA1 kit.

[0107] The mixture used was:

Ds aptamer	10μl
2x DNA Illumina ligase buffer	25μl
Annealed adapter oligo mix	10μl
Illumina DNA ligase	5μl

[0108] Reactions were mixed and incubated for 15 minutes at room temperature before gel electrophoresis in 2% agarose. The appropriately-sized band was excised from the gel and the DNA was extracted and amplified following the standard Illumina enrichment PCR protocol.

[0109] PCR amplicons were quantified using an Agilent Bioanalyzer 2100 and sequencing reactions were prepared by adding these amplified aptamers to a constant quantity of PhiX, with the intention of generating the Illumina GA1 cluster numbers / tile given in Table 3, to produce a 2x dilution series. Lane 4 is PhiX with no added aptamer and serves as a control lane.

[0110] For the avoidance of doubt, PhiX refers to the circular genome of the double stranded DNA PhiX174 bacteriophage which consists of 5386 nucleotides. In this study, the genome was processed according to the Illumina GA1 sequencing protocol to serve as normalization for the sequence load in the different lanes.

[0111] A single end flowcell was prepared and sequenced on an Illumina GA1 in accordance with the manufacturer's instructions.

Table 3: Actual ProNuc1 and PhiX sequence numbers used per tile for the eight different lanes on the Illumina GA1 flowcell. A lane consists of around 300 tiles.

Lane	ProNuc1/tile	ProNuc1/lane	PhiX
1	5000	1500000	20,000
2	2500	750000	20,000
3	1250	375000	20,000
4	0	0	20,000
5	625	187500	20,000
6	312	93600	20,000
7	156	46800	20,000
8	78	23400	20,000

Results:

[0112] Sequencing data was processed according to the manufacturer's instructions. For data analysis, different strategies were applied to evaluate the sequence read outs obtained in the different lanes: a) a search for exact matches to the aptamer sequence (exact), b) a search which allows for up to 3 amplification/sequencing errors (Agrep). It is clear that other processing methods can also be applied. Figure 8 plots the absolute numbers of IgE aptamers counted in the different lanes by looking for exact matches/matches with up to three errors retrieved in the different lanes, i.e. an exact aptamer sequence count and an error-tolerant aptamer sequence count (Agrep). The values at the X-axis correspond to the lane readouts of lanes 8 down to 1 (control lane 4 is excluded; see Table 3).

[0113] In contrast, Figure 9 plots the fraction of IgE aptamers counted by looking for exact matches in the total sequences obtained against the number of aptamers spiked in the PhiX sequence mixture. In both cases a strong linear correlation is shown confirming the use of Next Generation sequencing to quantify aptamers.

[0114] The above data shows that, when following a standard Illumina GA1 sequencing protocol, the count of whole ProNuc1 sequences correlates directly with the number of prepared ProNuc1 sequences spiked into the sample, thereby confirming that Next Generation sequencing can be used for quantification in accordance with the present invention.

[0115] Furthermore, the data show that the prepared ProNuc1 sequences can be sequenced using Next Gen-

PREOMICS EXHIBIT 1002
Page 348 of 1861

eration sequencing technology. In other words the identity of the ProNuc1 can be retrieved ("read").

Claims

1. A method for measuring the amount of at least one molecule in a biological sample, the method comprising:

   a) combining the sample with one or more aptamers and allowing one or more molecules in the sample to bind to the aptamer(s);
   b) separating bound from unbound molecules; and
   c) quantifying the molecule(s) bound to the or each aptamer,

   wherein quantification of the bound molecule(s) is carried out by sequencing at least part of the or each aptamer.

2. A method according to Claim 1, wherein the at least one molecule is a protein.

3. A method according to Claim 1 or Claim 2, wherein the identity of the molecule is known.

4. A method according to Claim 1 or Claim 2, wherein the identity of the molecule is unknown.

5. A method according to Claim 4, wherein the method further comprises determining the identity of the molecule.

6. A method according to any one of Claims 1 to 5, wherein the sequence of the or each aptamer is known.

7. A method according to any one of Claims 1 to 6, wherein each aptamer sequence carries a unique tag.

8. A method according to Claim 7, wherein the tag is the sequence of the aptamer.

9. A method according to Claim 7 or Claim 8, wherein the tag is part of the sequence of the aptamer.

10. A method according to any one of Claims 1 to 9, wherein sequencing is carried out on a single molecule array or a clonal single molecule array.

11. A method according to any one of Claim 1 to 10, wherein the method further comprises removing the aptamer(s) bound to the or each molecule and arraying the aptamer(s) on a surface.

12. A method according to Claim 11, wherein the method further comprises amplifying the arrayed aptamers.

13. A method according to any on of Claims 1 to 12, wherein the one or more aptamers comprise different sequences that bind to the same molecule.

14. A method according to any one of Claims 1 to 12, wherein the one or more aptamers comprise panels of aptamers sequences, each panel binding to a different molecule.

15. A method according to any one of Claims 1 to 14, wherein the aptamers are derived from polynucleotides.

16. A method according to Claim 15, wherein the aptamers are polynucleotides and have between about 30 and about 60 bases.

17. A method according to Claim 16, wherein the aptamers have about 40 bases.

18. A method according to any one of Claims 1 to 17, wherein the biological sample is a bodily fluid.

19. A method according to Claim 18, wherein the bodily fluid is blood or derived from blood.

20. A method according to Claim 19, wherein the bodily fluid is serum or plasma.

21. A method according to any one of Claims 1 to 20, wherein the method further comprises:

   d) combining a second biological sample with the bound aptamer fraction obtained in c);
   e) separating bound from unbound molecules;
   f) quantifying the molecule(s) bound to aptamers; and
   g) comparing quantities obtained in c) with those obtained in f),

   wherein quantification of aptamers bound to one or more molecules in the second sample is carried out by sequencing at least part of each aptamer.

22. A method according to any one of Claims 1 to 21, wherein the method further comprises comparing the quantity of the or each aptamer against a control or baseline quantity.

23. A method according to Claim 21 or Claim 22, wherein the second sample is obtained from an individual known to be in a diseased state.

24. A method according to Claim 21 or Claim 22, wherein the second sample is obtained from an individual

PREOMICS EXHIBIT 1002
Page 349 of 1861

known to be in a healthy state.

25. A method according to Claim 21 or Claim 22, wherein the second sample is obtained from an individual after drug treatment.

26. Use of a method as claimed in any one of Claims 1 to 25 to identify one or more biomarkers.

27. Use of a method as claimed in any one of Claims 1 to 25 to validate one or more biomarkers.

28. Use of a method as claimed in any one of Claims 1 to 25 as a diagnostic method.

Patentansprüche

1. Verfahren zum Messen der Menge von wenigstens einem Molekül in einer biologischen Probe, wobei das Verfahren folgendes umfasst:

    a) das Kombinieren der Probe mit einem oder mehreren Aptameren und das Erlauben eines oder mehrerer Moleküle in der Probe, sich an das/die Aptamer(e) zu binden;
    b) das Trennen der gebundenen von den ungebundenen Molekülen; und
    c) das Quantifizieren des Moleküls/der Moleküle, das/die an jedes Aptamer gebunden ist/sind,

    wobei die Quantifizierung des gebundenen Moleküls/der gebundenen Moleküle durch Sequenzierung von wenigstens einem Teil des oder jedes Aptamers durchgeführt wird.

2. Verfahren nach Anspruch 1, wobei das wenigstens eine Molekül ein Protein ist.

3. Verfahren nach Anspruch 1 oder Anspruch 2, wobei die Identität des Moleküls bekannt ist.

4. Verfahren nach Anspruch 1 oder Anspruch 2, wobei die Identität des Moleküls unbekannt ist.

5. Verfahren nach Anspruch 4, wobei das Verfahren ferner das Feststellen der Identität des Moleküls umfasst.

6. Verfahren nach einem der Ansprüche 1 bis 5, wobei die Sequenz des oder jedes Aptamers bekannt ist.

7. Verfahren nach einem der Ansprüche 1 bis 6, wobei jede Aptamersequenz eine einzigartige Markierung trägt.

8. Verfahren nach Anspruch 7, wobei die Markierung die Sequenz des Aptamers ist.

9. Verfahren nach Anspruch 7 oder Anspruch 8, wobei die Markierung Teil der Sequenz des Aptamers ist.

10. Verfahren nach einem der Ansprüche 1 bis 9, wobei die Sequenzierung an einer einzelnen Molekülanordnung oder einer klonalen einzelnen Molekülanordnung durchgeführt wird.

11. Verfahren nach einem der Ansprüche 1 bis 10, wobei das Verfahren ferner das Entfernen des Aptamers/der Aptamere umfasst, die an das oder jedes Molekül gebunden ist/sind sowie das Anordnen des Aptamers/der Aptamere an einer Oberfläche.

12. Verfahren nach Anspruch 11, wobei das Verfahren ferner die Verstärkung der angeordneten Aptamere umfasst.

13. Verfahren nach einem der Ansprüche 1 bis 12, wobei das eine Aptamer oder die mehreren Aptamere verschiedene Sequenzen umfassen, die sich an dasselbe Molekül binden.

14. Verfahren nach einem der Ansprüche 1 bis 12, wobei das eine Aptamer oder die mehreren Aptamere Aptamer-Sequenzelemente umfasst/umfassen, wobei sich jedes Element an ein verschiedenes Molekül bindet.

15. Verfahren nach einem der Ansprüche 1 bis 14, wobei die Aptamere von Polynukleotiden abgeleitet sind.

16. Verfahren nach Anspruch 15, wobei die Aptamere Polynukleotide sind und zwischen etwa 30 und etwa 60 Basen aufweisen.

17. Verfahren nach Anspruch 16, wobei die Aptamere etwa 40 Basen aufweisen.

18. Verfahren nach einem der Ansprüche 1 bis 17, wobei die biologische Probe eine Körperflüssigkeit ist.

19. Verfahren nach Anspruch 18, wobei die Körperflüssigkeit Blut oder von Blut abgeleitet ist.

20. Verfahren nach Anspruch 19, wobei die Körperflüssigkeit Serum oder Plasma ist.

21. Verfahren nach einem der Ansprüche 1 bis 20, wobei das Verfahren ferner folgendes umfasst:

    d) das Kombinieren einer zweiten biologischen Probe mit der in c) erhaltenen gebundenen Aptamerfraktion;
    e) das Trennen der gebundenen von den ungebundenen Molekülen;
    f) das Quantifizieren der an Aptamere gebundenen Molekül€; und

PREOMICS EXHIBIT 1002
Page 350 of 1861

g) das Vergleichen der in c) erhaltenen Mengen mit den in f) erhaltenen,

wobei die Quantifizierung von Aptameren, die an ein oder mehrere Moleküle in der zweiten Probe gebunden sind, durch Sequenzieren wenigstens eines Teils jedes Aptamers durchgeführt wird.

22. Verfahren nach einem der Ansprüche 1 bis 21, wobei das Verfahren ferner das Vergleichen der Menge des oder jedes Aptamers gegenüber einer Kontroll- oder Basismenge umfasst.

23. Verfahren nach Anspruch 21 oder Anspruch 22, wobei die zweite Probe von einer Person erhalten wird, die sich bekanntermaßen in einem Krankheitszustand befindet.

24. Verfahren nach Anspruch 21 oder Anspruch 22, wobei die zweite Probe von einer Person erhalten wird, die bekanntermaßen gesund ist.

25. Verfahren nach Anspruch 21 oder Anspruch 22, wobei die zweite Probe von einer Person nach einer Behandlung mit Arzneimitteln erhalten wird.

26. Verwendung eines Verfahrens nach einem der Ansprüche 1 bis 25 zum Identifizieren einer oder mehrerer Biomarkierungen.

27. Verwendung eines Verfahrens nach einem der Ansprüche 1 bis 25, um eine oder mehrere Biomarkierungen zu bestätigen.

28. Verwendung eines Verfahrens nach einem der Ansprüche 1 bis 25 als ein diagnostisches Verfahren.

Revendications

1. Procédé de mesure de la quantité d'au moins une molécule dans un échantillon biologique, le procédé comprenant :

a) combiner l'échantillon avec un ou plusieurs aptamères et amener une ou plusieurs molécules dans l'échantillon à se lier à l'aptamère ou aux aptamères ;
b) séparer les molécules liées des molécules non liées ; et
c) quantifier la (ou les) molécule(s) liée(s) à l'aptamère ou à chaque aptamère,

dans lequel une quantification de la (ou des) molécule(s) liée(s) est effectuée par le séquençage d'au moins une partie de l'aptamère ou de chaque aptamère.

2. Procédé selon la revendication 1, dans lequel la au moins une molécule est une protéine.

3. Procédé selon la revendication 1 ou la revendication 2, dans lequel l'identité de la molécule est connue.

4. Procédé selon la revendication 1 ou la revendication 2, dans lequel l'identité de la molécule est inconnue.

5. Procédé selon la revendication 4, dans lequel le procédé comprend en outre la détermination de l'identité de la molécule.

6. Procédé selon l'une quelconque des revendications 1 à 5, dans lequel la séquence de l'aptamère ou de chaque aptamère est connue.

7. Procédé selon l'une quelconque des revendications 1 à 6, dans lequel chaque séquence d'aptamère porte un marqueur unique.

8. Procédé selon la revendication 7, dans lequel le marqueur est la séquence de l'aptamère.

9. Procédé selon la revendication 7 ou la revendication 8, dans lequel le marqueur est une partie de la séquence de l'aptamère.

10. Procédé selon l'une quelconque des revendications 1 à 9, dans lequel un séquençage est effectué sur un réseau de molécules simples ou un réseau clonal de molécules simples.

11. Procédé selon l'une quelconque des revendications 1 à 10, dans lequel le procédé comprend en outre l'élimination de l'aptamère (ou des aptamères) lié(s) à la ou à chaque molécule et la disposition en réseau de l'aptamère ou des aptamères sur une surface.

12. Procédé selon la revendication 11, dans lequel le procédé comprend en outre l'amplification des aptamères disposés en réseau.

13. Procédé selon l'une quelconque des revendications 1 à 12, dans lequel le ou les aptamères comprennent différentes séquences qui se lient à la même molécule.

14. Procédé selon l'une quelconque des revendications 1 à 12, dans lequel le ou plusieurs aptamères comprennent des panels de séquences d'aptamères, chaque panel se liant à une molécule différente.

15. Procédé selon l'une quelconque des revendications 1 à 14, dans lequel les aptamères sont issus de polynucléotides.

16. Procédé selon la revendication 15, dans lequel les

PREOMICS EXHIBIT 1002
Page 351 of 1861

aptamères sont des polynucléotides et ont entre environ 30 et environ 60 bases.

17. Procédé selon la revendication 16, dans lequel les aptamères ont environ 40 bases.

18. Procédé selon l'une quelconque des revendications 1 à 17, dans lequel l'échantillon biologique est un fluide corporel.

19. Procédé selon la revendication 18, dans lequel le fluide corporel est le sang ou est issu du sang.

20. Procédé selon la revendication 19, dans lequel le fluide corporel est le sérum ou le plasma.

21. Procédé selon l'une quelconque des revendications 1 à 20, dans lequel le procédé comprend en outre :

d) combiner un second échantillon biologique avec la fraction d'aptamère lié obtenue en c) ;
e) séparer les molécules liées des molécules non liées ;
f) quantifier la ou les molécules liées aux aptamères ; et
g) comparer les quantités obtenues en c) avec celles obtenues en f),

dans lequel une quantification d'aptamères liés à une ou plusieurs molécules dans le second échantillon est effectuée par séquençage d'au moins une partie de chaque aptamère.

22. Procédé selon l'une quelconque des revendications 1 à 21, dans lequel le procédé comprend en outre la comparaison de la quantité de l'aptamère ou de chaque aptamère contre une quantité témoin ou de ligne de base.

23. Procédé selon la revendication 21 ou la revendication 22, dans lequel le second échantillon est obtenu à partir d'un individu connu comme étant dans un état de maladie.

24. Procédé selon la revendication 21 ou la revendication 22, dans lequel le second échantillon est obtenu à partir d'un individu connu comme étant dans un état sain.

25. Procédé selon la revendication 21 ou la revendication 22, dans lequel le second échantillon est obtenu à partir d'un individu après un traitement par un médicament.

26. Utilisation d'un procédé selon l'une quelconque des revendications 1 à 25 pour identifier un ou plusieurs biomarqueurs.

27. Utilisation d'un procédé selon l'une quelconque des revendications 1 à 25 pour valider un ou plusieurs biomarqueurs.

28. Utilisation d'un procédé selon l'une quelconque des revendications 1 à 25 comme méthode de diagnostic.

PREOMICS EXHIBIT 1002
Page 352 of 1861

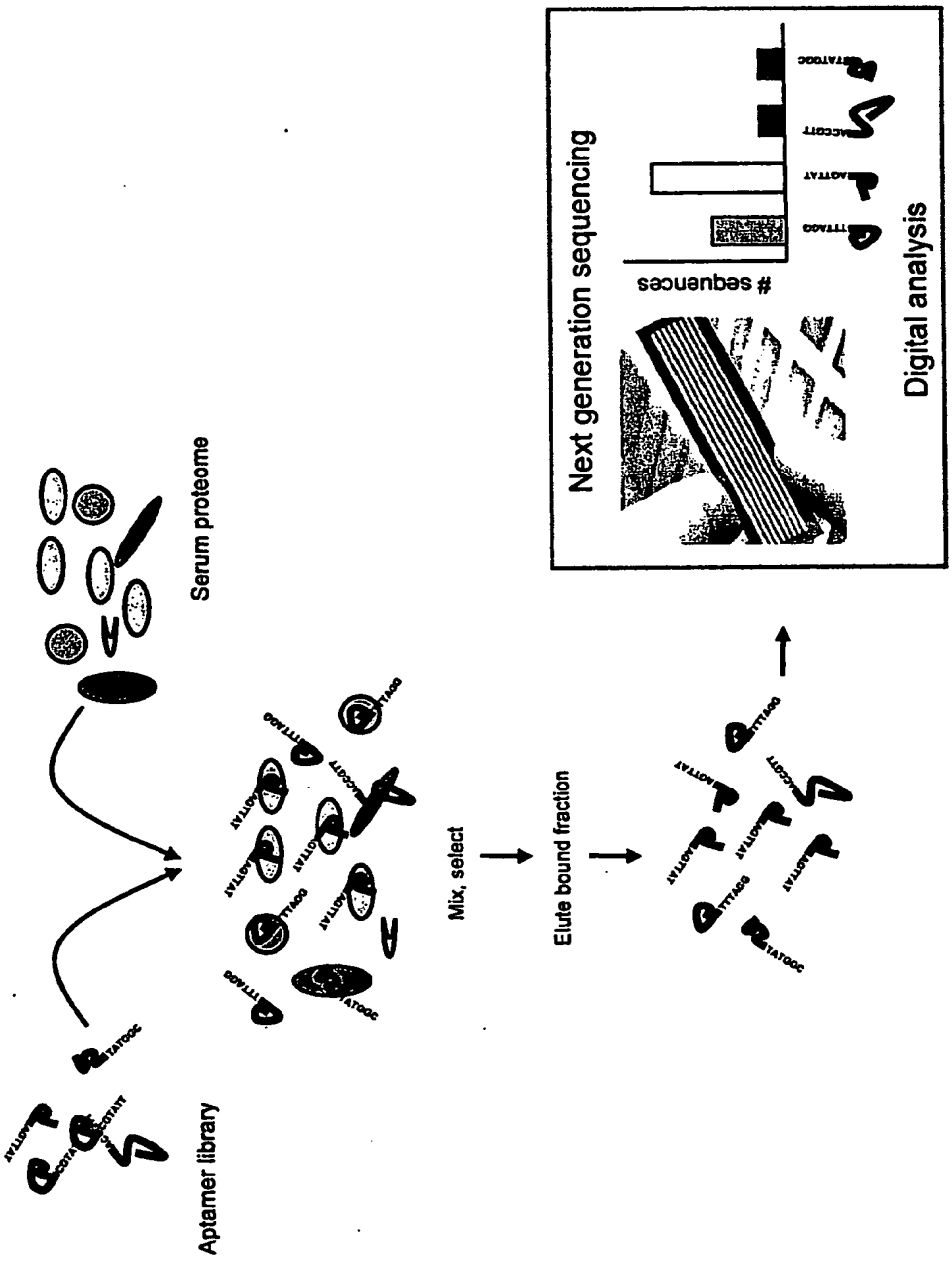


Figure 1

PREOMICS EXHIBIT 1002
Page 353 of 1861

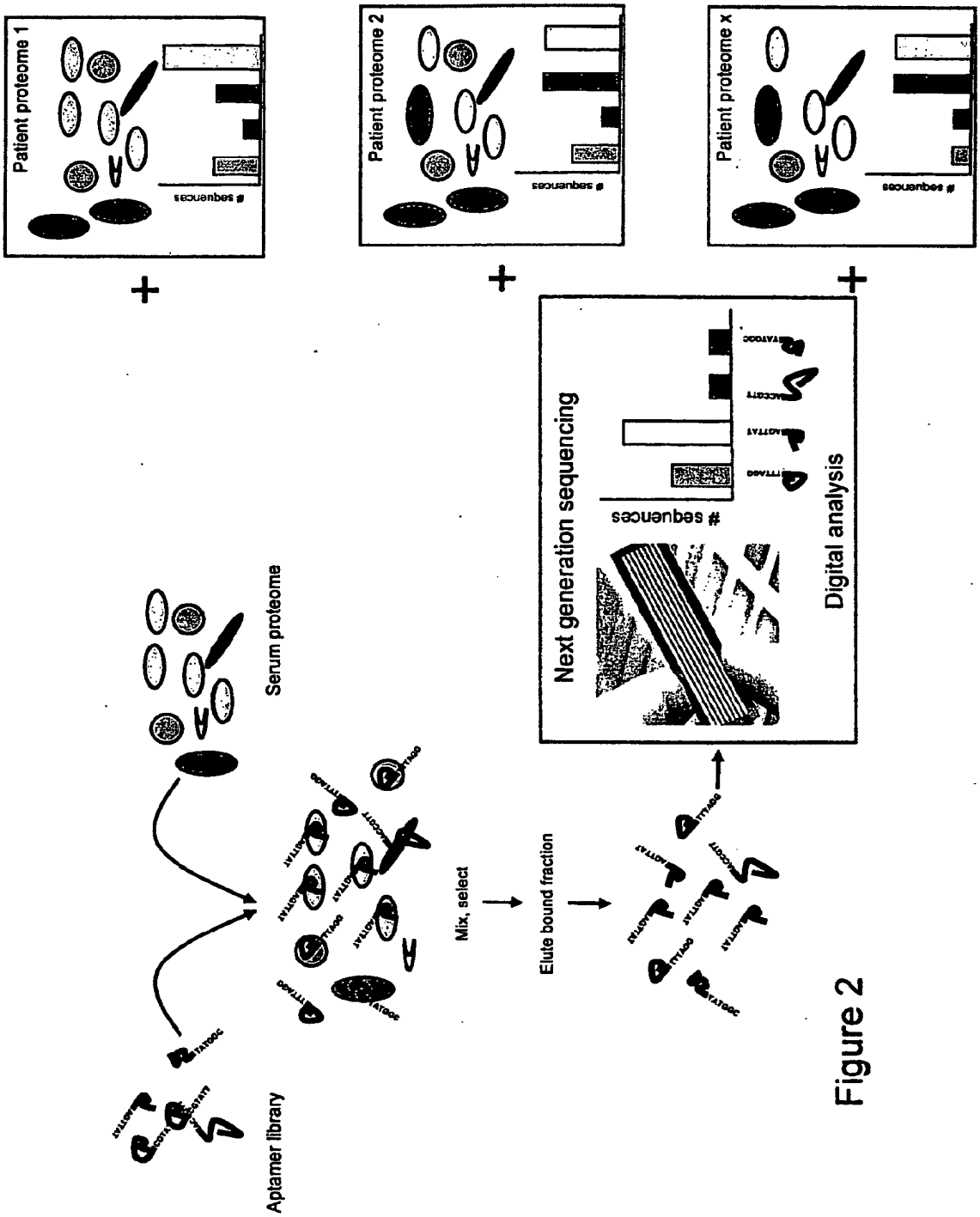


Figure 2

PREOMICS EXHIBIT 1002
Page 354 of 1861

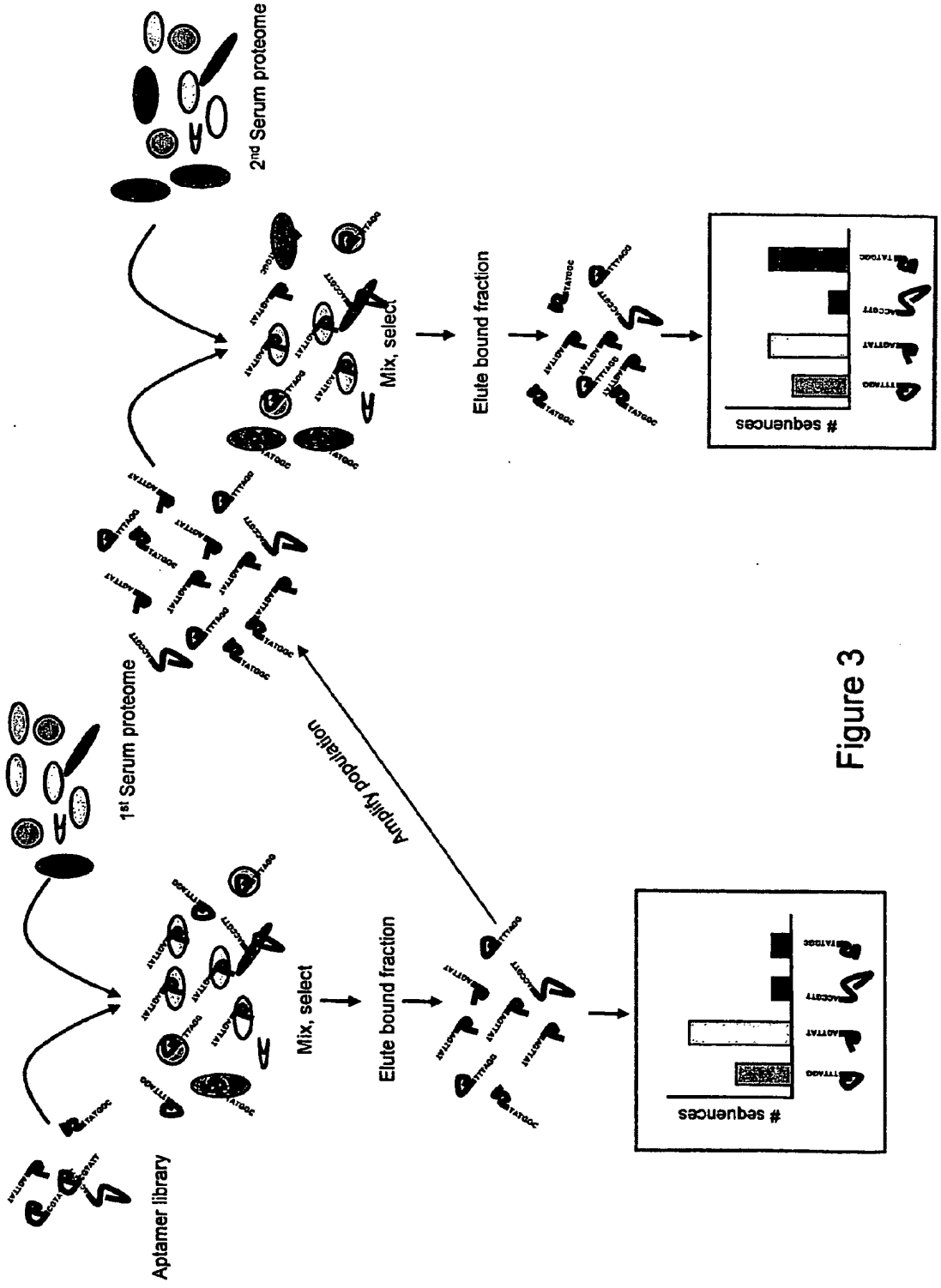


Figure 3

PREOMICS EXHIBIT 1002
Page 355 of 1861

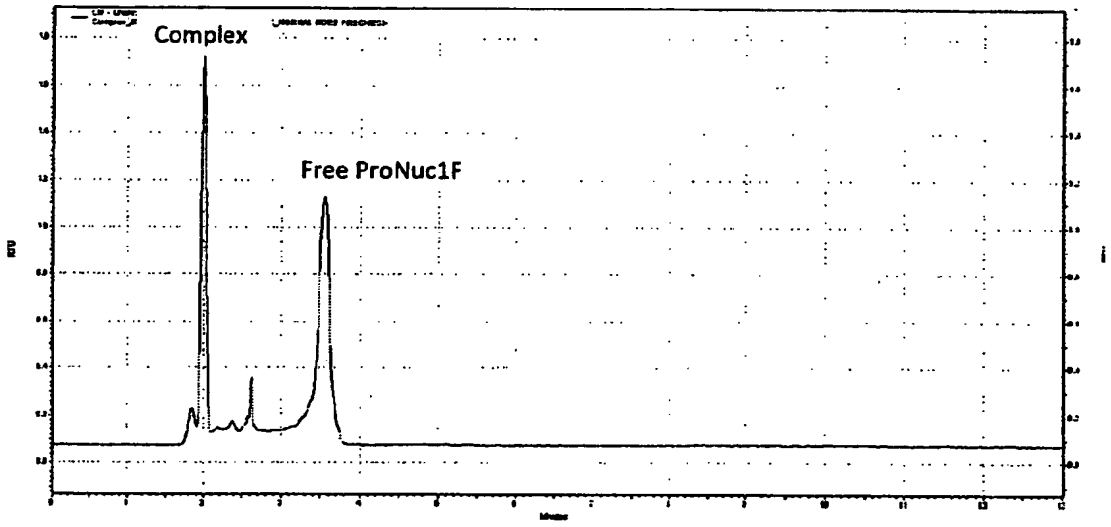


Figure 4

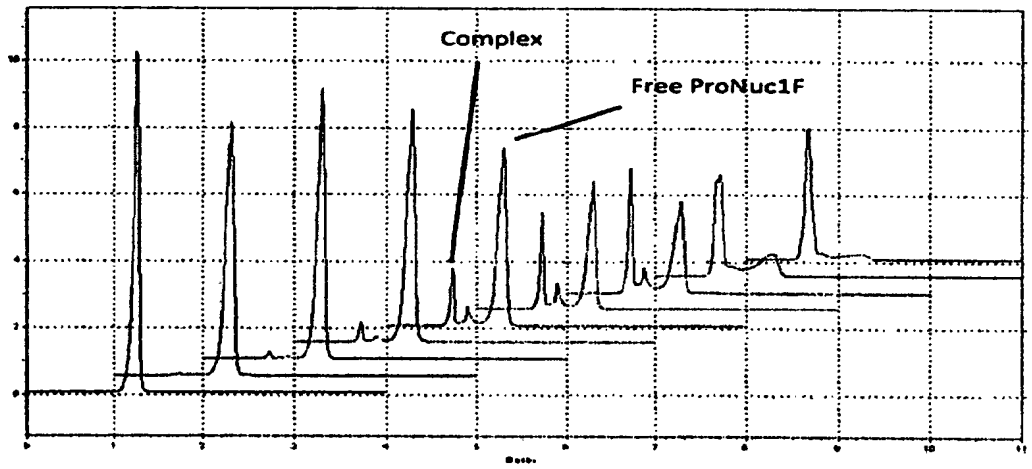


Figure 5

PREOMICS EXHIBIT 1002
Page 356 of 1861

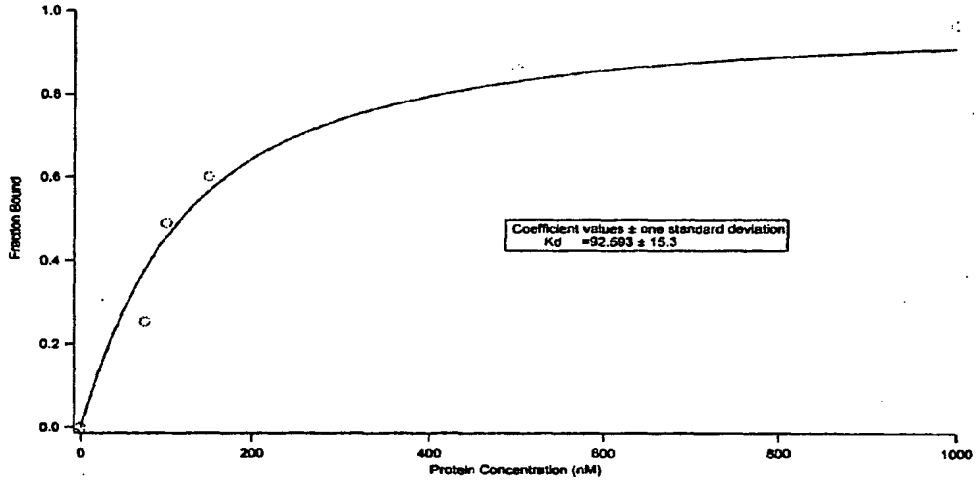


Figure 6

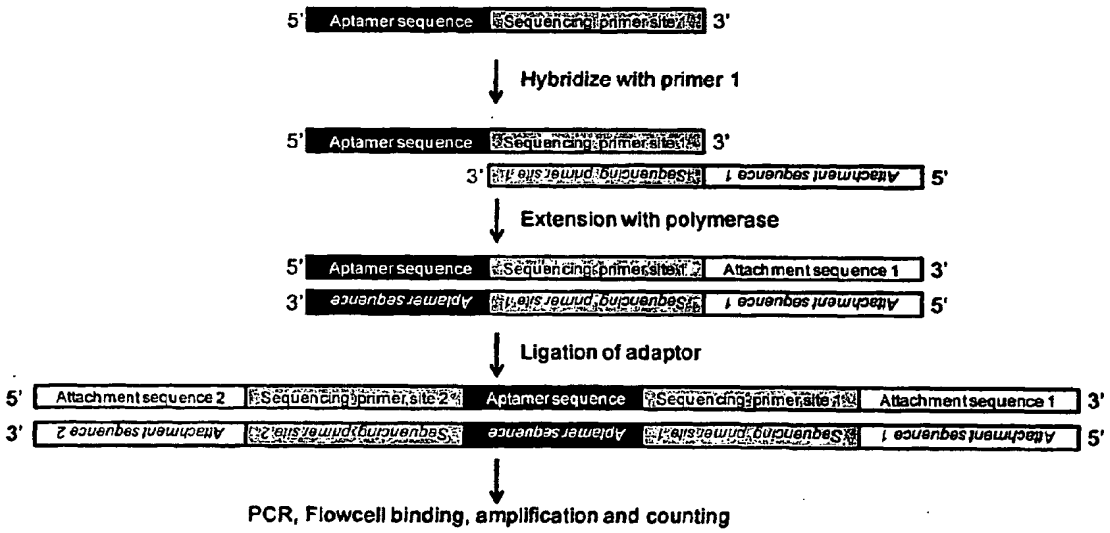


Figure 7

PREOMICS EXHIBIT 1002
Page 357 of 1861

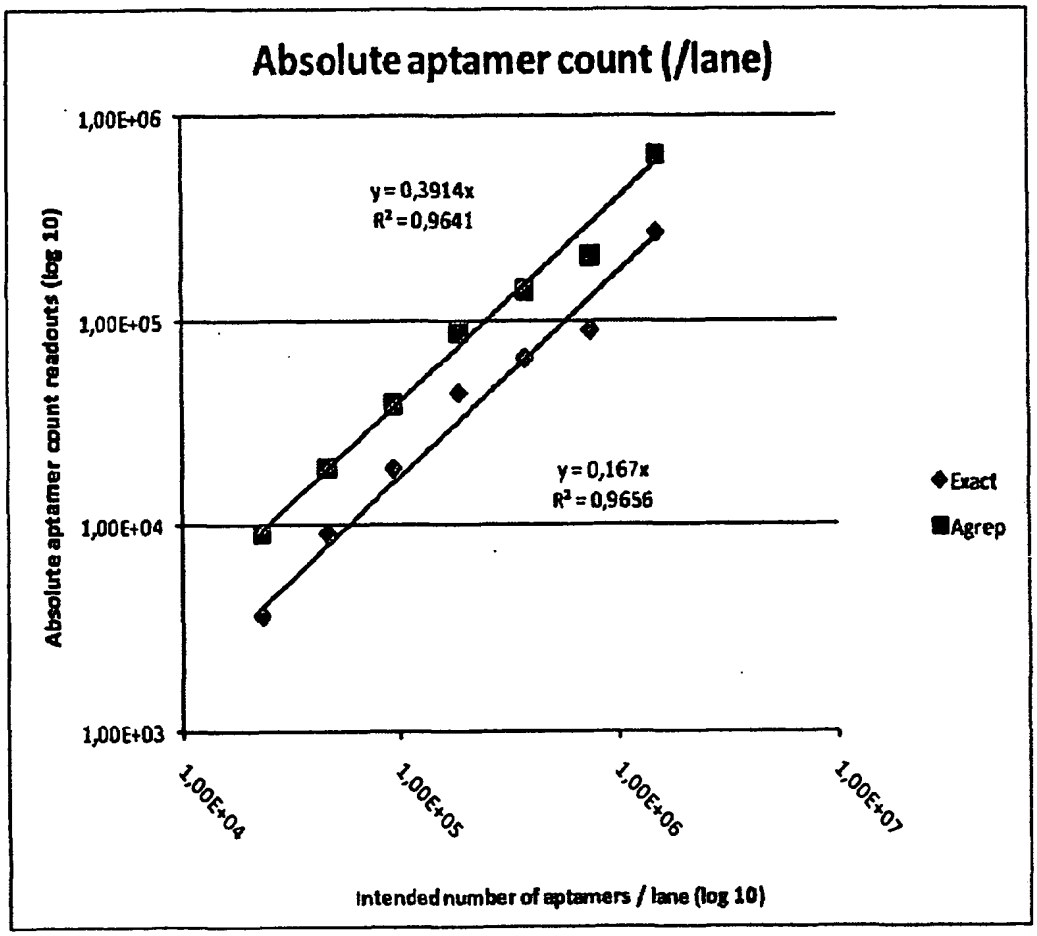


Figure 8

PREOMICS EXHIBIT 1002
Page 358 of 1861

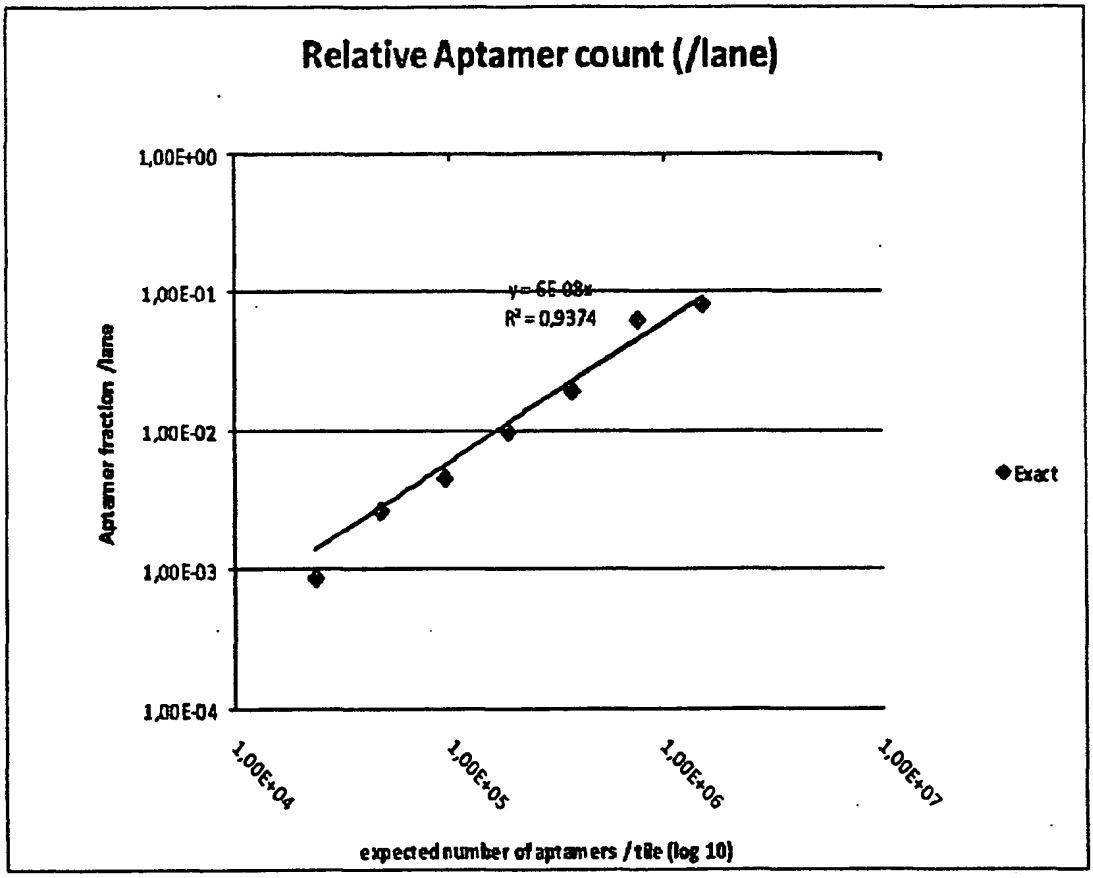


Figure 9

PREOMICS EXHIBIT 1002
Page 359 of 1861

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

- US 20070166741 A **[0021]**
- WO 00006770 A **[0031]**
- EP 07020629 A **[0035]**

Non-patent literature cited in the description

- **SCHRATTENHOLZ A ; GROEBE K.** *Electrophoresis.,* June 2007, vol. 28 (12), 1970-9 **[0003]**
- **QIAN W.J. et al.** *Mol. Cell. Prot.,* 2006, vol. 5 (10), 1727-1744 **[0005] [0009]**
- **LIU H ; SADYGOV RG ; YATES JR 3RD.** *Anal. Chem.,* 15 July 2004, vol. 76 (14), 4193-201 **[0007]**
- **ZOLG.** *Mol. Cell. Prot.,* 2006, vol. 5 (10), 1720-1726 **[0011] [0012]**
- **RIFAI et al.** *Nature Biotech,* 2006, vol. 24 (8), 971-83 **[0012]**
- **ANDERSON ; HUNTER.** *Mol. Cell Prot.,* 2005, vol. 5 (4), 573-588 **[0013]**
- **ZANGAR et al.** *Expert Review of Proteomics,* 2006, vol. 3 (1), 37-44 **[0013]**
- **BRANTON et al.** *Nature Biotechnology,* 2008, vol. 26, 1146-1153 **[0031]**
- **SMYTH GK.** *Bioinformatics,* 19 September 2007 **[0034]**
- **DRABOVICH et al.** *Anal. Chem.,* 2006, vol. 78, 3171-3178 **[0036]**
- **WIEGAND et al.** *Journal of Immunology,* 1996, vol. 157, 221-230 **[0084]**
- **KRYLOV et al.** *Journal of the American Chemical Society,* 2002, vol. 124, 13674-13675 **[0100]**
- *Analytical Chemistry,* 2003, vol. 75, 1382-1386 **[0100]**

PREOMICS EXHIBIT 1002
Page 360 of 1861

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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
2 September 2010 (02.09.2010)

PCT

(10) International Publication Number
WO 2010/097785 A1

(51) International Patent Classification:
G01N 33/543 (2006.01)

(21) International Application Number:
PCT/IE2010/000009

(22) International Filing Date:
26 February 2010 (26.02.2010)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/155,677 26 February 2009 (26.02.2009) US
EP09153864 27 February 2009 (27.02.2009) EP

(71) Applicant *(for all designated States except US)*: **UNIVERSITY COLLEGE DUBLIN, NATIONAL UNIVERSITY OF IRELAND, DUBLIN** [IE/IE]; Belfield, Dublin 4 (IE).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **DAWSON, Kenneth** [IE/IE]; c/o NovaUCD, Belfield Innovation Park, University College Dublin, Belfield, Dublin 4 (IE). **LYNCH, lseult** [IE/IE]; c/o NovaUCD, Belfield Innovation Park, University College Dublin, Belfield, Dublin 4 (IE). **LUNDQVIST, Martin** [SE/IE]; c/o NovaUCD, Belfield Innovation Park, University College Dublin, Belfield, Dublin 4 (IE). **CEDERVALL, Tommy** [SE/IE]; c/o NovaUCD, Belfield Innovation Park, University College Dublin, Belfield, Dublin 4 (IE).

(74) Agent: **PURDY, Hugh Barry**; PURDYLUCEY INTELLECTUAL PROPERTY, Suites 138/139, The Capel Building, Mary's Abbey, Dublin 7 (IE).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *with international search report (Art. 21(3))*

— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*

WO 2010/097785 A1

(54) **Title**: A METHOD FOR THE SELECTIVE CONCENTRATION OF A SPECIFIC LOW ABUNDANCE BIOMOLECULE

(57) **Abstract**: Provided herein is a method for the isolation or removal of a cellular component from a cell that comprises the steps of applying a pulse of nanoparticles to the cell, allowing the nanoparticles to traffic through the cell for a period of time sufficient to allow the nanoparticles locate to and interact with the cellular component to be isolated, and separation of the nanoparticles and isolated cellular component from the cell.

A METHOD FOR THE SELECTIVE CONCENTRATION OF A SPECIFIC LOW ABUNDANCE BIOMOLECULE

Technical Field

The invention relates to methods for the selective concentration, isolation or removal of specific biomolecules or biomolecule clusters, especially low abundance protein(s), from biological fluids and biological systems such as cells. The invention also relates to methods for the recovery or purification of low abundance biomolecules, and methods for the detection of biomarkers in biological fluids / biological systems.

Background to the Invention

Interactions between single proteins or simple mixtures of 2 or 3 proteins can be easily understood and predicted, based on charge or hydrohobicity interactions arguments, and indeed this would be considered intuitive for someone skilled in the field. Thus, to isolate a positive protein from a mixture containing 1 positive and 1 negative protein, one would clearly use a negative surface to attract the positive protein selectively. However, how this simple "model system" behaviour translates to real world systems such as the complex protein mixtures present in biological fluids in not at all obvious. As an example, human plasma contains 3,700 different proteins, each with different charge distributions and densities, different folding stabilities, and different hydrophobic/hydrophillic balances, and selectively targeting these cannot be predicted based on interaction studies on single proteins or simple protein mixtures. Here, many more features of the interaction come into play, such as protein abundance, binding affinity (related to protein stability), tendancy to bind to surfaces in general (e.g. albumin binds to all surfaces but has low specificity so is rapidly replaced by other proteins) and finally the electrostatic and hydrophobic characteristics will play a role. The higher abundance biomolecules, or those with some other simple basis of affinity will attach to the particle surface rapidly. However, in a complex mixture, it is found that over time (often many hours) these are displaced, and that the mutual crowding effects of many other biomolecules in a highly curved surface itself drives a highly selective effect, leading to a final 'hard corona'. This mutual crowding effect can be enhanced by additional surface structuring of the nanoparticle.

PREOMICS EXHIBIT 1002
Page 362 of 1861

Within complex biological fluids, terms such as low and high abundance are relative terms, and refer to the amount of a specific protein relative to the total amount of protein in the fluid. Thus, for example, albumin has a high abundance in plasma, accounting for almost 60% of the total protein content, whereas apolipoprotein E is a fairly low-abundance protein (2.32±0.10 μg apoE/mg total protein). A very low abundance protein (rare) would be, for example, a biomarker or a protein secreted in response to a perturbation.

The same is true for the biomolecules in cells, which are very tightly regulated and typically contained in specific organelles until needed. Some of the organelle-specific proteins would be high abundance, and others low or extremely low abundance. Biomarkers are typically released in response to a trigger or perturbation at the onset of a disease, and as such are usually very low abundance and often localised to a specific organelle.

Statements of Invention

Broadly, the invention relates to a method for the selective concentration of a cellular component, for example at least one specific low abundance biomolecule, biomolecule complex or cluster, or organelle, from a complex biological system such as a biological fluid or a cell, typically a complex biological system including one or more high abundance biomolecules. The method suitably comprises a step of providing a preparation of nanoparticles in which at least one physiochemical property of the nanoparticle surface is selectively modified to concentrate or more selectively bind at least one specific low abundance biomolecule, incubating the nanoparticle preparation with the complex biological system to enable the cellular component, for example at least one specific low abundance biomolecule, bind to the modified surface of the nanoparticle, and separating the bound and unbound components. In a preferred embodiment, the method involves an additional step of eluting (identifying or recovering) the at least one specific cellular component from the surface of the nanoparticles. Many low abundance proteins exist within functional clusters of biomolecules, and recovery of these functional biomolecule clusters intact by selectively binding to nanoparticles is also included in the invention.

The fact that the nanoparticle size is small also allows nanoparticles access to potential extraction sites not previously easily accessible, and thereby enables isolation procedures to be applied *in situ* to systems not previously considered. Subsequent removal of such particles can be achieved using magnetic field (nanoparticles with magnetic cores), affinity tags, or a

PREOMICS EXHIBIT 1002
Page 363 of 1861

variety of other approaches. For example, nanoparticles can gain access to biological cells, and a wide variety of organs in animals, via a variety of biologically regulated and other pathways. Intracellular or organ specific biomolecules (or small assemblies of such biomolecules), otherwise highly inaccessible, in low abundance and possessing low stability (decaying rapidly after cell or organ tissue lysis or other treatment) may thereby be accessed for the first time, allowing for isolation, and identification. This allows for both harvesting, and identification of such molecules, and assemblies. Access to such biomolecule assemblies can provide a new route to diagnostics by rapid extraction of rare intra-cellular or other biomarkers that are too unstable, or in too low abundance to be extracted and identified by classical methods. Thus, the method is also applicable to recovery of biomolecules from within specific sub-cellular organelles, or indeed to recover a specific organelle (collection of biomoleucles) from the cell. Thus, for example, selective binding to and recovery of apoptosomes, which are formed in cells undergoing programmed cell death, and contain many of the rare signalling proteins that control this cellular process, is included in the invention.

Thus, the invention also relates to a method for the isolation of a cellular component from a cell, for example a cellular component from a specific location in the cell (i.e. a lysosome), comprising the steps of applying a pulse of nanoparticles to the cell for a suitable period of time, allowing the nanoparticles to traffic through the cell for a period of time sufficient to allow the nanoparticles locate to a specific location within the cell and interact with the cellular component to be isolated, and separating the nanoparticles and isolated cellular component from the cell.

The method involves applying a pulse of nanoparticles to the cell. This generally means that the nanoparticles are incubated with the cells for a limited period of time to enable a pulse of nanoparticles enter the cell and begin trafficking around the cell as a discrete pulse or packet. This is somewhat similar to a train, in which the individual nanoparticles are similar to the carriages of the train, insofar as the nanoparticles in the pulse will travel throughout the cell on a defined trafficking route and in a group. If the pulse of nanoparticles is too long, the train will be too long, and nanoparticles will be located in different subcellular compartments at the same time. Thus, a limited pulse of nanoparticles will charge the cells with a discrete packet of nanoparticles, which will travel throughout the cell substantially together. Thus, at any given time, the nanoparticles will generally be primarily located at a single location (for example, the endosomes, or lysosomes). As the nanoparticles traffic throughout the cell, the biomolecule corona of the nanoparticles will change depending on the location within the cells

PREOMICS EXHIBIT 1002
Page 364 of 1861

of the nanoparticles at that point of time. This has enabled the Applicants to apply a pulse of nanoparticles to a cell, allow the nanoparticles traffic to a specific location within the cell where a cellular component is located (for example an organelle-specific low abundance protein), and then withdraw the nanoparticles (with the cellular component bound thereto) from the specific location. The nanoparticles may for example bind a panel of biomolecules, for example proteins, from the specific location, including for example a protein of interest which will therefore be isolated from the cell. The method of the invention also enables the isolation of biomolecule clusters, and organelles, from cell.

The period of time required to achieve a nanoparticle pulse depends on the nanoparticles, the physiochemical modification (if any) made to the nanoparticles, and the type of cells. Suitably, the pulse time (import incubation time) is less than 30, 25, 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2 or 1 minute.

The nanoparticles are typically dispersed in cell culture medium prior to being pulsed into cells. Separating or removing the nanoparticles from the cell may be achieved in a number of different ways. For example, the nanoparticles may have a magnetic core, and removal from the cell may be effected by use of magnetic fields. Alternatively, the nanoparticles may be provided with affinity tags, which facilitate removal of the nanoparticles from the cell or a cell extract. The nanoparticles may also be labelled with a detectable marker, be it fluorsecent, radioactive, or the like, and as such the method of the invention enables detection and quantification of biomolecuels of interest.

The cell or cells are generally lysed prior to extraction.

The cellular component is suitably an organelle, or a biomolecule, for example a protein, a polypeptide, a peptide, a polysaccharide, a sugar, a lipid, a lipoprotein, a metabolite, an oligonucleotide, or then like, or complexes or clusters of such molecules. Suitably, the biomolecule is a low abundance biomolecule. Typically, the biomolecule is a protein, such as a low abundance protein, a biomolecule complex or cluster, for example a protein complex or cluster, or a whole organelle. The term low abundance biomolecule should be understood to mean a biomolecule which is present in low amounts relative to the total amount of that type of biomolecule. Thus, for a specific protein in a biological fluid such as blood serum, the term low abundance typically means that the specific protein makes up less than 10%, preferably 5%, and more preferably less than 1%, ideally less than 0.1%, of the total protein in the serum

PREOMICS EXHIBIT 1002
Page 365 of 1861

(conc./conc.). Likewise, a high abundance protein would generally comprise at least 30%, preferably at least 50%, of the total protein in serum.

Typically, at least one physiochemical property of the nanoparticle surface is selectively modified to concentrate or more selectively bind at least one specific low abundance biomolecule. Examples of suitable physiochemical modifications are provided below.

The at least one physiochemical property of the nanoparticle surface that is selectively modified is typically selected from the group consisting of: surface curvature; surface charge; surface chemistry; surface functionalisation; and controlled surface morphology.

Generally, surface curvature is determined by the nanoparticle size. Thus, at a nanometre scale, as the size of the nanoparticle changes, the surface curvature of the particle changes, and this change of surface curvature affects the binding selectivity of the surface. For example, at a certain curvature, the surface of the particle may have a binding affinity for a specific type of biomolecule, whereas a different curvature will correlate with a binding affinity for a different biomolecule. Thus, the surface curvature may be tuned to facilitate selective binding (and therefore selective concentration) of specific biomolecules. Further tuning of the surface of the nanoparticle may be achieved by selective modification of one or more of the other physiochemical characteristics mentioned above.

In this specification, the term "selective concentration" should be understood as meaning that the modified nanoparticles are capable of selectively concentrating the specific low abundance biomolecule to a level of concentration that is greater than any other biomolecule concentrated on the nanoparticle. Thus, while the modified nanoparticle may selectively bind, and concentrate, other biomolecules present in the complex mixture of biomolecules, the specific low abundance biomolecule will be the most concentrated.

The method can also be considered as an iterative process, whereby a first pass is made to concentrate the 3,700 plasma proteins to a more limited panel of 20 proteins including the proteins(s) of interest, and a second step in which the protein of interest is further concentrated with respect to the other 19 proteins using the same or another (more tailored) nanoaprticle.

Thus, the invention also provides a method for the selective concentration of a specific biomolecule, typically a specific low abundance biomolecule, from a complex biological

PREOMICS EXHIBIT 1002
Page 366 of 1861

system comprising a multiplicity of biomolecules, comprising the steps of incubating a first nanoparticle preparation with the complex mixture of biomolecules to enable a plurality of biomolecules including the at least one specific biomolecule bind to the nanoparticle, separating the bound biomolecules including the specific biomolecule of interest from the unbound biomolecules, eluting the bound biomolecules from the nanoparticle preparation to provide a concentrated panel of biomolecules, and incubating the concentrated panel of biomolecules with at least one further nanoparticle preparation to further concentrate the specific biomolecule. Suitably, the at least one further nanoparticle preparation is different to the first nanoparticle preparation, for example different size, polymer, charge, hydrophobicity, surface curvature etc.

In cases where there is no or limited a priori knowledge of the nature of the biomolcule to be recovered, or its binding affinity, a panel of nanoparticles with different physicochemical and surface properties could be screened initially, and based on the most promising lead candidate (as identified by, for example, 1D SDS PAGE if the MW of the protein of interest is know, or by mass spectormtry) a more tailored nanoparticle selected for specific binding and concentration of the biomolecule of interest and a second recovery step performed. Thus, it is clear that the choice of nanoparticles physico-chemical and surface parameters may need to be optimised on a case-by-case basis for recovery of a specific biomolecule from a complex mixture contain many hundreds of different biomolecules.

Thus, the invention also provides a method for providing or identifying a nanoparticle suitable for the selective concentration of a specific biomolecule, typically a specific low abundance biomolecule, from a complex biological system comprising a complex mixture of biomolecules (for example a cell or a biological or non-biological fluid), comprising the steps of incubating the complex mixture of biomolecules with a plurality of nanoparticle preparations, each preparation differing in at least one physico-chemical or surface curvature parameter, identifying a lead nanoparticle preparation that provides the greatest concentration of the specific biomolecule of interest, and further modifying at least one physico-chemical and/or surface curvature parameter of the lead nanoparticle to tune the selectivity of the lead nanoparticle.

The terms complex biological system or complex mixture of biomolecules should be under stood to include cells, cell cultures, fermentation broths, peroducer cell cultures, cell lysates, biological fluids, serum, blood, plasma, urine, saliva etc. Generally, the terms implies that the

PREOMICS EXHIBIT 1002
Page 367 of 1861

system or mixture includes at least 50, 100, 200, 300, 400, 500, 600, 800, 900, 1000, 1100, 1200, 1300, 1400, 1500, 1600, 1700, 1800, 1900, or 2000 different biomolecules.

A more thorough screening approach is also outlined, whereby in Step 1 the biomolecule to be recovered is spotted onto dot-blots and used to screen several (100s) of different nanoparticles dispersed in PBS buffer in order to identify a candidate particle that binds strongly to the biomolecule to be recovered. Step 2 then involves using protein arrays containing 1000s of different proteins to determine that the particle does not bind significant amounts of other proteins that would compete for the particle surface with the biomolecule of interest. Step 3 is then to utilise the nanoparticles selected to recover the biomolecule from the complex mixture of which it is a small component, and then to recover the biomolecule, A final confirmation of the identity of the biomolecule can be performed again using any of the techniques listed below, or again via protein arrays.

Thus, the invention provides a further method of identifying a nanoparticle suitable for the selective concentration of a specific biomolecule, typically a specific low abundance biomolecule, from a complex biological system comprising a complex mixture of biomolecules, comprising the steps of reacting the specific biomolecule with a plurality of different nanoparticles to identify one or more lead nanoparticle that binds to the specific biomolecule most strongly, incubating the at least one thus identified lead nanoparticle with a plurality of biomolecules different to the specific biomolecule to assess non-specific binding, wherein the at least one lead nanoparticle having low non-specific binding is a nanoparticle suitable for the selective concentration of the specific biomolecule.

Surface morphology may be selectively modified by a method selected from the group consisting of: patterning the surface to provide areas of differing biomolecule affinity; introduction of porosity of different scales; engineered surface curvature on multiple length scales; and templating the surface with a template of a specific biomolecule.

Patterning of the surface is provided by a means selected from the group consisting of: forming the nanoparticle by block polymerisation in which the at least two blocks have different chemistries; and forming the nanoparticle using mixtures of at least two different polymers, phase separating the different polymers during polymerisation, and cross-linking the separated polymers following phase separation. Patterning may result in a first portion of the surface having a high binding affinity for a specific biomolecule, and a second portion of

PREOMICS EXHIBIT 1002
Page 368 of 1861

the surface having no or little binding affinity for the specific biomolecule and specific cell type. Thus, when the specific biomolecule binds to the surface of such a patterned nanoparticle, the effects is that the surface of the nanoparticle becomes patterned with the specific biomolecule. This clearly affects the surface morphology of the nanoparticle, but could also have the effect of modifying the surface curvature.

Surface curvature on different length scales

Engineered surface curvature on multiple length scales is provided, for example, by employing Pickering emulsions (Sacanna *et al.*, 2007) stabilised by finely divided particles for the synthesis of the nanoparticles. Typically, the finely divided particles are selected from the group consisting of but not limited to: silicates; aluminates; titanates; metal oxides such as aluminum, silicon, titanium, nickel, cobalt, iron, manganese, chromium, or vanadium oxides; carbon blacks; and nitrides or carbides, such as boron nitride, boron carbide, silicon nitride, or silicon carbide. Other suitable examples of finely divided particles suitable for Pickering emulsions will be known to those skilled in the art.

Templating the surface of the nanoparticle is provided, for example, by forming the nanoparticle in the presence of a biomolecule, wherein at least a portion of the biomolecule(s) will be exposed/embedded on a surface of the formed nanoparticle, and removing the biomolecules exposed/embedded on the surface of the nanoparticle to leave a template or imprint of the biomolecule on the surface of the nanoparticle. Such templated or imprinted nanoparticles will have a binding affinity for the biomolecules that are employed to make the template or imprint on the surface. Thus, they may be employed in the selective concentration of the template biomolecules, or in the selective concentration of a binding partner of the template biomolecule. In this specification, the term "template biomolecule" should be understood to mean the biomolecules that are employed to form a template, or imprint, on the surface of the nanoparticles.

The physiochemical properties of the nanoparticle surface may be modified by modification of the surface charge. For example, the surface charge of the nanoparticle surface may be modified to provide a controlled net positive surface charge, a controlled net negative surface charge, or a controlled zwitterionic charge. Typically, the surface charge on the surface is controlled at synthesis, although post-synthesis modification of the charge via surface functionalisation is also included, as described below. For example, in polymeric nanoparticles this may be achieved by use of different synthetic procedures (initiators, etc),

PREOMICS EXHIBIT 1002
Page 369 of 1861

different charged co-monomers, and in inorganic substances this may be achieved by different reaction conditions or formation of surface composites, as for example with mixed oxidation states.

The physiochemical properties of the nanoparticle surface may be modified by functionalisation of the surface. The surface functionalisation of the surface of the nanoparticles is preferably selected from the group consisting of: functionalisation with charged groups; functionalisation with groups that alter the hydrophobic or hydrophilic balance of the nanoparticle surface; and functionalisation with a ligand. In one embodiment, the ligand is selected from the group consisting of: oligopeptides (for example epitopes); sugars; nucleic acids and oligonucleotides (DNA, RNA, and mixtures thereof), antibodies or antibody fragments, whole proteins, protein complexes, protein-lipid complexes, enzymes etc.

Thus, by modification of a combination of the surface curvature and one or more of the physiochemical properties referenced above, the surface of the nanoparticles may be tuned for the selective concentration of any one of a vast number of different biomolecules. Thus, in one preferred embodiment, the surface curvature and at least one other physiochemical property of the nanoparticle surface is selectively modified to selectively bind the at least one specific low abundance biomolecule. In another embodiment, the surface curvature and at least two other physiochemical properties of the nanoparticle surface are selectively modified to selectively bind the at least one specific low abundance biomolecule.

Generally, the methods of the invention involve the selective concentration of a specific low abundance biomolecule, a collection of low abundance biomolecules that function as a unit (e.g. apolipoprotein complexes, apoptosomes, inflamasomes and others produced in response to a specific perturbant etc.), or site-specific biomolecules from within a cell or organism, such as lysosomal proteins. In one embodiment, the selective concentration of the biomolecule is carried out for the purpose of detecting the presence of the biomolecule in a sample, for example a biological fluid sample or a cellular organelle. Thus, the invention may be a method of detecting a low abundance biomolecule, such as for example a diagnostic or prognostic biomarker, in which the nanoparticles are modified for the selective concentration of the biomarker, and in which the biomarker is selectively concentrated and then detected after concentration, for example on a protein array with appropriate binding targets for the selected biomolecule. Various other techniques for detection of concentrated biomolecules will be known to those skilled in the art, such as for example, proteomics and mass spectrometry,

PREOMICS EXHIBIT 1002
Page 370 of 1861

chromatography, electrophoresis, and immunoprecipitation, as well as optical and fluorescence techniques such as could be utilised in point-of-care devices. Selective concentration generally involves a step of separating unbound and bound biomolecules (although this is not always required), and this may be achieved by means of a number of different separation technologies. Typically, separation is achieved by means of washing or by centrifugation. Preferably, the unbound biomolecule is separated by means of a plurality of centrifugation steps, or by a change in, for example, buffering conditions in the case of fixed or surface-based separation approaches. Array-based detection approaches, such as protein and antibody arrays, are also included, for which the existing uses for detection of protein-protein interactions are extended to include particle-bound-protein interactions with proteins or antibodies on the array (particle-protein-protein interactions) and subsequent detection by fluorescent or other methods.

In one embodiment of the invention, the selectively concentrated (bound) biomolecule is recovered or harvested from the nanoparticles. Typically, the bound biomolecule is recovered by means of centrifugation and separated ideally by PAGE. In one embodiment, there will be a plurality of bound biomolecules including the selectively concentrated biomolecule, in which case the bound biomolecules may be recovered in a step-wise process involving a plurality of centrifugation steps. Thus, each centrifugation step may provide a different fraction of the bound biomolecule or of the at least one specific low abundance biomolecule, which can be determined by SDS PAGE or by interactions studies with protein or antibody arrays and fluorescence detection.

In one embodiment of the invention, the steps of separation of unbound protein, and elution of bound protein, are carried out by elution in different buffers.

In one embodiment of the invention, the nanoparticle comprises a core and a separate shell encapsulating the core, wherein the core and shell are capable of being modified independently. This enables various modifications of the physiochemical properties of the surface (through modification of the shell) without having to modify the characteristics of the core. This provides a specific advantage for nanoparticle separation, where it may be desirable to form the core of a material that facilitates separation (for example, by means of the core being formed of a high density, or a magnetic, material), or for detection of binding where the core could be a fluorescent molecule whose signal changes upon binding of the selected biomolecule.

PREOMICS EXHIBIT 1002
Page 371 of 1861

In one embodiment, the nanoparticles are used to recover proteins from specific locations within a cell (specific organelles), organ or biological organism to which the nanoparticles preferentially locate. In this embodiment, nanoparticles are exposed to the biological system for a pre-determined time in a pre-determined manner in order to be taken up by the biological system and localised in the desired location, following which the nanoparticles are recovered (e.g. by lysing the cells and separating out the nanoparticles, or using an organelle separation strategy, or other approaches based on magnetic recovery etc.) in order to recover the organelle-specific proteins that were selectively pulled-down by the nanoparticles from that location.

In one embodiment of the invention, a given drug or nanoparticle is used to induce a functional response that results in formation of new functional biomolecule assemblies in cells, organs or animals, such as inflammation, apoptosis etc. which may then be selectively extracted using selective binding to nanoparticles applied as a short pulse at the appropriate time-point in the biological response.

In one embodiment, the nanoparticles are provided in the form of a solid phase during at least a portion of the process for selective concentration. The purpose of providing the nanoparticles in a solid phase during at least a part of the process is to facilitate removal of unbound biomolecule from bound biomolecule. In this regard, the solid phase may be provided by the nanoparticles being tethered to a support (for example a polystyrene support), or by the nanoparticles being crosslinked together, for example, by means of gellation.

Thus, in one embodiment, the nanoparticles are capable of being crosslinked, for example by gellation, under specific conditions or upon activation. This enables the nanoparticles to be crosslinked to facilitate the removal of unbound biomolecules. In one embodiment, the nanoparticles are capable of reversible crosslinking. Typically, the nanoparticles are modified by a method selected from the group consisting of: thiol group modification; pH induced crosslinking via hydrogen bonding (modification with -COOH and –OH groups); modification with lysine and arginine residues with subsequent production of methylene bridges (crosslinks) upon introduction of specific initiators such as formaldehyde; complementary biomolecules capable of hybridisation under specific conditions (.ie. DNA strands or other coil-forming polymers such as gelatin, chitosan and the like).

PREOMICS EXHIBIT 1002
Page 372 of 1861

In one preferred embodiment of the invention, gellation of the nanoparticles is reversible under specific conditions or upon activation. Thus, for example, the gelled nanoparticle preparation may be treated to reverse the gelation after removal of unbound biomolecule, whereupon the at least one specific low affinity biomolecule is optionally eluted from the nanoparticles and recovered in a purified and concentrated format. Gellation of the nanoparticle preparation may be activated, or reversed, by means of a specific, predetermined, cue, such as, for example, a change in temperature, pH, ionic strength, solvent, buffer, catalyst, oxidation state and the like.

The invention also provides a method for the selective concentration of at least one specific low abundance biomolecule from a complex mixture of biomolecules, typically including one or more high abundance biomolecules, comprising the steps of providing a preparation of nanoparticles in which the nanoparticles are capable of being crosslinked under specific conditions, incubating the nanoparticle preparation with the complex mixture of biomolecules to enable the at least one specific low abundance biomolecule bind to a surface of the nanoparticle, crosslinking the nanoparticle preparation either prior to, during, or after the incubation step, separating unbound biomolecule from the crosslinked nanoparticle preparation, and optionally eluting the at least one specific low abundance biomolecule from the surface of the nanoparticles.

Typically, the gellation of the nanoparticles is reversible, and wherein after removal of the unbound biomolecule, the nanoparticle preparation is treated to reverse the gelation whereupon the at least one specific biomolecule is optionally eluted from the nanoparticles.

Generally, the nanoparticles forming the nanoparticle preparation are selected from the group consisting of: natural or synthetic polymers; copolymers; and terpolymers (with the cores being composed of metals or inorganic oxides, including magnetic cores). As appropriate, the polymeric nanoparticles are selected form the group consisting of: polystyrene; poly(lysine); chitosan; dextran; poly(acrylamide), and its derivatives such as N-isopropylacrylamide, N-tert-butylacrylamide, N,N-dimethylacrylamide; Polyethylene glycol; poly(vinyl alcohol); gelatine; starch; and degradable (bio)polymers or silica.

In one embodiment, the complex mixture of biomolecules is a biological fluid, suitably selected from the group consisting of: plasma; serum; cell lysates; cytosolic fluid; gastric fluid; amniotic fluid; cerebrospinal fluid; lung lavage fluid; saliva; urine; or cell organelles;

PREOMICS EXHIBIT 1002
Page 373 of 1861

cells, organs, organisms or *in vivo* in animals or humans. However, the term should also be taken to include non-biological fluids that may contain low abundance biomolecules, or low abundance molecules, of interest. For example, the process of the invention may be employed to selectively concentrate low abundance, high value, bio-metabolites, for example high value metabolites obtainable from an industrial process such as bioethanol production.

Preferably, the specific low abundance biomolecule is a protein, polysaccharide, lipid or functioning until of these such as a lipoptotein complex or a "some" such as a lysosomes, endosomes, apoptosome, centrosome, etc. However, the method of the invention may also be employed to selectively concentrate other specific low abundance biomolecules including nucleic acid molecules (DNA, RNA, mRNA etc.), peptides, sugars, metabolites, and the like. Additionally, in certain circumstances, the processes of the invention may be employed to selectively concentrate low abundance molecules that are not from a biological source, such as a byproduct from an industrial process.

The invention also relates to a method of identifying the presence of a specific low abundance biomolecule in a sample such as a complex mixture of biomolecules, which method comprises a step of selectively concentrating the specific low abundance biomolecule using the method of the invention, and identifying the concentrated low abundance biomolecule. The low abundance biomolecule is typically identified using any suitable technique, for example proteomics and mass spectrometry, electrophoresis and mass spectrometry, or electrophoresis and western blotting, or alternatively, via fluorescence read-outs such as could be incorporated into point-of-care diagnostic devices, including for example, protein or antibody arrays.

In one embodiment, the low abundance biomolecule is a biomarker associated with a characteristic selected from the group consisting of: clinical assessment (including diagnosis or prognosis) of a disease, condition, or pathological status; drug response; adverse drug reaction; and response to treatment, either in plasma, cell lysate, cell cytosolic fluid, urine, saliva or recovered from a specific cellular organelle.

The invention also relates to the use of pulses of nanoparticles applied to cells (or organisms or animals) to induce an effect in a cell, and/or to track the cellular response to the presence of the nanoparticles by selectively binding, concentrating and recovering the biomarkers of response. Controlling the pulse length of the delivery of nanoparticles to cells controls where the particles will be along the cellular uptake and trafficking pathway, enabling recovery of

PREOMICS EXHIBIT 1002
Page 374 of 1861

proteins at a specific time-point and location along the trafficking pathway, such as selective recovery of proteins from the early endosome, the sorting endosomes, the late endosomes or the lysosomes, depending on the pulse length, and the time that has elapsed since the application of the pulse.

The invention also relates to a method for the purification of a specific low abundance protein comprising a step of selectively concentrating the specific low abundance biomolecule using the method of the invention, and eluting the low abundance biomolecule from the nanoparticle preparation. In one embodiment, elution is carried out in a step-wise manner.

The invention also relates to a method of preparing a nanoparticle preparation suitable for selectively concentrating a specific low abundance biomolecule comprising the step of modifying the surface curvature of the nanoparticle such that it has a specific binding affinity for the specific low abundance biomolecule. Preferably, the method includes a further step of selectively modifying a physiochemical property of the surface of the nanoparticle to concentrate or more selectively bind the at least one specific low abundance biomolecule, wherein the physiochemical property is selected from the group consisting of: surface charge; surface chemistry: surface functionalisation: and controlled surface morphology.

Suitably, surface morphology is selectively modified by a method selected from the group consisting of: patterning the surface to provide areas of differing biomolecule affinity; engineered surface curvature on multiple length scales; and templating the surface with a template of a specific biomolecule (as described above).

The invention also relates to a nanoparticle preparation that is modified to be capable of gelation under specific conditions. Typically, the nanoparticles are modified with a reactive group that, upon activation (or in response to a specific external cue or stimulus), is capable of cross-linking the nanoparticles to form a solid phase, typically a gel. In one embodiment, the reactive group is capable of activation in response to a condition or stimulus selected from the group consisting of: temperature; pH; ionic concentration; solvent; buffer; catalyst: oxidation state; and the like. Suitably, the reactive group is selected from the group consisting of: thiol group; -COOH and / or –OH groups; amino acid (or amino acid analogues) residues capable of forming of methylene bridges (for example lysine and arginine residues); complementary DNA strands, or other coil-forming polymers such as gelatin, chitosan, capable of hybridisation; and the like.

PREOMICS EXHIBIT 1002
Page 375 of 1861

The invention also relates to a nanoparticle preparation in which a surface of at least a portion of the nanoparticles is imprinted with a template of a specific biomolecule. This means that the surface of the nanoparticle comprises a hole or detent where the biomolecule was embedded, which hole or detent may retain some of the tertiary or secondary structural charactersiatics of the imprinted molecule, and which will preferentially bind that biomolecule in the same conformation. Such nanoparticles may be prepared by forming the nanoparticles in the presence of the specific biomolecule (in a specific conformation), wherein at least a portion of the formed nanoparticles have the specific biomolecule exposed/embedded in a surface of the nanoparticle, and treating the formed nanoparticles to remove the exposed/embedded biomolecule to leave an imprint of the specific biomolecule on the surface of the nanoparticle. In one preferred embodiment, the specific biomolecule is a protein, and wherein the embedded protein is removed from the formed nanoparticle by denaturation of the protein using for example SDS and subsequent washing.

The invention also relates to a nanoparticle having a patterned surface in which a first portion of the surface is modified to have a specific binding affinity for a specific patterning biomolecule, and a second portion of the surface is modified to have a different (typically less, or preferably little or no) binding affinity for the same specific patterning biomolecule compared with the first portion of the surface. Typically, the (patterned) nanoparticle is formed of a block co-polymer in which the different blocks have different binding affinities (for example, due to having different chemistries) for the specific biomolecule. The invention also relates to a patterned nanoparticle comprising a nanoparticle of the invention having a patterned surface and which is patterned with the specific patterning biomolecule, or a biomolecule that binds to the specific patterning biomolecule.

The invention also relates to a nanoparticle having engineered surface curvature on multiple length scales. Typically, the engineered surface curvature on multiple length scales is provided by employing Pickering emulsions stabilised by finely divided particles for the synthesis of the nanoparticles and polymerising these structures in place. Suitably, the finely divided particles are selected from the group consisting of but not limited to: silicates; aluminates; titanates; metal oxides such as aluminum, silicon, titanium, nickel, cobalt, iron, manganese, chromium, or vanadium oxides; carbon blacks; and nitrides or carbides, such as boron nitride, boron carbide, silicon nitride, or silicon carbide.

PREOMICS EXHIBIT 1002
Page 376 of 1861

The invention also relates to the use of a nanoparticle of the invention in the selective concentration of a specific low abundance biomolecule.

Nanoparticles are also employed to bind a selected protein during high throughput / recombinant protein expression (in e-coli for example as used for protein therapies), where a key challenge is that the high concentration of the expressed protein results in protein aggregation which is undesirable. The nanoparticles can bind the protein thereby reducing the local concentration and facilitating overall recovery of correctly-folded proteins at high concentration.

Thus, the invention also provides a method of producing a recombinant biomolecule such as a protein, which method employs a eukaryotic or prokaryotic producer cell engineered with a nucleic acid construct encoding the recombinant protein, the method comprising the steps of incubating the producer cells with a nanoparticle preparation suitable for selectively binding the recombinant protein, and recovering/separating the nanoparticle preparation and bound recombinant protein from the producer cells.

Suitably, the nanoparticle preparation is incubated with the producer cells during the production phase of cell growth. Typically, the method includes a step of eluting the recombinant protein from the nanoparticle preparation.

The invention also relates to a method of reducing recombinant protein aggregation during high-throughput recombinant protein production, the method comprising the step of incubating a nanoparticle preparation with the recombinant protein producer cells for a period of time, which nanoparticles are suitable for selective binding and/or concentration of the recombinant product, and recovering/separating the nanoparticle preparation and the bound/concentrated recombinant protein from the producer cells.

Brief Description of the Figures

Figure 1 is a schematic illustration of the use of surface curvature to selectively bind specific proteins, based on efficiency of packing at the surface.

PREOMICS EXHIBIT 1002
Page 377 of 1861

Figure 2 is a schematic illustration of the use of surface charge to add further selectivity beyond selection by packing parameters for selective biomolecule recovery and purification.

Figure 3: Left: Illustration of the protein concentrating effect of nanoparticles. Right: Comparison of Lanes a and b in the SDS-PAGE gel shows the concentrating effect of 9nm silica nanoparticles on plasma proteins. The more highly stained bands in lane b indicate significantly increased protein concentrations compared to the plasma in the absence of nanoparticles shown in lane a.

Figure 4 is a schematic illustration of the use of charged nanoparticles to selectively harvest proteins by controlling the electrostatic interactions.

Figure 5: Illustration of SDS-PAGE gels and main bands cut out for proteomic analysis to identify the bound proteins: 1 = 100 nm amine-modified, 2 = 50 nm amine-modified, 3 = 100 nm plain, 4 = 50 nm plain, 5 = 100 nm carboxyl-modified and 6 = 50 nm carboxyl-modified. The selected lanes for the different particle types were cut according to the pattern shown on the right of the gels.

Figure 6 is a schematic representation of some of the ways to change the surface chemistry on the nanoparticles which can then direct nanoparticle-protein interaction specificity with the one or more specific proteins, and for (reversible) crosslinking of the nanoparticles.

Figure 7. Fractionated lipoproteins and their binding to copolymer nanoparticles. SDS-PAGE of lipoprotein fractions and nanoparticles incubated in lipoprotein fractions. Lanes 1 to 4: Density fractions from human blood enriched in Chylomicron+VLDL, LDL, HDL and VHDL respectively. Lanes 5 to 8: Proteins adsorbed to 200 nm 50:50 NIPAM:BAM copolymer particles incubated in the density fractions loaded in lanes 1 to 4. Bound proteins were separated from unbound proteins by centrifugation and desorbed by SDS-PAGE loading buffer.

Figure 8: Illustration of the purity of apolipoprotein A1 harvested from human plasma using 50:50 NIPAM:BAM copolymer nanoparticles. Key: dark grey = commercial apolipoprotein A1; light grey = nanoparticle purified apolipoprotein A1.

PREOMICS EXHIBIT 1002
Page 378 of 1861

Figure 9: Selective protein binding to NIPAM:BAM copolymer particles of varying surface hydrophobicity and size.

Figure 10: 1D-PAGE of poly(vinyl alcohol) particles incubated with human plasma for 1 hour, and then separated from unbound proteins by centrifugation and washing 3 times. Bands were excised as shown by the arrows, and the proteins in each band identified by mass spectrometry. Several of the bands contained immunoglobulins as their principle component.

Figure 11: Schematic representation of the concept of (reversibly or irreversibly) gelling the nanoparticles into a separation matrix.

Figure 12: Schematic showing the principle of using the thermoresponsible behaviour of the NIPAM-BAM nanoparticles to control their gellation, via formation of attractive glasses as the temperature is increased above their lower critical solution temperature, at which point they become insoluble and prefer to aggregate then to remain dispersed. Lowering the temperature reverses the gellation.

Figure 13: Schematic showing the principle of depletion attraction, phase sepatation and gellation to create (reversibly) gelled structured from nanoparticles. Image on the right shows a fluorescence microscopy image of the type of structured high-surface area material that results from this process.

Figure 14: Illustration of the use of buffering conditions to initially bind selected proteins to gelled (or tethered) nanoparticles, and then changing the buffering conditions such that proteins are subsequently eluted and fractionated. Surface plasmon Resonance is used to quantify the binding and release of the proteins, and from the curves kinetic data and binding constants (affinities) can be determined.

Figure 15: Schematic of nanoparticles tethered to a surface for protein purification and separation. Protein solution is flowed over the surface, and the unbound proteins are removed by flowing buffer over the particles. Separation of the bound particles can then be achieved using additional washing steps under buffer conditions that suppress binding.

PREOMICS EXHIBIT 1002
Page 379 of 1861

Figure 16: Schematic of the concept of combined epitopes, surface curvature and surface physiochemical composition to selectively pull-down specific proteins from complex mixtures.

Figure 17: Comparison of the protein coronas around particles with different surface modifications: unmodified, PEG or RGD modification. PEG coating should reduce the protein binding, although does not seem to have this effect in any of the particle cases investigated. Modification of the surface with the tri-peptide epitope RGD results in new specific bands appearing, indicating that RDG-specific proteins bind.

Figure 18: Left: Confocal microscopy images of nanoparticles preferentially localised in a specific sub-cellular compartment (lysosomes), and Right: the use of nanoparticles to selectively pull-down (bind) proteins from various sub-cellular fractions, which are then identified either by PAGE and mass spectrometry, or on a protein array (Figure 21).

Figure 19: Illustration of the concepts of continuous (A) and pulse (B) delivery of nanoparticles to cells, and (C) pulse and chase delivery, where the first pulse is used to induce an effect, and the second pulse is used to fingerprint the cellular response by selectively binding organelle-specific biomolecules for subsequent identification and quantification.

Figure 20: Electron microscopy images of the time-resolved localisation of nanoparticles in cells following continuous exposure for up to 24 hours. During the earlier time-points, it is clear that the nanoparticles are located in several different organelles at each time-point, such that recovery of the nanoparticles would result in recovery of a mixture of proteins from each of the organelle types. However, by the longer times, all nanoparticles are located in lysosomes, which are their final destination. Thus, recovery of the particles at longer times will result in selective recovery of proteins specific to this organelle.

Figure 21: Top: Fluorescence microscopy imaging of lysosomes isolated from cells at different magnifications (A and B). Electron microscopy images of the localisation of nanoparticles in cells following exposure to a short (10 minute) pulse of nanoparticles. 50 nm green SiO_2 nanoparticles in early endosomes (C), multilamellar bodies (D) and multivescicular bodies (E), respectively. By recovering the nanoparticles at specific time-points, nanoparticles localised in each organelle fraction can be selectively recovered, and the selectively bound proteins recovered and identified.

PREOMICS EXHIBIT 1002
Page 380 of 1861

Figure 22: Illustration of the use of protein or antibody arrays for identification of proteins in the nanoparticles-protein corona. The data on the right shows the spots where nanoparticles bound to the array, indicating an interaction between the nanoparticles-bound protein and the array protein, i.e. identification of the bound protein via its interaction with a binding partner. The binding partner on the array is determined from the original spotting pattern for the array, and the fact that the spot lights up in duplicate is an internal verification that it is a genuine "hit" rather than a non-specific binding to the array.

Figure 23: Time-evolution of several of the markers of cellular apoptosis, triggered by the uptake of positively charged nanoparticles. NH_2-modified polystyrene nanoparticles induce caspase 3 and 7 activity and PARP-1 cleavage in 1321N1 cells. (A) YoPro-1/PI staining of 1321N1 incubated with nanoparticles shows an increase in the population of dead cells (early apoptotic and late apoptotic/necrotic) and a decrease in the viable cell population over a 24 hour period. (B) Analysis of the apoptosis-specific activity of caspases 3 and 7 reveals an increase over time, indicating that apoptosis is one of the mechanisms involved in cellular death induced by nanoparticles. Analysis of cellular ATP content indicates that there is also energy depletion in the cells. (C) Western-blot for the apoptosis-specific cleavage of Poly (ADP-ribose) Polymerase (PARP)-1 (116kDa) shows an increase of the cleaved protein (89kDa) over time in the presence of nanoparticles. Incubation with staurosporine (STS) for 6 hours was used as positive control for the experiment and GAPDH was used as the loading control.

Figure 24: Schematic of "domains" of interaction interspersed with domains of non-binding surface on a nanoparticle.

Figure 25: Schematic of the concept of increased nanoparticle surface area (increased area for protein binding and increased specificity) using Pickering emulsions for the synthesis of nanoparticles.

Figure 26: Schematic representation for Pickering emulsion polymerisation process

Figure 27: Schematic of globular (native) protein adsorption to nanoparticle surfaces as a method for templating protein adsorption motifs onto the surface of nanoparticles.

PREOMICS EXHIBIT 1002
Page 381 of 1861

Figure 28: Schematic representation of the interactions of proteins with surfaces (and amplified for the interactions of proteins with nanoparticles due to their significantly higher curvature).

Detailed Description of the Invention

This invention relates to the use of (polymeric) nanoparticles for the selective concentration, recovery and purification of proteins from complex mixtures, with particular emphasis on monomeric and correctly folded forms of the proteins. By tailoring the properties of the nanoparticles (surface curvature, hydrophobicity, charge, ligands etc.), the nature of the proteins selectively bound by the nanoparticles can be tailored. The key concepts are that (i) enormous amounts of surface area of specific kinds are presented by nanoscale particles, (ii) the curvature of the surface is a key separation parameter, and (iii) additional control of surface physicochemical characteristics, when combined with surface curvature are likely the most general control parameters to allow biomolecule separation in future. In addition, the particles are easily recovered and re-usable, leading to a simple one-step concentration or selection process.

Besides the intrinsic large surface area and sponge-like nature of many polymeric materials, additional properties of polymer materials such as the fact that many can be made stimuli-responsive (e.g. pH, enzyme or temperature responsive and /or crosslinkable) offer novel approaches to introduce additional levels of purification optimisation.

The fact that the protein conformation is not altered by the binding process is an incredible advantage, and results from the fact that we have done considerable work to optimise the biocompatibility of these materials. In many cases, the driving force for the binding of proteins to the surface is entropic, resulting in the release of surface-bound water from the particles, meaning that the confirmation of the proteins does not change, and that they can be recovered easily, and are intact and fully functional prior to recover. The quality of several test proteins seems to be higher than those supplied by laboratory supplies at present, and can be produced at lower cost via a simple process.

The fact that nanoparticles can reach sub-cellular locations in a highly regulated manner, with enormous specificity for specific sub-cellular organelles, utilising existing cellular trafficking pathways, opens up the potential for nanoparticles to selectively recover proteins from specific

PREOMICS EXHIBIT 1002
Page 382 of 1861

sub-cellular organelles for use in disease detection. The present invention includes the use of nanoparticles to selectively bind and recover organelle-specific proteins from inside cells, likely leading to significant earlier detection of changes in specific protein concentrations than other diagnostic approaches, and therefore to earlier disease detection.

The present invention offers many advantages over the existing approaches for protein concentration, purification and harvesting including simplicity, general applicability both *ex situ* and *in situ* in biological systems, reduced risk of protein degradation, higher and more cost effective yields, greater purity of the final protein, and general applicability to a range of biofluids, biological systems and proteins.

The present invention relates to the use of nanoparticles of controlled size (surface curvature) and, optionally, selectively modified surface (physiochemical) characteristics made from natural or synthetic polymers, copolymers, terpolymers and grafted or nanostructured surfaces for selective binding of proteins from complex biological solutions. Examples of such polymeric nanoparticles include (but is not restricted to) polystyrene, poly(lysine), chitosan, dextran, poly(acrylamide) and its derivatives (N-isopropylacrylamide, N-tert-butylacrylamide, N,N-dimethylacrylamide, etc.), Polyethylene glycol, poly(vinyl alcohol), degradable polymer particles and many others. Examples of biological fluids from which proteins can be selectively bound from include (but is not restricted to) plasma, serum, cell lysates, cytosolic fluid, gastric fluid, cerebrospinal fluid, lung lavage fluid, saliva, urine, and others. The methods of the invention cover all aspects of biomolecule (particularly protein) concentration, purification and harvesting, using nanoparticles of controlled surface curvature, physiochemical properties, free in solution, crosslinked to form gels, tethered to surfaces, and nanostructured surfaces.

Harvesting of specific selected proteins using a combination of surface curvature and changing the physiochemical properties of the nanoparticle surface – charge and chemistry.

The use of surface curvature (nanoparticle size) as a selection parameter is a fundamentally different approach and concept for biomolecule (particularly protein) separation that the traditional pull-down approach used in for example Protein A columns, as well as being significantly less expensive, more versatile and more re-usable. The "sorting" or "gating" parameter is based on a packing argument, whereby molecules are attracted to the surface in a

PREOMICS EXHIBIT 1002
Page 383 of 1861

manner that results in the most efficient packing at the surface, as illustrated schematically in Figure 2. The principle is that the surface does not want to be bare (nature abhors a vacuum, and a bare surface), and that the biomolecules (proteins) that bind will do so in a manner that allows efficient packing. In a situation where all proteins are attracted to the surface, the nanoparticle will seek out those proteins that reduce its surface exposure most efficiently, solely based on packing. Constant exchange between bound and free proteins will take place until an equilibrium situation is reached (based on packing efficiency).

Using the packing principle to perform a first selection or concentration of proteins, additional levels of protein selection can be built in by tailoring the properties of the nanoparticle surfaces to control their interactions with proteins using attractive and repulsive forces, such that only selected proteins adsorb to the nanoparticle surface. These interactions may also be tailored to ensure that the protein structure and conformation is not altered, making them an extremely useful tool for protein purification and protein therapeutics. The process of tailoring the surface properties (charge, chemistry, etc.) is shown schematically in Figure 1, where the repulsive interaction between two species of similar charge is used to prevent a positively charged particle from binding to the nanoparticle surface even through it has the correct size to optimise the packing at the surface. Such tailored interactions offer exquisite control over the nature and selectivity of the protein concentration, purification and harvesting using nanoparticles.

There are over 3700 different proteins in plasma, and typically nanoparticles bind only a handful of these in any significant concentration, offering a quick and cheap approach to concentrating proteins, as illustrated in Figure 3.

1A Using nanoparticle surface charge and electrostatic interactions to control selectivity of protein binding.

Proteins are typically quite highly charged, and much of their confirmation and binding is driven by electrostatic interactions. Nanoparticles with controlled surface charge (positive, negative, zwitterionic), and controlled surface charge density, can be used to control which proteins bind preferentially to the nanoparticle surfaces (see Figure 4).

Detailed investigations of protein binding to nanoparticles, using a series of nanoparticles of polystyrene nanoparticles of controlled surface curvature (50nm and 100nm) and controlled

PREOMICS EXHIBIT 1002
Page 384 of 1861

surface chemistry (plain, -COOH functionalised or -NH_2 functionalised), illustrated that subtle changes in surface curvature (nanoparticle size) and surface composition result in alterations in the nature of the proteins selectively bound, as shown in Figure 5 and Experimental Example 1, with a representative class of proteins (the apolipoproteins) shown in Table 1 illustrating where certain proteins are selectively bound to one size and charge of the polystyrene nanoparticles but not the others.

Lipoproteins	100nm NH_2	50nm NH_2	100nm plain	50nm plain	100nm COOH	50nm COOH
Apolipoprotein A-I	X	X	X	X	X	X
Apolipoprotein A-II		X				X
Apolipoprotein A-IV		X	X	X	X	X
Apolipoprotein B-100	X	X	X		X	
Apolipoprotein C-I	X	X			X	X
Apolipoprotein C-III		X		X		X
Apolipoprotein D		X	X	X		
Apolipoprotein E	X	X	X			X
Apolipoprotein F		X				
Apolipoprotein L1		X			X	
Beta-2-glycoprotein 1 (apolipoprotein H)	X		X	X	X	X
Clusterin (apolipoprotein J)	X	X	X	X	X	X

Table 1 Effect of nanoparticle size (surface curvature) and surface composition (charge) on the selective binding of apolipoproteins. Note that binding can be tuned by both size and surface composition, as indicated by the rows highlighted with the hatching (surface curvature effect), the horizontal lines (presence of surface charge), and in grey (surface curvature and presence of surface charge).

However, these polystyrene particles bind a very large number of proteins, making them less suitable for single protein purification, although they are capable of protein concentration. Changing the properties of the base polymer material dramatically changes the numbers of proteins bound, with, for example, NIPAM-BAM copolymers offering very high selectively for single proteins, depending on the copolymer ratio (Section 1B). Protein binding data for nanoparticles of gold, silica and many other nanoparticles is also available.

Nanoparticles can also selectively bind functional protein clusters, such as apolipoprotein complexes, and can distinguish between the different variants in this family, to selectively bind, for example, high density lipoproteins whilst not binding low density lipoproteins, as shown in Figure 6, and described in Experimental Example 4.

1B *Introduction of specific interactions to control protein-nanoparticle interactions (H bonding, thiol binding, hydrophobic binding etc.).*

PREOMICS EXHIBIT 1002
Page 385 of 1861

Introduction of surface groups with specific functionalities, such as lone pairs for Hydrogen Bonding, thiol groups for sulphur-bonding, and changing the hydrophobic/hydrophilic balance of the nanoparticle surface is used to control the specificity of the nanoparticle-protein interaction, and to harvest proteins in a controlled and specified manner (see Figure 7).

In this example the use of nanoparticles for protein purification is illustrated – copolymer particles composed of 50% N-isopropylacrylamide (NIPAM) and 50% N-tert-butylacrylamide (BAM) are mixed with human plasma, and bind Apolipoprotein A1 with a very high selectivity (Experimental Example 2). The specific method of concentration of Apolipoprotein A1 is illustrated in Table 2 below. Apolipoprotein A1 protein recovered from the nanoparticles is of very high purity, and is in fact much purer than the commercially available samples, and at a fraction of the cost of the commercial samples. This is illustrated in Figure 8, where Apolipoprotein A1 from Sigma is compared to Apolipoprotein A1 recovered from the NIPAM:BAM copolymeric nanoparticles. From the 1D PAGE gel (Figure 8A), it is clear that the commercial sample (lane 2) shows two distinct bands (corresponding to proteins of two different molecular weights) indicating the presence of protein variants in the sample. On the other hand, the sample in lane 3 which was the protein isolated via our polymeric nanoparticles, shows only one band, indicating a much higher degree of purity. The purity of the commercial and our purified samples are illustrated in Figure 8A-D, indicating that the one-step process outlined here using nanoparticles of controlled surface curvature and surface physicochemical properties already shows higher selectivity than the existing approaches currently on the market.

PREOMICS EXHIBIT 1002
Page 386 of 1861

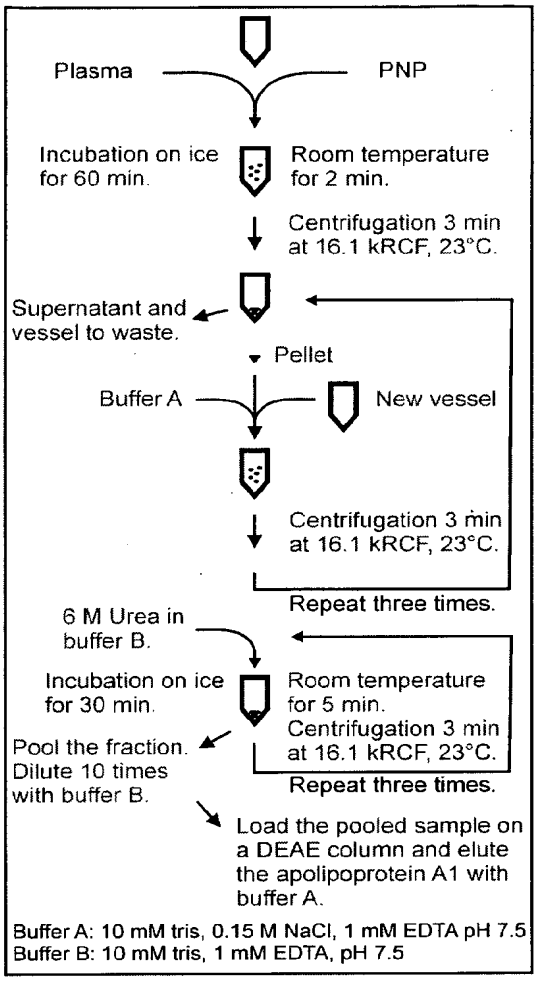


Table 2. Table showing steps involved the separation principle. The process is extremely simple, requiring only incubation and centrifugation and washing steps. Control of the protein selectivity comes from the surface characteristics of the nanoparticles used, and can be further functionalised to include a stimuli-responsive change in the interaction.

A detailed example of the use of surface curvature combined with surface hydrophobicity to control protein binding is given in Figure 9 and Experimental Example 3 below, where nanoparticles of NIPAM:BAM in different ratios (85:15, 65:35 and 50:50 NIPAM:BAM, with the hydrophobicity increasing as the amount of BAM increases) and of different sizes (70nm and 200nm) were incubated with human plasma and the bound proteins are shown using 1D SDS-PAGE gels. The plasma samples used here were pooled samples from 6 donors, and the fact that pregnancy zone protein was identified is a clear illustration of the selectivity and concentrating effect of the nanoparticles, as this protein is present only in very low concentrations in pregnant women, and would be greatly diluted in a pooled sample.

PREOMICS EXHIBIT 1002
Page 387 of 1861

Another example of selectivity binding and purification of proteins is given using Polyvinyl alcohol particles, which show a very high affinity for immunoglobulins, which have very significant use as protein therapeutics, as shown in Table 3, Figure 10 and Experimental Example 5.

Band	*UniProt ID*	*Name*
2-1	P01834	Ig κ chain C region
	P02768	Serum albumin precursor
	P99999	Cytochrome c
2-6	P01834	Ig κ chain C region
	P01857	Ig γ-1 chain C region
	P01859	Ig γ-2 chain C region
	P02768	Serum albumin precursor
	P99999	Cytochrome c
	Q9H4B7	Tubulin beta-1 chain
	P18206	Vinculin
	P02679	Fibrinogen γ chain precursor
2-9	P01871	Ig μ chain C region
	P02768	Serum albumin precursor

Table 3 Table showing the affinity of polyvinyl alcohol particles for Immunoglobulins (Ig) from the 3700 proteins present in human plasma.

Formats for separation (concentration) of specific selected proteins from complex biological solutions.

The use of nanoparticles for protein concentration / purification for therapeutic applications requires in certain cases development of new approaches to ensure that there are no residues of the nanoparticles remaining in the purified protein extracts. One option for achieving this objective is to provide a format for the nanoparticles where they are crosslinked into a gel format, such as used in column chromatography already.

2A Reversible nanoparticle gellation, controlled by, for example, buffering conditions

A very simple approach is to use the surface characteristics of the nanoparticles themselves to induce gellation under certain solution conditions (via buggering for example). For example, nanoparticles with thiol groups (-SH) at the surface can be gelled via –S-S- bonding under certain pH conditions. The use of this concept is illustrated for NIPAM:BAM copolymer particles with controlled numbers of –SH modifications at the nanoparticle surface. Selective binding of proteins via this format results in identical proteins being identified as with

PREOMICS EXHIBIT 1002
Page 388 of 1861

identical nanoparticles in the free format (where the separation from the unbound proteins is via controlled centrifugation and washing steps). In this case, unbound proteins are removed via a buffer washing step while the particles are still in the gelled state.

Recovery of the bound proteins is then possible by reversing the gellation and the usual centrifugation and recovery steps, as for the non-gelled nanoparticles. A detailed example of this is the modification of the NIPAM:BAM copolymer particles with –SH groups with very small (typically 2 or 3) –SH groups at the surface (determined by calculating the approximate number of monomers at the surface of a particle of a certain size, and adding enough of the –SH functionalisable monomer to the starting mixture to ensure that (averaged over the entire number of particles) 2-3 of the surface residues are –SH. By varying the concentration of the particles following cleaning, the onset of gellation, and the return to the ungelled state is controlled by the concentration of a redox activator, such as DTT. This concept is shown schematically in Figure 11.

Specific examples of reversible gellation of the nanoparticles for use as solid or gelled separation formats include exploiting the thermoreversible properties of NIPA:BAM copolymer particles, which can be caused to form a gel by increasing the temperature, whereby the particle-water interactions become unfavourable and the particles aggregate and can be cast to form a gelled solid phase which acts as the separation format. This is shown in Figure 12 and Experimental Example 6. The gellation is reversed by lowering the temperature again in the presence of an aqueous solution. The protein recovery and identification is then as before.

Another approach is reversible gellation via the use of a small molecule additive to induce a depletion attraction and cause the nanoparticles to aggregate into particle-rich and particle-poor pahses, which can then be cast onto a surface to create porous high surface area solid-like materials for separation, as shown in Figure 13 and Experimental Example 7. The gellation can be reversed by removing the small depleting molecule by changing the buffering conditions. The protein recovery and identification is then as before.

Another approach here is the use of magnetic cored nanoparticles, which can be induced to form a gelled structure by application of a magnet, and can be re-dispersed by removal of the magnetic field. The protein recovery and identification is then as before.

PREOMICS EXHIBIT 1002
Page 389 of 1861

2B Irreversible nanoparticle gellation – standard chromatography-type approaches

For applications in which the protein-nanoparticle interactions are easily modified, it is possible to have a single-step approach, where the binding and recovery steps are all performed in the gelled nanoparticle phase, such as occurs with classical column separation formats.

Here buffering conditions alone are used to recover the proteins bound to the nanoparticles. An additional advantage of this approach (and that of Section 2A) is that nanoparticles of different composition are used in the column, and the covalent interactions designed into the nanoparticle design. Suitable crosslinkers for the nanoparticles include (but are not limited to) bi-functionalised PEG spacers and other oligomeric species. A specific example of this is given by the on and off kinetics of proteins to nanoparticles shown in Figure 14 where the nanoparticles were tethered in a representative gelled format via thiol binding of the nanoparticles. Further details are given in 2C below, but the principles are the same here.

2C Surface tethering of the nanoparticles and use as a separation matrix

Another approach to ensuring that nanoparticle contamination of the purified protein is not an issue for regulatory approval of therapeutic uses of the recovered proteins is to tether the nanoparticles to a surface and to flow the biological fluid over the tethered nanoparticles – specific binding to the nanoparticles will occur as usual, and the unbound proteins can be removed by simple washing steps. Recovery of the selectively bound proteins can then be enacted using specific buffering conditions to desorb the proteins.

This approach has already been shown to work for both binding and release of selected proteins, as well as using full plasma, using the series of copolymer nanoparticles of controlled hydrophobicity, modified with exactly one –SH group which was then used to bind the copolymer nanoparticles to the surface of a gold sensor-chip for BIACore Surface Plasmon Resonance (SPR) determination of protein binding kinetics. The –SH modified particles were first allowed to flow over the sensorchip to allow binding of layer of the particles to the surface. Unbound particles were removed by flowing additional buffer over the sensorchips. These sensorchips were then exposed to complex biological fluids such as plasma until saturation binding was reached. Then the buffer solution was changed and the proteins were dissociated from the surface and recovered. Tailoring the particle size and surface functionality resulted in different on and off rates for the selectively bound proteins, as shown in Figure 15 and Experimental Example 8.

PREOMICS EXHIBIT 1002
Page 390 of 1861

Specific modifications of nanoparticle surfaces, including the use of epitopes, oligopeptides, sugars etc. to selectively target proteins

Many proteins have specific ligand-binding sites, which are spatially organised chemical units to which the protein binds via pattern recognition. This is also the mechanism by which cells control uptake and trafficking of proteins, and can be utilised to control uptake and sub-cellular localisation of nanoparticles. Development of nanoparticles with such motifs on the surface will be another very useful approach to direct protein binding to the surface, and can even be used to direct proteins that would normally not bind to polymer nanoparticle surfaces, and via this to direct protein-protein binding (Figure 16). A key advantage of the combined nanoparticle-based approach is that significantly less of the epitope modification is required to achieve a similar protein concentrating effect, meaning much lower costs.

3A *Use of simple conformational epitopes (peptides) to direct protein binding*

Preliminary data exists to show that the pattern of protein bind to nanoparticle surfaces can be controlled using even very simple tri-peptide epitopes, such as RGD, as shown in Figure 17 and Experimental Example 8. Here the polyvinyl alcohol particles described in Figure 10 have been surface-modified to contain a small amount of RGD peptides. The RGD tripeptide (Arg-Gly-Asp) is recognized as the active sequence of adhesive proteins (collagen, fibronectin, vitronectin and laminin) of the extracellular matrix which bind to the integrin receptors, a large family of transmembrane protein implicated in all cellular adhesion phenomena. The $\alpha v\beta 3$ and $\alpha v\beta 5$ integrins have an interesting expression pattern on endothelial cells during angiogenesis, where they are significantly over-expresssed on angiogenic endothelial cells within tumours. It is clear that a new band appears, indicating that an additional protein is bound with a high affinity as a consequence of the presence of the targeting ligand. Combining this with size and surface control, and optionally with PEGylation to reduce the overall binding, functionalisation with epitope / peptides / ligands etc. offers another route to increased specificity.

3B *As above but with spacer groups to control the distance between the nanoparticle surface and the oligopeptide*

Due to the fact that the surface itself will attract proteins, many of which are quite large, it may be desired to have some specific tags that protrude out from the surface of the nanoparticle, which can then bind specifically proteins via the pattern recognition approach

PREOMICS EXHIBIT 1002
Page 391 of 1861

described above. These proteins are then available for interaction with cell-surface receptors for controlled uptake by specific cells.

Using the uptake and sub-cellular localisation of nanoparticles to selectively recover specific organelles or organelle-specific proteins

The small size of nanoparticles means that they have the capability to reach all sub-cellular locations, including the nucleus, and to pass through all biological barriers, including the Blood-Brain Barrier (BBB), meaning that they have unparalleled access to sub-cellular locations for delivery of therapies and for detection of small changes in protein expression that may result from specific diseases. Data exists to show that nanoparticles can be directed in very controlled manners to specific organelles, and that these organelles can be separated from the whole cell, and their proteins identified (Figure 19, and Experimental Example 11). Significant differences in terms of the protein binding pattern have been observed depending on the cellular sub-fraction from which the particles were recovered (the example shown here is for SiO_2 nanoparticles incubated with whole cell lysate or with cytosolic fluid, both of which contain large numbers of proteins, a sub-set of which have selectively bound to the nanoparticles in each case), as shown in Figure 22. Recovery of the nanoparticles and subsequent identification of the bound proteins by any of the methods described above shows selective binding of a small sub-set of the organelle-specific proteins. Quantitative proteomics approaches, such as iTRAQ, can then be utilised to quantify the bound proteins, and to monitor changes in protein binding with time, or as a function of disease expression. Screening of the nanoparticles recovered from cells to identify the proteins bound can be achieved using proteins arrays (Figure 22).

4A *Utilising the fact that nanoparticles selectively localise to specific sub-cellular organelles to preferentially bind the proteins from that sub-cellular compartment selectively*

Nanoparticles have been shown to be taken up by cells and to have a final localisation in specific sub-cellular locations, thereby enabling pull-down of proteins specific to that sub-cellular location (Figure 18, Experimental Example 11). Data from time-resolved studies shows that the uptake process for nanoparticles (using 50nm SiO_2 nanoparticles as an example) is active, energy dependant, and leads to an intracellular load growing linearly in time. At shorter times, endosomal structures are occupied, and later lysosomes are populated (Figure 20, and Experimental Example 12) enabling nanoparticles to be recovered from specific sub-cellular fractions, and the associated low-abundance proteins to be recovered. Use

PREOMICS EXHIBIT 1002
Page 392 of 1861

of magnetic cores will facilitate magnetic recovery of these particles, and their associated / bound proteins, for subsequent identification of the proteins, as described above or using arrays of proteins (Figure 21).

4B *Utilising pulses of nanoparticles to target organelles in a time-resolved manner*

During continuous exposure to nanoparticles, nanoparticles are being taken up continuously, and thus at any given time, the nanoparticles can be found to simultaneously occupy all of the organelles involved in the uptake pathway, such as the early, sorting and late endosomes and the lysosomes, as shown in Figure 18. However, if the cells is only exposure to nanoparticles for a short time (i.e. a pulse of nanoparticles is applied, as shown schematically in Figure 19), all of the nanoparticles are at the same stage of uptake at each time-point, and as such the nanoparticles occupy only one organelle per time-point, as shown in Figure 21 and Experimental Example 12, enabling selective recovery of that specific organelle via application of a magnetic field to previous examples.

4C *Utilising pulses of nanoparticles to firstly perturb cells in a specific manner, and subsequently to selectively bind and recover biomolecules associated with the cellular response to the perturbant (biomarkers)*

A further expansion on the pulse concept, is to use a short pulse of a known antagonist (e.g.. drug or nanoparticles) to induce a response in cells, such as to trigger apoptosis for example, followed by application of a second pulse of nanoparticles to enter cells and selectively bind (pull-down) the functional protein clusters induced by the presence of the perturbant, as shown schematically in Figure 19 The time evolution of the cellular functional responses, such as apoptosis, can be determined as described in Experimental Example 13 and shown in Figure 23, and the "chase" pulse can be applied at the appropriate time to target functional protein clusters associated with the specific stage of the response. Again, the application of a magnetic field will be used to recover the magnetic-cored nanoparticles, and their bound functional protein clusters.

Engineered nanoparticle interfaces – radical new directions for selectively controlling protein binding to nanoparticles utilising surface curvature.

New applications of using nanoparticle surfaces as templates for protein binding are described. These applications result from the combined expertise in nanoparticle synthesis, nanoparticle-protein interactions and physical and biophysical chemistry of the inventors.

PREOMICS EXHIBIT 1002
Page 393 of 1861

5A *Nanoparticles with surface domains (patches) of differing interaction capacity*

A strategy in controlling (for example) cellular adhesion (which is actually mediated by protein binding) is to "pattern" the surface with areas that are conducive to protein binding, and areas that are non-protein binding, which is then translated into areas where cells adhere and areas where the cells do not adhere (see Figure 24 and below). A similar approach has been developed for nanoparticle surfaces with domains that are protein biding and domains that are non-binding, or domains that are selective for different proteins. This is achieved for example, using block-polymerisation, where the blocks have different chemistries, or via phase separation of mixtures of different polymers and cross-linking following the phase separation.

5B *Nanoparticles with additional surface area generated using novel synthetic routes, such as Pickering emulsions – engineered surface curvature on multiple length-scales*

Pickering emulsions are oil-in-water or water-in-oil emulsions stabilized using finely divided particles. Typical examples for the kinds of particles that have been used to stabilize such emulsions include are silicates, aluminates, titanates, metal oxides such as aluminum, silicon, titanium, nickel, cobalt, iron, manganese, chromium, or vanadium oxides, carbon blacks, or nitrides or carbides, such as boron nitride, boron carbide, silicon nitride, or silicon carbide. The particles have to meet several requirements in order to effectively stabilize emulsions. Most importantly, the particles should not be completely wetted by water for stabilization of o/w emulsions and not be completely wetted by oil when used for the stabilization of w/o emulsions.

Clearly, this approach for the modification of nanoparticles results in nanoparticles with dramatically increased surface area for protein purification and concentration, and surface curvature on multiple length-scales (see Figures 25 and 26 and Experimental Example 9). This approach offers new dimensions for gating of selected protein fractions via sophisticated packing parameters. The key step is the replacement of the inner solution by the polymerisable liquid, which is then caused to polymerise by application of the appropriate reaction conditions (temperature, initiator, etc.). These steps would be familiar to anyone trained in polymerization techniques. The subsequent use of these particles for separation and purification of proteins is as described for the previous examples.

5C *Templating increased surface area using intact proteins at the surface – generation*

PREOMICS EXHIBIT 1002
Page 394 of 1861

of nanoparticles with specific protein templates

Using single proteins, is be possible to develop strategies to bind the proteins to the particle surfaces with different conformations and different surface exposed amino acids, via combinations of electrostatic, H-bonding and other interactions. These are then templated by polymerising and gelling around them, and then removing the proteins via a denaturation step and washing step, such as in SDS urea. This will then leave particles with templates of the different conformationally bound proteins, which will likely preferentially bind that protein again in the same conformation (tertiary structure) even from complex mixtures of proteins where competitive binding might otherwise prevent that protein from binding due to a lower affinity for the surface compared to other proteins in the mixture. This approach is also used to promote specific protein-protein interactions to occur between the templated protein and its protein binding partners, which may have low affinity for the particle directly. In the best-case scenario, some of the secondary structure of the bound proteins may be templated into the nanoparticle surface as another example of surface curvature at multiple scales. See Figure 27.

5D *Templating using proteins in various states of degradation (for use in generation of antibodies against the different binding states)*

Proteins in solution are easily degraded by a host of well known routes, such as addition of increasing amounts of urea. The binding of the proteins in the various states of degradation is likely to be significantly different to that of the native (energetically stable state), as a result of the hydrophobic domains which are typically hidden in the core of the folded protein becoming exposed. Binding of such denatured and partially denatured proteins to the nanoparticle surface could have dramatic consequences for protein stability and fibrillation, and also for the immunogenicity of the proteins, as a result of exposure of new or "cryptic" epitopes (sequences of amino acids). This is shown schematically in Figure 28, where the different protein-binding scenarios are represented, ranging from binding with no conformational change, through to binding with complete denaturation of the protein.

Binding of selected proteins with varying degrees of denaturation to selected nanoparticles will be investigated. Use of such protein-bound nanoparticles on protein aggregation will be investigated as the potential to generate antibodies against the denatured proteins will also be investigated. Such denatured protein states could potentially be biomarkers for protein interaction with nanoparticles involving a conformational change, indicative of a change in protein function and/or protein-protein interactions.

PREOMICS EXHIBIT 1002
Page 395 of 1861

Experimental

Example 1

Polystyrene nanoparticles

Polystyrene latex beads were purchased from Sigma (amine modified 50nm and 100nm labeled with blue and orange fluorophores respectively) and from Polysciences (both unmodified (plain) and carboxyl-modified 50nm and 100nm, labeled with yellow-green fluorophore). All nanoparticles were used as received.

Human plasma

Blood was taken from 10 different seemingly healthy donors. Each donor donated blood for 10 x 3 ml tubes containing EDTA to prevent blood clotting. The blood donation was arranged such that the blood samples were labeled anonymously. They could not be traced back to a specific donor, however, it was possible to use plasma from just one of the donors for a specific experiment. The tubes were centrifuged, for 5 min at 800 RCF to pellet the red and white blood cells. The supernatant (the plasma) was transferred to labeled tubes and stored at -80°C until used. Upon thawing the plasma was centrifuged again for 2 min at 16.1 kRCF to further reduce the presence of red and white blood cells.

Polystyrene nanoparticle incubation with plasma

All experiments were conducted at least twice to ensure reproducibility of the particle-protein complex pellet sizes, general pattern and band intensities on the 1D gels. Particle suspensions of 1.66 mg/ml were incubated with different concentrations of human blood plasma in 10mM Phosphate, 0.15M NaCl, 1mM EDTA, pH7.5 for 1h (total volume 750µl). The ratio of total particle surface area to plasma concentration was kept the same for the two different particle sizes to ensure comparability between the results. The samples were centrifuged to pellet the particle-protein complexes. The pellet was resuspended in PBS, transferred to a new vial and centrifuged again to pellet the particle-protein complexes, this procedure was repeated three times. After the third washing step the supernatant did not contain any detectable amount of proteins. The proteins were eluted from the particles by adding SDS-sample buffer to the pellet and boiling the solution. The proteins were separated by 12% SDS-PAGE.

Determination of pellet weight

PREOMICS EXHIBIT 1002
Page 396 of 1861

Polystyrene nanoparticles were mixed with different amounts of human plasma as previously described. The samples were centrifuged to pellet the particle – protein complexes and the supernatants were discarded. The pellets were dried over night at 60 °C and then weighed with a Sartorius SE2 microbalance (Sartorius AG, Germany). The results represent the average over two individual series and the error bars represent the standard deviation.

Protein identification by Mass Spectrometry

Bands of interest from SDS-PAGE gels (12%) were excised and digested in-gel with trypsin according to the method of Shevchenko et al(19). The resulting peptide mixtures were resuspended in 0.1% formic acid and analyzed by electrospray liquid chromatography mass spectrometry (LC MS/MS). An HPLC (Surveyor, ThermoFinnigan, CA) was interfaced with an LTQ ion trap mass spectrometer (ThermoFinnigan, CA). Chromatography buffer solutions (Buffer A, 0.1% formic acid; Buffer B, 100% acetonitrile and 0.1% formic acid) were used to deliver a 72 min gradient (5 min sample loading, 32 min to 40% Buffer B, 2 min to 80%, hold 11 min, 1min to 0%, hold for 20 min, 1 min flow adjusting). A flow rate of 150 µl/min was used at the electrospray source. Spectra were searched using the SEQUEST algorithm(20) against the Indexed uniprot/swiss prot database (http://www.expasy.org; release 3 July 2007). The probability-based evaluation program, Bioworks Browser was used for filtering identifications; proteins with Xcorr (1,2,3) = (1.90, 2.00, 2.50) and a peptide probability of $1e^{-5}$ or better were accepted.

Example 2

70 nm NIPAM:BAM 50:50 polymer particles

N-isopropylacrylamide-co-N-*tert*-butylacrylamide (NIPAM:BAM) copolymer particles of 50 nm diameter with 50:50 ratio of the co-polymers were synthesized in SDS micelles by free radical polymerization. The procedure for the synthesis was as follows: 2.8 g monomers (in the appropriate wt/wt ratio), and 0.28 g crosslinker (N,N-methylenebisacrylamide) was dissolved in 190 mL MilliQ water with 0.8 g SDS and degassed by bubbling with N_2 for 30 min. Polymerisation was induced by adding 0.095 g ammonium persulfate initiator in 10 mL MilliQ water and heating at 70°C for 4 hours[2]. Particles were extensively dialysed against MilliQ water for several weeks, changing the water daily. Particles were lyophilized and stored in the fridge until used.

Plasma

PREOMICS EXHIBIT 1002
Page 397 of 1861

Human blood was withdrawn from seemingly healthy humans into vessels pre-treated with EDTA-solution. The blood vessels where centrifuged for 5 min at 800 RCF. The supernatants (the plasma) were transferred to new vessels and stored in -80°C freezer until time of use. Before use the plasma vessel were thawed and centrifuge for 3 min at 16.1 kRCF and the supernatant where transfer to a new vessel, or in the case where more than one plasma vessel where needed the supernatants were pooled.

Particle and plasma mixtures

Stock solution of particles was made by dissolving lyophilized co-polymer particles, to a concentration of 10 mg/mL, in 10 mM trizma base/HCl, pH 7.5 with 0.15 mM NaCl and 1 mM EDTA. The mixture was kept on ice until all the particles had dissolved (1-3 hours) before mixing with plasma. After mixing the particles and plasma the sample was kept on ice except during the centrifugation steps which were performed in room temperature.

In one experiment lyophilized co-polymer particles were dissolved directly in plasma solution by adding plasma to the particles and keeping the mixtures on ice for 1 hour.

Ion-exchange chromatography

Urea fractions (eluates form the nanoparticles) were pooled and passed through a 0.45 µm syringe filter. After the filtration the urea and salt concentrations were lowered by either dilutions (tenfold with 10 mM trizma base/HCl buffer pH 7.5 containing 1 mM EDTA) or by dialysis against 10 mM trizma base/HCl buffer pH 7.5, 1 mM EDTA. The sample was thereafter loaded on a HiTrap DEAE FF 1 ml column from GE Healthcare. The apoA1 were eluted with a stepwise elution profile. First step was elution with 10 mM trizma base/HCl buffer pH 7.5 containing 1 mM EDTA and 50 mM NaCl. The second step (the elution of apoA1) is performed with 10 mM trizma base/HCl buffer pH 7.5 containing 1 mM EDTA and 0.15 M NaCl. The ion-exchange chromatography was conducted at room temperature.

SDS-PAGE

In this study two different SDS-PAGE gels have been used. PIERCE, Precise™ Protein Gels 4-20% pre-made gel, commercial available. Also 12% gels were used which were cast at the occasion for use.

Commercial Apolipoprotein A-1

Apolipoprotein A-1 (A-0722) was obtained from Sigma-Aldrich.

PREOMICS EXHIBIT 1002
Page 398 of 1861

Biophysical analysis by UV-, CD- and fluorescence spectrometry

The commercial apolipoprotein A-1 (Sigma-Aldrich, A-0722) was diluted 10 times with 10 mM phosphate, 0.15 M NaCl, 1 mM EDTA, pH 7.5, to approximately 0.15 mg/ml which corresponds approximately to the concentration of the purified apoA1 as estimated from the absorbance at 280 nm.

The CD signals between 200 and 260 nm were recorded on a Jasco J-720 spectropolarimeter at 25 °C using a 1 mm quartz cuvette. The fluorescence emission spectrums, after excitation at 295 nm, was recorded between 310 and 400 nm on a Perkin-Elmer luminescence spectrometer LS 50 B in a 1 cm quartz cuvette and the UV absorbance spectrum was recorded between 250 and 300 in a 1 cm cuvette.

Trypsin digestion and identification with mass spectrometry

After the separation of proteins by SDS/PAGE (12%), bands were excised from the gel, reduced, alkylated, and digested with trypsin, and the resulting peptide mixtures were separated and analyzed by nanoscale liquid chromatography quadrupole time-of-flight MS/MS^3. Spectra were analyzed by MASCOT software to identify tryptic peptide sequences matched to the international protein index (IPI) database (www.ebi.ac.uk/IPI/IPIhelp.html).

Example 3

Nanoparticles

N-*iso*propylacrylamide-co-N-*tert*-butylacrylamide (NIPAM:BAM) copolymer particles of 70 and 200 diameter and with three different ratios of the co-monomers (85:15, 65:35 and 50:50 NIPAM:BAM) were synthesized in SDS micelles. The procedure for the synthesis was as follows: 2.8g monomers (in the appropriate wt/wt ratio), and 0.28g crosslinker (N,N-methylenebisacrylamide) was dissolved in 190 mL MilliQ water with either 0.8 g SDS (for the 70 nm particles) or 0.32 g SDS (for the 200 nm particles) and degassed by bubbling with N_2 for 30 minutes. Polymerisation was induced by adding 0.095 g ammonium persulfate initiator in 10 mL MilliQ water and heating at 70 °C for 4 hours (2). Particles were extensively dialysed against MilliQ water for several weeks, changing the water daily, until no traces of monomers, crosslinker, initiator or SDS could be detected by proton NMR (spectra were acquired in D_2O using a 500 MHz Varian Inova spectrometer). Particles were freeze-dried and stored in the fridge until used.

Particle handling

PREOMICS EXHIBIT 1002
Page 399 of 1861

Due to the inverse solubility of polyNIPAM, particle solutions were prepared by dissolving the particles on ice to ensure good solubility of the particles (i.e. to ensure that the solutions are below the lower critical solution temperature of the particles).

Gel filtration of copolymer nanoparticles and plasma proteins

Human plasma was drawn from healthy individuals into collection tubes with EDTA and stored in aliquots at -80 °C. Nanoparticles, 1mg, were mixed with 20 µl plasma in a final volume of 220 µl in10 mM Tris-HCl, pH 7.5, 0.15 M NaCl, 1 mM EDTA, and incubated on ice for 1 h. The mixture was loaded onto a 1.5 x 100 cm column packed with sephacryl S1000 SF, and eluted at 0.66 ml /min, with the same buffer. Fractions of 3.66 ml were collected and analysed by absorbance at 280 nm. Proteins in the fractions were precipitated with trichloroacetic acid (TCA), *i.e.* 0.9 ml of the fractions was mixed with 100 µl TCA and incubated on ice for at least 3 h, centrifuged, the supernatant removed and the pellet dried for 30 min at 60°C. The pellets were dissolved in SDS-PAGE loading buffer, and the pH adjusted with 1 µl 1 M Tris. The proteins were separated by 10 % SDS-PAGE. For the single protein experiments, 1mg copolymer particles were mixed with 100 µg HSA or fibrinogen in a final volume of 110 µl 10 mM Tris-HCl, pH 7.5, 0.15 M NaCl, 1 mM EDTA, and incubated on ice for 2 h. The mixtures were loaded onto a sephacryl S-1000 SF column, 90 x 1.5 cm, and eluted with the same buffer at 0.667 ml / min. Fractions were collected every 4 min (2.6 ml) and analysed by UV absorbance spectroscopy at 280 nm. Fractions were precipitated as described above and analysed by 10 % SDS-PAGE. For the experiments with isolated proteins on the particles 5 mg particles in 1 ml 10 mM Tris-HCl, pH 7.5, 0.15 M NaCl, and 1 mM EDTA were incubated with 400 µl plasma on ice. After 1 h the mixture was heated to 23°C to promote particles aggregation. The particles were pelleted by centrifugation, the supernatant discarded and the pellet washed three times with 0.5 ml buffer. The vials were changed after each washing step to minimize contamination of plasma proteins bound to the vials. The washed pellet was dissolved in 0.5 ml buffer A by incubation on ice. The sample was loaded onto a 95 cm S-1000 sephacryl SF column and eluted at 5°C with the same buffer. The fractions were incubated at 37°C to promote particle aggregation and the particle elution profile determined by measuring the absorbance at 280 nm. To visualize the proteins, each fraction was precipitated with TCA and separated on a 15% SDS-PAGE.

Protein identification by mass spectrometry

After the separation of proteins by SDS/PAGE (12%), bands were excised from the gel and identified as previously described (Berggard et al,. 2006). Briefly, the gel-bands were reduced

PREOMICS EXHIBIT 1002
Page 400 of 1861

and alkylated, digested with trypsin and the resulting peptide mixtures were separated and analyzed by nanoscale liquid chromatography quadrupole time-of-flight MS/MS. Spectra were analyzed by MASCOT software to identify tryptic peptide sequences matched to the International Protein Index (IPI) database (http://www.ebi.ac.uk/IPI/IPIhelp.html

Example 4

Particle synthesis

As in example 3.

Purification of lipoprotein particle fractions from plasma

Lipemic citrate plasma was ultra centrifuged repeatedly at 40 000 rpm in a Beckman centrifuge Optimal L-70K with rotor Ti 701, for 25h at 12°C. Before each centrifugation the density was adjusted with 5 M NaCl and saturated NaBr, both containing 0.04% EDTA. After each centrifugation step a lipoprotein fraction was collected from the top of the centrifuge tubes. The corresponding densities from which the fractions were collected are: 1.0068, 1.068, 1.21 and 1.25 g/ml for Chylomicron+VLDL, LDL, HDL and VHDL respectively. All fractions were dialysed against PBS-EDTA.

Bound protein and lipids from purified lipoprotein particle fractions

The proteins in the final four fractions separated as described (Chylomicron+VLDL, LDL, HDL and VHDL) were visualized by SDS-PAGE, Figure 4 lanes 1 to 4, and the proteins bound to copolymer nanoparticles from each lipoprotein particle fraction, lanes 5 to 8. The main proteins in the chylomicrons and VLDL fraction, lane 1, are from the top B-100 and/or B-48, (B-100 is not separated from its truncated variant B-48 in this system), HSA, apolipoprotein E and apolipoprotein A-1. The same proteins are present on the nanoparticles from the same fraction, lane 5, but the relative amounts of apolipoprotein E and A-I compared to B-100 are much greater on the nanoparticles. In the LDL fraction, lane 2, the main protein is as expected B-100, but visible amounts of albumin and apolipoproteins E and A-I are present. The relative amount of B-100 is much less on the copolymer nanoparticles incubated in the LDL-fraction, lane 6, indicating that lipoprotein particles with apolipoprotein E or A-I preferentially bind to the copolymer nanoparticles. In the HDL and VHDL fractions, lane 3 and 4, the major proteins are apolipoprotein A-I and HSA..On the copolymer nanoparticles incubated in the HDL and VHDL fractions, lane 7 and 8, apolipoprotein A-I dominates.

PREOMICS EXHIBIT 1002
Page 401 of 1861

Example 5

Human Blood Plasma

Blood samples were taken from various donors into sample tubes containing EDTA to prevent blood clotting. The samples were centrifuged at 800g for 5 min. The supernatant was recovered, aliquoted into 1ml portions, frozen and stored at −80 °C. Before each experiment, the plasma was thawed and centrifuged again at 16.1×10^3 g for 3 min. The supernatant was used for experiments.

Buffers

10mM Phosphate, 0.15M NaCl, 1mM EDTA (PBS buffer) was used as the buffer solutions for all protein studies, as it closely resembles physiological conditions (pH, ionic strength, etc.).

Polyvinyl alcohol particles

Polyvinyl alcohol particles are produced by rapid stirring of a telechelic Poly(vinyl alcohol) PVA solution under a gas atmosphere. Floating particles are separated and dialyzed against water. By varying the parameters (pH and temperature) during polymerization, the particle size and the thickness of the polymer shell can be controlled.

Tissue Culture Medium

MEM tissue culture medium (Invitrogen Corp.) was supplemented with 10% fetal bovine serum (FBS, Invitrogen Corp.), 1% penicillin/streptomycin (Invitrogen Corp.), 1% L-glutamine (Invitrogen Corp.), and 1% non-essential aminoacid (Hycrone), and stored at 37 °C.

1D Polyacrylamide Gel Electrophoesis (PAGE)

Polyacrylamide gels (12%) were used to separate the bound proteins.

Mass spectrometry

Bands of interest from SDS-PAGE gels (12%) were excised and digested in-gel with trypsin according to the method of Shevchenko et al[14]. The resulting peptide mixtures were re-suspended in 0.1% formic acid and analyzed by electrospray liquid chromatography mass spectrometry (LC MS/MS). An HPLC (Surveyor, ThermoFinnigan, CA) was interfaced with an LTQ ion trap mass spectrometer (ThermoFinnigan, CA). Chromatography buffer solutions (Buffer A, 0.1% formic acid; Buffer B, 100% acetonitrile and 0.1% formic acid) were used to

PREOMICS EXHIBIT 1002
Page 402 of 1861

deliver a 72 min gradient (5 min sample loading, 32 min to 40% Buffer B, 2 min to 80%, hold 11 min, 1min to 0%, hold for 20 min, 1 min flow adjusting). A flow rate of 150 µl/min was used at the electrospray source. Spectra were searched using the SEQUEST algorithm[15] against the Indexed uniprot/swiss prot database (http://www.expasy.org; release 3 July 2007). The probability-based evaluation program, Bioworks Browser was used for filtering identifications; proteins with Xcorr (1,2,3) = (1.90, 2.00, 2.50) and a peptide probability of $1e^{-5}$ or better were accepted.

Example 6

Particle synthesis

As in example 3.

Thermoreversible gellation

Tuning the interactions of the nanoparticles is possible simply by adjusting the temperature beyond the lower critical solution temperature of the particles, above which the polymer-polymer interactions become more favourable than the polymer-water contacts, and as such the particles self-associate to limit their contact with the surrounding water. As the particles are now "sticky" the normal packing parameters don't apply, and glasses or gels form at much lower packing fractions (so-called attractive glasses). For the NIPAM:BAM copolymers described here the transition temperatures vary from 34 °C for the 100% NIPAM particles, down to 10 °C for the 50:50 BAM:NIPAM particles. The gellation is reversed simply by lowering the temperature back below the LCST. Such inverse solubility makes these particles highly versatile.

Example 7

Particle synthesis

As in example 3.

Polymer Synthesis.

PREOMICS EXHIBIT 1002
Page 403 of 1861

Linear copolymers of similar composition, 50:50 poly(NtBAM-co-NiPAM), were prepared by free radical polymerization in benzene. A 10% w/v solution was used, resulting in a polymer of MW 1.5 _ 104, as determined by combined gel permeation chromatography (GPC) and light scattering.

Film Preparation from Particle Suspensions.

Dispersions of 4 wt % of the microgel particles in ethanol either alone or with various concentrations (between 0.4 and 0.8 wt %) of linear poly(NtBAM-co-NiPAM) 50:50 were spread over glass slides (LabTec II Chamber slides) and allowed to dry overnight (however, the evaporation of ethanol occurred rapidly in the first few minutes), resulting in "bumpy" surfaces. The amount of solution needed to ensure coverage of the slides was determined by varying the microgel particle concentration between 1 and 9 wt % with 0.4 wt % free polymer added.

Fluorescence Staining of the particle layers

After the surfaces had been cast, and left overnight to dry, they were stained with Fitc. A drop of cold fluorescein 5-isothiocyante isomer 1 (Fit-C, Molecular Probes) made up in DMSO (10% w/v) was added to the particle films at 4 °C, and they were left in the cold room for an hour or so. The plates were then warmed to 37 °C prior to washing thoroughly with warm water. The surfaces could now be visualized using fluorescence microscopy.

Fluorescence Microscopy.

The distribution of the microgel particles on the surface and the surface coverage by the particles was determined by fluorescence microscopy using a Zeiss Axioplan Imaging microscope with a _10 objective, and the images were obtained using a Zeiss Axiocam camera. Through the use of the _10 objective, it is possible to see the bulk structure with a view of 200 _ 200 ím, as well as the variations in the overall surface topography induced by adding free polymer and/or changing the microgel particle size.

Example 8

Thiol-linked Nanopartices

PREOMICS EXHIBIT 1002
Page 404 of 1861

NIPAM:BAM:acrylic acid copolymer nanoparticles were synthesized as above, with the addition of appropriate amounts of acrylic acid to obtain particles with on average less than one carboxyl group on the particle surface. Acrylic acid was distilled under reduced pressure before use to remove stabilizers. A stock solution of 1 mg/ml acrylic acid was prepared, and 10 µL (70 nm particles) or 1.4 µL (200 nm particles) of this solution was added to the monomer solution. Reaction proceeded at 70 °C for 4 hours followed by dialysis against MilliQ water for a couple of weeks. The covalent attachment of homocysteine to the acrylic acid groups involves the formation of amide bonds between the primary amino group of the amino acid and carboxylic acid (3). Briefly, 50 mL of the particle solution (after dialysis) was adjusted to pH 5 by small amounts of 5 M NaOH. 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide hydrochloride was added to a final concentration of 150 mM to activate the carboxylic acid moieties. After 1 hour of incubation with stirring at 4 °C, 0.4 g homocysteine was added and the pH was readjusted to 5. The reaction mixture was incubated for 5 h at room temperature under stirring, dialysed extensively against MilliQ water to ensure that no residual chemicals remained, and freeze-dried.

Conjugation of nanoparticles to gold surfaces for SPR studies

The SIA Au kit (BIAcore AB, Uppsala) was used for sensor chip preparation. Thiol-linked nanoparticles were dissolved at 0.2 mg/ml in 20 mM sodium phosphate buffer, 100 mM NaCl, pH 7.5) on ice and 120 µl was applied to a 10 x 10 mm gold surface for four hours or over night, before the surface was rinsed with H_2O, dried and assembled in a sensorchip cassette. The change in response units after coupling of the nanoparticles to gold reveals the amount of immobilized nanoparticles. Densely packed layers of 70 and 200 nm particles yields 35 kRU and 100 kRU, respectively, and the increase in response obtained in separate coupling trials ranged from 20 to 50 % of these numbers, indicating efficient coupling of the particles.

Surface Plasmon Resonance (SPR) experiments

SPR studies of protein associating to and dissociating from nanoparticles were performed using a BIAcore 3000 instrument (BIAcore AB, Uppsala). The flow buffer contained 10 mM Tris/HCl pH 7.4 with 3 mM EDTA, 150 mM NaCl and 0.005% Tween20, and was filtered (0.2 µm filter) and degassed for at least 30 minutes. Each sensorchip surface with attached particles was washed for at least 5 hours at a flow rate of 50-100 µl/min and then equilibrated at 10 µl/min for at least 30 minutes or until the baseline was stable. Single proteins or full plasma were alkylated with iodoacetamide to eliminate free thiol groups to avoid covalent coupling to the gold, using 25 mM iodoacetamide in 25 mM ammonium hydrogen carbonate

PREOMICS EXHIBIT 1002
Page 405 of 1861

buffer followed by gel filtration on a Nap-10 column (Amersham Biosciences) to remove excess reagent. The proteins were diluted in the flow buffer and injected over 30 minutes to study the association kinetics. After 30 minutes, buffer was flown over the sensorchip surface for 10-24 hours at 10 µl/min.

Association and dissociation data were fitted using equations 1 and 2, respectively

$$R(t) = C1 \left(k^{1on}/(k^{1on}+k^{1off}) \right) (1-\exp(-(k^{1on}+k^{1off})t)) + C2 \left(k^{2on}/(k^{2on}+k^{2off}) \right) (1-\exp(-(k^{2on}+k^{2off})t)) \quad (1)$$

$$R(t) = A1 \exp(-k^{1off} t) + A2 \exp(-k^{2off} t) \quad (2)$$

Data for single proteins were fitted assuming single processes, i.e. C2 and A2 were set to zero.

Example 9

RGD modification of Poly(vinyl alcohol) particles

Surface functionalisation of poly(vinyl alcohol) particles with RGD was carried out by means of the Schiff base coupling with residual aldehyde groups at the particle surface, followed by reductive amination. Qualitative analysis was performed by Elecrospray Ionisiation (ESI) mass spectrometry performed on a diluted suspension of RGD-functionalized MBs. Four peaks denote the presence of the RGD molecule on the surface: an RGD molecular peak at m/z = 343, an RGD Na^+, an RGD.2Na^+ and an RGD.3Na^+, at m/z = 365, 383, and 403, respectively. Quantitative determination of RGD tethered on the particle surface was performed by the microBCA colorimetric assay, and yielded a value of $5 \cdot 10^{-2}$ µmol RGD *per* mg of polyvinyl alcohol particles.

1D Polyacrylamide Gel Electrophoesis (PAGE)

Polyacrylamide gels (12%) were used to separate the bound proteins.

Mass spectrometry

Bands of interest from SDS-PAGE gels (12%) were excised and digested in-gel with trypsin according to the method of Shevchenko et al[14]. The resulting peptide mixtures were re-suspended in 0.1% formic acid and analyzed by electrospray liquid chromatography mass spectrometry (LC MS/MS). An HPLC (Surveyor, ThermoFinnigan, CA) was interfaced with an LTQ ion trap mass spectrometer (ThermoFinnigan, CA). Chromatography buffer solutions

PREOMICS EXHIBIT 1002
Page 406 of 1861

(Buffer A, 0.1% formic acid; Buffer B, 100% acetonitrile and 0.1% formic acid) were used to deliver a 72 min gradient (5 min sample loading, 32 min to 40% Buffer B, 2 min to 80%, hold 11 min, 1min to 0%, hold for 20 min, 1 min flow adjusting). A flow rate of 150 µl/min was used at the electrospray source. Spectra were searched using the SEQUEST algorithm[15] against the Indexed uniprot/swiss prot database (http://www.expasy.org; release 3 July 2007). The probability-based evaluation program, Bioworks Browser was used for filtering identifications; proteins with Xcorr (1,2,3) = (1.90, 2.00, 2.50) and a peptide probability of $1e^{-5}$ or better were accepted.

The invention is not limited to the embodiment hereinbefore described which may be varied in construction and detail without departing from the spirit of the invention.

Example 10

Modification of nano-SiO_2 particles

In a typical procedure, a definite amount of MPTMS was added to 30mL of ethanol/water mixture under agitation and hydrolyzed at 20°C for 6h. 10g of nano-SiO_2 particles were added to 80mL of ethanol/water mixture under agitation. The solution of hydrolyzed MPTMS was dropped to the nano-SiO_2 dispersion and reacted at 55°C for 24h. After reaction, the reaction mixture was centrifuged. The collected nano-SiO_2 particles were washed with water and ethanol respectively. The washed nano-SiO_2 particles were collected and dried under vacuum at 60°C for 10h (Zhang *et al.*, 2008)..

Preparation of PS/nano-SiO_2 composites

The synthesis strategy is presented schematically below. A representative preparation procedure is as follows. The modified nano-SiO_2 particles were ultrasonically dispersed into water for 15 min, and then pH value of the nano-SiO_2 dispersion was adjusted to a definite value using 1 M HCl or NaOH solution. AIBN was dissolved in styrene to form oil phase, and subsequently the oil phase was mixed with the nano-SiO_2 dispersion. A stable Pickering emulsion was generated using SCIENTZ 450 W digital sonifier for 6 min at 70% amplitude. The resulted Pickering emulsion was poured into a 100 mL three-neck flask equipped with a nitrogen inlet and a reflux condenser. The emulsion was agitated mildly (50 rpm) and polymerized at 78 °C for 24 h. The precipitates after filtration were washed with water and ethanol for three times respectively, and then a little amount of product was taken out and

PREOMICS EXHIBIT 1002
Page 407 of 1861

diluted using ethanol to prepare samples for TEM and SEM. The remaining product was dried at 60 °C under vacuum for 12 h (Zhang *et al.*, 2008).

Example 11

Recovery of organelle-specific proteins following nanoparticle uptake and sub-cellular localisation

Cell culture

A549 cells (passage 1-30 after defrosting from liquid nitrogen; original batches from ATCC, number CCL-185, at passage number 105 or 82) were cultured at 37°C in 5% CO_2 in Minimum Essential Medium (MEM, with additional L-Glutamine) supplemented with 10% Fetal Calf Serum (FCS, Gibco), 1% penicillin/streptomycin (Invitrogen Corp.), and 1% MEM non-essential amino acids (HyClone). Cells were confirmed to be mycoplasma negative using the mycoAlert kit (Lonza Inc. Allendale, NJ) and were tested monthly.

Nanoparticles

Polystyrene nanoparticles (YG plain polystyrene 50nm and YG carboxylate polystyrene 50nm from Polysciences; Green and Red carboxylate polystyrene 40nm from Invitrogen) were used without further modification or purification. These commercial fluorescent samples are commonly labelled by incorporation of a fluorescent marker into the glassy polymer matrix. A sample of the pure YG dye has been kindly provided by Polysciences. All stock solutions were stored at 4 °C.

Nanoparticle dispersions were prepared by diluting the concentrated nanoparticle stock solutions into the complete medium used for cell culture at room temperature, immediately prior to the experiments on cells, with an identical time delay between diluting and introducing to the cells for all experiments. The medium was kept at room temperature and not pre-warmed to 37°C to ensure better nanoparticle dispersions.

Nanoparticle size and charge determination in dispersion

Particle size distribution and zeta potential were determined using a photon correlation spectrophotometer (Malvern Zetasizer Nano ZS). Measurements were performed at 25 °C and

PREOMICS EXHIBIT 1002
Page 408 of 1861

37 °C, in different solvents, i.e. water, Dulbecco's Phosphate Buffered Saline (DPBS, without Ca^{2+} and Mg^{2+}, Gibco) and also in the complete MEM (cMEM) used for cell culture. The particle size distribution and zeta potential data are presented in Table S1.

Fluorescence microscopy

For confocal microscopy, 4.0 x 10^4 cells were seeded onto glass slides (Falcon, 4 well slides) and incubated for 24h prior to addition of particles. The cell number was set to ensure a cell density comparable to the flow cytometry experiments and, in order to keep all parameters affecting the experiment constant, the same protocols were used for exposure to particles, sample preparation and cell fixation. Thus, particle dispersions were prepared at room temperature just before addition to the cells and, after particle exposure, medium was removed and all samples were washed thrice with DPBS, fixed with 4% formalin solution neutral buffered, and the nucleus stained with 4',6-diamidino-2-phenylindole (DAPI blue), before analysis.

A confocal microscope (Carl Zeiss LSM 510 UVMETA, Thornwood. NY) was used to capture images of the intracellular environment and the sub-cellular localisation of the fluorescent polymeric nanoparticles. For multi-color microscopy, samples were excited with 364nm (blue channel) and 488nm (green channel) laser lines, and images were captured by multi-tracking to avoid bleed-through between the fluorophores. To achieve the necessary signal to noise ratio, while obtaining the thinnest possible optical slices, the pinhole diameters were set to less than 1 airy unit. After adjustment of the pinholes of both lasers to obtain the same optical slices, the optimal optical section that fulfilled our criteria was in the range 0.7 - 0.8 μm at magnification 63X. The gain and offset for the different channels were kept constant along the full series of experiments in order to allow quantitative comparison of the cell fluorescence intensities.

Immunostaining and co-localisation

For the co-localisation study, after exposure to nanoparticles and fixation as described above, cells were permeabilised with 0.1% Saponin (Sigma Aldrich) for 5 minutes before blocking non specific binding sites with 10% Bovine Serum Albumin (Sigma Aldrich) in Phosphate Buffered Saline Tween-20 (PBST) for 30 minutes. Cells were then incubated with anti Lamp1 or EEA1 antibodies (ABCAM) for 1 hour at room temperature. Cells were washed 3 times (3 minutes each) with PBS, and then incubated with Alexa488 conjugated secondary antibodies

PREOMICS EXHIBIT 1002
Page 409 of 1861

for 1 hour (Molecular Probes). The cells were washed as before (3 times for 3 minutes each with PBS) and then stained with DAPI for 3 minutes. The mounting medium was Mowiol 4-88 (Calbiochem) or 50% glycerol.

Co-localisation of lysosomes and early endosomes with nanoparticles was quantified with the 'Co-localisation' plug-in for ImageJ software (http://rsb.info.nih.gov/ij/). Backgrounds were subtracted from the raw images, and out of focus blur fluorescence coming from over the nuclei was excluded in order to avoid false positives. Co-localized pixels are shown in white in the images. The co-localisation images have been corrected to enhance the region of interest.

Example 12

Pulsed application of nanoparticles to cells and time-resolved localisation for recovery of proteins from specific organelles along the trafficking pathway

Nanoparticle characterization.

Silica dioxide (SiO_2) nanoparticles were purchased from G. Kisker- Products for Biotechnology (Steinfurt, Germany) at sizes of 50, 100nm and 300nm with green fluorescent labels. To confirm that the size of the nanoparticles matched the size as stated by the manufacturers, EM pictures of the dried nanoparticles were taken. Particle dispersions were characterized at concentrations of 100µg/ml in millipore water, PBS, and the cell culture media, using a Malvern Zetasizer Nano ZS90 (Worcestershire, UK) to measure the hydrodynamic radius by Dynamic light scattering (DLS) and the zeta potential (surface charge). The samples in cell culture media have been characterized for up to 24h of incubation at 37°C, in order to obtain a better description of the evolution of the protein corona formed upon contact with the serum and to study their stability against agglomeration during the full length of the exposure to cells. The emission and excitation spectra of the Fluorescent SiO_2 nanoparticles were produced using a Perkin-Elmer LS 50B fluorimeter (Perkin-Elmer, Waltham, Massachusetts).

Cell Culture.

A549 cells (passage 1-30 after defrosting from liquid nitrogen; original batches from ATCC, item number CCL-185, at passage number 105 or 82) were cultured at 37 °C in 5% CO_2 in

PREOMICS EXHIBIT 1002
Page 410 of 1861

Minimum Essential Medium (MEM, with additional L-Glutamine) supplemented with 10% Fetal Calf Serum (FCS, Gibco), 1% penicillin/streptomycin (Invitrogen Corp.), and 1% MEM non-essential amino acids (HyClone). Cells were confirmed to be mycoplasma negative using the mycoAlert kit (Lonza Inc. Allendale, NJ) and were tested monthly.

Cellular treatments and nanoparticle dispersion in cell medium.

Cells were plated at a density of 2.5 x 10^5 cells in a 6cm plate and allowed to adhere for 24 hours before exposure to 100ug/ml SiO_2 nanoparticle dispersions. Nanoparticle dispersions were prepared by diluting the concentrated nanoparticle stock solutions into the complete medium used for cell culture at room temperature, immediately prior to the experiments on cells, with an identical time delay between diluting and introducing to the cells for all experiments. The medium was kept at room temperature and not pre-warmed to 37°C to ensure better nanoparticle dispersions.

Cells were incubated with nanoparticles for the required times, depending on the experiment, and then the medium was discarded. After the required import incubation time, medium was removed and the samples were washed thrice with DPBS. Energy dependence of the uptake of the SiO_2 nanoparticles was determined by pre-incubating cells for 60 minutes at either 4°C or in media containing 5mg/ml sodium azide (Invitrogen). These energy-depleting conditions were maintained for the duration of the uptake experiments.

Electron Microscopy.

A549 cells treated as described above were fixed at room temperature in 2.5% glutaraldehyde in 0.1 M Sorensen phosphate buffer(pH 7.3) for 1h, rinsed with Sorensen phosphate buffer (pH 7.3), and then post-fixed for 1h in 1% osmium tetroxide in deionised water. After dehydrating the samples in increasing concentrations of ethanol (from 70% up to 100%), they were then immersed in an ethanol/Epon (1:1 vol/vol) mixture for 1h before being transferred to pure Epon and embedded at 37 °C for 2h. The final polymerization was carried out at 60 °C for 24h. Ultrathin Sections of 80nm, obtained with a diamond knife using an ultramicrotome Leica U6, were mounted on copper grids, and stained with uranyl acetate and lead citrate before being examined with an FEI TECNAI transmission electron microscope.

PREOMICS EXHIBIT 1002
Page 411 of 1861

Example 13

Nanoparticle characterization

50nm NH_2-modified fluorescent polystyrene nanoparticles (Sigma Aldrich) and 50nm unmodified and COOH-modified unlabelled polystyrene nanoparticles (Polysciences) were used without further modification. Size and ζ-potential were determined using a Malvern Zetasizer Nano Series. Polystyrene nanoparticles were diluted to 50μg/ml in PBS before measurement. Measurements were conducted at pH 7.0 and 25 °C.

Cell culture

Human brain astrocytoma 1321N1 cells were cultured at 37 °C in Dulbecco's modified Eagle medium (DMEM) supplemented with 10% Foetal Bovine Serum (Gibco) in a humidified atmosphere of 5% CO_2/95% air. Cells were routinely subcultured 1:5 by incubating them in 0.25% trypsin (Gibco) on reaching confluency.

Analysis of cell viability using YoPro-1/PI co-staining

Cells (3×10^4 cells/cm^2) grown for 24 hours were incubated with various concentrations of nanoparticles for 24h, or with a 50 μg/ml nanoparticle dispersion for varying amounts of time at 37 °C. Incubation with 1 μM of Staurosporine (Calbiochem) was used as a positive control for apoptosis, bearing in mind that the mechanism of action of the Staurosporine molecule is not likely to be similar to that of the nanoparticles, due to significant differences in how nanoparticles are processed compared to the passive diffusion of molecules. Following incubation, all cells, including those in the supernatant, were harvested. Cells were incubated for 5 minutes with 100nM YoPro-1 (Molecular Probes) and with 20□g/ml Propidium Iodide (PI) (Sigma Aldrich) for 1 minute on ice. Analysis was performed by Flow Cytometry using a Cyan ADP cytometer (DAKO) using 488nm excitation and measuring fluorescence emission at 530nm and 575nm. Viable cells exclude both dyes and are YoPro-1^-/PI^-. Cells in early phase apoptosis show increased permeability to YoPro-1 and remain impermeable to PI (YoPro-1^+/PI^-), while cells in late phase apoptosis or those undergoing secondary necrosis are permeable to both YoPro-1 and PI (YoPro-1^+/PI^+). Post-acquisition analysis was carried out using the Summit software (DAKO).

Measurement of Caspase 3/7 activity and cellular ATP content

PREOMICS EXHIBIT 1002
Page 412 of 1861

Measurement of Caspase 3/7 activity was carried out using the Caspase-Glo 3/7 assay (Promega) and cellular ATP content was determined using the CellTiter-Glo assay (Promega), according to the manufacturer's instructions. Briefly, cells were incubated in a 96-well plate with different concentrations of nanoparticles for 24h, or with a 50μg/ml nanoparticle dispersion for varying amounts of time at 37 °C. After incubation, an equal volume of the assay reagent was added to the cells and incubation was continued for a further 1 hour at room temperature. Luminescence was measured using a WALLAC VICTOR^{2TM}, 1520 Multilabel Counter. Results were normalised against the untreated control.

Preparation of whole cell extracts and Western-blotting

Following incubation with nanoparticles, 3×10^5 cells were washed with PBS and lysed for 5 minutes on ice with RIPA buffer (50mM Tris-Cl pH 7.5; 150mM NaCl; 1% NP-40; 0.1% SDS; 0.5% Sodium deoxycholate) supplemented with Protease inhibitor cocktail (Sigma Aldrich). Lysate was centrifuged at 14,000×g for 5 minutes at 4 °C and the supernatant containing the proteins was stored at -20 °C prior to Western-blot analysis. Proteins from each sample were subjected to electrophoresis in 10% SDS-PAGE. The proteins were transferred to an Optitran nitrocellulose membrane (Whatman) and detected using anti-PARP-1 and anti-GAPDH (Cell Signaling Technologies) diluted 1:1000, following incubation with Attophos fluorescent reagent (Roche).

References

Sacanna *et al.*, Thermodynamically stable pickering emulsions. Physical Review Letters, **98 (15)**, *158301* (2007)

Zhang *et al.*, Pickering emulsion polymerization: Preparation of polystyrene/nano-SiO2 composite microspheres with core-shell structure. Powder Technology, (2009).

PREOMICS EXHIBIT 1002
Page 413 of 1861

Claims

1. A method for the isolation or removal of a cellular component from a cell comprising the steps of applying a pulse of nanoparticles to the cell for a suitable period of time, allowing the nanoparticles to traffic through the cell for a period of time sufficient to allow the nanoparticles locate to and interact with the cellular component to be isolated, and separating the nanoparticles and isolated cellular component from the cell.

2. A method as claimed in Claim 1 in which the pulse time (import incubation time) is less than 30 minutes.

3. A method as claimed in Claim 1 in which the pulse time is less that 15 minutes.

4. A method as claimed in Claim 1 in which the pulse time is at most 10 minutes.

5. A method as claimed in any preceding Claim in which the cellular component is selected from a biomolecule, a cluster or complex of biomolecules, or an organelle.

6. A method as claimed in any preceding Claim in which the biomolecule, or the biomolecule cluster or complex, is an organelle-specific biomolecule or biomolecule cluster or complex.

7. A method as claimed in Claim 6 in which the biomolecule is a protein, or in which the cluster or complex is a protein cluster or complex.

8. A method as claimed in any of Claims 5 to 7 in which the biomolecule or cluster is a low abundance biomolecule or cluster.

9. A method as claimed in any preceding Claim in which the nanoparticles are modified to allow separation or removal of the nanoparticles from the cell.

10. A method as claimed in Claim 9 in which the nanoparticles have a magnetic core or an affinity tag to enable separation.

PREOMICS EXHIBIT 1002
Page 414 of 1861

11. A method as claimed in any preceding Claim in which the cell is lysed prior to separation of the nanoparticles from the cell.

12. A method as claimed in any preceding Claim in which the surface curvature and/or at least one physiochemical property of the nanoparticle surface is selectively modified to concentrate or more selectively bind at least one specific low abundance biomolecule.

13. A method as claimed in any preceding Claim in which the cell is provided in the form of a cell culture.

14. A method for assessing the effect of perturbation on a cell comprising the steps of perturbing the cell, removing a cellular component from the cell according to a method of any of Claims 1 to 13, and assessing changes in the cellular component in response to the perturbation.

15. A method as claimed in Claim 14 in which the cellular component is a specific protein or a panel of proteins from a specific location in the cell.

16. A method as claimed in Claim 15 in which the protein or panel of proteins is an organelle-specific panel or proteins.

17. A method as claimed in any of Claims 14 to 16 in which the cells are perturbed by applying an initial pulse of nanoparticles (perturbing pulse) to the cells prior to the loading pulse (extraction pulse).

18. A method for assessing the status of a disease, condition, pathology, or cell state, which disease, condition, pathology, or cell state is associated with the presence or level of a cellular component, the method comprising the steps of isolating or removing the cellular component associated with the disease, condition, pathology, or cell state by a method according to any of Claims 1 to 13, and correlating the presence or level of the cellular component with disease, condition, pathology, or cell state.

PREOMICS EXHIBIT 1002
Page 415 of 1861

19. A method as claimed in Claim 18 which is a method for diagnosis or prognosis of a disease or condition, wherein the cellular component is a diagnostic or prognostic variable of the disease or condition.

20. A method as claimed in Claim 19 in which the diagnostic or prognostic variable is an organelle-specific biomarker.

21. A method of assessing the effect of a candidate agent on a cell comprising the steps of administering the candidate agent to the cell, isolating a cellular component from the cell by a method according to any of Claims 1 to 13, and comparing the cellular component with a corresponding cellular component isolated from a cell not treated with the candidate agent.

22. A method as claimed in Claim 21 which is a method as assessing the effectiveness of a candidate drug for s specific disease or condition, and wherein the cellular component to be isolated is a biomarker of drug responsiveness for the disease or condition.

23. A method for the selective concentration of at least one specific low abundance biomolecule from a complex mixture of biomolecules including one or more high abundance biomolecules, comprising a step of providing a preparation of nanoparticles in which at least one physiochemical property of the nanoparticle surface is selectively modified to concentrate or more selectively bind the at least one specific low abundance biomolecule, incubating the nanoparticle preparation with the complex mixture of biomolecules to enable the at least one specific low abundance biomolecule bind to the modified surface of the nanoparticle, separating the bound and unbound biomolecules, and, optionally, eluting the at least one specific low abundance biomolecule from the surface of the nanoparticles.

24. A method as claimed in Claim 23 in which the at least one physiochemical property of the nanoparticle surface that is selectively modified is typically selected from the group consisting of: surface curvature; surface charge; surface chemistry; surface functionalisation; and controlled surface morphology.

PREOMICS EXHIBIT 1002
Page 416 of 1861

25. A method as claimed in Claim 24 in which surface curvature of the nanoparticle surface is tuned to facilitate selective binding (and therefore selective concentration) of a specific biomolecule.

26. A method as claimed in Claim 23 or 24 in which surface morphology is selectively modified by a method selected from the group consisting of: patterning the surface to provide areas of differing biomolecule affinity; introduction of porosity of different scales; engineered surface curvature on multiple length scales; and templating the surface with a template of a specific biomolecule.

27. A method as claimed in Claim 26 in which patterning of the surface is provided by a means selected from the group consisting of: forming the nanoparticle by block polymerisation in which the at least two blocks have different chemistries; and forming the nanoparticle using mixtures of at least two different polymers, phase separating the different polymers during polymerisation, and cross-linking the separated polymers during or following phase separation.

28. A method as claimed in Claim 26 in which engineered surface curvature on multiple length scales is provided by employing Pickering emulsions stabilised by finely divided particles for the synthesis of the nanoparticles.

29. A method as claimed in Claim 26 in which templating the surface of the nanoparticle is provided by forming the nanoparticle in the presence of a biomolecule, wherein at least a portion of the biomolecule(s) is exposed/embedded on the surface of the formed nanoparticle, and removing the biomolecules exposed/embedded on the surface of the nanoparticle to leave a template surface texture and moorphology.

30. A method as claimed in any of Claims 23 to 29 in which the physiochemical properties of the surface of the nanoparticles is modified by surface functionalisation selected from the group consisting of: functionalisation with charged groups; functionalisation with groups that alter the hydrophobic or hydrophilic balance of the nanoparticle surface; and functionalisation with a ligand.

PREOMICS EXHIBIT 1002
Page 417 of 1861

31. A method as claimed in any of Claims 23 to 30 in which the surface curvature and at least two other physiochemical properties of the nanoparticle surface are selectively modified to selectively bind the at least one specific low abundance biomolecule.

32. A method of detecting a low abundance biomarker, which method employs a method of any of Claims 23 to 31 in which the nanoparticle surface is modified for the selective concentration of the biomarker, and in which the biomarker is selectively concentrated and then detected / identified after concentration on the nanoparticle surface.

33. A method as claimed in Claim 32 in which the biomarker is detected and/or identified by a protein or antibody array.

34. A method for the purification and harvesting of a low abundance biomolecule, which method employs the method of any of Claims 23 to 31 to selectively concentrate the low abundance biomolecule, wherein the selectively concentrated (bound) biomolecule is recovered and harvested from the nanoparticles.

35. A method as claimed in any of Claims 23 to 34 in which the nanoparticle comprises a core and a separate shell encapsulating the core, wherein the shell is capable of being modified to alter the surface curvature and/or a physiochemical property of the surface independently of the core.

36. A method as claimed in any of Claims 23 to 34 in which the nanoparticles are provided in the form of a solid phase during at least a part of the process of selective concentration.

37. A method as claimed in Claim 36 in which the nanoparticles are capable of being gelled, and wherein the nanoparticle remain gelled during a step of removal of unbound biomolecule from bound biomolecule.

38. A method as claimed in Claim 36 in which the nanoparticles are capable of being reversibly gelled, and wherein the nanoparticles remain gelled during a step of removal

PREOMICS EXHIBIT 1002
Page 418 of 1861

of unbound biomolecules from bound biomolecule, wherein the gellation is reversed during a step of eluting the low abundance biomolecule from the nanoparticles.

39. A method for the selective concentration of at least one specific biomolecule from a complex mixture of biomolecules, typically including one or more high abundance biomolecules, comprising the steps of providing a preparation of nanoparticles in which the nanoparticles are capable of being crosslinked under specific conditions, incubating the nanoparticle preparation with the complex mixture of biomolecules to enable the at least one specific biomolecule bind to a surface of the nanoparticle, crosslinking the nanoparticle preparation either prior to, during, or after the incubation step, separating unbound biomolecule from the crosslinked nanoparticle preparation, and optionally eluting the at least one specific biomolecule from the surface of the nanoparticles.

40. A method as claimed in Claim 39 in which the gellation of the nanoparticles is reversible, and wherein after removal of the unbound biomolecule, the nanoparticle preparation is treated to reverse the gelation whereupon the at least one specific biomolecule is optionally eluted from the nanoparticles.

41. A method as claimed in Claim 39 or 40 in which the specific biomolecule is a low abundance biomolecule.

42. A method of identifying the presence of a specific low abundance biomolecule in a sample such as a complex mixture of biomolecules, which method comprises a step of selectively concentrating the specific low abundance biomolecule according to a method of any of Claims 23 to 31, and identifying the concentrated low abundance biomolecule.

43. A method as claimed in Claim 42 in which the low abundance biomolecule is identified by means of proteomics and mass spectrometry, electrophoresis and mass spectrometry, electrophoresis and western blotting, or protein or antibody arrays.

44. A method as claimed in Claim 42 or 43 in which the low abundance biomolecule is a biomarker associated with a characteristic selected from the group consisting of:

PREOMICS EXHIBIT 1002
Page 419 of 1861

clinical assessment (including diagnosis or prognosis) of a disease, condition, or pathological status; drug response; adverse drug reaction; and response to treatment.

45. A method for the purification of a specific low abundance protein comprising a step of selectively concentrating the specific low abundance biomolecule according to a method of any of Claims 23 to 31, and eluting the low abundance biomolecule from the nanoparticle preparation.

46. A method as claimed in Claim 45 in which elution is carried out in a step-wise manner.

47. A method of preparing a nanoparticle preparation suitable for selectively concentrating a specific low abundance biomolecule comprising the step of modifying the surface curvature of the nanoparticle such that it has a specific binding affinity for the specific low abundance biomolecule and selectively modifying a physiochemical property of the surface of the nanoparticle to concentrate or more selectively bind the at least one specific low abundance biomolecule, wherein the physiochemical property is selected from the group consisting of: surface charge; surface chemistry: surface functionalisation: and controlled surface morphology.

48. A method as claimed in Claim 47 in which surface morphology is selectively modified by a method selected from the group consisting of: patterning the surface to provide areas of differing biomolecule affinity; engineered surface curvature on multiple length scales; and templating the surface with a template of a specific biomolecule (as described above).

49. A nanoparticle preparation that is modified to be capable of forming a solid phase under specific conditions.

50. A nanoparticle preparation as claimed in Claim 49 in which the nanoparticles are modified with a reactive group that, upon activation, is capable of cross-linking the nanoparticles to form a solid phase.

PREOMICS EXHIBIT 1002
Page 420 of 1861

51. A nanoparticle preparation as claimed in Claim 49 or 50 in which the solid phase is a gel.

52. A nanoparticle preparation as claimed in Claim 50 or 51 in which the reactive group is activated by an external stimulus or cue.

53. A nanoparticle preparation as claimed in Claim 52 in which the external stimulus or cue is selected from the group consisting of: temperature; pH; ionic concentration; solvent; buffer; catalyst: oxidation state; and the like.

54. A nanoparticle preparation as claimed in any of Claims 50 to 53 in which the reactive group is selected from the group consisting of: thiol group; -COOH and / or –OH groups; amino acid or amino acid analogues capable of forming of methylene bridges; complementary DNA strands, or other coil-forming polymers capable of hybridisation.

55. A nanoparticle preparation in which a surface of at least a portion of the nanoparticles is imprinted with a template of a specific biomolecule.

56. A nanoparticle preparation as claimed in Claim 55 in which the biomolecule is a protein or a protein cluster or complex.

57. A method of preparing a nanoparticle preparation of any of Claims 55 or 56 comprising the steps of forming the nanoparticles in the presence of the specific biomolecule, typically in a specific conformation, such that at least a portion of the formed nanoparticles have the specific biomolecule exposed/embedded in a surface of the nanoparticle, and treating the formed nanoparticles to remove the exposed/embedded biomolecule to leave an imprint of the specific biomolecule on the surface of the nanoparticle.

58. A method as claimed in Claim 57 in which the specific biomolecule is a protein or a protein complex or cluster, and wherein the embedded protein or protein complex or cluster is removed from the formed nanoparticle by denaturation of the protein using a suitable denaturing agent.

PREOMICS EXHIBIT 1002
Page 421 of 1861

59. A nanoparticle having a patterned surface in which a first portion of the surface has a specific binding affinity for a specific patterning biomolecule, and a second portion of the surface has a different binding affinity for the same specific patterning biomolecule compared with the first portion of the surface.

60. A nanoparticle as claimed in Claim 59 in which the first portion of the surface has a binding affinity for the specific patterning biomolecule and the second portion of the surface has substantially no binding affinity for the specific patterning biomolecule.

61. A nanoparticle as claimed in Claim 59 or 60 in which the patterned nanoparticle is formed of a block co-polymer in which the different blocks have different binding affinities for the specific biomolecule..

62. A nanoparticle having engineered surface curvature on multiple length scales.

63. A nanoparticle as claimed in Claim 62 in which the engineered surface curvature on multiple length scales is provided by employing Pickering emulsions stabilised by finely divided particles for the synthesis of the nanoparticles and polymerising these structures in place.

64. A nanoparticle as claimed in Claim 63 in which the finely divided particles are selected from the group consisting of: silicates; aluminates; titanates; metal oxides such as aluminum, silicon, titanium, nickel, cobalt, iron, manganese, chromium, or vanadium oxides; carbon blacks; and nitrides or carbides, such as boron nitride, boron carbide, silicon nitride, or silicon carbide.

65. A method according to any of Claims 23 to 58 in which the complex mixture of biomolecules is a biological fluid.

66. A method according to Claim 65 in which the biological fluid is selected from the group consisting of: a culture of eukaryotic or prokaryotic (recombinant protein) producer cells; plasma; serum; cell lysates; cytosolic fluid; cell organelles; gastric fluid; amniotic fluid; cerebrospinal fluid; lung lavage fluid; saliva; and urine.

PREOMICS EXHIBIT 1002
Page 422 of 1861

67. A method according to any of Claims 23 to 58 in which the complex mixture of biomolecules is a non-biological fluid.

68. A method as claimed in Claim 67 in which the fluid is an industrial fermentation liquor or an industrial process stream.

69. A method according to Claim 67 or 68 in which the complex mixture of biomolecules comprises a low abundance metabolite.

70. A method as claimed in any of Claims 1 to 58 in which the specific biomolecule or specific low abundance biomolecule is selected from the group consisting of: a protein; a protein complex or cluster; a nucleic acid molecule; a polypeptide; a peptide; a polysaccharide; a sugar; a lipid; a lipoprotein complexe; and a metabolite.

71. A method of producing a recombinant protein, which method employs a eukaryotic or prokaryotic producer cell engineered with a nucleic acid construct encoding the recombinant protein, the method comprising the steps of incubating the producer cells with a nanoparticle preparation suitable for selectively binding the recombinant protein, and recovering/separating the nanoparticle preparation and bound recombinant protein from the producer cells.

72. A method as claimed in Claim 71 in which the nanoparticle preparation is incubated with the producer cells during the production phase of cell growth.

73. A method as claimed in Claim 7 or 72 including the step of eluting the recombinant protein from the nanoparticle preparation.

74. A method of reducing recombinant protein aggregation during high-throughput recombinant protein production, the method comprising the step of incubating a nanoparticle preparation with the recombinant protein producer cells for a period of time, which nanoparticles are suitable for selective binding and/or concentration of the recombinant product, and recovering/separating the nanoparticle preparation and the bound/concentrated recombinant protein from the producer cells.

PREOMICS EXHIBIT 1002
Page 423 of 1861

Figure 1

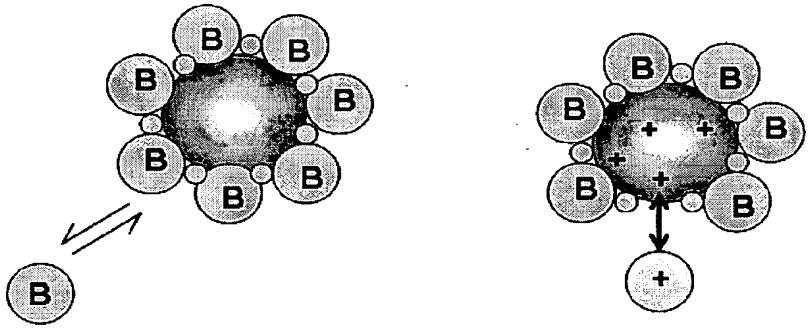


Figure 2

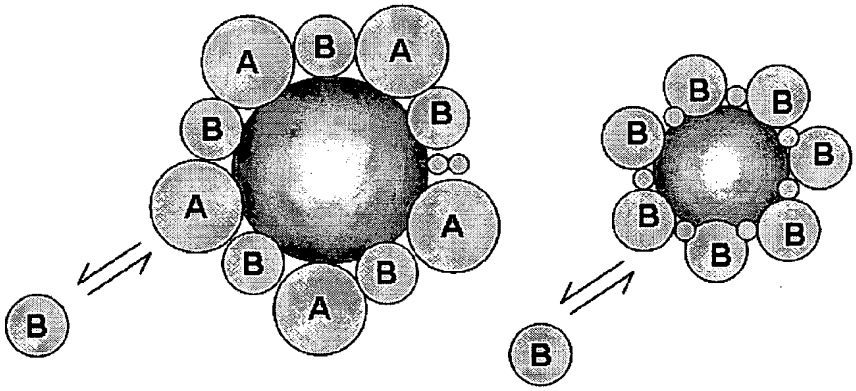
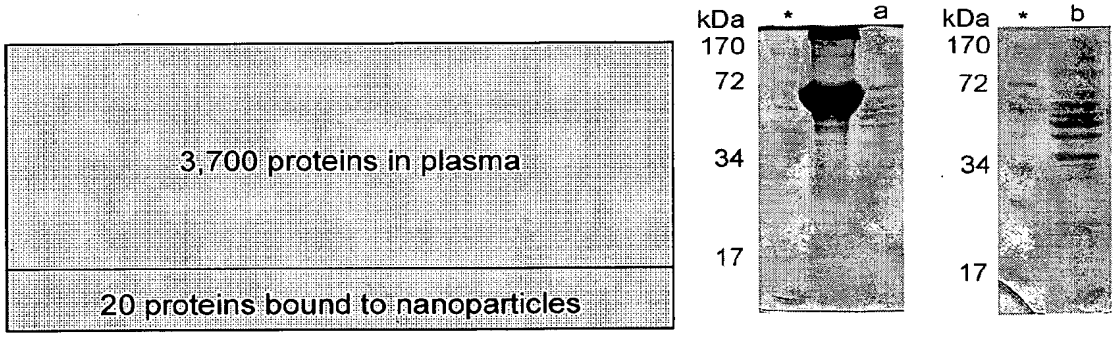


Figure 3



PREOMICS EXHIBIT 1002
Page 424 of 1861

Figure 4

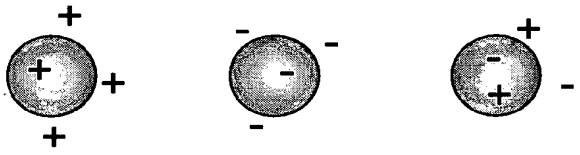


Figure 5

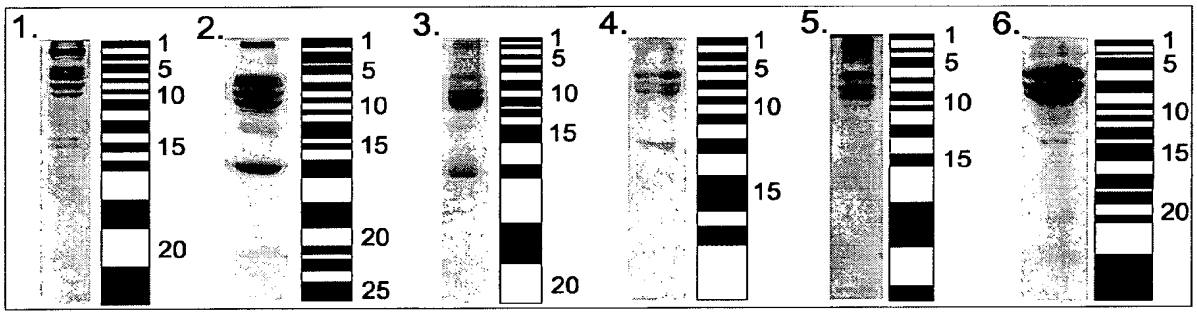


Figure 6

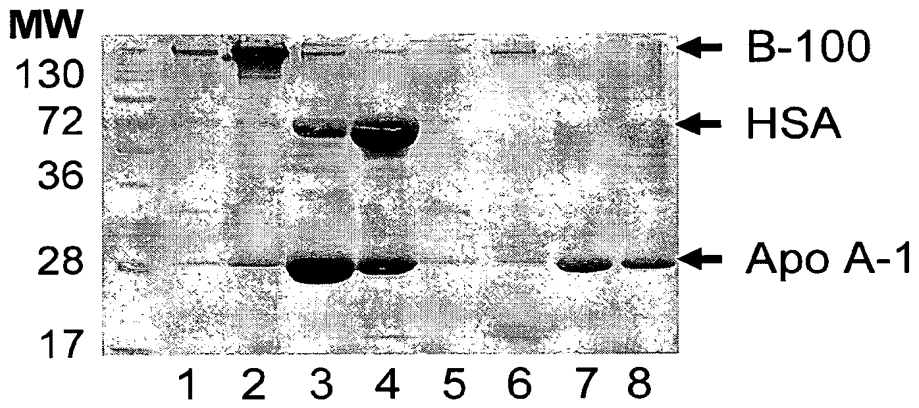


Figure 7

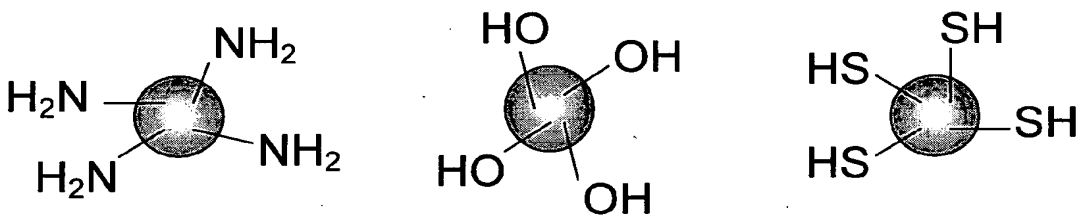


Figure 8

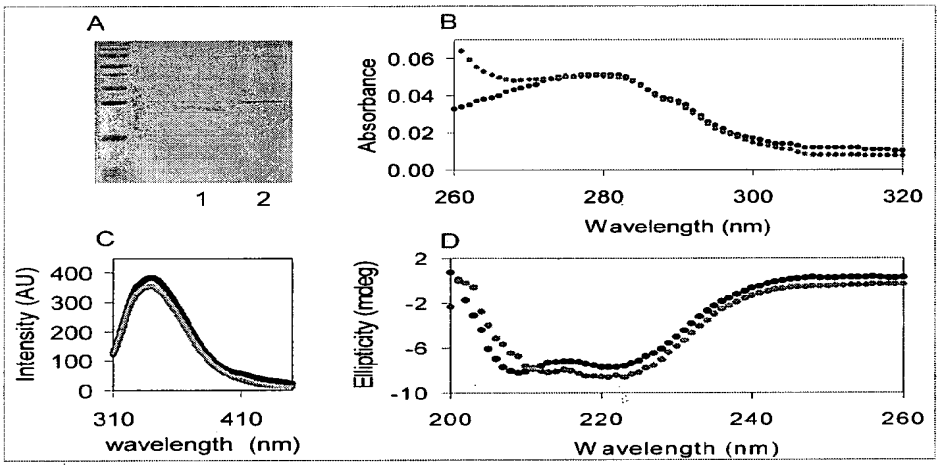
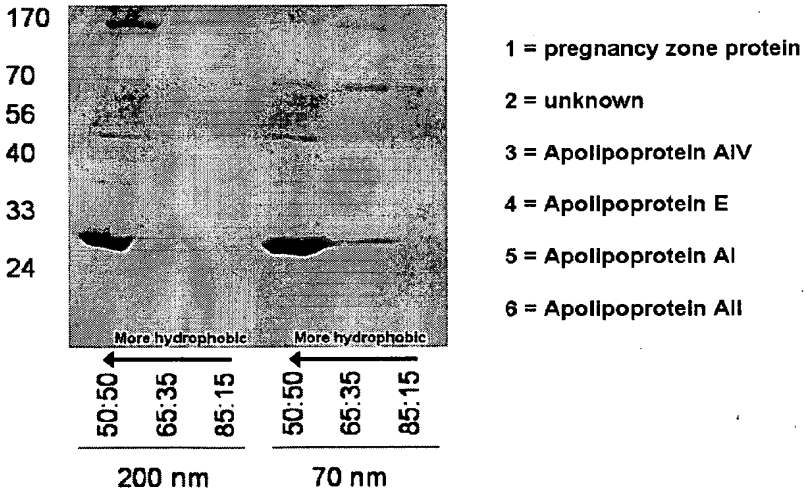


Figure 9



PREOMICS EXHIBIT 1002
Page 426 of 1861

Figure 10

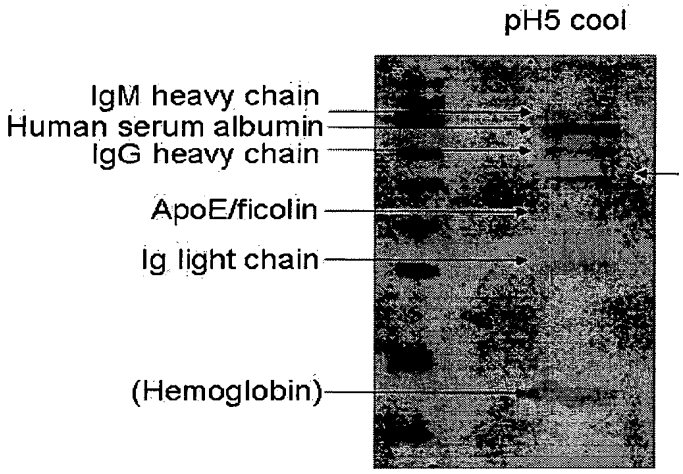


Figure 11

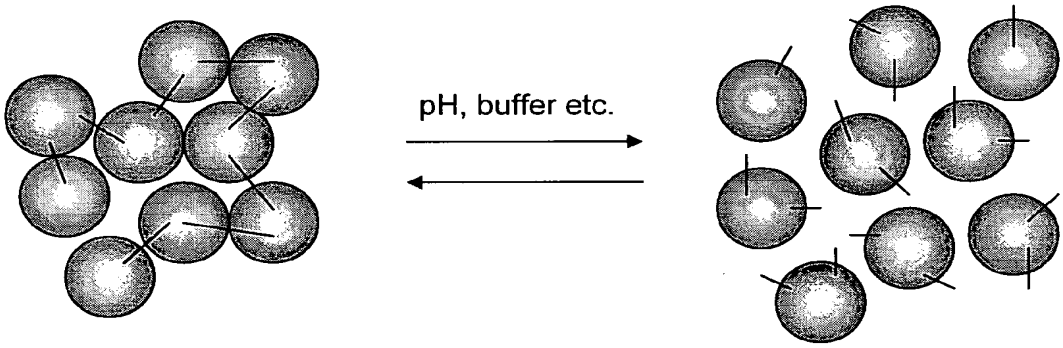
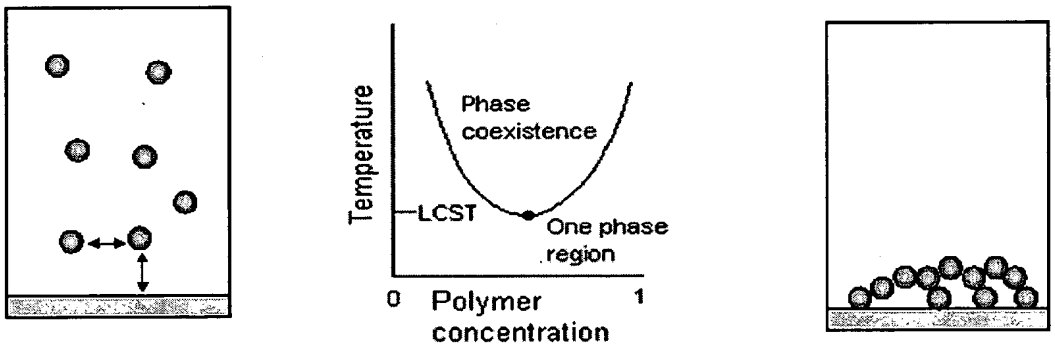


Figure 12



PREOMICS EXHIBIT 1002
Page 427 of 1861

Figure 13

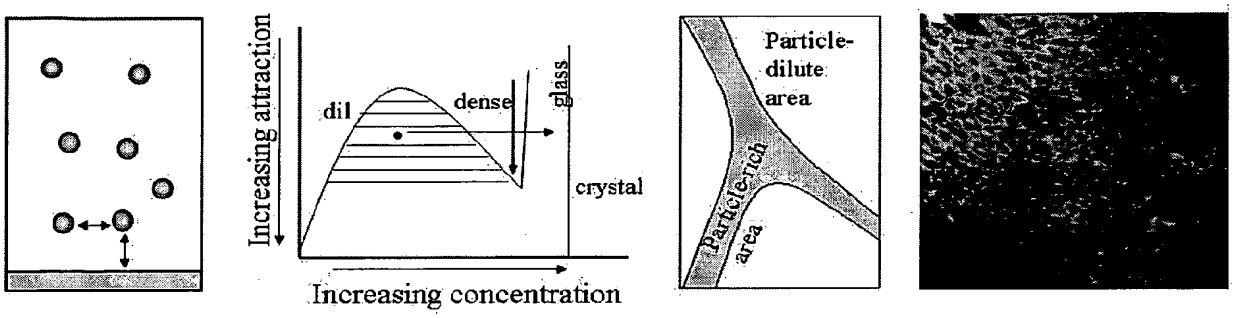
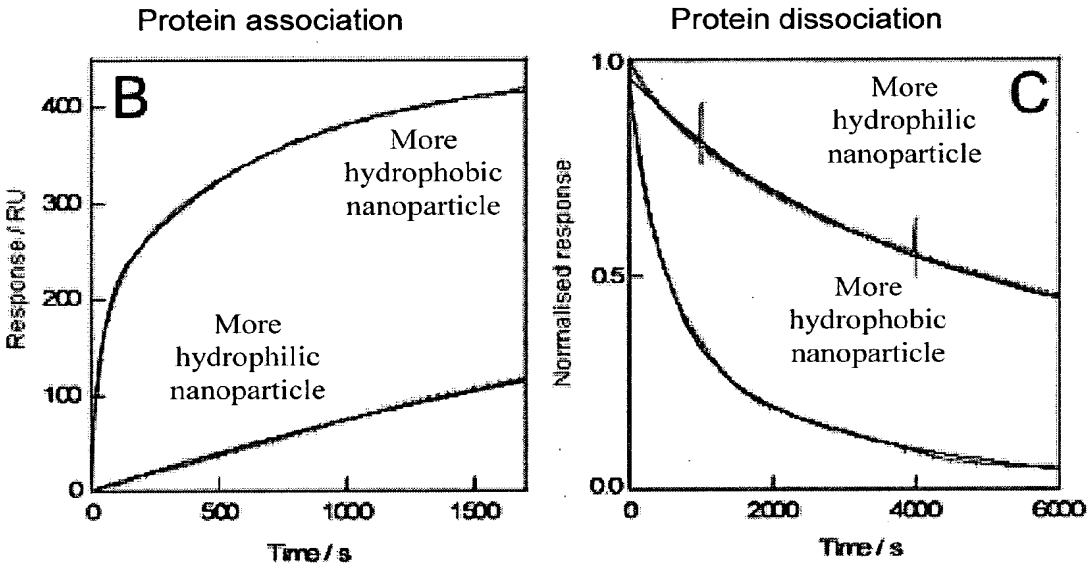


Figure 14



PREOMICS EXHIBIT 1002
Page 428 of 1861

Figure 15

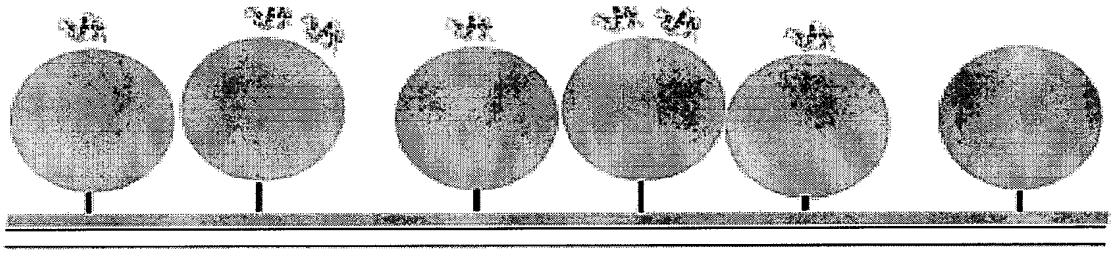


Figure 16

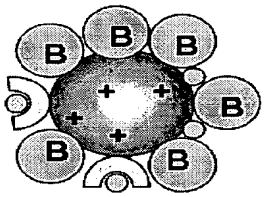


Figure 17

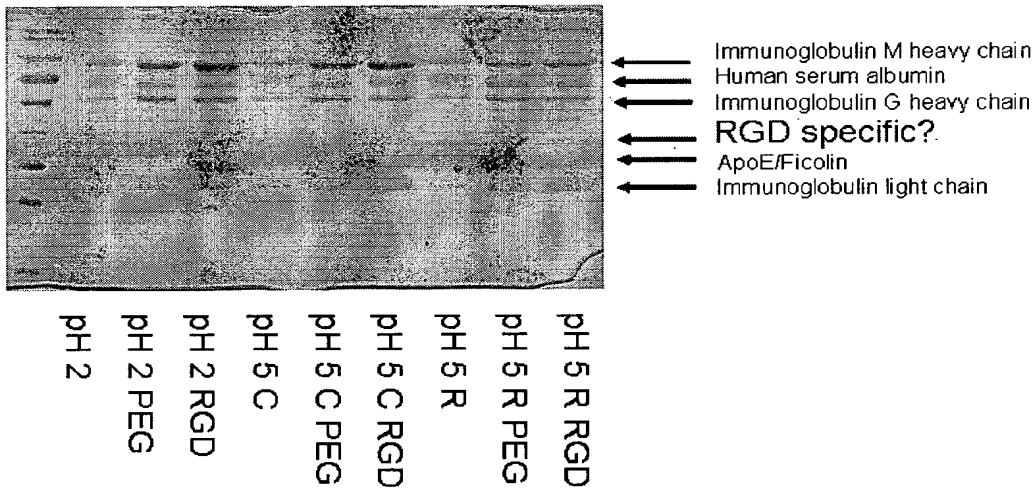


Figure 18

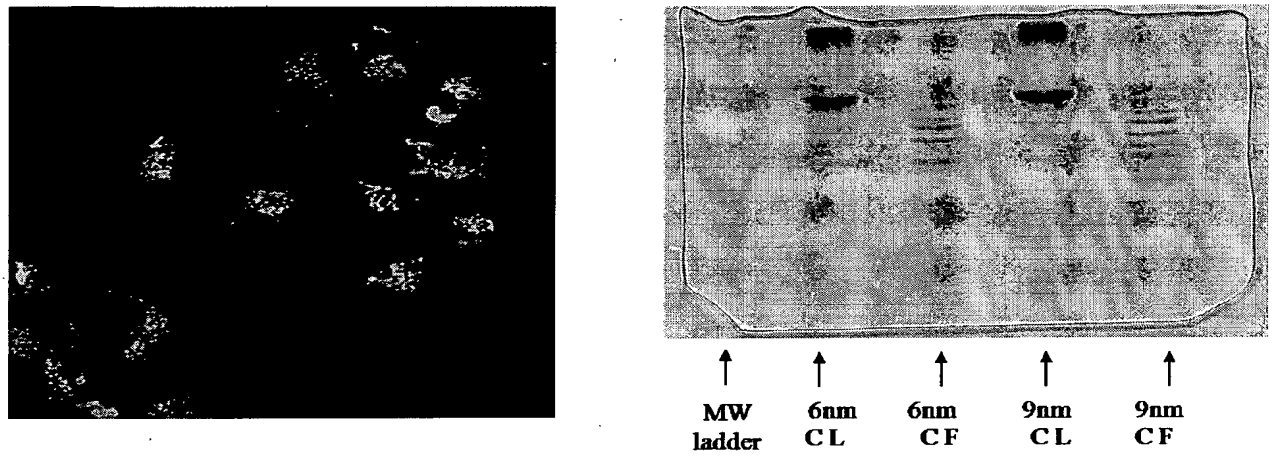
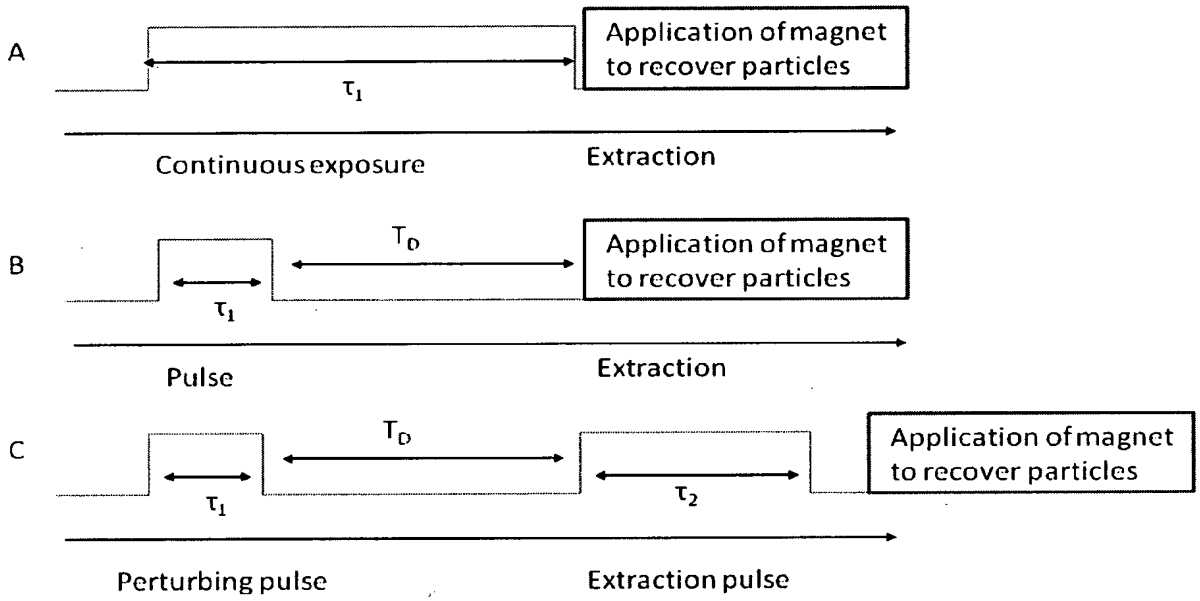


Figure 19



PREOMICS EXHIBIT 1002
Page 430 of 1861

Figure 20

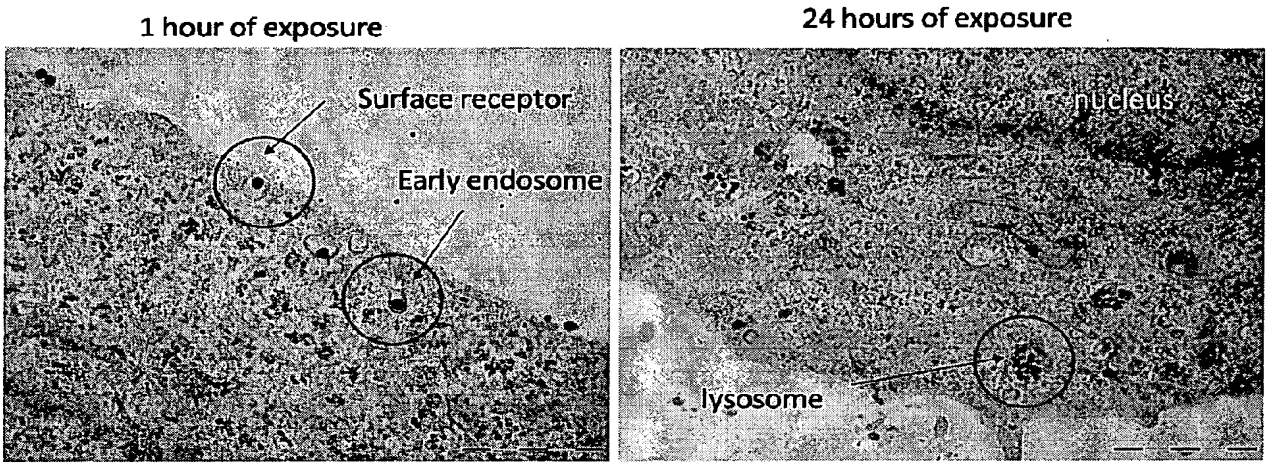
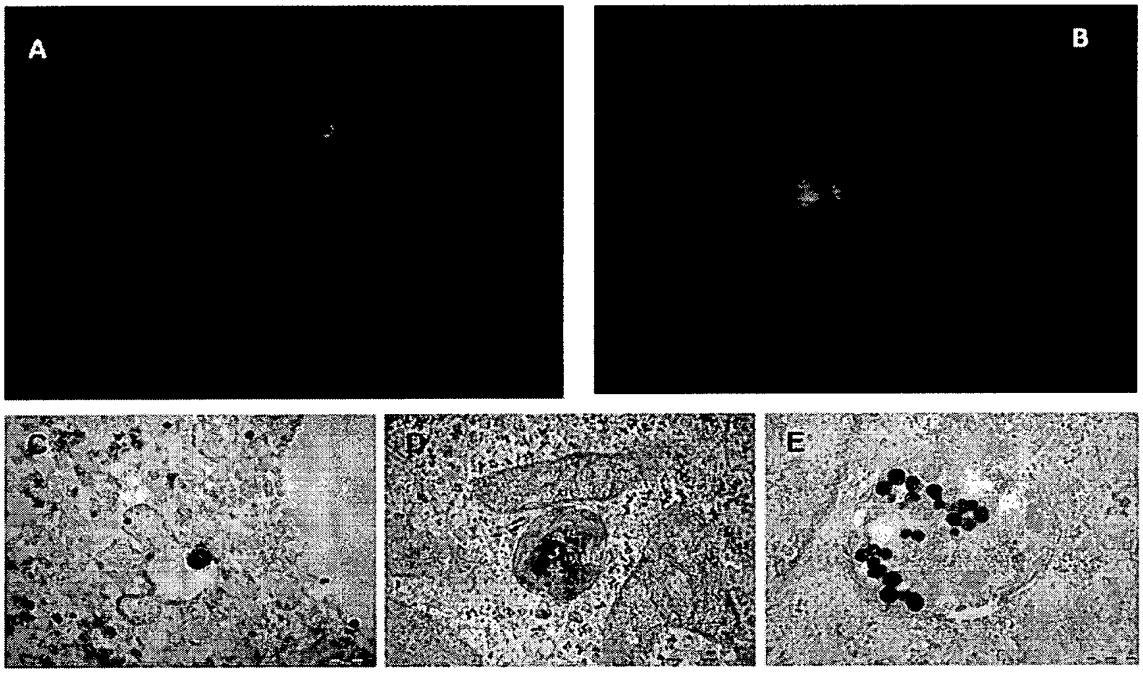


Figure 21



PREOMICS EXHIBIT 1002
Page 431 of 1861

Figure 22

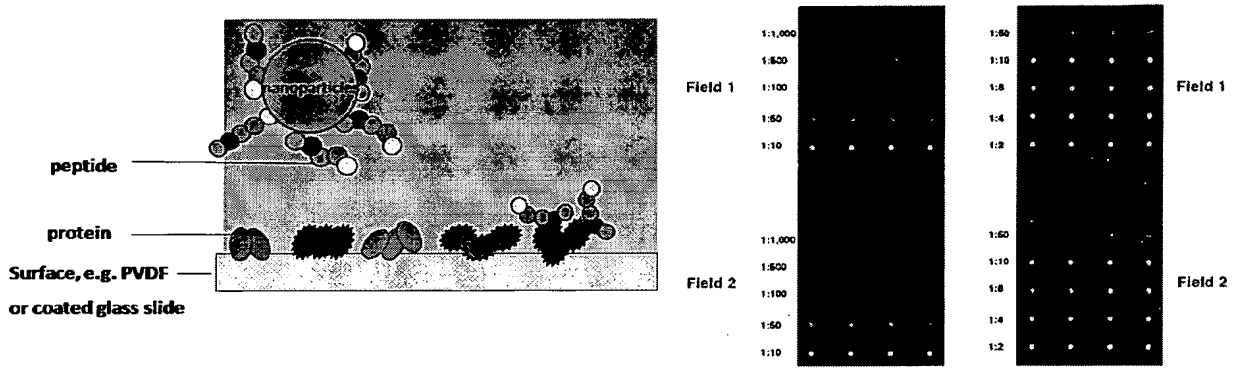


Figure 23

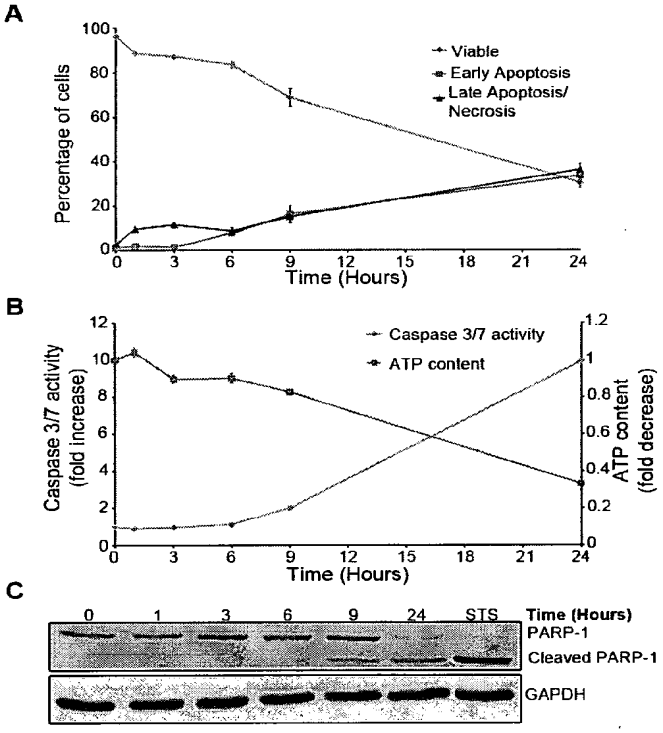


Figure 24



Figure 25

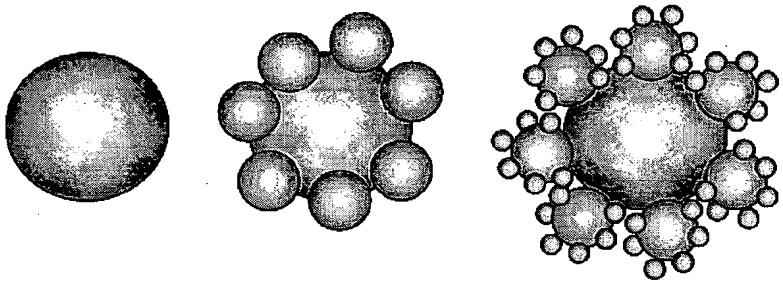


Figure 26

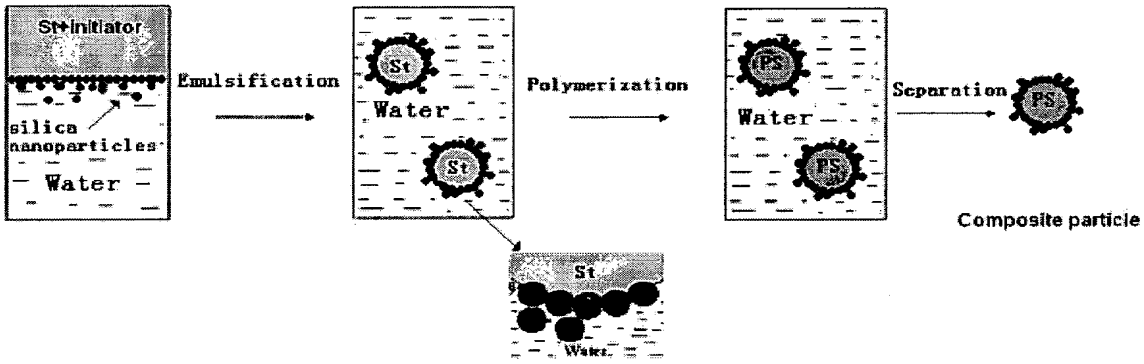
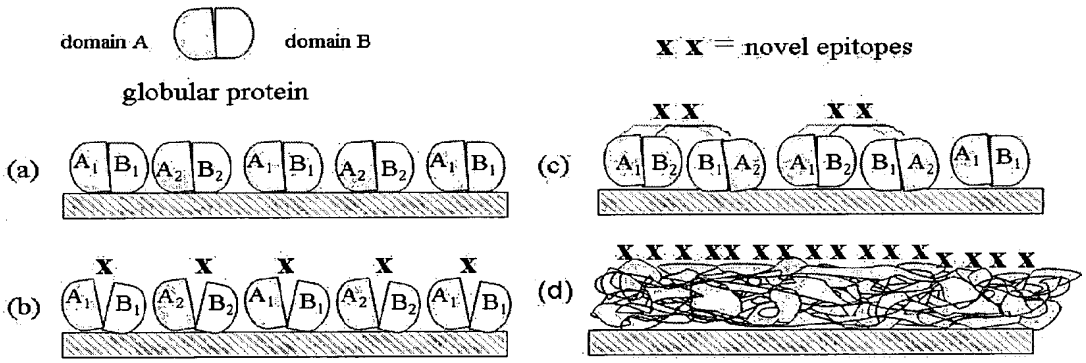


Figure 27



Figure 28



PREOMICS EXHIBIT 1002
Page 433 of 1861

INTERNATIONAL SEARCH REPORT

International application No
PCT/IE2010/000009

A. CLASSIFICATION OF SUBJECT MATTER
INV. G01N33/543
ADD.

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)
G01N

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, EMBASE, 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	SALAKLANG JATUPORN ET AL: "Superparamagnetic nanoparticles as a powerful systems biology characterization tool in the physiological context." ANGEWANDTE CHEMIE (INTERNATIONAL ED. IN ENGLISH) 2008 LNKD- PUBMED:18792049, vol. 47, no. 41, 2008, pages 7857-7860, XP002586198 ISSN: 1521-3773 the whole document figure 2	1-13, 18-22, 47,49, 62-64, 70-74

-/--

[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
9 June 2010	14/07/2010

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, Fax: (+31-70) 340-3016	Mauhin, Viviane

1

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

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WON J ET AL: "A magnetic nanoprobe technology for detecting molecular interactions in live cells" 1 July 2005 (2005-07-01), SCIENCE, AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE, WASHINGTON, DC; US LNKD-DOI:10.1126/SCIENCE.1112869, PAGE(S) 121 - 125 , XP003014064 ISSN: 0036-8075 the whole document	1-17,47, 49, 62-64, 70-74
X	GOTO YUSUKE ET AL: "Polymer nanoparticles covered with phosphorylcholine groups and immobilized with antibody for high-affinity separation of proteins" BIOMACROMOLECULES, ACS, WASHINGTON, DC, US LNKD- DOI:10.1021/BM701161D, vol. 9, no. 3, 1 March 2008 (2008-03-01), pages 828-833, XP002532655 ISSN: 1525-7797 Results and Discussion, first paragraph Results and Discussion: Characterisation of Phospholipid Polymer Nanoparticles, first paragraph Results and Discussion; BSA Adsorption on various Nanoparticles Results and Discussion: Affinity Selection of Target Protein	23-49, 67-70
X	ROACH PAUL ET AL: "Surface tailoring for controlled protein adsorption: effect of topography at the nanometer scale and chemistry." 29 March 2006 (2006-03-29), JOURNAL OF THE AMERICAN CHEMICAL SOCIETY 29 MAR 2006, VOL. 128, NR. 12, PAGE(S) 3939 - 3945 , XP002532656 ISSN: 0002-7863 Results and Discussion: Discussion Conclusions: last paragraph	23-49, 62-70
X	LUNDQVIST MARTIN ET AL: "Nanoparticle size and surface properties determine the protein corona with possible implications for biological impacts" September 2008 (2008-09), PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA, VOL. 105, NR. 38, PAGE(S) 14265-14270 , XP002532657 ISSN: 0027-8424 table 2 Conclusions	23-49, 62-66

-/--

1

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CEDERVALL TOMMY ET AL: "Detailed identification of plasma proteins adsorbed on copolymer nanoparticles." ANGEWANDTE CHEMIE (INTERNATIONAL ED. IN ENGLISH) 2007, vol. 46, no. 30, 2007, pages 5754-5756, XP002532658 ISSN: 1433-7851 the whole document	23-49, 62-66
X	WO 2008/131063 A1 (3M INNOVATIVE PROPERTIES CO [US]; JING NAIYONG [US]; SCHULTZ WILLIAM J) 30 October 2008 (2008-10-30) the whole document	47-64
X	KIM JASON S ET AL: "Highly-efficient purification of native polyhistidine-tagged proteins by multivalent NTA-modified magnetic nanoparticles" March 2007 (2007-03), BIOCONJUGATE CHEMISTRY, VOL. 18, NR. 2, PAGE(S) 333-341 , XP002586199 ISSN: 1043-1802 the whole document	71-73
X	WO 2007/140388 A2 (UNIV JOHNS HOPKINS [US]; HOH JAN [US]; HEINZ WILLIAM F [US]) 6 December 2007 (2007-12-06) the whole document	59-61
X	LUNDQVIST M ET AL: "Protein adsorption onto silica nanoparticles: conformational changes depend on the particles' curvature and the protein stability" 23 November 2004 (2004-11-23), LANGMUIR : THE ACS JOURNAL OF SURFACES AND COLLOIDS 23 NOV 2004, WASHINGTON DC USA, PAGE(S) 10639 - 10647 , XP002519286 the whole document	47,48, 62-64
X	KIRSTEN BORCHERS ET AL: "Microstructured layers of spherical biofunctional core-shell nanoparticles provide enlarged reactive surfaces for protein microarrays" ANALYTICAL AND BIOANALYTICAL CHEMISTRY, SPRINGER, BERLIN, DE, vol. 383, no. 5, 1 November 2005 (2005-11-01), pages 738-746, XP019327665 ISSN: 1618-2650 the whole document	47-64

-/--

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	REZWAN K ET AL: "Lysozyme and bovine serum albumin adsorption on uncoated silica and AlOOH-coated silica particles: the influence of positively and negatively charged oxide surface coatings" BIOMATERIALS, ELSEVIER SCIENCE PUBLISHERS BV., BARKING, GB, vol. 26, no. 21, 1 July 2005 (2005-07-01), pages 4351-4357, XP025280494 ISSN: 0142-9612 [retrieved on 2005-07-01] the whole document	47-64
X,P	WO 2009/047587 A1 (ECOLE POLYTECH [CH]; STEITZ BENEDIKT [CH]; PETRI-FINK ALKE [CH]; SALAK) 16 April 2009 (2009-04-16) the whole document	1,5-13, 47,49, 62-64,70
X,P	SHAN Z ET AL: "Bacteria capture, lysate clearance, and plasmid DNA extraction using pH-sensitive multifunctional magnetic nanoparticles" ANALYTICAL BIOCHEMISTRY, ACADEMIC PRESS INC, NEW YORK, vol. 398, no. 1, 10 November 2009 (2009-11-10), pages 120-122, XP026836327 ISSN: 0003-2697 [retrieved on 2009-11-10] the whole document	1,5-13, 47,49, 62-64, 70-73

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/IE2010/000009

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2008131063 A1	30-10-2008	EP 2140264 A1	06-01-2010
WO 2007140388 A2	06-12-2007	CA 2653792 A1	06-12-2007
		EP 2035541 A2	18-03-2009
		US 2009170723 A1	02-07-2009
		US 2007281310 A1	06-12-2007
WO 2009047587 A1	16-04-2009	NONE	

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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
23 December 2010 (23.12.2010)

PCT

(10) International Publication Number
WO 2010/148365 A2

(51) International Patent Classification:
G01N 33/53 (2006.01) *G01N 33/68* (2006.01)
G01N 33/564 (2006.01) *C40B 40/10* (2006.01)

(21) International Application Number:
PCT/US2010/039269

(22) International Filing Date:
18 June 2010 (18.06.2010)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/218,890 19 June 2009 (19.06.2009) US
61/249,147 6 October 2009 (06.10.2009) US

(71) Applicant *(for all designated States except US)*: **THE ARIZONA BOARD OF REGENTS, A BODY CORPORATE OF THE STATE OF ARIZONA FOR AND ON BEHALF OF ARIZONA STATE UNIVERSITY** [US/US]; 699 S. Mill Avenue, The Brickyard, Suite 601, Room 691 AA, Tempe, Arizona 85281 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **JOHNSTON, Stephen Albert** [US/US]; 8606 S. Dorsey Lane, Tempe, Arizona 85284 (US). **STAFFORD, Philip** [US/US]; 3419 E. Kachina Drive, Phoenix, Arizona 85044 (US).

(74) Agents: **LIEBESCHUETZ, Joe** et al.; Townsend and Townsend and Crew LLP, Two Embarcadero Center, 8th Floor, San Francisco, CA 94111 (US).

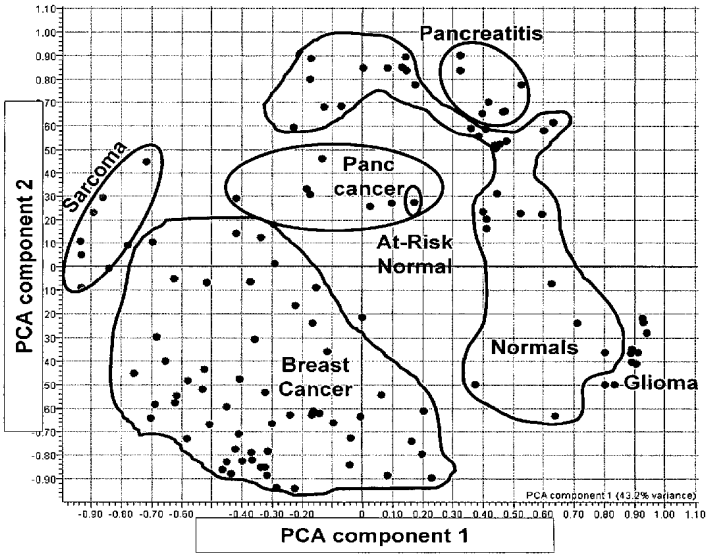
(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ,

[Continued on next page]

(54) Title: COMPOUND ARRAYS FOR SAMPLE PROFILING

Fig. 6



(57) Abstract: The invention provides arrays of compound for use in profiling samples. The arrays include compounds bind to components of the samples at relatively low affinities. The avidity of compounds binding to components of the samples can be increased by forming arrays such that multivalent components of the samples (e.g., antibodies or cells) can bind to more than one molecule of a compound at the same time. When a sample is applied to an array under such conditions, the compounds of the array bind to component(s) of the sample with significantly different avidities generating a profile characteristic of the sample.

WO 2010/148365 A2

TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:

— *as to non-prejudicial disclosures or exceptions to lack of novelty (Rule 4.17(v))*

Published:

— *without international search report and to be republished upon receipt of that report (Rule 48.2(g))*

COMPOUND ARRAYS FOR SAMPLE PROFILING

CROSS-REFERENCE TO RELATED CASES

[0001] The present application claims priority from USSN 61/218,890 filed June 19, 2009 and USSN 61/249,147 filed October 6, which are incorporated by reference in their entirety for all purposes.

BACKGROUND OF THE INVENTION

[0002] ELISA reactions and other simple immunoassays are commonly used for diagnosing disease. Many assays are configured to detect a single analyte. Therefore, when several differential diagnoses are possible, several different such assays are often conducted in parallel.

[0003] Existing approaches for broadly characterizing an immune response involve multiple standard ELISAs, use of library panning involving multiple rounds of selection, or printing of known proteins from pathogens or host proteins in an array format to detect antibodies to pathogens or autoantibodies. T-cells and B-cells have also been characterized by isolating and cloning specific regions of the T- and B-cell genome to sequence the recombination event. All of these processes are labor intensive and take time. They are not conducive to a standard clinical diagnostic protocol or early detection before specific analytes for which specific binding reagents are provided.

SUMMARY OF THE CLAIMED INVENTION

[0004] The invention provides methods of analyzing a sample, comprising: (a) contacting the sample with an array of immobilized different compounds occupying different areas of the array, wherein different molecules of the same compound within an area are spaced sufficiently proximate to one another for multivalent binding between at least two of the different molecules in the same area and a multivalent binding partner; and (b) detecting binding of the different compounds in the array to component(s) of the sample, such as antibodies. The sample can thus be characterized from the relative binding of compounds whose binding to the sample would be difficult to distinguish from each other and nonspecific binding under conditions of monovalent binding but which show significantly different binding from each other and nonspecific binding due to multivalent binding of the sample to the compounds thereby generating a binding profile characteristic of the sample.

[0005] In some methods, the sample is characterized from the relative binding a plurality of the sample that binds to the different compounds binding to the sample with association

PREOMICS EXHIBIT 1002
Page 441 of 1861

constants of 1 mm to 1 μM. In some methods, the sample is characterized from the relative binding of at least 10 or 100 compounds binding to the sample with association constants of 1 mm to 1 μM. In some methods, the sample is characterized by comparing a binding profile of the sample that includes the relative binding of the plurality of compounds with a reference binding profile. In some methods, the average spacing between molecules of a compound in an area of the array is less than 6 nm.

[0006] Optionally, the method further comprises identifying a component of the sample that binds to the different compounds. Optionally, the method further comprises detecting the identified component with a binding partner known to bind the component. Optionally, the identified component is detected with a plurality of different binding partners known to bind the component. Optionally, the binding partner is an antibody to the identified component. Optionally, the binding partner is a peptide known to bind the identified component. Optionally, the binding partner is one of the different compounds detected in step (b). Optionally, the binding partner is a synbody. Optionally, the binding partner is immobilized to a support. Optionally, the binding partner is immobilized to a support in an array. Optionally, the method further comprises forming a second array, the second array containing one or more of the different compounds in the array binding to identified component not all of the different compounds in the array. Optionally, the second array contains less than 5% of the different compounds in the array. Optionally, the method further comprises forming an array or other device comprising compounds determined to bind to the sample but not all of the different compounds in the array.

[0007] In some methods, the detecting step detects binding of the different compounds to an antibody or antibodies in the sample. In some methods, the detecting step detects binding of the different compounds to a biological entity displaying multiple copies of a protein from its outer surface. In some methods, the biological entity is cell displaying multiple copies of receptor from its outersurface. In some methods, the compound having strongest binding to the sample binds to the component of the sample to which it has strongest binding affinity with an avidity of 0.1 μM to 1 mM. In some methods, the compounds are peptides or small molecules. In some methods, the array has 500-50,000 peptides. In some methods, the peptides are 10-30 amino acid long. In some methods, the sequences of the peptides are randomly selected. In some methods, the different immobilized compounds are selected without regard to the sample and the array further comprises a plurality of compounds known to bind different proteins also occupying different areas of the array. In some methods, the plurality of compounds known to bind different proteins includes compounds known to bind

PREOMICS EXHIBIT 1002
Page 442 of 1861

at least 25%, 50 or 75% of different human proteins. In some methods, the different immobilized compounds including a plurality of compounds known to bind at least 25%, 50% or 75% of different human proteins and 500-50,000 random peptides. In some methods, the sequences of the peptides have less than 90% sequence identity to a known binding partner of the target. In some methods, the sequences of the peptides have less than 90% sequence identity to known proteins. In some methods, the average spacing between compounds in an area of the array is 2-4 nm. In some methods, the average spacing between compounds in an area of the array is 3 nm. In some methods, the contacting of the sample to the array is performed in the presence of a potential competitor of binding of the sample to the array. In some methods, the competitor is a known binding partner of a suspected component of the sample.

[0008] In some methods, the sample is a patient sample, and the competitor is a protein known to be associated with a disease affecting the patient. In some methods, the sample is a patient sample. In some methods, the sample contains a plurality of antibodies. In some methods, the patient is known or suspected to be suffering from a disease. In some methods, the patient is known to be at risk of a disease but is not showing symptoms of the disease. In some the disease is an autoimmune disease, an infectious disease, or a disease of the CNS. In some methods, the sample is a blood, urine, or CNS sample. In some methods, component(s) of the sample are labeled. In some methods, binding of the peptides to component(s) of the sample is detected using a secondary antibody. In some methods, the secondary antibody is an isotype-specific antibody. In some methods, binding of the peptides to component(s) of the sample is detected by spr or mass spectrometry.

[0009] Some methods further comprise affinity purifying a component of the sample using a peptide determined to bind to the sample. Some methods further comprise washing unbound component(s) of the sample from the array, and dissociating bound component(s) from the array. Some methods further comprise preparing an antibody library from the patient and using a peptide to which an antibody in the sample binds as an affinity reagent to screen the library. Some methods further comprise identifying a natural binding partner of the affinity purified antibody. Some methods further comprise comparing the sequence(s) of peptide(s) binding to the component(s) of the sample to a database of natural sequences to identify natural binding partner(s) of components of the sample. Some methods further comprise comparing a profile of different compounds binding to the sample with profiles of the different compounds associated with different diseases or different stages of a disease to diagnose a patient as having one of the diseases or stages of disease. Some methods further

PREOMICS EXHIBIT 1002
Page 443 of 1861

comprise comparing a profile of different compounds binding to the sample with a profile of the different compounds associated with lack of a disease to determine whether a disease is present. Some methods further comprise repeating the method for different samples from a plurality of patients with the same disease to develop a binding profile characteristic of the disease. Some methods further comprise repeating the method for different samples from a plurality of patients with different disease to develop a plurality of binding profiles characteristic of different diseases.

[0010] The invention further provides an array of immobilized different compounds occupying different areas of the array, wherein different molecules of the same compound within an area are spaced sufficiently proximate to one another for multivalent binding between at least two of the different molecules in the same area and a multivalent binding partner.

[0011] The invention further provides methods of analyzing a sample, comprising: contacting the sample with an array of immobilized different peptides occupying different areas of the array; detecting binding of the different peptides in the array to component(s) of the sample; and characterizing the sample from the relative binding of a plurality of peptides with apparent dissociation constants between 1 mM and 1 μM for the sample. Optionally, the sample is characterized from the relative binding of at least ten peptides with apparent dissociation constants between 1 mM and 1 μM.

[0012] The invention further provides a method of characterizing a plurality of different samples, comprising contacting the different samples with the same array or copies of the same array of immobilized different peptides occupying different areas of the array; and detecting different binding profiles of the different peptides to the different samples; wherein the samples are characterized from their respective binding patterns. Optionally, the plurality of different samples includes samples from patients with different disease symptoms. Optionally, the plurality of different samples includes samples of patients presenting with disease and lack of disease.

[0013] The invention further provides a method of analyzing a sample, comprising: contacting the sample with an array of immobilized different compounds occupying different areas of the array, wherein different molecules of the same compound within an area are spaced at an average distance of less than 4 nm apart in the same area; and detecting binding of the different compounds in the array to component(s) of the sample.

PREOMICS EXHIBIT 1002
Page 444 of 1861

[0014] The invention further provides a method of analyzing a sample, comprising (i) contacting a sample with a known binding partner of a component of the sample; and (ii) determining whether the binding partner binds to the sample compared with a control lacking the component; wherein the known binding partner is identified by a process comprising (a) contacting an initial sample with an array of immobilized different compounds occupying different areas of the array, wherein different molecules of the same compound within an area are spaced sufficiently proximate to one another for multivalent binding between at least two of the different molecules in the same area and a multivalent binding partner; and (b) detecting binding of the different compounds in the array to component(s) of the sample; (c) identifying a component of the sample that binds to the different compounds; and (d) identifying a known binding partner of the component for use in step (i).

[0015] Optionally, the component in step (i) is detected with a plurality of different binding partners known to bind the component. Optionally, the binding partner is an antibody to the component. Optionally, the binding partner is a peptide known to bind the component. Optionally, the binding partner is one of the different compounds detected in step (b). Optionally, the binding partner is the binding partner is immobilized to a support in an array in step (i).

[0016] The invention further provides a method of manufacturing a device for use in detecting a component of a sample comprising: (a) contacting the sample with an array of immobilized different compounds occupying different areas of the array, wherein different molecules of the same compound within an area are spaced sufficiently proximate to one another for multivalent binding between at least two of the different molecules in the same area and a multivalent binding partner; and (b) detecting binding of the different compounds in the array to component(s) of the sample; and .(c) forming a device including a known binding partner of a component to which binding of the different compounds is detected in step (b).

[0017] Some methods further comprise identifying the component of the sample that binds to the different compounds. Some methods further comprise forming the device including a plurality of different binding partners known to bind the component. Optionally, the binding partner is an antibody to the component. Optionally, the binding partner is a peptide known to bind the component. Optionally, the binding partner is one of the different compounds detected in step (b). Optionally, the binding partner is a synbody. Optionally, the binding partner is immobilized to a support. Optionally, the binding partner is immobilized to a support in an array. Optionally, step (c) comprises forming a second array, the second array

PREOMICS EXHIBIT 1002
Page 445 of 1861

containing one or more of the different compounds in the array binding t the second array contains less than 5% of the different compounds in the array.

[0018] The invention further provides a method of testing a vaccine, comprising contacting a blood sample of a subject immunized with a vaccine against a pathogenic microorganism with an array of immobilized different compounds occupying different areas of the array; detecting a pattern of binding of the sample to the different compounds in the array; and comparing the pattern of binding to the pattern of binding of one or more reference samples, wherein the reference samples are from subjects who have survived an infection with the virus, similarity of binding profile between the subject and reference samples providing an indication the vaccine is effective against the pathogenic microorganism. Optionally, the subjects have been immunized with a vaccine before exposure to the virus.

[0019] The invention further provides an array of immobilized different compounds occupying different areas of the array, wherein the different compounds include a plurality of compounds known to bind at least 25, 50 or 75% of the known human proteins and 500-1,000,000 or more random peptides, wherein different molecules of the same peptide within an area are spaced sufficiently proximate to one another for multivalent binding between at least two of the different molecules in the same area and a multivalent binding partner

[0020] The invention further provides a method of analyzing a sample, comprising contacting the sample with an array of immobilized different compounds occupying different areas of the arrays; detecting binding of the different compounds in the array to component(s) of the sample; and characterizing the sample from the relative binding of a plurality of compounds having binding strengths to the sample greater than but within three orders of magnitude of the mean plus three standard deviations of the binding strength of empty areas in the array. Optionally, the characterizing comprising comparing a binding profile of the sample including the plurality of compounds with a reference binding profile including the plurality of compounds. Optionally, the sample is characterized from the relative binding of at least 10 or 100 compounds having binding strengths to the sample greater than but within three orders of magnitude of the mean plus three standard deviations of the binding strength of empty areas in the array.

BRIEF DESCRIPTIONS OF THE FIGURES

[0021] Fig. 1shows binding of 800 peptides to various antibodies and the degree of separation of these antibodies.

PREOMICS EXHIBIT 1002
Page 446 of 1861

[0022] Fig. 2 shows the signal from peptides that bind to monoclonal antibodies in the presence or absence of competitor.

[0023] Fig. 3 shows the binding intensities to the 300 most informative peptides that distinguish mixtures of two different monoclonal antibodies with and without competitor. Different levels of white to black indicate strength of binding of peptide to antibody.

[0024] Fig. 4 shows the binding profiles of patients with either Valley Fever, Influenza, Influenza Vaccine recipients, of healthy volunteers indicating the distinct pattern of binding specific to a disease for the 100 informative peptides shown.

[0025] Fig. 5 shows the binding profile of the 88 most informative peptides that distinguish Asthma, Valley Fever (Cocci), Sarcoma, Breast Cancer, Glioma, Pancreatic Cancer, Pancreatitis, and healthy volunteers. The lower panel shows the principal components map for the same peptides and provides a view of the degree of separation between samples (in 2 dimensions).

[0026] Fig. 6 shows principal component analysis of patients having or at risk of pancreatic cancer in the context of other diseases (Breast Cancer, Sarcoma, Pancreatitis, Glioma).

[0027] Fig. 7 compares binding profiles of normal volunteers, Breast Cancer patients, and patients at risk of breast cancer who clinically have had a second, different primary tumor following an initial remission of Breast Cancer.

[0028] Fig. 8 shows a pattern for influenza in mice blocked using whole virus particles pre-adsorbed to antisera from infected mice. This same pattern was not blocked by an irrelevant virus.

[0029] Figs. 9A, B and C show that an antibody pulled-down using peptides from the immunosignaturing microarray can detect the influenza particles. Fig. 9A is that antibody detecting PR8 particles. Fig. 9B is a positive control showing where the antibody is detecting influenza particles. Fig. 9C is a negative control pull-down from the beads alone.

[0030] Fig. 10A shows a hierarchical clustering of 875 different individual samples and ~2000 different peptides indicating that patients with different diseases show a common pattern of binding per disease. Fig. 10B shows principal component maps for the same peptides and provides a view of the degree of separation between samples.

[0031] Fig. 11 shows an analyses of the number of different peptides binding to different numbers of antibodies.

PREOMICS EXHIBIT 1002
Page 447 of 1861

[0032] Fig. 12 compares the binding profile of several monoclonal and polyclonal antibodies to different peptides in an array.

DEFINITIONS

[0033] Specific binding refers to the binding of a compound to a target (e.g., a component of a sample) that is detectably higher in magnitude and distinguishable from non-specific binding occurring to at least one unrelated target. Specific binding can be the result of formation of bonds between particular functional groups or particular spatial fit (e.g., lock and key type) whereas nonspecific binding is usually the result of van der Waals forces. Specific binding does not however imply that a compound binds one and only one target. Thus, a compound can and often does show specific binding of different strengths to several different targets and only nonspecific binding to other targets. Preferably, different degrees of specific binding can be distinguished from one another as can specific binding from nonspecific binding. Specific binding often involves an apparent association constant of 10^3 M^{-1} or higher. Specific binding can additionally or alternatively be defined as a binding strength (e.g., fluorescence intensity) more than three standard deviations greater than background represented by the mean binding strength of empty control areas in an array (i.e., having no compound, where any binding is nonspecific binding to the support). Most informative compounds have binding strengths within 1000-fold of three standard deviations of the background level. The range of affinities or avidities of compounds showing specific binding to a monoclonal or other sample can vary by 1-4 and often 2.5-3.5 orders of magnitude. An apparent association constant includes avidity effects (sometimes also known as cooperative binding) if present (in other words, if a target shows multivalent binding to multiple molecules of the same compound the apparent association constant is a value reflecting the aggregate binding of the multiple molecules of the same compound to the target). The theoretical maximum of the avidity is the product of the multiple individual dissociation constants, but in practice the avidity is usually a value between the association constant of individual bonds and the theoretical maximum. When contacted with a random selection of monoclonal antibodies, a subset of informative compounds (e.g., 1-20 or 5-15%) have association constants in the range of 10^3 to 10^6 M^{-1} or 2×10^3 to 10^6 M^{-1} or 10^4-10^6 M^{-1} to at least one and sometimes several (e.g., at least 2, 5 or 10) different targets. A subset of all peptides or other compounds (e.g., at least 1%, at least 5% or 10%, 1-75%, 5-60%, 1-20% or 5-15% usually shows actual affinity constants of 10^3-10^6 M^{-1} to at least one and usually several targets (e.g., at least 2, 5 or 10). The same ranges of association constant apply to composite targets binding to the same compound in a complex sample. Of course different

PREOMICS EXHIBIT 1002
Page 448 of 1861

compounds in an array have different degrees of binding strength to components of a sample and some compounds can bind with higher or lower apparent association constants than these ranges.

[0034] Patients include humans, veterinary animals, such as cats, dogs, horses, farm animals, such as chickens, pigs, sheep, cattle and laboratory animals, such as rodents, e.g., mice and rats.

[0035] A binding profile of an array is a measure of the amount of component(s) of a sample bound to the different compounds of an array to a particular sample. The amount of component(s) bound reflects the amount of the components in the sample as well as the binding strength of components to the compounds. A binding profile can be represented for example as a matrix of binding strengths corresponding to the different compounds in an array. A binding profile typically includes binding strengths of a plurality of compounds (e.g., at least 2, 10, 50, 1000 or 1000 having association constants in a range of 1 mM to 1 μM to a sample or within a range of greater than but within a factor of 1000 of three standard deviations greater than the mean intensity of empty cells.

[0036] Binding strength can be measured by association constant, dissociation constant, dissociation rate, or association rate, or a composite measure of stickiness which may include one or more of these measures. The strength of a signal from a labeled component of a sample bound to immobilized compounds can provide a value for general stickiness. If a term used to define binding strength is referred to as "apparent" what is meant is a measured value without regard to multivalent bonding. For example, the measured value of an association constant under conditions of multivalent bonding includes a plurality of effects due to monovalent bonding among other factors. Unless otherwise specified binding strength can refer to any of these measures referred to above.

[0037] The term "nucleic acids" includes any and all forms of alternative nucleic acid containing modified bases, sugars, and backbones including peptide nucleic acids and aptamers, optionally, with stem loop structures.

[0038] The term "polypeptide" is used interchangeably with "peptide" and in its broadest sense to refer to a sequence of subunit natural amino acids, amino acid analogs including unnatural amino acids. Peptides include polymers of amino acids having the formula $H_2NCHRCOOH$ and/or analog amino acids having the formula $HRNCH_2COOH$. The subunits are linked by peptide bonds (i.e., amide bonds), except as noted. Often all subunits are connected by peptide bonds. The polypeptides may be naturally occurring, processed

PREOMICS EXHIBIT 1002
Page 449 of 1861

forms of naturally occurring polypeptides (such as by enzymatic digestion), chemically synthesized or recombinantly expressed. Preferably, the polypeptides are chemically synthesized using standard techniques. The polypeptides may comprise D-amino acids (which are resistant to L-amino acid-specific proteases), a combination of D- and L-amino acids, β amino acids, and various other "designer" amino acids (e.g., β-methyl amino acids, Cα-methyl amino acids, and Nα-methyl amino acids) to convey special properties. Synthetic amino acids include ornithine for lysine, and norleucine for leucine or isoleucine. Hundreds of different amino acid analogs are commercially available from e.g., PepTech Corp., MA. In general, unnatural amino acids have the same basic chemical structure as a naturally occurring amino acid, i.e., an a carbon that is bound to a hydrogen, a carboxyl group, an amino group, and an R group.

[0039] In addition, polypeptides can have non-peptide bonds, such as N-methylated bonds (-N(CH_3)-CO-), ester bonds (-C(R)H-C-O-O-C(R)-N-), ketomethylen bonds (-CO-CH_2-), aza bonds (-NH-N(R)-CO-), wherein R is any alkyl, e.g., methyl, carba bonds (-CH_2-NH-), hydroxyethylene bonds (-CH(OH)-CH_2-), thioamide bonds (-CS-NH-), olefinic double bonds (-CH=CH-), retro amide bonds (-NH-CO-), peptide derivatives (-N(R)-CH_2-CO-), wherein R is the "normal" side chain. These modifications can occur at any of the bonds along the peptide chain and even at several (2-3) at the same time. For example, a peptide can include an ester bond. A polypeptide can also incorporate a reduced peptide bond, i.e., R_1-CH_2-NH-R_2, where R_1 and R_2 are amino acid residues or sequences. A reduced peptide bond may be introduced as a dipeptide subunit. Such a polypeptide would be resistant to protease activity, and would possess an extended half-live *in vivo*. The compounds can also be peptoids (*N*-substituted glycines), in which the sidechains are appended to nitrogen atoms along the molecule's backbone, rather than to the α-carbons, as in amino acids.

[0040] The term "polysaccharide" means any polymer (homopolymer or heteropolymer) made of subunit monosaccharides, oligimers or modified monosaccharides. The linkages between sugars can include acetal linkages (glycosidic bonds), ester linkages (including phophodiester linkages), amide linkages, and ether linkages.

DETAILED DESCRIPTION

I. General

[0041] The invention provides arrays of compounds for use in profiling samples. The arrays include compounds binding to components of the samples at relatively low affinities. Although practice of the invention is not dependent on an understanding of mechanism, it is believed that under conditions of monovalent binding, different degrees of specific binding

PREOMICS EXHIBIT 1002
Page 450 of 1861

might be difficult to distinguish from each other and from nonspecific binding. However, the avidity of compounds binding to components of the samples can be increased by forming arrays such that components of the samples (e.g., antibodies or cells) can bind to more than one molecule of a compound at the same time. When a sample is applied to an array under such conditions, the compounds of the array bind to component(s) of the sample with significantly different affinities generating a profile characteristic of the sample. Such a profile usually includes some compounds having no specific binding to components of the sample and other compounds having different degrees of specific binding to components of the sample. Although such binding interactions are specific in the sense that overall binding profiles of an array are reproducible for replicates of the same sample and distinguishable between different samples, they are not necessarily unique in that compounds in the array usually show specific binding albeit of different degrees to a number of different components of a sample or different samples.

[0042] The avidity of informative compounds (i.e., those showing specific binding) in an array can be measured for monoclonal antibody samples. When measured against monoclonal antibodies selected from random (e.g., purchased from a commercial supplier as described in the Examples), informative compounds in some arrays often show apparent affinity association constants in a range of 10^4-10^9, 10^6-10^9, 10^4-10^7 10^4-10^6 M^{-1}. Association constants of such informative compounds are often within a range of 10^3-10^6 M^{-1} or 10^4-10^5 M^{-1}. When measured against a complex sample, similar ranges of apparent or actual association constants are observed; however, in this case, the constant are a composite of values for multiple different components within a sample binding to the same compound. Such affinities can be distinguished from nonspecific interactions. The proportion of informative compounds (i.e., compounds that show distinguishable binding among different targets) can vary depending on the composition of the array and the sample, but ranges of 1-75%, 5-60%, 1-20%, 5-15%, or 7-12% provide some guide Given the data in Example 1 showing that different monoclonal antibodies have their own signature, it might have seemed impossible to meaningfully resolve patient samples which may contain 10^8 or more different specificities of antibodies in the serum the array would be unresolveable. When an array is hybridized against a more complex sample, such as from a patient, the binding profile represents the aggregate effect of multiple components of a sample. Surprisingly despite the complexity of the samples, different samples are associated with different binding profiles. Also surprisingly, the intensity of binding profile often differs between patients with a disease or at risk of disease relative to normal patients. Relatively more compounds are informative

PREOMICS EXHIBIT 1002
Page 451 of 1861

for disease patients or patients at risk of disease relative to normals and binding intensities are relatively higher (e.g., biased toward the higher end of the range for a random selection of monoclonals) than for the normal patients (intensities biased toward the lower end of the range for a random selection of monoclonals).

[0043] The binding profile of such an array to a sample can be used to characterize a sample. For example, the binding profile can be compared with binding profiles known to be associated with different diseases or stages of diseases or lack of diseases. Alternatively or additionally, the binding an be analyzed , for example, by using a compound binding relatively strongly to a component of the sample to affinity purify an antibody from the sample, or by comparing the sequence of a peptide in the array known to bind strongly to a component of a sample with a protein database to identify a protein in the sample. Remarkably, the same array can generate different and informative profiles with many different samples representing different disease states, disease stages, lack of disease and the like. Moreover, a profile characteristic of disease or departure from a nondisease state can be detected very early in development of a disease before typical analytical markers of disease would be detectable by conventional methods, such as ELISA.

II. Compounds for Use in Arrays

[0044] Many different classes of compounds or combinations of classes of compounds can be used for the arrays and methods of the invention. Classes of compounds include nucleic acids and their analogs, polypeptides (broadly defined as above), polysaccharides, organic compounds, inorganic compounds, polymers, lipids, and combinations thereof. Combinatorial libraries can be produced for many types of compounds that can be synthesized in a step-by-step fashion. Such compounds include polypeptides, beta-turn mimetics, polysaccharides, phospholipids, hormones, prostaglandins, steroids, aromatic compounds, heterocyclic compounds, benzodiazepines, oligomeric N-substituted glycines and oligocarbamates. Large combinatorial libraries of the compounds can be constructed by the encoded synthetic libraries (ESL) method described in Affymax, WO 95/12608, Affymax, WO 93/06121, Columbia University, WO 94/08051, Pharmacopeia, WO 95/35503 and Scripps, WO 95/30642 (each of which is incorporated by reference for all purposes). Peptide libraries can also be generated by phage display methods. See, e.g., Devlin, WO 91/18980. The test compounds can be natural or synthetic. The test compounds can comprise or consist of linear or branched heteropolymeric compounds based on any of a number of linkages or combinations of linkages (e.g., amide, ester, ether, thiol, radical additions, and metal coordination), dendritic structures, circular structures, cavity structures

PREOMICS EXHIBIT 1002
Page 452 of 1861

or other structures with multiple nearby sites of attachment that serve as scaffolds upon which specific additions are made. The compounds can be naturally occurring or nonnaturally occurring. Many different classes of compounds other than nucleic acids can be used, but optionally if the compounds are nucleic acids, the sample components detected are not nucleic acids. In some arrays, the test compounds have a molecular weight of between 500 and 10,000 Da, and optionally 1000 to 4000 Da.

[0045] The number of compounds is a balance between two factors. The more compounds, the more likely an array to include members having detectable affinity for any target of interest. However, a larger number of compounds also increases the cost of synthesizing and analyzing an array. Arrays typically have at least 100 compounds. Arrays having between 500 and 25000 compounds provide a good compromise between likelihood of obtaining compounds with detectable binding to any target of interest and ease of synthesis and analysis. Arrays having, for example, 100 to 50,000 members or 500-500,000, or 1000-25,000 members can also be used. However, arrays having much larger numbers of members for example, 10^2-10^7 or 1000 to 5,000,000 or 500,000 to 2,000,000 can also be used.. Such arrays typically represent only a very small proportion of total structural space, for example less than 10^{-6}, 10^{-10}, or 10^{-15} in the case of peptides. Sequence space means the total number of permutations of sequence of a given set of monomers. For example, for the set of 20 natural amino acids there are 20^n permutations, where n is the length of a peptide. Although it is widely assumed that most if not all of the residues in a peptide epitope participate in binding to the a target, it is much more likely that between two and five residues in a 10-12mer epitope are involved in energetically favorable interactions with the target, the other residues are simply there to adjust the positions of the important residues, and to prevent inhibition of binding. Therefore, a relatively small number of peptides can provide a good representation of total sequence space, and include members capable of specific albeit low affinity interactions with a wide variety of targets. For example, 500-25,000 random peptides can cover evenly the entire shape space of an immune system (10^7 to 10^8 antibodies in humans).

[0046] More compounds in the array should allow higher resolution of the diversity of compounds in the complex sample. For example, an array of 1 million compound would begin to approach the complexity of antibodies in a person's serum and therefore should allow more resolution of complex samples. Yet, even with a much smaller number of compounds, one is able to resolve new immune responses from infection or immunization.

PREOMICS EXHIBIT 1002
Page 453 of 1861

[0047] For polymeric compounds, the lengths of polymers represent a compromise between binding affinity and ease of synthesis. There is some relationship between peptide length and binding affinity with increasing length increasing affinity. However, as peptide length increases the likelihood of binding a binding site on a target that interacts with the full peptide length decreases. Cost of synthesis also increases with increasing length as does the likelihood of insolubility. For peptide arrays, peptides having 8-35, 12-35, 15-25 or 9-20 residues are preferred. These ranges of monomer lengths can also be used for other polymers, although aptamers usually have longer lengths (e.g., up to 100 nucleotides).

[0048] The compounds (e.g., all or at least 80, 90 or 95%) are typically chosen without regard to the identity of a particular target or natural ligand(s) to the target. In other words, the composition of an array is typically not chosen because of a priori knowledge that particular compounds bind to a particular target or have significant sequence identity either with the target or known ligands thereto. A sequence identity between a peptide and a natural sequence (e.g., a target or ligand) is considered significant if at least 30% of the residues in the peptide are identical to corresponding residues in the natural sequence when maximally aligned as measured using a BLAST or BLAST 2.0 sequence comparison algorithm with default parameters described below, or by manual alignment and visual inspection (*see, e.g.*, NCBI web site ncbi.nlm.nih.gov/BLAST or the like).

[0049] Some compounds are randomly selected from total sequence space or a portion thereof (e.g., peptides in which certain amino acids are absent or under-represented). Random selection can be completely random in which case any compound has an equal chance of being selected from sequence space or partially random in which case the selection involves random choices but is biased toward or against certain monomers, such as amino acids. Random selection of peptides can be made for example by a pseudorandom computer algorithm. The randomization process can be designed such that different amino acids are equally represented in the resulting peptides, or occur in proportions representing those in nature, or in any desired proportions. Often cysteine residues are omitted from library members with the possible exception of a terminal amino acid, which provides a point of attachment to a support. In some libraries, certain amino acids are held constant in all peptides. For example, in some libraries, the three C-terminal amino acids are glycine, serine and cysteine with cysteine being the final amino acid at the C-terminus. A library chosen by random selection, once selected is of known identity and can be reproduced without repeating the initial random selection process. Nevertheless, the compounds in such library retain the same random relations with one another. For example, the peptides in random library that is

PREOMICS EXHIBIT 1002
Page 454 of 1861

subsequently reproduced retain a random distribution throughout sequence space (with the possible exception of cysteine residues, if this residue is omitted). Collections of compounds, such as peptides, that are randomly distributed over sequence space, are still considered random even if reproduced without repeating the initial random selection.

[0050] The principles for selecting peptides and other compounds for arrays in the present methods are analogous to those for selecting initial libraries of compounds in producing synthetic antibodies, as further described in WO/2008/048970 and WO2009/140039.

III. Making Arrays

[0051] Compounds can be presynthesized and spotted onto a surface of an array or can be synthesized in situ on an array surface (see, e.g., Cretich et al., Biomol. Eng. 2, 77-88 (2006); Min et al., Current Opinion in Chemical Biology 8, 554-558 (2004), Breitling, Mol. BioSyst., 5, 224-234 (2009) ,US 5,143,854; EP 476, 014, Fodor et al., 1993, Nature 364, 555-556; US 5,571,639, US 5,593,839, EP 624,059, US 6,620,584, EP 728,520). Customized arrays are also commercially available from suppliers such as Invitrogen or Pepscan. The surface is usually derivatized with a functional group that attaches to the compounds, optionally via linker. Compounds can be attached via covalent or noncovalent linkages. The array surface can be a single contiguous surface of a support. Alternatively an array can be formed by spotting or synthesizing different compounds on different particulate supports, such as beads. Peptides can be attached in either orientation (N or C) relative to the array. In general, the different compounds occupy different areas of a contiguous array or different particles in a particulate array. The identity of which compound occupies which area of an array or which particle is usually either known as a result of the synthesis process or determinable as a result of an encoding process. Encoding processes are commonly used for beads. The different areas in a contiguous array can be immediately adjoining as may arise when such arrays are the result of in situ synthesis, or separated, which is often the result of spotting.

[0052] An area or cell of an array is a unit of surface area from which a separate signal is detectable. In some arrays, each area of the array is occupied only by molecules of the same compound except for possibly a small degree of bleed over from one area to another, due for example, to imperfections in the array. In other arrays, some or all of the areas contain a pool of two or more different compounds. In such an array, the signal from an area containing a pool of two or more different compounds is the aggregate undivided signal from the compounds constituting the pool.

PREOMICS EXHIBIT 1002
Page 455 of 1861

[0053] Such arrays typically contain from 100-5,000,000 compounds (e.g., 100-1,000,000, 500, 100,000 or 500-25,000 compounds) as discussed above. These numbers of compounds can readily be accommodated in different regions of an array of the order of 1-5 cm^2 combined area.

[0054] Within any one area of a contiguous array or within anyone particle of a particle array many different molecules of the same compound are present. Because compounds are usually attached to a derivatized surface of a support or particle (e.g., a support or particle bearing a linker), the density of molecules within an area of an array or a particle can be controlled in part by the derivatization process, for example, the period of time and concentration of derivatizing agent used. The density of molecules can also be controlled by the attachment or in situ synthesis process by which a compound is attached to a support. The length of a coupling cycle and concentration of compound used in coupling can both affect compound density.

[0055] The density of different molecules of a compound within an area of an array or on a particle controls the average spacing between molecules of a compound (or compounds in the case of a pooled array, which in turn determines whether a compound is able to form multivalent bonds with a multivalent binding partner in a sample. If two molecules of a compound or compounds in the case of a pooled array, are sufficiently proximate to one another, both molecules can bind to the same multivalent binding partner (for example to the two arms of an antibody). For peptides of length 15-25 residues an average (mean) spacing of less than 0.1-6 nm, 1-4 nm, 2-4nm, e.g., 1, 2 or 3 nm is, for example, suitable to allow different regions of the same compound to undergo such multivalent bonding. Average (e.g., mean) spacings are typically less than 6 nm because spacings of 6 nm or more are do not allow simultaneous binding of two sites on the same target. For example, for peptides of lengths 15-25 residues, the two identical binding sites of one antibody could not span more than 6 nm to contact two peptides at once. The optimum spacing for multivalent interactions may vary depending on the compounds used and the components of the sample being analyzed.

[0056] The formation of multivalent bonds can be shown by several methods. For example, the binding of an array to an interact antibody (i.e., two binding sites) can be compared with an otherwise identical antibody fragment (e.g., a Fab fragment) having only one binding site. Stronger binding to the intact antibody than the antibody fragment (e.g., higher apparent association constant) indicates multivalent binding. Multivalent binding can also be shown by comparing the binding of an array of an immobilized compound to an intact

PREOMICS EXHIBIT 1002
Page 456 of 1861

antibody with two binding sites with the reverse format in which the antibody is immobilized and the compound is in solution. Stronger binding (e.g., higher apparent association constant) of the immobilized compound to the antibody in solution compared with immobilized antibody to the compound in solution provides an indication that the immobilized compound can form multivalent bonds to the antibody. If capacity of compounds in array to form multivalent bonds in such a procedure is tested, it is usually sufficient to test one or a few sample compounds from the array for such binding. If the compounds in the array are of similar type, e.g., peptides of the same length, and deposited or synthesized under the same conditions, it can be inferred that if one or a few compounds on an array (e.g., 1-10%) are capable of multivalent binding, then so are the others. It is also not necessary to test every array that is made. Association (i.e., affinity) constants of compounds can be measured by conventional methods using technologies like SPR, ELISA, Luminex and other solution-phase binding (e.g., monitoring changes in bound signal over time) when the antibody or other sample is immobilized and the compound is in solution. Conversely, apparent association constants can be measured when a compound is immobilized and antibody or other sample is in solution. Once suitable synthesis or deposit conditions have been established for achieving arrays capable of multivalent binding other arrays can by made under the same conditions without individualized testing.

[0057] Usually, different compounds are deposited or synthesized in different areas of an array under the same conditions, so that if one compound is spaced so that it is capable of multivalent binding, most or all compounds are. In some arrays, at least 10%, 50%, 75%, 90% or 100% of compounds in the array are spaced so as to permit multivalent binding with a multivalent binding partner. However, it is not necessary that all compounds be deposited or synthesized with the same spacing of molecules within an area of the array. For example, in some arrays, some compounds are spaced further apart so as not to permit or permit only reduced multivalent binding compared with other compounds in an array.

[0058] The spacing can be measured experimentally under given conditions of deposition by depositing fluorescently labeled compounds and counting photons emitted from an area of an array. The number of photons can be related to the number of molecules of fluorescein in such an area and in turn the number of molecules of compound bearing the label (see, e.g., US 5,143,854). Alternatively, the spacing can be determined by calculation taking into account the number of molecules deposited within an area of an array, coupling efficiency and maximum density of functional groups, if any, to which compounds are being attached. The spacing can also be determined by electron microscopy of an array.

PREOMICS EXHIBIT 1002
Page 457 of 1861

[0059] Arrays having larger spacing that do not permit multivalent interactions or do so to a reduced extent compared with spacing described above also have application in identifying high affinity interactions. This type of strategy can be used to identify peptides or other compounds, for example, that are very close structurally to the original epitope that raised the antibody response. Alternatively, for arrays of peptides from life space, this spacing facilitates identifying the true epitope.

[0060] The spacing between compounds can also be controlled using spaced arrays; that is, arrays on surfaces coated with nano-structures that result in more uniform spacing between compounds in an array. For example, NSB Postech amine slides coated with trillions of NanoCone apexes functionalized with primary amino groups spaced at 3-4 nm for a density of 0.05-0.06 per nm^2 can be used.

[0061] Array formats that can be used include microarrays, beads, columns, dipsticks optical fibers, nitrocellulose, nylon, glass, quartz, mica, diazotized membranes (paper or nylon), silicones, polyformaldehyde, cellulose, cellulose acetate, paper, ceramics, metals, metalloids, semiconductive materials, quantum dots, coated beads, other chromatographic materials, magnetic particles; plastics and other organic polymers such as polyethylene, polypropylene, and polystyrene; conducting polymers such as polypyrole and polyindole; micro or nanostructured surfaces, nanotube, nanowire, or nanoparticulate decorated surfaces; or porous surfaces or gels such as methacrylates, acrylamides, sugar polymers, cellulose, silicates, and other fibrous or stranded polymers.

[0062] An exemplary method of array preparation is as follows. A microarray is prepared by robotically spotting distinct polypeptides on a glass slide having an aminosilane functionalized surface. Each polypeptide has a C-terminal glycine-serine-cysteine as the three C-terminal residues and the remaining (17) residues determined by a pseudorandom computational process in which each of the 20 naturally occurring amino acids except cysteine had an equal probability of being chosen at each position. Polypeptides are conjugated to the aminosilane surface by thiol attachment of the C-terminal cysteine of the polypeptide to a maleimide (sulfo-SMCC, sulfosuccinimidyl 4-[*N*-maleimidomethyl]cyclohexane-1-carboxylate which is covalently bonded to the aminosilane surface. The polypeptides are chemically synthesized, dissolved in dimethyl formamide at a concentration that may range from about 0.1 mg/ml to about 2 mg/ml, and then diluted 4:1 with phosphate-buffered saline prior to spotting. The concentration of peptide or other compound determines the average spacing between peptide molecules within a region of the array. A concentration of 1 mg/ml gives an average spacing of about 0.5 nm. The spacing

PREOMICS EXHIBIT 1002
Page 458 of 1861

decreases non-linearly with dilution at lower concentrations. The printed slides stored under an argon atmosphere at 4°C until use.

[0063] An exemplary calculation of spacing is as follows: spot size: 150 μm, spot area: 17671 μm^2, nanoprint deposition volume: 200pL, peptide concentration: 1mg/ml, deposition amount: 200pg, # peptides deposited: $8x10^{10}$ per spot, $8x10^{10}$ peptides/17671 $\mu m^2 = 4.5x10^6$ peptides/μm^2, $2.2x10^{-7} um^2$ area needed by 1 peptide ($4.6x10^{-4}$ μm spacing).

[0064] As well as including compounds randomly or without regard to the sample being analyzed, arrays can include other compounds known to bind particular targets, such as proteins, in a sample. These compounds can be antibodies, synbodies or peptides among others. Usually, such interactions are high affinity (e.g., greater than 10^7, 10^8 or 10^9 M^{-1}). The number of such known binding partner compounds can be large, for example, there can be a different compound for at least 25, 50, 75, or 90% or substantially all of the known proteins expressed by a given genome, such as the human genome). The different known binding partner compounds occupy different areas of the array in similar fashion to randomly selected compounds. However, because the known binding partner compounds are in general capable of high affinity interactions, they can be used with or without an intermolecular spacing that permits multivalent interactions with the sample. Although one might think that inclusion of compounds selected at random or without regard to the sample being analyzed would be redundant in view of inclusion of known binding proteins to a large part or all of the encoded proteins in a genome, such is not the case because some diagnostic immune responses are the result of somatic mutation or non-protein components and not detected by binding proteins to encoded proteins.

IV. Samples and Components to be Analyzed

[0065] The arrays and methods of the invention can be used for analyzing any kind of sample containing or potentially containing analyte(s) of interest. Of particular interest are samples from human or veterinary patients or laboratory model animals. Such samples can be blood (including whole blood, red cells, plasma and the like), urine, feces, saliva, CNS fluid, other body fluids, hair, skin, biopsies and the like. A profile can be obtained from a small volume of sample, e.g., ≤ 1 μl. Some samples are from patients known or suspected to be suffering from a disease. The identity of the disease may or may not be known. Some samples are obtained from patients known to have been subjected to a risk of disease but in which symptoms of disease are not yet evident. The risk can be genetic (e.g., a particular gene or family history) or experiential (e.g., exposure to a toxic chemical or radiation).

PREOMICS EXHIBIT 1002
Page 459 of 1861

Samples can also be obtained from patients who have been vaccinated to analyze the resulting immune response.

[0066] Samples from patients can include a wide variety of components subject to potential analysis by an array. The components most amenable to detection are those capable of multivalent bonding to compounds in the array. Such components include antibodies, which can support multivalent bonding through their pairs of heavy and light chains (i.e., two binding sites per antibody) and cells, which can form multiple bonds through multiple copies of receptors displayed from their outer surfaces. Viruses can also form multivalent bonds through different copies of coat proteins on their outer surface. Samples from patients can include many different antibodies and/or different cells and/or other components.

[0067] Samples can be analyzed with little if any further processing or can be subject to further processing such that only selected components of the sample (e.g., antibodies or cells) are analyzed with the array.

V. Methods of Detection

[0068] Binding interactions between components of a sample and an array can be detected in a variety of formats. In some formats, components of the samples are labeled. The label can be a radioisotype or dye among others. The label can be supplied either by administering the label to a patient before obtaining a sample or by linking the label to the sample or selective component(s) thereof.

[0069] Binding interactions can also be detected using a secondary detection reagent, such as an antibody. For example, binding of antibodies in a sample to an array can be detected using a secondary antibody specific for the isotype of an antibody (e.g., IgG (including any of the subtypes, such as IgG1, IgG2, IgG3 and IgG4), IgA, IgM). The secondary antibody is usually labeled and can bind to all antibodies in the sample being analyzed of a particular isotype. Different secondary antibodies can be used having different isotype specificities. Although there is often substantial overlap in compounds bound by antibodies of different isotypes in the same sample, there are also differences in profile.

[0070] Binding interactions can also be detected using label-free methods, such as surface plasmon resonance (SPR) and mass spectrometry. SPR can provide a measure of dissociation constants, and dissociation rates. The A-100 Biocore/GE instrument, for example, is suitable for this type of analysis. FLEXchips can be used to analyze up to 400 binding reactions on the same support.

PREOMICS EXHIBIT 1002
Page 460 of 1861

[0071] Optionally, binding interactions between component(s) of a sample and the array can be detected in a competition format. A difference in the binding profile of an array to a sample in the presence versus absence of a competitive inhibitor of binding can be useful in characterizing the sample. The competitive inhibitor can be for example, a known protein associated with a disease condition, such as pathogen or antibody to a pathogen. A reduction in binding of member(s) of the array to a sample in the presence of such a competitor provides an indication that the pathogen is present.

[0072] The stringency can be adjusted by varying the salts, ionic strength, organic solvent content and temperature at which library members are contacted with the target.

VI. Applications

[0073] The arrays have a wide variety of applications in analyzing or characterizing clinical, veterinary, forensic, laboratory and other samples. As with conventional diagnostics, the arrays can be used to identify particular analytes within samples, for example, analytes associated with particular disease. However, the methods can also be used to provide a binding profile of different compounds characterizing a sample. The binding profile represents the aggregate interactions of the compounds with different components in the sample, and can be characteristic of a particular disease, stage of disease or lack of disease. The different components can be complex (e.g., at least 10, 100, 1000 or 1,000,000 different antibodies and/or different cells).

[0074] A binding profile typically includes compounds whose interactions with the sample are nonspecific as well as compounds whose interaction with the sample reflect specific but low affinity interactions (i.e., apparent or actual dissociation constant between 1 mM and 1 μM). Compounds with higher affinity interactions (i.e., dissociation constant less than 1 μM) may or may not be present. Such higher affinity interactions if present may arise by chance as a result of a compound in the array being a mimetic of a natural binding partner of a sample component or as a result of including a control in which a compound is a known binding partner of a component of a sample. However, a sample can usually be adequately characterized by the binding profile of compounds with low affinity interactions with the sample, optionally in combination with compounds lacking specific binding to components of the sample. For example, the identity and relative binding of at least 2, 5, 10 or 50 compounds capable of low affinity specific binding to components of the sample can often be used to characterize the sample. Such low affinities actions may in part be the result of compounds serving as mimetopes providing a linear epitope that (imperfectly) resemble an epitope against which an antibody in the same was raised (e.g., a complex 3D-structure).

PREOMICS EXHIBIT 1002
Page 461 of 1861

[0075] One application lies in analyzing samples from patients known or suspected to be suffering from a disease but in which the particular disease affecting the patient is not known. A conventional approach would be to perform separate assays for suspected diseases. By contrast, in the present methods, a single binding profile from the patient sample can be used to characterize the patient for many diseases, stage of disease or lack of disease. The binding profile can be used to characterize the sample for virtually any disease, including autoimmune disease, cancer, infectious diseases, and diseases of the CNS. Most if not all diseases involve some change s in antibodies, cells or other components present in patient samples, reflected in a binding profile. Some exemplary infectious diseases include bacterial, fungal and viral diseases, such as Valley Fever, Q-fever, Tularemia tularensis, Rickettsia rickettsii, HSV types I and II, HVB, HVC, CMV, Epstein Barr virus, JC virus, influenza, A, B or C, adenovirus, and HIV. Because different infections give different profiles, different infections in a patient having multiple infections can be detected simultaneously. Some exemplary cancers that can be diagnosed or prognosed using the methods of the invention include glioblastoma, breast cancer, multiple independent primary cancer and/or recurrence situation, pancreatic cancer, lung cancer, myeloma, ovarian cancer and esophageal cancer. Precancerous cells that are morphological distinguishable from normal cells but not yet cancerous can also be detected using the methods of the invention. Neurological diseases, such Alzheimer's disease, although not generally considered to be an autoimmune disease, results in some changes in antibodies present in a sample. The same is the case for chronic diseases, such as Asthma, Rheumatoid arthritis, Diabetes mellitus type 1, Psoriasis, Multiple Sclerosis and others.

[0076] Another application lies in analyzing samples from patients known or suspected to have a particular disease, but in which the stage, severity or prognosis for the disease is unclear. Again the binding profile can provide an indication of any of these factors.

[0077] Another application lies in analyzing samples from vaccinated patients to determine whether an adequate protective immune response is developing. The pattern of response in one patient can be compared, for example, with a patient who has been naturally infected with the pathogen and survived, a similarity of response pattern indicating the patient is likely to survive and a dissimilarity that the patient will get worse or die at least in the absence of alternate treatment. Alternatively, a profile of a patient or animal model immunized with a new vaccine (for example in a clinical or preclinical trial) can be compared with profiles of patients or control animals immunized with an existing vaccine known to be effective. In a further variation, patients being recruited for a clinical trial of a vaccine can be prescreened

PREOMICS EXHIBIT 1002
Page 462 of 1861

for binding profile. Those already having a binding profile similar to that of a patient immunized with a vaccine known to be effective or from a patient who has survived a natural infection can be eliminated from the trial because their inclusion might lead to a misleading placebo response.

[0078] Another application lies in screening samples from patients who have undergone organ transplant (particularly allotransplantation). The profile in a patient under test can be compared with profiles of patients undergoing organ transplant who have or have not undergone rejection following the transplant. Similarity of the profile between a patient under test and a patient who has previously undergone rejection (or an average profile of a collection of such patients) indicates that the patient is at risk or is undergoing rejection.

[0079] Another application lies in analyzing samples from a patient known to be at risk of a disease but in which symptoms of disease are not yet present. The risk can be genetic, such as a genetic mutation associated with disease or family history of the disease, or arise as a result of experience, for example, exposure to a toxic chemical, radiation, traumatic accident, stress, fatigue, chemotherapy, unprotected sex, or exposure to a subject with a contagious disease. Such a patient is naturally concerned about the possibility of acquiring a disease and early therapeutic intervention. The methods are particularly useful in crisis situations in which many subjects have had potential exposure to a risk. Conventional diagnostic assays often have a significant lag period before a disease can be developed. For example, conventional viral assays can take several months to develop detectable patient antibodies. Autoimmune diseases (e.g., lupus, type 1 diabetes, rheumatoid arthritis, multiple sclerosis) can take several years to develop specific autoantibody or T-cell responses to specific autoantigens. By contrast, the present methods can detect changes in a profile within a few days (e.g., less than 10, 5 or 3 days) of exposure to a risk, or infection. The changes in binding profile may reflect subtle changes in concentrations of many different components of a sample, few if any of which would be individually detectable. However, in the aggregate, the changes in binding profile of the compounds in the array indicate a change if the risk has started development of disease.

[0080] Another application lies in forensic analysis of a sample, for example, a sample recovered from a crime scene or a sample relevant to a paternity analysis. Comparison of a test sample with one or more references samples of known origin can provide an indication of the source of the test sample.

PREOMICS EXHIBIT 1002
Page 463 of 1861

[0081] Binding profiles can be used in a variety of ways in characterizing a sample. In some methods, a binding profile of a sample is compared with one or more reference binding profiles of the same compounds. A reference binding profile is a profile that characterizes a particular disease, stage of disease or lack of disease, and the like. Reference profiles are typically determined by averaging binding profiles of several samples (e.g., at least 2, 20, 50 or 100) each characterized for the same disease, stage of disease or lack of disease.

[0082] Comparison of a sample binding profile with a reference binding profile can involve comparing the different binding strengths of different compounds in an array to the respective samples to derive a value representing the overall similarity of the profiles. A measure of similarity on a scale of similarity is by implication an inverse measure of disimilarity and vice versa. Thus, a value representing the overall similarity includes a value representing the overall disimiliarity. However, mathematically disimiliarity matrices can be handled and analyzed distinctly from similarity matrices. Raw data from the sample being analyzed can of course be normalized before the comparison to eliminate any differences due to sample size, processing, concentration and the like, rather than relative representation of sample components. Standard ANOVA analyses can also block such nuisance factors, provided such factors are accounted for in the experimental design.

[0083] Various techniques can be used to derive a value based upon the comparison of a binding profile and a reference binding profile. A derived value can be used to measure the dissimilarity between the binding profile and the reference profile and be evaluated using a distance measure such as the Euclidean Distance (ED) metric. The ED metric is typically used for measuring the distance between two vectors of "n" elements. According to one implementation, if x = (x1, x2, x3, . . ., xN) and y = (y1, y2, y3, . . ., yN) are two points in Euclidean N-space, then the Euclidean distance between x and j may be computed as:

D_{xj} = SquareRoot(Summation $((x_i - y_i)^2)$)

[0084] The ED metric thus not a correlation (0 to 1), but a measurement of dissimilarity.

[0085] In the context of comparing a binding profile (defined by its binding values for each point in N-dimensional space, where N is the number of experimental points (conditions)) with a reference binding profile, a ED metric can be determined regardless of the complexity, number of peptides, or number of patients. Each profile being compared may be seen as a pattern: setting an explicit series of points across time, across dilutions, across disease states, across symptoms, etc., and the comparison described here looks for data that reflects this defined series of points.

PREOMICS EXHIBIT 1002
Page 464 of 1861

[0086] To standardize the difference between binding profiles being compared, the calculated ED measurement may be normalized by dividing by the square root of the number of conditions as follows:

Distance = | a-b | /square root of N

This is distinct from the aforementioned distance calculation by normalizing for the total number of conditions. This prevents the distance calculation from expanding too far given large numbers of samples.

[0087] Accordingly, calculating the Euclidean distance between two data points involves computing the square root of the sum of the squares of the differences between corresponding values. Because the ED metric is a measure of dissimilarity, the distance (d) may be converted, when needed, to a similarity measure as 1/(1+d). Distance, similarity, and dissimilarity are interchangeable to a certain degree but each is a uniquely useful given the calculations being applied. As the distance gets larger, the similarity gets smaller. This renders the original data useful for looking at differences in a non-biased and geometrical way. The computation is scalable with increasing number of experiments. In fact, the complexity of the pattern is inherently diminished to the calculation because it is in the denominator and is a square root.

[0088] Other distance metrics that can be used include Euclidean Squared, Pearson Correlation, Pearson Squared, Spearman Confidence or Correlation, and other like techniques.

[0089] Binding profiles can also be used in various analytical methods to further characterize the sample. For example, a compound in the array showing relatively strong binding to the sample (compared with other compounds in the array) can be used to affinity purify a component of the sample. The component can then be further characterized (e.g., by sequencing or immunoreactivity). The identity of the compound may be characteristic of a disease state (e.g., a pathogen , autoantibody or tumor associated antigen). If the component is not already known to be characteristic of a disease state, it can be used as a new target for developing therapies or diagnostics against the disease state. For example, autoantigens or peptides thereof, can be used in inducing tolerance of autoimmune disease. Alternatively, after washing off unbound cellular components, the cellular components binding to an array can be dissociated from the array, fractionated and analyzed in similar fashion. In a further variation, the identity of a compound in the array showing relatively strong binding to a sample can be used to identify a ligand of the component bound in the sample, and hence the component in the sample. For example, if the compounds of the array are peptides, the

PREOMICS EXHIBIT 1002
Page 465 of 1861

sequence of a peptide showing relatively strong binding to a sample can be compared with a database of protein sequences. Comparison can be pairwise between a database sequence and a peptide in the array or between a database sequence and a motif or consensus sequence from a plurality of peptides in the array. Sequence similarity to a protein in the database provides an indication that the protein is a ligand of the component in the sample to which the peptide showed strong binding. The identity of a ligand in turn provides at least an indication of potential molecules in the sample and in turn disease states characterized by such molecules.

[0090] The same array can be used in any of the applications described above and for virtually any disease or suspected disease state. The same array means either literally the same array, in which case the array may be washed between different samples, or different copies of an array of the same composition. The identity of which compounds in the array are most informative for a disease or other state being analyzed varies by state. Thus, having identified the most informative compounds for a particular disease, derivative arrays or other detection devices and kits can be made that have a reduced number of compounds including the most informative compounds. The derivative arrays are sometimes referred to as secondary arrays to distinguish them from primary arrays used in initial identification of binding compounds and sometimes a sample component bound by these compounds.

A further useful aspect of the present methods is that they can detect not only increased binding of compounds to cellular components in test samples relative to a control sample representing an undiseased subject (typically a human) but can also detect decreases. For example, some sample components, particularly antibodies, can be detected to decrease in a test sample, such as a disease or vaccinated sample or any other of the samples types mentioned, and other sample components increase.

VI. Derivative Analyses

[0091] As well as being useful in themselves for analyses of samples as discussed above, the present methods are also useful for determining derivative compounds and detection devices. In a simple form of such methods, a derivative device or other array in constructed containing one or more compounds known to be associated with a given disease, susceptibility to disease or other condition described above, and omission of other compounds from the primary array not found to be informative for this disease, susceptibility or other condition. In some such methods, only a small proportion of the compounds used in a primary array (e.g., less than 0.1%, 1% or 5% are retained). In other methods, a component of the sample bound by some of the compounds in a primary array is identified by any of the

PREOMICS EXHIBIT 1002
Page 466 of 1861

approaches discussed in the previous section. Having identified a component of the sample, one or more known binding partners of the component are also identified. The known binding partners can be compounds from the primary array, antibodies to the component or other compound, such as a synbody that is known to bind to the component. The known binding partner(s) can then be used to detect the sample component to which they are known to being by any otherwise conventional diagnostic assay. For example, if the known binding partner is an antibody, the assay can be an ELISA, immunoprecipitation, radioimmunoassay or the like. If a plurality of known binding partners are used, the known binding partners can be immobilized in an array format. The known binding partners can also be incorporated into diagnostic kits or diagnostic device (e.g., attached to a support). Such arrays, diagnostic devices and kits can be manufactured by conventional means. Of course, once the known binding partners of a component have been identified, it is not necessary to repeat the initial screening with the primary array for subsequent manufacture of such arrays, diagnostic devices and kits.

[0092] Although the invention has been described with reference to the presently preferred embodiments, various modifications can be made without departing from the invention. Unless otherwise apparent from the context any step, element, embodiment, feature or aspect of the invention can be used with any other. All publications (including GenBank Accession numbers and the like), patents and patent applications cited are herein incorporated by reference in their entirety for all purposes to the same extent as if each individual publication, patent and patent application was specifically and individually indicated to be incorporated by reference in its entirety for all purposes. If more than one version of a sequence is associated with a deposit number at different times, the version associated with the deposit number at the effective time of filing the application is meant. The effective time of filing means the earliest application from which priority is claimed disclosing the relevant accession number.

EXAMPLES

Example 1 Array Binding to Monoclonals

[0093] Triplicate copies of an array having 10,000 different random peptides made by robotically spotting distinct polypeptides on a glass slide having an aminosilane functionalized surface as described above were tested for binding to individual antibodies were hybridized at 100nM, 37°C, 8rpm, 1 hour and detected with 5nM secondary antibody. The antibodies tested included; (1) monoclonals, (2) polyclonals, (3) unknown epitope, (4) monoclonals with linear epitopes, (5) monoclonals with discontinuous epitopes, (6) anti-Fc

PREOMICS EXHIBIT 1002
Page 467 of 1861

antibodies, (7) polyreactive antibodies, (8) autoantibodies, (9) mixed monclonals, and (10) antibodies to glycans. Triplicate arrays were mathematically averaged, and the most informative 800 peptides were used to distinguish the relative differences. The binding of each antibody or antibody mixtures to the 800 peptides can be represented by a series of colored bands as shown in the right of Fig. 1. Each band represents binding to a different one of the 800 informative peptides in the array. Different colors can be used to represent the different strengths of binding, for example, red being the highest, blue the lowest and yellow intermediate. The apparent binding affinities of most of the informative peptides to the antibodies tested range from about 10^4 M^{-1} to 10^6 M^{-1} or 100 (background, e.g., an empty cell of an array lacking a peptide) to 65,500 in intensity unity. The left portion of Fig. 1 shows principal component analysis of the patterns shown in the right hand proportion of a figure. Principal component analysis represents the binding profiles of different antibodies as spots on a two-dimensional chart, such that the relative distance between the spots is a measure of the relatedness of patterns. The principal component analysis shows that each of the antibodies has a distinguishable binding profile. The analysis also shows that technical replicates are very reproducible per array and per peptide. The distinct binding patterns of different monoclonal and polyclonal antibodies are also shown in Fig. 12.

[0094] Fig. 11 shows an analyses of the number of different peptides binding to different numbers of antibodies. For example 1103 peptide bound to one and only one of the antibodies tested. 402 peptides bound to two of the antibodies tested. Six peptides bound to ten different antibodies and so forth. Thus, the different peptides have different degrees of promiscuity in binding to multiple targets. Such a range of promiscuities can be of assistance in using an array to characterized multiple samples.

[0095] Of over twenty types of monoclonal antibodies, each produced a distinct pattern on the array. These differences included both what peptides bind and the relative binding to each. Signals on the array varied over 3 logs of fluorescence intensity. The results show antibodies against sugars, non-natural protein sequences (e.g. frameshift peptides, translocation junctions, splice variants), self-proteins, post-translational modifications (PTM's), multi-species, and multi-class (IgG, IgA, IgM, IgE) are all detected. Because each antibody has a distinct pattern of binding, one might expect that the 10^8 repertoire of serum antibodies would create a monotonic pattern indistinguishable between different samples. Indeed a mathematical reconstruction of the binding of 30 monoclonals indicates that the distinctive patterns start to be lost. Remarkably as shown in subsequent examples, this is not the case in reality.

PREOMICS EXHIBIT 1002
Page 468 of 1861

Example 2 Dilution

[0096] A single monoclonal antibody, p53Ab1, was used to test the ability to discern a signature or binding profile in the presence of competing IgG. As before, 100nM p53Ab1 was hybridized to the arrays to detect its signature. We added 1, 10, and 100-fold pooled human IgG as competitor and could still distinguish the original signal. For the 2-color experiments, IgG was detected with a red fluor, p53 with a green fluor and the competing spots were detected by multi-wavelength scanning. Fig. 2 shows the signal from the most informative peptides in the different conditions tested. The signal from the monoclonal stands out against the complex background of the IgG serum. This is a key and unexpected aspect of the immunosignaturing technique. High affinity immune responses, like monoclonals mixed in serum, stand out on the arrays from the normal antibodies.

[0097] Two antibodies were selected against p53, Ab1 and Ab8. Each was titrated into the other yielding a gradient of signatures. Fig. 3 shows the binding intensities to the most informative peptides in which different colors are used to represent different binding strengths. As seen the strongest pattern, Ab8, is seen in the presence of another antibody and in the presence of competing IgG. Eight antibodies mixed together show a trend towards monotonic response. 100X IgG shows exactly this. Dilution shows that pattern of p53Ab1 dominates a poorer antibody, p53Ab8 and excess IgG. Physically pooled antibodies and antibody signals mathematically averaged show similar patterns. Secondaries and tertiaries were not detectable.

Example 3: Resolving Disease

[0098] We collected normal donor sera and sera from patients infected with Valley Fever and Influenza vaccine recipients. Valley Fever patients were previously typed against CF antigen for titer. Fig. 4 shows the binding profiles of different patients for the most informative patients with different color bands representing different binding intensities. In general, normal undiseased patients, showed a low intensity pattern. Patients infected with cocci showed a much higher intensity pattern even when no titer of cocci could be detected.. Patients inoculated with flu vaccine showed varying patterns likely corresponding to the strength of the immune response. These patterns were readily distinguishable from the normal patients and cocci infected patients.

[0099] In a further experiment, samples from about 200 cancer patients having various kinds of cancer including glioma, sarcoma and pancreatic cancer were tested. Fig. 5 shows the binding profile of the most informative peptides in comparison with cocci infected

PREOMICS EXHIBIT 1002
Page 469 of 1861

patients and asthma patients. Principal component analysis confirms that the patterns of patients with these different diseases are distinct, and the patterns of cancer patients show clustering by the type of cancer.

[0100] In a further experiment, samples from patients at risk of pancreatic cancer and/or having pancreatitis were compared with patients having pancreatic cancer, patients having sarcoma or normal patients. Principal component analysis (Fig. 6) shows that some of the patients at risk of pancreatic cancer have a binding profile clustering with patients having cancer. However, patients having pancreatitis cluster separately from patients having pancreatic cancer. Thus, binding profiles can be used as a predictor of cancer development in at risk patients and to distinguish pancreatitis from cancer.

[0101] Fig. 7 compares binding profiles of normal patients, breast cancer patients, and patients at risk of breast cancer. The binding profiles distinguish breast cancer patients from normals, and classify pre-symptomatic cancer patients as having a breast cancer pattern years before detection of a tumor. The binding profiles of pre-symptomatic samples also differ from normals because of their lack of variability (a hallmark of normals).

Example 4

[0102] A sublethal influenza A/PR/8/34 infection of 1×10^4 viral particles was used to impart protection against a later lethal challenge of the same strain in BALB/c mice. Lethal challenge using 2–5 mean lethal doses of influenza A/PR/8/34 occurred on day 35. The resulting immune response was protective as no overt clinical signs of infection or weight loss were observed following challenge. Portions of this example have been reported by the inventors in Lugutki et al, Vaccine (May 5, 2010) PMID: 20450869.

[0103] Serum from live influenza immunized mice were run on the arrays. At day 14, the influenza-specific IgG ELISA titer was 1:102,400 indicating virus specific antibodies were present. Difference in antibody binding to the CIM10K array could be detected 14 days after immunization confirming that a change in humoral immune status could be detected. Comparison of the 271 peptides recognized at least two-fold greater by live influenza immunized mice did not overlap with the 542 peptides recognized at least two-fold more than age matched naïve mice immunized 29 days previously with F. tularensis LVS. Therefore an array of 10,000 random peptides are sufficient generate immunosignatures can distinguish between two infections.

[0104] Mice were infected with influenza and 21-day antisera was tested for an immunosignature pattern. We found a specific influenza pattern in mice that could be

PREOMICS EXHIBIT 1002
Page 470 of 1861

blocked using whole virus particles pre-adsorbed to antisera from infected mice but not by an irrelevant virus (M13) (Fig. 8). Comparison of the absorbed sera in an ELISA revealed a reduction in IgG reactivity for the virus from an endpoint titer of 512,000 for the unabsorbed 387 sera to less than 500 for the virus absorbed sera. Reactivity for the possible to determine antibody reactivities that declined on immunization in a standard ELISA.

[0105] To determine whether these peptide populations were the result of the primary immunization or a combined effect of immunization and challenge, the recognition of these peptide populations were compared in immunized only mice. At 118 days post-immunization, immunized only mice had a mean fold increase from day 0 of 5.18±3.57 for the 283 time stable peptides compared to immunized and challenged mice at 98 days post-challenge which had a mean fold increase of 4.19±2.55 from day 0. At the same time point, the suppressed peptides had a similar fold change 350 from day 0 of 0.39±0.29 in the immunized mice and 0.36±0.098 in immunized and challenged mice. Taken together these results demonstrate that a lasting and complex immunosignature is generated upon initial infection and it is unique to the infecting organism.

[0106] We sought to determine whether the immunosignature on the CIM10K was transient or remained consistent over time and correlated to the antibody titer in a virus specific ELISA. The immunized animals were followed for 211 days post-challenge and the changes to peptide recognition levels by serum antibodies on the CIM10K array observed. In a standard ELISA against whole virus, the IgG titer rose to 819,200 by the day of challenge and remained elevated and unchanged during the post-challenge observation period. An "expression pro-file" reflecting the changes in bulk IgG ELISA titer was established in the Gene Spring software package. Comparison of the relative fluorescence intensities of recognized peptides to the profile identified peptides whose intensities over time matched the defined profile with a Pearson correlation of greater than 0.9. Profile analysis was similarly used to identify sets of peptides whose recognition was increased after immunization but diminished after challenge and those that were initially recognized at day 0 but were no longer recognized following immunization.. At 118 days post-immunization, immunized only mice had a mean fold increase from day 0 of 5.18±3.57 for the 283 time stable peptides compared to immunized and challenged mice at 98 days post-challenge which had a mean fold increase of 4.19±2.55 from day 0. At the same time point, the suppressed peptides had a similar fold change from day 0 of 0.39±0.29 in the immunized mice and 0.36±0.098 in immunized and challenged mice. Taken together these results demonstrate that a lasting and

PREOMICS EXHIBIT 1002
Page 471 of 1861

complex immunosignature is generated upon initial infection and it is unique to the infecting organism.

[0107] We sought to determine whether the antibody immunosignature on the CIM10K correlated to the immunizing dose of virus. Groups of BALB/c mice had increased weight loss and recovery time which correlated to increased infectious dose. Mice were allowed to recover to pre-infection weight and serum was collected at day post-infection. Bulk IgG endpoint titers against the virus were indistinguishable between groups with an endpoint titer of 102,400 for all groups. This suggests that the maximum detectable response in an ELISA was quickly reached at the lowest immunizing dose. Computer analysis of the CIM10K immunosignatures revealed 516 peptides that generally increased with infectious dose, of which a subset of 65 rose sharply. Comparison of the time stable peptides to the dose responsive peptides showed an overlap of 39 peptides, three times that predicted by chance. These results demonstrate that immunosignature has greater resolving power than whole virus ELISA in terms of distinguishing the infectious dose.

[0108] Antibody responses to infection or vaccine can be measured by ELISA for each of the responding isotypes. We sought to determine whether or not the antibody reactivity on the random CIM10K peptide array could detect changes in antibody isotype. The IgM, IgA, IgG1, IgG2a and IgG3 immunosignatures were determined using serum samples from day 0 and day 28. Peptides recognized with relative fluorescence intensities greater than array features containing only buffer were analyzed. Populations of peptides recognized by the different isotypes are presented as a modified Venn diagram in Fig. 5. Peptide overlap sets containing more members than that predicted by chance overlap for equivalent sized lists were limited to certain overlap regions suggesting potential class switching from IgM and overlap with IgA indicating the switch to IgA that is expected in containing an airways infection. Fitting an immune response to a viral infection, the IgG subtype response was predominantly IgG2a where nearly half the array was increasingly over two-fold recognized by day 28 serum. Conversely, only 144 peptides were over two-fold recognized by IgG1 antibodies. This demonstrates that the immunosignature can be subdivided based on antibody isotypes which reflects the pathogenesis of the responsible agent. In contrast to the ELISA protocol, each isotype assay only required 0.5 μl serum.

[0109] We sought to determine whether or not the immunosignature was consistent between biological replicates. Additional groups of BALB/c mice were separately obtained approximately 2 months apart and independently infected with the sublethal dose used for immunization. Weight loss was consistent for all groups and at day 28 each group had a bulk

PREOMICS EXHIBIT 1002
Page 472 of 1861

IgG titer of 19,200. To address the consistency of the immunosignature across independent infections, day 28 sera from the three infections were allowed to bind the CIM10K array and fluorescence intensities compared. For all peptides on the array the Pearson correlation was 0.94 between infections. The time stable peptides had Pearson correlations of 0.904, 0.936 and 0.912 between infections. Compared to the bioinformatic average of all naïves tested, the three infections had similar fold increase values for the 283 time stable peptides (2.23±1.22, 2.88±1.0, and 1.93±0.97). These results demonstrate that the immunosignature is consistent across biological replicates of infection and across technical replicates. There is no loss of consistency between biological replicates when using random peptides in place of the authentic virus antigen.

[0110] Naïve inbred mice provide a relatively empty canvas on which to develop an antibody repertoire. We sought to determine if a consistent immunosignature could be distinguished in a diverse human population. The immunosignature to the human seasonal 2006–2007 influenza vaccine was evaluated as our model. Individual sera from seven human donors were allowed to bind the array and analyzed for differences in peptide recognition between pre-immune and day 21. Donors were determined to have serum IgG antibodies in an ELISA against the 2006–2007 seasonal vaccine. For the human donors the median day 0 IgG titer was 1:3200 and the median day 21 titer was 1:12,800. Analysis identified 30 peptides that significantly increased ($p < 0.01$) on day 21 at least 1.3 times that of the pre-immune or donor background values. These 30 peptides were sufficient to clearly distinguish the immune and pre-immune classes using principle component analysis. Four principle components were identified of 38%, 23%, 13% and 9% variance, respectively. One patient was separated across the X-axis for both immune and pre-immune samples and may suggest a low responding immune system. The remaining patient samples were still distinguished by class when this patient was excluded from the analysis. These results demonstrate that the immunosignature of serum antibodies are informative.

[0111] It would be of great advantage to know what antigen had produced a diagnostic antibody. We have developed a protocol that allows this "backtracking" from the array peptide to the natural protein. Basically the reactive peptide is resynthesized on Tentagel Beads. The beads are reacted with the serum to affinity purify the antibody. The antibody is then released from the beads. This antibody can then be used against an extract to identify the protein. Another method would be to make a phage library of antibodies from the individual producing the immune response and isolate the phage-ab binding the peptide of interest.

PREOMICS EXHIBIT 1002
Page 473 of 1861

[0112] An array peptide that showed strong reactivity against influenza but not normal mouse sera was selected to pull down antibodies against influenza specifically. The pull-down was done using 10ug peptide chemically conjugated to Tentagel® beads. Antibodies were eluted with pH 2.0 glycine, then immediately neutralized. PR8 was immobilized on nitrocellulose and probed with the pulled-down antibodies. Figs. 9A, B and C show the pulled-down antibody detecting PR8 particles, a positive control anti-PR8 antibody picking up the virus particles, and a negative control pull-down from the beads alone. The data indicate we pulled down the appropriate anti-PR8 antibody with the peptide from the microarray that showed the strongest reactivity to PR8-infected mice.

[0113] The immunosignature can be used to classify the recipient based on immunogen. Protection against influenza infection in a murine model is largely based on the generation of neutralizing antibodies. To test the ability of the immunosignature to distinguish a protective antibody response from a non-protective antibody response, the archived serum samples were used to probe the CIM10K array. A random number generator was used to assign samples to either set A or set B so that each set contained four A/PR/8/34 infected groups, five KLH immunized groups and seven naïve groups. This way either set could be used as the training set and selected peptides used to predict the members of the second set. Peptides were selected from the CIM10K using the association algorithm in Genespring 7.3. For training on set A, 515 peptides with a median p value of 2.32×10^{-7} were identified. TSupport vector machines were used to test the ability of the 515 to predict the immunogen used. When set A was used to train for prediction of set B, 81.3% of the samples were predicted correctly. When set B was used to train for prediction of set A using the 515 peptides, 93.8% of the samples were predicted correctly.

[0114] As a test of the validity of the methodology to select predictive peptides, association algorithm was applied to set B to identify 518 peptides with median p values of 6.34×10^{-5} for set B and 2.00×10^{-5} for set A. Support vector machines trained on set B to predict set A identified 81.3% of the samples correctly and training on set A to predict set B identified 75% of the samples correctly. The identified lists of associated peptides from set A and set B overlapped by 122 peptides which is greater than predicted by chance overlap. The overlap list predicted 81.3% of the samples correctly regardless of which set was the training set or the test set.

[0115] Immunization with homologous and heterologous vaccines produce different antibody responses and degrees of protection to challenge with Influenza A/PR/8/34.

PREOMICS EXHIBIT 1002
Page 474 of 1861

[0116] We next sought to determine whether the immunosignature had the power to distinguish similarly composed vaccines and stratify them based on outcome. As our model vaccines we used inactivated A/PR/8/34 as the homologous "good" vaccine and the 2006/2007 & 2007/2008 human seasonal influenza vaccines as the heterologous "bad" vaccines. Efficacy was determined in a murine lethal challenge model of influenza A/PR/8/34 infection. At day 40, serum IgG titers were measured. Mock immunized mice showed no reactivity for either A/PR/8/34 or the seasonal vaccines. Mice immunized with inactivated virus had IgG titers for all three antigens with a lower titer for the 2007/2008 seasonal vaccine components. The seasonal vaccines showed cross reactivity for each other with low reactivity for the A/PR/8/34 virus. Cross reactivity between seasonal vaccines is expected due to the inclusion of A/Wisconsin/67/2005 and B/Malaysia/2506/2004 in both vaccine compositions. The heamagglutanin and neuraminidase sequences from A/PR/8/34 do have some homology to the A/New Caledonia/20/99 and A/Solomon Islands/3/2006 sequences but appear to have generated a unidirectional cross reactivity in the ELISA.

[0117] Pooled serum samples were also used to probe the CIM10K. The association test was applied to the array data and 367 peptides associated with immunogen were identified. These peptides are sufficient to distinguish the groups with an average intergroup p value of $3.9x10^{-5}$. This demonstrates that the CIM10K has the resolving power to distinguish closely related immunogens from each other based on the humoral immune response. To account for the heterosubtypic crossreactive antibodies observed in the ELISA the three inactivated vaccines were compared to mock immunized. This comparison yielded 116 peptides with a $p<0.01$ of which 54 were increased in the vaccines over mocks. Relative intensity levels of the 54 peptides were comparable between groups. This list of peptides was tested in support vector machines for the ability to separate the A/PR/8/34 and the KLH immunized mice. For the combined sets A and B, no crossvalidation error was observed suggesting this set is likely composed of influenza specific cross reactive antibodies.

[0118] At challenge on day 42, mice immunized with inactivated A/PR/8/34 survived without clinical signs of illness. Mice immunized with the seasonal vaccines became sick and had 60% survival in the 2006/2007 seasonal vaccine group and 70% survival in the 2007/2008 seasonal vaccine group. Surviving mice in the seasonal vaccine groups recovered to prechallenge weight This demonstrates that the vaccines are not equally protective in challenge with A/PR/8/34. Taken together these data demonstrate that the different vaccines generate distinctive immune responses as determined by whole virus ELISA and CIM10K and have different protective outcomes.

PREOMICS EXHIBIT 1002
Page 475 of 1861

[0119] The immunosignature of live immunization and subsequent challenge can predict the efficacy of inactivated vaccines. To evaluate if the immunosignature of a naturally acquired infection and resistance to subsequent infection could predict the efficacy of an inactivated or subunit vaccine, we examined microarray data from mice infected with a sublethal dose of A/PR/8/34 and subsequently challenged with a lethal dose of the same strain 35 days later. We first asked if the peptides that were recognized by serum antibodies present in immune mice at the time of challenge were capable of predicting vaccine efficacy. 50 peptides significantly (twofold increase with a $p<0.01$) increased from day 0 to day 28. Support vector machines trained on the immune and naïve mice (day 28 and day 0) for testing on the model vaccine trail were incapable of distinguishing the groups, suggesting an amplification of protective clones was required for survival in a subsequent challenge.

[0120] To test this hypothesis, we asked if peptides increasing between pre and post challenge could predict the efficacy of the inactive vaccines. Analysis identified 163 peptides as increasing greater than twofold from pre to post challenge with a p value less than 0.05. Support vector machines trained on the immune and naïve samples (day 28 and 0) were able to predict the challenge outcome in terms of success defined by no change in health status (inactive A/PR/8/34) or unsuccessful defined by illness (seasonal vaccines and mock). This demonstrated that the ability of the immunosignature to identify the antibody reactivities responsible for protection to a second exposure can predict the efficacy of a subunit or inactivated vaccine.

Example 5 Multi-disease classifier

[0121] About 875 individual samples from individuals with infectious disease, autoimmune disease, cancer, Alzheimer's and other diseases were analyzed independently on a 10,000 peptide array. The binding profiles were initially represented as a heat map as for other profiles (Fig. 10A). Principal component analysis of the binding profiles for different diseases is shown in Fig. 10B. Each of the diseases has a distinguishable profile. Although inflammation is a major component in many of the diseases, it is not a major contributor to their binding profiles, which remain predominantly distinct with a classifier error below 10%

Example 6

[0122] The following table summarizes the results from several tests on clinical samples. FP(false positive, FN (false negative), AUC (area under a ROC curve, a measure of accuracy of diagnosis).

PREOMICS EXHIBIT 1002
Page 476 of 1861

Class	Disease	FP	FN	AUC	# disease samples
Pre-disease (presymptomatic)	panIN (pre-panc. Cancer)	0	0	1	8 human
Vaccine and challenge	Influenza	0 0.001	0 0	1 0.999	15 human, 4 years 120 mice
	Pancreatic Cancer	0.2	0		8 human
Infectious disease	Valley Fever (Cocci)	0 0 0	0 0 0	1 1 1	40 human 9 dog 60 mice
	Q-Fever *(Coxiella)*	0	0	1	30 rabbit
	Tularemia tularensis	0	0	1	21 mice
	Rickettsia rickettsii	0	0	1	37 mice
	Glanders *(B. mallei)*	0	0	1	26 mice
Chronic	Asthma IgG Asthma IgE Asthma IgA	0 0 0	0 0.01 0	1 0.97 1	25 human 25 human 6 human
Autoimmune	Lupus	0	0	1	60 mice
	Type I diabetes	0.13	0.02	0.80	56 human
Cancer	Glioblastoma	0	0	1	10 human
	Breast	0.01	0.009	0.986	156 human
	Multiple primary	0	0	1	7 human
	Pancreatic cancer	0.06	0.02	0.92	122 human
	Lung	0	0.11	0.81	60 human
	Myeloma	0	0	1	15 human
	Ovarian	0	0	1	6 human
	Esophageal	0	0	1	2 human
Other	Transplant	0	0	1	6 human

Example 7 Epitope Mapping

[0123] Motif finding algorithms are able to find subtle patterns in sets of unaligned sequences. These algorithms may be classified in two main categories: deterministic and optimizing. Deterministic algorithms exhaustively search a sequence set for motifs fitting a well defined set of criteria. Some popular implementations of deterministic motif finding algorithms are TEIRESIAS or PRATT (Rigoutsos, *Bioinformatics.* 14, 55-67 (1998);. Jonassen (1997) *Comput. Appl. Biosci.* 13, 509-22).

[0124] The optimizing algorithms represent the motif probabilistically and try to maximize a scoring function. The optimization can be preformed stochastically such as using Gibbs motif sampling or by expectation maximization as implemented in MEME (Bailey (2006) *Nucleic Acids Res.* 1, W369-7316).

PREOMICS EXHIBIT 1002
Page 477 of 1861

[0125] An optimization approach is preferred when it is not known what criteria the motif should fulfill. The GLAM2 implementation of the Gibbs motif sampling algorithm will be used here because it allows for gaps (Frith (2008). *PLOS Comput Biol.* 4, e1000071.17).

[0126] An alternative to finding a motif among the peptides is to compare the peptides one at a time to the antigen sequence(s). The algorithm implemented in the RELIC MATCH 5 program compares each peptide sequence to the target protein sequence in five amino acid windows, and scores each window for similarity (18). The scores for all of the peptides are added up across the protein sequence to predict potential small molecule binding site. A similar approach can be used for predicting antibody recognition sites from dissimilar peptide sequences selected in a peptide microarray experiment.

[0127] Typically, epitope mapping is performed to identify the specific part of a protein target that is recognized by the antibody. A similar approach can be used to identify an unknown protein target of an antibody. This approach can be used in identifying the antigenic proteins in a pathogen, targets in an autoimmune disease, or discovering the cause of an unknown infection.

[0128] Antibodies with known epitopes were purchased from Labvision (Fremont, CA) and Abcam (Cambridge, MA). Mice were immunized with keyhole limpet hemocyanin (KLH) conjugated peptides and sera was obtained at day 35. After slides were passivated with 0.014% Mercaptohexanol, antibody was diluted to 100nM or sera was diluted 1:500 in 3% BSA, 0.05% Tween, PBS. Antibodies were incubated with slides for 1 hour at 37C in Agilent Chambers with rotation. Slides were washed three times with TBS, 0.05% Tween and three times with diH2O. The incubation and wash procedure was repeated with a biotinalyted secondary antibody (Bethyl Laboratoreis, Inc. Montgomery, TX), then with Alexa-555 labeled Streptavadin (Invitrogen, Carlsbad, CA).

Name	Isotype	Num of AA	pI	Hydropathicity
mAb1	IgG1	9	3.56	-0.9
mAb2	IgG1	9	8.75	-0.044
mAb3	IgG1	5	9.76	-0.02
mAb4	IgG2a IgG2b	6	5.84	0.517
mAb5	IgG1kappa	6	4.67	-0.583
pAb1 polyclonal	IgG	20	4.21	-0.185
pAb2 polyclonal	IgG	20	4.43	-1.245
pAb3 polyclonal	IgG	20	6.4	0.87
pAb4 polyclonal	IgG	20	8.64	-0.13
pAb5 polyclonal	IgG	20	3.49	-0.315

PREOMICS EXHIBIT 1002
Page 478 of 1861

[0129] Negative control arrays with no primary antibody or naïve mouse sera were also run for comparison. At least three replicate arrays were run for each antibody.

[0130] Antigen specific antibodies were absorbed from sera by binding to KLH immobilized on nitrocellulose membrane. A one by six centimeter nitrocellulose membrane was placed in a 15 ml conical tube with 1.0 mg/ml KLH in 2.0 ml PBS. The membrane was washed three times in TBST and incubated with 1.0% BSA in TBST for at least one hour or until used. After washing three times in TBST, the membrane was placed into 2.0 ml of sera diluted 1:500 in 3% BSA, 0.05% Tween, PBS buffer. Reactivity of sera for KLH was tested in an ELISA. Sera were considered absorbed when no reactivity for KLH was detected at the 1:500 dilution.

[0131] Negative control signals were subtracted from antibody signals to remove the contribution of the secondary binding. The top 500 peptides in fluorescent intensity were selected for each antibody. The number of times each peptide occurs in one of the top 500 peptides lists was tabulated. Peptides appearing in five or more lists were eliminated as they are likely Fc binders or other nonspecific interactions. Peptides from the array were compared to the epitope sequences to identify those with sequence similarity. The epitope was expressed as a GLAM2 motif and was used in GLAM2SCAN to search against the peptides from the array inserted in strings of cysteines, with an alphabet of equal amino acid frequencies. Peptides were sorted by the highest scoring match and lists of the best matching peptides were created. These lists were compared with lists of peptides that most strongly bind to each peptide and the proportion of overlap was examined. Test datasets were generated for the monoclonal antibodies by randomly selecting sequences from human Swissprot and then randomly selecting a window of that sequence the same length as the epitope sequence. Two hundred negative examples were generated for each monoclonal. One thousand random peptides were generated as the negative examples for the polyclonal antibodies with equal frequencies of the nineteen amino acids (cysteine was not included as in the arrays). All of these sequences were inserted within a string of seventeenc cysteines on each side to allow peptides to be aligned overhanging the test sequences.

[0132] Motifs were generated from the peptide lists using GLAM2, with a starting width of five amino acids, 1,000,000 iterations without improvement, 10 runs, and an alphabet of equal proportions of the 20 amino acids. GLAM2SCAN was used to search the corresponding test sets for sequences matching the motif with the alphabet set as the default protein alphabet for the monoclonal antibodies or equal amino acid frequencies for the polyclonal antibodies.

PREOMICS EXHIBIT 1002
Page 479 of 1861

[0133] GLAM2SCAN output is the score for each place the motif matches in the test sequence set. The test sequences were ranked by the highest score match within each sequence. The RELIC Fastaskan program was used to align the binding peptides to the test dataset.

[0134] The top 500 specific peptides were uploaded as the affinity selected peptides and the corresponding test dataset was uploaded as the FASTA file. Random peptides were not subtracted. Fastaskan compares each five amino acid window of the test sequence with the selected peptide sequence and summing scores of the alignments above a threshold. It outputs a score for each test sequence corresponding to the window of maximum similarity between the peptides and that sequence.

[0135] For both the GLAM2SCAN and the RELIC analysis, the rank of the true epitopes was compared to the test sequences using ROC analysis. A Matlab script to calculate the true positive and false positive rate for each score cutoff was obtained from world wide web //theoval.cmp.uea.ac.uk/~gcc/matlab/roc/ and modified to smooth tied scores. The area under the ROC curve was also calculated using a Matlab script from the same website. The area under the curve is used to predict the probability of finding an epitope in a database of a given size. We will assume positive and negative examples will be selected from a database of a fixed size without replacement weighted by the probability that a positive is chosen over a negative as estimated by the area under the curve.

[0136] T evaluate the peptide microarray platform, examples of known epitope antibodies were required to generate a test dataset. Five monoclonal antibodies with known linear epitopes, and five examples of anti-peptide polyclonal mouse sera raised against peptides selected from the array were used as the test set. The monoclonal antibodies were found to bind to a median of 64.1% (range 37.6% -74.9%) of the random peptides above the slide surface background and secondary only controls. Polyclonal sera showed similar peptide reactivity with a median of 63.6% (range 54.0% - 68.6%). Replicate slides had an average Pearson correlation of 0.785 for monoclonals and 0.764 for polyclonals. A heatmap showed that each antibody has a distinct binding pattern on the array. Although there is some overlap between the peptides bound by each antibody, about 22% of the top 500 peptides recognized by each antibody are not recognized by the other nine antibodies tested. The uniqueness of the peptides recognized by each antibody implies that the peptide sequences may contain information about antibody specificity.

PREOMICS EXHIBIT 1002
Page 480 of 1861

[0137] Each peptide sequence was scored for similarity against each protein sequence. Most of the peptides bound by the antibodies did not show strong sequence similarity to the epitope. However, there was some enrichment for sequence similar peptides among the binders. Most of the peptides bound are mimotopes rather than having any obvious similarity to the epitope.

[0138] To assess the predictive power of these sequences the alignment of the peptides to the epitopes was compared to their alignment with a set of negative examples. The RELIC alignment program was able to align binding peptides to all of the monoclonal epitopes and 62.7% of the negative examples. The true epitopes had an average score of 14.3 whereas the negative examples had an average score of 5.9. The ROC analysis found an area under the curve of 0.87 indicating that a true epitope has an 87% chance of having a higher score than randomly selected negative example. All of the polyclonals also had positive peptide alignment scores as well as 86.5% of the positive examples. The true epitopes had an average score of 14.7 whereas the negative examples had a score of 15.2. The ROC analysis (Fig. 5) indicates that a positive example has a 46% chance of having a higher score than a negative example based on the area under the curve. The monoclonal epitopes were predicted well by this method, whereas the polyclonal predictions were similar to chance. An algorithm capable of detecting subtle patterns may be able to garnish predictive power from these peptide sequences. Convergent motifs were identified for all of the antibodies using GLAM2. The motifs for the monoclonal antibodies ranged from three to five amino acids in width. The polyclonal motifs were four to five amino acids wide. The monoclonal motifs matched the epitope sequences with an average score of 3.5, whereas the negative examples had an average score of -3.7. Polyclonal motifs matched the immunizing peptide with an average score of 3.8 whereas the negative examples had an average score of 3.7. The ROC analysis demonstrates that the monoclonals epitopes have an 89.8% chance of being scored higher than the corresponding negative examples in the motif analysis while the polyclonals have a 67.9% chance of scoring higher than the negatives. The motif finding approach demonstrated predictive power on both datasets.

	binders	***aligner***	***both***	***expected***	***ratio***	***% of binders align***
AbCamHA	350	181	**6**	5.88	1.02	1.7%
DM1A	391	379	**21**	13.75	1.53	5.4%
LNKB2	354	96	**2**	3.15	0.63	0.6%
P53Ab1	379	188	**36**	6.61	5.45	9.5%

PREOMICS EXHIBIT 1002
Page 481 of 1861

	binders	*aligner*	*both*	*expected*	*ratio*	*% of binders align*
P53Ab8	369	365	**6**	12.50	0.48	1.6%
Neg1	258	722	**26**	17.28	1.50	10.1%
Rco4	258	755	**26**	18.07	1.44	10.1%
Rco3	263	710	**18**	17.32	1.04	6.8%
Rco2	274	742	**22**	18.86	1.17	8.0%
Rco1	267	699	**21**	17.31	1.21	7.9%
average	316.3	483.7	**18.4**	13.07	1.55	6.2%

[0139] To test if combining the two approaches may improve the predictive ability, the scores from the RELIC analysis and the GLAM2 analysis were each scaled to have a minimum score of zero and a maximum score of one and averaged. The ROC analysis was performed on the averaged scores. The area under the curve was 0.92 for the monoclonals and 0.69 for the polyclonals. Based on the probability estimated from the ROC analysis, there i about a 70% chance of finding a monoclonal epitope in the top ten windows out of a one hundred amino acid protein. There is about a 21% chance of correctly identifying a polyclonal epitope in a small virus among the top 100 hits out of a possible 1000 amino acid database, which is a two-fold enrichment.

Example 9: General Materials and Methods

[0140] **Preparation of the Random Peptide Microarrays.** The random CIM10,000 f microarray consists of 10,000 20-residue peptides of random sequence, with a C-terminal linker of Gly-Ser-Cys-COOH. Each peptide was manufactured by Alta Biosciences, Birmingham, UK based on amino acid sequences we provided. Random sequences were provided by custom software (Hunter, Preston and Uemura, Yusuke, The Biodesign Institute). Nineteen amino acids (cysteine was excluded) were selected completely at random in each of the first seventeen positions with GSC as the carboxy-terminus linker. The synthesis scale was 2-5 mg total at $\geq$ 70% purity with 2% of the peptides tested at random by mass spectrometry. Dry peptide was brought up in 100% *N,N'*dimethyl formamide until dissolved, then diluted 1:1 with purified water at pH 5.5 to a master concentration of 2 mg/ml. The original 96-deep-well plates were robotically transferred to 384-well potting plates, and the peptides were diluted to a final spotting concentration of 1mg/ml concentration in phosphate buffered saline at pH 7.2. High-quality pre-cleaned Gold Seal glass microscope slides were obtained from Fisher (Fair Lawn, NJ, cat# 3010). Each slide was treated with amino-silane, activated with sulfo-SMCC (Pierce Biotechnology, Rockford, IL) creating a quality-checked maleimide-activated surface designed to react with the

PREOMICS EXHIBIT 1002
Page 482 of 1861

peptide's terminal cysteine. During spotting, we employ a Telechem Nanoprint 60 using 48 Telechem series SMP2 style 946 titanium pins. Each pin is allowed to spot approximately 500 pL of 1mg/ml peptide per spot, which is estimated by allowing for pin trajectory, surface dwell time, and the amount of liquid each pin holds. The spotting environment is 25°C at 55% humidity. Each peptide is spotted twice per array; the arrays are spotted in an orange-crate packing pattern to maximize spot density. Fluorescent fiducials are applied asymmetrically using Alexa-647, Alexa-555, Alexa-488 and Alexa-350-labeled peptides. The fiducials are used to align each subarray during image processing. The printed slides are stored under argon at 4°C until used. Quality control consists of imaging the arrays by laser scanner (a Perkin-Elmer ProScanArray HT, Perkin Elmer, Wellesley, MA) at 647nm to image the spot morphology. If the batch passes this test, further testing of randomly selected slides with known proteins and antibodies allows QC of reproducibility. Array batches that fail to meet an array-to-array variability of 30% total CV are discarded. Data extraction is accomplished using GenePix Pro 6.0 (Molecular Devices Inc., Sunnyvale, CA).

[0141] Probing the Random Peptide Microarrays with Serum Antibodies. Slides were processed using a Tecan HS4800 Pro Hybridazation station with a protocol adapted to antibody binding. General settings were wash duration of 30 sec at 11.0 ml/min and sample agitation was set to high. Arrays were blocked with 1X PBS, 3% BSA, 0.05% Tween 20, 0.014% mercaptohexanol for 15 min at room at 23°C. Arrays were probed with 170 µl of serum diluted at 1:500 in incubation buffer (3% BSA, 1X PBS, 0.05% Tween 20) for 1 hour at 37°C. Slides were washed between with TBST. Bound IgG was detected using biotinylated anti-mouse IgG (Bethyl Laboratories, Montgomery, TX) for 1 hour at 37°C. The anti-alpha chain secondary antibody was detected using Alexafluor 649 labeled streptavidin at 5.0 nM in incubation buffer for 1 hour at 37°C. Final washes in TBST and distilled water were done to remove residual salt. Slides were dried by on board nitrogen flow for 5.0 min. Images were recorded using an Agilent 'C' type scanner at both 543 nm and 633 nm.

[0142] Analysis of Random Peptide Microarray Data. Statistical analysis of microarray data was done with GeneSpring 7.3.1 (Agilent, Inc., Palo Alto, CA) by importing image-processed data from GenePix Pro 6.0 (Axon Instruments, Union City, CA). Calculations based on the GenePix prepared gpr text files were done on the median signal intensity per spot. Poor quality spots were excluded from analysis by flagging as "absent" upon visual inspection. Prior to analysis, each array was normalized to the 50th percentile to eliminate array to array variation and signal intensities of less than 0.01 were set to 0.01. Values from triplicate arrays were averaged and used in the analysis. Informative peptides were

PREOMICS EXHIBIT 1002
Page 483 of 1861

determined by comparing groups on a peptide to peptide basis. Peptides with a relative fluorescence intensity of 500 or greater and a p value < 0.01 were selected. Further statistical analysis was conducted in MicroSoft Excel 2003 SP3 or in GraphPad Prism version 4.00 for windows (Graphpad Software, San Diego, CA). The principle component analysis feature in Gene Spring was used to distinguish serum samples based on selected gene lists. PCA was run using mean centering and scaling. Support vector machine analysis was run in GeneSpring using Fisher's Exact Test limited to 50 peptides, polynomial dot product order 1 and a diagonal scaling factor of 0 to generate cross validation errors and predict test sets.

[0143] Characterization of archived murine serum samples. To test whether an immunosignature correlates with vaccine efficacy, even if protection is driven by a small number of neutralizing antibodies, a sublethal influenza A/PR/8/34 infection was used as a model protective vaccine and KLH immunization as a model non-protective vaccine. KLH was chosen due to its ability to generate an equivalently robust humoral immune response. The average IgG titer against whole virus in A/PR/8/34 infected mice was 819,000 and 800 for KLH immunized mice. Lethal challenge using 2-5 mean lethal doses of influenza A/PR/8/34 occurred on day 35. The immune response generated by the sublethal infection was protective as no overt clinical signs of infection or weight loss were observed following challenge. Mice immunized with KLH contained either KLH alone or unrelated peptide-KLH conjugates adjuvanted with Alum. The average anti-KLH IgG titer was 819,000 in KLH immunized mice and 1,600 in A/PR/8/34 immunized mice. Following challenge at day 35, KLH immunization imparted no benefit over naives. This indicates that the model vaccines had equivalently strong immune responses for their respective immunogen with little ELISA cross-reactivity and no protection due to immunization with KLH.

PREOMICS EXHIBIT 1002
Page 484 of 1861

WHAT IS CLAIMED IS:

1. A method of analyzing a sample, comprising:

(a) contacting the sample with an array of immobilized different compounds occupying different areas of the array, wherein different molecules of the same compound within an area are spaced sufficiently proximate to one another for multivalent binding between at least two of the different molecules in the same area and a multivalent binding partner; and

(b) detecting binding of the different compounds in the array to component(s) of the sample, such as antibodies.

2. The method of claim 1, further comprising characterizing the sample from the relative binding of compounds whose binding to the sample would be difficult to distinguish from each other and nonspecific binding under conditions of monovalent binding but which show significantly different binding from each other and nonspecific binding due to multivalent binding of the sample to the compounds thereby generating a binding profile characteristic of the sample.

3. The method of claim 1, wherein the average spacing between different molecules of a compound in an area of the array is less than 6 nm.

4. The method of claim 1, further comprising characterizing the sample from the relative binding a plurality of the compounds binding to the sample with association constants of 1 mm to 1 μM.

5. The method of claim 4, wherein the plurality of compounds includes at least 10 or 100 compounds binding to the sample with association constants of 1 mm to 1 μM.

6. The method of claim 5, wherein the characterizing comprises comparing a binding profile of the sample that includes the relative binding of the plurality of compounds with a reference binding profile.

7. The method of claim 1, further comprising identifying a component of the sample that binds to the different compounds.

PREOMICS EXHIBIT 1002
Page 485 of 1861

8. The method of claim 6, further comprising detecting the identified component with a binding partner known to bind the component.

9. The method of claim 7, wherein the identified component is detected with a plurality of different binding partners known to bind the component.

10. The method of claim 7, wherein the binding partner is an antibody to the identified component or a peptide known to bind the identified component.

11. The method of claim 7, wherein the binding partner is one of the different compounds detected in step (b).

12. The method of claim 7, wherein the binding partner is a synbody.

13. The method of claim 7, wherein the binding partner is immobilized to a support.

14. The method of claim 13, wherein the binding partner is immobilized to a support in an array.

15. The method of claim 14, further comprising forming a second array, the second array containing one or more of the different compounds in the array binding to identified component not all of the different compounds in the array.

16. The method of claim 1, wherein the second array contains less than 5% of the different compounds in the array.

17. The method of claim 1, further comprising forming an array or other device comprising compounds determined to bind to the sample but not all of the different compounds in the array.

18. The method of claim 1, wherein the detecting step detects binding of the different compounds to an antibody or antibodies in the sample.

19. The method of claim 1, wherein the detecting step detects binding of the different compounds to a biological entity displaying multiple copies of a protein from its outersurface.

20. The method of claim 1, wherein the biological entity is cell displaying multiple copies of receptor from its outersurface.

PREOMICS EXHIBIT 1002
Page 486 of 1861

21. The method of claim 1, wherein the compound having strongest binding to the sample binds to the sample with a dissociation constant of 1 mM to 1 μM.

22. The method of claim 1, wherein the compounds are peptides or small molecules.

23. The method of claim 1, wherein the array has 500-50,000 peptides.

24. The method of claim 23, wherein the peptides are 10-30 amino acid long.

25. The method of claim 24, wherein the sequences of the peptides are randomly selected.

26. The method of claim 1, wherein the different immobilized compounds are selected without regard to the sample and the array further comprises a plurality of compounds known to bind different proteins also occupying different areas of the array. .

27. The method of claim 26, wherein the plurality of compounds known to bind different proteins includes compounds known to bind at least 25%, 50 or 75% of different human proteins.

28. The method of claim 27, wherein the different immobilized compounds including a plurality of compounds known to bind at least 25%, 50% or 75% of different human proteins and 500-50,000 random peptides.

29. The method of claim 23, wherein the sequences of the peptides have less than 90% sequence identity to a known binding partner of the target.

30. The method of claim 23, wherein the sequences of the peptides have less than 90% sequence identity to known proteins.

31. The method of claim 1, wherein the contacting of the sample to the array is performed in the presence of a potential competitor of binding of the sample to the array.

32. The method of claim 31, wherein the competitor is a known binding partner of a suspected component of the sample.

PREOMICS EXHIBIT 1002
Page 487 of 1861

33. The method of claim 1, wherein the sample is a patient sample, and the competitor is a protein known to be associated with a disease affecting the patient.

34. The method of claim 1, wherein the sample is a patient sample.

35. The method of claim 1, wherein the sample contains a plurality of antibodies.

36. The method of claim 34, wherein the patient is known or suspected to be suffering from a disease.

37. The method of claim 34, wherein the patient is known to be at risk of a disease but is not showing symptoms of the disease.

38. The method of claim 36 or 37, wherein the disease is an autoimmune disease.

39. The method of claim 36 or 37, wherein the disease is an infectious disease.

40. The method of claim 36 or 37, wherein the disease is a disease of the CNS.

41. The method of claim 33, wherein the sample is a blood, urine, or CNS sample.

42. The method of claim 1, wherein component(s) of the sample are labeled.

43. The method of claim 1, wherein binding of the peptides to component(s) of the sample is detected using a secondary antibody.

44. The method of claim 43, wherein the secondary antibody is an isotype-specific antibody.

45. The method of claim 1, wherein binding of the peptides to component(s) of the sample is detected by spr or mass spectrometry.

46. The method of claim 1, further comprising affinity purifying a component of the sample using a peptide determined to bind to the sample.

PREOMICS EXHIBIT 1002
Page 488 of 1861

47. The method of claim 1, further comprising washing unbound component(s) of the sample from the array, and dissociating bound component(s) from the array.

48. The method of claim 47, further comprising preparing an antibody library from the patient and using a peptide to which an antibody in the sample binds as an affinity reagent to screen the library.

49. The method of claim 48, further comprising identifying a natural binding partner of the affinity purified antibody.

50. The method of claim 23, further comprising comparing the sequence(s) of peptide(s) binding to the component(s) of the sample to a database of natural sequences to identify natural binding partner(s) of components of the sample.

51. The method of claim 50, further comprising comparing a profile of different compounds binding to the sample with profiles of the different compounds associated with different diseases or different stages of a disease to diagnose a patient as having one of the diseases or stages of disease.

52. The method of claim 50, further comprising comparing a profile of different compounds binding to the sample with a profile of the different compounds associated with lack of a disease to determine whether a disease is present.

53. The method of claim 50, further comprising repeating the method for different samples from a plurality of patients with the same disease to develop a binding profile characteristic of the disease.

54. The method of claim 50, further comprising repeating the method for different samples from a plurality of patients with different disease to develop a plurality of binding profiles characteristic of different diseases.

55. An array of immobilized different compounds occupying different areas of the array, wherein different molecules of the same compound within an area are spaced sufficiently proximate to one another for multivalent binding between at least two of the different molecules in the same area and a multivalent binding partner.

56. A method of analyzing a sample, comprising:

PREOMICS EXHIBIT 1002
Page 489 of 1861

contacting the sample with an array of immobilized different peptides occupying different areas of the array;

detecting binding of the different peptides in the array to component(s) of the sample; and

characterizing the sample from the relative binding of a plurality of peptides with dissociation constants between 1 mM and 1 μM for the sample.

57. The method of claim 56, wherein the sample is analyzed from the relative binding of at least ten peptides with dissociation constants between 1 mM and 1 μM.

58. A method of analyzing a sample, comprising

contacting the sample with an array of immobilized different compounds occupying different areas of the arrays;

detecting binding of the different compounds in the array to component(s) of the sample; and

characterizing the sample from the relative binding of a plurality of compounds having binding strengths to the sample greater than but within three orders of magnitude of the mean plus three standard deviations of the binding strength of empty areas in the array.

59. The method of claim 58, wherein the characterizing comprising comparing a binding profile of the sample including the plurality of compounds with a reference binding profile including the plurality of compounds.

60. The method of claim 58, wherein the sample is characterized from the relative binding of at least 10 or 100 compounds having binding strengths to the sample greater than but within three orders of magnitude of the mean plus three standard deviations of the binding strength of empty areas in the array.

61. A method of characterizing a plurality of different samples, comprising

contacting the different samples with the same array or copies of the same array of immobilized different peptides occupying different areas of the array; and

detecting different binding profiles of the different peptides to the different samples; wherein the samples are characterized from their respective binding patterns.

PREOMICS EXHIBIT 1002
Page 490 of 1861

62. The method of claim 61, wherein the plurality of different samples includes samples from patients with different disease symptoms.

63. The method of claim 61, wherein the plurality of different samples includes samples of patients presenting with disease and lack of disease

64. A method of analyzing a sample, comprising:

contacting the sample with an array of immobilized different compounds occupying different areas of the array, wherein different molecules of the same compound within an area are spaced at an average distance of less than 4 nm apart in the same area; and

detecting binding of the different compounds in the array to component(s) of the sample.

65. A method of analyzing a sample, comprising

(i) contacting a sample with a known binding partner of a component of the sample; and

(ii) determining whether the binding partner binds to the sample compared with a control lacking the component;

wherein the known binding partner is identified by a process comprising

(a) contacting an initial sample with an array of immobilized different compounds occupying different areas of the array, wherein different molecules of the same compound within an area are spaced sufficiently proximate to one another for multivalent binding between at least two of the different molecules in the same area and a multivalent binding partner; and

(b) detecting binding of the different compounds in the array to component(s) of the sample;

(c) identifying a component of the sample that binds to the different compounds; and

(d) identifying a known binding partner of the component for use in step (i)..

66. The method of claim 65, wherein the component in step (i) is detected with a plurality of different binding partners known to bind the component.

67. The method of claim 65, wherein the binding partner is an antibody to the component.

PREOMICS EXHIBIT 1002
Page 491 of 1861

68. The method of claim 65, wherein the binding partner is a peptide known to bind the component.

69. The method of claim 65, wherein the binding partner is one of the different compounds detected in step (b).

70. The method of claim 65, wherein the binding partner is a synbody.

71. The method of claim 65, wherein the binding partner is immobilized to a support in step (i).

72. The method of claim 65, wherein the binding partner is immobilized to a support in an array in step (i).

73. A method of manufacturing a device for use in detecting a component of a sample comprising;

(a) contacting the sample with an array of immobilized different compounds occupying different areas of the array, wherein different molecules of the same compound within an area are spaced sufficiently proximate to one another for multivalent binding between at least two of the different molecules in the same area and a multivalent binding partner; and

(b) detecting binding of the different compounds in the array to component(s) of the sample.

(c) forming a device including a known binding partner of a component to which binding of the different compounds is detected in step (b).

74. The method of claim 73, further comprising identifying the component of the sample that binds to the different compounds.

75. The method of claim 73, wherein step (c) comprising forming the device including a plurality of different binding partners known to bind the component.

76. The method of claim 73, wherein the binding partner is an antibody to the component.

77. The method of claim 73, wherein the binding partner is a peptide known to bind the component.

PREOMICS EXHIBIT 1002
Page 492 of 1861

78. The method of claim 73, wherein the binding partner is one of the different compounds detected in step (b).

79. The method of claim 73, wherein the binding partner is a synbody.

80. The method of claim 73, wherein the binding partner is immobilized to a support.

81. The method of claim 73, wherein the binding partner is immobilized to a support in an array.

82. The method of claim 73, wherein step (c) comprising forming a second array, the second array containing one or more of the different compounds in the array binding to component(s) of the sample and not all of the different compounds in the array.

83 The method of claim 82wherein the second array contains less than 5% of the different compounds in the array.

84 An array of immobilized different compounds occupying different areas of the array, wherein the different compounds include a plurality of compounds known to bind at least 25, 50 or 75% of the known human proteins and 500-50,000 random peptides, wherein different molecules of the same peptide within an area are spaced sufficiently proximate to one another for multivalent binding between at least two of the different molecules in the same area and a multivalent binding partner.

85 A method of testing a vaccine, comprising

contacting a blood sample of a subject immunized with a vaccine against a pathogenic microorganism with an array of immobilized different compounds occupying different areas of the array;

detecting a pattern of binding of the sample to the different compounds in the array; and

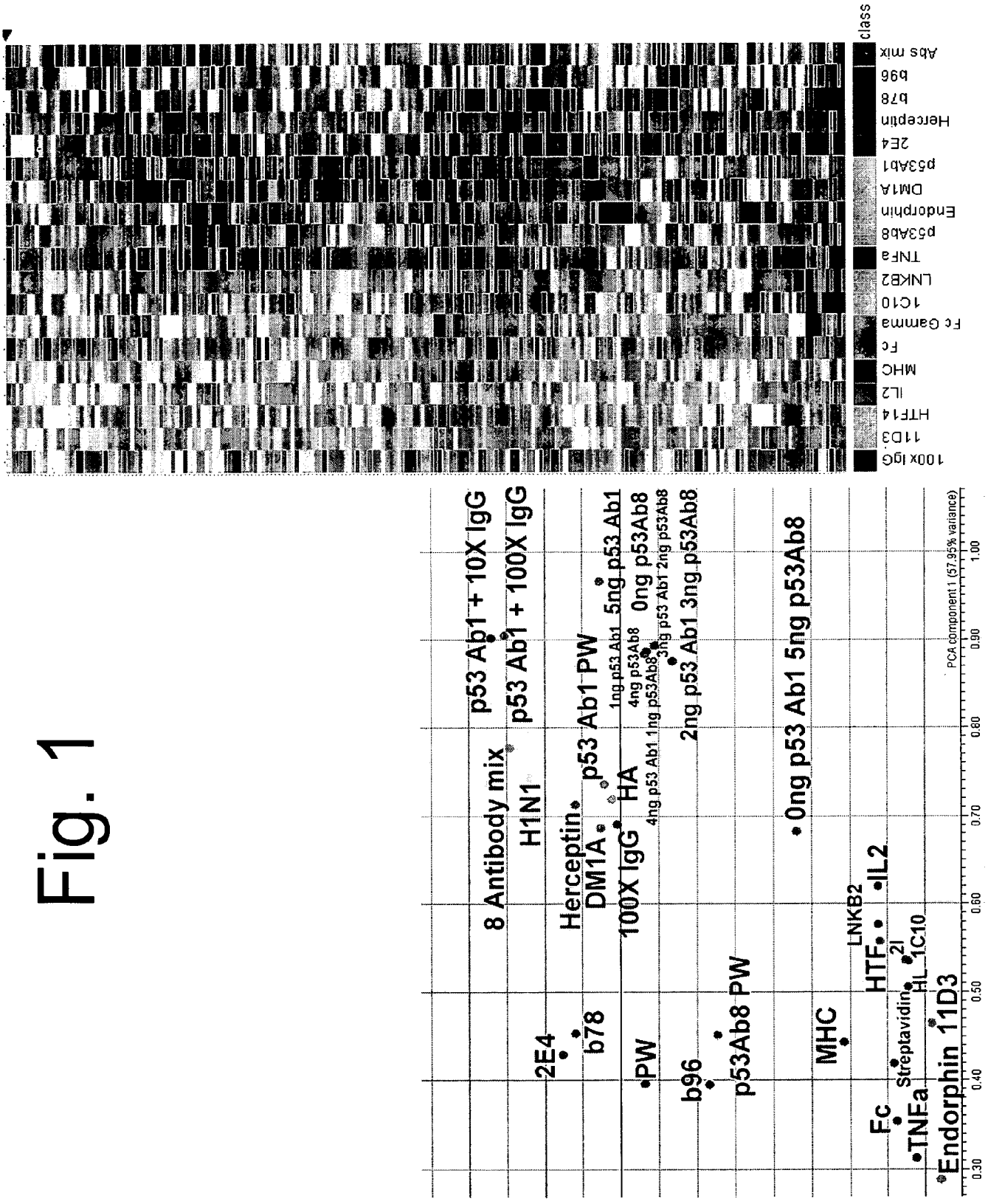
comparing the pattern of binding to the pattern of binding of one or more reference samples, wherein the reference samples are from subjects who have survived an infection with the virus, similarity of binding profile between the subject and reference samples providing an indication the vaccine is effective against the pathogenic microorganism.

PREOMICS EXHIBIT 1002
Page 493 of 1861

86. The method of claim 85, wherein the subjects have been immunized with a vaccine before exposure to the virus.

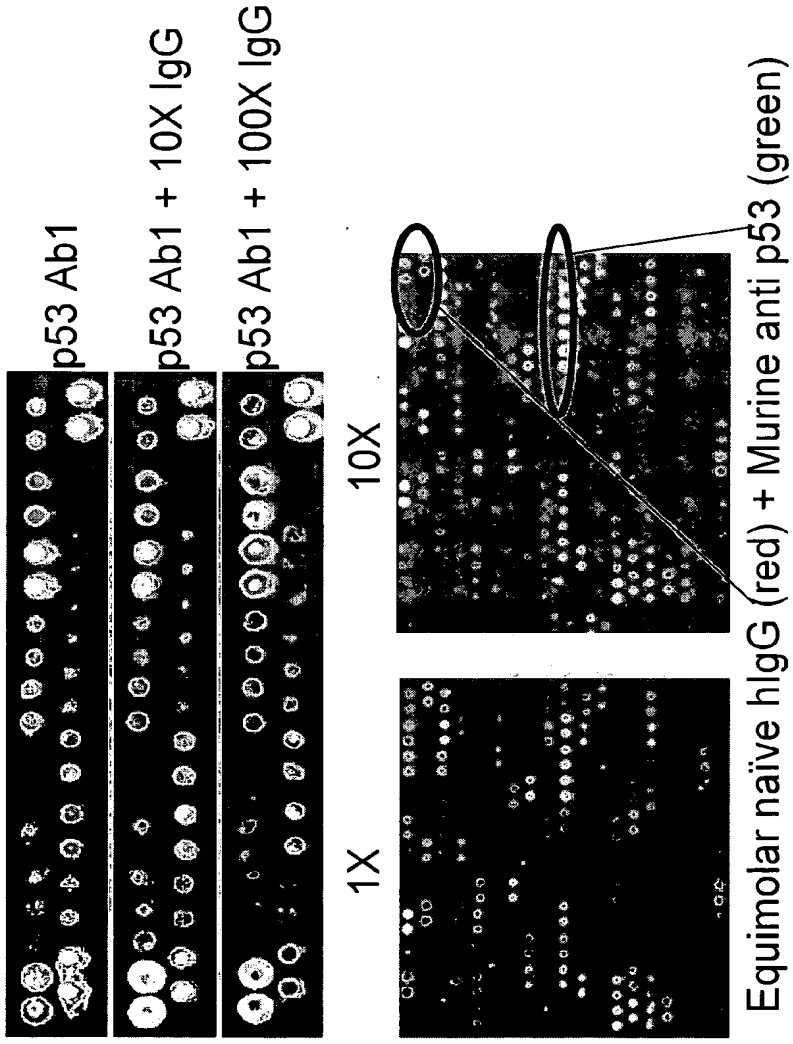
PREOMICS EXHIBIT 1002
Page 494 of 1861

Fig. 1



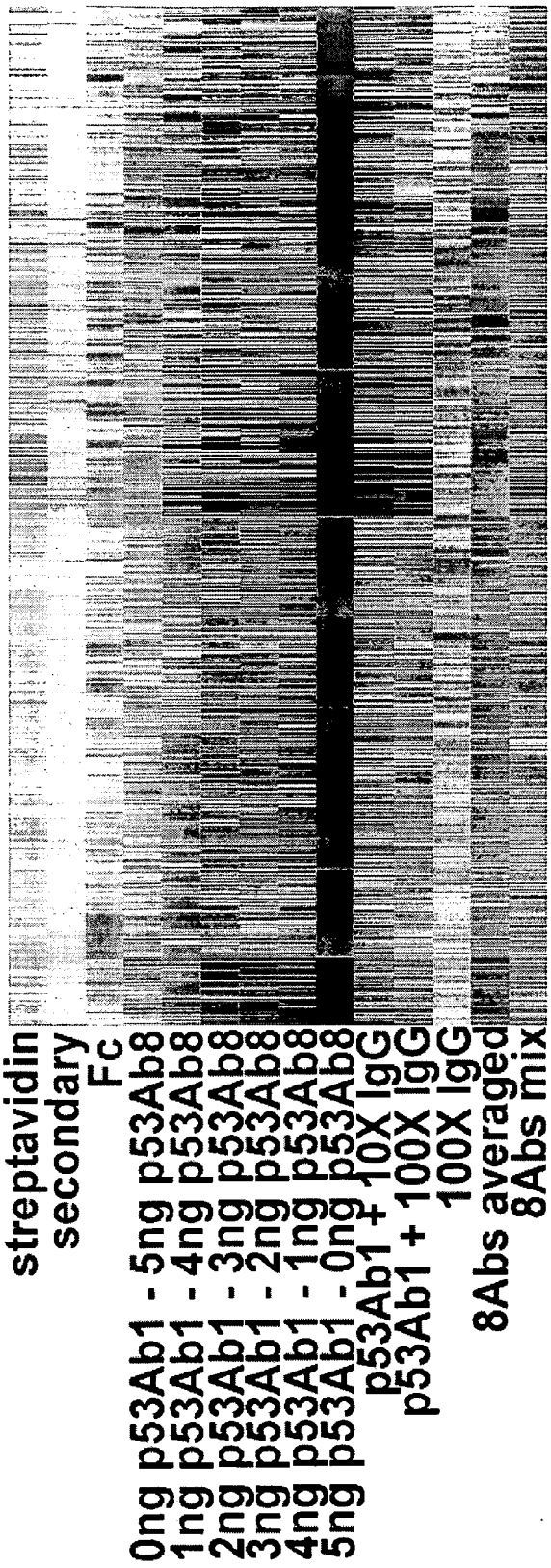
PREOMICS EXHIBIT 1002
Page 495 of 1861

Fig. 2



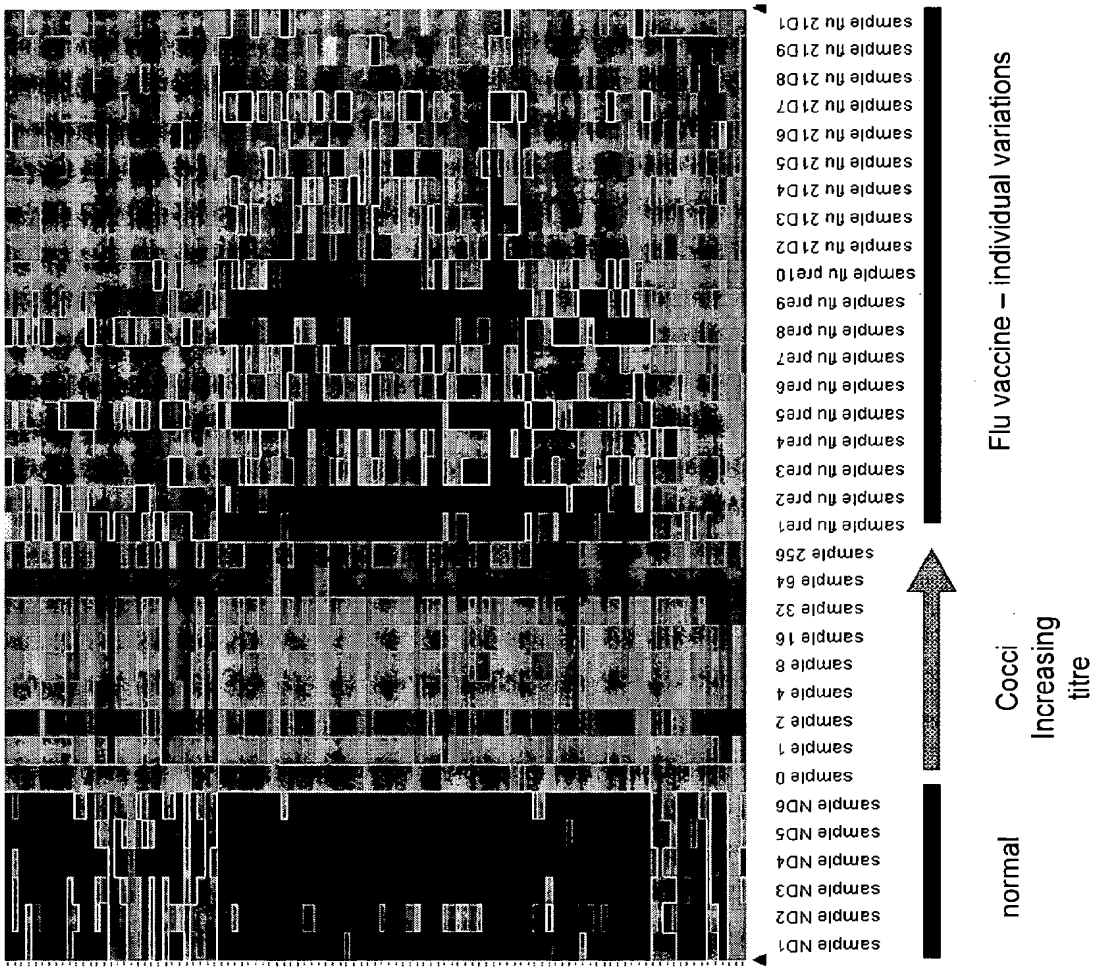
PREOMICS EXHIBIT 1002
Page 496 of 1861

Fig. 3



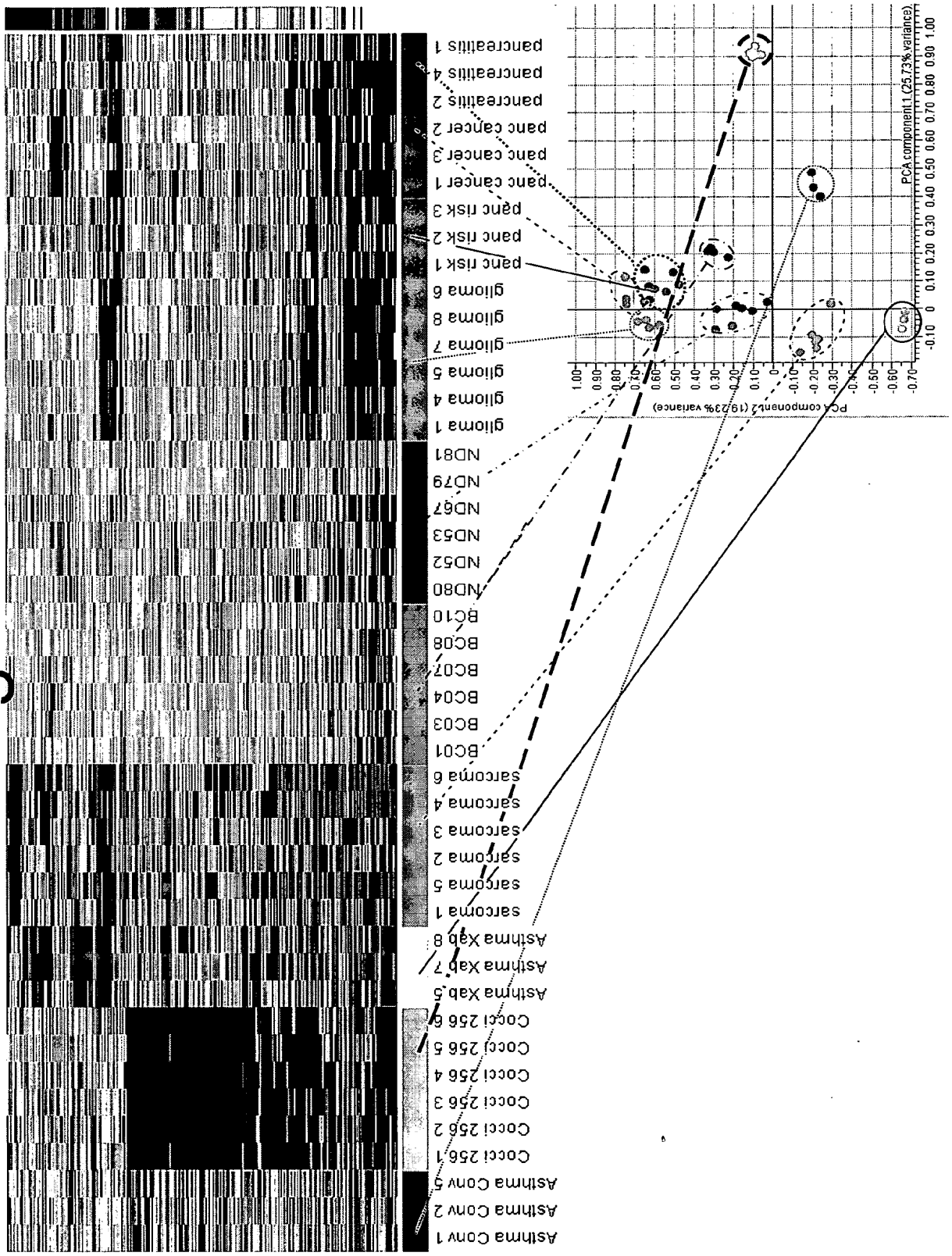
PREOMICS EXHIBIT 1002
Page 497 of 1861

Fig. 4



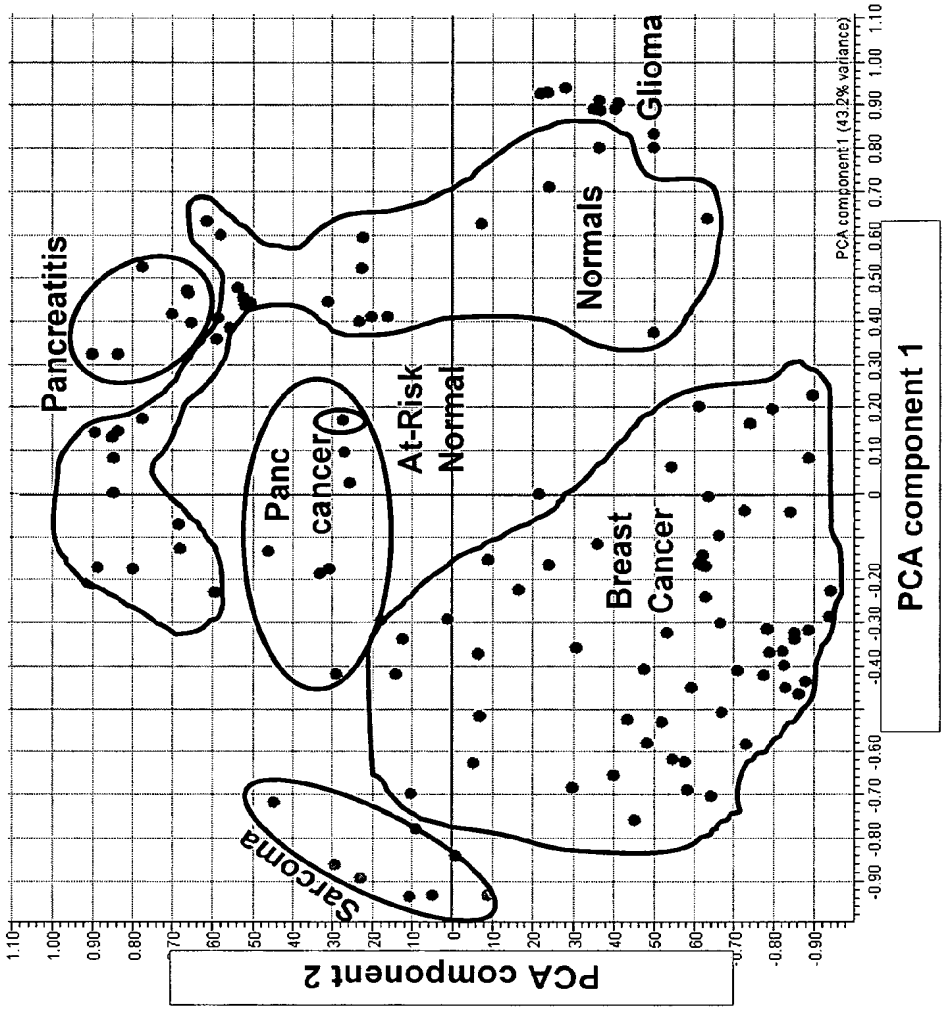
PREOMICS EXHIBIT 1002
Page 498 of 1861

Figure 5



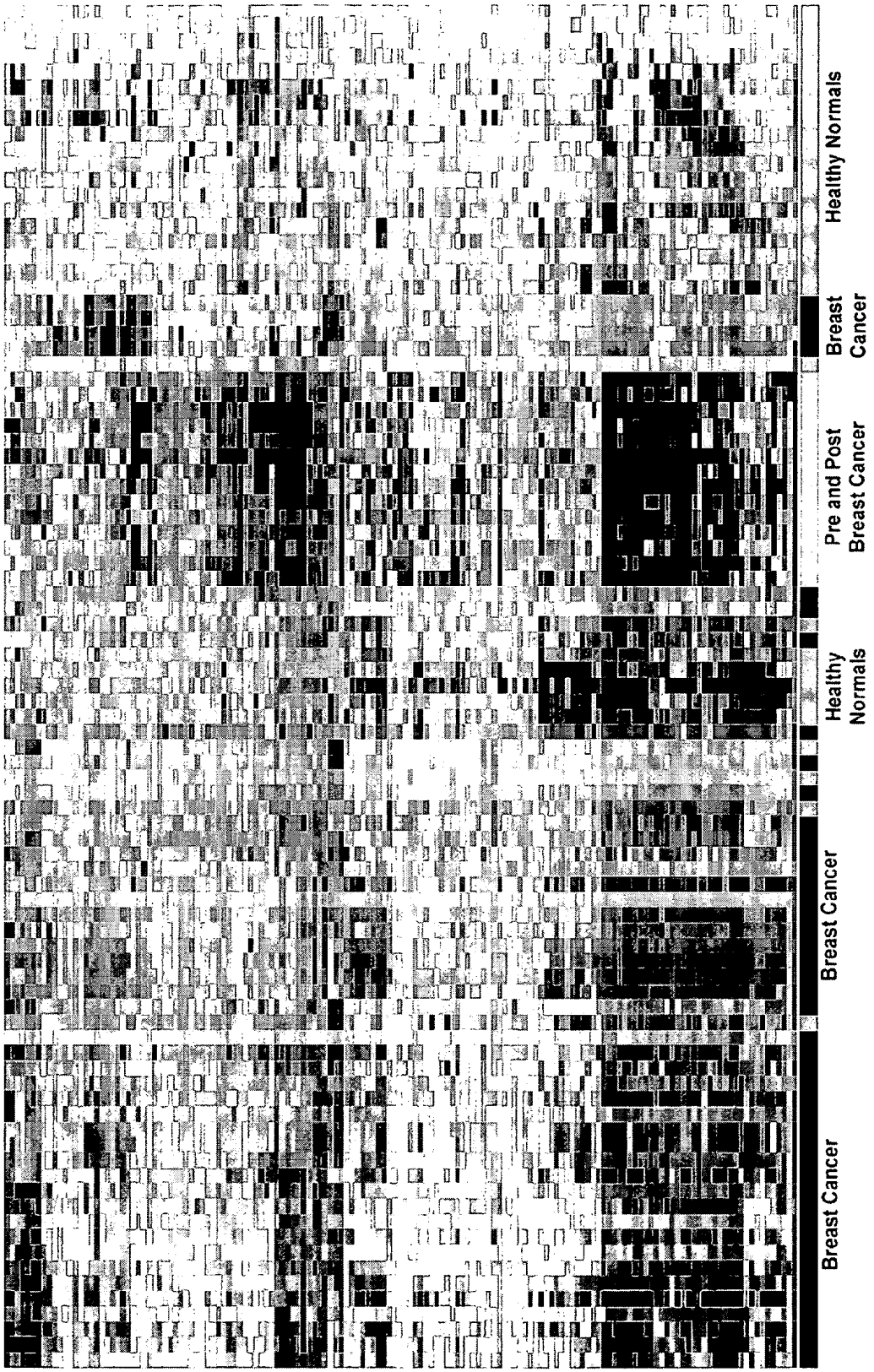
PREOMICS EXHIBIT 1002
Page 499 of 1861

Fig. 6



PREOMICS EXHIBIT 1002
Page 500 of 1861

Fig. 7



PREOMICS EXHIBIT 1002
Page 501 of 1861

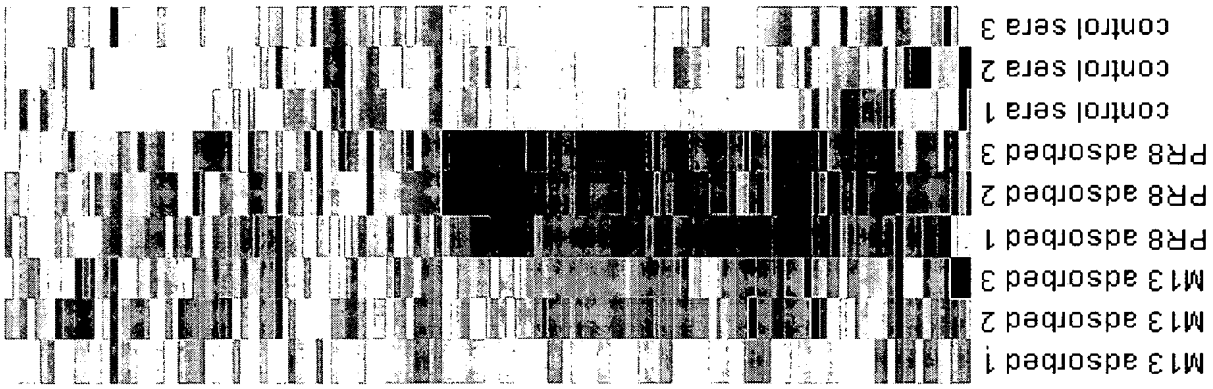
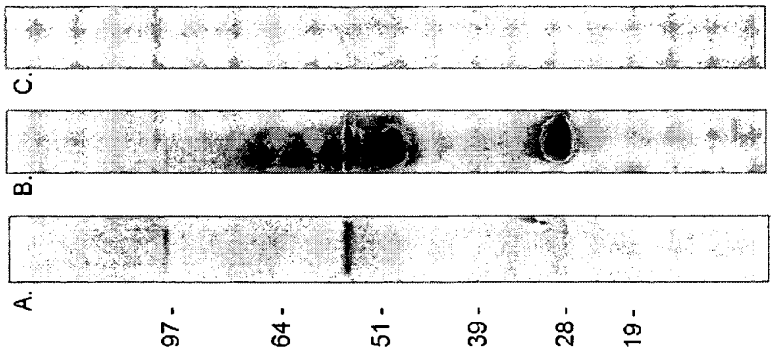


Fig. 8

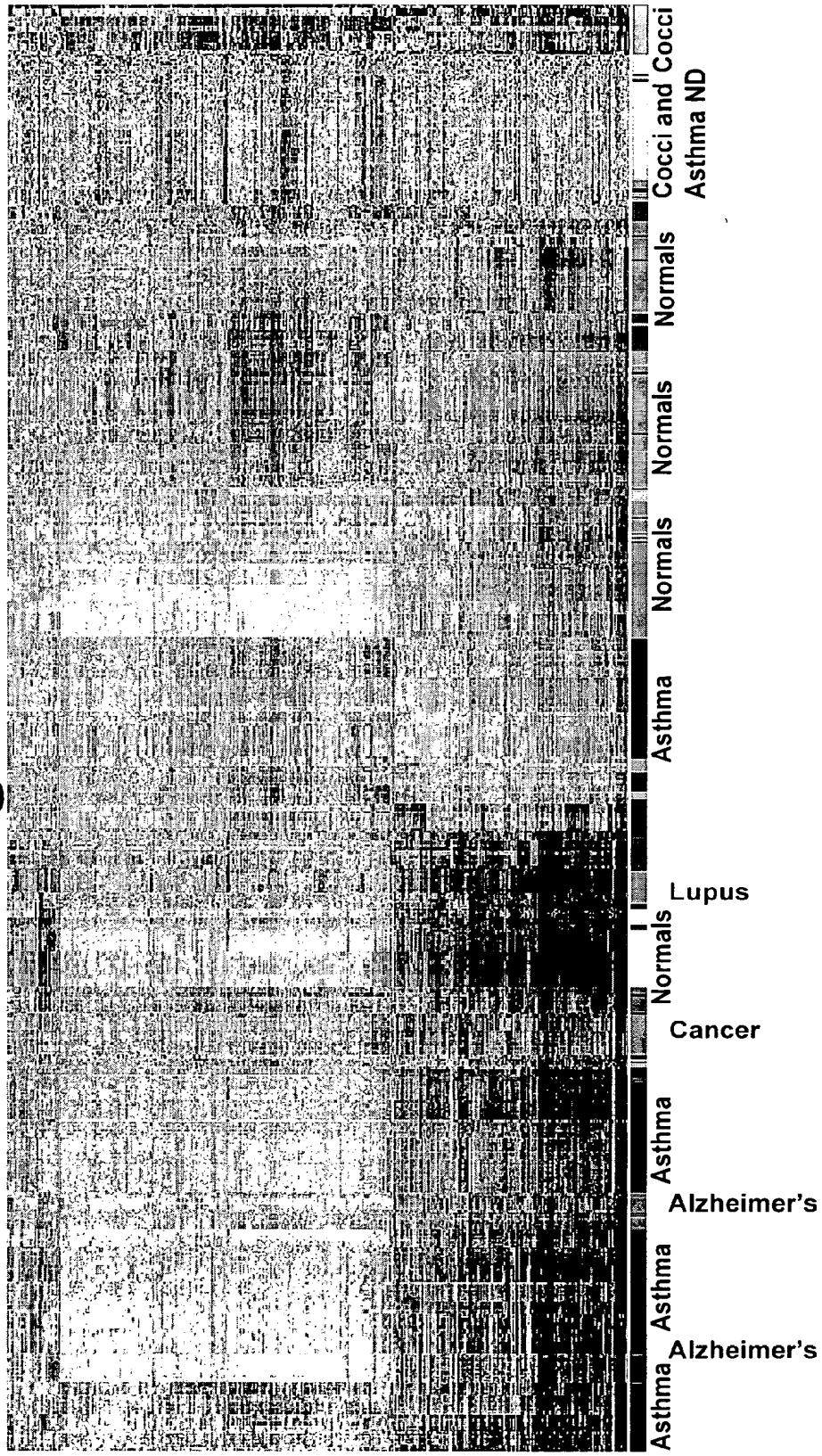
PREOMICS EXHIBIT 1002
Page 502 of 1861

Fig. 9



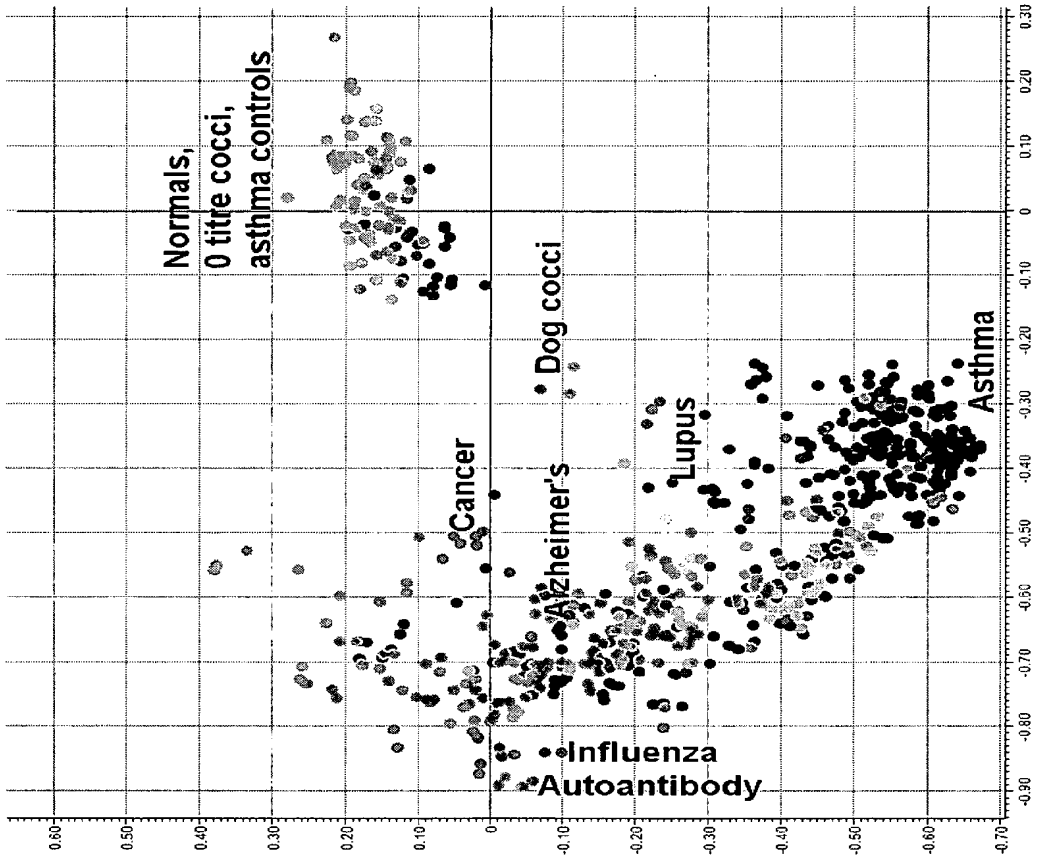
PREOMICS EXHIBIT 1002
Page 503 of 1861

Fig. 10A



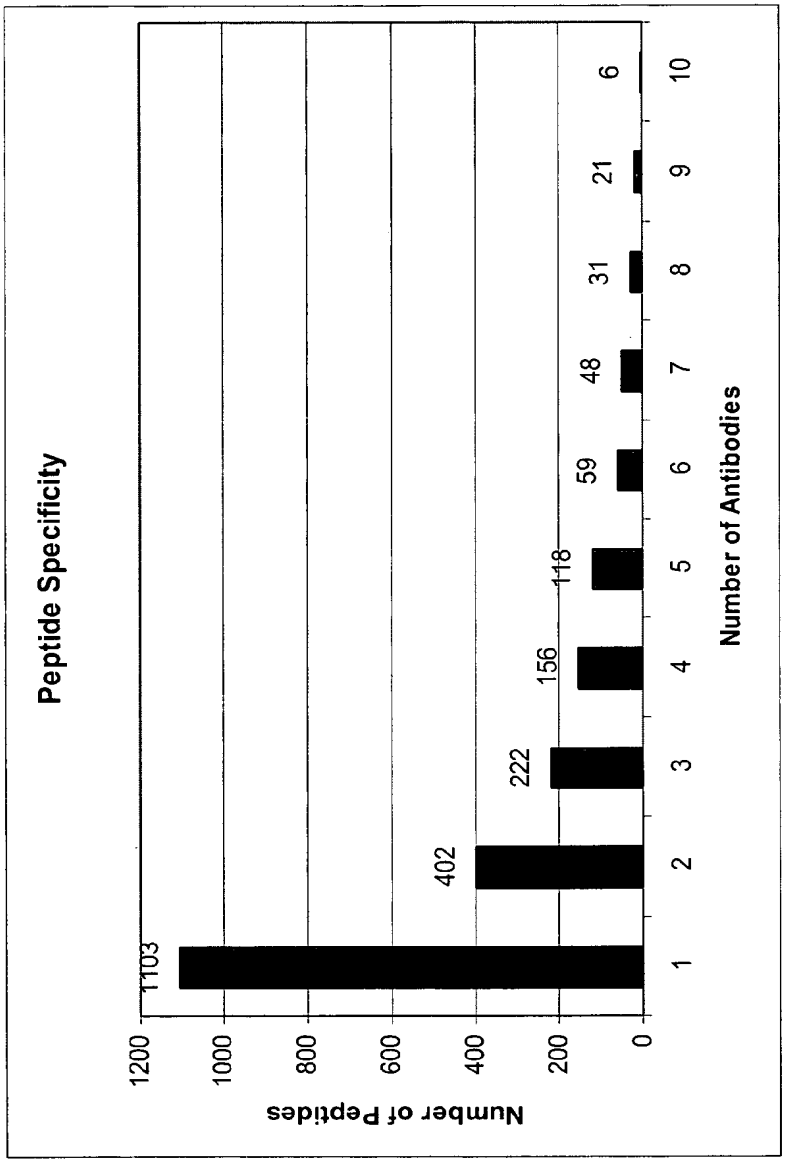
PREOMICS EXHIBIT 1002
Page 504 of 1861

Fig. 10B



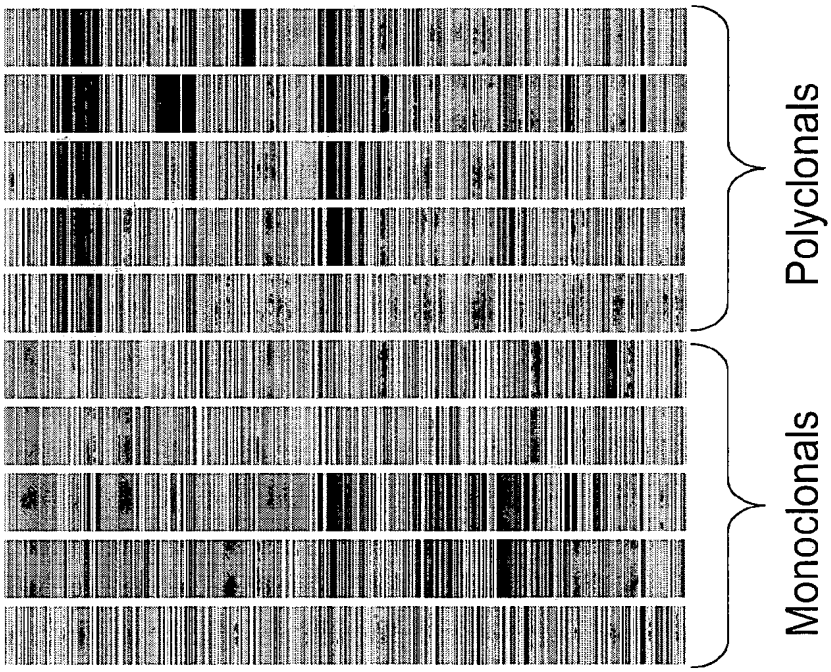
PREOMICS EXHIBIT 1002
Page 505 of 1861

Figure 11



PREOMICS EXHIBIT 1002
Page 506 of 1861

Figure 12



PREOMICS EXHIBIT 1002
Page 507 of 1861

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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
21 July 2011 (21.07.2011)

PCT

(10) International Publication Number
WO 2011/088128 A2

(51) International Patent Classification: Not classified

(21) International Application Number:
PCT/US2011/021002

(22) International Filing Date:
12 January 2011 (12.01.2011)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/294,877 14 January 2010 (14.01.2010) US
61/407,038 27 October 2010 (27.10.2010) US

(71) Applicant *(for all designated States except US)*: **UNIVERSITY OF CENTRAL FLORIDA RESEARCH FOUNDATION, INC.** [US/US]; 12201 Research Parkway, Suite 501, Orlando, FL 32826 (US).

(72) Inventor; and

(75) Inventor/Applicant *(for US only)*: **HUO, Qun** [CN/US]; 13507 Lakers Court, Orlando, FL 32828 (US).

(74) Agents: **VAN DYKE, Timothy, H.** et al.; Beusse Wolter Sanks Mora & Maire, 390 N. Orange Avenue, Suite 2500, Orlando, FL 32801 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

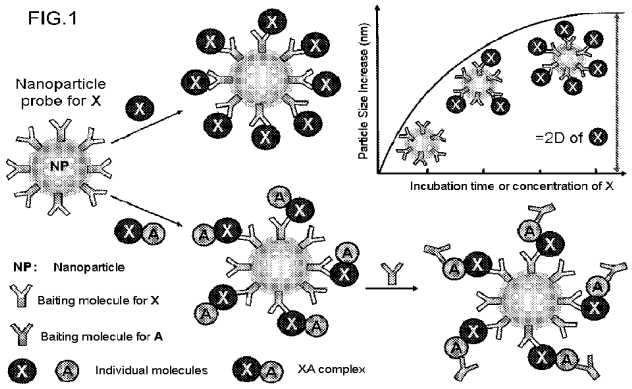
(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *without international search report and to be republished upon receipt of that report (Rule 48.2(g))*

(54) Title: METHODS FOR BIOMOLECULE AND BIOMOLECULE COMPLEX (BMC) DETECTION AND ANALYSIS AND THE USE OF SUCH FOR RESEARCH AND MEDICAL DIAGNOSIS

FIG.1



(57) Abstract: The present application pertains to improved methods of detecting biomolecules in a biological sample (or system), In particular, embodiments discussed herein allow for the detection of biomolecule complexes. The embodiments enable for the first time the elucidation of the significance of biomolecule complexes for certain disease states, which in turn enables the diagnosis of disease states based on the identity and complexing level of a biomolecule complex in a particular biological sample.

WO 2011/088128 A2

Methods for Biomolecule and Biomolecule Complex (BMC) Detection and Analysis and the Use of Such for Research and Medical Diagnosis

Cross-reference to related applications

[0001] This application is related to U.S. Provisional Application No. 61/294,877; filed January 14, 2010 and U.S. Provisional Application 61/407,038 filed October 27, 2010, and to which priority is claimed under 35 USC § 119.

Introduction

[0002] Biomolecular interactions such as protein-protein and protein-DNA interactions play an essential role in almost all cellular functions. Many intracellular biochemical processes are triggered by the assembly of biomolecules including proteins, DNAs and RNAs into biomolecule complexes (BMC), providing a means to control the myriad of biochemical processes for the efficient management of vital biological responses. The detection and analysis of these biomolecule complexes is not only critical for understanding the mechanisms of diseases, but also provide new methods and approaches for diagnosis and treatment of the diseases.

[0003] While a large number of bioanalytical techniques exist for detecting and analyzing the concentrations of individual biomolecules, techniques available for studying biomolecular interactions and for biomolecule complex detection and analysis are rather limited and inadequate to meet the demands and challenges of biomolecular research and medical diagnosis. Many traditional techniques such as co-immunoprecipitation, immunoaffinity chromatography, and yeast two-hybrid assays involve complicated assay procedures, require a large volume of samples (100s μL), and take hours to days to obtain the results. Furthermore, these traditional techniques can only reveal the identities of the binding partners, and are not capable of providing the kinetic binding information. More recent techniques, such as fluorescence-based techniques and surface plasmon resonance (SPR) enable both detection and kinetic binding study of biomolecular interactions and complex formation, but have limitations, such as labeling, that may affect or destroy the structure and binding activity of the biomolecules, or in the case of SPR, the need for a solid-phase based technique (not solution-based assay) and the high cost.

PREOMICS EXHIBIT 1002
Page 509 of 1861

Brief Description of the Drawings

[0004] FIG. 1 shows a schematic of a first embodiment of a target molecule detection system.

[0005] FIG. 2 shows a schematic of a second embodiment of a target molecule detection system.

[0006] FIG. 3 shows a schematic of a third embodiment of a target molecule detection system.

[0007] FIG. 4. (A) An illustration of an assay for protein complex detection and binding partner analysis from a cell nuclear extract. The assay was comprised of two steps: in the first step, AuNP immunoprobe for one target protein, here is EGFR, is mixed with a sample solution. The particle size increase caused by the binding of EGFR protein or protein complex to AuNP probes is monitored. After the binding event reaches an end point, the assay solution was divided into multiple portions and to each portion, a polyclonal antibody for each suspected protein binding partner was added to the assay solution. The particle size increase of the assay solution was then monitored and used to identify the binding partner. (B) Kinetic binding assay of EGFR-AuNP probe (or mouse IgG1-AuNP probe was used as a negative isotype control to the EGFR-AuNP probe) with EGFR protein and its complex from a pancreatic cancer cell line, *Panc-1* nuclear extracts. (C) and (D) Protein complex binding partner analysis whereby the polyclonal anti-Stat3, anti-Src or anti-EGFR antibody or the non-specific rabbit IgG (negative control) is added to the assay solution prepared from the anti-EGFR-AuNP probe (C) or non-specific mouse IgG1-AuNP probe (negative control) (D). Data are representative of 4 independent studies.

[0008] FIG. 5 (A) An illustration of gold nanoparticle immunoprobe cluster formation due to their binding with protein complexes or aggregates. (B) From the nonlinear dose-response curve, the concentration at which substantial GAPDH protein aggregation appears is identified. (C) and (D) The particle size distribution curves of the GAPDH assay solution at 25 and 50 μg/mL, respectively.

[0009] FIG. 6 (A) The size distribution curves of assay solutions from a normal prostate, prostate adenocarcinoma, and benign prostate hypertrophy (BPH) tissue lysate sample. Data presented in each graph are intensity distribution curves of an assay solution from ten measurements. (B) The polydispersity index (PdI) of the assay solutions from each sample. The numbers on top of the prostate cancer sample data are the Gleason sum score of the tissues. All data presented here are the average of ten measurements.

PREOMICS EXHIBIT 1002
Page 510 of 1861

[00010] FIG. 7 Mouse serum protein adsorption study and target protein analysis. Data are particle size increases of the assay solution of: (A) undiluted mouse serum adsorbed to AuNPs; (B) IgG assay of 10-fold diluted mouse serum; (B) VEGF assay of 10-fold diluted mouse serum. For each assay, 40 μL of AuNP solution was mixed with 2 μL of serum sample. After certain incubation time, an appropriate amount of anti-mouse IgG or anti-VEGF was added to the assay solution to analyze the level of IgG or VEGF adsorbed to AuNPs.

[00011] FIG. 8 VEGF assay results of 10-fold diluted (A) and undiluted (B) human serum samples from cancer and non-cancerous donors. The non-cancerous donors include normal and BPH patients. 40 μL of AuNP solution was mixed with 2 μL of serum sample. After certain incubation time, 2 μL anti-VEGF solution (0.5 mg/mL) was added. The particle size increase is the size difference measured before and after the addition of antibody. In A, the small legends on some samples such as T1c are the stage of the prostate cancer. Samples with a cancer stage of T2c or above are labeled with a cyan color. The average particle size change of all cancer samples minus the six early stage cancer samples is zero.

Detailed Description

[00012] The present patent application discloses several new methods to detect and analyze biomolecule complexes (BMC) and use of such for medical diagnosis. For the purpose of this application, biomolecule complex is defined as one biomolecule subunit such as proteins, DNAs, RNAs, carbohydrates bound with at least another biomolecule through covalent or non-covalent chemical bonding. An individual biomolecule or biomolecule subunit is defined as a monomer. Furthermore, two types of biomolecule complexes are specifically exemplified here: one is heterogeneous complex made from one biomolecule and at least another different biomolecule, and a second one is a homogenous complex made from at least two same biomolecules or biomolecule subunits. Also for the purpose of this application, each molecule component in the complex is defined as "binding partner". This application mainly uses proteins as an example of biomolecules, however, the methods disclosed here can be applied for any other types of biomolecule complexes such as protein-DNA, protein-RNA, DNA-DNA, DNA-RNA complexes. Naturally, the embodiments disclosed herein may be used to detect biomolecules that are not complexed or aggregated.

[00013] For the purpose of this application, nanoparticles in this application refer to particles with a diameter ranging from 1 nm to 10,000 nm, preferably 1-1000 nm. The new

PREOMICS EXHIBIT 1002
Page 511 of 1861

methods and system embodiments disclosed herein are an extension of the ground breaking work disclosed in PCT/US09/30087; filed January 5, 2009.

[00014] It should be borne in mind that the type of nanoparticle referred in the present patent application is not limited to metal nanoparticles. The particle size measurement technique in this application is not limited to dynamic light scattering. Furthermore in some embodiments as described below, other detection techniques may be used to detect the nanoparticle clustering caused by biomolecular complex interactions.

[00015] Biological samples that can be tested using the method embodiments taught herein include, but are not limited to, are tissue, tissue lysates, tissue fluids, cells, cell lystates, cell culture media, blood, urine, stool, semen, fluid secreted from breast, saliva, sputum, or circulating tumor cells; and/or products derived from the above biological samples after chemical, physical, or biological treatments. Reference to diseased or non-diseased samples herein refers to biological samples in or derived from a subject having a disease state or a non-diseased state, respectively.

[00016] FIGs 1-3 provide schematics of three embodiments for detecting biomolecules. In one embodiment (Figure 1), the invention involves a nanoparticle (preferably gold or silver nanoparticle) that is coated with a layer of "baiting" biomolecule. A "baiting molecule" as used herein means a molecule that can specifically bind with another biomolecule through non-covalent chemical interactions. Antibody is an example of "baiting molecule". Single strand DNA is another example of a "baiting molecule". The baiting molecule will bind with a specific target molecule X in the sample. This baiting molecule-coated nanoparticle is referred to herein as "nanoparticle probe". By contacting the nanoparticle probe solution with a sample solution, a target molecule X and/or its complexes will bind with the nanoparticle probe, causing nanoparticle probe size increase or nanoparticle clustering. Such nanoparticle size change in the assay solution can be measured by dynamic light scattering or other particle size analysis techniques to reveal the presence of target protein or protein complexes in the sample solution. The size measurement can be done continuously to reveal the binding kinetics as illustrated in the graph in Figure 1. The binding constant or binding energy between the baiting molecule and the target molecule or complex can be obtained by fitting the binding curve into a Langmuir adsorption model using Langmuir equation. When the binding reaches a saturated level, the size of the target molecule X and/or its complex may be revealed from the average particle size change or particle size distribution change. The particle size change can be used to obtain the concentration of molecule X in the sample solution. Furthermore, by designing the assay so that a saturated binding between molecule X

PREOMICS EXHIBIT 1002
Page 512 of 1861

and the nanoparticle probe is achieved, the net increase of the particle size at saturated binding should be equal to twice of the diameter (2D) of the target molecule X. From this analysis, the size of molecule X can be obtained. If target molecule X formed complexes with other biomolecules, then the nanoparticle size increase will exceed the 2D value of X. In this case, the obtained size information of molecule X can be used to deduce the complex status and level of molecule X.

[00017] As a variation to this embodiment, the same nanoparticle could contain more than one type of baiting molecules for different target molecules. For example, baiting molecule for X and A can be both immobilized on the nanoparticle and such nanoparticle can bind with more than one type of target molecule from sample solution.

[00018] Furthermore, by adding a second step to the assay, the binding partner of the biomolecular complex can be further revealed through a screening analysis. After the first step of the assay, a solution of baiting molecule that can specifically bind with an interested binding partner molecule A will be added to the assay solution. The nanoparticles may or may not need to be isolated from the rest of the assay solution before conducting the second step of the assay. If the interested potential binding partner A is present on the nanoparticles, the binding of the baiting molecule for target molecule A to the nanoparticle will cause a further increase of the nanoparticle size or nanoparticle cluster formation, therefore, confirming that molecule A and molecule X formed a biomolecular complex.

[00019] Alternatively, the method can include the successive exposure of the assay solution to baiting molecule for target molecule A without conducting particle analysis after exposure to the nanoparticle probe with the biological sample. Based on the particle characteristics, one can determine whether baiting molecule for target molecule A binds to the nanoparticle probe presumably already having target molecule X bound thereto. From this, it can be deduced that the biomolecule complex includes target molecule X complexed to target molecule A.

[00020] In a second embodiment (Figure 2), the assay is designed so that same or different nanoparticle probes will bind with a biomolecule complex from different sites of the complex to form nanoparticle clusters. In the illustration, two nanoparticle probes for two biomolecules X and A are shown. For illustration purpose, the complex shown here is a two-molecule complex, but the complex can include as many molecules as possible. The two nanoparticle probes may be mixed simultaneously or sequentially with the sample solution. When used simultaneously, the assay will lead to nanoparticle dimer or cluster formation if a biomolecule complex XA is present in the sample solution. The average particle size and/or

PREOMICS EXHIBIT 1002
Page 513 of 1861

particle size distribution will provide quantitative information on the complexing level of XA. Complexing level can mean both concentration of the complex XA and the size of the complex. For example, the presence of large biomolecule complex will lead to significant broadening of the particle size distribution curve. When used sequentially, either nanoparticle probe may be first mixed with sample solution to allow binding of the first biomolecule from the complex to the nanoparticle probe. Then a second probe is added to detect the second biomolecule in the complex. More nanoparticle probes may be added to the assay solution to detect additional binding partners in the complex.

[00021] As a special case of second embodiment, target molecule X and A are the same. In this case, biomolecular complex XA may also be called biomolecular aggregate (or may also be referred to as a homogenous complex). A larger aggregate can lead to larger particle size increase and/or broadened particle size distribution curve after the sample solution is in contact with at least one nanoparticle probe solution.

[00022] Reference to particle analysis may include, but is not limited to, determining a particle size change (including an individual particle size change), an average particle size change, particle size distribution change, polydispersity change of the size distribution, or measurement-to-measurement particle size variation, or combinations thereof. In a specific embodiment, particle analysis is conducted via DLS.

[00023] In a third embodiment, a nanoparticle without specific baiting molecule attached on the surface is mixed with a sample solution. Proteins and/or other biomolecules from the sample solution will be non-specifically adsorbed to the nanoparticle to form a biomolecule "corona". Gold or silver nanoparticles possess a charge that will allow them to non-specifically adsorb biomolecules. The size of this biomolecule corona, may be determined using dynamic light scattering or other suitable particle size analysis techniques. The size of this corona may be used to analyze a biological process, determine the disease status or reveal other physiological conditions of the sample donor. Furthermore, a second step assay may be conducted to reveal and quantify the individual biomolecules that are present in the biomolecule corona. If a target molecule X is present in the corona, the addition of an antibody or baiting molecule for molecule X to the assay solution will further increase the size of the particle, or cause particle cluster formation. Such particle size change can be correlated to the concentration or the size of molecule X in the sample solution. The concentration and/or size information of X can be used to analyze a biological process, determine the disease status or reveal other physiological conditions of the sample donor. Alternatively, if a target molecule D is not adsorbed to the nanoparticle, then the addition of a

PREOMICS EXHIBIT 1002
Page 514 of 1861

baiting molecule for D in the assay solution will not cause significant change of the nanoparticle size or formation of nanoparticle clusters.

[00024] **Examples**

[00025] In the following, three examples are given for each of the embodiments as described above. Additionally, a fourth example is provided to describe how the biomolecular complexes discovered from our research can be used for diagnostic applications.

[00026] *1. Detection of a heterocomplex EGFR/Stat3/Src from a pancreatic cancer cell line, Panc-1 cells*

[00027] Using the first embodiment (Figure 1), we recently successfully studied the assembly of signaling molecules into complexes that promote signal transduction. In particular, the new invention permitted the detection of hitherto unknown heterocomplex formed from three proteins: epidermal growth factor receptor (EGFR), Stat3, and Src in the nuclei of a pancreatic cancer cell line, *Panc-1*. EGFR is an important regulator of cellular responses, including growth and survival. The attention from biomolecular researchers on EGFR protein has been enormous in part because there remains a significant gap in the understanding of the key molecular partners that serve this pathway and how they contribute to the overall associated cellular responses. From the molecular and biological perspectives, EGFR is involved in a large number of sophisticated events that support cellular processes. EGFR is found in the plasma membrane, cytosol and nuclei of cells. A full understanding of all aspects of the EGFR protein, not only its expression level, but also its activity regulation, translocation, structural changes and the interactions with other proteins is essential for developing an effective model relevant to the functional role of this pathway in the overall biological phenotype. A study by J. Turkson et al. revealed that concurrent targeting of three proteins, EGFR, Stat3 and Src can enhance the therapeutic effect on pancreatic cancer. It was hypothesized based on their studies that EGFR, Stat3 and Src may have formed a heterocomplex in *Panc-1* nuclei and this complex functions as a transcript factor to regulate the cell growth of *Panc-1*. An assay based on the first embodiment confirmed the existence of such a complex in *Panc-1* cell nuclei.

[00028] A mouse monoclonal antibody, anti-EGFR, was conjugated to AuNP to form EGFR-AuNP probe. Upon mixing this probe solution with *Panc-1* nuclear lysate, the particle size increased rapidly from ca. 60 nm to 140 nm, while a non-specific isotype control mouse IgG1-conjugated to AuNP exhibits only about 10 nm of size increase after the assay (Figure 4B). The rapid increase of the nanoparticle probe size corresponds to specific binding of

PREOMICS EXHIBIT 1002
Page 515 of 1861

target protein EGFR to the probe, while the slow and small size increase of mouse IgG1-AuNP probe is indicative of non-specific interactions.

[00029] The net increase of the anti-EGFR-AuNP probes after deducting the size increase caused by non-specific interactions, is about 70 nm. This suggests that the size of the EGFR protein complex detected by the anti-EGFR-AuNP probe is approximately 35 nm. EGFR protein has a molecular mass of 170 KDa. The size of EGFR protein alone should be between 10-15 nm. Therefore, data shown in Figure 4B suggests that the EGFR detected by the nanoparticle probe is most likely in a complex form. In a second step assay, when a polyclonal anti-Stat3 or anti-Src was added to the assay solution, the nanoparticle size increased further by 30-35 nm; while the addition of isotype control rabbit IgG and anti-EGFR caused much less nanoparticle size increase (Figure 4C). The substantial size increase from the addition of anti-Stat3 and anti-Src suggests that Stat3 and Src are binding partners in the EGFR complex. The small size increase of the assay solution upon addition of anti-EGFR probe is due to the fact that Stat3 and Src in the complex blocked the binding sites of anti-EGFR to the nanoparticle-bound EGFR protein. The EGFR/Stat3/Src complex was further confirmed by co-immunoprecipitation followed by immunoblotting and laser scanning confocal microscopic analysis.

[00030] Co-immunoprecipitation is considered as the gold standard for protein complex detection from biological samples. However such analysis requires a substantial amount of samples (100s μL) and the analysis takes hours to days to complete. SPR technique also requires 10s-100s μL of sample per analysis. As a comparison, our new assay requires 1-2 μL of lysate samples and results are obtained in minutes. More importantly, neither co-immunoprecipitation, nor fluorescence technique, nor SPR is able to reveal the size information of the EGFR-Stat3-Src complex. Such information is particularly useful for protein complex analysis: it gives the first evidence of the presence of a protein complex, and can further reveal how many binding partners are involved in a protein complex. According to the assay results obtained so far, we believe the complex involves EGFR, Stat3 and Src only, because the estimated hydrodynamic diameter of such a triple-component complex according to each protein's size is around 35 nm, corresponds to the observed size from the assay.

[00031] *2. Protein aggregate detection and quantitative analysis*

[00032] Protein aggregate formation is a significant and challenging problem in biopharmaceutical development. To use proteins as therapeutic agents, they need to be formulated into high concentration solutions. Unfortunately, many proteins tend to aggregate

PREOMICS EXHIBIT 1002
Page 516 of 1861

at high concentrations. Once protein is aggregated, it can cause significant adverse effects by changing the pharmacokinetics of the drug or inducing unexpected immunogenicity. Protein aggregation is also closely associated to many diseases such as prion protein aggregates in Creutzfeldt-Jakob disease (CJD) and amyloid β–protein aggregates found in Alzheimer's disease. However, techniques that are available for direct detection and analysis of protein aggregation are extremely limited, especially for the detection of protein aggregates from complex biological samples and fluids in the presence of non-target proteins, biomolecules, and other colloidal particles. One widely used approach is by fluorescence techniques. Fluorescent probes are conjugated to the target protein by covalent bonding. Upon protein aggregation, fluorescent probe molecules will also aggregate, causing fluorescence signal change. A significant drawback of the fluorescence techniques is the need to do a fluorescence labeling of the target molecule. Other commonly used techniques for protein aggregate detection include size exclusion chromatography (SEC) and analytical ultracentrifugation. These techniques are only suitable for pure protein solution study, not suitable for detection of protein aggregates and complexes directly from real biological samples.

[00033] We have demonstrated that a new assay as illustrated in the second embodiment (Figure 2) can be used as a label-free technique for highly sensitive detection of protein aggregates. In Rayleigh scattering, the scattered light intensity is proportional to the sixth power of the radius ($I \propto R^6$) of a particle. When protein complexes or aggregates with a non-uniform size distribution are present in a sample solution, these complexes/aggregates will lead to the formation of large nanoparticle clusters as shown in Figure 5A. The scattered light intensity from the large AuNP clusters is thousands to millions times stronger than individual AuNPs (the detection limit for individual AuNPs by DLS can already reach pM-aM range). Protein aggregation can be detected with extremely high sensitivity through the amplification of AuNP probes. Results shown in Figure 5B-D are from our study on an enzyme GAPDH (glyceraldehyde 3-phosphate dehydrogenase, a protein used routinely as a loading control in Western Blot analysis). From the assay, we discovered that this protein has a strong tendency to form large aggregates at a concentration around 10-25 μg/mL. The protein aggregate formation leads to a non-linear and abnormal nanoparticle size increase around this concentration (Figure 5B), and significant broadening or even peak splitting of the size distribution curves (Compare Figure 5D with Figure 5C).

PREOMICS EXHIBIT 1002
Page 517 of 1861

[00034] Furthermore, by using this new assay, we discovered that a cancer biomarker protein, prostatic acid phosphatase (PAP) exists more in the form of biomolecular complexes in cancer tissue than in normal and benign prostate hyperplasia (BPH) tissue. For this assay, an anti-PAP antibody is conjugated to AuNPs to make a nanoparticle immunoprobe. As shown in Figure 6A, the particle size distribution curves from prostate cancer tissue lysate assay solution are substantially more broadened than BPH samples, and a significant amount of large nanoparticle clusters appeared at 100s nm size range. Such broad particle size distribution and the appearance of large nanoparticle clusters are explained by the presence of PAP protein in the form of large biomolecular complexes or aggregates. The level of particle size distribution curve broadening can be quantitatively expressed using a *polydispersity index* (PdI). As shown in Figure 6B, the PdI of assay solutions obtained from prostate cancer tissue lysates is clearly higher than the assay solutions of normal prostate and BPH tissue lysates. Interestingly, BPH tissue lysate assay solutions exhibit a narrower particle size distribution than normal tissue lysates, indicating that a different molecular mechanism is involved in BPH than in prostate cancer. Furthermore, it appears that more advanced prostate cancer with higher Gleason scores leads to a higher complexing level of PAP protein. To our best knowledge, this is the first time that such molecular profile differences were revealed from biological samples and the potential link of such differences with cancer status of a biological system was discovered.

[00035] *3. Serum protein and protein complex detection and profiling for diagnosis*

[00036] Using the third embodiment as illustrated in Figure 3, we discovered some very interesting and intriguing molecular differences between serum samples with and without prostate cancer from both mice models and human donors. Prostate cancer is one of the most prevalent types of cancer in men. It is estimated that 217,730 men will be diagnosed with and 32,050 men will die of prostate cancer in 2010. Despite the enormous research efforts and investigations, early detection and diagnosis of prostate cancer has remained a significant challenge in medicine due to multiple complicated factors. PSA test, a test that measures the level of prostate specific antigen in blood, has been widely used in the United States in the last two decades for initial screening of prostate cancer. However, several recent studies have led to significant doubts on the efficacy of PSA test. Over-diagnosis and treatment of low-risk prostate cancer has serious and long-lasting side effect: 33% of the patients who receive radical prostatectomy treatment will suffer erectile dysfunction that cannot be remedied by drugs such as Viagra. New biomarkers and test methods that can reliably distinguish prostate

PREOMICS EXHIBIT 1002
Page 518 of 1861

cancer from benign conditions and slow growing tumor from high risk cancer are of tremendous value in prostate cancer diagnosis and treatment.

[00037] Three mice models were used for this study: one was orthotopically injected with a fast growing prostate cancer cell line PC3; one with a slow growing tumor cell line LnCaP; and a third group of mice injected only with PBS saline solution as control. Figure 7 are the analysis results of mice serum samples. First, there is a dramatic difference in the "size" of the proteins or protein complexes adsorbed from serum to AuNPs (Fig. 7A). The particle sizes from mice with large and fast growing tumor from PC3 cells are substantially smaller than that from normal healthy mice and mice with small and slow growing tumor from LnCaP cells. When the serum samples were diluted 10-fold and the serum protein adsorption assay was repeated, all samples exhibited similar particle size increase of 15-18 nm after the assay.

[00038] The inventor then analyzed two target proteins: one is an abundant protein, mouse IgG, and another one is a well known cancer biomarker protein, VEGF (vascular endothelial growth factor) according to the assay process as illustrated in Figure 3. Again, some interesting differences were observed from these two assays. In the IgG assay, mice with small tumor grown from LnCaP cells exhibit higher level of IgG than the other two groups (Fig. 7B). In the VEGF assay, the normal healthy mice exhibit significantly higher level of AuNP-adsorbed VEGF than the mice with prostate cancer (Fig. 7C). Intrigued by the results observed from mice model study, we examined human serum samples obtained from subjects with prostate cancer. Three groups of human serum samples were studied: normal healthy donors; patients diagnosed with benign prostate hyperplasia (BPH); and patients diagnosed with prostate cancer (stages T1c to T3b). From the VEGF assay, same difference was observed from the cancer versus non-cancer samples (normal donor and BPH patients) (Figure 8): the non-cancer samples show higher level of AuNP-adsorbed VEGF level than the prostate cancer samples. T-test analysis of the assay data gave a p-value of 0.001 for both diluted and undiluted human serum samples, indicating that the difference between cancer and non-cancerous samples was statistically significant. Furthermore, it was found from the assay that the AuNP-adsorbed VEGF level is also cancer stage-dependent: the more advanced prostate cancer exhibits lower VEGF level than the early stage prostate cancer. Early stage prostate cancer such as T1c shows similar VEGF level as healthy donors. In conclusion, the analysis of both mice model and human donor samples revealed that the amount of AuNP-adsorbed VEGF is decreased in cancer serum samples. This discovery could potentially lead to a new blood test with improved accuracy for prostate cancer detection. Furthermore, by

PREOMICS EXHIBIT 1002
Page 519 of 1861

screening other proteins or biomolecular targets using this assay, it is expected that additional protein or biomolecular biomarkers that are unique to cancer and other human diseases can be discovered. This new assay established in our recent work may be used as a general tool for serum protein biomarker discovery.

[00039] *4. The use of biomolecular complexes as biomarkers for diagnostic applications*

[00040] There are several types of biomarkers that have been identified so far and can be used for prostate cancer detection and diagnosis, particularly, for distinguishing prostate cancer from non-malignant conditions such as BPH: (1) The increased complexing level of PAP protein in tissue and bodily fluids; (2) The decreased level of VEGF protein in blood samples. This VEGF protein is referred to VEGF protein that is adsorbed to a nanoparticle through complexing with other proteins or biomolecules present in the blood serum; (3) The size of the biomolecule corona formed on a nanoparticle surface after non-specific adsorption of biomolecules from a biological sample to the nanoparticles.

[00041] The claims for biomarkers may be extended to the detection and diagnosis of other diseases and conditions: (1) The size of the biomolecule corona formed on a nanoparticle surface after a non-specific adsorption of biomolecules from a biological sample to the nanoparticles; (2) The size of the biomolecule corona formed on a nanoparticle surface after a specific binding of biomolecules from a biological sample to the baiting molecule-conjugated nanoparticles; (3) Individual molecule component or biomolecule complexes of the biomolecule corona formed on a nanoparticle surface after a non-specific adsorption or specific binding of biomolecules from a biological sample to the nanoparticles that show significant difference between diseased and non-diseased samples.

PREOMICS EXHIBIT 1002
Page 520 of 1861

[00042] References

1. Liu, X.; Dai, Q.; Austin, L.; Coutts, J.; Knowles, G.; Zou, J.; Chen, H.; Huo, Q. A One-step homogeneous immunoassay for cancer biomarker detection using gold nanoparticle probes coupled with dynamic light scattering. J. Am. Chem. Soc. 2008, 130, 2780-2782.
2. Dai, Q.; Liu, X.; Coutts, J.; Austin, L.; Huo, Q. A one-step highly sensitive method for DNA detection using dynamic light scattering. J. Am. Chem. Soc. 2008, 130, 8138-8139.
3. Liu, X.; Huo, Q. A washing-free and amplification-free one-step homogeneous assay for protein detection using gold nanoparticle probes and dynamic light scattering. J. Immunol. Method 2009, 349, 38-44
4. Jans, H.; Liu, X.; Austin, L.; Maes, G.; Huo, Q. Dynamic light scattering as a powerful tool for gold nanoparticle bioconjugation and biomolecular binding study. Anal. Chem. 2009, 81, 9425-9432.
5. Austin, L.; Liu, X.; Huo, Q. An immunoassay for monoclonal antibody isotyping and quality analysis using gold nanoparticles and dynamic light scattering. 2010, American Biotechnology Laboratory, 2010, 22, No. 3, 8-12.
6. Bogdanovic, J.; Huo, Q. NanoDLSay: a new platform technology for biomolecular detection and analysis using gold nanoparticle probes coupled with dynamic light scattering. SPIE Proceedings, (2010), 7674 (Smart Biomedical and Physiological Sensor Technology), 767408/1-767408/9.
7. Bogdanovic, J.; Colon, J.; Baker, C.; Huo, Q. A label-free nanoparticle aggregation assay for protein complex/aggregate detection and analysis. Anal. Biochem. 2010, 405, 96-102.
8. Huo, Q. Protein complexes/aggregates as potential cancer biomarker revealed by a nanoparticle aggregation immunoassay. Colloids Surf. B. 2010, 78, 259-265.
9. Jaganathan, S.; Yue, P.; Bogdanovic, J.; Huo, Q.; Turkson, J. A functional nuclear epidermal growth factor receptor, Src and Stat3 heteromeric complex in pancreatic cancer cells. *Plos* One, 2010, submitted.
10. Huo, Q.; Cordero, A.; Bogdanovic, J.; Colon, J.; Baker, C.H.; Goodison, S.; Pensky, M. A facile nanoparticle immunoassay to detect multiple biomarkers in serum samples. Nanomedicine: Nanotechnology, Biology and Medicine, 2010, submitted.

[00043] It should be borne in mind that all patents, patent applications, patent publications, technical publications, scientific publications, and other references referenced herein are hereby incorporated by reference in their entirety to the extent not inconsistent with the teachings herein.

[00044] Reference to particular buffers, media, reagents, cells, culture conditions and the like, or to some subclass of same, is not intended to be limiting, but should be read to include all such related materials that one of ordinary skill in the art would recognize as being of interest or value in the particular context in which that discussion is presented. For example,

PREOMICS EXHIBIT 1002
Page 521 of 1861

it is often possible to substitute one buffer system or culture medium for another, such that a different but known way is used to achieve the same goals as those to which the use of a suggested method, material or composition is directed.

[00045] It is important to an understanding of the present invention to note that all technical and scientific terms used herein, unless defined herein, are intended to have the same meaning as commonly understood by one of ordinary skill in the art. The techniques employed herein are also those that are known to one of ordinary skill in the art, unless stated otherwise. For purposes of more clearly facilitating an understanding the invention as disclosed and claimed herein, the following definitions are provided.

[00046] While a number of embodiments of the present invention have been shown and described herein in the present context, such embodiments are provided by way of example only, and not of limitation. Numerous variations, changes and substitutions will occur to those of skilled in the art without materially departing from the invention herein. For example, the present invention need not be limited to best mode disclosed herein, since other applications can equally benefit from the teachings of the present invention. Also, in the claims, means-plus-function and step-plus-function clauses are intended to cover the structures and acts, respectively, described herein as performing the recited function and not only structural equivalents or act equivalents, but also equivalent structures or equivalent acts, respectively. Accordingly, all such modifications are intended to be included within the scope of this invention as defined in the following claims, in accordance with relevant law as to their interpretation.

PREOMICS EXHIBIT 1002
Page 522 of 1861

Claims

What is claimed is:

1. A method for detecting a target molecule in a biological sample, said method comprising obtaining a first assay product that has been produced by contacting a biological sample with a first nanoparticle probe, said first nanoparticle probe comprising a nanoparticle bound to a first baiting molecule specific to a first target molecule of interest; and detecting for presence of nanoparticle probe bound to said first target molecule of interest.

2. The method of claim 1, wherein said detecting comprises subjecting said first assay product to particle analysis.

3 The method of claim 2, wherein particle analysis comprises determining a particle size change, an average particle size change, particle size distribution change, polydispersity change of the size distribution, or measurement-to-measurement particle size variation, or combinations thereof.

4. The method of claim 2, wherein said particle analysis is by dynamic light scattering analysis

5. The method of claim 2, wherein the particle analysis comprising determining a particle size change over nanoparticle probe-sample incubation time or over first target molecule concentration is used to obtain the binding constant or binding energy between the first baiting molecule and the first target molecule

6. The method of claim 5, wherein the binding constant or binding energy is obtained by fitting the binding data into a Langmuir adsorption model.

7. The method of claim 2, further comprising determining whether a first target molecule is a monomer or biomolecule complex based on particle analysis..

8. The method of claim 7, further comprising determining the quantity of a biomolecule complex, quantity being absolute or relative quantity based on particle analysis; or determining the size of a biomoleucle complex based on particle analysis.

PREOMICS EXHIBIT 1002
Page 523 of 1861

9. The method of claim 1, wherein said target molecule is a biomolecule complex

10. The method of claim 8, further comprising utilizing a quantity of a biomolecule complex to analyze a biological process, or for detection, diagnosis and/or prognosis of a disease; or utilizing the size of a biomolecule complex to analyze a biological process, or for detection, diagnosis and/or prognosis of a disease.

11. The method of claim 7, wherein the biomolecule complex is a complex of PAP (prostatic acid phosphatase).

12. The method of claim 10, wherein the disease is cancer.

13. The method of claim 9, wherein said biomolecule complex comprises at least one first biomolecule or first biomolecule subunit associated with at least one other biomolecule.

14. The method of claim 13, wherein said at least one other biomolecule is the same as the first biomolecule or first biomolecule subunit.

15. The method of claim 1, further comprising contacting the first assay product with a second baiting molecule specific to a second target molecule of interest to produce a second assay product, and detecting the presence of said second target molecule of interest in said second assay product.

16. The method of claim 15, wherein said detecting comprises subjecting said second assay product to particle analysis.

17. A method for detecting a target molecule in a biological sample, said method comprising
(a) obtaining a first assay product that has been produced by contacting a biological sample with a first nanoparticle probe, said first nanoparticle probe comprising a nanoparticle bound to a first baiting molecule specific to a first target molecule of interest;
(b) subjecting said first assay product to a second nanoparticle probe comprising a second baiting molecule that is specific to a second target molecule of interest to produce a second

PREOMICS EXHIBIT 1002
Page 524 of 1861

assay product; and (c) detecting for presence of second nanoparticle probe bound to said second target molecule of interest in said second assay product; and

18. The method of claim 17, further comprising detecting presence of first nanoparticle probe bound to said first target molecule of interest in said first assay product before step (c).

19. A method for detecting at least one biomolecule in a biological sample, said method comprising obtaining a first assay product produced by contacting a biological sample with a nanoparticle, wherein said contacting forms a biomolecule corona on the nanoparticle surface, and conducting particle analysis of the first assay product; and further comprising (i) using biomolecule corona size information for detection, diagnosis and/or prognosis of a disease or (ii) contacting said first assay product with a baiting molecule that binds specifically to a target molecule of interest to form a second assay product, and detecting the presence and/or quantity of the target molecule of interest by particle analysis.

20. The method of claim 19, wherein said analysis comprises size analysis of the biomolecule corona.

21. The method of claim 19, wherein biomolecule corona size for a diseased biological sample is different than that for a non-diseased biological sample

22 The method of claim 19, further comprising correlating particle analysis information with a disease state.

23. The method of claim 19, wherein particle analysis detects a size change of the particle due to the binding of a baiting molecule to the target molecule of interest in the biomolecule corona.

24. The method of claim 19, further comprising correlating the presence and/or quantity of a target molecule of interest disease state, and optionally using said correlation for detection, diagnosis and/or prognosis of a disease.

PREOMICS EXHIBIT 1002
Page 525 of 1861

25. The method of claim 19, wherein said disease is cancer

26. The method of claim 19, wherein said target molecule is VEGF (vascular endothelial growth factor)

27. The method of claim 17, wherein said first baiting molecule and said second baiting molecule are the same.

PREOMICS EXHIBIT 1002
Page 526 of 1861

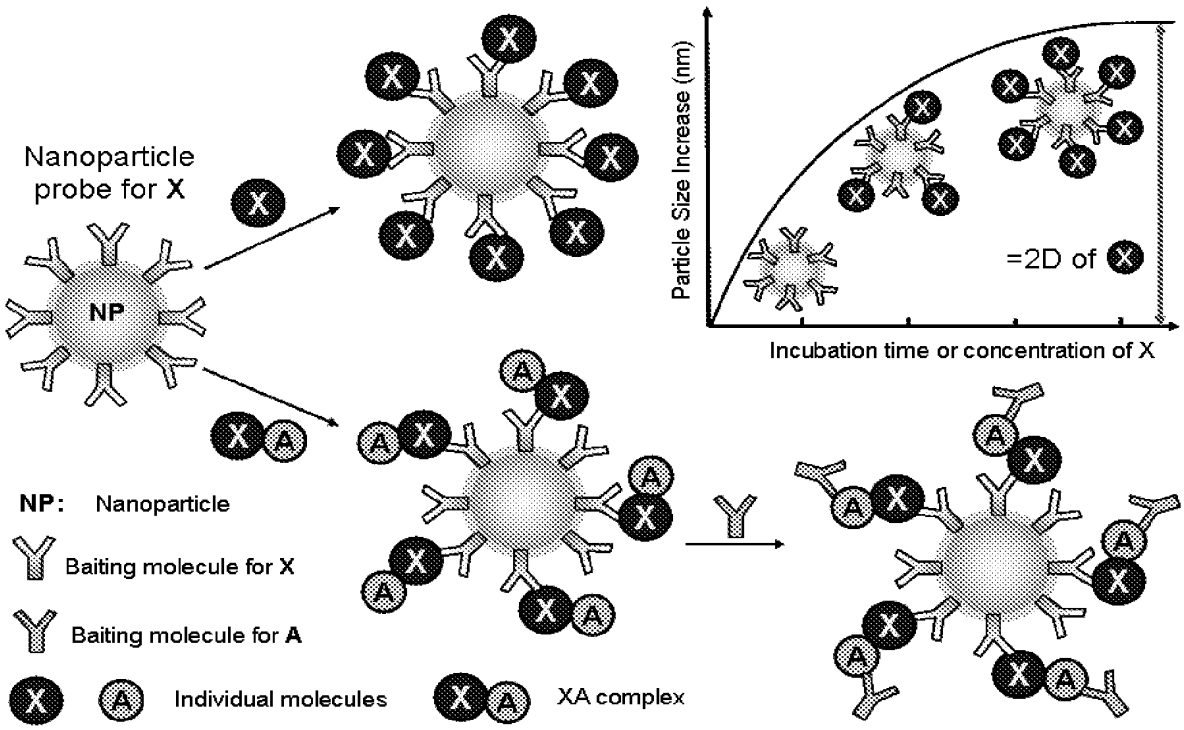


FIG.1

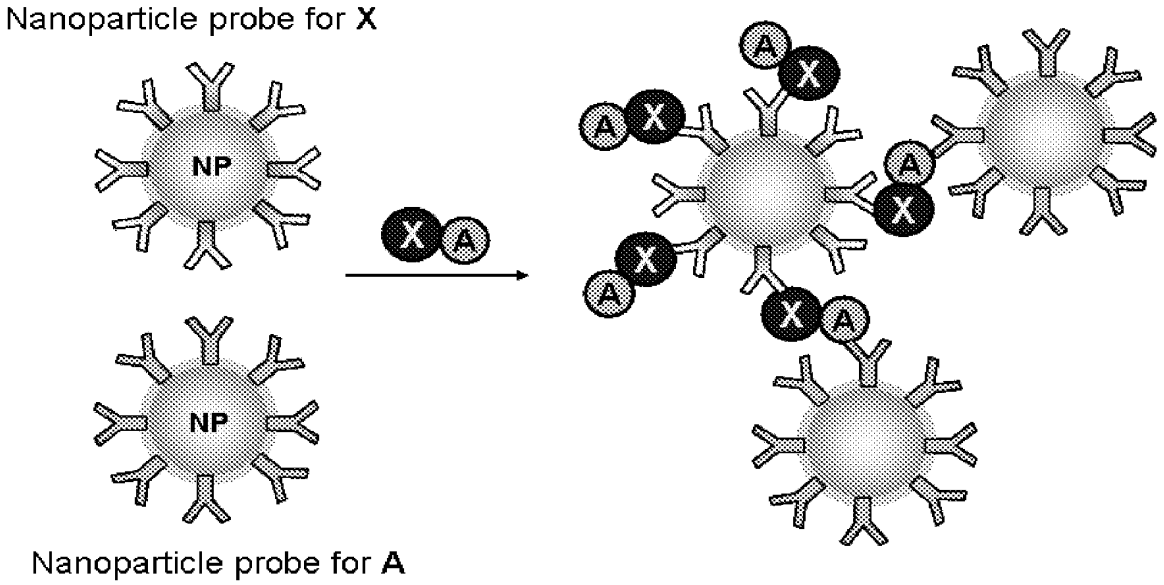


FIG. 2

PREOMICS EXHIBIT 1002
Page 527 of 1861

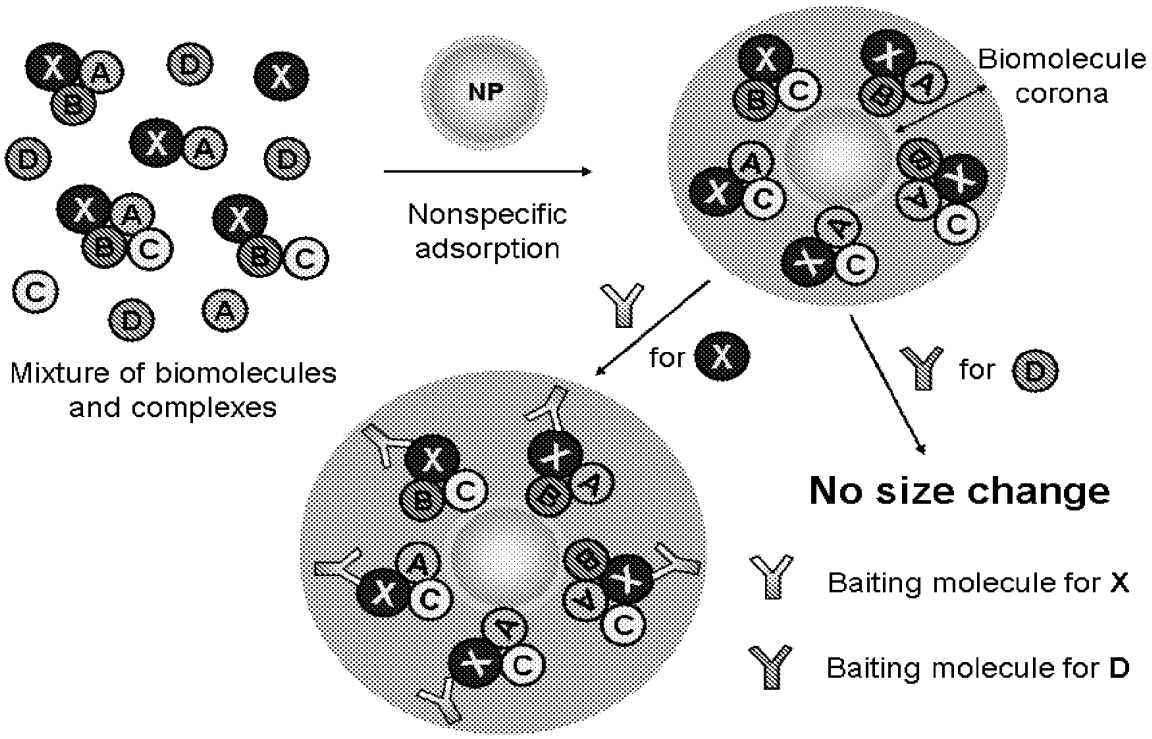

NP
Nonspecific adsorption
Biomolecule corona
Mixture of biomolecules and complexes
for X
for D
No size change
Baiting molecule for X
Baiting molecule for D

FIG. 3

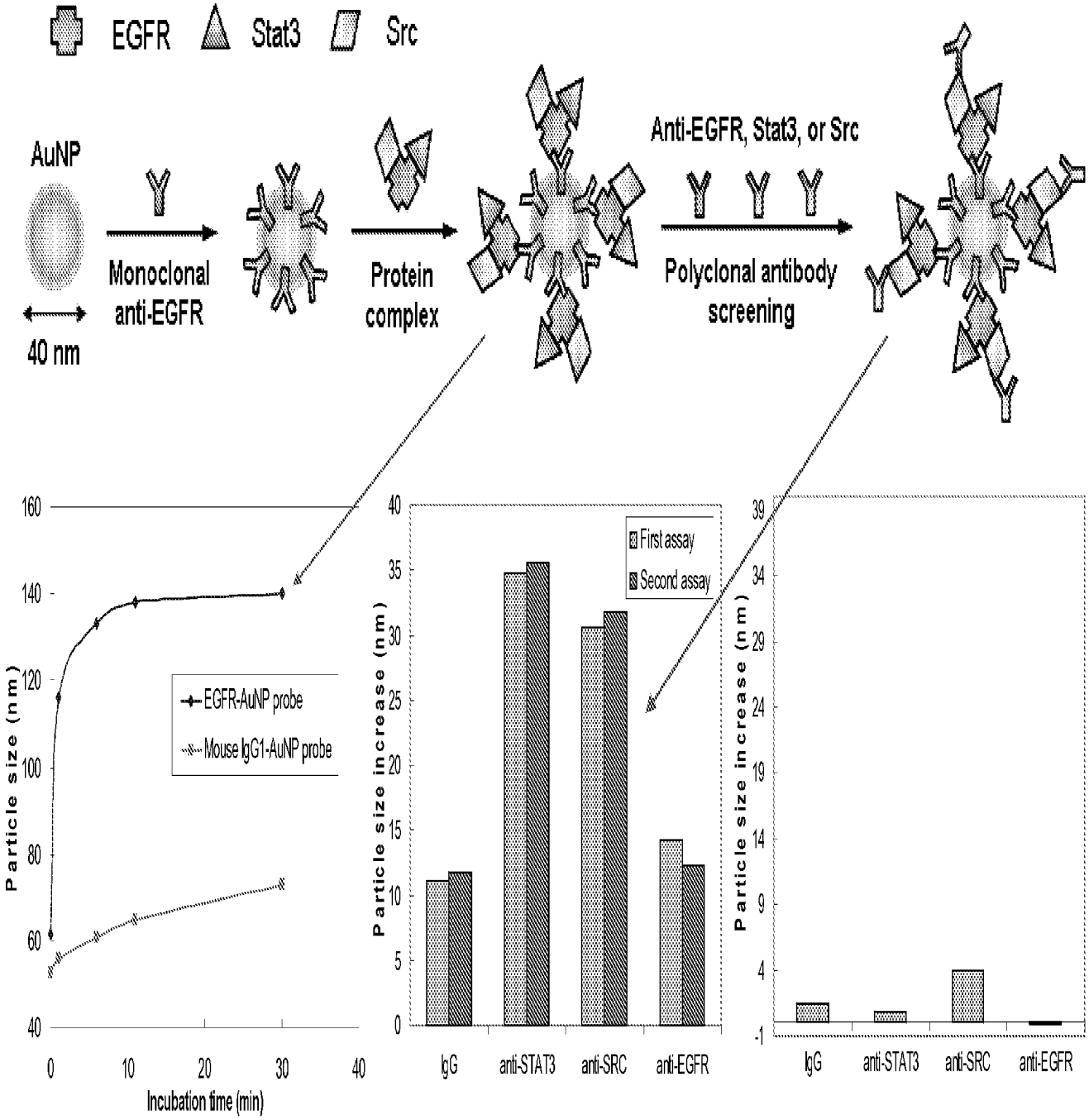
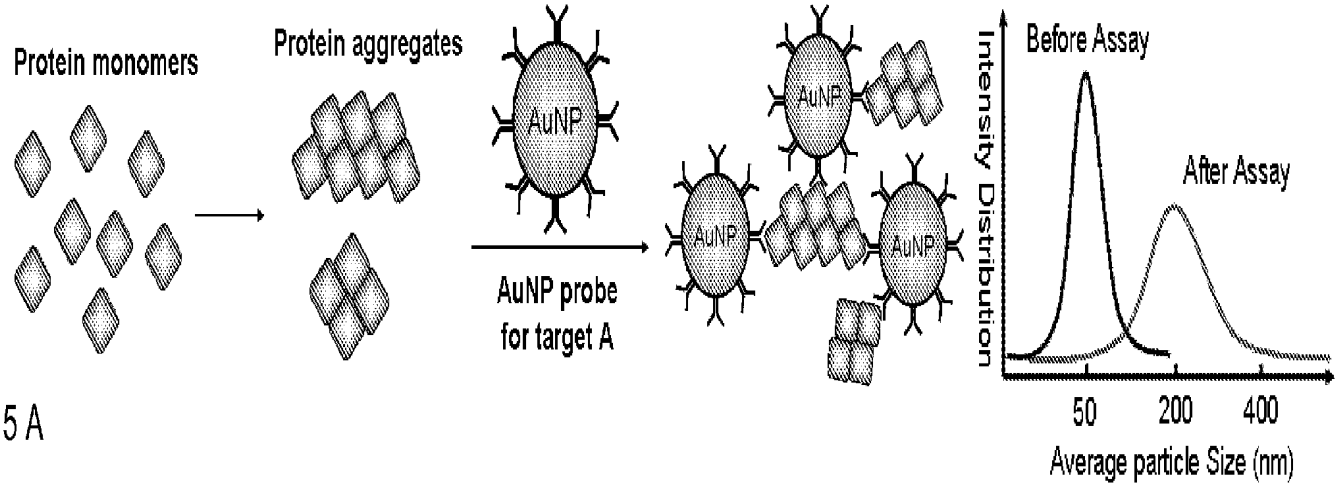
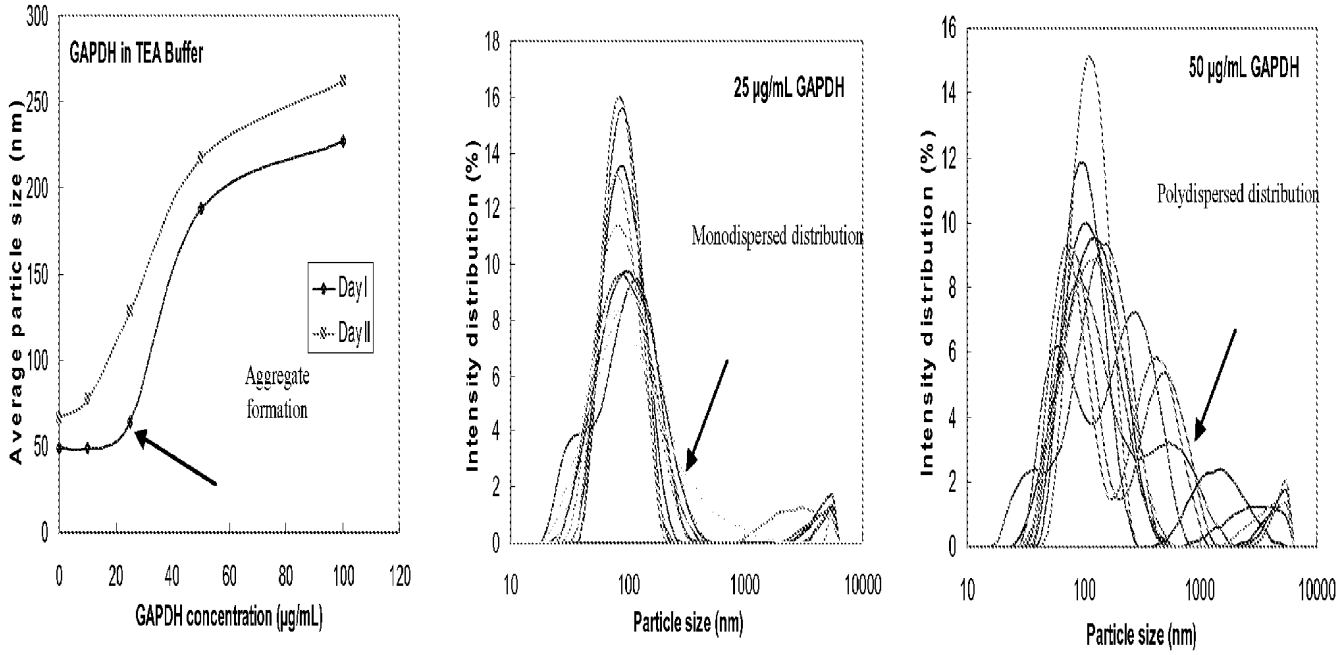


FIG. 4

PREOMICS EXHIBIT 1002
Page 529 of 1861


Protein monomers
Protein aggregates
AuNP
AuNP probe
for target A
Before Assay
After Assay
Intensity Distribution
50
200
400
Average particle Size (nm)

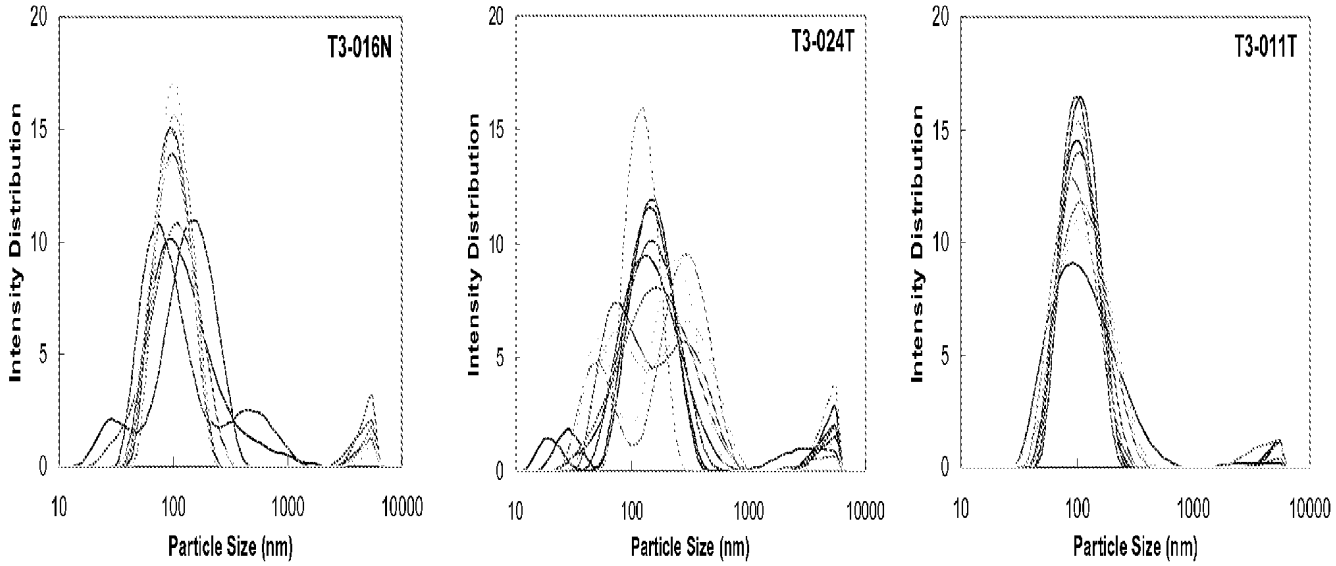
5 A


GAPDH in TEA Buffer
Average particle size (nm)
Day I
Day II
Aggregate formation
GAPDH concentration (μg/mL)
25 μg/mL GAPDH
Monodispersed distribution
Intensity distribution (%)
Particle size (nm)
50 μg/mL GAPDH
Polydispersed distribution
Intensity distribution (%)
Particle size (nm)

5 B

FIG. 5

6 A


T3-016N
T3-024T
T3-011T
Intensity Distribution
Particle Size (nm)
0
5
10
15
20
10
100
1000
10000

6 B

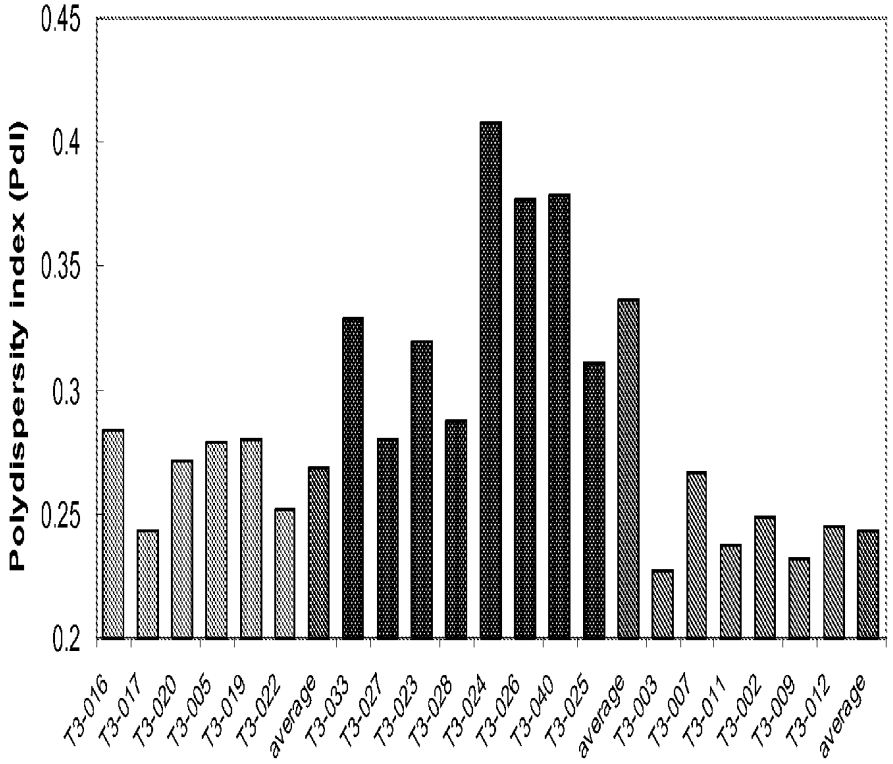
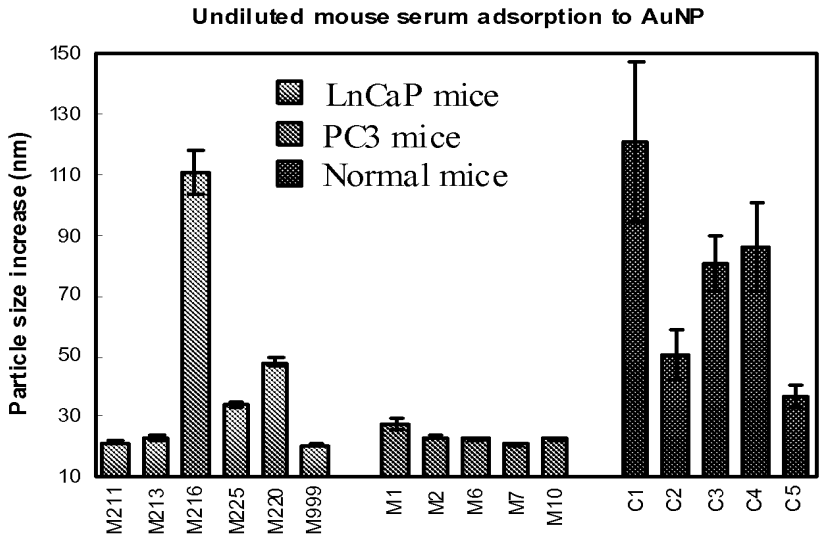
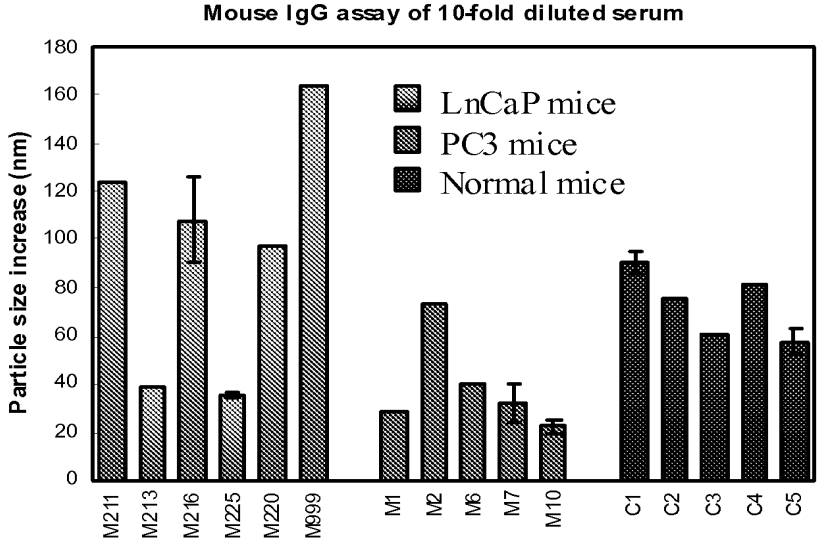

Polydispersity index (PdI)
0.2
0.25
0.3
0.35
0.4
0.45
T3-016
T3-017
T3-020
T3-005
T3-019
T3-022
average
T3-033
T3-027
T3-023
T3-028
T3-024
T3-026
T3-040
T3-025
average
T3-003
T3-007
T3-011
T3-002
T3-009
T3-012
average

FIG. 6

7 A



7 B



7 C

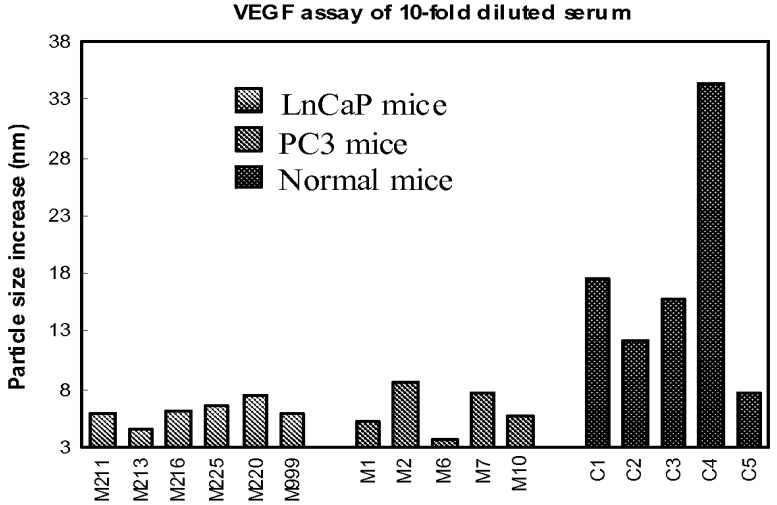
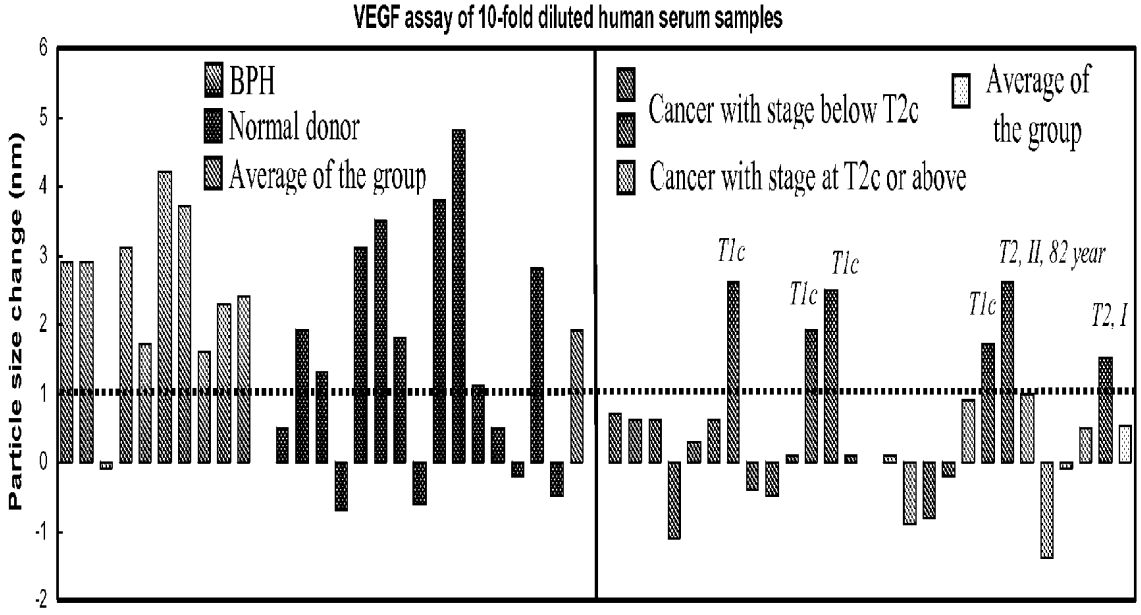


FIG. 7

PREOMICS EXHIBIT 1002
Page 532 of 1861

8 A



8 B

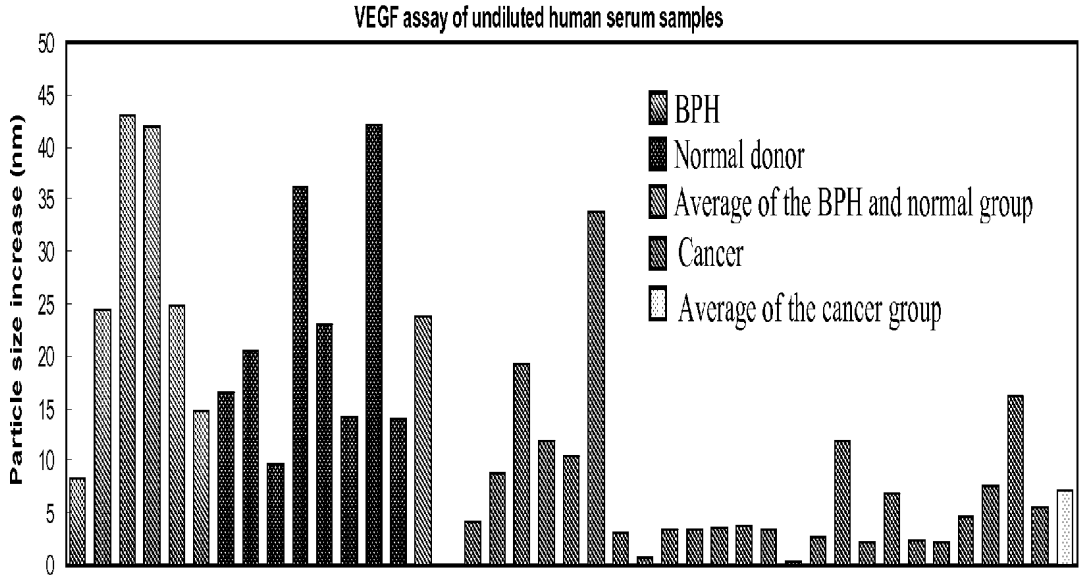


FIG. 8

PREOMICS EXHIBIT 1002
Page 533 of 1861

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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
24 May 2012 (24.05.2012)

WIPO | PCT

(10) International Publication Number
WO 2012/068226 A2

(51) International Patent Classification:
G01N 33/68 (2006.01) *G01N 21/65* (2006.01)
G01N 33/569 (2006.01) *B82Y 15/00* (2011.01)

(21) International Application Number:
PCT/US2011/060949

(22) International Filing Date:
16 November 2011 (16.11.2011)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
12/949,052 18 November 2010 (18.11.2010) US

(71) Applicants *(for all designated States except US)*: **ACADEMIA SINICA**; 128, Sec. 2, Academia Road, Nan-kang, Taipei (TW). **LIANG, Chi-Ming** [US/—]; 4th Fl., No. 8, Alley 2, Lane 61, Sec. 2 Academia Road, Nan-kang, Taipei (TW).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **WANG, Yuh-Lin**; 4F., No. 8, Alley 20, Lane 85, Changxing Street, Da'an District, Taipei City, 106 (TW). **LIU, Ting-Yu**; No. 48, Lane 57, Sec. 1, Jiansing Road, Hsinchu County, Sinfong Township, 304 (TW). **WANG, Huai-Hsien**; No. 14, Heping Street, Tainan County, Xinshi Township, 744 (TW).

(74) Agent: **TSAO, Y. Rocky**; 10 Fawcett Street, Cambridge, Massachusetts 02138 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *without international search report and to be republished upon receipt of that report (Rule 48.2(g))*

(54) **Title**: GLYCOPEPTIDE-FUNCTIONALIZED NANOPARTICLE ARRAYS FOR CAPTURING AND DETECTING BIOMOLECULES

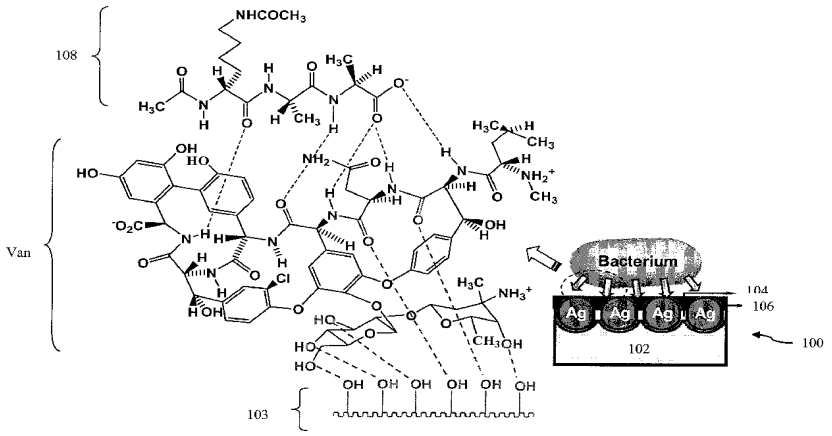


Figure 1

(57) **Abstract**: A surface-enhanced Raman spectroscopic (SERS) system for detecting a biomolecule. The system includes a substrate, an array of nanoparticles disposed on the substrate, each being partially embedded in the substrate and having a non-embedded surface, and a linking agent disposed on the non-embedded surface of each of the nanoparticles. The array of nanoparticles has a uniform interparticle gap of 1-50 nm and the linking agent is capable of binding to the biomolecule.

WO 2012/068226 A2

GLYCOPEPTIDE-FUNCTIONALIZED NANOPARTICLE ARRAYS FOR CAPTURING AND DETECTING BIOMOLECULES

BACKGROUND OF THE INVENTION

Surface-enhanced Raman spectroscopy (SERS) has been employed for label-free analysis of microorganisms and biomolecules to exploit its 10^6~10^{10} times enhancement in the Raman signal.

A variety of SERS substrates have been prepared by, e.g., disposing colloidal metal nanoparticles on a surface, roughening a metal surfaces to possess nanometer-scale features, or creating nanostructures on a surface by lithography. See, e.g., Demirel, M.C. et al., *Biointerphases* 4, 35-41 (2009); Stern, E. et al., *Nature Nanotech.*, 5, 138-142 (2010); Nie, S. et al., *Science*, 275, 1102-1106 (1997); Fang, Y. et al., *Science*, 321, 388-392 (2008); Li, J.F. et al., *Nature*, 464, 392-395 (2010); Tripp, R.A. et al., *Nanotoday*, 3, 31-37 (2008); Kao, P. et al., *Adv. Mater.*, 20, 3562-3565 (2008); Shachaf, C.M., et al., *PLoS ONE*, 4, e5206-e5217 (2009); and Qian, X. et al., *Nature Biotechnol.*, 26, 83-90 (2008).

Yet, there is still great need to develop new SERS substrates suitable for detecting biomolecules or microorganisms in a rapid, reliable, and uniform manner.

SUMMARY OF THE INVENTION

This invention is based at least in part on the unexpected discovery that certain SERS substrates, when coated with certain compounds (e.g., a glycopeptide antibiotic such as vancomycin), exhibit uniformly high sensitivity enhancement in detecting a biomolecule (e.g., a peptide on the cell wall of a bacterium). Thus, this invention relates to a SERS system and their use.

One aspect of this invention relates to a SERS system for detecting a biomolecule. The system includes a substrate, an array of nanoparticles disposed on the substrate, each being partially embedded in the substrate and having a non-embedded surface, and a linking agent disposed on the non-embedded surface of each of the nanoparticles. The linking agent is capable of binding to the biomolecule, e.g., a peptide on the cell wall of a bacterium. The term "nanoparticle" refers to the nanostructure in the shape of a particle having a diameter in the range of 10 nm to 100 nm and an aspect ratio between 1 and 5 (e.g., between 1 and 3).

This system may include one or more of the following features:

The system features (i)a substrate surface containing an array of wells having a uniform inter-well gap of 1-50 nm (e.g., less than 10 nm) and (ii) an array of nanoparticles being disposed in the array of wells, each well containing one nanoparticle. The wells are formed of aluminum oxide, titanium oxide, tantalum oxide, or niobium oxide. The term "inter-well gap" refers to the shortest distance between the outer rims of two neighbouring wells.

The array of nanoparticles in the system is formed of Ag, Au, or Cu. The non-embedded surface of nanoparticles is also formed of Ag, Au, or Cu. The inter-nanoparticle gaps can be uniform and the shortest distance can be 1-50 nm between the outer surfaces of two neighboring nanoparticles.

The linking agent used in the system can be any one of following glycopeptide antibiotics: vancomycin, teicoplanin, ristocetin, telavancin, bleomycin, ramoplanin, chloroeremomycin, and decaplanin. It is reversibly (e.g., non-covalently) attached to the non-embedded surface of each of the nanoparticles. Preferably, its molecular weight is not greater than 5 kDa, and, more preferably, not greater than 2 kDa. This linking agent forms, over the array of nanoparticles, a coating layer that has a thickness of 5-100 nm, e.g., 10-30 nm.

Another aspect of the present invention features a method for detecting the presence of a target biomolecule in a sample. The method includes providing a SERS system described above, contacting the SERS system with a sample suspected of containing the biomolecule, and detecting a SERS spectrum change (e.g., a change in the Raman spectrum shape or peak intensity) of the SERS system after the contacting step. When a SERS spectrum change is observed, it indicates that the sample contains the target biomolecule, which can be a peptide on the cell wall of a bacterium. This method may further include, after the detecting step, correlating a level of the SERS spectrum change with a concentration of the target biomolecule in the sample so as to quantify the target biomolecule. After the detecting step, the SERS system can be washed with a suitable solution to release the linking agent bound to the nanoparticles.

Still another aspect of the present invention features a method of identifying a bacteria. The method includes providing a SERS system described above, contacting the SERS system with a sample (e.g., a blood sample) containing or suspected of containing a

PREOMICS EXHIBIT 1002
Page 536 of 1861

bacteria, detecting a SERS spectrum of the SERS system, and comparing the SERS spectrum with a SERS spectrum of a bacteria-containing reference SERS system. The bacteria can be a drug-resistant bacteria, e.g., a bacteria resistant to the linking agent (which is a glycopeptides antibiotic) on the SERS system.

The details of several embodiments of the invention are set forth in the description below. Other features or advantages of the present invention will be apparent from the following drawings and actual examples, and also from the appended claims.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 is a schematic of a SERS active substrate used in this invention for capturing and detecting bacteria.

Figure 2 is a schematic diagram showing the process for fabricating metal-filled porous anodic alumina substrates. In panel f: S, D, and W are the interparticle spacing, particle diameter, and inter-well/interparticle gap, respectively.

Figure 3 illustrates Raman spectra of vancomycin in different forms.

Figure 4 illustrates the bacteria-capturing capability of a vancomycin-functionalized SERS substrate (coating solution: 10 mM vancomycin). Left panel: an optical image of a glass slide on which an Ag/AAO SERS substrate having a vancomycin-functionalized region (~300 μm in diameter) is placed; middle panel: a magnified view of vancomycin-functionalized region with the captured bacteria (*L. plantarum*); right panel: an SEM image of bacteria attached to the vancomycin-functionalized region.

Figure 5a illustrates the bacteria-capturing capability of various vancomycin-functionalized SERS substrates (coating solutions having 80 μM-50 mM of vancomycin; 10^9 cfu/ml bacteria seeding concentration) while Figure 5b illustrates the bacteria sensitivity of a vancomycin-functionalized SERS substrate (coating solution having 10 mM of vancomycin; various bacteria seeding concentrations of 10^3-10^9 cfu/ml).

Figure 6 illustrates relative SERS profiles (raw data after y-axis shift) from *E. coli* and *L. plantarum* bacteria on different substrates. A typical SERS of van-functionalized substrate is included for compoarison.

Figure 7 illustrates normalized SERS profiles from *E. coli* and *L. plantarum* bacteria on different substrates. Scale bar exhibits the counts per seconds (cps) of Raman intensity

PREOMICS EXHIBIT 1002
Page 537 of 1861

respectively.

Figure 8 illustrates SEM images and SERS spectra of pristine and Van-coated SERS substrates treated with VSE or VRE: (a) SERS spectra of Van-coated and pristine substrate treated with VSE or VRE, (b) SEM image of a pristine substrate treated with VSE, (c) SEM image of a pristine substrate treated with VRE, (d) SEM image of a Van-coated substrate treated with VSE, and (e) SEM image of a Van-coated substrate treated with VRE (all scale bars: 200 nm).

Figure 9 illustrates optical microscope images and SERS spectra of pristine and Van-coated substrates treated with a bacterium-spiked human blood sample: (a) SEM image of a pristine substrate treated with a bacteria-spiked human blood, (b) SEM image of a Van-coated substrate treated with a bacteria-spiked human blood, and (c) SERS spectra of a pristine substrate and a Van-coated substrate treated with a bacteria-spiked human blood (both scale bars: 5 μm).

DETAILED DESCRIPTION OF THE INVENTION

This invention relates to a SERS system and a method of using the system to capture and detect a target biomolecule such as a peptide on the cell wall of a bacterium. A schematic of the SERS system is illustrated in Figure 1. More specifically, as shown in the right panel of Figure 1, SERS system 100 includes a substrate 102 in which an array of wells 104 for holding metal nanoparticles (e.g., Ag nanoparticles) are formed, an array of nanoparticles partially embedded in the wells, and a layer of linking agent 106 covering the substrate and the nanoparticles. The linking agent 106 can bind to a biomolecule (e.g., protein, nucleic acid, polysaccharide, or lipid) or an organism (e.g., a bacterium, fungus or virus) via certain specific interactions (illustrated by gray arrows in Figure 1). As such, when system 100 is brought into contact with a sample suspected of containing the target biomolecule or organism, it can capture the target biomolecule or organism. Binding of the target biomolecule or organism to the linking agent 106 would affect the chemical environment of the linking agent, leading to a change in the SERS spectrum of SERS system 100 (including SERS signals from both the linking agent and the biomolecule or organism). The SERS spectrum change can therefore indicate the presence of the target biomolecule or organism in the sample.

In one embodiment, the linking agent 106 in the SERS substrate of this invention is vancomycin (denoted as "Van" in the left panel of Figure 1) and the target biomolecule is peptidoglycan or a certain fragment thereof such as D-Ala-D-Ala (denoted as "108" in Figure 1) on the cell wall of Gram-positive and Gram-negative bacteria. As illustrated in the left panel of Figure 1, vancomycin molecules bind to the terminal peptidoglycan, D-Ala-D-Ala, via hydrogen bonds, while vancomycin molecules also attach to the surface 103 of the SERS substrate 102 via hydrogen bonds.

The substrate 102 as shown in Figure 1 can be a metal substrate suitable for forming a nanometer-scale pattern of wells 104 via optical or e-beam lithography, electrochemical etching, or other etching methods. Examples of suitable metal include aluminum titanium, tantalum, niobium, tungsten, and zirconium. See, e.g., Wang et al., US Patent 7,453,565; Wang et al., *Adv. Mater.*, 18, 491-495 (2006); Liu, et al., *PLoS ONE*, 4, e5470-e5479 (2009); Singh et al., *ACS Nano*, 2 (12), 2453–2464 (2008); Shin & Lee, *Nano Lett.*, 8 (10), 3171-3173 (2008); and El-Sayed & Birss, *Nano Lett.*, 9 (4), 1350-1355 (2009). Preferably, aluminum is used as substrate 102. As illustrated by Figure 2, an anodic aluminum oxide (AAO) template having arrays of wells or pores with a substantially uniform inter-well gap (denoted as "W" in panel f) is prepared and used for fabricating arrays of Ag nanoparticles separated by the uniform inter-well gaps. First, as shown in panels a and b, an aluminum substrate (e.g., an Al foil) is finely polished and then anodized to form self-organized, hexagonally close-packed AAO wells. Next, as illustrated in panel c, the wells/pores are further enlarged by, e.g., etching the substrate in 5% phosphoric acid for a selected duration. This etching process also allows the selective control of the gap between the Ag nanoparticles deposited in the wells, as the interparticle gap equals the inter-well gap. After depositing Ag nanoparticles of the desired length, the upper part of the AAO film is etched away in phosphoric acid to increase the area of exposed Ag. See panels d through f.

After the substrate having an array of Ag nanoparticles partially embedded therein (i.e., "Ag/AAO substrate") is formed, the substrate is then placed into a solution containing a linking agent so that a layer of the linking agent can be coated onto the surface of the substrate and the exposed surfaces of the nanoparticles. In one example, Ag/AAO substrate is immersed in an aqueous solution of vancomycin hydrochloride with a concentration between 80 μM and 50 mM for a predetermined duration, thereby forming a vancomycin

layer with a thickness between 5 nm and 100 nm. As illustrated by Figure 1, vancomycin binds to the Ag/AAO substrate via hydrogen bonds.

In other embodiments, vancomycin can bind to a nanoparticle-embedded substrate via covalent bonds (e.g., thiol-bonds). Unexpectedly, the bacteria capturing capability of a SERS substrate having vancomycin attached via hydrogen bonding is better than that via covalent bonding.

As mentioned above, the linking agent can interact with the nanoparticle-embedded substrate in a reversible manner so that the substrate can be reused by releasing the linking agents bound to the target biomolecule and depositing on the released substrate a fresh layer of linking agents. The release can be achieved by washing the target biomolecule-bound SERS system with a solution containing a compound that interacts more strongly with the linking agent than the substrate, at a suitable concentration, so as to produce a free nanoparticle-embedded substrate (no longer bound to linking agent 106). This free substrate can then be re-deposited with a layer of linking agents and re-used in a subsequent assay for detecting presence of a target molecule.

The SERS system of this invention can also be used to quantify a target biomolecule as follows. SERS system 100 is brought into contact with a solution containing a target biomolecule capable of binding to linking agent 106 at a predetermined concentration. The level of the SERS spectrum change (e.g., a Raman peak intensity change) of system 100 after exposure to the solution is determined. A standard curve is then prepared based on the concentrations of the target biomolecule (e.g., log values) versus the levels of spectrum changes caused thereby. SERS system 100 is then exposed to a sample containing the target biomolecule. The level of the SERS spectrum change is measured. The concentration of the target molecule is then determined by comparing the level of the spectrum change with the standard curve.

In addition, the SERS system of this invention can be used to identify bacteria. An example follows. SERS system 100 is brought into contact with a bacteria solution (e.g., a bacteria-spiked blood sample). The SERS spectrum of the system is compared with the SERS spectra of the reference systems treated with various bacteria to determine the type of the bacteria based on the similarity between the sample system and each of the reference systems.

PREOMICS EXHIBIT 1002
Page 540 of 1861

Further, the SERS system of this invention can be used to distinguish drug-resistant bacteria from drug-susceptible bacteria. For example, SERS system 100 is brought into contact with a solution containing bacteria and a SERS spectrum of the system is compared with a SERS spectrum of a system treated with known corresponding drug-resistant or susceptible bacteria to determine the susceptibility.

Without further elaboration, it is believed that one skilled in the art can, based on the above description, utilize the present invention to its fullest extent. The following specific embodiments are, therefore, to be construed as merely illustrative, and not limitative of the remainder of the disclosure in any way whatsoever. All publications cited herein are incorporated by reference.

EXAMPLE 1: Fabrication of Vancomycin ("Van")-Functionalized SERS Substrate

SERS-active Ag/AAO substrate containing arrays of Ag nanoparticles partially embedded in AAO nanowells, was fabricated according to the methods described in Wang et al., *Adv. Mater.*, 18, 491-495 (2006). More specifically, high-purity (99.99%) annealed aluminum foil was electropolished in a mixture of $HClO_4$ and C_2H_5OH (volume ratio 1:5) until the root mean-square surface roughness of a typical 10 μm×10 μm area was 1 nm, as measured using an atomic force microscope operating in contact mode. The foil was then anodized in sulfuric acid (0.3 M) at 5°C using a voltage in the 10–30 V range to obtain AAO substrates with arrays of self-organized wells having a diameter of a few nanometers to a few hundred nanometers ("nanowells"). The nanowells in the AAO substrate were then enlarged by etching the substrate in 5% phosphoric acid to increase the diameters of the nanowells. By carefully controlling the etching process, arrays of nanowells with 5±2 nm inter-well gap were achieved. An electrochemical plating procedure was then employed to grow Ag nanoparticles into the AAO substrate. To grow Ag nanoparticles in the AAO nanowells, an alternating current
(9 V) electrochemical plating procedure was employed using a mixture of silver nitrate (0.006 M) and magnesium sulfate (0.165 M) as the electrolyte solution with a pH value of 2, set by the addition of sulfuric acid. Since the deposition of Ag occurred primarily inside the nanowells, it was possible to avoid the merging of Ag nanoparticles by confining them inside the wells.

Next, an Ag/AAO SERS substrate (1x1 cm^2) produced by the method described

above was immersed in an aqueous solution of vancomycin hydrochloride (Sigma, USA; with various concentrations of 80μM-50 mM) for 1 hour, and then dried in the air for 12 hours to produce vancomycin-functionalized Ag/AAO SERS substrates with different coating thicknesses varying from 5 nm to 100 nm.

EXAMPLE 2: Characterizations of Van-Functionalized SERS Substrate

Escherichia coli (*E. coli*, ATCC 11775) and *Lactobacillus plantarum* (*L. Plantarum*, ATCC 8014) were purchased from BCRC company in Taiwan. These bacteria were cultivated for 16 h at 37°C on nutrient, MRS, and Brain Heart Infusion agar bases, respectively. After subculturing, single colonies were collected using sterile plastic inoculating loops. Bacteria were suspended in 5 mL specific broth, grown for an additional 14 hours then sub-cultured at OD_{600}~0.5, taken as the beginning of the exponential growth phase. They were then washed and centrifuged three times with water and re-suspended in water.

Bacteria solution (1 mL, various concentrations: 10^4~10^9 cfu/ml) was dropped onto a Van-functionalized SERS substrate placed in a well of a 24-well cell culture plate. Then, the bacteria were incubated at 37°C under a 120 rpm shaking rate for 1 hour. Afterwards, the SERS substrate was washed with water 5 times before the Raman spectroscopy measurements were conducted.

Raman spectroscopy measurements were performed on a Raman microscopy (HR800, Jobin-Yvon) equipped with a HeNe laser at 632.8 nm (0.1 mW) and 50× objective lens. Individual single bacterium or clusters of bacteria were easily identified under this microscope system.

Raman signals were collected from the information-rich part of the spectrum between 400 and 1800 cm^{-1} using 60s acquisition time. The raw SERS readout datasets were processed and normalized to remove noises in the following three manners: (1) a median filter with noise estimation was applied to eliminate any sharp variations caused by cosmic rays; (2) a wavelet de-noising technique was used to smooth out high-frequency noise; and (3) iterative curve fitting to estimate and remove the background baseline due to the noise effect of environmental light.

Figure 3 displays SERS spectra of vancomycin on the Ag/AAO substrate, i.e., profile

PREOMICS EXHIBIT 1002
Page 542 of 1861

(a), and vancomycin in solution form, i.e., profiles (b). Compared with spectrum (b), containing a few sharp peaks, spectrum (a) only contained two broad bands. This indicates that vancomycin molecules were randomly immobilized on the Ag/AAO substrate by hydrogen bonds between their C=O moieties and the OH moieties on the surface of the Ag/AAO substrate, which resulted in the broadening and the reduced intensity of the Raman peaks corresponding to C=O stretching mode. See Lee, *J. Raman Spectrosc.*, 28, 45-51 (1997). As the broadband background could be easily treated by a simple background subtraction routine in the data acquisition system, the lack of sharp peaks in the SERS spectrum of the vancomycin-functionalized Ag/AAO SERS substrate was unexpectedly superior to SERS substrates functionalized by other molecules exhibiting sharp Raman peaks which interfere with the SERS spectrum of the biomolecule to be detected.

Scanning electron microscopy (SEM) was performed on DualBeam™ FIB/SEM system (FEI Nova™ 600, USA).

As shown in Figure 4, *L. plantarum* (Gram-positive) bacteria were hedged and concentrated on a Van-functionalized region of an Ag/AAO SERS substrate, demonstrating the much stronger bacteria capturing capability of the Van-functionalized region than the Van-free Ag/AAO region. Quantitatively, Figure 5a shows that the bacteria-capturing capability, defined as the ratio of the bacteria density (cells/mm^2) on a Van-functionalized substrate to the bacteria density on a Van-free Ag/AAO substrate, increased in a linear fashion as the vancomycin concentration of the coating solution increased for both *L. plantarum* and *E. coli* bacteria (Gram-negative). Unexpectedly, the bacteria-capturing capability increased by 700 times where the coating solution containing 50 mM vancomycin was used. Furthermore, the bacteria capturing capability also was evaluated under different bacteria seeding concentration (10^3~10^9 cfu/ml) on the Van-functionalized SERS substrates prepared from a 10 mM Van coating solution, as shown in Figure 5b. The results indicate that bacteria density increases as bacteria seeding concentration increases. Also, the results indicate that the Van-functionalized SERS substrate can be sensitive to a bacteria seeding concentration as low as 10^2 cfu/ml.

Figure 6 illustrates the relative SERS profiles (raw data after y-axis shift) of *E. coli* and *L. plantarum* bacteria on different SERS substrates. Clearly, the Van-functionalized SERS substrates (denoted as "Van-fun. Sub" in Figure 6) showed 4-5 times enhancement in

PREOMICS EXHIBIT 1002
Page 543 of 1861

the intensities of the major SERS peaks from the bacteria (733 cm^{-1} peak from *L. plantarum* and 654 cm^{-1} peak from *E. coli*) compared to that from Van-free Ag/AAO substrates (denoted as "Pristine Sub" in Figure 6).

As shown in Figure 7, the normalized SERS profile from the both bacteria on the Van-functionalized substrate (denoted as "Van-fun. Sub" in Figure 7) exhibits no significant change in the positions and relative intensities of major SERS peaks, as compared to that from the Van-free Ag/AAO substrate (denoted as "Pristine Sub" in Figure 7). However, some of secondary peaks in the bacteria cell wall become more obvious and shaper, especially in the negative bacteria with the thinner cell wall.

EXAMPLE 3: Capture of Drug-Resistant or Susceptible Bacteria Strains

Van-coated and pristine substitutes were tested following the method described in Example 2, except that both Van-resistant *Enterococcus* (VRE) and Van-susceptible *Enterococcus* (VSE) were used. Figure 8a shows SERS spectra of VRE and VSE on the pristine and Van-coated SERS substrates. While the SERS spectra of the pristine substrates treated with VSE andVRE were essentially identical, the SERS spectra of Van-coated substrates treated with VSE and VRE were distinguishable. Specifically, several extra peaks (labelled by dots) appeared in the SERS spectrum of VSE on the Van-coated substrate, and the spectrum of VRE on the Van-coated substrate was very similar to the spectra of both VRE and VSE on the pristine substrates. The results indicate that Van-coated substrates can be use to distinguish drug-resistant Gram-positive bacteria from drug-susceptible Gram-positive bacteria.

SEM images of VSE and VRE on both pristine and Van-coated substrates were also obtained. Figures 8b and 8c show the SEM images of the pristine substrates treated with VSE and VRE. The images were substantially identical. Figure 8d shows the SEM image of the Van-coated substrate treated with VSE, where the cell walls of many VSE were disrupted and secretion was clearly observed. Figure 8e shows the SEM image of the Van-coated substrate treated with VRE, where the cell walls remained intact.

EXAMPLE 4: Capture of Bacteria from Human Blood

The method described in Example 2 was conducted except that *L. plantarum* spiked

human blood, in stead of a bacteria solution, was used. Figures 9a and 9b show SEM images of the pristine and Van-coated substrates treated with the bacteria-spiked blood sample. The images demonstrate that (i) the sticking probability of blood cells (i.e., white blood cells, red blood cells, and platelets) on pristine substrates was significantly greater than that on Van-coated substrates; and (ii) the sticking probability of bacteria on pristine substrates was significantly lower than that on Van-coated substrates.

Figure 9c displays the SERS spectra of pristine and Van-coated substrates treated with the bacteria-spiked blood sample. It shows that (i) the SERS spectrum of the Van-coated substrate treated with the bacteria-spiked blood sample was similar to that of a pristine substrate treated with bacteria solution, and (ii) the SERS spectrum of a pristine substrate treated with the bacteria-spiked blood sample was very different from that of a Van-coated substrate treated with the bacteria-spiked blood sample, but was substantially identical to that of a pristine substitute treated with an un-spiked blood sample. These results indicate that, while the pristine substrate was overwhelmed by the blood cells, the Van-coated substrate was effective in capturing bacteria from blood, free from disturbance by blood cells.

OTHER EMBODIMENTS

All of the features disclosed in this specification may be combined in any combination. Each feature disclosed in this specification may be replaced by an alternative feature serving the same, equivalent, or similar purpose. Thus, unless expressly stated otherwise, each feature disclosed is only an example of a generic series of equivalent or similar features.

From the above description, one skilled in the art can easily ascertain the essential characteristics of the present invention, and without departing from the spirit and scope thereof, can make various changes and modifications of the invention to adapt it to various usages and conditions. Thus, other embodiments are also within the claims.

PREOMICS EXHIBIT 1002
Page 545 of 1861

What is claimed is:

1. A surface-enhanced Raman spectroscopic (SERS) system for detecting a biomolecule, the system comprising:

a substrate,

an array of nanoparticles disposed on the substrate, each being partially embedded in the substrate and having a non-embedded surface, and

a linking agent disposed on the non-embedded surface of each of the nanoparticles,

wherein the linking agent is capable of binding to the biomolecule.

2. The SERS system of claim 1, wherein the biomolecule is a peptide on the cell wall of a bacterium.

3. The SERS system of claim 2, wherein the linking agent is a glycopeptide antibiotic.

4. The SERS system of claim 3, wherein the glycopeptide antibiotic is selected from the group consisting of vancomycin, teicoplanin, ristocetin, telavancin, bleomycin, ramoplanin, chloroeremomycin, and decaplanin.

5. The SERS system of claim 1, wherein the linking agent is vancomycin.

6. The SERS system of claim 1, wherein the linking agent is reversibly attached to the non-embedded surface of each of the nanoparticles.

7. The SERS system of claim 1, wherein the linking agent is non-covalently attached to the non-embedded surface of each of the nanoparticles.

8. The SERS system of claim 1, wherein the linking agent has a molecular weight of not greater than 5 kDa.

9. The SERS system of claim 1, wherein the linking agent has a molecular weight of not greater than 2 kDa.

10. The SERS system of claim 1, wherein the linking agent forms a coating layer over the array of nanoparticles, the coating layer having a thickness of 5-100 nm.

11. The SERS system of claim 10, wherein the coating layer has a thickness of 10-30 nm.

12. The SERS system of claim 1, wherein the substrate has a surface containing an array of wells having a uniform inter-well gap of 1-50 nm and the array of nanoparticles are disposed in the array of wells, each well containing one nanoparticle.

13. The SERS system of claim 12, wherein the wells are formed of an oxide selected from the group consisting of aluminum oxide, titanium oxide, tantalum oxide, and niobium oxide.

14. The SERS system of claim 12, wherein the wells are formed of aluminum oxide.

15. The SERS system of claim 12, wherein the uniform inter-well gap is less than 10 nm.

16. The SERS system of claim 1, wherein the non-embedded surface of each of the array of nanoparticles is formed of a metal selected from the group consisting of Ag, Au, and Cu.

17. The SERS system of claim 1, wherein each of the array of nanoparticles is formed of a metal selected from the group consisting of Ag, Au, and Cu.

18. A method for detecting presence of a biomolecule in a sample, the method

PREOMICS EXHIBIT 1002
Page 547 of 1861

comprising:

providing a SERS system of claim 1;

contacting the SERS system with a sample suspected of containing the biomolecule, and

detecting a SERS spectrum change of the SERS system after the contacting step, wherein observation of a SERS spectrum change indicates presence of the biomolecule in the sample.

19. The method of claim 18, wherein the biomolecule is a peptide on the cell wall of a bacterium.

20. The method of claim 18, further comprising correlating a level of the SERS spectrum change with a concentration of the biomolecule in the sample.

21. A method of identifying bacteria, the method comprising:

providing a SERS system of claim 1,

contacting the SERS system with a sample containing or suspected of containing a bacteria,

detecting a SERS spectrum of the SERS system, and

comparing the SERS spectrum with a SERS spectrum of a bacteria-containing reference SERS system.

22. The method of claim 21, wherein the bacteria is drug-resistant bacteria.

23. The method of claim 21, wherein the sample is a blood sample.

PREOMICS EXHIBIT 1002
Page 548 of 1861

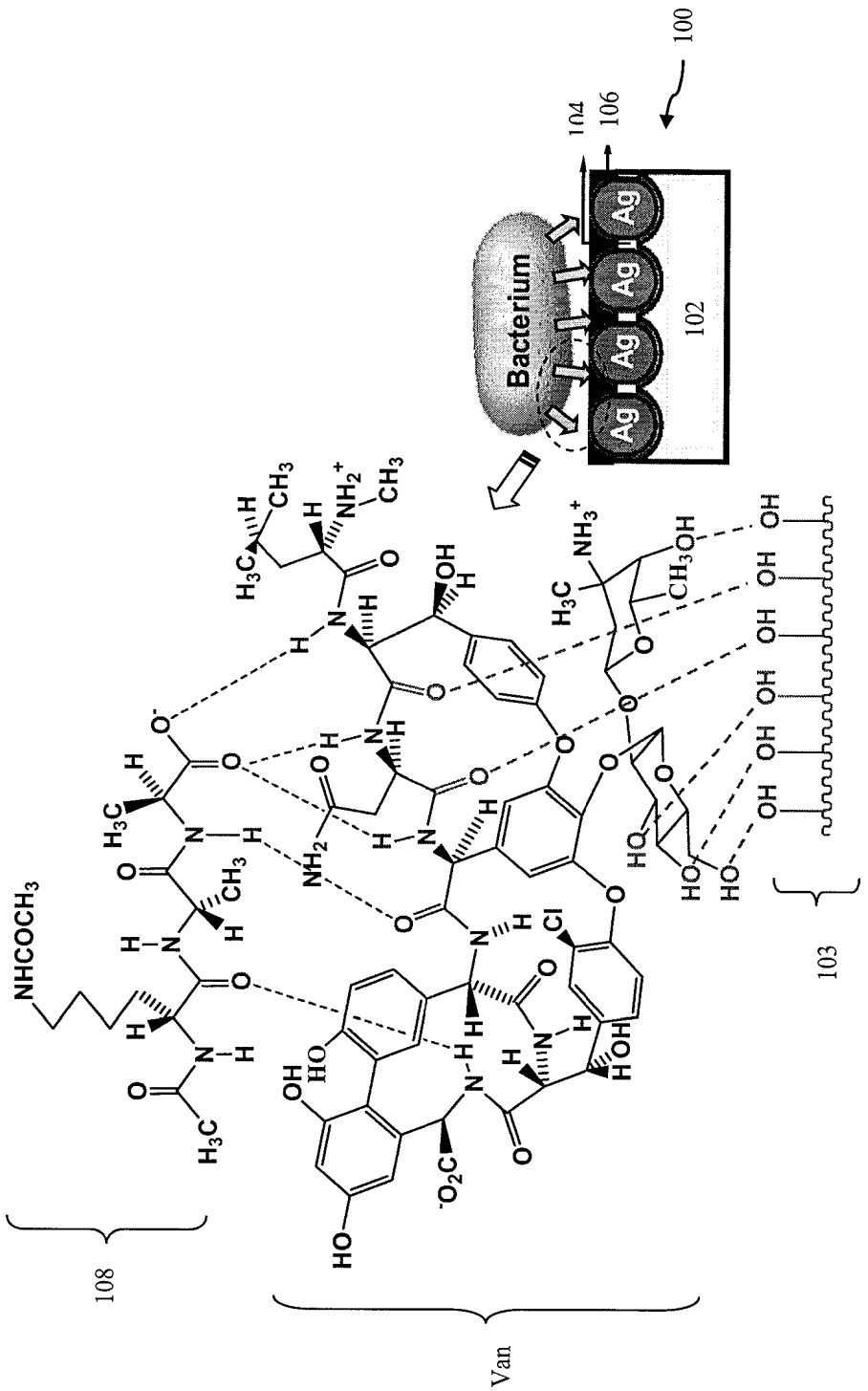


Figure 1

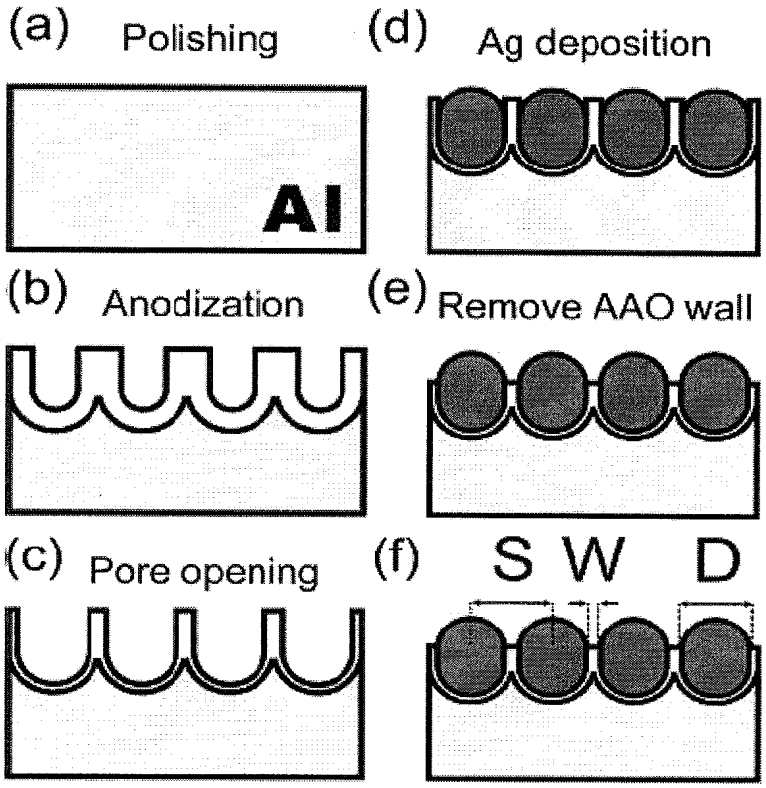

(a) Polishing
Al
(b) Anodization
(c) Pore opening
(d) Ag deposition
(e) Remove AAO wall
(f) S W D

Figure 2

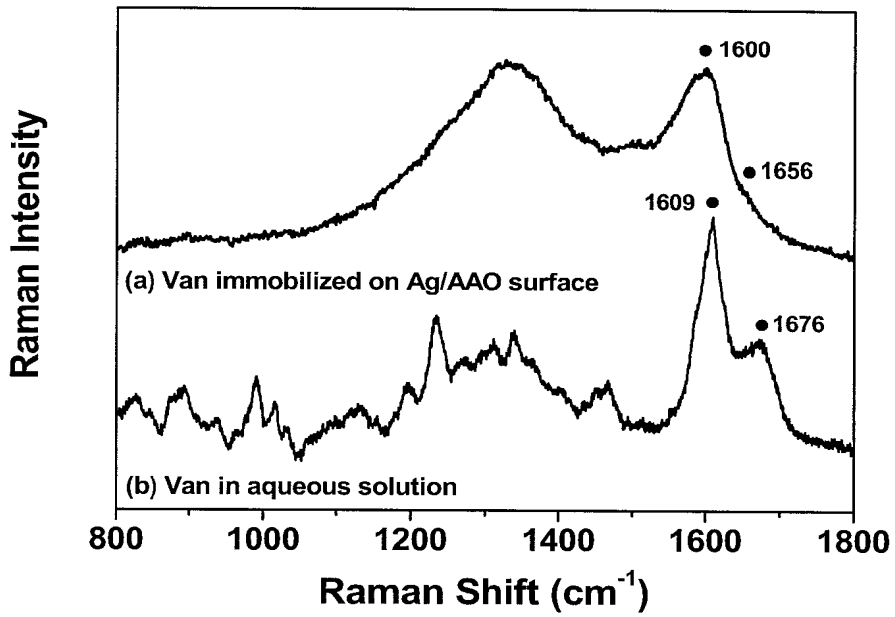

1600
1656
1609
(a) Van immobilized on Ag/AAO surface
1676
(b) Van in aqueous solution
Raman Intensity
800
1000
1200
1400
1600
1800
Raman Shift (cm-1)

Figure 3

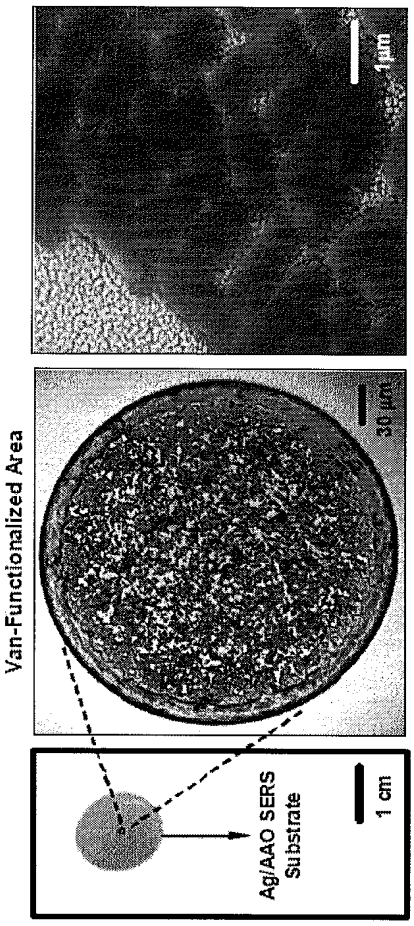


Figure 4

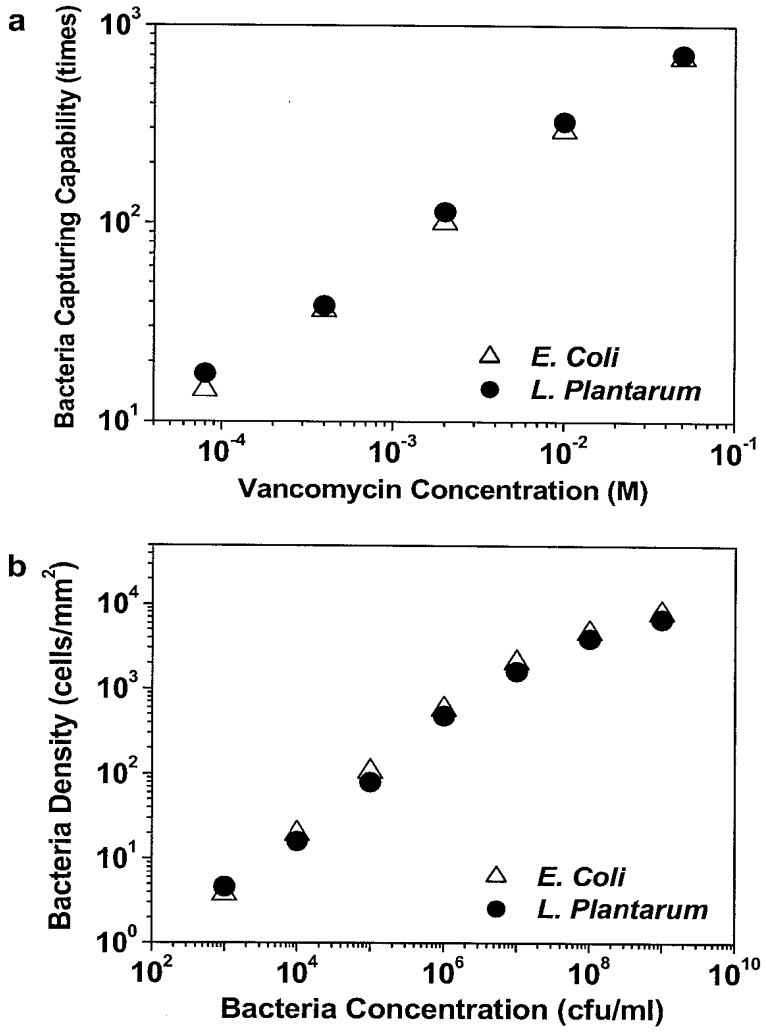

a
Bacteria Capturing Capability (times)
10^3
10^2
10^1
10^{-4}
10^{-3}
10^{-2}
10^{-1}
Vancomycin Concentration (M)
E. Coli
L. Plantarum
b
Bacteria Density (cells/mm^2)
10^4
10^3
10^2
10^1
10^0
10^2
10^4
10^6
10^8
10^{10}
Bacteria Concentration (cfu/ml)
E. Coli
L. Plantarum

Figure 5

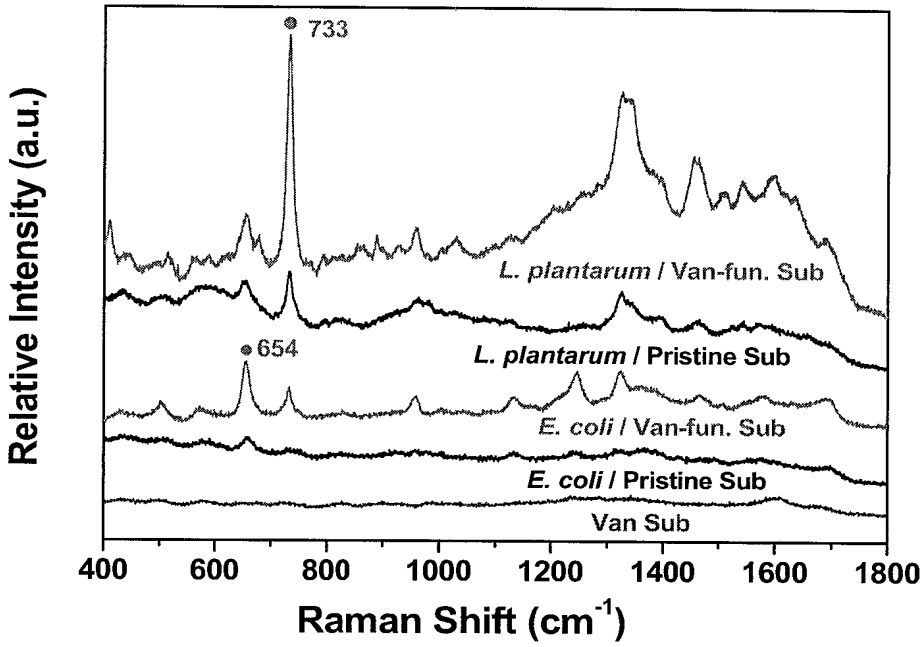

733
654
L. plantarum / Van-fun. Sub
L. plantarum / Pristine Sub
E. coli / Van-fun. Sub
E. coli / Pristine Sub
Van Sub
Relative Intensity (a.u.)
400
600
800
1000
1200
1400
1600
1800
Raman Shift (cm-1)

Figure 6

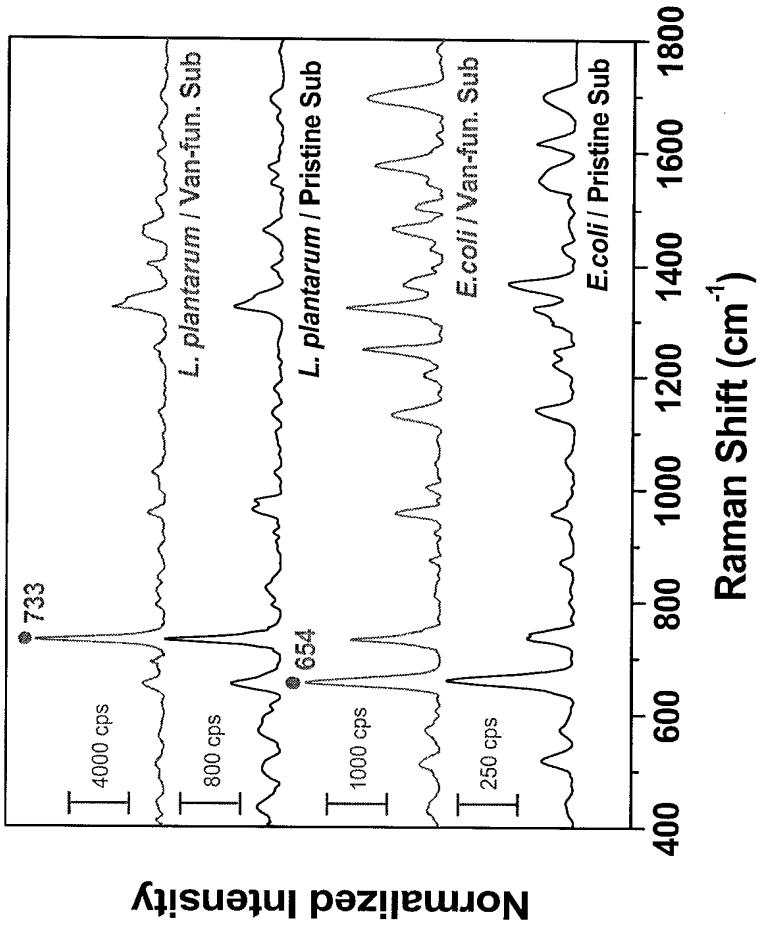

733
654
4000 cps
800 cps
1000 cps
250 cps
L. plantarum / Van-fun. Sub
L. plantarum / Pristine Sub
E.coli / Van-fun. Sub
E.coli / Pristine Sub
400
600
800
1000
1200
1400
1600
1800
Raman Shift (cm-1)
Normalized Intensity

Figure 7

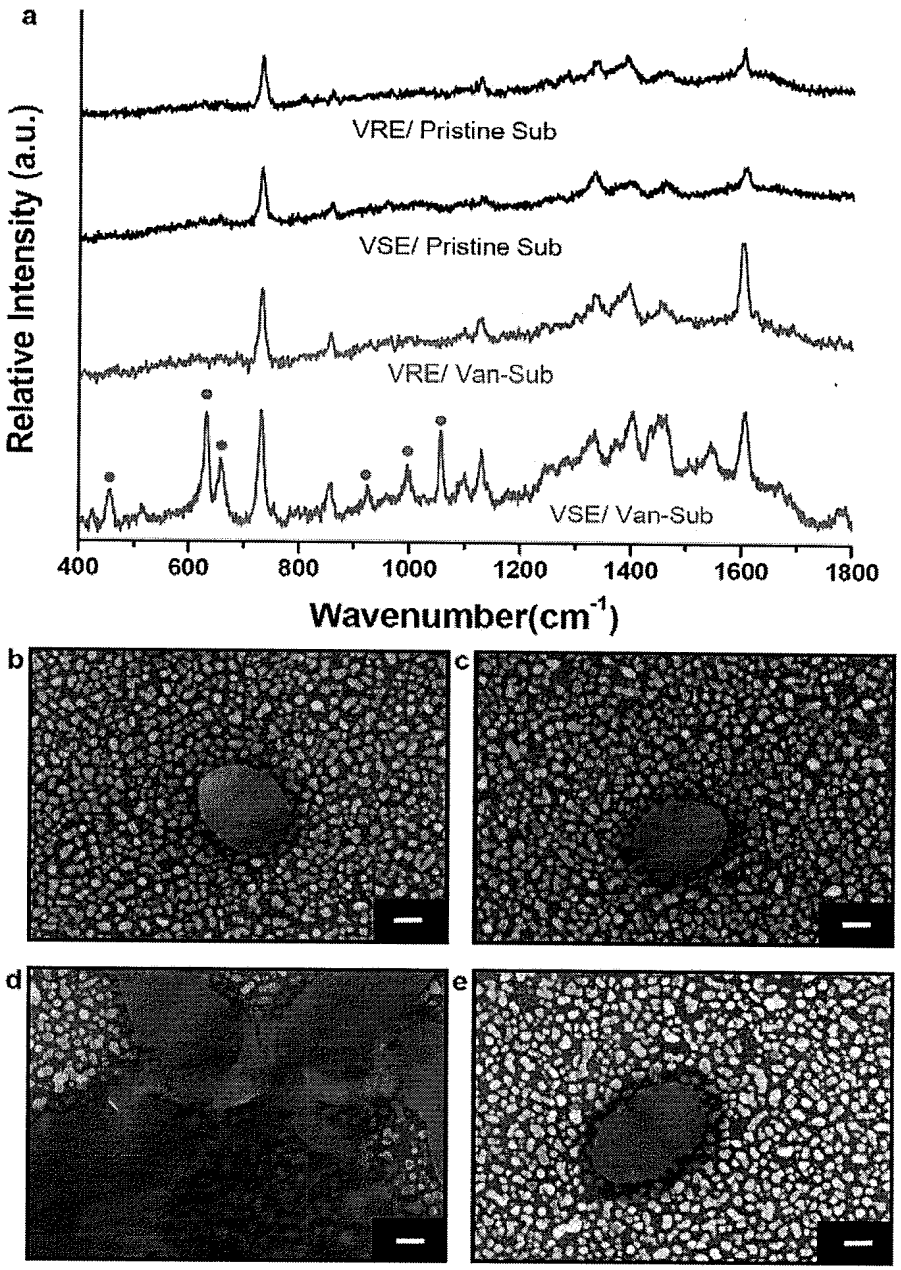

a
Relative Intensity (a.u.)
VRE/ Pristine Sub
VSE/ Pristine Sub
VRE/ Van-Sub
VSE/ Van-Sub
400
600
800
1000
1200
1400
1600
1800
Wavenumber(cm-1)
b
c
d
e

Figure 8

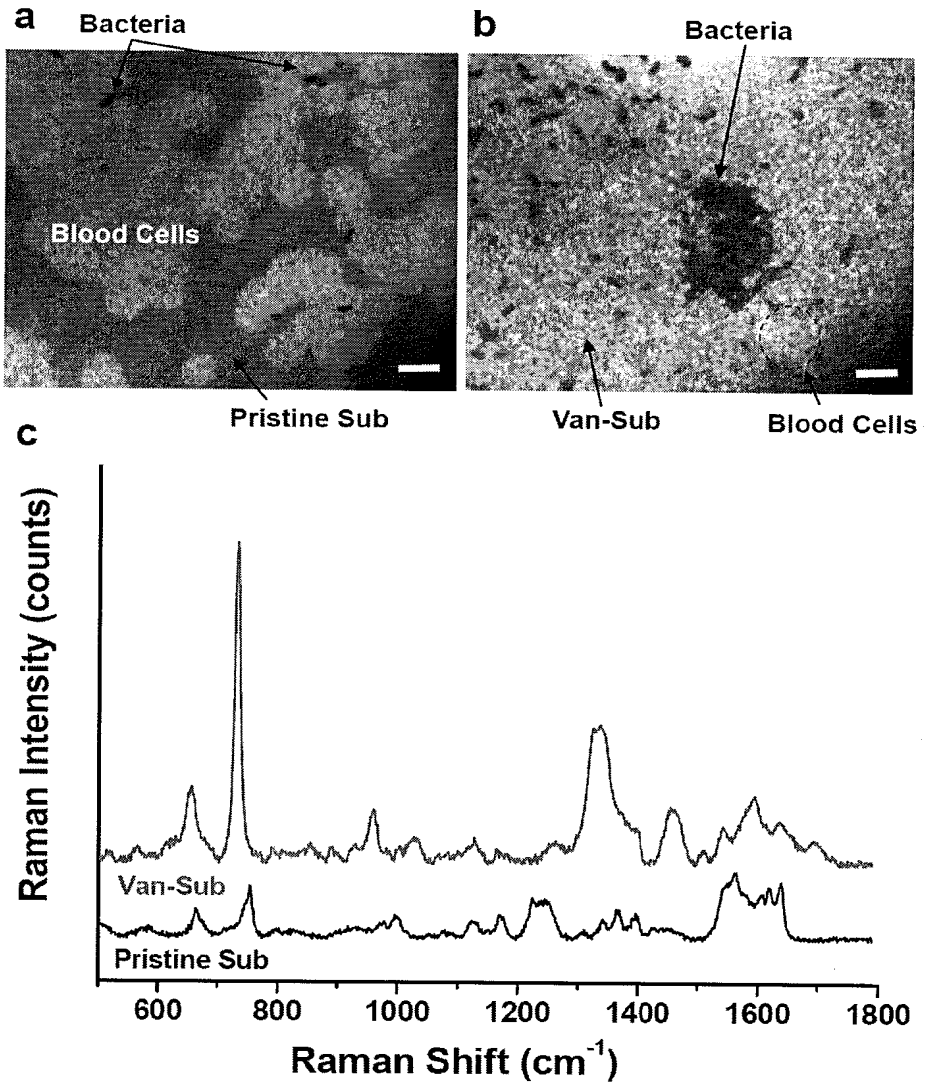

a
Bacteria
Blood Cells
Pristine Sub
b
Bacteria
Van-Sub
Blood Cells
c
Raman Intensity (counts)
Van-Sub
Pristine Sub
600
800
1000
1200
1400
1600
1800
Raman Shift (cm-1)

Figure 9

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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
15 March 2018 (15.03.2018)

WIPO | PCT

(10) International Publication Number
WO 2018/046542 A1

(51) International Patent Classification:
G01N 33/543 (2006.01) *A61K 9/127* (2006.01)
G01N 33/574 (2006.01)

(21) International Application Number:
PCT/EP2017/072349

(22) International Filing Date:
06 September 2017 (06.09.2017)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
1615128.4 06 September 2016 (06.09.2016) GB

(71) Applicant: **THE UNIVERSITY OF MANCHESTER** [GB/GB]; Oxford Road, Manchester M13 9PL (GB).

(72) Inventors: **KOSTARELOS, Kostas**; 20 Ambler Road, London N4 2QU (GB). **HADJIDEMETRIOU, Marilena**; Revekkas 1A Strovolos, Nicosia, 2039 (CY).

(74) Agent: **HGF LIMITED**; 4th Floor Merchant Exchange, 17-19 Whitworth Street West, Manchester M1 5WG (GB).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report (Art. 21(3))*

WO 2018/046542 A1

(54) Title: DETECTION OF CANCER BIOMARKERS USING NANOPARTICLES

(57) Abstract: The present invention relates to methods for identifying and detecting potential disease specific biomarkers from biofluids. In particular, the methods involve in vivo administration of nanoparticles to a subject in a diseased state or incubating nanoparticles in a biofluid sample taken from a subject in a diseased state and analysis of the biomolecule corona formed on said nanoparticles. In addition, the present invention also relates to methods that provide a means to distinguish between a healthy and diseased state in a subject, such as for example the presence of a tumor in a human subject.

DETECTION OF CANCER BIOMARKERS USING NANOPARTICLES

FIELD OF THE INVENTION

[0001] The present invention relates to methods for identifying and detecting potential disease specific biomarkers from biofluids. In particular, the methods involve administration of nanoparticles to a subject in a diseased state or incubation of nanoparticles in a biofluid sample taken from a subject in a diseased state and subsequent analysis of the biomolecule corona formed on said nanoparticles. In addition, the present invention also relates to methods that provide a means to distinguish between healthy and diseased states in a subject, such as for example the early detection of a tumor in a human subject or the presence or monitoring of the growth and/or response to treatment of a tumor in a human subject.

BACKGROUND OF THE INVENTION

[0002] A biomarker, or biological marker, generally refers to a qualitative and/or quantitative measurable indicator of some biological state or condition. Biomarkers are typically molecules, biological species or biological events that can be used for the detection, diagnosis, prognosis and prediction of therapeutic response of diseases. Most biomarker research has been focused on measuring a concentration change in a known/suspected biomarker in a biological sample associated with a disease. Such biomarkers can exist at extremely low concentrations, for example in early stage cancer, and accurate determination of such low concentration biomarkers has remained a significant challenge.

[0003] Research into nanoparticle-based technologies for biomarker detection has been carried out but so far has failed to provide suitable methods to accurately identify/discover and detect biomarkers. One particular problem is that currently available laboratory tests detect only a minute fraction of potential biomarkers, due to their extremely low concentration in biofluids. In addition to the 'swamping' effect, caused by "non-specific" high abundant molecules, this casuses significant difficulties. Indded, the issue of "signal-to-noise" exceeds the current capability of proteomic analysis and therefore limits the diagnostic information that can be obtained.

[0004] Furthermore, such methods are mainly used to detect already known disease-specific molecules. For example, nanoparticle-based immunoassays, such as sandwich-type assays have been developed, where nanoparticles are used to fluorophore-label the secondary antibody to a known disease-specific molecule. The presence of a target biomarker generates a fluorescent signal that is detected using microscopy or a fluorescence spectrophotometer. NanoDLSay™ is a particular nanoparticle-based immunoassay that detects specific known

biomarkers by measuring the size change of gold nanoparticle probes upon binding with target molecules. Gold nanoparticles conjugated to specific antibodies are mixed *in vitro* with biofluids. The binding of proteins with the gold nanoparticle immunoprobes can lead to nanoparticle cluster or aggregate formation. The average particle size increase of the assay solution is measured and correlated to the target protein concentration.

[0005] Nanoparticle-based immunoassays are inherently limited by the detection of a single molecule-biomarker. In order to increase the sensitivity and specificity of detection, biomarkers can, for example, be regarded as a panel of up- and down-regulated proteins, which differ in diseased and normal state. 'Multiplexing' is a difficult and highly expensive challenge with such technologies.

[0006] *In vitro* nanoparticle-based scavenger technologies such as Nanotrap® particle technology uses hydrogel nanoparticles with a porous outer shell that blocks the entry of high molecular weight proteins and an internal core which contains chemical affinity baits for interaction with low molecular weight proteins, after the introduction of nanoparticles into biological fluids. The limitation of this approach is that some of the high molecular weight proteins removed are known to act as 'carriers' for other proteins, possibly with biomarker potential.

[0007] Surprisingly, the inventors have found that analyzing the biomolecule corona formed on nanoparticles after following methods involving administration of nanoparticles to a subject in a diseased state or incubation of nanoparticles in a biofluid sample taken from a subject in a diseased state results in interaction of the nanoparticles with many more different types of biomolecules in comparison with analysis of biomolecules directly in biofluid samples. In one embodiment, the novel methods take advantage of the interaction of nanoparticles with biomolecules as a way to more comprehensively analyse the circulatory proteome and facilitate the detection of previously unknown disease-specific biomolecules.

[0008] In addition, inventors have also surprisingly found that particular methods of the invention can be employed to distinguish between healthy and disease states in a subject, for example the detection of the presence of a tumor or monitoring the growth and/or response to treatment of a tumor.

SUMMARY OF THE INVENTION

[0009] In one aspect, the present invention relates to a method of identifying a biomarker from a biofluid, wherein the method comprises:

a. administering a plurality of nanoparticles to a subject in a diseased state to allow a biomolecule corona to form on the surface of said nanoparticles or incubating a

PREOMICS EXHIBIT 1002
Page 560 of 1861

plurality of nanoparticles in a biofluid sample taken from a subject in a diseased state to allow a biomolecule corona to form on the surface of said nanoparticles;

b. isolating the nanoparticles and surface-bound biomolecule corona; and

c. analyzing the biomolecule corona to identify the said biomarker.

[00010] The methods result in an interaction between the nanoparticles and a greater number of different types of proteins than can be detected by direct analysis of biofluids taken from a subject in a diseased state. It is to be understood that the method involves identification of a biomarker that provides a measurable indicator of some biological state or condition. This includes, but is not limited to, the discovery of unique disease-specific biomolecules (those bioloecules that are only present in a diseased state) but also includes detection of changes in biomolecule(s) that are present in both healthy and diseased states, for example upregulation or down regulation of biomolecules in a diseased state when compared to the healthy state. It will be understood that in order to identify a potential disease-specific biomarker, comparison against a suitable non-diseased control reference can be required. In one particular embodiment, the methods involve identifying panels of biomarkers (multiplexing), which can lead to increased sensitivity and specificity of detection. In a further particular embodiment, the methods facilitate the detection of previously unknown unique disease-specific biomolecules. In yet a further particular embodiment, the methods allow identification or detection of a biomarker without the need for invasive tissue sampling, e.g. a biopsy.

[00011] The methods are applicable to a wide range of nanoparticles and allow the benefit of removal of unbound and highly abundant biomolecules to allow identification of low abundant biomarkers that would otherwise be undetected. In addition to identification of potential biomarkers, the methods can also be employed to monitor changes in biomarkers, for example in response to therapy and/or to assist in diagnosis.

[00012] In another aspect, the present invention relates to a method of detecting a diseased state in a subject, wherein the method comprises:

a. administering a plurality of nanoparticles to a subject to allow a biomolecule corona to form on the surface of the nanoparticles;

b. isolating the nanoparticles and surface-bound biomolecule corona; and

c. determining the total biomolecule content of the biomolecule corona, which is determinative of the presence of disease in said subject.

[00013] Surprisingly, inventors have found that particular methods can be employed to distinguish between healthy and diseased states in a subject. Such methods can, for example, be useful in the early detection of a diseased state such as the presence of a tumor in a human

PREOMICS EXHIBIT 1002
Page 561 of 1861

subject.

[00014] The methods disclosed herein are applicable to any disease state in which identification and/or monitoring of biomarkers would be beneficial. Furthermore, particular methods of the invention, which an be employed to distinguish between healthy and diseased states in a subject, are applicable to a wide range of diseases, including but not limited to, cancer and neurodegenerative diseases.

BRIEF DESCRIPTION OF DRAWINGS

Figure 1 shows the SDS-Page gel for protein recovered directly from plasma compared with the analysis conducted following either *in vivo* administration of nanoparticles in mice or after *in vitro* incubation with plasma taken from mice. Data is presented for healthy or tumor bearing mice (i.e. lung adenocarcinoma model).

Figure 2 shows the SDS-Page gel for protein recovered directly from plasma compared with the analysis conducted following either *in vivo* administration of nanoparticles in mice or after *in vitro* incubation with plasma taken from mice. Data is presented for healthy or tumor bearing mice (i.e. melanoma model).

Figure 3 shows the number of proteins identified following *in vivo* administration of nanoparticles in mice or after *in vitro* incubation when compared with plasma. The data demonstrates that a significantly higher number of proteins were identified when compared to plasma. Data generated using lung adenocarcinoma model.

Figure 4 provides a breakdown of the host proteins identified following *in vivo* administration (4a) and after *in vitro* incubation (4b) in the lung adenocarcinoma model, ranked by fold change in amount of that protein detected in the protein corona from healthy and diseased mice (proteins with greater than 5 fold change shown).

Figure 5 provides a breakdown of the proteins identified following *in vivo* administration (5a) and after *in vitro* incubation (5b) in the melanoma model ranked by fold change in amount of that protein detected in samples from healthy and diseased mice (proteins with greater than 5 fold change shown).

Figure 6 provides data from a BCA assay for the lung adenocarcinoma model. The level of total amount of protein detected in the protein corona from tumor-bearing animals following *in vivo* administration is significantly greater than after *in vitro* incubation. Furthermore, there is a clear distinction between tumor and healthy mice in regards to the level of total amount of protein detected in the protein corona following *in vivo* administration.

Figure 7 provides data from a BCA assay for the melanoma model. The level of total amount of protein detected in the protein corona from tumor-bearing animals following *in vivo*

PREOMICS EXHIBIT 1002
Page 562 of 1861

administration is significantly greater than after *in vitro* incubation. Furthermore, there is a clear distinction between tumor and healthy mice in regards to the level of total amount of protein detected in the protein corona following *in vivo* administration.

Figure 8 shows the SDS-Page gel for protein recovered directly from plasma before and after infusion with Caelx and compares this with analysis conducted following either *in vivo* administration of nanoparticles or after *in vitro* incubation of nanoparticles with plasma taken from the human patients and shows that the masking problem with protein recovered directly from plasma is not present.

Figure 9 provides a breakdown of the top 10 proteins identified in the nanoparticle protein corona based on Relative Protein Abundance (represents the percentage of each protein relative to the total amount of protein purified from the protein corona of the nanoparticles) following *in vivo* administration to 6 patients. Figure 9 also provides a list of the top 10 most abundant proteins found in plasma taken from one of the patients (Patient 1). The data demonstrates that the methods of the invention are able to extract out different types of proteins from the patients that are not readily detectable in plasma.

Figure 10 provides a breakdown of the top 200 proteins identified in the nanoparticle protein corona based on the RPA% following *in vivo* administration to 6 patients (RPA% data is based on the mean of the six patients). The data further demonstrates that the methods of the invention are able to extract out different types of proteins from the patients that are not readily detectable in plasma.

Figure 11 provides the results for the BCA assay for the human patient study. The level of total amount of protein detected in the protein corona from patients following *in vivo* administration is significantly greater than after *in vitro* incubation of the nanoparticles in plasma taken from such patients.

METHOD FOR IDENTIFYING NEW BIOMARKERS

[00015] Methods provided by the present invention include those described generally above, and are further illustrated by all of the particular method steps disclosed herein.

[00016] The present invention relates to a method of identifying a biomarker from a biofluid, wherein the method comprises:

a. administering a plurality of nanoparticles to a subject in a diseased state to allow a biomolecule corona to form on the surface of said nanoparticles or incubating a plurality of nanoparticles in a biofluid sample taken from a subject in a diseased state to allow a biomolecule corona to form on the surface of said nanoparticles;

b. isolating the nanoparticles and surface-bound biomolecule corona; and

PREOMICS EXHIBIT 1002
Page 563 of 1861

c. analyzing the biomolecule corona to identify the said biomarker.

[00017] The present invention relates to a method of identifying a biomarker from a biofluid. In a particular embodiment of the present invention, the biomarker comprises at least one biomolecule (for example, a protein, peptide, fatty acid, lipid, amino acid, sugar, amide or nucleic acid) and the biofluid is selected from plasma, urine, saliva, lacrimal, cerebrospinal and occular fluids. In a particular embodiment, the biofluid is plasma and the biomarker is at least one protein.

[00018] The present invention involves the use of a plurality of nanoparticles. The methods are applicable to any types of nanoparticles capable of attracting a biomolecule corona. In a particular embodiment, the nanoparticles are selected from liposomes, gold nanoparticles, polymeric nanoparticles, carbon nanotubes and graphene oxide nanoparticles.

[00019] Conveniently, the nanoparticles are liposomes. Liposomes are generally spherical vesicles comprising at least one lipid bilayer. Liposomes are often composed of phospholipids. In a particular embodiment, the liposomes are composed of phospholipid molecules and funcitonalised amphiphilic molecules (eg. PEGylated DSPE) that are able to self-assemble into unilamellar vesicles. Conveniently, the liposomes are able to encapsulate drug molecules in their inner aqueous phase.

[00020] The corona formed on the nanoparticles is a biomolecule corona. The term "biomolecule" in this context includes, but is not limited to, proteins, peptides, fatty acids, lipids, amino acids, amides, sugars and nucleic acids (such as for example different types of DNA or RNA). Conveniently, the biomolecule corona comprises a protein corona.

[00021] Step (a) of the method involves administering a plurality of nanoparticles to a subject in a diseased state to allow a biomolecule corona to form on the surface of said nanoparticles or incubating a plurality of nanoparticles in a biofluid sample taken from a subject in a diseased state to allow a biomolecule corona to form on the surface of said nanoparticles. Where the plurality of nanoparticles are administered to a subject in a diseased state to allow a biomolecule corona to form on the surface of said nanoparticles, administration can be by any route that allows the biomolecule corona to form. Suitable routes of administration include but are not limited to intravenous, oral, intracerebral (including spinal), intraperitoneal and intra-occular. Conveniently, the route of administration is by intravenous injection. The biomolecule corona typically forms within less than 10 minutes from administration.

[00022] In the case where the plurality of nanoparticles are incubated in a biofluid sample taken from a subject in a diseased state, such incubation can be carried out *in-vitro*. Conveniently, this involves incubating at 37°C the nanoparticles with plasma taken from the subject in the diseased state for a minimum period of 10 minutes, although it is possible that

PREOMICS EXHIBIT 1002
Page 564 of 1861

biomolecule corona can form immediately upon incubation. Conveniently, the mixture can be subject to agitation, for example by way of an orbital shaker set at approximately 250rpm to mimic *in vivo* conditions.

[00023] Once the biomolecule corona has formed on said nanoparticles, said particles are isolated. Any isolation technique that is capable of preserving the surface-bound biomolecule corona is suitable. Conveniently, the nanoparticles with surface-bound biomolecule corona are isolated from the biofluid and purified to remove unbound and highly abundant biomolecules (for example albumin and/or immunoglobulins, which can constitute 90% of the plasma proteome) to allow identification of lower abundant biomarkers. The method therefore allows minimization of any masking caused by the highly abundant proteins. Conveniently, the isolation is achieved by a method comprising size exclusion chromatography followed by ultrafiltration. The method offers particular benefits over centrifugation methodologies that favour interaction with proteins. In a particular embodiment, the method allows identification of low molecular weight protein biomarkers. Conveniently, the method allows identification of protein biomarkers with molecular weight of less than 80kDa. More conveniently, the method allows identification of protein biomarkers with molecular weight of less than 40kDa or less than 20kDa. In a particular embodiment of the invention, the beneficial sensistivity and high level of precision provided by the method allows the identification of intracellular protein disease related biomarkers that are present in low abundance and would otherwise be very difficult to identify.

[00024] Analysis of the biomolecule corona in order to identify biomarkers can be carried out using any suitable technique capable of detecting said biomarkers. In a particular embodiment of the invention, the biomolecule corona is analysed by gel electrophoresis, mass spectrometry, an immunoassay, UV-Vis. absorption, fluorescence spectroscopy, chromatography or NMR methodology. Conveniently, the biomolecule corona is analysed by mass spectrometry, which can allow qualitative and/or quantitative analysis of the biomolecule corona present on the nanoparticles. In a particular embodiment, certain methods may allow identification of unique biomolecules without the need for highly specialized and ultra-sensitive analytical mass spectrometry instrumentation such as an UltiMate® 3000 Rapid Separation LC (RSLC, Dionex Corporation, Sunnyvale, CA) coupled to a LTQ Velos Pro (Thermo Fisher Scientific, Waltham, MA) mass spectrometer. In one aspect of this embodiment, analysis of the biomolecule corona is carried out after administering a plurality of nanoparticles to a subject in a diseased state to allow a biomolecule corona to form on the surface of said nanoparticles and isolating the nanoparticles and surface-bound biomolecule corona. When compared to other methods, such methods can yield high levels of unique low abundant biomolecules and allow identification of such unique biomolecules without the need for highly specialized and ultra-sensitive analytical mass

PREOMICS EXHIBIT 1002
Page 565 of 1861

spectrometry instrumentation such as an UltiMate® 3000 Rapid Separation LC (RSLC, Dionex Corporation, Sunnyvale, CA) coupled to a LTQ Velos Pro (Thermo Fisher Scientific, Waltham, MA) mass spectrometer.

[00025] In addition to the identification of a single biomarker, the methods also provide the ability to identify panels of biomarkers (multiplexing). This approach can lead to increased sensitivity and specificity of detection.

[00026] In addition to the identification of new biomarkers, the methods also provide the ability to monitor changes in biomarkers for example in response to therapy. In one particular embodiment the therapy administered to the subject prior to testing is a drug molecule, such as for example, an anti-cancer compound. Suitable anti-cancer compounds include, but are not limited to, compounds with activity in cancers such as lung cancer, melanoma or ovarian cancer. Conveniently, the anti-cancer compound is doxorubicin.

[00027] In a particular embodiment, the invention relates to a method of identifying a biomarker from a biofluid, wherein the method comprises:

a. isolating a plurality of nanoparticles with surface-bound biomolecule corona from a biofluid sample taken from a subject in a diseased state; and
b. analyzing the biomolecule corona to identify the said biomarker.

METHOD OF DETECTING A DISEASED STATE IN A SUBJECT

[00028] The present invention also relates to a method of detecting a diseased state in a subject, wherein the method comprises:

a. administering a plurality of nanoparticles to a subject to allow a biomolecule corona to form on the surface of the nanoparticles;
b. isolating the nanoparticles and surface-bound biomolecule corona; and
c. determining the total biomolecule content of the biomolecule corona, which is determinative of the presence of disease in said subject.

[00029] Surprisingly, inventors have found that the analysis of the total biomolecule content of the biomolecule corona isolated after administering a plurality of nanoparticles to a subject to allow a biomolecule corona to form on the surface of the nanoparticles is indicative of the presence of disease in said subject. In a particular embodiment, the biomolecule is a protein.

[00030] The total biomolecule content of the biomolecule corona can be determined by any method capable of quantifying the level of said biomolecules in the surface-bound corona. In one embodiment, the biomeolecule method involves determining the total protein content and this is suitably determined by a bicinchoninic acid assay (BCA assay). In one particular embodiment, the subject is a human patient and the total protein content is at least 700, 800, 900, 1000, 1250,

PREOMICS EXHIBIT 1002
Page 566 of 1861

1500, 1800, 2000, 25000 or 3000Pb when measured using a BCA assay. In yet a further aspect of this embodiment, the disease state is cancer, such as for example ovarian cancer, lung cancer or melenoma.

[00031] Surprisingly, inventors have also found that the total protein content determined by the inventive method is greater than if determined by incubating the plurality of nanoparticles *in-vitro* with a biofluid taken from the subject. Figure 11 shows data to illustrate this surprising discovery. In a particular embodiment, the total protein content determined is at least between 1.2 and 5 fold higher than if determined by incubating the plurality of nanoparticles *in-vitro* with a biofluid isolated from the subject. Conveniently, total protein content determined is at least 1.5, 1.8, 2, 3, 4 or 5 fold higher than if determined by incubating the plurality of nanoparticles *in-vitro* with a biofluid isolated from the subject. Conveniently, the subject in this embodiment is a human.

[00032] Conveniently, the route of administration of the nanoparticles is by intravenous injection. The biomolecule corona typically forms within less than a few minutes from administration.

[00033] Any isolation technique that is capable of preserving the surface-bound biomolecule corona is suitable. Conveniently, the nanoparticles with surface-bound biomolecule corona are isolated from the biofluid and purified to remove unbound and highly abundant biomolecules (for example albumin) to allow identification of lower abundant biomarkers. The method therefore allows minimization of any masking caused by the highly abundant proteins. Conveniently, the isolation is achieved by a method comprising size exclusion chromatography followed by ultrafiltration.

[00034] In addition to a determination of the total biomolecule content of the biomolecule corona, analysis of the biomolecule corona can also reveal qualitative and quantitative information regarding specific potential biomarkers. Such analysis can be carried out using any suitable techniques of capable of detecting said biomarkers. In a particular embodiment of the invention, the biomolecule corona is analysed by mass spectrometry, an immunoassay, UV-Vis. absorption, fluorescence spectroscopy, chromatography or NMR methodology. Conveniently, the biomolecule corona is analysed by mass spectrometry, which can allow qualitative and quantitative analysis of the biomolecule corona present on the nanoparticles. In a particular embodiment, the methods allow identification of unique biomolecules without the need for highly specialized and ultra-sensitive analytical mass spectrometry instrumentation such as using an UltiMate® 3000 Rapid Separation LC (RSLC, Dionex Corporation, Sunnyvale, CA) coupled to a LTQ Velos Pro (Thermo Fisher Scientific, Waltham, MA) mass spectrometer.

[00035] A further aspect of the invention includes novel biomarkers that are linked with

PREOMICS EXHIBIT 1002
Page 567 of 1861

particular diseases, such as ovarian cancer, lung cancer or melanoma and could facilitate personalized healthcare and patient selection and stratification strategies for therapy. Exampes of such biomarkers are further described in the experimental examples below.

EXAMPLES

Further examples of the invention are described hereinbelow, by way of example only, with reference to the accompanying figures.

Example 1. Xenograft mice models

Six to eight week old male nude SCID beige mice were purchased from Charles River (UK). Five to six week old female C57BL/6 mice were purchased from Charles River (UK). Animal procedures were performed in compliance with the UK Home Office Code of Practice for the Housing and Care of Animals used in Scientific Procedures. Mice were housed in groups of five with free access to water and kept at temperature of 19-22 °C and relative humidity of 45-65%. Before performing the procedures, animals where acclimatized to the environment for at least 7 days.

Lung adenocarcinoma model:

Six to eight weeks old male nude SCID beige mice were intravenously injected (via tail vein) with a549-luc cells (5E6 cell/ 200ul of PBS).

Melanoma model:

Five to six week old female C57BL/6 mice were subcutaneously injected (to the left leg) with 0.5 $\times 10^6$ of B16F10-luc melanoma cells in a volume of 50 µl of PBS.

Preparation of liposome nanoparticles

Materials

Hydrogenated soy phosphatidylcholine (HSPC), 1,2-distearoyl-sn-glycero-3-phosphoethanolamine-N-[methoxy(polyethylene glycol)-2000 (DSPE-PEG2000) were purchased from Avanti Polar Lipids (USA), while cholesterol was purchased from Sigma (UK).

Preparation methods

Liposomes with a composition of (HSPC:CHOL:DSPE-PEG2000 (56.3:38.2:5.5) were prepared by thin lipid film hydration method followed by extrusion. Briefly, lipids were dissolved in chloroform:methanol mixture (4:1) in a total volume of 2 ml, using a 25 ml round bottom flask. Organic solvents were then evaporated using a rotary evaporator (Buchi, Switzerland) at 40 °C, at 150 rotations /min, 1 h under vacuum. Lipid films were hydrated with ammonium sulphate 250 mM (pH 8.5) at 60 °C to produce large multilammelar liposomes. Small unilamellar liposomes were then produced by extrusion though 800 nm and 200 nm polycarbonate filters (Whatman, VWR, UK) 10 times each and then 15 times through 100 nm and 80 nm extrusion filters

PREOMICS EXHIBIT 1002
Page 568 of 1861

(Whatman, VWR, UK) using a mini-Extruder (Avanti Polar Lipids, Alabaster, AL).

Protein corona formation after *in vivo* administration of nanoparticles in mice

Mice were anesthetized by inhalation of isoflurane and liposomes were administered intravenously via the lateral tail vein, at a lipid dose of 0.125mM/g body weight, used for preclinical studies. 10 minutes post- injection, blood was recovered by cardiac puncture using K2EDTA coated blood collection tubes. Plasma was prepared by inverting the collection tubes 10 times to ensure mixing of blood with EDTA and subsequent centrifugation for 12 minutes at 1300 RCF at 4 °C. Supernatant was collected into Protein LoBind Eppendorf Tubes. The plasma samples obtained from three mice were pooled together.

Protein corona formation after *in vitro* incubation with plasma taken from mice

The *in vitro* formed corona was allowed to form using a liposome concentration of 2.25mM, which is equivalent to the typical concentration of liposomes found in plasma after *in vivo* administration to mice as described above. For all *in vitro* protein binding studies, 2.25mM of liposomes (180µl of 12.5mM) were incubated with 820µl of mouse plasma for 10 min at 37°C in orbital shaker at 250rpm setting to mimic in vivo conditions.

Separation of corona-coated nanoparticles from unbound and weakly bound proteins

Nanoparticles recovered from the *in vivo* or *in vitro* experiments described above were separated from excess plasma proteins by size exclusion chromatography followed by membrane ultrafiltration. Immediately after *in vitro* or *in vivo* incubations, 1ml of plasma samples was loaded onto a Sepharose CL-4B (SIGMA-ALDRICH) column (15x1.5cm) equilibrated with HBS. Stewart assay in each chromatographic fraction (1ml) revealed that nanoparticles were eluted in fractions 4, 5 and 6. Fractions containing nanoparticles were then pooled together and concentrated to 500 µl by centrifugation using Vivaspin 6 column (10000 MWCO, Sartorious, Fisher Scientific) at 9000 rpm. Vivaspin 500 centrifugal concentrator (1 000 000 MWCO, Sartorious, Fisher Scientific) was then used at 9000 rpm, to further concentrate the samples to 100 µl and to ensure separation of protein-coated nanoparticles from the remaining large unbound proteins. Nanoparticles were then washed 3 times with 100 µl HBS to remove weekly bound proteins

PREOMICS EXHIBIT 1002
Page 569 of 1861

SDS-PAGE electrophoresis

Samples of protein corona-coated nanoparticles were mixed with 20 ul of Tris-Glycine SDS buffer, with 4ul of NuPAGE REDUCING and with water for a final volume of 40 µl and boiled for 5 minutes at 90°C. Samples were then loaded in 4-20% NOVEX Tris-Glycine Protein Gel (ThermoFisher Scientific). The gel was run for 25-40 minutes at 225V, until the proteins neared the end of the gel, in 10 times diluted Novex® Tris-Glycine SDS Running Buffer (ThermoFisher Scientific). Staining was performed with Imperial Gel Staining reagent (Sigma Life Science) for an hour followed by washing in distilled water for 2 days.

Figure 1 shows the SDS-Page gel for protein recovered directly from plasma (healthy and tumor, i.e. lung adenocarcinoma model) where a large signal is present as a result of highly abundant proteins such as albumin. In contrast, the SDS-Page gel for analysis conducted following either *in vivo* administration of nanoparticles in mice or after *in vitro* incubation with plasma taken from mice, does not exhibit the masking problem and the data reveals the presence of a number of lower abundant proteins.

Similar differences are also evident for the melanoma model (Figure 2).

Mass Spectrometry

Bands of interest were excised from the SDS-Page gel and dehydrated using acetonitrile followed by vacuum centrifugation. Dried gel pieces were reduced with 10 mM dithiothreitol and alkylated with 55 mM iodoacetamide. Gel pieces were then washed alternately with 25 mM ammonium bicarbonate followed by acetonitrile. This was repeated, and the gel pieces dried by vacuum centrifugation. Samples were digested with trypsin overnight at 37°C.
Digested samples were analysed by LC-MS/MS using an UltiMate® 3000 Rapid Separation LC (RSLC, Dionex Corporation, Sunnyvale, CA) coupled to Orbitrap Velos Pro (Thermo Fisher Scientific, Waltham, MA) mass spectrometer. Peptide mixtures were separated using a gradient from 92% A (0.1% FA in water) and 8% B (0.1% FA in acetonitrile) to 33% B, in 44 min at 300 nL min^{-1}, using a 250 mm x 75 µm i.d. 1.7 µM BEH C18, analytical column (Waters). Peptides were selected for fragmentation automatically by data dependant analysis. Data produced were searched using Mascot (Matrix Science UK), against the uniprot_concat2014_201410.fasta database. Data presented in Figure 3 was validated using Scaffold (Proteome Software, Portland, OR) and data presented in Figures 4a, 4b, 5a and 5b were validated using Progenesis QI for Proteomics.

PREOMICS EXHIBIT 1002
Page 570 of 1861

The Scaffold software (version Scaffold_4.4.5, Proteome Software Inc., Portland, OR) was used to validate MS/MS based peptide and protein identifications and for relative quantification based on spectral counting. Peptide identifications were accepted if they could be established at greater than 50.0% probability by the Peptide Prophet algorithm with Scaffold delta-mass correction. Protein identifications were accepted if they could be established at greater than 99.0% probability and contained at least 2 identified peptides. Protein probabilities were assigned by the Protein Prophet algorithm. Proteins that contained similar peptides and could not be differentiated based on MS/MS analysis alone were grouped to satisfy the principles of parsimony.

The improved ability to identify a greater number of proteins is also demonstrated in Figure 3 (lung adenocarcinoma model), which shows the number of proteins identified following *in vivo* administration of nanoparticles in mice or after *in vitro* incubation with mouse plasma when compared with plasma analysis (data presented is the result of analysis conducted with Scaffold software as described above). The data demonstrates that a significantly higher number of corona proteins identified when compared to plasma analysis.

Figure 4 provides a breakdown of the host proteins identified following *in vivo* administration (4a) and after *in vitro* incubation (4b) in the lung adenocarcinoma model, ranked by fold change in amount of that protein detected in the protein corona from healthy and diseased mice (proteins with greater than 5 fold change shown). The data demonstrates that the methods are able to identify changes in proteins between tumor and healthy groups. Particular human tumour specific proteins were also identified in the nanoparticle protein corona of diseased mice in this lung adenocarcinoma model that involves the use of the human dervived a549-luc cancer cell line. This demonstrates the ability of the method to detect proteins secreted by the human tumour. Data presented is the result of analysis conducted with Progenesis QI for Proteomics as described above.

Figure 5 provides a breakdown of the proteins identified following *in vivo* administration (5a) and after *in vitro* incubation (5b) in the melanoma model ranked by fold change in amount of that protein detected in samples from healthy and diseased mice. The data demonstrates that the methods are able to identify changes in proteins between tumor and healthy groups. Data presented is the result of analysis conducted with Progenesis QI for Proteomics as described above.

PREOMICS EXHIBIT 1002
Page 571 of 1861

The accession numbers of the proteins indicated in Figures 4a, 4b, 5a and 5b were assigned using uniprot_concat2014_201410.fasta database.

For the melanoma model study, surface-bound proteins were also classified according to their molecular mass as illustrated in Figure 14. The RPA% values for each molecular weight group represents the average of 3 biological replicates (n=3 mice/replicate). As illustrated in Figure 14, proteins with MW < 60 kDa accounted for approximately 70% of the protein coronas formed, in both healthy and tumor-inoculated mice. Remarkably, analysis of the in vivo protein coronas increased the identification of proteins with MW <40 kDa, in comparison with plasma control analysis.

For the lung carcinoma study, surface-bound proteins were classified according to their molecular mass. A heatmap of RPA (%) of proteins identified in the coronas formed onto liposomes intravenously infused in healthy and lung-carcinoma mice, as identified by LC-MS/MS is shown in Figure 15. Only proteins with RPA>0.5 % are shown. RPA (%) values represent the average of 3 biological replicates (n=3 mice/replicate). Protein coronas formed in SCID mice were mainly composed of low MW proteins (>60% of corona proteins had a MW<60 kDa).

Quantification of adsorbed proteins

Proteins associated with recovered nanoparticles were quantified by BCA Protein assay kit. Pb values, expressed as μg of protein/μM lipid were then calculated. For the BCA assay, a 6-point standard curve was generated by serial dilutions of BSA in HBS, with the top standard at a concentration of 2μg/ml. BCA reagent A and B were mixed at a ratio of 50:1 and 200μl of the BCA mixture were dispensed into a 96-well plate, in duplicates. Then, 25μl of each standard or unknown sample were added per well. The plate was incubated for 30 minutes at 37°C, after which the absorbance was read at 574nm on a plate reader (Fluostar Omega). Protein concentrations were calculated according to the standard curve. To quantify lipid concentration, 20 μl of each samples was mixed with 1ml of chloroform and 500 μl of Stewart assay reagent in an Eppendorf tube. The samples were vortexed for 20 seconds followed by 1 min of centrifugation at 13 000 RPM. 200 μl of the chloroform phase was transferred to a quartz cuvette. The optical density was measured on a using Cary 50 Bio Spectrophotometer (Agilent Technologies) at 485 nm. Lipid concentration was calculated according to a standard curve.

PREOMICS EXHIBIT 1002
Page 572 of 1861

Statistical analysis of the data was performed using IBM SPSS Statistics software. One-way analysis of variance (ANOVA) followed by the Tukey multiple comparison test were used and p values < 0.05 were considered significant.

Figure 6 provides the results for the BCA assay for the lung adenocarcinoma model. Surprsingly, the level of total amount of protein detected in tumor-bearing animals following *in vivo* administration is significantly greater than after *in vitro* incubation. Furthermore, there is a clear distinction between tumor and healthy mice in regards to the level of total amount of protein detected in the protein corona following *in vivo* administration.

A similar effect is observed for the for the melanoma model (Figure 7).

Example 2 Human experiments

Subjects

Eligible patients included women with recurrent ovarian cancer commencing liposomal doxorubicin (Caelyx) treatment for the first time.

Nanoparticles

Dynamic light scattering (DLS), ζ-potential measurements and negative stain transmission electron microscopy (TEM) data showing physicochemical characteristics of the PEGylated doxorubicin-encapsulated liposomes (Caelyx®) employed in this study before and after dosing are summarised in Figure 12.

Dosing and blood sample collection

Patients were intravenously infused with Caelyx (diluted in 5% dextrose) at a dose of 40 mg/m^2 for approximately 1.5 h. Collection of paired plasma samples (i.e. before and after cycle 1 infusion) were collected into commercially available anticoagulant-treated tubes (K2 EDTA BD Vacutainer). Plasma was then prepared by inverting the collection tubes 10 times to ensure mixing of blood with EDTA and subsequent centrifugation for 12 minutes at 1300 RCF at 4 °C. Following centrifugation supernatant was immediately collected into labelled Protein LoBind Eppendorf Tubes and samples were maintained on ice while handling.

Protein corona formation after *in vitro* incubation of nanaoparticles with plasma taken from patients before infusion

The *in vitro* formed corona was allowed to form using a liposome concentration of 0.3mM in plasma taken from patients before infusion, which is equivalent to the concentration of liposomes found in 1ml of recovered plasma after *in vivo* administration to humans as described above. For

PREOMICS EXHIBIT 1002
Page 573 of 1861

all *in vitro* protein binding studies, 0.3mM of liposomes (20µl of 17mM) were incubated with 980µl of human patient plasma for 10 min at 37°C in orbital shaker at 250rpm setting to mimic in vivo conditions.

Separation of corona-coated nanoparticles from unbound and weakly bound proteins

Nanoparticles recovered from the *in vivo* human studies or the *in vitro* incubation method described above were separated from excess plasma proteins by size exclusion chromatography followed by membrane ultrafiltration. Immediately after *in vitro* and *in vivo* incubations, 1ml of plasma samples was loaded onto a Sepharose CL-4B (SIGMA-ALDRICH) column (15x1.5cm) equilibrated with HBS. Stewart assay in each chromatographic fraction (1ml) revealed that nanoparticles were eluted in fractions 4, 5 and 6. Fractions containing nanoparticles were then pooled together and concentrated to 500µl by centrifugation using Vivaspin 6 column (10000 MWCO, Sartorious, Fisher Scientific) at 9000rpm. Vivaspin 500 centrifugal concentrator (1 000 000 MWCO, Sartorious, Fisher Scientific) was then used at 9000 rpm, to further concentrate the samples to 100 µl and to ensure separation of protein-coated nanoparticles from the remaining large unbound proteins. Nanoparticles were then washed 3 times with 100 µl HBS to remove weekly bound proteins.

SDS-PAGE electrophoresis

Samples of protein corona-coated nanoparticles were mixed with 20 µl of Tris-Glycine SDS buffer, with 4ul of NuPAGE REDUCING and with water for a final volume of 40 µl and boiled for 5 minutes at 90°C. Samples were then loaded in 4-20% NOVEX Tris-Glycine Protein Gel (ThermoFisher Scientific). The gel was run for 25-40 minutes at 225V, until the proteins neared the end of the gel, in 10 times diluted Novex® Tris-Glycine SDS Running Buffer (ThermoFisher Scientific). Staining was performed with Imperial Gel Staining reagent (Sigma Life Science) for an hour followed by washing in distilled water for 2 days.

Figure 8 shows the SDS-Page gel for protein recovered directly from plasma before and after infusion with Caelx. A large signal is present as a result of highly abundant proteins such as albumin. In contrast, the SDS-Page gel for analysis conducted following either *in vivo* administration of nanoparticles or after *in vitro* incubation of nanoparticles with plasma taken from the human patients shows that the masking problem is not present. The data shows the presence of a number of lower abundant proteins that would not be detectable by analyzing the plasma directly. Evidently, the amount of this lower abundant protein that would not be detectable by analyzing the plasma directly is higher for the samples analysed after *in vivo*

administration of nanoparticles than the samples analysed after *in vitro* incubation with plasma taken from the human patients. This benefit may further aid detection of potential unique biomolecules by reducing the need for particular highly specialized and ultra-sensitive analytical mass spectrometry instrumentation.

Mass Spectrometry

Bands of interest were excised from the SDS-Page gel and dehydrated using acetonitrile followed by vacuum centrifugation. Dried gel pieces were reduced with 10 mM dithiothreitol and alkylated with 55 mM iodoacetamide. Gel pieces were then washed alternately with 25 mM ammonium bicarbonate followed by acetonitrile. This was repeated, and the gel pieces dried by vacuum centrifugation. Samples were digested with trypsin overnight at 37°C.

Digested samples were analysed by LC-MS/MS using an UltiMate® 3000 Rapid Separation LC (RSLC, Dionex Corporation, Sunnyvale, CA) coupled to a Q Exactive™ Hybrid Quadrupole-Orbitrap™ (Thermo Fisher Scientific, Waltham, MA) mass spectrometer. Peptide mixtures were separated using a gradient from 92% A (0.1% FA in water) and 8% B (0.1% FA in acetonitrile) to 33% B, in 44 min at 300 nL min^{-1}, using a 250 mm x 75 μm i.d. 1.7 μM BEH C18, analytical column (Waters). Peptides were selected for fragmentation automatically by data dependent analysis. Data produced were searched using Mascot (Matrix Science UK), against the SwissProt_2016_04 databasedatabase. Data were validated using Scaffold (Proteome Software, Portland, OR).

The Scaffold software (version Scaffold_4.4.5, Proteome Software Inc., Portland, OR) was used to validate MS/MS based peptide and protein identifications and for relative quantification based on spectral counting. Peptide identifications were accepted if they could be established at greater than 50.0% probability by the Peptide Prophet algorithm with Scaffold delta-mass correction. Protein identifications were accepted if they could be established at greater than 99.0% probability and contained at least 2 identified peptides. Protein probabilities were assigned by the Protein Prophet algorithm. Proteins that contained similar peptides and could not be differentiated based on MS/MS analysis alone were grouped to satisfy the principles of parsimony. Semi quantitative assessment of the protein amounts was conducted using normalized spectral countings, NSCs, provided by Scaffold Software. The mean value of NSCs obtained in the three experimental replicates for each protein was normalized to the protein MW and expressed as a relative quantity by applying the following equation:

PREOMICS EXHIBIT 1002
Page 575 of 1861

$$MWNSC_k = \frac{(NSC/MW)_k}{\sum_{i=1}^{N}(NSC/MW)_i} \times 100 \qquad (1)$$

where, MWNSCk is the percentage molecular weight normalized NSC for protein k and MW is the molecular weight in kDa for protein k. This equation takes into consideration the protein size and evaluates the contribution of each protein reflecting its relative protein abundance (RPA).

Figure 9 provides a breakdown of the top 10 proteins identified with respect to percentage abundance following *in vivo* administration to 6 patients. The data demonstrates that the methods are able to extract out different types of proteins from the patients that are not readily detectable in plasma but are connected with cancer. Figure 10 provides an extended list of proteins identified.

The accession numbers of the proteins indicated in Figures 9 and 10 were assigned using SwissProt_2016_04 database.

Surface-bound proteins were also classified according to their molecular mass as illustrated in Figure 13. The percentage of relative protein abundance (%RPA) for each molecular weight group represent the average from 6 ovarian carcinoma patients. Plasma proteins with MW < 80 kDA accounted for almost 80% of the protein coronas formed. It is possible that the low MW proteins identified (Figure 13) have high affinity and interact directly with the surface of PEGylated liposomes and/or they are trapped between other corona-carrier proteins that are adhered to the nanoparticles surface.

Quantification of adsorbed proteins

Proteins associated with recovered nanoparticles were quantified by BCA Protein assay kit. Pb values, expressed as μg of protein/μM lipid were then calculated. For the BCA assay, a 6-point standard curve was generated by serial dilutions of BSA in HBS, with the top standard at a concentration of 2μg/ml. BCA reagent A and B were mixed at a ratio of 50:1 and 200μl of the BCA mixture were dispensed into a 96-well plate, in duplicates. Then, 25μl of each standard or unknown sample were added per well. The plate was incubated for 30 minutes at 37°C, after which the absorbance was read at 574nm on a plate reader (Fluostar Omega). Protein concentrations were calculated according to the standard curve. To quantify lipid concentration, 20 μl of each samples was mixed with 1ml of chloroform and 500 μl of Stewart assay reagent in

PREOMICS EXHIBIT 1002
Page 576 of 1861

an Eppendorf tube. The samples were vortexed for 20 seconds followed by 1 min of centrifugation at 13 000 RPM. 200 µl of the chloroform phase was transferred to a quartz cuvette. The optical density was measured on a using Cary 50 Bio Spectrophotometer (Agilent Technologies) at 485 nm. Lipid concentration was calculated according to a standard curve.

Statistical analysis of the data was performed using IBM SPSS Statistics software. One-way analysis of variance (ANOVA) followed by the Tukey multiple comparison test were used and p values < 0.05 were considered significant.

Figure 11 provides the results for the BCA assay for the human patient study. The level of total amount of protein detected in the patients following *in vivo* administration is significantly greater than after *in vitro* incubation of the nanoparticles in plasma taken from such patients.

PREOMICS EXHIBIT 1002
Page 577 of 1861

CLAIMS

1. A method of identifying a biomarker from a biofluid, wherein the method comprises:
   a. administering a plurality of nanoparticles to a subject in a diseased state to allow a biomolecule corona to form on the surface of said nanoparticles or incubating a plurality of nanoparticles in a biofluid sample taken from a subject in a diseased state to allow a biomolecule corona to form on the surface of said nanoparticles;
   b. isolating the nanoparticles and surface-bound biomolecule corona; and
   c. analyzing the biomolecule corona to identify the said biomarker.
2. A method of identifying a biomarker from a biofluid according to claim 1, wherein the method comprises:
   a. administering a plurality of nanoparticles to a subject in a diseased state to allow a biomolecule corona to form on the surface of said nanoparticles;
   b. isolating the nanoparticles and surface-bound biomolecule corona; and
   c. analyzing the biomolecule corona to identify the said biomarker.
3. A method of identifying a biomarker from a biofluid according to claim 1 or 2, wherein the nanoparticles are selected from liposomes, metallic nanoparticles (such as gold or silver), polymeric nanoparticles, fibre-shaped nanoparticles (such as carbon nanotubes and two dimensional nanoparticles such as graphene oxide nanoparticles.
4. A method of identifying a biomarker from a biofluid according to claim 1 or 2, wherein the nanoparticles are liposomes.
5. A method of identifying a biomarker from a biofluid according to any one of claims 1 to 4, wherein the nanoparticles with surface-bound biomolecule corona are isolated from the biofluid and purified to remove unbound and highly abundant biomolecules to allow identification of low abundant biomarkers.
6. A method according to claim 5, wherein the nanoparticles with surface-bound biomolecule corona are isolated from the biofluid and purified to remove unbound and highly abundant biomolecules by a method comprising size exclusion chromatography followed by ultrafiltration.
7. A method of identifying a biomarker from a biofluid according to any one of claims 1 to 6, wherein the biomolecule corona is analysed by mass spectrometry, an immunoassay, UV-Vis. absorption, fluorescence spectroscopy, chromatography or NMR methodology.

PREOMICS EXHIBIT 1002
Page 578 of 1861

8. A method of identifying a biomarker from a biofluid according to claim 7, wherein the biomolecule corona is analysed by mass spectrometry.

9. A method of identifying a biomarker from a biofluid according to any one of claims 1 to 8, wherein a change in a biomarker in response to therapy is monitored.

10. A method according to any one of claim 9, wherein the therapy administered to the subject prior to testing is a drug molecule.

11. A method according to any one of claim 10, wherein the drug molecule is an anti-cancer compound.

12. A method according to any one of claims 1 to 11, wherein the biomarker is a multiplex panel of disease-specific biomolecule biomarkers.

13. A method according to any one of claims 1 to 12, wherein the biomarker is a unique biomolecule, meaning that it is a biomolecule that would not have been detected if analysis was carried out directly on biofluid, such as plasma, isolated from the subject.

14. A method of detecting a diseased state in a subject, wherein the method comprises:
    a. administering a plurality of nanoparticles to a subject to allow a biomolecule corona to form on the surface of the nanoparticles;
    b. isolating the nanoparticles and surface-bound biomolecule corona; and
    c. determining the total biomolecule content of the biomolecule corona, which is determinative of the presence of disease in said subject.

15. A method of detecting a diseased state in a subject according to claim 14, wherein the biomolecule is a protein.

16. A method of detecting a diseased state in a subject according to claim 15, wherein the total protein content is determined by a bicinchoninic acid assay (BCA assay).

17. A method of detecting a diseased state in a subject according to claim 15 or 16, wherein the subject is a human and the total protein content is at least 700, 800, 900, 1000, 1250, 1500, 1800, 2000, 25000 or 3000Pb.

18. A method of detecting a diseased state in a subject according to claim 16 or 17, wherein the total protein content determined is greater than if determined by incubating the plurality of nanoparticles *in-vitro* with a biofluid taken from the subject.

19. A method according to any one of claims 14 to 18, wherein the total protein content determined is at least between 1.2 and 5 fold higher than if determined by incubating the plurality of nanoparticles *in-vitro* with a biofluid isolated from the subject.

PREOMICS EXHIBIT 1002
Page 579 of 1861

20. A method according to any one of claims 19, wherein the total protein content determined is at least 1.5, 1.8, 2, 3, 4 or 5 fold higher than if determined by incubating the plurality of nanoparticles *in-vitro* with a biofluid isolated from the subject.

21. A method according to anyone one of claims 1 to 20, wherein the nanoparticles are administered by intravenous injection.

22. A method according to anyone one of claims 1 to 21, wherein the said disease is cancer.

23. A method according to claims 22, wherein the said disease is lung cancer, melanoma or ovarian cancer.

PREOMICS EXHIBIT 1002
Page 580 of 1861

Figure 1

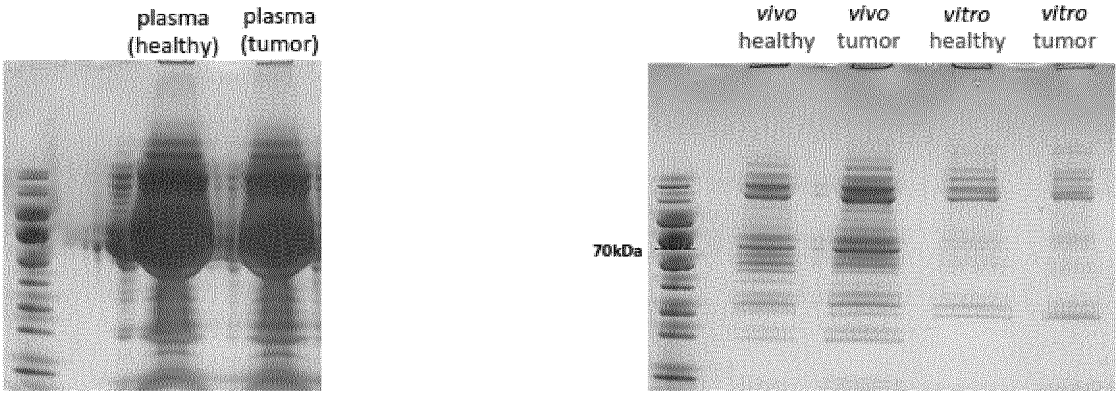

SDS-Page
plasma (healthy)
plasma (tumor)
vivo healthy
vivo tumor
vitro healthy
vitro tumor
70kDa

Figure 2

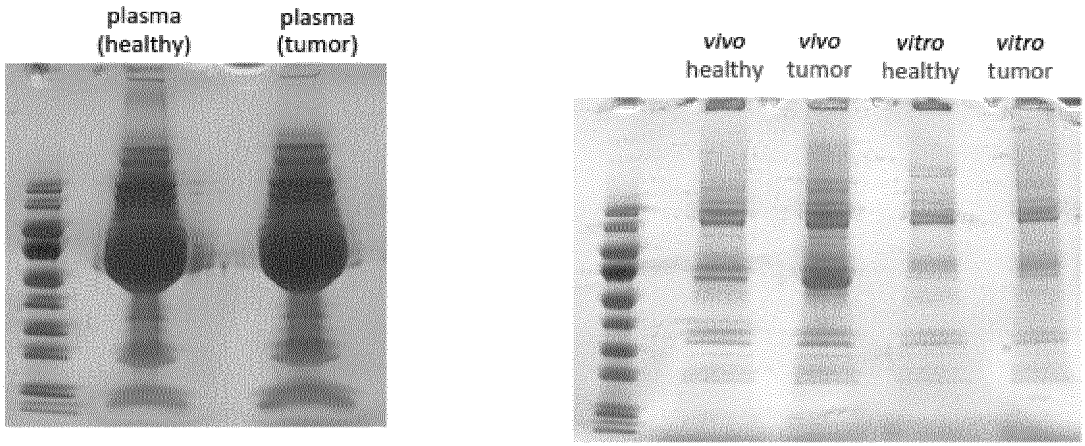

SDS-Page
plasma (healthy)
plasma (tumor)
vivo healthy
vivo tumor
vitro healthy
vitro tumor

Figure 3

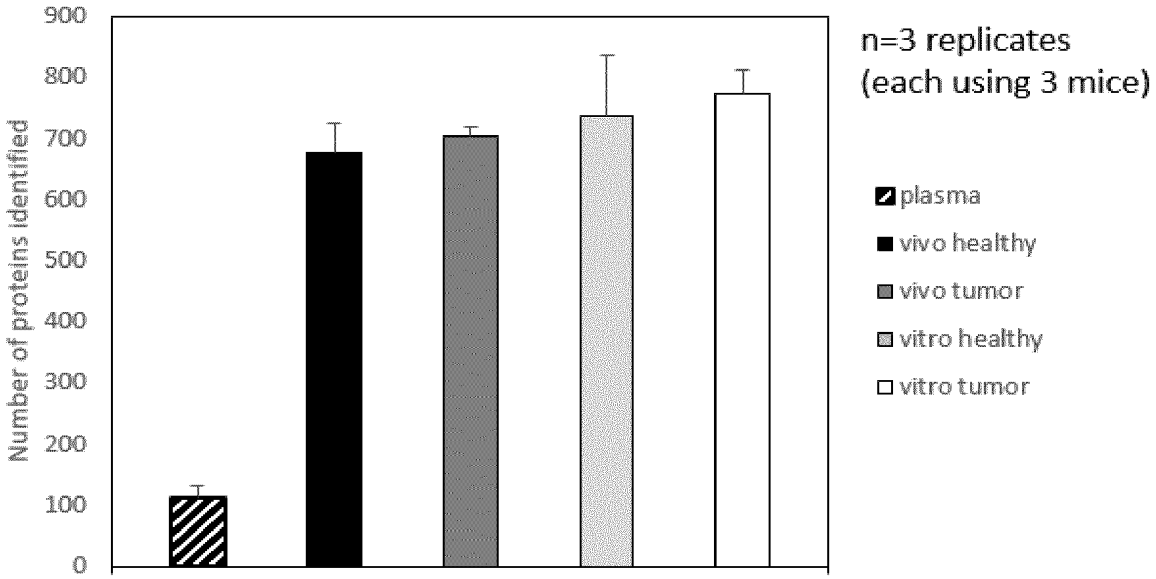

900
800
700
600
500
400
300
200
100
0
Number of proteins identified
n=3 replicates
(each using 3 mice)
plasma
vivo healthy
vivo tumor
vitro healthy
vitro tumor

Figure 4a

Max fold change	Highest mean condition	Lowest mean condition	Description	Accession
489.35	Healthy	Tumour	Protein Igkv15-103	tr|A0A087WNI5|A0A087WNI5_MOUSE
72.28	Tumour	Healthy	Protein Ighv1-81	tr|A0A075B5Y4|A0A075B5Y4_MOUSE;tr|A0A075B5U8|A0A075B5U8_MOUSE;tr|X5J5S9|X5J5S9_MOUSE
56.76	Tumour	Healthy	Ig kappa chain V-III region PC 7043	sp|P01665|KV3AD_MOUSE;sp|P01654|KV3A1_MOUSE;sp|P01664|KV3AC_MOUSE;tr|A0A087WNK6|A0A087WNK6_MOUSE;tr|A0A087WNK7|A0A087WNK7_MOUSE
46.47	Tumour	Healthy	Protein Igkv2-109	tr|A0A075B5K6|A0A075B5K6_MOUSE
41.48	Tumour	Healthy	Protein Ighv5-9-1	tr|A0A075B5Q6|A0A075B5Q6_MOUSE;tr|K7T935|K7T935_MOUSE
41.39	Tumour	Healthy	Complement factor B (Fragment)	tr|H3BK95|H3BK95_MOUSE
24.24	Tumour	Healthy	Protein Igkv4-79	tr|A0A075B5L8|A0A075B5L8_MOUSE
21.88	Tumour	Healthy	Ig kappa chain V-III region ABPC 22/PC 9245	sp|P01662|KV3AA_MOUSE
19.81	Tumour	Healthy	Aldehyde dehydrogenase family 16 member A1	tr|D3Z0B9|D3Z0B9_MOUSE;tr|F6RQF0|F6RQF0_MOUSE;tr|Q9CWT9|Q9CWT9_MOUSE
18.33	Healthy	Tumour	Ig gamma-2A chain C region secreted form	sp|P01864|GCAB_MOUSE
13.19	Healthy	Tumour	Protein Igha	tr|A0A075B6A3|A0A075B6A3_MOUSE
12.99	Tumour	Healthy	Pterin-mimicking anti-idiotope heavy chain variable region (Fragment)	tr|Q920E7|Q920E7_MOUSE
10.68	Tumour	Healthy	Ig mu chain C region	sp|P01872|IGHM_MOUSE
10.68	Tumour	Healthy	Ig mu chain C region	tr|A0A075B5P6|A0A075B5P6_MOUSE;tr|A0A075B6A0|A0A075B6A0_MOUSE
10.42	Tumour	Healthy	Tripeptidyl-peptidase 2	sp|Q64514|TPP2_MOUSE;tr|A0A087WQR6|A0A087WQR6_MOUSE
10.04	Tumour	Healthy	IgG1 heavy chain VDJ region (Fragment)	tr|X5J541|X5J541_MOUSE
9.38	Tumour	Healthy	Signal transducer and activator of transcription	tr|Q3TW11|Q3TW11_MOUSE
8.89	Tumour	Healthy	Protein 4732456N10Rik	tr|E9Q1Z0|E9Q1Z0_MOUSE
8.74	Healthy	Tumour	IgM heavy chain VDJ region (Fragment)	tr|X5J5N0|X5J5N0_MOUSE;sp|P01821|HVM45_MOUSE
8.33	Tumour	Healthy	Ig heavy chain V region 23	sp|P01748|HVM04_MOUSE
8.24	Tumour	Healthy	Betaine--homocysteine S-methyltransferase 1	sp|O35490|BHMT1_MOUSE
8.22	Tumour	Healthy	Glycogen phosphorylase, brain form	sp|Q8CI94|PYGB_MOUSE
7.12	Tumour	Healthy	IgM heavy chain VDJ region (Fragment)	tr|X5J5A3|X5J5A3_MOUSE
6.72	Healthy	Tumour	Fibronectin	tr|Q4KL80|Q4KL80_MOUSE
6.14	Tumour	Healthy	Protein 4.1	sp|P48193|41_MOUSE;tr|F7BUB8|F7BUB8_MOUSE
5.92	Tumour	Healthy	Fructose-bisphosphate aldolase	tr|A6ZI46|A6ZI46_MOUSE
5.72	Tumour	Healthy	Cytochrome P450, family 2, subfamily d, polypeptide 40	tr|Q5M8Q6|Q5M8Q6_MOUSE
5.71	Tumour	Healthy	Proteoglycan 4	tr|E0CZ58|E0CZ58_MOUSE
5.65	Healthy	Tumour	Carboxypeptidase	tr|G3X8T3|G3X8T3_MOUSE
5.48	Tumour	Healthy	B9-scFv	tr|A2NN81|A2NN81_MOUSE
5.43	Tumour	Healthy	H-2 class I histocompatibility antigen, alpha chain (Fragment)	sp|P01895|HA1Y_MOUSE
5.42	Healthy	Tumour	Putative uncharacterized protein	tr|Q9DBN0|Q9DBN0_MOUSE
5.23	Tumour	Healthy	Light chain variable region (Fragment)	tr|A2NVX0|A2NVX0_MOUSE
5.16	Tumour	Healthy	Liver carboxylesterase 1	sp|Q8VCC2|EST1_MOUSE

Figure 4b

Max fold change	Highest mean condition	Lowest mean condition	Description	Accession
74.82	Tumour	Healthy	Protein Igkv13-84	tr|A0A087WNI8|A0A087WNI8_MOUSE
55.62	Healthy	Tumour	Putative uncharacterized protein	tr|Q9D8L4|Q9D8L4_MOUSE;sp|P06327|HVM52_MOUSE;tr|A0A075B5T4|A0A075B5T4_MOUSE
25.83	Healthy	Tumour	Protein BC094916	tr|E9PVL3|E9PVL3_MOUSE
25.19	Tumour	Healthy	Ig kappa chain V-VI region XRPC 44	sp|P01675|KV6A1_MOUSE
19.86	Tumour	Healthy	Anti-H5N1 hemagglutinin monoclonal anitbody H5M9 heavy chain (Fragment)	tr|U5LP42|U5LP42_MOUSE
19.46	Tumour	Healthy	Putative uncharacterized protein	tr|Q3UEW2|Q3UEW2_MOUSE
15.26	Tumour	Healthy	Putative uncharacterized protein (Fragment)	tr|Q3UH12|Q3UH12_MOUSE
13.93	Tumour	Healthy	Filamin-A (Fragment)	tr|F7AVL7|F7AVL7_MOUSE
11.70	Tumour	Healthy	Calcium-activated chloride channel regulator 1	sp|Q9D7Z6|CLCA1_MOUSE
11.55	Healthy	Tumour	Mucin 13, epithelial transmembrane, isoform CRA_a	tr|Q3V1S6|Q3V1S6_MOUSE
10.31	Healthy	Tumour	Immnuoglobulin kappa light chain (Fragment)	tr|A2NVE9|A2NVE9_MOUSE
10.25	Tumour	Healthy	Cytochrome b5	tr|E0CY88|E0CY88_MOUSE
9.51	Tumour	Healthy	E2/E3 hybrid ubiquitin-protein ligase UBE2O	sp|Q6ZPJ3|UBE2O_MOUSE
9.45	Healthy	Tumour	Putative uncharacterized protein	tr|Q3UGS0|Q3UGS0_MOUSE
9.00	Healthy	Tumour	Fibronectin	tr|Q4KL80|Q4KL80_MOUSE
8.76	Healthy	Tumour	Putative uncharacterized protein	tr|Q3UEM7|Q3UEM7_MOUSE
8.70	Tumour	Healthy	Protein Ighv8-8	tr|A0A075B5W8|A0A075B5W8_MOUSE
8.65	Healthy	Tumour	Leukemia inhibitory factor receptor	sp|P42703|LIFR_MOUSE
8.56	Healthy	Tumour	Pterin-mimicking anti-idiotope heavy chain variable region (Fragment)	tr|Q920E7|Q920E7_MOUSE
8.43	Healthy	Tumour	Fibrinogen, B beta polypeptide, isoform CRA_a	tr|Q3TGR2|Q3TGR2_MOUSE
8.41	Tumour	Healthy	alpha-1,2-Mannosidase	tr|Q544T7|Q544T7_MOUSE
8.37	Healthy	Tumour	Keratin, type II cytoskeletal 72	sp|Q6IME9|K2C72_MOUSE
8.25	Tumour	Healthy	Sarcoplasmic/endoplasmic reticulum calcium ATPase 1	sp|Q8R429|AT2A1_MOUSE
8.17	Healthy	Tumour	Putative uncharacterized protein	tr|Q3V3W7|Q3V3W7_MOUSE;tr|Q3TQH1|Q3TQH1_MOUSE
8.02	Healthy	Tumour	Oxidation resistance protein 1	tr|D3YXF5|D3YXF5_MOUSE
7.69	Tumour	Healthy	Protein Ighv1-47	tr|A0A075B5V8|A0A075B5V8_MOUSE
7.68	Tumour	Healthy	Gamma actin-like protein	tr|Q9QZ83|Q9QZ83_MOUSE
7.63	Healthy	Tumour	Fibrinogen alpha chain	sp|E9PV24|FIBA_MOUSE;tr|Q08284|Q08284_MOUSE
7.49	Tumour	Healthy	Ribosomal protein L15	tr|Q3U7D2|Q3U7D2_MOUSE
7.17	Tumour	Healthy	40S ribosomal protein S8	tr|Q5M9L9|Q5M9L9_MOUSE;tr|Q3UA25|Q3UA25_MOUSE
7.02	Healthy	Tumour	Desmin	sp|P31001|DESM_MOUSE
7.00	Tumour	Healthy	Serum amyloid A protein	tr|Q64454|Q64454_MOUSE
6.87	Tumour	Healthy	Uncharacterized protein	tr|F6YVP7|F6YVP7_MOUSE
6.72	Healthy	Tumour	Protein Igkv4-79	tr|A0A075B5L8|A0A075B5L8_MOUSE
6.70	Tumour	Healthy	UbiE1	tr|Q1XG79|Q1XG79_MOUSE;tr|H3BJI7|H3BJI7_MOUSE
6.25	Healthy	Tumour	Putative uncharacterized protein	tr|Q3U6I5|Q3U6I5_MOUSE
6.19	Tumour	Healthy	Charged multivesicular body protein 4b	sp|Q9D8B3|CHM4B_MOUSE
6.06	Healthy	Tumour	Putative uncharacterized protein (Fragment)	tr|Q3TRH4|Q3TRH4_MOUSE
6.03	Tumour	Healthy	60S ribosomal protein L36	tr|Q5M9L1|Q5M9L1_MOUSE
5.92	Tumour	Healthy	Microsomal glutathione S-transferase 1	tr|D3YU60|D3YU60_MOUSE
5.90	Healthy	Tumour	C4a protein	tr|B9EIU2|B9EIU2_MOUSE
5.61	Tumour	Healthy	3-keto-steroid reductase	tr|Q8C5N9|Q8C5N9_MOUSE
5.56	Tumour	Healthy	Myosin-9	sp|Q8VDD5|MYH9_MOUSE;tr|A2VCK0|A2VCK0_MOUSE;tr|Q14B05|Q14B05_MOUSE;tr|Q3UFT0|Q3UFT0_MOUSE;tr|Q7TQJ6|Q7TQJ6_MOUSE;tr|Q811J9|Q811J9_MOUSE
5.43	Tumour	Healthy	Alpha-1 type I procollagen (Fragment)	tr|Q60785|Q60785_MOUSE
5.40	Tumour	Healthy	Putative uncharacterized protein	tr|Q3U944|Q3U944_MOUSE
5.39	Healthy	Tumour	N-acetylglucosaminyl-phosphatidylinositol biosynthetic protein	sp|Q64323|PIGA_MOUSE
5.32	Healthy	Tumour	Glutathione peroxidase 3	sp|P46412|GPX3_MOUSE;tr|D3Z2Y7|D3Z2Y7_MOUSE
5.28	Healthy	Tumour	Protein Ces1a	tr|E9PYP1|E9PYP1_MOUSE
5.27	Tumour	Healthy	7-alpha-hydroxycholest-4-en-3-one 12-alpha-hydroxylase	sp|O88962|CP8B1_MOUSE
5.23	Tumour	Healthy	Syntaxin-binding protein 3	sp|Q60770|STXB3_MOUSE
5.12	Tumour	Healthy	Bile acyl-CoA synthetase	sp|Q4LDG0|S27A5_MOUSE
5.10	Tumour	Healthy	Ribosomal protein L7A	tr|Q6P1A9|Q6P1A9_MOUSE
5.07	Healthy	Tumour	60S ribosomal protein L18 (Fragment)	tr|G3UYV6|G3UYV6_MOUSE

Figure 5a

Max fold change	Highest mean condition	Lowest mean condition	Description	Accession
Infinity	Healthy	Tumour	Beta-actin (Fragment)	tr|A1E281|A1E281_MOUSE
50.25	Tumour	Healthy	Hemoglobin subunit beta-2	sp|P02089|HBB2_MOUSE;tr|A8DV59|A8DV59_MOUSE;tr|D4N6N2|D4N6N2_MUSMC
43.39	Healthy	Tumour	Fibrinogen alpha chain	sp|E9PV24|FIBA_MOUSE
42.75	Healthy	Tumour	Putative uncharacterized protein	tr|Q3UEM7|Q3UEM7_MOUSE
17.95	Healthy	Tumour	Cytochrome P450 2F2	sp|P33267|CP2F2_MOUSE
17.24	Healthy	Tumour	Fibrinogen, B beta polypeptide, isoform CRA_a	tr|Q3TGR2|Q3TGR2_MOUSE
14.18	Healthy	Tumour	Keratin (Fragment)	tr|Q61764|Q61764_MOUSE
12.48	Healthy	Tumour	Precursor polypeptide (AA -10 to 121) (413 is 1st base in codon) (Fragment)	tr|A2NV19|A2NV19_MOUSE
12.23	Tumour	Healthy	Alpha-1-antitrypsin 1-1	sp|P07758|A1AT1_MOUSE;tr|D3YZL0|D3YZL0_MOUSE
11.36	Healthy	Tumour	Protein Ighv1-20	tr|A0A075B5U6|A0A075B5U6_MOUSE
9.86	Healthy	Tumour	Keratin, type II cytoskeletal 4	sp|P07744|K2C4_MOUSE
9.69	Healthy	Tumour	Putative uncharacterized protein	tr|Q9DCD9|Q9DCD9_MOUSE
9.31	Tumour	Healthy	Carboxypeptidase Q	sp|Q9WVJ3|CBPQ_MOUSE
9.13	Healthy	Tumour	EP3-31 light chain variable region (Fragment)	tr|B5UB71|B5UB71_MOUSE
8.87	Tumour	Healthy	Ugt2b36 protein	tr|B7ZP03|B7ZP03_MOUSE;tr|D3YUP6|D3YUP6_MOUSE
8.06	Healthy	Tumour	Protein Ighv1-85	tr|A0A075B5Y6|A0A075B5Y6_MOUSE
8.02	Healthy	Tumour	Beta-actin-like protein 2	sp|Q8BFZ3|ACTBL_MOUSE
7.96	Healthy	Tumour	Protein Igkv8-24	tr|A0A087WNJ5|A0A087WNJ5_MOUSE;tr|A2N1N1|A2N1N1_MOUSE
7.50	Healthy	Tumour	VH7183-DSP2-JH3-CH1 protein (Fragment)	tr|Q65ZL8|Q65ZL8_MOUSE
7.41	Healthy	Tumour	Heat shock protein 90kDa beta (Grp94), member 1	tr|Q3UAD6|Q3UAD6_MOUSE;tr|D3Z1R1|D3Z1R1_MOUSE;tr|F7C312|F7C312_MOUSE;tr|Q3TUD6|Q3TUD6_MOUSE;tr|Q8C2R2|Q8C2R2_MOUSE;tr|Q9CW12|Q9CW12_MOUSE
6.98	Healthy	Tumour	Ankyrin-1	tr|G3UY11|G3UY11_MOUSE
6.95	Tumour	Healthy	Igh protein	tr|Q6PIP8|Q6PIP8_MOUSE;sp|P01863|GCAA_MOUSE
6.75	Healthy	Tumour	Cytochrome P450 2D9	sp|P11714|CP2D9_MOUSE
6.68	Healthy	Tumour	Tubulin alpha-4A chain (Fragment)	tr|A0A087WQS4|A0A087WQS4_MOUSE
6.63	Tumour	Healthy	Ig kappa chain V-VI region XRPC 44	sp|P01675|KV6A1_MOUSE
6.21	Healthy	Tumour	Skeletal muscle alpha-actin mRNA (Fragment)	tr|Q61264|Q61264_MOUSE
5.92	Tumour	Healthy	LOC676837 protein	tr|B2RY37|B2RY37_MOUSE
5.88	Tumour	Healthy	Putative uncharacterized protein	tr|Q3TK73|Q3TK73_MOUSE
5.61	Healthy	Tumour	Moesin	sp|P26041|MOES_MOUSE;tr|Q05DU4|Q05DU4_MOUSE;tr|Q3UF33|Q3UF33_MOUSE
5.59	Tumour	Healthy	Putative uncharacterized protein	tr|Q8BK73|Q8BK73_MOUSE
5.41	Tumour	Healthy	Erythrocyte band 7 integral membrane protein	sp|P54116|STOM_MOUSE;tr|Q8BPA1|Q8BPA1_MOUSE
5.37	Healthy	Tumour	Ankyrin-1	tr|D3Z5M4|D3Z5M4_MOUSE;tr|B7ZW98|B7ZW98_MOUSE;tr|F7D1P5|F7D1P5_MOUSE
5.25	Tumour	Healthy	Transient receptor potential cation channel subfamily M member 8	sp|Q8R4D5|TRPM8_MOUSE

Figure 5b

Max fold change	Highest mean condition	Lowest mean condition	Description	Accession
Infinity	Tumour	Healthy	Putative uncharacterized protein	tr|Q3V2B9|Q3V2B9_MOUSE
144.58	Healthy	Tumour	Putative uncharacterized protein (Fragment)	tr|Q9CUN8|Q9CUN8_MOUSE
69.50	Healthy	Tumour	Ig kappa chain V-III region PC 2154	sp|P01674|KV3AM_MOUSE
59.33	Tumour	Healthy	Actin, cytoplasmic 2 (Fragment)	tr|G3UZ07|G3UZ07_MOUSE
51.27	Healthy	Tumour	Putative uncharacterized protein	tr|Q3V3W7|Q3V3W7_MOUSE;tr|Q3TQH1|Q3TQH1_MOUSE
47.96	Healthy	Tumour	Putative uncharacterized protein	tr|Q3UER8|Q3UER8_MOUSE
47.04	Healthy	Tumour	Coagulation factor XIII, beta subunit	tr|Q8K0W0|Q8K0W0_MOUSE
41.08	Healthy	Tumour	Putative uncharacterized protein	tr|Q3TB69|Q3TB69_MOUSE
41.07	Healthy	Tumour	Translocation protein SEC62	sp|Q8BU14|SEC62_MOUSE
38.82	Healthy	Tumour	Fibrinogen alpha chain	sp|E9PV24|FIBA_MOUSE
28.06	Healthy	Tumour	Putative uncharacterized protein (Fragment)	tr|Q3UAS8|Q3UAS8_MOUSE
27.46	Healthy	Tumour	Fibrinogen, B beta polypeptide, isoform CRA_a	tr|Q3TGR2|Q3TGR2_MOUSE
25.57	Tumour	Healthy	NAD(P) dependent steroid dehydrogenase-like	tr|Q8VE30|Q8VE30_MOUSE
25.20	Healthy	Tumour	Signal peptide peptidase-like 2A	sp|Q9JJF9|SPP2A_MOUSE
24.18	Tumour	Healthy	Steryl-sulfatase	sp|P50427|STS_MOUSE
23.09	Tumour	Healthy	Peptidyl-prolyl cis-trans isomerase FKBP11	sp|Q9D1M7|FKB11_MOUSE
22.20	Tumour	Healthy	Bax inhibitor 1 (Fragment)	tr|E0CX98|E0CX98_MOUSE
21.52	Tumour	Healthy	MCG140784	tr|Q792Z1|Q792Z1_MOUSE
20.93	Healthy	Tumour	26S protease regulatory subunit 6A	tr|A2AGN7|A2AGN7_MOUSE
19.73	Tumour	Healthy	Pterin-mimicking anti-idiotope heavy chain variable region (Fragment)	tr|Q920E7|Q920E7_MOUSE
19.10	Tumour	Healthy	Igh protein	tr|Q3KQK2|Q3KQK2_MOUSE
17.65	Tumour	Healthy	3 beta-hydroxysteroid dehydrogenase type 5	sp|Q61694|3BHS5_MOUSE
17.50	Tumour	Healthy	3 beta-hydroxysteroid dehydrogenase/Delta 5-->4-isomerase type 6 OS=Mus musculus GN=Hsd3b6 PE=2 SV=4	sp|O35469|3BHS6_MOUSE
16.91	Tumour	Healthy	Fn1 protein (Fragment)	tr|Q99KD0|Q99KD0_MOUSE
15.06	Tumour	Healthy	Microsomal glutathione S-transferase 1	tr|E9QJW0|E9QJW0_MOUSE;tr|D3YVR3|D3YVR3_MOUSE
13.89	Tumour	Healthy	Very-long-chain (3R)-3-hydroxyacyl-CoA dehydratase 2	sp|Q9D3B1|HACD2_MOUSE
13.71	Tumour	Healthy	Protein Igkv4-86	tr|A0A075B5L5|A0A075B5L5_MOUSE
13.26	Healthy	Tumour	Complement component C8 beta chain	sp|Q8BH35|CO8B_MOUSE
13.22	Healthy	Tumour	Keratin (Fragment)	tr|Q61764|Q61764_MOUSE
13.10	Healthy	Tumour	Annexin (Fragment)	tr|F7ANV6|F7ANV6_MOUSE
13.00	Tumour	Healthy	Cytochrome c oxidase subunit 2	tr|A0A023J6F3|A0A023J6F3_MOUSE;tr|A0A023J5Y1|A0A023J5Y1_MOUSE
12.02	Tumour	Healthy	BTB (POZ) domain containing 1	tr|Q5M8N6|Q5M8N6_MOUSE
11.79	Tumour	Healthy	Putative uncharacterized protein (Fragment)	tr|Q3U3J9|Q3U3J9_MOUSE
11.54	Tumour	Healthy	Treml1 protein	tr|B7ZNI7|B7ZNI7_MOUSE
11.42	Tumour	Healthy	Putative uncharacterized protein	tr|Q9CY59|Q9CY59_MOUSE
11.42	Healthy	Tumour	Dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit STT3B OS=Mus musculus GN=Stt3b PE=1 SV=2	sp|Q3TDQ1|STT3B_MOUSE
11.01	Healthy	Tumour	Plectin	sp|Q9QXS1|PLEC_MOUSE
10.41	Tumour	Healthy	Cytochrome P450 2D9	sp|P11714|CP2D9_MOUSE
10.39	Healthy	Tumour	Guanine nucleotide-binding protein subunit alpha-12	sp|P27600|GNA12_MOUSE;tr|Q2NKI3|Q2NKI3_MOUSE
10.38	Tumour	Healthy	Membrane protein, palmitoylated	tr|Q542P4|Q542P4_MOUSE;tr|B7ZCL9|B7ZCL9_MOUSE;tr|B7ZCM0|B7ZCM0_MOUSE;tr|Q3THA8|Q3THA8_MOUSE
10.20	Tumour	Healthy	Serum amyloid A-4 protein	sp|P31532|SAA4_MOUSE
10.13	Tumour	Healthy	Glycine N-methyltransferase	tr|Q5I0T9|Q5I0T9_MOUSE
10.04	Tumour	Healthy	Cytochrome P450, family 2, subfamily b, polypeptide 13	tr|A6H6J2|A6H6J2_MOUSE
9.95	Tumour	Healthy	Cytochrome P450 CYP2C44 OS=Mus musculus GN=Cyp2c44 PE=2 SV=1	tr|Q6IEF7|Q6IEF7_MOUSE;tr|Q3UEM4|Q3UEM4_MOUSE;tr|Q8CIE7|Q8CIE7_MOUSE
9.92	Healthy	Tumour	TBC1 domain family member 8B	sp|A3KGB4|TBC8B_MOUSE
9.87	Tumour	Healthy	Putative uncharacterized protein	tr|Q3UGZ4|Q3UGZ4_MOUSE;tr|Q68FM2|Q68FM2_MOUSE
9.87	Tumour	Healthy	MKIAA0302 protein (Fragment)	tr|Q6A087|Q6A087_MOUSE
9.81	Tumour	Healthy	Signal sequence receptor, delta, isoform CRA_b	tr|Q3TVJ8|Q3TVJ8_MOUSE
9.69	Tumour	Healthy	Cis-retinol androgen dehydrogenase 1	tr|O54909|O54909_MOUSE
9.53	Tumour	Healthy	Polymerase I and transcript release factor	sp|O54724|PTRF_MOUSE;tr|Q3TER0|Q3TER0_MOUSE

Max fold change	Highest mean condition	Lowest mean condition	Description	Accession
9.48	Tumour	Healthy	Carbonic anhydrase 3	sp|P16015|CAH3_MOUSE
9.45	Tumour	Healthy	Dimethylaniline monooxygenase [N-oxide-forming] 1	sp|P50285|FMO1_MOUSE;tr|D3Z0T2|D3Z0T2_MOUSE;tr|Q3UNX7|Q3UNX7_MOUSE;tr|Q8C9C1|Q8C9C1_MOUSE
9.09	Tumour	Healthy	Solute carrier family 22 member 18	sp|Q78KK3|S22AI_MOUSE
9.05	Tumour	Healthy	Protein Cyp2c68	tr|K7N6C2|K7N6C2_MOUSE
8.96	Tumour	Healthy	Putative uncharacterized protein	tr|Q3TEK2|Q3TEK2_MOUSE
8.92	Tumour	Healthy	Delta(14)-sterol reductase	tr|E9Q4M8|E9Q4M8_MOUSE
8.74	Tumour	Healthy	Putative uncharacterized protein	tr|Q3TDT8|Q3TDT8_MOUSE
8.63	Tumour	Healthy	Putative uncharacterized protein	tr|Q3U7E6|Q3U7E6_MOUSE;tr|E0CXD9|E0CXD9_MOUSE
8.54	Tumour	Healthy	Sulfotransferase	tr|Q9R1S5|Q9R1S5_MOUSE
8.52	Healthy	Tumour	Cytoplasmic dynein heavy chain (Fragment)	tr|O08821|O08821_MOUSE
8.50	Healthy	Tumour	Retinol-binding protein 3	sp|P49194|RET3_MOUSE
8.38	Tumour	Healthy	Iodotyrosine dehalogenase 1	sp|Q9DCX8|IYD1_MOUSE
8.28	Tumour	Healthy	Myadm protein	tr|Q0VE46|Q0VE46_MOUSE
8.20	Tumour	Healthy	Putative uncharacterized protein	tr|Q3TXE5|Q3TXE5_MOUSE
8.11	Tumour	Healthy	Fructose-1,6-bisphosphatase 1 OS=Mus musculus GN=Fbp1 PE=4 SV=1	tr|E9Q0T7|E9Q0T7_MOUSE
7.85	Tumour	Healthy	Cytochrome P450 2C29	sp|Q64458|CP2CT_MOUSE;tr|H3BLM0|H3BLM0_MOUSE
7.69	Tumour	Healthy	Putative uncharacterized protein	tr|Q3TPA9|Q3TPA9_MOUSE
7.52	Healthy	Tumour	Ankyrin-2	sp|Q8C8R3|ANK2_MOUSE
7.48	Tumour	Healthy	Solute carrier family 43 member 3	sp|A2AVZ9|S43A3_MOUSE
7.41	Tumour	Healthy	Basal cell adhesion molecule	sp|Q9R069|BCAM_MOUSE
7.40	Tumour	Healthy	Duffy blood group antigen	tr|Q7TSL2|Q7TSL2_MOUSE
7.40	Tumour	Healthy	Protein Ugt2b34	tr|Q8K154|Q8K154_MOUSE;tr|E9PZ30|E9PZ30_MOUSE
7.38	Tumour	Healthy	CSC1-like protein 1	sp|Q91YT8|CSCL1_MOUSE
7.37	Tumour	Healthy	Uncharacterized protein	tr|D3Z6R0|D3Z6R0_MOUSE
7.36	Tumour	Healthy	UDP-glucuronosyltransferase 2A3	sp|Q8BWQ1|UD2A3_MOUSE
7.26	Healthy	Tumour	Keratin, type II cytoskeletal 4	sp|P07744|K2C4_MOUSE
7.11	Healthy	Tumour	Ephrin B1, isoform CRA_a	tr|Q544L9|Q544L9_MOUSE
7.03	Healthy	Tumour	ADP-ribosylation factor 5	sp|P84084|ARF5_MOUSE
7.01	Tumour	Healthy	Hemoglobin subunit beta-2	sp|P02089|HBB2_MOUSE;tr|A8DV59|A8DV59_MOUSE;tr|D4N6N2|D4N6N2_MUSMC
6.98	Tumour	Healthy	3 beta-hydroxysteroid dehydrogenase type 7	sp|Q9EQC1|3BHS7_MOUSE
6.92	Healthy	Tumour	Gamma heavy chain variable region (Fragment)	tr|Q5F2I8|Q5F2I8_MOUSE
6.90	Healthy	Tumour	DNA damage-binding protein 1	sp|Q3U1J4|DDB1_MOUSE
6.85	Tumour	Healthy	Chloride intracellular channel protein 3	tr|A2AJ28|A2AJ28_MOUSE
6.82	Tumour	Healthy	Putative uncharacterized protein	tr|Q9DC85|Q9DC85_MOUSE
6.81	Tumour	Healthy	3-keto-steroid reductase	tr|Q8C5N9|Q8C5N9_MOUSE
6.79	Tumour	Healthy	Methyltransferase-like protein 7B	sp|Q9DD20|MET7B_MOUSE
6.78	Tumour	Healthy	Ankyrin-1	tr|G8JL84|G8JL84_MOUSE;sp|Q02357|ANK1_MOUSE;tr|D6RJ51|D6RJ51_MOUSE;tr|F7D1P5|F7D1P5_MOUSE;tr|Q61304|Q61304_MOUSE
6.76	Tumour	Healthy	Putative uncharacterized protein	tr|Q3UF03|Q3UF03_MOUSE
6.75	Tumour	Healthy	Putative uncharacterized protein	tr|Q3TFP8|Q3TFP8_MOUSE
6.64	Tumour	Healthy	Glycolipid transfer protein domain-containing protein 2	sp|Q8K0R6|GLTD2_MOUSE
6.58	Healthy	Tumour	Protein Ahnak	tr|E9Q616|E9Q616_MOUSE;tr|A0JLR7|A0JLR7_MOUSE
6.58	Healthy	Tumour	AHNAK (Fragment)	tr|Q6UL10|Q6UL10_MOUSE;tr|Q61484|Q61484_MOUSE
6.54	Healthy	Tumour	Cytochrome P450 2A5 (Fragment)	tr|F6SHL3|F6SHL3_MOUSE
6.49	Healthy	Tumour	Protein Ighv7-1	tr|A0A075B5S2|A0A075B5S2_MOUSE
6.48	Tumour	Healthy	Collagen alpha-1(XIV) chain (Fragment)	tr|F7D5Y4|F7D5Y4_MOUSE
6.43	Healthy	Tumour	60S ribosomal protein L7 (Fragment)	tr|F6XI62|F6XI62_MOUSE
6.41	Healthy	Tumour	Protein 9530053A07Rik	tr|E9PVG8|E9PVG8_MOUSE
6.39	Tumour	Healthy	Probable N-acetyltransferase CML2	sp|Q8CHQ9|CMLO2_MOUSE
6.35	Tumour	Healthy	Transmembrane 9 superfamily member 4	sp|Q8BH24|TM9S4_MOUSE
6.28	Healthy	Tumour	Putative uncharacterized protein	tr|Q9CY10|Q9CY10_MOUSE
6.22	Tumour	Healthy	Annexin A5 OS=Mus musculus GN=Anxa5 PE=1 SV=1	sp|P48036|ANXA5_MOUSE
6.19	Healthy	Tumour	von Willebrand factor OS=Mus musculus GN=Vwf PE=1 SV=2	sp|Q8CIZ8|VWF_MOUSE
6.12	Tumour	Healthy	Uncharacterized protein OS=Mus musculus GN=Gm5422 PE=4 SV=1	tr|J3KMQ2|J3KMQ2_MOUSE
6.12	Tumour	Healthy	Solute carrier family 2, facilitated glucose transporter member 2 OS=Mus musculus GN=Slc2a2 PE=1 SV=3	sp|P14246|GTR2_MOUSE

Max fold change	Highest mean condition	Lowest mean condition	Description	Accession
6.07	Healthy	Tumour	Vacuolar protein sorting-associated protein 13A	sp|Q5H8C4|VP13A_MOUSE
6.06	Tumour	Healthy	Cytochrome P450 2D26	sp|Q8CIM7|CP2DQ_MOUSE
6.02	Tumour	Healthy	Putative uncharacterized protein	tr|Q3TJK3|Q3TJK3_MOUSE
5.98	Tumour	Healthy	Putative uncharacterized protein	tr|Q3TK73|Q3TK73_MOUSE
5.96	Tumour	Healthy	IgM heavy chain VDJ region (Fragment)	tr|X5J5L2|X5J5L2_MOUSE
5.90	Healthy	Tumour	Alpha-2-macroglobulin (Fragment) OS=Mus musculus GN=Pzp PE=4 SV=1	tr|D3YUI3|D3YUI3_MOUSE
5.89	Healthy	Tumour	Mannose binding lectin (C), isoform CRA_b	tr|Q3UEK1|Q3UEK1_MOUSE
5.87	Tumour	Healthy	Putative uncharacterized protein	tr|Q3TV47|Q3TV47_MOUSE
5.85	Tumour	Healthy	UDP-glucuronosyltransferase 3A2	sp|Q8JZZ0|UD3A2_MOUSE
5.83	Healthy	Tumour	Eukaryotic translation initiation factor 3 subunit C	sp|Q8R1B4|EIF3C_MOUSE
5.79	Tumour	Healthy	Ras-related protein Rab-1A	sp|P62821|RAB1A_MOUSE;tr|Q3UB66|Q3UB66_MOUSE
5.77	Tumour	Healthy	17-beta-hydroxysteroid dehydrogenase 13	tr|A8Y5N4|A8Y5N4_MOUSE
5.76	Tumour	Healthy	UDP-glucuronosyltransferase 1-6	sp|Q64435|UD16_MOUSE
5.73	Tumour	Healthy	Long-chain-fatty-acid--CoA ligase 5	sp|Q8JZR0|ACSL5_MOUSE;tr|Q9D9S0|Q9D9S0_MOUSE
5.68	Tumour	Healthy	Ig kappa chain V-III region PC 2413	sp|P01657|KV3A5_MOUSE
5.68	Healthy	Tumour	RAB14 protein variant	tr|Q50HX1|Q50HX1_MOUSE
5.59	Healthy	Tumour	Mesothelin	sp|Q61468|MSLN_MOUSE
5.57	Tumour	Healthy	Anti-VIPase light chain variable region (Fragment)	tr|Q8K1F2|Q8K1F2_MOUSE
5.54	Healthy	Tumour	Cytoplasmic dynein 1 heavy chain 1	sp|Q9JHU4|DYHC1_MOUSE;tr|Q3UA73|Q3UA73_MOUSE;tr|Q80U36|Q80U36_MOUSE;tr|Q8C6T5|Q8C6T5_MOUSE
5.51	Tumour	Healthy	3 beta-hydroxysteroid dehydrogenase/Delta 5-->4-isomerase type 3	sp|P26150|3BHS3_MOUSE;tr|B1ARN9|B1ARN9_MOUSE
5.48	Healthy	Tumour	Ankyrin-1 (Fragment)	tr|G3UXF4|G3UXF4_MOUSE
5.46	Tumour	Healthy	14-3-3 protein zeta/delta	sp|P63101|1433Z_MOUSE
5.45	Healthy	Tumour	Ig kappa chain V-IV region S107B	sp|P01680|KV4A1_MOUSE
5.40	Tumour	Healthy	Igh protein	tr|Q5U472|Q5U472_MOUSE
5.39	Tumour	Healthy	NADPH--cytochrome P450 reductase	sp|P37040|NCPR_MOUSE;tr|E9PVT9|E9PVT9_MOUSE;tr|E9Q997|E9Q997_MOUSE;tr|F6R7H8|F6R7H8_MOUSE;tr|Q05DV1|Q05DV1_MOUSE
5.37	Healthy	Tumour	Vesicle transport through interaction with t-SNAREs 1B homolog	tr|Q91XH6|Q91XH6_MOUSE
5.32	Tumour	Healthy	Vitronectin	sp|P29788|VTNC_MOUSE
5.31	Tumour	Healthy	Cytochrome P450 2B9	sp|P12790|CP2B9_MOUSE
5.29	Tumour	Healthy	Cytochrome P450 4A14	sp|O35728|CP4AE_MOUSE
5.28	Tumour	Healthy	Putative uncharacterized protein	tr|Q3UEJ7|Q3UEJ7_MOUSE
5.27	Healthy	Tumour	Ig kappa chain V-VI region XRPC 44	sp|P01675|KV6A1_MOUSE
5.25	Tumour	Healthy	Sarcoplasmic/endoplasmic reticulum calcium ATPase 1	sp|Q8R429|AT2A1_MOUSE;tr|Q3V1C4|Q3V1C4_MOUSE
5.23	Healthy	Tumour	Serine (Or cysteine) peptidase inhibitor, clade F, member 2, isoform CRA_c	tr|Q5ND36|Q5ND36_MOUSE;tr|E9PXE0|E9PXE0_MOUSE;tr|Q5ND35|Q5ND35_MOUSE
5.23	Healthy	Tumour	Heat shock-related 70 kDa protein 2	sp|P17156|HSP72_MOUSE
5.21	Tumour	Healthy	Transmembrane channel-like protein	tr|B1ATB3|B1ATB3_MOUSE
5.21	Tumour	Healthy	Hsd17b2 protein	tr|Q91X35|Q91X35_MOUSE
5.20	Tumour	Healthy	Myosin-9 OS=Mus musculus GN=Myh9 PE=1 SV=4	sp|Q8VDD5|MYH9_MOUSE;tr|A2VCK0|A2VCK0_MOUSE;tr|Q7TQJ6|Q7TQJ6_MOUSE;tr|Q811J9|Q811J9_MOUSE;tr|Q8BIE2|Q8BIE2_MOUSE
5.19	Healthy	Tumour	Putative uncharacterized protein	tr|Q3UHL7|Q3UHL7_MOUSE;tr|Q3UR38|Q3UR38_MOUSE
5.14	Tumour	Healthy	Putative uncharacterized protein	tr|Q3UYU1|Q3UYU1_MOUSE
5.14	Tumour	Healthy	Regucalcin	sp|Q64374|RGN_MOUSE
5.12	Healthy	Tumour	Igh protein	tr|Q4V9Z4|Q4V9Z4_MOUSE
5.10	Tumour	Healthy	MCG17387, isoform CRA_a	tr|Q5FWB6|Q5FWB6_MOUSE;tr|D3YVM5|D3YVM5_MOUSE
5.06	Tumour	Healthy	Cytochrome b-5, isoform CRA_a	tr|G5E850|G5E850_MOUSE
5.05	Tumour	Healthy	Tubulin alpha-4A chain	sp|P68368|TBA4A_MOUSE
5.04	Healthy	Tumour	Aminopeptidase A	tr|Q52JJ6|Q52JJ6_MOUSE
5.03	Tumour	Healthy	Hemoglobin beta chain subunit	tr|B1Q450|B1Q450_MOUSE
5.02	Healthy	Tumour	Coagulation factor X	tr|Q3TBR2|Q3TBR2_MOUSE
5.01	Tumour	Healthy	Starch-binding domain-containing protein 1	sp|Q8C7E7|STBD1_MOUSE
5.00	Tumour	Healthy	Oncoprotein-induced transcript 3 protein	sp|Q8R4V5|OIT3_MOUSE

Figure 6

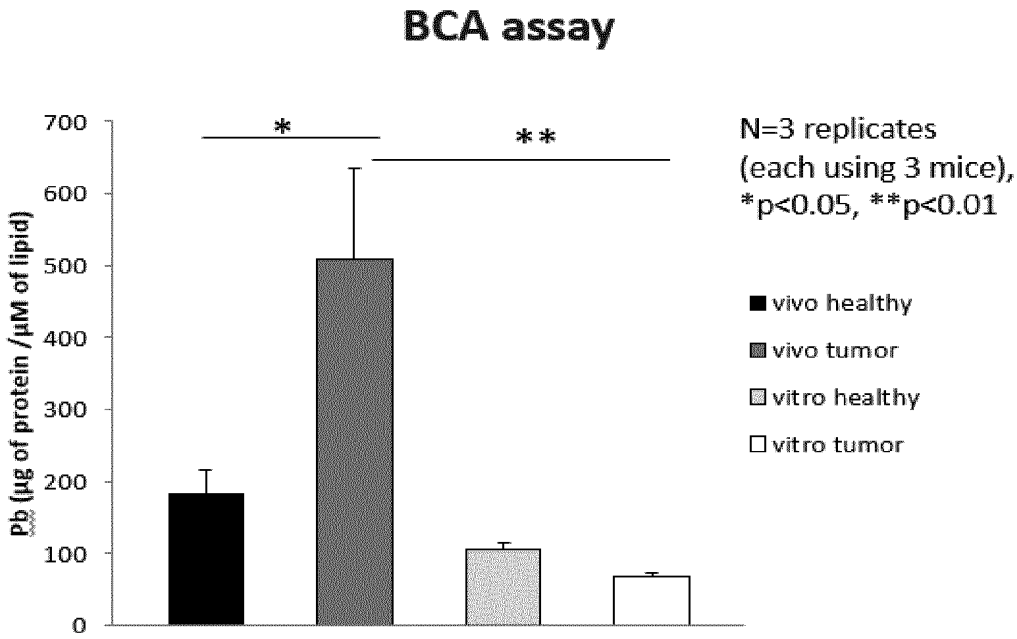

BCA assay
Pb (µg of protein /µM of lipid)
700
600
500
400
300
200
100
0
*
**
N=3 replicates
(each using 3 mice),
*p<0.05, **p<0.01
vivo healthy
vivo tumor
vitro healthy
vitro tumor

Figure 7

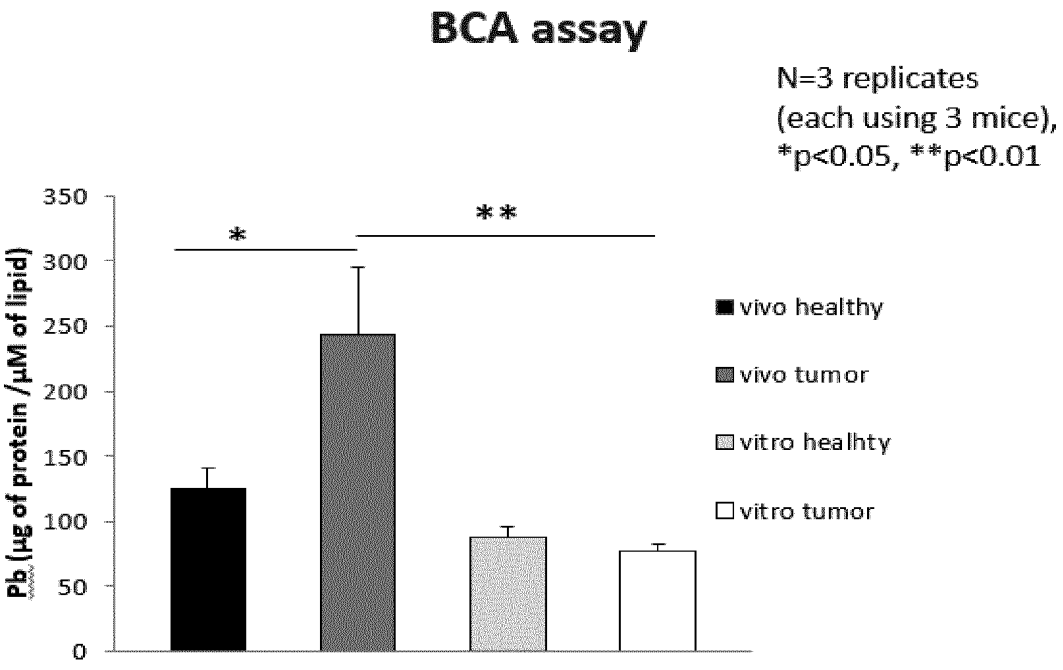

BCA assay
N=3 replicates
(each using 3 mice),
*p<0.05, **p<0.01
Pb (µg of protein /µM of lipid)
350
300
250
200
150
100
50
0
*
**
vivo healthy
vivo tumor
vitro healhty
vitro tumor

Figure 8

SDS-Page

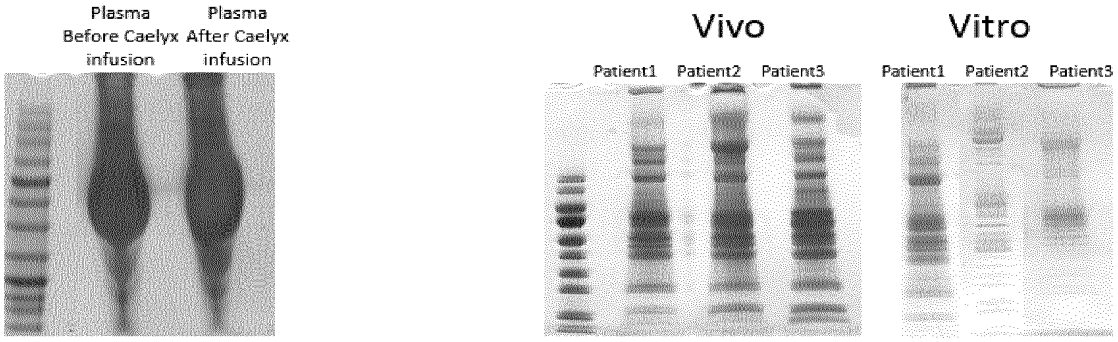


Figure 9

Rank	RPA%	Plasma Proteins by Abundance
1	17.38	Serum Albumin
2	3.92	monoclonal antibody CR8033 light chain
3	3.84	Fab light chain
4	3.77	light chain O12/KAPPA
5	3.59	Putative uncharacterized protein DKFZp686C11235
6	3.48	MSL-109 Light Chain
7	3.48	679-14-14E06 Fab fragment light chain
8	3.41	B3 (IGKV4-1) light chain
9	2.88	IGH protein
10	2.79	cDNA FLJ78387

	Rank	RPA%	Identified Proteins	Accession Number
Patient 1	1	7.68	Full-length cDNA clone CS0DD006YL02 of Neuroblastoma	Q86TT1_HUMAN
	2	5.42	Ig mu chain C region(A0A087X2C0)	A0A087X2C0_HUMAN
	3	5.17	Ig mu chain C region(A0A087WYJ9)	A0A087WYJ9_HUMAN
	4	5.17	Fibrinogen beta chain	FIBB_HUMAN (+1)
	5	3.49	Alpha-2-macroglobulin	A2MG_HUMAN
	6	3.25	Fibrinogen alpha chain	FIBA_HUMAN
	7	2.55	C4b-binding protein alpha chain	C4BPA_HUMAN
	8	1.92	Fibrinogen gamma chain	FIBG_HUMAN
	9	1.84	Protein IGHV3-72	A0A087WW89_HUMAN
	10	1.65	Haptoglobin	H0Y300_HUMAN
Patient 2	1	5.47	Serum albumin	ALBU_HUMAN
	2	4.90	Full-length cDNA clone CS0DD006YL02 of Neuroblastoma	Q86TT1_HUMAN
	3	3.95	Fibrinogen beta chain	FIBB_HUMAN (+1)
	4	3.53	Ig mu chain C region(A0A087X2C0)	A0A087X2C0_HUMAN
	5	3.40	Ig mu chain C region(A0A087WYJ9)	A0A087WYJ9_HUMAN
	6	3.12	Alpha-2-macroglobulin	A2MG_HUMAN
	7	2.70	Fibrinogen alpha chain	FIBA_HUMAN
	8	2.33	Putative uncharacterized protein DKFZp686C11235	Q6MZV7_HUMAN
	9	2.10	C4b-binding protein alpha chain	C4BPA_HUMAN
	10	1.91	Fibrinogen gamma chain	FIBG_HUMAN
Patient 3	1	6.91	Full-length cDNA clone CS0DD006YL02 of Neuroblastoma	Q86TT1_HUMAN
	2	4.81	Ig mu chain C region(A0A087X2C0)	A0A087X2C0_HUMAN
	3	4.72	Serum albumin	ALBU_HUMAN
	4	4.66	Ig mu chain C region(A0A087WYJ9)	A0A087WYJ9_HUMAN
	5	3.60	Fibrinogen beta chain	FIBB_HUMAN (+1)
	6	3.28	Fibrinogen alpha chain	FIBA_HUMAN
	7	2.38	Putative uncharacterized protein DKFZp686C11235	Q6MZV7_HUMAN
	8	2.22	Fibrinogen gamma chain	FIBG_HUMAN
	9	2.11	IGH protein	Q6GMX6_HUMAN
	10	2.03	Ig gamma-1 chain C region	A0A087WV47_HUMAN (+1)
Patient 4	1	7.99	Full-length cDNA clone CS0DD006YL02 of Neuroblastoma	Q86TT1_HUMAN
	2	5.69	Ig mu chain C region(A0A087X2C0)	A0A087X2C0_HUMAN
	3	4.68	Ig mu heavy chain disease protein	MUCB_HUMAN
	4	4.55	Alpha-2-macroglobulin	A2MG_HUMAN
	5	3.64	Lipoprotein B (Fragment)	S5FLF7_HUMAN
	6	3.42	APOB protein	Q7Z7Q0_HUMAN
	7	3.37	Apolipoprotein B (Including Ag(X) antigen)	C0JYY2_HUMAN
	8	3.05	Fibrinogen beta chain	FIBB_HUMAN (+1)
	9	2.77	Fibrinogen alpha chain	FIBA_HUMAN
	10	2.68	Apolipoprotein B variant (Fragment)	Q59HB3_HUMAN
Patient 5	1	12.44	Full-length cDNA clone CS0DD006YL02 of Neuroblastoma	Q86TT1_HUMAN
	2	8.99	Ig mu chain C region(A0A087X2C0)	A0A087X2C0_HUMAN
	3	8.65	Ig mu chain C region(A0A087WYJ9)	A0A087WYJ9_HUMAN
	4	3.79	Apolipoprotein B (Including Ag(X) antigen)	C0JYY2_HUMAN
	5	3.79	APOB protein	Q7Z7Q0_HUMAN
	6	3.48	Alpha-2-macroglobulin	A2MG_HUMAN
	7	2.85	Protein IGHV3-72	A0A087WW89_HUMAN
	8	2.28	Immunglobulin heavy chain variable region (Fragment)	Q0ZCH9_HUMAN
	9	1.96	Fibrinogen alpha chain	FIBA_HUMAN
	10	1.95	C4b-binding protein alpha chain	C4BPA_HUMAN
Patient 6	1	8.00	Full-length cDNA clone CS0DD006YL02 of Neuroblastoma	Q86TT1_HUMAN
	2	5.76	Ig mu chain C region(A0A087X2C0)	A0A087X2C0_HUMAN
	3	5.46	Ig mu chain C region(A0A087WYJ9)	A0A087WYJ9_HUMAN
	4	4.92	APOB protein	Q7Z7Q0_HUMAN
	5	4.52	Apolipoprotein B (Including Ag(X) antigen)	C0JYY2_HUMAN
	6	4.45	Alpha-2-macroglobulin	A2MG_HUMAN
	7	3.93	Fibrinogen alpha chain	FIBA_HUMAN
	8	3.43	Fibrinogen beta chain	FIBB_HUMAN (+1)
	9	2.00	Actin, cytoplasmic 2	ACTG_HUMAN (+1)
	10	1.66	C4b-binding protein alpha chain	C4BPA_HUMAN

Figure 10

Rank	RPA% (Mean)	Identified Proteins	Accession Number
1	7.987	Full-length cDNA clone CS0DD006YL02 of Neuroblastoma of Homo sapiens (human)	Q86TT1_HUMAN
2	5.700	Ig mu chain C region	A0A087X2C0_HUMAN
3	4.557	Ig mu chain C region	A0A087WYJ9_HUMAN
4	3.475	Alpha-2-macroglobulin	A2MG_HUMAN
5	3.453	Fibrinogen beta chain	FIBB_HUMAN (+1)
6	2.982	Fibrinogen alpha chain	FIBA_HUMAN
7	2.638	Apolipoprotein B (including Ag(X) antigen)	C0JYY2_HUMAN
8	2.468	Serum albumin	ALBU_HUMAN
9	2.223	APOB protein	Q7Z7Q0_HUMAN
10	1.903	Protein IGHV3-72	A0A087WW89_HUMAN
11	1.737	C4b-binding protein alpha chain	C4BPA_HUMAN
12	1.595	Fibrinogen gamma chain	FIBG_HUMAN
13	1.353	Haptoglobin (Fragment)	H3BS21_HUMAN
14	1.345	Haptoglobin	H0Y300_HUMAN
15	1.140	IGH@ protein	Q6GMX6_HUMAN
16	1.122	Ig gamma-1 chain C region	A0A087WV47_HUMAN (+1)
17	1.113	Uncharacterized protein	A8K008_HUMAN
18	1.055	Uncharacterized protein	Q6N089_HUMAN
19	1.042	Actin, cytoplasmic 2	ACTG_HUMAN (+1)
20	1.042	IgG H chain	S6B291_HUMAN
21	0.957	GCT-A10 heavy chain variable region (Fragment)	A0A120HG46_HUMAN
22	0.957	Ig gamma-1 chain C region GN=IGHG1	A0A087WYC5_HUMAN
23	0.942	cDNA FLJ14473 fis, clone MAMMA1001080, highly similar to Homo sapiens SNC73 protein (SNC73) mRNA	Q96K68_HUMAN
24	0.927	Haptoglobin-related protein	HPTR_HUMAN
25	0.915	Putative uncharacterized protein DKFZp686G11190	Q6MZQ6_HUMAN
26	0.883	Myosin-reactive immunoglobulin heavy chain variable region (Fragment)	Q9UL90_HUMAN
27	0.868	Putative uncharacterized protein DKFZp686C15213	Q6MZU6_HUMAN
28	0.860	Ig gamma-3 chain C region	A0A087WXL8_HUMAN
29	0.828	Apolipoprotein E isoform 1 (Fragment)	A0A0S2Z3D5_HUMAN (+1)
30	0.785	Putative uncharacterized protein DKFZp686C11235 GN=DKFZp686C11235	Q6MZV7_HUMAN
31	0.780	Ig mu heavy chain disease protein	MUCB_HUMAN
32	0.772	IGK@ protein	Q6PIL8_HUMAN
33	0.768	Complement C3	CO3_HUMAN (+1)
34	0.690	Immunglobulin heavy chain variable region (Fragment)	Q0ZCH9_HUMAN
35	0.637	Beta-globin	D9YZU5_HUMAN (+1)
36	0.625	MS-D4 heavy chain variable region (Fragment)	A0A0X9UWK7_HUMAN
37	0.623	IgG L chain	S6BGD6_HUMAN
38	0.620	Immunoglobulin heavy chain variant (Fragment)	Q9NPP6_HUMAN
39	0.612	Anti-FactorVIII scFv (Fragment)	A2KBC6_HUMAN
40	0.607	Lipoprotein B (Fragment)	S5FLF7_HUMAN
41	0.602	Lambda-chain (AA -20 to 215)	A2NUT2_HUMAN
42	0.587	Mutant hemoglobin alpha 2 globin chain	A0A0K2BMD8_HUMAN (+2)
43	0.585	Protein IGHV3-74 (Fragment)	A0A0B4J1X5_HUMAN
44	0.575	Protein S isoform 1 (Fragment)	A0A0S2Z4K3_HUMAN (+2)
45	0.553	Myosin-reactive immunoglobulin heavy chain variable region (Fragment)	Q9UL88_HUMAN
46	0.505	Apolipoprotein C-III	A3KPE2_HUMAN (+2)
47	0.502	Lectin galactoside-binding soluble 3 binding protein isoform 1 (Fragment)	A0A0S2Z3Y1_HUMAN (+1)
48	0.502	cDNA FLJ53691, highly similar to Serotransferrin	B4E1B2_HUMAN
49	0.495	cDNA, FLJ94213, highly similar to Homo sapiens pregnancy-zone protein (PZP), mRNA	B2R950_HUMAN (+1)
50	0.480	Thrombospondin-1	TSP1_HUMAN
51	0.470	Protein IGHV5-51 (Fragment)	A0A0C4DH38_HUMAN
52	0.458	Vinculin, isoform CRA_c	A0A024QZN4_HUMAN (+2)
53	0.455	Fibronectin 1, isoform CRA_n	A0A024R462_HUMAN
54	0.447	Apolipoprotein B variant (Fragment)	Q59HB3_HUMAN
55	0.437	Keratin 1	H6VRF8_HUMAN (+3)
56	0.420	Filamin-A	FLNA_HUMAN
57	0.413	Putative uncharacterized protein DKFZp686L19235 GN=DKFZp686L19235	Q6MZV6_HUMAN
58	0.412	Ig heavy chain V-III region BUT	HV306_HUMAN
59	0.410	IBM-B2 heavy chain variable region (Fragment)	A0A125QYY9_HUMAN
60	0.410	Alpha-1-antitrypsin	A1AT_HUMAN (+1)
61	0.408	Talin-1	TLN1_HUMAN
62	0.403	Actinin, alpha 1, isoform CRA_a	A0A024R694_HUMAN (+1)
63	0.397	Apolipoprotein A-I, isoform CRA_a	A0A024R3E3_HUMAN (+1)
64	0.392	Transferrin variant (Fragment)	Q53H26_HUMAN
65	0.388	von Willebrand factor	VWF_HUMAN
66	0.387	Ig gamma-4 chain C region (Fragment)	A0A0G2JPD4_HUMAN (+1)
67	0.383	von Willebrand factor	L8E853_HUMAN
68	0.377	Immunoblobulin light chain (Fragment)	Q0KKI6_HUMAN (+1)
69	0.370	Rheumatoid factor RF-ET9 (Fragment)	A2J1N6_HUMAN
70	0.362	Coagulation factor XIII A chain	F13A_HUMAN
71	0.353	Apolipoprotein C-I, isoform CRA_a	A0A024R0T8_HUMAN (+2)
72	0.345	Rheumatoid factor RF-ET6 (Fragment)	A2J1N5_HUMAN

73	0.322	Complement factor H	CFAH_HUMAN
74	0.303	Keratin, type I cytoskeletal 10 GN=KRT10	K1C10_HUMAN
75	0.302	Protein APOC4-APOC2	K7ER74_HUMAN
76	0.290	MS-D3 heavy chain variable region (Fragment)	A0A0X9T7T4_HUMAN
77	0.288	Ig heavy chain V-III region WEA	HV302_HUMAN
78	0.287	Uncharacterized protein	A0A5E4_HUMAN
79	0.275	Keratin, type I cytoskeletal 9	K1C9_HUMAN
80	0.270	cDNA FLJ75066, highly similar to Homo sapiens complement component 1, r subcomponent (C1R), mRNA	A8K5J8_HUMAN (+1)
81	0.265	Epididymis tissue protein Li 173	E9KL26_HUMAN (+1)
82	0.263	Myosin, heavy polypeptide 9, non-muscle, isoform CRA_a	A0A024R1N1_HUMAN (+1)
83	0.260	Cryocrystalglobulin CC1 heavy chain variable region (Fragment)	B1N7B6_HUMAN
84	0.245	Putative uncharacterized protein DKFZp686I04196 (Fragment)	Q6N093_HUMAN
85	0.235	Heat shock cognate 71 kDa protein	HSP7C_HUMAN (+1)
86	0.233	Apolipoprotein C-IV	A5YAK2_HUMAN
87	0.225	Hemopexin	HEMO_HUMAN
88	0.222	Epididymis secretory protein Li 52	V9HWG7_HUMAN (+1)
89	0.220	Actin, alpha skeletal muscle	ACTS_HUMAN (+1)
90	0.215	Keratin, type II cytoskeletal 2 epidermal	K22E_HUMAN
91	0.215	Putative uncharacterized protein DKFZp686M08189	Q6MZX9_HUMAN
92	0.210	Fermitin family homolog 3	URP2_HUMAN
93	0.208	Complement C4-B	CO4B_HUMAN
94	0.207	Pyruvate kinase PKM	KPYM_HUMAN (+1)
95	0.205	GCT-A1 heavy chain variable region (Fragment)	A0A125U0V2_HUMAN
96	0.202	Gelsolin	GELS_HUMAN
97	0.192	cDNA FLJ32131 fis, clone PEBLM2000267, highly similar to Tubulin alpha-ubiquitous chain	B3KPS3_HUMAN (+2)
98	0.188	CD5 antigen-like	CD5L_HUMAN
99	0.185	cDNA FLJ56821, highly similar to Inter-alpha-trypsin inhibitor heavy chain H1	B7Z549_HUMAN (+1)
100	0.183	Immunglobulin heavy chain variable region (Fragment)	Q0ZCH6_HUMAN
101	0.183	Ig heavy chain V-III region GAL	HV320_HUMAN
102	0.178	Actinin alpha 4 isoform 1 (Fragment)	A0A0S2Z3G9_HUMAN (+1)
103	0.178	Complement component 1, q subcomponent, C chain, isoform CRA_a	A0A024RAA7_HUMAN (+1)
104	0.177	Tubulin alpha-4A chain	TBA4A_HUMAN
105	0.170	cDNA FLJ53487, highly similar to Coagulation factor XIII A chain (EC 2.3.2.13)	B4E2L8_HUMAN
106	0.167	cDNA FLJ35730 fis, clone TESTI2003131, highly similar to ALPHA-1-ANTICHYMOTRYPSIN	B3KS79_HUMAN
107	0.165	Profilin-1	PROF1_HUMAN
108	0.165	Immunoglobulin J chain	IGJ_HUMAN
109	0.165	Immunglobulin heavy chain variable region (Fragment)	Q0ZCF6_HUMAN
110	0.165	cDNA FLJ53025, highly similar to Complement C4-B	B7Z1F8_HUMAN
111	0.157	Moesin	MOES_HUMAN (+1)
112	0.157	78 kDa glucose-regulated protein	GRP78_HUMAN (+1)
113	0.157	Inter-alpha (Globulin) inhibitor H2	A2RTY6_HUMAN (+3)
114	0.157	IGL@ protein	Q6PIK1_HUMAN
115	0.152	Apolipoprotein M	APOM_HUMAN
116	0.152	cDNA FLJ57038, highly similar to Filamin-A	B4E2F9_HUMAN
117	0.150	Complement component 1, q subcomponent, B chain, isoform CRA_a	A0A024RAB9_HUMAN (+3)
118	0.147	Beta 5-tubulin	Q5SU16_HUMAN (+1)
119	0.145	IgG H chain	S6B2A6_HUMAN
120	0.143	Rheumatoid factor RF-IP12 (Fragment)	A2J1M8_HUMAN
121	0.140	Protein IGHV3-73	A0A0G2JN55_HUMAN
122	0.135	Lipoprotein, Lp(A)	Q1HP67_HUMAN
123	0.132	Keratin, type I cytoskeletal 16	K1C16_HUMAN
124	0.130	Heat shock protein HSP 90-alpha	HS90A_HUMAN (+1)
125	0.125	Platelet factor 4	PLF4_HUMAN
126	0.125	Putative uncharacterized protein DKFZp686C02220 (Fragment)	Q6N091_HUMAN
127	0.123	Apolipoprotein D (Fragment)	C9JF17_HUMAN
128	0.120	Keratin, type I cytoskeletal 14	K1C14_HUMAN
129	0.118	Tubulin beta-1 chain	TBB1_HUMAN
130	0.118	Inter-alpha (Globulin) inhibitor H4 (Plasma Kallikrein-sensitive glycoprotein)	B2RMS9_HUMAN (+1)
131	0.118	Protein disulfide-isomerase	B3KQT9_HUMAN (+2)
132	0.117	Protein SAA2-SAA4	A0A096LPE2_HUMAN
133	0.112	CP protein	A5PL27_HUMAN (+3)
134	0.110	Alpha-1B-glycoprotein	A1BG_HUMAN (+1)
135	0.110	Integrin beta-3	ITB3_HUMAN (+1)
136	0.110	Heparin cofactor 2	HEP2_HUMAN
137	0.108	Band 3 anion transport protein	B3AT_HUMAN (+3)
138	0.107	Ig heavy chain V-III region CAM	HV307_HUMAN
139	0.105	Apolipoprotein L, 1, isoform CRA_b	A0A024R1G8_HUMAN (+11)
140	0.102	Glycoprotein Ib (Platelet), alpha polypeptide	A0A0C4DGZ8_HUMAN (+2)
141	0.102	Complement C5	CO5_HUMAN
142	0.102	Endoplasmin	ENPL_HUMAN (+2)
143	0.100	Rheumatoid factor C6 light chain (Fragment)	A0N5G1_HUMAN
144	0.098	Clusterin	CLUS_HUMAN
145	0.097	Adenylyl cyclase-associated protein	D3DPU2_HUMAN
146	0.095	Complement C4-A	A0A0G2JPR0_HUMAN
147	0.095	Complement C1s subcomponent	C1S_HUMAN
148	0.093	cDNA FLJ54184, highly similar to Tropomyosin alpha-4 chain	B4DVY2_HUMAN (+2)

149	0.093	Glyceraldehyde-3-phosphate dehydrogenase	G3P_HUMAN (+1)
150	0.092	Heat shock protein 90kDa alpha (Cytosolic), class B member 1, isoform CRA_a	A0A024RD80_HUMAN (+2)
151	0.090	C4b-binding protein beta chain	C4BPB_HUMAN
152	0.090	Anti-Influenza A hemagglutinin heavy chain variable region (Fragment)	G1FM90_HUMAN
153	0.088	Integrin alpha-IIb	ITA2B_HUMAN
154	0.088	Keratin, type II cytoskeletal 5	K2C5_HUMAN
155	0.087	Ig heavy chain V-III region GA	HV308_HUMAN
156	0.087	Vitronectin	D9ZGG2_HUMAN (+1)
157	0.087	Plasminogen	PLMN_HUMAN
158	0.087	GCT-A2 heavy chain variable region (Fragment)	A0A125U0V4_HUMAN
159	0.085	Ficolin (Collagen/fibrinogen domain containing) 3 (Hakata antigen), isoform CRA_b	Q6UXM4_HUMAN (+1)
160	0.083	IgGFc-binding protein	A0A087WXI2_HUMAN (+1)
161	0.078	Vitamin D-binding protein	D6RF35_HUMAN
162	0.077	Calreticulin	CALR_HUMAN (+2)
163	0.077	cDNA FLJ53075, highly similar to Kininogen-1	B4DPP8_HUMAN (+1)
164	0.075	14-3-3 protein zeta/delta	1433Z_HUMAN (+1)
165	0.075	Transitional endoplasmic reticulum ATPase	TERA_HUMAN (+1)
166	0.073	Uncharacterized protein (Fragment)	A0A0J9YX35_HUMAN
167	0.072	Platelet glycoprotein V	GPV_HUMAN
168	0.072	cDNA FLJ41552 fis, clone COLON2004478, highly similar to Protein Tro alpha1 H,myeloma	Q6ZW64_HUMAN
169	0.070	Platelet basic protein	CXCL7_HUMAN (+1)
170	0.068	Protein IGHV3-13 (Fragment)	A0A0A0MS11_HUMAN
171	0.067	Complement component 9, isoform CRA_a	A0A024R035_HUMAN (+1)
172	0.065	Proteoglycan 4, isoform CRA_a	A0A024R930_HUMAN (+2)
173	0.065	Protein disulfide-isomerase	A0A024R8S5_HUMAN (+1)
174	0.065	Catalase	CATA_HUMAN
175	0.065	Epididymis luminal protein 180 (Fragment)	B6EDE2_HUMAN
176	0.065	Alpha-1-acid glycoprotein 1	A1AG1_HUMAN (+1)
177	0.063	Heat shock 70 kDa protein 1B	A0A0G2JIW1_HUMAN (+3)
178	0.063	Prenylcysteine oxidase 1	PCYOX_HUMAN
179	0.063	Integrin-linked protein kinase	A0A0A0MTH3_HUMAN (+2)
180	0.063	Carboxypeptidase N subunit 2	CPN2_HUMAN
181	0.063	Anti-streptococcal/anti-myosin immunoglobulin kappa light chain variable region (Fragment)	Q96SA9_HUMAN
182	0.063	GCT-A5 light chain variable region (Fragment)	A0A0X9UWL5_HUMAN
183	0.062	Zyxin GN=ZYX	ZYX_HUMAN
184	0.062	V2-17 protein (Fragment)	Q5NV90_HUMAN
185	0.062	Myosin-reactive immunoglobulin heavy chain variable region (Fragment)	Q9UL89_HUMAN
186	0.060	cDNA FLJ54622, highly similar to Prothrombin (EC 3.4.21.5)	B4DDT3_HUMAN (+1)
187	0.060	Multimerin-1	MMRN1_HUMAN
188	0.058	Tropomyosin 3 isoform 1 (Fragment)	A0A0S2Z4G4_HUMAN (+1)
189	0.058	Ig gamma-1 chain C region	A0A087X1C7_HUMAN
190	0.057	Alpha-2-antiplasmin	A2AP_HUMAN
191	0.057	Ig heavy chain variable region (Fragment)	A0A068LRW6_HUMAN (+2)
192	0.057	Enolase 1, (Alpha), isoform CRA_a	A0A024R4F1_HUMAN (+1)
193	0.057	Apolipoprotein A-IV	APOA4_HUMAN
194	0.057	MS-D1 light chain variable region (Fragment)	A0A0X9TD47_HUMAN
195	0.055	Uncharacterized protein	B4E1Z4_HUMAN
196	0.055	Calpain 1, (Mu/I) large subunit, isoform CRA_a	A0A024R580_HUMAN (+2)
197	0.055	GCT-A8 heavy chain variable region (Fragment)	A0A0X9V9C4_HUMAN
198	0.055	Serpin peptidase inhibitor, clade C (Antithrombin), member 1, isoform CRA_a	A0A024R944_HUMAN (+1)
199	0.052	Protein AMBP	AMBP_HUMAN
200	0.050	Fructose-bisphosphate aldolase A	ALDOA_HUMAN (+2)

Figure 11

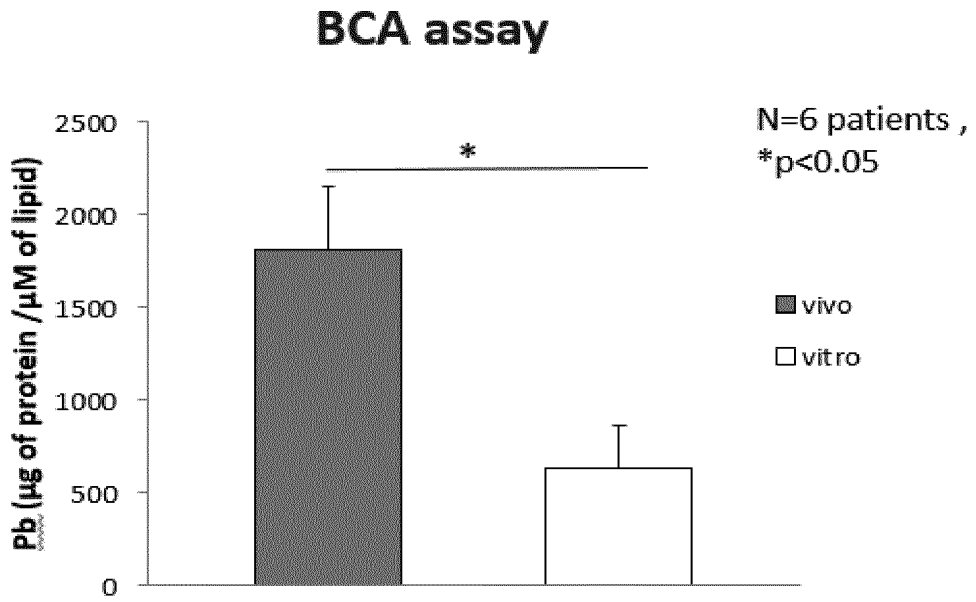
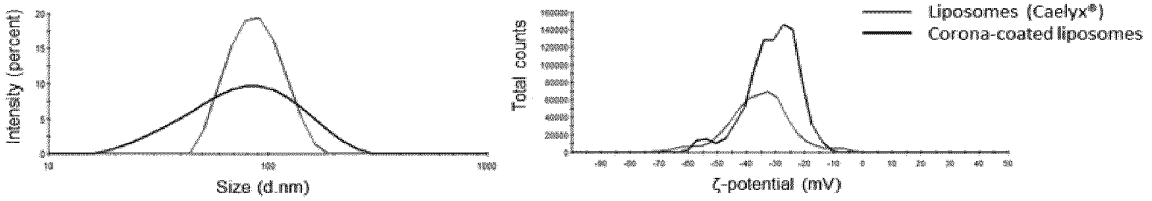
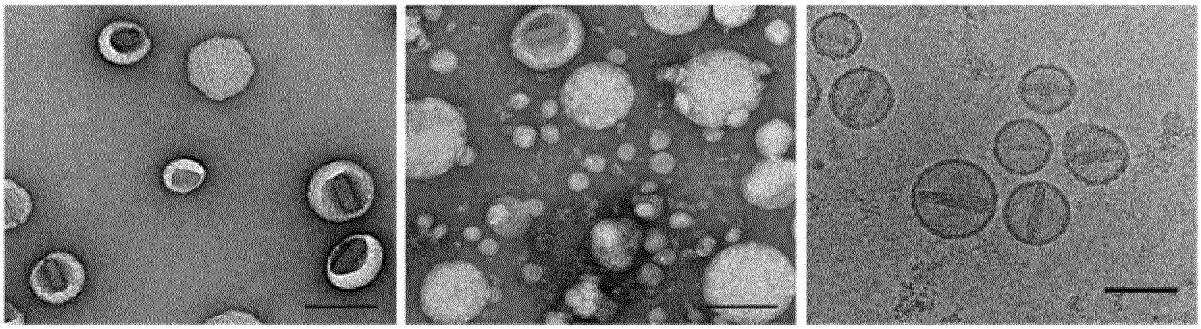

BCA assay
N=6 patients ,
*p<0.05
*
vivo
vitro
Pb (µg of protein /µM of lipid)
2500
2000
1500
1000
500
0

Figure 12



	Mean hydrodynamic diameter (nm)	ζ-potential (mV)	PDI
Liposomes-Caelyx® [HSPC/CHOL/mPEG2000-DSPE(56.2:38.5:5.3)]	82.63 ± 1.391	-33.8 ± 1.74	0.087 ± 0.023
Corona-coated liposomes (Patient1)	72.81 ±1.120	-30.7 ± 1.87	0.275 ± 0.021
Corona-coated liposomes (Patient2)	89.56 ± 1.662	-36.2 ± 0.85	0.243 ± 0.011
Corona-coated liposomes (Patient3)	80.80 ± 0.782	-33.0 ± 2.04	0.234 ± 0.006
Corona-coated liposomes (Patient4)	86.95 ± 3.469	-37.5 ± 0.76	0.228 ± 0.008
Corona-coated liposomes (Patient5)	79.78 ± 0.550	-30.8 ± 0.06	0.274 ± 0.006
Corona-coated liposomes (Patient6)	98.28 ± 2.012	-39.8 ± 2.80	0.229 ± 0.002



Liposomes (Caelyx®) by TEM | Liposome: ***in vivo*** corona by TEM | Liposome: ***in vivo*** corona by cryo-EM

Figure 13

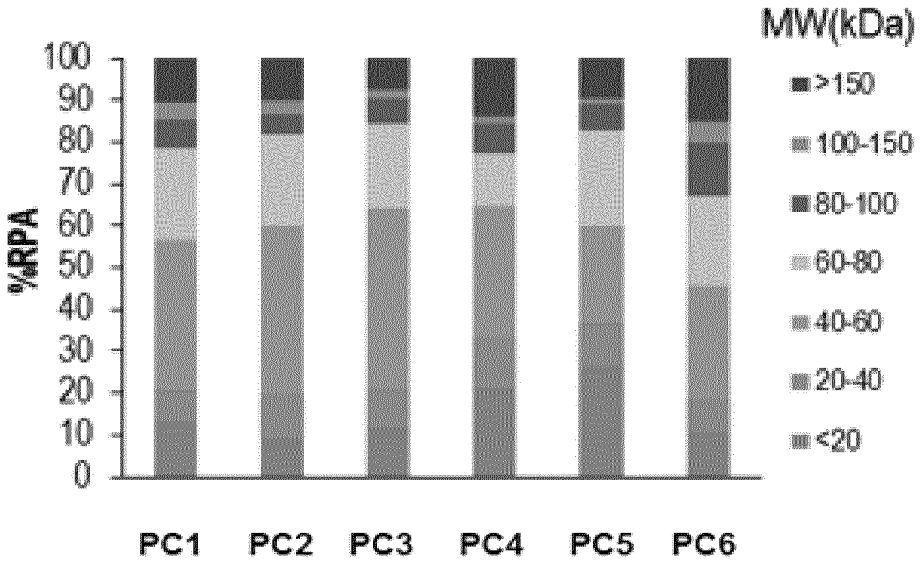

MW(kDa)
>150
100-150
80-100
60-80
40-60
20-40
<20
%RPA
100
90
80
70
60
50
40
30
20
10
0
PC1
PC2
PC3
PC4
PC5
PC6

Figure 14

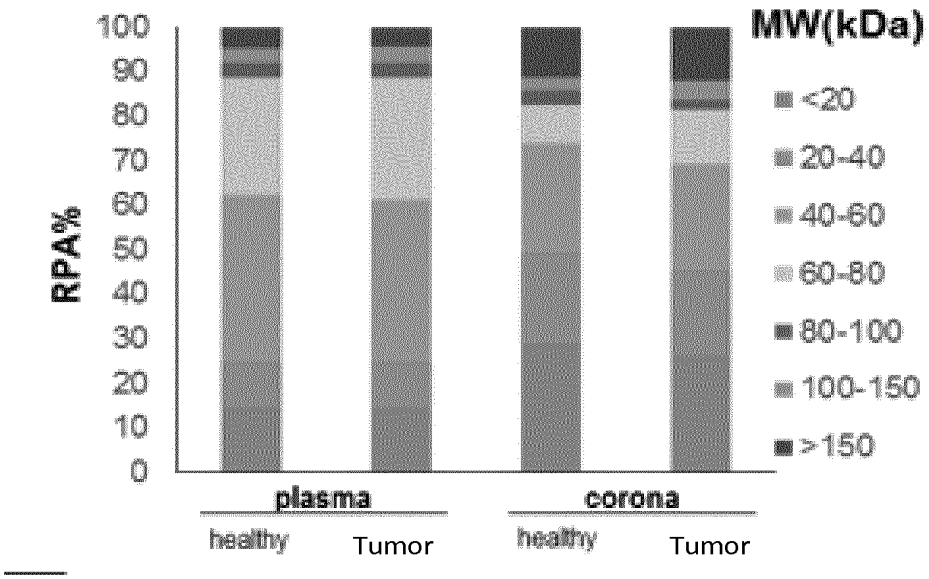
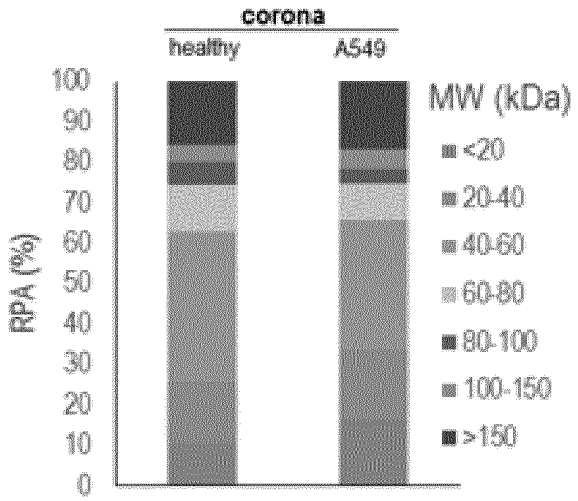

RPA%
100
90
80
70
60
50
40
30
20
10
0
MW(kDa)
<20
20-40
40-60
60-80
80-100
100-150
>150
plasma
corona
healthy
Tumor
healthy
Tumor

Figure 15


corona
healthy
A549
RPA (%)
100
90
80
70
60
50
40
30
20
10
0
MW (kDa)
<20
20-40
40-60
60-80
80-100
100-150
>150