Developing High-Affinity Protein Capture Agents and Nanotechnology-Based Platforms for *In Vitro* Diagnostics

Thesis by

Rosemary Dyane Rohde

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

2009

(Defended May 27, 2009)

This thesis is dedicated to my Mom for her unconditional and everlasting love and support.

Acknowledgements

The work described in this thesis would not have been possible without guidance from many people. First, I would like to thank my advisor, Jim Heath. Jim is an inspiration— his intensity, creativity, and dedication to science is admirable. I never left his office without being motivated and excited to further my research. I am thankful for the space, funding, and intellectual support he has provided me. I am excited to continue to work with him at our new start-up called Integrated Diagnostics. Thank you to the rest of my faculty committee: Geoffrey Blake, Jack Beauchamp, and Christina Smolke. It has been an honor to work with such a distinguished group of scientists. I would especially like to thank my lab mate, Heather Agnew, who is a great scientist— hard working, passionate, and dedicated. I am truly grateful to have worked with her over the years. She is also an inspiration and has helped me grow into the scientist I am today. Special thanks to Arundhati Nag, Dr. Steven Millward, Dr. Woon-Seok Yeo, and Dr. Ryan Bailey for useful discussions, assistance with experiments, and support. I also want to thank all the bright undergraduates: Abdul Ahad Tariq, Vanessa M. Burns, Russell J. Krom, Marissa Barrientos, Eric M. Johnson, and Andreea D. Stuparu, who worked with me throughout the years. Thanks to everyone in the Heath group for caring about my graduate career and me. Your ideas, support, and suggestions are greatly appreciated. Thanks to Kevin Kan for fixing my computer several times. Thanks to Diane Robinson and Gabe Kwong for all their advice and support. I would like to thank the people at The Scripps Research Institute: Dr. Jason E. Hein, Dr. Suresh M. Pitram, Dr. Valery Fokin, Dr. M. G. Finn, and Dr. Barry K. Sharpless for our scientific conversations and collaborations. Special thanks to Dr. Michael McAlpine, Dr. Mario Blanco, Dr. Bert

Lai, and Dr. Bruce Brunschwig for their teachings and help with experiments.

I greatly appreciate the Bill and Melinda Gates Foundation for supporting my education, including my undergraduate studies at UCLA. I thank the Genomics Facility for use of the Genepix 4200 array scanner. ESI- and MALDI-MS analysis was completed by the Protein/Peptide MicroAnalytical Laboratory. Fluorescence polarization experiments were done in the Beckman Institute Laser Resource Center. SPR experiments were performed in the Protein Expression Center, under the direction of Jost Vielmetter. I thank Carl S. Parker for assistance with dot blot experiments.

Importantly, I thank my loved ones who are always a constant source of support. My wonderful mom, Lily Coloma, taught me the value of unconditional love. The sacrifices she made for me will never be forgotten. Thanks to George Porcari who has always been there for me, whether I needed help moving, a visit, or a phone call. I thank my aunts, Rosana Coloma-Diaz, Jenny Member, and Raquel Coloma, who are strong, hardworking beautiful women. I thank my uncles, Joe Diaz and Gregory Rohde, for their support. Thanks to all my family in Peru, especially my Abuelita Lily Coloma, for their love. I thank all my friends for their motivation and encouragement. I especially thank Moises Robles, Christine Baron, and Ronnie Bryan, who have helped me developed into the person I am today. Thanks to everyone who proofread my thesis. Thanks to all the GSC, Net Impact, and SASS students, staff, and administration that have worked with me over the years to help improve the quality of life at Caltech. I will miss Caltech and its community.

Abstract

Blood protein diagnostics has the potential to revolutionize health care by providing the relevant clinical measurements that can provide the foundation for predictive, preventive, and personalized medicine.[1] Blood bathes all organs in the body, with a circulation time of just a few minutes.[2] Those organs secrete proteins into the blood, some of which contain information relevant to the health or disease status of the organ.[3–5] Thus, the blood potentially provides a window into the health state of each individual, and the organ-specific secreted proteins can provide a molecular fingerprint of disease. The hypothesis is that each organ (and its associated diseases) has a unique fingerprint that can be read from the blood with an appropriately multiplexed diagnostic platform. These fingerprints potentially provide insight that can be harnessed for early disease diagnostics, since the protein levels associated with the organ-specific fingerprint will be altered by the onset of disease and by the stages of disease progression. A challenge is that capturing each organ-specific blood fingerprint will require the assessment of the levels of many blood protein biomarkers. Capturing the fingerprint from all of the 50 or so major organs (or organ regions) may require the measurement of from several hundred to a thousand or more proteins. Meeting such a challenge requires new technologies at all levels— from devices designed to process and deliver the blood proteins for measurement,[2,6,7] to sensitive measurement approaches,[8–12] to affinity agents that can be utilized to capture the relevant biomarker proteins,[13–17] to computational approaches that can process large numbers of measurements into a result that can be interpreted by a physician.[3,4,18–20]

To realize this vision of personalized healthcare, disease diagnostics measurements must be extremely inexpensive. Currently the assessment of blood-based protein biomarkers can take from between several hours to several days, depending upon the resources of a specific healthcare clinic. Consider, for example, a prostate-specific antigen (PSA) ELISA (enzyme-linked immunosorbent assay) test that helps detect prostate cancer in men.[21] In this test, a patient's serum is diluted and applied to a plate to which PSA antigens have been attached. If antibodies to PSA are present in the serum, they may bind to these PSA antigens. The plate is then washed to remove all other components of the serum. A specially prepared "secondary antibody" — an antibody that specifically binds to human antibodies — is then applied to the plate, followed by another wash. Prior to this step, the secondary antibody is chemically linked to an enzyme. Thus, after application of the secondary antibody, the plate will contain immobilized enzyme in proportion to the amount of secondary antibody bound to the plate. A substrate for the enzyme is applied, and catalysis by the enzyme leads to a change in color or fluorescence. ELISA results are reported as a number; the most subjective aspect of this test is determining the "cut-off" point between a positive and negative result. This ELISA assay is representative of a typical test for a protein biomarker but it does have its disadvantages. First, it requires large amounts (~ 5–10 ml) of blood to be collected from the patient. Second, the long time period separating the point of blood collection from the time of measurement has associated costs in terms of both measurement accuracy (degradation of blood-based proteins) and labor costs. Third, the antibodies that serve as protein capture agents are expensive to develop, purchase, and store, and can be degraded by any number of subtle physical, chemical, and biological influences. Fourth, most

diagnostics measurements are pauciparameter. Fifth, the sensitivity, specificity, and/or dynamic range of the biomarker detection strategies are often limited.

In this thesis, I describe projects that were aimed at addressing specific aspects of some of these problems. Nanoelectronic sensors, such as silicon nanowires (SiNWs), [12,22] can provide quantitative measurements of protein biomarkers in real time. Another advantage of SiNWs is that they are label-free sensors, so no secondary antibody is needed to detect the binding event. The goal is to fabricate large arrays of SiNW circuits, each of which can be individually functionalized by a different capture agent. When the blood protein binds to the specific capture agent, both the electrical conductance of the nanowire and the electrical capacitance between the nanowire and the serum is changed. These changes correlate to the amount of the protein in the blood, and thus permit a label-free, real-time measurement. One technical challenge for nanoelectronic protein sensors is to develop chemistry that can be applied for selectively encoding the nanowire surfaces with capture agents, thus making them sensors that have selectivity for specific proteins or other biomolecules. Furthermore, because of the nature of how the sensor works, it is desirable to achieve this spatially selective chemical functionalization without having the silicon undergo oxidation. The native oxide on silicon (SiO_2) has a low isoelectric point, meaning that under physiological conditions (= pH 7.4), surfaces are negatively charged.[23] These surface charges can potentially limit the sensitivity of certain nanoelectronic biomolecular sensor devices through Debye screening of the biomolecular probe/target binding event to be sensed. Furthermore, the native oxide of Si can detrimentally impact carrier recombination rates.[24] For high-surface-area devices, such as SiNWs, this can likely result in a degradation of electrical properties.

A general method for the non-oxidative functionalization of single-crystal silicon (111) is described in Chapter 1. The general approach is to start with a silicon-on-insulator (SOI) wafer. A SOI wafer is comprised of a thin, single-crystal silicon film (~ 30–50 nm) on top of a thick, insulating SiO_2 layer (on the order of microns). It is from this SOI layer that the nanowire sensors are fabricated. The silicon film, unless specially treated, is naturally passivated with a thin (1–2 nm thick) native oxide (SiO_2) layer as described above. To remove the native oxide, the silicon surface is fully acetylenylated (-C≡CH), thus preventing the growth of oxidation. Additionally, the -C≡CH group also provides a chemical handle for additional functionalization via the 'click' reaction[25,26] between an azide containing benzoquinone (masked with a primary amine) and the surface-bound alkyne. The benzoquinone is electrochemically reduced, exposing the amine terminus. During this process, minimal oxidation is present. Molecules presenting a carboxylic acid have been immobilized to the exposed amine sites. This strategy provides a general platform that can incorporate organic and biological molecules on Si (111) with minimal oxidation of the silicon surface. This method can further be extended towards the selective biopassivation of capture agent arrays of nanoelectronic sensor devices.

The development of these devices is, in part, driven by early diagnosis, differential treatment, monitoring, and personalized medicine— all of which are increasingly requiring quantitative, rapid, and multiparameter measurement capabilities on ever smaller amounts of tissues, cells, serum, etc.[1–5] To begin achieving this goal, a large number of protein biomarkers need to be captured and quantitatively measured to create a diagnostic panel. One of the greatest challenges towards making protein-

biomarker-based *in vitro* diagnostics inexpensive involves developing capture agents to detect the proteins. In the content of this thesis, a capture agent is a biochemical molecule that has specificity for a distinct target molecule and can be naturally derived and/or synthetic. A major thrust of this thesis is to develop multi-valent, high-affinity and high-selectivity protein capture agents using *in situ* click chemistry.[27–30] *In situ* click chemistry is a tool that utilizes the protein itself to catalyze the formation of a biligand from individual azide and alkyne ligands that are co-localized. Large libraries of peptides are used to form the body of these ligands, also providing high chemical diversity and protease stability, with minimal synthetic effort. Peptide-based moderate-affinity lead compounds can be isolated from a 2-generation screen of the protein against a one-bead one-compound (OBOC) library. By modifying that peptide with appropriate alkyne or azide functionalities, that peptide becomes an anchor (1°) ligand, and part of the capture agent. Simultaneously screening the protein against this single anchor ligand and a large library of click-complementary secondary (2°) ligands constitutes our strategy for identifying biligand capture agents. The protein target holds the two peptide-units in close proximity promoting the covalent coupling between the azide and alkyne moieties. This process can be repeated— the biligand capture agent can serve as the new anchor unit and the same OBOC library can be employed to identify a triligand, tetraligand, and so forth. The addition of each ligand to the capture agent causes the affinity and selectivity to increase dramatically for its cognate protein. In particular, we describe the production of a triligand capture agent that exhibits 45 nM and 64 nM affinities against human and bovine carbonic anhydrase II (bCAII and hCAII) protein, respectively, and can be used in a dot blot test to detect those proteins at the ≥ 20 ng level from 10%

serum. *In situ* click chemistry screens are shown to yield results identical to more traditional OBOC screens, but the *in situ* screens permit orders of magnitude more chemical space to be sampled. Moreover, the resulting multiligand protein capture agents can be produced in gram-scale quantities using conventional synthetic methods with designed control over chemical and biochemical stability and water solubility.

Chapter 2 will cover the synthesis of the azide containing artificial amino acids at the multi-gram quantity scale. The construction of large (up to 20 million elements on 100 million beads) peptide libraries for screening, including bulk peptide synthesis and on-bead click reaction will be discussed. Bead-based library screening procedures will be reviewed. Analysis of lead compounds by Edman degradation will be introduced, including calibrating peptide-sequencing equipment so artificial amino acids can be identified.

Chapter 3 will cover screening procedures and results for the anchor ligand, biligand, and triligand in detail. Binding affinity measurements using fluorescence polarization and surface plasmon resonance (SPR) are reported. The sensitivity and selectivity of the multi-ligand (biligand and triligand) capture agents for CAII proteins in complex environments are demonstrated through the use of dot blot experiments in 10% serum. The advances of this approach are multifold and will be discussed in Chapter 3. This is a general and robust method for inexpensive, high-throughput capture agent discovery that can be utilized to capture the relevant biomarker proteins for blood protein diagnostics.

References:

1. Hood, L.; Heath, J. R.; Phelps, M. E.; Lin, B. Y. *Science* **2004,** *306*, 640–643.
2. Fan, R.; Vermesh, O.; Srivastava, A.; Yen, B.; Qin, L.; Ahmad, H.; Kwong, G.; Liu, C.; Gould, J.; Hood, L.; Heath, J. R. *Nature Biotechnology* **2008**, *26*, 1373–1378.
3. Davidson, E. H., et al. *Science* **2002,** *295*, (5560), 1669–1678.
4. Kitano, H. *Science* **2002,** *295*, (5560), 1662–1664.
5. Ludwig, J. A.; Weinstein, J. N. *Nature* **2005**, *5*, 845–856.
6. Toner, M.; Irimia, D. *Annu. Rev. Biomed. Eng*. **2005**, *7*, 77–103.
7. Yang, S.; Undar, A.; Zahn, J. D. *Lab Chip* **2006**, *6*, 871–880.
8. Fritz, J., et al. *Science* **2000**, *288*, 316.
9. Cui, Y.; Wei, Q.; Park, H.; Lieber, C. M. *Science* **2001**, *293*, 1289.
10. Chen, R. J., et al. *Proc. Natl. Acad. Sci. U.S.A*. **2003**, *100*, 4984.
11. Patolsky, F.; Lieber, C. M. *Mater. Today* **2005**, *8*, 20.
12. Bunimovich, Y. L.; Shin, Y.-S.; Yeo, W.-S.; Amori, M.; Kwong, G.; Heath, J. R. *J. Am. Chem. Soc.* **2006**, *128*, 16323–16331.
13. Lee, J. F.; Hesselberth, J. R.; Meyers, L. A.; Ellington, A. D. *Nucleic Acids Res.* **2004**, *32*, D95–D100.
14. Fabian, M. A., et al. *Nat. Biotechnol.* **2005**, *23*, 329–336.
15. Smith, G. P.; Petrenko, V. A. *Chem. Rev.* **1997**, *97*, 391–410.
16. Lam, K. S.; Lebl, M.; Krchňák, V. *Chem. Rev.* **1997**, *97*, 411–448.
17. Agnew, H. D.; Rohde, R. D.; Millward, S. W.; Nag, A.; Yeo, W.-S.; Hein, J. E.; Pitram, S. M.; Tariq, A.-A.; Burns, V. M.; Krom, R. J.; Fokin, V. V.; Sharpless, B. K.; Heath, J. R. *Angew. Chem. Int. Ed.* **2009**, *accepted*.
18. Cui, J., Liu, Q., Puett, D., Xu, Y. *Bioinformatics* **2008**, *24*, 2370–2375.
19. Davidson, E. H., et al. *Science* **2002,** *295*, 1669–1678.
20. Kitano, H. *Science* **2002,** *295*, 1662–1664.
21. Kuriyama M.; Wang, M.C.; Papsidero, L.D.; et al. *Cancer Research* **1980**, *40*, 4658–4662.

22. Zheng, G.; Patolsky, F.; Cui, Y.; Wang, W. U.; Lieber, C. M. *Nature Biotechnol.* **2005**, *23*, 1294, and references therein.

23. Hu, K.; Fan, F.-R. F.; Bard, A. J.; Hillier, A. C. *J. Phys. Chem. B* **1997**, *101*, 8298.

24. Yablonovitch, E.; Allara, D. L.; Chang, C. C.; Gmitter, T.; Bright, T. B. *Phys. Rev. Lett.* **1986**, *57*, 249.

25. Kolb, H. C.; Finn, M. G., Sharpless, K. B. *Angew. Chem., Int. Ed.* **2001**, *40*, 2004.

26. Bock, V. D.; Hiemstra, H.; van Maarseveen, J. H. *Eur. J. Org. Chem.* **2006**, *51*, and references therein.

27. Lewis, W. G.; Green, L. G.; Grynszpan, F.; Radić, Z.; Carlier, P. R.; Taylor, P.; Finn, M. G.; Sharpless, K. B. *Angew. Chem. Int. Ed.* **2002**, *41*, 1053–1057.

28. Manetsch, R.; Krasiński, A.; Radić, Z.; Raushel, J.; Taylor, P.; Sharpless, K. B.; Kolb, H. C. *J. Am. Chem. Soc.* **2004**, *126*, 12809–12818.

29. Bourne, Y.; Kolb, H. C.; Radić, Z.; Sharpless, K. B.; Taylor, P.; Marchot, P. *Proc. Natl. Acad. Sci. USA* **2004***, 101*, 1449–1454.

30. Mocharla, V. P.; Colasson, B.; Lee, L. V.; Röper, S.; Sharpless, K. B.; Wong, C.-H.; Kolb, H. C. *Angew. Chem. Int. Ed.* **2005**, *44*, 116, and references therein.

Table of Contents

List of Figures

List of Schemes

List of Tables