Chapter 1: Introduction

1.1 Introduction

As scientists develop increasing insight in the sciences of genomics and proteomics, a molecular picture of disease development and progression is gradually evolving[1]. It is essential to translate this vast resource of scientific knowledge into clinical platforms, so that one can envisage a future, each individual has his/her genetic and proteomic 'fingerprints' or biomarker patterns recorded. Monitoring this fingerprint will help not only in treatment, but also in disease prevention. While this picture of personalized medicine is a lofty goal, it is not an unattainable one if broken down into smaller, more attainable, goals. One of the challenges that must be addressed in realizing the goal is the development of a rapid, inexpensive and minimally invasive way of measuring biomarkers[2]. In my thesis I focus on a small but vital part of this challenge, developing a rapid, inexpensive way of measuring protein biomarkers.

A biomarker is defined as an indicator of a specific biological state, which can be used to evaluate the risk of contraction or monitor the presence of a disease. Biomarkers can also provide guidance to targeted therapy and assess the response of therapy. While all the three different types of biomolecules, DNA, RNA and proteins, can act as biomarkers, protein biomarkers deserve a special mention[3]. Changes in the genetic level, such as DNA mutation, are manifested in changes in the protein levels. Additionally, proteins can undergo post-translational modifications, like phosphorylation, ubiquitination or glycolysation, that play key roles in modulating various biological processes like cell cycle, signal transduction[4], which are almost always perturbed in disease development and progression. Current traditional diagnostics platforms use a limited number of biomarkers, typically 1-2, for prognosis e.g. measuring LDL for cholesterol. These assays typically use monoclonal antibodies in a Enzyme-Linked ImmunoSorbent Assay (ELISA) format. While monoclonal antibodies have levels of affinity and selectivity, they are expensive, especially if instead of one or two biomarkers one aims to measure a large panel of protein

biomarkers. The lack of stability of monoclonal antibodies with variation of temperature and humidity also adds significant additional costs to their shipping, handling, and storage. The third limiting aspect of monoclonal antibodies is their batch-to-batch variation in composition, which can alter the quantitative character of protein assays. Despite all these drawbacks, antibodies still remain the gold standard protein detection in different clinical and diagnostics platforms due to their high affinity and selectivity.

To address the high demand for protein detection agents with antibody like sensitivity in diagnostics platforms while overcoming the inherent disadvantages of antibodies, scientists have tried to develop artificial alternatives to antibodies. There have been two approaches for this development. One is a protein engineering approach using an alternative small and robust protein scaffold with favorable folding and stability, and easy bacterial expression. Notable examples are affibody[5] that is based on Z domain of protein A and DARPins that use ankyrin repeat motifs[6]. The second approach is chemical development of artificial protein capture agentsusing a bottom up approach. Artificial protein capture agents have been synthesized using a variety of using biological building blocks like nucleic acids[7], peptides[8], and carbohydrates[9]. Chemical approaches such as phage display[10,11], nucleic acid aptamers[12], mRNA display[13], OBOC library[14], have been used in to develop artificial antibody like molecules that have been used in detection, imaging, therapeutics etc. Phage display library technology has been extensively used to isolate both linear[15] and cyclic peptide[16] based protein capture agents. However phage display libraries are comprised of L-amino acids, so that peptides isolated from the library are protease sensitive.
An alternative to the phage display library is the One-bead one-compound (OBOC)[17] peptide or peptide mimetic libraries. While in OBOC libraries one can easily modulate chemical stability by incorporating artificial amino acids, OBOC libraries have smaller chemical diversities than phage libraries. Peptide ligand discovery is made typically with pentameric to octameric peptide

libraries, with lower diversity than phage display libraries. This is because in OBOC libraries the on bead peptides have to be highly pure, and coupling of each amino acid decreases the purity.

To exploit the chemical flexibility of comprehensive, OBOC libraries of oligopeptides to develop high affinity, high selectivity protein capture agents, we married the OBOC peptide ligand discovery process with "in situ click chemistry". "Click chemistry" describes a class of chemical reactions that quickly generate substances by joining small modular units[18]. Copper catalyzed Azide Alkyne Cycloaddition (CuAAC) is the quintessential 'click' reaction, where Cu acts as a catalyst in the reaction of azide and alkyne and create a 1,2,3 triazole linkage. A variant of click chemistry, 'in situ click chemistry'[19-21], is particularly relevant to our discussion. In 'in situ click chemistry' the biological target, rather than Cu, acts as the reaction catalyst. The target acts as a scaffold, bringing together complementary building block reagents to create the triazole linkage. In the demonstration of "in situ click" strategy, libraries of small molecules, based upon components of a known inhibitor, were used as building blocks and the assembled compound was this known inhibitor, targeting the active site of the target enzyme. It was observed that the reaction had high selectivity, selecting not only the right building blocks, but also the right orientation between the building blocks, to create a specific regioisomer.

Iterative In Situ Click Chemistry (IISCC)[22,23] is a further development of this approach that we have developed for developing peptide based protein capture agents. Unlike the original approach that expanded upon a known inhibitor, in this approach we use large comprehensive libraries. This allows us to sample a much larger chemical space. In this process, the target is screened against OBOC peptide libraries to produce a peptide binder, typically having low micromolar affinity for the target protein. This peptide is referred to as the 1° ligand or monoligand. The 1° ligand is modified with an alkyne or azide to form a monoligand anchor. The protein is treated with the monoligand anchor and then screened against an OBOC library containing the complementary click handle. The protein acting as a catalyst, selects a second

binder (2° ligand), and clicks it with the anchor peptide to form a biligand. The biligand, modified with a new click handle (biligand anchor), is used, along with the protein, in the next round of ligand discovery. The process is then iterated to form a triligand, etc., until we achieve the desired levels of selectivity and affinity. The iterative nature of ligand discovery allows us to overcome the length limitation of OBOC peptide libraries. In the first demonstration of the IISCC technology, Agnew et al developed a triligand PCC Agent against bovine carbonic anhydrase II (bCAII) demonstrated that the developed PCC Agents could replace the antibody for carbonic anhydrase in protein assays like dot blots and ELISA. The target Carbonic anhydrase II was chosen as it has earlier been extensively characterized to study ligand-protein interactions. The protein has two distinct binding pockets and hence bind bivalent ligands[21,24]

Chapter 1 of my thesis explores the IISCC technology to develop branched PCC agents against bovine carbonic anhydrase II. The reported PCC agents for bCAII had not been allowed to branch on the peptide backbone but was grown off at the C terminal or the N terminal. Hence the process did not exploit the architectural flexibility that arises from the modular nature of the PCC agents. If a PCC agent is allowed to evolve via branching on the 1° or 2° ligand, the architectural space scanned vastly increases. Peptide dendrimers containing branching amino acid cores and surface peptide chains have been reported to have higher binding affinities[25]. Constraining a peptide by branching or cyclizing it increases the geometric constraints and reduces the conformational entropy of the molecule[26] which results in higher affinity. In the current work, starting from the same 1° ligand for bovine carbonic anhydrase, we allow development of branched PCC agents where one peptide arm is built off the backbone of another peptide arm. We discovered that branched ligands developed this way have affinities 2-3 times better than the original linear triligand reported. In ‘In situ click chemistry’ with protein as a catalyst, the protein may yield either anti (1,4), or syn (1,5) regioisomers a mixture of the two as the product. The multiligands that are produced in bulk and characterized, are synthesized

through CuAAC reaction, which means that in the characterized ligands the peptide arms are anti to each other and linked through a 1,4-triazole linker. 1,5 regioisomers are difficult to synthesize in bulk. We overcome this technical difficulty by developing a mimic of the 1,5 triazole linker. This allows us to characterize the PCC agents where the individual peptide arms are syn to each other. In the case of bCAII, we observe that the PCC agent with the 1,5 triazole linker mimic has higher affinity than the 1,4 triazole linked PCC agent.

While one of the major advantages of the PCC capture agent is the generality of the approach, the original process of development does not allow one to target a specific region of the protein, unless one can start with a literature molecule that is known to bind to the region of interest such as the active site. Without this knowledge it is almost impossible to develop PCC agents that can distinguish between subtle variations of the protein structure. However, in diagnostics platforms it is often observed that the salient protein marker parameter is not the quantity of the protein but changes in the protein structure. Disease prognosis may require detection of phosphorylation, glycosylation and other post-translational modifications. Certain non-antibody protein capture agents have been developed to recognize specific post-translational modifications like acylation[27]. Nevertheless, there is no general strategy that can be used to develop these reagents. In chapter 3, we outline a general strategy to chemically develop a peptide based capture agent specifically targeting a certain region of the protein. We describe the epitope targeting strategy and the development of peptide based PCC ligands that target the phosphorylated C-terminal hydrophobic motif of the protein Akt2. Akt, an AGC kinase, plays a central regulatory role in growth factor signaling, and serves as a key node in the phosphatidylinositol 3-kinase (PI3k) signaling pathway[28,29]. In the fully active Akt2 protein, two residues T309 and S474 are phosphorylated. T309 is located near the ATP and substrate binding pocket. It is the primary phosphorylation site required for the activity of the kinase. S474 is located in the C terminal of Akt2, and is part of unstructured region of the protein. However,

S474 is known to be important for the kinase activity of the protein and phosphorylation at S474 enhances the kinase activity 10-fold[30]. In the context of current studies showing that ATP competitive inhibitors of Akt can cause hyperphosphorylation of the protein[31], we thought it was worthwhile to find binders to the C terminal kinase regulatory domain, as such binders would have a high potential to regulate kinase activity. In this epitope targeting strategy, the target protein region is chemically modified to incorporate an azide/alkyne handle. The target, the C-terminal polypeptide fragment of Akt2 (amino acids 450-481) with the phosphorylated Serine 474 (p-S474), is modified with a dinuclear Zn (II) DPA type complex, which has an appended biotin label and an azide handle. The dinuclear Zn (II) DPA selectively binds to the phosphate anion and provides an initial in situ click reaction site adjacent to the phosphorylated residue azide functionality close to the hydrophobic motif. This phosphopeptide/metal complex is then subjected to an in situ click screen against a large one-bead-one-compound (OBOC) library of acetylene-containing 6-mer peptides to identify an initial (1^o) peptide ligand. For an in situ click reaction, the binding target itself (here the 32-mer peptide) provides a highly selective catalytic scaffold for promoting the click reaction between an azide and an alkyne to produce a triazole linkage. The process of sequential in situ click chemistry is then utilized to expand the 1^o ligand into a biligand, and then the biligand into two triligands.

In chapter 4, we describe the assays used to choose the peptide arms, and characterize the final developed ligands in additional assays. The three triligand candidates have low nanomolar binding affinities. One triligand is selective for the Akt2 and over its homologous isoforms Akt1 and Akt3. We also explore their effect on the kinase activity of Akt. Remarkably, the two triligands, sharing the same biligand fragment, were found to have opposite effects on the kinase activity of the Akt2 protein. While one of the ligands enhances the kinase activity of the protein, the other ligand inhibits the kinase activity.

In the epitope targeting strategy, we are trying to detect peptide-peptide interaction, which is typically weaker than peptide-protein interaction and tougher to detect. If the 1° ligand chosen has a high affinity for its peptide epitope, it significantly simplifies the epitope targeted PCC agent development process. This led us to explore cyclic peptides, which have higher affinity and selectivity than their corresponding linear counterparts. For demonstrating of this process, cyclic ligands were developed against the C terminal fragment from the protein Akt2, amino acids 450-481. In chapter 5 we describe the development of OBOC CuAAC cyclized peptide libraries, which are then screened against the peptide epitope to isolate a cyclic peptide monoligand. We demonstrate that the isolated monoligand binds to the peptide epitope and the full-length protein. The cyclic ligand has a low nanomolar binding affinity, rather than a micromolar binding affinity that is characteristic of a linear 1° ligand. The ligand is also found to inhibit the kinase activity of the protein. Future studies shall be done to develop the biligand and to study the effects in live cells.

REFERENCES:

(1) Hood, L.; Heath, J. R.; Phelps, M. E.; Lin, B. *Science* **2004**, *306*, 640.

(2) Heath, J. R.; Davis, M. E. *Annual review of medicine* **2008**, *59*, 251.

(3) Rifai, N.; Gillette, M. A.; Carr, S. A. *Nature biotechnology* **2006**, *24*, 971.

(4) Ludwig, J. A.; Weinstein, J. N. *Nat Rev Cancer* **2005**, *5*, 845.

(5) Orlova, A.; Magnusson, M.; Eriksson, T. L. J.; Nilsson, M.; Larsson, B.; Hoiden-Guthenherg, I.; Widstrom, C.; Carlsson, J.; Tolmachev, V.; Stahl, S.; Nilsson, F. Y. *Cancer Research* **2006**, *66*, 4339.

(6) Binz, H. K.; Amstutz, P.; Kohl, A.; Stumpp, M. T.; Briand, C.; Forrer, P.; Grutter, M. G.; Pluckthun, A. *Nature Biotechnology* **2004**, *22*, 575.

(7) Famulok, M.; Mayer, G.; Blind, M. *Acc Chem Res* **2000**, *33*, 591.

(8) Kunz, C.; Borghouts, C.; Buerger, C.; Groner, B. *Mol Cancer Res* **2006**, *4*, 983.

(9) Horlacher, T.; Seeberger, P. H. *Chemical Society Reviews* **2008**, *37*, 1414.

(10) Smith, G. P.; Petrenko, V. A. *Chem Rev* **1997**, *97*, 391.

(11) Wu, P.; Leinonen, J.; Koivunen, E.; Lankinen, H.; Stenman, U. H. *Eur J Biochem* **2000**, *267*, 6212.

(12) Sharma, A. K.; Kent, A. D.; Heemstra, J. M. *Anal Chem* **2012**, *84*, 6104.

(13) Takahashi, T. T.; Austin, R. J.; Roberts, R. W. *Trends in Biochemical Sciences* **2003**, *28*, 159.

(14) Lam, K. S.; Salmon, S. E.; Hersh, E. M.; Hruby, V. J.; Kazmierski, W. M.; Knapp, R. J. *Nature* **1991**, *354*, 82.

(15) Koivunen, E.; Gay, D. A.; Ruoslahti, E. *Journal of Biological Chemistry* **1993**, *268*, 20205.

(16) Katz, B. A. *Biochemistry* **1995**, *34*, 15421.

(17) Lam, K. S.; Lebl, M.; Krchnák, V. *Chem Rev* **1997**, *97*, 411.

(18) Kolb, H. C.; Finn, M. G.; Sharpless, K. B. *Angewandte Chemie-International Edition* **2001**, *40*, 2004.

(19) Manetsch, R.; Krasiński, A.; Radić, Z.; Raushel, J.; Taylor, P.; Sharpless, K. B.; Kolb, H. C. *Journal of the American Chemical Society* **2004**, *126*, 12809.

(20) Mamidyala, S. K.; Finn, M. G. *Chem Soc Rev* **2010**, *39*, 1252.

(21) Mocharla, V. P.; Colasson, B.; Lee, L. V.; Roper, S.; Sharpless, K. B.; Wong, C. H.; Kolb, H. C. *Angew Chem Int Ed Engl* **2004**, *44*, 116.

(22) 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, K. B.; Heath, J. R. *Angewandte Chemie-International Edition* **2009**, *48*, 4944.

(23) Millward, S. W.; Henning, R. K.; Kwong, G. A.; Pitram, S.; Agnew, H. D.; Deyle, K. M.; Nag, A.; Hein, J.; Lee, S. S.; Lim, J.; Pfeilsticker, J. A.; Sharpless, K. B.; Heath, J. R. *Journal of the American Chemical Society* **2011**, *133*, 18280.

(24) Jain, A.; Whitesides, G. M.; Alexander, R. S.; Christianson, D. W. *Journal of Medicinal Chemistry* **1994**, *37*, 2100.

(25) Adesida, A. B.; Aojula, R. R.; Aojula, H. S.; Clarke, D. J. *Vaccine* **1999**, *18*, 315.

(26) Doig, A. J.; Sternberg, M. J. *Protein Sci* **1995**, *4*, 2247.

(27) Lin, L.; Fu, Q.; Williams, B. A.; Azzaz, A. M.; Shogren-Knaak, M. A.; Chaput, J. C.; Lindsay, S. *Biophysical journal* **2009**, *97*, 1804.

(28) Vivanco, I.; Sawyers, C. L. *Nat Rev Cancer* **2002**, *2*, 489.
(29) Manning, B. D.; Cantley, L. C. *Cell* **2007**, *129*, 1261.
(30) Yang, J.; Cron, P.; Good, V. M.; Thompson, V.; Hemmings, B. A.; Barford, D. *Nat Struct Biol* **2002**, *9*, 940.
(31) Okuzumi, T.; Fiedler, D.; Zhang, C.; Gray, D. C.; Aizenstein, B.; Hoffman, R.; Shokat, K. M. *Nat Chem Biol* **2009**, *5*, 484.