Title
Development of Chemical Crosslinkers to Trap-Substrate Pairs

Permalink
https://escholarship.org/uc/item/6d64n3vz

Author
Riel-Mehan, Megan

Publication Date
2013

Peer reviewed|Thesis/dissertation

Development of Chemical Crosslinkers to Trap Kinase-Substrate
Pairs

by

Megan Riel-Mehan

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Chemistry and Chemical Biology

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

Dedication and Acknowledgments Page

I would like to thank Kevan Shokat for his guidance on this project and his mentorship and support for my career goals. If I ever have the chance to be a mentor or boss in my future, I am going do my best to live up to his example. I want to thank Dustin Maly and Sasha Statsuk for their work on this project and Sasha for being a mentor to me as well. The entire Shokat lab has been a great place to grow as a scientist. Mike Lopez has been an invaluable source of ideas and support, I am very glad I have been sitting next to him for the last five years. Greg Hamilton and Jon Ostrem came up with an idea that tuned my project around. Sam Myers and Dan Le have been great colleagues, friends, and comrades in arms. This thesis is dedicated to my brother, Michael, who has always gone first in the world and been gracious enough to turn around and help me on my way.

Sasha Statsuk established the crosslinking assay used in Chapter 2, he designed the synthesis of compound **1** and the western blots shown Figure 2(A-C) of that chapter were preformed by him. With this exception all experiments herein were preformed by Megan M. Riel-Mehan under the direction of Kevan M. Shokat.

Abstract

The sheer number of kinases and phosphoproteins makes the important task of identifying kinase-substrate pairs very difficult. New phophosphoproteins are being identified by proteomic techniques, yet there is currently no robust method for finding a kinase that preformed a particular phosphorylation event. I report in this thesis on such a method that covalently traps a kinase to the corresponding substrate through a crosslinking ATP mimetic to enable identification of upstream kinases by mass spectrometry.

Table of Contents

List of Tables

Chapter 3

List of Figures and Illustrations

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Chapter 1

Use of activity based probes and chemical genetics to study kinase

activity

A fundamental and essential characteristic of living cells is the ability to respond to their environment by integrating external stimuli into a directed response, such as whether or not to grow, move, differentiate, or attack neighboring cells. They do so through a tightly regulated and complex system of signaling molecules, which transfer information throughout the cell via a cascade of chemical reactions known as post-translational modifications (PTMs). These modifications can be quickly installed and quickly removed by counteracting enzymes, allowing for rapid changes and responses within the cell. Much of current cancer biology, immunology, epigenetics and neuroscience research is focused on identifying the participants in these chemical reactions and discovering how their concerted action produces larger cellular consequences.[1][2][3]

Protein phosphorylation represents one of the most ubiquitous types of post-translation modification[4]. The transfer of the gamma phosphate of adenosine triphosphate (ATP) to a serine, threonine or tyrosine residue on a protein is catalyzed by kinases. This transfer of a negatively charged moiety can induce conformational changes in the substrate, resulting in modulation of enzyme activity, changes in binding partners or both. Misregulation of kinases is responsible for many diseases, and thus they are an attractive area for study both for better understanding of biology and for therapeutic agents[5]. The scale of the human kinome, which is comprised of approximately 500 members, makes studying any one kinase very challenging[6]. Large scale efforts to map human kinase networks include phosphopeptide enrichment coupled with mass spectrometric identification. These reveal the vast array of protein kinase substrates, currently

numbered at greater than 100,000[7], but lack the ability to measure phosphorylation events in a temporal and spatial manner.
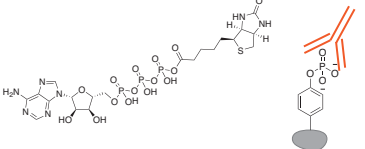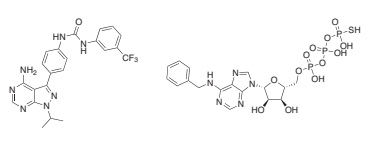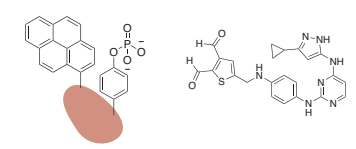
To understand cellular networks, kinase phosphorylation cascades need to be specifically delineated. However, vastly disparate amounts of information exist about how kinases transmit different signals and different techniques are needed given these varying degrees of a priori knowledge. Chemical biological techniques have found particular utility in addressing these problems and many have been summarized in few recent reviews[8-10]. Here we look at what techniques are applicable given the amount of information one has on hand (**Figure 1.1**). When little is known about how individual kinases or substrates mediate a biological process in a given system, methods to look at global changes are useful. These methods can give information about the proteins involved but not any information about the connectivity of the network. If a specific kinase is the point of interest there are methods to discover both the biological processes the kinase is involved in and its downstream substrate(s). Lastly, it is often easier to identify a protein that is phosphorylated in a given condition than a kinase involved in a pathway, and therefore techniques that elucidate the signaling network starting from the substrate are needed.

Techniques to monitor or control kinase events en mass

Often in the early stages of studying a system, a new cell type, or the effect of a perturbation, little may be known about the individual proteins involved controlling the biological outcomes, so tools that give a global picture of the system are needed. These

Figure 1.1. Illustration of kinome complexity and chemical biology tools to study signaling networks. Mapping of kinase signaling networks includes identify active kinases, finding phosphorylation sites and determining kinase-substrate pairs. Unbiased techniques reveal active kinases and phosphorylated proteins. Specific kinase inhibitors can target a known kinase to uncover physiological roles of the kinase by monitoring changes in phenotype and downstream phosphorylation events. Using a substrate as bait can identify kinases responsible for a specific phosphorylation of the substrate.

tools can answer questions about overall changes in phosphorylation, without knowing the specific substrates, or they can reveal which enzymes are activated under the conditions being studied. These tools are often powerful at illuminating a direction to follow for further study.

One way to look at overall changes in a system under different conditions is using phospho-specific antibodies. Phosphotyrosine specific antibodies can detect and enrich phosphotyrosine containing proteins and peptides[11]. Unfortunately, general phosphoserine/threonine antibodies are typically lower affinity, and as such are not as useful in this manner[12].

In order for signaling networks to quickly respond to stimuli, kinases need to rapidly switch between an active state and an inactive state for the given condition and the desired biological outcome. In order to study a signaling network in a particular state, such as cell division, it is essential to know which kinases have switched their activity status to on or off. Activity-based probes (ABPs) have been developed to target the active state of enzymes, usually by labeling the enzyme with a tag for either visualization or identification. Inactive kinases are often in a conformational state that renders them incapable of binding ATP, so kinase ABPs are composed of a kinase ATP pocket binding element with a reactive group and a tag. ABPs have been used to study the kinome under different cellular conditions, or to screen for reversible inhibitors. Patricelli and co-workers reported a biotinylated derivative of the acyl phosphate ADP, which has been shown to react with more than 400 human kinases via their conserved catalytic lysine residue[13,14].They employ this APB as a screen for kinase inhibitors by incubating small

molecules in whole cell lysates followed by ATP-botin. Only a kinase with an unoccupied ATP pocket will get biotinylated, therefore potential targets for the inhibitor are any kinase that is not isolated when the small molecule is present but is isolated in the control group. Other ABPs make use of known promiscuous kinase inhibitor scaffolds, such as the natural product Wortmanin. Wortmanin tethered to a flurophore has been used to label PI3 kinases and Plk1 and Plk3[15].

The two previous strategies mentioned give little information about kinase network connectivity, the knowledge of which is essential for fully understanding a network. Kinases phosphorylate their substrates by directly binding to them. Therefore all active kinases have an inherent binding affinity for their substrates, but unfortunately this binding affinity is usually too weak to identify kinase-substrate pairs by techniques such as affinity purification that rely on protein-protein interactions. One way to get around this problem is to convert the weak protein-protein interaction to a covalent bond through a crosslinking reaction. Protein crosslinking reagents have provided a rich source of information about protein-protein interactions in various biological system, and a few labs have been attempting to develop such a molecule for kinase-substrate pairs. Cole and co-workers reported a photo-crosslinking molecule which are derivatized with an aryl azide moiety at the gamma phosphate and another azide moiety at the C^8 position on the adenine ring[16]. This compound has only ever crosslinked kinase-substrate pairs in purified conditions. More recently, Suwal and Pflum have reported another aryl azide based photo-crosslinker[17]. These molecules react with the hydroxyl group on the substrate, much like ATP-based ABPs, leaving behind a photo-crosslinker, which then is

able to react with the kinase upon exposure to UV irradiation. This crosslinking reaction was demonstrated between casein kinase II and casein with low yield. Given the number of kinases and substrates, a crosslinker that is designed to react with every kinase and substrate pair may produce too many products in a cellular assay, although such a compound may be useful for trapping complexes for the purpose of crystallization.

Techniques for a known kinase of interest

Since the development of imatinib in the late 1990s, creating new inhibitors that target kinases have been the goal of many researchers. Aside from their use is as therapeutic agents, small molecule inhibitors can help answer questions about a specific kinase's role in the cell. Most inhibitors target the ATP pocket, which is highly conserved in kinases, and therefore designing a specific inhibitor can be a challenge, but despite this difficulty kinase inhibitors have been successfully designed to target a wide variety of kinases and have helped elucidate their biological roles[18]. Irreversible kinase inhibitors can be converted to chemical probes by the addition of a tag. In one such approach, Taunton and co-workers took a molecule they designed and synthesized to be an irreversible RSK kinase inhibitor, and converted the molecule to a kinase specific ABP by adding a alkyne linker to enable labeling by a tag[19].

If the interest is purely biological and not therapeutic, instead of tailoring an inhibitor to the kinase of interest, it is possible to alter the kinase to accept a particular inhibitor. The challenge then becomes ensuring the kinase activity and/or structure is preserved. This has been accomplished in many kinases by mutating the residue in the

ATP pocket that resides above the N^6 atom of ATP to a n glycine or alanine residue and thus opening up a hydrophobic pocket. The resulting analog-sensitive (AS) kinase can be targeted by inhibitors with corresponding bulky groups designed to occupy the newly accessible pocket, while wild type kinases are unable to bind the inhibitor due to steric clash with the larger naturally occurring gatekeeper residues. This method has been successfully used in both yeast and mammalian cells to elucidate kinase functions[20,21].

The power of this method is its broad application due to the conserved nature of the ATP pocket. However, there are kinases that do not tolerate a small gatekeeper residue. To build on this method, Shokat and coworkers have demonstrated the use of a cysteine residue in place of the gatekeeper residue to add a new chemical reactive group (handle) instead of an enlarged additional pocket[22].

The techniques mentioned above all inhibit the kinase of interest which can indirectly help identify substrates of a specific kinase, but they cannot distinguish between substrates that are directly phosphorylated and those that are simply downstream. Tools for identifying the substrates of a given kinase either test the ability of a kinase to phosphorylate different individual substrates, or attempt to trace the activity of a the kinase in a mixture of possible substrates. In vitro substrate discovery is carried out by incubating the kinase with the possible substrates to see if it is possible for the phosphorylation to occur. The first obvious limitation of this technique is the low throughput nature, since each putative substrate needs to be incubated individually. Additionally, kinases that phosphorylate a substrate in vitro do not necessarily do so in vivo, either due to different subcellular localization or additional binding partners. To get

around the first limitation protein kinase assays can be conducted on arrays imprinted with candidate substrates for a hight throughput format. This has been mostly carried out with peptide substrates, which is one way kinase phosphorylation consensus motifs have been established.

One of the early techniques for in vitro substrate identification with full length proteins is Kinase Substrate Tracking and Elucidation (KESTREL)[23]. In this approach, a kinase of interest in added into cell lysate with a kinase assay buffer containing γ^{32}P-ATP. The proteins in the sample are separated by electrophoresis and radioactive bands correspond to putative targets. This technique is hampered with a high level of background signal since there are already kinases present in the lysate. This drawback was overcome by heat-inactivating the lysate prior to the assay[24]. Despite improvements, this technique has not been able to identify more than a few substrates per kinase. Protein chips engineered to contain an array of yeast GST-fusion proteins have been used in a kinase-radiolabeling assay, which led to the identification of hundres of potential new substrates for the yeast kinome[25]. However, the site of modification cannot be identified using this technique.

To discover more relevant substrates and identify the site of phosphorylation, the AS kinases can be used in conjunction with an ATP analogs. Just as the bulky inhibitors are only accepted by the AS kinase containing the pocket to accommodate them, AS kinases can accept an ATP analog molecule with a bulky group built from the N6 nitrogen of ATP , while wild-type kinases cannot. By replacing the gamma phosphate with a thiophosphate, this orthogonal reaction is traceable in the presence of all other

phosphorylation reactions. Phosphorylation with the analog results in the attachment of a

thiophosphate group to substrates of the AS kinase, which can be isolated either by

antibody-based enrichment or covalent capture using iodoacetyl agarose, and identified

by mass spectrometry[26]. Although this method has been most widely applied to in vitro

study, through permeabilzation with digitonin, ATPɣS can enter cells and be used for in

vivo substrate identification[27].

Techniques with a known substrate

Given the growing size of the known phophoproteome, it is often the case that a

researcher knows a site is phosphorylated but does not know the kinase or kinases that

reacted with their protein of interest. When this is the case, many of the techniques

outlined above are not applicable. The ability to match an upstream kinase with a

phosphorylation site is still an unsolved problem. One approach is to match the sequence

motif flanking the phosphorylation residue to what is known about the substrate

preferences of protein kinases, based on their ability to phosphorylate synthetic peptides

in vitro and or their known cellular substrates. Bioinformatic algorithms have been

developed for this purpose[28-30]. These algorithms are only a first step, since they are an

oversimplification of complicated kinase-substrate relationships.

If the goal is to monitor the phosphorylation of the substrate of interest, then

antibodies and substrate based biosensors can used. Antibodies raised against either a

particular phosphorylation site or against a residue specific type of phosphorylation have

been invaluable in studying signaling cascades. Engineered substrates can report

phosphorylation events through fluorescence. As stated above, the addition of a phosphate group can induce conformational changes in a substrate, and by introducing reporter molecules onto the substrate these conformational changes can be taken advantage of. Synthetic peptide-based biosensors have been used successfully to monitor activity of many kinases in including PKC, Cdk2, PKA, Akt, MK2, and Pim2[31]. Phosphorylation increases the sensors affinity for the Mg^{2+} ion present in the active site of the kinase, and this chelation increases the fluorescence of Sox, an unnatural amino acid present on the bio-sensor. In the same vein, genetically encoded protein biosensors have been developed. They have the ability to track phosphorylation changes in real-time. There are a few different designs but they all utilize some conformational change in the biosensor upon phosphorylation as to trigger either a Forster resonance energy transfer (FRET) signal, luminescence or fluorescence[32,33]. Although these types of sensors have been mainly utilized to study a known kinase's activity, if a full length substrate is used as the sensor, knowing the upstream kinase or kinases is unnecessary for being able to follow the phosphorylation of the biosensor.

Large scale screens similar to the substrate screens mentioned above have been developed to discover kinases that can phosphorylate a substrate of interest. In this approach a library of kinases are each incubated with a peptide containing the phosphorylation site. Similar to the kinase screens, this is a labor intensive process and requires that the kinases are able to be expressed recombinantly.

Finally, the approach that has the most promise for direct kinase identification is protein crosslinking. The advantage of crosslinking is the kinases and substrate can be

incubated together, eliminating the need for individual reactions for each putative kinase. Additionally, when preformed with cell lysate, other binding partners that may be important for in vivo phosphorylation are also present in the mixture. Crosslinking techniques using a specific phosphoprotein as bait have been developed in our lab. One of the key differences between our method and other crosslinking methods described above is that the reaction will occur only with the substrate of interest, reducing the amount of uninteresting crosslinking products. This is analogous to the reverse method of using an AS kinase to create a bio-orthogonal chemical reaction. This method remains the only example of a substrate-kinase crosslinking in cell lysate, but optimization is still need to make the method robust enough to be carried out in an unbiased manner.

References

(1) Krueger, K. E.; Srivastava, S. Posttranslational protein modifications: current implications for cancer detection, prevention, and therapeutics. *Mol. Cell Proteomics* **2006**, *5*, 1799–1810.

(2) Sidoli, S.; Cheng, L.; Jensen, O. N. Proteomics in chromatin biology and epigenetics: Elucidation of post-translational modifications of histone proteins by mass spectrometry. *J Proteomics* **2012**, *75*, 3419–3433.

(3) Charlot, C.; Dubois-Pot, H.; Serchov, T.; Tourrette, Y.; Wasylyk, B. A review of post-translational modifications and subcellular localization of Ets transcription factors: possible connection with cancer and involvement in the hypoxic response. *Methods Mol. Biol.* **2010**, *647*, 3–30.

(4) Khoury, G. A.; Baliban, R. C.; Floudas, C. A. Proteome-wide post-translational modification statistics: frequency analysis and curation of the swiss-prot database. *Sci. Rep.* **2011**, *1*, –.

(5) Lapenna, S.; Giordano, A. Cell cycle kinases as therapeutic targets for cancer. *Nature Reviews Drug Discovery* **2009**, *8*, 547–566.

(6) Manning, G. The Protein Kinase Complement of the Human Genome. *Science* **2002**, *298*, 1912–1934.

(7) Hornbeck, P. V.; Chabra, I.; Kornhauser, J. M.; Skrzypek, E.; Zhang, B. PhosphoSite: A bioinformatics resource dedicated to physiological protein phosphorylation. *PROTEOMICS* **2004**, *4*, 1551–1561.

(8) H S Lu, C.; Liu, K.; Tan, L. P.; Yao, S. Q. Current chemical biology tools for studying protein phosphorylation and dephosphorylation. *Chemistry* **2012**, *18*, 28–39.

(9) Martić, S.; Kraatz, H.-B. Chemical biology toolkit for exploring protein kinase catalyzed phosphorylation reactions. *Chemical Science* **2013**, *4*, 42–59.

(10) Knight, J. D.; Pawson, T.; Gingras, A.-C. Profiling the kinome: current capabilities and future challenges. *J Proteomics* **2012**.

(11) Bergström Lind, S.; Molin, M.; Savitski, M. M.; Emilsson, L.; Aström, J.; Hedberg, L.; Adams, C.; Nielsen, M. L.; Engström, A.; Elfineh, L.; Andersson, E.; Zubarev, R. A.; Pettersson, U. Immunoaffinity enrichments followed by mass spectrometric detection for studying global protein tyrosine phosphorylation. *J. Proteome Res.* **2008**, *7*, 2897–2910.

(12) Grønborg, M.; Kristiansen, T. Z.; Stensballe, A.; Andersen, J. S.; Ohara, O.; Mann, M.; Jensen, O. N.; Pandey, A. A Mass Spectrometry-based Proteomic Approach for Identification of Serine/Threonine-phosphorylated Proteins by Enrichment with Phospho-specific Antibodies. *Mol. Cell Proteomics 1*, 517–527.

(13) Patricelli, M. P.; Nomanbhoy, T. K.; Wu, J.; Brown, H.; Zhou, D.; Zhang, J.; Jagannathan, S.; Aban, A.; Okerberg, E.; Herring, C.; Nordin, B.; Weissig, H.; Yang, Q.; Lee, J.-D.; Gray, N. S.; Kozarich, J. W. In Situ Kinase Profiling Reveals Functionally Relevant Properties of Native Kinases. *Chemistry & Biology* **2011**, *18*, 699–710.

(14) Patricelli, M. P.; Szardenings, A. K.; Liyanage, M.; Nomanbhoy, T. K.; Wu, M.; Weissig, H.; Aban, A.; Chun, D.; Tanner, S.; Kozarich, J. W. Functional interrogation of the kinome using nucleotide acyl phosphates. *Biochemistry* **2007**, *46*, 350–358.

(15) Liu, Y.; Jiang, N.; Wu, J.; Dai, W.; Rosenblum, J. S. Polo-like kinases inhibited by wortmannin. Labeling site and downstream effects. *J. Biol. Chem.* **2007**, *282*, 2505–2511.

(16) Parang, K.; Kohn, J. A.; Saldanha, S. A.; Cole, P. A. Development of photo-crosslinking reagents for protein kinase-substrate interactions. *FEBS Letters* **2002**, *520*, 156–160.

(17) Suwal, S.; Pflum, M. K. H. Phosphorylation-Dependent Kinase-Substrate Cross-Linking. *Angewandte Chemie* **2010**, *122*, 1671–1674.

(18) Dar, A. C.; Shokat, K. M. The evolution of protein kinase inhibitors from antagonists to agonists of cellular signaling. *Annual review of biochemistry* **2011**, *80*, 769–795.

(19) Cohen, M. S.; Zhang, C.; Shokat, K. M.; Taunton, J. Structural bioinformatics-based design of selective, irreversible kinase inhibitors. *Science* **2005**, *308*, 1318–1321.

(20) Shokat, K. M.; Bishop, A. C.; Ubersax, J. A.; Petsch, D. T.; Matheos, D. P.; Gray, N. S.; Blethrow, J.; Shimizu, E.; Tsien, J. Z.; Schultz, P. G.; Rose, M. D.; Wood, J. L.; Morgan, D. O. A chemical switch for inhibitor-sensitive alleles of any protein kinase

. *Nature* **2000**, *407*, 395–401.

(21) Bishop, A. C.; Shah, K.; Liu, Y.; Witucki, L.; Kung, C.-Y.; Shokat, K. M. Design of allele-specific inhibitors to probe protein kinase signaling. *Current Biology* **1998**, *8*, 257–266.

(22) Garske, A. L.; Peters, U.; Cortesi, A. T.; Perez, J. L.; Shokat, K. M. Chemical genetic strategy for targeting protein kinases based on covalent complementarity. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 15046–15052.

(23) Knebel, A. A novel method to identify protein kinase substrates: eEF2 kinase is phosphorylated and inhibited by SAPK4/p38delta. *EMBO J* **2001**, *20*, 4360–4369.

(24) Troiani, S.; Uggeri, M.; Moll, J.; Isacchi, A.; Kalisz, H. M.; Rusconi, L.; Valsasina, B. Searching for Biomarkers of Aurora-A Kinase Activity: Identification of in Vitro Substrates through a Modified KESTREL Approach. *J. Proteome Res.* **2005**, *4*, 1296–1303.

(25) Ghaemmaghami, S.; Huh, W.-K.; Bower, K.; Howson, R. W.; Belle, A.; Dephoure, N.; O'Shea, E. K.; Weissman, J. S. Global analysis of protein expression in yeast. *Nature* **2003**, *425*, 737–741.

(26) Hertz, N. T.; Wang, B. T.; Allen, J. J.; Zhang, C.; Dar, A. C.; Burlingame, A. L.; Shokat, K. M. *Chemical Genetic Approach for Kinase-Substrate Mapping by Covalent Capture of Thiophosphopeptides and Analysis by Mass Spectrometry*; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2009.

(27) Banko, M. R.; Allen, J. J.; Schaffer, B. E.; Wilker, E. W.; Tsou, P.; White, J. L.; Villén, J.; Wang, B.; Kim, S. R.; Sakamoto, K. Chemical genetic screen for AMPKα2 substrates uncovers a network of proteins involved in mitosis. *Molecular Cell* **2011**, *44*, 878–892.

(28) Obenauer, J. C. Scansite 2.0: proteome-wide prediction of cell signaling interactions using short sequence motifs. *Nucleic Acids Research* **2003**, *31*, 3635–3641.

(29) Linding, R.; Jensen, L. J.; Pasculescu, A.; Olhovsky, M.; Colwill, K.; Bork, P.; Yaffe, M. B.; Pawson, T. NetworKIN: a resource for exploring cellular phosphorylation networks. *Nucleic Acids Research* **2007**, *36*, D695–D699.

(30) Miller, M. L.; Jensen, L. J.; Diella, F.; Jørgensen, C.; Tinti, M.; Li, L.; Hsiung, M.; Parker, S. A.; Bordeaux, J.; Sicheritz-Ponten, T.; Olhovsky, M.; Pasculescu, A.; Alexander, J.; Knapp, S.; Blom, N.; Bork, P.; Li, S.; Cesareni, G.; Pawson, T.; Turk, B. E.; Yaffe, M. B.; Brunak, S.; Linding, R. Linear Motif Atlas for Phosphorylation-Dependent Signaling. *Science Signaling* **2008**, *1*, ra2.

(31) Shults, M. D.; Carrico-Moniz, D.; Imperiali, B. Optimal Sox-based fluorescent chemosensor design for serine/threonine protein kinases. *Analytical Biochemistry* **2006**, *352*, 198–207.

(32) Herbst, K. J.; Allen, M. D.; Zhang, J. Luminescent Kinase Activity Biosensors Based on a Versatile Bimolecular Switch. *J. Am. Chem. Soc.* **2011**, *133*, 5676–5679.

(33) Kawai, Y.; Sato, M.; Umezawa, Y. Single color fluorescent indicators of protein phosphorylation for multicolor imaging of intracellular signal flow dynamics. *Analytical chemistry* **2004**, *76*, 6144–6149.

Chapter 2

Investigating factors that effect a three component reaction that

traps kinase and substrate pairs

Abstract

The ubiquitous nature of protein phosphorylation makes it challenging to map kinase substrate relationships, which is needed to construct an understanding of signaling networks. To identify a kinase that is responsible for a known phosphorylation event we have been developing a chemical reaction able to crosslink an orthogonal tagged substrate to its corresponding upstream kinase. The substrate, which is engineered to contain a cysteine in place of the phosphorylable residue and the kinase, with its conserved catalytic lysine, react with *ortho* aromatic dialdehyde based crosslinkers to produce a covalently linked complex. The generality of this technique was demonstrated with a panel of kinases and corresponding peptide substrates. The overall reaction is sensitive to changes in the binding affinity of the crosslinker and the complexity of the reaction mixture.

Introduction.

Protein phosphorylation is one of the most dominant types of covalent post-translational modifications and serves to rapidly transmit information flow in cells and organisms[1,2]. The phosphorylation signaling network is tightly regulated to ensure proper biological responses to stimuli. One way this regulation occurs is through protein-protein interactions between kinases and substrate. Identification of the kinase-substrate pairs is therefore essential to reveal physiological functions of kinases. Several years ago we, and others, developed methods to identify substrates of protein kinases[3]. The reverse problem of mapping unknown kinases into known phosphorylation sites, however, remains a challenge.

The originally developed *o*-phthaldialdehyde based cross-linker covalently crosslinked cysteine containing substrate peptides with their corresponding kinases *in vitro* but not in cell lysates[4]. We subsequently improved the crosslinker by replacing adenosine based directing group of the crosslinker with the promiscuous kinase inhibitor scaffold, while highly reactive *o*-phthaldialdehyde moiety was replaced with less reactive thiophene-2,3-dialdehyde. The resulting kinase-substrate cross-linker covalently cross-linked cysteine containing peptide substrates with recombinant kinases in the presence of competing cell lysate proteins[5]. In this three component, kinase-substrate cross-linking reaction, kinase serves as *macro*-molecular component, which displays specific binding interactions with two *small*-molecular components crosslinker **1** and the corresponding cysteine containing peptide substrate, leading to the formation of a single cross-linked product (**Figure 2.1**). We explore factors that affect efficiency and chemoselectivity of this three-component kinase-substrate cross-linking reaction. The observed

Figure 2.1. Scheme of crosslinking reaction. The conserved lysine reacts with the aldehyde on the crosslinker to form an imine. The cysteine residue from the pseudo-substrate peptide then attacks to form a covalent product.
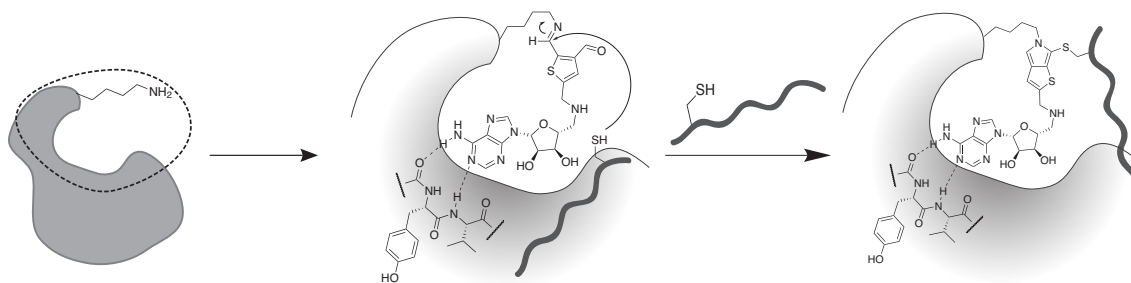
chemoselective kinase-substrate crosslink is dependent on specific kinase-substrate binding interactions, and is enhanced when detergent or cell lysate is added to the reaction media. In the absence of detergent or cell lysate, kinase-substrate cross-linking reaction displayed low chemo-selectivity, probably due to the non-specific kinase-substrate binding. We have subsequently found that the affinity of directing group in the kinase-substrate crosslinker determines the efficiency and the selectivity of the kinase-substrate crosslinking. Kinase substrate cross-linker that is based on high affinity kinase inhibitor scaffold produces more covalently crosslinked kinase-substrate product, while a crosslinker that is based on low affinity adenosine scaffold is less efficient at producing a kinase-substrate product. Taken together our findings indicate tunable and chemo-selective nature of the three-component kinase-substrate crosslinking reaction. The three component kinase-substrate crosslinking reaction can be exploited for selective covalent trapping of kinase-substrate complexes in native proteomes, and their subsequent identification via mass spectrometry.

Results and Discussion

We used Casein Kinase 2 (CKII), Protein Kinase A (PKA), and tyrosine kinase c-SRC as model kinases. Their substrates peptides were derived from known consensus motifs with a cysteine residue in place of the phosphorylatable residue: biotin-$(PEG)_2$-RRADDDDCDD-OH (**Cys-CKIItide**), biotin-$(PEG)_2$-GGLRRACLG-OH (**Cys-PKAtide**), and biotin-$(PEG)_2$-ICGEFRR-OH (**Cys-cSrctide**).

We have previously reported a new thiophene-2,3-dialdehyde based cross-linker (**1**) which is capable of covalent cross-linking recombinant kinase Akt with its putative peptide substrate in the presence of cell lysate. To test the generality of our method we used other serine and tyrosine kinases with their corresponding cysteine containing peptide substrates. Tyrosine kinase c-Src, serine/threonine kinase PKA, and CKII were successfully crosslinked in the absence or presence of cell lysate (**Figure 2.2A-C**). Given that the affinity of kinases for ATP and kinase inhibitor ligands varies within the kinase family, we expect that the efficiency of the crosslinking reaction should be different for each kinase.

Interestingly, treatment of cell lysate alone with the cysteine containing CKII kinase peptide substrate and crosslinker **1**, produced one major labeled band of approximate 50 kDa molecular weight (**Figure 2.2C, lane 3**) Given the high abundance of CKII kinase in HeLa cell lysate, we attributed this labeled band to endogenously labeled CKII kinase. Treatment of HeLa cell lysate with cysteine containing CKII kinase peptide substrate and crosslinker **1**, followed by incubation with streptavidin beads, washing and elution, confirmed the presence of CKII kinase in the elution fraction

Figure 2.2. General and selective features of the three-component kinase-substrate cross-linking reaction. All reactions were resolved with SDS-PAGE, transferred to nitrocellulose and probed for biotin with a streptavidin-HRP conjugate. (A-C) Cross-linking of PKA, cSrc, and CKII kinases and their peptide substrates with the crosslinker **1** in the absence or presence of HeLa cell lysates. Cysteine containing substrate peptides, kinase and HeLa cell lysate were treated with cross-linker and incubated for 30 min at r.t. (D) Pulldown of CKII from HeLa cell lysate with crosslinker **1**. (E) Control reaction without crosslinker.

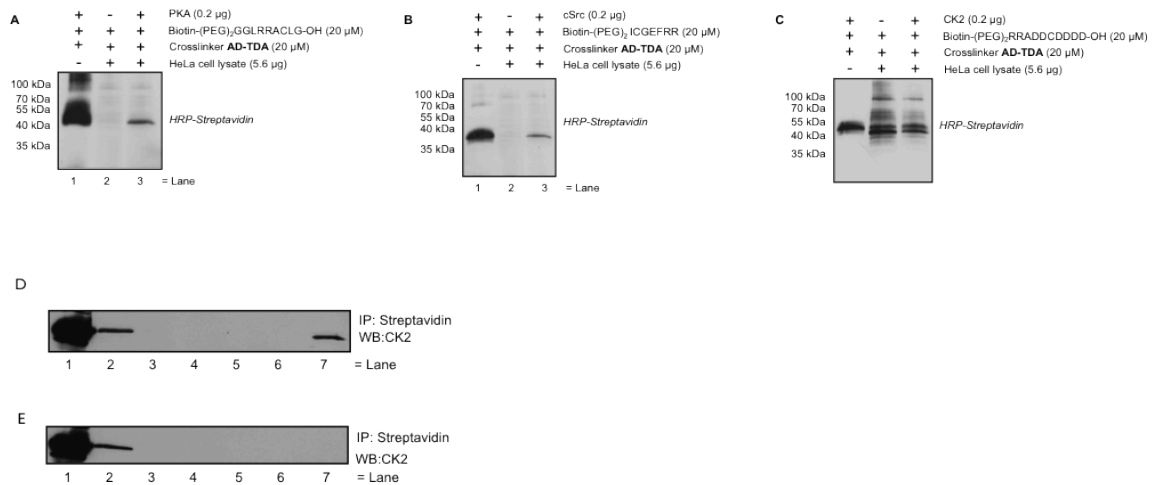**Figure 2.3**. CKII pulldowns. Lysates were diluted up to 2 mL. Reacted crosslinker and

tev-CysCKIItide for 2 hours (lane 1). Ran on a PD10 column (lane 2). Incubated with

Avidin beads overnight and saved flow though (lane 3). Washed 6 times (lane 4). Eluted

with tev protease overnight (lane 5) and the beads were boiled in SDS loading buffer for

5 mins (lane 6). Silver stain of elution fractions are shown below.
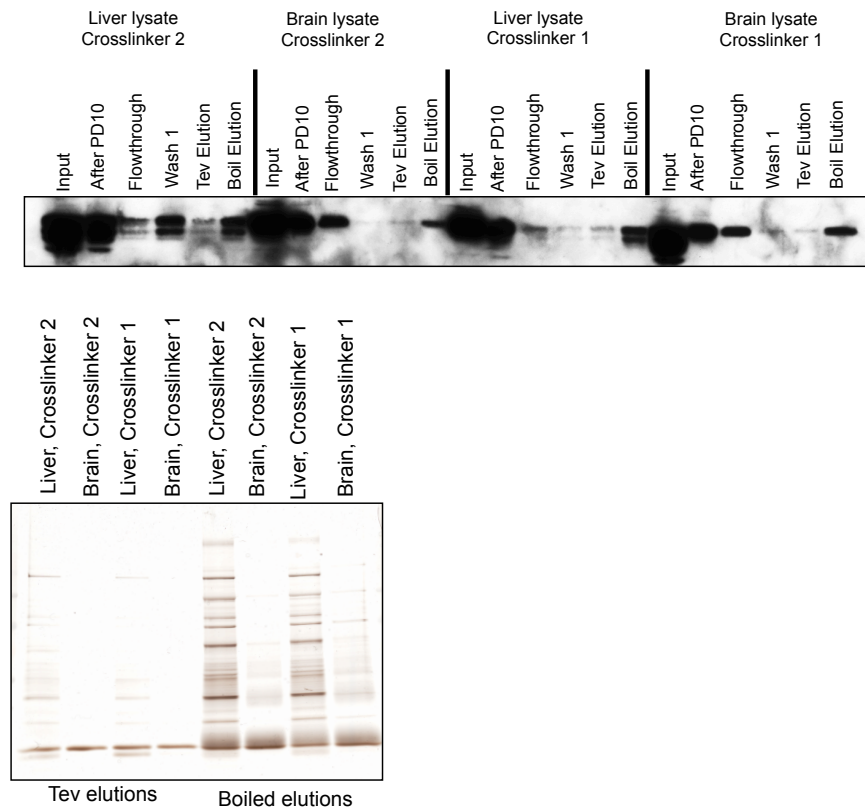
(**Figure 2.2D**), but not in the control experiment (**Figure 2.2E**). However, all attempts to pull down CKII and identify it by mass spectroscopy were unsuccessful (**Figure 2.3**).

With the goal of increasing the yield of the reaction we wanted to determine what other factors would effect the chemo-selectivity. We knew from previous experiments that a denaturing amount of detergent inhibited the crosslinking reaction, but we asked what the effect of a small amount of detergent would have on the reaction. With no other competitors present, PKA and Akt will both crosslink to a PKA substrate peptide (**Figure 2.3**). With the addition of 0.01% Triton-X we see an overall reduction in activity, and a complete ablation of the mismatched crosslinking product between Akt and CysPKAtide (**Figure 2.4B**). We see a similar effect with increasing concentration of cell lysate; in the absence of cell lysate CysPKAtide will react with either PKA or tyrosine kinase c-Src, but this mismatched product does not occur in the presence of other cellular proteins (**Figure 2.4C**). Taken together, the results indicate that by creating a more competitive binding environment, either by shielding the protein with detergent, or by having other proteins present, the requirements for a successful crosslink to occur become more stringent and the overall reaction is more dependent on the kinase-substrate binding affinity.

We postulated that our current crosslinker **1** may be occupying the pocket with too great avidity, and thus allowing the kinase to react with any substrate it came into contact with. Since the specificity of the reaction relies on the kinase substrate binding affinity, we then focused our attention on the crosslinker **1** itself. We envisioned that two main factors of the crosslinker would affect the efficiency of the kinase substrate crosslinking

Figure 2.4. Effect of detergent and cell lysate on chemoselectivity. (A) Western blot

showing the effect of detergent on the crosslinking reaction (B) Increasing cell lysate

concentration decreases the over all yield of the reaction, but does so in a chemoselective
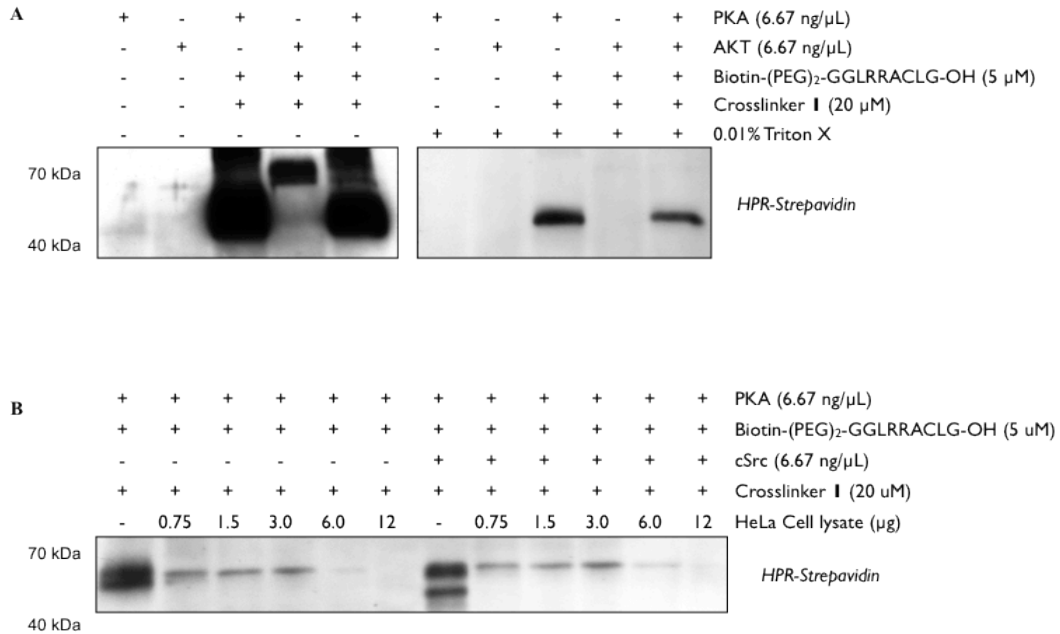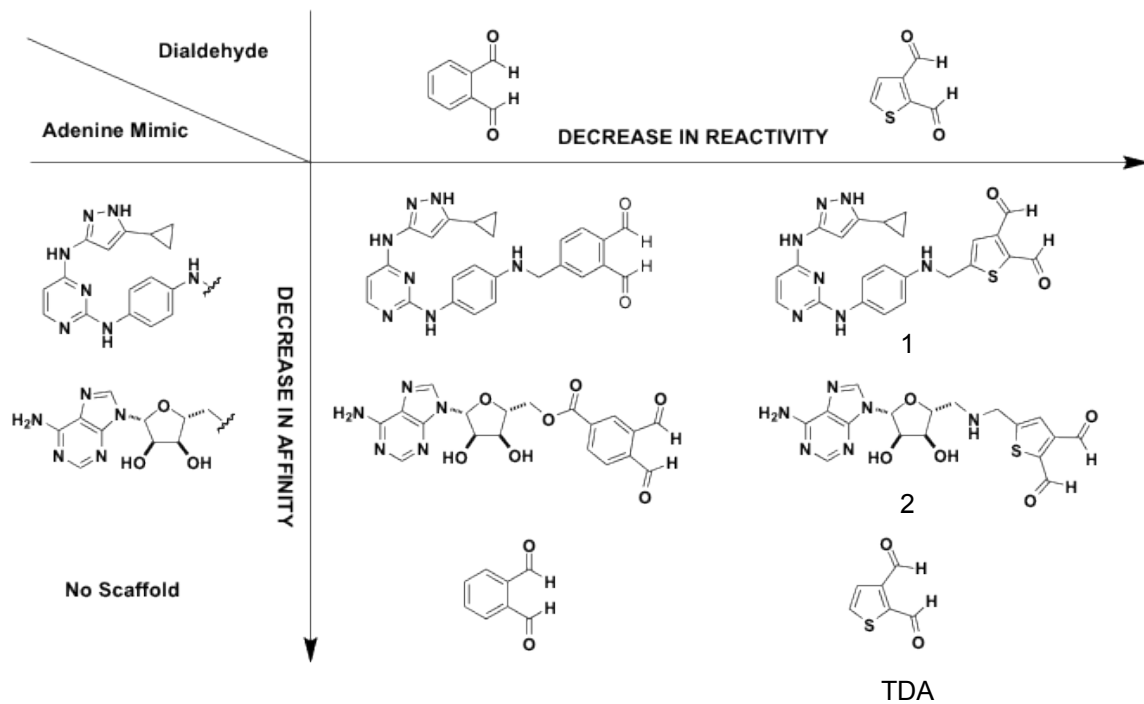
manner.

Figure 2.5. Synthesized crosslinkers organized by reactivity and affinity for the kinase ATP pocket.

reaction: the affinity of crosslinker **1** to the kinase active site and the reactivity of

thiophene-2,3-dialdehyde group itself (**Figure 2.5**. We prepared a thiophene-2,3-

dialdehyde based crosslinker **2**, where high affinity promiscuous pyrazolopyrimidine

kinase inhibitor scaffold was replaced with the adenosine directing group, thinking that

by reducing the affinity of the crosslinking molecule the specificity of the reaction may

increase.

   Treatment of CKII and Cys-CKIItide with crosslinkers **1** or **2** produced

corresponding cross-linked kinase bands but crosslinker **1** produced significantly more

intensely labeled CKII band (both in the absence or presence of HeLa cell lysate)

compared to adenosine based crosslinker **2** (**Figure 2.6**). Notably, although adenosine

based crosslinker **2** is less efficient at kinase-substrate crosslinking compared to the

pyrazolopyrimidine based crosslinker **1**, it also displays higher signal to noise ratio

producing less background labeling of endogenous cellular proteins in HeLa cell lysates.

Complete removal of the ATP binding site directing group in crosslinkers **1** or **2** further

decreased the efficiency of the three-component kinase-substrate cross-linking reaction.

Thiophene-2,3-dialdehyde alone produces a barely observable cross-linking of CKII

kinase *in vitro* or in the presence of cell lysates, nor did it produce observable labeling of

endogenous cell lysate proteins. In our previous finding we observed that *o*-

phthaldialdehyde alone or *o*-phthaldialdehyde containing kinase-substrate cross-linkers

produced similar patterns of labeled protein bands in HeLa cell lysate, irrespective of the

presence or absence of the directing group in the kinase-substrate crosslinker. Thus, the

replacement of *o*-phthaldialdehyde group with the less reactive thiophene-2,3-dialdehyde

Figure 2.6. Comparison of thiophene-2,3-dialdehyde with crosslinker **1** and **2**. (A) CKII

and Cys-CKIItide were incubated with one of the three crosslinkers for 30 mins. After gel

electrophoresis, proteins were detected with HRP-Stepavidin western blot analysis. (B)

CKII and Cys-CKIItide were incubated with either crosslinker 1 or 2 in the presence and

absence of cell lysate.
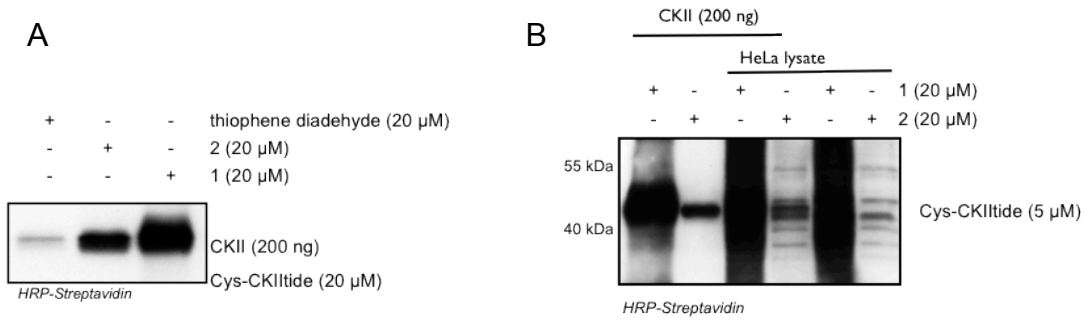
moiety made the three component kinase-substrate cross-linking reaction controllable and dependent on the kinase active site directing group.

We tested the generality of crosslinker **2** by asking if it could crosslink PKA, c-Src and CKII to their respective substrate peptides in the absence and presence of cell lysate. The bands corresponding to the crosslinked product in the absence of cell lysate were strong in all three cases (**Figure 2.7A-C, lane 1**). In the presence of cell lysate the yield of the c-Src and CysSrtide product was very reduced in comparison to crosslinker 1. In the case of CKII a band corresponding to the correct molecular weight appeared in both the conditions with and without added CKII kinase, again indicating we may be labeling endogenous kinase.

To test the specificity of crosslinker **2** we compared the ability of both crosslinkers to crosslink CysCKIItide with a panel of kinases. Other than the correct CKII CysCKIItide crosslinking product, **2** only crosslinked the peptide to two other kinases. In comparison, **1** produced a crosslinked product with all but Csk and Aurora A, demonstrating the increased specificity of the crosslinking reaction gained by decreasing the kinase affinity of the kinase binding element(**Figure 2.7D**).

In summary, we investigated factors affecting the outcome of the three-component kinase-substrate crosslinking reaction. We have shown that the three-component kinase-substrate crosslinking reaction is chemoselective and can be carried out in the presence of competing cellular proteins and protein kinases. Chemoselectivity of the three component kinase-substrate cross-linking reaction is governed by specific kinase-substrate interactions, while the affinity of thiophene-2,3-dialdehyde cross-linking element to the

Figure 2.7. (A-C) Cross-linking of PKA and cSrc kinases and their peptide substrates with the crosslinker **2** in the absence or presence of HeLa cell lysates. Cysteine containing substrate peptides, kinase and HeLa cell lysate were treated with cross-linker and incubated for 30 min at r.t. (D) Specificity of the kinase-substrate crosslinking reaction. Each kinase was incubated with CysCKIItide and crosslinker 1 (gel 1) or crosslinker 2 (gel 2).
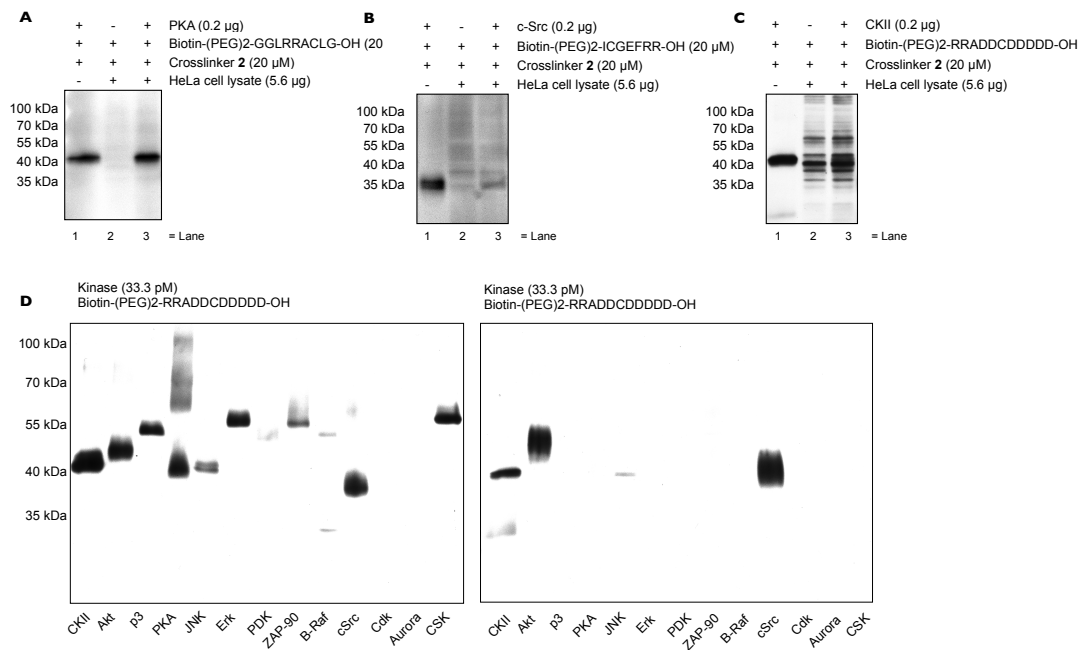
kinase active site affects the yield of the kinase-substrate crosslinking reaction. Our

findings can potentially be applied to design other *multi*-component chemical systems to

covalently trap proximal cysteine-lysine pairs, using dialdehyde chemistry.

References

(1) Khoury, G. A.; Baliban, R. C.; Floudas, C. A. *Sci. Rep.* **2011**, *1*, –.

(2) Manning, G. *Science* **2002**, *298*, 1912–1934.

(3) Bishop, A. C.; Buzko, O.; Shokat, K. M. *Trends Cell Biol.* **2001**, *11*, 167–172.

(4) Maly, D. J.; Allen, J. A.; Shokat, K. M. *J. Am. Chem. Soc.* **2004**, *126*, 9160–9161.

(5) Statsuk, A. V.; Maly, D. J.; Seeliger, M. A.; Fabian, M. A.; Biggs, W. H.; Lockhart, D. J.; Zarrinkar, P. P.; Kuriyan, J.; Shokat, K. M. *J. Am. Chem. Soc.* **2008**, *130*, 17568–17574.

General Methods

Unless noted, chemical reagents and solvents were used without further purification from commercial sources. Reaction mixtures were magnetically stirred. This layer chromatography were preformed on Merck pre-coated silica gel F-254 plates (0.25 mm). Concentration in vacuo was generally preformed using a Buchi rotary evaporator. Purification was carried out on an HPLC from Waters Co. using an xBridge prep C18 5µm 30x250 mm column. Solvent A: H_2O with 0.1% formic acid; Solvent B: ACN with 0.1% formic acid

Nuclear magnetic resonance spectra was recorded on a Varian 400 MHz instrument. Proton NMR spectra were recorded in ppm using the residual solvent signal as an internal standard: $CDCl_3$ (7.26 ppm), or d_6-DMSO (2.49 ppm). Carbon NMR were recorded in ppm relative to solvent signal: $CDCl_3$ (77.07 ppm), or d_6-DMSO (39.5 ppm).

Synthesis of compound 1. Crosslinker **1** was synthesized as previously reported[1]

Synthesis of compound 4. Amine **4** was synthesized as previously reported[2].

Synthesis of compound 5. 4,5-di(1,3-dioxolan-2-yl)thiophene-2-carbaldehyde was synthesized as previously described[1]. To a solution of adenosine amine **4** (0.140 g, 0.45 mmol) in DCM (7 mL) was added 4,5-di(1,3-dioxolan-2-yl)thiophene-2-carbaldehyde (0.098 mg, 0.38 mmol) and NaCNBH4 (0.024 mg, 0.38 mmol) and stirred overnipht at room temperature. Solution was concentrated and purified by flash chromoatgraphy (MeOH:DCM, 0:100 to 8:92). ^{1}H-NMR (400 MHz, CD$_3$OD) δ 8.21 (s, 1H), 8.11 (s, 1H), 7.94 (s, 1H) 6.22 (s, 1H), 6.10 (s, 1H), 5.91-5.88 (m, 1H), 5.42 (d, 1H, J = 8 Hz), 5.01 (d, 1H, J = 8 Hz), 4.22 (s, 1H), 4.03-3.94 (m, 4H), 3.90-3.68 (m,4H), 3.58 (s, 1H), 2.86 (d, 2H, J = 8 Hz), 1.58 (s, 3H), 1.27 (s, 3H). ESI-MS m/z [M+H]+ calculated 547.19, found 547.10.

Crosslinker 2. To a solution of **5** (0.12 g, 0.02 mmol) in water:AcCN 1:1 (1mL) was added TFA (3.2 µL, 0.04 mmol) and stirred overnight. Solution was concentrated, dissolved in water:AcCN 1:1 and purified by HPLC to yield 6.67 mg (80%) of desired 2. ^{1}H-NMR (400 MHz, DMSO) δ 10.45 (s, 1H), 10.38 (s, 1H), 9.20 (br, 2H), 8.36 (s, 1H), 8.10 (s, 1H) 7.65 (s, 1H), 6.25 (m, 1H), 5.36 (m, 1H), 4.67 (s, 2H), 4.06-3.97 (m, 4H). ESI-MS m/z [M+H]+ calculated 419.11, found 5419.02.

General protocol for crosslinking reaction. Kinases were purchased from Millipore and stored at -80˚C. Peptides were synthesized using a Liberty1 microwave peptide synthesizer. Kinase (200 ng), peptide (final concentration 5 µM), and cross-linker (final

concentration 20 μM) were added to reaction buffer (25 mM HEPES pH= 6.5, 150 mM NaCl, 2 mM MgCl2, 20 μM BME) for a total volume of 30 μL and incubated at rt. After 30 min, 6 μL of 6x loading buffer was added to quench the reactions. The sample mixtures (10 μL) were resolved by 12% SDS-PAGE, and transferred to nitrocellulose paper.

Immunoblot Analysis. The blots were blocked with 5% BSA in TBST-1 (Tris-Buffered saline, pH=7, with 0.01% Tween-20) for 5 mins at rt. The blots were next incubated with a solution of HRP-strepavidin (10000:1 dilution) for 1 hr. The blots were washed 4 times (10 min each) with TBS-T. The blots were developed with Pico SuperSignal (Piecre), and chemiluminescence was imaged on an Alpha Innotech.

HeLa Lysate. HeLa-S3 pellets (0.75x109 cells) were obtained from the National Cell Culture Center (NCCC). After thawing, 10 mL of RIPA lysis buffer (50 mM Tris pH, 150 mM NaCl, 1.0% NP-40, 0.1% SDS, with protease inhibitors (Roche, complete) was added and the cells were rotated at 4˚C for 15 minutes. After centrifugation at 20,000 g for 10 minutes the supernatant was decanted and used for kinase crosslinking experiments or stored at -80˚C.

Immunoprecipiation of crosslinked products. Brain (1 mL, 20 mg/mL) and liver (0.5 mL, 221 mg/mL) lysates were obtained from Nick Hertz. Lysates were diluted to 2 mL with reaction buffer and tev-CysCKIItide (5 μM) and crosslinker (20 μM) were added

and incubated while shaking for 1 hour. The reaction was loaded on a PD10 column, and

3.5 mL were collected. High capacity NeutrAvidin Agarose resin (62 µL slurry) was

prepared according to manufacturer instructions incubated with the resulting solution

overnight at 4˚C. After washing the beads they were incubated with AcTEV Protease

(Invitrogen) under recommended conditions. After removing the supernatant, the beads

were boiled in SDS-PAGE loading buffer for 5 mins. Both the AcTEV eluted fractions

and the boiled elution fractions were subjected to SDS-PAGE analysis.

Supplemental References

(1) Statsuk, A. V.; Maly, D. J.; Seeliger, M. A.; Fabian, M. A.; Biggs, W. H.; Lockhart, D. J.; Zarrinkar, P. P.; Kuriyan, J.; Shokat, K. M. *J. Am. Chem. Soc.* **2008**, *130*, 17568–17574.

(2) Kolb, M.; Danzin, C.; Barth, J. *Journal of Medicinal Chemistry* **1982,** *25*, 550-556

Chapter 3

Using probes to investigate imine formation between

kinases and crosslinker molecules

Abstract

In order to facilitate the elucidation of kinase signaling networks we have previously designed a chemical reaction which covalently links a kinase to its substrate via iso-indole formation. The yield of final adduct is critical to the viability of identifying kinase/substrate pairs, and in its current form, the crosslinking reaction had limited yield. This crosslinking approach was analyzed to understand the step which limited final product formation with previous crosslinkers. The analysis revealed the imine formation was almost quantitative but the subsequent nucleophilic cysteine reaction was extremely inefficient.

Introduction

The protein kinase catalyzed transfer of a phosphate from ATP to protein

substrates constitutes the major form of information transfer in biology. With 518 human

kinases and an estimated 100,000 or more phosphorylation sites the phosphoproteome

represents a complex network of enzyme-substrate relationships[1]. Identification of

kinase-substrate pairs is essential to reveal physiological functions of kinases, however,

the transient nature of the kinase substrate interaction makes this a challenging task. The

ability to trap covalent complexes of kinases and their substrates should facilitate

proteomics based approaches for unambiguously identifying the kinase responsible for a

particular phosphorylation event[2].

We previously reported a three-component chemical reaction that covalently

cross-links protein substrates of interest with their upstream kinases[3,4]. The method

design has features similar to the phosphorylation reaction, wherein the small molecule

cofactor is a promiscuous kinase binder and the specificity of the reaction is dictated by

the kinase-substrate binding affinity. Similar to ATP, our crosslinking molecules contain a

kinase binding directing group and a reactive group able to perform chemistry. The

reactive moiety on our first two generations of crosslinkers is an aromatic *ortho*-

dialdehyde, which has the ability to react with thiols and amines to form a stable iso-indol

product[5]. In our reaction the catalytic lysine on the kinase acts as the amine and the thiol

is provided by a cysteine residue in place of the phosphorylatable serine, threonine or

tyrosine on the substrate.

The originally developed *o*-phthaldialdehyde and subsequent thiophene

dialdehyde based cross-linker covalently cross-linked cysteine containing substrate peptides with their corresponding kinases. However, the reaction was too inefficient to perform successful pull-downs of the kinase. To better understand why the reaction yield was so low when the protein was one of the three components in comparison to the yield when all three components are small molecules we studied each step of the reaction separately. In this chapter, I describe the development of an LC-MS based assay to quantify the production of each intermediate of the reaction: first the imine formation between the crosslinker and the kinase, and then the attack by a thiol to form the final product. The analysis revealed an unexpectedly poor yield of nucleophillic thiol attack by the cysteine residue on the substrate, but not by a small molecule thiol. This suggested that once the crosslinker and the peptide are both bound to the kinase they create a rigid structure that is unable to adopt a productive orientation in space for the cysteine residue to attack.

Results and Discussion

We used the tyrosine kinase c-Src as a model kinase for our experiments in this report. Our lab has a robust expression system established making it ideal for the LC-MS experiments, which require a large amount of protein. c-Src with an engineered cysteine replacing the gatekeeper residue was used for all experiments (T338C) which has been previously shown have the same activity as while type c-Src and provided us with an additional chemical handle for competition experiments[6]. As a substrate mimic we used a peptide containing a consensus motif for c-Src with the tyrosine residue replaced by a cysteine residue and a biotin tag (biotin-PEG_2-I-C-G-E-F-R-R, **CysSrctide**).

To determine the why the reaction has such a low yield we turned to labeling experiments analyzed by LC-MS. To study the first step of the reaction we trapped the intermediate by incubating purified kinase with the cross-linker molecule followed by reduction of the reversible imine to a secondary amine using sodium borohydride. This has the added benefit of reducing the aldehyde on the crosslinker to quench the reaction. In these experiments we compared three different crosslinkers all with the same reactive moiety: thiophene dialdehyde (TDA) which does not contain kinase directing group, an adenosine based crosslinker **2**, and our previously reported crosslinker that has a binding group derived from a highly promiscuous nM affinity kinase inhibitor **1** (**Figure 3.1B**). We conducted a time course measuring the formation of the kinase-crosslinker complex on the LC-MS by quantifying the formation of a peak corresponding to the expected molecular weight of the kinase crosslinker product (**Table 3.1**). We found that a kinase-binding group is necessary for the formation of a detectable level of imine, and that the

Figure 3.1. Quantification of imine formation. (A) Illustration of imine formation between crosslinker 2 and the catalytic lysine residue. (B) Chart of crosslinker structures. (C) Time course of imine formation.
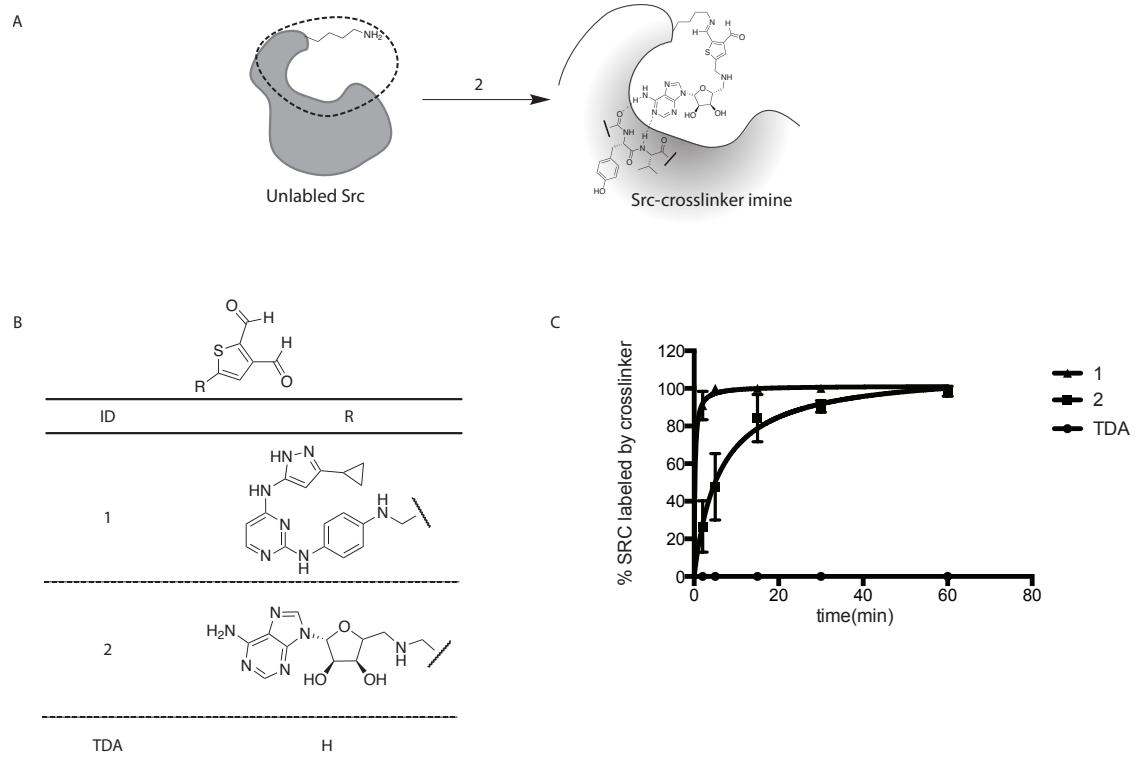
Table 3.1 Example raw data for labeling experiments. Samples were analyzed by LC-MS and the peaks corresponding to the mass of cSrc were deconvoluted and integrated using MassLynx software. Traces show a decrease in peak corresponding to unlabeled cSrc and an increase in the peak corresponding to the expected molecular weight of cSrc with crosslinker **2** over time.

MW cSrc	MW Crosslinker 2	Expected MW of cSrc + Crosslinker 2
32,698 kDa	418 g/mol	33,100 kDa

src 2uM adtda 10 uM 30mins c 43 (0.867)

1: Scan ES+
4.02e6

100

%

33096.9
65881504

33077.8
11381471

33143.9
8706000

31012.1
1121362

31388.4
3640626

32255.1
1879510

33949.0
3495940

34226.8
3273342

35158.6
1060752

35470.3
1069350

0

mass

30000   31000   32000   33000   34000   35000

src 2uM adtda 10 uM 50mins c 43 (0.867)

1: Scan ES+
5.01e6

100

%

33094.0
5012097

33114.0
2338599

33162.0
732819

30667.0
307197

31426.0
583725

32271.0
691706

34245.0
844626

34500.0
369400

35151.0
427881

35813.0
306103

0

mass

30000   31000   32000   33000   34000   35000

48

relative affinity of the binding group for the kinase dictates the rate of formation (**Figure 3.1C**).

To confirm that these compounds are reacted with the kinase in the manner we expected we needed to test three things, first that the imine is forming with the catalytic lysine residue, second that binding to the ATP pocket is necessary for imine formation, and third that the imine is reversible. We utilized two other kinase reactive probes to answer these questions. The first is a probe designed to label active kinases through their catalytic lysine, ATP-biotin (**Figure 3.2A**)[7,8]. The molecule contains biotin linked to ATP through a phospho-ester bond on the gamma phosphate. Once the probe binds to the ATP pocket the lysine attacks the ester forming an amide linkage to biotin. The second probe is AG93E, which was developed in the lab to react with the engineered gatekeeper cysteine residue (**Figure 3.2C**).

Since ATP-biotin is designed to react with the catalytic lysine, we assumed this compound would block the chemical addition of our cross-linking molecules. We first reacted the kinase with the ATP-biotin to completion, and then added our cross-linkers. ATP-biotin adequately inhibits **2**, however **1** is still able to react, either with a surface lysine or with another lysine in the ATP pocket (**Figure 3.2B**). Using AG93E we were able to ask which of these two options was more likely. If compound **1** is binding in the pocket and reacting with a different lysine residue, then occluding the pocket with another molecule should reduce labeling by the compound. On the other hand, if it is reacting with another surface lysine residue then blocking the ATP binding pocket should have no affect on labeling. We reacted the c-Src T338C with AG93E to completion

Figure 3.2. Using probes to investigate imine mechanism. (A) ATP-biotin. (B)

Competition assay between ATP-biotin and crosslinking molecules. (C) AG93E. (D)

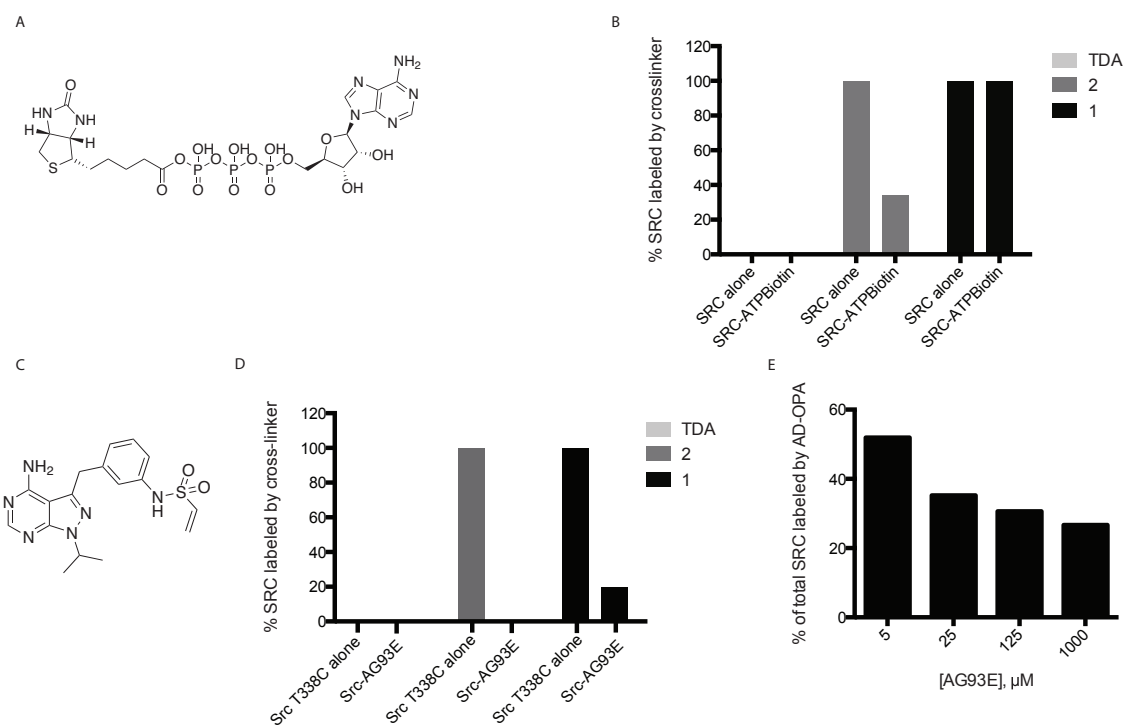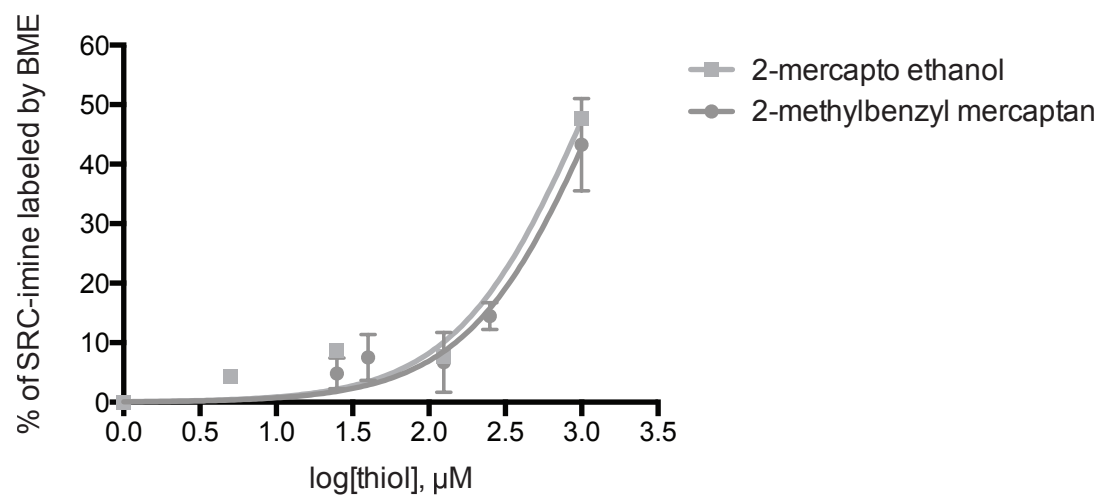Competition between AG93E and crosslinking molecules. (E) Demonstration of imine

reversibility.

**Figure 3.3**. Reaction of c-Src-imine with small molecule thiols.

followed by incubating with the crosslinkers. We found AG93E does block both the crosslinkers more effectively than ATP-biotin but still did not completely block labeling by compound 1 (**Figure 3.2D**). The reduction in labeling by compound **1** when the pocket is occupied by AG93E indicates that most of the labeling we saw when the catalytic lysine residue was rendered inactive was occurring within the ATP pocket, but due to the hydrophobic nature of the kinase directing group, crosslinker **1** may also be accessing other lysine residues on the kinase when given no other option.

We next used AG93E to test the reversibility of these compounds. We reacted the kinase first with the cross-linker to completion and then added AG93E in increasing concentrations and monitored by LCMS. After one hour the percentage of the kinase labeled by the cross-linker decreased and the percentage of the kinase labeled by AG93E increased (**Figure 3.2E**).

Since we had eliminated the imine formation as the limiting factor of the reaction, we needed to determine what aspect of the second step was unfavorable. We first asked if the kinase-imine could react with a small molecule thiol similarly to the parent reaction of OPA with free amines and thiols. We formed the kinase-imine with crosslinker 2 and then incubated this complex with different small molecule thiols. We found that this complex reacted with both 2-mercaptoethanol and 2-methylbenzyl mercaptan at a remarkably similar rate of addition (**Figure 3.3**). This shows that the crosslinking product is stable and detectable by LC-MS.

The corresponding reaction between the kinase-imine and the cysteine containing substrate peptide is not detectable by LC-MS. Like the crosslinkers, the peptide contains

Figure 3.4. Model of c-Src-imine reaction with thiols. Cysteine containing peptide binds in an orientation that makes it unable to access the imine to produce the desired product while a free thiol can react to form the product.
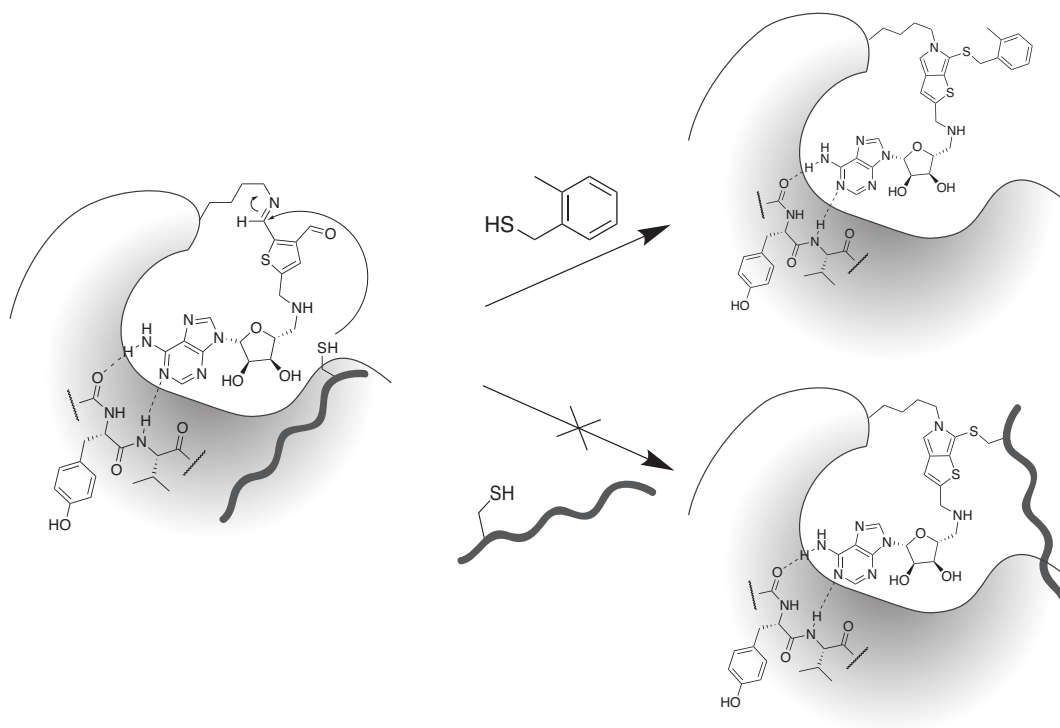
two elements, a kinase binding sequence (the amino acids flanking the cystine residue) and a reactive group (the cysteine residue). Because the small molecule thiols can react with Src-imine, the kinase binding sequence must be responsible for the lack of product. Once the peptide binds to the kinase the cysteine residue must be constrained in a position that makes it unable to react with the imine, while a free thiol dissolved in solution is able to adopt a productive orientation (**Figure 3.4**). We address this issue in Chapter 5 by designing a new type of crosslinking molecule which installs a flexible electrophile onto the kinase while the kinase binding element of the crosslinker serves as a leaving group.

References

(1) Khoury, G. A.; Baliban, R. C.; Floudas, C. A. *Sci. Rep.* **2011**, *1*, –.

(2) H S Lu, C.; Liu, K.; Tan, L. P.; Yao, S. Q. *Chemistry* **2012**, *18*, 28–39.

(3) Maly, D. J.; Allen, J. A.; Shokat, K. M. *J. Am. Chem. Soc.* **2004**, *126*, 9160–9161.

(4) Statsuk, A. V.; Maly, D. J.; Seeliger, M. A.; Fabian, M. A.; Biggs, W. H.; Lockhart, D. J.; Zarrinkar, P. P.; Kuriyan, J.; Shokat, K. M. *J. Am. Chem. Soc.* **2008**, *130*, 17568–17574.

(5) Zuman, P. *Chem. Rev.* **2004**, *104*, 3217–3238.

(6) Garske, A. L.; Peters, U.; Cortesi, A. T.; Perez, J. L.; Shokat, K. M. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 15046–15052.

(7) Patricelli, M. P.; Szardenings, A. K.; Liyanage, M.; Nomanbhoy, T. K.; Wu, M.; Weissig, H.; Aban, A.; Chun, D.; Tanner, S.; Kozarich, J. W. *Biochemistry* **2007**, *46*, 350–358.

(8) Patricelli, M. P.; Nomanbhoy, T. K.; Wu, J.; Brown, H.; Zhou, D.; Zhang, J.; Jagannathan, S.; Aban, A.; Okerberg, E.; Herring, C.; Nordin, B.; Weissig, H.; Yang, Q.; Lee, J.-D.; Gray, N. S.; Kozarich, J. W. *Chemistry & Biology* **2011**, *18*, 699–710.

General Methods

All mass spectrometric analysis was preformed using a LCT-Premier LC/ electrospray ionization-MS instrument (Waters). Protein masses were deconvoluted using the Max-Ent algorithm within the MassLynx software

General procedure for crosslinking reactions with c-Scr T338C. Kinase labeling experiments were preformed by incubating 10 µL of c-Src T338C (final concentration 2 µM) in kinase reaction buffer (25 mM HEPES pH=6.5 for thiophene based compounds, PBS pH=8 for acrylates) with 1 µL crosslinker (5 µM final concentration, final DMSO concentration 0.025%). Reaction was quenched with 3 µL NaBH$_4$ in cold MeOH (final concentration 20 µM). Reaction was analyzed by LC mass spectrometry.

Thiol reactions. The protein was labeled with a crosslinker in conditions described above for 15 minutes and no quench step was preformed. The product was incubated with 1 µL of 2-methyl benzyl mercaptan or 2-mercaptoethanol for 40 minutes. The reaction was analyzed by LC mass spectrometry.

Chapter 4

Applying kinase-substrate crosslinking technology to investigate the

interplay between phosphorylation and methylation on Histone H3

Abstract

Understanding how a kinase recognizes its substrate is a difficult problem with many complicating factors. Traditionally, this has been studied by screening peptides to find consensus sequences for a particular kinase. However, this technique ignores other important factors, such as protein-protein interactions, the posttranslational state of the substrate, and other binding proteins. One substrate for which these complicating factors are particularly important is histones. In addition to phosphorylation, histones are subject to a wide variety of posttranslational modifications that can recruit other binding partners, and how this affects kinase recognition and histone kinase activity is not well understood. One very interesting type of cross talk is the seemingly antagonistic relationship between histone phosphorylation and methylation on neighboring residues, which may serve as a general switching mechanism between chromatin states. The approaches to studying histone crosstalk have been limited by the difficulty of installing specific modifications on histones. We have developed a dialdehyde based kinase-substrate cross-linking method and applied it to histone H3 and Aurora B. This method can be combined with post-translational modification analogs to elucidate the mechanisms of interaction between phosphorylation and methylation on histone H3.

Introduction

In understanding a phosphorylation cascade, it is important to take into consideration the complex hierarchy in substrate-enzyme recognition. At the first level, the residues directly around this amino acid help determine which kinase or kinases will react with that site. These consensus sequences have been determined for many kinases, either by the identification of the in vivo substrate, or by screening peptide sequences [1]. However, even if a kinase demonstrates activity towards a peptide, it does not necessarily confirm that sequence is phosphorylated in vivo, and conversely, if a protein contains an identified phosphorylation motif of a particular kinase, it only implies that it might be phosphorylated by that kinase. Obviously, there are many more complicating factors determining the in vivo kinase specificity and activity then just the amino acid sequence around the phosphorylated site. The next level of regulation is the protein-protein interaction. For example, based on kcat/Km, p38 MAPK phosphorylates the protein substrate, MAPKAP2, 150-fold more efficiently than the peptide corresponding to the minimal motif [2].

The post-translational state of the substrate is also important; some kinases require the substrate to be pre-phosphorylated or primed by another kinase before the substrate is recognized. In addition to phosphorylation many other types of post-translational modifications or combinations of PTMs can serve to increase the diversity of the substrate pool. Additionally, these PTMs can recruit other proteins, which can possibly also affect the activity or recognition of the kinase for this substrate. There are many proteins known to be heavily modified, but this type of interaction between PTMs have

perhaps been best studied in the context of histones.

Phosphorylation of histones has been proposed to be acting in a binary switch mechanism with methylation[3]. One such pair of residues are H3Ser10ph and H3Lys9me3 (**Figure 4.1A**). When lysine 9 is trimethylated, heterochromatin protein 1 (HP1), which forms compact chromatin, binds to this mark[4], but HP1 does not bind when H3 is phosphorylated at serine 10[5]. In this way, phosphorylation acts as a switch between the bound and unbound states of HP1. There are still many questions to be answered about the mechanism of this 'methyl/phos switch', and an apparently similar mechanism may be occurring at other residues on the histone tail where a methylation mark is next to a phospho-acceptor (**Figure 4.1B**). A second set of marks occur at residues Thr3 and Lys4. Threonine 3 is phosphorylated by Haspin kinase, an atypical human kinase[6]. Interestingly, instead of the methylation mark being associated with regions of heterochromatin, as in the case of K9 and K27 methylation, K4me3 is found at active loci. Chromodomain-helicase-DNA-binding protein 1 (CHD1) has been identified as a reader protein that binds to the methylation mark at Lys4[7] (**Figure 4.1C**). So, even though these sets of marks have similar components (methylation, phosphorylation and a methylation binding protein) the outcome is different. It has been suggested that the order of the phosphorylation site and methylation site might be indicative of how these marks effect transcription.

Histone modifications have mainly been studied either in a global way, such as chromatin immunoprecipitation coupled to gene arrays (ChIP on CHIP), or on an individual basis, such as installing a modification analog. Although both types of

Figure 4.1. A. Simplified scheme of proposed methyl/phos switch at H3K9S10 and H3K27S28. HP1 (green circle) binds to methyl marks only when the substrate is not phosphorylated. Other factors can recognize phosphorylation (purple circle). B. Methyl marks on histone H3 tail next to known phospho acceptors.
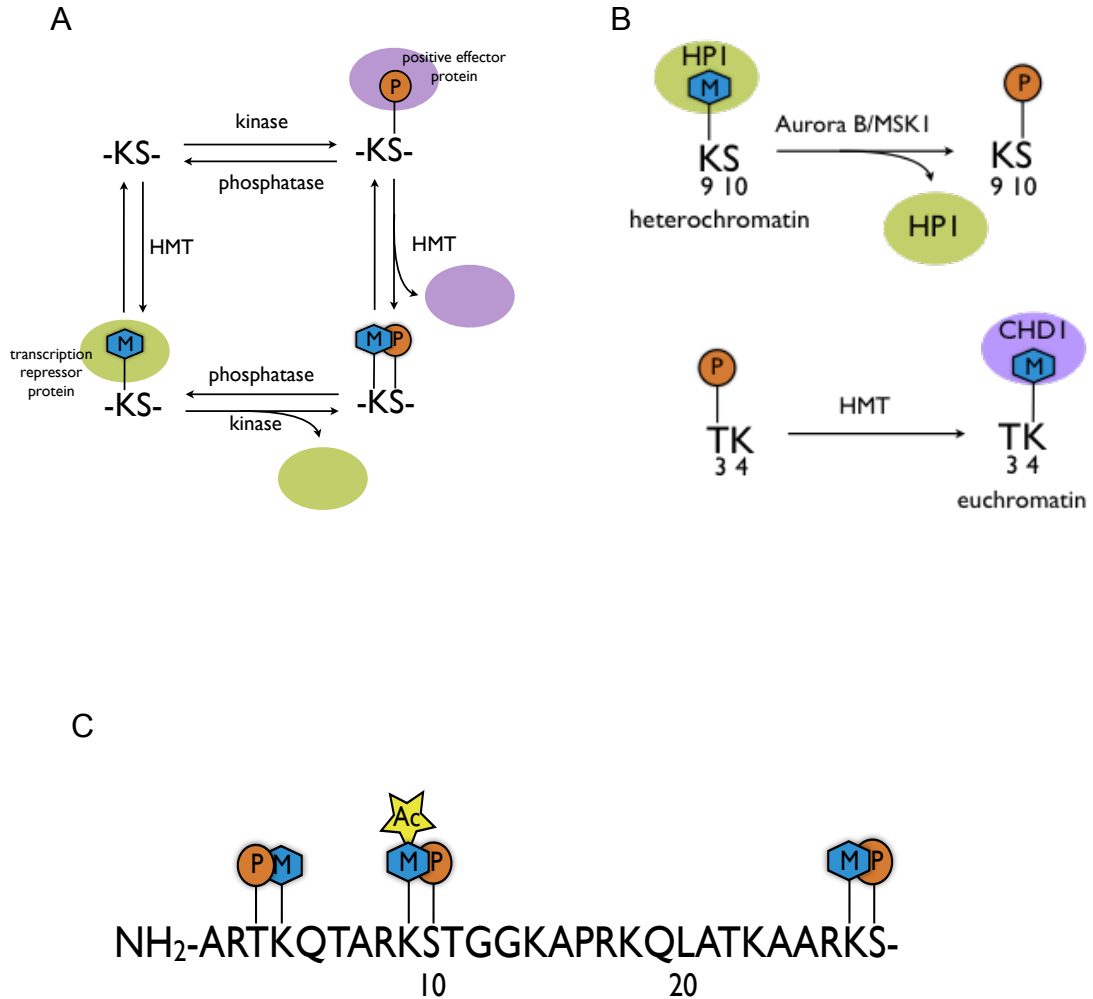
techniques have been valuable in identifying modifications, their shortcoming is in the study of the interactions between PTMs. Most early work on PTM crosstalk has been done in vitro on peptides, or through genetic alterations. One advancement in the ability to investigate these types of questions was the development of expressed protein ligation. This technology was first applied to histones by Shogren-Knaak who showed conflicting results with previous work done with H3 tail peptides, demonstrating the importance of the full proteins for biologically relevant results [8].

If these regulation mechanisms do exist, then the installation of a mark on one part of the histone tail should affect the ability of a histone modifying enzyme to install a second mark. These studies have been carried out on peptides containing different modifications, but with conflicting results, and the few reactions that have been carried out on full proteins do not always agree with the peptide results. Additionally, the recently published crystal structures of Haspin kinase reveal unusual stabilizing elements that seem to cause the kinase to be constitutively active[9,10], but it has been shown that threonine 3 is not very phosphorylated in interphase, despite the fact that the kinase is continually expressed and localized to the nucleus[6]. Therefore, it is necessary to do experiments that begin to add the complicating components to the system to better understand the interrelationships.

The in vitro data for Haspin activity suggests a reduced activity towards methylated substrates, but a recent report found that H3T3ph is always found with H3K4me3 and H3R8me2[11]. However, SILAC mass spectrometry experiments did not identify these combinatorial marks (Bonenfant et al.; 2007). Also, immunofluorescence experiments

have shown that H3K4me2, H3K4me3, and H3T3ph have distinct localizations at centromeres in mitosis[12]. If the different localization of these marks contributes to distinct zones of different chromatin function, then perhaps the interplay between K4 methylation and Haspin activity plays a roll in contributing to these zones of PTMs.

There have been kinase activity assays done on peptides with the three methylation marks which show that Aurora B has equal kinase activity towards all three marks [13], while others show an inhibitory effect of methylation[14]. All combinations of K9 methylation and S10 phosphorylation have been identified durning mitosis by mass spectrometry[15], although if the main roll of S10 phosphorylation is the dissociation of HP1, then the primary substrate of Aurora B should be trimethylated at H3K9. Once again, this points to a gap between the in vitro experiments on peptides and the in vivo observations, and analysis of these components in a more biologically relevant system will supply much needed data.

MSK1 phosphorylates S10 in interphase in response to the upstream activation of the ERK and p38 MAP kinase signaling cascades. Since this is a transient activation, perhaps it occurs in conjunction with K4 methylation. It has been shown that there is no change in K4 methylation at the hsp70 promotor before and after heat shock, but phosphorylation at S10 is increased[16]. This perhaps points to an interplay between these two switching mechanisms, allowing for more finely tuned control of the two genes. If this is true, K9 methylation might not have a large inhibitory effect on MSK1, but K4 might have a stimulatory affect.

The same study that established HP1 does not bind to the peptide with serine 10

phosphorylated showed that HP1 does not effect the activity of aurora B, even at high concentrations[13]. Maintaining the hypothesis that these two modifications are interrelated, this result does not seem to be an accurate reflection of what is occurring in the cell. Most likely, the peptide cannot recapitulate the true interaction between the two proteins. In alignment with this, it has been found that HP1 has a higher affinity for the protein then the peptide.

If the model in Figure 1 is correct HP1 should have a different effect on the ability of the kinase to phosphorylate the substrate than CHD1. By the model, the natural substrate for Aurora B is when HP1 bound, while Haspin kinase phosphorylates without the reader bound, or at sites without K4 methylation at all, so perhaps Aurora B will have a great ability of accessing the substrate in the presence of the reader protein. In peptide kinase assays, it has been shown that Haspin is less active towards H3K4me3[9], which is perhaps in agreement with the proposed model.

One of the paradoxes of H3 phosphorylation is that it seems to have different outcomes during mitosis and interphase. Stimulation of cells by growth factors induces phosphorylation of H3Ser10 by MSK1/2 at early intermediate genes, and is associated with the expression of those genes. One question here is, once the kinase is activated, how does it differentiate between histones? Perhaps the difference between the mitotic marker and the phosphorylation at transcriptionally active loci has more to do with the other PTMs then just the phosphorylation itself.

Additionally, Haspin is the only kinase identified to date that phosphorylates Thr3. Since these two phosphorylation marks (Thr3ph and Ser10ph) act similarly in mitosis, it

seems likely that Thr3 is also phosphorylated in interphase by a kinase other than Haspin. To discover if this in the case, I need a method to monitor the kinase acting at a specific site.

In order to track a single phosphorylation in the presence of all the other cellular factors a chemical reaction is needed which is easier to trace than phosphorylation while still staying as true to the endogenous reaction as possible. In the endogenous reaction all the specificity is imparted by the substrate, while ATP is a promiscuous binder. Using this reaction as a model, our system uses a promiscuous ATP pocket binder with a reactive moiety in place of the phosphate that is able to chemically crosslink the 'bait' substrate to the relevant kinase, effectively tagging the kinase for subsequent pull down and identification. All kinases have a conserved lysine near the active site, and this residue acts as a chemical handle to crosslink with the introduced cysteine residue on the substrate through a reaction involving amines and thiols. Specifically, the reaction of *ortho*-phalthaldehyde with an amine and a thiol to form an isoindole[17].

The first demonstration of this technique was done with an adenosine-based cross-linker with OPA as the reactive moiety referred to as AD-OPA[18]. Although this molecule was able to cross-link peptides to their corresponding kinases, and demonstrated some selectivity between different peptides, this molecule was not selective in the presence of cell lysate. To address this problem we reduced the reactivity of the cross-linking group by replacing OPA with a thiophenedialdehyde (TDA). To balance the decreased reactivity, the adenosine group was replaced by a higher affinity, promiscuous hinge binder[19]. Both versions of the cross-linker with OPA as the reactive group were non-

selective, indicating that OPA is so reactive the directing group has no effect on the

overall outcome of the reaction. On the other hand, both versions of the cross-linker with

TDA are able to cross-link a single kinase in the presence of cell lysate.

Results and Discussion

We used a histone H3 construct previously developed with the one naturally occurring cysteine residue mutated to an alanine[20]. Using quickchange technique and an established Histone expression protocol, we expressed and purified Histone H3 S10C[21]. We first wanted to ask if we could successfully crosslink Histone H3 S10C to Aurora kinases. We tested both Aurora A and Aurora B, crosslinker **1** and crosslinker **2**. As a control we used Histone H3 WT as the bait substrate. The reactions were analyzed by western blot using *anti*-Aurora antibodies (**Figure 4.2A-B**). We observed a higher molecular weight band only in the reactions containing Aurora B, either crosslinker, and Histone H3 S10C. This indicated to us that a crosslink had occurred in the specific manner we expected. We were surprised that the reactions containing Aurora A did not show a crosslink product as Aurora A has been reported to phosphorylate Histone H3 on Ser10 as both have been shown to phosphorylate Histone H3 in vitro. We therefore preformed an in vitro kinase assay comparing Aurora A and Aurora B in their ability to phosphorylate *Xenopus* Histone H3 (**Figure 4.2C**). With our components, Aurora B had more activity towards the histone, showing that the crosslinking reaction may be recapitulating the phosphorylation activity.

We next wanted to test if we could crosslink Histone H3 S10C to endogenous Aurora B. We incubated Histone H3 S10C with HeLa cell lysate, and added either crosslinker 1 or 2 to the reaction. The reactions were again analyzed by western blot using an antibody against Aurora B (**Figure 4.2D**). When all three components necessary for the crosslinker reaction to occur are present there is an increase in a band

Figure 4.2. Crosslinking experiments with Aurora kinases and Histone H3. (A) Western blot of crosslinking reaction between Aurora A and both crosslinkers (B) Western blot of crosslinking reaction between Aurora B and both crosslinkers. (C) In vitro kinase assay with Aurora A and Aurora B with Histone H3. (D) Western blot of reaction between crosslinkers, histone H3 S10C and HeLa cell lysate.
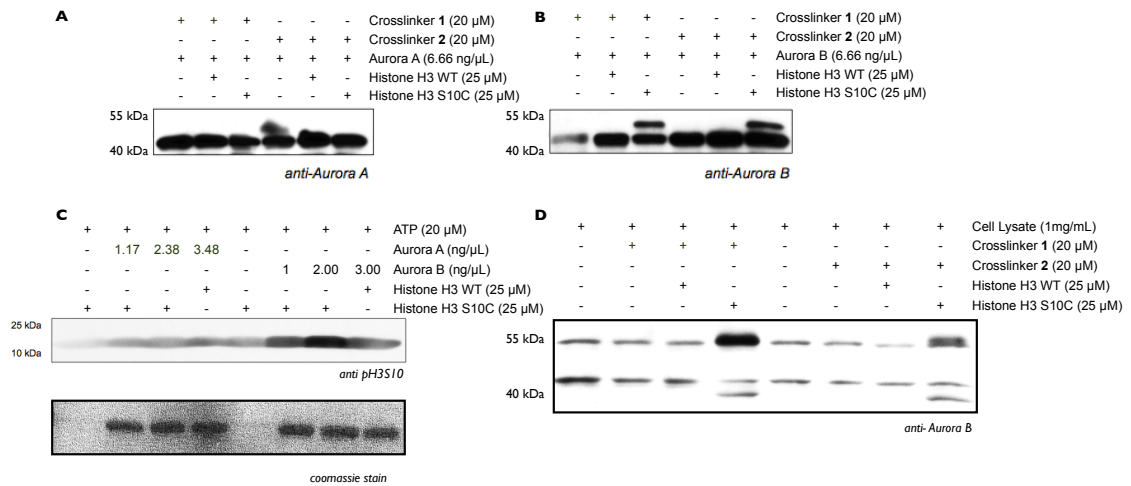
corresponding to a 15 kDa increase to the molecular weight of Aurora B. In the control

lanes (lanes 1 and 5) there is an additional band corresponding to 55 kDa in addition to

the expected molecular weight of 45 kDa for Aurora B.

With this encouraging result we expressed and purified a 6-His-tagged Histone H3

S10C to enable pull-downs of endogenous kinases (**Figure 4.3**). We reacted each

crosslinker in cell lysate with 6-His Histone H3 S10C and preformed a pulldown using

cobalt affinity resin. The samples were analyzed by western blot and mass spectrometry

(**Figure 4.4, Appendix I**). Although there is a weak band in the elution fractions of the

western blot, no kinase known to or likely to phosphorylate Histone H3 was identified by

mass spectrometry. The one positive result from this experiment is the mass spectrometric

analysis did identify histone peptides in the experiments that included a crosslinker but

not in the control experiment. The in-gel digest was preformed by cutting out the area of

the gel corresponding to a protein molecular weight 40-60 kDa, so the fact that Histone

H3 was present at a higher molecular weight region indicates that it did get crosslinked to

a larger protein, but either at not a high enough yield to identify or to a protein of little

interest. The confirms that the pulldown and in gel digest were preformed successfully,

but that the signal to noise of the reaction is still too low.

Despite the pulldown experiments, the fact that crosslinking reaction reflects

corresponding kinase activity is encouraging for using this as a tool to study

phosphorylation and methylation interplay in histones. By combining crosslinking with

expressed protein ligation it would be possible to install both a methylation analog and a

cysteine residue on a full length protein. This pseudo-substrate could be used to study the

Figure 4.3. Purification of histone H3 S10C. Coomassi stained gel of elution fractions. Fractions 7-10 were pooled. LC-MS trace of histone H3 S10C shows expected molecular weight.
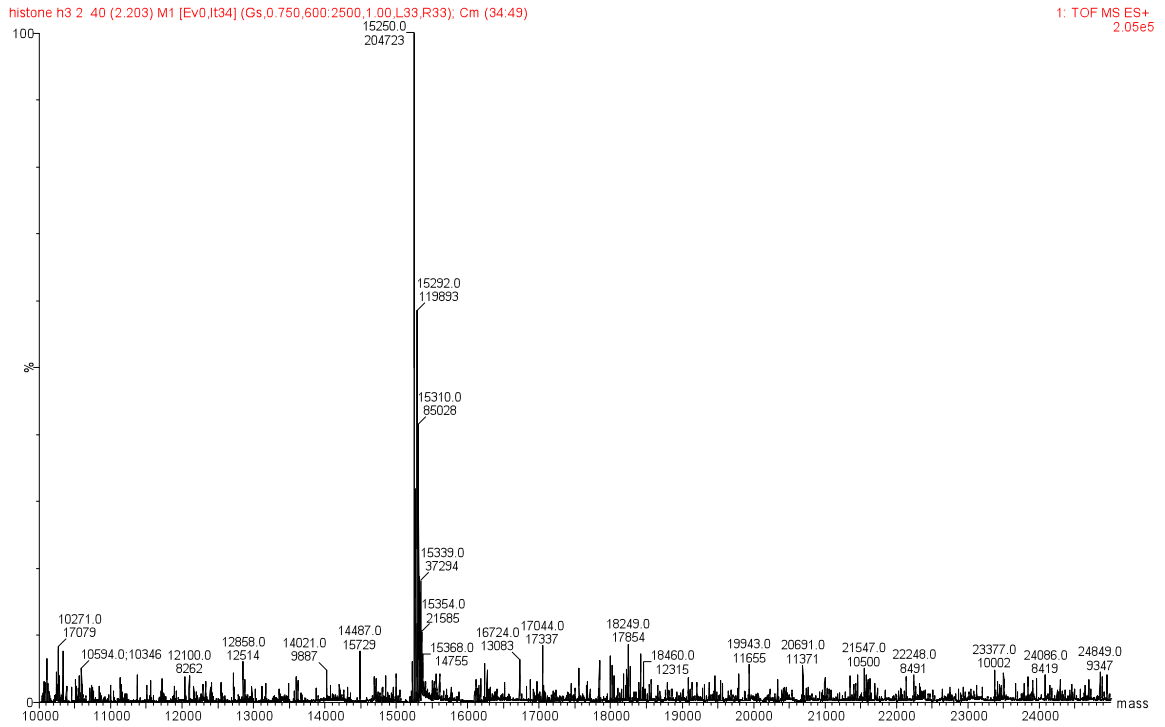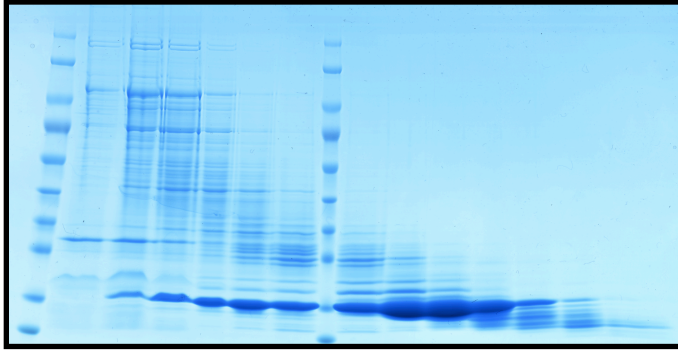
Figure 4.4. Crosslinking and purification of 6-His Histone H3 S10C. Crosslinkers and 6-His Histone H3 S10C was incubated with cell lysate for 3 hours. Reactions were washed over cobalt affinity resin and flow through was saved. 6-His Histone H3 S10C was cleaved off the beads with tev protease. The beads were then incubated with imidazole to elute remaining protein.
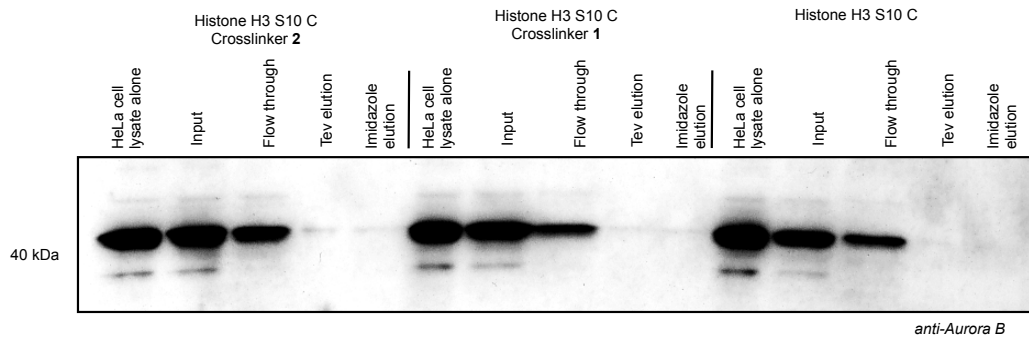
changes in binding affinity of different kinases for Histone H3 with a variety of

methylation marks, with the assumption that crosslinking yield corresponds to kinase

activity. This is a promising application of crosslinking technology beyond kinase-

substrate pair identification.

References

(1) Adams, J. A. *Chem. Rev.* **2001**, *101*, 2271–2290.

(2) Hawkins, J.; Zheng, S.; Frantz, B.; LoGrasso, P. *Arch. Biochem. Biophys.* **2000**, *382*, 310–313.

(3) Fischle, W.; Wang, Y.; Allis, C. D. *Nature* **2003**, *425*, 475–479.

(4) Jacobs, S. A.; Khorasanizadeh, S. *Science Signaling* **2002**, *2002*, tw115–tw115.

(5) Hirota, T.; Lipp, J. J.; Toh, B.-H.; Peters, J.-M. *Nature* **2005**, *438*, 1176–1180.

(6) Dai, J. *Genes & Development* **2005**, *19*, 472–488.

(7) Sims, R. J. *Journal of Biological Chemistry* **2005**, *280*, 41789–41792.

(8) Shogren-Knaak, M. A. *Journal of Biological Chemistry* **2003**, *278*, 15744–15748.

(9) Eswaran, J.; Patnaik, D.; Filippakopoulos, P.; Wang, F.; Stein, R. L.; Murray, J. W.; Higgins, J. M. G.; Knapp, S. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 20198–20203.

(10) Villa, F.; Capasso, P.; Tortorici, M.; Forneris, F.; de Marco, A.; Mattevi, A.; Musacchio, A. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 20204–20209.

(11) Markaki, Y.; Christogianni, A.; Politou, A. S.; Georgatos, S. D. *Journal of Cell Science* **2009**, *122*, 2809–2819.

(12) Sullivan, B. A.; Karpen, G. H. *Nat Struct Mol Biol* **2004**, *11*, 1076–1083.

(13) Fischle, W.; Tseng, B. S.; Dormann, H. L.; Ueberheide, B. M.; Garcia, B. A.; Shabanowitz, J.; Hunt, D. F.; Funabiki, H.; Allis, C. D. *Nature* **2005**, *438*, 1116–1122.

(14) Rea, S.; Eisenhaber, F.; O'Carroll, D.; Strahl, B. D.; Sun, Z. W.; Schmid, M.; Opravil, S.; Mechtler, K.; Ponting, C. P.; Allis, C. D.; Jenuwein, T. *Nature* **2000**, *406*, 593–599.

(15) Garcia, B. A.; Barber, C. M.; Hake, S. B.; Ptak, C.; Turner, F. B.; Busby, S. A.; Shabanowitz, J.; Moran, R. G.; Allis, C. D.; Hunt, D. F. *Biochemistry* **2005**, *44*, 13202–13213.

(16) Valls, E.; Sánchez-Molina, S.; Martínez-Balbás, M. A. *J. Biol. Chem.* **2005**, *280*, 42592–42600.

(17) Zuman, P. *Chem. Rev.* **2004**, *104*, 3217–3238.

(18) Maly, D. J.; Allen, J. A.; Shokat, K. M. *J. Am. Chem. Soc.* **2004**, *126*, 9160–9161.

(19) Statsuk, A. V.; Maly, D. J.; Seeliger, M. A.; Fabian, M. A.; Biggs, W. H.; Lockhart, D. J.; Zarrinkar, P. P.; Kuriyan, J.; Shokat, K. M. *J. Am. Chem. Soc.* **2008**, *130*, 17568–17574.

(20) Simon, M. D.; Chu, F.; Racki, L. R.; la Cruz, de, C. C.; Burlingame, A. L.; Panning, B.; Narlikar, G. J.; Shokat, K. M. *Cell* **2007**, *128*, 1003–1012.

(21) Luger, K.; Rechsteiner, T. J.; Richmond, T. J. *Meth. Enzymol.* **1999**, *304*, 3–19.

General Methods

Unless noted, chemical reagents and solvents were used without further purification from commercial sources. Reaction mixtures were magnetically stirred. This layer chromatography were preformed on Merck pre-coated silica gel F-254 plates (0.25 mm). Concentration in vacuo was generally preformed using a Buchi rotary evaporator.

Nuclear magnetic resonance spectra was recorded on a Varian 400 MHz instrument. Proton NMR spectra were recorded in ppm using the residual solvent signal as an internal standard: CDCl$_3$ (7.26 ppm), or d_6-DMSO (2.49 ppm). Carbon NMR were recorded in ppm relative to solvent signal: CDCl$_3$ (77.07 ppm), or d_6-DMSO (39.5 ppm).

Synthesis of compound 1. Crosslinker **1** was synthesized as previously reported [1].

Synthesis of compound 4. Amine **4** was synthesized as previously reported [2].

Synthesis of compound 5. 4,5-di(1,3-dioxolan-2-yl)thiophene-2-carbaldehyde was synthesized as previously described [1]. To a solution of adenosine amine **4** (0.140 g, 0.45 mmol) in DCM (7 mL) was added 4,5-di(1,3-dioxolan-2-yl)thiophene-2-carbaldehyde (0.098 mg, 0.38 mmol) and NaCNBH4 (0.024 mg, 0.38 mmol) and stirred overnight at room temperature. Solution was concentrated and purified by flash chromoatgraphy (MeOH:DCM, 0:100 to 8:92). ^{1}H-NMR (400 MHz, CD$_3$OD) δ 8.21 (s, 1H), 8.11 (s, 1H), 7.94 (s, 1H) 6.22 (s, 1H), 6.10 (s, 1H), 5.91-5.88 (m, 1H), 5.42 (d, 1H, J = 8 Hz), 5.01 (d, 1H, J = 8 Hz), 4.22 (s, 1H), 4.03-3.94 (m, 4H), 3.90-3.68 (m,4H), 3.58 (s, 1H), 2.86 (d,

2H, J = 8 Hz), 1.58 (s, 3H), 1.27 (s, 3H). ESI-MS m/z [M+H]+ calculated 547.19, found

547.10.

Crosslinker 2. To a solution of **5** (0.12 g, 0.02 mmol) in water:AcCN 1:1 (1mL) was

added TFA (3.2 μL, 0.04 mmol) and stirred overnight. Solution was concentrated,

dissolved in water:AcCN 1:1 and purified by HPLC to yield 6.67 mg (80%) of desired 2.

^{1}H-NMR (400 MHz, DMSO) δ 10.45 (s, 1H), 10.38 (s, 1H), 9.20 (br, 2H), 8.36 (s, 1H),

8.10 (s, 1H) 7.65 (s, 1H), 6.25 (m, 1H), 5.36 (m, 1H), 4.67 (s, 2H), 4.06-3.97 (m, 4H).

ESI-MS m/z [M+H]+ calculated 419.11, found 5419.02.

Histone expression. *Xenopus* histones were expressed in E. coli and purified by gel

filtration as previously described[3]. All histone H3 constructs contain a C110A mutation,

including those labeled as wild-type here. Ser-to-Cys mutants constructed using

Quikchange mutagenesis (Stratagene) according to manufacturer's instructions.

HeLa Lysate. HeLa-S3 pellets (0.75x109 cells) were obtained from the National Cell

Culture Center (NCCC). After thawing, 10 mL of RIPA lysis buffer (50 mM Tris pH, 150

mM NaCl, 1.0% NP-40, 0.1% SDS, with protease inhibitors (Roche, complete) was

added and the cells were rotated at 4°C for 15 minutes. After centrifugation at 20,000 g

for 10 minutes the supernatant was decanted and used for kinase crosslinking

experiments or stored at -80°C.

General protocol for crosslinking reaction. Aurora A and Aurora B were purchased

from Millipore and stored at -80˚C. Kinase (200 ng), histone (final concentration 5 μM),

and cross-linker (final concentration 20 μM) were added to reaction buffer (25 mM

HEPES pH= 6.5, 150 mM NaCl, 2 mM MgCl2, 20 μM BME) for a total volume of 30

μL and incubated at rt. After 30 min, 6 μL of 6x loading buffer was added to quench the

reactions. The sample mixtures (10 μL) were resolved by 12% SDS-PAGE, and

transferred to nitrocellulose paper.

Kinase Assay. Aurora B and Aurora A were purchased from Millipore. Kinases were

diluted with 20mM MOPS/NaOH pH=7.0, 1 mM EDTD, 0.01% Brij-35, 5% glycerol,

0.1% 2-mercaptoethanol, 1 mg/mL BSA. To a reaction tube with 5 x Reaction Buffer (40

mM MOPS/NaOH pH=7.0, 1mM EDTA), 5 μL histone (750 μM stock), 2.5 μ

(3.2-8.7ng) of the kinase), and 2.5 μL H_2O was added 10 μL diluted ATP mixture (25 mM

AgAc and 0.25 mM ATP). The mixture was incubated for 10 minuets at 30˚C. Reaction

was quenched with SDS loading buffer and proteins were resolved by SDS-PAGE and

transferred to nitrocellulose paper.

Immunoblot Analysis. The blots were blocked with 5% BSA in TBST-1 (Tris-Buffered

saline, pH=7, with 0.01% Tween-20) for 5 mins at rt. The blots were next incubated with

a solution of HRP-strepavidin (10000:1 dilution) for 1 hr. The blots were washed 4 times

(10 min each) with TBS-T. The blots were developed with Pico SuperSignal (Piecre), and

chemiluminescence was imaged on an Alpha Innotech.

Immunoprecipiation of crosslinked products. To HeLa cell lysate were were added 6His-tagged histone H3 S10C (5 µM) and crosslinker (20 µM) and incubated while shaking for 3 hours. Talon affinity resin (16.5 µL slurry) was prepared according to manufacturer instructions incubated with the resulting solution overnight at 4˚C. After washing the beads they were incubated with AcTEV Protease (Invitrogen) under recommended conditions. Proteins were resolved by SDS-PAGE and subjected to in-gel digest and MS analysis.

References

(1) Statsuk, A. V.; Maly, D. J.; Seeliger, M. A.; Fabian, M. A.; Biggs, W. H.; Lockhart, D. J.; Zarrinkar, P. P.; Kuriyan, J.; Shokat, K. M. *J. Am. Chem. Soc.* **2008**, *130*, 17568–17574.

(2) Kolb, M.; Danzin, C.; Barth, J. *Journal of Medicinal Chemistry* **1982**, *25*, 550-556.

(3) Luger, K.; Rechsteiner, T. J.; Richmond, T. J. *Meth. Enzymol.* **1999**, *304*, 3–19.

Chapter 5

Development of new ATP based kinase-substrate crosslinkers

Abstract

In order to facilitate the elucidation of kinase signaling networks we have previously designed a chemical reaction which covalently links a kinase to its substrate The yield of final adduct is critical to the viability of identifying kinase/substrate pairs. We report a newly designed the crosslinker using a more flexible prosthetic acrylamide arm delivered to the catalytic lysine of the kinase to facilitate more efficient crosslink formation. This new kinase substrate crosslinker has significant potential for studying kinase substrate pairs.

Introduction

The protein kinase catalyzed transfer of a phosphate from ATP to protein substrates constitutes a major form of information transfer in biology. With 518 human kinases and an estimated 20,000 or more phosphorylation sites, the phosphoproteome represents a complex network of enzyme-substrate relationships. Identification of individual kinase-substrate pairs is essential to reveal the specific physiological functions of kinases, however, the transient nature of these kinase substrate interactions makes this a challenging task. Chemical genetic techniques exist to determine the substrates of a known kinase, but no general method has yet been developed to identify the kinase for a known phosphorylation site. A method to covalently trap complexes of kinases and their substrates should facilitate proteomics based approaches for unambiguously identifying the kinase responsible for a particular phosphorylation event.

To this end we have developed a three-component chemical reaction capable of covalently linking an engineered "bait" substrate peptide to its kinase. This bio-orthogonal chemical reaction is designed to mimic aspects of the native phosphorylation reaction. All active kinases use the same phosphate donor, ATP, and a very conserved set of residues to perform the phosphor-transfer reaction chemistry. What allows for specificity from this generic chemical template is the very tightly regulated kinase substrate binding specificity. We wanted to maintain this simultaneous generality and specificity of protein phosphorylation in our method design so the orthogonal chemical reaction would only occur between the bait substrate peptide of interest and its upstream kinase without biasing the set of kinases we could detect. To do this we designed

crosslinkers that are promiscuous kinase binders and contain a reactive group able to covalently react with the engineered substrate. The substrate is modified to contain a cysteine residue in place of the phosphorylatable serine, threonine or tyrosine, generating a reactive chemical handle within the native peptide sequence and preserving substrate recognition by its target kinase(s).

For our newly designed crosslinker molecule reported herein, we took inspiration from a kinase activity-based probe ActivX, which uses an ATP scaffold to deliver a biotin tag to the conserved catalytic lysine of the kinase. We postulated that a similar molecule could instead deliver an electrophile. To test this we replaced biotin on ActivX with an acrylate moiety. The new crosslinker, **1** reacts with the conserved lysine residue on the kinase via a substitution reaction in which the kinase binding element is released from the kinase. The installed acrylamide arm is then available for nucleophilic attack by the cysteine residue on the substrate.

In this chapter we compare our new crosslinking molecule to our previously reported crosslinker quantitatively using an LC-MS based assay. We find that the two versions of crosslinkers are each able to covalently label the kinase via the catalytic lysine. However, when we incubate the kinase-crosslinker complex with the cysteine containing peptide substrate the new crosslinker improves the yield of the desired three component product 40-fold.

Figure 5.1. Scheme of crosslinking reaction

Unlabled Src

1

ADP
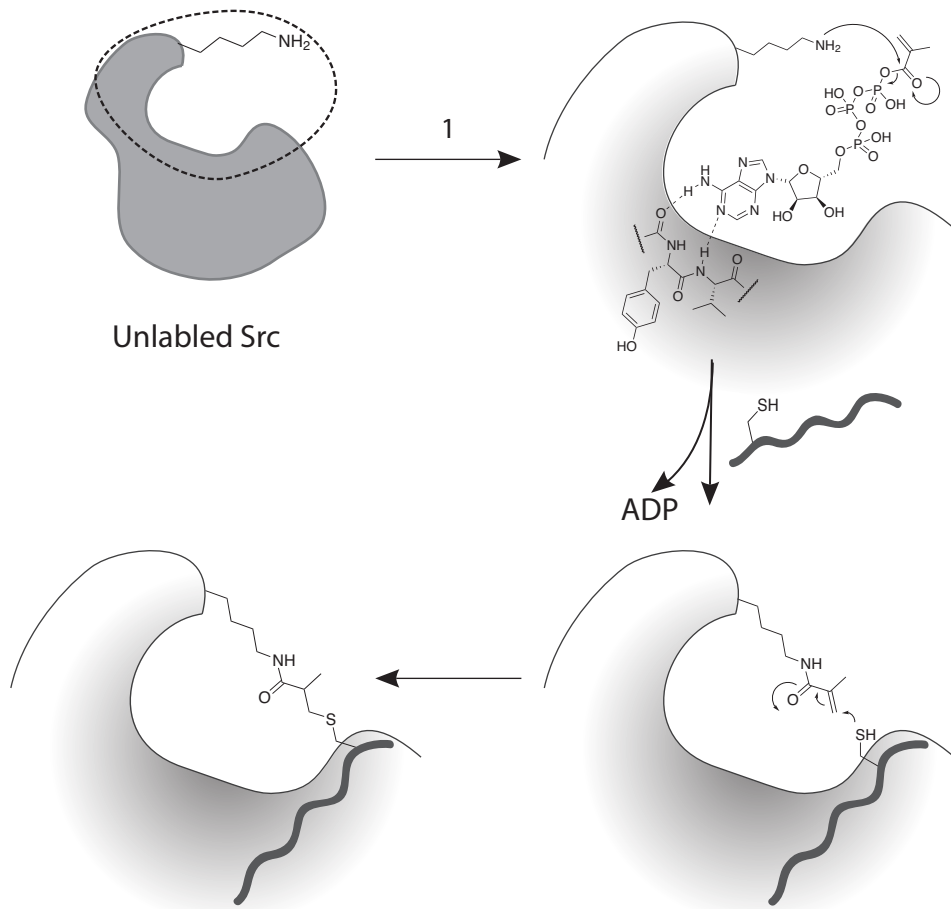
Results and Discussion

To compare the reaction progress of the different three-component reaction, we turned to labeling experiments analyzed by LC-MS. We compared **1** to our previously reported crosslinker **2**, which has a binding group derived from the highly promiscuous nM affinity kinase inhibitor ASC24 (**Figure 2A**). We used the tyrosine kinase c-Src as the model kinase for crosslinking experiments. The availability of a robust expression system enabled use of mass spectrometry to measure covalent adducts. We replaced the gatekeeper residue of c-Src with an engineered cysteine (T338C) which has been previously shown have the same activity as wild type c-Src and provided us with an additional chemical handle for competition experiments (supplemental). As a substrate mimic we used a peptide containing a consensus motif for c-Src with the tyrosine residue replaced by a cysteine residue and a biotin tag (biotin-PEG$_2$-I-C-G-E-F-R-R, **CysSrctide**). In this system, the cysteine is replacing a tyrosine residue instead of serine residue as in our previous model system. This peptide reacts with c-Src in our previously reported crosslinking reaction as measured by western blot (supplemental).

We conducted a time course measuring the formation of the kinase-crosslinker complex. Crosslinker **1** reacted with the catalytic lysine to form Src-acrylamide at a slower rate than crosslinker **2** reacted to form Src-imine (Figure 1B). This is consistent with our model that the binding affinity of the kinase binding element increases the effective concentration of the reactive group in proximity to the lysine.

We next wanted to ask the how each kinase-crosslinker complex reacted with thiols. We tested each one with two different thiol containing molecules, 2-methylbenzyl

mercaptan and CysSrctide. The comparison between these two molecules allows us to probe what effects the binding affinity the peptide for the kinase has on the overall reaction. We treated T338C Src-imine with 2-methyl benzyl mercaptan, with increasing concentration of the thiol, the formation of the three component product increased, showing that not only can the thiol access the imine, but that the fused aromatic complex is stable and detectable by the LC-MS. We performed the same experiment with T338C Src-acrylamide as the starting material, and found the thiol reacted less with this complex, converting at maximum 20% of the protein in solution to the final three component product.

We next treated the T338C Src-imine and Src-acrylamide with CysSrctide in increasing concentrations. We expected the binding affinity of peptide for the kinase would increase the yield of the reaction compared to 2-methyl benzyl mercaptan, however with the T388C Src imine, even at 600 µM, the formation of a kinase-**2**-peptide complex was undetectable. In contrast, the reaction between T338C Src-acrylamide and CysSrctide rapidly led to product and in excess over the reaction with 2-methylbenzyl mercaptan. This validates our original plan to use the kinase-substrate interaction as the specificity determinate; due to the higher affinity of the peptide for the kinase than a free thiol, the peptide product is more favorable. This is encouraging result, as it might indicate that the reaction will be selective for correct kinase-substrate pairings.

We hypothesize that the failure to observe the three component product using crosslinker **2** is due to geometric constraints imposed by the position of the crosslinker in the kinase active site that is unable to adopt a productive orientation for attack by the

Figure 5.2. Addition of crosslinking molecule to the kinase. (A) Structures of crosslinking molecules aligned with structure of ActivX. Kinase binding elements are shown on the right, and reactive group on the left. (B) Time course of kinase labeling by crosslinker **2** and crosslinker **1**.
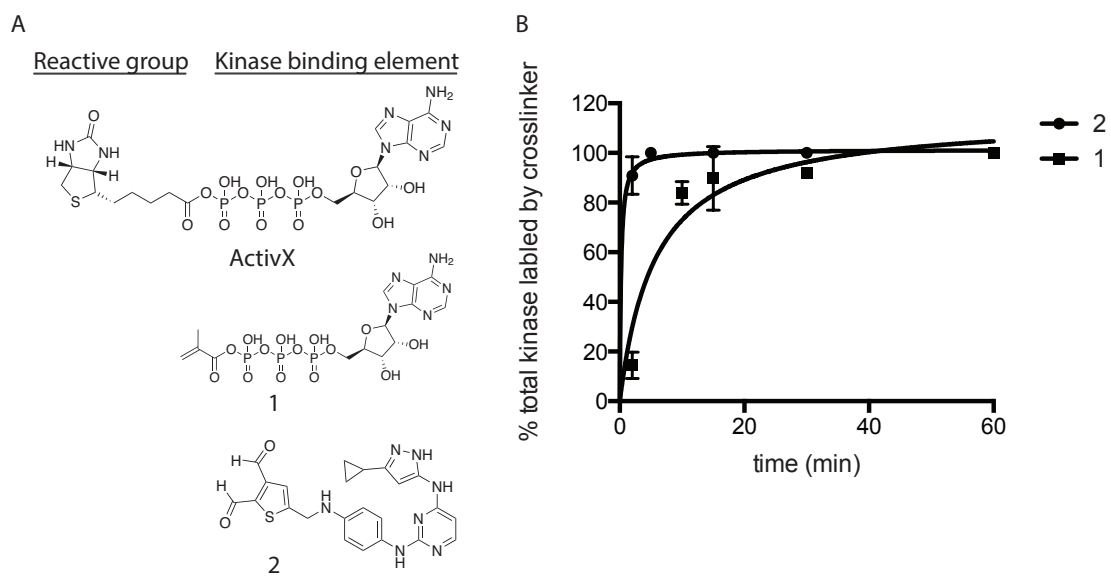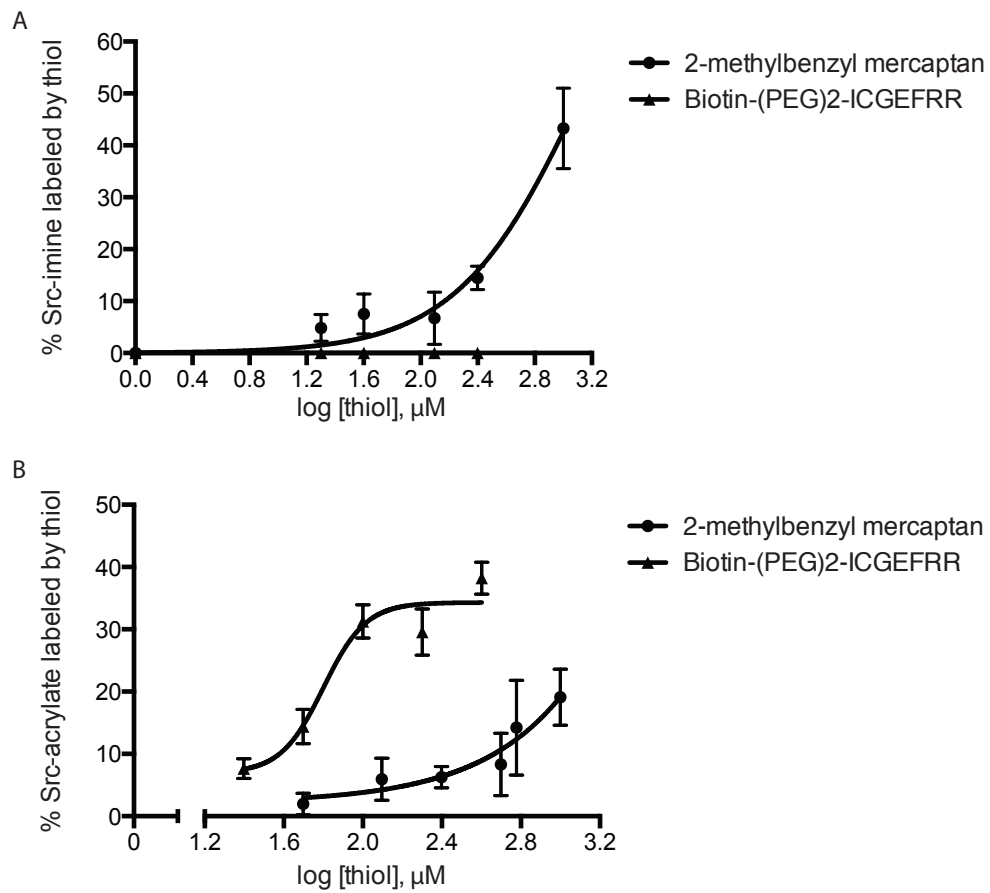
Figure 5.3. Reactivity of kinase-crosslinker with thiol molecules. (A) Src-imine formed by a reaction with c-Src and thiophene dialdehyde based crosslinker was reacted with 2-methylbenzyl mercaptan or peptide. (B) Src-acrylate formed by a reaction with c-Src and crosslinker **1** was reacted with 2-methylbenzyl mercaptan or peptide.

peptide thiol. By altering the design of the crosslinker to allow the kinase binding element to be released after the first covalent step, we engineered more flexibility into the reaction and this improved the yield of our desired product. These are promising tools for trapping and investigating substrate kinase interactions and with further experimentation and optimization may be useful for kinase identification of a known phosphorylation event.

Our most recent molecule needs further testing and optimization, but this represents the latest advancement in the progress towards the goal of a selective, robust kinase-substrate crosslinking method. These tools can be employed in the discovery of unknown kinase-substrate pairs, the investigation of kinase affinity in the context of other PTM markers on the substrate, and to trap complexes for crystallization. All of which will further our understanding of kinase signaling networks.

General Methods

Unless noted, chemical reagents and solvents were used without further purification from commercial sources. Reaction mixtures were magnetically stirred. This layer chromatography were preformed on Merck pre-coated silica gel F-254 plates (0.25 mm). Concentration in vacuo was generally preformed using a Buchi rotary evaporator.

Nuclear magnetic resonance spectra was recorded on a Varian 400 MHz instrument. Proton NMR spectra were recorded in ppm using the residual solvent signal as an internal standard: CDCl$_3$ (7.26 ppm), or d_6-DMSO (2.49 ppm). Carbon NMR were recorded in ppm relative to solvent signal: CDCl$_3$ (77.07 ppm), or d_6-DMSO (39.5 ppm).

Preparation of ATP-triethylammonium salt. Dowex 50W resin (4.0g) was incubated in saturated triethylamine (50 mL) for 1 hour. The liquid was removed and the beads were washed with H20 (pH=9.0) (2 x 5 mL). ATP (4.0 g) dissolved in H20 (5 mL) was poured over beads and incubated for 30 mins. The flow through was removed and saved. ATP-TEA was eluted with H20 (2 x 5 mL). The flow through and elution fractions were frozen and lyophilized.

Synthesis of 1. To a solution of methyl acrylic anhydride (74 µL, 0.50 mmol) in 3 mL of a dioxane/DMF/DMSO mixture (1:1:1) was added trethylamine (50 µL) and a solution of ATP triethylammonium salt (0.40 mg, 0.49 mmol) in anhydrous DMSO (3 mL) and stirred overnight at room temperature. The reaction was quenched with water (4 mL) and the solution was quickly extracted with ethyl acetate (3 x 4 mL). The aqueous layer was

immediately frozen and lyophilized. The yellow solid was suspended in water and transferred to a pre-equilibrated lichrosorb RP-18 column (Agilent), and eluted with water. Fractions containing product were frozen and lyophilized to yield 0.224 g (78%) of a white powder.

Synthesis of other ATP-acrylic phosphates. To a solution of an acrylic acid (0.49 mmol) in dry DMF:Dioxane:DMSO 1:1:1 (3 mL) was added triethylamine (278 µL, 2 mmol), and anhydrous isobutyl chloroformate (194 µL, 1.5 mmol) and stirred at 0˚C for 20 min. The reaction was removed from ice bath and stirred at room temperature for 1 hour and 55 min. A solution of ATP triethylammonium salt (0.400 mg, 0.49 mmol) in DMSO (3 mL) was added. Reacted overnight at room temperature, and workout was the same as for molecule **1**.

General procedure for crosslinking reactions with c-Scr T338C. Kinase labeling experiments were preformed by incubating 10 µL of c-Src T338C (final concentration 2 µM) in kinase reaction buffer (25 mM HEPES pH=6.5 for thiophene based compounds, PBS pH=8 for acrylates) with 1 µL crosslinker (5 µM final concentration, final DMSO concentration 0.025%). Reaction was quenched with 3 µL NaBH$_4$ in cold MeOH (final concentration 20 µM). Reaction was analyzed by LC mass spectrometry.

Thiol reactions. The protein was labeled with a crosslinker in conditions described above for 15 minutes and no quench step was preformed. The product was incubated with

1 μL of 2-methyl benzyl mercaptan or peptide for 40 minutes. The reaction was analyzed by LC mass spectrometry.

Appendix I

Proteins identified in the Presence of crosslinkers and Histone H3 S10C

Control: 6His-Histone H3 S10C, no crosslinker

Rank	Acc #	Uniq pep	% Cov	Best Disc Score	Protein MW	Species	Protein Name
[1]	P63038	2	9.1	2.23	60956	MOUSE	60 kDa heat shock protein, mitochondrial
[2]	P35579	1	0.9	3.75	226534.2	HUMAN	Myosin-9
[3]	Q8C196	2	2.9	1.75	164619.5	MOUSE	Carbamoyl-phosphate synthase [ammonia], mitochondrial
[4]	P08113	1	2.4	2.78	92476.5	MOUSE	Endoplasmin
[5]	P07724	1	2.1	1.2	68693.2	MOUSE	Serum albumin
[6]	P21579	1	5.2	1.05	47573.5	HUMAN	Synaptotagmin-1
[7]	P02768	1	3.1	1.04	69367.4	HUMAN	Serum albumin
[8]	Q6MZU6	1	7.1	0.92	51099.3	HUMAN	Putative uncharacterized protein DKFZp686C15213
[9]	A2BDB0	1	6.1	0.77	41793.2	XENLA	Actin, cytoplasmic 2
[10]	O35490	1	4.7	0.65	45020.9	MOUSE	Betaine--homocysteine S-methyltransferase 1

6His Histone H3 S10C + crosslinker **1**

Rank	Acc #	Uniq pep	% Cov	Best Disc Score	Protein MW	Species	Protein Name
[1]	P63038	2	9.1	3.18	60956	MOUSE	60 kDa heat shock protein, mitochondrial
[2]	Q6MZU6	1	7.1	2.06	51099.3	HUMAN	Putative uncharacterized protein DKFZp686C15213
[3]	P02768	1	3.1	1.41	69367.4	HUMAN	Serum albumin
[4]	O08599	1	2.5	1.35	67569.3	MOUSE	Syntaxin-binding protein 1
[5]	O35490	1	4.7	1.31	45020.9	MOUSE	Betaine--homocysteine S-methyltransferase 1
[6]	P62799	1	7.8	0.44	11367.4	XENLA	Histone H4

6His Histone H3 S10C + crosslinker **2**
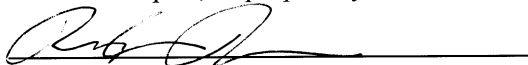
Rank	Acc #	Uniq pep	% Cov	Best Disc Score	Protein MW	Species	Protein Name
[1]	P63038	2	9.1	1.76	60956	MOUSE	60 kDa heat shock protein, mitochondrial
[2]	P08113	1	2.4	1.53	92476.5	MOUSE	Endoplasmin
[3]	P02768	1	3.1	1.3	69367.4	HUMAN	Serum albumin
[4]	O60814	1	11.9	0.84	13890.3	HUMAN	Histone H2B type 1-K
[5]	P62799	3	27.2	0.36	11367.4	XENLA	Histone H4

Publishing Agreement

It is the policy of the University to encourage the distribution of all theses, dissertations, and manuscripts. Copies of all UCSF theses, dissertations, and manuscripts will be routed to the library via the Graduate Division. The library will make all theses, dissertations, and manuscripts accessible to the public and will preserve these to the best of their abilities, in perpetuity.

I hereby grant permission to the Graduate Division of the University of California, San Francisco to release copies of my thesis, dissertation, or manuscript to the Campus Library to provide access and preservation, in whole or in part, in perpetuity.

_____   6/11/13
Author Signature Date