DETECTION OF POLYPEPTIDE INTERACTIONS VIA PERIODATE TRIGGERED DOPA CROSSLINKING

APPROVED BY SUPERVISORY COMMITTEE

Name	**Signature**
Thomas Kodadek Ph.D.	
George DeMartino Ph.D.	
Phil Thomas Ph.D.	
Masahide Kikkawa M.D., Ph.D.	

Acknowledgments

I would like to utilize this space to individually recognize and thank the people who played instrumental roles during my doctoral work.

Tom Kodadek (Mentor): Over the course of my doctoral work Tom has provided what I consider to be the ideal environment in which to conduct research. Namely, he has provided a wealth of financial resources, intellectual ideas, enthusiasm for my project, and intellectual freedom. Amazingly, he has done all of this without even a hint of " result nagging", a daily habit of many advisors. Rather, he has provided nothing but constant support and encouragement. I have learned more over the last five years than in any other point in my life. Tom was, in short, a fantastic mentor.

Chase Archer: Chase, a fellow graduate student in the lab, is a first class biochemist and scientific researcher. Chase was instrumental in the work mapping the interaction between the Gal4 activation domain to the 26S proteasome and has been a fantastic collaborator throughout the cross-linking project. I am extremely grateful for his help over the last several years.

Bo Liu: Bo is an instructor in the Kodadek lab whose efforts have enabled me to carry out significant portions of my doctoral work. Namely, he was responsible for cracking two scientific challenges that were rate limiting within my cross-linking project. First, he mastered a difficult synthetic protocol and taught me how to synthesize the FLAsH-DOPA molecule. Second, he taught me how to synthesize

peptide nucleic acids. Finally, he “debugged” the PNA assay in Chapter 2 that enabled the publication of an additional manuscript. Many Thanks.

Thomas Gillette: Tom is an assistant professor in the DeMartino lab at UTSW whose insight into many of the scientific problems I encountered and creative approaches to solve those problems have been extremely valuable. Tom collaborated with me on projects at the beginning and end of my doctoral work. Thanks for the help and support.

Melissa O’Neal: Melissa had the unfortunate opportunity to have a bench next to Chase and I over the last couple of years and endured a daily stream of bad jokes and raids on her candy drawer. I blame you for my graduate school weight gain. Seriously, Melissa is an excellent molecular biologist who has been extremely helpful when my project ventured into yeast genetics and molecular biology in general.

Reddy Moola: Thanks for helping me to get started in the Kodadek lab.

George DeMartino, Phil Thomas, Masahide Kikkawa: I wanted to thank these men for agreeing to be on my committee and performing all the requirements necessary to fulfill that agreement.

Marie Schluterman: Thanks babe.

DETECTION OF POLYPEPTIDE INTERACTIONS VIA PERIODATE TRIGGERED DOPA CROSSLINKING

by

LYLE JACKSON BURDINE

DISSERTATION

Presented to the Faculty of the Graduate School of Biomedical Sciences

The University of Texas Southwestern Medical Center at Dallas

In Partial Fulfillment of the Requirements

For the Degree of

DOCTOR OF PHILOSOPHY

The University of Texas Southwestern Medical Center at Dallas

Dallas, Texas

July, 2006

DETECTION OF POLYPEPTIDE INTERACTIONS VIA PERIODATE TRIGGERED DOPA CROSSLINKING

Publication No._____

Lyle Jackson Burdine B.A.

The University of Texas Southwestern Medical Center at Dallas, 2006

Supervising Professor: Thomas Kodadek Ph.D.

Protein-Protein interactions mediate most biological function, yet elucidation of the molecular architecture within a cell still remains a formidable challenge for molecular biologist. We have developed a novel, bioorthogonal cross-linking chemistry based upon periodate mediated oxidation of the artificial amino acid 3,4-Dihydroxyphenylalanine to a resultant ortho-quinone. This ortho-quinone was proven capable of capturing either cysteine, lysine, histidine, or a peptidyl α-amine in templated chemical reactions. After elucidating the chemistry, we describe utilization

of this methodology to map peptide-protein interactions between the 26S proteasome and activation domains as well as the Arp 2/3 complex and the CA peptide. Finally, we present the creation of a chimeric molecule consisting of the biarsencial fluorescent reporter FLAsH conjugated to 3,4-Dihydroxyphenylalanine as a route to deliver 3,4-dihydroxyphenylalanine site specifically to a protein of interest and probe for protein-protein interactions partners in cell lysates.

TABLE OF CONTENTS

Chapter 1- Protein-Protein Interaction Detection Methodology

Chapter 2-Development of DOPA Cross-linking Chemistry

Chapter 4-Application of DOPA Chemistry Towards the Detection of Protein-Protein Interactions

Chapter 5-Discussion

PRIOR PUBLICATIONS

Burdine L. & Kodadek T. Target Identification in Chemical Genetics: The (Often) Missing Link. Chemistry and Biology Vol. 11, May 2004, pgs 1-5

Burdine L. Lin H., Gilette T., Kodadek T. Periodate-Triggered Cross-Linking of DOPA-Containing Peptide-Protein Complexes. JACS 2004 Sep 22;126(37)

Kreishman-Deitrick M, Goley ED, Burdine L, Denison C, Egile C, Li R, Murali N, Kodadek TJ, Welch MD, Rosen MK. NMR analyses of the activation of the Arp2/3 complex by neuronal Wiskott-Aldrich syndrome protein. Biochemistry. 2005 Nov 22;44(46):15247-56.

Archer CT*, Burdine L*, Kodadek T. Identification of Transactivation Domain-Binding Proteins In the 26S Proteasome By Periodate-Triggered Cross-linking Molecular BioSystems, 2005, 1, 366 - 372

Burdine L, Liu B, Gilette T, Archer CT., Kodadek T. An Efficient Label Transfer Reaction For the Detection of Protein-Protein Interactions That Does Not Require Covalent Protein Modification (In Revision)

Liu, B, Burdine L., Kodadek T. The Mechanism of Periodate-Mediated Cross-linking of Dihydroxy-ortho-phenylalanine (DOPA)-Containing Molecules to Proteins (Submitted to the Journal of the American Chemical Society)

LIST OF FIGURES

LIST OF TABLES

LIST OF DEFINITIONS

ORF-Open Reading Frame

DOPA-3,4-Dihydroxyphenylalanine

Interactome – Sum total of cellular molecular interactions.

Chemical Cross-linking – Chemical ligation of protein binding partners.

AMU-Atomic Mass Unit

MALDI-TOF – Matirx Assisted Laser Desorption Ionization-Time of Flight

GAL4 AD – Activation domain of the Gal4 transcription factor in yeast

Gal80 – Binds the Gal4 AD and functions as a repressor protein of the galactose system

26S – 2.5 MegaDalton complex involved in protein degradation

FLAsH- biarsenical fluorescent dye that associates with a CCPGCC motif in proteins

FRP- FLAsH Receptor Peptide, i.e. CCPGCC

NA-HRP- NeutrAvidin – Horse Radish Peroxidase

APIS-Base of the 19S proteasome that functions during transcription

PNA-Peptide Nucleic Acid

BP-Binding Peptide

Chapter One

Detection of Protein-Protein Interactions

Introduction to Protein-Protein Interactions

Genomic sequencing coupled with transcriptome analysis provides an inventory of proteins within a cell. The function of most proteins in the inventory, however, remains unknown. Fortunately, a starting point in the assignment of protein function is found in the realization that an individual protein's function is often carried out in the context of a multi-protein complex[4].

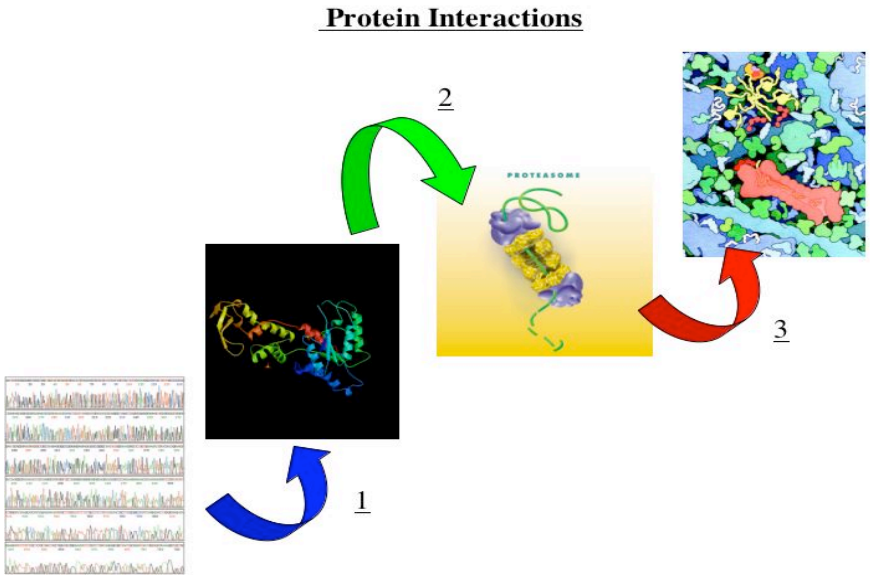


Figure 1.1 Protein Interactions. Blue Arrow-Sequenced genomes provide a cellular parts list. Green Arrow-Protein parts often assemble into tight molecular complexes such as the proteasome. Red Arrow-Large molecular complexes such as the proteasome (drawn in red) transiently interact with a variety of proteins inside of cells.

For example, estimates predict between 16,000 to 26,000 interactions amongst the approximately 6000 proteins of yeast[5, 6]. Furthermore, approximately 25% of all proteins function in more than a single complex[7]. Given this molecular connectivity, most labs working in biology at the molecular level exert a considerable amount of effort looking for interaction partners for their particular protein of interest.

Methods to Detect Protein-Protein Interactions

Several techniques exist to detect protein-protein interactions. In **Figure 1.2** they have been graphed according to the degree of throughput versus structural information obtained with the technique.

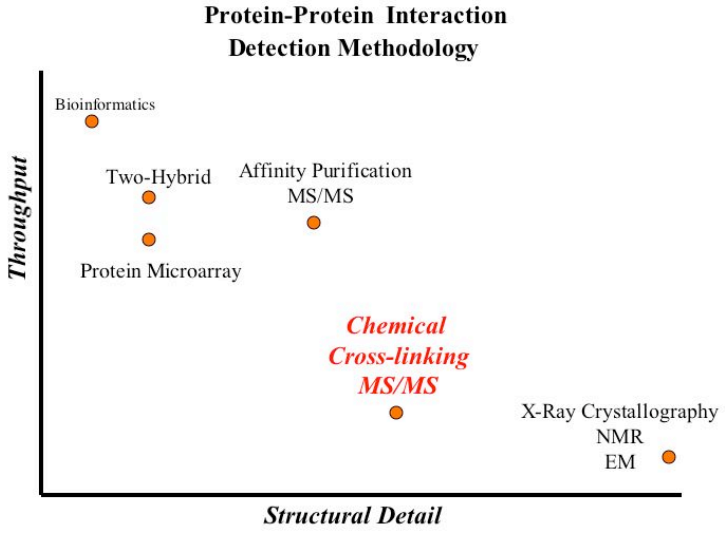


Figure 1.2 Graph detailing Throughput versus Structural Detail of various Protein-Protein Interaction Methodologies.

On the left side of the graph in **Figure 1.2** are those techniques that can and have been, for the most part automated, thus allowing thousands of potential protein interaction partners to be screened within relatively short periods of time. These methodologies often provide a list of potential interaction candidates that may

interact with a protein of interest. Further validation, however, is always necessary when utilizing these techniques. On the right side of the figure reside the techniques that are capable of providing data detailing the architecture of macro-molecular complexes and interaction interfaces. This type of information is invaluable in the design of small molecule therapeutics tailored to "fit" into interaction interfaces and therebye disrupt protein-protein interactions. Each technique possesses technical strengths and weaknesses in the detection of protein-protein interactions with the most widely utilized detailed in subsequent sections. Many biochemical research groups employ all the listed techniques to some degree in their research endeavors.

Computational Inference of Protein-Protein Interactions

The vast amount of genomic sequence information being obtained from an ever-increasing number of organisms allows for two types of sequence analysis that can posit protein-protein interaction candidates. First, many protein structural features such as conserved structural motifs or sites of post-translational modification are apparent at the nucleotide level. For example, if a protein of interest contains a RING finger domain then one can infer that this protein most likely interacts with other partners involved the ubiquitin-proteasome pathway. Second, many interacting proteins co-evolve and it is therefore possible to predict candidate protein binding partners through comparison of phylogenetic trees[8-13]. The combination of static genome sequences and protein evolutionary history with dynamic microarray and proteomic data sets permits inferences of functional and

often physical interactions amongst proteins inside of cells. The computational inference of protein-protein interactions should continue to improve but currently this method of calculating protein interaction pairs is dependent upon experimental data born out from the remainder of the techniques described in this chapter. To date results obtained from papers in the field generally only confirm existing interaction results. If *a priori* prediction of individual protein structures becomes a viable option, then computational inference of macromolecular structures should replace more labor intensive wet lab techniques for detecting protein-protein interactions.

Yeast-Two Hybrid

Since it's invention by the Fields laboratory in 1987, the molecular biology community has made extensive use of the yeast two-hybrid methodology to uncover protein-protein interactions. Working under the assumptions that proteins can function as chimeras, that non-nuclear protein domains could be targeted to the nucleus, that activation domains could be displayed properly by interacting chimeras, and that protein-protein interactions can reconstitute a signal to begin transcription, a general approach was carried out as pictured in **Figure 1.3**[14, 15]

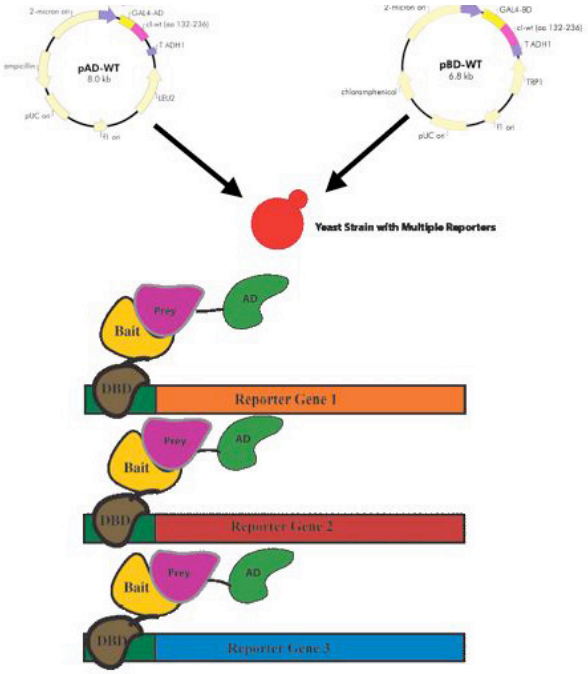


Figure 1.3 Traditional yeast-2-hybrid

The initial form the yeast-two hybrid technique involved fusing a protein of interest to the Gal4 DNA binding domain in a bait plasmid and screening against a library of potential interaction partners that are fused to the Gal4 activation domain in a prey plasmid. Sufficient interaction between the bait and prey reconstitutes the Gal4 transcription factor and initiates transcription of genes possessing a Gal4 promoter. It has been estimated that a minimum dissociation constant of 1μM is necessary to reconstitute the functional transcription factor[16]. However, it is not possible to quantify the association strength between two protein domains based solely on two-hybrid experiments. This is due to the binary nature of the reporter signal. For example, when a beta-galactosidase reporter gene is used in the two-hybrid assay, blue yeast colonies are those containing interaction pairs while white colonies consist of yeast cells not expressing domains that interact at a detectable level. Approaches have been developed to select for only strong and weak

interactions within the assay. For example, the HIS3 gene encodes imidazole glycerol phosphate dehydratase an enzyme involved in the histidine biosynthetic pathway. This enzyme can be inhibited with 3-aminotrizole. Therefore, when 3-aminotrizole is added onto the growth plate or into the culture tube only those cells that express high enough levels of HIS3 to overcome the growth inibition will produce colonies[17]. Alternatively, a reporter gene encoding URA3, Orotidine-5'-phosphate decarboxylase, can be utilized to select for cells that do not express high levels of URA3 through the addition of 5-fluoro-orotic acid to the growth culture medium. URA3 converts 5-fluoro-orotic acid into 5-Fluorouracil a poisonous nucleotide analogue[18]. As a result, only low expressing cells are able to survive with 5-fluoro-orotic acid in the media. Unquestionably, unbiased selection of only those prey proteins that interact with the bait from a large library of potential candidates is the major advantage of the technique. Although the original version of the technique suffered from a high false positive, subsequent iterations using multiple single-copy reporter genes and low copy bait and prey expression plasmids have significantly reduced the number of false positives[5]. There are of course several classes of proteins not amenable to two-hybrid analysis in the traditional form. For example, proteins that display a functional transcriptional activation domain or recruit in components of the transcription machinery activate transcription independent of the prey plasmid. Likewise, membrane proteins are not readily transported into the nucleus and are often recalcitrant to traditional yeast two-hybrid

approaches. Finally, many protein-protein interactions are dependent on post-translational modifications such as phosphorylation or glycosylation. Naturally, these contingent interactions are also not picked up with traditional two-hybrid methodology[19, 20]. Several variations of the yeast two-hybrid approach have been developed over the last decade to overcome the limitations of the traditional version. For example, Varshavsky and Johnsson developed that two-hybrid alternative that allows for the detection of interactions amongst membrane bound proteins[21] (**Figure 1.4**). Ubiquitin can be divided into an N-terminal portion, amino acids 1-35, and a C-terminal portion consisting of amino acids 35-76. If the two halves are in close proximity, they reconstitute the native ubiquitin structure which can be recognized by ubiqutin binding proteases or de-ubiquitinases. In the split-ubiquitn method a protein of interest is fused to the C-terminal portion of ubiquitin which is in turn fused in frame to a transcription factor (**Figure 1.4**). A library of potential protein interaction partners are then fused to the N-terminal half of ubiquitin. If two membrane proteins are in close proximity, the two halves reconstitute the ubiquitin structure. The reconstituted ubiquitin is then clipped N-terminal to the fused transcription factor by a ubquitin protease, there-bye freeing the transcription factor for translocation into the nucleus. The transcription factor can then activate a reporter gene signaling the protein-protein interaction that occurred in the cell membrane[21]. Fields an co-workers utilized this technique to create one of the first large scale interaction maps of the membrane proteome in yeast[22].

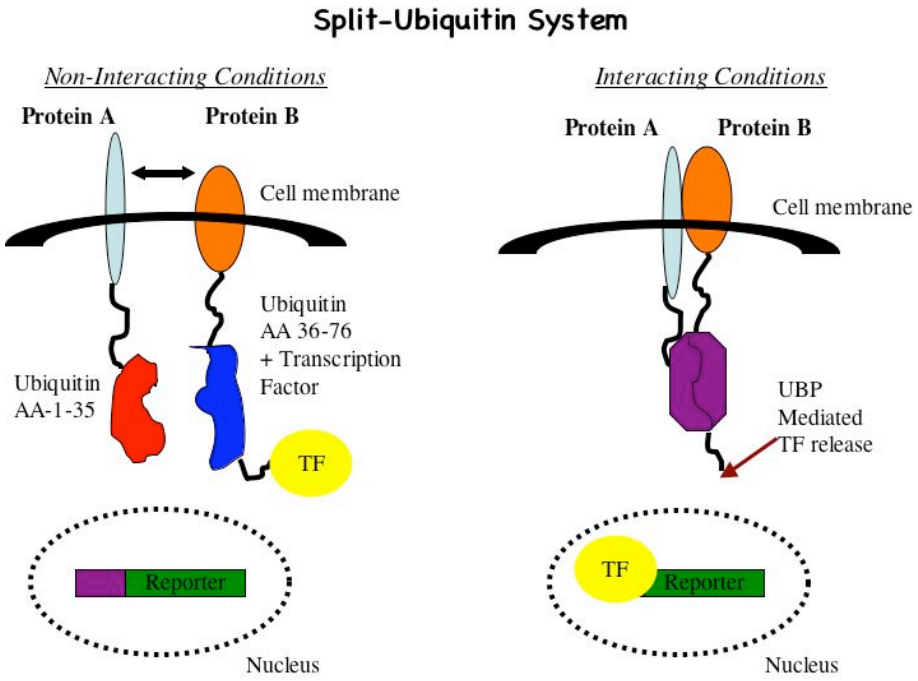


Figure 1.4: Diagram of the split ubiquitin system.

Other variations of the yeast-two hybrid have also been engineered to genetically solve specific questions concerning protein/biomolecular interactions in living cells not approachable through traditional two-hybrid methodology. For example yeast one hybrid screens are designed to search for sequence specific DNA binding domains in proteins, while three-hybrid screens have been developed to search for RNA-protein and small-molecule protein interactions. In summary, the ability to search vast amounts of genetic information for protein-biomolecular interaction information has revolutionized cell biology[23].

High-throughput yeast two-hybrid experiments attempting to probe all interactions amongst large ORF collections have begun to reveal topological characteristics of protein interaction networks not readily apparent in more limited

approaches. For example, there appears to be a tendency for highly connected "hub" proteins to interact with less connected fringe proteins "spokes". It has been hypothesized that hub proteins often constitute large stable macromolecular machines/structures that are organized in various combinations through transient interactions with fringe proteins[24, 25]. Although data from two-hybrid experiments have begun to piece together a static skeleton of binary protein-protein interactions within cells this technique misses several types of interactions most notably those occurring within the membranes of cells.

Affinity Purification/Mass Spectrometry

Affinity purification of a protein of interest is perhaps the most common method used to detect protein-protein interactions. Affinity purification involves genetically inserting an epitope tag into a protein of interest that can be bound with high affinity to a solid support containing an appropriate binding partner, e.g. an antibody. Proteins associated with the tagged protein are also "pulled out" from the biological milieu when incubated with binding beads **Figure 1.5**. Since individual protein-protein interactions within a multi-protein complex range from low to high with respect to affinity, some associations within a complex survive the washing process during the purification while other interactions are lost.

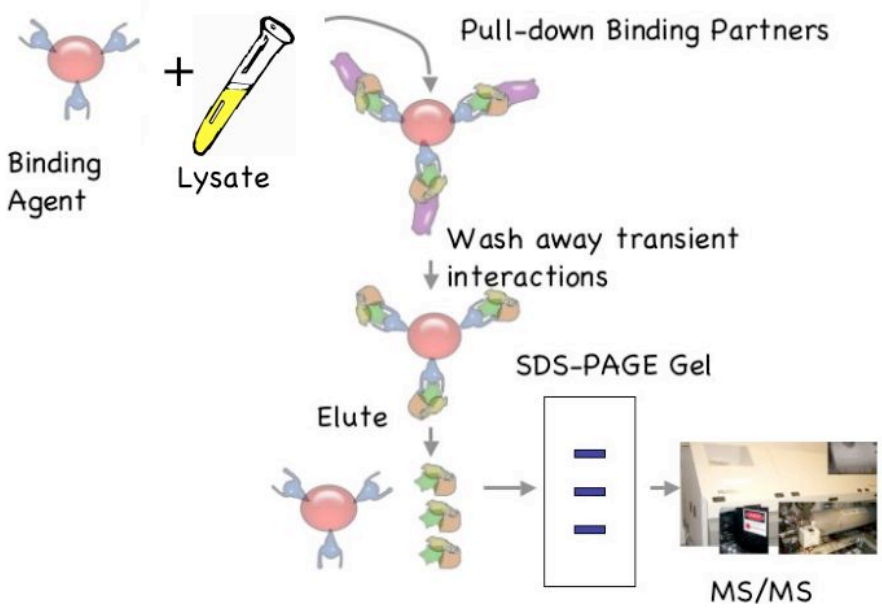


Figure 1.5 Affinity purification scheme

Two complementary technological advances have enabled a tremendous increase in the affinity pull-down utilization. First mass spectrometry has made the identification of affinity co-purified proteins far simpler than was previously the case. State of the art mass spectrometers can often detect peptides obtained from tryptic digest of proteins in gels at nanograms levels[26, 27]. Once detected, the molecular ion of an abundant peptide can then be trapped in a second round of MS and fragmented to reveal the amino acid sequence of that peptide. Databases of genomic information, a magnificent second advance, permit the fragmented amino acid sequences to be reverse translated and searched for confirmed protein coding sequences.

The most comprehensive affinity purification data set to date hypothesizes that there exist approximately 800 core protein machines in yeast that are moved, rearranged, partnered, and adapted to carry out biological activity[28]. Although this work provided only a snapshot of the ever changing molecular architecture in yeast, it does provide a useful reference for invesigators. Error rates or false prediction of interactions for such experiments have been estimated at 30%[29]. This requires that investigators trying to answer a specific biological question to re-confirm all predicted interactions derived from a high-throughput affinity purification data set. Furthermore, modified peptides and peptides with that ionize inefficiently are often missed by mass spectrometry experiments. In addition, affinity purification protocols are extremely limited when applied to integral membrane proteins due to the fact that adequate membrane solublization for subsequent analysis often disrupts transient interactions amongst proteins embedded within the membrane. There have been techniques developed to address these shortfalls. For example Tackett and Chait developed a technique coined I-DIRT to differentiate between specific and non-specific protein-protein interaction partners during an affinity purification experiment[30]. I-DIRT takes advantage of the isotopic labeling strategies pioneered by Aebersold and Mann over the last five years[31]. In the I-DIRT technique isotopically heavy and light cells are grown, lysed, and the lysates are mixed together. Importantly, only the cells grown in the light medium contain the affinity tagged protein. Upon immuno-isolation both light and heavy proteins will be pulled-

down via the affinity resin and run out onto a gel. After trypsin digestion the peptides are analyzed by MS/MS. Those peptides that contain a large ratio of light signal as compared to heavy signal are indicative of specific interactions while peptides demonstrating an even ratio of light and heavy signal strength are deemed background contaminants as showed in **Figure 1.6**. This type of methodology will become increasingly valuable when analyzing protein-protein interactions in complex biological mixtures.

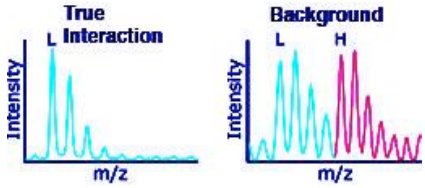


Figure 1.6. Example of a true interaction signal versus background signal in an I-DIRT experiment[30].

As productive as affinity purification has been in the past and will be in the future, affinity purification techniques are unable to detail the molecular architecture of a complex nor are "pull down" experiments able to decipher how a protein machine "moves" to carry out biological function. These are two primary reasons driving the development of additional protein-protein interaction techniques.

Chemical Cross-linking

Protein cross-linking approaches ligate reactive moieties on neighboring protein surfaces together through a variety of chemistries. This approach results in transformation of a protein-protein interaction from non-covalent to covalent. Cross-

linking cements complex molecular architecture for subsequent biochemical analysis under denaturing conditions, and permits conclusions to be drawn regarding the three dimensional structure of a complex of interest. Currently, chemical cross-linking is constrained by the types of chemistry used to ligate the reactive moieties on the surface of interacting proteins, how those chemistries are applied on a relevant biological timescale, and by the analysis of cross-linked products. Commonly utilized cross-linking strategies and the chemistries utilized to carry out those strategies are discussed in subsequent sections.

Bi-functional Cross-linking Strategies

The most common type of protein cross-linking approach consists of two reactive electrophilic groups connected with a linker arm. Variations of this general strategy are presented

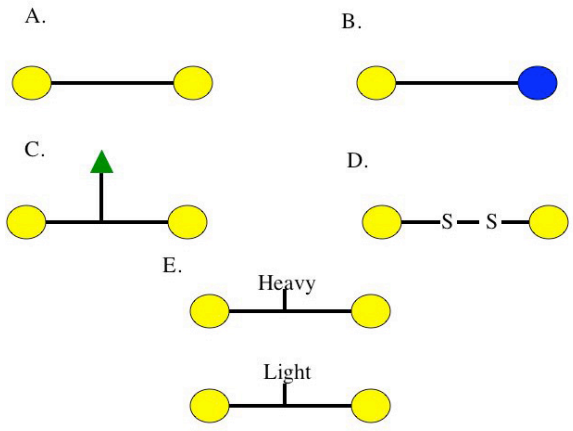


Figure 1.7. Variations of Bi-functional Cross-linkers

The goal of the bi-functional cross-linker approach is to capture nucleophilic groups that are in close proximity at the protein interaction interface. The two

electrophilic groups can be the identical (**Figure1.7A**) and capture a single type of nucleophile or different as in the case of (**Figure 1.7B**) and capture distinct classes of nucleophiles. Often a handle or tag is added to the bi-functional cross-linker to aid in the recovery or detection of the cross-linked products, (**Figure 1.7C**). For example, a biotin moiety can be inserted in the linker-arm of a bi-functional cross-linker to allow for facile isolation of cross-linked products on streptavidin beads[32]. Often times a cleavage site is incorporated into the linker arm to facilitate label transfer of the cross-linking agent from a purified protein of interest to an interacting protein partner **Figure 1.7D**)[33-36]. This approach avoids difficulties in ionization observed when analyzing cross-linked peptides by mass spectrometry. Finally, a 50/50 mixture of isotopically labeled heavy and light bi-functional chemical cross-linkers allows for identification of cross-linked products in the mass spectrometer by creating a characteristic doublet signal separated by the mass difference of the heavy and light labels[36]. In theory, computer algorithms should be able to identify cross-linked products by scanning the signal peak list for the signature doublet.

Bi-functional cross-linking approaches have yielding modest results to date in comparison with other protein-protein interaction detection methodology. The reason for the modest success can be attributed to four problems. First, for reasons still unclear, cross-linked peptides do not “fly” nearly as well as non-branched counter parts in the mass spectrometer. The upshot of this fact is a reduction in sensitivity of the technique to a point where it becomes necessary to work at artificially high or

non-biological protein concentrations. The second problem can be attributed to the fact that most cross-linkers are bio-incompatible. In general, most chemical cross-linking reactions are adaptations of chemical reactions that work well or were designed to operate in an organic or a “dry” environment with a limited and defined number of reactants. The aqueous nature of a biological environment often results in cross-linker hydrolysis while the complexity and transient nature of protein interactions often results in incorrect interaction data. This lack of appropriate biochemical cross-linking chemistry leads directly to the third and fourth problems encountered in experiments utilizing bi-funcitonal cross-linkers, namely, false positives and false negatives. A false positive result reports an interaction between two proteins that do not truly interact, while a false negative result does not report an interaction between proteins that do interact.

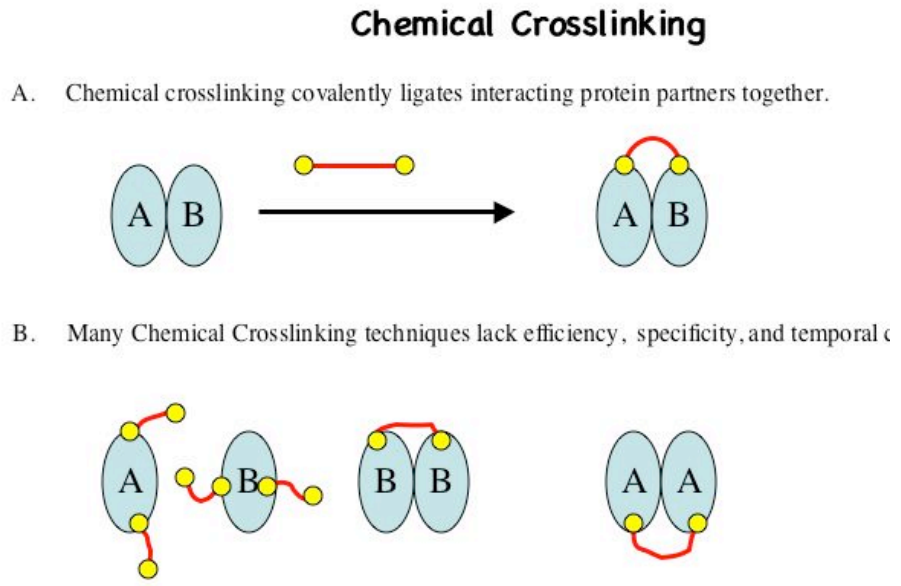


Figure 1.8: Bi-functional cross-linkers are utilized to ligate interacting proteins together. Often times however, the reactions fail due to inefficient cross-linking, non-specificity, and a lack of temporal control.

Comparison of cross-linking data reporting the molecular architecture of a multi-protein complex prior to the publication of the crystal structure has revealed extremely high error rates[37]. For example, cross-linking data obtained with bi-functional cross-linkers and detailing the architecture of several now well known protein machines such as RNA polymerase II, the Arp 2/3 complex, and the 20S proteasome were eventually shown to possess respective error rates of 41, 14, and 36%.[38-44] Obviously, this type of error limits the utility of the methodology.

Cross-linking in a true biological system requires a chemical reaction that can withstand bulk nucleophiles, water, and low concentrations of the complex of interest. Subsequent sections define the most common electrophiles employed during bi-functional cross-linking experiments.

Common Amine Reactive Chemistry

The ε-amino group of lysine is one of the most prominent nucleophiles present on a protein's surface. Given this fact, lysines are most effectively captured with bi-functional through sulfo-NHS ester groups. Sulfo-NHS esters are water-soluble reagents that will react with primary amines to form amide bonds. Reactions are typically allowed to proceed from 30 minutes to one hour. Although straight-forward in theory, sulfo-NHS ester cross-linking has been utilized with only modest

success for a variety of reasons. First, this technique requires at least micrograms of a protein complex that must be stable in phosphate buffered saline under non-reducing conditions. Second, although lysines are present on the protein surface studies suggest that the more common amino acids found at protein-protein interaction interfaces are tyrosine and serine[45, 46]. Finally, this type of chemistry is not applicable when trying to observe native interactions that exist inside the cell or inside cell lysates due to potent nucleophiles, i.e. glutathione, that attack the activated ester bond.

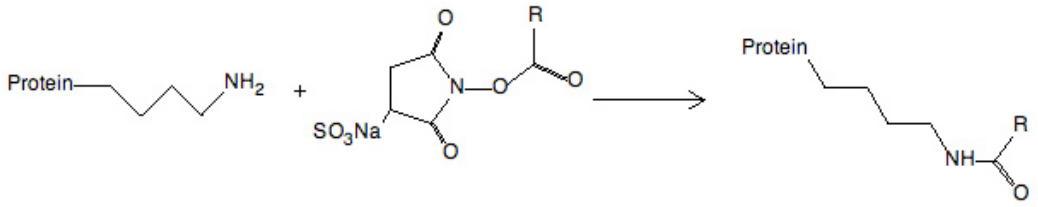


Figure 1.9. Sulfo-NHS-Ester reaction with a free amine.

Thiol Reactive Chemistry

Another reactive nucleophile that is often solvent exposed on the protein surface is cysteine. A malemide group, the most common cysteine modifying reagent (**Figure 1.10**) usually makes up one-half of a hetero-bi-functional cross-linker, and is used to site-specificly tether cross-linking moieties into a purified protein. The labeled protein can then be added to a mix of potential protein binding partners and the reaction is allowed to proceed with the goal of capturing an interacting protein interface. This is usually accomplished with the sulfo-NHS chemistry mentioned previously or photo-activated chemistry detailed in the next

section. Like malemides, a 2-thiopyridyl moiety (**Figure 1.10**) also permits cysteine-specific incorporation of a cross-linking reagent into a protein of interest by taking advantage of this group to form disulfide linkages. This chemistry is reversible under mild reducing conditions effectively opening up the way for label transfer protocols[34].

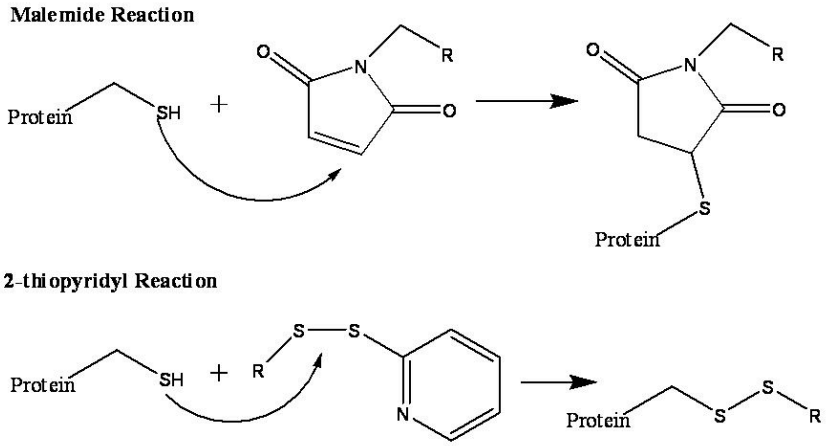


Figure 1.10: Common Thiol Reactivity Chemistry

UV Crosslinking Chemistry

Phenyl azides, benzophenones, and diazirine rings are the three types of UV-sensitive functional groups that are employed as photo-activatable cross-linking reagents (**Figure 1.11**). Upon irradiation, these inert moieties are transformed into highly reactive groups capable of capturing interacting proteins through C-H bond insertion or reaction with nucleophilic groups. The ability to trigger a cross-linking reaction is a significant advantage over previously addressed cross-linking

approaches because it allows for cross-linking information to be gathered in a temporal fashion.

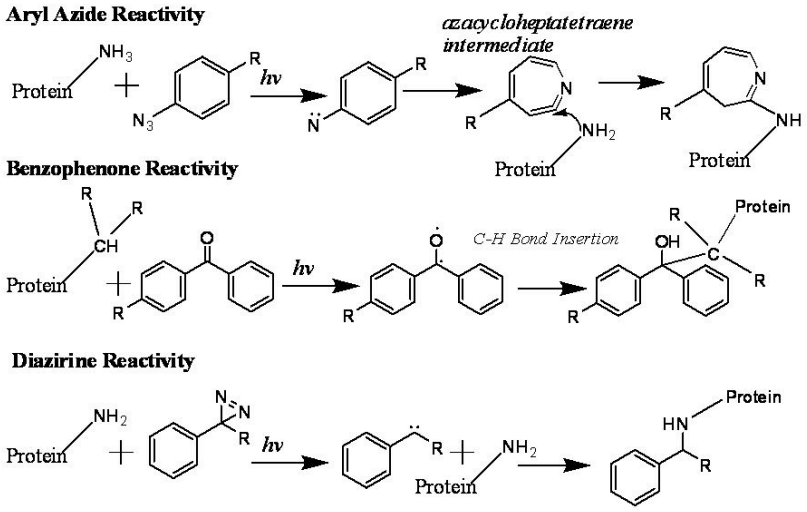


Figure 1.11: Common UV-Sensitive Photo-cross-linking reagents.

Aryl azides are generally illuminated with a wave-length shorter than 360 nm and undergo formation of an azacycloheptatetraene intermediate (**Figure 1.11**) through rapid intramolecular ring expansion of the nitrene. Azacycloheptatetraenes are highly electrophilic, but significantly less reactive species than typical nitrenes or carbenes[47]. Aryl azides usually require 10-30 minutes of irradiation to produce detectable cross-linking yields. These yields are typically very low, .1%-1% of the interacting complex is retained in a cross-link. For this reason, radioactive tags are often tethered to the cross-linker to enhance signal detection. Perhaps the best example of a chemical cross-linking reagent utilizing an aryl azide was developed by the Ebright group with the synthesis of a bi-functional cross-linker that tethers an aryl azide to an exposed cysteine on the the protein's surface via the aforementioned 2-thiopyridyl group (**Figure 1.12**). Cysteines can be engineered into

a protein of interest at defined sites, the cross-linker attached via thiol exhange, and the reaction triggered to reveal novel binding partners. Although several groups have utilized this technique to probe protein-protein interactions two high-profile articles demonstrate both the strengths and weaknesses of the bi-functional cross-linker and are diagramed in **Figure 1.12**[3, 48]. In the first example, **Figure 1.12A**, Pickart and co-workers attempted to determine the tetra-ubiquitin binding site within the 26S proteasome. To accomplish this, tetra-ubiquitin chains were built with a mutant form of ubiqutin that contains a proline 37 to cysteine mutation. This form of ubiquitin was built into tetrameric ubiquitin chains and labeled at cysteine 37 with the radio-iodinated aryl azide. After UV irradiation and cross-linking the samples were reduced and separated on a standard SDS-PAGE gel. Under reducing conditions the radio-label was transferred from the cysteine 37 on ubiquitin to the now-cross-linked site of association. From these data the site of interaction was determined to be Rpt5 a triple-ATPase subunit that resides in a hexameric ring at the base of the 19S proteasome. In a technically similar set of experiments Ptashne and co-workers utilized the same cross-linking chemistry to help determine the protein interaction partner of the Gal4 activation domain in the context of a yeast mediator sub-complex consisting of Srb2, 4, 5, and 6 (**Figure 1.12B**). Incubation of a labeled Gal4 activation domain with the tetrameric complex revealed that the aryl azide cross-linked to Srb4. Importantly, this cross-linking signal was competed away in the presence of Gal80.

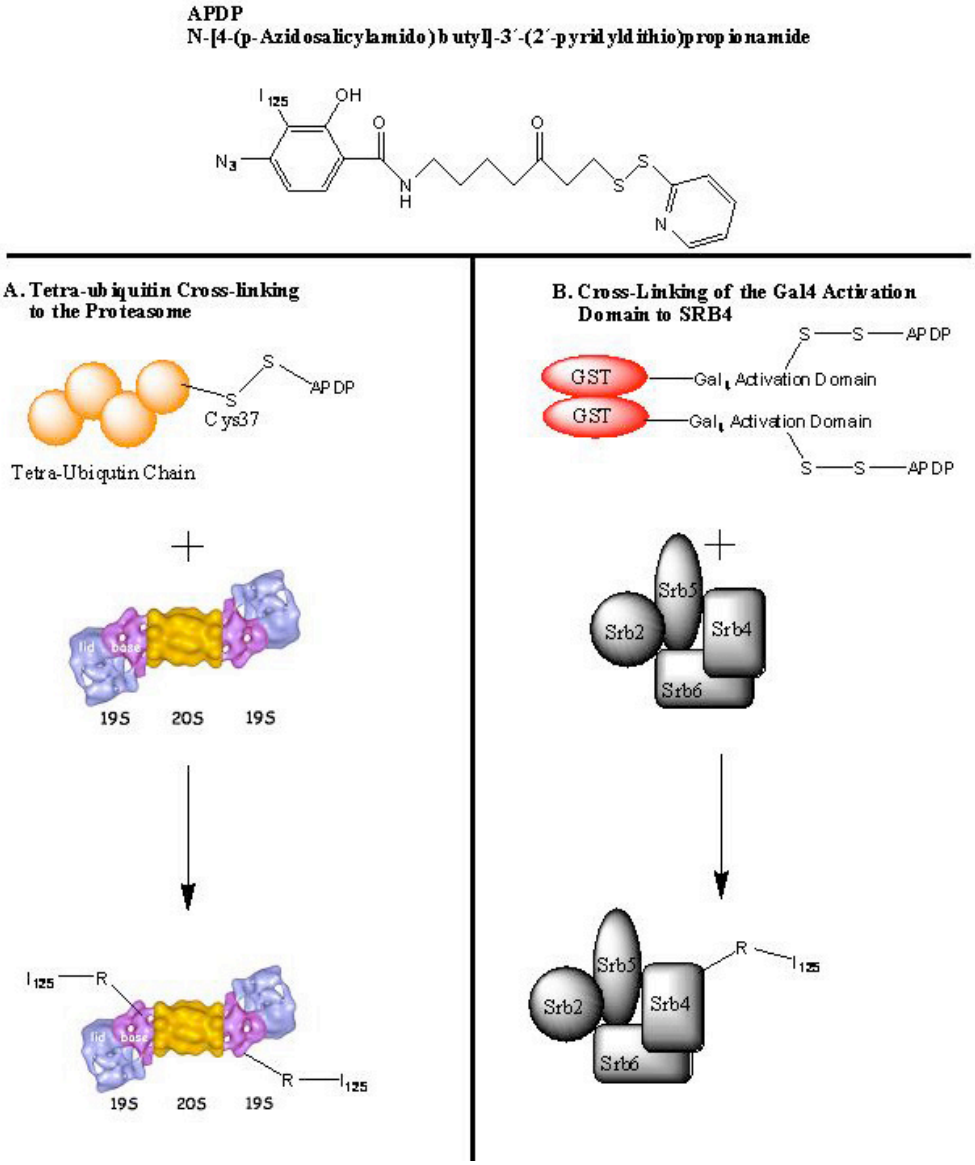


Figure 1.12: Examples of aryl azide cross-linking A.) Cross-linking of tetraubiquitin to the 26S proteasome[2]. B.)Cross-linked the Gal4 activation domain to a sub-complex of Mediator[3]

Although both data sets produced answers to specific biological questions not easily addressed by conventional techniques, these experiments also highlight the weaknesses of traditional bi-functional cross-linkers. First, both sets of experiments were carried out in a non-reducing, relative homogenous molecular environment that does not mimic the native environmental conditions inside the cell. Second, both experiments required that protein of interest be purified and labeled with the cross-linker prior to the experiments an onerous requirement for most proteins within the cell. Finally, cross-linking yields from both sets of experiments were less than 1%, hence the requirement for the radio-label. The combination of low-yields and the artificial environment cast an element of uncertainty upon the obtained results. Specifically, the question of whether the reported interactions actually occur within cells remains open for debate.

In an attempt to move protein cross-linking technology into cells, Thiele and coworkers synthesized three amino acids that are isosteric to isoleucine, leucine and methionine but contain a photoactivatable diazirine ring and yield a reactive carbene after the light-induced loss of nitrogen[1]. Surprisingly, “photo-leucine” and “photo-methionine” were incorporated into proteins when added to the cell media without detectable cytotoxic effects (**Figure 1.13**). After three minutes of UV irradiation at 350 nm, cells were lysed and individual protein cross-links were assayed by western blotting for a protein of interest. The salient result of the work was the observation

that SCAP AND Insig, resident ER proteins, cross-linked to a progesterone binding membrane protein. In short, this method provides a means to test predicted protein-protein interactions within the membranes of live cells. Although cross-linking yields were extremely low, incorporation rates were less than 1%, and cross-linked products could only be detected by western blotting this work was a significant accomplishment given the current state of the field.

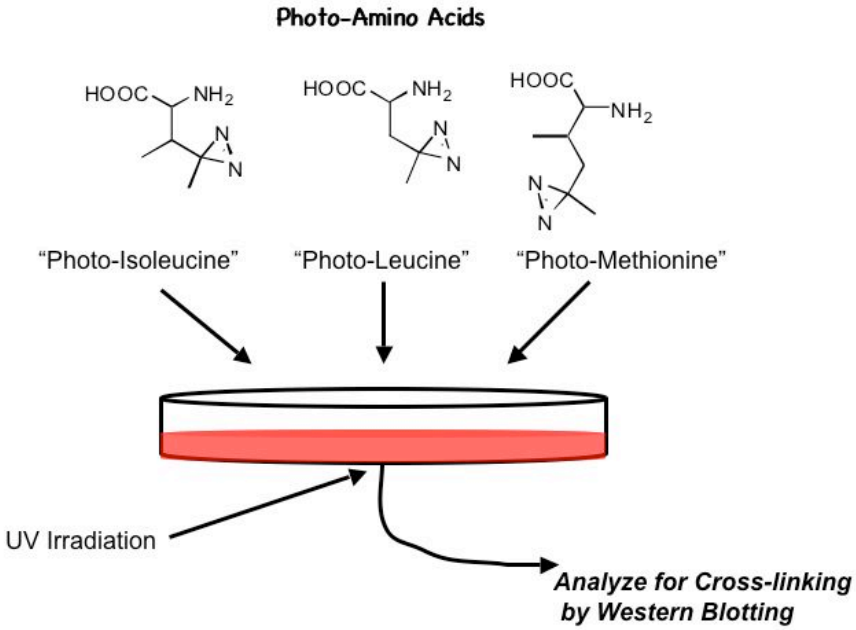


Figure 1.13: In vivo incorporation of isosteric diazirine amino acids into cells for photo-cross-linking[1].

Finally, benzophenones have been considered to be the most advantageous of all photocrosslinkers in use for two reasons. First, benzophenones are chemically more stable than aryl azides and most diazirines. Second, benzophenones can be manipulated in ambient light and activated at 350-360 nm[49-52]. Upon irradiation at 350 nm benzophenones are capable of inserting into C-H bonds with irradiation

times typically being on the order of 30 minutes. Despite various claims of C-H bond insertion a 1998 survey of studies using UV triggered cross-linkers has revealed that primary sites of modification are typically lysine, tyrosine, and tryptophan and that C-H bond insertion was a relatively rare event when using these compounds[53].

An exceptionally interesting application of benzophenone cross-linking comes from work done in the field of in vivo artificial amino acid incorporation[54-63]. The Schultz lab has created several orthogonal t-RNA/ aminoacyl-t-RNA synthetase pairs that incorporate select artifical amino acids at TAG amber codons in E.Coli and yeast. Schultz and co-workers engineered the system to insert benzophenones into the homo-dimer GST and demonstrated in-vivo crosslinking at a calculated yield of 50%. The site selective incorporation of cross-linkers into proteins of interest opens up the opportunity to study protein-protein interactions in their native state and should continue to prove invaluable in future protein-protein cross-linking applications.

Oxidative Cross-linking Chemistry

Over the last several years the Kodadek lab has developed a new class of zero Å cross-linking reagents that initiate coupling of interacting proteins through oxidative chemical reactions. The most efficient version of this process employs a Ru(III) complex as the oxidant. This species is generated by brief photolysis of ruthenium(II) tris-bipyridyl dication ($Ru(II)(bpy)_3^{2+}$ with visible light in the presence of an electron acceptor such as ammonium persulfate (APS) or $Co(III)(NH_3)_6^{3+}$ (**Figure**

1.14). Reaction times are quick, consisting of a 0.5 second irradiation using a high-intensity mercury lamp with a 400 nm cut-on filter and cross-linking yields of 10-90% are usually observed for a typical reaction as long as aromatic groups such as tyrosine or tryptophan are present at the protein interface. Importantly, proteins not closely associated in solution do not cross-link at physiological relevant protein

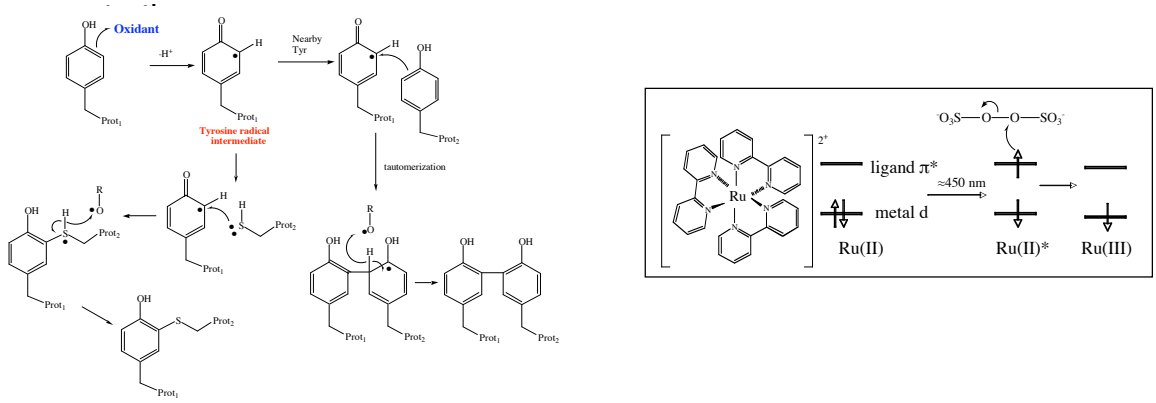


Figure 1.14: Photo-Induced Cross-linking of Unmodified Proteins (PICUP)

Denison and Kodadek applied this technique towards the goal of mapping the molecular architecture of the 20S proteasome. Most of the interactions between the 28 proteins of this multi-protein complex were detected by this technique without the observation cross-linking artifacts amongst non-interacting proteins. The data produced from this experiment could not however be interpreted without the aid of the 20S crystal structure for two reasons. First, cross-linked peptides were not observed in the mass spectrometer. Second, the efficiency of the reaction created multiple cross-linked species that contained more than simple binary interactions, thus making it difficult to determine the molecular connectivity of the complex.

Both UV triggered reactions and PICUP chemical reactions are inherently complex and the final products are difficult to characterize at the peptide level. Many products tend to rearrange and under-go secondary reactions. This contributes to the overall modification of the proteins involved in the cross-linking reaction. Unpredictable rearrangements are a fundamental problem when trying to analyze cross-linked products by mass spectrometry, due to the precise mass requirements necessary to identify a protein or peptide of interest. There are, of course, several solutions to this problem such as more specific elucidation of the mechanistic details of current photo-activated chemistry within biological systems and the creation of more efficient iterations of approaches. Perhaps more importantly, the generation of algorhithms for efficient mass spectrometric analysis of cross-linked products should aid in the identification of crosslinked protein interfaces[47].

The Advantage of Multiple Approaches

It is often difficult to assess to what extent any one large scale two-hybrid, affinity pull-down experiment captures the entire “interactome” within a cell or even if a simple protein-protein cross-linking experiments captures all of the actual interaction partners for a protein of interest. For this reason, a variety of interaction techniques are needed to provide a complete picture of the biological molecular architecture operating inside cells.

Chapter 2

Development of DOPA Cross-linking Chemistry

Introduction

Elucidating the molecular organization of the cell in real time and space is a formidable challenge facing molecular biologists. Chemical cross-linking is a potentially powerful method for use in this challenge, allowing for the anlaysis of protein-protein and protein-biomolecule interactions. To complement existing methodologies, most of which are limited in many ways, we have developed a series of oxidative coupling reactions that are capable of cross-linking many protein-protein and protein-peptide complexes rapidly and in good yield[37, 64-66]. However, a limitation of this approach is that when one examines a large multi-protein complex or carries out experiments in complex mixtures, a large number of multiply cross-linked species are produced, complicating the separation and analysis of products[37]. For many applications, such as identifying the target of bioactive molecules[67], it would be desirable to be able to focus the oxidative reaction to a single molecule of interest without perturbing any other proteins present. As mentioned in Chapter 1 there have been a few cross-linking chemistries aimed at achieving this goal[34]. However, these approaches suffer from a general incompatibly with biological systems. For example, cross-linking yields are often poor in aqueous environments, while the cross-linking require reaction times on the order of hours as opposed to the more relevant biological

time scale of seconds. We describe here a useful method of this sort that allows molecules containing an ortho-dihydroxyphenyl ring to be linked covalently to associated proteins in a process that is orthogonal to the functional groups present in most proteins. Our hypothesis, based on the known oxidation chemistry of 3,4-dihydroxyphenylalanine (DOPA)-containing proteins common in mollusks and insects[68-74], was that this side chain could be oxidized selectively using sodium periodate to produce an ortho-quinone intermediate that could be attacked by nearby nucleophiles, resulting in a stable cross-link (**Figure 2.1**). Given that proteins are mostly unaffected by sodium periodate and that DOPA chemistry naturally occurs in many biological organisms, we reasoned that this approach would function within biological systems with minimal perturbation.[70, 74]

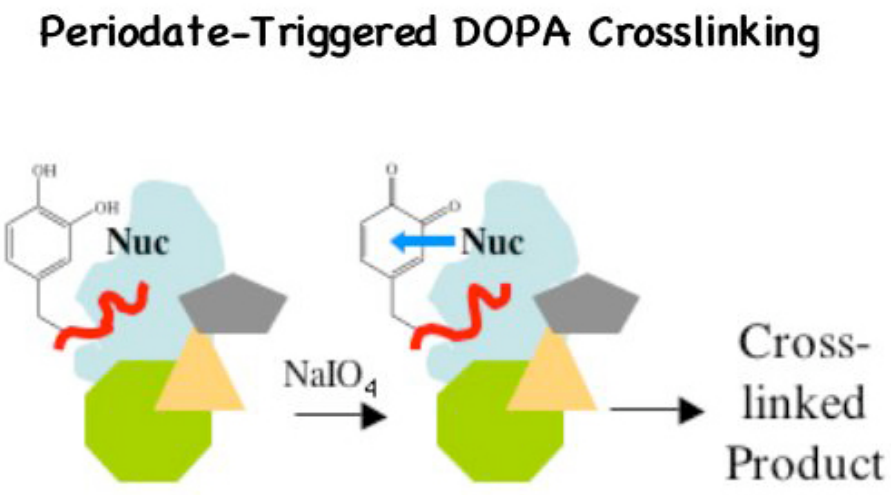


Figure 2.1: DOPA Cross-linking Hypothesis. Upon binding to a multi-protein complex, a DOPA containing molecule can be selectively oxidized with periodate to form an ortho-quinone intermediate. This intermediate in turn should be open to attack from neighboring nucleophiles. Importantly, bulk nucleophiles in the milieu should prohibit spurious cross-linking.

Specific Cross-linking within a Model Peptide-Protein Complex

As a model to test this hypothesis we used a complex between the yeast Gal80 transcriptional repressor and a 20-residue peptide that had been selected by phage display to bind Gal80 (NH_2-YDQDMQNNTFDDLFWKEGHR-COOH; K_D = 300 nM)[75]. The binding properties of this peptide to Gal80 and other transcription proteins are of interest because the peptide functions as an artificial activation domain in vivo and in vitro when fused to suitable sequence-specific DNA-binding domains[75-77]. We first analyzed whether the chemical reaction was triggered in the hypothesized manner by inserting tyrosine, DOPA, and Dimethoxy-phenylalanine at the same position within the Gal80 binding peptide. Periodate is known to associate with diols to form a periodic ester product (**Figure 2.2**). Tyrosine lacks adjacent hydroxyl groups, while the hydroxyl groups are methylated in dimethoxy-phenylalanine (**Figure 2.2**). Therefore, only the DOPA containing peptide should be competent to undergo oxidation by sodium periodate to generate the reactive ortho-quinone and undergo the cross-linking reaction.

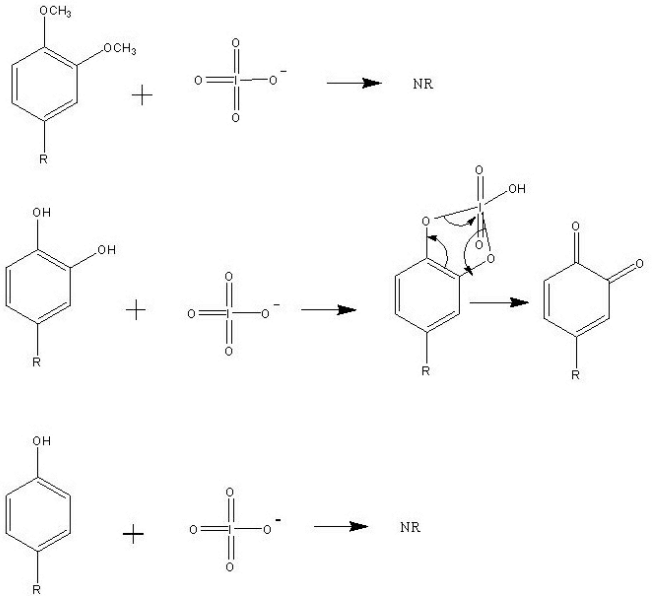


Figure 2.2: Selective oxidation of DOPA with periodate. NR=No Reaction

Two Gal80-binding peptide analogues were synthesized by replacing the N-terminal tyrosine of the peptide with either DOPA or 3,4-dimethoxyphenylalanine (**Figure 2.3**). All three peptides were extended at the N terminus with two glycines and a biotinylated glutamic acid to allow cross-linked products to be detected by blotting with a NeutrAvidin-horse radish peroxidase (HRP) conjugate (**Figure 2.3**). A six-histidine (His_6)-tagged Gal80 protein was incubated with each of the three peptides. Thirty seconds after periodate addition the reaction was quenched using a buffer containing 100 mM DTT. The products were analyzed by denaturing gel electrophoresis and blotting with NeutrAvidin-HRP. Cross-linking occurred only between the DOPA-containing peptide and Gal80p in the presence of periodate (**Figure 2.3**, lanes 1 and 2). No reaction was observed when the peptide contained the native tyrosine (lanes 5 and 6) or a dimethoxy-substituted ring (lanes 3 and 4) or when periodate was omitted (lane 7). These results, particularly the lack of reactivity of the dimethoxyphenylalanine-containing molecule, are consistent with an ortho-quinone intermediate that then cross-links to nearby residues on Gal80.

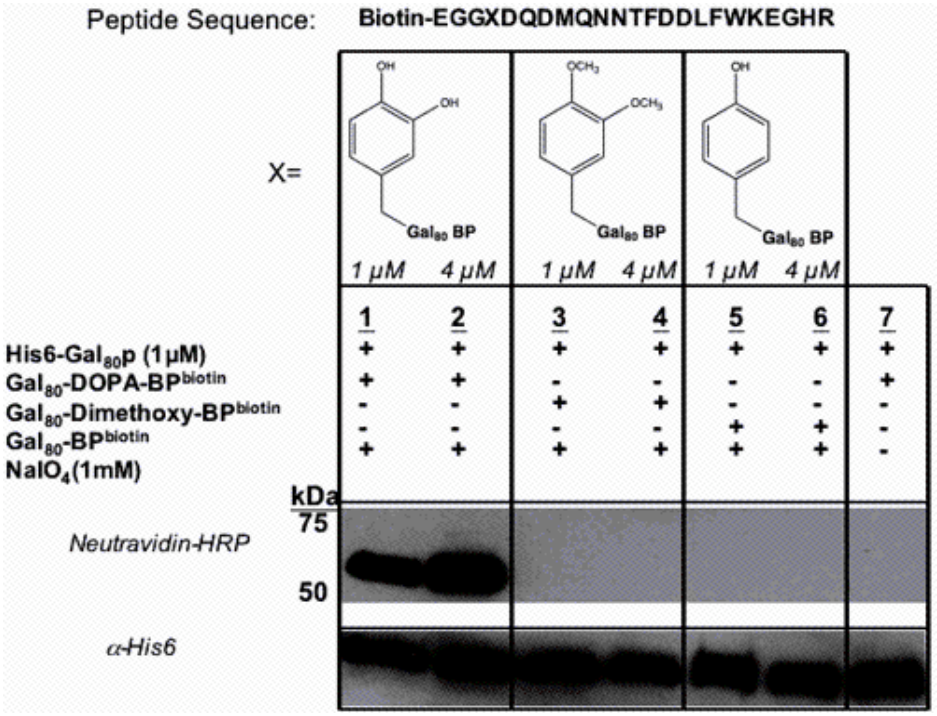


Figure 2.3: Figure 1 Lanes 1 and 2: Chemical cross-linking reaction between His6-Gal80p and Gal80-DOPA-BP. Lanes 3 and 4: Chemical cross-linking reaction between His6-Gal80p and Gal80-dimethoxy-BP. Lanes 5 and 6: Chemical cross-linking reaction between His6 Gal80p and the Gal80-Y-binding peptide. Lane 7: His6-Gal80p +Gal80-DOPA-BP without the addition of periodate. Probing the blot with anti-His6 antibody showed that equivalent amounts of protein were loaded and that the His_6 tag had not been damaged by the cross-linking reaction.

It is important to also note that although Gal80 exist in solution as a homo-dimer , no cross-linking was observed between the individual Gal80 subunits. This observation highlights the focused nature of the cross-linking reaction. A problem with many cross-linking reactions is that they can produce products between molecules not tightly associated in solution as mentioned in Chapter 1. To assess this issue in the context of this chemistry, one microgram of His_6-Gal80 was mixed with a large excess, 150 micrograms, of 10 other proteins listed

in the experimental protocols, and the experiment was repeated. As shown in **Figure 2.4**, lane 6, only a single cross-linked product was obtained. Indeed, the result was almost identical to that of the control experiment lacking the competitor proteins (lane 3), except that the yield was reduced by over 50%. This was most likely due to some competition of the peptide-Gal80 complex by the other proteins, which was not surprising since activating peptides typically bind nonspecific proteins promiscuously. As a control, the same experiment was done with a biotinylated DOPA-containing peptide (biotin-KG-DOPA-AHNRLIYMQD) not known to associate with any of the proteins in solution (lane 8). No reaction was observed between the DOPA-containing control peptide and any of the proteins present. We conclude that this reaction reports only stable intermolecular associations.

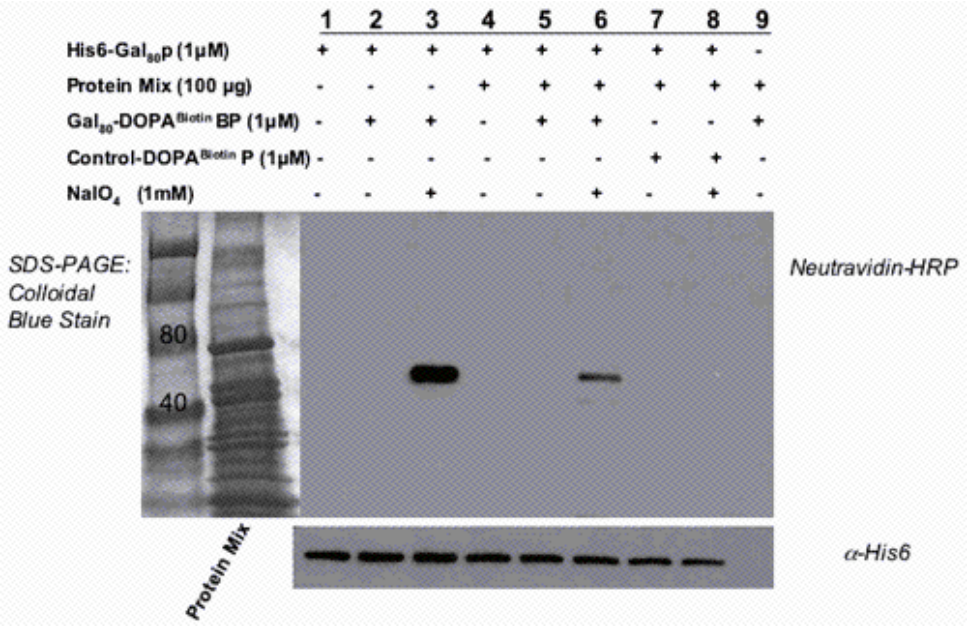


Figure 2.4: Cross-linking between His6-Gal80p and the Gal80-DOPA-binding peptide is only seen in the presence of periodate (lane 3). The addition of the protein mix shown on a colloidal stained SDS-PAGE gel to the left does not significantly interfere with the cross-linking reaction as shown in lane 6. No cross-linked products were detected when

the control-DOPA peptide was subjected to the same reaction (lane 8). The bands in the top NeutrAvidin-HRP blot overlay with the bands in the bottom panel (-His6), proving that Gal80p is cross-linking to Gal80-DOPA-BP despite the presence of the protein mix.

To evaluate this chemistry in a more demanding setting, we examined the interaction of the Gal80-binding peptide with the 19S regulatory particle of the proteasome. Recently, evidence from a number of laboratories, including our own, have shown that the proteasome is intimately involved in transcription[78]. Proteasome-mediated turnover is required for the function of some transcription factors. In a mechanistically distinct process, the Saccharomyces cerevisiae Gal4 transcription factor activation has been shown to recruit a fragment of the proteasome (coined APIS) including the ATPases, but excluding the proteolytic core, to activated promoters via it's AD[78]. This event is critical for subsequent promoter escape and elongation by polII. Genetic evidence implicates Rpt4/Sug2 and Rpt6/Sug1 as the targets of the Gal4 AD within the 26S proteasome and/or the APIS complex[79-84]. These data are consistent with biochemical results showing that a GST-Gal4 AD fusion protein binds in vitro translated Sug1 and Sug2and that the Gal4 AD binds Sug2 in a two-hybrid-like assay in yeast[79-84]. However, these interactions remain to be validated in the context of the native complex. In the case of Gal4, interaction of the AD with transcription and proteasomal proteins is blocked by the repressor protein Gal80, which binds tightly and specifically to the Gal4 AD under non-inducing conditions[78, 85-90]. These binding events are important for efficient Gal4-mediated transcription. Since the Gal80-binding peptide also functions as an activation domain in vivo[91], we

wondered whether it would exhibit the same binding properties, even though it had not been selected to bind the 19S complex. As shown in **Figure 2.5**, the DOPA-containing Gal80-binding peptide cross-linked cleanly to two proteins of about 50 kDa. These were identified as Sug 1 and Sug 2 by Western blotting with highly specific antibodies and overlaying the Western and NeutrAvidin-HRP blots (**Figure 2.6**). When this experiment was repeated in the presence of a 50-fold excess of the tyrosine-containing Gal80-binding peptide, the DOPA-dependent cross-linked product was reduced drastically (**Figure 2.5, lane 3**), showing that the native and DOPA-containing peptides compete for the same sites on the 19S complex. An interesting sidelight to this experiment was the detection of a small amount of a third cross-linked product (lane 2). Western blotting revealed that this species represented cross-linking of the peptide to Rpn2 (**Figure 2.6**). Since Rpn2 is thought to interact with Sug1 and Sug2, these experiments suggest that the three proteins form a pocket in the 19S complex responsible for binding activation domains.

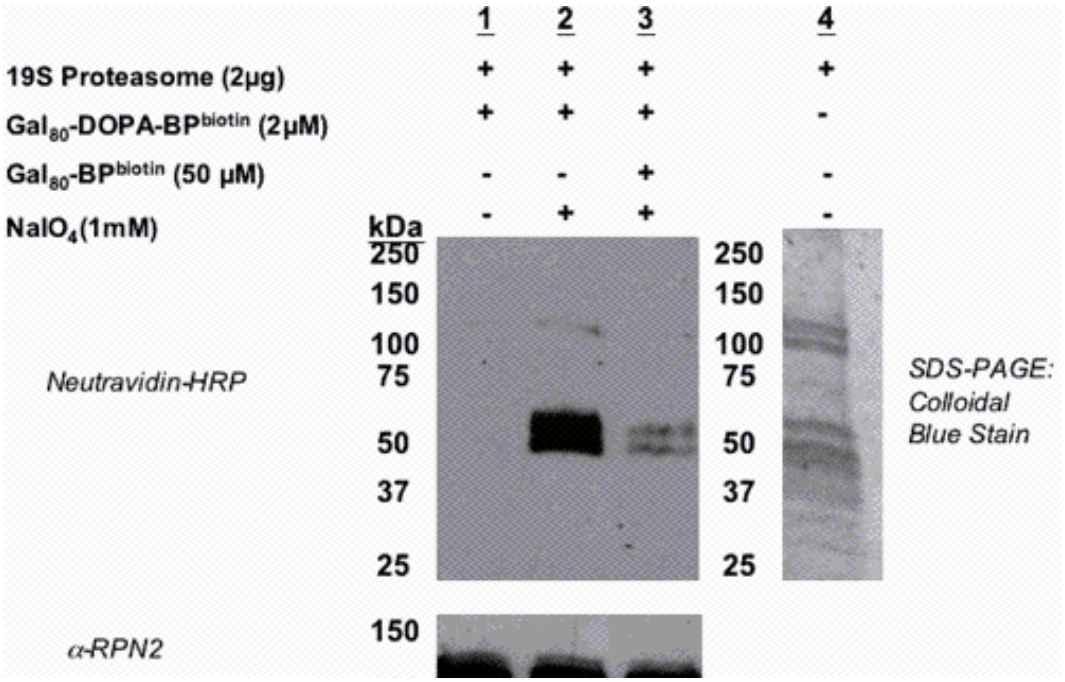


Figure 2.5: Peptide-19S complex cross-linking. Lane 1: 19S + the Gal80-DOPA-BP in NE buffer without periodate. Lane 2: 19S + the Gal80-DOPA-BP in NE buffer with periodate. Lane 3: 19S + the Gal80-DOPA-BP + excess Gal80-BP+ periodate in NE Buffer. Lane 4: Collodial blue stained SDS-PAGE gel showing the amount of 19S added to each reaction. The bottom -RPN2 blot serves as a loading control. The blot was stripped and reprobed with anti-Sug1, anti-Sug2, anti-Cim5. These blots are presented in the Supporting Information.

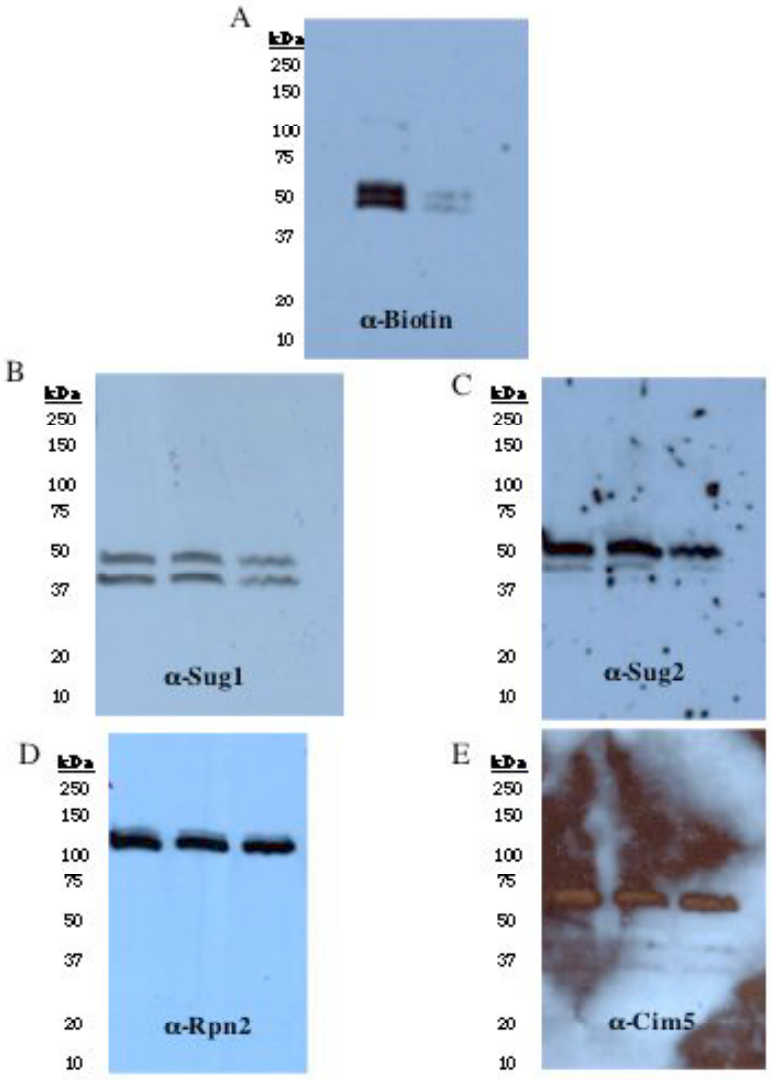


Figure 2.6: . The same blot from Figure 2.5 stripped and reprobed with anti-Sug1, anti-Sug2, and anti-Cim5 (A) Crosslinking reaction of the 19S Proteasome with the Gal80-DOPA Binding Peptide after probing the blot with Neutravidin-HRP. (B) Blot was stripped and re-probed with Sug1 antibody. (C) Blot was stripped and re-probed with and antibody against Sug2. (D) Blot was stripped and reprobed with an antibody against RPN2 (E) Blot was stripped and re-probed with and antibody against Cim5 .

In conclusion, an operationally simple and highly specific cross-linking chemistry has been developed. Our preliminary studies suggest that the ortho-quinone resulting from periodate oxidation of the DOPA residue is an important intermediate in the reaction, but the chemistry of how this species couples to closely associated proteins remains to be determined. Whatever the detailed mechanism, the reaction should be useful for the study of peptide-protein interactions and more generally for small-molecule-protein complexes where the former contains the appropriate oxidizable ring. The chemistry is orthogonal to the 20 common amino acids present in most proteins, and thus the oxidizing power can be focused specifically to the molecule of interest without the production of multimeric species containing protein-protein as well as peptide-protein cross-links. Furthermore, protein-engineering methods developed over the past few years now make it feasible to place such selectively oxidizable residues into specific positions[92] of native proteins, thus making this approach potentially applicable to the study of protein-protein interactions.

Development of an Experimental System to Study DOPA Reactivity

In the aforementioned set of experiments we simply utilized DOPA as a peptide-protein cross-linking reagent. However, this artificial redox-active amino acid is found in a wide variety of natural biochemical reactions ranging from structural maintenance of insect exo-skeletons to neurotransmission in humans[68, 73, 93-95]. Furthermore, a number of syndromes, such as Parkinson's disease, may

involve cross-linking of oxidized DOPA to neuronal proteins such as Parkin or alpha-synuclein. Although biochemical reports indicate that cysteine is the most likely amino acid responsible for DOPA-protein coupling in the case of Parkin[96], however, it is unclear how DOPA might react with alpha-synuclein since this protein does not contain a cystiene[97, 98]. For these reasons we sought to elucidate the relative reactivity of various amino acid side chains with ortho-quinone intermediate. This work was done in collaboration with a post-doctoral researcher in the lab, Bo Liu, who was instrumental in the synthesis of the peptide nucleic acid constructs, optimization of the assay conditions, and analysis of the cross-linked products.

Previous results indicated that the DOPA residue only seems capable of cross-linking with proteins in very close proximity, i.e. protein-protein interaction partners. This type of templated reaction is quite different than the standard chemical reaction, because although DOPA is present at a relatively low total concetration in solution, the effective concentration of DOPA to the cognate protein interface is quite high[99-101]. Therefore, to study the DOPA reactivity in conditions similar to that in nature, we used short complementary peptide nucleic acid (PNA)[102] strands to bring individual amino acid residues and DOPA close to each other through annealing, subsequent to triggering the reaction with sodium periodate. The amino acid-DOPA cross-linking products were then detected by high performance liquid chromatography (HPLC) and the cross-linked product analyzed with mass spectroscopy.

A DOPA amino acid was conjugated to the C-terminal of an octameric PNA strand via an acetyl-ethyleneglycol-ethyl-amine (AEEA) linker (**Figure 2.7**). Various other amino acids were also conjugated to the N-terminus of a complimentary PNA strand via an AEEA linker (**Figure 2.7**). 1.7 mM amino acid-PNA2 and 1.0 mM PNA1-DOPA conjugates were mixed in a 200-mL neutral pH buffer containing 100 mM sodium phosphate, 138 mM NaCl, 2.7 mM KCl, and 0.5 mM dithiothretol (DTT). The mixture was heated at 99 °C for 1 min and cooled slowly to room temperature to allow PNA conjugates to anneal. 1 mM sodium periodate was added to trigger the cross-linking reaction, which was allowed to proceed for 90 seconds. The mixture was then injected on a C-18 reverse phase analytical column for HPLC analysis. Under these conditions, sodium periodate eluted from the column almost immediately, effectively quenching the reaction.

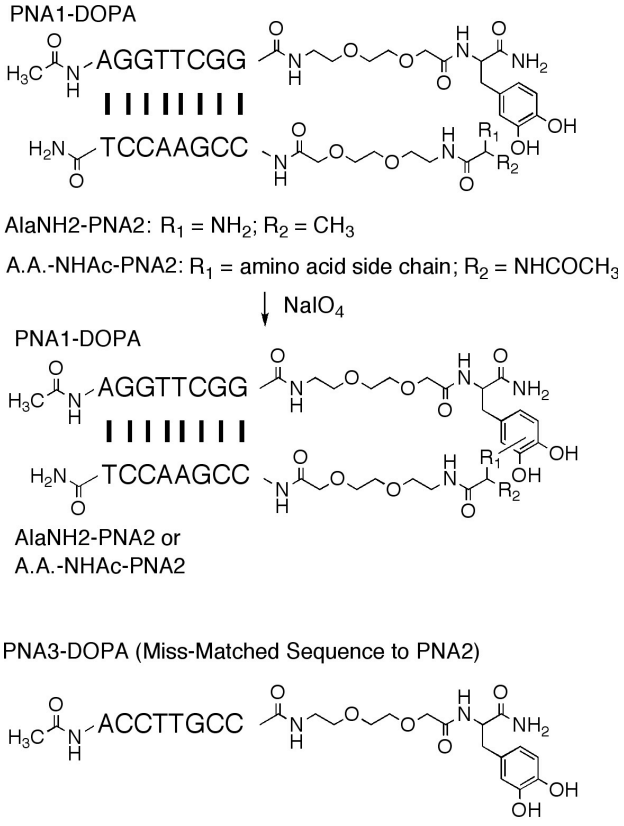


Figure 2.7: Methodology used to analyze DOPA reactivity. Various amino acids were conjugated to the C-terminus of an octameric PNA. A complementary PNA was synthesized with DOPA at the N-terminus. With this method DOPA is positioned in close proximity to the amino acid side chain in question. A non-complementary DOPA containing PNA was also synthesized as a negative control.

Coupling of the Quinone Intermediate with Nitrogen Nucleophiles

We first examined potential coupling between oxidized DOPA and an α-amino group . Alanine was conjugated to PNA2 to form AlaNH2-PNA2 (**Figure 2.8**). In the absence of sodium periodate induction, the annealed complex AlaNH2-PNA2 / DOPA-PNA1 yielded two well-separated peaks on HPLC (peak 1 and 2, **Figure 2.8**). The identities of the two peaks were confirmed by MALDI-TOF (**Figure 2.8 B**), demonstrating that the PNA duplex was unstable under these experimental conditions. When 1 mM sodium periodate was added to the complex, a new peak (peak 3, **Figure 2.8A**) emerged at longer retention time (~ 26 min), accompanied by reduced intensities of peak 1 and 2. We identified peak 3 as the a-amino-DOPA cross-linking product. Peak 3 gave a clean mass spectrum with the predicted mass value of 4951.71 corresponding to the PNA 1+2 cross-linked product. Judging from the HPLC trace (**Figure 2.8A**) this reaction proceeded to > than %80 yield, as estimated from the relative peak heights of peak 1 (present in 1.7-fold molar excess) to peak 3. To determine if PNA-PNA hybridization was required for the formation of this product, we repeated the experiment with the mismatched PNA's 2 and 3 (**Figure 2.9B**) conjugated to alanine and DOPA, respectively. In this case, when sodium periodate was added, no cross-linking product was detected (**Figure 2.9B**).

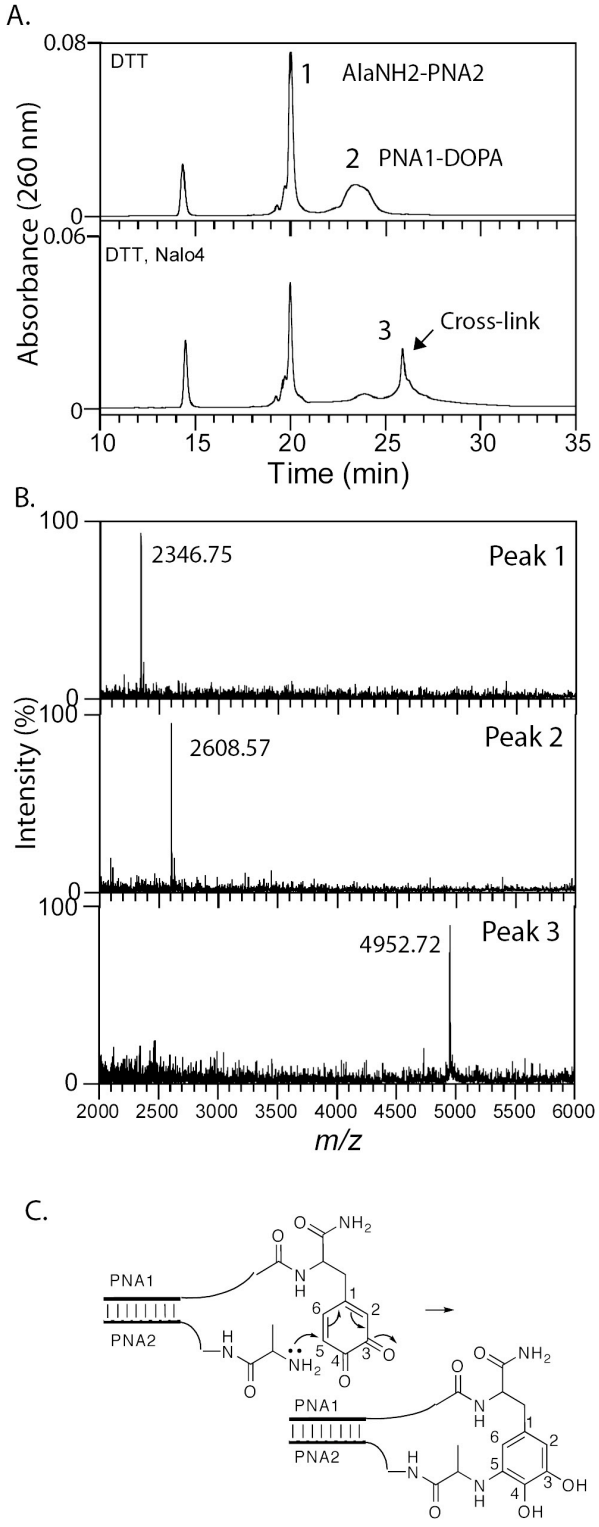


Figure 2.8: Crosslinking of α-amine with DOPA. A.) HPLC chromatograms demonstrating that the cross-linked product (3) is only formed in the presence of periodate. B.) Mass of peaks 1,2, and 3. C.) Model of α-amine reactivity with DOPA.

To determine if the nuclophilic α-amino group is essential for cross-linking, as anticipated, we carried out a nearly identical experiment using the matched PNA's, with the exception being that the α-amino group attached to strand 2 was acetylated (**Figure 2.9**). As shown in Figure 2.9A, addition of periodate to the PNA duplex did not produce a new peak, suggesting that clean covalent coupling of the two strands did not occur. However, we not that the peak corresponding to the DOPA conjugated PNA strand was diminished substantially, suggesting that it underwent other reactions that did not produce a single product. This result suggest that clean coupling requires a potent nucleophile.

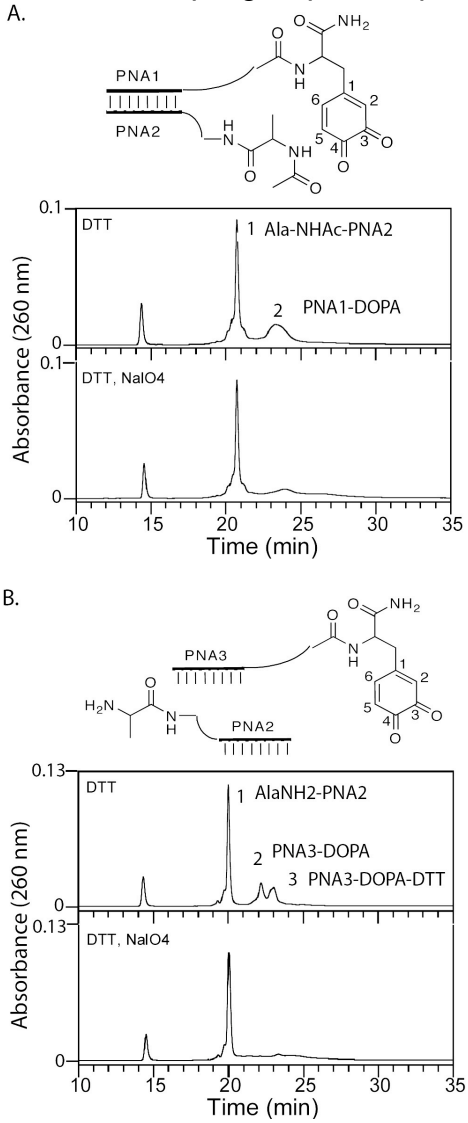


Figure 2.9: A.) Cross-linking results between a complementary acetylated α-amino PNA and DOPA-PNA. B.) Cross-linking results of the reaction between non-complemenatry DOPA and an alanine-α-amino PNA.

Cross-linking of Lysine and Histidine Side-chains to DOPA ortho-quinone

With the analytical system established, we turned to study whether other prominent nucleophiles on the the protein surface, lysine and histidine, were potential reactive partners with oxidized DOPA. These amino acids, with the α-amino group blocked by acetylation, were conjugated to PNA2 and hybridized to PNA1-DOPA. Upon sodium periodate treatment, the Lys-NHAc-PNA2 / PNA1-DOPA complex yielded a distinct cross-linking peak on HPLC (peak 3, **Figure 2.10A**). The mass of the cross-link product was measured as 5052.25 (**Figure 2.10B**). This mass value was consistent with a Michael addition mechanism, differing by only .5 mass units from the predicted product. The cross-linking yield was lower than that observed for coupling of the α-amino group of alanine, possibly due to the lower pKa value of the ε-amino group.

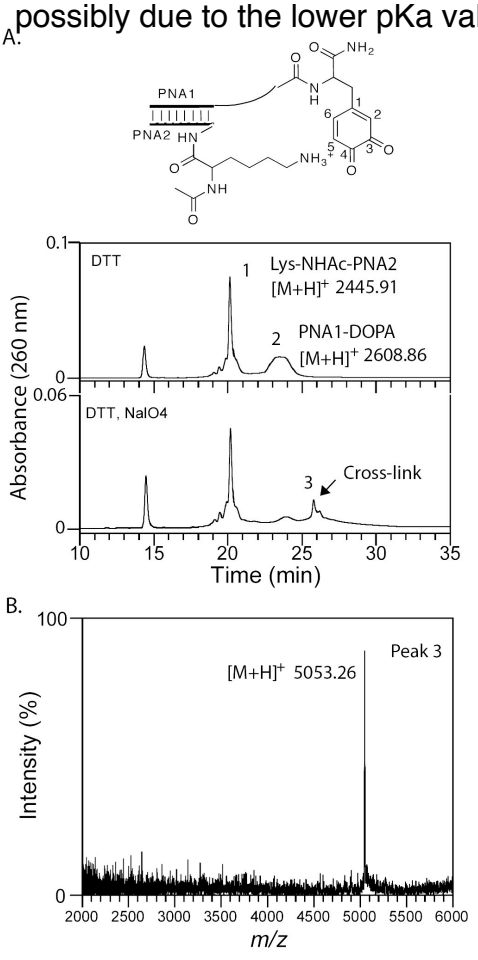


Figure 2.10: A.) Strucutre of Lysine-PNA/DOPA-PNA annealed complex (top). Chromatogram of cross-linking reactions (+/-) periodate B.) Mass of lysine/DOPA cross-linked product (peak 3).

Likewise, the His-NHAc-PNA2 / DOPA-PNA1 complex showed that the imidazole side chain of histidine was also capable of reacting with DOPA under oxidative conditions (**Figure 2.11A**). However, the cross-linking yield was the lowest among the three groups examined. The cross-link peak (peak 3, **Figure 2.11A**) gave a mass value of 5061.56 (**Figure 2.11B**) that was within .4 mass units of the predicted Michael product.

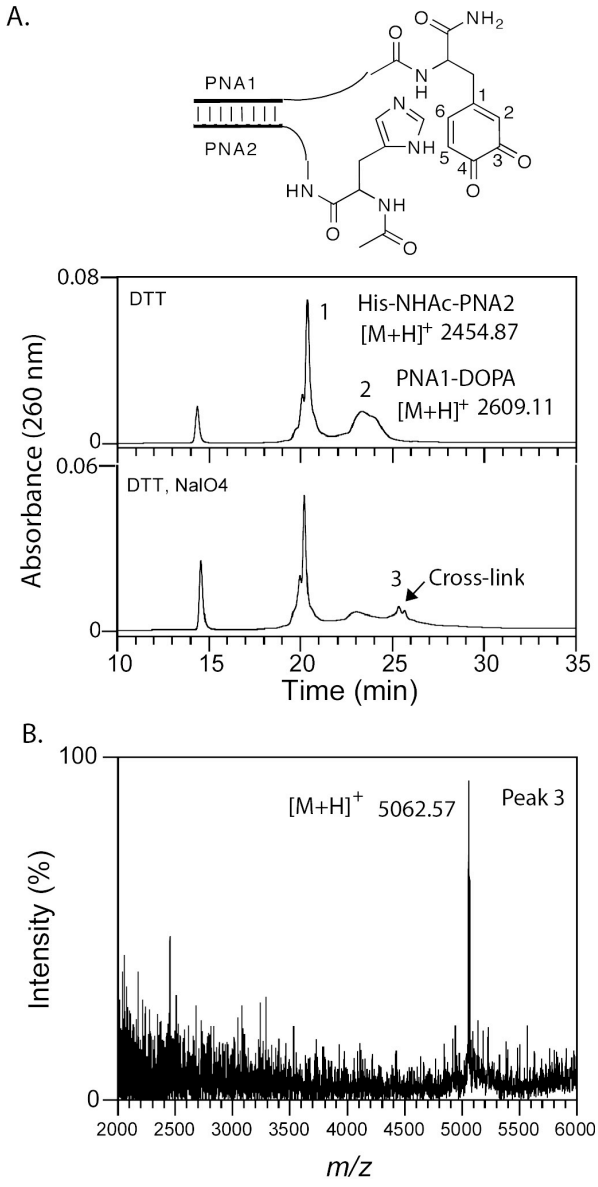


Figure 2.11: A.) DOPA-PNA1/His-PNA2 complex reaction +/- periodate. B.) Mass of cross-linked His/DOPA product (peak 3).

Analysis of Cysteine Reactivity with Oxidized DOPA

It has been known for some time that oxidized DOPA reacts efficiently with cysteine *in vitro* and *in vivo*[93, 96, 103, 104]. In order to compare the reactivity of cysteine to the reactivity of the other nucleophiles tested in the experiment we generated a cysteine containing PNA. Cys-NHAc-PNA2 and PNA1-DOPA were heated and annealed in a phosphate buffered saline containing 0.5 mM DTT. Even without sodium periodate induction, however, a distinct cross-link peak was detected (peak 3, **Figure 2.12A**). Peak 3 gave a clean mass value of 5026.40, which was within 1 amu of the predicted Michael addition product. In a reducing environment, DOPA maintains predominately the catechol state, and only trace amounts of ortho-quinone exists. However, under these conditions cysteine was capable of trapping the small amount DOPA that did spontaneously oxidize to the ortho-quinone. When Cys-NHAc-PNA2 was mixed with PNA3-DOPA, a conjugate with a mismatched PNA sequence in 0.5 mM DTT, no cross-linking product was detected (**Figure 2.12C**). Upon sodium periodate induction, the Cys-NHAc-PNA2 / PNA1-DOPA gave a broad peak at the cross-link product retention time accompany to significant losses of the intensities of the starting PNA conjugates. Unfortunately, we were unable to cleanly identify the identities of the cross-link products. A possible reason for this complication is that under oxidation condition, DTT interfered with cysteine-DOPA product formation.

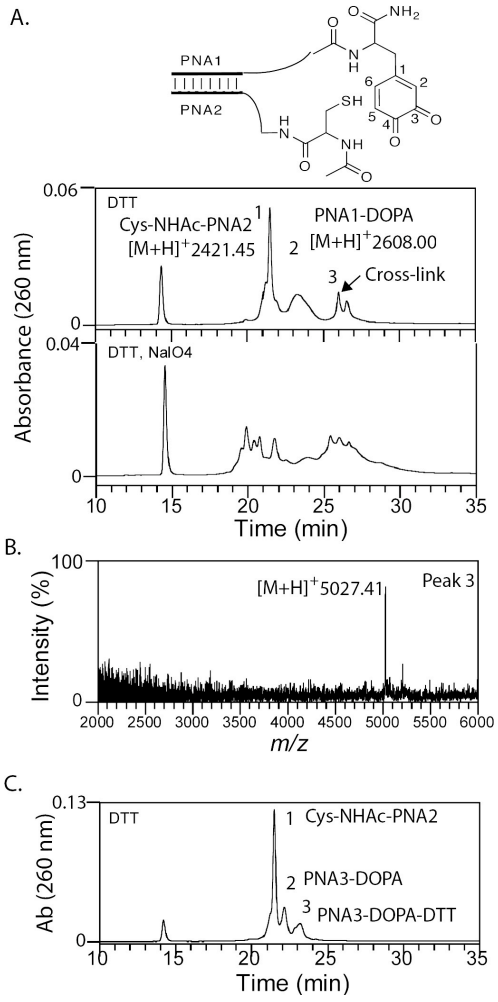


Figure 2.12: A.) DOPA-PNA1/Cys-PNA2 complex +/- periodate. The chromatograms demonstrate cross-linking even the absence of periodate (top: peak 3).

B.) Mass of peak 3.

C.) Cross-link product formation was not observed between cysteine and DOPA when the PNA stands were non-complimentary.

To evaluate the reactivity of other functional amino acid side chains with oxidized DOPA, we made a number of amino acid-PNA2 conjugates, including Arg-NHAc-PNA2, Glu-NHAc-PNA2, Met-NHAc-PNA2, Ser-NHAc-PNA2, Thr-NHAc-PNA2, Trp-NHAc-PNA2, and Tyr-NHAc-PNA2. No reaction was observed between these amino acid side chains and DOPA.

Among all the amino acids studied, only the α-amino group, ε-group, imidizole, and the thiol of cysteine were competent to attack DOPA ortho-quinone

and produced cross-linked products with retention times of 26 minutes. Cysteine was expceptionally reactive cross-linking with DOPA even the absence of periodate. Importantly, these experiments demonstrated that this reaction must be templated to occur to an appreciable extent.

Aliphatic 1,2-Diols Do Not Interfere with Cross-linking

Sodium periodate is a selective oxidant acting upon compounds having two hydroxyl groups attached to adjacent carbons. Periodate mediated oxidation of an aliphatic diol results in cleavage of carbon-carbon bond joining the alcohol to form dual aldehydes. Likewise, periodate mediated oxidation of ortho-dihydroxylphenol forms an electrophilic ortho-quinone intermediate[105, 106]. Given that carbohydrates make up at least 6% of the total cellular dry mass[107], we sought to answer the question of whether aliphatic diols and catechols compete for periodate mediated oxidation (**Figure 2.13)**.

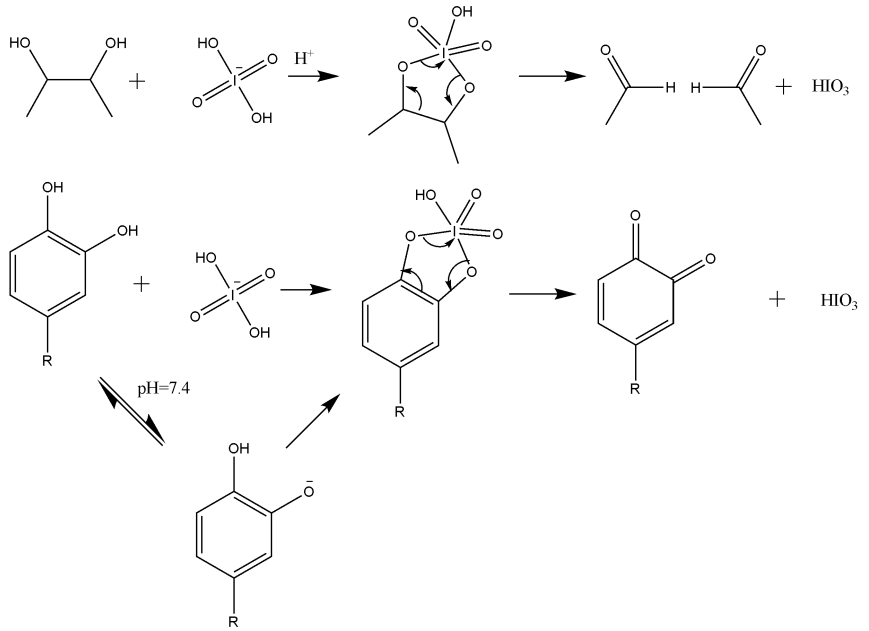


Figure 2.13: Periodate mediated oxidation of an aliphatic diol (top) and catechol moiety (bottom). Notice that periodate mediated cleavage of an aliphatic diol only occurs under acidic conditions.

To test this question we repeated our previous experiments on Gal80 protein (Gal80p) and Gal80-binding peptide (Gal80BP) in solutions containing ethylene glycol or lactose, compounds known to readily undergo periodate mediated oxidative cleavage under the appropriate conditions. Gal80BP (NH_2-YDQDMQNNTFDDLFWKEGHR-COOH), selected by phage display[108] against Gal80p binds to Gal80p with the equilibrium dissociate constant (K_D) of 300 nM.[109] The N-terminal tyrosine on Gal80BP was replaced with DOPA, and the peptide was extended at the N-terminal with two glycines and a biotinylated glutamic acid (Gal80-DOPA-BP) to allow cross-linked product to be detected by blotting with a NeutrAvidin-horse radish peroxidase (HRP) conjugate. A six-histidine (His_6)-tagged Gal80p was mixed with the Gal80-DOPA-BP in a reaction buffer containing 0.5 mM DTT and ethylene glycol or lactose. The mixtures were briefly treated with 1 mM sodium periodate for 30 s, and the reaction was quenched by a protein loading buffer containing 100 mM DTT. As shown in **Figure 2.14A**, at up to 0.1 M lactose concentrations (100 fold higher than sodium periodate concentration), there was no detectable reduction of Gal80p-Gal80BP cross-linking (lane 3-8, **Figure 2.14A**). Similarly, adding ethylene glycol had no effect on Gal80p-Gal80BP crossing (lane 3-8, **Figure 2.14B**). Because excess lactose and ethylene glycol did not quench periodate-triggered cross-linking reaction, we concluded that under these cross-linking conditions, periodate

mediated oxidation of DOPA to a resultant ortho-quinone dominates periodate mediated oxidative cleavage of aliphatic diols.

Periodate mediated oxidative cleavage of aliphatic diols requires acid catalysis[110]. Under acidic conditions the electrophilicity of the periodic species is increased to the point where-bye attack by an aliphatic hydroxyl group is permissible. Without acid catalysis however, this reaction should not occur to appreciably extent. In contrast to the high pKa of an aliphatic diol which is often greater than 14 (CRC Handbook of Chemistry and Physics), the pKa of the 1st catechol proton is estimated between 9.2-9.4 dependending on the context of the group[111]. Given this fact, a small amount of phenoxide ion should be present in solution at neutral pH. We predicted the phenoxide species facilitates the attack of the weakly electrophillic periodate catalyst and trigger the reaction at neutral pH. If this hypothesis is true then periodate mediated oxidation of DOPA should dominate in the neutral to basic pH range and aliphatic diols should only be able to compete for the periodate catalyst under acidic conditions. To test this hypothesis, we performed the His_6-Gal80p - Gal80-DOPA-BP cross-linking experiments at different in the biological pH range of 4-9. Cross-linking was observed at neutral or slightly alkali pH (lane 8-13, **Figure 2.14C**). Very little cross-linking product was detected below pH 6.5 (lane 3-7, **Figure 2.14C**). The slight reduction in cross-linking yield observed at a pH over 8.5 can be attributed to two different effects. At a pH above 9 it is possible that Gal80 may be starting to lose its structure, there-bye reducing the interaction affinity. Alternative the

basic charge of the bulk solvent may be masking the acidic charge on the activation domain interfering with the charge-charge interactions that drive Gal80/AD association.

Although these reactions appear to be very similar, the two reactions occur within different portions of the pH spectrum. The likely cause of this can be attributed to the distinct routes towards formation of the cyclic periodic ester intermediate. For DOPA, formation of the cyclic periodic ester intermediate is likely catalyzed by deprotonation of the catechol while aliphatic diols are capable of forming a cyclic periodic ester only under acidic conditions, or as a result of increased electrophilicity of the periodate ion. This result also suggest a phenoxide ion is necessary to catalyze the periodate induced oxidation of the ortho-quinone.

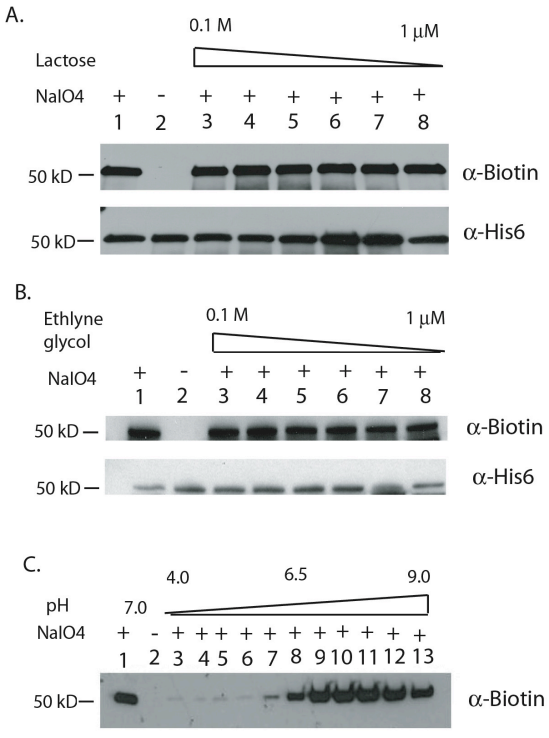


Figure 2.14*:* Chemical cross-linking between His6-Gal80p and Gal80-DOPA-BP in buffers containing lactose (A) and ethylene glycol (B), and in buffers at different pH (C).

From this work we were able to conclude which amino acid residues on proteins attacked the ortho-quinone. We also observed that this reaction was optimal at physiological pH and orthogonal to sugars. We were able to mimick peptide-protein and protein-protein by confining DOPA and individual amino acids in close range through annealing two complementary PNA strands. Among all amino acids and N-terminal amino group, we identified four residues capable to attack DOPA ortho quinone, namely, a-amino, Lys, His, and Cys. Cysteine was exceptional reactive, while the reactivity of the rest three could be described as a-amino > Lys > His.

Experimental Protocols

Materials for Gal80/Gal80 Binding peptide Cross-linking Experiments. Fmoc-DOPA(acetonid)-OH and Fmoc-Glu(biotinyl-PEG)-OH were purchased from Novabiochem (San Diego, CA). Fmoc-3,4-Dimethoxyphenylalanine was purchased from Synthetech (Albany, Oregon). All other Fmoc amino acid derivatives and resins were purchased from SynPep (Dublin, CA). Sodium Periodate and proteins within the protein mix were purchased from the Sigma-Aldrich Chemical Company (Milwaukee,WI) and used without further purification. Super Signal® West Pico and Super Signal® West Dura kits were purchased from Pierce (Rockford, IL). Immobilon Transfer-P membrane was purchased from Millipore (Billerica, MA). Crosslinking reactions were analyzed by SDS-PAGE and western blotting using standard protocols.

Peptide synthesis was performed on a Ranin Symphony 12 channel

FMOC synthesizer. Analytical gradient reversed phase HPLC was performed on a Waters Breeze HPLC system with a Vydac C18 analytical column. Flow rate was 1ml/min and detection was at 214 nm. Preparative HPLC was performed on the same instrument with a Vydac C18 preprative column. Flow rate was 10 ml/min. HPLC runs used linear gradients of 0.1% TFA and 90% acetonitrile plus 0.1% TFA. Mass spectrometry was performed on all synthesized peptides with the MALDI-Voyager DE Pro instrument.

19S and Gal80 purification

The Gal80 protein was purified as previously described via a His6 tag. The 19S proteasome was purified as previously described via the Cim5-Flag tagged subunit. Purified proteins were analyzed by SDS-PAGE using standard protocols.

Synthesis of Gal80 peptide and analogues:

All peptides were synthesized on rink amide resin for crosslinking experiments. Peptides were cleaved from the resin with a mixture of TFA:water:tri-isopropyl silane at a ratio of 27:1.5:1.5. Crude peptides were precipitated and washed with anhydrous ether before being dissolved in 8M Guanidinium-HCL and purified over preprative HPLC. The pure peptides were characterized as desired products by MALDI-MS.

Gal80-Binding Peptide (Tyrosine Analogue): Biotin-

EGGYDQDMQNNTFDDLFWKEGHR

Observed Mass: 3027.7 Expected Mass:3028.02

Gal80-Binding Peptide (DOPA Analogue): Biotin-

EGG(DOPA)DQDMQNNTFDDLFWKEGHR

Observed Mass: 3045.04 Expected Mass: 3045.42

Gal80-Binding Peptide (Dimethoxy Analogue):

Biotin-EGG(Dimethoxyphenylalanine)DQDMQNNTFDDLFWKEGHR

Observed Mass: 3273.14 Expected Mass: 3274.00

Control DOPA peptide: Biotin-KG(DOPA)AHNRLIYMQD

Observed Mass: 1852.00 Expected Mass: 1850.32

Note: DOPA containing peptides should be stored under Argon to prevent gradual atmospheric oxidation of the peptide.

Protein Mix

Ribonuclease A from bovine pancrease

Aldolase from rabbit muscle

_-galactosidase from bovine liver

_-chymotrypsinogen A from bovine pancrease

Ovalbumin form chicken egg

Trypsin inhibitor from soybean

Ubiquitin from bovine red blood cell

Carbonic anhydrase from bovine erythrocyte

Histone type II A from calf thymus

Bovine serum albumin

All proteins are dissolved in NE buffer separately at a concentration of 10mg/ml, and then mixed at equal volumes to make the final mixture solution.

NE Buffer: 20 mM HEPES, pH=7.9, 20% Glycerol, 100mM KCL,12.5 mM EDTA pH=8 .5mM DTT

Gal80 and 19S Cross-linking to the Gal80 Binding peptide

The protein target (Gal80 or 19S) at 1_M was incubated with the indicated peptide (1_M) for 5 minutes in NE Buffer at room temperature. After 5 minutes sodium periodate was added to a final concentration of 1 mM and the reaction was quenched 30 seconds later by the addition of 6X protein loading buffer containing 100mM DTT. The final volume was 36_. The samples were then run on an SDS-PAGE gradient gel (4-20%) and transferred to PVDF membrane for western blotting. The blot was blocked with 1X Uniblock (Aspen Bio Inc.) for 1 hour. After 1 hour, a Neutravidin-HRP stock (1mg/ml) was diluted 1:5000 in Uniblock and placed on the membrane at 4°C for 1 hour. Biotinylated proteins were detected by incubating the blot with the Super Signal reagent and processing according to standard protocols. After detection of the biotin signal, blots were stripped and re-probed with protein specific antibodies to determine which proteins had crosslinked with the DOPA peptide. Supplemental Figure 1 displays a series of blots demonstrating that the biotinylated bands overlay with iso-forms of Sug1 and Sug2 but not Cim5.

PNA Synthesis. Peptide nucleic acid (PNA) monomers Fmoc-T-OH, Fmoc-C(Bhoc)-OH, Fmoc-G(Bhoc)-OH, Fmoc-A(Bhoc)-OH, and Fmoc-AEEA-OH; carboxyl activators *O*-(7-azabenzo-triazol-1-yl)-1,1,3,3-tetramethyluronium

hexafluorophosphate (HATU), and 1-hydroxy-7-azabenzotriazole (HoAt) were purchased from Applied Biosystems. Diisopropylethylamine (DIPEA) and 2,6-lutidine were from Sigma-Aldrich. Fmoc-XAL-PEG-PS resin was from NOVAbiochem. Fmoc-8-amino-3,6-dioxaoctanoic acid (Fmoc-acetyl-ethyleneglycol-ethylamine, Fmoc-AEEA-OH) was from Peptides Inernational. Fmoc-DOPA(acetonid)-OH was from NOVAbiochem. All Fmoc amino acid monomers were from BACHEM and Advanced Chemtech.

PNA1-DOPA, AA-PNA2, and PNA3-DOPA were synthesized manually in a 25 mL reaction vessel (ChemGlass) at 5 mmol scale using standard fluorenylmethoxy-carbonyl (Fmoc) chemistry. Fmoc-XAL-PEG-PS resin was used as the solid phase. Fmoc deblocking was done by treating the resin in 20 % piperidine in dimethylformamide (DMF) (5 min × 2). Fmoc-DOPA(acetonid)-OH, Fmoc-AEEA-OH, and Fmoc-amino acid couplings were carried out in 2 mL of anhydrous dimethylformamide (DMF) solution containing 25 μmol of each monomer, 25 μmol of HATU, 25 μmol of HoAt, 6.4 μL of DIPEA, and 2.6 μL of 2,6-lutidine. For each PNA monomer coupling, 25 μmol of each monomer, 25 μmol of HATU, 25 μmol of HoAt, 6.4 μL of DIPEA, and 2.6 μL of 2,6-lutidine were dissolved in 2 mL of anhydrous *N*-methyl-2-pyrolidinone (NMP) and applied to the resin. A stream of nitrogen passed through the reaction solution from the bottom of the reaction vessel to mix the resin with the reaction mixture. Each coupling was allowed at room temperature for 30 min. When a PNA synthesis was finished, the resin was treated with 1 mL of 20 % acetic anhydride in DMF

containing 6.4 mL of DIPEA. The resin was thoroughly washed with DMF, dichloromethane, and methanol, dried in vacuum, and cleaved in 400 mL of cleavage mixture containing 80 % trifluoroacetic acid and 20 % *m*-cresol. The resin was then filtered with an Ultrafree-MC centrifugal filter unit (Millipore). To the filtrate was added 2 mL of cold ethyl ether. The white precipitant was collected by centrifugation and purified by HPLC on a C18 semi-preparative column. MALDI-TOF $[M+H]^+$: PNA1-DOPA calculated 2608.04, found 2608.89; PNA3-DOPA calculated 2488.02, found 2489.09; Cys-NHAc-PNA2 calculated 2419.97, found 2420.92; Ala-NHAc-PNA2 calculated 2388.00, found 2388.94; Lys-NHAc-PNA2 calculated 2445.06, found 2445.82; His-NHAc-PNA2 calculated 2454.02, found2455.28; Tyr-NHAc-PNA2 calculated 2480.03, found 2481.21; Arg-NHAc-PNA2 calculated 2473.07, found 2474.29; AlaNH_2-PNA2 calculated 2345.99, found 2346.95; Ser-NHAc-PNA2 calculated 2404.00, found 2404.85; Met-NHAc-PNA2 calculated 2448.00, found 2448.60; Trp-NHAc-PNA2 calculated 2503.04, found 2504.80; Thr-NHAc-PNA2 calculated 2418.01, found 2418.91; Glu-NHAc-PNA2 calculated 2446.01, found 2446.69.

Periodate Induced DOPA-amino acid Cross-link Reactions. All PNA samples in water, or 0.5 mM dithiothreitol (DTT) were prepared freshly before reactions and analyses. The concentrations of PNA1-DOPA and PNA3-DOPA stock solutions were controlled below 4 mM to achieve maximum DOPA activity (PNA-DOPA conjugates tended to aggregate in high concentration stock solutions). In a 200 mL reaction volume was added 1 mM PNA1-DOPA (or

PNA3-DOPA), 1.7 mM amino acid-PNA2, 10 mM sodium phosphate, 138 mM NaCl, 2.7 mM KCl, and 0.5 mM DTT. All reactions were carried out at pH 7.0 except that for His-NHAc-PNA2, was carried out at pH 6.4. The solution was heated at 99 °C for 1 min and slowly cooled to room temperature in 1 hr. 1 mM sodium periodate was added at room temperature for 90 s and the sample was injected in a Waters HPLC system on a C18 analytical column. The elution gradient was 1 % acetonitrile / min.

Periodate-Triggered Gal80p-Gal80BP Cross-linking in Lactose and Ethylene Glycol, and In Various pH. Biotinylated Gal80-DOPA-BP and His_6-tagged Gal80p were prepared as described before.[112] The peptide (1 mM) and the protein (1 mM) were incubated for 5 min at room temperature in reaction buffers containing 25 mM tris, pH 7.4, 150 mM NaCl, 25 mM $MgCl_2$, and 0.5 mM DTT, and lactose (or ethylene glycol) at various concentrations. For pH dependent cross-linking experiments, the reaction buffer contained 10 mM sodium phosphate, 137 mM NaCl, 2.7 mM KCl, and 0.5 mM DTT at various pH. Sodium periodate was added to a final concentration of 1mM. The reaction was quenched 30 s later by the addition of 6× protein-loading buffer containing 100 mM DTT. The final sample volume was 30 mL. The samples were then run on an SDS-PAGE gradient gel (4 - 20%). Peptide-protein cross-links were detected by western blotting as describe before.

Chapter 3

Utilization of DOPA Cross-linking in the Detection of Peptide/Protein Interactions

Introduction

Protein–protein interactions mediate most biological functions. Existing methodologies for studying protein–protein interactions are powerful, but nonetheless have important limitations. Immunoprecipitation or affinity purification coupled with mass spectrometry is useful for cataloguing proteins that associate into stable complexes. But proteins associated more loosely with these complexes or those which do so only under certain conditions may be lost. Furthermore, individual protein–protein interactions are not revealed by this technique. Yeast two-hybrid experiments are often used to map interaction networks, but require that the protein of interest be studied outside of the environment of its native complex. For several years, we have worked towards the development of new chemical cross-linking methods to address these methodological gaps. In particular, we have reported recently an efficient cross-linking reaction based on the periodate-mediated activation of 3,4-dihydroxyphenylalanine (DOPA).[112] The resultant ortho-quinone intermediate is an electrophile capable of forming adducts with histidine, cysteine, and lysine.[113, 114] Cross-linking of DOPA-containing molecules to closely associated proteins often occurs in high yield, but little or no coupling of non-associated proteins is observed even in complex solutions containing thousands of proteins. The DOPA residue can be incorporated into peptides and proteins in a straightforward

fashion and thus appears promising as a method for probing the interactions of biomolecules in the context of large protein assemblies. In this chapter, we report the application of this chemistry to an important, recently recognized problem in transcription enzymology, the identification of transcriptional activator-binding proteins in the proteasome.

Activation Domain Interaction with the 26S Proteasome

Transcriptional activators generally consist of a DNA-binding domain (DBD) responsible for promoter recognition and an activation domain (AD) that recruits the protein machinery necessary for transcription, including coactivators, the RNA polymerase II (polII) holoenzyme, and chromatin remodeling and modification complexes. However, the precise nature of these activation domain interactions with the transcription and chromatin remodeling machinery remain poorly understood. A host of potential activation domain binding partners have been reported in various in vitro experiments but the physiological relevance of many of these contacts is open to question[115-121].

Recent evidence indicates that the proteasome,[122, 123] is intimately involved in transcription. As mentioned in Chapter 2 we are particularly interested in how the components of the proteasome are functioning at the Gal4 promoter in concert with the Gal4 transcription factor.[78, 82, 122-126;87, 90]. Genetic evidence implicates Rpt4/Sug2 and Rpt6/Sug1 as the targets of the Gal4 AD within the 26S proteasome and/or the APIS complex[79-81, 83, 84, 127]. These data are consistent with biochemical results showing that a GST-Gal4 AD fusion protein binds *in vitro*

translated Sug1 and Sug2and that the Gal4 AD binds Sug2 in a two-hybrid-like assay in yeast. However, these interactions remain to be validated in the context of the native complex.

In this chapter, we employ periodate-triggered cross-linking as a tool to probe the interactions of synthetic, DOPA-containing AD peptides with the 2 MDa 26S proteasome. In addition to the 34 residue native Gal4 AD we also examine a 20 residue peptide that was selected by phage display for binding to Gal80 and shown subsequently to function as an AD when fused to the Gal4 DBD[75, 128, 129]. We demonstrate that the Sug1 and Sug2 proteins are indeed the direct targets of these peptides in the 26S proteasome, validating the previous biochemical and genetic data. These results show that periodate-triggered cross-linking of DOPA-containing molecules is an effective method for the characterization of protein receptors in large complexes.

Chemical cross-linking of the Gal4-activation domain to Gal80

Standard solid phase peptide synthesis was employed to create a biotinylated, DOPA-containing Gal4 AD construct for use in the cross-linking experiments. To validate the utility of this reagent, we initially tested cross-linking of the DOPA-Gal4 AD to the Gal80 transcriptional repressor (**Figure 3.1**). DOPA-Gal4 AD (1 μM) was mixed with purified, recombinant His6-Gal80 protein (1 μM) and 5 mM of $NaIO_4$ was then added. After a 30 second incubation followed by quenching, the proteins were separated by SDS-PAGE gel and the AD-containing products were detected by probing a blot with NeutrAvidin-HRP (NA-HRP) (**Figure 3.1B**). The observed signal matched the expected DOPA-Gal4

AD-His6-Gal80 cross-linked product.

We then employed the biotinylated, DOPA-Gal80 binding peptide (DOPA-Gal80 BP) in the same experiment. As expected, DOPA-Gal80 BP also cross-linked to His6-Gal80, albeit with lower yield than the DOPA-Gal4 AD peptide (Figure 3.1B) As a control, we then repeated the experiment with a small DOPA peptide containing a sequence not known to bind to Gal-80, i.e. DOPA-Control peptide (Control in Figure 3.1B). No cross-linking of this control peptide to His6-Gal80 was observed.

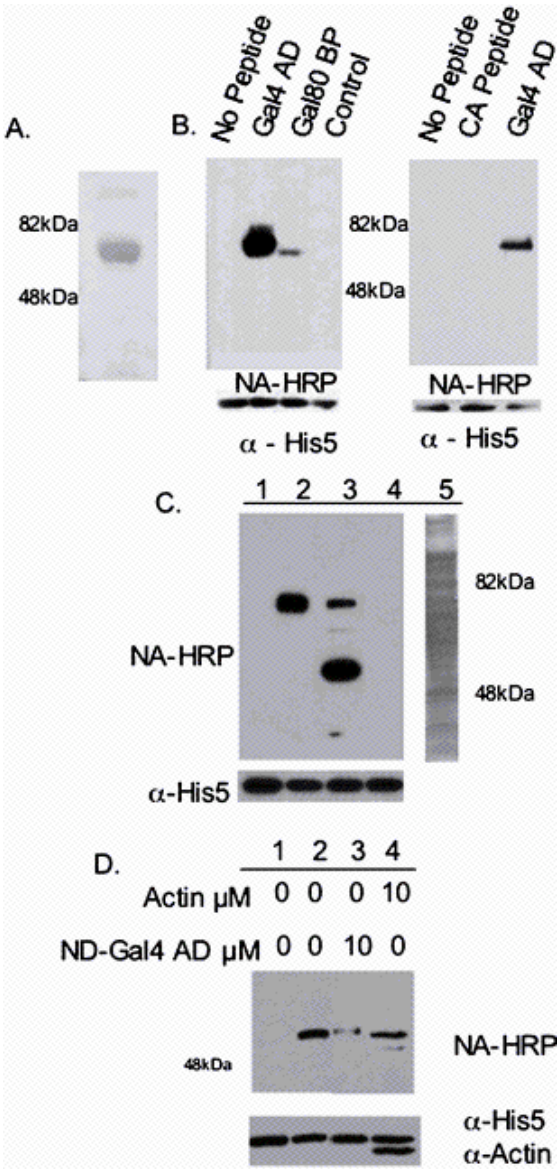


Figure 3.1: (A) Colloidal Blue stained SDS-PAGE of His6-Gal80. Apparent molecular weights are indicated. (B) NeutrAvidin (NA)-HRP was used to probe for the biotin tag incorporated in each peptide after SDS-PAGE and transfer to PVDF membrane. A major cross-linked product is detected between His6-Gal80 with the DOPA-Gal4 AD peptide and the DOPA-Gal80 BP. The Control peptide and the CA peptide show no product. The no peptide lane is a control showing the specificity of the NA-HRP. The -His5 blot shows the amount of the His6-Gal80 in each reaction. (C) Cross-linking between the DOPA-Gal4 AD peptide and His6-Gal80 protein in the presence of non-specific proteins. NA-HRP blot of a cross-linking reaction between 1 μM DOPA-Gal4 AD peptide and 1 μM His6-Gal80 showing the major product of the DOPA-Gal4 AD peptide His6-Gal80 reaction (lane 2). Lane 1 is the same reaction without addition of DOPA-Gal4 AD peptide. 1 μM Gal4 AD peptide added into a lysate of E. coli expressing the His6-Gal80 protein produces only 2 major cross-linked products (lane 3). No product is formed without the periodate trigger (lane 4). Colloidal Blue stained gel of E. coli lysate reaction mix (lane 5) used in lanes 3 and 4. The amount of the His6-Gal80 protein is indicated by the -His5 blot of the reaction mix. (D) A NA-HRP blot shows the results of a cross-linking reaction between the DOPA-Gal80 BP and His6-Gal80 with the presence of specific competitor proteins. 10 μM of ND-Gal4 AD peptide decreases the yield of cross-linked product (compare lane 2 and lane 3), but 10 μM of actin did not decrease yields of cross-linked products (compare lane 2 and lane 4). The amount of His6-Gal80 and actin are shown by the -His5 and -actin blots, respectively.

We also tried a control experiment using the CA DOPA peptide (CA peptide in **Figure 3.1B**), a peptide that cross-links very efficiently with the Arp3 protein in purified Arp2/3 complex (data shown in subsequent section). No cross-linked product was observed when His6-Gal80 and the CA peptide were treated with periodate (**Figure 3.1B**), demonstrating that specific binding is required for a cross-linked product to form.

To further probe the specificity of the DOPA-containing peptides to Gal80, we examined this reaction in the context of a crude E. coli extract containing hundreds of proteins, none of which should interact with the Gal4 AD or the Gal80 BP specifically (**Figure 3.1C**) Lanes 1 and 2 of **Figure 3.1C** simply repeat the experiments done in **Figure 3.1B**, purposes of comparison. In lane 3, DOPA-Gal4 AD was present at a final concentration of 1 μM in the presence of 200 μg of protein derived from an E. coli cell extract prepared from a strain expressing the His6-Gal80 protein (see lane 5 for a Commassie Blue-stained gel of this protein mixture). Upon addition of periodate, only two major biotin containing products were detected in (**Figure 3.1C**, lane 3), one with the apparent molecular mass expected for the Gal4 AD-His6-Gal80 cross-linked product and another unknown product that was of lower apparent mass. No cross-linking was detected in the absence of periodate (Lane 4).

We then employed a competitive cross-linking experiment to ask if both peptides share a common binding surface on Gal80 through competition cross-linking experiments. An excess of a non-DOPA containing Gal4 AD peptide (ND-

Gal4 AD) or, as a control, actin, were added to a solution containing 1 μM of DOPA-Gal80 BP with 1 μM of His6-Gal80 protein. 10 μM of ND-Gal4 AD decreased cross-linking between 1 μM of DOPA-Gal80 BP and His6-Gal80 as assayed by Western blotting (**Figure 3.1D** Lanes 2 vs. 3). Addition of actin had no effect on the cross-linking reaction (lane 4). The fact that the ND-Gal4 AD competed the DOPA-Gal80 BP-Gal80 crosslink, but the control protein actin did not, suggests that the native Gal4 AD and the Gal80 BP recognize overlapping surfaces of Gal80, though allosteric competition cannot be ruled out. We conclude from these data that the DOPA-containing Gal4 AD and Gal80 BP are competent substrates for periodate-mediated cross-linking reactions.

DOPA-Gal4 AD cross-linking to the 26S proteasome

With the utility of these DOPA-containing peptides demonstrated from the Gal80 cross-linking studies described above, we turned to using this technique to study AD/proteasome interactions. It has been demonstrated that a sub-complex of the proteasome containing all six ATPases and probably other factors must be recruited to many activated promoters through direct interactions with transactivators. We therefore used periodate-triggered cross-linking to attempt to identify the specific targets of the AD in the proteasome. This work was carried out in collaboration with a fellow graduate student in the lab, Chase Archer, who performed the two-dimensional mapping experiments and produced the purified 26S proteasome detailed in the subsequent sections.

1 μM of biotinylated, DOPA-containing peptide was equilibrated with 60 nM of 26S proteasome and the cross-linking reaction was initiated with periodate. Probing the blot with NA-HRP revealed one major cross-linked, Gal4 AD-containing product with an apparent mass of approximately 50 kDa (**Figure 3.2B**) At this exposure, no Gal80 BP containing product was apparent, but a longer exposure of the same blot (**Figure 3.2C**) shows that the Gal80 BP was incorporated into a cross-linked product of approximately the same apparent mass, but in much lower yield. The Control peptide did not cross-link to the 26S proteasome detectably.

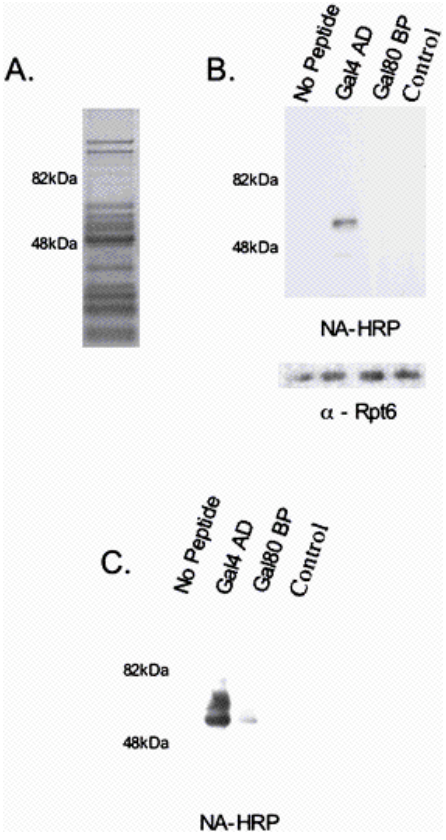


Figure 3.2: (A) Colloidal Blue stained SDS-PAGE of 26S proteasome. Apparent molecular weights are indicated. (B) NA-HRP blot of 1 μM peptides cross-linked to 60 nM 26S proteasome. The DOPA-Gal4 AD peptide gives a cross-linked product of approximately 50 kDa with the 26S proteasome. The DOPA-Gal80 BP or the Control peptide show no cross-linked product with the proteasome. The no peptide control shows the NA-HRP blot shows no cross reactivity to proteins contained in the 26S proteasome. The -Rpt5 blot indicates the amount of 26S proteasome loaded. (C) A longer exposure of 2(B) showing a cross-linked product between DOPA-Gal80 BP and the 26S proteasome at approximately 50 kDa. The DOPA-Gal80 BP cross-links several fold less well than the Gal4 AD peptide. The Control peptide does not show any cross-linked product.

Identification of the cross-linked products

The products of both the DOPA-Gal80-BP and DOPA-Gal4 AD-26S proteasome cross-linking reactions had electrophoretic mobilities consistent with coupling of the activating peptides to one of the proteasomal ATPases, all of which have similar masses. We used 2D gel electrophoresis and Western blotting to differentiate between these proteins and identify the major cross-linked species for both domains.

DOPA-Gal4 AD was cross-linked to the 26S proteasome. The proteins were then separated by isoelectric focusing (IEF) using a pH 3–10 gradient, and separated in a second dimension by SDS-PAGE. A NA-HRP probed blot (**Figure 3.3A**) showed one major cross-linked product was formed between DOPA-Gal4 AD and a 26S subunit of approximately 50 kDa with a pI of 5.5. We then stripped and reprobed the blot with several antibodies against components of the 26S proteasome. The images in **Figure 3.3B** were false-colored blue for the cross-linked product (NA-HRP) and red for the 26S subunit. Overlaying these images (**Figure 3.3B**) unequivocally identified Rpt4/Sug2 as the 26S subunit that was cross-linked to the Gal4 AD peptide.

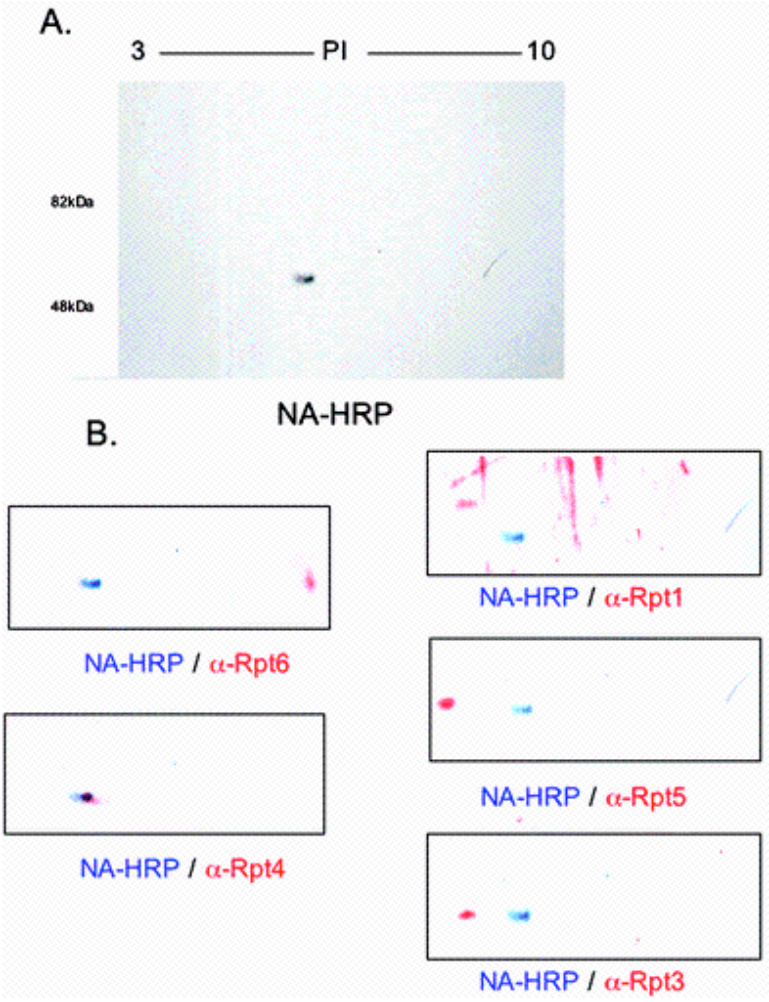


Figure 3.3:
Identification of Cross-linked Product of 26S proteasome and N-term DOPA-Gal4 AD peptide. (A) Cross-linking reaction of DOPA-Gal4AD peptide and 26S proteasome was separated by 2D gel electrophoresis and blotted with NeutrAvidin-HRP. The blot shows the major cross-linked product of the reaction with an apparent molecular weight of approximately 50 kDa and an isoelectric point of about 5.5. (B) Western blots of cross-linked product and 26S subunits from membrane in (A) were overlayed. The NeutrAvidin-HRP blot was false colored blue and 26S subunit blots were false colored red. The 26S subunit is indicated under each overlay. This analysis identified Rpt4/Sug2 as the major product of the 26S proteasome and N-term DOPA-Gal4 AD peptide cross-linking reactions.

The chemical cross-linking and 2D mapping experiment was repeated to identify the cross-linked product between Gal80 BP and the proteasome. The NA-HRP blot (**Fig. 3.4A**) showed a product with an apparent mass similar to that of the Gal4 AD-Rpt4 product, but a much more basic pI. Overlay of the cross-linked product (false-colored blue) and Western blots using antibodies against subunits of the 26S proteasome (false colored red) (**Fig. 3.4B**) identified Rpt6/Sug1 as the 26S subunit that forms a cross-linked product with Gal80 BP. No other proteasome subunits overlaid with the cross-linked products.

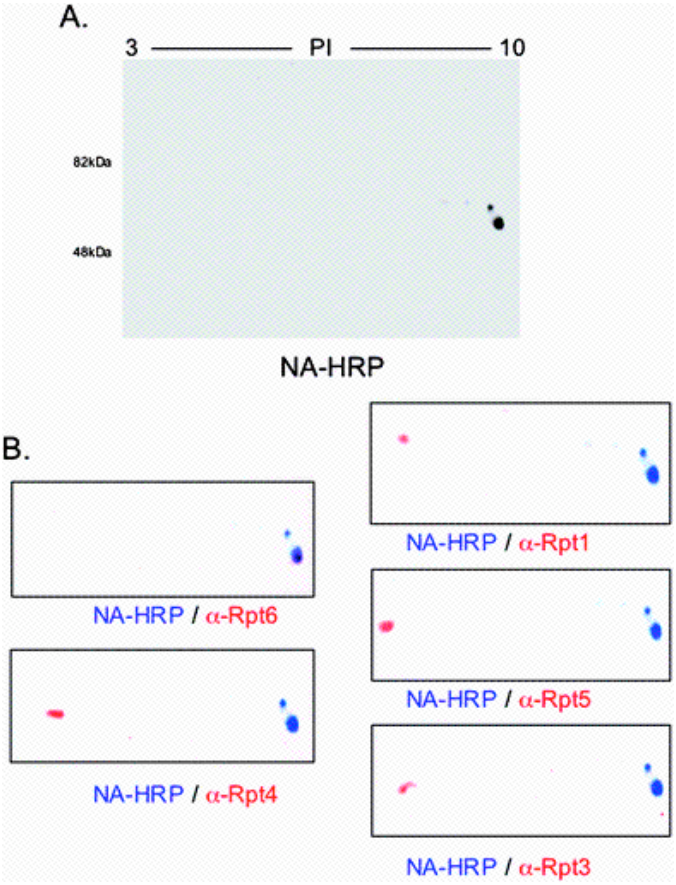


Figure 3.4: (A) Cross-linking reaction of DOPA-Gal80 BP and 26S proteasome separated by 2D gel electrophoresis and blotted with NA-HRP. The blot shows the major cross-linked product of the reaction with an apparent molecular weight of approximately 50 kDa and a isoelectric point of almost 10. (B) Western blots of cross-linked product and 26S subunits from membrane in (A) were overlayed. The false coloring and overlaying were done as in Fig. 3B. Rpt6/Sug1 was identified as the subunit cross-linked by the Gal80 BP.

The finding that the DOPA-Gal4 AD and DOPA-Gal80 BP cross-link to different proteasomal ATPases was surprising since both recognize the same surface of Gal80 and therefore presumably have similar protein-binding preferences. Since there is genetic and biochemical evidence that Rpt6/Sug1 is also recognized by the Gal4 AD, we considered the possibility that both DOPA-containing peptides might contact Rpt4 and Rpt6, but that only one of these contacts resulted in a cross-linked product due to the position of the DOPA residue in the peptide. To test this model, we synthesized a Gal4 AD derivative containing the DOPA residue at the C-terminus and repeated the cross-linking experiments with the 26S proteasome. The NA-HRP blot (**Figure 3.5A**) now showed a product with an approximate molecular weight of 50 kDa but with a much more basic pI than the cross-linked product created with the N-terminal

DOPA Gal4-AD. Overlay of the cross-linked product and Western blots using antibodies against subunits of the 26S proteasome (**Figure 3.5B**) identified Rpt6/Sug1 as the 26S subunit that forms a product with the C-terminal DOPA Gal4-AD.

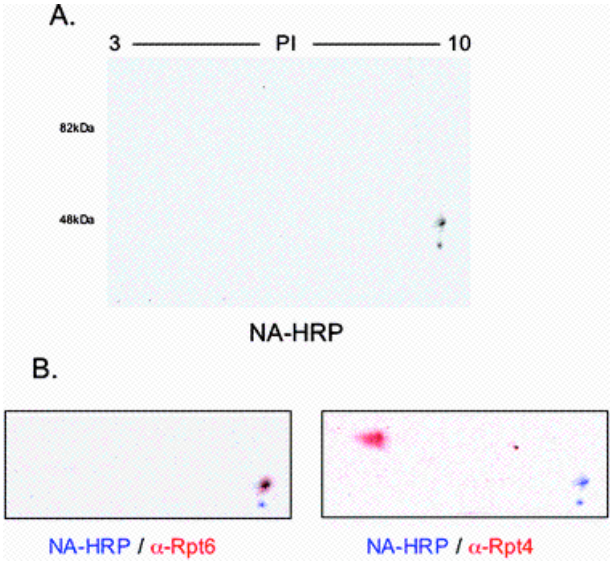


Figure 3.5 : (A) Cross-linking reaction of C-term DOPA-Gal4AD peptide and 26S proteasome was separated by 2D gel electrophoresis and blotted with NA-HRP. The blot shows the major cross-linked product of the reaction with an apparent molecular weight of approximately 50 kDa and an isoelectric point of about 10. (B) Western blots of cross-linked product and 26S subunits from membrane in (A) were overlayed. The NA-HRP blot was false colored blue and 26S subunit blots were false colored red. The 26S subunit is indicated under each overlay. This analysis identified Rpt6/Sug1 as the major product of the 26S proteasome and C-term DOPA Gal4 AD peptide cross-linking reactions.

To further confirm that the Gal80-BP and the Gal4-AD bind to the 26S proteasome in a similar fashion, we again turned to a competition experiment. A solution containing DOPA-Gal80 BP and 26S proteasome was equilibrated with either increasing amounts of the ND-Gal4 AD or actin. The NA-HRP blot (**Figure 3.6**) of the products formed after periodate-triggered cross-linking showed that increasing amounts of ND-Gal4 AD peptide decreased the signal intensity of cross-linked product (lanes 2–6). Addition of 20 μM actin however, had no affect on the formation of cross-linked product (lane 8). The Rpt6 and actin blots served as loading controls for the 26S proteasome and actin, respectively. Lane 1 is a

chemical cross-linking reaction without the Gal80 BP. These results are consistent with a model in which both peptides bind Rpt4 and Rpt6, though we cannot completely rule out the alternative model where each peptide competes with the binding of the other by an allosteric mechanism.

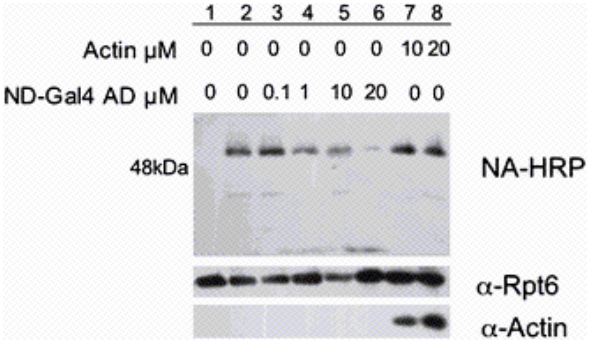


Figure 3.6: A NA-HRP blot shows the results of a cross-linking reaction between DOPA-Gal80 BP and the 26S proteasome in the presence of competitor proteins lacking a DOPA residue. Increasing amounts of ND-Gal4 AD peptide reduced the amount of DOPA-Gal80 BP-Rpt6/Sug1 cross-linked product (compare lane 2 and lane 6). Actin did not affect the cross-linking reaction (compare lane 2 and lane 8). Lane 1 is the reaction lacking the DOPA-Gal80 BP. -Rpt6 and -Actin Western blots serve as controls for the amount of 26S proteasome and actin, respectively.

We have used periodate-triggered cross-linking of DOPA-containing peptides that correspond to acidic transactivation domains to identify their direct binding partners in the proteasome. We anticipate that this technique will be of utility in identifying AD-binding proteins in other large transcription complexes, such as the Mediator, and more generally in the characterization of protein–protein interactions in the context large, multi-protein assemblages.

Observation of Inhibition of Activation Domain Cross-linking to the Proteasome Upon Mono-Ubiquitin addition *in Trans*

Work by the Kodadek and Johnston laboratories as well as others has put forth a “timer” model of transcriptional activation positing in part that mono-

ubiquitylation is a necessary licensing event for high-level gene expression at certain promoters. In support of this model, work by the Kodadek and Johnston groups have demonstrated that mono-ubiquitylation of the Gal4 transcription factor is required to initiate transcription of Gal4 inducible genes. Furthermore, it has been demonstrated that failure to mono-ubquitylate Gal4 results in the rapid removal of the transcription factor from the promoter (Ferdous et al. submitted, Archer et al. submitted). This removal or "stripping" of a transcription factor from the promoter has also been shown to be dependent upon the Gal4 activation domain interaction with Rpt6/Sug1 and Rpt4/Sug2 (Feredous et al. Archer et al.). Mono-ubiquitylation, therefore, has been predicted to disrupt the Gal4 activation domain 19S interaction. We investigated whether mono-ubiquitin effected the interaction and there-bye the observed cross-linking between the Gal4 activation domain and the 19S proteasome (**Figure 3.7**).

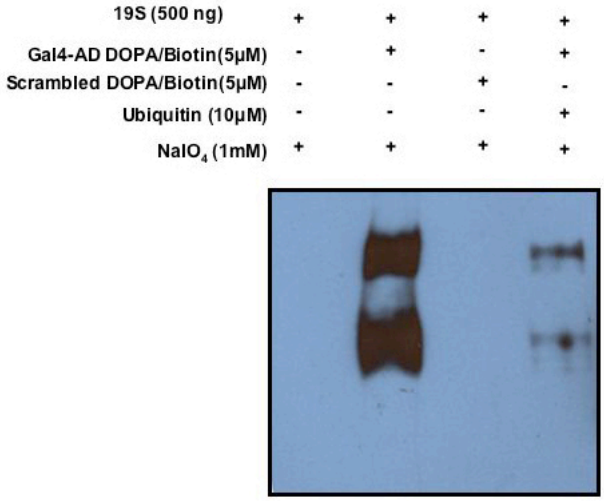


Figure 3.7: The 19S proteasome was crosslinked to the Gal4 activation domain in the presence and absence of monubiquitin (Lanes 4 and 2, respectively). Lanes 1 and 3 served as negative controls. Cross-linking was assayed by western blotting against biotin.

Similar to results presented in Chapter 2, cross-linking of the Gal4 activation domain solely to the 19S proteasome resulted in cross-links to Sug1, Sug2, and Rpn1/2 (Figure 3.7, Lane 2). The presence of 10μM mono-ubiquitin however, almost entirely abolished the cross-linking signal (Lane 4). Although preliminary, this result was important for two reasons. First, this result was suggestive of an as yet undefined mono-ubiquitin binding site within the 19S proteasome. Second, this result also suggested that DOPA mediated protein-protein cross-linking might be useful in observing allosteric interactions, a traditionally difficult phenomena to capture under physiological relevant conditions.

Identification of the Gal4/Gal80 Interaction Interface

Although the interaction between Gal4 and Gal80 has been extensively studied and utilized over the last twenty years no crystal or solution structure detailing the interaction interface between the two proteins has been published. Given that the DOPA chemical reaction was both specific and bioorthogonal we set out to determine whether it would be possible to map the activation domain of Gal4/Gal80 interface. To accomplish this goal we first sought to design the minimal activation domain sequence necessary for binding to Gal80 (**Figure 3.8**). Previous work by the Johnston and Kodadek labs indicated that the C-terminal 23 amino acid stretch of the Gal4 activation domain was necessary for activation domain activity in the native protein[80, 83]. Using the DOPA cross-linking assay we split the activation domain at methionine 853 and tested the relative efficiency with which these two AD fragments cross-linked to Gal80 relative to the full length activation domain and the Gal80 binding peptide. As demonstrated in **Figure 3.8**, the Gal4AD (853-875) cross-linked to Gal80 much more efficiently than the Gal80 binding peptide and at an equal efficiency to the full length AD, this is most likely due to the higher affinity of the Gal4 AD in comparison to the Gal80 BP. The Gal4^{D} peptide, however, did not cross-link to Gal80. We concluded from this result that the Gal4 AD (853-875) would be a suitable reagent for mapping the Gal80/Gal4 AD interface.

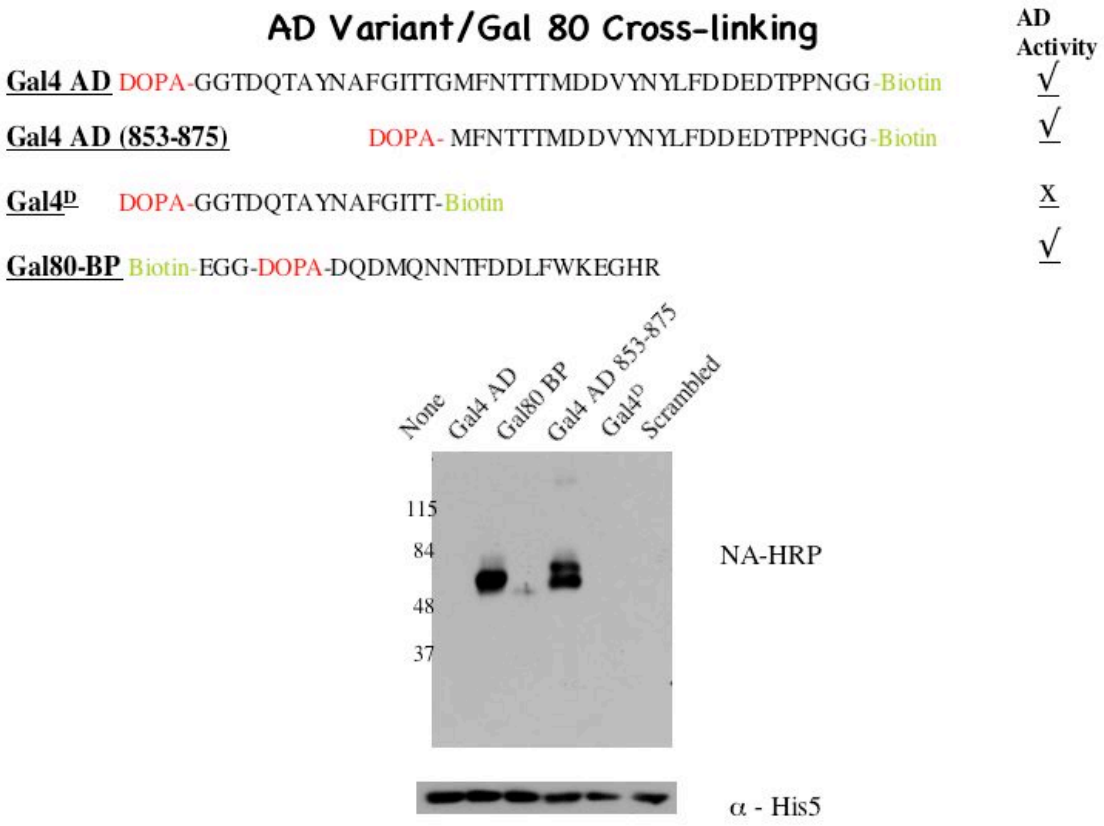


Figure 3.8: Test of the relative cross-linking efficiencies amongst different transcriptional activation domains.

Previous work has indicated that cross-linked peptides possess poor "flight" properties, i.e ionize poorly, in the mass spectrometer and are therefore difficult to detect. In addition, the accurate mass range of the Voyger Pro MALDI-TOF mass spectrometer, the workhorse instrument in our laboratory, spans from 800 to 4000 Daltons. For these reason we devised an analytical strategy to map the interaction interface between Gal80 and the Gal4 activation domain (**Figure 3.9**).

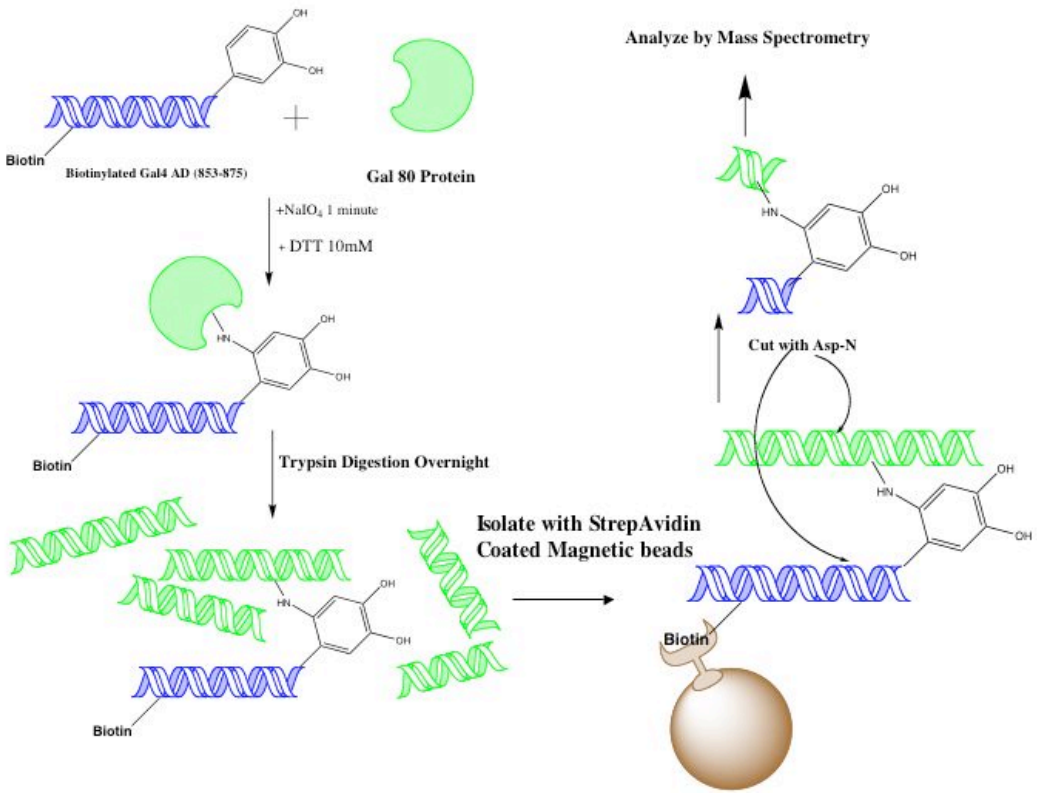


Figure 3.9: Strategy for Isolation of Gal4 AD-Gal 80 Interface.

Briefly, after cross-linking of the Gal4 AD (853-875) to Gal80, the cross-linking reaction was quenched and digested over-night with trypsin. It is important to note that the Gal4 AD (853-875) does not contain a tryptic digestion site and remains intact and ligated to a portion of Gal80. We then isolated the cross-linked product via a biotin tag on the C-terminus of the Cut AD using magnetic streptavidin coated beads. In order to both release the cross-linked product and to reduce the size of the branched peptide, Asp-N was added to release the N-terminal portion of the Gal4 AD 853-875 from magnetic streptavidin beads, Asp-N cuts 9 amino acids from the site of cross-linking. After release, the solution was processed for MALDI-TOF analysis according to standard

procedure[130]. Two sets of spectra, (+ and – sodium periodate), were collected on a Voyger Pro MALDI-TOF and representative spectra are presented in **Figure 3.10**. The spectral data were linked and searched for unique peaks in the cross-linked spectrum. Only one unique peak was found in the cross-linked spectrum with a mass of 2339.47 $[M+H]^+$. This mass was searched against a table of potential cross-linked products between Gal80 and the Gal4 AD while allowing for a mass error +/- 2 amu or a mass error of .08%. It was unclear to us what effects the branched structure of the peptide might have on the cross-linked peptide, hence the +/-2 amu mass error range. Only a single combination of products from the Gal80 and Gal4 AD tyrptic/Asp-N digestion possessed a calculated mass within the error range for the experiment. This product matched with a peptide that corresponded to amino acids 291-300 of Gal80 (**Figure 3.10** Blue Sequence).

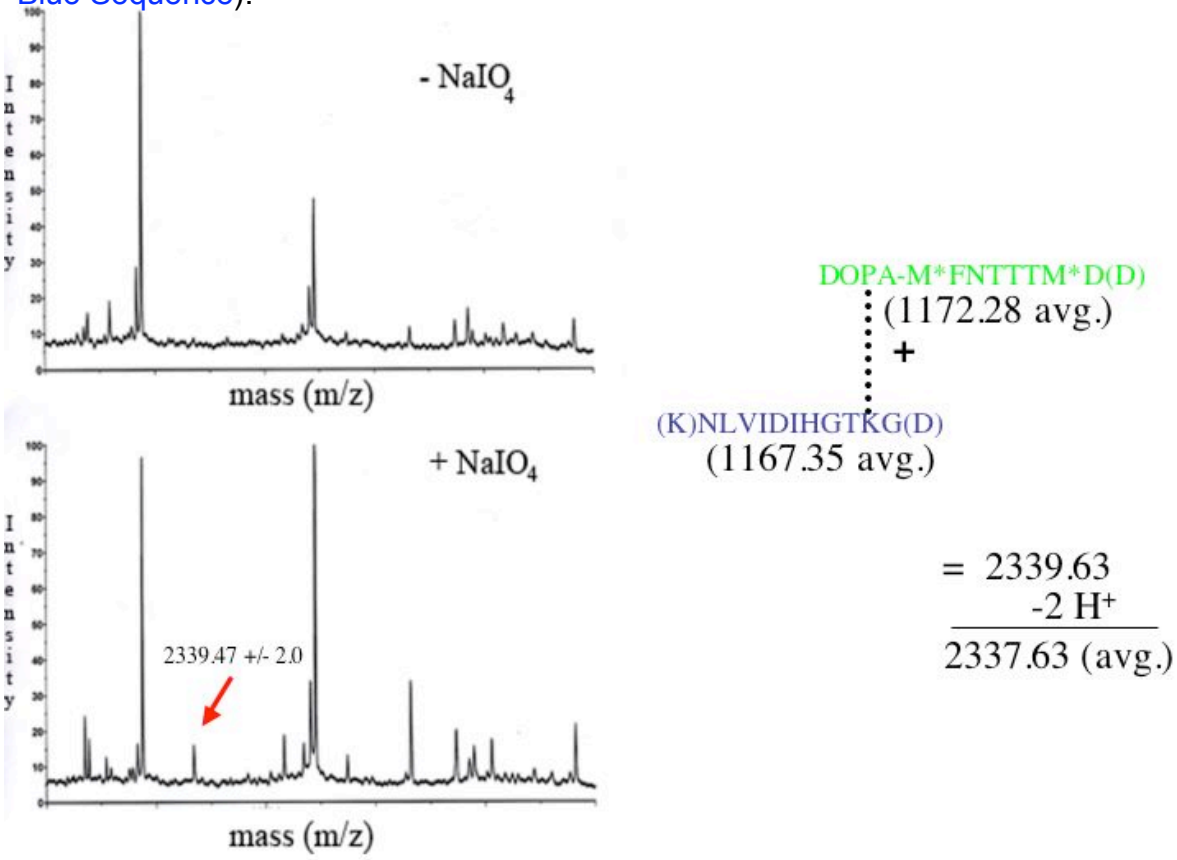


Figure 3.10: Mass spectra of Cut Gal4 Activation Domain/ Gal80 Cross-linking.Top Sepctrum-Control reaction without periodate. Bottom Sepctrum-Reaction triggered with periodate.

As expected, the signal from the cross-linked product was extremely weak despite the use of 25 micrograms of Gal80. Of this amount, at least 30% is cross-linked to the Gal4 Activation Domain (**Figure 3.11**).

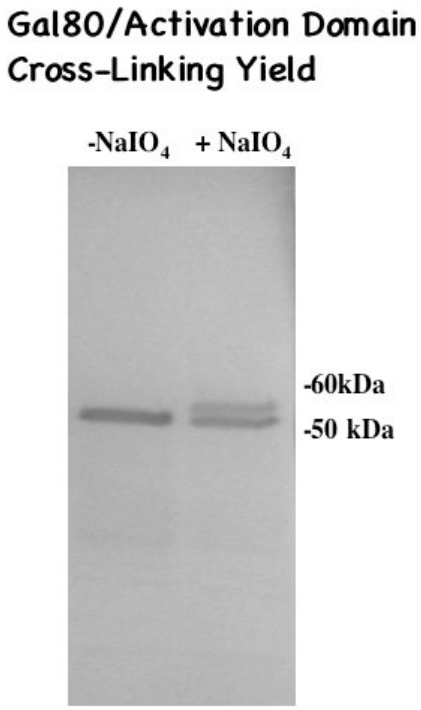


Figure 3.11: Gal80/Gal4 Activation Domain Cross-linking yield. 5 micrograms of Gal80 was added to transcription buffer and cross-linked to an equal molar amount of N-terminal DOPA containing Gal4 Activation Domain. After cross-linking the reactions was quenched and run on a 4-20% SDS-PAGE Gel and Stained with Collodial Blue Stain. Top-Band=Cross-linked Product

We were not therefore, able to fragment the product to confirm the sequence. However, these data are in agreement with mutagenesis studies[131-137] of the Gal80 protein which demonstrate that mutation of Histidine 293 to Alanine specifically ablates association with the Gal4 Activation domain, while the mutation of Glycine 301 to Arginine creates a Gal80 super-repressor[131-137], i.e.

increases the affinity between the Gal4 activation domain and the Gal80 repressor. In conclusion, mapping protein-protein interactions with the aid of chemical cross-linking is possible, however analytical advances are still needed to make the technique generally practical.

Creation of a Ubiquitin Cross-linker

Ubiquitin functions as a molecular protein tag in eukaryotic cells[138]. Thousands of proteins within the cell are involved in shuttling ubiquitin and ubiquitin chains onto and off of target proteins[139]. Unfortunately, however, it remains difficult to capture the transient interactions between ubiquitin and ubiquitin associated proteins[140]. Given this fact, it would be useful to design labeled forms of ubiquitin that could specifically and irreriversibly capture those transient interactions at physiologically relevant concentrations. We were particularly interested in designing a form of ubiquitin that could utilized to identify the aforementioned mono-ubiquitin binding site within the 19S proteasome.

To accomplish this task we incorporated the cross-linker 3,4-Dihydroxyphenylalanine (DOPA) into ubiquitin[112]. As mentioned in the previously, periodate mediated oxidation of DOPA to the reactive ortho-quinone creates an electrophile capable of forming adducts with cystiene, histidine, and lysine Michael addition. Although the reaction readily occurs within biological solutions, the DOPA containg polypeptide must be interacting with a binding partner to drive the cross-linking reaction when individual protein concentrations are in the nM to low mM range.

DOPA can be incorporated into proteins of interest either genetically or synthetically[92]. In order to incorporate DOPA and a biotin tag site specifically into the ubiquitin molecule we chose to construct ubiquitin via solid phase synthesis. We took advantage of recent advances in microwave assisted solid phase synthesis to increase product yield and decrease the time required to synthesize the synthetic form of ubiquitin.

Ubquitin contains a single, surface exposed tyrosine residue at position 59 in the polypeptide chain as pictured in **Figure 3.12**[141]. In order to achieve minimal structural perturbation we substituted Y59 for DOPA. We also tagged the N-terminus of ubiquitin with a biotinylated Glutamic acid providing an additional means of cross-linked product detection. This type of modification does not significantly effect ubiquitin structure or function[142, 143].

Standard solid phase FMOC chemistry was employed to synthesize the ubiquitin crosslinker with the following additions[144]. Each residue was double-coupled at 65°C under a microwave power of 25W for 5 minutes. The total time required for synthesis of the 77 amino acid ubiquitin cross-linker was slightly less than 2 days. After cleavage from the resin the molecule was precipitated with ice-cold ether, dried, and resuspended in 8M Guanidinium-HCl for subsequent purification on a C4 semi-prep reserve phase HPLC column. A fraction containing the pure product eluted at 47 minutes in a concentration of 37% ACN with .1% TFA. The overall yield of the reaction was 4%, supplying 8 milligrams of the desired product, a sufficient amount for biological experimentation.

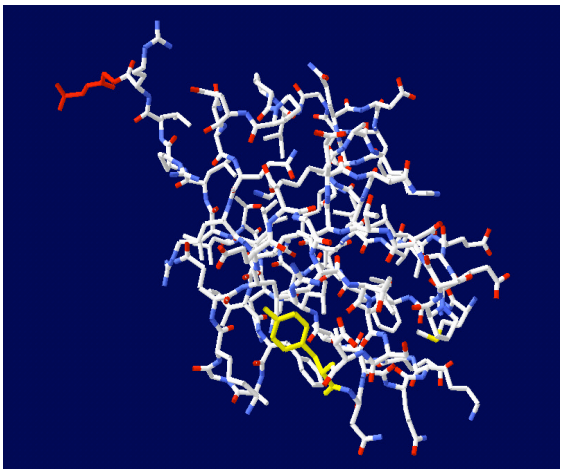


Figure 3.12: Tyrosine 59 of ubiquitin (yellow) was replaced with 3,4-Dihydroxyphenylalanine to create a form of ubiquitin competent to cross-link to binding partners. The C-terminus of ubiquitin is high-lighted in red.(10)

Given the extremely large size of this peptide we performed a trypsin digestion on the chemically synthesized ubiquitin and compared it to native ubiquitin. The expected masses are list in **Table 3.1** below followed by a figure of the linked spectra **Figure 3.13**.

Table 3.1: Expected and Observed Masses of Native and Synthetic Ubiquitin Peaks.
Z=Biotinylated Glutamic Acid
X=3,4 Dihydroxyphenylalanine

	Native Ubiquitin			**Synthetic Ubiquitin**	
Amino Acid Residue #	*Expected Mass*	*Observed Mass*	Amino Acid Residue #	*Expected Mass*	*Observed Mass*
1-6 MQIFVK	765.4	764.9	1-7 ZMQIFVK	1191.4	1191.01
49-54 QLEDGR	717.3	717	49-54 QLEDGR	717.3	717
34-42 EGIPPDQQR	1039.5	1039	34-42 EGIPPDQQR	1039.5	1039
64-72 ESTLHLVLR	1067.6	1067.1	64-72 ESTLHLVLR	1067.6	1067.1
55-63 TLSDYNIQK	1081.5	1082.1	55-63 TLSDXNIQK	1098.5	1099.4
30-42 IQDKEGIPPDQQR	1523.7	1523.2	30-42 IQDKEGIPPDQQR	1523.7	1523.2
12-27 TITLEVEPSDTI ENVKAK	1787.9	1787.2	12-27 TITLEVEPSDTI ENVKAK	1787.9	1787.2
1-76 Intact Native Ubiquitin	8565	8562	1-77 Intact Synthetic Ubiquitin	9009	9007

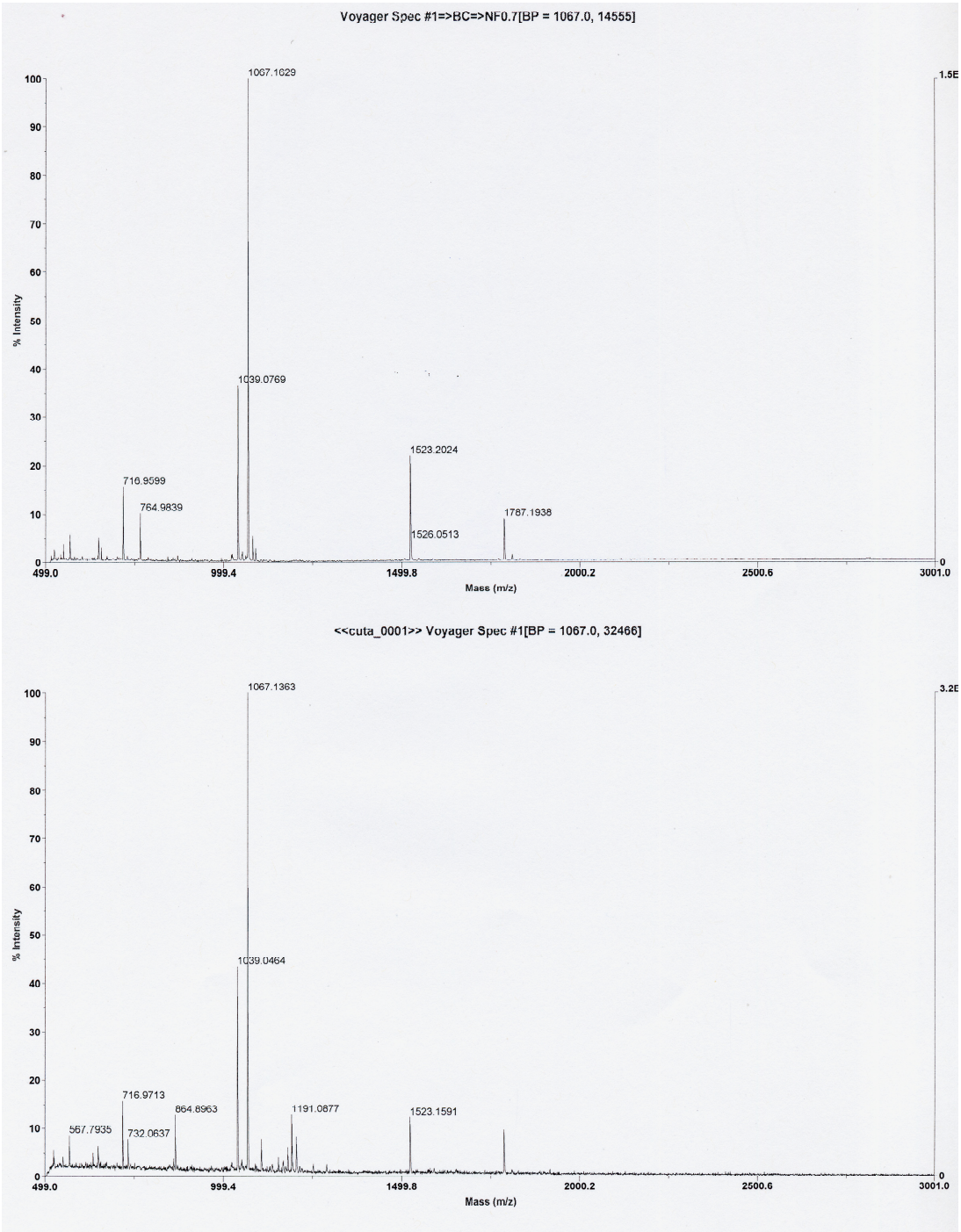


Figure 3.13: Mass Spectrometric Analysis of Synthetic Biotin-DOPA-Ubiquitin

Top Spectra: Native Ubiqutin Tryptic Digest
Bottom Spectra: Synthetic Ubiquitin Tryptic Digest

E2-25K is a class II ubiquitin conjugating enzyme composed of a ubiquitin association domain (UBA) at the C-terminus downstream from the enzyme's catalytic core[33, 145, 146]. Given the known binding interaction between ubiquitin and E2-25K we first tested whether the ubiquitin cross-linker could be coupled to the enzyme. The cross-linking reaction was performed at both 25 and 96°C in the presence of periodate and a GST tagged version E2-25K (**Figure 3.14**). Upon addition of periodate at 25 °C a significant amount of cross-linked product was observed on a Collodial Blue Stained gel. This product was recognized by both NeutrAvidin-HRP and a monoclonal antibody against ubiquitin. Importantly, under denaturing conditions at 96°C no crosslinking was observed between ubiquitin and GST-E2-25K highlighting the inability to form a cross-link product in the absence of a stable protein-protein interaction, an important chemical characteristic as this technique is advanced into more complex biological solutions.

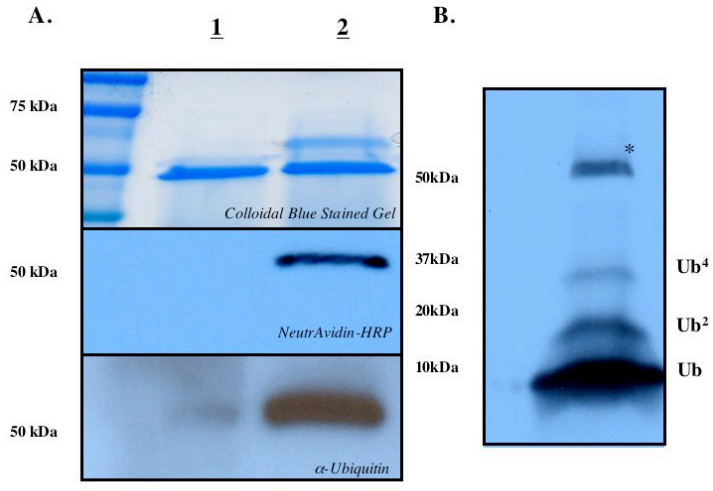


Figure 3.14: A.) Lane 1: 10mM DOPA-Ub was incubated with 1mM GST-E2-25K in cross-linking buffer with 1mM NaIO4 at 96°C. Lane 2: The same reaction was performed at 25°C. Cross-linked product was detected only in lane 2 (Top Panel). This cross-linked product was detectable with NeutrAvidin-HRP (Middle Panel) and α-ubiquitin (Bottom Panel). B.) In the presence of E2-25K, E1, and ATP DOPA-Ub is capable of being ligated into tetrameric ubiquitin chains (Ub4).

To further test the utility of the ubiquitin cross-linker we compared the cross-linking of synthetic ubiquitin to GST-E2-25K and the abundant cytoskeletal protein actin in Figure 3.15A. In the absence of periodate no cross-linking was observed for GST-E2-25K at a ubiqutin cross-linker concentration of 10mM (Lane 1). However, as shown previously a very strong biotinylated cross-linked product was apparent at the same ubiquitin concentration in the presence of GST-E2-25K and periodate (Lane 5). Importantly no cross-linked product was detected between actin and the ubiquitin cross-linker at any concentration (Lanes 6-9). Western blots against GST and actin were performed as loading controls. The E2-25K/ubiquitin cross-link was also detectable in the presence of E.Coli extract (Figure 3.15B Lanes 1 and 2). This interaction was competed away with exogenous ubiqutin (Lane 3).

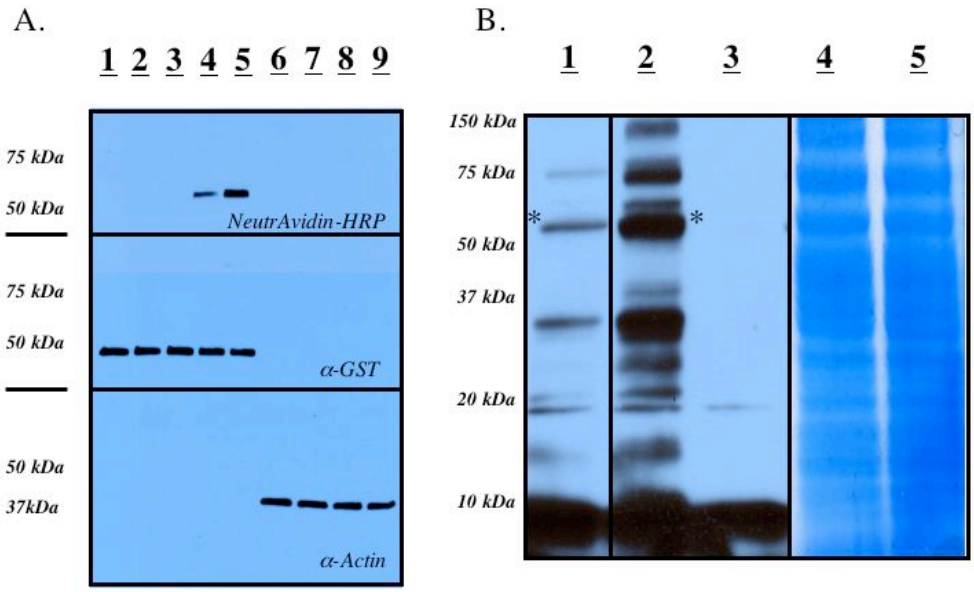


Figure 3.15: A.) DOPA-Ub was incubated at various concentrations with either GST-E2-25K 1μM (Lanes 1-5) or Actin 1μM (Lanes 6-9). Cross-links to GST-E2-25K were not detected upon incubation with 10nM and 100nM DOPA-Ub (Lanes 2 and 3). However cross-linked product was detected at 1μM and 10μM (Lanes 4 and 5). Cross-linking to actin was not observed (Lanes 6-9). B.) Cross-linking of DOPA-Ub (10μM) to GST-E2-25K (250ng) in E.Coli lysate (150μg). Lane 1. 30s exposure of Neutravidin-HRP blot. Lane 2. 5 minute exposure of NeutrAvidin-HRP blot. Lane 3. Competition of the cross-linking reaction with 200μM ubiquitin in E.coli lysate. Lanes 4: 150μg of E.Coli lysate + GST-E2-25K. Lane 5: 150μg of E.Coli lysate without GST-E2-25K.

In summary we have developed a novel class of capture agents for enzymes involved in ubiquitination and deubiquitination through recent advances in solid-phase chemistry and chemical cross-linking. Upon treatment with sodium periodate the ubiquitin molecule cross-links rapidly (approximately 10s) to a known binding partner at physiological concentrations in a biological buffer containing 1mM DTT. In the future this novel class of reagents should be extremely useful in identifying enzymes responsible for adding and removing ubiquitin and ubiquitin like molecules at the protein surface. Unfortunately, this form of ubiquitin did not cross-link detectably to the 26 or 19S proteasome. As explained in the next section, a negative result from a cross-linking experiment is not informative. However, it is possible to speculate on the reasons for the lack of cross-linking between ubiquitin and the proteasome. One reason for this lack of results could be that the DOPA residue is not placed at the interface between ubiquitin and the 26S and must be moved to pick up this particular interaction. Alternative, the face containing the DOPA residue may not interact with the proteasome. DOPA positional effects on the variability of cross-linking yields is considered in the next section.

Context Dependent Cross-linking Efficiency between Peptides and Proteins

Cross-linking between a binding partner and a molecule of interest containing DOPA should occur more or less efficiently depending on the proximity of the DOPA residue to reactive nucleophiles on the surface of a protein binding partner. To test this context dependent hypothesis we inserted the DOPA residue at three separate locations within the Src kinase substrate peptide $p34^{cdc2(aa\ 6\text{-}20)}$ and tested the ability of the three peptide analogues to crosslink to a His-6 tagged Src kinase (**Figure 3.16**)[147]. It is important to note that we did not test whether Src could phosphorylate this peptide sequence. Upon incubation with His-6 Src kinase (30nM) and periodate treatment *Src Substrate 1* and *Src Substrate 2* failed to crosslink to the kinase at the indicated concentrations (Lanes 1 & 2). Substitution of the endogenous tyrosine phosphorylation site with DOPA however did result in a crosslinked product (Lane 3). Importantly, *Src Substrate 3* did not crosslink to the Syk family tyrosine kinase Zap-70 (Lane 4) or Src at 96°C (Lane 5). Taken together these results indicate that although an interaction event is necessary to initiate cross-linking, the absence of crosslinked product does not rule out an association between two proteins. The site of DOPA incorporation within a molecule of interest and that location's positional relationship to the nucleophilic side-chains on the surface of a binding partner are criticial determinates in obtaining crosslinked product. In the future, the ability to observe a binding event between a kinase and substrate at relatively low concentrations independent of activity may prove useful in

screening for or validating pharmaceutical agents that are thought to bind in the substrate binding pocket of a kinase.

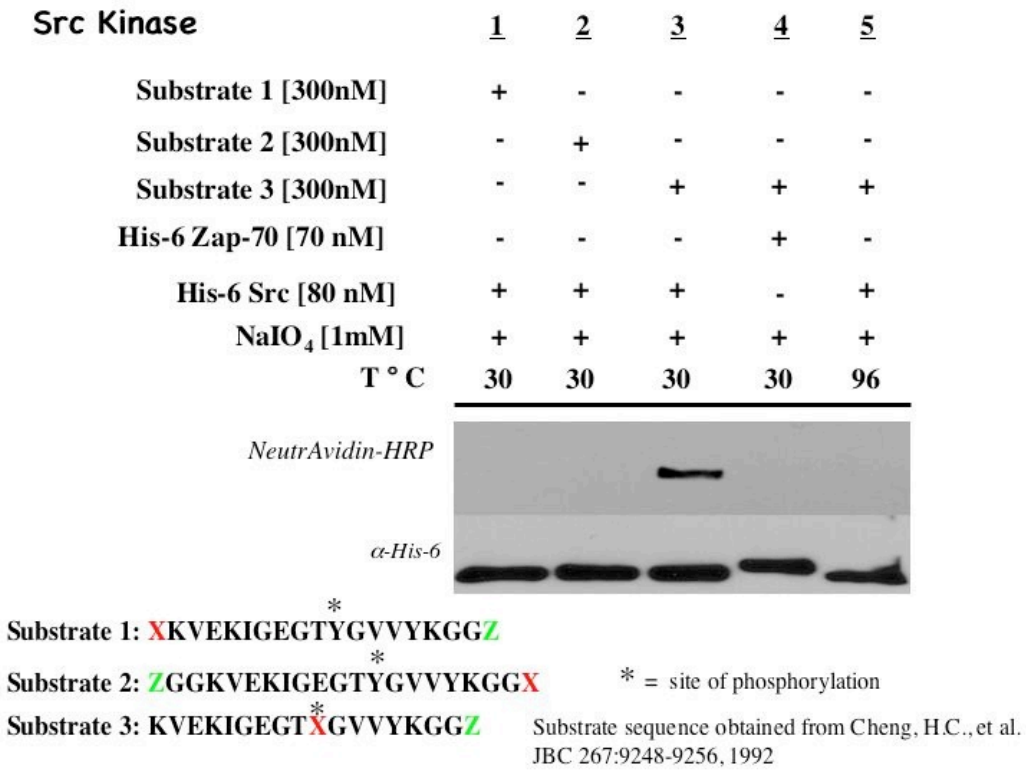


Figure 3.16: Reactions were performed in a total volume of 40 microliters (kinase Reactions buffer): 100mM Tris-HCL, pH=7.2, 125mM MgCL2, 25mM MnCl2, 2mM, EGTA, 1mM DTT. All other conditions were standard, 30 seconds of cross-linking, quench with 6X protein loading buffer. Samples were run on a 4-20% gradient gel (Bio-rad) and transferred to PVDF for western blotting.

Mapping the Interaction between the Arp 2/3 Complex and the N-terminus of the CA peptide.

The interaction between the Arp 2/3 complex and the CA peptide was another system in which we again attempted to map the interaction partner of a peptide within a multi-protein complex. The Arp2/3 complex is a constitutive 240 kDa assembly of seven polypeptides: actin-related proteins 2 and 3 and five novel proteins [referred to as Arp complex (ARPC) 1 though 5] (**Figure 3.17**).

The purified Arp2/3 complex promotes nucleation poorly, but its activity can be significantly enhanced by the cooperative effects of preexisting filaments and nucleation promoting factors (NPFs), such as members of the Wiskott-Aldrich syndrome protein family[148-150]. Proteins in the Wiskott-Aldrich syndrome protein (WASP) family activate the Arp2/3 complex through a conserved C-terminal element termed the VCA domain. The VCA can be divided into three distinct functional segments: a verprolin-homology sequence, which binds actin monomers and may recruit them to the Arp2-Arp3 dimer, a hydrophobic central region required for Arp2/3 activation, and an acidic tail required for high-affinity binding of VCA to the Arp2/3 complex. In efforts to map activator contact sites on the Arp2/3 complex, VCA peptides have been chemically cross-linked to Arp3, Arp2, ARPC1, and ARPC3 using the zero-length cross-linking reagent 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC). C-Terminally truncated VCA peptides retain the ability to cross-link to Arp2 and ARPC1, but cross-linking to Arp3 is severely decreased, suggesting that Arp3 may contain the binding site for the acidic region[148-150]. However, recent direct binding data suggest that the VCA C-terminus may contact the ARPC1 subunit. The relative positions of the N-terminal V and C portions of VCA have not yet been examined experimentally. The position of the C region is particularly important because its contacts are necessary to drive an activation step during the nucleation process.

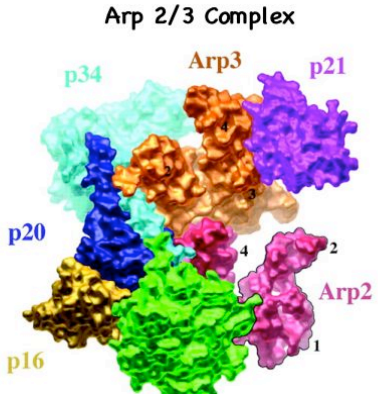
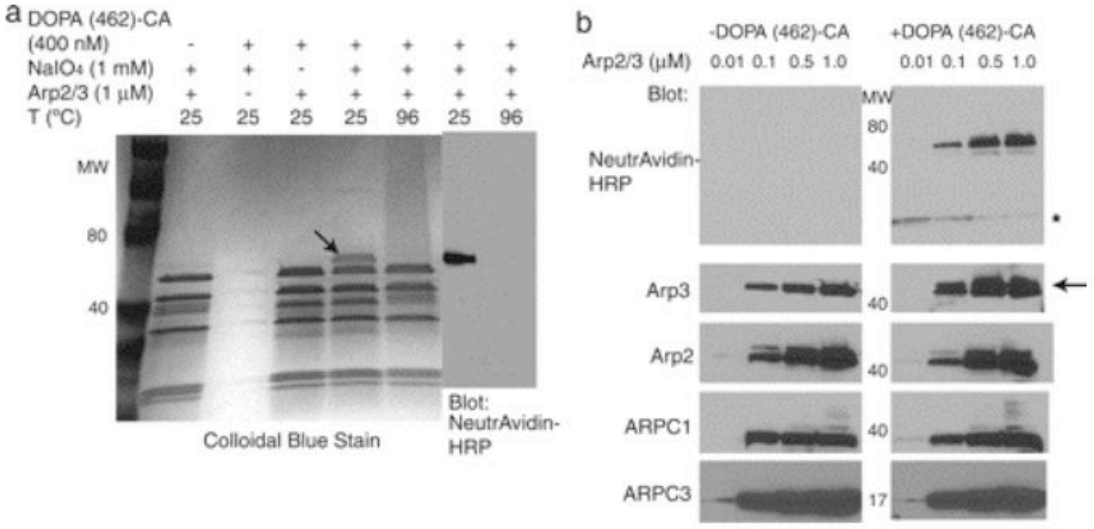


Figure 3.17: Crystal structure of the Arp 2/3 Complex

We sought to identify the Arp2/3 subunit that directly interacts with the N-terminus of the CA peptide. To accomplish this goal, we inserted the site-specific chemical cross-linker 3,4-dihydroxyphenylalanine (DOPA) into position 462 of a biotinylated CA peptide [to give DOPA(462)-CA peptide]. In **Figure 3.18A** the DOPA(462)-CA peptide was added to a final concentration of 400 nM in a solution containing 1 M Arp2/3 complex. After equilibrating the solution for 10 min, sodium periodate was added to a final concentration of 1 mM, and cross-linking was allowed to proceed for 30 s before the reaction was quenched. A strong cross-linked product was detected in the presence of periodate. This migrated slightly higher than Arp3 on an SDS-PAGE gel (**Figure 3.18A**, lane 4). The product was not produced under denaturing conditions at 96 °C (lane 5), in the absence of periodate (lane 3), in the absence of Arp2/3 (lane 2), or in the absence of the DOPA(462)-CA peptide. Western blotting with NeutrAvidin-horseradish peroxidase (HRP) revealed that the cross-linked product was biotinylated (lane 6) while mass spectrometric analysis of the cross-linked product indicated that the band of interest contained Arp3 (data not shown). To further confirm the existence of a direct interaction between Arp3 and the N-

terminus of the CA peptide, the DOPA(462)-CA peptide was incubated at a final concentration of 20 nM with increasing concentrations of the Arp2/3 complex, cross-linked, and analyzed by Western blotting against biotin and several proteins within the complex (**Figure 3.18B**). In the presence of the DOPA(462)-CA peptide a biotinylated product was detected when incubated with 100 nM, 500 nM, and 1 M Arp2/3 complex (top panel). As expected, the cross-linked product was recognized with α-Arp3 but not α-Arp2, α-ARPC1, or α-ARPC3, confirming that the DOPA(462)-CA peptide cross-links specifically with Arp3.

Arp 2/3 CA-peptide Interaction*



Collaboration with Rosen Lab

Figure 3.18: (a) Periodate-mediated cross-linking of the DOPA(462)-CA peptide to the bovine Arp2/3 complex. Cross-linking reactions were separated on SDS-PAGE gels and either stained with Collodial Blue (lanes 1-5) or transferred to a PVDF membrane and probed for biotinylated products (lanes 6 and 7). The arrow denotes the cross-linked product. (b) Periodate-mediated cross-linking of the DOPA(462)-CA peptide to the bovine Arp 2/3 complex. The DOPA(462)-CA peptide was present at 20 nM and incubated with 10 nM, 100 nM, 500 nM, and 1 M of the Arp 2/3 complex. Cross-linked products were separated on SDS-PAGE gels and detected by Western blotting for the subunits as indicated. The star denotes free peptide. The arrow indicates product cross-linked to Arp3.

Despite the presence of crystal and NMR structures it was still not possible to map the interaction between the CA-peptide and the Arp 2/3 complex. Insertion of DOPA at the N-terminus of this peptide provided a convenient means to identify Arp3 as the interaction partner of this peptide. This interaction was mapped at under physiological relevant concentrations and provided another example of the utility of DOPA chemistry in mapping peptide-protein interactions occurring within a multi-protein complex.

Conclusions

It is often desirable to identify the protein target of a bio-active peptide. In this chapter we have proven that DOPA cross-linking is an effective solution to this problem. Initially, we inserted DOPA into the Gal4 activatin domain and after cross-linking we determined the receptors to be Rpt4 and Rpt6. The cross-linking methodology was able to even provide low resolution information detailing how the activation domain bound into the protein pocket created by Rpt4 and Rpt6. Specifically, the N-terminus of the activation domain bound to Rpt4 while the C-terminal bound to Rpt6. Competition cross-linking experiments demonstrated that the Gal4 activation domain shared a binding interface with an artificial activation domain, the Gal80 binding peptide. In agreement with aforementioned genetic experiments, this set of results indicated that a general acidic activation domain binding pocket exist within the base of the 19S proteasome. We next demonstrated that mono-ubiquitin was able to modulate this interaction indicating that perhaps the 19S base contains an unidentified

binding pocket for mono-ubiquitin. In an attempt to map that binding pocket, we next created a synthetic form of ubiquitin that contained DOPA at position 59 in place of the native tyrosine. Although, this form of ubiquitin cross-linked effectively to the ubiquitin binding protein E2-25K, this form of ubiquitin did not cross-link to the 19S. This result provided an impetus to create alternative methods for DOPA delivery that will be discussed in Chapter 4. Finally, DOPA cross-linking chemistry was used to map the CA-peptide binding partner within the Arp 2/3 complex. Traditional technologies for identifying how a bio-active peptide fits into a multi-protein complex had failed to identify where the N-terminus of the CA-peptide bound within the Arp 2/3 complex. Utilizing the DOPA cross-linking methodology, we were able to discover that Arp-3 associated with the N-terminus of the CA-peptide.

Experimental methods

Fmoc-DOPA(acetonid)-OH and Fmoc-Glu(biotinyl-PEG)-OH were purchased from Novabiochem (San Diego, CA). All other Fmoc amino acid derivatives and resins were purchased from SynPep (Dublin, CA). Sodium Periodate and DOPAC (3,4-dihydroxyphenylacetic acid) were purchased from the Sigma-Aldrich Chemical Company (Milwaukee, WI). Super Signal® West Pico and Super Signal® West Dura kits were purchased from Pierce (Rockford, IL). Immobilon Transfer-P membrane was purchased from Millipore (Billerica, MA). Cross-linking reactions were analyzed by SDS-PAGE and Western blotting using standard protocols. Peptide synthesis was performed on a Ranin Symphony 12 channel FMOC synthesizer. Analytical gradient reversed phase HPLC was performed on a Waters Breeze HPLC system with a Vydac C18 analytical column. Flow rate was 1 ml min^{-1} and detection was at 214 nm. Preparative HPLC was performed on the same instrument with a Vydac C18 preparative column. Flow rate was 10 ml min^{-1}. HPLC runs used linear gradients of 0.1% TFA and 90% acetonitrile plus 0.1% TFA. Mass spectrometry was performed on all synthesized peptides with the MALDI-Voyager DE Pro instrument.

Peptide synthesis

All peptides were synthesized on rink amide resin for cross-linking experiments. Importantly, each residue of the Gal4 AD was double-coupled to drive up the final product yield. Furthermore, DOPAC was added as an N-terminal cap to the end of the N-term DOPA Gal4 AD peptide as opposed to DOPA. Peptides were

cleaved from the resin with a mixture of TFA water tri-isopropyl silane at a ratio of 27 1.5 1.5. Crude peptides were precipitated and washed with anhydrous ether before being dissolved in 8 M Guanidinium-HCL and purified over preprative HPLC. The pure peptides were characterized as desired products by MALDI-MS. After purification the peptides were resuspended in Buffer A (25 mM Tris pH 7.4, 150 mM NaCl, 5 mM $MgCl_2$, 10% Glycerol) containing 1 mM DTT aliquoted and stored at –80 °C.__

N-term Dopa Gal-4 AD

DOPAC-TDQTAYNAFGITTGMFNTTTMDDVYNLPDDEDTPPNPKKEGGEBiotin

Observed Mass: 5287.91 Predicted Mass: 5289.33

C-term Dopa Gal-4 AD

Biotin-GGTDQTAYNAFGITTGMFNTTTMDDVYNYLFDDEDTPPN-DOPA

Observed Mass: 5054.6 Predicted Mass: 5050.61

Gal-4 ND AD

TDQTAYNAFGITTGMFNTTTMDDVYNLPDDEDTPPNPKKEGGE-Biotin

Observed Mass: 5136.6 Predicted Mass: 5138.33

Gal80-binding peptide

Biotin-EGG(DOPA)DQDMQNNTFDDLFWKEGHR

Observed Mass: 3045.04 Predicted Mass: 3045.42

Control DOPA peptide

Biotin-KG(DOPA)AHNRLIYMQD

Observed Mass: 1852.00 Predicted Mass: 1850.32

CA DOPA peptide

P(DOPA)PTSGIVGALMEVMQKRSKAIHSSDEDEDEDDEDEFEDDDEWEDGG-Biotin

Observed Mass: 5962.49 Predicted Mass: 5963.2

Protein purification

His6-Gal80 protein was purified as previously described from *E. coli* strain BL21. Affinity-purified 26S proteasome was purified from *S. cerevisiae* strain RJD1144 (*MATa* his3200 leu2-3, 112 lys2-801 trp63 PRE1 FLAG::Ylplac211 (URA3) as described previously with modifications. Actin was purchased from Sigma.

Antibodies and Western blotting

26S proteasome antibodies were produced in rabbits. Mouse -Actin (Sigma), NeutrAvidin-HRP (Pierce), mouse -His5 (Qiagen), goat -rabbit-HRP (BioRad), and goat -mouse-HRP (BioRad) were purchased. NeutrAvidin-HRP blots were done as described previously.[1] All other antibodies were used in 4% milk/TBST using standard Western blotting procedures at the following dilutions: rabbit -Rpt1 (1 2500), rabbit Rpt3 (1 5000), rabbit -Rpt4 (1 5000), rabbit -Rpt5 (1 5000), rabbit -Rpt6 (1 4000), mouse -His5 (1 2500), mouse -Actin (1 1000), goat -rabbit-HRP (1 5000), and goat -mouse-HRP (1 5000).

Cross-linking reactions

All cross-linking reactions were done in 1/2 Nuclear Extract Buffer (1/2 NEB) (10 mM HEPES (pH 7.9), 10% Glycerol, 50 mM KCl, 6.25 mM $MgCl_2$, 0.1 mM EDTA, 1 mM DTT. 30 μL reactions contained either 26S proteasome at 60 nM or His6-Gal80 at 1 μM with 1 μM of the indicated peptide. Protein and peptide were incubated 10 min at room temperature until $NaIO_4$ was added at a final concentration of 5 mM to start reaction. After 1 min the cross-linking reaction was quenched with 6_ protein loading buffer containing 100 mM DTT. The samples were separated using standard SDS-PAGE. Cross-linking experiments in the presence of *E. coli* lysate were done using strain BL21. Cell pellets from 100 mL cultures were resuspended in 1/2 NE buffer and Complete EDTA-free protease inhibitor tablets (Roche) and lysed by sonication. 1 μM His6-Gal80 and 1 μM DOPA-Gal4 AD were added to 200 μg of *E. coli* lysate. Reactions were done as described above.

Competition experiments

The competition experiments were done using Gal80 BP with either 26S proteasome or His6-Gal80. ND-Gal4 AD peptide or actin was added at the concentrations indicated and cross-linking was done as described above.

Identification of cross-linked product using 2D gel electrophoresis

Cross-linking was done as above, but reactions were done in a 5-fold larger volume (150 μL) using the same final concentrations. Reactions were quenched in 100 mM DTT and proteins were precipitated using 450 μL of cold acetone 2 h at –20 °C. The precipitates were spun down 10 min at 14000 rpm and supernatant removed. The pellet was resuspended in 125 μL rehydration buffer (7 M urea, 2 M thiourea, 4% CHAPS, 65 mM DTT, and IPG buffer pH 3–10 (10 μL IPG buffer: 1 mL rehydration buffer) (Amersham Pharmacia)) for 2 h at room temperature. Samples were loaded on a pH 3–10 IPG strip (Biorad) and separated according to pI on an IPGphor Isoelectric Focusing System (Amersham Pharmacia) according to the manufacturers instructions. After IEF, samples were equilibrated in equilibration buffer (50 mM Tris (pH 8.8) 6 M Urea, 30% Glycerol, 2% SDS, 65 mM DTT) for 15 min at room temperature and separated by SDS-PAGE. Samples were analyzed by blotting as described above. After a blot the membrane was stripped with stripping buffer (62.5 mM Tris (pH 7.4), 100 mM DTT, 2% SDS) for 30 min at 55 °C. Membranes were checked with substrate to insure complete removal of previous antibody signal before next blot. Blot films were scanned and false colored in Adobe Photoshop. Blot pairs were aligned and merged to produce the overlay images.

Synthesis of Biotin-DOPA-Ubiquitin:

The Biotin-DOPA-Ubiquitin protein was synthesized using standard Fmoc-Solid Phase Peptide Synthesis (SPPS) on a CEM Liberty microwave-assisted peptide synthesizer. Amino acid monomers and other reagents (DOPA, Biotinylated Glutamic Acid) were purchased from NovaBiochem. Monomers were dissolved in DMF (EMD Biosciences) at a concentration of 0.2M. A 20% (V/V) solution of piperidine in DMF was used as the deprotectant. Activation of the peptides was by 0.5M HBTU/HOBT and 2.0M NMM/DMF, with a final reaction chamber concentration of 0.2M. Rink Amide AM resin (NovaBiochem) with a loading capacity of 0.68mM/g was used for the synthesis. The coupling reactions were carried out at a 0.10 mMol scale and 0.05 mMol of resin was used for an 8x molar excess of monomers. To ensure complete Fmoc removal, the resin was microwaved with 20% piperidine two times. Each amino acid was double-coupled with a 5 minute mixing time: the microwave was turned on and increased the coupling temperature to 65˚C using 25w of power. The microwave then shut off and the mixture was bubbled with nitrogen for the remainder of the 5 minutes (see Fig. 1 for exact protocol). To cleave the peptide off the resin, a 95%TFA/2.5%H_2O/2.5%Triisopropylsilane solution was used. The peptide was cleaved on the CEM Liberty using an 18-minute temperature controlled microwave method which used short bursts of 25w to keep the temperature at a maximum of 38˚C. After cleavage, the peptide was precipitated using ether and dry ice. The solid product was lyophilized over night subsequently purified using

a C4 reverse-phase column (fVydac) on a Waters Breeze HPLC system. Fractions were collected and analyzed on an AppliedBiosystems Voyager DE Pro MALDI-TOF mass spectrometer. Fractions containing the correct mass were pooled, frozen, and lyophilized for subsequent use in biological experiments.

Sequence:
ZMQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRT
LSDXNIQKESTLHLVLRLRGG
Z=Biotinylated Glutamic Acid
X=DOPA

Supplemental Table 1: Synthetic Protocol on CEM Liberty

	operation	parameter	volume	drain
1	Clean resin dip tube	DMF	12ml	yes
2	wash-top	DMF	7ml	yes
3	add deprotection	20% piperidine	7ml	no
4	microwave method	30sec 50w 50˚C		yes
5	Clean resin dip tube	DMF	12ml	yes
6	add deprotection	20% piperidine	7ml	no
7	microwave method	180sec 50w 65˚C		yes
8	wash-top	DMF	7ml	yes
9	wash-bottom	DMF	7ml	yes
10	wash-top	DMF	7ml	yes
11	Clean resin dip tube	DMF	12ml	yes
12	wash-top	DMF	7ml	yes
13	add amino acid		2.5ml	no
14	add activator	0.5M HBTU/HOBT	1ml	no
15	add activator base	2.0M NMM/DMF	0.5ml	no
16	microwave method	300sec 25w 65˚C		yes
17	Clean resin dip tube	DMF	12ml	yes
18	add amino acid		2.5ml	no
19	add activator	0.5M HBTU/HOBT	1ml	no
20	add activator base	2.0M NMM/DMF	0.5ml	no
21	microwave method	coupling TempCtrl 5min		yes
22	wash-top	DMF	7ml	yes
23	wash-bottom	DMF	7ml	yes
24	wash-top	DMF	7ml	yes

Ub Chains were synthesized according to the following protocol. In a total volume of 20μl was placed 100ng of E1 ub-activating enzyme (Boston Scientific), 500ng of GST-E2-25K, 2.5 μg, and 4μl of 5X ubiquitination buffer (250mM Tris-HCl pH=7.5, 12.5 mM MgCl2, 2.5mM DTT, 10mM ATP), and the remaining volume filled with ddH20. The reaction was incubated for 7 hours at room temperature then stopped with 6X Protein loading Buffer and run onto a 4-20% SDS-PAGE gel. The gel was transferred according to standard western blotting procedures and probed with NeutrAvidin-HRP to detect ubiquitin chains.

The synthetic ubiquitin was able to be coupled into dimeric and tetrameric products as evidenced from the blot above. Indicating that the structure of the synthetic ubiqutin is comparable to the structure of the native molecule. However, these yields are generally less than what is achieved when forming ubiquitin chains with native ubiquitin. Under standard atmospheric conditions the gradual oxidation of the DOPA residue over long periods of time most likely inactivates the E2 enzyme by irreversibly cross-linking to the UBA domain, as evidenced in the blot above. Current investigation is underway into optimization of the ubiquitin chain formation with the synthetic DOPA ubiqutin substrate.

Chapter 4

Application of DOPA Chemistry Towards the Detection of Protein-Protein Interactions

Introduction

Examining how proteins transiently interact to carry out biological function at physiological concentrations and within a relevant biological timeframe remains an analytical challenge for biochemist. In addition, it remains difficult to analyze protein-protein interactions in the midst of the complex molecular environment of a cell. We have sought to develop chemical reactions capable of revealing direct protein-protein interactions by rapidly cross-linking interacting proteins together within the complex biological milieu. As mentioned previously, we have utilized the artificial amino acid 3,4-Dihydroxyphenylalanine, DOPA, as a chemical cross-linker to discover peptide targets within multi-protein complexes (**Chapter 3**). Periodate mediated oxidation of DOPA to the resultant orthoquinone captures cysteine, lysine, histidine, and α-amines on the surface of protein interaction partners via a Michael type addition (**Chapter 2**). This reaction can be carried out in cell lysates and often results in high product yields with little or no cross-linking to non-interacting proteins (**Chapter 3**). All materials for experiments utilizing DOPA reported in previous chapters were generated by incorporating the redox-active amino acid into peptides and small proteins, i.e. ubiquitin, through standard solid phase synthesis. One major advantage of the approach is this ability to incorporate novel chemical functionality, i.e. DOPA, into peptides of interest. Under optimal conditions, each coupling reaction occurs on

average with an efficiency of approximately 99%. Therefore, a peptide of 30 amino acids can be synthesized with a yield of 73.4%. Obviously, this yield drops as the desired peptide becomes longer. For example, assuming a coupling efficiency of 99%, a 70 amino acid peptide will be produced at a 49% yield and only 37% of the theoretical product amount is produced when trying to synthesize a 100 amino acid peptide. In reality, many couplings do not take place at 99% efficiency, further reducing the amount of desired product. Solid phase synthesis therefore, limits the size of molecule in which DOPA can be incorporated, and an alternative strategy must be developed in order to introduce DOPA into proteins of interest.

Besides avoiding the limitations of solid phase synthesis, there are additional reasons for pursuing alternative strategies of DOPA incorporation. As mentioned in Chapter 1, most proteins function within the context of a multi-protein complex. For well studied complexes, such as the proteasome, protein complex assembly is a regulated process occurring in modular stages[151]. Once assembled, core complexes are often stable. It is likely, therefore, that attempts to insert a protein of interest into a purified core complex might prove difficult if not done during the natural course of complex assembly. Insertion of the DOPA functionality into cells should allow for interactions to be captured in the native environment and perhaps provide a more realistic picture of the molecular architecture within the cell. We have considered three options, biosynthetic,

chemical, and genetic methods, for DOPA delivery with each being considered in detail in the subsequent sections.

Artificial Incorporation of DOPA into a Protein of Interest

Perhaps the most elegant solution to the problem of artificial amino acid incorporation into proteins of interest has been developed by the Schultz laboratory (**Figure 4.1**) over the last 15 years[59-63, 152, 153]. This group developed methodology capable of introducing over 30 different artificial amino acids into proteins, *in vivo*. These amino acids are all variants of tyrosine that introduce unique chemical functionality into the cell. One can imagine the exponential impact this type of approach will have on the synthetic chemistry/biology field as researchers become capable of moving beyond the standard 20 amino acids encoded within the natural world in the construction of novel molecules and materials.

The Schultz lab evolved orthogonal tRNA/synthetase pairs capable of specifically recognizing the artificial amino acid, charging of the tRNA with that amino acid, and site-specific incorporation of the artificial amino acid at a suppressor codon within the protein coding sequence. All of this occurs in the context of normal protein translation and has been applied in *E.coli*, yeast, and mammalian cells[153]. Most importantly, the technique was used to incorporate DOPA into recombinant myoglobin within *E.Coli* with the end goal of affording unique redox properties to myoglobin[92]. The limitations to this approach were primarily due to the cellular DOPA toxicity when present in solution at milli-molar

concentrations. This toxicity problem may be overcome however, with the use of o-amino-tyrosine, a DOPA analogue that is easily oxidized with periodate yet remains quiescent under standard atmospheric conditions. Taking into account aforementioned work, site-specific DOPA incorporation into proteins of interest should afford biochemical researchers an additional tool to map the molecular architecture of multi-protein complexes.

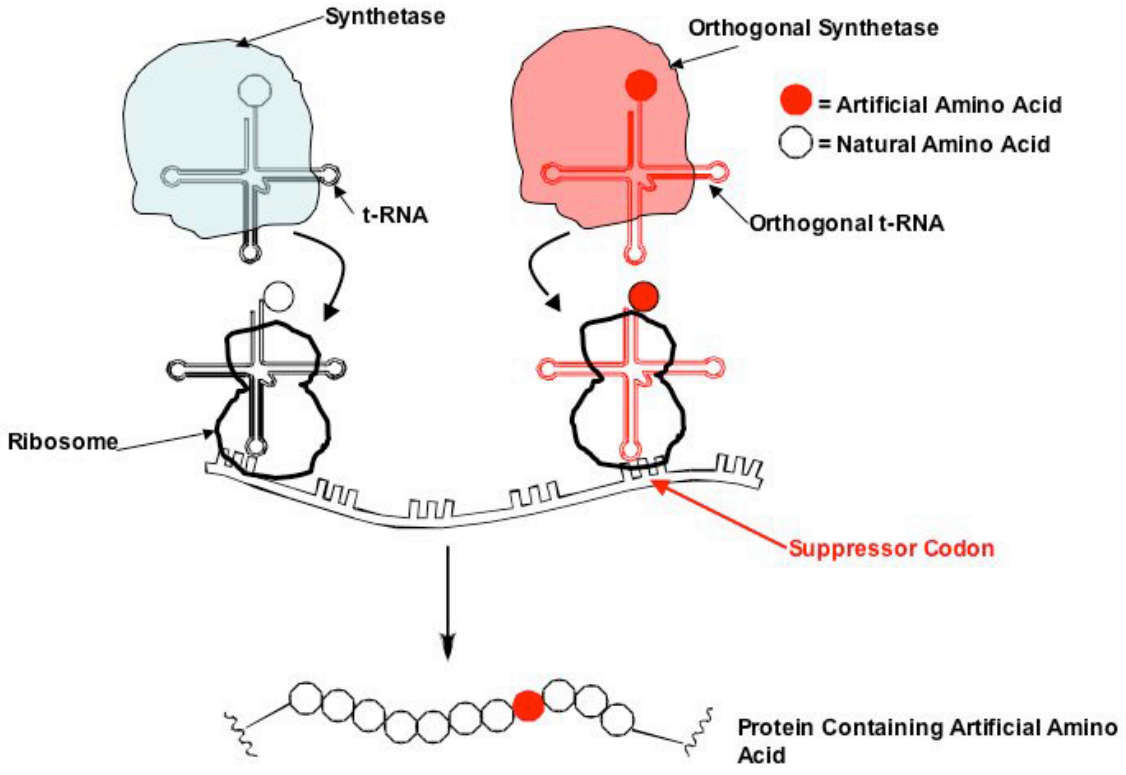


Figure 4.1: The Schultz laboratory developed a system for artificial amino acid incorporation of tyrosine derivates by generating an orthogonal t-RNA/synthetase pair capable of inserting the artificial amino acid at suppressor codons. The end result of this technology is the facile creation of proteins with novel chemical functionality.

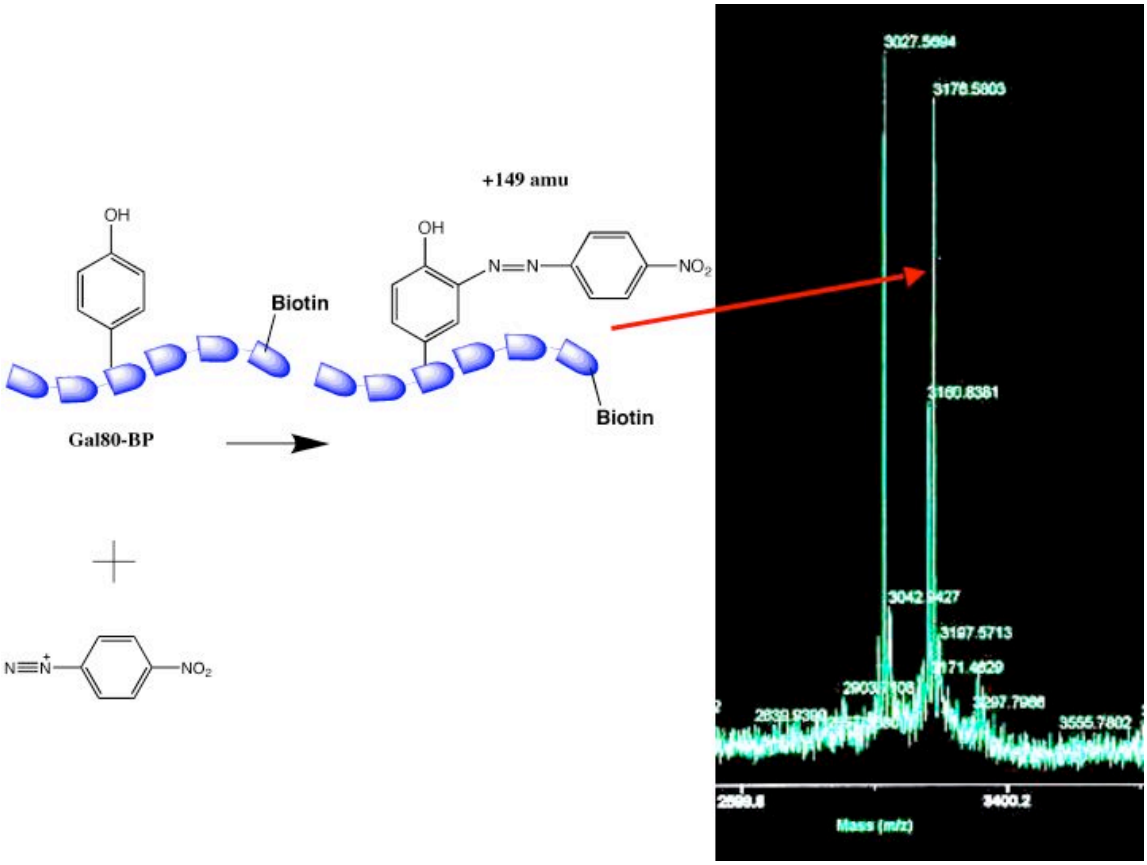
Conversion of Solvent Accessible Tyrosine to Ortho-amino-tyrosine

An alternative strategy for DOPA incorporation into proteins of interest would be to transform a specific, surface exposed tyrosine into the DOPA cross-linking moiety. Unfortunately, a convenient chemical methodology that can achieve this transformation without sacrificing the overall structure of the protein does not exist. The Francis laboratory, however, has published a synthetic protocol that transforms surface exposed tyrosine residues into ortho-amino tyrosine[154] (Figure 4.2). Ortho-amino tyrosine is sterically and electronically similar to DOPA, and like DOPA, readily undergoes periodate mediated oxidation to form an aza-quinone analogue. Based upon our previous observations of DOPA chemistry, we hypothesized that ortho-amino tyrosine would be a useful chemical cross-linking moiety.

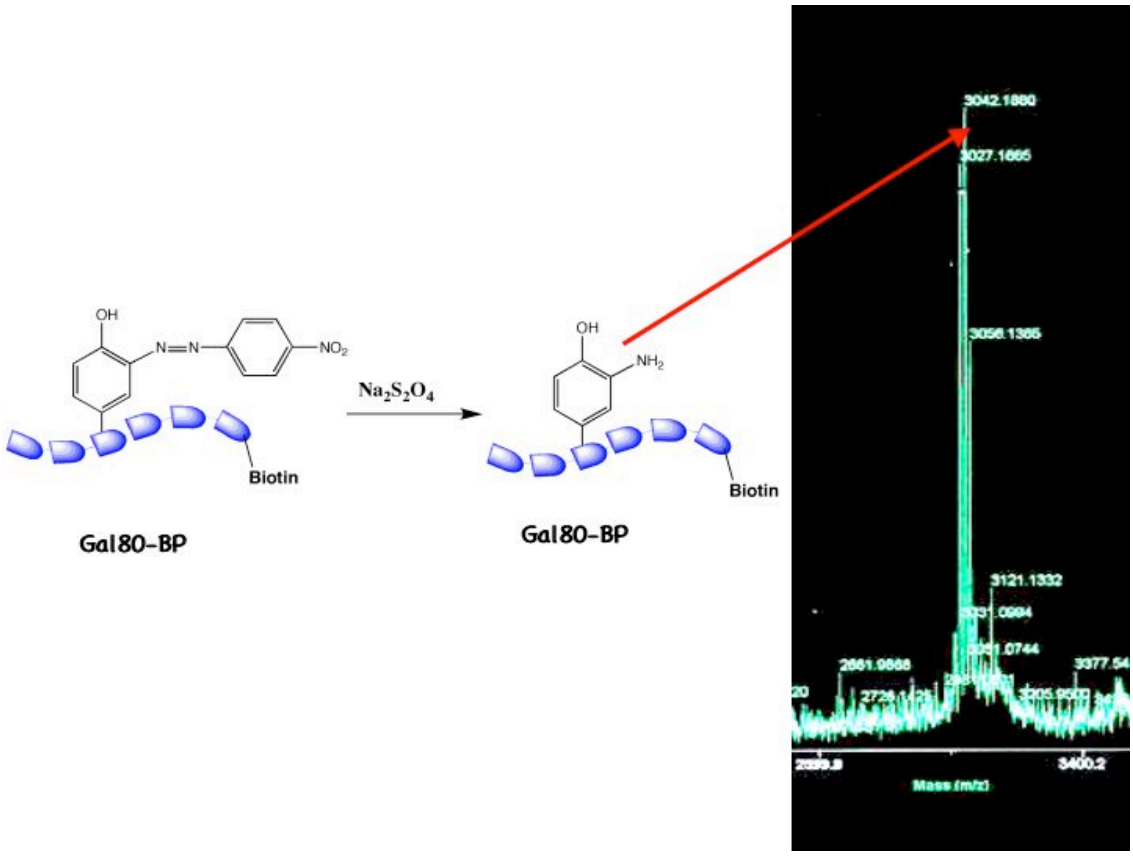
To test this hypothesis, again we relied on the aforementioned interaction between the Gal80 Binding Peptide and the Gal80 protein (**Chapters 2 and 3**). The Gal80-BP ,(Biotin-EGG(Y)DQDMQNNTFDDLFWKEGHR, m/z 3027 amu) , contains a tyrosine four residues from the N-terminus. We set out to convert this tyrosine to ortho-amino tyrosine as pictured in **Figure 4.2**. Initially, 1 mg of the Gal80 binding peptide underwent a diazonium coupling reaction at pH=9 in 150 mM $NaHPO_4$ buffer at 4° for 15 minutes. The initially colorless reaction turned a deep brown/violet color indicating the formation of a diazonium adduct with a maximal absorbance at approximately 360 nm (**Figure 4.2**). The reaction mix was then separated usung a gel filtration chromatography (NAP-10) and the components were analyzed by mass spectrometry. The spectrum in **Figure 4.2A**

shows that the mass of the starting peptide had increased by 149 amus consistent with diazonium coupling to tyrosine. Twenty milligrams of sodium dithionite was then added to the fraction containing the Gal80 BP-diazonium adduct. The solution was vortexed and left at room temperature. After two hours, the reactions was again passed through a gel filtration column and collected for analysis. As demonstrated in **Figure 4.2B** a single product was identified containing the expected mass of 3042 amu, indicative of the loss of nitroaniline and the formation of ortho-amino tyrosine at the single tyrosine on the Gal80-BP.

A.



B.



C.

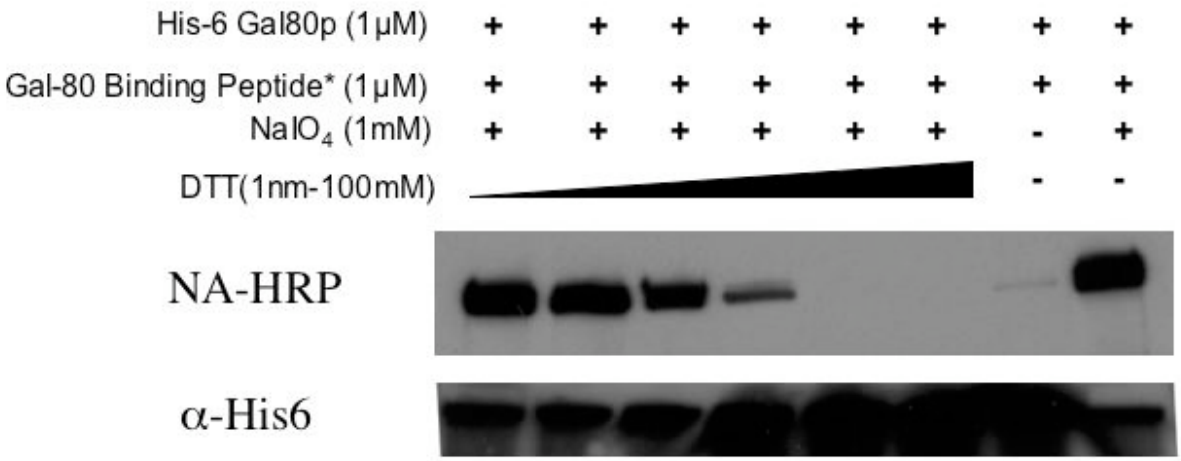


Figure 4.2: A) Diazonium Coupling reaction of the Gal80 binding peptide and the diazonium salt. The mass spectrum indicated the formation of the appropriate product with a mass of +149 amu added to the Gal80-BP B) Reduction of the diazonium adduct

to form ortho-amino tyrosine. The mass spectrum indicated an increased mass of 15 amu indicative of ortho-amino tyrosine formation on the Gal80-BP C) Ortho-amino tyrosine cross-linking experiment. 1μM of the Gal80 protein was incubated with 1μM of the Gal80-BP that contained ortho-amino tyrosine. This peptide cross-linked to Gal80 as indicated by the NA-HRP blot in the presence of periodate. The cross-linking was resistant to DTT concentrations below 10mM but was quenched at DTT concentrations 10mM and above. The α-His6 blot served as a loading control for Gal80 addition.

After forming the ortho-amino tyrosine derivative of the Gal80-BP, we next tested whether this peptide was competent to cross-link to Gal80. 1μM of the peptide was mixed with the Gal80 protein in buffer containing various DTT concentrations ranging from 1nM to 100mM (left to right in Figure 4.2C). Cross-linking was triggered with the addition of 1mM sodium periodate. After quenching the reaction with 6X protein loading buffer, the samples were run on a SDS-PAGE gel, transferred, and probed for the presence of a biotin signal overlaying the Gal80 protein. Robust cross-linking was observed between the peptide and the Gal80 protein up to a DTT concentration of 10mM. The cross-linking yields obtained with the ortho-amino tyrosine peptide were comparable to those obtained with the DOPA peptide (far right lane Figure 4.2C). These results demonstrate that it is possible to chemically transform a solvent accessible tyrosine into a chemical cross-linking group. If large amounts of protein or peptide, milligram quantities, are available this method could prove useful for mapping protein-protein interactions. However, this approach does possess significant drawbacks. First, and perhaps most importantly, this system does not have the potential to ever be applicable in vivo. Large amounts of protein or peptide must be purified, chemically modified, and then added to potential binding protein interaction partners prior to analysis. Second, all solvent

accessible tyrosine groups will be modified with this protocol. This is a significant drawback if the end goal is to analyze the molecular interactions at one particular portion of the protein. Finally, as evidenced by the mass spectrum presented in **Figure 4.2**, the amination reaction is not quantitative. The lack of complete conversion of tyrosine into ortho-amino tyrosine creates two populations of peptide that both bind at the some spot on an interacting protein. As a result, an internal competition experiment is created, most likely reducing cross-linking yields in many cases. In conclusion, this approach yields little advantage over previously mentioned cross-linking approaches utilizing DOPA chemistry in some applications, in some applications however, this could be a valuable approach if large amounts of the protein or peptide of interest are available for modification.

Genetically Targeted DOPA Delivery

Rather than pursue either of the previously mentioned options, we decided to focus on an alternative solution to the DOPA delivery problem. This involved tethering DOPA to a small molecule that could be delivered non-covalently to a peptide tag fused to the protein of interest. Specifically we created a chimeric molecule consisting of a DOPA moiety conjugated to the biarsenical ligand FLAsH. FLAsH is a fluorescent molecule designed by the Tsien lab in 1998 that binds to the tetra-cysteine motif CCPGCC or Flash Receptor Peptide, with a dissociation constat of approximately 10 pM[155-160]. This small molecule has been utilized to label proteins containing the –CCPGCC- sequence *in vivo*. We reasoned that genetic incorporation of the tetracysteine motif into a protein of

interest would provide a “landing spot” for a tethered DOPA moiety that upon oxidation with periodate could capture protein binding partners via surface exposed nucleophiles. The cross-linking reaction could then be quenched rapidly by reducing agents. We chose to use the reductant 2,3-dimercapto-1-propanol as it is quite soluble in aqueous solutions and reverses FlAsH association with the tetra-cysteine motif far more effectively than standard biological reducing agents such as 2-mercaptoethanol[155-160]. Upon reversal, FLAsH is transferred to a proximal protein binding partner that can be quickly visualized on a fluorescently scanned SDS-PAGE gel (**Figure 4.3**).

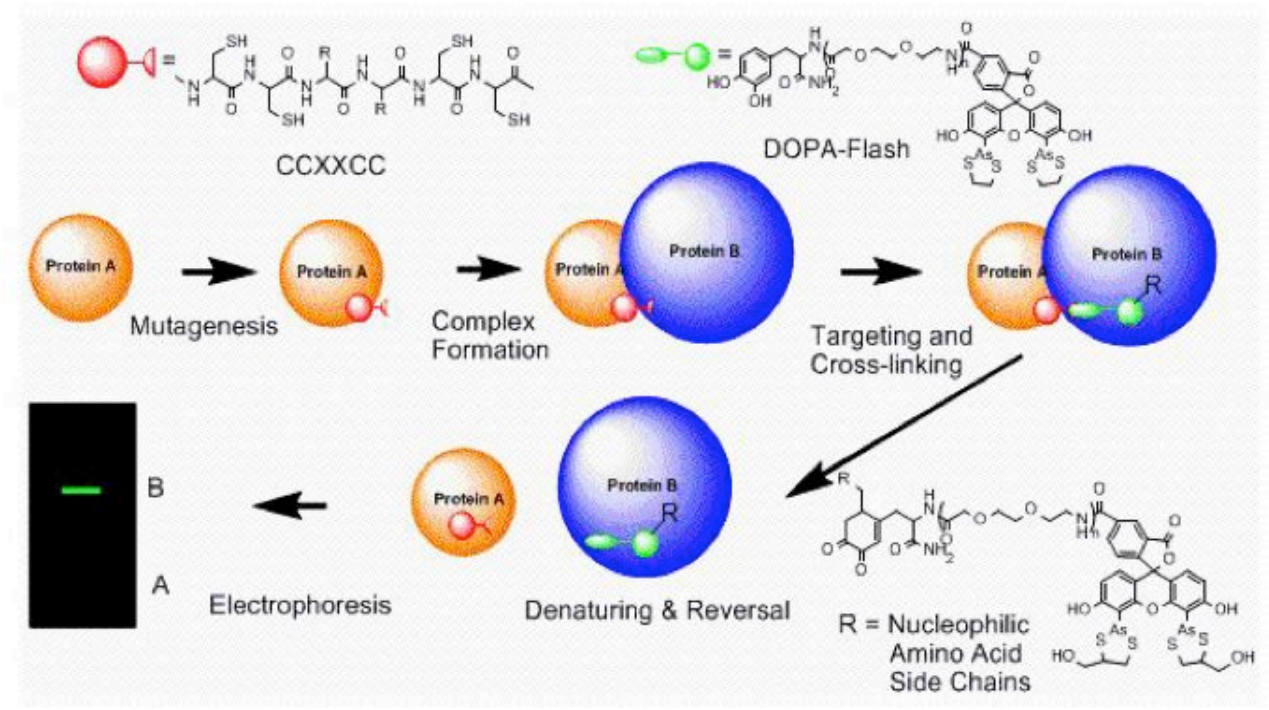


Figure 4.3: Schematic of FLAsH-DOPA cross-linking Strategy

To initially test this strategy we relied again on the Gal4 AD/Gal80 interaction. The Gal4 AD (**Chapter 3**) was synthesized with the CCPGCC tag at the C-terminal of the peptide (**Figure 4.4**), while His6-Gal80 was expressed in BL21 cells. After induction and cell lysis and removal of cellular debris, the Gal4AD-CCPGCC was added to the lysate at 1 and 10μM (**Figure 4.4** Lanes 1 and 2). FLAsH-DOPA was added to the lysate at a final concentration of 10μM (**Figure 4.4** Lanes 1,2, and 3). The reaction mix was allowed to equilibrate for thirty minutes prior to the initiation of the reaction with 10 mM sodium periodate. After cross-linking, reaction samples were quenched with standard protein loading buffer to maintain the protein-peptide ligation via the FLAsH-DOPA molecule. The samples were then run on a denaturing 4-20% SDS-PAGE gel which was then scanned for a fluorescent signal at 532 nM (**Figure 4.4A**) on a Typhoon 9200 (GE Healthcare). Cross-linking to Gal80 was detected in lanes 1 and 2, but was not observed in the absence of the activation domain despite the presence of FLAsH-DOPA (Lane 3). No signal was detected in the absence of FLASH-DOPA (Lane 4). To further characterize the product, we took advantage of the presence of a biotin tag at the N-terminus of the Gal4 AD construct. After blotting with streptavidin, a biotinylation signal was observed that over-laid with both Gal80 protein and the fluorescent signal, providing further evidence that the detected signal was a result of cross-linking the Gal4 AD to Gal80. We concluded from this set of experiments that it was possible to use this small chimeric

molecule to deliver DOPA specifically to a group of interacting proteins and probe for interaction partners.

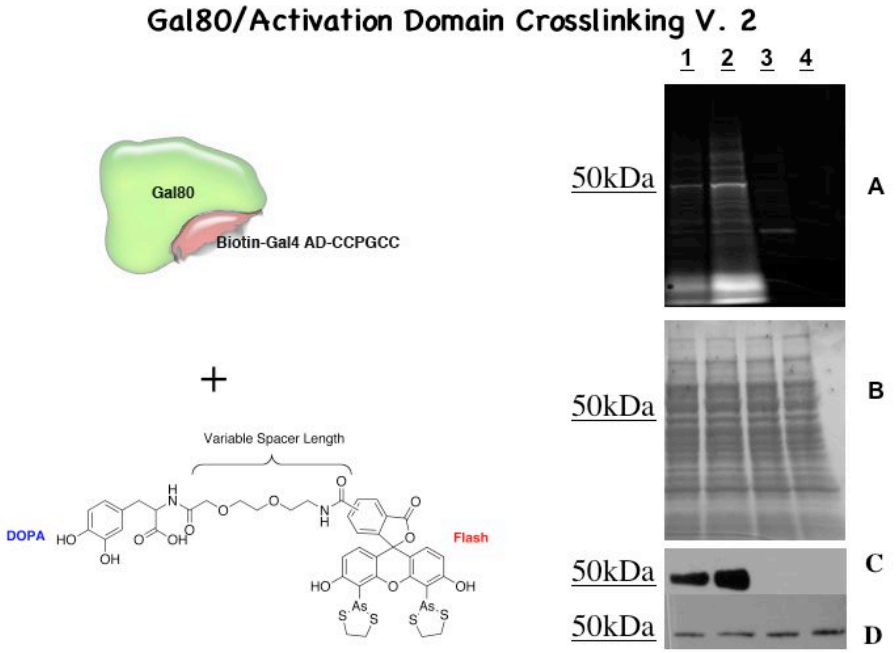


Figure 4.4: Cross-linking of the Cut-AD-CCPGCC peptide with Gal80. Lane 1: Cross-linking of the Gal 4 AD (1μM) to Gal80 with FLAsH-DOPA in cell lysate. Lane 2: Cross-linking of the Gal4 AD (10μM) to Gal80 in E.Coli lysate with FLAsH-DOPA. Lane 3: Control Reaction lacking the Gal4 AD.Lane 4: Control Reaction lacking FLAsH-DOPA A.) Fluorescent Scan at 532 nM to detect cross-linked products. B.) Stained gel demonstrating equal amounts of lysate in every lane. C.) NA-HRP western blot. D.) α-His6 blot (Gal80 loading control)

A significant issue when developing methods to probe protein-protein interactions is to determine whether said method can detect both strong and weak interactions. The Gal4 AD/Gal80 is extremely tight, with an estimated dissociation constant of approximately 4 nM[75]. To test whether the FLAsH delivery strategy could detect a much weaker protein-protein interaction, we attempted to crosslink S5a, a component of the mammalian proteasome, to the

ubiquitin-like (Ubl) domain of human RAD23A. These two recombinant proteins interact with a dissociation constant of approximately 3 μM[161] outside of the context of the 26S proteasome. An FRP site was engineered into the N-terminus of the Ubl domain and purified recombinant FRP-Ubl was incubated at various concentrations with 10 mM FLAsH-DOPA and 150 nM S5a **Figure 4.5**. After a 30 minute equilibration, the reaction was triggered by the addition of periodate (10 mM). Thirty seconds later the reaction was quenched with protein loading buffer containing 100 mM 2,3-dimercapto-1-propanol. The samples were then boiled, separated by SDS-PAGE and the gel was scanned for a fluorescent signal. Fluorescently labeled S5a began to be observable at .3 μM and gradually increased in intensity up to 9.6 μM (**Figure 4.5**). Label transfer was not observable at the lower FRP-Ubl concentrations of .05 or .1 μM, far below the Kd, or in the absence of FRP-Ubl despite the presence of equivalent amounts of S5a and DOPA-FLAsH in every sample. We concluded therefore, that this reaction was capable of reporting weak protein-protein interactions with sufficient sensitivity to be useful as an analytical method. Interestingly, we were unable to reproduce this interaction within the context of the mammalian proteasome. This indicates that perhaps Rad23 and S5a interact

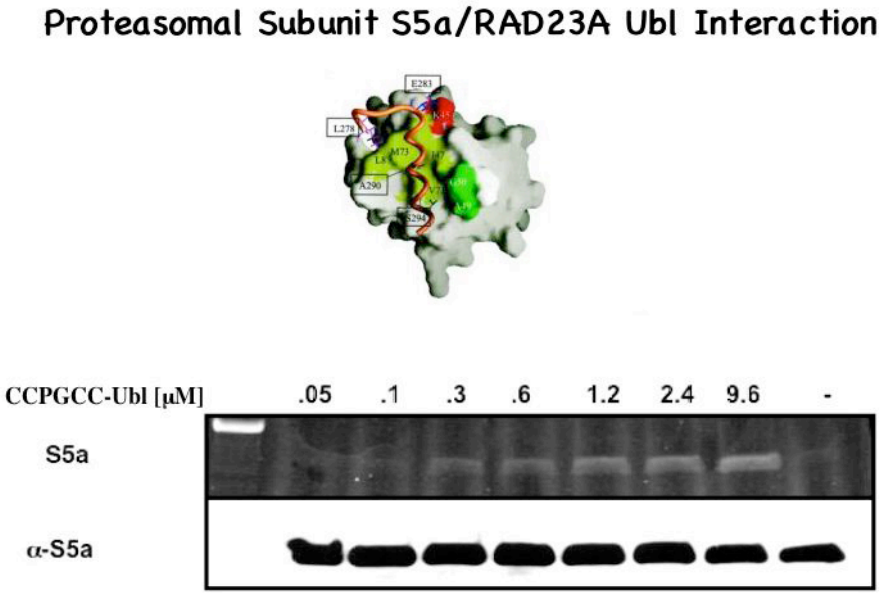


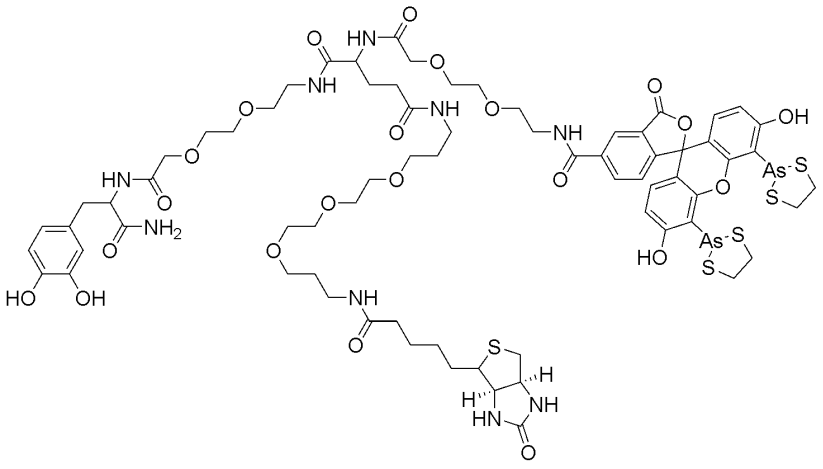
Figure 4.5: Ubl containing an N-terminal FRP site was incubated at the indicated concentrations with 10 mM FLAsH-DOPA compound and 150 nM S5a. Samples were allowed 30 minutes to equilibrate prior to the addition of periodate (10 mM). After 30s the reaction was quenched with 100 mM dimercaptopropanol in protein loading buffer and run on an SDS-PAGE gel. After fluorescent scanning the gel was probed for S5a as a loading control. Reprinted with permission Fujiwara, K. et al. J. Biol. Chem. 2004; 279:4760-4767

A Tripartite Chemical Cross-linker

The minimal amount of a CCPGCC-tagged protein necessary to detect via FLAsH labeling is approximately 1 picomole. For example, roughly 50 nano-grams of a protein with a molecular weight of 50,000 Daltons is necessary to visualize fluorescently on an SDS-PAGE gel. Although this level of detection is often sufficient, many cross-linking reactions do not yield pico-mole amounts of cross-linked products for two reasons. First, often times the protein-protein interaction partners of a protein of interest are not present in the cell in high enough amounts to produce a strong fluorescent signal when labeled with

FLAsH-DOPA. Second, the label transfer efficiency of the FLAsH-DOPA reaction is contigent upon several variables. Namely, the ability of the reactive ortho-quinone to trap a surface nucleophile depends upon the distance between the two groups and type of nucleophile on the surface of the protein. For example, in Chapter 2 we demonstrated that oxidized DOPA reacts much more efficiently with cystiene as compared to lysine. Likewise, lysine was demonstrated to couple to oxidized DOPA more efficiently than histidine. Finally, we demonstrated that if a nucleophile is not in a relatively close proximity, then no reaction will occur.

For these reason, we decided to add a biotin group to the chimeric FLAsH-DOPA molecule to increase the limits of detection of cross-linked products. This molecule was synthesized by standard solid phase synthetic methods as described in the experimental protocols with the exception being that a biotinylated-glutamic acid group was used in the middle of the molecule in replacement of an AEEA group (**Figure 4.6**). After synthesis and purification, a molecule was isolated at the appropriate mass 1734.6 amu.



Chemical Formula: $C_{71}H_{93}As_2N_9O_{22}S_5$
Exact Mass: 1733.35
Molecular Weight: 1734.71 Observed Mass:1734.9
Elemental Analysis: C, 49.16; H, 5.40; As, 8.64; N, 7.27; O, 20.29; S, 9.24

Figure 4.6: Structure of the DOPA-Biotin-FLAsH Chemical Cross-linking Molecule.

As a test to see whether this molecule would be useful in cross-linking reactions we generated a recombinant transcription factor GST-Gal4DBD-VP16-FRP that contained the CCPGCC motif on the C-terminus of the VP16 activation domain. It had been shown previously that this viral activation domain associates specifically with the AAA class ATPase Sug1 in yeast[84]. Sug1 was expressed in E.coli under control of the lac-operon. GST-DBD-VP16 was added in equal amounts to lysates in which Sug1 was induced (+IPTG) or uninduced (**Figure 4.7**) along with 10μM of the FLAsH-Biotin-DOPA molecule. After a 30 minute incubation, periodate was added to a final concentration of 5mM and the cross-linking reaction was allowed to proceed for 1 minute before quenching the reaction with 5X protein loading buffer containing 100mM EDT. Following SDS-PAGE the NA-HRP blot revealed a product slightly below 50 kilodaltons that was present only in the induced sample. This product was shown to overlay with the Sug-1 protein in a blot probed with α-Sug1. No cross-linking was observed in the un-induced sample while the sample containing only the GST-DBD-VP16-FRP protein revealed a minor product above 50 kDa, which was attributed to self-labeling due to GST dimerization. After probing for biotin and Sug1, the blot was stained with Collodial Blue Stain to reveal equal amounts of lysate were loaded into each lane. This experiment demonstrates this technique to be useful in the detection of protein-protein interactions in cell lysates.

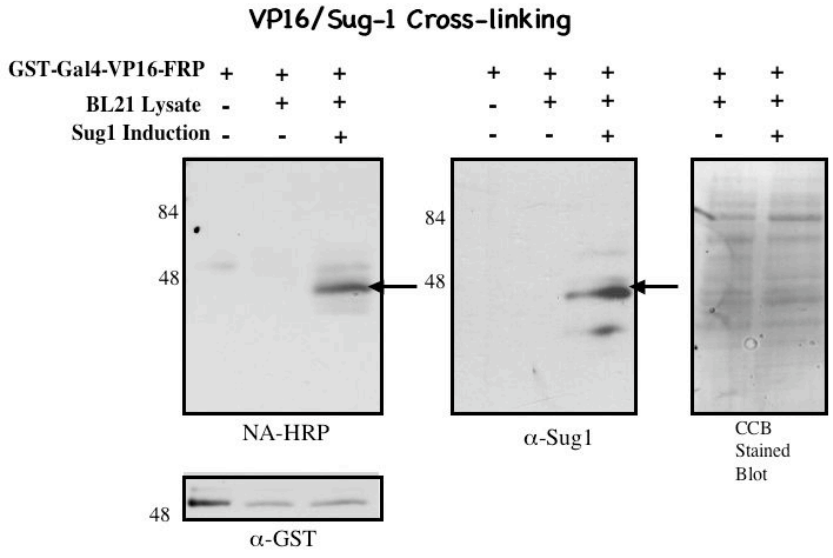


Figure 4.7: Recombinant GST-Gal4-VP16-FRP was added to cell lysates that were either induced to express Sug1p or non-induced. The western blot was probed with NA-HRP to detect biotinylated products, then with α-Sug1, and finally stained as a loading control. A GST blot was performed to reveal equilivant amounts of GST-DBD-VP16-FRP in each sample.

Conclusions

To address current analytical limitations of protein-protein interaction detection technology we designed a chimeric molecule consisting of a DOPA cross-linking moiety tethered to the biarsencial dye, FLAsH. FLAsH localizes the cross-linking moiety, DOPA, specifically to a protein of interest there bye templating the capture of interacting proteins within a linker dependent search area. This work tests the applicability of this approach by first demonstrating that FLAsH can be transferred to an interacting protein at physiological relevant concentrations and second by demonstrating that this reaction can be carried out in cellular lysates without significant background labeling. In addition we also generated a biotinylated version of the chimeric molecule that significantly lowers the limits of detection of cross-linked products. Efforts are underway to apply this technique in three general areas not readily approached by conventional technology. First, it may be possible to probe for protein-protein interactions in living cells considering that FLAsH is cell . Second, this reagent should be useful

to analyze allosteric or context protein-protein interactions in multi-protein complexes. Finally, the FLAsH-DOPA reagent has the unique property of being transferred from a protein of interest to the interaction partner. In theory, this label transfer procedure should make identification of the interaction site on the protein interaction partner rather simple. Simply, the peptide that is cross-linked to DOPA should carry a fluorescent label.

Experimental Methods

Synthesis of Gal80 Binding Peptide containing O-amino-tyrosine

Formation of the Diazonium Salt: 5λ if ab aqueous solution of p-toluene sulfonic acid monohydrate (160 mg/ml) was added to a 10 λ solution of p-nitroaniline (20mg/ml) at 4°C. The solution was vortexed for 1 minute prior to the addition of 5λ of an aqueous solution of sodium nitrite (32 mg/ml). The solution was vortexed for 1 minute and diazotization was carred out at 4°C for 1 hour.

Diazonium Coupling Reaction: To a 1 ml solution of the Gal80-Binding Peptide (Biotinylated) pH=9 was added 6λ of the diazonium salt. The solution was vortexed for 1 minute and the reaction was allowed to proceed for 15 minutes at 4°C. After 15 minutes the reaction solution was passed over a NAP-10 gel

filitration column and the brown/violet fraction was collected. This fraction was then treated with 20 milligrams of sodium dithionite for 2 hours at room temperature. After 2 hours the reaction solution was again passed over a NAP10 column and the fractions were collected. After determining the peptide containing fraction the solution was lyophilized and reconstituted to the appropriate concentration for cross-linking experiments. Cross-linking experiments were carried out in an identical fashion to that described in Chapters 2 and 3.

Synthesis of DOPA-FLAsH Conjugate:

Chemicals. 4(5)-Carboxyfluorescein was purchased from Fluka. Mercury acetate and arsenic chloride were from Sigma-Aldrich. Fmoc-8-amino-3,6-dioxaoctanoic acid (Fmoc-AEEA-OH), O-(7-azabenzo-triazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HATU), and 1-hydroxy-7-azabenzotriazole (HoAt) were from Applied Biosystems. Fmoc-DOPA(acetonid)-OH was from NOVAbiochem. Rink Amide AM resins (0.57 mmol/g) were from NOVAbiochem.

Synthesis. 4(5)-Carboxy-[4', 5'-bis(acetoxymecuri)]fluorescein. 2.0 g (6.0 mmol) mercury acetate was dissolved in 45 mL of glacial acetic acid in a 500 mL threeneck round-bottom flask. The solution was heated to 50 °C. 1.0 g (2.7 mmol) of 4(5)- carboxyfluorescein was dissolved in 150 mL 2N NaOH, and added

to the mercury acetatesolution dropwise in 1 h. The final mixture was stirred at room temperature for another hour. The mixture was centrifuged at 5,000 rpm for 15 min and the supernatant was discarded. The red precipitant was washed with 100 mL of water and centrifuged again. The supernatant was discarded and 20 mL of water was added to the wash the precipitant.

After centrifugation, the supernatant was almost colorless. The precipitant was air-dried for 3 days to give a 1.8 g of red powder, 76% yield. mp > 360 °C. 1H NMR (300 MHz, D2O, pD 13) ¶ƒ 8.17 (s, 0.3H), 8.03 (d, J = 4.1 Hz, 0.3H), 7.91 (d, J = 3.9 Hz, 0.7H), 7.80 (d, J = 4.1 Hz 0.7H), 7.66 (s, 0.3H), 7.12 (d, J = 4.7 Hz, 1H), 7.06 (d, J = 4.5 Hz, 1H), 6.86 (d, J = 3.9 Hz, 0.7H), 6.66 (d, J = 4.7 Hz, 2H), 1.91 (s, 3H)

4(5)-Carboxyl-[4',5'-bis(1,3,2-dithioarsolan-2-nl)]fluorescein. 2 g (2.2 mmol) of 4(5)-carboxy-[4', 5'-bis(acetoxymecuri)]fluorescein were placed in a 100 mL round bottle flask with 33 mL of anhydrous N-methyl-2-pyrrolidinone (NMP, Aldrich). While stirring, 3.2 mL (44 mmol) of arsenic chloride, 20 mg of palladium acetate, and 3.3 mL (18 mmol) of N,N-diisopropylethylamine (DIPEA) were added. The reaction was allowed at room temperature for 16 h. The final solution was poured into a mixture of 500 mL of acetone and 500 mL of 0.25 M, pH 7 phosphate buffer. 6.0 mL of 1,2- ethanedithiol was added and the solution was stirred for 30 min. The mixture was

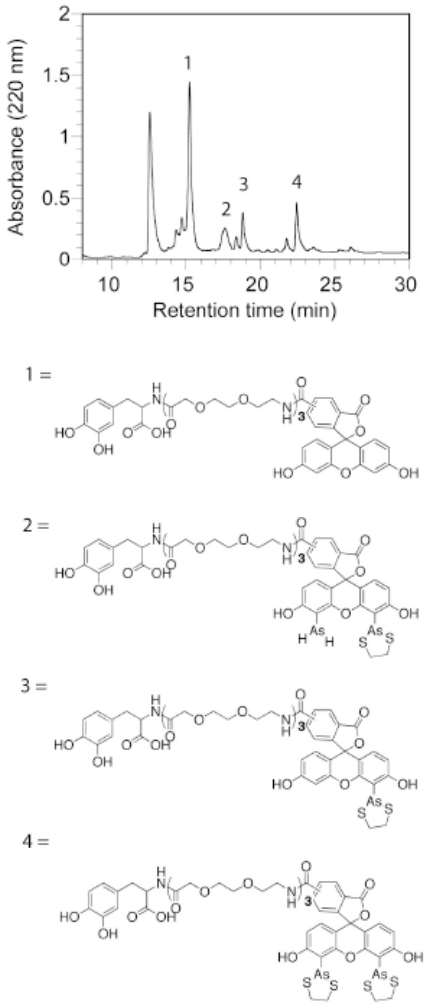
extracted with chloroform (3 X 500 mL). The combined chloroform was dried by

magnesium sulfate, and the solvent was removed by evaporation under vacuum. The red gooey residue was purified on silica column. The elution gradient was from 2% methanol in chloroform to 10% methanol in chloroform. The last fraction was collected, and the solvent was removed. The red gooey residue was triturated in 15 mL of ethanol. 15 mL of water was added and the pH was adjusted to pH 3.5 with 6N HCl to get an orange precipitant. The mixture was centrifuged at 1,000 rpm for 15 min and the supernatant was discarded. The orange solid was dried under vacuum over night. The final yield was 0.85 g, 54%. 115-116 °C dissociated. 1H NMR (300 MHz, CDCL3): ¶*ƒ* 9.94 (S, 2H), 8.72 (s, 0.7H), 8.39 (d, J = 8.7 Hz, 0.7 H), 8.35 (d, J = 8.7 Hz, 0.3H), 8.07 (d, J = 8.6 Hz, 0.3H), 7.90 (s, 0.3H), 7.28 (d, J = 8.7 Hz, 0.7H), 6.61 (d, J = 8.7 Hz, 1.4H), 6.59 (d, J = 8.7 HZ, 0.6H), 6.52 (d, J = 8.7 Hz, 1.4H), 6.50 (d, J = 8.7 Hz, 0.6H), 3.80 – 3.45 (m, 8H). N-Methyl-2-pyrrolidinone coexists with the product. ¶*ƒ* 3.41 (J = 7.0 Hz, 2H), 2.87 (S, 3H), 2.43 (J = 8.2 HZ), 2.04 (p, J = 6.9 Hz, J = 7.8 Hz, 2H). ESI MS for C25H19O7S4As2+: calculated 708.8446, found 708.8401.

DOPA-AEEAn-Flash. Rink amide AM resins (5 μmol capacity) were swelled in 5 mL of N,N-dimethylformamide (DMF) in a 25 mL peptide synthesis vessel (ChemGlass) for 45 min. Fmoc-DOPA(acetonid)-OH and Fmoc-AEEA-OH were added to the resins following the standard Fmoc chemistry. For each coupling step, 25 mgmol of each monomer, 10 mg HATU, 3.4 mg HoAT, 6.4 £gL of DIPEA, and 2.8 £gL of 2,6- lutidine were mixed in 1.5 mL of anhydrous DMF and added to the resins. Each reaction lasted 30 min at room temperature. The resins

were washed with N-methyl-2- pyrrolidinone (NMP, 5 mL ° 4). A mixture of 10 mg 4(5)-carboxyl-[4',5'-bis(1,3,2- dithioarsolan-2-nl)]fluorescein, 5 mg HATU, 1.7 mg HoAT, 3.4 μL of DIPEA, and 1.4 £gL of 2,6-lutidine in 1.5 mL of NMP was added to the resins. The reaction was allowedat room temperature for 45 min. The resins then were washed with NMP (5 mL ° 4), dichloromethane (DCM, 5 mL ° 6), and dried under vacuum for 1 h. The dry resins were treated with 300 £gL of trifluoroacetic acid (TFA) / m-cresol (80:20 v/v) for 80 min at room temperature. 2 mL of ethyl ether was added to get an orange precipitant. The precipitant was collected by centrifugation and purified by HPLC on a C-18 analytical column (11 % yield). DOPA-AEEA3-Flash MALDI-TOF: [M+1] calculated 1321.97,found 1321.86. DOPA-AEEA6-Flash MALDI-TOF: [M+1] calculated 1756.97, found 1756.41.

Supplemental Figure 4.1:



Engineering and Purificaiton of CCPGCC containing Proteins:

The CCPGCC sequence was inserted into expression vectors contatining Glutathione-S-Transferase or a His-6-thrombin-Ubl domain under control of an IPTG inducible promoter with the aid of a Stratagene Quickchange Kit. Briefly, page purified oligonucleotides were designed containing 14 base pairs of

homologous sequence to the protein of interest flanking each side of CCPGCC insertion (nucleotide insertion sequence: TGCTGCCCGGGCTGCTGC). A standard quickchange PCR reaction was run and the sample was digested with DpnI to digest the parental plasmid. The CCPGCC-protein coding plasmid was then transformed into DH5α cells for plamid purification and sequencing and then BL21 cells for protein purification. GST-CCPGCC was used in lysates while the Ubl protein was purified via a His-6 tag that was subsequently excised with biotinylated thrombin overnight at 4°C. (Note: The His-6 tag is a potent nucleophilic group for DOPA when in close proximity and must be removed when adjacent to the CCPGCC site.) Biotinlylated thrombin was removed after the reaction with the aid of StreptAvidin M-280 beads.

Label Transfer Protocol:

FLAsH-DOPA was resuspended in the following buffer : 80mM PIPES, 2mM MgCl2, 1mM EDTA, 5mM BME, and .1mM DTT. The concentration was was measured at a lamda max= 507 nm with an extinction coefficient of 4.1 X 104 m-1 cm-1. Aliquots of FLAsH-DOPA were stored in the dark at -80°C. Flash-DOPA 10μM was added to Buffer A (25mM TRIS, 10% Glycerol, 10mM MgCl2, 150mM NaCl, 1mM BME)

containing indicated concentrations of CCPGCC-Ubl and S5a at 150nM (gift from DeMartino Lab-UTSW). After a 30 minute equilibration the reaction was triggered with 5mM periodate and allowed to proceed for 30 seconds befored being quenched with standard protein loading buffer containing 100mM DMP. The

samples were boiled and run on a pre-cast 4-20% SDS-PAGE gel and scanned for fluorescent signal on a Typhoon scanner (Amersham) at 526 nm laser excitation wavelength. Western blotting and gel staining was performed according to procedures detailed in Chapters 2 and 3.

Chapter 5

Discussion

The Templating of Chemical Reactions in Biological Systems

Most chemical reactions within the cell occur within the context of a multi-protein complex. For example, two of the most well studied protein machines to date, the ribosome and the proteasome (**Figure 5.1**), make and break, respectively, peptide bonds.

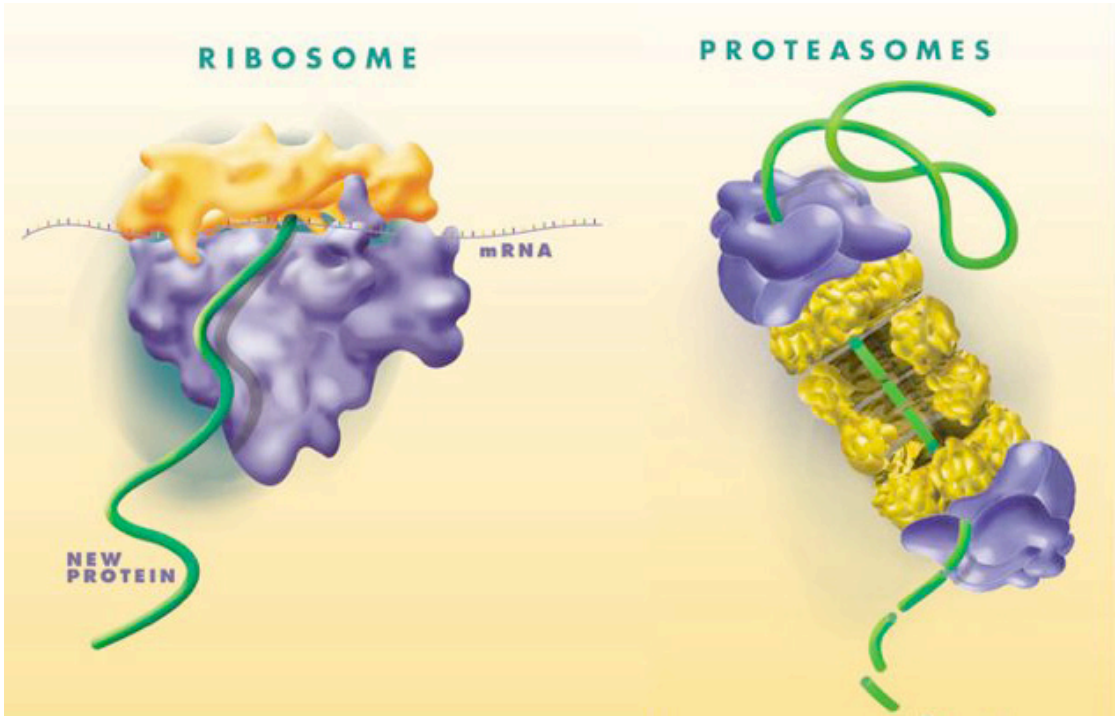


Figure 5.1: The Ribosome and Proteasome are two examples of large protein "machines" that make and break amide bonds. Most chemical reactions within the cell occur within a confined environment of a multi-protein complex effectively templating the reaction. Figures reprinted with permission from the Department of Energy Educational website. http://www.doegenomes.org

In general, multi-protein complexes like these gather reactants from the cellular milieu, orient those reactants at high effective concentrations, and catalyze a chemical reaction under strict temporal control. Given these facts, it seems to logically follow that any biochemical reaction, naturally occurring or otherwise, should also be carried out under templating conditions. By site specifically incorporating or delivering the cross-linking moiety DOPA, we took advantage of the natural interaction between a peptide or protein of interest and a target protein to catalyze DOPA adduct formation and there-bye protein-protein cross-linking. From our experiments it was clear that whenever an interaction was not present, cross-linking did not occur. In my opinion this work broadly demonstrates that reaction templating is necessary when carrying out an abiological reaction such as protein-protein cross-linking within a biological system.

The Ideal Chemical Cross-linking Reagent

In addition to confining the protein cross-linking chemistry specifically to a protein interaction interface, I think there are several other characteristics that would make up an ideal protein-protein cross-linking reagent. First, true protein-protein interaction information will only be derived from experiments performed in living cells. Therefore, the first requirement of an ideal chemical cross-linking reagent would be that the reactive moiety be of such a nature that it could be delivered or incorporated into a protein of interest in the living cell. Once inside the cell, however, the reagent must remain quiescent until activated by an

external, cell permeable signal, e.g. light. Upon activation the cross-linking chemistry should react efficiently and quickly, within seconds, but only be reactive for a short time and within a confined space. These attributes would allow the cross-linking chemistry to capture specifically the very transient, non-equilibrium state interactions amongst proteins that are not currently reported by any technique available. Finally, after the high-yield formation of a stable bond between the protein of interest and the protein interaction partner, the cross-linking reagent should contain a reporter signal to aid in product detection.

The DOPA chemistry system we developed in the lab fulfills many of the requirements listed above, however, it is not without limitations. Far and away, the largest limitation of the DOPA cross-linking approach was the context dependent nature of the cross-linking for any given reaction (Chapter 3 page 91). This experiment demonstrated that an interaction between two proteins may occur, yet not be reported by the DOPA cross-linking reaction. This most likely occurs when the oxidized DOPA moiety is not in an appropriate location to capture a surface nucleophile on a neighboring protein, but is rather, attacked by an intra-molecular nucleophile or a bulk solvent nucleophile. As a result, DOPA cross-linking experiments, like all protein cross-linking chemistries developed to date, may generate false negatives, a serious flaw with any experimental technique that must be controlled within the experiment.

In Vivo Analysis of Genetically Targeted DOPA Cross-linking

The ultimate goal of the DOPA cross-linking project is real-time protein-protein cross-linking *in vivo*. As mentioned previously, there are several reasons suggesting that this goal is obtainable with the aforementioned technology. First, the DOPA chemistry was demonstrated to be both bioorthogonal and biocompatible. As a result, neither the reducing environment inside the cell nor the presence of intracellular sugars should effect the cross-linking. Second, FLAsH, the targeting moiety of the chimeric cross-linker, was proven effective at labeling CCPGCC-tagged proteins *in vivo*.[155-160] The next experiment within this project, therefore, is to test whether either the FLAsH-DOPA molecule and/or the FLAsH-Biotin-DOPA molecule are cell permeable and competent to carry out *in vivo* cross-linking.

A strategy for carrying out this experiment is presented in **Figure 5.2**. Briefly, two Gateway vectors were obtained from Invitrogen that permit the tagging of any protein of interest with the FLAsH binding sequence, CCPGCC, at either the N or C terminus of the protein. Furthermore, Invitrogen distributes most mammalian genes in Gateway clone collections that facilitate insertion into the CCPGCC tagging vectors. With this infrastructure in place, the time required in the lab to create a mammalian protein expression vector encoding a protein whose *in vivo* interactions could be analyzed with the FLAsH-DOPA cross-linking approach would be approximately two days. In fact, I have created two vectors containing CCPGCC-tagged PSMC5, the mammalian homologue of yeast

Rpt6/Sug1, under control of the CMV promoter. After transfection of these two clones and 1.5 days of cell growth, FLAsH-DOPA can be added to the cells and allowed to bind to the recombinant protein. After 1 hour, the cells should be washed and periodate added to the medium of experimental samples. As a starting point the reaction should be allowed to proceed for ten minutes. After the addition of periodate the cells are removed from the plate, and lysed under denaturing conditions. The lysate can then be analyzed for the presence of FLAsH-DOPA labeling of PSMC5 and for cross-linked products in the periodate treated samples.

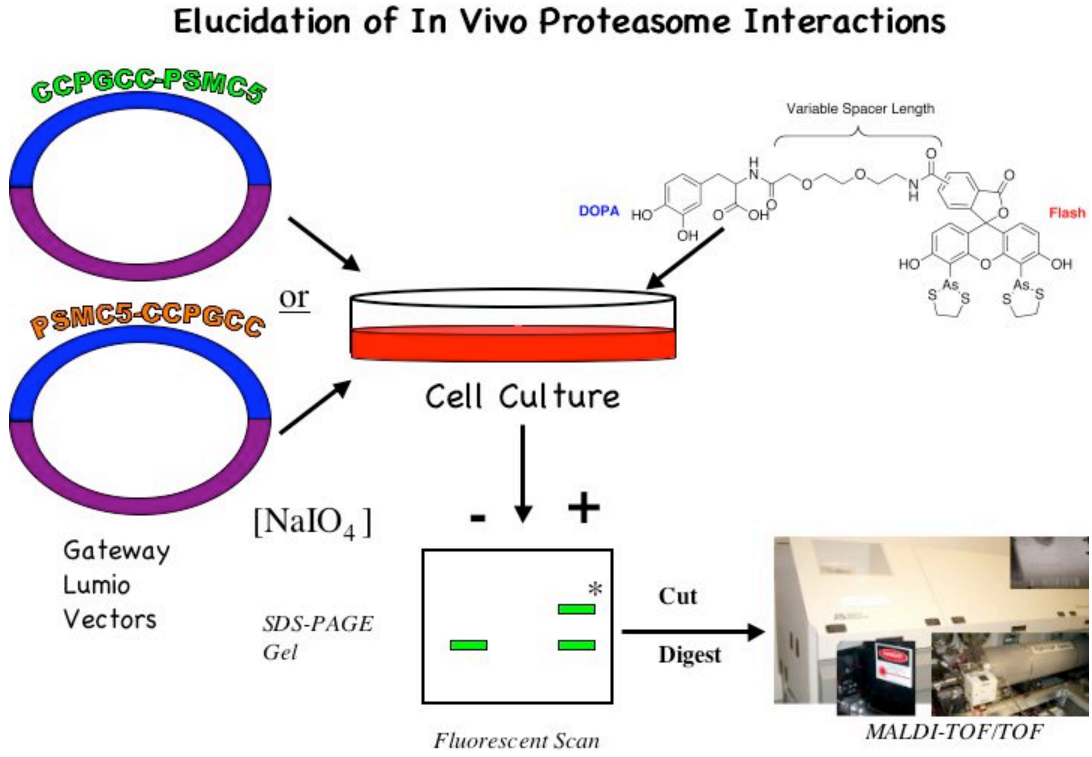


Figure 5.2: The PSMC5 gene, an ATPase component in the mammalian proteasome, was inserted in the the Lumio Gateway vector that contains the CCPGCC tag at either the C or N terminus. This mammalian expression vector can then be transformed into cells. FLAsH-DOPA can be then be added to the culture. Following FLAsH-DOPA addition the cells are washed and lysed and the lysates run on an SDS-PAGE gel. The presence of a fluorescent signal is indicative of FLAsH-DOPA premeablility, while the presence of a cross-linked product is indicative that the reaction was successful. If present in sufficient amounts, this product can by analyzed by mass spectrometry.

There are several caveats associated with this experiment, and the standard controls utilized in previous cross-linking experiments should also be applied in this case. In particular, there are significant, related barriers to the success of this approach that are worth considering in detail. First, as mentioned previously, 1 picomole of protein is necessary for fluorescent detection upon FLAsH labeling. This amount translates to approximately, 50 ng, of PSMC5. This amount of recombinant protein is easily obtained with standard mammalian cell culture procedures. However, if one assumes a modest cross-linking yield of 10%, then it becomes necessary to start with 500 ng of PSMC5 to produce the necessary 50 ng of detectable cross-linked product (top band in periodate treated sample **Figure 5.2**). Certainly, obtaining 500 ng of a protein of interest from mammalian cell culture is a more onerous task, but is ultimately obtainable with multiple plates of cell culture. The requirement for multiple plates of cells in turn requires large amounts of FLAsH-DOPA to be utilized during the cross-linking experiment. This is perhaps the most serious drawback to the approach given that the synthesis of this compound is relatively poor and proceeds with a yield of less than 10%. It seems likely that alternative methods to target DOPA to a genetically encoded sequence or the development of an improved FLAsH-DOPA synthetic scheme would be fruitful approaches in the future.

References:

1. Suchanek, M., Radzikowska, A. & Thiele, C. Photo-leucine and photo-methionine allow identification of protein-protein interactions in living cells. Nat Methods 2, 261-7 (2005).
2. Lam, Y. A., Lawson, T. G., Velayutham, M., Zweier, J. L. & Pickart, C. M. A proteasomal ATPase subunit recognizes the polyubiquitin degradation signal. Nature 416, 763-7 (2002).
3. Koh, S. S., Ansari, A. Z., Ptashne, M. & Young, R. A. An activator target in the RNA polymerase II holoenzyme. Mol Cell 1, 895-904 (1998).
4. Alberts, B. The cell as a collection of protein machines: preparing the next generation of molecular biologists. Cell 92, 291-4 (1998).
5. Cusick, M. E., Klitgord, N., Vidal, M. & Hill, D. E. Interactome: gateway into systems biology. Hum Mol Genet 14 Spec No. 2, R171-81 (2005).
6. Vidal, M. Interactome modeling. FEBS Lett 579, 1834-8 (2005).
7. Aloy, P. et al. Structure-based assembly of protein complexes in yeast. Science 303, 2026-9 (2004).
8. Lee, I., Date, S. V., Adai, A. T. & Marcotte, E. M. A probabilistic functional network of yeast genes. Science 306, 1555-8 (2004).
9. Marcotte, E. M. Assembling a jigsaw puzzle with 20,000 parts. Genome Biol 4, 323 (2003).
10. Marcotte, E. M., Xenarios, I. & Eisenberg, D. Mining literature for protein-protein interactions. Bioinformatics 17, 359-63 (2001).
11. McDermott, J. & Samudrala, R. Enhanced functional information from predicted protein networks. Trends Biotechnol 22, 60-2; discussion 62-3 (2004).
12. Xenarios, I. et al. DIP: the database of interacting proteins. Nucleic Acids Res 28, 289-91 (2000).
13. Tan, S. H., Zhang, Z. & Ng, S. K. ADVICE: Automated Detection and Validation of Interaction by Co-Evolution. Nucleic Acids Res 32, W69-72 (2004).
14. Chien, C. T., Bartel, P. L., Sternglanz, R. & Fields, S. The two-hybrid system: a method to identify and clone genes for proteins that interact with a protein of interest. Proc Natl Acad Sci U S A 88, 9578-82 (1991).
15. Fields, S. & Song, O. A novel genetic system to detect protein-protein interactions. Nature 340, 245-6 (1989).
16. Phizicky, E. M. & Fields, S. Protein-protein interactions: methods for detection and analysis. Microbiol Rev 59, 94-123 (1995).
17. Kishore, G. M. & Shah, D. M. Amino acid biosynthesis inhibitors as herbicides. Annu Rev Biochem 57, 627-63 (1988).
18. Boeke, J. D., LaCroute, F. & Fink, G. R. A positive selection for mutants lacking orotidine-5'-phosphate decarboxylase activity in yeast: 5-fluoro-orotic acid resistance. Mol Gen Genet 197, 345-6 (1984).
19. Fields, S. High-throughput two-hybrid analysis. The promise and the peril. Febs J 272, 5391-9 (2005).
20. LaCount, D. J. et al. A protein interaction network of the malaria parasite Plasmodium falciparum. Nature 438, 103-7 (2005).

21. Johnsson, N. & Varshavsky, A. Split ubiquitin as a sensor of protein interactions in vivo. Proc Natl Acad Sci U S A 91, 10340-4 (1994).
22. Stagljar, I. & Fields, S. Analysis of membrane protein interactions using yeast-based technologies. Trends Biochem Sci 27, 559-63 (2002).
23. Vidal, M. & Legrain, P. Yeast forward and reverse 'n'-hybrid systems. Nucleic Acids Res 27, 919-29 (1999).
24. Rual, J. F. et al. Toward improving Caenorhabditis elegans phenome mapping with an ORFeome-based RNAi library. Genome Res 14, 2162-8 (2004).
25. Rual, J. F. et al. Towards a proteome-scale map of the human protein-protein interaction network. Nature 437, 1173-8 (2005).
26. Wang, R. & Chait, B. T. High-accuracy mass measurement as a tool for studying proteins. Curr Opin Biotechnol 5, 77-84 (1994).
27. Zhao, Y. et al. Mapping protein-protein interactions by affinity-directed mass spectrometry. Proc Natl Acad Sci U S A 93, 4020-4 (1996).
28. Gavin, A. C. et al. Proteome survey reveals modularity of the yeast cell machinery. Nature 440, 631-6 (2006).
29. Howell, J. M., Winstone, T. L., Coorssen, J. R. & Turner, R. J. An evaluation of in vitro protein-protein interaction techniques: assessing contaminating background proteins. Proteomics 6, 2050-69 (2006).
30. Tackett, A. J. et al. I-DIRT, a general method for distinguishing between specific and nonspecific protein interactions. J Proteome Res 4, 1752-6 (2005).
31. Gingras, A. C., Aebersold, R. & Raught, B. Advances in protein complex analysis using mass spectrometry. J Physiol 563, 11-21 (2005).
32. Trester-Zedlitz, M. et al. A modular cross-linking approach for exploring protein interactions. J Am Chem Soc 125, 2416-25 (2003).
33. Chen, Y. & Ebright, R. H. Phenyl-azide-mediated photocrosslinking analysis of Cro-DNA interaction. J Mol Biol 230, 453-60 (1993).
34. Ebright, Y. W., Chen, Y., Kim, Y. & Ebright, R. H. S-[2-(4-azidosalicylamido)ethylthio]-2-thiopyridine: radioiodinatable, cleavable, photoactivatible cross-linking agent. Bioconjug Chem 7, 380-4 (1996).
35. Naryshkin, N., Kim, Y., Dong, Q. & Ebright, R. H. Site-specific protein-DNA photocrosslinking. Analysis of bacterial transcription initiation complexes. Methods Mol Biol 148, 337-61 (2001).
36. Petrotchenko, E. V., Olkhovik, V. K. & Borchers, C. H. Isotopically coded cleavable cross-linker for studying protein-protein interaction and protein complexes. Mol Cell Proteomics 4, 1167-79 (2005).
37. Denison, C. & Kodadek, T. Toward a general chemical method for rapidly mapping multi-protein complexes. J Proteome Res 3, 417-25 (2004).
38. Mullins, R. D., Heuser, J. A. & Pollard, T. D. The interaction of Arp2/3 complex with actin: nucleation, high affinity pointed end capping, and formation of branching networks of filaments. Proc Natl Acad Sci U S A 95, 6181-6 (1998).
39. Mullins, R. D., Stafford, W. F. & Pollard, T. D. Structure, subunit topology, and actin-binding activity of the Arp2/3 complex from Acanthamoeba. J Cell Biol 136, 331-43 (1997).
40. Kopp, F. et al. Subunit arrangement in the human 20S proteasome. Proc Natl Acad Sci U S A 94, 2939-44 (1997).

41. Dahlmann, B., Kopp, F., Kristensen, P. & Hendil, K. B. Identical subunit topographies of human and yeast 20S proteasomes. Arch Biochem Biophys 363, 296-300 (1999).
42. Ishiguro, A., Nogi, Y., Hisatake, K., Muramatsu, M. & Ishihama, A. The Rpb6 subunit of fission yeast RNA polymerase II is a contact target of the transcription elongation factor TFIIS. Mol Cell Biol 20, 1263-70 (2000).
43. Ishiguro, A. et al. Two large subunits of the fission yeast RNA polymerase II provide platforms for the assembly of small subunits. J Mol Biol 279, 703-12 (1998).
44. Yasui, K., Ishiguro, A. & Ishihama, A. Location of subunit-subunit contact sites on RNA polymerase II subunit 3 from the fission yeast Schizosaccharomyces pombe. Biochemistry 37, 5542-8 (1998).
45. Fellouse, F. A., Barthelemy, P. A., Kelley, R. F. & Sidhu, S. S. Tyrosine plays a dominant functional role in the paratope of a synthetic antibody derived from a four amino acid code. J Mol Biol 357, 100-14 (2006).
46. Fellouse, F. A. et al. Molecular recognition by a binary code. J Mol Biol 348, 1153-62 (2005).
47. Brunner, J. New photolabeling and crosslinking methods. Annu Rev Biochem 62, 483-514 (1993).
48. Young, P., Deveraux, Q., Beal, R. E., Pickart, C. M. & Rechsteiner, M. Characterization of two polyubiquitin binding sites in the 26 S protease subunit 5a. J Biol Chem 273, 5461-7 (1998).
49. Andersen, J. F. et al. Photoaffinity labeling of methyl farnesoate epoxidase in cockroach corpora allata. Insect Biochem Mol Biol 25, 713-9 (1995).
50. Dorman, G. & Prestwich, G. D. Benzophenone photophores in biochemistry. Biochemistry 33, 5661-73 (1994).
51. Lee, H. K., Denner-Ancona, P., Sakakibara, J., Ono, T. & Prestwich, G. D. Photoaffinity labeling and site-directed mutagenesis of rat squalene epoxidase. Arch Biochem Biophys 381, 43-52 (2000).
52. Olszewski, J. D. et al. Tethered benzophenone reagents for the synthesis of photoactivatable ligands. Bioconjug Chem 6, 395-400 (1995).
53. Knorre, D. G. & Godovikova, T. S. Photoaffinity labeling as an approach to study supramolecular nucleoprotein complexes. FEBS Lett 433, 9-14 (1998).
54. Chin, J. W., Martin, A. B., King, D. S., Wang, L. & Schultz, P. G. Addition of a photocrosslinking amino acid to the genetic code of Escherichiacoli. Proc Natl Acad Sci U S A 99, 11020-4 (2002).
55. Chin, J. W. et al. Addition of p-azido-L-phenylalanine to the genetic code of Escherichia coli. J Am Chem Soc 124, 9026-7 (2002).
56. Mehl, R. A. et al. Generation of a bacterium with a 21 amino acid genetic code. J Am Chem Soc 125, 935-9 (2003).
57. Santoro, S. W., Wang, L., Herberich, B., King, D. S. & Schultz, P. G. An efficient system for the evolution of aminoacyl-tRNA synthetase specificity. Nat Biotechnol 20, 1044-8 (2002).
58. Wang, L. & Schultz, P. G. A general approach for the generation of orthogonal tRNAs. Chem Biol 8, 883-90 (2001).

59. Wang, L. & Schultz, P. G. Expanding the genetic code. Chem Commun (Camb), 1-11 (2002).
60. Wang, L., Zhang, Z., Brock, A. & Schultz, P. G. Addition of the keto functional group to the genetic code of Escherichia coli. Proc Natl Acad Sci U S A 100, 56-61 (2003).
61. Johnson, A. E. & Cantor, C. R. Affinity labeling of multicomponent systems. Methods Enzymol 46, 180-94 (1977).
62. Johnson, A. E., Miller, D. L. & Cantor, C. R. Functional covalent complex between elongation factor Tu and an analog of lysyl-tRNA. Proc Natl Acad Sci U S A 75, 3075-9 (1978).
63. Johnson, A. E., Woodward, W. R., Herbert, E. & Menninger, J. R. Nepsilon-acetyllysine transfer ribonucleic acid: a biologically active analogue of aminoacyl transfer ribonucleic acids. Biochemistry 15, 569-75 (1976).
64. Amini, F., Denison, C., Lin, H. J., Kuo, L. & Kodadek, T. Using oxidative crosslinking and proximity labeling to quantitatively characterize protein-protein and protein-Peptide complexes. Chem Biol 10, 1115-27 (2003).
65. Fancy, D. A. et al. Scope, limitations and mechanistic aspects of the photo-induced cross-linking of proteins by water-soluble metal complexes. Chem Biol 7, 697-708 (2000).
66. Kreishman-Deitrick, M. et al. NMR analyses of the activation of the Arp2/3 complex by neuronal Wiskott-Aldrich syndrome protein. Biochemistry 44, 15247-56 (2005).
67. Burdine, L. & Kodadek, T. Target identification in chemical genetics: the (often) missing link. Chem Biol 11, 593-7 (2004).
68. Burzio, L. A. & Waite, J. H. Reactivity of peptidyl-tyrosine to hydroxylation and cross-linking. Protein Sci 10, 735-40 (2001).
69. Burzio, L. A. & Waite, J. H. Cross-linking in adhesive quinoproteins: studies with model decapeptides. Biochemistry 39, 11147-53 (2000).
70. Waite, J. H. Reverse engineering of bioadhesion in marine mussels. Ann N Y Acad Sci 875, 301-9 (1999).
71. Klug, C. A., Burzio, L. A., Waite, J. H. & Schaefer, J. In situ analysis of peptidyl DOPA in mussel byssus using rotational-echo double-resonance NMR. Arch Biochem Biophys 333, 221-4 (1996).
72. Taylor, S. W., Ross, M. M. & Waite, J. H. Novel 3,4-di- and 3,4,5-trihydroxyphenylalanine-containing polypeptides from the blood cells of the ascidians Ascidia ceratodes and Molgula manhattensis. Arch Biochem Biophys 324, 228-40 (1995).
73. Taylor, S. W., Molinski, T. F., Rzepecki, L. M. & Waite, J. H. Oxidation of peptidyl 3,4-dihydroxyphenylalanine analogues: implications for the biosynthesis of tunichromes and related oligopeptides. J Nat Prod 54, 918-22 (1991).
74. Waite, J. H. & Tanzer, M. L. The bioadhesive of Mytilus byssus: a protein containing L-dopa. Biochem Biophys Res Commun 96, 1554-61 (1980).
75. Han, Y. & Kodadek, T. Peptides selected to bind the Gal80 repressor are potent transcriptional activation domains in yeast. J Biol Chem 275, 14979-84 (2000).

76. Liu, B., Han, Y., Corey, D. R. & Kodadek, T. Toward synthetic transcription activators: recruitment of transcription factors to DNA by a PNA-peptide chimera. J Am Chem Soc 124, 1838-9 (2002).
77. Van Hoy, M., Leuther, K. K., Kodadek, T. & Johnston, S. A. The acidic activation domains of the GCN4 and GAL4 proteins are not alpha helical but form beta sheets. Cell 72, 587-94 (1993).
78. Gonzalez, F., Delahodde, A., Kodadek, T. & Johnston, S. A. Recruitment of a 19S proteasome subcomplex to an activated promoter. Science 296, 548-50 (2002).
79. Russell, S. J., Gonzalez, F., Joshua-Tor, L. & Johnston, S. A. Selective chemical inactivation of AAA proteins reveals distinct functions of proteasomal ATPases. Chem Biol 8, 941-50 (2001).
80. Russell, S. J. & Johnston, S. A. Evidence that proteolysis of Gal4 cannot explain the transcriptional effects of proteasome ATPase mutations. J Biol Chem 276, 9825-31 (2001).
81. Russell, S. J., Reed, S. H., Huang, W., Friedberg, E. C. & Johnston, S. A. The 19S regulatory complex of the proteasome functions independently of proteolysis in nucleotide excision repair. Mol Cell 3, 687-95 (1999).
82. Salghetti, S. E., Muratani, M., Wijnen, H., Futcher, B. & Tansey, W. P. Functional overlap of sequences that activate transcription and signal ubiquitin-mediated proteolysis. Proc Natl Acad Sci U S A 97, 3118-23 (2000).
83. Swaffield, J. C., Bromberg, J. F. & Johnston, S. A. Alterations in a yeast protein resembling HIV Tat-binding protein relieve requirement for an acidic activation domain in GAL4. Nature 360, 768 (1992).
84. Swaffield, J. C., Melcher, K. & Johnston, S. A. A highly conserved ATPase protein as a mediator between acidic activation domains and the TATA-binding protein. Nature 374, 88-91 (1995).
85. Sulahian, R., Johnston, S. A. & Kodadek, T. The proteasomal ATPase complex is required for stress-induced transcription in yeast. Nucleic Acids Res 34, 1351-7 (2006).
86. Gillette, T. G., Gonzalez, F., Delahodde, A., Johnston, S. A. & Kodadek, T. Physical and functional association of RNA polymerase II and the proteasome. Proc Natl Acad Sci U S A 101, 5904-9 (2004).
87. Ferdous, A., Kodadek, T. & Johnston, S. A. A nonproteolytic function of the 19S regulatory subunit of the 26S proteasome is required for efficient activated transcription by human RNA polymerase II. Biochemistry 41, 12798-805 (2002).
88. Sun, L., Johnston, S. A. & Kodadek, T. Physical association of the APIS complex and general transcription factors. Biochem Biophys Res Commun 296, 991-9 (2002).
89. Chang, C. et al. The Gal4 activation domain binds Sug2 protein, a proteasome component, in vivo and in vitro. J Biol Chem 276, 30956-63 (2001).
90. Ferdous, A., Gonzalez, F., Sun, L., Kodadek, T. & Johnston, S. A. The 19S regulatory particle of the proteasome is required for efficient transcription elongation by RNA polymerase II. Mol Cell 7, 981-91 (2001).
91. Han, J. D. et al. Evidence for dynamically organized modularity in the yeast protein-protein interaction network. Nature 430, 88-93 (2004).

92. Alfonta, L., Zhang, Z., Uryu, S., Loo, J. A. & Schultz, P. G. Site-specific incorporation of a redox-active amino acid into proteins. J Am Chem Soc 125, 14662-3 (2003).
93. Fornstedt, B., Rosengren, E. & Carlsson, A. Occurrence and Distribution of 5-S-Cysteinyl Derivatives of Dopamine, Dopa and Dopac in the Brains of 8 Mammalian-Species. Neuropharmacology 25, 451-454 (1986).
94. Montine, T. J., Farris, D. B. & Graham, D. G. Covalent Cross-Linking of Neurofilament Proteins by Oxidized Catechols as a Potential Mechanism of Lewy Body Formation. Journal of Neuropathology and Experimental Neurology 54, 311-319 (1995).
95. Pandell, A. J. Mechanism of the Fe(Iii)-Catalyzed Peracetic-Acid Oxidation of Catechol - a Biomimetic Reaction for Pyrocatechase. Journal of Organic Chemistry 48, 3908-3912 (1983).
96. Stokes, A. H., Hastings, T. G. & Vrana, K. E. Cytotoxic and genotoxic potential of dopamine. Journal of Neuroscience Research 55, 659-665 (1999).
97. Conway, K. A., Rochet, J. C., Bieganski, R. M. & Lansbury, P. T., Jr. Kinetic stabilization of the alpha-synuclein protofibril by a dopamine-alpha-synuclein adduct. Science 294, 1346-9 (2001).
98. Rochet, J. C. et al. Interactions among alpha-synuclein, dopamine, and biomembranes: some clues for understanding neurodegeneration in Parkinson's disease. J Mol Neurosci 23, 23-34 (2004).
99. Orgel, L. E. Molecular replication. Nature 358, 203-9 (1992).
100. Wu, T. & Orgel, L. E. Nonenzymatic template-directed synthesis on oligodeoxycytidylate sequences in hairpin oligonucleotides. J Am Chem Soc 114, 317-22 (1992).
101. Bohler, C., Nielsen, P. E. & Orgel, L. E. Template switching between PNA and RNA oligonucleotides. Nature 376, 578-81 (1995).
102. Li, X. Y. & Liu, D. R. DNA-Templated organic synthesis: Nature's strategy for controlling chemical reactivity applied to synthetic molecules. Angewandte Chemie-International Edition 43, 4848-4870 (2004).
103. LaVoie, M. J., Ostaszewski, B. L., Weihofen, A., Schlossmacher, M. G. & Selkoe, D. J. Dopamine covalently modifies and functionally inactivates parkin. Nature Medicine 11, 1214-1221 (2005).
104. Rosengren, E., Lindereliasson, E. & Carlsson, A. Detection of 5-S-Cysteinyldopamine in Human-Brain. Journal of Neural Transmission 63, 247-253 (1985).
105. Pieken, W. A. & Kozarich, J. W. Selective Synthesis and Hydrolysis of Dimethyl Cis,Cis-3-Halomuconates. Journal of Organic Chemistry 54, 510-512 (1989).
106. Zhou, Q. B. & Rokita, S. E. A general strategy for target-promoted alkylation in biological systems. Proceedings of the National Academy of Sciences of the United States of America 100, 15452-15457 (2003).
107. Prescher, J. A. & Bertozzi, C. R. Chemistry in living systems. Nat Chem Biol 1, 13-21 (2005).
108. Smith, G. P. & Petrenko, V. A. Phage display. Chemical Reviews 97, 391-410 (1997).

109. Han, Y. & Kodadek, T. Peptides selected to bind the Ga180 repressor are potent transcriptional activation domains in yeast. Journal of Biological Chemistry 275, 14979-14984 (2000).
110. Blank, M. L., Cress, E. A., Stephens, N. & Snyder, F. On the analysis of long-chain alkane diols and glycerol ehters in biochemical studies. J Lipid Res 12, 638-40 (1971).
111. Schweigert, N., Zehnder, A. J. & Eggen, R. I. Chemical properties of catechols and their molecular modes of toxic action in cells, from microorganisms to mammals. Environ Microbiol 3, 81-91 (2001).
112. Burdine, L., Gillette, T. G., Lin, H. J. & Kodadek, T. Periodate-triggered cross-linking of DOPA-containing peptide-protein complexes. Journal of the American Chemical Society 126, 11442-11443 (2004).
113. Kerwin, J. L. et al. Mass spectrometric analysis of catechol-histidine adducts from insect cuticle. Anal Biochem 268, 229-37 (1999).
114. Li, W. W., Heinze, J. & Haehnel, W. Site-specific binding of quinones to proteins through thiol addition and addition-elimination reactions. J Am Chem Soc 127, 6140-1 (2005).
115. Ansari, A. Z. et al. Transcriptional activating regions target a cyclin-dependent kinase. Proc Natl Acad Sci U S A 99, 14706-9 (2002).
116. Ansari, A. Z., Ogirala, A. & Ptashne, M. Transcriptional activating regions target attached substrates to a cyclin-dependent kinase. Proc Natl Acad Sci U S A 102, 2346-9 (2005).
117. Ansari, A. Z., Reece, R. J. & Ptashne, M. A transcriptional activating region with two contrasting modes of protein interaction. Proc Natl Acad Sci U S A 95, 13543-8 (1998).
118. Maheswaran, S., Lee, H. & Sonenshein, G. E. Intracellular association of the protein product of the c-myc oncogene with the TATA-binding protein. Mol Cell Biol 14, 1147-52 (1994).
119. Melcher, K. & Johnston, S. A. GAL4 interacts with TATA-binding protein and coactivators. Mol Cell Biol 15, 2839-48 (1995).
120. Stringer, K. F., Ingles, C. J. & Greenblatt, J. Direct and selective binding of an acidic transcriptional activation domain to the TATA-box factor TFIID. Nature 345, 783-6 (1990).
121. Stringer, K. F., Ingles, C. J. & Greenblatt, J. Direct and selective binding of an acidic transcriptional activation domain to the TATA-box factor TFIID. Nature 345, 783-786 (1990).
122. Baumeister, W., Cejka, Z., Kania, M. & Seemuller, E. The proteasome: a macromolecular assembly designed to confine proteolysis to a nanocompartment. Biol Chem 378, 121-30 (1997).
123. Baumeister, W., Walz, J., Zuhl, F. & Seemuller, E. The proteasome: paradigm of a self-compartmentalizing protease. Cell 92, 367-80 (1998).
124. Conaway, R. C., Brower, C. S. & Conaway, J. W. Emerging roles of ubiquitin in transcription regulation. Science 296, 1254-8 (2002).
125. Lonard, D. M., Nawaz, Z., Smith, C. L. & O'Malley, B. W. The 26S proteasome is required for estrogen receptor-alpha and coactivator turnover and for efficient estrogen receptor-alpha transactivation. Mol Cell 5, 939-48 (2000).

126. Lonard, D. M. & O'Malley, B. W. Expanding functional diversity of the coactivators. Trends Biochem Sci 30, 126-32 (2005).
127. Russell, S. J., Sathyanarayana, U. G. & Johnston, S. A. Isolation and characterization of SUG2. A novel ATPase family component of the yeast 26 S proteasome. J Biol Chem 271, 32810-7 (1996).
128. Jeong, C. J. et al. Evidence that Gal11 protein is a target of the Gal4 activation domain in the mediator. Biochemistry 40, 9421-7 (2001).
129. Ma, L., Rohatgi, R. & Kirschner, M. W. The Arp2/3 complex mediates actin polymerization induced by the small GTP-binding protein Cdc42. Proc Natl Acad Sci U S A 95, 15362-7 (1998).
130. Archambault, V. et al. Genetic and biochemical evaluation of the importance of Cdc6 in regulating mitotic exit. Mol Biol Cell 14, 4592-604 (2003).
131. Melcher, K. Mutational hypersensitivity of a gene regulatory protein: Saccharomyces cerevisiae Gal80p. Genetics 171, 469-76 (2005).
132. Melcher, K. & Xu, H. E. Gal80-Gal80 interaction on adjacent Gal4p binding sites is required for complete GAL gene repression. Embo J 20, 841-51 (2001).
133. Hashimoto, H., Kikuchi, Y., Nogi, Y. & Fukasawa, T. Regulation of expression of the galactose gene cluster in Saccharomyces cerevisiae. Isolation and characterization of the regulatory gene GAL4. Mol Gen Genet 191, 31-8 (1983).
134. Matsumoto, K., Toh-e, A. & Oshima, Y. Genetic control of galactokinase synthesis in Saccharomyces cerevisiae: evidence for constitutive expression of the positive regulatory gene gal4. J Bacteriol 134, 446-57 (1978).
135. Nogi, Y. & Fukasawa, T. Nucleotide sequence of the yeast regulatory gene GAL80. Nucleic Acids Res 12, 9287-98 (1984).
136. Nogi, Y., Matsumoto, K., Toh-e, A. & Oshima, Y. Interaction of super-repressible and dominant constitutive mutations for the synthesis of galactose pathway enzymes in Saccharomyces cerevisiae. Mol Gen Genet 152, 137-44 (1977).
137. Yocum, R. R. & Johnston, M. Molecular cloning of the GAL80 gene from Saccharomyces cerevisiae and characterization of a gal80 deletion. Gene 32, 75-82 (1984).
138. Pickart, C. M. & Eddins, M. J. Ubiquitin: structures, functions, mechanisms. Biochim Biophys Acta 1695, 55-72 (2004).
139. Huibregtse, J. M. UPS shipping and handling. Cell 120, 2-4 (2005).
140. Ohno, A. et al. Structure of the UBA domain of Dsk2p in complex with ubiquitin molecular determinants for ubiquitin recognition. Structure 13, 521-32 (2005).
141. Layfield, R. et al. Chemically synthesized ubiquitin extension proteins detect distinct catalytic capacities of deubiquitinating enzymes. Anal Biochem 274, 40-9 (1999).
142. Vijay-Kumar, S., Bugg, C. E. & Cook, W. J. Structure of ubiquitin refined at 1.8 A resolution. J Mol Biol 194, 531-44 (1987).
143. Vijay-Kumar, S., Bugg, C. E., Wilkinson, K. D. & Cook, W. J. Three-dimensional structure of ubiquitin at 2.8 A resolution. Proc Natl Acad Sci U S A 82, 3582-5 (1985).
144. Ramage, R. et al. Synthetic, structural and biological studies of the ubiquitin system: the total chemical synthesis of ubiquitin. Biochem J 299 (Pt 1), 151-8 (1994).

145. Chen, Z. & Pickart, C. M. A 25-kilodalton ubiquitin carrier protein (E2) catalyzes multi-ubiquitin chain synthesis via lysine 48 of ubiquitin. J Biol Chem 265, 21835-42 (1990).
146. Chen, Z. J., Niles, E. G. & Pickart, C. M. Isolation of a cDNA encoding a mammalian multiubiquitinating enzyme (E225K) and overexpression of the functional enzyme in Escherichia coli. J Biol Chem 266, 15698-704 (1991).
147. Cheng, H. C., Nishio, H., Hatase, O., Ralph, S. & Wang, J. H. A synthetic peptide derived from p34cdc2 is a specific and efficient substrate of src-family tyrosine kinases. J Biol Chem 267, 9248-56 (1992).
148. Higgs, H. N., Blanchoin, L. & Pollard, T. D. Influence of the C terminus of Wiskott-Aldrich syndrome protein (WASp) and the Arp2/3 complex on actin polymerization. Biochemistry 38, 15212-22 (1999).
149. Higgs, H. N. & Pollard, T. D. Regulation of actin polymerization by Arp2/3 complex and WASp/Scar proteins. J Biol Chem 274, 32531-4 (1999).
150. Higgs, H. N. & Pollard, T. D. Activation by Cdc42 and PIP(2) of Wiskott-Aldrich syndrome protein (WASp) stimulates actin nucleation by Arp2/3 complex. J Cell Biol 150, 1311-20 (2000).
151. Zwickl, P., Kleinz, J. & Baumeister, W. Critical elements in proteasome assembly. Nat Struct Biol 1, 765-70 (1994).
152. Wang, L., Brock, A., Herberich, B. & Schultz, P. G. Expanding the genetic code of Escherichia coli. Science 292, 498-500 (2001).
153. Wang, L., Xie, J. & Schultz, P. G. Expanding the Genetic Code. Annu Rev Biophys Biomol Struct 35, 225-249 (2006).
154. Hooker, J. M., Kovacs, E. W. & Francis, M. B. Interior surface modification of bacteriophage MS2. J Am Chem Soc 126, 3718-9 (2004).
155. Adams, S. R. et al. New biarsenical ligands and tetracysteine motifs for protein labeling in vitro and in vivo: synthesis and biological applications. J Am Chem Soc 124, 6063-76 (2002).
156. Hoffmann, C. et al. A FlAsH-based FRET approach to determine G protein-coupled receptor activation in living cells. Nat Methods 2, 171-6 (2005).
157. Martin, B. R., Giepmans, B. N., Adams, S. R. & Tsien, R. Y. Mammalian cell-based optimization of the biarsenical-binding tetracysteine motif for improved fluorescence and affinity. Nat Biotechnol 23, 1308-14 (2005).
158. Sosinsky, G. E. et al. Tetracysteine genetic tags complexed with biarsenical ligands as a tool for investigating gap junction structure and dynamics. Cell Commun Adhes 10, 181-6 (2003).
159. Tour, O., Meijer, R. M., Zacharias, D. A., Adams, S. R. & Tsien, R. Y. Genetically targeted chromophore-assisted light inactivation. Nat Biotechnol 21, 1505-8 (2003).
160. Tsien, R. Y. Building and breeding molecules to spy on cells and tumors. FEBS Lett 579, 927-32 (2005).
161. Fujiwara, K. et al. Structure of the ubiquitin-interacting motif of S5a bound to the ubiquitin-like domain of HR23B. J Biol Chem 279, 4760-7 (2004).

VITAE

Lyle Jackson Burdine was born in Dallas, Texas on October 23, 1975 to Gail Holmes Burdine and Alvie Neil Burdine of Arlington. He has two brothers, Landry Neil Burdine of Fort Worth age 29 and Logan Alvie Burdine of Fort Worth age 25. He attended Butler Elementary, Shackelford Junior High, and Lamar High Schoool in Arlington, Texas. He then attended Duke University from 1994-1999 on a football scholarship where he obtained a double major in chemistry and philosophy. He entered UT Southwestern Medical School in 2001 as a Ph.D. student in the Division of Basic Sciences and a member of the Kodadek Laboratory. He is married to Marie Kay Schluterman and currently has one daughter, Ella Virginia Burdine, born January 1, 2006.