Title
Development and Application of Oxidative Coupling Bioconjugation Reactions with ortho-Aminophenols

Author
Obermeyer, Allie

Publication Date
2013

Peer reviewed|Thesis/dissertation

Development and Application of Oxidative Coupling Bioconjugation Reactions with ortho-Aminophenols

By

Allie Caitlin Obermeyer

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy
in
Chemistry
in the
Graduate Division
of the
University of California, Berkeley

Committee in charge:

Professor Matthew B. Francis, Chair
Professor Robert G. Berman
Professor Danielle Tullman-Ercek

Fall 2013

Development and Application of Oxidative Coupling Bioconjugation Reactions with ortho-Aminophenols

By: Allie Caitlin Obermeyer

Abstract

Development and Application of Oxidative Coupling Bioconjugation Reactions with ortho-Aminophenols

by

Allie Caitlin Obermeyer

Doctor of Philosophy in Chemistry
University of California, Berkeley
Professor Matthew B. Francis, Chair

The synthetic modification of proteins plays an important role in the fields of chemical biology and biomaterials science. As applications of protein-based materials continue to become more complex, improved methods for the covalent modification of proteins are needed. Although many methods for the modification of native and artificial amino acids exist, they often require long reaction times or lengthy syntheses of reactive substrates. This work describes the development and application of a suite of bioconjugation reactions that utilize *ortho*-aminophenols. The oxidative coupling of aniline residues with *o*-aminophenol substrates was optimized. Potassium ferricyanide was identified as an alternative, mild oxidant for this coupling. These new conditions enabled the use of the oxidative coupling reaction in the presence of free cysteines and glycoslated substrates. Aminophenols were also discovered to react with native residues on protein substrates in addition to artificial aniline moieties. Cysteine and the N-terminus were identified as the reactive residues. The oxidative coupling of *o*-aminophenols with the N-terminus was optimized to achieve high levels of modification on peptide and protein substrates. The oxidative coupling of anilines and *o*-aminophenols was applied to the synthesis of a targeted, virus-like particle and to the detection of protein tyrosine-nitration. Overall, these updated and novel oxidative coupling methods expand the utility *ortho*-aminophenols for the modification of proteins.

Dedicated to my family, without whom I am nowhere.

Table of Contents

Chapter 5: Oxidative coupling for the detection of protein tyrosine-nitration

Acknowledgments

There are so many people to thank for their guidance, help, and kindness during my time at Berkeley. These past five years have been some of the most memorable and rewarding and most of this can be attributed to the people I have had the privilege of working with.

I would first like to acknowledge those that sparked my interest in chemistry and shared with me the excitement of research. My high school chemistry teacher, Julie Winter, had boundless enthusiasm for chemistry and shear energy that were infectious. I also owe a huge debt of gratitude to my undergraduate advisor, Seiichi Matsuda, who encouraged me to pursue independent research and allowed me the freedom to explore my interests. He has always been incredibly kind and supportive and without his help I am confident I would not be writing this.

I have many people to thank for my great experience at Berkeley, but perhaps none as much as Matt Francis. Matt has consistently been a supportive and positive mentor. His enthusiasm and positivity have helped even out the natural ups and downs associated with research. Despite his busy schedule, he manages to make time for anyone that needs it and this has been quite valuable both to me and many other graduate and undergraduate students in the department. In addition, he has managed to create a work environment that is both fun and productive.

There have been many members of the Francis group who have helped me along the way, and it is tradition to acknowledge everyone in the group with whom I overlapped. Two older students, Pat and Aaron, set the tone when I joined the group. Pat provided an excellent example of how to diligently pursue a research problem as well as patiently train new students. Aaron and I share an uncanny number of similarities - birth date, hometown in Michigan, and finally lab bench/desk. Although I did not "dance my Ph.D.", I hope I have lived up to the precedent that Aaron set. Aaron was a phenomenal scientist and incredible teacher - the students at UC Irvine are lucky to have him.

The next class was just Nick. However, he was able to accomplish as much as at least two students. He had an incredible burst of productivity at the end of his time in the group, and his example helped minimize my panic as I neared graduation. The following class was a big one - Zac, Michel, Sonny, Kanna, Gary and Wesley. Zac was a true force of nature and was one of the hardest working people I have ever met. Michel has many talents that are now sorely missed in lab; his ability to fix anything enabled everyone in the group to do experiments. Sonny did some outstanding work on DNA mediated assembly and his cheerful demeanor was always appreciated. I spent many years with Wesley in 743, and graduate school would not have been the same without him. Wesley taught me a lot about MS2, targeting, and cute animals on the internet. He also somehow survived being in a room with all women for many years, so I am sure that he can do anything he sets his mind to. Finally, Gary and Kanna were great friends both in and outside of lab. While Kanna may have been "100% Bertozzi," he was a constant source of sound scientific and presentation advice. Gary was similarly a fountain of knowledge and always the life of the party. He was willing to challenge people to make them better, or to find out what you "would rather."

The class immediately before mine demonstrated how to survive and be successful in grad school. Both Chris and Leah were immensely helpful throughout my time in the Francis group and have become close friends. Chris laid the groundwork for all of the *ortho*-aminophenol oxidative coupling discussed here. He did some truly heroic work on the oxidative coupling reaction

and always managed to maintain his cool. I am so happy that I followed Leah from Rice, not just to Berkeley but to the Francis group and then 743 Latimer. Leah is a fantastic scientist and person, and the group has not been the same since she left. Leah's perpetually positive attitude, scientific insights and gregarious laugh made graduate school not just bearable but at times also fun. She also showed me how to efficiently set-up experiments, and I am sad that I was never able to steal all of her ridiculous organizational skills. I am so lucky I got to know and be great friends with Leah. Thank you Leah, you go girl, your the bomb.

Now on to my fellow classmates, Mike, Dan, Troy, Amy, and Kristen. Mike was another joint student, but luckily not "100% Groves." Mike was a great scientific advisor and was able to answer all of my questions about "lasers" and "photons" in a way that I could maybe comprehend. He seemed to effortlessly manage being a joint student and fortunately his "coyle-spills" were limited to the Sierras. Dan has been tackling the light harvesting projects in the group with great gusto, and I have always been impressed with his persistence and quick wit. I hope he continues to share his POTDs with me after I leave, they have made this last year much more pleasant. Troy knows just about everything and I really appreciate him keeping my mom occupied these past few years with many games of Words with Friends. Amy impressed me from day one, and I am glad that we didn't have to fight to get into the group for I surely would have lost. Amy is filled with inventive ideas and is capable of realizing them. She also brought much needed biological expertise to the group. I am excited to see what she does next, hopefully one day I will be lucky enough to collaborate with her on something! Last, but not least Kristen. We spent nearly all of our time these past five years together in 743. Kristen is one of the most driven and talented people I have met. She was a great sounding board for experimental design, font choice, and fashion. Although it eventually got old, I could not think of a better person to be confused with on a regular basis. I am excited that we will be reunited on the East Coast soon!

The next class was another big one. It has been fun to watch Kareem develop over his time here. He came up with a crazy idea for a photochemical oxidative coupling, but then was actually able to make it work! He has always been very sweet and I hope only good things come his way. Jelly was also full of ideas and able to make them happen. If anybody can make antibody conjugates that selectively kill cells, it will be Jelly. JGlo/SSS/Jeff lived in the DTE group, but his vast knowledge of MS2 always came in handy. Nobody, myself included, thought that he would be able to disassemble and reassemble MS2 but he remarkably managed to do that and so much more. I cannot wait to see what else he thinks of to put inside capsids! Katherine has natural scientific abilities as evidenced by her work on cellulases and delicious baked goods. Stacy was an honorary member of 743 as she occupied a small corner of the room for far too long. I am in awe of how she was able to create (and finish!) a project that nobody in the group was really able to help her with. She has an incredible work ethic and positive, yet realistic, attitude. She was also a constant source of advice and an outstanding lunch companion. I hope she will continue to share adorable pictures of dogs, and hope that some of them will be of her own dogs.

Abby ended up in a class all by herself, and I am not sure how well that would have worked out for anyone else. Abby has a positive intensity that is unmatched and if she wants something done it happens. While she is sometimes loud, she is always excited and happy for others and willing to share in their success. She has spearheaded the lab's work on peptoids and continues to come up with interesting applications for metal binding peptoids. Abby has become one of my besties and

she has made life in and out of lab fantastic these past few years. I hope that she and Mike make the trek to the East Coast soon, and not just so that I can continue to steal all of her nail polish.

Lastly, all of the youngsters. Jake may be quiet but he is also amazingly capable. He has consistently kept the new mass spec functioning, a nearly impossible task. Ioana has tackled the *in vivo* targeting of MS2 with amazing ferocity and I hope that some day she is able to get MS2 inside tumors. Kanwal has deftly managed to take on many bioconjugation projects. She has incredible enthusiasm, instincts and fashion sense - all of which I am sure will serve her well. Despite the drastic change in gender balance, Jim and Richard have been excellent labmates in 743. They both hit the ground running and were valuable resources and fun to be around from the beginning. Jim has expansive synthetic knowledge and is always willing to help, or at the very least gossip. He has also provided me with some of the most vivid memories from my time in the group and for that I am appreciative. Richard also liked to cause mischief, but he is also an incredible scientist. He has an impressive knowledge of both physical and organic chemistry, unrivaled by anyone in the group. He has also shared his coding expertise with the group, making a much improved, functional group website and a speedy plotting program that has been critical in helping me finish up this past year. I am confident that the two of them will be successful both in the Francis group and beyond. Jenna and Jess have been great recent additions to the group. They have both jumped right in and figured out how to navigate the group. I am excited to see what they both end up accomplishing. I am sorry I did not get to know Joel better, but I am sure he will go on to do great things.

We have not had too many post docs in the Francis group, but they have made up for it by being fantastic and providing much needed expertise. Praveena was always positive and friendly and a good companion when learning how to ski. Michelle established most of the biological facilities for the group, which was no small feat. Henrik has been a cheerful addition to the group, and his willingness to explore many project areas has been inspiring. Two post docs recently joined 743, Adel and Meera. Adel has been nothing but nice and patient with me. He has enthusiastically thrown himself into the group, actively learning about what everyone is working on. This allows him to help anyone that needs it, be it synthetic advice or general experimental design. He is incredibly talented and is working on so many things that I cannot wait to see what he accomplishes in the group. He will make a fantastic advisor some day (soon!). Although Meera has not been here long, she is already off to a running start. She is fun, hard-working, and excited to learn new things - three qualities that will ensure her success in the group and beyond.

I also had the pleasure of working with an outstanding undergraduate, Bryce Jarman. It was a joy to work alongside him and see him grow into a wonderful scientist. He was supremely intelligent and perceptive, skills that will serve him well in graduate school. He also tolerated all of my crazy lab practices and data processing requirements, for which I am extremely grateful. In addition to Bryce, I had the opportunity to work with two great rotation students, Cameron and Am. They were both fun to work with and extremely bright, while Cameron did not join the group hopefully I will be one for two. I would also like to thank friends outside of lab for their constant support and for providing a distraction from lab. In particular I would like to acknowledge the ladies of book club and the organizers of SLAM. I am impressed by each of you and grateful for your help in making me a better person.

Lastly, I need to thank my family. My mom and dad have been incredibly supportive and encouraging both before and during graduate school. I am so appreciative of all of their help and

willingness to listen to me complain about lab. I am also lucky to have two wonderful sisters, both of whom have supported me throughout my time at Berkeley. From helping me with Excel to answering my constant texts and sending encouraging gchats, I could not have made it to this point without their love and support. And finally, I need to thank John, my best friend and the love of my life. Not only has he put up with all of my (hopefully grad school induced) craziness, but he has always believed in me and encouraged me to try and achieve my dreams. He has pushed me to become the best person and scientist I can be, despite not having taken any science classes in college. He has always been there for me and supported me in any way that he can. There would be no pages after this one without him.

Chapter 1
Site-specific protein modification

Abstract

The synthetic modification of proteins plays an important role in the fields of materials science and chemical biology. Conjugation of synthetic molecules to proteins enables the incredible diversity of protein structure and function to be harnessed in new protein-based materials. These protein bioconjugates have been used for varied applications from the investigation of biological function to use in new biomaterials such as targeted therapeutics. The synthesis of these materials requires well-defined protein bioconjugation reactions. However, these reactions must take place in water, at ambient temperature, near neutral pH and in the presence of a wide array of functional groups. Methods for the modification of both natural and artificial amino acids are reviewed herein. Site-specific modification using enzymatic tags, non-canonical amino acids, and the N-terminus are highlighted.

1.1 Importance of protein-based materials

The modification of proteins plays an important role in the fields of chemical biology and materials science. The introduction of synthetic components to proteins enables the study of native protein function,[1] the modulation of the inherent properties of the protein,[2] and the construction of novel protein-based materials (Figure 1.1).[3] For example, modification of proteins with fluorophores enables tracking of their location within cells or even organs.[1,4,5] Conjugation of a second, spectrally unique fluorophore enables fluorescence resonance energy transfer (FRET) to monitor protein folding[6,7] or protein-protein interactions.[8,9] In addition, synthetic modification of proteins with small molecule drugs or polymers can be used to create potent biologic therapeutics with improved pharmacokinetics.[10,11] Attachment of polymers to proteins combines the well-defined materials properties of the polymer with the structure and function of the protein. These protein-polymer materials based on elastin, cellulase, and metallothionein, have found use in tissue engineering, enzyme recycling,[12] and water remediation, respectively.[13]

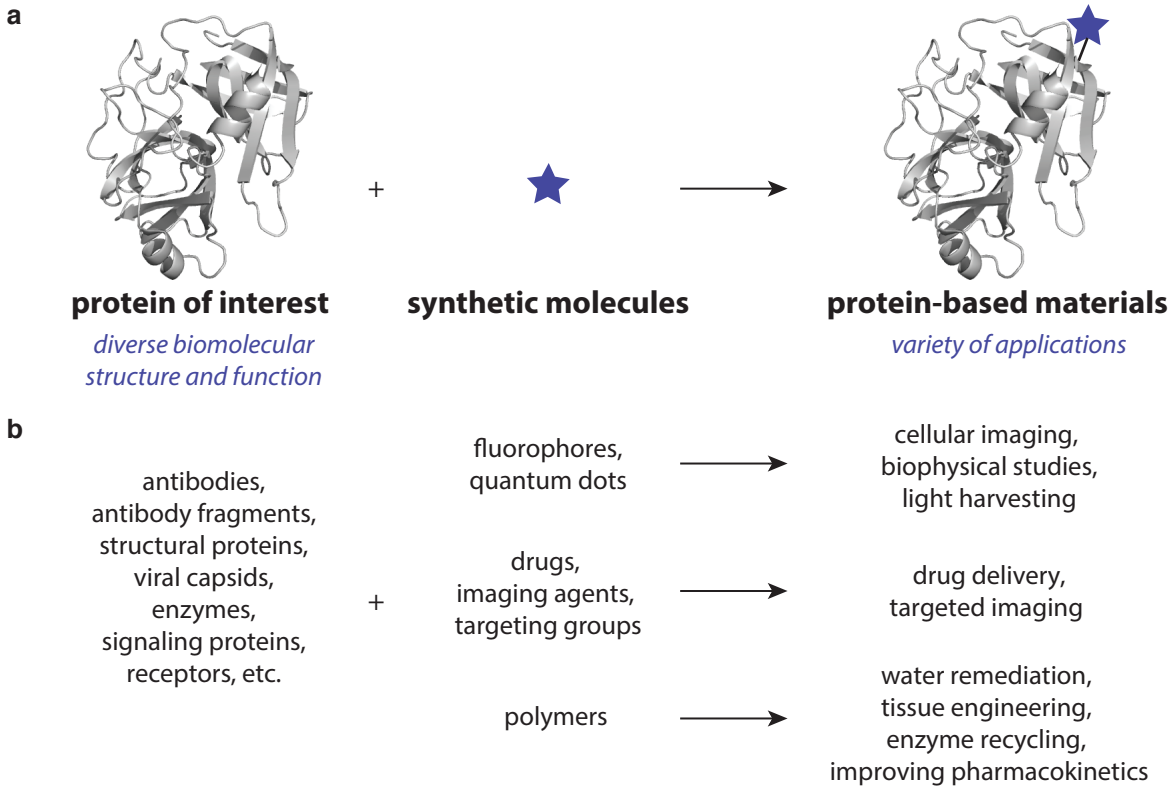


Figure 1.1. (a) The attachment of synthetic components to proteins of interest allows for the synthesis of protein-based materials. (b) The resulting materials take advantage of the diverse biomolecular structure and function of proteins and can be used for a variety of useful and promising applications.

A unique class of proteins, antibodies, have the ability to bind nearly any target antigen. This makes them of great interest for synthetic modification (Figure 1.2a). Attachment of fluorophores to antibodies enables their use in immunofluorescence; these chemically modified antibodies have become essential tools for the investigation of cellular biology. The advancement of these biomolecules as therapeutics has opened up new opportunities for "magic bullet" therapies using syn-

thetically modified antibodies. Combination of the targeting and pharmacokinetic properties of antibodies with the potent cytotoxic nature of small molecule drugs can create highly effective antibody-drug conjugates (ADCs).[11] The first ADC was approved by the FDA in 2011 (Adcetris) and there are currently over 20 ADCs in clinical trials.[14] Additionally, the attachment of poly(ethylene glycol) (PEG) chains to antibodies and antibody fragments has been shown to decrease their immunogenicity and increase their circulation time.[15] This modification results in improved therapeutic properties and even greater potential for these hybrid materials.

Work in the Francis group relies on the synthetic modification of proteins with interesting, defined structure for a variety of applications. As one example, we have created an artificial light harvesting complex by conjugation of donor and acceptor chromophores to the rod-like tobacco mosaic virus.[16,17] Utilizing the hollow, spherical bacteriophage MS2 viral capsid, we have constructed an array of targeted imaging agents by selectively attaching imaging agents to the interior surface of the capsid and targeting groups to the exterior surface.[18] In addition, we have utilized proteins with unique native function for applications in enzyme recycling[12] and water remediation.[13] Using a protein (metallothionein) to cross-link hydrogels allowed for both the detection and removal of heavy metal toxins in water.[13]

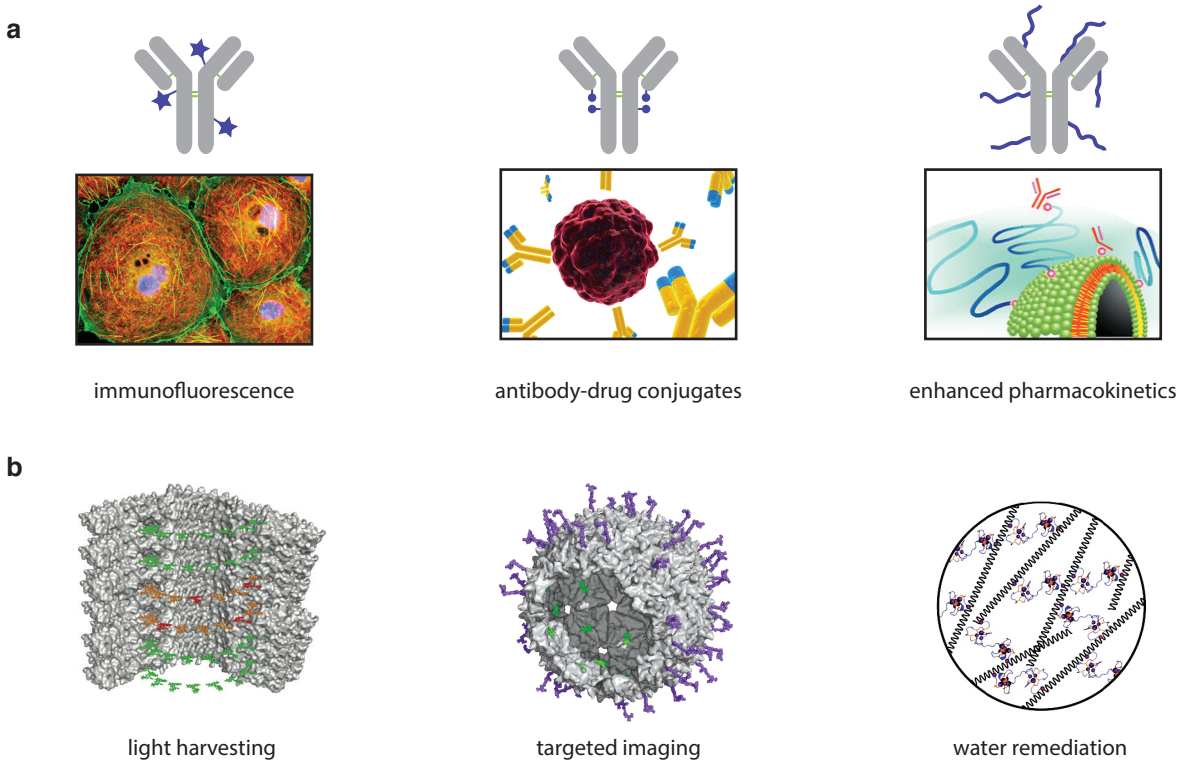


Figure 1.2. Examples of applications of synthetically modified proteins. (a) Antibodies can be modified with fluorophores for cellular imaging (immunofluorescence), with drug molecules to create antibody-drug congjuate therapeutics, or with poly(ethylene glycol) chains for enhanced pharmacokinetics. (b) Within the Francis group, proteins are modified for a variety of applications. Tobacco mosaic virus (TMV) is modified with fluorophores for light harvesting, bacteriophage MS2 is modified with imaging agents and targeting groups for targeted imaging, and toxin-binding enzymes (metalloproteins) are modified with polymers for water remediation. Images from: rockland-inc.com, Ornes, S. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, 13695, and www.peg-drug.com.

As demonstrated by these examples, the incredible diversity of protein structure and function can be utilized to create novel, highly functional materials, provided the protein can be readily modified with synthetic components. Specific, protein-compatible chemical reactions are needed to attach proteins to synthetic molecules, polymeric materials, and surfaces. Therefore, the improvement of existing methods and development of new methodology for well-defined protein modification can facilitate the synthesis of these protein-based materials.

1.2 Modification of native amino acids

Several methods exist for the modification of native amino acids (Figure 1.3). These methods typically target the nucleophilic side chains of lysine and cysteine. At high pH (~8.0), the side-chain amino group of lysine can be modified with activated esters (e.g. NHS esters), sulfonyl chlorides, and iso(thio)cyanates to form amides, sulfonamides, or (thio)ureas.[19–22] Additionally, imine formation between lysine and functionalized aldehydes followed by reductive amination forms a stable linkage to lysine.[23] This approach results in a secondary amine, which minimally perturbs the isoelectric point (pI) of the protein. The high abundance of lysine[24] on protein surfaces ensures that nearly any protein can be modified using these methods. However, the location and number of modifications are hard to control given this high abundance. In addition, these conditions are typically not selective for lysine as they can also modify the N-terminal amine.

The carboxylic acid side chains of aspartate and glutamate are also highly abundant on protein surfaces. These acids can be coupled to amines using water soluble activating agents, such as 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC).[20–22] As amino groups are also natively present on the surface of proteins, this method can also result in the undesired cross-linking of proteins.

The relatively low abundance of cysteine on protein surfaces allows for better control over the level and location of modification. The thiolate side chain can be modified using disulfide exchange to form a mixed disulfide. This strategy creates a labile modification that can be reversed upon addition of common reductants such as dithiothreitol, β-mercaptoethanol, tris(2-carboxyethyl) phosphine (TCEP), or glutathione. Cysteines can be alkylated with alkylhalides (such as iodoacetamides) or via Michael addition with α,β-unsaturated carbonyl compounds (such as maleimides). Recently, several new methods for the modification of cysteine have been developed. Mono- and dibromomaleimides have been developed as reagents for reversible cysteine modification.[25] The dibromo reagents can be inserted into disulfide bonds to modify the protein without disrupting the contribution of the disulfide bond to the protein structure. At high pH (10-11) cysteine can be transformed into dehydroalanine using *O*-mesitylenesulfonylhydroxylamine.[26] This new α,β-unsaturated carbonyl moiety can then be modified with thiol reagents giving a thioether linkage. The introduction of a C-C double bond on the protein surface, followed by thiol-ene chemistry has gained popularity as the genetic introduction of alkenes has become facile.[27,28] However, the modification of native cysteines via thiol-ene chemistry has also been demonstrated.[29] Cysteines have also been converted to allyl sulfides for their subsequent use in cross-metathesis.[30,31]

More recently, methods have been developed for the modification of the low-abundant aromatic amino acids tyrosine and tryptophan. Early methods for the modification of tyrosine relied on electrophilic aromatic substitution with reagents such as iodine and nitrous acid. At low temperature, diazonium salts can be used to modify the phenolic side chain of tyrosine resulting in an azo linkage. Expanding on this electrophilic aromatic substitution chemistry, our group has developed a three-component Mannich reaction between tyrosine, aldehydes and anilines.[32] Cyclic

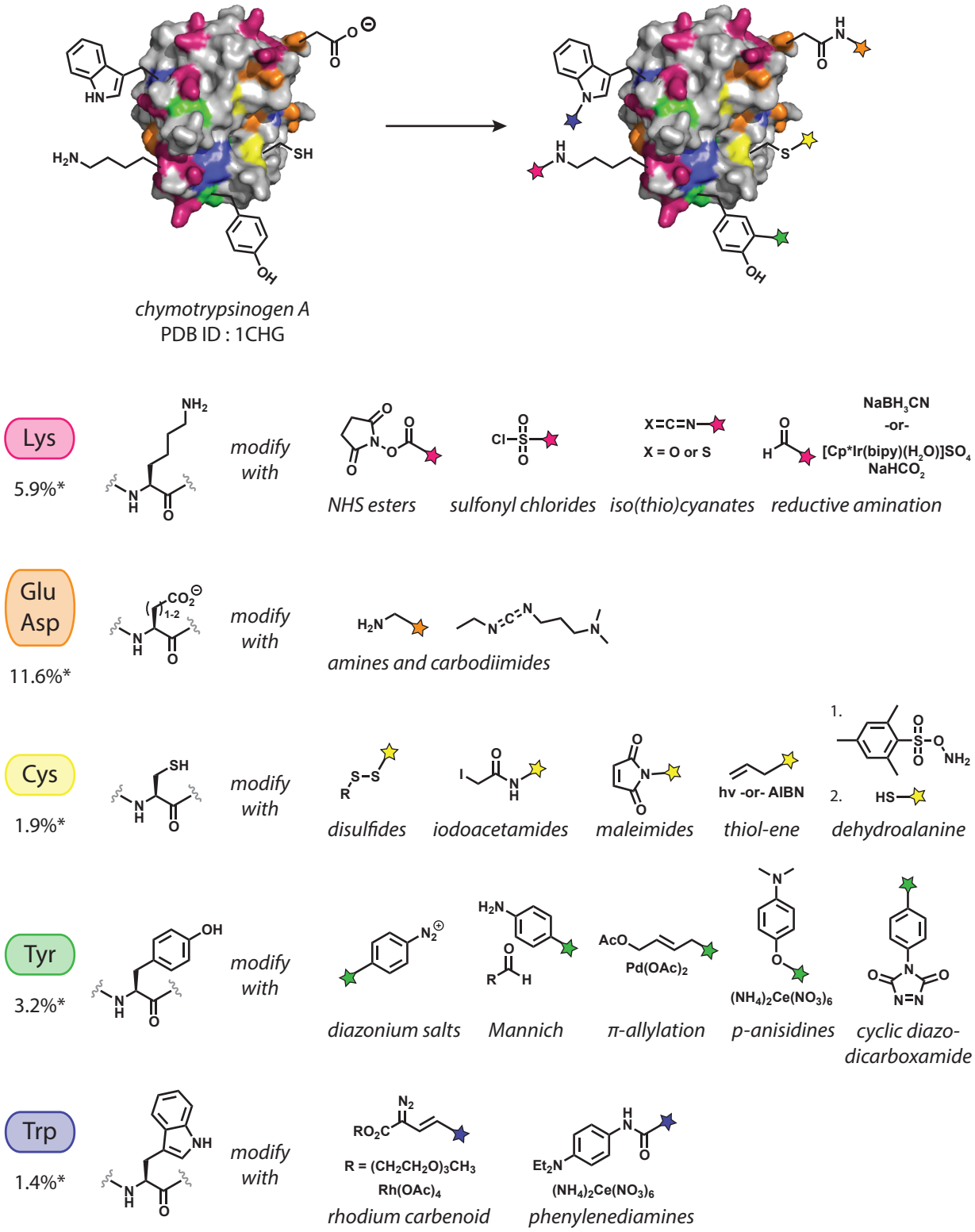


Figure 1.3. Bioconjugation methods for native amino acid side chains that occur in either high or low abundance on a protein surface (* = average abundance in proteins).[24]

diazodicarboxamides have also been advanced for the selective modification of tyrosine.[33,34] Transition metal and lanthanide based methods have been developed for the modification of tyrosine and tryptophan. Tyrosine can be modified with palladium through π-allyl chemistry.[35] The phenolic side chain of tyrosine can be cross-linked with other phenols using a Ni(II) catalyst and co-oxidant. Recent work in our group has also shown that in the presence of ceric ammonium nitrate (CAN) tyrosine can be coupled to electron-rich aromatic rings, such as *N,N*-dialkyl-*p*-anisidines.[36] Ce(IV) can also be used to couple *N,N*-dialkyl-*p*-phenylenediamines to both tyrosine and tryptophan residues.[36] Tryptophan residues can be selectively modified with rhodium carbenoids generated *in situ* from diazo compounds and $[Rh_2(OAc)_4]$.[37] Addition of *N*-(*tert*-butyl)hydroxylamine allows the reaction to be carried out at near neutral conditions.[38]

These methods collectively provide means for residue-specific modification and allow for relatively facile modification of native proteins. However, to achieve site-selectivity it is often necessary to genetically introduce a uniquely reactive residue. Recent work in the field has focused on the development of alternative methods to site-selectively modify proteins. Three general approaches have been explored for site-selective modification and will be discussed in the following sections.

1.3 Enzymatic methods for site-selective modification of proteins

Many enzymes have the native function of attaching small molecules to protein substrates. By harnessing the capability of these enzymes, they can be used to artificially label proteins of interest.[39] Elucidation of the enzyme's recognition sequence allows for the genetic introduction of this sequence as an artificial peptide (or small protein) tag to the protein of interest. Additional understanding of the enzyme's tolerance of perturbations to its natural substrate enables the attachment of functionalized small molecules to the tagged protein of interest. Depending on the tolerance of the enzyme, synthetic (or biological) molecules of interest can be directly conjugated to tagged proteins or the enzymes can be used to attach a small reactive probe for secondary conjugation with the desired synthetic molecule (Figure 1.4).

A commonly used method for enzymatic labeling relies on the enzyme transglutaminase. Transglutaminase natively catalyzes the cross-linking of lysine and glutamine residues to form an isopeptide bond. The enzyme, however, is capable of coupling non-peptidic alkyl amines to glutamine residues located in flexible loops. This has been used to modify proteins *in vitro* and more recently on the surface of cells.[40,41] To obtain site-selective modification, a transglutaminase from guinea pig was paired with "Q-tagged" proteins.[42] This particular enzyme demonstrates specificity for the glutamine substrate (peptide recognition sequences PNPQLPF, PKPQQFM, GQQQLG) but is highly promiscuous with regard to the amine substrate.

Another method for site-selective enzymatic modification relies on bacterial sortases, commonly Sortase A derived from *S. aureus*.[43] This enzyme catalyzes the cleavage of the Thr-Gly amide bond in the sequence LPXTG and then subsequently catalyzes the formation of a new amide bond between the C-terminal threonine and the N-terminus of an oligoglycine peptide. This method has been used to perform ligations of peptides, proteins, and carbohydrates.[44,45] In addition, this enzymatic method has been used for labeling *ex vivo*, on cell surfaces, and in both yeast and mammalian cells.[46–48]

An alternative method for enzymatic tagging of protein substrates relies on the fusion of the protein of interest with an enzyme that is covalently modified by its substrate. Johnsson and coworkers developed a method termed "SNAP tag" that fuses a human DNA repair protein (O^6-al-

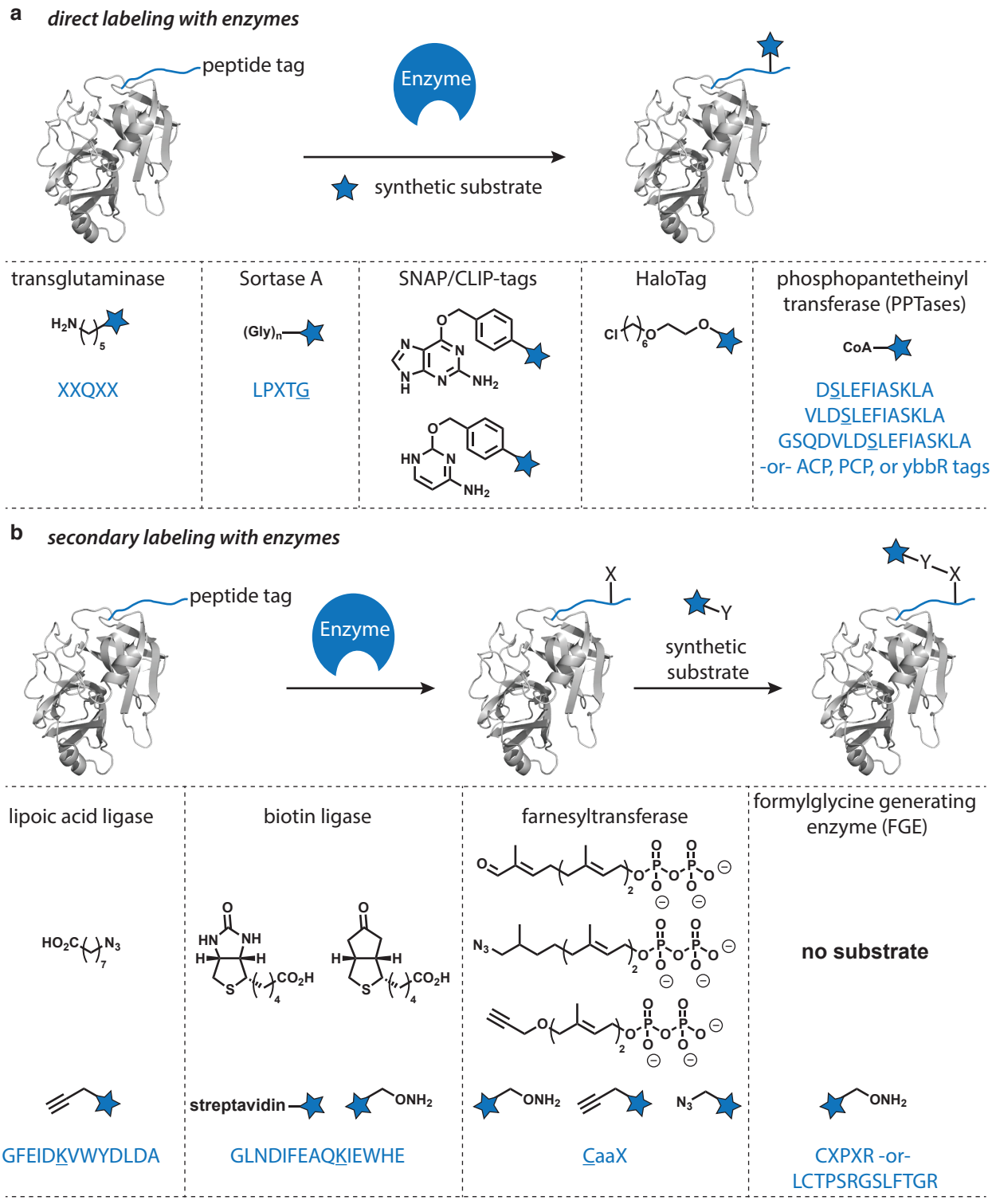


Figure 1.4. Methods for protein labeling with enzymes. (a) Direct methods for tagging artificial peptides with synthetic substrates using enzymes. Transglutaminase, Sortase A, SNAP/CLIP tags, HaloTags and PPTases are capable of direct labeling of proteins. (b) Methods for secondary tagging of artificial peptides with synthetic substrates using enzymes. Lipoic acid ligase, biotin ligase, farnesyltransferase, and formylglycine generating enzyme (FGE) are capable of appending bioorthogonal reactive groups on proteins. These reactive handles can be modified subsequently with functionalized synthetic substrates.

kylguanine-DNA alkyltransferase, hAGT) with the protein of interest.[49,50] The DNA repair protein normally repairs 6-oxo-alkylated guanosine DNA lesions by transferring the alkyl group to a cysteine residue. Co-opting this reactivity and incubating the protein fusion with O^6-benzylguanosine derivatives results in effective transfer of the benzyl group to the protein fusion. A similar, orthogonal method ("CLIP tag") has been developed for alkylated cytosine residues.[51] Additionally, Promega has developed an engineered enzyme termed "HaloTag" that is covalently modified with alkyl chlorides.[52] Mutations introduced to the active site of a haloalkane dehalogenase (DhaA) prevent catalytic turnover. The reactive intermediate is trapped to form a covalent linkage between the alkyl chloride and the enzyme.

Phosphopantetheinyl transferases (PPTases, Sfp and AcpS) naturally catalyze the transfer of a phosphopantetheinyl (Ppant) group from Coenzyme A (CoASH) to a serine residue on acyl carrier protein (ACP) or peptide carrier protein (PCP). It was discovered that the PPTases accept a wide variety of modifications at the thiol moiety of Coenzyme A and this promiscuity can be used to attach synthetic molecules to ACP or PCP tagged proteins.[53–56] One shortcoming of this method was the relatively large size of ACP/PCP (80-100 amino acids). To address this fact, Walsh and coworkers used phage display to identify a series of small peptide tags (11-17 amino acids) that are still recognized by PPTases.[57,58] These peptide fragments can be inserted at the N- or C-terminus or in flexible loops and be site-selectively modified with CoA derivatives.

Ting and co-workers repurposed biotin ligase from *E. coli* (BirA) to biotinylate proteins containing a 15-residue acceptor peptide.[59] After biotinylation, the modification can be probed with streptavidin-labeled quantum dots for fluorescence detection. It was also determined that BirA accepts a ketone-containing analogue of biotin, allowing for secondary labeling with aminooxy or hydrazide molecules (Section 1.4). To incorporate biotin analogues with other reactive handles, Ting and co-workers screened a panel of biotin ligases and identified one from *P. horikoshii* that could incorporate azide and alkyne-containing biotin substrates.[60] This allows for further modification using robust bioorthogonal chemistries (see Section 1.4).

Lipoic acid ligase from *E. coli* (LplA) has also been developed by the Ting group for non-native modification of proteins.[61] LplA normally catalyzes the acylation of lysine with lipoic acid. An azide bearing short-chain fatty acid was found to be accepted as a substrate for LplA. This azide could be modified in a secondary conjugation reaction with triarylphosphines or alkynes (see Section 1.4). Investigation of a natively lipoylated protein, enabled the identification of a small peptide (22 amino acids) that was sufficient for effective labeling by LplA. In addition, lipoic acid ligase was shown to be orthogonal to biotin ligase by the concurrent labeling of two independent proteins in the same cell.[61]

Another method for the site-selective modification of proteins relies on farnesyltransferase. Many proteins are natively farnesylated at a C-terminal cysteine, within a "CaaX-box" (where a is an aliphatic amino acid and X is one of a variety of amino acids). Many different isoprenoid pyrophosphate analogs are accepted as farnesyltransferase substrates, including alkyne,[62] azide,[63,64] aldehyde,[65,66] and biotin[67] functionalized farnesyl pyrophosphate mimetics.

Formylglycine-generating enzyme (FGE) catalyzes the conversion of a cysteine residue to formylglycine. The enzyme recognizes a cysteine within the six residues CXPXR, but inclusion of a longer 13 amino acid sequence results in improved levels of conversion to the aldehyde.[68,69] Overexpression of FGE with the tagged protein results in direct production of the aldehyde labeled protein.[70,71] The resulting aldehyde can be modified with alkoxyamine or hydrazide reagents (see Section 1.4). These enzymatic methods provide means to selectively modify protein substrates.

However, many of these methods rely on bioorthogonal reactions to attach synthetic components to the protein of interest.

1.4 Unnatural amino acids for site-selective protein modification

The ability to introduce unnatural amino acids into proteins has dramatically expanded modes of reactivity for proteins. Several methods have been developed for the *in vivo* incorporation of artificial amino acids in *E. coli* as well as in mammalian cells. Many modified analogs of amino acids are tolerated by the translational machinery in *E. coli*.[72] Simple addition of a non-canonical amino acid to the growth media allows for its non-selective incorporation into proteins. Yields can be dramatically improved by using *E. coli* that are auxotrophic for the amino acid being targeted for replacement.[73–76] Additionally, over-expression of the aminoacyl-tRNA synthetase (aaRS) for the amino acid being replaced can help improve incorporation of the unnatural amino acid.

Methods for the site-selective incorporation of non-canonical amino acids have also been developed. These methods utilize the amber stop codon (UAG) to incorporate a 21st amino acid. Schultz and co-workers have evolved orthogonal aaRSs and tRNAs that recognize the artificial amino acid and the amber stop codon, respectively.[77–79] Orthogonal pairs have been selected for dozens of unnatural amino acids. The acylated tRNA is accepted by the translational machinery and selectively incorporates the unnatural amino acid in response to the UAG codon. Co-expression of the aaRS, tRNA, and the gene of interest containing the amber mutation results in production of the protein of interest with the unnatural amino acid incorporated.[80] The artificial amino acid is generally supplemented in the growth medium, although a strain of *E. coli* has been engineered that can both biosynthesize and incorporate the unnatural amino acid *p*-aminophenylalanine.[81] More recent reports of site-selective unnatural amino acid incorporation rely on the aaRS and tRNA for pyrrolysine, enabling the incorporation of pyrrolysine analogs.[27,82–87] The aaRS seems to tolerate relatively large perturbations to the amino acid without requiring re-engineering, allowing for the incorporation of larger unnatural moieties into proteins.

Now that proteins containing unnatural functionality are readily biosynthesized, reactions that are orthogonal to the native reactivity are rapidly being developed (Figure 1.5). One of the earliest methods for bioorthogonal labeling of proteins involved reaction with the carbonyl group of ketones and aldehydes.[88,89] Introduction of a ketone or aldehyde to a protein generates a uniquely electrophilic site for modification as most of the native amino acids react as nucleophiles. While the carbonyl moiety reacts with amines to form imines, the equilibrium for this reaction lies on the side of the carbonyl in water.[90] However, reaction with α-effect amines, such as alkoxyamines and hydrazides, results in the formation of a stable product in water.[91,92] Further improving the stability of these products, Bertozzi and co-workers have developed an aminooxy reagent that first undergoes oxime formation followed by a Pictet-Spengler type reaction to form a product that is stable to acid and dilute conditions.[93]

One of the most commonly used reagents for biomolecule modification is the azide. The azide is both inert and relatively small in size, facilitating its wide adoption as a bioorthogonal functional group. Critical to this adoption was the development of mild reactions that covalently modify the azide. In seminal work, Bertozzi *et al*. demonstrated the utility of the organic azide as a bioorthogonal functional group by reaction with modified triphenylphosphines.[94] By altering the well-known Staudinger reduction to trap the aza-ylide intermediate, they were able to form an amide bond between the reduced azide and the oxidized phosphine.[95] This Bertozzi-Staudinger ligation has been

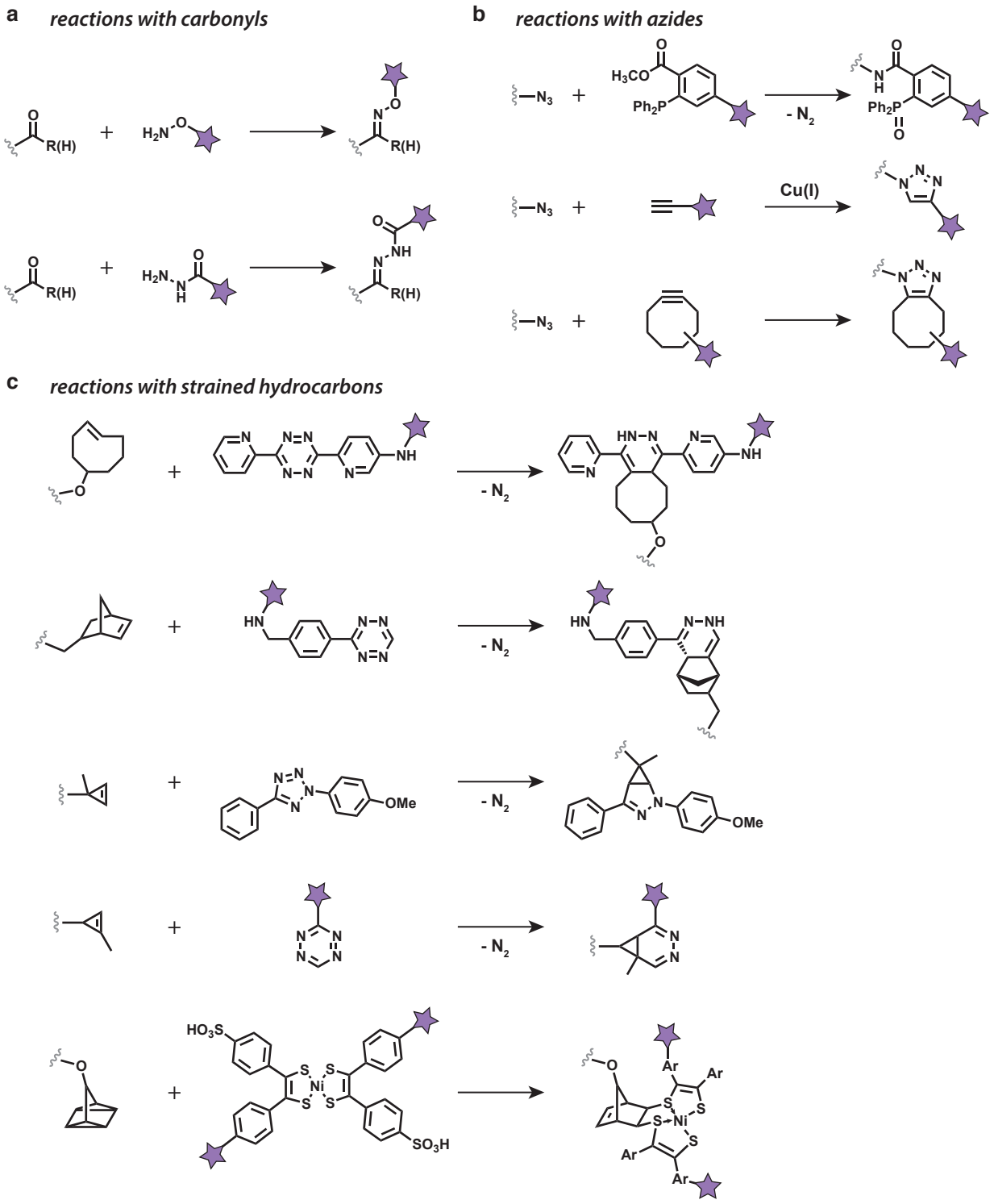


Figure 1.5. Introduction of artificial, bioorthogonal functionality to proteins allows for selective modification of the unnatural moiety. (a) Introduced aldehydes and ketones can be modified with α-effect amines. (b) Azides can be modified with triaryl phosphines *via* a Bertozzi-Staudinger ligation or with alkynes *via* a dipolar cycloaddition. (c) Strained hydrocarbons can be modified with a variety of reactive reagents. Many of these reactions can be used in conjunction with one another.

used to label proteins *in vitro*, on live cells, and even in live mice.[96] The azide was also long known to react with alkynes in a 1,3-dipolar cycloaddition, however only at elevated temperatures.[97] Meldal[98] and Sharpless[99] concurrently discovered that Cu(I) catalyzes this cycloaddition (named by Sharpless and now referred to generally as a "click" reaction) dramatically expanding the biocompatibility of the reaction. Since this discovery, the copper catalyzed azide-alkyne click reaction has been widely adopted and has been used in varied applications from organic and polymer synthesis,[100] to construction of pharmacophore libraries,[101] and biomolecule labeling.[102] Given the toxic nature of Cu(I), the copper catalyzed cycloaddition has limitations for reactions with live cells.[21] Bertozzi and co-workers developed a variant of the azide-alkyne cycloaddition that proceeds under biological conditions without the addition of Cu(I) as a catalyst.[103] Building on an earlier report of "explosive" reactivity between cyclooctyne and benzyl azide,[104] they were able to label cell surface azides by a strain-promoted cycloaddition with cyclooctyne. Since this initial report, many other strained alkynes have been developed for the selective labeling of azides.[105,106] However, some of these strained alkynes lose selectivity for the azide and show background reactivity with thiols.[107,108]

Recent work has dramatically expanded the use of strained hydrocarbons in bioconjugation reactions. Concurrent reports from Fox and Hildebrand of an inverse-electron demand Diels Alder reaction between tetrazines and strained alkenes (*trans*-cyclooctene or norbornene) revealed the utility of tetrazines for rapid, bioconjugation reactions.[109,110] The reaction of tetrazines has since been expanded to include other strained alkenes, such as 1,3-disubstituted cyclopropenes,[111] as well as unstrained alkenes.[112] Lin and co-workers first reported the 1,3-dipolar cycloaddition between alkenes and photochemically generated nitrile imines.[113] A combination of computational and experimental work has shown that this cycloaddition also occurs with 3,3-disubstituted cyclopropenes.[114,115] The highly strained quadricyclane was recently reported to undergo selective ligation with Ni bis(dithiolene) reagents in aqueous conditions as well as in the presence of cell lysate.[116] After the initial development of these strain-promoted bioorthogonal reactions, site-selective genetic incorporation of many of these reactive pairs was accomplished allowing for the synthesis of well-defined bioconjugates with these techniques.[114,117–120]

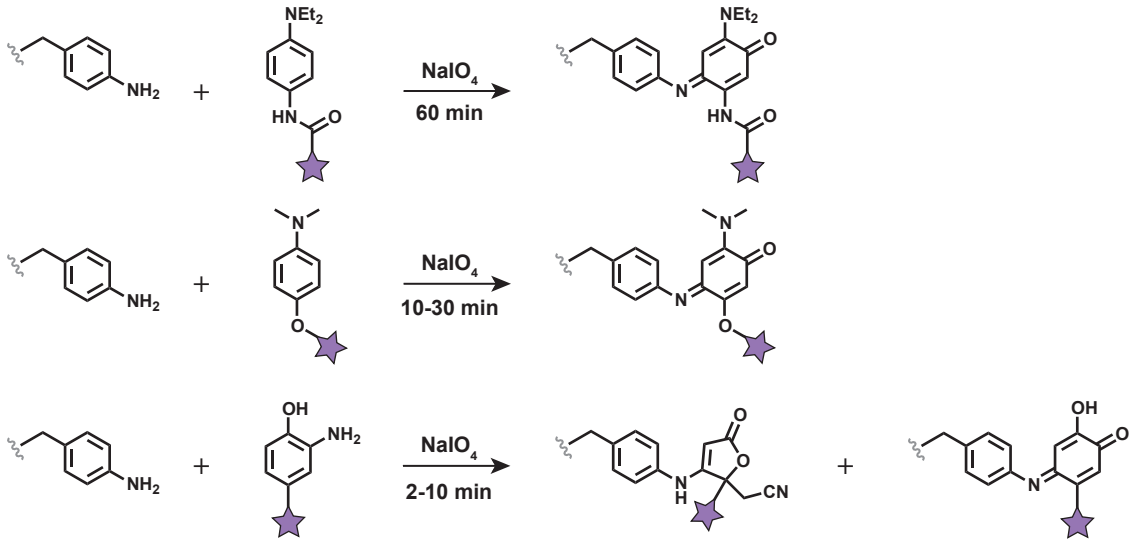


Figure 1.6. Strategies for the modification of aniline residues. Introduced *p*-aminophenylalanine residues can be modified with various electron rich aromatic rings in the presence of sodium periodate.

Our group has developed bioorthogonal reactions for the selective modification of anilines with electron-rich aromatic rings (Figure 1.6). A well-defined one-to-one coupling of *N,N*-dialkyl-*p*-phenylenediamine and aniline was accomplished with sodium periodate in just 60 min.[121,122] The coupling could be used to attach peptides, DNA, and PEG to aniline containing proteins. *N,N*-dialkyl-*p*-anisidines were also found to undergo a similar coupling in the presence of sodium periodate, but with faster reaction rates. More recently, an even faster variant of the reaction was identified. Anilines were found to rapidly couple to *o*-aminophenols using sodium periodate as an oxidant.[123] While this reaction was originally reported as a hetero-Diels Alder reaction with an acrylamide-functionalized aniline,[124] recent characterization with small molecule analogs revealed that the reaction forms a mixture of two products. These products arise from the 1,4-addition of aniline to the oxidized aminophenol. Further oxidation results in the minor iminoquinone product and the major ring-contracted butenolide product. This reaction was used to attach PEG, peptides and radionuclides to proteins.

Collectively, these methods offer a powerful way to selectively modify proteins. However, some applications, such as the modification of isolated, native proteins for tissue engineering, are not compatible with the introduction of an artificial amino acid. These limitations have led to the recent development of methods to site-selectively modify native positions on proteins.

1.5 Modification of the N-terminus for site-selective bioconjugation

Perhaps the easiest way to achieve site-selective modification is to target a unique position on a protein. The N-terminus occurs only once on all proteins and has reactivity that can be differentiated from the side-chain amino group of lysine. The pKa of the N-terminal amine is lower than that of the average lysine side-chain amine. This enables selective modification of the N-terminus by control of the reaction pH. However, achieving site-selective modification this way is difficult. More selective reactions have been developed that take advantage of the N-terminal amine and its proximity to the α-carbon and side-chain functionality of certain amino acids (Figure 1.7).

The most general method for selective N-terminal modification involves a biomimetic oxidation of the N-terminus via transamination. Transamination of the N-terminus was first reported in 1956 using pyridoxal at elevated temperatures (100 °C).[125] A subsequent report by Dixon improved the biocompatibility of the reaction by performing it at room temperature using glyoxylate, acetate, and Cu(I).[126] Further improvements were made to the reaction by our group, allowing the reaction to proceed under physiological conditions without the addition of a metal catalyst. Both pyridoxal-5'-phosphate and *N*-methylpyridinium-4-carboxaldehyde (Rapoport's salt) have been identified by our group as reagents capable of transforming the N-terminal amine into a reactive aldehyde or ketone.[127,128] Each of these reagents show an inherent preference for particular N-termini; however conditions can be screened to achieve acceptable levels of modification for nearly any N-terminal residue.[129,130] The resulting transaminated terminus can be modified with aminooxy reagents or other α-effect amines.

N-terminal Ser and Thr can also be converted to reactive aldehydes by sodium periodate oxidation.[131] When at the N-terminus these residues present a unique 1,2-aminoalcohol. Treatment with sodium periodate results in cleavage of the aminoalcohol and the resulting glyoxamide can be treated with alkoxyamine reagents to selectively label the N-terminus. Both of these methods require initial oxidation of the N-terminus to a reactive carbonyl moiety followed by reaction with a second, functionalized reagent.

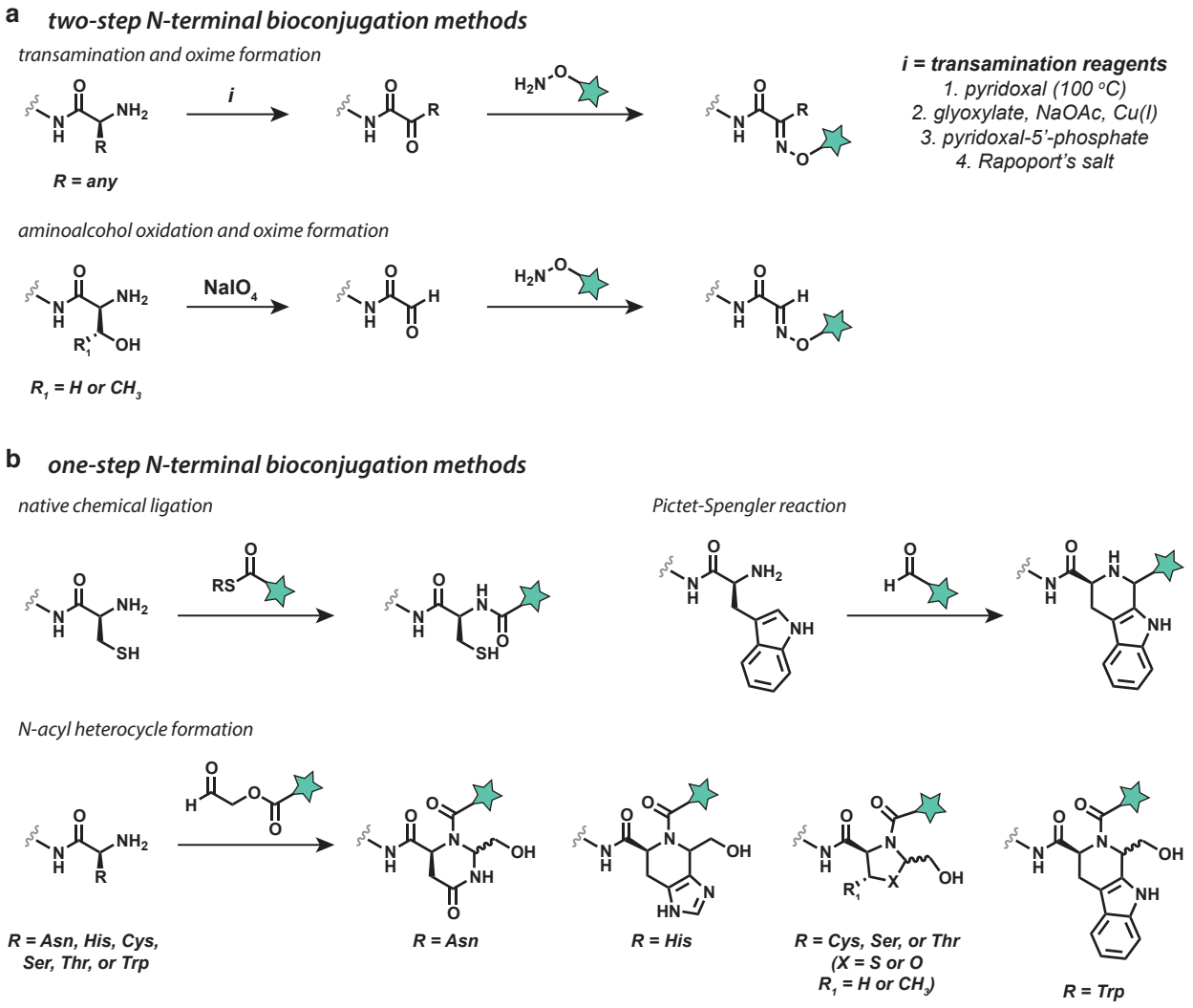


Figure 1.7. Methods for the site-selective modification of the N-terminus. (a) Oxidation of the N-terminus results in a reactive glyoxamide or pyruvamide that can be modified in a second step with alkoxyamine reagents. Several reagents are known for the selective transamination of virtually any N-terminal residue. Additionally, sodium periodate can be used to oxidize N-terminal Ser or Thr. (b) Certain N-terminal residues can be modified directly. N-terminal Cys can be modified with thioester reagents by native chemical ligation. N-terminal Trp can be modified with aldehydes via a Pictet-Spengler reaction. Acyl-aldehydes can be used to modify N-terminal Asn, His, Cys, Ser, Thr or Trp residues.

As an alternative to these two-step procedures several single-step modification strategies for the N-terminus have been developed. N-terminal Trp can be modified with aldehydes via a Pictet-Spengler reaction.[132] While the initial formation of the imine at the N-terminus is readily reversible, the neighboring nucleophilic indole side chain irreversibly intercepts the electrophile. This method provides a route for direct modification of the N-terminus. However, for proteins heterologously expressed in *E. coli* the N-terminal Met is not cleaved when Trp is the penultimate residue. Several other N-termini form stable products with aldehyde probes. Asn, His, Cys, Ser, Thr, and Trp have been shown to form stable N-acyl heterocycles or bicyclic lactams.[133] Inclusion of an acyl group on the aldehyde probe allows for formation of N-acyl heterocyclic or bicyclic lactam products with the

acyl group transferred to the N-terminal nitrogen.

A notable method for the modification of N-terminal cysteines is via reaction with thioester reagents in a process known as native chemical ligation.[134] Thioesterification between the N-terminal Cys and the thioester reagent followed by irreversible S-N acyl transfer to the N-terminal amine results in formation of a native amide bond. The specificity of this ligation method has allowed for its use in not only the modification of proteins but also in the total and semisynthesis of proteins.[135–138] Recent work in the Danishefsky group has used iterative native chemical ligation steps to synthesize the glycoprotein erythropoietin.[139]

1.6 Conclusions and thesis overview

Recent work has dramatically expanded the number of methods to modify biomolecules. Powerful, new methods for the modification of native and artificial amino acids have been developed. In conjunction, improved strategies for the incorporation of non-canonical functionality into proteins have also been developed facilitating the use of bioorthogonal chemistries. However, many of these new bioorthogonal strategies rely on substrates that are relatively unstable and require lengthy syntheses. We sought to advance alternative methods for selective protein modification. These methods would be reliant on oxidation to generate reactive, bioorthogonal substrates.

This work focuses on the development and improvement of site-selective bioconjugation reactions as well as the application of these reactions to protein-based materials and detection of post-translational modification. Briefly, an oxidative coupling reaction between anilines and *o*-aminophenols was optimized to increase compatibility with cysteine-containing and glycosylated protein substrates. We also found conditions that allowed the efficient coupling of *o*-aminophenols and native amino acids, including cysteine and N-terminal proline. These oxidative coupling reactions were used to construct a protein-based material for the optical detection of fibrous clots. Additionally, it was shown that the aniline-based oxidative coupling could be accomplished in cell lysate. This enabled the application of this reaction to the detection of post-translational tyrosine-nitration.

1.7 References

1. O'Hare, H. M.; Johnsson, K.; Gautier, A. *Curr. Opin. Struct. Biol.* **2007**, *17*, 488–494.
2. Zalipsky, S. *Bioconjugate Chem.* **1995**, *6*, 150–165.
3. Witus, L. S.; Francis, M. B. *Acc. Chem. Res.* **2011**, *44*, 774–783.
4. Giepmans, B. N. G.; Adams, S. R.; Ellisman, M. H.; Tsien, R. Y. *Science* **2006**, *312*, 217–224.
5. Lavis, L. D.; Raines, R. T. *ACS Chem. Biol.* **2008**, *3*, 142–155.
6. Deniz, A. A.; Laurence, T. A.; Beligere, G. S.; Dahan, M.; Martin, A. B.; Chemla, D. S.; Dawson, P. E.; Schultz, P. G.; Weiss, S. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 5179–5184.
7. Schuler, B.; Eaton, W. A. *Curr. Opin. Struct. Biol.* **2008**, *18*, 16–26.
8. Bücherl, C.; Aker, J.; de Vries, S.; Borst, J. W. *Methods Mol. Biol.* **2010**, *655*, 389–399.
9. Periasamy, A.; Wallrabe, H.; Chen, Y.; Barroso, M. *Methods Cell Biol.* **2008**, *89*, 569–598.
10. Jevsevar, S.; Kunstelj, M.; Porekar, V. G. *Biotechnol J* **2010**, *5*, 113–128.
11. Alley, S. C.; Okeley, N. M.; Senter, P. D. *Curr. Opin. Chem. Biol.* **2010**, **14**, 529–537.
12. Mackenzie, K. J.; Francis, M. B. *J. Am. Chem. Soc.* **2013**, *135*, 293–300.
13. Esser-Kahn, A. P.; Iavarone, A. T.; Francis, M. B. *J. Am. Chem. Soc.* **2008**, *130*, 15820–15822.

14. Flygare, J. A.; Pillow, T. H.; Aristoff, P. *Chem Biol Drug Des* **2013**, *81*, 113–121.
15. Chapman, A. P. *Adv. Drug Deliv. Rev*. **2002**, *54*, 531–545.
16. Miller, R. A.; Presley, A. D.; Francis, M. B. *J. Am. Chem. Soc*. **2007**, *129*, 3104–3109.
17. Dedeo, M. T.; Duderstadt, K. E.; Berger, J. M.; Francis, M. B. *Nano Lett*. **2010**, *10*, 181–186.
18. Tong, G. J.; Hsiao, S. C.; Carrico, Z. M.; Francis, M. B. *J. Am. Chem. Soc.* **2009**, *131*, 11174–11178.
19. Francis, M. B. In Chemical Biology; Schreiber, S. L.; Kapoor, T. M.; Wess, G., Eds.; Wiley-VCH Verlag GmbH, 2008; pp. 593–634.
20. Tilley, S. D.; Joshi, N. S.; Francis, M. B.; Begley, T. P. In Wiley Encyclopedia of Chemical Biology; John Wiley & Sons, Inc., 2007.
21. Sletten, E. M.; Bertozzi, C. R. *Angew. Chem. Int. Ed*. **2009**, *48*, 6974–6998.
22. Baslé, E.; Joubert, N.; Pucheault, M. *Chem. Biol.* **2010**, *17*, 213–227.
23. McFarland, J. M.; Francis, M. B. *J. Am. Chem. Soc*. **2005**, *127*, 13490–13491.
24. Doolittle, R. F. In Prediction of Protein Structure and the Principles of Protein Conformation; Fasman, G. D., Ed.; Springer US, 1989; pp. 599–623.
25. Smith, M. E. B.; Schumacher, F. F.; Ryan, C. P.; Tedaldi, L. M.; Papaioannou, D.; Waksman, G.; Caddick, S.; Baker, J. R. *J. Am. Chem. Soc.* **2010**, *132*, 1960–1965.
26. Bernardes, G. J. L.; Chalker, J. M.; Errey, J. C.; Davis, B. G. *J. Am. Chem. Soc*. **2008**, *130*, 5052–5053.
27. Li, Y.; Yang, M.; Huang, Y.; Song, X.; Liu, L.; Chen, P. R. *Chem. Sci*. **2012**, *3*, 2766–2770.
28. Li, Q.-F.; Yang, Y.; Maleckis, A.; Otting, G.; Su, X.-C. *Chem. Commun.* **2012**, *48*, 2704–2706.
29. Li, F.; Allahverdi, A.; Yang, R.; Lua, G. B. J.; Zhang, X.; Cao, Y.; Korolev, N.; Nordenskiöld, L.; Liu, C.-F. *Angew. Chem. Int. Ed*. **2011**, *50*, 9611–9614.
30. Lin, Y. A.; Chalker, J. M.; Floyd, N.; Bernardes, G. J. L.; Davis, B. G. *J. Am. Chem. Soc*. **2008**, *130*, 9642–9643.
31. Chalker, J. M.; Lin, Y. A.; Boutureira, O.; Davis, B. G. *Chem. Commun*. **2009**, 3714–3716.
32. Joshi, N. S.; Whitaker, L. R.; Francis, M. B. *J. Am. Chem. Soc.* **2004**, *126*, 15942–15943.
33. Ban, H.; Gavrilyuk, J.; Barbas, Carlos F. *J. Am. Chem. Soc.* **2010**, *132*, 1523–1525.
34. Ban, H.; Nagano, M.; Gavrilyuk, J.; Hakamata, W.; Inokuma, T.; Barbas, C. F. *Bioconjugate Chem*. **2013**, *24*, 520–532.
35. Tilley, S. D.; Francis, M. B. *J. Am. Chem. Soc*. **2006**, *128*, 1080–1081.
36. Seim, K. L.; Obermeyer, A. C.; Francis, M. B. *J. Am. Chem. Soc*. **2011**, *133*, 16970–16976.
37. Antos, J. M.; Francis, M. B. *J. Am. Chem. Soc*. **2004**, *126*, 10256–10257.
38. Antos, J. M.; McFarland, J. M.; Iavarone, A. T.; Francis, M. B. *J. Am. Chem. Soc*. **2009**, *131*, 6301–6308.
39. Rashidian, M.; Dozier, J. K.; Distefano, M. D. *Bioconjugate Chem*. **2013**, *24*, 1277–1294.
40. Dutton, A.; Singer, S. J. *Proc. Natl. Acad. Sci. U.S.A.* **1975**, *72*, 2568–2571.
41. Sato, H.; Ikeda, M.; Suzuki, K.; Hirayama, K. *Biochemistry* **1996**, *35*, 13072–13080.
42. Lin, C.-W.; Ting, A. Y. *J. Am. Chem. Soc*. **2006**, *128*, 4542–4543.
43. Mazmanian, S. K.; Liu, G.; Ton-That, H.; Schneewind, O. *Science* **1999**, *285*, 760–763.
44. Samantaray, S.; Marathe, U.; Dasgupta, S.; Nandicoori, V. K.; Roy, R. P. *J. Am. Chem. Soc*. **2008**, *130*, 2132–2133.
45. Mao, H.; Hart, S. A.; Schink, A.; Pollok, B. A. J. Am. Chem. Soc. 2004, 126, 2670–2671.
46. Popp, M. W.; Antos, J. M.; Grotenbreg, G. M.; Spooner, E.; Ploegh, H. L. *Nat. Chem. Biol.* **2007**, *3*, 707–708.

47. Tsukiji, S.; Nagamune, T. *ChemBioChem* **2009**, *10*, 787–798.
48. Strijbis, K.; Spooner, E.; Ploegh, H. L. *Traffic* **2012**, *13*, 780–789.
49. Keppler, A.; Pick, H.; Arrivoli, C.; Vogel, H.; Johnsson, K. *Proc. Natl. Acad. Sci. U.S.A*. **2004**, *101*, 9955–9959.
50. Keppler, A.; Gendreizig, S.; Gronemeyer, T.; Pick, H.; Vogel, H.; Johnsson, K. *Nat. Biotechnol.* **2003**, *21*, 86–89.
51. Gautier, A.; Juillerat, A.; Heinis, C.; Corrêa, I. R., Jr; Kindermann, M.; Beaufils, F.; Johnsson, K. *Chem. Biol*. **2008**, *15*, 128–136.
52. Los, G. V.; Encell, L. P.; McDougall, M. G.; Hartzell, D. D.; Karassina, N.; Zimprich, C.; Wood, M. G.; Learish, R.; Ohana, R. F.; Urh, M.; Simpson, D.; Mendez, J.; Zimmerman, K.; Otto, P.; Vidugiris, G.; Zhu, J.; Darzins, A.; Klaubert, D. H.; Bulleit, R. F.; Wood, K. V. *ACS Chem. Biol*. **2008**, *3*, 373–382.
53. Joshi, A. K.; Zhang, L.; Rangan, V. S.; Smith, S. J. Biol. Chem. 2003, 278, 33142–33149.
54. Yin, J.; Liu, F.; Li, X.; Walsh, C. T. *J. Am. Chem. Soc*. **2004**, *126*, 7754–7755.
55. George, N.; Pick, H.; Vogel, H.; Johnsson, N.; Johnsson, K. *J. Am. Chem. Soc*. **2004**, *126*, 8896–8897.
56. Johnsson, N.; George, N.; Johnsson, K. *ChemBioChem* **2005**, *6*, 47–52.
57. Yin, J.; Straight, P. D.; McLoughlin, S. M.; Zhou, Z.; Lin, A. J.; Golan, D. E.; Kelleher, N. L.; Kolter, R.; Walsh, C. T. *Proc. Natl. Acad. Sci. U.S.A*. **2005**, *102*, 15815–15820.
58. Zhou, Z.; Cironi, P.; Lin, A. J.; Xu, Y.; Hrvatin, S.; Golan, D. E.; Silver, P. A.; Walsh, C. T.; Yin, J. *ACS Chem. Biol*. **2007**, *2*, 337–346.
59. Chen, I.; Howarth, M.; Lin, W.; Ting, A. Y. *Nat. Methods* **2005**, *2*, 99–104.
60. Slavoff, S. A.; Chen, I.; Choi, Y.-A.; Ting, A. Y. *J. Am. Chem. Soc*. **2008**, *130*, 1160–1162.
61. Fernández-Suárez, M.; Baruah, H.; Martínez-Hernández, L.; Xie, K. T.; Baskin, J. M.; Bertozzi, C. R.; Ting, A. Y. *Nat. Biotechnol*. **2007**, *25*, 1483–1487.
62. Duckworth, B. P.; Zhang, Z.; Hosokawa, A.; Distefano, M. D. *ChemBioChem* **2007**, *8*, 98–105.
63. Duckworth, B. P.; Xu, J.; Taton, T. A.; Guo, A.; Distefano, M. D. *Bioconjugate Chem*. **2006**, *17*, 967–974.
64. Kho, Y.; Kim, S. C.; Jiang, C.; Barma, D.; Kwon, S. W.; Cheng, J.; Jaunbergs, J.; Weinbaum, C.; Tamanoi, F.; Falck, J.; Zhao, Y. *Proc. Natl. Acad. Sci. U.S.A*. **2004**, *101*, 12479–12484.
65. Rashidian, M.; Song, J. M.; Pricer, R. E.; Distefano, M. D. *J. Am. Chem. Soc.* **2012**, *134*, 8455–8467.
66. Rashidian, M.; Dozier, J. K.; Lenevich, S.; Distefano, M. D. *Chem. Commun*. **2010**, *46*, 8998–9000.
67. Nguyen, U. T. T.; Guo, Z.; Delon, C.; Wu, Y.; Deraeve, C.; Fränzel, B.; Bon, R. S.; Blankenfeldt, W.; Goody, R. S.; Waldmann, H.; Wolters, D.; Alexandrov, K. *Nat. Chem. Biol.* **2009**, *5*, 227–235.
68. Carrico, I. S.; Carlson, B. L.; Bertozzi, C. R. *Nat. Chem. Biol*. **2007**, *3*, 321–322.
69. Rush, J. S.; Bertozzi, C. R. *J. Am. Chem. Soc*. **2008**, *130*, 12240–12241.
70. Wu, P.; Shui, W.; Carlson, B. L.; Hu, N.; Rabuka, D.; Lee, J.; Bertozzi, C. R. *Proc. Natl. Acad. Sci. U.S.A*. **2009**, *106*, 3000–3005.
71. Rabuka, D.; Rush, J. S.; deHart, G. W.; Wu, P.; Bertozzi, C. R. *Nat. Protocols* **2012**, *7*, 1052–1067.
72. Cowie, D. B.; Cohen, G. N. *Biochim. Biophys. Acta* **1957**, *26*, 252–261.
73. Van Hest, J. C. M.; Kiick, K. L.; Tirrell, D. A. *J. Am. Chem. Soc*. **2000**, *122*, 1282–1288.

74. Kiick, K. L.; Saxon, E.; Tirrell, D. A.; Bertozzi, C. R. *Proc. Natl. Acad. Sci. U.S.A*. **2002**, *99*, 19–24.
75. Link, A. J.; Tirrell, D. A. *J. Am. Chem. Soc*. **2003**, *125*, 11164–11165.
76. Hendrickson, T. L.; de Crécy-Lagard, V.; Schimmel, P. *Annu. Rev. Biochem*. **2004**, *73*, 147–176.
77. Wang, L.; Brock, A.; Herberich, B.; Schultz, P. G. *Science* **2001**, *292*, 498–500.
78. Santoro, S. W.; Wang, L.; Herberich, B.; King, D. S.; Schultz, P. G. *Nat. Biotechnol.* **2002**, *20*, 1044–1048.
79. Chen, P. R.; Groff, D.; Guo, J.; Ou, W.; Cellitti, S.; Geierstanger, B. H.; Schultz, P. G. *Angew. Chem. Int. Ed.* **2009**, *48*, 4052–4055.
80. Hammill, J. T.; Miyake-Stoner, S.; Hazen, J. L.; Jackson, J. C.; Mehl, R. A. *Nat. Protocols* **2007**, *2*, 2601–2607.
81. Mehl, R. A.; Anderson, J. C.; Santoro, S. W.; Wang, L.; Martin, A. B.; King, D. S.; Horn, D. M.; Schultz, P. G. *J. Am. Chem. Soc.* **2003**, *125*, 935–939.
82. Srinivasan, G.; James, C. M.; Krzycki, J. A. *Science* **2002**, *296*, 1459–1462.
83. Blight, S. K.; Larue, R. C.; Mahapatra, A.; Longstaff, D. G.; Chang, E.; Zhao, G.; Kang, P. T.; Green-Church, K. B.; Chan, M. K.; Krzycki, J. A. *Nature* **2004**, *431*, 333–335.
84. Krzycki, J. A. *Curr. Opin. Microbiol*. **2005**, *8*, 706–712.
85. Polycarpo, C. R.; Herring, S.; Bérubé, A.; Wood, J. L.; Söll, D.; Ambrogelly, A. *FEBS Lett*. **2006**, *580*, 6695–6700.
86. Mukai, T.; Kobayashi, T.; Hino, N.; Yanagisawa, T.; Sakamoto, K.; Yokoyama, S. *Biochem. Biophys. Res. Commun.* **2008**, *371*, 818–822.
87. Nozawa, K.; O'Donoghue, P.; Gundllapalli, S.; Araiso, Y.; Ishitani, R.; Umehara, T.; Söll, D.; Nureki, O. *Nature* **2009**, *457*, 1163–1167.
88. Cornish, V. W.; Hahn, K. M.; Schultz, P. G. *J. Am. Chem. Soc*. **1996**, *118*, 8150–8151.
89. Mahal, L. K.; Yarema, K. J.; Bertozzi, C. R. *Science* **1997**, *276*, 1125–1128.
90. Jencks, W. P. *J. Am. Chem. Soc*. **1959**, *81*, 475–481.
91. Dirksen, A.; Hackeng, T. M.; Dawson, P. E. *Angew. Chem. Int. Ed.* **2006**, *45*, 7581–7584.
92. Dirksen, A.; Dawson, P. E. *Bioconjugate Chem.* **2008**, *19*, 2543–2548.
93. Agarwal, P.; Weijden, J. van der; Sletten, E. M.; Rabuka, D.; Bertozzi, C. R. *Proc. Natl. Acad. Sci. U.S.A*. **2013**, *110*, 46–51.
94. Saxon, E.; Bertozzi, C. R. *Science* **2000**, *287*, 2007–2010.
95. Staudinger, H.; Meyer, J. *Helvetica Chimica Acta* **1919**, *2*, 635–646.
96. Sletten, E. M.; Bertozzi, C. R. *Acc. Chem. Res*. **2011**, *44*, 666–676.
97. *Proc. Chem. Soc.* **1961**, 357–396.
98. Tornøe, C. W.; Christensen, C.; Meldal, M. *J. Org. Chem*. **2002**, *67*, 3057–3064.
99. Rostovtsev, V. V.; Green, L. G.; Fokin, V. V.; Sharpless, K. B. *Angew. Chem. Int. Ed*. **2002**, *41*, 2596–2599.
100. Meldal, M.; Tornøe, C. W. *Chem. Rev*. **2008**, *108*, 2952–3015.
101. Agalave, S. G.; Maujan, S. R.; Pore, V. S. *Chemistry – An Asian Journal* **2011**, *6*, 2696–2718.
102. Dirks, A. (Ton) J.; Cornelissen, J. J. L. M.; van Delft, F. L.; van Hest, J. C. M.; Nolte, R. J. M.; Rowan, A. E.; Rutjes, F. P. J. T. QSAR & Combinatorial *Science* **2007**, *26*, 1200–1210.
103. Agard, N. J.; Prescher, J. A.; Bertozzi, C. R. *J. Am. Chem. Soc*. **2004**, *126*, 15046–15047.
104. A. T. Blomquist, L. H. L. 2002.
105. Jewett, J. C.; Bertozzi, C. R. *Chem. Rev.* **2010**, *39*, 1272–1279.

106. Debets, M. F.; van Berkel, S. S.; Dommerholt, J.; Dirks, A. (Ton) J.; Rutjes, F. P. J. T.; van Delft, F. L. *Acc. Chem. Res.* **2011**, *44*, 805–815.
107. Van Geel, R.; Pruijn, G. J. M.; van Delft, F. L.; Boelens, W. C. *Bioconjugate Chem.* **2012**, *23*, 392–398.
108. Debets, M. F.; Hest, J. C. M. van; Rutjes, F. P. J. T. *Org. Biomol. Chem*. **2013**, *11*, 6439–6455.
109. Blackman, M. L.; Royzen, M.; Fox, J. M. *J. Am. Chem. Soc.* **2008**, *130*, 13518–13519.
110. Devaraj, N. K.; Weissleder, R.; Hilderbrand, S. A. *Bioconjugate Chem.* **2008**, *19*, 2297–2299.
111. Patterson, D. M.; Nazarova, L. A.; Xie, B.; Kamber, D. N.; Prescher, J. A. *J. Am. Chem. Soc*. **2012**, *134*, 18638–18643.
112. Niederwieser, A.; Späte, A.-K.; Nguyen, L. D.; Jüngst, C.; Reutter, W.; Wittmann, V. Angew. Chem. Int. Ed. Engl. 2013, 52, 4265–4268.
113. Song, W.; Wang, Y.; Qu, J.; Madden, M. M.; Lin, Q. *Angew. Chem. Int. Ed.* **2008**, *47*, 2832–2835.
114. Yu, Z.; Pan, Y.; Wang, Z.; Wang, J.; Lin, Q. *Angew. Chem. Int. Ed.* **2012**, *51*, 10600–10604.
115. Kamber, D. N.; Nazarova, L. A.; Liang, Y.; Lopez, S. A.; Patterson, D. M.; Shih, H.-W.; Houk, K. N.; Prescher, J. A. *J. Am. Chem. Soc*. **2013**, *135*, 13680–13683.
116. Sletten, E. M.; Bertozzi, C. R. *J. Am. Chem. Soc*. **2011**, *133*, 17570–17573.
117. Lang, K.; Davis, L.; Torres-Kolbus, J.; Chou, C.; Deiters, A.; Chin, J. W. *Nat. Chem* **2012**, *4*, 298–304.
118. Lang, K.; Davis, L.; Wallace, S.; Mahesh, M.; Cox, D. J.; Blackman, M. L.; Fox, J. M.; Chin, J. W. *J. Am. Chem. Soc*. **2012**, *134*, 10317–10320.
119. Plass, T.; Milles, S.; Koehler, C.; Schultz, C.; Lemke, E. A. *Angew. Chem. Int. Ed*. **2011**, *50*, 3878–3881.
120. Seitchik, J. L.; Peeler, J. C.; Taylor, M. T.; Blackman, M. L.; Rhoads, T. W.; Cooley, R. B.; Refakis, C.; Fox, J. M.; Mehl, R. A. *J. Am. Chem. Soc*. **2012**, *134*, 2898–2901.
121. Hooker, J. M.; Esser-Kahn, A. P.; Francis, M. B. *J. Am. Chem. Soc*. **2006**, *128*, 15558–15559.
122. Carrico, Z. M.; Romanini, D. W.; Mehl, R. A.; Francis, M. B. *Chem. Commun*. **2008**, 1205–1207.
123. Behrens, C. R.; Hooker, J. M.; Obermeyer, A. C.; Romanini, D. W.; Katz, E. M.; Francis, M. B. *J. Am. Chem. Soc*. **2011**, *133*, 16398–16401.
124. Hooker, J. M.; Kovacs, E. W.; Francis, M. B. *J. Am. Chem. Soc*. **2004**, *126*, 3718–3719.
125. Cennamo, C.; Carafoli, B.; Bonetti, E. P. *J. Am. Chem. Soc*. **1956**, *78*, 3523–3527.
126. Dixon, H. B. *Biochem. J*. **1964**, *92*, 661–666.
127. Gilmore, J. M.; Scheck, R. A.; Esser-Kahn, A. P.; Joshi, N. S.; Francis, M. B. *Angew. Chem. Int. Ed*. **2006**, *45*, 5307–5311.
128. Witus, L. S.; Netirojjanakul, C.; Palla, K. S.; Muehl, E. M.; Weng, C.-H.; Iavarone, A. T.; Francis, M. B. *J. Am. Chem. Soc*. **2013**.
129. Scheck, R. A.; Dedeo, M. T.; Iavarone, A. T.; Francis, M. B. *J. Am. Chem. Soc*. **2008**, *130*, 11762–11770.
130. Witus, L. S.; Moore, T.; Thuronyi, B. W.; Esser-Kahn, A. P.; Scheck, R. A.; Iavarone, A. T.; Francis, M. B. *J. Am. Chem. Soc.* **2010**, *132*, 16812–16817.
131. Geoghegan, K. F.; Stroh, J. G. *Bioconjugate Chem*. **1992**, *3*, 138–146.
132. Li, X.; Zhang, L.; Hall, S. E.; Tam *, J. P. *Tetrahedron Letters* **2000**, *41*, 4069–4073.
133. Tam, J. P.; Yu, Q.; Miao, Z. *Peptide Science* **1999**, *51*, 311–332.
134. Dawson, P. E.; Muir, T. W.; Clark-Lewis, I.; Kent, S. B. *Science* **1994**, *266*, 776–779.

135. Yamazaki, T.; Otomo, T.; Oda, N.; Kyogoku, Y.; Uegaki, K.; Ito, N.; Ishino, Y.; Nakamura, H. *J. Am. Chem. Soc*. **1998**, *120*, 5591–5592.
136. Dawson, P. E.; Kent, S. B. *Annu. Rev. Biochem*. **2000**, *69*, 923–960.
137. Camarero, J. A.; Shekhtman, A.; Campbell, E. A.; Chlenov, M.; Gruber, T. M.; Bryant, D. A.; Darst, S. A.; Cowburn, D.; Muir, T. W. *Proc. Natl. Acad. Sci. U.S.A*. **2002**, *99*, 8536–8541.
138. Schwarzer, D.; Cole, P. A. *Curr. Opin. Chem. Biol.* **2005**, *9*, 561–569.
139. Wang, P.; Dong, S.; Brailsford, J. A.; Iyer, K.; Townsend, S. D.; Zhang, Q.; Hendrickson, R. C.; Shieh, J.; Moore, M. A. S.; Danishefsky, S. J. *Angew. Chem. Int. Ed.* **2012**, *51*, 11576–11584.

Chapter 2
Oxidative coupling to anilines on proteins

Abstract

This chapter describes the development of an alternative oxidative coupling strategy for the modification of aniline residues. Using a small molecule based screen, ferricyanide was identified as a mild and efficient oxidant for the coupling of anilines and *o*-aminophenols on protein substrates. This reaction is compatible with thiols and 1,2-diols, allowing its use in the creation of complex bioconjugates for use in biotechnology and materials applications. This methodology was applied to the dual modification of a viral capsid. The coupling reaction was also demonstrated on a glycoprotein substrate.

Portions of the work described in this chapter have been reported in a separate publication.

2.1 Importance of site-selective bioconjugation reactions

The synthetic modification of proteins is a critical aspect of chemical biology and biomaterials science. Synthetically modified proteins are used to study biochemical function,[1] modulate pharmacokinetics,[2] and construct new materials with applications in drug-delivery[3–5] and targeted imaging.[6,7] Many of these applications require consistent, well-defined modifications that do not perturb the native protein structure or function. The precise modification of proteins, however, presents a significant chemical challenge as the modification must occur at ambient temperature, near neutral pH and in the presence of a wide variety of unprotected functional groups.

As discussed in the Chapter 1, several methods for the site-selective modification of proteins have been developed.[8–10] These methods either rely on natural amino acids found in low abundance, typically targeting the nucleophilic thiolate side chain of cysteine,[8,11,12] or they rely on engineered artificial amino acids.[10] The latter approach can result in completely site selective modification, but it depends critically on bioorthogonal[13] chemical reactions that can modify the functional groups with exquisite selectivity. Ketones,[14,15] azides,[16,17] strained alkenes,[18,19] alkynes,[20] and anilines[21] are commonly targeted in these reactions. These strategies are particularly useful for preparing proteins that are modified in multiple locations for use in biophysical and materials applications. In our own work,[22] the modification of proteins at two distinct locations is required for varied applications from light harvesting[23,24] to drug-delivery[4,5] and water remediation.[25] Each of the complex biomolecule targets in these studies requires bioorthogonal methods that are compatible with cysteine chemistry.

Recent work in our group has explored the oxidative coupling reaction of aniline side chains with electron-rich aromatic coupling partners, such as *o*-aminophenols.[21,26,27] These reactions proceed with high coupling rates even when low concentrations of reactants are used. Previous studies demonstrated the use of sodium periodate as an oxidant for highly efficient coupling. However the ability of periodate to oxidize other moieties on proteins, notably cysteines and 1,2-diols found in glycans, may limit the scope of these coupling reactions. Additionally, the periodate-mediated coupling forms a mixture of two products, which complicates analysis and may prevent its use in applications such as antibody-drug conjugates that require well-defined linkages. In this work, we report a substantially milder oxidant, potassium ferricyanide, that is capable of coupling anilines and *o*-aminophenols on protein substrates without oxidizing thiols or 1,2-diols (Figure 2.1). Most notably, a single, stable reaction product is obtained with this oxidant. We demonstrate the use of this reaction in conjunction with thiol-maleimide chemistry as well as on a complex glycoprotein substrate. These optimized conditions should allow for the wide adoption of this coupling chemistry to virtually any bioconjugation target. In addition, the ease of substrate synthesis, as well as the low cost and facile removal of potassium ferricyanide, provide an opportunity for using this chemistry on larger scale.

2.2 Identification of alternative oxidants with a small molecule screen

Our efforts to identify milder reaction conditions began with an HPLC-based assay to screen the efficiency of different oxidants to couple *p*-toluidine to electron-rich coupling partners, such as 2-amino-*p*-cresol. Reactions were run with 0.1 mM of each coupling partner and 1 mM oxidant at near neutral pH (6.5-7.2) to mimic the conditions compatible for use on most biomolecules. Reactions were quenched with excess tris(2-carboxyethyl)phosphine (TCEP, 5-10 mM), and an internal

standard, *p*-toluenesulfonic acid (*p*-TsOH, 0.2 mM), was added for reliable quantitation. Many oxidants screened (Figure 2.1) were capable of coupling the small molecules, and more importantly generated compound **1** as the sole product. The oxidants formed the [A+B] product with varying levels of conversion, requiring anywhere from 15 min to 18 h to reach completion (Figure 2.1). We found Ag(I), Ce(IV), Cu(II), and Fe(III) were particularly promising and potentially suitable for protein substrates.

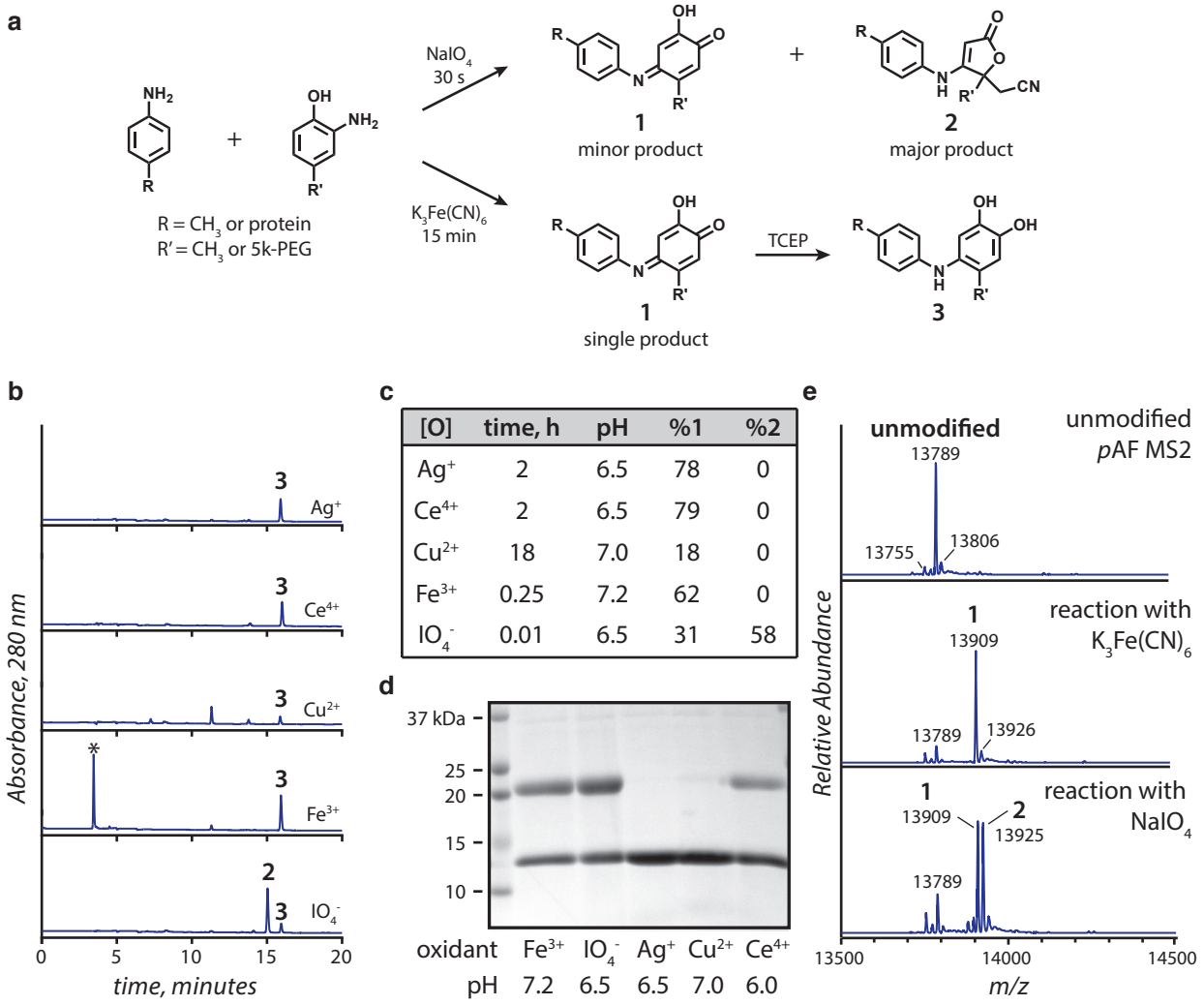


Figure 2.1. (a) Schematic for the oxidative coupling of anilines and *o*-aminophenols. (b) Reverse-phase HPLC traces of the reactions with alternative oxidants. Oxidants were screened by HPLC to determine suitable alternatives to $NaIO_4$. The reaction was quenched by the addition of TCEP, which results in the reduction of iminoquinone product **1** to the corresponding hydroquinone **3**. The asterisk denotes a peak that corresponds to the reduced oxidant. (c) Conversion was calculated using *p*-TsOH as an internal standard. (d) The same oxidants as in (b) were tested for their ability to mediate the coupling reaction on proteins using T19*p*AF MS2 and *o*-aminophenol 5k-PEG as model substrates. (e) LC-MS analysis of MS2 monomers modified with 2-amino-*p*-cresol shows that only one product is formed when ferricyanide is used as the oxidant, whereas the use of periodate resulted in a mixture of two different products.

2.3 Application to proteins

We next determined the applicability and biocompatibility of these oxidants on a protein sub-

strate containing one aniline side chain (*p*-aminophenylalanine, *p*AF) in the primary sequence (Figure 2.1). The aniline side chain (T19*p*AF) was introduced into the coat protein of the MS2 viral capsid using amber stop codon suppression (Figure 2.2a).[28–30] The MS2 capsid self assembles from 180 sequence identical monomers to form a hollow spherical structure.[31,32] The capsid provides a scaffold for the construction of targeted imaging agents or drug delivery vehicles. Use of the capsids in this capacity for applications in cardiovascular imaging will be discussed in Chapter 4. For this methodology development, reactions were performed on the assembled capsid, but analysis was carried out on the monomers after disassembly. $K_3Fe(CN)_6$ and $Ce(NH_4)_2(NO_3)_6$ (CAN) maintained high reactivity on the protein substrate, whereas $AgCO_2CF_3$ and $CuSO_4$ showed poor reactivity on proteins. This was likely due to precipitation and slower reaction rates, respectively. While Ag(I) was found to be compatible with some buffers (such as HEPES), the strict requirement to remove all chloride ions limits the potential of this oxidant. Cu(II) exhibited the lowest levels of conversion for both the small molecule and protein substrates. Of the oxidants tested, Cu(II) had the lowest oxidation potential, and appeared to be unable to oxidize the *o*-aminophenol coupling partner efficiently.

Ferricyanide ($K_3Fe(CN)_6$) was chosen for further development, as it was a mild oxidant and was compatible with proteins and most buffers.[33–36] Coupling a small molecule, 2-amino-*p*-cresol (80 μM) to the protein substrate (20 μM) with ferricyanide (1 mM) for 20 min showed complete modification to a single product by LC-MS (Figure 2.1c). When the reaction was run under the same conditions with periodate, an equal mixture of two products was observed.

Initial observations from the HPLC reactivity screen indicated that ferricyanide formed the same iminoquinone minor product (**1**) as the periodate reaction. When using $K_3Fe(CN)_6$, HPLC analysis of the reactivity of both 2-amino-*p*-cresol and 4-methylcatechol with *p*-toluidine identified the same product. Two dimensional NMR analyses and high-resolution mass spectrometry were used to confirm that ferricyanide formed product **1** (Figure 2.11). Iminoquinone **1** could be reduced to corresponding hydroquinone **3** by TCEP, as was the case with the $NaIO_4$ associated product.[21] This reduction was observed on small molecule and protein substrates; however, over time the hydroquinone was observed to re-oxidize to the iminoquinone in air.

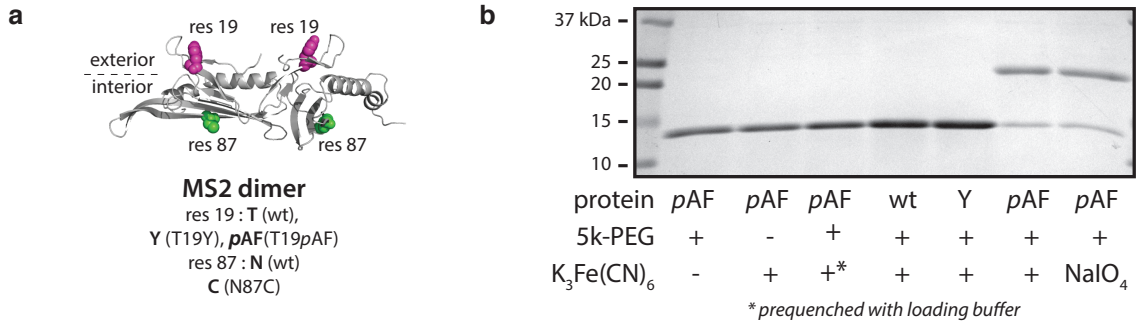


Figure 2.2. (a) Representation of the MS2 coat protein dimer, indicating the location of residues 19 and 87 based on PDB ID: 2MS2. MS2 substrates tested contained a Thr, Tyr, or *p*AF residue at position 19 and an Asn or Cys at position 87. (b) The coupling reaction was quenched by the addition of loading buffer (with DTT) and analyzed by SDS-PAGE. The lower band corresponds to unmodified MS2 monomers and the upper band corresponds to MS2 monomers modified with 5k-PEG. The coupling only took place when all three components were present.

The selectivity of the reaction was tested using several mutants of the MS2 coat protein: wild-type (wt), T19Y, and T19*p*AF (Figure 2.2). Coupling was only observed when the oxidant, *o*-ami-

nophenol, and aniline all were present (Figure 2.2b). In addition, the yield was comparable to the reaction mediated by periodate. At higher pH, some background reactivity with native amino acids was observed. However, the coupling was selective for the *p*AF residue when the reaction pH was held between 6.0 and 6.5. This is likely because anilines are deprotonated under these conditions, rendering them uniquely nucleophilic. Additionally, it was found that addition of 1-10 mM imidazole could further prevent any undesired non-specific reactivity (Figure 2.3). The nature of this side reactivity will be discussed in detail in Chapter 3 as we have optimized conditions for efficient coupling to native amino acids.

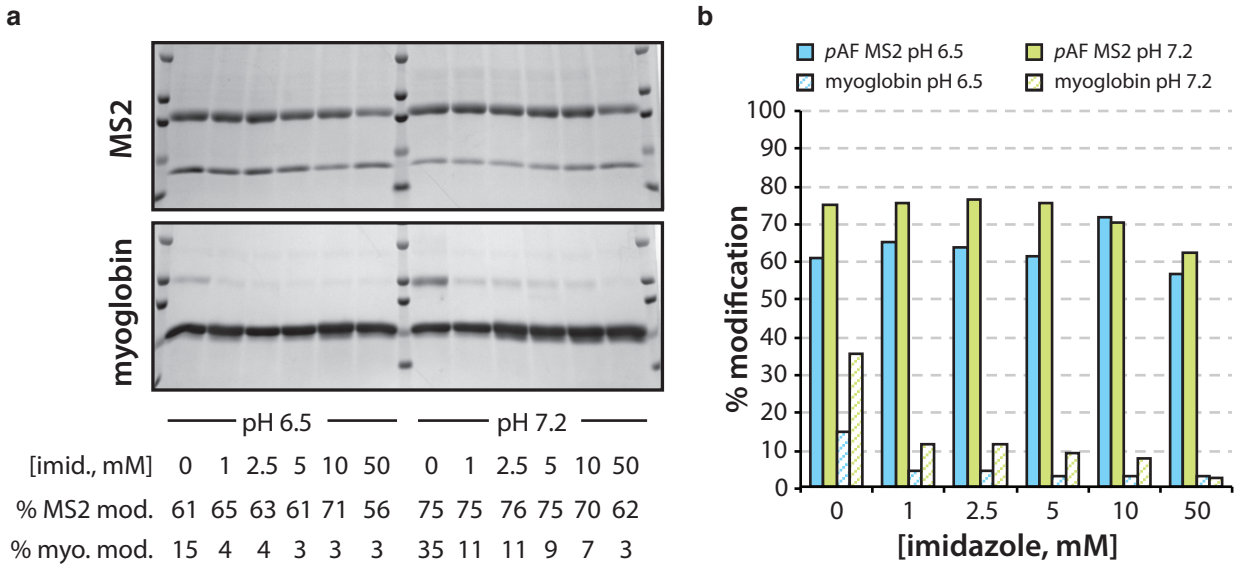


Figure 2.3. Background reactivity with native amino acids was observed. (a) SDS-PAGE analysis of *p*AF MS2 and myoglobin modified with *o*-aminophenol 5k-PEG at either pH 6.5 or 7.2 revealed low levels of background reactivity. Imidazole was added to inhibit the background reactivity. (b) Graph of the background reactivity shows that 1-10 mM imidazole impedes the reaction with native amino acids while allowing the desired oxidative coupling to occur.

2.4 Optimization of the reaction conditions on proteins

Other iron(III) sources were evaluated for their ability to perform the coupling (Figure 2.4). Most ferric salts tested were poorly soluble in water or rapidly formed insoluble iron oxides. Ferrocenium ion was also able to couple aminophenols and anilines, albeit with lower yield, because it was only temporarily stable in water.[37] Optimization of the equivalents, time, pH and buffer of the coupling reaction revealed that high levels of conversion (>75%) could be achieved with only 2.5-5 equiv of the aminophenol substrate in 15-20 min. In addition, neither the reaction pH nor the buffer salt was found to have an effect on the efficiency of the reaction (Figure 2.5). However, caution should be taken when running the reaction at higher pH as the reaction may lose selectivity for the aniline side chain under more basic conditions. Under the optimized conditions, a high level of modification (~85%, 150 modifications per capsid) could be achieved when 20 μM protein and 100 μM *o*-aminophenol polyethylene glycol (PEG) were incubated with 1 mM $K_3Fe(CN)_6$ at pH 6.0 for 20 min. To reach these high levels of conversion, it was critical to purify the *o*-aminophenol substrate thoroughly as well as store the purified substrate at -20 °C before use (see the Materials

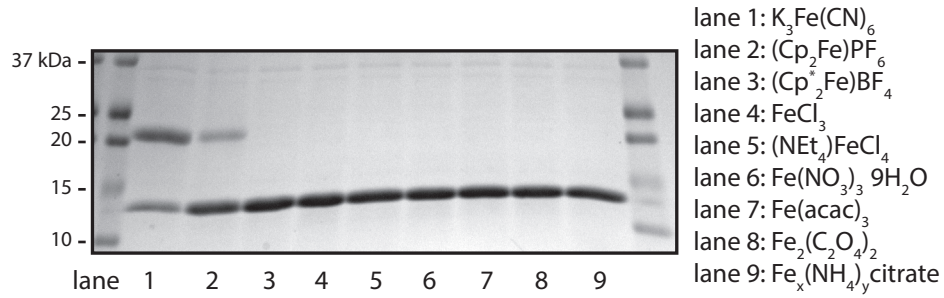


Figure 2.4. Several iron(III) sources were tested for their ability to couple *o*-aminophenols and anilines. Most Fe(III) sources were either poorly soluble or unstable in water. Only ferricyanide and ferrocenium were competent oxidants for the coupling.

and Methods section for details on purification).

Despite the mild nature of ferricyanide, we wanted to confirm that excess oxidant could be completely removed from the bioconjugation reaction.[36,38,39] Using standard biomolecule purification techniques, such as gel filtration and ion exchange, it was possible to remove all detectable iron (< 0.1 μM, Figure 2.6). Chelex resin, which is commonly used to remove divalent cations from aqueous solutions, did not remove excess ferricyanide. This is likely due to the inability of the car-

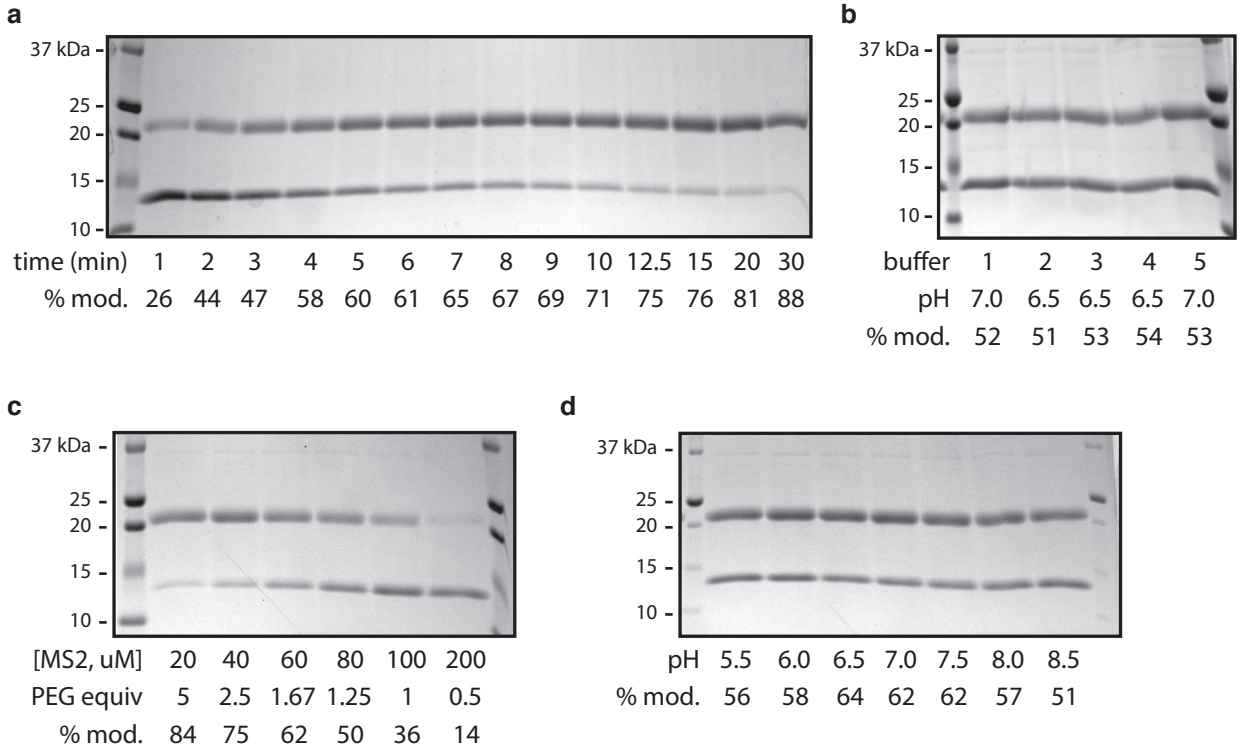


Figure 2.5. Gel analyses of reaction optimization screens on model substrate *p*AF MS2. (a) A screen of the reaction time indicated that the reaction reaches completion within 20-30 min at pH 6.5. (b) Additionally, a screen of common buffer salts showed that the coupling reaction is insensitive to the buffer salt (1 = phosphate, 2 = HEPES, 3 = MOPS, 4 = Tris, 5 = TEA). (c) Optimization of the equiv of *o*-aminophenol revealed that high levels of conversion could be achieved with 2.5-5 equiv *o*-aminophenol. (d) A screen of the buffer pH (from 5.5-8.5) demonstrated that the reaction is not sensitive to pH. The varying levels of conversion between the screens can be attributed to the different levels of purity of the *o*-aminophenol PEG.

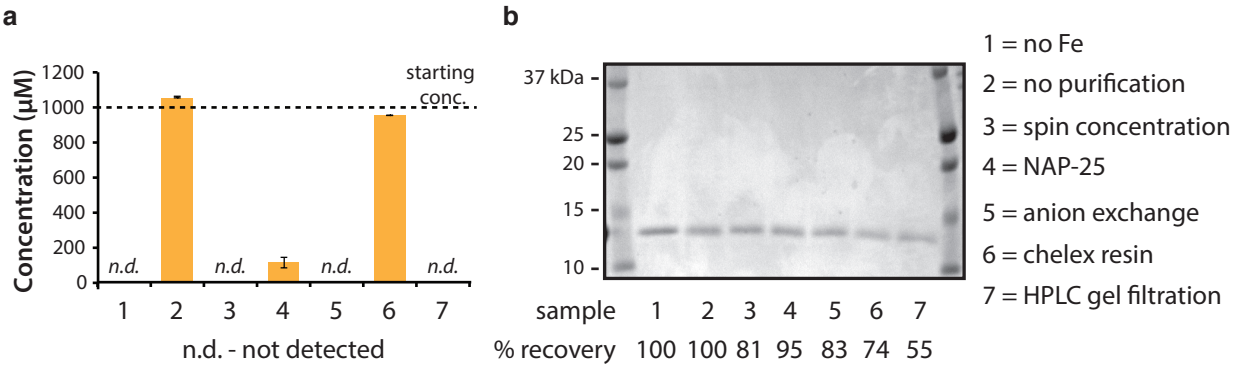
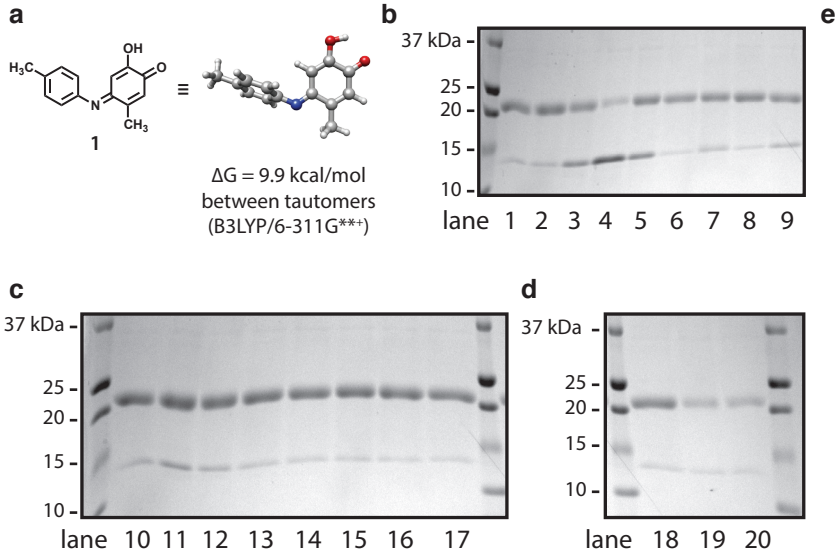


Figure 2.6. (a) ICP-OES analysis of iron removal after the reaction. Standard methods of biomolecule purification sufficiently removed excess iron. Chelex resin, however, was not effective as it was not able to bind Fe^{3+} that remain coordinated to the cyanide ligands. (b) SDS-PAGE analysis of the samples subjected to ICP-OES.

boxylic acid groups on the resin to displace the strongly coordinating cyanide ligands.

2.5 Computational studies

Computational studies of **1** indicated a strong preference (~10 kcal/mol) for the iminoquinone tautomer shown relative to the *o*-quinone structure (Figure 2.7a). This was confirmed by NMR, as only **1** was observed. Despite the presence of the imine moiety, iminoquinone **1** was found to be resistant to hydrolysis and relatively stable with respect to nucleophilic attack. We attribute this stability to a combination of two factors: (1) the imine carbon is rendered poorly electrophilic by



e

lane	conditions	% modified
1	initial	82
2	benzyl-ONH_2	87
3	hydrazine	57
4	*p*-anisidine	22
5	*p*-toluidine	61
6	glutathione	90
7	cysteine	78
8	TCEP	89
9	dithionite	73
10	initial	87
11	pH 4.0	81
12	pH 5.0	84
13	pH 6.0	88
14	pH 7.0	90
15	pH 8.0	92
16	pH 9.0	93
17	pH 10.0	90
18	initial	88
19	37 °C	79
20	50 °C	74

Figure 2.7. (a) Computational studies and NMR analysis indicated that product **1** has a strong preference for the tautomer shown. (b,c,d,e) The stability of the product was tested on MS2 after modification with *o*-aminophenol 5k-PEG. The capsids were treated with 10 mM additive for 18 h. SDS-PAGE, followed by Coomassie staining and densitometry analysis (±5% accuracy), indicated a high degree of stability in the presence of reducing agents, acidic and basic pH, and high temperature. Treatment with nucleophilic amines, however, resulted in signifcant reversal of the addition.

a high degree of electron donation from the alcohol lone pairs and (2) DFT calculations indicated that the aromatic ring of the aniline is oriented perpendicular to the rest of the structure, which aligns the sp^2-hybridized nitrogen lone pair with the π-system. This latter effect results in a significant reduction in the nitrogen atom basicity, and thus protonation of the imine group is disfavored. The stability of the product was assayed by subjecting purified, modified protein (~80% modified) to a range of nucleophilic and reducing additives, as well as acidic and basic pH (10 mM for 18 h at room temperature) and increased temperature. Only in the presence of competing nucleophilic amines, such as hydrazine, *p*-anisidine, and *p*-toluidine, was appreciable loss of product observed (Figure 2.7). Exposure to a wide range of physiologically relevant pH values (4.0-10.0), glutathione, or increased temperature (37 and 50 °C) for 18 h did not result in product loss, indicating the relative stability of the product (Figure 2.7). Additionally, no loss in product was observed after 7 days of storage at room temperature at neutral pH.

2.6 Compatibility of ferricyanide coupling with cysteine

The ferricyanide-mediated coupling was also evaluated for its compatibility with cysteine chemistry. In addition to the important biological role of cysteine, it is frequently a target for protein modification.[11,12] We tested the ability of the ferricyanide-mediated coupling to be used in conjunction with cysteine-maleimide chemistry following the scheme outlined in Figure 2.8a. A cysteine introduced to the interior surface of bacteriophage MS2 capsids[7] was first modified with a fluorescent maleimide (Alexa Fluor 488 C_5-maleimide) and then subjected to the oxidative coupling conditions. PEG and DNA *o*-aminophenol containing substrates were coupled to the fluorescently labeled viral capsid (Figure 2.8b,c). In 20 min, up to 150 copies of the aminophenol substrate could be attached to the capsids while using only 5 equiv of the coupling partner.

We also confirmed that unmodified cysteines were still reactive after the oxidative coupling step. T19*p*AF N87C MS2 capsids were first reacted with *o*-aminophenol 5k-PEG using either potassium ferricyanide or sodium periodate, and were subsequently treated with a fluorescent maleimide. After exposure to periodate, the cysteine no longer reacted with the maleimide; however, after oxidative coupling with ferricyanide, the cysteine was successfully labeled with the fluorophore (Figure 2.8d). To rule out the possibility that this reactivity was seen because the 5k-PEG substrate was too large to diffuse into the interior of the capsids, we also verified that the cysteine maintained reactivity after oxidative coupling with a small molecule aminophenol (Figure 2.8e). A fluorescent rhodamine aminophenol substrate (**7**) was synthesized and reacted with T19*p*AF N87C MS2 capsids. A spectrally separated fluorescent maleimide (Alexa Fluor 680 C_2-maleimide) was then used to assay the reactivity of the cysteine. The modified protein was analyzed by SDS-PAGE with two-color fluorescence detection. Only when ferricyanide was used as the oxidant was the thiol moiety still reactive after the oxidative coupling step (Figure 2.8e).

2.7 Compatibility of ferricyanide oxidative coupling with 1,2-diols

The mild nature of ferricyanide increases the compatibility of the oxidative coupling reaction with a broader scope of protein targets. Glycosylated proteins are attractive targets for modification with synthetic molecules, with antibody-drug conjugates serving as a prominent example.[3,40,41] While sodium periodate can be used to modify glycoproteins, it is also known to oxidize the 1,2-diols found in sugars.[42,43] To test the ability of ferricyanide to modify glycosylated proteins with-

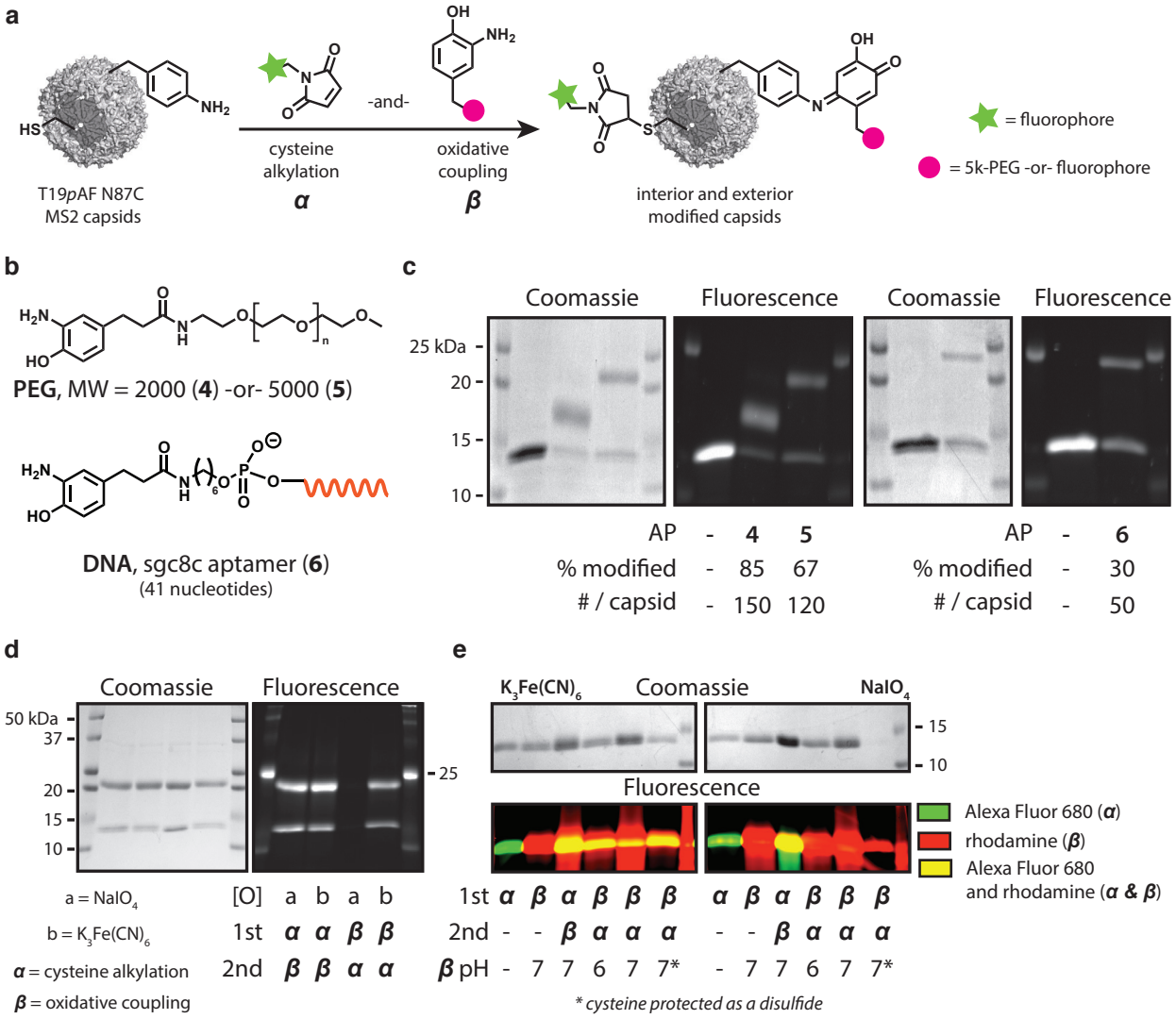


Figure 2.8. (a) Schematic for the modification of both cysteine and *p*AF residues on the interior and exterior surfaces of bacteriophage MS2 capsids, respectively. (b) Aminophenol coupling partners tested with the ferricyanide-mediated oxidative coupling. (c) Coomassie-stained and fluorescent images of SDS-PAGE gels of functionalized MS2. The MS2 capsids were first labeled with Alexa Fluor 488 C_5-maleimide followed by modification with the *o*-aminophenol substrates shown in (b). The shifts in apparent molecular weight correspond to the addition of PEG or DNA. (d) MS2 capsids were first treated with a fluorescent maleimide for cysteine alkylation (**α**) or *o*-aminophenol PEG for oxidative coupling (**β**). The modified proteins were then subjected to the other conditions (either **α** or **β**). The modified proteins were analyzed by SDS-PAGE. The cysteine was not reactive if it was first treated with periodate. (e) Coomassie-stained and fluorescent images of SDS-PAGE gels of MS2 modified with a fluorescent maleimide (**α**) and a fluorescent *o*-aminophenol (**β**) under various conditions. The oxidative coupling reaction was carried out with $K_3Fe(CN)_6$ or $NaIO_4$.

out this side reaction, an aniline moiety was first site-selectively introduced on the N-terminus of an engineered antibody fragment (Fc). The Fc was transaminated using pyrodixal-5'-phosphate (PLP), generating a uniquely reactive ketone at each N-terminus (Figure 2.9a).[43–45] This ketone was then modified with an alkoxyamine-functionalized aniline for subsequent oxidative coupling. The aniline-Fc, **10**, was successfully PEGylated with a 5 kDa-PEG. The Fc was approximately 40% modified when either periodate or ferricyanide was used to couple the *o*-aminophenol PEG to the Fc (Figure 2.9b). The lower level of modification was likely due to increased sterics as the N-termini

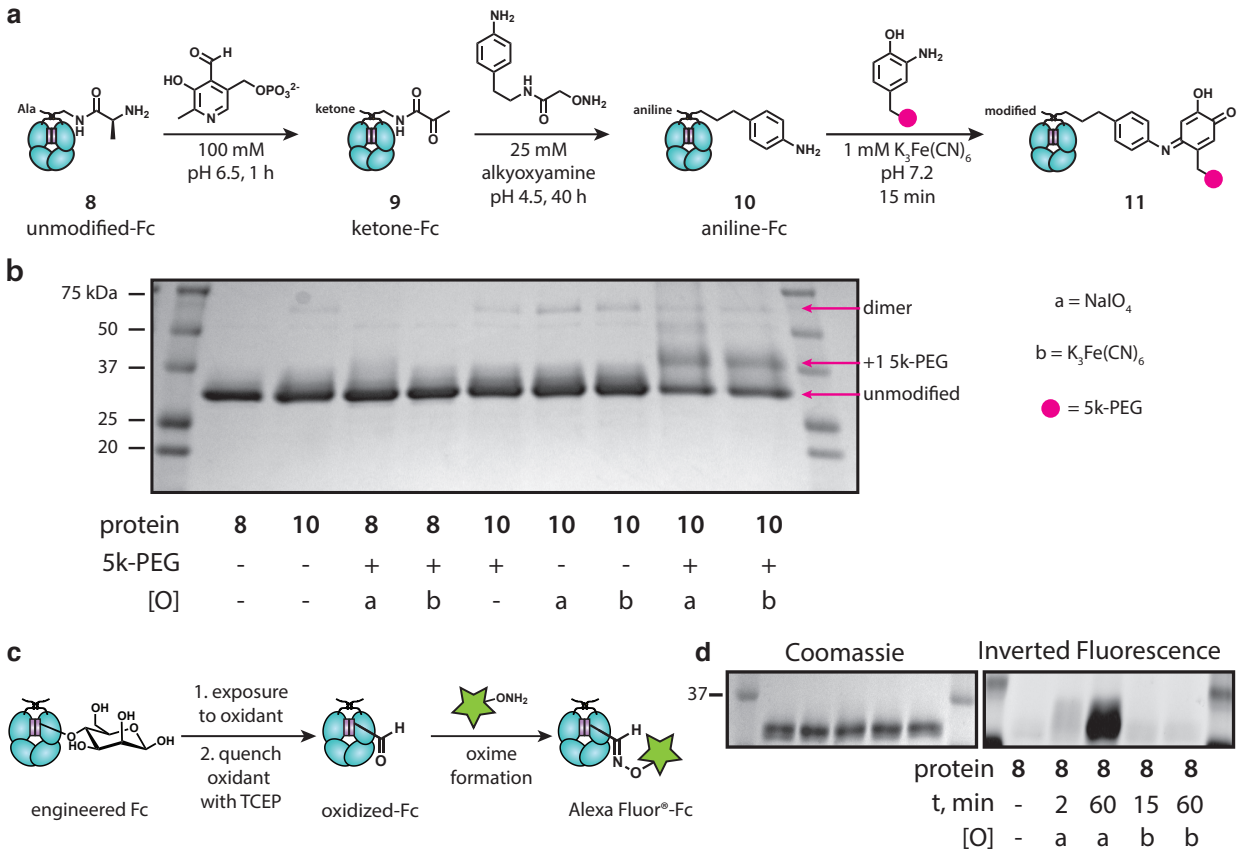


Figure 2.9. (a) Schematic for the modification of anilines introduced via transamination of the N-terminus. The terminus was first converted to a ketone with pyridoxal-5'-phosphate (PLP). This ketone was modified with an alkoxyamine-functionalized aniline. The aniline then underwent oxidative coupling with *o*-aminophenol subtrates. (b) An engineered antibody fragment (Fc) was subjected to the conditions shown in (a). Analysis of the modification with *o*-aminophenol 5k-PEG by SDS-PAGE demonstrated the ability of ferricyanide to modify a complex glycoprotein substrate. (c) The Fc sample was treated with either periodate or ferricyanide followed by a fluorescent alkoxyamine to probe for reactive, oxidized sugars.

of the two Fc chains are in close proximity to each other.

The undesired glycan oxidation was probed by treating the Fc with either $NaIO_4$ or $K_3Fe(CN)_6$ followed by a fluorescent alkoxyamine (Figure 2.9c).[43] After the oxidants were quenched with TCEP at the indicated time, the oxidant treated protein was incubated with an aminooxy dye to probe for reactive aldehydes formed as a byproduct of sugar oxidation. Exposure to periodate for 2 min resulted in a slight amount of oxidation of the glycan, with extensive oxidation observed after 1 h of exposure. No oxidation of the Fc was observed after treatment with ferricyanide even after 1 h of incubation. Glycopeptide standards were also used to assess the reactivity of periodate and ferricyanide toward glycosylated substrates. Treatment of a glycosylated erythropoietin fragment (EPO, 117-131) with $K_3Fe(CN)_6$ (1 mM) resulted in no oxidation of the GalNAc residue over the course of 1 h (Figure 2.10a). Incubation of $NaIO_4$ (1 mM) with glycosylated-EPO, however, resulted in rapid oxidation of the *O*-linked sugar with complete oxidation observed in 1 h. To investigate the compatibility of ferricyanide with 1,2-diols further, several additives were screened for their effect on reactivity. Mannose, glucose, and glycerol had little to no effect on the ferricyanide-mediated coupling even at concentrations of 1 M (Figure 2.10b). Periodate reactivity was significantly

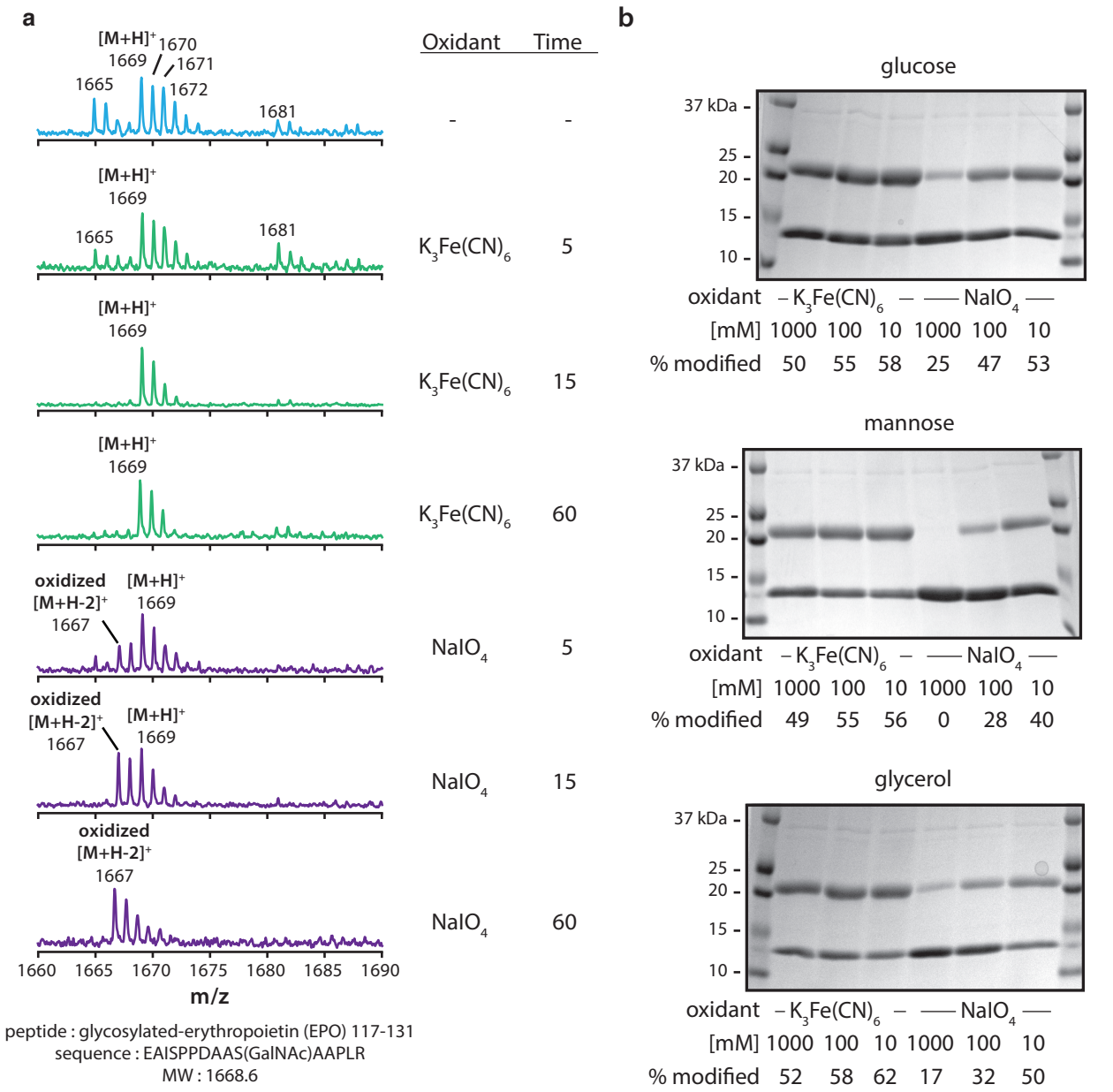


Figure 2.11. (a) Time course of glycopeptide oxidation with periodate and ferricyanide. MALDI-TOF MS shows that ferricyanide does not oxidize the glycopeptide standard over a period of 1 h. Exposure to $NaIO_4$ results in sugar oxidation after 5 min, with complete oxidation observed after 1 h. (b) The effect of 1,2-diols on both the ferricyanide and periodate-mediated oxidative couplings was analyzed by SDS-PAGE. Ferricyanide is compatibile with mannose, glucose, and glycerol, whereas the periodate-mediated reaction is inhibited at high concentrations of all three additives.

quenched in the presence of a vast excess of 1,2-diols. However, at lower concentrations (10 mM) these additives could be used to protect glycoproteins from over-oxidation by periodate while still allowing the desired oxidative coupling to take place, as previously reported.[43]

To optimize this oxidative coupling methodology, we used a small molecule based screen to identify alternative oxidants for the coupling of *o*-aminophenols and anilines. Ferricyanide efficiently performed the coupling on both small molecules and proteins. The updated coupling reaction formed a single product with excellent conversion using only a few equivalents of the *o*-

aminophenol coupling partner. The biomolecule compatibility of this reaction was demonstrated with several protein and aminophenol substrates. Importantly, the ferricyanide-mediated coupling of *o*-aminophenols and anilines was completely orthogonal to both thiol and 1,2-diol moieties, facilitating its use with cysteine chemistry and glycosylated substrates. These attributes make the reaction particularly appropriate for applications that require installation of two synthetic components onto a protein. As an example, we demonstrated the dual modification of a viral capsid using both cysteine-maleimide chemistry as well as the ferricyanide-mediated coupling. The oxidative coupling reaction can also now reliably be used on glycoproteins, such as antibodies or antibody fragments, without any undesired oxidation of the glycan. While the periodate-mediated coupling is both unfailing and useful, the ferricyanide-mediated reaction can successfully be used in situations when $NaIO_4$ is not compatible with the substrate. Taken together, both oxidative coupling strategies offer valuable new ways to create nearly any complex bioconjugation target.

2.8 Materials and methods

General methods

Unless otherwise noted, the chemicals and solvents used were of analytical grade and were used as received from commercial sources. $K_3Fe(CN)_6$ was purchased from Sigma Aldrich and used without further purification. Analytical thin layer chromatography (TLC) was performed on EM Reagent 0.25 mm silica gel 60-F_{254} plates with visualization by ultraviolet (UV) irradiation at 254 nm and/or potassium permanganate stain. Purifications by flash chromatography were performed using EM silica gel 60 (230-400 mesh). The eluting system for each purification was determined by TLC analysis. Chromatography solvents were used without distillation. All organic solvents were removed under reduced pressure using a rotary evaporator. Water (dd-H_2O) used as reaction solvent was deionized using a Barnstead NANOpure purification system (ThermoFisher, Waltham, MA). Centrifugations were performed with an Eppendorf Mini Spin Plus (Eppendorf, Hauppauge, NY).

Instrumentation and sample analysis

NMR. ^{1}H and ^{13}C spectra were measured with a Bruker AVB-400 (400 MHz, 100 MHz) or a Bruker AV-600 (600 MHz, 150 MHz) spectrometer, as noted. ^{1}H NMR chemical shifts are reported as δ in units of parts per million (ppm) relative to $CDCl_3$ (δ 7.26, singlet). Multiplicities are reported as follows: s (singlet), d (doublet), t (triplet), dd (doublet of doublets), br (broad) or m (multiplet). Coupling constants are reported as a J value in Hertz (Hz). The number of protons (n) for a given resonance is indicated as nH and is based on spectral integration values. ^{13}C NMR chemical shifts are reported as δ in units of parts per million (ppm) relative to $CDCl_3$ (δ 77.2, triplet).

Inductively Coupled Plasma Optical Emission Spectroscopy. ICP-OES was performed on a Perkin Elmer 5300 DV optical emission ICP with a standard nebulizer system in the College of Natural Resources at UC Berkeley. Scandium was used as an internal standard for each sample. Each sample was measured 3 times.

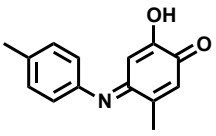
Mass Spectrometry. Matrix assisted laser desorption-ionization time-of-flight mass spectrometry (MALDI-TOF MS) was performed on a Voyager-DE system (PerSeptive Biosystems, USA) and data

were analyzed using Data Explorer software. Peptide samples were co-crystallized with α-cyano-4-hydroxycinnamic acid in 1:1 acetonitrile (MeCN) to H_2O with 0.1% trifluoroacetic acid (TFA). Protein bioconjugates were analyzed using an Agilent 1200 series liquid chromatograph (Agilent Technologies, USA) that was connected in-line with an Agilent 6224 Time-of-Flight (TOF) LC/MS system equipped with a Turbospray ion source.

High Performance Liquid Chromatography. HPLC was performed on Agilent 1100 Series HPLC Systems (Agilent, USA). Sample analysis for all HPLC experiments was achieved with an in-line diode array detector (DAD) and in-line fluorescence detector (FLD). Analytical reverse-phase HPLC of small molecules was accomplished using a C18 stationary phase and a H_2O/ MeCN with 0.1% TFA gradient mobile phase. Quantitation was accomplished using calibration curves for *p*--toluenesulfonic acid, purified reduced hydroquinone product **3**, and purified butenolide product **2**.

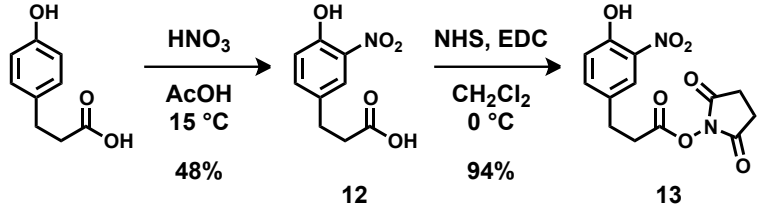
Gel Analyses. For protein analysis, sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) was carried out on a Mini-Protean apparatus (Bio-Rad, Hercules, CA), using a 10-20% precast linear gradient polyacrylamide gel (Bio-Rad). The sample and electrode buffers were prepared according to Laemmli.[46] All protein electrophoresis samples were heated for 5-10 min at 95 °C in the presence of 1,4-dithiothreitol (DTT) to ensure reduction of disulfide bonds. Gels were run for 75-90 minutes at 120 V to separate the bands. Commercially available markers (Bio-Rad) were applied to at least one lane of each gel for assignment of apparent molecular masses. Visualization of protein bands was accomplished by staining with Coomassie Brilliant Blue R-250 (Bio-Rad). For fluorescent protein conjugates, visualization was accomplished on Typhoon 9410 variable mode imager (Amersham Biosciences) prior to gel staining. Gel imaging was performed on an EpiChem3 Darkroom system (UVP, USA). ImageJ was used to determine the level of modification by optical densitometry.

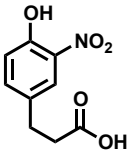
Small molecule synthesis



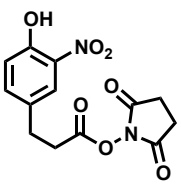
Synthesis of small molecule coupled product (1). To 90 mL of 10 mM phosphate buffer, pH 6.5 was added 2-amino-*p*-cresol (12.4 mg, 0.1 mmol, purified by sublimation and recrystallization from toluene) in 200 μL of MeCN and *p*-toluidine (10.7 mg, 0.1 mmol, purified by sublimation) in 200 μL of MeCN. To the stirred solution was added potassium ferricyanide (330 mg, 1 mmol, as a solution in 10 mL dd-H_2O). The reaction mixture was stirred at rt for 30 min, after which the solvent was removed *in vacuo* and the resulting solid was taken up in $CDCl_3$ for characterization. ^{1}H NMR (600 MHz, $CDCl_3$): δ 7.20 (d, 2H, *J* = 8.2), 6.75 (d, 2H, *J* = 8.0), 6.63 (s, 1H), 6.36 (s, 1H), 2.37 (s, 3H), 2.35 (s, 3H). ^{13}C NMR (150 MHz, $CDCl_3$): δ 182.96, 158.70, 153.29, 151.55, 147.93, 135.41, 129.67, 127.44, 120.43, 101.16, 20.98, 18.49. HRMS (ESI) calculated for $C_{14}H_{14}O_2N$ ($[M+H]^+$) 228.1019, found 228.1018 m/z.

Scheme 1. Synthesis of *o*-nitrophenol NHS ester 13.

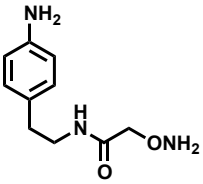




Synthesis of nitrophenol acid (12). To a solution of 3-(4-hydroxyphenyl)propionic acid (5 g, 30.1 mmol) in 25 mL acetic acid at 15 °C was added a solution of fuming nitric acid (1.6 mL, 33.8 mmol) in acetic acid (4 mL). The solution immediately turned orange. After 15 min the reaction was quenched by addition to ice water. The precipitate was filtered and dried. The yellow precipitate was recrystallized from 1:1 ethanol:water to afford 3.07 g of a yellow powder (48% yield). ^{1}H NMR (400 MHz, $CDCl_3$): δ 10.47 (br s, 1H), 7.94 (d, 1H, *J* = 1.9), 7.45 (dd, 1H, *J* = 8.6, 1.9), 7.09 (d, 1H, *J* = 8.6), 2.94 (t, 2H, *J* = 7.4), 2.68 (t, 2H, *J* = 7.4). ^{13}C NMR (100 MHz, $CDCl_3$): δ 178.22, 153.86, 138.11, 133.47, 132.60, 124.22, 120.27, 35.19, 29.37. HRMS (ESI) calculated for $C_9H_8O_5N$ ($[M-H]^-$) 210.0408, found 210.0410 m/z.



Synthesis of nitrophenol NHS ester (13). To a solution of 3-(4-hydroxy-3-nitrophenyl)propionic acid (**12**) (1 g, 4.7 mmol) and *N*-hydroxysuccinimide (0.65 g, 5.7 mmol) in CH_2Cl_2 at 0 °C was added *N*-(3-dimethylaminopropyl)-*N*'-ethylcarbodiimide hydrochloride (1.09 g, 5.7 mmol). The reaction mixture was stirred for 2 h and then diluted with CH_2Cl_2 and washed with water. The combined organic layers were dried over sodium sulfate and the solvent was removed *in vacuo*. The reaction afforded 1.38 g of a yellow solid (94% yield). ^{1}H NMR (400 MHz, $CDCl_3$): δ 10.49 (s, 1H), 7.98 (d, 1H, *J* = 2.1), 7.48 (dd, 1H, *J* = 8.6, 2.1), 7.12 (d, 1H, *J* = 8.6), 3.05 (t, 2H, *J* = 7.3), 2.93 (t, 2H, *J* = 7.3), 2.83 (s, 4H). ^{13}C NMR (100 MHz, $CDCl_3$): δ 169.06, 167.56, 154.09, 138.00, 133.54, 131.40, 124.52, 120.49, 32.48, 29.43, 25.70. HRMS (ESI) calculated for $C_{13}H_{11}O_7N_2$ ($[M-H]^-$) 307.0572, found 307.0568 m/z.



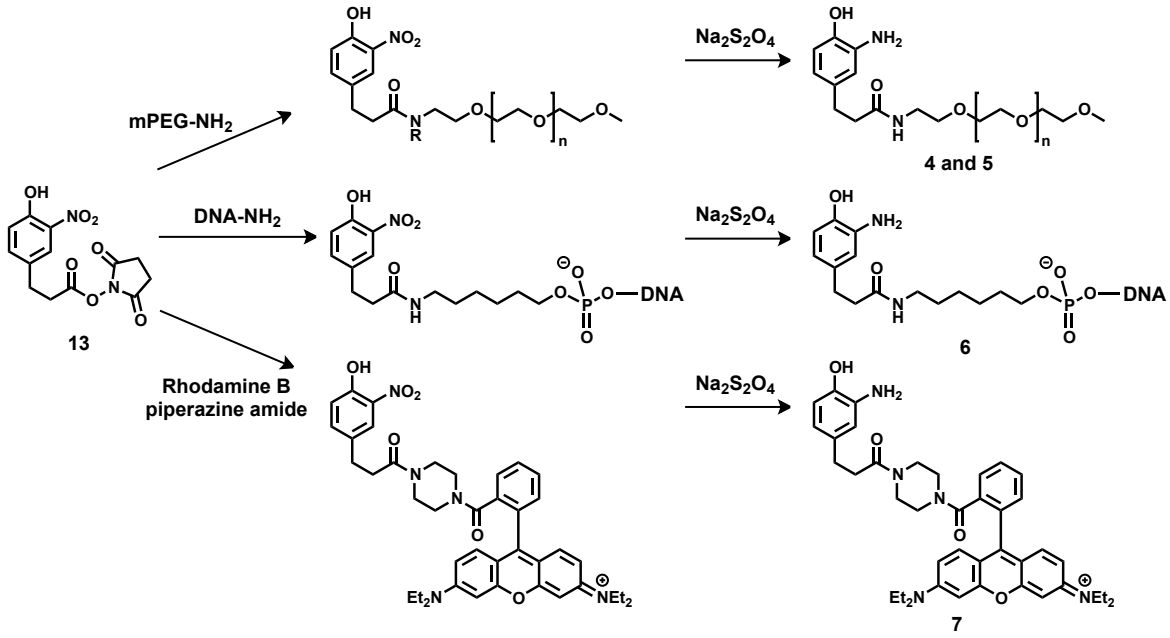
Synthesis of aniline alkoxyamine (14). The alkoxyamine was synthesized as previously reported.[43] To a solution of (Boc-aminooxy)acetic acid (280 mg, 1.5 mmol) and *N*-hydroxysuccinimide (168 mg, 1.5 mmol) in CH_2Cl_2 was added *N*,*N*-dicyclohexylcarbodiimide (363 mg, 1.7 mmol). The reaction mixture was stirred at rt for 15 min and then was filtered through Celite and a 0.22 micron PVDF syringe filter. To the filtered solution was added triethylamine (550 µL, 4 mmol) and 2-(4-aminophenyl)ethylamine (430 mg, 3.1 mmol). After 1 h of stirring the solvent was removed *in vacuo*. Purification on silica gel with ethyl acetate as the mobile phase yielded 294 mg (65% yield) of the product. ^{1}H NMR (600 MHz, $CDCl_3$): δ 6.94 (d, 2H, *J* = 8.2), 6.56 (d, 2H, *J* = 8.2), 4.19 (s, 2H), 3.40 (m, 2H), 2.68 (m, 2H), 1.41 (s, 9H). ^{13}C NMR (150 MHz, $CDCl_3$): δ 169.06, 157.78, 144.75, 129.41, 128.60, 115.29, 82.41, 75.73, 40.77, 34.62, 28.04. HRMS (ESI) calculated for $C_{15}H_{23}O_4N_3$ ($[M+H]^+$) 310.1761, found 310.1765 m/z. The Boc group was removed by exposure to a 1:1 mixture of trifluoroacetic acid (TFA):CH_2Cl_2 for 10 min. The solvent was removed under a stream of N_2 and the resulting oil was put under vacuum overnight to remove residual TFA. The alkoxyamine was stored as a 200 mM solution in water. Before use with proteins, the solution was diluted with phosphate buffer to adjust the pH to 4.0-6.0.

Aminophenol substrate synthesis

Synthesis of *o*-aminophenol PEG (4: 2k-PEG, 5: 5k-PEG). To a solution of mPEG-NH_2 (MW=5000, 115 mg, 0.023 mmol) in CH_2Cl_2 (1 mL) was added triethylamine (7 µL, 0.05 mmol) and **13** (99 mg, 0.32 mmol as a 0.4 M solution in DMSO). The solution was stirred for 2 h at rt and then the solvent

was removed *in vacuo*. Excess **13** was precipitated by the addition of water. The resulting solution was filtered through a 0.22 micron filter and three Sephadex size exclusion columns (one NAP-10 and two NAP-25, GE Healthcare) according to the manufacturer's instructions. The modified PEG was further purified by precipitation from CH_2Cl_2 with Et_2O three times. Reduction of the nitrophenol was accomplished with sodium dithionite. To a 1 mM solution of the modified PEG was added an equal volume of a freshly prepared 120 mM solution of sodium dithionite in 100 mM phosphate buffer, pH 6.5. Excess dithionite was removed by purification with three Sephadex size exclusion columns. The purified aminophenol PEG was lyophilized and resuspended in 10 mM phosphate buffer, pH 7.2. The concentration was adjusted to 1 mM by measuring the aminophenol absorbance at 290 nm. The final solution of *o*-aminophenol PEG was stored in single-use aliquots at -20 °C until use. An analogous procedure was followed for the synthesis of a 2k version of the

Scheme 2. Synthesis of *o*-aminophenol substrates.



o-aminophenol PEG.

Synthesis of *o*-aminophenol DNA aptamer (6). To a solution of 5' amine modified DNA (sequence : 5'(NH_2) - ATCTAACTGCTGCGCCGCCGGGAAAATACTGTACGGTTAGA - 3') in 50 mM phosphate buffer, pH 8.0 (312 µM, 4 mg/mL) was added an equal volume of **13** (38.4 mM, in 4:1 DMF:DMSO). The reaction mixture was briefly vortexed and incubated at rt for 1.5 h. The reaction mixture was filtered through a 0.22 micron filter to remove any precipitate. Excess small molecule was removed by filtration through a Sephadex size exclusion NAP-5 column, pre-equilibrated with 10 mM phosphate buffer, pH 6.5. The nitrophenol was reduced by the addition of sodium dithionite (5.3 µmol, 53 µL of a freshly prepared 100 mM solution in 100 mM phos buffer pH 6.5) with an incubation at rt for 20 min. Removal of the dithionite and any remaining small molecule was accomplished by two successive filtrations on Sephadex columns. The purified *o*-aminophenol DNA was lyophilized and resuspended in 30 µL of 10 mM phosphate buffer, pH 7.2 to an approximate concentration of 1 mM *o*-aminophenol. The modified DNA was stored at -20 °C until use.

Synthesis of *o*-aminophenol rhodamine (7). To a solution of Rhodamine B piperazine amide[47] (50 mg, 0.1 mmol) in DMF (1 mL) was added diisopropylethylamine (35 μL, 0.2 mmol) followed by **13** (30 mg, 0.1 mmol). After stirring for 2 h the solvent was removed *in vacuo*. The product was purified on silica gel with a gradient of CH_2Cl_2:CH_3OH (0 to 10% CH_3OH). HRMS (ESI) calculated for $C_{41}H_{45}O_6N_5$ ($[M+H]^+$) 704.3433, found 704.3436 m/z. To a 1 mg portion of **7** in 100 μL of 100 mM phosphate buffer, pH 6.5 was added an equal volume of a freshly prepared 100 mM solution of sodium dithionite in 100 mM phosphate buffer, pH 6.5. After 20 min, the excess dithionite was removed by purification on a C18 Sep-Pak (Waters) according to the manufacturer's instructions. The eluent was concentrated to dryness. The rhodamine aminophenol was resuspended in DMF and stored frozen as a 50 mM solution.

Protein expression and purification

Expression and purification of wt and T19Y MS2. The empty capsids were expressed and purified as previously reported.[48] Briefly, the DNA for the wild type and T19Y MS2 coat protein were cloned into pBAD/*myc* plasmids. The plasmids were transformed into DH10B cells for expression. The cells were grown in 1 L LB with 100 μg/L ampicillin with 0.05% arabinose for 18 h at 37 °C. The cells were collected by centrifugation and resuspended in 20 mM taurine buffer, pH 9.0 supplemented with 200 μg DNase I and RNase A and 24 mg $MgCl_2$. The cells were lysed by sonication. The cell debris was removed by centrifugation. The clarified lysate was applied to a DEAE-Sephadex column (GE Healthcare). Fractions containing MS2 were combined and precipitated with 10% w/v PEG-6k with 0.5 M NaCl at 4 °C overnight. The precipitated protein was collected by centrifugation, resuspended in 10 mM phosphate buffer, pH 7.2, and further purified on a Sephacryl-500 gel filtration column (GE Healthcare).

Expression and purification of T19*p*AF and T19*p*AF N87C MS2. The MS2 capsids with the unnatural amino acid were expressed and purified as previously reported.[30] The plasmid containing the amber stop codon mutation in the MS2 coat protein was co-tranformed into DH10B cells with the pDULE plasmid containing the *p*AF aminoacyl tRNA synthetase and tRNA. The expression was carried out in minimal media, following the published protocol.[49] The purification was as described above for the wt and T19Y mutants; however the purification procedure was completed two times to afford pure capsids. Approximately 10-15 mg of purified protein was recovered per L of expression culture.

Expression and purification of AKT-Fc. The antibody fragment, AKT-Fc, was expressed in HEK-293 cells and purified as previously reported.[43]

Protein modification

Transamination and oxime formation of AKT-Fc. To a solution of AKT-Fc in 25 mM phosphate buffer, pH 6.5 (50 μL, 50 μg) was added an equal volume of a 200 mM solution of pyridoxal-5'-phosphate (PLP) in phosphate buffer, pH 6.5. The resulting solution was incubated for 1 h at 37 °C. Excess PLP was removed by a NAP-5 Sephadex size exclusion column (GE Healthcare). The transaminated protein was exchanged into 25 mM phosphate buffer, pH 5.0 by repeated centrifugal filtration against a 10 kDa MWCO membrane. To the transaminated Fc was added an equal volume

of a 50 mM solution of aniline alkoxyamine **14**. The solution was incubated at rt for 48 h.

Disulfide exchange with Ellman's reagent. Free cysteines were protected from potential oxidation by disulfide formation with Ellman's reagent (5,5'-dithiobis-(2-nitrobenzoic acid), DTNB). To a solution of T19*p*AF N87C MS2 (100 µL of a 50 µM solution) was added DTNB (100 µL of a 10 mM solution in 100 mM phosphate buffer, pH 7.2 with 1 mM EDTA). The reaction mixture was incubated at rt for 1 h and then the excess DTNB was removed by repeated (3-6 times) centrifugal filtration against a 100 kDa MWCO membrane. To reduce the disulfide, 1 µL of a 0.5 M solution of TCEP, pH 7.0 was added to the protein sample.

General procedure for oxidative coupling on proteins. To a solution of aniline containing protein (*p*AF MS2, aniline-Fc, 5-20 µM) in 10 mM phosphate buffer, pH 6.5 was added 5-10 equiv of the *o*-aminophenol coupling partner (50-100 µM). The solution was briefly vortexed and then 5-10 equiv (relative to the *o*-aminophenol) of potassium ferricyanide (as a 10 mM solution in 10 mM phosphate buffer, pH 6.5) was added. After 20 min, the reaction mixture was purified on a NAP-5 Sephadex size exclusion column (GE Healthcare) according to the manufacturer's instructions and eluted in 10 mM phosphate buffer, pH 7.2. Any remaining ferricyanide and unreacted *o*-aminophenol coupling partner were removed using a 0.5 mL centrifugal filter with an appropriate molecular weight cut off (MWCO, Millipore). The samples were concentrated to 50 µL and then diluted 10-fold with 10 mM phosphate buffer, pH 7.2. This process was repeated 5-10 times depending on the concentration of *o*-aminophenol used. Modification was monitored by SDS-PAGE or LC-MS.

Cysteine alkylation with fluorescent maleimides. A solution of T19*p*AF N87C MS2 (100 µM) in 10 mM phosphate buffer, pH 8.0 was incubated with 1 equiv of an Alexa Fluor maleimide (Alexa Fluor 680 C_2-maleimide or Alexa Fluor 488 C_5-maleimide, 20 mM in DMSO) for 2 h. Unreacted dye was removed with a NAP-10 Sephadex size exclusion column (GE Healthcare) and by repeated centrifugal filtration against a 100 kDa MWCO membrane.

Fc oxidation and oxime formation. To a solution of AKT-Fc (9 µL, 3 µg) in 10 mM phosphate buffer was added either sodium periodate or potassium ferricyanide (1 µL of a 10 mM solution). The reaction mixture was incubated for 2-60 min before the addition of 1 µL of 0.5 M TCEP, pH 7.0. After quenching the oxidant, 1 µL of a 6 mM solution of Alexa Fluor 488 C_5-aminooxyacetamide was added. The oxime formation reaction was incubated at rt for 4 h before analysis by SDS-PAGE.

Other procedures

Glycopeptide oxidation. To a solution of glycosylated peptides (9 µL of a 3 µM solution of glycosylated erythropoietin, MUC5AC 3, and carcinoembryonic antigen CGM2, Protea Biosciences, USA) was added either sodium periodate or potassium ferricyanide (1 µL of a 10 mM solution). The reaction mixtures were incubated for 5-60 min and were then co-crystallized with matrix on a MALDI plate for analysis. Signal was only observed for glycosylated erythropoietin (Figure 2.11).

Removal of ferricyanide and quantification with ICP-OES. To a solution of T19*p*AF N87C MS2 (1.2 mL, 120 nmol) in 10 mM phosphate buffer, pH 7.2 (4.68 mL) was added 2-amino-*p*-cresol (60 µL, 0.6 µmol) followed by potassium ferricyanide (60 µL, 6 µmol). The reaction mixture was incu-

bated for 15 min at rt and then divided in to 6 portions. The first portion was not subjected to any purification. The second portion was purified by repeated centrifugal filtration (8 times) against a 0.5 mL 100 kDa MWCO membrane. The third portion was purified by using a Sephadex NAP-25 size exclusion column. The fourth portion was purified by a 1 h incubation with 0.5 g of anion exchange resin (Amberlite® IRA-900 chloride form, Sigma Aldrich) pre-rinsed with dd-H_2O. The fifth portion was purified in the same manner as the fourth portion with the exception that Chelex® 100 resin (Bio-Rad) was used instead of the anion exchange resin. The sixth and final portion was purified by HPLC using a Yarra gel filtration column (Phenomenex). After purification, all of the samples were diluted to 5 mL for analysis by ICP-OES. A reaction was also run without ferricyanide as a negative control.

DFT calculations. Jaguar (Maestro v. 9.3.5) was used to calculate the gas phase Gibbs free energies of the two tautomers of the small molecule product **1**. The basis set used was B3LYP/6-311G^{**+}.

2.9 References

1. O'Hare, H. M.; Johnsson, K.; Gautier, A. *Curr. Opin. Struct. Biol.* **2007**, *17*, 488–494.
2. Zalipsky, S. *Bioconjugate Chem*. **1995**, *6*, 150–165.
3. Alley, S. C.; Okeley, N. M.; Senter, P. D. *Curr. Opin. Chem. Biol*. **2010**, *14*, 529–537.
4. Wu, W.; Hsiao, S. C.; Carrico, Z. M.; Francis, M. B. *Angew. Chem. Int. Ed.* **2009**, *48*, 9493–9497.
5. Stephanopoulos, N.; Tong, G. J.; Hsiao, S. C.; Francis, M. B. *ACS Nano* **2010**, *4*, 6014–6020.
6. Day, J. J.; Marquez, B. V.; Beck, H. E.; Aweda, T. A.; Gawande, P. D.; Meares, C. F. *Curr. Opin. Chem. Biol.* **2010**, *14*, 803–809.
7. Tong, G. J.; Hsiao, S. C.; Carrico, Z. M.; Francis, M. B. *J. Am. Chem. Soc*. **2009**, *131*, 11174–11178.
8. Baslé, E.; Joubert, N.; Pucheault, M. *Chem. Biol.* **2010**, *17*, 213–227.
9. Sletten, E. M.; Bertozzi, C. R. *Angew. Chem. Int. Ed.* 2009, 48, 6974–6998.
10. Kim, C. H.; Axup, J. Y.; Schultz, P. G. *Curr. Opin. Chem. Biol*. **2013**, *17*, 412–419.
11. Crankshaw, M. W.; Grant, G. A. In Current Protocols in Protein Science; John Wiley & Sons, Inc., 2001.
12. Chalker, J. M.; Bernardes, G. J. L.; Lin, Y. A.; Davis, B. G. *Chem Asian J* **2009**, *4*, 630–640.
13. Hang, H. C.; Yu, C.; Kato, D. L.; Bertozzi, C. R. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 14846–14851.
14. Cornish, V. W.; Hahn, K. M.; Schultz, P. G. *J. Am. Chem. Soc.* 1996, 118, 8150–8151.
15. Dirksen, A.; Dawson, P. E. *Bioconjugate Chem.* **2008**, *19*, 2543–2548.
16. Kiick, K. L.; Saxon, E.; Tirrell, D. A.; Bertozzi, C. R. *Proc. Natl. Acad. Sci. U.S.A*. **2002**, *99*, 19–24.
17. Jewett, J. C.; Bertozzi, C. R. *Chem. Rev.* **2010**, *39*, 1272–1279.
18. Blackman, M. L.; Royzen, M.; Fox, J. M. *J. Am. Chem. Soc*. **2008**, *130*, 13518–13519.
19. Devaraj, N. K.; Weissleder, R.; Hilderbrand, S. A. *Bioconjugate Chem.* **2008**, *19*, 2297–2299.
20. Li, N.; Lim, R. K. V.; Edwardraja, S.; Lin, Q. *J. Am. Chem. Soc*. **2011**, *133*, 15316–15319.
21. Behrens, C. R.; Hooker, J. M.; Obermeyer, A. C.; Romanini, D. W.; Katz, E. M.; Francis, M. B. *J. Am. Chem. Soc.* **2011**, *133*, 16398–16401.
22. Witus, L. S.; Francis, M. B. *Acc. Chem. Res*. **2011**, *44*, 774–783.

23. Miller, R. A.; Presley, A. D.; Francis, M. B. *J. Am. Chem. Soc.* **2007**, *129*, 3104–3109.
24. Dedeo, M. T.; Duderstadt, K. E.; Berger, J. M.; Francis, M. B. *Nano Lett*. **2010**, *10*, 181–186.
25. Esser-Kahn, A. P.; Iavarone, A. T.; Francis, M. B. *J. Am. Chem. Soc.* **2008**, *130*, 15820–15822.
26. Hooker, J. M.; Kovacs, E. W.; Francis, M. B. *J. Am. Chem. Soc.* **2004**, *126*, 3718–3719.
27. Hooker, J. M.; Esser-Kahn, A. P.; Francis, M. B. *J. Am. Chem. Soc.* **2006**, *128*, 15558–15559.
28. Wang, L.; Brock, A.; Herberich, B.; Schultz, P. G. *Science* **2001**, *292*, 498–500.
29. Mehl, R. A.; Anderson, J. C.; Santoro, S. W.; Wang, L.; Martin, A. B.; King, D. S.; Horn, D. M.; Schultz, P. G. *J. Am. Chem. Soc*. **2003**, *125*, 935–939.
30. Carrico, Z. M.; Romanini, D. W.; Mehl, R. A.; Francis, M. B. *Chem. Commun.* **2008**, 1205–1207.
31. Strauss Jr., J. H.; Sinsheimer, R. L. *J. Mol. Biol.* **1963**, *7*, 43–54.
32. Valegård, K.; Liljas, L.; Fridborg, K.; Unge, T. *Nature* **1990**, *345*, 36–41.
33. Antonini, E.; Brunori, M.; Wyman, J. *Biochemistry* **1965**, *4*, 545–551.
34. Corbett, J. F. *J. Chem. Soc. B* **1969**, 207–212.
35. Gardlik, S.; Rajagopalan, K. V. *J. Biol. Chem*. **1991**, *266*, 4889–4895.
36. Yamagishi, Y.; Ashigai, H.; Goto, Y.; Murakami, H.; Suga, H. *ChemBioChem* **2009**, *10*, 1469–1472.
37. Tabbì, G.; Cassino, C.; Cavigiolio, G.; Colangelo, D.; Ghiglia, A.; Viano, I.; Osella, D. *J. Med. Chem*. **2002**, *45*, 5786–5796.
38. Hadjipetrou, L. P.; Gray-Young, T.; Lilly, M. D. *J. Gen. Microbiol.* **1966**, *45*, 479–488.
39. Liu, C.; Sun, T.; Zhai, Y.; Dong, S. *Talanta* **2009**, *78*, 613–617.
40. Alley, S. C.; Anderson, K. E. *Curr. Opin. Chem. Biol*. **2013**, *17*, 406–411.
41. Caravella, J.; Lugovskoy, A. *Curr. Opin. Chem. Biol*. **2010**, *14*, 520–528.
42. Hamilton, J. K.; Smith, F. *J. Am. Chem. Soc*. **1956**, *78*, 5910–5912.
43. Netirojjanakul, C.; Witus, L. S.; Behrens, C. R.; Weng, C.-H.; Iavarone, A. T.; Francis, M. B. *Chem. Sci*. **2013**, 4, 266–272.
44. Gilmore, J. M.; Scheck, R. A.; Esser-Kahn, A. P.; Joshi, N. S.; Francis, M. B. *Angew. Chem. Int. Ed.* **2006**, *45*, 5307–5311.
45. Witus, L. S.; Moore, T.; Thuronyi, B. W.; Esser-Kahn, A. P.; Scheck, R. A.; Iavarone, A. T.; Francis, M. B. *J. Am. Chem. Soc.* **2010**, *132*, 16812–16817.
46. Laemmli, U. K. *Science* **1970**, *227*, 680–685.
47. Nguyen, T.; Francis, M. B. *Org. Lett.* **2003**, *5*, 3245–3248.
48. Hooker, J. M.; O'Neil, J. P.; Romanini, D. W.; Taylor, S. E.; Francis, M. B. *Mol. Imaging Biol.* **2008**, *10*, 182–191.
49. Hammill, J. T.; Miyake-Stoner, S.; Hazen, J. L.; Jackson, J. C.; Mehl, R. A. *Nat. Protocols* **2007**, *2*, 2601–2607.

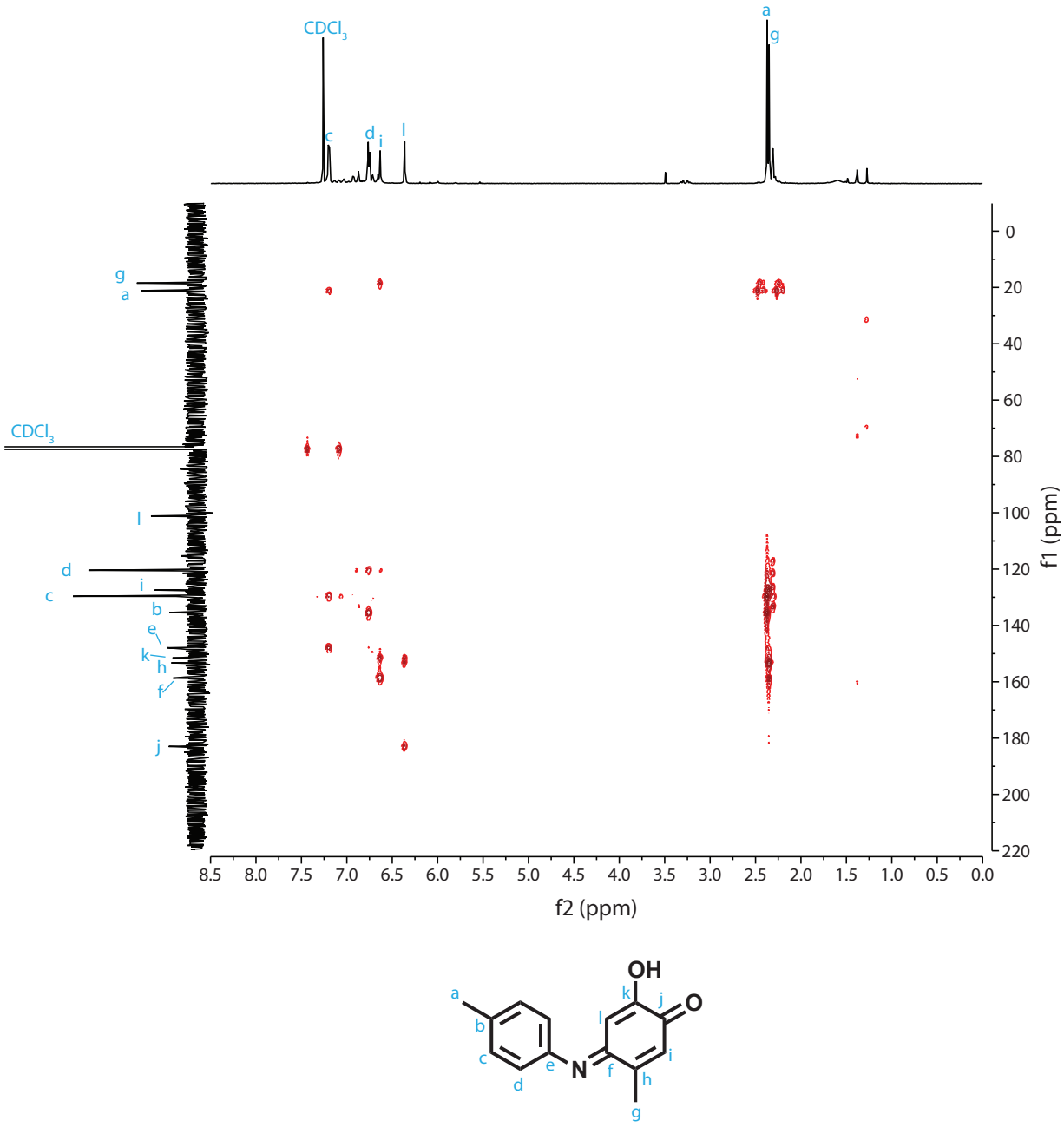


Figure 2.11. HMBC spectrum of **1** in $CDCl_3$.

Chapter 3
Oxidative Coupling to Native Amino Acids

Abstract

The oxidative coupling of electron-rich aromatic molecules, such as *o*-aminophenols, to aniline residues has been used to site-selectively modify proteins for a variety of applications (as discussed in Chapter 2). However, these methods all require introduction of an artificial amino acid, *p*-aminophenylalanine. This chapter describes the discovery and optimization of oxidative coupling conditions for the modification of native amino acids. The N-terminal amino group was identified as the reactive functional group on peptide and protein substrates. The effect of the N-terminal residue on reactivity was assessed and the secondary amine of N-terminal proline was identified as the most reactive residue. The reaction was characterized on small molecule, peptide, and protein substrates.

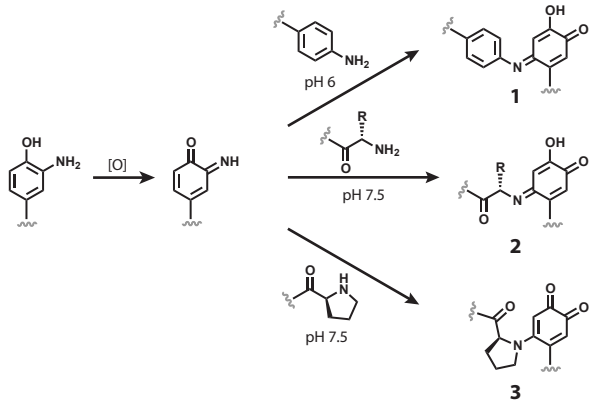
3.1 Site-specific modification of proteins

As outlined in the previous two chapters, the synthetic modification of proteins enables the construction of hybrid materials that can be used in applications from the study of protein function[1] to the development of potent, targeted therapeutics.[2] The synthesis of these materials requires a suite of bioconjugation reactions suited for the application at hand. These reactions must take place in mild, aqueous conditions in the presence of the native functional groups present on the protein surface. Common methods for protein modification target the nucleophilic side-chains of lysine and cysteine.[3–5] However, these strategies can result in complex product mixtures, as lysine is often found in high abundance on the protein surface,[6] and cysteine is often critical for native structure and function. While these limitations may be acceptable for some applications, often site-selective modification of the protein is required.

As discussed in Chapter 2, one approach for the site-selective modification of proteins relies on the introduction of an artificial amino acid with reactivity that is orthogonal to the native amino acids. A number of powerful methods have been developed for the selective modification of azide,[7–13] alkyne,[9–12] alkene,[14–17] carbonyl,[18,19] and aniline[20,21] moieties. However, the difficulty of introducing a non-canonical amino acid can limit the application of these methods. Complementary approaches rely on site-selective modification of native amino acids by enzymes[22–27] or by targeting the protein termini.[28,29] A robust method for the modification of C-terminal thioesters with N-terminal cysteines, termed "native chemical ligation" was developed by Kent and co-workers. This method has been used for the semi- and total synthesis of complex protein substrates, including the synthesis of a single glycoform of human erythropoietin. As an alternative approach, we and others have developed methods for the selective modification of the N-terminal amino group.[30–36] These methods, however, typically require long reaction times, large excesses of reagent, and at least two steps for the conjugation of synthetic molecules. We have long sought a method that both targets the N-terminus and directly modifies the terminus in one step.

The oxidative coupling reaction described in the previous chapter features rapid addition of anilines to iminoquinone intermediates generated by $NaIO_4$ or $K_3Fe(CN)_6$ (Scheme 1). The chemoselectivity of the coupling most likely derives from the ability of aniline to act as a nucleophile at pH values (6.0-6.5) that leave other protein residues protonated, and thus non-nucleophilic.[21] In the course of developing the reaction with $K_3Fe(CN)_6$, we discovered that native amino acids can participate in the reaction at elevated pH. Additionally, the addition of *o*-aminophenols and catechols to native amino acids was known dating back to 1948.[37–42] Based on these reports and our success with aniline-based couplings, we sought to develop a robust method for the modification of native amino acids with *o*-aminophenols.

Scheme 1. Oxidative coupling with *o*-aminophenols.



Covered in this chapter, we examined the reactivity of *o*-aminophenols with native amino acids.We found that at pH 7.5 N-terminal amino groups can be selectively targeted for oxidative coupling to *o*-aminophenols and *o*-catechols. We identified conditions that preferentially modify the N-terminus with fast sec-

ond order kinetics. Peptide substrates were used to screen reaction conditions and identify the site of modification. A panel of N-terminal variant peptides was screened to determine the sequence specificity of the reaction and proline was identified as the optimal N-terminal amino acid. The reaction was applied to protein substrates and showed high levels of conversion with an N-terminal proline residue. This new bioconjugation reaction enables the facile, rapid modification of proteins creating a well-defined, stable linkage.

3.2 Discovery of native amino acid reactivity

We have previously identified *o*-aminophenols and *o*-catechols as efficient coupling partners for the oxidative modification of aniline residues as covered in Chapter 2.[21] The mild oxidant, $K_3Fe(CN)_6$, reported there expands the scope, convenience, and scalability of the reaction. In the course of optimizing the oxidative coupling, we observed low amounts of background reactivity with native amino acids. This side reactivity could be prevented by performing the bioconjugation at pH 6.0-6.5 (Figure 2.3). Addition of imidazole to the reaction also inhibited the reactivity with native amino acids. While we had previously identified conditions that enabled the selective modification of aniline side chains, we hypothesized that conditions could also be found to modify native residues.

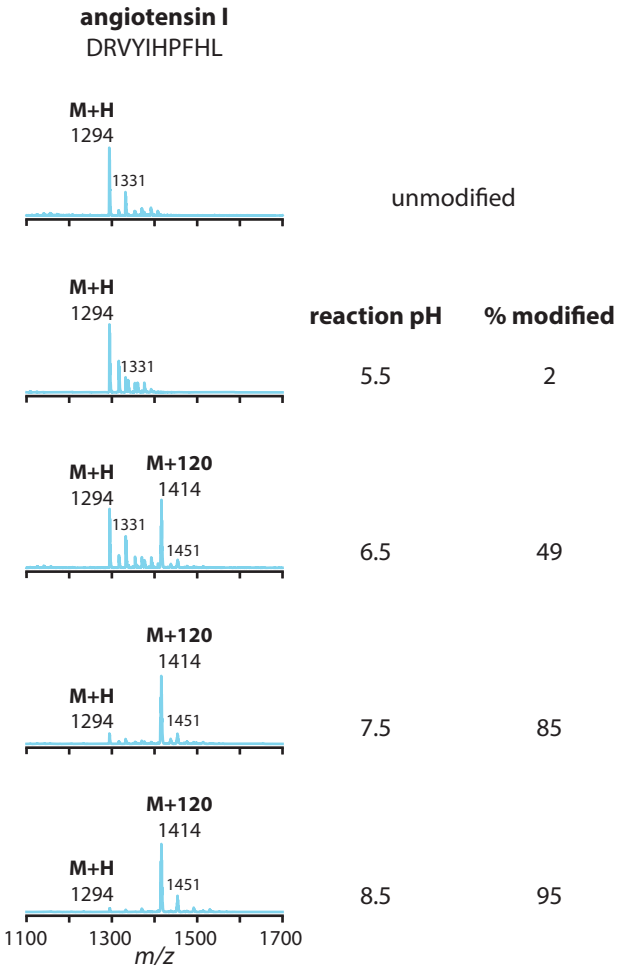


Figure 3.1. The effect of pH on the reactivity of the N-terminus with *o*-aminophenols was tested. Angiotensin I (100 µM) was reacted with 2-amino-*p*-cresol (100 µM) in the presence of $K_3Fe(CN)_6$ (1 mM) in 10 mM phosphate buffer at a pH range of 5.5-8.5 (expected mass addition = +120).

3.3 Screening reactivity on peptide substrates

The ability of native amino acids to undergo oxidative coupling with *o*-aminophenols was first assayed using

peptide substrates. Angiotensin I and melittin were chosen initially as they contained many potentially reactive amino acids, including Lys, Arg, His, Trp, and Tyr. The peptides were reacted with 2-amino-*p*-cresol using $K_3Fe(CN)_6$ as the oxidant. The reaction pH was varied from 5.5 to 8.5 and the reactions were analyzed by MALDI-TOF MS (Figure 3.1). The level of modification increased with the basicity of the reaction, with near quantitative modification of angiotensin I at pH 7.5 and higher. Throughout these initial investigations it was noted that angiotensin and melittin showed significant differences in reactivity, with angiotensin consistently demonstrating better conversion. MS/MS analysis of modified angiotensin was used to identify the reactive functional group and revealed that the N-terminal residue was responsible for the observed reactivity (Figure 3.2). As further confirmation of the site-selectivity, several peptide substrates were screened for reactivity (Figure 3.3). The only peptide that did not react had a blocked, pyroglutamate N-terminus.

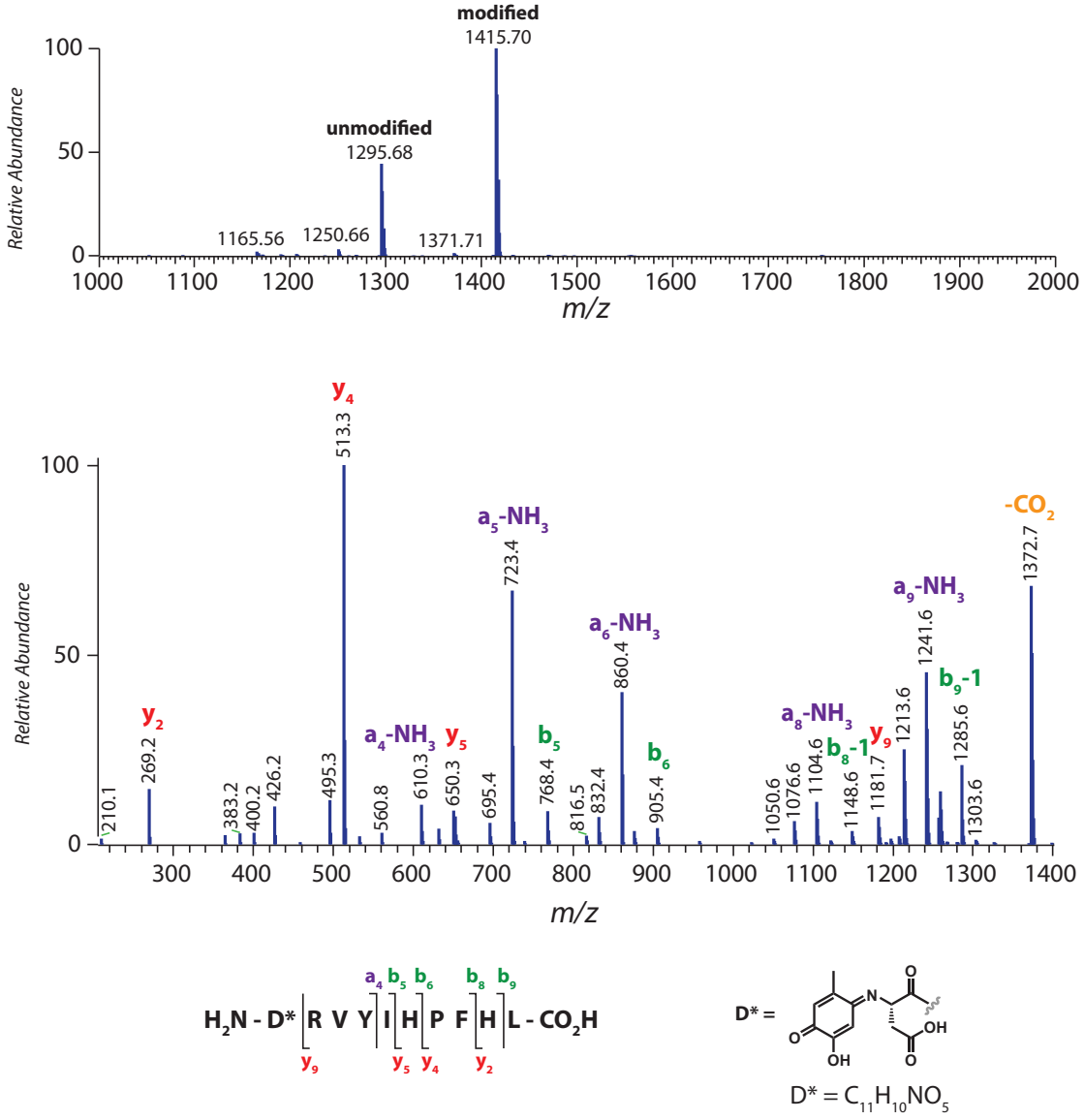


Figure 3.2. MS/MS analysis of modified angiotensin I. The *y* ions are shown in red, the *b* ions are shown in green, and the *a* ions (with neutral losses of ammonia) are shown in purple. The analysis is consistent with modification at the N-terminus.

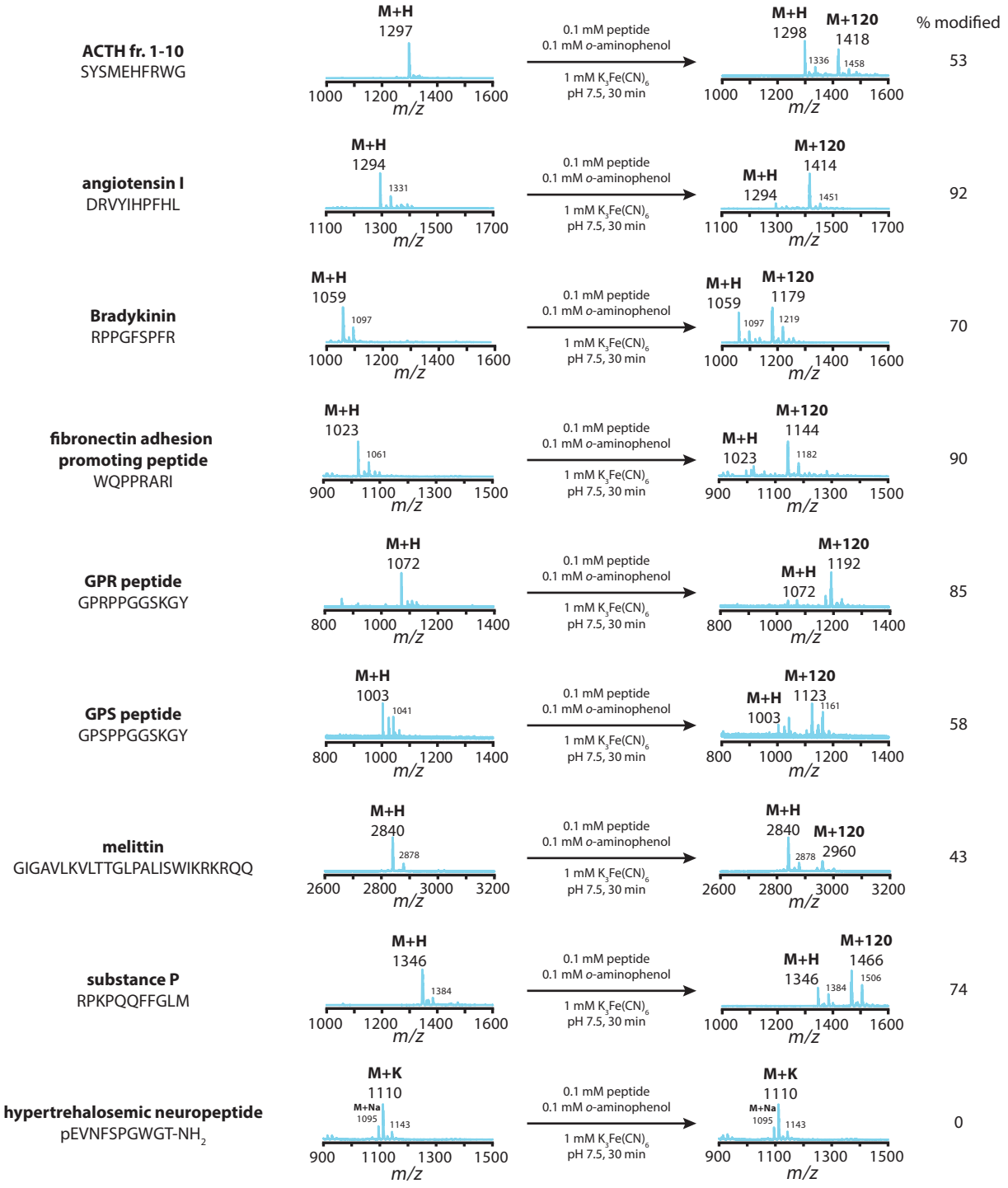


Figure 3.3. Several peptides were screened for their ability to react with *o*-aminophenols in the presence of ferricyanide. Reactions were run with approximately 0.1 mM peptide, 0.1 mM *o*-aminophenol and 1 mM ferricyanide in phosphate buffer, pH 7.5. All of the peptides showed some degree of modification (expected mass addition = +120), with the exception of the hypertrehalosemic neuropeptide. Potassium adducts were frequently observed (+38). The only consistent functional group amongst the peptides that modified but was not present in the peptide that did not was a free N-terminus.

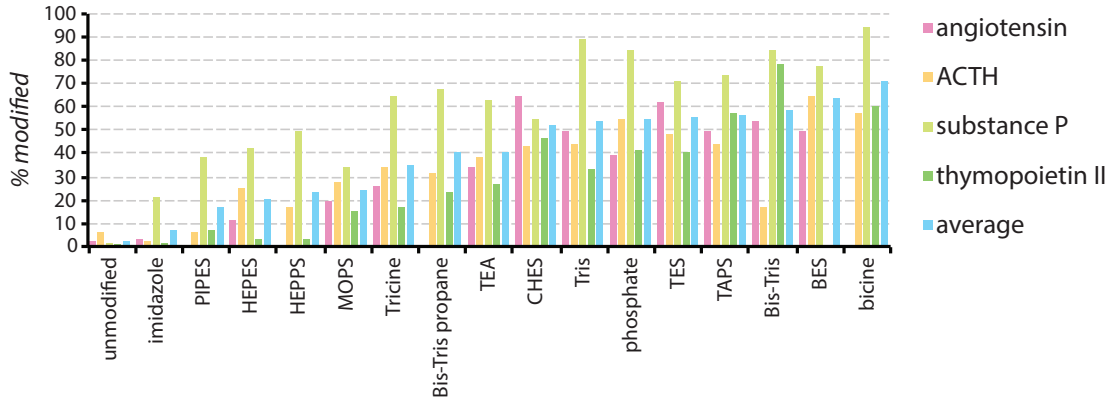


Figure 3.4. Peptides were screened in several different buffers at pH 7.5 to determine the buffer salt compatibility of the reaction. Most buffers showed no effect on the reaction, however a few buffers inhibited the reaction. The buffers that diminished reactivity generally contained a morpholino or piperazine moiety. Reactions were run with approximately 0.1 mM peptide, 0.1 mM *o*-aminophenol, and 1 mM ferricyanide at pH 7.5 for 30 minutes.

The oxidative coupling conditions were then optimized on peptide substrates. Several buffer salts were screened for compatibility with the reaction. While most buffers did not effect the reaction, imidazole and buffers containing a morpholine or piperazine ring (PIPES, HEPES, HEPPS, and MOPS) significantly impeded the reaction (Figure 3.4). This is possibly due to buffer impurities (such as secondary amines) that inhibit reactivity or competitively react with the oxidized aminophenol. In addition, it was shown that the peptides could be modified using $NaIO_4$ as the oxi-

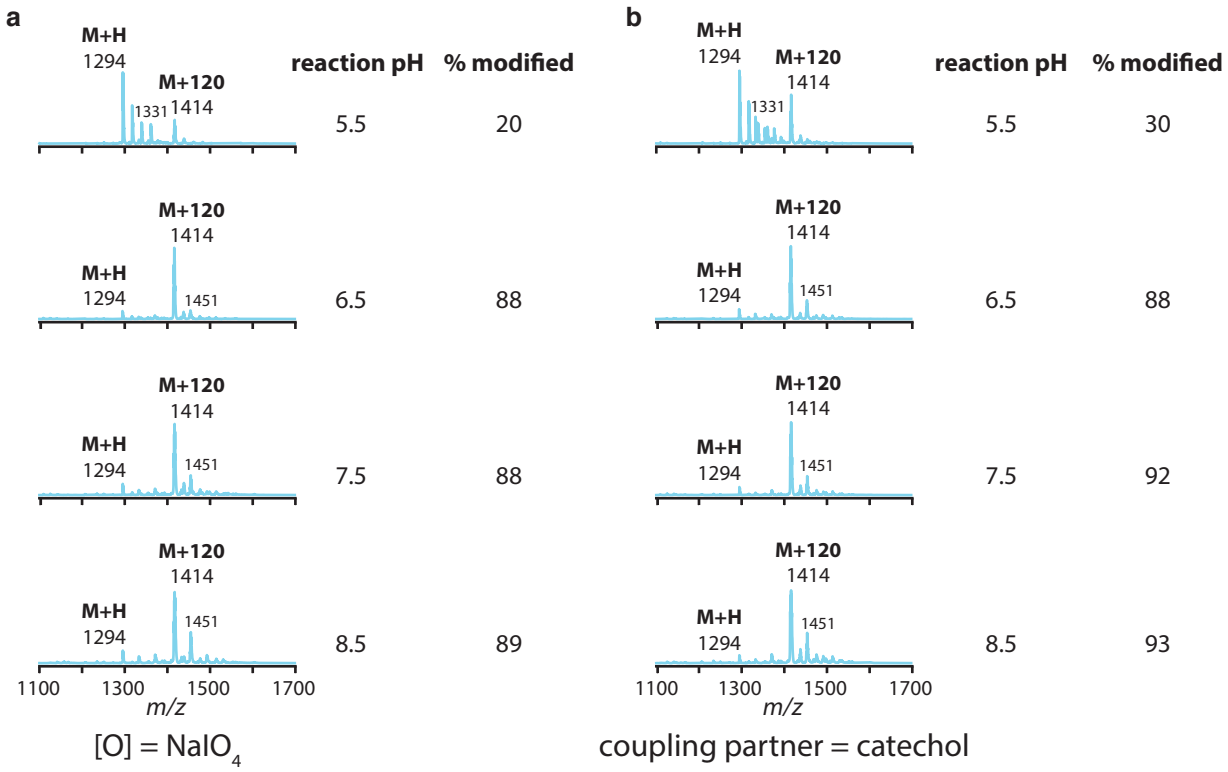


Figure 3.5. The effect of pH on oxidative coupling to the N-terminus was tested. (a) Angiotensin I (100 μM) was reacted with 2-amino-*p*-cresol (100 μM) in the presence of $NaIO_4$ (1 mM) in 10 mM phosphate buffer at a pH range of 5.5-8.5. (b) Angiotensin I (100 μM) was reacted with 4-methylcatechol (100 μM) in the presence of $K_3Fe(CN)_6$ (1 mM) in 10 mM phosphate buffer at a pH range of 5.5-8.5 (expected mass addition = +120).

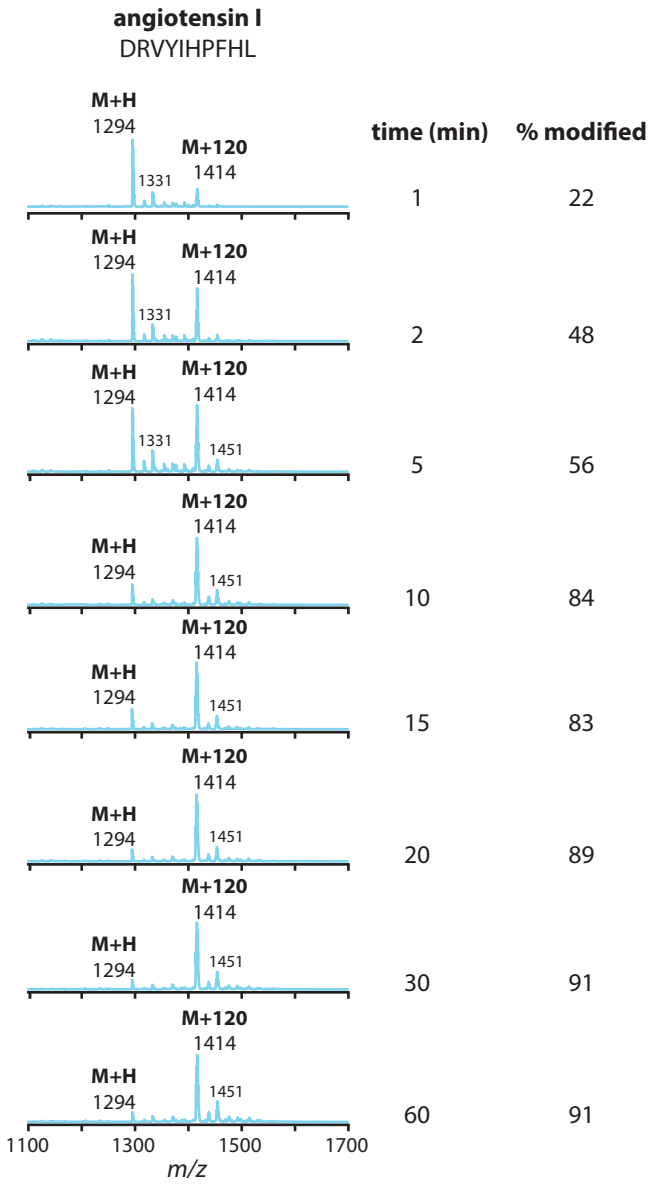


Figure 3.6. The time course of the reaction was assayed using angiotensin I as a substrate. The reactions were co-crystallized with matrix on a MALDI plate at the indicated time (expected mass addition = +120).

dant and/or 4-methylcatechol as the coupling partner (Figure 3.5). Both of these reactions showed the same dependence on pH; however, moderate levels of modification were observed at acidic pH (5.5-6.5) when using these alternative coupling conditions. These observations led us to hypothesize that the peptides were reacting with the *o*-quinone formed *in situ*. The time course of the reaction was also investigated (Figure 3.6). The reaction reached completion after only 20 min. We next sought to investigate the effect of the N-terminal residue on reactivity.

3.4 Positional scanning

Given the differential reactivity observed on peptide substrates, we synthesized peptides with varied N-terminal residues (XADSWAG) to determine the specificity of the reaction for the N-terminal amine. The base sequence was selected to increase the mass of the peptide and impart water solubility as well as to include a UV absorbant amino acid for facile quantitation. The peptides were synthesized on the solid phase using standard Fmoc synthesis, cleaved from the resin and purified by HPLC. After purification, the peptides were resuspended in phosphate buffer, pH 7.5, adjusted to a concentration of 1 mM, and stored at -20 °C until use. To assess the effect of the N-terminal residue on reactivity, the peptides (100 µM) were treated with 2 equiv of a small molecule *o*-aminophenol, 2-amino-*p*-cresol (200 µM) in the presence of $K_3Fe(CN)_6$ (5 mM) in phosphate buffer, pH 7.5 (Figure 3.7a). The reactions were performed in triplicate and the modification was monitored by LC-MS (see Figure 3.7b for representative MS data for the modified peptides). Most N-terminal amino acids showed high levels of conversion (60-90%), but proline stood out as the only residue

that reached nearly complete modification (90-100%).

Optimization of the equivalents of *o*-aminophenol demonstrated that conversion was highest when using 2-5 equiv of the coupling partner (Figure 3.7c). Using more than 5 equiv of the aminophenol resulted in lower levels of modification of the peptide. This is most likely due to the ability of the aminophenol to react with itself at higher concentrations (appx. 1 mM). When using 10 equiv of the *o*-aminophenol we observed a product with a mass that corresponded to the condensation of 3 aminophenols (344 Da).

We also investigated possible differences in rates of reactivity between N-termini. The reaction of 2-amino-*p*-cresol with three different peptides was monitored over the course of 1 h (Figure 3.7d). The peptides (100 μM) were treated with 2 equiv of the *o*-aminophenol (200 μM) in the presence of ferricyanide (5 mM) and aliquots were quenched with excess tris(2-carboxyethyl) phosphine (TCEP) at the indicated time points. The proline terminal peptide not only reached the highest conversion, but also did so in a significantly shorter time than the other termini. Despite efforts to optimize conditions for all N-termini, proline still stood out as the most reactive terminus.

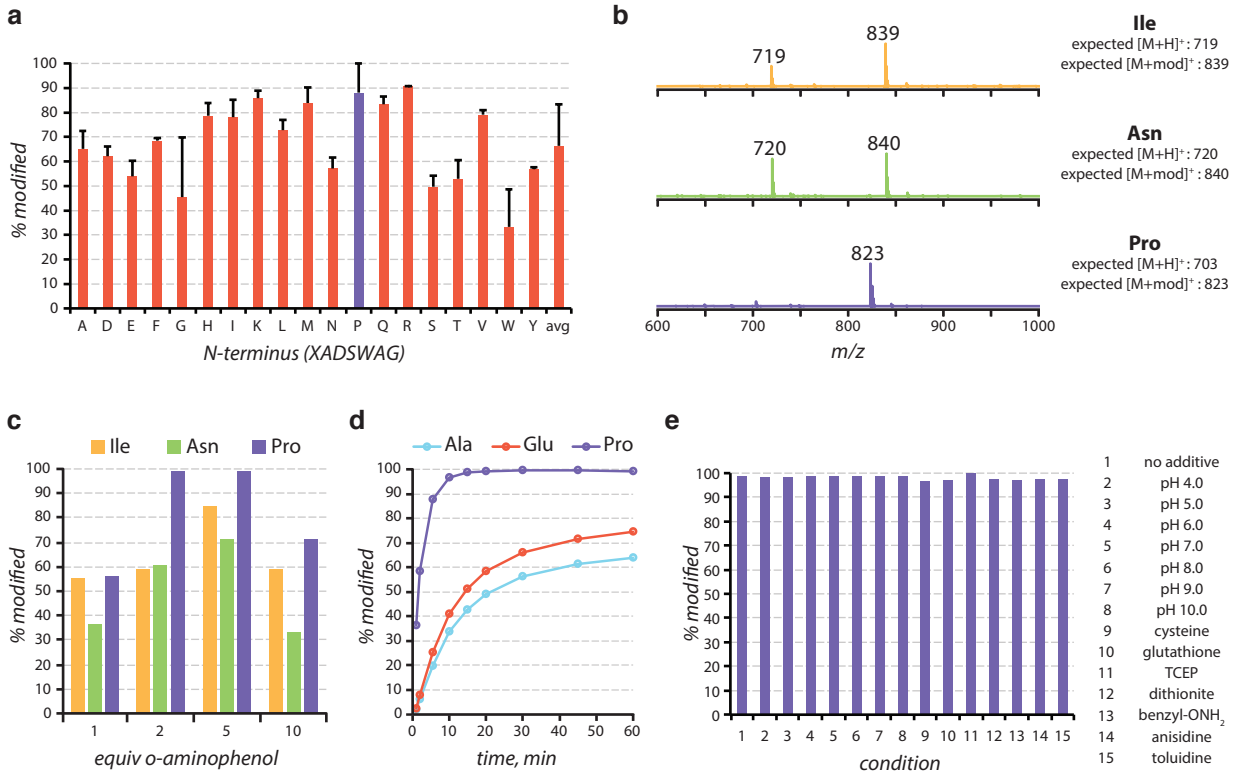


Figure 3.7. (a) A positional scan of the N-terminal amino acid was evaluated. Peptides with the sequence XADSWAG were tested for reactivity with 2-amino-*p*-cresol. The reactions were run with 100 μM peptide, 200 μM aminophenol, and 5 mM ferricyanide at pH 7.5 and analyzed by LC-MS. The same peptides were used for (c) and (d). (b) Representative MS traces used to quantify the percent modification. (c) Reactivity was highest when 2-5 equiv of *o*-aminophenol were used. When more equiv were used, trimerization of the aminophenol predominated. (d) The reaction on peptides required 15-60 min to reach completion, depending on the N-terminal amino acid. (e) The stability of the proline product was assayed by LC-MS. Modified PADSWAG was treated with 10 mM additive (reductant or nucleophile) or buffer (acidic or basic pH) for 8-18 h.

3.5 Reaction characterization with small molecule mimics

The reaction of N-terminal amines with *o*-aminophenols was characterized using small molecule mimics. The methyl esters of phenylalanine (H-Phe-OMe) and proline (H-Pro-OMe) were coupled to 2-amino-*p*-cresol using ferricyanide at pH 7.5. The crude products were characterized by two-dimensional NMR analyses and high-resolution mass spectrometry. Crude ^{1}H NMR spectra revealed that the products were formed with very high and clean conversion (Figure 3.14). The primary amine of H-Phe-OMe formed an iminoquinone product analogous to the one formed with aniline coupling partners (Figure 3.15). The secondary amine of proline prevents the formation of the *p*-iminoquinone, and thus forms *o*-quinone product **3** (Scheme 1, Figure 3.16). Given the different linkage obtained with proline, we verified the stability of the product to a variety of conditions. The proline terminal peptide, PADSWAG, was first modified with 2-amino-*p*-cresol. After purification, the modified peptide (100 μM) was exposed to reductants, nucleophiles and acidic and basic pH (10 mM additive or buffer). After 8-18 h of treatment the peptides were analyzed by LC-MS. No loss of product was observed under any of the conditions tested, demonstrat-

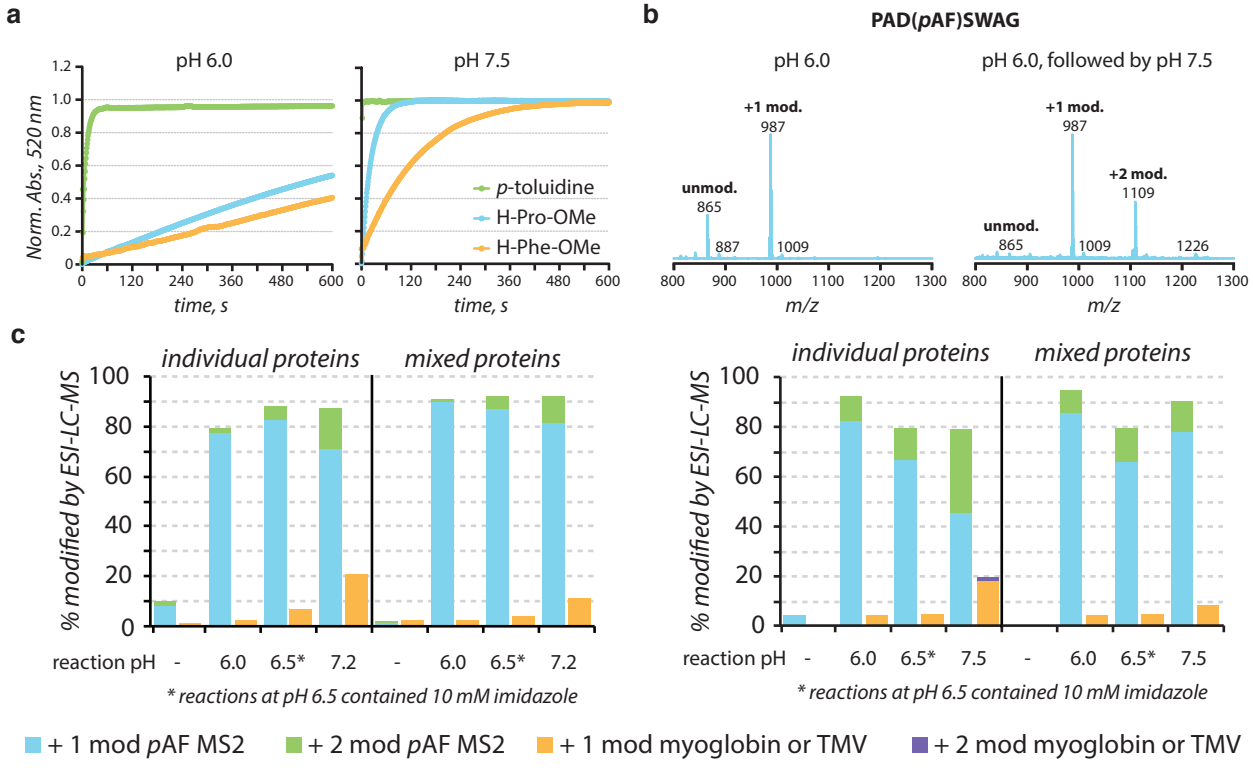


Figure 3.8. Characterization of the oxidative coupling reactions using small molecules. (a) Amine coupling partners were reacted with 4-methylcatechol as a model substrate. The reaction was followed by monitoring the product absorbance at 520 nm. Reactions were run under pseudo-first order conditions with 100 μM catechol, 1 mM amine, 100 mM ferricyanide, in 50 mM phosphate buffer. (b) A peptide containing both an N-terminal proline as well as a *p*-aminophenylalanine residue (PAD(*p*AF)SWAG) was tested for reactivity with 2 equiv of 2-amino-*p*-cresol at pH 6. An aliquot of the reaction was quenched and analyzed by LC-MS. The remainder of the reaction was purified and then reacted with 2 equiv of the aminophenol at pH 7.5 and then analyzed by LC-MS. (c) The relative reactivity of *p*AF residues and N-terminal residues was evaluated under several conditions. Proteins (20 μM) were reacted with 4 equiv of 2-amino-*p*-cresol (80 μM) with 1 mM ferricyanide in 10 mM buffer at the indicated pH for 30 min (reactions with imidazole contained 10 mM of the additive). Reactions with myoglobin are shown on the left and reactions with AG-TMV are shown on the right.

ing the hydrolytic stability of the product (Figure 3.7e).

In the process of characterizing the reaction products, it was observed that upon addition of ferricyanide the reaction mixture quickly turned a deep red. This provided means to monitor the progress of reaction. The different amine coupling partners – *p*-toluidine, H-Pro-OMe, and H-Phe-OMe – were reacted with 4-methylcatechol and the absorbance of the resulting solution was monitored at 520 nm to determine their relative rates of reactivity (Figure 3.8). The catechol substrate was used for these studies to simplify the rate law by eliminating the imine hydrolysis step. The reactions were run under pseudo-first order conditions with 0.1 mM catechol and 1 mM amine coupling partner in the presence of 10 mM ferricyanide. When the reaction was carried out at pH 6.0, only the aniline coupling partner underwent rapid coupling with the catechol. However, at pH 7.5 all three amines reacted efficiently with the catechol. The aniline coupling partner demonstrated the fastest coupling (< 30 s). The reaction with the proline analog reached completion nearly as rapidly (appx. 2 min), while the reaction with the primary aliphatic amine of phenylalanine required slightly longer reaction times (appx. 10 min). This demonstrates that the reaction can exhibit very high selectivity for aniline residues when it is carried out at low pH or for short reaction times.

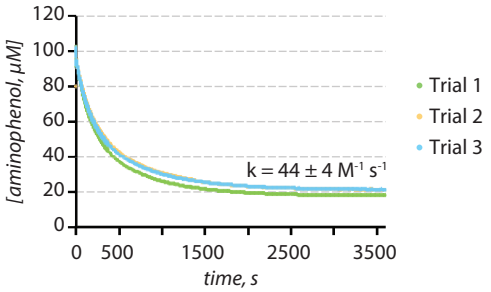


Figure 3.9. Plot of calculated aminophenol concentration vs. time used to determine the second-order rate constant.

In addition by quantifying the formation of product by absorbance, we were able to measure the second-order rate constant for the proline-based coupling (Figure 3.9). Reaction of 1 equiv of H-Pro-OMe (0.1 mM) and 1 equiv of 4-methylcatechol (0.1 mM) with 100 equiv of $K_3Fe(CN)_6$ (10 mM) was performed in triplicate at 25 °C. The second order rate constant for the coupling of H-Pro-OMe and 4-methylcatechol with ferricyanide was determined to be 44 ± 4 M^{-1} s^{-1}. While proline reacted rapidly with the electron-rich coupling partner, the small molecule studies indicated that aniline should react faster. The rate for the aniline reaction was too fast under these conditions to determine the second-order rate constant accurately. We next tested the relative reactivity of anilines and N-terminal amines on peptide and protein substrates to verify this preferential reactivity.

3.6 Preferential reactivity with aniline

Given the differences in reactivity observed at pH 6.0 and 7.5, we hypothesized that it was possible to modify the aniline side chain of *p*-aminophenylalanine (*p*AF) and the N-terminal proline amine sequentially. To test this hypothesis we synthesized a peptide containing both reactive moieties (PAD(*p*AF)SWAG). Only one modification was observed when the peptide was treated with 2-amino-*p*-cresol at pH 6.0 (Figure 3.8b). After modification at pH 6.0, the peptide was purified and subsequently treated with 2-amino-*p*-cresol at pH 7.5. Reaction at higher pH enabled a second modification of the peptide substrate.

The preferential reactivity with aniline side chains was also probed using protein substrates. The differential reactivity was investigated by comparing the reactivity of a protein containing a *p*AF residue to proteins without the artificial amino acid. The *p*AF residue was introduced into the coat protein of bacteriophage MS2, as described in the prior chapter. Myoglobin and a mutant of the tobacco mosaic virus (TMV) coat protein were used as native protein substrates. Reac-

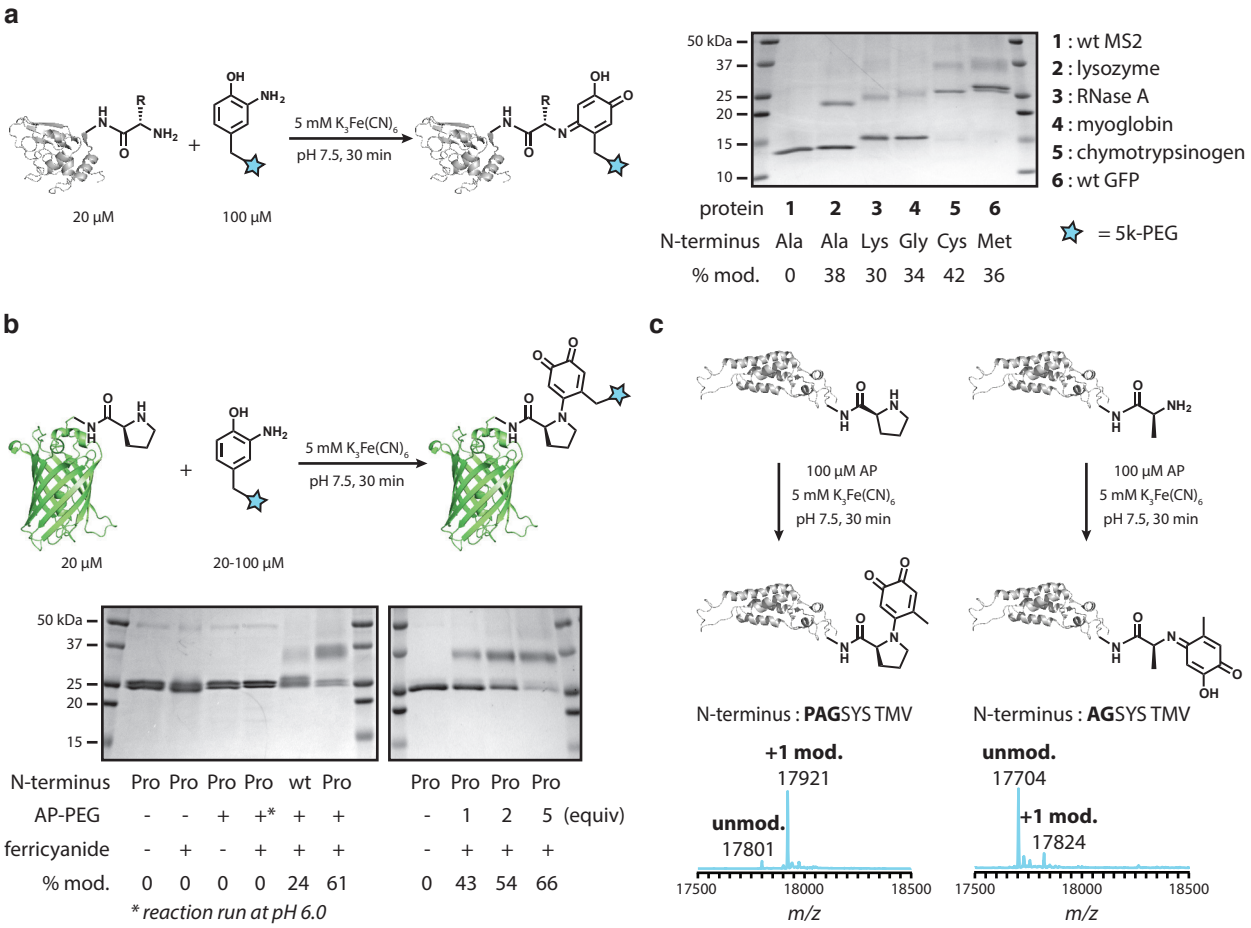


Figure 3.10. Protein modification with *o*-aminophenols. (a) The N-terminus of proteins was PEGylated using *o*-aminophenol-functionalized 5 kDa PEG and ferricyanide. Modification of wild type proteins with 5 kDa *o*-aminophenol-PEG was monitored by SDS-PAGE. (b) A proline was introduced to the N-terminus of GFP. Reactivity with *o*-aminophenol-PEG was monitored by SDS-PAGE. The proline terminal variant showed much higher levels of modification than the wild-type protein. No modification of N-terminal proline-GFP occurred at pH 6.0. (c) Mutants of TMV were reacted with 5 equiv of 2-amino-*p*-cresol with 1 mM $K_3Fe(CN)_6$ for 30 min and analyzed by LC-MS.

tions with 2-amino-*p*-cresol were either performed on the isolated, individual proteins or with the aniline containing protein mixed with the native protein substrate (Figure 3.8c). Addition of the aniline containing protein to the native protein decreased the N-terminal reactivity. In addition, MS2 showed significantly higher reactivity at all of the pHs tested, confirming that the oxidative coupling with *o*-aminophenols reacts preferentially with aniline residues.

3.7 Application of N-terminal oxidative coupling to proteins

The oxidative coupling reaction was first tested on proteins with native N-termini. Several protein substrates were reacted with *o*-aminophenol-functionalized 5 kDa PEG to evaluate the reaction on proteins (Figure 3.10a). The native proteins showed moderate levels of reactivity, which could be attributed to non-accessible N-termini or simply to the less reactive N-terminal residues. To test if proline-terminal proteins were more reactive, a proline residue was introduced to the N-terminus of GFP and the TMV coat protein. The N-termini were also slightly extended from the

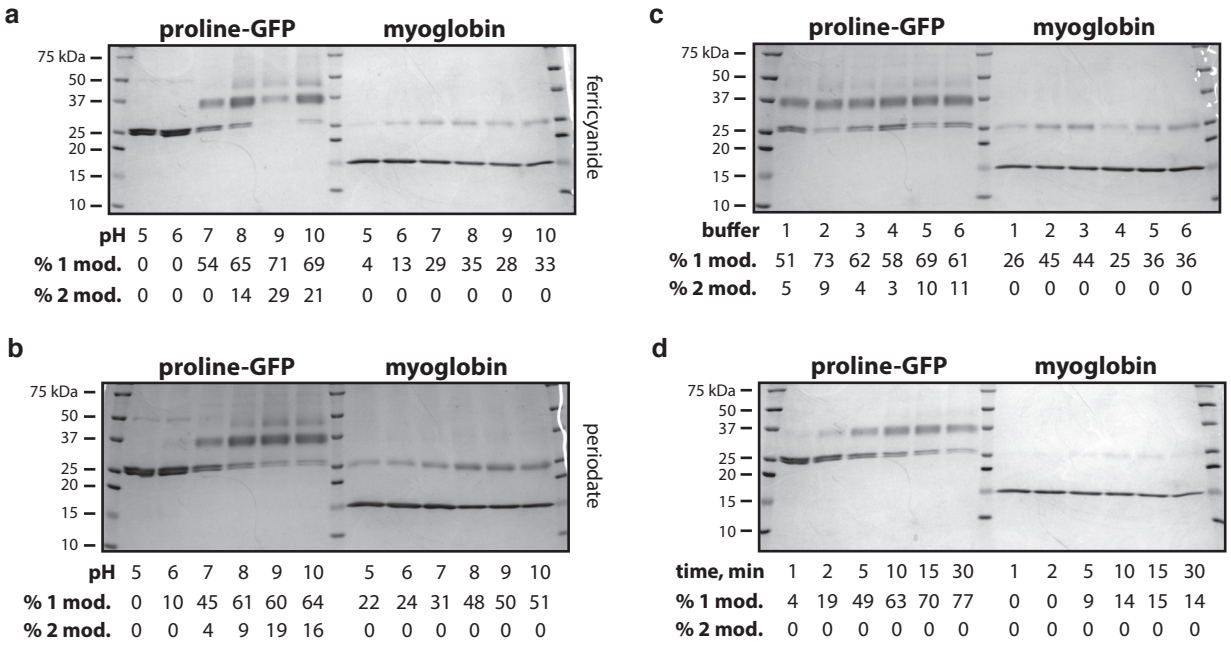


Figure 3.11. Optimization of the oxidative coupling on proteins was quantified by SDS-PAGE. The reactivity was assayed on a protein with the optimized proline terminus (proline-GFP) and one with its native terminus (myoglobin). Reactions were run with 20 μM protein, 200 μM *o*-aminophenol PEG and 5 mM oxidant and were quenched by the addition of loading buffer after 30 min. Both (a) ferricyanide and (b) periodate were tested as oxidants for this reaction. (c) Several buffer salts were tested for their effect on the reactivity (1 = MOPS, 2 = phosphate, 3 = TES, 4 = HEPES, 5 = Tris, 6 = bicine). (d) The reaction was monitored at various time points to determine the time course of the reaction. Aliquots of the reaction were taken out at the indicated time point and quenched by the addition of loading buffer.

native sequences to improve the accessibility of the N-termini (addition of PKT for GFP and PAG for TMV). The N-terminal extensions were chosen based on prior work on the optimization of pyrdoxial-5'-phosphate (PLP) mediated transamination.[31,32] The proline-GFP was treated with a variety of conditions to determine the specificity of the reaction (Figure 3.10b). Only at basic pH in the presence of both the *o*-aminophenol substrate and the oxidant was modification observed. Additionally, the proline-terminal variant showed significantly improved reactivity compared to that of the wild-type N-terminus. These high levels of modification were maintained even when only 1-2 equiv of the *o*-aminophenol PEG were used.

Reaction conditions for both native and proline terminal proteins were optimized by screening reactivity on myoglobin and proline-GFP. The reaction time, buffer, and pH were screened (Figure 3.11). Similar to the results obtained with peptide substrates, the reaction reached completion after appx. 15-30 min. In addition, most buffer salts tested were compatible with the reaction with the exception of buffers containing morpholine (MOPS) or piperazine moieties (HEPES) as was observed with peptide substrates. These buffers decreased the level of modification slightly, but did not completely inhibit reactivity. The effect of reaction pH was tested using both $K_3Fe(CN)_6$ and $NaIO_4$ as oxidants. Little reactivity was observed at acidic pH, and reactivity increased between pH 7.0 and 8.0.

N-terminal mutants of TMV were also evaluated for their reactivity with *o*-aminophenols. Two N-terminal mutants (PAG and AG) were treated with 5 equiv of 2-amino-*p*-cresol (0.1 mM) with

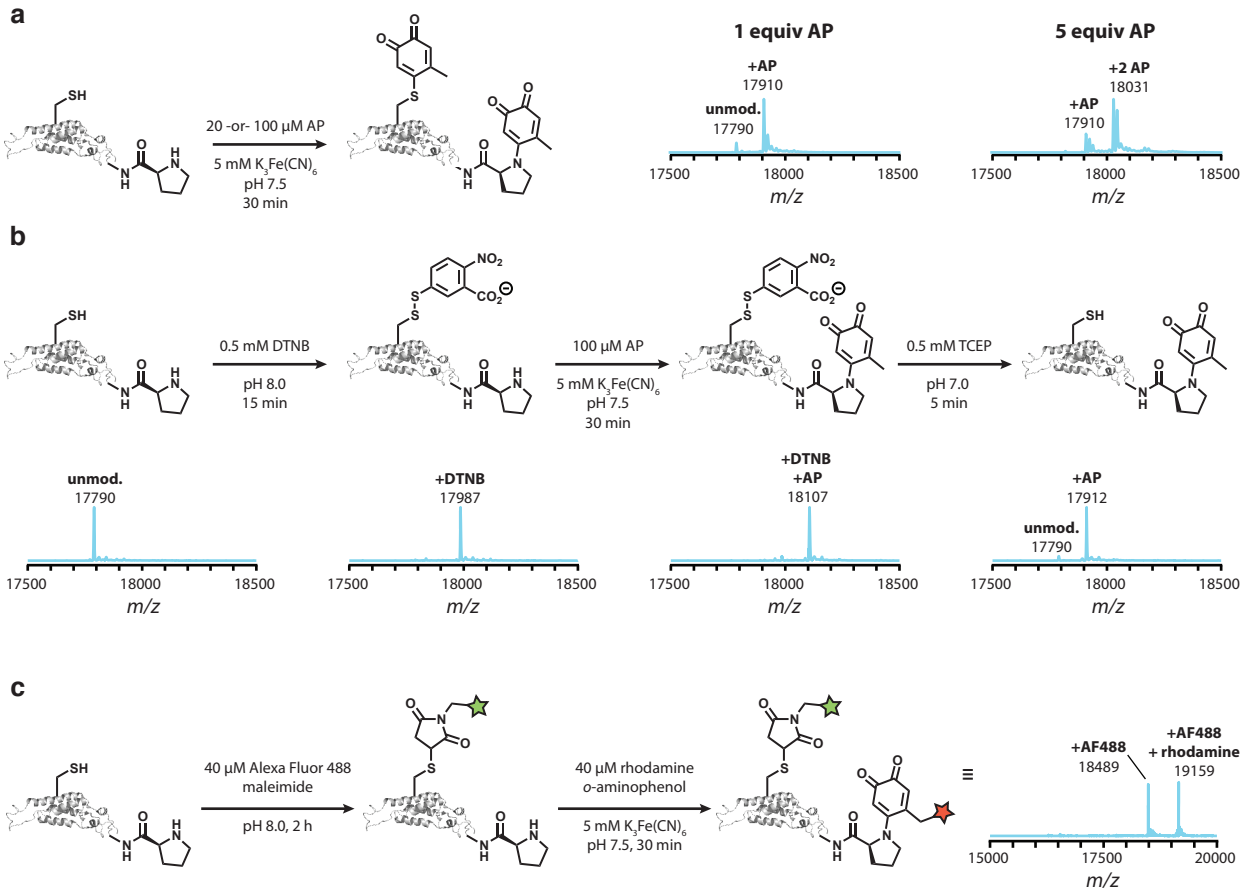


Figure 3.12. Compatibility with cysteine residues. (a) PAG S123C TMV was reacted with either 1 or 5 equiv of 2-amino-*p*-cresol in the presence of ferricyanide and analyzed by LC-MS. (b) PAG S123C TMV was reacted with small molecule substrates and analyzed by LC-MS. Cysteine residues were protected as a disulfide using Ellman's reagent (DTNB) before oxidative coupling. Subsequent reduction of the disulfide resulted in selective modification of the N-terminus. (c) PAG S123C TMV was labeled with two fluorophores. The cysteine was first alkylated with an Alexa Fluor maleimide. The N-terminal proline was then modified with a rhodamine-functionalized *o*-aminophenol.

1 mM $K_3Fe(CN)_6$ for 30 min. Analysis by LC-MS demonstrated that the proline-terminal mutant reached near complete modification while the alanine-terminal mutant showed low levels of conversion under these coupling conditions (Figure 3.10c). The compatibility of the reaction with cysteine residues was also tested using TMV. A single cysteine residue (S123C) was introduced into the TMV coat protein with a proline N-terminus (PAG S123C TMV). This mutant was reacted with 2-amino-*p*-cresol and analyzed by LC-MS (Figure 3.12). The cysteine residue also reacted with the *o*-aminophenol resulting in two modifications. The N-terminal proline could be selectively modified if the cysteine was first capped (Figure 3.12). The cysteine residue was protected in a disulfide bond by reaction with 5,5'-dithiobis-(2-nitrobenzoic acid) (DTNB, Ellman's reagent). After the oxidative coupling step the disulfide bond was readily reduced by TCEP, leaving the free cysteine and the modified N-terminus. This modification strategy allowed for the efficient modification of the N-terminus while maintaining a free cysteine residue. Alternatively, the cysteine residue was modified with a maleimide reagent, followed by modification at the N-terminus with an *o*-aminophenol reagent. This strategy allowed for direct, dual modification of the protein at both the cysteine residue and the N-terminus. Two fluorophores paired for Förster resonance energy transfer

(FRET, Alexa Fluor 488 C_5-maleimide and *o*-aminophenol functionalized rhodamine B) were conjugated to TMV using this strategy to create a templated array of chromophores for light harvesting applications.

The oxidative coupling reaction was also compared to the reaction of protein amino groups with activated esters. This acylation methodology is commonly employed and can be targeted to the N-terminus by controlling the reaction pH. The reactions were compared on creatine kinase, a protein with a native proline N-terminus. Reaction with 1-5 equiv of *o*-aminophenol PEG resulted in high levels of modification of creatine kinase (~50-60%) while reaction with 1-5 equiv of *N*-hydroxysuccinimide (NHS) PEG resulted in low levels of modification (5-15%, Figure 3.13). Only when a vast excess of the NHS PEG was used were high levels of modification achieved. These conditions also tended to result in multiple additions of the PEG substrate. Some over-modification was observed under the oxidative coupling conditions, but lowering the reaction pH slightly or using fewer equivalents of the *o*-aminophenol substrate prevented the over-modification while still maintaining high levels of single modification.

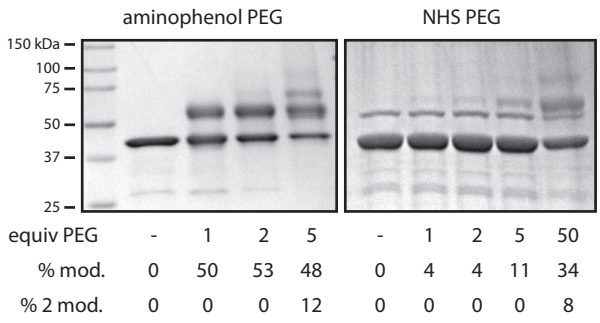


Figure 3.13. Modification of the N-terminus of creatine kinase by aminophenol PEG was compared to the reaction with NHS PEG and analyzed by SDS-PAGE.

In this chapter we identified conditions for the oxidative coupling of *o*-aminophenols to native amino acids. This method for site-selective bioconjugation serves as a complementary approach to the methodology outlined in Chapter 2. Peptides were initially screened to determine the reactive amino acids and optimize the reaction conditions. The N-terminus was identified as the primary target of modification. A positional scan of N-terminal residues revealed that proline N-termini showed improved levels of modification. The kinetics of the oxidative coupling of amines and *o*-catechols was probed using small molecule mimics of the coupling partners. The reaction was applied to protein substrates with native N-termini as well as engineered proline termini. Native termini could be modified with low yield, but the proline-terminal proteins showed high levels of modification. The oxidative coupling reaction discussed in this chapter was also used to modify TMV in conjunction with cysteine-maleimide chemistry to create a light harvesting array. This reaction allows for the rapid, facile attachment of *o*-aminophenol-functionalized molecules to any protein of interest. This variant of the oxidative coupling reaction expands the utility of the methodology as it can be applied to any protein substrate with a free N-terminus.

3.8 Materials and methods

Unless otherwise noted, the chemicals and solvents used were of analytical grade and were used as received from commercial sources. Analytical thin layer chromatography (TLC) was performed on EM Reagent 0.25 mm silica gel 60-F_{254} plates with visualization by ultraviolet (UV) irradiation at 254 nm and/or potassium permanganate stain. Purifications by flash chromatography were performed using EM silica gel 60 (230-400 mesh). The eluting system for each purification was determined by TLC analysis. Chromatography solvents were used without distillation. All organic solvents were removed under reduced pressure using a rotary evaporator. Water (dd-H_2O) used as reaction solvent was deionized using a Barnstead NANOpure purification system (ThermoFisher,

Waltham, MA). Centrifugations were performed with an Eppendorf Mini Spin Plus (Eppendorf, Hauppauge, NY).

Instrumentation and sample analysis

NMR. ^{1}H and ^{13}C spectra were measured with a Bruker AVB-400 (400 MHz, 100 MHz) or a Bruker AV-600 (600 MHz, 150 MHz) spectrometer, as noted. ^{1}H NMR chemical shifts are reported as δ in units of parts per million (ppm) relative to $CDCl_3$ (δ 7.26, singlet) or CD_3CN (δ 1.94, quintet). Multiplicities are reported as follows: s (singlet), d (doublet), t (triplet), dd (doublet of doublets), br (broad) or m (multiplet). Coupling constants are reported as a J value in Hertz (Hz). The number of protons (n) for a given resonance is indicated as nH, and is based on spectral integration values. ^{13}C NMR chemical shifts are reported as δ in units of parts per million (ppm) relative to $CDCl_3$ (δ 77.2, triplet) or CD_3CN (δ 118.26, singlet).

Mass Spectrometry. High-resolution electrospray ionization (ESI) and liquid chromatography with tandem mass spectrometry detection (LC-MS/MS) mass spectra were obtained at the UC Berkeley QB3/Chemistry Mass Spectrometry Ficlity. Matrix assisted laser desorption-ionization time-of-flight mass spectrometry (MALDI-TOF MS) was performed on a Voyager-DE system (PerSeptive Biosystems, USA) and data were analyzed using Data Explorer software. Peptide samples were co-crystallized with α-cyano-4-hydroxycinnamic acid in 1:1 acetonitrile (MeCN) to H_2O with 0.1% trifluoroacetic acid (TFA). Synthetic peptides and protein bioconjugates were analyzed using an Agilent 1200 series liquid chromatograph (Agilent Technologies, USA) that was connected in-line with an Agilent 6224 Time-of-Flight (TOF) LC/MS system equipped with a Turbospray ion source.

High Performance Liquid Chromatography. HPLC was performed on Agilent 1100 Series HPLC Systems (Agilent Technologies, USA) outfitted with an Agilent 1200 series automatic fraction collector. Sample analysis for all HPLC experiments was achieved with an inline diode array detector (DAD) and inline fluorescence detector (FLD). Semi-preparative reverse-phase HPLC of peptides was accomplished using a C18 stationary phase and a H_2O/ MeCN with 0.1% TFA gradient mobile phase.

Gel Analyses. For protein analysis, sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) was carried out on a Mini-Protean apparatus (Bio-Rad, Hercules, CA), using a 10-20% precast linear gradient polyacrylamide gel (Bio-Rad). The sample and electrode buffers were prepared according to Laemmli.[43] All protein electrophoresis samples were heated for 5-10 min at 95 °C in the presence of 1,4-dithiothreitol (DTT) to ensure reduction of disulfide bonds. Gels were run for 75-90 minutes at 120 V to separate the bands. Commercially available markers (Bio-Rad) were applied to at least one lane of each gel for assignment of apparent molecular masses. Visualization of protein bands was accomplished by staining with Coomassie Brilliant Blue R-250 (Bio-Rad). Gel imaging was performed on an EpiChem3 Darkroom system (UVP, USA). ImageJ was used to determine the level of modification by optical densitometry.

Peptide synthesis

General procedure for solid-phase peptide synthesis. Peptides were synthesized using standard

Fmoc-based chemistry on Tentagel S-OH resin (Advanced ChemTech, Louisville, KY). The base peptide sequence (ADSWAG) was synthesized on 1 g of resin and split into 20 x 50 mg aliquots for the addition of the final amino acid. The side chain protecting groups used were: Asn(Trt), Cys(Trt), Asp(tBu), Glu(tBu), His(Trt), Lys(Boc), pAF(Boc), Gln(Trt), Arg(Pbf), Ser(tBu), Thr(tBu), Trp(Boc), Tyr(tBu). The C-terminal amino acid (10 equiv) was preactivated at 0 °C with 5 equivalents of diisopropylcarbodiimide (DIC) and then coupled to the resin with 0.1 equivalents of *N,N*-dimethylaminopyridine (DMAP) as a catalyst. Deprotection of the Fmoc groups was performed with a 20 min incubation in a 20% v/v piperidine in dimethylformamide (DMF) solution. Coupling reactions were carried out using 20 equivalents of amino acid with 10 equivalents of 2-(6-chloro-1-*H*-benzotriazole-1-yl)-1,1,3,3-tetramethylaminium hexafluorophosphate (HCTU)[44] and 20 equivalents of *N,N*-diisopropylethylamine (DIPEA) in DMF for 20 min. Side-chain deprotection was accomplished using a 1-2 h incubation with a 95:2.5:2.5 ratio of TFA to H_2O to triisopropylsilane (TIPS). Peptides were cleaved from the resin by a 30-45 min incubation with a 100 mM sodium hydroxide solution. The resulting basic solution was neutralized with 100 mM phosphate buffer, pH 6.5. The cleaved peptides were purified using reverse-phase HPLC with a gradient of water/CH_3CN with 0.1% TFA. The organic solvent was removed on a vacuum centrifuge courtesy of Prof. Carolyn Bertozzi and the remaining water was removed by lyophilization. The lyophilized peptides were dissolved in 10 mM phosphate buffer, pH 7.5 and the concentration was adjusted to 1 mM based on absorbance.

Peptide modification

General methods. The small molecule coupling partners (2-amino-*p*-cresol and 4-methylcatechol) were purchased from Sigma Aldrich and purified by sublimation. The purified small molecules were stored frozen (at -80 °C) as a 500 mM solution in CH_3CN and diluted to 1 mM before use. $K_3Fe(CN)_6$ was purchased from Sigma Aldrich and used without further purification. Commercial peptides were dissolved in dd-H_2O to an approximate concentration of 1 mM. The dissolved peptides were stored between 4 °C and -20 °C to prevent degradation. When using high ionic strength buffers, peptide solutions were purified using Millipore ZipTip C18 Pipette Tips as specified by the supplier before analysis by MALDI-TOF MS.

General method for the modification of commercial peptides. To a solution of peptide (1 nmol, final concentration 100 μM) in buffer was added 2-amino-*p*-cresol or 4-methylcatechol (1-5 nmol, from a 1 mM stock containing 0.2% CH_3CN, final concentration 100-500 μM). To these reagents was added 1 μL of a solution of $K_3Fe(CN)_6$ (10 or 50 mM stock solution, final concentration 1-5 mM) to start the reaction. The reaction mixture was immediately vortexed and allowed to incubate at room temperature for the indicated time. The peptides were either purified or co-crystallized with matrix on the MALDI plate at the indicated time to stop the reaction. The percent modification was determined using MALDI-TOF MS. Given the high variability in signal in MALDI-TOF MS these values were used only for comparison.

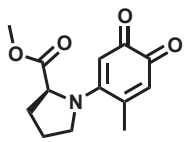
Angiotensin modification for MS/MS analysis. To a solution of angiotensin (10 nmol, final concentration 100 μM) in 50 mM phosphate buffer, pH 7.5 was added 2-amino-*p*-cresol (10 nmol, final concentration 100 μM). $K_3Fe(CN)_6$ was added (10 μL of a 10 mM solution, final concentration 1 mM) to start the reaction. After 30 min, the reaction mixture was purified on a C18 Sep-Pak

(Waters, conditioned with CH_3OH, equilibrated and washed with H_2O with 0.1% TFA, and eluted with CH_3CN). The solvent was removed under reduced pressure and the peptide was redissolved in dd-H_2O for analysis.

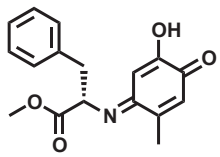
General method for the modification of synthetic peptides. To a solution of peptide (2 nmol, final concentration 100 μM) in buffer was added 2 equiv of 2-amino-*p*-cresol (4 nmol, final concentration 200 μM) followed by addition of 100 nmol of $K_3Fe(CN)_6$. The solution was immediately mixed after addition of the oxidant. Reactions were incubated at room temperature for the indicated time. The reactions were quenched by addition of buffered tris(2-carboxyethyl)phosphine hydrochloride (0.5 M solution of TCEP pH 7.0, 1 μmol) and analyzed by LC-MS.

Method for the sequential modification of PAD(*p*AF)SWAG. To a solution of peptide (20 nmol, final concentration 100 μM) in phosphate buffer, pH 6.0 was added 2-amino-*p*-cresol (40 nmol, final concentration 200 μM) followed by $K_3Fe(CN)_6$ (1 μmol, 5 mM final concentration). The reaction mixture was vortexed immediately following addition of the oxidant. After 30 min of incubation at room temperature the reaction mixture was purified on a C18 Sep-Pak (conditioned with CH_3OH, equilibrated and washed with H_2O with 0.1% TFA, and eluted with CH_3CN). The solvent was removed under reduced pressure and the peptide was redissolved in 10 mM phosphate buffer, pH 7.5. The purified peptide was modified for a second time following the general method described above.

Small molecule synthesis

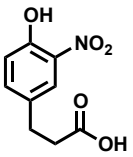


Synthesis of proline product (3). To 150 mL of buffered dd-H_2O (10 mM phosphate buffer, pH 7.5) was added 4-methylcatechol (24.2 mg, 0.2 mmol, purified by sublimation and recrystallization from toluene) in 400 μL of MeCN and L-proline methyl ester hydrochloride (33.1 mg, 0.2 mmol) in 1 mL 100 mM phosphate buffer pH 9.0. To the stirred solution was added potassium ferricyanide (690 mg, 2.1 mmol, as a solution in 50 mL of buffer). The reaction mixture was stirred at rt for 20 min. The reaction mixture was extracted with 3 portions of 100 mL 30% 2-propanol in $CHCl_3$. The combined organic layers were dried over sodium sulfate and the solvent was removed *in vacuo*. Over 95% of the material was recovered by mass. The resulting deep purple solid was taken up in $CDCl_3$ for characterization ^{1}H NMR (600 MHz, $CDCl_3$): δ 6.21 (s, 1H), 5.49 (s, 1H), 4.55 (m, 1H), 3.80 (m, 1H), 3.77 (s, 3H), 3.71 (m, 1H) 2.43 (m, 1H), 2.33 (s, 3H), 2.04 (m, 1H), 2.12 (m, 2H). ^{13}C NMR (150 MHz, $CDCl_3$): δ 24.06, 24.57, 30.62, 53.11, 53.60, 63.81, 103.45, 132.12, 146.41, 157.62, 171.53, 176.37, 182.16. HRMS (ESI) calculated for $C_{13}H_{15}NO_4$ ($[M+H]^+$) 250.1074, found 250.1073.

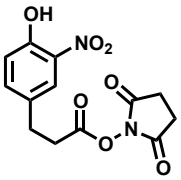


Synthesis of phenylalanine product (2). To 90 mL of buffered dd-H_2O (10 mM phosphate buffer, pH 7.5) was added 4-methylcatechol (12.4 mg, 0.1 mmol, purified by sublimation and recrystallization from toluene) in 200 μL of MeCN and L-phenylalanine methyl ester hydrochloride (21.6 mg, 0.1 mmol) in 1 mL of 100 mM phosphate buffer pH 9.0. To the stirred solution was added potassium ferricyanide (330 mg, 1 mmol, as a solution in 10 mL dd-H_2O). The reaction mixture was stirred at rt for 30 min. The solvent was then removed *in vacuo* and the resulting solid was taken up in CD_3CN for characterization. ^{1}H NMR (600 MHz, CD_3CN): δ 7.31 (m, 2H, J = 8.2),

7.26 (m, 1H) 7.21 (m, 2H, J = 8.0), 6.21 (s, 1H), 5.37 (s, 1H), 4.55 (m, 1H), 3.73 (s, 3H), 3.32 (m, 1H), 3.19 (m, 1H), 2.09 (s, 3H). ^{13}C NMR (150 MHz, CD_3CN): δ 183.48, 171.74, 155.24, 146.88, 136.88, 130.73, 130.10, 129.81, 128.95, 98.07, 57.67, 53.43, 37.34, 18.17.



Synthesis of nitrophenol acid (4). The nitrophenol was synthesized as previously reported.[45] Briefly, to a solution of 3-(4-hydroxyphenyl)propionic acid (5 g, 30.1 mmol) in 25 mL acetic acid at 15 °C was added a solution of fuming nitric acid (1.6 mL, 33.8 mmol) in acetic acid (4 mL). The solution immediately turned orange. After 15 min the reaction was quenched by addition to ice water. The precipitate was filtered and dried. The yellow precipitate was recrystallized from 1:1 ethanol:water to afford 3.07 g of a yellow powder (48% yield). ^{1}H NMR (400 MHz, $CDCl_3$): δ 10.47 (br s, 1H), 7.94 (d, 1H, J = 1.9), 7.45 (dd, 1H, J = 8.6, 1.9), 7.09 (d, 1H, J = 8.6), 2.94 (t, 2H, J = 7.4), 2.68 (t, 2H, J = 7.4). ^{13}C NMR (100 MHz, $CDCl_3$): δ 178.22, 153.86, 138.11, 133.47, 132.60, 124.22, 120.27, 35.19, 29.37. HRMS (ESI) calculated for $C_9H_8O_5N$ ($[M]^-$) 210.0408, found 210.0410 m/z.



Synthesis of nitrophenol NHS ester (5). The nitrophenol NHS ester was synthesized as previously reported.[45] To a solution of **4** (1 g, 4.7 mmol) and *N*-hydroxysuccinimide (0.65 g, 5.7 mmol) in CH_2Cl_2 at 0 °C was added *N*-(3-dimethylaminopropyl)-*N*'-ethylcarbodiimide hydrochloride (1.09 g, 5.7 mmol). The reaction mixture was stirred for 2 h and then diluted with CH_2Cl_2 and washed with water. The combined organic layers were dried over sodium sulfate and the solvent was removed *in vacuo*. The reaction afforded 1.38 g of a yellow solid (94% yield). ^{1}H NMR (400 MHz, $CDCl_3$): δ 10.49 (s, 1H), 7.98 (d, 1H, J = 2.1), 7.48 (dd, 1H, J = 8.6, 2.1), 7.12 (d, 1H, J = 8.6), 3.05 (t, 2H, J = 7.3), 2.93 (t, 2H, J = 7.3), 2.83 (s, 4H). ^{13}C NMR (100 MHz, $CDCl_3$): δ 169.06, 167.56, 154.09, 138.00, 133.54, 131.40, 124.52, 120.49, 32.48, 29.43, 25.70. HRMS (ESI) calculated for $C_{13}H_{12}O_7N_2$ ($[M-H]^-$) 307.0572, found 307.0568 m/z.

Analysis of kinetics using UV-Vis spectroscopy

Pseudo-first order reactions. To 50 mM phosphate buffer (either pH 6.0 or 7.5) was added 10 μL of a 10 mM solution of 4-methylcatechol (final concentration 100 μM) and 100 μL of a 10 mM solution of the coupling partner (*p*-toluidine, H-Pro-OMe, or H-Phe-OMe, final concentration 1 mM) for a volume of 900 μL. This solution was used as a blank before measurement. The absorbance at 520 nm was monitored every 1 s for 3600 s using a Cary 50 spectrophotometer. Within the first 1 s of measurement, 100 μL of a 100 mM solution of $K_3Fe(CN)_6$ (final concentration 10 mM) was added for a final reaction volume of 1 mL. A control reaction was run that excluded the amine coupling partner to verify the increase in absorbance at 520 nm did not arise from simple oxidation of the catechol substrate. The data were normalized to account for differences in extinction coefficients of the products.

Second order reactions. To 50 mM phosphate buffer pH 7.5 was added 10 μL of a 10 mM solution of 4-methylcatechol (final concentration 100 μM) and 10 μL of a 10 mM solution of H-Pro-OMe (final concentration 100 μM) for a final volume of 900 μL. This solution was used as a blank before measurement. The absorbance at 520 nm was monitored every 0.1 s for 3600 s at 25 °C using a Cary 50 spectrophotometer outfitted with a temperature control unit (courtesy of Prof. Carolyn

Bertozzi). Within the first 5 s of measurement, 100 μL of a 100 mM solution of $K_3Fe(CN)_6$ (final concentration 10 mM) was added. The reactions were performed in triplicate. A calibration curve was constructed by measuring the absorbance of reactions set up under pseudo-first order conditions with 10, 25, 50, 100, 250, 500 and 1000 μM catechol. The second order rate constants were calculated using the Cary software.

Aminophenol substrate synthesis

Synthesis of *o*-aminophenol rhodamine (6). To a solution of rhodamine B piperazine amide[46] (50 mg, 0.1 mmol) in DMF (1 mL) was added diisopropylethylamine (35 μL, 0.2 mmol) followed by **5** (30 mg, 0.1 mmol). After stirring for 2 h the solvent was removed *in vacuo*. The product was purified on silica gel with a gradient of CH_2Cl_2:CH_3OH (0 to 10% CH_3OH). HRMS (ESI) calculated for $C_{41}H_{45}O_6N_5$ ($[M+H]^+$) 704.3433, found 704.3436 m/z. To a 1 mg portion of the nitrophenol-functionalized rhodamine in 100 μL of 100 mM phosphate buffer, pH 6.5 was added an equal volume of a freshly prepared 100 mM solution of sodium dithionite in 100 mM phosphate buffer, pH 6.5. After 20 min, the excess dithionite was removed by purification on a C18 Sep-Pak according to the manufacturer's instructions. The eluent was concentrated to dryness. The rhodamine aminophenol was resuspended in DMF and stored frozen as a 50 mM solution.

Synthesis of *o*-aminophenol 5k-PEG. The *o*-aminophenol PEG substrate was synthesized as following a previously published protocol.[45] Briefly, to a solution of mPEG-NH_2 (MW=5000, 115 mg, 0.023 mmol) in CH_2Cl_2 (1 mL) was added triethylamine (7 μL, 0.05 mmol) and **5** (99 mg, 0.32 mmol as a 0.4 M solution in DMSO). The solution was stirred for 2 h at rt and then the solvent was removed *in vacuo*. Excess **5** was precipitated by the addition of water. The resulting solution was filtered through a 0.22 micron filter and three Sephadex size exclusion columns (one NAP-10 and two NAP-25, GE Healthcare) according to the manufacturer's instructions. The modified PEG was further purified by precipitation from CH_2Cl_2 with Et_2O three times. Reduction of the nitrophenol was accomplished with sodium dithionite. To a 1 mM solution of the modified PEG was added an equal volume of a freshly prepared 120 mM solution of sodium dithionite in 100 mM phosphate buffer, pH 6.5. Excess dithionite was removed by purification with three Sephadex size exclusion columns. The purified aminophenol PEG was lyophilized and resuspended in 10 mM phosphate buffer, pH 7.2. The concentration was adjusted to 1 mM by measuring the aminophenol absorbance at 290 nm. The final solution of *o*-aminophenol PEG was stored in single use aliquots at -20 °C until use.

Protein expression and purification

Generation of the proline terminal mutants. The QuikChange II Site Directed Mutagenesis Kit (Stratagene, La Jolla, CA) was used to introduce mutations to GFP and TMV.

The N-terminus of a previously generated GFP mutant was changed from alanine to proline using the primers :

sense : 5' - GATATACATATGCCCAAAACGGGCGAGGAGCTGTCCACC - 3'
antisense : 5' - GGTGAACAGCTCCTCGCCCGTTTTGGGCATATGTATATC - 3'

The N-terminus of TMV mutants was extended from the native N-terminus of SYS to PAGSYS using the primers :

sense : 5' - GAAGGAGATATACATATGCCTGCCGGCAGCTATAGCATTACC -3'
antisense : 5' - TGCTATAGCTGCCGGCAGGCATATGTATATCTCCTTCTTAAG - 3'

The TMV coat protein already had the following mutations : K53R, K68R (RR-TMV) and either S123C or T104K.

Expression and purification of TMV. The plasmids were transformed into BL21 DE3 RIL Codon+ cells for expression. The cells were grown in 1 L of LB with 100 μg/L of ampicillin at 37 °C until induction, at which point the temperature was lowered to 30 °C. The cells were induced by addition of 100 μL of 0.3 M isopropyl β-D-1-thiogalactopyranoside (IPTG) at an OD600 of ca. 0.6. The cells were collected by centrifugation at 7000 rpm for 20 min. The cells were resuspended in lysis buffer (20 mM TEA, pH 7.2) supplemented with 0.2 mg DNase and RNase A and 25 mg of $MgCl_2$. The cells were lysed by sonication and the cell debris was removed by centrifugation at 14000 rpm for 45 min. The clarified lysate was treated with an equal volume of 100% $(NH_4)_2SO_4$ and rotated at 4 °C for 10 min to precipitate the protein. The precipitated protein was collected by centrifugation at 16000 g for 45 min and then resuspended in 10 mL lysis buffer. The resuspended protein was applied to a DEAE-Sephadex column and purified using the following gradient (buffer A: 20 mM TEA, pH 7.2, buffer B: 20 mM TEA, pH 7.2 with 1 M NaCl): 0 min (0% B), 60 min (30% B), 61 min (100% B), 90 min (100% B), 91 min (0% B), 120 min (0 %B) with a flow rate of 3 mL/min. The fractions containing pure TMV were combined and then buffer exchanged into 10 mM phosphate buffer, pH 7.5 using spin concentrators with a 100 kDa MWCO.

Expression and purification of wt-GFP and Pro-GFP. The plasmids were transformed into T7 Express *lysY/I^q* cells for expression. The cells were grown in 1 L of LB with 100 μg/L of ampicillin at 37 °C until induction, at which point the temperature was lowered to 16-25 °C. The cells were induced by addition of 1 mL of 0.3 M IPTG and grown for ~18 h post induction. The cells were collected by centrifugation at 8000 rpm for 15 min and then resuspended in 10-15 mL lysis buffer (20 mM Tris, 150 mM NaCl, 5 mM EDTA, pH 8.0). The cells were lysed by sonication. The lysate was clarified by centrifugation at 9000 rpm for 1 h. The lysate was applied to 5-7 mL of chitin resin prerinsed with 100 mL of lysis buffer. The resin was rinsed with 100 mL wash buffer (20 mM Tris, 500 mM NaCl 1 mM EDTA, pH 8.0) and then incubated with 20 mL cleavage buffer (wash buffer with 50 mM sodium 2-mercaptoethanesulfonate (MESNa)) for 48 h. The cleaved protein was eluted with an additional 20 mL wash buffer and then buffer exchanged into 10 mM phosphate buffer, pH 7.5 using spin concentrators with a 10 kDa molecular weight cutoff (MWCO).

Protein Modification

Disulfide exchange with Ellman's reagent. Free cysteines were protected from potential modification by disulfide formation with Ellman's reagent (5,5'-dithiobis-(2-nitrobenzoic acid), DTNB). To a solution of PAG S123C TMV (780 μL of a 100 μM solution) was added DTNB (20 μL of a 20 mM solution in 100 mM phosphate buffer, pH 7.2 with 1 mM EDTA). The reaction mixture was incu-

bated at rt for 15 min and then the excess DTNB was removed by repeated (3-6 times) centrifugal filtration against a 100 kDa MWCO membrane. To reduce the disulfide, approximately 25 equiv of TCEP (as a 0.5 M solution, pH 7.0) was added to the protein sample.

General procedure for oxidative coupling on proteins. To a solution of proline terminal protein (PAG TMV, Pro-GFP, creatine kinase, 5-20 μM) in 10 mM phosphate buffer, pH 7.5 was added 2-10 equiv of the *o*-aminophenol coupling partner (20-100 μM). The solution was briefly vortexed and then 10-20 equiv (relative to the *o*-aminophenol) of $K_3Fe(CN)_6$ (as a 50 mM solution in 10 mM phosphate buffer, pH 7.5) was added. After 20-30 min, the reaction mixture was purified using a Sephadex size exclusion column (GE Healthcare) according to the manufacturer's instructions or using a 0.5 mL centrifugal filter with an appropriate molecular weight cut off (MWCO, Millipore). Modification was monitored by SDS-PAGE or LC-MS.

Cysteine alkylation with fluorescent maleimides. A solution of PAG S123C TMV (100 μM) in 10 mM phosphate buffer, pH 8.0 was incubated with 1 equivalent of an Alexa Fluor maleimide (Alexa Fluor 488 C_5-maleimide, 20 mM in DMSO) for 2 h. Unreacted dye was removed with a NAP-10 Sephadex size exclusion column (GE Healthcare) and by repeated centrifugal filtration against a 100 kDa MWCO membrane.

Modification with NHS PEG. A solution of creatine kinase (100 μM) in 10 mM phosphate buffer, pH 7.5 was incubated with NHS-functionalized PEG (tBoc-PEG-succinimidyl carboxymethyl, MW = 5000, Laysan Bio., Inc) for 2 h at room temperature. The PEG was dissolved in dd-H_2O immediately prior to use. The reaction mixture was analyzed by SDS-PAGE without any prior purification.

3.9 References

1. O'Hare, H. M.; Johnsson, K.; Gautier, A. *Curr. Opin. Struct. Biol.* **2007**, *17*, 488–494.
2. Alley, S. C.; Okeley, N. M.; Senter, P. D. *Curr. Opin. Chem. Biol.* **2010**, *14*, 529–537.
3. Tilley, S. D.; Joshi, N. S.; Francis, M. B.; Begley, T. P. In Wiley Encyclopedia of Chemical Biology; John Wiley & Sons, Inc., 2007.
4. Francis, M. B. In Chemical Biology; Schreiber, S. L.; Kapoor, T. M.; Wess, G., Eds.; Wiley-VCH Verlag GmbH, 2008; pp. 593–634.
5. Baslé, E.; Joubert, N.; Pucheault, M. *Chem. Biol*. **2010**, *17*, 213–227.
6. Doolittle, R. F. In Prediction of Protein Structure and the Prin-ciples of Protein Conformation; Fasman, G. D., Ed.; Springer US, 1989; pp. 599–623.
7. Saxon, E.; Bertozzi, C. R. *Science* **2000**, *287*, 2007–2010.
8. Kiick, K. L.; Saxon, E.; Tirrell, D. A.; Bertozzi, C. R. *Proc. Natl. Acad. Sci. U.S.A*. **2002**, *99*, 19–24.
9. Tornøe, C. W.; Christensen, C.; Meldal, M. *J. Org. Chem*. **2002**, *67*, 3057–3064.
10. (Rostovtsev, V. V.; Green, L. G.; Fokin, V. V.; Sharpless, K. B. *Angew. Chem. Int. Ed*. **2002**, *41*, 2596–2599.
11. Agard, N. J.; Prescher, J. A.; Bertozzi, C. R. *J. Am. Chem. Soc*. **2004**, *126*, 15046–15047.
12. Jewett, J. C.; Bertozzi, C. R. *Chem. Soc. Rev*. **2010**, *39*, 1272–1279.
13. Debets, M. F.; van Berkel, S. S.; Dommerholt, J.; Dirks, A. (Ton) J.; Rutjes, F. P. J. T.; van Delft, F. L. *Acc. Chem. Res*. **2011**, *44*, 805–815.

14. Blackman, M. L.; Royzen, M.; Fox, J. M. *J. Am. Chem. Soc*. **2008**, *130*, 13518–13519.
15. Devaraj, N. K.; Weissleder, R.; Hilderbrand, S. A. *Bioconjugate Chem*. **2008**, *19*, 2297–2299.
16. Patterson, D. M.; Nazarova, L. A.; Xie, B.; Kamber, D. N.; Prescher, J. A. *J. Am. Chem. Soc. 2012*, *134*, 18638–18643.
17. Niederwieser, A.; Späte, A.-K.; Nguyen, L. D.; Jüngst, C.; Reut-ter, W.; Wittmann, V. *Angew. Chem. Int. Ed*. **2013**, *52*, 4265–4268.
18. Cornish, V. W.; Hahn, K. M.; Schultz, P. G. *J. Am. Chem. Soc*. **1996**, *118*, 8150–8151.
19. Dirksen, A.; Dawson, P. E. *Bioconjugate Chem.* **2008**, *19*, 2543–2548.
20. Hooker, J. M.; Esser-Kahn, A. P.; Francis, M. B. *J. Am. Chem. Soc*. **2006**, *128*, 15558–15559.
21. Behrens, C. R.; Hooker, J. M.; Obermeyer, A. C.; Romanini, D. W.; Katz, E. M.; Francis, M. B. *J. Am. Chem. Soc*. **2011**, *133*, 16398–16401.
22. Rashidian, M.; Dozier, J. K.; Distefano, M. D. *Bioconjugate Chem*. **2013**, *24*, 1277–1294.
23. Lin, C.-W.; Ting, A. Y. *J. Am. Chem. Soc*. **2006**, *128*, 4542–4543.
24. Fernández-Suárez, M.; Baruah, H.; Martínez-Hernández, L.; Xie, K. T.; Baskin, J. M.; Bertozzi, C. R.; Ting, A. Y. *Nat. Biotechnol*. **2007**, *25*, 1483–1487.
25. Chen, I.; Howarth, M.; Lin, W.; Ting, A. Y. *Nat. Methods* **2005**, *2*, 99–104.
26. Mao, H.; Hart, S. A.; Schink, A.; Pollok, B. A. *J. Am. Chem. Soc*. **2004**, *126*, 2670–2671.
27. Yin, J.; Liu, F.; Li, X.; Walsh, C. T. *J. Am. Chem. Soc*. **2004**, *126*, 7754–7755.
28. Dawson, P. E.; Muir, T. W.; Clark-Lewis, I.; Kent, S. B. *Science* **1994**, *266*, 776–779.
29. Dawson, P. E.; Kent, S. B. *Annu. Rev. Biochem*. **2000**, *69*, 923–960.
30. Gilmore, J. M.; Scheck, R. A.; Esser-Kahn, A. P.; Joshi, N. S.; Francis, M. B. *Angew. Chem. Int. Ed.* **2006**, *45*, 5307–5311.
31. Scheck, R. A.; Dedeo, M. T.; Iavarone, A. T.; Francis, M. B. *J. Am. Chem. Soc*. **2008**, *130*, 11762–11770.
32. Witus, L. S.; Moore, T.; Thuronyi, B. W.; Esser-Kahn, A. P.; Scheck, R. A.; Iavarone, A. T.; Francis, M. B. *J. Am. Chem. Soc*. **2010**, *132*, 16812–16817.
33. Witus, L. S.; Netirojjanakul, C.; Palla, K. S.; Muehl, E. M.; Weng, C.-H.; Iavarone, A. T.; Francis, M. B. *J. Am. Chem. Soc*. **2013**.
34. Geoghegan, K. F.; Stroh, J. G. *Bioconjugate Chem*. **1992**, *3*, 138–146.
35. Li, X.; Zhang, L.; Hall, S. E.; Tam *, J. P. *Tetrahedron Lett*. **2000**, *41*, 4069–4073.
36. Tam, J. P.; Yu, Q.; Miao, Z. *Pept. Sci*. **1999**, *51*, 311–332.
37. Jackson, H.; Kendal, L. P. *Biochem. J*. **1949**, *44*, 477–487.
38. Mason, H. S.; Peterson, E. W. *J. Biol. Chem*. **1955**, *212*, 485–493.
39. King, C. M.; Kriek, E. *Biochim. Biophys. Acta* 1965, *111*, 147–153.
40. Gutmann, H. R.; Nagasawa, H. T. *J. Biol. Chem*. 1960, *235*, 3466–3471.
41. King, C. M.; Gutmann, H. R.; Chang, S. F. *J. Biol. Chem*. **1963**, *238*, 2199–2205.
42. King, C. M.; Chang, S. F.; Gutmann, H. R. *J. Biol. Chem*. **1963**, *238*, 2206–2212.
43. Laemmli, U. K. *Science* **1970**, *227*, 680–685.
44. C. A. Hood, G. Fuentes, H. Patel, K. Page, M. Menakuru, J. H. Park, *J. Pept. Sci*. **2008**, *14*, 97–101.
45. Obermeyer, A. C.; Jarman, J. B.; Netirojjanakul, C.; El Muselmany, K.; Francis, M. B.; *Angew. Chem. Int. Ed*. **2013**, early view.
46. Nguyen, T.; Francis, M. B. *Org. Lett*. **2003**, *5*, 3245–3248.

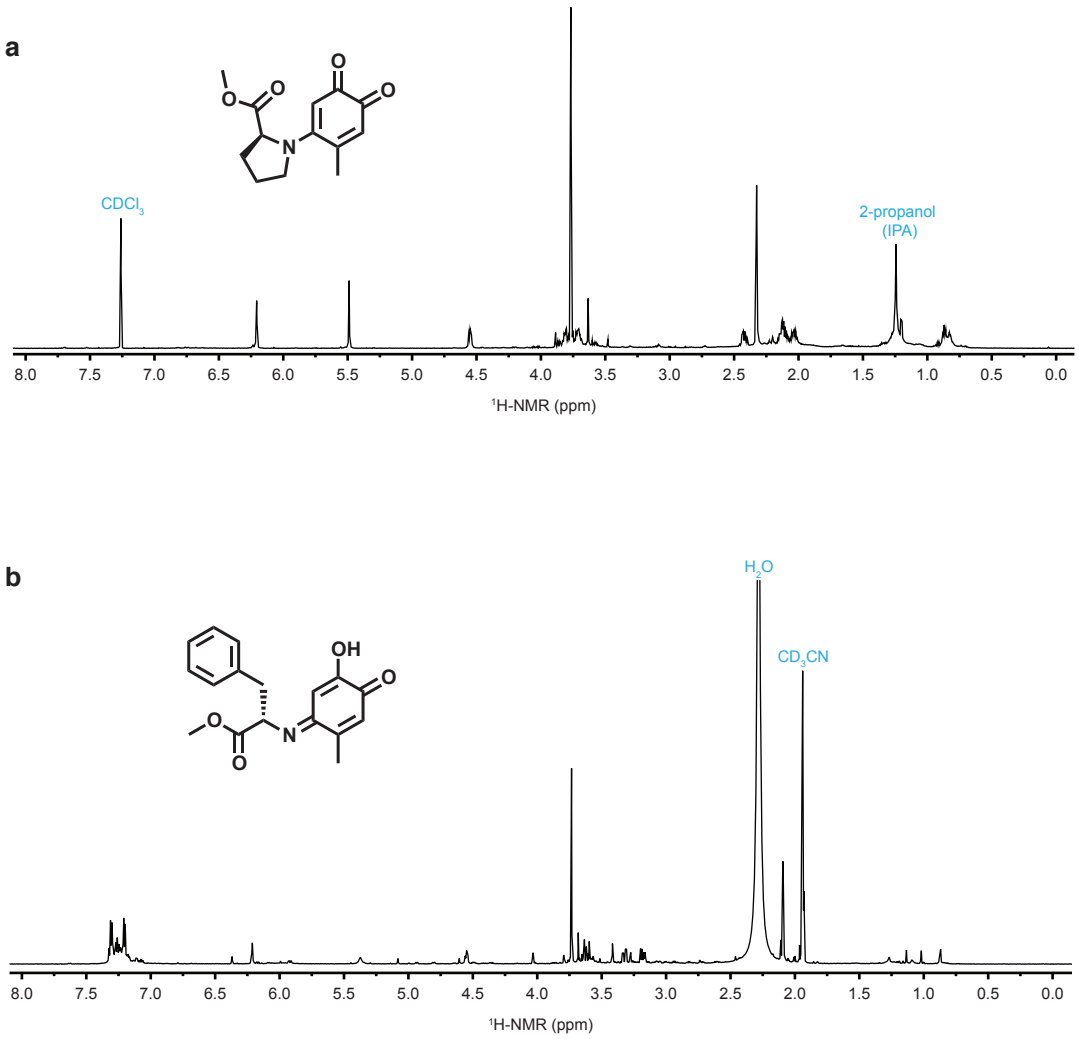


Figure 3.14. Crude ^{1}H NMR of the small molecule products. (a) ^{1}H NMR of the crude product of the ferricyanide-mediated coupling of 4-methylcatechol and H-Pro-OMe. (b) ^{1}H NMR of the crude product of the ferricyanide-mediated coupling of 4-methylcatechol and H-Phe-OMe. See figures 3.15 and 3.16 for peak assignments.

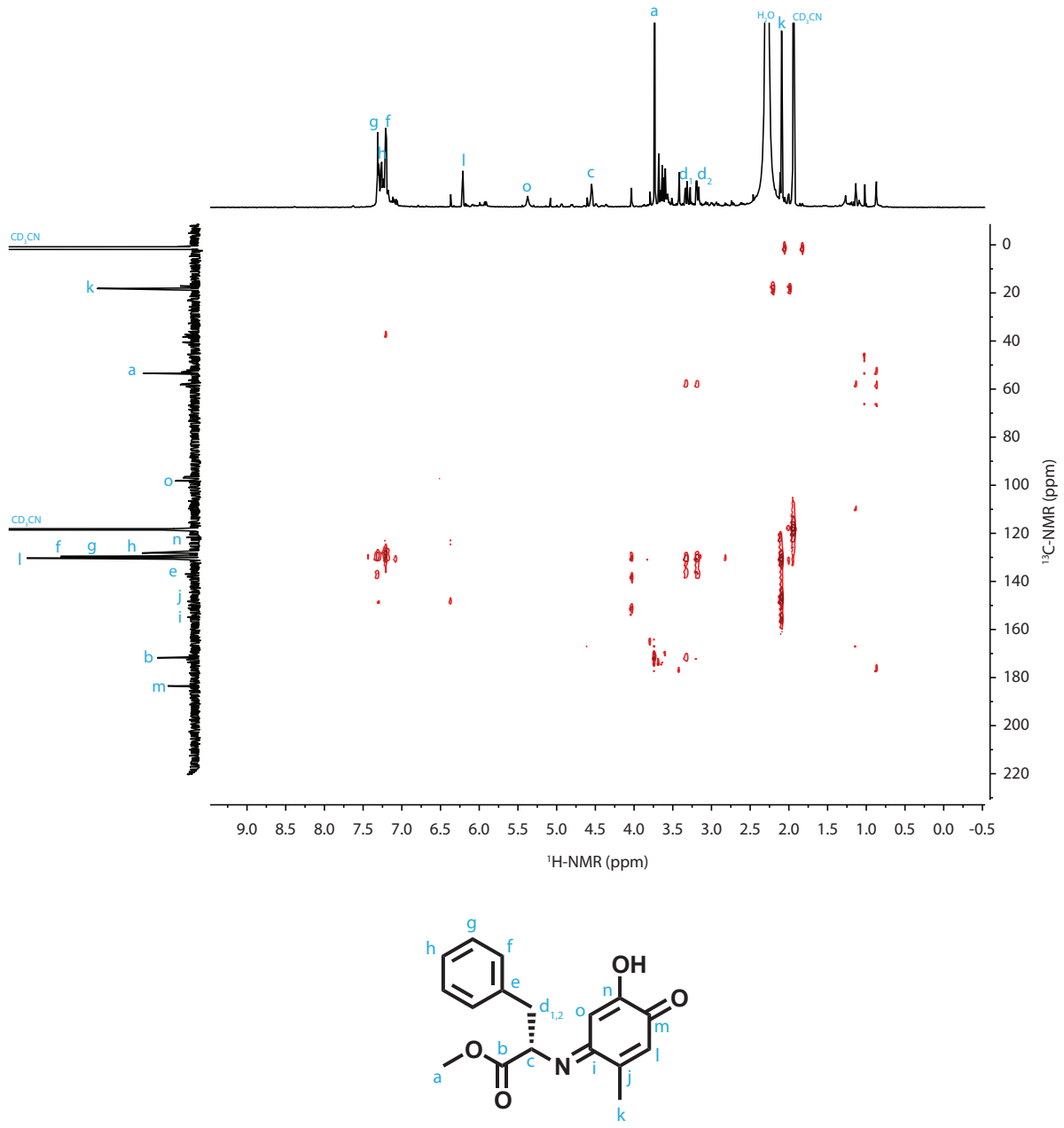


Figure 3.15. HMBC spectrum of **2** in CD_3CN.

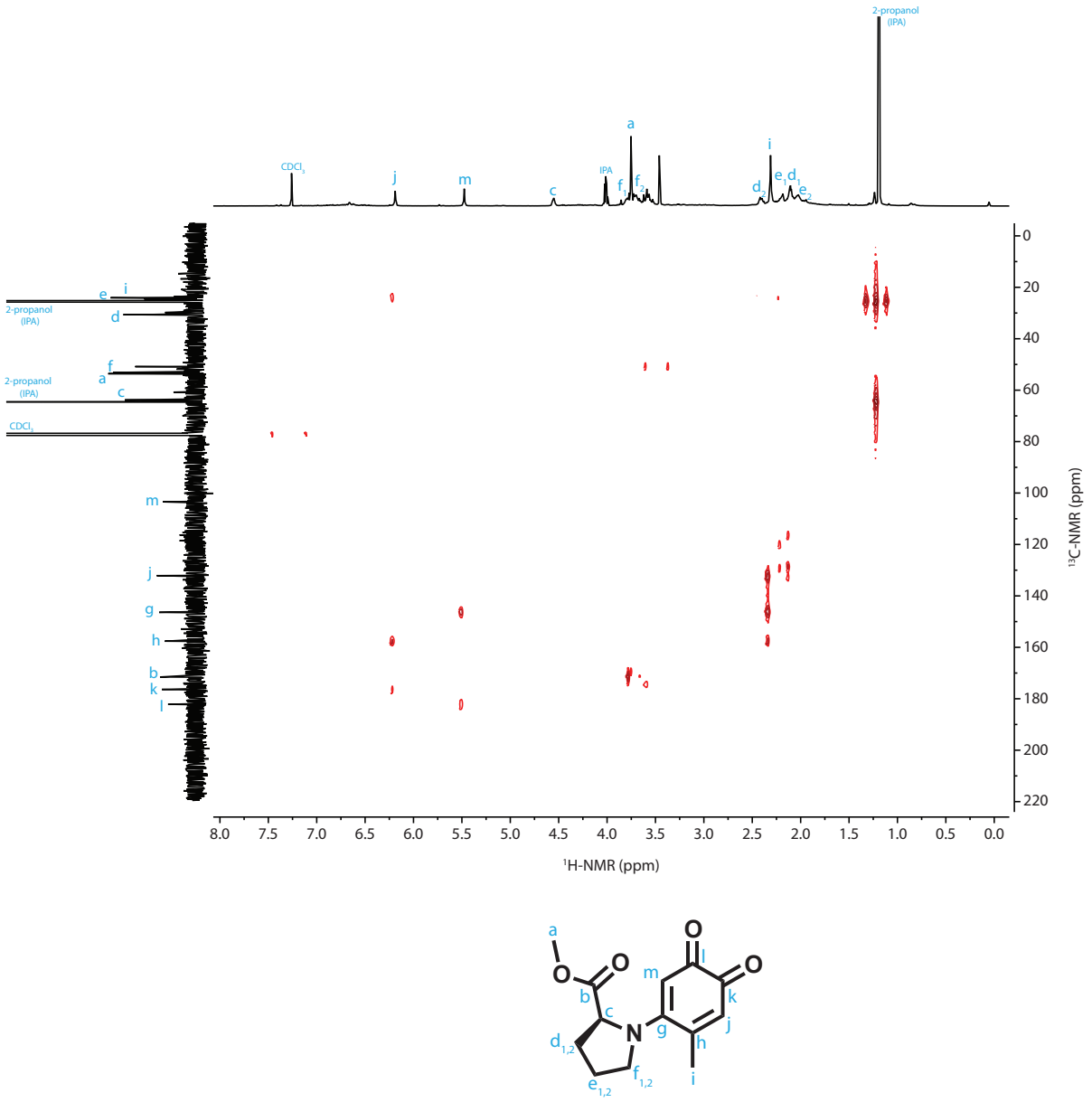


Figure 3.16. HMBC spectrum of **3** in $CDCl_3$.

Chapter 4
MS2-based agents for detection of atherosclerosis

Abstract

Cardiovascular diseases present a serious threat to human health, with nearly one in three deaths in the United States attributable to these afflictions. The development of atherosclerosis targeted imaging agents has the capability to provide molecular information about pathological clots, potentially improving detection, risk stratification, and therapy of cardiovascular diseases. Nanocarriers are a promising platform for the development of molecular imaging agents as they can be modified to have external targeting ligands and internal functional cargo. In this chapter, we report the synthesis and use of chemically functionalized bacteriophage MS2 capsids as protein-based nanoparticles for fibrin imaging and initial studies for vascular cellular adhesion molecule 1 (VCAM1) imaging. The capsids were modified using an oxidative coupling reaction, conjugating ~90 copies of targeting peptides to the exterior of each protein shell. The ability of the multivalent, targeted capsids to bind fibrin was first demonstrated by determining the impact on thrombin-mediated clot formation. The modified capsids out-performed the free peptides and were shown to inhibit clot formation at effective concentrations over ten-fold lower than the monomeric peptide alone. The installation of near-infrared fluorophores on the interior surface of the capsids enabled optical detection of binding to fibrin clots. The targeted capsids bound to fibrin, exhibiting higher signal-to-background than control, non-targeted MS2-based nanoagents. The VCAM1 peptide was successfully attached to the capsid, however the modified capsids did not bind to endothelial cells over-expressing VCAM1.

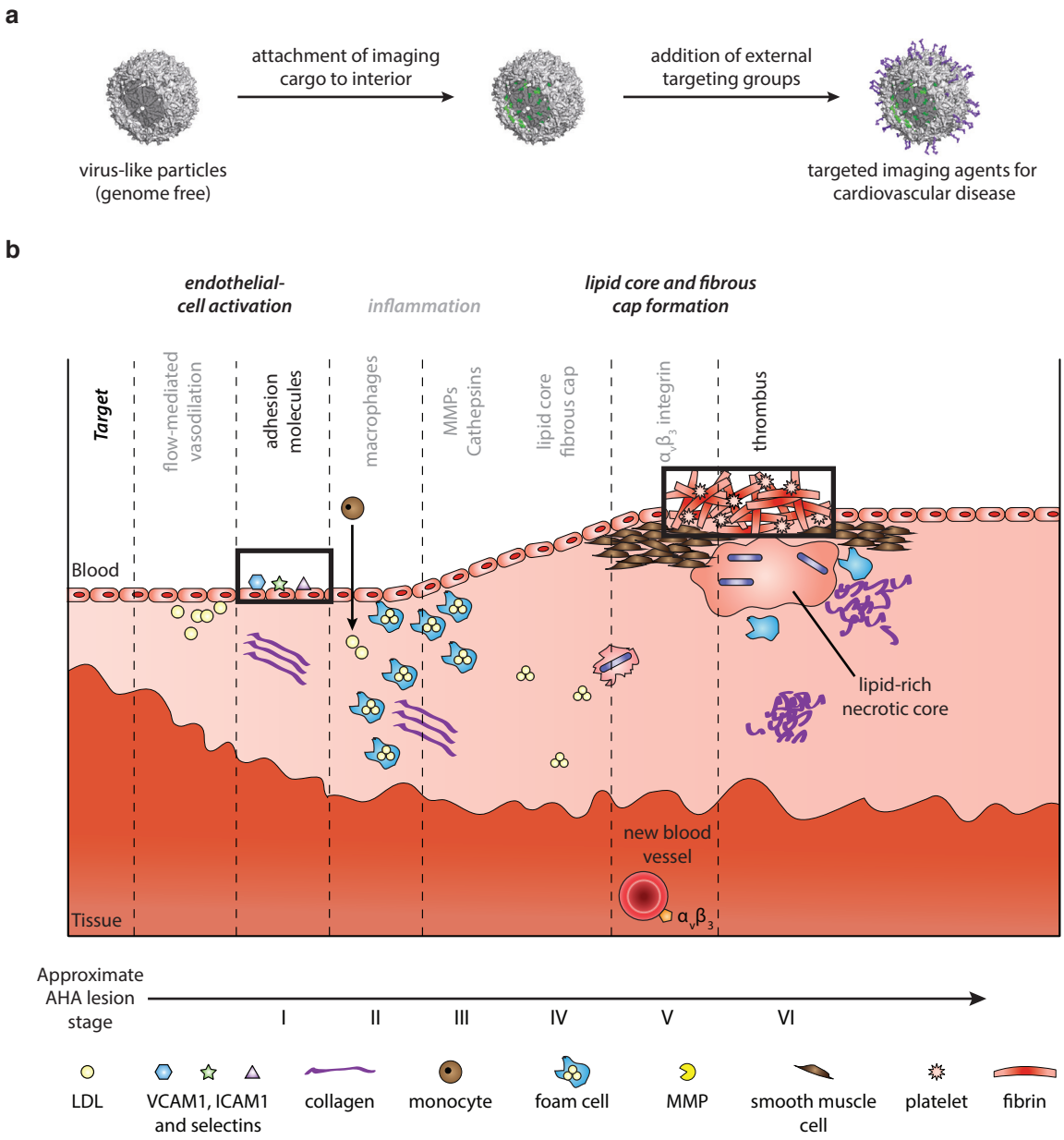
4.1 Targeted imaging of cardiovascular disease

Cardiovascular diseases are the leading cause of mortality in the United States, currently accounting for one in three deaths overall.[1] The noninvasive identification and molecular characterization of atherosclerosis could lead to improved detection of many of these conditions, and could provide a useful means to monitor the effectiveness of different treatment methods (Figure 4.1).[2–6] A set of targeted agents that can specifically associate with markers of various disease stages would enable the delivery of PET radiotracers, MRI contrast enhancement agents, or long wavelength responsive chromophores for molecular imaging. Two molecular markers were selected for the development of these targeted imaging agents. Vascular cellular adhesion molecule 1 (VCAM1) was chosen for the detection of early stage endothelial cell activation and fibrin was chosen for the detection of thrombi present in advanced atherosclerosis.

Nanomaterials could have several distinct advantages for the targeting and imaging of molecular markers of cardiovascular disease (Figure 4.1). Chief among them is their large surface area, relative to small molecule diagnostic agents, which could allow the presentation of multiple copies of the targeting groups for enhanced binding avidity. Nanoscale carriers can also be tailored to house multiple copies of the reporter moieties, allowing greater signal-to-noise ratios to be achieved. With these concepts in mind, many synthetic nanoparticles have been developed for the purpose of targeted imaging,[7] including synthetic polymers,[8] liposomes,[9] dendrimers,[10,11] and inorganic nanoparticles.[12,13] Another approach to preparing nanocarrier platforms relies on functionalized biomolecular assemblies, such as viruses and virus-like particles (VLPs),[14–17] as well as mammalian protein cages including heat shock cages,[18] ferritins,[19–21] and vaults.[22] These protein-based nanoparticles have the natural advantage of being monodisperse, non-toxic, and biodegradable.

Recent work has shown the utility of biomolecular assemblies, such as VLPs, for targeted delivery to cancer[23–25] and sites of inflammation.[14,18,19,26] In these examples, target binding peptide sequences have been incorporated directly into the protein monomers,[18] or synthetic targeting agents have been installed using chemical bioconjugation techniques.[25] Although both strategies can be quite effective, the latter approach has the advantage of increased modularity, and it can be used with virtually any targeting agent, including aptamers,[27,28] peptoids,[29,30] and engineered protein binders, such as single-chain antibody fragments.[31,32] Our group has previously used this strategy to functionalize bacteriophage MS2 VLPs for targeted delivery.[24]

Prior work on MS2-based delivery agents has shown that MS2 capsids can be heterologously expressed in *E. coli*, where they self-assemble from 180 sequence-identical monomers to form 27 nm diameter icosahedral capsids. The genomic material can be removed, yielding hollow capsids that are stable toward a wide range of temperature, pH, and ionic strength conditions, and are amenable to genetic mutation and sequence insertions.[33–35] The assembled capsids contain 32 pores (~ 2 nm) that permit the diffusion of small molecules to the interior of the VLPs. These pores, in conjunction with the mutation of a residue on the interior surface of the capsid to a uniquely reactive cysteine, have been used for the conjugation of up to 180 maleimide-functionalized cargo molecules inside each carrier.[24] The subsequent attachment of targeting groups to the exterior surface has been accomplished with high efficiency by targeting an unnatural amino acid, *p*-aminophenylalanine (*p*AF).[34] The side chain aniline group of this amino acid can be modified with *o*-aminophenol-containing targeting groups using the periodate or ferricyanide-mediated oxidative coupling.[36] Other strategies for preparing targeted MS2-based carriers have relied on the attachment of synthetic peptides to exterior lysine residues.[25,37]



c *advantages of virus-like particle based imaging agents*

- multivalent display of targeting ligands for improved binding
- multiple cargo attachment sites for increased specific activity or signal-to-noise
- modular design allows many combinations of imaging groups and targeting ligands
- excellent capsid stability under biological conditions, *in vitro* and *in vivo*
- protein-based capsids are non-toxic, biodegradable and monodisperse
- exterior polymers can reduce immunogenicity and shield surface charge

Figure 4.1. Construction of nanoscale diagnostic imaging agents using viral capsids. (a) Schematic for the construction of the imaging agents using virus-like particles based on bacteriophage MS2. (b) Work discussed in this chapter targets two molecular markers of atherosclerosis, adhesion molecules and thrombi. Targets from both early and late-stages of the disease were selected (figure adapted from reference 3). Advantages of virus-like particle based imaging agents are summarized in (c).

In this chapter, the use of this oxidative coupling strategy to synthesize bacteriophage MS2-based VLPs that are conjugated to protein and peptide-based targeting groups is described. These structures also contain multiple copies of fluorescent dyes for facile detection in imaging experiments. The fibrin-targeted capsids were found to inhibit fibrin clot formation at a concentration significantly lower than the free peptide targeting group, demonstrating the benefits of using multivalent nanoscale structures for target binding. Additionally, functionalized MS2 capsids were used for the *in vitro* optical imaging of fibrin clots. The VCAM1-targeted capsids did not bind to their cellular target in initial experiments and were not pursued further.

4.2 Fibrin targeting using single chain antibody fragments

Our first attempt at targeting fibrin relied on single chain antibody fragments (scFvs). Antibody fragments are promising targeting agents because they maintain the selectivity of their parent antibodies, are not highly charged, are readily expressed in *E. coli*, and can be generated by phage display to bind nearly any target. A single chain antibody fragment derived from a fibrin-binding antibody (59D8) was previously reported to successfully target fibrin clots *in vitro*.[38] We investigated the use of this scFv as a targeting group for fibrin imaging. The gene for the scFv was synthesized from overlapping primers using PCR and then subsequently cloned into the pET 22b (+) vector.

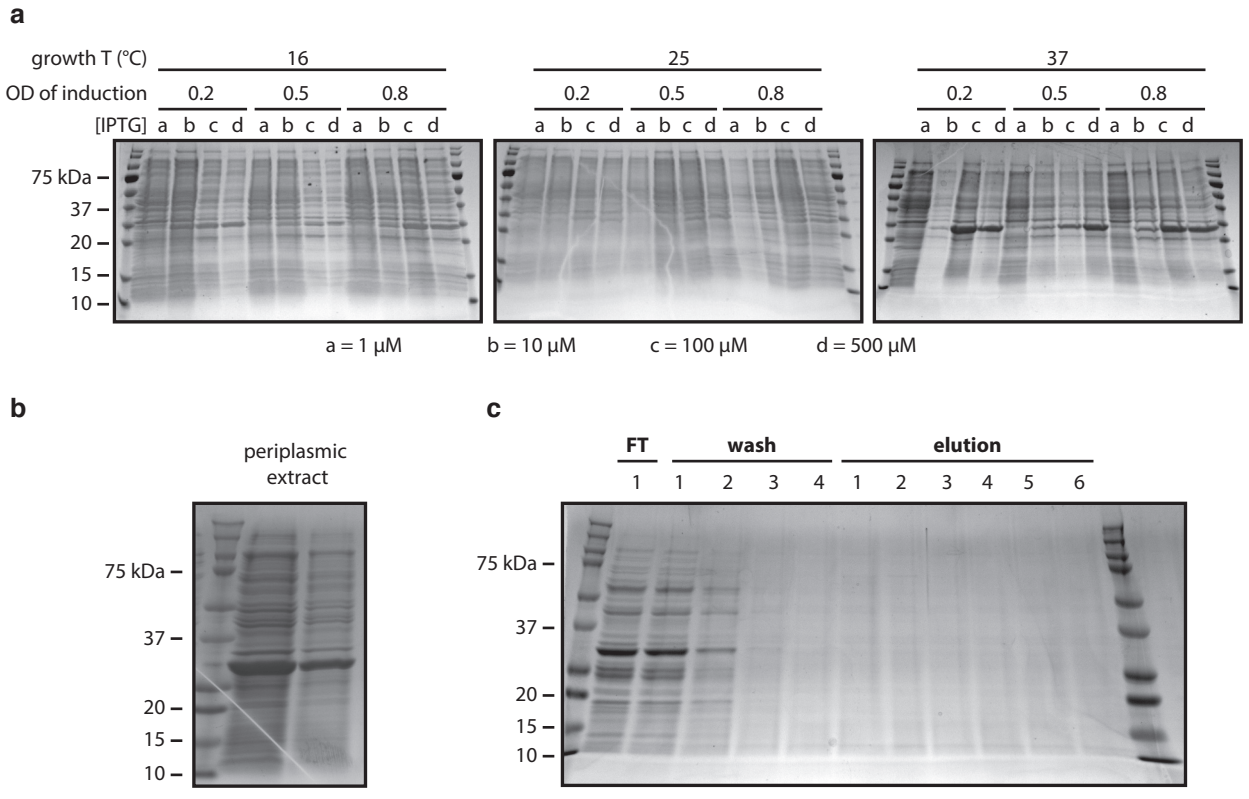
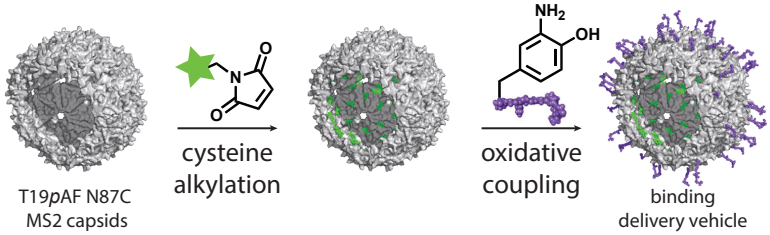


Figure 4.2. Expression and purification of scFv constructs. (a) Conditions for the expression of the scFv in *E. coli* were optimized and analyzed SDS-PAGE of the total cell lysate. (b) The scFv was partially purified by extraction from the periplasm. (c) Purification using Ni-NTA was not successful. The His-tagged protein did not bind to the resin under any native conditions.

The protein was designed to include a pelB leader at the N-terminus for export to the periplasm and a His_6-tag at the C-terminus for purification with Ni-NTA. Expression of the scFv in *E. coli* was optimized by screening expression temperature, optical density (OD) of induction, and concentration of inducer (IPTG). Several expression conditions were identified for the over-expression of the scFv in the total cell lysate (Figure 4.2a), but expression at 37 °C with induction of protein expression in early log phase (OD = 0.2) using 100 μM IPTG was selected for all future experiments. The scFv was successfully expressed in *E. coli* and the periplasmic proteins were extracted by cold osmotic shock (Figure 4.2b). However, we were unable to purify the scFv using the affinity of the His_6-tag for Ni-NTA. The scFv did not bind to the Ni-NTA resin under native conditions (Figure 4.2c). The scFv was purified using Ni-NTA under denaturing conditions, but we were unable to refold the protein. Several protein sequences were evaluated, but none were successfully purified. Due to these difficulties with the protein-based targeting group, we identified an alternative fibrin-binding targeting group for the synthesis of the MS2-based imaging agents.

a



b

peptide	sequence*	target
GPR	H_2N-GP**R**PPGGSKGY(C)-CO_2H	fibrin
GPS	H_2N-GP**S**PPGGSKGY(C)-CO_2H	non-binding control
VCAM1	H_2N-VHPKQHRGGSKGY-CO_2H	VCAM1

*The C-terminal residue was varied depending on the desired application and required modification.

c

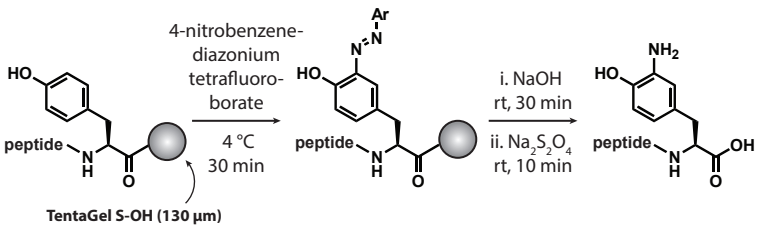


Figure 4.3. Design of a viral capsid-based targeting system. (a) A scheme is shown for the synthesis of multivalent viral capsids for cardiovascular disease imaging. (b) Peptides used for attachment to MS2 or binding to fibrin or VCAM1. Peptides with a cysteine at the C-terminus were modified with a fluorescent dye and used for direct binding, while peptides with a C-terminal tyrosine were coupled to MS2 and tested for binding. (c) Scheme for *o*-aminophenol peptide synthesis. The peptides were synthesized using standard Fmoc chemistry using a base-labile linkage to the resin. After side-chain deprotection, the C-terminal tyrosine residue was coupled to 4-nitrobenzenediazonium tetrafluoroborate. The azo peptide was then cleaved from the resin and reduced to the *o*-aminophenol by sodium dithionite.

4.3 Fibrin binding peptide design and synthesis

We next turned to a peptide targeting group (GPR) with an affinity for fibrin to create the thrombus targeted VLP.[39] The GPR peptide was derived from the N-terminus of the α-chain of fibrin (Gly-Pro-Arg), and was discovered to bind to a pocket in the C-terminal region of the γ-chain.[40,41] The originally discovered peptide was shown to inhibit thrombin-mediated fibrin clotting by competitively binding to the fibrin polymerization pocket. Subsequent optimization of the original peptide has identified an extended peptide with improved fibrin-binding and increased proteolytic resistance (Gly-Pro-Arg-Pro-Pro).[42] This updated peptide has been used to image pulmonary emboli in swine and more recently has been attached to cross-linked iron oxide particles for the *ex vivo* imaging of intravascular thrombi in mice using fluo-

rescence reflectance imaging.[43,44] The GPR peptide-based targeting strategy was chosen because this peptide has been used successfully for *in vivo* imaging of thrombi and was synthetically tractable.[43–45,38,46,47]

The scheme for the synthesis of the fibrin targeted delivery vehicle consisted of the selective modification of the interior of MS2 with optical imaging agents, followed by the coupling of the targeting peptide to the exterior surface (Figure 4.3). A GPR peptide bearing an *o*-aminophenol for coupling to MS2 was prepared using solid phase peptide synthesis (Figure 4.3).[48] The sequence Gly-Gly-Ser-Lys-Gly-Tyr was added to the peptide to increase the spacing between the binding residues and the capsid, as well as provide a functional handle for modification (Tyr). For peptide samples used in optical imaging experiments, a cysteine residue was also included at the C-terminus to allow modification with the same near-infrared maleimide dye that was introduced in the VLPs (see below). This allowed the binding behavior of the free peptides and the peptide-MS2 conjugates to be compared directly. A non-binding peptide analog was also synthesized as a negative control. It has been reported that a single amino acid change in the GPR peptide (Arg to Ser) drastically decreases its affinity for fibrin.[39] This peptide (GPS) was used as a negative control both as a free peptide and after conjugation to MS2 (GPS-MS2).

4.4 Coupling fibrin binding peptides to the MS2 capsid

To attach the peptide targeting groups to the MS2 capsids, we used a previously described double mutant of the capsid protein, Thr19*p*AF/Asn87Cys (T19*p*AF/N87C) and the $NaIO_4$-mediated oxidative coupling of anilines and *o*-aminophenols (Figure 4.4).[24,36] The aniline functionality was introduced onto the exterior surface of the capsids using the amber stop codon suppression method developed by the Schultz group.[49,50] The *o*-aminophenol functionality was then incorporated into the peptide targeting group by chemically modifying a tyrosine residue after synthesis on the solid phase using standard Fmoc chemistry.[48] Removal of the side-chain protecting groups, azo coupling, and subsequent reduction with $Na_2S_2O_4$ afforded the desired *o*-aminophenol-containing peptides (Figure 4.5). The peptide modification was confirmed by mass spectrometry (MS), and the site of modification was verified through MS/MS analysis (Figure 4.5).

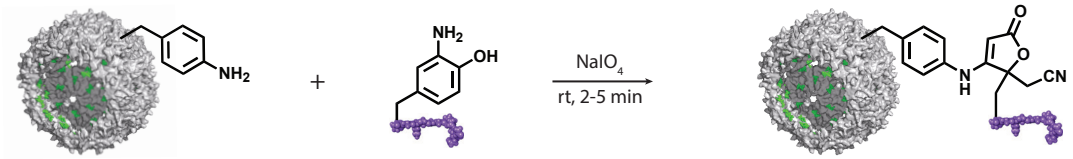


Figure 4.4. The periodate-mediated oxidative coupling reaction takes place between *o*-aminophenol peptides and aniline containing MS2 capsids.

The ability of the modified peptides to participate in the periodate-mediated oxidative coupling was confirmed by reacting the *o*-aminophenol-containing peptides with 1 equiv of *p*-toluidine and 10 equiv of $NaIO_4$, followed by MS characterization. The *o*-aminophenol containing peptides were then coupled to aniline-containing MS2 capsids using the periodate-mediated oxidative coupling shown in Figure 4.4. This bioconjugation strategy was chosen because it is fast, chemoselective and has been used successfully on complex biomolecular assemblies.[36] Additionally, this work was started before the discovery of ferricyanide as an alternative oxidant for the coupling. By altering the reaction time from 30 s to 10 min and the number of peptide equivalents from 1 to 40, the

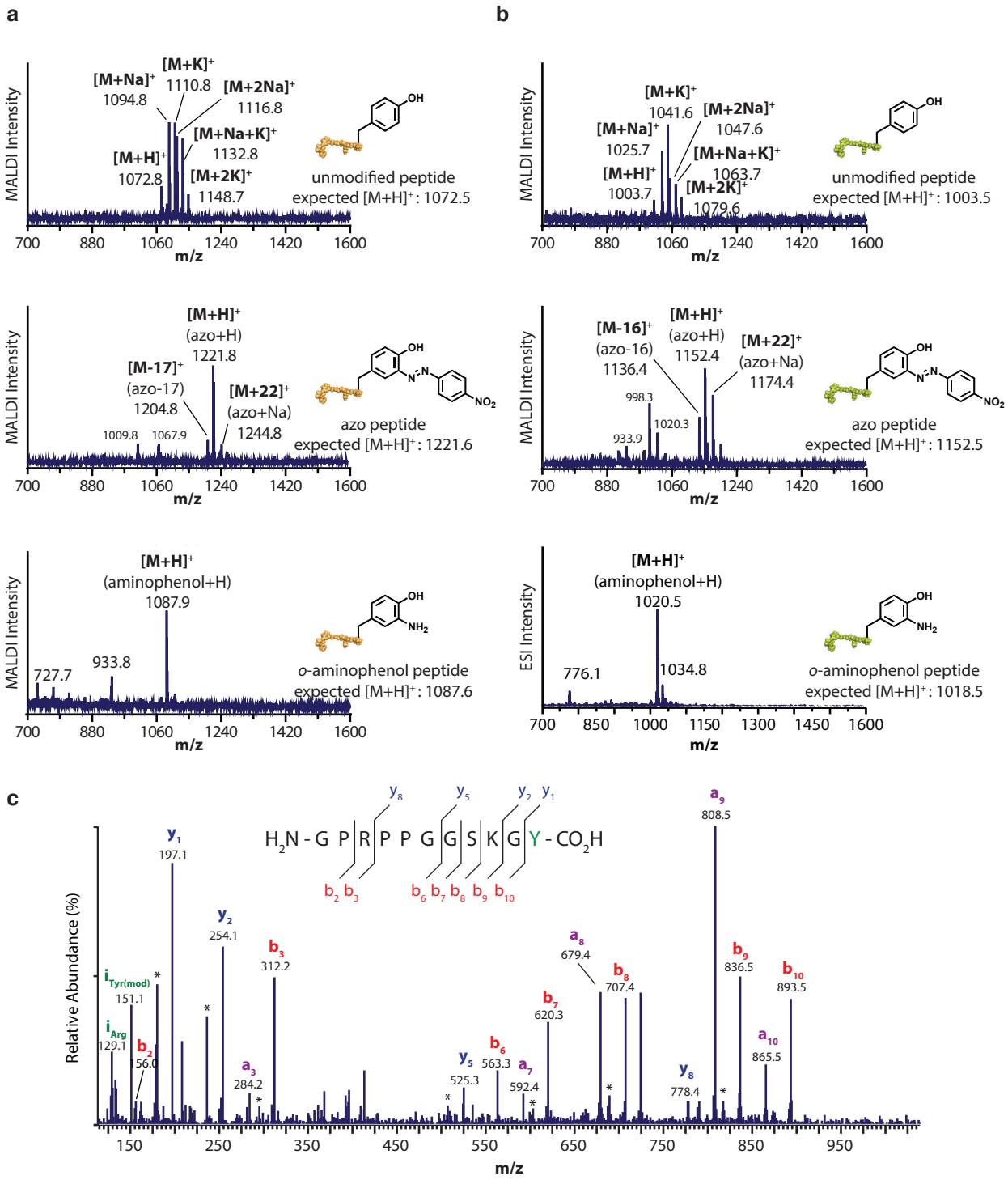


Figure 4.5. Conversion of (a) GPR and (b) GPS peptide C-terminal tyrosine to an *o*-aminophenol. MALDI-TOF MS was used to monitor the modification of (top) the unmodified peptide to (middle) the azo-peptide with 4-nitrobenzenediazonium tetrafluoroborate, and then subsequent sodium dithionite reduction to (bottom) the *o*-aminophenol peptide. (c) MS/MS analysis of *o*-aminophenol modified GPR peptide. Fragments shown in blue are *y* ions (modified fragments), fragments shown in red are *b* ions (unmodified fragments), fragments shown in green are imminium ions (both modified and unmodified fragments), and fragments shown in purple are *a* ions (unmodified fragments). Products of neutral losses of molecules of water or ammonia are denoted by asterisks. The analysis is consistent with the modification of the C-terminal tyrosine.

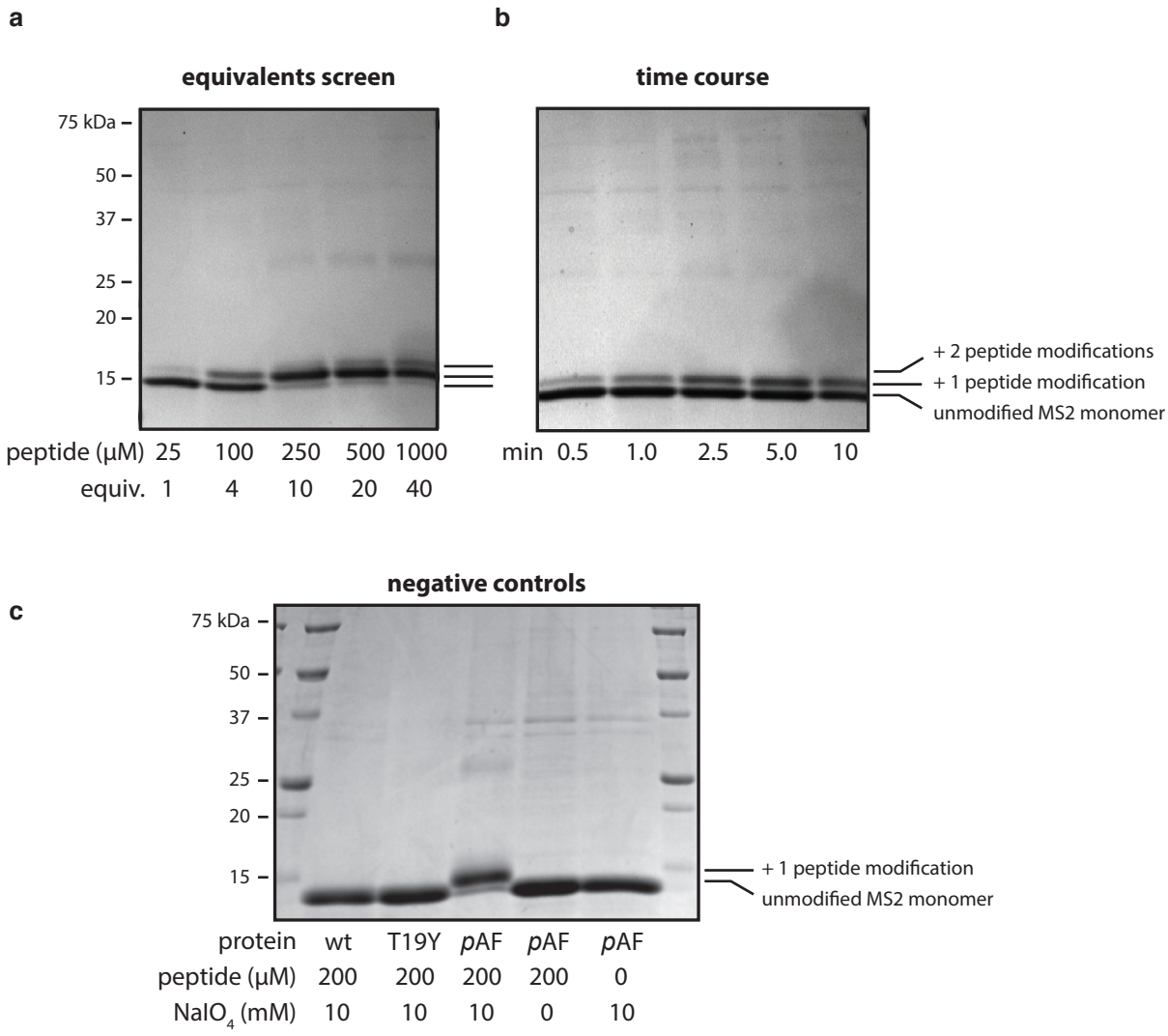


Figure 4.6. Oxidative coupling optimization screen with GPR peptide. (a) Coupling of varying concentrations of GPR peptide with 25 μM T19*p*AF N87C MS2 in the presence of 1 mM sodium periodate in 10 mM phosphate buffer, pH 7.2 for 10 min. Varying the peptide concentration allowed for control over the level of modification. The secondary modification is presumed to be due to a second addition of the *o*-aminophenol to the exterior aniline as this has occasionally been observed with small molecules or to the N-terminus of MS2 as discussed in Chapter 3. (b) A time course experiment with 25 μM T19*p*AF N87C MS2, 100 μM GPR peptide, and 1 mM sodium periodate in 10 mM phosphate buffer, pH 7.2. Each time point was quenched by the addition of 5 μL of loading buffer. The reaction reached maximum conversion after 2.5-5 min. (c) Negative controls confirm the specificity of the reaction for the aniline side-chain. Additionally, both coupling partners and periodate were necessary for modification to be observed.

extent of peptide labeling could be varied from 18 to 150 peptides/capsid (Figure 4.6). Some background reactivity was observed (+2 peptide modifications) when a large excess of peptide was used. The reactions were carried out at pH 7.2, which could account for some of the second addition observed. The periodate-mediated coupling was optimized to install approximately 90 peptides on the surface of each capsid in 5 min, corresponding to ~50% modification of MS2 monomers as determined using optical densitometry of SDS-PAGE gels (Figure 4.7). This reaction used 10 equiv of the *o*-aminophenol peptide and 100 equiv of $NaIO_4$. No coupling was observed for control reactions with capsids lacking the *p*AF groups (both wt-MS2 and T19Y MS2), confirming the che-

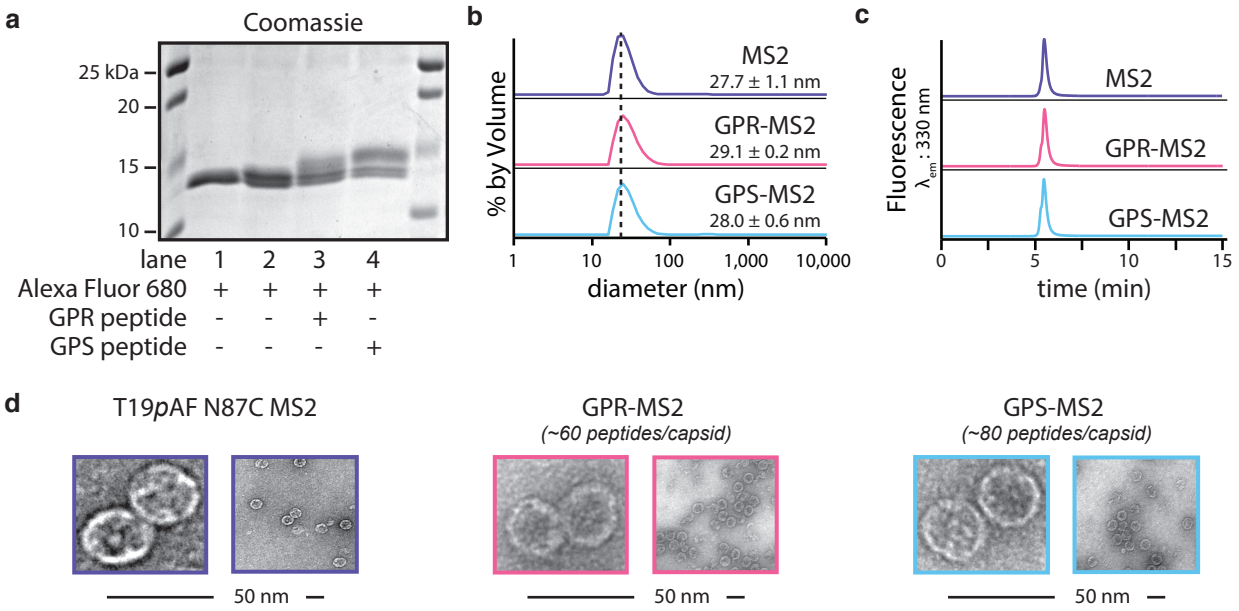


Figure 4.7. Characterization of MS2 conjugates. (a) Alexa Fluor 680 and peptide-MS2 conjugates were analyzed by SDS-PAGE, with visualization of fluorescent and Coomassie-stained bands (shown). Lanes 2-4 were disassembled MS2 monomers labeled with Alexa Fluor 680 (lane 1 is unmodified MS2 monomers). In lane 3, the GPR peptide was added and in lane 4 the GPS peptide was added. The upper bands represent the fraction of the MS2 monomers conjugated to the peptides. (b) dynamic light scattering, (c) size-exclusion chromatography (fluorescence: λ_{ex}=280 nm, λ_{em}=330 nm) and (d) transmission electron micrographs (left close-up and right wide-field) of MS2, GPR-MS2, and GPS-MS2 confirmed that the capsids remained intact after modification.

moselectivity of the method (Figure 4.6). The coupling reactions were terminated by removal of the $NaIO_4$ via gel filtration through Nap 5 Sephadex columns. After modification, the assembly state of the capsids was confirmed by transmission electron microscopy (TEM), dynamic light scattering (DLS), and size-exclusion chromatography (SEC, Figure 4.7). Unmodified MS2 capsids were measured to have a diameter of 27.7 ± 1.1 nm using DLS, while the modified capsids were determined to have diameters of 29.1 ± 0.2 nm and 28.0 ± 0.6 nm for GPR-MS2 and GPS-MS2, respectively. DLS also confirmed that the capsids had a normal size distribution.[24,51]

4.5 Binding to fibrin

The ability of the peptides to bind to fibrin, both before and after conjugation to MS2, was first verified through the use of a clotting inhibition assay (Figure 4.8). Agents capable of binding to fibrin were expected to slow thrombin-mediated clotting because the GPR peptide competitively binds to fibrin at the site necessary for initial fibrin aggregation. Fibrin clotting times in the presence of each of the agents were determined by measuring the extent of light scattering at 350 nm.[52] In agreement with previously published reports, free GPR was found to inhibit thrombin induced clotting at high concentrations (≥ 300 μM), but free GPS did not do so at any of the concentrations assayed (Figure 4.8b).[39,43] Unmodified MS2 capsids and capsids labeled with ~50% peptide (GPR-MS2 and GPS-MS2) were also assayed for their ability to inhibit clotting. At 222 nM MS2 capsid (40 μM in MS2 monomer and ~20 μM in peptide), GPR-MS2 slowed aggregation, while none of the other agents showed an effect (Figure 4.8c). Clotting times more than doubled when treated

with GPR-MS2 (51 ± 29 min), whereas they remained unchanged when treated with free GPR (18 ± 13 min), free GPS (13 ± 6 min), MS2 (16 ± 9 min) or GPS-MS2 (16 ± 6 min). GPR-MS2 slowed fibrin polymerization at a peptide concentration approximately ten-fold less than that observed for the free peptide. We found that the multivalent display of the binding moiety resulted in increased avidity of the targeting peptide. While others have observed varying effects with a multivalent display of similar peptides (GPRP), perhaps both the spacing between the peptides and the size of the multivalent object play a role in the interaction of fibrin/fibrinogen with these peptides in multivalent displays.[53,54] Similar increases in avidity have been observed with different multivalent objects.[25,55–58]

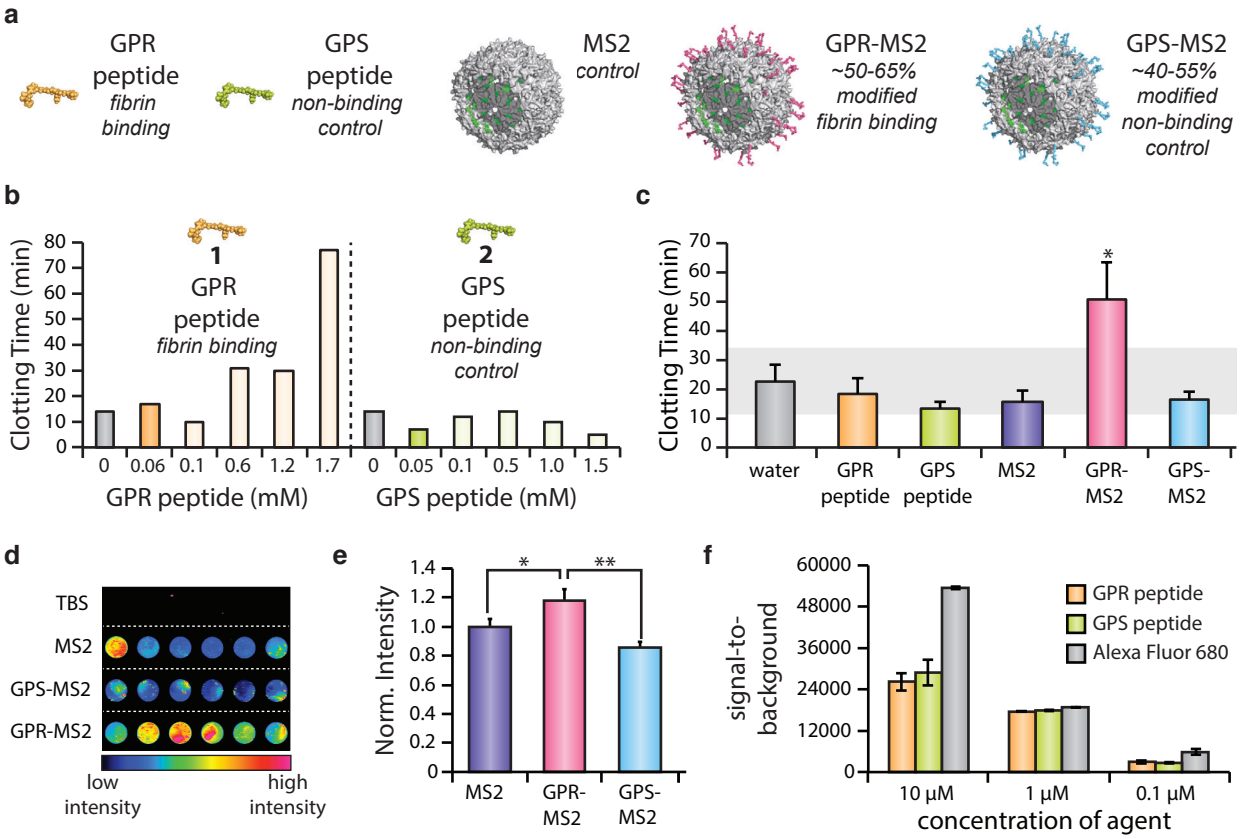


Figure 4.8. Fibrin targeting with GPR-MS2 conjugates. (a) Five different agents were tested for fibrin binding properties. (b) Inhibition of thrombin induced fibrin clotting with varying concentrations of free GPR and GPS peptides. Only at high concentrations (≥ ~ 300 μM) of GPR peptide was an effect observed. The concentrations used in part (c) (10-60 μM) showed no effect for either peptide. (c) Binding to fibrin monomers was confirmed with a clotting inhibition assay. Clotting times are plotted as the average of several trials (n=4-6), shown with the standard error. The grey bar indicates the 95% confidence interval for the water control. At a concentration of 20 μM peptide, (222 nM MS2 capsid), only GPR-MS2 increased the thrombin induced clotting time of fibrin (Welch's t-test, $p < 0.05$). (d) A representative near-infrared fluorescence image of fibrin clots in a 96-well plate indicated the in vitro binding of agents to the clots. (e) Binding was quantified, showing that the differences in relative spot intensity (normalized to the MS2 capsid intensity) between the agents were statistically significant (* $p < 0.05$; ** $p < 0.01$). Normalized intensities are shown with the standard error. (f) Alexa Fluor 680 labeled peptides and Alexa Fluor 680 were assayed for their binding to fibrin clots. At all concentrations tested (10, 1, and 0.1 μM) the two peptides showed no difference in binding. Additionally, the free dye associated non-specifically to the clots as well as, or better than, the peptides modified with the fluorescent dye.

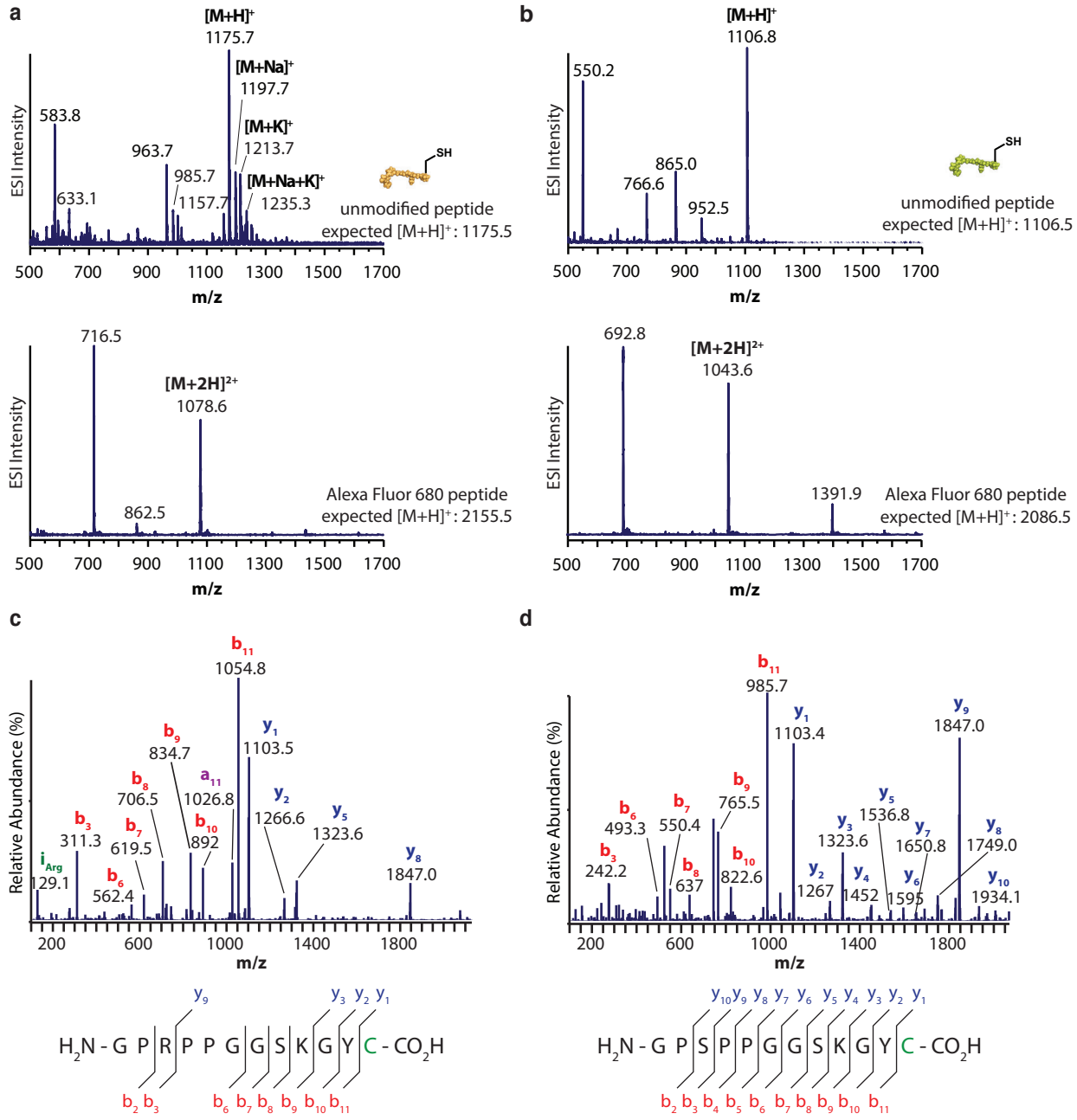


Figure 4.9. Modification of (a) GPR and (b) GPS peptide C-terminal cysteine with Alexa Fluor 680 C_2-maleimide. ESI-MS confirmed the mass of the unmodified peptide (top) and showed the mass of Alexa Fluor 680 modification of the peptide (bottom). MS/MS analysis of Alexa Fluor 680 modified GPR (c) and GPS (d) peptide. Fragments shown in blue are *y* ions (modified fragments), fragments shown in red are *b* ions (unmodified fragments), fragments shown in green are imminium ions (both modified and unmodified fragments), and fragments shown in purple are *a* ions (unmodified fragments). The analysis is consistent with the modification of the C-terminal cysteine.

4.6 Imaging fibrin clots *in vitro*

The potential for MS2 VLPs to image fibrous clots was also tested *in vitro*. The introduction of a reactive cysteine to the interior surface of MS2 allowed for the conjugation of maleimide functionalized imaging agents. Previous work has shown that the capsids can be labeled with fluores-

cent dyes for optical imaging, Gd^{3+} and ^{129}Xe for MRI, and ^{64}Cu for PET using this strategy.[24,59–61] For this work, MS2 capsids were modified with ~60 copies of Alexa Fluor 680 to create near-infrared optical imaging agents. GPR or GPS peptides were then installed on the exterior surface using the oxidative coupling strategy described above.[36] This resulted in ~45-55% labeling for GPR-MS2 (~100 peptides/capsid) and GPS-MS2 (~80 peptides/capsid).

To prepare well-defined clots for use in the binding assays, purified bovine fibrinogen was added to a series of Eppendorf tubes. Thrombin and $CaCl_2$ were added to induce aggregation. After 1 h, the MS2-based agents were added to the fibrin clots at a capsid concentration of 111 nM (20 μM in MS2 monomer, approximately 10 μM in targeting peptide). After 3 h of exposure at room temperature, the clots were rinsed three times with tris-buffered saline (TBS) and then transferred to a 96-well plate. The samples were imaged with an optical scanner using the 700 nm channel (Figure 4.8d). Clots treated with the fibrin-targeted MS2 capsids (GPR-MS2) showed greater signal than clots incubated with control agents (GPS-MS2 and MS2, Figure 4.8e). The signal intensity was normalized to the unmodified MS2 capsids to allow for comparison between experiments. While the MS2 capsids showed some background interaction with the clot, conjugation of the non-binding peptide to the exterior (GPS-MS2) greatly diminished the binding of the capsids. When the fibrin binding peptide was displayed on the exterior of MS2 (GPR-MS2), however, the capsids showed some enhanced binding relative to both the unmodified (MS2) and non-binding (GPS-MS2) capsids. The ability to bind at such low concentrations parallels results using single-chain antibody fragments, antibodies and other peptides.[38,46,47,62,63] However, in this case, significant amounts of imaging cargo can be simultaneously delivered for improved signal-to-noise ratios.

The free peptides were also labeled with Alexa Fluor 680 and tested for comparison. Before use, the modified peptides were HPLC purified and characterized by MS and MS/MS analysis (Figure 4.9). The binding of the GPR and GPS peptides was compared to that of the Alexa Fluor 680 dye at concentrations ranging from 0.1 to 10 μM (Figure 4.8f). At all concentrations tested, the peptides and dye non-specifically bound the fibrin clots. Additionally, the quantified signal-to-background ratio revealed that the free dye associated with the clots as well as, or better than, the free peptides. The binding of the free peptides was expected to be drastically perturbed by the Alexa Fluor 680 dye because the unlabeled peptides and dye have similar molecular weights. This did not affect the MS2-based agents, however, because the optical imaging agent was sequestered inside the capsid where it could not interact with fibrin.

4.7 VCAM1 binding peptide synthesis and conjugation to MS2

While MS2 capsids labeled with the GPR peptide were capable of binding to fibrin, found in advanced atherosclerotic lesions, we also sought to develop VLPs targeted to molecular markers present in the early stages of cardiovascular diseases. The activation of endothelial cells and the over-expression of adhesion molecules occurs early in the development of atherosclerotic lesions (Figure 4.1). A peptide known to bind to the adhesion molecule VCAM1 was used to construct a VLP targeted to early stage atherosclerosis. The peptide (VHPKQHR) was initially identified from an *in vivo* phage display screen in $apoE^{-/-}$ mice.[64,65] The same peptide design was used for the VCAM1 peptide and the fibrin targeted peptide; the solubilizing linker (GGSKG) and reactive handle (Y) were appended to the C-terminus of the peptide to enable attachment to MS2 capsids. The C-terminal tyrosine was converted to an *o*-aminophenol using an azo intermediate (Figure 4.10). To construct the targeted imaging agent, a fluorescent maleimide (Alexa Fluor 488 C_5-maleimide)

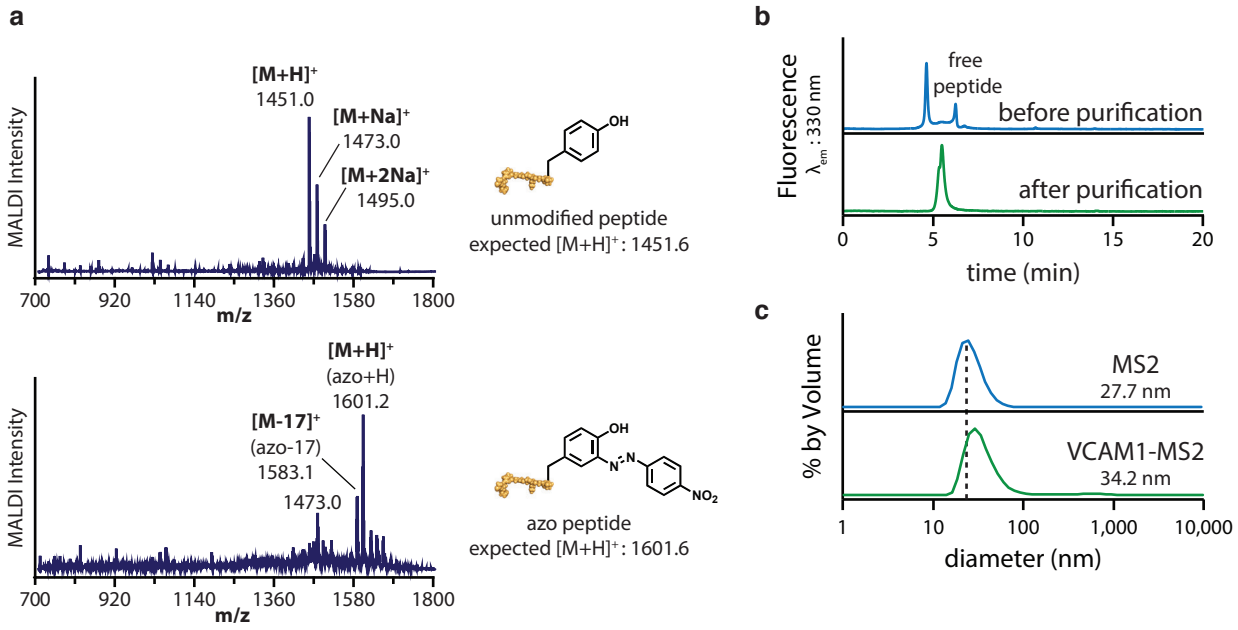


Figure 4.10. Conversion of (a) VCAM-1 peptide C-terminal tyrosine to an *o*-aminophenol. MALDI-TOF MS was used to monitor the modification of (top) the unmodified peptide to (bottom) the azo-peptide with 4-nitrobenzenediazonium tetrafluoroborate. (b) Size exclusion chromatography of MS2 capsids modified with the VCAM-1 peptide before and after purification. (c) Dynamic light scattering measurements of the peptide modified capsids showed assembled capsids with an increased hydrodynamic volume.

was first conjugated to the interior cysteine of T19*p*AF N87C MS2 capsids. The peptide was attached to the fluorescent capsids using the $NaIO_4$-mediated oxidative coupling of *o*-aminophenols and anilines. The purity of the modified samples was evaluated using size exclusion chromatography (Figure 4.10). In addition, the assembly state was verified using size exclusion chromatography and dynamic light scattering. After modification the capsids remained assembled and showed a slight increase in hydrodynamic volume (Figure 4.10).

4.8 Evaluation of binding to VCAM1

The work on VCAM1 binding was carried out in collaboration with an undergraduate, Bryce Jarman. The ability of the modified capsids to bind to VCAM1 was evaluated using flow cytometry (Figure 4.11a). Human umbilical vein endothelial cells (HUVECs) were cultured and VCAM1 expression was stimulated by the addition of TNF-α 8-24 h before the experiment, and non-stimulated cells were used as a negative control. After stimulation, the cells were harvested and divided into 500,000 cell aliquots for each binding experiment. The cells were then incubated with the agents: VCAM1-peptide labeled MS2, VCAM1-peptide, and an α-VCAM1 antibody, at 4 °C for 1 h. The peptide and peptide labeled capsids were directly fluorescently labeled. Binding of the α-VCAM1 antibody was detected by incubation with a fluorescent secondary antibody. After binding, the cells were washed and then analyzed by flow cytometry. The α-VCAM1 antibody showed selective binding to the stimulated cells, but the free peptide and peptide-modified MS2 showed only low levels of non-specific binding to the cells (Figure 4.11b). The inability of the peptide-modified MS2 to selectively bind even at relatively high concentrations (10 μM) prevented us from pursuing these

agents further.

In this work, we describe the synthesis of new VLPs for the targeted imaging of cardiovascular disease. A multivalent scaffold was used to achieve increased binding of a fibrin-binding peptide. An efficient oxidative coupling strategy was used to install many copies of a given functional group in a short period of time. The VCAM1 targeted capsids were not capable of binding to the adhesion molecule. However, the MS2-based delivery vehicle was shown to bind fibrin, a key molecular

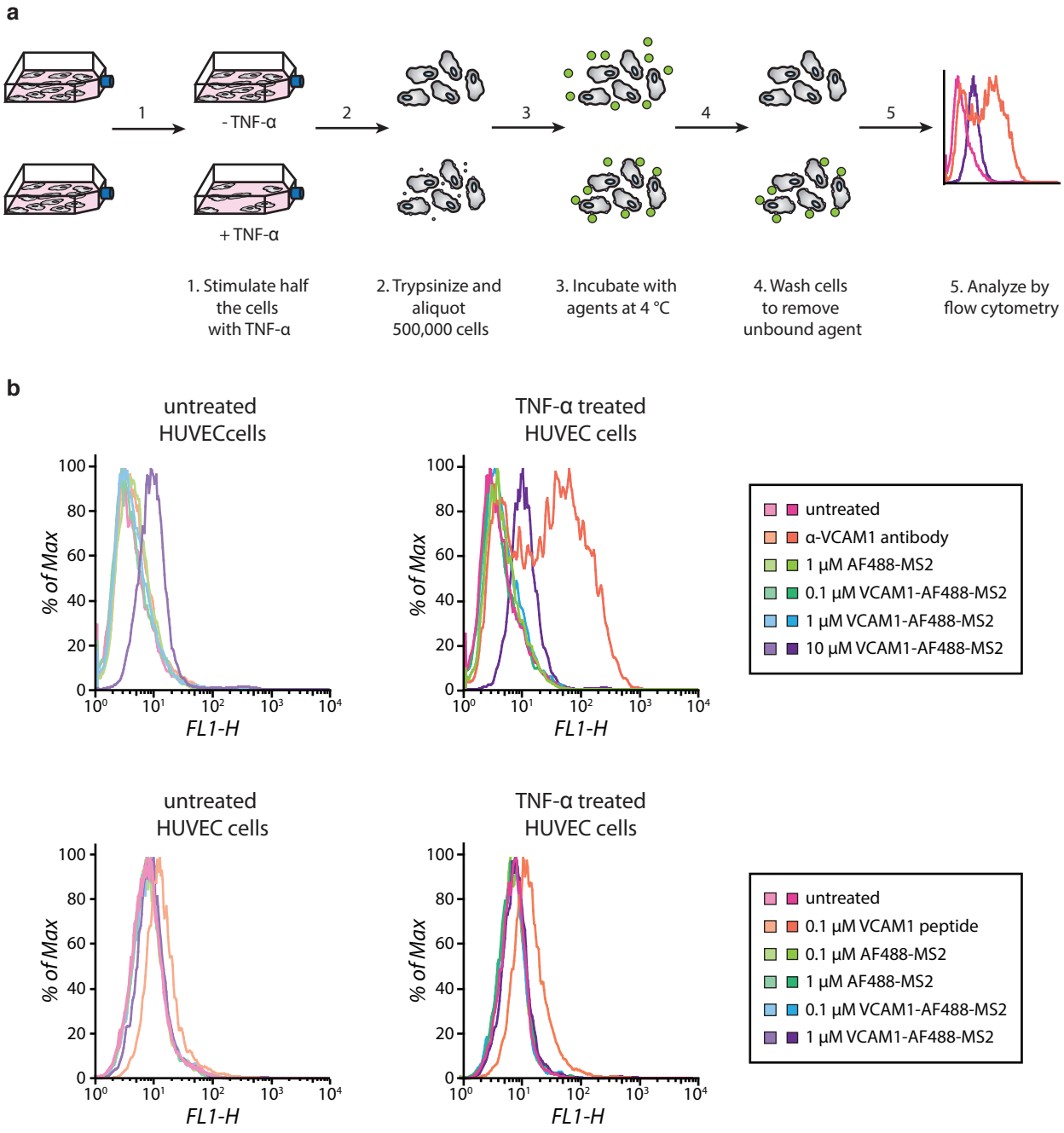


Figure 4.11. (a) Schematic for evaluation of binding to VCAM1 expressing cells. (b) Flow-cytometry of untreated HUVEC cells and HUVEC cells stimulated with TNF-α.

target in blood clots, with greater efficiency than that of free peptides evaluated at the same concentration. The multivalent display of the peptide on the MS2 capsid improved the ability of the GPR peptide to inhibit thrombin-mediated clotting. The attachment of near-infrared dyes allowed for *in vitro* optical detection of fibrin binding.

4.9 Materials and Methods

Unless otherwise noted, the chemicals and solvents used were of analytical grade and were used as received from commercial sources. Water (dd-H_2O) used as reaction solvent and in biological procedures was deionized using a Barnstead NANOpure purification system (ThermoFisher, Waltham, MA).

Instrumentation and sample analysis

Mass Spectrometry. Matrix assisted laser desorption-ionization time-of-flight mass spectrometry (MALDI-TOF MS) was performed on a Voyager-DE system (PerSeptive Biosystems, USA) and data were analyzed using Data Explorer software. Peptide samples were co-crystallized with α-cyano-4-hydroxycinnamic acid in 1:1 acetonitrile (MeCN) to H_2O with 0.1% trifluoroacetic acid (TFA). Electrospray ionization mass spectrometry (ESI-MS) of peptides was performed using an Agilent 1100 series LC pump outfitted with either an Agilent 6224 Time-of-Flight (TOF) LC/MS system or an API 150EX system (Applied Biosystems, USA) equipped with a Turbospray ion source. LC-ESI-MS of proteins and tandem mass spectrometry (MS/MS) of peptides were obtained from the UC Berkeley QB3/Chemistry Mass Spectrometry Facility. Protein bioconjugates were analyzed using an Agilent 1200 series liquid chromatograph (Agilent Technologies, USA) that was connected in-line with an LTQ Orbitrap XL hybrid mass spectrometer equipped with an Ion Max electrospray ionization source (ESI; Thermo Fisher Scientific, Waltham, MA). MS/MS analysis was accomplished with a using a Waters nanoAcquity ultra performance liquid chromatograph (UPLC)/quadrupole time-of-flight (Q-TOF) Premier instrument.

High Performance Liquid Chromatography. HPLC was performed on Agilent 1100 Series HPLC Systems (Agilent Technologies, USA). Sample analysis for all HPLC experiments was achieved with an inline diode array detector (DAD) and inline fluorescence detector (FLD). Analytical and preparative reverse-phase HPLC of peptides and proteins was accomplished using a C18 stationary phase and a MeCN/H_2O with 0.1% TFA gradient mobile phase. Analytical size exclusion chromatography of peptides and proteins was accomplished using a BioSep 4000S stationary phase with isocratic flow of aqueous buffer (10 mM phosphate buffer, pH 7.2).

Gel Analyses. For protein analysis, sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) was carried out on a Mini-Protean apparatus (Bio-Rad, Hercules, CA), using a 10-20% precast linear gradient polyacrylamide gel (Bio-Rad). The sample and electrode buffers were prepared according to Laemmli. All protein electrophoresis samples were heated for 5-10 min at 95 °C in the presence of 1,4-dithiothreitol (DTT) to ensure reduction of disulfide bonds. Gels were run for 75-90 minutes at 120 V to separate the bands. Commercially available markers (Bio-Rad) were applied to at least one lane of each gel for assignment of apparent molecular masses. Visualization of protein bands was accomplished by staining with Coomassie Brilliant Blue R-250 (Bio-Rad).

For fluorescent protein conjugates, visualization was accomplished on a Typhoon 9410 variable mode imager (Amersham Biosciences) courtesy of Prof. Carolyn Bertozzi prior to gel staining. Coomassie stained protein gels were imaged on an EpiChem3 Darkroom system (UVP, USA).

Dynamic Light Scattering. DLS measurements were obtained using a Malvern Instruments Zetasizer Nano ZS. Data plots were calculated from an average of three measurements, each of which consisted of 12 runs of 45 s each. Measurement data are presented as a volume plot, which weights larger dimensions by a factor of 10^3 more than smaller dimensions. Samples were measured in 10 mM phosphate buffer, pH 7.2.

Transmission Electron Microscopy. TEM images were obtained at the UC Berkeley Electron Microscope Lab (www.em-lab.berkeley.edu) using a FEI Tecnai 12 transmission electron microscope with 120 kV accelerating voltage. Protein samples were prepared for TEM analysis by pipetting 5 µL of the samples onto Formvar-coated copper mesh grids (400 mesh, Ted Pella, Redding, CA), after 3 min of equilibration the samples were then wicked with filter paper. The samples were then rinsed with dd-H_2O. Subsequently, the grids were exposed to 8 µL of a 1% (w/v) aqueous solution of uranyl acetate for 90 s as a negative stain. After excess stain was removed, the grids were allowed to dry in air.

Peptide synthesis and modification

General procedure for solid-phase peptide synthesis. Peptides were synthesized using standard Fmoc-based chemistry on Tentagel S-OH resin (Advanced ChemTech, Louisville, KY). The side chain protecting groups used were: Arg(Pbf), Cys(Trt), Lys(Boc), Ser(tBu), Tyr(tBu). The C-terminal amino acid (10 equiv) was preactivated at 0 °C with 5 equivalents of diisopropylcarbodiimide (DIC) and then coupled to the resin with 0.1 equivalents of *N*,*N*-dimethylaminopyridine (DMAP) as a catalyst. Deprotection of the Fmoc groups was performed with a 20 min incubation in a 20% v/v piperidine in dimethylformamide (DMF) solution. Coupling reactions were carried out using 20 equivalents of amino acid with 10 equivalents of 2-(6-chloro-1-*H*-benzotriazole-1-yl)-1,1,3,3-tetramethylaminium hexafluorophosphate (HCTU)[48] and 20 equivalents of *N*,*N*-diisopropylethylamine (DIPEA) in DMF for 20 min. Side-chain deprotection was accomplished using a 1-2 h incubation with either a 95:2.5:2.5 ratio of TFA to H_2O to triisopropylsilane (TIPS) (peptides without cysteine) or a 90:5:2.5:2.5 ratio of TFA to phenol to H_2O to TIPS (peptides containing cysteine). Peptides were cleaved from the resin by a 30-45 min incubation with a 100 mM sodium hydroxide solution. The resulting basic solution was neutralized with 100 mM phosphate buffer, pH 6.5.

General procedure for azo coupling to peptides. To a portion of resin-bound peptide (appx. 20 mg) in 900 µL of 100 mM phosphate buffer, pH 9.0 was added 100 µL of a saturated solution of 4-nitrobenzenediazonium tetrafluoroborate (300 mM, Sigma-Aldrich) in MeCN at 4 °C. After rotation for 30 min at 4 °C, the resin was rinsed with MeCN and DMF until no color remained in solution. The peptides were then cleaved from the resin, and the modification was confirmed by MALDI-TOF MS (Figure 4.5 and 4.10).

General procedure for dithionite reduction of azo peptides. To a solution of azo-modified peptide (appx. 1 mM) in 100 mM phosphate buffer, pH 6.5 was added an equal volume of a freshly

prepared solution of sodium dithionite (115 mM) in 100 mM phosphate buffer, pH 7.2. After 10 min, the solution was applied to a C18 Sep-Pak pre-conditioned with methanol then 0.1% aqueous TFA. After washing with aqueous 0.1% TFA and 5% MeCN in aqueous 0.1% TFA, the peptide was eluted with MeCN. After removing the MeCN under reduced pressure the solid peptide was dissolved in 100 µL of 10 mM phosphate buffer, pH 6.5. The *o*-aminophenol peptide was characterized by MALDI-TOF or LC-ESI MS (Figure 4.5) and stored at 4 °C. The site of modification was confirmed for the GPR peptide by MS/MS analysis (Figure 4.5).

General procedure for maleimide modification of peptides. To a 1.75 mM solution of cleaved peptide in 10 mM phosphate buffer, pH 8.0 was added 1.1 equivalents of tris(2-carboxyethyl)phosphine hydrochloride (0.5 M solution, pH 7.0). After 15 min, 0.6 equivalents of maleimide (Alexa Fluor 680 C_2-maleimide, 20 mM in DMSO) was added and the solution was briefly vortexed. The peptides were incubated in the dark at room temperature overnight. The peptides were then purified by HPLC using a C18 Gemini column (5 micron, 250 x 7.8 mm, Phenomenex, Torrance, CA). The purified peptides were characterized by LC-ESI-MS and MS/MS (Figure 4.9).

Protein expression, purification and modification

Cloning, expression and purification of fibrin binding scFv. The following primers were used to synthesize the gene for the scFv. The gene was stitched together using 10 nM of each primer except the F0 and F798 primers, which were used at a concentration of 400 nM.

R0 TGGacatgacgcggacg
F0 cgtccgcgtcatgtCCATGGCAAGTACAACTGCAAC
R17 CCAGATCACCACCGGATTGTTGCAGTTGTACTTGCCA
F36 AATCCGGTGGTGATCTGGTTAAGCCAGGTGGCTCTC
R54 CGGCGCAAGACAGTTTCAGAGAGCCACCTGGCTTAA
F72 TGAAACTGTCTTGCGCCGCGTCTGGTTTCTCTTTCTCTTC
R90 CACCCAGCTCATACCGTAAGAAGAGAAAGAGAAACCAGACG
F112 TTACGGTATGAGCTGGGTGCGTCAGACCCCGGATAAA
R131 CCACCCATTCCAGGCGTTTATCCGGGGTCTGACG
F149 CGCCTGGAATGGGTGGCGTCTATCTCTTCCGGTGG
R165 GGATAGTACGTGTGGCGGCCACCGGAAGAGATAGACG
F184 CCGCCACACGTACTATCCGGACAGCGTGAAGGGC
R202 ACGGCTGATCGTGAAACGGCCCTTCACGCTGTCC
F218 CGTTTCACGATCAGCCGTGACAACGCTAAAAACACCCT
R236 AGAGAAGACATCTGCAGGTACAGGGTGTTTTTAGCGTTGTC
F256 GTACCTGCAGATGTCTTCTCTGAAATCCGAAGACACTGCG
R277 GGCGTGCACAAAAGTACATCGCAGTGTCTTCGGATTTC
F296 ATGTACTTTTGTGCACGCCAGGAAGGTGATTACGATGATTGG
R315 CGGTGGTACCCTGACCCCAATCATCGTAATCACCTTCCT
F338 GGTCAGGGTACCACCGTGACCGTGTCCTCCGG
R354 GCCAGAACCACCGCCACCGGAGGACACGGTCA
F370 TGGCGGTGGTTCTGGCGGTGGCGGCTCTGGTG
R386 TCGATATCGCTACCACCACCACCAGAGCCGCCACC

F402 GTGGTGGTAGCGATATCGAACTGACCCAGTCTCCTCT
R421 CCAATAATCACGGACAGGGTCAGAGGAGACTGGGTCAGT
F439 GACCCTGTCCGTGATTATTGGTCAGCCTGCCTCCATTT
R460 GAGACTGGGAGGATTTGCAAGAAATGGAGGCAGGCTGA
F477 CTTGCAAATCCTCCCAGTCTCTGCTGTATTCTGACGGTACTAC
R498 CAGCAGCCAGTTCAGGTAGGTAGTACCGTCAGAATACAGCA
F520 CTACCTGAACTGGCTGCTGCAGCGTCCGGGTCAGTC
R539 AGATCAGACGCTTCGGGGACTGACCCGGACGCTG
F556 CCCGAAGCGTCTGATCTATCTGGTATCCAAAGTTGATTCTG
R573 TGAAGCGATCTGGCACACCAGAATCAACTTTGGATACCAGAT
F597 GTGTGCCAGATCGCTTCACCGGTTCTGGTAGCGGC
R615 CTAATTTTCAGGGTGAAGTCAGTGCCGCTACCAGAACCGG
F632 ACTGACTTCACCCTGAAAATTAGCCGCGTGGAGGCGGAA
R655 CAGTAGTAAACGCCCAGGTCTTCCGCCTCCACGCGG
F671 GACCTGGGCGTTTACTACTGCTGGCAGGGCACGCAT
R691 GGAACCGAAAGTAAACGGGAAATGCGTGCCCTGCCAG
F707 TTCCCGTTTACTTTCGGTTCCGGCACCAAACTGGAACTG
R728 CAGCCGCTGCACGTTTCAGTTCCAGTTTGGTGCC
F746 AAACGTGCAGCGGCTGGTTCTTGTATCGAAGGCCGC
R762 CGAGGCCGCCAGAGGAGCGGCCTTCGATACAAGAAC
F782 TCCTCTGGCGGCCTCGAGgatcccactgggct
F798 agcccagtgggatcCT

The isolated gene insert was cloned into the pET22b(+) vector using NcoI and XhoI restriction sites. The scFv was transformed into T7 Express *lysY/I^q E. coli* and the expression was optimized in 5 mL cultures. The cells were grown in 1 L LB with 100 μg/L ampicillin with 100 μM IPTG for 18 h at 37 °C. The periplasmic proteins were extracted and purified following the Qiagen protocol. Briefly, the cells were collected by centrifugation and resuspended in 30 mM Tris-Cl, 20% sucrose, pH 8.0 at appx. 80 mL per gram of wet *E. coli*. The cells were kept on ice and EDTA (as a 500 mM solution) was added to a final concentration of 1 mM. The cells were gently agitated for 5-10 min. The cell suspension was centrifuged at 8000 g for 20 min and the pellet was resuspended in an equal volume of ice-cold 5 mM $MgSO_4$. Following 10 min of gentle stirring the suspension was centrifuged at 8000 g for 20 min again. The supernatant was collected and buffer exchanged into lysis buffer (50 mM NaH_2PO_4, 300 mM NaCl, 10 mM imidazole, pH 8.0). The periplasmic extract was applied to pre-equilibrated Ni-NTA resin (Qiagen). The resin was washed with wash buffer (50 mM NaH_2PO_4, 300 mM NaCl, 20 mM imidazole, pH 8.0) and then protein was eluted with elution buffer (50 mM NaH_2PO_4, 300 mM NaCl, 250 mM imidazole, pH 8.0). However the scFv eluted in the flow-through or the wash steps.

Expression and purification of T19*p*AF N87C MS2. The protein was expressed and purified following a previously published protocol.[34] After purification, approximately 10 mg of protein were obtained per L of culture.

General procedure for oxidative coupling of *o*-aminophenol peptides to *p*AF MS2. To a solution of T19*p*AF N87C MS2 (10 μM) in 10 mM phosphate buffer, pH 6.5 was added 5-10 equivalents of

o-aminophenol peptide (50-100 μM). The solution was briefly vortexed and then 10 equivalents (relative to the *o*-aminophenol) of sodium periodate was added. After 5 min, the reaction mixture was purified on a Nap 5 Sephadex size exclusion column (GE Healthcare) according to the manufacturer's instructions and eluted in 10 mM phosphate buffer, pH 7.2. Any remaining periodate and free peptide were removed using a 0.5 mL centrifugal filter with a molecular weight cut off (MWCO) of 100 kDa (Millipore). The samples were concentrated to 50 μL and then diluted 10-fold with 10 mM phosphate buffer, pH 7.2. This process was repeated 5-10 times depending on the concentration of peptide used. Modification was monitored by SDS-PAGE and quantified using optical densitometry (Figure 4.6).

General procedure for dual modification of T19*p*AF N87C MS2. Dual modification always started with maleimide modification of the interior cysteine. A solution of T19*p*AF N87C MS2 (100 μM) in 10 mM phosphate buffer, pH 7.2 was incubated with 1 equivalent of Alexa Fluor 680 C_2-maleimide (38 mM in DMSO) for 2 h. Unreacted dye was removed with a Nap 10 Sephadex size exclusion column (GE Healthcare) and by repeated centrifugal filtration against a 100 kDa MWCO membrane. Exterior modification with an *o*-aminophenol peptide was performed as described above. Modification was monitored by SDS-PAGE (Figure 4.6).

Other procedures

Inhibition of fibrin clot formation. Bovine fibrinogen (≥75% clottable, Sigma Aldrich, USA) was dissolved in 0.9% saline solution at 37 °C at a concentration of 12 mg/mL. Bovine thrombin (100 U/mL) was dissolved in phosphate buffered saline (PBS) and added into wells of a 96-well polystyrene half area plate (Costar, Corning, NY) to a final concentration of 0.25 U/well. Assayed agents (H_2O (control), GPR peptide, GPS peptide, MS2, GPR-MS2, and GPS-MS2) were added to the wells (72.5 μL/well, n=4-6) to a final concentration of 20 μM peptide (40 μM in MS2 monomer for MS2, GPR-MS2, and GPS-MS2 for an effective peptide concentration of 20 μM and capsid concentration of 222 nM). Immediately preceding absorbance measurements, 25 μL of the fibrinogen solution was added to each well (0.3 μg/well). The absorbance at 350 nm was monitored once a minute for 75 min using a SpectraMax M3 microplate reader (Molecular Devices, Sunnyvale, CA), courtesy of Prof. Carolyn Bertozzi. Additional concentrations of peptide (10 μM to 1.7 mM) were evaluated for their ability to inhibit fibrin clot formation in a similar manner (Figure 4.8). The clotting time was determined to be the time at which the absorbance first approached the horizontal asymptote. Statistical significance was evaluated by a one-tailed Student's t-test with unequal variance.

Binding to fibrin clots. Bovine fibrinogen (≥75% clottable, Sigma Aldrich, USA) was dissolved in 50 mM Tris, pH 7.4, 150 mM sodium chloride (TBS) and dialyzed against TBS with 5 mM sodium citrate (TBS-citrate). The resulting fibrinogen solution was adjusted to 4 mg/mL by measuring the absorbance at 280 nm (a 1 mg/mL fibrinogen solution has an absorbance of 1.512 OD units).[66] The fibrinogen solution (50 μL) was aliquoted into Eppendorf tubes. Subsequently, calcium chloride (10 μL of a 70 mM solution) and 0.4 U of bovine thrombin (10 U/mL in TBS) were added to each tube to initiate clotting. After incubation at room temperature for 1 h, 100 μL of Alexa Fluor 680 labeled agent (MS2, GPR-MS2 or GPS-MS2) was added to each clot (n=3, with 4-6 replicates per experiment) for a final concentration of 20 μM MS2 monomer (111 nM MS2 capsid and approximately 10 μM peptide). The agents were allowed to bind for 3 h at room temperature and then the

clots were washed 3 times with 200 μL TBS. The clots were transferred to a 96-well polystyrene plate (Costar, Corning, NY) and centrifuged at 4000 rpm for 10 min. The plate was imaged using an Odyssey CLx infrared imaging system (Li-Cor, Lincoln, NE) courtesy of Prof. Jay Groves. The resulting images were quantified using ImageJ 1.42q software (NIH) using the Microarray Profile plugin to select regions-of-interest. The image was pseudo-colored in ImageJ. Additionally, Alexa-Fluor 680 labeled peptide and free AlexaFluor 680 C_2-maleimide (10, 1, and 0.1 μM) were tested for their binding to fibrin clots (Figure 4.8). Statistical significance was evaluated by a one-tailed Student's t-test with unequal variance.

Culturing of human umbilical vein endothelial cells (HUVECs). The cells were purchased from ATCC and grown in Medium 200 supplemented with low serum growth supplement and 1% penicillin and streptomycin. To passage the cells, the cells were first rinsed with phosphate buffered saline (PBS) and then treated with trypsin and incubated at room temperature for approximately 3-5 min. The cells were collected, counted and plated at a density of 2,500 cells/cm^2 in culture flasks pre-treated with gelatin. The cells were grown at 37 °C with an atmosphere of 95% air and 5% CO_2.

Flow cytometry. A FACSCalibur flow cytometer (BD Biosciences, USA) equipped with a 488 laser was used for all flow cytometry measurements, usage courtesy of Prof. Carolyn Bertozzi. Following the harvesting and counting of cells as above, cells were resuspended in flow cytometry buffer (1% FBS in DPBS). The cells were aliquotted into Eppendorf tubes at 100 μL (500,000 cells) per tube and kept on ice. 100 μL of agent in flow cytometry buffer (FCB) was added and incubated at 4 °C. After 1 h, each sample was diluted to 1 mL with FCB, and the tubes were centrifuged at 2,000 rpm for 5 min. The supernatant was removed, and the cells were washed two times with 1 mL of FCB. The cells were resuspended in 200 μL of FCB. Data were analyzed using FlowJo analysis software (Tree Star Inc.). Gating was performed by applying the autogating tool in FlowJo onto the major population of cells in the FSC x SSC (forward versus side scatter plots) of untreated samples; treated samples were subject to the same gating as the untreated population.

4.10 References

1. Roger, V. L.; Go, A. S.; Lloyd-Jones, D. M.; Benjamin, E. J.; Berry, J. D.; Borden, W. B.; Bravata, D. M.; Dai, S.; Ford, E. S.; Fox, C. S.; Fullerton, H. J.; Gillespie, C.; Hailpern, S. M.; Heit, J. A.; Howard, V. J.; Kissela, B. M.; Kittner, S. J.; Lackland, D. T.; Lichtman, J. H.; Lisabeth, L. D.; Makuc, D. M.; Marcus, G. M.; Marelli, A.; Matchar, D. B.; Moy, C. S.; Mozaffarian, D.; Mussolino, M. E.; Nichol, G.; Paynter, N. P.; Soliman, E. Z.; Sorlie, P. D.; Sotoodehnia, N.; Turan, T. N.; Virani, S. S.; Wong, N. D.; Woo, D.; Turner, M. B. *Circulation* **2012**, *125*, e2–e220.
2. Choudhury, R. P.; Fuster, V.; Fayad, Z. A. *Nat. Rev. Drug Discov.* **2004**, *3*, 913–925.
3. Sanz, J.; Fayad, Z. A. *Nature* **2008**, *451*, 953–957.
4. Choudhury, R. P.; Fisher, E. A. *Arterioscler. Thromb. Vasc. Biol.* **2009**, *29*, 983–991.
5. McCarthy, J. R.; Jaffer, F. A. *Nanomedicine* **2011**, *6*, 1291–1293.
6. McCarthy, J. R. *Curr. Cardiovasc. Imaging Reports* **2010**, *3*, 42–49.
7. Chacko, A.-M.; Hood, E. D.; Zern, B. J.; Muzykantov, V. R. *Curr. Opin. Colloid Interface Sci.* **2011**, *16*, 215–227.
8. Li, D.; Patel, A. R.; Klibanov, A. L.; Kramer, C. M.; Ruiz, M.; Kang, B.-Y.; Mehta, J. L.; Beller, G. A.; Glover, D. K.; Meyer, C. H. *Circ. Cardiovasc. Imaging* **2010**, *3*, 464–472.

9. Kamaly, N.; Miller, A. D. *Int. J. Mol. Sci*. **2010**, *11*, 1759–1776.
10. Parrott, M. C.; Benhabbour, S. R.; Saab, C.; Lemon, J. A.; Parker, S.; Valliant, J. F.; Adronov, A. *J. Am. Chem. Soc.* **2009**, *131*, 2906–2916.
11. Lee, C. C.; MacKay, J. A.; Fréchet, J. M. J.; Szoka, F. C. *Nat. Biotechnol*. **2005**, *23*, 1517–1526.
12. McCarthy, J. R.; Weissleder, R. *Adv. Drug Deliv. Rev*. **2008**, *60*, 1241–1251.
13. Jayagopal, A.; Russ, P. K.; Haselton, F. R. *Bioconjugate Chem*. **2007**, *18*, 1424–1433.
14. Yildiz, I.; Shukla, S.; Steinmetz, N. F. *Curr. Opin. Biotechnol*. **2011**, *22*, 901–908.
15. Manchester, M.; Singh, P. *Adv. Drug Deliv. Rev*. **2006**, *58*, 1505–1522.
16. Lee, L. A.; Wang, Q. *Nanomedicine* **2006**, *2*, 137–149.
17. Li, F.; Zhang, Z.-P.; Peng, J.; Cui, Z.-Q.; Pang, D.-W.; Li, K.; Wei, H.-P.; Zhou, Y.-F.; Wen, J.-K.; Zhang, X.-E. *Small* **2009**, *5*, 718–726.
18. Uchida, M.; Kosuge, H.; Terashima, M.; Willits, D. A.; Liepold, L. O.; Young, M. J.; McConnell, M. V.; Douglas, T. *ACS Nano* **2011**, *5*, 2493–2502.
19. Kitagawa, T.; Kosuge, H.; Uchida, M.; Dua, M.; Iida, Y.; Dalman, R.; Douglas, T.; McConnell, M. *Mol. Imaging Biol*. **2012**, *14*, 315–324.
20. Terashima, M.; Uchida, M.; Kosuge, H.; Tsao, P. S.; Young, M. J.; Conolly, S. M.; Douglas, T.; McConnell, M. V. *Biomaterials* **2011**, *32*, 1430–1437.
21. Uchida, M.; Flenniken, M. L.; Allen, M.; Willits, D. A.; Crowley, B. E.; Brumfield, S.; Willis, A. F.; Jackiw, L.; Jutila, M.; Young, M. J.; Douglas, T. *J. Am. Chem. Soc.* **2006**, *128*, 16626–16633.
22. Kickhoefer, V. A.; Garcia, Y.; Mikyas, Y.; Johansson, E.; Zhou, J. C.; Raval-Fernandes, S.; Minoofar, P.; Zink, J. I.; Dunn, B.; Stewart, P. L.; Rome, L. H. *Proc. Natl. Acad. Sci. U.S.A*. **2005**, *102*, 4348–4352.
23. Pokorski, J. K.; Hovlid, M. L.; Finn, M. G. *ChemBioChem* **2011**, *12*, 2441–2447.
24. Tong, G. J.; Hsiao, S. C.; Carrico, Z. M.; Francis, M. B. *J. Am. Chem. Soc*. **2009**, *131*, 11174–11178.
25. Ashley, C. E.; Carnes, E. C.; Phillips, G. K.; Durfee, P. N.; Buley, M. D.; Lino, C. A.; Padilla, D. P.; Phillips, B.; Carter, M. B.; Willman, C. L.; Brinker, C. J.; Caldeira, J. do C.; Chackerian, B.; Wharton, W.; Peabody, D. S. *ACS Nano* **2011**, *5*, 5729–5745.
26. Plummer, E. M.; Thomas, D.; Destito, G.; Shriver, L. P.; Manchester, M. *Nanomed*. **2012**, *7*, 877–888.
27. Ellington, A. D.; Szostak, J. W. *Nature* **1990**, *346*, 818–822.
28. Tuerk, C.; Gold, L. *Science* **1990**, *249*, 505–510.
29. Kodadek, T.; Bachhawat-Sikder, K. *Mol. Biosyst*. **2006**, *2*, 25–35.
30. Udugamasooriya, D. G.; Dineen, S. P.; Brekken, R. A.; Kodadek, T. *J. Am. Chem. Soc*. **2008**, *130*, 5744–5752.
31. Binz, H. K.; Stumpp, M. T.; Forrer, P.; Amstutz, P.; Plückthun, A. *J. Mol. Biol*. **2003**, *332*, 489–503.
32. Binz, H. K.; Amstutz, P.; Kohl, A.; Stumpp, M. T.; Briand, C.; Forrer, P.; Grütter, M. G.; Plückthun, A. *Nat. Biotechnol*. **2004**, *22*, 575–582.
33. Hooker, J. M.; Kovacs, E. W.; Francis, M. B. *J. Am. Chem. Soc*. **2004**, *126*, 3718–3719.
34. Carrico, Z. M.; Romanini, D. W.; Mehl, R. A.; Francis, M. B. *Chem. Commun.* **2008**, 1205–1207.
35. Mastico, R. A.; Talbot, S. J.; Stockley, P. G. *J. Gen. Virol.* **1993**, *74* (Pt 4), 541–548.
36. Behrens, C. R.; Hooker, J. M.; Obermeyer, A. C.; Romanini, D. W.; Katz, E. M.; Francis, M. B. *J. Am. Chem. Soc.* **2011**, *133*, 16398–16401.

37. Wei, B.; Wei, Y.; Zhang, K.; Wang, J.; Xu, R.; Zhan, S.; Lin, G.; Wang, W.; Liu, M.; Wang, L.; Zhang, R.; Li, *J. Biomed. Pharmacother.* **2009**, *63*, 313–318.
38. Peter, K.; Graeber, J.; Kipriyanov, S.; Zewe-Welschof, M.; Runge, M. S.; Kübler, W.; Little, M.; Bode, C. *Circulation* **2000**, *101*, 1158–1164.
39. Laudano, A. P.; Doolittle, R. F. *Proc. Natl. Acad. Sci. U.S.A.* **1978**, *75*, 3085–3089.
40. Pratt, K. P.; Côté, H. C.; Chung, D. W.; Stenkamp, R. E.; Davie, E. W. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 7176–7181.
41. Yang, Z.; Mochalkin, I.; Doolittle, R. F. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 14156–14161.
42. Thakur, M. L.; Pallela, V. R.; Consigny, P. M.; Rao, P. S.; Vessileva-Belnikolovska, D.; Shi, R. *J. Nucl. Med.* **2000**, *41*, 161–168.
43. Aruva, M. R.; Daviau, J.; Sharma, S. S.; Thakur, M. L. *J. Nucl. Med.* **2006**, *47*, 155–162.
44. McCarthy, J. R.; Patel, P.; Botnaru, I.; Haghayeghi, P.; Weissleder, R.; Jaffer, F. A. *Bioconjugate Chem.* **2009**, *20*, 1251–1255.
45. Yu, X.; Song, S. K.; Chen, J.; Scott, M. J.; Fuhrhop, R. J.; Hall, C. S.; Gaffney, P. J.; Wickline, S. A.; Lanza, G. M. *Magn. Reson. Med.* **2000**, *44*, 867–872.
46. Overoye-Chan, K.; Koerner, S.; Looby, R. J.; Kolodziej, A. F.; Zech, S. G.; Deng, Q.; Chasse, J. M.; McMurry, T. J.; Caravan, P. *J. Am. Chem. Soc.* **2008**, *130*, 6025–6039.
47. Kolodziej, A. F.; Nair, S. A.; Graham, P.; McMurry, T. J.; Ladner, R. C.; Wescott, C.; Sexton, D. J.; Caravan, P. *Bioconjugate Chem.* **2012**, *23*, 548–556.
48. Hood, C. A.; Fuentes, G.; Patel, H.; Page, K.; Menakuru, M.; Park, J. H. *J. Pept. Sci.* **2008**, *14*, 97–101.
49. Wang, L.; Brock, A.; Herberich, B.; Schultz, P. G. *Science* **2001**, *292*, 498–500.
50. Santoro, S. W.; Wang, L.; Herberich, B.; King, D. S.; Schultz, P. G. *Nat. Biotechnol.* **2002**, *20*, 1044–1048.
51. Comellas-Aragonès, M.; Sikkema, F. D.; Delaittre, G.; Terry, A. E.; King, S. M.; Visser, D.; Heenan, R. K.; Nolte, R. J. M.; Cornelissen, J. J. L. M.; Feiters, M. C. *Soft Matter* **2011**, *7*, 11380–11391.
52. Latallo, Z. S.; Fletcher, A. P.; Alkjaersig, N.; Sherry, S. *Am. J. Physiol.* **1962**, *202*, 681–686.
53. Soon, A. S. C.; Lee, C. S.; Barker, T. H. *Biomaterials* **2011**, *32*, 4406–4414.
54. Watson, J. W.; Doolittle, R. F. *Biochemistry* **2011**, *50*, 9923–9927.
55. Nahrendorf, M.; Keliher, E.; Panizzi, P.; Zhang, H.; Hembrador, S.; Figueiredo, J.-L.; Aikawa, E.; Kelly, K.; Libby, P.; Weissleder, R. *JACC Cardiovasc. Imaging* **2009**, *2*, 1213–1222.
56. Sigal, G. B.; Mammen, M.; Dahmann, G.; Whitesides, G. M. *J. Am. Chem. Soc.* **1996**, *118*, 3789–3800.
57. Gestwicki, J. E.; Cairo, C. W.; Strong, L. E.; Oetjen, K. A.; Kiessling, L. L. *J. Am. Chem. Soc.* **2002**, *124*, 14922–14933.
58. Woller, E. K.; Walter, E. D.; Morgan, J. R.; Singel, D. J.; Cloninger, M. J. *J. Am. Chem. Soc.* **2003**, *125*, 8820–8826.
59. Garimella, P. D.; Datta, A.; Romanini, D. W.; Raymond, K. N.; Francis, M. B. *J. Am. Chem. Soc.* **2011**, *133*, 14704–14709.
60. Meldrum, T.; Seim, K. L.; Bajaj, V. S.; Palaniappan, K. K.; Wu, W.; Francis, M. B.; Wemmer, D. E.; Pines, A. *J. Am. Chem. Soc.* **2010**, *132*, 5936–5937.
61. Farkas, M. E.; Aanei, I. L.; Behrens, C. R.; Tong, G. J.; Murphy, S. T.; O'Neil, J. P.; Francis, M. B. *Mol. Pharm.* **2013**, *10*, 69–76.
62. Hui, K. Y.; Haber, E.; Matsueda, G. R. *Science* **1983**, *222*, 1129–1132.

63. Knight, L. C.; Maurer, A. H.; Ammar, I. A.; Shealy, D. J.; Mattis, J. A. *J. Nucl. Med*. **1988**, *29*, 494–502.
64. Kelly, K. A.; Nahrendorf, M.; Yu, A. M.; Reynolds, F.; Weissleder, R. *Mol. Imaging Biol.* **2006**, *8*, 201–207.
65. Nahrendorf, M.; Jaffer, F. A.; Kelly, K. A.; Sosnovik, D. E.; Aikawa, E.; Libby, P.; Weissleder, R. *Circulation* **2006**, *114*, 1504–1511.
66. Johnson, P.; Mihalyi, E. *Biochim. Biophys. Acta BBA - Biophys. Photosynth*. **1965**, *102*, 467–475.

Chapter 5
Oxidative coupling for the detection of protein tyrosine-nitration

Abstract

Tyrosine residues are occassionally post-translationally nitrated. A new method to detect this post-translational modification is outlined in this chapter. The ferricyanide-mediated oxidative coupling reaction described in Chapter 2 was used to detect nitrotyrosine residues. The reaction was demonstrated to be selective for the reduced amino-tyrosine species. Additionally, both reduction of the nitrophenol moeity and oxidative coupling to the resulting aminophenol moiety were successfully demonstrated in cellular lysate.

5.1 Post-translational tyrosine nitration

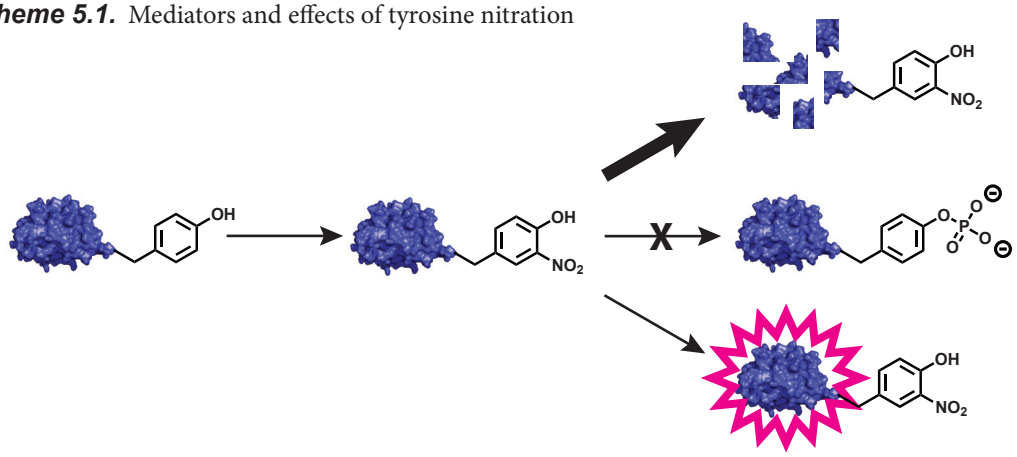
As discussed in the previous chapters, bioconjugation reactions can be used to construct protein-based materials. However, bioconjugation reactions can also be used to detect and monitor biological processes. In this Chapter, application of the ferricyanide-mediated oxidative coupling reaction (described in Chapter 2) to the detection of tyrosine nitration is explored.

The nitration of tyrosine residues is mediated by reactive nitrogen species, such as peroxynitrite anion ($OONO^-$) and nitrogen dioxide ($\cdot NO2$).[1] Nitration of tyrosine lowers the phenolic pK_a by 2-3 units and adds a bulky substituent ortho to the hydroxyl. Both of these effects can alter protein function and conformation. The biological effect of this nitration can be a gain-of-function, loss of activity, sensitivity to proteolytic degradation, or interference with tyrosine phosphorylation (Scheme 5.1).[2] For example, after nitration of a tyrosine in cytochrome *c*, it gains strong peroxidase activity. While tyrosine nitration is known to alter biological properties of proteins, there is debate over whether tyrosine nitration is merely a biological marker of oxidative stress or if the nitration mediates the downstream biological effects.[3] Nitrotyrosine, either as a free amino acid or within a protein, has been detected in many diseases (>50). In particular, there is a strong association between nitrotyrosine levels and coronary artery disease risk.[1] In addition, nitrotyrosine has been observed in association with lung disease, diabetes, and neurodegenerative diseases.[4] Improved methods of detection of tyrosine nitration could lead to a better understanding of the biological effects and disease implications of this post-translational modification.

5.2 Methods for detection of 3-nitro-tyrosine

Most studies of tyrosine nitration have relied on commercially available anti-NO_2-Tyr anti-

Scheme 5.1. Mediators and effects of tyrosine nitration



mediators of nitration	effect on proteins	implications for disease
• oxidative stress	• increased proteolysis	• strong association with incidence of coronary artery disease
• reactive nitrogen species	• decreased phosphorylation	• linked to diabetes
• $\cdot NO/O_2^{\cdot -} \longrightarrow OONO^-$	• enzyme activation	• implicated in age-related diseases

bodies with detection via western blot after 1D or 2D electrophoretic separation (Figure 5.1a).[5] To control for potential non-specific binding of the antibody, parallel experiments are run with a dithionite-medated reduction of the nitro-tyrosine to the corresponding amino-tyrosine. This dithionite-mediated reduction has also been utilized to chemically detect nitrated proteins (Figure 5.1b). The amino-tyrosine residue can be selectively labeled by first capping the native amines of lysines and the N-terminus followed by reduction of nitro-tyrosine residues and a second acylation step.[6] The selectivity of this method relies on the complete modification of native amino groups in the first step. More recently, an alternative chemical detection method was developed that relies on the unique reactivity the *o*-aminophenol functionality of reduced nitro-tyrosine.[7] Reaction of the amino-tyrosine with salicylaldehyde and aqueous $AlCl_3$ results in a fluorescent complex (λ_{ex} = 412 nm, λ_{em} = 520 nm). While the imine formation with salicylaldehyde is not selective for the amine of amino-tyrosine, coordination of Al^{3+}, and thus fluorescent signal, relies on the *o*-hydroxyl found only in the amino-tyrosine group.

The success of these chemical detection methods, coupled with our discovery of the relatively mild conditions for the oxidative coupling of *o*-aminophenols (discussed in Chapter 2), led us to investigate the use of ferricyanide-mediated oxidative coupling of anilines and *o*-aminophenols for the detection of tyrosine nitration.[8] Reduction of the nitro-tyrosine to the corresponding amino-tyrosine (*o*-aminophenol), followed by oxidative coupling with a functionalized aniline probe would

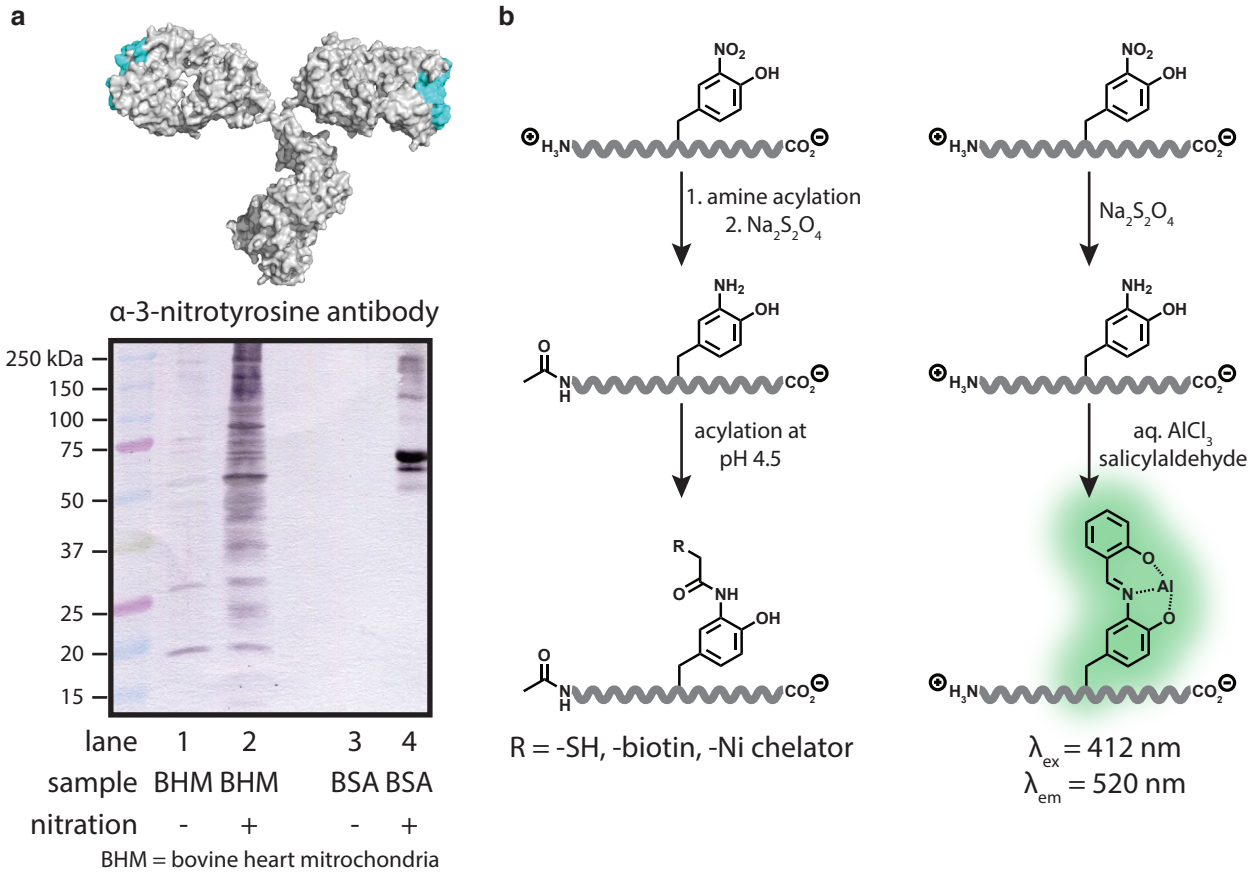


Figure 5.1. (a) Antibody based detection of tyrosine nitration. (b) Chemical methods for the detection of tyrosine nitration.

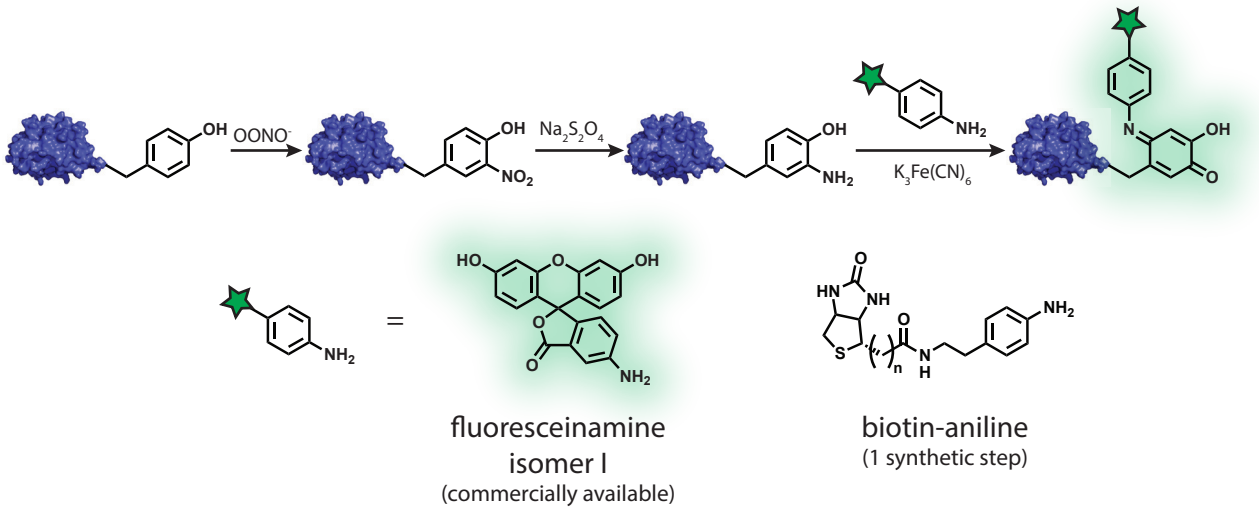


Figure 5.2. Schematic for an oxidative coupling based method for the detection of tyrosine nitration. Fluorescent or biotinylated aniline probes enable detection by fluorescence or western blot.

allow for the rapid, selective detection of 3-nitro-tyrosine (Figure 5.2). To avoid background reactivity with native amino acids, the reduction and oxidative coupling reactions were run at pH 6.0. Two aniline probes were utilized for these studies for detection by either fluorescence imaging or western blot. Initial studies were carried out with a commercially available fluorescent aniline substrate, fluoresceinamine isomer I, and subsequent studies relied on a synthetic biotinylated aniline substrate. The experiments discussed in the next sections were carried out in collaboration with an undergraduate, Bryce Jarman, and a rotation student, Rapeepat Sangsuwan.

5.3 Introduction of artificial nitro-tyrosine residues

To test the ability of the ferricyanide-mediated oxidative coupling reaction to detect tyrosine nitration, the nitrophenol moiety was first introduced to protein substrates. Tyrosine residues can be directly nitrated with tetranitromethane; however, for these initial investigations surface accessible lysine residues were acylated with an activated ester to introduce the nitrophenol functionality (Figure 5.3a). This

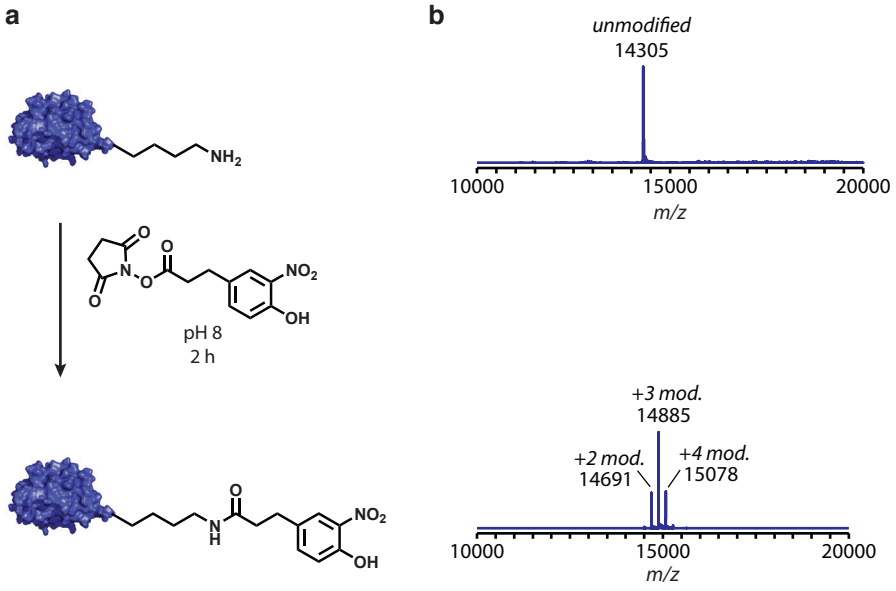


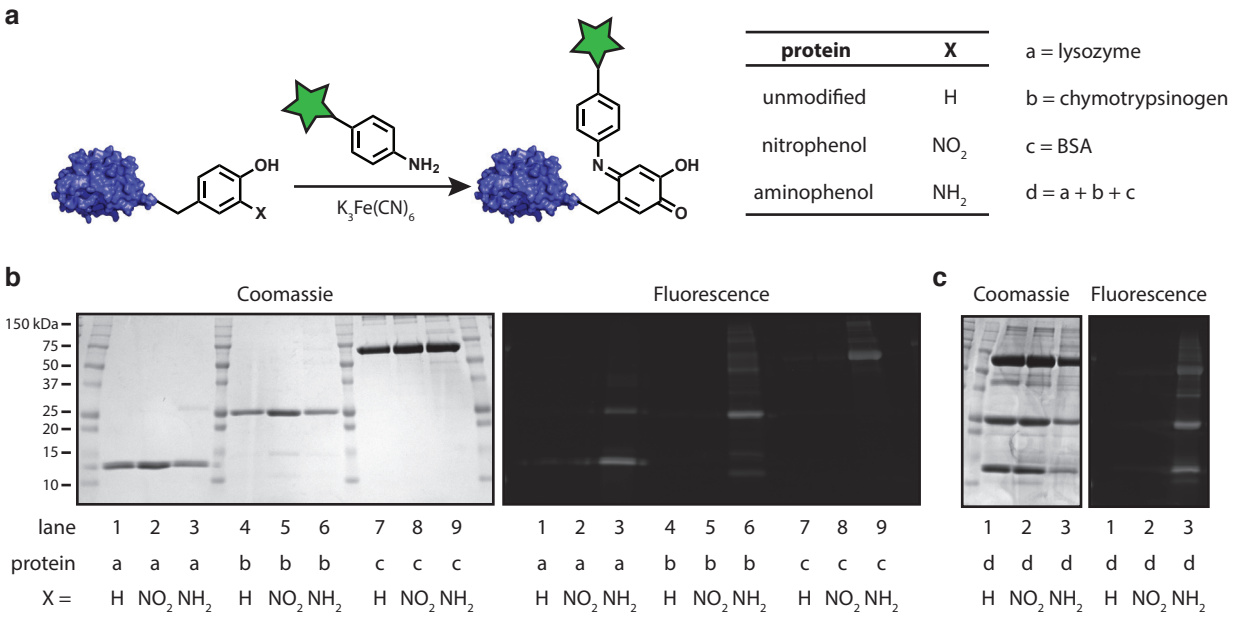
Figure 5.3. (a) Schematic for the introduction of nitro-tyrosine like residues on protein substrates. (b) LC-MS analysis of unmodified lysozyme (top) and lysozyme modified with the nitrophenol functionality (bottom).

method was chosen as the level of modification could easily be controlled by varying the equiv of activated ester. Additionally, the mass addition was large enough to monitor by LC-MS (Figure 5.3b). Several proteins — lysozyme, α-chymotrypsinogen A, and bovine serum albumin — were acylated for use as model substrates for the detection of nitro-tyrosine.

5.4 Specificity and sensitivity of the oxidative coupling reaction

The specificity of the oxidative coupling reaction for the aminophenol modified protein was first verified. The nitrophenol modified proteins were treated with $Na_2S_2O_4$ to generate aminophenol modified proteins. The resulting aminophenol proteins were purified to remove excess reductant. Unmodified, nitrophenol modified, and aminophenol modified proteins were then treated with 10 equiv of fluoresceinamine (100 μM) in the presence of 1 mM ferricyanide (Figure 5.4). Slightly longer reaction times, 1 h, were used to maximize the level of modification. The reaction mixtures were purified to remove excess, unreacted fluorophore. The purified reactions were then analyzed by SDS-PAGE and labeling was monitored by fluorescence detection. Only the aminophenol modified proteins were successfully labeled with the fluorescent aniline, demonstrating the feasibility of using this reaction to selectively detect native tyrosine nitration.

While the oxidative coupling with fluoresceinamine was specific to the aminophenol functionalized protein, the reactions were run with relatively high concentrations (10 μM) of protein. In biological samples, nitrated proteins are more likely to be present at nanomolar concentrations. To determine if the reaction would be sensitive enough to detect nitro-tyrosine residues at biologically relevant concentrations, varying concentrations of aminophenol lysozyme were subjected to the



conditions : 10 μM lysozyme, 100 μM aniline, 1 mM ferricyanide, 100 mM phosphate buffer, pH 6.0, 60 min

Figure 5.4. (a) Schematic for testing the selectivity of the oxidative coupling for *o*-aminophenol residues. (b,c) Unmodified, nitrophenol, and aminophenol proteins were reacted with 10 equiv of fluoresceinamine in the presence of ferricyanide at pH 6.0 for 60 min and then analyzed by SDS-PAGE. Fluorescent and Coomassie-stained images show that the reaction is selective for the aminophenol functional group.

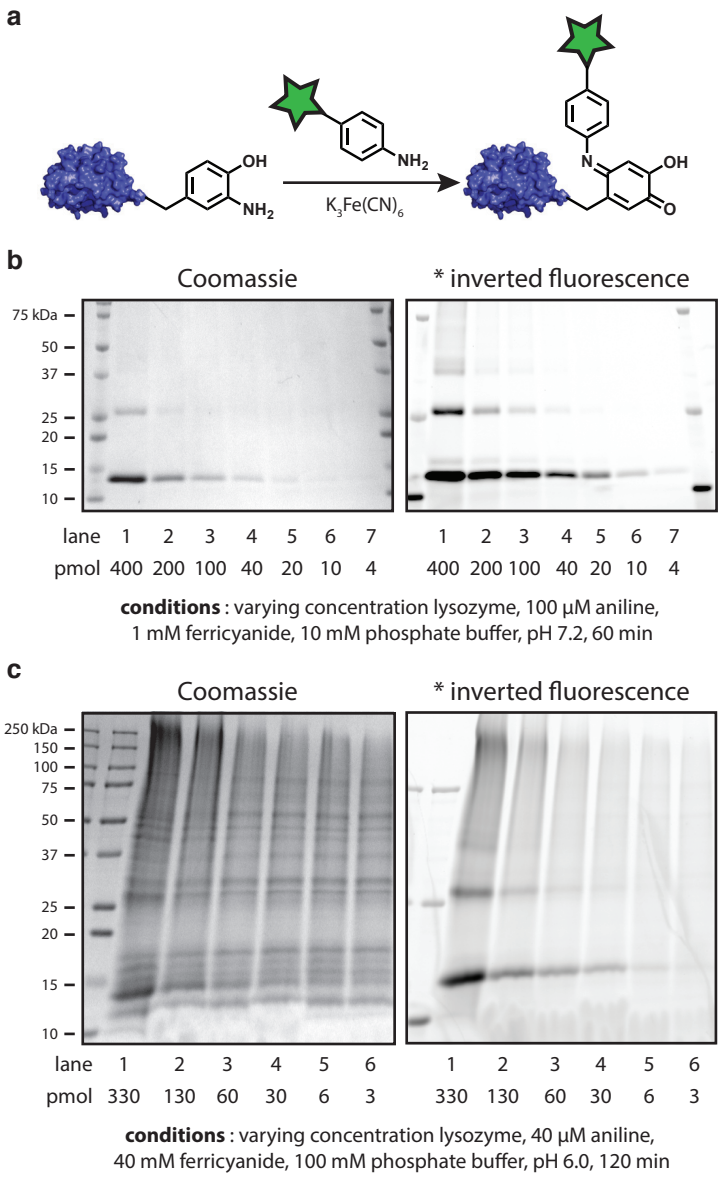


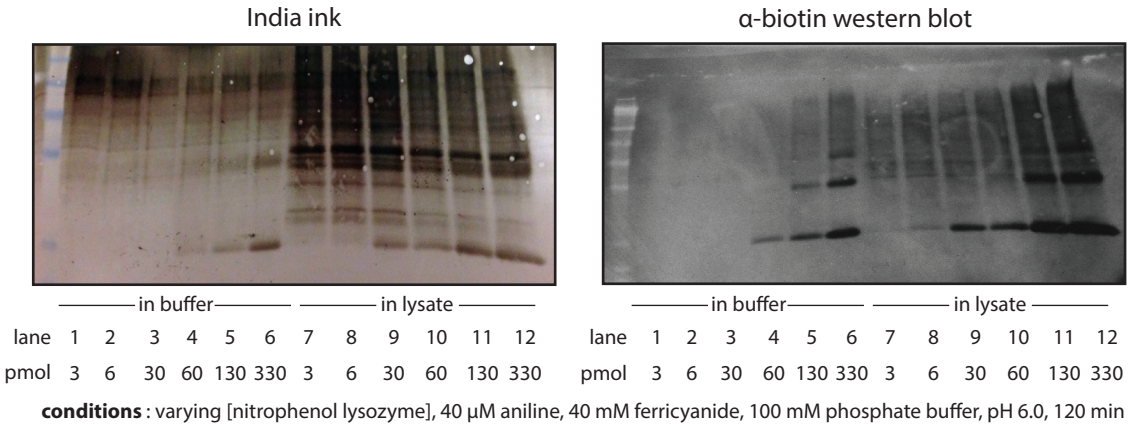
Figure 5.5. (a) Schematic for the oxidative coupling reaction on aminophenol containing proteins. (b) The detection limit was probed on purified, aminophenol modified lysozyme. Reactions were analyzed by SDS-PAGE with fluorescence detection. (c) The detection limit was also evaluated in the prescence of cell lysate. Reduction of nitrophenol modified lysozyme and the oxidative coupling were performed in 'one-pot' and then analyzed by SDS-PAGE with fluorescence and Coomassie detection.

reaction conditions. Aminophenol functionalized lysozyme could be detected even with only a few pmol of protein substrate, equating to concentrations in the nanomolar range (Figure 5.5a).

5.5 Detection of artificial nitro-tyrosine in cell lysate

The oxidative coupling reaction was able to modify proteins selectively, even at low concentrations; however, the reaction would need to maintain selectivity and sensitivity for aminophenol functionalized proteins in a complex mixture of cellular components in order to serve as a useful detection method. To assess the viability of detecting nitrated proteins in complex cellular milieu, nitrophenol modified lysozyme was added to cell lysate. The $Na_2S_2O_4$-mediated reduction was then carried out in the protein mixture. The reduced lysate was then treated with fluoresceinamine and $K_3Fe(CN)_6$. No purification was performed between the reduction and oxidation steps to minimize the number of steps and thus minimize the loss of the biological sample. After the oxidative coupling step the small molecule reactants were removed by precipitation of the protein samples. The precipitated proteins were then resuspended in buffer and analyzed by SDS-PAGE. Excitingly, the exogenous nitrophenol lysozyme was modified with the fluorophore.

Conditions for the reduction and oxidative coupling steps in cell lysate were then optimized. The concentration of dithionite, ferricyanide, and fluoresceinamine were varied, as was the time of the oxidative coupling step. Signal-to-background was optimized when 10 µg of cell lysate was treated with 2 mM $Na_2S_2O_4$ for 1 h followed by reaction with 20 µM fluoresceinamine and 20 mM $K_3Fe(CN)_6$ for 2 h. Using these conditions, as little as 6-30 pmol of nitro-tyrosine modified lyso-



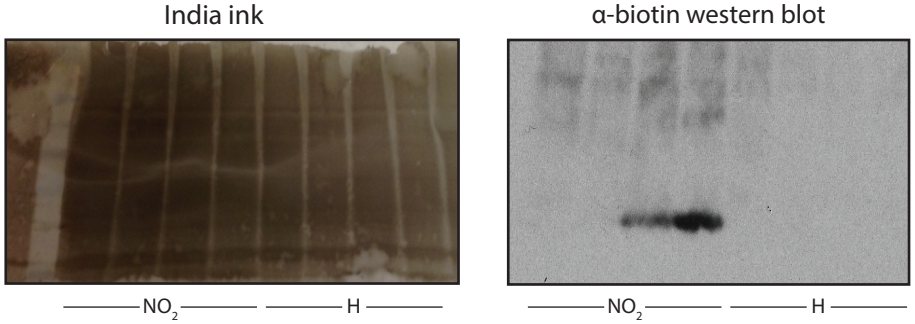
conditions : varying [nitrophenol lysozyme], 40 μM aniline, 40 mM ferricyanide, 100 mM phosphate buffer, pH 6.0, 120 min

Figure 5.6. Detection of nitrophenol lysozyme via an anti-biotin western blot. The nitrophenol protein was reduced with 4 mM $Na_2S_2O_4$ for 1 h and then treated with 40 μM biotin-aniline and 40 mM $K_3Fe(CN)_6$ for 2 h either in the presence (in lysate) or absence (in buffer) of 20 μg cell lysate. After purification, the reactions were analyzed by western blot.

zyme could be detected (Figure 5.5). However, some background non-specific association of the fluorophore with the cellular lysate was also observed.

To avoid this background association as well as possibly improve the detection limit, a biotinylated aniline substrate was tested. The aniline substrate was synthesized in one step from commercially available starting materials and was used without further purification. Detection of nitrophenol modified lysozyme was tested both in buffer and cell lysate using the optimized conditions described above. The modification was analyzed by western blot (Figure 5.6). In these initial studies the detection limit with the biotin aniline matched that of the fluorescent aniline, but optimization of the western blot conditions could improve the detection limit. Additionally, the selectivity for nitrophenol-containing proteins in a complex mixture was confirmed. Both unmodified and nitrophenol modified lysozyme were added to cell lysate and then subjected to the optimized conditions. Only the reactions containing the nitrophenol modified lysozyme were detected by the anti-biotin western blot (Figure 5.7).

The use of the ferricyanide-mediated oxidative coupling reaction for the detection of post-translational tyrosine nitration was preliminarily investigated. Several protein substrates were modified to introduce artificial nitrophenol moieties. The ni-



conditions : varying [nitrophenol lysozyme], 40 μM aniline, 40 mM ferricyanide, 100 mM phosphate buffer, pH 6, 120 min

Figure 5.7. Detection of nitrophenol lysozyme via an anti-biotin western blot. Lysozyme, either modified with a nitrophenol moiety (NO_2) or unmodified (H) was reduced with 4 mM $Na_2S_2O_4$ for 1 h and then treated with 40 μM biotin-aniline and 40 mM $K_3Fe(CN)_6$ for 2 h in the presence of 20 μg cell lysate. After purification, the reactions were analyzed by western blot.

trophenol functional group was selectively detected by reduction to the corresponding aminophenol and then oxidative coupling to an aniline probe. The reduction and oxidative coupling reactions were also successfully carried out in cell lysate.

5.6 Materials and methods

General methods

Unless otherwise noted, the chemicals and solvents used were of analytical grade and were used as received from commercial sources. $K_3Fe(CN)_6$ was purchased from Sigma Aldrich and used without further purification. Analytical thin layer chromatography (TLC) was performed on EM Reagent 0.25 mm silica gel 60-F_{254} plates with visualization by ultraviolet (UV) irradiation at 254 nm and/or potassium permanganate stain. Purifications by flash chromatography were performed using EM silica gel 60 (230-400 mesh). The eluting system for each purification was determined by TLC analysis. Chromatography solvents were used without distillation. All organic solvents were removed under reduced pressure using a rotary evaporator. Water (dd-H_2O) used as reaction solvent was deionized using a Barnstead NANOpure purification system (ThermoFisher, Waltham, MA). Centrifugations were performed with an Eppendorf Mini Spin Plus (Eppendorf, Hauppauge, NY).

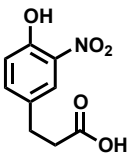
Instrumentation and sample analysis

NMR. ^{1}H and ^{13}C spectra were measured with a Bruker AVB-400 (400 MHz, 100 MHz) or a Bruker AV-600 (600 MHz, 150 MHz) spectrometer, as noted. ^{1}H NMR chemical shifts are reported as δ in units of parts per million (ppm) relative to $CDCl_3$ (δ 7.26, singlet). Multiplicities are reported as follows: s (singlet), d (doublet), t (triplet), dd (doublet of doublets), br (broad) or m (multiplet). Coupling constants are reported as a J value in Hertz (Hz). The number of protons (n) for a given resonance is indicated as nH and is based on spectral integration values. ^{13}C NMR chemical shifts are reported as δ in units of parts per million (ppm) relative to $CDCl_3$ (δ 77.2, triplet).

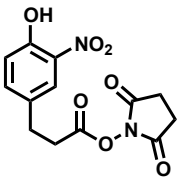
Mass Spectrometry. Protein bioconjugates were analyzed using an Agilent 1200 series liquid chromatograph (Agilent Technologies, USA) that was connected in-line with an Agilent 6224 Time-of-Flight (TOF) LC/MS system equipped with a Turbospray ion source.

Gel Analyses. For protein analysis, sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) was carried out on a Mini-Protean apparatus (Bio-Rad, Hercules, CA), using a 10-20% precast linear gradient polyacrylamide gel (Bio-Rad). The sample and electrode buffers were prepared according to Laemmli.[9] All protein electrophoresis samples were heated for 5-10 min at 95 °C in the presence of 1,4-dithiothreitol (DTT) to ensure reduction of disulfide bonds. Gels were run for 75-90 minutes at 120 V to separate the bands. Commercially available markers (Bio-Rad) were applied to at least one lane of each gel for assignment of apparent molecular masses. Visualization of protein bands was accomplished by staining with Coomassie Brilliant Blue R-250 (Bio-Rad). For fluorescent protein conjugates, visualization was accomplished on Typhoon 9410 variable mode imager (Amersham Biosciences) prior to gel staining. Gel imaging was performed on an EpiChem3 Darkroom system (UVP, USA). ImageJ was used to determine the level of modification by optical densitometry.

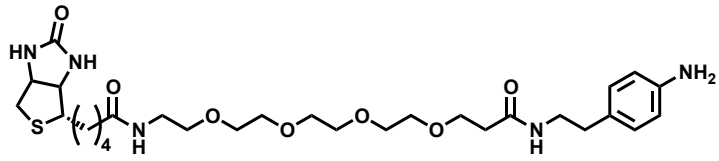
Small molecule synthesis



Synthesis of nitrophenol acid (1). To a solution of 3-(4-hydroxyphenyl)propionic acid (5 g, 30.1 mmol) in 25 mL acetic acid at 15 °C was added a solution of fuming nitric acid (1.6 mL, 33.8 mmol) in acetic acid (4 mL). The solution immediately turned orange. After 15 min the reaction was quenched by addition to ice water. The precipitate was filtered and dried. The yellow precipitate was recrystallized from 1:1 ethanol:water to afford 3.07 g of a yellow powder (48% yield). ^{1}H NMR (400 MHz, $CDCl_3$): δ 10.47 (br s, 1H), 7.94 (d, 1H, *J* = 1.9), 7.45 (dd, 1H, *J* = 8.6, 1.9), 7.09 (d, 1H, *J* = 8.6), 2.94 (t, 2H, *J* = 7.4), 2.68 (t, 2H, *J* = 7.4). ^{13}C NMR (100 MHz, $CDCl_3$): δ 178.22, 153.86, 138.11, 133.47, 132.60, 124.22, 120.27, 35.19, 29.37. HRMS (ESI) calculated for $C_9H_8O_5N$ ([M-H]$^-$) 210.0408, found 210.0410 m/z.



Synthesis of nitrophenol NHS ester (2). To a solution of 3-(4-hydroxy-3-nitrophenyl)propionic acid (**1**) (1 g, 4.7 mmol) and *N*-hydroxysuccinimide (0.65 g, 5.7 mmol) in CH_2Cl_2 at 0 °C was added *N*-(3-dimethylaminopropyl)-*N*'-ethylcarbodiimide hydrochloride (1.09 g, 5.7 mmol). The reaction mixture was stirred for 2 h and then diluted with CH_2Cl_2 and washed with water. The combined organic layers were dried over sodium sulfate and the solvent was removed *in vacuo*. The reaction afforded 1.38 g of a yellow solid (94% yield). ^{1}H NMR (400 MHz, $CDCl_3$): δ 10.49 (s, 1H), 7.98 (d, 1H, *J* = 2.1), 7.48 (dd, 1H, *J* = 8.6, 2.1), 7.12 (d, 1H, *J* = 8.6), 3.05 (t, 2H, *J* = 7.3), 2.93 (t, 2H, *J* = 7.3), 2.83 (s, 4H). ^{13}C NMR (100 MHz, $CDCl_3$): δ 169.06, 167.56, 154.09, 138.00, 133.54, 131.40, 124.52, 120.49, 32.48, 29.43, 25.70. HRMS (ESI) calculated for $C_{13}H_{11}O_7N_2$ ([M-H]$^-$) 307.0572, found 307.0568 m/z.



Synthesis of aniline-biotin substrate (3). To a solution of EZ-Link NHS-PEG_4-Biotin (25 mg, 0.042 mmol, Thermo Scientific) in CH_2Cl_2 was added 2-(4-aminophenyl)ethylamine (6.1 µL, 0.046 mmol) and triethylamine (12 µL, 0.084 mmol). The reaction mixture was stirred for 4 h and then diluted with CH_2Cl_2 and washed with water. The combined organic layers were dried over sodium sulfate and the solvent was removed *in vacuo*. The reaction afforded 25 mg of a pale yellow oil (quantitative yield). HRMS (ESI) calculated for $C_{29}H_{48}O_7N_5S$ ([M+H]$^+$) 610.3274, found 610.3275 m/z.

Protein modification

General procedure for lysine acylation. To 200 µL of 50 µM protein in 50 mM phosphate buffer pH 8.0 was added 0.5 µL nitrophenol NHS-ester as a 100 mM solution in DMSO. The rection was incubated at room temperature for 1 h and then purified by repeated centrifugal filtration using a 0.5 spin filter with the appropriate molecular weight cut off (MWCO, Millipore).

General procedure for oxidative coupling in buffer. To a solution of nitrophenol modified protein (lysozyme, chymotrypsinogen, or BSA, 1-10 µM) in 72 µL 10 mM phosphate buffer, pH 6.0 was added 4 µL of a 100 mM solution of $Na_2S_2O_4$ in 100 mM phosphate buffer, pH 6.0. The solution was briefly vortexed and incubated at room temperature for 1 h. To the solution of reduced

protein was added 4 μL of a 1 mM solution of aniline probe (fluoresceinamine I or **3**) and 20 μL of a 200 mM solution of $K_3Fe(CN)_6$. The solution was briefly vortexed and then incubated at room temperature for 2 h. The reaction mixtures were purified by using a 0.5 mL centrifugal filter with the appropriate MWCO. The samples were concentrated to 50 μL and then diluted 10-fold with 10 mM phosphate buffer, pH 6.0. This process was repeated 3-6 times. After purification the reactions were analyzed by SDS-PAGE.

Oxidative coupling in cell lysate. The procedure outlined above for the oxidative coupling in buffer was performed in the presence of 20 μg mammalian cell lysate (untreated Jurkat or Ramos cell lysate). Reaction mixtures were purified by precipitation with MeOH/$CHCl_3$ instead of centrifugal filtration.

Other procedures

Preparation of cell lysate. Cells were collected and then washed twice in 10 mL of DPBS (with centrifugation in between each wash step to pellet cells at 1300 rpms for 3 min). The cell pellet was resuspended in lysis buffer (1% NP-40, 150 mM NaCl, 20 mM Tris pH 7.4, and protease inhibitors (Calbiochem, set III, diluted 1:100 or EDTA-free complete tablets)) and then disrupted by sonication (5 s on, 5 s off, for 1 min of sonication); the volume of lysis buffer was two times the volume of the cell pellet or 200 μL, whichever was larger. Insoluble material was removed by centrifugation (12000 rpms for 10 min at 4 °C). Protein concentration was determined using a BCA assay (Piece).

Precipitation of proteins. To a sample of protein in 100 μL buffer was added 400 μL cold MeOH followed by 100 μL cold $CHCl_3$. The sample was vortexed after addition of each portion of organic solvent. The solution was diluted to 1 mL with 300 μL of cold dd-H_2O and then vortexed. The aqueous and organic layers were separated by centrifugation at 14,000 g for 2 min. The aqueous layer was removed by pipet. To the remaining $CHCl_3$ was added 400 μL MeOH and the solution was vortexed. The protein was pelleted by centrifugation at 14,000 g for 3 min. The organic solvent was removed by pipet and the residual solvent was allowed to evaporate at room temperature. The protein was resuspended in 1% Triton-X before analysis.

Synthesis of peroxynitrite. To a stirring solution of 1.8 mL of 0.55 M NaOH in 2 mL isopropanol was sequentially added 0.093 mL 30% H_2O_2 and 7 μL isoamyl nitrite. The reaction was incubated at room temperature for 10 min. The peroxynitrite was extracted with 4 x 2 volumes of CH_2Cl_2 and then treated with appx. 0.5 g of MnO_2. The solution was stirred for 5 min and then filtered to remove the MnO_2. After purging the solution with N_2, the concentration was determined by measuring the absorbance at 302 nm ($\varepsilon = 1670\ M^{-1}\ cm^{-1}$).

Western blot protocol. Protein samples were run on a SDS-PAGE gel and then the contents of the gel were wet-transferred to nitrocellulose membranes in tris-glycine buffer (120 min, 60 V). The blots were blocked overnight at 4 ºC in 20 mL tris-buffered saline with 0.1% Tween-20 (TBST) containing bovine serum albumin (5% w/v). To the blot in blocking solution was added 0.2 μL α-biotin antibody conjugated to horse-radish peroxidase (1:100,000 dilution). The antibody was incubated with the blot for 30 min at room temperature. The blot was then washed with TBST (3 x 15 min) followed by detection using chemiluminescence using SuperSignal West Pico Chemiluminescent

Substrate (Pierce).

5.7 References

1. Radi, R. *Proc. Natl. Acad. Sci. U.S.A*. **2004**, *101*, 4003–4008.
2. Monteiro, H. P.; Arai, R. J.; Travassos, L. R. *Antioxid. Redox Signaling* **2008**, *10*, 843–890.
3. Souza, J. M.; Peluffo, G.; Radi, R. *Free Radic. Biol. Med*. **2008**, *45*, 357–366.
4. Abello, N.; Kerstjens, H. A. M.; Postma, D. S.; Bischoff, R. *J. Proteome Res*. **2009**, *8*, 3222–3238.
5. Bachi, A.; Dalle-Donne, I.; Scaloni, A. Chem. Rev. 2013, 113, 596–698.
6. Nikov, G.; Bhat, V.; Wishnok, J. S.; Tannenbaum, S. R. *Anal. Biochem*. **2003**, *320*, 214–222.
7. Wisastra, R.; Poelstra, K.; Bischoff, R.; Maarsingh, H.; Haisma, H. J.; Dekker, F. J. *ChemBioChem* **2011**, *12*, 2016–2020.
8. Obermeyer, A. C.; Jarman, J. B.; Netirojjanakul, C.; El Muselmany, K.; Francis, M. B.; *Angew. Chem. Int. Ed*. **2013**, early view.
9. Laemmli, U. K. *Science* **1970**, *227*, 680–685.