**Title**
Oxidative Coupling of ortho-Aminophenols and Anilines for the Application of Labeling Proteins with Fluorine-18



**Author**
Behrens, Christopher Richard

**Publication Date**
2012

Peer reviewed|Thesis/dissertation

# Oxidative Coupling of *ortho*-Aminophenols and Anilines for the Application of Labeling Proteins with Fluorine-18

By

Christopher Richard Behrens

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 Christopher J. Chang

Professor Andrew O. Jackson

Spring 2012

Abstract

Oxidative Coupling *ortho*-Aminophenols and Anilines for the

Application of Labeling Proteins with Fluorine-18

by

Christopher Richard Behrens

Doctor of Philosophy in Chemistry

University of California, Berkeley

Professor Matthew Francis, Chair

Chemically modified proteins have been of great use in both the pharmaceutical industry and in academic research. Development of new reactions is necessary to further increase the diversity of protein-based biological agents available to researchers and physicians. In particular, high efficiency reactions that can site-specifically modify proteins are sought by the research community. In this thesis, a new reaction that oxidiatively couples *o*-aminophenols and anilines at very high rates was presented, as the MS2 viral capsid coat protein was fully modified site-specifically with a small molecule in under 30 s. The products of the reaction were characterized using small molecule model compounds and a reaction mechanism was posulated based on the available information.

Further, the reaction was used to synthesize of variety of protein-peptide or protein-polymer conjugates. Over 100 copies of 2k or 5k molecular weight poly(ethyleneglycol) (PEG) were attached to a 28 nm diameter MS2 virus-like particle in under 2 min using this new oxidative coupling reaction. In addition, several small peptides were also attached in the same way with similar results. The reaction was also utilized for modification of other proteins such as lysozyme to show its general applicability.

Finally, the reaction was applied to labeling proteins with radioactive Fluorine-18 for the purpose of positron emission tomography (PET). Fluorine-18 was incorporated into a small molecule prosthetic group to form [$^{18}$F]-*p*-fluoroaniline ([$^{18}$F]-FA). $^{18}$F-FA was successfully coupled to both lysozyme and MS2 virus-like particles containing aminophenol functionality. Again the reaction reached completion in under 2 min and gave excellent yields.

It is our hope that this reaction will be utilized by other protein chemists looking for a highly efficient way to assemble challenging protein bioconjugates.

This manuscript is dedicated to my parents, Richard and Catherine Behrens,
who always pushed me to be my best. Also to the love of my life, Nicole Shinbori,
for all of her support and encouragement.
Without them none of this would have been possible.

# Table of Contents

Acknowledgements

I'd first like to thank Matt for being a great advisor who let me pursue some of the crazy ideas I came up with. His passion for chemistry will always stick with me even when I move on to other research fields. I'd also like to thank Jacob and Dante for paving the way on my project. Without their contributions and input my road would have been significantly tougher. Thanks to Aaron for showing us all what it takes to be successful in academia while still being a great guy and making all of us laugh. My two longtime labmates Gary and Zac couldn't have been more different but both brought important characteristics to the lab. Gary and I had countless conversations and jokes about a variety of inappropriate topics and this served to bring levity to the long five years of graduate school. Gary provided valuable insight but was also brutally honest when needed – these characteristics made him a great friend inside and outside of lab. Zac on the other hand brought a very professional approach to the lab. His tireless work ethic helped make everybody else more productive and I had some of my best conversations about science with him. These older students served as my role models and I was lucky to follow in their footsteps.

Unlike most other Francis group classes, my year had only two students. My fellow classmate Leah is one of the nicest people I've ever met. She is constantly thoughtful of everybody, not just her friends. She's been very successful in lab and whatever organization she decides to eventually work for is going to be extremely lucky to have her. She's been a great friend and her relentlessly positive attitude will always be remembered.

I hope I've been able to aid the younger students as much as the older students helped me. There's too many of them to name them all, but just know that I enjoyed working with all of them, no matter how grumpy I became in the last year or two. Dan, Troy, Amy, and Mike were a lot of fun to be around for the last four years. They definitely made the lab an enjoyable place to work. Allie was also a pleasure to be around both in and out of lab. Her help was very important in forming the hypothesis for the mechanism of the aminophenol-aniline oxidative coupling. She's great at everything and I'm sure she's going to excel at whatever she decides to do. Kristen and Jelly both always impressed me as good scientists. They too have a bright future in science should they pursue it. Abigail is the kid who started hanging out with the old kids as soon as she joined the lab. She's either going to develop into a Francis group leader or become incredibly bitter within the next year – hopefully the former! Last but not least, Tony and Richard… I teased you two quite a bit but you've actually really impressed me. I expect to visit in a few years and see big things from you guys. The group is in good hands heading into the future.

Outside of the group, I'd like to thank in particular Jim O'Neil. He was my mentor at LBNL for all of the radiochemistry work and always was eager to help me out in any way he possibly could. One of my major regrets in graduate school was not spending more time working with him. Jim's coworkers, Mustafa and Nick, were always very helpful and welcomed me into their lab.

Apologies to anybody I skipped, but you were all great and made my graduate career enjoyable.

# Chapter 1 - Introduction to Chemical Modification of Proteins

## Section 1.1: Importance of Chemical Modification of Proteins

The chemical modification of proteins is a large and diverse field that has grown exponentially in recent years due to development of analytical technology such as mass spectrometry. The ability to easily assay the extent of a protein modification reaction has given chemists the necessary tool to test for new reactivity. Reactions for protein modification have subsequently produced numerous bioconjugates used in medicine and research.

Chemical modifications of proteins have made a profound contribution to the field of medicine. Many pharmaceutical formulations contain proteins that have been chemically modified prior to administration. These modifications can make the drug more potent or increase circulation halflife. Both of these effects can allow for a smaller drug dose, thus decreasing side effects and cost of therapy. The most common chemical modification used in the pharmaceutical industry is the attachment of polyethyleneglycol (PEG). Numerous FDA-approved drugs utilize PEG modification in their formulations (Figure 1-1). PEG provides several advantageous properties to its parent protein. Reduced elimination by the reticuloendothelial system, prevention of enzymatic degradation, and avoidance of the body's immune response are the three most important of these.[1] Once sufficient PEG is attached to a protein it forms a layer surrounding the biomolecule that protects it from adsorption to serum proteins and immunoglobulin recognition, thus creating a "stealth" agent.[2] These properties are vital to the success of several protein-based drugs and could not be used without the use of chemical modifications. There is not currently a known method for the PEGylation of proteins through biological transformations.

Besides PEGylation, antibody-drug conjugates are another class of chemically modified proteins of interest to the pharmaceutical industry. Though there is only one antibody-drug conjugate currently approved by the FDA (Adcetris) there are countless others in late phase clinical trials. Antibody-drug conjugates aim to target a powerful cytotoxic drug to a specific receptor that is upregulated in tumor cells. This avoids healthy cell toxicity and thus lessens the drug side-effects. Using Adcetris as an example, it is synthesized by first reducing the native disulfides in an anti-CD30 antibody. Then the resulting thiols are modified with monomethyl auristatin E using a cleavable linker with maleimide functionality.[3] The resulting antibody-drug conjugate is stable until it enters tumor cells where the cytotoxic agent is freed by cleavage of the linker by cathepsin. Again, this construct could not be synthesized through purely biological means, as chemistry is necessary to attach the cytotoxic agent.

Though the contributions of chemically modified proteins to medicine are significant, they pale in comparison to the contributions in research. The majority of the biological experiments that track proteins use some sort of fluorescent or radioactive label that was chemically attached to the protein of interest. Other uses include attaching affinity reagents (biotin) or hapten molecules to a carrier protein.[4] There are numerous reviews that describe the multitude of ways to fluorescently label a protein,[5] radiolabel a protein (discussed further in Chapter 3),[6] or couple haptens to a carrier protein.[7] All of these methods rely heavily on protein modification reactions to synthesize the bioconjugates necessary to investigate biological questions.

| PEG drug description | Company | Indication | Year of approval |
|---|---|---|---|
| Adagen (11–17×5 kDa mPEG per adenosine deaminase) | Enzon Inc. (USA & Europe) | severe combined immunodeficiency | 1990 (USA) |
| Oncospar (5 kDa mPEG-L-asparaginase) | Enzon Inc. (USA)/ Rhône–Poulenc Rorer (Europe) | acute lymphoblastic leukemia | 1994 (USA) |
| Doxil/Caelyx (SSL formulation of doxorubicin) | Alza Corp. (USA)/ Schering-Plough Corp. (Europe) | Kaposi's sarcoma, ovarian cancer, breast cancer, multiple myeloma | 1995 (USA) 1999 (USA) all 1996 (EU) |
| PEG-Intron (2×20 kDa mPEG-interferon-α-2a) | Schering- Plough Corp. (USA & EU) | chronic hepatitis C | 2000 (EU) 2001 (USA) |
| Pegasys (12 kDa mPEG-interferon-α-2b) | Hoffmann-La Roche (USA & EU) | chronic hepatitis C | 2002 (USA & EU) |
| Neulasta (20 kDa mPEG-G-CSF) | Amgen Inc. (USA & EU) | febrile neutropenia | 2002 (USA & EU) |
| Somavert (4–6×5 kDa mPEG per structurally modified HG receptor antagonist) | Pfizer (USA & EU) | acromegaly | 2002 (EU) 2003 (USA) |
| Macugen (2×20 kDa mPEG- anti-VEGF- aptamer) | Pfizer (EU)/OSI Pharm. Inc. and Pfizer (USA) | age-related macular degeneration | 2004 (USA) 2006 (EU) |
| Cimzia (2×40 kDa mPEG- anti-TNFα) | UCB S. A. (USA & EU) | Crohn's disease, rheumatoid arthritis | 2008 (USA) 2009 (USA) 2009 (EU) |

[a] mPEG: methoxypoly(ethylene glycol), SSL: sterically stabilized liposome, G-CSF: granulocyte-colony stimulating factor, HG: human growth, VEGF: vascular endothelial growth factor, TNF: tumor necrosis factor.

**Figure 1-1**. List of approved drugs that have been chemically modified with PEG. Figure courtesy of Knop, K., Hoogenboom, R., Fischer, D., and Schubert, U.S. *Angew. Chem. Int. Ed.* **2010**, *49*, 6288–6308.

## Section 1.2: Examples of Protein Modification Reactions with Native Amino Acid Side Chains

The 20 naturally occurring amino acids have a wide variety of side chain groups. Since most proteins contain at least one of each of the amino acids, the protein environment is a very complex mix of charged species, hydrogen bonds, and functional groups. When organic chemists approach molecules with such complexity they usually devise a creative plan to mask the reactive functionality with protecting groups. This is not feasible for a chemist trying to modify proteins. In addition to the complexity, there are other challenges in protein modification. Nearly all reactions on proteins are performed in aqueous solution and near neutral pH. The number of known reactions that are specific to a particular functional group and work in aqueous solution near neutral pH is relatively small. Figure 1-2 lists the amino acid side chains that are most often targeted for modification. Most of the side chains listed are nucleophilic with the exception of the carboxylic acid side chains which are converted into electrophiles for modification. Also, the conditions are specific for each reagent to give the reaction specificity. For example, the modification of imidazole (histidine side chain) with diethyl pyrocarbonate is performed at pH 4-5 because at higher pH lysine would also be modified.[8] Because of the chemical complexity of proteins, changing the pH or number of equivalents of a reagent can drastically change which side chains are modified. Two of these reactions have become so common that the functional groups necessary for modification have been attached to many commercially available molecules that scientists have interest

| side chain or group | reagent or procedure | optimum reaction pH, side chain selectivity, and other comments |
|---|---|---|
| amino (Lys + $\alpha$) | amidination (ethyl acetimidate) | pH ~9, no other side chains react, positive charge maintained, other imido esters are available, extent of modification may be determined with TNBS |
| | reductive alkylation (formaldehyde + $NaBH_4$ or $NaBH_3CN$) | pH ~9 with $NaBH_4$, pH ~7 with $NaBH_3CN$; reaction is much slower under the latter conditions; no other side chains react; positive charge maintained; other aldehydes and reducing agents may be used; extent of modification may be determined by amino acid analysis, the incorporation of radiolabel, or with TNBS |
| | acylation (acetic anhydride) | pH ~8 and above, Tyr residues also modified, elimination of positive charge, extent of modification may be determined with TNBS |
| | (succinic anhydride) | same as above, Tyr residues undergo slow deacylation above pH ~5, replaces positive charges with negative charges |
| | trinitrobenzenesulfonate | pH ~8 and above, also reacts slowly with thiol groups, eliminates positive charge and introduces large hydrophobic substituent, extent of reaction may be determined spectrophotometrically |
| carboxyl (Asp + Glu) | water-soluble carbodiimide + nucleophile (EDC + glycine ethyl ester) | pH ~4.5–5, some side reactions with Tyr and thiol groups, other carbodiimides are available, many other nucleophiles (amines) may be used to either maintain or alter the charge, extent of reaction may be determined by amino acid analysis or from incorporation of radiolabel |
| guanidino (Arg) | dicarbonyls [2,3-butanedione, phenylglyoxal, and (*p*-hydroxyphenyl)glyoxal] | pH ~7 or higher, reaction promoted by borate buffer, no major side reactions; partially reversible upon dialysis, eliminates positive charge, extent of reaction can be determined from incorporation of radiolabel or by amino acid analysis, other dicarbonyl compounds can also be used (i.e., cyclohexanedione, glyoxal, etc.). |
| imidazole (His) | diethyl pyrocarbonate (ethoxyformic anhydride) | pH ~4–5, side reactions with Lys kept to minimum by low pH, extent of modification may be determined by spectrophotometric measurement, reversed in the presence of $NH_2OH$ |
| indole (Trp) | *N*-bromosuccinimide | usually pH ~4 or lower, higher pH values can be used; thiol groups are rapidly oxidized; Tyr and His react more slowly; extent of modification may be determined spectrophotometrically or by amino acid analysis |
| | 2-hydroxy-5-nitrobenzyl bromide | pH <7.5, slight reaction with thiols, strong visible absorbance, can be used to determine the extent of reaction |
| phenol (Tyr) | iodination ($I_3^-$, chloramine T + $I^-$, ICl, lactoperoxidase + $I^-$, and $H_2O_2$) | pH ~8 or higher, many different procedures and reagents, His also reacts but usually to a lesser extent, thiol groups are rapidly oxidized, both mono and diiodo derivatives are formed, the extent of reaction can be estimated spectrophotometrically or by amino acid analysis, widely used for radiolabeling of proteins |
| | tetranitromethane | pH ~8 or slightly higher, thiol groups are also rapidly oxidized, some nitration of Trp, extent of reaction may be determined spectrophotometrically or by amino acid analysis |
| thiol (Cys-SH) | carboxymethylation (iodo- and bromoacetate and iodo- and bromoacetamide) | pH ~7 or higher; no effect on other residues under appropriate conditions; Lys, His, Tyr and Met react slowly with excess reagent and long reaction times; extent of reaction may be determined with DTNB, by the incorporation of radiolabel, or by amino acid analysis |
| | *N*-ethylmaleimide | pH ~6 or higher, reaction with Lys and His are much slower at pH 7 and usually of no importance, the extent of reaction may be determined from incorporation of radiolabel or by amino acid analysis |
| | 5,5′-dithiobis(2-nitrobenzoic acid) (Ellman's reagent) | pH ~7 or higher, no other side chains react, reversible in presence of excess low MW thiol, the extent of modification can be determined spectrophotometrically |
| thioether (Met) | oxidation ($H_2O_2$) | pH ~1 and higher, thiol groups also react very rapidly, reversed by treatment with low MW thiols, extent of modification may be determined by amino acid analysis after alkaline hydrolysis or by carboxymethylation followed acid hydrolysis |

**Figure 1-2**. List of most commonly modified amino acid side chains along with the necessary reagents and conditions. Notice that many of the conditions are necessary to avoid modifying other side chains. Figure courtesy of Means, G.E., and Feeney, R.E. *Bioconj. Chem.* **1990**, *1*, 2–12.

in attaching to proteins. *N*-hydroxysuccinimidyl (NHS) esters and maleimides have become the mainstays of simple protein modification.

NHS esters are incredibly useful because they can be used to modify nearly any protein. They mainly react with amines (Figure 1-3a), both the N-terminus and lysine side chains, but have also been shown to modify tyrosines under specific conditions.[9] Most proteins have an N-terminus (though many already have a post-translational modification and are thus unreactive towards NHS esters) and most water soluble proteins possess several lysines because of their role as a solubilizing side chain. NHS esters are also easy to synthesize as they can be made from nearly any carboxylic acid. They also store well assuming they are protected from moisture as water will hydrolyze the ester thus rendering the reagent unreactive. All of these qualities combined make the NHS ester the most broad protein modification reagent used today.

However, the broad reactivity is not always advantageous. For many bioconjugates, the goal is to specifically label the protein in a particular position. For this purpose the maleimide has become the reactive handle of choice because it reacts with cysteines that are not often found on

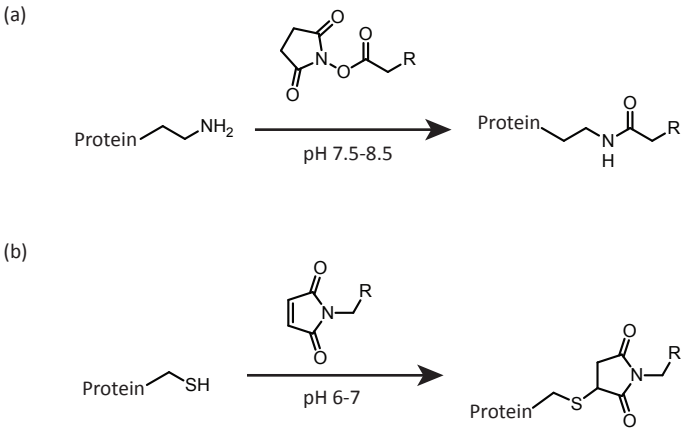



**Figure 1-3**. The two most common reagents for modifying native amino acid side chains. (a) NHS esters react with amines to form amide bonds. This reaction is usually performed at pH 7.5-8.5. At lower pH the lysines react too slowly because they are mostly protonated. (b) Maleimides react with thiols to form thioethers. This reaction is usually performed at pH 6-7. At higher pH lysines can start to react with the maleimide.

the surface of proteins (Figure 1-3b). With the advent of polymerase chain reaction and facile DNA sequencing, it has become quite trivial to make point mutations in recombinantly expressed proteins. Thus it has become exceedingly common to see a protein with a cysteine mutated in at varying positions. Since cysteines are normally either not solvent accessible or in a disulfide bond the newly mutated cysteine is often the only reactive one. Modification of this single cysteine results in a homogeneous mixture of site-specifically modified proteins. This can be useful when labeling an enzyme with a fluorescent dye. Labeling all lysines could impact the active site of the enzyme and change its properties. Modifying the enzyme at a known residue far from the active site will most likely not have an effect. Maleimides are more challenging to synthesize than NHS esters but there is such high demand for them that most fluorescent dyes can be purchased with maleimide functionality already installed. Several examples of maleimide reactions with a mutated cysteine can be found later in the text.

## Section 1.3: Examples of Protein Modification Reactions with Bio-orthogonal Functional Groups

Though cysteine mutation followed by maleimide modification is probably the most common strategy for site-specific modification, it is not always feasible. Therefore, a new class of protein modification reactions have been developed that utilize chemistry not normally found in nature. Carolyn Bertozzi termed this class of reactions “bio-orthogonal”, meaning that they show no reactivity towards native residues and side chains.

The oldest class of such reactions is the aldehyde or ketone condensation with alkoxyamines or hydrazines to form oximes and hydrazones, respectively (Figure 1-4a). Neither of these functional groups form stable covalent bonds with any amino acid side chains. However, for this reaction to be of use, one of these groups must be incorporated into the protein. Numerous clever genetic and enyzmatic approaches have been developed to incorporate aldehydes into proteins. Work in the Bertozzi group developed a short amino acid sequence recognized by an enzyme that can convert a cysteine within the recognition sequence to an aldehyde.[10] Another approach developed by the Schultz group incorporated *p*-acetylphenylalanine as an unnatural amino acid by

evolving a tRNA/tRNA synthetase pair that hijacked the amber stop codon for incorporation of this 21st amino acid.[11] Also, to add a chemical method to this class of reactions, work in the Francis group led to the development of a reaction between pyridoxyl-5-phosphate (PLP) and the N-terminal amine.[12] This reaction specifically transaminates the N-terminal amine to a ketone or aldehyde. Though it shows varied yields based on the N-terminal sequence, further work has isolated several new sequences that give consistently high yields on a variety of recombinant proteins.[13,14]

Once the aldehyde or ketone is incorporated into the protein, it can be specifically modified using an alkoxyamine or hydrazine. In-depth work by Jencks determined the mechanism of this class of condensation reactions.[15] Of importance to this discussion is the effect of pH on the rate of reaction. Two steps of the reaction affect the overall rate, with higher pH favoring the deprotonation of the nucleophile while lower pH favoring the protonation of the carbonyl oxygen to make it a better leaving group. The combination of these effects causes the reaction to reach its maximum rate at pH 4.5, and moving in either direction away from that number causes a drop in rate. This can be problematic for many proteins that are not stable for long periods at pH 4.5. The reaction works reasonably well if large excesses of the nucleophile can be added to the protein. These reactions are also sometimes incubated for 1-2 days to achieve good yield.

The stability of the oxime or hydrazone product is also of interest. Work by the Raines group investigated the relative stability of oximes and hydrazones and found that the hydrolysis rate of oximes was 1000-fold slower than hydrazones.[16] Therefore, oximes are generally used when a permanent linkage is desired while the facile hydrolysis has been widely used for drug conjugates.[17] At low pH the hydrolysis rate of the hydrazone increases, thus making it an attractive release strategy for drug delivery. When the drug conjugate enters the endosome and encounters a low pH environment, the hydrazone linkage is hydrolyzed and the drug is released into the target cell.

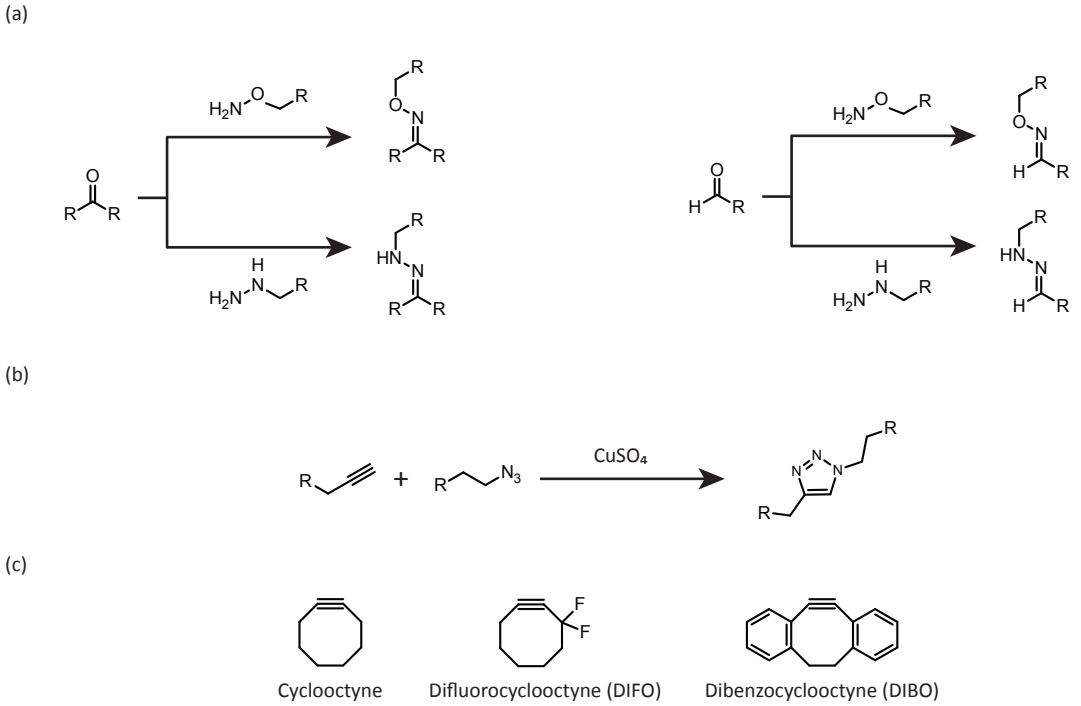



**Figure 1-4**. Bio-orthogonal protein modification reactions. (a) Ketone or aldehyde condensation with alkoxyamines or hydrazines results in oximes and hydrazones, respectively. (b) Alkyne-azide [3+2] cycloaddition is catalyzed by copper. (c) Examples of ring-strained alkynes that react with azides with no copper catalyst present.

The other major class of bio-orthogonal reactions is the azide-alkyne "click" reaction (Figure 1-4b). Seminal work by the Sharpless, Fokin, and Finn groups led to the development of a copper-catalyzed [3+2] cycloaddition reaction between azides and alkynes.[18] This reaction had been known for some time but it required heating that would not be compatible with protein modification. The breakthrough centered around finding a catalyst that accelerated the reaction enough that it no longer needed to be heated. Not only did this work directly lead to many interesting applications,[19,20] but it spawned an entire class of reactions.

The next major contribution to this class of reactions was the use of ring-strain-promoted cycloadditions by the Bertozzi group.[21] They first utilized a cyclooctyne as the alkyne substrate, and the 18 kcal/mol of ring strain made this reaction fast enough to not require copper catalysis (Figure 1-4c). The absence of copper made this reaction useful for applications on live cells, as the copper used in the Sharpless case is toxic. This reaction was still relatively slow and ineffective in dilute cellular environments. Thus an improved, more reactive version of the cyclooctyne was synthesized. Difluorocyclooctyne used two fluorine atoms to lower the LUMO which made the molecule more reactive. Other groups have also developed a large number of ring-strain-promoted cycloaddition reagents that use different methods for lowering the LUMO.[23,24]

As in the aldehyde case, the reactive groups for the [3+2] cycloaddition are commonly incorporated through genetic methods. The Schultz group has developed an unnatural amino acid incorporation strategy for both *p*-azidophenylalanine[25] and *p*-propargyloxyphenylalanine.[26] The Bertozzi group has also developed several methods for incorporation of azides and alkynes by engineering cells to process azide or alkyne modified sugars. The result is cells labeled with azides or alkynes in the cell surface glycoproteins.[27]

The reactions described in this section have been widely used for a variety of applications. There is however room for more such bio-orthogonal reactions provided they fill a niche in protein modification that has not been solved already.

## Section 1.4: The Unexplored Niche in Protein Modification Reactions: Ultra-High Efficiency Site-Specific Reactions

The contributions from Sharpless were impressive when judged by both novelty and impact on the field. The great improvements in rate at ambient temperature produced by the copper catalysis still resulted in reaction times of 12-24 h for the small molecule couplings demonstrated in the original paper.[18] As described above, new reactions for protein modification must fill a niche to be a valuable addition to the field. While the [3+2] cycloaddition shows excellent selectivity and compatibility with a variety of functional groups, its efficiency does leave room for improvement. Therefore, the niche we are trying to fill is that of the ultra-high efficiency yet still site-specific reaction.

Most reactions with a large energy difference between reactants and products have relatively high energy reactants. While this large difference can help in pushing the reaction towards completion, the downside of high energy reactants is the possibility that they could travel down undesirable pathways. If a large amount of undesired product is formed then the reaction is not efficient. One example of this concept is the NHS ester. The activated ester modifies amines rapidly at pH 8-9, but suffers from additional undesired reactions. Histidine side chain groups (imidazoles) can react with activated ester and then quickly hydrolyze back to imidazole and the carboxylic acid, thus catalyzing the hydrolysis of the ester. At pH 8.6 and 4 °C, the estimated halflife of an

NHS ester is approximately 10 min due to hydrolysis.[28] It is therefore necessary to identify a reaction with a large energy difference that proceeds down one predominant reaction pathway.

There are several scenarios that an ultra-high efficiency reaction would be necessary to construct the desired protein conjugate. They all involve situations where the two substrates have limited collisions with each other, therefore necessitating a high efficiency reaction. The first scenario is coupling two large, sterically hindered substrates. Assuming there is one reactive group on each substrate, it is unlikely that when the two collide the reactive groups will be properly positioned for the reaction to occur. An example of this scenario would be coupling two proteins together; and there is a specific example from work in the Francis group. Wesley Wu has attempted to build a construct that consists of a 28 nm viral capsid with approximately 20 copies of an antibody-like protein coupled around the outside. This is a challenging construct to synthesize and requires a very efficient reaction to do so.

Another scenario that would require an efficient reaction is when one of the substrates is precious or expensive. Some drug, peptide, or sugar molecules are synthesized through many chemical steps prior to attachment to proteins and as a result they cannot be added in a large excess. At these low concentrations, reactions such as the ketone/alkoxyamine condensation are extremely slow. A more efficient reaction results in higher levels of protein modification with low concentrations of reagents.

The final scenario is similar to the previous one in that it also involves low concentrations of reagents. In this case, however, the reason for low concentrations is not cost, but instead is the solubility of the reagents. Some structures, such as organic dyes, are quite insoluble in water but have no solubility issues once attached to a protein. Another example from the Francis group illustrates the benefit of a highly efficient reaction. Work by Jake Jaffe in collaboration with Kathryn Strobel from the Doug Clark group has pursued a filamentous phage construct with highly insoluble organic dyes attached. Due to the structure of the phage a site-specific reaction is necessary and mutant cysteines cannot be used. As in the previous example, the dye is not soluble enough to reach the substrate concentrations necessary for the ketone/alkoxyamine condensation reaction. Therefore, an ultra-high efficiency, yet site-specific, reaction is necessary.

Though the bioconjugation reaction toolbox has made great progress in the last 10 years, there are still areas that could use a more effective reaction. This thesis will detail the efforts to identify and characterize such a reaction, and then utilize it for several well-suited applications.

## Section 1.5: References


1. Knop, K., Hoogenboom, R., Fischer, D., and Schubert, U.S. *Angew. Chem., Int. Ed*. **2010**, *49*, 6288–6308.

2. Neoh, K.G., and Kang, E.T. *Polym. Chem.* **2011**, *2*, 747–759.

3. Francisco, J.A., Cerveny, C.G., Meyer, D.L., Mixan, B.J., Klussman, K., Chace, D.F., Rejniak, S.X., Gordon, K.A., DeBlanc, R., Toki, B.E., et al. *Blood.* **2003**, *102*, 1458–1465.

4. Brinkley, M. *Bioconj. Chem.* **1992**, *3*, 2-13.

5. Giepmans, B.N.G., Adams, S.R., Ellisman, M.H., and Tsien, R.Y. *Science*. **2006**, *312*, 217–224.

6. Tolmachev, V., and Stone-Elander, S. *Biochim. Biophys. Acta.* **2010**, *1800*, 487–510.

7. Inman, J.K., Merchant, B., Claflin, L., and Tacey, S.E. *Immunochemistry*. **1973**, *10*, 165–174.

8. Means, G.E., and Feeney, R.E. *Bioconj. Chem.* **1990**, *1*, 2–12.

9. Leavell, M.D., Novak, P., Behrens, C.R., Schoeniger, J.S., and Kruppa, G.H. *J. Am. Soc. Mass Spect.* **2004**, *15*, 1604–1611.

10. Carrico, I.S., Carlson, B.L., and Bertozzi, C.R. *Nat. Chem. Biol.* **2007**, *3*, 321–322.

11. Brustad, E.M., Lemke, E.A., Schultz, P.G., and Deniz, A.A. *J. Am. Chem. Soc.* **2008**, *130*, 17664–17665.

12. Gilmore J.M., Scheck R.A., Esser-Kahn A.P., Joshi N.S., Francis M.B. *Angew. Chem.*, *Int. Ed.* **2006**, *45*, 5307-5311.

13. Scheck R.A., Dedeo M.T., Iavarone A.T., Francis M.B. *J. Am. Chem. Soc.* **2008**, *130*, 11762-11770.

14. 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*, 16812-7.

15. Jencks WP. *J. Am. Chem. Soc.* **1959**, *81*, 475–481.

16. Kalia, J., and Raines, R.T. *Angew. Chem. Int. Ed.* **2008**, *47*, 7523–7526.

17. Lee, C.C., Gillies, E.R., Fox, M.E., Guillaudeu, S.J., Fréchet, J.M.J., Dy, E.E., and Szoka, F.C. *Proc. Natl. Acad. Sci.* **2006**, *103*, 16649–16654.

18. Rostovtsev, V.V., Green, L.G., Fokin, V.V., Sharpless, K.B. *Angew. Chem. Int. Ed.* **2002**, *41*, 2596–2599.

19. Speers, A.E., Adam, G.C., and Cravatt, B.F. *J. Amer. Chem. Soc.* **2003**, *125*, 4686-4687.

20. Lutz, J. *Angew. Chem. Int. Ed.* **2007**, *46*, 1018–1025.

21. Agard, N.J., Prescher, J.A., and Bertozzi, C.R. *J. Am. Chem. Soc.* **2004**, *126*, 15046–15047.

22. Codelli, J.A., Baskin, J.M., Agard, N.J., and Bertozzi, C.R. *J. Am. Chem. Soc.* **2008**, *130*, 11486–11493.

23. van Berkel, S. S., Dirks, A. J., Debets, M. F., van Delft, F. L., Cornelissen, J. J. L. M., Nolte, R. J. M., and Rutjes, F. P. J. T. *Chem. Bio. Chem.* **2007**, *8*, 1504-1508.

24. Ning, X., Guo, J., Wolfert, M. A., and Boons, G.-J. *Angew. Chem., Int. Ed.* **2008**, *47*, 2253-2255.

25. Chin, J.W., Santoro, S.W., Martin, A.B., King, D.S., Wang, L., and Schultz, P.G. *J. Am. Chem. Soc.* **2002**, *124*, 9026–9027.

26. Deiters, A., and Schultz, P.G. *Bioorgan. Med. Chem. Lett.* **2005**, *15*, 1521–1524.

27. Saxon, E., and Bertozzi, C.R. *Science.* **2000**, *287*, 2007–2010.

28. Hermanson, G. T. *Bioconjugate Techniques*, 2nd ed.; Academic Press:London, 2008.

# Chapter 2 – Oxidative Coupling of Aminophenols to Anilines

## Section 2.1: Discovery and Characterization of Aminophenol-Aniline Oxidative Coupling

In search of a more efficient class of protein modification reaction, work by Jacob Hooker and other early members of the Francis group focused on utilizing the unique oxidation properties of electron rich aromatic rings such as anilines. Mild, water soluble oxidants such as sodium periodate cause rapid polymerization of anilines[1] but have little reactivity with native amino acid side chains. Work by the Kodadek lab used sodium periodate to cross link proteins modified with dihydroxylphenylalanine while native side chains showed no reactivity. Therefore, sodium periodate could be used to site-selectively couple aniline substrates to anilines that were incorporated into proteins. Hooker *et al.* developed this strategy into tan oxidative coupling reaction that formed a covalent bond between a *N*,*N*-dialkylphenylenediamine and a *para*-substituted aniline (Figure 2-1).[1] The alkyl substituted phenylenediamine prevented modification of the *ortho* position due to sterics, thus forcing the aniline to add at the *meta* position which resulted in a single product. The reaction proceeded under a wide pH range, thus allowing modification of proteins with varied pI's; and native side chain groups did not react under the mild oxidative conditions used for the reaction; therefore the reaction was site-specific. This allowed for precise construction of protein-based systems and materials. However, many systems are not compatible with the 30-60 min exposure to periodate necessary to complete the reaction. We attempted to improve upon this by investigating similar oxidative coupling reactions for their reaction speed and efficiency as well as compatibility with a wide range of substrates.

While investigating diazonium chemistry on proteins, Hooker observed that the conversion of a tyrosine to an azo compound, followed by reduction with sodium dithionite, gave an *ortho*-aminophenol that was reactive towards anilines in the presence of sodium periodate (Scheme 2-1, Figure 2-2).[3] Unfortunately, when attempted with small molecule model compounds (4-*tert*-butyl-2-aminophenol, *p*-toluidine, and sodium periodate), several products were formed and the isolated major product could not be identified with NMR. As a result this reaction was set aside while the *N*-acylphenylenediamine-based oxidative coupling reaction was further pursued.

When the aminophenol-aniline reaction was revisited, the small molecule model com-

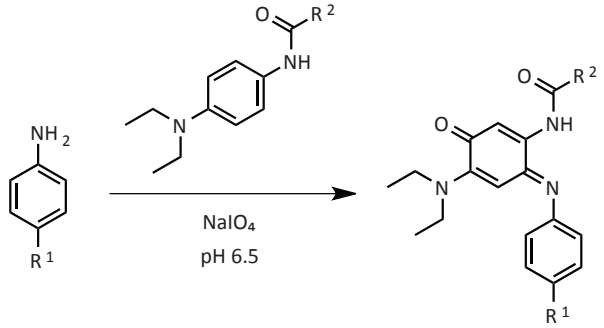
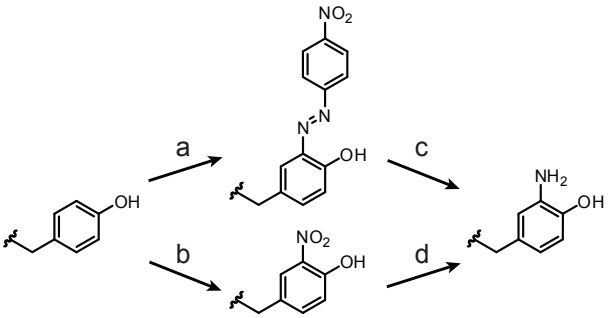
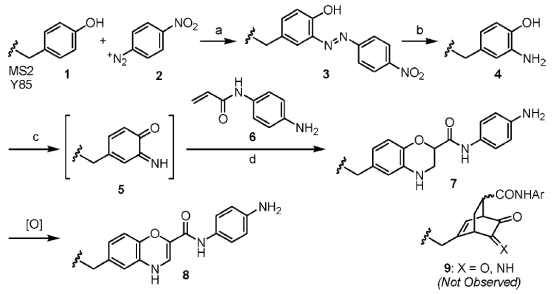



**Figure 2-1**. The previously reported coupling between anilines and phenylene diamines occurs within 30-60 min. Examples of $R^1$ and $R^2$ include proteins, peptides, and DNA aptamers.

**Scheme 2-1.** Conversion of Tyrosine to *ortho*-Aminotyrosine[a]




[a]Conditions: (a) 4-nitrobenzenediazonium tetrafluoroborate, pH 9, 4 °C, 15 min; (b) $C(NO_2)_4$, pH 8, 90 min; (c) $Na_2S_2O_4$, pH 7, 1 h; (d) $Na_2S_2O_4$, pH 7, 5 min.




[a] (a) 2 (5 equiv), pH 9, 4 °C, 15 min; (b) $Na_2S_2O_4$ (100 mM), pH 7.2, rt, 2 h, 80−85% protein recovery, 2 steps; (c) $NaIO_4$ (100 $\mu$M), pH 6.5, followed by (d) **6** (10 mM), 2 h, 75% protein recovery.

**Figure 2-2**. This scheme describes the diazonium/dithionite reaction developed by Hooker that created an aminophenol on tyrosine-85 of MS2. The phenylene diamine acrylamide singly modified the aminophenol after oxidation with periodate. The final product shown above was incorrectly characterized; the aniline, not the acrylamide, was the necessary functional group for reactivity. Nonetheless this observation was the first example of reactivity between aminophenols and anilines in the presence of sodium periodate. Figure from Hooker J.M., Esser-Kahn, A.P., Francis, M.B., *J. Am. Chem. Soc.*, **2006**, *128*, 15558-15559.

pound reaction was repeated. The same major product was isolated and it corresponded to the same NMR that could not be unambiguously identified (Figure 2-3a). This result led to an adjustment in the model compounds: 4-methyl-2-aminophenol was used instead. The aminophenol and aniline reactants were added at a 1:1 ratio with a ten-fold excess of sodium periodate in pH 6.5 phosphate buffer. The reaction was repeated with the aminophenol and aniline reactants at concentrations ranging from 100 μM to 5 mM. In all of these cases the reaction reached completion within several minutes. All of these conditions gave a major product that was clearly different from the *t*-butyl case, but as before it could not be identified with NMR (Figure 2-3b). To unambiguously confirm the structure of the isolated product, we turned to x-ray crystallography. The isolated product was dissolved in acetonitrile at a high concentration then diluted into toluene to a concentration of approximately 1 mg/mL. An open vial of this toluene solution was placed in a larger, sealed vial containing hexanes. After several days at room temperature, sufficient hexanes had diffused into the toluene solution to cause small crystals to form. X-ray diffraction of these crystals identified the major product (**1**) as the structure shown in Figure 2-3c (Crystallographic data in Chapter 2 Experimental Section). The structure matched the $^1$H NMR data; in particular, the newly formed chiral center explained the splitting of the methylene protons at δ 3.0-3.1 (Figure 2-3b). There was also evidence of a nitrile in both the $^{13}$C NMR spectrum (signal at 115 ppm) and IR spectrum (absorption band at 2218 cm$^{-1}$). The $^{13}$C NMR spectrum can be found in the experimental section.

Reversed-phase HPLC analysis of the crude reaction showed that the isolated product was the dominant species, but there was a small amount of a minor product (Figure 2-3d). When this reaction was applied to proteins as described in future sections, there was a much more significant amount of the minor product observed. Therefore we decided to further investigate the structure of this minor product. First, the minor product could not be converted to the major product by additional sodium periodate or through any other method. Work by Allie Obermeyer suggested that the minor product was formed through the same first step of aniline addition. However, the mass showed that there was no water addition; also the mass increased by two upon addition of tris-(2-carboxyethyl)phosphine (TCEP) and decreased back to its original mass once the TCEP was removed and the compound was exposed to air. Together these observations suggested a product that contained a quinone or iminoquinone. By obtaining the high resolution mass we determined that the product was the quinone, presumably formed from hydrolysis of the iminoquinone. NMR data, though difficult to interpret because of tautomerization and reduction potential, supported this conclusion and the structure (**2**) is shown in Figure 2-3e. Comparing these observations to our proposed mechanism was crucial towards determining the structure of the minor product.

My proposed mechanism for the formation of the two products is shown in Figure 2-4. The first step is oxidation of the aminophenol to an iminoquinone followed by addition of the aniline

in the position para to the imine. After tautomerization to regain aromaticity, the aniline-substituted aminophenol is then more electron rich than the aminophenol starting material. We would expect this species to oxidize immediately back to the iminoquinone. At this point there are two competing reactions that result in the two products. First, the imine can be hydrolyzed to form **2**.

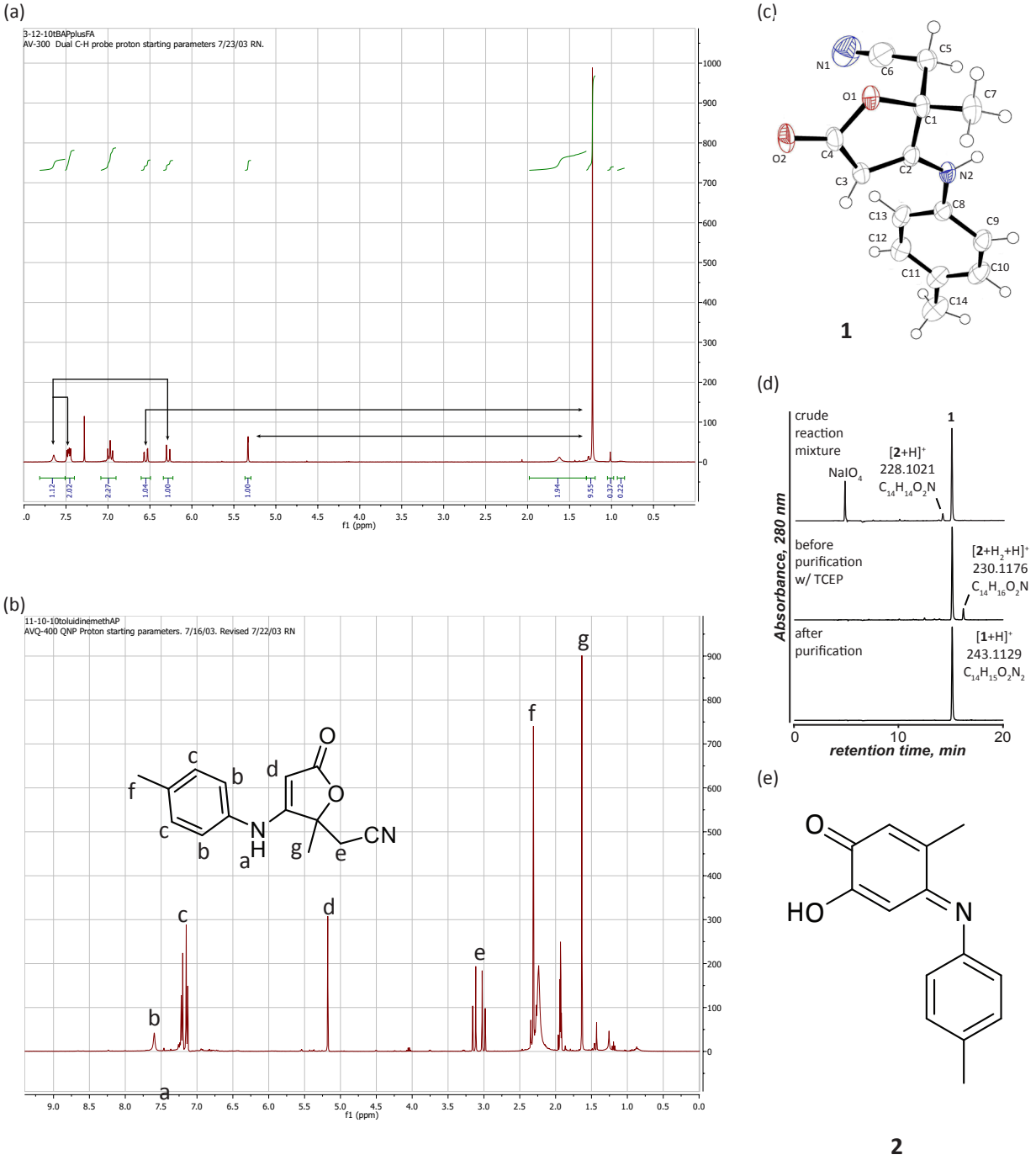



**Figure 2-3**. (a) $^{1}$H NMR spectra of the isolated major product of the reaction between 4-*tert*-butyl-2-aminophenol, *p*-toluidine, and sodium periodate. Arrows indicate the observed NOESY interactions. (b) $^{1}$H NMR spectra of **1**, the isolated major product of the reaction between 4-methyl-2-aminophenol, *p*-toluidine, and sodium periodate. (c) Product **1** was characterized using X-ray diffraction. (d) Reversed-phase HPLC analysis of the crude reaction with 4-methyl-2-aminophenol showed an uneven mixture of **1** and **2**. Addition of TCEP added two protons to **2** but **1** was unaffected. After TCEP was removed **2** rapidly re-oxidized upon exposure to air. (e) Based on the molecular formula and behavior in the presence of TCEP, **2** was the most likely structure of the minor product.

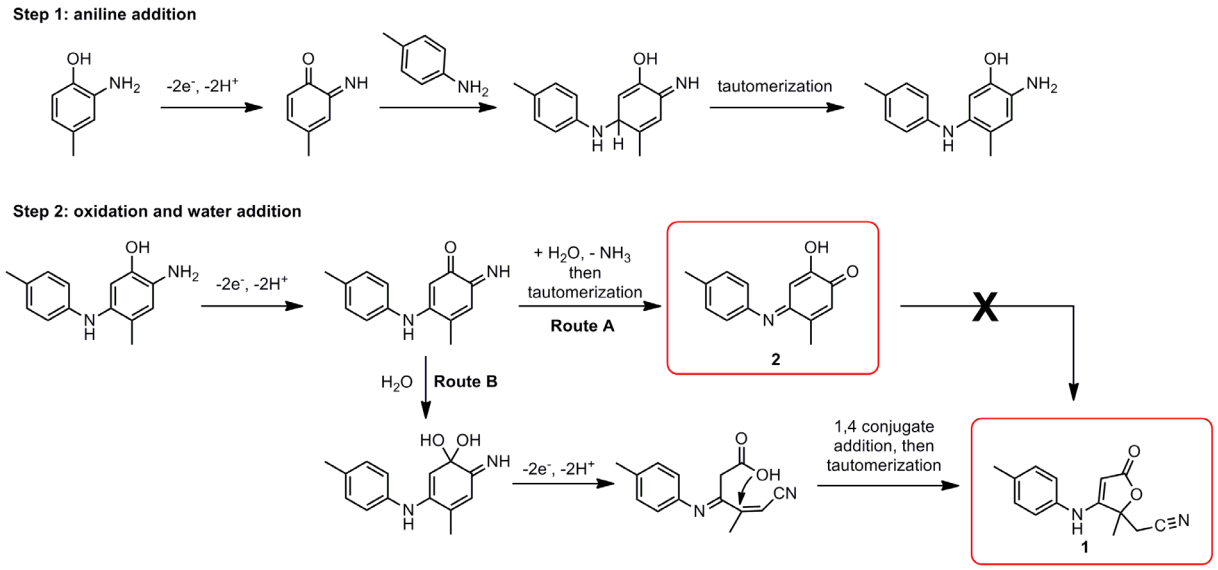



**Figure 2-4**. Proposed mechanism for aniline addition to aminophenol under oxidative conditions. Step 1 shows the aniline addition and step 2 shows two possible water addition routes that result in either product **1** or **2**. **2** cannot be converted to **1** once formed.

The periodate can no longer cleave the carbon-carbon bond between the carbonyls so **2** cannot be converted into **1**. The formation of **1** is slightly more complex. Water is added to the carbonyl to form a hydrate which can then form the five-membered ring with periodate necessary to cleave the carbon-carbon bond between the imine and carbonyl. The resulting carboxylic acid then undergoes a 1,4 conjugate addition to the acrylonitrile group followed by tautomerization to give **1**. The exact order of these steps is still under investigation.

After identifying the products and postulating a mechanism of this new reaction, we set forth to apply it to protein modification. It was first important to show that the reaction behaved similarly on protein substrates as it did in the small molecule studies. A cyclic, five amino acid peptide was used as a model peptide substrate. The tyrosine in the RGDyE (the lower case "y" denotes a D amino acid which makes cyclization more favorable) peptide was converted to an aminophenol by nitration with tetranitromethane followed by reduction with sodium dithionite (Scheme 2-1, Figure 2-5a-d). To this peptide *p*-toluidine was coupled in the presence of sodium periodate and analyzed by mass spectrometry. There were two distinct masses observed that correspond to the same mass changes observed for **1** and **2** (Figure 2-5e). Upon addition of TCEP, the peak at 740 gained two mass units while the peak at 755 remained unchanged. These changes were consistent with those of **1** and **2** in the presence of TCEP. Given the identical mass differences and similar behavior in the presence of TCEP, we can assume that the two products observed on the peptide substrate had the same structure as **1** and **2**. One key difference, however, was the relative amounts of the two products. To achieve full conversion to the modified product an excess of toluidine was added and that could have affected the product ratio. The only trend observed was that more electron-poor anilines resulted in a larger amount of the product analogous to **2**. Examples of this can be found in Chapter 3 on the topic of the CTT peptide.

To test this new chemistry for protein modification we wanted to mimic the small molecule studies but with either the aniline or aminophenol on a protein. Therefore we turned to MS2 viral capsids because of our group's familiarity with them. We have significant experience with recombinant expression of the MS2 coat protein in *E. coli*.[3,4] 180 copies of the coat protein monomer self-assemble into a spherical capsid approximately 28 nm in diameter. It has no genetic material,

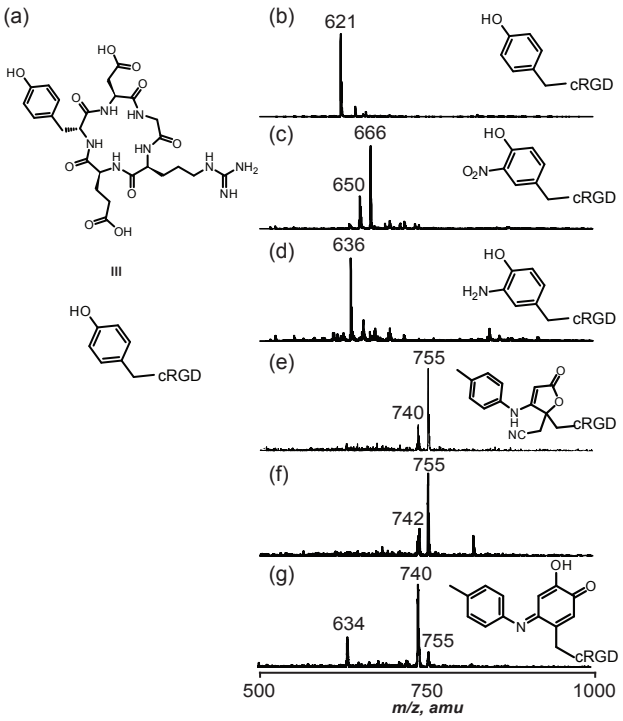



**Figure 2-5**. Modification of cyclic RGDyE peptide to aminophenol cyclic RGD. (a) Cyclic RGDyE and abbreviated structure. (b) MALDI-TOF MS of unmodified RGDyE. Expected: $[M+H]^+ = 621$. (c) MS after reaction with tetranitromethane at pH 7.8 for 1 h. Expected: $[M+H]^+ = 666$. The peak at 650 amu (loss of 16 from main peak) is an artifact always seen in the MALDI-TOF MS of *o*-nitrophenol substrates. (d) MS after reduction of nitro-RGD with sodium dithionite at pH 6.5 for 2 min. Expected: $[M+H]^+ = 636$. (e) LC-ESI-MS of coupling between cRGD-aminophenol and toluidine (200 μM) in the presence of sodium periodate (5 mM) at pH 6.5. Expected: $[M+H]^+ = 755$, $[m+H]^+ = 740$. (f) Addition of TCEP to sample e immediately prior to injection on LC-ESI-MS shows an addition of two mass units to the minor product, thus mimicking the behavior of **2**. Expected: $[M+H]^+ = 755$, $[m+H_2+H]^+ = 742$. (g) MS of coupling between cRGD-aminophenol and toluidine (200 μ M) at pH 6.5 in the presence of a reaction-limiting concentration of sodium periodate. The mass at 740 m/z corresponds to the mass change of **2**.

so the inside of the sphere is hollow; therefore, both the inner and outer surface can be modified. Access to the inner surface is achieved by diffusion through the 1-2 nm diameter pores of the capsid.[5] To utilize this new reaction, it was necessary to incorporate anilines or aminophenols into the protein. There is one particularly solvent accessible tyrosine in the native sequence. Tyrosine-85 can be selectively modified with diazonium salts.[3] On the outer surface residue 19 was found to be accommodating to several mutations and sufficiently solvent accessible for chemical modification. Threonine-19 was mutated to the amber stop codon to facilitate incorporation of *para*-aminophenylalanine (T19*p*AF MS2) using a tRNA/aminoacyl tRNA synthetase pair developed by the Schultz laboratory.[4]

First, wildtype MS2 was chemically modified to produce an *ortho*-aminophenol at position 85. There are two possible methods for this chemical transformation – they are illustrated in Scheme 2-1. For MS2 it was found that the diazonium salt route gave more singly modified protein than the nitration route. The sodium dithionite conditions are similar for the two methods, but the azobenzene functional group requires a longer reaction time to reduce than the *ortho*-nitrophenol. Modification at each step was analyzed with MALDI-TOF MS (Figure 2-6b-d). MS2 with aminophenol functionality at position 85 was reacted with *p*-toluidine and sodium periodate to give the oxidative coupling product (Figure 2-6e). Unfortunately, due to side reactions caused by the diazonium salt (possibly deaminations of lysine) the peak in the MS spectrum was significantly broadened. Due to this broadening it is not possible to determine the relative ratio of the two oxidative coupling products. However, we can conclude that the reaction reached full conversion to the MS2+toluidine product in two minutes.

To further investigate the oxidative coupling of small molecules to MS2, we instead turned to the T19*p*AF MS2 capsids because they required no chemical modification to prepare them for oxidative coupling. 4-methyl-2-aminophenol was coupled to T19*p*AF MS2 in the presence of sodium periodate at pH 6.5. In less than 5 min there was quantitative conversion to the MS2+aminophenol product. To further investigate the rate of reaction a time course assay was designed. At certain

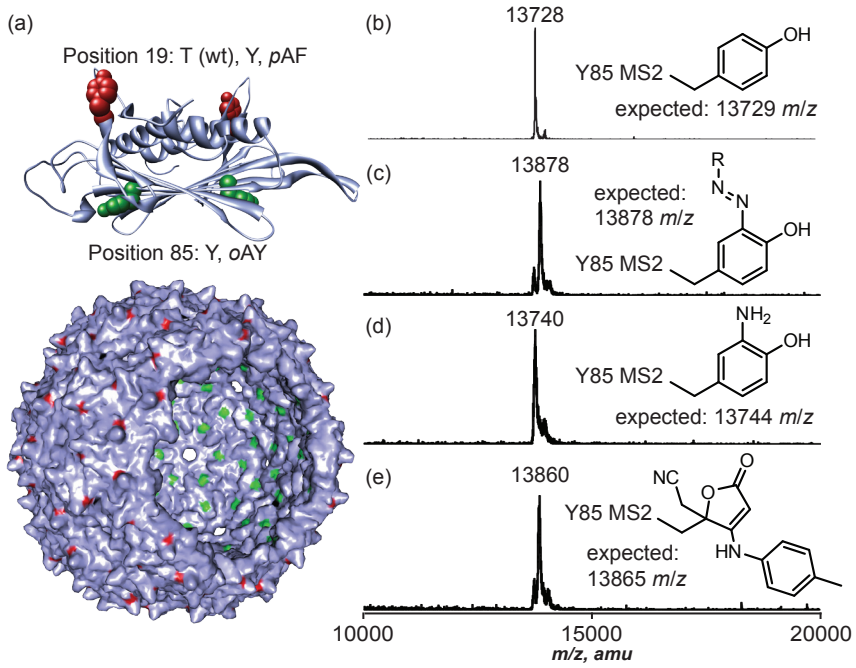



**Figure 2-6**. (a) Structure of the MS2 viral capsid (based on wt structure PDBID: 2MS2) indicating the locations of position 19 (outer surface) and 85 (inner surface). Following chemical modification, the capsids were dissociated and analyzed using MALDI-TOF MS: (b) wtMS2, (c) wtMS2 after reaction with 4-nitrobenzenediazonium tetrafluoroborate, (d) azo-modified MS2 after reduction with sodium dithionite, and (e) MS2-*o*-aminotyrosine-85 coupled to *p*-toluidine in the presence of 5 mM periodate (RT, 2 min).

time points a portion of the reaction mixture was removed and quenched with TCEP, then analyzed with MALDI-TOF MS. The reaction was near completion in 15 s and was complete at 30 s (Figure 2-7). There was no additional modification even after 20 min. As a control, T19Y MS2 (one atom difference from T19*p*AF MS2) was reacted under the same conditions and showed no modification after 20 min. The broadening of the peaks in the MS spectrum was due to the inability of the MALDI-TOF MS to differentiate between the two expected products. Therefore a sample was submitted for high-resolution MS and it showed an approximately 2:1 ratio of the ring-contracted product to the quinone product (Figure 2-8).

In conclusion, a new oxidative coupling reaction between *ortho*-aminophenols and anilines has been optimized and characterized. The two products have been structurally identified and fit into a plausible mechanism using small molecule studies. The reaction is also capable of modifying proteins with extremely high efficiency. At concentrations in the low micromolar, aniline-containing MS2 coat protein monomers were modified with 4-methyl-2-aminophenol to give quantitative conversion in 30 s. This reaction is not only fast, but also site-specific. Changing the aniline to a tyrosine (a one atom change) makes the protein unreactive towards the aminophenol. The reaction proceeds at pH 6.5 at room temperature, thus making it ideal for protein modification. In the next section we will use this reaction to couple more challenging substrates, such as two macromolecules, in order to make useful bioconjugates.

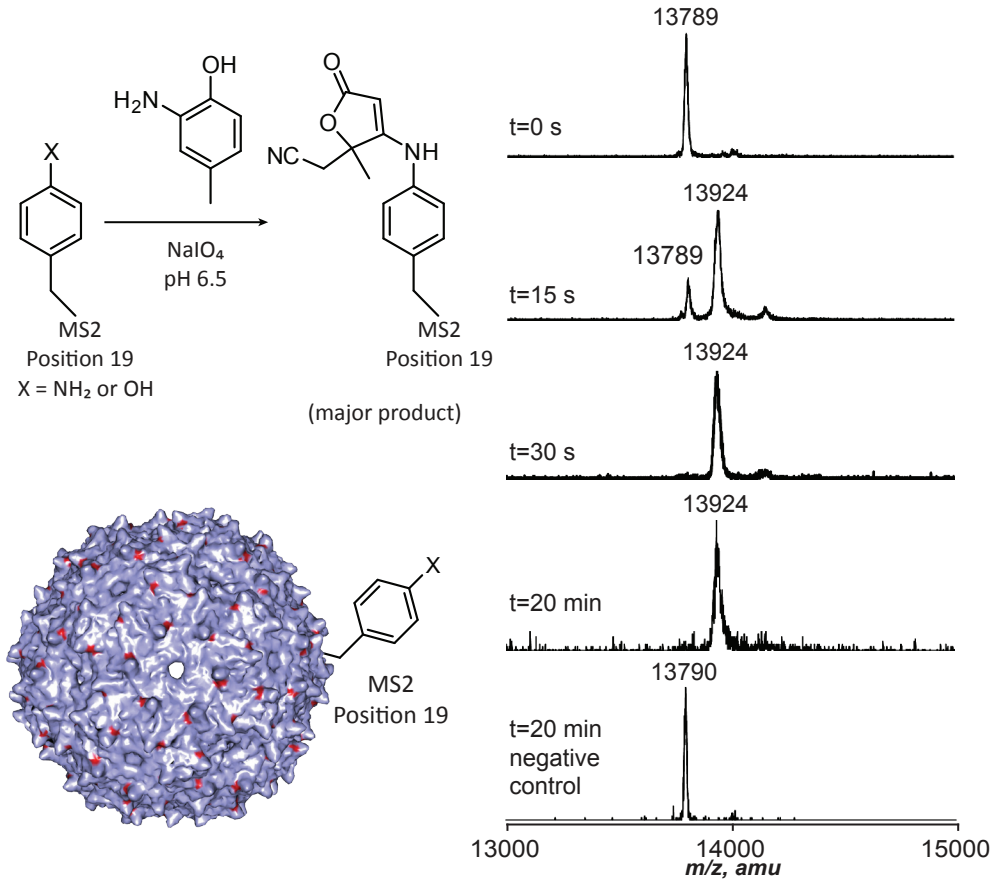



**Figure 2-7**. Targeting an artificial amino acid through oxidative coupling. T19*p*AF MS2 (X = $NH_2$, 30 μM) was reacted with 2-amino-4-methylphenol (100 μM) in the presence of sodium periodate (1 mM). Aliquots were removed at various time points and quenched by the addition of tris(2-carboxyethyl)phosphine (TCEP). Single protein modification was complete by 30 s. High-resolution ESI-MS analysis revealed that the major product's structure corresponded to **1**. In addition there was a significant minor product corresponding to addition product **2**. These signals do not resolve using MALDI-MS (see Figure 2-8 for high-resolution ESI spectra). No additional modification was detected after 20 min. A negative control reaction using T19Y MS2 (X = OH) showed no modification under otherwise identical conditions.

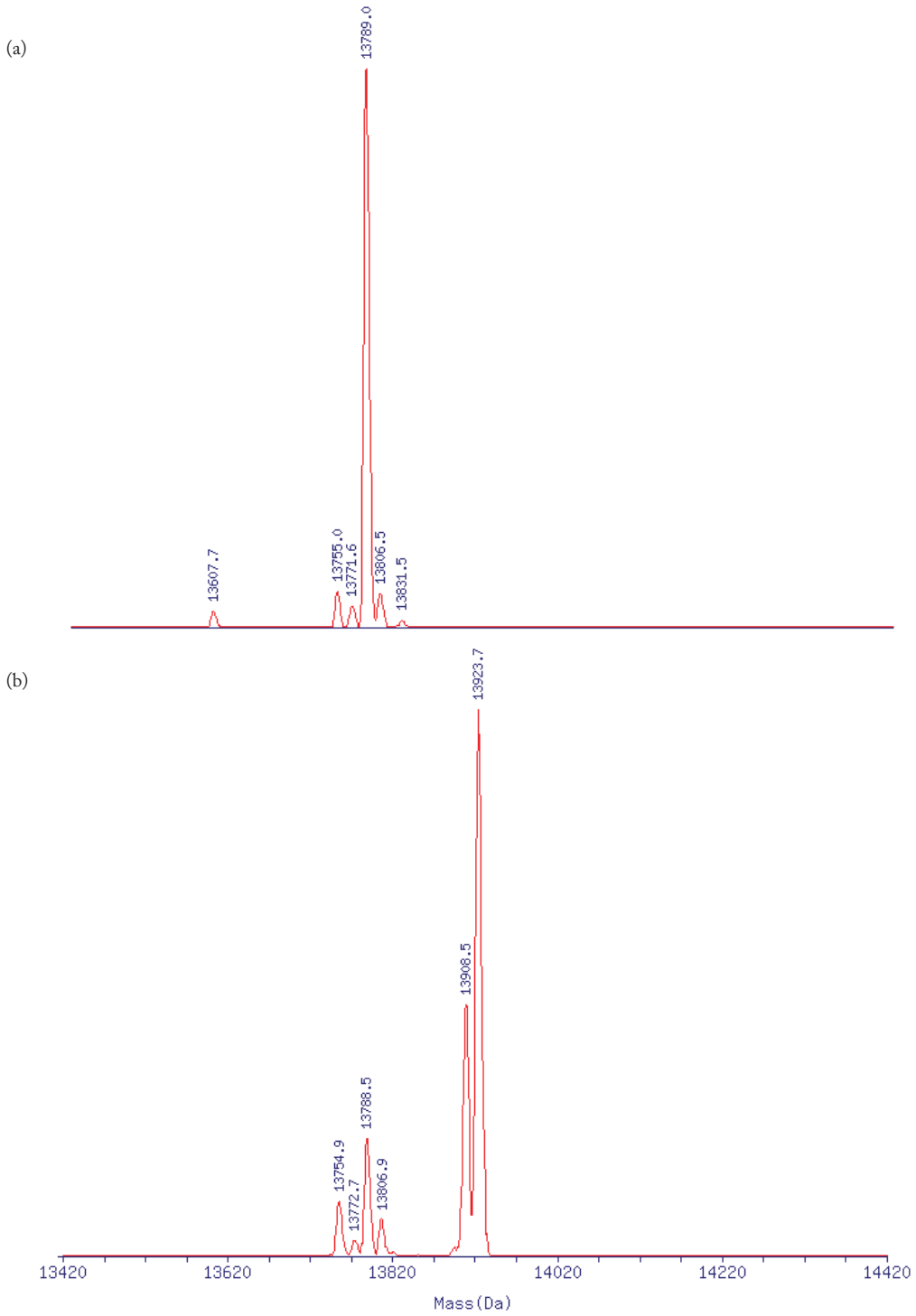



**Figure 2-8**. (a) Unmodified T19*p*AF MS2 capsids were disassembled into monomers and analyzed via ESI-MS. (b) T19*p*AF MS2 (30 μM) was reacted with 2-amino-4-methylphenol (100 μM) in the presence of $NaIO_4$ (1 mM) for 30 s before quenching with 50 mM tris(2-carboxyethyl)phosphine (TCEP). The product was purified with size exclusion chromatgraphy, the capsids were disassembled into monomers, and the protein was analyzed via ESI-MS. Note that the protein was exposed to air after removal of TCEP so the minor product corresponding to **2** is in the oxidized state in this mass spectrum.

## Section 2.2: Macromolecule Bioconjugates via Oxidative Coupling

Because of the high efficiency and mild conditions of this new reaction, it is particularly well suited for assembling complex protein bioconjugates. As mentioned in Chapter 1 there are several bioconjugation cases that require a high efficiency reaction. The simplest case to test is covalently linking two sterically hindered substrates. Because of its common use in the pharmaceutical industry to reduce immunogenicity of proteins,[6] polyethylene glycol (PEG) was selected as a test compound.

The synthesis of PEG molecules functionalized with aminophenols or anilines began with 2k or 5k molecular weight PEG with a methyl group at one end and a primary amine at the other. The PEG-aminophenol was synthesized by coupling the amine to an activated ester containing an *o*-nitrophenol functionality. After purifying away the PEG from any remaining small molecules, the nitrophenol was reduced to the aminophenol with sodium dithionite. The PEG-aniline was synthesized through a similar procedure but instead coupled an activated ester with *N*-Boc-aniline functionality. Deprotection of the Boc group with trifluoroacetic acid afforded the PEG-aniline. Detailed synthetic procedures for these molecules can be found in the Chapter 2 Experimental Section.

To first test the PEG coupling on a common protein, lysozyme was chemically modified with anilines. 3-(4-aminophenyl)propionic acid was converted to an activated ester using 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide and *N*-hydroxysuccinimide in DMF. Without purification, the activated ester was added to a solution of lysozyme. As shown in the MS spectrum in Figure 2-9b, several residues (presumably lysines) were modified to have aniline functionality. The lysozyme-aniline was then reacted with PEG2k-aminophenol in the presence of sodium periodate for two minutes. The reaction was quenched with gel loading buffer containing excess dithiothreitol (DTT) and the resulting protein conjugate was analyzed via SDS-PAGE. The lane with all three components (aniline, aminophenol, and periodate) shows between one to four PEG molecules attached to each lysozyme (Figure 2-9c). The negative controls showed no significant modification.

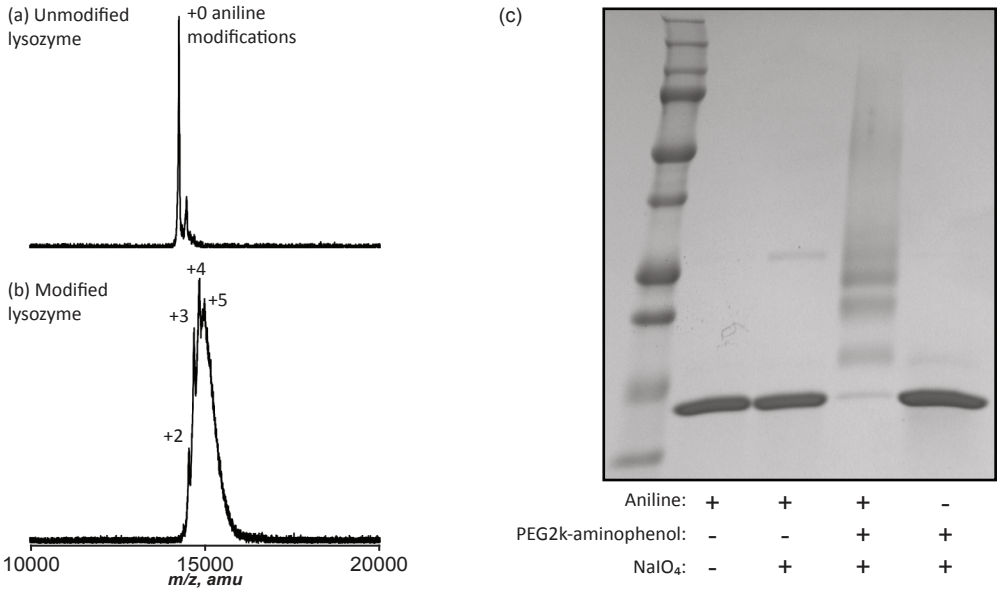



**Figure 2-9**. (a) MALDI-TOF of wildtype lysozyme. (b) MALDI-TOF of lysozyme modified with 3-(4-aminophenyl)propionic acid showed 2-5 aniline modifications. (c) SDS-PAGE analysis of lysozyme (20 μM) reacted with PEG2k-aminophenol (100 μM) in the presence of $NaIO_4$ (2 mM) for 2 min. The reaction only occurred if both the aniline and the aminophenol were present.

Once the conditions for PEG modification of lysozyme were optimized, we turned back to MS2 as a target for PEG modification. As mentioned previously, MS2 is a modular scaffold for many biological applications in our lab. It is important to be able to attach cargo to the inner surface and targeting moieties to the outer surface. Therefore, using this new chemistry as a means to attach a variety of macromolecules to the outside of MS2 is a clear goal with many useful applications. As in the lysozyme case, PEG macromolecules were used to simulate attaching large targeting moieties.

The reaction between MS2-aminophenol-85 and *p*-toluidine as described above was repeated but with PEG2k-aniline replacing toluidine. As shown in Figure 2-10c, Lane 7, there was no modification of the aminophenol at position 85. We hypothesized that the PEG was too large to diffuse through the 1-2 nm pores of the capsid and therefore could not react with residues on the inner surface. To test this idea, T19Y MS2 was chemically modified using the diazonium salt/sodium dithionite procedure described above. The resulting MS2 with *o*-aminophenols at positions 19 and 85 was singly modified with PEG2k-aniline (Figure 2-10c, Lane 5), thereby confirming our hypothesis that the PEG could not access the inner surface of the capsid. The experiment was repeated with PEG5k-aniline and similar results were observed except with a lower amount of conversion to the PEGylated product (Figure 2-10c, Lane 4). PEG5k-aniline addition resulted in 40% PEGylation while PEG2k-aniline resulted in 50% conversion. The lower level of conversion for the 5k PEG suggested that the incomplete PEGylation was due to steric constraints around the surface of the MS2 capsid.

Similar PEGylation was possible with the reactive groups swapped between the substrates. T19*p*AF MS2 was reacted with PEG-aminophenol under the same conditions as the previous PEG experiments. Both 5k and 2k PEG showed excellent conversion, with approximately 65% modified in the 5k case and 75% for the 2k (Figure 2-10c, Lanes 1-2), while the negative control with T19Y MS2 showed no modification (Figure 2-10c, Lane 3). These numbers are both higher than the case with the aminophenol on the MS2, likely because the chemical modification of tyrosine to aminophenol is incomplete whereas the incorporation of *p*-aminophenylalanine is close to 100%. When also considering the diazonium salt deamination issue mentioned previously, we feel that the best method to modify MS2 capsids is to use T19*p*AF MS2 and attach substrates with aminophenol functionality.

With a fast and reproducible gel shift assay worked out, it was then possible to assay different conditions. First, the reaction between T19*p*AF MS2 and PEG5k-aminophenol was repeated in phosphate buffer with a pH range from 5.0 to 9.0. The results were identical at each pH within the range. Next the compatibility of this reaction with glycoproteins was investigated. Sodium periodate can also be used as an oxidant for the cleavage of vicinal diols in sugars. Therefore, we repeated the coupling of PEG5k-aminophenol with T19*p*AF MS2 in the presence of varying amounts of glucose. We found no significant difference in coupling between samples with 0 or 20 mM glucose (Figure 2-11). Though this experiment did not analyze the state of the vicinal diols, it did show that small amounts of sugars will not consume enough oxidative equivalents of periodate to prevent the aminophenol-aniline coupling.

After showing that PEG could be successfully attached to MS2, we also wanted to try attaching peptide substrates. The first peptide substrate used was the cyclic RGD peptide described previously. The tyrosine of the peptide was chemically modified with tetranitromethane followed by sodium dithionite to give the aminophenol. In this case it was difficult to purify the peptide after the reduction step because it was not retained by any resin particularly well. Instead, after

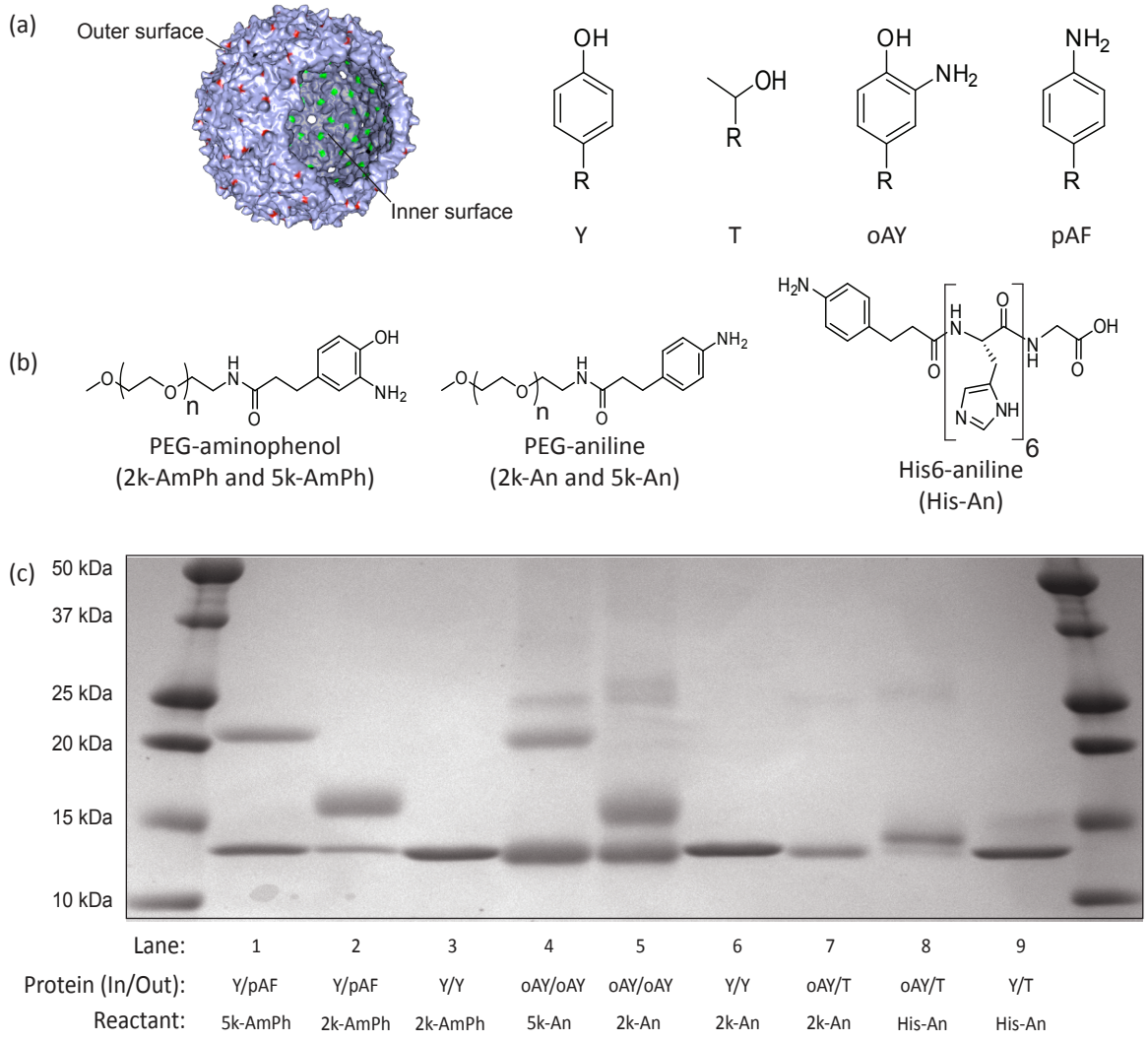



**Figure 2-10**. A panel of reactants was exposed to MS2 capsids containing one of the four reactive groups listed in (a). Each capsid sample contained one group on its inner surface and one on its outer surface. Each reaction contained 30 µM protein, a coupling partner selected from (b) at 200-500 µM, and $NaIO_4$ (1 mM for lanes 1-7, 5 mM for lanes 8 and 9). Each reaction time was 2 min, after which the samples were quenched with gel loading buffer containing excess DTT and analyzed via SDS-PAGE.

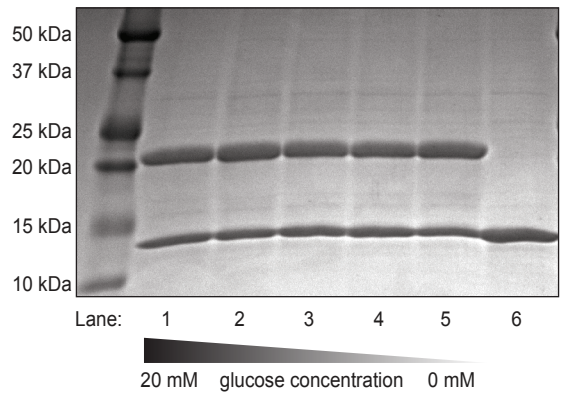



**Figure 2-11**. Attachment of PEG aminophenol (5k molecular weight) to T19*p*AF MS2 in the presence of varying amounts of glucose. PEG5k-aminophenol (200 µM) was exposed to T19*p*AF MS2 (30 µM) for 5 min in the presence of sodium periodate (5 mM) and 20 mM, 5 mM, 2 mM, or 200 µM glucose (Lanes 1-4, respectively). None show an appreciable difference from Lane 5 which had no glucose. Lane 6 is a negative control with 5 mM sodium periodate and 5 mM glucose but no PEG5k-aminophenol.

reduction the peptide was added directly to T19*p*AF MS2 and extra sodium periodate was added to quench the remaining dithionite. SDS-PAGE analysis of the MS2-cRGD conjugate showed that approximately 75% of the MS2 monomers were modified with the peptide in only 2 min (Figure 2-12).

During her rotation working under my supervision, Abby Knight synthesized another biologically relevant peptide for attachment to the outer surface of MS2. Bombesin is a small peptide that binds the gastrin-releasing peptide receptor and this receptor has been found upregulated in breast and prostate cancer.[7] The active portion of the peptide (QWAVGHLM) was synthesized using standard Fmoc chemistry on solid phase beads using a Wang linker. To the N-terminus was attached a 6 carbon linker (6-aminocaproic acid), and that amine was capped with 3-(4-hydroxy-3-nitrophenyl)propionic acid. After reduction to the aminophenol using sodium dithionite, the peptide was reacted with T19*p*AF MS2 in the presence of sodium periodate. SDS-PAGE analysis of the conjugate showed a small gel shift and 20-40% modification depending on the concentration of peptide added (Figure 2-13a). MALDI-TOF MS also showed the desired peak (Figure 2-13c).

To test for attachment of a peptide to MS2 with the reactive groups swapped, a seven amino acid test peptide was synthesized on solid phase beads using a Wang linker and standard Fmoc chemistry. The N-terminal amine was capped with an activated ester containing *N*-Boc-aniline functionality similar to the PEG-aniline synthesis described above. Upon cleavage of the peptide from the resin using trifluoroacetic acid the Boc group was liberated giving a peptide-aniline (Figure 2-14b). This aniline was coupled to MS2-aminophenol-85 using standard oxidative coupling conditions. The construct was purified using a Nap5 size exclusion column then analyzed using SDS-PAGE (Figure 2-10c, Lane 8). There is a small but noticeable gel shift corresponding to the MS2-peptide conjugate. Because the peptide was smaller than the PEG2k-aniline, it was able to

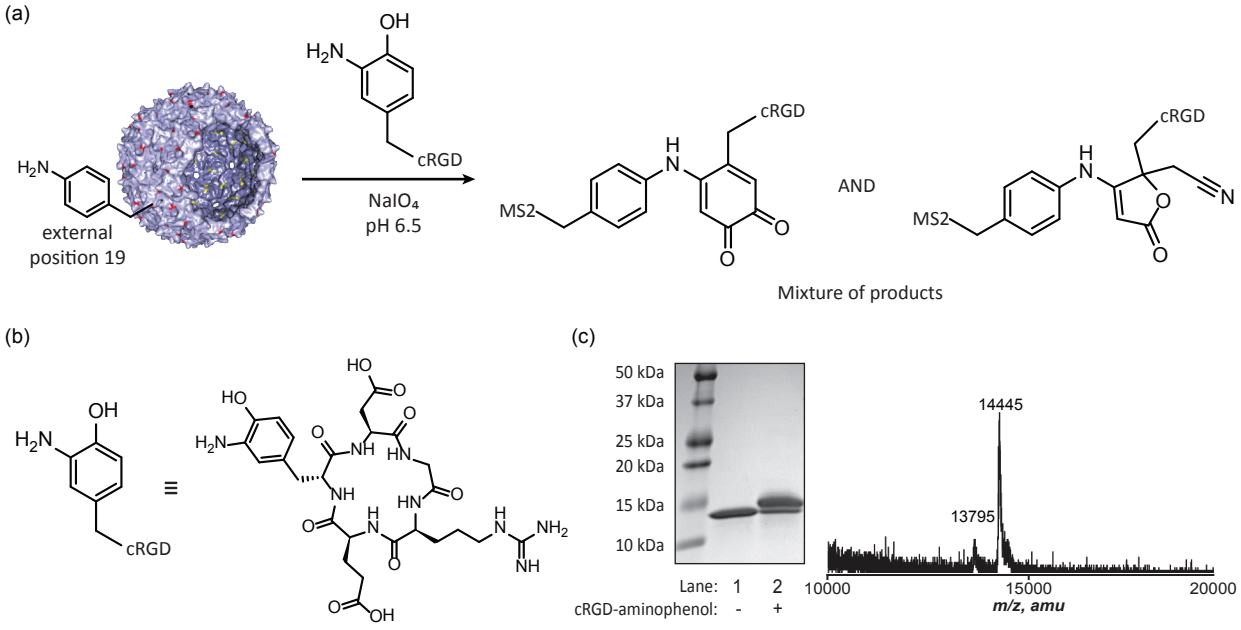



**Figure 2-12**. Attachment of a cyclic RGD peptide to T19pAF MS2. (a) cRGD-aminophenol (200 µM) was exposed to T19*p*AF MS2 (30 µM) in the presence of sodium periodate (5 mM) for 2 min. (b) The full structure of cRGD-aminophenol is shown. (c) MS2 capsids were disassembled into monomers for analysis via SDS-PAGE and MALDI-TOF. Expected masses are m/z 13795 for T19*p*AF MS2 and m/z 14443 for MS2+cRGD-aminophenol.

diffuse through the capsid pores and modify the inner surface at position 85 while the PEG-aniline could not (Figure 2-10c, Lane 7). Identical reaction conditions between the peptide and wildtype MS2 showed no modification (Figure 2-10c, Lane 9). Figure 2-14c shows a MALDI-TOF MS spectrum of the MS2+$His_6$ conjugate.

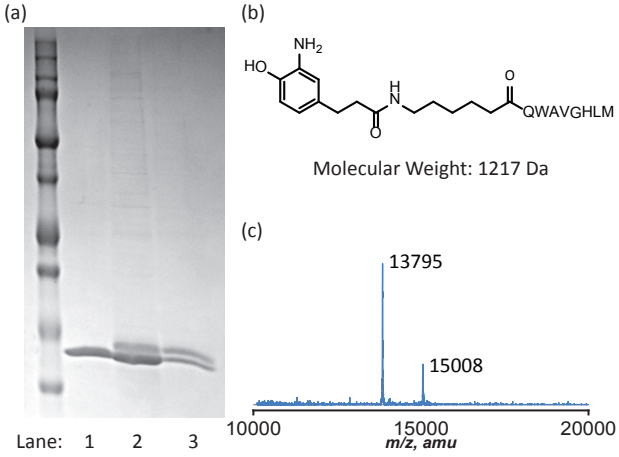



**Figure 2-13**. (a) Bombesin-aminophenol was coupled to T19pAF MS2, then analyzed via SDS-PAGE. Lane 1 is unreacted T19pAF. Lanes 2 and 3 are MS2+bombesin at 20% and 40% modification, respectively. Differing levels of modification are achieved by varying the concentration of bombesin-aminophenol. (b) Structure of bombesin-aminophenol. (c) MALDI-TOF MS of sample from Lane 2 in (a) cospotted with sinapinic acid matrix.

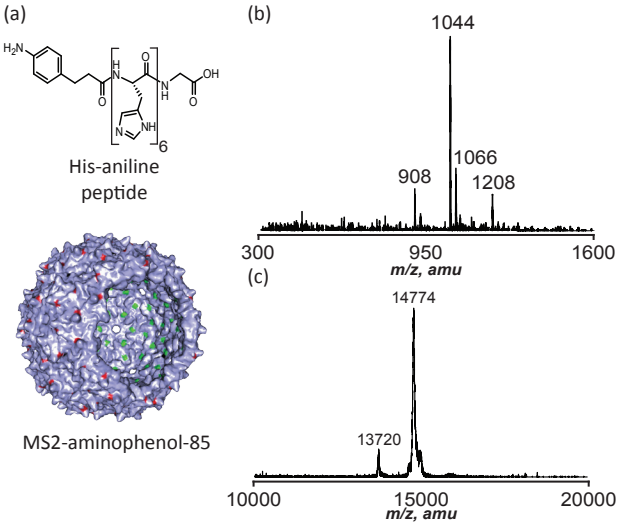



**Figure 2-14**. (a) His-aniline peptide and MS2-*o*-aminophenol-85 with 85 position highlighted. (b) MALDI-TOF of His-aniline peptide after cleavage from resin. Expected: $[M+H]^+$ = 1045, $[M+Na]^+$ = 1067 m/z. (c) MS2-*o*-aminophenol-85 (50 µM) was reacted with His-aniline peptide (500 µM) in the presence of sodium periodate (5 mM) for 2 minutes before purification by size exclusion chromatography (NAP-5 column). SDS-PAGE analysis of this sample appears in Figure 2-10c, lane 8.

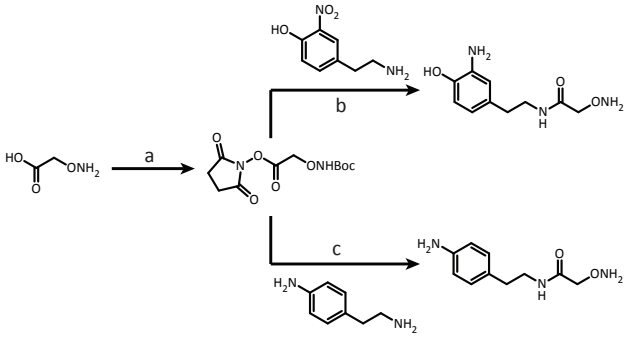
## Section 2.3: Filamentous Phage Modification with Oxidative Coupling

This oxidative coupling method has been successfully utilized on several proteins by other members of the lab and by myself, but one in particular showed interesting properties. M13 or fd phage are commonly used bacteriophage in phage display library experiments.[8,9] They are composed of several proteins but the two of main importance are pIII and pVIII. Five copies of the pIII protein are found at one end of the pencil-shaped phage, and the library members are fused to this protein for phage display panning. The pVIII protein is the coat protein composing the shaft of the phage and the number of copies varies based upon the length of the phage genome. The phage commonly used have approximately 2000-4000 copies of the pVIII protein, thus making it an appealing protein to target for chemical modification of the virus. Work by Zac Carrico in the Francis group (unpublished as of 4/8/2012) resulted in a method for the selective transamination of the N-terminus of pVIII but not pIII using modification with pyridoxal-5-phosphate.[10,11,12] The resulting ketone could then be site-specifically modified using alkoxyamino compounds.

To attach oxidative coupling reactive groups the the pVIII ketone, alkoxyamine linkers were necessary. Scheme 2-2 details the synthesis of an aniline-alkoxyamine and aminophenol-alkoxyamine. Briefly, aniline or nitrophenol amines were reacted with the succinimidyl ester of (Boc-aminooxy)acetic acid followed by deprotection with trifluoroacetic acid. After careful neutralization of the residual TFA from the deprotection step, the alkoxyamines were added to the transaminated phage in as large excess as possible, usually 200-1000 fold excess. The oxime formation reaction[13] is fastest at pH 4.5 but due to concerns about phage solubility the reaction was performed between pH 5 and 6. The constructs were purified by Nap-5 (the concentration of phage should be <100 μM to avoid loss). MALDI-TOF MS was used to characterize the modification (Figure 2-15).

Characterizing the PEGylation of pVIII was more difficult than the previous proteins. The pVIII protein is only 5 kDa and because of its amphiphilic properties it is not easily analyzed with

**Scheme 2-2.** Synthesis of Aminophenol and Aniline Alkoxyamines[a]




[a]Conditions: (a) 1. $Boc_2O$, $Et_3N$; 2. DCC, NHS; (b) 1. 2-(4-hydroxy-3-nitrophenyl)ethylamine, $Et_3N$; 2. Pd/C, $H_2$; 3. TFA; (c) 2-(4-aminophenyl)ethylamine, $Et_3N$; 2. TFA

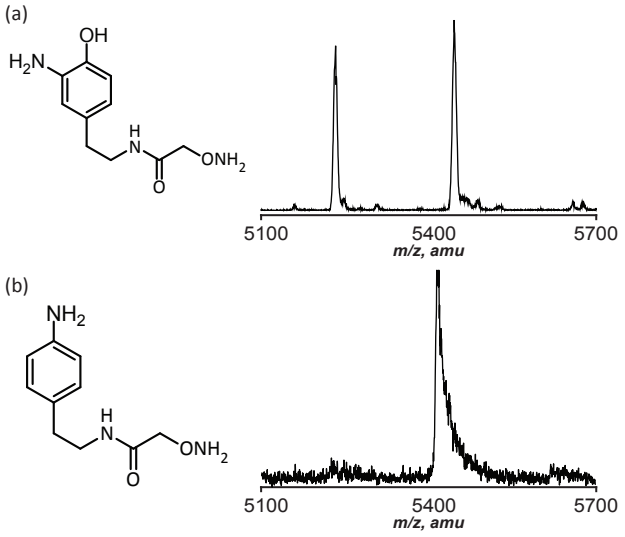



**Figure 2-15**. The pVIII proteins of M13 phage were transaminated with pyridoxal-5-phosphate (PLP) and then reacted with 200-1000 fold excesses of either aminophenol-alkoxyamine (a) or aniline-alkoxyamine (b) at pH 5-6 for 24 hr. Longer reaction times will result in greater amounts of modification. After purification with Nap-5 the level of modification was determined by MALDI-TOF MS.

SDS-PAGE. Instead reversed-phase HPLC was used to analyze the PEGylation reactions. In the method used, the pVIII protein eluted at around 10 min and because it is so much more abundant than the other phage proteins it is the only detected protein in the chromatogram. Transamination followed by modification with aminophenol or aniline alkoxyamines only slightly broadened the elution profile of the pVIII (Figure 2-16a,c). As a negative control, sodium periodate was added to these phage samples without any oxidative coupling partner. As shown in Figure 2-16b, the phage-aminophenol sample showed a later eluting peak suggesting that an undesired side reaction occurred, possibly pVIII crosslinking. This result strongly suggested that the phage-aminophenol construct was not a suitable substrate for the oxidative coupling reaction. However, the phage-aniline reaction showed no such behavior, as the elution profile looked identical upon addition of sodium periodate (Figure 2-16d). Therefore, the phage-aniline was reacted with PEG5k-aminophenol and periodate for 2 min, then was Nap-5 purified and spin concentrated. HPLC analysis

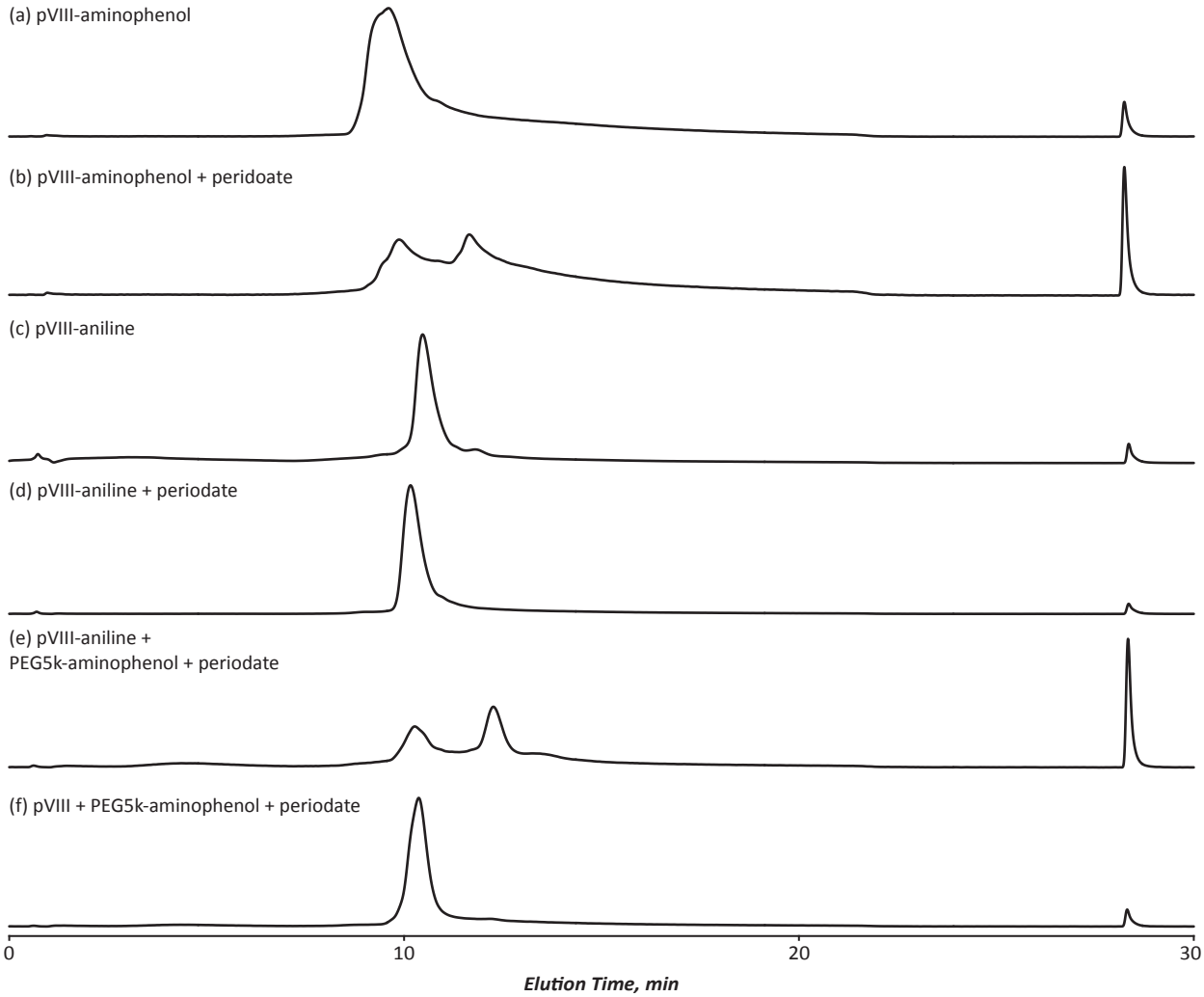



**Figure 2-16**. HPLC analysis of pVIII coat protein from fd or M13 bacteriophage. 10 µL injections on Zorbax-CN reversed-phase column running $H_2O$:MeCN + 0.1% TFA gradient. All samples were Nap-5 purified to quench reaction and spin concentrated to approximately 100 µM. Y-axis is fluorescence (Ex. 280 nm, 330 nm detection). The peak at approximately 28 min is present in all runs, even blank runs, and is likely caused by a sudden change in the gradient. (a) Transaminated pVIII reacted with aminophenol-alkoxyamine. (b) Sample from a exposed to sodium periodate for 2 min. (c) Transaminated pVIII reacted with aniline-alkoxyamine. (d) Sample from c exposed to sodium periodate for 2 min. (e) pVIII-aniline was reacted with PEG5k-aminophenol in the presence of sodium periodate for 2 min. (f) Negative control of e, transaminated pVIII with no aniline was reacted with PEG5k-aminophenol in the presence of sodium periodate for 2 min.

showed formation of a later eluting peak, suggesting that the PEG5k had been attached to a large portion of the pVIII proteins (Figure 2-16e). A second negative control where the phage-aniline and PEG5k-aminophenol were mixed without periodate showed no new peak (Figure 2-16f), thus confirming that the new peak was both PEG5k-aminophenol and periodate dependant.

Although there are simpler methods for attaching PEG to phage, namely direct attachment of PEG-alkoxyamines, these experiments have shown that oxidative coupling can also be used. As mentioned in Chapter 1, there are significant benefits to a reaction as efficient as the one described in this chapter. While PEG is very water soluble, there are many less soluble substrates that cannot be added in a large excess to the protein substrate. This idea was tested by the graduate students Kathryn Strobel and Jake Jaffe. They wanted to attach a somewhat insoluble dye to the pVIII protein but it was not soluble enough to be added in the large excess necessary to drive oxime formation. They attempted both versions of the oxidative coupling described in the previous paragraphs, and found that the phage-aniline plus aminophenol substrate worked well while the opposite reaction failed. These results corroborate the PEG experiments shown in Figure 2-16. Their experiments also support the reasoning behind why very efficient reactions like this oxidative coupling reaction are an important addition to the protein bioconjugation toolbox.

## Section 2.4: Experimental Section

**Aminophenol/Aniline coupling small molecule model reaction.**

*Small molecule coupling reaction between* p-*toluidine and 2-amino-4-methylphenol*

2-Amino-4-methylphenol (123 mg, 1 mmol) and *p*-toluidine (107 mg, 1 mmol) were dissolved in two separate portions of acetonitrile (2 mL each) and then added to 2 L of 10 mM phosphate buffer, pH 6.5. Sodium periodate (2.14 g, 10 mmol) was dissolved in 20 mL of phosphate buffer and then added in one portion with vigorous stirring. The solution immediately turned red. After 5 min, sodium chloride was added to 1 M and the solution was extracted with 1 L portions of ethyl acetate until the organic layers showed no more color. The combined organic layers were washed with brine and dried over sodium sulfate before removing the solvent under reduced pressure. The resulting solid was purified via silica gel chromatography using a mobile phase of 1:1 ethyl acetate:hexanes ($R_f \approx 0.25$) to afford 95 mg of 1 as a gray powder (40% yield). ESI-MS: 242.7 (242.27 expected). $^{1}$H NMR (300 MHz, $CD_3CN$) δ 1.6 (s, 3H), 2.3 (s, 3H), 3.0-3.1 (dd, 2H), 5.2 (s, 1H), 7.1 (d, 2H), 7.2 (d, 2H), 7.5 (br s, 1H). $^{13}$C NMR (100 MHz, $CD_3CN$) δ 19.82, 22.68, 27.43, 80.12, 83.86, 115.64, 120.53, 129.93, 134.61, 136.95, 166.34, 171.83. $^{1}$H spectrum can be found in Figure 2-3b and $^{13}$C NMR spectrum can be found at the end of the experimental section. IR: 3250, 3048, 2218, 1713, 1616 $cm^{-1}$.

*Small molecule coupling reaction between p-toluidine and 2-hydroxy-4-methylphenol*

In the above reaction, byproduct **2** was not produced in sufficient yield for characterization using NMR. Instead the reaction between *p*-toluidine and 2-hydroxy-4-methylphenol was used to produce **2** directly. Co-injection on reversed phase HPLC showed the major product of this reaction matched the **2** byproduct observed in the above reaction. The mass spectra of the two compounds were identical, and both compounds exhibited the same gain of 2 Da when exposed to TCEP. The synthesis of **2** was achieved as follows:

2-hydroxy-4-methylphenol (12 mg, 0.097 mmol) and *p*-toluidine (9 mg, 0.084 mmol) were dissolved in acetonitrile (0.2 mL) and then diluted with 6 mL of water. Sodium periodate (52 mg, 0.24 mmol) was dissolved in 2 mL of water and then added to the stirring solution in one portion. The solution immediately turned a deep purple. After 5 minutes, the water was removed under reduced pressure. The resulting red solid was taken up directly in $CDCl_3$ for analysis. IR (thin film): 3268, 3034, 2923, 2854, 1712, 1674, 1637, 1608, 1515 $cm^{-1}$. $^{1}$H NMR (600 MHz, $CDCl_3$) δ 7.19 (d, 2H), 6.77 (br s, 2H), 6.61 (br s, 1H), 6.34 (br s, 1H), 4.29 (br s, 1H), 2.37 (s, 3H), 2.34 (d, 3H). 13C NMR (150 MHz, $CDCl_3$) δ 18.46, 21.13, 101.09, 120.53, 127.50, 129.72, 135.52, 147.87, 153.14, 158.73, 182.92. One $^{13}$C NMR signal was not detected, possibly due to rapid tautomerization of the molecule. HRMS (ESI) calculated for $C_{14}H_{14}O_2N$ ($[M+H]^+$) 228.1021, found 228.1024. HRMS (ESI) calculated for $C_{14}H_{13}O_2N\ ^{23}Na$ ($[M+Na]^+$) 250.0839, found 250.0840.

**Crystal formation for X-ray analysis.**

**1** was dissolved in a minimal amount of acetonitrile and then diluted to approximately 1 mg/mL with toluene. Crystals formed overnight upon vial-in-vial vapor diffusion using hexanes as the diluting solvent.

A colorless rod measuring 0.12 x 0.10 x 0.10 mm in size was mounted on a Cryoloop with Paratone oil. Data were collected in a nitrogen gas stream at 100(2) K using phi and omega scans. The crystal-to-detector distance was 60 mm and the exposure time was 5 s per frame using a scan

width of 1.0°. Data collection was 98.3% complete to 67.00° in θ. A total of 17398 reflections were collected covering the indices, -16<=*h*<=16, -8<=*k*<=7, -16<=*l*<=16. 2246 reflections were found to be symmetry independent, with an $R_{int}$ of 0.0248. Indexing and unit cell refinement indicated a primitive, monoclinic lattice. The space group was found to be P2(1)/c (No. 14). The data were integrated using the Bruker SAINT software program and scaled using the SADABS software program. Solution by direct methods (SIR-2008) produced a complete heavy-atom phasing model consistent with the proposed structure. All non-hydrogen atoms were refined anisotropically by full-matrix least-squares (SHELXL-97). All hydrogen atoms were placed using a riding model. Their positions were constrained relative to their parent atoms using the appropriate HFIX command in SHELXL-97.

**Production of genome-free MS2 capsids.**

The DNA encoding the MS2 coat protein sequence was inserted into the pBAD/Myc-His vector. The plasmid was transformed into New England Biosciences 5α *E.coli* and grown in LB media with 100 μg/mL ampicillin and 0.05% arabinose for induction. After 18 h of growth at 37 °C the cells were centrifuged and separated from the media. The following steps were carried out at 4 ºC: First, the cells were resuspended in pH 9 taurine buffer, lysed via sonication, and then the cell lysate was centrifuged and the supernatant was loaded on a DEAE-Sephadex column (GE Healthcare). Fractions containing MS2 capsids were combined and the protein was precipitated by adding $(NH_4)_2SO_4$ to a final concentration of 50%. The protein pellet was next resuspended in a minimal volume of pH 6.5 phosphate buffer and applied to a Sephacryl-500 column (GE Healthcare). Fractions containing MS2 capsids were combined and concentrated using 100k MWCO spin concentrators (Millipore).

Both wild type (19T and 85Y) MS2 and T19Y MS2 were produced using this procedure. T19*p*AF MS2 was produced as described.[4]

**Diazonium/dithionite modification of MS2 to produce MS2-*o*-aminophenol-85.**

4-Nitrobenzenediazonium tetrafluoroborate (Sigma-Aldrich) was dissolved in acetonitrile at a concentration of 200 mM and stored on ice. To a 495 μL solution of wild type MS2 capsids (100 μM in monomer) in 75 mM phosphate buffer, pH 9, was added 5 μL of the diazonium solution (resulting in a final concentration of 2 mM diazonium salt). The solution was allowed to react for 15 min while being cooled on ice before purification on a Nap-5 Sephadex size exclusion column equilibrated with pH 6.5 phosphate buffer. For other proteins, we recommend varying the concentrations of the diazonium salt to obtain the optimal level of conversion and maximum protein recovery.

To the above solution of azo-modified MS2 was added sodium dithionite[14] to a final concentration of 20 mM. This solution was allowed to react for 90 minutes at room temperature before purification on a Nap-10 Sephadex size exclusion column (GE Healthcare) equilibrated with pH 6.5 phosphate buffer. For MS2 capsids, the total protein recovery was ~70% after both reaction steps and SEC columns.

Note: It is important that a fresh bottle of sodium dithionite that has been stored at 4 °C is used. Reduction using some samples of sodium dithionite that had been stored at room temperature resulted in additional unidentified products (such as [M+80]) being formed in varying amounts.

**Generation of MS2 T19Y mutants.**

The QuikChange II Site Directed Mutagenesis Kit (Stratagene, La Jolla, CA) was used to mutate the threonine at position 19 to a tyrosine using the following forward and reverse primers:

Forward: 5' - ggaactggcgacgtgtatgtcgcccca - 3'
Reverse: 5' - ttgcttggggcgacatacacgtcgcca - 3'

The manufacturer's instructions were followed for these procedures. Incorporation of the point mutation was confirmed by sequencing of the colonies that were obtained after the mutation procedure.

**General procedure for coupling anilines to *o*-aminophenols on proteins (written specifically for the oxidative coupling of aniline peptide to MS2-aminophenol-85).**

To a solution of MS2-aminophenol-85 (50 μM) in dilute pH 6.5 phosphate buffer was added a solution of aniline-$His_6$-Gly peptide to a final concentration of approximately 500 μM. Sodium periodate (Sigma-Aldrich) was dissolved to a concentration of 50 mM in dilute pH 6.5 phosphate buffer. It was then added to the protein solution to reach a final concentration of 5 mM, and the reaction was allowed to proceed for 2 min at RT. In cases where a reaction quench was desired, a 100 mM solution of tris(carboxyethyl)phosphine (TCEP) at pH 6.5 was added to a final concentration of 50 mM. The resulting protein samples were purified on a Nap 5 Sephadex size exclusion column (GE Healthcare). If any unreacted peptide remained, it was removed by diluting the protein solution with phosphate buffer and concentrating the resulting solution using a 100k molecular weight cut-off spin concentrator (Millipore).

**Synthesis of PEG-o-aminophenol, PEG-aniline, and aniline peptide.**

*Monomethoxy-polyethyleneglycol*-o-*aminophenol (5k-PEG*-o-*aminophenol).*

Dicyclohexyl-carbodiimide (27 mg, 0.13 mmol) was added to a flask containing 3-(4-hydroxy-3-nitrophenyl)propionic acid (23 mg, 0.11 mmol, synthesized via literature procedure[15]) and *N*-hydroxysuccinimide (13 mg, 0.11 mmol) dissolved in $CH_2Cl_2$. The reaction was stirred for 15 min at room temperature, after which the solution was filtered to remove the DCU precipitate. The resulting solution was added to a second flask containing monomethoxy-polyethyleneglycol-amine (Sigma-Aldrich, 5000 molecular weight, 500 mg, 0.10 mmol) and triethylamine (0.32 mmol). After 2 h of stirring at RT the solvent was removed under reduced pressure and the resulting solid was dissolved in 15 mL of pH 6.5 100 mM phosphate buffer. The resulting suspension was then filtered to remove insoluble material. Sodium dithionite (1 mmol) was then added and solution immediately turned colorless. After 5 min of stirring, NaCl was added to 2 M and the aqueous layer was extracted three times with $CH_2Cl_2$. The combined organic extracts were washed once with brine and dried over $Na_2SO_4$. The solvent was removed under reduced pressure to afford 250 mg of product as a solid (48% yield for PEG5k-aminophenol). The extent of PEG modification with *o*-aminophenol was quantified by the absorbance at 290 nm and compared to standards of 2-amino-4-methylphenol. This value was used to determine the amount of functionalized PEG that was added to the coupling experiments. PEG2k-aminophenol was prepared using the same procedure and purification protocol. The product was obtained as a solid (~25% yield). The extent of conversion for the 2k product was determined as described for the 5k version, above.

*Monomethyoxy-polyethyleneglycol-aniline (5k-PEG-aniline).*

Dicyclohexylcarbodiimide (27 mg, 0.13 mmol) was added to a flask containing 3-(4-(*N*-boc-amino)phenyl)propionic acid (29 mg, 0.11 mmol) and *N*-hydroxysuccinimide (13 mg, 0.11 mmol) dissolved in $CH_2Cl_2$. The reaction was stirred for 15 min at room temperature, after which the solution was filtered to remove the DCU precipitate. The resulting material was added to a second flask containing monomethoxy-polyethyleneglycol-amine (Sigma-Aldrich, 5000 molecular weight, 500 mg, 0.10 mmol) and triethylamine (0.32 mmol). After two hours of stirring at room temperature, the solvent was removed under reduced pressure. The resulting solid was dissolved in 15 mL of 1 M NaCl and filtered to remove insoluble material. The aqueous layer was extracted three times with $CH_2Cl_2$ and the combined organic extracts were washed once with brine and dried over $Na_2SO_4$. The resulting solid was dissolved in 10 mL of $CH_2Cl_2$ and 10 mL of trifluoroacetic acid and allowed to react for 1 h at room temperature. The solvents were next removed under reduced pressure. The residue was dissolved in phosphate buffer and adjusted to pH 8. NaCl was added to 2 M and the aqueous layer was extracted with $CH_2Cl_2$ and dried as described above. Removal of the solvent under reduced pressure afforded 240 mg of solid (46% yield). The extent of PEG modification with aniline was quantified by absorbance at 287 nm and compared to standards of *p*-toluidine. This value was used to determine the amount of functionalized PEG that was added to the coupling experiments. PEG2k-aniline was prepared using the same procedure and purification protocol. The product was obtained as a solid (~25% yield). The extent of conversion for the 2k product was determined as described for PEG5k-aniline, above.

*Aniline-$His_6$-Gly peptide.*

The peptide was synthesized using standard Fmoc chemistry using a Wang resin preloaded with glycine (EMD Chemicals). HCTU was used as a coupling agent. As a final step, the N-terminus of the peptide was modified with the NHS ester of 3-(4-(*N*-boc-amino)phenyl)propionic acid, which was prepared as described in the PEG-aniline synthesis described above. After cleavage from the resin the peptide was precipitated by the addition of ether, isolated via filtration, and then dissolved in water for use in the coupling reactions. The peptide was characterized using MALDI-TOF with sinapinic acid as the matrix (Figure 2-13).

**Synthesis of cyclic RGD peptide aminophenol and subsequent coupling of *p*-toluidine.**

Cyclic peptide RGDyE (Peptides International, structure in Figure 2-12) was dissolved in DMSO to a concentration of 50 mg/mL, then diluted to 5 mM in pH 6.5 phosphate buffer (10 mM). A 50 µL portion of this peptide solution was diluted into 400 µL of 50 mM pH 7.8 ammonium bicarbonate buffer. To this solution was added 50 µL of a 150 mM solution of tetranitromethane dissolved in ethanol (Scheme 2-1). After 1.5 h reaction at room temperature trifluoroacetic acid was added to achieve a pH of 3 and the solution was loaded on to a C18 Sep Pack (Waters) preconditioned with acetonitrile then 0.1% aqueous TFA. After washing with aqueous 0.1% TFA the peptide was eluted with acetonitrile. After removing the acetonitrile under reduced pressure the solid peptide was dissolved in 100 µL pH 6.5 phosphate buffer (10 mM). The nitrated peptide was characterized by MALDI-TOF MS using sinapinic acid as the matrix. In addition to the expected peak at 666 amu there was a second peak at 650 amu (-16 amu from the main peak) that was an artifact seen on MALDI-TOF MS for all nitrophenol containing substrates (Figure 2-5c). Sodium dithionite dissolved in pH 6.5 phosphate buffer (200 mM) was added to 20 µL of the above peptide solution to a final concentration of 5 mM. After 15 minutes at room temperature 2 µL were removed for MS characterization and the remainder was added to a solution of *p*-toluidine (200 µM).

Sodium periodate was added to a final concentration of 5 mM and after two minutes of reaction the solution was purified via C18 Zip Tip (Millipore) in the same manner as the C18 Sep Pack above. The RGD-toluidine coupling product was characterized via LCMS. In addition to the expected mass of 755 amu, a second peak at 740 amu was observed. This mass corresponded to product **2**, in which the aniline and aminophenol have formed a covalent bond but imine hydrolysis occurred before the iminoalcohol cleavage. This peak becomes the major peak when reaction-limiting concentrations of sodium periodate are used (Figure 2-5g). Also, when TCEP was added, the mass corresponding to 2 increased by two mass units, thus mimicking the behavior of **2** (Figure 2-5e and 2-5f, for with and without TCEP, respectively).

**Oxidative coupling in the presence of glucose.**

Sodium periodate can also be used as an oxidant for the cleavage of vicinal diols in sugars. As a preliminary test for the compatibility of this oxidative coupling reaction with carbohydrates, we repeated the coupling of PEG5k-aminophenol with T19*p*AF MS2 in the presence of varying amounts of glucose. We found no significant difference in coupling between samples with 0 or 20 mM glucose (Figure 2-11). Though this experiment did not analyze the state of the vicinal diols, it did show that small amounts of sugars will not consume enough oxidative equivalents of periodate to prevent the aminophenol-aniline coupling.

**Reversed phase HPLC of small molecule coupling reaction.**

The isolated product **1** and crude reaction of 2-amino-4-methylphenol and *p*-toluidine were analyzed by reversed-phase HPLC on an Agilent 1100 Series HPLC System (Agilent Technologies, Santa Clara, CA). Reversed-phase liquid chromatography was accomplished on a Gemini 5 μm C18 column (150 mm × 4.6 mm, Phenomenex). The mobile phase (0.5 mL/min flowrate) was a linear gradient over 20 minutes from 15% to 95% acetonitrile:water containing 0.1% TFA.

**Coupling of *p*-toluidine to (MS2-aminophenol-85).**

MS2-aminophenol-85 was produced according to the procedure above and was then reacted with *p*-toluidine in the presence of sodium periodate. The capsids were disassembled before analysis by MALDI-TOF MS (Figure 2-6).

**Preparation of lysozyme-aniline and subsequent coupling to PEG-*o*-aminophenol**

3-(4-aminophenyl)propionic acid (Aldrich) was dissolved to 50 mM in DMF. *N*-hydroxysuccinimide was added to a concentration of 100 mM, after which 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC, Fluka) was added to a concentration of 100 mM. After 10 min, a 50 μL aliquot of the solution was added to 950 μL of a 10 μM solution of lysozyme (from chicken egg white, Sigma) in 50 mM pH 8.5 phosphate buffer. The resulting solution was allowed to react at room temperature for 1 h, at which time the lysozyme was purified via a Nap-10 size exclusion column equilibrated with pH 6.5 phosphate buffer. After concentrating the lysozyme down to a volume of 1.0 mL using a 10kDa molecular weight cutoff spin concentrator, it was again purified through a Nap-10 size exclusion column. MALDI-TOF analysis indicated a range of two to five aniline modifications on each lysozyme (Figure 2-9b).

The resulting lysozyme-aniline (20 μM) was reacted with PEG2k-aminophenol (100 μM), in the presence of $NaIO_4$ (2 mM) for 2 min. The reaction was quenched with DTT and was analyzed via SDS-PAGE (Figure 2-9c).

**Preparation of fd Phage**

A colony of bacteria containing fd-tet-scFv was inoculated into 5 mL 2TY with 15 μg/mL tetracycline and grown overnight at 30 °C with shaking (300 rpm). The overnight culture was added into 1 L of 2TY with 15 μg/mL tetracycline and grown at 30 °C with shaking for 12-18 h. The bacteria were then centrifuged at 6,000 x g in 500 mL centrifuge tubes in a GS3 rotor for 30 minutes. The supernatant was transferred to new 500 mL centrifuge tubes and the phage were precipitated by adding 1/10-1/5 volume of PEG/NaCl solution (20% PEG8k, 2.5 M NaCl) and left on ice for 1 hour. The phage were visible as a clouding of the supernatant. After 1 h the phage were pelleted in 500 mL bottles by centrifuging in a GS3 rotor at 4,000 rpm for 15 minutes at 4 °C. The supernatant was discarded and the 'dry' pellet was centrifuged again for 30 seconds to bring down the last drops of supernatant, and the liquid was removed. The pellet was suspended in 1/10 volume of PBS and centrifuged at 10,000 rpm for 15 min to pellet bacterial debris. The resulting supernatant was transferred to a new tube. These precipitation steps can be repeated to further purify the phage. Absorbance at 269 nm was used to quantify the concentration of phage. For the 7222 base pair genome construct used for these experiments, 1 absorbance unit at 269 nm is expected for a phage solution at 37.3 μM in pVIII. For example, a 100 μM pVIII phage solution will give an absorbance at 269 nm of 2.68. Changing the size of the genome will give a different absorbance coefficient because the number of copies of pVIII is dependent on the size of the genome.

**PLP-mediated N-terminal Transamination of pVIII Protein of fd Phage**

To a solution of phage (100 μM pVIII) was added pyridoxyl-5-phosphate (PLP) to a final concentration of 100 mM in pH 6.5 phosphate. After overnight incubation at room temperature the phage were purified using multiple iterations of the precipitation protocol described in the previous paragraph, until the supernatant was no longer colored with PLP.

**Synthesis of Aminophenol and Aniline Alkoxyamines**

Transaminated pVIII could be site-specifically modified at the N-terminus using alkoxyamine functional groups. An aminophenol-alkoxyamine and aniline-alkoxyamine were synthesized for this purpose (Scheme 2-2):

*Synthesis of aminophenol-alkoxyamine*

To tyramine was added dropwise one equivalent of fuming nitric acid at 4 °C using trifluoroacetic acid as the solvent and this resulted in quantitative conversion to *o*-nitrotyramine. Separately, to (Boc-aminooxy)acetic acid (280 mg, 1.5 mmol) dissolved in methylene chloride was added dicyclohexylcarbodiimide (362 mg, 1.7 mmol) and *N*-hydroxysuccinimide (168 mg, 1.5 mmol). After 15 min with stirring, the precipitate was filtered through Celite, then a 0.22 μM PVDF syringe filter. *O*-nitrotyramine (273 mg, 1.5 mmol) was added to the succinimidyl ester along with sufficient triethylamine to deprotonate the phenol. After 1 h stirring at room temperature the solution was filtered to remove any precipitate and the solution was washed with 0.1M $NaHSO_4$, brine and dried over $Na_2SO_4$. The nitrophenol was dissolved in a small amount of methylene chloride and purified on a silica column using 65:35 ethyl acetate:hexanes mobile phase to afford 350 mg purified product for a 65% yield. The purified nitrophenol-Boc-alkoxyamine (250 mg, 0.7 mmol) was reduced with a small scoop of Pd/C and $H_2$ supplied by a rubber ball. After 2 h stirring the reaction was filtered through Celite to remove the catalyst, concentrated under reduced

pressure, then purified on a silica column using ethyl acetate as the mobile phase. This afforded 100 mg of the final product for a 31% yield. The loss of product occurred before the silica column purification, so it is suspected that either the starting material or product stuck to the activated carbon. $^1$H NMR (400 MHz, CDCl3) δ 1.5 (s, 9H), 2.7 (t, 2H), 3.5 (q, 2H), 4.3 (s, 2H), 6.5 (d, 1H), 6.6 (s, 1H), 6.7 (d, 1H), 7.7 (s, 1H), 8.0 (br s, 1H). The Boc group was removed with 1:1 trifluoroacetic acid:methylene chloride for 10 min, the solvent was removed under a stream of nitrogen, and the resulting oil was placed under vacuum overnight. It was then dissolved to 100 mM in water and frozen until used. It was necessary to neutralize the residual trifluoroacetic acid in this solution before addition to proteins to avoid protein precipitation.

*Synthesis of aniline-alkoxyamine*

To (Boc-aminooxy)acetic acid (280 mg, 1.5 mmol) dissolved in methylene chloride was added dicyclohexylcarbodiimide (362 mg, 1.7 mmol) and *N*-hydroxysuccinimide (168 mg, 1.5 mmol). After 15 min with stirring, the precipitate was filtered through Celite, then a 0.22 μM PVDF syringe filter. To the remaining solution was added 2-(4-aminophenyl)ethylamine (200 mg, 1.5 mmol) and triethylamine (400 mg, 4 mmol). After 1 h stirring the solution was concentrated under reduced pressure and applied to a silica column. Purification using ethyl acetate as the mobile phase afforded approximately 200 mg of the product for 45% yield. $^1$H NMR (400 MHz, CDCl3) δ 1.4 (s, 9H), 2.7 (t, 2H), 3.4 (q, 2H), 3.5 (br s, 2H), 4.2 (s, 2H), 6.6 (d, 2H), 7.0 (d, 2H), 8.1 (br s, 1H), 8.5 (s, 1H). The Boc group was removed with 1:1 trifluoroacetic acid:methylene chloride for 10 min, the solvent was removed under a stream of nitrogen, and the resulting oil was placed under vacuum overnight. It was then dissolved to 100 mM in water and frozen until used. It was necessary to neutralize the residual trifluoroacetic acid in this solution before addition to proteins to avoid protein precipitation.

**Oxime Formation with Transaminated pVIII with Aminophenol and Aniline Alkoxyamines**

The proper volume alkoxyamine stock solutions were dispensed into a tube and were neutralized with small volumes of concentrated buffer until the solution was approximately pH 5. The phage solution was added for a final concentration of 50 μM pVIII and 10-20 mM alkoxyamine. The solution was incubated at room temperature overnight, then the phage were Nap-5 or Nap-10 purified. MALDI-TOF MS of the phage showed the percentage of the pVIII coat proteins that were modified (Figure 2-15).

**Attachment of PEG5k to pVIII Using Oxidative Coupling and HPLC Analysis**

To a pH 6.5 phosphate buffered solution of phage-aminophenol or phage-aniline (20 μM pVIII) was added the complentary PEG5k-aminophenol or PEG5k-aniline (100 μM) and sodium periodate (1 mM). After 2 min the reaction was purified by Nap-5 and spin concentrated against a 100,000 molecular weight cut off membrane until the concentration of pVIII was 100 μM. 10 μL of 100 μM pVIII was injected onto a Zorbax-CN 150 x 4.6 mm analytical HPLC column. A gradient of $H_2O$:MeCN both with 0.1% TFA was used to separate the proteins. The pVIII peak could be visualized using absorbance at 280 nm or fluorescence (280 nm excitation, 330 nm emission).

## $^{13}$C NMR Spectra of 1

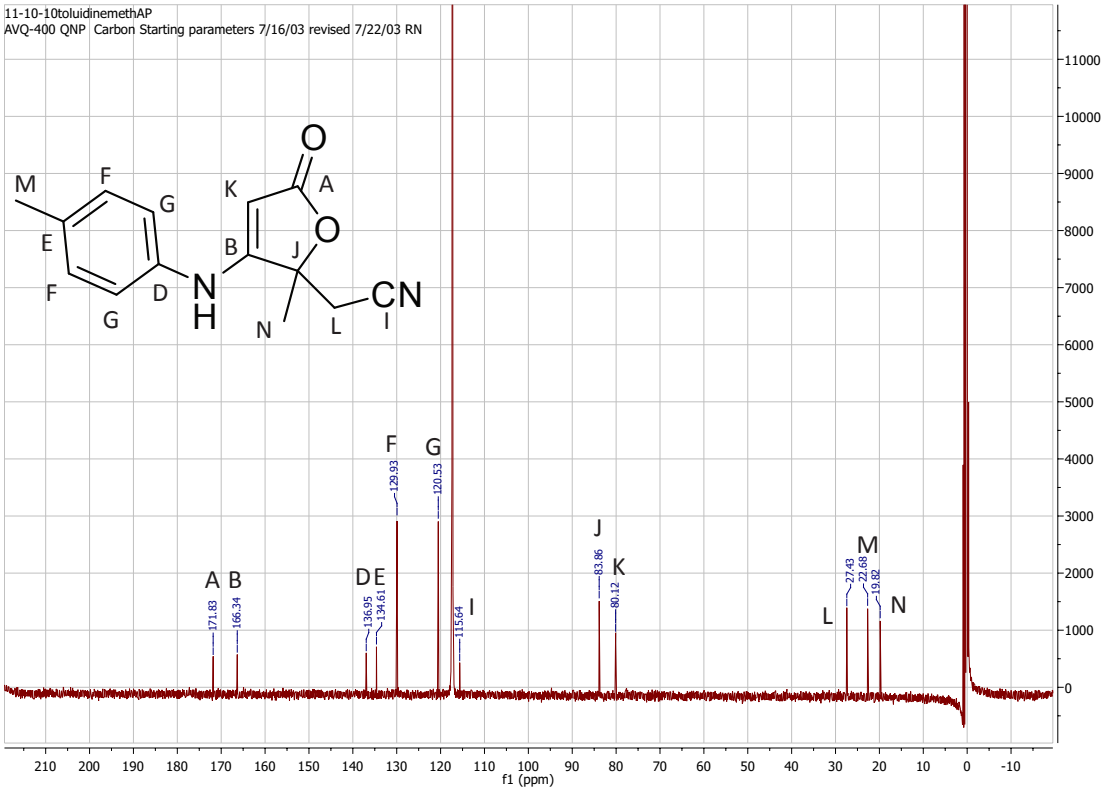
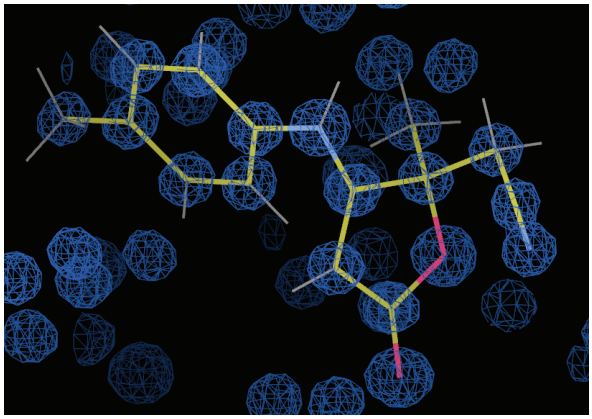



**Figure 2-17.** $^{13}$C NMR spectrum of **1.**

## Structure of 1 Overlay with Electron Density Map

(a)



(b)

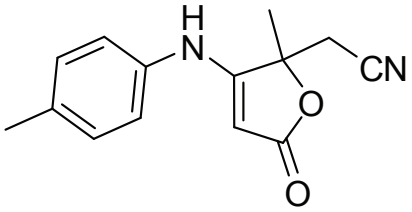



**Figure 2-18**. (a) **1** structure overlay with electron density map from x-ray diffraction data. (b) Structure of **1**.

**Crystallographic Characterization Data**

Table 1. Crystal data and structure refinement for **1**.

| | | |
|---|---|---|
| X-ray ID | **1** | |
| Empirical formula | C14 H14 N2 O2 | |
| Formula weight | 242.27 | |
| Temperature | 100(2) K | |
| Wavelength | 1.54178 Å | |
| Crystal system | Monoclinic | |
| Space group | P2(1)/c | |
| Unit cell dimensions | a = 14.1010(12) Å | α= 90°. |
| | b = 6.9117(5) Å | β= 112.570(5)°. |
| | c = 14.0777(12) Å | γ = 90°. |
| Volume | 1266.96(18) $Å^3$ | |
| Z | 4 | |
| Density (calculated) | 1.270 $Mg/m^3$ | |
| Absorption coefficient | 0.701 $mm^{-1}$ | |
| F(000) | 512 | |
| Crystal size | 0.12 x 0.10 x 0.10 $mm^3$ | |
| Crystal color/habit | colorless rod | |
| Theta range for data collection | 3.39 to 67.64°. | |
| Index ranges | -16<=h<=16, -8<=k<=7, -16<=l<=16 | |
| Reflections collected | 17398 | |
| Independent reflections | 2246 [R(int) = 0.0248] | |
| Completeness to theta = 67.00° | 98.3 % | |
| Absorption correction | Semi-empirical from equivalents | |
| Max. and min. transmission | 0.9332 and 0.9206 | |
| Refinement method | Full-matrix least-squares on $F^2$ | |
| Data / restraints / parameters | 2246 / 0 / 165 | |
| Goodness-of-fit on $F^2$ | 1.050 | |
| Final R indices [I>2sigma(I)] | R1 = 0.0674, wR2 = 0.1819 | |
| R indices (all data) | R1 = 0.0735, wR2 = 0.1872 | |
| Largest diff. peak and hole | 0.682 and -0.304 $e.Å^{-3}$ | |

Table 2. Atomic coordinates ( x $10^4$) and equivalent isotropic displacement parameters ($Å^2$x $10^3$) for **1**. U(eq) is defined as one third of the trace of the orthogonalized $U^{ij}$ tensor.

| | x | y | z | U(eq) |
|---|---|---|---|---|
| C(1) | 1035(2) | 6562(3) | 9080(2) | 32(1) |
| C(2) | 2165(2) | 6529(3) | 9760(2) | 29(1) |
| C(3) | 2482(2) | 4657(4) | 9937(2) | 30(1) |
| C(4) | 1627(2) | 3432(3) | 9422(2) | 31(1) |
| C(5) | 832(2) | 7422(4) | 8019(2) | 38(1) |
| C(6) | 1352(3) | 6375(5) | 7484(2) | 47(1) |
| C(7) | 373(2) | 7526(4) | 9582(2) | 40(1) |
| C(8) | 3726(2) | 8466(3) | 10631(2) | 31(1) |
| C(9) | 4047(2) | 9883(4) | 11390(2) | 36(1) |
| C(10) | 5086(2) | 10239(4) | 11916(2) | 41(1) |
| C(11) | 5832(2) | 9183(4) | 11716(2) | 40(1) |
| C(12) | 5490(2) | 7773(4) | 10956(2) | 40(1) |
| C(13) | 4459(2) | 7416(4) | 10414(2) | 37(1) |
| C(14) | 6958(3) | 9555(5) | 12282(3) | 54(1) |
| N(1) | 1775(3) | 5490(5) | 7072(2) | 64(1) |
| N(2) | 2660(2) | 8206(3) | 10066(2) | 33(1) |
| O(1) | 764(1) | 4537(2) | 8930(1) | 32(1) |
| O(2) | 1538(2) | 1684(2) | 9349(2) | 37(1) |

Table 3. Bond lengths [Å] and angles [°] for **1**.

| | | | |
|---|---|---|---|
| C(1)-O(1) | 1.445(3) | C(8)-C(9) | 1.390(4) |
| C(1)-C(2) | 1.513(4) | C(8)-C(13) | 1.390(4) |
| C(1)-C(7) | 1.522(4) | C(8)-N(2) | 1.417(3) |
| C(1)-C(5) | 1.529(4) | C(9)-C(10) | 1.386(4) |
| C(2)-N(2) | 1.336(3) | C(9)-H(9) | 0.9500 |
| C(2)-C(3) | 1.360(3) | C(10)-C(11) | 1.395(4) |
| C(3)-C(4) | 1.423(4) | C(10)-H(10) | 0.9500 |
| C(3)-H(3) | 0.9500 | C(11)-C(12) | 1.389(4) |
| C(4)-O(2) | 1.215(3) | C(11)-C(14) | 1.501(4) |
| C(4)-O(1) | 1.378(3) | C(12)-C(13) | 1.380(4) |
| C(5)-C(6) | 1.433(4) | C(12)-H(12) | 0.9500 |
| C(5)-H(5A) | 0.9900 | C(13)-H(13) | 0.9500 |
| C(5)-H(5B) | 0.9900 | C(14)-H(14A) | 0.9800 |
| C(6)-N(1) | 1.154(4) | C(14)-H(14B) | 0.9800 |
| C(7)-H(7A) | 0.9800 | C(14)-H(14C) | 0.9800 |
| C(7)-H(7B) | 0.9800 | N(2)-H(2) | 0.8800 |
| C(7)-H(7C) | 0.9800 | | |

| | |
|---|---|
| O(1)-C(1)-C(2) | 103.49(19) |
| O(1)-C(1)-C(7) | 108.5(2) |
| C(2)-C(1)-C(7) | 113.5(2) |
| O(1)-C(1)-C(5) | 107.0(2) |
| C(2)-C(1)-C(5) | 112.3(2) |
| C(7)-C(1)-C(5) | 111.4(2) |
| N(2)-C(2)-C(3) | 132.2(2) |
| N(2)-C(2)-C(1) | 118.9(2) |
| C(3)-C(2)-C(1) | 108.8(2) |
| C(2)-C(3)-C(4) | 108.5(2) |
| C(2)-C(3)-H(3) | 125.7 |
| C(4)-C(3)-H(3) | 125.7 |
| O(2)-C(4)-O(1) | 117.7(2) |
| O(2)-C(4)-C(3) | 132.4(2) |
| O(1)-C(4)-C(3) | 109.8(2) |
| C(6)-C(5)-C(1) | 112.1(2) |
| C(6)-C(5)-H(5A) | 109.2 |
| C(1)-C(5)-H(5A) | 109.2 |
| C(6)-C(5)-H(5B) | 109.2 |
| C(1)-C(5)-H(5B) | 109.2 |
| H(5A)-C(5)-H(5B) | 107.9 |
| N(1)-C(6)-C(5) | 178.1(3) |
| C(1)-C(7)-H(7A) | 109.5 |
| C(1)-C(7)-H(7B) | 109.5 |
| H(7A)-C(7)-H(7B) | 109.5 |
| C(1)-C(7)-H(7C) | 109.5 |
| H(7A)-C(7)-H(7C) | 109.5 |
| H(7B)-C(7)-H(7C) | 109.5 |
| C(9)-C(8)-C(13) | 119.1(3) |
| C(9)-C(8)-N(2) | 118.9(2) |
| C(13)-C(8)-N(2) | 121.9(2) |
| C(10)-C(9)-C(8) | 120.1(3) |
| C(10)-C(9)-H(9) | 120.0 |
| C(8)-C(9)-H(9) | 120.0 |
| C(9)-C(10)-C(11) | 121.5(2) |
| C(9)-C(10)-H(10) | 119.2 |
| C(11)-C(10)-H(10) | 119.2 |
| C(12)-C(11)-C(10) | 117.2(3) |
| C(12)-C(11)-C(14) | 120.8(3) |
| C(10)-C(11)-C(14) | 122.0(3) |
| C(13)-C(12)-C(11) | 122.1(3) |
| C(13)-C(12)-H(12) | 118.9 |
| C(11)-C(12)-H(12) | 118.9 |

| | |
|---|---|
| C(12)-C(13)-C(8) | 120.0(2) |
| C(12)-C(13)-H(13) | 120.0 |
| C(8)-C(13)-H(13) | 120.0 |
| C(11)-C(14)-H(14A) | 109.5 |
| C(11)-C(14)-H(14B) | 109.5 |
| H(14A)-C(14)-H(14B) | 109.5 |
| C(11)-C(14)-H(14C) | 109.5 |
| H(14A)-C(14)-H(14C) | 109.5 |
| H(14B)-C(14)-H(14C) | 109.5 |
| C(2)-N(2)-C(8) | 126.9(2) |
| C(2)-N(2)-H(2) | 116.5 |
| C(8)-N(2)-H(2) | 116.5 |
| C(4)-O(1)-C(1) | 109.27(19) |

Symmetry transformations used to generate equivalent atoms:

Table 4. Anisotropic displacement parameters ($Å^2 x\ 10^3$)for **1**. The anisotropic displacement factor exponent takes the form: $-2\pi 2[\ h^2 a^{*2} U^{11} + ... + 2\ h\ k\ a^*\ b^*\ U^{12}\ ]$

| | $U^{11}$ | $U^{22}$ | $U^{33}$ | $U^{23}$ | $U^{13}$ | $U^{12}$ |
|---|---|---|---|---|---|---|
| C(1) | 43(1) | 14(1) | 37(1) | 0(1) | 12(1) | 1(1) |
| C(2) | 43(1) | 17(1) | 29(1) | -1(1) | 14(1) | -1(1) |
| C(3) | 39(1) | 20(1) | 31(1) | 1(1) | 12(1) | 4(1) |
| C(4) | 44(1) | 18(1) | 33(1) | 2(1) | 18(1) | 4(1) |
| C(5) | 49(2) | 19(1) | 37(1) | 3(1) | 8(1) | -1(1) |
| C(6) | 58(2) | 40(2) | 41(2) | 4(1) | 17(1) | -9(1) |
| C(7) | 47(2) | 22(1) | 53(2) | -4(1) | 22(1) | 2(1) |
| C(8) | 45(1) | 18(1) | 30(1) | 2(1) | 14(1) | -4(1) |
| C(9) | 52(2) | 22(1) | 36(1) | -2(1) | 20(1) | -4(1) |
| C(10) | 58(2) | 31(2) | 37(1) | -9(1) | 20(1) | -17(1) |
| C(11) | 52(2) | 36(2) | 38(1) | -4(1) | 25(1) | -16(1) |
| C(12) | 48(2) | 34(2) | 46(2) | -6(1) | 28(1) | -10(1) |
| C(13) | 51(2) | 26(1) | 37(1) | -7(1) | 21(1) | -8(1) |
| C(14) | 54(2) | 61(2) | 53(2) | -13(2) | 28(2) | -22(2) |
| N(1) | 83(2) | 63(2) | 56(2) | -6(2) | 37(2) | -6(2) |
| N(2) | 46(1) | 15(1) | 34(1) | -1(1) | 12(1) | 2(1) |
| O(1) | 42(1) | 14(1) | 38(1) | -1(1) | 12(1) | 0(1) |
| O(2) | 52(1) | 14(1) | 49(1) | 0(1) | 23(1) | 1(1) |

Table 5. Hydrogen coordinates ( x $10^4$) and isotropic displacement parameters ($Å^2$x $10^3$) for **1**.

| | x | y | z | U(eq) |
|---|---|---|---|---|
| H(3) | 3160 | 4238 | 10338 | 36 |
| H(5A) | 84 | 7406 | 7605 | 45 |
| H(5B) | 1063 | 8786 | 8100 | 45 |
| H(7A) | -352 | 7430 | 9122 | 60 |
| H(7B) | 566 | 8892 | 9712 | 60 |
| H(7C) | 479 | 6878 | 10234 | 60 |
| H(9) | 3554 | 10608 | 11547 | 43 |
| H(10) | 5294 | 11227 | 12424 | 49 |
| H(12) | 5982 | 7032 | 10805 | 48 |
| H(13) | 4251 | 6453 | 9893 | 44 |
| H(14A) | 7261 | 8506 | 12774 | 81 |
| H(14B) | 7055 | 10787 | 12652 | 81 |
| H(14C) | 7295 | 9613 | 11790 | 81 |
| H(2) | 2283 | 9262 | 9899 | 39 |

## Section 2.5: References


1. Hooker J.M., Esser-Kahn, A.P., Francis, M.B., *J. Am. Chem. Soc.*, **2006**, *128*, 15558-15559.

2. Burdine, L.; Gillette, T. G.; Lin, H.J.; Kodadek, T. *J. Am. Chem. Soc.* **2004**, *126*,11442-11443.

3. Hooker J.M., Kovacs E.W., Francis, M.B. *J. Am. Chem. Soc.*, **2004**, *126*, 3718-3719.

4. Carrico Z.M., Romanini D.W., Mehl R.A., Francis M.B. *Chem. Commun.* **2008**, *10*, 1205-1207.

5. Valegard K., Liljas L., Fridborg K., Unge T., *Nature.* **1990**, *345*, 36 – 41.

6. PEG immunogenicity ref

7. Maina, T., Nock, B., and Mather, S. *Cancer Imaging.* **2006**, *6*, 153–157.

8. Ladner, R.C., Sato, A.K., Gorzelany, J., and de Souza, M. *Drug. Discov. Today*. **2004**, *9*, 525–529.

9. Sidhu, S.S., *Biomol. Eng.* **2001**, *18*, 57–63.

10. Gilmore J.M., Scheck R.A., Esser-Kahn A.P., Joshi N.S., Francis M.B. *Angew. Chem.*, *Int. Ed.* **2006**, *45*, 5307-5311.

11. Scheck R.A., Dedeo M.T., Iavarone A.T., Francis M.B. *J. Am. Chem. Soc.* **2008**, *130*, 11762-11770.

12. 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*, 16812-7.

13. Jencks WP. *J. Am. Chem. Soc.* **1959**, *81*, 475–481.

14. Gorecki, M.; Wilchek, M.; Patchornik, A. *Biochim. Biophys. Acta.* **1971**, *229*, 590-595.

15. Wright, J.B. *J. Heterocycl. Chem.* **1972**, *9*, 681-682.

# Chapter 3 - A Novel Method for Fluorine-18 Labeling of Proteins Using Oxidative Coupling

## Section 3.1: Introduction to Radioactivity and Positron Emission Tomography

Since the beginning of the molecular biology era, the labeling of proteins with radioactive isotopes has been a widely used technique for tracking proteins, DNA, and other biological molecules. It was particularly useful for early molecular biology studies because it was by far the most sensitive detection technique at the time. Another major advantage of radioisotopes were their similarity to biologically relevant atoms. Radioactive isotopes of elements such as hydrogen,[1] carbon,[2] phosphorus,[3] and sulfur[4] can all be incoporated into biological macromolecules synthetically or by living cells. One of the first important uses of radioactive isotopes in molecular biology was the Hershey-Chase experiment in 1952.[5] To determine whether proteins or DNA were responsibile for transferring genetic information Hershey and Chase added radioactive sulfur and phosphorous to bacteria before infecting them with bacteriophage. The infected bacteria produced bacteriophage with radioactive sulfur present in the protein coat and radioactive phosphorus present in the DNA. After re-infecting new bacteria and physically removing the bacteriophage from the membrane of the bacteria, the phosphorus was found inside the cell while the sulfur remained outside (Figure 3-1). These infected bacteria also produced new bacteriophage that contained some of the phosphorus indicating that the radiolabeled DNA had been passed on to the next generation of bacteriophage. Experiments like this one would have been impossible without the aid of radioactive isotopes to track and quantify biological molecules.

As a wide range of radioisotopes became more accessible, medical researchers began to show interest in using them to track molecules and processes in the body. Imaging the body posed significant challenges because of tissue absorption of the energy given off by the radionuclide. Beta emitters such as phosphorus-32 cannot easily penetrate tissue and were therefore not useful as *in vivo* imaging agents. Instead it was necessary to use radionuclides that emit gamma photons

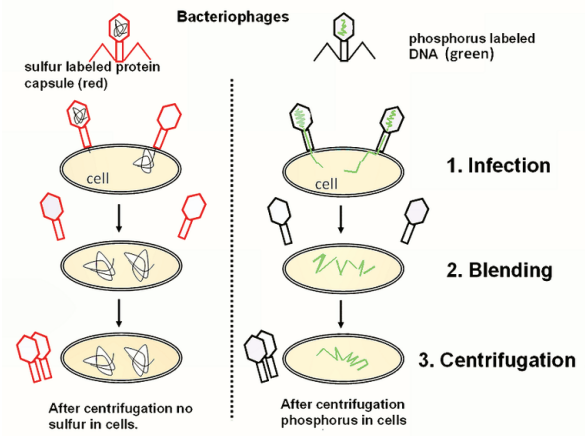



**Figure 3-1**. Experiment performed by Hershey and Chase in 1952. Viruses were labeled with radioactive sulfur and phosphorous to differentially track protein and DNA. During infection, the proteins labeled with sulfur remained outside the cell while the phosphorous-labeled DNA entered the cell and were incorporated into daughter viruses.

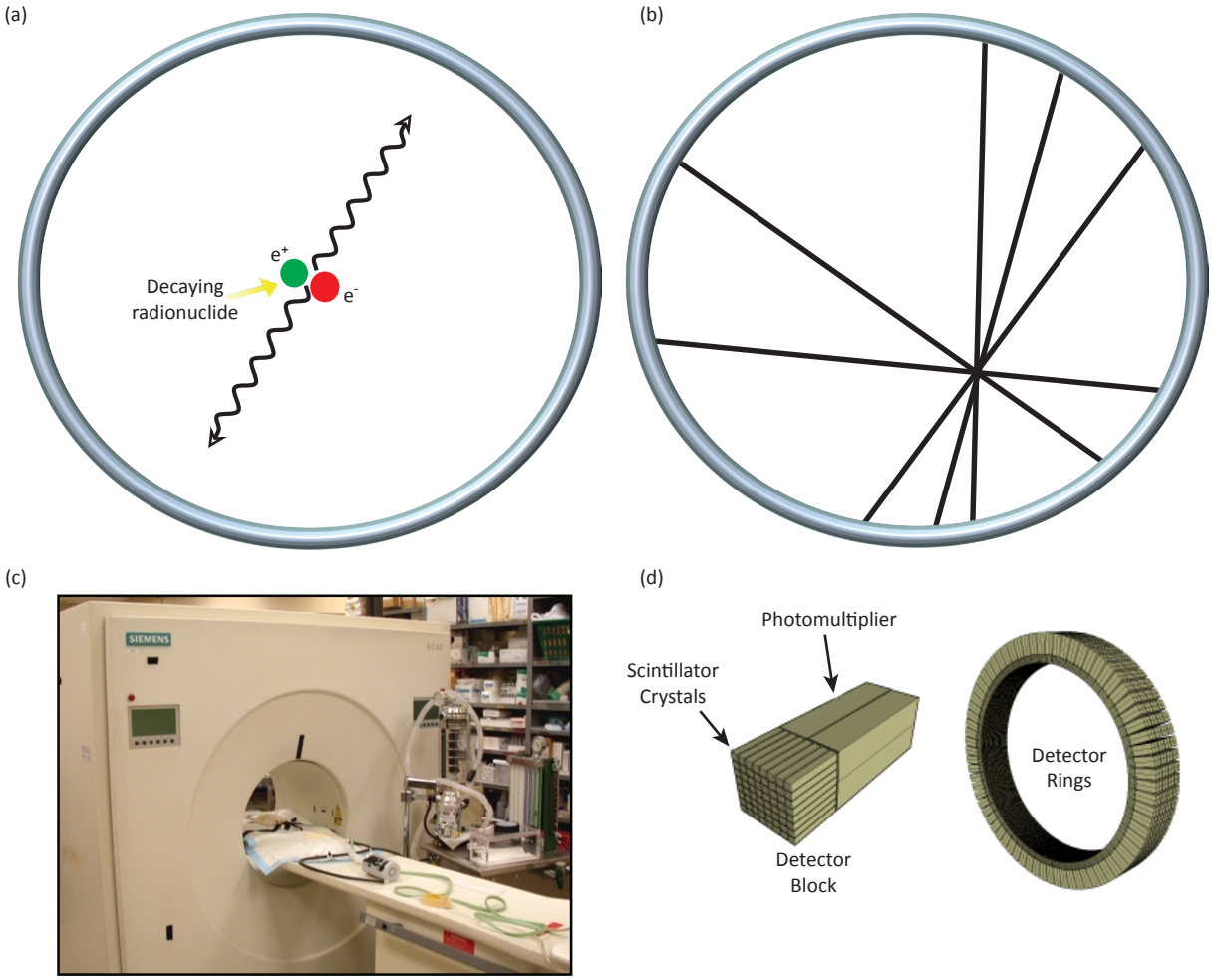



**Figure 3-2**. Positron emission tomography (PET) basics. (a) Schematic of a radionuclide emitting a positron which travels a short distance before matching energies with an electron and producing two 511 keV incident gamma rays through an annhilation event. (b) Schematic with line segments representing the trajectory of pairs of incident gamma rays. When many pairs of incident gamma rays have been detected, the overlapping point of the line segments can be pinpointed as the site of the annhilation event. (c) Photograph of a human PET scanner. Notice that only one part of the body, such as the head, can be imaged at once. Extending the field of view requires more detector rings which are expensive. (d) The scintillator crystals convert radiation into a photon and the photomultiplier converts the photon into an electrical signal that can be detected and recorded. One ring of detector blocks would result in a thin slice that would look like a two-dimensional picture. Coupling together many detector rings in the z-direction expands the field of view to give a true three-dimensional image.

since those are high enough energy to penetrate tissue but low enough to be detected by a scintillator crystal. Some radionuclides (most commonly $^{99m}$Tc and $^{123}$I) directly emit gamma photons which can be used for Single-Photon Emission Computed Tomography (SPECT). This manuscript, however, will focus on Positron Emission Tomography (PET) because it is now the more commonly used technique. As the name suggests, the radionuclides required for PET are positron emitters.

The positron emitted by these nuclides is short lived though. Once it loses enough energy to match the energy of an electron, an annhilation event occurs. The product of this event is two incident gamma photons. Due to conservation of momentum the two photons are emitted 180º apart. This is a key property of PET: if the circular detector picks up two simultaneous signals, then it can infer that the positron originated from somewhere on the line segment drawn between the two points of detection. If multiple segments overlap then it is likely that a positron emitted from

| Radionuclide | $^{11}C$ | $^{13}N$ | $^{15}O$ | $^{18}F$ | $^{45}Ti$ | $^{60}Cu$ | $^{61}Cu$ | $^{62}Cu$ | $^{64}Cu$ | $^{66}Ga$ | $^{76}Br$ | $^{82}Rb$ | $^{86}Y$ | $^{94m}Tc$ | $^{124}I$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $Energy_{avg}$ [a] (MeV) | 0.386 | 0.492 | 0.735 | 0.250 | 0.439 | 0.977 | 0.499 | 1.315 | 0.278 | 1.744 | 1.184 | 1.475 | 0.666 | 1.072 | 0.818 |
| $Range_{avg}$ [b] (mm) | 1.52 | 2.05 | 3.28 | 0.83 | 1.78 | 4.50 | 2.09 | 6.21 | 0.97 | 8.37 | 5.55 | 7.02 | 2.93 | 4.98 | 3.70 |

**Figure 3-3**. Average energies of emitted positrons for various radionuclides used for imaging. Also shown is the estimated average distance that the positron can travel through tissue before annhilating. This distance is the best spatial resolution possible with the radionuclide. Table from Tai and Laforest, *Annu. Rev. Biomed. Eng.*, **2005**, *7*, 255-285.

the point of intersection (Figure 3-2b). Once millions of these intersection points are measured computer software can piece them together to form two-dimensional images. Arrays of circular detectors give depth in the third dimension by stacking multiple two-dimensional cross-sections together to give a three-dimensional image.

The distance the positron travels through tissue before the annhilation event is also important. Since only the point of origin of the incident gamma photons can be measured, the distance the positron travels before the annhilation event is the limit of resolution of that imaging agent. This distance is different for each positron emitting isotope. Figure 3-3 lists the average energies and tissue range for various radionuclides. Notice that Fluorine-18 and Copper-64 have the shortest tissue penetration distance and therefore have the highest inherent resolution.

In addition to tissue penetration distance, another important attribute of radionuclides is their halflife. Nuclides with extremely long halflives are not useful because they emit an insufficient amount of gamma photons for imaging. On the other end of the spectrum, nuclides with very short halflives are difficult to incorporate into molecules and transport to patients before the activity has decayed. Of the commonly used organic radionuclides ($^{18}F$, $^{11}C$, $^{13}N$, $^{15}O$) $^{18}F$ has the most synthetic versatility because of its 110 min halflife (Figure 3-4). This near two hour halflife also allows for unbound imaging agent to clear out of the system in some cases. However, when imaging with antibodies the optimal target-to-background ratio can reach its maximum 1-2 days after injection. For these cases it is necessary to use longer lived halflife nuclides such as $^{64}Cu$. The chemistry for attaching radiometals is relatively simple — a chelating group such as 1,4,7,10-tetraazacyclododecane-1,4,7,10-tetraacetic acid (DOTA) is attached to the molecule of interest ahead of time and then mixed with the radiometal once it is created.[6] As a lab in the Chemistry Depart-

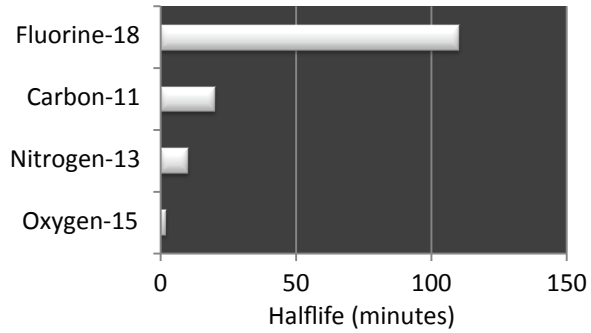
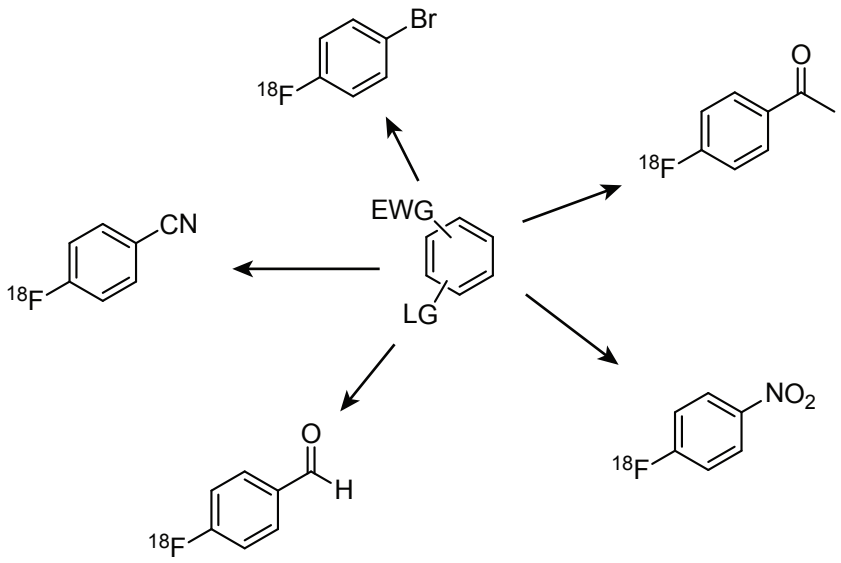



**Figure 3-4**. Halflifes of various radionuclides. $^{15}O$ and $^{13}N$ are rarely incorporated into complex molecules because their halflives are too short. $^{11}C$ can be incorporated into many molecules but must be made on site. $^{18}F$ has a sufficiently long halflife (110 min) that it can be made in bulk at a central site and shipped to nearby hospitals for use. A common example of such a product is $^{18}F$-fluorodeoxyglucose (FDG), which is the most used radiotracer for PET.

**Scheme 3-1.** Suitable Substrates for Aromatic Nucleophilic Substitution with $^{18}$F-Fluoride[a]




[a]EWG = Electron Withdrawing Group, LG = Leaving Group. Possible leaving groups include: nitro, trialkylammonium tosylate, halide, or iodonium salt.

ment we felt it would be more synthetically challenging to develop new methods for $^{18}$F incorporation into proteins.

There are relatively few reactions for incorporation of $^{18}$F into molecules. $^{18}$F-$F_2$ reacts with a wide range of substrates but the reaction suffers from a lack of selectivity and has problems with low specific activity because of the $^{19}$F-$F_2$ added during the incorporation.[7] Recently Hooker and Ritter have developed a palladium-based electrophilic fluorination reaction that utilizes the versatility of electrophilic fluorination while avoiding the poor specific activity of $^{18}$F-$F_2$.[8] However, the overwhelming majority of $^{18}$F incorporation has been through nucleophilic substition using $^{18}$F-$F^-$. Aliphatic nucleophilic substitution is utilized in the most used PET imaging agent, $^{18}$F-fluorodeoxyglucose. Because of its popularity that reaction has been optimized to give quite good yields. However, in general, aromatic nucleophilic substitution gives higher $^{18}$F-$F^-$ incorporation and as a result this is the most widely used strategy. The precursor molecules used for aromatic nucleophilic substitution consist of a leaving group (usually a nitro or trimethylaminium salt) and a strong electron withdrawing group in the *para* position. Scheme 3-1 shows several possible substrates for aromatic nucleophilic substitution with different electron-withdrawing groups. After $^{18}$F incorporation these electron-withdrawing groups can be converted to other chemical functionality of interest. These precursors are called "prosthetics" and are a strategy for conveying chemical functionality to the naked fluoride ion. As an example, Hooker demonstrated that *p*-nitrobenzaldehyde can react with $^{18}$F-$F^-$ to give [$^{18}$F]-*p*-fluorobenzaldehyde. The aldehyde can then be conjugated to an alkoxylamine modified protein through oxime formation (Figure 3-5a).[9] It would have been difficult to modify the protein directly with fluoride but the prosthetic provided a way to attach $^{18}$F to molecules through aldehyde functionality.

Several other prosthetic approaches have been developed to label proteins. As described in Chapter 1, most proteins are not stable to extreme temperatures and chemical conditions. Therefore the prosthetics used must fall into the categories of chemical protein modification reactions mentioned earlier. One approach utilized maleimide functionality to modify cysteines (Figure 3-5b). The prosthetic was synthesized by first incorporating $^{18}$F-$F^-$ into *p*-trimethylaminobenzaldehyde

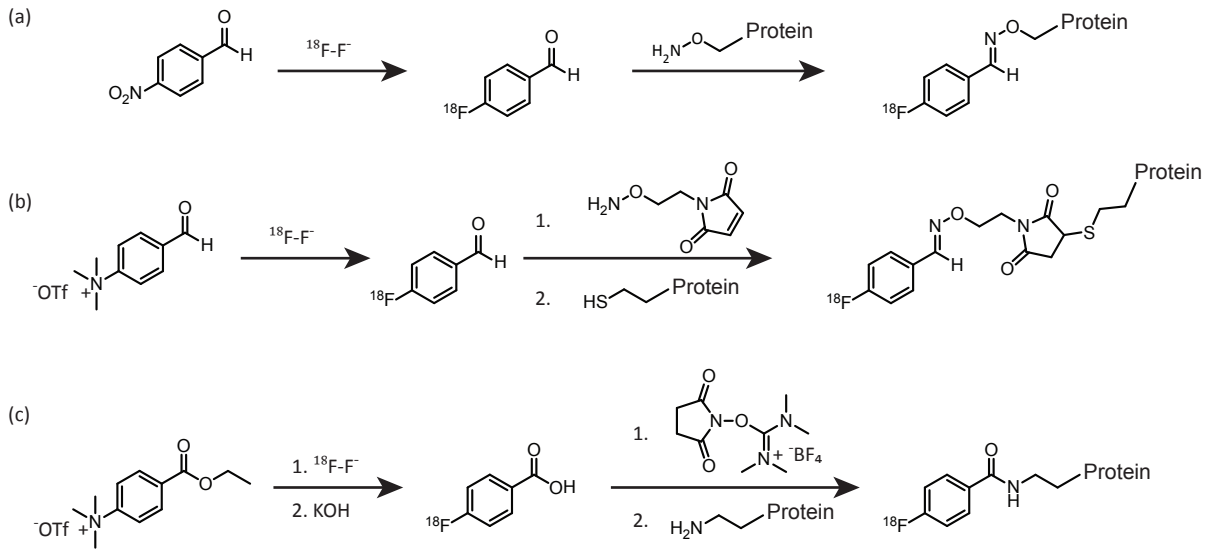



**Figure 3-5**. Synthesis schemes for various published $^{18}$F prosthetic groups. (a) $^{18}$F-fluorobenzaldehyde was synthesized from the corresponding nitrobenzaldehyde precursor, then it was coupled to proteins containing alkoxyamine functional groups. Hooker J.M. *et al.*, *Mol. Imaging. Biol.*, **2008**, *10*, 182-191. (b) $^{18}$F-fluorobenzaldehyde was synthesized from a different precursor, trimethylanilinium-benzaldehyde triflate, in this case. It was then formed an oxime with a maleimide-oxyamine linker and the resulting $^{18}$F-maleimide was conjugated to a protein containing thiols. Berndt, M., Pietzsch, J., Wuest, F., *Nucl Med Biol.*, **2007**, *34*, 5-15. (c) $^{18}$F-fluorobenzoic acid was synthesized from ethyl trimethylaniliniumbenzoate triflate. *O*-(*N*-succinimidyl)-*N*,*N*,*N'*,*N'*-tetramethyluronium tetrafluoroborate was used to convert the carboxylic acid to the corresponding succinimidyl ester which was then reacted with primary amines on proteins. Grierson, J.R. *et al.*, *Bioconjugate Chem.*, **2004**, *15*, 373-379.

triflate to form [$^{18}$F]-*p*-fluorobenzaldehyde. Then a maleimide-alkoxyamine bifunctional linker was added to form an oxime with the aldehyde. The resulting $^{18}$F-labeled maleimide ($^{18}$F-FBAM) was used to modify proteins rapidly under mild conditions.[10] Another approach utilized an activated ester to modify protein lysines. Similar to the previous case, [$^{18}$F]-*p*-fluorobenzoic acid was synthesized from the corresponding ethyl ester of the trimethylanilinium salt (Figure 3-5c). The carboxylic acid was then converted into a succinimidyl ester ($^{18}$F-FBA-OSu) which could be used to modify a wide range of amines including lysines.[11] While these approaches have been successfully used to modify proteins, they have several shortcomings.

The synthesis of the $^{18}$F-maleimide was completed in only 69 min, but the decay corrected yield was quite low, at 29%. The low overall yield was due to the poor yield of the oxime formation step. Usually oxime formation can be driven to higher yields by using high concentrations of reagents, but in the case of radiochemistry the radionuclides are in such low concentrations that only very efficient reactions will be successful. On the other hand, the synthesis of the $^{18}$F-succinimidyl ester gave an excellent 77% decay corrected yield, but required three synthetic steps and 118 min. In addition the reaction between succinimidyl esters and lysines suffers from the competing hydrolysis of the ester which renders that molecule useless for attaching $^{18}$F to the protein. Most radiochemistry syntheses fall into one of these two categories. Low yielding reactions make it difficult to produce enough labeled compound for experiments while long syntheses expose the chemist to more radiation and can also result in low final yields if using a radionuclide with a short halflife. Therefore, our goal was to develop an $^{18}$F prosthetic that is synthesized in a short time in high yield and has an efficient coupling reaction to its protein target.

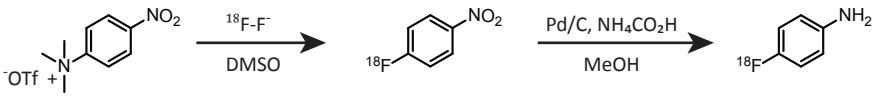
## Section 3.2: Synthesis of [$^{18}$F]-FA, a $^{18}$F Prosthetic Group for Use in Oxidative Coupling

The oxidative coupling reaction described in Chapter 2 seemed like the ideal reaction for incorporation of $^{18}$F into proteins. When choosing between the aniline or aminophenol for $^{18}$F incorporation, the aniline moiety was the clear first choice because of its simplicity. There is a clear retrosynthetic scheme from [$^{18}$F]-*p*-fluoroaniline ([$^{18}$F]-FA) back to either *p*-dinitrobenzene or *p*-nitro-*N*,*N*,*N*-trimethylanilinium triflate (Scheme 3-2). To ensure that *p*-fluoroaniline behaved similarly to the methyl substituted aniline (*p*-toluidine) used in the model studies, those studies were repeated using *p*-fluoroaniline instead. Similar to the previous studies, 4-methyl-2-aminophenol and *p*-fluoroaniline were dissolved in acetonitrile then diluted into water buffered at pH 6.5. The final concentrations were sub-mM, not quite as low as the expected radionuclide concentration but as low as feasibly achievable in order to isolate enough material for NMR studies. Addition of sodium periodate in a ten fold excess to a rapidly stirring solution resulted in an immediate color change that became progressively darker and more red until reaching its final color at 30 s. Isolation of the major product resulted in $^{1}$H and $^{13}$C NMR spectra that matched the structure of the product isolated in the experiments described in Chapter 2 (Figure 3-6a,b). The only major difference was the splitting of the aromatic protons from the aniline. Fluorine-19 has a similar splitting constant to protons when *ortho* and a smaller splitting constant in the *meta* position. Therefore the protons adjacent to the fluorine on the aromatic ring appear as a triplet and the protons *meta* to the fluorine appear as a doublet of doublets. For futher confirmation, a crystal of the isolated major product was grown for x-ray diffraction analysis (Figure 3-6c). In addition, the HPLC assays described in Chapter 2 were repeated with *p*-fluoroaniline and showed very similar results. The minor product again showed a later retention time upon addition of TCEP (Figure 3-6d). Therefore we felt comfortable assigning the products of this reaction as the analgous structures to those of **1** and **2** from Chapter 2 (Figure 2-3).

After confirming that *p*-fluoroaniline gave the same products as *p*-toluidine, we set forth to synthesize [$^{18}$F]-FA. Work by Dante Romanini, Jacob Hooker, and Jim O'Neil led to the synthetic design shown in Scheme 3-2. Their retrosynthetic scheme backtracked through [$^{18}$F]-*p*-fluoronitrobenzene and began with either *p*-dinitrobenzene or *p*-nitro-*N*,*N*,*N*-trimethylanilinium triflate. They reasoned that HPLC purification could be avoided by using the *p*-nitro-*N*,*N*,*N*-trimethylanilinium triflate precursor because the unreacted portion could be easily removed using a cation exchange resin. They also designed the synthesis to end with the [$^{18}$F]-FA in a buffered aqueous solution so that it could be directly added to proteins. The synthesis was carried out in a lead-shielded "cave" to minimize radiation exposure.

The experimental setup has several goals. First, it should reduce the radiation exposure

**Scheme 3-2.** Synthesis of $^{18}$F-fluoroaniline from *p*-Trimethylaniliniumnitrobenzene[a]




[a]Conditions: $^{18}$F-F$^{-}$ was added to *p*-trimethylaniliniumnitrobenzene triflate dissolved in DMSO. The reaction was heated at 95 °C for 5 min. The $^{18}$F-nitrobenzene was purified away from the precursor by passage through a cation exchange column and trapped on a C18 cartridge. It was eluted from the C18 cartridge with a methanolic solution of ammonium formate into a vial containing palladium-on-carbon catalyst. The reduction was stirred vigorously and heated at 95 °C for 5 min. The catalyst was removed by filtration to give $^{18}$F-fluoroaniline.

of the chemist by minimizing time inside the cave, maximizing distance from the radiation, and keeping body parts outside of the cave as much as possible. Next, having an established experimental setup helps with reproducibility between runs. Finally, having a well designed setup can

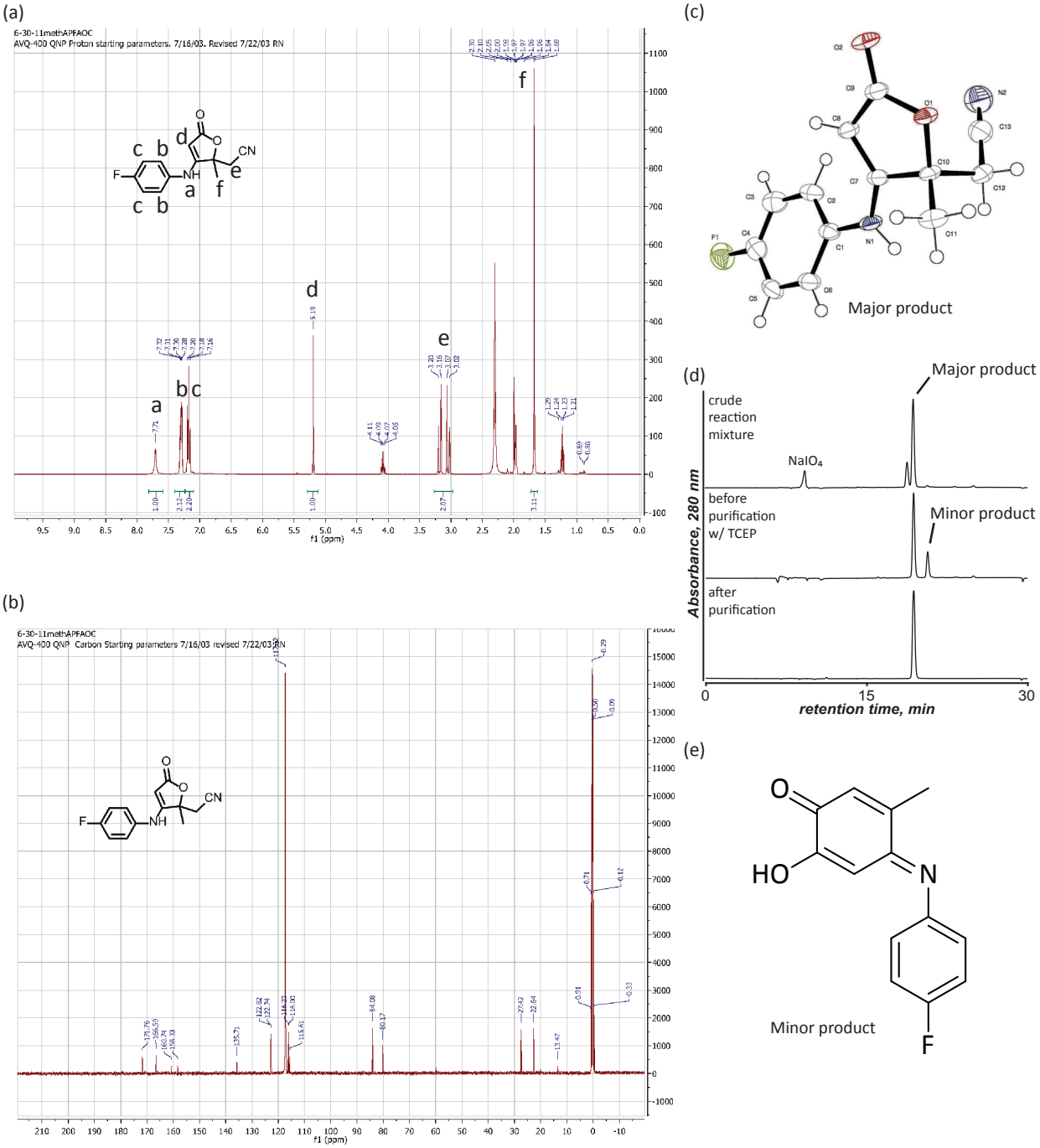



**Figure 3-6**. (a) $^{1}$H NMR spectra of the isolated major product of the reaction between 4-*tert*-butyl-2-aminophenol, *p*-fluoroaniline, and sodium periodate. (b) $^{13}$C NMR spectra of the isolated major product of the reaction between 4-methyl-2-aminophenol, *p*-fluoroaniline, and sodium periodate. (c) The major product was also characterized using X-ray diffraction. (d) Reversed-phase HPLC analysis of the crude reaction showed nearly identical behavior to the corresponding experiments in Chapter 2 (Figure 2-3) that used *p*-toluidine instead of *p*-fluoroaniline. Upon addition of TCEP the minor product changed retention time but the major product remained unchanged. The main difference is the ratio of major to minor product, as the fluoroaniline reaction showed a greater quantity of the minor product. This is likely caused by the electron withdrawing effect of the fluorine on the aniline. (e) Because the HPLC behavior is identical to the experiment in Figure 2-3, we feel comfortable assigning the minor product in this reaction the same structure as **2**.

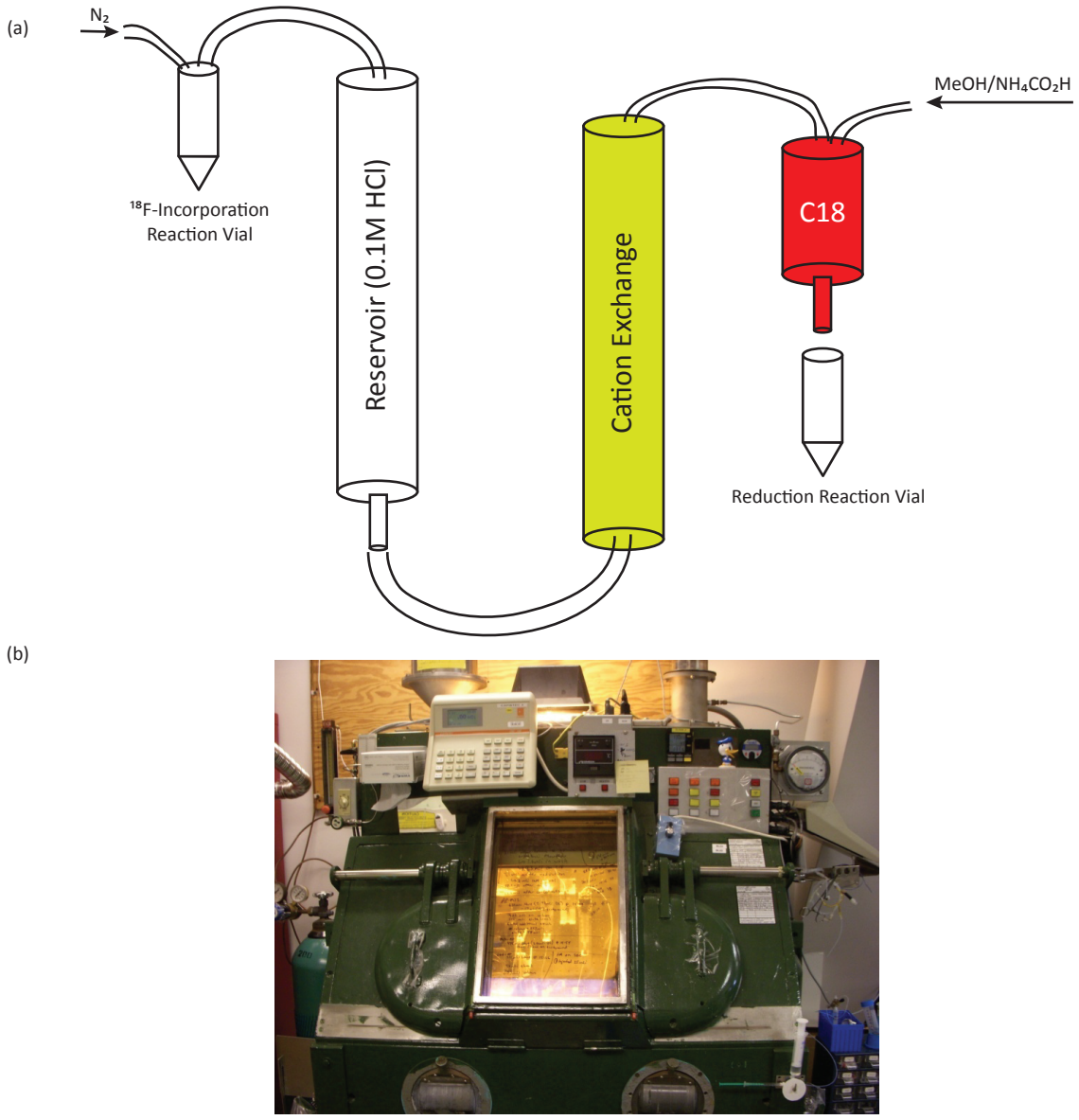



**Figure 3-7**. (a) Schematic of the purification of $^{18}$F-fluoronitrobenzene ([$^{18}$F]-FNB). Nitrogen gas is pushed into the sealed reaction vial which pushes the small reaction volume into a large reservoir filled with 0.1M HCl. This serves to dilute the DMSO so the product will be trapped on the C18 cartridge. Passage through the cation exchange column traps the unreacted precursor due to its positive charge, the [$^{18}$F]-FNB is trapped on the C18 cartridge and any unreacted $^{18}$F-fluoride passes through both columns and is discarded. A methanolic solution of ammonium formate is used to directly elute from the C18 cartridge, thus bypassing the cation exchange column. The eluted [$^{18}$F]-FNB is pushed directly into the reduction reaction vial which has been preloaded with a stir bar and Pd/C. These steps serve to minimize the radiation exposure during the synthesis. (b) Photograph of the lead cave that contains the experimental setup. Notice all of the electronics attached to the outside of the cave, as these serve to execute simple parts of the synthesis remotely so no body parts need to be inside the cave. Through the window the reservoir and purification columns depicted in the schematic are visible - they are mounted on the rear wall of the cave to maximize the distance between the radioactivity and the chemist when it is necessary to have hands inside the cave.

reduce the overall time for synthesis, thus reducing radiation exposure and finishing the reaction with more activity. A schematic of the setup is shown in Figure 3-7 along with a photograph of the lead-shielded cave used for the synthesis.

The first step for any $^{18}$F synthesis is making the actual $^{18}$F. This was accomplished by bombardment of $^{18}$O enriched water with protons. Because the nuclear reaction $^{18}$O(p,n)$^{18}$F is efficient,

relatively low energy protons from small medical cyclotrons are capable of producing significant amounts of $^{18}F$.[12] At the end of bombardment the $^{18}F$ is in aqueous solution. The large energy of association between water and fluoride ion make the fluoride a poor nucleophile. Two steps are performed to increase the nucleophilicity of the $^{18}F$-fluoride. First, a mixture of the cryptand Kryptofix K2.2.2 (Figure 3-8) and potassium bicarbonate is added to deprotonate the fluoride and give it a diffuse postive counterion to lessen the energy of association. Also, as much of the water from solution is removed as possible. We accomplished this through azeotropic drying with acetonitrile under vacuum, nitrogen gas flow, and heat. The resulting $[K/K2.2.2]^+[^{18}F]^-$ was ready for a nucleophilic substitution reaction.

Next, the precursor, *p*-nitro-*N*,*N*,*N*-trimethylanilinium triflate (synthesis and characterization shown in Experimental section), was dissolved in DMSO and added to the $[K/K2.2.2]^+[^{18}F]^-$. After 5 min reaction at 100 ºC, the $[^{18}F]$-*p*-fluoronitrobenzene ($[^{18}F]$-FNB) required purification before proceeding to the reduction step. As shown in the experimental setup in 3-7, a 25 mL reservoir was connected to a homemade cation exchange column packed with AG 50W-X8 resin. Connected in serial after the cation exchange column was a C18 Sep-Pac Plus (Waters). These items were attached to the back wall of the cave to minimize exposure during purification. The reservoir was filled with 0.1 M HCl prior to the experiment and served to dilute the DMSO such that the $[^{18}F]$-FNB could be trapped on the C18 cartridge. After the reaction mixture was pushed into the reservoir using nitrogen gas, it was then pushed through the cation exchange column and C18 cartridge. The unreacted $[^{18}F]$-$F^-$ passed through both purification devices while the positively charged precursor was trapped on the cation exchange resin. Only the desired product, $[^{18}F]$-FNB, was trapped on the C18 cartridge. The decay-corrected radiochemical yield of the $[^{18}F]$-$F^-$ incorporation was 70-80%.

The next reaction is a reduction of the nitro group of $[^{18}F]$-FNB. There are many reactions that have been shown to work for this transformation, but work by Dante Romanini found that palladium-on-carbon (Pd/C) along with ammonium formate gave quantitative conversion to the desired product in 5 min. Therefore the $[^{18}F]$-FNB was eluted from the C18 cartridge with 10 mg/mL ammonium formate in methanol into a reaction vial preloaded with a stir bar and 2 mg of Pd/C. After 5 min of reaction at 100 ºC with vigorous stirring the reaction was passed through a nylon filter with glass pre-filter to remove the Pd/C catalyst. The vial was washed with methanol and a small amount of dilute aqueous sulfuric acid. It was necessary to acidify the solution prior to evaporation of the methanol because the $[^{18}F]$-FA was quite volatile if not protonated. After removal of methanol the $[^{18}F]$-FA remained in the small volume water added in the previous step. After adding phosphate buffer and 6 M sodium hydroxide the $[^{18}F]$-FA was ready to use in a pH 6.5 phosphate buffered solution. The final product was achieved in 60-70% decay corrected radiochemical yield in 45 min from the beginning of synthesis. The radiochemical purity of the final product was always greater than 95% according to reversed-phase HPLC (Figure 3-9).

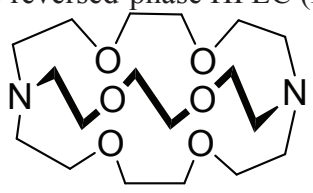



**Figure 3-8**. Structure of Krytofix 222, the kryptand used for sequestering the potassium cation. This minimizes the association energy between the cation and anion which enhances the nucleophilicity of fluoride ion.

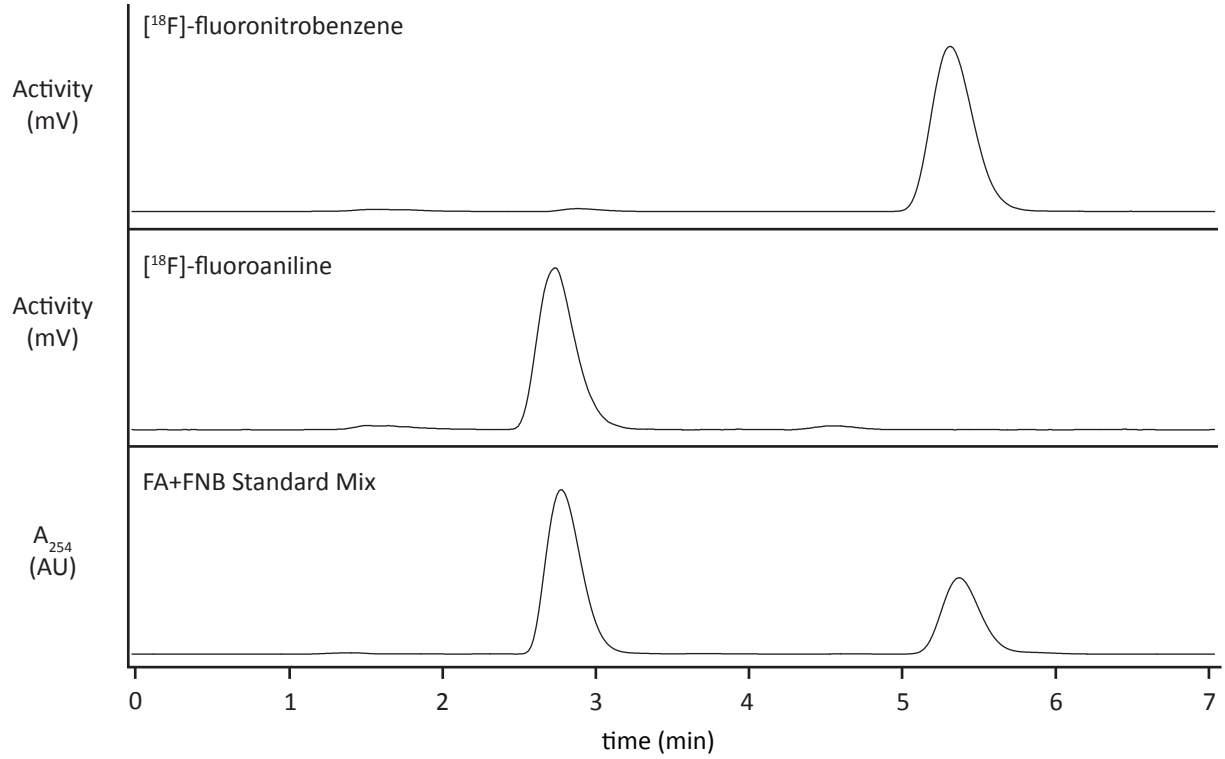
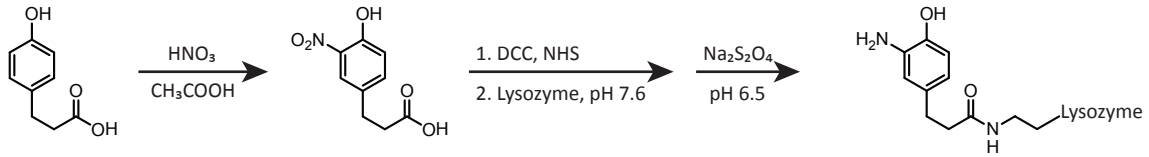



**Figure 3-9**. The synthesis was monitored after each reaction by reversed phase HPLC. An isocratic flow with 50:50 MeCN:0.12 M ammonium formate was run on a 100 x 4.6 mm C18-Luna column with a 1 mL/min flowrate. The top two traces are of the synthesized radioactive molecules and measure only radioactivity. The desired products comprised the majority of the radioactivity in the samples, and they were always of greater than 95% radiochemical purity. The mixture of the two standards in the bottom trace was visualized by UV absorbance at 254 nm.

## Section 3.3: Labeling Proteins with [$^{18}$F]-FA

To couple a protein to [$^{18}$F]-FA, it was necessary to construct a protein-aminophenol. As in Chapter 2, lysozyme was first used as a trial protein. Lysozyme has multiple lysine side chains that can be modified using succinimidyl esters. To install aminophenol functional groups on lysozyme, a succinimidyl ester containing an *o*-nitrophenol group was synthesized and added to the protein solution without purification (Scheme 3-3). The protein was purified with a Nap-5 size exclusion column (GE Healthcare). To the purified lysozyme-nitrophenol protein was added fifty equivalents of sodium dithionite to reduce the nitrophenol to the desired aminophenol. After another Nap-5 purification, MALDI-TOF MS analysis of the protein showed the expected multiple aminophenol

**Scheme 3-3.** Installing Aminophenol Functionality on Lysozyme[a]




[a]Conditions: Fuming nitric acid was added dropwise to a stirring solution of 3-(4-hydroxyphenyl)propionic acid in acetic acid on ice. Because acetic acid freezes at 4 °C it was necessary to remove the solution from the ice periodically. After 1 hr reaction the solution was diluted in a large quantity of water and crystals of the nitrated product formed overnight. Activation of the carboxylic acid was achieved with a 15 min reaction using DCC and NHS. The urea precipitate was filtered off and the succinimidyl ester was added directly to the lysozyme solution without purification. After 30 min reaction in pH 7.6 phosphate buffer, the modified lysozyme was purified on a Nap-5 column. The lysozyme-nitrophenol was then reduced to the aminophenol by addition of sodium dithionite. After 2 min reaction the lysozyme was again purified via Nap-5.

modifications (Figure 3-10a-b).

Both lysozyme-aminophenol and lysozyme-nitrophenol were exposed to [$^{18}$F]-FA in the presence of 5 mM sodium periodate at pH 6.5 for 2 min. The lysozyme-aminophenol showed excellent incorporation of $^{18}$F as determined by Nap-5 purification and comparing the activity of the eluent (protein fraction) to the activity of the column (small molecule fraction). To verify that the activity was actually incoporated into the protein, the two samples were analyzed via SDS-PAGE. Once the gel had sufficiently separated the proteins, it was exposed to a phosphorimaging plate to capture a picture of the radioactive spots on the gel. The lysozyme-aminophenol lane showed a strong signal at the expected mass of lysozyme (14 kDa) and the lysozyme-nitrophenol lane showed almost no radioactivity. The corresponding Coomassie stained gel showed equivalent amounts of protein in the two lysozyme lanes (Figure 3-10c). There was a small amount of activity (<5%) in the lysozyme-aminophenol sample at the mass corresponding to a lysozyme dimer. Unreacted lysozyme-aminophenol showed the same behavior when analyzed on a gel with Coomassie stain. A small amount of the protein either dimerized while stored in solution or alternatively heating the sample prior to SDS-PAGE analysis caused dimerization. Nonetheless this experiment showed that proteins can be efficiently labeled with [$^{18}$F]-FA when they contain aminophenols but not when there are no aminophenols present.

This successful proof of principle experiment was sufficient to continue on to more inter-

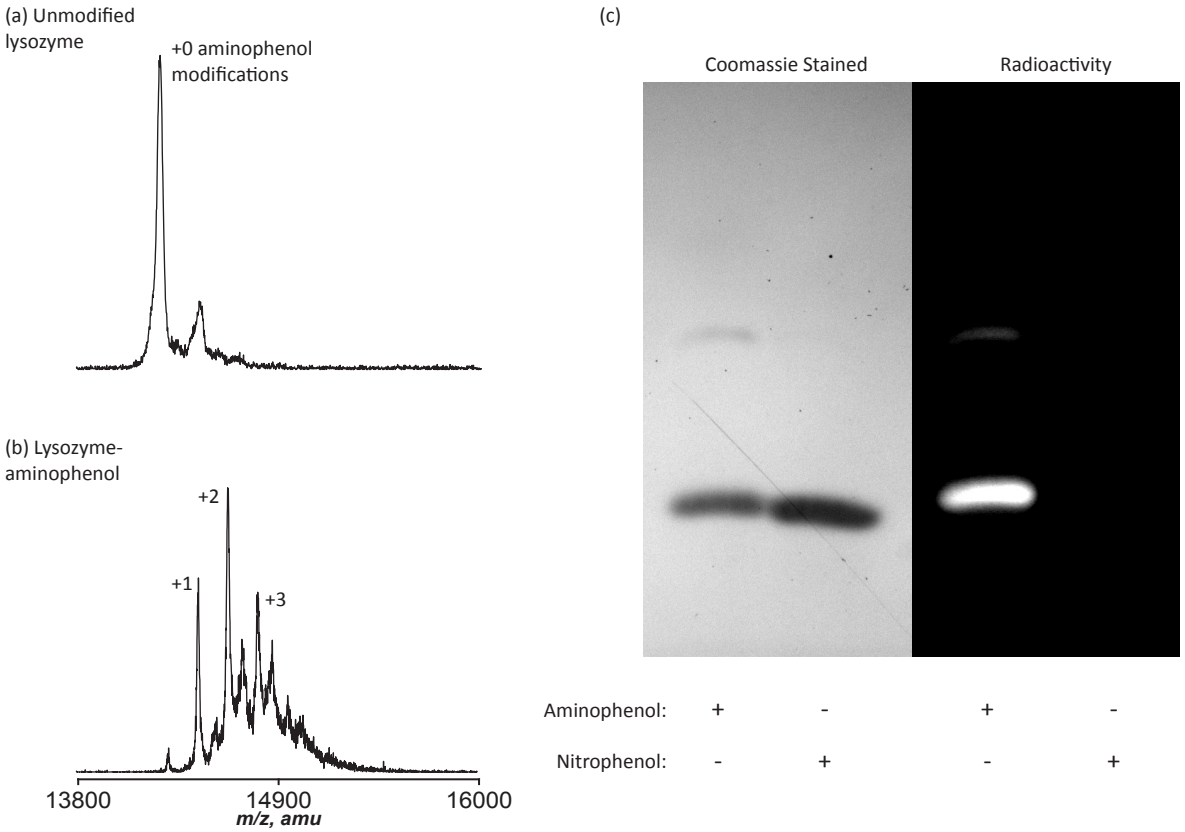



**Figure 3-10**. Modification of lysozyme with $^{18}$F-fluoroaniline. (a) MALDI-TOF MS of unmodified lysozyme using sinapinic acid as the matrix. (b) After modification of lysozyme with 3-(4-hydroxy-3-nitrophenyl)propionic acid via succinimidyl ester and subsequent reduction to the aminophenol with sodium dithionite, MALDI-TOF MS showed multiple aminophenol modifications. (c) Lysozyme-nitrophenol and lysozyme-aminophenol were exposed to $^{18}$F-fluoroaniline and sodium periodate for 2 min before Nap-5 purification. The samples were analyzed via SDS-PAGE with the gel imaged for radioactivity before being stained with Coomassie dye. Lysozyme-aminophenol showed significant incorporation of radioactivity while lysozyme-nitrophenol showed little to none.

esting protein substrates. Since our long term goal was to make $^{18}$F-labeled MS2 capsids for use as breast cancer imaging agents, we next turned our focus to MS2 radiolabeling. Chapter 4 will describe using MS2 capsids as an imaging agent while this section will solely focus on $^{18}$F-labeling and characterizing the capsids.

Work by Dante Romanini showed that the MS2-aminophenol-85 capsids described in Chapter 2 (Figure 2-6) could be $^{18}$F-labeled using this method but there were significant problems with reproducibility between batches of MS2. This likely occured because the two reactions used to install the aminophenol did not go to completion so there was a varied conversion to aminophenol in each batch. The deamination side reaction caused by the diazonium salt made it difficult to accurately determine the percent conversion to aminophenol. In search of a more reliable method for aminophenol installation, we turned to maleimide chemistry. MS2 capsids have native cysteines but they are not solvent accessible and therefore are not easily modified. A non-native cysteine was mutated in at position 87 and the N87C MS2 capsids were expressed in high yield and were shown to be modified to completion exactly once with maleimide reagents.[13] The proposed method for aminophenol installation was similar to the one used to label lysozyme in the previously described experiment. Instead of making a succinimidyl ester, a maleimide was attached to *o*-nitrophenol functionality. Briefly, tyramine was nitrated, then *o*-nitrotyramine was reacted with *N*-ε-(malemidocaproyloxy) succinimide ester to give the desired product. The detailed synthetic procedure and NMR characterization can be found in the Chapter 3 Experimental section. The nitrophenol-maleimide was dissolved in DMSO and 20 equivalents were added to N87C MS2. After 1 h reaction at pH 6.5, the MS2 was purified via Nap-5 column and analyzed with MALDI-TOF MS. There appeared to be quantitative conversion to the singly modified product. The MS2-nitrophenol was then exposed to ten equivalents of sodium dithionite for 2 min at pH 6.5 before purification via Nap-5 column. MALDI-TOF MS analysis again showed near quantitative conversion to the aminophenol product (Figure 3-11).

With a reliable and reproducible method for installing aminophenols on MS2, we then attempted to repeat the $^{18}$F labeling in hopes that it would also be more reproducible. In pH 6.5

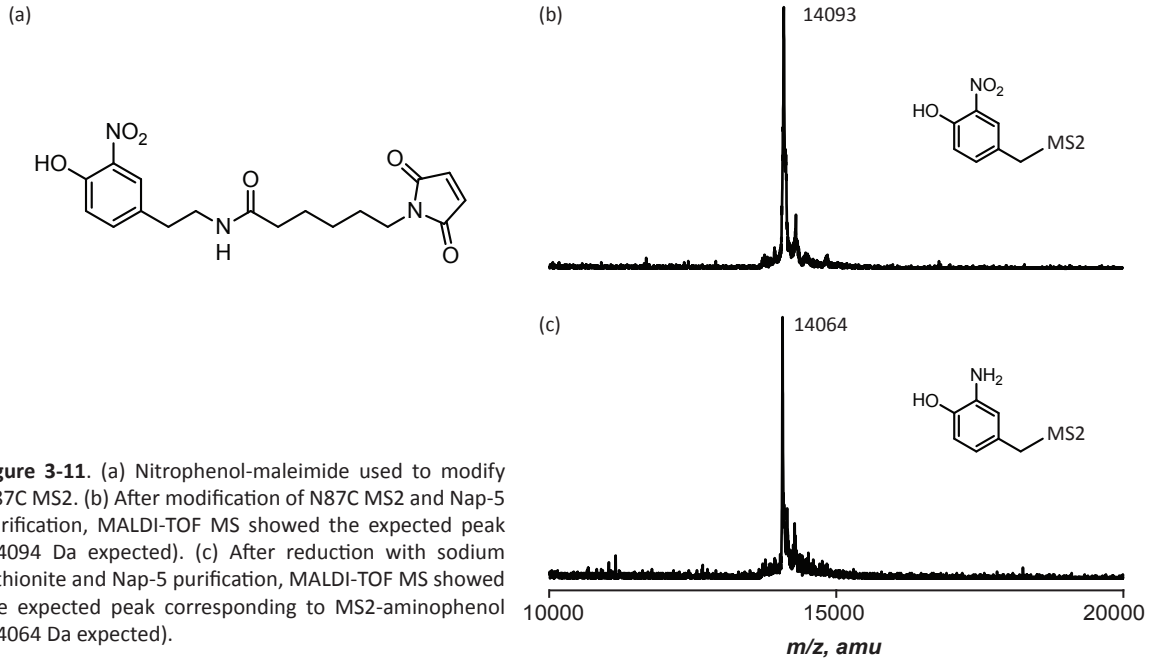



**Figure 3-11**. (a) Nitrophenol-maleimide used to modify N87C MS2. (b) After modification of N87C MS2 and Nap-5 purification, MALDI-TOF MS showed the expected peak (14094 Da expected). (c) After reduction with sodium dithionite and Nap-5 purification, MALDI-TOF MS showed the expected peak corresponding to MS2-aminophenol (14064 Da expected).

buffered phosphate, a small amount of [$^{18}$F]-FA (200 µCi) was added to 40 µL of 100 µM of MS2-aminophenol, then sodium periodate was added to a final concentration of 5 mM. After the standard 2 min reaction time, the protein was purified via Nap-5 to remove unreacted [$^{18}$F]-FA. Size exclusion chromatography of the reaction before and after purification showed that the Nap-5 did an excellent job of removing unreacted [$^{18}$F]-FA (Figure 3-12). Measuring the radioactivity in the protein elution fraction versus the activity remaining on the Nap-5 column (small molecule fraction) gave a rough estimate of the radiochemical yield. Since the peak areas on the SEC runs shown in Figure 3-12 correlated with the Nap-5 calculation, the Nap-5 method was thereafter used to report radiochemical yield for protein modification reactions. The above reaction gave 50% radiochemical yield, meaning that 50% of the radioactivity was incorporated into the desired product.

In an attempt to improve the radiochemical yield, the [$^{18}$F]-FA was HPLC purified at the end of the synthesis. We had hypothesized that residual starting material, *p*-nitro-*N*,*N*,*N*-trimethylanilinium triflate, could be carried through the synthesis and after reduction would be a competing aniline for the aminophenol groups on the protein. HPLC purification removed all anilines besides *p*-fluoroaniline. The above labeling experiment was repeated with an identical amount of radioactivity. The experiment with HPLC purified [$^{18}$F]-FA gave a much improved 87% radiochemical yield with 2 min of coupling time.

At this point, we attempted to calculate the specific activity of the protein. Specific activity

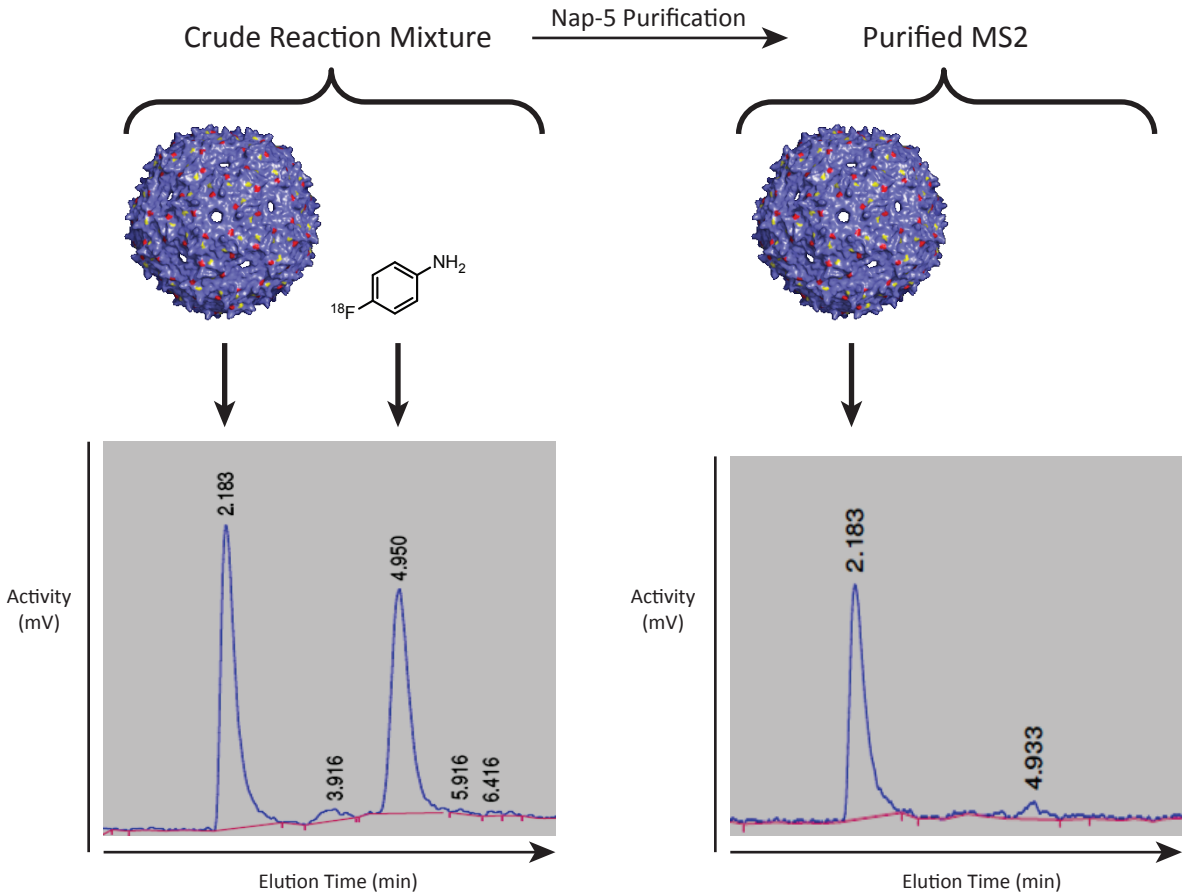



**Figure 3-12**. Size exclusion chromatography HPLC traces show that Nap-5 purification successfully separates radiolabeled MS2 away from unreacted [$^{18}$F]-FA. 20 µL injection on 250 x 4.6 mm Zorbax GF-250 column (Agilent) running at 1 mL/min with mobile phase of 5 mM phosphate buffer pH 6.5.

| Reaction | Amount $^{18}$F-FA (mCi) | Calculated Amount $^{18}$F-FA (pmol) | Calculated Amount Total FA (nmol) | Limiting Reagent | Radiochemical Yield (%) | $^{18}$F/capsid ratio |
|---|---|---|---|---|---|---|
| 1 | 0.2 | 0.12 | 0.12 | FA | 87 | 1/218 |
| 2 | 2.0 | 1.2 | 1.2 | FA | 80 | 1/24 |
| 3 | 20 | 12 | 12 | MS2 | 22 | 1/9 |

**Figure 3-13**. Each reaction contained 4 nmol of MS2-aminophenol monomer (22 pmol MS2 capsid). Total fluoroaniline (FA) amount was calculated based on the specific activity of [$^{18}$F]-FA (approximately 2mCi/nmol or 1 $^{18}$F per 1000 total F). The specific activity of the capsids illustrates the benefit of labeling a multimeric protein - in reaction 3 when MS2 is the limiting reagent the resulting specific activity is more than 100 times greater than [$^{18}$F]-FA (1/9 versus 1/1000).

is the percentage of radioactive particles compared to the total number of particles. Specific activity of the radiotracer is important especially in cases where a specific receptor is being targeted. If the specific activity is too low, then the receptor will be saturated with non-radioactive compound, thus making it difficult to image the desired target. A major benefit of using MS2 capsids is their multivalency. Because 180 aminophenol groups are installed on the inside of the capsids, there is 180 times the chance of incorporating a radioactive *p*-fluoroaniline. There are two specific activity measurements that will be discussed. First is the specific activity of the prosthetic group, [$^{18}$F]-FA. Because of the natural level of $^{19}F^-$ present in water, the typical specific activity of radiofluorinated compounds is 2-10 mCi/nmol which corresponds to approximately 1:1000 $^{18}$F:$^{19}$F.[12] The specific activity of the synthesized [$^{18}$F]-FA was measured by quantifying total *p*-fluoroaniline concentration using HPLC and standards. These values matched up so the 1:1000 approximation was used in calculations.

As mentioned above, the specific activity of MS2 capsids should be 180-fold higher than [$^{18}$F]-FA because of the multivalency effect. Based off the 1:1000 approximation of [$^{18}$F]-FA, the capsids should have a $^{18}$F/capsid ratio of approximately 1:5. However, when the above sample's specific activity was measured, it resulted in a $^{18}$F/capsid ratio of 1:218. This ratio corresponded to only a 5-fold increase due to the multivalency of the capsids. We then realized that the MS2 capsids were in excess in this reaction, and were therefore unnecessarily lowering the specific activity. To test this idea, an equal amount of MS2 capsids were reacted with increasing amounts of [$^{18}$F]-FA until the capsids were the limiting reagent in the reaction. The results are shown in Figure 3-13. There are two clear trends evident. When the excess of MS2 capsids is decreased the radiochemical yield also decreases. Also, when an excess of [$^{18}$F]-FA is used, then the specific activity is much higher. In this case (Figure 3-13, reaction 3), a $^{18}$F/capsid ratio of 1:9 was achieved, which was a 110-fold improvement over the specific activity of [$^{18}$F]-FA. Of course the downside of these conditions is the much lower 22% radiochemical yield, which is expected with so little capsid present. These experiments show that the conditions can be catered based on the desired end product. If a large amount of radioactivity is desired but specific activity is not as important (a very abundant receptor or target, for example), then conditions similar to Reaction 1 can be used. Alternatively, if high specific activity is desired, then conditions similar to Reaction 3 can be used.

Another experiment was performed to explore the specific activity of this system. A series of reactions were set up with an identical amount of MS2 capsid and [$^{18}$F]-FA in each. Then differing amounts of non-radioactive $^{19}$F-*p*-fluoroaniline were added to artificially lower the specific activity of the [$^{18}$F]-FA. As expected, the radiochemical yield decreased as more $^{19}$F-*p*-fluoroaniline was added because more of the aminophenol reactive groups on the protein were consumed by the non-radioactive *p*-fluoroaniline (Figure 3-14).

Having characterized the $^{18}$F-labeling of MS2 capsids and explored the multivalency effect,

the next step was to create agents for *in vivo* imaging. This will be further discussed in Chapter 4.

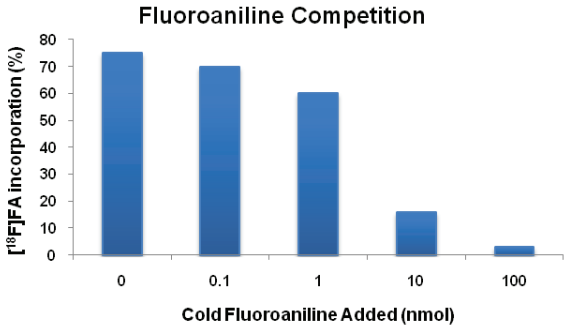



**Figure 3-14**. Varying amounts of non-radioactive (cold) fluoroaniline were added to five identical coupling reactions between MS2-aminophenol and [$^{18}$F]-FA. The percent incorporation of radioactivity into the MS2 was measured for each reaction. The graph shows that [$^{18}$F]-FA contained between 1-10 nmol of anilines.

## Section 3.4: Fluorine-18 Labeling of Peptides

While the labeling method described here is more important for protein modification because of the lack of alternatives, it can also be used to label small peptides. Peptides are generally easier to label because they are stable to more extreme conditions. The ease of synthesis of [$^{18}$F]-FA and its efficient coupling to proteins led us to try labeling a peptide as well. Through a collaboration with Lawrence Berkeley National Laboratories, a collaborator looking to study heart hypertrophy was interested in labeling a peptide that bound matrix metalloproteinases (MMP's). High levels of MMP's are indicative of repair and expansion of the extracellular matrix which is a symptom of heart hypertrophy. Identifying heart hypertrophy at an early stage could be an early diagnosis for heart disease.

We found several references that suggested this idea was feasible. First, a paper describing a phage display library search for MMP binding proteins provided a 10 amino acid cyclic peptide sequence (CTTHWGFTLC) that was a potent inhibitor of MMP-2 and MMP-9.[14] Another reference showed that an N-terminal extension provided a handle for chemical modification but the peptide retained proper function.[15] The sequence from the second reference was changed slightly and will be referred to as CTT-2. To introduce aminophenol functionality to the peptide, a 6-carbon spacer (6-aminocaproic acid) was added to the N-terminus, followed by modification with the nitrophenol-succinimidyl ester described previously (Scheme 3-3). The peptide was cleaved from the resin and isolated via ether precipitation. Reduction with sodium dithionite followed by HPLC purification afforded CTT-2-aminophenol. The peptide contains a disulfide bond to force it into a cyclic conformation. The formation of this disulfide was achieved by dissolving the peptide in 20% DMSO, pH 4.5, and incubating at room temperature overnight. After dilution with 0.1% TFA the peptide was purified on a C18 cartridge, then purified via HPLC to give the final product. The peptide was then characterized with MALDI-TOF MS (Figure 3-15a). A mass change of -32 Da was observed between CTT-2-nitrophenol and the final product (Figure 3-15c). This corresponded to the -30 Da change from the nitro group reduction and the -2 Da change from the disulfide bond

formation. To test if the peptide was suitable for radiolabeling, it was first coupled to nonradioactive *p*-fluoroaniline in the presence of sodium periodate and HPLC purified. MALDI-TOF MS analysis of the product showed two peaks that corresponded to the two oxidative coupling products (Figure 3-15b). In this case, the imine hydrolysis product previously found to be the minor product was instead the major product (Figure 3-6 and Figure 2-3). Further experiments are necessary to determine what factors affect the ratio of the two products.

Radiolabeling of CTT-2 loosely followed the procedure used for the MS2 capsids. In an effort to reduce the total time of synthesis, the [$^{18}$F]-FA was not HPLC purified before use. The other key difference was the method of purification after labeling. As a small peptide, CTT-2 could be purified under denaturing conditions. Acidifying the crude reaction before addition to a C18 cartridge (Waters) caused the unreacted [$^{18}$F]-FA to pass through while the peptide was trapped (Figure 3-16). This method of purification allowed for elution into a small amount of ethanol. After dilution with saline this method afforded a facile way to prepare large amounts of labeled peptide

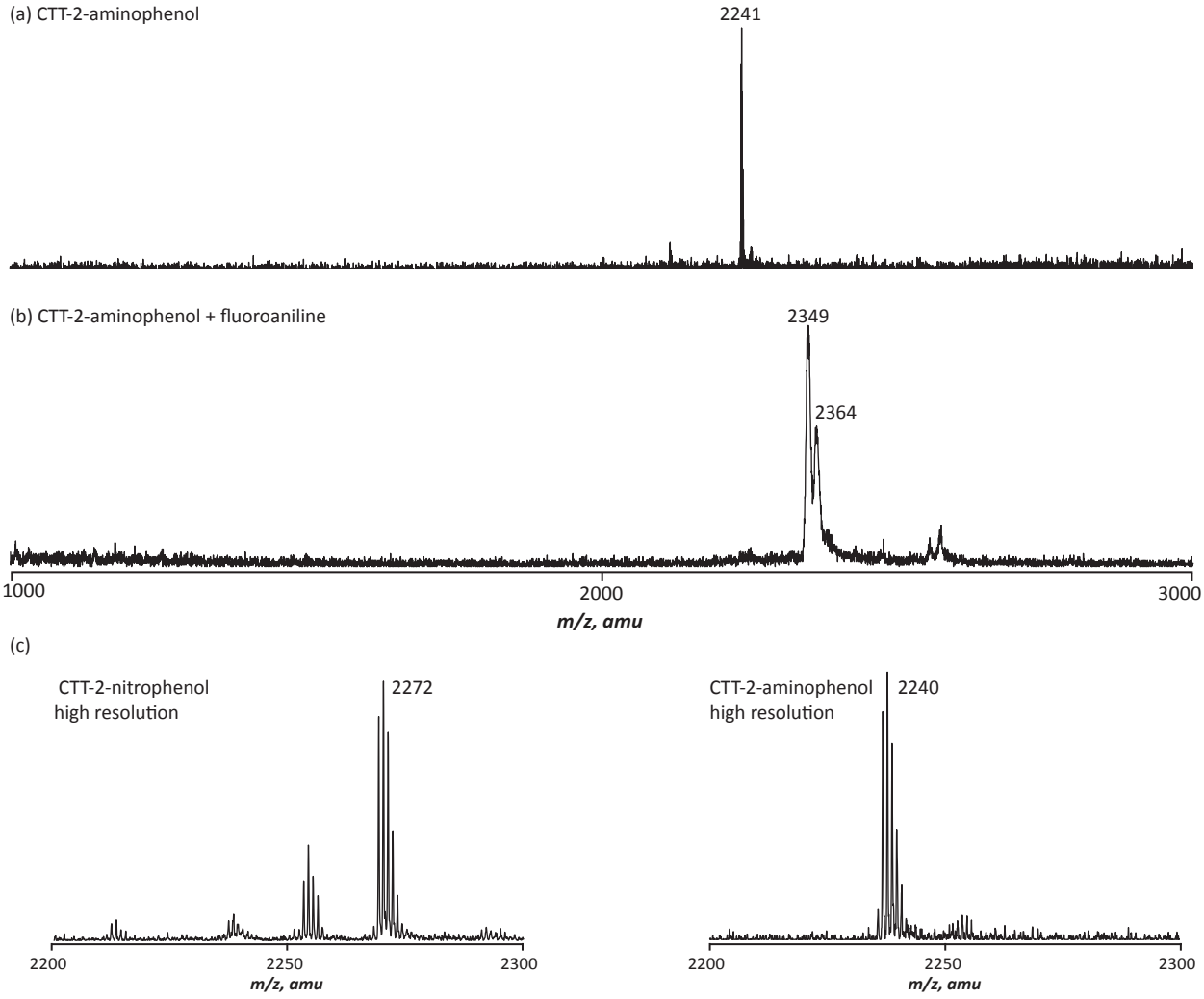



**Figure 3-15**. MALDI-TOF MS of CTT-2 peptide. All samples were co-spotted with sinapinic acid as the matrix. (a) Purified CTT-2-aminophenol after disulfide bond formation. Expected $[M+H]^+$ = 2241. (b) HPLC purified conjugate of CTT-2-aminophenol and p-fluoroaniline after 2 min reaction in the presence of sodium periodate. Expected $[M+H]^+$ = 2349, 2364. (c) High resolution MALDI-TOF MS of CTT-2-nitrophenol and CTT-2-aminophenol showed a mass difference of -32 Da, corresponding to the nitro group reduction (-30 Da) and disulfide bond formation (-2 Da). Because the isotope patterns of the two look identical, one can conclude that there is only one species present, thus suggesting that the disulfide bond formation is 100% complete.

in a volume small enough for injection into an animal. In preparation for an animal experiment, 12 mCi of [$^{18}$F]-FA was added to the peptide followed by sodium periodate to a final concentration of 5 mM. After purification, 8 mCi of labeled peptide was isolated (65% radiochemical yield), diluted in saline, and passed through a 0.2 μM filter for sterilization. The purity of the sample was verified by HPLC. This sample was transferred to the collaborator, who intended to split the sample for injection into one healthy rat and one rat genetically engineered to develop a hypertrophied heart. Unfortunately, the experiment could not be completed before the end of the animal model lifespan. This experiment will be repeated in the future.

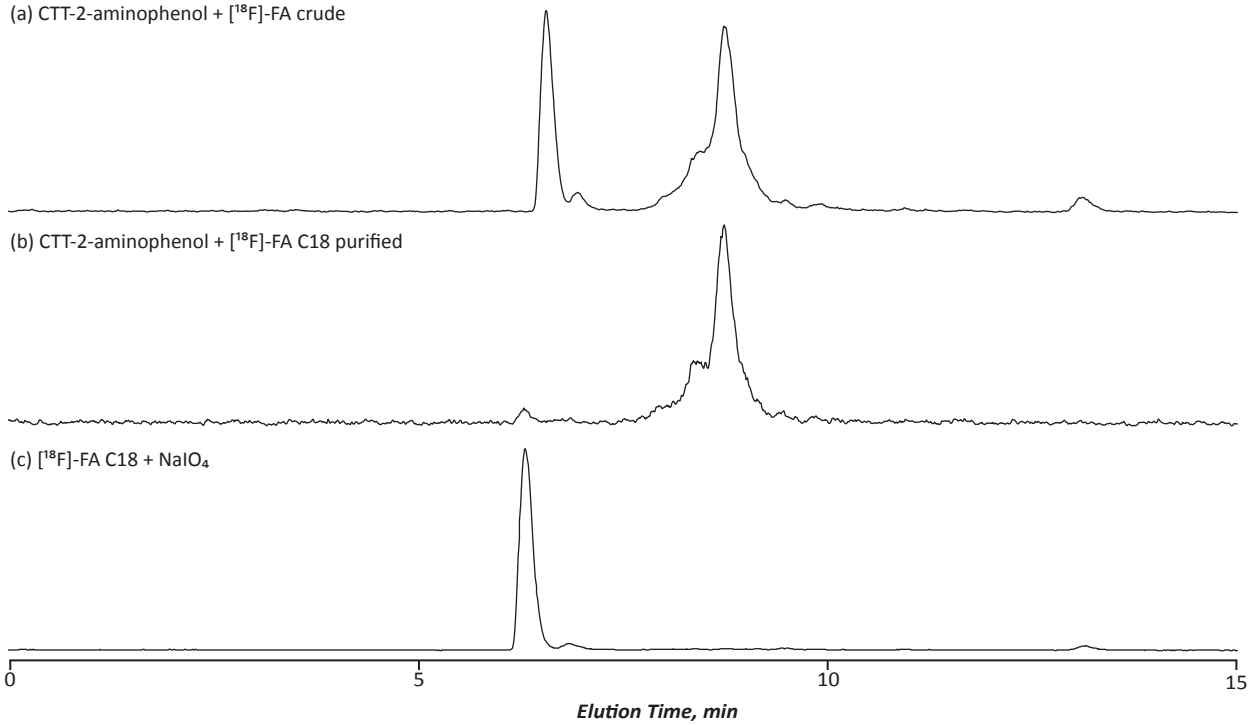



**Figure 3-16**. Analytical HPLC spectra measuring radioactivity. 20 μL injections on 250 mm x 4.6 mm C18-Gemini (Phenomenex). (a) Crude mixture after 2 min reaction between CTT-2-aminophenol (8 min elution) and [$^{18}$F]-FA (6 min elution) in the presence of sodium periodate. (b) After purification on C18 Sep-pak (Waters), the unreacted [$^{18}$F]-FA was removed. (c) [$^{18}$F]-FA was reacted with sodium periodate (no peptide) and injected to show where unreacted [$^{18}$F]-FA eluted.

## Section 3.5: Experimental Section

### Characterization of Products from Oxidative Coupling of *p*-Fluoroaniline and 2-Amino-4-methylphenol

2-Amino-4-methylphenol (123 mg, 1 mmol) and *p*-fluoroaniline (111 mg, 1 mmol) were dissolved in two separate portions of acetonitrile (2 mL each) and then added to 2 L of 10 mM phosphate buffer, pH 6.5. Sodium periodate (2.14 g, 10 mmol) was dissolved in 20 mL of phosphate buffer and then added in one portion with vigorous stirring. The solution immediately turned red. After 5 min, sodium chloride was added to 1 M and the solution was extracted with 1 L portions of ethyl acetate until the organic layers showed no more color. The combined organic layers were washed with brine and dried over sodium sulfate before removing the solvent under reduced pressure. The resulting solid was purified via silica gel chromatography using a mobile phase of

1:1 ethyl acetate:hexanes ($R_f \approx 0.25$). $^1$H NMR (400 MHz, $CD_3CN$) δ 1.84 (s, 3H), 3.05 (d, *J*=20 Hz, 1H), 3.18 (d, *J*=16 Hz), 5.19 (s, 1H), 7.18 (dd, *J*=8 Hz, 2H), 7.30 (dd, *J*=8 Hz, 2H), 7.71 (s, 1H). $^{13}$C NMR (100 MHz, $CD_3CN$) δ 22.64, 27.42, 80.17, 84.08, 115.61, 116.00, 116.23, 122.74, 122.82, 135.71, 158.33, 166.59, 171.76. Spectra can be found in Figure 3-6.

For crystal formation and HPLC analysis of these compounds, the methods described in the Chapter 2 experimental section were followed. Crystal structure and HPLC traces can be found in Figure 3-6. Crystallography details can be found at end of this experimental section.

**Synthesis of Radiolabeling Precursor, *p*-Nitro-*N*,*N*,*N*-trimethylanilinium Triflate**

To a flame dried flask containing a solution of 4-dimethylaminonitrobenzene (500 mg, 3 mmol) in distilled methylene chloride was added one equivalent of methyl triflate (493 mg, 3 mmol). After stirring overnight, the product was filtered off and washed with methylene chloride. 400 mg were recovered for a yield of 40%. $^1$H NMR (400 MHz, $CD_3CN$) δ 3.6 (s, 9H), 8.0 (d, 2H), 8.4 (d, 2H).This procedure is slightly different from the published literature procedure.[16]

**Synthesis of *o*-Nitrophenol Protein Modification Reagents**

*Synthesis of* o-*Nitrophenol Succinimidyl Ester*

Dicyclohexyl-carbodiimide (27 mg, 0.13 mmol) was added to a flask containing 3-(4-hydroxy-3-nitrophenyl)propionic acid (23 mg, 0.11 mmol, synthesized via literature procedure[17] and *N*-hydroxysuccinimide (13 mg, 0.11 mmol) dissolved in $CH_2Cl_2$. The reaction was stirred for 15 min at room temperature, after which the solution was filtered to remove the DCU precipitate. The solvent was removed under reduced pressure and a stock solution of 100 mM was prepared by dissolving the product in DMSO. The stock solution was frozen until used.

*Synthesis of* o-*Nitrophenol Maleimide*

To tyramine was added dropwise one equivalent of fuming nitric acid at 4 °C using trifluoroacetic acid as the solvent and this resulted in quantitative conversion to *o*-nitrotyramine. *O*-nitrotyramine (50 mg, 0.27 mmol) was dissolved in 10 mL DMF and treated with one equivalent of succinimidyl-6-*N*-maleimidohexanoate[18] along with sufficient triethylamine to reach pH 8. Multiple equivalents of triethylamine were required due to residual trifluoroacetic acid from the nitration step. After 45 min, 20 mL of 0.1M $NaHSO_4$ was added to the reaction. The product was extracted with methylene chloride, dried over $Na_2SO_4$, and purified on a silica column using ethyl acetate as the mobile phase. The isolated yield was 33%. A stock solution was prepared at 100 mM by dissolving the product in DMSO. This stock solution was stored frozen when not in use, and the reagent was still effective several years after synthesis. $^1$H NMR (300 MHz, CDCl3) δ 1.2 (m, 2H), 1.6 (m, 4H), 2.1 (t, 2H), 2.8 (t, 2H), 3.5 (t, 4H), 5.6 (br s, 1H), 6.6 (s, 2H), 7.0 (d, 1H), 7.5 (d, 1H), 8.0 (s, 1H), 10.5 (br s, 1H).

**Installing Aminophenol Functionality on Lysozyme**

To a 50 μM solution of lysozyme (from chicken egg white, Sigma) in 50 mM pH 8.5 phosphate buffer was added the *o*-nitrophenol succinimidyl ester to a final concentration of 1 mM. The resulting solution was allowed to react at room temperature for 1 h, at which time the lysozyme was purified via a Nap-10 size exclusion column equilibrated with pH 6.5 phosphate buffer. To the eluted protein solution was added sodium dithionite to a final concentration of 5 mM. After 2 min reaction the lysozyme was again purified via Nap-10. MALDI-TOF MS analysis indicated a range

of two to five aniline modifications on each lysozyme (Figure 3-10b).

**Preparation of N87C MS2 Capsids**

N87C MS2 capsids were prepared as described.[13]

**Labeling of N87C MS2 with *o*-Nitrophenol Maleimide and Subsequent Reduction**

To 980 μL of a 100 μM solution of N87C in 20 mM pH 6.5 phosphate buffer was added 20 μL of a 100 mM solution of *o*-nitrophenol maleimide. After 1 h at room temperature the MS2 was purified via Nap-10. To the resulting solution was added sodium dithionite to a final concentration of 1 mM. After 2 min the MS2 was again purified via Nap-10. MALDI-TOF MS showed a single modification (Figure 3-11).

**Radiochemical Synthesis**

Potassium(Kryptofix) [$^{18}$F]-fluoride was produced by azeotropic drying of target water with 0.5 mL of a solution of 10 mg/mL 4,7,13,16,21,24-Hexaoxa-1,10-diazabicyclo[8.8.8]-hexacosane (Kryptofix-222) and 5 mg/mL potassium bicarbonate in 10% $H_2O$/90% acetonitrile. To the resulting dry mixture was added precursor solutions in DMSO.

*[$^{18}$F]-4-fluoronitrobenzene ([$^{18}$F]-FNB)*

$K_{222}/K^{18}F$ was resuspended in 300 μL of a solution of *p*-nitro-*N*,*N*,*N*-trimethylanilinium triflate (0.8-1.2 mg) in anhydrous dimethylsulfoxide. The reaction vial was sealed, and the mixture was heated to 95 °C for 5 minutes. The reaction was diluted into a reservoir containing 20 mL of 0.1M HCl then the reaction vial was washed with 2 x 1 mL of water which was also added to the reservoir. Using nitrogen gas pressure, the reservoir solution was passed through a column of AG 50W-X8 cation exchange resin (Biorad, USA)(3 mm x 10 mm) and directly through a C18 Sep-Pak Plus reversed-phase cartridge (Waters). The columns were washed with 10 mL of 0.1M HCl, and the product was eluted from the Sep-pak with 2 mL of a 10 mg/mL solution of ammonium formate in methanol. Typical recovered yield was 70-80% (n = 10).

*[$^{18}$F]-4-fluoroaniline*

The methanolic solution of [$^{18}$F]-FNB was eluted from the Sep-Pak directly into a conical vial containing 2 mg 10% Pd/C and a stirbar. This vial was sealed and heated to 95 °C with stirring for 5 minutes. The vial was removed from heat, and the reaction mixture was filtered through an in-line 0.45 μm Whatman filter with glass prefilter into a test tube equipped with a distillation head. The reaction vial and filter were rinsed with a mixture of 2 mL of methanol and 300 μL of 1% $H_2SO_4$. The combined solutions were stripped of methanol under vacuum to provide crude [$^{18}$F]fluoroaniline in acidic solution. The pH of this solution was adjusted to 7 by addition of 400 μL of 1M $K_2HPO_4$ (pH 8). The resulting solution was then HPLC purified and the isolated fraction was acidified and solvent was removed as described above. Radiochemical purity was greater than 98% and the yield was typically 60-65% over two steps (n = 10). The time of synthesis starting from dry [$^{18}$F]-fluoride was 40 min without HPLC purification and 70 min with HPLC purification.

**Typical Protein Labeling Experiment with [$^{18}$F]-FA Using MS2 as Example**

*Labeling of MS2-aminophenol with [$^{18}$F]-FA*

Labeling with [$^{18}$F]fluoroaniline was accomplished by placing in a 1.5 mL conical centri-

fuge tube a solution of 100 μM aminophenol-containing MS2 (40 μL). To this solution was added 50 mM $K_2HPO_4$, pH 6.5 (10 μL) and the aqueous solution of [$^{18}$F]fluoroaniline (20-50 μL). A freshly prepared solution of 50 mM sodium periodate (10 μL) was then added and the reaction was incubated at room temperature for 2 minutes. The reactions were analyzed by SDS-PAGE and size exclusion HPLC.

Preparative reactions were scaled up accordingly and purified with NAP5 or NAP10 gel filtration columns. Yields were determined by measuring the radioactivity in the eluant of the column (containing protein) and that remaining in the column.

**Synthesis of CTT-2-aminophenol**

The CTT-2 peptide was synthesized on solid phase resin with a Rink Amide linker (EMD Biosciences). Standard Fmoc chemistry was used. Reactants added for each coupling step were 10 equivalents of amino acid, 10 equivalents of *o*-(6-chlorobenzotriazol-1-yl)-*N*,*N*,*N*′,*N*′-tetramethyluronium hexafluorophosphate (HCTU), and 20 equivalents of diisopropylethylamine (DIPEA). The first amino acid coupling to the Rink linker was incubated for 45 min at room temperature, then the resin was washed and fresh reagents were added for a second 45 min incubation. All other amino acid coupling steps used 10 min incubation times. 20% piperidine in DMF was used for deprotecting Fmoc groups. The deprotection reaction was incubated for 5 min then fresh deprotection solution was added and incubated for additional 2 min. Between all steps the resin was washed 3 times with DMF. The following amino acid sequence was synthesized using these conditions: $H_2$N-GRENFHGCTTHWGFTLC-Rink linker. 6-Fmoc-aminocaproic acid was then coupled using standard coupling conditions, followed by coupling of 3-(4-hydroxy-3-nitrophenyl) propionic acid. The coupling of the nitrophenol led to a significant amount of precipitation by the end of the reaction. The resin was then incubated with deprotection solution which dissolved the precipitate and it was washed away. The beads were also washed with methanol and DMF. The peptide was cleaved from the resin and the side chain protecting groups were deprotected using a 88% TFA, 5% $H_2O$, 5% phenol, and 2% triisopropylsilane (TIPS). After 2 h incubation the cleavage solution was collected and dripped slowly into cold ether to precipitate the peptide. The ether solution was centrifuged and the ether was removed. Resuspension of the peptide in fresh cold ether followed by another centifugation gave CTT-2-nitrophenol (Figure 3-15c).

CTT-2-nitrophenol was dissolved in 35:65 MeCN:$H_2O$ to a concentration of about 0.5 mg/mL, and the pH was adjusted to 6-7. 10 equivalents of sodium dithionite were added and after 2 min the solution was filtered and injected onto a semi-prep HPLC column (C18-Gemini, Phenomenex). The major fluorescence peaks (excitation 280 nm, emission 330 nm) were collected manually and analyzed with MALDI-TOF MS to determine fractions of interest. The fraction containing the desired CTT-2-aminophenol was adjusted to pH 5 using a concentrated sodium acetate buffer. To form the disulfide necessary for cyclizing the peptide, DMSO was added to 20% and the peptide was incubated at room temperature overnight.[19] The peptide solution was diluted 10-fold with 0.1% TFA and additional TFA was added if the pH was not below 2. This solution was purified on a large C18 cartridge and eluted in MeCN. The MeCN was diluted down to 30% with 0.1% TFA and the sample was HPLC purified on the column used for the previous step. The C18 cartidge purification step was for the convenience of reducing sample volume to avoid multiple HPLC injections. This would not be necessary if this synthesis is done on a small scale. MALDI-TOF MS analysis showed the desired product (Figure 3-15a). Reflector mode MALDI-TOF MS showed a mass difference of -32 Da between the CTT-2-nitrophenol and the cyclized CTT-2-aminophenol

thus verifying that both the reduction and cyclization went to completion (Figure 3-15c,d).

**Oxidative Coupling of *p*-Fluoroaniline to CTT-2-aminophenol**

CTT-2-aminophenol (100 μM) and *p*-fluoroaniline (250 μM) were mixed and sodium periodate was added to a final concentration of 1 mM. After 2 min the reaction solution was diluted with 0.1% TFA until the pH was below 2. This solution was HPLC purified on the column described above and fluorescent fractions were manually collected. MALDI-TOF MS showed the desired coupling product (Figure 3-15b).

Labeling of CTT-2-aminophenol with [$^{18}$F]-FA followed the same general procedure but was instead purified on a C18 Sep-pak (Waters) to more easily concentrate the sample. Ethanol was used to elute the peptide in this case as this is the preferred solvent used for *in vivo* experiment compatibility.

**Crystal formation for X-ray analysis.**

The isolated major product described in Figure 3-6 was dissolved in a minimal amount of acetonitrile and then diluted to approximately 1 mg/mL with toluene. Crystals formed overnight upon vial-in-vial vapor diffusion using hexanes as the diluting solvent.

A colorless blade 0.12 x 0.08 x 0.04 mm in size was mounted on a Cryoloop with Paratone oil. Data were collected in a nitrogen gas stream at 100(2) K using phi and omega scans. Crystal-to-detector distance was 60 mm and exposure time was 5 seconds per frame using a scan width of 1.0°. Data collection was 99.2% complete to 67.00° in θ. A total of 11013 reflections were collected covering the indices, -15<=*h*<=15, -8<=*k*<=5, -16<=*l*<=17. 2090 reflections were found to be symmetry independent, with an $R_{int}$ of 0.0182. Indexing and unit cell refinement indicated a primitive, monoclinic lattice. The space group was found to be P2(1)/c (No. 14). The data were integrated using the Bruker SAINT software program and scaled using the SADABS software program. Solution by direct methods (SIR-2008) produced a complete heavy-atom phasing model consistent with the proposed structure. All non-hydrogen atoms were refined anisotropically by full-matrix least-squares (SHELXL-97). All hydrogen atoms were placed using a riding model. Their positions were constrained relative to their parent atom using the appropriate HFIX command in SHELXL-97.

**Crystallographic Characterization Data**

Table 1. Crystal data and structure refinement for francis01.

| | | |
|---|---|---|
| X-ray ID | francis01 | |
| Sample/notebook ID | CRB-1 | |
| Empirical formula | C13 H11 F N2 O2 | |
| Formula weight | 246.24 | |
| Temperature | 100(2) K | |
| Wavelength | 1.54178 Å | |
| Crystal system | Monoclinic | |
| Space group | P2(1)/c | |
| Unit cell dimensions | a = 12.7853(3) Å | α= 90°. |
| | b = 6.9961(2) Å | β= 115.406(2)°. |
| | c = 14.4223(4) Å | γ = 90°. |
| Volume | 1165.28(5) $Å^3$ | |
| Z | 4 | |
| Density (calculated) | 1.404 $Mg/m^3$ | |
| Absorption coefficient | 0.902 $mm^{-1}$ | |
| F(000) | 512 | |
| Crystal size | 0.12 x 0.08 x 0.04 $mm^3$ | |
| Crystal color/habit | colorless blade | |
| Theta range for data collection | 3.83 to 67.78°. | |
| Index ranges | -15<=h<=15, -8<=k<=5, -16<=l<=17 | |
| Reflections collected | 11013 | |
| Independent reflections | 2090 [R(int) = 0.0182] | |
| Completeness to theta = 67.00° | 99.2 % | |
| Absorption correction | Semi-empirical from equivalents | |
| Max. and min. transmission | 0.9648 and 0.8995 | |
| Refinement method | Full-matrix least-squares on $F^2$ | |
| Data / restraints / parameters | 2090 / 0 / 164 | |
| Goodness-of-fit on $F^2$ | 1.195 | |
| Final R indices [I>2sigma(I)] | R1 = 0.0638, wR2 = 0.2035 | |
| R indices (all data) | R1 = 0.0681, wR2 = 0.2065 | |
| Largest diff. peak and hole | 0.670 and -0.291 $e.Å^{-3}$ | |

Table 2. Atomic coordinates ( x $10^4$) and equivalent isotropic displacement parameters ($Å^2$x $10^3$) for francis01. U(eq) is defined as one third of the trace of the orthogonalized $U^{ij}$ tensor.

| | x | y | z | U(eq) |
|---|---|---|---|---|
| C(1) | 898(3) | 8746(4) | -760(2) | 29(1) |
| C(2) | 55(3) | 7551(4) | -712(3) | 33(1) |
| C(3) | -1101(3) | 7828(5) | -1359(3) | 40(1) |
| C(4) | -1412(3) | 9326(5) | -2031(3) | 37(1) |
| C(5) | -609(3) | 10583(5) | -2075(2) | 39(1) |
| C(6) | 560(3) | 10280(4) | -1450(2) | 34(1) |
| C(7) | 2638(3) | 6790(4) | 169(2) | 25(1) |
| C(8) | 2278(3) | 4950(4) | -54(2) | 27(1) |
| C(9) | 3242(3) | 3719(4) | 484(2) | 27(1) |
| C(10) | 3911(3) | 6791(4) | 895(2) | 27(1) |
| C(11) | 4655(3) | 7761(4) | 449(3) | 36(1) |
| C(12) | 4145(3) | 7614(4) | 1950(2) | 33(1) |
| C(13) | 3521(3) | 6602(5) | 2430(3) | 41(1) |
| N(2) | 3021(3) | 5759(5) | 2795(3) | 55(1) |
| N(1) | 2087(2) | 8476(3) | -118(2) | 29(1) |
| O(1) | 4208(2) | 4783(3) | 1035(2) | 28(1) |
| O(2) | 3320(2) | 1983(3) | 530(2) | 34(1) |
| F(1) | -2550(2) | 9582(3) | -2670(2) | 52(1) |

Table 3. Bond lengths [Å] and angles [°] for francis01.

| | |
|---|---|
| C(1)-C(2) | 1.389(5) |
| C(1)-C(6) | 1.399(4) |
| C(1)-N(1) | 1.413(4) |
| C(2)-C(3) | 1.382(5) |
| C(2)-H(2) | 0.9500 |
| C(3)-C(4) | 1.366(5) |
| C(3)-H(3) | 0.9500 |
| C(4)-F(1) | 1.359(4) |
| C(4)-C(5) | 1.374(5) |
| C(5)-C(6) | 1.391(5) |
| C(5)-H(5) | 0.9500 |
| C(6)-H(6) | 0.9500 |
| C(7)-N(1) | 1.345(4) |
| C(7)-C(8) | 1.358(4) |
| C(7)-C(10) | 1.511(4) |
| C(8)-C(9) | 1.428(4) |
| C(8)-H(8) | 0.9500 |
| C(9)-O(2) | 1.218(4) |
| C(9)-O(1) | 1.367(4) |
| C(10)-O(1) | 1.447(3) |
| C(10)-C(11) | 1.518(4) |
| C(10)-C(12) | 1.532(4) |
| C(11)-H(11A) | 0.9800 |
| C(11)-H(11B) | 0.9800 |
| C(11)-H(11C) | 0.9800 |
| C(12)-C(13) | 1.446(5) |
| C(12)-H(12A) | 0.9900 |
| C(12)-H(12B) | 0.9900 |
| C(13)-N(2) | 1.150(5) |
| N(1)-H(1) | 0.8800 |

| | |
|---|---|
| C(2)-C(1)-C(6) | 119.2(3) |
| C(2)-C(1)-N(1) | 121.8(3) |
| C(6)-C(1)-N(1) | 119.0(3) |
| C(3)-C(2)-C(1) | 120.7(3) |
| C(3)-C(2)-H(2) | 119.7 |
| C(1)-C(2)-H(2) | 119.7 |
| C(4)-C(3)-C(2) | 119.2(3) |
| C(4)-C(3)-H(3) | 120.4 |
| C(2)-C(3)-H(3) | 120.4 |
| F(1)-C(4)-C(3) | 118.9(3) |
| F(1)-C(4)-C(5) | 119.2(3) |
| C(3)-C(4)-C(5) | 121.9(3) |

| | |
|---|---|
| C(4)-C(5)-C(6) | 119.3(3) |
| C(4)-C(5)-H(5) | 120.4 |
| C(6)-C(5)-H(5) | 120.4 |
| C(5)-C(6)-C(1) | 119.7(3) |
| C(5)-C(6)-H(6) | 120.1 |
| C(1)-C(6)-H(6) | 120.1 |
| N(1)-C(7)-C(8) | 132.8(3) |
| N(1)-C(7)-C(10) | 118.6(2) |
| C(8)-C(7)-C(10) | 108.6(2) |
| C(7)-C(8)-C(9) | 108.5(3) |
| C(7)-C(8)-H(8) | 125.7 |
| C(9)-C(8)-H(8) | 125.7 |
| O(2)-C(9)-O(1) | 118.6(3) |
| O(2)-C(9)-C(8) | 131.4(3) |
| O(1)-C(9)-C(8) | 110.0(2) |
| O(1)-C(10)-C(7) | 103.7(2) |
| O(1)-C(10)-C(11) | 108.6(2) |
| C(7)-C(10)-C(11) | 113.0(3) |
| O(1)-C(10)-C(12) | 107.3(2) |
| C(7)-C(10)-C(12) | 112.5(3) |
| C(11)-C(10)-C(12) | 111.2(3) |
| C(10)-C(11)-H(11A) | 109.5 |
| C(10)-C(11)-H(11B) | 109.5 |
| H(11A)-C(11)-H(11B) | 109.5 |
| C(10)-C(11)-H(11C) | 109.5 |
| H(11A)-C(11)-H(11C) | 109.5 |
| H(11B)-C(11)-H(11C) | 109.5 |
| C(13)-C(12)-C(10) | 112.3(3) |
| C(13)-C(12)-H(12A) | 109.1 |
| C(10)-C(12)-H(12A) | 109.1 |
| C(13)-C(12)-H(12B) | 109.1 |
| C(10)-C(12)-H(12B) | 109.1 |
| H(12A)-C(12)-H(12B) | 107.9 |
| N(2)-C(13)-C(12) | 178.2(4) |
| C(7)-N(1)-C(1) | 126.4(3) |
| C(7)-N(1)-H(1) | 116.8 |
| C(1)-N(1)-H(1) | 116.8 |
| C(9)-O(1)-C(10) | 109.2(2) |

Symmetry transformations used to generate equivalent atoms:

Table 4. Anisotropic displacement parameters ($Å^2 x\ 10^3$)for francis01. The anisotropic displacement factor exponent takes the form: $-2\pi^2[\ h^2 a^{*2} U^{11} + ... + 2\ h\ k\ a^*\ b^*\ U^{12}\ ]$

| | $U^{11}$ | $U^{22}$ | $U^{33}$ | $U^{23}$ | $U^{13}$ | $U^{12}$ |
|---|---|---|---|---|---|---|
| C(1) | 47(2) | 17(1) | 26(1) | 0(1) | 21(1) | 5(1) |
| C(2) | 44(2) | 23(2) | 37(2) | 7(1) | 23(2) | 6(1) |
| C(3) | 47(2) | 34(2) | 46(2) | 1(2) | 27(2) | 3(2) |
| C(4) | 43(2) | 37(2) | 32(2) | -2(1) | 17(1) | 10(2) |
| C(5) | 62(2) | 27(2) | 26(2) | 6(1) | 16(2) | 11(2) |
| C(6) | 55(2) | 21(2) | 30(2) | 2(1) | 22(2) | 1(1) |
| C(7) | 40(2) | 16(1) | 24(1) | 1(1) | 18(1) | -1(1) |
| C(8) | 37(2) | 16(1) | 27(2) | -1(1) | 14(1) | -2(1) |
| C(9) | 39(2) | 17(1) | 29(2) | -1(1) | 19(1) | -2(1) |
| C(10) | 41(2) | 11(1) | 30(2) | -1(1) | 14(1) | -1(1) |
| C(11) | 45(2) | 20(2) | 47(2) | 3(1) | 24(2) | -2(1) |
| C(12) | 45(2) | 19(2) | 32(2) | -4(1) | 14(1) | -3(1) |
| C(13) | 57(2) | 31(2) | 34(2) | -4(1) | 18(2) | 3(2) |
| N(2) | 76(2) | 52(2) | 43(2) | -2(2) | 31(2) | -3(2) |
| N(1) | 43(2) | 12(1) | 34(1) | 0(1) | 19(1) | -3(1) |
| O(1) | 38(1) | 13(1) | 32(1) | 0(1) | 14(1) | -1(1) |
| O(2) | 46(1) | 12(1) | 47(1) | -1(1) | 23(1) | -2(1) |
| F(1) | 50(1) | 53(1) | 46(1) | 1(1) | 13(1) | 12(1) |

Table 5. Hydrogen coordinates ( x $10^4$) and isotropic displacement parameters (Å$^2$x $10^3$) for francis01.

| | x | y | z | U(eq) |
|---|---|---|---|---|
| H(2) | 275 | 6533 | -230 | 39 |
| H(3) | -1673 | 6986 | -1336 | 48 |
| H(5) | -849 | 11646 | -2528 | 47 |
| H(6) | 1126 | 11111 | -1491 | 41 |
| H(8) | 1515 | 4549 | -493 | 32 |
| H(11A) | 5475 | 7593 | 918 | 53 |
| H(11B) | 4471 | 9128 | 360 | 53 |
| H(11C) | 4500 | 7190 | -217 | 53 |
| H(12A) | 4985 | 7544 | 2401 | 40 |
| H(12B) | 3916 | 8977 | 1873 | 40 |
| H(1) | 2509 | 9512 | 116 | 35 |

## Section 3.6: References


1. Tack, B.F., Dean, J., Eilat, D., Lorenz, P.E., and Schechter, A.N. *J. Biol. Chem.* **1980**, *255*, 8842–8847.

2. Yuile, C.L., Lucas, F.V., Jones, C.K., Chapin, S.J., and Whipple, G.H. *J. Exp. Med.* **1953**, *98*, 173–194.

3. Wyttenbach, A., and Tolkovsky, A.M. *Mol. Cell. Proteomics.* **2006**, *5*, 553–559.

4. Izumi, Y., Sugisaki, K., Tani, Y., and Ogata, K. *Biochim. Biophys. Acta.* **1973**, *304*, 887–890.

5. Hershey, A.D., Chase, M. *J. Gen. Physiol.* **1952**, *36*, 39-56

6. Liu, S. *Adv. Drug Deliv. Rev.* **2008**, *60*, 1347–1370.

7. Miller, P.W., Long, N.J., Vilar, R., and Gee, A.D. *Angew. Chem. Int. Edit.* **2008**, *47*, 8998-9033.

8. Lee, E., Kamlet, A.S., Powers, D.C., Neumann, C.N., Boursalian, G.B., Furuya, T., Choi, D.C., Hooker, J.M., and Ritter, T. *Science.* **2011**, *334*, 639–642.

9. Hooker J.M., O'Neil J.P., Romanini D.W., Taylor S.E., Francis M.B. *Mol Imaging Biol.* **2008**, *10*, 182-91.

10. Berndt, M., Pietzsch, J., and Wuest, F. *Nuc. Med. Biol.* **2007**, *34*, 5–15.

11. Grierson, J.R., Yagle, K.J., Eary, J.F., Tait, J.F., Gibson, D.F., Lewellen, B., Link, J.M., and Krohn, K.A. *Bioconj. Chem.* **2004**, *15*, 373–379.

12. Cai, L., Lu, S., Pike, V.W., Cai, L., Lu, S., and Pike, V.W. *Eur. J. Org. Chem.* **2008**, *2008*, 2853–2873.

13. Wu W., Hsiao S.C., Carrico Z.M., Francis M.B. *Angew. Chem., Int. Ed.* **2009**, *48*, 9493-9497

14. Koivunen, E., Arap, W., Valtanen, H., Rainisalo, A., Medina, O.P., Heikkila, P., Kantor, C., Gahmberg, C.G., Salo, T., Konttinen, Y.T., et al. *Nat. Biotechnol.* **1999**, *17*, 768–774.

15. Penate Medina, O., Haikola, M., Tahtinen, M., Simpura, I., Kaukinen, S., Valtanen, H., Zhu, Y., Kuosmanen, S., Cao, W., Reunanen, J., et al. *Journal of Drug Delivery,* **2011**, 1–9.

16. Haka, M.S., Kilbourn, M.R., Leonard Watkins, G., and Toorongian, S.A. *J. Labelled Cmpd. Rad.* **1989**, *27*, 823–833.

17. Wright, J.B. *J. Heterocyclic Chem.* **1972**, *9*, 681-682.

18. Nielsen, O. and O. Buchardt *Synthesis* **1991**, *1991*, 819-821

19. Tam, J.P., Wu, C.R., Liu, W., and Zhang, J.W. *J. Am. Chem. Soc.* **1991**, *113*, 6657-6662.

# Chapter 4 - Using Oxidative Coupling to Synthesize MS2 Based Tumor Imaging Agents

## Section 4.1: Introduction to Tumor Imaging Agents

In the future, as medicine continues to develop, the leading causes of death in the US will become cancer related and much of the financial strain on the medical system will be due to cancer detection and treatment.[1] Because successful treatment is contingent on proper diagnosis, quick and inexpensive detection methods will be a major need in the near future. Patient survival rates vary depending on the type and location of cancer, but early and accurate diagnoses are important in all cases. Conventional tools for diagnosis include X-rays, CT (computed tomography) scans, and biopsies, but new tools are being developed that will give doctors more information about the cancer – resulting in better treatment. Many research groups are working to develop new technologies for cancer imaging with the hopes of providing higher resolution pictures at earlier stages of tumor development. Improvement in the resolution comes from two factors: first, new contrast agents release a signal much stronger than the background released by the body and second, specific targeting of these contrast agents attempts to place them only in cancerous tissue.[2,3,4,5]

The conventional methods available for tumor imaging such as X-rays and CT scans rely on differential absorption of electromagnetic radiation by tissues. For example, a tumor in the lung can be detected with a chest X-ray because the tumor tissue absorbs the radiation much differently than the surrounding healthy lung tissue. These methods are effective for determining the presence and location of large tumors. However, for many types of cancer, detection at this stage makes it difficult to treat because the tumor may have already begun to metastasize. Increasing the contrast between tumor tissue and the surrounding healthy tissue will help to visualize tumors at smaller sizes – thus resulting in earlier diagnosis and surgery that may possibly preempt metastasis. In addition to the benefit of early detection, imaging early stage cancer also gives the opportunity to study metastasis in animal tumor models. Learning how a cancer begins to spread would eventually help to better treat cancer. There are several classes of molecules that are being utilized as contrast agents for tumor visualization, and when injected into the body they increase the signal to background ratio for any tissue in which they are found. The most successful to date are fluorescent and positron emission tomography (PET) agents.

Fluorescent imaging agents are relatively inexpensive and simple to make but suffer from one main drawback — attenuation of the fluorescence signal through tissue. At this point most agents emit a wavelength of light that has relatively low transmission through biological tissue, thus making it difficult to image any organs more than a few centimeters from the surface. This is not a problem when imaging mouse animal models, as the animals are so small that all of their organs are sufficiently close to the skin surface. Even with small animals there is a stronger signal from organs closer to the surface so fluorescent imaging is not reliable for quantifying the relative amounts of agent in various organs. The ease of making fluorescent imaging agents makes them a common first pass in the early stages of research. Promising fluorescent agents can be radiolabeled for more in depth studies and for use in larger animals and humans. There is also significant ongoing research in search of effective near-infrared fluorescent molecules that could emit at a wavelength that shows little attenuation through tissue.[6]

There are several methods for producing fluorescent imaging agents. The most common

method is to attach a small molecule dye. Many new dyes have been designed to exhibit better properties for imaging, such as less bleaching and quenching. Another way to make fluorescent agents is through the use of quantum dots. They are a more robust imaging agent as they avoid many of the photobleaching problems that small molecule dyes have. One example of fluorescent imaging agents in the literature involved conjugation of epidermal growth factor (EGF) to a near-infrared quantum dot with the goal of targeting the epidermal growth factor receptor (EGFR).[7] The authors installed maleimide functional groups on the quantum dot and attached the EGF through the reduced native cysteines. Their agent showed binding to a cell line overexpressing EGFR and binding was blocked upon addition of a competing antibody for EGFR. Their experiments imaging immunocompromised mice with human tumors implanted in their hind flanks showed accumulation of the EGF-quantum dot agent. Interestingly, the agent reached a maximum tumor-to-blood ratio 4 h after injection and slowly dissipated at longer timepoints. While not directly relevant to the experiments described in this chapter, this paper showed the experiments necessary to prepare an imaging agent. The agent must be synthesized, then its binding to the target cell line must be verified, and finally the *in vivo* properties are investigated under several conditions at varying timepoints. It is a difficult and time-consuming process that has been a challenge for our group.

In contrast to fluorescent imaging agents, PET imaging agents are more difficult to prepare

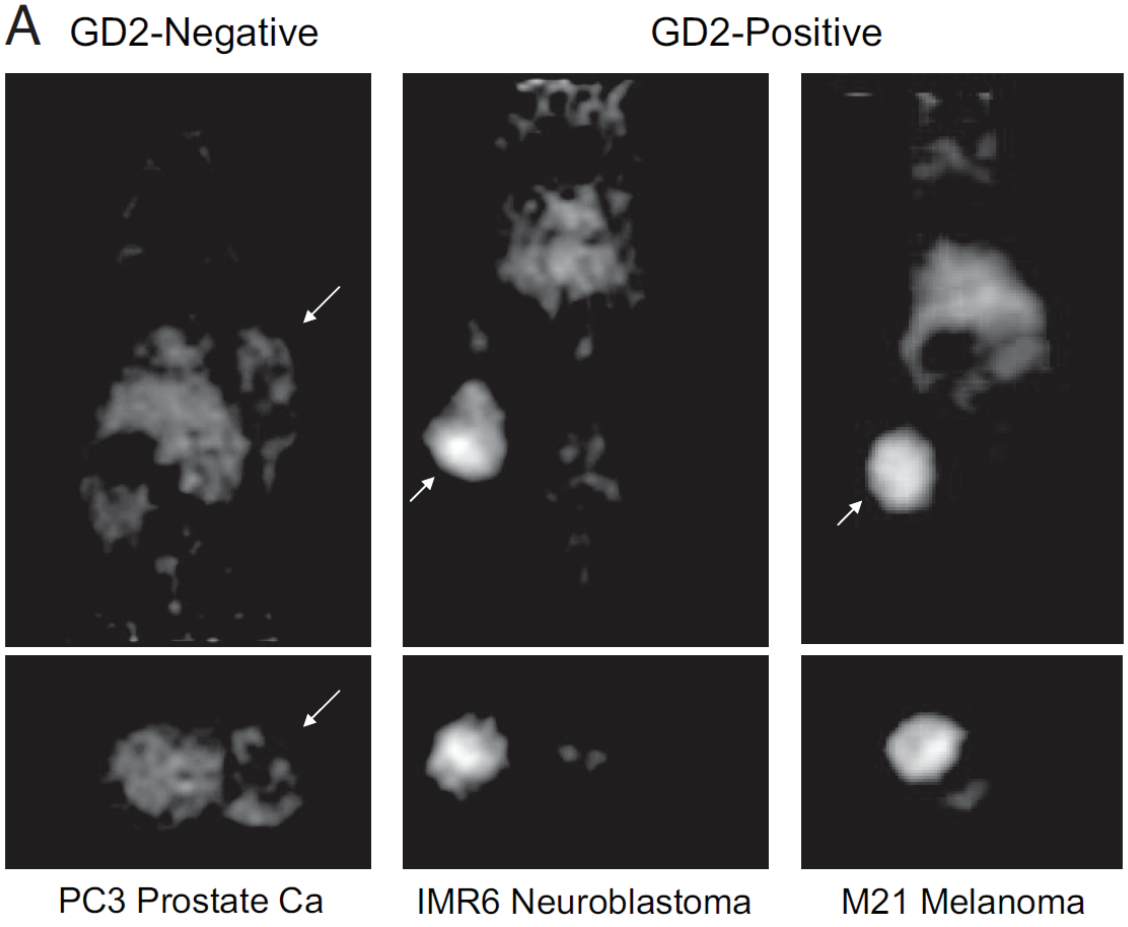



**Figure 4-1**. Anti-GD2 monoclonal antibody labeled with $^{64}$Cu was used to image three mouse xenograft models with PET. The GD2-negative prostate cancer cell line showed no accumulation of the imaging agent while the GD2-positive neuroblastoma and melanoma lines showed excellent tumor-to-background contrast ratios. Figure from Voss *et al. Proc Natl Acad Sci.* **2007**, *104*, 17489-17493.

but are a more powerful imaging technique. These agents are labeled with short-lived radioactive isotopes, as described in Chapter 3, so they must be synthesized immediately before use. Preparing these agents requires access to a medical cyclotron to produce the radionuclides, and expensive lead shielding and automation equipment to protect the chemist during synthesis. For PET imaging using agents derived from biological molecules (for example, antibodies and their many analogs) the most common radionuclides used are $^{18}$F and $^{64}$Cu for reasons mentioned in Chapter 3. Also discussed in Chapter 3 were several methods for labeling proteins with $^{18}$F, so those will not be revisited here. Instead, it would be more interesting to examine an example in literature of $^{64}$Cu labeling.

One group attached a strong copper chelator to an antibody and subsequently labeled the antibody with $^{64}$Cu. The antibody targeted disialogangliosides which are overexpressed on the surface of neuroblastoma and melanoma tumor cells. The agents showed excellent tumor-to-background contrast in the mouse xenografts of the positive cell lines and showed little accumulation in the negative control cell lines (Figure 4-1). This is an example of an ideal result for an *in vivo* imaging agent.[8]

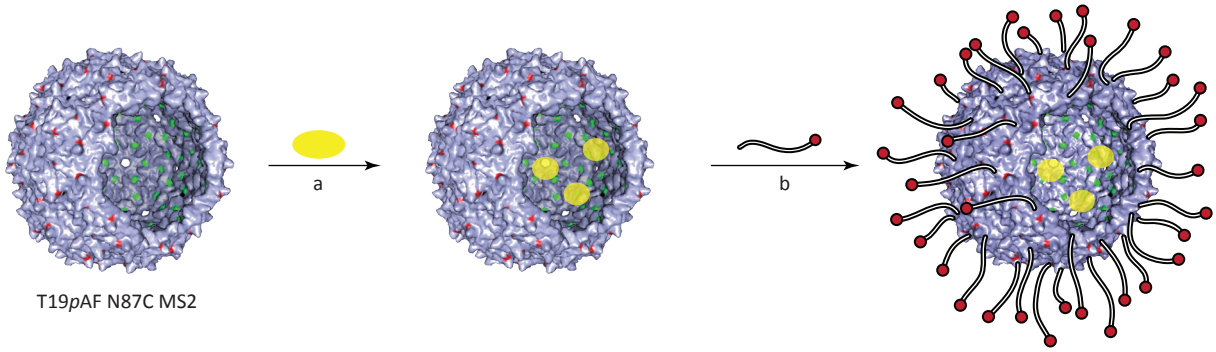
After presenting examples of the process of creating a tumor imaging agent and the ideal *in vivo* result, we will present our efforts towards the Francis group's goal of imaging breast cancer.

## Section 4.2: Building MS2-Based Fluorescent Imaging Agents

As described in the previous chapters, MS2 capsids have two modifiable surfaces. The inner surface of the hollow capsids can be accessed by diffusion through pores. Experiments in Chapter 2 determined that PEG2k could not diffuse through the pore but a small 8-mer peptide could diffuse through the pore and modify a residue on the inner surface. On the other hand, the outer surface was modified with PEG2k, PEG5k, and several small peptides. The final goal of the research into chemical modification of MS2 capsids is a scaffold approach where cargo can be loaded on the inside of the capsid, protected from the environment, and the outside can be modified with the probe of interest.

To build MS2-based fluorescent imaging agents, the inside of the capsid will be modified with a small molecule dye while the outside is modified with PEG or a peptide that binds an interesting target. To achieve this goal a new mutant of MS2 was necessary. The two most efficient methods of modifying MS2, oxidative coupling at *p*AF-19 and maleimide modification of cysteine-87, were combined into a single mutant, T19*p*AF N87C MS2.[9] These capsids were prepared in

**Scheme 4-1**. Assembling MS2-Based Fluorescent Imaging Agents[a]




[a]Conditions: (a) 20 equivalents of a maleimide fluorescent dye are added to T19*p*AF N87C MS2 at pH 6.5 for 1 hr before Nap-5 purification. (b) 5 equivalents of PEG or peptide aminophenols are added along with 25 equivalents of sodium periodate at pH 6.5 for a 2 min reaction before Nap-5 purification.

much the same way as the T19*p*AF mutant.

The scheme for synthesis of these constructs was maleimide modification of the cysteine followed by oxidative coupling at *p*AF-19 (Scheme 4-1). The dye selected for initial experiments was IRDye 680LT maleimide (LI-COR) as recommended by Ella Jones, an expert in small animal fluorescent imaging. The dye emits in the near-infrared (NIR) region and shows low levels of bleaching and quenching, thus making it particularly well suited for this application. To verify that this synthetic approach would be successful, the dye needed to survive exposure to periodate. Two identical samples of the dye at 10 µM were prepared and to one was added sodium periodate to a final concentration of 5 mM. The other sample received only phosphate buffer. The samples were monitored by absorbance at 680 nm and the sample exposed to periodate showed a significant decrease in absorbance over 25 min while the negative control showed no change (Figure 4-2). Since the aminophenol-aniline oxidative coupling only requires 2 min reaction times, the dye should be fine during that timeframe. However, this experiment showed the utility of this improved oxidative coupling reaction. The older reaction between anilines and phenylenediamine derivatives described by Hooker required 30 min reactions, which would not be compatible with this substrate.

After successful testing of IRDye 680 LT, the dye was then coupled to T19*p*AF N87C MS2 capsids. 20 equivalents of the dye were added at pH 6.5 and after 1 h incubation at room temperature the capsids were Nap-5 purified. For initial animal experiments three capsid-based

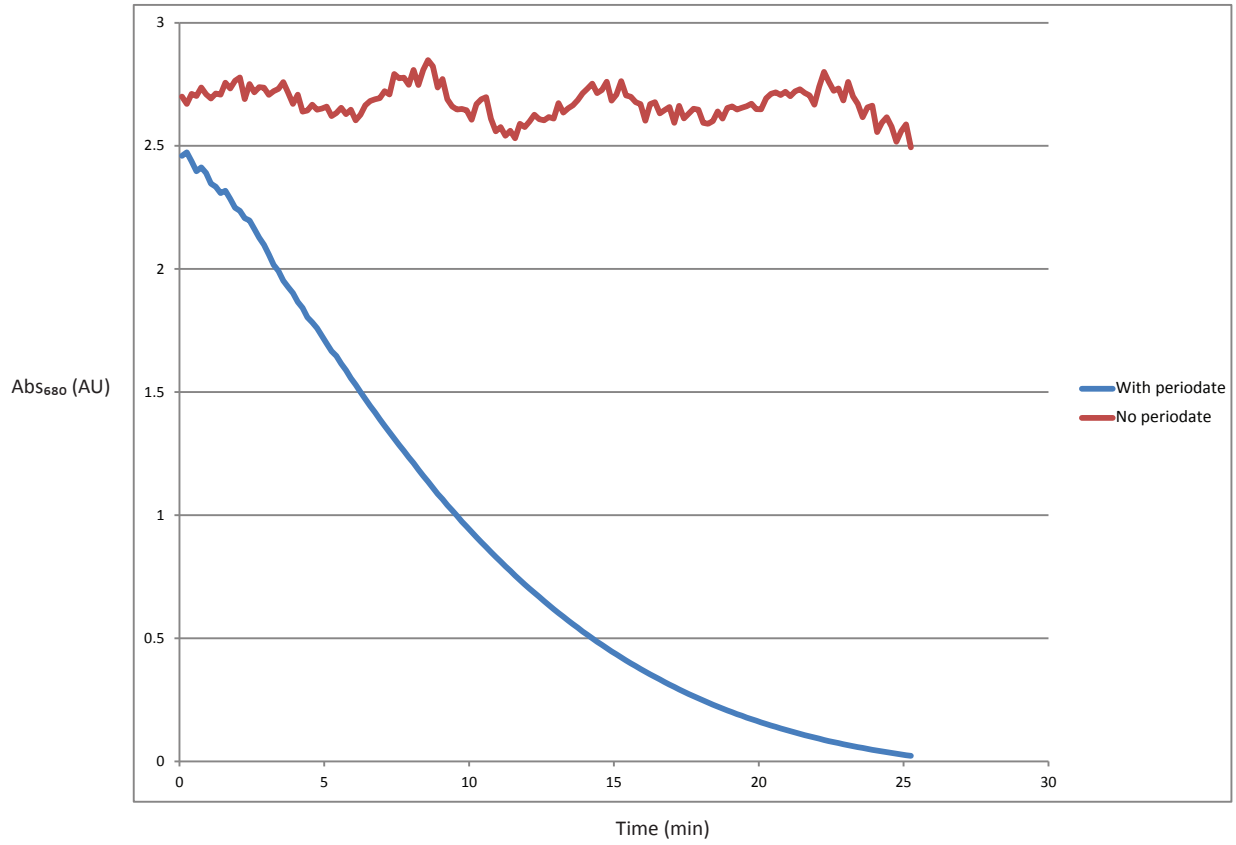



**Figure 4-2**. IRDye 680LT maleimide (LI-COR) was investigated for its stability towards sodium periodate. 100 µM dye was exposed to 5 mM sodium periodate and the absorbance at 680 nm was monitored. The dye exposed to periodate gradually showed a decrease in absorbance over 25 min while the negative control with no periodate showed no decrease in absorbance. Note that the noisy signal for the negative control is due to the dye concentration being too high and the instrument was near its maximum detection point. This experiment shows the utility of the improved oxidative coupling reaction - the old reaction required 30 min reactions that would be incompatible with this substrate.

agents were prepared with differing outer surfaces: unmodified, PEG5k modified, and cyclic RGD peptide (cRGD) modified. The PEG5k and cRGD agents were assembled by reacting the MS2-IRDye with aminophenol substrates in the presence of sodium periodate for 2 min before quenching by Nap-5 purification. The samples were analyzed using SDS-PAGE and the gel was imaged using the fluorescence of the IRDye (Figure 4-3b). The gel showed the expected shifts caused by the oxidative coupling of PEG5k or cRGD, thus suggesting that the samples were synthesized as planned in Scheme 4-1. Further verification of the assembly states of the capsids was performed using size exclusion chromatography on a GF-250 HPLC column (Agilent). The intact capsids

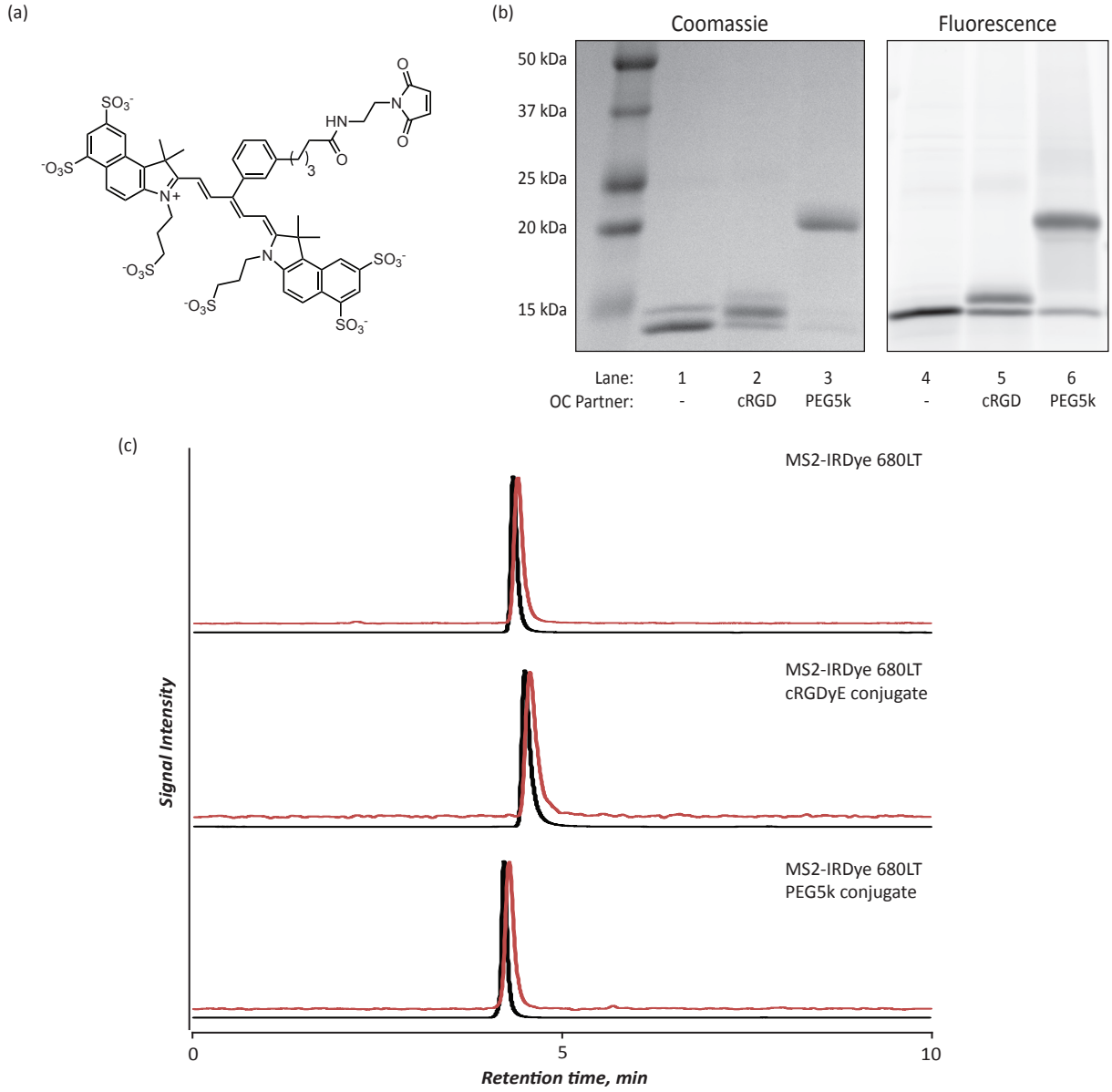



**Figure 4-3**. (a) Structure of IRDye 680LT maleimide (LI-COR). (b) SDS-PAGE analysis of MS2-IRDye samples conjugated to either PEG5k-aminophenol (Lanes 3 and 6) or cRGD-aminophenol (Lanes 2 and 5), while Lanes 1 and 4 have no outer surface modification. The coomassie-stained gel shows a small gel shift for the attachment of the dye (Lane 1 is easiest to see). For the fluorescence image the gel was visualized on the Typhoon Variable Mode Imager, with excitation at 680 nm and emission of 700 nm. (c) Size exclusion chromatography of MS2 capsids on GF-250 column (Agilent) to verify that disassembly did not occur. Intact capsids elute at 4-5 min, while disassembled monomers elute at 8-10 min. The black trace is the absorbance at 280 nm and the red trace is the fluorescence (680 nm excitation, 700 nm emission) of IRDye 680LT (LI-COR), which has been attached to the inner surface of the capsids. The retention time difference between the black and red traces corresponds to the distance between the two detectors.

eluted between 4-5 min, which from past experience is when intact MS2 capsids elute (Figure 4-3c). MS2 monomers elute between 8-10 min. With the construction of the imaging agents verified, the agents were ready for injection into animal models.

Healthy mice were imaged with these agents by labmate Michelle Farkas and collaborators at UCSF. After sterilizing the prepared MS2 samples (0.22 μm sterile filter), they were injected into the anesthetized mice and images were acquired at various timepoints over 48 h. Two significant observations were gleaned from the images produced by these preliminary experiments. Most importantly, the PEG-MS2 sample showed a more diffuse distribution throughout the body while the MS2 and MS2-cRGD samples showed much faster clearance into the liver (Figure 4-4). It is impossible to draw any definite conclusions based on an experiment with this few animals, but it does appear that PEG-MS2 has a longer circulation halflife than the other two samples. This would concur with the literature precedent that PEG attachment to a protein can increase its circulation halflife.[10] Also of note is the substantial increase in fluorescence in the 8 h and later timepoints. Unfortunately this represents a major fundamental flaw in this experiment. Experiments by Wesley Wu determined that the fluorescence of the dyes on the inside of MS2 was quenched nearly 100-fold from their fluorescence as free dyes in solution. This explained the unexpected results at the later timepoints — as the capsids were broken apart by the animal's filtering organs the newly freed dyes became much more fluorescent than the dyes present in capsid form. This made it extremely difficult to analyze the data past the timepoints at which the capsids began to disintegrate.

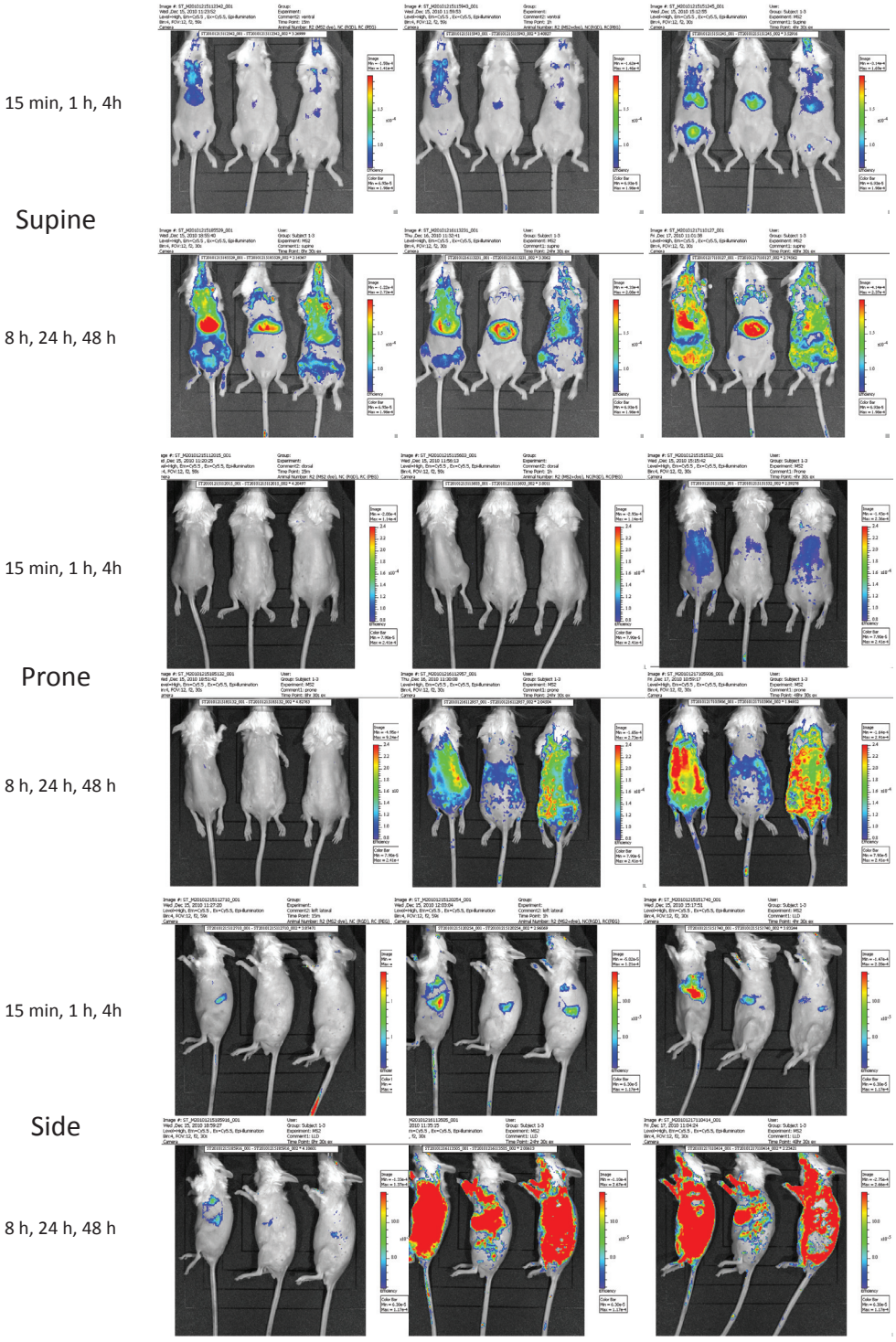



**Figure 4-4**. Fluorescent images of healthy mice injected with three MS2-based imaging agents: MS2, MS2-cRGD, and MS2-PEG5k. The animals are presented in that order from left to right in each frame. Animals were imaged at six timepoints: 15 min, 1h, 4h, 8h, 24h, and 48h and are designated by the caption on the left side. In an effort to reduce the effects of tissue absorbance, the animals were imaged in three positions at each timepoint: supine, prone, and side.

## Section 4.3: Building MS2-Based PET Imaging Agents

In contrast to the fluorescent imaging agents, the MS2-based PET imaging agents were significantly more complicated to synthesize. After attachment of the radionuclide, the agent must be immediately used as it is difficult to further modify or purify a radioactive substance and the short halflife makes it necessary to use within a reasonable timeframe. However, the first example of using MS2 as an imaging agent was a radiolabeled capsid with no external functionality attached. The first instance of this experiment was performed in healthy rats by Jacob Hooker using the $^{18}$F-fluorobenzaldehyde conjugation to alkoxyamine-functionalized MS2 capsids described in Chapter 3.[11] The results of this experiment showed that naked MS2 capsids were cleared from the heart by the 30 min timepoint and most of the activity remained in the liver before gradually moving through the digestive system towards the bladder over the next couple hours (Figure 4-5).

To improve upon the length of time in circulation, we set out to synthesize PEG-modified [$^{18}$F]-MS2. Unfortunately the same oxidative coupling reaction was used to attach [$^{18}$F]-FA and PEG. It would require difficult protecting group chemistry to do these reactions in a step-wise manner. The first attempt at solving this problem was to attach both simultaneously. Since PEGylation via the oxidative coupling only required 2 min reaction times, it was certainly feasible to do the reaction during or after radiolabeling. The MS2 construct necessary for this experiment was the one with an aminophenol on both the inner and outer surface. T19Y MS2 reacted with *p*-nitrobenzenediazonium tetrafluoroborate followed by reduction with sodium dithionite (Scheme 2-1) afforded MS2 with aminophenol functionality at position 85 (inside) and position 19 (outside). This MS2 constuct was modified with PEG5k-aniline in the presence of sodium periodate and analyzed via SDS-PAGE (Figure 2-10c, Lane 4). The same reaction was repeated but [$^{18}$F]-FA was added in along with the PEG5k-aniline before the sodium periodate was added. The reaction was quenched and purified on a Nap-5 column and 68% of radioactivity was incorporated

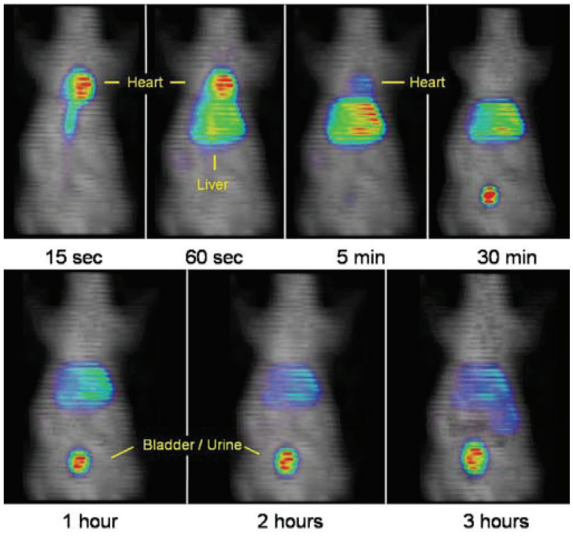



**Figure 4-5**. PET images of healthy rats injected with MS2-[$^{18}$F]-fluorobenzaldehyde. At 30 min the agent had been mostly cleared from the bloodstream into the liver. After 1 h the activity has begun to clear from the liver into the bladder. Figure from Hooker, J.M. *et al. Mol. Imaging. Biol.* **2008**, *10*, 182-191.

into the protein. The resulting capsids were analyzed with SDS-PAGE and the radioactive protein bands were imaged using a phosphorimager. A radioactive band correlating to the PEGylated MS2 monomer was clearly visible (Figure 4-6) while a control reaction where no PEG was added did not show this band. The presence of this band signified that a portion of the MS2 monomers were successfully modified with both a PEG5k-aniline and a [$^{18}$F]-FA.

This method of labeling MS2 was used to generate samples for experiments in healthy Sprague-Dawley rats. Two samples were prepared using the method described in the previous paragraph, except PEG2k-aniline was used instead of PEG5k-aniline. Also, the bare MS2 capsid was prepared from wildtype MS2 instead of T19YMS2, so there was only an aminophenol on the inside of the capsid at position 85. These two samples were injected into rats and imaged using PET. The results were nearly opposite of those found in the fluorescent imaging case described previously. As can be seen in Figure 4-7, immediately after injection much of the activity is found in the heart (upper red spot). At the 25-30 min timepoint, however, the rat on the left (MS2-PEG2k) shows nearly complete clearance to the liver. This looked more similar to the bare MS2 capsid than the MS2-PEG from the fluorescence experiment. The rat on the right (bare MS2) did not show complete clearance to the liver, even out to as long as 3 h, which looked most like the MS2-PEG example in the fluorescence images.

These confounding results were puzzling but one possible explanation was the poor state of the capsids following diazonium modification. Recall in Chapter 2 that the diazonium modification of MS2 caused possible deamination side reactions. Unfortunately the effect of this reaction on capsids (potential disassembly) was not realized until after these animal experiments were performed. It was apparent at that point that a new strategy for PEG and [$^{18}$F]-FA labeling of MS2. In-

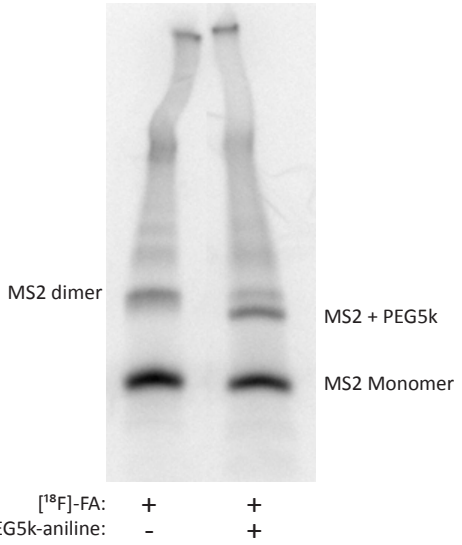



**Figure 4-6**. SDS-PAGE analysis of MS2 reaction with [$^{18}$F]-FA and PEG5k-aniline followed by imaging the gel's radioactivity using a phophorimager. MS2 with aminophenol functional groups at position 85 and 19 was mixed with both PEG5k-aniline and [$^{18}$F]-FA prior to addition of sodium periodate. In the reaction containing PEG5k-aniline a new protein band corresponding to MS2-PEG5k appeared.

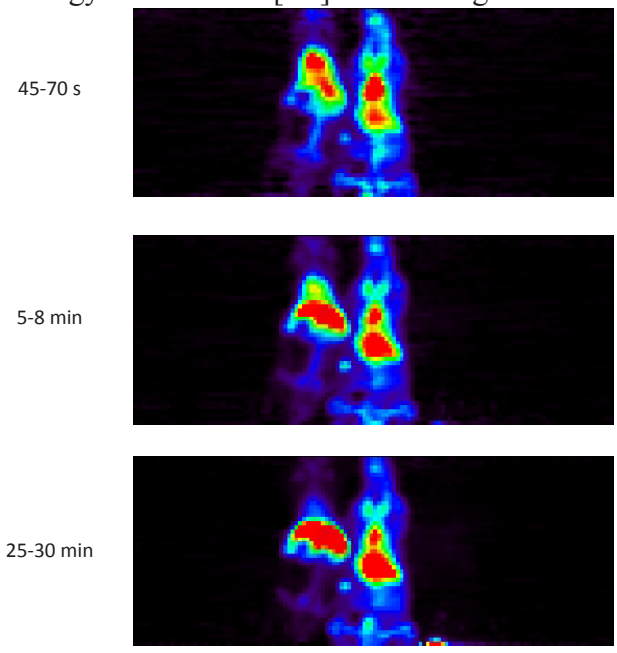



**Figure 4-7**. Coronal slices of PET imaging data at different timepoints. Two rats were imaged: The rat on the left was injected with [$^{18}$F]-MS2-PEG2k while the rat on the right was injected with [$^{18}$F]-MS2 with no PEG. The two animals are slightly offset but the two red regions correspond to the heart (upper) and liver (lower). Immediately following injection most of the activity is in the heart region and gradually moves into the liver.

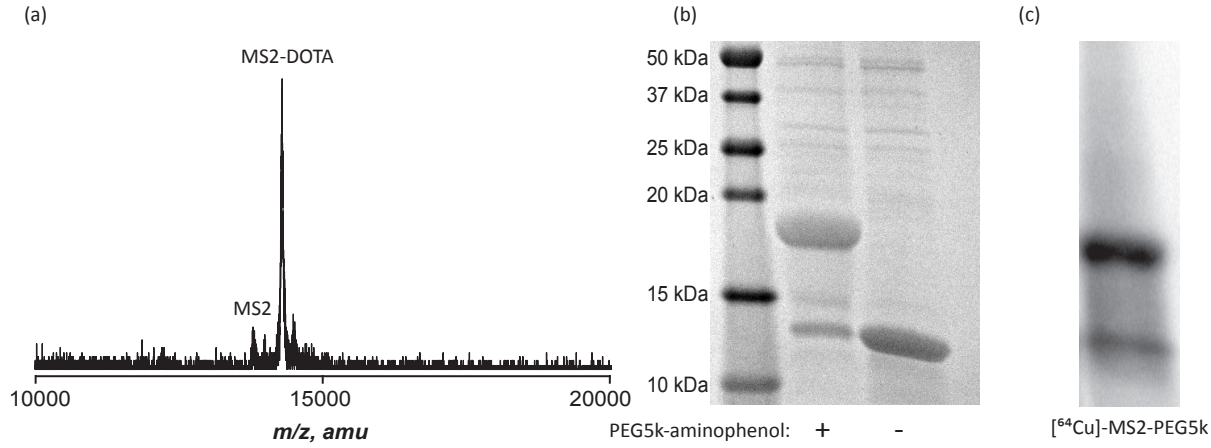



**Figure 4-8**. (a) MALDI-TOF MS of T19pAF N87C MS2 modified with DOTA-maleimide. Sample was cospotted with sinapinic acid. (b) Coomassie stained SDS-PAGE analysis of MS2-DOTA with and without PEG5k-aminophenol attached. (c) The DOTA-MS2-PEG5k was mixed with $^{64}$Cu and analyzed with SDS-PAGE. The resulting gel was imaged for radioactivity using a phosphorimager.

stead of simultaneously coupling the two, they could be coupled step-wise. As mentioned before, this would require a protecting group of some sort. Our original method for installing aminophenol functionality on MS2 for radiolabeling as described in Chapter 3 already contained an inherent protecting group. The nitrophenol intermediate could serve as a protecting group while oxidative coupling attaches PEG to the outer surface. Then the nitrophenol could be reduced with sodium dithionite and [$^{18}$F]-FA could be coupled. These experiments are still currently ongoing.

Separate experiments in tumored animals by Michelle Farkas suggested that we might not be looking at long enough timepoints to see the desired *in vivo* properties. Therefore we set forth to synthesize the same MS2-based imaging agents as before except using $^{64}$Cu to label the capsids. As mentioned in Chapter 2, the most common metal chelator used for $^{64}$Cu is DOTA. DOTA-maleimide was purchased and used to modify the cysteine of T19*p*AF N87C MS2 capsids using the standard maleimide conditions described previously. MALDI-TOF MS analysis of the MS2 monomers showed nearly complete single modification (Figure 4-8a). The DOTA-MS2 was then reacted with PEG5k-aminophenol and analyzed with SDS-PAGE to show the expected PEGylated monomer protein band (Figure 4-8b).

The procedure for labeling DOTA-protein complexes with $^{64}$Cu is simple and merely involves neutralizing the 0.1 M HCl solution of $^{64}$Cu and adding it to the protein.[12] After 1 h incubation at 37 °C, the protein was Nap-5 purified to remove free $^{64}$Cu. The resulting protein was analyzed with SDS-PAGE which showed two radioactive bands, one for unmodified monomer and the other for the PEG5k-modified MS2 monomer (Figure 4-8c).

This PEG5k-MS2-[$^{64}$Cu]-DOTA reagent was compared *in vivo* to a reagent produced in a similar fashion by Wesley Wu. Instead of PEG5k on the outer surface of MS2, Wesley's reagent had attached a protein called a DARPin that was capable of binding the HER2 receptor.[13] The two reagents were injected into mouse models bearing MCF7clone18[14] and MDA-MB-231 tumors,[15] the animals were sacrificed after 48 h, and their organs were counted for radioactivity. Of particular interest was the level of PEG5k-MS2 remaining in the blood at 48 h compared to the other MS2 construct (Figure 4-9). The PEG5k-MS2 sample showed 5.0 %ID/g while the other MS2 construct showed 1.6 %ID/g. It appeared that the PEG protected the capsid from being as readily removed from the bloodstream, thus increasing its circulation time. There were also much lower levels of capsid in the liver, the usual end point for MS2 capsids. Unlike the [$^{18}$F]-MS2-PEG experiment

described earlier, these capsids have been verified to be stable and in the proper assembly state. These promising results have led to another round of experiments currently ongoing that aim to definitively determine whether PEGylation of the capsids leads to a significant change in circulation time and biodistribution.

The original goal of my project was to image breast cancer, but that turned out to be an extremely complex problem for both scientific and logistical reasons. I therefore focused on the development of several types of imaging agents with differing pros and cons for each. The group now possesses the necessary chemistry to build imaging agents for multiple modalities through several different assembly strategies. It is my hope that a future student can harness these developments to investigate breast cancer in more detail.

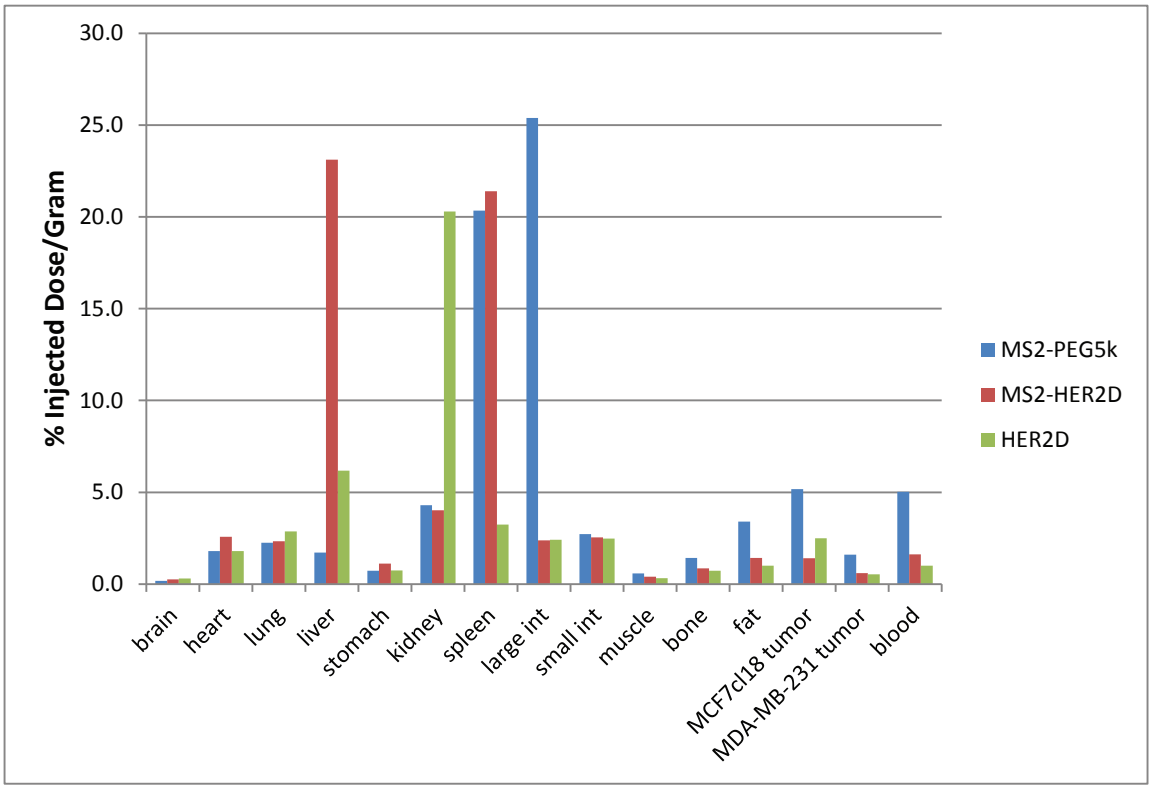



**Figure 4-9**. Biodistribution data for three tumored mice injected with MS2-PEG5k, MS2-HER2D (HER2 targeting DARPin), and HER2D (no MS2). The animals were sacrificed 48 h post-injection. Of most interest are the differences between the two MS2 samples. Compared to the MS2-HER2D sample, the MS2-PEG5k showed much less activity in the liver and substantially more activity in the tumor and blood. Only one animal was tested for each reagent, so these results must be repeated before any definite conclusions can be drawn.

## Section 4.4: Experimental Section

### Production of T19*p*AF N87C MS2 Capsids

T19*p*AF N87C MS2 was produced as described.[9]

### Oxidation of IRDye 680LT Maleimide with Sodium Periodate

To a 10 μM solution of the dye in 20 mM pH 6.5 phosphate buffer was added sodium periodate to a final concentration of 5 mM. Immediately upon addition of the sodium periodate the solution was transferred to a 1 cm pathlength quartz cuvette and the absorbance at 680 nm was monitored every 10 s for a total of 25 min. For a negative control, this procedure was repeated with an identical sample of the dye but no buffer was added instead of periodate.

### Labeling of T19*p*AF N87C MS2 with IRDye 680LT Maleimide

MS2 capsids with the unnatural amino acid *p*-aminophenylalanine at position 19 and cysteine at position 87 were modified with IRDye 680LT maleimide (LI-COR). To MS2 capsids (100 μM) in pH 6.5 phosphate buffer was added IRDye 680LT maleimide to a final concentration of 1 mM. After 1 h of reaction at room temperature the protein was purified via a Nap 5 size exclusion column, followed by a Nap 10 size exclusion column (GE Healthcare). The solution was then concentrated using a 100k molecular weight cutoff spin concentrator (Millipore). After concentration the solution was diluted with pH 6.5 phosphate buffer and concentrated again. This process was repeated until the flow through from the spin concentrator showed no signs of free dye. The capsids were then disassembled and analyzed via SDS-PAGE gel with both fluorescence detection and Coomassie staining (Figure 4-3b).

### Size Exclusion Chromatography of Unmodified MS2 Capsids and MS2 Capsid Conjugates

The capsids were analyzed by size exclusion chromatography to verify their assembly state after conjugation to peptides or PEG chains. MS2 was labeled with IRDye 680LT (LI-COR) on the inner surface using maleimide chemistry and then further modified on the outer surface via oxidative coupling chemistry. The MS2-IRDye 680LT (no outer surface modification), MS2-IRDye 680LT/cRGDyE conjugate, and MS2-IRDye 680LT/PEG5k conjugate were injected onto a GF-250 size exclusion column (Agilent). The mobile phase (0.5 mL/min flowrate) was 100 mM phosphate buffer pH 6.5. Each sample eluted near the void volume of the column, which suggested all samples were still in the assembled capsid form (Figure 4-3c).

### Simultaneous Labeling of MS2 with PEG5k-Aniline and [$^{18}$F]-FA

To 40 μL 100 μM MS2-aminophenol-85/19 was added 50 μL [$^{18}$F]-FA and 4 μL 10 mM PEG5k-aniline. After thoroughly mixing the three reagents sodium periodate was added to a final concentration of 5 mM. Following 2 min incubation at room temperature the reaction was quenched and purified on a Nap-5 column.

### Labeling of T19*p*AF N87C MS2 with DOTA Maleimide

MS2 capsids with the unnatural amino acid *p*-aminophenylalanine at position 19 and cysteine at position 87 were modified with DOTA maleimide. To MS2 capsids (100 μM) in pH 6.5 phosphate buffer was added DOTA maleimide to a final concentration of 2 mM. After 1 h of reaction at room temperature the protein was purified via a Nap 5 size exclusion column.

**$^{64}$Cu Labeling of PEG5k-MS2-DOTA Capsids**

$^{64}$Cu arrived from the supplier in a 0.1M HCl solution. 100 mM ammonium citrate pH 6 was added until the copper solution was pH 5-7. To 500 µL of 50 µM PEG5k-MS2-DOTA was added 5 mCi of $^{64}$Cu solution. The protein solution was incubated at 37 °C for 1 h, then Nap-5 purified to afford 2.5 mCi of labeled capsid. The capsid labeling was verified by SDS-PAGE (Figure 4-8c).

## Section 4.5: References


1. "Cancer Facts & Figures 2012" by American Cancer Society, **2012.**

2. Bulte, J.W.M., Kraitchman, D.L. *NMR Biomed.* **2004**, *17*, 484–499.

3. Rhyner, M.N., Smith, A.M., Gao, X., Mao, H., Yang, L., and Nie, S. *Nanomedicine-UK* **2006**, *1*, 209–217.

4. Sevick-Muraca, E.M., Houston, J.P., and Gurfinkel, M. *Curr. Opin. Chem. Biol.* **2002**, *6*, 642–650.

5. Torchilin, V.P. *Adv. Drug Deliv. Rev.* **2002**, *54*, 235–252.

6. Bremer, C., Ntziachristos, V., and Weissleder, R. *Eur. Radiol.* **2003**, *13*, 231–243.

7. Diagaradjane, P., Orenstein-Cardona, J.M., E. Colón-Casasnovas, N., Deorukhkar, A., Shentu, S., Kuno, N., Schwartz, D.L., Gelovani, J.G., and Krishnan, S. *Clin. Cancer Res.* **2008**, *14*, 731–741.

8. Voss, S.D., Smith, S.V., DiBartolo, N., McIntosh, L.J., Cyr, E.M., Bonab, A.A., Dearling, J.L.J., Carter, E.A., Fischman, A.J., Treves, S.T., et al. *Proc. Natl. Acad. Sci.* **2007**, *104*, 17489-17493.

9. Tong G.J., Hsiao S.C., Carrico Z.M., Francis M.B. *J. Am. Chem. Soc.* **2009**, *131*, 11174-11178.

10. Fishburn, C.S. *J. Pharm. Sci.* **2008**, *97*, 4167–4183.

11. Hooker J.M., O'Neil J.P., Romanini D.W., Taylor S.E., Francis M.B. *Mol Imaging Biol.* **2008**, *10*, 182-91.

12. Wu, A.M., Yazaki, P.J., Tsai, S.-W., Nguyen, K., Anderson, A.-L., McCarthy, D.W., Welch, M.J., Shively, J.E., Williams, L.E., Raubitschek, A.A., et al. *Proc. Natl. Acad. Sci.* **2000**, *97*, 8495–8500.

13. Zahnd, C., Pecorari, F., Straumann, N., Wyler, E., and Plückthun, A. *J. Biol. Chem.* **2006**, *281*, 35167–35175.

14. Surmacz, E., and Burgaud, J.L. *Clin. Cancer Res.* **1995**, *1*, 1429–1436.

15. Cailleau, R., Olivé, M., and Cruciger, Q. *In Vitro Cell Dev-Pl.* **1978**, *14*, 911–915.