UC Berkeley
UC Berkeley Electronic Theses and Dissertations

Title
Using bioconjugation reactions for protein detection, intracellular delivery, and the study of protein conformation

Permalink
https://escholarship.org/uc/item/9jq5r01q

Author
Sangsuwan, Rapeepat

Publication Date
2018

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

Using bioconjugation reactions for protein detection, intracellular delivery, and the study of protein conformation

By

Rapeepat Sangsuwan

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 Evan W. Miller
Professor Andreas Martin

Fall 2018

Using bioconjugation reactions for protein detection, intracellular delivery, and the study of protein conformation

Copyright ©2018

By: Rapeepat Sangsuwan

Abstract

Using bioconjugation reactions for protein detection, intracellular delivery, and the study of protein conformation

by

Rapeepat Sangsuwan

Doctor of Philosophy in Chemistry

University of California, Berkeley

Professor Matthew B. Francis, Chair

The synthetic modification of proteins has significantly contributed in the fields of chemical biology and material science. Choosing an appropriate chemical strategy that involves the controlled coupling of biomolecules in aqueous solution and in a well-defined manner is key. This work describes the use of suitable bioconjugation reactions for a variety of applications. Firstly, oxidative coupling of *o*-aminophenols and anilines was used for the detection of nitrotyrosine-containing proteins. In this work, we demonstrated that the chemoselectivity of the tandem nitro reduction/oxidative coupling sequence was sufficiently high to label nitrated proteins among thousands of others in complex mixtures. Second, we constructed protein-polymer conjugates using 2-pyridinecarboxaldehyde reaction for intracellular protein delivery. The polymers were endowed with 2-pyridinecarboxaldehyde (2PCA) groups, which modify proteins selectively at their N-terminal positions through imidazolidinone formation. The polymer-modified proteins were delivered into the cytosolic compartment of various cancer cells and the function of the proteins was retained during the process. Importantly, mechanistic studies have revealed that the internalization pathways of the constructs are likely to occur through a membrane fusion mechanism, thus at least partially circumventing the protein inactivation that occurs through endocytosis. Finally, we adapted 2PCA and oxidative coupling reactions to create a Förster resonance energy transfer construct of an intrinsically disordered protein (IDP) to monitor protein folding. A library of IDP with different lengths and densities of polyethylene glycol (PEG) through lysine modifications was constructed. We demonstrated that the size and density of the PEG chains had little to no effect on the conformation of IDPs.

Dedicated to my grandparents Khean and Chalouy Patan, and Thue and Pakhui Sae-Lim

Table of Contents

Contents **Page**

Chapter 1: Site-specific protein modification

1.1. Importance of protein-based materials 2
1.2. Bioconjugation reactions for targeting native amino acid residues 3
1.3. Bioconjugation reactions for targeting unnatural amino acids 5
1.4. Bioconjugation reactions for targeting a protein N-terminus 8
1.5. Conclusions and thesis overview 11
1.6. References 11

Chapter 2: Direct detection of nitrotyrosine-containing proteins using an aniline-based oxidative coupling strategy

2.1. Introduction 15
2.2. Results and discussion 15
2.2.1. Design of oxidative coupling mediated detection probes for Y-NO_2 15
2.2.2. Confirmation of Y-NO_2 labeling chemoselectivity 17
2.2.3. Detection of Y-NO_2 in cell lysates 20
2.3. Conclusions 21
2.4. Materials and methods 22
2.4.1. Reagents 22
2.4.2. Instrumentation and sample analyses 22
2.4.3. Synthetic procedures 23
2.4.4. Protein modification procedures 23
2.4.5. Additional experimental procedures 23
2.5. Supporting information 25
2.6. References 32

Chapter 3: Cytosolic delivery of proteins using amphiphilic polymers with 2-pyridinecarboxaldehyde groups for site-selective attachment

3.1. Introduction 35
3.2. Results and discussion 36
3.2.1. Polymer-protein conjugate design and polymer synthesis 36
3.2.2. 2PCA conjugation to GFP 38
3.2.3. Intracellular uptake of protein-polymer conjugates and uptake mechanism studies 40
3.2.4. Reactivity of 2PCA with RNase A 43
3.2.5. Enzymatic activity and cytotoxic effects of RNase A conjugates 43
3.3. Conclusions 43
3.4. Materials and methods 44
3.4.1. Reagents 44
3.4.2. Instrumentation and sample analyses 45
3.4.3. Synthetic procedures 45

3.4.4. Protein modification procedures 46
3.4.5. Additional experimental procedures 47
2.5. Supporting information 50
2.6. References 62

Chapter 4: Conformational change study of PEGylated intrinsically disordered proteins

4.1. Introduction 66
4.2. Results and discussion 67
4.2.1. Truncated IDP construct design, purification, and characterization 67
4.2.2. Synthesis of PEGylated IDP 68
4.2.3. Fluorescence anisotropy of PEGylated IDPs 70
4.2.3.1. The effect of salt concentrations and PEG lengths on the anisotropy of PEGylated IDPs 70
4.2.3.2. The effect of charges and ionic strength of buffers on the anisotropy of PEGylated IDPs 72
4.2.3.3. The effect of temperature, pH, and glucose on the anisotropy of PEGylated IDPs 73
4.2.4. Fluorescence correlation spectroscopy study of PEGylated IDPs 75
4.2.5. Synthesis of fluorescently labelled IDP for Förster Resonance Energy Transfer (FRET) study 75
4.2.6. FRET study 78
4.3. Conclusions 78
4.4. Materials and Methods 78
4.4.1. Reagents 78
4.4.2. Instrumentation and sample analyses 79
4.4.3. Synthetic procedures 79
4.4.4. Protein modification procedures 79
4.4.5. Additional experimental procedures 80
4.5. Supporting information 81
4.6. References 82

Acknowledgements

I am eternally grateful for Professor Matt Francis for being an extremely supportive advisor. I had a lot of fun and positive experiences working in his lab. Working with Matt provided an excellent introduction to chemical biology-based research and protein chemistry. Not only is he incredibly intelligent, he is also one of the nicest people I know. I highly value Matt's positivity and enthusiasm and he has directly contributed to my success in graduate school. I am also grateful to his guidance outside the sciences and his soup he made for me during the time I faced the medical problem, the most difficult time in my life.

I also thank the other members of my dissertation committee, Professors Evan Miller and Professor Andreas Martin for their time, feedback, questions during exams and assistance in improving doctoral projects.

I also have really appreciated how supportive all members of the Francis group have been in the last five years. They supported me at practice talks and quals, I don't think I would have learned or accomplished nearly as much without all their help. I am thankful for the discussion we had for proteins or experimental design. Thank you to all my labmates over the past five years. I especially would like to thank Allie Obermeyer who I have worked with during my rotation in the Francis lab and she has trained me all kinds of chemical biology skills. I still remember the time I run the first protein gel with her. Working with her has encouraged me become a better scientist and I rely on her for everything from answers to trivial chemistry questions, to discussions on current projects. I also thank Joel Finblom, Alan Marmelstein, Rachel Bisiewicz and Emmy Tian for the feedback on my scientific writing. I also had the pleasure of working with an outstanding undergraduate, Heather Ruiz and Emmy tian. They two are extremely talented and they impressed me so much by their commitment and hard work. Their contributions to my project and labwork were greatly appreciated. In addition to being great in lab, they were also affable and fun to be around and it was a joy to work with them and see them grow into a wonderful scientist.

I would like to thank my best friend, Phum Tachachartvanich, who is always beside me no matter how hard my life is. He is the one who sits and listens to me when I cry, and he is also the one who laughs with me. He is a talented cell biologist and he was a constant source of sound scientific and presentation advice. He has pushed me to become the best person and scientist I can be. I couldn't have done it without him.

Lastly, I would like to express my gratitude and love to my family–my niece, Emaline Sangsuwan, my sister in-law, Pohnsuree Sangsuwan, my brothers, Nontapan and Panupong Sangsuwan, and my parents, Prakong and Anuwat Sangsuwan. My family is the pillar that holds my life. No matter how far I am from home, I feel their support and their love. Not only has them put up with all of my craziness, but they have always believed in me and encouraged me to try and achieve my dreams. There would be no pages after this one without them.

Chapter 1

Site-specific protein modification

Abstract

The synthetic modification of proteins plays a crucial role in chemical biology, medicine, and materials science. Attaching synthetic molecules to proteins can be used in numerous applications, from the investigation of biological function to therapeutics. Bioconjugation reactions must take place in mild conditions and are required to achieve high levels of control to produce well-defined bioconjugates with optimal properties. This chapter reviews the targeting of natural amino acids and recent advancements in unnatural amino acid and enzyme-based protein modification strategies.

1.1. Importance of protein-based materials

The chemical modification of proteins has potentially contributed in the fields of chemical biology and material science (Figure 1.1.).[1-2] Introducing small molecules enables the study of protein conformation,[3] controls the protein properties, and allows the construction of protein-based materials.[4-5] Applications employing bioconjugations include the proteins with fluorescent dyes to track protein localization and conjugating drugs to a protein vehicle for use in drug delivery.[10-11] Additionally, the coupling of proteins with polymers has found use in drug delivery, cell attachment and improving the properties of proteins.[5, 12] As of promising proteins for bio-applications, antibodies have the binding ability to specific targets. Fluorescently labelled antibodies serve as a useful tool to investigate cellular biology. Incorporation of small molecule

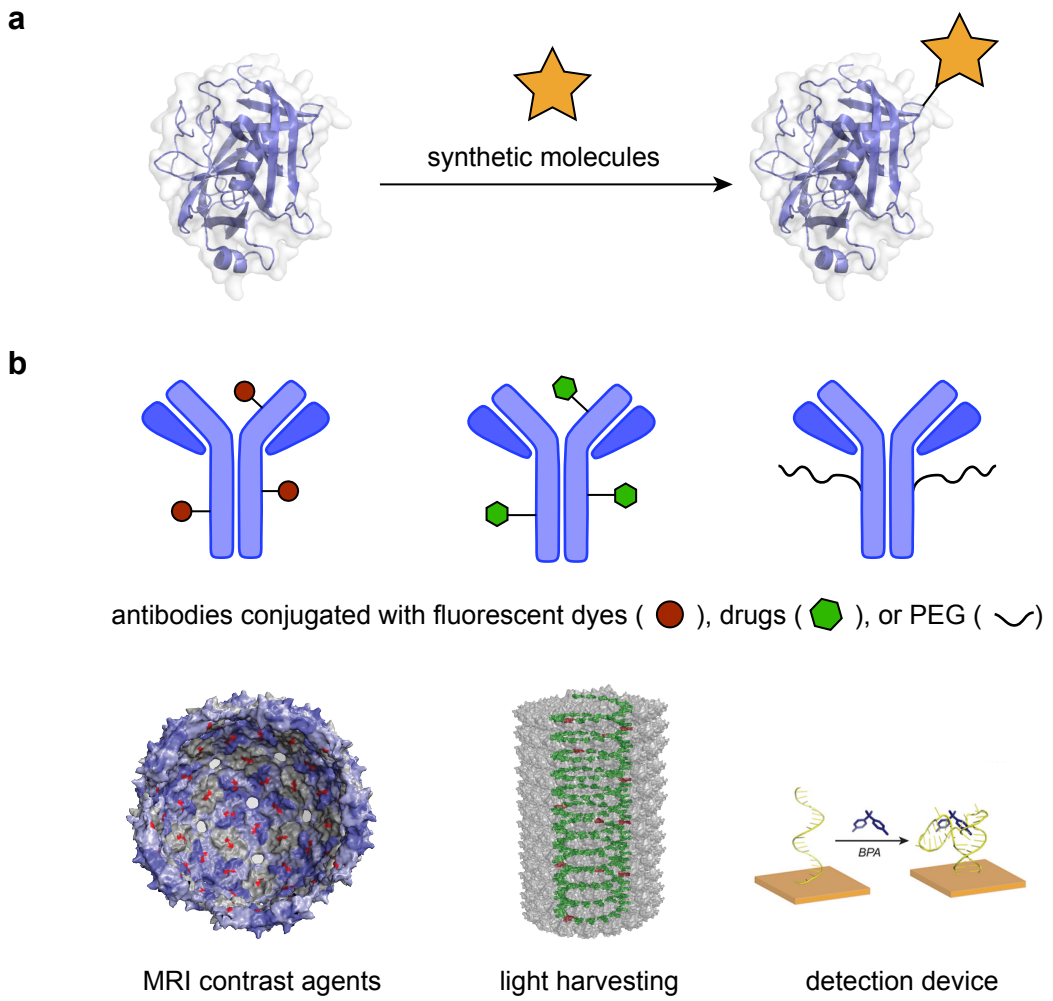


Figure 1.1. Attachment of synthetic components to proteins. (a) Protein modification allows for the synthesis of protein-based materials. (b) The resulting materials take advantage of the diverse biomolecular structure and function of proteins and can be used for various applications including antibodies modified with fluorophores, drug molecules, or polyethylene glycol (PEG) chains. Within the Francis group, proteins have been modified for a variety of applications including MRI contrast agents,[6-7] light harvesting,[8] and small molecule detection.[9]

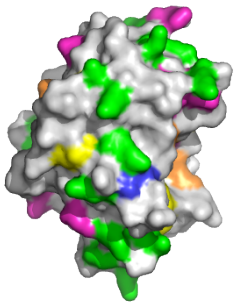
drugs into antibodies advances the use of these biomolecules for therapeutic delivery. For example, brentuximab vedotin became commercially available in USA in 2017 and was used to treat Hodgkin's lymphoma and systemic anaplastic large cell lymphoma (Figure 1.1b).[13] In addition, incorporation of polyethylene glycol into antibodies helps increase bioavailability in blood circulation, resulting in improved properties of the materials.[14]

Within the Francis Group, the synthetic modification of proteins has been applied to construct many materials for a variety of applications (Figure 1.1b). As an example, we constructed MRI contrast agents by the attachment of multiple gadolinium (III) complexes to the interior surfaces of an icosahedral virus (MS2).[7] We also have developed an artificial light-harvesting system by coupling donor and acceptor fluorophores to the rod-like tobacco mosaic virus.[8] Of additional focus and interest is the immobilization of oligonucleotides to gold surfaces, which allowed an electrochemical detection of the endocrine disruptor bisphenol A.[9]

The benefits of these materials, as demonstrated by these examples, are the incredible diversity of protein structures and functions which can be used to create new, highly functional materials. Specific protein bioconjugation reactions are required to incorporate synthetic molecules to protein materials. One additional frontier is the use of conditions that allow for the precise spatial control within a sample of those reactions, which guarantees a high degree of selectivity to maintain biological efficacy and to create uniform properties from batch to batch.

1.2. Bioconjugation reactions for targeting native amino acid residues

Several methods have emerged for the modification of native amino acids (Figure 1.2). Among the most common, techniques typically utilize the nucleophilicity of lysine and cysteine. Lysine can be modified with activated esters (e.g. NHS esters), sulfonyl chlorides, and iso(thio)cyanates to form amides, sulfonamides, or (thio)ureas (Figure 1.3).[2, 15-17] In addition, lysine can be coupled with aldehydes to form an imine and further generate a amine linkage through reductive amination. Provided the high abundance, these methods lack the site specificity necessary to allow for highly precise modifications. These conditions also are typically not only selective for lysine but also N-terminal amino group.



PDB ID: 1CHG

residue	average abundance
Lys	5.7 %
Glu/Asp	5.7 %
Cys	1.9 %
Tyr	3.2 %
Trp	1.4 %

Figure 1.2. Native amino acid side chains targeted for protein modifications

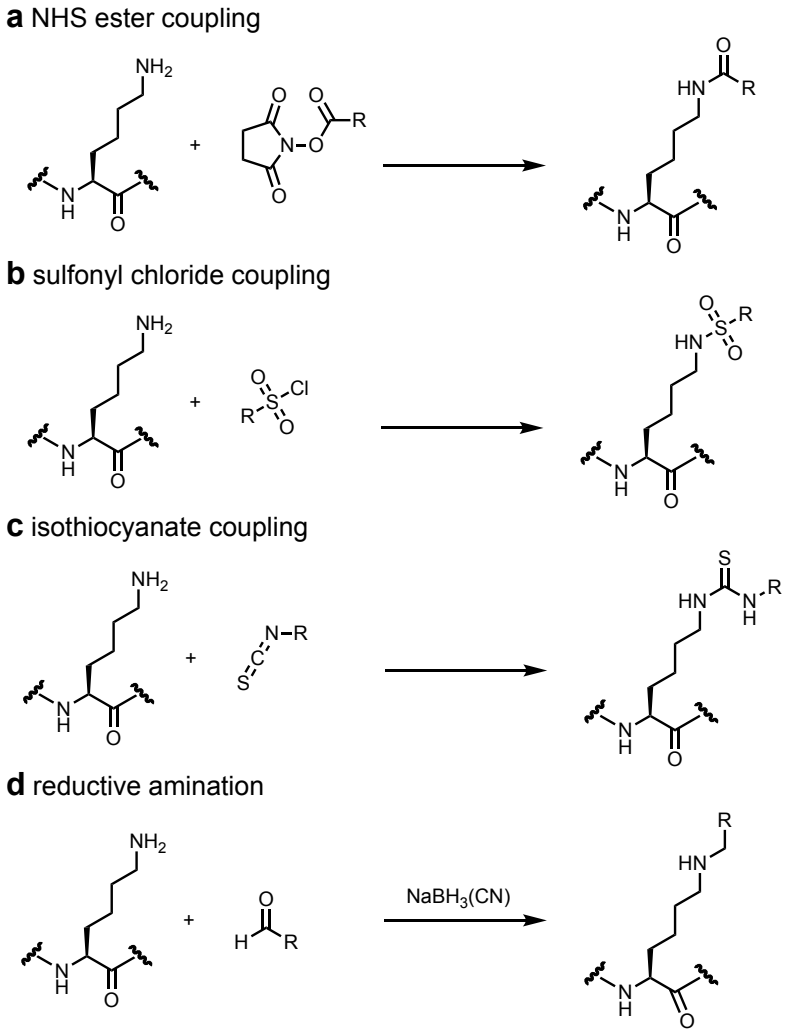


Figure 1.3. Bioconjugation methods for targeting lysine residues.

Other residues of interest are aspartic and glutamic acid, which can be modified using water-soluble activating agents, such as 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC).[17] This method can also result in the undesired cross-linking of proteins, as the amino groups of lysines can also be present on the surface of proteins (Figure 1.4).

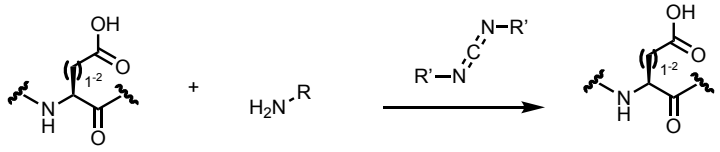


Figure 1.4. Bioconjugation methods for targeting aspartic acid and glutamic acid residues.

Cysteine, with the advantage of its low abundance on the protein surface, allows for better control over the level and location of modification. The most commonly used techniques for cysteine modification include disulfide exchange, alkylation of cysteines with iodoacetamide

reagents, and maleimide (Figure 1.5).[2, 17] Often, cysteines play a key role in enzymatic activity; therefore, modifying these residues may interfere with protein function.[18]

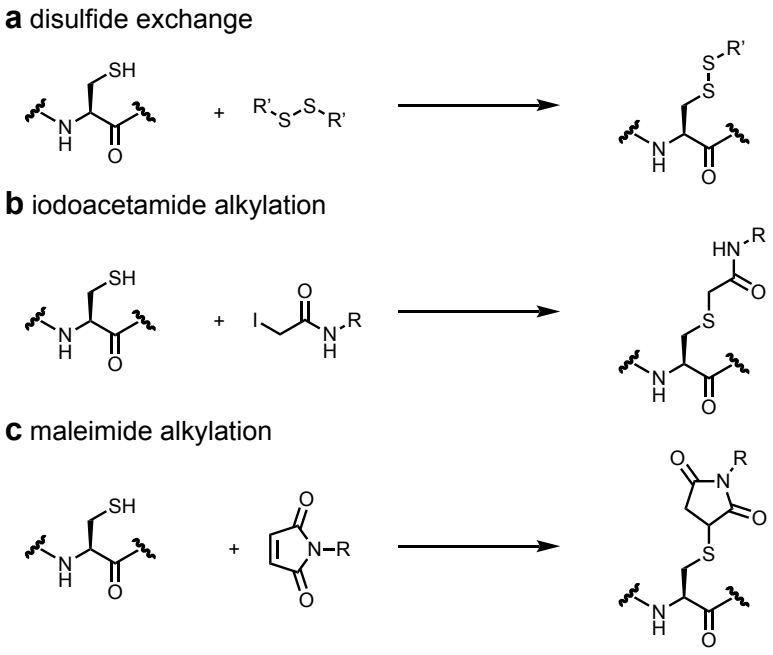


Figure 1.5. Bioconjugation methods for targeting cysteine residues.

The modification of aromatic-containing amino acid residues including tyrosine and tryptophan has been developed. These modifications ensure enhanced selectivity, as tyrosine and especially tryptophan display on the protein surface with low frequency. Of the common methods (Figure 1.6), tyrosine is modified through Mannich-type reactions,[19] water-soluble π-allylpalladium complexes,[20] cyclic diazodicarboxamide reagents,[21] or anisidine derivatives in the presence of cerium(IV) ammonium nitrate.[22] Tryptophan residues can be selectively reacted with rhodium carbenoids generated from diazo compounds and $[Rh_2(OAc)_4]$.[23] Cerium (IV) ammonium nitrate was also used to catalyze the coupling between tryptophan and phenylene diamine compounds (Figure 1.7).[22]

Even though these methods provide residue-specific modification and allow for relatively facile modification of native proteins, they often generate multiple modifications. Therefore, these methods often yield heterogeneous product mixtures. Recent efforts have focused on the development of strategies for site-selective modification of proteins, which can create predictable and well-defined properties of proteins.

1.3. Bioconjugation reactions for targeting unnatural amino acids

Site-selective modification of proteins can generally be divided into two different approaches. The first is the use of enzymatic means to react with a unique motif. The enzymatic-

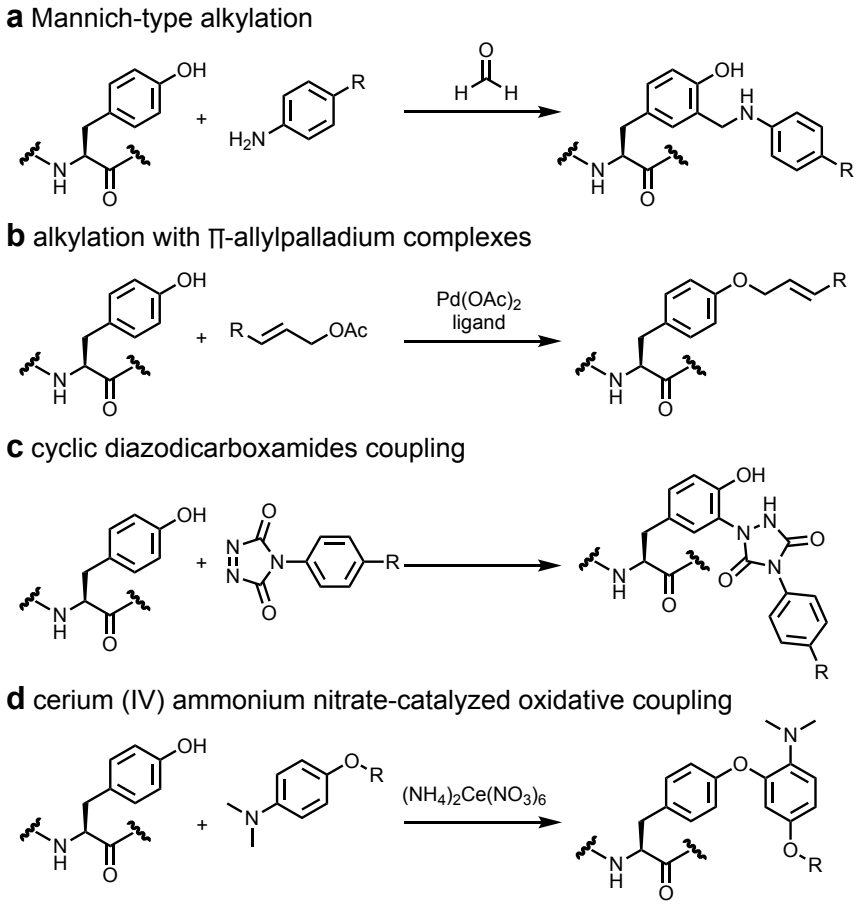


Figure 1.6. Bioconjugation methods for targeting tyrosine residues.

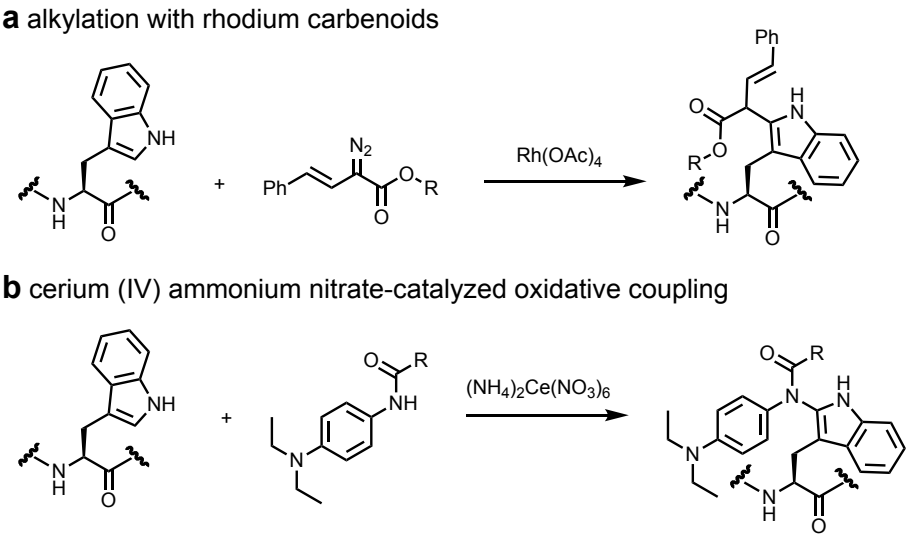


Figure 1.7. Bioconjugation methods for targeting tryptophan residues.

based reactions have been useful due to the exceptional specificity of enzymes. The motif/tag can be installed through genetic mutation at numerous locations within the protein, given that the proteins has tolerance for structural variation. As an example, bacterial Sortase A derived from *S. aureus* is used to catalyze the cleavage of the amide bond between threonine and glycine in the sequence LPXTG through a nucleophilic cysteine residue in the active site (Figure 1.8). This facilitates the formation of an amide linkage between thioester and oligoglycine derivatives.

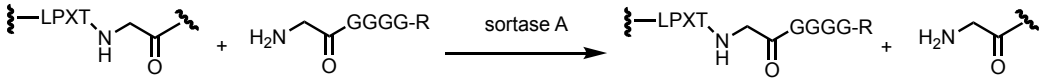


Figure 1.8. Sortase-catalyzed ligation for direct labeling of proteins

A second method for site-selective modification of proteins is the incorporation of unnatural amino acid residues. This is achieved by using the amber stop codon suppression method, which incorporates artificial amino acids in *E. coli* as well as in mammalian cells.[24-27] The unnatural amino acid can either have an effect of its own or may introduce a reactive handle to a protein that is orthogonal to the naturally occurring residues. The method requires the installation of mutated aminoacyl tRNA that has been covalently attached to an unnatural amino acid residue by an evolved aminoacyl tRNA synthetase. They are targeted at rarely-used stop codons that do not interfere with normal protein synthesis. The desired location of the incorporation can be achieved based on the position of the amber stop codon within the genetic sequence.

Methods for the site-selective incorporation of unnatural amino acids have also been reported. The introduction of azido and alkyne groups have been some of the most common reagents used due to their inertness and small size, allowing for wide adoption in many different substrates. Bertozzi *et al*. reported a copper-catalyzed [3+2] cycloaddition reaction as the so-called "Click" reaction of azido and alkyne groups (Figure 1.9a).[28] A variant of the azide-alkyne cycloaddition has been developed for a Cu-free click reaction to overcome the limitation of the toxicity of Cu(I) to biological systems (Figure 1.9b).[29-30] Azido groups have also been utilized to couple with functionalized phosphines to form the aza-ylide intermediate through the Bertozzi-Staudinger ligation (Figure 1.9c).[31]

The Francis group has developed specific bioconjugation reactions of anilines and electron-rich aromatic rings using an oxidant, known as "oxidative coupling" reactions (Figure 1.10). The first two of these reactions were carried out with sodium periodate for the coupling of *N*,*N*-dialkyl-*p*-phenylenediamine or *N,N*-dialkyl-*p*-anisidines and aniline.[32-33] The second variant was later developed for a faster reaction, using sodium periodate as an oxidant for the coupling of anilines with *o*-aminophenols to generate iminoquinone intermediates.[34] These reactions feature mild conditions and proceed with high coupling rates, even when low concentrations of reactants are used. However, the ability of periodate to oxidize other moieties on proteins, including cysteines and 1,2-diols found in glycans, may limit the scope of these coupling reactions. Additionally, the periodate-mediated coupling produces a mixture of two products, which complicates analysis and may prevent its use in applications that require well-defined linkages. Therefore, a milder oxidant, potassium ferricyanide, was used instead without oxidizing thiols or 1,2-diols.[35] Notably, a single, stable reaction product is obtained with this oxidant. This reaction was used to attach proteins to a nanoparticle or a targeting peptide for bioapplications.[36-37] The proposed mechanism of the reaction is depicted in Figure 1.11.[38]

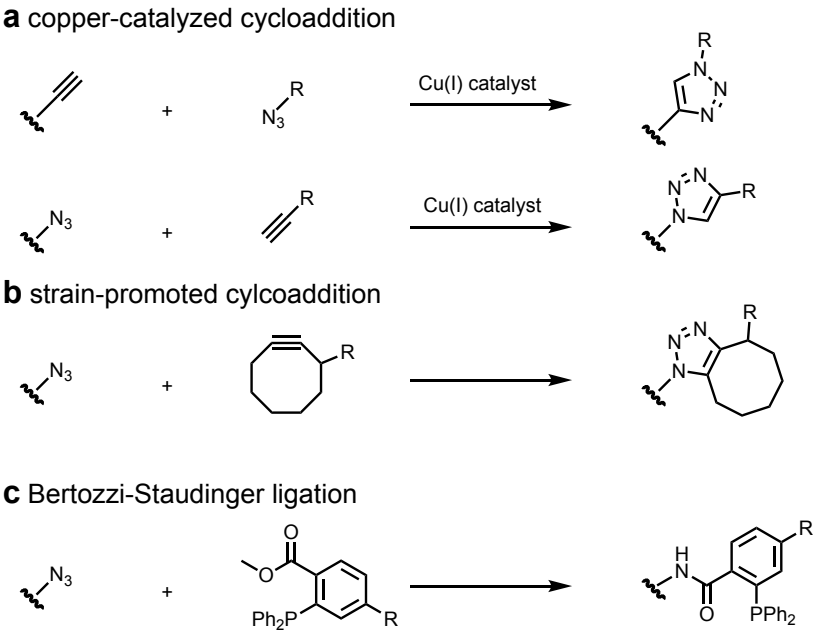


Figure 1.9. Bioconjugation methods for targeting unnatural amino acid residues. Introduction of azides to proteins allows for selective modification.

1.4. Bioconjugation reactions for targeting a protein N-terminus

The N-terminus is a unique site and has a different reactivity from the side-chain amino group of lysine. Many strategies have targeted the protein N-terminus in multiple ways (Figure 1.12). N-terminal tryptophan has been carried out via a Pictet-Spengler reaction.[39] The modification of N-terminal cysteines has been explored through reaction with thioester reagents in a process known as native chemical ligation.[40] N-terminal serine and threonine residues have been modified through periodate oxidation to yield reactive aldehydes.[41-42] Our group reported the method for selective N-terminal modification using pyridoxal-5'-phosphate and N-methylpyridinium-4-carboxaldehyde (Rapoport's salt) for the oxidation of the N-terminus to a reactive carbonyl moiety.[43-46] The aldehyde or ketone product can be further modified with aminooxy reagents to generate oxime ligation. Our group also reported an oxidative coupling reaction between aminophenols and N-terminal proline residues as a two-step reaction.[47] After further advancement, we developed a 2-pyridinecarboxaldehyde (2PCA) reaction as a simple, one-step method suitable for modifying a wide range of protein N-termini (Figure 1.13).[48] The reaction generates an imine intermediate that is subsequently attacked by an amide nitrogen to yield the imidazolidinone product. The 2PCA reaction is highly useful, as it does not require varying degrees of protein engineering that can reduce protein expression levels and disturb protein function.

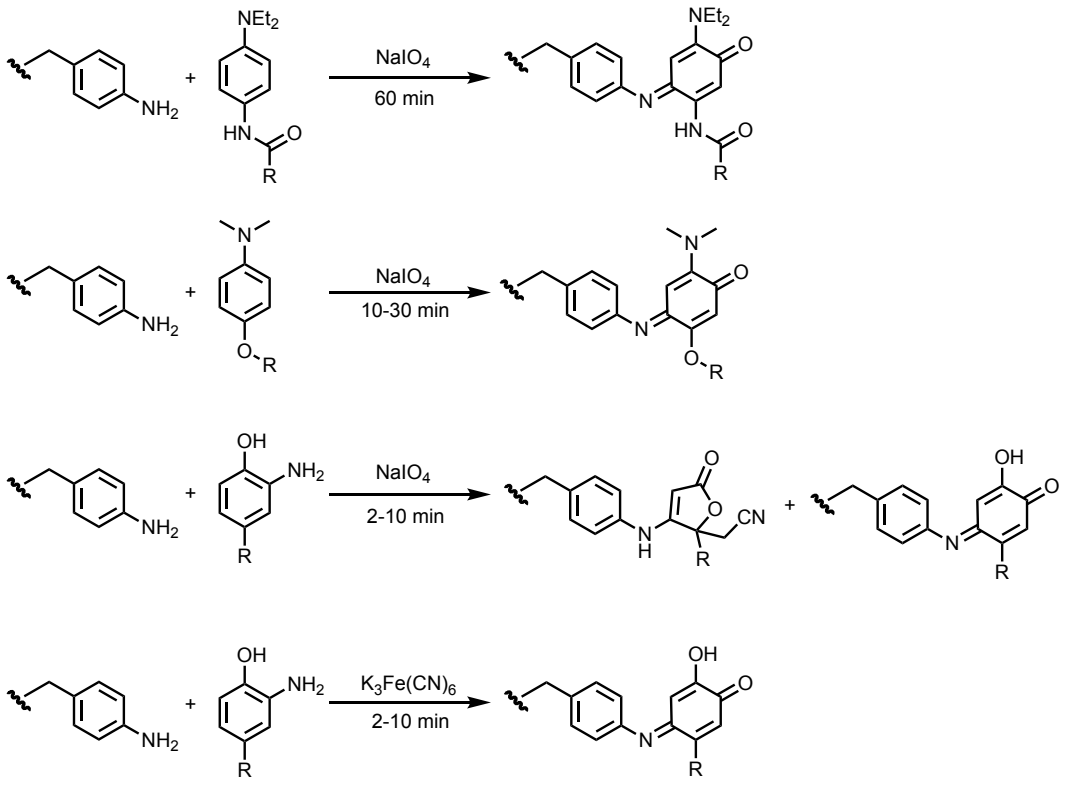


Figure 1.10. Strategies for the modification of aniline residues. Introduced *p*-aminophenylalanine residues can be modified with various electron-rich aromatic rings in the presence of oxidants.

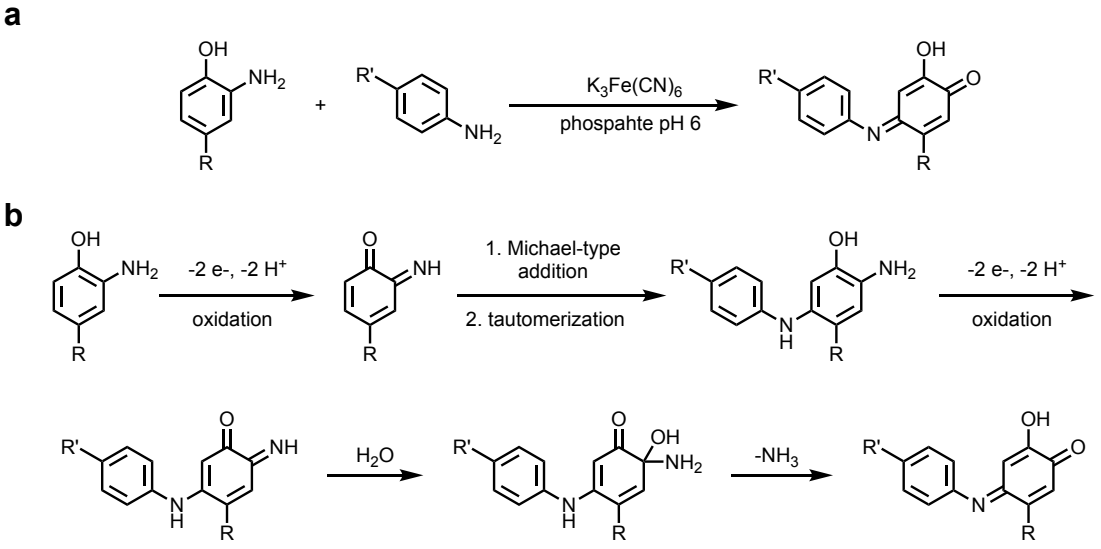


Figure 1.11. Proposed mechanism of ferricyanide-catalyzed oxidative coupling. (a) an oxidative coupling reaction between aminophenol and aniline in the presence of ferricyanide. (b) the proposed mechanism of the oxidative coupling.[38]

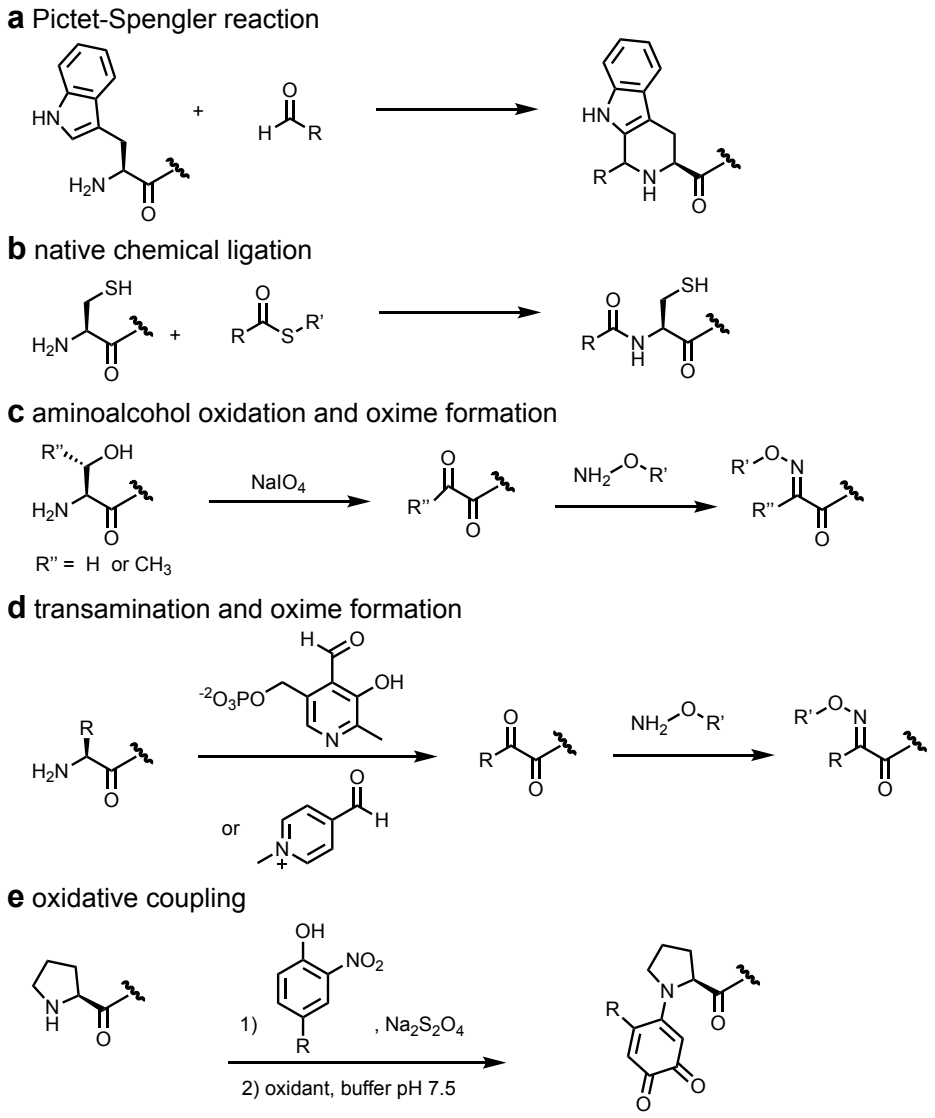


Figure 1.12. Bioconjugation methods for targeting N-terminal amino acid residues.

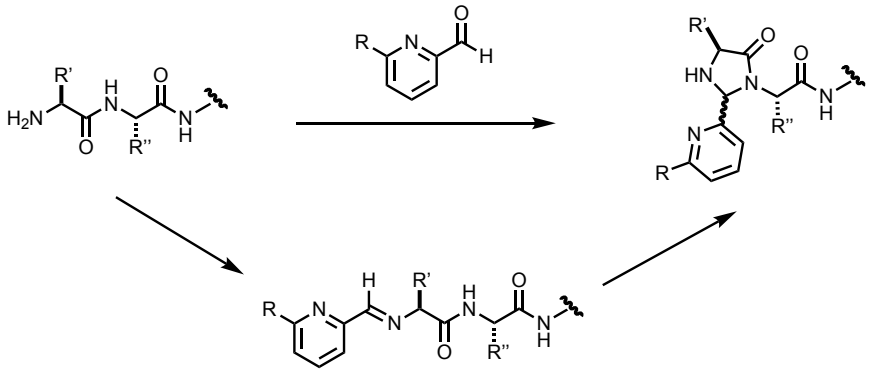


Figure 1.13. 2-pyridinecarboxaldehyde reaction for N-terminal specific modification.

1.5. Conclusions and thesis overview

As for the fields of chemical biology, bioengineering, biophysics, and materials science progress, the selection of an appropriate modification strategy is key for the construction of well-defined conjugates. Ideally, the reaction should proceed in aqueous media under mild pH and temperature conditions to maintain the structural integrity and function of the protein. In addition, the reaction should achieve the high degree of chemoselectivity and site specificity.

This work focuses on the application of site-selective bioconjugation reactions for various applications. The following chapters report the application of the oxidative coupling reaction for the detection of protein nitration, a post-translational modification. Further, the 2PCA reaction was used for attaching a polymer with proteins for cytosolic protein delivery. Finally, a protein construct of Förster resonance energy transfer was produced using the oxidative coupling and 2PCA reactions.

1.6. References

1. Spicer, C. D.; Davis, B. G., Selective chemical protein modification. *Nat. Commun.* **2014,** *5*, 4740.
2. Boutureira, O.; Bernardes, G. a. J., Advances in chemical protein modification. *Chem. Rev.* **2015,** *115* (5), 2174-2195.
3. Nanda, J. S.; Lorsch, J. R., Labeling of a protein with fluorophores using maleimide derivitization. *Methods Enzymol.* **2014,** *536*, 79-86.
4. Nischan, N.; Hackenberger, C. P. R., Site-specific PEGylation of Proteins: Recent Developments. *J. Org.* **2014,** *79* (22), 10727-10733.
5. Dumas, A.; Spicer, C. D.; Gao, Z.; Takehana, T.; Lin, Y. A.; Yasukohchi, T.; Davis, B. G., Self-liganded Suzuki–Miyaura coupling for site-selective protein PEGylation. *Angew. Chem.* **2013,** *125* (14), 4008-4013.
6. Jeong, K.; Netirojjanakul, C.; Munch, H. K.; Sun, J.; Finbloom, J. A.; Wemmer, D. E.; Pines, A.; Francis, M. B., Targeted molecular imaging of cancer cells using MS2-based 129Xe NMR. *Bioconjug. Chem.* **2016,** *27* (8), 1796-1801.
7. Garimella, P. D.; Datta, A.; Romanini, D. W.; Raymond, K. N.; Francis, M. B., Multivalent, high-relaxivity MRI contrast agents using rigid cysteine-reactive gadolinium complexes. *J. Am. Chem. Soc.* **2011,** *133* (37), 14704-14709.
8. Delor, M.; Dai, J.; Roberts, T. D.; Rogers, J. R.; Hamed, S. M.; Neaton, J. B.; Geissler, P. L.; Francis, M. B.; Ginsberg, N. S., Exploiting Chromophore–Protein Interactions through Linker Engineering To Tune Photoinduced Dynamics in a Biomimetic Light-Harvesting Platform. *J. Am. Chem. Soc.* **2018**.
9. Furst, A. L.; Smith, M. J.; Francis, M. B., Direct electrochemical bioconjugation on metal surfaces. *J. Am. Chem. Soc.* **2017,** *139* (36), 12610-12616.
10. Alley, S. C.; Okeley, N. M.; Senter, P. D., Antibody–drug conjugates: targeted drug delivery for cancer. *Curr. Opin. Chem. Biol.* **2010,** *14* (4), 529-537.
11. Lavis, L. D.; Raines, R. T., Bright ideas for chemical biology. *ACS Chem. Biol.* **2008,** *3* (3), 142-155.
12. Lee, J. P.; Kassianidou, E.; MacDonald, J. I.; Francis, M. B.; Kumar, S., N-terminal specific conjugation of extracellular matrix proteins to 2-pyridinecarboxaldehyde functionalized polyacrylamide hydrogels. *Biomaterials* **2016,** *102*, 268-276.
13. Seattle Genetics, I. Classical Hodgkin lymphoma (cHL). (accessed 10/18/12018).

14. Pelegri-O'Day, E. M.; Lin, E.-W.; Maynard, H. D., Therapeutic protein–polymer conjugates: advancing beyond PEGylation. *J. Am. Chem. Soc.* **2014,** *136* (41), 14323-14332.
15. Hermanson, G. T., *Bioconjugate techniques*. Academic press: 2013.
16. Sletten, E. M.; Bertozzi, C. R., Bioorthogonal chemistry: fishing for selectivity in a sea of functionality. *Angew. Chem. Int. Ed.* **2009,** *48* (38), 6974-6998.
17. Basle, E.; Joubert, N.; Pucheault, M., Protein chemical modification on endogenous amino acids. *Chem. Biol.* **2010,** *17* (3), 213-227.
18. Verma, S.; Dixit, R.; Pandey, K. C., Cysteine proteases: modes of activation and future prospects as pharmacological targets. *Front. Pharmacol.* **2016,** *7*, 107.
19. Joshi, N. S.; Whitaker, L. R.; Francis, M. B., A three-component Mannich-type reaction for selective tyrosine bioconjugation. *J. Am. Chem. Soc.* **2004,** *126* (49), 15942-15943.
20. Tilley, S. D.; Francis, M. B., Tyrosine-selective protein alkylation using π-allylpalladium complexes. *J. Am. Chem. Soc.* **2006,** *128* (4), 1080-1081.
21. Ban, H.; Gavrilyuk, J.; Barbas III, C. F., Tyrosine bioconjugation through aqueous ene-type reactions: a click-like reaction for tyrosine. *J. Am. Chem. Soc.* **2010,** *132* (5), 1523-1525.
22. Seim, K. L.; Obermeyer, A. C.; Francis, M. B., Oxidative modification of native protein residues using cerium (IV) ammonium nitrate. *J. Am. Chem. Soc.* **2011,** *133* (42), 16970-16976.
23. Antos, J. M.; Francis, M. B., Selective tryptophan modification with rhodium carbenoids in aqueous solution. *J. Am. Chem. Soc.* **2004,** *126* (33), 10256-10257.
24. Cornish, V. W.; Mendel, D.; Schultz, P. G., Probing protein structure and function with an expanded genetic code. *Angew. Chem. Int. Ed.* **1995,** *34* (6), 621-633.
25. Noren, C. J.; Anthony-Cahill, S. J.; Griffith, M. C.; Schultz, P. G., A general method for site-specific incorporation of unnatural amino acids into proteins. *Science* **1989,** *244* (4901), 182-188.
26. Mehl, R. A.; Anderson, J. C.; Santoro, S. W.; Wang, L.; Martin, A. B.; King, D. S.; Horn, D. M.; Schultz, P. G., Generation of a bacterium with a 21 amino acid genetic code. *J. Am. Chem. Soc.* **2003,** *125* (4), 935-939.
27. Wang, L.; Schultz, P. G., Expanding the genetic code. *Angew. Chem. Int. Ed.* **2005,** *44* (1), 34-66.
28. Kolb, H. C.; Sharpless, K. B., The growing impact of click chemistry on drug discovery. *Drug Discov. Today* **2003,** *8* (24), 1128-1137.
29. Letelier, M. E.; Lepe, A. M.; Faúndez, M.; Salazar, J.; Marín, R.; Aracena, P.; Speisky, H., Possible mechanisms underlying copper-induced damage in biological membranes leading to cellular toxicity. *Chem.-Biol. Interact.* **2005,** *151* (2), 71-82.
30. Agard, N. J.; Prescher, J. A.; Bertozzi, C. R., A strain-promoted [3+ 2] azide− alkyne cycloaddition for covalent modification of biomolecules in living systems. *J. Am. Chem. Soc.* **2004,** *126* (46), 15046-15047.
31. Saxon, E.; Bertozzi, C. R., Cell surface engineering by a modified Staudinger reaction. *Science* **2000,** *287* (5460), 2007-2010.
32. Hooker, J. M.; Esser-Kahn, A. P.; Francis, M. B., Modification of aniline containing proteins using an oxidative coupling strategy. *J. Am. Chem. Soc.* **2006,** *128* (49), 15558-15559.
33. Carrico, Z. M.; Romanini, D. W.; Mehl, R. A.; Francis, M. B., Oxidative coupling of peptides to a virus capsid containing unnatural amino acids. *Chem. Commun.* **2008,** (10), 1205-1207.

34. Behrens, C. R.; Hooker, J. M.; Obermeyer, A. C.; Romanini, D. W.; Katz, E. M.; Francis, M. B., Rapid chemoselective bioconjugation through oxidative coupling of anilines and aminophenols. *J. Am. Chem. Soc.* **2011,** *133* (41), 16398-16401.
35. Obermeyer, A. C.; Jarman, J. B.; Netirojjanakul, C.; El Muslemany, K.; Francis, M. B., Mild bioconjugation through the oxidative coupling of ortho-aminophenols and anilines with ferricyanide. *Angew. Chem. Int. Ed.* **2014,** *53* (4), 1057-1061.
36. Capehart, S. L.; ElSohly, A. M.; Obermeyer, A. C.; Francis, M. B., Bioconjugation of gold nanoparticles through the oxidative coupling of ortho-aminophenols and anilines. *Bioconjug. Chem.* **2014,** *25* (10), 1888-1892.
37. Obermeyer, A. C.; Capehart, S. L.; Jarman, J. B.; Francis, M. B., Multivalent viral capsids with internal cargo for fibrin imaging. *PLoS One* **2014,** *9* (6), e100678.
38. ElSohly, A. M.; Francis, M. B., Development of oxidative coupling strategies for site-selective protein modification. *Acc. Chem. Res.* **2015,** *48* (7), 1971-1978.
39. Li, X.; Zhang, L.; Hall, S. E.; Tam, J. P., A new ligation method for N-terminal tryptophan-containing peptides using the Pictet–Spengler reaction. *Tetrahedron Lett.* **2000,** *41* (21), 4069-4073.
40. Dawson, P. E.; Muir, T. W.; Clark-Lewis, I.; Kent, S., Synthesis of proteins by native chemical ligation. *Science* **1994,** *266* (5186), 776-779.
41. Geoghegan, K. F.; Stroh, J. G., Site-directed conjugation of nonpeptide groups to peptides and proteins via periodate oxidation of a 2-amino alcohol. Application to modification at N-terminal serine. *Bioconjug. Chem.* **1992,** *3* (2), 138-146.
42. Ning, X.; Temming, R. P.; Dommerholt, J.; Guo, J.; Ania, D. B.; Debets, M. F.; Wolfert, M. A.; Boons, G. J.; van Delft, F. L., Protein Modification by Strain-Promoted Alkyne–Nitrone Cycloaddition. *Angew. Chem. Int. Ed.* **2010,** *49* (17), 3065-3068.
43. Gilmore, J. M.; Scheck, R. A.; Esser-Kahn, A. P.; Joshi, N. S.; Francis, M. B., N-terminal protein modification through a biomimetic transamination reaction. *Angew. Chem. Int. Ed.* **2006,** *45* (32), 5307-5311.
44. Witus, L. S.; Netirojjanakul, C.; Palla, K. S.; Muehl, E. M.; Weng, C.-H.; Iavarone, A. T.; Francis, M. B., Site-specific protein transamination using N-methylpyridinium-4-carboxaldehyde. *J. Am. Chem. Soc.* **2013,** *135* (45), 17223-17229.
45. Dixon, H. B.; Fields, R., [33] Specific modification of NH2-terminal residues by transamination. In *Methods Enzymol.*, Elsevier: 1972; Vol. 25, pp 409-419.
46. Dixon, H., N-terminal modification of proteins—a review. *J. Protein Chem.* **1984,** *3* (1), 99-108.
47. Obermeyer, A. C.; Jarman, J. B.; Francis, M. B., N-terminal modification of proteins with o-aminophenols. *J. Am. Chem. Soc.* **2014,** *136* (27), 9572-9579.
48. MacDonald, J. I.; Munch, H. K.; Moore, T.; Francis, M. B., One-step site-specific modification of native proteins with 2-pyridinecarboxyaldehydes. *Nat. Chem. Biol.* **2015,** *11* (5), 326.

Chapter 2

Direct detection of nitrotyrosine-containing proteins using
an aniline-based oxidative coupling strategy

Abstract

A convenient two-step method is described for the detection of nitrotyrosine-containing proteins. First, nitrotyrosines are reduced to aminophenols using sodium dithionite. Following this, an oxidative coupling reaction is used to attach anilines bearing fluorescence reporters or affinity probes. Features of this approach include fast reaction times, pmol-level sensitivity, and excellent chemoselectivity.

This chapter is based on the following publication.
Sangsuwan, R.; Obermeyer, A. C.; Tachachartvanich, P.; Palaniappan, K. K.; Francis, M. B., Direct detection of nitrotyrosine-containing proteins using an aniline-based oxidative coupling strategy. *Chem. Commun.* **2016,** *52* (65), 10036-10039.

2.1. Introduction

Tyrosine nitration is a covalent post-translational protein modification that is associated with oxidative stress.[1-3] The formation of this group typically involves peroxynitrite, which is generated by the reaction between superoxide anions (O_2^-) and nitric oxide (NO).[4] Other reactive nitrogen species, such as NO_2^+, NO_2Cl, and NO_2, are also suggested as possible sources of tyrosine nitration in living cells.[5,6] This modification is of current interest due to its potential to modulate changes in protein function, location, and turn over.[7] Recently, it was shown that 3-nitrotyrosine (Y-NO_2) has been implicated in a large number of disorders, such as Parkinson's disease, Alzheimer's disease, and over 50 others.[7]

The continued study of this posttranslational modification creates a need for reliable analytical methods that can be used to study protein nitration, and, indeed, several approaches have been developed for this purpose.[8] These include Western blotting with anti-Y-NO_2 antibodies,[9] direct analysis using mass spectrometry,[10] and chemical detection methods based on biotinylation,[11] resin capture,[12,13] and fluorophore formation.[14-16]. In the latter cases,[11-16] the Y-NO_2 groups are reduced to aminophenols (Y-NH_2) with dithionite[17] prior to probe attachment. In our lab, aminophenols have been central to a series of chemoselective oxidative coupling (OC) strategies (section 1.3),[18,19] which have involved the facile addition of anilines to iminoquinone intermediates resulting from oxidation with periodate[18,20] or ferricyanide, Figure 2.1, step 2.[21] A closely related strategy has been reported for Y-NO_2 detection by the Schöneich group,[15,16] in which benzylamine derivatives are coupled to the same intermediate. Their strategy leads to the formation of a new fluorophore and the simultaneous attachment of a diol handle for use in boron-based affinity chromatography. However, these nucleophiles require pH 9 for the addition to occur, and are in competition with other amines in solution and on the protein itself. This can lead to multiple products and potential crosslinking that can complicate sample analysis.

Based on our experience with the OC strategy for the preparation of complex bioconjugates,[21] we hypothesized that suitably-functionalized anilines would offer higher coupling yields at lower pH values, and the formation of single products rather than mixtures. The reaction exhibits very high rates and chemoselectivities at pH 6.0 to 6.5, and is compatible with glycoproteins,[22] nucleic acids,[20] and complex molecular structures (section 1.3).[19] However, the chemoselectivity of the aniline-based OC strategy has not been tested for use in protein mixtures, and the concentration limits of the reaction have not yet been determined. In this report, we show that this aniline-based OC method is indeed sufficiently selective to label proteins with Y-NH_2 groups in crude lysates with no additional protecting group strategies, leading to pmol-level detection limits.

2.2. Results and discussion

2.2.1. Design of oxidative coupling mediated detection probes for Y-NO_2

The overall strategy for Y-NO_2 detection involved the development of the two-step protocol outlined in Figure 2.1. In the first step, Y-NO_2 groups are reduced to Y-NH_2 in the presence of sodium dithionite. This is then followed by an OC with a suitably functionalized aniline probe: either fluoresceinamine (**2a**) or a biotinylated aniline (**2b**). In much of our previous work, sodium periodate was used as an oxidant for the OC reaction due to its high coupling rates. However, ferricyanide was used for this application because (1) it forms a single reaction product,

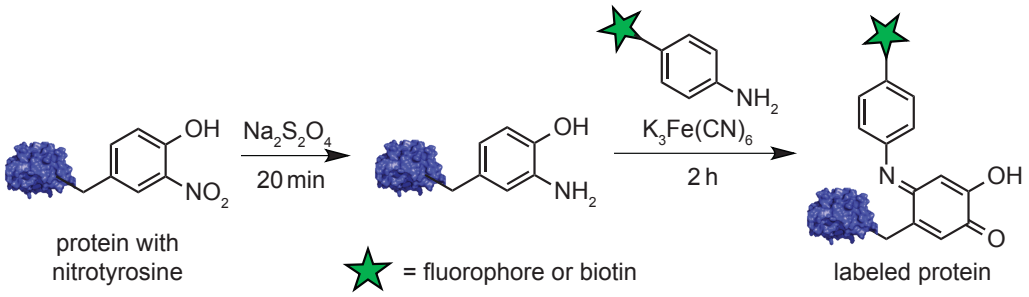


Figure 2.1. Detection of nitrotyrosine-containing proteins using a two-step protocol.

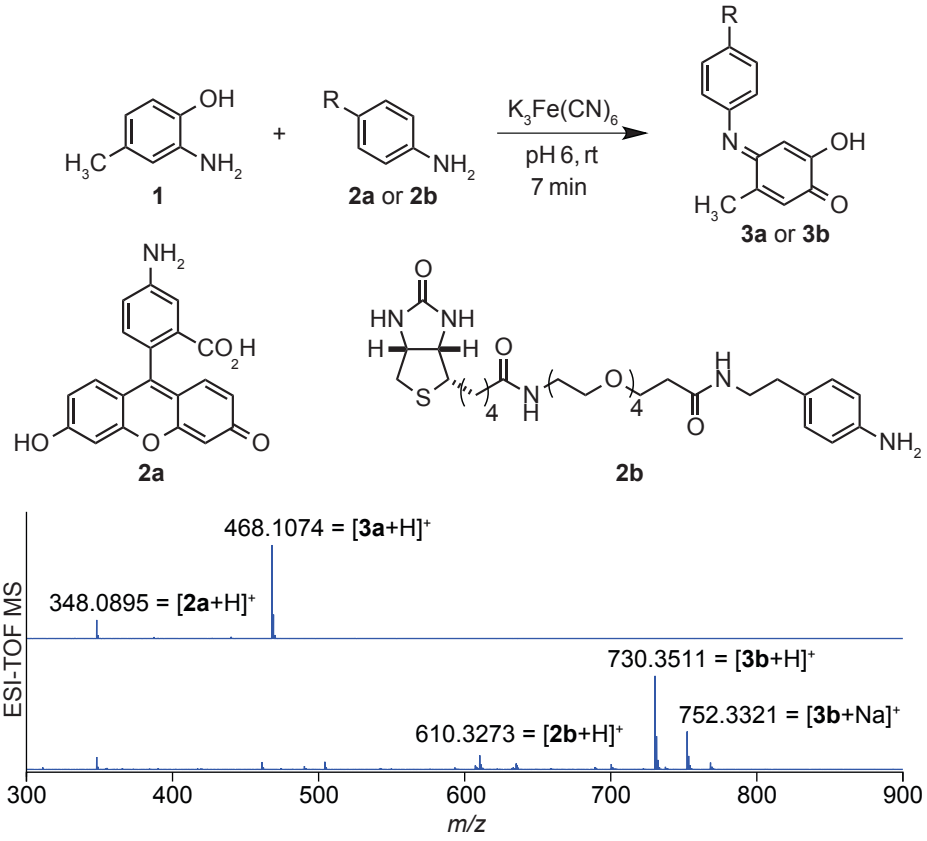


Figure 2.2. Oxidative coupling reactions between *o*-aminophenols and functionalized anilines. The ESI-TOF MS analysis of the reactions of *o*-aminophenol **1** with fluoresceinamine (**2a**) or biotin aniline (**2b)** shows clean conversion to the respective oxidative coupling products (expected $[M+H]^+$ *m/z* = 468.1078 for **3a** and 730.3480 for **3b**). The reactions were run with 200 µM functionalized anilines, 3 µM **1**, and 80 mM ferricyanide in 100 mM phosphate buffer, pH 6.

while periodate generates a mixture of two products, and (2) it does not oxidize thiols, 1,2-diols, tryptophans, N-termini, or other functionalities on protein surfaces (section 1.3).[21] For the aniline component, fluoresceinamine was selected due to the convenience of its commercial availability. In addition, the corresponding biotin aniline (**2b**) was synthesized in one step from EZ-Link NHS-

PEG4-Biotin and 2-(4-aminophenyl)ethylamine. This substrate can be detected using Western blotting techniques, and it can also be used to enrich samples for proteomic analysis. To verify that the OC of **2a** and **2b** would proceed to form the anticipated products, the coupling reactions were performed using a small molecule Y-NH_2 mimic (2-amino-*p*-cresol, **1**). After only 7 min, both of the functionalized anilines cleanly coupled to **1** in the presence of ferricyanide, as determined by ESI-TOF MS (Figure 2.2).

2.2.2. Confirmation of Y-NO_2 labeling chemoselectivity

After confirming the reactivity patterns of the two aniline probes, their ability to label nitrated protein substrates was investigated. First, Y-NO_2 was incorporated site-selectively into the protein monomers of MS2 viral capsids (MS2-T19(Y-NO_2)) using amber stop codon suppression.[20,23-25] Interestingly, ESI-TOF MS analysis of the initially obtained protein indicated that some Y-NO_2 groups had been reduced during expression (Figure 2.3b [iv]). To convert the remaining Y-NO_2 groups to Y-NH_2 moieties, the sample was treated with 10 mM sodium dithionite[17] for 20 min and then purified by ultrafiltration. ESI-TOF MS analysis revealed complete reduction at this point (Figure 2.3b [v]). The specificity of the OC reaction was then evaluated by comparing wild-type (wt) MS2 coat proteins with those bearing Y-NH_2 groups. Following exposure of the samples to aniline probes **2a** or **2b** and ferricyanide at room temperature for 2 h, the reaction mixtures were analyzed using ESI-TOF MS (Figure 2.3). Coupling products were obtained only for the Y-NH_2 cases, with comparable results being obtained for both anilines (Figure 2.3b [vi, vii]). No reactivity was observed for wt MS2 with either aniline (Figure 2.3b [ii–iii]).

As a second, non-site-selective incorporation method, tyrosines in other proteins (bovine serum albumin (BSA), chymotrypsinogen, and lysozyme) were nitrated with tetranitromethane and then reduced with dithionite equivalently (supporting information Figures S2.1-S2.2 for mass spectra and details). Samples of unmodified protein, Y-NO_2 containing proteins, purified Y-NH_2 containing proteins, and Y-NH_2 containing proteins that still contained the dithionite reductant were exposed to both aniline probes (**2a** and **2b**) in the presence of ferricyanide. The reaction mixtures were then purified using ultrafiltration to remove the excess, unreacted anilines. The reaction products were analyzed by SDS-PAGE, and the labeling was detected by fluorescence imaging or immunoblotting, as appropriate (Figure 2.4). Only the Y-NH_2 modified proteins were successfully labeled with the anilines, with no significant differences being observed between the purified and unpurified samples. These results demonstrate the chemoselectivity of this two-step process. They also indicate that the dithionite reductant need not be removed from the samples before performing the OC step so long as excess ferricyanide is used.

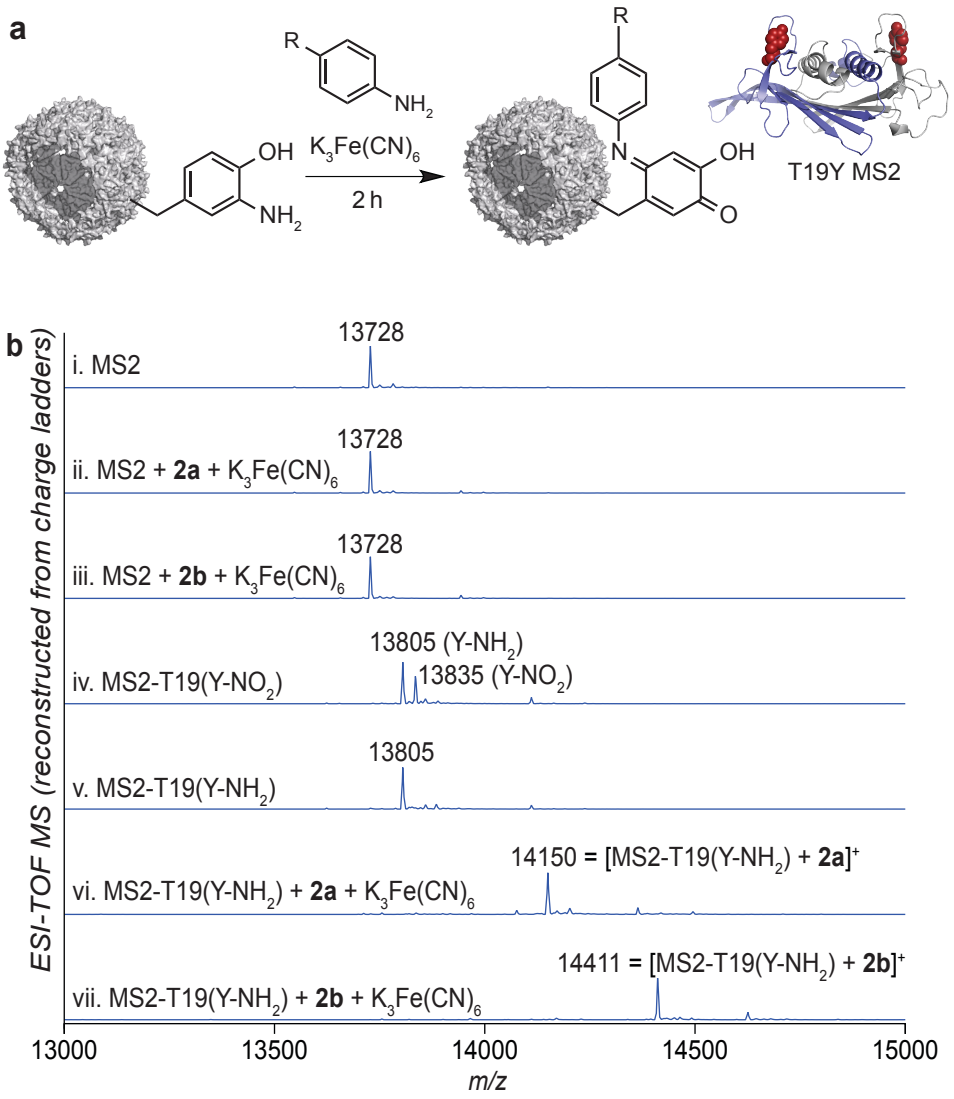


Figure 2.3. Chemoselectivity tests for the oxidative coupling reaction. (a) Two mutants of the MS2 coat protein were used: the wild-type (T19) and a variant containing an aminotyrosine side chain (T19(Y-NH_2), shown). (b) ESI-TOF MS analyses are shown for the MS2 proteins before and after oxidative coupling with functionalized anilines (**2a** and **2b**). Y-NO_2 = nitrotyrosine, Y-NH_2 = aminotyrosine. The reactions were run with 200 μM **2**, 10 μM protein (based on monomer), and 80 mM ferricyanide in 100 mM phosphate, pH 6.

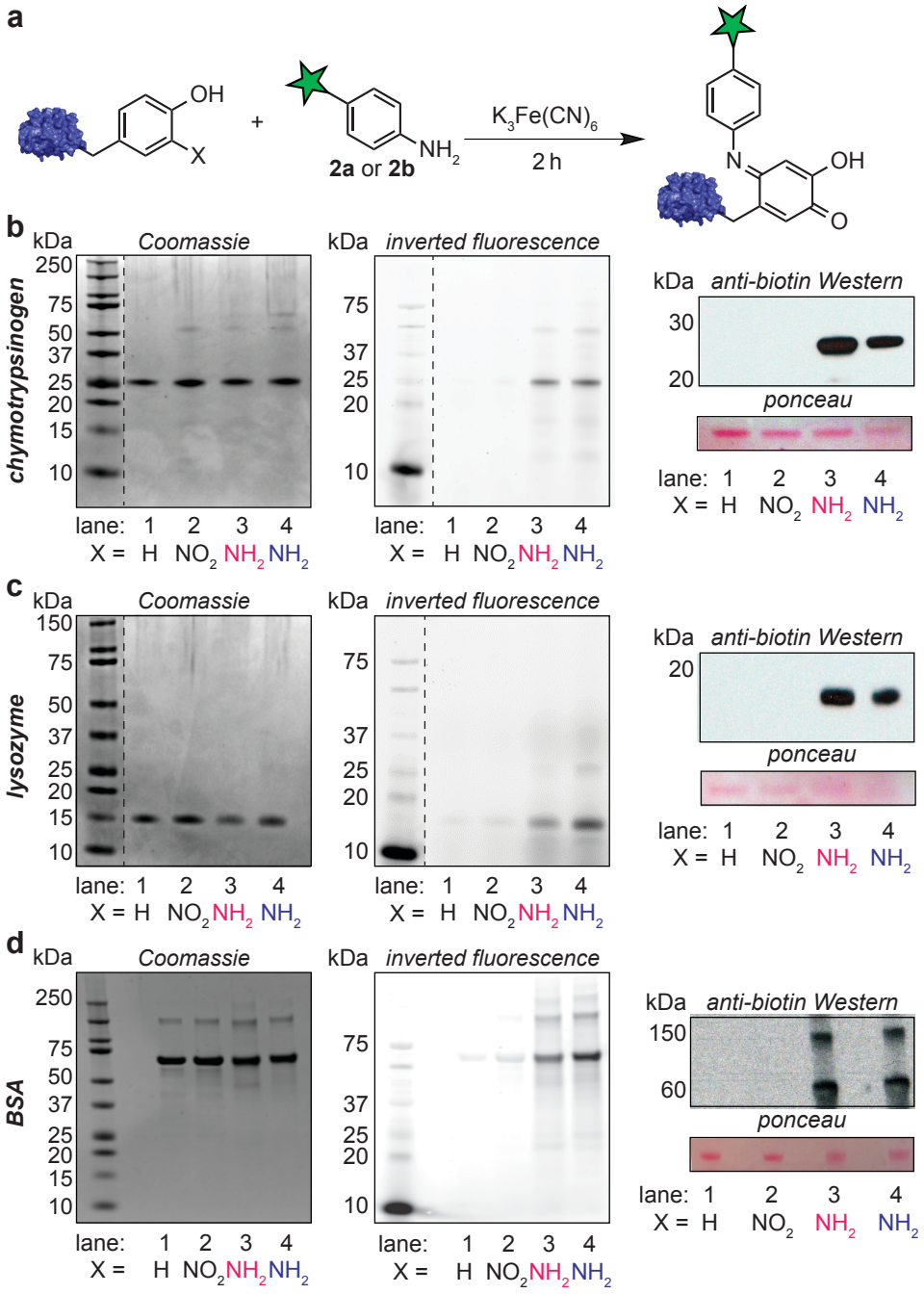


Figure 2.4. Testing the specificity of the oxidative coupling reaction for *o*-aminophenol residues on additional proteins. (a) Reactions were run with 200 µM **2a** or **2b**, 10 µM protein, and 80 mM ferricyanide in 100 mM phosphate buffer, pH 6, for 2 h. (b-d) Following SDS-PAGE, samples labeled with **2a** were visualized using fluorescence imaging (shown as an inverted image) and Coomassie stain. Samples labeled with **2b** were visualized through Western blotting with anti-biotin-HRP probes. The X labels refer to the tyrosine substituents, as defined in the structure in (a). NH_2 groups in pink indicate protein samples that were purified after dithionite reduction. NH_2 groups in blue indicate samples that were not purified between the steps. The vertical dashed lines indicate where two portions of the same gel were juxtaposed.

2.2.3. Detection of Y-NO_2 in cell lysates

Although the OC reactions were found to be specific to Y-NH_2 containing proteins, the initial reactions were run using single-component samples and relatively high concentrations of protein (10 μM). In biological samples, Y-NO_2 containing proteins are only observed at nanomolar concentrations,[26] and always in the presence of many other species. In order to adapt the method for use in cell lysates, the reaction conditions were optimized to maximize detection sensitivity (supporting information Figures S2.3-S2.4 for details). In these experiments, Y-NO_2 modified chymotrypsinogen (0.8 μM) was added to 20 μg of centrifuged Ramos cell lysate, and the mixture was subsequently reacted with excess $Na_2S_2O_4$ (0.25-4 mM, 1.25-20 equiv. relative to **2a**). Next, 25-1000 μM of aniline **2a** (31–500 equiv. relative to Y-NO_2 protein) was added in the presence of excess ferricyanide (5-80 mM, 25-400 equiv. relative to **2a**). The effects of the dithionite reduction and OC reaction times were also explored. To quench the reactions, the samples were precipitated two times by adding a $CHCl_3$-MeOH mixture.[27]

It was found that the conversion was highest using 2 mM $Na_2S_2O_4$ for 20 min, followed by the reaction with 200 μM **2a** and 80 mM $K_3Fe(CN)_6$ for 2 h. Increasing the reduction time beyond 30 min was not beneficial. It is possible that with extended reduction times sodium dithionite decomposition products add to the Y-NH_2 group and prevent the desired reaction with the aniline probe. Using more than 250 equiv. of **2a** resulted in higher levels of background labeling. We hypothesized that this was largely due to non-specific associations between the dye and the proteins. To address this issue, several methods of purification were tested, including ultrafiltration by spin concentration, protein precipitation, and a combination of the two. The use of ultrafiltration purification helped minimize the background signal (especially for high molecular weight proteins, supporting information Figures S2.6-S2.7).

To explore the compatibility with other fluorophores, additional aniline dyes based on Oregon Green 488 and TAMRA were synthesized. Both of these anilines reacted successfully with Y-NO_2 modified proteins, but they exhibited nonspecific background signals in the analysis of cell lysates even in the absence of ferricyanide. Nonetheless, the simple synthesis of the fluorescent-aniline probes suggests that this method could be generalized to accommodate many different detection strategies. Altogether, the fluoresceinamine probe was chosen for further studies, and the ultrafiltration purification step was utilized before gel analysis. Using these conditions, the amount of Ramos lysate could be increased to 72 μg for a 100 μL reaction volume while still maintaining Y-NO_2 selectivity.

To investigate the sensitivity of the detection strategy in complex cellular mixtures, known concentrations of Y-NO_2 modified proteins were added to Ramos cell lysates. The two-step procedure was then performed in the same reaction tube using the optimized conditions described above. The excess small molecules were removed by ultrafiltration after the OC step, and the reaction mixtures were analyzed by SDS-PAGE. For fluorescence detection with probe **2a**, 5-16 pmol of Y-NO_2 modified protein could be detected, depending on the target (Figure 2.5 and supporting information Figure S2.5). The limit for unambiguous detection for biotinylated aniline **2b** was established at 20 pmol after visualization with an anti-biotin-HRP Western blotting probe. The specificity for Y-NO_2 modified proteins was further verified by evaluating lysate samples that contained native versions of the protein targets. Only the Y-NO_2 versions of the proteins were detected by fluorescence and anti-biotin antibodies. For comparison purposes, an anti-Y-NO_2 antibody was used with a HRP-secondary antibody detection probe. This system showed visible

bands as low as 2 pmol (Figure 2.5). Although the direct antibody detection method can achieve lower detection limits than probe **3a**, the OC labeling method still offers advantages of fewer steps, lower costs, and significantly shorter times. The ability to attach biotin probes also opens avenues for protein enrichment, proteomic studies, and immunohistochemical staining.

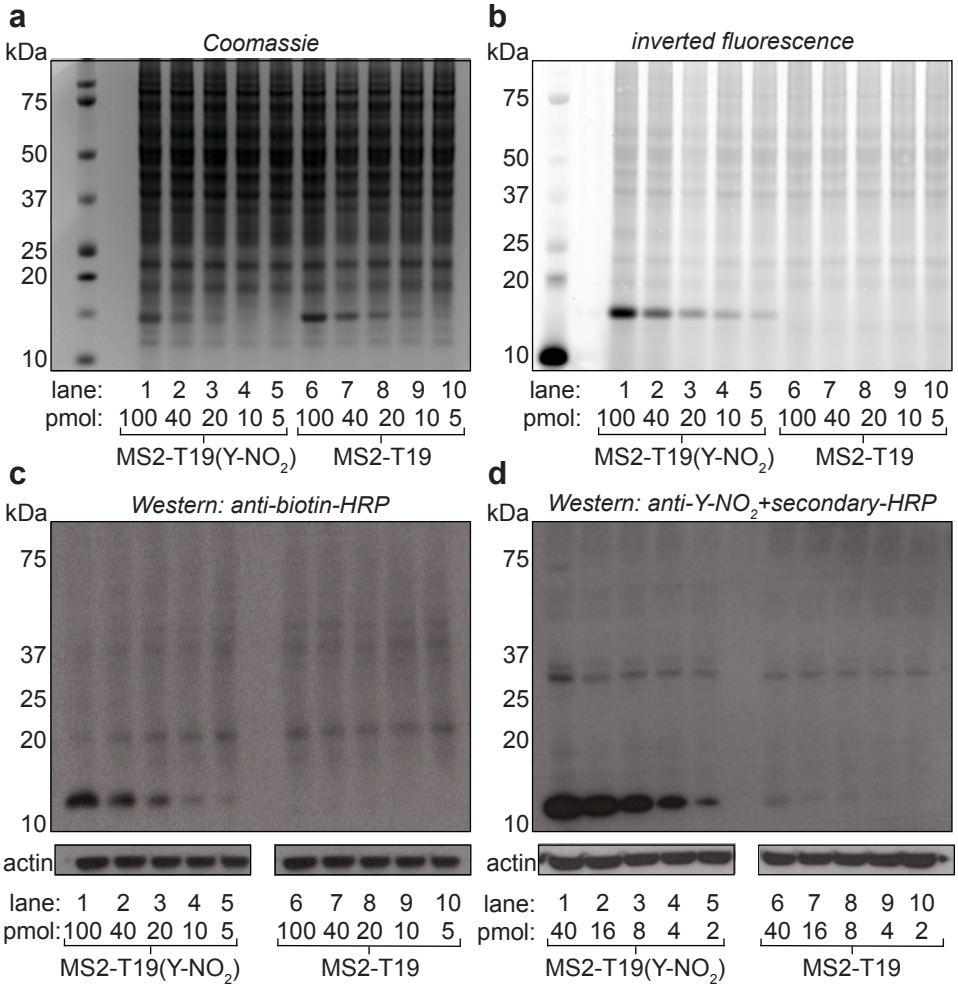


Figure 2.5. Detection of MS2-T19(Y-NO_2) in cell lysates (24 µg Ramos lysate loaded per lane). Portions of MS2-T19(Y-NO_2) were added to 72 µg of cell lysate in 100 µL of phosphate buffer, pH 6. The samples were reduced with 2 mM sodium dithionite for 20 min, and then treated with 200 µM **2a** and 80 mM ferricyanide for 2 h. After ultrafiltration, the reactions were analyzed by SDS-PAGE with (a) Coomassie staining and (b) fluorescence detection (inverted image shown). (c) The same procedure was used with **2b**, and the samples were visualized via Western blotting with an anti-biotin-HRP probe. (d) For comparison purposes, MS2-T19Y-NO_2 was also added to lysates and visualized via Western blotting with anti-nitrotyrosine and secondary antibody-HRP.

2.3. Conclusions

We have demonstrated the use of a fast aniline-based OC bioconjugation reaction for the detection of nitrotyrosine-containing proteins in cell lysates. The tandem reaction sequence occurs

in a single sample vessel in less than 2.5 h, and is likely to be compatible with a variety of different anilines. One probe used in this study (fluoresceinamine) is commercially available, and the easily-synthesized biotin tagged substrate offers both detection and the potential for enrichment of tagged proteins. This OC methodology therefore provides a valuable advance for monitoring changes in tyrosine nitration due to oxidative stress, and will likely facilitate future studies of the biological role of this posttranslational modification.

2.4. Materials and methods

2.4.1. Reagents

Unless otherwise noted, the chemicals and solvents used were of analytical grade and were used as received from commercial sources. Fluorescinamine isomer I, 5(6)-carboxytetramethyl-rhodamine *N*-succinimidyl ester, tetranitromethane, 2-(4-aminophenyl)ethylamine and $K_3Fe(CN)_6$ were purchased from Aldrich and used without further purification. Oregon Green 488 carboxylic acid, succinimidyl ester, 5-isomer was purchased from life technologies and used without further purification. Anhydrous sodium dithionite ($Na_2S_2O_4$) was purchased from Aldrich and stored in a desiccator at room temperature until use. Over the years we have noted that the storage of this reagent at reduced temperature leads to its degradation (possibly due to the condensation of water on the cold material when the bottle is opened). If stored improperly, unidentified adducts can form during the reduction step. Analytical thin layer chromatography (TLC) was performed on EM Reagent 0.25 mm silica gel 60-F254 plates with visualization by ultraviolet (UV) irradiation at 254 nm and/or potassium permanganate stain. All organic solvents were removed under reduced pressure using a rotary evaporator. Water (dd-H_2O) used as reaction solvent was deionized using a Barnstead NANOpure purification system (Thermo Fisher). Centrifugations were performed with an Eppendorf Mini Spin Plus (Eppendorf, Hauppauge).

2.4.2. Instrumentation and sample analyses

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

Gel analyses. For protein analysis, sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) was carried out on a mini gel apparatus (Novex), using a 4-12% NuPAGE 4-12% Bis-Tris gel (Novex). The sample and electrode buffers were prepared according to Laemmli. All protein electrophoresis samples were heated for 5-10 min at 95 °C in the presence of 1,4-dithiothreitol (DTT) to ensure reduction of disulfide bonds. Gels were run for 35 min at 200 V to separate the bands. Commercially available markers (Bio-Rad) were applied to at least one lane of each gel for assignment of apparent molecular masses. Visualization of protein bands was accomplished by staining with Coomassie Brilliant Blue R-250 (Bio-Rad). Gel imaging was performed on a Gel Doc EZ System (Bio-Rad). For fluorescent protein conjugates, visualization was accomplished on Typhoon 9410 variable mode imager (Amersham Biosciences) prior to gel staining. ImageJ was used to determine the level of modification by optical densitometry.

2.4.3. Synthetic procedures

*Synthesis of aniline-biotin substrate (**2b**).* To a solution of EZ-Link NHS-PEG_4-Biotin (25 mg, 0.042 mmol, Thermo Scientific) in CH_2Cl_2 was added 2-(4-aminophenyl)ethylamine (6.1 µL, 0.046 mmol) and triethylamine (12 µL, 0.084 mmol). The reaction mixture was stirred for 4 h and then diluted with CH_2Cl_2 and washed with water. The combined organic layers were dried over sodium sulfate and the solvent was removed *in vacuo*. The reaction afforded 25 mg of pale-yellow oil (quantitative yield). HRMS (ESI) calculated for $C_{29}H_{48}O_7N_5S$ ($[M+H]^+$) 610.3269, found 610.3259 m/z.

Synthesis of Oregon green 488 and TAMRA aniline. To a solution of the appropriate NHS ester dye (0.0098 mmol) in DMF was added 2-(4-aminophenyl)ethylamine (0.0118 mmol) and triethylamine (0.02 mmol). The reaction mixture was stirred for 18 h, after which it was evaporated to remove DMF. The reaction was purified by reversed phase HPLC. HRMS (ESI) of Oregon Green 488 aniline calculated for $C_{29}H_{19}F_2N_2O_6$ ($[M]^+$) 529.1217, found 529.1209 m/z. HRMS (ESI) of TAMRA aniline calculated for $C_{33}H_{33}N_4O_4^+$ ($[M]^+$) 549.2496, found 549.2501 m/z.

2.4.4. Protein modification procedures

General procedure for oxidative coupling in buffered solution. To a solution of nitrophenol modified protein (bovine serum albumin (BSA), chymotrypsinogen, lysozyme, or MS2) in 100 mM phosphate buffer, pH 6.0 was added 2 µL of a freshly prepared 100 mM solution of $Na_2S_2O_4$ in 100 mM phosphate buffer, pH 6.0 (see note about dithionite storage in the reagents Section). The solution was briefly vortexed and incubated at room temperature for 20 min. To the solution of reduced protein was added 10 µL of a 2 mM solution of aniline probe and 10 µL of a 800 mM solution of $K_3Fe(CN)_6$, for a final total volume of 100 µL. The solution was briefly vortexed and then incubated at room temperature for 2 h. The reaction mixtures were purified by using a 0.5 mL centrifugal filter with the appropriate molecular weight cut off (MWCO, Millipore). The reaction mixtures were purified by repeated (3-6 times) centrifugation before analysis by SDS-PAGE.

Oxidative coupling in cell lysates. The procedure outlined above for the oxidative coupling in buffered solution was performed in the presence of 72 µg of mammalian cell lysate (untreated, from Jurkat or Ramos cells). The reaction mixtures were purified by repeated (3-6 times) centrifugation before analysis by SDS-PAGE.

Protein nitration by tetranitromethane. Proteins (200 µM of BSA, chymotrypsinogen, or lysozyme) were dissolved in phosphate buffer saline (PBS, Gibco). Aliquots of the protein solution (0.25 ml) were diluted with an equal volume of PBS. 10 µL of tetranitromethane dissolved in ethanol (4% v/v) was then added to the protein solution. Afterwards, the reaction mixtures were gently shaken at room temperature (~25 °C) for 1 h. The modified proteins were purified by illustra NAP-5 columns (GE Healthcare).

2.4.5. Additional experimental procedures

Cell culture and preparation of cell lysates. Cell lines including Jurkat and Ramos were cultured in RPMI 1640 medium (Gibco) supplemented with 2 mM L-glutamine (Gibco), 100 I.U./ml penicillin, 100 µg/ml streptomycin (Gibco), and 10% (v/v) fetal bovine serum (FBS) (Gibco). All cell lines were maintained at 37 °C in a humidified atmosphere containing 95% air

and 5% CO_2. Cells were collected and then washed twice in 10 mL of DPBS, with centrifugation at 1300 rpms for 3 min after each wash step to pellet the cells. The cell pellet was resuspended in lysis buffer (1% NP-40, 150 mM NaCl, 20 mM Tris at pH 7.4, and protease inhibitors (Calbiochem, set III, diluted 1:100 or EDTA-free complete tablets)) and then disrupted by sonication (5 s on, 5 s off, for 1 min of sonication). The volume of lysis buffer was two times the volume of the cell pellet or 200 µL, whichever was larger. Insoluble material was removed by centrifugation (12000 rpms for 10 min at 4 °C). Protein concentration was determined using a BCA assay (Pierce).

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

Western blot protocol. Protein samples were run on a SDS-PAGE gel and then the contents of the gel were dry-transferred to nitrocellulose membranes using iBlot 2 Dry Blotting System (Novex, P0:7 min, 20-25 V). The blots were blocked overnight at 4 ^oC in 10 mL tris-buffered saline with 0.1% Tween-20 (TBST) containing bovine serum albumin (5% w/v). To the blot in blocking solution was added an anti-biotin antibody conjugated to horseradish peroxidase (#7075, cell signaling at 1:5000 dilution). The antibody was incubated with the blot for 30 min at room temperature. The blot was then washed with TBST (3 x 15 min) followed by detection using chemiluminescence using SuperSignal West Pico Chemiluminescent 98 Substrate (Pierce).

For the anti-nitrotyrosine antibody, a similar method was used: to the membrane was added anti-nitrotyrosine antibody (ab110282, abcam at 1:2000 dilution) in blocking solution. The antibody was incubated overnight at 4 ^oC. After washing with TBST, the membranes were incubated with a Goat anti-Mouse IgG H&L (ab97040, abcam at 1:50000 dilution) secondary antibody conjugated to horseradish peroxidase in blocking solution.

2.5. Supporting information

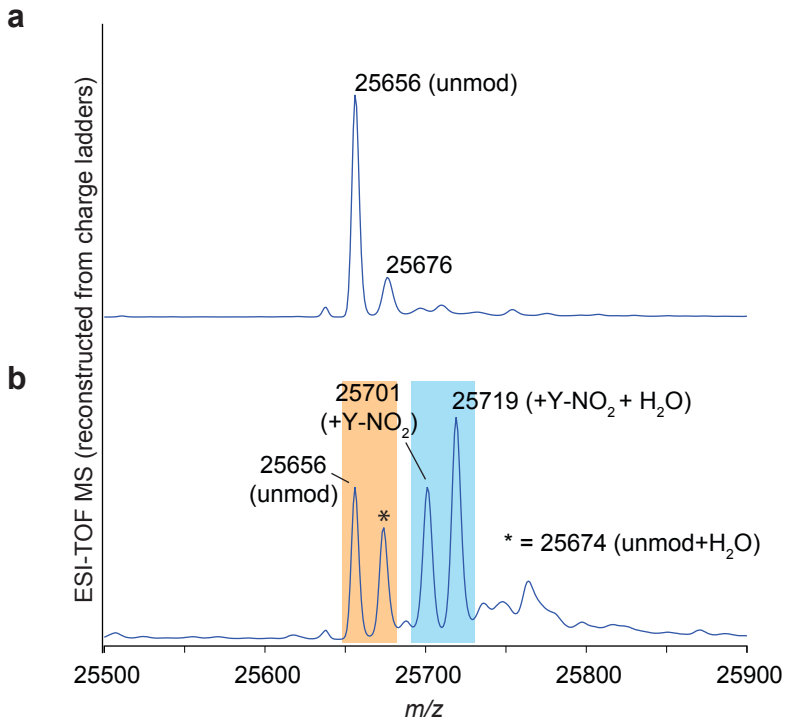


Figure S2.1. ESI-TOF MS analysis of (a) unmodified and (b) tetranitromethane modified chymotrypsinogen (Y-NO_2 chymotrypsinogen). Deconvoluted charge ladders are shown. The percent modification was calculated by integration of the shaded regions: 37% unmod (orange), 63% Y-NO_2 chymotrypsinogen (blue).

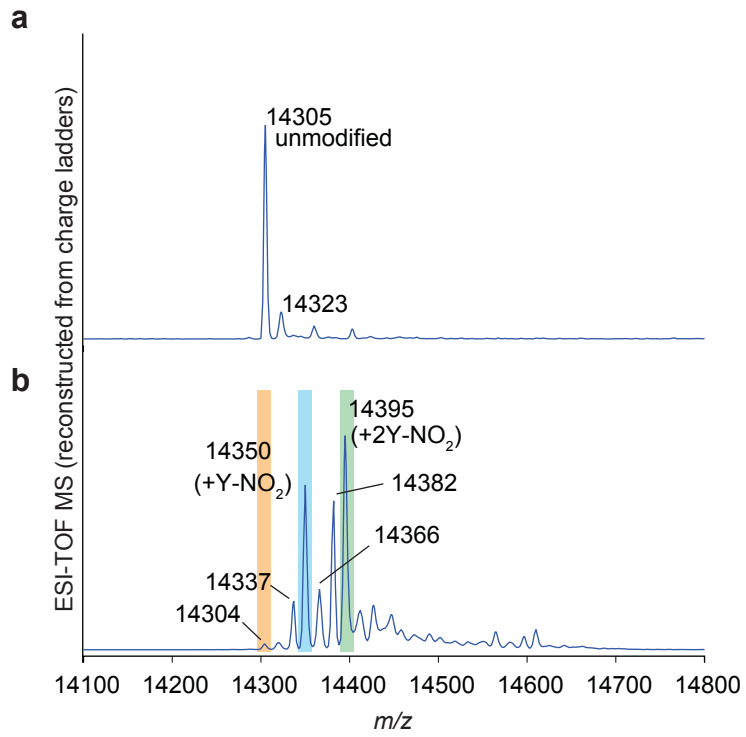
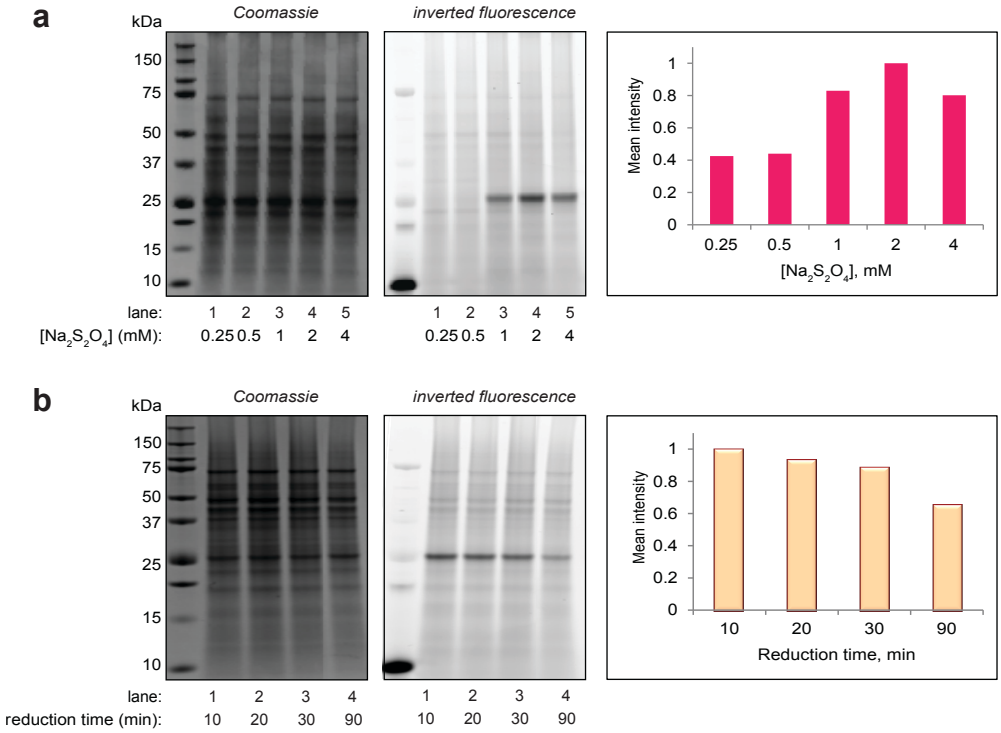


Figure S2.2. ESI-TOF MS analysis of (a) unmodified and (b) tetranitromethane modified lysozyme (Y-NO_2 lysozyme). Deconvoluted charge ladders are shown. Both singly- and doubly-modified products were observed. The percent modification was calculated by integration of the shaded regions: 2% unmod (orange), 40% Y-NO_2 lysozyme (blue), and 58% 2(Y-NO_2) lysozyme (green).



FigureS2.3. Optimization of the dithionite-mediated Y-NO_2 reduction step in cell lysates using Y-NO_2 chymotrypsinogen as a model protein. (a) To determine the optimal number of dithionite equivalents, 100 µL reduction reactions were run with 0.8 µM protein and 20 µg of Ramos lysate for 20 min with varying concentrations of sodium dithionite (0.25-4 mM). After reduction, the reaction mixtures were treated with **2a** (200 µM) in the presence of ferricyanide (80 mM) in 100 mM phosphate buffer, pH 6 for 2 h. (b) To determine the optimal reaction time, 100 µL reduction reactions were run with 0.4 µM protein in 20 µg of cell lysate with 1 mM sodium dithionite for 10-90 min. After reduction, the reaction mixtures were treated with **2a** (400 µM) in the presence of ferricyanide (80 mM) in 100 mM phosphate buffer, pH 6 for 2 h. After purification, the reaction mixtures were analyzed by SDS-PAGE with fluorescence detection. The mean intensities of the modified protein bands were quantified using ImageJ, and these values were used to determine the optimal dithionite concentration.

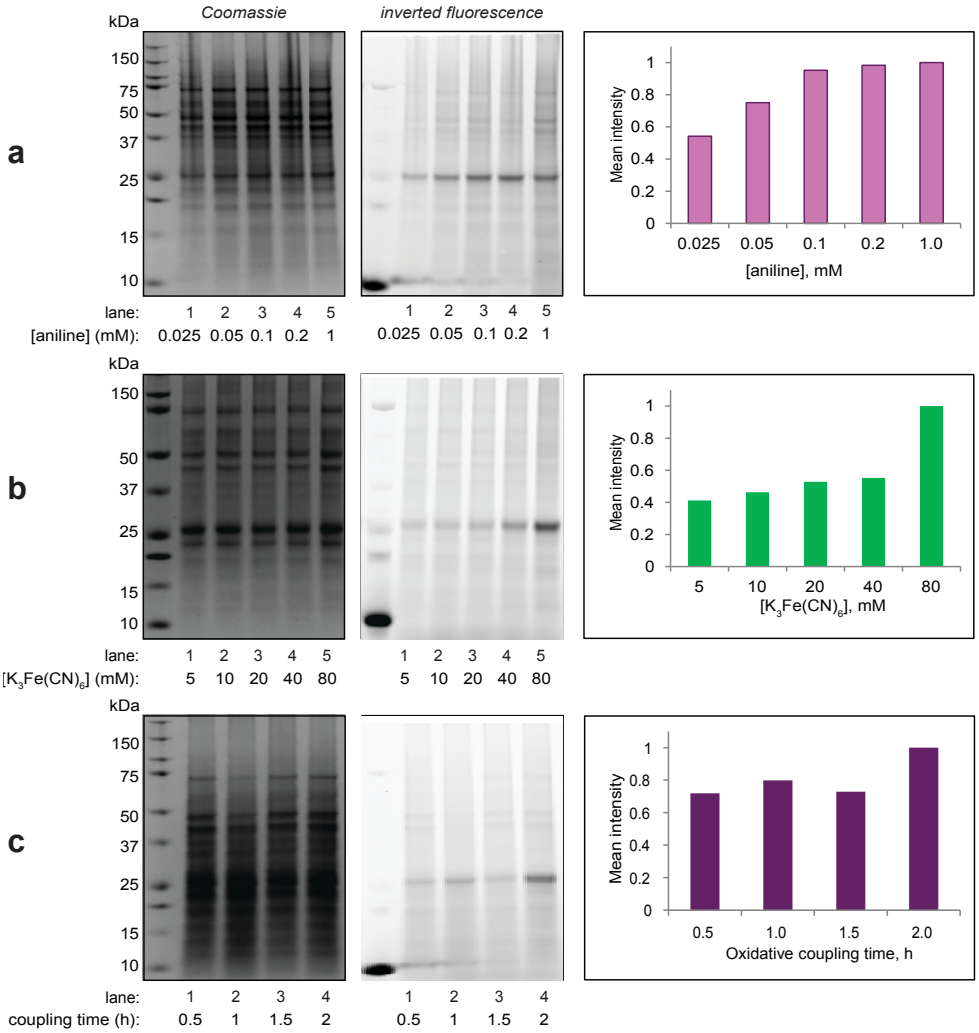


Figure S2.4. Optimization of the oxidative coupling in cell lysates with Y-NO_2 chymotrypsinogen as a model protein. The equivalents of the aniline coupling partner (a), the concentration of oxidant (b), and the duration of oxidative coupling (c) were optimized separately. (a) To determine the optimal equivalents of aniline coupling partner, 100 µL reactions were run with 0.8 µM protein and 20 µg of Ramos cell lysate with 1 mM sodium dithionite for 20 min. After reduction, the reaction mixtures were treated with varying concentrations of **2a** (25-1000 µM) in the presence of ferricyanide (80 mM) in 100 mM phosphate buffer, pH 6 for 2 h. (b) To optimize the oxidant concentration, 100 µL reactions were run with 0.8 µM protein and 20 µg of cell lysate with 2 mM sodium dithionite for 20 min. After reduction, the reaction mixtures were treated with varying concentrations of **2a** (200 µM) in the presence of ferricyanide (5-80 mM) in 100 mM phosphate buffer, pH 6 for 2 h. (c) To optimize the oxidative coupling time, 100 µL reactions were run with 0.8 µM protein and 20 µg of cell lysate with 1 mM sodium dithionite for 20 min. After reduction, the reaction mixtures were treated with varying concentrations of **2a** (400 µM) in the presence of ferricyanide (80 mM) in 100 mM phosphate buffer, pH 6 for 0.5-2 h. After purification, the reaction mixtures were analyzed by SDS-PAGE with fluorescence detection. The mean intensities of the modified protein bands were quantified using ImageJ, as shown in the graphs.

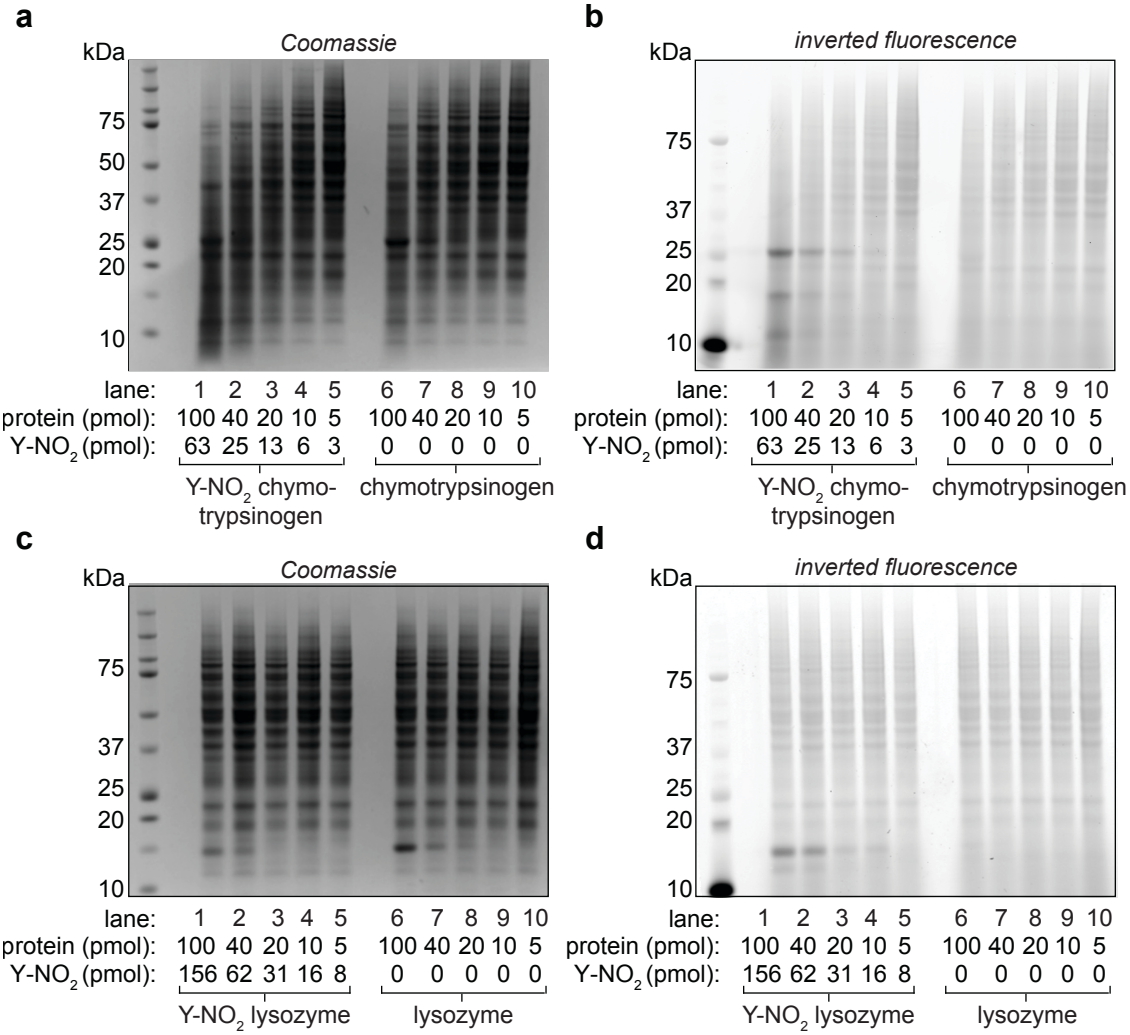


Figure S2.5. Detection of Y-NO_2 containing proteins (Y-NO_2 chymotrypsinogen and Y-NO_2 lysozyme) in cell lysates (24 µg Ramos lysate loaded per lane). Portions of each protein were added to 72 µg of cell lysate in 100 µL of phosphate buffer, pH 6. The samples were reduced with 2 mM sodium dithionite for 20 min, and then treated with 200 µM **2a** and 80 mM ferricyanide for 2 h. After ultrafiltration, the reactions were analzed by SDS-PAGE with (a, c) Coomassie staining and (b, d) fluorescence detection (inverted image shown).

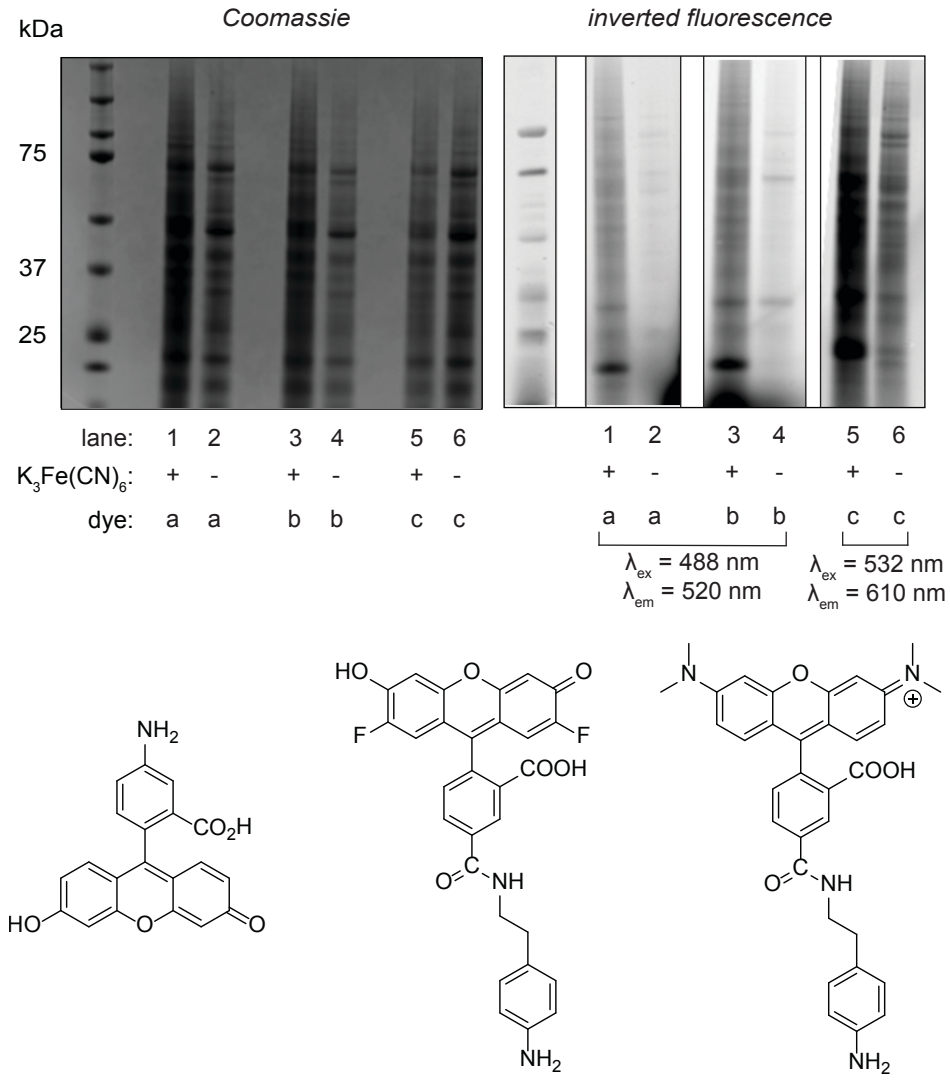


Figure S2.6. Screen of alternative fluorescent aniline dyes. The efficiencies of three aniline dyes were compared by reacting a 100 µL portion of 1 µM MS2-T19(Y-NO_2) and 72 µg cell lysate with 2 mM sodium dithionite for 20 min, followed by the addition of 200 µM aniline dye either with (+) or without (–) 80 mM ferricyanide in 100 mM phosphate buffer, pH 6. The reaction mixtures were purified by ultrafiltration and analyzed by SDS-PAGE with fluorescence detection.

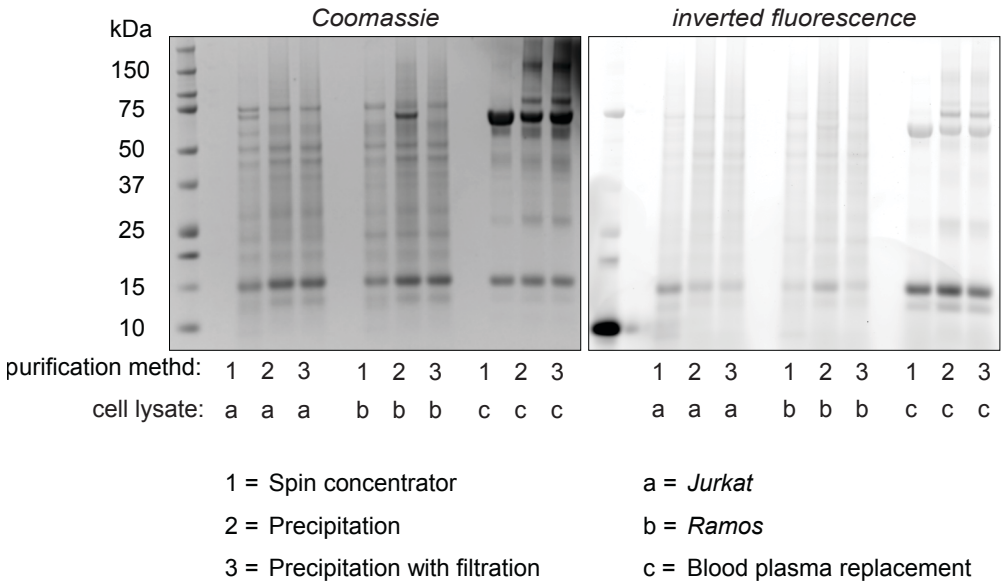


Figure S2.7. Several methods for the purification of reactions performed in complex biological samples were tested. To compare the different purification methods the oxidative coupling-based detection of Y-NO_2 lysozyme was performed in various biological mixtures including (a) Jurkat cell lysates, (b) Ramos cell lysates, and (c) a commercial blood plasma replacement. The detection reactions were performed with 10 μM Y-NO_2 lysozyme in 72 μg biological mixture with 2 mM sodium dithionite for 20 min and then 200 μM **2a** with 80 mM ferricyanide in 100 mM phosphate buffer, pH 6 for 2 h. The reaction mixtures were purified by ultrafiltration (with a 3 kDa molecular weight cutoff spin concentrator), by two rounds of precipitation with MeOH/$CHCl_3$, or by two rounds of precipitation followed by ultrafiltration. The resulting samples were analyzed by SDS-PAGE with fluorescence detection.

2.6. References

1. Abello, N.; Kerstjens, H. A.; Postma, D. S.; Bischoff, R., Protein tyrosine nitration: selectivity, physicochemical and biological consequences, denitration, and proteomics methods for the identification of tyrosine-nitrated proteins. *J. Proteome Res.* **2009,** *8* (7), 3222-3238.
2. Ischiropoulos, H., Protein tyrosine nitration—an update. *Arch. Biochem. Biophys.* **2009,** *484* (2), 117-121.
3. Souza, J. M.; Peluffo, G.; Radi, R., Protein tyrosine nitration—functional alteration or just a biomarker? *Free Radical Biol. Med.* **2008,** *45* (4), 357-366.
4. Beckman, J. S.; Ischiropoulos, H.; Zhu, L.; van der Woerd, M.; Smith, C.; Chen, J.; Harrison, J.; Martin, J. C.; Tsai, M., Kinetics of superoxide dismutase- and iron-catalyzed nitration of phenolics by peroxynitrite. *Arch. Biochem. Biophys.* **1992,** *298* (2), 438-445.
5. Vandervliet, A.; Eiserich, J. P.; Oneill, C. A.; Halliwell, B.; Cross, C. E., Tyrosine Modification by Reactive Nitrogen Species: A Closer Look. *Arch. Biochem. Biophys.* **1995,** *319* (2), 341-349.
6. van der VLIET, A.; Eiserich, J. P.; Shigenaga, M. K.; CROSS, C. E., Reactive nitrogen species and tyrosine nitration in the respiratory tract: epiphenomena or a pathobiologic mechanism of disease? *Am. J. Respir. Crit. Care Med.* **1999,** *160* (1), 1-9.
7. Greenacre, S. A. B.; Ischiropoulos, H., Tyrosine nitration: Localisation, quantification, consequences for protein function and signal transduction. *Free Radical Res.* **2001,** *34* (6), 541-581.
8. Bachi, A.; Dalle-Donne, I.; Scaloni, A., Redox proteomics: chemical principles, methodological approaches and biological/biomedical promises. *Chem. Rev.* **2012,** *113* (1), 596-698.
9. Haddad, I. Y.; Pataki, G.; Hu, P.; Galliani, C.; Beckman, J. S.; Matalon, S., Quantitation of nitrotyrosine levels in lung sections of patients and animals with acute lung injury. *J. Clin. Invest.* **1994,** *94* (6), 2407-2413.
10. Söderling, A. S.; Ryberg, H.; Gabrielsson, A.; Lärstad, M.; Torén, K.; Niari, S.; Caidahl, K., A derivatization assay using gaschromatography/negative chemical ionization tandem mass spectrometry to quantify 3-nitrotyrosine in human plasma. *J. Mass Spectrom.* **2003,** *38* (11), 1187-1196.
11. Nikov, G.; Bhat, V.; Wishnok, J. S.; Tannenbaum, S. R., Analysis of nitrated proteins by nitrotyrosine-specific affinity probes and mass spectrometry. *Anal. Biochem.* **2003,** *320* (2), 214-222.
12. Prokai, L.; Guo, J.; Prokai-Tatrai, K., Selective chemoprecipitation to enrich nitropeptides from complex proteomes for mass-spectrometric analysis. *Nat. Protoc.* **2014,** *9* (4), 882.
13. Nuriel, T.; Whitehouse, J.; Ma, Y.; Mercer, E. J.; Brown, N.; Gross, S. S., ANSID: A Solid-Phase Proteomic Approach for Identification and Relative Quantification of Aromatic Nitration Sites. *Front. Chem.* **2016,** *3*, 70.
14. Wisastra, R.; Poelstra, K.; Bischoff, R.; Maarsingh, H.; Haisma, H. J.; Dekker, F. J., Antibody-Free Detection of Protein Tyrosine Nitration in Tissue Sections. *ChemBioChem* **2011,** *12* (13), 2016-2020.
15. Dremina, E. S.; Li, X.; Galeva, N. A.; Sharov, V. S.; Stobaugh, J. F.; Schöneich, C., A methodology for simultaneous fluorogenic derivatization and boronate affinity enrichment of 3-nitrotyrosine-containing peptides. *Anal. Biochem.* **2011,** *418* (2), 184-196.
16. Sharov, V.; Pal, R.; Dremina, E.; Michaelis, E.; Schöneich, C., Fluorogenic tagging of protein 3-nitrotyrosine with 4-(aminomethyl) benzene sulfonate in tissues: A useful

alternative to Immunohistochemistry for fluorescence microscopy imaging of protein nitration. *Free Radical Biol. Med.* **2012,** *53* (10), 1877-1885.
17. Jaffe, C. L.; Lis, H.; Sharon, N., New cleavable photoreactive heterobifunctional cross linking reagents for studying membrane organization. *Biochemistry* **1980,** *19* (19), 4423-4429.
18. Hooker, J. M.; Esser-Kahn, A. P.; Francis, M. B., Modification of aniline containing proteins using an oxidative coupling strategy. *J. Am. Chem. Soc.* **2006,** *128* (49), 15558-15559.
19. ElSohly, A. M.; Francis, M. B., Development of oxidative coupling strategies for site-selective protein modification. *Acc. Chem. Res.* **2015,** *48* (7), 1971-1978.
20. Behrens, C. R.; Hooker, J. M.; Obermeyer, A. C.; Romanini, D. W.; Katz, E. M.; Francis, M. B., Rapid chemoselective bioconjugation through oxidative coupling of anilines and aminophenols. *J. Am. Chem. Soc.* **2011,** *133* (41), 16398-16401.
21. Obermeyer, A. C.; Jarman, J. B.; Netirojjanakul, C.; El Muslemany, K.; Francis, M. B., Mild bioconjugation through the oxidative coupling of ortho-aminophenols and anilines with ferricyanide. *Angew. Chem. Int. Ed.* **2014,** *53* (4), 1057-1061.
22. ElSohly, A. M.; Netirojjanakul, C.; Aanei, I. L.; Jager, A.; Bendall, S. C.; Farkas, M. E.; Nolan, G. P.; Francis, M. B., Synthetically Modified Viral Capsids as Versatile Carriers for Use in Antibody-Based Cell Targeting. *Bioconjug. Chem.* **2015,** *26* (8), 1590-1596.
23. Wang, L.; Brock, A.; Herberich, B.; Schultz, P. G., Expanding the genetic code of Escherichia coli. *Science* **2001,** *292* (5516), 498-500.
24. Mehl, R. A.; Anderson, J. C.; Santoro, S. W.; Wang, L.; Martin, A. B.; King, D. S.; Horn, D. M.; Schultz, P. G., Generation of a bacterium with a 21 amino acid genetic code. *J. Am. Chem. Soc.* **2003,** *125* (4), 935-939.
25. Carrico, Z. M.; Romanini, D. W.; Mehl, R. A.; Francis, M. B., Oxidative coupling of peptides to a virus capsid containing unnatural amino acids. *Chem. Commun.* **2008,** (10), 1205-1207.
26. Ryberg, H.; Caidahl, K., Chromatographic and mass spectrometric methods for quantitative determination of 3-nitrotyrosine in biological samples and their application to human samples. *J. Chromatogr. B* **2007,** *851* (1–2), 160-171.
27. Wessel, D.; Flügge, U.-I., A method for the quantitative recovery of protein in dilute solution in the presence of detergents and lipids. *Anal. Biochem.* **1984,** *138* (1), 141-143.

Chapter 3

Cytosolic delivery of proteins using amphiphilic polymers with 2-pyridinecarboxaldehyde groups for site-selective attachment

Abstract

Protein-based drugs are a promising class of therapeutics, but poor membrane permeability typically limits their application to extracellular receptors. Delivery strategies that can transport functional proteins to reach intracellular targets are needed, but with many current approaches, biomolecules become entrapped in the endosomes. This greatly reduces the effective concentrations of therapeutic agents at the target sites. Herein, we report a bioconjugation-based approach for intracellular protein delivery by site-selectively attaching amphiphilic polymers to the N-terminal positions of proteins using 2-pyridinecarboxaldehyde groups. The reaction is simple and features mild, aqueous conditions with no required genetic engineering of the proteins. Imaging studies demonstrate that the polymer-protein conjugates are successfully delivered into the cytosol of various cancer cell lines, mainly through a membrane fusion mechanism. When conjugated to the delivery polymers, the activity of modified RNase A was retained and notably promoted cytotoxicity in cancer cells upon delivery to the cytosol. This work therefore provides a promising platform for protein-based material delivery for therapeutic applications.

3.1. Introduction

Protein-based drugs are gaining significant interest due to their potential to engage macromolecular targets that are difficult to address using small molecules.[1-3] Since insulin was approved as the first recombinant protein therapeutic,[4] multiple engineered proteins have been developed for the treatment of diseases including hepatitis,[5] hemophilia,[6] and cancers.[7-10] Generally, proteins are membrane impermeable due to their large size,[11] which restricts their therapeutic use to extracellular targets. Proteins that are taken up intracellularly most often enter through the endocytosis pathway, meaning that after internalization, the proteins are likely to be trapped in vesicles fated for lysosomal degradation.[12] Only a small number of therapeutic agents can escape from the endosomes to access the cytosol, and in most cases this severely reduces the biological activity that can be achieved for intracellular targets.[12] The development of efficient methods that deliver greater amounts of therapeutics to intracellular targets is thus critically important for the advancement of new biomolecule-based therapies.

A number of protein delivery technologies have emerged to address this limitation, such as cationic oligomers,[13] microgels,[14] nanoparticles,[15-17] and the direct incorporation of peptide transduction domains.[18-19] In the context of polymers, several approaches have been explored for assisted cytosolic delivery.[20-21] As one particularly promising lead, negatively charged and pH responsive polyanion-based amphiphiles have previously been developed for delivery applications through endosomal escape mechanisms.[20, 22-25] In a recent example, an amphiphilic polymer was used to coat quantum dots, which were transferred into the living cells.[26-32] This polymer has a water-soluble hydrophilic backbone partly substituted with hydrophobic alkyl chains that can embed into cell membranes.[26-32] Moreover, the polymaleic anhydride backbone used in this study is commercially available and provides a convenient route for incorporating additional synthetic groups into the polymer. Based on this precedent, we became interested in endowing these polymers with synthetic groups capable of site-specific bioconjugation, thus expanding their use to include a wide range of proteins of interest.

Existing bioconjugation reactions have contributed to the synthesis of many protein-polymer materials, though these methods often rely on nonspecific conjugations that can interfere with active sites or modify critical amino acid residues, resulting in an alteration of protein function. Ideally, one desires a well-defined protein bioconjugation reaction that occurs in water under mild pH and temperature conditions while also providing a high degree of selectivity to maintain biological efficacy. The reaction between the N-termini of proteins and 2-pyridinecarboxaldehyde (2PCA) has been reported by our laboratory as a one-step, site-specific method for protein modification, Figure 3.1.[33] This reaction features mild conditions and does not require genetic engineering of the protein if the N-terminal residue is exposed. Importantly, the N-terminal chemoselectivity of the reaction has been confirmed using protein digestion and MS/MS analyses.[33] This reaction has been successfully used to conjugate extracellular matrix (ECM) proteins to 2PCA-functionalized polyacrylamide hydrogels for cell adhesion, and the properties of the modified ECM proteins were still maintained after modification.[34] Herein, we apply this N-terminal 2PCA modification reaction to prepare protein-polymer conjugates for cytosolic delivery (Figure 3.1). In initial studies**,** we attached green fluorescein protein (GFP) to polymers with different molar ratios of 2PCA groups and investigated the intracellular delivery of the conjugates into multiple types of cancer cells. The optimized construct showed substantial levels of delivery of the protein to the cytoplasm. Mechanistic studies further suggested that the conjugate likely

avoids endosomal entrapment, instead entering the cell through a direct membrane interaction pathway. We used this polymer system to deliver ribonuclease A (RNase A), which is a protein that promotes cell death through cytosolic RNA degradation. We found that RNase A-polymer conjugates significantly decreased the viability of cancer cells upon delivery, suggesting that this strategy could likely be used in a number of future therapeutic applications.

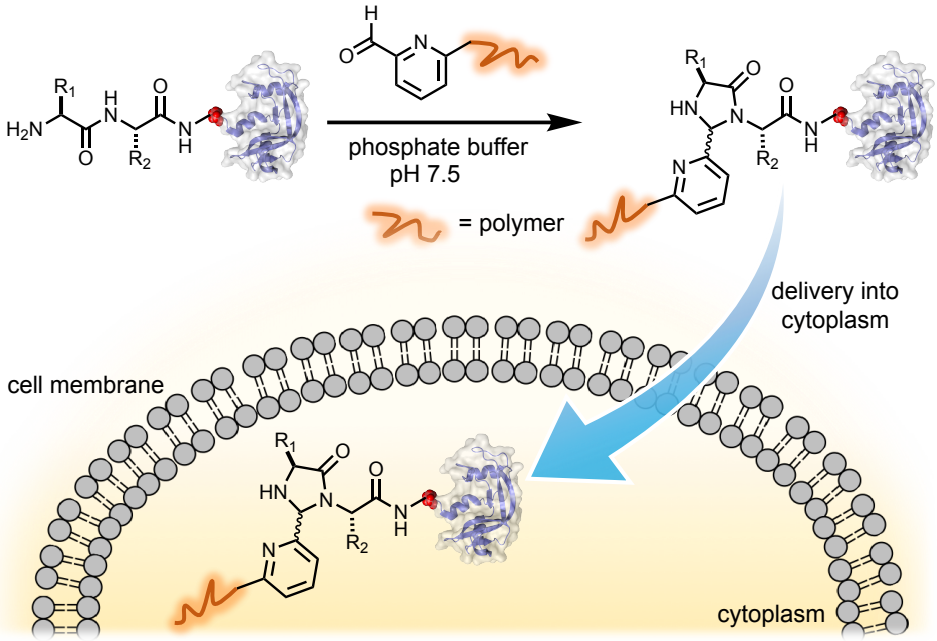


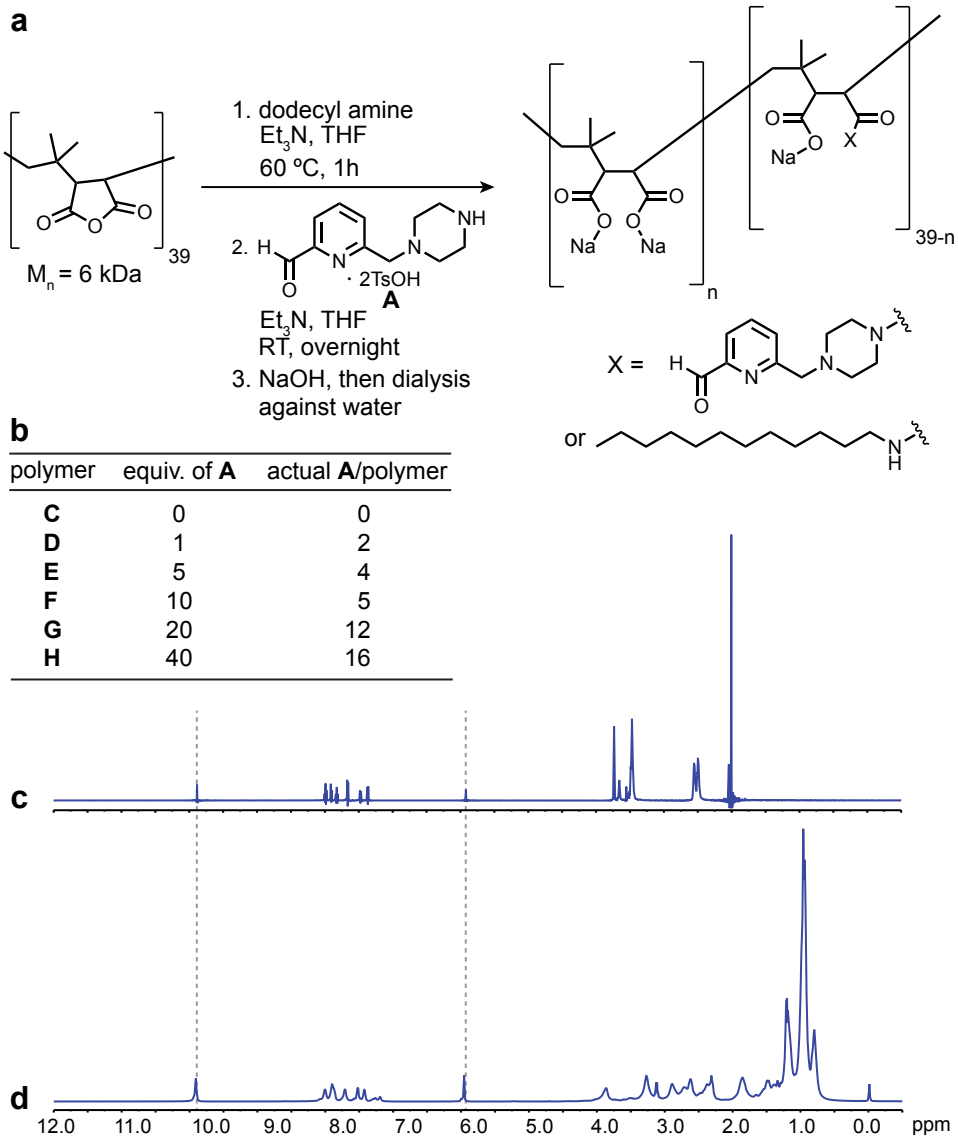
Figure 3.1. Schematic illustration of protein-polymer conjugate preparation using 2-pyridinecarboxaldehyde for N-terminal modification.

3.2. Results and discussion

3.2.1. Polymer-protein conjugate design and polymer synthesis

The polymer synthesis was modified from previous methods (Figure 3.2).[26-32] The maleic anhydride ring opening was carried out in the presence of reactive amines (dodecylamine and piperazinyl 2PCA, **A**) and trimethylamine. Following this, the remaining anhydride groups were hydrolyzed using aqueous NaOH. The reaction mixtures were subsequently dialyzed against water and lyophilized. The polymer products were found to be completely water soluble. We hypothesized that different amounts of attached 2PCA would give different reactivities for the protein bioconjugation reaction, and this could influence the polymer-protein conjugate internalization. We therefore varied the ratio of **A** from 0-40 equivalents with respect to the polymer strand to form polymers **C**-**H**, as indicated in Figure 3.2b. NMR analysis of polymers **D**-**H** displayed characteristic aldehyde peaks, which aligned well with those of a 2PCA derivative **B**, Figure 3.2c-d and supporting information Figure S3.1. The peak at 6 ppm indicated the corresponding aldehyde hydrate, which is always observed for these compounds in aqueous samples. The NMR peak intensities (compared to the aliphatic protons) at 10 ppm increased

from polymer **C** to **H**, suggesting that a higher amount of 2PCA moiety was incorporated into the polymer backbone (supporting information Figure S3.1). Elemental analysis was used to determine



polymer	equiv. of **A**	actual **A**/polymer
C	0	0
D	1	2
E	5	4
F	10	5
G	20	12
H	40	16

Figure 3.2. Synthesis and characterization of 2PCA-containing polymers. (a) A ring-opening of poly(maleic anhydride) was performed using dodecylamine (6 equiv.) followed by 2PCA-piperazine reagent **A** in the presence of triethylamine. (b) For the second amine addition, the amount of **A** was added at 0, 1, 5, 10, 20, and 40 equiv. relative to the individual polymer strands. Using elemental analysis, the resulting levels of **A** incorporation were calculated. ^{1}H NMR spectra are shown for (c) the small-molecule 2PCA **B** and (d) polymer **H** in D_2O. The peaks corresponding to the aldehydes and their corresponding hydrates are indicated with dotted lines. For clarity, the water signal has been suppressed.

the percentage of C, H, and N atoms, and this information was used to estimate the numbers of 2PCA groups and alkyl chains (Figure 3.2b and supporting information Tables S3.1 and S3.2). The results suggested that increasing the amount of 2PCA influenced the magnitude of monomer modification, as expected. Small differences in the numbers of modified monomers were observed among **D-H**, which can be explained by the small molar ratios of **A** and the monomers compared to the polymers. For use in characterization and control experiments, a fluoresceinamine-substituted polymer and a 2PCA-5k polyethylene glycol (PEG) polymer[33] (**I**) were also synthesized (see supporting information for details).

3.2.2. 2PCA conjugation to GFP

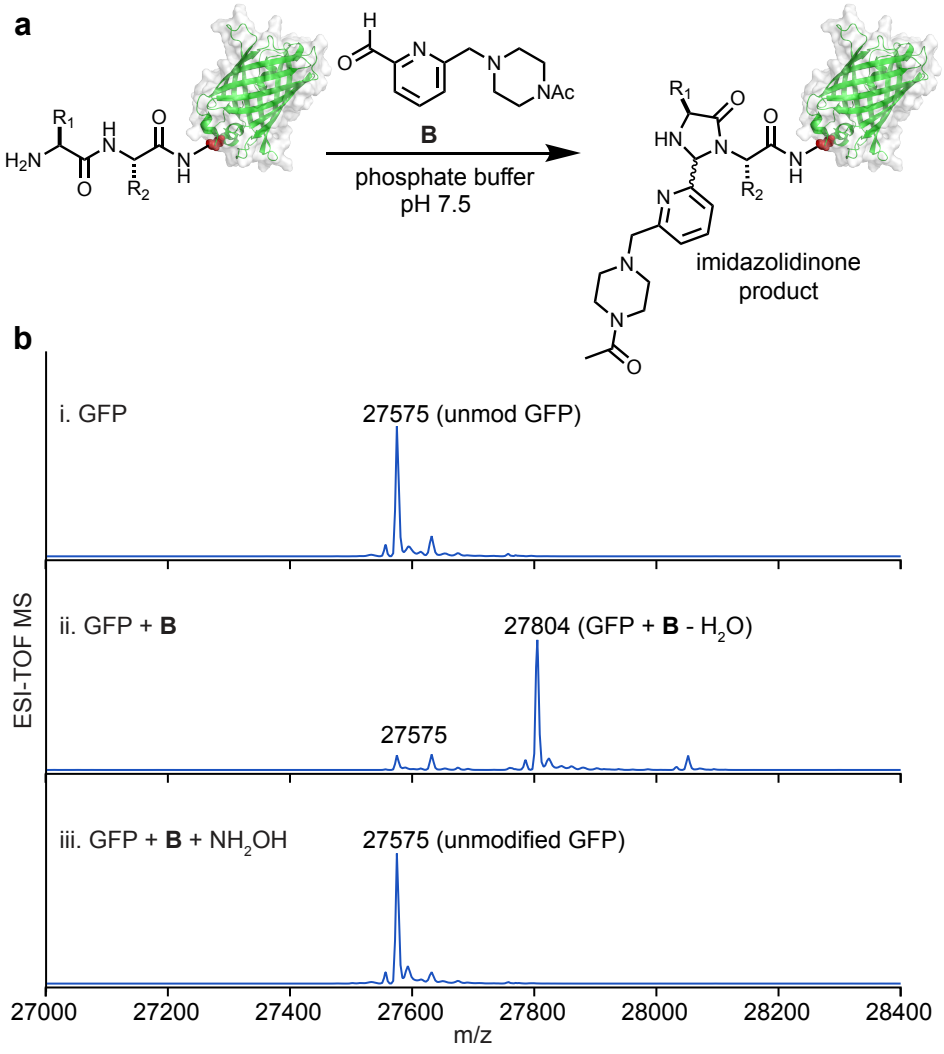


Figure 3.3. Coupling protein cargo to 2PCA-substituted polymers. (a) 2PCA reacts with N-terminal residues to form imidazolidinones. (b) ESI-TOF MS analyses are shown for GFP before and after the reaction with 2PCA reagent **B** or for **B** in the presence of NH_2OH as a coupling inhibitor. The reactions were run in 10 mM of **B** and 10 μM of protein, with or without 20 mM of NH_2OH in 50 mM phosphate buffer, pH 7.5 at 37 °C for 18 h.

The 2PCA reaction generates an imidazolidinone product through the nucleophilic attack on the initially formed N-terminal imine by the neighboring amide nitrogen on the protein backbone.[33] The chemoselectivity and reactivity of the 2PCA reaction was confirmed using 2PCA small molecule **B**, which cleanly coupled to GFP as seen in Figure 3.3. GFP was chosen as the model protein because it also provided a fluorescent marker that would allow us to track intracellular protein localization. To verify whether the attachment occurred through the 2PCA reaction mechanism, hydroxylamine (NH_2OH) was reacted with an aldehyde group of 2PCA, which resulted in oxime formation. As expected, no N-terminal reactivity was observed in the presence of NH_2OH, as determined by ESI-TOF MS (Figure 3.3b).

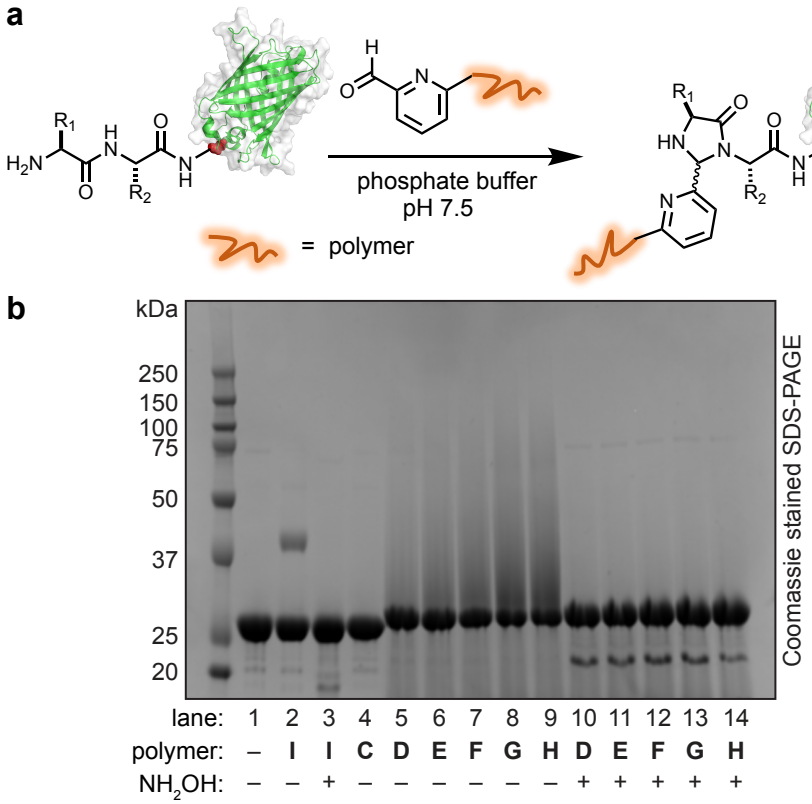


Figure 3.4. Coupling protein cargo to 2PCA-substituted polymers. (a) Coupling reactions were run with 20 mg/mL of polymer samples and 100 μM of GFP with or without 200 mM of NH_2OH in 50 mM phosphate buffer, pH 7.5 for 18 h at 37 °C. (b) Following SDS-PAGE, samples were visualized using Coomassie stain. **C** = polymer containing dodecylamine (no 2PCA), **D**–**H** = 2PCA-polymers (see Figure 3.2 for functionalization levels), **I** = 2PCA-5k PEG.

We next optimized the conditions for protein-polymer conjugation according to previous work.[33] In general, the protein concentration did not significantly affect the reaction, and when 2PCA was present in millimolar concentrations, the reaction proceeded smoothly.[33] To confirm reactivity in the polymer context, GFP was mixed with the polymers (**C**-**H** or 2PCA-5k PEG **I**) for 18 h at 37 °C, and the products were analyzed by SDS-PAGE. The selectivity of the 2PCA groups was confirmed (Figure 3.4) using polymer **C** (no 2PCA) and 2PCA-5k PEG **I** as negative and positive controls, respectively. As expected, polymer **C** did not conjugate to GFP, as no modified protein band was observed by SDS-PAGE, and polymer **I** afforded a single new high-molecular-

weight band by SDS-PAGE, representing PEGylated GFP protein. Co-incubation with NH_2OH did not show any 2PCA reactivity, which was consistent with the selectivity test performed with 2PCA mimic **B** (Figure 3.3b). Polymers **D**-**H** afforded a smeared band of modified proteins on SDS-PAGE due to the polydisperse nature of the polymers. Consistent with the findings for the PEG-2PCA reagent, no higher-molecular-weight bands were observed in the presence of NH_2OH. The GFP-polymer conjugates were then purified to remove excess polymer via Ni-NTA affinity chromatography, and the remaining 2PCA aldehyde groups were inactivated with hydroxylamine before being carried forward to the intracellular delivery studies.

3.2.3. Intracellular uptake of protein-polymer conjugates and uptake mechanism studies

The internalization of the GFP-polymer conjugates was investigated using the HeLa cell line as an initial model. Following 2 h of exposure at 10 μM, confocal microscopy analysis indicated that the GFP molecules conjugated to polymer **H** were efficiently delivered into the cytosolic compartment, while GFP alone did not show detectable levels of delivery (Figure 3.5a [i and ii] and z-stacking images in supporting information Figure S3.2). Increasing the equivalents of 2PCA on the polymers clearly increased the degree of conjugation and promoted greater cellular uptake of the proteins, as observed by the increased fluorescence (supporting information Figure S3.3). However, there was little difference observed in internalization levels for the different polymers among **D**, **E**, and **F** or between **G** and **H**, which agreed well with the elemental analysis results that all polymers had a similar amount of 2PCA on the polymer backbone (Figure 3.2b). The delivery of the fluorescein-tagged polymer at 5 μM was confirmed as a positive control (Figure 3.5a [iii]). Polymer **H** was used for further study as it showed the most efficient delivery levels.

Flow cytometry was also used to investigate the protein delivery, showing a >100-fold difference in cellular uptake between unmodified GFP and the GFP-polymer conjugate, Figure 3.5b. An additional toxicity study of both the conjugate and the free polymer was carried out to ensure that the concentrations used for the cell treatment did not interfere with cell viability (supporting information Figure S3.4). In addition to HeLa cells, the protein conjugate could be delivered into other cancer cell types (lung [A549], breast [MCF7], prostate [PC-3], and brain [U-87 MG]), with similar efficiencies, demonstrating that the approach is likely to be widely applicable (supporting information Figures S3.5–S3.8).

The uptake mechanism of the conjugates was also assessed (Figure 3.6). Inhibitors of clathrin dependent endocytosis (chlorpromazine and wortmannin)[35-39] did not significantly attenuate intracellular internalization of the GFP-polymer conjugate. Other inhibitors of caveolae dependent endocytosis and macropinocytosis uptake pathways also did not block the translocation of GFP-polymer conjugate (supporting information Figure S3.9). To confirm that the polymer did not compromise the integrity of the cell membranes, cells were co-treated with free GFP and free polymer that was inactivated with hydroxylamine, and no intracellular internalization of GFP was seen. Interestingly, cholesterol depletion using methyl-β-cyclodextrin, an inhibitor of membrane fusion uptake and endocytosis, substantially lowered the degree of GFP-polymer delivery.[40-41]

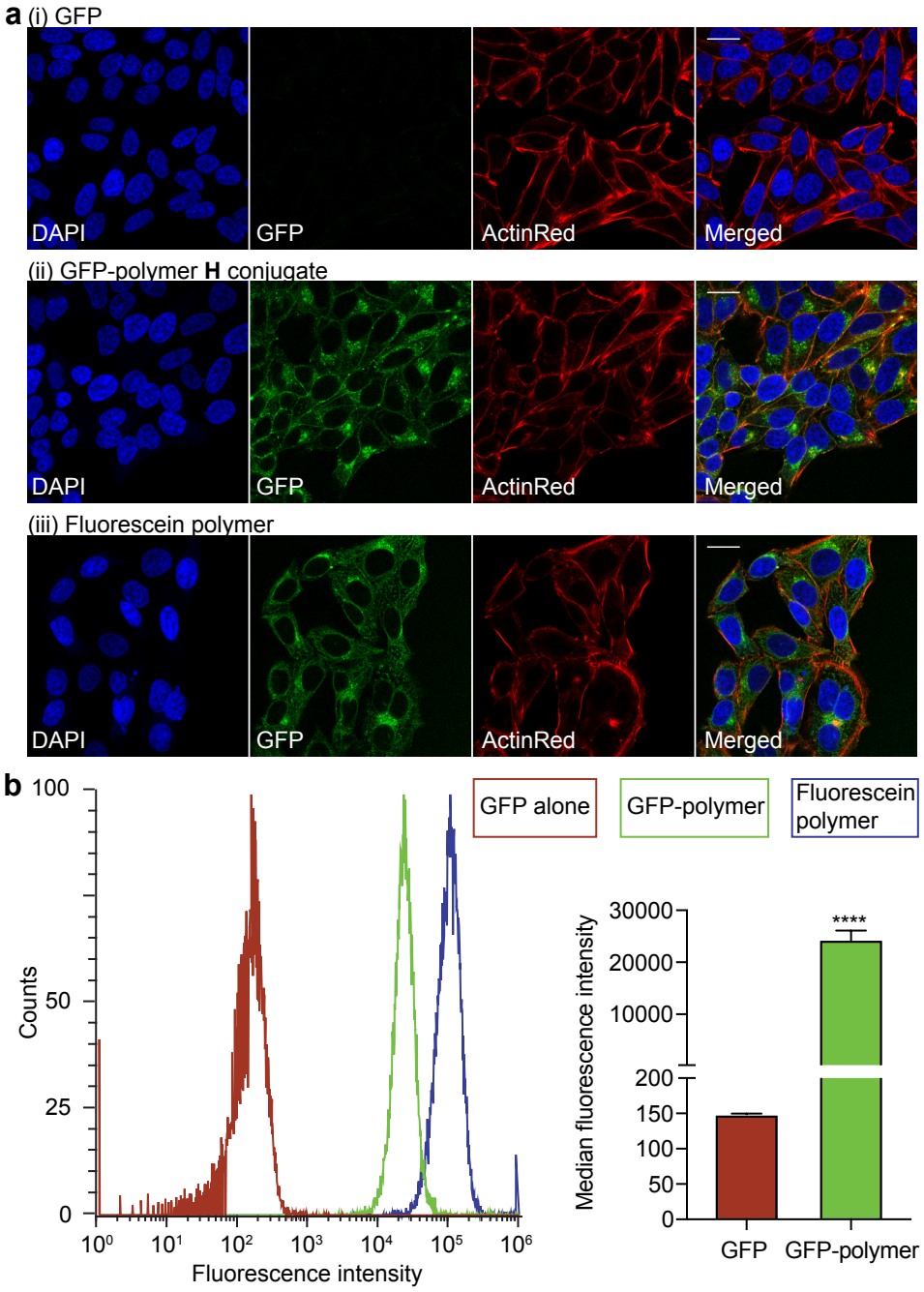


Figure 3.5. Uptake of GFP-labeled polymers into living cells. (a) Confocal microscopy images are shown for HeLa cells after exposure to GFP-polymer conjugates. Nuclear and cytoplasmic stains were performed on fixed cells using DAPI and ActinRed, respectively. The scale bars represent 20 μm. (b) HeLa cells were treated with GFP alone (10 μM), GFP conjugated to polymer **H** (10 μM), or fluorescein-labeled polymer (5 μM) for 2 h at 37 °C. Flow cytometry data are shown, along with the medians of the fluorescence intensity values. Error bars represent the standard errors resulting from three independent experiments. ****$p < 0.0001$ indicates a significant difference between GFP-polymer-treated cells and unmodified GFP-treated cells.

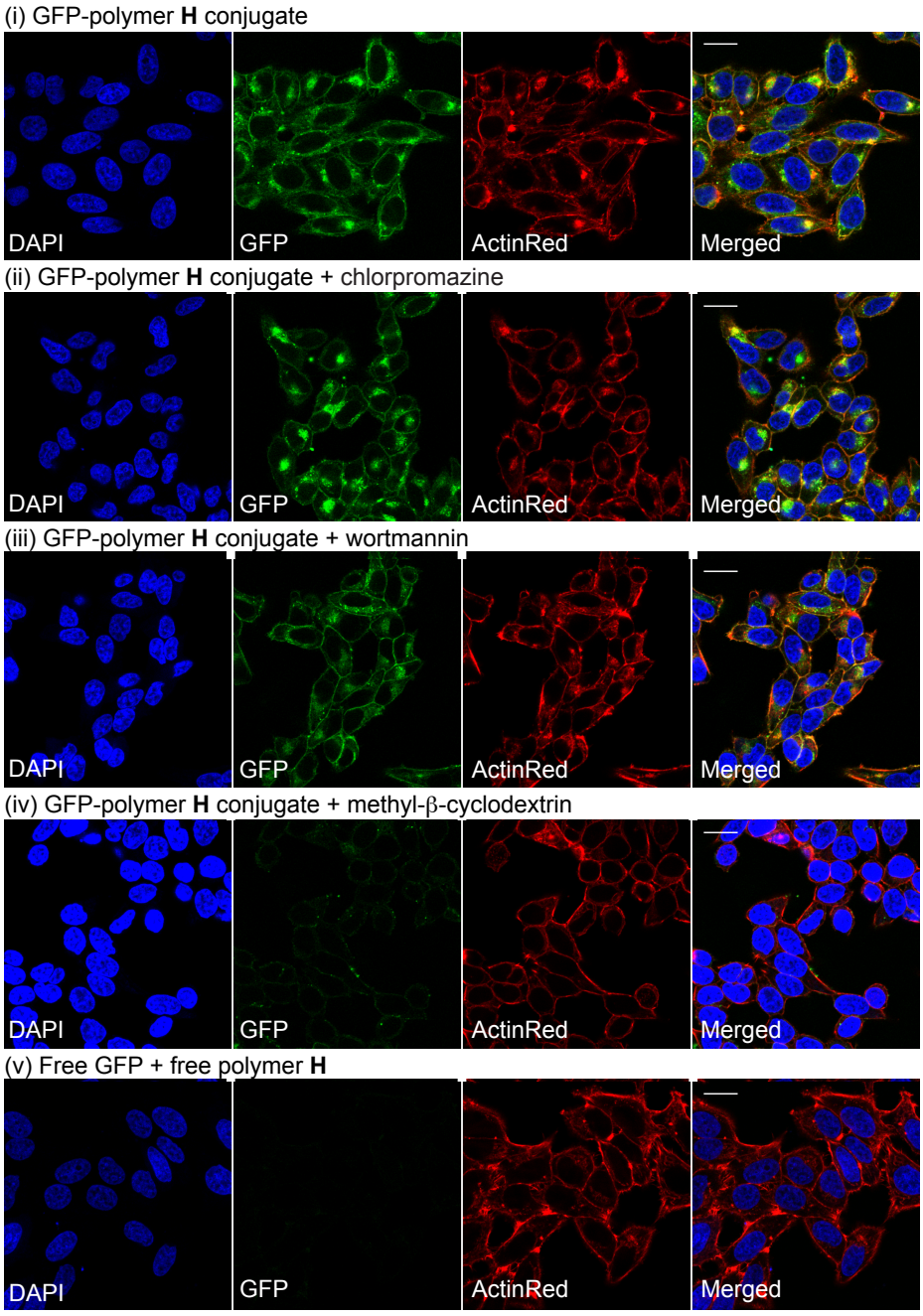


Figure 3.6. Probing the uptake mechanism of GFP-polymer conjugates in HeLa cells using confocal microscopy. HeLa cells were pretreated with corresponding inhibitors for 30 min prior to a 2 h co-treatment with GFP conjugated to polymer **H** and inhibitors (5.7 mM of methyl-*β*-cyclodextran, 10 μM of chlorpromazine, or 0.35 μM of wortmannin). As a control experiment, cells were co-incubated with 10 μM of free GFP and 10 μM of polymer **H** (with no attached protein). Nuclear and cytoplasmic stains were performed on fixed cells using DAPI and ActinRed, respectively. The scale bars represent 20 μm.

Collectively, these results suggest that at least some of the uptake of the GFP-polymer conjugate did not occur through a membrane disruption process following clathrin/caveolae dependent endocytosis, but instead occurred through a membrane fusion mechanism that is independent of endocytosis. This is a particularly significant finding, as such a pathway would avoid the acidification and proteolysis events that can inactivate proteins following uptake into vesicles.

3.2.4. Reactivity of 2PCA with RNase A

The results of the GFP-polymer delivery study indicated that our approach facilitated the internalization of proteins from the extracellular environment into the cytosolic compartment of the cells. Next, a study was performed to elucidate the integrity and functionality of delivered modified proteins. RNase A catalyzes the degradation of RNA strands and promotes cancer cell death if it is internalized into the cytosol.[42-44] It therefore served as a model functional protein for testing cytosolic delivery. The selectivity of the 2PCA reaction was first assessed using ESI-TOF MS and SDS-PAGE (supporting information Figures S3.10 and S3.11). The data showed that 2PCA mimic **B** did conjugate to the N-terminus of RNase A. With NH_2OH quenching, there was no reactivity observed, consistent with SDS-PAGE analyses and indicating that the reaction was again specific for this protein.

3.2.5. Enzymatic activity and cytotoxic effects of RNase A conjugates

The enzymatic activity of RNase A and the RNase A-polymer conjugate was assessed as previously described.[45] The absorbance of a solution containing RNA oligos and RNase A or the RNase-polymer conjugate was monitored at 300 nm at 1 min intervals, and the decrease in absorption intensity as a result of the RNA degradation was observed (Figure 3.7). Both RNase A and the RNase-polymer conjugate showed similar enzymatic activity, which indicated that polymer conjugation did not affect the activity of protein.

We next confirmed the integrity and functionality of the RNase-polymer conjugate in cell culture. HeLa cells were exposed to solutions of unconjugated polymer **H**, free RNase A, and the RNase A-polymer conjugate at 37 °C. Cell viability was then determined using the MTT assay. After a 6 h treatment, the RNase A-polymer conjugate exhibited a notable cytotoxic effect, while the control (no treatment), polymer **H**, and unmodified RNase A did not significantly lower cell viability, Figure 3.7b. It is worth noting that the RNase A-polymer conjugate elicited significant cytotoxicity in a dose-dependent manner, indicating that the method successfully modified and delivered functional protein intracellularly.

3.3. Conclusions

In the present work, we demonstrate a polymer designed to facilitate protein-based drug delivery. The polymer has both hydrophobic and hydrophilic properties, 2PCA moieties for conjugation, and excellent water solubility. The 2PCA conjugation is simple and occurs specifically at the N-termini of proteins, which provides site-selective functionalization for many different biomolecules in a location that minimally affects structure and function. We also show that this strategy both delivers proteins into the cytosolic compartment and preserves their integrity and functionality, which are key for therapeutics. More importantly, the internalization pathway of

the construct is likely to be mediated through a membrane fusion mechanism without membrane disruption, at least partially circumventing the protein degradation that occurs through endocytosis. The delivery is not cell-type specific, making it possible to apply this polymer conjugation technology for a wide range of applications in different cell types. This work therefore provides a promising platform for protein-based material delivery for therapeutic applications.

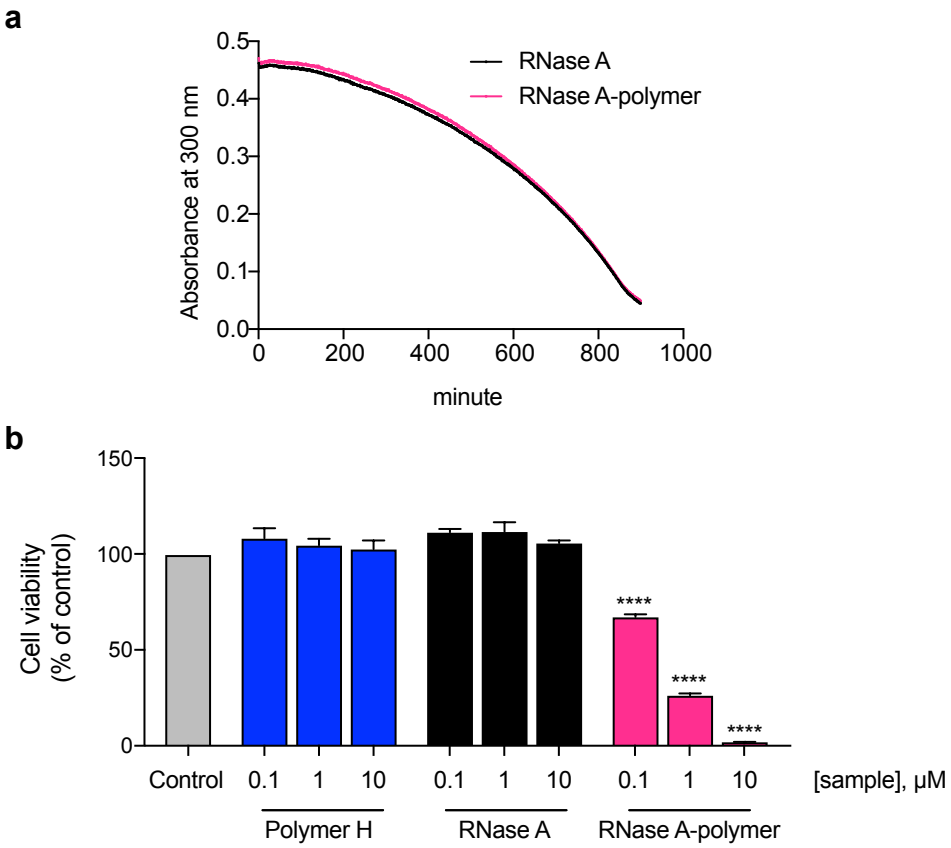


Figure 3.7. Delivery of RNase A to cells as a model cytotoxic cargo molecule. (a) The enzymatic activity of RNase A is unaffected by polymer attachment. The absorbance at 300 nm was recorded at 1 min intervals. (b) HeLa cells were incubated alone (control), with polymer **H**, or with the conjugate of RNase A to polymer **H** for 6 h at 37 °C. Viability was measured using an MTT assay and is reported as the percentage of surviving cells compared to the untreated control. Error bars represent the standard error of three independent experiments. ****$p < 0.0001$ indicates a significant difference between RNase A-polymer **H**-treated cells and unmodified RNase A-treated cells.

3.4. Materials and methods

3.4.1. Reagents

Unless otherwise noted, the chemicals and solvents used were of analytical grade and were used as received from commercial sources. Acetic anhydride, poly(isobutylene-*alt*-maleic anhydride), dodecylamine, methoxypolyethylene glycol 5,000 acetic acid NHS ester, and

ribonucleic acid from torula yeast were purchased from Aldrich and used without further purification. Fluorescein (FAM) amine, isomerically pure dye (5-isomer) was purchased from Lumiprobe and used without further purification. Analytical thin layer chromatography (TLC) was performed on EM Reagent 0.25 mm silica gel 60-F254 plates with visualization by ultraviolet (UV) irradiation at 254 nm. All organic solvents were removed under reduced pressure using a rotary evaporator. Water (dd-H_2O) used as reaction solvent was deionized using a Barnstead NANOpure purification system (Thermo Fisher). Centrifugations were performed with an Eppendorf Mini Spin Plus (Eppendorf).

3.4.2. Instrumentation and sample analyses

NMR. ^{1}H and ^{13}C spectra were measured with a Bruker AVQ400 (400 MHz, 100 MHz) and AV600 (600 MHz, 150 MHz) spectrometers.

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

Gel analyses. For protein analysis, sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) was carried out on a mini gel apparatus (Novex), using a 4-12% NuPAGE 4-12% Bis-Tris gel (Novex). The sample and electrode buffers were prepared according to Laemmli.[46] All protein electrophoresis samples were heated for 5-10 min at 95 °C in the presence of 2-mercaptoethanol to ensure reduction of the disulfide bonds. Gels were run for 35 min at 200 V to separate the bands. Commercially available markers (Precision Plus Protein Kaleidoscope Prestained Protein Standards, Bio-Rad) were applied to at least one lane of each gel for assignment of apparent molecular masses. Visualization of protein bands was accomplished by staining with Coomassie Brilliant Blue R-250 (Bio-Rad). Gel imaging was performed on a Gel Doc EZ System (Bio-Rad).

Elemental analyses. Polymer samples were analyzed by the microanalytical facility (UC Berkeley). The Perkin Elmer 2400 Series II combustion analyzer is equipped for determination of %C, %H and %N.

Statistical analyses. All data are reported as mean values ± SEM for at least three independent experiments. Statistical comparisons between each treatment and the corresponding controls were determined by a one-way analysis of variance (ANOVA), followed by Dunnett' s multiple-comparison test.

3.4.3. Synthetic procedures

*Synthesis of 6-((4-acetylpiperazin-1-yl)methyl)picolinaldehyde (**B**).* The bistosylate salt of 6-(1-piperazinylmethyl)-2-pyridinecarboxaldehyde was synthesized as previously reported.[33] To a solution of 6-(1-piperazinylmethyl)-2-pyridinecarboxaldehyde bistosylate salt (**A**, 0.5 g, 0.9 mmol) in CH_2Cl_2 was added acetic anhydride (0.2 mL, 1.8 mmol) and triethylamine (0.5 mL, 3.6 mmol). The reaction mixture was stirred for 20 min, then diluted with dichloromethane and washed

with water. The aqueous layer was extracted three times with dichloromethane. The combined organic layers were dried over sodium sulfate and the solvent was removed *in vacuo*. The reaction was purified by silica gel chromatography using hexane/ethyl acetate and the fractions were dried under vacuum to afford 0.17 g of a pale-yellow oil (80% yield). ^{1}H NMR (400 MHz, Chloroform-*d*) δ 10.01 (s, 1H), 7.87 – 7.77 (m, 2H), 7.65 (dd, J = 6.2, 2.7 Hz, 1H), 3.74 (s, 2H), 3.60 (t, J = 5.1 Hz, 2H), 3.45 (t, J = 5.1 Hz, 2H), 2.48 (dt, J = 16.6, 5.1 Hz, 4H), 2.03 (d, J = 1.5 Hz, 3H). ^{13}C NMR (100 MHz, Chloroform-*d*) δ 193.46, 168.97, 159.04, 152.34, 137.55, 127.43, 120.41, 63.84, 53.29, 52.88, 46.19, 41.31, 21.32. HRMS (ESI) calculated for $C_{13}H_{18}N_3O_2^+$ ($[M+H]^+$) 248.1394, found 248.1393 *m/z*.

Synthesis of dodecylamine-, 2-pyridinecarbozaldehyde-, and fluoresceinamine-modified poly(isobutylene-alt-maleic anhydride) derivatives. The synthesis was modified from the previously reported protocol.[26-32] A flame dried 250 mL round bottom flask was charged with 100 mg of poly(isobutylene-alt-maleic anhydride) (average molecular weight 6000 g/mol, 0.017 mmol) in 30 mL of dry THF. Dodecylamine (18 mg, 0.1 mmol, 6 equivalents relative to the polymer strands) and dry triethylamine (28 μL, 0.2 mmol) were dissolved in 5 mL of dry THF and added dropwise. The solution was stirred 1 h at 60 °C. The temperature was reduced and varying amounts of **A** (9.2, 46, 92, 183, and 366 mg: 0.017, 0.083, 0.17, 0.33, and 0.66 mmol, respectively, or 1, 5, 10, 20, 40 equivalents relative to the polymer strands, respectively) along with triethylamine (2 equivalents relative to **A**). For the fluoresceinamine-labelled polymer, fluoresceinamine isomerically pure dye (5-isomer) (16.9 mg, 0.033 mmol) along with triethylamine (9 μL, 0.066 mmol) dissolved in THF (15 mL) were added dropwise. The solution was allowed to stir overnight at room temperature. The solvent and excess triethylamine were removed under reduced pressure and the resulting residue was dissolved in aqueous 27.5 mM NaOH. The solution was dialyzed against several changes of deionized water over multiple days and lyophilized to yield 130-139 mg of the polymer products.

Synthesis of 2PCA poly(ethylene glycol, 5 kDa). The synthesis was based on a previous report, starting with MeO-PEG_{5k}-NHS ester.[33] Compound **A** (3 equivalents relative to the PEG chains) and triethylamine (9 equivalents) were used in the procedure. The reaction was concentrated under reduced pressure and extracted with dichloromethane and brine. The aqueous layer was extracted three times with dichloromethane. The combined organic layers were dried over sodium sulfate, filtered, and concentrated under reduced pressure. The resulting product was purified by precipitation from diethyl ether

3.4.4. Protein modification procedures

*General procedure for the reaction of 2PCA **B** with proteins.* A 5 μL aliquot was taken from a 1 mM solution of protein (final concentration of 100 μM) and added to 40 μL of 50 mM phosphate buffer at pH 7.5 with or without 20 mM NH_2OH (which was only used in control experiments). To the resulting solution was added a 5 μL aliquot of a 100 mM solution of **B** in H_2O (final concentration of 10 mM). The reaction was briefly agitated to ensure proper mixing and incubated at 25 °C or 37 °C without further mixing. The reaction product was purified using a 0.5 mL Amicon Ultra-4 centrifugal filter units with an appropriate molecular weight cut off (EMD Millipore). Modification was monitored by LC-MS.

General procedure for the reaction of 2-pyridinecarboxaldehyde polymers with proteins. A 100 μL aliquot was taken from a 1 mM solution of protein (final concentration of 100 μM) and added to 400 μL of 50 mM phosphate buffer at pH 7.5. To the resulting solution was added a 500 μL of 40 mg/mL solution of polymer in the same buffer. The reaction was briefly agitated to ensure proper mixing and incubated at 25 °C or 37 °C without further agitation. The reaction mixtures containing sfGFP-polymer conjugates were purified by HisPur Ni-NTA spin columns (Thermo Fisher) according to the manufacturer's instructions. For RNase A-polymer conjugates, the reactions were purified using Pierce strong cation exchange spin columns (Thermo Fisher) according to the manufacturer's instructions. The purified proteins were treated with 5 mM NH_2OH for 10 min at room temperature and then desalted against 10 mM phosphate buffer at pH 7.5 by using PD MidiTrap G-10 columns (GE Healthcare). Protein concentration was determined using a BCA assay (Thermo Fisher).

3.4.5. Additional experimental procedures

Construction of the GFP plasmid. A gene encoding superfolder GFP (sfGFP) in a pBAD vector was elongated at the N-terminus to include a Pro-His_6 sequence using Golden Gate cloning procedured. The gene encoding the Pro-sfGFP was then ligated into a pET28b vector, which contained a kanamycin resistant gene.

sfGFP Amino acid sequence:

(M)PHHHHHHRKGEELFTGVVPILVELDGDVNGHKFSVRGEGEGDATNGKLTLKFICTTG
KLPVPWPTLVTTLTYGVQCFARYPDHMKQHDFFKSAMPEGYVQERTISFKDDGTYKTR
AEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNFNSHNVYITADKQKNGIKANFKIRH
NVEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSVLSKDPNEKRDHMVLLEFVTAA
GITHGMDELYK

sfGFP encoding gene sequence:

ATGCCTCATCACCATCACCATCACCGTAAAGGCGAAGAGCTGTTCACTGGTGTCGTCC
CTATTCTGGTGGAACTGGATGGTGATGTCAACGGTCATAAGTTTTCCGTGCGTGGCGA
GGGTGAAGGTGACGCAACTAATGGTAAACTGACGCTGAAGTTCATCTGTACTACTGG
TAAACTGCCGGTACCTTGGCCGACTCTGGTAACGACGCTGACTTATGGTGTTCAGTGC
TTTGCTCGTTATCCGGACCATATGAAGCAGCATGACTTCTTCAAGTCCGCCATGCCGG
AAGGCTATGTGCAGGAACGCACGATTTCCTTTAAGGATGACGGCACGTACAAAACGC
GTGCGGAAGTGAAATTTGAAGGCGATACCCTGGTAAACCGCATTGAGCTGAAAGGCA
TTGACTTTAAAGAAGACGGCAATATCCTGGGCCATAAGCTGGAATACAATTTTAACAG
CCACAATGTTTACATCACCGCCGATAAACAAAAAAATGGCATTAAAGCGAATTTTAAA
ATTCGCCACAACGTGGAGGATGGATCTGTGCAGCTGGCTGATCACTACCAGCAAAAC
ACTCCAATCGGTGATGGTCCTGTTCTGCTGCCAGACAATCACTATCTGAGCACGCAAA
GCGTTCTGTCTAAAGATCCGAACGAGAAACGCGATCATATGGTTCTGCTGGAGTTCGT
AACCGCAGCGGGCATCACGCATGGTATGGATGAACTGTACAAA

Protein expression and purification. The expression vector encoding sfGFP was transformed into BL21-CodonPlus (DE3)-RIL *E. coli* cells, then plated on LB agar containing

kanamycin (50 μg/mL). On the following day, a single colony was used to inoculate 10 mL of LB medium, and the resulting culture was grown overnight at 37 °C and 200 rpm in a shaking incubator (12-16 h). On the following day, 10 mL of starter culture was used to inoculate 1 L of 2xYT medium in a 4 L glass flask with kanamycin at a final concentration of 50 μg/mL. Flasks were shaken at 220 rpm at 37 °C in a shaking incubator until cells reached log phase (OD_{600} = 0.6–0.8). The production of sfGFP was induced by adding IPTG to a final concentration of 0.2 mg/mL (0.8 mM), and cells were grown for 18 h at 25 °C in a shaking incubator at 220 rpm. Cells were harvested by centrifugation for 15-20 min at 6,000 rpm at 4 °C. Cell pellets, which were bright green/yellow in color, were resuspended in 15 mL of lysis buffer per 1 L of liquid growth, which was 20 mM phosphate buffer, pH 7.5, containing 500 mM NaCl and 10 mM imidazole, 0.1 mM PMSF and protease inhibitors (Pierce protease inhibitor tablet, Thermo Fisher). Cells were lysed with a sonicator (2 s on, 4 s off 10 min, 70 % amplitude) and the lysate was subjected to centrifugation immediately for 1 h at 14,000 rpm and 4 °C. The supernatant was filtered through 0.45-μm syringe filters from EMD Millipore and any solid, pelleted material was discarded. Filtered cell lysates were purified by chromatography on a Ni–NTA column (Histrap HP, 5 mL from GE Healthcare) and eluted using 20% of the elution buffer for 7 column volumes to remove impurities. Elution with 100% elution buffer over 10 column volumes recovered the pro-sfGFP. The binding and wash buffers were 20 mM phosphate buffer, pH 7.5, containing 500 mM NaCl and 10 mM imidazole. The Elution Buffer was 20 mM phosphate buffer, pH 7.5, containing 500 mM NaCl and 250 mM imidazole. Eluted fractions were collected, pooled, and desalted against 10 mM phosphate buffer at pH 7.5 by using Econo-Pac 10DG desalting prepacked gravity flow columns (Bio-Rad).

Cell culture. All cell lines (HeLa, A549, PC-3, MCF7, and U-87 MG) were cultured in DMEM (Gibco) supplemented with 2 mM L-glutamine (Gibco), 100 I.U./mL penicillin, 100 μg/mL streptomycin (Gibco), and 10% (v/v) fetal bovine serum (FBS, Gibco). For the PC-3 cell line, the media were additionally supplemented with MEM non-essential amino acids solution (1X, Gibco). For MCF7 and U-87 MG cell lines, the media were additionally supplemented with 1 mM of sodium pyruvate and MEM non-essential amino acids solution (1X, Gibco). All cell lines were maintained at 37 °C in a humidified atmosphere containing 95% air and 5% CO_2.

Flow cytometry analysis. HeLa cells were treated with sfGFP (10 μM), sfGFP-polymer conjugate (10 μM), or fluorescein-coupled polymer (5 μM) for 2 h at 37 °C. The treated cells were then washed trice with Dulbecco's phosphate buffered saline (DPBS), followed by cell dissociation buffer (Gibco) at 37 °C for 5 min. The released cells were then pelleted and resuspended in DPBS to a density of 1 x 10^6 cells/mL and analyzed by flow cytometry (Attune NxT flow cytometer, Thermo Fisher). For each sample, 20,000 cells were counted. Gating was performed by applying the auto-gating function on the major population of cells in the FSC x SSC (forward versus side scatter plots) of untreated samples; the treated samples were subject to the same gating as the untreated populations.

Confocal microscopy. Cells were washed with DPBS and trypsinized, then the trypsin was quenched with complete growth media. The resulting cells were resuspended in growth media at a concentration of 50,000 cells/mL, and 0.5 mL of the cell suspension was added to each section of an 8-well chambered coverglass (LabTek). Cells were allowed to grow at 37 °C with 5% CO_2 for 72 h. Media were aspirated, and the cells were washed twice with 0.5 mL of complete media. Samples were then added to each well at a final concentration of 1-10 μM 2PCA protein-polymer

conjugate, 0.5-10 μM of fluorescein-coupled polymer, or 1-10 μM unmodified protein in a total of 200 μL of media. The dishes were incubated at 37 °C with 5% CO_2 for 2 h, and then the cells were washed twice with 0.5 mL DPBS. The cells were fixed with DAPI (1 μg/mL, Thermo Fisher) in 4% formaldehyde in DPBS for 10 min at room temperature. The cells were washed trice with DPBS (5-min incubation), and incubated with 0.1% triton X in DPBS at room temperature for 15 min, followed by three washes with DPBS. The cells were stained with an ActinRed ready probe (Thermo Fisher) according to the manufacturer's instructions. Images were acquired on a Zeiss LSM710 laser-scanning microscope, usage courtesy of Professor Christopher Chang (U.C. Berkeley).

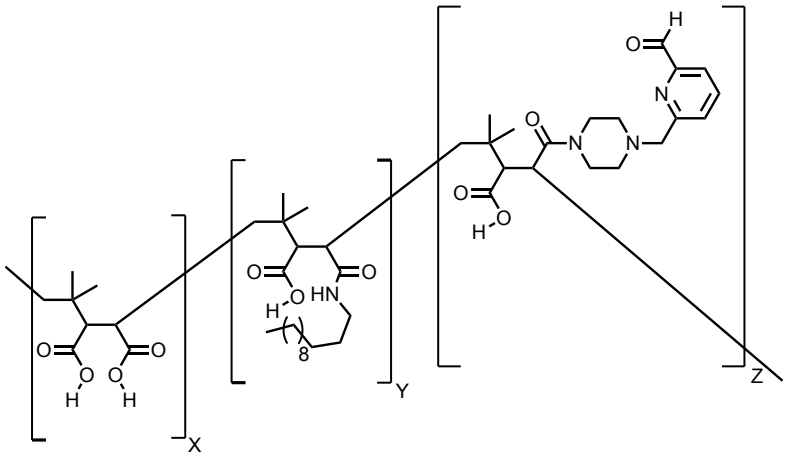
Internalization mechanism study. Cells were grown and seeded in an 8-well chambered coverglass (LabTek) setup as described above. Cells were pretreated with inhibitors (5.7 mM methyl-β-cyclodextrin, 10 μM chlorpromazine, 0.35 μM wortmannin, 50 μM dynasore, or 30 μM amiloride) at 37 °C for 30 min. Following this, was a 2 h of treatment (37 °C) with 10 μM protein with or without 10 μM polymers, as reported previously.[35-38, 40] Nuclear and cytoplasmic staining were performed using DAPI and ActinRed, as described above.

RNase A Activity Assay. The protocol used in these experiments was reported previously.[47] A 50 μL aliquot was taken from a 1 mg/mL solution of RNase A or its polymer conjugate and added to 50 μL of RNA from a 4 mg/mL 100 mM sodium acetate, pH 5.0 at 25 °C in a 96-well plate (Thermo Fisher). The absorbance at 300 nm was monitored over time by an Infinite 200 Pro plate reader (Tecan, Switzerland).

Cell Viability assays. The protocol was modified from a previously described procedure.[48-49] Cells were trypsinized and diluted to a density of 50,000 cells/mL. An aliquot of 200 μL of cell stock was placed in a well of a 96 well plate (Corning) to reach a density of 10,000 cells/well. The plate was incubated at 37 °C in a 5% CO_2 atmosphere for 18 h. Media were removed from the plate, and 100 μL portions of the appropriate sample stocks in media were added. The cells were incubated at 37 °C in a 5% CO_2 atmosphere for a determined time (2-18 h). The media containing the sample was removed from the well and 100 μL of 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyl-tetrazoliumbromide (MTT) in serum free media (0.5 mg/mL, Amresco) was added to each well. The samples were incubated for 4 h. An Infinite 200 Pro plate reader (Tecan) was used to measure the absorbance at 570 nm.

3.5. Supporting information

Table S3.1. Calculated elemental analysis of monomer variants.[a]



monomer	% C[a]	% N[a]
X	53.16	0
Y	70.11	4.30
Z	62.59	12.17

[a] ChemDraw Professional 17.1 software was used to calculate the percentage of C and N.

Table S3.2. Elemental analysis of polymer variants.

polymer	% C	% N
C	46.75	0.98
D	52.27	1.70
E	51.95	2.28
F	51.90	2.36
G	53.10	4.59
H	56.88	6.12

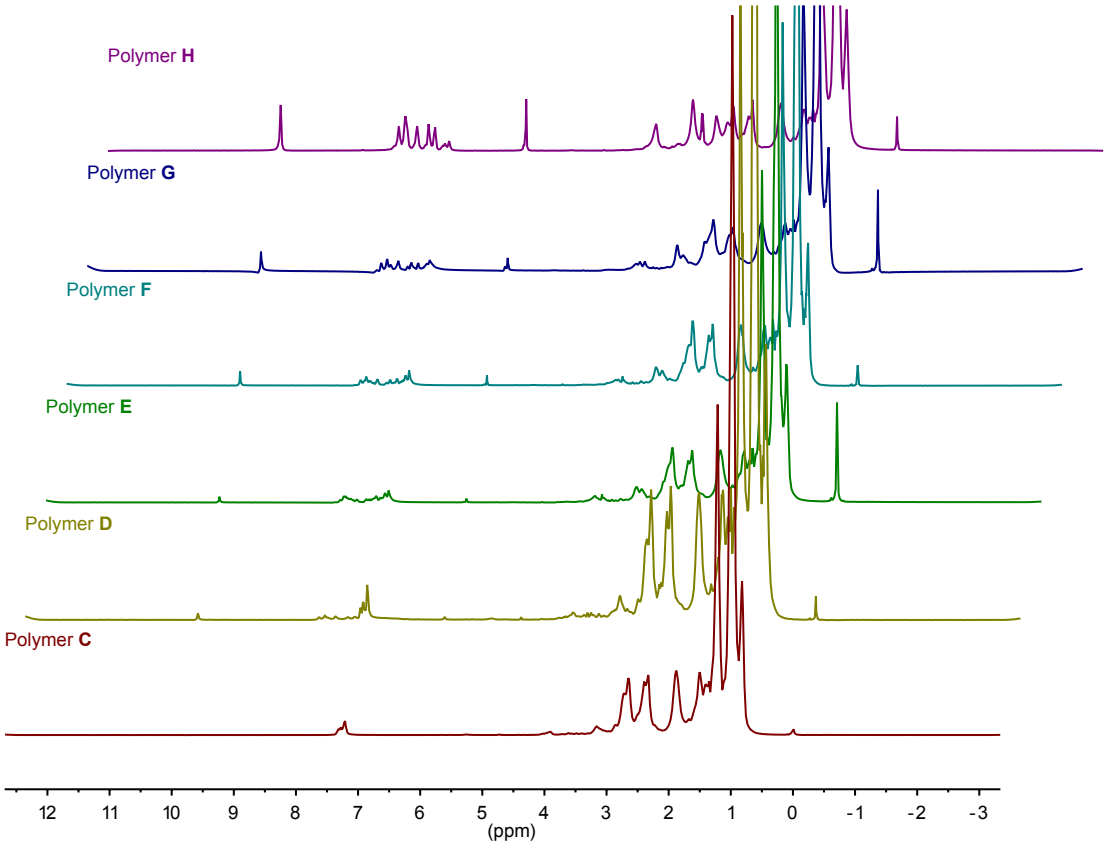


Figure S3.1. ^{1}H NMR in D_2O of polymers **C**-**H**. For clarity, the water signal has been suppressed.

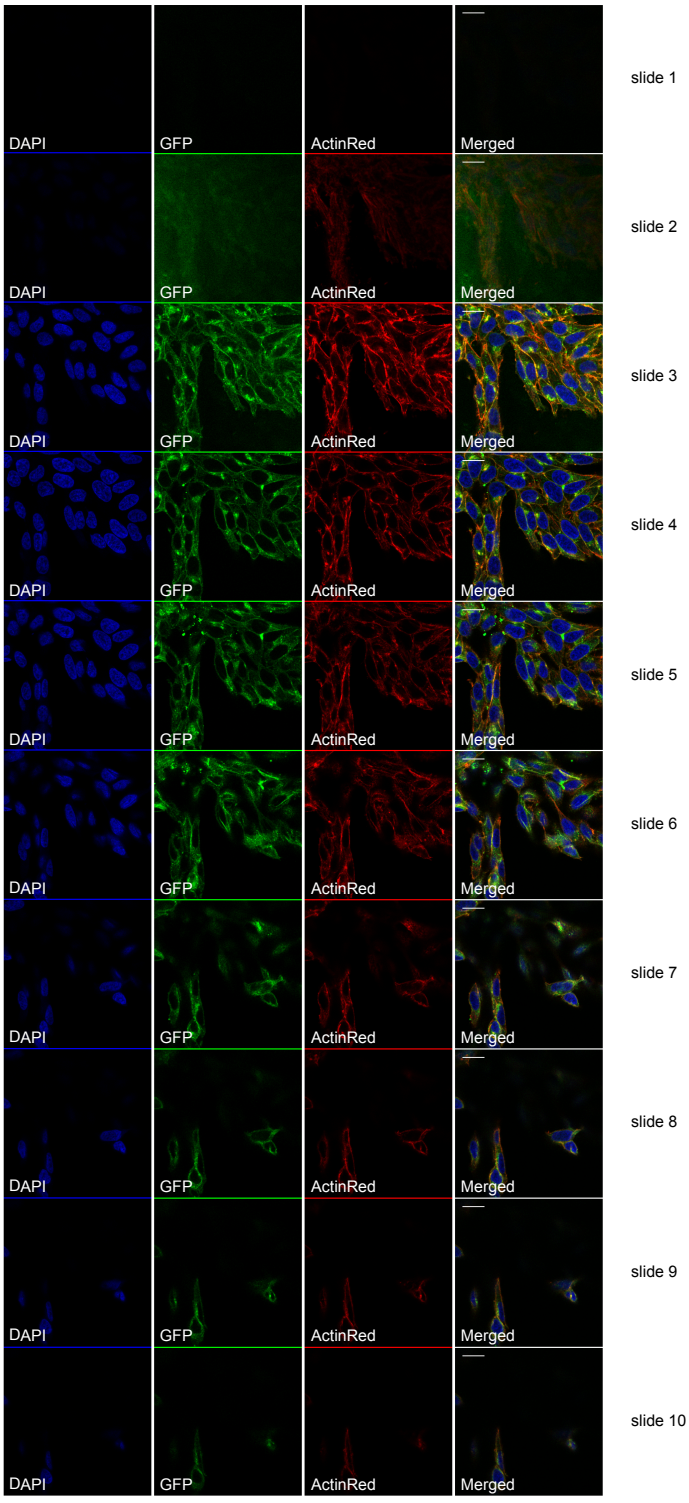


Figure S3.2. Z-stacking images of the conjugate of GFP to polymer **H** in HeLa cells. Nuclear and cytoplasmic stains were performed on fixed cells using DAPI and ActinRed, respectively. The scale bars represent 20 μm. Cells were treated with the conjugate of GFP to polymer **H** (10 μM) for 2 h at 37 ºC.

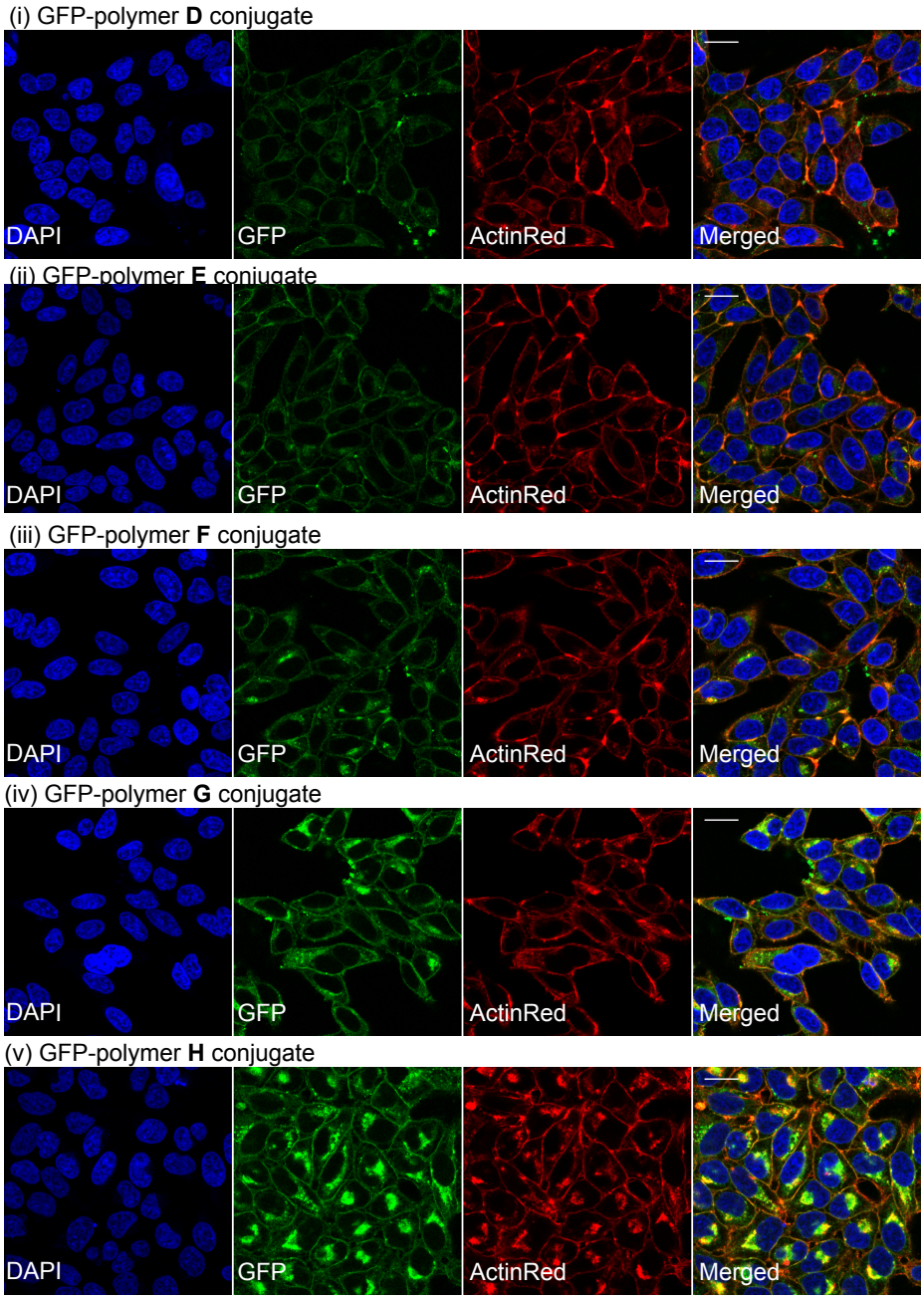


Figure S3.3. Confocal microscopy study of the conjugates of GFP to polymer **D**-**H** in HeLa cells. Nuclear and cytoplasmic stains were performed on fixed cells using DAPI and ActinRed, respectively. The scale bars represent 20 μm. Cells were treated with the conjugates of GFP to polymer **D**-**H** (10 μM) for 2 h at 37 ºC.

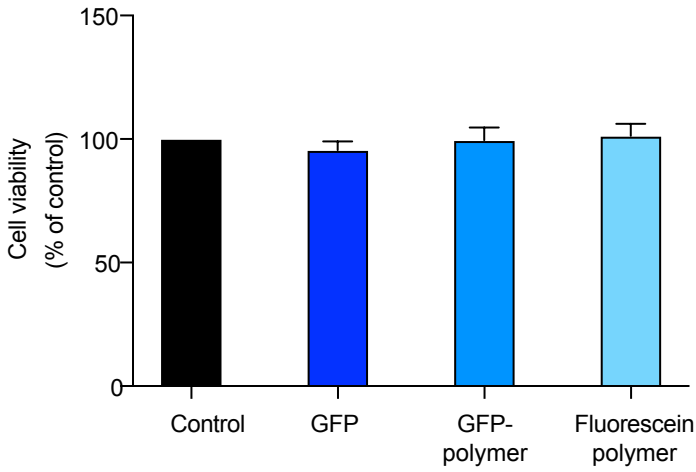


Figure S3.4. Cytotoxic effect of the conjugate of GFP to polymer **H**. HeLa cells were incubated with control (no treatment), GFP (10 μM), the conjugate of GFP to polymer **H** (10 μM), or fluorescein-labelled polymer (5 μM) for 2 h at 37 ºC. Viability was measured using an MTT assay, and is reported as the percentage of surviving cells compared to the untreated control. Error bars represent the standard error of three independent experiments.

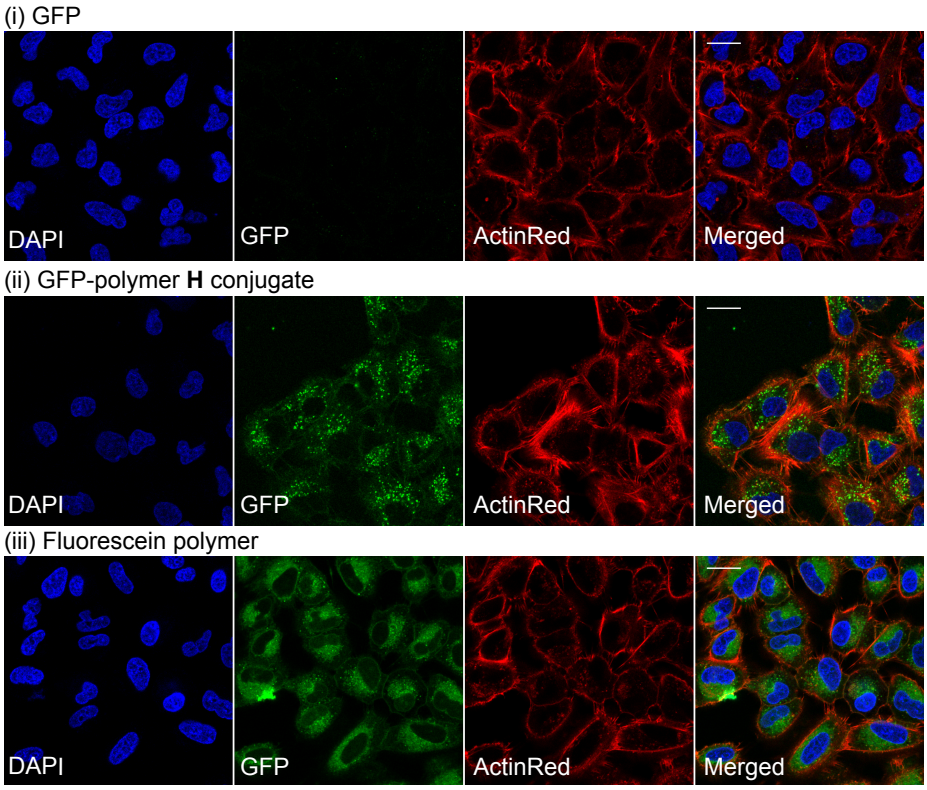


Figure S3.5. Confocal microscopy study of the conjugate of GFP to polymer **H** in A549 cells. Nuclear and cytoplasmic stains were performed on fixed cells using DAPI and ActinRed, respectively. The scale bars represent 20 μm. Cells were treated with GFP alone (10 μM), GFP conjugated to polymer **H** (10 μM), or fluorescein-labeled polymer (5 μM) for 2 h at 37 ºC.

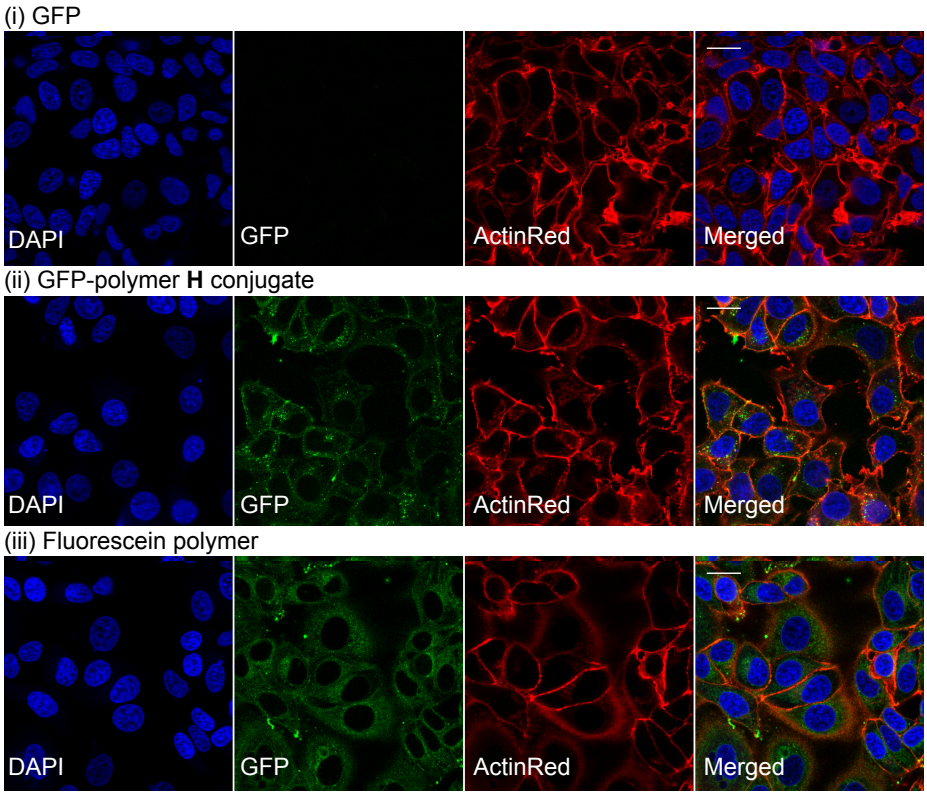


Figure S3.6. Confocal microscopy study of the conjugate of GFP to polymer **H** in MCF7 cells. Nuclear and cytoplasmic stains were performed on fixed cells using DAPI and ActinRed, respectively. The scale bars represent 20 μm. Cells were treated with GFP alone (10 μM), GFP conjugated to polymer **H** (10 μM), or fluorescein-labeled polymer (5 μM) for 2 h at 37 ºC.

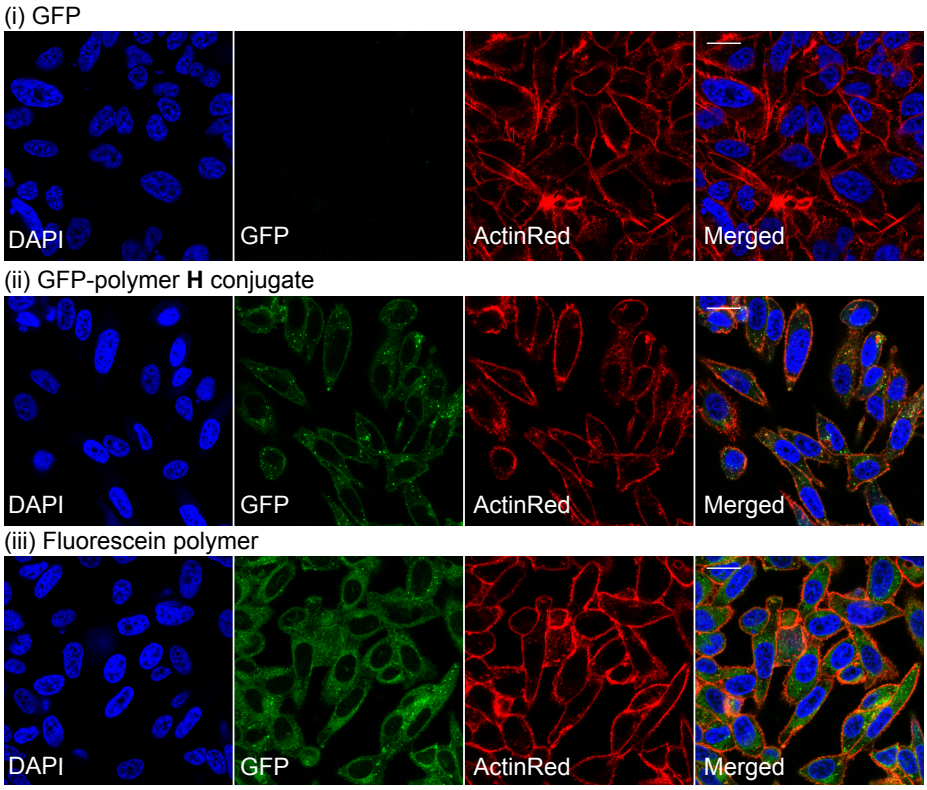


Figure S3.7. Confocal microscopy study of the conjugate of GFP to polymer **H** in PC-3 cells. Nuclear and cytoplasmic stains were performed on fixed cells using DAPI and ActinRed, respectively. The scale bars represent 20 μm. Cells were treated with GFP alone (10 μM), GFP conjugated to polymer **H** (10 μM), or fluorescein-labeled polymer (5 μM) for 2 h at 37 ºC.

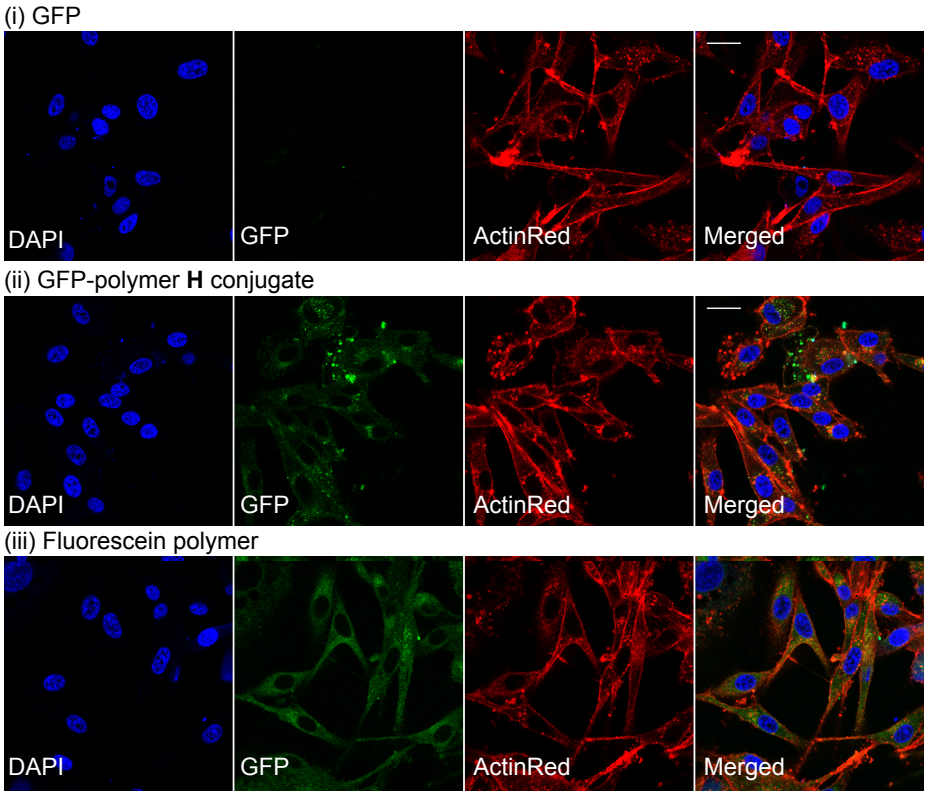


Figure S3.8. Confocal microscopy study of the conjugate of GFP to polymer **H** in U-87 MG cells. Nuclear and cytoplasmic stains were performed on fixed cells using DAPI and ActinRed, respectively. The scale bars represent 20 μm. Cells were treated with GFP alone (1 μM), GFP conjugated to polymer **H** (1 μM), or fluorescein-labeled polymer (0.5 μM) for 2 h at 37 ºC.

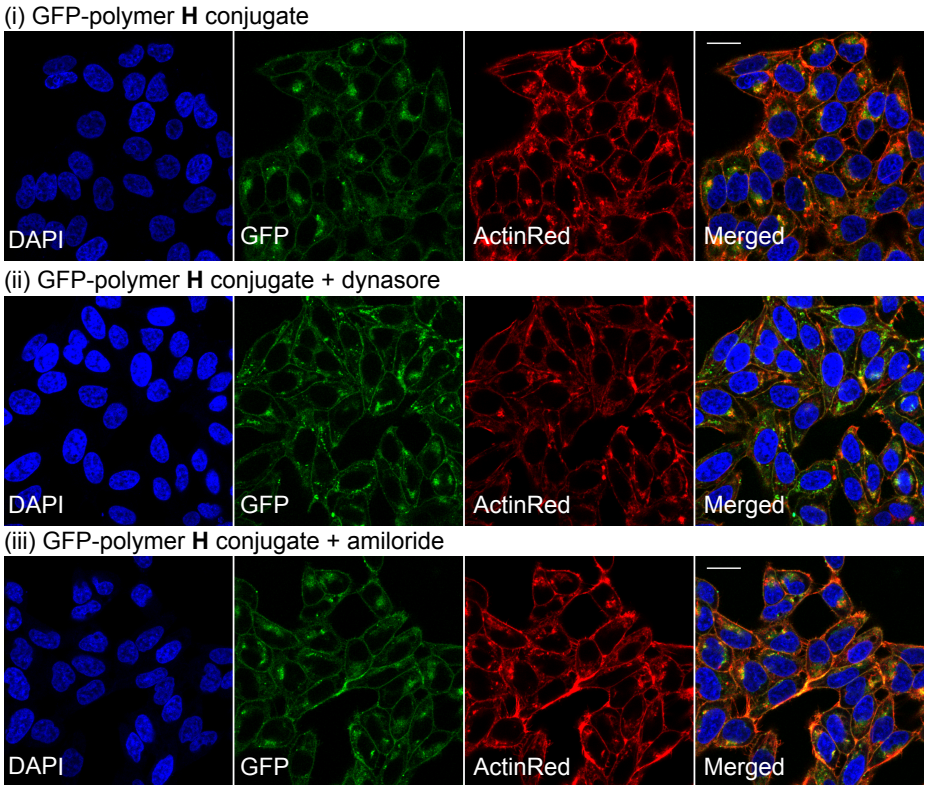


Figure S3.9. Probing uptake mechanism of the conjugate of GFP to polymer **H** in HeLa cells using confocal microscopy. HeLa cells were pretreated with corresponding inhibitors for 30 min prior to 2 h co-treatment of the conjugate of GFP to polymer **H** and inhibitors (50 μM dynasore or 30 μM amiloride). Nuclear and cytoplasmic stains were performed on fixed cells using DAPI and ActinRed, respectively.

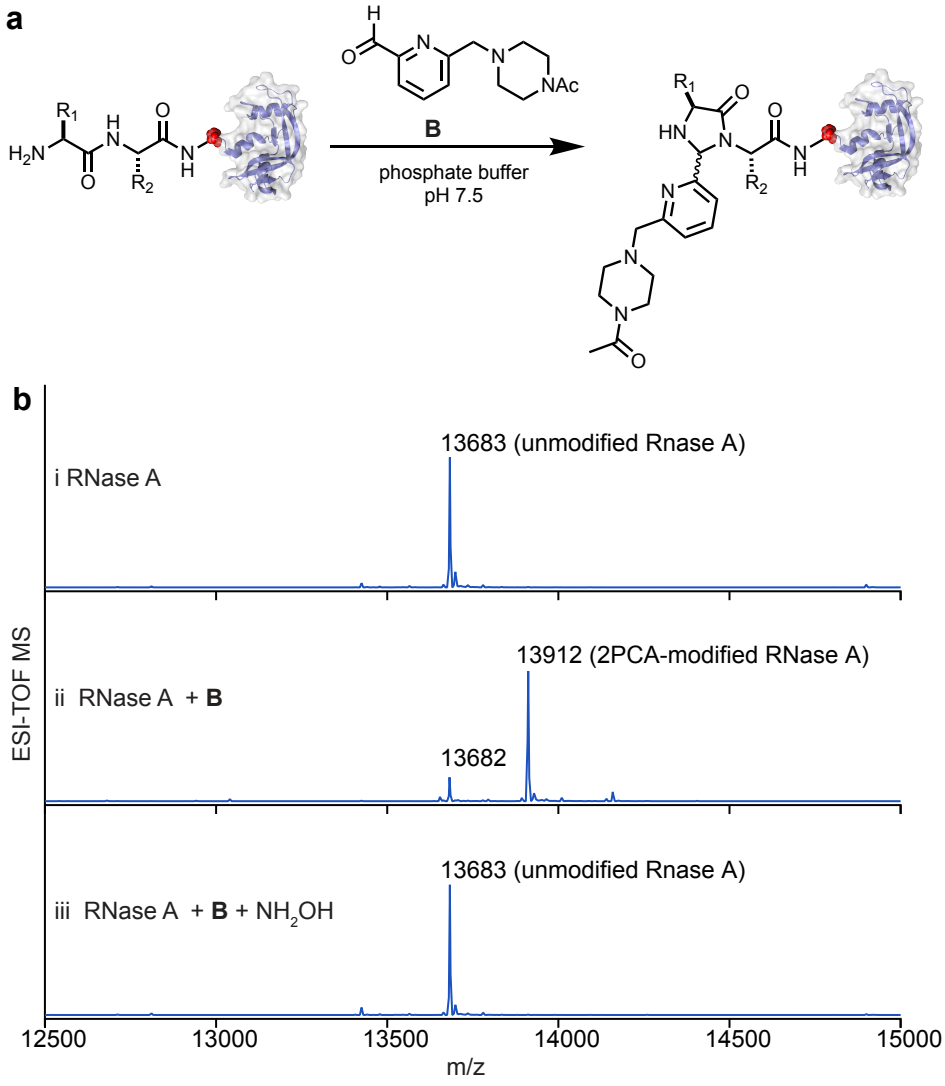


Figure S3.10. Coupling protein cargo to a 2PCA molecule. (a) 2PCA reacts with N-terminal residues to form imidazolidinones. (b) ESI-TOF MS analyses are shown for the RNase A before and after the reaction with 2PCA reagent **B**, or **B** in the presence of NH_2OH as a coupling inhibitor. The reactions were run with 10 mM **B** and 10 μM protein, with or without 20 mM NH_2OH in 50 mM phosphate pH 7.5 at RT for 18 h.

3.6. References

1. Leader, B.; Baca, Q. J.; Golan, D. E., Protein therapeutics: a summary and pharmacological classification. *Nat. Rev. Drug Discov.* **2008,** *7* (1), 21-39.
2. Bruno, B. J.; Miller, G. D.; Lim, C. S., Basics and recent advances in peptide and protein drug delivery. *Ther. Deliv.* **2013,** *4* (11), 1443-1467.
3. Craik, D. J.; Fairlie, D. P.; Liras, S.; Price, D., The Future of Peptide-based Drugs. *Chem. Biol. Drug Des.* **2013,** *81* (1), 136-147.
4. Golan, D. E.; Tashjian, A. H.; Armstrong, E. J., *Principles of pharmacology: the pathophysiologic basis of drug therapy*. Lippincott Williams & Wilkins: 2011.
5. Lindsay, K. L., Therapy of hepatitis C: overview. *Hepatology* **1997,** *26* (S3), 71S-77S.
6. Roth, D. A.; Kessler, C. M.; Pasi, K. J.; Rup, B.; Courter, S. G.; Tubridy, K. L., Human recombinant factor IX: safety and efficacy studies in hemophilia B patients previously treated with plasma-derived factor IX concentrates. *Blood* **2001,** *98* (13), 3600-3606.
7. Administration, U. S. F. a. D. Information on Cetuximab (marketed as Erbitux). (accessed 8/25/2018).
8. Goldman, S. C.; Holcenberg, J. S.; Finklestein, J. Z.; Hutchinson, R.; Kreissman, S.; Johnson, F. L.; Tou, C.; Harvey, E.; Morris, E.; Cairo, M. S., A randomized comparison between rasburicase and allopurinol in children with lymphoma or leukemia at high risk for tumor lysis: Presented in part at the American Society of Hematology Conference in Miami Beach, FL, December 1998. *Blood* **2001,** *97* (10), 2998-3003.
9. Hurwitz, H.; Fehrenbacher, L.; Novotny, W.; Cartwright, T.; Hainsworth, J.; Heim, W.; Berlin, J.; Baron, A.; Griffing, S.; Holmgren, E., Bevacizumab plus irinotecan, fluorouracil, and leucovorin for metastatic colorectal cancer. *New Engl. J. Med.* **2004,** *350* (23), 2335-2342.
10. Vogel, C. L.; Cobleigh, M. A.; Tripathy, D.; Gutheil, J. C.; Harris, L. N.; Fehrenbacher, L.; Slamon, D. J.; Murphy, M.; Novotny, W. F.; Burchmore, M., Efficacy and safety of trastuzumab as a single agent in first-line treatment of HER2-overexpressing metastatic breast cancer. *J. Clin. Oncol.* **2002,** *20* (3), 719-726.
11. Yang, N. J.; Hinner, M. J., Getting across the cell membrane: an overview for small molecules, peptides, and proteins. In *Site-Specific Protein Labeling*, Springer: 2015; pp 29-53.
12. Kou, L.; Sun, J.; Zhai, Y.; He, Z., The endocytosis and intracellular fate of nanomedicines: Implication for rational design. *Asian J. Pharm.* **2013,** *8* (1), 1-10.
13. Liu, X.; Zhang, P.; He, D.; Rödl, W.; Preiß, T.; Rädler, J. O.; Wagner, E.; Lächelt, U., pH-Reversible Cationic RNase A Conjugates for Enhanced Cellular Delivery and Tumor Cell Killing. *Biomacromolecules* **2015,** *17* (1), 173-182.
14. Murthy, N.; Thng, Y. X.; Schuck, S.; Xu, M. C.; Fréchet, J. M., A novel strategy for encapsulation and release of proteins: hydrogels and microgels with acid-labile acetal cross-linkers. *J. Am. Chem. Soc.* **2002,** *124* (42), 12398-12399.
15. Maier, K.; Martin, I.; Wagner, E., Sequence defined disulfide-linked shuttle for strongly enhanced intracellular protein delivery. *Mol. Pharm.* **2012,** *9* (12), 3560-3568.
16. Mout, R.; Ray, M.; Tay, T.; Sasaki, K.; Yesilbag Tonga, G.; Rotello, V. M., General Strategy for Direct Cytosolic Protein Delivery via Protein-Nanoparticle Co-engineering. *ACS nano* **2017,** *11* (6), 6416-6421.
17. Mout, R.; Yesilbag Tonga, G.; Wang, L.-S. S.; Ray, M.; Roy, T.; Rotello, V. M., Programmed Self-Assembly of Hierarchical Nanostructures through Protein-Nanoparticle Coengineering. *ACS nano* **2017,** *11* (4), 3456-3462.

18. Nischan, N.; Herce, H. D.; Natale, F.; Bohlke, N.; Budisa, N.; Cardoso, M. C.; Hackenberger, C. P., Covalent attachment of cyclic TAT peptides to GFP results in protein delivery into live cells with immediate bioavailability. *Angew. Chem. Int. Ed.* **2015,** *54* (6), 1950-1953.
19. Akishiba, M.; Takeuchi, T.; Kawaguchi, Y.; Sakamoto, K.; Yu, H.-H.; Nakase, I.; Takatani-Nakase, T.; Madani, F.; Gräslund, A.; Futaki, S., Cytosolic antibody delivery by lipid-sensitive endosomolytic peptide. *Nat. Chem.* **2017,** *9* (8), 751.
20. Guo, S.; Huang, L., Nanoparticles escaping RES and endosome: challenges for siRNA delivery for cancer therapy. *J. Nanomater.* **2011,** *2011*, 11.
21. Wu, Y.; Smith, A. E.; Reineke, T. M., Lipophilic Polycation Vehicles Display High Plasmid DNA Delivery to Multiple Cell Types. *Bioconjug. Chem.* **2017,** *28* (8), 2035-2040.
22. Convertine, A. J.; Diab, C.; Prieve, M.; Paschal, A.; Hoffman, A.; Johnson, P.; Stayton, P. S., pH-responsive polymeric micelle carriers for siRNA drugs. *Biomacromolecules* **2010,** *11* (11), 2904-2911.
23. Kyriakides, T. R.; Cheung, C. Y.; Murthy, N.; Bornstein, P.; Stayton, P. S.; Hoffman, A. S., pH-sensitive polymers that enhance intracellular drug delivery in vivo. *J. Controlled Release* **2002,** *78* (1-3), 295-303.
24. Jones, R. A.; Cheung, C. Y.; Fiona, E.; STAYTON, P. S.; HOFFMAN, A. S.; WILSON, M. R., Poly (2-alkylacrylic acid) polymers deliver molecules to the cytosol by pH-sensitive disruption of endosomal vesicles. *Biochem. J.* **2003,** *372* (1), 65-75.
25. Cheung, C. Y.; Murthy, N.; Stayton, P. S.; Hoffman, A. S., A pH-sensitive polymer that enhances cationic lipid-mediated gene transfer. *Bioconjug. Chem.* **2001,** *12* (6), 906-910.
26. Jańczewski, D.; Tomczak, N.; Han, M.-Y.; Vancso, G. J., Synthesis of functionalized amphiphilic polymers for coating quantum dots. *Nat. Protoc.* **2011,** *6* (10), 1546.
27. Lin, C. A. J.; Sperling, R. A.; Li, J. K.; Yang, T. Y.; Li, P. Y.; Zanella, M.; Chang, W. H.; Parak, W. J., Design of an amphiphilic polymer for nanoparticle coating and functionalization. *Small* **2008,** *4* (3), 334-341.
28. Jańczewski, D.; Tomczak, N.; Khin, Y. W.; Han, M.-Y.; Vancso, G. J., Designer multi-functional comb-polymers for surface engineering of quantum dots on the nanoscale. *Eur. Polym. J.* **2009,** *45* (1), 3-9.
29. Jańczewski, D.; Tomczak, N.; Han, M.-Y.; Vancso, G. J., Introduction of quantum dots into PNIPAM microspheres by precipitation polymerization above LCST. *Eur. Polym. J.* **2009,** *45* (7), 1912-1917.
30. Jańczewski, D.; Tomczak, N.; Song, J.; Long, H.; Han, M.-Y.; Vancso, G. J., Fabrication and responsive behaviour of Quantum Dot/PNIPAM micropatterns obtained by template copolymerization in water. *J. Mater. Chem.* **2011,** *21* (18), 6487-6490.
31. Tomczak, N.; Jańczewski, D.; Tagit, O.; Han, M. Y.; Vancso, G. J., Surface engineering of quantum dots with designer ligands. *Surface Design: Applications in Bioscience and Nanotechnology* **2009**, 341-361.
32. Tagit, O.; Jańczewski, D.; Tomczak, N.; Han, M. Y.; Herek, J. L.; Vancso, G. J., Nanostructured thermoresponsive quantum dot/PNIPAM assemblies. *Eur. Polym. J.* **2010,** *46* (7), 1397-1403.
33. MacDonald, J. I.; Munch, H. K.; Moore, T.; Francis, M. B., One-step site-specific modification of native proteins with 2-pyridinecarboxyaldehydes. *Nat. Chem. Biol.* **2015,** *11* (5), 326.
34. Lee, J. P.; Kassianidou, E.; MacDonald, J. I.; Francis, M. B.; Kumar, S., N-terminal specific conjugation of extracellular matrix proteins to 2-pyridinecarboxaldehyde functionalized polyacrylamide hydrogels. *Biomaterials* **2016,** *102*, 268-276.

35. Saha, K.; Kim, S. T.; Yan, B.; Miranda, O. R.; Alfonso, F. S.; Shlosman, D.; Rotello, V. M., Surface functionality of nanoparticles determines cellular uptake mechanisms in mammalian cells. *Small* **2013,** *9* (2), 300-305.
36. Gratton, S. E.; Ropp, P. A.; Pohlhaus, P. D.; Luft, J. C.; Madden, V. J.; Napier, M. E.; DeSimone, J. M., The effect of particle design on cellular internalization pathways. *Proc. Natl. Acad. Sci. USA* **2008,** *105* (33), 11613-11618.
37. Zhang, L. W.; Monteiro-Riviere, N. A., Mechanisms of quantum dot nanoparticle cellular uptake. *Toxicol. Sci.* **2009,** *110* (1), 138-155.
38. Plummer, E. M.; Manchester, M., Endocytic Uptake Pathways Utilized by CPMV Nanoparticles. *Mol. Pharm.* **2013,** *10* (1), 26-32.
39. Bandmann, V.; Müller, J. D.; Köhler, T.; Homann, U., Uptake of fluorescent nano beads into BY2-cells involves clathrin-dependent and clathrin-independent endocytosis. *FEBS Lett.* **2012,** *586* (20), 3626-3632.
40. Yang, S.-T.; Kreutzberger, A. J.; Lee, J.; Kiessling, V.; Tamm, L. K., The role of cholesterol in membrane fusion. *Chem. Phys. Lipids* **2016,** *199*, 136-143.
41. Juliano, R. L., The delivery of therapeutic oligonucleotides. *Nucleic Acids Res.* **2016,** *44* (14), 6518-6548.
42. Leland, P. A.; Raines, R. T., Cancer chemotherapy–ribonucleases to the rescue. *Chem. Biol.* **2001,** *8* (5), 405-413.
43. Raines, R. T., Ribonuclease a. *Chem. Rev.* **1998,** *98* (3), 1045-1066.
44. Shlyakhovenko, V. A., Ribonucleases in tumor growth. *Exp. Oncol.* **2009,** *31*, 127-133.
45. Kunitz, M., A spectrophotometric method for the measurement of ribonuclease activity.
46. Laemmli, U. K., Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **1970,** *227* (5259), 680.
47. Kunitz, M., A spectrophotometric method for the measurement of ribonuclease activity. *J. biol. Chem* **1946,** *164*, 563-568.
48. Suriyo, T.; Tachachartvanich, P.; Visitnonthachai, D.; Watcharasit, P.; Satayavivad, J., Chlorpyrifos promotes colorectal adenocarcinoma H508 cell growth through the activation of EGFR/ERK1/2 signaling pathway but not cholinergic pathway. *Toxicology* **2015,** *338*, 117-129.
49. Tachachartvanich, P.; Sangsuwan, R.; Ruiz, H. S.; Sanchez, S. S.; Durkin, K. A.; Zhang, L.; Smith, M. T., Assessment of the Endocrine-Disrupting Effects of Trichloroethylene and Its Metabolites Using in *Vitro* and in *Silico* Approaches. *Environ. Sci. Technol.* **2018,** *52* (3), 1542-1550.

Chapter 4

Conformational change study of PEGylated intrinsically disordered proteins

Abstract

Proteins are of great interest in the field of drug delivery. Using a protein-based carrier is advantageous because these materials are nontoxic and are easily degraded by the body. Despite many advances in developing protein-based carriers, using these agents efficiently requires fundamental knowledge of biophysical properties of materials, which has an effect on *in vivo* biodistribution. Intrinsically disordered protein (IDP) is a flexible linear protein that contains numerous lysine residues, which are the sites of protein modification. The role of IDP as an delivery scaffold and a structural change through chemical modifications were not thoroughly explored. We decorated the external surface of IDP with polyethylene glycol (PEG) chains in order to vary the size and charge of the protein as well as prolong its circulation. The physical properties of IDPs conjugated with PEG chains were investigated. The size and density of the PEG chains had little to no effect on the conformation of IDPs.

The studies described in this chapter were done in collaboration with Alexander C. Hoepker and Jessica Lee.

4.1. Introduction

Protein-based nanomaterials have gained substantial popularity, especially in the fields of materials science and chemical biology.[1-3] They provide many advantages, including (1) proteins are biocompatible and biodegradable, (2) proteins have the possibility of less opsonization by the reticuloendothelial system through an aqueous steric barrier, (3) proteins can be easily prepared as monodisperse protein carriers, and (4) synthetic molecules can be introduced using site-specific modification. A key challenge in the success of drug delivery is the escape of nanoparticles from the blood stream through epithelial barriers.[4] In recent years, studies have been focused on the influence of physicochemical properties on the behavior of a particle in living cells. Gratton *et al.* compared spherical and rod-shaped PEG-based hydrogels[5] and showed that nonspherical particles were transported through different mechanisms than spherical particles. As an example for the effect of surface properties on biodistribution, polyethylene glycol (PEG) is frequently used to shield the charge of nanoparticles and reduce their interactions with serum proteins and the reticuloendothelial system.[6-7]

Intrinsically disordered proteins (IDPs) do not have well-defined three-dimensional structures and are structurally flexible due to the abundance of disorder-promoting amino acids such as Gln, Ser, Glu, Lys, and Pro.[8-9] Most IDPs lack a hydrophobic core and are highly charged. Due to their high flexibility, the protein crystallography cannot be used to characterize them, resulting in limited knowledge of their structural conformations. However, we hypothesize that the high flexibility of IDP can be employed to modulate the conformation of IDP through chemical modifications. Various modification degrees might lead to an array of scaffolds with morphologies that range from a disordered strand to a flexible rod. This would allow us to fine-tune the physical properties of IDP. PEG was chosen due to its advantages mentioned earlier. We conjugated lysines of an IDP with various numbers of PEG chains of different sizes and characterized the biophysical properties of these modified IDPs for future use in drug delivery system. Based on a fully disordered IDP from a neurofilament heavy side-arm (NFH-SA), we have designed a human-derived IDP sequence as a 161 amino acid-long monodisperse scaffold (see details below).

Full length IDP construct:

SEFTSMSTHIKVKSEEKIKVVEKSEKETVIVEEQTEEIQVTEEVTEEEDKEAQGEEEEEAE
EGGEEAATTSPPAEEAASPEKETKSPVKEEAKSPAEAKSPAEAKSPAEAKSPAEVKSPAV
AKSPAEVKSPAEVKSPAEAKSPAEAKSPAEVKSPATVKSPGEAKSPAEAKSPAEVKSPVE
AKSPAEAKSPASVKSPGEAKSPAEAKSPAEVKSPATVKSPVEAKSPAEVKSPVTVKSPAE
AKSPVEVKSPASVKSPSEAKSPAGAKSPAEAKSPVVAKSPAEAKSPAEAKPPAEAKSPAE
AKSPAEAKSPAEAKSPAEAKSPVEVKSPEKAKSPVKEGAKSLAEAKSPEKAKSPVKEEIK
PPAEVKSPEKAKSPMKEEAKSPEKAKTLDVKSPEAKTPAKEEAKRPADIRSPEQVKSPA
KEEAKSPEKEETRTEKVAPKKEEVKSPVEEVKAKEPPKKVEEEKTPATPKTEVKESKKD
EAPKEAQKPKAEEKEPLTEKPKDSPGEAKKEEAKEKKAAAPEEETPAKLGVKEEAKPKE
KAEDAKAKEPSKPSEKEKPKKEEVPAAPEKKDTKEEKTTESKKPEEKPKMEAKAKEED
KGLPQEPSKPKTEKAEKSSSTDQKDSQPSEKAPEDKLYWW

Truncated IDP construct:

ASRGSPSAEAKSPAEAKSPAEVKSPAVAKSPAEVKSPAEVKSPAEAKSPAEAKSPAEVKS
PATVKSPGEAKSPAEAKSPAEVKSPVEAKS PAEAKSPASV KSPGEAKSPAEAKSPAEVK

SPATVKSPVEAKSPAEVKSPVTVKSPAEAKSPVEVKSPFSCA

4.2. Results and discussion

4.2.1. Truncated IDP construct design, purification, and characterization

This sequence has a calculated theoretical length of 50 nm when stretched linearly. A unique cysteine was introduced near the C-terminus as a site-specific conjugation handle. In general, IDP expression has low yield due to degradation in live cells.[10] A 6xHistidine-tag (brown) along with maltose-binding protein (blue) and a 10xAsparagine (N10) linker (white) were incorporated in the IDP construct for purification purposes and for increased expression levels,[11-13] respectively (Figure 4.1).

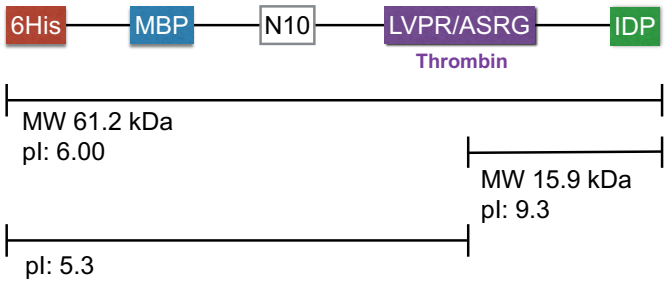


Figure 4.1. Truncated IDP expression construct containing a 6xHistidine-tag (6His, brown), maltose-binding protein (MBP, blue), a 10xAsparagine (N10, white) linker, and an alternative thrombin cleavage site (purple).

A thrombin cleavage site was introduced to cleave the fusion tags from IDP (Figure 4.1). The thrombin cleavage site has the sequence Leu-Val-Pro-Arg-Gly-Ser, and the cleavage occurs between the arginine and the glycine residues. However, glycine is poorly applicable for N-terminal modifications developed in the Francis lab. Therefore, an alternative thrombin cleavage site was alternated to Leu-Val-Pro-Arg-Ala-Ser, which will result in an N-terminal alanine after thrombin cleavage.[14] The cleaved target IDP has a molecular weight of 15.9 kDa and a theoretical pI of 9.3, while the uncleaved protein (MBP-IDP) has a molecular weight of 61.5 kDa and a theoretical pI of 6.0 (Figure 4.1). The details of cloning and expression are described in the materials and methods section. Briefly, the lysate obtained from the overexpression of IDP in *E. coli* was subjected to Ni-NTA purification to isolate the 6His-MBP-IDP, which was then analyzed by LC-MS (Figure 4.2). The protein sample was treated with thrombin to cleave the monomeric IDP from 6His-MBP. The monomeric IDP was then labelled with Alexa Flour 488 (A488) maleimide prior to purification with ion-exchange chromatography since the IDP does not have UV reactive amino acids. After ion-exchange chromatography, the monomeric IDP conjugated with A488 was obtained and the mass analyses were confirmed (Figure 4.2)

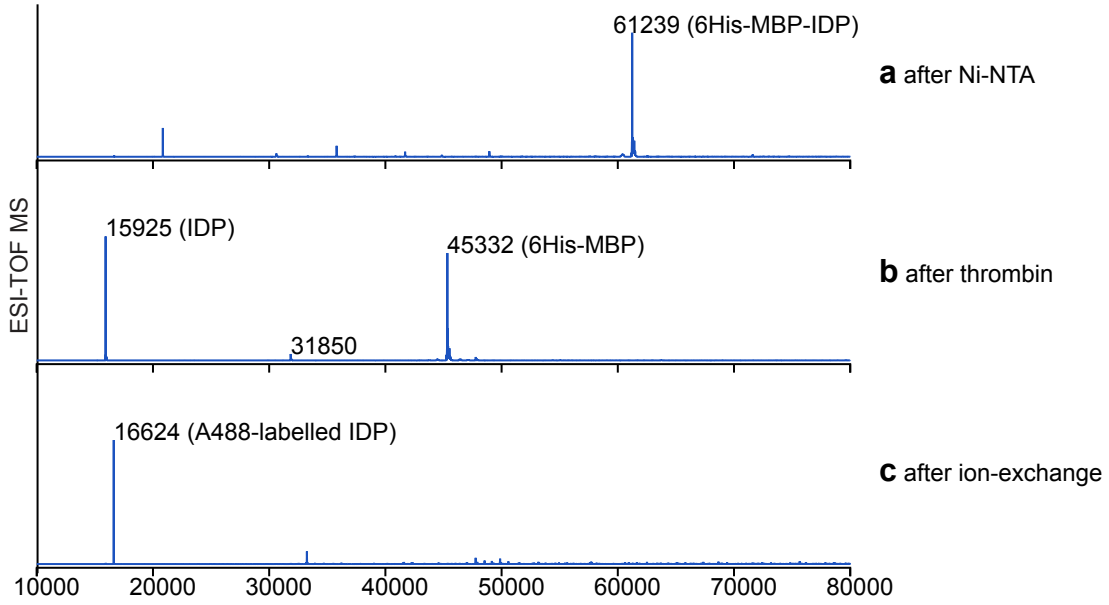


Figure 4.2. ESI-TOF MS spectra of IDP samples. (a) MS spectra of Ni-NTA-purified 6His-MBP-IDP (b) MS spectra of the IDP sample after the treatment with thrombin (c) MS spectra of ion exchange-purified IDP samples conjugated with Alexa Flour 488 (A488).

4.2.2. Synthesis of PEGylated IDP

IDP has 25 lysine residues, which can be modified using NHS chemistry. We generated a library of IDP with four different amine-reactive PEGs—a monodisperse 0.5 kDa PEG NHS-ester, a 1.2 kDa PEG NHS-ester, and a 4 kDa branched PEG NHS-ester. The IDP labelled with A488 was conjugated with various concentrations of PEGylating agents, giving rise to a library of IDP with different extents of PEG addition as shown in mass analyses (Figure 4.3). The average number of PEG chains added, and the effective protein charge were calculated and are listed in Table 4.1. As expected, the degree of PEG modification increased with increasing concentrations of monodisperse PEGylation reagents. As for the length of PEG reagents, fewer PEG chains were attached to IDP as the size of the PEG chain increased because of steric bulk. For the shortest length, 0.5 kDa PEG, up to 26 PEG chain additions could be observed as confirmed by mass spectrometry (Figure 4.3), indicating that our library spans a wide extent of PEGylation for a future biophysical study of how PEGylation affects IDP conformation.

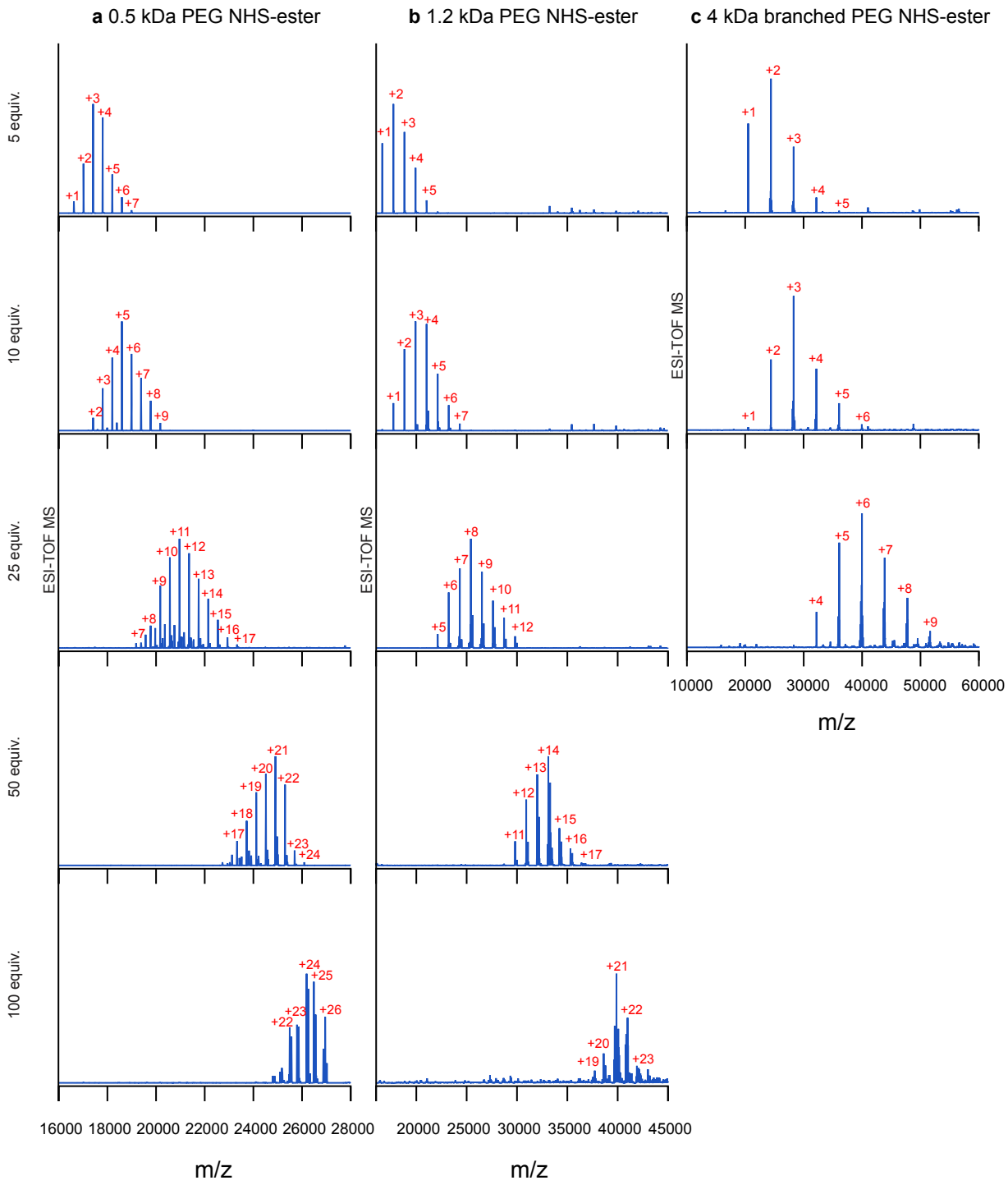


Figure 4.3. ESI-TOF MS spectra of IDP samples modified with various equivalents (with respect to IDP) of monodisperse 0.5 kDa PEG NHS-ester, 1.2 kDa PEG NHS-ester, and 4 kDa branched PEG NHS-ester. The numbers of PEG chain modifications are indicated in red. The MS spectra of IDP modified with 50 and 100 equivalents of 4 kDa branched PEG NHS-ester could not be deconvoluted.

Table 4.1. The average number of conjugated PEG chains and the effective protein charge (Z_{eff}) of IDP conjugated with various equivalents of monodisperse 0.5 kDa PEG NHS-ester, 1.2 kDa PEG NHS-ester, and 4 kDa branched PEG NHS-ester. The average number of PEG chains added was calculated using the integration of modified peaks in mass analyses (Figure 4.3). The Z_{eff} was calculated by subtracting the average number of conjugated PEG chains from the theoretical protein charge at pH 7.5, which is 5.6.

0.5 kDa PEG	Average number of conjugated PEG	Z_{eff}
5 equiv.	2.5	3.1
10 equiv.	5.3	0.3
25 equiv.	11.7	-6.1
50 equiv.	21.4	-15.8
100 equiv.	25.2	-20.4
1.2 kDa PEG	Average number of conjugated PEG	Z_{eff}
5 equiv.	1.54	4.1
10 equiv.	3.46	2.1
25 equiv.	8.2	-2.6
50 equiv.	15.4	-9.8
100 equiv.	21.3	-15.7
4 kDa branched PEG	Average number of conjugated PEG	Z_{eff}
5 equiv.	1.0	4.6
10 equiv.	2.1	3.5
25 equiv.	6.4	-0.8

4.2.3. Fluorescence anisotropy of PEGylated IDPs

Fluorescence anisotropy was used to study protein conformational dynamics. The lower anisotropy value, the higher the tumbling rate the protein has.[15]

4.2.3.1. The effect of salt concentrations and PEG lengths on the anisotropy of PEGylated IDPs

As shown in Figure 4.4, the degree of PEG modification slightly increased the tumbling rates of PEGylated IDPs. The length of PEG did not affect the anisotropy values. This suggests that the number of lysines modified with PEG influenced the tumbling rate more than the length of PEG. Increasing the concentrations of NaCl slightly increased the tumbling rates of PEGylated IDPs with a lower degree of modification (5, 10, equiv.), while the samples with higher degree of modification (50 and 100 equiv.) tumbled slower as the concentration of NaCl increased. The rate of tumbling did not change in the samples with 25 equiv. PEG reagents.

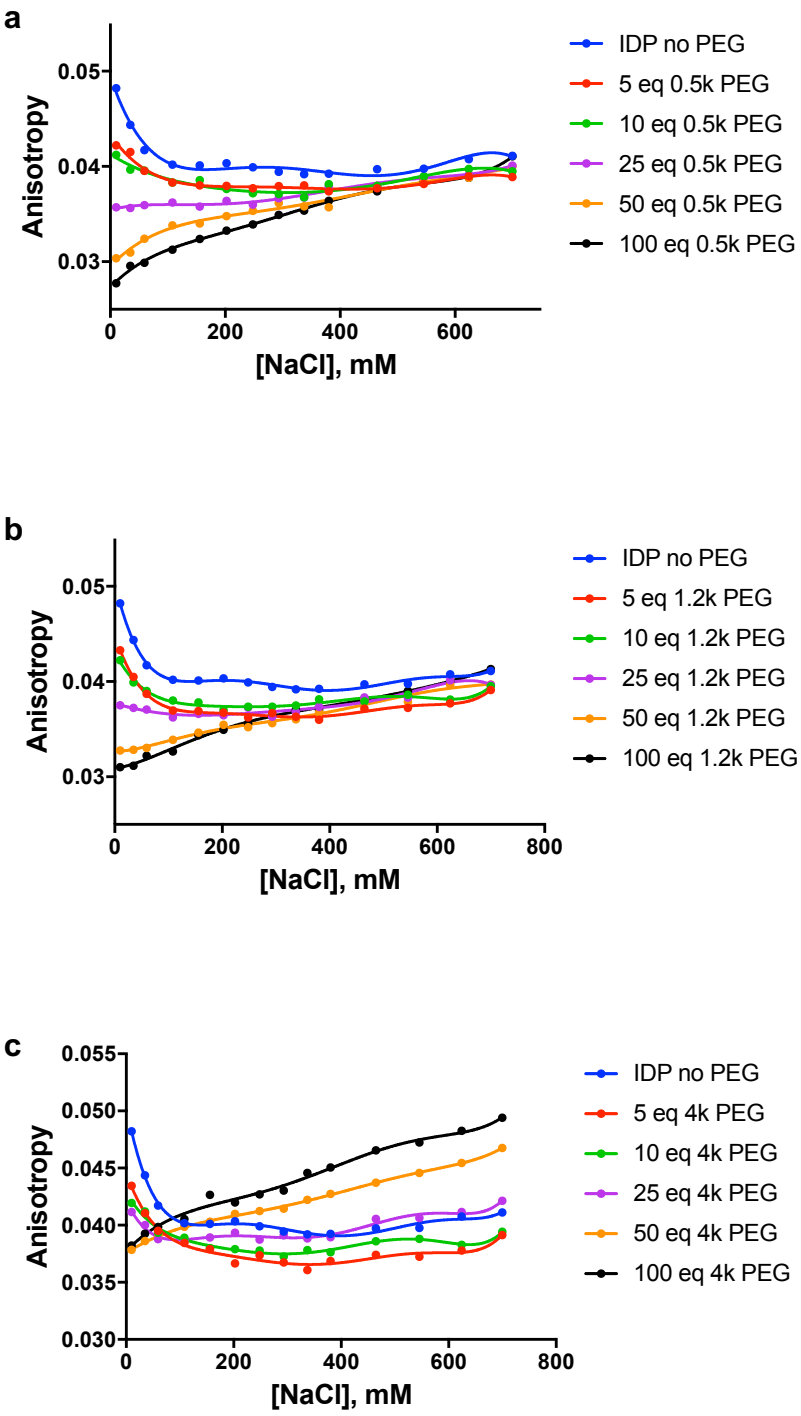


Figure 4.4. Anisotropy variations of PEGylated IDPs on treatment with the incremental concentration of NaCl. IDP samples modified with various equivalents (with respect to IDP) of monodisperse 0.5 kDa PEG NHS-ester (a), 1.2 kDa PEG NHS-ester (b), and 4 kDa branched PEG NHS-ester (c).

4.2.3.2. The effect of charges and ionic strength of buffers on the anisotropy of PEGylated IDPs

To investigate whether the charge of IDP has an influence on the tumbling rates, we alkylated all the lysine residues with NHS-acetate. The LC-MS confirmed that all lysines and the N-terminus of IDPs were capped (Figure 4.5). As shown in Figure 4.6, capped-lysine samples had a similar behavior as PEGylated IDPs whose all lysines were conjugated. The result indicated that charges of lysines had a small effect on the tumbling rates. In addition, the buffer strength did not affect the tumbling rate (Figure 4.6)

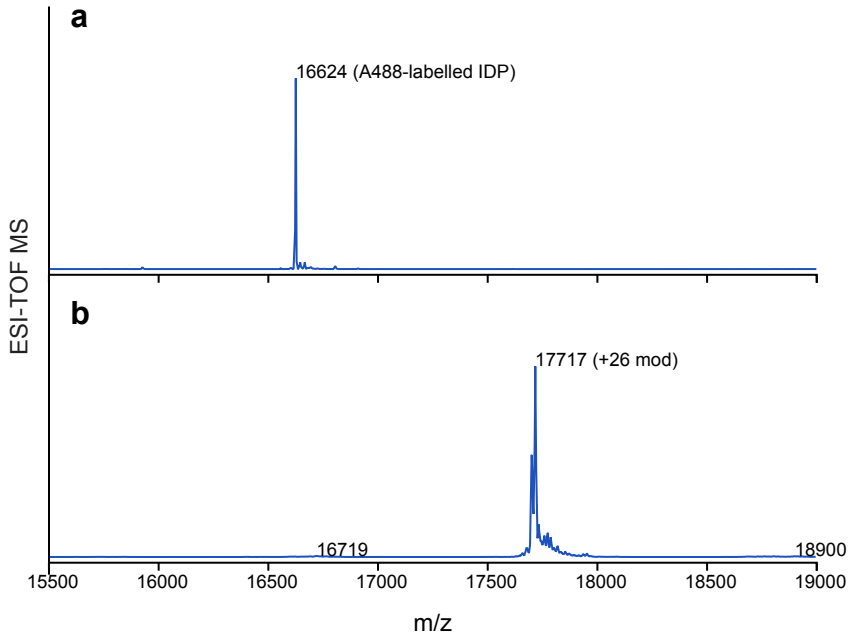


Figure 4.5. ESI-TOF MS spectra of IDP samples modified with NHS-acetate. (a) MS spectra of IDP samples labelled with A488. (b) MS spectra of IDP samples after the modification with NHS-acetate. The reaction was run with100 μM IDP and 10 mM of NHS-acetate in 10 mM phosphate buffer, pH 7.5.

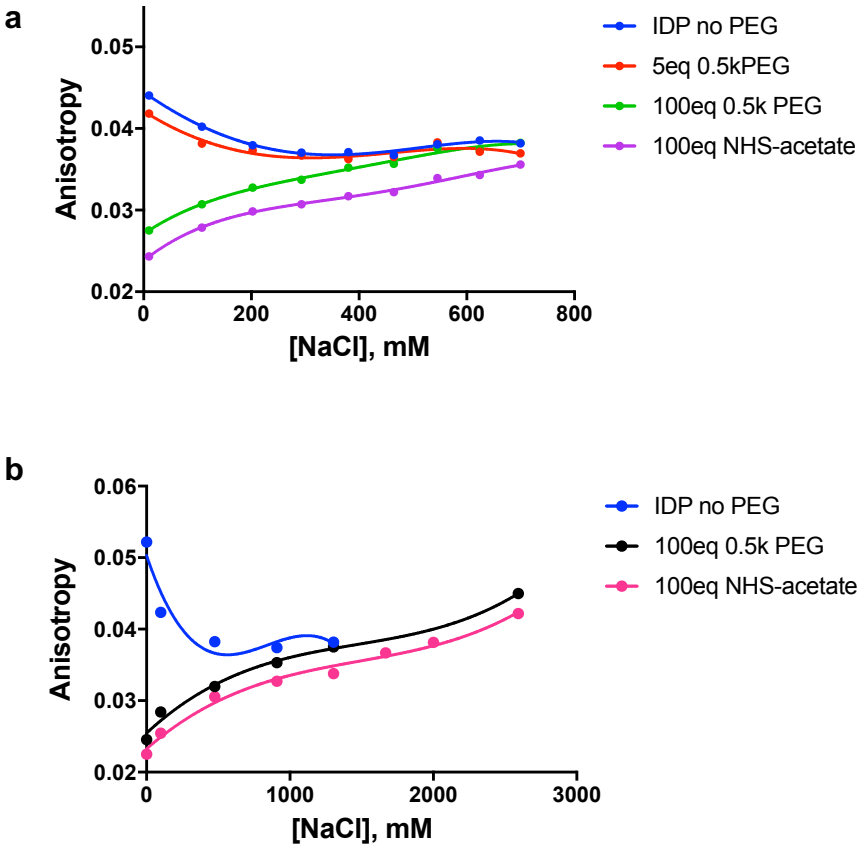


Figure 4.6. Anisotropy variations of IDP samples on treatment with incremental concentrations of NaCl. The buffers used were 10 mM phosphate, pH 7.5 (a) and 0.5 mM phosphate buffer, pH 7.5 (b).

4.2.3.3. The effect of temperature, pH, and glucose on the anisotropy of PEGylated IDPs

Increasing the temperature increased the tumbling rates of all IDP samples (Figure 4.7a). This can be explained by the increased kinetic energy from heat. Changes in pH and glucose concentration had a small effect on the tumbling rates (Figure 4.7b-c)

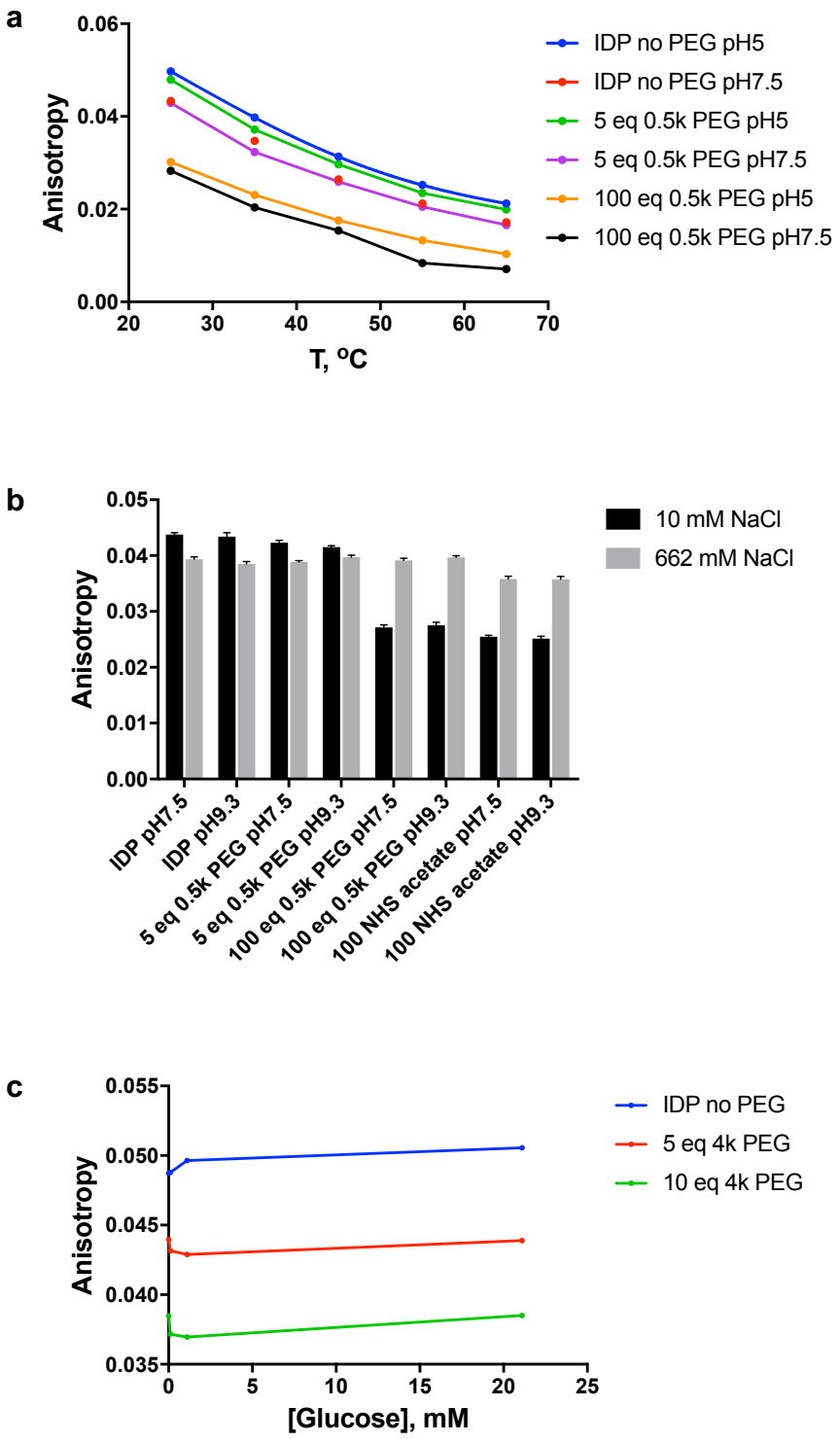


Figure 4.7. Anisotropy variations of IDP samples on different treatment conditions. (a) The effect of temperature on the anisotropy. (b) The effect of pH on the anisotropy. (c) The effect of glucose on the anisotropy.

4.2.4. Fluorescence correlation spectroscopy study of PEGylated IDPs

The fluorescence correlation spectroscopy was investigated and diffusion coefficients and hydrodynamic radii of PEGylated IDPs were calculated (Figure 4.8). The effects of salt concentrations and PEG modifications on the diffusion coefficients corresponding to hydrodynamic radii of PEGylated IDPs were small.

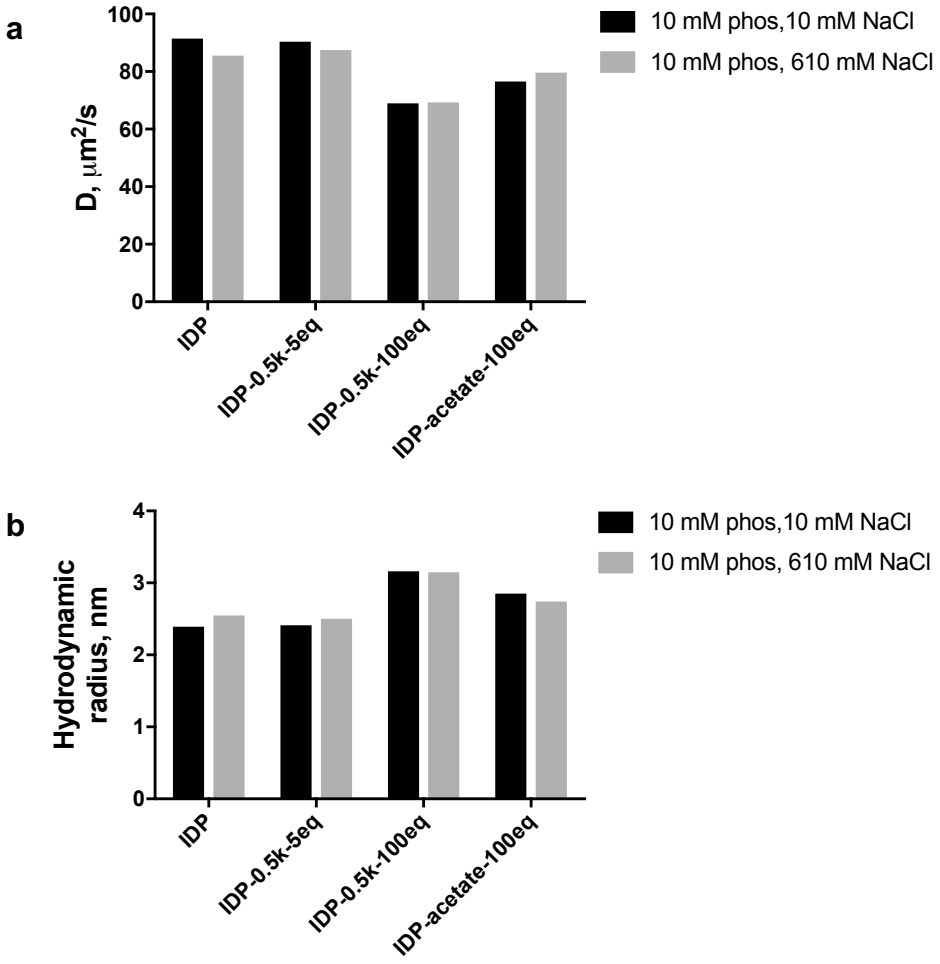


Figure 4.8. Fluorescence correlation spectroscopy study of IDP samples on treatment with NaCl. Diffusion coefficients (a) and hydrodynamic radii (b) of IDP samples were calculated.

4.2.5. Synthesis of fluorescently labelled IDP for Förster Resonance Energy Transfer (FRET) study

The IDP constructs conjugated with A488 were coupled with sulforhodamine-labelled catechol or Alexa Flour 594-labelled 2-pyridinecarboxaldehyde (A594-labelled 2PCA) using the oxidative coupling or 2PCA reaction mentioned in section 1.4. The catechol version of oxidative coupling was developed in our lab as a one-step reaction. The oxidative coupling reaction gave multiple modifications of IDPs. We hypothesized that the lysine residues produced the multiple modifications. Optimizing the conditions by reducing the equivalents of catechol and shortening

the reaction times did not improve the selectivity of the reaction (Figure 4.9). Moreover, decreasing pH reduced the percentage of modified IDPs (data not shown). This was due to the decreasing amounts of unprotonated N-terminal residues as decreasing pH. The percentage of a single modification at a minimal modification of the lysine residues was 14 (Figure 4.9d).

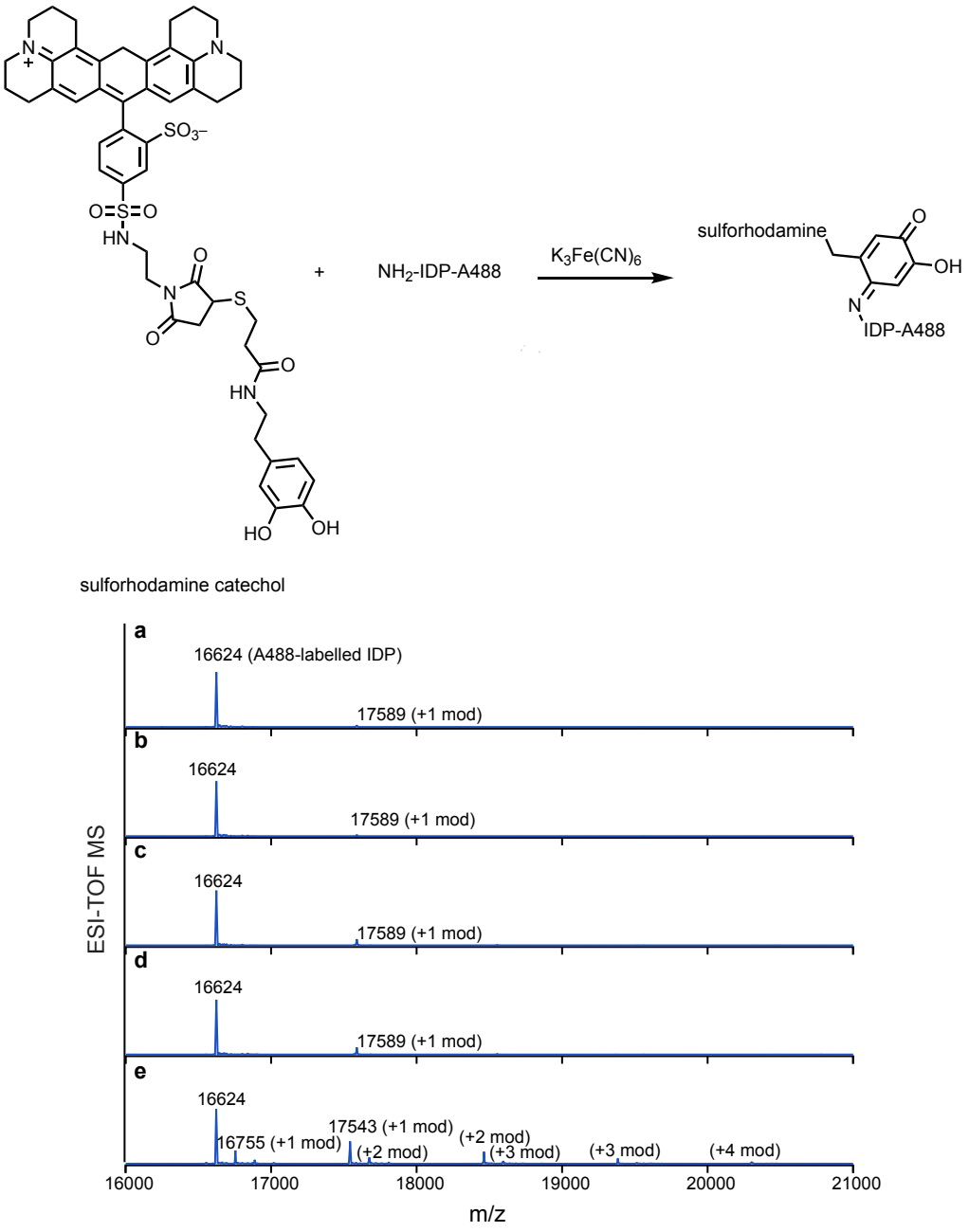


Figure 4.9. Oxidative coupling reaction between IDPs and sulforhodamine catechol. IDPs (20 μM) were reacted with sulforhodamine catechol and ferricyanide in 100 mM phosphate buffer at pH 7.5 at room temperature and analyzed by ESI-TOF mass spectrometry. (a-b) 60 μM sulforhodamine catechol and 600 μM ferricyanide were used for 30 min (a) and 1 h (b). (c-e) 200 μM sulforhodamine catechol and 2000 μM ferricyanide were used for 30 min (c), 1 h (d), and 2 h (e).

The 2PCA reaction was used as an alternative method. A594-labelled 2PCA was mixed with IDPs labelled with A488 and the reaction was analyzed by mass spectrometry (Figure 4.10). The 2PCA reaction also gave multiple modifications of IDPs. We hypothesized that imine formations of lysine and 2PCA occurred. However, adding hydroxylamine at the end of the reaction did not inhibit the formation of imine. For the 2PCA reaction, the percentage of a single modification at a minimal modification of the lysine residues was 30 (Figure 4.10c). Since the 2PCA reaction gave a better yield for the modification, its FRET construct was used further for the PEGylation or acetylation and FRET study.

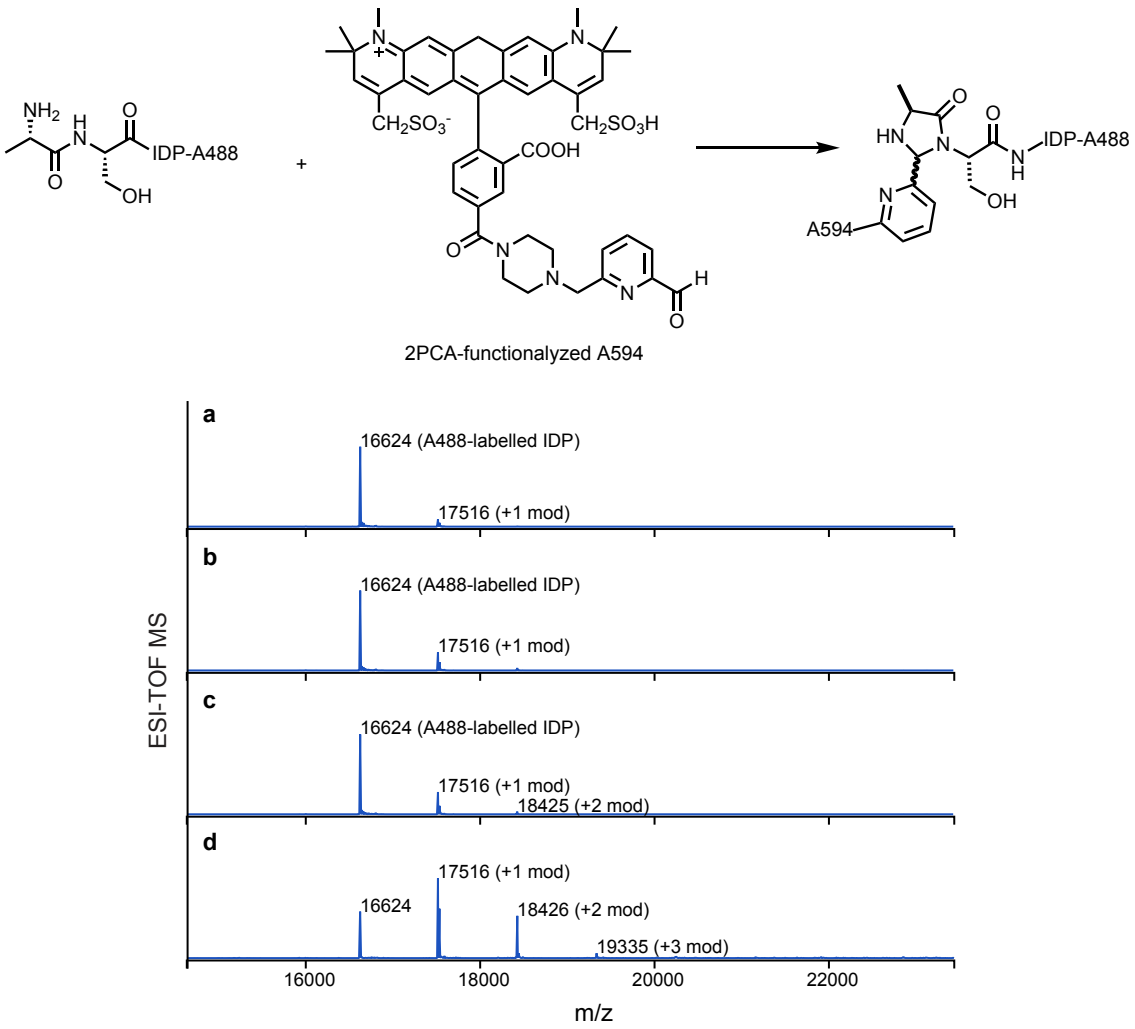


Figure 4.10. A 2PCA reaction between IDPs labelled with Alexa Flour 488 and 2PCA-functionalyzed Alexa Flour 594. IDPs (100 μM) were reacted with 2PCA labelled with Alexa Flour 594 (3 mM) in 100 mM phosphate buffer at pH 7.5 at 37 ºC for 30 min (a), 1 h (b), 2 h (c), and 48 h (d), then analyzed by ESI-TOF mass spectrometry.

4.2.6. FRET study

The FRET study was assessed, and FRET efficiency was calculated and expressed in Table 4.2. Overall, there was no change in FRET efficiency, indicating no conformational change in IDP samples. Even though there were some changes in the anisotropy study, the differences in anisotropy values were small. The amount of fluorescent dyes incorporated in IDPs also affected the FRET experiment. The ideal FRET construct requires more than 80% modification of both dyes. Further optimization of N-terminal modification of the second dye is needed to ensure the accuracy of the FRET study.

Table 4.2. FRET efficiency of IDP samples.

samples	10 mM NaCl	500 mM NaCl
5 equiv. 0.5 k PEG	0.012	0.0253
100 equiv. 0.5 k PEG	0.0128	0.034
100 equiv. NHS acetate	0.034	0.0284

4.3. Conclusions

We have successfully cloned, expressed and purified a 16 kDa IDP, and constructed a PEGylated IDP library with different sizes of PEG chains. Biophysical characterization of the library, including fluorescence anisotropy, fluorescence correlation spectroscopy, and Förster resonance energy transfer uncovered little to no effect of size and density PEG chains on the conformation of IDP. The optimization of the N-terminal modification of IDPs and fluorescent dyes is recommended to incorporate more dyes on to the IDPs, which helps confirm the accuracy of FRET study.

4.4. Materials and Methods

4.4.1. Reagents

Unless otherwise noted, the chemicals and solvents used were of analytical grade and were used as received from commercial sources. Dopamine, 3-mercaptopropanoic acid, 2-merptoethanol, and TCEP were purchased from Aldrich and used without further purification. Sulforhodamine maleimide, A488 maleimide, A594 NHS-ester were purchased from Thermo Fisher. Monodisperse 0.5 kDa PEG NHS-ester (cat# 10260), 1.2 kDa PEG NHS-ester (cat# 10304) and 4 kDa branched PEG NHS-ester (cat# 10454) were purchased from Quanta BioDesign. Analytical thin layer chromatography (TLC) was performed on EM Reagent 0.25 mm silica gel 60-F254 plates with visualization by ultraviolet (UV) irradiation at 254 nm. All organic solvents were removed under reduced pressure using a rotary evaporator. Water (dd-H_2O) used as reaction solvent was deionized using a Barnstead NANOpure purification system (Thermo Fisher). Centrifugations were performed with an Eppendorf Mini Spin Plus (Eppendorf).

4.4.2. Instrumentation and sample analyses

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

Fluorometer. Fluorescence measurements were performed using 3x3 mm Suprasil quartz cuvettes (Hellma Analytics, Germany) with a Fluoromax-4 spectrofluorometer equipped with a Peltier temperature controller and automatic excitation and emission polarizers (HORIBA Jobin Yvon Inc., Edison, NJ).

4.4.3. Synthetic procedures

Synthesis of sulforhodamine catechol. To a solution of dopamine (15.3 mg, 0.1 mmol) in DMF was added 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (15.5 mg, 0.1 mmol), 3-mercaptopropanoic acid (10.6 mg, 0.1 mmol) and triethylamine (0.138 mL, 1 mmol). The reaction mixture was stirred and monitored by TLC for 20 h and then evaporated to remove DMF. The reaction was purified by reversed phase HPLC. The purified mercapto product (0.4 mg, 1.7 μmol) in DMF was then reacted with sulforhodamine maleimide (1 mg, 1.4 μmol) in the presence of 5 mM TCEP. The reaction was purified by reversed phase HPLC. HRMS (ESI) of sulforhodamine catechol calculated for $C_{49}H_{53}N_5O_7S_2$ ($[M]^+$) 887.3381, found 887.3564m/z.

Synthesis of A594-labelled 2PCA. 6-(1-Piperazinylmethyl)-2-pyridinecarboxaldehyde bistosylate salt was synthesized as previously reported.[16] To a solution of 6-(1-piperazinylmethyl)-2-pyridinecarboxaldehyde bistosylate salt (0.83 mg, 1.5 μmol) in DMF was added A594 NHS-ester (1 mg, 1.2 μmol) and triethylamine (0.83 μL, 6 μmol). The reaction mixture was stirred for 18 h and then evaporated to remove DMF. The reaction was purified by reversed phase HPLC. HRMS (ESI) of A594-labelled 2PCA calculated for $C_{47}H_{49}N_5O_{10}S_2$ ($[M]^+$) 907.2921, found 907.3041 m/z.

4.4.4. Protein modification procedures

Coupling A594 or sulforhodamine with A488-labelled IDP. The solution of A488-labelled IDP (0.2 mM) in 50 mM phosphate buffer, pH 7.5, was incubated with 3 mM A594-labelled 2PCA for 16 h at room temperature in the dark. Excess dye was removed using a desalting column (10DG, BioRad) and 10 mM phosphate buffer, pH 7.5. For the oxidative coupling reaction between IDPs and sulforhodamine catechol, IDPs (20 μM) were reacted with sulforhodamine catechol (3 or 10 equiv. with respect to IDP) and ferricyanide (30 or 100 equiv. with respect to IDP) in 100 mM phosphate buffer at pH 7.5 at room temperature.

PEGylation of IDPs. PEG NHS-ester reagents (20 mM stock solutions in H_2O) were made fresh immediately before PEGylation reactions. PEG-NHS- or TFP- ester reagent (5-100 equiv. with respect to IDP) was added to a 50 μL solution of A488 labeled-IDP (152 μM) in 20 mM HEPES, pH 7.5, 50 mM NaCl. This buffer was also added to each reaction to reach a final volume of 150 μL. The reactions were incubated at room temperature in the dark for 12-18 h before purification and buffer exchange using spin concentrators with a 10 kDa molecular weight cut-off (Amicon Ultra, EMD Millipore). The resulting PEGylated IDP was analyzed by LC/MS.

4.4.5. Additional experimental procedures

Plasmid Preparation for IDP encoding pSKB3 vector. The inherent repetitive sequence of IDP made the synthesis of one contiguous gene block impossible. Instead, two gene blocks (gBlocks: IDT Technologies) had to be synthesized (IDP-1 and IDP2; see supporting information for respective sequences) with a 32 bp consensus sequence that allows for Gibson assembly. 100 ng of each gBlock and 10 µL of 2x Gibson master mix[17] (a gift from the Chang lab) was adjusted with water to a volume of 20 µL and was incubated at 50 °C for 60 min. After DNA cleanup with QiaQuick (Qiagen), the assembly product was PCR amplified (VENT polymerase from NEB, T_m=61 °C) with forward and reverse primers: 5'- ATA ATA GCT AGC TTA GTT CCT CGT GCC TGG CGT G -3' and 5'- TAT TAT CTC GAG CTA TTA GGC ACA CCA GTA CGG AGA TTT C -3'. The IDT insert contained NheI and XhoI restriction sites, that were double digested, heat inactivated at 80 °C for 5 min and ligated (QuickLigase, NEB) with an 5'-terminal MBP containing pSKB3 vector with a 5' containing MBP site. Plating on Kanamycin agar plates yielded individual colonies that were cultured, DNA purified (NucleoSpin, Macherey-Nagel) and sequenced (Quintara BioSciences). The resulting vector yielded the sequence shown in Figure 4.1.

Protein expression and purification. Plasmids were transformed into *E. coli* BL21 (DE3) competent cells. Starter cultures (20 mL LB, 50 mg/L Kanamycin) were grown from single colonies, grown overnight at 37 °C, and used to inoculate 1 L of TB media (50 mg/L Kanamycin). Cultures were grown to an OD ~ 0.5, cooled for 20 min at 25 °C, induced with 0.5 mM IPTG, and expressed overnight (~ 18 h) at 25 °C. Cells were harvested by centrifugation for 15 min at 4,000 rcf (g) at 4 °C. The pellet was transferred to a 50 mL Falcon tube in PBS buffer, and spun down for 10 min at 4000 rcf (g). The resulting pellet (~ 5 g) was lysed in 30 mL of lysis buffer (20 mM HEPES, pH 7.5, 300 mM NaCl, 10 mM imidazole = buffer A) supplemented with one tablet of EDTA-free SigmaFast protease inhibitor (Aldrich), 2 mM PMSF, and 10 mg lysozyme. The resuspended sample was lysed with an Avestin C3 homogenizer followed by 20 min of centrifugation at 24,000 rcf (g) at 4 °C. The supernatant was filtered through a 40 µm Steriflip filter (Millipore), and loaded onto a 5 mL NiNTA column (Protino, Macherey-Nagel) connected to an Akta purifier that was pre-equilibrated with buffer A. The column was washed with 50 mL (10 CV) of 20 mM HEPES, pH 7.5, 300 mM NaCl, 10 mM imidazole, 10 mM 2-merptoethanol. The protein was eluted with 20 mM HEPES, pH 7.5, 300 mM NaCl, 250 mM imidazole, 10 mM 2-merptoethanol. Imidazole was removed by exchanging against 20 mM HEPES, pH 7.5, 100 mM NaCl with a 10DG desalting column (BioRad), and subsequently digested with 1 mg of thrombin protease (high purity from Bovine, MP Biomedicals). Complete digestion was achieved at room temperature after 1 h, as confirmed by LC/MS (ESI/TOF). The protein sample was added A488 maleimide (4 equiv. with respect to IDP), in the presence of 5 mM TCEP (0.5 M, Aldrich). The reaction was briefly agitated to ensure proper mixing and incubated at 25 °C for 1 h. The protein mixture was diluted with salt-free HEPES buffer (20 mM, pH 7.5) to 50 mL (5 mM NaCl) before loading it onto a 1 mL HiTrap SP HP cation exchange column connected to an Akta purifier (GE Healthcare). The column was pre-equilibrated with 20 mM HEPES, 10 mM 2-merptoethanol, washed with 10 mL (10 CV) of 20 mM HEPES, pH 7.5, 10 mM 2-merptoethanol and eluted with a gradient from 0-1 M NaCl (50 mL total volume). IDP eluted at around 200 mM NaCl. The final A488-labelled IDP sample was obtained with a final desalting column (10DG, BioRad) to obtain the protein at a final concentration of 270 µM in 20 mM HEPES (pH=7.5), 50 mM NaCl. The protein was >95% pure by SDS-PAGE and LCMS. The purified protein was flash frozen with liquid N_2 in 20 µL aliquots.

Fluorescence Anisotropy and FRET experiment. Samples were diluted to an OD ≤0.03 and filtered through a 0.22 μm centrifugal filter prior to analysis. All measurements were collected with a 0.1 s integration time. Fluorescence anisotropy experiments were performed using 5.0 nm excitation and 1.0 nm emission slit widths at 25.0 °C, unless otherwise noted. Anisotropy measurements were reported as the mean and standard deviation of a minimum of 9 measurements. The solution of IDP samples was prepared in 1 μM PEGylated IDP conjugated with A488 in 10 mM phosphate buffer, unless otherwise noted. NaCl titration was carried out by adding 5 M NaCl incrementally.

4.5. Supporting information

Open reading frame sequence of MBP-IDP encoding pSKB3 vector:

CC**ATG**GCCAGCAGCCATCATCATCATCACGATTACGATATCCCAACGACCGAA
AACCTTTACTTCCAGGGATCCGAAAACCTTTACTTCCAGGGATCCAAAATCGAAGAA
GGTAAACTGGTAATCTGGATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGT
CGGTAAGAAATTCGAGAAAGATACCGGAATTAAAGTCACCGTTGAGCATCCGGATA
AACTGGAAGAGAAATTCCCACAGGTTGCGGCAACTGGCGATGGCCCTGACATTATC
TTCTGGGCACACGACCGCTTTGGTGGCTACGCTCAATCTGGCCTGTTGGCTGAAATC
ACCCCGGACAAAGCGTTCCAGGACAAGCTGTATCCGTTTACCTGGGATGCCGTACGT
TACAACGGCAAGCTGATTGCTTACCCGATCGCTGTTGAAGCGTTATCGCTGATTTAT
AACAAAGATCTGCTGCCGAACCCGCCAAAAACCTGGGAAGAGATCCCGGCGCTGGA
TAAAGAACTGAAAGCGAAAGGTAAGAGCGCGCTGATGTTCAACCTGCAAGAACCGT
ACTTCACCTGGCCGCTGATTGCTGCTGACGGGGGTTATGCGTTCAAGTATGAAAACG
GCAAGTACGACATTAAAGACGTGGGCGTGGATAACGCTGGCGCGAAAGCGGGTCTG
ACCTTCCTGGTTGACCTGATTAAAAACAAACACATGAATGCAGACACCGATTACTCC
ATCGCAGAAGCTGCCTTTAATAAAGGCGAAACAGCGATGACCATCAACGGCCCGTG
GGCATGGTCCAACATCGACACCAGCAAAGTGAATTATGGTGTAACGGTACTGCCGA
CCTTCAAGGGTCAACCATCCAAACCGTTCGTTGGCGTGCTGAGCGCAGGTATTAACG
CCGCCAGTCCGAACAAAGAGCTGGCAAAAGAGTTCCTCGAAAACTATCTGCTGACT
GATGAAGGTCTGGAAGCGGTTAATAAAGACAAACCGCTGGGTGCCGTAGCGCTGAA
GTCTTACGAGGAAGAGTTGGTGAAAGATCCACGTGTTGCCGCCACTATGGAAAACG
CCCAGAAAGGTGAAATCATGCCGAACATCCCGCAGATGTCCGCTTTCTGGTATGCCG
TGCGTACTGCGGTGATCAACGCCGCCAGCGGTCGTCAGACTGTCGATGAAGCCCTG
AAAGACGCGCAGACTAATTCGAGCTCGAACAACAACAACAATAACAATAACAACA
ACCTCGGGGCTAGCTTAGTTCCTCGTGCCTGGCGTGGCTCCCCGTGGGCAGAGGCCA
AGAGTCCAGCGGAAGCTAAGTCGCCAGCCGAAGTCAAGTCGCCCGCCGTCGCGAAA
AGCCCCGCAGAGGTGAAATCCCCGGCCGAAGTCAAATCGCCGGCAGAAGCGAAATC
CCCGGCAGAAGCAAAAAGTCCTGCTGAGGTCAAATCGCCAGCAACCGTCAAATCCC
CTGGAGAGGCAAAATCTCCGGCAGAAGCCAAGTCCCCTGCCGAAGTGAAGTCACCT
GTCGAAGCCAAGTCGCCGGCCGAAGCGAAGAGCCCAGCGAGCGTGAAAAGTCCTG
GTGAGGCTAAGTCCCCGGCGGAAGCGAAATCTCCAGCGGAAGTAAAGAGTCCGGCC
ACCGTTAAATCCCCGGTAGAGGCCAAAAGCCCTGCGGAAGTTAAATCGCCGGTGAC
GGTCAAATCACCCGCGGAAGCGAAGTCCCCGGTGGAGGTGAAATCTCCGTACTGGT
GTGCC**TAA**TAGCTCGAG

Underlined restriction sites in order of sequence: NcoI/BamHI/NheI/XhoI
Bold: Start and Stop codons

Green: 6His affinity tag encoding sequence
Orange: TEV encoding sequence
Blue: Maltose Binding Protein encoding sequence
Purple: N_{10} linker encoding sequence
Orange: Thrombin encoding sequence
Red: IDP encoding sequence
Cyan: Cysteine encoding sequence

gBlock IDP-1:

ATAATAGCTAGCTTAGTTCCTCGTGCCTGGCGTGGCTCCCCGTGGGCAGAGGCCAAG
AGTCCAGCGGAAGCTAAGTCGCCAGCCGAAGTCAAGTCGCCCGCCGTCGCGAAAAG
CCCCGCAGAGGTGAAATCCCCGGCCGAAGTCAAATCGCCGGCAGAAGCGAAATCCC
CGGCAGAAGCAAAAAGTCCTGCTGAGGTCAAATCGCCAGCAACCGTCAAATCCCCT
GGAGAGGCAAAATCTCCGGCAGAAGCCAAGTCCCCTGCCGAAGTGAAGTCAC

gBlock IDP-2:

AGAAGCCAAGTCCCCTGCCGAAGTGAAGTCACCTGTCGAAGCCAAGTCGCCGGCCG
AAGCGAAGAGCCCAGCGAGCGTGAAAAGTCCTGGTGAGGCTAAGTCCCCGGCGGA
AGCGAAATCTCCAGCGGAAGTAAAGAGTCCGGCCACCGTTAAATCCCCGGTAGAGG
CCAAAAGCCCTGCGGAAGTTAAATCGCCGGTGACGGTCAAATCACCCGCGGAAGCG
AAGTCCCCGGTGGAGGTGAAATCTCCGTACTGGTGTGCCTAATAGCTCGAGATAATA
Underlined: Consensus sequence used for Gibson Assembly

4.6. References

1. Yoo, J.-W.; Irvine, D. J.; Discher, D. E.; Mitragotri, S., *Nat. Rev. Drug Discov.* **2011,** *10* (7), 521.
2. Elzoghby, A. O.; Samy, W. M.; Elgindy, N. A., *J. Controlled Release* **2012,** *161* (1), 38-49.
3. Vázquez, E.; Villaverde, A., *Microb. Cell Fact.* **2010,** *9* (1), 101.
4. Gaumet, M.; Vargas, A.; Gurny, R.; Delie, F., *Eur. J. Pharm. Biopharm.* **2008,** *69* (1), 1-9.
5. Gratton, S. E.; Ropp, P. A.; Pohlhaus, P. D.; Luft, J. C.; Madden, V. J.; Napier, M. E.; DeSimone, J. M., *Proceedings of the National Academy of Sciences* **2008,** *105* (33), 11613-11618.
6. Otsuka, H.; Nagasaki, Y.; Kataoka, K., *Adv. Drug Del. Rev.* **2012,** *64*, 246-255.
7. Hak, S.; Helgesen, E.; Hektoen, H. H.; Huuse, E. M.; Jarzyna, P. A.; Mulder, W. J.; Haraldseth, O.; Davies, C. d. L., *ACS nano* **2012,** *6* (6), 5648-5658.
8. Wright, P. E.; Dyson, H. J., *J. Mol. Biol.* **1999,** *293* (2), 321-331.
9. Dunker, A. K.; Lawson, J. D.; Brown, C. J.; Williams, R. M.; Romero, P.; Oh, J. S.; Oldfield, C. J.; Campen, A. M.; Ratliff, C. M.; Hipps, K. W., *J. Mol. Graphics Model.* **2001,** *19* (1), 26-59.
10. Hwang, P. M.; Pan, J. S.; Sykes, B. D., *Protein Expression Purif.* **2012,** *85* (1), 148-151.
11. di Guanº, C.; Li, P.; Riggs, P. D.; Inouye, H., *Gene* **1988,** *67* (I988), 2I-30.
12. Terpe, K., *Appl. Microbiol. Biotechnol.* **2003,** *60* (5), 523-533.
13. Pryor, K. D.; Leiting, B., *Protein Expression Purif.* **1997,** *10* (3), 309-319.

14. Gallwitz, M.; Enoksson, M.; Thorpe, M.; Hellman, L., *PLoS One* **2012,** *7* (2), e31756.
15. Fluorescence Anisotropy. In *Principles of Fluorescence Spectroscopy*, Lakowicz, J. R., Ed. Springer US: Boston, MA, 2006; pp 353-382.
16. MacDonald, J. I.; Munch, H. K.; Moore, T.; Francis, M. B., *Nat. Chem. Biol.* **2015,** *11* (5), 326.
17. Gibson, D. G.; Young, L.; Chuang, R.-Y.; Venter, J. C.; Hutchison III, C. A.; Smith, H. O., *Nat. Methods* **2009,** *6* (5), 343.