Genetically encoded 3,4-ethylenedioxythiophene (EDOT) functionality for fabrication of protein-based conductive polymers

Thesis by

Maiko Obana

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy in Chemistry

CALIFORNIA INSTITUTE OF TECHNOLOGY

Pasadena, California

2021

Defended October, 2020

Maiko Obana
ORCID: 0000-0003-4150-0055

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my thesis advisor, Prof. David Tirrell, for his guidance and support. In the beginning of my graduate studies, I struggled to find a project I would like to pursue for my PhD. He kindly listened to me about my research interest and guided me to work on a project I really enjoyed throughout my PhD. I also appreciate his respect to his students. Dave always allows us to think on our own. Although I was unconfident about myself when I started at Caltech, I learned to think myself as an independent scientist over the course of my PhD. Thank you for everything, Dave.

I would like to thank to the members of my thesis committee, Prof. Jacqueline Barton, Prof. Julia Kornfield, and Prof. Maxwell Robb, for their support and advice. Committee meetings were always fun and helpful, and I appreciate many constructive comments on my project from all the different perspectives.

I would like to thank to past and present members of the Tirrell group, for a supportive and fun environment. I learned not only technical skills, but also a mindful attitude from my colleagues. I would especially like to thank Dr. Bradley Silverman for working with me on the colloidal assembly project in my first year. He kindly and patiently taught me molecular biology techniques when I literally knew nothing about cloning and protein expression as a first-year student. The project could not have been done without his careful image analysis and constructive discussions. Thanks to Dr. Samuel Ho and Stephanie Breunig for their advice on my research in the Protein Modification subgroup as well as working with me as safety coordinators. They are always easy to talk to and I really appreciate their willingness to provide any help anytime. Thanks to Dr. Shannon Stone and Dr. Lawrence Dooling for helpful advice, especially for helpful discussions when I was struggling to find a project to work on during my PhD.

I would also like to thank Prof. Edmund Tse, a former postdoc in the Barton group, for technical assistance with electrochemistry experiments and discussions. Thanks to Dr. Mona Shahgholi at the Mass Spectrometry Laboratory for her help with mass spectrometry analysis. Thanks to Dr. Daisuke Saito, Dr. Naoyuki Suzuki, and Dr. Yasuaki Nakayama,

who are former postdocs in the Stoltz group and the Reisman group, for the helpful advice in amino acid synthesis. Thanks to Jiajun Du for assistance with Raman spectroscopy, Kai Narita for assistance with conductivity measurement, and Dr. Jeong Hoon Ko for a lot of helpful discussions throughout my research. Thanks to people in the Barton group, especially Dr. Rebekah Silva for always welcoming me to use their potentiostat. Thanks to the Nakajima Foundation for the financial support over the five years of my Ph.D. I also would like to thank Dr. Takahiro Fukino, Prof. Keiichiro Kushiro, and my undergraduate advisor Prof. Takuzo Aida. Without their support and encouragement, I could not have taken on the challenge to study abroad for my PhD.

Thanks to all my friends, especially to Tomoyuki Oniyama, Silvia Kim, Jieun Shin. Making friends with people like you was one of the greatest parts of graduate school. Thanks to all my friends in the Japanese and the Korean Tennis Clubs, and the Caltech Japanese Student Association. Especially thanks to Mooseok Jang, Albert Chung, and Sonjong Hwang for training tennis with me every week (or sometimes five days a week). Thanks to Yoshihiro Nakayama, Andy Brown, and all of the Pasadena Ultimate Frisbee friends. Ultimate is the greatest thing I learned during graduate school besides science.

Finally, I would like to thank my family. I grew up in a little rural town in Japan, so allowing me to study in Tokyo for college and then the US for graduate school must have been difficult for my parents. But they accepted me whatever I wanted to do, and they were always supportive. Thanks to my older sister, who has been a great friend throughout my life. Thanks to my parents-in-law Christine and Richard, for welcoming me into their family. Thanks to my husband Kevin, who always gives me tremendous support in my life and research. Thank you for everything.

ABSTRACT

Genetic code expansion provides powerful strategies to improve the properties of protein-based materials. One novel application of this technique is to genetically incorporate an electroactive functional group into proteins, which can be subsequently polymerized into conductive polymers, enabling fabrication of various protein–conductive polymer hybrids that are widely applicable in bioelectronics. To this end, we developed a technique to incorporate an amino acid bearing 3,4-ethylenedioxythiophene (EDOT), the monomer precursor of a well-known conductive polymer PEDOT.

In Chapter 1, we review the basics of protein-based materials and genetic code expansion technology. We also highlight some examples where genetic code expansion was used for the development of protein-based materials. Finally, we discuss applications of protein and peptide-based materials in bioelectronics.

In Chapter 2, we describe our effort to incorporate an amino acid bearing EDOT group (EDOT-Ala) designed as an analogue of aromatic canonical amino acids. We synthesized EDOT-Ala in three steps of organic reaction, and evaluated the activity of known aminoacyl-tRNA synthetase (aaRS) variants for EDOT-Ala. In addition, we performed evolution of aaRS for EDOT-Ala using two different evolution techniques: cell viability-based approach and phage-assisted approach. Although the evolution experiment did not yield an aaRS variant that can incorporate EDOT-Ala, the results presented in this chapter provide valuable information for engineering of aaRS and incorporation of non-canonical amino acids (ncAA) with bulky functional groups.

In Chapter 3, we describe the incorporation of another EDOT-functionalized amino acid (EDOT-Lys) designed as an analogue of a canonical amino acid pyrrolysine (Pyl). When we co-expressed a GFP reporter and a mutant pyrrolysyl-tRNA synthetase (PylRS) in *E. coli* in the presence of EDOT-Lys, the cells exhibited strong fluorescence as an indication of successful incorporation of EDOT-Lys into GFP. We further confirmed the incorporation using MALDI-TOF mass spectrometry.

In Chapter 4, we describe the electropolymerization of a model protein XTEN that carries genetically incorporated EDOT-Lys (XTEN-E49am). We performed electropolymerization of XTEN-E49am in the presence of a self-doping EDOT monomer (EDOT-S) by cyclic voltammetry. The solution formed dark blue solids immediately after the potential cycles. The composition of the product was determined by FT-IR spectroscopy, suggesting that one protein is found per 12.5 monomer units of PEDOT. In addition, we investigated the effect of amino acids located adjacent to EDOT-Lys. Although the presence of cysteine (Cys), lysine (Lys), methionine (Met), arginine (Arg), and tryptophan (Trp) located adjacent to EDOT-Lys had an impact on the electropolymerization of model peptides, XTEN proteins carrying these adjacent residues (XTEN-E49am-G50Z; Z = Cys, Lys, Met, Arg, Trp) were electropolymerized with EDOT-S without noticeable effect from these adjacent residues, indicating that the EDOT-Lys residues in proteins undergo electropolymerization with EDOT-S in different chemical environments.

In Chapter 5, we describe the oxidative chemical polymerization of XTEN proteins and model peptides. When XTEN-E49am was polymerized with EDOT-S by addition of ammonium persulfate (APS) and iron(III) chloride ($FeCl_3$), the solution yielded dark blue solids. To evaluate the reactivity of the EDOT-Lys residue in the protein, we reacted the protein with an end-capped EDOT derivative (EDOT-cap). MALDI-TOF mass spectrometry revealed the appearance of new peaks corresponding to the addition of one or two EDOT-caps to the protein, suggesting that EDOT-Lys residue in the protein can react with EDOT derivatives. We also investigated the effect of adjacent amino acids using a series of model peptides. In polymerization using APS without $FeCl_3$ catalyst, peptides carrying basic amino acids (His, Lys, Arg) adjacent to EDOT-Lys showed enhanced polymerization compared to the ones carrying neutral and acidic adjacent residues. All the tested peptides polymerized well when $FeCl_3$ was added as a catalyst. The results presented in this chapter provide valuable insights into synthesis of protein–PEDOT conjugates via oxidative chemical polymerization.

PUBLISHED CONTENT AND CONTRIBUTIONS

(1) Obana, M.*; Silverman, B. R.*; Tirrell, D. A. Protein-Mediated Colloidal Assembly. *J. Am. Chem. Soc.* **2017,** 139, 14251–14256. DOI: 10.1021/jacs.7b07798.

M. O. contributed to the experimental work, data analysis, conception of the project, and preparation of the manuscript.

*Co-first authors.

(2) Obana, M.; Tse, E. M. C.; Barton, J. K.; Tirrell, D. A. Genetically Encoded 3,4-Ethylenedioxythiophene (EDOT) Functionality for Fabrication of Protein-Based Conductive Polymers. *Manuscript to be submitted.*

M. O. contributed to the experimental work, data analysis, conception of the project, and preparation of the manuscript.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF TABLES

Chapter 1

PROTEIN-BASED MATERIALS

1.1. Introduction

With unique structural and functional properties, proteins have attracted significant attention as a source of biomaterials. Diverse protein-based materials have been developed by adapting natural proteins and by rational and computational design. The properties of protein-based materials can be further improved by protein modification via post-translational and co-translational methods. Genetic code expansion is a powerful protein modification technique that enables incorporation of non-canonical amino acids (ncAAs) into proteins. This technique is especially useful in introducing new functions to protein-based materials while taking advantage of inherent properties of the proteins. One novel application of genetic code expansion is to introduce an electroactive functional group into proteins in order to form conductive polymers. The overall goal of this thesis project is to incorporate an electroactive ncAA into proteins for the fabrication of protein-based conductive materials. In this introductory chapter, we describe protein-based materials and the improvement of their properties by genetic code expansion. We also discuss protein- and peptide-based materials for applications in bioelectronics.

1.2. Protein-Based Materials

Proteins are highly attractive sources of biomaterials because of their unique structural and functional properties. They have evolved to perform a vast array of functions in living organisms, including catalysis, recognition, and structural support. By adapting the sequences of proteins from nature, various protein-based materials have been developed.[1–4] Typical examples are based on structural proteins such as elastin, silk, and collagen elastin (Figure 1.1).[2] These structural proteins are not only used in their full natural forms, but their repeated motifs are also used as building blocks to produce biomaterials. For example, the hydrophobic domains of elastin are adapted to synthesize elastin-like polypeptides (ELPs), recapturing the elasticity and thermal responsiveness of the original domains.[3]

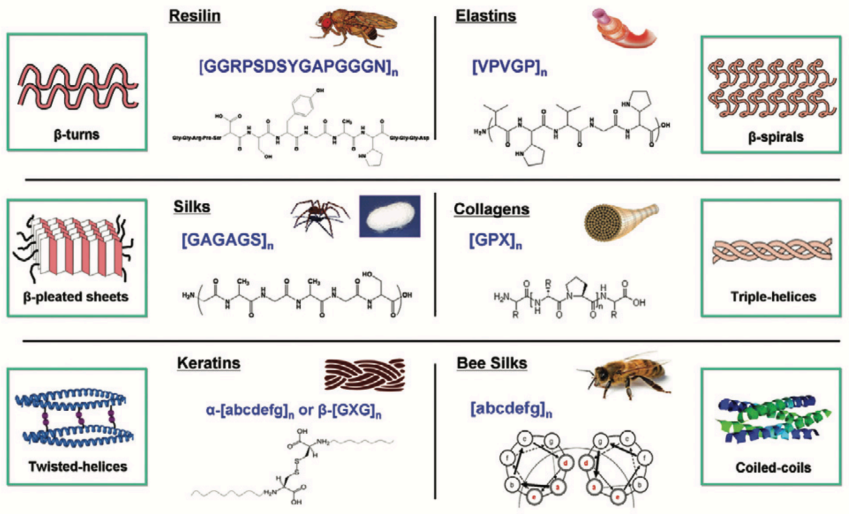


Figure 1.1. Natural proteins commonly used for protein-based materials. Adapted from ref. 2.

With the better understanding of the relationship between the sequences and the properties of proteins in recent decades, various protein-based materials have been developed by rational and computational designs.[1,5–8] Secondary motifs such as α-helices and β-sheets are the critical elements in facilitating the formation of higher-order protein assemblies and serve as the building blocks in designing protein-based

materials.[6,7] While each interaction is usually too weak to achieve sufficient affinity to assemble proteins, a bundle of the interactions exhibits a much stronger force, facilitating protein association. Various motifs and sequences have been used to form unique protein assemblies such as filaments, cages, and cross-linked networks.[1,3,6,7]

When necessary, protein modification via post-translational or co-translational methods is employed to improve the properties of protein-based materials.[9–12] Bioconjugation at the reactive residues such as cysteine and lysine is commonly used for labeling proteins with functional molecules.[12,13] Enzymatic modification also provides a powerful strategy to functionalize proteins with high sequence specificity and mild reaction conditions.[12,14,15] Various enzymes such as transglutaminases, peroxidases, sortase, and biotin ligase are employed to modify proteins at specific residues or sequences. Proteins can be also modified co-translationally by a technique known as genetic code expansion, which utilizes non-canonical amino acids (ncAAs) to introduce novel functionalities into proteins.[9,10] We discuss genetic code expansion and its application for the development of protein-based materials in the next section.

1.3. Genetic Code Expansion for Protein-Based Materials

In contrast to conventional synthetic polymers, proteins can be prepared with perfect control of size, sequence, and stereochemistry through gene expression. However, their chemical diversity is limited to the functional groups that comprise the side chains of twenty canonical amino acids. To circumvent this limitation, protein chemists have developed a powerful strategy to introduce novel functionality into proteins by incorporation of non-canonical amino acids (ncAAs).[10,16–19] Since the seminal work by Cowie and Cohen in the 1950s to incorporate selenomethionine as a replacement for methionine,[20] hundreds of ncAAs have been incorporated into proteins.[10,16–19]

The key to incorporation of ncAAs is the aminoacylation step in protein biosynthesis (Figure 1.2).[10,16–19] During aminoacylation, enzymes known as aminoacyl-tRNA synthetases (aaRSs) charge amino acids to their cognate tRNAs, and the aminoacylated tRNAs are subsequently used for translation at the ribosome, integrating the amino acids into polypeptide chains. Because the ribosome has relaxed substrate specificity, the limiting step to incorporate ncAAs is recognition by aaRS during aminoacylation. Once ncAAs are charged to tRNAs, they can be incorporated into the protein during translation.

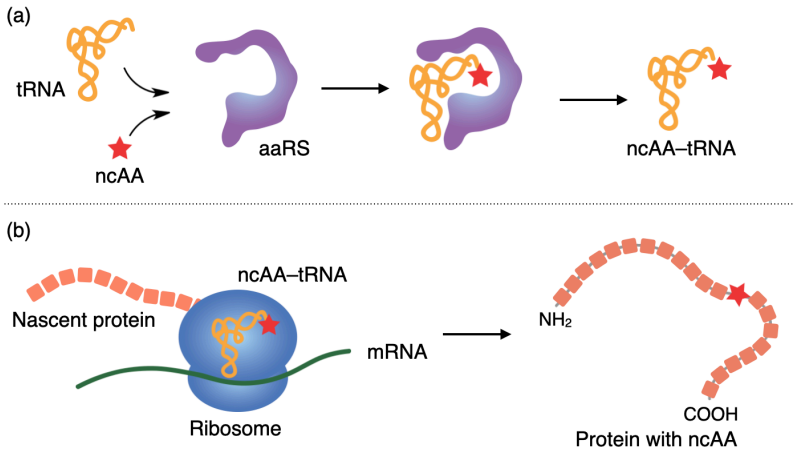


Figure 1.2. Protein biosynthesis and incorporation of ncAA. (a) Aminoacylation. (b) Translation.

There are two major approaches to incorporate ncAAs: the residue-specific method and the site-specific method.[10,16–19] The residue-specific method globally replaces one of the canonical amino acids with a non-canonical analogue. The cognate tRNA for a specific canonical amino acid is charged with ncAA by aaRS, and the ncAA is assigned to the codons for the canonical amino acid. To avoid the competition with the canonical amino acid, incorporation is often performed using an auxotrophic cell strain and minimal media lacking the canonical amino acid. In the site-specific method, ncAA is only incorporated at a specific codon, usually one of the three stop codons. A nonsense suppressor, an engineered tRNA that recognizes a stop codon, is charged with ncAA by aaRS and inserts the ncAA at the

target site. These two approaches are complementary to each other and applied for various purposes including bioorthogonal reactions, proteomic study, cross-linking, study of posttranslational modification, and spectroscopic probes.[10,16–19]

Various studies have shown that ncAAs can be effectively employed for the development of protein-based materials.[9,10] Typical examples are functionalization of protein-based materials using ncAAs as reactive handle. Strable and coworkers incorporated azidohomoalanine (Aha) and homopropargylglycine (HPG) into virus-like particles and labeled the particles with a variety of reagents via copper(I)-catalyzed azide–alkyne cycloaddition (CuAAC).[21] In another report, they functionalized the virus particles with gadolinium complexes using Aha and studied the plasma clearance and tissue distribution of the particles in mice.[22] *p*-Aminophenylalanine (NH_2-Phe) was also employed for functionalization of MS2 virus capsid with DNA aptamers via oxidative coupling, and the modified virus capsid was successfully targeted to cancer cells.[23] For the functionalization of proteins with a synthetic polymer, incorporation of a phenylalanine derivative carrying the initiator of atom transfer radical polymerization (ATRP) has been demonstrated.[24] The subsequent polymerization successfully produced protein–polymer conjugates. The use of ncAA is not limited to proteins expressed in bacterial hosts. Teramoto and coworkers developed transgenic silkworms (*Bombyx mori*) and incorporated *p*-azidophenylalanine (N_3-Phe) into silk fibroin.[25,26] The resulting silk fibers were able to react with alkyne-funtionalized fluorophore via CuAAC.

NcAAs with a photo-reactive functional group have also been employed to control material properties in response to UV irradiation. Carrico and coworkers demonstrated the lithographic patterning of artificial extracellular matrix proteins using N_3-Phe as photo-crosslinking agent.[27] The elastic moduli of the films were controlled by the amount of N_3-Phe incorporated into the protein,[27,28] and mammalian cells adhered selectively to the region of the surface patterning of the protein film.[27] This technique was later used for the generation of protein surface

gradients in the microfluidic device.[29] Another example is mussel adhesive protein incorporated with a photo-caged 3,4-dihydroxyphenylalanine (DOPA) derivative.[30] Cleavage of nitrobenzyl groups by UV irradiation unmasked catechol side chains of DOPA and enhanced the adhesive property of the protein. Incorporation of ncAAs appears to be especially useful in introducing new functions to proteins while taking advantage of inherent properties of the proteins.

1.4. Bioelectronics Applications of Proteins and Peptides

With favorable mechanical properties, designability and biocompatibility, materials based on proteins and peptides are especially attractive for bioelectronics applications.[31–33] In order to introduce electrical conductivity, inorganic and organic conductors are often integrated with protein and peptide-based materials, producing biohybrid materials.[33–38] One of the major roles of proteins and peptides in the conductive biohybrid materials is to serve as structural support.[34] Protein fibers such as amyloids, silk, and elastin have been used as templates for assembly of metal nanoparticles to form protein nanowires.[39–44] Various virus particles have also been used to assemble metals and semiconductors.[45] Materials assembled from proteins and peptides offer advantages such as precise and tunable spatial control and the versatility of surface chemistry, and can be applied for biosensors, supercapacitors, and organic solar cells.[40,43,46]

In addition, proteins and peptides can be genetically programmed in the cells to serve as essential components of engineered living materials.[47–49] For example, Chen and coworkers engineered *E. coli* to produce curli amyloid fibrils in defined patterns and formed the network of gold nanoparticles in the biofilm.[47] Using multiple variants of the protein component CsgA, they demonstrated the assembly of nanoparticles into different patterns such as nanowires and nanorods. Nguyen and coworkers also developed biofilm-based materials using engineered curli nanofibers.[48] They fused a variety of functional peptides with the CsgA protein and demonstrated the assembly of silver nanoparticles and the cellular adhesion to

stainless steel. These examples illustrate the utility of protein and peptide assemblies to integrate inorganic conductive materials into the living system.

Protein- and peptide-based materials are also combined with carbon-based materials and conductive polymers.[36,37] Silk fibers, elastin-like polypeptide (ELP) hydrogels, and keratin-based fibers have been integrated with carbon nanotubes (CNTs), graphene, and graphene oxide to introduce electrical conductivity.[50–59] Amyloid fibrils and silks are also coated with conductive polymers such as polyaniline (PANI) and poly(3,4-ethylenedioxythiophene) (PEDOT) to form conductive nanowires.[60–65] The electrode composed of PEDOT-coated silk fiber was employed for electrophysiological recording in rats and successfully recorded the brain and cardiac activity.[66] Due to good mechanical compatibility with soft biological tissues, the composites of conductive polymers and proteins are especially suitable for the interface materials in bioelectronics.

Besides the structural support of biohybrid materials, proteins and peptides are also responsible for interactions with biological systems. Proteins and peptides in conductive biohybrids create a biocompatible interface that supports cell adhesion, and thereby stabilize cell proliferation and differentiation. Extracellular matrix (ECM) proteins such as collagen and fibronectin are often combined with carbon-based materials to improve their cell adhesive property.[57,67–72] Peptide motifs derived from ECM proteins (e.g. RGD, IKVAV) have also been introduced to conductive polymers in order to enhance cell adhesion, growth, and differentiation.[73–76] These materials are able to conduct electrical current through biological samples and especially useful for cell culture materials and neural probes, enabling electrical stimulation of the cells and recording the cellular activity.[77,78] The ability of peptides to specifically interact with biomolecules is also exploited for biosensors. Peptides binding to various biomolecules such as DNA, proteins, viruses, and small molecules have been combined with conductive materials and shown to detect the target molecules on the electrode.[79–82] Because

peptides can be engineered to bind a wide range of substances,[83–85] these techniques provide versatile platforms for diagnosis and monitoring.

1.5. Conclusion

In addition to diverse structural and functional properties, protein-based materials can be further improved using protein engineering techniques. Genetic code expansion is a powerful strategy to introduce new functionalities into proteins and has contributed to the development of protein-based materials. One novel application of this technique is to introduce an electroactive functional group into proteins that can be subsequently polymerized into conductive polymers. The resulting protein–conductive polymer hybrids are useful for various bioelectronics applications such as cell culture materials for electrical stimulation, interface for bioelectronic devices, and conductive living materials.

1.6. References

(1) H. Smith, K.; Montes, E. T.; Poch, M.; Mata, A. Integrating Top-down and Self-Assembly in the Fabrication of Peptide and Protein-Based Biomedical Materials. *Chem. Soc. Rev.* **2011**, *40* (9), 4563–4577.

(2) Hu, X.; Cebe, P.; Weiss, A. S.; Omenetto, F.; Kaplan, D. L. Protein-Based Composite Materials. *Mater. Today* **2012**, *15* (5), 208–215.

(3) Gagner, J. E.; Kim, W.; Chaikof, E. L. Designing Protein-Based Biomaterials for Medical Applications. *Acta Biomater.* **2014**, *10* (4), 1542–1557.

(4) Desai, M. S.; Lee, S. W. Protein-Based Functional Nanomaterial Design for Bioengineering Applications. *Wiley Interdiscip. Rev. Nanomedicine Nanobiotechnology* **2015**, *7* (1), 69–97.

(5) Tu, R. S.; Tirrell, M. Bottom-up Design of Biomimetic Assemblies. *Adv. Drug Deliv. Rev.* **2004**, *56* (11), 1537–1563.

(6) Luo, Q.; Hou, C.; Bai, Y.; Wang, R.; Liu, J. Protein Assembly: Versatile Approaches to Construct Highly Ordered Nanostructures. *Chem. Rev.* **2016**, *116* (22), 13571–13632.

(7) Yang, Y. J.; Holmberg, A. L.; Olsen, B. D. Artificially Engineered Protein Polymers. *Annu. Rev. Chem. Biomol. Eng.* **2017**, *8*, 549–575.

(8) Yeates, T. O. Geometric Principles for Designing Highly Symmetric Self-

Assembling Protein Nanomaterials. *Annu. Rev. Biophys.* **2017**, *46*, 23–42.

(9) Connor, R. E.; Tirrell, D. A. Non-Canonical Amino Acids in Protein Polymer Design. *Polym. Rev.* **2007**, *47* (1), 9–28.

(10) Johnson, J. A.; Lu, Y. Y.; Van Deventer, J. A.; Tirrell, D. A. Residue-Specific Incorporation of Non-Canonical Amino Acids into Proteins: Recent Developments and Applications. *Curr. Opin. Chem. Biol.* **2010**, *14* (6), 774–780.

(11) Witus, L. S.; Francis, M. B. Using Synthetically Modified Proteins to Make New Materials. *Acc. Chem. Res.* **2011**, *44* (9), 774–783.

(12) Spicer, C. D.; Pashuck, E. T.; Stevens, M. M. Achieving Controlled Biomolecule-Biomaterial Conjugation. *Chem. Rev.* **2018**, *118* (16), 7702–7743.

(13) Boutureira, O.; Bernardes, G. J. L. Advances in Chemical Protein Modification. *Chem. Rev.* **2015**, *115* (5), 2174–2195.

(14) Sunbul, M.; Yin, J. Site Specific Protein Labeling by Enzymatic Posttranslational Modification. *Org. Biomol. Chem.* **2009**, *7* (17), 3361–3371.

(15) Rashidian, M.; Dozier, J. K.; Distefano, M. D. Enzymatic Labeling of Proteins: Techniques and Approaches. *Bioconjug. Chem.* **2013**, *24* (8), 1277–1294.

(16) Link, A. J.; Mock, M. L.; Tirrell, D. A. Non-Canonical Amino Acids in Protein Engineering. *Curr. Opin. Biotechnol.* **2003**, *14* (6), 603–609.

(17) Wang, L.; Schultz, P. G. Expanding the Genetic Code. *Angew. Chemie - Int. Ed.* **2004**, *44* (1), 34–66.

(18) Liu, C. C.; Schultz, P. G. Adding New Chemistries to the Genetic Code. *Annu. Rev. Biochem.* **2010**, *79*, 413–444.

(19) Dumas, A.; Lercher, L.; Spicer, C. D.; Davis, B. G. Designing Logical Codon Reassignment-Expanding the Chemistry in Biology. *Chem. Sci.* **2015**, *6* (1), 50–69.

(20) Cowie, D. B.; Cohen, G. N. Biosynthesis by Escherichia Coli of Active Altered Proteins Containing Selenium Instead of Sulfur. *BBA - Biochim. Biophys. Acta* **1957**, *26* (2), 252–261.

(21) Strable, E.; Prasuhn, D. E.; Udit, A. K.; Brown, S.; Link, A. J.; Ngo, J. T.; Lander, G.; Quispe, J.; Potter, C. S.; Carragher, B.; Tirrell, D. A.; Finn, M. G. Unnatural Amino Acid Incorporation into Virus-like Particles. *Bioconjug. Chem.* **2008**, *19* (4), 866–875.

(22) Prasuhn, D. E.; Singh, P.; Strable, E.; Brown, S.; Manchester, M.; Finn, M. G. Plasma Clearance of Bacteriophage Qβ Particles as a Function of Surface Charge. *J. Am. Chem. Soc.* **2008**, *130* (4), 1328–1334.

(23) Stephanopoulos, N.; Tong, G. J.; Hsiao, S. C.; Francis, M. B. Dual-Surface Modified Virus Capsids for Targeted Delivery of Photodynamic Agents to Cancer Cells. *ACS Nano* **2010**, *4* (10), 6014–6020.

(24) Peeler, J. C.; Woodman, B. F.; Averick, S.; Miyake-Stoner, S. J.; Stokes, A. L.; Hess, K. R.; Matyjaszewski, K.; Mehl, R. A. Genetically Encoded Initiator for Polymer Growth from Proteins. *J. Am. Chem. Soc.* **2010**, *132* (39), 13575–13577.

(25) Teramoto, H.; Kojima, K. Production of Bombyx Mori Silk Fibroin Incorporated with Unnatural Amino Acids. *Biomacromolecules* **2014**, *15* (7), 2682–2690.

(26) Teramoto, H.; Amano, Y.; Iraha, F.; Kojima, K.; Ito, T.; Sakamoto, K. Genetic Code Expansion of the Silkworm Bombyx Mori to Functionalize Silk Fiber. *ACS Synth. Biol.* **2018**, *7* (3), 801–806.

(27) Carrico, I. S.; Maskarinec, S. A.; Heilshorn, S. C.; Mock, M. L.; Liu, J. C.; Nowatzki, P. J.; Franck, C.; Ravichandran, G.; Tirrell, D. A. Lithographic Patterning of Photoreactive Cell-Adhesive Proteins. *J. Am. Chem. Soc.* **2007**, *129* (16), 4874–4875.

(28) Nowatzki, P. J.; Franck, C.; Maskarinec, S. A.; Ravichandran, G.; Tirrell, D. A. Mechanically Tunable Thin Films of Photosensitive Artificial Proteins: Preparation and Characterization by Nanoindentation. *Macromolecules* **2008**, *41* (5), 1839–1845.

(29) Zhang, K.; Sugawara, A.; Tirrell, D. A. Generation of Surface-Bound Multicomponent Protein Gradients. *ChemBioChem* **2009**, *10* (16), 2617–2619.

(30) Hauf, M.; Richter, F.; Schneider, T.; Faidt, T.; Martins, B. M.; Baumann, T.; Durkin, P.; Dobbek, H.; Jacobs, K.; Möglich, A.; Budisa, N. Photoactivatable Mussel-Based Underwater Adhesive Proteins by an Expanded Genetic Code. *ChemBioChem* **2017**, *18* (18), 1819–1823.

(31) Torculas, M.; Medina, J.; Xue, W.; Hu, X. Protein-Based Bioelectronics. *ACS Biomater. Sci. Eng.* **2016**, *2* (8), 1211–1223.

(32) Panda, S. S.; Katz, H. E.; Tovar, J. D. Solid-State Electrical Applications of Protein and Peptide Based Nanomaterials. *Chem. Soc. Rev.* **2018**, *47* (10), 3640–3658.

(33) Wang, C.; Xia, K.; Zhang, Y.; Kaplan, D. L. Silk-Based Advanced Materials for Soft Electronics. *Acc. Chem. Res.* **2019**, *52* (10), 2916–2927.

(34) Dickerson, M. B.; Sandhage, K. H.; Naik, R. R. Protein- and Peptide-Directed Syntheses of Inorganic Materials. *Chem. Rev.* **2008**, *108* (11), 4935–4978.

(35) Hardy, J. G.; Scheibel, T. R. Composite Materials Based on Silk Proteins. *Prog. Polym. Sci.* **2010**, *35* (9), 1093–1115.

(36) Calvaresi, M.; Zerbetto, F. The Devil and Holy Water: Protein and Carbon Nanotube Hybrids. *Acc. Chem. Res.* **2013**, *46* (11), 2454–2463.

(37) Li, C.; Mezzenga, R. The Interplay between Carbon Nanomaterials and Amyloid Fibrils in Bio-Nanotechnology. *Nanoscale* **2013**, *5* (14), 6207–6218.

(38) Wang, F.; Yang, C.; Hu, X. Advanced Protein Composite Materials. *ACS Symp. Ser.* **2014**, *1175*, 177–208.

(39) Scheibel, T.; Parthasarathy, R.; Sawicki, G.; Lin, X. M.; Jaeger, H.; Lindquist, S. L. Conducting Nanowires Built by Controlled Self-Assembly of Amyloid Fibers and Selective Metal Deposition. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100* (8), 4527–4532.

(40) Sasso, L.; Suei, S.; Domigan, L.; Healy, J.; Nock, V.; Williams, M. A. K.; Gerrard, J. A. Versatile Multi-Functionalization of Protein Nanofibrils for Biosensor Applications. *Nanoscale* **2014**, *6* (3), 1629–1634.

(41) Lin, Y.; Xia, X.; Wang, M.; Wang, Q.; An, B.; Tao, H.; Xu, Q.; Omenetto, F.; Kaplan, D. L. Genetically Programmable Thermoresponsive Plasmonic Gold/Silk-Elastin Protein Core/Shell Nanoparticles. *Langmuir* **2014**, *30* (15), 4406–4414.

(42) Guo, C.; Hall, G. N.; Addison, J. B.; Yarger, J. L. Gold Nanoparticle-Doped Silk Film as Biocompatible SERS Substrate. *RSC Adv.* **2015**, *5* (3), 1937–1942.

(43) Das, C.; Krishnamoorthy, K. Flexible Microsupercapacitors Using Silk and Cotton Substrates. *ACS Appl. Mater. Interfaces* **2016**, *8* (43), 29504–29510.

(44) Peng, Z.; Peralta, M. D. R.; Cox, D. L.; Toney, M. D. Bottom-up Synthesis of Protein-Based Nanomaterials from Engineered β-Solenoid Proteins. *PLoS One* **2020**, *15* (2), 1–17.

(45) Flynn, C. E.; Lee, S. W.; Peelle, B. R.; Belcher, A. M. Viruses as Vehicles for Growth, Organization and Assembly of Materials. *Acta Mater.* **2003**, *51* (19), 5867–5880.

(46) Liu, Y.; Qi, N.; Song, T.; Jia, M.; Xia, Z.; Yuan, Z.; Yuan, W.; Zhang, K. Q.; Sun, B. Highly Flexible and Lightweight Organic Solar Cells on Biocompatible Silk Fibroin. *ACS Appl. Mater. Interfaces* **2014**, *6* (23), 20670–20675.

(47) Chen, A. Y.; Deng, Z.; Billings, A. N.; Seker, U. O. S.; Lu, M. Y.; Citorik, R. J.; Zakeri, B.; Lu, T. K. Synthesis and Patterning of Tunable Multiscale Materials with Engineered Cells. *Nat. Mater.* **2014**, *13* (5), 515–523.

(48) Nguyen, P. Q.; Botyanszki, Z.; Tay, P. K. R.; Joshi, N. S. Programmable Biofilm-Based Materials from Engineered Curli Nanofibres. *Nat. Commun.* **2014**, *5*, 1–10.

(49) Kalyoncu, E.; Ahan, R. E.; Olmez, T. T.; Safak Seker, U. O. Genetically

Encoded Conductive Protein Nanofibers Secreted by Engineered Cells. *RSC Adv.* **2017**, *7* (52), 32543–32551.

(50) Kang, M.; Jin, H. J. Electrically Conducting Electrospun Silk Membranes Fabricated by Adsorption of Carbon Nanotubes. *Colloid Polym. Sci.* **2007**, *285* (10), 1163–1167.

(51) Wang, E.; Desai, M. S.; Lee, S. W. Light-Controlled Graphene-Elastin Composite Hydrogel Actuators. *Nano Lett.* **2013**, *13* (6), 2826–2830.

(52) Wang, E.; Desai, M. S.; Heo, K.; Lee, S. W. Graphene-Based Materials Functionalized with Elastin-like Polypeptides. *Langmuir* **2014**, *30* (8), 2223–2229.

(53) Liang, B.; Fang, L.; Hu, Y.; Yang, G.; Zhu, Q.; Ye, X. Fabrication and Application of Flexible Graphene Silk Composite Film Electrodes Decorated with Spiky Pt Nanospheres. *Nanoscale* **2014**, *6* (8), 4264–4274.

(54) Annabi, N.; Shin, S. R.; Tamayol, A.; Miscuglio, M.; Bakooshli, M. A.; Assmann, A.; Mostafalu, P.; Sun, J. Y.; Mithieux, S.; Cheung, L.; Tang, X.; Weiss, A. S.; Khademhosseini, A. Highly Elastic and Conductive Human-Based Protein Hybrid Hydrogels. *Adv. Mater.* **2016**, *28* (1), 40–49.

(55) Jeon, J. W.; Cho, S. Y.; Jeong, Y. J.; Shin, D. S.; Kim, N. R.; Yun, Y. S.; Kim, H. T.; Choi, S. B.; Hong, W. G.; Kim, H. J.; Jin, H. J.; Kim, B. H. Pyroprotein-Based Electronic Textiles with High Stability. *Adv. Mater.* **2017**, *29* (6), 1–6.

(56) Liu, Y.; Tao, L. Q.; Wang, D. Y.; Zhang, T. Y.; Yang, Y.; Ren, T. L. Flexible, Highly Sensitive Pressure Sensor with a Wide Range Based on Graphene-Silk Network Structure. *Appl. Phys. Lett.* **2017**, *110* (12).

(57) Chi, N.; Wang, R. Electrospun Protein-CNT Composite Fibers and the Application in Fibroblast Stimulation. *Biochem. Biophys. Res. Commun.* **2018**, *504* (1), 211–217.

(58) Yin, Z.; Jian, M.; Wang, C.; Xia, K.; Liu, Z.; Wang, Q.; Zhang, M.; Liang, X.; Wang, H.; Liang, X.; Long, Y.; Yu, X.; Zhang, Y. Splash-Resistant and Light-Weight Silk-Sheathed Wires for Textile Electronics. *Nano Lett.* **2018**, *18* (11), 7085–7091.

(59) Cataldi, P.; Condurache, O.; Spirito, D.; Krahne, R.; Bayer, I. S.; Athanassiou, A.; Perotto, G. Keratin-Graphene Nanocomposite: Transformation of Waste Wool in Electronic Devices. *ACS Sustain. Chem. Eng.* **2019**, *7* (14), 12544–12551.

(60) Herland, A.; Björk, P.; Nilsson, K. P. R.; Olsson, J. D. M.; Åsberg, P.; Konradsson, P.; Hammarström, P.; Inganäs, O. Electroactive Luminescent Self-Assembled Bio-Organic Nanowires: Integration of Semiconducting Oligoelectrolytes within Amyloidogenic Proteins. *Adv. Mater.* **2005**, *17* (12), 1466–1471.

(61) Hamedi, M.; Herland, A.; Karlsson, R. H.; Lnganäs, O. Electrochemical Devices Made from Conducting Nanowire Networks Self-Assembled from Amyloid Fibrils and Alkoxysulfonate PEDOT. *Nano Lett.* **2008**, *8* (6), 1736–1740.

(62) Müller, C.; Jansson, R.; Elfwing, A.; Askarieh, G.; Karlsson, R.; Hamedi, M.; Rising, A.; Johansson, J.; Inganäs, O.; Hedhammar, M. Functionalisation of Recombinant Spider Silk with Conjugated Polyelectrolytes. *J. Mater. Chem.* **2011**, *21* (9), 2909–2915.

(63) Meier, C.; Lifincev, I.; Welland, M. E. Conducting Core-Shell Nanowires by Amyloid Nanofiber Templated Polymerization. *Biomacromolecules* **2015**, *16* (2), 558–563.

(64) Elfwing, A.; Bäcklund, F. G.; Musumeci, C.; Inganäs, O.; Solin, N. Protein Nanowires with Conductive Properties. *J. Mater. Chem. C* **2015**, *3* (25), 6499–6504.

(65) Bäcklund, F. G.; Elfwing, A.; Musumeci, C.; Ajjan, F.; Babenko, V.; Dzwolak, W.; Solin, N.; Inganäs, O. Conducting Microhelices from Self-Assembly of Protein Fibrils. *Soft Matter* **2017**, *13* (25), 4412–4417.

(66) Tsukada, S.; Nakashima, H.; Torimitsu, K. Conductive Polymer Combined Silk Fiber Bundle for Bioelectrical Signal Recording. *PLoS One* **2012**, *7* (4).

(67) Cai, N.; Wong, C. C.; Gong, Y. X.; Tan, S. C. W.; Chan, V.; Liao, K. Modulating Cell Adhesion Dynamics on Carbon Nanotube Monolayer Engineered with Extracellular Matrix Proteins. *ACS Appl. Mater. Interfaces* **2010**, *2* (4), 1038–1047.

(68) Namgung, S.; Kim, T.; Baik, K. Y.; Lee, M.; Nam, J. M.; Hong, S. Fibronectin-Carbon-Nanotube Hybrid Nanostructures for Controlled Cell Growth. *Small* **2011**, *7* (1), 56–61.

(69) Bongo, M.; Winther-Jensen, O.; Himmelberger, S.; Strakosas, X.; Ramuz, M.; Hama, A.; Stavrinidou, E.; Malliaras, G. G.; Salleo, A.; Winther-Jensen, B.; Owens, R. M. PEDOT:Gelatin Composites Mediate Brain Endothelial Cell Adhesion. *J. Mater. Chem. B* **2013**, *1* (31), 3860–3867.

(70) Kim, T.; Sridharan, I.; Zhu, B.; Orgel, J.; Wang, R. Effect of CNT on Collagen Fiber Structure, Stiffness Assembly Kinetics and Stem Cell Differentiation. *Mater. Sci. Eng. C* **2015**, *49*, 281–289.

(71) Radunovic, M.; De Colli, M.; De Marco, P.; Di Nisio, C.; Fontana, A.; Piattelli, A.; Cataldi, A.; Zara, S. Graphene Oxide Enrichment of Collagen Membranes Improves DPSCs Differentiation and Controls Inflammation Occurrence. *J. Biomed. Mater. Res. - Part A* **2017**, *105* (8), 2312–2320.

(72) Ryan, A. J.; Kearney, C. J.; Shen, N.; Khan, U.; Kelly, A. G.; Probst, C.; Brauchle, E.; Biccai, S.; Garciarena, C. D.; Vega-Mayoral, V.; Loskill, P.; Kerrigan, S. W.; Kelly, D. J.; Schenke-Layland, K.; Coleman, J. N.;

O'Brien, F. J. Electroconductive Biohybrid Collagen/Pristine Graphene Composite Biomaterials with Enhanced Biological Activity. *Adv. Mater.* **2018**, *30* (15), 1–8.

(73) Cui, X.; Lee, V. A.; Raphael, Y.; Wiler, J. A.; Hetke, J. F.; Anderson, D. J.; Martin, D. C. Surface Modification of Neural Recording Electrodes with Conducting Polymer/Biomolecule Blends. *J. Biomed. Mater. Res.* **2001**, *56* (2), 261–272.

(74) Zhang, L.; Stauffer, W. R.; Jane, E. P.; Sammak, P. J.; Cui, X. T. Enhanced Differentiation of Embryonic and Neural Stem Cells to Neuronal Fates on Laminin Peptides Doped Polypyrrole. *Macromol. Biosci.* **2010**, *10* (12), 1456–1464.

(75) Povlich, L. K.; Cho, J. C.; Leach, M. K.; Corey, J. M.; Kim, J.; Martin, D. C. Synthesis, Copolymerization and Peptide-Modification of Carboxylic Acid-Functionalized 3,4-Ethylenedioxythiophene (EDOTacid) for Neural Electrode Interfaces. *Biochim. Biophys. Acta - Gen. Subj.* **2013**, *1830* (9), 4288–4293.

(76) Maione, S.; Gil, A. M.; Fabregat, G.; Del Valle, L. J.; Triguero, J.; Laurent, A.; Jacquemin, D.; Estrany, F.; Jiménez, A. I.; Zanuy, D.; Cativiela, C.; Alemán, C. Electroactive Polymer-Peptide Conjugates for Adhesive Biointerfaces. *Biomater. Sci.* **2015**, *3* (10), 1395–1405.

(77) Cui, X.; Wiler, J.; Dzaman, M.; Altschuler, R. A.; Martin, D. C. In Vivo Studies of Polypyrrole/Peptide Coated Neural Probes. *Biomaterials* **2003**, *24* (5), 777–787.

(78) Zhu, B.; Luo, S. C.; Zhao, H.; Lin, H. A.; Sekine, J.; Nakao, A.; Chen, C.; Yamashita, Y.; Yu, H. H. Large Enhancement in Neurite Outgrowth on a Cell Membrane-Mimicking Conducting Polymer. *Nat. Commun.* **2014**, *5* (May), 1–9.

(79) Khatayevich, D.; Page, T.; Gresswell, C.; Hayamizu, Y.; Grady, W.; Sarikaya, M. Selective Detection of Target Proteins by Peptide-Enabled Graphene Biosensor. *Small* **2014**, *10* (8), 1505–1513.

(80) Tara Bahadur, K. C.; Tada, S.; Zhu, L.; Uzawa, T.; Minagawa, N.; Luo, S. C.; Zhao, H.; Yu, H. H.; Aigaki, T.; Ito, Y. In Vitro Selection of Electrochemical Peptide Probes Using Bioorthogonal TRNA for Influenza Virus Detection. *Chem. Commun.* **2018**, *54* (41), 5201–5204.

(81) Lee, K.; Yoo, Y. K.; Chae, M. S.; Hwang, K. S.; Lee, J.; Kim, H.; Hur, D.; Lee, J. H. Highly Selective Reduced Graphene Oxide (RGO) Sensor Based on a Peptide Aptamer Receptor for Detecting Explosives. *Sci. Rep.* **2019**, *9* (1), 1–9.

(82) Li, W.; Gao, Y.; Zhang, J.; Wang, X.; Yin, F.; Li, Z.; Zhang, M. Universal DNA Detection Realized by Peptide Based Carbon Nanotube Biosensors. *Nanoscale Adv.* **2020**, *2* (2), 717–723.

(83) Lebaron, R. G.; Athanasiou, K. A. Extracellular Matrix Cell Adhesion Peptides: Functional Applications in Orthopedic Materials. *Tissue Eng.* **2000**, *6* (2), 85–103.

(84) Waugh, D. S. Making the Most of Affinity Tags. *Trends Biotechnol.* **2005**, *23* (6), 316–320.

(85) Zhao, X.; Li, G.; Liang, S. Several Affinity Tags Commonly Used in Chromatographic Purification. *J. Anal. Methods Chem.* **2013**, *2013* (Table 1).

C h a p t e r 2

INCORPORATION OF EDOT-ALA INTO PROTEINS

2.1. Abstract

Poly(3,4-ethylenedioxythiophene) is a π-conjugated polymer that exhibits high electrical conductivity, electrochemical stability, and biocompatibility, and therefore, is applied in various fields including bioelectronics. We envisioned that genetic incorporation of the monomer precursor of PEDOT will facilitate the fabrication of protein-based conductive materials. To this end, we sought to genetically incorporate an ncAA bearing EDOT group by harnessing *E. coli* protein biosynthesis machinery. We prepared an amino acid EDOT-Ala in three steps of organic reaction. To incorporate EDOT-Ala into proteins, we examined the activity of known aaRS variants such as an endogenous *E. coli* TrpRS, yeast PheRS(T415G), and *Mm*PylRS(N346A/C348A), which were previously reported to activate bulky amino acids. However, we were unable to find any aaRS variant that recognizes EDOT-Ala. We then performed evolution of *Mm*PylRS(N346A/C348A) and *p*-NFRS to alter their specificity towards EDOT-Ala via two approaches: cell viability-based evolution and phage-assisted evolution. Unfortunately, neither of these strategies yielded an aaRS that incorporates EDOT-Ala. We believe that a broader search of aaRS and further efforts on evolution should be able to discover an aaRS variant that is able to incorporate EDOT-Ala.

2.2. Introduction

Protein-based materials are attractive candidates for applications in bioelectronics. In order to introduce electrical conductivity into protein-based materials, electrical conductors such as metals, carbon-based materials, and conductive polymers are often combined with proteins to create conductive biohybrid materials.[1–6] We envisioned that genetic incorporation of an electroactive functional group into proteins could be used to form a protein–conductive polymer conjugate and will provide a new strategy to produce protein-based conductive materials. To this end, we prepared a non-canonical amino acid (ncAA) bearing 3,4-ethylenedioxythiophene (EDOT) and performed incorporation into proteins. EDOT is the monomer precursor of the conductive polymer PEDOT, which is recognized as the most successful conductive polymer in both fundamental research and practical applications.[7–12] Due to its attractive features such as high conductivity, electrochemical stability, and biocompatibility, PEDOT has been used for various bioelectronics applications such as tissue engineering,[13–18] neural probes,[19–22] biosensors,[9,23–27] and drug delivery.[28–30]

The key requirement for incorporation of an ncAA is an aminoacyl-tRNA synthetase (aaRS) that can recognize the target ncAA and ligate it to the cognate tRNA.[31–33] Because of this requirement, ncAAs are usually designed as analogues of canonical amino acids or previously reported ncAAs, and the aaRS variants for the known amino acids are used as an initial set of aaRSs to evaluate their activities for the target ncAA. When the tested aaRS variants show little or no activity for the target ncAA, new aaRS variants should be prepared by mutagenesis. Directed evolution[34–36] and rational and computational design[37–42] are effective approaches to alter the specificity of aaRS towards target ncAAs.

The EDOT group consists of bulky functional groups such as a thiophene ring and a dioxane ring. Therefore, aaRS with a large binding pocket should be used to incorporate the ncAA bearing an EDOT group. A major class of bulky amino acids is derivatives of phenylalanine (Phe). Besides the canonical amino acid Tyrosine (Tyr), a large number of substituted Phe analogues have been successfully incorporated by phenylalanyl t-RNA

synthetase (PheRS),[37,43] tyrosyl t-RNA synthetase (TyrRS),[44] and pyrrolysyl t-RNA synthetase (PylRS)[44,45] derived from different organisms.

Here, we designed an amino acid with the EDOT group (designated EDOT-Ala) as an analogue of aromatic canonical amino acids and sought to incorporate the ncAA using aaRS variants with large binding pockets. We performed evolution of the aaRS by two different evolution approaches: cell viability-based method[34,35] and phage-assisted method.[36] Although EDOT-Ala was not incorporated by evolved aaRS variants, the results presented in this chapter provide valuable insights into incorporation of EDOT-functionalized amino acids into proteins.

2.3. Results and Discussion

2.3.1. Synthesis of an amino acid EDOT-Ala

We synthesized an ncAA bearing the EDOT group termed EDOT-Ala (Figure 2.1). In EDOT-Ala, the dioxane ring of EDOT lies one methylene group apart from the α-carbon, which is the same position as the indole group of Trp and the phenol group of Tyr and Phe (Figure 2.1). Based on the structural similarity of EDOT-Ala and these large natural amino acids, we hypothesized that EDOT-Ala is recognized by endogenous or engineered aaRSs for these amino acids. Kwon *et al.* have demonstrated that an engineered yeast PheRS (*y*PheRS(T415G)) efficiently incorporates Trp analogues at the amber stop codon when co-expressed with a mutant yeast phenylalanine amber suppressor tRNA ($ytRNA^{Phe}_{CUA\_UG}$).[43] Moreover, the Phe analogues with various substituents have also been incorporated using a TyrRS/tRNA pair from the archaebacteria *Methanococcus jannaschii* (*Mj*TyrRS/$tRNA^{Tyr}_{CUA}$) and PylRS/tRNA pairs from *Methanosarcina mazei* (*Mm*PylRS/$tRNA^{Pyl}_{CUA}$) and *Methanosarcina barkeri* (*Mb*PylRS/$tRNA^{Pyl}_{CUA}$).[44] In addition to the existence of aaRS variants that activate bulky amino acids, previous studies have proven that active aaRS for target ncAAs can be produced by directed evolution[34–36] and rational and computational design.[37–42] Therefore, we envisioned that EDOT-Ala should be incorporated either by one of the reported aaRS variants or by their engineered variants.

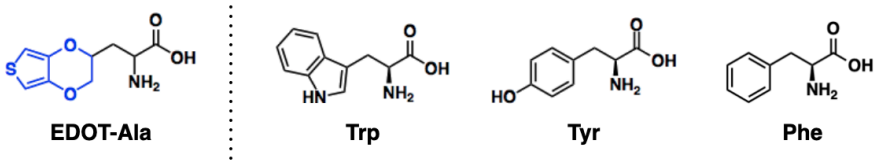


Figure 2.1. Chemical structures of EDOT-Ala and natural amino acids Trp, Tyr, and Phe.

The synthetic scheme of EDOT-Ala is shown in Figure 2.2. EDOT-Ala was synthesized via alkylation of *N*-(diphenylmethylene)glycine ethyl ester (1) and subsequent deprotection of the Schiff base under acidic conditions.[46] A diphenylmethylene protecting group was employed because the deprotection proceeds rapidly at room temperature.[46] The choice of the protecting group was critical for the synthesis of EDOT-Ala, because the EDOT group is susceptible to harsh acidic conditions. The synthesis was completed in three steps from the commercial compounds, and EDOT-Ala was obtained as a racemic mixture with a relatively good yield. Because previous studies have demonstrated the successful incorporation of ncAAs regardless of the use of a racemic mixture,[47,48] we proceeded to the incorporation experiment without isolation of the L-type of EDOT-Ala. The details of the synthesis and NMR data are provided in Section 2.4.2 and 2.6.

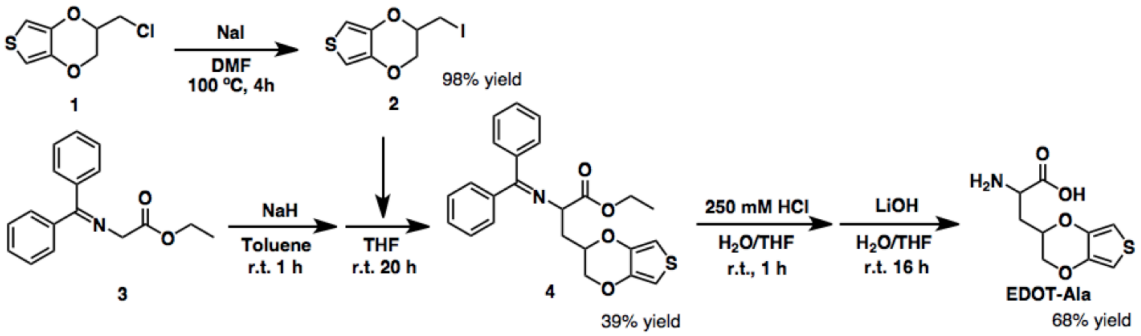


Figure 2.2. Synthesis of EDOT-Ala.

2.3.2. Incorporation of EDOT-Ala Using Known aaRS Variants

Natural TrpRS has a large binding pocket to accommodate the bulky amino acid Trp, which carries an indole group in the side chain. Therefore, we anticipated that endogenous

E. coli TrpRS should be able to recognize EDOT-Ala. To test this possibility, we performed incorporation of EDOT-Ala via the global replacement approach. We chose an artificial coiled coil protein SYNZYP3[49] as a test protein, which carries no Trp residue in the main sequence. We prepared a plasmid encoding a single Trp residue at position 4, a copy of SYNZYP3, and *N*- and *C*-terminal 6xHisTags for affinity purification (Figure 2.3a). The test protein was placed under the control of bacteriophage T5 promoter to enable induction by isopropyl- β -Dthiogalactopyranoside (IPTG).

We expressed SYNZIP3 in a Trp-auxotrophic *E. coli* strain derived from DH10B. After culturing the cells to the mid-log phase in M9 minimal media supplemented with 20 canonical amino acids, the media was shifted to M9 media supplemented with 2 mM EDOT-Ala and 19 amino acids without Trp, and expression was induced for 5 hours by addition of IPTG. Then cells were lysed, and the protein was purified by Ni-NTA chromatography and subjected to SDS-PAGE. Unfortunately, no bands corresponding to SYNZIP3 appeared for the sample expressed in the presence of EDOT-Ala, whereas the positive control expressed in the presence of Trp showed the bands for SYNZIP3.

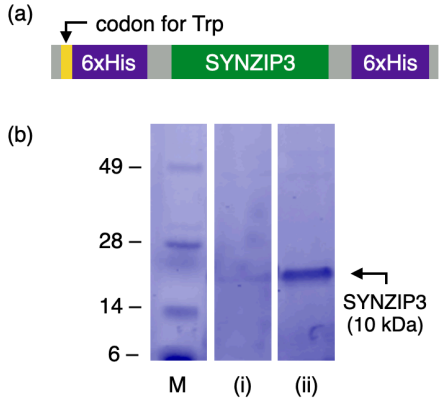


Figure 2.3. Incorporation of EDOT-Ala by endogenous *E. coli* TrpRS. (a) Sequence of SYNZIP3 with a single Trp codon. (b) Coomassie-stained SDS-PAGE for purified proteins. Proteins were expressed in M9 minimal media supplemented with (i) 19 canonical amino acids (no Trp) with EDOT-Ala and (ii) 20 canonical amino acids. M: marker.

Next, we investigated the activity of an engineered yeast PheRS (yPheRS(T415G)) that was previously reported to activate Phe and Trp analogues (p-bromophenylalanine (pBrF), 6-chlorotryptophan (6ClW), 6-bromotryptophan (6BrW), and benzothienylalanine (BT)).[43] The pair of yPheRS(T415G) and a mutant yeast phenylalanine amber suppressor tRNA ($ytRNA^{Phe}_{CUA\_UG}$) is orthogonal to the endogenous *E. coli* translational apparatus, and introduction of the pair into an *E. coli* expression host allowed site-specific incorporation of the analogues at amber stop codons with high fidelity.[43] Therefore, we anticipated that the yPheRS(T415G)/$ytRNA^{Phe}_{CUA\_UG}$ pair could be used to incorporate EDOT-Ala.

To study this possibility, we expressed recombinant murine dihydrofolate reductase (mDHFR) as a test protein following a previously reported procedure.[43] We used a Phe/Trp double auxotrophic strain of *E. coli* as a host cell and performed expression in M9 minimal media supplemented with 18 amino acids (40 mg/L), low concentrations of Phe (0.03 mM) and Trp (0.01 mM), and 3 mM EDOT-Ala. After expression, cells were lysed and subjected to SDS-PAGE. We observed the protein bands corresponding to the mDHFR in all induced samples. (Figure 2.4). However, the sample expressed in the presence of EDOT-Ala only showed a faint band comparable to the negative control expressed in the presence of 5-bromotryptophan (5BrW). Because the previous study indicated that 5BrW is not recognized by the yPheRS(T415G) and that Phe predominantly occupies the amber codon under this experimental condition, it is unlikely that yPheRS(T415G) has high activity for EDOT-Ala. Although there is a possibility that some portion of the expressed protein contained EDOT-Ala, we did not further investigate the product by mass spectrometry in order to conserve the amino acid from a large-scale expression and proceeded to explore other aaRS variants.

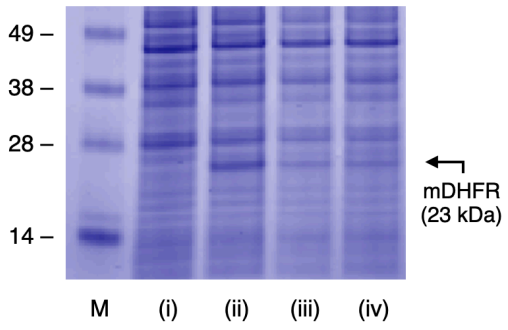


Figure 2.4. Incorporation of EDOT-Ala by the yPheRS(T415G) into mDHFR. Coomassie-stained SDS-PAGE of (i) uninduced culture and induced cultures in M9 minimal media supplemented with (ii) 20 canonical amino acids (40 mg/L), (iii) 18 amino acids (40 mg/L), Phe (0.03 mM), Trp (0.01 mM), 5-BrTrp (1 mM), and (iv) 18 amino acids (40 mg/L), Phe (0.03 mM), Trp (0.01 mM), and EDOT-Ala (3 mM). M: marker.

The variants of aaRS from methanogenic archaea, *Mj*TyrRS, *Mm*PylRS, and *Mb*PylRS have been used for site-specific incorporation of a variety of amino acids via amber suppression. In particular, *Mm*PylRS(N346A/C348A) developed by Liu *et al.* facilitated efficient incorporation of various Phe derivatives in *E. coli* host cells.[50–53] Therefore, we investigated the activity of *Mm*PylRS(N346A/C348A) to incorporate EDOT-Ala.

We chose green fluorescent protein (GFP) as a test protein, which enables evaluation of the aaRS activity based on cellular fluorescence. To avoid potential complications arising from the disturbance of fluorescent activity of GFP by incorporation of EDOT-Ala, we placed an amber codon upstream of the main GFP sequence (GFP-S4am; Figure 2.5a). Incorporation was performed in the BL21(DE3) strain of *E. coli* cells expressing the *Mm*PylRS(N346A/C348A)/$tRNA^{Pyl}$ pair. The cells were cultured in 2xYT media supplemented with 2 mM EDOT-Ala, induced for 18 hours, and subjected to flow cytometry. For comparison, we also prepared the uninduced cells and induced cells without ncAA or with one of the Phe analogues previously reported to be recognized by *Mm*PylRS(N346A/C348A).[51] In the previous study, three Phe analogues, 3-trifluoromethylphenylalanine (3-CF_3F), 3-chlorophenylalanine (3-ClF), and 3-cyanophenylalanine (3-CNF), yielded 39, 16, and 7 mg of a superfolder GFP (sfGFP) per one litter scale of *E. coli* culture in modified glycerol minimal media.[51] In our flow cytometry experiment, the induced cells cultured with EDOT-Ala showed a significantly larger median fluorescence than uninduced cells (Figure 2.5b). However, the distribution of the cell population was very similar to those of the cells cultured without ncAA or with 3-cyanophenylalanine (3-CNF). These results imply that the observed fluorescence from the cells cultured with EDOT-Ala could originate from background fluorescence due to the incorporation of canonical amino acids such as Phe and Tyr.

To avoid the mis-incorporation of Phe and Tyr, we performed GFP expression in M9 minimal media without supplementing these amino acids. The resulting cells showed significantly less background fluorescence than the cells expressed in 2xYT (Figure 2.5c). However, the median fluorescence in the cells cultured with EDOT-Ala was only slightly larger than those cultured without ncAA. When we analyzed the purified protein by mass spectrometry, the GFP expressed in the presence of EDOT-Ala showed a predominant peak corresponding to the incorporation of Tyr at the amber codon, and no peak for GFP with EDOT-Ala was observed (*data not shown*). This is attributed to the endogenous production of Tyr by the BL21(DE3) strain of *E. coli* cells. Although we attempted to express GFP using a Phe/Tyr double auxotrophic strain of *E. coli*, the cell growth in M9 minimal media without Phe and Tyr was significantly hindered, and the protein was not expressed well.

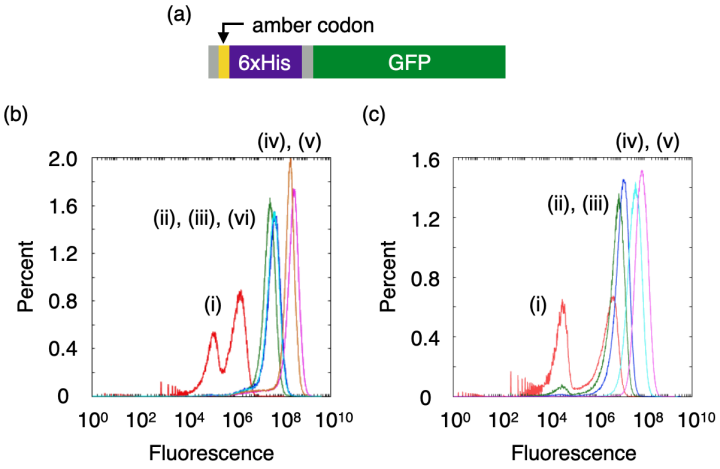


Figure 2.5. Incorporation of EDOT-Ala by *Mm*PylRS(N346A/C348A). (a) Sequence of the GFP reporter protein (GFP-S4am). (b) Flow cytometry analysis of GFP expression in *E. coli* host cells cultured in 2xYT media. (i) Uninduced (red), (ii) no ncAA (green), (iii) EDOT-Ala (blue), (iv) $3CF_3F$ (magenta), (v) 3-ClF (orange), and (vi) 3-CNF (aqua). (c) Flow cytometry analysis of GFP expression in *E. coli* host cells cultured in M9 minimal media supplemented with 18 canonical amino acids without Phe and Tyr. (i) Uninduced (red), (ii) no ncAA (green), (iii) EDOT-Ala (blue), (iv) $3CF_3F$ (magenta), and (v) 3-CNF (aqua).

2.3.3. Evolution of aaRS by Cell Viability-Based Selection

Because we were unable to incorporate EDOT-Ala using the known aaRSs such as an endogenous *E. coli* TrpRS, yeast PheRS(T415G), and *Mm*PylRS(N346A/C348A), we next sought to develop aaRS by directed evolution. Protein evolution has been proven to be a powerful tool to create aaRS variants for target ncAAs.[34–36] One of the major approaches utilizes positive and negative selections based on cell viability upon expression of an antibiotic resistance gene and a bacterial toxin.[34,35] The positive selection plasmid encodes a chloramphenicol acetyltransferase (CAT) gene under constitutive promoter, which confers the resistance to an antibiotic chloramphenicol (Figure 2.6a).[35] Because an amber codon is placed in the middle of the CAT gene, the plasmid confers resistance to chloramphenicol only when an amino acid is incorporated at the amber codon. Thus, the plasmid allows viability-based screening of the cells carrying the aaRS gene that activates the target ncAA. The negative selection plasmid encodes a bacterial toxin barnase with two amber codons (Figure 2.6b).[34] When the aaRS has off-target activity towards canonical amino acids, barnase is expressed and exerts lethal effect to the host *E. coli* cells. Following these established screening methods, we sought to create an aaRS that is capable of incorporating EDOT-Ala.

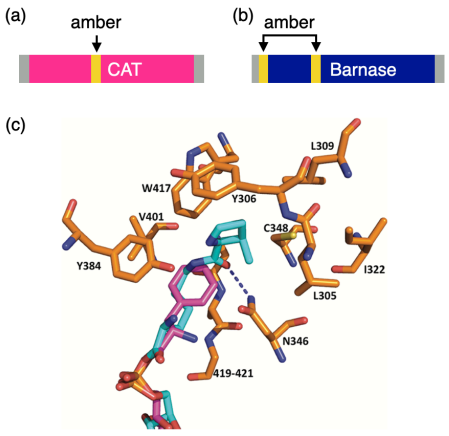


Figure 2.6. Evolution of *Mm*PylRS(N346A/C348A) for incorporation of EDOT-Ala. Plasmids used for (a) positive and (b) negative selection of active aaRS variants. (c) Crystal structure of the active site of *Mm*PylRS. Adapted from ref. 50.

We chose *Mm*PylRS(N346A/C348A) as a template for evolution because of its ability to incorporate various Phe analogues.[50–53] The plasmid toolset was readily available for evolution of *Mm*PylRS variants, and no laborious cloning was required to customize the plasmids to perform experiments. Although this template itself did not show any activity towards EDOT-Ala (see *2.3.2. Incorporation of EDOT-Ala using known aaRS variants*), we reasoned that mutations at the active site would alter the specificity to enable incorporation of EDOT-Ala.

To this end, we constructed a site-saturation mutagenesis library of *Mm*PylRS(N346A/C348A) targeting the five residues at the amino acid binding pocket: L305, Y306, L309, Y384, and W417 (Figure 2.6). These residues were previously targeted to alter the specificity of *Mm*PylRS variants.[54,55] The plasmid library was transformed into *E. coli* cells harboring the positive selection plasmid encoding chloramphenicol acetyltransferase (CAT) gene with an amber codon at position 112. The transformed cells were grown overnight on LB agar plates supplemented with chloramphenicol and EDOT-Ala. Over 100 colonies were formed in 72 hours, and we collected 62 colonies to extract plasmids. The plasmids were transformed into *E. coli* cells harboring the negative selection plasmid encoding a toxic protein barnase with amber codons at positions 2 and 44. When the cells were grown on LB agar plates without EDOT-Ala, 8 colonies appeared after 18 hours. We collected the colonies, extracted plasmids, and subjected them to another cycle of positive selection. We did not observe any colony formation in 72 hours. Therefore, it is likely that the survived population in the first positive selection was due to the mutation of amber codon to sense codon in CAT gene or degradation of chloramphenicol over the long culture time. The other possible explanation is that the aaRS variants have low off-target activity and the cells were able to survive in the negative selection. Future steps in this experiment would be the optimization of selection stringency based on concentration of EDOT-Ala and number of amber codons in positive and negative selection plasmids as well as targeting a broader range of residues in *Mm*PylRS(N346A/C348A).

2.3.4. Evolution of aaRS by Phage-Assisted Approach

Recently, a novel evolution technique has been developed by Liu *et al.* using an engineered M13 bacteriophage.[36,56] This technique, termed phage-assisted continuous evolution (PACE) or non-continuous evolution (PANCE), enabled the production of highly active and selective aaRSs for their target ncAAs.[36] We reasoned that the technique could be used to create an aaRS variant that recognizes EDOT-Ala.

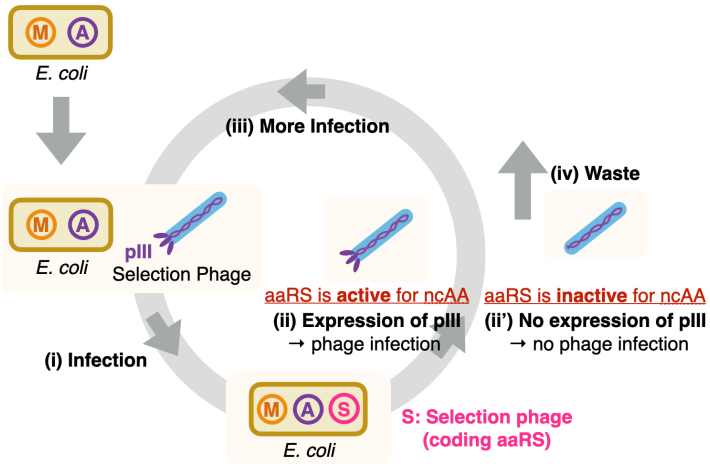


Figure 2.7. Phage-assisted protein evolution. M: mutagenesis plasmid, A: accessory plasmid, S: selection phage.

The general scheme of phage-assisted evolution is described in Figure 2.7. Phage-assisted evolution utilizes propagation of an engineered M13 bacteriophage termed selection phage (SP). In each evolution cycle, SP is allowed to propagate based on the activity of aaRS and diluted for the next cycle. SP carries the gene of aaRS replacing a phage gene *gIII*, which encodes an essential protein (pIII) required for the progeny phage infectivity. Due to the lack of *gIII*, SP does not propagate by itself. Instead, *E. coli* cells harboring an accessory plasmid coding *gIII* are used to assist the propagation of SP. To link the activity of aaRS and phage propagation, expression of pIII is regulated by the event of amber suppression. In low stringency strategy, amber codons are placed in T7 RNA polymerase (T7 RNAP) coded on the complementary plasmid (CP), and *gIII* is placed under control of T7 promoter on the accessory plasmid (AP) (Figure 2.8a). This way, expression of one copy of T7

RNAP can produce multiple pIII to impart infectivity to SP. In the high stringency strategy, amber codons are placed directly in *gIII* encoded on AP (Figure 2.8b). Therefore, expression of a copy of pIII requires multiple events of amber suppression to provide high stringency in the selection.

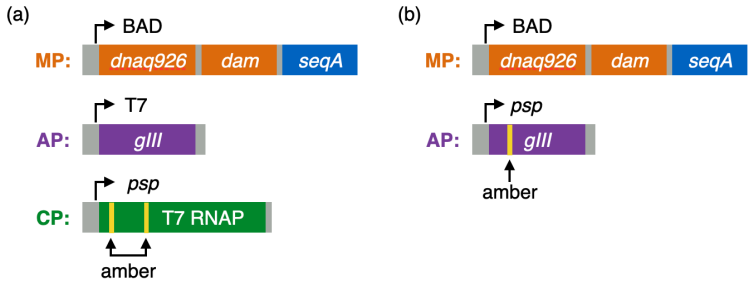


Figure 2.8. Design of plasmids used for positive selections in phage-assisted evolution. (a) Low stringency selection and (b) high stringency selection. MP: mutagenesis plasmid, AP: accessory plasmid, CP: complementary plasmid. *Psp*: phage-shock promoter.

Library diversity of aaRS is generated *in vivo* using the mutagenesis plasmid (MP).[36,57] MP encodes *dnaQ926* (dominant-negative variant of the *E. coli* DNA Pol III proofreading domain), *dam* (DNA adenine methylase that impairs mismatch repair), and *seqA* (negative regulator of DNA replication), which elevate the error rate during DNA replication with minimal background mutations under the uninduced condition. By combining MP-induced *in vivo* mutagenesis and SP-based selection for active aaRS, aaRS variants with high activity towards the target ncAA are enriched without laborious cloning steps between evolution cycles.[36]

Negative selection relies on the production of a dominant-negative variant of pIII termed pIII-neg (Figure 2.9a).[58] Because pIII-neg impairs the infectivity of the emergent phage, propagation of SP is hindered. To link the activity of aaRS with expression of pIII-neg, the gene of pIII-neg is placed under control of T7 promoter on CP, and T7 RNAP with amber codons is coded on AP (Figure 2.9b).

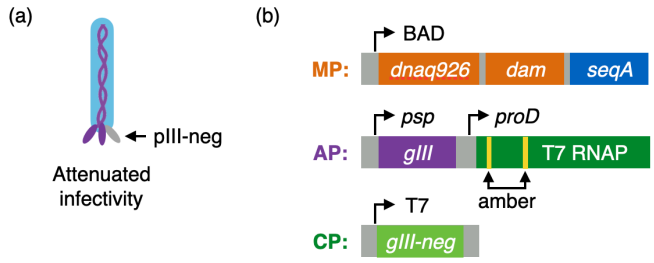
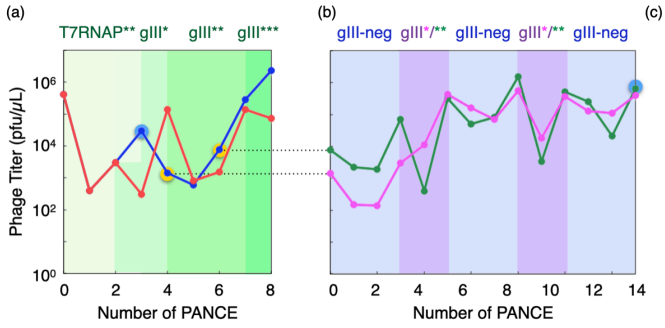


Figure 2.9. Negative selection in phage-assisted evolution. (a) SP carrying pIII-neg shows attenuated infectivity. (b) Design of plasmids for negative selection.

To develop an aaRS for EDOT-Ala, we performed phage-assisted evolution using a mutant *Mj*TyrRS (*p*-NFRS) as a template. *p*-NFRS was originally evolved from *Mj*TyrRS for incorporation of *p*-nitrophenylalanine (*p*-NF),[59] but has activity towards Phe and *p*-iodophenylalanine (*p*-IF).[36] A previous study demonstrated that phage-assisted evolution of *p*-NFRS significantly enhanced its activity and specificity towards *p*-IF.[36] Because *Mj*TyrRS mutants are widely used to incorporate various Phe derivatives, we reasoned that phage-assisted evolution of *p*-NFRS can produce an aaRS that recognizes EDOT-Ala.

We performed 8 positive selection cycles in 2xYT media supplemented with EDOT-Ala, increasing stringency after several cycles (Phase 1) (Figure 2.10a). The phage titer significantly decreased after the initial cycle, which is attributed to the limited diversity and low activity of aaRS library. The phage titer increased over the subsequent positive selection cycles and decreased when the selection stringency was increased. Selection was continued to Phase 2 alternating three negative selection cycles and two positive selection cycles (Figure 2.10b). We observed a decreased phage titer after switching to positive or negative selections. Six plaques were sequenced at each cycle to trace the evolution of *p*-NFRS. At the early stage of the evolution, we observed no mutation or several mutations in each sample, and none of the samples shared common mutations (Figure 2.10c). At the late stage of the evolution, we found two converged mutations in all six samples: L110F and A167T. These residues are located in close proximity to the binding pocket of *Mj*TyrRS (Figure 2.11a), and A167 was previously reported to affect the activity of *Mj*TyrRS.[42,60,61]



(c)

	Batch 1, Cycle 3	Batch 1, Cycle 14
1	No mutation	L110F, L121F, A167T
2	G238R	L110F, A167T, V236G
3	K51L, V168I	D87Y, L110F, A167T, N239Y
4	No mutation	I62V, N74K, L110F, A167T
5	K40R	Y102N, L110F, A167T, L198S, A233V
6	No mutation	E3D, L110F, L127F, P152L, A167T

Figure 2.10. Evolution of *p*-NFRS for EDOT-Ala in 2xYT. (a, b) Change in the phage titer in (a) Phase 1: eight positive selection cycles of batch 1 (blue) and batch 2 (red), and (b) Phase 2: fourteen positive and negative selection cycles of batch 1 (green) and batch 2 (magenta). Three negative cycles and two positive cycles were alternated. The type of APs used for the selection are indicated above the graphs. (* indicates the number of amber codons.) Starting SP solutions for phase 2 are indicated by dot lines. (c) Mutations in aaRS found in cycle 3 of batch 1 in phase1 (left) and cycle 14 of batch 1 in phase 2.

To determine whether EDOT-Ala is activated by this mutant, we cloned *p*-NFRS(L110F/A167T) into a plasmid and co-expressed in *E. coli* with GFP bearing an amber codon at position 4 (Figure 2.11b). After induction for 16 hours, the cells cultured in 2xYT with EDOT-Ala exhibited strong fluorescence. However, the cells cultured without ncAA also showed fluorescence at the comparable level. This result indicates that *p*-NFRS(L110F/A167T) has high off-target activity towards canonical amino acids.

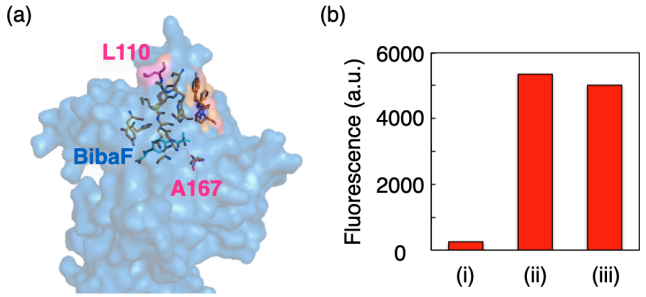
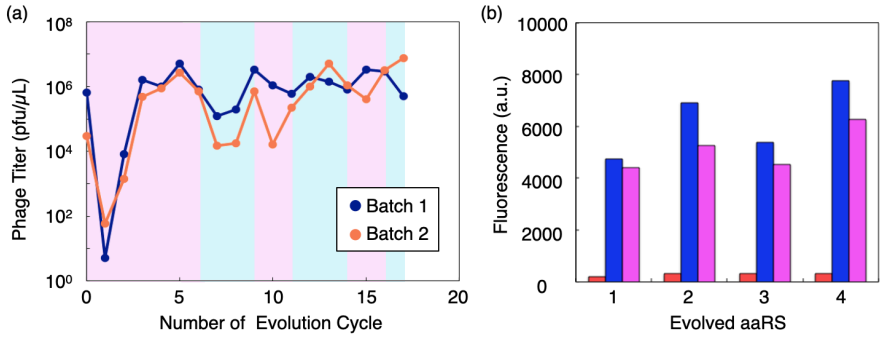


Figure 2.11. Crystal structure of *Mj*TyrRS mutant (PDB: 4PBR) highlighting the active site residues and its substrate 4-(2'-bromoisobutyramido)-phenylalanine (BibaF). (b) Fluorescence

measurement of the cells co-expressing *p*-NFRS(L110F/A167T) and GFP-4am. (i) Uninduced, (ii) no ncAA, (iii) EDOT-Ala.

To minimize the possibility that the aaRSs gain the off-target activity towards canonical amino acids, we proceeded the evolution in M9 minimal media supplemented with 17 amino acids (no Phe, Tyr, Trp) and EDOT-Ala. We performed a total of 17 evolutions, starting with 6 successive positive selection cycles and then alternating 3 negative selection cycles and 2 positive selection cycles (Figure 2.12a). Over the course of evolution, we observed the convergence in mutations in the sequence of aaRS (Figure 2.12c). We chose four aaRS variants from different cycles of evolution and analyzed their activity by co-expression with GFP bearing an amber codon at position 4. Unfortunately, all the variants showed strong background fluorescence in the absence of EDOT-Ala (Figure 2.12b). The mass peaks of all the purified proteins corresponded with the incorporation of Tyr at the amber codon. We also performed a phage-assisted evolution using another template *Mm*PylRS(N346A/N348C), but the evolved mutants acquired off-target activity towards canonical amino acids and did not incorporate EDOT-Ala (*data not shown*).

The major challenge in the phage-assisted evolution for EDOT-Ala was the evolution of aaRS to acquire off-target activity for canonical amino acids. Because initial templates had little or no activity for EDOT-Ala, it was especially difficult to evolve the aaRS to target EDOT-Ala in the presence of other amino acids to compete with. To circumvent this problem, future experiments should start with cycles without selection pressure to generate a diverse aaRS library. Another approach could be to construct a site-saturation mutagenesis library of SP by targeting key residues of aaRS and start selection cycles with the diversified library. These approaches will allow a broader search of aaRS mutants and increase the probability to generate an aaRS variant for EDOT-Ala.



(c)

	Mutations in evolved *mj*TyrRS	*
1	K101I, L110F, A167T	Tyr
2	K101I, L110F, P158H, A167T	Tyr
3	A167T	Tyr
4	E28K, L55S, L110F, D113A, P158T, A167T, V236G	Tyr

Figure 2.12. Evolution of *p*-NFRS for EDOT-Ala in M9 minimal media supplemented with 17 canonical amino acids (no Phe, Tyr, Trp) and EDOT-Ala (Phase 3). (a) Changes in the phage titer during six positive selection cycles followed by alternating negative and positive cycles. Three negative cycles and two positive cycles were alternated. Low stringency method was used for positive selection. SP solutions from cycle 3 of batch 1 in phase 1 and cycle 14 of batch 1 in phase 2 were used as starting SP libraries. (b) Fluorescence measurement of the cells co-expressing *p*-NFRS mutants and GFP-4am. Uninduced (red), no ncAA (blue), EDOT-Ala (magenta). Mutations in each variant are indicated in (c). (c) Common mutations found during evolution cycles. * indicates the amino acid incorporated at amber codon analyzed by ESI-MS after protein purification.

2.3.5. Discussion

We initially hypothesized that EDOT-Ala should be incorporated by one of the aaRS variants that recognizes the analogues of Phe, Tyr, and Trp due to the structural similarity between EDOT-Ala and these analogues. However, we were unable to find aaRS variants for EDOT-Ala through testing the activity of known aaRS and through their evolution. One critical difference between EDOT-Ala and these analogues would be that EDOT-Ala carries a non-planar dioxane ring whereas Phe, Tyr, and Trp carry an aromatic ring that can form π-stacking interactions with the residues of aaRS. Although various ncAAs have been incorporated using PheRS, TyrRS, and PylRS from different species, many of them carry an aromatic ring at the same position as Phe, Tyr, and Trp. Major exceptions are analogues

of Pyl, which carry longer side chains than Phe, Tyr, and Trp. Various Pyl analogues with large substituents have been incorporated using engineered PylRS, and therefore, we investigated the incorporation of EDOT-functionalized Pyl analogue in the next chapter.

2.4. Conclusion

Novel ncAA, EDOT-Ala was synthesized in three steps from commercial compounds. We sought to genetically incorporate EDOT-Ala using the known aaRS variants that have activity towards bulky amino acids (i.e. endogenous *E. coli* TrpRS, yPheRS(T415G), *Mm*PylRS(N346A/C348A)), but these variants did not show activity for EDOT-Ala. We then performed evolution of aaRS variants via cell viability-based selection and phage-assisted evolution. However, the selected mutants showed large off-target activity for canonical amino acids and did not incorporate EDOT-Ala. For future experiments, we propose to perform evolution with a more diverse aaRS library as a starting point, which increases the probability to generate aaRS for EDOT-Ala.

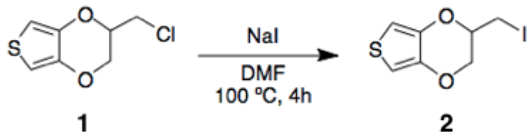
2.5. Experimental Procedures

2.5.1. General

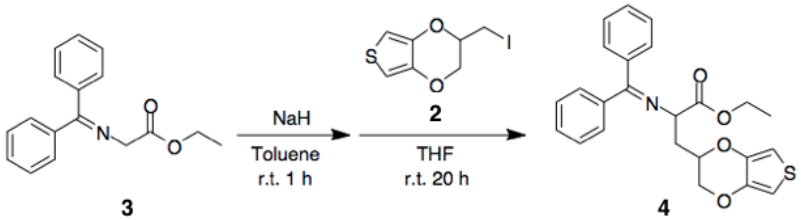
All the reagents for organic synthesis were purchased from Sigma Aldrich and Tokyo Chemical Industry Co., Ltd. (TCI), and used as received without further purification. For column chromatography, Biotage Isolera Spektra equipped with a SNAP HP-Sil 10 g cartridge or a SNAP ULTRA C18 12 g cartridge was used. ^{1}H (500 MHz) and ^{13}C NMR (126 MHz) spectra were recorded on a Varian Inova 500 spectrometer. Analytical HPLC analyses for small compounds were performed with an Agilent 1290 Infinity Series HPLC instrument at the Center for Catalysis and Chemical Synthesis in the Caltech Beckman Institute. High-resolution mass spectrometry (HRMS) was performed with an LCT Premier XE Electrospray TOF Mass Spectrometer with electrospray ionization (ESI) at the Caltech CCE Multiuser Mass Spectrometry Laboratory. Matrix-assisted laser desorption/ionization time-of-flight mass (MALDI-TOF MS) spectrometry measurements were carried out on a Bruker Daltonics autoflex^{TM} speed MALDI-TOF/TOF spectrometer at the Caltech CCE

Multiuser Mass Spectrometry Laboratory. A Tecan Safire II plate reader was used for the fluorescence measurement in 96-well plates.

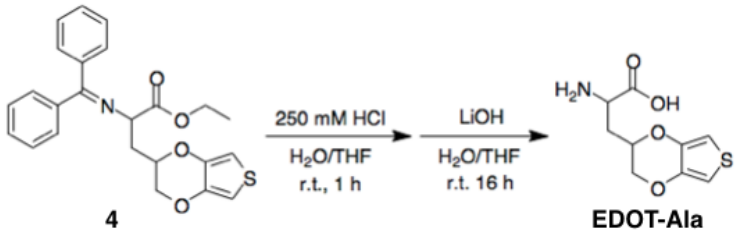
2.5.2. Synthesis and Characterization of Compounds



Iodomethyl EDOT (2). Chloromethyl EDOT (**1**) (1.20 g, 6.32 mmol) and sodium iodide (6.62 g, 44.2 mmol) were dissolved in anhydrous DMF (15 mL) under Ar atmosphere. The reaction mixture was heated to 100 ºC and stirred for 4 h. After cooling to room temperature, water (50 mL) was added and the reaction was extracted with DCM (3 x 20 mL). The combined organic layers were dried over $MgSO_4$, and concentrated *in vacuo*, yielding 1.7 g (98%) of a white solid. ^{1}H NMR (500 MHz, $CDCl_3$) δ 6.37 (d, J = 3.7 Hz, 1H), 6.36 (d, J = 3.7 Hz, 1H), 4.31 (dd, J = 11.5, 2.1 Hz, 1H), 4.28 – 4.23 (m, 1H), 4.15 (dd, J = 11.6, 6.2 Hz, 1H), 3.37 – 3.29 (m, 2H).; ^{13}C NMR (126 MHz, $CDCl_3$) δ 141.17, 140.89, 100.23, 100.12, 72.96, 67.40, 0.83. HRMS (ESI) calculated for $C_7H_7IO_2S$ $[M+H]^+$: m/z = 282.9290; found: 282.9307.



***N*-(Diphenylmethylene)-EDOT-Ala-ethyl ester (4).** *N*-(Diphenylmethylene)glycine ethyl ester (**3**) (1.37 g, 4.25 mmol) was dissolved in anhydrous toluene (4 mL) under Ar atmosphere. Sodium hydride (60% dispersion in mineral oil, 170 mg, 4.25 mmol) was added, and the reaction mixture was stirred at room temperature for 1 h. Iodomethyl EDOT (1.00 g, 3.54 mmol) was dissolved in anhydrous THF (4 mL) and slowly added to the reaction mixture. The solution was stirred at room temperature for 20 h. A small amount of ethanol was added to quench unreacted sodium hydride. Water (50 mL) was added and the reaction was extracted with DCM (3 x 20 mL). The combined organic layers were dried over $MgSO_4$, and concentrated *in vacuo*, yielding 580 mg (39%) of a white solid. ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.64-7.16 (m, 10H), 6.28 (dd, J = 3.7, 0.8 Hz, 1H), 6.16-6.11 (dd, J = 3.6, 0.8 Hz, 1H), 5.30 (d, J = 0.8 Hz, 1H), 4.47-3.83 (m, 5H), 2.25 (m, 2H), 1.27 (m, 3H); ^{13}C NMR (126 MHz, $CDCl_3$) δ 196.93, 172.61, 141.69, 132.57, 130.21, 129.04, 128.98, 128.88, 128.77, 128.65, 128.42, 128.40, 128.26, 128.19, 126.67, 99.55, 99.47, 70.34, 68.69, 68.19, 61.40, 60.97, 34.03, 14.33. HRMS (ESI) calculated for $C_{24}H_{23}NO_4S$ $[M+H]^+$: m/z = 422.1426; found: 422.0992.



EDOT-Ala. Compound 4 (1.00 g, 2.37 mmol) was dissolved in THF (10 mL), and 500 mM HCl in water (10 mL) was added to the solution. After stirring at room temperature for 1 h, lithium hydroxide (1.00 g, 41.7 mmol) was added, and the reaction mixture was stirred at room temperature for 16 h. The product was purified by C18-reversed phase chromatography (water:methanol 100:0 to 70:30), yielding 370 mg (68%) of a white solid. ^{1}H NMR (500 MHz, MeOD) δ 6.54–6.46 (m, 2H), 4.46–4.30 (m, 1H), 4.31–4.19 (m, 1H), 3.99 (m, 1H), 3.63 – 3.50 (m, 1H), 2.14 – 1.76 (m, 2H); ^{13}C NMR (126 MHz, MeOD) δ 180.74, 143.13, 143.08, 100.37, 100.36, 100.06, 72.66, 69.62, 53.80, 36.36. HRMS (ESI) calculated for $C_9H_{11}NO_4S$ $[M+H]^+$: *m/z* = 230.0487; found: 230.0468.

2.5.3. Cloning

General. The plasmids were constructed by standard recombinant DNA technology using *E. coli* DH10B strain. DNA oligomers for PCRs were purchased from IDT (Coralville, IA). Bacteria were grown on LB/agar plates and LB liquid media with the following antibiotic concentrations unless indicated otherwise: 100 μg/mL ampicillin, 100 μg/mL carbenicillin, 50 μg/mL kanamycin, 100 μg/mL spectinomycin, and 40 μg/mL chloramphenicol. The plasmid pQE80-SYNZIP3 was a kind gift from Dr. Bradley R. Silverman. The plasmid pEVOL-*Mm*PylRS(N346A/C348A)-PylT[50] was a kind gift from Prof. Wenshe R. Liu. The plasmids for cell viability-based selections[34,35] were kind gifts from Prof. Peter G. Schultz. SP and plasmids required for phage-assisted evolution[36] were kind gifts from Prof. David R. Liu. Plasmids and their corresponding coding sequences are presented in *2.7. DNA and Protein Sequences*.

pQE80-SYNZIP3-S4W. The pQE80-SYNZIP3 plasmid confers resistance to ampicillin and carries *N*- and *C*-terminal 6xHis-Tags and SYNZIP3 under the T5 promoter. A Trp codon was inserted by site-directed mutagenesis of Ser at position 4, yielding the plasmid pQE80-SYNZIP-S4W.

pQE80-GFP-S4am. The previously reported plasmid pQE-GFPrmAM[62] was subjected to site-directed mutagenesis to introduce two mutations: S4am and am251oc (am: amber, oc: ochre), yielding the plasmid pQE80-GFP-4am.

Construction of the pBK-*Mm*PylRS(Y306A/Y384F) library. The *Mm*PylRS(N346A/C348A) gene was amplified from pEVOL-*Mm*PylRS(N346A/C348A)-PylT and cloned into the pBK vector, replacing the gene of an aaRS variant to yield the plasmid pBK-*Mm*PylRS(Y306A/Y384F). To construct the site-saturation mutagenesis library, NNK (N = A, T, G, C, K = G, T) mutations were introduced at five sites (L305, Y306, L309, Y384, W417) by overlap extension PCR following the previous study.[55] The primers used to introduce NNK mutations are presented in *2.8. Primers and Plasmids*. The gene library was digested with the restriction enzymes *Nde*I and *Nsi*I, purified on 1% agarose gel, and ligated into the pBK vector that was digested by *Nde*I and *Pst*I, yielding the pBK-*Mm*PylRS(Y306A/Y384F) library. The ligation products were electroporated into *E. coli* Top10 cells, and the cells were rescued in SOC medium for 1 hour at 37 ºC, grown in 50 mL of 2xYT at 37 ºC to an optical density at 600 nm (OD600) of 1.0. To evaluate the transformation efficiency, the rescued SOC culture was serially diluted and plated on LB agar plate with kanamycin. The calculated library size was 1.0×10^6 independent transformants. Ten colonies were picked from the plate to confirm mutations. No significant mutation bias was observed at the targeted sites.

2.5.4. Protein Expression and Purification

Incorporation of EDOT-Ala into SYNZIP3 by *E. coli* TrpRS.

To incorporate EDOT-Ala into SYNZIP3, a Trp auxotrophic strain of *E. coli* (KY35) derived from DH10B was transformed with pQE-SYNZIP3-S4W. *E. coli* cultures were grown in M9 minimal media (5 mL) supplemented with 0.4% glycerol, 25 mg/L thiamine, 1 mM $MgSO_4$, 0.1 mM $CaCl_2$, 20 amino acids (40 mg/L, respectively), and 100 mg/L ampicillin at 37 °C to an OD600 of 0.5-0.7. Then the cells were sedimented by centrifugation, washed twice with 0.9% NaCl, and resuspended in M9 minimal media (5 mL) supplemented with 0.4% glycerol, 25 mg/L thiamine, 1 mM $MgSO_4$, 0.1 mM $CaCl_2$,

19 amino acids without Trp (40 mg/L, respectively), 100 mg/L ampicillin, and 3 mM EDOT-Ala. Protein expression was induced by addition of 1 mM isopropyl-β-thiogalactopyranoside (IPTG) for 5 hours, and the cells were then harvested by centrifugation.

Incorporation of EDOT-Ala into DHFR by *y*PheRS(T415G).

A Phe/Trp double auxotrophic strain of *E. coli* harboring the plasmids coding the yPheRS(T415G)/ytRNA$^{Phe}_{CUA_UG}$ pair and mDHFR with an amber codon at position 38 was used for the expression.[43] *E. coli* cultures were grown in M9 minimal media (5 mL) supplemented with 0.4% glycerol, 25 mg/L thiamine, 1 mM $MgSO_4$, 0.1 mM $CaCl_2$, 20 amino acids (40 mg/L, respectively), and antibiotics (100 mg/L ampicillin and 50 mg/L kanamycin) at 37 °C to an OD600 of 0.5-0.7. Then the cells were sedimented by centrifugation, washed twice with 0.9% NaCl, and resuspended in M9 minimal media (5 mL) supplemented with 0.4% glycerol, 25 mg/L thiamine, 1 mM $MgSO_4$, 0.1 mM $CaCl_2$, 18 amino acids (40 mg/L, respectively), 0.01 mM Trp, 0.03 mM Phe, antibiotics (100 mg/L ampicillin and 50 mg/L kanamycin) and 3 mM EDOT-Ala. Protein expression was induced by addition of 1 mM IPTG for 5 hours, and the cells were harvested by centrifugation.

Incorporation of EDOT-Ala into GFP by *Mm*PylRS(N346A/C348A).

To incorporate EDOT-Ala into GFP, BL21(DE3) or KY35 strain of *E. coli* was transformed with pQE-GFP-S4am. For expression in the BL21(DE3) strain, *E. coli* cultures were grown in 2xYT media at 37 °C to an OD600 of 0.5-0.7. Then 2 mM EDOT-Ala was added, and protein expression was induced by addition of 1 mM IPTG for 16 hours.

For expression in the KY35 strain, *E. coli* cultures were grown in M9 minimal media (5 mL) supplemented with 0.4% glycerol, 25 mg/L thiamine, 1 mM $MgSO_4$, 0.1 mM $CaCl_2$, 20 amino acids (40 mg/L), and ampicillin (100 mg/L) at 37 °C to an OD600 of 0.5-0.7. Then the cells were sedimented by centrifugation, washed twice with 0.9% NaCl, and resuspended in M9 minimal media (5 mL) supplemented with 0.4% glycerol, 25 mg/L thiamine, 1 mM $MgSO_4$, 0.1 mM $CaCl_2$, 17 amino acids without Phe, Tyr and Trp (40

mg/L, respectively), 100 mg/L ampicillin, and 2 mM EDOT-Ala. Protein expression was induced by addition of 1 mM IPTG for 5 h.

Protein Purification. After expression, cells were harvested by centrifugation, resuspended in lysis buffer (0.1 M Na_2HPO_4, 10 mM imidazole; pH 8.0), and lysed by sonication. Lysates were cleared by centrifugation and incubated with Ni-NTA agarose. The resin was washed with lysis buffer and wash buffer (0.1 M Na_2HPO_4, 25 mM imidazole; pH 8.0). Protein was eluted with elution buffer (0.1 M Na_2HPO_4, 250 mM imidazole; pH 8.0), dialyzed against water, and lyophilized for storage.

SDS-PAGE. After protein expression, cells were lysed by resuspension in a lysis buffer (100 mM Tris·HCl, pH 8.0, 1% SDS). Lysates were mixed with an SDS loading buffer (0.05% bromophenol blue, 0.1 M dithiothreitol (DTT), 10% glycerol, 2% SDS, 50 mM Tris·HCl; pH 8.0), heated to 95 °C for 5 min, and separated via SDS-PAGE. The gel was stained with Colloidal Blue (Life Technologies) and imaged on a Typhoon gel imager (GE Healthcare).

Flow Cytometry. After expression of GFP, cells were harvested by centrifugation, washed twice with PBS and strained through a 40 μm filter to remove aggregates. Flow cytometry was performed on a MoFlo XDP cell sorter equipped with an argon ion laser emitting at 488 nm. Data was analyzed in MATLAB using a custom software EasyFlow.[63]

2.5.5. Cell Viability-Based Selection

The selection followed the reported procedure.[34,35] The positive selection plasmid pRep-PylT-GFP confers tetracycline resistance and encodes CAT with an amber codon at position 112, T7 RNA with amber codons at position 6 and 114, GFP under the T7 promoter, and the amber suppressor $tRNA^{Pyl}$ under the constitutive *lpp* promoter. The negative selection plasmid pNeg-PylT conferes ampicillin resistance and encodes a bacterial toxin barnase with amber codons at positions 2 and 44 and the amber suppressor $tRNA^{Pyl}$ under the constitutive *lpp* promoter.

For the positive selection, the pBK-*Mm*PylRS(Y306A/Y384F) library was transformed into *E. coli* TOP10 electrocompetent cells (Invitrogen) with the pRep plasmid. Cells were grown on LB agar plates supplemented with 12 μg/mL tetracycline, 50 μg/mL kanamycin, 68 μg/mL chloramphenicol, and 2 mM EDOT-Ala. After 72 hours of incubation at 37 ºC, colonies were collected and grown in LB media (10 mL) supplemented with 12 μg/mL tetracycline and 25 μg/mL kanamycin. The plasmids were extracted and purified with Zyppy Plasmid Miniprep Kit (Zymo Research).

For the negative selection, the pBK-*Mm*PylRS(Y306A/Y384F) library collected from the positive selection was transformed into *E. coli* TOP10 electrocompetent cells (Invitrogen) with the pNeg plasmid. Cells were grown on LB agar plates supplemented with 50 μg/mL kanamycin, 200 μg/mL ampicillin, and 0.2% arabinose. After 18 hours of incubation at 37 ºC, the colonies were collected to extract plasmids. The plasmids were used for another cycle of positive selection.

2.5.6. Phage-Assisted Evolution

We performed phage-assisted evolution via the non-continuous method.[36] The plasmids used for the experiment are listed in *2.8. Primers and Plasmids*. S1030 cells harboring the required accessory plasmid and complementary plasmid were transformed with a mutagenesis plasmid. The cells were rescued in SOC media with 0.4% glucose for 1 hour at 37 ºC and grown on LB agar plates supplemented with 0.4% glucose and antibiotics. Single colonies of the transformed cells were picked and grown in LB media supplemented with 0.4% glucose and antibiotics for 14 hours at 37 ºC. The LB cultures were diluted 50-fold into 2xYT media (5 mL) supplemented with 0.4% glucose and antibiotics and grown to an OD600 of 0.5-0.7. Then cells were harvested by centrifugation and resuspended in 2xYT media (5 mL) supplemented with 5 mM arabinose and 2 mM EDOT-Ala. The cultures were inoculated with SP solution (0.2 pfu) and incubated for 12 hours at 37 ºC. The cells were centrifuged, and the supernatant was filtered with 0.22 μm syringe filter. The obtained SP solutions were stored at 4 ºC and used for the next cycle of evolution.

Negative selection was performed in the same procedure without addition of EDOT-Ala in the media.

For positive selection in minimal media, the overnight LB cultures of the host cell were diluted 50-fold into M9 minimal media (5 mL) supplemented with 0.4% glucose, 25 mg/L thiamine, 1 mM $MgSO_4$, 0.1 mM $CaCl_2$, 20 canonical amino acids (40 mg/L, respectively), and antibiotics, and grown to an OD600 of 0.5-0.7. Then cells were harvested by centrifugation and resuspended in M9 minimal media (5 mL) supplemented with 5 mM arabinose, 25 mg/L thiamine, 1 mM $MgSO_4$, 0.1 mM $CaCl_2$, 17 canonical amino acids without Tyr, Phe, Trp (40 mg/L, respectively), 2 mM EDOT-Ala, and antibiotics. The cultures were inoculated with SP solution (0.2 pfu) and incubated for 12 hours at 37 ºC. The cells were centrifuged, and the supernatant was filtered with 0.22 μm syringe filter. The obtained SP solutions were stored at 4 ºC and used for the next cycle of evolution.

Plaque Assay. S1059 cells were grown in 2xYT media supplemented with antibiotics to an OD600 of 0.6–0.8. The SP solution was serially diluted at 10-fold, yielding six total samples. Each solution (2 μL) was added to 100 μL of the cell culture, and 900 μL of top agar (0.7% bacteriological agar in 2xYT) dissolved at 55 ºC was mixed with the infected cells. The mixtures were plated on 2xYT agar plates without antibiotics. After incubation for 12 hours at 37 ºC, the number of plaques in each sample was counted and the phage titer of the original sample was calculated using the dilution factor.

2.6. NMR Spectra

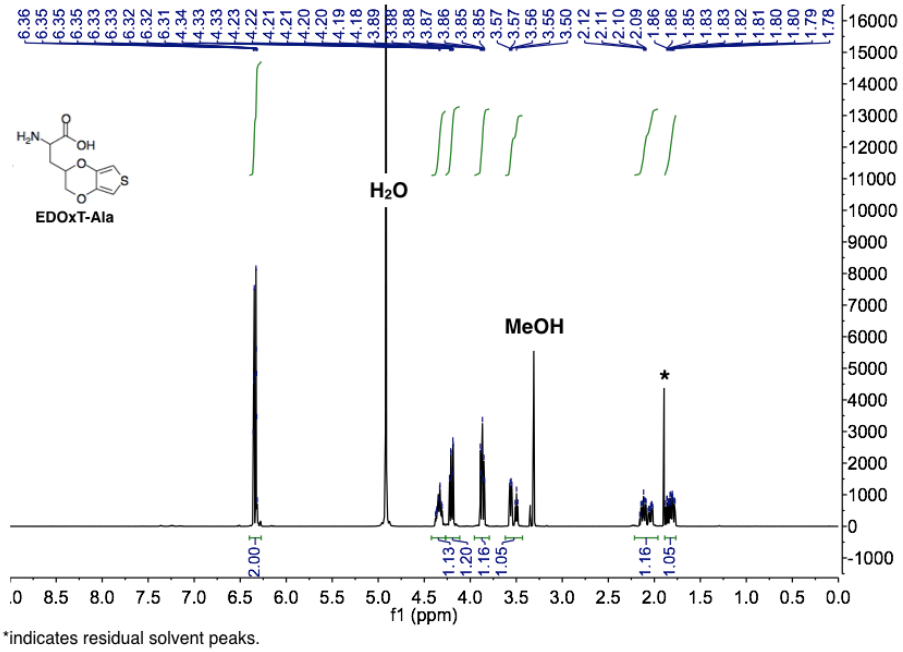


Figure S2.1. ^{1}H NMR spectrum (500 MHz) of EDOT-Ala in MeOD at 22 ºC.

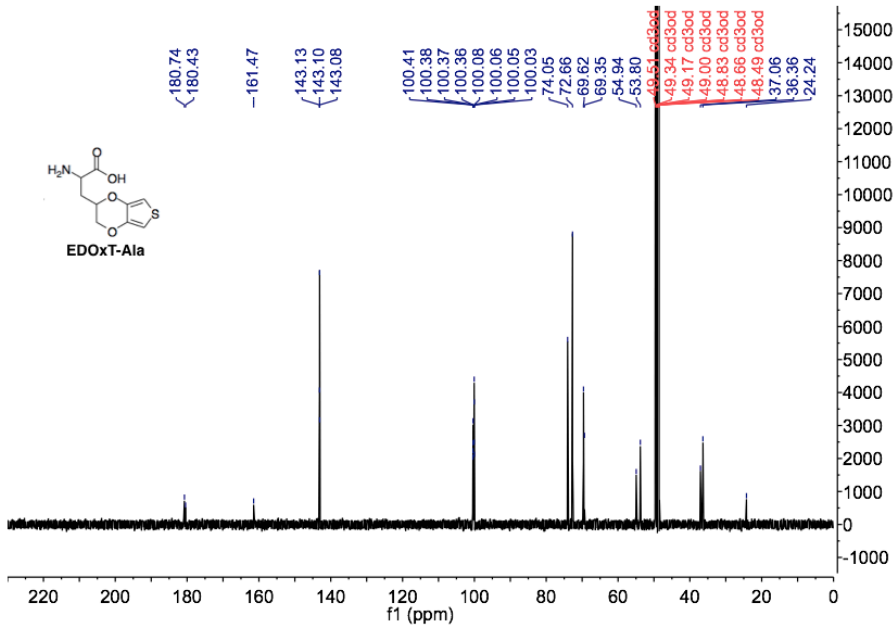


Figure S2.2. ^{13}C NMR spectrum (126 MHz) of EDOT-Ala in MeOD at 22 ºC.

2.7. DNA and Protein Sequences

SYNZIP3:

atgagaggatggcatcaccatcaccatcacggatccgtcgacggcagcggatcgggttcaggcagtggtagcg
ggtcgaatgaagtcacgactttggaaaatgatgcggcgttcatcgaaaatgagaacgcgtacttagagaagga
aatcgcgcgcctgcgcaaagaaaaagcggcattacgtaaccgcctggcccataaaaagggcagcggttcgggg
agcggttctggaagcctcgagcatcaccatcaccatcacaagctttgctag

MRGWHHHHHHGSVDGSGSGSGSGSGSNEVTTLENDAAFIENENAYLEKEIARLRKEKAALRNRLAHKKGSGSG
SGSGSLEHHHHHHKLC

*The single codon for Trp is highlighted in red. *N*- and *C*-terminal 6xHis-Tags are highlighted in blue.

GFP-S4am:

atgagaggatagcatcaccatcaccatcacggatccttgagtaaaggagaagaacttttcactggagtcgtcc
caattcttgttgaattagatggtgatgttaatgggcacaaattttctgtcagaggagagggtgaaggtgatgc
cacatacggaaaaattacccttaaattgatttgcactactggaaaactacctgttccatggccaacacttgtc
actacttgcggttatggtgttcaatgctttgcgcgttatccggatcatctgaaacggcatgacttttcaaga
gtgcctttcccgaaggttatgtacaggaacgcactatatctttcaaagatgacgggaagttcaagacgcgtgc
tgaagtcaagtttgaaggtgataccattgttaatcgtatcaagttaaaaggcattgattttaaagaagatgga
aacattctcggacacaaactcgagtacaactataactcacacgatgtatacatcacggcagacaaacaaaaga
ctggaatcaaagctaacttcaaaattcgccacaacgttgaagatggttccgttcaactggcagaccattatca
acaaaatactccaattggcgatggccctgtccgtttaccagacaaccattacctgttgacacaatctgtcatt
tcgaaagatcccaacgaaaagcgtgaccacgcggtccttcatgagtttgtaactgctgctgggattacacatg
gcatcgatgagctctacaaataa

MRGXHHHHHHGSLSKGEELFTGVVPILVELDGDVNGHKFSVRGEGEGDATYGKITLKLICTTGKLPVPWPTLV
TTCGYGVQCFARYPDHLKRHDFFKSAFPEGYVQERTISFKDDGKFKTRAEVKFEGDTIVNRIKLKGIDFKEDG
NILGHKLEYNYNSHDVYITADKQKTGIKANFKIRHNVEDGSVQLADHYQQNTPIGDGPVRLPDNHYLLTQSVI
SKDPNEKRDHAVLHEFVTAAGITHGIDELYK

*The amber codon is highlighted in red. *N*-terminal 10xHis-Tag is highlighted in blue.

2.8. Primers and Plasmids

Table S2.1. Sequences of primers for construction of a PylRS gene library.[55]

Lib-NdeI-F	gaatcccatatggataaaaaaccactaaacactctg
Lib-1-R	ggccctgtcaagcttgcg**mnn**gtagttmnnmnngtttggagcaagcatggg
Lib-2-F	cagatgggatcgggatgcacacg
Lib-3-R	tacatcaagggtatcccc**mnn**gaccatgcaggaatcgcc
Lib-4-F	ggggataccccttgatgtaatgcacggagac
Lib-5-R	ccgaaacctgcccctat**mnn**gggtttatcaatacccca
Lib-6-F	ataggggcaggtttcgggctcgaacgcc
Lib-PstI~NsiI-R	gggatttctaccaacctgtaaatgcattttcaaac

*N represents A, T, G, C, and M represents C, A.

Table 2.2. Plasmids used for phage-assisted evolution.[36]

Selection/stringency	Type	Name	Antibiotic	ORF 1		ORF 2		ORF 3	
				Prom	Gene	Prom	Gene	Prom	Gene
Positive-low	AP	pDB007+	Carb	P(T7)	gIII, luxAB	P(proK)	tyrT		
	CP	pDB023f1	Spec	P(psp)	T7RNAP(S12∗, S203∗)				
	MP	MP4	Cm	P(psp)	dnaQ926, dam, seqA	P(C)	araC		
Positive-high	AP	pDB026a,e,f	Carb	P(psp)	gIII(P29∗,P83∗, Y184∗), luxAB	P(proK)	tyrT		
	MP	MP4	Cm	P(psp)	dnaQ926, dam, seqA	P(C)	araC		
Negative	AP	pDB007ns2a	Carb	P(psp)	gIII	P(proK)	tyrT	P(tet)	T7RNAP(S12∗,S203∗)
	CP	pDB016	Spec	P(T7)	gIII-neg				
	MP	MP4	Cm	P(psp)	dnaQ926, dam, seqA	P(C)	araC		

*indicates an amber codon.

2.9. References

(1) Flynn, C. E.; Lee, S. W.; Peelle, B. R.; Belcher, A. M. Viruses as Vehicles for Growth, Organization and Assembly of Materials. *Acta Mater.* **2003**, *51* (19), 5867–5880.

(2) Tsukada, S.; Nakashima, H.; Torimitsu, K. Conductive Polymer Combined Silk Fiber Bundle for Bioelectrical Signal Recording. *PLoS One* **2012**, *7* (4).

(3) Calvaresi, M.; Zerbetto, F. The Devil and Holy Water: Protein and Carbon Nanotube Hybrids. *Acc. Chem. Res.* **2013**, *46* (11), 2454–2463.

(4) Meier, C.; Lifincev, I.; Welland, M. E. Conducting Core-Shell Nanowires by Amyloid Nanofiber Templated Polymerization. *Biomacromolecules* **2015**, *16* (2), 558–563.

(5) Bäcklund, F. G.; Elfwing, A.; Musumeci, C.; Ajjan, F.; Babenko, V.; Dzwolak, W.; Solin, N.; Inganäs, O. Conducting Microhelices from Self-Assembly of Protein Fibrils. *Soft Matter* **2017**, *13* (25), 4412–4417.

(6) Wang, C.; Xia, K.; Zhang, Y.; Kaplan, D. L. Silk-Based Advanced Materials for Soft Electronics. *Acc. Chem. Res.* **2019**, *52* (10), 2916–2927.

(7) Groenendaal, L.; Jonas, F.; Freitag, D.; Pielartzik, H.; Reynolds, J. R. Poly(3,4-Ethylenedioxythiophene) and Its Derivatives: Past, Present, and Future. *Adv. Mater.* **2000**, *12* (7), 481–494.

(8) Kirchmeyer, S.; Reuter, K. Scientific Importance, Properties and Growing Applications of Poly(3,4-Ethylenedioxythiophene). *J. Mater. Chem.* **2005**, *15* (21), 2077–2088.

(9) Rozlosnik, N. New Directions in Medical Biosensors Employing Poly(3,4-Ethylenedioxy Thiophene) Derivative-Based Electrodes. *Anal. Bioanal. Chem.*

2009, *395* (3), 637–645.

(10) Sun, K.; Zhang, S.; Li, P.; Xia, Y.; Zhang, X.; Du, D.; Isikgor, F. H.; Ouyang, J. Review on Application of PEDOTs and PEDOT:PSS in Energy Conversion and Storage Devices. *J. Mater. Sci. Mater. Electron.* **2015**, *26* (7), 4438–4462.

(11) Mantione, D.; del Agua, I.; Sanchez-Sanchez, A.; Mecerreyes, D. Poly(3,4-Ethylenedioxythiophene) (PEDOT) Derivatives: Innovative Conductive Polymers for Bioelectronics. *Polymers (Basel).* **2017**, *9* (8).

(12) Donahue, M. J.; Sanchez-Sanchez, A.; Inal, S.; Qu, J.; Owens, R. M.; Mecerreyes, D.; Malliaras, G. G.; Martin, D. C. Tailoring PEDOT Properties for Applications in Bioelectronics. *Mater. Sci. Eng. R Reports* **2020**, *140* (August 2019).

(13) Herland, A.; Persson, K. M.; Lundin, V.; Fahlman, M.; Berggren, M.; Jager, E. W. H.; Teixeira, A. I. Electrochemical Control of Growth Factor Presentation to Steer Neural Stem Cell Differentiation. *Angew. Chemie - Int. Ed.* **2011**, *50* (52), 12529–12533.

(14) Bongo, M.; Winther-Jensen, O.; Himmelberger, S.; Strakosas, X.; Ramuz, M.; Hama, A.; Stavrinidou, E.; Malliaras, G. G.; Salleo, A.; Winther-Jensen, B.; Owens, R. M. PEDOT:Gelatin Composites Mediate Brain Endothelial Cell Adhesion. *J. Mater. Chem. B* **2013**, *1* (31), 3860–3867.

(15) Zhu, B.; Luo, S. C.; Zhao, H.; Lin, H. A.; Sekine, J.; Nakao, A.; Chen, C.; Yamashita, Y.; Yu, H. H. Large Enhancement in Neurite Outgrowth on a Cell Membrane-Mimicking Conducting Polymer. *Nat. Commun.* **2014**, *5* (May), 1–9.

(16) Pires, F.; Ferreira, Q.; Rodrigues, C. A. V.; Morgado, J.; Ferreira, F. C. Neural Stem Cell Differentiation by Electrical Stimulation Using a Cross-Linked PEDOT Substrate: Expanding the Use of Biocompatible Conjugated Conductive Polymers for Neural Tissue Engineering. *Biochim. Biophys. Acta - Gen. Subj.* **2015**, *1850* (6), 1158–1168.

(17) Maione, S.; Gil, A. M.; Fabregat, G.; Del Valle, L. J.; Triguero, J.; Laurent, A.; Jacquemin, D.; Estrany, F.; Jiménez, A. I.; Zanuy, D.; Cativiela, C.; Alemán, C. Electroactive Polymer-Peptide Conjugates for Adhesive Biointerfaces. *Biomater. Sci.* **2015**, *3* (10), 1395–1405.

(18) Wang, S.; Guan, S.; Xu, J.; Li, W.; Ge, D.; Sun, C.; Liu, T.; Ma, X. Neural Stem Cell Proliferation and Differentiation in the Conductive PEDOT-HA/Cs/Gel Scaffold for Neural Tissue Engineering. *Biomater. Sci.* **2017**, *5* (10), 2024–2034.

(19) Cui, X.; Martin, D. C. Electrochemical Deposition and Characterization of Poly(3,4-Ethylenedioxythiophene) on Neural Microelectrode Arrays. *Sensors and Actuators* **2003**, *89*, 92–102.

(20) Ludwig, K. A.; Uram, J. D.; Yang, J.; Martin, D. C.; Kipke, D. R. Chronic Neural Recordings Using Silicon Microelectrode Arrays Electrochemically Deposited with a Poly(3,4-Ethylenedioxythiophene) (PEDOT) Film. *J. Neural Eng.* **2006**, *3* (1), 59–70.

(21) Luo, S. C.; Ali, E. M.; Tansil, N. C.; Yu, H. H.; Gao, S.; Kantchev, E. A. B.; Ying, J. Y. Poly(3,4-Ethylenedioxythiophene) (PEDOT) Nanobiointerfaces: Thin, Ultrasmooth, and Functionalized PEDOT Films with in Vitro and in Vivo Biocompatibility. *Langmuir* **2008**, *24* (15), 8071–8077.

(22) Castagnola, V.; Descamps, E.; Lecestre, A.; Dahan, L.; Remaud, J.; Nowak, L. G.; Bergaud, C. Parylene-Based Flexible Neural Probes with PEDOT Coated Surface for Brain Stimulation and Recording. *Biosens. Bioelectron.* **2015**, *67*, 450–457.

(23) Lin, K. C.; Tsai, T. H.; Chen, S. M. Performing Enzyme-Free H2O2 Biosensor and Simultaneous Determination for AA, DA, and UA by MWCNT-PEDOT Film. *Biosens. Bioelectron.* **2010**, *26* (2), 608–614.

(24) Jiang, F.; Yue, R.; Du, Y.; Xu, J.; Yang, P. A One-Pot “green” Synthesis of Pd-Decorated PEDOT Nanospheres for Nonenzymatic Hydrogen Peroxide Sensing.

Biosens. Bioelectron. **2013**, *44* (1), 127–131.

(25) García, M.; Orozco, J.; Guix, M.; Gao, W.; Sattayasamitsathit, S.; Escarpa, A.; Merkoçi, A.; Wang, J. Micromotor-Based Lab-on-Chip Immunoassays. *Nanoscale* **2013**, *5* (4), 1325–1331.

(26) Hai, W.; Goda, T.; Takeuchi, H.; Yamaoka, S.; Horiguchi, Y.; Matsumoto, A.; Miyahara, Y. Specific Recognition of Human Influenza Virus with PEDOT Bearing Sialic Acid-Terminated Trisaccharides. *ACS Appl. Mater. Interfaces* **2017**, *9* (16), 14162–14170.

(27) Tara Bahadur, K. C.; Tada, S.; Zhu, L.; Uzawa, T.; Minagawa, N.; Luo, S. C.; Zhao, H.; Yu, H. H.; Aigaki, T.; Ito, Y. In Vitro Selection of Electrochemical Peptide Probes Using Bioorthogonal TRNA for Influenza Virus Detection. *Chem. Commun.* **2018**, *54* (41), 5201–5204.

(28) Abidian, M. R.; Kim, D. H.; Martin, D. C. Conducting-Polymer Nanotubes for Controlled Drug Release. *Adv. Mater.* **2006**, *18* (4), 405–409.

(29) Boehler, C.; Kleber, C.; Martini, N.; Xie, Y.; Dryg, I.; Stieglitz, T.; Hofmann, U. G.; Asplund, M. Actively Controlled Release of Dexamethasone from Neural Microelectrodes in a Chronic in Vivo Study. *Biomaterials* **2017**, *129*, 176–187.

(30) Carli, S.; Fioravanti, G.; Armirotti, A.; Ciarpella, F.; Prato, M.; Ottonello, G.; Salerno, M.; Scarpellini, A.; Perrone, D.; Marchesi, E.; Ricci, D.; Fadiga, L. A New Drug Delivery System Based on Tauroursodeoxycholic Acid and PEDOT. *Chem. - A Eur. J.* **2019**, *25* (9), 2322–2329.

(31) Dougherty, D. A. Unnatural Amino Acids as Probes of Protein Structure and Function. *Curr. Opin. Chem. Biol.* **2000**, *4* (6), 645–652.

(32) Link, A. J.; Mock, M. L.; Tirrell, D. A. Non-Canonical Amino Acids in Protein Engineering. *Curr. Opin. Biotechnol.* **2003**, *14* (6), 603–609.

(33) Wang, L.; Schultz, P. G. Expanding the Genetic Code. *Angew. Chemie - Int. Ed.* **2004**, *44* (1), 34–66.

(34) Wang, L.; Schultz, P. G. A General Approach for the Generation of Orthogonal TRNAs. *Chem. Biol.* **2001**, *8* (9), 883–890.

(35) Santoro, S. W.; Wang, L.; Herberich, B.; King, D. S.; Schultz, P. G. An Efficient System for the Evolution of Aminoacyl-TRNA Synthetase Specificity. *Nat. Biotechnol.* **2002**, *20* (10), 1044–1048.

(36) Bryson, D. I.; Fan, C.; Guo, L. T.; Miller, C.; Söll, D.; Liu, D. R. Continuous Directed Evolution of Aminoacyl-TRNA Synthetases. *Nat. Chem. Biol.* **2017**, *13* (12), 1253–1260.

(37) Datta, D.; Wang, P.; Carrico, I. S.; Mayo, S. L.; Tirrell, D. A. A Designed Phenylalanyl-TRNA Synthetase Variant Allows Efficient in Vivo Incorporation of Aryl Ketone Functionality into Proteins. *J. Am. Chem. Soc.* **2002**, *124* (20), 5652–5653.

(38) Zhang, D.; Vaidehi, N.; Goddard, W. A.; Danzer, J. F.; Debe, D. Structure-Based Design of Mutant Methanococcus Jannaschii Tyrosyl-TRNA Synthetase for Incorporation of O-Methyl-L-Tyrosine. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99* (10), 6579–6584.

(39) Bullock, T. L.; Rodríguez-Hernández, A.; Corigliano, E. M.; Perona, J. J. A Rationally Engineered Misacylating Aminoacyl-TRNA Synthetase. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105* (21), 7428–7433.

(40) Sun, R.; Zheng, H.; Fang, Z.; Yao, W. Rational Design of Aminoacyl-TRNA Synthetase Specific for p-Acetyl-l-Phenylalanine. *Biochem. Biophys. Res. Commun.* **2010**, *391* (1), 709–715.

(41) Hauf, M.; Richter, F.; Schneider, T.; Faidt, T.; Martins, B. M.; Baumann, T.;

Durkin, P.; Dobbek, H.; Jacobs, K.; Möglich, A.; Budisa, N. Photoactivatable Mussel-Based Underwater Adhesive Proteins by an Expanded Genetic Code. *ChemBioChem* **2017**, *18* (18), 1819–1823.

(42) Baumann, T.; Hauf, M.; Richter, F.; Albers, S.; Möglich, A.; Ignatova, Z.; Budisa, N. Computational Aminoacyl-TRNA Synthetase Library Design for Photocaged Tyrosine. *Int. J. Mol. Sci.* **2019**, *20* (9).

(43) Kwon, I.; Tirrell, D. A. Site-Specific Incorporation of Tryptophan Analogues into Recombinant Proteins in Bacterial Cells. *J. Am. Chem. Soc.* **2007**, *129* (34), 10431–10437.

(44) Dumas, A.; Lercher, L.; Spicer, C. D.; Davis, B. G. Designing Logical Codon Reassignment-Expanding the Chemistry in Biology. *Chem. Sci.* **2015**, *6* (1), 50–69.

(45) Wan, W.; Tharp, J. M.; Liu, W. R. Pyrrolysyl-TRNA Synthetase: An Ordinary Enzyme but an Outstanding Genetic Code Expansion Tool. *Biochim. Biophys. Acta - Proteins Proteomics* **2014**, *1844* (6), 1059–1070.

(46) Ansari, A. M.; Ugwu, S. O. Efficient Synthesis of α-Substituted Amino Acid Ester: Alkylation and Hydrogenation Removal of Schiff's Base Protecting Group. *Synth. Commun.* **2008**, *38* (14), 2330–2340.

(47) Xie, J.; Liu, W.; Schultz, P. G. A Genetically Encoded Bidentate, Metal-Binding Amino Acid. *Angew. Chemie - Int. Ed.* **2007**, *46* (48), 9239–9242.

(48) Peeler, J. C.; Woodman, B. F.; Averick, S.; Miyake-Stoner, S. J.; Stokes, A. L.; Hess, K. R.; Matyjaszewski, K.; Mehl, R. A. Genetically Encoded Initiator for Polymer Growth from Proteins. *J. Am. Chem. Soc.* **2010**, *132* (39), 13575–13577.

(49) Reinke, A. W.; Grant, R. A.; Keating, A. E. A Synthetic Coiled-Coil Interactome Provides Heterospecific Modules for Molecular Engineering. *J. Am. Chem. Soc.* **2010**, *132* (17), 6025–6031.

(50) Wang, Y. S.; Fang, X.; Wallace, A. L.; Wu, B.; Liu, W. R. A Rationally Designed Pyrrolysyl-TRNA Synthetase Mutant with a Broad Substrate Spectrum. *J. Am. Chem. Soc.* **2012**, *134* (6), 2950–2953.

(51) Wang, Y. S.; Fang, X.; Chen, H. Y.; Wu, B.; Wang, Z. U.; Hilty, C.; Liu, W. R. Genetic Incorporation of Twelve Meta -Substituted Phenylalanine Derivatives Using a Single Pyrrolysyl-TRNA Synthetase Mutant. *ACS Chem. Biol.* **2013**, *8* (2), 405–415.

(52) Tuley, A.; Wang, Y. S.; Fang, X.; Kurra, Y.; Rezenom, Y. H.; Liu, W. R. The Genetic Incorporation of Thirteen Novel Non-Canonical Amino Acids. *Chem. Commun.* **2014**, *50* (20), 2673–2675.

(53) Tharp, J. M.; Wang, Y. S.; Lee, Y. J.; Yang, Y.; Liu, W. R. Genetic Incorporation of Seven Ortho-Substituted Phenylalanine Derivatives. *ACS Chem. Biol.* **2014**, *9* (4), 884–890.

(54) Yanagisawa, T.; Ishii, R.; Fukunaga, R.; Kobayashi, T.; Sakamoto, K.; Yokoyama, S. Multistep Engineering of Pyrrolysyl-TRNA Synthetase to Genetically Encode Nε-(o-Azidobenzyloxycarbonyl) Lysine for Site-Specific Protein Modification. *Chem. Biol.* **2008**, *15* (11), 1187–1197.

(55) Wang, Y. S.; Russell, W. K.; Wang, Z.; Wan, W.; Dodd, L. E.; Pai, P. J.; Russell, D. H.; Liu, W. R. The de Novo Engineering of Pyrrolysyl-TRNA Synthetase for Genetic Incorporation of l-Phenylalanine and Its Derivatives. *Mol. Biosyst.* **2011**, *7* (3), 714–717.

(56) Esvelt, K. M.; Carlson, J. C.; Liu, D. R. A System for the Continuous Directed Evolution of Biomolecules. *Nature* **2011**, *472* (7344), 499–503.

(57) Badran, A. H.; Liu, D. R. Development of Potent in Vivo Mutagenesis Plasmids with Broad Mutational Spectra. *Nat. Commun.* **2015**, *6*, 1–10.

(58) Carlson, J. C.; Badran, A. H.; Guggiana-Nilo, D. A.; Liu, D. R. Negative Selection and Stringency Modulation in Phage-Assisted Continuous Evolution. *Nat. Chem. Biol.* **2014**, *10* (3), 216–222.

(59) Tsao, M. L.; Summerer, D.; Ryu, Y.; Schultz, P. G. The Genetic Incorporation of a Distance Probe into Proteins in Escherichia Coli. *J. Am. Chem. Soc.* **2006**, *128* (14), 4572–4573.

(60) Cooley, R. B.; Karplus, P. A.; Mehl, R. A. Gleaning Unexpected Fruits from Hard-Won Synthetases: Probing Principles of Permissivity in Non-Canonical Amino Acid-TRNA Synthetases. *ChemBioChem* **2014**, *15* (12), 1810–1819.

(61) Amiram, M.; Haimovich, A. D.; Fan, C.; Wang, Y. S.; Aerni, H. R.; Ntai, I.; Moonan, D. W.; Ma, N. J.; Rovner, A. J.; Hong, S. H.; Kelleher, N. L.; Goodman, A. L.; Jewett, M. C.; Söll, D.; Rinehart, J.; Isaacs, F. J. Evolution of Translation Machinery in Recoded Bacteria Enables Multi-Site Incorporation of Nonstandard Amino Acids. *Nat. Biotechnol.* **2015**, *33* (12), 1272–1279.

(62) Szychowski, J.; Mahdavi, A.; Hodas, J. J. L.; Bagert, J. D.; Ngo, J. T.; Landgraf, P.; Dieterich, D. C.; Schuman, E. M.; Tirrell, D. A. Cleavable Biotin Probes for Labeling of Biomolecules via Azide-Alkyne Cycloaddition. *J. Am. Chem. Soc.* **2010**, *132* (51), 18351–18360.

(63) Antebi, Y. E.; Reich-Zeliger, S.; Hart, Y.; Mayo, A.; Eizenberg, I.; Rimer, J.; Putheti, P.; Pe'er, D.; Friedman, N. Mapping Differentiation under Mixed Culture Conditions Reveals a Tunable Continuum of T Cell Fates. *PLoS Biol.* **2013**, *11* (7).

Chapter 3

INCORPORATION OF EDOT-LYS INTO PROTEINS

3.1. Abstract

Genetic incorporation of ncAA with EDOT functionality will extend the potential of protein-based materials for their applications in bioelectronics. Because our attempts to incorporate EDOT-Ala into cellular proteins were unsuccessful, we synthesized another ncAA, which was termed EDOT-Lys. When a mutant *Mm*PylRS was co-expressed with the reporter protein GFP, *E. coli* cells showed strong fluorescence as an indication of successful incorporation of EDOT-Lys. The incorporation was fully confirmed by fluorescence measurements, SDS-PAGE, and MALDI-TOF mass spectrometry. The technique developed here will open up the possibility to fabricate protein-PEDOT conjugates through genetically incorporated EDOT functionality.

3.2. Introduction

Genetic incorporation of non-canonical amino acids (ncAAs) provides a powerful tool to introduce desirable functionalities into proteins. One novel application of this technique is to incorporate an electroactive functional group into proteins, which can be subsequently polymerized into conductive polymers. To this end, we sought to genetically incorporate an ncAA bearing 3,4-ethylenedioxythiophene (EDOT) functionality by harnessing the *E. coli* protein biosynthesis machinery. In the previous chapter, we synthesized an amino acid EDOT-Ala as an analogue of aromatic canonical amino acids. However, incorporation of EDOT-Ala was unsuccessful despite our effort to evolve cognate aminoacyl t-RNA synthetases (aaRSs) (see *Chapter 2*).

In addition to the derivatives of phenylalanine (Phe), analogues of pyrrolysine (Pyl) have also been used to introduce bulky functional groups into proteins.[1,2] Pyl is the 22nd canonical amino acid (the 21st being selenocysteine) that is found in the genetic code of some methanogenic archaea.[1] Pyl is co-translationally inserted in response to the amber stop codon. Due to the orthogonality in *E. coli* and eukaryotes, the PylRS/$tRNA_{CUA}$ pairs have been widely used to incorporate ncAAs. Pyl carries a 4-methylpyrroline unit linked to the side chain amine group of lysine via an amide bond. Therefore, the pyrrolysyl tRNA synthetase (PylRS) has a large binding pocket to accommodate the methylpyrroline group. Mutagenesis of PylRS has further extended the binding pocket to accommodate Pyl analogues with bulky functional groups. Examples include Pyl analogues with carboxybenzyl (Cbz),[3] cyclooctyne,[4] bicyclononyne (BCN),[5] norbornene,[6] furan,[7] and coumarin groups.[8] Many of these ncAAs have been incorporated by an engineered PylRS from *Methanosarcina mazei* (*Mm*PylRS(Y306A/Y384F)) developed by Yokoyama and coworkers.[3] Therefore, we envisioned that designing an ncAA with EDOT group as a Pyl analogue would be an effective approach to incorporate EDOT functionality into proteins.

3.3. Results and Discussion

3.3.1. Synthesis of EDOT-Lys

We designed and synthesized a Pyl analogue bearing an EDOT group, which we termed EDOT-Lys (Figure 3.1). Because Pyl analogues with bulky functional groups have been successfully incorporated by variants of *Mm*PylRS and *Mb*PylRS,[1,2] we expected that EDOT-Lys could be incorporated by reported PylRS variants or by evolved variants of them.

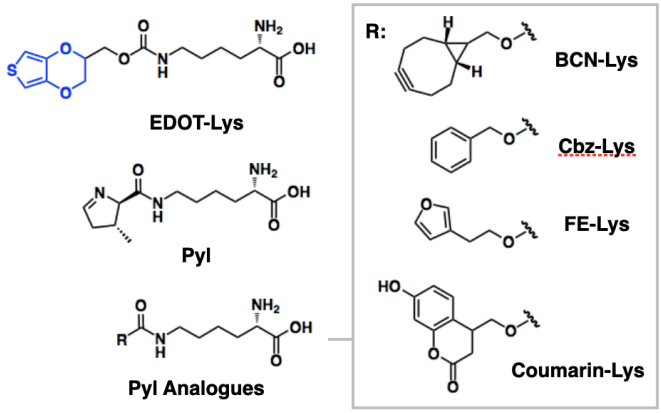


Figure 3.1. Chemical structures of EDOT-Lys, Pyl, and previously reported Pyl analogues.[3,5,7,8]

The synthetic scheme of EDOT-Lys is shown in Figure 3.2. For the synthesis of EDOT-Lys, we activated hydroxymethyl EDOT (5) with *N,N'*-disuccinimidyl carbonate (DSC) to obtain EDOT-NHS (6), which was treated with Fmoc-protected lysine. Deprotection of the Fmoc group with piperidine yielded EDOT-Lys. All the reactions proceeded under mild conditions with high yields (overall 53%), and EDOT-Lys was readily synthesized on gram scales. The details of the amino acid synthesis and NMR data are provided in Sections 3.4.2 and 3.6.

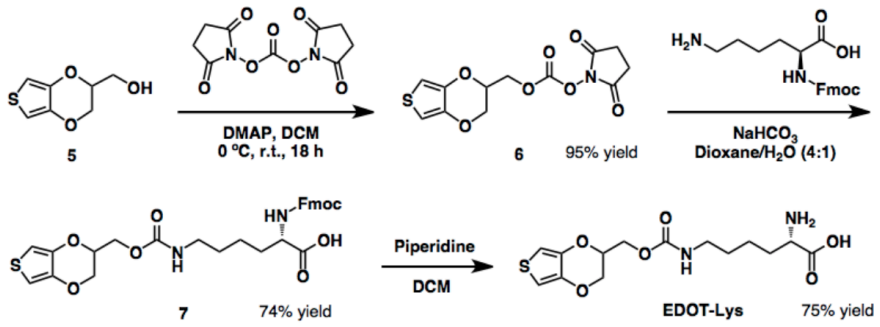


Figure 3.2. Synthesis of EDOT-Lys.

3.3.2. Incorporation of EDOT Functionality into GFP

To incorporate EDOT-Lys into recombinant proteins, we employed the *Mm*PylRS with two mutations at the active site (Y306A/Y384F), which was developed by Yokoyama and coworkers.[2,3] We chose the green fluorescent protein (GFP) as a test protein for investigation of the capacity of the *Mm*PylRS(Y306A/Y384F)/tRNA^{Pyl} pair to incorporate EDOT-Lys into recombinant proteins in *E. coli*. GFP was expressed from the plasmid pET16b-GFP-25am under the control of bacteriophage T7 promoter (Novagen). An amber codon was placed at position 25, which is upstream the GFP gene (Figure 3.3a), which allowed evaluation of the incorporation based on the measurement of whole-cell fluorescence. The *Mm*PylRS(Y306A/Y384F)/tRNA^{Pyl} pair was expressed from the modified pUltra plasmid.[9] The *Mm*PylRS(Y306A/Y384F) gene was placed under the IPTG-inducible *tacI* promoter and the tRNA^{Pyl} was placed under the constitutive *pro*K promoter. We co-transformed *E. coli* BL21(DE3) with the plasmids for the expression of GFP and the *Mm*PylRS(Y306A/Y384F)/tRNA^{Pyl} pair, and expressed the proteins in the absence and presence of 1 mM EDOT-Lys.

Cellular fluorescence as well as SDS-PAGE analysis of the lysate revealed the expression of GFP-25am in the presence of EDOT-Lys, whereas no expression was observed in the absence of EDOT-Lys (Figure 3.3b, c). In addition, we expressed GFP-25am in a large scale and purified by Ni-NTA chromatography (typical yield: 20 mg/L). The purified protein was digested by trypsin and analyzed by MALDI-

TOF and tandem mass spectrometry (Figure 3.4). The mass spectra of trypsin-digested GFP-25am showed the peaks corresponding to the peptides carrying EDOT-Lys. No peak was observed for the off-target incorporation of canonical amino acids. These results indicate the successful incorporation of EDOT-Lys.

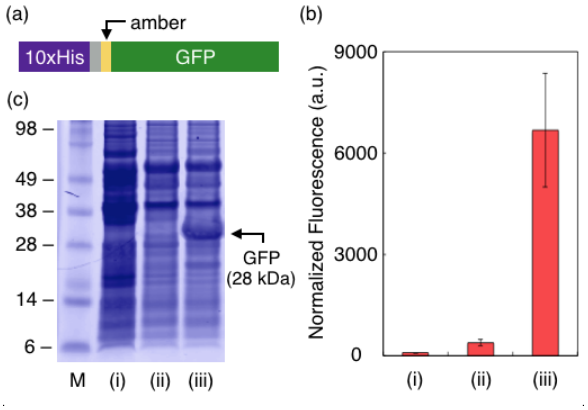


Figure 3.3. Incorporation of EDOT-Lys into GFP. (a) Sequence of GFP with amber codon. (b) Fluorescence intensity of *E. coli* measured at 510 nm (n = 4) and (c) Coomassie-stained SDS-PAGE. M: marker, (i) uninduced, (ii) no ncAA, (iii) EDOT-Lys.

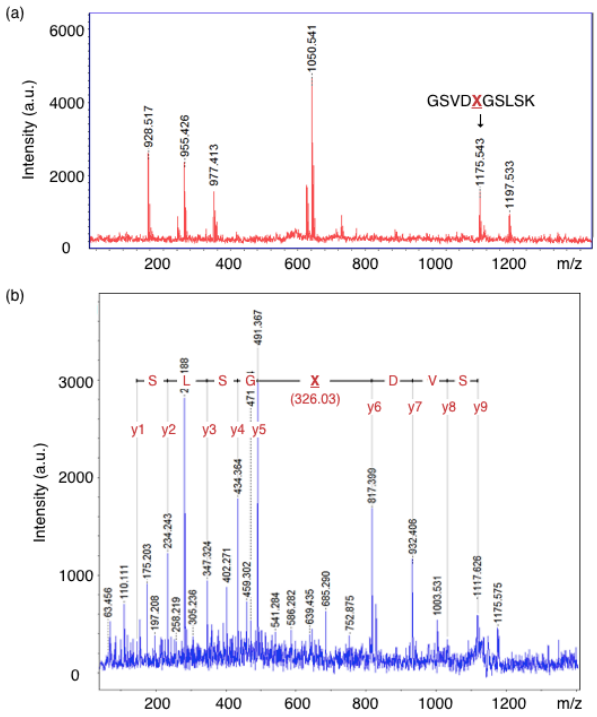


Figure 3.4. (a) MALDI mass spectra for trypsin-digested GFP-25am. The expected mass peak for the peptide with EDOT-Lys (GSVDXGSLSK; X = EDOT-Lys) was observed; calculated for $C_{48}H_{79}N_{12}O_{20}S$ $[M+H]^+$: *m/z* = 1175.525; found: 1175.543. (b) Tandem mass spectrum of the parent peak for GSVDXGSLSK. α-cyano-4-hydroxycinnamic acid (CHCA) was used as a matrix.

3.4. Conclusion

We prepared an amino acid EDOT-Lys as a Pyl analogue and successfully incorporated it into a GFP reporter using an *Mm*PylRS mutant. Incorporation was confirmed by expression of GFP as well as the analysis of the expressed protein by MALDI-TOF mass spectrometry. We believe that developing a technique to incorporate EDOT functionality into proteins is an important step towards fabrication of protein-based conductive materials.

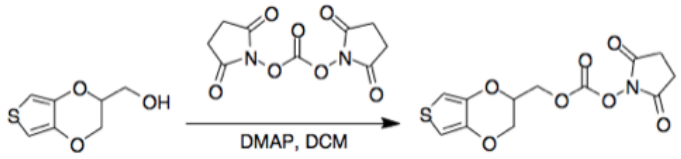
3.5. Experimental Procedures

3.5.1. General

All the reagents for organic synthesis were purchased from Sigma Aldrich and Tokyo Chemical Industry Co., Ltd. (TCI), and used as received without further purification. For column chromatography, Biotage Isolera Spektra equipped with a SNAP HP-Sil 10 g cartridge or a SNAP ULTRA C18 12 g cartridge was used. ^{1}H (500 MHz) and ^{13}C NMR (126 MHz) spectra were recorded on a Varian Inova 500 spectrometer. Analytical HPLC analyses for small compounds were performed with an Agilent 1290 Infinity Series HPLC instrument at the Center for Catalysis and Chemical Synthesis in the Caltech Beckman Institute. High-resolution mass spectrometry (HRMS) was performed with an LCT Premier XE Electrospray TOF Mass Spectrometer with electrospray ionization (ESI) at the Caltech CCE Multiuser Mass Spectrometry Laboratory. Matrix-assisted laser desorption/ionization time-of-flight mass (MALDI-TOF MS) spectrometry measurements were carried out on a Bruker Daltonics autoflex^{TM} speed MALDI-TOF/TOF spectrometer at the Caltech

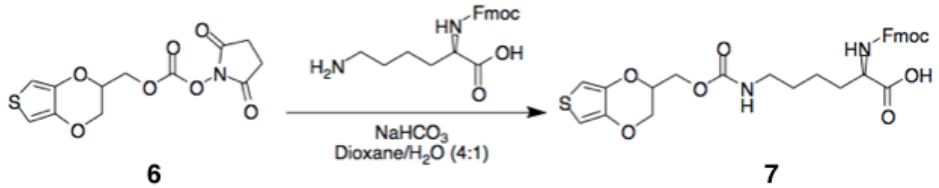
CCE Multiuser Mass Spectrometry Laboratory. A Tecan Safire II plate reader was used for the fluorescence measurement in 96-well plates.

3.5.2. Synthesis and Characterization of Compounds



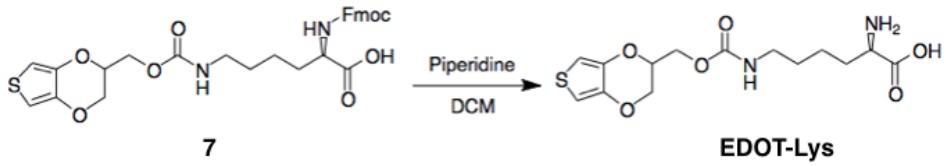
5 6

EDOT-NHS (6). Hydroxymethyl EDOT (5) (800 mg, 4.65 mmol) and 4-dimethylaminopyridine (85 mg, 0.70 mmol) were dissolved in anhydrous DCM (2 mL) under Ar atmosphere. *N,N'*-disuccinimidyl carbonate (2.10 g, 4.92 mmol) was added, and the reaction mixture was stirred at room temperature for 18 h. A solution of 5% citric acid (40 mL) was added and the reaction was extracted with DCM (3 x 20 mL). The combined organic layers were washed with saturated $NaHCO_3$ (2x 70 mL) and brine (2x100 mL), dried over $MgSO_4$, and concentrated *in vacuo*, yielding 1.4 g (95%) of a white solid. ^{1}H NMR (500 MHz, $CDCl_3$) δ 6.40 (d, J = 3.6 Hz, 1H), 6.37 (d, J = 3.7 Hz, 1H), 4.58-4.46 (m, 3H), 4.29-4.24 (m, 2H), 2.85 (s, 4H); ^{13}C NMR (126 MHz, $CDCl_3$) δ 168.47, 151.62, 141.02, 140.48, 100.72, 100.50, 70.71, 68.33, 65.06, 25.60. HRMS (ESI) calculated for $C_{12}H_{11}NO_7S$ $[M+H]^+$: m/z = 314.0329; found: 314.0314.



6 7

Fmoc-EDOT-Lys (7). Nα-Fmoc-L-lysine hydrochloride (646 mg, 1.60 mmol) and sodium bicarbonate (268 mg, 3.19 mmol) were dissolved in dioxane:H_2O (4:1, 10 mL). To the solution was added EDOT-NHS (6) (500 mg, 1.60 mmol) and the resulting mixture was allowed to stir at room temperature for 16 h. After completion of the reaction, 5% citric acid (80 mL) was added and the reaction was extracted with DCM (3 x 20 mL). The combined organic layers were dried over $MgSO_4$ and concentrated *in vacuo* to afford Fmoc-EDOT-Lys (7) as yellow oil (quantitative). The product was used for the next step without further purification.



7 EDOT-Lys

EDOT-Lys. Fmoc-EDOT-Lys (7) (800 mg, 1.41 mmol) was dissolved in 20% piperidine in DCM (15 mL) and allowed to stir at room temperature for 18 h. After the reaction, the solution was preciptated in diethyl ether (60 mL) and washed with diethyl ether (x3) and ice-cold MeOH to afford 365 mg of a white solid (75%). ^{1}H NMR (500 MHz, D_2O) δ 6.41 (s, 1H), 6.41 (s, 1H), 4.42 – 3.95 (m, 4H), 3.72 – 3.42 (m, 2H), 3.08 (q, *J* = 5.9 Hz, 2H), 1.91 – 1.21 (m, 6H); ^{13}C NMR (126 MHz, HFIP-d_2) δ 175.90, 160.62, 142.19, 142.18, 102.69, 102.68, 74.34, 67.69, 65.21, 58.49, 41.95, 31.94, 30.73, 23.76. HRMS (ESI) calculated for $C_{14}H_{20}N_2O_6S$ $[M+H]^+$: *m/z* = 345.1115; found: 345.1121.

3.5.3. Cloning

General. The plasmids were constructed by standard recombinant DNA technology using an *E. coli* DH10B strain. DNA oligomers for PCRs were purchased from IDT (Coralville, IA). Bacteria were grown on LB/agar plates and LB liquid media with the following antibiotic concentrations: 100 μg/mL spectinomycin and 100 μg/mL ampicillin. The pUltra vector[9] was the kind gift from Prof. Peter G. Schultz. The plasmid encoding *Mm*PylRS(N346A/C348A)[10] was a kind gift from Prof. Wenshe R. Liu. Plasmids and their corresponding coding sequences are presented in *3.7. DNA and Protein Sequences*.

pUltra-*Mm*PylRS(Y306A/Y384F). The pUltra vector confers resistance to spectinomycin and encodes aaRS under an IPTG-inducible *tacI* promoter and $tRNA_{CUA}$ under a constitutive *proK* promoter. The aaRS was replaced with *Mm*PylRS(N346A/C348A) gene, yielding the plasmid pUltra-*Mm*PylRS(N346A/C348A). A series of mutations (Y306A, A346N, A348C, and Y384F) were introduced by site-directed mutagenesis, yielding the plasmid pUltra-*Mm*PylRS(Y306A/Y384F).

pET16b-GFP-25am. The pET16b vector (Novagen) confers resistance to ampicillin and carries an *N*-terminal 10xHis-Tag and cloning sites under the T7 promoter. The GFP gene was amplified from a previously reported plasmid pQE-GFPrmAM,[11] and inserted to a cloning site by Gibson Assembly. An amber stop codon was introduced at the position 25.

3.5.4. Protein Expression and Purification

To incorporate EDOT-Lys into GFP, an *E. coli* strain BL21(DE3) was co-transformed with pUltra-*Mm*PylRS(Y306A/Y384F) and pET16b-GFP-25am. *E. coli* cultures were grown in 2xYT media at 37 °C to an optical density at 600 nm (OD600) of 0.5-0.7. Then 1 mM EDOT-Lys was added, and protein expression was induced by addition of 1 mM isopropyl-β-thiogalactopyranoside (IPTG). After 15 h, we diluted the cells with phosphate buffered saline (PBS), and analyzed the whole-cell fluorescence by plate reader. For SDS-PAGE, cells were lysed by resuspension in a lysis buffer (100 mM Tris·HCl, pH 8.0, 1% SDS). Lysates were mixed with an SDS loading buffer (0.05% bromophenol blue, 0.1 M dithiothreitol (DTT), 10% glycerol, 2% SDS, 50 mM Tris·HCl; pH 8.0), heated to 95 °C for 5 min, and separated via SDS-PAGE. The gel was stained with Colloidal Blue (Life Technologies) and imaged on a Typhoon gel imager (GE Healthcare).

For mass spectrometry, the cells were harvested by centrifugation after expression, resuspended in lysis buffer (0.1 M Na_2HPO_4, 10 mM imidazole; pH 8.0), and lysed by sonication. Lysates were cleared by centrifugation, and incubated with Ni-NTA agarose. The resin was washed with a lysis buffer and a wash buffer (0.1 M Na_2HPO_4, 25 mM imidazole; pH 8.0). Protein was eluted with an elution buffer (0.1 M Na_2HPO_4, 250 mM imidazole; pH 8.0), dialyzed against water, and lyophilized for storage.

3.5.5. MALDI-TOF Mass Spectrometry

The dry protein powder was dissolved in 100 mM NH_4HCO_3 (pH 8.0) to a final concentration of 1 mg/mL. Trypsin (Sigma-Aldrich; 10 U/mL) was added, and the solution was incubated at 37°C for 3 h. Proteolysis was quenched by addition of a final concentration of 0.1% trifluoroacetic acid (TFA) and the peptides were purified by C18 ZipTip (Millipore) according to the manufacturer's protocol. The peptides were eluted in 50% acetonitrile with 0.1% TFA and analyzed by MALDI-TOF mass spectrometry.

3.6. NMR Spectra

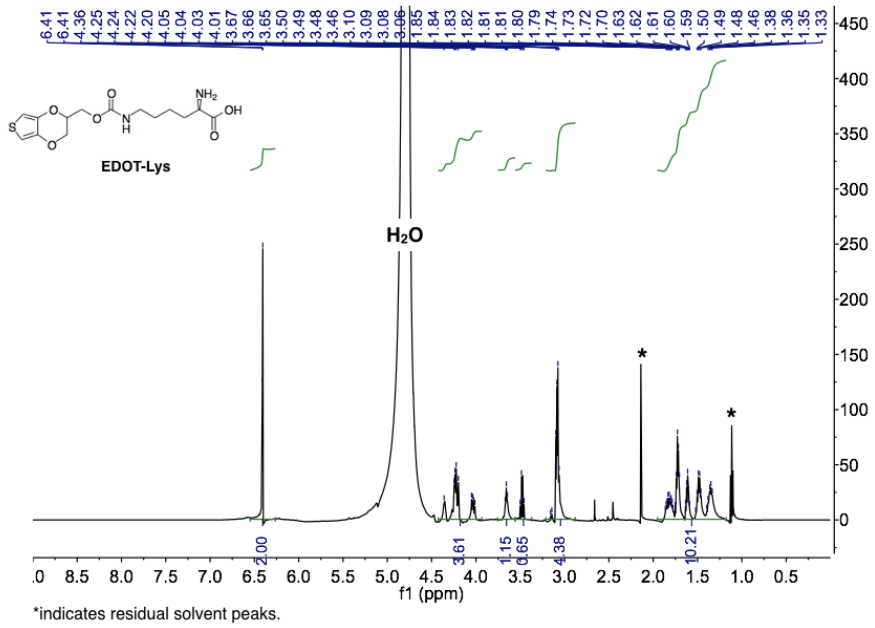


Figure S3.1. ^{1}H NMR spectrum (500 MHz) of EDOT-Lys in D_2O at 22 ºC.

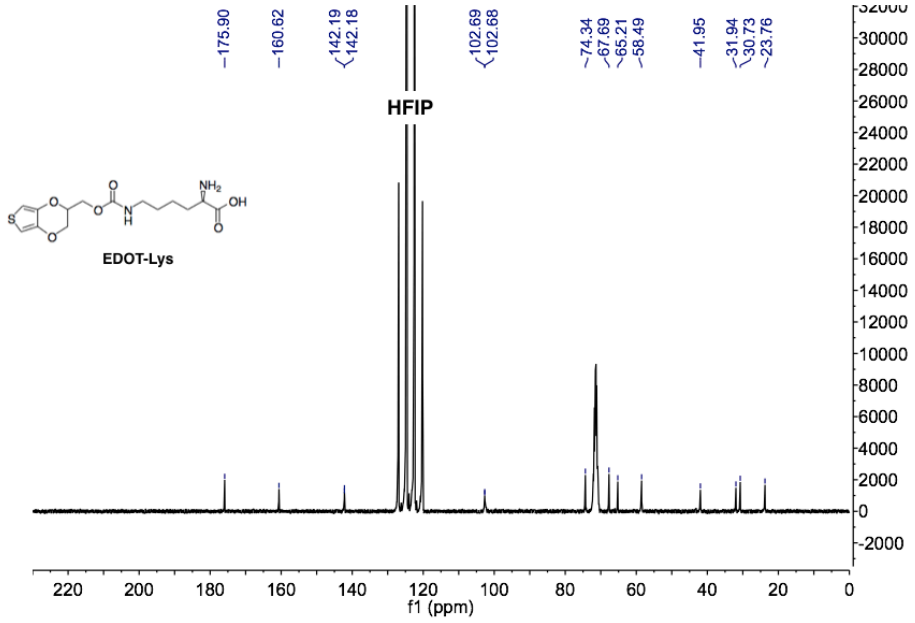


Figure S3.2. ^{13}C NMR spectrum (126 MHz) of EDOT-Lys in HFIP-d_2 at 22 ºC.

3.7. DNA and Protein Sequences

*Mm*PylRS(Y306A/Y384F):

atggataaaaaaccactaaacactctgatatctgcaaccgggctctggatgtccaggaccggaacaa
ttcataaaataaaacaccacgaagtctctcgaagcaaaatctatattgaaatggcatgcggagacca
ccttgttgtaaacaactccaggagcagcaggactgcaagagcgctcaggcaccacaaatacaggaag
acctgcaaacgctgcagggtttcggatgaggatctcaataagttcctcacaaaggcaaacgaagacc
agacaagcgtaaaagtcaaggtcgtttctgcccctaccagaacgaaaaaggcaatgccaaaatccgt
tgcgagagccccgaaacctcttgagaatacagaagcggcacaggctcaaccttctggatctaaattt
tcacctgcgataccggtttccacccaagagtcagtttctgtcccggcatctgtttcaacatcaatat
caagcatttctacaggagcaactgcatccgcactggtaaaagggaatacgaaccccattacatccat
gtctgcccctgttcaggcaagtgcccccgcacttacgaagagccagactgacaggcttgaagtcctg
ttaaacccaaaagatgagatttccctgaattccggcaagcctttcagggagcttgagtccgaattgc
tctctcgcagaaaaaaagacctgcagcagatctacgcggaagaaagggagaattatctggggaaact
cgagcgtgaaattaccaggttctttgtggacaggggttttctggaaataaaatccccgatcctgatc
cctcttgagtatatcgaaaggatgggcattgataatgataccgaactttcaaaacagatcttcaggg
ttgacaagaacttctgcctgagacccatgcttgctccaaaccttgccaactacctgcgcaagcttga
cagggccctgcctgatccaataaaaatttttgaaataggcccatgctacagaaaagagtccgacggc
aaagaacacctcgaagagtttaccatgctgaacttctgccagatgggatcgggatgcacacgggaaa
atcttgaaagcataattacggacttcctgaaccacctgggaattgatttcaagatcgtaggcgattc
ctgcatggtctttggggataccttgatgtaatgcacggagacctggaactttcctctgcagtagtc
ggacccataccgcttgaccgggaatggggtattgataaaccctggataggggcaggtttcgggctcg
aacgccttctaaaggttaaacacgactttaaaaatatcaagagagctgcaaggtccgagtcttacta
taacgggatttctaccaacctgtaa

MDKKPLNTLISATGLWMSRTGTIHKIKHHEVSRSKIYIEMACGDHLVVNNSRSSRTARALRHHKYRK
TCKRCRVSDEDLNKFLTKANEDQTSVKVKVVSAPTRTKKAMPKSVARAPKPLENTEAAQAQPSGSKF
SPAIPVSTQESVSVPASVSTSISSISTGATASALVKGNTNPITSMSAPVQASAPALTKSQTDRLEVL
LNPKDEISLNSGKPFRELESELLSRRKKDLQQIYAEERENYLGKLEREITRFFVDRGFLEIKSPILI
PLEYIERMGIDNDTELSKQIFRVDKNFCLRPMLAPNLANYLRKLDRALPDPIKIFEIGPCYRKESDG
KEHLEEFTMLNFCQMGSGCTRENLESIITDFLNHLGIDFKIVGDSCMVFGDTLDVMHGDLELSSAVV
GPIPLDREWGIDKPWIGAGFGLERLLKVKHDFKNIKRAARSESYYNGISTNL

*The two mutations (Y306A and Y384F) are highlighted in red.

GFP-25am:

atgggc**catcatcatcatcatcatcatcatcac**agcagcggccatatcgaaggtcgtggatccg
tcgac**tag**ggttccttgagtaaaggagaagaacttttcactggagtcgtcccaattcttgttgaatt
agatggtgatgttaatgggcacaaattttctgtcagaggagagggtgaaggtgatgccacatacgga
aaaattacccttaaattgatttgcactactggaaaactacctgttccatggccaacacttgtcacta
cttgcggttatggtgttcaatgctttgcgcgttatccggatcatctgaaacggcatgacttttcaa
gagtgcctttcccgaaggttatgtacaggaacgcactatatctttcaaagatgacgggaagttcaag
acgcgtgctgaagtcaagtttgaaggtgataccattgttaatcgtatcaagttaaaaggcattgatt
ttaaagaagatggaaacattctcggacacaaactcgagtacaactataactcacacgatgtatacat
cacggcagacaaacaaaagactggaatcaaagctaacttcaaaattcgccacaacgttgaagatggt
tccgttcaactggcagaccattatcaacaaaatactccaattggcgatggccctgtccgtttaccag
acaaccattacctgttgacacaatctgtcatttcgaaagatcccaacgaaaagcgtgaccacgcggt
ccttcatgagtttgtaactgctgctgggattacacatggcatcgatgagctctacaaataa

MG**HHHHHHHHHH**SSGHIEGRGSVD**X**GSLSKGEELFTGVVPILVELDGDVNGHKFSVRGEGEGDATYG
KITLKLICTTGKLPVPWPTLVTTCGYGVQCFARYPDHLKRHDFFKSAFPEGYVQERTISFKDDGKFK
TRAEVKFEGDTIVNRIKLKGIDFKEDGNILGHKLEYNYNSHDVYITADKQKTGIKANFKIRHNVEDG
SVQLADHYQQNTPIGDGPVRLPDNHYLLTQSVISKDPNEKRDHAVLHEFVTAAGITHGIDELYK

*The amber codon is highlighted in red. *N*-terminal 10xHis-Tag is highlighted in blue.

3.8. References

(1) Dumas, A.; Lercher, L.; Spicer, C. D.; Davis, B. G. Designing Logical Codon Reassignment-Expanding the Chemistry in Biology. *Chem. Sci.* **2015**, *6* (1), 50–69.

(2) Yanagisawa, T.; Kuratani, M.; Seki, E.; Hino, N.; Sakamoto, K.; Yokoyama, S. Structural Basis for Genetic-Code Expansion with Bulky Lysine Derivatives by an Engineered Pyrrolysyl-TRNA Synthetase. *Cell Chem. Biol.* **2019**, *26* (7), 936-949.e13.

(3) Yanagisawa, T.; Ishii, R.; Fukunaga, R.; Kobayashi, T.; Sakamoto, K.; Yokoyama, S. Multistep Engineering of Pyrrolysyl-TRNA Synthetase to

Genetically Encode Nε-(o-Azidobenzyloxycarbonyl) Lysine for Site-Specific Protein Modification. *Chem. Biol.* **2008**, *15* (11), 1187–1197.

(4) Plass, T.; Milles, S.; Koehler, C.; Schultz, C.; Lemke, E. A. Genetically Encoded Copper-Free Click Chemistry. *Angew. Chemie - Int. Ed.* **2011**, *50* (17), 3878–3881.

(5) Borrmann, A.; Milles, S.; Plass, T.; Dommerholt, J.; Verkade, J. M. M.; Wießler, M.; Schultz, C.; van Hest, J. C. M.; van Delft, F. L.; Lemke, E. A. Genetic Encoding of a Bicyclo[6.1.0]Nonyne-Charged Amino Acid Enables Fast Cellular Protein Imaging by Metal-Free Ligation. *ChemBioChem* **2012**, *13* (14), 2094–2099.

(6) Lang, K.; Davis, L.; Torres-Kolbus, J.; Chou, C.; Deiters, A.; Chin, J. W. Genetically Encoded Norbornene Directs Site-Specific Cellular Protein Labelling via a Rapid Bioorthogonal Reaction. *Nat. Chem.* **2012**, *4* (4), 298–304.

(7) Schmidt, M. J.; Summerer, D. Red-Light-Controlled Protein-RNA Crosslinking with a Genetically Encoded Furan. *Angew. Chemie - Int. Ed.* **2013**, *52* (17), 4690–4693.

(8) Luo, J.; Uprety, R.; Naro, Y.; Chou, C.; Nguyen, D. P.; Chin, J. W.; Deiters, A. Genetically Encoded Optochemical Probes for Simultaneous Fluorescence Reporting and Light Activation of Protein Function with Two-Photon Excitation. *J. Am. Chem. Soc.* **2014**, *136* (44), 15551–15558.

(9) Chatterjee, A.; Sun, S. B.; Furman, J. L.; Xiao, H.; Schultz, P. G. A Versatile Platform for Single- and Multiple-Unnatural Amino Acid Mutagenesis in Escherichia Coli. *Biochemistry* **2013**, *52* (10), 1828–1837.

(10) Wang, Y. S.; Fang, X.; Wallace, A. L.; Wu, B.; Liu, W. R. A Rationally Designed Pyrrolysyl-TRNA Synthetase Mutant with a Broad Substrate

Spectrum. *J. Am. Chem. Soc.* **2012**, *134* (6), 2950–2953.

(11) Szychowski, J.; Mahdavi, A.; Hodas, J. J. L.; Bagert, J. D.; Ngo, J. T.; Landgraf, P.; Dieterich, D. C.; Schuman, E. M.; Tirrell, D. A. Cleavable Biotin Probes for Labeling of Biomolecules via Azide-Alkyne Cycloaddition. *J. Am. Chem. Soc.* **2010**, *132* (51), 18351–18360.

Chapter 4

ELECTROPOLYMERIZATION OF XTEN WITH AN EDOT GROUP

4.1. Abstract

Genetic incorporation of EDOT-Lys could be a useful approach to create protein-based conductive materials. To demonstrate the utility of EDOT-Lys for electrochemical synthesis of protein–PEDOT conjugates, we incorporated EDOT-Lys into a model protein XTEN and performed electropolymerization by cyclic voltammetry. When electropolymerized with a water-soluble, self-doping EDOT monomer (EDOT-S), XTEN carrying EDOT-Lys (XTEN-E49am) formed a blue precipitate in the solution, whereas XTEN carrying no EDOT-Lys (XTEN-noam) did not form any precipitate after electropolymerization. Notably, MALDI-TOF mass spectra with an internal standard protein revealed that the large amount of XTEN-E49am was removed from the solution upon electropolymerization, suggesting that XTEN-E49am was integrated into the solid fraction of the product. Further analysis by FT-IR spectroscopy revealed that the solid product contains both XTEN-E49am and PEDOT-S. In addition, we investigated the effect of the chemical environment surrounding EDOT-Lys using a series of model peptides. Electropolymerization did not proceed in the applied potential range when Cys, Lys, Met, and Trp were located adjacent to EDOT-Lys. The peptide carrying Arg next to EDOT-Lys revealed a higher oxidation potential compared to the peptides. Nonetheless, electropolymerization of XTEN-E49am carrying these adjacent amino acids (XTEN-E49am-G50Z; Z = Cys, Lys, Met, Arg, Trp) with EDOT-S proceeded well without noticeable effects from the adjacent amino acids. The technique presented here provides a versatile strategy for the synthesis of protein–PEDOT conjugates and opens new avenues for protein-based conductive materials that are useful for various bioelectronics applications.

4.2. Introduction

The genetic code expansion technique has been revolutionizing the approach to develop protein-based materials.[1,2] One novel application of this technique is creating a protein-based conductive polymer using an electroactive non-canonical amino acid (ncAA). To explore this possibility, we incorporated an amino acid bearing 3,4-ethylenedioxythiophene (EDOT) group (designated EDOT-Lys) into proteins (*see Chapter 3*). In this chapter, we describe the polymerization of proteins through the pendant EDOT-Lys residue via electropolymerization.

PEDOT is often synthesized via one of the two major approaches: oxidative chemical polymerization and electropolymerization.[3,4] Electropolymerization has advantages over oxidative chemical polymerization in that it only requires short reaction time and a small volume of the sample, while chemical polymerization is preferred for large scale synthesis.[3,4] Figure 4.1 describes the proposed mechanism of polymerization of the EDOT monomer.[5,6] Upon application of an oxidation potential or addition of an oxidant, EDOT monomers form radical cations. Two radical cations couple together to form a dimer of EDOT, and the repeat of this process elongates the polymer chain to form PEDOT. The oxidation of a neutral oligomer to the radical cation proceeds much faster than that of the EDOT monomer, and therefore, the initial monomer oxidation is the rate-limiting step.[5]

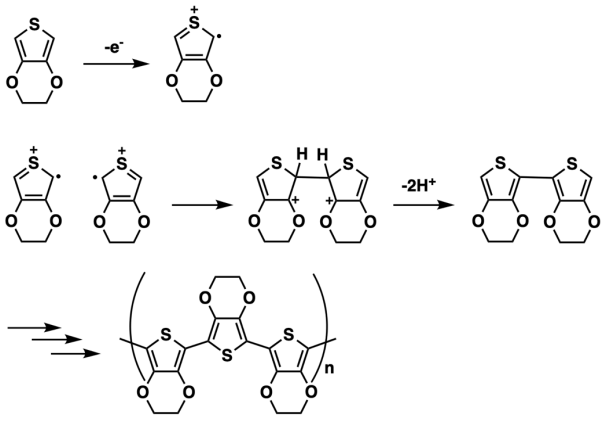


Figure 4.1. Mechanism of polymerization of the EDOT monomer.

In electropolymerization, an oxidation potential or current is applied to a monomer solution through the working electrode connected to an electrochemical instrument. The type of electrolytes, solvents, and reaction time affect the property of the polymer product.[7,8] Although the EDOT monomer is not soluble in water, previous studies have shown that electropolymerization of EDOT proceeds in aqueous solutions.[9–11] Electropolymerization of EDOT in the presence of protein has also been demonstrated.[12–14] Functionalized EDOT monomers such as EDOT with sulfonate group (EDOT-S) are also used to synthesize water-soluble PEDOT,[15,16] which is useful for biological applications.[17–21]

4.3. Results and Discussion

4.3.1. Electropolymerization of XTEN with an EDOT Group

To evaluate the utility of EDOT-Lys for polymerization, we employed an unstructured protein XTEN as a model protein. XTEN is a soluble, unstructured, artificial protein developed to prolong the *in vivo* retention of therapeutic proteins.[22–24] Conjugation with XTEN (XTENylation) increases the hydrodynamic diameter of target molecules in a similar manner as poly(ethylene glycol).[22–24] XTEN is composed of six hydrophilic amino acids (Ala, Glu, Gly, Pro, Ser, and Thr).[22] We introduced a C-terminal 6xHis-Tag to an XTEN construct and mutated a Glu residue at position 49 to an amber codon (XTEN-E49am; ~10 kDa). The protein was expressed in *E. coli* BL21(DE3) harboring pUltra-*Mm*PylRS(Y306A/Y384F) in the presence of EDOT-Lys. A control protein with no amber codon (XTEN-noam) was also prepared.

Electropolymerization was performed in aqueous solutions of the proteins (0.8 mM) with a supporting electrolyte $NaClO_4$ (100 mM), a self-doped EDOT derivative EDOT-S (25 mM),[15,16] and 5% trifluoroacetic acid (TFA). The solutions were dropped on the fluorine-doped tin oxide (FTO) working electrode. The potential was applied by cyclic voltammetry (CV) (E_{low} = +0.65 V; E_{high} = +1.5 V; scan rate = 100 mV/s; 5 cycles) in a three-electrode cell equipped with an Ag/AgCl

as the reference electrode and a Pt wire as the counter electrode. We observed the onset of oxidation current at 1.1 V in all three samples containing (i) EDOT-S, (ii) EDOT-S and XTEN-noam, and (iii) EDOT-S and XTEN-E49am (Figure 4.2). The solutions turned a characteristic blue color as an indication of the formation of PEDOT chains. The consumption of EDOT-S monomer was calculated as 97% by NMR spectra of EDOT-S before and after electropolymerization (Figure S4.1). The current decreased over the CV cycles, which is attributed to the deposition of the polymer on the FTO electrodes. MALDI mass spectrometry of the electropolymerized XTEN-E49am revealed the appearance of new peaks corresponding to the addition of EDOT-S monomer units to the protein, whereas only the original protein peak was observed for XTEN-noam (Figure 4.3, Figure S4.2). Notably, the solution of (iii) EDOT-S and XTEN-E49am formed a precipitate immediately after electropolymerization (Figure 4.3e), and large aggregates (over the size of 100 μm) were observed under the optical microscope (Figure 4.4c, f). No aggregates or only small aggregates were formed after electropolymerization in the solutions containing (i) EDOT-S and (ii) EDOT-S and XTEN-noam (Figure 4.4a, b, d, e).

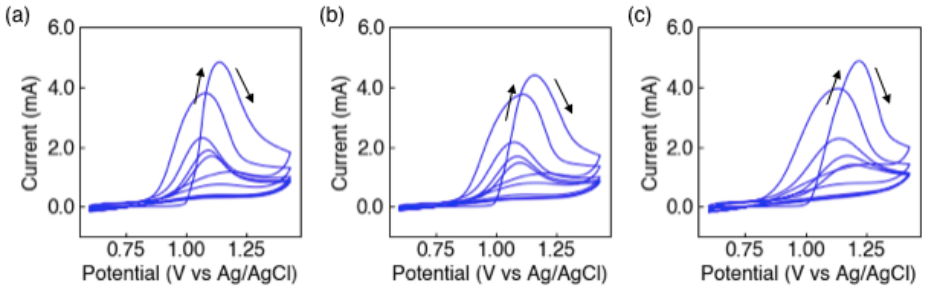


Figure 4.2. Electropolymerization of (a) EDOT-S, (b) EDOT-S and XTEN-noam, and (c) EDOT-S and XTEN-E49am by cyclic voltammetry on FTO electrodes (E_{low} = +0.65 V; E_{high} = +1.5 V; scan rate = 100 mV/s; 5 cycles). Samples are prepared at [XTEN] = 0.8 mM, [EDOT-S] = 25 mM, [$NaClO_4$] = 100 mM in H_2O with 5% TFA. Arrows indicate the scan direction.

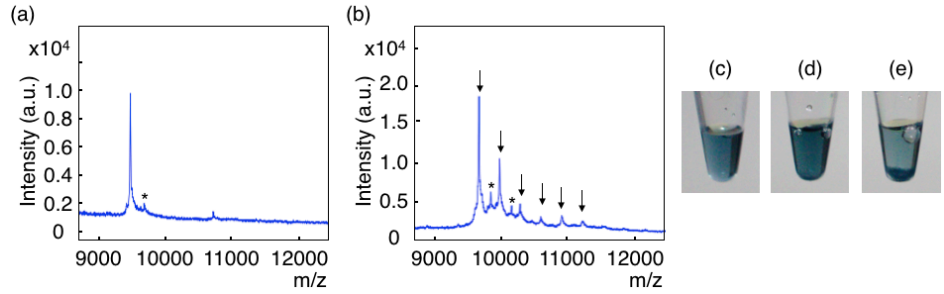


Figure 4.3. MALDI-TOF mass spectra of (a) XTEN-noam and (b) XTEN-E49am after electropolymerization with EDOT-S. Samples were collected by dissolving the deposited films on the FTO electrodes into water. Arrows indicate the peak positions corresponding to the addition of EDOT-S to the protein (XTEN-E49am–nEDOT-S; n = 0-5), and * indicates the proteins without Met cleavage. (c, d, e) Sample solutions after electropolymerization in 0.2 mL PCR tubes. (c) EDOT-S, (d) EDOT-S with XTEN-noam, and (e) EDOT-S with XTEN-E49am.

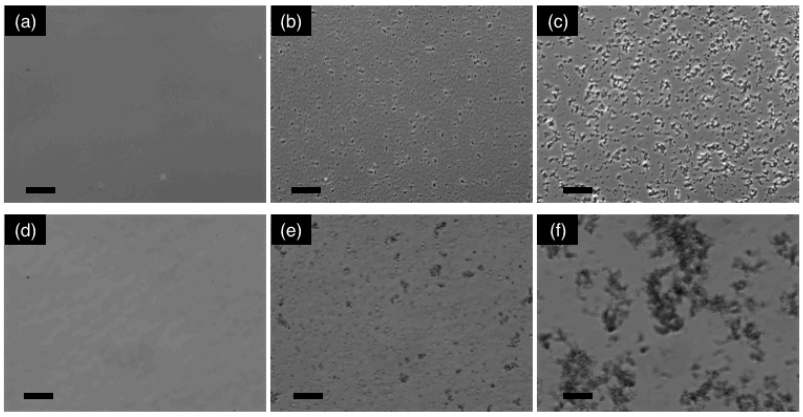


Figure 4.4. Optical microscope images of sample solutions after electropolymerization. (a, e) EDOT-S, (b, f) EDOT-S with XTEN-noam, and (c) EDOT-S with XTEN-E49am. Scale bar: (a-c) 100 µm, (d-f) 25 µm. Solutions were cast between two glass cover slips with a 120 µm spacer.

To further confirm that the aggregation is formed specifically from the protein with the pendant EDOT group, we prepared an internal standard protein for the analysis by mass spectrometry. The standard protein was prepared by deleting six residues in XTEN-noam (designated as sXTEN-noam). sXTEN-noam has a smaller molecular weight than XTEN-noam and its mass peak is well-separated from the one of XTEN-E49am. Therefore, sXTEN-noam can be used as a reference in mass spectra to evaluate the amount of XTEN-E49am remaining in the supernatant after

electropolymerization. We prepared a solution of EDOT-S with a 1:1 mixture of sXTEN-noam and XTEN-E49am (0.4 mM, respectively), and electropolymerized on FTO. The solution formed a precipitate after electropolymerization. We collected the supernatant and analyzed the ratio of peak intensities of sXTEN-noam and XTEN-E49am by MALDI-TOF mass spectrometry to evaluate the ratio of sXTEN-noam and XTEN-E49am after electropolymerization (Figure 4.5). Notably, the ratio of protein peaks (XTEN-E49am/sXTEN-noam) decreased by 87% after polymerization (n = 3). This indicates that XTEN-E49am polymerized with EDOT-S and precipitated as the dark blue solid, while sXTEN-noam largely remained in the supernatant.

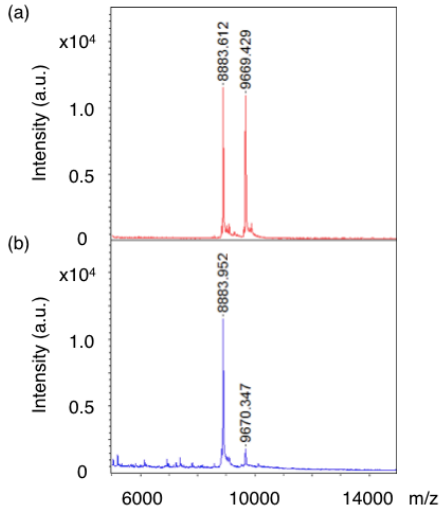


Figure 4.5. MALDI-TOF mass spectra for a 1:1 mixture of sXTEN-noam and XTEN-E49am (a) before and (b) after electropolymerization with EDOT-S. After electropolymerization, the supernatant was collected from the rection mixture and subjected to the analysis. Sinapic acid (SA) was used as a matrix.

Generality of the protein polymerization was investigated by introducing EDOT-Lys at a different position. XTEN carrying EDOT-Lys at position 5 (XTEN-K5am) formed the dark blue precipitate when electropolymerized with EDOT-S (Figure S4.3a-c). MALDI-TOF mass spectrometry with the internal standard sXTEN-noam revealed that 83% (n = 3) of XTEN-K5am was depleted from the solution upon electopolymerization (Figure S4.3d).

4.3.2. FT-IR Spectroscopy for the Polymer Products

We further analyzed the polymer products by FT-IR spectroscopy (Figure 4.6a). The solid product and the supernatant were separated by centrifugation, and the solid was washed with water to remove soluble PEDOT-S. The supernatant, solid product, and electropolymerized PEDOT-S were air dried at room temperature, mixed with KBr to form pellets, and subjected to FT-IR spectroscopy. Electropolymerized PEDOT-S showed peaks at around 1,683, 1,637, 1,180, 1,144, 1,120, 1,109, 1,084, and 1,051 cm^{-1}. The strongest peaks at 1,120 and 1,084 cm^{-1} are attributed to the asymmetric and symmetric S=O stretching modes in the sulfonic acid side chain.[25] The peak at 1,683 cm^{-1} is attributed to the stretching modes of C=C in the thiophene ring.[26] On the other hand, the spectrum of pure XTEN-E49am showed characteristic amide peaks at 1,653 (amide I) and 1,539 (amide II) cm^{-1}. The supernatant and the solid products from electropolymerization of EDOT-S and XTEN-E49am showed peaks originating from PEDOT-S. In addition, the solid product showed a strong protein peak at 1522 cm^{-1}, which suggests that XTEN-E49am was incorporated into the solid polymer product. The supernatant only showed a small peak at 1522 cm^{-1}, which is consistent with the depletion of XTEN-E49am in the soluble fraction.

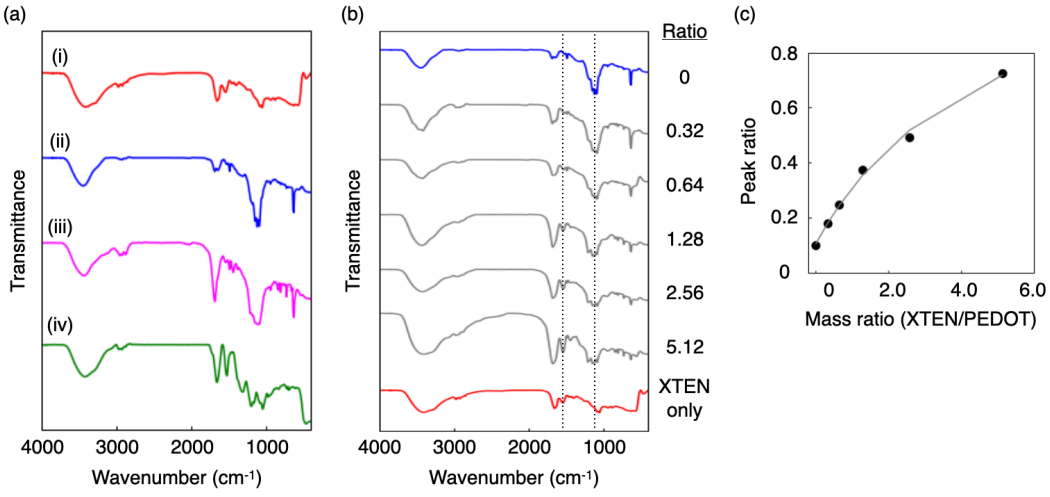


Figure 4.6. FT-IR spectroscopy of the polymer products. (a) FT-IR spectra of (i) XTEN-E49am, (ii) electropolymerized PEDOT-S, (iii) the supernatant, and (iv) the solid product in the reaction mixture of EDOT-S and XTEN-E49am. (b) FT-IR spectra of mixtures of XTEN-E49am and PEDOT-S mixed at the indicated mass ratio. Dot lines indicate the characteristic peaks of XTEN-E49am (1522 cm^{-1}) and PEDOT-S (1084 cm^{-1}). (c) The intensity ratios of peaks of XTEN-E49am (1522 cm^{-1}) and PEDOT-S (1084 cm^{-1}) plotted against mass ratio (XTEN-E49am/PEDOT-S). The curve was fit with the following functional form: y = (ax + b)/(cx + d). (a = 2642, b = 1059, c = 2027, d = 9937).

In order to determine the protein and PEDOT-S composition of the solid product, we prepared a calibration curve using the mixture of a fixed amount of PEDOT-S and variable amount of XTEN-E49am (Figure 4.6b). We observed a stronger protein peak at 1522 cm^{-1} relative to the peak of PEDOT-S (1,084 cm^{-1}) as the amount of XTEN-E49am increased. The ratios of peak intensities (1522 cm^{-1}/1,082 cm^{-1}) were plotted against mass ratios of XTEN-E49am and PEDOT-S (Figure 4.6c) to create a calibration curve. The mass composition of XTEN-E49am and PEDOT-S in the solid product was calculated as 73:27. This indicates that one protein is found per 13 monomer units of the PEDOT chain.

4.3.3. Conductivity of the XTEN–PEDOT Conjugate

We studied the electrical conductivity of the solid product that was electropolymerized from the mixture of EDOT-S and XTEN-E49am. In order to measure the sheet resistance of the sample in a two-electrode configuration,[27] we deposited two parallel-aligned Pt electrodes on a glass substrate (Figure 4.7a). The purified solid product was dispersed in water, cast between the two Pt electrodes, and air dried overnight (Figure 4.7b). A potential sweep (E_{low} = 0 V; E_{high} = +0.1 V) between the two electrodes showed linear increase in the current (Figure 4.7c), which gave a sheet resistance of (2.6 ± 0.7) x 10^2 kΩ (n = 4). Conductivity is given by $\sigma = \frac{1}{\rho} = \frac{L}{RWt}$, where ρ is resistivity, R is resistance, L is the spacing between the two electrodes, W is the width of the sample, and t is the thickness of the sample, and calculated as (4.6 ± 1.7) x 10^{-3} S/cm (n = 4).

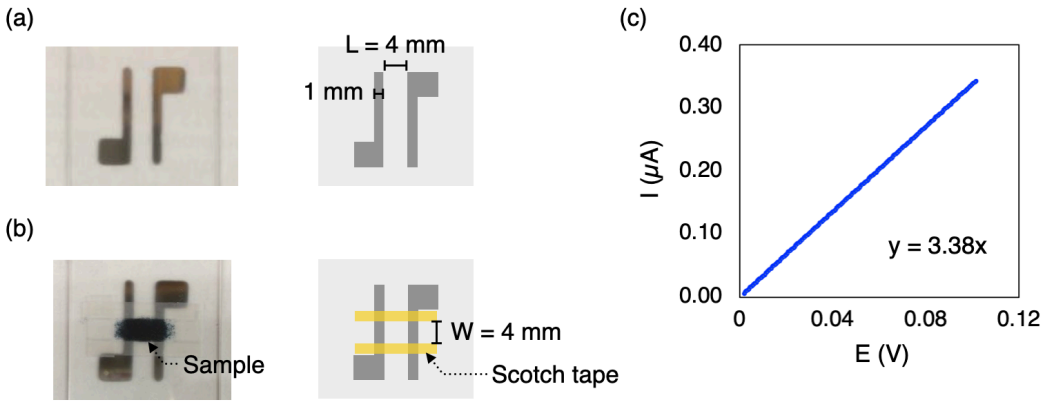


Figure 4.7. Conductivity measurement. (a) Two parallel-aligned Pt electrodes deposited on a glass substrate. Thickness of the Pt electrodes is 10 nm. (b) The solid product of electropolymerization was cast in the defined area between the two electrodes. The area was defined by the scotch tape attached with a spacing of 4 mm. Sample thickness was 8.9 ± 1.3 μm (n = 4). (c) A representative I-E curve of the solid product from electropolymerization of XTEN-E49am with EDOT-S. Average resistance and conductivity were calculated as (2.6 ± 0.7) x 10^2 kΩ and (4.6 ± 1.7) x 10^{-3} S/cm (n = 4), respectively.

4.3.4. Effect of Adjacent Amino Acids

The chemical environment around EDOT-Lys might have an impact on electropolymerization. To investigate the effect of the amino acid adjacent to EDOT-Lys, we prepared a series of short peptides as a model and evaluated their electropolymerization. Peptides with EDOT-Lys are readily synthesized from Fmoc-protected peptides carrying Lys residue via the reaction with EDOT-NHS. We prepared a series of model peptides with the sequence of ST**X**Z̲S (X = EDOT-Lys, Z = Cys, Glu, Gly, His, Lys, Met, Pro, Gln, Arg, and Trp) and performed electropolymerization by cyclic voltammetry (Figure 4.8). Peptides carrying Glu, Gly, His, Pro and Gln adjacent to EDOT-Lys showed the onset of an oxidation current at ~1.1 V and the solutions turned blue upon oxidation. When Cys, Lys, or Met was located adjacent to EDOT-Lys, the peptides did not show a current response, and no color change was observed in the solution. The peptide with Trp showed the initiation of the current production at ~0.9 V, which corresponds to the oxidation potential of the indole group in Trp. The solution revealed no color

change, indicating that polymerization was inhibited by the presence of Trp adjacent to EDOT-Lys. Interestingly, the peptide carrying Arg showed a higher oxidation potential (~1.3V) compared to the other peptides that proceeded polymerization (~1.1 V).

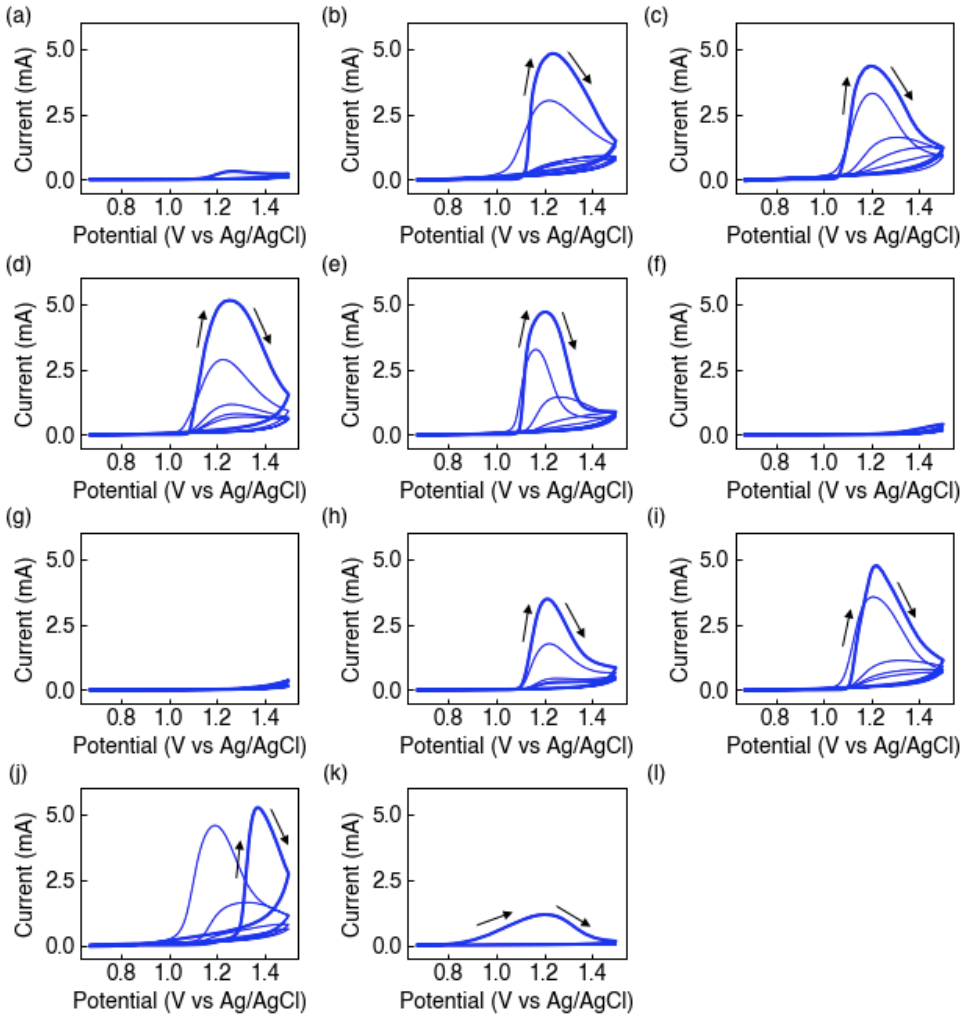


Figure 4.8. Electropolymerization of model peptides. (a) Chemical structure of model peptides (ST<u>X</u>**Z**S; X = EDOT-Lys, Z = indicated amino acid). Cyclic voltammograms of model peptides carrying the adjacent amino acid Z = (a) Cys, (b) Asp, (c) Glu, (d) Gly, (e) His, (f) Lys, (g) Met, (h) Pro, (i) Gln, (j) Arg, and (k) Trp. Peptide solutions were prepared at 50 mM in H_2O with 100 mM $NaClO_4$ and 5% TFA. The potential was swept between +0.65 V and +1.5 V (vs Ag/AgCl) at a scan rate of 100 mV/s.

Because the amino acids Cys, Lys, Met, Arg, and Trp showed influence on electropolymerization of the model peptides, we analyzed the effect of these amino acids in the context of polymerization of the XTEN protein and EDOT-S. We prepared a series of protein variants XTEN-E49am-G50Z (Z = Cys, Lys, Met, Arg, Trp) and performed electropolymerization with EDOT-S by cyclic voltammetry (Figure S4.4). Notably, all the samples showed an oxidation current with the onset of current production at ~1.1 V, similar to electropolymerization of XTEN-E49am with EDOT-S. The solutions turned dark blue during the potential cycles, which indicates the formation of a PEDOT chain. Dark blue precipitates were formed in all the samples (Figure 4.9a), and aggregates were observed under an optical microscope (Figure 4.9b). Polymerization of the XTEN variants without EDOT-Lys (XTEN-noam-G50Z (Z = Cys, Lys, Met, Arg, Trp)) formed no precipitate in the solutions (Figure 4.9a). These results indicate that proteins carrying EDOT-Lys can polymerize with EDOT-S regardless of the presence of adjacent Cys, Lys, Met, Arg, and Trp residues, whereas these residues inhibited peptide polymerization or increased the oxidation potential (Figure 4.8). A possible explanation of this is that in protein polymerization, the EDOT-S monomer supplemented to assist protein polymerization enabled formation of radical species to initiate the reaction. In contrast, EDOT-S was not added in the peptide polymerization experiment, and the initial oxidation step was hampered by the adjacent Cys, Lys, Met, Arg, and Trp residues.

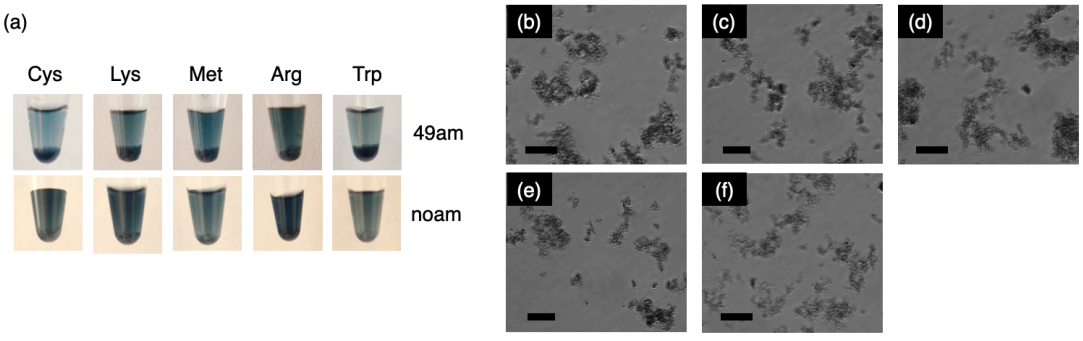


Figure 4.9. Electropolymerization of XTEN-E49am-G50Z. (a) Solutions after electropolymerization of XTEN-E49am-G50Z or XTEN-noam-G50Z (Z = Cys, Lys, Met,

Arg, Trp) with EDOT-S. (b-f) Optical microscope images of sample solutions after electropolymerization of XTEN-E49am-G50Z with EDOT-S. Z = (b) Cys, (c) Lys, (d) Met, (e) Arg, and (f) Trp. Scale bar: 25 μm. Solutions were cast between two glass cover slips with a 120 μm spacer.

4.4. Conclusion

We performed electropolymerization of XTEN proteins carrying EDOT-Lys. The solution containing XTEN-E49am and EDOT-S monomer yielded dark blue solids upon the application of the oxidation potential. MALDI-TOF mass spectrometry with an internal standard protein revealed that 85% of the protein was consumed in the reaction mixture. FT-IR spectroscopy indicated that the solid product contained both protein and PEDOT-S (one protein in 12.5 monomer units). The effect of adjacent amino acids was investigated in the model peptide and XTEN proteins. These results suggest that genetic incorporation of EDOT-Lys can be used for the synthesis of protein–PEDOT conjugates and provides a versatile strategy to fabricate protein-based conductive materials.

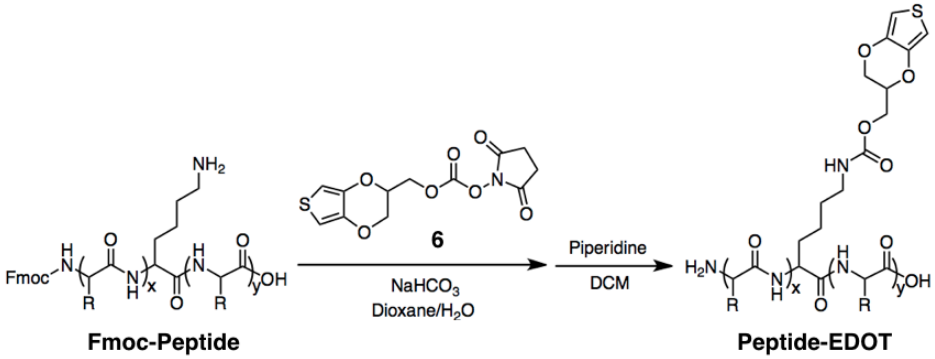
4.5. Experimental Procedures

4.5.1. General

All the reagents for organic synthesis were purchased from Sigma Aldrich and used as received without further purification. Fmoc-protected peptides were custom synthesized by GenScript. For column chromatography, Biotage Isolera Spektra equipped with a SNAP ULTRA C18 12 g cartridge was used. ^{1}H (500 MHz) and ^{13}C NMR (126 MHz) spectra were recorded on a Varian Inova 500 spectrometer. High-resolution mass spectrometry (HRMS) was performed with an LCT Premier XE Electrospray TOF Mass Spectrometer with electrospray ionization (ESI) at the Caltech CCE Multiuser Mass Spectrometry Laboratory. Matrix-assisted laser desorption/ionization time-of-flight mass (MALDI-TOF MS) spectrometry measurements were carried out on a Bruker Daltonics autoflex^{TM} speed MALDI-

TOF/TOF spectrometer at the Caltech CCE Multiuser Mass Spectrometry Laboratory. IR absorption data was collected on a Thermo Nicolet Nexus FT-IR spectrometer using an LN_2-cooled MCT-A detector. The sample chamber was purged with nitrogen.

4.5.2. Synthesis and Characterization of the Peptides



Peptides with EDOT group. General procedure for synthesis of peptides with EDOT group is as follows: Fmoc-protected peptide with Lys (100 mg) and sodium bicarbonate (48 mg, 0.57 mmol) were dissolved in dioxane:H_2O (4:1, 1 mL). EDOT-NHS (6) (60 mg, 0.19 mmol) was added to the solution and the resulting mixture was allowed to stir at room temperature for 18 h. Then, DCM (0.6 mL) and piperidine (0.4 mL) were added and the solution was stirred for 16 h. After the completion of the reaction, the solution was poured into excess ether. The white precipitate was washed with ether three times, air dried, and purified by C18 reversed-phase silica column chromatography (H_2O:methanol 98:2 to 0:100).

STXCS. HRMS (ESI) calculated for $C_{27}H_{42}N_6O_{13}S_2$ $[M+H]^+$: *m/z* = 723.2329; found: 723.2302
STXDS. HRMS (ESI) calculated for $C_{28}H_{42}N_6O_{15}S$ $[M+H]^+$: *m/z* = 735.2507; found: 735.2501
STXES. HRMS (ESI) calculated for $C_{29}H_{44}N_6O_{15}S$ $[M+H]^+$: *m/z* = 749.2663; found: 749.2664
STXFS. HRMS (ESI) calculated for $C_{33}H_{46}N_6O_{13}S$ $[M+H]^+$: *m/z* = 767.2922; found: 767.2941
STXGS. HRMS (ESI) calculated for $C_{26}H_{40}N_6O_{13}S$ $[M+H]^+$: *m/z* = 677.2452; found: 677.2452
STXHS. HRMS (ESI) calculated for $C_{30}H_{44}N_8O_{13}S$ $[M+H]^+$: *m/z* = 757.2827; found: 757.2816
STXKS. A peptide Fmoc-STK(FmocK)S was used as starting material. HRMS (ESI) calculated for $C_{30}H_{49}N_7O_{13}S$ $[M+H]^+$: *m/z* = 748.3187; found: 748.3171

STXMS. HRMS (ESI) calculated for $C_{29}H_{46}N_6O_{13}S_2$ $[M+H]^+$: *m/z* = 751.2642; found: 751.2642
STXPS. HRMS (ESI) calculated for $C_{29}H_{44}N_6O_{13}S$ $[M+H]^+$: *m/z* = 717.2765; found: 717.2755
STXQS. HRMS (ESI) calculated for $C_{29}H_{45}N_7O_{14}S$ $[M+H]^+$: *m/z* = 748.2823; found: 748.2808
STXRS. HRMS (ESI) calculated for $C_{30}H_{49}N_9O_{13}S$ $[M+H]^+$: *m/z* = 776.3249; found: 776.3256
STXWS. HRMS (ESI) calculated for $C_{35}H_{47}N_7O_{13}S$ $[M+H]^+$: *m/z* = 806.3031; found: 806.3049
STXYS. HRMS (ESI) calculated for $C_{33}H_{46}N_6O_{14}S$ $[M+H]^+$: *m/z* = 783.2871; found: 783.2883

*The peptides (Z = F, Y) were insoluble in aqueous solutions and not used for the electropolymerization experiment.

4.5.3. Cloning, Protein Expression, and Purification

Cloning. The plasmids were constructed by standard recombinant DNA technology using *E. coli* DH10B strain. DNA oligomers for PCRs were purchased from IDT (Coralville, IA). Bacteria were grown on LB/agar plates and LB liquid media with the following antibiotic concentrations: 100 μg/mL spectinomycin and 100 μg/mL ampicillin. The plasmid encoding XTEN was a kind gift from Dr. Peter Rapp. Plasmids and their corresponding coding sequences are presented in *4.7. DNA and Protein Sequences*.

pET16b-XTEN Plasmids. The pET16b vector (Novagen) confers resistance to ampicillin and carries an *N*-terminal 10xHis-Tag and cloning sites under the T7 promoter. The XTEN gene was cloned to a cloning site by Gibson Assembly. *N*-terminal 10xHis-Tag was removed, and 6xHis-Tag was introduced to the *C*-terminus. A modified FLAG-Tag was introduced to the *N*-terminus to improve protein expression. An amber stop codon was introduced at the indicated position by site-directed mutagenesis.

Protein Expression and Purification. To express XTEN carrying EDOT-Lys, *E. coli* strain BL21(DE3) was co-transformed with pUltra-*Mm*PylRS(Y306A/Y384F) and a pET plasmid encoding XTEN. *E. coli* cultures were grown in 2xYT media at

37 °C to an optical density at 600 nm (OD600) of 0.5-0.7. Then 1 mM EDOT-Lys was added, and protein expression was induced by addition of 1 mM IPTG. After 15 h, the cells were harvested by centrifugation, resuspended in a lysis buffer (0.1 M Na_2HPO_4, 10 mM imidazole; pH 8.0), and lysed by sonication. Lysates were cleared by centrifugation, and incubated with Ni-NTA agarose. The resin was washed with a lysis buffer and wash buffer (0.1 M Na_2HPO_4, 25 mM imidazole; pH 8.0). Protein was eluted with elution buffer (0.1 M Na_2HPO_4, 250 mM imidazole; pH 8.0), dialyzed against water, and lyophilized for storage.

To express the control XTEN without EDOT-Lys, *E. coli* strain BL21(DE3) was transformed with a pET plasmid encoding XTEN. *E. coli* cultures were grown in 2xYT media at 37 °C to an optical density at 600 nm (OD600) of 0.5-0.7. Then protein expression was induced by addition of 1 mM IPTG. After 6 h, the cells were harvested by centrifugation, resuspended in lysis buffer (0.1 M Na_2HPO_4, 10 mM imidazole; pH 8.0), and lysed by sonication. Lysates were cleared by centrifugation and incubated with Ni-NTA agarose. The resin was washed with lysis buffer and wash buffer (0.1 M Na_2HPO_4, 25 mM imidazole; pH 8.0). Protein was eluted with elution buffer (0.1 M Na_2HPO_4, 250 mM imidazole; pH 8.0), dialyzed against water, and lyophilized for storage.

4.5.4. Electropolymerization

Electropolymerization was performed by cyclic voltammetry in a three-electrode configuration using a fluorine-doped tin oxide (FTO) working electrode, Ag/AgCl reference electrode, and Pt counter electrode. The Ag/AgCl electrode was placed in a pipette tip filled with 1% agarose gel in 3M NaCl. The FTO electrodes were cleaned by sonication in acetone, ethanol, and deionized water, and dried in the air. The SecureSeal Imaging Spacer (d = 9 mm; Grace Bio Labs) was attached to the FTO electrodes to define the conductive surface area (0.025 m^2). Sample solution (40 μL) was cast on the FTO electrode, and the Ag/AgCl and the Pt electrodes were immersed in the solution. Five potential cycles were applied between +0.65 V and

+1.5 V vs Ag/AgCl at a scan rate of 100 mV/s. All the experiments were carried out at room temperature.

4.5.5. FT-IR Spectroscopy

IR absorption measurement. The reaction mixture of electropolymerization of EDOT-S (25 mM) and XTEN-E49am (0.9 mM) was centrifuged, and the precipitate and the supernatant were collected separately. The precipitate was washed with water three times to remove water-soluble PEDOT-S. The supernatant and the solid product were air dried at room temperature, mixed with KBr, and subjected to FT-IR spectroscopy.

Composition analysis. Standard samples were prepared at 0:1, 0.32:1, 0.64:1, 1.28:1, and 2.56:1 mass ratio of XTEN-E49am and electropolymerized PEDOT-S. Samples were dried in the air at room temperature and mixed with KBr to form pellets. The absorbance values of characteristic peaks of XTEN-E49am (1522 cm^{-1}) and PEDOT-S (1084 cm^{-1}) were used to create a calibration curve.

4.5.6. Conductivity Measurement

Two-electrode sheet resistance measurement was performed following a reported procedure with minor modifications.[27] Two parallel-aligned Pt electrodes (thickness = 10 nm) were deposited on a glass substrate using a plastic mask prepared by a 3D printer. After electropolymerization of XTEN-E49am with EDOT-S, the solid products were isolated by centrifugation and washed with water five times to remove water-soluble PEDOT-S. The solid products were resuspended in water and cast between the two Pt electrodes deposited on the glass substrate. The samples were dried in the air at room temperature overnight and a potential sweep (E_{low} = 0 V; E_{high} = +0.1 V) was applied to measure the current response. Four independent samples were analyzed. Sample thickness was measured using a DekTakXT stylus profiler (Bruker).

4.6. Supporting Figures

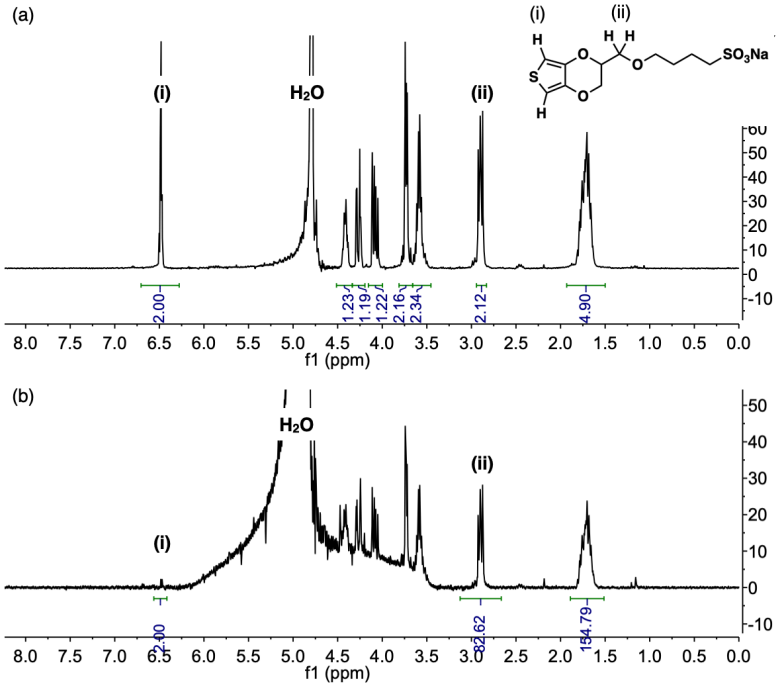


Figure S4.1. Study of monomer consumption by NMR spectroscopy. ^{1}H NMR spectrum (500 MHz) of EDOT-S in D_2O (a) before and (b) after electropolymerization. Monomer consumption was calculated as 97% based on the reduction in the integration value of (i) the thiophene peaks relative to that of (ii) the methylene peaks. Related to Figure 4.2a. Electropolymerization was conducted by cyclic voltammetry on FTO electrodes (E_{low} = +0.65 V; E_{high} = +1.5 V; scan rate = 100 mV/s; 5 cycles). The sample for electropolymerization was prepared at [EDOT-S] = 25 mM, [$NaClO_4$] = 100 mM in D_2O with 5% TFA-d.

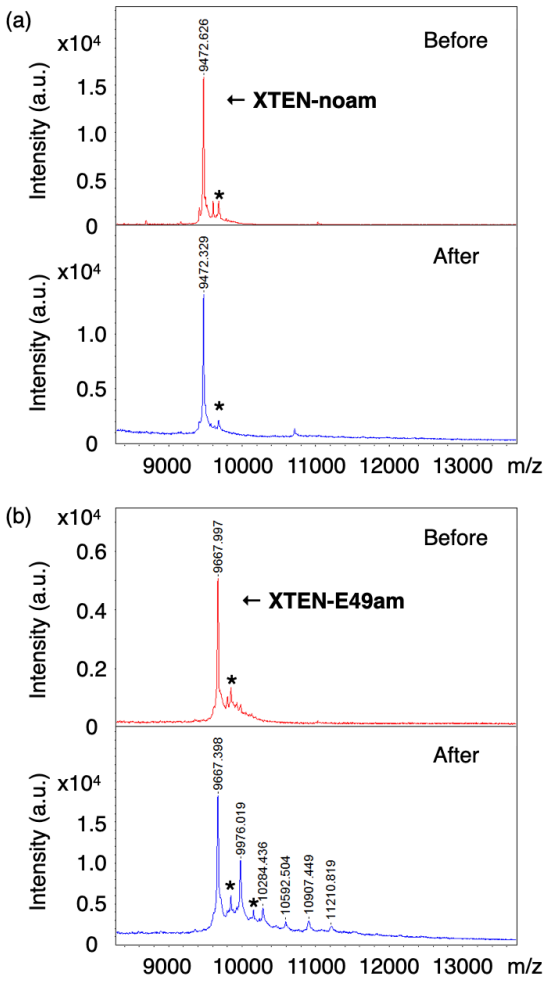


Figure S4.2. MALDI-TOF mass spectra for (a) XTEN-noam and (b) XTEN-E49am before and after electropolymerization with EDOT-S. The major peaks correspond to the molecular weight of methionine-cleaved proteins (*indicates the peaks from proteins without methionine cleavage). Samples after electropolymerization were collected by dissolving the deposited films on the FTO electrodes into water. Sinapic acid (SA) was used as a matrix.

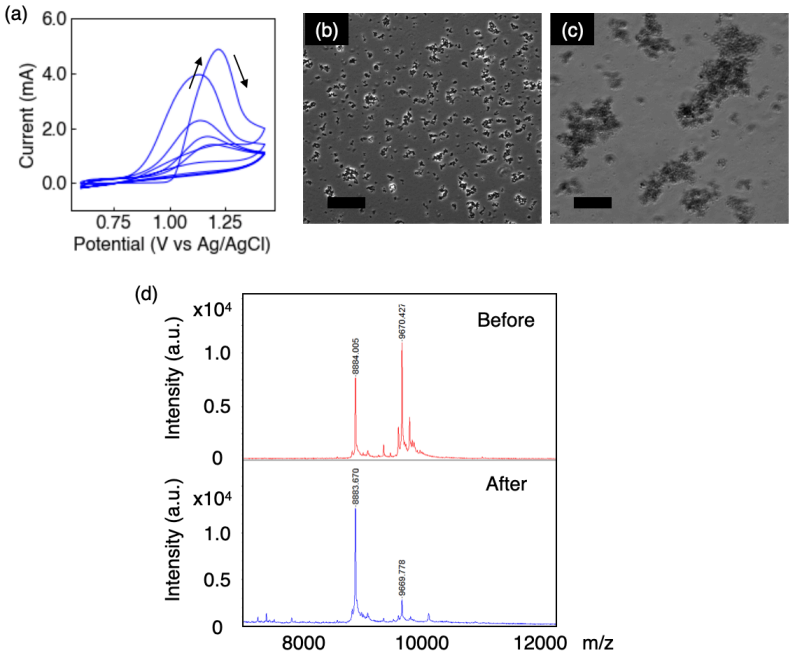


Figure S4.3. Electropolymerization of EDOT-S with XTEN-49am-K5am by cyclic voltammetry on FTO electrodes. (a) Cyclic voltammogram (E_{low} = +0.65 V; E_{high} = +1.5 V; scan rate = 100 mV/s; 5 cycles). Sample was prepared at [XTEN-K5am] = 0.8 mM, [EDOT-S] = 25 mM, [$NaClO_4$] = 100 mM in H_2O with 5% TFA. Arrows indicate the scan direction. (b, c) Optical microscope images of the solution after electropolymerization of EDOT-S with XTEN-49am-K5am. Scale bar: (b) 100 µm, (c) 25 µm. (d) MALDI-TOF mass spectra for a 1:1 mixture of sXTEN-noam and XTEN-K5am before and after electropolymerization with EDOT-S. After electropolymerization, the supernatant was collected from the reaction mixture and subjected to the analysis. Sinapic acid (SA) was used as a matrix.

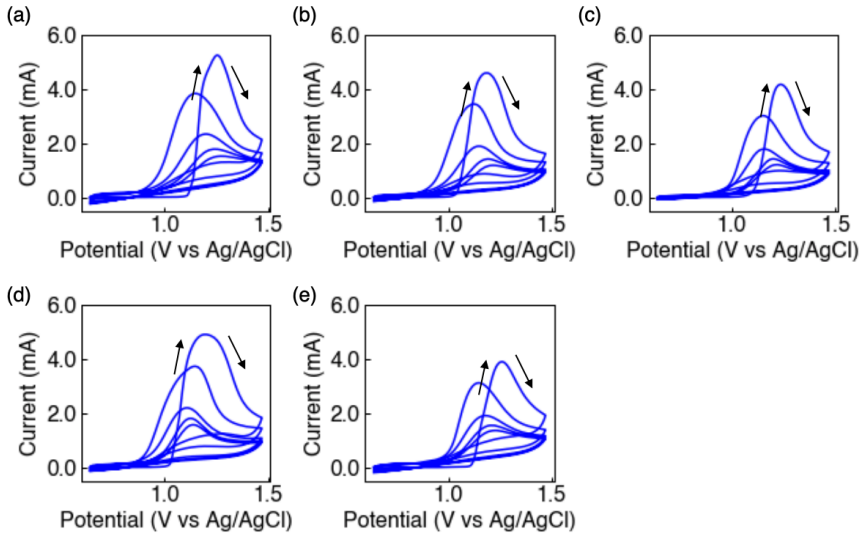


Figure S4.4. Electropolymerization of EDOT-S with XTEN-E49am-G50Z by cyclic voltammetry on FTO electrodes (E_{low} = 0.65 V; E_{high} = +1.5 V; scan rate = 100 mV/s; 5 cycles). Z = (a) Cys, (b) Lys, (c) Met, (d) Arg, and (e) Trp. Samples are prepared at [XTEN] = 0.8 mM, [EDOT-S] = 25 mM, [$NaClO_4$] = 100 mM in H_2O with 5% TFA. Arrows indicate the scan direction.

4.7. DNA and Protein Sequences

XTEN-noam:

atgggccatgac**aaa**gacgatgacgacaagggatccgtcgacgcgggttccggctccccggctggca
gcccgaccagcactgaagagggcacgagcgagtcggcgaccccggagtctggtccgggcacctccac
cgaaccgtct**gag**ggcagcgcaccgggtagcccggccggtagccctaccagcaccgaagagggtacc
agcacggaaccgagcgaaggctcggcaccg**ggtacgagcaccgagatc**gaaggtcgtctcgag**catc**
atcaccaccaccataagcttaattaa

MGHD**K**DDDDKGSVDAGSGSPAGSPTSTEEGTSESATPESGPGTSTEPS**E**GSAPGSPAGSPTSTEEGT
STEPSEGSAP**GTSTEI**EGRLE**HHHHHH**KLN

*The positions 5 and 49 are highlighted in red. *C*-terminal 6xHis-Tag is highlighted in blue. The six amino acids deleted in sXTEN-noam are highlighted with underlining.

4.8. References

(1) Connor, R. E.; Tirrell, D. A. Non-Canonical Amino Acids in Protein Polymer Design. *Polym. Rev.* **2007**, *47* (1), 9–28.

(2) Johnson, J. A.; Lu, Y. Y.; Van Deventer, J. A.; Tirrell, D. A. Residue-Specific Incorporation of Non-Canonical Amino Acids into Proteins: Recent Developments and Applications. *Curr. Opin. Chem. Biol.* **2010**, *14* (6), 774–780.

(3) Groenendaal, L.; Jonas, F.; Freitag, D.; Pielartzik, H.; Reynolds, J. R. Poly(3,4-Ethylenedioxythiophene) and Its Derivatives: Past, Present, and Future. *Adv. Mater.* **2000**, *12* (7), 481–494.

(4) Mantione, D.; del Agua, I.; Sanchez-Sanchez, A.; Mecerreyes, D. Poly(3,4-Ethylenedioxythiophene) (PEDOT) Derivatives: Innovative Conductive Polymers for Bioelectronics. *Polymers (Basel).* **2017**, *9* (8).

(5) Kirchmeyer, S.; Reuter, K. Scientific Importance, Properties and Growing Applications of Poly(3,4-Ethylenedioxythiophene). *J. Mater. Chem.* **2005**, *15* (21), 2077–2088.

(6) Ramalingam, K.; Panchu, S. J.; Salunke, A. S.; Muthukumar, K.; Ramanujam, A.; Muthiah, S. Free-Standing Graphene/Conducting Polymer Hybrid Cathodes as FTO and Pt-Free Electrode for Quasi-State Dye Sensitized Solar Cells. *ChemistrySelect* **2016**, *1* (15), 4814–4822.

(7) Poverenov, E.; Li, M.; Bitler, A.; Bendikov, M. Major Effect of Electropolymerization Solvent on Morphology and Electrochromic Properties of PEDOT Films. *Chem. Mater.* **2010**, *22* (13), 4019–4025.

(8) Seki, Y.; Takahashi, M.; Takashiri, M. Effects of Different Electrolytes and Film Thicknesses on Structural and Thermoelectric Properties of Electropolymerized Poly(3,4-Ethylenedioxythiophene) Films. *RSC Adv.* **2019**, *9* (28), 15957–15965.

(9) Kanungo, M.; Srivastava, D. N.; Kumar, A.; Contractor, A. Q. Conductimetric Immunosensor Based on Poly(3,4-Ethylenedioxythiophene). *Chem. Commun.* **2002**, *2* (7), 680–681.

(10) Wen, Y.; Xu, J.; He, H.; Lu, B.; Li, Y.; Dong, B. Electrochemical Polymerization of 3,4-Ethylenedioxythiophene in Aqueous Micellar Solution Containing Biocompatible Amino Acid-Based Surfactant. *J. Electroanal. Chem.* **2009**, *634* (1), 49–58.

(11) Bodart, C.; Rossetti, N.; Hagler, J.; Chevreau, P.; Chhin, D.; Soavi, F.; Schougaard, S. B.; Amzica, F.; Cicoira, F. Electropolymerized Poly(3,4-

Ethylenedioxythiophene) (PEDOT) Coatings for Implantable Deep-Brain-Stimulating Microelectrodes. *ACS Appl. Mater. Interfaces* **2019**, *11* (19), 17226–17233.

(12) Asplund, M.; von Holst, H.; Inganäs, O. Composite Biomolecule/PEDOT Materials for Neural Electrodes. *Biointerphases* **2008**, *3* (3), 83–93.

(13) Xiao, Y.; Li, C. M.; Wang, S.; Shi, J.; Ooi, C. P. Incorporation of Collagen in Poly(3,4-Ethylenedioxythiophene) for a Bifunctional Film with High Bio- and Electrochemical Activity. *J. Biomed. Mater. Res. - Part A* **2010**, *92* (2), 766–772.

(14) Xu, F.; Ren, S.; Gu, Y. A Novel Conductive Poly (3, 4-Ethylenedioxythiophene)-BSA Film for the Construction of a Durable HRP Biosensor Modified with NanoAu Particles. *Sensors (Switzerland)* **2016**, *16* (3), 10–13.

(15) Stéphan, O.; Schottland, P.; Le Gall, P. Y.; Chevrot, C.; Mariet, C.; Carrier, M. Electrochemical Behaviour of 3,4-Ethylenedioxythiophene Functionalized by a Sulphonate Group. Application to the Preparation of Poly(3,4-Ethylenedioxythiophene) Having Permanent Cation-Exchange Properties. *J. Electroanal. Chem.* **1998**, *443* (2), 217–226.

(16) Zotti, G.; Zecchin, S.; Schiavon, G.; Groenendaal, L. B. Electrochemical and Chemical Synthesis and Characterization of Sulfonated Poly (3,4-Ethylenedioxythiophene): A Novel Water-Soluble and Highly Conductive Conjugated Oligomer. *Macromol. Chem. Phys.* **2002**, *203* (13), 1958–1964.

(17) Hamedi, M.; Herland, A.; Karlsson, R. H.; Lnganäs, O. Electrochemical Devices Made from Conducting Nanowire Networks Self-Assembled from Amyloid Fibrils and Alkoxysulfonate PEDOT. *Nano Lett.* **2008**, *8* (6), 1736–1740.

(18) Herland, A.; Persson, K. M.; Lundin, V.; Fahlman, M.; Berggren, M.; Jager, E. W. H.; Teixeira, A. I. Electrochemical Control of Growth Factor Presentation to Steer Neural Stem Cell Differentiation. *Angew. Chemie - Int. Ed.* **2011**, *50* (52), 12529–12533.

(19) Müller, C.; Jansson, R.; Elfwing, A.; Askarieh, G.; Karlsson, R.; Hamedi, M.; Rising, A.; Johansson, J.; Inganäs, O.; Hedhammar, M. Functionalisation of Recombinant Spider Silk with Conjugated Polyelectrolytes. *J. Mater. Chem.* **2011**, *21* (9), 2909–2915.

(20) Stavrinidou, E.; Gabrielsson, R.; Gomez, E.; Crispin, X.; Nilsson, O.; Simon, D. T.; Berggren, M. Electronic Plants. *Sci. Adv.* **2015**, *1* (10).

(21) Kim, J. Y.; Nagamani, S.; Liu, L.; Elghazaly, A. H.; Solin, N.; Inganäs, O. A DNA and Self-Doped Conjugated Polyelectrolyte Assembled for Organic Optoelectronics and Bioelectronics. *Biomacromolecules* **2020**, *21* (3), 1214–1221.

(22) Schellenberger, V.; Wang, C. W.; Geething, N. C.; Spink, B. J.; Campbell, A.; To, W.; Scholle, M. D.; Yin, Y.; Yao, Y.; Bogin, O.; Cleland, J. L.; Silverman, J.; Stemmer, W. P. C. A Recombinant Polypeptide Extends the in Vivo Half-Life of Peptides and Proteins in a Tunable Manner. *Nat. Biotechnol.* **2009**, *27* (12), 1186–1190.

(23) Podust, V. N.; Sim, B. C.; Kothari, D.; Henthorn, L.; Gu, C.; Wang, C. W.; McLaughlin, B.; Schellenberger, V. Extension of in Vivo Half-Life of Biologically Active Peptides via Chemical Conjugation to XTEN Protein Polymer. *Protein Eng. Des. Sel.* **2013**, *26* (11), 743–753.

(24) Podust, V. N.; Balan, S.; Sim, B. C.; Coyle, M. P.; Ernst, U.; Peters, R. T.; Schellenberger, V. Extension of in Vivo Half-Life of Biologically Active Molecules by XTEN Protein Polymers. *J. Control. Release* **2016**, *240*, 52–

66.

(25) Cutler, C. A.; Bouguettaya, M.; Kang, T. S.; Reynolds, J. R. Alkoxysulfonate-Functionalized PEDOT Polyelectrolyte Multilayer Films: Electrochromic and Hole Transport Materials. *Macromolecules* **2005**, *38* (8), 3068–3074.

(26) Susanti, E.; Wulandari, P.; Herman. Effect of Localized Surface Plasmon Resonance from Incorporated Gold Nanoparticles in PEDOT:PSS Hole Transport Layer for Hybrid Solar Cell Applications. *J. Phys. Conf. Ser.* **2018**, *1080* (1).

(27) Kulhánková, L.; Tokarský, J.; Ivánek, L.; Mach, V.; Peikertová, P.; Matějka, V.; Mamulová Kutláková, K.; Matoušek, J.; Čapková, P. Enhanced Electrical Conductivity of Polyaniline Films by Postsynthetic DC High-Voltage Electrical Field Treatment. *Synth. Met.* **2013**, *179*, 116–121.

Chapter 5

CHEMICAL POLYMERIZATION OF XTEN WITH AN EDOT GROUP

5.1. Abstract

PEDOT is often synthesized either by electropolymerization or oxidative chemical polymerization. While electropolymerization has advantages such as a short reaction time and the requirement of only a small sample volume, chemical polymerization does not require electrochemical instruments and is suitable for large scale synthesis. In this chapter, we describe chemical polymerization of the XTEN protein and model peptides carrying an EDOT group. When XTEN carrying EDOT-Lys was polymerized in the presence of an EDOT-S monomer, the reaction mixture yielded a dark blue precipitate. To demonstrate the reactivity of the pendant EDOT group in the protein, we performed oxidation reaction XTEN carrying EDOT-Lys with an end-capped EDOT derivative (EDOT-cap). Mass spectrometry showed the peaks of protein species that correspond to the XTEN with one or two EDOT-caps, indicating that XTEN can react with the EDOT-cap through the pendant EDOT group. We further studied the effect of the chemical environment on polymerization using a series of model peptides. In the reaction without an $FeCl_3$ catalyst, we found that neighboring basic residues (His, Lys, Arg) enhanced polymerization. These results provide valuable information for fabrication of protein-PEDOT conjugates via oxidative chemical polymerization.

5.2. Introduction

PEDOT and its derivatives are often synthesized by two major approaches: electrochemical polymerization and oxidative chemical polymerization.[1,2] Electrochemical polymerization has advantages such as short reaction time and the requirement of only a small volume of sample solutions. On the other hand, oxidative chemical polymerization does not require electrochemical instruments and is advantageous in the production of polymers at large scales. In this chapter, we studied chemical polymerization of the XTEN protein and model peptides carrying EDOT-Lys.

In the standard method of oxidative chemical polymerization, ammonium persulfate (APS), iron(III) chloride ($FeCl_3$), and iron tosylate ($Fe(OTs)_3$) are commonly used as oxidants.[1,2] In particular, polymerization by persulfates is usually performed in aqueous solution and suitable for polymerization in the presence of biomolecules.[3–5] Addition of a catalytic amount of Fe salts accelerates the reaction.[2,3]

Progress of the polymerization is often visible to the eye because of the characteristic dark blue color of PEDOT and its derivatives.[6] In UV-vis spectroscopy, they exhibit absorption bands at 600 nm (neutral), 800 nm (polaron), and <1200 nm (bipolaron).[7] Polaron is the expected state when the polymer is lightly doped, whereas the bipolaron is the expected state when the polymer is heavily doped.[8] Because the pristine PEDOT is insoluble in any common solvent, typical solution-based analytical techniques (e.g. gel permeation chromatography (GPC), NMR spectroscopy) cannot be used for the characterization of PEDOT. This limitation hampers the detailed analysis of some properties of PEDOT, such as molecular weight.[6,7] The general length of PEDOT chain is believed to be about 10-20 repeating units[9] based on the studies using electron microscopy[10] and MALDI-TOF mass spectrometry[11] as well as the analysis of PEDOT derivatives.[12,13]

5.3. Results and Discussion

5.3.1. Chemical Polymerization of XTEN with an EDOT Group

To study the oxidative chemical polymerization of proteins containing EDOT-Lys, we prepared a model protein XTEN-E49am (XTEN with EDOT-Lys at position 49). XTEN-E49am was mixed with the self-doping EDOT monomer EDOT-S[12,14] in 100 mM HCl, and polymerization was initiated by addition of APS as an oxidant and $FeCl_3$ as a catalyst (Figure 5.1a). The solution gradually turned blue within 10 min. After stirring at room temperature for 18 hours, we observed dark blue precipitates in the solution, whereas polymerization of the EDOT-S monomer and the EDOT-S monomer with XTEN-noam did not form any precipitate. In UV-vis spectra of the reaction mixtures, the electropolymerized product from the EDOT-S monomer exhibited a broad absorption peak in the 600-1000 nm region with increasing absorption towards longer wavelengths, which is characteristic of heavily doped PEDOT species (Figure 5.1b).[7] The product from the mixture of EDOT-S and XTEN-noam showed less absorption in the 600-1000 nm region compared to the product from EDOT-S, suggesting that oxidation was diminished in the presence of the XTEN protein. The product from EDOT-S and XTEN-E49am showed marginal absorption in the 600-1000 nm region, because the dark blue solid product was insoluble and precipitated in the solution (Figure 5.1c).

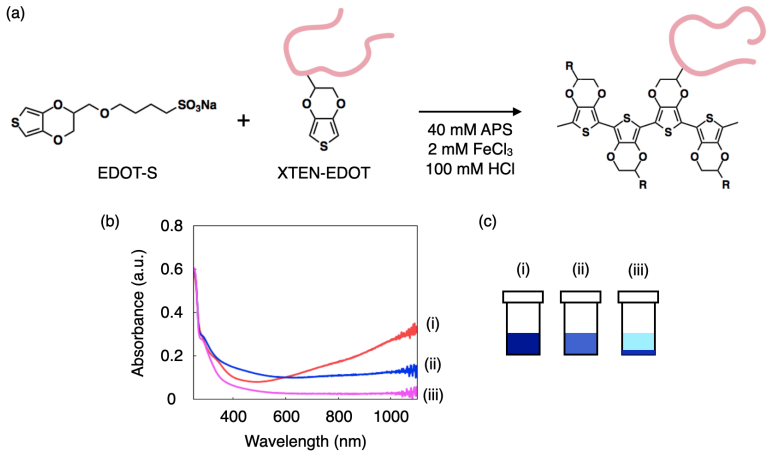


Figure 5.1. Oxidative chemical polymerization of EDOT-S and XTEN. (a) Schematic illustration of chemical polymerization of EDOT-S and XTEN. (b) UV-vis spectra of (i)

EDOT-S, (ii) EDOT-S and XTEN-noam, and (iii) EDOT-S and XTEN-E49am after chemical polymerization. The reaction mixtures were stirred at room temperature for 18 hours, and the supernatant was analyzed by UV-vis spectroscopy.

5.3.2 Reactivity Test with End-capped EDOT

The polymer product from EDOT-S and XTEN-E49am was insoluble in water (pH 2 and pH 9) and in common organic solvents such as DMSO, DMF, THF, methanol, ethanol, hexafluoroisopropanol (HFIP), and acetonitrile. The insoluble nature of the product hampers characterization by solution-based analytical techniques,[7] and therefore, proving the reactivity of EDOT-Lys is challenging. To circumvent this problem, we performed a reactivity test using an end-capped EDOT derivative (EDOT-cap). EDOT-cap carries a reactive α proton on one end and a triethylene glycol-substituted phenyl protecting group on the other end. Therefore, reaction with XTEN carrying EDOT-Lys will yield protein species with the addition of one or two EDOT-caps (Figure 5.2a), which should have better solubility than XTEN-PEDOT and can be analyzed by MALDI-TOF mass spectrometry. The unreacted EDOT-cap and the dimerized byproduct diEDOT-cap are soluble in ethanol while the XTEN protein is insoluble in ethanol. This enables the purification of protein species via ethanol precipitation.

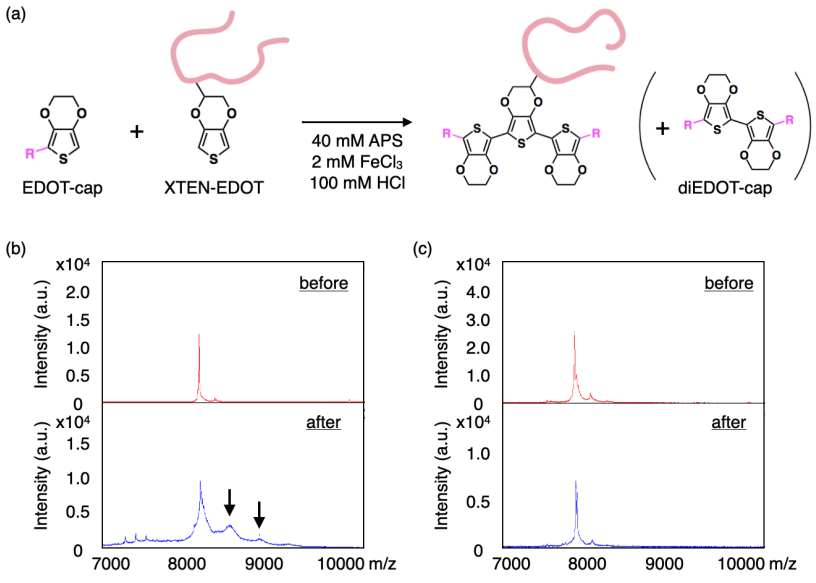


Figure 5.2. Reactivity test with end-capped EDOT (EDOT-cap). (a) Schematic illustration of the reactivity test with EDOT-cap. R: -(C_6H_4)-(OC_2H_4)$_3CH_3$. (b, c) MALDI-TOF mass spectra of (b) Q-XTEN-A15am and (c) Q-XTEN-noam before and after reaction with EDOT-cap. Sinapic acid was used as matrix. Arrows indicate the mass peaks corresponding to the addition of one or two EDOT-caps to Q-XTEN-A15am.

The oxidation reaction was performed in 100 mM HCl using APS as an oxidant and $FeCl_3$ as a catalyst. The solution turned red due to some extension of the π-conjugated system. After 18 hours, the unreacted EDOT-cap and diEDOT-cap were removed by ethanol wash and proteins were analyzed by MALDI-TOF mass spectrometry. XTEN carrying EDOT-Lys (Q-XTEN-A15am) showed the appearance of two protein peaks that correspond to the addition of one or two EDOT-caps to Q-XTEN-A15am (Figure 5.2b). By contrast, XTEN carrying no EDOT-Lys (Q-XTEN-noam) showed no peak that corresponds to the addition of EDOT-cap (Figure 5.2c). These results indicate that XTEN carrying the EDOT-Lys reacted with EDOT-cap through the pendant EDOT group.

5.3.3. Chemical Polymerization of Model Peptides

The chemical environment around the EDOT group can influence the chemical polymerization of proteins. To investigate the effect of adjacent amino acids, we prepared a series of model peptides with the sequence of ST**X**Z̲S (X = EDOT-Lys, Z = Cys, Glu, Gly, His, Lys, Met, Pro, Gln, Arg, Trp) and performed chemical polymerization (Figure 5.3). When the peptides were polymerized by APS without a catalytic amount of $FeCl_3$, peptides with adjacent basic residues (Z = His, Lys, Arg) showed strong absorption peaks at 780 nm, which is characteristic to the formation of PEDOT chains. On the other hand, the peptides with adjacent non-basic residues (Z = Asp, Glu, Gly, Met, Pro, Gln) showed only a little absorption at 780 nm. These results indicate that basic residues enhance chemical polymerization. When the peptides were polymerized by APS with $FeCl_3$ catalyst, all the peptides (Z = Asp, Glu, Gly, His, Lys, Met, Pro, Gln, Arg) showed broad absorption in the 780-1100 nm region, which indicates the formation of highly doped PEDOT.[7]

MALDI-TOF mass spectra of the model peptides (Z = Cys, Asp, Glu, Gly, His, Lys, Met, Pro, Gln, Arg, Trp) showed peaks corresponding to the oligomers of model peptides (n = 1-10) (Figure 5.4). Some peptides with aromatic amino acids (Z = Phe, Tyr) were insoluble in water and only showed the peaks of small oligomers (n = 1-5).

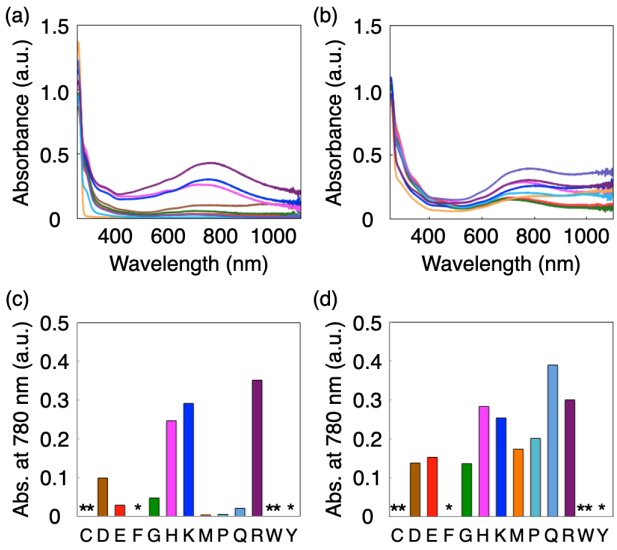


Figure 5.3. Oxidative chemical polymerization of model peptides. (a, b) UV-vis spectra of model peptides (ST**X**ZS, Z = indicated amino acid) chemically polymerized by APS (a) with or (b) without an $FeCl_3$ catalyst. Z = D (sienna), E (red), G (green), H (magenta), K (blue), M (orange), P (deep sky blue), Q (slate blue), R (purple). (c, b) Absorbance at 780 nm of the samples polymerized (c) with or (d) without $FeCl_3$. *: peptides were insoluble in water. **: polymer products were insoluble in water.

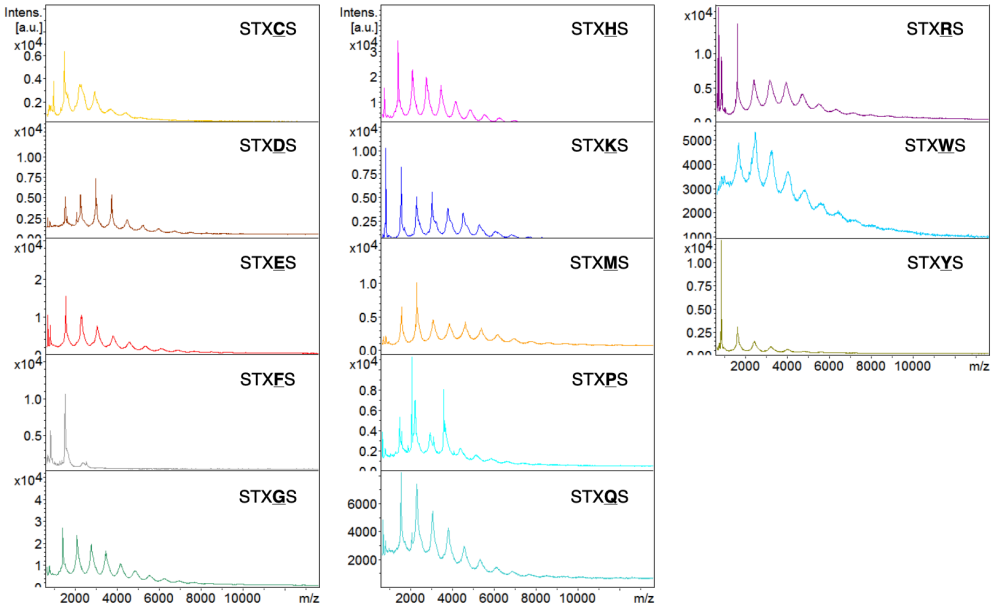


Figure 5.4. MALDI-TOF mass spectra of chemically polymerized peptides. Polymerization was performed by APS without an $FeCl_3$ catalyst. α-cyano-4-hydroxycinnamic acid (CHCA) was used as a matrix. Phe and Tyr: peptides were insoluble in water. Cys and Trp: polymer products were insoluble in water.

5.4. Conclusion

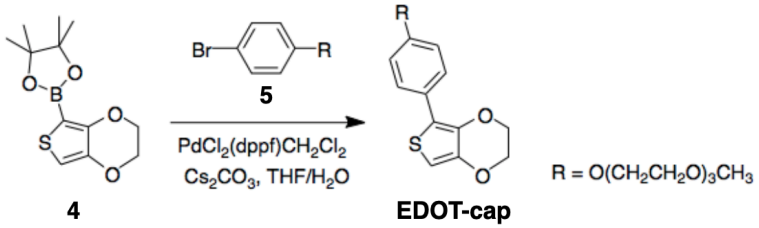
We investigated the oxidative chemical polymerization of the XTEN protein and model peptides. The polymerization of XTEN carrying EDOT-Lys with EDOT-S monomer formed a dark blue precipitate in the solution. A reactivity test using end-capped EDOT revealed that EDOT-Lys in the XTEN protein can react with the EDOT derivative. In addition, study of the effect of the chemical environment on polymerization revealed that peptides with a basic residue adjacent to EDOT-Lys enhance polymerization by APS without an $FeCl_3$ catalyst. The results presented here provide valuable insight into the synthesis of protein–PEDOT conjugates via oxidative chemical polymerization.

5.5. Experimental Procedures

5.5.1 General

All the reagents for organic synthesis were purchased from Sigma Aldrich and Tokyo Chemical Industry Co., Ltd. (TCI), and used as received without further purification. For column chromatography, Biotage Isolera Spektra equipped with a SNAP HP-Sil 10 g cartridge was used. ^{1}H (500 MHz) and ^{13}C NMR (126 MHz) spectra were recorded on a Varian Inova 500 spectrometer. Analytical HPLC analyses for small compounds were performed with an Agilent 1290 Infinity Series HPLC instrument at the Center for Catalysis and Chemical Synthesis in the Caltech Beckman Institute. High-resolution mass spectrometry (HRMS) was performed with an LCT Premier XE Electrospray TOF Mass Spectrometer with electrospray ionization (ESI) at the Caltech CCE Multiuser Mass Spectrometry Laboratory. Matrix-assisted laser desorption/ionization time-of-flight mass (MALDI-TOF MS) spectrometry measurements were carried out on a Bruker Daltonics autoflex™ speed MALDI-TOF/TOF spectrometer at the Caltech CCE Multiuser Mass Spectrometry Laboratory. Absorption spectra were recorded using a Varian Cary 50 UV–visible spectrophotometer.

5.5.2. Synthesis and Characterization of the Compounds



EDOT-cap. 1-Bromo-4-{2-[2-(2-methoxy-ethoxy)-ethoxy]-ethoxy}-benzene (5) was synthesized following the reported procedure.[15] EDOT boronic acid pinacol ester (4) (100 mg, 0.37 mmol), compound 5 (119 mg, 0.37 mmol), and cesium carbonate (365 mg, 1.12 mmol) were dissolved in THF:H_2O (7:1, 8 mL). After addition of $PdCl_2(dppf)CH_2Cl_2$ (31 mg, 0.04 mmol), the solution was degassed for 20 min with Ar and heated to 90 ºC. After 18 h, the reaction mixture was cooled to room temperature, poured into brine (50 mL), and extracted with DCM (3 x 20 mL). The combined organic layers were dried over $MgSO_4$, concentrated *in vacuo*,

and purified by silica gel chromatography (hexane:ethyl acetate 95:5 to 50:50) yielding 64 mg (45%) of yellow oil. ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.61 (d, *J* = 8.7 Hz, 2H), 6.91 (d, *J* = 8.8 Hz, 2H), 6.23 (s, 1H), 4.26 (m, 4H), 4.14 (dd, J = 5.7, 4.1 Hz, 2H), 3.86 (dd, J = 5.6, 4.1 Hz, 2H), 3.78 – 3.51 (m, 8H), 3.38 (s, 3H).; ^{13}C NMR (126 MHz, $CDCl_3$) δ 157.74, 142.33, 137.31, 127.45, 126.25, 117.50, 114.94, 114.94, 96.57, 72.09, 70.99, 70.80, 70.71, 69.88, 67.64, 64.84, 64.62, 59.15. HRMS (ESI) calculated for $C_{19}H_{24}O_6S$ $[M+H]^+$: *m/z* = 381.1372; found: 381.2972.

5.5.3. Cloning, Protein Expression, and Purification

Cloning. The plasmids were constructed by standard recombinant DNA technology using *E. coli* DH10B strain. DNA oligomers for PCRs were purchased from IDT (Coralville, IA). Bacteria were grown on LB/agar plates and LB liquid media with the following antibiotic concentrations: 100 μg/mL spectinomycin and 100 μg/mL ampicillin. The plasmid encoding XTEN was a kind gift from Dr. Peter Rapp. Plasmids and their corresponding coding sequences are presented in *5.7. DNA and Protein Sequences*.

pQE80-XTEN Plasmids. The pQE80 vector (Novagen) confers resistance to ampicillin and carries an *N*-terminal 6xHis-Tag and cloning sites under the T5 promoter. The XTEN gene was cloned to a cloning site by Gibson Assembly. An amber stop codon was introduced at position 15 by site-directed mutagenesis.

Protein Expression and Purification. To express the XTEN carrying EDOT-Lys, *E. coli* strain BL21(DE3) was co-transformed with pUltra-*Mm*PylRS(Y306A/Y384F) and a pQE plasmid encoding XTEN. *E. coli* cultures were grown in 2xYT media at 37 °C to an optical density at 600 nm (OD600) of 0.5-0.7. Then 1 mM EDOT-Lys was added, and protein expression was induced by addition of 1 mM IPTG. After 15 h, the cells were harvested by centrifugation, resuspended in a lysis buffer (0.1 M Na_2HPO_4, 10 mM imidazole; pH 8.0), and lysed by sonication. Lysates were cleared by centrifugation and incubated with Ni-NTA agarose. The resin was washed with a lysis buffer and wash buffer (0.1 M

Na_2HPO_4, 25 mM imidazole; pH 8.0). Protein was eluted with elution buffer (0.1 M Na_2HPO_4, 250 mM imidazole; pH 8.0), dialyzed against water, and lyophilized for storage.

To express the control XTEN without EDOT-Lys, *E. coli* strain BL21(DE3) was transformed with a pQE plasmid encoding XTEN. *E. coli* cultures were grown in 2xYT media at 37 °C to an optical density at 600 nm (OD600) of 0.5-0.7. Then protein expression was induced by addition of 1 mM IPTG. After 6 h, the cells were harvested by centrifugation, resuspended in the lysis buffer (0.1 M Na_2HPO_4, 10 mM imidazole; pH 8.0), and lysed by sonication. Lysates were cleared by centrifugation and incubated with Ni-NTA agarose. The resin was washed with a lysis buffer and wash buffer (0.1 M Na_2HPO_4, 25 mM imidazole; pH 8.0). Protein was eluted with elution buffer (0.1 M Na_2HPO_4, 250 mM imidazole; pH 8.0), dialyzed against water, and lyophilized for storage.

5.5.4. Chemical Polymerization and Reactivity Test

For oxidative chemical polymerization of proteins, EDOT-S (25 mM) and XTEN protein (1 mM) were dissolved in 100 mM HCl. To initiate polymerization, APS (40 mM) and $FeCl_3$ (2 mM) were added, and the reaction mixtures were stirred at room temperature for 18 hours. The reaction mixtures were diluted 80-fold with water, and absorption spectra were recorded using a UV-vis spectrophotometer.

For reactivity test using EDOT-cap, XTEN protein (1 mM) and EDOT-cap (25 mM) were dissolved in 100 mM HCl. To initiate oxidation, APS (40 mM) and $FeCl_3$ (2 mM) were added, and the reaction mixtures were stirred at room temperature for 18 hours. Then excess amount of ethanol was added to the solution to precipitate protein. After removal of the supernatant, protein was redissolved in a small volume of water and washed with ethanol three more times. The resultant protein product was redissolved in water and subjected to MALDI-TOF mass spectrometry.

For oxidative chemical polymerization of model peptides, model peptides (50 mM) were dissolved in 100 mM HCl. To initiate polymerization, APS (40 mM) or APS (40 mM) and $FeCl_3$ (2 mM) were added, and the reaction mixtures were stirred at room temperature for 20 hours. The reaction mixtures were diluted 80-fold with water, and absorption spectra were recorded using a UV-vis spectrophotometer. For mass spectrometry, the reaction mixtures were dialyzed against water overnight and analyzed by MALDI-TOF mass spectrometry.

5.6. NMR Spectra

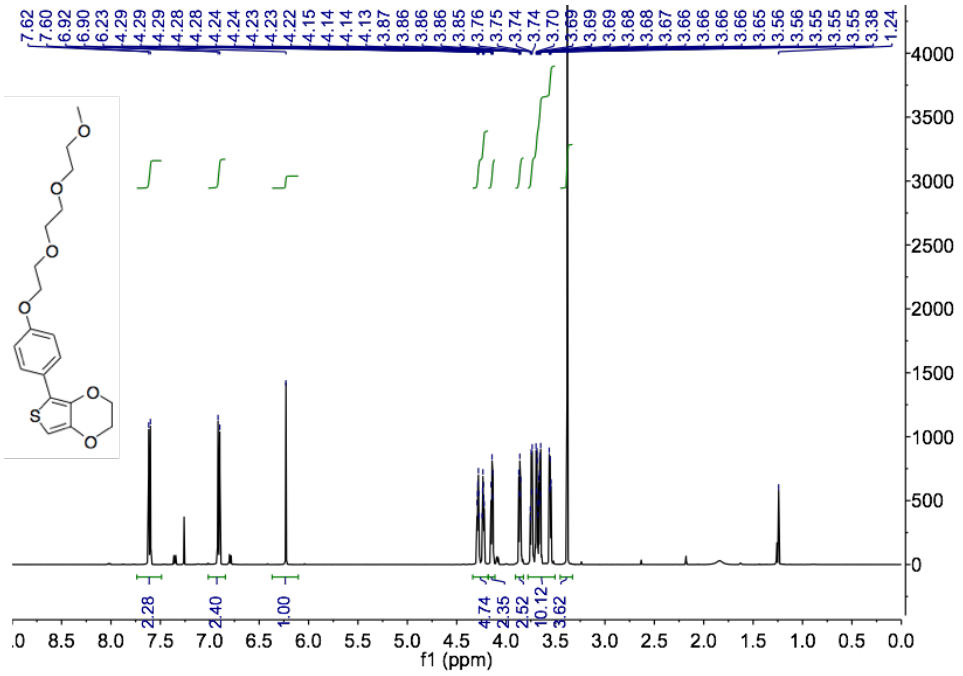


Figure S5.1. ^1H NMR spectrum (500 MHz) of EDOT-cap in $CDCl_3$ at 22 ºC.

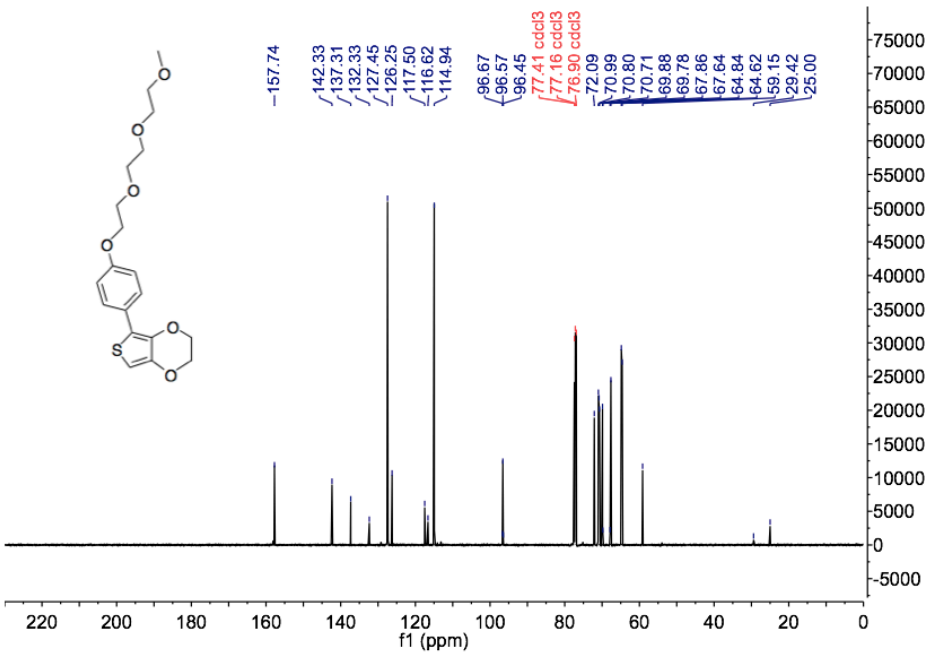


Figure S5.2. ^{13}C NMR spectrum (126 MHz) of EDOT-cap in $CDCl_3$ at 22 ºC.

5.7. DNA and Protein Sequences

Q-XTEN-noam:

atgagaggatcg**catcaccatcaccatcac**ggatccgtcgac**gcg**ggttccaagtccccggctggca
gcccgaccagcactgaagagggcacgagcgagtcggcgaccccggagtctggtccgggcacctccac
cgaaccgtctgagggcagcgcaccgggtagcccggccggtagccctaccagcaccgaagagggtacc
agcacggaaccgagcgaaggctcggcaccgggtacgagcaccgagtaa

MRGS**HHHHHH**GSVD**A**GSKSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGT
STEPSEGSAPGTSTE

*The position 15 is highlighted in red. *N*-terminal 6xHis-Tag is highlighted in blue.

5.8. References

(1) Groenendaal, L.; Jonas, F.; Freitag, D.; Pielartzik, H.; Reynolds, J. R. Poly(3,4-Ethylenedioxythiophene) and Its Derivatives: Past, Present, and Future. *Adv. Mater.* **2000**, *12* (7), 481–494.

(2) Mantione, D.; del Agua, I.; Sanchez-Sanchez, A.; Mecerreyes, D. Poly(3,4-Ethylenedioxythiophene) (PEDOT) Derivatives: Innovative Conductive Polymers for Bioelectronics. *Polymers (Basel).* **2017**, *9* (8).

(3) Ner, Y.; Invernale, M. A.; Grote, J. G.; Stuart, J. A.; Sotzing, G. A. Facile Chemical Synthesis of DNA-Doped PEDOT. *Synth. Met.* **2010**, *160* (5–6), 351–353.

(4) Horikawa, M.; Fujiki, T.; Shirosaki, T.; Ryu, N.; Sakurai, H.; Nagaoka, S.; Ihara, H. The Development of a Highly Conductive PEDOT System by Doping with Partially Crystalline Sulfated Cellulose and Its Electric Conductivity. *J. Mater. Chem. C* **2015**, *3* (34), 8881–8887.

(5) PuiggalÍ-Jou, A.; Del Valle, L. J.; Alemán, C. Encapsulation and Storage of Therapeutic Fibrin-Homing Peptides Using Conducting Polymer Nanoparticles for Programmed Release by Electrical Stimulation. *ACS Biomater. Sci. Eng.* **2020**, *6* (4), 2135–2145.

(6) Kirchmeyer, S.; Reuter, K. Scientific Importance, Properties and Growing Applications of Poly(3,4-Ethylenedioxythiophene). *J. Mater. Chem.* **2005**, *15* (21), 2077–2088.

(7) Gueye, M. N.; Carella, A.; Faure-Vincent, J.; Demadrille, R.; Simonato, J. P. Progress in Understanding Structure and Transport Properties of PEDOT-Based Materials: A Critical Review. *Prog. Mater. Sci.* **2020**, *108* (October 2019), 100616.

(8) Hwang, J.; Tanner, D. B.; Schwendeman, I.; Reynolds, J. R. Optical Properties of Nondegenerate Ground-State Polymers: Three Dioxythiophene-Based Conjugated Polymers. *Phys. Rev. B - Condens. Matter Mater. Phys.* **2003**, *67* (11), 10.

(9) Franco-Gonzalez, J. F.; Zozoulenko, I. V. Molecular Dynamics Study of

Morphology of Doped PEDOT: From Solution to Dry Phase. *J. Phys. Chem. B* **2017**, *121* (16), 4299–4307.

(10) Lapitan, L. D. S.; Tongol, B. J. V.; Yau, S. L. In Situ Scanning Tunneling Microscopy Imaging of Electropolymerized Poly(3,4-Ethylenedioxythiophene) on an Iodine-Modified Au(1 1 1) Single Crystal Electrode. *Electrochim. Acta* **2012**, *62*, 433–440.

(11) Zhang, J.; Ellis, H.; Yang, L.; Johansson, E. M. J.; Boschloo, G.; Vlachopoulos, N.; Hagfeldt, A.; Bergquist, J.; Shevchenko, D. Matrix-Assisted Laser Desorption/Ionization Mass Spectrometric Analysis of Poly(3,4-Ethylenedioxythiophene) in Solid-State Dye-Sensitized Solar Cells: Comparison of In Situ Photoelectrochemical Polymerization in Aqueous Micellar and Organic Media. *Anal. Chem.* **2015**, *87* (7), 3942–3948.

(12) Zotti, G.; Zecchin, S.; Schiavon, G.; Groenendaal, L. B. Electrochemical and Chemical Synthesis and Characterization of Sulfonated Poly (3,4-Ethylenedioxythiophene): A Novel Water-Soluble and Highly Conductive Conjugated Oligomer. *Macromol. Chem. Phys.* **2002**, *203* (13), 1958–1964.

(13) Yano, H.; Kudo, K.; Marumo, K.; Okuzaki, H. Fully Soluble Self-Doped Poly(3,4-Ethylenedioxythiophene) with an Electrical Conductivity Greater than 1000 S Cm -1. *Sci. Adv.* **2019**, *5* (4), 1–10.

(14) Persson, K. M.; Karlsson, R.; Svennersten, K.; Löffler, S.; Jager, E. W. H.; Richter-Dahlfors, A.; Konradsson, P.; Berggren, M. Electronic Control of Cell Detachment Using a Self-Doped Conducting Polymer. *Adv. Mater.* **2011**, *23* (38), 4403–4408.

(15) Li, W. S.; Yamamoto, Y.; Fukushima, T.; Saeki, A.; Seki, S.; Tagawa, S.; Masunaga, H.; Sasaki, S.; Takata, M.; Aida, T. Amphiphilic Molecular

Design as a Rational Strategy for Tailoring Bicontinuous Electron Donor and Acceptor Arrays: Photoconductive Liquid Crystalline Oligothiophene-C60 Dyads. *J. Am. Chem. Soc.* **2008**, *130* (28), 8886–8887.