

**PROTEIN ENGINEERING TO ENABLE SITE-SPECIFIC  
BIOCONJUGATION THROUGH CATECHOL-BORONATE COUPLING  
CHEMISTRY**

by

Soumili Chattopadhyay

A thesis submitted to the Faculty of the University of Delaware in partial fulfillment of the requirements for the degree of Master of Chemical Engineering

Spring 2021

© 2021 Soumili Chattopadhyay  
All Rights Reserved

**PROTEIN ENGINEERING TO ENABLE SITE-SPECIFIC  
BIOCONJUGATION THROUGH CATECHOL-BORONATE COUPLING  
CHEMISTRY**

by

Soumili Chattopadhyay

Approved: \_\_\_\_\_  
Aditya M. Kunjapur, Ph.D.  
Professor in charge of thesis on behalf of the Advisory Committee

Approved: \_\_\_\_\_  
Eric Furst, Ph.D.  
Chair of the Department of Chemical and Biomolecular Engineering

Approved: \_\_\_\_\_  
Levi T. Thompson, Ph.D.  
Dean of the College of Engineering

Approved: \_\_\_\_\_  
Louis F. Rossi, Ph.D.  
Vice Provost for Graduate and Professional Education and  
Dean of the Graduate College

## ACKNOWLEDGMENTS

With great pleasure I would like to thank my advisor, Prof. Aditya M. Kunjapur for his constant supervision, encouragement and constructive criticism. I must also express my gratitude towards him, for allowing me the requisite level of freedom to conduct my research without which perhaps, the work may not have been that enjoyable. It was a nice experience to be a member of Dr. Kunjapur's research lab. Execution of this project wouldn't have been possible without his valuable suggestion at every step.

At the same time, I would like to express my gratitude towards the graduate students Neil Butler, Michaela Jones, Sunny Sen and Morgan Sulzbach who had provided me with constant encouragement and support throughout my stay at Kunjapur Lab. I sincerely thank all the undergraduate students for helping me in the time of need. I would specially thank Ishika Govil for all her valuable contribution in the docking experiments of the project.

Last but not the least, the main persons responsible for my success in life are my parents, Mili and Sourav Chatterjee and my fiancé. They have been my constant source of encouragement, inspiration and motivation. They have shown immense confidence and faith in me which made me to persevere everyday to reach my destination. I want to thank them for their love, support, suggestions and advice. Finally, I would like to dedicate this thesis to the love of my life, my fiancé, Snehashis Choudhury.

## TABLE OF CONTENTS

|                      |      |
|----------------------|------|
| LIST OF TABLES.....  | vii  |
| LIST OF FIGURES..... | viii |
| ABSTRACT.....        | xi   |

### Chapter

|       |   |    |
|-------|---|----|
| 1     | INTRODUCTION.....   | 1  |
| 1.1   | Motivation.....   | 1  |
| 1.2   | Non-standard amino acid (nsAA).....   | 3  |
| 1.3   | Genetic code Expansion.....   | 5  |
| 1.4   | Bio-orthogonal conjugation using nsAA.....  | 8  |
| 1.4.1 | Reactions involving aldehydes and ketones.....  | 9  |
| 1.4.2 | Reactions involving azides.....   | 10 |
| 1.4.3 | Diels – Alder reaction scheme.....  | 10 |
| 1.5   | Aim and outline of the thesis.....  | 13 |
| 2     | STUDYING MISINCORPORATION OF L-3,4-<br>DIHYDROXYPHENYLALANINE (DOPA) IN PLACE OF<br>TYROSINE..... | 14 |
| 2.1   | Background.....   | 14 |
| 2.2   | Materials and Methods.....  | 15 |
| 2.2.1 | Strain and Chemicals.....   | 15 |
| 2.2.2 | Cell culture conditions.....  | 15 |
| 2.2.3 | Cell lysis.....   | 15 |
| 2.2.4 | NBT assay.....  | 16 |
| 2.3   | Results and discussion.....   | 18 |
| 2.4   | Summary.....  | 21 |

|       |  |    |
|-------|--|----|
| 3     | INVESTIGATING SITE-SPECIFIC INCORPORATION OF CATECHOL AMINO ACID WITH EXTENDED SIDECHAIN IN PROTEINS VIA ORTHOGONAL TRANSLATION MACHINERY..... | 22 |
| 3.1   | Background.....  | 22 |
| 3.2   | Materials and Methods.....   | 24 |
| 3.2.1 | Strains, plasmids and chemicals.....   | 24 |
| 3.2.2 | Transformation of C321. ΔA E. coli strain.....   | 25 |
| 3.2.3 | Fluorescence-based nsAA incorporation assay.....   | 25 |
| 3.2.4 | Protein over-expression.....   | 26 |
| 3.2.5 | Protein purification.....  | 27 |
| 3.2.6 | Analysis of protein.....   | 28 |
| 3.2.7 | NBT assay on cell lysate.....  | 28 |
| 3.3   | Results and Discussion.....  | 29 |
| 3.3.1 | Investigating misincorporation of C5-DOPA in E. coli proteome .....  | 29 |
| 3.3.2 | Optimization of the fluorescence-based assay for C5-DOPA incorporation.....  | 29 |
| 3.3.3 | Screening of an orthogonal aaRS-tRNA library for C5-DOPA incorporation in a reporter protein.....  | 34 |
| 3.3.4 | Confirmation of incorporation of C5-DOPA in single-site UAG sfGFP .....  | 37 |
| 3.4   | Summary.....   | 43 |
| 4     | INVESTIGATING CATECHOL-BORONATE COUPLING CHEMISTRY IN CONTEXT OF EXTENDED SIDE-CHAIN CATECHOLS.....  | 44 |
| 4.1   | Background.....  | 44 |
| 4.2   | Materials and Methods.....   | 49 |
| 4.2.1 | Chemicals.....   | 49 |
| 4.2.2 | ARS assay – small molecule conjugation.....  | 49 |
| 4.2.3 | 3-(Dansylamino)phenylboronic acid (Dns-PBA) and catechol   |    |

|       |  |    |
|-------|--|----|
|       | small- molecule conjugation.....   | 50 |
| 4.2.4 | Dansyl-phenyl boronic acid and C5-DOPA protein<br>conjugation.....                                   | 50 |
| 4.3   | Results and Discussion.....  | 51 |
| 4.3.1 | Optimizing ARS assay for catechols conjugation.....  | 51 |
| 4.3.2 | Expanding the library of boronic acid derivative.....  | 54 |
| 4.3.3 | Optimization of catechol-boronate chemistry in-vitro.....  | 60 |
| 4.4   | Summary.....   | 64 |
| 5     | CONCLUSIONS AND DIRECTIONS FOR FUTURE WORK .....   | 66 |
| 5.1   | Conclusions.....   | 66 |
| 5.2   | Directions for future work.....  | 67 |
| 5.2.1 | Optimizing the shake flask culture conditions for C5-DOPA<br>incorporation study.....                | 67 |
| 5.2.2 | Investigating incorporation study for other extended side chain<br>catechols.....                    | 68 |
| 5.2.3 | Generation and screening of synthetase library for selective C5-<br>DOPA aminoacylation.....         | 68 |
| 5.2.4 | Characterization of catechol-boronate chemistry for coupling in<br>fermentation-like conditions..... | 71 |
|       | REFERENCES.....  | 72 |

## LIST OF TABLES

|                  |   |    |
|------------------|---|----|
| <b>Table 1.1</b> | List of bio-orthogonal reactions.....   | 12 |
| <b>Table 3.1</b> | Expected and observed molecular weight (MW) for LC-MS analysis.<br>(NT– Not Tested) ..... | 39 |

## LIST OF FIGURES

- Figure 1.1:** Site-specific incorporation of nsAAs into a protein. Engineered OTS that consists of a stop codon (amber) suppressor orthogonal synthetase/tRNA pair is used to introduce a nsAA into the amber codon of the mRNA template of protein of interest.....7
- Figure 2.1:** Misincorporation of DOPA in place of tyrosine (a) Structure of tyrosine and DOPA. (b) NBT stained dot-blot of cell lysates with 0, 1, and 3 mM DOPA supplemented in media for detection of catechol presence in proteins. (c) Docking simulation of tyrosine and DOPA as substrates with EcTyrRS performed using AutoDock Vina ..... 17
- Figure 2.2:** Determining the sensitivity of NBT-dot blot. The negative control for this experiment is untreated shown on the extreme right... .. 20
- Figure 2.3:** Growth curve for MG1655 grown in presence and absence of DOPA. Calculation of doubling time is based on mid-exponentialphase ..... 20
- Figure 3.1:** a) Structure of C5-DOPA. b) NBT-dot blot of cell lysate with various concentrations of C5-DOPA. c) Docking simulation of Tyrosine (pink) and C5-DOPA (yellow) as substrates in EcTyrRS using Autodock Vina ..... 31
- Figure 3.2:** FL/OD graph for plasmid-based sfGFP reporter. Experiments were performed on separate days and in triplicate..... 32
- Figure 3.3:** Switching from plasmid-based sfGFP reporter to genome integrated sfGFP reporter system. The reporter contains 1 UAG, 2 UAG and 3 UAG codons ..... 32
- Figure 3.4:** Schematic of the nsAA incorporation machinery and “post-translational proofreading” system in a recoded E. coli C321.ΔA strain with genome integrated sfGFP bearing N-terminal 1 UAG codon for site-specific incorporation of nsAA..... 33

|  |    |
|--|----|
| <b>Figure 3.5:</b> FL/OD graph for genome integrated sfGFP reporter. Reproducibility of the incorporation is shown using DopaRS (A) and pAcFRS (B) synthetases for DOPA and C5-DOPA. Experiments were performed on separate days and in triplicate.....  | 33 |
| <b>Figure 3.6:</b> Preliminary results of synthetase screen for 12 different aaRS against C5-DOPA. A small subset of the screened synthetases for sfGFP with 1 stop codon is shown in (a) and (b). The selected synthetases with 2 and 3 stop codons are shown in (c) and (d), respectively.....   | 36 |
| <b>Figure 3.7:</b> LC-MS analysis for Ub-M-GFP. This is one of the controls for the experiment. For this control the strain did not contain PTP plasmid. Hence there was no expression of Ubiquitinase. Observed molecular weight = 37140 Da. Figure 3.3: FL/OD graph for plasmid-based sfGFP reporter. Experiments were performed on separate days and in triplicate..... | 40 |
| <b>Figure 3.8:</b> LC-MS analysis for Ub-S-GFP. This is one of the controls for the experiment. For this control the strain harbored the PTP plasmid. Observed molecular weight = 34296 Da. Figure 3.3: FL/OD graph for plasmid-based sfGFP reporter. Experiments were performed on separate days and in triplicate.....   | 40 |
| <b>Figure 3.9:</b> LC-MS analysis for pAcFRS_1UAG sfGFP. Observed molecular weight = 37302 Da.....   | 41 |
| <b>Figure 3.10:</b> LC-MS analysis for pAcFRS.2t.1_1UAG sfGFP. Observed molecular weight = 34296 Da.....   | 41 |
| <b>Figure 3.11:</b> LC-MS analysis for ScWRS_1UAG sfGFP. Observed molecular weight = 33468 Da.....   | 42 |
| <b>Figure 4.1:</b> Reaction scheme for reversible and irreversible catechol-boronate conjugation.....  | 47 |
| <b>Figure 4.2:</b> Scheme of the ARS assay.....  | 48 |
| <b>Figure 4.3:</b> Change in fluorescence intensity for ARS-PBA assay in presence of competing catechol, DOPA, C5-DOPA.....  | 52 |
| <b>Figure 4.4:</b> Calculation of equilibrium constant based on ARS binding with PBA.....  | 52 |
| <b>Figure 4.5:</b> Calculation of equilibrium constant based on DOPA binding with PBA.....   | 53 |

|   |    |
|---|----|
| <b>Figure 4.6:</b> Calculation of equilibrium constant based on C5-DOPA binding with PBA.....                   | 53 |
| <b>Figure 4.7:</b> Reaction of Dns-PBA with DOPA and C5-DOPA .....  | 55 |
| <b>Figure 4.8:</b> Fluorescence intensity for DOPA and Dns-PBA conjugation reaction .....                       | 56 |
| <b>Figure 4.9:</b> Fluorescence spectra for DOPA and Dns-PBA conjugation reaction..                             | 57 |
| <b>Figure 4.10:</b> Fluorescence intensity for C5-DOPA and Dns-PBA conjugation reaction .....                   | 58 |
| <b>Figure 4.11:</b> Fluorescence spectra for C5-DOPA and Dns-PBA conjugationreaction .....                      | 59 |
| <b>Figure 4.12:</b> Fluorescence intensity change for protein samples upon conjugation with Dns-PBA.....        | 61 |
| <b>Figure 4.13:</b> Fluorescence intensity for varying protein concentration upon conjugation with Dns-PBA..... | 62 |
| <b>Figure 4.14:</b> LC-MS analysis for tyrosinase treated sfGFP in absence of Dns-PBA.....                      | 63 |
| <b>Figure 4.15:</b> LC-MS analysis for tyrosinase treated sfGFP in presence of 0.1 mM Dns-PBA .....             | 63 |
| <b>Figure 5.1:</b> Engineering and selection of candidate synthetase for C5-DOPA incorporation .....            | 70 |

## ABSTRACT

Over the past few decades, the customizability provided to biomolecules through protein conjugation has enabled a diverse set of applications, which include protein-based therapeutics, biomaterials, and biosensors. However, current chemical modification techniques suffer from lack of specificity that results in heterogeneous mixtures, limited time stability in vivo, and perturbations to the protein function. Synthetic biology researchers have overcome some of the challenges by utilizing genetic code expansion technology to enable precise formation of protein conjugates.

In this project, we seek to develop a bio-orthogonal protein conjugation platform in *Escherichia coli* (*E. coli*) based on catechol-boronate coupling chemistry. This chemistry proceeds in aqueous and mild conditions and can be tuned to be reversible or irreversible according to the end application. In order to perform catechol-boronate conjugation in proteins, we have site-specifically inserted catechol amino acids into proteins using non-standard amino acid (nsAA) incorporation technique. Although a catechol amino acid: L-3,4-dihydroxyphenylalanine (DOPA) is a proven candidate for ribosomal incorporation, an unaddressed challenge is the replacement of tyrosine with DOPA due to the structural similarity. We have investigated in this thesis whether a DOPA analogue having extended side chain (C5-DOPA) with lower resemblance to tyrosine can be site-specifically incorporated into proteins. We have screened a library of published

OTSs from different families for specificity and activity on C5-DOPA. Based on our results top performing candidates will be engineered as next step for improved specificity to incorporate C5-DOPA at multiple instances in a reporter protein.

## Chapter 1

### INTRODUCTION

#### 1.1 Motivation

Owing to their improved function and specificity, proteins have become an irreplaceable biological tool for applications that pertain to therapeutics (Elzahhar et al. 2019), in vivo diagnostics (Esteves-Villanueva, Trzeciakiewicz, and Martic2014), and imaging (Zhao et al. 2019). By using bioconjugation, these proteins can be selectively paired to a non-native partner to meet broader end goals. For example, in a biosensing application, a conjugate could serve as a signal transducer, or, for drug targeting purposes, the conjugate could serve as a drug payload (Elzahhar et al.2019). These applications have tremendous economic value as shown by the market size for protein labeling, which was estimated to be \$1.8 billion in 2016 (“Protein Labeling Market Size , Share & Trends Analysis Report by Product (Reagents, Kits, Services), By Application (Immunological Techniques, Fluorescence, Microscopy), By Labeling Method, and Segment Forecasts, 2018-2024” 2018); similarly, the antibody-drug conjugate (ADCs) market size is expected to reach \$9.93 billion by 2025 (Grand View Research 2019).

Current methods for protein chemical modification can be primarily categorized as thiol-Michael addition (forms covalent bonds at all surface accessible cysteines), Staudinger ligation, metal catalyzed click conjugations and various cycloadditions (Hermanson 2013). Despite the success of these modification

strategies, they are inhibited by challenges including poor site-selectivity, sluggish kinetics, and heterogeneous product distribution, leading to inseparable mixtures and lower yield, which limit their end usage for controlled and well-defined conjugation applications (Stephanopoulos and Francis 2011; Chalker et al. 2009; Ryan et al. 2011). To overcome these obstacles and label proteins at predefined sites, synthetic biologists have successfully developed tools such as site-specific non-standard amino acid (nsAA) incorporation (Vanbrunt et al. 2015; Pickens et al. 2018; Jakob, Gust, and Grohmann 2019; Kwon and Yang 2017; Lim et al. 2014), sortase-mediated ligation (Proft 2010; Schmohl and Schwarzer 2014) and SpyCatcher/SpyTag chemistry (Reddington and Howarth 2015; Jia et al. 2017). Furthermore, to ensure that native structure of the proteins are unperturbed during these chemical reactions, it is important to meet the following bio-orthogonality criteria: (i) No interference with native cellular biochemical reactions, (ii) Considerable reactivity under mild and physiological conditions, and (iii) Conjugation reactions employed should not cause toxicity (Sletten and Bertozzi 2011; Borrmann and van Hest 2014). In this project, we will focus on developing bioconjugates based on a site-specific, nsAA incorporation strategy because it is an ideal pathway for introducing new functionality in proteins to enable desired conjugation reactions in complex biological media that meet the bio-orthogonality criteria.

## **1.2 Non-standard Amino Acid (nsAA)**

nsAAs play a significant role in genetic code expansion. Translational incorporation of nsAA provides unique features to proteins. Like standard amino acids, an nsAA consists of an amine group, a carboxyl group, and a specific side chain. nsAA possesses a specific side chain that provides structural diversity and unique functionalities to a protein. Till date, 200 nsAAs have been known to be successfully incorporated into proteins (Dumas et al. 2015). nsAAs with spectroscopic properties have been utilized for determining protein structure and function. These probes provide further information on protein dynamics, localization, and protein-protein interactions. The additional chemistries introduce functionalities for the purposes of biocontainment, photocrosslinking, bioconjugation and fluorescence. Successful site-specific labeling and modification of proteins will open new routes for protein conjugates especially ADCs. Ribosomal incorporation of an nsAA into proteins has relied on one of the two approaches: residue-specific incorporation (Singh-Blom, Hughes, and Ellington 2013; Johnson et al. 2010) and site-specific incorporation (Wang, Xie, and Schultz 2006). Residue-specific methods are employed to induce global modifications in the protein sequence, which modifies the physicochemical properties of a protein considerably. This method utilizes the flexibility of the enzyme which is responsible for activation of the natural amino acid to tRNA, amino-acyl tRNA synthetase (aaRS) for recognition of the unnatural substrate (Singh-Blom, Hughes, and Ellington 2013). For residue-specific incorporation, the nsAA of interest should be an analogue of

one of the twenty natural amino acid that will be replaced in the protein and the availability of the respective natural amino acid should be limited in the media. Limited availability of the natural amino acid during *in-vivo* protein translation in *Escherichia coli* (*E. coli*) is ensured by using an auxotrophic strain and chemically defined media such as M9 minimal media. In the past decade, this approach has been utilized to incorporate over a hundred of nsAAs finding application from evolving proteome wide incorporation to survival of an organism in the presence of a particular nsAA (Bacher and Ellington 2001).

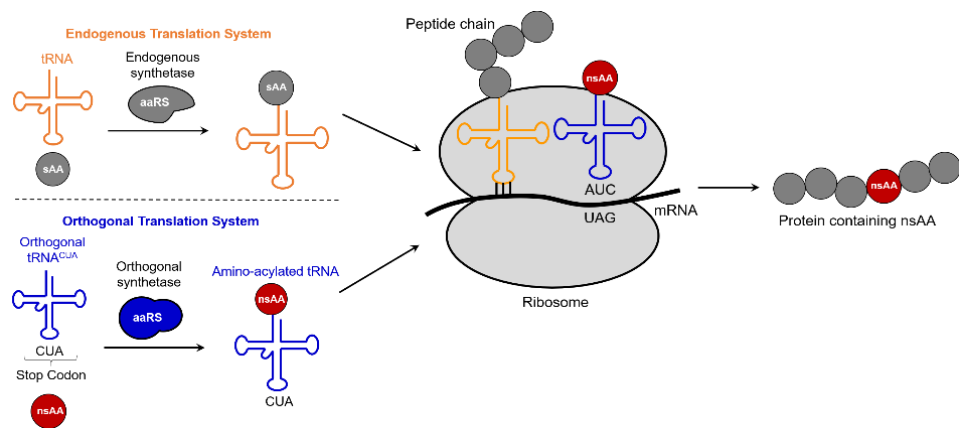
### 1.3 Genetic code Expansion

In 2001, Schultz Lab found a more common way of incorporating an nsAA into a protein of interest in live *E. coli* cells (Wang, Xie, and Schultz 2006). This technique named site-specific incorporation of nsAA, hijacks an orthogonal translation system (OTS) from a phylogenetically distant organism to incorporate an nsAA at a specific codon (Chin 2017). The OTS machinery consists of an aaRS and a complementary tRNA that is orthogonal to the natural amino acids and host translational machinery (**Figure 1.1**). The orthogonal aaRS/tRNA pair selectively charges the nsAA of interest onto the tRNA at reassigned low frequency stop codons called ‘amber’ (UAG) stop codons. A major limitation of this technique is the low activity of the orthogonal aaRS for the non-natural analog. This problem is generally met by altering the specificity of a synthetase enzyme towards a nsAA of interest over its canonical counterpart. For example, over the years tyrosyl-synthetase has been engineered to accommodate a broad scope of tyrosine analogs.

Site-specific incorporation of nsAAs in *E. coli*, has primarily been engineered using: the pyrrolysyl-tRNA synthetase pair (PylRS) from *Methanosarcina* genus (Wan, Tharp, and Liu 2014), the tyrosyl-tRNA synthetase pair (MjTyrRS) from *Methanocaldococcus jannaschii* (Ryu and Schultz 2006) and the tryptophanyl-tRNA synthetase pair (ScWRS) from *Saccharomyces cerevisiae* (Hughes and Ellington 2010). MjTyrRS is extensively used in bacterial hosts for incorporation of tyrosine analogs but PylRS variants covers broader host organisms including both prokaryotes and eukaryotes (Ding et al. 2020; Hohl et al.

2019; Fekner and Chan 2011). To widen the range of chemically diverse nsAAs that could be genetically incorporated into proteins, orthogonal aaRS/tRNA pair from other organism and based on different amino acids have been developed. Lysyl-tRNA synthetase/tRNA<sup>Lys</sup> (Anderson et al. 2004) and glutamyl-tRNA synthetase/tRNA<sup>Glu</sup> (Santoro et al. 2003) have been developed from *Pyrococcus horikoshii*, whereas tryptophanyl-tRNA synthetase/tRNA<sup>Trp</sup> (Chatterjee et al. 2013) have been derived from *Saccharomyces cerevisiae*.

An underlying limitation of this technology lies in the natural phenomenon of recognition of UAG stop codon by release factor 1, which leads to termination of growing peptide chain (Capecchi and Klein 1970). In the presence of an OTS, the UAG stop codon can either be suppressed by an nsAA or terminated by release factor 1. This competition has two major implications (i) low efficiency of nsAA incorporation and, (ii) generation of pool of heterogenous product. Extensive effort has been made to alleviate this challenge by partially or fully recoding the *E. coli* genome, where all the 321 UAG stop codons have been synonymously substituted for UAA stop codons. Furthermore, release factor 1 has been deleted from this strain eliminating competition with nsAA for UAG codon. In this project, we will use the C321.  $\Delta$ A *E. coli* strain for most of the experiments.



**Figure 1.1:** Site-specific incorporation of nsAAs into a protein. Engineered OTS that consists of a stop codon (amber) suppressor orthogonal synthetase/tRNA pair is used to introduce a nsAA into the amber codon of the mRNA template of protein of interest.

#### 1.4 Bio-orthogonal conjugation using nsAA

The last decade was characterized by significant progress in bioconjugation strategies for biomolecules like proteins. Owing to their improved function, specificity and efficacy, proteins have become an irreplaceable member of the therapeutic family rendering the ability to treat multitude of diseases including cancers, autoimmune diseases, and metabolic disorders. Conjugation to a small molecule confers a protein to modify its properties which leads to enhanced circulation time, specificity for drug delivery. With these improved properties, a significant increase in market size of protein-based products such as antibody-drug conjugates, nanobodies, enzymes, cytokines, hormones and inhibitors can be expected in near future. Due to tailorable conjugation chemistries and possibility of carrying out the reaction in mild conditions, protein conjugates are a powerful tool for immunodiagnostic tests. Proteins like Horseradish Peroxidase (HRP) and Alkaline Phosphatase (AP) with chemiluminescent detection properties are an ideal choice for antigen or antibody detection in ELISA, Western blot, immunofluorescence, immunohistochemistry and flow cytometry (Moral and Siahaan 2017).

At present, chemical modification of proteins includes but is not limited to the following bioconjugation reactions (**Table 1.1**): thiol-Michael addition to form covalent bonds at all surface accessible cysteines, Staudinger ligation, carbonyl condensations, metal catalyzed click conjugations and various cycloadditions. These methods typically suffer from poor control over the site of modification in a

protein sequence. For eg. N-hydroxysuccinimidyl (NHS) esters have been known to react with nucleophilic residues on protein surface esp. Lysine residues and yield a covalently modified protein. In most of the proteins, this reaction results in production of heterogeneous products with modification at different sites due to the presence of abundant lysine residues on the surface. Reaction yield, purity, functional characterization, and efficacy become questionable factors in these reaction schemes, when one is seeking for a therapeutic application.

For protein modification at pre-defined sites, scientists have come up with methods including nsAA incorporation, sortase-mediated ligation and SpyCatcher/SpyTag chemistry (Lieser et al. 2020). In this project, we will focus on bio-orthogonal conjugation strategies based on unique functionalities present in protein due to genetic incorporation of nsAA rather than derivatizing natural amino acids. In this way, we aim to produce homogenous products with highly defined conjugation chemistry. Use of versatile bio-orthogonal ligation technique by incorporating nsAA with unique functionalities not available in live cells, allows us to site-selectively modify proteins in complex biological mixtures, with stoichiometric control of conjugated molecules, fast reaction kinetics and reducing the chances of perturbing the protein folding.

#### **1.4.1 Reactions involving aldehydes and ketones**

To skirt around the challenge of common thiol chemistry related to obtaining undesired disulfide bond formation and labile thioester bonds, reactions based on

oxime formation has been approached (Rose 1994). Formation of stable oximes and hydrazone products for *in-vitro* site-specific protein labeling have been achieved upon reaction of hydrazides and alkoxyamines with genetically encoded aldehydes and ketones (Canne et al. 1995; Gaertner et al. 1992; Gong and Pan 2015). However, due to the requirement of harsh acidic conditions ~ pH 4 and slow reaction kinetics, this reaction scheme is unfavorable for intracellular conjugation.

#### **1.4.2 Reactions involving azides**

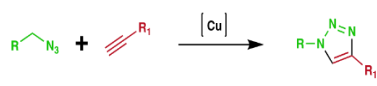
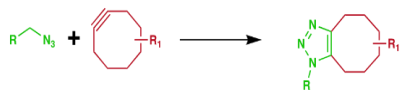
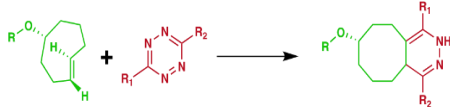
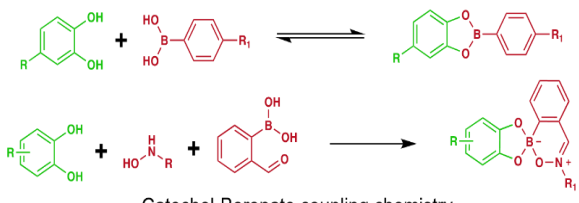
The Bertozzi group in 2000 developed a reaction scheme named Staudinger reactions which involves reaction of an azide containing group with phosphine reagent. The purpose of this chemistry is to form stable amide bonds between protein bearing alkyl azide functional group and phosphine derivatives with ester groups on their aromatic rings (Lin et al. 2005; Nilsson, Kiessling, and Raines 2000). Later scheme of this ligation technique named “traceless Staudinger Ligation” has been developed. The updated scheme involves an acyl group attached to the phosphine derivative, which prevents the formation of toxic by-product phosphine oxide in the final bioconjugate (Saxon, Armstrong, and Bertozzi 2000).

#### **1.4.3 Diels – Alder reaction scheme**

In context of bioconjugation, thiol-maleimide reaction has certain drawbacks such as unstable thiosuccinimide products and requirement of protein manipulation prior to conjugation reaction. The Diels-Alder (DA) bioconjugation platform is based on reaction of cyclic dienes with maleimide groups. The DA

reaction is classified as click chemistry, is traceless and does not require any metal catalyst (St. Amant et al. 2019). This is one of the most commonly utilized bioconjugation strategy for these same reasons.

**Table 1.1:** List of bio-orthogonal reactions

|   | Bio-orthogonal Reaction   | Rate Constant<br>k<br>[M <sup>-1</sup> s <sup>-1</sup> ] | Reversibility<br>of<br>reaction            | <i>in vivo</i> ? |
|---|---|--|--|------------------|
| A |  <p>Cu<sup>I</sup>-catalyzed alkyne-azide cycloadditions</p>               | 10 – 200   | Irreversible                               | No               |
| B |  <p>strain-promoted alkyne-azide cycloadditions</p>                        | 10 <sup>-2</sup> – 1                                     | Irreversible                               | Yes              |
| C |  <p>Diels-Alder cycloadditions between tetrazines and strained alkenes</p> | 1 – 10 <sup>6</sup>                                      | Irreversible                               | Yes              |
| D |  <p>Catechol-Boronate coupling chemistry</p>                              | Not reported   | Tunable<br>(reversible or<br>irreversible) | Yes              |

## 1.5 Aim and outline of the thesis

Our research group works on biosynthesis and genetic incorporation of non-standard amino acids (nsAAs). The focus of this thesis is primarily on site-specific incorporation of a novel catechol nsAA, 2-amino-5- (3,4-dihydroxyphenyl) pentanoic acid (C5-DOPA) which is a structural analogue of L-3,4-dihydroxy phenylalanine (DOPA). This thesis also covers single site and multi-site incorporation of C5-DOPA. A part of the thesis is focused on the application of catechol-containing protein with long term vision like conjugation in living cell or fermentation-like conditions.

Chapter 2 investigates the misincorporation of DOPA in an *E. coli* strain, MG1655 using a redox-staining method, which can detect catechols in proteins.

Chapter 3 illustrates the results obtained by screening of a list of synthetases for C5-DOPA incorporation in sfGFP reporter system with varying UAGs in its sequence.

Chapter 4 demonstrates the small molecule catechol-boronate chemistry and shows the development so far in protein conjugation based on site-specifically incorporated C5-DOPA and a fluorescent boronic acid derivative.

## Chapter 2

### STUDYING MISINCORPORATION OF L-3,4-DIHYDROXYPHENYLALANINE (DOPA) IN PLACE OF TYROSINE

#### 2.1 Background

Catechol functional groups in nsAAs can perform diverse and tunable covalent coupling reactions in cell culture conditions and exhibit bioadhesive feature. The natural catechol-containing amino acid, DOPA, is metabolically accessible (Fordjour et al. 2019; Muñoz et al. 2011) and is a proven candidate for ribosomal incorporation (Alfonta et al. 2003a). However, due to structural similarity between tyrosine (4-hydroxyphenylalanine) and DOPA, EcTyrRS can accept DOPA as its substrate. Successful residue-specific incorporation of DOPA in place of tyrosine has cast doubt on the specificity of EcTyrRS for tyrosine over DOPA. Although EcTyrRS exhibits a 200-fold higher  $K_M$  for DOPA, efficient DOPA incorporation (>90%) has been achieved in non-auxotrophic *E. coli* strain using a cell-free translation system (Ozawa et al. 2005a). Global ribosomal incorporation of DOPA in place of tyrosine results in formation of oxidized proteins, which eventually leads to protein degradation and arrests/retards cell growth (Rodgers et al. 2004; 2002). Global incorporation can be ascribed to the similarity in structure of tyrosine and DOPA (**Figure 2.1(a)**), and hence is a concern for a process seeking specificity and high protein yield with purity. In this sub-aim, we seek to study the misincorporation event for DOPA in MG1655 *E. coli* strain.

## **2.2 Materials and Methods**

### **2.2.1 Strain and Chemicals**

MG1655 *E. coli* strain was used for the DOPA misincorporation study. L-Ascorbic acid was obtained from Alfa Aesar (50-81-7). L-3,4-dihydroxyphenylalanine (59-92-7), Glycine (56-40-6) and Sodium tetraborate decahydrate (1303-96-4) was purchased from Sigma Aldrich. 100 mM of DOPA stock solution was made in 25 mM L-Ascorbic acid solution and 50-100 mM NaOH. LB-Lennox medium (LBL: Tryptone 10g/L, Yeast extract 5g/L, Sodium Chloride 10g/L) was purchased from Fisher Scientific. Nitrotetrazolium blue chloride (298-83-9) was purchased from TCI Chemicals Pvt. Ltd.

### **2.2.2 Cell culture conditions**

For culturing, MG1655 *E. coli* strain was grown in LB-Lennox medium supplemented with 3 mM of DOPA. Negative control of this experiment was culture grown in the absence of DOPA. All the cultures were incubated at 37 °C in a shaking plate incubator with 1000 rpm. After 22 hours, cultures were spun down and washed twice with 1X PBS solution. Cell pellets obtained were stored in -20 °C before lysis.

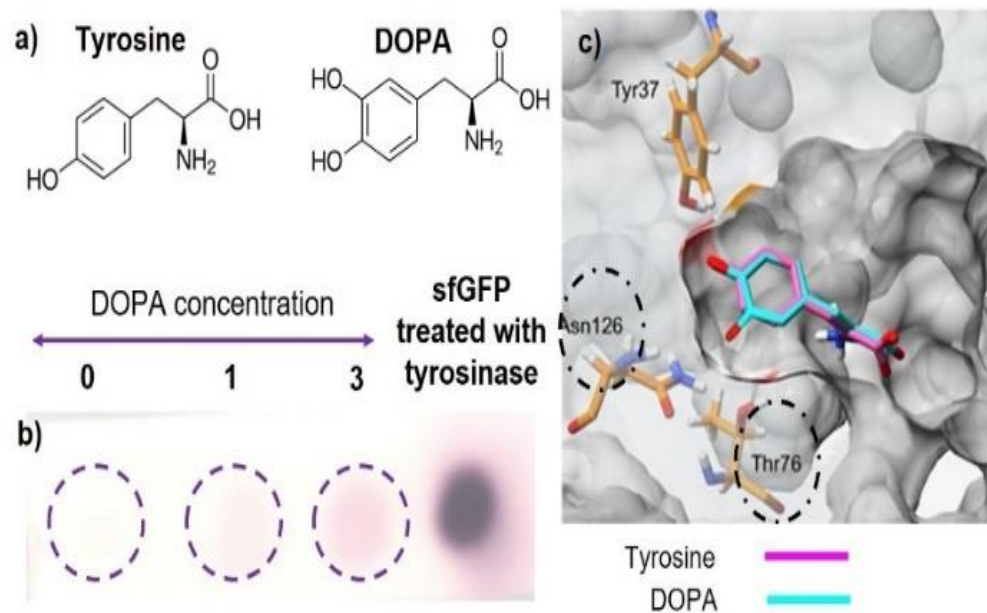
### **2.2.3 Cell lysis**

Next day, pellets were resuspended in lysis buffer (25 mM HEPES, 300 mM NaCl, 10 mM imidazole, 10% glycerol). A resuspension ratio of cell weight (g): liquid (mL) of 1:3 was maintained for all the samples. A common method for sonication was used based on 5 seconds on/10 seconds off for 50 seconds total at 75% amplitude. Total of 5 cycles were performed with 5 mins gap between each

cycle. Throughout the process, samples were kept in ice bucket. After the lysis, cells were centrifuged for 60 mins at 4 °C and supernatant was collected for further protein purification.

#### **2.2.4 NBT assay**

2 uL of 0.5 mg/mL samples were dropped on PVDF membrane 0.45 um. Staining was done with 0.5 mg/mL NBT in 2M Na-glycinate buffer pH 10. After 45 min incubation, the membrane was washed for 10 min with 0.2M Sodium borate pH 8 solution twice followed by water wash for 30 mins.



**Figure 2.1:** Misincorporation of DOPA in place of tyrosine **(a)** Structure of tyrosine and DOPA. **(b)** NBT stained dot-blot of cell lysates with 0, 1, and 3 mM DOPA supplemented in media for detection of catechol presence in proteins. **(c)** Docking simulation of tyrosine and DOPA as substrates with EcTyrRS performed using AutoDock Vina.

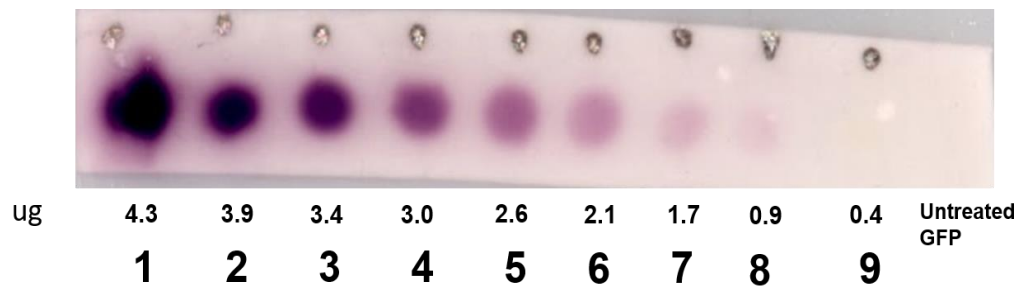
### 2.3 Results and discussion

Herein we have utilized a qualitative assay based on nitroblue tetrazolium dye (NBT) which undergoes color change upon reaction with catechol (MA et al. 1991). Specifically, a standard protein dot-blot protocol with NBT staining was used for rapid detection of catechol containing proteins. In order to have a positive control for this assay, we have purified an in-house sfGFP treated with mushroom tyrosinase for native tyrosine oxidation to catechol. As a negative control, we have used purified sfGFP containing no catechol residues.

In our experiment, we have used this staining procedure to detect incorporation of supplemented DOPA at 1 mM and 3 mM concentrations into proteins of *E. coli* cells (**Figure 2.1(b)**). The positive staining in these samples suggest that there is incorporation of DOPA in the cellular proteome. We have further confirmed this observation by molecular docking using Autodock Vina, where we have performed a docking simulation for tyrosine and DOPA in the active site of EcTyrRS. We observed that both tyrosine and DOPA are docked similarly in the active site of the enzyme (**Figure 2.1(c)**). This result along with NBT dot blot data leads to the hypothesis that DOPA can be accepted as a substrate by the native enzyme.

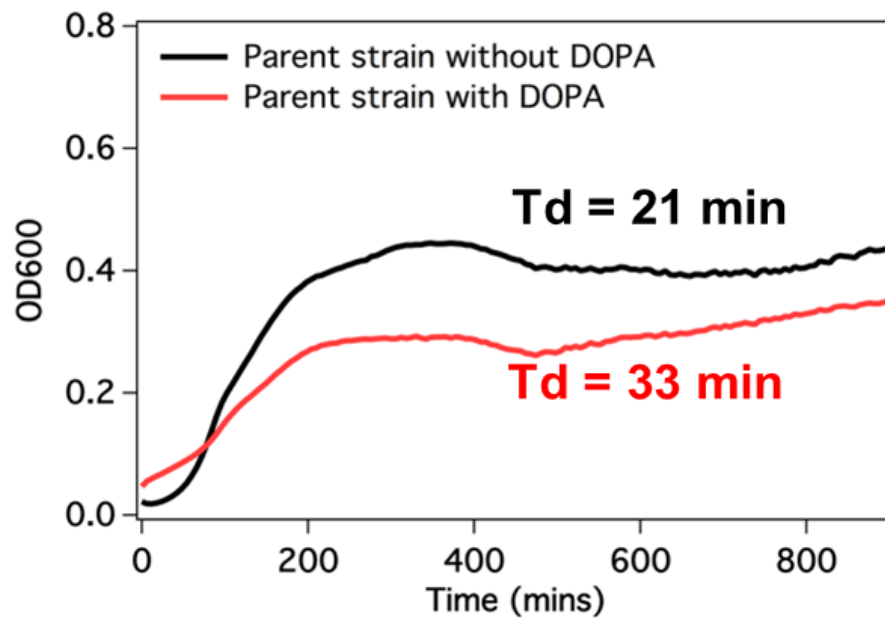
Sensitivity of the NBT staining method was determined by using varying concentration of the tyrosinase treated sfGFP. We varied the concentration from 4.3 ug to 0.4 ug as shown in **Figure 2.2**. Although this is a qualitative technique to detect catechols but it is a quick method to detect catechol presence in any sample.

To study the impact of catechols in media on cell health, we have incubated cells with varying concentrations of DOPA. Previous work in the literature have reported that 3mM concentration of DOPA in the media is toxic to cells, hence affecting their growth. Here, we have analyzed the doubling time of cells from 0-5 mM DOPA and the doubling time result for 3 mM DOPA compared to cells not grown in DOPA is shown in **Figure 2.3**. In presence of DOPA, doubling time of *E. coli* has increased to 33 min as compared to 21 min for untreated cells.



**Figure 2.2:** Determining the sensitivity of NBT-dot blot. The negative control for this experiment is untreated shown on the extreme right.

---



**Figure 2.3:** Growth curve for MG1655 grown in presence and absence of DOPA. Calculation of doubling time is based on mid-exponential phase.

---

## **2.4 Summary**

Due to the structural similarity between the two amino acids, the native enzyme “EcTyrRS” mischarges DOPA onto tRNA of tyrosine and results in substitution of tyrosine residues by DOPA in proteins, which was demonstrated using redox staining NBT dot blot assay. So, the problem in this scenario is when one is looking for a therapeutic application with DOPA at a specific site but obtaining a final product where multiple tyrosine are getting replaced by DOPA. This will have major implications in terms of safety and efficacy of the product.

## Chapter 3.

# INVESTIGATING SITE-SPECIFIC INCORPORATION OF CATECHOL AMINO ACID WITH EXTENDED SIDECHEIN IN PROTEINS VIA ORTHOGONAL TRANSLATION MACHINERY

### 3.1 Background

The incorporation of nsAA during translation process have expanded the chemical repertoire of protein which cannot be attained with natural amino acids. Catechol-based amino acids are versatile chemical compounds exhibiting adhesive (Jeong et al. 2020), redox (Alfonta et al. 2003b), metal chelating (Ayyadurai et al. 2011; B. J. Kim et al. 2015), cross-linking and conjugation properties (Umeda et al. 2009; 2010; Burdine et al. 2004; Xu et al. 2014). Catechol amino acids are formed naturally by post-translational modification of tyrosine but the possibility of genetic encoding of these nsAAs will expand the ability to engineer a protein with unique properties.

L-DOPA is an important catechol amino acid showing multiple biochemical properties. L-DOPA is a precursor of dopamine which can pass through the blood-brain barrier and is usually used for the treatment of Parkinson's disease (Pinder 1970). This amino acid is also found in mussel adhesion proteins (MAPs), and hence responsible for adhesive properties of mussels in wet conditions (Waite and Tanzer 1981; H. Lee, Scherer, and Messersmith 2006; Nicklisch and Waite 2012). Conventional process of obtaining DOPA in a protein is either by encoding tyrosine

at desired location in the sequence followed by treatment with tyrosinase enzyme that converts tyrosine to DOPA (Silverman and Roberto 2007; B. P. Lee et al. 2011) or by residue-specific incorporation of DOPA (Ozawa et al. 2005b). But this process results in a heterogenous product mixture. To obtain site-specific incorporation of DOPA, engineered orthogonal aaRS/tRNA pair has been used. The orthogonal system developed so far for DOPA consists of altered *Methanococcus jannaschii* aminoacyl-tRNA synthetase that charges a mutant tyrosine amber suppressor tRNA. Five residues of the active site (Tyr 32, Glu 107, Asp 158, Ile 159, Leu 162) in the synthetase that are within 6.5 Å of the aryl ring of tyrosine were randomly mutated to alter the specificity for DOPA (Alfonta et al. 2003a). The incorporated DOPA was used to enable electron transfer in proteins.

Due to similar structure of DOPA and tyrosine, in Chapter 2 we have demonstrated the misincorporation problem of DOPA into the *E. coli* proteome. To address this challenge, in this aim, we will investigate whether a DOPA-analogue (C5-DOPA) can be incorporated into proteins. Owing to its extended side-chain, we hypothesize that C5-DOPA will be less likely to be accepted as a substrate by EcTyrRS. Site-directed incorporation of C5-DOPA has never been probed before, however, we posit that an engineered OTS derived from MjTyrRS (Alfonta et al. 2003a) with activity on DOPA could be a starting point. Additionally, to introduce multiple functional handles in a protein that will enable bioconjugate formation, we will also pursue incorporation of C5-DOPA at multiple residues using multi-UAG reporter systems.

## 3.2 Materials and Methods

### 3.2.1 Strains, plasmids and chemicals

MG1655 and C321.  $\Delta A$  *E. coli* strains were used for the C5-DOPA incorporation study. For studies with MG1655 *E. coli* strain, sfGFP reporter was in pZE plasmid under TET-promoter and kanamycin resistance. For studies with C321.  $\Delta A$  *E. coli* strain, Ub-UAG-sfGFP reporter construct was genome integrated under TET-promoter. Both the *E. coli* strains harbors well characterized pEVOL plasmid containing orthogonal aaRS/tRNA<sup>CUA</sup> pair. aaRS expression is controlled by arabinose inducible P<sub>BAD</sub> promoter system whereas tRNA expression is driven by a constitutive promoter. In addition to pEVOL plasmids, C321.  $\Delta A$  strain was also transformed with “posttranslational proofreading- UBPI/ClpS-V65I” (PTP) system in a pZE plasmid under TET-promoter system. Plasmid extraction and purification was performed using QIAprep® Spin Miniprep Kit (27104). Measurement of plasmid DNA concentration was performed using Nanodrop instrument at 260 nm. It was ensured that 260/280 absorbance ratio was ~1.8 for plasmid purity.

2-amino-5-(3,4-dihydroxyphenyl)pentanoic acid (2352060-41-2) was purchased from Chemspace Delivering Discovery Solutions. 100 mM of C5-DOPA stock solution was made in 25 mM L-Ascorbic acid solution and 50-100 mM NaOH. Ammonium acetate (631-61-8) was purchased from Fisher Chemicals. NaCl (7647-14-5), Glycerol (56-81-5) and Imidazole (288-32-4) were purchased from Sigma Aldrich. 2-[4-(2-hydroxyethyl)piperazin-1-yl]ethanesulfonic acid (HEPES) (7365-45-9) was purchased from TCI.

### **3.2.2 Transformation of C321. $\Delta A$ *E. coli* strain**

Prior to electroporation, cells were made electrocompetent. In-house standard protocol for making electrocompetent cells was used. 2 mL of overnight culture was grown from fresh glycerol stock. The following day, fresh inoculation was initiated with the saturated overnight culture and was incubated at 34°C with 250 rpm to reach mid-exponential phase of  $OD_{600} = 0.4-0.6$ . Cells were transferred to Eppendorf tubes and spun down at 7,000 rcf for 2 min to obtain the pellet. Supernatant was discarded and the cells were resuspended in 1 mL of chilled 10% glycerol. The process was repeated thrice with increasing speed: 8000, 9000, and 10000 rcf. After the last wash round, electrocompetent cells were finally resuspended in 100  $\mu$ L of chilled 10% glycerol.

For transformation of plasmids in electrocompetent cells, electroporation technique was used. 20  $\mu$ L of electrocompetent cells were mixed with 20  $\mu$ L of chilled 10% glycerol. For transforming two plasmids, 1.5  $\mu$ L of each plasmid were used. For electroporation, 1 mm gap cuvettes were used with 1.8 kV, 200 ohms and 25  $\mu$ F settings. Recovery was done by adding 500  $\mu$ L of LB and incubating the cells for 1 hour at 34°C followed by plating the cells with kanamycin and chloramphenicol resistance and incubating overnight.

### **3.2.3 Fluorescence-based nsAA incorporation assay**

For culture studies with C321.  $\Delta A$  *E. coli* strain, cells were grown to confluence overnight in LB media with required antibiotics. Next day, experimental cultures were seeded at 1:50 dilution in 2xYT media supplemented with chloramphenicol,

kanamycin, arabinose and C5-DOPA in sterile 96-deep well culture plates. Chloramphenicol and kanamycin antibiotics were used at 0.5x concentration to maintain pEVOL and pZE plasmids respectively. Cultures were incubated at 34 °C in a shaking plate incubator with 1000 rpm. sfGFP and UBPI/ClpS expression was induced after 5 hours by addition of anhydrotetracycline, and cultures were further incubated at the same conditions for an additional 18-22 hours prior to measurement. Cells were centrifuged and pellets were washed twice with 1x PBS and finally resuspended in PBS. sfGFP fluorescence was measured on a Spectramax spectrophotometric plate reader with excitation and emission wavelengths of 488 and 525 nm. Fluorescence data was normalized by the OD<sub>600</sub> reading to obtain FL/OD.

#### **3.2.4 Protein over-expression**

C321.  $\Delta A$  *E. coli* strain was used as host for the over-expression of sfGFP-1 UAG reporter protein. Prior to over-expression in shake flask, cells were grown to confluence overnight in LB media with 0.5x of each kanamycin and chloramphenicol for maintenance of plasmids. The following day, experimental cultures were seeded at 1:50 ratio in 100 mL of 2xYT media supplemented with chloramphenicol, kanamycin, arabinose and 1.5 mM C5-DOPA. Chloramphenicol and kanamycin antibiotics were used at 0.5x concentration to maintain pEVOL and pZE plasmids respectively. Cultures were grown at 34 °C in an incubator shaker and after 5 hours, sfGFP and UBPI/ClpS expression was induced with anhydrotetracycline addition. Cultures were further incubated at the same conditions

for an additional 18-22 hours. Post cell culture, cells were spun down and washed twice with 1x PBS solution. Cell pellets obtained were stored in -20 °C before lysis.

### **3.2.5 Protein purification**

For cell lysis, frozen pellets were resuspended in lysis buffer (25 mM HEPES, 10 mM Imidazole, 300 mM NaCl and 10% glycerol). A resuspension ratio of cell weight (g): liquid (mL) of 1:3 was maintained for all the samples. A common method for sonication was used based on 5 seconds on/10 seconds off for 50 seconds total at 75% amplitude. Total of 5 cycles were performed with 5 mins gap between each cycle. Throughout the process, samples were kept in ice bucket. After the lysis, cells were centrifuged for 60 mins at 4 °C and supernatant was collected for further protein purification.

Prior to Ni-NTA based chromatographic purification, concentration of the cell lysate was measured using standard Bradford Assay protocol. His-tagged sfGFP was purified using immobilized metal affinity chromatography (IMAC) using pre-charged Ni Sepharose™ High Performance, 5 mL column in ÄKTA Avant. Maximum loading on the column was maintained to be ~200 mg to avoid overloading. For binding of the his-tagged protein, equilibration buffer of composition similar to lysis buffer was used, whereas for elution, lysis buffer composition but with 300 mM Imidazole was used. Purified sfGFP protein was checked for fluorescence activity under blue light.

### **3.2.6 Analysis of protein**

Concentration of the purified protein was measured using Nanodrop instrument, where extinction coefficient of  $83,300 \text{ M}^{-1} \text{ Cm}^{-1}$  and molecular weight of 26,800 Da was used. To further verify the over-expression of sfGFP, SDS-PAGE with Coomassie Staining was performed.

ESI-Q-Tof LC-MS analysis for confirmation of C5-DOPA incorporation was performed at the University of Delaware, Center Core Facility. Samples were buffer exchanged into 10 mM Ammonium acetate buffer, which is a volatile buffer suitable for LC-MS analysis. 200 uL of samples at 10 uM was required for the LC-MS analysis.

### **3.2.7 NBT assay on cell lysate**

2 uL of 0.5 mg/mL samples were dropped on 0.45  $\mu\text{m}$  PVDF membrane. Staining was done with 0.5 mg/mL NBT in 2M Na-glycinate buffer pH 10. After 45 min incubation, the membrane was washed for 10 min with 0.2M Sodium borate pH 8 solution twice followed by water wash for 30 mins. Samples containing catechol will be stained purple, intensity of the band will depend on the number of catechol residues in the sample.

### 3.3 Results and Discussion

#### 3.3.1 Investigating misincorporation of C5-DOPA in *E. coli* proteome

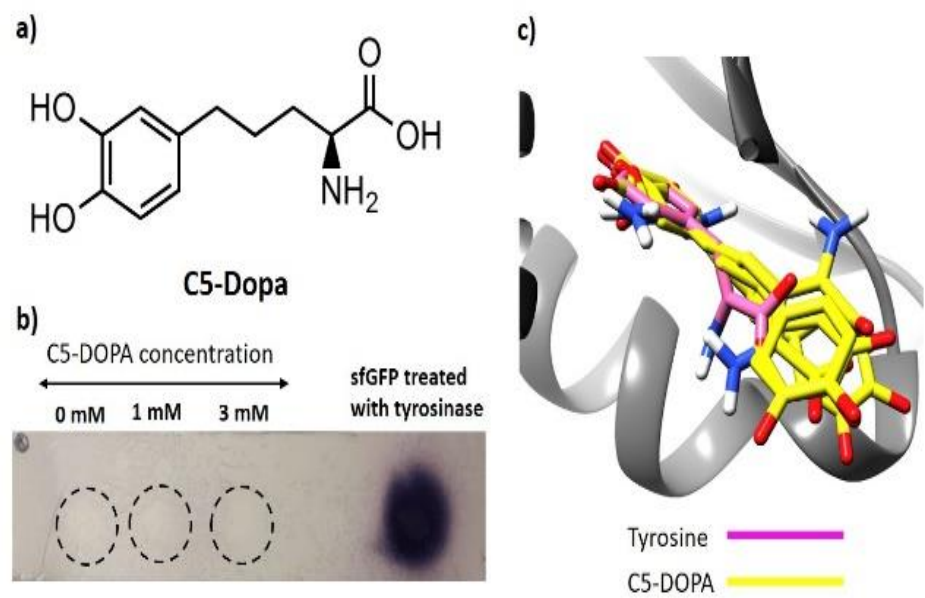
Due to the larger sidechain of C5-DOPA relative to Tyr and DOPA, we hypothesize that EcTyrRS will not accept C5-DOPA as its substrate (**Figure 3.1(a)**). To better understand this, we have performed molecular docking of C5-DOPA into the active-site of EcTyrRS. The docking data showed that C5-DOPA is oriented in opposite direction to tyrosine when forced to find the lowest free energy conformation within a confined geometry, which suggests that C5-DOPA is not considered as a substrate by the native enzyme as shown in **Figure 3.1(c)**. To confirm our hypothesis, we performed an NBT-dot blot on cell lysates of *E. coli* cells grown in various concentrations of C5-DOPA (**Figure 3.1(b)**). The protein dot-blot indicates that C5-DOPA is not accepted as a substrate by EcTyrRS. Thus, if an orthogonal aminoacyl-tRNA synthetase can be identified that will affix orthogonal tRNA to C5-DOPA, then C5-DOPA can be used for high fidelity catechol conjugation applications.

#### 3.3.2 Optimization of the fluorescence-based assay for C5-DOPA incorporation

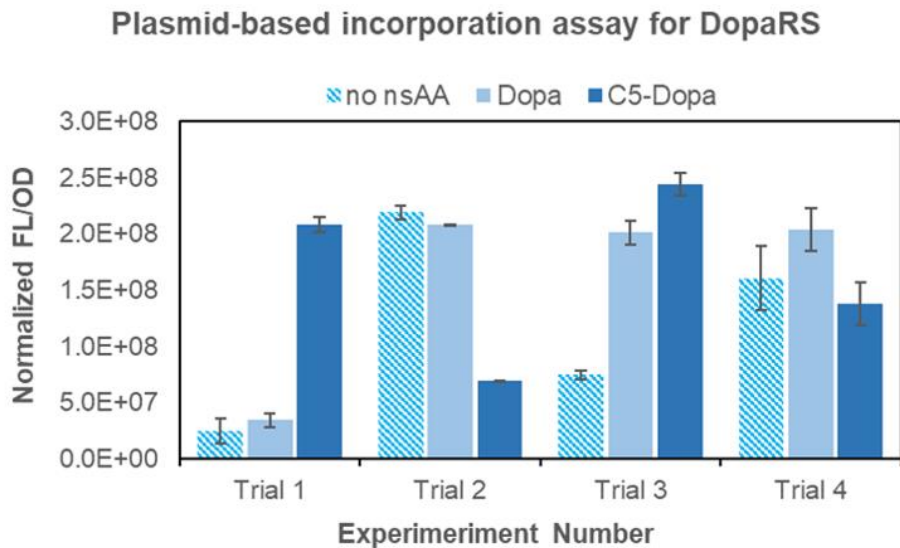
For genetic encoding of C5-DOPA we have initially used a plasmid-based reporter system. In this system, MG1655 *E. coli* strain was transformed with two plasmids, pEVOL and pZE. pEVOL plasmid was responsible for carrying the orthogonal aaRS/tRNA gene whereas pZE plasmid harbored the reporter protein with amber (UAG) codons in it to be suppressed by C5-DOPA incorporation. Due to varying plasmid copy number from cell-to-cell, and reliance of this assay on

fluorescence signal from sfGFP to indicate nsAA incorporation, the plasmid-based reporter system did not yield reproducible results. In **Figure 3.2**, fluorescence signal FL/OD for DOPA and C5-DOPA using DopaRS/tRNA<sup>Tyr</sup> pair has been compared with the negative control no nsAA condition for four different trials. FL/OD for the control and the actual samples varied from experiment-to- experiment, hence we switched the sfGFP reporter from plasmid to genome-integrated reporter system (**Figure 3.3**).

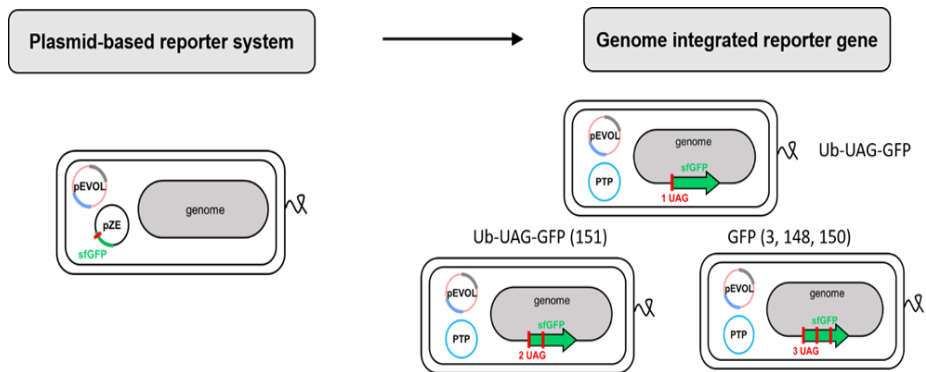
In genome-integrated sfGFP reporter system, the *E. coli* strain C321.  $\Delta A$  harbors a pEVOL plasmid and a posttranslational proofreading plasmid genomically integrated sfGFP reporter contains either one (situated at N-terminal), two, or three UAG codons situated at the N-terminus and a C-terminal 6X-His tag (**Figure 3.4**). Additionally, to make the incorporation process more efficient, we have used a “posttranslational proofreading- UBP1/ClpS-V65I” (PTP) system that reduces the background incorporation of standard amino acids (Kunjapur et al. 2018). **Figure 3.5** shows the results for two synthetases DopaRS and pAcFRS for C5-DOPA incorporation using this system. Compared to the plasmid-based system, the genome-integrated reporter system along with PTP plasmid can achieve reproducible incorporation of C5-DOPA.



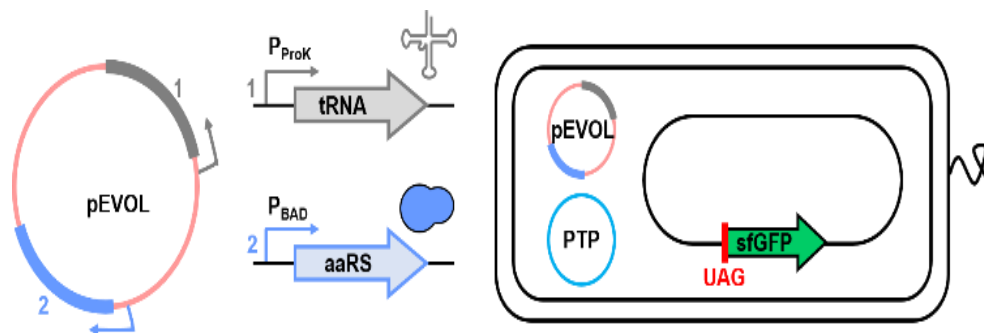
**Figure 3.1:** a) Structure of C5-DOPA. b) NBT-dot blot of cell lysate with various concentrations of C5-DOPA. c) Docking simulation of Tyrosine (pink) and C5-DOPA (yellow) as substrates in EcTyrRS using Autodock Vina.



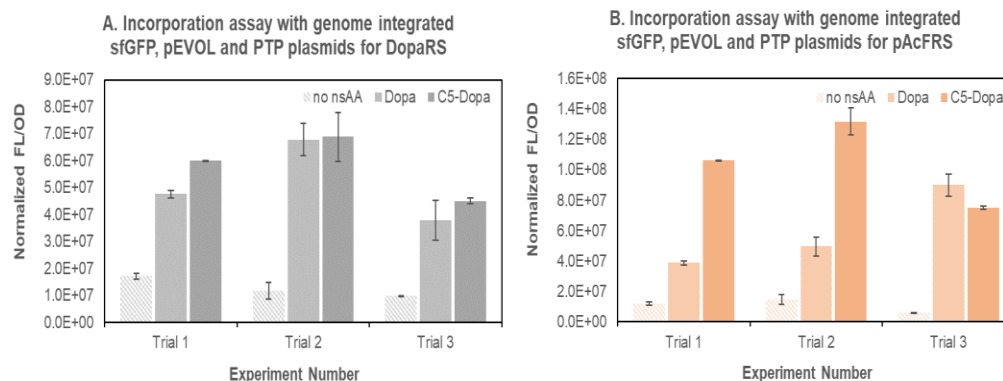
**Figure 3.2:** FL/OD graph for plasmid-based sfGFP reporter. Experiments were performed on separate days and in triplicate.



**Figure 3.3:** Switching from plasmid-based sfGFP reporter to genome integrated sfGFP reporter system. The reporter contains 1 UAG, 2 UAG and 3 UAG codons.



**Figure 3.4:** Schematic of the nsAA incorporation machinery and “post-translational proofreading” system in a recoded *E. coli* C321.ΔA strain with genome integrated sfGFP bearing N-terminal 1 UAG codon for site-specific incorporation of nsAA.



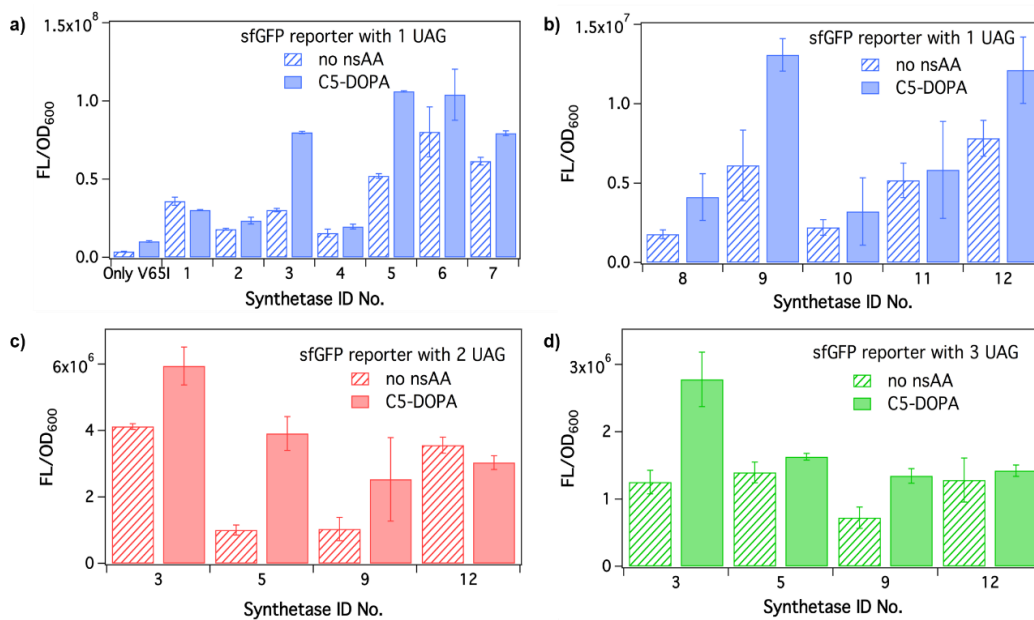
**Figure 3.5:** FL/OD graph for genome integrated sfGFP reporter. Reproducibility of the incorporation is shown using DopaRS (A) and pAcFRS (B) synthetases for DOPA and C5-DOPA. Experiments were performed on separate days and in triplicate.

### 3.3.3 Screening of an orthogonal aaRS-tRNA library for C5-DOPA incorporation in a reporter protein

Within this sub-aim, we seek to find optimal synthetase-tRNA pairs for C5-DOPA incorporation. To accomplish this, we will screen a small library of OTS (aaRS-tRNA pairs) that consists of well-characterized mutants from the MjTyrRS (Ryu and Schultz 2006; Alfonta et al. 2003a; S. Kim et al. 2018), PylRS (Wan, Tharp, and Liu 2014), and ScWRS (Hughes and Ellington 2010) families. Screening will be performed using a recoded *E. coli* C321.ΔA strain with a genomically integrated sfGFP reporter as mentioned earlier. Suppression of the N-terminal UAG with nsAA and full-length expression of the protein will result in fluorescence, which will be quantified using excitation and emission wavelengths of 485 and 525 nm, respectively, in a plate reader. At this point, library members that showed the largest differential in fluorescence signal over OD<sub>600</sub> between no nsAA and C5-DOPA was considered as the top candidates that later would be utilized for synthetase engineering.

**Figure 3.6 (a, b)** shows the data for incorporation of C5-DOPA in a sfGFP reporter with a single UAG codon. FL/OD<sub>600</sub> data suggested that aaRS-3 (DopaRS), aaRS-5 (pAcFRS), aaRS-9 (ScWRS), and aaRS-12 (CouRS) demonstrated activity on C5-DOPA. This result indicates that these synthetases could be used to incorporate C5-DOPA into specific sites in proteins. Furthermore, we utilized this subset to probe C5-DOPA incorporation into a multi-

UAG sfGFP reporter (**Figure 3.6 (c, d)**). For confirmation of C5-DOPA incorporation into sfGFP reporter we have performed ESI-Q-Tof LC-MS analysis.



**Figure 3.6:** Preliminary results of synthetase screen for 12 different aaRS against C5-DOPA. A small subset of the screened synthetases for sfGFP with 1 stop codon is shown in (a) and (b). The selected synthetases with 2 and 3 stop codons are shown in (c) and (d), respectively.

### 3.3.4 Confirmation of incorporation of C5-DOPA in single-site UAG sfGFP

To confirm the incorporation of C5-DOPA in sfGFP reporter protein, we have attempted LC-MS analysis (**Figure 3.7 - 3.11**). The expected and observed MWs of the samples are tabulated in **Table 3.1**. The control for this experiment was Ub-M-sfGFP and Ub-S-sfGFP with methionine and serine amino acid in place of UAG stop codon. In presence of the ubiquitinase enzyme (UBP1), it was expected that these constructs will be stabilized and would yield finally M-sfGFP and S-sfGFP protein constructs.

For expression of M-sfGFP control, the *E. coli* strain did not contain PTP plasmid which was responsible for UBP1 expression followed by degradation of ubiquitin part from the protein fusion construct. We have observed a mass of 37,140.0 Da for the protein fusion Ub-M-sfGFP. For expression of the second control, S-sfGFP, the *E. coli* strain was transformed with PTP plasmid. Due to degradation of the ubiquitin part from the protein fusion construct, an expected mass would be 28593.02 Da. We have observed a MW of 34,296 Da for this construct.

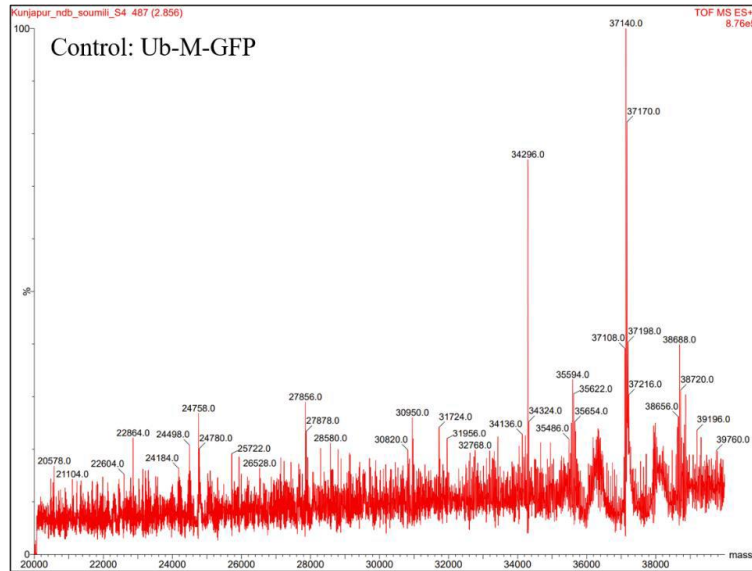
The reported MW of sfGFP is 26,800 Da, whereas for sfGFP with N-degron and incorporated C5-DOPA at the N-terminal of the fusion expected MW was 28,712.93 Da. Expression of 1-UAG sfGFP protein in presence of pAcFRS synthetase pair and C5-DOPA, resulted in a protein of observed MW 37,302.0 Da. Similar analysis was performed for pAcFRS.2t.1-1 and ScWRS synthetases, where the observed MW was 34,296 Da and 33,468 Da respectively.

A difference in the expected and observed MW was observed during the analysis. This change in MW can be attributed to multiple reasons. First, lower expression of UBP1 enzyme which was responsible for cleavage of the ubiquitin from the Ub-UAG-sfGFP protein construct or partial cleavage of the ubiquitin might have resulted in higher MW products. Second, since catechols are susceptible to get modified whenever a change in solvent is encountered, this may lead to unwanted mass change in LC-MS.

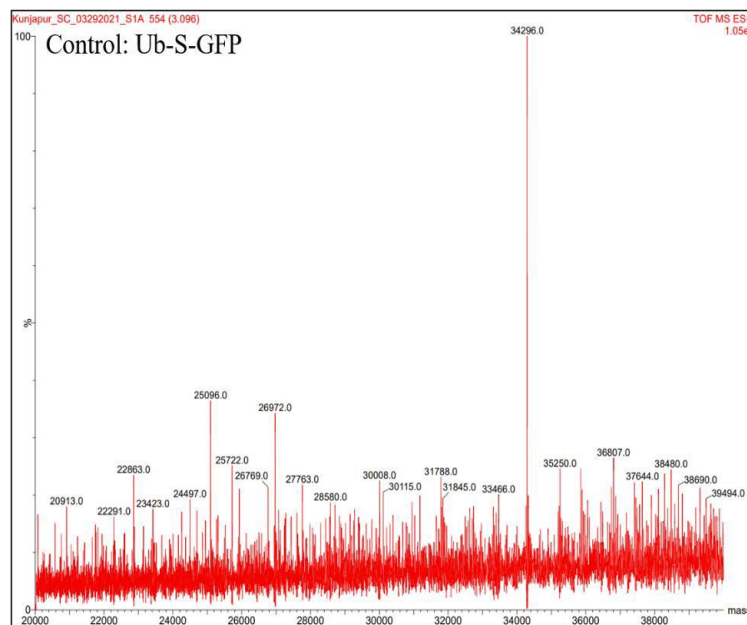
**Table 3.1:** Expected and observed molecular weight (MW) for LC-MSanalysis.

(NT – Not Tested)

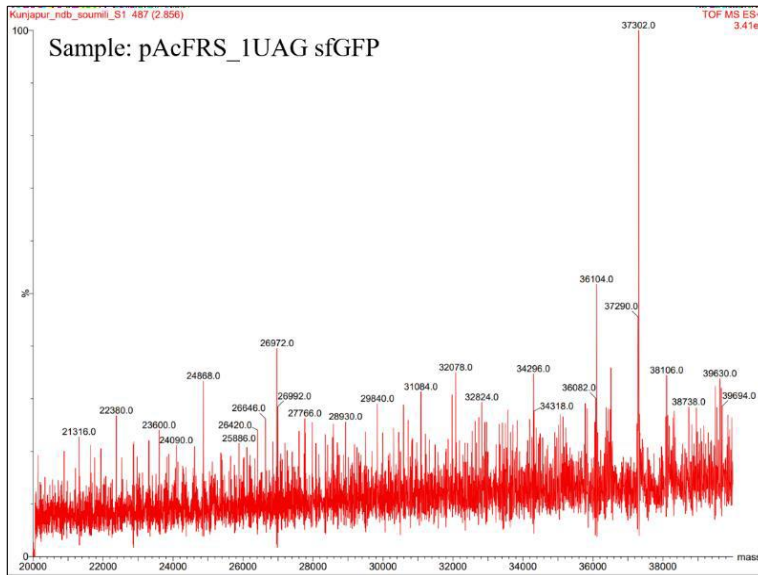
| <b>Sample Name</b> | <b>Expected MW (Da)</b> | <b>Observed MW (Da)</b> |
|--------------------|-------------------------|-------------------------|
| Ub-M-sfGFP         | 37193.86                | 37,140.0                |
| M-sfGFP            | 28637.14                | NT                      |
| Ub-S-sfGFP         | 37149.74                | NT                      |
| S-sfGFP            | 28593.02                | 34,296                  |
| Ub-C5-DOPA-sfGFP   | 37269.65                | NT                      |
| C5-DOPA-sfGFP      | 28712.93                | 37,302                  |



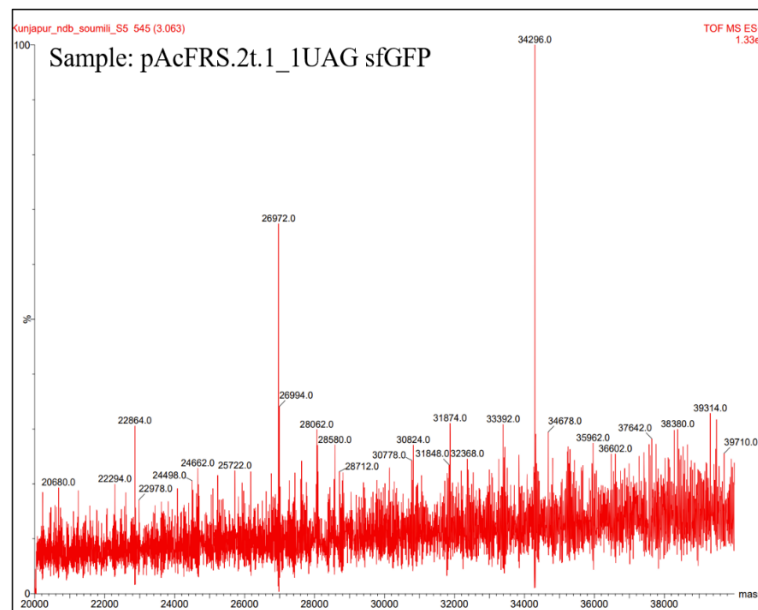
**Figure 3.7:** LC-MS analysis for Ub-M-GFP. This is one of the controls for the experiment. For this control the strain did not contain PTP plasmid. Hence there was no expression of Ubiquitinase. Observed molecular weight = 37140 Da.



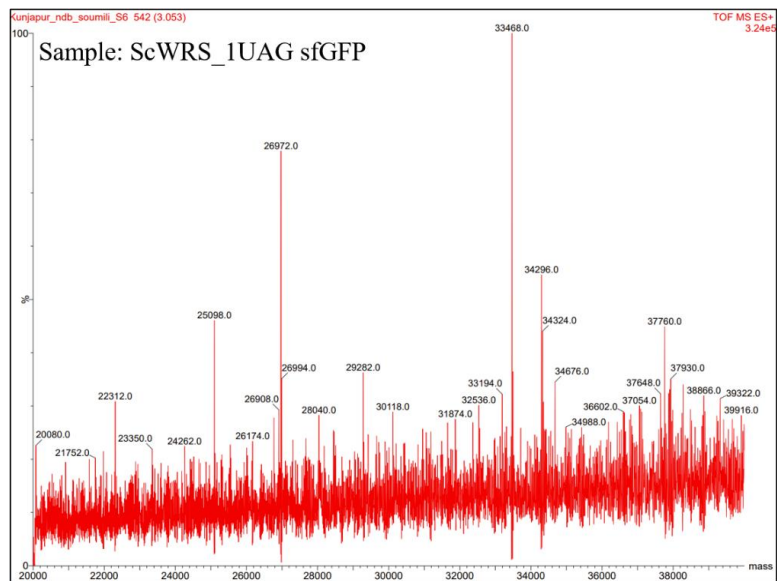
**Figure 3.8:** LC-MS analysis for Ub-S-GFP. This is one of the controls for the experiment. For this control the strain harbored the PTP plasmid. Observed molecular weight = 34296 Da.



**Figure 3.9:** LC-MS analysis for pAcFRS\_1UAG sfGFP.  
Observed molecular weight = 37302 Da.



**Figure 3.10:** LC-MS analysis for pAcFRS.2t.1\_UAG sfGFP.  
Observed molecular weight = 34296 Da.



**Figure 3.11:** LC-MS analysis for ScWRS\_1UAG sfGFP.  
Observed molecular weight = 33468 Da.

---

### 3.4 Summary

In this Chapter, we were able to demonstrate the suitability of extended sidechain catechol amino acid for ribosomal incorporation during translation process. C5-DOPA did not show any misincorporation in place of tyrosine and thus can be used in place of DOPA for biosynthesis of bioconjugates or biomaterial. For C5-DOPA incorporation, we relied on a recoded *E. coli* strain with genome integrated sfGFP with UAG codons which provided reproducible data. This strain also harbored a proofreading plasmid that was responsible for degrading the proteins incorporated with standard amino acids like tyrosine at the N-terminal UAG. Finally, we were able to show that using the fluorescence assay that MjTyrRS and ScWRS synthetase have activity towards C5-DOPA and can be used for engineering for efficient multi-site incorporation.

## **Chapter 4.**

### **INVESTIGATING CATECHOL-BORONATE COUPLING CHEMISTRY IN CONTEXT OF EXTENDED SIDE-CHAIN CATECHOLS**

#### **4.1 Introduction and background**

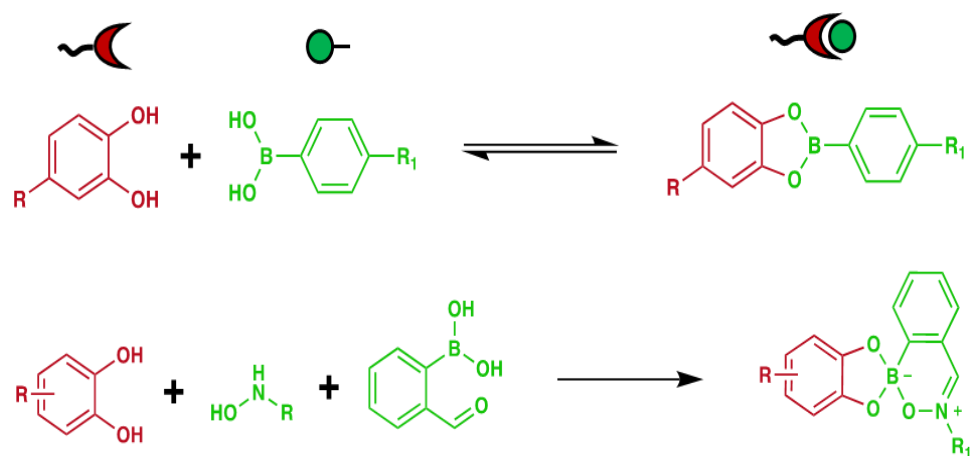
Protein engineering has tremendous opportunity in improving the human health both from the perspective of drug delivery and therapeutics. Many previous therapeutic methods have utilized the biospecificity and diversity of protein functions to treat many diseases, including cancers and autoimmune diseases. Furthermore, proteins are ideal candidates for drug encapsulation and delivery because of their intrinsic properties like biocompatibility and biodegradability. The efficacy of such proteins is limited owing to their poor stability, impermeability, and fast degradation rate. To further enhance target specificity, delivery rate and cell permeation, different biomolecules have been conjugated to proteins (Singh-Blom, Hughes, and Ellington 2013). Typically, naturally occurring amino acids like lysine and cysteine have been utilized for conjugating functional moieties (Axup et al. 2012). However, this technique has several shortcomings including non-site specificity, heterogeneity and alteration of protein native structure, which often leads to compromised bioactivity for therapeutic and drug delivery applications. Site-specific protein bioconjugation is a unique biological tool to build highly tunable protein constructs to enable target-specific drug delivery, prolonging

bioactivity and minimizing side-effects in cellular environments (Lieser et al. 2020). Boronate-diol conjugation chemistry has been established for several biomedical applications, such as in vivo carbohydrate and dopamine sensing (Elsherif et al. 2018; Liu et al. 2019) as well as for controlled and targeted release of drugs such as insulin (Mansoor et al. 2019). But boronate complexation with catechols has been sparsely studied for protein conjugation. This conjugation scheme offers several advantages, such as control over reversibility, reactivity under physiological conditions, and pH-dependent binding characteristics. To the best of our knowledge, bioconjugation between proteins incorporated with C5-DOPA and boronic acid derivatives has never been established before. Specifically, in this aim, we will assess the viability of this coupling chemistry for our incorporated catechol substrates in proteins and different boronic acid derivatives that can be relevant for live cell imaging and drug delivery purpose. The final outcome of this aim will be to establish C5-DOPA incorporated protein and boronic acid conjugation platform, with a longer-term goal of enabling live-cell conjugation.

Within this aim, first we will focus on establishing a small-molecule, diol-boronate reaction for our catechol substrates (DOPA and C5-DOPA) under specific solvent conditions (**Figure 4.1**). To our benefit, a tripartite system has been developed that includes the catechol-based, fluorescent reporter Alizarin Red S (ARS), boronic acid derivative, and a competing diol-containing compound (Springsteen and Wang 2001; Gennari et al. 2017) (**Figure 4.2**). We will use this system as a reference assay to investigate the reaction parameters including

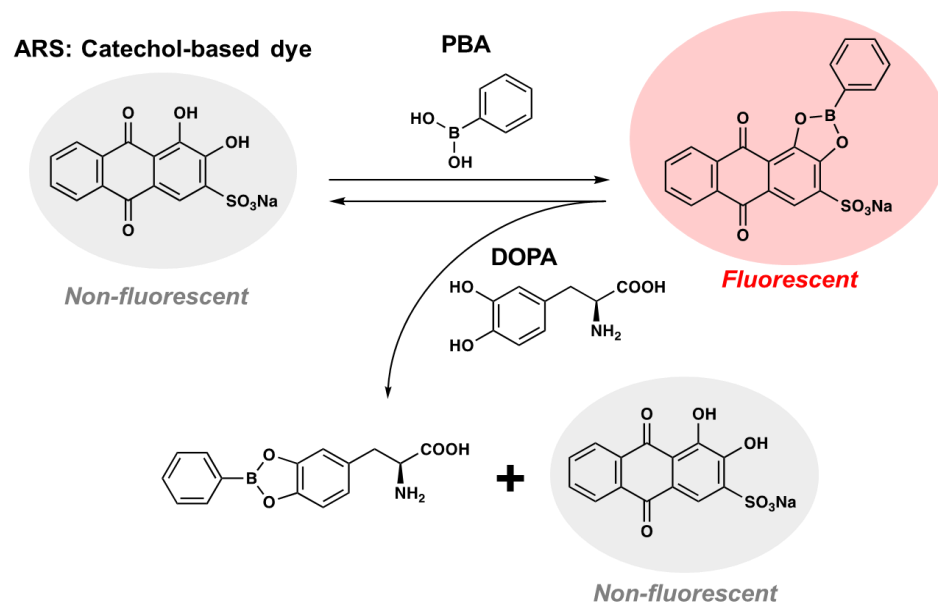
equilibrium constant and pH sensitivity between our catechol substrates and different boronic acid derivatives. In brief, ARS coupled to boronic acid derivative exhibits fluorescence at 572 nm and, after introduction of a competing catechol, a fraction of ARS-boronic acid decouples to result in a fluorescence drop that is proportional to apparent  $K_{eq}$ .

To expand our library of boronic acid derivatives with an eye towards applications in biosensing and drug delivery, we will screen small molecules fluorescent boronic acid probes. Our initial experiments are focused on the commercially available compounds of 3-(Dansylamino) phenylboronic acid (Dns-PBA). After establishing the small-molecule chemistry, we will optimize the reaction conditions for catechol containing proteins. This work is going to act as foundation for catechol-boronate chemistry using site-specific nsAA incorporation technology.



**Figure 4.1:** Reaction scheme for reversible and irreversible catechol-boronate conjugation.

---



**Figure 4.2:** Scheme of the ARS assay.

---

## **4.2 Materials and Methods**

### **4.2.1 Chemicals**

Alizarin Red S dye (ARS) was purchased from Alfa Aesar (130-22-3). 3-(Dansylamino)phenylboronic acid was purchased from Sigma Aldrich (75806-94-9). Phenylboronic acid (PBA) was purchased from TCI (98-80-6). Sodium phosphate monobasic anhydrous (7558-80-7) and sodium phosphate dibasic anhydrous (7558-79-4) was purchased from VWR. 1 mM of ARS stock was made in water for all the experiments and stored in 2-8 °C in dark tubes for 1 week. 100 mM Phenylboronic acid stock was made in water. 100 mM of DOPA and C5-DOPA stock solution was made in 50-100 mM NaOH.

### **4.2.2 ARS assay – small molecule conjugation**

ARS-PBA assay was setup in 50 mM phosphate buffer at pH 7.5 by mixing 10  $\mu$ M of ARS in varying concentration of PBA – 0, 1, 2.5, 5, 7.5, 10. For the fluorescence measurement, a buffer blank and 10  $\mu$ M ARS in buffer blank was used as the control. Samples were incubated for 15 min at room temperature before measuring fluorescence intensity by exciting the samples at 468 nm and analyzing emission at 572 nm in 96-well plates using Spectramax spectrophotometer instrument. All the readings were taken in duplicate.

$K_{eq}$  to analyze the ARS-PBA and ARS-PBA-catechol interaction from this assay was calculated based on fluorescence intensity data (Brooks, Deng, and Sumerlin 2018). For ARS-PBA interaction, the value of  $1/(FL-FL_0)$  vs.  $1/[phenylboronate]$  where FL = fluorescent intensity at the specified phenylboronate

concentration and  $FL_0$  is the fluorescent intensity in the absence of phenylboronate was plotted. Finally, the  $K_{eq}$  is obtained by dividing the intercept of the line by the slope of the line. For ARS-PBA-catechol interaction, in 50 mM phosphate buffer pH 7.5 with 10  $\mu$ M ARS and 0.25 mM PBA concentration and varying concentration of DOPA – 0, 1, 2.5, 5, 7.5, 10 mM was used. Concentrations were selected such that if PBA concentration is denoted as  $n$ , catechol concentration was varied from  $4n - 40n$ .  $K_{eq}$  was then calculated as shown in the literature.

#### **4.2.3 3-(Dansylamino)phenylboronic acid (Dns-PBA) and catechol small-molecule conjugation**

Dns-PBA conjugation assay was setup in 50 mM phosphate buffer pH 7.5 with 0.5 mM Dns-PBA mixed in varying concentration of catechols-0, 1, 2.5, 5, 7.5, 10. Fluorescence intensity was measured by exciting the samples at 337 nm and obtaining fluorescence at 528 and 490 nm.

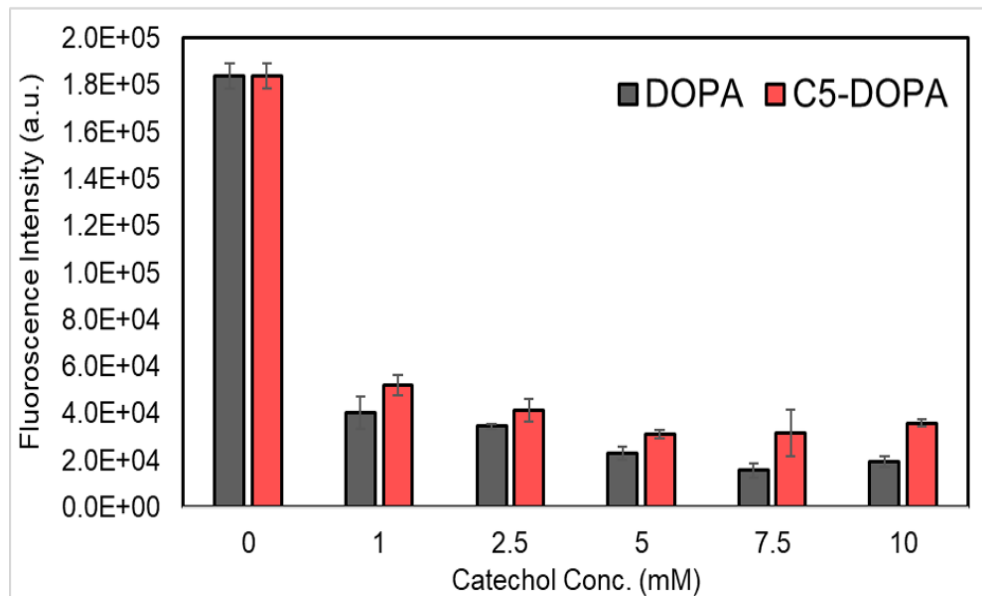
#### **4.2.4 Dansyl-phenyl boronic acid and C5-DOPA protein conjugation**

Dns-PBA protein conjugation assay was setup in 50 mM phosphate buffer pH 7.5 with 0.1 mM Dns-PBA mixed in protein solution. Fluorescence intensity was measured by exciting the samples at 337 nm and obtaining fluorescence at 528 and 490 nm.

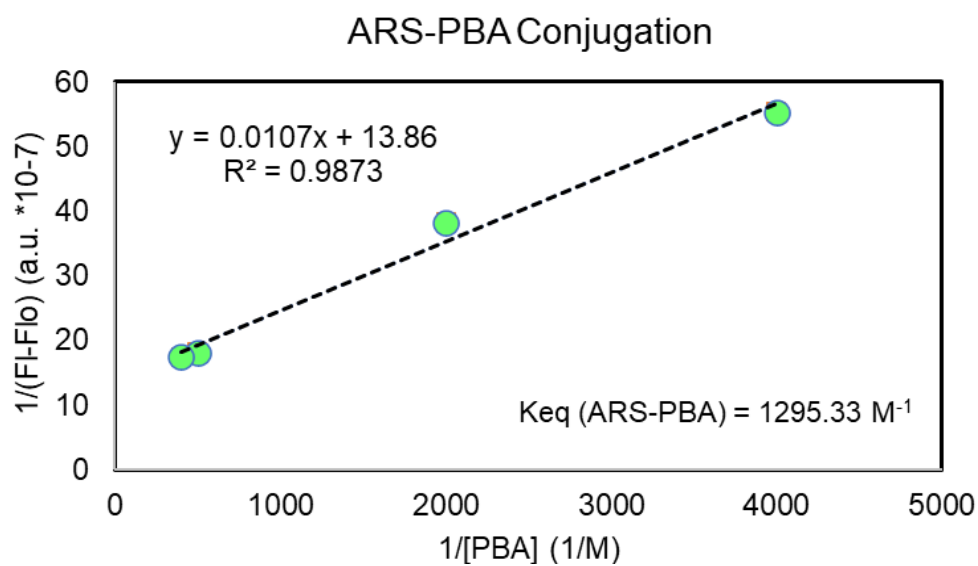
## 4.3 Results and Discussion

### 4.3.1 Optimizing ARS assay for catechols conjugation

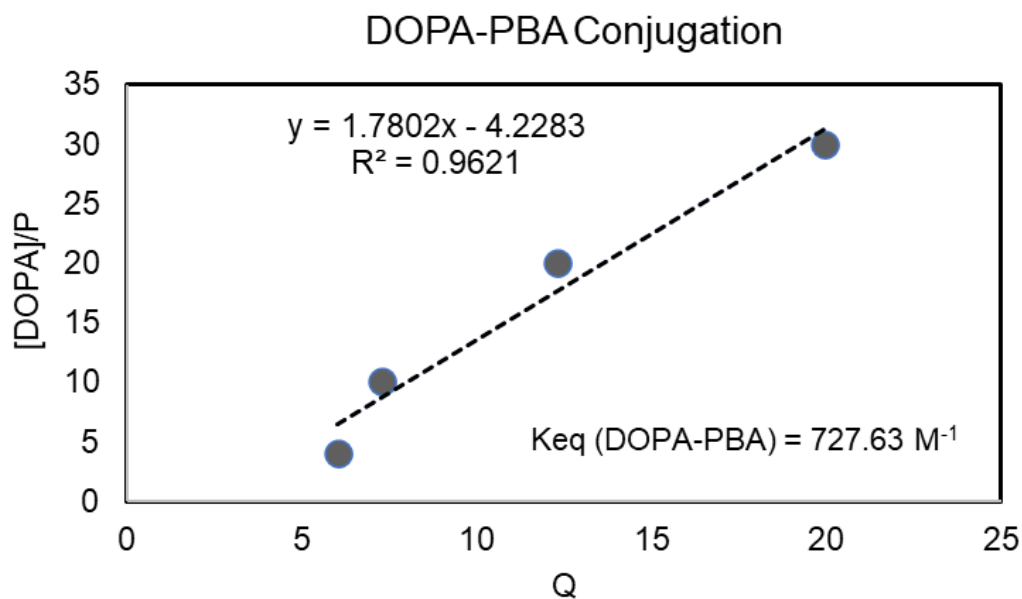
ARS assay was performed to determine the equilibrium constants ( $K_{eq}$ ) of catechols with phenylboronic acid (PBA). **Figure 4.2** showed the conjugation of DOPA and C5-DOPA with PBA. Due to decoupling of ARS from PBA in presence of a competing catechol, there was a decrease in fluorescence intensity for samples containing catechols (**Figure 4.3**). In **Figure 4.4**,  $K_{eq}$  for ARS and PBA was calculated from the plot of  $1/(FL-FL_0)$  vs.  $1/[phenylboronate]$ . Observed  $K_{eq}$  was  $1295.33\text{ M}^{-1}$  which is comparable to the  $K_{eq}$  value obtained for ARS-PBA in literature (Yan et al. 2004). To calculate the  $K_{eq}$  between ARS-PBA-catechol we adopted the method mentioned in the literature (Brooks, Deng, and Sumerlin 2018). Based on the method, we observed  $K_{eq}$  for DOPA and C5-DOPA as  $727.83\text{ M}^{-1}$  (**Figure 4.5**) and  $238.26\text{ M}^{-1}$  (**Figure 4.6**) respectively.  $K_{eq}$  value for DOPA is comparable to the literature value of  $830\text{ M}^{-1}$  (Yan et al. 2004). The  $K_{eq}$  value reported for glucose is  $4.6\text{ M}^{-1}$  at pH 7.5 implying that glucose have lower binding affinity towards boronic acid derivatives at physiological pH compared to both the catechols.



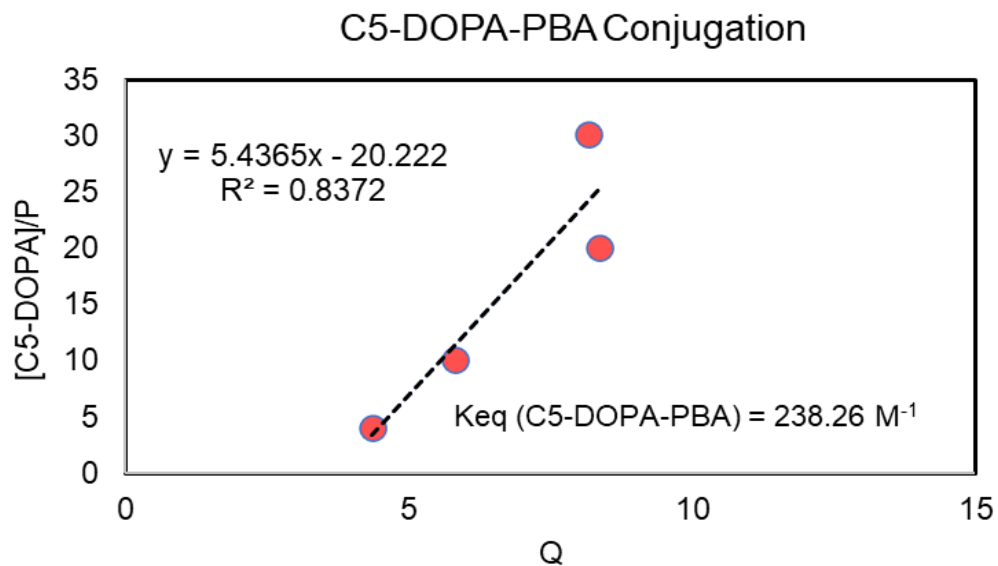
**Figure 4.3:** Change in fluorescence intensity for ARS-PBA assay in presence of competing catechol, DOPA, C5-DOPA.



**Figure 4.4:** Calculation of equilibrium constant based on ARS binding with PBA.



**Figure 4.5:** Calculation of equilibrium constant based on DOPA binding with PBA.

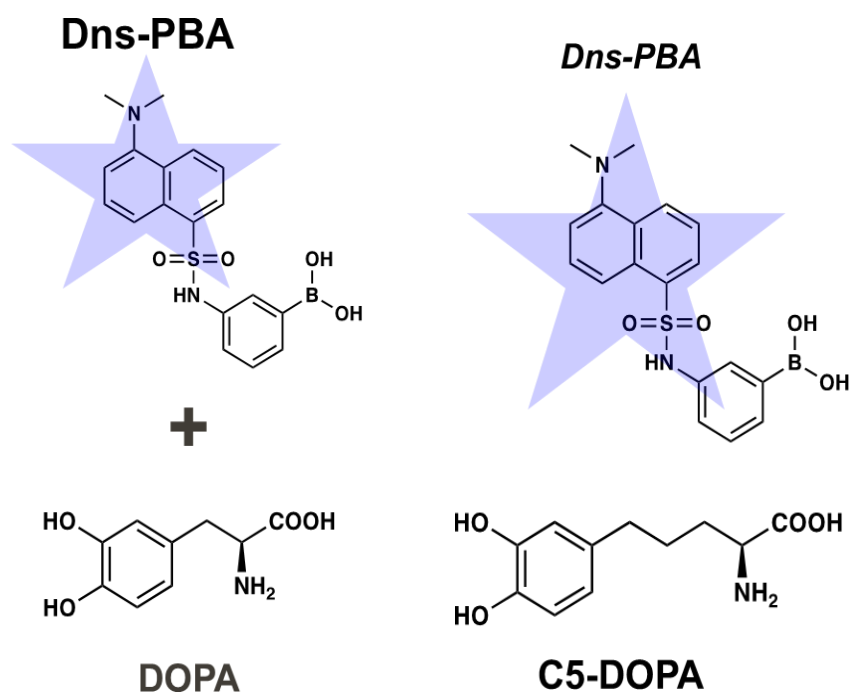


**Figure 4.6:** Calculation of equilibrium constant based on C5-DOPA binding with PBA.

### 4.3.2 Expanding the library of boronic acid derivative

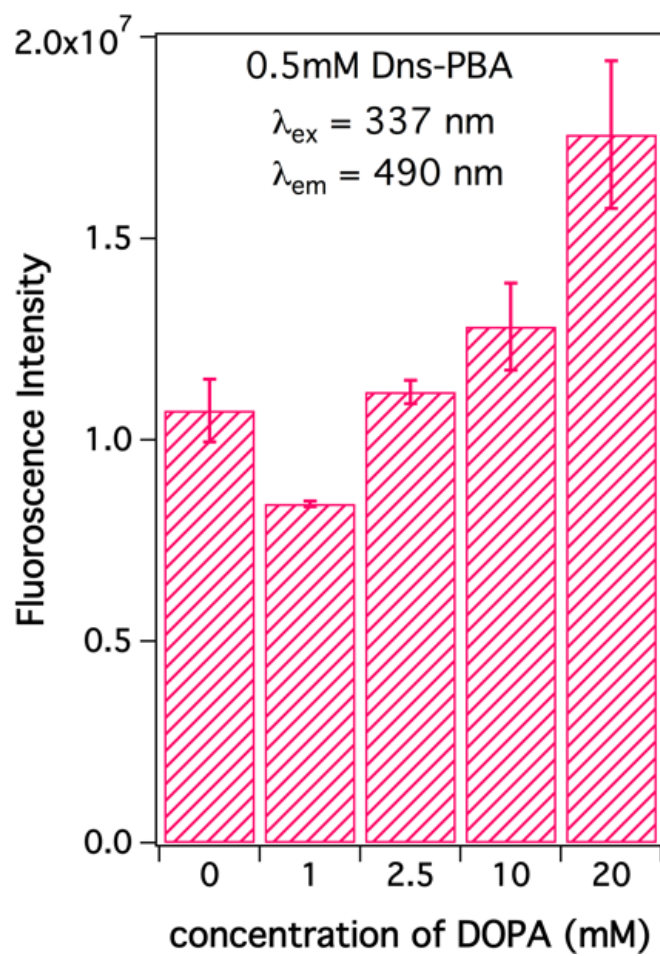
ARS assay is an indirect way of evaluating the conjugation between catechols and boronic acid derivative. Hence, we seek to utilize a fluorescently labeled boronic acid derivative that shows a change in the fluorescence intensity upon conjugation. Specifically we will use 3-(Dansylamino)phenylboronic acid (Dns-PBA) that has excitation of 337 nm and emission at 528 nm in methanol and 490 nm for protein adduct or conjugate. Herein we have tested both of our catechols propensity to bind and form conjugate with Dns-PBA (**Figure 4.7**).

**Figure 4.8 - 4.9** shows the result for DOPA and Dns-PBA conjugation. Here, we have titrated Dns-PBA with varying concentrations of DOPA. After reaction, we have measured fluorescence by exciting the samples at 337 nm and emission at 490 nm. We observe that with increasing DOPA concentration there is an increase in intensity and blue-shift in the spectra which suggests conjugate formation. In the same way, we have also performed the conjugation reaction for C5-DOPA with the same conditions and the result is shown in **Figure 4.10 – 4.11**. For C5-DOPA, although we have not observed a significant shift in the spectra with varying concentration of C5-DOPA but we did observe an increase in the fluorescence intensity at 490 nm which suggests the formation of conjugate.



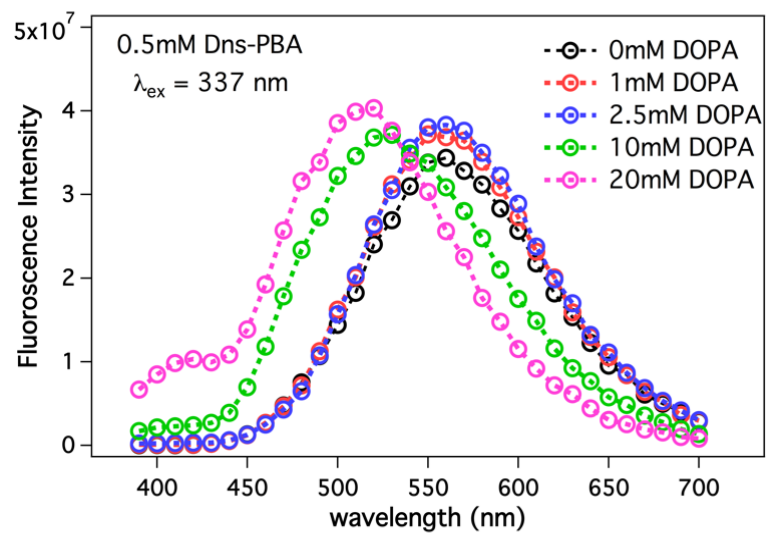
**Figure 4.7:** Reaction of Dns-PBA with DOPA and C5-DOPA.

---



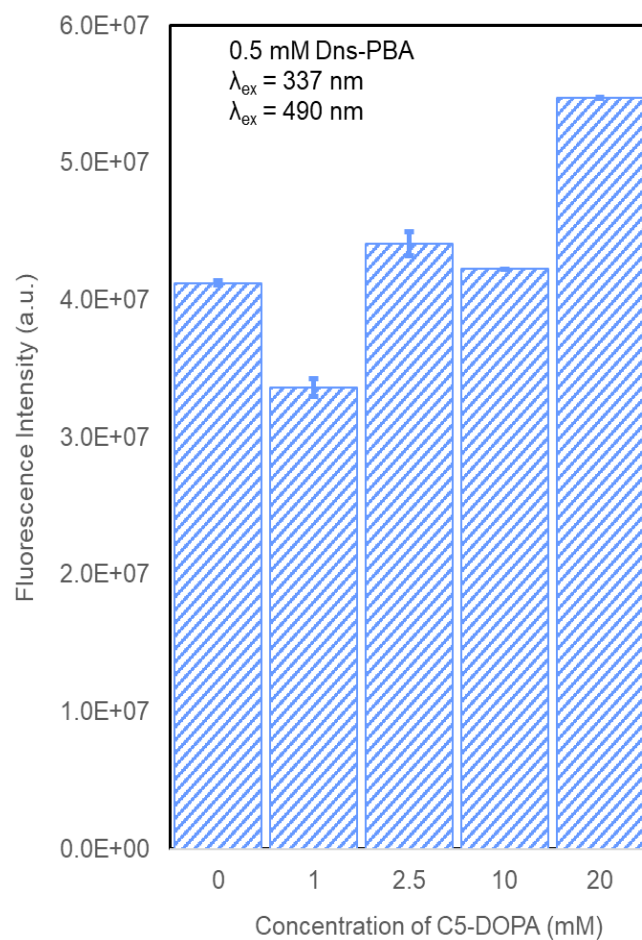
**Figure 4.8:** Fluorescence intensity for DOPA and Dns-PBA conjugation reaction

---



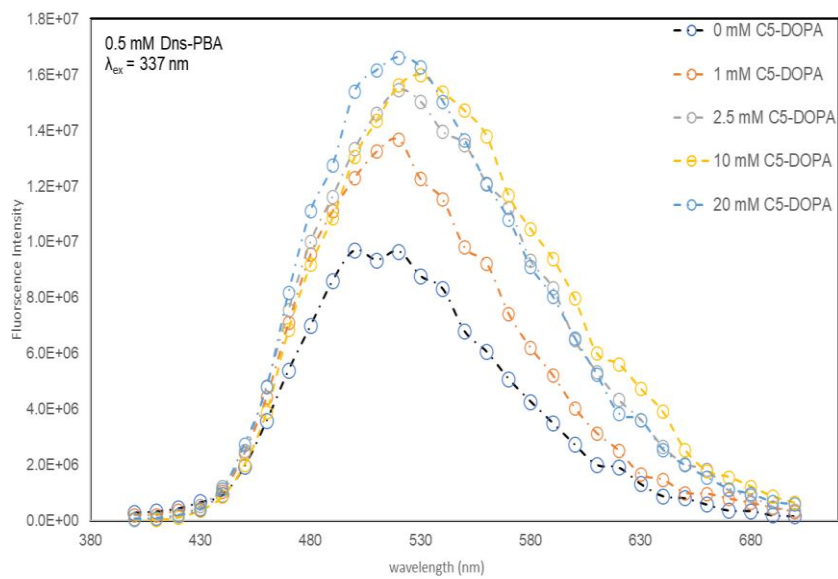
**Figure 4.9:** Fluorescence spectra for DOPA and Dns-PBA conjugation reaction

---



**Figure 4.10:** Fluorescence intensity for C5-DOPA and Dns-PBA conjugation reaction

---



**Figure 4.11:** Fluorescence spectra for C5-DOPA and Dns-PBA conjugation reaction

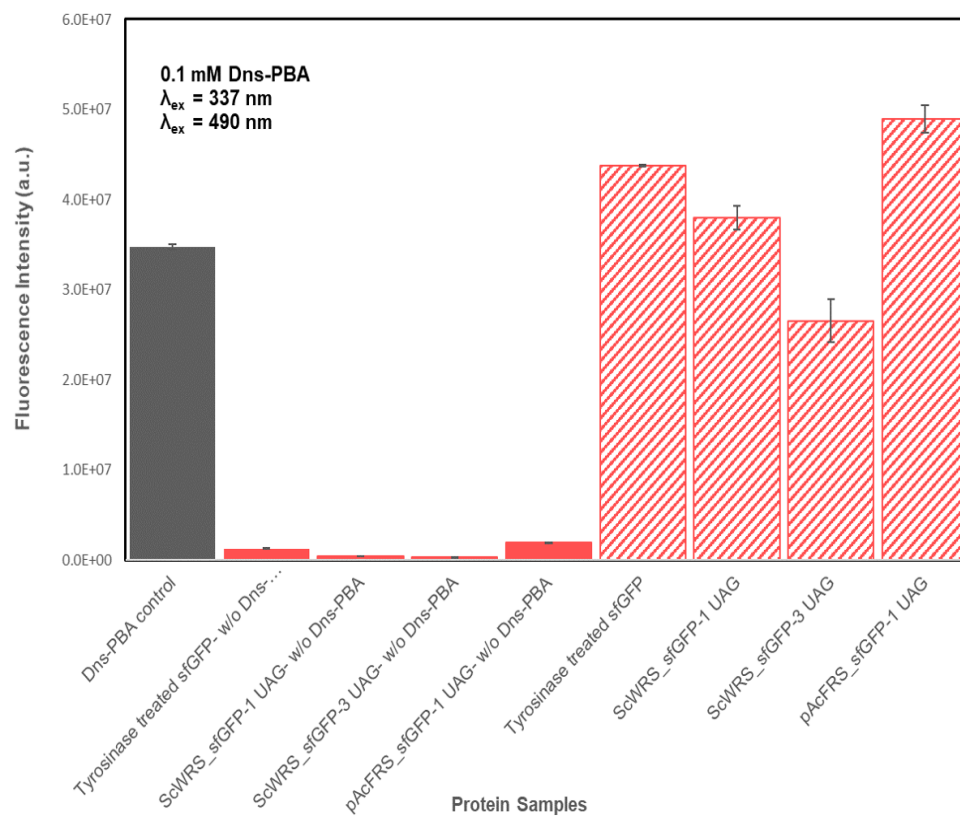
---

### 4.3.3 Optimization of catechol-boronate chemistry *in-vitro*

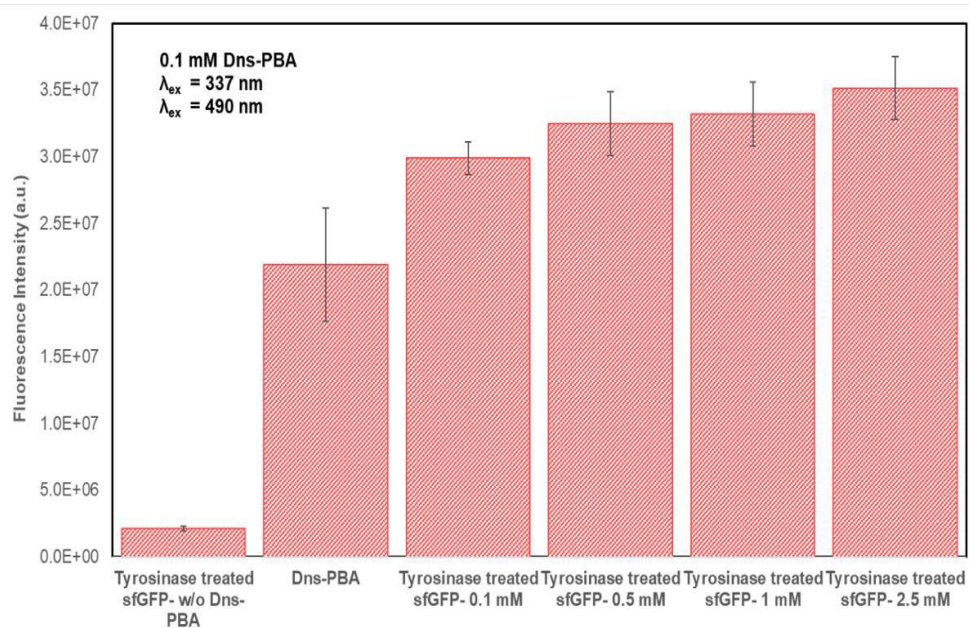
Conjugation of Dns-PBA with purified protein was performed under the same conditions as small molecule except a lower concentration of 0.1 mM of Dns-PBA was used. This reaction was carried out for 1 hour and the reaction was stopped by filtering the samples using 10 KDa centrifugal filter. We observed a slight increase in the fluorescence intensity for protein samples (I) tyrosinase treated sfGFP, (II) pAcFRS – 1 UAG\_C5-DOPA incorporated compared to the Dns-PBA control (**Figure 4.12**).

To understand better whether the conjugation is taking place, we buffer exchanged the tyrosinase treated sfGFP to a volatile buffer, ammonium acetate that is LC-MS compatible. It is expected that upon conjugation with boronic derivative, Dns-PBA (MW = 370.23 g/mol) there will be an increase in mass of the protein as compared to unconjugated protein. **Figure 4.14** shows the protein sample without Dns-PBA and **Figure 4.15** shows the protein sample treated with Dns-PBA for 1 hour. We observed an increase of 1808 Da change in the molecular weight of the protein.

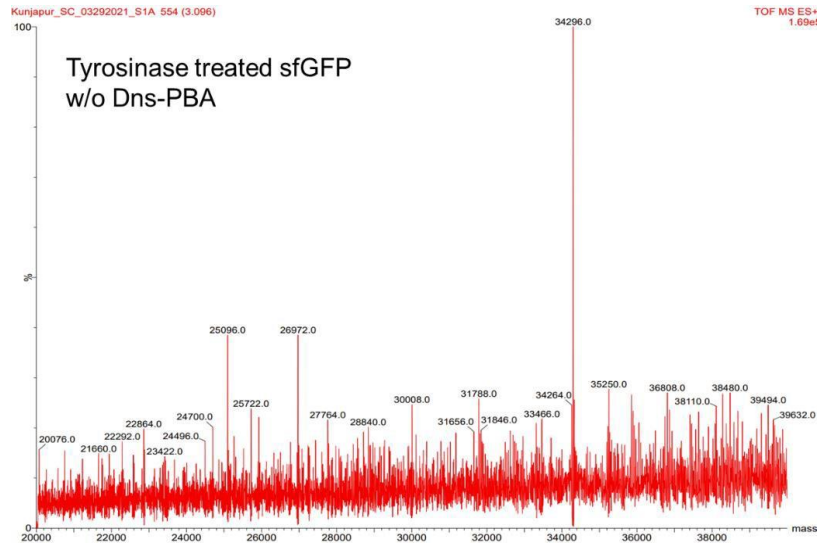
To further optimize the protein conjugation reaction, we have titrated 0.1 mM Dns-PBA with varying concentrations of tyrosinase treated sfGFP and incubated all the samples at room temperature for 1 hour. **Figure 4.13** shows the result for protein conjugation with tyrosinase treated sfGFP, it can be observed that for all the protein samples there is an increase in fluorescence intensity but there is no significant increase in intensity across the samples.



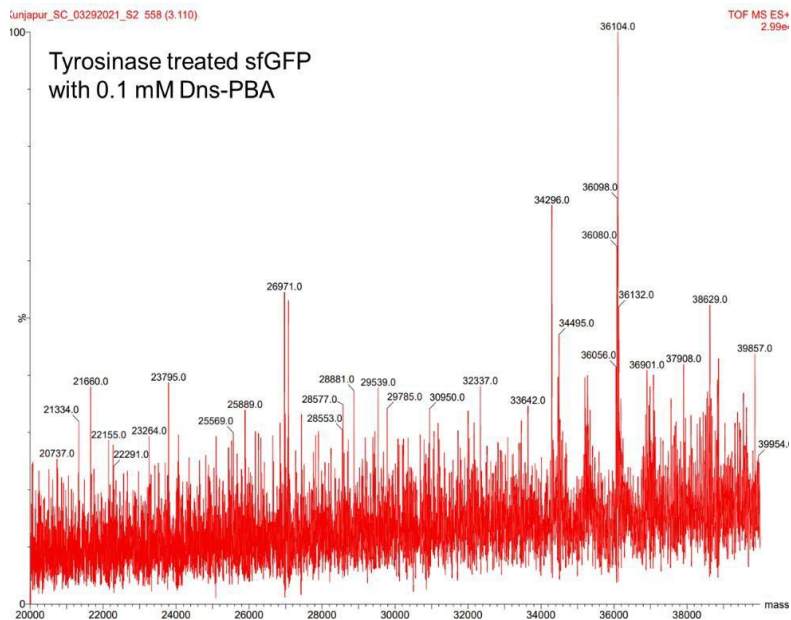
**Figure 4.12:** Fluorescence intensity change for protein samples upon conjugation with Dns-PBA



**Figure 4.13:** Fluorescence intensity for varying protein concentration upon conjugation with Dns-PBA



**Figure 4.14:** LC-MS analysis for Tyrosinase treated sfGFP in absence of Dns-PBA



**Figure 4.15:** LC-MS analysis for Tyrosinase treated sfGFP in presence of 0.1mM Dns-PBA

#### 4.4 Summary

This chapter includes our attempt on establishing catechol-boronate reaction for our catechol substrates DOPA and C5-DOPA. First we investigated the  $K_{eq}$  values for the catechol substrates using ARS assay. The  $K_{eq}$  values were obtained to be  $727.63M^{-1}$  and  $238.26M^{-1}$  for DOPA and C5-DOPA, respectively. Although the observed  $K_{eq}$  for C5-DOPA is lower than DOPA but it is still higher than reported value of glucose implying that C5-DOPA has higher binding affinity towards PBA as compared to glucose. Further, we attempted on establishing the conjugation reaction for a fluorescent boronic acid derivative, Dns-PBA and both the catechol substrates. When excited at 337 nm, Dns-PBA in conjugation with diols shows maximum fluorescence at 490 nm. To study this effect, we have used varying concentrations of catechols in a solution containing fixed amount of Dns-PBA and measured fluorescence intensity. For both the catechols, we have observed an increase in intensity when conjugated to Dns-PBA. The blue shift in fluorescence spectra was observed for higher concentrations of DOPA, unlike C5-DOPA, hence deeper understanding of the reaction conditions for C5-DOPA needs to be obtained. To extend the reaction to proteins, we have utilized protein samples containing single and multiple instances of catechols. A slight increase in fluorescence intensity was observed for tyrosinase treated sfGFP and pAcFRS-1UAG sfGFP compared to the control, Dns-PBA. Further to confirm the conjugate formation, we have attempted LC-MS analysis of tyrosinase treated sfGFP with and without Dns-PBA. We have observed a difference of 1808 Da in MW of tyrosinase treated sfGFP with

and without Dns-PBA, which suggests the formation of conjugates between catechols present in the protein and Dns-PBA.

## Chapter 5.

### CONCLUSIONS AND DIRECTIONS FOR FUTURE WORK

#### 5.1 Conclusions

In this thesis, we have described a study on incorporation of catechol-based nsAA into protein of interest and how the modified protein can be utilized for bioconjugation. Our study acts as a foundation for site-specific multi-site incorporation of catechol-based extended sidechain amino acid and using it for long term purpose of live cell conjugation with boronic acid derivative.

In Chapter 2, we investigated background misincorporation of DOPA into the *E. coli* proteins. We focused on a redox staining assay for catechol detection in the bacterial proteome. The result aligns with the literature evidence of DOPA misincorporation in mammalian cells. We also performed molecular docking to understand the mechanism behind misincorporation of DOPA in place of tyrosine.

In Chapter 3, we investigated an alternative nsAA to DOPA for site-specific incorporation into proteins. C5-DOPA is an extended sidechain catechol amino acid, which has never been explored before for ribosomal incorporation. Hence, we performed a detailed screening of the orthogonal translation machinery for activity on C5-DOPA. We performed the screening on single and multi-site UAG sfGFP reporters with an eye on end applications such as bioconjugation and biomaterials production.

In Chapter 4, we explored the reactivity of boronic acid derivatives on DOPA and C5-DOPA. To the best of our knowledge, catechol-boronate reaction has never been explored for C5-DOPA. Using an indirect method, ARS assay, we evaluated the  $K_{eq}$  for C5-DOPA and compared it to DOPA. Observed  $K_{eq}$  value for ARS-PBA and PBA-DOPA are  $1295.33 \text{ M}^{-1}$  and  $727.63 \text{ M}^{-1}$  respectively is comparable to the reported values. Further, we have explored the conjugation propensity of catechol with a boronic acid with a fluorescent functional group. We were able to establish the conjugation with the small molecules of catechol and attempted the same reaction with proteins containing catechols. On the basis of Chapter 3 screening experiments, pAcFRS – 1 UAG sfGFP\_C5-DOPA came out to be the top candidate. We used the same sample for conjugation reaction and observed slight increase in the fluorescence intensity after incubation.

## **5.2 Directions for future work**

### **5.2.1 Optimizing the shake flask culture conditions for C5-DOPA incorporation study**

In Chapter 3, we have shown the data for LC-MS analysis of purified Ub-M-sfGFP, S-sfGFP and C5-DOPA-sfGFP. The observed MW for these products does not match their expected MW. To understand the unexpected change in MW of the purified products, further investigations are required.

We have observed an increased MW of the purified products, which can be due to lower expression of UBP1 protease enzyme. A plausible reason can be partial

degradation of ubiquitin by the enzyme leading to formation of heterogeneous products. Shake flask culture conditions such as induction time, inducer concentration, strain type etc. should be optimized for UBP1 expression. Alongside, culture conditions, optimum C5-DOPA and ascorbic acid concentration should be determined for shake flask cultures. To confirm catechol incorporation in sfGFP, western blot followed by NBT staining should be performed along with LC-MS.

### **5.2.2 Investigating incorporation study for other extended side chain catechols**

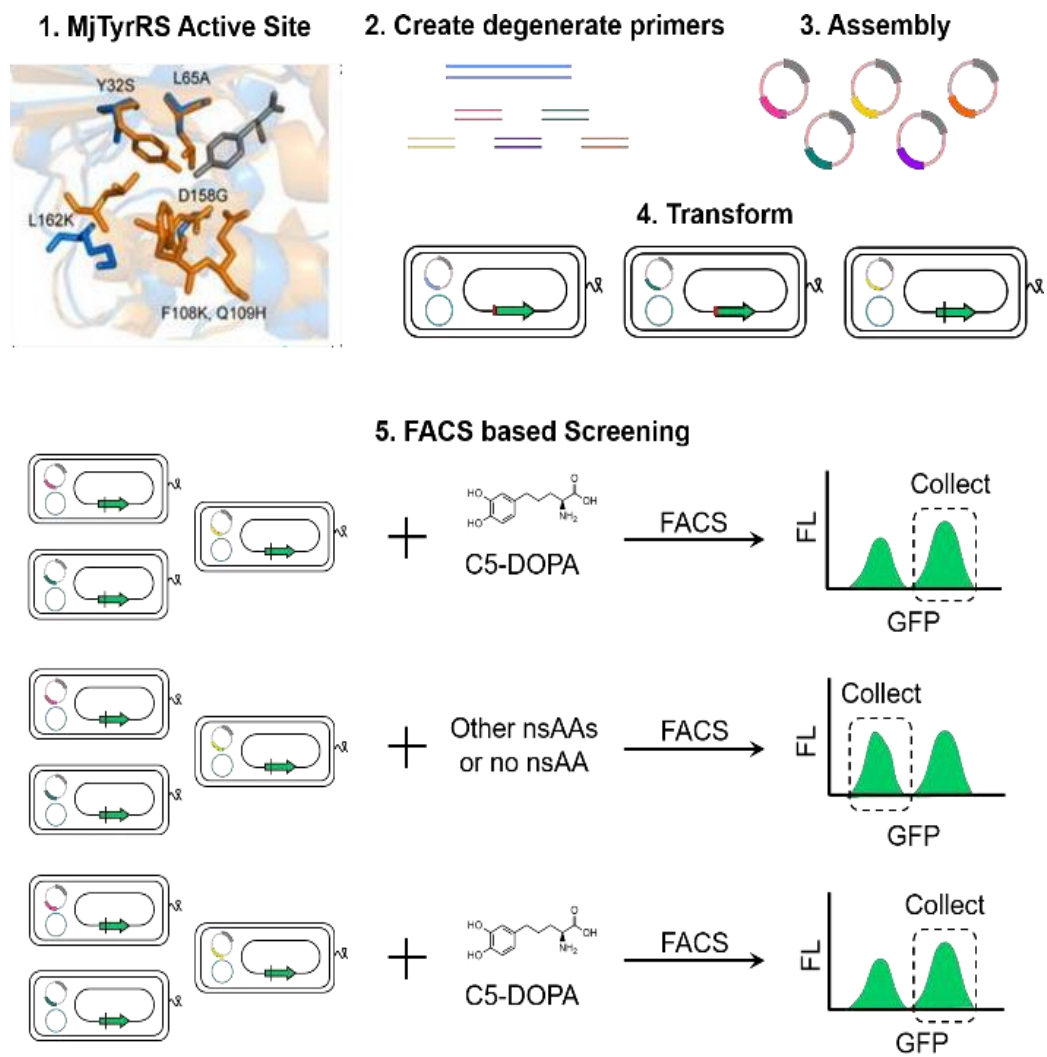
With the success of C5-DOPA incorporation based on fluorescence incorporation assay, a parallel step can be to perform screening of similar extended nsAAs for incorporation. Since these extended side chain catechols have never been tested for site-specific incorporation into proteins, this will provide an opportunity to explore different applications relevant for these nsAAs.

### **5.2.3 Generation and screening of synthetase library for selective C5-DOPA aminoacylation**

To establish multi-site nsAA incorporation in proteins, we believe that the orthogonal aaRS/tRNACUA pairs should exhibit optimal activity and maximum specificity. Based on the results obtained in Chapter 3, a candidate synthetase can be selected for engineering. From the preliminary incorporation data, we have observed that MjTyrRS variants accept C5-DOPA as a substrate. Based on the information on MjTyrRS active site residues (**Figure 15.1**) [64], site-directed mutagenesis at the following positions: 32, 158, 159, and 162 can be performed.

Mutagenesis can be performed by ordering primers that contain degenerate codons at each selected residue followed by Gibson assembly to clone the library in the pEVOL plasmid.

To validate that the engineered orthogonal synthetase show improved specificity for C5-DOPA, single and multi-UAG sfGFP will be expressed in the presence of C5--DOPA under the same culture conditions mentioned in Chapter 3. After the cell lysis and purification of sfGFP, samples will be analyzed for size confirmation with SDS-PAGE and western-blot followed by NBT-staining to detect the presence of catechol in proteins. Samples can be analyzed using ESI-Q-Tof MS to confirm C5-DOPA incorporation at multiple sites in the protein.



**Figure 5.1.** Engineering and selection of candidate synthetase for C5-DOPA incorporation.

#### **5.2.4 Characterization of catechol-boronate chemistry for coupling in fermentation-like conditions**

With the knowledge of small-molecule catechol-boronate chemistry, this conjugation chemistry can be expanded in proteins to different media compositions (PBS, M9 minimal media, bacteria cell lysate) to mimic live cell conditions. As the boronic acid derivatives are prone to interact with glucose present in the cellular environment, we are particularly interested in evaluating the catechol-boronate coupling reaction in presence of varying amounts of glucose. Initially, the equilibrium constants of boronic acid coupling with glucose can be measured in different pH and media conditions such as LB, M9 minimal media and phosphate buffer using the ARS assay. This result can be compared to the equilibrium constants established Chapter 4 for our catechol substrates (DOPA and C5-DOPA). Although we have performed initial protein conjugation experiments, those results need to be optimized and repeated. For optimization of the reaction conditions for macromolecules like protein containing one or multiple nsAAs, parameters like pKa of the boronic acid, solvent type, pH, temperature, time and concentration of the reactants needs to be optimized. Post reaction, the unconjugated reagents can be separated out using a 10 KDa filter and will assess the conjugate formation and yield using HPLC and by ESI-Q-Tof MS analysis.

## REFERENCES

- Alfonta, Lital, Zhiwen Zhang, Sean Uryu, Joseph A. Loo, and Peter G. Schultz. 2003a. "Site-Specific Incorporation of a Redox-Active Amino Acid into Proteins." *Journal of the American Chemical Society* 125 (48): 14662–63. <https://doi.org/10.1021/ja038242x>.
- Alfonta, Lital, Zhiwen Zhang, Sean Uryu, Joseph A Loo, and Peter G Schultz. 2003b. "Site-Specific Incorporation of a Redox-Active Amino Acid into Proteins." *Journal of the American Chemical Society* 125 (48): 14662–63. <https://doi.org/10.1021/ja038242x>.
- Amant, Andre H St., Fengying Huang, Jia Lin, Daniel Lemen, Chacko Chakiath, Shenlan Mao, Christine Fazenbaker, et al. 2019. "A Reactive Antibody Platform for One-Step Production of Antibody–Drug Conjugates through a Diels–Alder Reaction with Maleimide." *Bioconjugate Chemistry* 30 (9): 2340–48. <https://doi.org/10.1021/acs.bioconjchem.9b00436>.
- Anderson, J Christopher, Ning Wu, Stephen W Santoro, Vishva Lakshman, David S King, and Peter G Schultz. 2004. "An Expanded Genetic Code with a Functional Quadruplet Codon." *Proceedings of the National Academy of Sciences of the United States of America* 101 (20): 7566 LP – 7571. <https://doi.org/10.1073/pnas.0401517101>.
- Axup, Jun Y., Krishna M. Bajjuri, Melissa Ritland, Benjamin M. Hutchins, Chan Hyuk

- Kim, Stephanie A. Kazane, Rajkumar Halder, et al. 2012. "Synthesis of Site-Specific Antibody-Drug Conjugates Using Unnatural Amino Acids." *Proceedings of the National Academy of Sciences of the United States of America* 109 (40): 16101–6. <https://doi.org/10.1073/pnas.1211023109>.
- Ayyadurai, Niraikulam, Nadarajan Saravanan Prabhu, Kanagavel Deepankumar, Sun-Gu Lee, Heon-Ho Jeong, Chang-Soo Lee, and Hyungdon Yun. 2011. "Development of a Selective, Sensitive, and Reversible Biosensor by the Genetic Incorporation of a Metal-Binding Site into Green Fluorescent Protein." *Angewandte Chemie International Edition* 50 (29): 6534–37. <https://doi.org/https://doi.org/10.1002/anie.201008289>.
- Bacher, J. M., and A. D. Ellington. 2001. "Selection and Characterization of Escherichia Coli Variants Capable of Growth on an Otherwise Toxic Tryptophan Analogue." *Journal of Bacteriology* 183 (18): 5414–25. <https://doi.org/10.1128/JB.183.18.5414-5425.2001>.
- Borrmann, Annika, and Jan C M van Hest. 2014. "Bioorthogonal Chemistry in Living Organisms." *Chemical Science* 5 (6): 2123–34. <https://doi.org/10.1039/C3SC52768A>.
- Brooks, William L A, Christopher C Deng, and Brent S Sumerlin. 2018. "Structure–Reactivity Relationships in Boronic Acid–Diol Complexation." *ACS Omega* 3 (12): 17863–70. <https://doi.org/10.1021/acsomega.8b02999>.
- Burdine, Lyle, Thomas G Gillette, Hai-Jun Lin, and Thomas Kodadek. 2004. "Periodate-Triggered Cross-Linking of DOPA-Containing Peptide–Protein

- Complexes.” *Journal of the American Chemical Society* 126 (37): 11442–43.  
<https://doi.org/10.1021/ja045982c>.
- Canne, Lynne E, Adrian R Ferre- D’Amare, Stephen K Burley, and Stephen B H Kent. 1995. “Total Chemical Synthesis of a Unique Transcription Factor-Related Protein: CMyc-Max.” *Journal of the American Chemical Society* 117 (11): 2998–3007. <https://doi.org/10.1021/ja00116a005>.
- Capecchi, M R, and H A Klein. 1970. “Release Factors Mediating Termination of Complete Proteins.” *Nature* 226 (5250): 1029–33.  
<https://doi.org/10.1038/2261029a0>.
- Chalker, Justin M., Gonçalo J.L. Bernardes, Yuya A. Lin, and Benjamin G. Davis. 2009. “Chemical Modification of Proteins at Cysteine: Opportunities in Chemistry and Biology.” *Chemistry - An Asian Journal* 4 (5): 630–40.  
<https://doi.org/10.1002/asia.200800427>.
- Chatterjee, Abhishek, Han Xiao, Peng-Yu Yang, Gautam Soundararajan, and Peter G Schultz. 2013. “A Tryptophanyl-TRNA Synthetase/TRNA Pair for Unnatural Amino Acid Mutagenesis in E. Coli.” *Angewandte Chemie International Edition* 52 (19): 5106–9. <https://doi.org/https://doi.org/10.1002/anie.201301094>.
- Chin, Jason W. 2017. “Expanding and Reprogramming the Genetic Code.” *Nature* 550 (7674): 53–60. <https://doi.org/10.1038/nature24031>.
- Ding, Wenlong, Hongxia Zhao, Yulin Chen, Bin Zhang, Yang Yang, Jia Zang, Jing Wu, and Shixian Lin. 2020. “Chimeric Design of Pyrrolysyl-TRNA Synthetase/TRNA Pairs and Canonical Synthetase/TRNA Pairs for Genetic Code Expansion.” *Nature*

- Communications* 11 (1): 3154. <https://doi.org/10.1038/s41467-020-16898-y>.
- Dumas, Anaelle, Lukas Lercher, Christopher D Spicer, and Benjamin G Davis. 2015. “Designing Logical Codon Reassignment - Expanding the Chemistry in Biology.” *Chemical Science* 6 (1): 50–69. <https://doi.org/10.1039/c4sc01534g>.
- Elsherif, Mohamed, Mohammed Umair Hassan, Ali K Yetisen, and Haider Butt. 2018. “Glucose Sensing with Phenylboronic Acid Functionalized Hydrogel-Based Optical Diffusers.” *ACS Nano* 12 (3): 2283–91. <https://doi.org/10.1021/acsnano.7b07082>.
- Elzahhar, Perihan, Ahmed S F Belal, Fatema Elamrawy, Nada A Helal, and Mohamed Ismail Nounou. 2019. “Bioconjugation in Drug Delivery: Practical Perspectives and Future Perceptions BT - Pharmaceutical Nanotechnology: Basic Protocols.” In , edited by Volkmar Weissig and Tamer Elbayoumi, 125–82. New York, NY: Springer New York. [https://doi.org/10.1007/978-1-4939-9516-5\\_11](https://doi.org/10.1007/978-1-4939-9516-5_11).
- Esteves-Villanueva, Jose O, Hanna Trzeciakiewicz, and Sanela Martic. 2014. “A Protein-Based Electrochemical Biosensor for Detection of Tau Protein, a Neurodegenerative Disease Biomarker.” *Analyst* 139 (11): 2823–31. <https://doi.org/10.1039/C4AN00204K>.
- Fekner, Tomasz, and Michael K Chan. 2011. “The Pyrrolysine Translational Machinery as a Genetic-Code Expansion Tool.” *Current Opinion in Chemical Biology* 15 (3): 387–91. <https://doi.org/https://doi.org/10.1016/j.cbpa.2011.03.007>.
- Fordjour, Eric, Frederick Komla Adipah, Shenghu Zhou, Guocheng Du, and Jingwen Zhou. 2019. “Metabolic Engineering of Escherichia Coli BL21 (DE3) for de Novo

- Production of l-DOPA from d-Glucose.” *Microbial Cell Factories* 18 (1): 74.  
<https://doi.org/10.1186/s12934-019-1122-0>.
- Gaertner, Hubert F, Keith Rose, Ron Cotton, David Timms, Roger Camble, and Robin E Offord. 1992. “Construction of Protein Analogs by Site-Specific Condensation of Unprotected Fragments.” *Bioconjugate Chemistry* 3 (3): 262–68.  
<https://doi.org/10.1021/bc00015a010>.
- Gennari, Arianna, Chirag Gujral, Erwin Hohn, Enrique Lallana, Francesco Cellesi, and Nicola Tirelli. 2017. “Revisiting Boronate/Diol Complexation as a Double Stimulus-Responsive Bioconjugation.” *Bioconjugate Chemistry* 28 (5): 1391–1402. <https://doi.org/10.1021/acs.bioconjchem.7b00080>.
- Gong, Yukang, and Lifeng Pan. 2015. “Recent Advances in Bioorthogonal Reactions for Site-Specific Protein Labeling and Engineering.” *Tetrahedron Letters* 56 (17): 2123–32. <https://doi.org/https://doi.org/10.1016/j.tetlet.2015.03.065>.
- Grand View Research. 2019. “Antibody Drug Conjugate Market Size Worth USD 9.93 Billion By 2025.” <https://www.grandviewresearch.com/press-release/global-antibody-drug-conjugates-market>.
- Hermanson, Greg T. 2013. *Bioconjugate Techniques*. San Diego, UNITED STATES: Elsevier Science & Technology. <http://ebookcentral.proquest.com/lib/udel-ebooks/detail.action?docID=1323702>.
- Hohl, Adrian, Ram Karan, Anastassja Akal, Dominik Renn, Xuechao Liu, Seema Ghorpade, Michael Groll, Magnus Rueping, and Jörg Eppinger. 2019. “Engineering a Polyspecific Pyrrolysyl-TRNA Synthetase by a High Throughput

- FACS Screen.” *Scientific Reports* 9 (December). <https://doi.org/10.1038/s41598-019-48357-0>.
- Hughes, Randall A, and Andrew D Ellington. 2010. “Rational Design of an Orthogonal Tryptophanyl Nonsense Suppressor tRNA.” *Nucleic Acids Research* 38 (19): 6813–30.
- Jakob, Leonhard, Alexander Gust, and Dina Grohmann. 2019. “Evaluation and Optimisation of Unnatural Amino Acid Incorporation and Bioorthogonal Bioconjugation for Site-Specific Fluorescent Labelling of Proteins Expressed in Mammalian Cells.” *Biochemistry and Biophysics Reports* 17 (November 2018): 1–9. <https://doi.org/10.1016/j.bbrep.2018.10.011>.
- Jeong, Ye, Byeongseon Yang, Byungseop Yang, Mincheol Shin, Jihyoun Seong, Hyung Joon Cha, and Inchan Kwon. 2020. “Enhanced Production of Dopa-Incorporated Mussel Adhesive Protein Using Engineered Translational Machineries.” *Biotechnology and Bioengineering* 117 (March). <https://doi.org/10.1002/bit.27339>.
- Jia, Lili, Kosuke Minamihata, Hirofumi Ichinose, Kouhei Tsumoto, and Norihiro Kamiya. 2017. “Polymeric SpyCatcher Scaffold Enables Bioconjugation in a Ratio-Controllable Manner.” *Biotechnology Journal* 12 (12): 1700195. <https://doi.org/https://doi.org/10.1002/biot.201700195>.
- Johnson, Jeremiah A, Ying Y Lu, James A Van Deventer, and David A Tirrell. 2010. “Residue-Specific Incorporation of Non-Canonical Amino Acids into Proteins: Recent Developments and Applications.” *Current Opinion in Chemical Biology* 14

(6): 774–80. <https://doi.org/https://doi.org/10.1016/j.cbpa.2010.09.013>.

Kim, Bum Jin, Hogyun Cheong, Byeong Hee Hwang, and Hyung Joon Cha. 2015.

“Mussel-Inspired Protein Nanoparticles Containing Iron(III)–DOPA Complexes for PH-Responsive Drug Delivery.” *Angewandte Chemie International Edition* 54 (25): 7318–22. <https://doi.org/https://doi.org/10.1002/anie.201501748>.

Kim, Sanggil, Bong Hyun Sung, Sun Chang Kim, and Hyun Soo Lee. 2018. “Genetic

Incorporation of L-Dihydroxyphenylalanine (DOPA) Biosynthesized by a Tyrosine Phenol-Lyase.” *Chemical Communications* 54 (24): 3002–5. <https://doi.org/10.1039/C8CC00281A>.

Kunjapur, Aditya M, Devon A Stork, Erkin Kuru, Oscar Vargas-Rodriguez, Matthieu

Landon, Dieter Söll, and George M Church. 2018. “Engineering Posttranslational Proofreading to Discriminate Nonstandard Amino Acids.” *Proceedings of the National Academy of Sciences* 115 (3): 619 LP – 624. <https://doi.org/10.1073/pnas.1715137115>.

Kwon, Inchan, and Byungseop Yang. 2017. “Bioconjugation and Active Site Design of

Enzymes Using Non-Natural Amino Acids.” *Industrial & Engineering Chemistry Research* 56 (23): 6535–47. <https://doi.org/10.1021/acs.iecr.7b00612>.

Lee, Bruce P., P.B. Messersmith, J.N. Israelachvili, and J.H. Waite. 2011. “Mussel-

Inspired Adhesives and Coatings.” *Annual Review of Materials Research* 41 (1): 99–132. <https://doi.org/10.1146/annurev-matsci-062910-100429>.

Lee, Haeshin, Norbert F Scherer, and Phillip B Messersmith. 2006. “Single-Molecule

Mechanics of Mussel Adhesion.” *Proceedings of the National Academy of*

- Sciences* 103 (35): 12999 LP – 13003. <https://doi.org/10.1073/pnas.0605552103>.
- Lieser, Rachel M., Daniel Yur, Millicent O. Sullivan, and Wilfred Chen. 2020. “Site-Specific Bioconjugation Approaches for Enhanced Delivery of Protein Therapeutics and Protein Drug Carriers.” *Bioconjugate Chemistry* 31 (10): 2272–82. <https://doi.org/10.1021/acs.bioconjchem.0c00456>.
- Lim, Sung In, Yukina Mizuta, Akinori Takasu, Yong Hwan Kim, and Inchan Kwon. 2014. “Site-Specific Bioconjugation of a Murine Dihydrofolate Reductase Enzyme by Copper(I)-Catalyzed Azide-Alkyne Cycloaddition with Retained Activity.” *PLoS ONE* 9 (6). <https://doi.org/10.1371/journal.pone.0098403>.
- Lin, Fiona L, Helen M Hoyt, Herman van Halbeek, Robert G Bergman, and Carolyn R Bertozzi. 2005. “Mechanistic Investigation of the Staudinger Ligation.” *Journal of the American Chemical Society* 127 (8): 2686–95. <https://doi.org/10.1021/ja044461m>.
- Liu, Xiaoxia, Miaomiao Tian, Wenmei Gao, and Jinzhong Zhao. 2019. “A Simple, Rapid, Fluorometric Assay for Dopamine by In Situ Reaction of Boronic Acids and *Cis*-Diol.” Edited by Chih-Ching Huang. *Journal of Analytical Methods in Chemistry* 2019: 6540397. <https://doi.org/10.1155/2019/6540397>.
- MA, Paz, Rudolf Fluckiger, Boak A, Herbert Kagan, and Gallop PM. 1991. “Paz MA, Fluckiger R, Boak A, Kagan HM, Gallop PM. Specific Detection of Quinoproteins by Redox-Cycling Staining. *J Biol Chem*. 1991 Jan 15; 266(2):689-92.” *Journal of Biological Chemistry* 266 (February): 689. [https://doi.org/10.1016/S0021-9258\(17\)35225-0](https://doi.org/10.1016/S0021-9258(17)35225-0).

- Mansoor, Shazia, Pierre P D Kondiah, Yahya E Choonara, and Viness Pillay. 2019. "Polymer-Based Nanoparticle Strategies for Insulin Delivery." *Polymers* 11 (9): 1380. <https://doi.org/10.3390/polym11091380>.
- Moral, Mario E G, and Teruna J Siahaan. 2017. "Conjugates of Cell Adhesion Peptides for Therapeutics and Diagnostics Against Cancer and Autoimmune Diseases." *Current Topics in Medicinal Chemistry* 17 (32): 3425–43. <https://doi.org/10.2174/1568026618666180118154514>.
- Muñoz, Ana Joyce, Georgina Hernández-Chávez, Ramon de Anda, Alfredo Martínez, Francisco Bolívar, and Guillermo Gosset. 2011. "Metabolic Engineering of Escherichia Coli for Improving L-3,4-Dihydroxyphenylalanine (l-DOPA) Synthesis from Glucose." *Journal of Industrial Microbiology & Biotechnology* 38 (11): 1845. <https://doi.org/10.1007/s10295-011-0973-0>.
- Nicklisch, Sascha C T, and J Herbert Waite. 2012. "Mini-Review: The Role of Redox in Dopa-Mediated Marine Adhesion." *Biofouling* 28 (8): 865–77. <https://doi.org/10.1080/08927014.2012.719023>.
- Nilsson, Bradley L, Laura L Kiessling, and Ronald T Raines. 2000. "Staudinger Ligation: A Peptide from a Thioester and Azide." *Organic Letters* 2 (13): 1939–41. <https://doi.org/10.1021/ol0060174>.
- Ozawa, Kiyoshi, Madeleine J. Headlam, Dmitri Mouradov, Stephen J. Watt, Jennifer L. Beck, Kenneth J. Rodgers, Roger T. Dean, Thomas Huber, Gottfried Otting, and Nicholas E. Dixon. 2005a. "Translational Incorporation of L-3,4-Dihydroxyphenylalanine into Proteins." *FEBS Journal* 272 (12): 3162–71.

<https://doi.org/10.1111/j.1742-4658.2005.04735.x>.

Ozawa, Kiyoshi, Madeleine J Headlam, Dmitri Mouradov, Stephen J Watt, Jennifer L Beck, Kenneth J Rodgers, Roger T Dean, Thomas Huber, Gottfried Otting, and Nicholas E Dixon. 2005b. “Translational Incorporation of L-3,4-Dihydroxyphenylalanine into Proteins.” *The FEBS Journal* 272 (12): 3162–71. <https://doi.org/https://doi.org/10.1111/j.1742-4658.2005.04735.x>.

Pickens, Chad J, Stephanie N Johnson, Melissa M Pressnall, Martin A Leon, and Cory J Berkland. 2018. “Practical Considerations, Challenges, and Limitations of Bioconjugation via Azide–Alkyne Cycloaddition.” *Bioconjugate Chemistry* 29 (3): 686–701. <https://doi.org/10.1021/acs.bioconjchem.7b00633>.

Pinder, R M. 1970. “Possible Dopamine Derivatives Capable of Crossing the Blood-Brain Barrier in Relation to Parkinsonism.” *Nature* 228 (5269): 358. <https://doi.org/10.1038/228358a0>.

Proft, Thomas. 2010. “Sortase-Mediated Protein Ligation: An Emerging Biotechnology Tool for Protein Modification and Immobilisation.” *Biotechnology Letters* 32 (1): 1–10. <https://doi.org/10.1007/s10529-009-0116-0>.

“Protein Labeling Market Size , Share & Trends Analysis Report by Product (Reagents, Kits, Services), By Application (Immunological Techniques, Fluorescence, Microscopy), By Labeling Method, and Segment Forecasts, 2018-2024.” 2018. *Grand View Research*.

Reddington, Samuel C, and Mark Howarth. 2015. “Secrets of a Covalent Interaction for Biomaterials and Biotechnology: SpyTag and SpyCatcher.” *Current Opinion in*

- Chemical Biology* 29: 94–99.  
<https://doi.org/https://doi.org/10.1016/j.cbpa.2015.10.002>.
- Rodgers, Kenneth J., Peter M. Hume, Rachael A. Dunlop, and Roger T. Dean. 2004. “Biosynthesis and Turnover of DOPA-Containing Proteins by Human Cells.” *Free Radical Biology and Medicine* 37 (11): 1756–64.  
<https://doi.org/10.1016/j.freeradbiomed.2004.08.009>.
- Rodgers, Kenneth J., Hongjie Wang, Shanlin Fu, and Roger T. Dean. 2002. “Biosynthetic Incorporation of Oxidized Amino Acids into Proteins and Their Cellular Proteolysis.” *Free Radical Biology and Medicine* 32 (8): 766–75.  
[https://doi.org/10.1016/S0891-5849\(02\)00768-2](https://doi.org/10.1016/S0891-5849(02)00768-2).
- Rose, Keith. 1994. “Facile Synthesis of Homogeneous Artificial Proteins.” *Journal of the American Chemical Society* 116 (1): 30–33.  
<https://doi.org/10.1021/ja00080a004>.
- Ryan, Chris P, Mark E B Smith, Felix F Schumacher, Dina Grohmann, Danai Papaioannou, Gabriel Waksman, Finn Werner, James R Baker, and Stephen Caddick. 2011. “Tunable Reagents for Multi-Functional Bioconjugation: Reversible or Permanent Chemical Modification of Proteins and Peptides by Control of Maleimide Hydrolysis.” *Chemical Communications* 47 (19): 5452–54.  
<https://doi.org/10.1039/C1CC11114K>.
- Ryu, Youngha, and Peter Schultz. 2006. “Efficient Incorporation of Unnatural Amino Acids into Proteins in Escherichia Coli.” *Nature Methods* 3 (May): 263–65.  
<https://doi.org/10.1038/nmeth864>.

- Santoro, Stephen W, J Christopher Anderson, Vishva Lakshman, and Peter G Schultz. 2003. "An Archaeobacteria-Derived Glutamyl-TRNA Synthetase and TRNA Pair for Unnatural Amino Acid Mutagenesis of Proteins in Escherichia Coli." *Nucleic Acids Research* 31 (23): 6700–6709. <https://doi.org/10.1093/nar/gkg903>.
- Saxon, Eliana, Joshua I Armstrong, and Carolyn R Bertozzi. 2000. "A 'Traceless' Staudinger Ligation for the Chemoselective Synthesis of Amide Bonds." *Organic Letters* 2 (14): 2141–43. <https://doi.org/10.1021/ol006054v>.
- Schmohl, Lena, and Dirk Schwarzer. 2014. "Sortase-Mediated Ligations for the Site-Specific Modification of Proteins." *Current Opinion in Chemical Biology* 22: 122–28. <https://doi.org/https://doi.org/10.1016/j.cbpa.2014.09.020>.
- Silverman, Heather G, and Francisco F Roberto. 2007. "Understanding Marine Mussel Adhesion." *Marine Biotechnology (New York, N.Y.)* 9 (6): 661–81. <https://doi.org/10.1007/s10126-007-9053-x>.
- Singh-Blom, Amrita, Randall A Hughes, and Andrew D Ellington. 2013. "Residue-Specific Incorporation of Unnatural Amino Acids into Proteins In Vitro and In Vivo BT - Enzyme Engineering: Methods and Protocols." In , edited by James C Samuelson, 93–114. Totowa, NJ: Humana Press. [https://doi.org/10.1007/978-1-62703-293-3\\_7](https://doi.org/10.1007/978-1-62703-293-3_7).
- Sletten, Ellen M, and Carolyn R Bertozzi. 2011. "From Mechanism to Mouse: A Tale of Two Bioorthogonal Reactions." *Accounts of Chemical Research* 44 (9): 666–76. <https://doi.org/10.1021/ar200148z>.
- Springsteen, Greg, and Binghe Wang. 2001. "Alizarin Red S. as a General Optical

- Reporter for Studying the Binding of Boronic Acids with Carbohydrates.”  
*Chemical Communications* 1 (17): 1608–9. <https://doi.org/10.1039/b104895n>.
- Stephanopoulos, Nicholas, and Matthew B. Francis. 2011. “Choosing an Effective Protein Bioconjugation Strategy.” *Nature Chemical Biology* 7 (12): 876–84. <https://doi.org/10.1038/nchembio.720>.
- Umeda, Aiko, Gabrielle Nina Thibodeaux, Kathryn Moncivais, Faqin Jiang, and Zhiwen Jonathan Zhang. 2010. “A Versatile Approach to Transform Low-Affinity Peptides into Protein Probes with Cotranslationally Expressed Chemical Cross-Linker.” *Analytical Biochemistry* 405 (1): 82–88. <https://doi.org/10.1016/j.ab.2010.05.026>.
- Umeda, Aiko, Gabrielle Nina Thibodeaux, Jie Zhu, YungAh Lee, and Zhiwen Jonathan Zhang. 2009. “Site-Specific Protein Cross-Linking with Genetically Incorporated 3,4-Dihydroxy-L-Phenylalanine.” *Chembiochem: A European Journal of Chemical Biology* 10 (8): 1302–4. <https://doi.org/10.1002/cbic.200900127>.
- Vanbrunt, Michael P., Kurt Shanebeck, Zachary Caldwell, Jeffrey Johnson, Pamela Thompson, Thomas Martin, Huifang Dong, et al. 2015. “Genetically Encoded Azide Containing Amino Acid in Mammalian Cells Enables Site-Specific Antibody-Drug Conjugates Using Click Cycloaddition Chemistry.” *Bioconjugate Chemistry* 26 (11): 2249–60. <https://doi.org/10.1021/acs.bioconjchem.5b00359>.
- Waite, J HERBERT, and MARVIN L Tanzer. 1981. “Polyphenolic Substance of *Mytilus Edulis*: Novel Adhesive Containing L-Dopa and Hydroxyproline.” *Science* 212 (4498): 1038 LP – 1040. <https://doi.org/10.1126/science.212.4498.1038>.

- Wan, Wei, Jeffery M Tharp, and Wenshe R Liu. 2014. "Pyrrolysyl-TRNA Synthetase: An Ordinary Enzyme but an Outstanding Genetic Code Expansion Tool." *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* 1844 (6): 1059–70. <https://doi.org/https://doi.org/10.1016/j.bbapap.2014.03.002>.
- Wang, Lei, Jianming Xie, and Peter G Schultz. 2006. "Expanding The Genetic Code." *Annual Review of Biophysics and Biomolecular Structure* 35 (1): 225–49. <https://doi.org/10.1146/annurev.biophys.35.101105.121507>.
- Xu, Jianqing, Drew Tack, Randall A Hughes, Andrew D Ellington, and Jeffrey J Gray. 2014. "Structure-Based Non-Canonical Amino Acid Design to Covalently Crosslink an Antibody–Antigen Complex." *Journal of Structural Biology* 185 (2): 215–22. <https://doi.org/http://dx.doi.org/10.1016/j.jsb.2013.05.003>.
- Yan, Jun, Greg Springsteen, Susan Deeter, and Binghe Wang. 2004. "The Relationship among PKa, PH, and Binding Constants in the Interactions between Boronic Acids and Diols—It Is Not as Simple as It Appears." *Tetrahedron* 60 (49): 11205–9. <https://doi.org/https://doi.org/10.1016/j.tet.2004.08.051>.
- Zhao, Ning, Kouta Kamijo, Philip D. Fox, Haruka Oda, Tatsuya Morisaki, Yuko Sato, Hiroshi Kimura, and Timothy J. Stasevich. 2019. "A Genetically Encoded Probe for Imaging Nascent and Mature HA-Tagged Proteins in Vivo." *Nature Communications* 10 (1). <https://doi.org/10.1038/s41467-019-10846-1>.