The Development Of Intrinsically Fluorescent Unnatural Amino Acids For In Vivo Incorporation Into Proteins

Itthipol Sungwienwong

University of Pennsylvania, itthipol@sas.upenn.edu
The Development Of Intrinsically Fluorescent Unnatural Amino Acids For In Vivo Incorporation Into Proteins

Abstract

The amino acid acridone-2-ylalanine (Acd) can be a valuable probe of protein dynamics either alone or as part of a Förster resonance energy transfer (FRET) or photo-induced electron transfer (eT) probe pair. We have previously reported the genetic incorporation of Acd by an aminoacyl tRNA synthetase (RS). However, this RS, developed from a library of permissive RSs, also incorporates N-phenyl-amino-phenylalanine (Npf), a trace byproduct of one Acd synthetic route. We have performed negative selections in the presence of Npf and analyzed the selectivity of the resulting AcdRSs by in vivo protein expression and detailed kinetic analyses of the purified RSs. We find that selection conferred a ~50-fold increase in selectivity for Acd over Npf, eliminating incorporation of Npf contaminants, and allowing one to use a high yielding Acd synthetic route for improved overall expression of Acd-containing proteins. More generally, our report also provides a cautionary tale on the use of permissive RSs, as well as a strategy for improving selectivity for the target amino acid.

In spite of its utility for studying proteins by fluorescence spectroscopy, Acd can potentially be improved by making it longer wavelength or brighter. We reported the synthesis of Acd core derivatives and their photophysical characterization. We also performed ab initio calculations of the absorption and emission spectra of Acd derivatives, which agree well with experimental measurements. The amino acid aminoacridonylalanine (Aad) was synthesized in forms appropriate for genetic incorporation and peptide synthesis. We show that Aad is a superior FRET acceptor to Acd in a peptide cleavage assay, and that Aad can be activated by an aminoacyl tRNA synthetase for genetic incorporation. Together, these results show that we can use computation to design enhanced Acd derivatives which can be used in peptides and proteins.

Finally, the Aad synthesis has been improved and it will be further tested in vivo incorporations into proteins, and alkylated Aad core analogs show improved brightness making their use as amino acids promising.

Degree Type
Dissertation

Degree Name
Doctor of Philosophy (PhD)

Graduate Group
Chemistry

First Advisor
Ernest J. Petersson

Subject Categories
Chemistry

This dissertation is available at ScholarlyCommons: https://repository.upenn.edu/edissertations/3037
THE DEVELOPMENT OF INTRINSICALLY FLUORESCENT UNNATURAL AMINO ACIDS
FOR IN VIVO INCORPORATION INTO PROTEINS

Itthipol Sungwienwong

A DISSERTATION

in

Chemistry

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2018

Supervisor of Dissertation

________________________
E. James Petersson

Associate Professor of Chemistry

Graduate Group Chairperson

________________________
David W. Christianson, Roy and Diana Vagelos Professor of Chemistry and Chemical Biology

Dissertation Committee:

David M. Chenoweth, Associate Professor of Chemistry

Tobias Baumgart, Professor of Chemistry

Ivan J. Dmochowski, Professor of Chemistry
THE DEVELOPMENT OF INTRINSICALLY FLUORESCENT UNNATURAL AMINO ACIDS FOR IN VIVO INCORPORATION INTO PROTEINS

COPYRIGHT

2018

Ittipol Sungwienwong
ACKNOWLEDGEMENTS

First and foremost, I would like to thank my advisor, Professor E. James Petersson for providing me with an opportunity to join the laboratory and for his generous guidance throughout my graduate career. He has supported me not only in scientific research but also in general life. I am deeply appreciative of his mentorship. Without him, I would not be able to achieve success in my career. I would also like to thank my dissertation committee members; Professor David M. Chenoweth, Professor Tobias Baumgart, and Professor Ivan Dmochowski for their kind support, helpful feedback, and their encouragement throughout these years.

I would like to acknowledge all members of the Petersson group and alumni, who have been great colleagues and friends. It was my pleasure to work with these people; Dr. Lee Speight, Dr. Rebecca Wissner, Dr. Solongo Ziraldo, Dr. Yanxin Wang, Dr. Xing Chen, Dr. Chris Walters, Dr. Yun Huang, Dr. Conor Haney, Dr. Naoya Ieda, Miklos Szantai-Kis, John Ferrie, Anand Muthusamy, Taylor Barrett, Joomyung Jun, Chunxiao Liu, Buyan Pan, Christina Cleveland, E. Keith Keenan, Tiberiu Mihaila, Jimin Yoon. Furthermore, I would like to thank my collaborators for their helpful support; Professor Rahul Kohli and Zachary Hostetler (Department of Medicine, Perelman School of Medicine, University of Pennsylvania), Professor Ryan Mehl (Oregon State University), Professor John Perona (Portland State University), and Professor Jeffery Saven (Penn Chemistry).

Finally, I would like to acknowledge my family and friends in Thailand. I am quite thankful for my parents for allowing me to study abroad and inspiring me to pursue a PhD
degree. I would like to thank the Thai community at the University of Pennsylvania and Drexel University for mental support and invaluable friendship. I would like to make special thanks to Nisalak Trongsiriwat, Parawee Kasitinon, and Dr.Weerapha Panaddasirisuk for their priceless companionship. And last, I need to thank my scholarship support from the Royal Thai government and my future workplace (Srinakharinwirot University, Bangkok, Thailand) that provided me with the great opportunity to accomplish this doctoral degree.
ABSTRACT

THE DEVELOPMENT OF INTRINSICALLY FLUORESCENT UNNATURAL AMINO ACIDS
FOR IN VIVO INCORPORATION INTO PROTEINS

Itthipol Sungwienwong
Professor E. James Petersson

The amino acid acridon-2-ylalanine (Acd) can be a valuable probe of protein
dynamics either alone or as part of a Förster resonance energy transfer (FRET) or photo-
induced electron transfer (eT) probe pair. We have previously reported the genetic
incorporation of Acd by an aminoacyl tRNA synthetase (RS). However, this RS, developed
from a library of permissive RSs, also incorporates N-phenyl-amino-phenylalanine (Npf),
a trace byproduct of one Acd synthetic route. We have performed negative selections in
the presence of Npf and analyzed the selectivity of the resulting AcdRSs by in vivo protein
expression and detailed kinetic analyses of the purified RSs. We find that selection
conferred a ~50-fold increase in selectivity for Acd over Npf, eliminating incorporation
of Npf contaminants, and allowing one to use a high yielding Acd synthetic route for
improved overall expression of Acd-containing proteins. More generally, our report also
provides a cautionary tale on the use of permissive RSs, as well as a strategy for improving
selectivity for the target amino acid.

In spite of its utility for studying proteins by fluorescence spectroscopy, Acd can
potentially be improved by making it longer wavelength or brighter. We reported the
synthesis of Acd core derivatives and their photophysical characterization. We also
performed *ab initio* calculations of the absorption and emission spectra of Acd derivatives, which agree well with experimental measurements. The amino acid aminoacridonylalanine (Aad) was synthesized in forms appropriate for genetic incorporation and peptide synthesis. We show that Aad is a superior FRET acceptor to Acd in a peptide cleavage assay, and that Aad can be activated by an aminoacyl tRNA synthetase for genetic incorporation. Together, these results show that we can use computation to design enhanced Acd derivatives which can be used in peptides and proteins.

Finally, the Aad synthesis has been improved and it will be further tested *in vivo* incorporations into proteins, and alkylated Aad core analogs show improved brightness making their use as amino acids promising.
TABLE OF CONTENTS

ACKNOWLEDGEMENT........................................................................................................III

ABSTRACT ............................................................................................................................... V

LIST OF TABLES AND SCHEMES....................................................................................... IV

LIST OF ILLUSTRATIONS .................................................................................................... X

CHAPTER 1 INTRODUCTION ..................................................................................................1

1.1 Background .................................................................................................................... 2
1.2 Expanded Genetic Code ............................................................................................. 5
1.3 Analysis of Aminoacyl tRNA Synthetases (RSs) ...................................................... 9
1.4 Genetic Incorporation of Fluorescent Amino Acids .................................................. 12
1.5 Fluorescent Labeling Using Unnatural Amino Acids with Bioorthogonal Reactivity ... 19
1.6 Acriflavone as a Tunable Fluorescent Scaffold for Generating Fluorescent Amino Acids ............................................................................................................................. 22

CHAPTER 2 THE OPTIMIZATION OF A PERMISSIVE AMINOACYL tRNA SYNTHETASE FOR A TARGET UNNATURAL AMINO ACID................................................................. 24

2.1 Background .................................................................................................................... 25
2.2 Synthetase Selection .................................................................................................. 29
2.3 AcdRS Optimization ................................................................................................ 32
2.4 Comparison of AcdRS Active Sites .......................................................................... 33
LIST OF TABLES

Table 2.1 AcdRS *in vitro* enzymology and *in vivo* MS selectivity

Table 2.2 Sequence of top performing AcdRSs.

Table 2.3 CaM 113 trypsin digest intensity scaling

Table 3.1 Calculated and observed photophysical parameters of acridone derivatives

Table 3.2 Screening of acridone nitration conditions

Table 3.3 Calculated and observed photophysical parameters of acridone derivatives.

LIST OF SCHEMES

Scheme 2.1 The synthesis of acridon-2-ylalanine (Acd).

Scheme 3.1 Synthesis of nitroacridones and aminoacridones

Scheme 3.2 Synthesis of acridone cores and benzoacridones

Scheme 4.1 The alternative synthesis of 2-aminoacridone (Aad)
LIST OF ILLUSTRATIONS

Figure 1.1 Chromophore size

Figure 1.2 Site-specific incorporation of unnatural amino acids

Figure 1.3 Positive and negative selection strategies

Figure 1.4 The active site of TyrRS and a mutant UaaRS evolved to incorporate o-Iodotyrosine

Figure 1.5 Structure of 7-hydroxycoumarin and three coumarin lysine analogues

Figure 1.6 Structure of NAP, Anap, danLys, dansylalanine, and terphenyl amino acids.

Figure 1.7 Labeling proteins via incorporation of unnatural amino acids

Figure 1.8 Effects of flexible linkers on the range of chromophore distances in a FRET experiment.

Figure 1.9 The structure of acridonylalanine (Acd) and 2-aminoacridonylalanine (Aad)

Figure 2.1 AcdRS selection

Figure 2.2 In vivo AcdRS selectivity

Figure 2.3 CaM\textsubscript{113} trypsin digest data for AcdRS selectivity analysis

Figure 2.4 tRNA aminoacylation kinetics

Figure 2.5 AcdRS homology models

Figure 2.6 Fluorescence measurements of RSs with GFP reporter

Figure 2.7 In vivo AcdRS selectivity of CaM

Figure 2.8 In vivo AcdRS selectivity of αS

Figure 2.9 In vivo AcdRS selectivity of LexA

Figure 2.10 AcdRS purification SDS-PAGE gels

Figure 2.11 AcdRS malachite green activity assay
Figure 2.12 Representative primary data for aminoacylation

Figure 3.1 FRET experiments

Figure 3.2 Fluorescent amino acids based on 7-methoxycoumarin and acridone cores

Figure 3.3 Absorption and emission spectra of acridone derivatives

Figure 3.4 Calculated acridone spectra

Figure 3.5 Aad is a superior FRET acceptor for Mcm

Figure 3.6 Aad Activation by AcdRS1

Figure 3.7 Screening of acridone nitration conditions

Figure 3.8 Screening of 2- and 4-nitroacridone reduction conditions

Figure 3.9 $^1$H and $^{13}$C NMR spectra of Boc$_2$-Aad in MeOD

Figure 3.10 Acridone (5) spectra

Figure 3.11 2-Aminoacridone (8) spectra

Figure 3.12 4-Aminoacridone (9) spectra

Figure 3.13 2-Nitroacridone (6) and 4-Nitroacridone (7) spectra

Figure 3.14 2-Fluoroacridone (15), 2-methoxyacridone (S1), and 4-methoxyacridone (16) spectra

Figure 3.15 4-Benzoaacidine (18), 4-fluorobenzoacidine (19), and 4-methoxybenzoacidine (20) spectra

Figure 4.1 Cross-coupling screening results

Figure 4.2 Aad activation by AcdRS2b (A9)

Figure 4.3 Structure of 2-(dimethylamino)acridin-9(10H)-one (Dad) and 2-(azetidin-1-yl)acridin-9(10H)-one (AzAcD)
Figure 4.4 Absorption (top) and emission (bottom) spectrum of Aad, Dad, and AzA
CHAPTER 1 : Introduction
Background

Anfinsen’s hypothesis is a classical postulate in molecular biology stating that the three-dimensional structure of a native protein in its standard physiological environment (solvent, pH, ionic strength, presence of other components such as metal ions or prosthetic groups, temperature, and other factors) is determined only by the totality of inter-atomic interactions within the protein and hence by the amino acid sequence.\(^1\) While it has been clearly shown that many proteins violate this hypothesis – for example, by requiring chaperones for folding\(^2\)\(^-\)\(^4\) – nonetheless, studies of protein structure should imitate the native state in a physiologically-relevant local environment as much as possible. Several methods have been advanced to study protein architecture such as X-ray crystallography\(^5\), nuclear magnetic resonance (NMR) spectroscopy\(^6\), and cryo-electron microscopy (cryo-EM)\(^7\). Structures revealed through these methods can solve biological questions and provide insight into protein interactions, folding pathways, and enzymatic activities. However, many dynamic mechanisms remain obscure due to the limitations of these methods. The X-ray crystallography technique intrinsically requires the protein to be packed into a fine crystal lattice which freezes out most dynamics. Cryo-EM provides greater opportunity for flexibility, but still requires freezing the protein and deposition on an imaging grid. NMR spectroscopy permits solution phase characterization and the opportunity to observe dynamics, but NMR is limited to small proteins and cannot be used for real-time observations of protein motions. Thus, there is a need to complement these high resolution, relatively static structural methods with techniques that can provide information on protein motions in physiologically-relevant situations.
Fluorescence spectroscopy is one valuable technique that can be used to analyze changes in protein structure and protein/protein interactions\textsuperscript{8-9}. Fluorescent experiments can be performed with either intrinsically fluorescent proteins or fluorescently labelled proteins. Protein conformational changes can be monitored by observation of changes in the fluorescence of a single probe or through interactions of multiple (usually pairs) of probes. A change in fluorescence of a single probe can result from a change in the local environment of the fluorophore due to a conformational change or binding event. Fluorescence polarization (FP) varies with the reorientation time of the chromophore and can report on changes in the flexibility of a protein region or, more commonly, on a change in tumbling rate due to interaction with another macromolecule. Experiments using pairs of probes can be used to track protein structural change on the nanosecond timescale using distance-dependent chromophore interactions either through Förster resonance energy transfer (FRET) or quenching by photo-induced electron transfer (PeT). When using a fluorescent probe, one needs to be concerned with its ability to disrupt protein structure and function. These disruptions can result from the size and/or location of the fluorophores. The size of commonly used fluorescent labels ranges from 29820 Å\textsuperscript{3} for the green fluorescent protein (GFP) to 9 Å\textsuperscript{3} for thioamide substitution\textsuperscript{8}. (Figure 1.1) There are a variety of methods for introducing probes into proteins, which vary in complexity from simple cloning of the cDNA for GFP into a protein expression plasmid to the chemical synthesis of a protein in order to install a synthetic fluorophore. Among these methods, genetic code expansion offers a near ideal combination of small probe size, potential flexibility in probe properties, and ease of probe introduction.
Figure 1.1 Chromophore size. Space-filling renderings of several chromophores, based on either crystal structures or ab initio calculations. The green fluorescent protein (GFP) was rendered from PDB ID 1GFL9. The benzylated SNAP tag protein was rendered from PDB ID 3L00. The geometry of tetracysteine-bound FlAsH was optimized at the HF/6-31G level, all others were optimized at the AM1 level (Figure adapted from Speight L.C., Moumita S., Petersson, E. J. (2014) Minimalist Approaches to Protein Labelling: Getting the Most Fluorescent Bang for Your Steric Buck. Australian Journal of Chemistry, 67, 686-700.)
Expanded Genetic Code

**Figure 1.2** Site-specific incorporation of unnatural amino acids.

Protein translation uses transfer RNAs (tRNAs), which are aminoacylated with their canonical amino acids by aminoacyl-tRNA synthetase enzymes (RSs), to read triplet codons in messenger RNAs (mRNAs) when paired with the anticodon of the tRNA. The ribosome facilitates both the decoding of triplet codons on mRNAs by tRNAs, and the polymerization of the corresponding amino acids in order for the mRNA to be translated into a polypeptide chain. All proteins are ribosomally synthesized with the limitation of 20 canonical amino acids building blocks. While a wide variety of post-translational modifications occur on proteins, these are generally restricted in sequence context so that using them to modify a protein requires the introduction of a tagging sequence which may alter the protein fold. Over the past three decades, approaches have been developed to
expand the genetic code, enabling the site-specific incorporation of unnatural amino acids (Uaas, also known as non-canonical amino acids or ncAAs) into proteins synthesized in both prokaryotic and eukaryotic cells (Figure 1.2).

The site-specific incorporation of a Uaa into proteins has provided new insights to study biological processes which are difficult and challenging to address by classical approaches. Peter Schultz developed a genetic code expansion approach that uses an orthogonal RS/tRNA pair to direct the incorporation of a Uaa into proteins in response to an unassigned codon (commonly the so-called amber stop codon, UAG) introduced at the desired site in a gene of interest, which is called amber suppression. Researchers can now incorporate Uaas into proteins in bacteria (typically Escherichia coli or E. coli), Saccharomyces cerevisiae, mammalian cells (a few experiments in mice have even been reported), Caenorhabditis elegans, Drosophila melanogaster. To be “orthogonal”, the RS/tRNA pair must meet many requirements. The specificity of the RS active site needs to be optimized so that it can transfer the desired Uaa, but not any of the canonical amino acids to the orthogonal tRNA. The orthogonal tRNA must be recognized only by the URS, and none of the endogenous RSs for the natural amino acids. Finally, the URS must not recognize any of the endogenous tRNAs. To meet these complex requirements, Schultz developed a strategy of importing an RS/tRNA from another organism where the tRNA structure was known to be substantially different from that of the host organism. Initially, most amber suppression experiments used orthogonal RS/tRNA pairs based on TyrRS and LeuRS evolved from Methanocaldococcus jannaschii (e.g. MfTyrRS, MjLeuRS) or E. coli (e.g. EcLeuRS). These pairs require active site evolution to facilitate binding with the
desired Uaa of interest. Recently the pyrrolysyl (Pyl) RS/tRNA pairs (from *Methanosarcina barkeri* and *Methanosarcina mazei*) have grown in popularity as Pyl is found in archaea, but not in bacterial or mammalian cells.\textsuperscript{13-16} Thus, PylRS/tRNA pairs are orthogonal in both cell types and only require evolution of the active site for the desired Uaa. The methods for directed evolution of the active site have been well-developed and are usually performed in *E. coli* or yeast cells.

**Figure 1.3** Positive and negative selection strategies for the generation of an orthogonal synthetase/tRNA pair with altered amino acid specificity for an unnatural amino acid (blue star).

Typically, several rounds of positive and negative selection are used to select an active site variant that allows the incorporation of the Uaa but none of the 20 canonical amino acids (**Figure 1.3**).\textsuperscript{17} The positive selection step relies on the ability of the RS/tRNA pair to facilitate read-through of an amber codon in an essential gene (such as antibiotic
This positive selection provides RSs in surviving colonies that can incorporate the Uaa and/or a natural amino acid. These RSs are then subjected to a negative selection procedure performed in the absence of the Uaa with the amber codon now encoded in toxic a gene (such as a toxic ribonuclease or barnase). The RSs that use the canonical amino acid will read through the amber codon and lead to cell death. After several rounds of positive and negative selection, only mutant RSs that can selectively incorporate the Uaa will survive. This method has been used to generate RSs for more than 100 Uaas, which have been widely employed in the production of both recombinant proteins for in vitro experiments and in live cell experiments. However, some limitations on protein yield exist.

The efficiency of protein expression with Uaa incorporation can be improved by removing release factor 1 (RF1 for E. coli and eRF1 for eukaryotes). The decoding of the UAG amber codon normally competes with RF1, which leads to termination of protein synthesis and decreases protein yield. This issue has been dealt with in several ways. In one case, the Chin laboratory’s evolution of an orthogonal ribosome that no longer interacts with RF1 enables the efficient incorporation of a Uaa in response to amber codons at multiple sites. In a second example, it has been shown in some bacterial strains (such as E. coli C321) that the deletion of RF1 can improve the performance of Uaa incorporation in response to the quadruplet codon UAGA. This quadruplet system has been evolved from Methanocaldococcus jannaschii-derived frame-shift suppressor tRNA/RS pair that can enhance UAGN (N = any nucleotide) suppression efficiency. In eukaryotes, the release factor eRF1, which usually terminates the translation on all three stop codons, has been
engineered to provide a significant increase in Uaa incorporation in response to the UAG codon without increasing read-through of other stop codons. With the combination of this engineered eRF1 and an optimized PylRS/tRNA expression system for mammalian cells, one can increased the yield of protein containing a single site of Uaa substitution 17- to 20-fold and the proteins yield for multiple site substitutions are comparable to a protein yields from a gene that does not contain a UAG stop codon. With these and other methods to improve protein yields, Uaa mutagenesis has become a robust technique that is being applied by an ever-growing number of laboratories.

**Analysis of Aminoacyl tRNA Synthetases (RSs)**

Since most UaaRSs are achieved through directed evolution as described above, there are a limited number of cases where detailed analysis of the structure and function of the resulting UaaRS have been performed. RSs catalyze a two-step reaction that provides aminoacylated tRNA for protein synthesis. The first step (1) involves the condensation of amino acid (Aa, or in this case, the Uaa) and ATP to form the enzyme-bound aminoacyl adenylate intermediate (Aa-AMP or Uaa-AMP) and release pyrophosphate (PP\(_i\)). The second step (2) is the reaction between this adenylated intermediate and the 2' or 3'-terminal hydroxyl (OH) group on tRNA to form aminoacyl-tRNA (Aa-tRNA or Uaa-tRNA).

\[
\text{Aa} + \text{ATP} \rightarrow \text{Aa-AMP} + \text{PP}_i \quad (1)
\]
\[
\text{Aa-AMP} + \text{tRNA} \rightarrow \text{Aa-tRNA} + \text{AMP} \quad (2)
\]
The ability of RSs to differentiate an amino acid substrate from other amino acids is almost solely responsible for the accuracy of protein synthesis. Indeed, it is the lack of ribosomal proof-reading of Uaa-tRNA that permits Uaa mutagenesis to work. Thus, an analysis of the initial Uaa-AMP formation step can provide significant insight into UaaRS incorporation. In general, the analysis of RS function has mainly relied on steady state kinetic pyrophosphate exchange and aminoacylation assays. In the pyrophosphate exchange reaction, the RS is added to the reaction buffer containing amino acid, Mg•ATP, and \[^{32}\text{P}\]-pyrophosphate in prepared reaction buffer. The \[^{32}\text{P}\]-ATP formed is then adsorbed on activated charcoal or thin layer chromatography support. The standard pyrophosphate exchange assay is typically performed at 37 °C in a reaction containing \[^{32}\text{P}\] Na•PP\(_i\), RS, amino acid in concentrations ranging from 0.2 to 10 times the Michaelis constant (\(K_m\)), and ATP. The PP\(_i\) exchange reaction is usually performed in the absence of tRNA. Therefore, the \(k_{cat}\) and \(K_m\) Michaelis-Menten values derived from this method are not necessarily identical to the actual rate of amino acid activation in the presence of tRNA. The steady-state tRNA aminoacylation reaction is traditionally investigated by monitoring the formation of \(^{3}\text{H}\) or \(^{14}\text{C}\) aa-tRNA over time. This steady-state assay is usually run at 37 °C in buffer containing RS, amino acid, ATP, and tRNA. The limitation of the steady state aminoacylation assay is that amino acid concentrations are practically limited to concentrations of about 1 mM, due to the relatively low specific activity of \(^3\text{H}\)- or \(^{14}\text{C}\)-labeled amino acids. Alternatively, the \[^{32}\text{P}\]-labeled tRNA can be used together with unlabeled amino acid.
Previous kinetic activities of RS charging of a few Uaas have been studied. For example, the Schultz group has analyzed the kinetics of adenylate formation of O-methyl tyrosine and tyrosine with ATP catalyzed by mutant TyrRSs using the pyrophosphate exchange assay. The $K_m$ value for tyrosine ($5833 \pm 902 \, \mu M$) is approximately 13-fold higher than that for O-methyl-L-tyrosine ($443 \pm 93 \, \mu M$). The $k_{cat}$ for tyrosine ($1.8 \pm 0.2 \times 10^{-3} \, s^{-1}$) is eight-fold less than that for O-methyl-L-tyrosine ($14 \pm 1 \times 10^{-3} \, s^{-1}$).

**Figure 1.4** The active site of TyrRS and a mutant UaaRS evolved to incorporate o-iodotyrosine. The amino acid-binding site of the wild-type TyrRS L-tyrosine (Top left), 37V195C TyrRS 3-iodo-L-tyrosine (Top right), and 37V195C TyrRS L-tyrosine complexes (Bottom right). Bottom left: The structure of o-iodotyrosine (Oiy).
To alter the amino acid specificity of the orthogonal TyrRS so that it can charge the mutRNA$^{\text{Tyr}_{\text{CUA}}}$ with a desired unnatural amino acid, a library of TyrRS mutants was generated and screened. Based on the crystal structure of the homologous TyrRS (Figure 1.4) from *Bacillus stearothermophilus*, five residues in the active site of *M. jannaschii* TyrRS that are within 6.5 Å of the para position of the aryl ring of bound tyrosine were mutated. Crystal structures of the corresponding evolved synthetases can be used to rationalize how specificity was changed.

Kamtekar *et al.*$^{28}$ reported the 3.2-Å resolution crystal structure of the *Methanococcus maripaludis* phosphoseryl-tRNA synthetases (SepRSs). The kinetic constants of these *M. Maripaludis* SepRS mutants for phosphoserine charging activation has measured by ATP-PP$_i$ exchange assay. Umehara *et al.* reported the directed evolution of PylRS from *Methanosarcina mazei* to generate N-acetyl lysyl-tRNA synthetases (AcKRSs). The kinetic activity of these AcKRSs has also been determined by ATP-PP$_i$ exchange assays.$^{29}$ Limited structural information is available for PylRSs.$^{30}$

**Genetic Incorporation of Fluorescent Amino Acids**

Genetically encoded fluorescent Uaas have been used to label proteins in both prokaryotes and eukaryotes, but the number of fluorescent Uaas developed is surprisingly small, given the notable advantages of Uaa labeling: it provides perfect site-specificity and minimal perturbation of protein structure and function. Schultz and coworkers reported the generation of an orthogonal RS/tRNA pair that selectively incorporates 7-hydroxycoumarin amino acid (Hco) into proteins in *E.coli*.$^{31}$ Hco has interesting properties
including a high fluorescence quantum yield (0.63), a large Stoke’s shift, small size, and sensitivity to pH and solvent polarity. The incorporation of Hco into proteins was carried out by using an evolved MjTyrRS/tRNA pair. Protein expression was shown in sperm whale myoglobin and protein mutants were used as a probe to study urea-induced chemical denaturation of holomyoglobin. The fluorescence intensity of the hydroxycoumarin moiety is sensitive to solvent polarity, and it showed correlation between protein unfolding and an increase in fluorescence of Hco. Hco has also been site-specifically incorporated into signal transducer and activator of transcription 3 (STAT3) in order to use the labeled protein as a fluorescent reporter of the phosphorylation status of STAT3. Hco was genetically incorporated into the STAT3β isoform in E. coli at Trp residue 564. A large fluorescence change was observed when the STAT3 probe was phosphorylated by Src kinase in vitro and when it was incubated with endogenously activated STAT3. This method enables the investigation of protein phosphorylation on other STAT protein substrates and other SH2 domain-containing proteins because Trp564 is conserved in all seven mammalian STAT proteins. It also illustrates a general strategy for using fluorescent Uaas to make sensor proteins.

Lou et al. reported the site-specific incorporation of three new coumarin lysine derivatives (Figure 1.5) into proteins in bacterial and mammalian cells using an engineered PyI RS. The genetically encoded coumarin lysines were utilized as optochemical probes for protein localization and light activation of protein function in live cells.
These three new coumarin lysine derivatives showed unique photochemical activities. Bromine atom substitution at the 6-position of 3 enables decaging not only with UV, but also near IR excitation using two photon methods. The extension of the linker chain in 4 can prohibit photolysis. Therefore, coumarin lysines 2 and 3 can be used as both fluorescent and light-activated probes for optochemical control of protein function using UV or near-IR light, while coumarin lysine 4 serves as a stable fluorescent probe that does not decage under UV excitation. To demonstrate the application of these coumarin lysines for optical control, photoregulation of firefly luciferase was achieved in live cells by caging a lysine residue, and excellent turn-on ratios were observed for 2 and 3. As expected, the stable coumarin amino acid 4 did not undergo photolysis. Furthermore, two-photon and single-photon excitation of EGFP tagged with 4 was also demonstrated by using different excitation wavelengths (365, 405, and 760 nm) for the sequential activation of protein function in live cells. While caged lysine 3 could be activated using two-photon irradiation at 760 nm, lysine 2 was stable under these conditions but could be decaged with blue light of 405 nm. This example illustrates how the long lysine linker allows one to attach

Figure 1.5 Structure of 7-hydroxycoumarin (1, Hco) and three coumarin lysine analogues (2-4)
substrates of varying bulk and yet remain compatible with PylRS charging for protein incorporation.

Schultz reported the site-specific incorporation of a second fluorophore, *L*-3-(2-naphthyl)alanine (5, NAP), shown in Figure 1.6, into proteins in *E. coli* using *M. jannaschii* TyrRS mutants. A mutant of mouse dihydrofolate reductase containing NAP at Tyr163 was generated and characterized to confirm the ability of the NAPRS/tRNA pair to site-specifically incorporate NAP in response to an amber stop codon.

Lee *et al.* reported the incorporation of 3-(6-acetylnaphthalen-2-ylamino)-2-aminopropanoic acid (6, Anap) into proteins using an orthogonal RS/tRNA pair derived from the *E. coli* LeuRS and engineered in a two-step process for the bulky Uaa (Figure 1.6). Anap has absorption and emission maxima at 360 and 490 nm, respectively. Its extinction coefficient is 17,500 cm⁻¹ M⁻¹ and its quantum yield is 0.48 in EtOH (with excitation at 360 nm). The Anap amino acid was site-specifically incorporated in *E. coli* glutamine-binding protein and it was used to probe ligand-induced local conformational changes in proteins through solvatochromic effects without the need for a FRET pair. The small size of Anap and ability to introduce it by simple mutagenesis at defined sites should make it a useful probe of protein structure, molecular interactions, protein folding, and localization.
Chatterjee and Schultz also demonstrated the incorporation of Anap in mammalian cells in response to the TAG codon with high efficiency using AnapRS for expression of enhanced green fluorescent protein (EGFP) with an Anap mutation at a permissive site (Tyr40TAG) and a C-terminal histidine tag into Human Embryonic Kidney 293 (HEK293) or Chinese Hamster Ovary (CHO) cells. They further demonstrated that Anap can be used to image the subcellular localization of proteins in live mammalian cells by using confocal microscopy. Anap fluorescence was localized in the nucleus of histone H3. In addition to conventional fluorescence microscopy, Anap fluorescence can also be visualized by two-photon excitation, which provides deeper tissue penetration, efficient light detection, and reduced phototoxicity. Two-photon excitation of the Anap-mCherry double-labeled histone H3 revealed excellent overlap between Anap (excitation, 730 nm, two-photon; emission, 420–500 nm) and mCherry (excitation, 543 nm; emission, 600–700 nm) fluorescence in the nucleus.
Anap has also been used as a FRET pair in mammalian cells. Chatterjee and coworkers developed a genetically encoded FRET that uses Anap as the donor fluorophore and enhanced green fluorescent protein (EGFP) as the acceptor.\(^{37}\) They co-expressed the EGFP- Y39TAG mutant in HEK293T human embryonic kidney cells along with an orthogonal AnapRS/tRNA pair. This expression system can efficiently perform with Anap concentrations as low as 1 µM in the media. The calculate Förster radius of the Anap-EGFP FRET pair is 49 Å. They demonstrated the use of this FRET pair to monitor proteolysis reactions and protein conformational change in the Ca\(^{2+}\)-dependent binding of the calmodulin protein to its substrate peptide M13.

Another application of the Anap fluorophore is that it can be used to incorporate in the Shaker voltage-gated potassium channel (Kv) at key regions that were previously inaccessible to labeling, and it can be used to study dynamic structural information collected from the cytosolic side of the channel which was not possible with the other chemical labeling method because of the large number of unwanted binding sites in the cytosol.\(^{38}\)

The dansyl fluorophore has been shown to be ribosomally incorporated into β-galactosidase by Chamberlin. ε-Dansyllysine 7 (ε-DanLys) along with the other two fluorescent amino acids; 5-hydroxytryptophan, 7-azatryptophan, have been site-specifically incorporated \textit{in vivo} using chemically charged tRNAs (note that this method has not been discussed above due it slack of adoptability by other laboratories in comparison to the UaaRS methods).\(^{39}\)
The fluorescent unnatural amino acid 2-amino-3-(5-(dimethylamino)naphthalene-1-sulfonamide)propanoic acid (8, Dansylalanine) was genetically encoded in *Saccharomyces cerevisiae* with high fidelity and good yield using an orthogonal RS/tRNA pair.\(^{40}\) This environmentally sensitive fluorophore was selectively inserted into human superoxide dismutase and used to monitor unfolding process in the presence of guanidinium chloride.

Lei Wang also reported the incorporation of dansylalanine 8 into a voltage-dependent membrane lipid phosphatase, specifically the voltage-sensitive domain (VSD) of *Ciona intestinalis* voltage-sensitive phosphatase (CiVSP). They described a lentiviral-based gene delivery method to stably incorporate Uaas into proteins expressed in HCN-A94 neural stem cells. This genetically encoded fluorescent Uaa optically reported on the conformational change of the voltage-sensitive domain in response to membrane depolarization in differentiated neurons.\(^ {41}\)

The incorporation of fluorescent unnatural amino acid 4-biphenyl-L-phenyl alanine (9, Terphenyl) in Figure 1.6 into GFP has been achieved by Douglas Young and coworkers.\(^{42}\) Terphenyl moieties have relatively high quantum yields ($\phi = 0.49$), long fluorescent lifetimes ($\tau = 4.38$ ns) and novel emission spectra with large Stokes shifts ($\lambda_{\text{ex}}/\lambda_{\text{em}} = 280/342$ nm). This terphenyl moiety has also been found to be environmentally sensitive and widely used in various applications including two-photon laser scanning microscopy, femtosecond fluorescence spectroscopy, and $\alpha$-helical secondary structure investigations. Previously, Sidney Hecht’s group demonstrated that various terphenyl derivatives could be incorporated into dihydrofolate reductase from *E. coli* for FRET
studies by using chemically acylated tRNA methods.\textsuperscript{43-44} The terphenyl amino acid has been shown by Young to be site-specifically incorporated into GFP at multiple sites and these GFP mutants has been shown to impact the protonation state of the terphenyl fluorophore, potentially expanding the utility of the protein as a biosensor.\textsuperscript{42} Together, these studies illustrate the wide array of potential applications of fluorescent amino acids.

**Fluorescent Labeling Using Unnatural Amino Acids with Bioorthogonal Reactivity**

An alternative approach for protein labelling can be performed by site-specific incorporation of Uaas with “biorthogonal” handles into proteins by genetic code expansion followed by subsequent labeling of these proteins by specific chemoselective reactions. (Figure 1.7a) These biorthogonal reactions need to proceed under biologically/physiologically compatible conditions and yet react only with the desired partner and not any of the endogenous biomolecule functional groups. Several examples of these reactions have been shown including reaction of ketones/aldehydes with hydrazines/hydroxylamines, the classic copper-catalyzed azide-alkyne cycloaddition reaction (CuAAC), the strain-promoted azide-alkyne cycloaddition reaction (SPAAC), Staudinger ligations, certain palladium-catalyzed cross-coupling reactions, “photoclick” reactions, inverse-electron-demand Diels-Alder reactions.\textsuperscript{17} These chemoselective reactions usually have rate constant in the order of $10^{-4}$ – $10^{4}$ M\textsuperscript{-1} s\textsuperscript{-1} as shown in Figure 1.7b.
Figure 1.7 Labeling proteins via incorporation of unnatural amino acids that can be chemoselectively labeled. (a) An unnatural amino acid bearing a unique bioorthogonal functionality is introduced site-specifically into a protein via genetic code expansion and then chemoselectively labeled with an externally added chemical probe. (b) Rate constants of chemoselective reactions for which one of the partners can be genetically encoded in form of an unnatural amino acid. (Adapted figure reproduces from Lang, K., Chin, J.W. Cellular Incorporation of Unnatural Amino Acids and Bioorthogonal Labeling of Proteins, Chem. Rev. 2014, 114, 4764–4806.)

While Uaas containing bioorthogonal amino acids can be used to install a wide variety of fluorophores, there can be some limitations of the utility of these probes in studying protein folding. Direct genetic encoding of a fluorescent Uaa allows one to place a chromophore on the interior of a protein and to label proteins that cannot be reversibly unfolded and refolded. Having the chromophore attached by a short sidechain rather than a “click” chemistry linker also reduces the positional uncertainty of the FRET probe. For example, in the example of measuring FRET between two probes attached to the protein
α-synuclein (Figure 1.8), commonly used linkers would lead to a 30 Å variance in the distance between the probes. Thus, our work has focused on small fluorophores, relatively closely tied to the protein backbone (unlike Lys derivatives 2-4 or 7) and with rationally tunable fluorescent properties (unlike NAP and Anap).

Figure 1.8 Effects of flexible linkers on the range of chromophore distances in a FRET experiment.
Acridone as a Tunable Fluorescent Scaffold for Generating Fluorescent Amino Acids

The research described in this thesis builds off of previous experiments in which the amino acid acridonylalanine (Acd, Figure 1.9) was shown to be a genetically incorporable fluorescent Uaa. This was initially demonstrated by Sisido and coworkers using in vitro translation of chemically aminoacylated tRNA. Subsequently, our laboratory collaborated with Ryan Mehl to develop an RS-based system for Acd incorporation. However, this AcdRS was prone to misincorporation of a trace byproduct of the Acd synthesis, and thus Chapter 2 describes its optimization. These optimization efforts included collaborations with the Saven laboratory to computationally model different generations of AcdRS resulting from directed evolution experiments, as well as analysis of purified versions of the RSs, using established methods which are briefly described above. While Acd is a valuable blue fluorescent amino acid, many aspects can be improved, including its brightness and modifying its absorption and emission to make red-shifted variants. Thus, Chapter 3 describes these efforts, resulting in aminoacridoylalanine (Aad, Figure 1.9). To arrive at Aad, a series of analogs of the acridone core were prepared and analyzed using photophysical techniques. These data were also compared to calculations performed by George Petersson which showed excellent correlation with the experimental data, making us optimistic about predicting Acd derivatives in the future. Finally, Chapter 4 describes ongoing efforts to improve the synthesis of Aad and rationally modulate its fluorescence to increase its brightness.
Figure 1.9 The structure of acridonylalanine (Acd) and 2-aminoacridonylalanine (Aad)
CHAPTER 2: The Optimization of a Permissive Aminoacyl tRNA Synthetase for a Target Unnatural Amino Acid

**Background**

It is now well-established that protein folding and dynamics play essential roles in health and disease. For example, the small protein calmodulin (CaM) undergoes a dramatic conformational rearrangement to carry out its calcium sensor function in eukaryotic cells.\textsuperscript{47-48} In a second example, the bacterial repressor-protease LexA uses a complex sequence of RecA-induced structural change, self-proteolysis, and dissociation of subunits to sense DNA damage and activate genes that ultimately lead to antibiotic resistance.\textsuperscript{49-50} Finally, the conformational flexibility of the neuronal protein \textalpha-synuclein (\textalpha\text{S}) is a liability, as it leads \textalpha\text{S} to misfold and form amyloid fibrils that contribute to the pathogenesis of Parkinson’s disease.\textsuperscript{51-52} Fluorescence spectroscopy is a powerful tool for studying such processes, as it allows one to observe protein motions in real time under physiological conditions, including measurements in live cells.\textsuperscript{53-54} One can even obtain low resolution structural information using distance-dependent chromophore interactions such as Förster resonance energy transfer (FRET) and quenching by photo-induced electron transfer (eT).\textsuperscript{9} To appropriately model protein motions, one needs a set of probes that are capable of accurately reporting on distance changes without disrupting the fold and function of the protein of interest.\textsuperscript{8}

Recent developments in genetic code expansion and biorthogonal chemistry have made the site-specific incorporation of unnatural amino acids (Uaas) and installation of fluorophores through post-translational modification straightforward, even in cells or lysates.\textsuperscript{17} Unfortunately, the fluorophores used in these approaches are often relatively large and are attached by long flexible linkers, which have a non-trivial impact on the
observed FRET measurements. Additionally, larger fluorophores cannot be introduced co-
translationally to allow packing into the interior of a folded protein. Thus, they will be
restricted to surface-accessible positions, limiting the regions of the protein for which
conformational changes can be studied.

Smaller probes that are more closely tied to the backbone are better able to report
on conformational changes of the protein. The Petersson laboratory has developed small
fluorescent probes and quenchers that should be non-perturbing to proteins and which are
closely tied to the protein backbone. These include thioamide substitutions of the backbone
itself and intrinsically fluorescent Uaas such as acridon-2-ylalanine (Acd or 1).46 These
new approaches are significant because, of the >100 Uaas that have been genetically
encoded in E. coli, only four are fluorescent. 18, 31, 35-36, 40, 46, 55-56

Acd is a blue-wavelength fluorescent amino acid that is a useful fluorophore
because of its small size (222 Å³), near unity quantum yield in water (Φ = 0.95), unusually
long lifetime (τ ~ 15 ns) and high photostability (< 5% degradation after 3 h irradiation).45,
57-58 Previous work in the Petersson laboratory has shown that Acd can be efficiently
quenched by a thioamide through an eT mechanism.59 We have also shown that it can be a
valuable FRET acceptor from Trp or methoxycoumarin, and a donor to more red-shifted
dyes such as nitrobenzoxadiazole (NBD) or fluorescein.46

Prior to our work, Sisido and coworkers had shown that Acd can be ribosomally
incorporated into proteins by in vitro translation using the PURE system. Since protein
yields from this system can be quite low due to its small scale, the Mehl and Petersson
laboratories developed an in vivo system for Acd incorporation using methods pioneered
by Schultz. These methods require the generation of an aminoacyl tRNA synthetase (RS) that is selective for the Uaa and a tRNA that can be selectively charged by the ncRS to deliver the tRNA to an unassigned codon, typically the amber stop codon UAG (tRNA_{CUA}). An RS was selected from a library of permissive *M. janaschii* (*Mj*) tyrosyl RS mutants that had previously been shown to incorporate bulky aromatic amino acids such as 4-(2′-bromoisobutyramido) phenylalanine (Brb) and *p*-benzoyl phenylalanine (*Bzf*). The most active mutant from this library (G2) was used to express Acd-containing variants of CaM, triose phosphate isomerase, and αS. Here, we will refer to this mutant as AcdRS1.

Previous Acd syntheses reported by Lankiewicz and Sisido used a NaOH deprotection step and acid-catalyzed Friedel-Crafts cyclization in polyphosphoric acid (PPA). However, the yields from this route were only 23% from the *p*-nitrophenylalanine starting material, and even this material presumably had substantial racemization due to hydroxide treatment. We improved the synthesis by starting from the natural amino acid Tyr (2) using a Buchwald-Hartwig coupling to *O*-methyl anthranilate (4) to form 5. (Scheme 2.1) We also found heating in sulfuric acid with the addition of water to be a preferable way to induce Friedel-Crafts cyclization, as it resulted in deprotection of the Boc group and methyl esters as well. We obtained an 86% overall yield from Tyr using the route shown in Scheme 2.1 with *H*₂*SO₄* used in the final step (no racemization observed by high performance liquid chromatography, HPLC). Unfortunately, this route also produced a decarbonylated by-product, *N*-phenyl-aminophenylalanine (*Npf* or 6) in trace amounts. We found that *Npf* was in fact incorporated much more efficiently by AcdRS1.
than Acd. Thus, although Npf was present at < 1% according to HPLC analysis, roughly 50% of the expressed protein contained Npf.

We were able to initially address this problem by converting 5 to Acd using LiOH deprotection followed by PPA cyclization. This eliminated Npf formation, but limited us to a 44% overall yield from Tyr with complete racemization (a 22% yield of the requisite L-Acd form). While we were able to express proteins containing exclusively Acd, this was not a very satisfactory solution, and we sought to obtain an AcdRS that was sufficiently selective against Npf that we could use the higher yielding \( \text{H}_2\text{SO}_4 \) synthetic route. We also recognized that the paradigm of using a permissive RS library to obtain an initial, working RS and subsequently optimizing it for a target amino acid would be a strategy that could be employed by other laboratories if we were successful.

\[
\text{Scheme 2.1} \text{ The synthesis of acridon-2-ylalanine (Acd).}
\]
Synthetase Selection

A typical selection experiment to evolve an RS for a Uaa consists of rounds of positive selection (performed in the presence of the Uaa and the 20 canonical amino acids) and negative selection (performed in the presence of only the 20 canonical amino acids). To develop an AcdRS specific for Acd only, we performed these standard selections as well as negative selections where Npf was included in the selection media to eliminate those AcdRS mutants charging tRNA with Npf. This resulted in two additional RSs, AcdRS2a (clone G11, from traditional negative selection) and AcdRS2b (clone A9, from Npf counter-selection), which both showed good selectivity in an initial screen in which the Uaa s were incorporated into green fluorescent protein (GFP).

We have performed detailed studies of their in vivo selectivities in expressions of CaM, αS, and LexA and found that AcdRS2b has superior selectivity for Acd when compared to AcdRS2a. We have also expressed and purified AcdRS1 and AcdRS2b to measure their Acd and Npf activation kinetics. We are able to rationalize their selectivities in terms of the X-ray crystal structure of AcdRS1 and a homology model of AcdRS2b. Our study provides an improved Acd incorporation method for fluorescent labeling of proteins, and also validates a general strategy for how one may optimize a permissive RS to eliminate incorporation of an unwanted contaminant.
**Figure 2.1** AcdRS selection. Left: Images of *E. coli* agar plates used in rounds of positive (+ Acd) and negative (+ Npf) selection. Media for both plates also contain sources of the 20 canonical amino acids. Middle: Fluorescence of suspensions of *E. coli* cells expressing GFP with a TAG codon at position 150 using the indicated RS clone and amino acid mixture. Emission was measured at 528 nm with excitation at 485 nm. Top Right: Image of the AcdRS1 (G2) active site with radical polymerization initiator Brb bound. (PDB ID: 4PBR) A favourable hydrogen bond between the carboxylate of Glu65 and the aniline N-H of Brb can be seen. Bottom Right: Sequences of *Mj* RS clones used for incorporation of Acd. Additional sequences of clones from GFP-based screening are given in the experimental method section (**Figure 2.6** and **Table 2.2**).

CaM, LexA and αS constructs with UAG mutations were expressed in *E. coli* along with plasmids encoding AcdRS1 and its cognate tRNA$_{CUA}$ species. We analyzed the selectivity of AcdRS1 based on matrix-assisted laser desorption ionization (MALDI) mass spectrometry (MS) data of both intact and trypsin-digested, purified proteins, including CaM, LexA, and αS. Data for incorporation at position 113 in CaM are shown in **Figure 2.2, 2.3, and 2.7**; additional data for αS and LexA are shown in **Figure 2.8 and 2.9**, respectively. In all cases, when the
proteins are expressed using media containing pure Npf or Acd produced using the PPA route, a single peak for the intended product is obtained in the MALDI spectra. When the proteins are expressed in media containing Acd produced using the H$_2$SO$_4$ route, we observe a roughly 45:55 Npf/Acd ratio of the two CaM species, even though Npf is present only in trace quantities. We also considered the possibility that Npf could be generated \textit{in vivo} by several possible routes. For example, a carboxy-lyase such as YigC could convert 7 (also a < 1% contaminant in the H$_2$SO$_4$ Acd synthesis) into Npf,$^{63}$ or some fraction of Acd could be converted to Npf. Our PPA Acd expression data allow us to exclude metabolic processing of Acd, since we see no Npf incorporation in this case. When we use media in which we intentionally include 1% or 10% Npf with Acd, we detect products containing Npf/Acd ratios of 65:35 and > 95:5, respectively. (Figure 2.8) These data and similar data for other proteins indicated to us that the selectivity of our nominal AcdRS—designed to incorporate Acd specifically—in fact favored Npf incorporation by approximately 100-fold. Despite the incongruity between its intended use and its actual fidelity, one should keep in mind that AcdRS1 was selected based on expression yields of GFP in the presence of media containing Acd from the H$_2$SO$_4$ route (and therefore containing ~1% Npf). Similar contamination, barely detectable by HPLC analysis, may be present in other stocks of home-synthesized or commercial batches of Uaas used in RS selection.
**AcdRS Optimization**

**RS Selection to Remove Npf Activity.**

*(Note: Synthetases selection experiments were performed by Robert J. Blizzard, Joseph J. Porter, and Professor Ryan A. Mehl at Department of Biochemistry and Biophysics, Oregon State University)*

To reduce the incorporation of Npf, we screened mutant *Mj* TyrRS libraries with positions in the amino acid binding pocket randomized using a GFP expression screen common to the Mehl laboratory.\(^{61,64}\) Two rounds of positive and negative selection were performed according to standard protocols, with the desired Uaa (i.e., Acd) present in the media for rounds of positive selection and only the 20 canonical amino acids present in the media for rounds of negative selection. In parallel, we also performed a similar two round selection experiment with a novel negative selection step that included the undesired Npf in the media. Thus, in this second “counter-selection” protocol, we explicitly selected against Npf incorporation rather than relying on high activity for Acd to be mutually exclusive of activity for Npf. The *Mj* RS libraries were based around the G2 (AcdRS1) and F9 clones, both of which incorporate Acd, and have very similar sequences. The level of Npf misincorporation is noticeably higher for F9 than for G2, highlighting the idea that small sequence changes can have a large impact on selectivity.

From these two selection protocols, eleven RS clones were identified that showed high levels of Acd incorporation. Of these clones, only A9 derived from the Npf counter-selection. Among the ten clones from the standard negative
selection, the G11 clone demonstrated the highest level of selectivity as measured by the GFP fluorescence of cell suspensions expressed in media containing either 1 mM Acd or 1 mM Npf. (Figure 2.1 and 2.6) Therefore, the G11 and A9 RSs were cloned into the pDule2 vector for expression of other proteins, and are referred to as AcdRS2a and AcdRS2b, respectively.

**Comparison of AcdRS Active Sites.**

The availability of an X-ray crystal structure of AcdRS1 (G2) allowed us to examine the sequences of various library members to understand how changes might confer increased selectivity for Acd relative to Npf. Although we were never able to obtain suitably diffracting crystals of an Acd complex with AcdRS1, the Mehl group has previously published a structure of this RS with Brb bound. (Figure 2.1, Top right) Recall that “AcdRS1” is a permissive RS, as can be seen in an analysis of its activation of a variety of aromatic Uaa in Cooley et al. Brb, Npf, and Acd share the feature of a nitrogen atom in the *para* position of the phenylalanine ring, which seems to be a key recognition determinant as it can hydrogen bond with Glu65. Both of the RS clones that give the highest selectivity for Acd, AcdRS2a (G11) and AcdRS2b (A9), have a mutation at this position, but one that maintains the potential hydrogen bond acceptor functionality. Other library members with Pro, Trp, Ile, or Val at position 65 had less than ten-fold selectivity for Acd over Npf in the GFP assay. (See experimental method section for details) Further analysis of AcdRS1
(G2), AcdRS2a (G11), and AcdRS2b (A9) selectivity based on homology modelling and docking studies is given below.

Analysis of AcdRS Selectivity

In Vivo Characterization of AcdRS Selectivity.

To more rigorously investigate the selectivity of AcdRSs 2a and 2b, we expressed proteins in E. coli, purified them, and analyzed incorporation by MALDI MS of whole proteins and trypsin digests. As above, each experiment was carried out under five media conditions, varying in the amino acid provided and/or the synthetic route by which it was obtained: with H$_2$SO$_4$ Acd, PPA Acd, Npf, 1% Npf/Acd, and 10% Npf/Acd. Data for incorporation at position 113 in CaM are shown in Figure 2.2 and Figure 2.3; data for αS and LexA are shown in the Figure 2.8 and Figure 2.9. As anticipated from the GFP screening data, we found that both AcdRS2a and 2b have improved selectivity against Npf. However, AcdRS2a still gives a mixture of Npf- and Acd-containing protein when H$_2$SO$_4$ Acd is used in the growth media. In contrast, proteins expressed using AcdRS2b contain only Acd, even when challenged with 10% Npf in the media. This level of selectivity was observed in CaM, αS, and LexA. (Figure 2.2 and Figures 2.7 – 2.9) It should also be noted that AcdRS2b selectivity is observed in a variety of media. CaM and αS are expressed in minimal media or Luria broth (LB) with isopropyl β-D-1-thiogalactopyranoside (IPTG) induction, while LexA is expressed in richer,
arabinose auto-induction media. Based on these data, we selected the AcdRS2b (A9) for further characterization.

**Figure 2.2 In vivo AcdRS selectivity.** CaM (UAG codon at 113) was expressed in minimal media containing 1 mM Uaa: Acd, synthesized either using the H$_2$SO$_4$ route or PPA route, or Npf. Expression was performed with one of the three AcdRSs indicated. Significant incorporation of Npf is seen for AcdRS2a (G11) when using H$_2$SO$_4$ Acd, but only Acd-containing protein is seen with AcdRS2b (A9).

For a more rigorous, quantitative analysis of selectivity, we analyzed the CaM trypsin digest data by normalizing the intensity of the peak for the 108-116 fragment, containing Acd or Npf at position 113, to the intensity of the peak for the 117-127 fragment. (Figure 2.3) CaM$_{117-127}$ should be produced in a 1:1 ratio with CaM$_{108-116}$ when the protein is completely digested by trypsin, and this can be confirmed by varying the digest time and observing that the intensity ratios do not change (data not shown). Normalization using an internal standard is essential to interpreting the intensity data correctly. The CaM$_{108-116}$Acd$_{113}$ fragment (A in Figure 2.3) ionizes 5.5-fold better than the CaM$_{117-127}$ fragment (* in Figure 2.3),
while the CaM\textsubscript{108-116}Npf\textsubscript{113} fragment (N in Figure 2.3) ionizes 6.1-fold worse than the CaM\textsubscript{117-127} fragment. This can be seen by examining the peak ratios for PPA Acd (i.e., Acd only) and Npf expressions. After peak scaling, one obtains an MS-based Acd/Npf selectivity ratio (MS Sel) of \(2.6 \times 10^{-3}\) for AcdRS1 and 0.20 for AcdRS2b, calculated as follows:

\[
\text{MS Sel} = \frac{\text{Scaled Acd}}{\text{Scaled Npf}} / \frac{\text{Acd/Npf ratio in media}}
\]

The AcdRS1 MS Sel value is in good agreement with our estimate of 100-fold selectivity for Npf based on the less quantitative whole protein MALDI MS data. While the AcdRS2b MS Sel value may seem surprisingly low given the absence of any obvious CaM\textsubscript{108-116}Npf\textsubscript{113} peak in the MALDI spectra, it is important to keep in mind that the RS is only being challenged with at most 10% Npf in the media and that CaM\textsubscript{108-116}Npf\textsubscript{113} ionizes 30-fold worse than CaM\textsubscript{108-116}Acd\textsubscript{113}. Using these Acd/Npf selectivity ratios, we determined that selection resulted in a 76-fold improvement in selectivity for AcdRS2b relative to AcdRS1. We note that the effective \textit{in vivo} Acd selectivities for both AcdRS1 and AcdRS2b are less than one and are probably influenced by preferential uptake of Npf into cells.
Figure 2.3 CaM\textsubscript{113} trypsin digest data for AcdRS selectivity analysis. CaM (UAG codon at 113) was expressed in minimal media containing 1 mM Uaa: Acd, synthesized either using the H\textsubscript{2}SO\textsubscript{4} route or PPA route, Npf, or either a 99:1 or 90:10 mixture of PPA Acd and Npf. Expression was performed with one of the three AcdRSs indicated. The peaks for the (M+H)+ masses of the CaM\textsubscript{108-116}Acd\textsubscript{113} fragment (A, 1179.8 Da), CaM\textsubscript{108-116}Npf\textsubscript{113} fragment (N, 1153.8 Da), and CaM\textsubscript{117-127} fragment (*) 1349.9 Da) are indicated. The intensities of the CaM\textsubscript{108-116}Acd\textsubscript{113} and CaM\textsubscript{108-116}Npf\textsubscript{113} fragment peaks were normalized using the intensity of the CaM\textsubscript{117-127} fragment peak.
In Vitro Characterization of RS Activity.

To better understand AcdRS1 and 2b selectivity, we generated His-tagged variants of the enzymes, then expressed and purified them for in vitro activity assays. Charging of tRNA by RSs is a two-step process, where the first step is the RS-catalyzed reaction of amino acid with ATP to form an aminoacyl-adenylate intermediate (Aa-AMP), releasing inorganic pyrophosphate (PPi); and the second step is the reaction of this enzyme-bound adenylate with the 2’ or 3’ hydroxyl group on A76 at the 3’ end of the tRNA.65

\[
\text{Aa + ATP} \rightarrow \text{Aa-AMP + PPi} \quad (1)
\]

\[
\text{Aa-AMP + tRNA} \rightarrow \text{Aa-tRNA + AMP} \quad (2)
\]

While some prior studies have used assays that measure only the first step of the aminoacylation reaction to demonstrate that in vitro activities are consistent with Uaa incorporation,10,66-67 it has been shown by the Perona laboratory that full tRNA aminoacylation assays correlate well with in vivo observations of RS activity.68 This is expected, since amino acid incorporation into protein in vivo can only occur upon aminoacylation. Thus, aminoacylation of an in vitro transcribed \(^{32}\)P-labeled tRNA\(_{CUA}\) was measured under single-turnover conditions at a variety of Acd or Npf concentrations for AcdRS1 and AcdRS2b. Plots of the first-order rate constants as a function of amino acid concentration were used to determine \(k_{\text{obs}}\) and \(K_d\) for each amino acid and RS combination, as previously described.69 (Figure 2.4 and Table
2.1) For these measurements, $k_{\text{obs}}$ corresponds to the microscopic rate constant for the chemical steps of aminoacylation, or to a closely linked first-order rearrangement that follows tRNA binding and precedes aminoacylation.

Comparative kinetic analysis of AcdRS1 and AcdRS2a reveal catalytic preferences that correlate well with measurements of amino acid incorporation \textit{in vivo} and mass spectrometry of proteins incorporating Acd and Npf. AcdRS1 aminoacylates tRNA$_{CUA}$ with Acd three-fold slower than with Npf, while AcdRS2b aminoacylates tRNA$_{CUA}$ with Acd five-fold faster than with Npf (Table 2.1). These changes in $k_{\text{obs}}$ are primarily responsible for the change in selectivity between these two RSs, as there are only relatively minor differences in $K_d$ for the amino acids. The enzymological selectivity (Enz Sel), calculated as

\[ \text{Enz Sel} = \frac{k_{\text{obs}}(\text{Acd})/K_d(\text{Acd})}{k_{\text{obs}}(\text{Npf})/K_d(\text{Npf})} \]

changes from about a two-fold Npf preference by AcdRS1 to 16-fold Acd selectivity for AcdRS2b. Overall, AcdRS2b is improved in selectivity for Acd relative to Npf by 34-fold, in reasonable agreement with the 76-fold specificity shift derived from the \textit{in vivo} measurements made based on trypsin digest peak intensities. Note that while the relative change in selectivity is consistent between the \textit{in vitro} and \textit{in vivo} data, the actual selectivities are quite different, presumably because the \textit{in vivo} MS Sel is the product of Enz Sel and factors affecting the availabilities of the amino acids inside the cells. We estimate that the effective Npf concentration in cells is
about 100-fold greater than the Acd concentration. This could be the result of preferential uptake of Npf, or sequestration of Acd by binding to other targets. Such targets could include DNA or RNA, both of which are known targets of acridone-based intercalator molecules.\textsuperscript{70} Thus, to prevent Npf misincorporation, AcdRS2b must be significantly Acd-selective, even though Npf contaminants are only present at about 1% (10 µM) in the media.

**Figure 2.4** tRNA Aminoacylation kinetics. Plot of $k_{obs}$ for reactions of AcdRS1 and AcdRS2b as a function of Acd or Npf concentration. Reaction rates were determined using thin layer chromatography following nuclease digestion of aminoacylated tRNA as described in the ESI. Sample primary data are shown in Fig. 2.12 Rate constants were determined from progress curves for each enzyme at varying amino acid concentrations.

*(Note: These tRNA aminoacylation kinetic assays were tested by Professor John J. Perona at Portland State University)*
Table 2.1 AcdRS *In vitro* enzymology and *in vivo* MS selectivity (Sel) parameters

<table>
<thead>
<tr>
<th>AcdRS</th>
<th>( k_{\text{obs}}(\text{Acd}) \times 10^{-3} \text{ s}^{-1} )</th>
<th>( K_a(\text{Acd}) ) µM</th>
<th>( k_{\text{obs}}(\text{Npf}) \times 10^{-3} \text{ s}^{-1} )</th>
<th>( K_a(\text{Npf}) ) µM</th>
<th>Enz Sel</th>
<th>MS Sel</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (G2)</td>
<td>8.0 ± 2.0</td>
<td>45.1 ± 18.8</td>
<td>24.0 ± 1.0</td>
<td>64.1 ± 10.8</td>
<td>0.47</td>
<td>2.6 x 10^{-3}</td>
</tr>
<tr>
<td>2b (A9)</td>
<td>28.0 ± 2.0</td>
<td>39.9 ± 6.8</td>
<td>5.0 ± 1.0</td>
<td>~114 ± 10.0</td>
<td>16</td>
<td>0.20</td>
</tr>
</tbody>
</table>

Modeling of AcdRSs complexed with Npf and Acd.

*(Note: Modeling experiments were done by José A. Villegas and Professor Jeffery G. Saven at Department of Chemistry, University of Pennsylvania)*

To inform the differences in Acd and Npf charging activity, models were built based upon the crystallographic structure of AcdRS1 complexed with Brb (PDB ID: 4PBR).\(^{64}\) In generating a model of AcdRS1 (G2), the phenyl ring on the Brb ligand guided the placement of Acd and Npf within the binding pocket. Hydrogen atoms were added to the protein, and the ligand structure was energy minimized within the fixed protein structure. For Acd, a void volume remains between the molecule and the G\(_{32}\)-G\(_{34}\) β-strand. In each G2 complex, Glu\(_{65}\) is poised to accept a hydrogen-bond from the N-H of the ligand’s side chain. For Npf, the additional free volume allows the molecule to relax to a twisted conformation, alleviating repulsive internal steric contacts between the ligand’s two phenyl rings.

Using the same protein structure, computational protein design methods developed in the Saven laboratory were used to generate a model of A9,\(^{71-74}\) followed by energy minimization of the atomic coordinates of the ligands and the mutated side chains. In A9, the Gly\(_{32}\)Ala mutation forces Npf to adopt a more planar
conformation of the ligand’s two phenyl rings in order to form a hydrogen bond with Asp$_{65}$. Instead, Npf relaxes to avoid this high-energy distortion, which not only prevents hydrogen bonding of the Npf aniline N-H with Asp$_{65}$, but also forces the Asp$_{65}$ sidechain to move. While it is possible that the actual conformations adopted in the active site are different in the presence of tRNA, these models are consistent with AcdRS2b’s increased activity toward Acd relative to Npf. Thus, it appears that the combination of the Gly$_{32}$Ala and Glu$_{65}$Asp mutations provides the increase in Acd selectivity for AcdRS2b (A9).

**Figure 2.5** AcdRS homology models. Acd or Npf were energy minimized in the active site of AcdRS1 (G2), taken directly from PDB structure file 4PBR, or a model of AcdRS2b (A9). Npf is accommodated in the AcdRS1 active site where the phenyl rings can become non-planar with respect to each other. In the A9 active site, Npf twisting is restricted by the Ala$_{32}$ sidechain, and a hydrogen bond cannot be made with Asp$_{65}$.
Conclusion

Our elucidation of the process by which trace Npf in the Acd media was incorporated into expressed proteins and our subsequent optimization of AcdRS1 provide a cautionary tale to the field of genetic code expansion, as well as a potentially general solution to the problem. Explicit inclusion of the unwanted Npf in the counter-selection step allowed us to obtain AcdRS2b, which had much higher selectivity toward its cognate Uaa than AcdRS2a, obtained through standard rounds of positive and negative selection (against only the 20 natural amino acids). Trace contaminants in batches of Uaas, synthesized either in-house or bought from commercial vendors, may indeed be utilized by many RSs in the literature which have never been purposefully selected for or against charging of these Uaas. Moreover, variations of in vivo availability (resulting from differential uptake, for example) may exacerbate these problems, as appears to be the case with Npf.

Permissive RSs in libraries, which provide a very efficient way of obtaining RSs for new Uaas, are likely to be susceptible to this problem. Indeed, RSs with open binding sites that accept bulky amino acids such as Acd and Brb may be particularly prone to this difficulty. However, using such a permissive RS with an open pocket is sometimes necessary to finding an RS for an amino acid such as Acd. In some cases, RS selection for bulky Uaas has required an initial round of selection using a more moderate sized amino acid to expand the binding pocket before selection for charging of a bulky aromatic Uaa. Such RSs may be prone to misincorporation of amino acids similar to that intermediate amino acid. Here, we
have seen that small sequence changes (i.e. Gly->Ala and Glu->Asp) can result in large increases in *in vitro* and *in vivo* selectivity for permissive RSs. Thus, we expect that the counter-selection strategy used here will be generally useful to the genetic code expansion community.

**Experimental Methods**

**General Information**

**Materials**

L-Tyrosine, thionyl chloride, di-tert-butyl dicarbonate (Boc anhydride), methyl 2-aminobenzoate, and phenylsepharose CL-4B resin were purchased from Sigma-Aldrich (St. Louis, MO, USA). *N*-Phenyl-bis(trifluoromethane sulfonimide) was purchased from Oakwood Chemical (West Columbia, SC, USA). *E. coli* BL21(DE3) cells were purchased from Stratagene (La Jolla, CA, USA). *E. coli* ElectroMAX DH10B cells were purchased from Invitrogen (Grand Island, NY, USA). Milli-Q filtered (18 MΩ) water was used for all solutions (Millipore; Billerica, MA, USA). Bradford reagent assay kits were purchased from BioRAD (Hercules, CA, USA). Amicon Ultra centrifugal filter units (3 kDa MWCO) were purchased from EMD Millipore. All other reagents and solvents were purchased from Fisher Scientific or Sigma-Aldrich unless otherwise specified. DNA sequencing was performed at the University of Pennsylvania DNA sequencing facility.
**Instruments**

Low resolution electrospray ionization mass spectra (ESI-LRMS) were obtained on a Waters Acquity Ultra Performance LC connected to a single quadrupole detector (SQD) mass spectrometer. UV-Vis absorption spectra were acquired on a Hewlett-Packard 8452A diode array spectrophotometer (currently Agilent Technologies; Santa Clara, CA, USA). Nuclear magnetic resonance (NMR) spectra were obtained on a Bruker DRX 500 MHz instrument (Billerica, MA, USA). Matrix assisted laser desorption/ionization with time-of-flight detector (MALDI-TOF) mass spectra were acquired on a Bruker Ultraflex III instrument. Analytical HPLC was performed on an Agilent 1100 Series HPLC system. Preparative HPLC was performed on a Varian Prostar HPLC system (currently Agilent Technologies). HPLC columns were purchased from W. R. Grace & Compnay (Columbia, MD, USA).

**Chemical synthesis of acridon-2-ylalanine (Acd)**

**Methyl (S)-2-((tert-butoxycarbonyl)amino)-3-((trifluoromethyl)sulfonyl)oxy phenyl) propanoate (Boc-L-Tyr(OTf)-OMe, 3)**. L-Tyrosine (2) (5.07 g, 27.99 mmol) was reacted with SOCl₂ 10.21 mL (139.94 mmol) in 50 mL of cooled MeOH. The reaction was allowed to stir overnight at room temperature under Ar. The reaction mixture was concentrated under vacuum and washed with MeOH. 6.44 g of crude L-tyrosine methyl ester (99% yield) was obtained after drying under vacuum overnight. The crude product was further dissolved in 120 mL tetrahydrofuran (THF)/H₂O (3:1). Na₂CO₃ 3.37 g (31.80 mmol) was then added, and the reaction was stirred for 10 min. Then 6.98 g Boc₂O was
added, and the solution was stirred at room temperature overnight. 100 mL H$_2$O was added to the solution and then the reaction was acidified with 3 M HCl to pH 3. The organic layer was extracted with ethyl acetate (EtOAc), dried with MgSO$_4$, and concentrated under reduced pressure. The product can be recrystallized by dissolution in THF/hexane (3:1). $N$-Boc-L-tyrosine methyl ester (4.50 g, 15.24 mmol) was further dissolved in 30 mL CH$_2$Cl$_2$. Triethylamine (6.38 mL, 45.72 mmol), 4-(dimethylamino) pyridine (0.186 g, 1.524 mmol), and $N$-phenyl bis(trifluoro-methanesulfonimide) (7.62g, 21.34mmol) were then added, respectively. The reaction was stirred overnight. The reaction mixture was concentrated under reduced pressure and worked up with NH$_4$Cl and then extracted by EtOAc. The crude product was purified by flash chromatography, eluting with 8:2 hexanes/EtOAc. (R$_f$ = 0.26). The purified product Boc-L-Tyr-OTf-OMe (3) was obtained as colorless oil 5.77 g (88.62% yield). ESI-LRMS and NMR spectra matched previous reports.$^{46}$

(S)-Methyl 2-((4-(2-((tert-butoxycarbonyl)amino)-3-methoxy-3-oxopropyl)phenyl)amino) benzoate (5). 75 mL degassed toluene was added to Boc-L-Tyr-OTf-OMe (12) (3.000 g, 7.02 mmol) in a dried round-bottom flask, followed by methyl 2-aminobenzoate (4) (1200 µL, 9.27 mmol). The solution was degassed with Ar for 5 min. Then palladium(II) acetate (0.082 g, 0.365 mmol), and racemic 2,2'- bis(diphenyl-phosphino)-1,1'-binaphthyl (0.054 g, 0.087 mmol) were added to the flask. Cesium carbonate (6.88 g, 21.1 mmol) was ground and added to the flask. The flask was then fitted with a reflux condenser and heated to 135 °C for 23 h. After the solution was allowed to cool to ambient
temperature, the contents were filtered through a short plug of silica gel using CH$_2$Cl$_2$ to transfer the material to the silica (200 mL), and then ethyl acetate (300 mL) was used to elute the product. The clarified solution was then concentrated under reduced pressure and further purified by flash column chromatography (9:1 Hexane/EtOAc) afforded 2.759 g of compound 5 (6.44 mmol, 91.7%) as a yellow oil. (R$_f$ = 0.2 in 85:15 Hexane/EtOAc) ESI-LRMS and NMR spectra matched previous reports.$^{46}$

**Acridon-2-ylalanine (Acd, 1) (H$_2$SO$_4$ route).** A solution of 13.5 M sulfuric acid (12 mL) was added to a flask containing 5 (1.02 g, 2.38 mmol). The flask was then fitted with a reflux condenser and heated to 115 °C for 16 h in an oil bath. 80 mL water was then added to the flask and allowed to stir for 15 min. The reaction was then removed from the hot oil bath and allowed to cool down. Upon reaching ambient temperature, the solution was cooled to 4 °C and allowed to stand for 2 h. 100 g of ion-exchange resin (Dowex® 50WX8 hydrogen form, strongly acidic cation exchange resin) was made into a slurry with 1.8 M aqueous H$_2$SO$_4$ and applied to a flash chromatography column. The resin was washed with 350 mL 1.8 M aqueous H$_2$SO$_4$, 2 L of water, 1 L of 1.5 M aqueous NH$_4$OH, and 4 L of water. Following these washes, the resin was dried by passing air through the column. The cooled Acd solution was then vacuum filtered on a Büchner funnel to remove precipitated material, and the clarified solution was applied to the washed and dried ion-exchange resin. The resulting resin slurry was shaken in the chromatography column for 5 min before the solution was drained. This solution was then reapplied to the dried resin and shaken for an additional 5 min. The twice-passed solution was then set aside. The loaded resin was
washed with 4 L water before the compound of interest was eluted with 1.45 L of 1.5 M NH₄OH. The solution was concentrated to 50 mL by rotary evaporation, and then lyophilized to dryness, yielding a crop of Acd (1) as a yellow powder (0.6375 g 2.26 mmol 94.9%). The ion-exchange resin was recycled by washing with 4 L of water and dried until further use. To maximize yield, the twice-passed solution was reapplied to washed and dried ion-exchange resin, and the process was repeated to yield a second crop of Acd (0.0232 g of 0.082 mmol 3.4%).

**Acridon-2-ylalanine (Acd, 1) (PPA route).** Compound 5 2.16 g (5.04 mmol) was added to a round-bottom flask. Then 50 mL THF was added and the solution cooled to 4 °C. LiOH (3.31 g, 138.15 mmol) was dissolved with 150 mL water. The LiOH solution was then slowly added to the reaction. The mixture was stirred under Ar at 4 °C overnight. The pH of the reaction mixture was adjusted to 3 with 6 M HCl. Crude product was extracted with CH₂Cl₂ (3 x 100 mL), dried with MgSO₄, and concentrated under reduced pressure. The resulting crude oil was re-dissolved in 25 mL CH₂Cl₂ and cooled to 4 °C for 16 h. A first crop of recrystallized product was collected by vacuum filtration and washed with 100 mL of cold CH₂Cl₂ (1.63 g, 4.08 mmol, 80.8 %). The product was obtained as a white powder (82.6 % yield). The product was further reacted with polyphosphoric acid (PPA).

66.62 g of polyphosphoric acid (PPA) was added into a round-bottom flask with a stir bar. A round-bottom flask was heated to 135 °C in an oil bath. The (S)-2-((4-(2-((tert-butoxycarbonyl)amino)-2-carboxylethyl)phenyl)amino)benzoic acid (1.43 g, 3.57 mmol)
was added to the flask and stirred for 2 h. Then 50 mL water was slowly added (2 mL portions over 10 min) and the reaction allowed to cool to 60 °C. After stirring for 1 h at 60 °C, the reaction was cooled to ambient temperature. Insoluble impurities were removed by vacuum filtration and the clarified solution was adjusted the pH to 4 by addition of 8 M NaOH. Then the solution was cooled to 4 °C for precipitation about 16 h. Crude product was collected by vacuum filtration. After drying, the crude material was re-suspended in 50 mL water and brought into solution by adjusting the pH to 9.0. Insoluble impurities were removed by vacuum filtration. The pH of the clarified solution was then adjusted to 5.5 with 6 M HCl. The yellow precipitate was collected by vacuum filtration and dried overnight affording 0.66 g of Acd (1) (1.86 mmol, 65.3% yield).
Figure 2.6 Fluorescence measurements of RSs with GFP reporter. Grey, green, and blue represent fluorescence from colonies induced in media containing no Uaa, 1 mM Acd, or 1 mM Npf, respectively. Expressions of 500 µL were grown for 48 hours before 2-fold dilution of suspended cells directly from culture with PBS. Fluorescence measurements were collected using a BIOTEK® Synergy 2 Microplate Reader.
Table 2.2 Sequence of top performing AcdRSs. The A9 and G11 clones were moved into the pDule2 plasmid for protein expression.

<table>
<thead>
<tr>
<th>Variant</th>
<th>Residue number</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>32</td>
</tr>
<tr>
<td>WT</td>
<td>Y</td>
</tr>
<tr>
<td>G2</td>
<td>G</td>
</tr>
<tr>
<td>A9</td>
<td>A</td>
</tr>
<tr>
<td>B11</td>
<td>G</td>
</tr>
<tr>
<td>E2</td>
<td>G</td>
</tr>
<tr>
<td>H12</td>
<td>A</td>
</tr>
<tr>
<td>G11</td>
<td>A</td>
</tr>
<tr>
<td>E6</td>
<td>G</td>
</tr>
<tr>
<td>E7</td>
<td>G</td>
</tr>
<tr>
<td>E10</td>
<td>G</td>
</tr>
<tr>
<td>G8</td>
<td>A</td>
</tr>
<tr>
<td>G10</td>
<td>A</td>
</tr>
</tbody>
</table>

In Vivo Characterization of AcdRSs Selectivity

Acd Mutant Calmodulin Protein Expression. For calmodulin (CaM) mutant expression, pCaM-L113TAG and pDule2-AcdRS plasmids was used to transform E. coli BL21(DE3) cells. The pDule2-AcdRS1, pDule2-AcdRS2a, and pDule2-AcdRS2b plasmids encode AcdRS1(G2), AcdRS2a(G11) and AcdRS2b(A9), respectively. The pCaM-L113TAG plasmid confers ampicillin (Amp) resistance and the pDule2-AcdRS plasmids confer streptomycin (Strep) and spectinomycin (Spec) resistance. Transformed cells were selected on the basis of Amp and Strep or Amp and Spec resistance. Single colonies were used to inoculate 5 mL of LB media supplemented with Amp and Strep (100 µg/mL each).
To an autoclaved solution containing 42 mM Na$_2$HPO$_4$, 22 mM KH$_2$PO$_4$, 19 mM NH$_4$Cl, and 86 mM NaCl (M9 salts), the following autoclaved solutions were added per liter of M9 salts: 1 mL of 2 M MgSO$_4$, 1 mL of 15 mg/mL FeCl$_2$ (in 1.0 M HCl), 1 mL of 15 mg/mL ZnCl$_2$ (in acidified H$_2$O), 6.25 mL of 40% glucose, 100 µL of 1 M CaCl$_2$ and 2 mL of 10% BactoTM Yeast Extract. The primary 5 mL culture was incubated at 37 °C with shaking at 250 rpm for 4 h. Upon reaching saturation, the primary culture was added to 1 L of M9 minimal media supplemented with Amp and Strep (or Spec). The 100 mL culture was incubated at 37 °C with shaking at 250 rpm until the OD$_{600}$ reached 0.7. At this point, the different amounts of unnatural amino acid were added to the culture depending on the following conditions:

A: a solution of 28.23 mg Acd (1) from H$_2$SO$_4$ route in 5 mL sterile water
B: a solution of 28.23 mg Acd (1) from PPA route
C: a solution of 25.63 mg Npf (6)
D: a solution of 0.26 mg Npf (6) and 27.95 mg Acd (1)
E: a solution of 2.56 mg Npf (6) and 25.41 mg Acd (1)

Acd was solubilized with 5 drops 10 M NaOH. The protein expression was then induced with IPTG. The culture was incubated at 37 °C for an additional 16 h. The cells were harvested at 5000 x g for 15 min, and the resulting pellet was suspended in 15 mL of 3-(N-morpholino) propanesulfonic acid (MOPS) resuspension buffer (50 mM MOPS, 100 mM KCl, 1 mM EDTA, pH 7.5). Following sonication, the cell lysate was allowed to cool.
on ice for 5 min. CaCl$_2$ was added to the sonicated lysate to a final concentration of 5 mM prior to centrifugation for 20 min at 30,000 x g, 4 °C.

CaM was purified from the cleared cell lysate using a phenyl-sepharose (PhS) CL-4B column with EDTA as eluent. Using a total resin bed volume of 10 mL, the column was first equilibrated with 4 column volumes of PhS Buffer A (50 mM Tris base, 1 mM CaCl$_2$, pH 7.5). After the cleared cell lysate was loaded and allowed to pass through the resin, the column was washed with 4 column volumes of PhS Buffer A, 4 column volumes of high-salt PhS Buffer B (50 mM Tris base, 0.5 M NaCl, 0.1 mM CaCl$_2$, pH 7.5), and an additional 2 column volume washes of PhS Buffer A to restore low-salt conditions. CaM was eluted with PhS Buffer C (10 mM Tris base, 10 mM EDTA, pH 7.5) and collected in 1 mL fractions. The presence of protein was detected by SDS-PAGE. Fractions containing protein were combined and dialyzed against water for 16 h at 4 °C.
Figure 2.7  *In vivo* AcdRS selectivity. CaM (UAG codon at 113) was expressed in minimal media containing 1 mM Uaa: Acd, synthesized either using the H$_2$SO$_4$ route or PPA route, Npf, or either a 99:1 or 90:1 mixture of PPA Acd and Npf. Expression was performed with either AcdRS1 (G2) or AcdRS2b (A9). AcdRS2b (A9) can tolerate at least 10% Npf in the growth media with no apparent Npf incorporation.
**Acdd Mutant α-Synuclein Protein Expression.**

For α-synuclein (αS) mutant expression, *E. coli* BL21-Gold (DE3) cells were transformed with the αS-F94TAG plasmid and a pDule2 plasmid containing the AcdRS and tRNA_{CUA} pair. Cells were selected for resistance to both Amp (100 µg/mL) and Strep (100 µg/mL) or Spec (50 µg/mL). Single colonies were used to inoculate 4 mL of LB media. The primary culture was grown at 37 °C with shaking at 250 rpm for 4 h. The 100 mL culture was incubated at 37 °C with shaking at 250 rpm until the OD_{600} reached 0.8 AU. At this point, the different amounts of unnatural amino acid were added to the culture grown under conditions A-E as in the CaM protocol above. The cells were harvested at 5000 x g for 15 min and the resulting pellet was resuspended in 20 mM Tris, pH 8.0 with 1 mM PMSF and sonicated. Following sonication, the cell lysate was boiled for 15 min prior to centrifugation for 20 min at 13,200 x g, 4 °C. The cleared lysate was dialyzed overnight against 20 mM Tris, pH 8.0 at 4 °C prior to purified using a HiTrap Q HP column (GE Healthcare). FPLC fractions were dialyzed against 20 mM Tris, pH 8.0 and stored at 4 °C.
Figure 2.8 In vivo AcdRS selectivity. αS (UAG codon at 94) was expressed in LB media containing 1 mM Uaa: Acd, synthesized either using the H$_2$SO$_4$ route or PPA route, Npf, or either a 99:1 or 90:1 mixture of PPA Acd and Npf. Expression was performed with either AcdRS1 (G2) or AcdRS2b (A9). AcdRS2b (A9) can tolerate at least 10% Npf in the growth media with no apparent Npf incorporation. * indicates MALDI matrix adduct. αSNpf$_{94}$, Calc’d: 14482; αSAcd$_{94}$, Calc’d: 14508.

Trypsin Digest Analysis of Acd Mutants.

Protein Acd mutants were precipitated using 1:4 8.75 M trichloroacetic acid/protein sample and incubated at 4 °C for 15 minutes. The precipitate was centrifuged for 15 min at 13,200 rpm to pellet protein. The protein pellet was then washed three times with cold acetone to remove trace trichloroacetic acid. Trace acetone was removed by incubating protein pellets in a 95 °C water bath for 5 min open to the atmosphere. Protein pellets were
then re-suspended in 6 M guanidinium hydrochloride with 50 mM Tris pH 8.0, and denatured by boiling at 95 °C for 10 minutes. Protein samples were then diluted to 0.75 M guanidinium hydrochloride with 50 mM Tris pH 7.6 and 1 mM calcium chloride. Sequencing grade modified trypsin (0.6 µg, Promega) was used to digest samples for 24 hours at 37 °C. Trypsin digest aliquots (1 µL) were combined with α-cyano-4-hydroxycinnamic acid (1 µL of a saturated solution in 1:1 H₂O/CH₃CN with 1 % TFA) and analyzed by MALDI-MS.

**Table 2.3** CaM 113 trypsin digest intensity scaling.

<table>
<thead>
<tr>
<th>Uaas</th>
<th>G2</th>
<th>A9</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>m/z</td>
<td>Peak Height</td>
</tr>
<tr>
<td>Acd (H₂SO₄)</td>
<td>1153.8</td>
<td>391</td>
</tr>
<tr>
<td></td>
<td>1179.8</td>
<td>8166</td>
</tr>
<tr>
<td></td>
<td>1349.9</td>
<td>5661</td>
</tr>
<tr>
<td>Acd (PPA)</td>
<td>1153.7</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td>1179.8</td>
<td>7574</td>
</tr>
<tr>
<td></td>
<td>1349.9</td>
<td>1091</td>
</tr>
<tr>
<td>Npf</td>
<td>1153.8</td>
<td>518</td>
</tr>
<tr>
<td></td>
<td>1179.8</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>1349.9</td>
<td>3500</td>
</tr>
<tr>
<td>1% Npf/Acd</td>
<td>1153.8</td>
<td>1224</td>
</tr>
<tr>
<td></td>
<td>1179.8</td>
<td>20839</td>
</tr>
<tr>
<td></td>
<td>1349.9</td>
<td>7968</td>
</tr>
<tr>
<td>10% Npf/Acd</td>
<td>1153.8</td>
<td>3034</td>
</tr>
<tr>
<td></td>
<td>1179.8</td>
<td>754</td>
</tr>
<tr>
<td></td>
<td>1350.0</td>
<td>10266</td>
</tr>
</tbody>
</table>

CaM₁₀₈₋₁₁₆Acd₁₁₃/CaM₁₁₇₋₁₁₂ Scaling Factor: 5.47
CaM₁₀₈₋₁₁₆Npf₁₁₃/CaM₁₁₇₋₁₁₂ Scaling Factor: 0.165

CaM₁₀₈₋₁₁₆Npf₁₁₃/CaM₁₁₇₋₁₁₂ Background Factor: 0.0156
CaM₁₀₈₋₁₁₆Acd₁₁₃/CaM₁₁₇₋₁₁₂ Background Factor: 0.0122
LexA Cloning/Expression/Purification.

(Note: LexA proteins were expressed and purified by Zachary M. Hostetler. Trypsin digestions and MALDI-MS characterization were performed by Itthipol Sungwienwong.)

To create a C-terminal His-tagged LexA expression construct, a recombinant *E. coli* *lexA* gene containing the S119A catalytically-inactive mutation was PCR amplified with forward75

A: 5’-GGCAGCCATATGAAAGCGTTAACGG-3’

and reverse

B: 5’-AATCTCGAGCCAGTCGCCGTTGC-3’

primers. This amplified *lexA* gene was sub-cloned into the NdeI and XhoI sites in pET-41a(+) plasmid (EMD Millipore) to produce pET41-LexA-S119A-HIS. To generate an amber stop codon in the S60 position, the pET41-LexA-S119A-HIS plasmid was amplified with forward

C: 5’-ATTGTTTATGGCGCATCAGCGGAATTCGTC-3’

and reverse

D: 5’-GTGATGCGCCCTAAAACAAATTTCAATAACGCC-3’

primers designed for site-directed mutagenesis with Phusion polymerase, creating pET41-LexA-S60TAG-S119A-HIS.76

Chemically competent BLR(DE3) cells (EMD Millipore) were transformed with either of the above pET41-LexA plasmids (Kn resistance) and either of the pDule2-AcdRS1a or pDule2-AcdRS2b plasmids (streptomycin resistance) and selected on LB + Kn (30 µg/mL) + Strep (100 µg/mL) plates. Isolated colonies were grown overnight to saturation in liquid LB + Kn + Strep. Overnight cultures were inoculated at a 1:100 ratio
into 25 mL MD-5051 auto-inducing medium containing 150 µg/mL Kn, 100 µg/mL Strep, and 1 mM acridone (H₂SO₄ route). Cultures were grown for 24 hours in sterile 250 mL polypropylene centrifuge tubes (Corning) in a shaking 37 °C incubator. Cells were harvested at 4000 x g for 30 min at 4 °C.

Cell pellets were resuspended in 4 mL of wash buffer (50 mM sodium phosphate, 300 mM NaCl, 10 mM imidazole, pH 7.1) containing 1x BugBuster Protein Extraction Reagent (EMD Millipore), 0.5 mg/mL lysozyme, 25 U/µL benzonase, and 1x cOmplete EDTA-free protease inhibitor cocktail (Roche). After incubation at room temperature for 30 min, lysed cells were centrifuged at 20000 x g for 30 min at 4 °C. LexA proteins were purified from the clarified cell lysate using 0.5 mL HisPur cobalt spin columns per manufacturer’s instructions (Thermo Scientific), using 300 mM imidazole as eluent. LexA proteins were detected by running purification samples on 15% SDS-PAGE gels. Elutions containing LexA were pooled and buffer-exchanged into low salt buffer (20 mM Tris-HCl, 25 mM NaCl, pH 7.6) using Vivaspin 500 concentration devices.
Figure 2.9 *In vivo* AcdRS selectivity. LexA (UAG codon at 60) was expressed in auto-induction media containing 1 mM Acd, synthesized using the H$_2$SO$_4$ route. Expression was performed with either AcdRS1 (G2) or AcdRS2b (A9). Although the whole protein peaks (left) are very broad, trypsin digest producing the LexA$_{53-64}$ fragment (right) indicates the presence of only Npf in the G2 sample and only Acd in the A9 sample. * indicates MALDI matrix adduct. LexA$_{53-64}$Npf$_{60}$, Calc’d: 1366.5, Obsv’d: 1366.6; LexA$_{53-64}$Acd$_{60}$, Calc’d: 1392.5, Obsv’d: 1392.4

**AcdRS Cloning/Expression/Purification.**

*(Note: These experiments were done together with Zachary M. Hostetler)*

Synthetic genes (IDT) for AcdRS1 and AcdRS2b were amplified with forward

**E:** 5'-GAGCGGATAAACAATCCCTCTTAG-3'

and reverse

**F:** 5'-GTGGTGTTGCTCGAGTCTCTTTTCTCTTAG-3'

primers. The resulting PCR product was digested with NdeI and XhoI and sub-cloned into a pBAD24 plasmid modified with a C-terminal His tag, resulting in the three pBAD24-AcdRS-HIS constructs. Chemically competent BL21(DE3) cells were transformed with one of the three above pBAD24-AcdRS-HIS plasmids (Amp resistance) and selected on
LB + Amp (100 µg/mL) plates. Isolated colonies were inoculated into 200 mL arabinose auto-inducing medium containing 100 µg/mL carbenicillin (Carb). Cultures were grown for 24 hours in 2 L flasks in a shaking 37°C incubator. Cells were harvested at 4000g for 15 min at 4°C.

Cell pellets were resuspended in 30 mL of wash buffer (50 mM HEPES-KOH, 500 mM NaCl, 10 mM imidazole, pH 7.9) containing 1 mM PMSF and 1x cOmplete EDTA-free protease inhibitor cocktail (Roche). Cell suspensions were lysed using a French pressure cell (FA-032, Thermo Electron Corp) and then centrifuged at 20000g for 30 min at 4°C (Sorvall RC 6+). Clarified lysates were loaded with a BioLogic LP System (Bio-Rad) onto a column containing 4 mL of packed HisPur resin (Thermo Scientific), and proteins were eluted with a linear imidazole gradient. AcdRS proteins were detected by running elution fractions on 12% SDS-PAGE gels. Elutions containing AcdRS were pooled and dialyzed three times against 2 L of 20 mM Tris-HCl pH 8.5, 50 mM NaCl, 10 mM 2-mercaptoethanol, and 20% glycerol. Dialyzed samples were loaded onto a 5 mL HiTrap Heparin HP column (GE Healthcare) and proteins were separated with a linear NaCl gradient. FPLC fractions were analyzed by 12% SDS-PAGE, and those containing AcdRS proteins were concentrated using Amicon Ultra centrifugal filter units (Sigma), buffer-exchanged (50 mM HEPES-KOH pH 7.5, 20 mM NaCl, 10 mM MgCl₂, 1 mM DTT, and 50% glycerol), and stored at -80°C.
**AcdRS Activity Test.**

The activity of each AcdRS enzyme was verified by monitoring pyrophosphate release during the initial condensation reaction of Acd or Npf with ATP. To measure pyrophosphate, an enzyme-coupled, colorimetric assay recently described in the literature was used. Briefly, pyrophosphate generated during each reaction is hydrolyzed by excess inorganic pyrophosphatase (PPIase), and free orthophosphate is incubated with malachite green under acidic conditions (Malachite Green Phosphate Assay Kit, Sigma) to generate a malachite green phosphomolybdate complex (600-660 nm absorbance). Several concentrations of each AcdRS enzyme (0, 0.5, 1.0, and 2.0 µM) were incubated in reaction buffer (50 mM HEPES-KOH pH 7.5, 20 mM NaCl, 10 mM MgCl₂, 1 mM DTT) with 2 U/ml PPIase (NEB), 1 mM of either Acd or Npf, and with or without 0.2 mM ATP for 35 min at 37 °C. Reactions were quenched with chilled EDTA to a final concentration of 10 mM. Samples were developed per manufacturer’s instructions in a clear 96-well plate (Corning Costar), and absorbances at 620 nm were read in an Infinite F200 plate reader (Tecan). Standard curves were used to interpolate observed phosphate concentrations, and the data describing calculated pyrophosphate release rate versus AcdRS concentration were fit to a straight line.
Figure 2.10 AcdRS purification. Left: SDS-PAGE gel showing purification samples of AcdRS1 (G2, lanes 2-7) and AcdRS2b (A9, lanes 8-13) taken throughout the cell lysis and HisPur affinity resin steps. Right: SDS-PAGE gel showing final, purified samples of both AcdRS enzymes following the purification scheme described in the supplemental text. For the lanes in each gel, an equal volume of sample and Laemmli buffer (Bio-Rad) were mixed, boiled at 95°, centrifuged, and loaded onto 12% SDS-PAGE gels, which were run for 50 min at 200 V.
Figure 2.11 AcdRS malachite green activity assay. Plots showing rate of inorganic pyrophosphate formation versus AcdRS concentration for AcdRS1 (G2, left) and AcdRS2b (A9, right). Data are from reactions at 37° containing either 1 mM Acd (blue) or 1 mM Npf (green) and in the presence (solid points) or absence (open points) of 0.2 mM ATP. The linear dependence of activity on AcdRS concentration was demonstrated by fitting a straight line through the data points.
Figure 2.12 Representative primary data for aminoacylation. Left: TLC plate showing reactions of the AcdRS2b(A9) enzyme with saturating levels of orthog-tRNA and Acid at varying times. There is almost complete aminoacylation by the end of the timecourse (about 5 min). Right: Timecourse for the AcdRS2b(A9) reaction fit to derive $k_{obs}$. 
CHAPTER 3: The Design of Fluorescent Unnatural Amino Acids Based on the Acridone Scaffold

Introduction

Fluorescence spectroscopy can be a valuable tool for studying the structural dynamics of proteins and protein/protein interactions. There are several common types of protein experiments that employ fluorescence spectroscopy: folding/conformational change experiments, binding experiments, and proteolysis experiments (Figure 3.1). Changes in fluorescence can be used to track protein structural change on the ns timescale using distance dependent chromophore interactions: either Förster resonance energy transfer (FRET) or quenching by photo-induced electron transfer (eT). Such studies require fluorescent probes that enable accurate measurement on a variety of distance ranges between the two chromophores. FRET ranges are characterized by the Förster radius, \( R_0 \), the distance of half-maximal energy transfer for any chromophore pair. For example, the common FRET pair fluorescein (Fam)/ tetramethylrhodamine (Tmr) has an \( R_0 \) of 47 Å and is useful to measure distances in the 30 to 90 Å range. Since many inter-residue distances in proteins are shorter than this, one needs to complement the Fam/Tmr FRET pair with other probe pairs that are better suited to shorter interactions. With this in mind, the Petersson laboratory has developed a methoxycoumarinylalanine (Mcm, 1)/ acridonylalanine (Acd, 2) FRET pair for monitoring distances in the 15 to 40 Å range, and thioamide/Mcm or thioamide/Acd eT quenching pairs for short distance (<15 Å) measurements.
Figure 3.1 FRET experiments. Left: Protein conformational changes, protein–protein interactions, and proteolytic cleavage can be monitored by changes in the intra- or intermolecular distance between two FRET (Förster resonance energy transfer) probes.

In addition to being better suited to short distance ranges than Fam or Tmr, Acd is small enough to be directly genetically incorporated, rather than post-translationally attached, and is less likely to disrupt protein folding.\(^8\) This allows one to place a chromophore on the interior of a protein and to label proteins that cannot be reversibly unfolded and refolded (e.g., our published labeling of LexA).\(^88\) Having the chromophore attached by a short sidechain rather than a Cys maleimide or “click” chemistry triazole linker also reduces the positional uncertainty of the FRET probe with respect to the protein backbone. Thus, distance measurements from FRET should more reliably report on changes in protein conformation. We and others have previously shown that Acd can be a
valuable probe for protein study because of its small size (222 Å$^3$), high quantum yield in water ($\Phi = 0.95$), unusually long fluorescence lifetime ($\tau \sim 15$ ns), and high photostability.$^{45-46, 57}$ We have developed an engineered aminoacyl tRNA synthetase (RS)/orthogonal tRNA pair for selective Acd incorporation by unnatural mutagenesis.$^{46}$ This has allowed us to label proteins and peptides with methoxycoumarin/Acd FRET pairs either through Mcm incorporation by solid phase peptide synthesis (SPPS) or by attachment of methoxycoumarin-maleimide to a Cys residue in a protein.$^{86}$ While many aspects of the Mcm/Acd FRET pair are optimal, such as significant spectral overlap and a high Mcm extinction coefficient at 325 nm where Acd has a minimum, one disadvantage is the small Acd Stokes shift which leads to significant overlap of their emission spectra. This overlap necessitates a challenging deconvolution of the Mcm/Acd spectra in order to determine FRET efficiencies and distance measurements. Thus, an Acd derivative with a larger Stokes shift would be desirable for Mcm FRET.

In addition to improving Mcm FRET, we also wish to alter other fluorescent properties of Acd, such as red-shifting excitation and emission, increasing the extinction coefficient, and altering solvatochromic effects to make brighter derivatives that are better suited to microscopy or single-molecule fluorescence applications. Sisido and coworkers$^{58}$ have previously shown that some of these effects can be achieved by simply homoligating Acd with a benzene ring (benzoacidonylalanine or Bad, 3) to extend the $\pi$-system. However, we wish to make derivatives that can still be incorporated by the ribosome, and Sisido’s laboratory also showed that Bad was not incorporable during in vitro translation with chemically-charged tRNA. Therefore, we will use crystal structures and
computational models of our evolved AcdRS and “rules” for ribosomal permissivity established in previous in vitro translation studies\textsuperscript{46, 89} to restrict Acd substitutions to positions that will allow in vivo tRNA charging and incorporation into proteins.

![Figure 3.2](image)

**Figure 3.2** Fluorescent amino acids based on 7-methoxycoumarin and acridone cores.

Identifying Acd derivatives by making amino acid analogs is synthetically laborious and unnecessary given that Acd and Bad spectroscopic properties are identical to the properties of the respective chromophore cores\textsuperscript{45, 57-58}. Thus, to improve Acd fluorescence, we set out to make a series of acridone (5) core derivatives in order to identify derivatives with sufficiently improved properties to warrant synthesis of the amino acid form for incorporation into peptides. Previous studies of acridone derivatives have shown that many of the spectroscopic properties can easily be modulated through substituent effects, providing strong precedent for our work\textsuperscript{13, 58, 90-96}. Moreover, focusing on the acridone core makes computational modeling more tractable, with the potential to further narrow the scope of synthetic work by predicting absorption and emission spectra.

Here, we prepare a series of acridone derivatives in order to validate the accuracy of our electronic structure calculations and identify a derivative, aminoacridonyllalanine.
(Aad, 4) with substantially red-shifted emission. In addition, our calculations help to explain the origin of fine structure in acridone spectra, an explanation which is in conflict with the conclusions of previous computational and experimental Acd spectroscopy studies, but is consistent with the larger body of acridone literature. Finally, we synthesize Aad and perform initial trials toward its genetic incorporation.

**Results and Discussion:**

**Acridone derivative syntheses**

To synthesize acridone cores that were not commercially available, we used two general strategies, functionalization of acridone and cross-coupling of pre-functionalized aryl units. Nitration of acridin-9(10H)-one 5 was performed with HNO₃ in CH₃COOH at 50 °C, giving a mixture 2-nitroacridin-9(10H)-one (6) and 4-nitroacridin-9(10H)-one (7) (variations in reaction conditions produced a 74% yield of 6 or 38% yield of 7, see Experimental Methods section for details). Reduction of 6 and 7 using Na₂S and NaOH in refluxing EtOH/water gave 2-aminoacridin-9(10H)-one (8) and 2-aminoacridin-9(10H)-one (9), respectively, in quantitative yields (Na₂S used for small scale, Pd/C also used to reduce 7, see experimental details section for details).²⁷ (Scheme 3.1)

![Scheme 3.1 Synthesis of nitroacridones (6 and 7) and aminoacridones (8 and 9).](image)
We also used Buchwald-Hartwig cross-coupling amination between methyl 2-aminobenzoate (10) and aryl bromides 11-12, followed by Friedel-Crafts cyclization, to generate several acridone derivatives. Cross-coupling was accomplished using Pd(OAc)$_2$, racemic 2,2'-bis(diphenyl-phosphino)-1,1'-binaphthyl (rac-BINAP), and Cs$_2$CO$_3$ to generate methyl 2-(phenylamino)benzoates 13 and 14 in 86 and 78% yield, respectively. Then the acridone cores were achieved by saponification with LiOH, followed by Friedel–Crafts cyclization, to give 4-fluoroacridin-9(10H)-one (15, 93%) and 4-methoxyacridin-9(10H)-one (16, 96%). (Scheme 3.2)

![Scheme 3.2 Synthesis of acridones 15 and 16 and benzoacridones 18-20.](image)

The same strategy was used to form benzoacridone derivatives. Buchwald-Hartwig cross-coupling with methyl 3-amino-2-naphthoate (17) and aryl bromides (11-12 and bromobenzene) gave the corresponding methyl 3-(phenylamino)-2-naphthoates in 97-98% yields (see experimental details section). Then the benzo[\(b\)]acridin-12(5\(H\))-one derivatives
were formed by hydrolysis and Friedel–Crafts ring-closing reactions in 85-89% yields. (Scheme 3.2)

These methods provided access to many acridone derivatives for this study and are sufficiently general to provide access to most desired derivatives for future studies. In particular, the cross-coupling route allows us to access derivatives, such as 3-aminoacridone, that may be difficult to synthesize using nucleophilic aromatic substitution, as the directing effects of the ortho carbonyl and amino groups on each benzene ring in the acridone core would strongly favor reactions at the 2- or 4- positions.

Spectroscopic characterization

Following production of the various Acd derivatives, we determined the absorbance and emission profiles, as well as the extinction coefficient, of each core. These data are summarized in Figure 3.3 and Table 3.1, and the spectra of each compound are reported separately in the experimental method section. The absorption and emission spectra are normalized in Figure 3.3 for clarity, raw spectra are shown in the experimental method section and extinction coefficients and emission intensities are given in Table 3.1. Since we ultimately endeavor to utilize these derivatives as fluorescent unnatural amino acids, we attempted to perform the spectroscopic characterization in phosphate-buffered saline (PBS), pH 7.4. However, due to the relatively low solubility of the benzoacridone compounds in water, all measurements were performed in 1:1 acetonitrile/PBS. Extinction coefficients were obtained through serial dilutions of stocks of 300 µM, which is
approximately the solubility limit of benzoacridone in acetonitrile. Fluorescence emission spectra were acquired under the same solvent conditions.

The parent acridone (5) absorbance spectrum features two major peaks in the near UV region with maxima at 382 and 398 nm as well as additional features below 300 nm. Derivatization of the acridone core resulted in changes of the absorbance profile that were both functional group and position dependent. For example, introducing a methoxy group in the 2 position in S1 resulted in a ~15 nm red-shift in the absorbance maximum, with minimal modulation of the line shape, while introduction at the 4 position in 16 resulted in a very minimal shift, with the multi-peak profile becoming less defined. Amino modification at either the 2 or 4 position produces a singular broad feature that is significantly red-shifted. Moreover, the absorbance profile of 2-aminoacridone (8) is pH sensitive (see Figure 3.11). The spectrum takes on a single broad absorbance feature very similar to that of 4-aminoacridone (9) at high pH, which is dramatically reduced at low pH. It is important to note that both the unmodified and the 4-amino Acd chromophores are pH insensitive (see Figure 3.10 and 3.12). Lastly, we observe that extension of the conjugated system in the case of benzoacridone (18) results in an expected shift of the absorbance to higher wavelengths, but displays an unexpected reduction in the extinction coefficient. When modified with either fluoro (19) or O-methyl (20) substituents, the absorbance profile is minimally shifted and the peaks become less well-resolved, similar to what was observed when modifying the parent acridone scaffold (15 and 16).

Acridone substitution also elicited changes in the emission profile and Stokes shift. The unmodified acridone emission features two major peaks at 412 and 435 nm with a
minor peak around 460 nm. As in the absorbance profiles, the maxima of these peaks move to lower energy upon modification with the O-methyl group (S1 or 16) with the multi-peak profile becoming less defined. Similarly, introduction of the amine functionality resulted in a significant red shift as well as a reduction of the multi-peak line shape to a single broad emission. Both compounds 8 and 9 exhibited a dramatic increase in Stokes shift of ~100 nm compared to the acridone core. Both also displayed a sensitivity to pH, manifested as a decrease in emission intensity with decreasing pH, compared to the unmodified core whose emission is insensitive to pH changes in the 2-10 range. Finally, the benzo modified acridone analogs all feature a nearly identical emission profile ~100 nm red shifted from the emission of the parent acridone compound. The combination of benzo and fluoro or methoxy substitution further red-shifts the absorption spectrum, but does not appreciably change the emission spectrum.
Figure 3.3 Absorption and emission spectra of acridone derivatives. Spectra determined in 1:1 acetonitrile/PBS, pH 7.4. Spectra are shown normalized to enable comparison of changes in absorption and emission maxima.

Electronic structure calculations

(Note: Calculations were performed by Professor George A. Petersson)

In order to understand the nature of the substituent effects on acridone fluorescence, we performed ab initio electronic structure calculations of the absorption and emission spectra for all of the core derivatives. The calculations employed the APF-D density
functional as implemented in the Gaussian16™ suite of programs with the 6-311+G(2d,p) basis set which has been recommended for calculations of fluorescence spectra.\textsuperscript{99-101} The close agreement between the calculated (Calc.) and experimental (Obs.) spectra for the parent species, acridone, clearly justifies this choice. (\textbf{Figure 3.4}) For microscopy and spectroscopy applications, the useful transition is not the expected intense C=O \( n \rightarrow \pi^* \) excitation at 253/255 nm (Calc./Obs.). Rather, it is the weaker \( \pi \rightarrow \pi^* \) excitation with peaks at 366/376 nm (Calc./Obs.) and 384/402 nm (Calc./Obs.). The emission spectrum shows a \( \sim 10 \) nm Stokes shift with \( \pi^* \rightarrow \pi \) emission peaks at 396/412 nm (Calc./Obs.) and 418/436 nm (Calc./Obs.). We note that although the agreement of the calculated and observed extinction coefficients is excellent for the \( n \rightarrow \pi^* \) transition, the calculations overestimate the intensity of the \( \pi \rightarrow \pi^* \) absorption. However, what is most important is that throughout the series of acridone derivatives, the wavelengths and intensities of the \( \pi \rightarrow \pi^* \) absorption and emission track very well with the experimental data, showing that the calculations can clearly predict acridone derivative properties to guide synthetic efforts. (\textbf{Table 3.1})
Figure 3.4 Calculated acridone spectra. Top Left: Experimental and calculated, Franck-Condon corrected spectra, with vertical transitions used to determine the calculated spectrum shown. Bottom Left: Primary molecular orbitals (MOs) involved in the $n \rightarrow \pi^*$ ($\sim 250$ nm) and $\pi \rightarrow \pi^*$ ($\sim 380$ nm) transitions of acridone. Right: Jablonski diagram showing strategy for using frozen ground or excited state solvation in geometry optimization and frequency calculations to determine electronic spectra.
For acridone, we performed more sophisticated calculations, combining Franck-Condon integrals from vibrational calculations to generate spectra representing acridone absorption and emission spectra in aqueous solution. Bonding in excited states is generally not as strong as in ground states, which reduces the zero-point-energy (ZPE) from molecular vibrations. For example, the ZPE for the lowest-lying $\pi \rightarrow \pi^*$ excited state of acridone (115.76 kcal/mol) is 2.21 kcal/mol less than the ZPE for the ground state (117.97 kcal/mol). This reduces the vertical excitation energy to 78.20 kcal/mol ($\lambda_{\text{ex}} = 366$ nm). The electronically excited state then relaxes through internal conversion to its vibrational ground state, lowering the energy by 3.85 kcal/mol. This is followed by solvent relaxation.

### Table 3.1 Calculated and observed photophysical parameters of acridone derivatives.\(^a\)

<table>
<thead>
<tr>
<th>Compound</th>
<th>Calculated</th>
<th></th>
<th>Observed</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\lambda_{\text{ex/e}}$</td>
<td>$\lambda_{\text{em/Int.}}$</td>
<td>$\lambda_{\text{ex/e}}$</td>
<td>$\lambda_{\text{em/Int.}}$</td>
<td></td>
</tr>
<tr>
<td>Acd 5</td>
<td>383/2.43</td>
<td>411/1.00</td>
<td>398/0.70</td>
<td>412/1.00</td>
<td></td>
</tr>
<tr>
<td>2-NO$_2$ 6</td>
<td>408/1.87</td>
<td>-</td>
<td>402/0.37</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>4-NO$_2$ 7</td>
<td>464/2.06</td>
<td>-</td>
<td>434/0.60</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>2-NH$_2$ 8</td>
<td>419/0.47</td>
<td>508/0.21</td>
<td>420/0.52</td>
<td>527/0.16</td>
<td></td>
</tr>
<tr>
<td>4-NH$_2$ 9</td>
<td>400/0.61</td>
<td>492/0.59</td>
<td>400/0.59</td>
<td>540/0.01</td>
<td></td>
</tr>
<tr>
<td>4-F 15</td>
<td>365/1.19</td>
<td>397/0.48</td>
<td>378/0.51</td>
<td>412/0.65</td>
<td></td>
</tr>
<tr>
<td>2-OMe S1</td>
<td>400/0.95</td>
<td>444/0.47</td>
<td>396/0.82</td>
<td>447/0.63</td>
<td></td>
</tr>
<tr>
<td>4-OMe 16</td>
<td>375/0.62</td>
<td>421/0.35</td>
<td>384/0.69</td>
<td>431/0.37</td>
<td></td>
</tr>
<tr>
<td>Bz 18</td>
<td>446/0.81</td>
<td>495/0.31</td>
<td>438/0.58</td>
<td>507/0.42</td>
<td></td>
</tr>
<tr>
<td>4-F-Bz 19</td>
<td>446/0.61</td>
<td>496/0.24</td>
<td>448/0.54</td>
<td>508/0.21</td>
<td></td>
</tr>
<tr>
<td>4-OMe-Bz 20</td>
<td>456/0.67</td>
<td>512/0.25</td>
<td>454/0.61</td>
<td>512/0.14</td>
<td></td>
</tr>
</tbody>
</table>

\(a\)Lowest energy absorption and highest intensity emission data reported here. A full list of absorption and emission peaks are found in Table 3.3 in experimental method section.

\(b\)Extinction coefficients (ε) reported as $10^4$ M$^{-1}$cm$^{-1}$.

\(c\)Emission intensity (Int.) normalized the intensity of the highest emission peak of acridone for both calculated and observed spectra.
which lowers the energy of the excited state by an additional 1.19 kcal/mol, and raises the ground state energy by 0.94 kcal/mol (Figure 3.4). The vertical emission (68.61 kcal/mol or \( \lambda_{\text{em}} = 418 \) nm) is followed by relaxation through internal conversion to the vibrational ground state, lowering the energy by 3.61 kcal/mol. Solvent relaxation then lowers the energy of the ground state by an additional 0.94 kcal/mol. We found that these Franck-Condon corrected spectra have excellent agreement with the observed acridone spectra, particularly for the features in the \( \pi \rightarrow \pi^* \) absorption and emission peaks. We note that previous computational studies of acridone by Oshima, accompanied by gas phase electronic spectra measurements, concluded that the various vibronic modes that they observed for the \( \pi \rightarrow \pi^* \) absorption and emission were the result of different hydration clusters.\textsuperscript{102-104} While this explanation may be valid for gas phase measurements, solution phase measurements show a similar spectral shape for acridone absorption and emission in polar, hydrogen-bonding solvents like water, and in non-polar solvents like THF.\textsuperscript{57} Thus, the explanation arising from our Franck-Condon corrected calculations (determined with continuum solvation), that the multiple peaks at \( \sim 380 \) nm arise from internal acridone vibrational modes, seems to fit better with solution phase data.

**Synthesis of aminoacridonylalanine for protein studies**

We chose to generate the amino acid form of 2-aminooacidone (8) since its emission was as red-shifted as any of the benzo derivatives, but its smaller size and greater polarity make it more favorable for eventual genetic incorporation. While the absorption and
emission spectra of 2- and 4-aminoacridone are relatively similar, the 2-amino isomer was chosen because it is brighter and our modeling studies (below) suggest that it is likely to be genetically incorporable. Since the 2-aminoacridone core shows a larger Stokes shift than the acridone core with comparable spectral overlap and selective excitability (i.e., the maximum in Mcm absorbance coincides with a minimum in Acd or Aad absorbance) it can function as a superior FRET partner for Mcm. Our spectral overlap calculations indicate that acridonylalanine (Acd, 2) and aminoacridonylalanine (Aad, 4) have similar $R_0$ values (Acd: 25.3 Å, Aad: 26.5 Å), but the emission of Aad at 530 nm will be easier to monitor in ratiometric FRET experiments and no deconvolutions with Mcm emission will be necessary (Figure 3.5). Although 2-aminoacridone is commonly used in glycan experiments, and acyl 2-aminoacidone peptides have been used previously in protease sensors, to our knowledge, amino acid forms such as Aad have not been made previously or used in FRET applications.92,105
Figure 3.5 Aad is a Superior FRET Acceptor for Mcm. Top: Absorbance (solid lines) and fluorescence emission (dashed lines) of Mcm (purple), Acd (blue), and Aad (orange) with spectral overlap shaded. Both Acd and Aad have low absorbance at the excitation maximum of Mcm (325 nm) and comparable spectral overlap (shaded areas), but Aad emission can more easily be distinguished from Mcm emission. Bottom Left: Normalized changes in FRET ratio (380/480 nm for Acd, 380/530 nm for Aad) of 5 µM Xxx-Leu-Leu-Lys-Ala-Ala-Ala-Mcm (Xxx = Acd or Aad) peptides upon rapid mixing with 1 µM trypsin protease. Data points from three trials shown to demonstrate reproducibility, with average shown as a smooth curve. Bottom Right: Fluorescence emission spectra, with excitation at 325 nm, of Acd or Aad peptides before and after proteolysis. Spectra normalized based on area under the curve.
Acd (2) was synthesized as we have previously reported. Nitration of Acd was carried out similarly to nitration of acridone, affording a single nitro isomer in 39% isolated yield (>90% conversion based on integration of the chromatogram). The nitration reaction mixture was basified and sodium sulfide added to perform the reduction to give Aad (4) in 86% yield. (see experimental details section) The position of the amino group was confirmed by NMR and by comparison of the absorption and emission spectra to the spectra of 2- and 4-aminoacridone (8 and 9). Protected forms of Acd and Aad were generated for solid phase peptide synthesis, and the peptides Xxx-Leu-Leu-Lys-Ala-Ala-Ala-Mcm (Xxx = Acd or Aad) were synthesized. The peptide sequence was chosen as a protease substrate based on previous studies in our laboratory so that cleavage should result in a decrease in FRET.106 Cleavage by the proteases trypsin and cathepsin B was tested by exciting at 325 nm and monitoring emission at 380 nm (Mcm) and either 480 nm (Acd) or 530 nm (Aad). After confirming that trypsin proteolysis proceeded at a reasonable rate in a multi-well plate assay, we measured cleavage rates in a stopped-flow fluorometer. As one can see in Figure 3.5, while both peptides give usable data, the change in FRET ratio is two-fold greater for the Aad peptide than the Acd peptide.

Site-specific genetic incorporation of Aad requires a tRNA synthetase, or RS, that efficiently and selectively adenylates Aad and then charges an orthogonal tRNA with this activated Aad ester.17 As a first step toward achieving an AadRS, we tested Acd and Aad adenylation by AcdRS1 (also referred to as clone G2), a permissive Methanocaldococcus jannaschii TyrRS derivative that incorporates Acd, and also incorporates N-phenyl p-aminophenylanine (Npf) and 4-(2′-bromoisobutyramido) phenylalanine (Brb).46, 61, 64, 88
Using a purified form of AcdRS1 in a Malachite green assay that we have previously used to characterize the relative activity of the enzyme toward Acd and Npf, we find that Aad is activated by AcdRS1, although less efficiently than Acd.\textsuperscript{88} (Figure 3.6) These results can be understood with a simple model in which Aad is aligned with Acd in our published AcdRS1 docking model, based on the crystal structure of AcdRS1 with Brb.\textsuperscript{64, 88} As one can see in Figure 3.6, the Aad exocyclic amine makes contact with the sidechains of Val164 and Ala167. Thus, mutation of these residues to smaller or more polar residues may improve Aad charging. Generating a library of randomized mutants of the 164-167 loop, followed by selection of the most active clones, is also likely to result in an efficient AadRS for \textit{in vivo} incorporation of Aad.
Figure 3.6 Aad Activation by AcdRS1. Left: Acd and Aad adenylation measured based on inorganic pyrophosphate (PPi) formation in an enzyme-coupled Malachite green assay. Control experiments lacking ATP show that the low levels of Aad activity are significant. Right: Aad docked in the AcdRS1 active site by alignment with Acd in an energy-minimized model generated from AcdRS1 PDB coordinates 4PBR.
Conclusions

We can draw several conclusions from our results. Firstly, as anticipated from previous literature reports, we are able to modulate the fluorescence of the acridone core through relatively simple substitutions. These substitutions, introduced through direct modification of acridone or through cross-coupling and cyclization, are compatible with eventual usage in generating Acd amino acid derivatives. Secondly, the close correlation (% differences $\lambda_{ex}$: 2.35, $\lambda_{em}$: 2.40) between our calculated and observed absorption and emission spectra give us confidence that we can predict the spectra for acridone derivatives, providing guidance for future synthetic efforts. This will be particularly valuable for targeting multiply-substituted derivatives, where the number of possibilities is geometrically larger and the synthesis will be more challenging. Finally, Aad appears to be superior to Acd for several fluorescence applications. As demonstrated here, the red-shifted emission of Aad provides superior dynamic range in ratiometric FRET measurements with an Mcm donor. Furthermore, its absorption and emission wavelengths make Aad amenable to use in microscopy applications using typical filter sets for cyan fluorescent protein excitation and yellow fluorescent protein emission. Aad can also serve as a FRET donor to red wavelength dyes such as indodicarbocyanine or sulforhodamine G.\textsuperscript{107} Using the strategy outlined here, we can further improve the spectroscopic properties of Aad, such as its relatively low quantum yield. While Aad can be incorporated into peptides, and native chemical ligation can be used for incorporation into proteins if desired, our ultimate goal is genetic incorporation of Acd derivatives. Our initial in vitro activity results for Aad are very promising, and computational models based on the crystal structure
of AcdRS1 indicate that we may be able to obtain an efficient AadRS with just one or two mutations. These rationally-designed efforts, as well as screening of random mutant libraries, are currently under way. If we can successfully generate an AadRS, this would enable the furthest red-shifted fluorescent amino acid to be genetically incorporated in living cells.

**Experimental Methods**

**General Information**

**Materials**

L-Tyrosine, thionyl chloride, di-tert-butyl dicarbonate (Boc₂O), methyl 2-amino benzoate, and phenylsepharose CL-4B resin were purchased from Sigma-Aldrich (St. Louis, MO, USA). \(N\)-Phenyl-bis(trifluoromethane sulfonimide) was purchased from Oakwood Chemical (West Columbia, SC, USA). Acridin-9(10\(H\))-one, 2-aminoacridin-9(10\(H\))-one, 2-methoxyacridin-9(10\(H\))-one, and 3-Methyl-2-aminonaphthoic acid were purchased from Sigma-Aldrich. \(N^\alpha\)-Fmoc-7-methoxycoumarin-4-yl-alanine (Mcm) was purchased from Bachem (Torrence, CA, USA). All other Fmoc protected amino acids and peptide coupling reagents, as well as Rink amide resin, were purchased from EMD Millipore (Billerica, MA, USA). Piperidine was purchased from American Bioanalytical (Natick, MA, USA). Milli-Q filtered (18 MΩ) water was used for all solutions (Millipore; Billerica, MA, USA). All other reagents and solvents were purchased from Fisher Scientific or Sigma-Aldrich unless otherwise specified.
Instruments

Low resolution electrospray ionization mass spectra (ESI-LRMS) were obtained on a Waters Acquity Ultra Performance LC connected to a single quadrupole detector (SQD) mass spectrometer (Milford, MA, USA). Accurate mass measurement analyses were conducted on either a Waters GCT Premier, time-of-flight, GCMS with electron ionization (EI), or an LCT Premier XE, time-of-flight, LCMS with electrospray ionization (ESI). The signals were measured against an internal lock mass reference of perfluorotributylamine (PFTBA) for EI-GCMS, and leucine enkephalin for ESI-LCMS. UV-Vis absorption spectra were acquired on a Hewlett-Packard 8452A diode array spectrophotometer (currently Agilent Technologies; Santa Clara, CA, USA). Fluorescence spectra were acquired on a Photon Technologies International (PTI) QuantaMaster40 fluorometer (currently Horiba Scientific, Edison, NJ, USA). Nuclear magnetic resonance (NMR) spectra were obtained on a Bruker DRX 500 MHz instrument (Billerica, MA, USA). Analytical HPLC was performed on an Agilent 1100 Series HPLC system. Preparative HPLC was performed on a Varian Prostar HPLC system (currently Agilent Technologies). HPLC columns were purchased from W. R. Grace & Company (Columbia, MD, USA). Matrix-assisted laser desorption/ionization mass spectra (MALDI-MS) were collected with a Bruker Ultraflex III MALDI-TOF/TOF mass spectrometer (Billerica, MA). Stopped flow fluorescence measurements were made on a KinTek AutoSF-120 instrument (Snow Shoe, PA, USA).
Chemical Synthesis

2-Nitroacridin-9(10H)-one (6) and 4-nitroacridin-9(10H)-one (7)

Acridone (5, 50 mg, 0.25 mmol) in 36% acetic acid (250 µL) was stirred vigorously at 50 °C. Then glacial acetic acid (500 µL) and 70% nitric acid (250 µL) were added (Table S1, entry 2). The reaction mixture was stirred for 1 h at 50 °C. The mixture was then poured into ice-water (10 mL), neutralized with K₂CO₃, and extracted with EtOAc three times, resulting in a yellow precipitate in the aqueous layer. The yellow solid in the aqueous layer was filtered off, washed thoroughly with water and EtOAc, and dried overnight. Dried solid was then recrystallized in boiling ethanol to give 6 (30 mg, 49% isolated yield). The organic layer was extracted, dried with anhydrous Na₂SO₄, and concentrated in vacuo. The crude mixture was purified by flash column chromatography to give 7 (6 mg, 10% isolated yield).

2-Nitroacridin-9(10H)-one (6) HRMS (ESI) calculated for C₁₃H₉N₂O₃⁺ is 241.0613, [M+H]⁺ found 241.0627. ¹H NMR (500 MHz, DMSO-d₆) δH: 12.34 (s, 1H), 8.98 (d, 1H, J = 2.6 Hz), 8.42 (dd, 1H, J = 9.5, 2.7 Hz), 8.23 (d, 1H, J = 8.2 Hz), 7.79 (t, 1H, J = 7.5 Hz), 7.60 (dd, 2H, J = 25.3, 9.0 Hz), 7.34 (t, 1H, J = 7.3 Hz) ¹³C NMR (100 MHz, DMSO-d₆) δC: 177.1, 145.7, 141.8, 141.1, 135.0, 127.8, 126.7, 123.7, 123.4, 121.5, 120.0, 119.8, 119.2
4-Nitroacridin-9(10H)-one (7) HRMS (ESI) calculated for C$_{13}$H$_{9}$N$_{2}$O$_{3}$ $^+$ is 241.0613 [M+H]$^+$ found 241.0625. $^1$H NMR (500 MHz, DMSO-$d_6$) $\delta$: 11.47 (s, 1H), 8.69 – 8.66 (m, 2H), 8.22 (dd, 1H, $J$ = 8.1, 1.5 Hz), 8.07 (dt, 1H, $J$ = 8.4, 0.7 Hz), 7.81 (ddd, 1H, $J$ = 8.5, 7.0, 1.6 Hz), 7.44 – 7.37 (m, 2H). $^{13}$C NMR (500 MHz, DMSO-$d_6$) $\delta$: 176.5, 140.8, 136.0, 135.5, 135.3, 135.1, 132.2, 126.4, 123.9, 123.7, 121.3, 120.6, 119.9

Table 3.2 Screening of acridone nitration conditions

<table>
<thead>
<tr>
<th>Entry</th>
<th>AcOH (%)</th>
<th>HNO$_3$ (%)</th>
<th>5 (%)</th>
<th>6 (%)</th>
<th>7 (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>75</td>
<td>17.5</td>
<td>5</td>
<td>74</td>
<td>21</td>
</tr>
<tr>
<td>2</td>
<td>59</td>
<td>17.5</td>
<td>10</td>
<td>61</td>
<td>29</td>
</tr>
<tr>
<td>3</td>
<td>43</td>
<td>17.5</td>
<td>7</td>
<td>68</td>
<td>25</td>
</tr>
<tr>
<td>4</td>
<td>27</td>
<td>17.5</td>
<td>98</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>75</td>
<td>12.5</td>
<td>11</td>
<td>50</td>
<td>38</td>
</tr>
<tr>
<td>6</td>
<td>59</td>
<td>12.5</td>
<td>10</td>
<td>55</td>
<td>35</td>
</tr>
<tr>
<td>7</td>
<td>43</td>
<td>12.5</td>
<td>87</td>
<td>10</td>
<td>3</td>
</tr>
<tr>
<td>8</td>
<td>27</td>
<td>12.5</td>
<td>98</td>
<td>1.5</td>
<td>0.5</td>
</tr>
</tbody>
</table>

Each reaction was performed on 50 mg (0.25 mmol) 5 following the general experimental procedure described above. Yields calculated based on the area of corresponding HPLC chromatogram peak (100% = area of 5 + 6 + 7).

(Note: The screening of nitration conditions was done by Joomyung V. Jun)
Figure 3.7 Screening of acridone nitration conditions.
2-Aminoacridin-9(10H)-one (8) and 4-aminoacridin-9(10H)-one (9)

Reduction of 2-nitroacridone (6) to form 2-aminoacridone (8) was performed to test conditions for the reduction of nitroacridonylalanine (7) to form aminoacridonylalanine (4).

Condition 1 (without NaOH): To a stirred suspension of 2-nitroacridin-9(10H)-one (6, 4.0 mg, 0.016 mmol) in ethanol (1.3 mL), a solution of sodium sulfide nonahydrate (110 mg, 0.46 mmol) in water (2.0 mL) was added. The reaction was refluxed for 20 h. Resulting analytical HPLC data of the crude, showed that there was no reaction.

Condition 2 (with NaOH): Subsequent reduction to 2-aminoacridin-9(10H)-one (8) was tried under alkaline condition (condition 2). To a stirred suspension of 2-aminoacridin-9(10H)-one (8), a premixed solution of sodium sulfide nonahydrate (110 mg, 0.46 mmol) and sodium hydroxide (26 mg, 0.65 mmol) in water (2.0 mL) was added. After refluxing the reaction mixture for 20 h, 2-nitroacridone (6) was almost completely reduced to 2-aminoacridone (8, 84 % HPLC yield) with a trace amount of starting material.
4-nitroacridin-9(10H)-one (7, 4.0 mg, 0.016 mmol) was also reduced under basic solution (condition 2) to give 4-aminoacridin-9(10H)-one (9, 80 % HPLC yield). HPLC yields were obtained by using 2-aminoacridone or 4-aminoacridone as a standard.

Figure 3.8. Screening of 2- and 4-Nitroacridone reduction conditions.

(Note: The screening of 2- and 4-nitroacridone reduction conditions was done by Joomyung V. Jun)
4-4-aminoacridin-9(10H)-one (9)

4-Nitroacridin-9(10H)-one (7.0.040 g) was dissolved in 10 mL MeOH and 5 mL CH₂Cl₂. Pd/C (20 mg) was added to the flask and it was capped with rubber septum. The reaction flask was evacuated and back-filled with H₂ gas from a balloon three times. The reaction was stirred at room temperature overnight. Then, the crude reaction was filtered through Celite to remove the Pd/C. The Celite was washed with CH₂Cl₂. The filtrate was concentrated and purified by flash column chromatography (5:4:1 CH₂Cl₂/EtOAc/MeOH). 4-aminoacridin-9(10H)-one (9) was obtained in 29% isolated yield (10 mg). HRMS (ESI) m/z calculated for C₁₃H₁₃NO₂⁺ [M+H]⁺ is 211.0871, [M+H]⁺ found 211.0861. ¹H NMR (500 MHz, DMSO-d₆) δH: 10.69 (1 H, s), 8.16 (1 H, d, J 8.1), 7.69 – 7.63 (2 H, m), 7.50 (1 H, p, J 3.6), 7.19 (1 H, tt, J 6.7, 2.9), 6.98 – 6.93 (2 H, m), 5.56 (2 H, s). ¹³C NMR (126 MHz, DMSO-d₆) δC: 177.36, 141.03, 137.26, 133.35, 130.04, 126.27, 121.85, 121.80, 121.22, 120.38, 117.95, 116.23, 113.81.
4-Fluoroacridin-9(10H)-one (15)

1-Bromo-2-fluorobenzene (11, 219 µL, 2.0 mmol) was added to 10 mL degassed toluene in a dried round-bottom flask, followed by methyl 2-aminobenzoate (311 µL, 2.4 mmol). The solution was degassed with Ar for 5 min. Then palladium(II) acetate (0.018 g, 0.08 mmol), and racemic 2,2'- bis(diphenyl-phosphino)-1,1'-binaphthyl (rac-BINAP, 0.100 g, 0.16 mmol) were added to the flask. Cesium carbonate (0.912 g, 2.80 mmol) was ground and added to the flask. The flask was then fitted with a reflux condenser and heated to 115 °C for 24 h. After the solution was allowed to cool to ambient temperature, the contents were filtered through a short plug of silica gel using CH₂Cl₂ to transfer the material to the silica (50 mL), and then ethyl acetate (100 mL) was used to elute the product. The clarified solution was then concentrated under reduced pressure and further purified by flash column chromatography (9:1 Hexane/EtOAc) to afford 0.4211 g (86%) of methyl 2-((2-fluorophenyl)amino)benzoate (13).

Methyl 2-((2-fluorophenyl)amino)benzoate (13) was dissolved in 4 mL THF in a flask and then a LiOH solution (0.206 g in H₂O 2 mL) was added. The reaction mixture was refluxed at 80 °C for 1 h and then cooled to room temperature. The organic solvent was evaporated and the pH of the aqueous phase was adjusted to 6.0 with 6 M HCl. The
precipitate was collected and dried under vacuum overnight giving 2-((2-
fluorophenyl)amino)benzoic acid (0.2056 g, 52%).

7.0 g of polyphosphoric acid (PPA) was added to a round-bottom flask with a stir
bar and heated to 135 °C in an oil bath. 2-((2-Fluorophenyl)amino)benzoic acid (0.2056
g, 0.89 mmol) was added to the flask and stirred for 16 h. Then, 10 mL water was slowly
added and the reaction allowed to cool to 60 °C. After stirring for 1 h at 60 °C, the reaction
was cooled to ambient temperature. Insoluble impurities were removed by vacuum
filtration and the pH of the clarified solution was adjusted to 6.5 by addition of 8 M NaOH.
Then, the solution was cooled to 4 °C for precipitation for about 4 h. The yellow precipitate
was collected by vacuum filtration and dried overnight, affording 0.1758 g of 4-
fluoroacridin-9(10H)-one (15, 93%). HRMS (ESI) m/z calculated for C_{13}H_{8}FNO is
213.0590, [M+H]^+ found 213.0585. ^1^H NMR (500 MHz, DMSO-d_6) δ_H: 11.64 (1 H, s),
8.20 (1 H, dd, J 8.2, 1.5), 8.01 (1 H, dd, J 8.2, 1.3), 7.78 (1 H, d, J 8.2), 7.72 (1 H, ddd, J
8.4, 6.8, 1.6), 7.64 (1 H, ddd, J 11.6, 7.9, 1.4), 7.29 – 7.24 (1 H, m), 7.19 (1 H, td, J 7.9,
4.8). ^1^C NMR (126 MHz, DMSO-d_6) δ_C: 176.49 (d, J 2.8), 151.52 (d, J 246.4), 141.11,
134.10, 130.71 (d, J 13.3), 126.31, 122.85 (d, J 1.9), 123.62 – 120.56 (m), 120.97, 120.62
(d, J 6.9), 119.92 – 116.55 (m).

4-Methoxyacridin-9(10H)-one (16)
Degassed toluene (10 mL) was added to 2-bromoanisole (12, 249 µL, 2.0 mmol) in a dried round-bottom flask, followed by methyl 2-aminobenzoate (311 µL, 2.4 mmol). The solution was degassed with Ar for 5 min. Then, palladium(II) acetate (0.018 g, 0.08 mmol), and rac-BINAP (0.100 g, 0.16 mmol) were added to the flask. Cesium carbonate (0.912 g, 2.80 mmol) was ground and added to the flask. The flask was then fitted with a reflux condenser and heated to 115 °C for 24 h. After the solution was allowed to cool to ambient temperature, the contents were filtered through a short plug of silica gel using CH₂Cl₂ to transfer the material to the silica (50 mL), and then ethyl acetate (100 mL) was used to elute the product. The clarified solution was then concentrated under reduced pressure and further purified by flash column chromatography (9:1 Hexane/EtOAc) afforded methyl 2-((2-methoxyphenyl)amino) benzoate (14, 0.400 g, 78%).

A solution of 13.5 M sulfuric acid (5 mL) was added to a flask containing methyl 2-((2-methoxyphenyl)amino) benzoate (14, 0.40 g, 1.55 mmol). The flask was then fitted with a reflux condenser and heated to 115 °C for 3 h in an oil bath. 15 mL water was then added to the flask and allowed to stir for 15 min. The reaction was then removed from the hot oil bath and allowed to cool down. The precipitate formed was then collected with a Buchner funnel. The filtrate was extracted with CH₂Cl₂ (3 x 40 mL). The organic phase was dried with MgSO₄ and evaporated. The crude product was then purified by flash column chromatography (99.5:0.5 CH₂Cl₂/EtOH) to afford 0.334 g (96%) of 4-methoxyacridin-9(10H)-one (16). HRMS (ESI) calculated for C₁₄H₁₁NO₂ is 225.0790, [M+H]⁺ found 225.0808. ¹H NMR (500 MHz, DMSO-d₆) δH: 11.18 (1 H, s), 8.18 (1 H, dd, J 8.1, 1.7), 7.90 (1 H, d, J 8.5), 7.77 (1 H, dd, J 8.2, 1.3), 7.67 (1 H, ddd, J 8.5, 6.8, 1.7),
7.30 (1 H, dd, J 7.9, 1.4), 7.22 (1 H, ddq, J 8.1, 6.9, 1.1), 7.16 (1 H, td, J 8.0, 2.1), 4.01 (3 H, d, J 2.1). $^1$H NMR (126 MHz, DMSO-$d_6$) $\delta$H: 176.96, 148.21, 141.04, 133.47, 132.23, 126.11, 121.63, 121.49, 120.92, 118.68, 117.51, 112.74, 56.62.

$^1$C NMR (126 MHz, CDCl$_3$) $\delta$C: 168.20, 145.91, 140.97, 137.28, 133.36, 129.14, 128.69, 125.97, 125.08, 122.42, 114.67, 109.89, 51.83.

**Methyl 3-amino-2-naphthoate (17)**

![Chemical reaction](image)

3-Amino-2-naphthoic acid 1.0010 g (5.35 mmol) was added to an oven-dried flask and dissolved in methanol (15 mL). Concentrated H$_2$SO$_4$ 5 mL was then gradually added. The flask was connected to a condenser, and the reaction was refluxed at 100 °C overnight. After cooling the reaction, saturated NaCO$_3$ was added to the crude mixture, and extracted with CH$_2$Cl$_2$ (3 x 50mL). The combined organic layer was dried with MgSO$_4$, filtered, evaporated, and dried under vacuum overnight. Methyl 3-amino-2-naphthoate (17) was obtained as a white solid (0.8085 g, 75%). HRMS (ESI) calculated for C$_{12}$H$_{11}$NO$_2^+$ is 201.0790, [M+H]$^+$ found 201.0785. $^1$H NMR (500 MHz, CDCl$_3$) $\delta$H: 8.48 (1 H, d, J 2.3), 7.72 – 7.66 (1 H, m), 7.53 – 7.46 (1 H, m), 7.38 (1 H, ddt, J 8.6, 5.2, 1.7), 7.20 – 7.13 (1 H, m), 6.94 (1 H, d, J 2.3), 5.55 (2 H, s), 3.93 (3 H, s). $^{13}$C NMR (126 MHz, CDCl$_3$) $\delta$C: 168.20, 145.91, 140.97, 137.28, 133.36, 129.14, 128.69, 125.97, 125.08, 122.42, 114.67, 109.89, 51.83.
Methyl 3-amino-2-naphthoate (17) (0.241 g, 1.20 mmol) was dissolved in 10 mL degassed toluene in a dried sealed tube and bromobenzene (S2, 105 µL, 1.0 mmol) was added. Then palladium(II) acetate (0.009 g, 0.04 mmol), and rac-BINAP (0.050 g, 0.08 mmol) were added to the tube. Cesium carbonate (0.456 g, 1.40 mmol) was added to the tube. The reaction mixture was capped and heated to 115 °C for 24 h. After the solution was allowed to cool to ambient temperature, the contents were filtered through a short plug of silica gel using CH₂Cl₂ to transfer the material to the silica and then ethyl acetate (100 mL) was used to elute the product. The clarified solution was then concentrated under reduced pressure and further purified by flash column chromatography (9:1 Hexane/EtOAc) to afford 0.2311 g (83%) of methyl 3-(phenyl amino)-2-naphthoate (S3).

Methyl 3-(phenylamino)-2-naphthoate was dissolved in THF (4 mL) and MeOH (2 mL) in a 25 mL flask and then a LiOH solution (0.0527 g in 2 mL H₂O) was added. The reaction mixture was refluxed at 90 °C for 1 h and then cooled to room temperature. The organic solvent was evaporated, and the pH of the aqueous phase was adjusted to 6.0 with 6 M HCl. The precipitate was collected and dried under vacuum overnight, giving 3-(phenylamino)-2-naphthoic acid (0.1131 g, 98%).

6.56 g PPA was added to a 25 mL round-bottom flask with a stir bar. The flask was heated to 125 °C in an oil bath. Then 3-(phenylamino)-2-naphthoic acid (0.1131 g,
0.43 mmol) was added to the flask and it was stirred for 2 h. Water (10 mL) was slowly added and the reaction allowed to cool to 60 °C. After stirring for 1 h, the reaction was cooled to ambient temperature. Insoluble impurities were removed by vacuum filtration and the pH of the clarified solution was adjusted to 7.0 by addition of 10 M NaOH. The yellow precipitate was collected by vacuum filtration and dried overnight affording 0.090 g of benzol[b]acridin-12(5H)-one (18, 85%). HRMS (ESI) calculated for C_{17}H_{11}NO^+ is 245.0841, [M+H]^+ found 245.0863. ^1H NMR (500 MHz, Methanol-\textit{d}_4) \delta H: 8.97 (1 H, s), 8.34 (1 H, dd, J 8.1, 1.5), 8.05 (1 H, d, J 8.5), 7.91 – 7.88 (2 H, m), 7.72 (1 H, ddd, J 8.6, 7.0, 1.6), 7.55 (1 H, ddd, J 8.3, 6.5, 1.2), 7.48 (1 H, d, J 8.3), 7.40 (1 H, ddd, J 8.0, 6.6, 1.2), 7.22 (1 H, ddd, J 8.1, 6.9, 1.1). ^13C NMR (126 MHz, DMSO-\textit{d}_6) \delta C: 178.20, 142.18, 138.90, 135.86, 134.21, 129.77, 128.53, 127.85, 127.37, 126.72, 126.56, 124.84, 121.42, 120.28, 118.4, 117.2, 112.5.

4-Fluorobenzo[b]acridin-12(5H)-one (19)

Methyl 3-amino-2-naphthoate (17) (0.241 g, 1.20 mmol) was dissolved in 10 mL degassed toluene in a dried sealed tube, followed by addition of 1-bromo-2-fluorobenzene (11) (109 µL, 1.0 mmol). Then palladium(II) acetate (0.009 g, 0.04 mmol), and rac-BINAP
(0.050 g, 0.08 mmol) were added to the tube. Cesium carbonate (0.456 g, 1.40 mmol) was added to the tube. The reaction mixture was capped and heated to 115 °C for 24 h. After the solution was allowed to cool to ambient temperature, the contents were filtered through a short plug of silica gel using CH₂Cl₂ to transfer the material to the silica and then ethyl acetate (100 mL) was used to elute the product. The clarified solution was then concentrated under reduced pressure and further purified by flash column chromatography (9:1 Hexane/EtOAc) to afford 0.2648 g (90%) of 3-((2-fluorophenyl)amino)-2-naphthoate (S4).

Methyl 3-(phenylamino)-2-naphthoate was dissolved in THF (4 mL) and MeOH (2 mL) in a 25 mL flask and then a LiOH solution was added (0.0527 g in 2 mL H₂O). The reaction mixture was refluxed at 90 °C for 1 h and then cooled to room temperature. The organic solvent was evaporated and the pH of the aqueous phase was adjusted to 6.0 with 6 M HCl. The precipitate was collected and dried under vacuum overnight giving 3-(phenylamino)-2-naphthoic acid 0.1131 g (98%).

PPA (6.56 g) was added to a 25 mL round-bottom flask with a stir bar. The flask was heated to 125 °C in an oil bath. 3-(phenylamino)-2-naphthoic acid 0.1131 g (0.43 mmol) was added to the flask and stirred for 2 h. Then, 10 mL water was slowly added and the reaction allowed to cool to 60 °C. After stirring for 1 h, the reaction was cooled to ambient temperature. Insoluble impurities were removed by vacuum filtration and the pH of the clarified solution was adjusted to 7.0 by addition of 10 M NaOH. The yellow precipitate was collected by vacuum filtration and dried overnight affording 0.090 g of benzo[b]acridin-12(5H)-one (19, 85 %). HRMS (EI) calculated for C₁₇H₁₀FNO is
263.0746, found 263.0753. $^1$H NMR (500 MHz, DMSO-$d_6$) $\delta$: 11.52 (1 H, s), 8.91 (1 H, s), 8.18 (1 H, s), 8.14 (1 H, d, $J$ 8.3), 8.04 (1 H, d, $J$ 8.1), 7.93 (1 H, d, $J$ 8.4), 7.66 (2 H, dd, $J$ 11.6, 7.7), 7.58 (1 H, dd, $J$ 8.5, 6.6), 7.43 (1 H, q, $J$ 7.4, 6.2), 7.15 (1 H, td, $J$ 7.9, 4.8).

$^{13}$C NMR (126 MHz, DMSO-$d_6$) $\delta$: 177.84, 152.18, 150.23, 137.86, 136.21, 131.77, 130.05, 129.05, 128.37, 127.72, 126.96, 124.84, 122.42, 121.48, 119.73, 118.94, 113.46.

4-Methoxybenzo[b]acridin-12(5H)-one (20)

![Chemical structure](image)

Methyl 3-amino-2-naphthoate (17, 0.241 g, 1.20 mmol) was dissolved in 10 mL degassed toluene in a dried sealed tube, followed by adding 2-bromoanisole (12, 125 µL, 1.0 mmol). Then palladium(II) acetate (0.009 g, 0.04 mmol), and rac-BINAP (0.050 g, 0.08 mmol) were added to the tube. Cesium carbonate (0.456 g, 1.40 mmol) was added to the tube. The reaction mixture was capped and heated to 115 °C for 24 h. After the solution was allowed to cool to ambient temperature, the contents were filtered through a short plug of silica gel using CH$_2$Cl$_2$ to transfer the material to the silica and then ethyl acetate (100 mL) was used to elute the product. The clarified solution was then concentrated under reduced pressure and further purified by flash column chromatography (90:10 Hexane/EtOAc) to afford 0.2094 g (68%) of methyl 3-((2-methoxyphenyl)amino)-2-naphthoate (S5).
Methyl 3-((2-methoxyphenyl)amino)-2-naphthoate (S5, 0.1943 g, 0.632 mmol) was dissolved in THF (5 mL) and MeOH (3 mL) in a 25 mL flask and then a LiOH solution was added (0.0757 g in 3 mL H₂O). The reaction mixture was refluxed at 90 °C for 1 h and then cooled to room temperature. The organic solvent was evaporated and the pH of the aqueous phase was adjusted to 6.0 with 6 M HCl. The precipitate was collected and dried under vacuum overnight giving 3-((2-methoxyphenyl)amino)-2-naphthoic acid 0.1623 g (98%).

PPA (6.90 g) was added to a 25 mL round-bottom flask with a stir bar. The round-bottom flask was heated to 125 °C in an oil bath. 3-((2-methoxyphenyl)amino)-2-naphthoic acid 0.1513 g (0.516 mmol) was added to the flask and stirred for 2 h. Then 10 mL water was slowly added and the reaction allowed to cool to 60 °C. After stirring for 1 h at 60 °C, the reaction was cooled to ambient temperature. Insoluble impurities were removed by vacuum filtration and the pH of the clarified solution was adjusted to 7.0 by addition of 10 M NaOH. The yellow precipitate was collected by vacuum filtration and dried overnight affording 01261 g of 4-methoxybenzo[b]acridin-12(5H)-one (20, 89%). HRMS (EI) calculated for C₁₈H₁₃NO₂ is 275.0946, found 275.0948. ¹H NMR (500 MHz, DMSO-d₆) δ_H: 11.06 (1 H, s), 8.88 (1 H, s), 8.32 (1 H, s), 8.12 (1 H, d, J 8.3), 7.89 (1 H, d, J 8.4), 7.80 (1 H, d, J 8.0), 7.58 – 7.51 (1 H, m), 7.43 – 7.36 (1 H, m), 7.34 – 7.28 (1 H, m), 7.12 (1 H, t, J 7.9), 4.03 (3 H, s). ¹³C NMR (126 MHz, DMSO-d₆) δ_C: 178.33, 147.85, 138.00, 135.99, 133.39, 129.99, 128.73, 128.18, 127.30, 126.96, 124.56, 121.58, 120.05, 119.82, 117.84, 113.50, 56.64.
(S)-2-((tert-butoxycarbonyl)amino)-3-(9-oxo-9,10-dihydroacridin-2-yl)propanoic acid (S6)

(S)-2-amino-3-(9-oxo-9,10-dihydroacridin-2-yl) propanoic acid (AcD, 2, 0.050 g, 0.18 mmol), synthesized as previously described, was dissolved in EtOH (15 mL). $^{46}$ Na$_2$CO$_3$ (0.029 g, 0.27 mmol) and di-tert-butyl dicarbonate (Boc$_2$O, 0.059 g, 0.27 mmol) was added to the flask. The reaction was stirred at room temperature for 16 h. The reaction mixture was evaporated and quenched with 0.1 M HCl. The crude product was extracted with CH$_2$Cl$_2$ three times. The combined organic layers were washed with saturated NaCl solution, dried with MgSO$_4$, and evaporated in vacuo. (S)-2-((tert-butoxycarbonyl)amino)-3-(9-oxo-9,10-dihydroacridin-2-yl) propanoic acid, Boc-AcD, S6 was obtained (0.064 g, 93%).

(S)-2-amino-3-(7-amino-9-oxo-9,10-dihydroacridin-2-yl)propanoic acid (4)
To 180 mg (0.64 mmol) of Acd (2) in a 40 mL scintillation vial, glacial acetic acid (3 mL) and 70% nitric acid (1 mL) were added. The reaction mixture was stirred for 2 h at 50 °C. The mixture was cooled to room temperature and cold brine (20 mL) was added to precipitate (S)-2-amino-3-(7-nitro-9-oxo-9,10-dihydroacridin-2-yl)propanoic acid (S7). The crude yellow precipitate (81 mg, 39% isolated yield; >90% conversion based on integration of the HPLC chromatogram, but isolation of pure material was limiting) was reduced without further purification. 81 mg S7 (0.247 mmol, 1 equiv) was dissolved in 2 mL ethanol and 1 mL water in a 20 mL scintillation vial. In a separate vial, NaOH (333 mg, 8.3 mmol, 33 equiv) and sodium sulfide nonahydrate (500 mg, 2.08 mmol, 8.4 equiv) were mixed in 2 mL of water. The premixed solution of NaOH and sodium sulfide was added to the reaction vial and stirred overnight at 80 °C. After the reaction was complete, the crude solution was purified by ion-exchanged resin and lyophilized to give (S)-2-amino-3-(7-amino-9-oxo-9,10-dihydroacridin-2-yl)propanoic acid (S)-2-amino-3-(7-amino-9-oxo-9,10-dihydroacridin-2-yl) propanoic acid (Aad, 4) (63 mg, 86 % yield). HRMS (ESI) calculated for C_{16}H_{16}N_{3}O_{3}^{+} is 298.1192, found 298.1191. $^{1}$H NMR (500 MHz, D$_2$O) $\delta$H: 7.31 (1 H, d, $J$ 2.1), 6.99 (1 H, dd, $J$ 8.5, 2.0), 6.84 – 6.75 (2 H, m), 6.61 (2 H, dd, $J$ 19.7, 8.6), 4.72 (1 H, d, $J$ 1.1), 3.65 (1 H, dd, $J$ 7.6, 5.6), 2.84 (1 H, dd, $J$ 14.5, 5.6), 2.72 (1 H, dd, $J$ 14.5, 7.6). $^{13}$C NMR (126 MHz, D$_2$O) $\delta$C: 176.34, 160.89, 140.29, 137.79, 133.85, 133.49, 128.21, 125.25, 124.46, 119.41, 117.96, 117.55, 117.43, 107.71, 56.19, 37.25.
(S)-2-((tert-butoxycarbonyl)amino)-3-((tert-butoxycarbonyl)amino)-9-oxo-9,10-dihydro acridin-2-yl)propanoic acid (Boc₂-Aad, S8)

To 90 mg (0.32 mmol, 1 equiv) Acd (2) in a 40 mL scintillation vial, glacial acetic acid (1.5 mL) and 70% nitric acid (0.5 mL) were added. The reaction mixture was stirred for 2 h at 50 °C. The mixture was cooled down room temperature and cold water (5 mL) was added to quench the nitration reaction. For reduction, ethanol (5 mL) was added to the crude material followed by addition of premixed NaOH (512 mg, 12.8 mmol, 40 equiv) and sodium sulfide nonahydrate (769 mg, 3.2 mmol, 10 equiv) in water (2 mL). The reaction vial was stirred overnight at 80 °C. After the reaction was judged complete by chromatography analysis, the crude solution was cooled to room temperature and quenched by adding saturated NaHCO₃ (5 mL). The crude mixture was transferred to a 100 mL round bottom flask. NaHCO₃ (900 mg) was added to the reaction mixture and it was stirred at room temperature. Once the pH of the reaction mixture reached pH 8-9, Boc₂O (2.1 g, 9.6 mmol, 30 equiv) was added, and the reaction stirred at room temperature overnight. The mixture was then slightly acidified with a saturated NH₄Cl solution. The aqueous mixture was extracted with ethyl acetate three times, then the organic layer was collected and dried with Na₂SO₄. The organic layer was concentrated in vacuo and purified via silica
column chromatography (0%-20% MeOH in CH₂Cl₂) to remove excess t-butanol. The product was obtained and lyophilized to give Boc₂-Aad (S₈, 79 mg, 50% yield over three steps).

HRMS (ESI) calculated for C₂₆H₃₂N₃O₇⁺ is 498.2240, found 498.2230. ¹H NMR (500 MHz, MeOD) δH: 8.28 (1 H, s), 8.16 (1 H, s), 7.83 (1 H, d, J 8.4 Hz), 7.61 (1 H, d, J 9.1 Hz), 7.45-7.41 (2 H, m), 4.39 (1 H, br. s), 3.28 (1 H, m), 3.04 (1 H, br. s), 1.55 (9 H, s), 1.33 (9 H, s), 1.40 (tBuOH residual peak), 3.35 (MeOH solvent residual peak) ¹³C NMR (126 MHz, MeOD) δC: 179.94, 158.19, 155.92, 141.52, 138.92, 136.86, 135.30, 132.84, 128.37, 122.33, 127.87, 122.33, 121.49, 119.38, 118.82, 115.67, 80.91, 58.07, 39.33, 29.25 (31.22 residual tBuOH). See Figure 3.7.
Figure 3.9 Boc₂-Aad (S8) $^1$H and $^{13}$C NMR spectra in MeOD.
Spectroscopic Characterization of Substituted Acridones

Prior to spectroscopic characterization, stocks of all acridone and benzoacridone compounds were prepared in acetonitrile at a concentration of 200 or 300 µM. All absorbance and fluorescence measurements were taken in 1:1 CH₃CN/buffer. Buffers include citrate buffer (89.1 mM citric acid 21.8 mM Na₂HPO₄) pH 2.6, phosphate buffered saline (PBS, 137 mM NaCl, 2.7 mM KCl, 10 mM Na₂HPO₄, 2 mM KH₂PO₄) pH 7.4, and 3-(cyclohexylamino)-1-propanesulfonic acid (CAPS) buffer, pH 10.0. The absorbance and emission spectra of all compounds were acquired in 1:1 CH₃CN/PBS. The sensitivity of the absorbance and emission to pH was assessed for the parent acridone core as well as the 4-NH₂ and 2-NH₂ derivatives in 1:1 CH₃CN/citrate and 1:1 CH₃CN/CAPS. Absorbance measurements were acquired at concentrations of 75 and 7.5 µM for proper visualization of the spectral profile above and below ~300 nm (4-aminoacridone absorbance measurements were taken at 7.5 and 0.75 µM).

Fluorescence measurements were acquired under identical solution conditions at concentrations of 7.5 µM for all compounds, except for the 4-NH₂, 2-NO₂, and 4-NO₂ derivatives of Acd which were obtained at 75 µM due to a lack of brightness. Spectra were collected using excitation wavelengths matching the maximum absorbance wavelength for each analog measured. The collection window started 15 nm from the excitation wavelength and extended to 650 nm. This window was truncated to no less than a total range of 150 nm for more blue shifted chromophores. All slit widths were set to 1.5 nm and spectra were acquired with an integration time of 0.25 sec/nm. Higher signal-to-noise
spectra for the 4-NH$_2$, 2-NO$_2$, and 4-NO$_2$ derivatives were collected keeping all other setting the same but adjusting all slit widths to 3 nm.

Extinction coefficients were calculated from absorbance measurements collected on a Tecan M1000 plate reader (Mannedorf, Switzerland). Samples containing 100, 80, 60, 40 and 20 µM chromophore were prepared in 1:1 CH$_3$CN/PBS solution with a total sample volume of 100 µL. Following brief vortexing, samples were loaded into a Corning CoStar black, clear bottom, 96-well plate.

Absorbance and emission spectra are reported in Figures 3.8 – 3.13. Emission spectra are shown normalized to the acridone emission at 412 nm to approximate quantum yields. The most prominent peak positions are reported along with extinction coefficients and these relative emission intensities in Table 3.2. Calculated spectra (described below) are shown for comparison.
Table 3.3 Calculated and Observed Photophysical Parameters of Acridone Derivatives.

<table>
<thead>
<tr>
<th>Compound</th>
<th>Calculated</th>
<th>Observed</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\lambda_{ex}/\varepsilon^a$</td>
<td>$\lambda_{em}/\text{Int.}^b$</td>
</tr>
<tr>
<td>Acd 5</td>
<td>253/11.4</td>
<td>411/1.00</td>
</tr>
<tr>
<td></td>
<td>364/1.51</td>
<td>434/0.45</td>
</tr>
<tr>
<td></td>
<td>383/2.43</td>
<td>434/0.45</td>
</tr>
<tr>
<td>2-NO$_2$ 6</td>
<td>235/5.32</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>282/2.53</td>
<td>296/0.87</td>
</tr>
<tr>
<td></td>
<td>341/1.03</td>
<td>356/0.76</td>
</tr>
<tr>
<td>4-NO$_2$ 7</td>
<td>243/6.62</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>292/2.56</td>
<td>266/0.94</td>
</tr>
<tr>
<td></td>
<td>331/1.27</td>
<td>328/0.58</td>
</tr>
<tr>
<td></td>
<td>464/2.06</td>
<td>434/0.60</td>
</tr>
<tr>
<td>2-NH$_2$ 8</td>
<td>252/4.93</td>
<td>509/0.21</td>
</tr>
<tr>
<td></td>
<td>278/2.96</td>
<td>278/2.46</td>
</tr>
<tr>
<td></td>
<td>419/0.47</td>
<td>420/0.52</td>
</tr>
<tr>
<td>4-NH$_2$ 9</td>
<td>252/5.74</td>
<td>492/0.59</td>
</tr>
<tr>
<td></td>
<td>299/1.20</td>
<td>314/0.38</td>
</tr>
<tr>
<td></td>
<td>400/0.61</td>
<td>400/0.59</td>
</tr>
<tr>
<td>4-F 15</td>
<td>248/12.8</td>
<td>397/0.48</td>
</tr>
<tr>
<td></td>
<td>365/1.19</td>
<td>378/0.51</td>
</tr>
<tr>
<td>2-OMe S1</td>
<td>251/10.0</td>
<td>444/0.47</td>
</tr>
<tr>
<td></td>
<td>271/3.61</td>
<td>268/5.38</td>
</tr>
<tr>
<td></td>
<td>400/0.95</td>
<td>396/0.82</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4-OMe 16</td>
<td>249/6.62</td>
<td>421/0.35</td>
</tr>
<tr>
<td></td>
<td>375/0.62</td>
<td>384/0.69</td>
</tr>
<tr>
<td>Bz 18</td>
<td>234/1.38</td>
<td>495/0.31</td>
</tr>
<tr>
<td></td>
<td>269/21.7</td>
<td>272/12.7</td>
</tr>
<tr>
<td></td>
<td>302/1.40</td>
<td>296/2.07</td>
</tr>
<tr>
<td></td>
<td>446/0.81</td>
<td>438/0.58</td>
</tr>
<tr>
<td>4-F-Bz 19</td>
<td>268/15.3</td>
<td>496/0.24</td>
</tr>
<tr>
<td></td>
<td>446/0.61</td>
<td>448/0.54</td>
</tr>
<tr>
<td>4-OMe-Bz 20</td>
<td>272/14.0</td>
<td>512/0.25</td>
</tr>
<tr>
<td></td>
<td>456/0.67</td>
<td>454/0.61</td>
</tr>
</tbody>
</table>

$^a$Extinction coefficients (ε) reported as $10^4$ M$^{-1}$cm$^{-1}$.

$^b$Emission intensity (Int.) normalized the intensity of the highest emission peak of acridone for both calculated and observed spectra.
Figure 3.10 Acridone (5) spectra. Experimental (solid lines) and calculated (dashed lines) spectra and with vibronic transitions to/from lowest excited state (black bars) of 5 at pH 7.4, and experimental spectra at pH 2.6 and 10.0.
Figure 3.11 2-Aminoacridone (8) spectra. Experimental (solid lines) and calculated (dashed lines) spectra and with vibronic transitions to/from lowest excited state (black bars) of 8 at pH 7.4, and experimental spectra at pH 2.6 and 10.0.
Figure 3.12 4-Aminoacridone (9) spectra. Experimental (solid lines) and calculated (dashed lines) spectra and vertical transitions (black bars) of 9 at pH 7.4, and experimental spectra at pH 2.6 and 10.0.
Figure 3.13 2-Nitroacridone (6) and 4-Nitroacridone (7) spectra. Experimental (solid lines) and calculated (dashed lines) spectra and vertical transitions (black bars) of 6 and 7 at pH 7.4.
Figure 3.14 2-Fluoroadcridone (15), 2-Methoxyacridone (S1), and 4-Methoxyacridone (16) spectra. Experimental (solid lines) and calculated (dashed lines) spectra and vertical transitions (black bars) of 15, S1, and 16 at pH 7.4.
Figure 3.15 4-Benzacridone (18), 4-Fluorobenzoacridone (19), and 4-Methoxybenzoacridone (20) spectra. Experimental (solid lines) and calculated (dashed lines) spectra and vertical transitions (black bars) of 18, 19, and 20 at pH 7.4.
CHAPTER 4: Toward the Synthesis and \textit{in vivo} Incorporation of Acridone Derivatives into Proteins
Introduction

In order to improve the properties of the fluorescent amino acid acridonylalanine (Acd) and provide other amino acids with complementary properties, we have prepared a series of acridone derivatives and studied their photophysical properties.\(^\text{108}\) 2-Aminoacridone has a red-shifted emission and larger Stokes shift than acridone. Therefore, it provides a superior Förster resonance energy transfer (FRET) partner for 7-methoxycoumarin (Mcm) in fluorescence spectroscopy studies. The amino acid form, aminoacridonylalanine (Aad, 1), has also been synthesized to test applications in FRET-based proteolysis experiments and preliminary \textit{in vitro} tRNA charging activity. In these studies, Aad was synthesized through the nitration reaction of Acd, followed by reduction with Na\(_2\)S under basic conditions. However, the nitration reaction gave multiple products corresponding to the 5-nitro, 7-nitro (i.e. Aad), and 5-,7-dinitro derivatives of Acd. Since it was difficult to control the selectivity of the reaction or separate these side products, we pursued an alternate synthetic route, shown in Scheme 4.1.

![Scheme 4.1](image-url)  
\textbf{Scheme 4.1} The alternative synthesis of 2-aminoacridone (1, Aad)
Improving the Synthesis of 2-aminoacridone amino acid (Aad)

The alternate route uses the same tert-butoxycarbonyl (Boc) protected tyrosine methyl ester, triflated for cross-coupling (2), that we used to produce Acd. This compound is then subjected to a C-N cross-coupling reaction with an aniline derivative that already bears the amino group for Aad formation, followed by our usual acid-catalyzed ring-closing reaction. We utilized the Buchwald-Hartwig amination strategy to couple Boc-Tyr(OTf)-OMe (2) and tert-butyl (4-amino-3-cyanophenyl)carbamate (3). The cross-coupling product (4) contains a cyano group which can then be hydrolyzed to the carboxylic acid by H₂SO₄. The sulfuric conditions not only hydrolyze the cyano group, but also catalyze Friedel-Crafts ring-closing, and cleave both the Boc group and methyl ester to achieve the Aad product.

We took advantage of University of Pennsylvania High Through Experimentation (HTE) Center to screen Buchwald-Hartwig Pd-mediated cross-coupling amination conditions. Different palladium catalysts, bases, solvents were tested. We selected Buchwald Pd pre-catalysts as a choice of catalytic system. For these pre-catalysts, palladium has already been bound to the biarylphosphine ligands and their use ensures the efficient formation of the active catalytic species so that we do not need to vary Pd-to-ligand ratios. These catalysts are also air, moisture, and thermally-stable and exhibit good solubility in organic solvent. 3.0 equivalents of bases NaOtBu and Cs₂CO₃ were used relative to starting material (2). The reaction was stirred in either dioxane or toluene at 100 °C for 16 h. The screening results are shown in Figure 4.1.
Figure 4.1 Cross-coupling Screening results. Coupling of compounds 2 and 3 carried in 24 well plates with varying solvents, bases, and Pd pre-catalyst. Conversion determined as the ratio of the product peak and a standard peak in HPLC chromatograms. Catalysts defined in Experimental Methods section.

From these cross-coupling reaction screening results, we selected the best yielding condition which is the reaction with BrettPhos-Pd G1, using Cs₂CO₃ as base in either toluene or dioxane. We then increased the scale of the reaction to test this HTE result, followed by acid-catalyzed ring-closing reaction to obtain Aad in 32% overall yield.

Improving the Synthetase Activation of Aad

To be used for site-specific in vivo protein labeling, Aad must be charged onto an orthogonal tRNA by an aminoacyl tRNA synthetase (RS). Our initial test of aminoacyl adenylate activation of Aad by the permissive, first generation RS AcdRS1 (G2) showed that Aad was activated at 12% of the level of Acd activation. Further testing of other RS constructs showed that a second generation RS evolved for improved Acd incorporation, AcdRS2b (A9), activated Aad at 36% of the level of Acd, making it promising for in vivo
incorporation of Aad in bacteria without further modification. Unfortunately, initial tests for Aad incorporation in calmodulin and α-synuclein yielded very low protein quantities with no indication of Aad fluorescence (data not shown). Therefore, we are currently pursuing random selection of RS clones with Aad incorporation activity in collaboration with Ryan Mehl at Oregon State University as well as rationally designed mutagenesis of residues 164-167, which have been identified as potentially blocking Aad binding in the A9 active site in homology models based on the AcdRS1 (G2) crystal structure.

![Aad Activation by AcdRS2b (A9). Left: Acd and Aad adenylation measured based on inorganic pyrophosphate (PP) formation in an enzyme-coupled Malachite green assay. Right: Aad docked in the AcdRS2b active site by alignment with Acd in an energy-minimized model generated from AcdRS1 PDB coordinates 4PBR.](image)

**Figure 4.2** Aad Activation by AcdRS2b (A9). Left: Acd and Aad adenylation measured based on inorganic pyrophosphate (PP) formation in an enzyme-coupled Malachite green assay. Right: Aad docked in the AcdRS2b active site by alignment with Acd in an energy-minimized model generated from AcdRS1 PDB coordinates 4PBR.
**Additional Acridone Core Derivatives**

We have shown that the 2-aminoacridone core has valuable properties such as a red-shift emission and a large Stokes shift. However, it also suffers from a low quantum yield relative to Acd (0.16 vs. 0.95 in water). Therefore, we were interested in achieving brighter Aad analogs. It has been shown for many fluorophores that dialkylation of a pendant amino group can modulate absorption and emission spectra by sterically restricting orbital alignment and increase the quantum yield by reducing internal charge transfer that can lead to non-radiative relaxation. Recently, Lavis and coworkers have shown that this effect can be improved by using an azetidine ring to further restrict orbital alignment leading to quenching. To determine whether either of these effects could be exploited to improve the brightness of Aad, we synthesized two aminoacridone derivatives: 2-(dimethylamino)acridin-9(10H)-one (5) and 2-(azetidin-1-yl)acridin-9(10H)-one (6), shown in Figure 4.3.

![Figure 4.3 Structure of 2-(dimethylamino)acridin-9(10H)-one (5, Dad) and 2-(azetidin-1-yl)acridin-9(10H)-one (6, AzAcd)]
To synthesize the Dad (5), we have utilized the reductive amination reaction by reacting 2-aminoacridone (Aad) with formaldehyde (CH$_2$O), followed by sodium cyanoborohydride (NaBH$_3$CN) to give the dimethylated product, 2-(dimethylamino)acridin-9(10H)-one (5) along trimethylated 2-(dimethylamino)-10-methylacridin-9(10H)-one as side product, which can then be separated by column chromatography. To synthesize azetidinyl acridone (AzAc, 6), we performed Pd-mediated cross-coupling reaction between 2-bromoacridin-9(10H)-one (11) and azetidine hydrochloride as described in experimental method section. 2-Bromoacridin-9(10H)-one was previously synthesized by cross-coupling reaction of 1,4-dibromobenzene (8) and methyl 2-aminobenzoate (9) to provide product methyl 2-((4-bromophenyl)amino)benzoate (10). Then saponification with LiOH followed by Friedel–Crafts cyclization was used to provide 2-bromoacridin-9(10H)-one (11).

The photophysical properties were then characterized by measuring absorption and emission spectra and compared to 2-aminoacridone (Aad, 7) in Figure 4.4. The absorption maximum of Dad (5) and AzAc (6) are not significantly shifted from aminoacridone (7). The emission maximum of Dad (5) has red-shifted about 15 nm compared to Aad (7). However, the fluorescence intensity of Dad (5) and AzAzd (6) and 3- and 4-fold higher than Aad, respectively, at equivalent concentrations. Therefore, these two aminoacridone derivatives demonstrated that modification of the amino group can indeed improve the brightness of fAad. Thus, we will synthesize Dad and AzAc amino acid analogs for future in vivo incorporation.
Conclusions

We have further developed the set of acridone fluorophores to achieve a straightforward synthesis of Aad, which has good prospects for \textit{in vivo} incorporation, as well as alkylated Aad core analogs, one of which has slightly improved spectroscopic properties. Future studies will focus on other methods for improving the brightness of Aad as well as genetic selection experiments to achieve a highly active RS. A synthesis of an
amino acid analog of Dad and selection of an RS for its incorporation would provide a further red-shifted probe, which could function as a FRET acceptor for Acd.

**Experimental Method**

**General Information**

**Materials**

Di-tert-butyl dicarbonate (Boc₂O), 2,5-diaminobenzonitriles, 1,4-dibromobenzene azetidine hydrochloride, BrettPhos-Pd G1, Cesium carbonate were purchased from Sigma-Aldrich (St. Louis, MO, USA). N-Phenyl-bis(trifluoromethane sulfonimide) was purchased from Oakwood Chemical (West Columbia, SC, USA). Milli-Q filtered (18 MΩ) water was used for all solutions (Millipore, Billerica, MA, USA). All other reagents and solvents were purchased from Fisher Scientific or Sigma-Aldrich unless otherwise specified.

**Instruments**

Low resolution electrospray ionization mass spectra (ESI-LRMS) were obtained on a Waters Acquity Ultra Performance LC connected to a single quadrupole detector (SQD) mass spectrometer (Milford, MA, USA). Accurate mass measurement analyses were conducted on either a Waters GCT Premier, time-of-flight, GCMS with electron ionization (EI), or an LCT Premier XE, time-of-flight, LCMS with electrospray ionization (ESI). The signals were measured against an internal lock mass reference of perfluorotributylamine (PFTBA) for EI-GCMS, and leucine enkephalin for ESI-LCMS. UV-Vis absorption spectra were acquired on a Hewlett-Packard 8452A diode array spectrophotometer.
(currently Agilent Technologies; Santa Clara, CA, USA). Fluorescence spectra were acquired on a Photon Technologies International (PTI) QuantaMaster40 fluorometer (currently Horiba Scientific, Edison, NJ, USA). Nuclear magnetic resonance (NMR) spectra were obtained on a Bruker DRX 500 MHz instrument (Billerica, MA, USA).

**Chemical Synthesis**

**Synthesis of 2-(dimethylamino)acridin-9(10H)-one**

2-aminoacridone (7) 15.0 mg (0.071 mmol) was dissolved in concentrated CH₃COOH 4 mL. The p-formaldehyde (CH₂O) 21.3 mg (0.71 mmol) was added to reaction flask. Reaction mixture was stirred at room temperature for 15 min. (Color changed from yellow to dark red.) Then NaBH₃CN 17.9 mg (0.285 mmol) was added to the flask. The mixture was allowed to stir at room temperature overnight. Crude mixture was neutralized by 6M NaOH and extracted with EtOAc 3 x 20 mL. The combined organic layer was dried with Na₂SO₄ and evaporated. Then purified by flash chromatography column (EtOAC/CH₂Cl₂/MeOH = 5: 4.5: 0.5) to separate di- and tri- methylated product. The 2-(dimethylamino)acridin-9(10H)-one (5) was obtained 4 mg as yellow solid. (24 % yield)

¹H NMR (500 MHz, Methanol-d₄) δH: 8.33 (dd, J = 8.3, 1.5 Hz, 1H), 7.65 (ddd, J = 8.5,
6.9, 1.5 Hz, 1H), 7.56 (q, J = 1.7 Hz, 1H), 7.51 – 7.45 (m, 3H), 7.22 (ddd, J = 8.1, 6.8, 1.1 Hz, 1H), 2.99 (d, J = 1.4 Hz, 6H). $^{13}$C NMR (126 MHz, Methanol-$d_4$) $\delta C$: 180.86 – 174.68 (m), 146.39, 140.48, 134.16, 132.66, 125.72, 123.19, 121.24, 120.44, 119.44, 117.86, 116.78, 105.40, 40.01.

**Synthesis of 2-bromoacridin-9(10H)-one**

![Synthesis reaction diagram](image)

Degassed toluene (10 mL) was added to 1,4-Dibromobenzene (8) 0.236 g (1.0 mmol) in a dried round-bottom flask, followed by methyl 2-aminobenzoate (9) (162 µL, 1.25 mmol). The solution was degassed with Ar for 5 min. Then, palladium(II) acetate (0.025 g, 0.1 mmol), and rac-BINAP (0.093 g, 0.15 mmol) were added to the flask. Cesium carbonate (0.997 g, 3.0 mmol) was added to the flask. The flask was then fitted with a reflux condenser and heated to 60 °C for 24 h. After the solution was allowed to cool to ambient temperature, the contents were filtered through a short plug of silica gel using CH$_2$Cl$_2$ to transfer the material to the silica (20 mL), and then ethyl acetate (100 mL) was used to elute the product. The clarified solution was then concentrated under reduced pressure and further purified by flash column chromatography (9:1 Hexane/EtOAc) afforded methyl 2-((4-bromophenyl)amino) benzoate (10) 0.146 g (48% yield).
Methyl 2-((4-bromophenyl)amino)benzoate 0.0864 g (0.28 mmol) was dissolved in THF 5 mL. LiOH 0.0335 g (1.40 mmol) was dissolved in H$_2$O 4 mL and added to starting material. Reaction was stirred at room temperature overnight. Then 6M HCl was added to crude mixture until pH ~3. The reaction mixture was then extracted with CH$_2$Cl$_2$ 3 x 10 mL. The combined organic layer was dried with Na$_2$SO$_4$ and evaporated. Then the product was recrystallized by dissolving in CH$_2$Cl$_2$ 2 mL and allowed to recrystallize and dried under vacuum overnight. The 2-((4-bromophenyl)amino)benzoic acid was obtained as white crystal 0.078 mg (95 % yield). $^1$H NMR (500 MHz, Methanol-d$_4$) $\delta$: 7.96 (dt, J = 8.0, 1.5 Hz, 1H), 7.41 (dq, J = 8.7, 2.2, 1.6 Hz, 2H), 7.33 (ddt, J = 8.4, 7.0, 1.4 Hz, 1H), 7.24 – 7.18 (m, 1H), 7.15 – 7.09 (m, 2H), 6.75 (ddq, J = 8.2, 6.0, 1.1 Hz, 1H). $^{13}$C NMR (126 MHz, Methanol-d$_4$) $\delta$: 170.23, 146.93, 140.28, 133.67, 131.91, 122.66, 117.38, 114.66, 113.67, 112.74.

6.90 g PPA was added to a 25 mL round-bottom flask with a stir bar. The flask was heated to 125 °C in an oil bath. Then 2-((4-bromophenyl)amino) benzoic acid (0.0825 g, 0.28 mmol) was added to the flask and it was stirred for overnight. Water (20 mL) was slowly added and the reaction allowed to cool to 60 °C. After stirring for 30 min, the reaction was cooled to ambient temperature. Insoluble impurities were removed by vacuum filtration and the pH of the solution was adjusted to 7.0 by addition of 8 M NaOH. Then the precipitate was purified by flash column chromatography 80:20 CH$_2$Cl$_2$/EtOAc. The fractions containing product ($R_f$ of product = 0.63) were combined, evaporated and dried overnight. 40 mg of 2-bromoacridin-9(10H)-one (11) were obtained as greenish yellow solid (52 % yield).
Synthesis of 2-(azetidin-1-yl)acridin-9(10H)-one

To a solution of degassed toluene in a sealed tube, 2-bromoacridin-9(10H)-one (11, 15 mg, 0.055 mmol) was added. Then, azetidine hydrochloride (15.4 mg, 0.165 mmol), Brettphos-Pd G1 (6.4 mg, 0.008 mmol), and Cs$_2$CO$_3$ (107 mg, 0.33 mmol) were added. The reaction tube was capped and stirred at 100 °C for 16 h. After the solution was allowed to cool to ambient temperature, the contents were filtered through a short plug of silica gel using CH$_2$Cl$_2$ to transfer the material to the silica (20 mL), and then ethyl acetate (100 mL) was used to elute the product. The clarified solution was concentrated under reduced pressure and further purified by flash column chromatography (7:3 CH$_2$Cl$_2$/EtOAc) to afford 0.5 mg of 2-(azetidin-1-yl)acridin-9(10H)-one (6). $^1$H NMR (500 MHz, Methanol-d$_4$) $\delta$: 8.32 (dd, J = 8.2, 1.5 Hz, 1H), 7.66 (ddd, J = 8.4, 6.8, 1.5 Hz, 1H), 7.48 (ddd, J = 10.2, 7.9, 2.6 Hz, 2H), 7.30 (d, J = 2.7 Hz, 1H), 7.22 (ddd, J = 8.1, 6.8, 1.1 Hz, 1H), 7.11 (dd, J = 8.9, 2.7 Hz, 1H), 3.94 (t, J = 7.2 Hz, 4H), 2.43 – 2.35 (m, 2H).
HTE Screening

Cross-coupling reactions between methyl (S)-2-(((tert-butoxycarbonyl)amino)-3-(4-(((trifluoromethyl)sulfonyl)oxy)phenyl)propanoate (2, 1.0 eq) and tert-butyl (4-amino-3-cyanophenyl)carbamate (3, 1.5 eq) were screened by using 24-well plate. Experiments were set up inside a glovebox under a nitrogen atmosphere. Starting materials were dosed as a solution (0.1 M) in two different solvents (dioxane in row A and B, toluene in row C and D) into glass vials with stir bar. Then 10 mol% of Pd precatalysts including XPhos Pd G1, tBuXPhos Pd G1, RuPhos Pd G1, SPhos Pd G1, BrettPhos Pd G1, and XantPhos Pd G1 were added into column 1-6, respectively. 3.0 equivalents of base either Cs₂CO₃ (row A and C) or NaOtBu (row B and D) were then dosed. The 24-well plate was then sealed with screws and stirred for 16 h at 100 °C. To work up the reactions, the plate was opened to air and 500 µL of a solution of product standard (used as internal standard to measure UPLC yields) in acetonitrile (0.002 mol/L) was added into each vial. The plate was covered again, and the vials stirred for 10 min to ensure good homogenization. Into a separate 24-well LC block was added 700 µL of acetonitrile, followed by 25 µL of the diluted reaction mixtures. The LC block was then sealed with a silicon-rubber storage mat and mounted on an automated UPLC instrument for analysis.
Synthesis of \((S)-2\text{-amino}-3\text{-}(7\text{-amino}-9\text{-oxo}-9,10\text{-dihydroacridin}-2\text{-yl})\text{propanoic acid}\)

Methyl \((S)-2\text{-}((\text{tert}-\text{butoxycarbonyl})\text{amino})\text{-}3\text{-}(4\text{-}((\text{tert}-\text{butoxycarbonyl})\text{amino})\text{-}2\text{-}\text{cyanophenyl})\text{amino})\text{-}\text{phenyl})\text{propanoate (2)} (0.427 g, 1.0 mmol) was dissolved in degassed dioxane 5 mL in sealed tube. Then \text{tert}-\text{butyl} (4\text{-amino-3-cyanophenyl})\text{carbamate} 0.350 g (1.50 mmol), BrettPhos-Pd G1 80 mg (0.1 mmol) and cesium carbonate 0.3259 g (3.0 mmol) were added to the tube. The reaction mixture was capped and heated to 100 °C for 16 h. After the solution was allowed to cool to ambient temperature, the contents were filtered through a short plug of silica gel using CH\(_2\)Cl\(_2\) to transfer the material to the silica and then ethyl acetate (50 mL) was used to elute the product. The clarified solution was then concentrated under reduced pressure and further purified by flash column chromatography (50:50 Hexane/EtOAc) to afford 0.1642 g (33 %) of methyl \((S)-2\text{-}((\text{tert}-\text{butoxycarbonyl})\text{amino})\text{-}3\text{-}((4\text{-}((\text{tert}-\text{butoxycarbonyl})\text{amino})\text{-}2\text{-}\text{cyanophenyl})\text{amino})\text{ phenyl})\text{propanoate (4)}. \(^1\text{H} \text{NMR} (500 \text{ MHz, Chloroform-d}) \delta_H: 7.61 \text{ (s, 1H)}, 7.28 \text{ (d, J = 9.1 Hz, 1H)}, 7.12 \text{ (d, J = 9.1 Hz, 1H)}, 7.07 - 6.99 \text{ (m, 4H)}, 6.55 \text{ (s, 1H)}, 6.14 \text{ (s, 1H)}, 5.00 \text{ (d, J = 8.3 Hz, 1H)}, 4.68 - 4.42 \text{ (m, 1H)}, 3.70 \text{ (s, 3H)}, 3.09 - 2.92 \text{ (m, 2H)}, 1.48 \text{ (s, 9H)}, 1.39 \text{ (s, 9H)}. \(^{13}\text{C} \text{NMR} (126 \text{ MHz, Chloroform-d}) \delta_C: 172.25, 155.01, 152.81, 151.73, 142.68, 139.58, 130.86, 120.39,
116.54 (d, J = 124.7 Hz), 99.62, 80.39 (d, J = 126.3 Hz), 60.31, 54.41, 52.16, 37.71, 28.23, 14.10.

A solution of 13.5 M sulfuric acid (5 mL) was added to a flask containing methyl 
(S)-2-((tert-butoxycarbonyl)amino)-3-(4-((4-(tert-butoxycarbonyl)amino)-2-cyano 
phenyl)amino)phenyl)propanoate 0.1642 g (0.32 mmol). The flask was then fitted with a 
reflux condenser and heated to 115 °C for 2 days in an oil bath. 40 mL of water was then 
added to the flask and allowed to stir for 15 min. The reaction was removed from the hot 
oil bath and allowed to cool down to room temperature. 50 g of ion-exchange resin 
(Dowex® 50WX8 hydrogen form, strongly acidic cation exchange resin) was made into a 
slurry with 1.8 M aqueous H₂SO₄ and applied to a flash chromatography column. The resin 
was washed with 350 mL 1.8 M aqueous H₂SO₄, 2 L of water, 1 L of 1.5 M aqueous 
NH₄OH, and 4 L of water. Following these washes, the resin was dried by passing air 
through the column. The cooled Aad solution was then vacuum filtered on a Büchner 
funnel to remove precipitated material, and the clarified solution was applied to the washed 
and dried ion-exchange resin. The resulting resin slurry was shaken in the chromatography 
column for 5 min before the solution was drained. This solution was then reapplied to the 
dried resin and shaken for an additional 5 min. The twice-passed solution was then set 
aside. The loaded resin was washed with 4 L water before the compound of interest was 
eluted with 1.45 L of 1.5 M NH₄OH. The solution was concentrated to 50 mL by rotary 
evaporation, and then lyophilized to dryness, yielding 40 mg of Aad as a yellow powder 
42% yield. ESI-LRMS and NMR spectra matched previous reports.108
Spectroscopic Characterization of Substituted Aminoacridones

Prior to spectroscopic characterization, stocks of all acridone compounds were prepared in acetonitrile at a concentration of 200 or 300 µM. All absorbance and fluorescence measurements were taken in 1:1 CH$_3$CN/PBS buffer (PBS, 137 mM NaCl, 2.7 mM KCl, 10 mM Na$_2$HPO$_4$, 2 mM KH$_2$PO$_4$, pH 7.4). Absorbance measurements were acquired at concentrations of 75 and 7.5 µM for proper visualization of the spectral profile above and below ~300 nm.

Fluorescence measurements were acquired under identical solution conditions at concentrations of 7.5 µM for all compounds. Spectra were collected using excitation wavelengths matching the maximum absorbance wavelength for each analog measured. The collection window started 15 nm from the excitation wavelength and extended to 650 nm. This window was truncated to no less than a total range of 150 nm for more blue shifted chromophores. All slit widths were set to 1.5 nm and spectra were acquired with an integration time of 0.25 sec/nm.

Extinction coefficients were calculated from absorbance measurements collected on a Tecan M1000 plate reader (Mannedorf, Switzerland). Samples containing 100, 80, 60, 40 and 20 µM chromophore were prepared in 1:1 CH$_3$CN/PBS solution with a total sample volume of 100 µL. Following brief vortexing, samples were loaded into a Corning CoStar black, clear bottom, 96-well plate. Emission spectra are shown normalized to the acridone emission at 412 nm to approximate quantum yields.
Amino Acid Activation by Acd Aminoacyl tRNA Synthetase (AcdRS)

(These experiments were performed in collaboration with Zach Hosteteler.)

AcdRS Expression and Purification. We had previously generated recombinant Acd aminoacyl tRNA synthetase (AcdRS2b) by overexpressing polyHis-tagged AcdRS2b from a pBAD construct in arabinose auto-induction media and purifying the enzyme using both cobalt and heparin affinity columns. For our amino acid activation assay, we used an enzyme aliquot from that AcdRS2b preparation, which we had exchanged into buffer (50 mM HEPES-KOH pH 7.5, 20 mM NaCl, 10 mM MgCl$_2$, 1 mM DTT, 50% glycerol) and stored at -80 °C.

AcdRS Activity Test Using a Malachite Green Assay. As previously described, we tested the activity of AcdRS2b with either Acd or Aad using an adapted form of an enzyme-coupled, colorimetric assay. Formation of inorganic pyrophosphate (PPI) during the amino acid activation step of the reaction is coupled to the activity of inorganic pyrophosphatase (PPiase), which splits PPI into two inorganic phosphate (Pi) molecules. Under acidic conditions, free Pi reacts with Malachite green to produce a colored Malachite green phosphomolybdate complex (absorbing at 600-660 nm). Since our AcdRS2b freezer stock only retained partial activity (data not shown), we chose conditions that drove the amino acid activation reaction toward completion. We prepared reactions in triplicate containing reaction buffer (50 mM HEPES-KOH pH 7.5, 20 mM NaCl, 10 mM MgCl$_2$, 1 mM DTT), 0.2 mM ATP, 2 U/mL PPiase (NEB), 250 µM amino acid substrate (Acd or Aad), and 4 µM AcdRS2b enzyme. To control for background phosphate signal in our
reactions, we excluded AcdRS2b enzyme; to demonstrate the ATP-dependence of the amino acid activation reaction, we also prepared reactions excluding ATP. Reactions were incubated at 37 °C for 100 min, then quenched with three reaction volumes of ice-cold EDTA solution to a final concentration of 10 mM. Quenched reactions were developed with Malachite green per kit instructions (Malachite Green Phosphate Assay Kit, Sigma #MAK307) in a 96-well plate (Corning Costar). Absorbance at 620 nm was measured in an Infinite F200 plate reader (Tecan), and a P\textsubscript{i} standard curve was used to calculate phosphate concentration in each reaction. Data were reported as mean and standard deviation (n=3) of amount of PP\textsubscript{i} formed after subtracting the background signal from the no enzyme control reactions.

**Aad Docking Model.** The structure of 2-aminoacridone was optimized in Gaussian at the AM1 level of theory with a 6-31G basis set. The optimized structure was aligned with the sidechain of Acd docked into a homology model of AcdRS2b built using the crystal structure coordinates, as previously described.\textsuperscript{64, 88} The 2-aminoacridone molecule was aligned to position the amino group as in Aad. Steric clashes with the sidechains of Val164 and Ala167 were observed.


93. Schulman, S. G.; Sturgeon, R. J., VARIATIONS OF FLUORESCENCE EFFICIENCIES OF 9-(10H)-ACRIDONE AND ITS 4-METHOXY DERIVATIVE

149


101. M. J. Frisch; G. W. Trucks; H. B. Schlegel; G. E. Scuseria; M. A. Robb; J. R. Cheeseman; G. Scalmani; V. Barone; G. A. Petersson; H. Nakatsuji; X. Li; M. Caricato; A. Marenich; J. Bloino; B. G. Janesko; R. Gomperts; B. Mennucci; H. P. Hratchian; J. V. Ortiz; A. F. Izmaylov; J. L. Sonnenberg; D. Williams-Young; F. Ding; F. Lipparini; F. Egidi; J. Goings; B. Peng; A. Petrone; T. Henderson; D. Ranasinghe; V. G. Zakrzewski; J. Gao; N. Rega; G. Zheng; W. Liang; M. Hada; M. Ehara; K. Toyota; R. Fukuda; J. Hasegawa; M. Ishida; T. Nakajima; Y. Honda; O. Kitao; H. Nakai; T. Vreven; K. Throssell; J. A. Montgomery, J.; J. E. Peralta; F. Ogliaro; M. Bearpark; J. J. Heyd; E. Brothers; K. N. Kudin; V. N. Staroverov; T. Keith; R. Kobayashi; J. Normand; K. Raghavachari; A. Rendell; J. C. Burant; S. S. Iyengar; J. Tomasi; M. Cossi; J. M. Millam; M. Klene; C. Adamo; R. Cammi; J. W. Ochterski; R. L. Martin; K. Morokuma; O. Farkas; J. B. Foresman; D. J. Fox, Gaussian 09, Revision A.09. Gaussian, Inc.: Wallingford CT, 2016.


110. Biscoe, M. R.; Fors, B. P.; Buchwald, S. L., A New Class of Easily Activated Palladium Precatalysts for Facile C–N Cross-Coupling Reactions and the Low


