Novel Strategies for Producing Proteins with Non-proteinogenic Amino Acids

by

Ryan Christopher Nangreave

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved December 2013 by the
Graduate Supervisory Committee:

Sidney M. Hecht, Chair
Hao Yan
Ian Gould

ARIZONA STATE UNIVERSITY

December 2013

ABSTRACT

The biological and chemical diversity of protein structure and function can be greatly expanded by position-specific incorporation of non-natural amino acids bearing a variety of functional groups. Non-cognate amino acids can be incorporated into proteins at specific sites by using orthogonal aminoacyl-tRNA synthetase/tRNA pairs in conjunction with nonsense, rare, or 4-bp codons. There has been considerable progress in developing new types of amino acids, in identifying novel methods of tRNA aminoacylation, and in expanding the genetic code to direct their position.

Chemical aminoacylation of tRNAs is accomplished by acylation and ligation of a dinucleotide (pdCpA) to the 3′-terminus of truncated tRNA. This strategy allows the incorporation of a wide range of natural and unnatural amino acids into pre-determined sites, thereby facilitating the study of structure-function relationships in proteins and allowing the investigation of their biological, biochemical and biophysical properties.

Described in Chapter 1 is the current methodology for synthesizing aminoacylated suppressor tRNAs. Aminoacylated suppressor $tRNA_{CUA}$s are typically prepared by linking pre-aminoacylated dinucleotides (aminoacyl-pdCpAs) to 74 nucleotide (nt) truncated tRNAs (tRNA-C_{OH}) via a T4 RNA ligase mediated reaction. Alternatively, there is another route outlined in Chapter 1 that utilizes a different pre-aminoacylated dinucleotide, AppA. This dinucleotide has been shown to be a suitable substrate for T4 RNA ligase mediated coupling with abbreviated tRNA-C_{OH}s for production of 76 nt aminoacyl-$tRNA_{CUA}$s. The synthesized suppressor tRNAs have been shown to participate in protein synthesis *in vitro*, in an S30 (*E. coli*) coupled transcription-translation system in which there is a UAG codon in the mRNA at the position corresponding to Val10.

Chapter 2 describes the synthesis of two non-proteinogenic amino acids, L-thiothreonine and L-*allo*-thiothreonine, and their incorporation into predetermined positions of a catalytically competent dihydrofolate reductase (DHFR) analogue lacking cysteine. Here, the elaborated proteins were site-specifically derivitized with a fluorophore at the thiothreonine residue.

The synthesis and incorporation of phosphorotyrosine derivatives into DHFR is illustrated in Chapter 3. Three different phosphorylated tyrosine derivatives were prepared: bis-nitrobenzylphosphoro-L-tyrosine, nitrobenzylphosphoro-L-tyrosine, and phosphoro-L-tyrosine. Their ability to participate in a protein synthesis system was also evaluated.

ACKNOWLEDGEMENTS

My sincerest gratitude is extended to my research advisor, Professor Sidney M. Hecht. His ability to guide and instill independent thought in me has truly facilitated my growth as a chemist and as an individual and for this I am extremely grateful. I would also like to thank Professors Hao Yan and Yan Liu for their guidance throughout the years and help with the subtleties of graduate school. I would also like to acknowledge Professor Ian Gould who played a large role in my choice to attend Arizona State University and for his help in my progression through the program.

Research is an ongoing quest for knowledge and an attempt to understand the unknown, and the desire to pursue the unidentified that I developed during my time with Dr. Hecht is something I'm especially thankful for. I must extend my sincerest gratitude to Dr. Larisa Dedkova. She offered her guidance and support quite often and truly exemplifies the best qualities of a mentor, colleague and friend. I would also like to thank Dr. Noureddine Fahmi for his leadership and firm understanding of synthetic organic chemistry, as well as Dr. Shengxi Chen for his help with all biochemical aspects of my PhD studies. Dr. Pablo Arce, Dr. Manikandidas M.M., and Dr. Rakesh Paul offered a great deal of insight into many complicated research issue that I encountered and for that I thank them.

I extend my gratitude to Rumit Maini for continually inspiring me to search for new and interesting research ideas and keeping my mind fresh, and Trevor Bozeman for his continual assistance in the lab, which undoubtedly assisted in my graduation. I would also like to thank my dear friend Basab Roy who has helped me rediscover many interests that I once had, and given me a different outlook on life. My thanks are also

extended to the current and former members of the Sidney Hecht group.

I would finally like to thank my family. To my parents who instilled me with the values and ideas that lead me to this point in my life - without their patience, guidance, acceptance and love I would surely not have accomplished this. To my wife and my daughter who support me unconditionally and inspire me to grow every day.

TABLE OF CONTENTS

Page

LIST OF ABBREVIATIONS

aa	amino acid
AcOH	acetic acid
AMP	adenosine-5′-monophosphate
AppA	diadenosyl-5′,5′-diphosphate
ATP	adenosine-5′-triphosphate
aq	aqueous
Bn	benzyl
Boc	t-butoxycarbonyl
Boc_2O	t-butoxycarbonyl anhydride
BOP-Cl	Bis(2-oxo-3-oxazolidinyl)phosphonic chloride
^oC	degrees Celsius
cat	catalytic
cm	centimeter
CPM	7-diethylamino-3-(4′-maleimidylphenyl)-4-methylcoumarin
Cys	cysteine
δ	chemical shift (ppm)
d	doublet
DCC	*N*,*N*′-dicyclohexylcardodiimide
dec	decomposition
dd	doublet of doublets
ddd	doublet of doublet of doublets
DEAE	diethylaminoethyl

DHFR	dihydrofolate reductase
DIPEA	diisopropylethylamine
DMAP	dimethylaminopyridine
DMF	dimethylformamide
DMSO	dimethylsulfoxide
DMT	4,4′-dimethoxytrityl
DNA	deoxyribonucleic acid
DTT	dithiothreitol
EDTA	ethylenedinitrilotetraacetic acid
ESI	electrospray ionization
Fig	figure
Fmoc	*N*-(9-fluorenylmethoxy-carbonyloxy)
FmocOSu	*N*-(9-fluorenylmethoxy-carbonyloxy)succinimide
g	gram
h	hour
^{1}H NMR	proton nuclear magnetic resonance
HATU	2-(7-aza-1H-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate
HBTU	*O*-benzotriazole-*N*,*N*,*N*′,*N*′-tetramethyl-uronium-hexafluoro-phosphate
HPLC	high pressure liquid chromatography
Hz	Hertz
J	coupling constant
L	liter

M	molar
m	multiplet
MALDI-TOF	matrix assisted laser desorption ionization - time of flight
MBL	mannose-binding lectin
mCPBA	metachloroperbenzoic acid
Met	methionine
MHz	mega Hertz
min	minutes
mL	milliliter
mM	millimolar
mmol	millimole(s)
mpYRNA8	plasmid containing the gene encoding 75 nucleotide $tRNA_{CUA}$
mRNA	messenger ribonucleic acid
μmol	micromole(s)
NHS	*N*-hydroxysuccinimide
nm	nanometer
nt	nucleotide
NMR	nuclear magnetic resonance
NTA	nitrilotriacetic acid
NVOC	6-nitroveratryoxycarbonyl
NVOCCl	6-nitroveratryoxycarbonyl chloride
pdCpA	5′-*O*-phosphoryl-2′-deoxycytidylyl(3′→5′)adenosine
PTC	peptidyltransferase center

pTyr	phosphorotyrosine
pYRNA8	plasmid containing the gene encoding 74 nucleotide $tRNA_{CUA}$
R_f	ratio of fronts
RNA	ribonucleic acid
s	singlet
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
satd	saturated
soln	solution
t	triplet
TBA	tetrabutylammonium
TFA	trifluoroacetic acid
THF	tetrahydrofuran
TLC	thin layer chromatography
tRNA	transfer RNA
UV	ultraviolet
Val	valine

LIST OF TABLES

LIST OF FIGURES

Figure Page

LIST OF SCHEMES

Scheme Page

CHAPTER 1

INTRODUCTION

1.1. Protein biosynthesis

Proteins are biochemical macromolecules that play a central role in a vast array of biological processes in living organisms; they catalyze metabolic reactions, transport molecules within cells, replicate DNA, and perform many other vital functions. Proteins contain one or more chains of amino acids (polypeptides) that are folded into specific three-dimensional shapes. In most cases, the identity and position of the amino acid residues and their side chains within a protein determine its properties and function.

Protein biosynthesis is the process by which cells generate new proteins. The biosynthetic pathway for proteins involves a number of events including transcription of nuclear DNA to mRNA, translation of mRNA, and polymerization of amino acids to afford protein products. Protein synthesis is carried out by an elegant complex composed of the ribosome and several accessory protein factors, as well as mRNA and aminoacylated tRNA molecules. All ribosomes are comprised of two subunits of unequal size. In *E. coli*, the small subunit is referred to as the 30S subunit, while the large subunit is called the 50S subunit. Together, these subunits are responsible for decoding the genetic message and catalyzing peptide bond formation.[1] It has recently been shown that mutation of the 50S subunit, specifically the 23S component, can affect the ability of the ribosomal machinery to accommodate β-L-amino and D-amino acids for incorporation into nascent polypeptides.[2-4]

Protein translation proceeds via four distinct steps, namely, (1) tRNA activation,

(2) initiation of translation, (3) peptide elongation, and (4) peptide termination. tRNA molecules are the decoders of genetic information. tRNAs provide the link between the sequence of nucleotides in mRNA and the resulting sequence of amino acids in a polypeptide. Every cell contains at least 20 tRNA molecules, one for each naturally occurring amino acid, and each tRNA recognizes at least one codon (triplet sequence) in the genetic code. In an aminoacylation event (tRNA activation), a particular amino acid is covalently attached to the 3′ end of a tRNA molecule by means of an aminoacyl-tRNA synthetase (aaRS) and ATP. Aminoacylation proceeds in two discrete steps. First, an amino acid is activated by the formation of a reactive aminoacyl-adenylate intermediate. The intermediate is tightly (but non-covalently) bound to the aminoacyl-tRNA synthetase. Subsequent hydrolysis of pyrophosphate promotes an essentially irreversible reaction. Next, the aminoacyl group is transferred from the intermediate to tRNA. Upon binding to its cognate tRNA, the amino acid is transferred to either the 2′ or 3′-OH group of the terminal adenylate residue of tRNA, depending on the specific synthetase catalyzing the reaction. There are two classes of aminoacyl-tRNA synthetases, class I and class II, that aminoacylate the A_{76} 2′- or 3′-OH group of the tRNA, respectively. If the amino acid is initially attached to the 2′-OH of the tRNA, it will ultimately migrate to afford a 3′-*O*-acyl isomer that can serve as a substrate for protein synthesis.[5]

Initiation of protein synthesis involves the formation of a translation complex at the beginning of the mRNA coding sequence. The complex contains the ribosomal subunits, the mRNA template, an initiation tRNA molecule, and several accessory protein initiation factors. The first codon translated is typically AUG. Every cell contains at least

two methionyl-tRNA^{Met} molecules that recognize the AUG codon, one of which is used exclusively at initiation codons and is referred to as the initiator tRNA. In prokaryotes, the 30S ribosomal subunit binds to the mRNA at a purine-rich region of the mRNA template just upstream of the initiation codon. This region, referred to as the Shine-Dalgarno sequence, is complementary to a pyrimidine-rich region at the 3′ end of the 16S rRNA component. Formation of the initiation complex requires several initiation factors; prokaryotes contain 3 factors whose role is to ensure that the aminoacylated initiator tRNA is correctly positioned at the initiation codon.

At the close of the initiation step, the mRNA is positioned such that the next codon can be translated during the elongation stage of protein synthesis. The elongation cycle can be summarized by three major events, (1) codon recognition, (2) peptide bond formation and (3) translocation, which are assisted by elongation factors. The recognition event between the next codon in the mRNA and the anti-codon loop of the corresponding aminoacylated tRNA results in binding of the tRNA to the A-site of the ribosome, and transfer of the amino acid from the tRNA in the P-site to the tRNA transiently bound in the A-site. The subsequent translocation event results in movement of the tRNAs in the A- and P-sites to the P- and E-sites, respectively. The next aminoacylated-tRNA, corresponding to the next codon in the mRNA, then binds to the vacant A-site. Elongation proceeds until termination is signaled by the presence of any one of three stop codons: UAG, UAA and UGA. The presence of stop codons in the ribosomal A-site initiates binding of protein release factors due to the lack of cognate tRNAs present in the cell. These factors trigger the release of the nascent polypeptide from the ribosome and the dissociation of the ribosomal subunits.[6]

1.2. Suppressor tRNA technology

There are several methods that can be used to circumvent the termination event, including mRNA frame-shifting and the use of modified or cellular tRNAs capable of recognizing stop codons (natural suppressors). A suppression event refers to translation of a stop codon, thereby circumventing the effect of release factors that signal the termination of protein synthesis in the ribosome. Goodman and co-workers realized that the genetic code could be expanded by suppressing the UAG stop codon via modification of the nucleobases of a tyrosyl-tRNA anticodon loop. Briefly, the authors illustrated that mutation of the su_{III} gene changed the anticodon loop of a minor Tyr-tRNA species in *E. coli* from GUA to CUA. They demonstrated that the amino acid tyrosine could be inserted into a protein by suppressing the UAG stop codon.[7] Ten years later Hecht and co-workers pioneered the "chemical aminoacylation" of tRNAs, signifying a breakthrough in *in vitro* protein synthesis.[8] Here, a nitrophenylsulfenyl-*N*-protected amino acid was used to aminoacylate a P^1,P^2-bis(5′-adenosyl)diphosphate, which was subsequently ligated (by T4 RNA ligase) to a truncated tRNA lacking the terminal 3′-adenosine moiety. This technology has allowed researchers to mis-aminoacylate tRNAs, *i.e.* attach a variety of natural and unnatural amino acids to non-cognate tRNAs, potentially enabling the preparation of proteins containing non-proteinogenic amino acids. Several years later Hecht and co-workers improved upon this methodology by aminoacylating a dinucleotide, pCpA, with an *N*-protected amino acid and enzymatically ligating the aminoacylated dinucleotide to a 74 nucleotide tRNA (tRNA-C_{OH}) lacking the terminal two nucleotides at the 3′-end. However, because the aminoacyl-tRNA contained an *N*-protected amino acid they could not be used to produce modified proteins.[9] *N*-acetyl

protection of the amino group precludes the ability of misacylated tRNAs to be accepted in the A-site. Brunner and co-workers circumvented this problem by protecting the amine as a di-*tert*-butoxycarbonyl (*t*-Boc) derivative, which is an acid labile protecting group typically removed with trifluoroacetic acid (TFA).[10] Unfortunately, the preparation and subsequent aminoacylation of the dinucleotide afforded only modest amounts of 3′-*O*-aminoacylated pCpA. In an effort to overcome this issue, Schultz and co-workers prepared a deoxycytidine dinucleotide, pdCpA, which not only simplified the synthesis of the dinucleotide but also enhanced aminoacylation yields.[11] pdCpA is now widely used for preparing mono-aminoacylated $tRNA_{CUA}$s, a suppressor tRNA bearing one amino acid.

Interestingly, Stepanov and co-workers described the ability of a phenylalanyl-tRNA synthetase to attach two phenylalanine amino acids onto the 3′-end of its cognate tRNA.[12,13] However, the ability of the doubly charged tRNAs ability to participate in protein biosynthesis was not studied, nor well understood. The Hecht laboratory later demonstrated the ability of bisaminoacylated tRNAs to participate in protein synthesis.[14-16] The ability of bisaminoacylated tRNAs to function in protein synthesis suggests that the prokaryotic elongation factor EF-Tu•GTP is not only capable of recognizing tRNAs with two amino acids attached, but also of binding to and translocating tRNA from the cytoplasm to the ribosomal A-site. Additionally, they demonstrated that the nascent polypeptide, bound to tRNA in the P-site, is transferred to the aminoacyl moeity at the 3′-position of the bisaminoacylated $tRNA_{CUA}$ in the A-site. They demonstrated this by utilizing a bisaminoacylated tRNA bearing two different amino acids at the 2′- and 3′-positions of the adenosyl-ribose.[17,18]

CHAPTER 2

A NEW STRATEGY FOR THE SYNTHESIS OF BISAMINOACYLATED tRNAs

2.1. Introduction

Chemically aminoacylated suppressor tRNAs have allowed the introduction of various natural and non-natural amino acids into proteins at pre-determined sites as illustrated in Figure 2.1.[19-21]

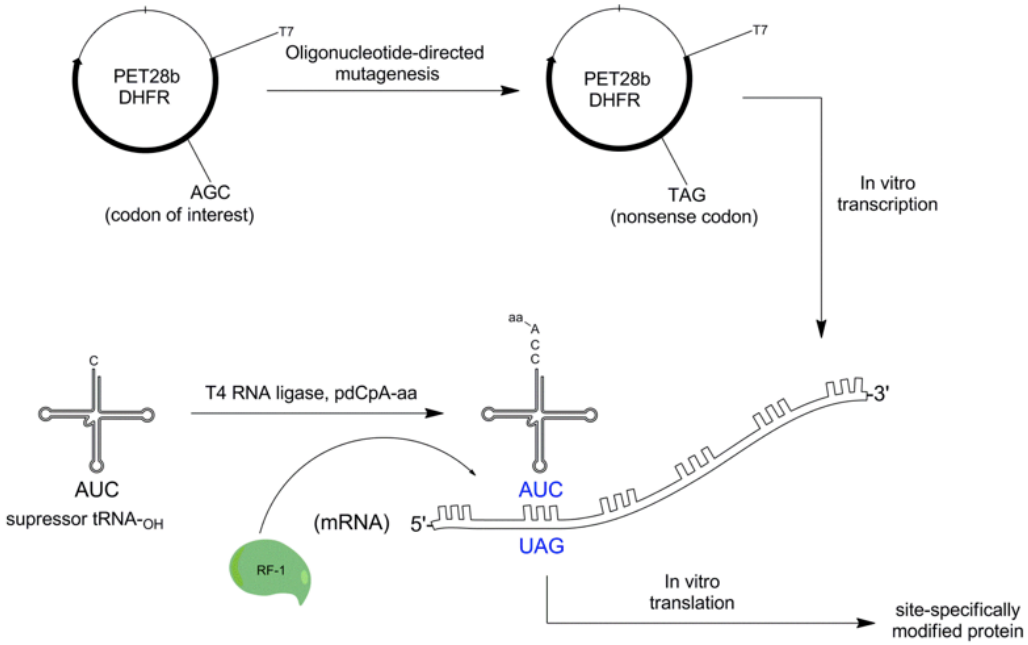


Figure 2.1. Utilization of $tRNA_{CUA}$s for protein elaboration.

Briefly, a UAG stop codon is introduced at a desired position within a gene of interest by oligonucleotide site-directed mutagenesis, and the gene is subsequently inserted into a plasmid expression vector.[20,22] Similarly, a gene coding for a complementary phenylalanyl-tRNA, containing an anticodon loop that recognizes and binds to the UAG codon in the modified mRNA, is inserted into an expression vector. Transcription of both plasmids results in the production of modified mRNA and $tRNA_{CUA}$. When used together in a cell free protein synthesis system, the aminoacylated $tRNA_{CUA}$ binds to the stop codon, suppresses any release factors, and enables peptide elongation to continue.[9,16,20,22-25]

This strategy has facilitated study of the biochemical and biophysical properties of numerous proteins.[7,26] While monoaminoacylated tRNAs are ordinarily used in protein synthesis systems, Lavrik and co-workers described a phenylalanyl-tRNA synthetase from *Thermus thermophilus* that incorporates more than one molecule of phenylalanine into $tRNA^{Phe}$, producing bis-(2',3'-*O*-phenylalanyl)-tRNAs.[12,13] It is notable that bisaminoacylated tRNAs exhibit higher chemical stability than their monoaminoacylated tRNA counterparts (Figure 2.2), and are able to participate in two cycles of peptide bond formation (Figure 2.3).[14-16,24]

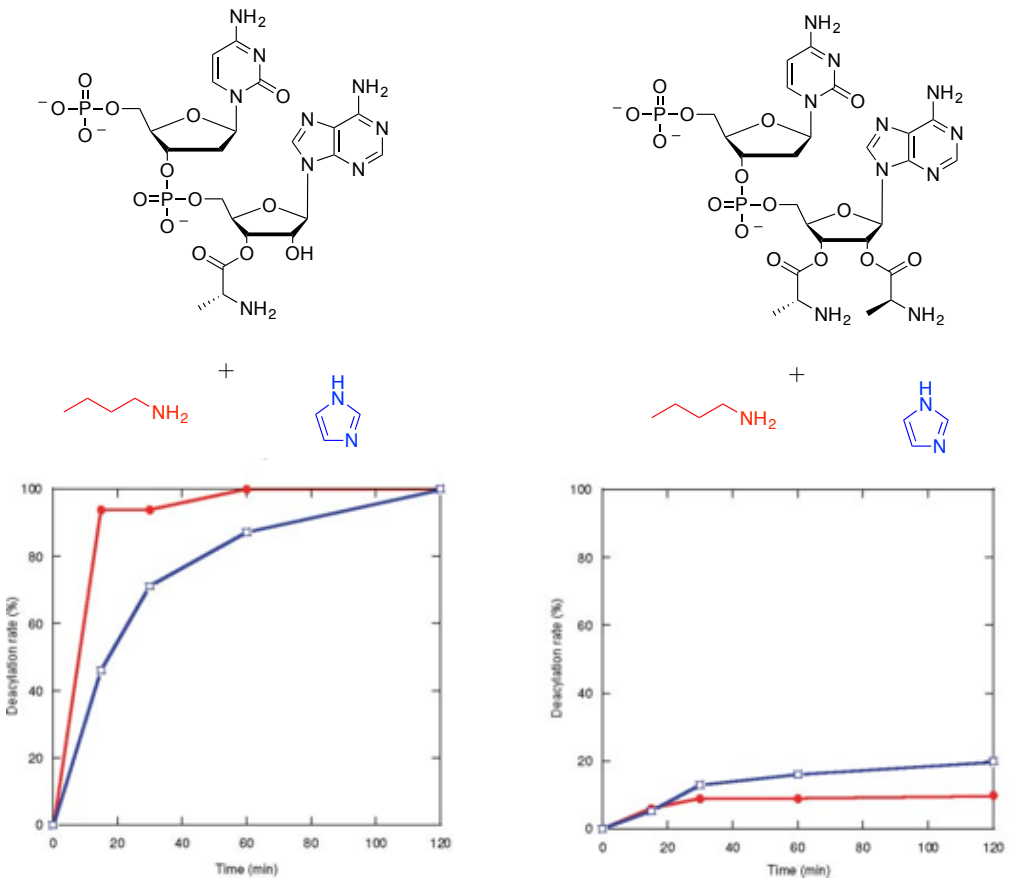


Figure 2.2. Relative stabilities of mono- and bisaminoacylated tRNA models.[15]

In the presence of two different nucleophiles, butylamine and imidazole, bisaminoacylated pdCpAs, which may be regarded as models for bisaminoacyl-tRNAs, were converted to their deacetylated derivatives much more slowly than the corresponding monoaminoacylated molecules (Figure 2.2).

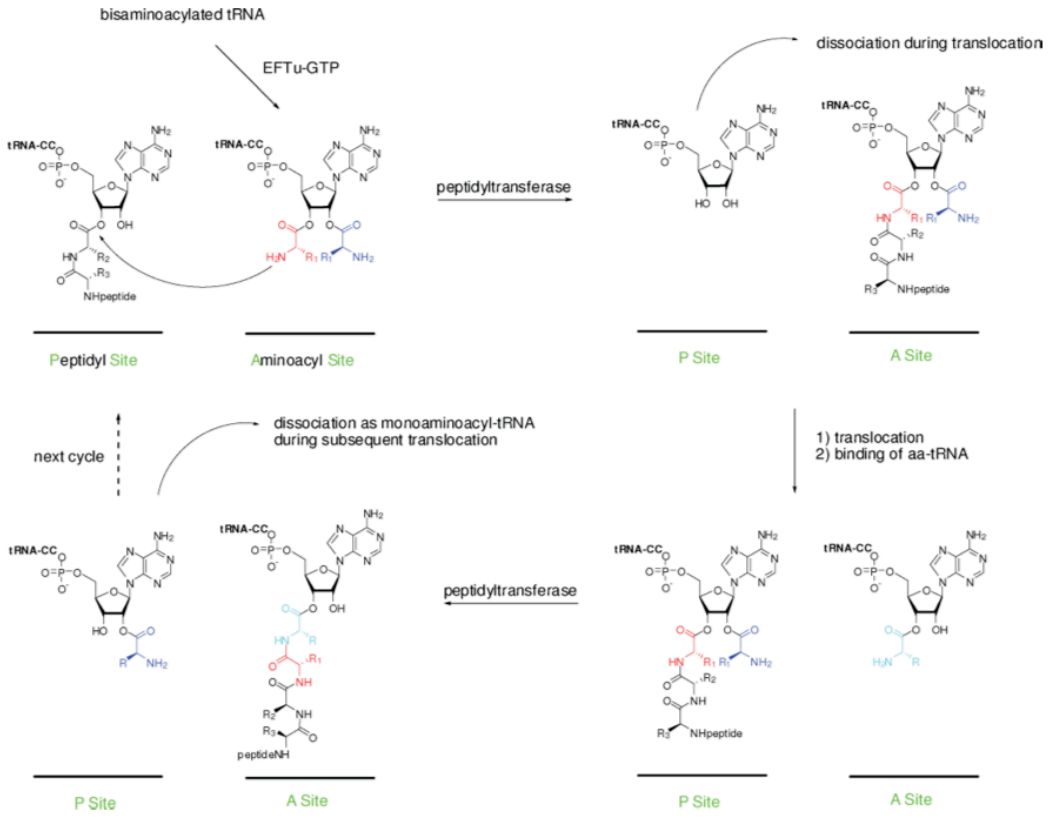


Figure 2.3. Utilization of bisaminoacylated $tRNA_{CUA}$s in protein synthesis.[15]

In addition, under synthesis conditions that are limiting for aminoacylated suppressor tRNAs, twice as much protein is expressed when utilizing bisaminoacylated $tRNA_{CUA}$s.[17] Considering these results, it is reasonable to assume that the advantages of bisaminoacylated $tRNA_{CUA}$s can be exploited for *in vitro* translation experiments and other protein studies. It is important to note that preparation of suppressor tRNAs bearing two amino acids is dependent on the preparation of the corresponding dinucleotide precursors. The synthesis of bisaminoacylated pdCpA bearing amino acids with relatively small side chains is expected to proceed in reasonable yield. However, it seems likely that steric bulk from the nitrogenous bases of the dinucleotide, together with the amino acid

side chains, will hinder the preparation of bisaminoacylated pdCpAs with amino acids having large side chains.[27] Careful consideration and revision of the methodology could facilitate general use of this technology and lead to the production of suppressor tRNAs bearing two non-natural amino acids with large side chains.

The proposed strategy to expand the production bisaminoacylated derivatives of pdCpA described herein is based on a 'chemical aminoacylation' technology reported by Hecht and co-workers.[8] Hecht demonstrated that a nitrophenylsulfenyl-*N*-protected amino acid could be used to aminoacylate a P^1,P^2-bis(5′-adenosyl)diphosphate. The aminoacylated dinucleotide was then ligated to a truncated tRNA lacking the terminal 3′-adenosine nucleotide by T4 RNA ligase. Based on this result, we predicted that bisaminoacylation of 5′-AMP with amino acids having bulky side chains would proceed more easily than pdCpA, and that the resulting bisaminoacylated AMP would still be converted to a substrate recognized by T4 RNA ligase for attachment to an abbreviated $tRNA_{CUA}$. To test this hypothesis we aminoacylated pdCpA (Figure 2.4) and AMP (Figure 2.5) with amino acids bearing relatively large side chains (Figure 2.6) and analyzed the relative amount of mono- and bisaminoacylated products.

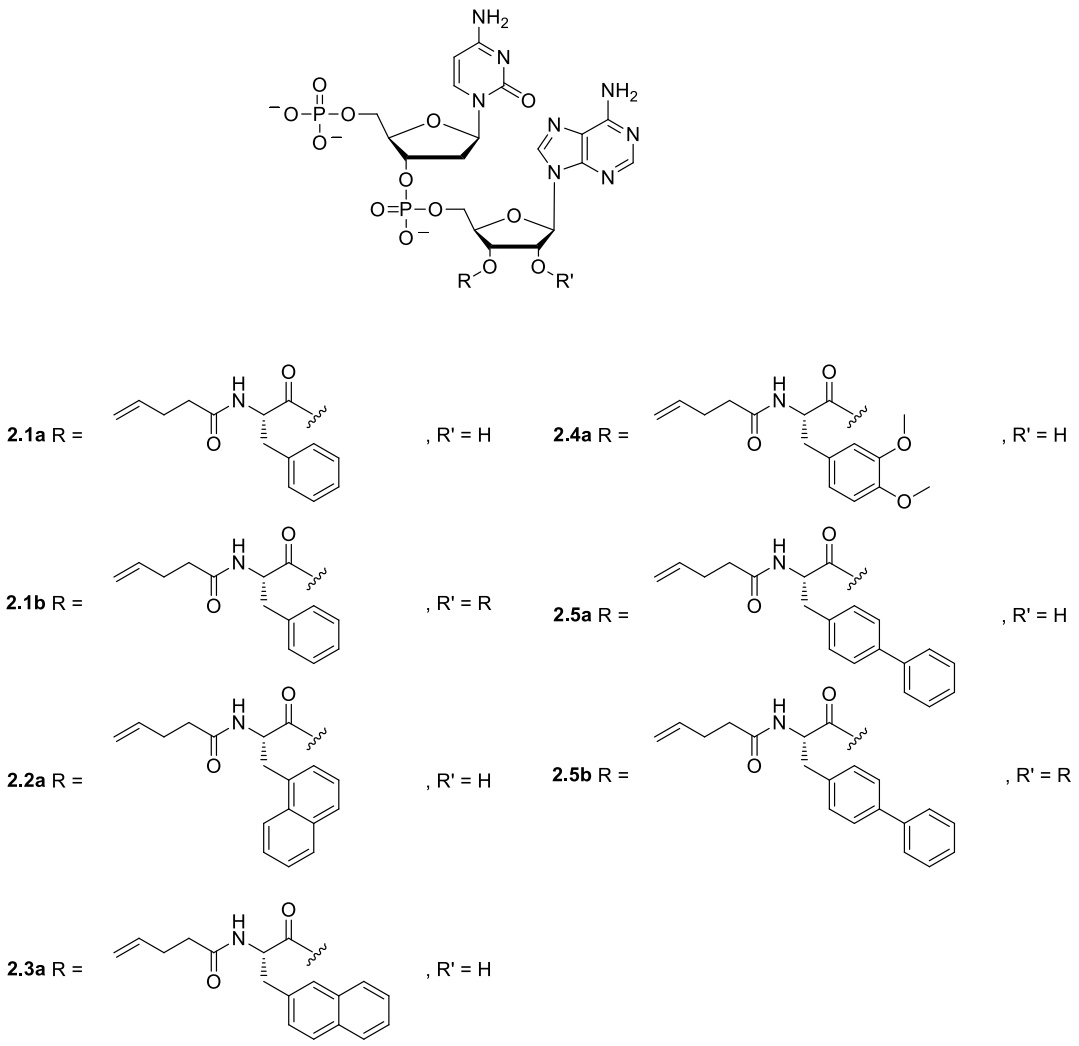


Figure 2.4. Mono- and bisaminoacylated pdCpAs.

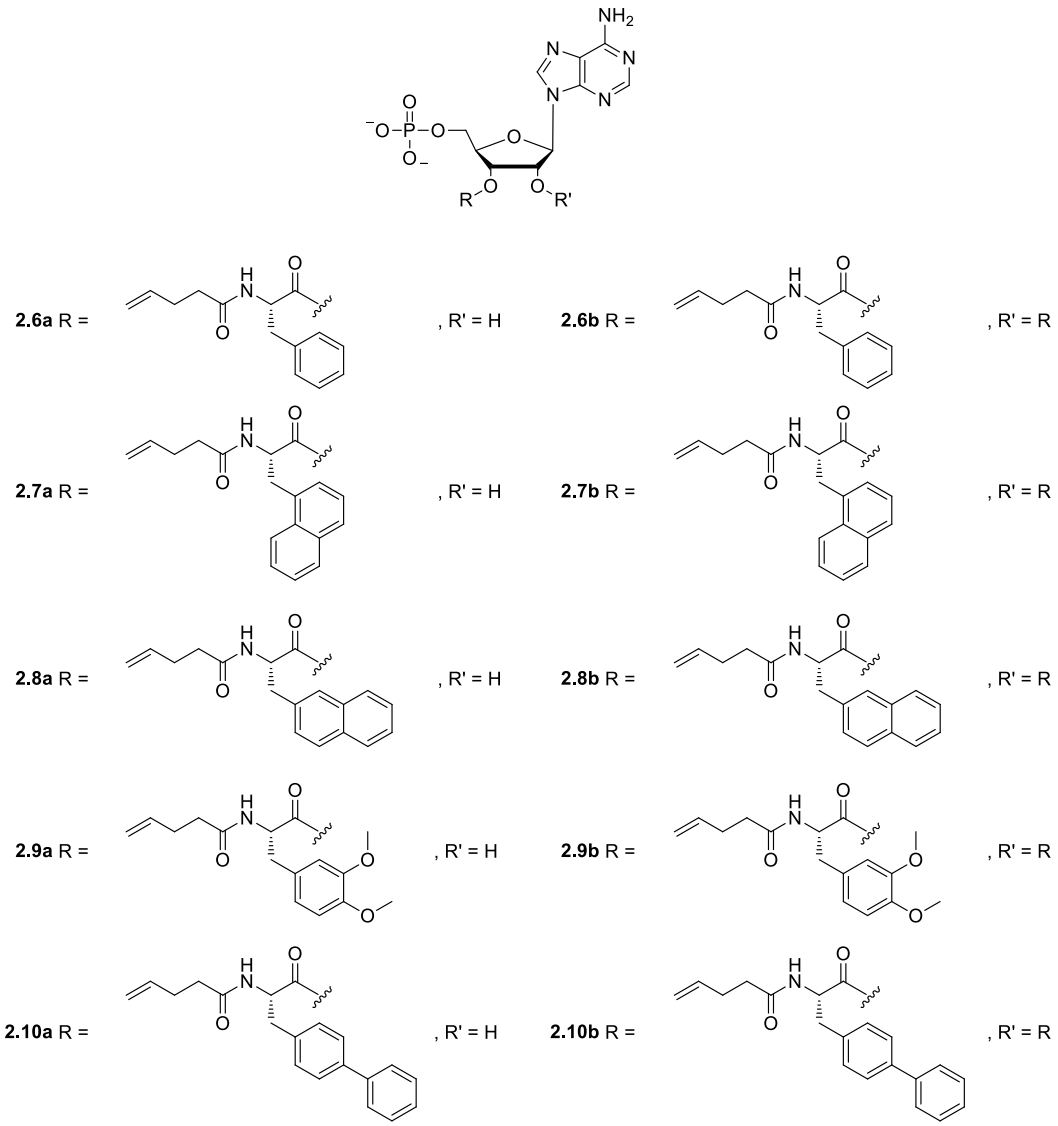


Figure 2.5. Mono- and bisaminoacylated AMPs.

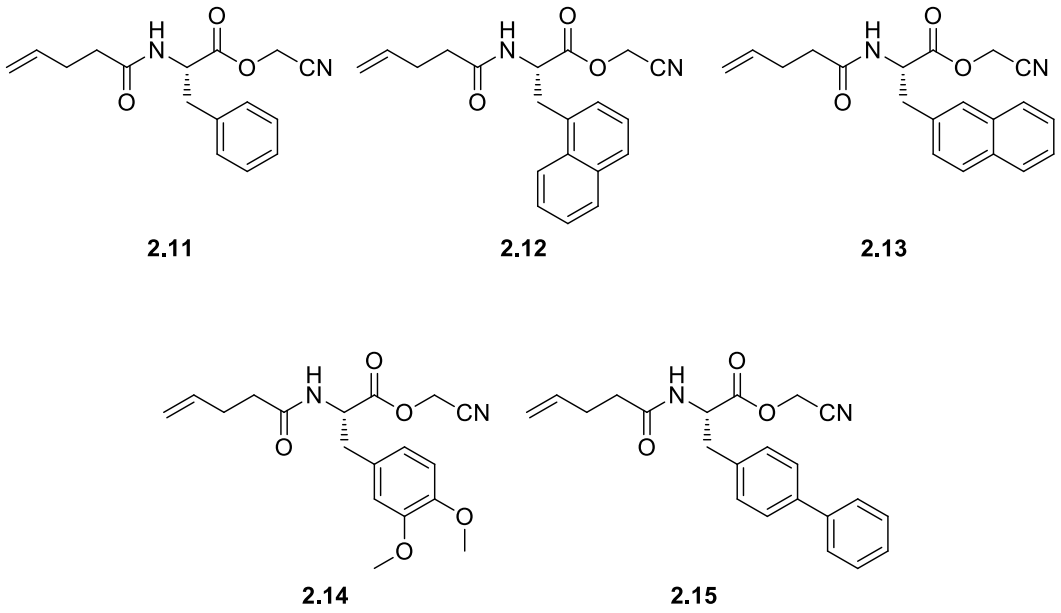
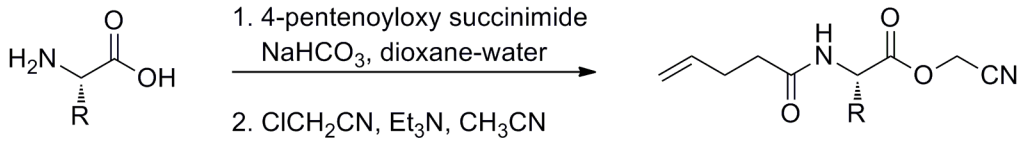


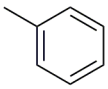
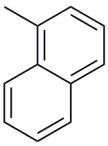
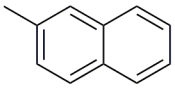
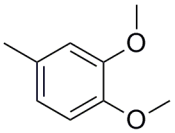
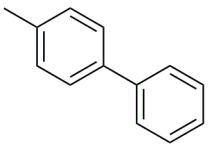
Figure 2.6. Amino acids analogues prepared for aminoacylation of pdCpA and AMP.

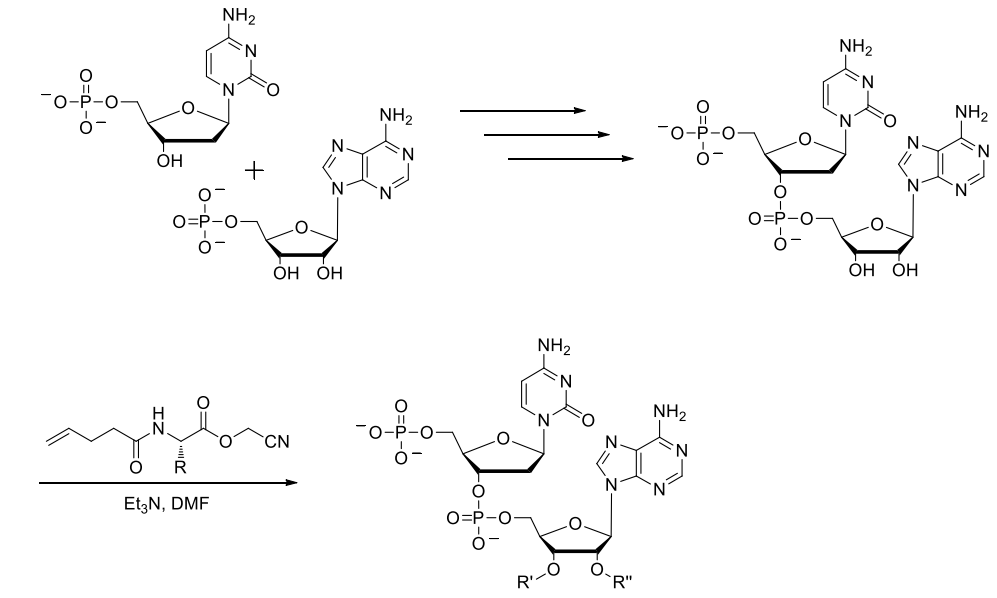
2.2. Results

In order to synthesize aminoacylated tRNAs, the free amino acid must first be protected, followed by activation of the carboxylic acid as the cyanomethyl ester. Pentenoyl protected amines were prepared from the corresponding free amino acids by treatment with 4-pentenoyloxy succinimide ester; the resulting crude products were extracted, isolated, and subsequently used without further purification. *N*-pentenoyl-L-amino acids were then treated with chloroacetonitrile and triethylamine to yield products **2.11-2.15**, in yields ranging from 56% to 88% (Table 2.1).

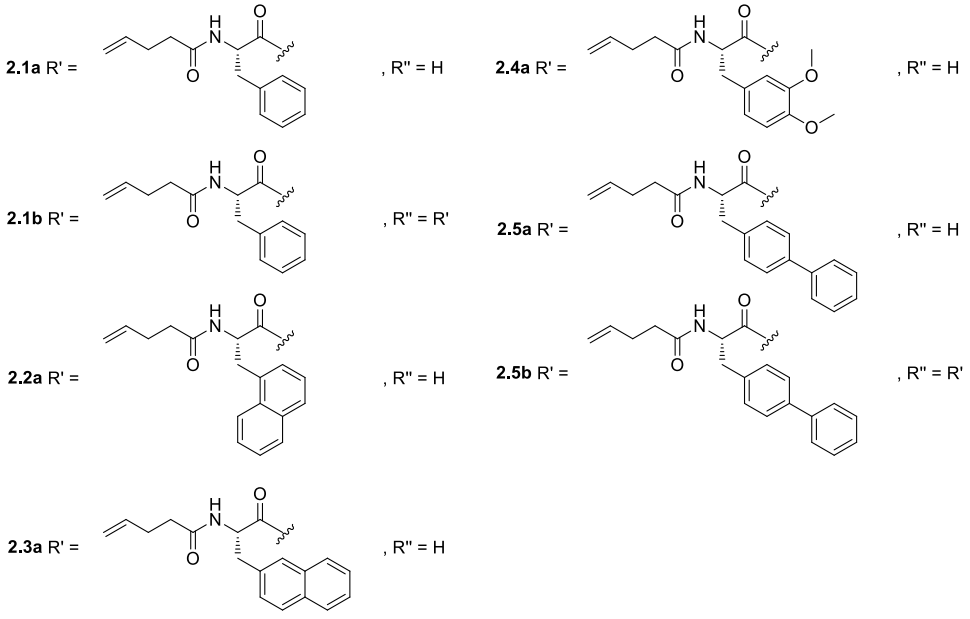
Table 2.1. Pentenoyl protection and cyanomethyl ester formation of amino acids.



R		entry
	88%	2.18
	65%	2.19
	77%	2.20
	70%	2.21
	56%	2.22

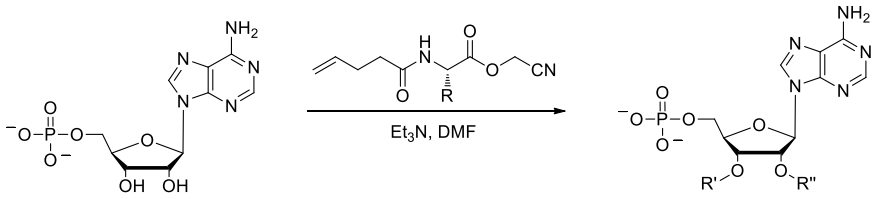


60%	**2.1a**
30%	**2.1b**
70%	**2.2a**
65%	**2.3a**
84%	**2.4a**
50%	**2.5a**
40%	**2.5b**

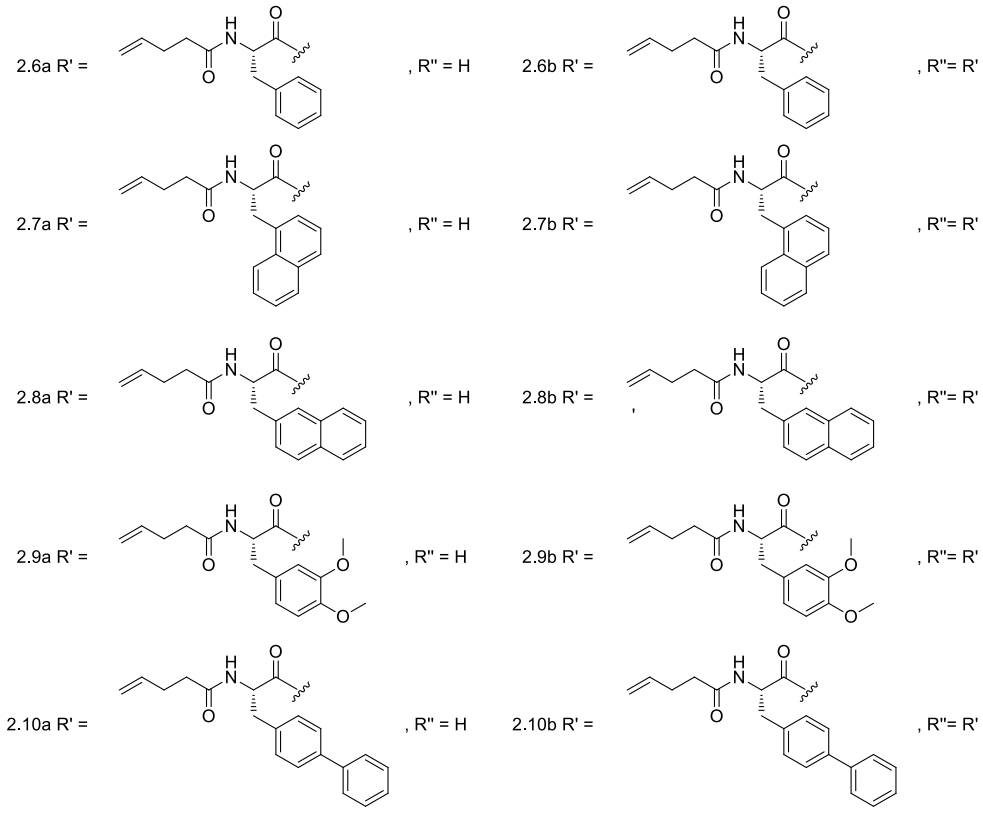


Scheme 2.1. Synthesis of mono- and bisaminoacylated-pdCpAs.

The syntheses of mono- and bisaminoacylated dinucleotide pdCpAs and mononucleotide AMP are illustrated in Schemes 2.1 and 2.2, respectively. Coupling amino acids to pdCpA in DMF, in the presence of Et_3N, affords the mono- and in some instances bisaminoacylated pdCpAs. The monoaminoacylated derivatives were prepared in relatively high yields, 50-84%. However, there were only two amino acids, phenylalanine and biphenylalanine, that afforded the bisaminoacylated species. Moreover, generation of the bisaminoacylated dinucleotide proceeded in modest yields, 30% and 40%, respectively. While there are presumably many factors involved in the successful production of dinucleotides bearing two amino acids, the steric hindrance resulting from the amino acid side chains and the orientation of the dinucleotide are certainly relevant. Thus, as the size of the amino acid side chains increased we observed the production of fewer bisaminoacylated species. Typically, the reactions were monitored by reverse phase HPLC, and purified after the consumption of pdCpA using a linear gradient of acetonitrile in 45 mM ammonium acetate (pH 4.5, 45 minutes). Collection and lyophilization of the appropriate fractions gave the aminoacylated dinucleotide as a white solid. Interestingly, if required, the two pdCpAs isomers (2′- and 3′-aminoacyl pdCpAs) can be separated following a previously described purification method.[18]



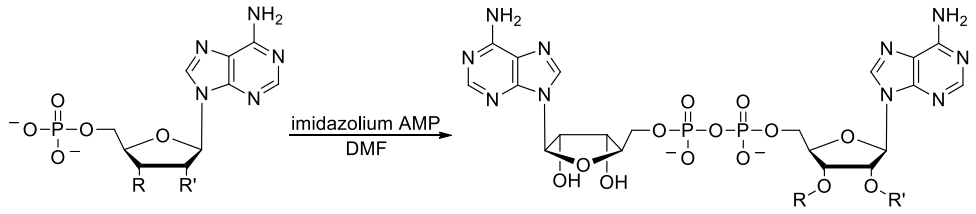
30%	2.6a
67%	2.6b
34%	2.7a
54%	2.7b
28%	2.8a
57%	2.8b
25%	2.9a
67%	2.9b
21%	2.10a
77%	2.10b



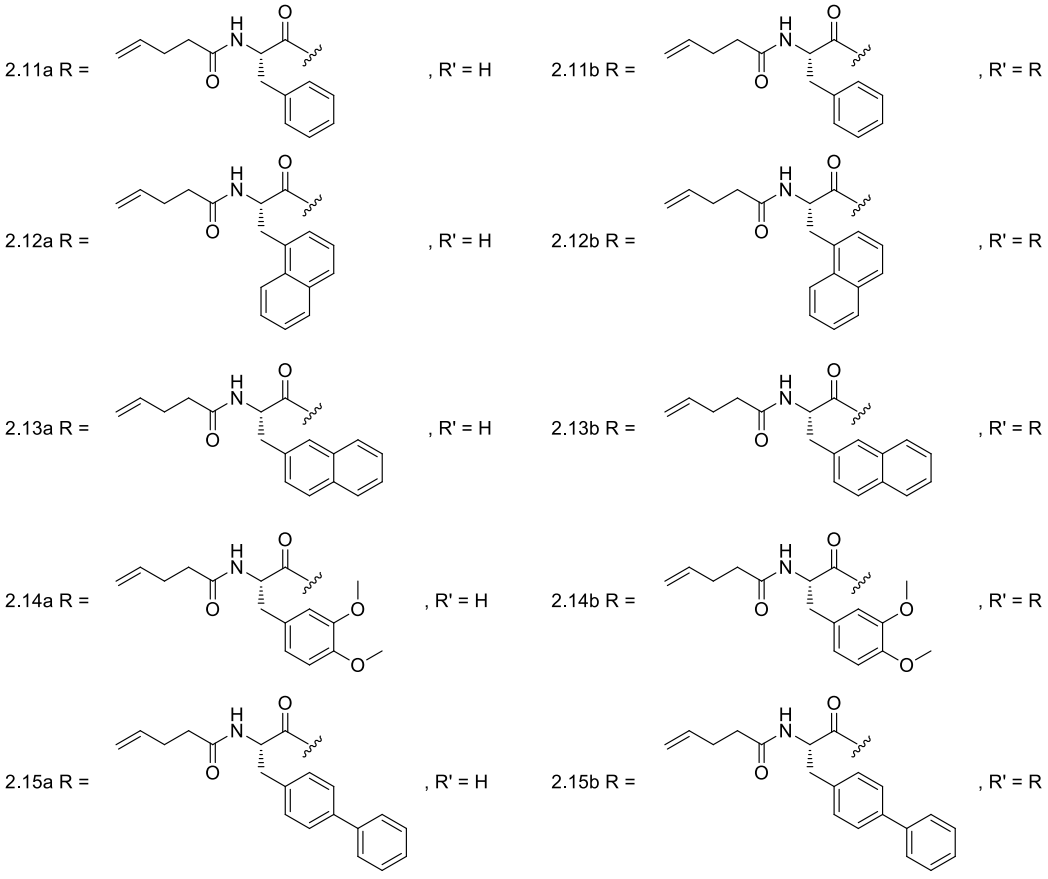
Scheme 2.2. Synthesis of mono- and bisaminoacylated AMPs.

Utilizing the mono nucleotide adenosine-5′-monophosphate for aminoacylation with amino acids **2.11**-**2.15** afforded both the mono- and bisaminoacylated nucleotides,

regardless of the amino acid side chain, in yields ranging from 67-77%. While the aminoacylated dinucleotide pdCpA-aa is a substrate for T4 RNA ligase, the aminoacyl AMP analogue is not recognized by the enzyme and must be converted to its AMP pyrophosphate derivative, AppA-aa. Treatment of each of the five monoaminoacylated species, **2.6a**-**2.10a**, and five bisaminoacylated AMPs, **2.6b**-**2.10b**, with imidazolium AMP afforded the respective aminoacylated-AppAs in yields ranging from 81% to 94% (Scheme 2.3). Gratifyingly, preparation of a usable substrate for the generation of suppressor tRNAs, *i.e.* aminoacyl-AppA derivatives, did not seem to be affected by mono- or bisaminoacylation.



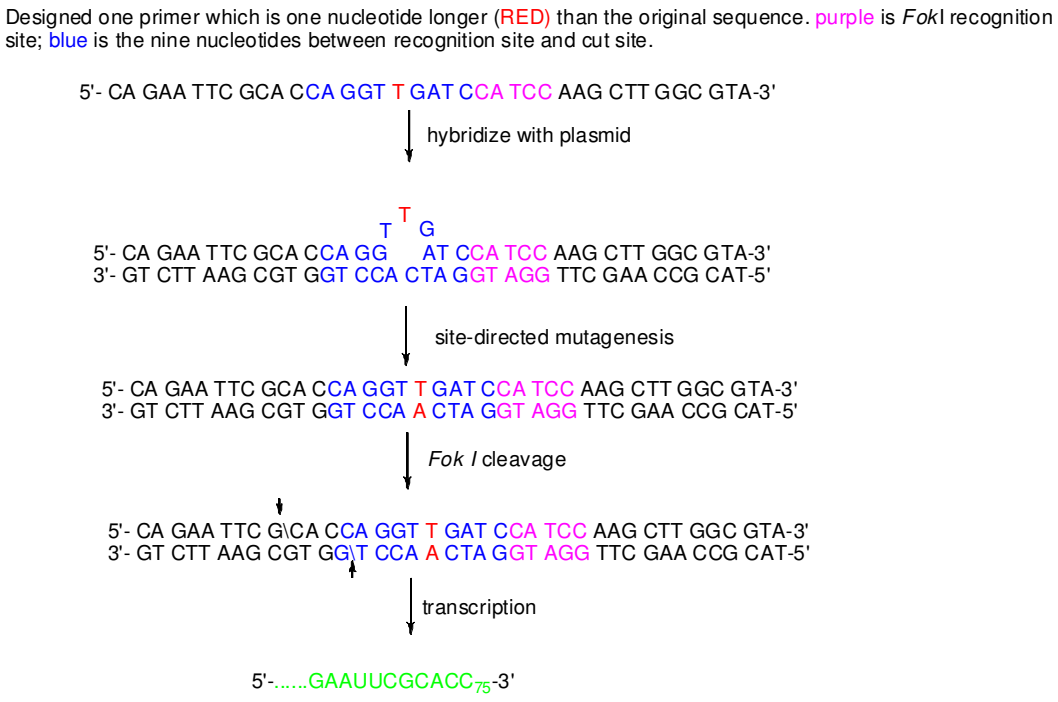
87%	2.11a
90%	2.11b
94%	2.12a
86%	2.12b
93%	2.13a
87%	2.13b
85%	2.14a
81%	2.14b
79%	2.15a
81%	2.15b



Scheme 2.3. Synthesis of mono- and bisaminoacylated AppAs.

Preparation of AppA derivatives is dependent on the preparation of imidzaolium AMP. AppA derivatives were prepared in a similar fashion as previously reported.[28,29] Generally, the imidazolium derivative is prepared *in situ* with triphenylphosphine, 4,4′-dipyridyl disulfide, triethylamine and methylimidazole. By using imidazole rather than the methylated derivative, the reagent could be isolated by precipitation (in 95% yield) and stored for future use.

Suppressor tRNA-C_{OH}, in combination with pdCpA, has been used extensively for the preparation of misacylated tRNAs.[11,17,18,21,25,27,30-32] However, the suppressor tRNA-CC_{OH} had never been described, and thus, had to be prepared. Oligonucleotide directed mutagenesis of a pYRNA8 plasmid (containing the gene for tRNA-C_{OH}) allowed the insertion of one nucleotide, T, between the recognition and cleavage sites of restriction enzyme *Fok*I (Scheme 2.4). This work was performed by Dr. Shengxi Chen.



Scheme 2.4. Strategy for inserting one nucleotide into a specific position in pYRNA8.

The primer that was used for insertion of the desired nucleotide, 5′-CAG AAT TCG CAC CAG GTT GAT CCA TCC AAG CTT GGC GTA A-3′, was incubated with ATP and T4 polynucleotide kinase (PNK) so that it would be phosphorylated at the 5′-end.[33] A polymerase chain reaction (PCR) was carried out in a mixture of pYRNA8 plasmid DNA, phosphorylated primer, dNTPs, *Pfu* DNA polymerase and *Taq* DNA ligase. Pre-incubation of the mixture at 65 °C for 5 minutes ensured that all nicks in the double helices would be repaired by the ligase; the thermal cycle was then carried out, followed by a post-incubation hold at 75 °C for 7 minutes. The restriction enzyme *Dpn*I was subsequently added to the mixture, which was then incubated at 37 °C for 60 minutes to eliminate the methylated and hemi-methylated wild type DNA template. Ten μL of the final sample was transformed into 100 μL of *E. coli* competent cell line DH5α. Linearization of the mutant plasmid pYRNA8 with *Fok*I to obtain the truncated (75 nt) yeast suppressor $tRNA^{Phe}_{CUA}$-CC_{OH} was performed by incubation at 37 °C for 4 hours. The supernatant was carefully decanted, washed with 70% ethanol, and dried in air (Figure 2.7).

Figure 2.7. Digestion of mpYRNA8 and pYRNA8 plasmids with *Fok*I.

Transcription of the digested plasmid was carried out with an Ampliscribe T7 transcription kit containing 7.5 mM each of ATP, CTP, GTP and UTP, 10 mM dithiothreitol, linearized DNA template and T7 RNA polymerase. The reaction was run at 37 °C for 6 hours. Following incubation, the mixture was treated with RNase free DNase I and incubated at 37 °C for 15 minutes. Next, the mixture was incubated in 3 M NaOAc and ethanol. The DNA was precipitated by incubation at –20 °C for 30 minutes,

followed by centrifugation (14,000×g for 30 minutes at 4 °C).

The elaborated tRNA-CC_{OH} transcript was dissolved in 0.1 M NaOAc, pH 5.2, and applied to a 400-µL DEAE-Sepharose CL-6B column. The column was subsequently washed with 0.1 M NaOAc, pH 5.2. Elution with a sodium chloride gradient afforded the desired $tRNA_{CUA}$s . The collected fractions were then treated with 2-propanol, incubated at 4 °C for 30 minutes, and centrifuged at 14,000×g for 30 minutes. The pellets were washed with 70% ethanol and reconstituted in RNase-free water for subsequent analysis in an 8% denaturing polyacrylamide gel (100 V, 2 hours). The fractions (0.6 and 0.7 M NaCl) affording the tRNA-CC_{OH} were pooled and stored at –80 °C. (Figure 2.8).

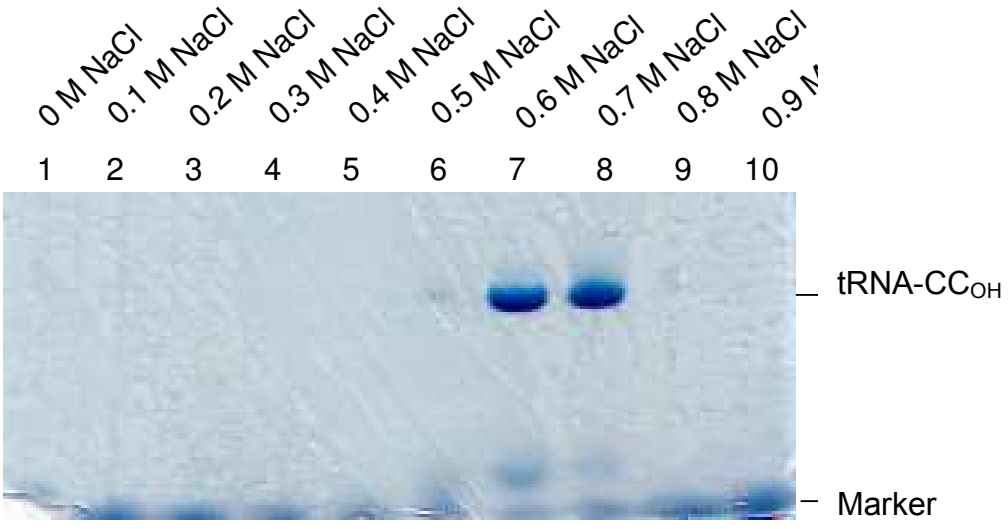


Figure 2.8. Elution of abbreviated suppressor tRNA-CC_{OH} on a DEAE-Sepharose Cl-6B column with NaOAc gradient.

Mono and bis-aminoacyl-tRNAs were prepared by a T4 RNA ligase-mediated ligation reaction between the *N*-protected, chemically synthesized aminoacylated pdCpA derivatives and the abbreviated suppressor tRNA-C_{OH} lacking the last two nucleotides at the 3′-end of the tRNA[11,22]. Similarly, ligation of aminoacylated AppA derivatives employed the abbreviated suppressor tRNA-CC_{OH} lacking the last nucleotide, adenosine, at the 3′-end. The ligation reactions were carried out in 100 mM HEPES, pH 7.5, with ATP (only in the reaction containing aminoacylated pdCpA derivatives + suppressor tRNA-C_{OH}), 15 mM $MgCl_2$, the appropriate suppressor tRNA, 2.0 A_{260} units of *N*-pentenoyl- protected mono- or bisaminoacyl-pdCpA or AppA derivatives, 5- to 10-fold molar excess, 15% DMSO and T4 RNA ligase. After incubation at 37 °C for 30 minutes, the tRNAs were precipitated and dissolved in 50 μL of water. The efficiencies of ligation were estimated by denaturing polyacrylamide gel electrophoresis at pH 5.2 (100 V, 2 hours). Estimating the ligation efficiency for AppA derivatives proved difficult; resolving the single nucleotide addition to the abbreviated tRNA and verification of the ligation reaction was illustrated through expression experiments.

Deprotection of suppressor tRNAs was achieved by using I_2 (25 mM,1:1 H_2O–THF). The reaction mixtures were incubated at 25 °C for 10 minutes. The tRNAs were precipitated by the addition of ethanol and centrifuged at 14,000 × g (30 minutes, 4 °C). The supernatants were then carefully decanted and the pellets were washed with 50 μL of 70% ethanol and then reconstituted in water.

In vitro translation reactions were performed using a bacterial S-30 extract from *E. coli* strain BL21(DE3). The reaction mixtures were incubated at 37 °C for 45 minutes.

Aliquots from *in vitro* translation mixtures, employing mono- and bisaminoacylated tRNAs prepared from aminoacyl-AppA precursors, were analyzed by SDS-PAGE, as shown in Figures 2.9 and 2.10, respectively.

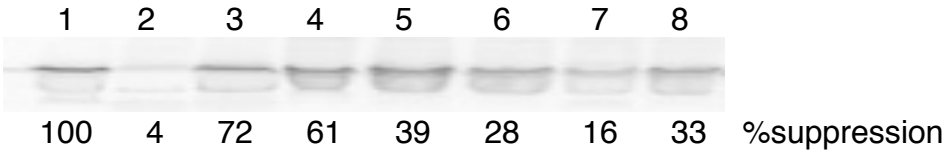


Figure 2.9. *In vitro* synthesis of DHFR utilizing monoacylated $tRNA_{CUA}$s to suppress a UAG codon at position 10 of DHFR mRNA. Lane 1, wild-type mRNA; Lanes 2-8, modified mRNA; Lane 2, no $tRNA_{CUA}$; lane 3, L-phenylalanyl-$tRNA_{CUA}$ (prepared from pdCpA); lane 4, L-phenylalanyl-$tRNA_{CUA}$; lane 5, 3,4-dimethoxy-L-phenylalanyl-$tRNA_{CUA}$; lane 6, 3-(1-naphthyl)-L-alanyl-$tRNA_{CUA}$; lane 7, 3-(2-naphthyl)-L-alanyl-$tRNA_{CUA}$; lane 8, L-4,4′-biphenylalanyl-$tRNA_{CUA}$.

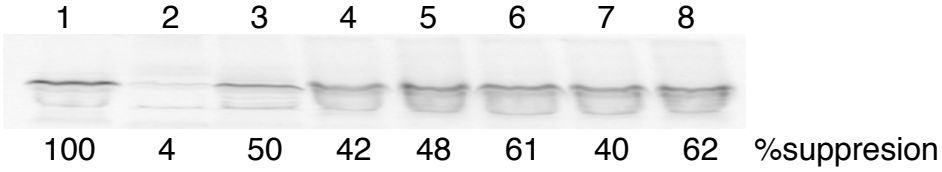


Figure 2.10. *In vitro* synthesis of DHFR utilizing bisaminoacylated $tRNA_{CUA}$s to suppress a UAG codon at position 10 of DHFR mRNA. Lane 1, wild-type mRNA; Lanes 2-8 modified mRNA; lane 2, no $tRNA_{CUA}$; lane 3, L-phenylalanyl-$tRNA_{CUA}$ (prepared from pdCpA); lane 4, bis-L-phenylalanyl-$tRNA_{CUA}$; lane 5, bis-L-3,4-dimethoxyphenylalanyl-$tRNA_{CUA}$; lane 6, bis-3-(1-naphthyl)-L-alanyl-$tRNA_{CUA}$; lane 7, bis-3-(2-naphthyl)-L-alanyl-$tRNA_{CUA}$; lane 8, bis-L-4,4′-biphenylalanyl-$tRNA_{CUA}$.

2.3. Discussion

Early studies demonstrated that the dinucleotide pdCpA is capable of being esterified at both the 2' and 3'-OH groups to form a bisaminoacyl derivative that can ultimately be activated tRNAs in tandem, as outlined in Fig. 2.1.[17] Additionally, it was shown that bisactivated tRNAs are able to participate in protein synthesis much more efficiently than their monoactivated counterparts. When *in vitro* protein translation experiments were carried out under conditions limiting for suppressor tRNA, twice as much protein was produced by bisaminoacylated tRNAs than monoaminoacylated tRNAs. Additionally, tRNAs bearing two amino acids also exhibit enhanced stability compared to their monoaminoacyl counterparts.[14-17,24] Molecular modeling studies suggest the enhanced stability of the bisaminoacylated tRNAs arises from the interaction of a hydronium ion between the aminoacyl moieties and the N-3 of the adenine nucleobase.[14]

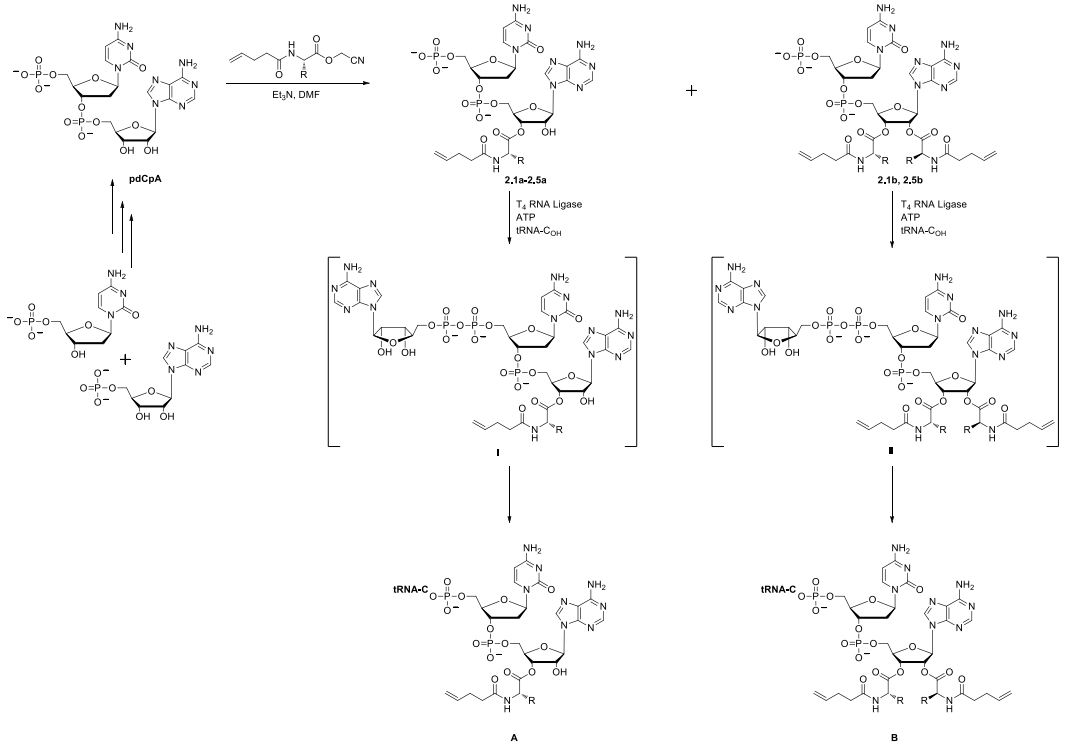
The advantages of bisaminoacylated tRNAs make them attractive compounds for use in *in vitro* and *in vivo* protein synthesis systems. However, compound **B** (Schemes 2.5 and 2.6) is difficult to obtain, as the tRNA is prepared from pdCpAs bearing large side chain amino acids. This has prevented the general use of these compounds and requires revision. Therefore, it is important to find a more facile route for the production of pre-bisaminoacylated dinucleotides capable of ligation to abbreviated suppressor tRNAs.

Misacylated suppressor tRNAs can be prepared by reacting chemically aminoacylated pdCpA with tRNAs lacking the 3'-terminal dinucleotide pCpA, via a T4

RNA ligase-mediated condensation.[11] T4 RNA ligase is an ATP dependent enzyme that catalyzes 3'→5' phosphodiester bond formation in RNA. The reaction sequence is initiated by a phosphoryl transfer reaction with ATP; in the presence of T4 RNA ligase an adenylated enzyme is generated, followed by subsequent transfer of adenosine to an RNA donor (pNn) to form a phosphoroanhydride (A5'pp5'N). Finally, the adenylate donor reacts with a 3' terminal hydroxyl group of the RNA molecule (abbreviated tRNA) to form a 3'→5' phosphodiester linkage. Hecht and co-workers first described "chemical aminoacylation" in which the phosphoroanhydride intermediate of the T4 reaction was utilized.[8] They demonstrated that the aminoacylated dinucleotide AppA was able to participate in the ligation reaction with T4 RNA ligase and abbreviated tRNA-CC_{OH} to produce misacylated tRNAs.[8]

There are two widely used methods for the misacylation of tRNAs: chemical and enzymatic aminoacylation. Some variations of these methods have been described. In addition, chemical aminoacylation in cationic micelles via ultrasonic agitation with pdCpA and amino acid, as well as a flexizyme technique have been applied.[34-41] Duffy and co-workers established a method of aminoacylation via lanthanum(III) and amino acid phosphate esters.[42] However, none of these methods produce significant amounts of bisaminoacylated products.[41,43] Meanwhile, chemical aminoacylation has been used to attach a variety of amino acids to the 3′-end of tRNA via pre mono- or bisaminoacylated pdCpAs. However, this method is not efficient enough for the production of a variety bisaminoacylated species.[27] It is hypothesized that the bisaminoacyl derivative is difficult to produce due to steric interactions between pdCpA and the amino acid (exemplified in Table 2.2).

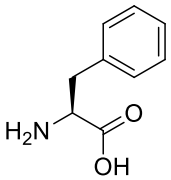
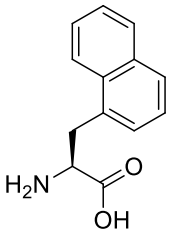
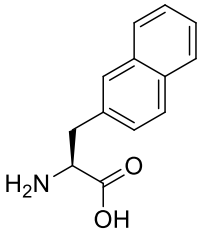
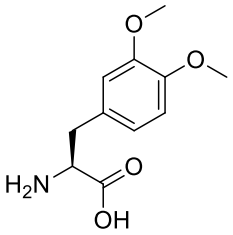
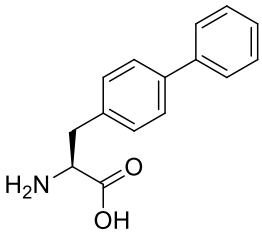
In the chemical route to produce bisaminoacylated pdCpAs, the larger amino acid side chains result in lower yields of the bisaminoacylated products. Additionally, the synthesis of the precursor pdCpA is rather laborious (scheme 2.5).



Scheme 2.5. Preparation of mono- and bisaminoacylated $tRNA_{CUA}$s by chemical aminoacylation and T4 RNA ligase mediated ligation using pdCpA and tRNA-C_{OH}.

However, once the key intermediates, **2.1a**-**2.5a** and **2.1b** and **2.5b,** required for the production of bisaminoacylated tRNAs have been prepared (Figure 2.4, Scheme 2.5), they are readily used in T4 RNA ligase-catalyzed condensation reactions with the abbreviated suppressor tRNA (tRNA-C_{OH}) lacking the 3'-terminal pCpA dinucleotide.

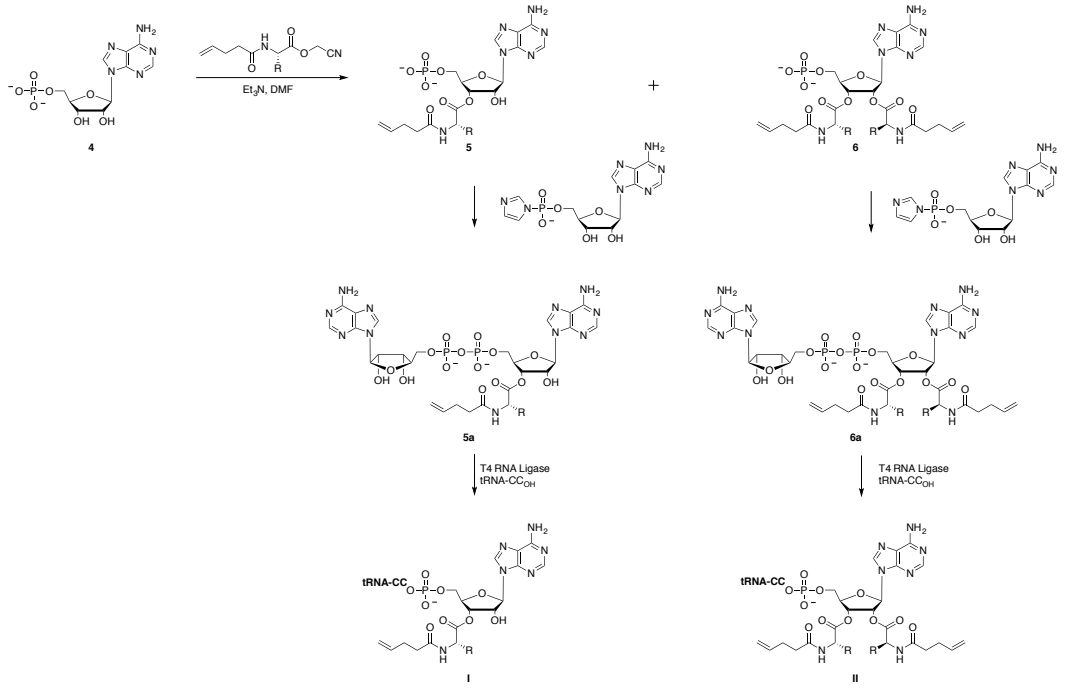
Table 2.2. Mono and bisaminoacylation of pdCpA and AMP.

	pdCpA		AMP	
	monoacyl	bisacyl	monoacyl	bisacyl
	60%	30%	30%	67%
	70%	0%	34%	54%
	65%	0%	28%	57%
	84%	0%	25%	67%
	50%	40%	21%	77%

We anticipated that as a substrate for aminoacylation, the mononucleotide AMP would relieve a portion of the steric restriction due to the absence of the second

nucleobase present in pdCpA. We expected that this would result in enhanced bisacylation. Hecht and co-workers already demonstrated that aminoacylated AppA can act as an aminoacyladenylate donor in a T4 RNA ligation reaction with abbreviated tRNA for the production of misacylated tRNAs.[8] Therefore, the fidelity of acylation with pdCpA and AMP containing several bulky amino acids was examined to determine whether this new route would be more effective for obtaining bisactivated tRNAs (Table 2.2).

AMP is tolerant to structural modification and retains the ability to form AppA derivatives.[44] Therefore, we proposed a strategy in which AMP would first be aminoacylated, and then converted to the mono- and bisaminoacylated AppA derivatives, **2.11a-2.15a** and **2.11b-2.15b**. In the presence of T4 RNA ligase and an abbreviated tRNA (tRNA-CC_{OH}) lacking the 3′-terminal adenosine moiety, this will afford mono- and bisaminoacylated tRNAs, **A** and **B** (Scheme 2.6)



Scheme 2.6. Preparation of mono- and bisaminoacylated $tRNA_{CUA}$s by successive aminoacylation of AMP and T4 RNA ligase ligation of aminoacyl-AppA with abbreviated tRNA-CC_{OH}.

Transfer RNAs **I** and **II** (Schemes 2.5 and 2.6) should be accessible from either aminoacylated precursors, pdCpA or AMP. However, efficient aminoacylation with amino acids having relatively large side chains may require the utilization of AMP.

While there are many protecting groups utilized for amines, here it is essential to select one amenable to deprotection under conditions compatible with the stability of the final aminoacyl-tRNA product. The nitroveratryloxycarbonyl (NVOC) protecting group is commonly used as it can be removed by simple UV irradiation.[32] As the 4-pentenoyl group can be removed under mild conditions (aqueous I_2) without disturbing the integrity of other functional groups, we decided to employ the latter methodology.[25,45]

Amino acid derivatives were prepared as previously described[25] by first protecting the amine as the pentenoyl derivative, followed by activation of the acid as the cyanomethyl ester. The resulting amino acids (**2.11**-**2.15**) were then used to study the extent of aminoacylation of both the dinucleotide pdCpA and mononucleotide AMP (Table 2.2). All pdCpA and AMP derivatives were purified by C_{18} reverse phase HPLC using a gradient of 0 → 63% acetonitrile in 50 mM NH_4OAc, pH 4.5. In all cases the starting material was consumed, however, when aminoacylating pdCpA there was little or no production of the bisaminoacylated derivative as seen from the HPLC spectrum.

The *N*-pentenoyl protected bisaminoacylated dinucleotide derivatives were ligated to abbreviated suppressor $tRNA_{CUA}$ transcripts lacking the cytidine and adenosine (pdCpA) or adenosine (AppA) moieties normally present at the 3'-terminus of all tRNAs. The tRNA transcript was obtained by *in vitro* run-off transcription of a *Fok1*-linearized plasmid encoding the abbreviated $tRNA_{CUA}$ (Fig. 2).[20] The ligation reaction was catalyzed by T4 RNA ligase.[8,11] The formed *N*-pentenoyl bisacylated tRNAs were deprotected by treatment with aqueous iodine as described previously.[25,45]

To investigate the extent of aminoacylation of pdCpA and AMP the esterification was carried out with a variety of *N*-protected amino acids bearing bulky side chains, including L-phenylalanine, L-1-naphthylalanine, L-2-naphthylalanine, 3-(3,4-dimethoxyphenyl)-L-alanine, and L-biphenylalanine.

Initially, pdCpA and AMP were subjected to aminoacylation with phenylalanine to determine the extent of aminoacylation with a natural amino acid bearing a large side chain. As anticipated, aminoacylation of AMP yielded nearly two-fold more bisaminoacylated product than the dinucleotide pdCpA, 67% and 30%, respectively.

Consequently, a variety of non-natural amino acids bearing bulky side chains were employed for the same esterification reaction and the extent of aminoacylation was monitored over a period of three days. As expected, pdCpA afforded no bisaminoacylated product when 3-(3,4-dimethoxyphenyl)-L-alanine, L-1-naphthylalanine, or L-2-naphthylalanine were used. However, *N*-protected phenylalanine and biphenylalanine esterified both the 2' and 3'-OH groups of the ribose sugar. Conversely, the esterification of AMP with the various amino acids was more facile. Aminoacylation resulted in greater yields of the bisaminoacylated product as compared to the monoacylated adduct, in yields ranging from 55% to 77%.

It was substantially easier to obtain the bisacylated derivatives when using AMP as compared to pdCpA. While the monoacylation yields of pdCpA are quite high, it is presumably more difficult to acylate the second free ribose-OH which results in little bisacylated product. Chemical acylation of the sugar is believed to occur at the 2'-OH position.[46] However, the acylated nucleotide undergoes rapid equilibration to a mixture of 2' (3')-*O*-acyladenosine, primarily affording the 3'-*O*-acylated compound.[18] The mononucleotide, 5'-AMP, is most likely relieved of the steric interference caused by the presence of the cytidine base in pdCpA due to the free rotation about the C5'-*O* bond of adenosine, allowing more facile migration to occur. This suggests that monocleotide AMP can better tolerate structural modification than the dinucleotide pdCpA. This hypothesis is further supported by work describing the introduction of bulky groups at the C-8 position of adenosine without compromising its ability to form 5'→5' diphosphates.[44]

The preparation of AppA was realized by modifying a method published by Kanivarioti and co-workers.[28] Treatment of aminoacylated AMP with excess

imidazolium AMP sodium salt proceeded in very good yields (scheme 2.6). This dinucleotide resembles an obligatory intermediate formed in T4 RNA ligase reactions. Therefore, no ATP is required for the ligation reaction (Scheme 2.6).[47]

Ligation efficiencies of aminoacyl-pdCpAs with tRNA-C_{OH}s are easily estimated by SDS-PAGE. However, ligation of the AppA derivative and the 75nt tRNA (tRNA-CC_{OH}) was less obvious. While the product of the ligation reaction contained an additional nucleotide and a hydrophobic amino acid, there was no apparent difference in mobility between the product and the 75 nt tRNA in SDS-PAGE analysis.

Verifying the addition of a single nucleotide was rather difficult due to an ambiguous result from SDS-PAGE analysis, even though hydrophobic amino acids were attached to the tRNAs.

The ability of the bisacylated tRNAs to participate in protein synthesis was investigated *in vitro* in an *E. coli* S30 coupled transcription-translation system in which the DHFR gene has a UAG codon at a position corresponding to Val10. Five different amino acids were incorporated from two sets of aminoacyl tRNAs (Scheme 2.4). The syntheses of full length DHFR in the presence of mono- and bisaminoacylated-tRNAs are illustrated in Figures 2.7 and 2.8, respectively. Both mono- and bisaminoacylated-tRNAs were shown to effectively suppress the UAG codon, resulting in production of comparable amounts of full length DHFR. Some variations were observed while employing mono- versus bisaminoacylated-tRNAs for incorporating the same amino acid. In the case of L-1-naphthylalanine, L-2-naphthylalanine and L-biphenylalanine, the use of bisaminoacylated-tRNAs resulted in the production of approximately 2-3 times more DHFR.

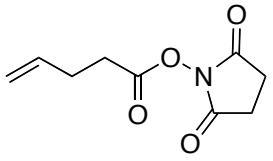
2.4. Experimental

2.4.1. General methods and materials

L-Phenylalanine was purchased from Acros Organics, 3-(3,4-dimethoxyphenyl)-L-alanine was purchased from Aldrich Chemical Company and all other amino acids were purchased from Peptech Corporation. Reagents and solvents for chemical synthesis were purchased from Aldrich Chemical Co. or Sigma Chemical Co. and were used without further purification. All reactions involving air or moisture-sensitive reagents or intermediates were carried out under argon with anhydrous or distilled solvents. Analytical TLC was performed using Silicycle silica gel, 60Å F_{254} plates (0.25 mm), and was visualized by UV irradiation (254 nm). Flash chromatography was performed using Silicycle silica gel (40-60 mesh). ^{1}H and ^{13}C NMR spectra were obtained using a Varian 400 MHz NMR spectrometer. Chemical shifts are reported in parts per million (ppm, δ) referenced to the residual ^{1}H of the solvent ($CDCl_3$, δ 7.26). ^{13}C NMR spectra were referenced to the residual ^{13}C resonance of the solvent ($CDCl_3$, δ 77.16). Splitting patterns are designated as follows: s, singlet; d, doublet; dd, doublet of doublets; t, triplet; q, quartet; m, multiplet. High-resolution mass spectra were obtained at the Arizona State University CLAS High Resolution Mass Spectrometry Facility or Michigan State University Mass Spectrometry Facility. HPLC purification was performed with a Waters 600 pump, Varian ProStar 340 detector, and Grace Econosil C_{18} column (250 x 10 mm, 5 μm). The tetra-*n*-butylammonium (TBA) salt of pdCpA and AMP derivatives were prepared using Dowex 50W×8, 200-400 mesh activated in its tetra-*n*-butylammonium form. Gels were visualized and quantified by phosphorimager analysis, which was

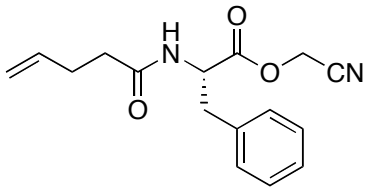
carried out using an Amersham Biosciences Storm 820 equipped with ImageQuant version 5.0 software. The pixel density of the image was directly related to the amount of radioactivity present in the sample by using a calibrated phosphorimager screen.

2.4.2. Synthesis of *N*-4-pentenoyl cyanomethylester amino acids



4-pentenoyloxy succinimide (2.16)[25]

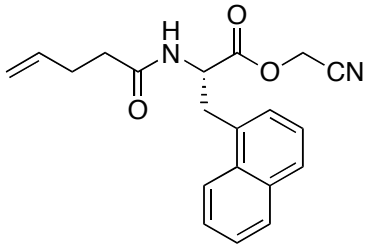
To a solution containing 5.0 mL (49.0 mmol) of pentenoic acid and 5.64 g (49.0 mmol) of *N*-hydroxysuccinimide in 100 mL of CH_2Cl_2, was added 10.3 g (50.0 mmol) of *N*,*N*′-dicyclohexylcarbodiimide. After stirring at room temperature for 2 hours, the reaction mixture was filtered and the filtrate was concentrated under diminished pressure. Crystallization of the crude product from ether–petroleum ether afforded 4-pentenoyloxy succinimide (**2.16**) as colorless needles: yield 7.63 g (79%); R_f 0.43 (1:1 ethyl acetate–hexane); mp 45-47 °C (47-48 °C).



***N*-4-pentenoyl-L-phenylalanine cyanomethyl ester (2.11)**[25]

To a solution containing 0.50 g (3.02 mmol) of phenylalanine and 0.51g (6.04 mmol) of $NaHCO_3$ in 18 mL of 1:1 dioxane–H_2O, was added 0.71 g (3.62 mmol) of **2.16**. The reaction mixture was stirred at room temperature for 16 hours. Eighty mL of 1 N

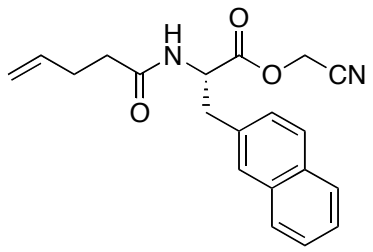
$NaHSO_4$ was then added to the reaction mixture. The aqueous layer was extracted with three 100-mL portions of ethyl acetate. The combined organic extract was dried ($MgSO_4$) and concentrated under diminished pressure. The crude product was then dissolved in 6 mL acetonitrile and to the solution was added 2.1 mL (15.1 mmol) of triethylamine and 0.93 mL (15.1 mmol) of chloroacetonitrile. The reaction mixture was stirred at room temperature for 20 hours. Eighty mL of ethyl acetate was then added to the reaction mixture. The organic layer was washed with three 75-mL portions of 1 N $NaHSO_4$. The organic layer was dried ($MgSO_4$) and concentrated under diminished pressure. The crude residue was purified on a silica gel column (45 x 2 cm); elution with 1:1 ethyl acetate–hexanes afforded *N*-pentenoyl-L-phenylalanine cyanomethyl ester (**2.11**) as a colorless solid: yield 0.76 g (88%); mp 52-53 °C (40-41 °C); R_f 0.45 (1:1 ethyl acetate–hexane); ^{1}H NMR δ 2.26-2.38 (m, 4H), 3.09-3.19 (m, 2H), 4.68, 4.79 (ABq, 2H, J = 15.6 Hz), 4.91-5.06 (m, 3H), 5.72-5.84 (m, 2H), 7.12-7.14 (m, 2H) and 7.29-7.35 (m, 3H); ^{13}C NMR δ 29.2, 35.3, 37.6, 48.8, 52.8, 105.0, 113.7, 115.9, 127.5, 128.9, 129.1, 134.9, 136.6, 170.4 and 172.0.



_N_-4-pentenoyl-L-1-naphthylalanine cyanomethyl ester (2.12) [41]

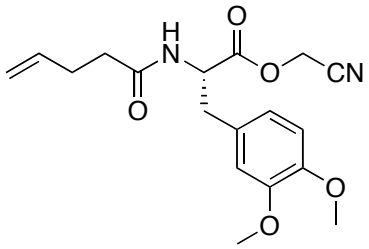
To a solution containing 0.10 g (0.46 mmol) of L-1-naphthylalanine and 0.80 g (0.92 mmol) of $NaHCO_3$ in 3 mL of 1:1 dioxane–H_2O, was added 0.11 g (0.56 mmol) of **2.16**. The reaction mixture was stirred at room temperature for 16 hours. Twelve mL of 1 N

$NaHSO_4$ was then added to the reaction mixture. The aqueous layer was extracted with three 15-mL portions of ethyl acetate. The combined organic extract was dried ($MgSO_4$) and concentrated under diminished pressure. The crude product was then dissolved in 1 mL of acetonitrile and to the solution was added 0.32 mL (2.3 mmol) of triethylamine and 0.14 mL (2.3 mmol) of chloroacetonitrile. The reaction mixture was stirred at room temperature for 20 hours. Twelve mL of ethyl acetate was then added to the reaction mixture. The organic layer was washed with three 12-mL portions of 1 N $NaHSO_4$. The organic layer was dried ($MgSO_4$) and concentrated under diminished pressure. The crude residue was purified on a silica gel column (30 x 2 cm); elution with 1:1 ethyl acetate–hexanes afforded *N*-pentenoyl-L-1-naphthylalanine cyanomethyl ester (**2.12**) as an off-white solid: yield 0.11 g (70%); mp 95-97 °C; R_f 0.45 (1:1 ethyl acetate–hexane); ^{1}H NMR δ 2.18-2.29 (m, 4H), 3.50-3.60 (m, 2H), 4.55, 4.63 (ABq, 2H, J = 15.7 Hz), 4.92-5.03 (m, 3H), 5.67-5.74 (m, 1H), 6.02 (d, 1H, J = 6.9 Hz), 7.27 (d, 1H, J = 7.0 Hz), 7.41 (t, 1H, J = 7.9 Hz), 7.48-7.57 (m, 2H), 7.78 (d, 1H, J = 8.2 Hz), 7.88 (d, 1H, J = 7.9 Hz) and 8.03 (d, 1H, J = 8.3 Hz); ^{13}C NMR δ 29.1, 34.8, 35.2, 48.8, 52.9, 113.7, 115.8, 123.1, 125.4, 126.0, 126.6, 127.5, 128.4, 129.0, 131.4, 131.9, 133.9, 136.6, 170.6 and 172.2; mass spectrum (APCI) *m*/*z* 337.1561 $(M+H)^+$ ($C_{20}H_{21}N_2O_3$ requires *m/z* 337.1552).



_N_-4-pentenoyl-L-2-naphthylalanine cyanomethyl ester (2.13) [41,43]

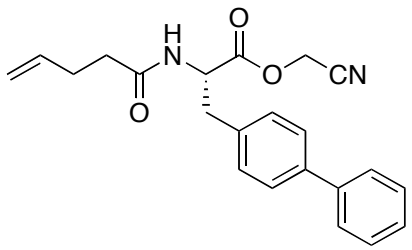
To a solution containing 57 mg (0.27 mmol) of L-2-naphthylalanine and 0.47 mg (0.54 mmol) of $NaHCO_3$ in 2 mL of 1:1 dioxane–H_2O, was added 63 mg (0.32 mmol) of **2.16**. The reaction mixture was stirred at room temperature for 16 hours. Two mL of 1 N $NaHSO_4$ was then added to the reaction mixture. The aqueous layer was extracted with three 2-mL portions of ethyl acetate. The combined organic extract was dried ($MgSO_4$) and concentrated under diminished pressure. The crude product was then dissolved in 0.5 mL acetonitrile and to the solution was added 187 μL (1.3 mmol) of triethylamine and 83 μL (1.3 mmol) of chloroacetonitrile. The reaction mixture was stirred at room temperature for 20 hours. Two mL of ethyl acetate was then added to the reaction mixture. The organic layer was washed with three 2-mL portions of 1 N $NaHSO_4$. The organic layer was dried ($MgSO_4$) and concentrated under diminished pressure. The crude residue was purified on a silica gel column (15 x 2 cm); elution with 1:1 ethyl acetate–hexanes afforded *N*-pentenoyl-L-2-naphthylalanine cyanomethyl ester (**2.13**) an off–white solid: yield 0.058 g (65%); mp 89-90 °C; R_f 0.45 (1:1 ethyl acetate–hexane); ^{1}H NMR δ 2.25-2.34 (m, 4H), 3.27-3.33 (m, 2H), 4.67, 4.80 (ABq, 2H, J = 16 Hz), 4.78-5.04 (m, 3H), 5.69-5.79 (m, 1H), 5.91 (d, 1H, J = 7.6 Hz), 7.48-7.50 (m, 2H), 7.60 (s, 1H) and 7.80-7.84 (m, 3H); ^{13}C NMR δ 29.2, 35.3, 37.7, 48.9, 52.9, 113.8, 115.8, 126.1, 126.4, 126.9, 127.6, 127.7, 128.0, 128.7, 132.5, 132.6, 133.4, 136.6, 170.4 and 172.2; mass spectrum (APCI) *m/z* 337.1546 $(M+H)^+$ ($C_{20}H_{21}N_2O_3$ requires *m/z* 337.1552).



***N*-4-pentenoyl-3-(3,4-dimethoxyphenyl)-L-alanine cyanomethyl ester (2.14)**

To a solution containing 0.20 g (0.89 mmol) of 3-(3,4-dimethoxyphenyl)-L-alanine and 0.15 g (1.78 mmol) of $NaHCO_3$ in 5.5 mL of 1:1 dioxane–H_2O, was added 0.21 g (1.1 mmol) of **2.16**. The reaction mixture was stirred at room temperature for 16 hours. Twenty-four mL of 1 N $NaHSO_4$ was then added to the reaction mixture. The aqueous layer was extracted with three 30-mL portions of ethyl acetate. The combined organic extract was dried ($MgSO_4$) and concentrated under diminished pressure. The crude product was then dissolved in 6 mL acetonitrile and to the solution was added 0.62 mL (4.45 mmol) of triethylamine and 0.27 mL (4.45 mmol) of chloroacetonitrile. The reaction mixture was stirred at room temperature for 20 hours. Twenty-four mL of ethyl acetate was then added to the reaction mixture. The organic layer was then washed with three 22-mL portions of 1 N $NaHSO_4$. The organic layer was dried ($MgSO_4$) and concentrated under diminished pressure. The crude residue was purified on a silica gel column (30 x 2 cm); elution with 1:1 ethyl acetate–hexanes afforded *N*-pentenoyl-3-(3,4-dimethoxyphenyl)-L-alanine cyanomethyl ester (**2.14**) as a colorless solid: yield 0.24 g (77%); mp 113-115 °C; R_f 0.50 (1:1 ethyl acetate–hexane); ^{1}H NMR δ 2.20-2.30 (m, 4H), 2.95-3.06 (m, 2H), 3.78 (s, 6H), 4.62, 4.74 (ABq, 2H, J = 16 Hz), 4.79 (q, 1H, J = 6.4 and 14 Hz), 4.90-4.99 (m, 2H), 5.65-75 (m, 1H), 6.19 (d, 1H, J = 7.6 Hz), 6.60-6.63 (m, 2H) and 6.74 (d, J = 8.1 Hz, 1H); ^{13}C NMR δ 29.0, 35.0, 36.8, 48.6, 52.7, 55.6, 55.7, 111.2,

112.0, 113.8, 115.4, 121.1, 127.3, 136.4, 148.1, 148.9, 170.3 and 172.0; mass spectrum (ESI) *m/z* 347.1593 $(M+H)^+$ ($C_{18}H_{23}N_2O_5$ requires *m/z* 347.1607).

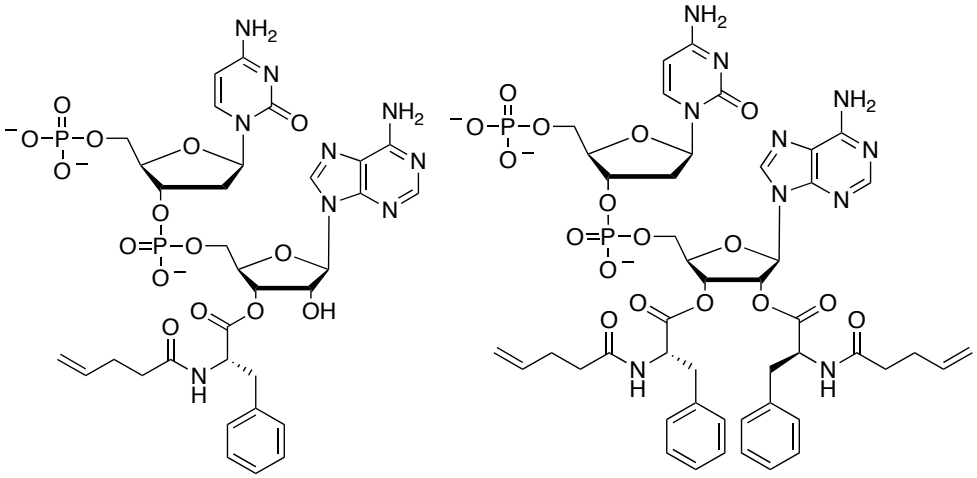


***N*-4-pentenoyl-L-4,4'-biphenylalanine cyanomethyl ester (2.15)**

To a solution containing 0.21 g (0.92 mmol) of L-4,4'-biphenylalanine and 0.16 g (1.84 mmol) of $NaHCO_3$ in 5.5 mL of 1:1 dioxane–H_2O, was added 0.20 g (1.02 mmol) of **2.16**. The reaction mixture was stirred at room temperature for 16 hours. Twenty-four mL of 1 N $NaHSO_4$ was then added to the reaction mixture. The aqueous layer was extracted with three 30-mL portions of ethyl acetate. The combined organic extract was dried ($MgSO_4$) and concentrated under diminished pressure. The crude product was then dissolved in 6 mL acetonitrile and to the solution was added 0.64 mL (4.61 mmol) of triethylamine and 0.28 mL (4.61 mmol) of chloroacetonitrile. The reaction mixture was stirred at room temperature for 20 hours. Twenty-four mL of ethyl acetate was then added to the reaction mixture. The organic layer was then washed with three 23-mL portions of 1 N $NaHSO_4$. The organic layer was dried ($MgSO_4$) and concentrated under diminished pressure. The crude residue was purified on a silica gel column (30 x 2 cm); elution with 1:1 ethyl acetate–hexanes afforded *N*-pentenoyl-L-phenylalanine cyanomethyl ester (**2.11**) as a colorless solid: yield 0.19 g (56%); mp 129-133 °C; R_f 0.45 (1:1 ethyl acetate–hexane); ^{1}H NMR δ 2.28-2.40 (m, 4H), 3.14-3.24 (m, 2H), 4.71, 4.82

(ABq, 2H, J = 15.6 Hz), 4.88-5.06 (m, 3H), 5.73-5.83 (m, 1H), 5.88 (d, 1H, J = 7.6 Hz), 7.19 (d, 2H, J = 7.2 Hz), 7.33-7.37 (m, 1H), 7.43 (t, 2H, J = 7.8) and 7.55-7.59 (m, 4H); ^{13}C NMR δ 29.2, 35.4, 37.3, 48.9, 52.8, 113.7, 115.9, 127.0, 127.4, 127.6, 128.8, 129.6, 133.9, 136.6, 140.4, 140.5, 170.4 and 172.0; mass spectrum (APCI) *m/z* 363.1723 $(M+H)^+$ ($C_{22}H_{23}N_2O_3$ requires *m/z* 363.1709).

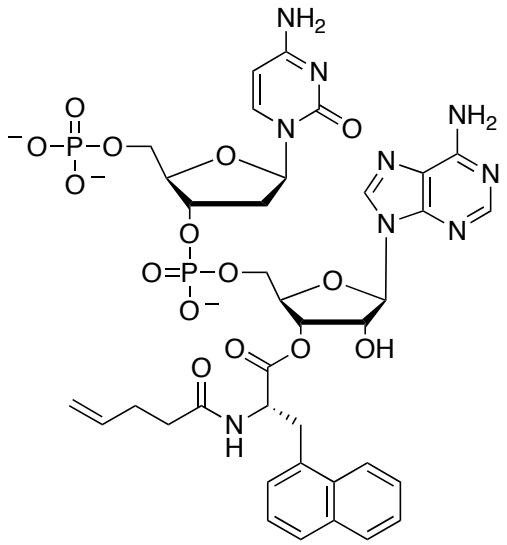
2.4.3. Synthesis of aminoacyl-pdCpA derivatives



Mono-2'(3')-*O*-(L-phenylalanyl)-pdCpA (2.1a) and Bis-2',3'-*O*-(L-phenylalanyl)-pdCpA (2.1b)

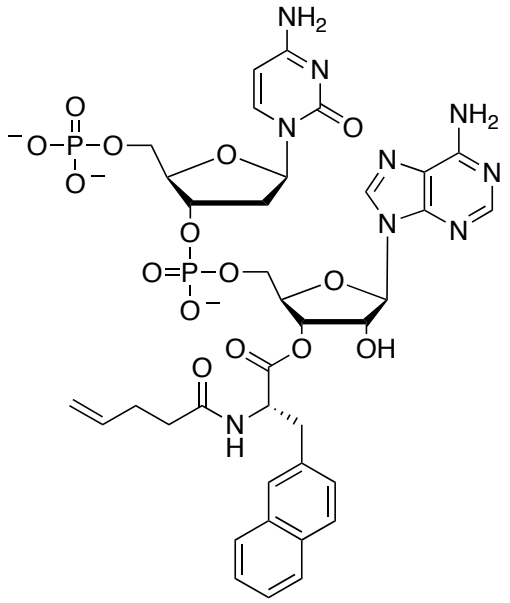
To a conical vial containing 5.0 mg (3.7 μmol) of TBA-pdCpA in freshly distilled DMF was added 13 mg (46 μmol) of **2.11** and 4.4 μL (32 μmol) of triethylamine. After stirring for three days at room temperature the reaction was purified by reverse phase HPLC using a gradient of 1%→65% acetonitrile in 50 mM ammonium acetate buffer (pH = 4.5) over 45 minutes to afford monoaminoacylated pdCpA **(2.1a)**: yield 1.9 mg (60%); t_R 24

minutes; mass spectrum (MALDI) *m/z* 866.2 $(M+H)^+$ (theoretical *m/z* 866.6); and bisaminoacylated pdCpA **(2.1b)**: yield 1.2 mg (30%); t_R 30.6 minutes; mass spectrum (ESI) *m/z* 1095.3405 $(M+H)^+$ ($C_{47}H_{57}N_{10}O_{17}P_2$ requires *m/z* 1095.3378).



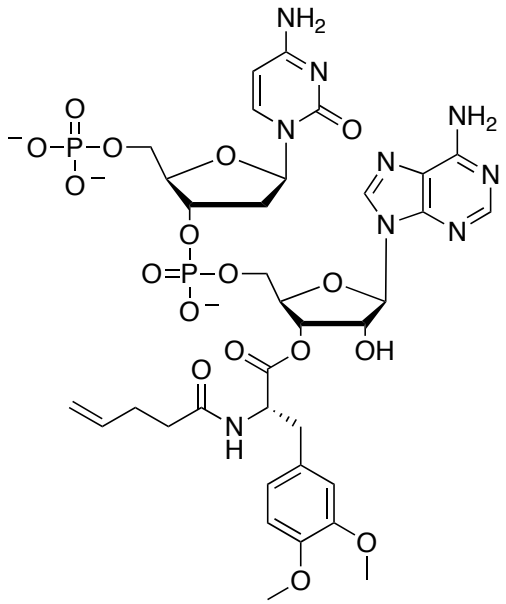
Mono-2'(3')-*O*-(L-2-naphthylalanyl)-pdCpA (2.2a)[43]

To a conical vial containing 5.0 mg (3.7 μmol) of TBA-pdCpA in freshly distilled DMF was added 16 mg (46 μmol) of **2.12** and 4.4 μL (32 μmol) of triethylamine. After stirring for three days at room temperature the reaction was purified by reverse phase HPLC using a gradient of 1%→65% acetonitrile in 50 mM ammonium acetate buffer (pH = 4.5) over 45 minutes to afford monoaminoacylated pdCpA **(2.2a)**: yield 2.4 mg (70%); t_R 24.0 minutes; mass spectrum (ESI) *m/z* 916.2410 $(M+H)^+$ ($C_{37}H_{44}N_9O_{15}P_2$ requires *m/z* 916.2432).



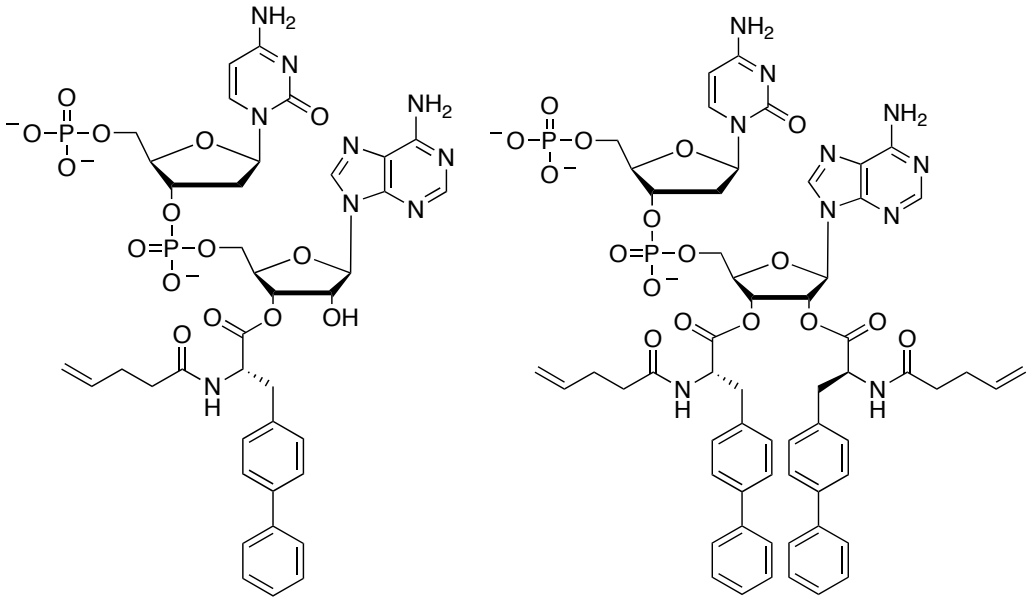
Mono-2'(3')-*O*-(L-2-naphthylalanyl)-pdCpA (2.3a)[41,43]

To a conical vial containing 8.3 mg (6.1 μmol) of TBA-pdCpA in freshly distilled DMF was added 26 mg (77 μmol) of **2.13** and 7.3 μL (52 μmol) of triethylamine. After stirring for three days at room temperature the reaction was purified by reverse phase HPLC using a gradient of 1%→65% acetonitrile in 50 mM ammonium acetate buffer (pH = 4.5) over 45 minutes to afford monoaminoacylated pdCpA **(2.3a)**: yield 3.6 mg (65%); t_R 22.0 minutes; mass spectrum (ESI) *m/z* 916.2422 $(M+H)^+$ ($C_{37}H_{44}N_9O_{15}P_2$ requires *m/z* 916.2432).



Mono-2'(3')-*O*-(3-(3,4-dimethoxyphenyl)-L-alanyl)-pdCpA (2.4a)

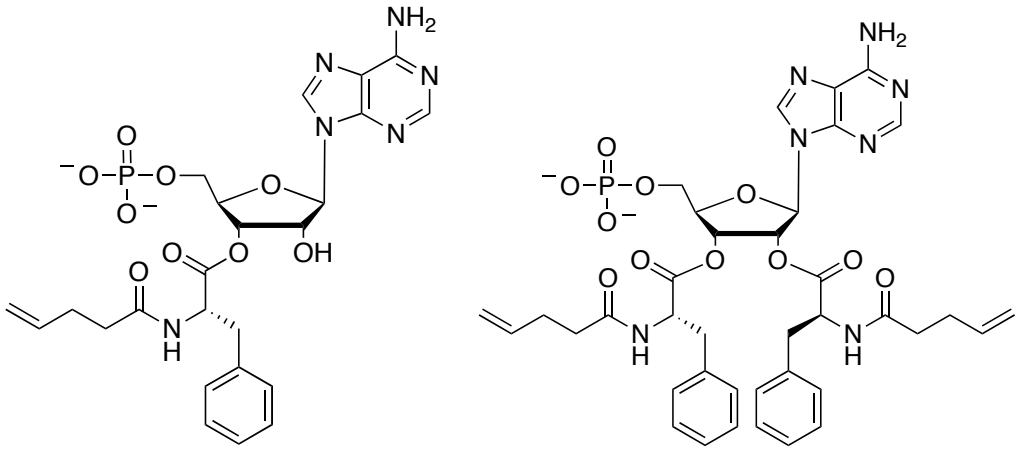
To a conical vial containing 5.0 mg (3.7 μmol) of TBA-pdCpA in freshly distilled DMF was added 16 mg (46 μmol) of **2.11** and 4.4 μL (32 μmol) of triethylamine. After stirring for three days at room temperature the reaction was purified by reverse phase HPLC using a gradient of 1%→65% acetonitrile in 50 mM ammonium acetate buffer (pH = 4.5) over 45 minutes to afford monoaminoacylated pdCpA **(2.4a)**: yield 2.9 mg (84%); t_R 28.6 minutes; mass spectrum (ESI) *m/z* 926.2451 $(M+H)^+$ ($C_{37}H_{44}N_9O_{15}P_2$ requires *m/z* 926.2487).



Mono-2'(3')-*O*-(L-4,4'-biphenylalanyl)-pdCpA (2.5a) and Bis-2',3'-*O*-(L-4,4'-biphenylalanyl)-pdCpA (2.5b)

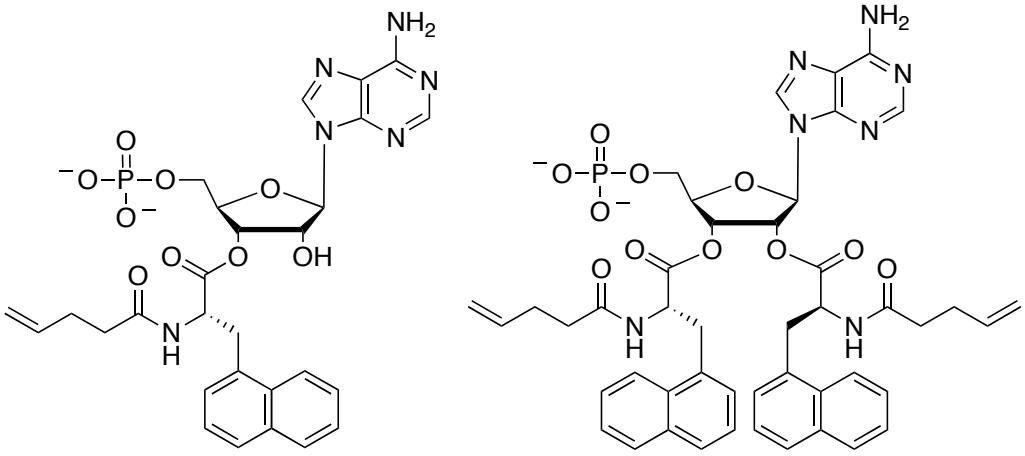
To a conical vial containing 5.0 mg (3.7 μmol) of TBA-pdCpA in freshly distilled DMF was added 17 mg (46 μmol) of **2.11** and 4.4 μL (32 μmol) of triethylamine. After stirring for three days at room temperature the reaction was purified by reverse phase HPLC using a gradient of 1%→65% acetonitrile in 50 mM ammonium acetate buffer (pH = 4.5) over 45 minutes to afford monoaminoacylated pdCpA (**2.5a)**: yield 1.7 mg (50%); t_R 27.9 minutes; mass spectrum (ESI) *m/z* 940.2 $(M-H)^-$ (theoretical *m/z* 940.2); and bisaminoacylated pdCpA **(2.5b)**: yield 1.8 mg (40%); t_R 37.6 minutes; mass spectrum (ESI) *m/z* 1245.4 $(M-H)^-$ (theoretical *m/z* 1245.4).

2.4.4. Synthesis of aminoacyl-AMP derivatives



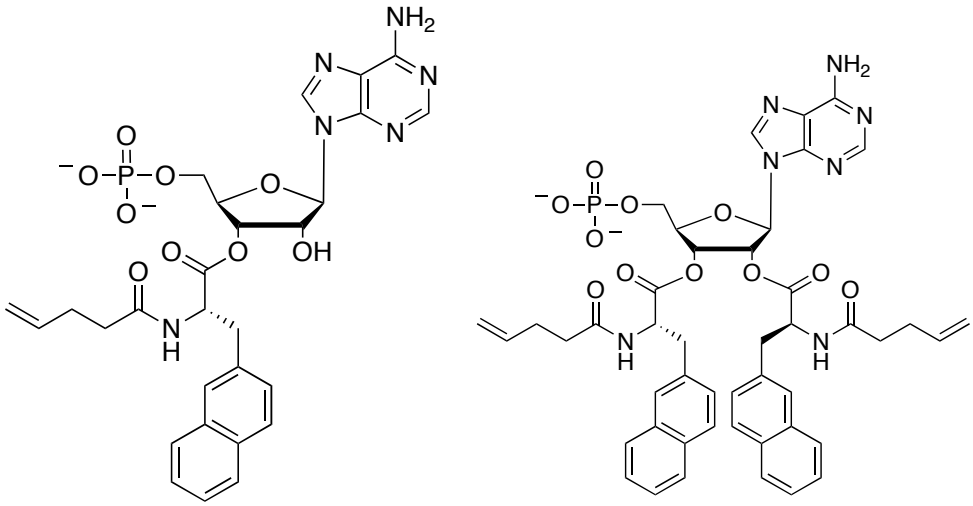
Mono-2'(3')-*O*-phenylalanyl-AMP (2.6a) and Bis-2',3'-phenylalanyl-AMP (2.6b)

To a conical vial containing 6.0 mg (7.2 μmol) TBA-AMP in freshly distilled DMF was added 26 mg (91 μmol) of **2.11** and 8.6 μL (62 μmol) of triethylamine. After stirring for three days at room temperature the reaction was purified by reverse phase HPLC using a gradient of 1%→65% acetonitrile in 50 mM ammonium acetate buffer (pH = 4.5) over 45 minutes to afford monoaminoacylated AMP **(2.6a**): yield 1.3 mg (30%); t_R 20.6 minutes; mass spectrum (ESI) *m/z* 599.1622 $(M+Na)^+$ ($C_{24}H_{30}N_6NaO_9P$ requires *m/z* 599.1631); and bisaminoacylated AMP **(2.6b)**: yield 3.9 mg (67%); t_R 26.0 minutes; mass spectrum (ESI) *m/z* 806.2880 $(M+H)^+$ ($C_{38}H_{45}N_7O_{11}P$ requires *m/z* 806.2915).



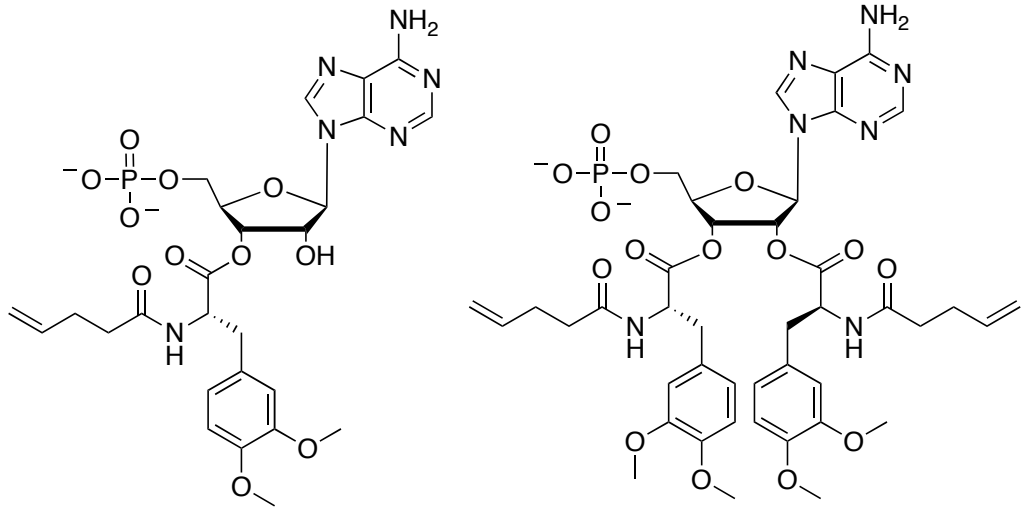
Mono-2'(3')-*O*-(L-1-naphthylalanyl)-AMP (2.7a) and Bis-2',3'-*O*-(L-1-naphthylalanyl)-AMP (2.7b)[41,43]

To a conical vial containing 6.0 mg (7.2 μmol) of TBA-AMP in freshly distilled DMF was added 31 mg (91 μmol) of **2.12** and 8.6 μL (62 μmol) of triethylamine. After stirring for three days at room temperature the reaction was purified by reverse phase HPLC using a gradient of 1%→65% acetonitrile in 50 mM ammonium acetate buffer (pH = 4.5) over 45 minutes to afford monoaminoacylated AMP **(2.7a)**: yield 1.5 mg (34%); t_R 24.5 minutes; mass spectrum (ESI) *m/z* 627.1946 $(M+H)^+$ ($C_{28}H_{32}N_6O_9P$ requires *m/z* 627.1968); and bisaminoacylated AMP **(2.7b)**; yield 3.5 mg (54%); t_R 31.0 minutes; mass spectrum (ESI) *m/z* 906.3204 $(M+H)^+$ ($C_{46}H_{49}N_7O_{11}P$ requires *m/z* 906.3228).



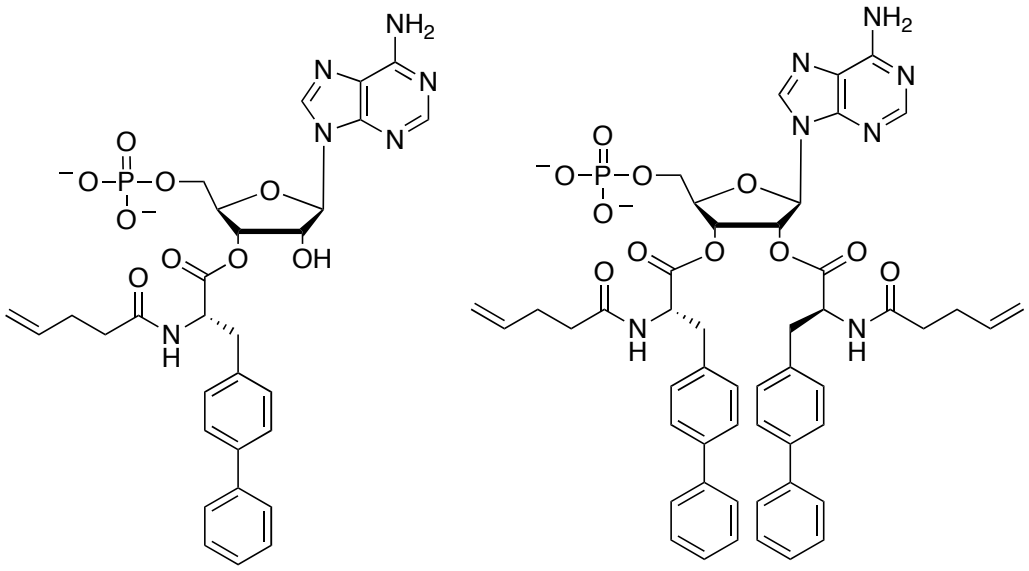
Mono-2'(3')-*O*-(L-2-naphthylalany)-AMP (2.8a)[41,43] and Bis-2',3'-*O*-(L-2-naphthylalanyl)-AMP (2.8b)

To a conical vial containing 5.0 mg (6.0 μmol) of TBA-AMP in freshly distilled DMF was added 26 mg (76 μmol) of **2.13** and 7.2 μL (52 μmol) of triethylamine. After stirring for three days at room temperature the reaction was purified by reverse phase HPLC using a gradient of 1%→65% acetonitrile in 50 mM ammonium acetate buffer (pH = 4.5) over 45 minutes to afford monoaminoacylated AMP **(2.8a)**: yield 1.1 mg (28%); t_R 22.5 minutes; mass spectrum (ESI) *m/z* 627.1960 $(M+H)^+$ ($C_{28}H_{32}N_6O_9P$ requires 627.1968); and bisaminoacylated AMP **(2.8b)**: yield 3.1 mg (57%); t_R 30.8 minutes; mass spectrum (ESI) *m/z* 906.3 $(M+H)^+$ (theoretical *m/z* 906.3).



Mono-2'(3')-*O*-(3-(3,4-dimethoxyphenylalanyl)-AMP (2.9a) and Bis-2'(3')-*O*-(3-(3,4-dimethoxyphenylalanyl)-AMP (2.9b)

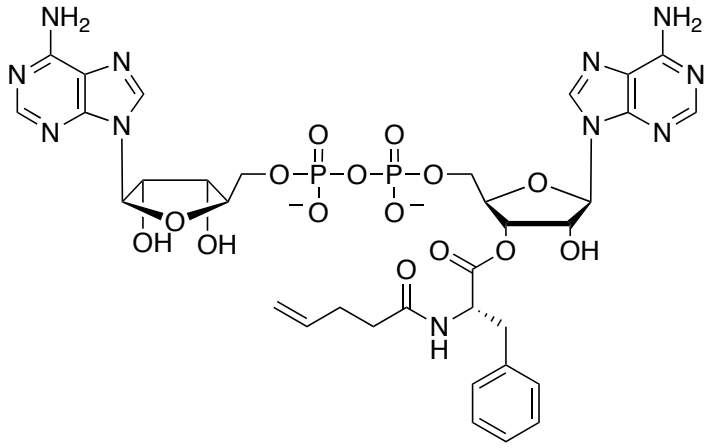
To a conical vial containing 6.0 mg (7.2 μmol) of TBA-AMP in freshly distilled DMF was added 32 mg (91 μmol) of **2.14** and 8.6 μL (62 μmol) of triethylamine. After stirring for three days at room temperature the reaction was purified by reverse phase HPLC using a gradient of 1%→65% acetonitrile in 50 mM ammonium acetate buffer (pH = 4.5) over 45 minutes to afford monoaminoacylated AMP **(2.9a)**: yield 1.1 mg (25%); t_R 22.1 minutes; mass spectrum (ESI) *m/z* 637.2039 $(M+H)^+$ ($C_{26}H_{34}N_6O_{11}P$ requires *m/z* 637.2023); and bisaminoacylated AMP **(2.9b)**: yield 4.6 mg (67%); t_R 29 minutes; mass spectrum (ESI) *m/z* 948.3187 $(M+Na)^+$ ($C_{42}H_{54}N_7NaO_{15}P$ requires *m/z* 948.3157).



Mono-2'(3')-*O*-(L-4,4'-biphenylalanine)-AMP (2.10a) and Bis-2',3'-*O*-(L-4,4'-biphenylalanyl)-AMP (2.10b)

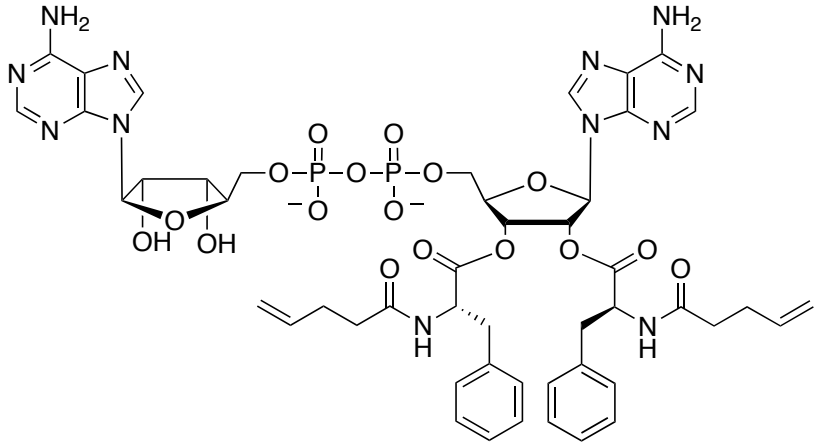
To a conical vial containing 4.0 mg (4.8 μmol) of TBA-AMP in freshly distilled DMF was added 22 mg (61 μmol) of **2.12** and 5.7 μL (41 μmol) of triethylamine. After stirring for three days at room temperature the reaction was purified by reverse phase HPLC using a gradient of 1%→65% acetonitrile in 50 mM ammonium acetate buffer (pH = 4.5) over 45 minutes to afford monoaminoacylated AMP **(2.10a)**: yield 0.7 mg (21%); t_R 23.8 minutes; HRMS (ESI) *m/z* 675.1954 $(M+Na)^+$ ($C_{30}H_{35}N_6NaO_6P$ requires 675.1944); and bisaminoacylated AMP **(2.10b)**: yield 3.6 mg (77%); t_R 32.2 minutes; mass spectrum (ESI) *m/z* 980.3327 $(M+Na)^+$ ($C_{50}H_{53}N_7NaO_{11}P$ requires *m/z* 980.3360).

2.4.5. Synthesis of aminoacyl-AppA analogues



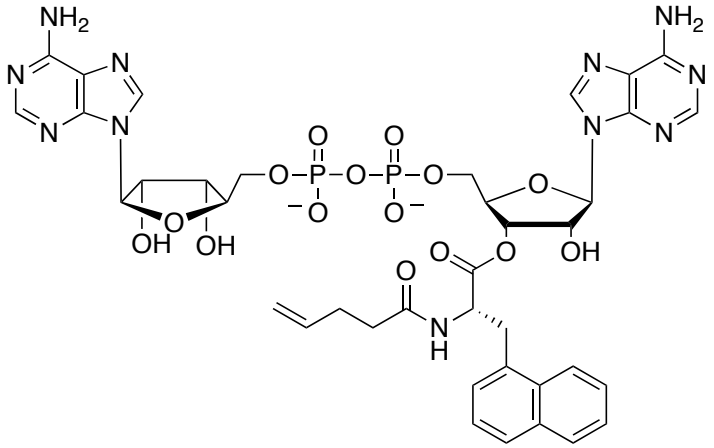
Mono-2'(3')-*O*-(L-phenylalanyl)-AppA (2.11a)

To a conical vial containing 2.2 mg (3.8 μmol) of **2.6a** in freshly distilled DMF was added 21 mg (50 μmol) of imidazolium AMP sodium salt (I-AMP). After stirring for three days at room temperature the reaction was purified by reverse phase HPLC using a gradient of 1%→65% acetonitrile in ammonium acetate buffer (pH = 4.5) over 45 minutes to afford compound **2.11a**: yield 3.0 mg (87%); t_R 16.0 minutes; mass spectrum (ESI) *m/z* 906.2343 $(M+H)^+$ ($C_{34}H_{42}N_{11}O_{15}P_2$ requires *m/z* 906.2337).



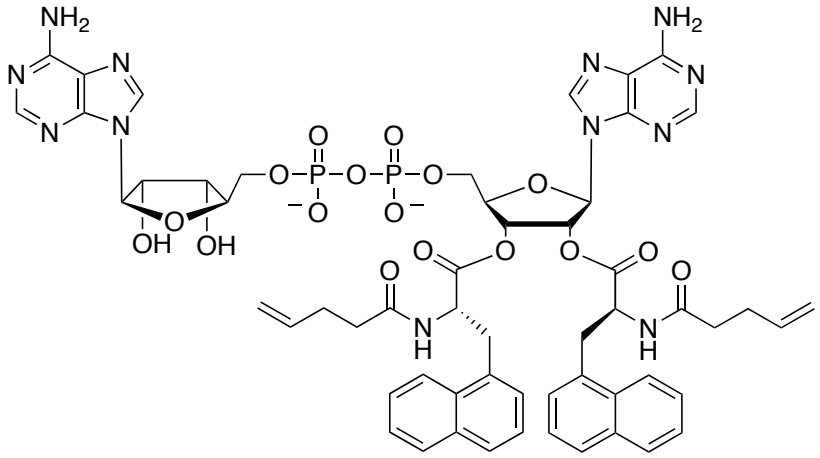
Bis-2',3'-*O*-(L-phenylalanyl)-AppA (2.11b)

To a conical vial containing 1.6 mg (1.2 μmol) of **2.6b** in freshly distilled DMF was added 6.8 mg (16 μmol) of I-AMP. After stirring for three days at room temperature the reaction was purified by reverse phase HPLC using a gradient of 1%→65% acetonitrile in ammonium acetate buffer (pH = 4.5) over 45 minutes to afford compound **2.11b**: yield 1.3 mg (90%); t_R 23.4 minutes; mass spectrum (ESI) *m/z* 1135.3459 $(M+H)^+$ ($C_{56}H_{61}N_{12}O_{17}P_2$ requires *m/z* 1135.3440).



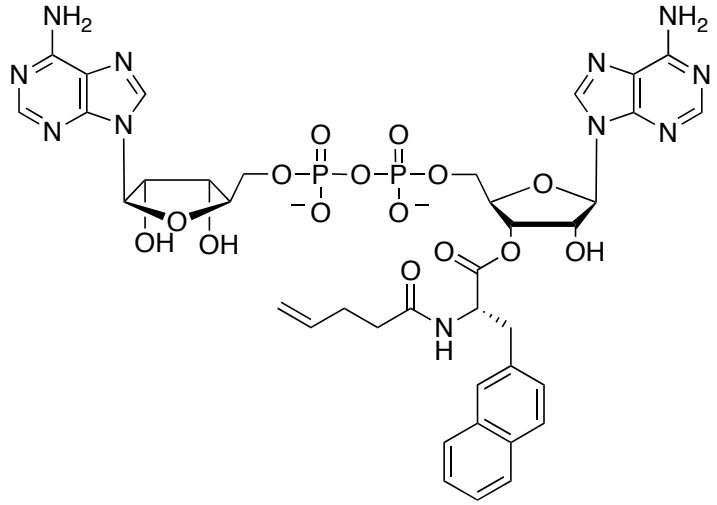
Mono-2'(3')-*O*-(L-1-naphthylalanyl)-AppA (2.12a)

To a conical vial containing 2.0 mg (3.2 μmol) of **2.7a** in freshly distilled DMF was added 17 mg (42 μmol) of I-AMP. After stirring for three days at room temperature the reaction was purified by reverse phase HPLC using a gradient of 1%→65% acetonitrile in ammonium acetate buffer (pH = 4.5) over 45 minutes to afford compound **2.12a**: yield 2.9 mg (94%); t_R 20.5 minutes; HRMS (ESI) *m/z* 956.2466 $(M+H)^+$ ($C_{38}H_{44}N_{11}O_{15}P_2$ requires *m/z* 956.2494).



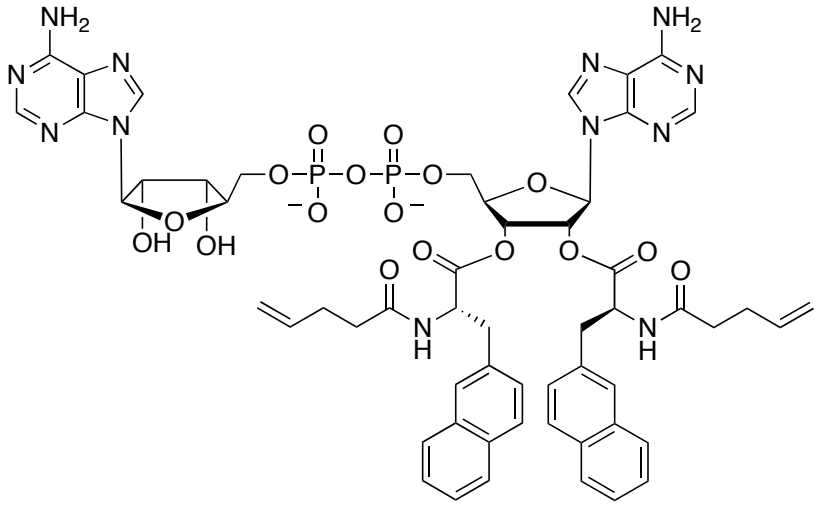
Bis-2',3'-*O*-(L-1-naphthylalanyl)-AppA (2.12b)

To a conical vial containing 3.8 mg (3.8 μmol) of **2.7b** in freshly distilled DMF was added 23 mg (55 μmol) of I-AMP. After stirring for three days at room temperature the reaction was purified by reverse phase HPLC using a gradient of 1%→65% acetonitrile in ammonium acetate buffer (pH = 4.5) over 45 minutes to afford compound **2.12a**: yield 4.5 mg (86%); t_R 26.0 minutes; mass spectrum (ESI) *m/z* 1235.3717 $(M+H)^+$ ($C_{56}H_{61}N_{12}O_{17}P_2$ requires *m/z* 1235.3753).



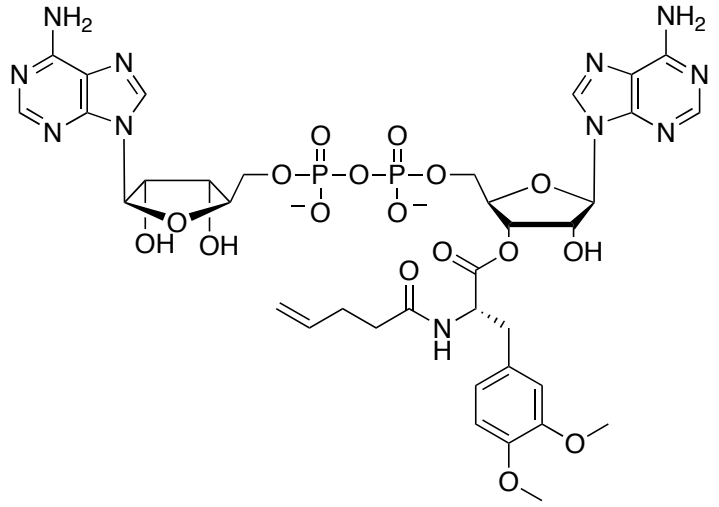
Mono-2'(3')-*O*-(L-2-naphthylalanyl)-AppA (2.13a)

To a conical vial containing 1.0 mg (1.6 μmol) of **2.8a** in freshly distilled DMF was added 8.7 mg (21 μmol) of I-AMP. After stirring for three days at room temperature the reaction was purified by reverse phase HPLC using a gradient of 1%→65% acetonitrile in ammonium acetate buffer (pH = 4.5) over 45 minutes to afford compound **2.13a**: yield 1.4 mg (93%); t_R 18.5 minutes; mass spectrum (ESI) *m/z* 956.2458 $(M+H)^+$ ($C_{38}H_{44}N_{11}O_{15}P_2$ requires *m/z* 956.2494).



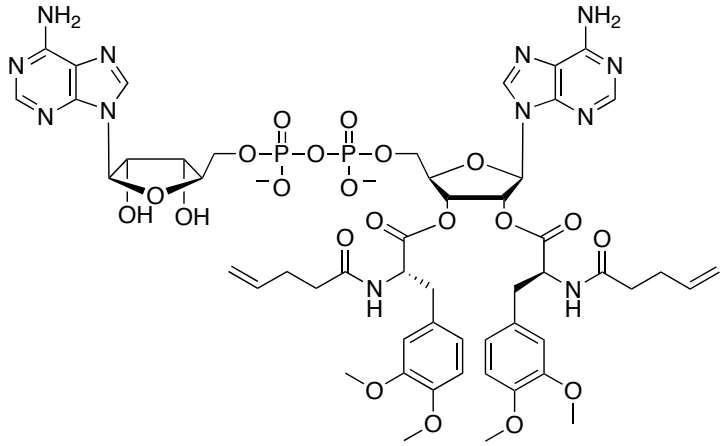
Bis-2',3'-*O*-(L-2-naphthylalanyl)-AppA (2.13b)

To a conical vial containing 3.0 mg (3.3 μmol) of **2.8b** in freshly distilled DMF was added 18 mg (43 μmol) of I-AMP. After stirring for three days at room temperature the reaction was purified by reverse phase HPLC using a gradient of 1%→65% acetonitrile in ammonium acetate buffer (pH = 4.5) over 45 minutes to afford compound **2.13b**: yield 3.6 mg (87%); t_R 24.6 minutes; mass spectrum (ESI) *m/z* 1235.3712 $(M+H)^+$ ($C_{56}H_{61}N_{12}O_{17}P_2$ requires *m/z* 1235.3753).



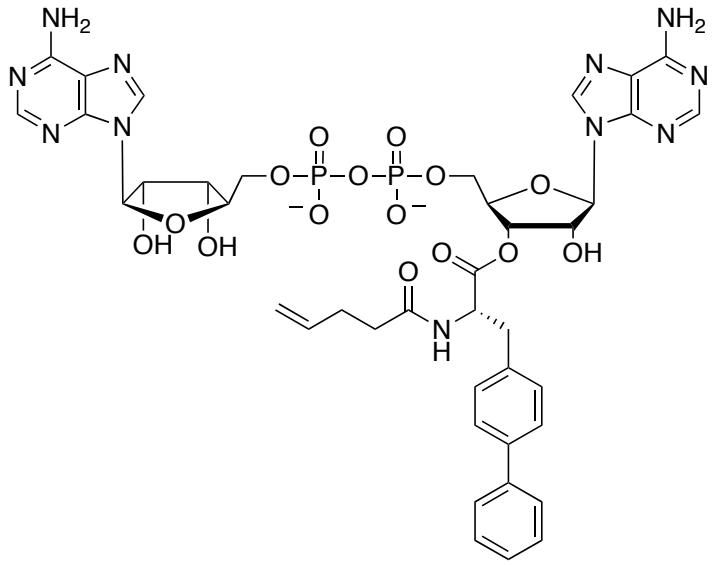
Mono-2'(3')-*O*-(3-(3,4-dimethoxyphenyalanyl)-AppA (2.14a)

To a conical vial containing 1.8 mg (2.8 μmol) of **2.9a** in freshly distilled DMF was added 15 mg (37 μmol) of I-AMP. After stirring for three days at room temperature the reaction was purified by reverse phase HPLC using a gradient of 1%→65% acetonitrile in ammonium acetate buffer (pH = 4.5) over 45 minutes to afford compound **2.14a**: yield 2.3 mg (85%); t_R 19.0 minutes; mass spectrum (ESI) *m/z* 966.2539 $(M+H)^+$ ($C_{36}H_{46}N_{11}O_{17}P_2$ requires *m/z* 966.2549).



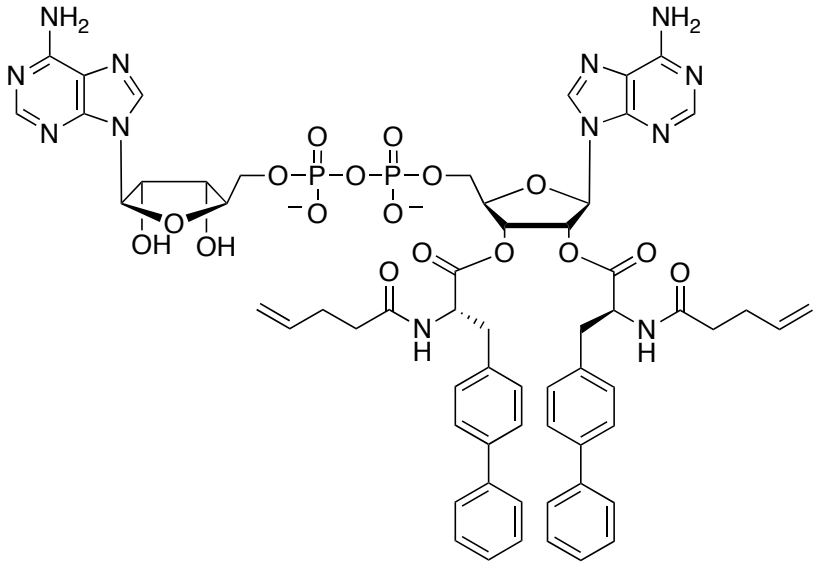
Bis-2',3'-*O*-(3-(3,4-dimethoxyphenylalanyl)-AppA (2.14b)

To a conical vial containing 4.0 mg (4.3 μmol) of **2.7b** in freshly distilled DMF was added 24 mg (56 μmol) of I-AMP. After stirring for three days at room temperature the reaction was purified by reverse phase HPLC using a gradient of 1%→65% acetonitrile in ammonium acetate buffer (pH = 4.5) over 45 minutes to afford compound **2.14b**: yield 4.4 mg (81%); t_R 22.4 minutes; mass spectrum (ESI) *m/z* 1255.3921 $(M+H)^+$ ($C_{52}H_{65}N_{12}O_{21}P_2$ requires *m/z* 1255.3862).



Mono-2'(3')-*O*-(L-4,4'-biphenylalanyl)-AppA (2.15a)

To a conical vial containing 1.2 mg (1.8 μmol) of **2.10a** in freshly distilled DMF was added 10 mg (24 μmol) of I-AMP. After stirring for three days at room temperature the reaction was purified by reverse phase HPLC using a gradient of 1%→65% acetonitrile in ammonium acetate buffer (pH = 4.5) over 45 minutes to afford compound **2.15a**: yield 1.3 mg (81%); t_R 20.7 minutes; mass spectrum (ESI) *m/z* 982.2695 $(M+H)^+$ ($C_{40}H_{46}N_{11}O_{15}P_2$ requires *m/z* 982.2650).



Bis-2',3'-*O*-(L-4,4'-biphenylalanyl)-AppA (2.15b)

To a conical vial containing 2.3 mg (2.4 μmol) of **2.10b** in freshly distilled DMF was added 13 mg (31 μmol) of I-AMP. After stirring for three days at room temperature the reaction was purified by reverse phase HPLC using a gradient of 1%→65% acetonitrile in ammonium acetate buffer (pH = 4.5) over 45 minutes to afford compound **2.15b**: yield 2.3 mg (79%); t_R 25.6 minutes; mass spectrum (ESI) *m/z* 1287.4072 $(M+H)^+$ ($C_{56}H_{61}N_{12}O_{17}P_2$ requires *m/z* 1287.4066).

2.4.6. Preparation of mono and bis-aminoacyl-tRNAs

Mono and bis-aminoacyl-tRNAs were prepared by a T4 RNA ligase-mediated ligation of the protected, chemically synthesized pdCpA and AppA derivatives (**2.1**-**2.5** and **2.11**-**2.15**, respectively), with the abbreviated suppressor tRNA-C_{OH}. The ligation reaction was carried out in 100 μL (total volume) of 100 mM HEPES, pH 7.5, containing 0.5 mM ATP, 15 mM $MgCl_2$, 100 μg of suppressor tRNA-C_{OH}, 2.0 A_{260} unit of *N*-pentenoyl-

protected mono- or bis-aminoacyl-pdCpA derivatives (5- to 10-fold molar excess), 15% DMSO and 200 units of T4 RNA ligase. After incubation at 37 ºC for 30 minutes, the reactions were quenched by the addition of 10 μL of 3 M NaOAc, pH 5.2, followed by 300 μL of ethanol. The reaction mixtures were incubated at -20 ºC for 30 minutes and centrifuged at 14,000 x g (30 minutes, 4 ºC). The supernatants were carefully decanted. The pellets were washed with 50 μL of 70% ethanol and dissolved in 50 μL of H_2O. The efficiencies of ligation when utilizing pdCpA were estimated by gel electrophoresis, pH 5.2 (100 V, 2 hours). To a solution containing 100 μg of aminoacyl-tRNA in 50 μL of water was added 12 μL of 25 mM I_2 (6.3 mg iodine was dissolved in 1 mL of 1:1 H_2O–THF). The reaction mixtures were incubated for 10 minutes at 25 ºC, then 6 μL of 3 M NaOAc was added. The tRNAs were precipitated by the addition of 200 μL of ethanol, then centrifuged at 14,000 x g (30 minutes, 4 ^{o}C) and the supernatants were carefully decanted. The pellets were washed with 50 μL of 70% ethanol and then dissolved in 30 μL of water.

2.4.7. *In vitro* translation of DHFR

In vitro translation was performed using a bacterial S-30 extract from *E. coli* strain BL21(DE3). The reaction mixture (20 μL total volume) containing 2 μg of pETDH10 plasmid, containing mutant DHFR gene (TAG codon in position corresponding to Val10), 8 μL of premix (35 mM Tris-acetate, pH 7.0, 190 mM potassium glutamate, 30 mM ammonium acetate, 2.0 mM dithiothreitol, 11 mM magnesium acetate, 20 mM phosphoenolpyruvate, 0.8 mg/mL of *E.coli* tRNA, 0.8 mM isopropyl β-D-

thiogalactopyranoside, 20 mM ATP and GTP, 5 mM CTP and UTP, 4 mM cAMP), 250 μM of each of the 20 amino acids, 5μCi of [^{35}S]-methionine, 10 μg/μL rifampicin, 30 μL of bacterial S-30 extract from *E.coli* strain BL21(DE3) and 0.6 μg of aminoacyl-$tRNA_{CUA}$ was incubated at 37 ºC for 45 minutes. As a control, the *in vitro* translation was also carried out in the presence of mono-aminoacyl-$tRNA_{CUA}$. Aliquots from *in vitro* translation mixtures were analyzed by SDS-PAGE following by quantification of the radioactive bands by phosphorimager analysis. Suppression efficiency was calculated as the percentage of the protein produced via nonsense codon suppression relative to the production of wild-type protein.

CHAPTER 3

SYNTHESIS OF CAGED DMT-THIOTHREONYL-DHFR ANALOGUES

3.1. Introduction

Many biologically significant peptides and peptidomimetics contain non-proteinogenic amino acids with functionalized side chains, and significant effort has been expended developing synthetic routes for the preparation of such compounds.[48-52] In fact, a common way to realize post-translational protein modifications is via the use of reagents that react with specific amino acid side chains. The introduction of a limited number of non-natural amino acids into predetermined positions in proteins can be an effective method to target specific sites for modification. In particular, a non-natural amino acid having a single thiol functional group can be introduced into a target protein and subsequently transformed by reagents that react with S atoms, such as maleimide esters.

The thiothreonines prepared in this study were incorporated into a modified DHFR protein lacking cysteine residues (csDHFR), facilitating site-specific protein modification at introduced thiothreonine and *allo*-thiothreonine residues via treatment with a thiol specific fluorophore, 7-diethylamino-3-(4'-maleimidylphenyl)-4-methylcoumarin (CPM). The fluorescent prosthetic group was introduced at the level of pdCpA, suppressor tRNA, or protein to define the optimal methodology for introducing specific functionalities into proteins.

As part of our ongoing efforts to enable the synthesis of modified proteins containing mechanistically interesting non-proteinogenic amino acids,[22,23,53-63] we focused on the synthesis of β-methylcysteine (thiothreonine). Thus, the syntheses of L-thiothreonine and L-*allo*-thiothreonine (Figure 3.1) were undertaken utilizing a novel strategy.

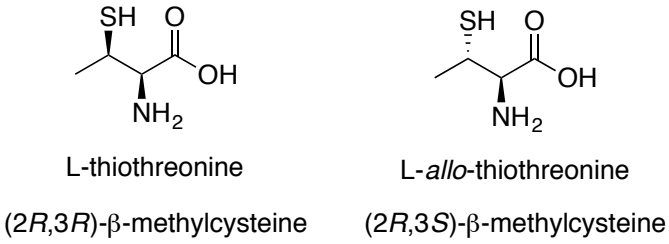
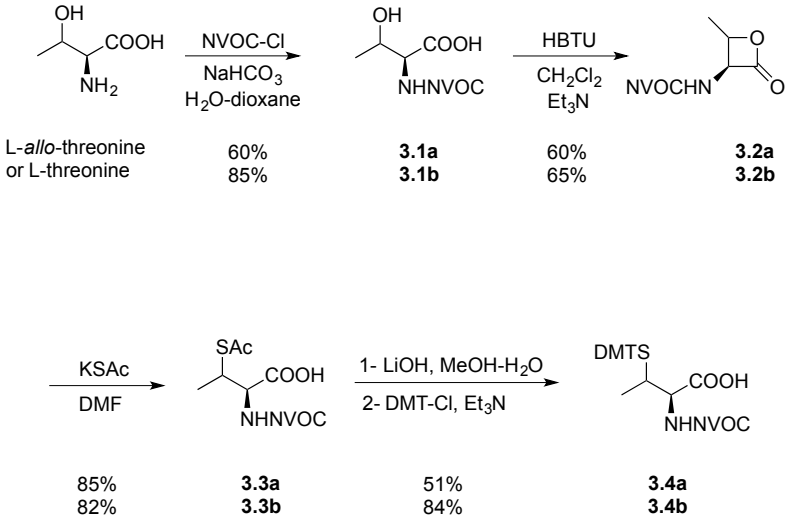


Figure 3.1. Structures of the β-methylcysteine (thiothreonine) stereoisomers synthesized.

3.2. Results

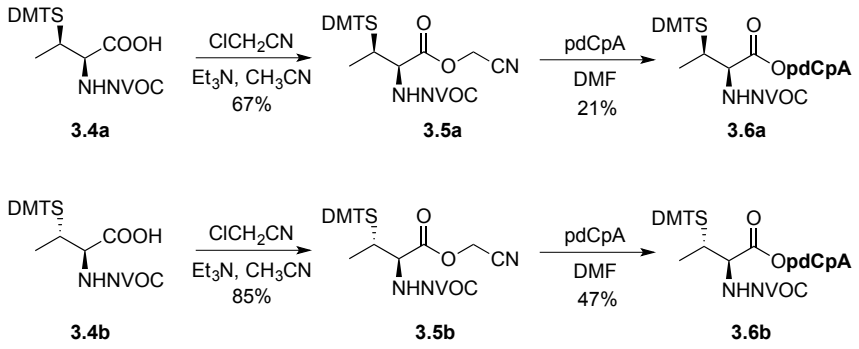
Dr. Nour Eddine Fahmi performed a substantial number of the chemical transformations for the preparation of the following molecules, and I extend the sincerest gratitude for his assistance with this project. The synthesis of protected L-thiothreonines (**3.4a** and **3.4b**, Scheme 3.1) commenced with the preparation of the NVOC derivative of either L-*allo*- or L-thiothreonine, by treatment of the respective amino acid with 6-nitroveratryl chloroformate (NVOCCl), the latter of which was prepared according to the procedure described by Katritzky et al.[64] Thus, L-*allo*-threonine or L-threonine was dissolved in a 1:1 mixture of dioxane and water in the presence of sodium bicarbonate, to which NVOCCl was added. Upon stirring for 17 hours at room temperature, NVOC-protected threonine derivatives, **3.1a** and **3.1b**, were obtained following work-up in 60

and 85% yields, respectively. **3.1a** or **3.1b** was then solubilized in dichloromethane and treated with triethylamine and HBTU and left to stir at room temperature under argon. Following stirring for 20 hours, extractive work-up followed by silica gel column chromatography gave the desired lactones **3.2a** and **3.2b** in 60 and 65% yield, respectively. Treatment of **3.2a** and **3.2b** with potassium thioacetate in DMF at room temperature for 23 hours afforded thioacetates **3.3a** and **3.3b,** regio- and stereospecifically in 85 and 82% yields, respectively, after purification. Following opening of the lactone the acetate group was hydrolyzed using lithium hydroxide and the resulting free thiol was treated with 4,4'-dimethoxytrityl chloride in the presence of triethylamine to give the orthogonally protected L-thiothreonine and L-*allo*-thiothreonine **3.4a** and **3.4b** in 51 and 84% yields, respectively.



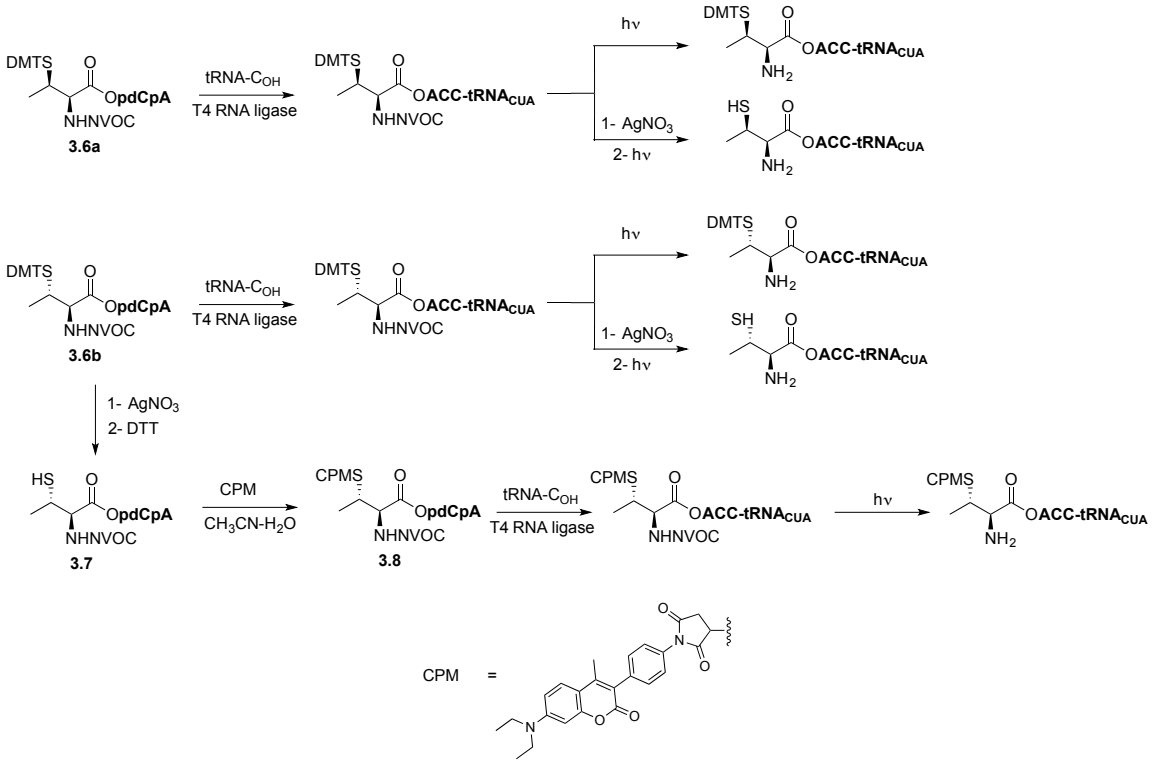
Scheme 3.1. Synthesis of *N*,*S*-orthogonally protected L-thiothreonine and L-*allo*-thiothreonine derivatives **3.4a** and **3.4b**.

Conversion of **3.4a** and **3.4b** to their corresponding cyanomethyl ester derivatives was accomplished by treatment with chloroacetonitrile in dry acetonitrile in the presence of triethylamine to afford compounds **3.5a** and **3.5b** in 67 and 85% yields, respectively. Treatment of the cyanomethyl esters with the tris-(tetrabutylammonium) salt of pdCpA in anhydrous DMF afforded the corresponding aminoacylated pdCpAs **3.6a** and **3.6b** in 21 and 47% yields, respectively (Scheme 3.2).



Scheme 3.2. Synthesis of *N*,*S*-orthogonally protected L-thiothreonine and L-*allo*-thiothreonine derivatives of pdCpA **3.6a** and **3.6b**.

Removal of the DMT group at the pdCpA level was performed on protected L-*allo*-thiothreonyl pdCpA ester **3.14** using silver nitrate in aqueous acetonitrile, followed by treatment with dithiothreitol to afford the free thiol (Scheme 3.4). The dinucleotide was labeled on the sulfur atom with a coumarin fluorophore [7-diethylamino-3-(4′-maleimidophenyl)-4-methylcoumarin] in acetonitrile−water, and the fluorescent dinucleotide **3.16** was obtained in 64% yield after purification by HPLC (Scheme 3.3 and Figure 3.2).



Scheme 3.3. Synthesis of thiothreonyl- and *allo*-thiothreonyl-tRNAC$_{CUA}$s.

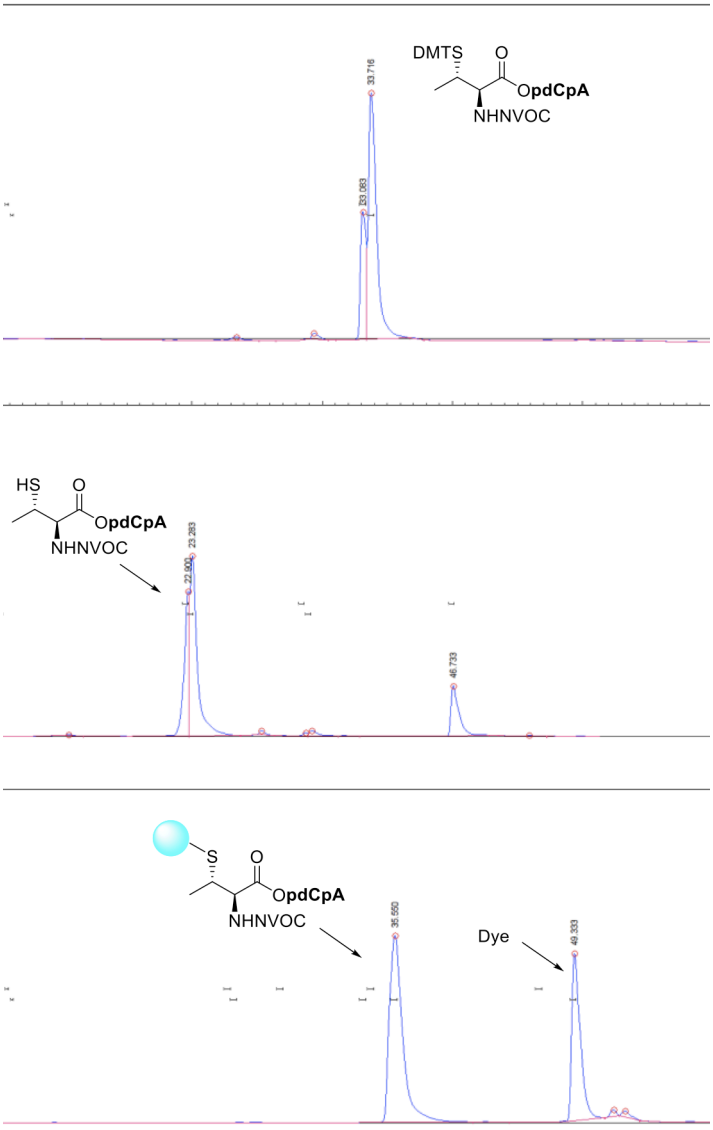


Figure 3.2. HPLC analysis of pdCpA derivatives of *allo*-thiothreonine.

Following treatment of **3.6b** with silver nitrate, two new peaks at 22.9 and 23.3 minutes (corresponding to the free thiol) were observed, as well as the disappearance of the peaks at 33.1 and 33.7, indicating the reaction was complete and quantitative. Moreover, coupling of the dye, mediated by maleimide chemistry, was very efficient, proceeding quantitatively as seen by the disappearance of thepeakes at 33.1 and 33.7 minutes and the

appearance of a new peak at 35.6 minutes.

Importantly, the technology developed herein allows for the introduction of the conjugated adduct (the dye coupled thiol) at various stages leading to the production of protein, as well as at the protein level itself.

3.3. Discussion

We reported the preparation of L-thiothreonine and L-*allo*-thiothreonine stereoisomers utilizing a novel ring opening reaction involving threonine-derived lactones. Lactone ring opening, using LiSH and NaSH as nucleophiles, afforded a β-isothiouronium salt, a plausible precursor to a β-methylcysteine derivative. This strategy was therefore applied for the synthesis of NVOC protected β-methylcysteine using potassium thioacetate as the nucleophile for ring opening of the β-lactone at the β-carbon atom (Scheme 3.1).

L-*allo*-threonine and L-threonine were employed as starting materials and were treated initially with NVOCCl, leading to *N*-protection and affording compounds **3.1a** and **b,** followed by treatment with HBTU to produce β-lactone derivatives **3.2a** and **3.2b**. Opening of the β-lactone ring was facilitated by potassium thioacetate to afford the thioacetate derivatives **3.3a** and **3.3b**. Transformation of the thiol protecting group by hydrolysis of the acetate with lithium hydroxide, followed by treatment the resulting free thiol with 4,4′-dimethoxytrityl chloride, was achieved in high yield. The conversion of protected compounds **3.4a** and **3.4b** to the corresponding cyanomethyl esters (**3.5a** and **3.5b)** was achieved by treatment with chloroacetonitrile in dry acetonitrile in the presence of triethylamine.

The treatment of the active esters **3.5a** and **3.5b** with the tris-(tetrabutylammonium) salt of pdCpA in anhydrous DMF gave the corresponding aminoacylated pdCpAs **3.6a** and **3.6b**, respectively. Removal of the DMT group at the pdCpA level was performed on protected L-*allo*-thiothreonyl pdCpA ester **3.6b** using silver nitrate in aqueous acetonitrile, followed by treatment with dithiothreitol, DTT (Scheme 3.3 and Figure 3.2).

The dinucleotide was labeled on the sulfur atom with a coumarin fluorophore [7-diethylamino-3-(4′-maleimidophenyl)-4-methylcoumarin] in acetonitrile−water and the fluorescent dinucleotide **3.8** was obtained following HPLC purification. Suppressor $tRNA_{CUA}$s activated with either L-thiothreonine or *allo*-L-thiothreonine were prepared as outlined in Scheme 3.3. The aminoacylated pdCpAs were ligated to an abbreviated suppressor $tRNA_{CUA}$ transcript ($tRNA_{CUA}$-C_{OH}, lacking the terminal cytidine and adenosine moieties present at the 3′-terminus of all tRNAs) via a T4 RNA ligase mediated ligation reaction to obtain the N and S-protected thiothreonyl-$tRNA_{CUA}$s. The ligation efficiencies were evaluated via denaturing PAGE analysis as shown in Figure 3.3.[65]

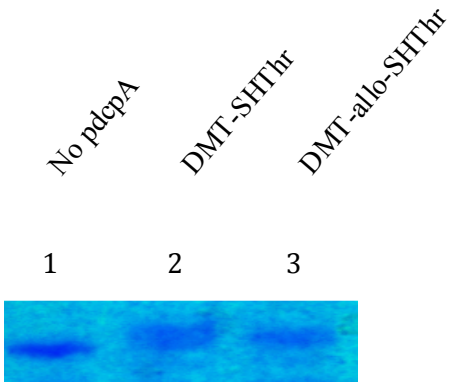


Figure 3.3. Ligation of thiothreonyl-pdCpA derivatives to $tRNA_{CUA}$-C_{OH}. Lane 1, 74 nt abbreviated suppressor $tRNA_{CUA}$; Lane 2, 76 nt thiothreonyl-$tRNA_{CUA}$; Lane 3, 76 nt allo-thiothreonyl-$tRNA_{CUA}$.

The *N*-protected aminoacylated tRNAs were then irradiated with UV light at 0 °C for 5 minutes to afford the activated tRNAs having amino acids with free α-amines. Evaluation of the deprotection efficiency of the DMT group was investigated prior to photolysis. A portion of the fully protected tRNAs were subjected to treatment with silver nitrate, followed by treatment with dithiothreitol, affording thiothreonyl-$tRNA_{CUA}$s having free sulfhydryl groups. Removal of the DMT group was verified via coupling with the coumarin fluorophore and visualized by irradiating with 365 nm wavelength light (Figure 3.4).

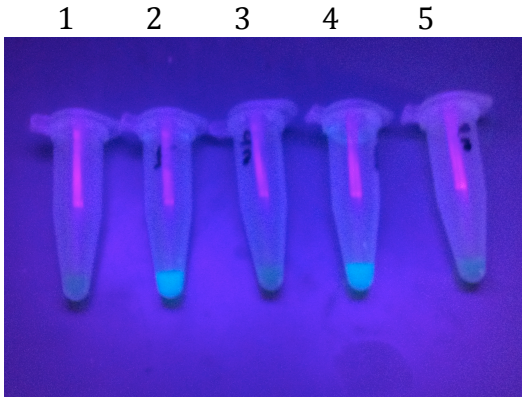


Figure 3.4. Coupling of the NVOC-protected thiothreonyl-tRNA analogues with CPM after DMT group removal.

Allo-thiothreonyl-tRNA$_{CUA}$ derivatized with the same fluorophore was also prepared by a T4 RNA ligase-mediated coupling of dinucleotide 14 with tRNA-C$_{OH}$ (Scheme 3.4). Following irradiation with UV light to remove the NVOC protecting group, the CPM-derivatized *allo*-thiothreonyl-tRNA$_{CUA}$ was employed in a protein synthesis system, as described below.

The ability of thiothreonyl-tRNA$_{CUA}$s to participate in protein synthesis was evaluated in a bacterial cell free protein synthesis system. A plasmid encoding modified *E. coli* DHFR where the Cys85 and Cys152 codons were replaced with serine codons was used. This modified DHFR has been found to retain nearly full enzymatic activity (Figure 3.5). This work was performed by Dr. Shengxi Chen.

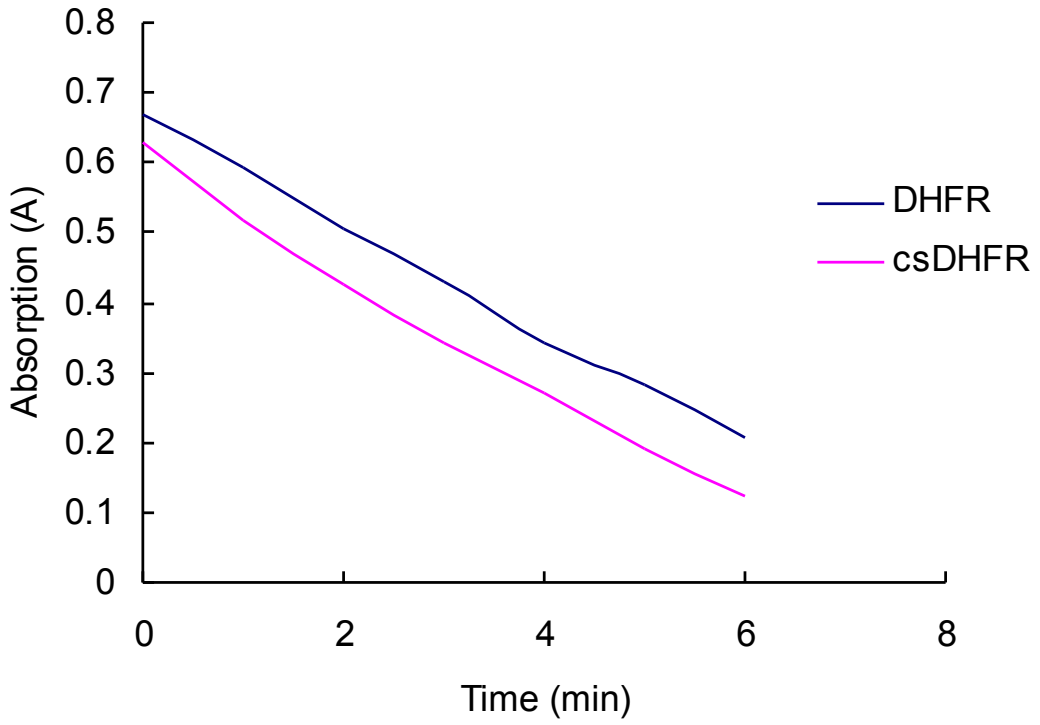
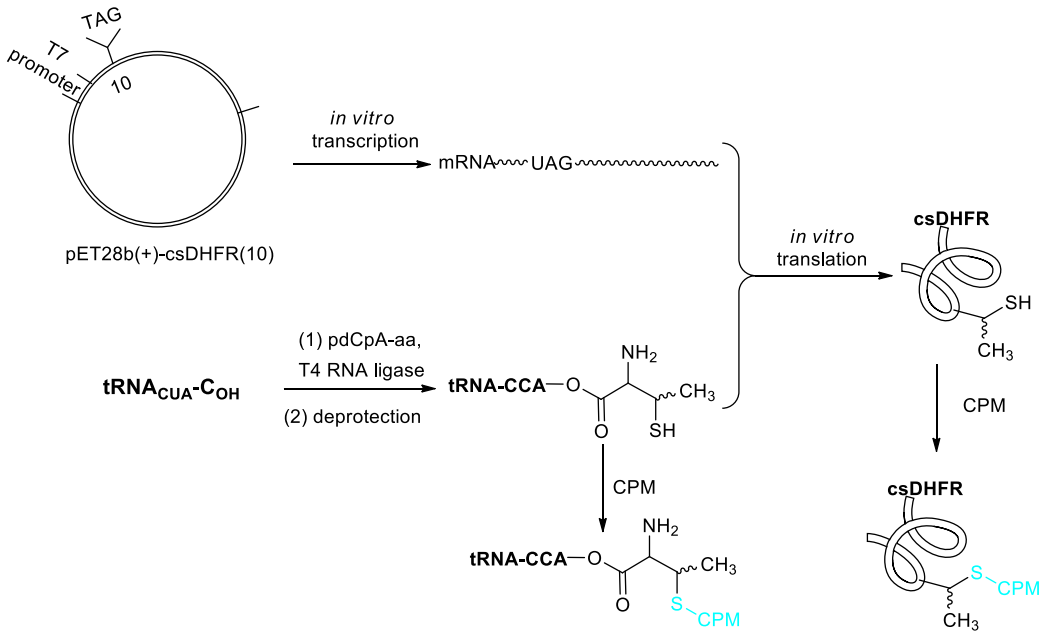


Figure 3.5. Enzymatic activity of csDHFR compared to that of wild-type DHFR.

Additionally, a stop codon (TAG) was incorporated into the plasmid at DHFR position 10 to facilitate incorporation of the thiothreonine derivatives from the misacylated suppressor tRNAs (Scheme 3.4).



Scheme 3.4. Strategy employed for incorporation of thiothreonines into csDHFR and coupling with CPM.

Both L-thiothreonyl-tRNA$_{CUA}$ and L-*allo*-thiothreonyl-tRNA$_{CUA}$ were found to participate in protein synthesis, affording DHFRs with 43% and 45% suppression efficiencies, respectively, relative to the amount of csDHFR produced from the corresponding mRNA lacking any stop codon (Figure 3.6). The DMT-protected thiothreonine and *allo*-thiothreonine were incorporated in yields of 28% and 26%, respectively. In comparison, the incorporation of the fluorescently labeled *allo*-thiothreonine from CPM-derivatized *allo*-thiothreonyl-tRNA$_{CUA}$ (prepared by ligation of **3.8** and tRNA-C$_{OH}$) proceeded in low (~5%) yield. In replicate experiments, the incorporation of CPM-*allo*-thiothreonine into csDHFR ranged from 2–8%.

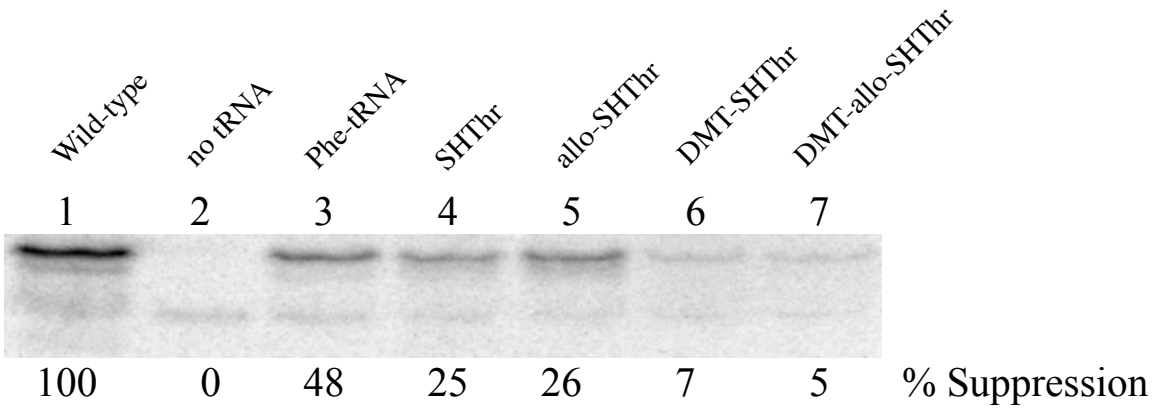


Figure 3.6. *In vitro* incorporation of thiothreonine derivatives into csDHFR at position 10.

All of the proteins were purified successively on Ni-NTA and DEAE-Sepharose columns, as illustrated for the thiothreonine-containing DHFR in Figure 3.7. These experiments were performed by Dr. Shengxi Chen.

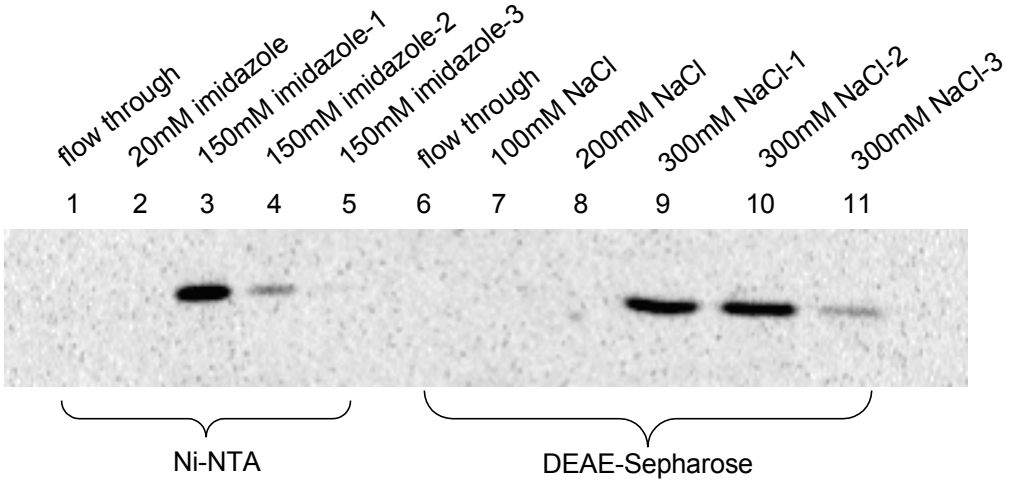


Figure 3.7. Purification of thiothreonyl-DHFR by successive chromatography on Ni-NTA and DEAE Sepharose columns.

Coupling of the purified proteins containing malemide fluorophore labeled (position 10) thiothreonine and *allo*-thiothreonine was visualized by irradiating with 365 nm wavelength light (Figure 3.8). These experiments were performed by Dr. Shengxi Chen. As is clear from the figure, the fluorophore was readily introduced into the nascent csDHFRs containing thiothreonine and *allo*-thiothreonine. As anticipated, the csDHFRs elaborated with DMT-protected thiothreonine and *allo*-thiothreonine did not react detectably with the coumarin fluorophore.

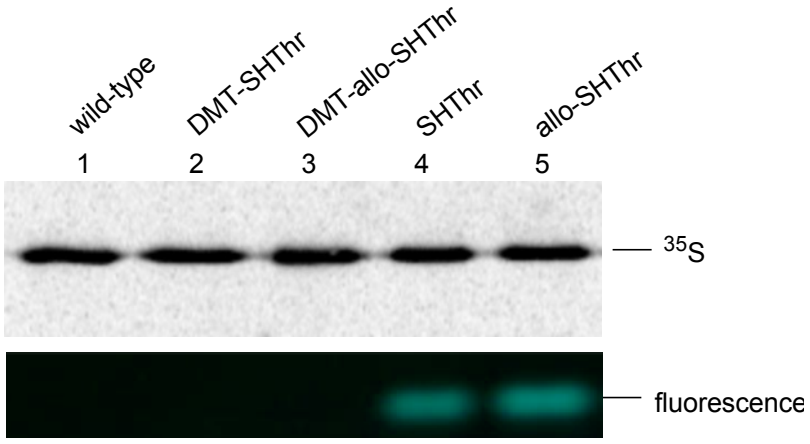


Figure 3.8. Coupling of thiothreonyl-csDHFR analogues with CPM. All of the protein samples were purified chromatographically prior to treatment with the fluorophore.

In the present study, we used a DMT group to protect the thiol groups of L-thiothreonine and L-*allo*-thiothreonine. The DMT group can be removed at the pdCpA and tRNA levels, and both the DMT-protected and unprotected thiothreonyl- and *allo*-thiothreonyl-$tRNA_{CUA}$s were used to incorporate the thiolated amino acids at position 10 of csDHFR. The nascent csDHFRs containing unprotected thiothreonine and

allo-thiothreonine readily reacted with CPM, affording protein that was fluorescently labeled at a single position. However, the DMT group could not be removed from the thiothreonines at the protein level due to protein precipitation under the conditions used for deprotection. The free thiol groups of thiothreonine and *allo*-thiothreonine could be modified with the fluorophore CPM at the pdCpA, tRNA and protein levels. The CPM-protected *allo*-thiothreonyl-$tRNA_{CUA}$, however, proved to be a poor substrate in the peptidyltransferase reaction, such that the preferred method for elaborating the fluorescent protein was by direct derivatization of the initially formed csDHFR containing a free SH group in the amino acid at position 10.

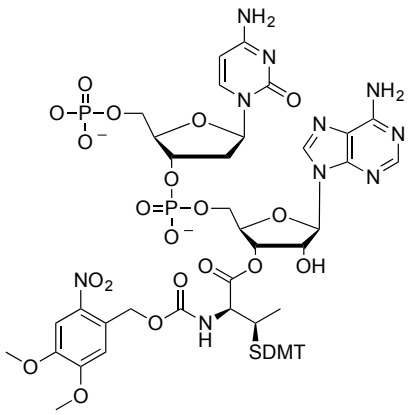
3.4. Experimental

3.4.1. General materials and methods

Reagents and solvents for chemical synthesis were purchased from Aldrich Chemical Co. or Sigma Chemical Co. and used without further purification. Ni-NTA agarose was obtained from Qiagen Inc. (Valencia, CA). DNA oligonucleotides were ordered from Integrated DNA Technologies (Coralville, IA). DEAE-Sepharose, ammonium persulfate, acrylamide, *N*, *N′*-methylene-bis-acrylamide, acetic acid, potassium glutamate, ammonium acetate, dithiothreitol, magnesium acetate, phospho(enol)pyruvate, *Escherichia coli* tRNA, isopropyl β-D-thiogalactopyranoside (IPTG), ATP, GTP, CTP, UTP, cAMP, amino acids, rifampicin, formamide and 7-diethylamino-3-(4′-maleimidophenyl)-4-methylcoumarin (CPM) were obtained from Sigma-Aldrich (St. Louis, MO). Tris and SDS were obtained from Bio-Rad Laboratories (Hercules, CA). [^{35}S]-methionine (1000 Ci/mmol, 10 μCi/μL) was purchased from PerkinElmer Inc.

(Boston, MA). Protease inhibitor (complete, EDTA-free) was obtained from Boehringer Mannheim Corp. (Indianapolis, IN). T4 RNA ligase and T4 polynucleotide kinase were purchased from New England Biolabs Inc. (Ipswich, MA). Phosphorimaging analysis was performed using an Amersham Biosciences Storm 820 equipped with ImageQuant version 5.2 software from Molecular Dynamics. UV spectral measurements were performed using a Perkin-Elmer Lamdba 20 UV/vis spectrometer. Fluorescence was monitored at 365 nm. The mutant csDHFR analogues modified at position 10 with thiothreonine were prepared using the general strategy shown in Scheme 5.[22,23,31,53-63,66]

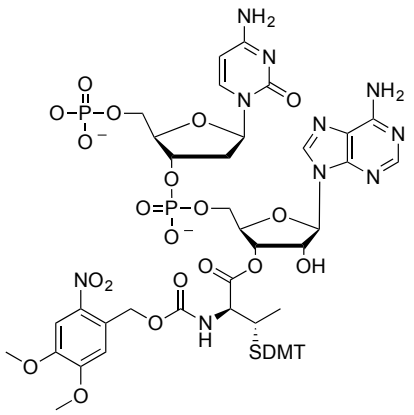
3.4.2. Preparation of thiothreonyl-pdCpA derivatives



***S*-(4,4′-Dimethoxytrityl)-*N*-(6-nitroveratryloxycarbonyl)-L-thiothreonyl pdCpA ester (3.6a)**

To a conical vial containing 16.0 mg (22.4 μmol) *S*-(4,4′-dimethoxytrityl)-*N*-(6-nitroveratryloxycarbonyl)-L-thiothreonine cyanomethyl ester (**3.5a**) was added a solution containing 5.8 mg (4.3 μmol) of the tris-(tetrabutylammonium) salt of pdCpA in 90 μL of anhydrous DMF followed by 10 μL of triethylamine. The reaction mixture was stirred at

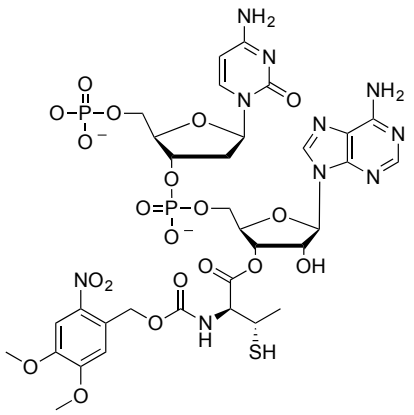
room temperature under argon atmosphere for 24 hours. A 5-μL aliquot of the reaction mixture was diluted with 55 μL of 1:1 CH_3CN/50 mM NH_4OAc at pH 4.5 and was analyzed by HPLC on a C_{18} reverse phase column (250 × 10 mm). The column was washed with 1 → 65% CH_3CN in 50 mM NH_4OAc at pH 4.5 over a period of 45 minutes at a flow rate of 3.5 mL/minute (monitored at 260 nm). The remaining reaction mixture was diluted to a total volume of 0.6 mL with CH_3CN and purified using the same C_{18} reverse-phase column. *S*-(4,4′-Dimethoxytrityl)-*N*-(6-nitroveratryloxycarbonyl)-L-thiothreonyl pdCpA ester (**3.6a**) (retention time 33.8 and 34.4 minutes) was recovered from the appropriate fractions as a colorless solid by lyophilization: yield 1.2 mg (21%); mass spectrum (ESI), *m/z* 1295.3136 $(M+H)^+$ ($C_{54}H_{61}N_{10}O_{22}P_2S$ requires *m/z* 1295.3158).



_S_-(4,4'-Dimethoxytrityl)-_N_-(6-nitroveratryloxycarbonyl)-L-_allo_-thiothreonyl pdCpA ester (3.6b)

To a conical vial containing 16.0 mg (22.4 μmol) of *S*-(4,4'-dimethoxytrityl)-*N*-(6-nitroveratryloxycarbonyl)-L-*allo*-thiothreonine cyanomethyl ester (**3.5b**) was added a solution of 5.8 mg (4.3 μmol) of the tris-(tetrabutylammonium) salt of pdCpA in 90 μL of anhydrous DMF, followed by 10 μL of triethylamine. The reaction mixture was stirred

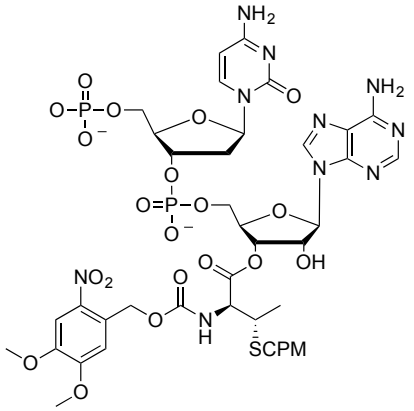
at room temperature under argon atmosphere for 24 hours. A 5-μL aliquot of the reaction mixture was diluted with 55 μL of 1:1 CH_3CN/50 mM NH_4OAc at pH 4.5 and was analyzed by HPLC on a C_{18} reverse phase column (250 × 10 mm). The column was washed with 1 → 65% CH_3CN in 50 mM NH_4OAc at pH 4.5 over a period of 45 minutes at a flow rate of 3.5 mL/minute (monitoring at 260 nm). The remaining reaction mixture was diluted to a total volume of 0.6 mL with CH_3CN and purified using the same C_{18} reverse phase column. *S*-(4,4'-dimethoxytrityl)-*N*-(6-nitroveratryloxycarbonyl)-L-*allo*-thiothreonyl pdCpA ester (**3.6b**) (retention time 34.0 and 34.5 minutes) was recovered from the appropriate fractions as a colorless solid by lyophilization: yield 2.6 mg (47%); mass spectrum (ESI), *m/z* 1295.3203 $(M+H)^+$ ($C_{54}H_{61}N_{10}O_{22}P_2S$ requires 1295.3158).



_N_-(6-nitroveratryloxycarbonyl)-L-_allo_-thiothreonyl pdCpA ester (3.7)

To a solution containing 1.6 mg (1.23 μmol) of *S*-(4,4'-dimethoxytrityl)-*N*-(6-nitroveratryloxycarbonyl)-L-*allo*-thiothreonyl pdCpA ester (**3.6b**) in 0.60 mL of 3:1 acetonitrile/water (v/v) was added a solution containing 3.13 mg (18.5 μmol) of $AgNO_3$ in 40 μL of water. The reaction mixture was stirred at room temperature for 1 hour, then 9.5 mg (61.6 μmol) of dithiothreitol in 50 μL of water was added and stirring was

continued for 90 minutes. The mixture was centrifuged at 15,000 × g at 4 °C for 5 minutes and the supernatant was collected and purified by HPLC on a C_{18} reverse phase column (250 × 10 mm). The column was washed with 1 → 65% CH_3CN in 50 mM NH_4OAc at pH 4.5 over a period of 45 min at a flow rate of 3.5 mL/min (monitoring at 260 nm). *N*-(6-Nitroveratryloxycarbonyl)-L-*allo*-thiothreonyl pdCpA ester (**3.7**) (retention time 22.9 and 23.3 min) was recovered from the appropriate fractions as a colorless solid by lyophilization and was used immediately in the next step: yield 1.0 mg (82%); mass spectrum (MALDI), *m/z* 993.4 $(M+H)^+$ (theoretical *m/z* 993.2).



_S_-[7-Diethylamino-3-(4'-maleimidophenyl)-4-methylcoumarin]-_N_-(6-nitroveratryloxycarbonyl)-L-_allo_-thiothreonyl pdCpA ester (3.8)

To a solution containing 1.0 mg (1.0 μmol) of *N*-(6-nitroveratryloxycarbonyl)-L-*allo*-thiothreonyl pdCpA ester (**3.7**) in 375 μL of 2:1 acetonitrile/water (v/v) was added 0.8 mg (2.0 μmol) of 7-diethylamino-3-(4'-maleimidophenyl)-4-methylcoumarin in 75 μL of acetonitrile. The mixture was stirred at room temperature for 3 hours, then purified by HPLC on a C_{18} reverse phase column (250 × 10 mm). The column was washed with 1 → 65% CH_3CN in 50 mM NH_4OAc at pH 4.5 over a period of 45 minutes at a flow rate of

3.5 mL/minute (monitored at 260 nm). *S*-[7-Diethylamino-3-(4'-maleimidophenyl)-4-methylcoumarin]-*N*-(6-nitroveratryloxycarbonyl)-L-*allo*-thiothreonyl pdCpA ester (**3.8**) (retention time of 34-35 min) was recovered from the appropriate fractions as a yellow solid by lyophilization: yield 0.9 mg (64%); mass spectrum (MALDI), *m/z* 1395.6 $(M+H)^+$ ($C_{57}H_{65}N_{12}O_{24}P_2S$ requires 1395.3).

3.4.3. Ligation of suppressor $tRNA_{CUA}$-C_{OH} with thiothreonine analogues and deprotection of DMT and NVOC groups

Activation of suppressor $tRNA_{CUA}$ was carried out in 100 μL (total volume) of 100 mM Hepes buffer, pH 7.5, containing 2.0 mM ATP, 15 mM $MgCl_2$, 100 μg of suppressor tRNA-C_{OH}, 2.0 A_{260} units of protected aminoacyl-pdCpA (5-10 fold molar excess), 15% DMSO and 200 units of T4 RNA ligase. After incubation at 37 ^oC for 1 hour, the reaction was quenched by the addition of 10 μL of 3 M NaOAc, pH 6.3, followed by 300 μL of ethanol. The reaction mixture was incubated at –20 ^oC for 30 minutes, then centrifuged at 15,000 × g at 4 ^oC for 30 minutes. The supernatant was carefully decanted and the tRNA pellet was washed with 100 μL of 70% ethanol and dissolved in 100 μL of RNase free H_2O. The efficiency of ligation was estimated by 8% denaturing PAGE (pH 5.2).[65] The DMT protecting group of *N*-(6-nitroveratryloxycarbonyl)-thiothreonyl-tRNA or *N*-(6-nitroveratryloxycarbonyl)-*allo*-thiothreonyl-tRNA was removed by treatment with 50 mM $AgNO_3$ at room temperature for 30 minutes, followed by treatment with 60 mM DTT at room temperature for 30 minutes.[67] The reaction mixture was centrifuged at 15,000 × g at 4 ^oC for 10 minutes, and then the supernatant was carefully decanted. The precipitated pellet was washed with 100 μL of 0.3 M NaOAc, pH 6.3. The combined

aqueous supernatant was treated with 600 μL of ethanol to precipitate the tRNA. The tRNA pellet was washed with 100 μL of 70% ethanol and then dissolved in 30 μL of RNase free H_2O. The NVOC-protected aminoacyl-tRNA was cooled to 2 ^oC and irradiated with a 500 W mercury-xenon lamp for 5 minutes.[32,45] After irradiation, the deblocked aminoacylated suppressor tRNAs were using *in vitro* suppression experiments without further purification.

Also prepared were samples of the thiothreonyl- and *allo*-thiothreonyl-tRNAs in which the DMT group was not removed prior to photolysis of the NVOC protecting group.

3.4.4. Coupling of NVOC-thiothreonyl-tRNA analogues with CPM

The NVOC-thiothreonyl-tRNA coupling was carried out in 50 μL (total volume) of 100 mM NaOAc, pH 6.3, containing 90 μg of thiothreonyl-tRNA analogue and 1 mM of CPM. The reaction was incubated at room temperature for 1 hour and was quenched by the addition of 3 μL of 3 M NaOAc, pH 6.3, followed by 150 μL of ethanol. The reaction mixture was incubated at –20 ^oC for 30 minutes, then centrifuged at 15,000 × g at 4 ^oC for 30 minutes. The supernatant was carefully decanted and the tRNA pellet was washed with 50 μL of 70% ethanol and dissolved in 30 μL of RNase free H_2O. The fluorescence was monitored at ~ 470 nm following irradiation at 365 nm.

3.4.5. Ligation of suppressor tRNA-C_{OH} with CPM-*allo*-thiothreonyl-pdCpA and deprotection of NVOC groups

Suppressor tRNA aminoacylation was carried out in 100 mL (total volume) of 100 mM

Hepes buffer, pH 7.5, containing 2.0 mM ATP, 15 mM $MgCl_2$, 100 μg of suppressor tRNA-C_{OH}, 2.0 A_{260} units of NVOC-protected CPM-*allo*-thiothreonyl-pdCpA, 15% DMSO and 200 units of T4 RNA ligase. After incubation at 37 ^{o}C for 1 hour, the reaction was quenched by the addition of 10 μL of 3 M NaOAc, pH 5.2, followed by 300 μL of ethanol. The reaction mixture was incubated at –20 ^{o}C for 30 minutes, then centrifuged at 15,000 × g at 4 ^{o}C for 30 minutes. The supernatant was carefully decanted and the tRNA pellet was washed with 100 μL of 70% ethanol and dissolved in 30 μL of RNase free H_2O. The NVOC-protected CPM-*allo*-thiothreonyl-tRNA was cooled to 2 ^{o}C and irradiated with a 500 W mercury-xenon lamp for 5 minutes.

3.4.6. *In vitro* translation of csDHFR analogues

The wild-type csDHFR plasmid was obtained by site-directed mutation as described as previously, using wild-type DHFR plasmid as the template (prepared by Dr. Chen).[33] The DNA primer for the mutation at position 85 is 5'-GATGAAGCCATCGCGGCGTCTGGTGACGTACCAGAAATC-3'; the primer for the mutation at position 152 is 5'-CAGAACTCTCACAGCTATAGCTTTGAGATTCTGGAGC-3'. The mutant csDHFR (TAG at position 10) was obtained by the same site-directed mutation using the wild-type csDHFR plasmid as the template and a sequence of 5'-GTCTGATTGCGGCGTTAGCGTAGGATCGCGTTATCGGCATG-3' as the primer. The *in vitro* expression mixture (300 μL total volume) contained 30 μg of mutant DHFR (TAG at position 10) plasmid DNA, 120 μL of premix (35 mM Tris-acetate, pH 7.0, 190 mM potassium glutamate, 30 mM ammonium acetate, 2.0 mM dithiothreitol, 11 μM

maganesium acetate, 20 mM phospho(enol)pyruvate, 0.8 mg/mL of *E. coli* tRNA, 0.8 mM IPTG, 20 mM ATP and GTP, 5 mM CTP and UTP and 4 mM cAMP), 100 mM of each of the 20 amino acids, 30 mCi of [^{35}S]-L-methionine, 10 μg/μL rifampicin, 90 μg of deprotected misacylated $tRNA_{CUA}$ and 90 μL of S-30 extract from *E. coli* strain BL21(DE3). The reaction mixture was incubated at 37 ^{o}C for 45 minutes. Plasmid DNA containing the gene for wild-type csDHFR was used as the positive control, and an abbreviated tRNA (tRNA-C_{OH}) lacking any amino acid was used as the negative control. An aliquot containing 2 μL of reaction mixture was removed, treated with 2 μL of loading buffer and heated at 90 ^{o}C for 2 minutes. This was analyzed by 15% SDS-PAGE at 100 V for 2 hour.

CHAPTER 4

SYNTHESIS OF CAGED DHFRS WITH PROTECTED PHOSPHOROTYROSINES

4.1. Introduction

Protein phosphorylation is a universal, post-translational mechanism for activation or deactivation of proteins and for the regulation of cell signaling events, differentiation, metabolism, neuronal communication and signal transduction of many life forms.[68-71] Two amino acids are generally implicated in protein phosphorylation events, namely serine and threonine.[72] It has been shown, however, that tyrosine, histidine, aspartate and glutamate are also targets of phosphorylation.[72,73] Additionally, radiolabeling studies have suggested that nearly 30% of eukaryotic proteins are subject to phosphorylation.[72] The signals that initiate and terminate phosphorylation are regulated by kinases and phosphoprotein phosphatases, respectively[74,75], and abnormal phosphorylation events have been implicated in many diseases.[76-78] The importance of kinases in cell signaling is consistent with their relative abundance in the eukaryotic family; they constitute nearly 2% of the entire genome. This realization has resulted in their broad study.[71,79,80] The discovery that glycogen is activated by means of phosphorylation prompted intensive research in an attempt to understand the conversion between the following two states: the active glucose-1-phosphate (phosphorylase a) and the inactive glycogen (phosphorylase b).[81,82] It was formulated that a protein kinase, phosphorylase kinase (PhK), phosphorylated two Ser residues, thereby inducing a conformational shift that rendered the protein active. In fact, one of the requirements for

activity is phosphorylation of PhK, an event that is carried out by another phosphorylase kinase, protein kinase A (PKA). Accordingly, the idea of protein phosphorylation as a regulatory process was soon realized.[68,83-85] Protein phosphorylation has been shown to support a number of physiological processes, and defects in these signaling pathways have been linked to pathogenic mechanisms such as cancer and other human diseases. Moreover, protein-tyrosine kinases (PTKs) have been shown to function in growth control; phosphorylation of tyrosine residues is a major mechanism of transmembrane signaling.[86-88] Biochemical analysis of protein phosphorylation events and their effects are typically demonstrated through site-directed mutagenesis and replacement of known, post-translationally modified Ser, Thr or Tyr residues with Ala for Ser/Thr and Phe for Tyr.[89]

Chemical approaches for examining protein phosphorylation and phosphorylated protein function include the use of photolabile caged analogs of amino acids and the introduction of amino acid derivatives.

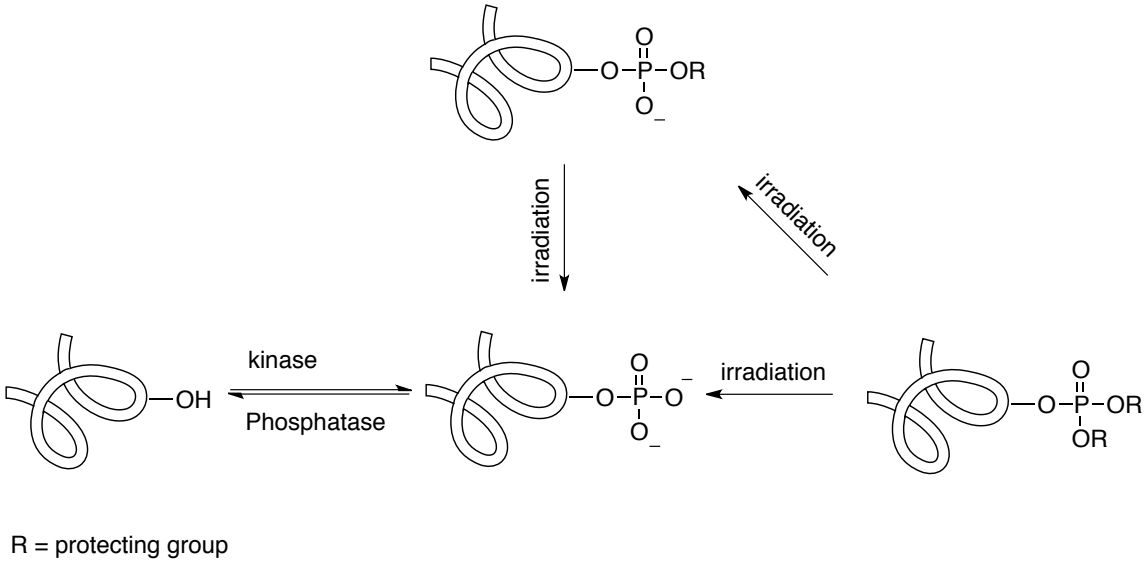


Figure 4.1. Phosphoregulation of proteins. Kinases and phosphatases add or remove phosphate groups in amino acids, respectively. Alternatively, a caged phosphate can be removed by irradiation to reveal the phosphate.

Caged compounds are typically photolabile molecules or probes that functionally encapsulate biomolecules in an inactive form by masking the functional groups required for activity. Irradiation of the biologically masked species removes the photolabile molecule and releases the active species, the phosphate, into the system. There are several attributes required for a protecting group to be effective for caging: (1) it doesn't function as an agonist or an antagonist, (2) following irradiation the molecule is released rapidly and in high yield, (3) the byproducts of release should be inert, and (4) the photolytic conversion should be accessible under physiological conditions.[90] The predominant structural motif for biological investigations is derived from the *O*-nitrobenzyl functionality. The *O*-nitrobenzyl group is inert within living cells, has a relatively good quantum yield of uncaging, is excited with nearly visible light which will not cause

radiative damage to cells and releases byproducts that are water soluble.[91]

Caged, phosphorylated amino acids are ideal probes for studying phosphorylated protein interactions and cell signaling pathways as they offer both spatial (the site of the protein) and temporal control over targeted effector molecules. We, therefore, sought to develop a variety of caged phosphorylated tyrosine derivatives for the purpose of aminoacylating tRNAs and evaluating their incorporation into proteins. Interestingly, ribosomal acceptance of the free phosphotyrosine, the mono-ester (singly protected), and the di-ester (doubly protected) varied drastically. Similar observations have been previously documented, with a variety of charged amino acids slowing the ribosomal synthetic machinery. *In vitro* expression systems utilizing amber stop codons are also biased against charged amino acids. Specifically, a variety of aspartic acids residues were utilized for protein synthesis and only an allyl protected aspartic acid was able to participate in protein synthesis most efficiently.[66] While there have been several studies on the preparation of caged phosphoproteins, there are limited reports that illustrate the ability of the ribosome to incorporate a variety of phosphorylated amino acids in varying states of caging, *i.e.* the free phosphate versus the singly and the doubly caged adducts, compounds **4.1**, **4.2** and **4.3,** respectively.

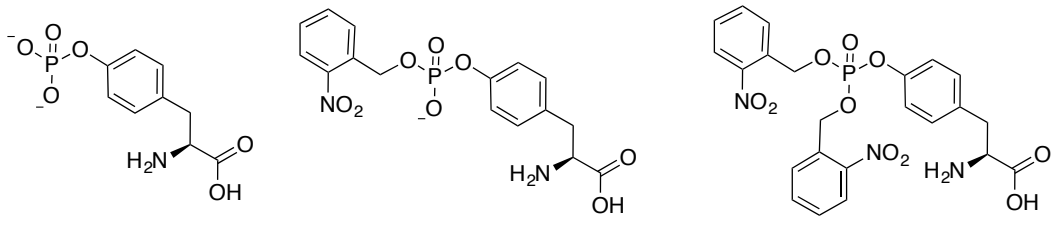


Figure 4.2. Synthesized phosphorylated tyrosine derivatives.

Ribosomal acceptance of bulky non-natural amino acids tends to diminish with increasing amino acid side chain size. Therefore, the steric bulk resulting from masking the phosphate charge could potentially result in poor incorporation of the amino acid.

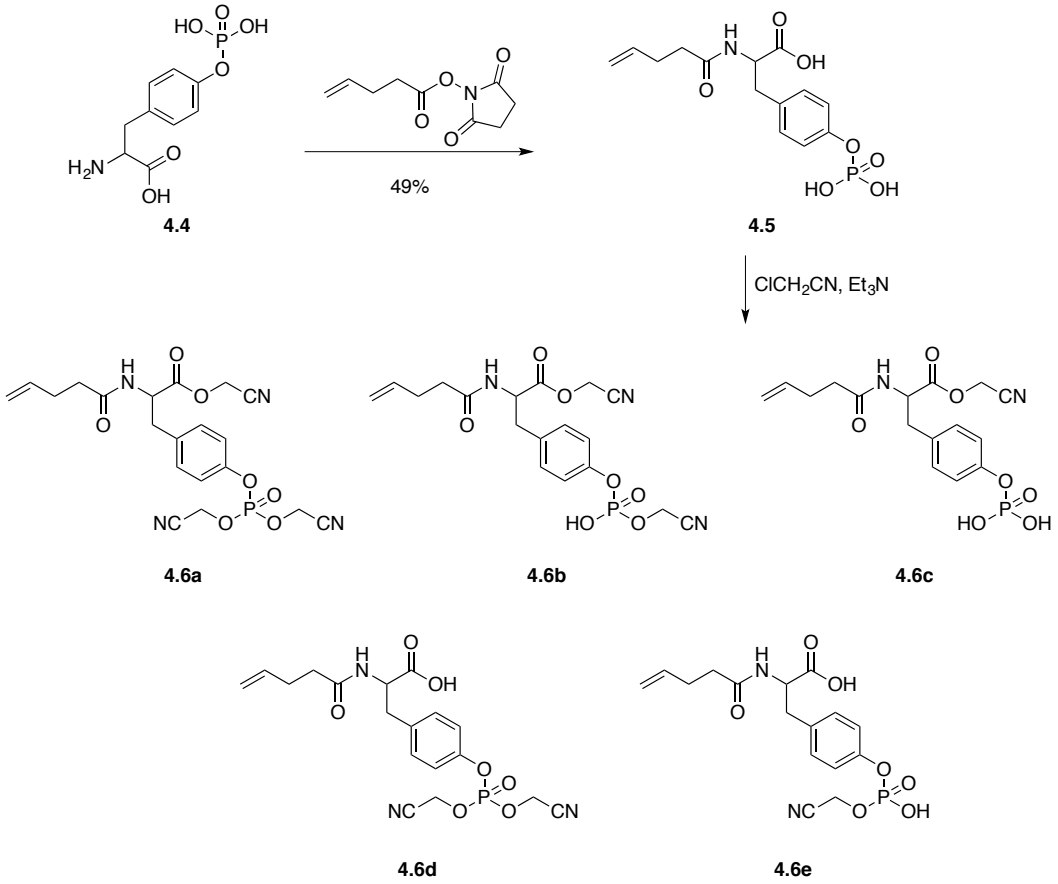
A more widely accepted and applied method involves the use of phosphate-specific antibodies. Phosphate-specific antibodies are capable of binding and detecting phosphorylation-induced conformational changes in proteins; however, it's still unclear whether the antibodies recognize the phosphorylated amino acid or the altered protein conformation.

Two-dimensional gels are considered the easiest method for detection,[92,93] as the electrophoretic mobility should be altered due to the addition of a phosphate group. However, two-dimensional gels typically require a large amount of sample handling. Consequently, it is difficult to reproduce results and the method suffers from a smaller dynamic range than other methods. Mass spectrometry is a powerful tool when more comprehensive assessments of protein phosphorylation are required. However, it too has its limitations; for a thorough review see Mann et al.[94] In order to verify the identity/incorporation of the protected p-Tyr into proteins, we selected mass spectrometry for our identification method.

4.2. Results

The synthesis of phosphorotyrosyl-tRNA was initiated by the preparation of *N*-pentenoyl phosphorotyrosine, via treatment of phosphorotyrosine **4.4** with pentenoic acid succinimide ester to afford **4.5** (Scheme 4.1). The *N*-protected amino acid was then

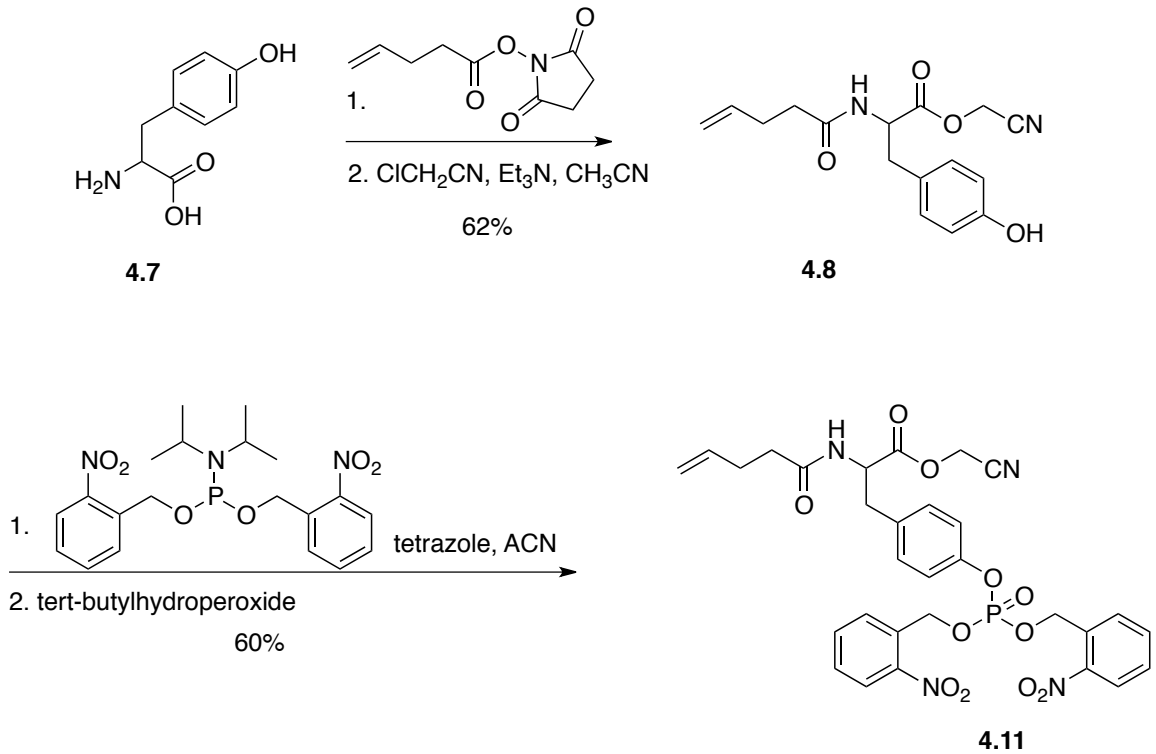
treated with chloroacetonitrile and triethylamine in an attempt to activate the C-terminus as the cyanomethyl ester derivative. However, when attempting to activate the acid we observed multiple additions via MALDI-TOF MS analysis, putatively leading to compounds **4.6a-4.6e** as outlined in scheme 4.1.



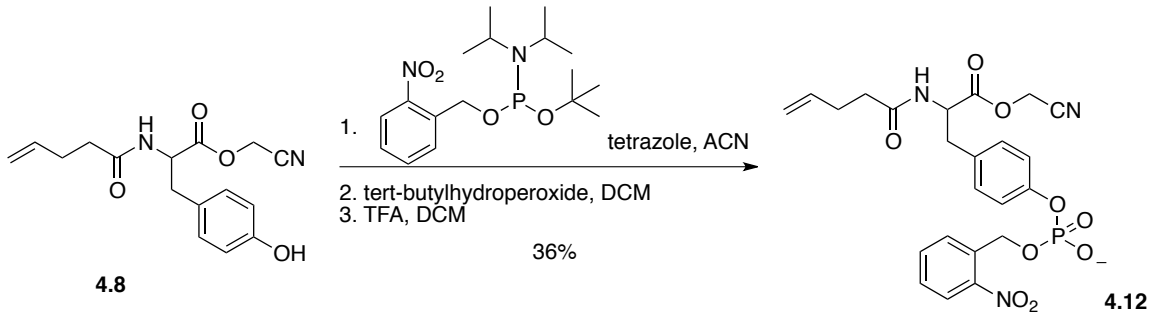
Scheme 4.1. Synthesis of protected phosphotyrosines activated as cyanomethyl esters.

Activating the carboxylic acid in the presence of a free phosphate group has the potential of resulting in a complex mixture of products and was, therefore, revised. We instead initiated the synthesis from free tyrosine, **4.7**. Again, the amine was protected as

the pentenoyl derivative by treatment with pentenoic acid succinimide ester, followed by activation as the cyanomethyl ester via chloroacetonitrile, to afford compound **4.8** in 48% yield. Treatment of **4.8** with phosphitylating agent **4.9** in the presence of tetrazole, followed by oxidation with *tert*-butylhydroperoxide afforded compound **4.11** in 60% yield over two steps. Similarly, compound **4.12** was prepared in 36% yield by phosphorylation of **4.8** with phosphitylating agent **4.10** in the presence of tetrazole, followed by subsequent oxidation with *tert*-butylhydroperoxide and deprotection with TFA (Schemes 4.2 and 4.3).[53,95]

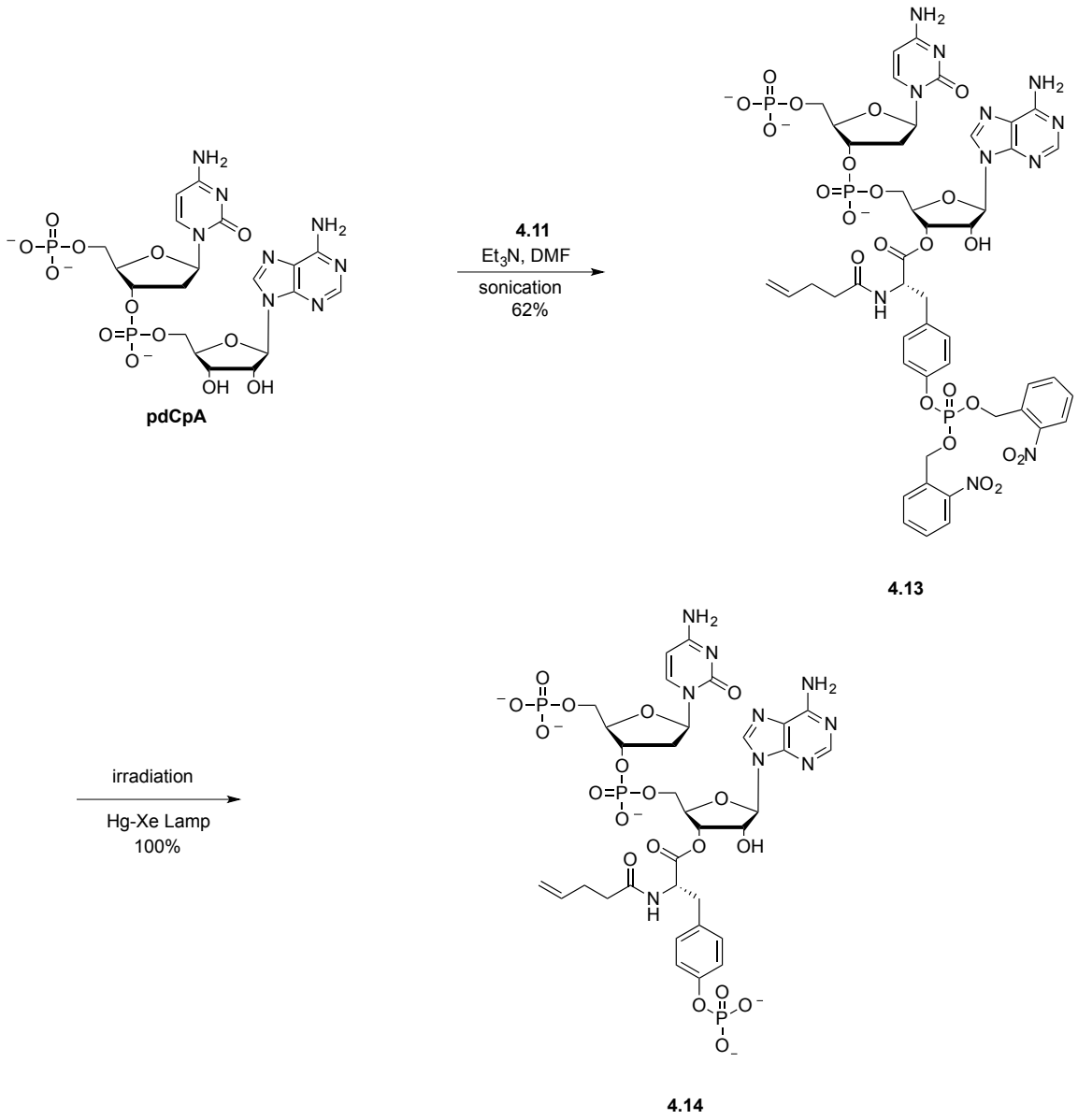


Scheme 4.2. Route employed for the synthesis of doubly caged phosphorylated tryrosine.



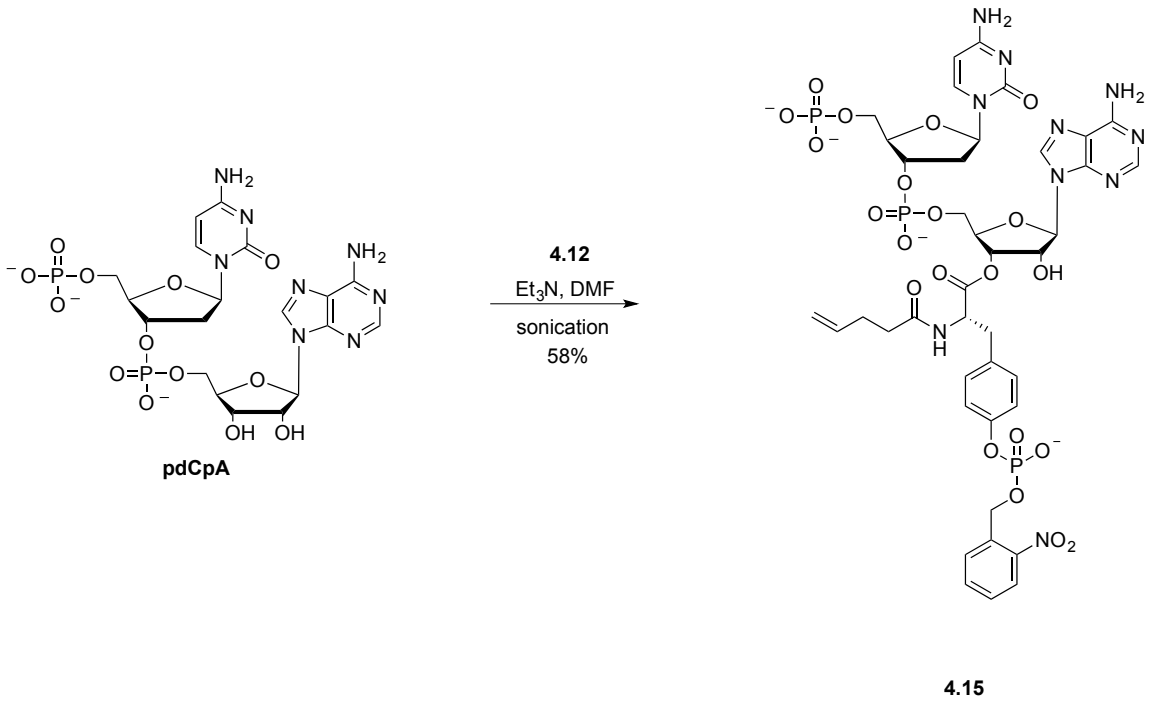
Scheme 4.3. Route employed for the synthesis of singly caged phosphorylated tryrosine.

Aminoacylation of the dinucleotide pdCpA with compounds **4.11** and **4.12,** in DMF in the presence of triethylamine (Schemes 4.4 and 4.5), proceeded quite slowly and resulted in challenging HPLC purification. Surprisingly, sonicating the reaction under similar conditions afforded compounds **4.13** and **4.15** in better yields, 62 and 58%, respectively, and shorter reaction times, ranging from 1-12 hours.



Scheme 4.4. Route employed for the synthesis di-nitrobenzylphosphotyrosyl-pdCpA and phosphotyrosyl-pdCpA.

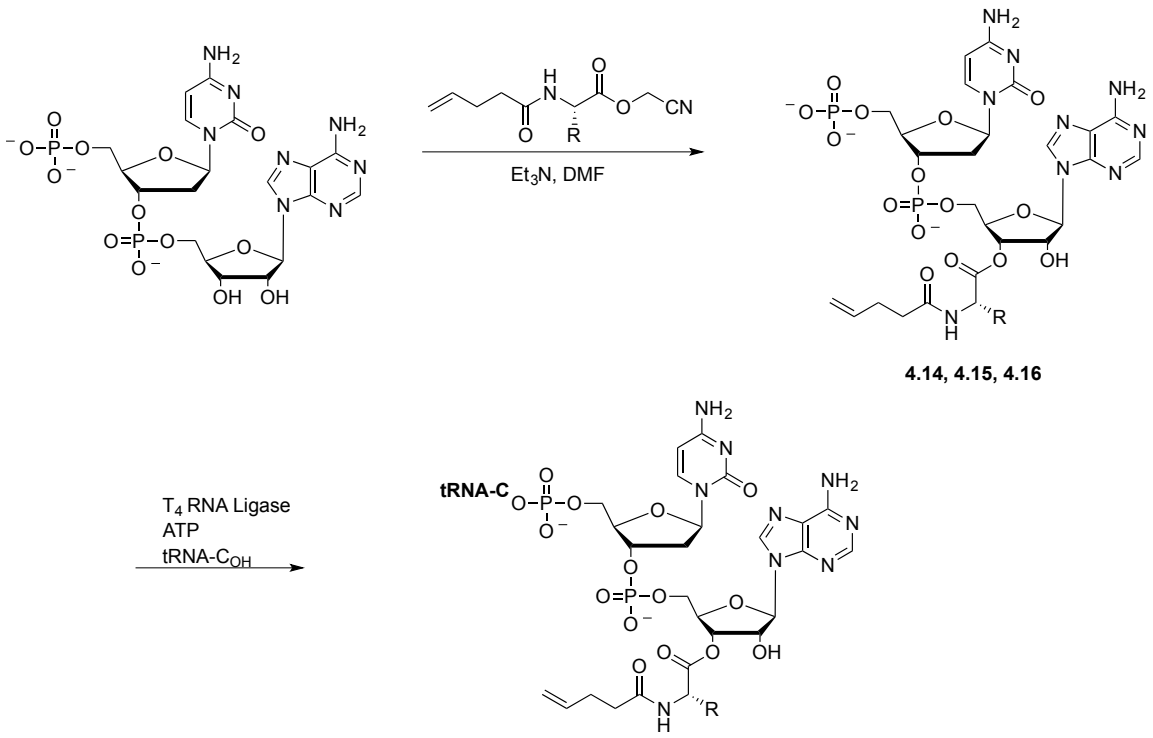
Moreover, following aminoacylation of pdCpA with compound **4.11**, irradation of **4.13** afforded the free phosphate derivative, compound **4.14**, in quantitative yield.



Scheme 4.5. Synthesis of mono-nitrobenzylphosphotyrosyl-pdCpA.

The following biological experiments were carried out by Dr. Larissa Dedkova and Rumit Maini: preparation of abbreviated $tRNA_{CUA}$, ligation of abbreviated tRNA and aminoacylated pdCpAs, *in vitro* protein translation of DHFR and proteolytic digestion of DHFR with trypsin. The preparation of abbreviated suppressor tRNAs was performed by transcription of plasmid pYRNA8,[22] followed by digestion with *Fok*I to afford tRNA-C_{OH}.[11] The elaborated, abbreviated tRNA was then purified on a DEAE-Sepharose column by eluting with a sodium chloride gradient.

Aminoacylation of the abbreviated $tRNA_{CUA}$s was carried out by a T4 RNA ligase mediated reaction with the *N*-protected, chemically synthesized pdCpA derivatives to afford the activated suppressor tRNAs (Scheme 4.6 and Figure 4.3)



Scheme 4.6. Preparation of protected phosphotyrosyl-tRNA$_{CUA}$s by T4 RNA ligase mediated ligation of amionoacyl-pdCpA and abbreviated tRNA-CC$_{OH}$.

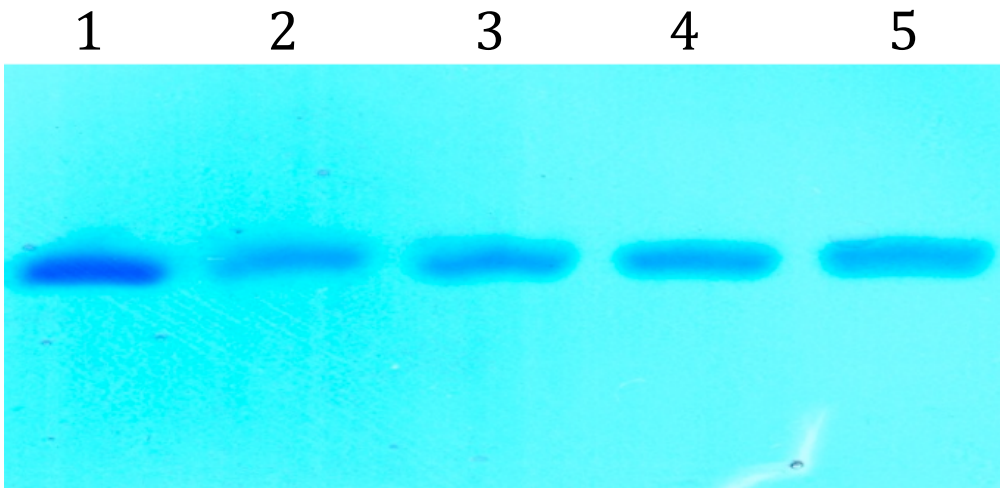


Figure 4.3. Ligation of phosphorotyrosyl-pdCpA derivatives to tRNA$_{CUA}$-C$_{OH}$. (analyzed on an acidic 8% polyacrylamide-7M urea gel)[65]. Lane 1, nonacylated tRNA$_{CUA;}$ Lane 2,

NVOC-tyrosyl-pdCpA;[53] Lane 3, phosphorotyrosyl-pdCpA (**4.15**); Lane 4, mono-nitrobenzyl-phosphorotyrosyl-pdCpA (**4.16**); Lane 5, dinitrobenzyl-phosphorotyrosyl-pdCpA (**4.14**).

The resulting ligation reactions were analyzed by polyacrylamide gel electrophoresis under acidic conditions.[65] Gel shift analysis of the ligation reaction as shown by bands with retarded migration (lanes 2-5), compared to the abbreviated tRNA (lane 1), suggests nearly quantitative ligation efficiencies.

The ability of the aminoacylated suppressor $tRNA_{CUA}$s to participate in protein synthesis was estimated by the synthesis of full length DHFR encoded by two plasmids, pETDH10 and pETDH49, that contained TAG codons at the positions corresponding to Val10 and Ser49, respectively. The *in vitro* experiments were performed using bacterial S-30 extract from *E. coli* strain BL21 (DE3). The reaction mixtures were incubated at 37 °C for 45 minutes. Aliquots from the *in vitro* translation mixtures employing pETDH10 and pETDH49 were analyzed via SDS-PAGE and the results are shown in Figures 4.4 and 4.5, respectively.

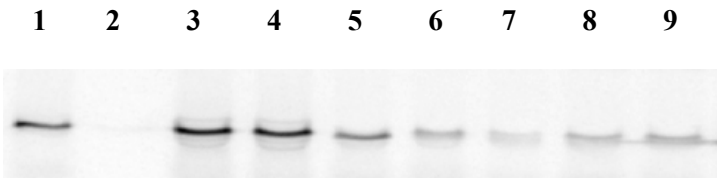


Figure 4.4. *In vitro* synthesis of DHFR utilizing monoaminoacylated $tRNA_{CUA}$s to suppress a UAG codon at position 10 of DHFR mRNA. Lane 1, tyrosyl-$tRNA_{CUA}$

(prepared from NVOC-tyrosyl-pdCpA);[53] Lane 2, nonacylated $tRNA_{CUA}$; Lane 3, phenylalanyl-$tRNA_{CUA}$; Lane 4, phenylalanyl-$tRNA_{CUA}$, prepared from phenylalanyl-pdCpA, irradiated with UV light prior to ligation; Lane 5, phosphorotyrosyl-$tRNA_{CUA}$ (both hydroxyl groups protected with nitrobenzyl groups); Lane 6, phosphorotyrosyl-$tRNA_{CUA}$ (prepared from phosphorotyrosyl-pdCpA, treated with UV light following ligation); Lane 7, phosphorotyrosyl-$tRNA_{CUA}$ (prepared from phosphorotyrosyl-pdCpA, treated with UV light prior ligation); Lane 8, phosphorothreonyl-$tRNA_{CUA}$ (one phosphate –OH group is protected with a nitrobenzyl group).

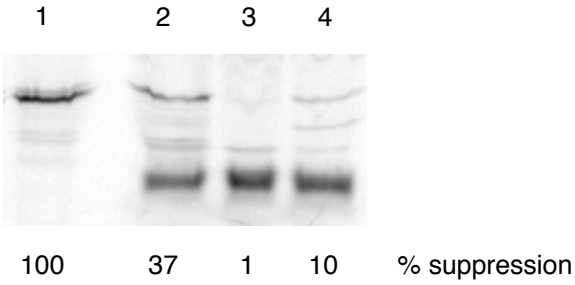


Figure 4.5. *In vitro* synthesis of DHFR utilizing di-nitrobenzyl-phosphorotyrosyl-$tRNA_{CUA}$s to suppress a UAG codon at position 49 of DHFR mRNA. Lane 1, wild-type mRNA; Lanes 2-4, modified mRNA containing a UAG codon at position 49; Lane 2, threonyl-$tRNA_{CUA}$, Lane 3, no $tRNA_{CUA}$, Lane 4, di-nitrobenzyl-phosphotyrosyl-$tRNA_{CUA}$

Analyses to confirm the introduction of phosphorotyrosine were carried out in two ways. First, we visualized proteolytic digestion of DHFR utilizing endoproteinase GluC.

Briefly, radiolabeled DHFR samples were prepared by *in vitro* translation reactions in the presence of ^{35}S-methionine and purified by successive chromatographic purifications on Ni-NTA and DEAE-Sepharose columns. The purified DHFR samples were then transferred to 50 μL of 50 mM Tris-HCl, pH 8.0, containing 0.5 mM Glu-Glu, by the use of a YM-10 ultrafiltration device. Digestion of the protein was carried out in 20 μL reactions; 20 ng of each DHFR sample (~ 10^7 cpm, as verified by liquid scintillation counting) was mixed with 2 μg of GluC endoproteinase and incubated at 25 ºC for 16 hours. Following incubation, the samples were quenched by the addition of 2 μL of 1.1 % formic acid and the final samples (5 -10 μL) were analyzed by 20% Tris-tricine polyacryalamide gel electrophoresis[27], followed by phosphorimager analysis, as shown in Figure 4.6.

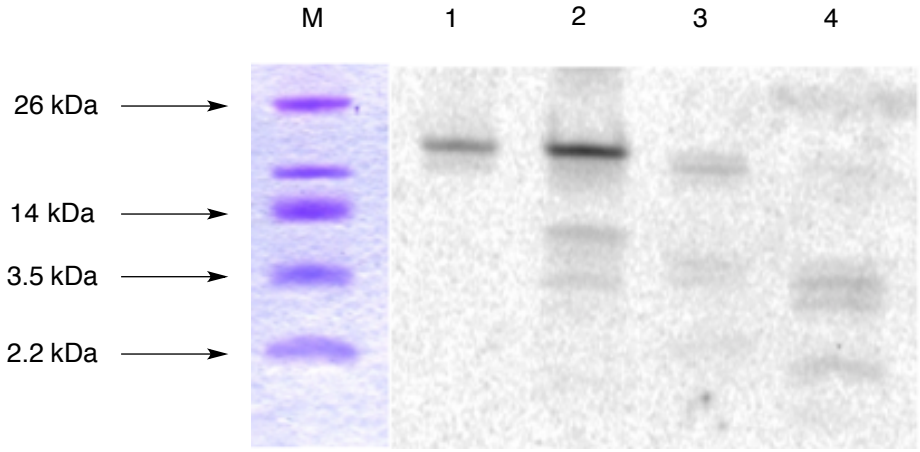


Figure 4.6. GluC endoproteinase digestion of Ser49dinitrobenzyl-phosphorotyrosyl DHFR with varying concentrations of GluC endoproteinase. Lane M, molecular weight marker; Lane 1, non-digested DHFR; Lane 2, digested sample of DHFR having di-nitrobenzyl-phosphorotryosine in position 49 and treated with 0.5 μg/mL GluC

endoproteinase; Lane 3, digested sample of DHFR having di-nitrobenzyl-phosphorotryosine in position 49 and treated with 1 μg/mL GluC endoproteinase; Lane 4, digested sample of DHFR having di-nitrobenzyl-phosphorotryosine in position 49 and treated with 2 μg/mL GluC endoproteinase.

One aliquot of the digested protein was then irradiated to uncage the peptide of interest, while another aliquot was left untreated. Both aliquots were then analyzed via reverse phase-HPLC by co-injection with the synthetically prepared peptide standard. Fractions were collected from both HPLC purifications and analyzed by liquid scintillation counting. It should be noted that an authentic peptide, GMGIVRD(phosphoroT)ALAAILSIM (Figure 4.7) was synthesized via solid phase peptide synthesis (SPPS) as a standard and was co-injected (Figure 4.8) with all samples. Initial attempts to synthesize the peptide utilizing an automated system were unsuccessful; the reason for that still remains unclear.

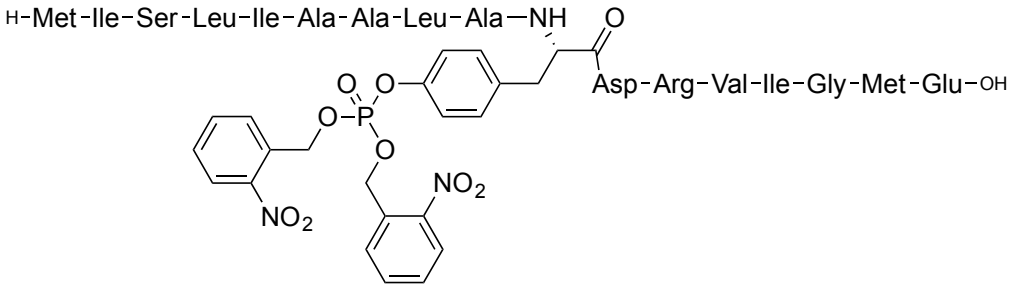


Figure 4.7. Structure of peptide synthesized by SPPS.

We reasoned that introducing the fully protected phosphorylated tyrosine hindered the coupling event, due to the bulkiness of the side chain. Therefore, the peptide

was synthesized by an automated system and removed from the capillary prior to the introduction of the phosphorylated tyrosine, GMGIVRD. Following removal of the resin from the capillary, the coupling of Fmoc-phosphorotyrosine was carried out on the resin in a sealed tube and was allowed to react overnight under normal solid phase coupling conditions. Quantitative Fmoc cleavage indicated the coupling reaction was complete and the resin was placed back into the capillary; the remainder of the peptide was synthesized on the automated system. All *in vitro* synthesized proteins were purified by NiNTA columns, desalted, and concentrated prior to further analysis.

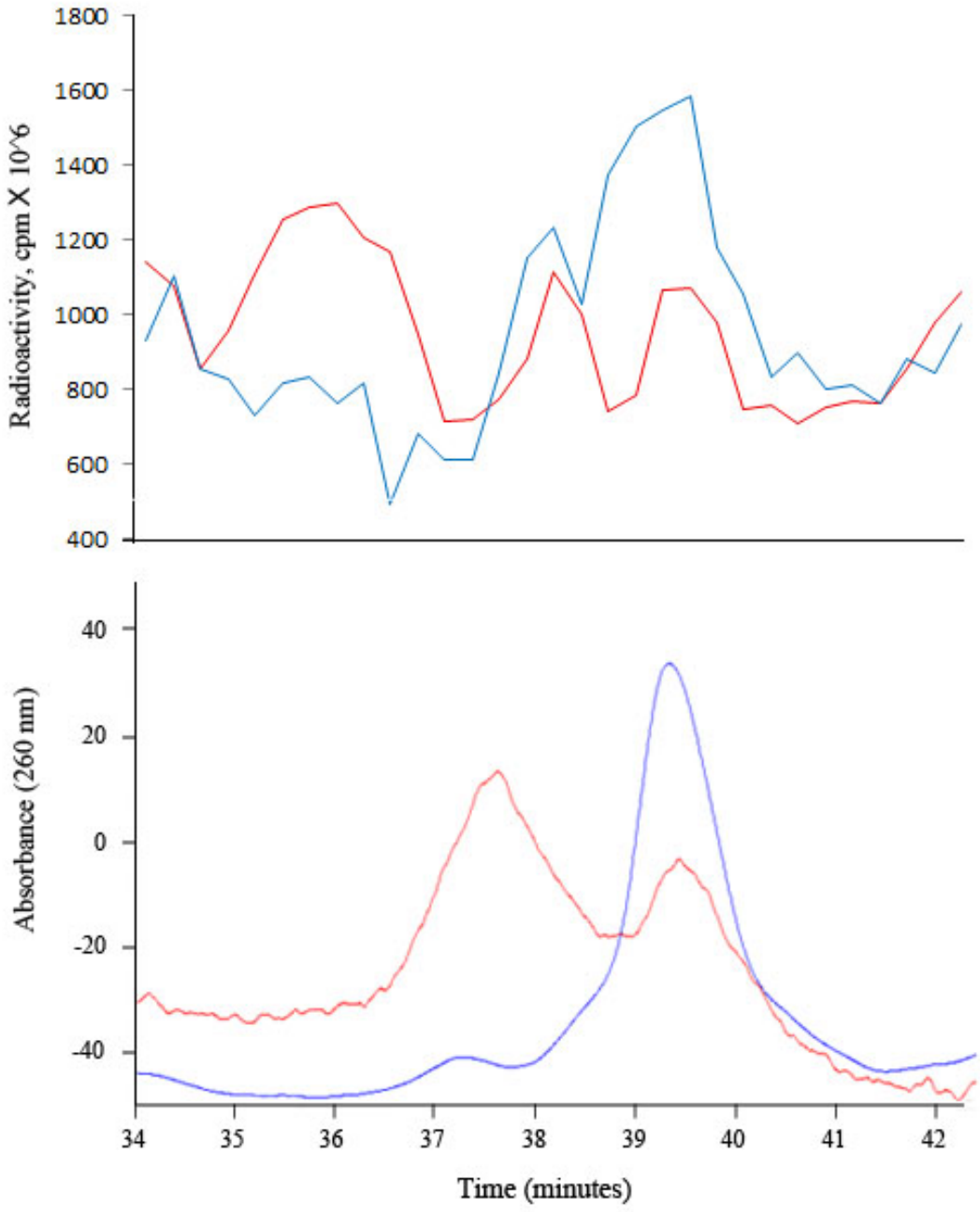


Figure 4.8. HPLC profiles of the authentic peptide before (blue) and after (red) irradiation, visualized by absorbance at 260 nm (bottom) or by radioactivity (upper).

A mobility shift from the apex of each peak is quite apparent and was expected. Uncaging of the peptide liberates the free phosphate, resulting in a slightly more polar peptide. Notably, not all of the peptide was uncaged (deprotected), as can be seen by the results shown in Figure 4.8. Moreover, it should be further noted that the completely uncaged peptide does not absorb at 260 nm. The red trace, corresponding to the irradiated peptide, can be described as follows: the peak at 39 minutes can be putatively identified as the bis-nitrobenzylphospho peptide and the peak at 38 minutes as the mono-nitrobenzylphospho peptide. Analysis of the *in vitro* synthesized DHFR did not afford enough product to be identified by UV detection. Characterization was therefore performed by liquid scintillation counting of fractions, as shown by the lower panel in Figure 4.8.

The sensitivity of the scintillation analysis offers additional information about the products after irradiation of the peptides. The appearance of a new peak at 36.4 minutes (red), is apparent following irradiation. This peak can potentially be identified as the product resulting from the loss of two nitrobenzyl groups, a fact that would not be detectable by UV.

Further confirmation of ribosomal incorporation of the modified amino acids, as directed by the UAG codon in the mRNA, was illustrated by MALDI-TOF MS characterization. MALDI-TOF MS analysis had to be performed on the protein containing a point mutation at position 49 because the His tag at the *N*-terminus of DHFR (used for purification) diminishes the response of *N*-terminal peptides in MALDI-TOF MS.

DHFR samples for MALDI-TOF MS analysis were prepared from non-

radiolabeled reaction mixtures employing pETDH49 plasmid with a TAG codon at position 49 of the DHFR gene. Following expression, the reaction mixture was subjected to a 12% SDS-polyacrylamide gel and subsequently stained with Coomassie R-250 for DHFR identification. The gel was de-stained until the background was clear. The DHFR band was identified, extracted from the gel, and proteolytically digested by trypsin.

A second mutant, Ser49Phe, prepared from misacylated tRNA was prepared in a similar fashion for comparative MALDI-TOF MS analysis. The resulting peptide fragments were analyzed via MALDI-TOF MS by preparing solutions of the digested protein on a dihydroxybenzoic acid (DHB) matrix. The spectrometric data for the wild-type, Ser49Phe and Ser49dinitrobenzylphosphoTyr DHFRs are shown below in Figures 4.9, 4.10 and 4.11, respectively.

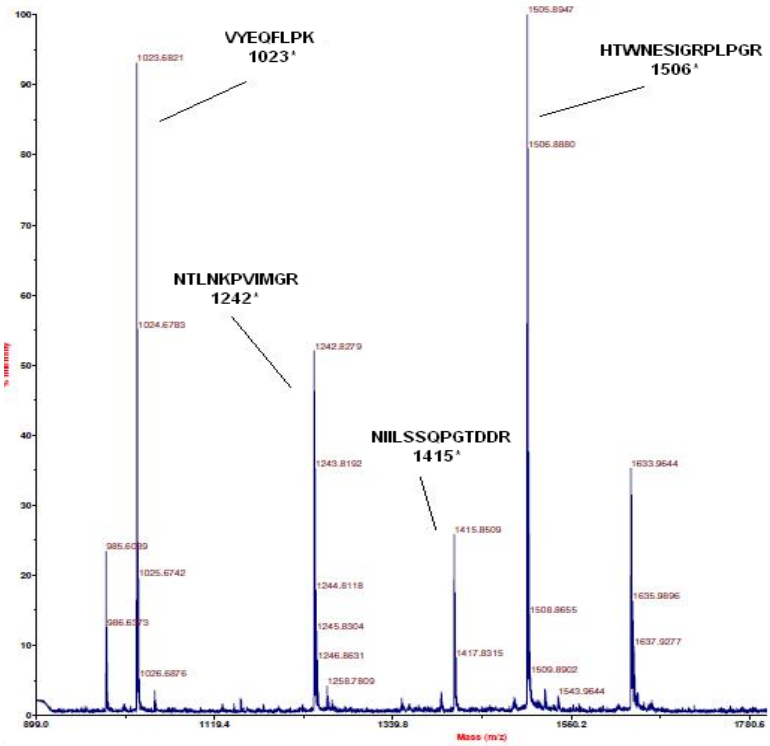


Figure 4.9. MALDI-TOF MS of tryptic fragments of wild-type DHFR.

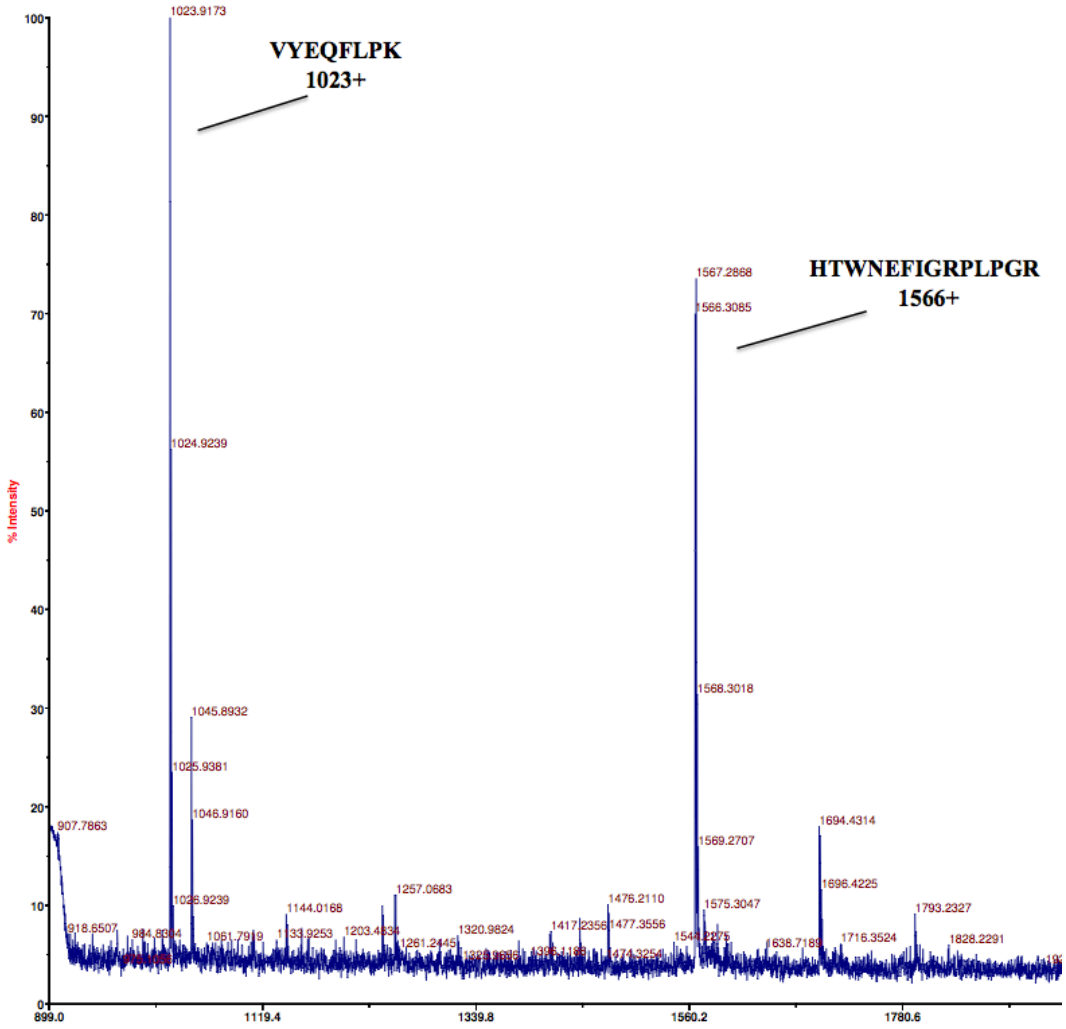


Figure 4.10. MALDI-TOF MS of tryptic fragments of Ser49Phe DHFR.

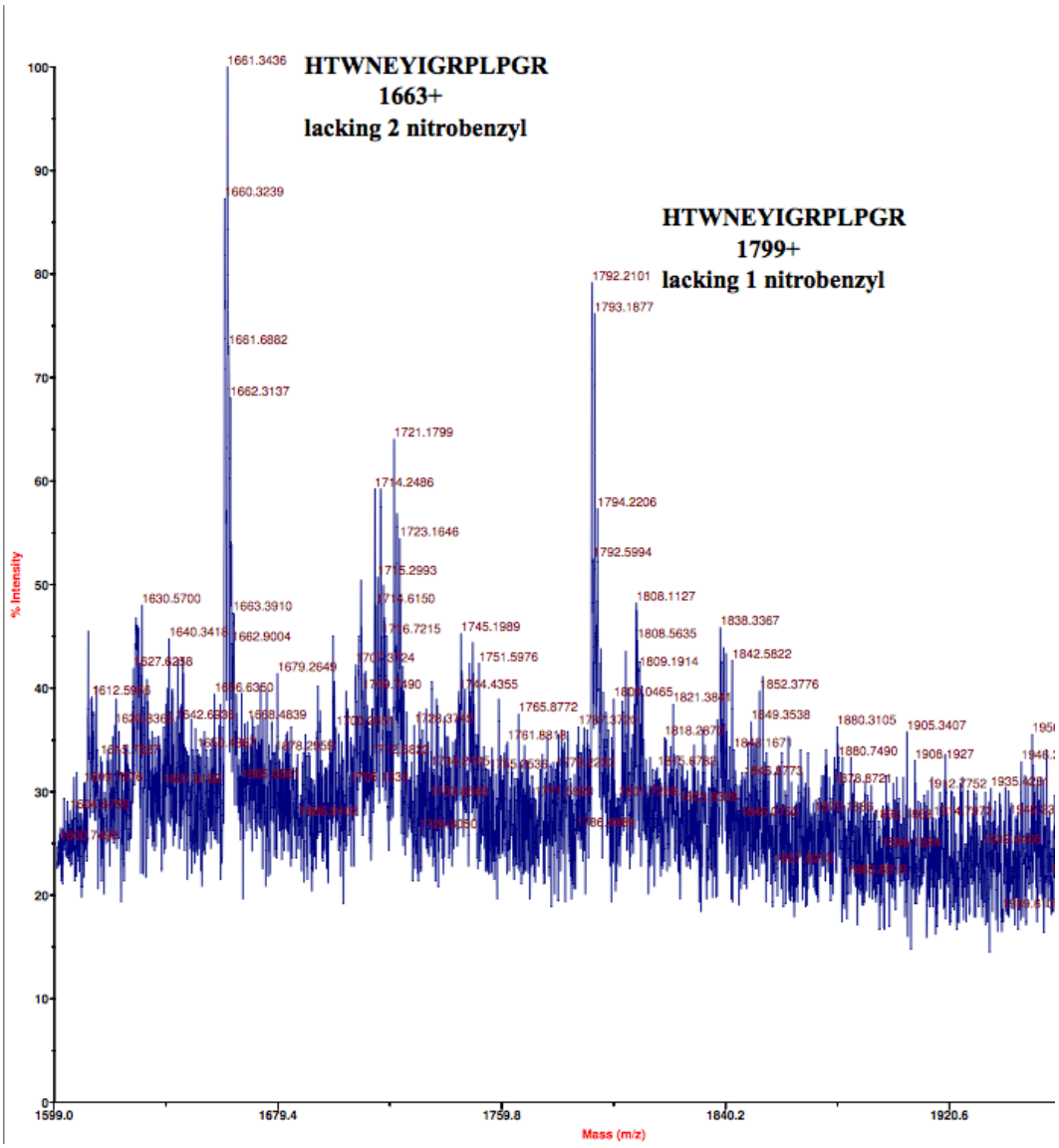


Figure 4.11. MALDI-TOF MS of tryptic fragments of Ser49dinitrobenzyl-phosphotyrosyl DHFR.

Comparative analysis of the peptide fragments from the two mutants and wild-type DHFR further verified the incorporation of the desired amino acid at the predetermined position. For Ser49Phe DHFR, a tryptic fragment containing phenylalanine at the fifth position of the peptide is expected to have a molecular mass of 1566. Accordingly, as seen in figure 4.10, there was a molecular ion peak observed at *m/z*

1567 that is not present in the tryptic digest of wild-type DHFR. Moreover, the same molecular ion peaks corresponding to the remaining six tryptic digests containing amino acids 13-106 of DHFR could be seen in both wild-type and Ser49Phe DHFR. Analogously, Ser49dinitrobenzylphosphoTyr DHFR showed molecular ion peaks corresponding to *m/z* 1792 and 1661,which supports the incorporation of the di-nitrobenzylphosphotyrosine resulting from the loss of one and both nitrobenzyls (Figure 4.11). A summary of the tryptic digest patterns are outlined in Table 4.1.

Table 4.1.
MALDI-TOF MS analysis of tryptic digests of wild-type and modified DHFR samples.

Position	Peptide sequence	MALDI-TOF MS analysis, molecular mass, Da					
		Wild-type		Modified DHFR 1		Modified DHFR 2	
		Est	MS	Est	MS	Est	MS
34-44	NTLNKPVIMGR	1242	1242.8	1242	1242.6	1242	1242.8
45-57	HTWE**S**IGRPLPGR	1506	1505.8	**1566**	**1566.3**	***1797**	**1793.2**
59-71	NIILSSQPGTDDR	1415	1415.8	1415	1415.6	1415	1415.8
72-76	VTWVK	632	632.5	632	634.3	632	632.4

Modified DHFR 1: phenylalanine in position 49; modified DHFR 2: phosphorotyrosine analogue **4.3** in position 49. Position 49 is indicated in red. *The molecular mass resulting from the loss of one nitrobenzyl protecting group on the phosphate.

Table 4.1 presents the anticipated molecular mass and observed molecular ion peaks resulting from proteolytic digestion of wild-type and modified DHFRs 1 and 2, containing Phe or **4.3**, respectively, with trypsin. The wild-type peptide and mutant containing Phe at position 49 showed ion peaks at 1506 (anticipated molecular mass 1506) and 1566 (anticipated molecular mass 1566), respectively, which is in good agreement with the expected molecular masses. The molecular ion peak observed for

modified DHFR 2 did not correspond to the anticipated molecular mass. The observed ion for modified DHFR 2, however, is in close agreement with the peptide resulting from loss of one nitrobenzyl protecting group on the phosphate. Similar occurrences were observed when preparing compounds **4.1**-**4.3**.

4.3. Discussion

The synthetic route leading to the three phosphorylated tyrosines (**4.1**-**4.3**, Figure 4.3), as well as the aminoacylated dinucleotides (**4.14**-**4.16**, Scheme 4.4 and 4.5) for the preparation of phosphotyrosyl-$tRNA_{CUA}$ variants (Scheme 4.6) was very efficient because they all share a common intermediate, compound **4.7**. Phosphorylation of **4.8** with phosphitiling agents **4.9** or **4.10**, afforded caged phosphotyrosine dervatives as the mono- or diesters, **4.11** and **4.12**, bearing one or two nitrobenzyl groups, respectively. Yields and reaction times for phosphorylation of the phenolic alcohol are in good agreement with previously published reports.[95] Initially, oxidation of the phosphite was performed with iodine. However, purification proved to be rather difficult and was simplified by employing *tert*-butylhydroperoxide. Aminoacylation of pdCpA with **4.11** and **4.12** afforded aminoacylated pdpCpAs **4.14** and **4.16** and irradiation of **4.14** afforded the third aminoacylated pdCpA **4.15**, containing a free phosphate. The efficiency of the aminoacylation reaction was greatly enhanced by sonicating the reaction mixture. Interestingly, phosphorotyrosyl-pdCpA, mono-nitrobenzyl-phosphorotyrosyl-pdCpA and di-nitrobenzyl-phosphorotyrosyl-pdCpA were ligated to abbreviated $tRNA_{CUA}$s very efficiently, in yields greater than 80% (Figure 4.4). The extent to which the phosphorylated tyrosine derivatives were protected on the phosphate group appears to

have a direct effect on the ability of the tRNA to suppress the nonsense codon in *in vitro* protein synthesis (Figure 4.5). The ability of the di-nitrobenzyl-phosphotyrosyl-$tRNA_{CUA}$ to participate in DHFR synthesis in our *in vitro* system via suppression of a UAG codon at position 10 of the corresponding mRNA was reasonably efficient, ~31% relative to the suppression efficiency of Phe-$tRNA_{CUA}$. Mono-nitrobenzyl-phosphorotyrosyl-$tRNA_{CUA}$ and phosphorotyrosyl-$tRNA_{CUA}$ were able to suppress the UAG codon with efficiencies of only 11 and 5%, respectively, as compared to phenylalanyl-$tRNA_{CUA}$ (Figure 4.5). It should be noted that phenylalanyl-$tRNA_{CUA}$ typically exhibits suppression efficiencies in the range of 60-90%. The inability of charged aminoacylated $tRNA_{CUA}$s to efficiently participate as suppressor tRNAs in *in vitro* protein synthesis has already been reported. For example, it was shown that aspartyl-tRNAs suppress nonsense codons much more efficiently when the amino acid side chain is protected. This effect is also demonstrated by the phosphoryltyrosyl-tRNAs prepared from the pdCpA derivatives discussed here. Phosphorylated tyrosyl-tRNAs bearing two protecting groups on the phosphate were able to suppress a UAG codon more efficiently than those containing one or no protecting groups.

To provide evidence of the successful incorporation of the phosphorylated tyrosines described here into DHFR, two independent experiments were performed: co-elution on HPLC and identification of peptide fragments from digestion of *in vitro* synthesized proteins as compared to a synthetically prepared standard; and MALDI-TOF MS characterization. The synthetic standard was prepared by modified solid phase peptide synthesis. After synthesis, the peptide was cleaved from the resin and purified by RP-HPLC. Irradiation of the synthetically prepared peptide standard afforded a

substantial decrease in the amount of peptide bearing two nitrobenzyl groups, and an increase in the mono-nitrobenzylated peptide, as evidenced in Figure 4.8. Irradiation resulting in a peptide that lost both nitrobenzyl groups would not be detectable at 260 nm, and liquid scintillation counting was required to identify the peptide prepared in the *in vitro* system, as can be seen in Figure 4.8.

Further verification of the incorporation of phosphorotyrosine is evident by the mass spectrometric data. Due to the relatively small difference in molecular weight between wild type DHFR and Ser49phosphoTyr DHFR, analysis of the whole protein is not advisable. To circumvent this problem, the protein was proteolytically digested by trypsin and the resulting fragments were analyzed by mass spectrometry. Initial attempts to digest the protein with GluC endoproteinase proved inadequate, as the enzyme is less efficient than trypsin and the molecular ion peaks corresponding to the peptide fragments could not be observed by MALDI-TOF MS. Therefore, following purification by Ni-NTA, DEAE-Sephadex, and SDS-polyacrylamide gel electrophoresis, the proteins were treated with trypsin according to a published "in-gel" protocol.[96] The resulting peptide fragments were subjected to MALDI-TOF MS analysis, (Figures 4.9, 4.10 and 4.11), performed by Dr. Larissa Dedkova and Rumit Maini. Two samples (wild-type DHFR and Ser49Phe DHFR) were used as controls for MALDI-TOF MS analysis of the mutant DHFR containing di-nitrobenzyl-phosphorotyrosine (the only phosphorotyrosyl employed for MALDI-TOF MS analysis due to expression yields). Comparative analysis of the peptide fragments from the two mutant and wild-type DHFRs further verified the incorporation of the desired amino acid at the predetermined position. Regarding Ser49Phe DHFR, a tryptic fragment containing phenylalanine at the fifth position of the

peptide is expected to have a molecular mass of 1566. Accordingly, as seen in Figure 4.10, there was a molecular ion peak observed at *m/z* 1567 that was not present in the tryptic digest of wild-type DHFR. Moreover, the same molecular ion peaks corresponding to the other tryptic peptide fragments containing amino acids 13-106 of DHFR were present for both wild-type and Ser49Phe DHFR. Analogously, Ser49dinitrobenzylphosphoTyr DHFR showed molecular ion peaks corresponding to *m/z* 1792 and 1661 (Figure 4.11) which is in support of the successful incorporation of the di-nitrobenzylphosphorotyrosine resulting from the loss of one or both nitrobenzyl groups.

Accordingly, incorporation of deprotected phosphorylated tyrosines should be amenable to the same type of analysis. However, the poor efficiency with which they participate in ribosomally mediated protein synthesis precludes our ability to characterize the products. Recently, Hecht and co-workers described ribosomal selection for the synthesis of DHFR containing β-amino acids, with suppression efficiencies up to 18%.[30] If this technology were developed to incorporate negatively charged amino acids it would be possible to pursue the incorporation of a variety of phosphorylated amino acids.

4.4. Experimental

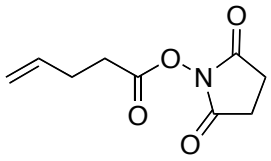
4.4.1. General methods and materials

L-Tyrosine was purchased from Alfa Aesar. All other reagents and solvents for chemical synthesis were purchased from Sigma Aldrich Co. and used without further purification. All reactions involving air- or moisture-sensitive reagents or intermediates were performed under argon. Flash chromatography was performed using Silicycle silica gel (40–60 mesh). Analytical TLC was performed using EM silica gel 60 F_{254} plates

(0.25 mm) and was visualized by UV irradiation (254 nm). ^{1}H and ^{13}C NMR spectra were obtained using a 400 MHz Varian NMR instrument. Chemical shifts are reported in parts per million (ppm, δ) referenced to the residual ^{1}H resonance of the solvent ($CDCl_3$, δ 7.26; CD_3OD, δ 3.31, DMSO-d_6, δ 2.50). ^{13}C NMR spectra were referenced to the residual ^{13}C resonance of the solvent ($CDCl_3$, δ 77.16; CD_3OD, δ 49.00, DMSO-d_6, δ 39.52). Splitting patterns are designated as follows: s, singlet; d, doublet; dd, doublet of doublets; t, triplet; q, quartet; quint, quintet; m, multiplet; br, broad. HPLC purification was performed with a Waters 600 pump coupled with a Varian *ProStar* 340 detector and a Grace Econosil C_{18} column (250 × 10 mm, 5 μm). The tetra-*n*-butylammonium (TBA) salts of pdCpA were prepared using Dowex 50WX8, 200-400 mesh activated in its TBA form. High-resolution mass spectra were obtained at the Arizona State University CLAS High Resolution Mass Spectrometry Laboratory or at the Michigan State University Mass Spectrometry Facility. Ni-NTA agarose was obtained from Qiagen Inc. (Valencia, CA). DNA oligonucleotides were purchased from Integrated DNA Technologies (Coralville, IA). DEAE-Sepharose, ammonium persulfate, acrylamide, *N*, *N′*-methylene-bis-acrylamide, acetic acid, potassium glutamate, ammonium acetate, dithiothreitol, magnesium acetate, phospho(enol)pyruvate, *Escherichia coli* tRNA, isopropyl β-D-thiogalactopyranoside (IPTG), ATP, GTP, CTP, UTP, cAMP, amino acids, rifampicin, formamide and 7-diethylamino-3-(4′-maleimidophenyl)-4-methylcoumarin (CPM) were obtained from Sigma-Aldrich (St. Louis, MO). Tris and SDS were obtained from Bio-Rad Laboratories (Hercules, CA). [^{35}S]-methionine (1000 Ci/mmol, 10 μCi/μL) was purchased from PerkinElmer Inc. (Boston, MA). Protease inhibitor (complete, EDTA-free) was obtained from Boehringer Mannheim Corp. (Indianapolis, IN). T4 RNA ligase

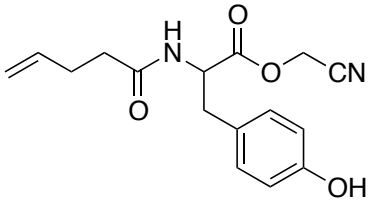
and T4 polynucleotide kinase were purchased from New England Biolabs Inc. (Ipswich, MA). Phosphorimaging analysis was performed using an Amersham Biosciences Storm 820 equipped with ImageQuant version 5.2 software from Molecular Dynamics. UV spectral measurements were made using a Perkin-Elmer Lamdba 20 UV/vis spectrometer.

4.4.2. Synthesis of *N*-petenoyl-phosphorotyrosine cyanomethyl esters



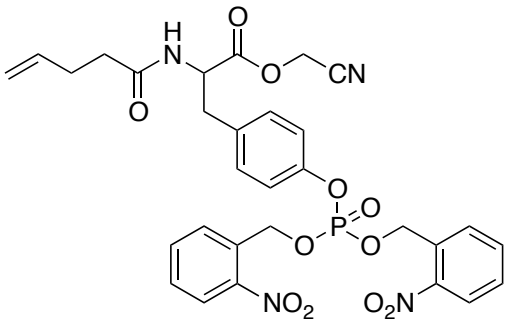
4-pentenoyloxysuccinimide ester(4.19)[25]

To a solution containing 5.0 mL (49.0 mmol) of pentenoic acid and 5.64 g (49.0 mmol) of *N*-hydroxysuccinimide in 100 mL CH_2Cl_2 was added 10.3 g (50.0 mmol) of *N*,*N*′-dicyclohexylcarbodiimide. After stirring at room temperature for 2 hours, the reaction mixture was filtered and the filtrate was concentrated under diminished pressure. Crystallization of the crude product from ether-petroleum ether afforded 4-pentenoyloxy succinimide ester as colorless crystals: yield 7.63 g (79%); R_f 0.43 (1:1 ethyl acetate–hexane); mp 45-47 °C (47-48 °C).



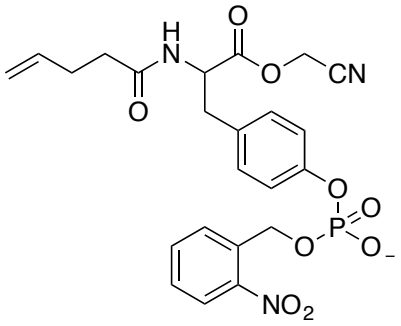
***N*-pentenoyl-L-tyrosine cyanomethyl ester (4.8)**

To a solution containing 1.00 g (5.52 mmol) of tyrosine and 1.50 g (11.0 mmol) of K_2CO_3 in 15 mL of H_2O (a few drops of 6 N NaOH were added for solubility) was added 1.30 g (6.62 mmol) of 4-pentenoyloxy succinimide ester (**4.19)**. The reaction mixture was stirred at room temperature for 16 hours. 80 mL of 1 N $NaHSO_4$ was then added to the reaction mixture. The aqueous layer was extracted with three 125-mL portions of ethyl acetate. The combined organic extract was dried ($MgSO_4$) and concentrated under diminished pressure. The crude product was then dissolved in 10 mL of acetonitrile; to the solution was added 4.6 mL (33.1 mmol) of triethylamine and 2.1 mL (33.1 mmol) of chloroacetonitrile. The reaction mixture was stirred at room temperature for 20 hours then diluted with 150 mL of ethyl acetate. The organic layer was washed with three 75-mL portions of sat. $NaHCO_3$. The organic layer was dried ($MgSO_4$) and concentrated under diminished pressure. The crude residue was purified on a silica gel column (15 cm x 5 cm); elution with 1:1 ethyl acetate–hexanes afforded *N*-pentenoyl-L-tyrosine cyanomethyl ester **(4.8)** as a colorless solid: yield 1.03 g (62%); mp. 52-53 °C; R_f 0.45 (1:1 ethyl acetate–hexane); ^{1}H NMR δ 2.26-2.38 (m, 4H), 3.09-3.19 (m, 2H), 4.68, 4.79 (ABq, 2H, J = 15.6 Hz), 4.91-5.06 (m, 3H), 5.72-5.84 (m, 2H), 7.12-14 (m, 2H) and 7.29-7.35 (m, 3H); ^{13}C NMR δ 14.1, 21.1, 35.2, 36.6, 49.0 53.3, 60.6, 114.1, 115.9, 126.1, 130.3, 136.5, 156.0, 170.5, 171.7 and 173.3, mass spectrum (ESI) *m/z* 325.1 $(M+Na)^+$ (theoretical *m/z* 325.1).



***N*-4-pentenoyl-phosphoro(1,2-dinitrobenzyl)-L-tyrosine cyanomethyl ester (4.11)**

To a stirred solution of 416 mg (0.95 mmol) of *O*-dinitrobenzyl-*N*,*N*-diisopropyl phosphoramidite and 92.0 mg (1.31 mmol) of tetrazole in 5 mL of anhydrous acetonitrile was added 107 mg (0.35 mmol) of *N*-pentenoyl-L-tyrosine cyanomethyl ester (**4.8)**. Following stirring for 30 minutes at r.t., a 5 M solution of *tert*-butylhydroperoxide in decane (262 μl, 1.31 mmol) was added. Following stirring for 10 minutes, the reaction mixture was diluted with 50 mL ethyl acetate and washed with three 30 mL portions of sat $NaHCO_3$. The organic layer was dried ($MgSO_4$) and concentrated under diminished pressure. The residue was purified by silica gel chromatography (15 cm x 5 cm); elution with 1:1 ethyl acetate–hexane afforded the phosphate product **4.11** as a yellow oil: yield 138 mg (60%). ^{1}H NMR δ 2.28-2.35 (m, 4H), 3.07-3.18 (m, 2H), 4.74 (ABq, 2H, J = 12Hz), 4.88-4.93 (m, 1H), 4.97-5.05 (m, 2H), 5.61-5.63 (m, 4H), 5.72-5.82 (m, 1H), 6.04 (d, 1H, J = 8 Hz), 7.11 (d, 2H, J = 8 Hz), 7.18 (d, 2H, J = 12 Hz), 7.51 (t, 2H, J = 8 Hz), 7.66-7.74 (m, 4H) and 8.13 (d, 2H, J = 8 Hz); ^{13}C NMR δ 29.6, 35.7, 49.3, 63.9, 114.4, 117.3, 120.9, 120.8, 125.3, 128.8, 129.0, 131.0, 134.4, 147.2, 151.9, 170.7 and 172.5; mass spectrum (MALDI) *m/z* 540.3 $(M+H-C_7H_7NO_2)^+$ (theoretical *m/z* 540.1).



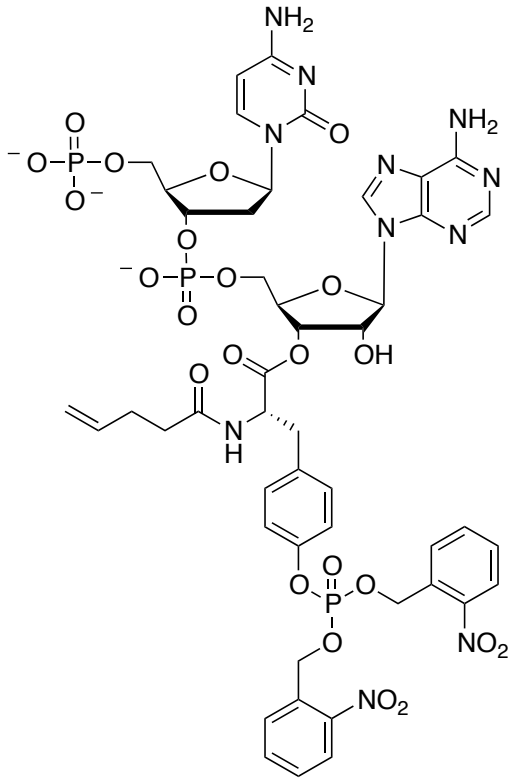
***N*-4-pentenoyl-phospho(1-nitrobenzyl-2-*tert*-butyl)-L-tyrosine cyanomethyl ester (4.12)**

To a stirred solution containing 272 mg (0.83 mmol) of *O*-1-nitrobenzyl-*O*′-*tert*-butyl-*N*,*N*-diisopropyl phosphoramidite and 70 mg (0.99 mmol) of tetrazole in 2 mL of anhydrous ACN was added 100 mg (0.33 mmol) of *N*-4-pentenoyl-L-Tyrosine cyanomethyl ester. Following stirring for two days the reaction mixture was concentrated then re-dissolved in 10 mL DCM and washed with two 4-mL portions saturated $NaHCO_3$, dried ($MgSO_4$) and concentrated under reduced pressure. The phosphite was then dissolved in 2 mL anhydrous DCM and t-BuOOH (200 μL, 5M solution in decane) was added dropwise. Following stirring for 30 minutes, the reaction was diluted with 3 mL of DCM and washed with two 3-mL portions of saturated $NaHCO_3$. The organic layer was dried ($MgSO_4$) and concentrated under reduced pressure. The crude phosphate was then dissolved in 2 mL of anhydrous DCM followed by addition of 300 μL TFA. Following stirring for 30 minutes, the reaction was concentrated and the crude mixture was purified via preparitve HPLC (93:7 to 0:100 water/acetonitrile/0.1% TFA over 35 minutes, t_R = 24.3 min) to afford the product as a brown oil: yield 62 mg (36%); ^{1}H NMR δ 2.3 (s, 4H), 3.01-3.13 (m, 2H), 4.72 (ABq, 2H, J = 16 Hz), 4.85 (bs, 1H), 4.95-5.03 (m, 2H), 5.51-5.55 (m, 2H), 5.68-5.80 (m, 1H), 6.35 (bs, 1H), 7.07 (d, 2H, J = 8 Hz), 7.13 (d, 2H, J = 8

Hz), 7.47 (t, 1H, J = 16 Hz), 7.65 (t, 1H, 16 Hz), 7.76 (d, 1H, J = 8 Hz) and 8.11 (d, 1H, J = 4Hz); ^{13}C NMR δ 29.1, 35.0, 36.7, 48.9, 52.9, 66.2, 66.3, 113.7, 115.9, 120.3, 120.4, 124.9, 128.2, 128.7, 130.5, 132.5, 134.2, 136.3, 170.0 and 173.0; mass spectrum (ESI) *m/z* 518.1328 $(M+H)^+$ ($C_{23}H_{25}N_3O_9P$ requires 518.1329).

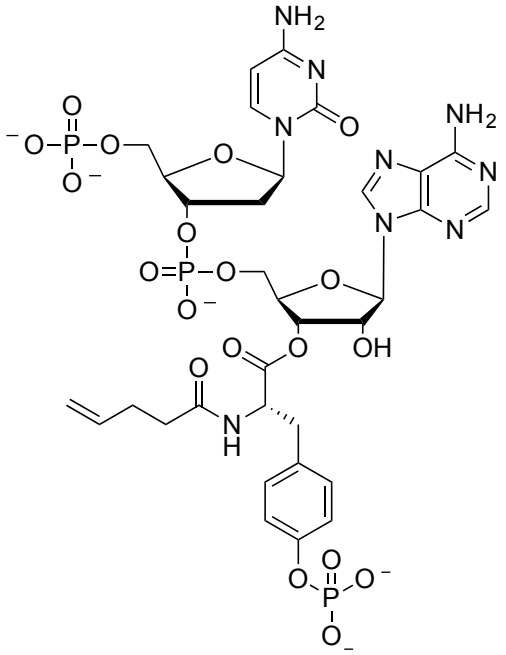
4.4.3. Synthesis of phosporotyrosyl-pdCpA derivatives

To a conical vial containing 1 equivalent of the tris(tetrabutylammonium) salt of pdCpA (TBA-pdCpA) in anhydrous DMF was added 5 eq. of *N*-4-pentenoyl amino acid cyanomethylester or NVOC-amino acid cyanomethylester and sonicated at room temperature. Following consumption of pdCpA, the reaction was purified by reverse phase HPLC using a gradient of 1%→65% acetonitrile in 50 mM ammonium acetate buffer (pH = 4.5) over 45 minutes. The appropriate fractions were collected and lyophilized to afford the product.



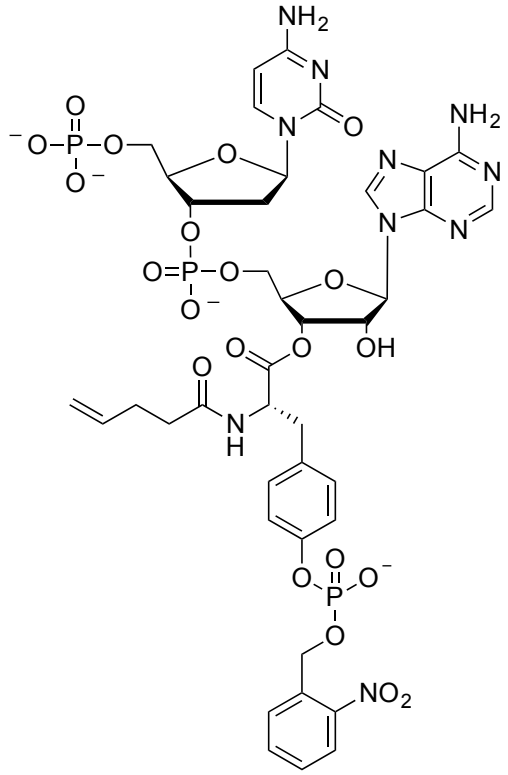
Mono-2'(3')-*O*-(dinitrobenzylphospho)-L-tyrosyl-pdCpA (4.14)

The reaction was carried out as described above starting from 10.0 mg (7.4 μmol) of TBA-pdCpA and 24 mg (36.7 μmol) of compound **4.11** to afford monoaminoacylated pdCpA **(4.14)**: t_R 30 minutes; mass spectrum (ESI) *m/z* 1230.2360 $(M-H)^-$ ($C_{47}H_{51}N_{11}O_{23}P_3$ requires *m/z* 1230.2372).



Mono-2'(3')-*O*-phospho-L-tyrosyl-pdCpA (4.15)

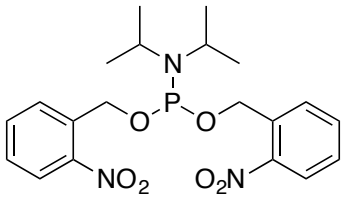
1.3 mg (1.19 μmol) of **4.14** in 100 μL DMSO was irradiated using a Hg-Xe lamp at 0 °C for 5 minutes. Following irradiation HPLC analysis showed all starting was consumed and the solution was used directly in the ligation step without further purification. The new peak (t_R 13 minutes) gave mass spectrum (ESI) *m/z* 960.1712 $(M-H)^-$ ($C_{33}H_{41}N_9O_{19}P_3$ requires *m/z* 960.1737).



Mono-2'(3')-*O*-(nitrobenzylphospho)-L-tyrosyl-pdCpA (4.16)

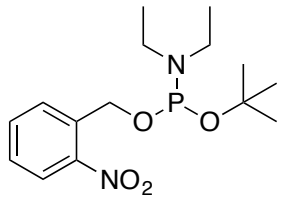
The reaction was carried out as described above starting from 5.0 mg (3.7 μmol) of TBA-pdCpA and 9.5 mg (18.4 μmol) of **4.13** to afford monoaminoacylated pdCpA (**4.16**): t_R 18 and 19 minutes; mass spectrum (ESI) *m/z* 1095.2013 $(M-H)^-$ ($C_{40}H_{46}N_{10}O_{21}P_3$requires *m/z* 1095.2057).

4.4.4. Synthesis of phosphoramidites



Bis(2-nitrobenzyl)-*N,N*-diisopropylphsophoramidite (4.17)[53]

To a 0 °C cooled solution of dichloro(diisopropylamino)phosphine (5.0 g, 24.7, mmol) in 20 mL of THF was added Hunigs Base (12.8 mL, 73 mmol) followed by 2-nitrobenzyl alcohol (7.56 g, 49.4 mmol) in 50 mL of THF. The reaction was allowed to stir at room temperature for 3 h. The precipitated solid was filtered and washed with 100 mL ethyl acetate. The organic phase was then wash with sat. $NaHCO_3$ and NaCl. The filtrate was dried ($MgSO_4$) and concentrated under diminished pressure to afford the product as a white solid:yield 5.3 g (53%); ^{1}H NMR δ 1.23 (d, 12 Hz, J = 6.8 Hz), 3.67-3.77 (m, 2H), 5.14 (dd, 2H, J = 7.6, 16.4 Hz), 7.42 (t, 2H), 7.64 (t, 2H), 7.85 (d, 2H, J = 7.6Hz) and 8.08 (d, 2H, J = 7.6 Hz); mass spectrum (ESI) *m/z* 436.1638 $(M+H)^+$ ($C_{20}H_{26}N_3O_6P$ requires *m/z* 436.1638).

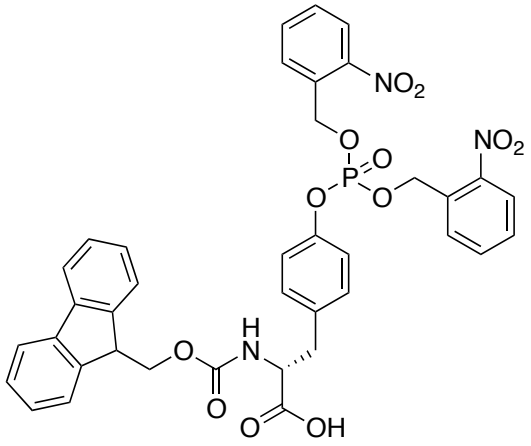


_O_-2-nitrobenzyl-_O_-t-butyl-_N,N_-diisopropylphsophoramidite (4.18)

To a 0 °C solution of PCl_3 (4.37 mL, 50 mmol) in 150 mL of diethylether was added diethylamine (34 mL, 300 mmol) and the mixture was stirred for 8 h. The reaction mixture was filtered and the filtrate was concentrated under diminished pressure. The residue was redissolved in 100 mL of THF and to the solution was added tetrazole (3.34 g 49 mmol) followed by *tert*-butanol (4.75 mL, 50 mmol). The mixture was allowed to stir at room temperature for 5 hours under argon. Nitrobenzyl alcohol (1 eq) was added

to the reaction and the mixture was left to stir overnight under argon. the solvent was concentrated under diminished pressure and the residue purified by silica gel flash chromatography (12 x 3 cm); elution with 97:3 hexane–triethylamine afforded the product was a yellow oil: yield 5.1g (31%); ^{1}H NMR δ 1.06-1.10 (m, 6H), 1.38 (s, 9H), 2.05-3.17 (m, 4H), 5.01-5.05 (m, 2H), 7.41 (t, 1H), 7.64 (t, 1H), 7.90 (d, 1H, J = 7.6 Hz) and 8.08 (d, 1H, J = 8 Hz); ^{13}C NMR δ 15.2, 31.0, 37.9, 61.6, 75.4, 124.7, 127.7, 129.0, 136.8 and 147.0; mass spectrum (MALDI) *m/z* 328.6 $(M+H)^{+}$ (theoretical *m/z* 328.4).

4.4.5. Synthesis of Fmoc-dinitrobenzyl-phosphorotyrosine



Fmoc-dintrobenzylphosphotyrsoine (4.19)

To a solution of 369 mg (5.26 mmol) of tetrazole and 868 mg (2.63) of *O*-dinitrobenzyl-*N*,*N*-diisopropyl phosphoramidite in 25 mL of ACN was added 805 mg (1.75 mmol) of Fmoc-Tyr-OtBu. Following stirring for 12 hours the reaction was concentrated and the crude residue was solubilized in 40 mL of ethyl acetate. The organic phase was washed with three 40 mL portions of sat. $NaHCO_3$, dried ($MgSO_4$) and concentrated. The crude

phosphite was dissolved in 5 mL of anhydrous ACN and treated with t-butyl hyrdoperoxide (1.05 mL,5 M solution in decane). Following consumption of the starting material the reaction mixture was concentrated and the residue was dissolved in 40 mL ethyl acetate. The organic phase was washed with three 40 mL portions of sat. $NaHCO_3$, dried ($MgSO_4$) and concentrated. The crude residue was solubilized 5 mL of a 40% TFA solution in anhydrous ACN and stirred until the starting material was consumed. The reaction was then concentrated and re-solubilized in 40 mL ethyl acetate. The organic solution was washed with three 40 mL portions of saturated $NaHCO_3$, dried ($MgSO_4$) and concentrated. The crude residue was purified by preparative HPLC 95-5 water/ACN to 60% acetonitrile over 40 minutes (t_R = 30 minutes); ^{1}H NMR δ 1.13-1.26 (m, 3H), 3.02-3.06 (m, 2H), 4.03-4.10 (m, 2H), 4.26-4.28 (m, 1H), 4.35-4.39 (m, 1H), 5.53-5.54 (m, 3H), 6.60 (d, 2H, J = 7.6 Hz), 6.84 (d, 2H, J = 7.6 Hz), 7.15-7.18 (m, 6h), 7.24-7.27 (m, 5H) and 7.36-7.44 (m, 7H) ^{13}C NMRδ 14.3, 20.8, 28.8, 37.1, 47.2, 54.6, 60.6, 120.1, 125.2, 127.2, 128.6, 129.2, 131.1, 131.7, 133.6, 134.3, 141.4, 143.8, 146.8, 149.2, 155.9, 174.2 and 176.6; mass spectrum (MALDI) *m/z* 753.9 $(M+H)^+$ (theoretical *m/z* 753.7).

REFERENCES

(1) Ramakrishnan, V. *Cell* **2002**, *108*, 557.

(2) Dedkova, L. M.; Fahmi, N. E.; Golovine, S. Y.; Hecht, S. M. *J. Am. Chem. Soc.* **2003**, *125*, 6616.

(3) Dedkova, L. M.; Fahmi, N. E.; Golovine, S. Y.; Hecht, S. M. *Biochemistry* **2006**, *45*, 15541.

(4) Dedkova, L. M.; Fahmi, N. E.; Paul, R.; del Rosario, M.; Zhang, L.; Chen, S.; Feder, G.; Hecht, S. M. *Biochemistry* **2012**, *51*, 401.

(5) Carter, C. W., Jr. *Annu. Rev. Biochem* **1993**, *62*, 715.

(6) Wang, B.; Zhou, J.; Lodder, M.; Anderson, R. D., 3rd; Hecht, S. M. *J. Biol. Chem.* **2006**, *281*, 13865.

(7) Schmeing, T. M.; Ramakrishnan, V. *Nature* **2009**, *461*, 1234.

(8) Goodman, H. M.; Abelson, J.; Landy, A.; Brenner, S.; Smith, J. D. *Nature* **1968**, *217*, 1019.

(9) Hecht, S. M.; Alford, B. L.; Kuroda, Y.; Kitano, S. *J. Biol. Chem.* **1978**, *253*, 4517.

(10) Heckler, T. G.; Chang, L. H.; Zama, Y.; Naka, T.; Chorghade, M. S.; Hecht, S. M. *Biochemistry* **1984**, *23*, 1468.

(11) Baldini, G.; Martoglio, B.; Schachenmann, A.; Zugliani, C.; Brunner, J. *Biochemistry* **1988**, *27*, 7951.

(12) Robertson, S. A.; Noren, C. J.; Anthony-Cahill, S. J.; Griffith, M. C.; Schultz, P. G. *Nuc. Acids Res* **1989**, *17*, 9649.

(13) Stepanov, V. G.; Moor, N. A.; Ankilova, V. N.; Vasil'eva, I. A.;

Sukhanova, M. V.; Lavrik, O. I. *Biochim. Biophys. Acta* **1998**, *1386*, 1.

(14) Stepanov, V. G.; Moor, N. A.; Ankilova, V. N.; Lavrik, O. I. *FEBS Lett.* **1992**, *311*, 192.

(15) Duca, M.; Trindle, C. O.; Hecht, S. M. *J. Am. Chem. Soc.* **2011**, *133*, 11368.

(16) Duca, M.; Chen, S.; Hecht, S. M. *Org. Biomol. Chem.* **2008**, *6*, 3292.

(17) Duca, M.; Chen, S.; Hecht, S. M. *Methods* **2008**, *44*, 87.

(18) Maloney, D. J.; Ghanem, N.; Zhou, J.; Hecht, S. M. *Org. Biomol. Chem.* **2007**, *5*, 3135.

(19) Hecht, S. M. *Acc. Chem. Res.* **1992**, *25*, 545.

(20) Karginov, V. A.; Mamaev, S. V.; Hecht, S. M. *Nucl. Acids Res.* **1997**, *25*, 3912.

(21) Mamaev, S. V.; Laikhter, A. L.; Arslan, T.; Hecht, S. M. *J. Am. Chem. Soc.* **1996**, *118*, 7243.

(22) Karginov, V. A.; Mamaev, S. V.; An, H. Y.; VanCleve, M. D.; Hecht, S. M.; Komatsoulis, G. A.; Abelson, J. N. *J. Am. Chem. Soc.* **1997**, *119*, 8166.

(23) Chen, S. X.; Hecht, S. M. *Bioorg. Med. Chem.* **2008**, *16*, 9023.

(24) Duca, M.; Maloney, D. J.; Lodder, M.; Wang, B.; Hecht, S. M. *Bioorg. Med. Chem.* **2007**, *15*, 4629.

(25) Lodder, M.; Wang, B. X.; Hecht, S. M. *Methods* **2005**, *36*, 245.

(26) Choudhury, A. K.; Golovine, S. Y.; Dedkova, L. M.; Hecht, S. M.

Biochemistry **2007**, *46*, 4066.

(27) Nangreave, R. C.; Dedkova, L. M.; Chen, S.; Hecht, S. M. *Org. Lett.* **2011**, *13*, 4906.

(28) Kanavarioti, A.; Lu, J.; Rosenbach, M. T.; Hurley, T. B. *Tet. Lett.* **1991**, *32*, 6065.

(29) Mukaiyama, T.; Hashimoto, M. *J. Am. Chem. Soc.* **1972**, *94*, 8528.

(30) Maini, R.; Nguyen, D. T.; Chen, S.; Dedkova, L. M.; Chowdhury, S. R.; Alcala-Torano, R.; Hecht, S. M. *Bioorg. Med. Chem.* **2013**, *21*, 1088.

(31) Noren, C. J.; Anthony-Cahill, S. J.; Griffith, M. C.; Schultz, P. G. *Science* **1989**, *244*, 182.

(32) Robertson, S. A.; Ellman, J. A.; Schultz, P. G. *J. Am. Chem. Soc.* **1991**, *113*, 2722.

(33) Sawano, A.; Miyawaki, A. *Nucl. Acids Res.* **2000**, *28*, E78.

(34) Murakami, H.; Ohta, A.; Goto, Y.; Sako, Y.; Suga, H. *Nucleic Acids Symp. Ser.* **2006**, 35.

(35) Murakami, H.; Suga, H. *Tanpakushitsu kakusan koso. Protein, Nucleic Acid, Enzyme* **2006**, *51*, 2496.

(36) Niwa, N.; Yamagishi, Y.; Murakami, H.; Suga, H. *Bioorg. Med. Chem. Lett.* **2009**, *19*, 3892.

(37) Ohuchi, M.; Murakami, H.; Suga, H. *Curr. Opin. Chem. Biol.* **2007**, *11*, 537.

(38) Passioura, T.; Suga, H. *Chemistry* **2013**, *19*, 6530.

(39) Ramaswamy, K.; Saito, H.; Murakami, H.; Shiba, K.; Suga, H. *J. Am.*

Chem. Soc. **2004**, *126*, 11454.

(40) Sako, Y.; Murakami, H.; Suga, H. *Tanpakushitsu kakusan koso. Protein, Nucleic Acid, Enzyme* **2007**, *52*, 1643.

(41) Hashimoto, N.; Ninomiya, K.; Endo, T.; Sisido, M. *Chem. Comm.* **2005**, 4321.

(42) Duffy, N. H.; Dougherty, D. A. *Org. Lett.* **2010**, *12*, 3776.

(43) Ninomiya, K.; Kurita, T.; Hohsaka, T.; Sisido, M. *Chem. Comm.* **2003**, 2242.

(44) Collier, A.; Wagner, G. K. *Chem. Comm.* **2008**, 178.

(45) Lodder, M.; Golovine, S.; Hecht, S. M. *J. Org. Chem.* **1997**, *62*, 778.

(46) Griffin, B. E.; Reese, C. B.; Stephenson, G. F.; Trentham, D. R. *Tet. Lett.* **1966**, *36*, 4349.

(47) Hoffmann, P. U.; McLaughlin, L. W. *Nucleic Acids Res.* **1987**, *15*, 5289.

(48) Sardina, F. J.; Rapoport, H. *Chem. Rev.* **1996**, *96*, 1825.

(49) Fulop, F. *Chem. Rev.* **2001**, *101*, 2181.

(50) Beck, G. *Synlett* **2002**, 837.

(51) Najera, C. *Synlett* **2002**, 1388.

(52) Ager, D. J.; Li, T.; Pantaleone, D. P.; Senkpeil, R. F.; Taylor, P. P.; Fotheringham, I. G. *J. Mol. Catal. B: Enzym.* **2001**, *11*, 199.

(53) Arslan, T.; Mamaev, S. V.; Mamaeva, N. V.; Hecht, S. M. *J. Am. Chem.*

Soc. **1997**, *119*, 10877.

(54) Killian, J. A.; Van Cleve, M. D.; Shayo, Y. F.; Hecht, S. M. *J. Am. Chem. Soc*. **1998**, *120*, 3032.

(55) Short, G. F.; Lodder, M.; Laikhter, A. L.; Arslan, T.; Hecht, S. M. *J. Am. Chem. Soc*. **1999**, *121*, 478.

(56) Short, G. F.; Laikhter, A. L.; Lodder, M.; Shayo, Y.; Arslan, T.; Hecht, S. M. *Biochemistry* **2000**, *39*, 8768.

(57) Baird, T.; Wang, B. X.; Lodder, M.; Hecht, S. M.; Craik, C. S. *Tetrahedron* **2000**, *56*, 9477.

(58) Wang, B. X.; Brown, K. C.; Lodder, M.; Craik, C. S.; Hecht, S. M. *Biochemistry* **2002**, *41*, 2805.

(59) Anderson, R. D.; Zhou, J.; Hecht, S. M. *J. Am. Chem. Soc*. **2002**, *124*, 9674.

(60) Gao, R.; Zhang, Y.; Choudhury, A. K.; Dedkova, L. M.; Hecht, S. M. *J. Am. Chem. Soc*. **2005**, *127*, 3321.

(61) Gao, R.; Zhang, Y.; Dedkova, L.; Choudhury, A. K.; Rahier, N. J.; Hecht, S. M. *Biochemistry* **2006**, *45*, 8402.

(62) Yakovleva, L.; Chen, S. X.; Hecht, S. M.; Shuman, S. *J. Biol. Chem*. **2008**, *283*, 16093.

(63) Chen, S. X.; Zhang, Y.; Hecht, S. M. *Biochemistry* **2011**, *50*, 9340.

(64) Katritzky, A. R.; Xu, Y. J.; Vakulenko, A. V.; Wilcox, A. L.; Bley, K. R. *J. Org. Chem*. **2003**, *68*, 9100.

(65) Varshney, U.; Lee, C. P.; Rajbhandary, U. L. *J. Biol. Chem*. **1991**, *266*, 24712.

(66) Karginov, V. A.; Mamaev, S. V.; An, H.; Van Cleve, M. D.; Hecht, S. M.; Komatsoulis, G. A.; Abelson, J. N. *J. Am. Chem. Soc.* **1997**, *119*, 8166.

(67) Connolly, B. A.; Rider, P. *Nuc. Acids Res.* **1985**, *13*, 4485.

(68) Fischer, E. H.; Krebs, E. G. *Fed. Proc.* **1966**, *25*, 1511.

(69) Hunter, T. *Cell* **2000**, *100*, 113.

(70) Pawson, T.; Nash, P. *Gene Dev.* **2000**, *14*, 1027.

(71) Manning, G.; Whyte, D. B.; Martinez, R.; Hunter, T.; Sudarsanam, S. *Science* **2002**, *298*, 1912.

(72) Ubersax, J. A.; Ferrell, J. E., Jr. *Nature reviews. Molecular cell biology* **2007**, *8*, 530.

(73) Ciesla, J.; Fraczyk, T.; Rode, W. *Acta Biochim. Pol.* **2011**, *58*, 137.

(74) Hanks, S. K.; Hunter, T. *FASEB J.* **1995**, *9*, 576.

(75) Pereira, S. R.; Vasconcelos, V. M.; Antunes, A. *Crit. Rev. Toxicol.* **2011**, *41*, 83.

(76) Alonso, A. D.; Grundke-Iqbal, I.; Barra, H. S.; Iqbal, K. *PNAS* **1997**, *94*, 298.

(77) Motevalli, M.; Goldschmidt-Clermont, P. J.; Virgil, D.; Kwiterovich, P. O., Jr. *J. Biol. Chem.***1997**, *272*, 24703.

(78) Cohen, P. *Eur. J. Biochem.* **2001**, *268*, 5001.

(79) Zhu, H.; Klemic, J. F.; Chang, S.; Bertone, P.; Casamayor, A.; Klemic, K. G.; Smith, D.; Gerstein, M.; Reed, M. A.; Snyder, M. *Nature Gen.* **2000**, *26*, 283.

(80) Caenepeel, S.; Charydczak, G.; Sudarsanam, S.; Hunter, T.; Manning, G. *PNAS* **2004**, *101*, 11707.

(81) Bishop, J. S.; Larner, J. *J. Biol. Chem.* **1967**, *242*, 1354.

(82) de Wulf, H.; Hers, H. G. *Eur. J. Biochem.* **1968**, *6*, 545.

(83) Fischer, E. H.; Krebs, E. G. *J. Biol. Chem.* **1955**, *216*, 121.

(84) Krebs, E. G.; Graves, D. J.; Fischer, E. H. *J. Biol. Chem.* **1959**, *234*, 2867.

(85) Fischer, E. H. *Biol. Chem.* **2010**, *391*, 131.

(86) Maxwell, S. A.; Sacks, P. G.; Gutterman, J. U.; Gallick, G. E. *Cancer Res.* **1989**, *49*, 1130.

(87) Tiganis, T. *IUBMB life* **2002**, *53*, 3.

(88) Shi, C.; Ma, Y.; Liu, H.; Zhang, Y.; Wang, Z.; Jia, H. *Mol. Cellular Biochem.* **2012**, *362*, 65.

(89) Tarrant, M. K.; Cole, P. A. *Ann. Rev. Biochem.* **2009**, *78*, 797.

(90) Ellis-Davies, G. C. *Cold Spring Harbor protocols* **2009**, *2009*.

(91) Kaplan, J. H.; Forbush, B., 3rd; Hoffman, J. F. *Biochemistry* **1978**, *17*, 1929.

(92) Kovarova, H.; Hajduch, M.; Livingstone, M.; Dzubak, P.; Lefkovits, I. *Journal of Chromatography. B, Analytical technologies in the biomedical and life sciences* **2003**, *787*, 53.

(93) Deng, Z.; Bu, S.; Wang, Z. Y. *Methods Mol. Biol.* **2012**, *876*, 47.

(94) Mann, M.; Ong, S. E.; Gronborg, M.; Steen, H.; Jensen, O. N.; Pandey, A.

Trends in Biotechnology **2002**, *20*, 261.

(95) Rothman, D. M.; Petersson, E. J.; Vazquez, M. E.; Brandt, G. S.; Dougherty, D. A.; Imperiali, B. *J. Am. Chem. Soc.* **2005**, *127*, 846.

(96) Huynh, M. L.; Russell, P.; Walsh, B. *Methods Mol. Biol.* **2009**, *519*, 507.