Study of Ribosomes having Modifications in the Peptidyltransferase Center Using Non-α-L-Amino Acids and Synthesis and Biological Evaluation of Topopyrones

by

Rumit Maini

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

Approved July 2013 by the
Graduate Supervisory Committee:

Sidney M. Hecht, Chair
Ian Gould
Hao Yan

ARIZONA STATE UNIVERSITY

August 2013

ABSTRACT

The ribosome is a ribozyme and central to the biosynthesis of proteins in all organisms. It has a strong bias against non-α-L-amino acids, such as α-D-amino acids and β-amino acids. Additionally, the ribosome is only able to incorporate one amino acid in response to one codon. It has been demonstrated that reengineering of the peptidyltransferase center (PTC) of the ribosome enabled the incorporation of both α-D-amino acids and β-amino acids into full length protein.

Described in Chapter 2 are five modified ribosomes having modifications in the peptidyltrasnferase center in the 23S rRNA. These modified ribosomes successfully incorporated five different β-amino acids (**2.1** – **2.5**) into *E. coli* dihydrofolate reductase (DHFR). The second project (Chapter 3) focused on the study of the modified ribosomes facilitating the incorporation of the dipeptide glycylphenylalanine (**3.25**) and fluorescent dipeptidomimetic **3.26** into DHFR. These ribosomes also had modifications in the peptidyltransferase center in the 23S rRNA of the 50S ribosomal subunit. The modified DHFRs having β-amino acids **2.3** and **2.5**, dipeptide glycylphenylalanine (**3.25**) and dipeptidomimetic **3.26** were successfully characterized by the MALDI-MS analysis of the peptide fragments produced by "in-gel" trypsin digestion of the modified proteins. The fluorescent spectra of the dipeptidomimetic **3.26** and modified DHFR having fluorescent dipeptidomimetic **3.26** were also measured.

The type I and II DNA topoisomerases have been firmly established as effective molecular targets for many antitumor drugs. A "classical" topoisomerase I or II poison acts by misaligning the free hydroxyl group of the sugar moiety of DNA and preventing the reverse transesterfication reaction to religate DNA. There have been only two classes

of compounds, saintopin and topopyrones, reported as dual topoisomerase I and II poisons. Chapter 4 describes the synthesis and biological evaluation of topopyrones. Compound **4.10**, employed at 20 μM, was as efficient as 0.5 μM camptothecin, a potent topoisomerase I poison, in stabilizing the covalent binary complex (~30%). When compared with a known topoisomerase II poison, etoposide (at 0.5 uM), topopyorone **4.10** produced similar levels of stabilized DNA–enzyme binary complex (~34%) at 5 μM concentration.

ACKNOWLEDGEMENTS

I would like to extend my sincere gratitude to my research advisor Professor Sidney M. Hecht who has been an outstanding mentor over the last five years. I am extremely grateful to him for providing me with an opportunity to work and grow in a stimulating research environment. I would also like to thank him for allowing me to explore opportunities in the field of protein biochemistry along with synthetic organic chemistry. I must thank Professors Seth Rose and Ian Gould for their guidance throughout the years and help with the complexities of graduate school. I shall also acknowledge and thank Professor Hao Yan for accepting my request to be on my committee on such short notice.

I must extend my sincere gratitude to Dr. Larisa Dedkova for being an amazing mentor, colleague and friend. I am grateful to her for training me as a proficient biochemist. Dr. Simon J. Leiris was my first mentor in my graduate research and I am thankful to him for imparting strong synthetic organic chemistry skills to me. I am grateful to Dr. Shengxi Chen, Dr. Pablo Arce, Dr. Manikandidas M.M., and Dr. Rakesh Paul for their continuous input regarding many complicated issues related to my research. Thanks to Dr. Paul A. Zaleski for his work on the biological study of topoyrone derivatives.

I extend my gratitude to Dr. Ryan Nangreave for being a great lab mate and an amazing friend; research would not have been as much fun without you. I would also like to thank two wonderful ladies, Dr. Jeanette Nangreave and Emma Nangreave for their constant support over the years. I thank my dear friend, Trevor Bozeman, for his continual assistance in and outside the lab and for introducing me to various aspects of

the American culture. I would especially like to thank my dearest friend Caridad Rodriguez, who has been a pillar in my life here in the US. I have learned a lot of from her and I could not have done this without her constant support. I would like to thank Basab Roy for giving me rides to lab on numerous occasions. My thanks also go to all the current and former members of the Sidney Hecht group.

I would finally like to thank my parents and my sister for their unconditional love and support. I am grateful for their patience and guidance and letting me follow my dreams. I would not have achieved this without their encouragement. I would finally like to dedicate this work to my *nani* (grandmother) and *nana* (late grandfather) who have always showered me with their unconditional love.

TABLE OF CONTENTS

Page

LIST OF ABBREVIATIONS

aaRS	aminoacyl-tRNA synthetase
AMP	adenosine-5'-monophosphate
ATP	adenosine-5'-triphosphate
aq	aqueous
bs	broad singlet
BSA	bovine serum albumin
^{13}C NMR	carbon-13 nuclear magnetic resonance spectroscopy
^{o}C	degrees Celsius
cat	catalytic
cm	centimeter
δ	chemical shift (ppm)
d	doublet
dec	decomposition
DEAE	diethylaminoethyl
DHF	7,8-dihydrofolate
DHFR	dihydrofolate reductase
DIPEA	diisopropylethylamine
DMAP	dimethylaminopyridine
DMF	dimethylformamide
DMSO	dimethylsulfoxide
DNA	deoxyribonucleic acid
DTT	dithiothreitol

E. coli	*Escherichia coli*
EDTA	ethylenedinitrilotetraacetic acid
ESI	electrospray ionization
GDP	guanosine-5'-diphosphate
GTP	guanosine-5'-triphosphate
^{1}H NMR	proton nuclear magnetic resonance spectroscopy
HBTU	*O*-benzotriazole-*N*,*N*,*N*′,*N*′-tetramethyl-uronium-hexafluoro-phosphate
HPLC	high performance liquid chromatography
Hz	Hertz
IPTG	isopropyl-β-D-thiogalactopyranoside
J	coupling constant
L	liter
M	molar
m	multiplet
MALDI-MS	matrix assisted laser desorption ionization – mass spectrometry
MHz	mega Hertz
mL	milliliter
mM	millimolar
mmol	millimole(s)
mRNA	messenger ribonucleic acid
uM	micromolar
μmol	micromole(s)
nm	nanometer

NADPH	nicotinamide adenine dinucleotide phosphate
NMR	nuclear magnetic resonance
pdCpA	5′-*O*-phosphoryl-2′-deoxycytidyl(3′→5′)adenosine
PTC	peptidyltransferase center
p-TsCl	*p*-toluenesulfonyl chloride
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
t	triplet
TBA	tetrabutylammonium
$TBAF{\cdot}H_2O$	tetrabutylammoniumfluoride monohydrate
TFA	trifluoroacetic acid
THF	5,6,7,8-tetrahydrofolate
TLC	thin layer chromatography
Topo	topoisomerase
tRNA	transfer RNA
UV	ultraviolet
Val	valine

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

Figure Page

LIST OF SCHEMES

Chapter 1

THE RIBOSOME AND PROTEIN SYNTHESIS: GENERAL INTRODUCTION

1.1. The ribosome

Proteins are one of the most versatile classes of polymers that nature synthesizes to perform important biological functions. Scientists have long been intrigued by the structural and functional diversity of proteins. Proteins are comprised of different amino acid monomers, which are linked by peptide bonds. The cellular machinery which synthesizes proteins, one amino acid at a time, is called the ribosome.[1]

The ribosome is made up of a complex array of ribosomal RNAs (rRNAs) and proteins. It synthesizes proteins with high speed and excellent fidelity using messenger RNAs (mRNAs) and transfer RNAs (tRNAs).[2] The bacterial ribosome is comprised of 50S (large) and 30S (small) ribosomal subunits. The large subunit consists of 23S RNA, 5S RNA and approximately 30 proteins, whereas 16S RNA and 20 proteins together form the 30S subunit.[1] Each ribosomal subunit has three sites for tRNA binding, designated as the A-site (site for aminoacyl-tRNA), P-site (site for peptidyl-tRNA) and E-site (site for deacylated tRNA set to exit the ribosome) (Figure 1.1).[3] In bacteria, the 30S ribosomal subunit binds to the mRNA upstream from the start codon, and the 16S ribosomal RNA base pairs with its complementary Shine–Dalgarno sequence in the bound mRNA.[1] In addition, the 30S ribosomal subunit binds to the anticodon stem-loop of tRNA and contributes to the fidelity of protein synthesis by monitoring codon-anticodon base pairing during the decoding process.[3] The 50S ribosomal subunit is responsible for catalyzing peptide bond formation.[5] The α-amino moiety of the aminoacyl group of

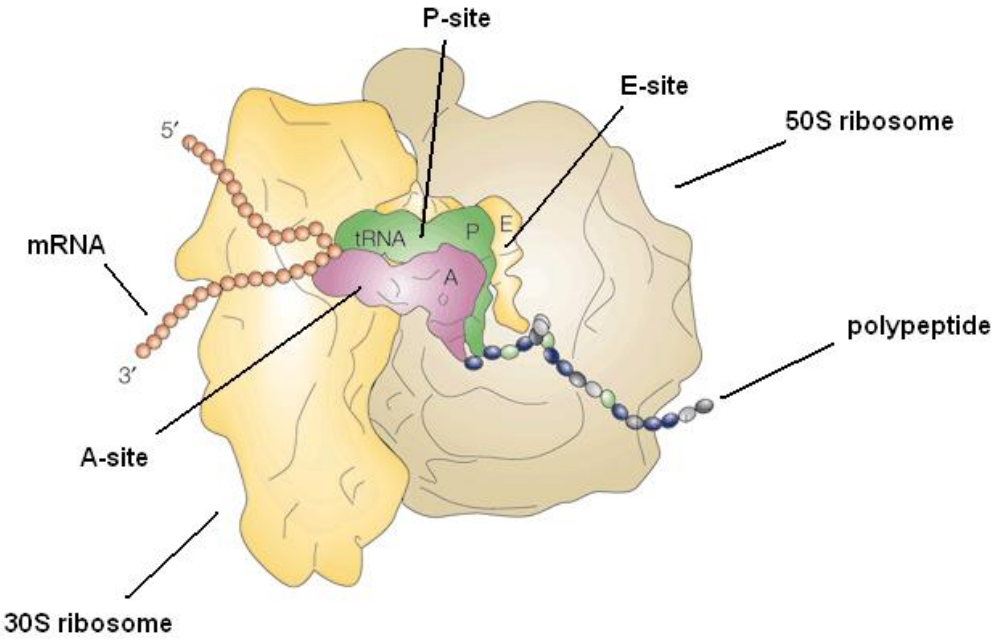
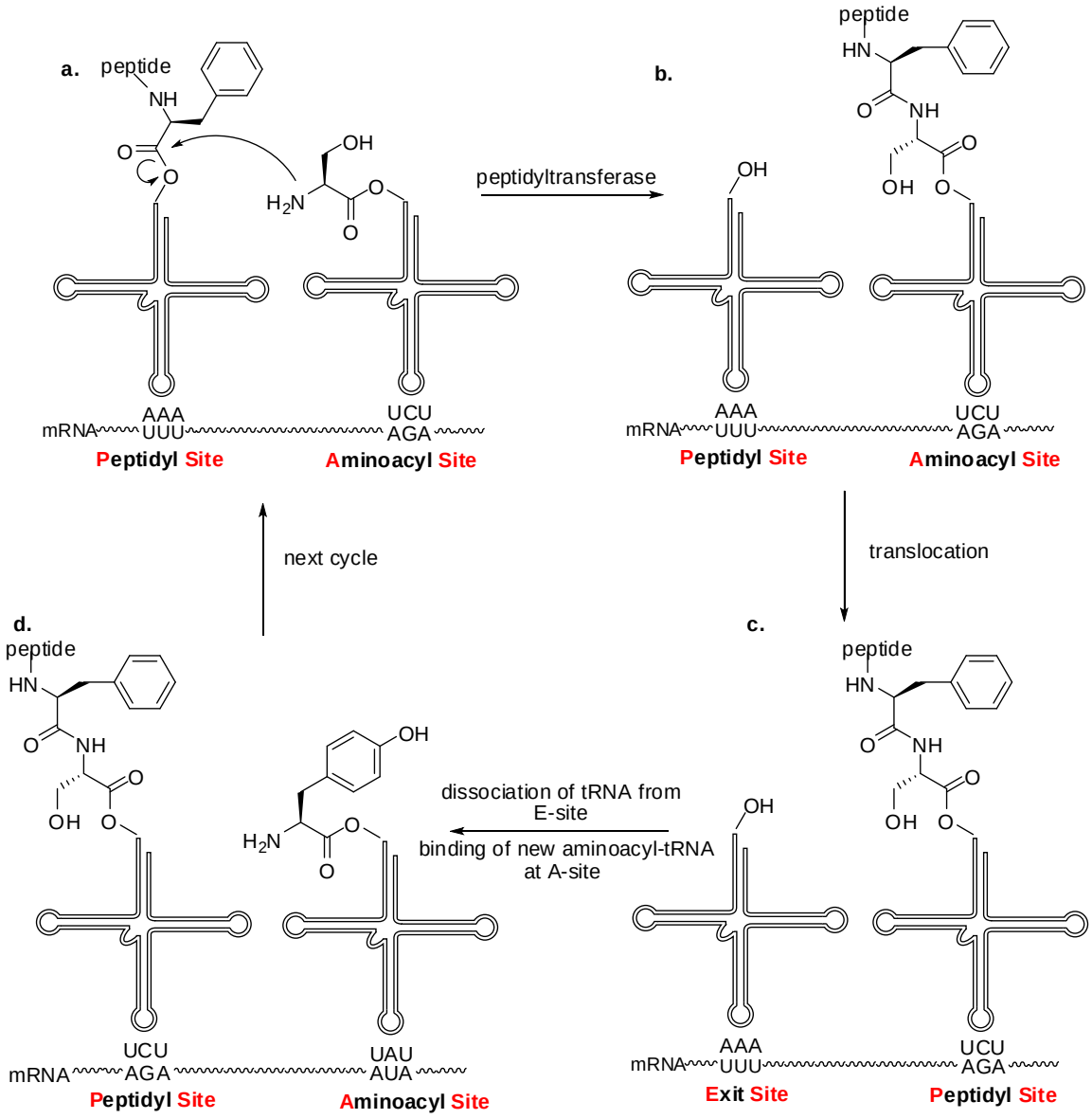


Figure 1.1. Schematic diagram of a ribosome bound to an mRNA showing the A, P and E-sites.[4]

tRNA, located in the A-site, mediates a nucleophilic attack on the carbonyl carbon of the ester bond connecting the nascent peptide to the peptidyl-tRNA (Scheme 1.1a). As a consequence, the growing polypeptide is transferred onto the tRNA occupying the A-site, adding a new amino acid to the growing chain (Scheme 1.1b). This event triggers the hydrolysis of GTP, followed by translocation of the mRNA relative to the ribosome. The peptidyl-tRNA in the A-site and the deacylated tRNA in the P-site translocate to the P-site and E-site, respectively (Scheme 1.1c). Following this event a new cycle begins, in which the next aminoacyl-tRNA binds to mRNA in the A-site and the deacylated tRNA dissociates from the E-site (Scheme 1.1d).

Scheme 1.1. Formation of a peptide bond in the PTC.[1]



During mRNA translation, the overall error rates range from 6×10^{-4} to 5×10^{-3} which is a result of the net accumulation of errors in several steps, mainly transcription (10^{-4}), aminoacyl-tRNA synthesis (10^{-4}) and ribosomal decoding (10^{-4}).[6,7] The aminoacyl-tRNA synthetase (aaRS) activates an amino acid with ATP to form an aminoacyladenylate, followed by the transfer of the activated amino acid to its cognate

tRNA to generate an aminoacyl-tRNA.[1] Each tRNA is a specific substrate for its cognate synthetase; the tRNA provides a large contact surface area during recognition, and many aminoacyl-tRNA synthetases have a unique proofreading mechanisms to assure the fidelity of aminoacylation.[1] The intrinsic proofreading mechanism of an aminoacyl-tRNA is termed "editing" and was first described by Baldwin and Berg in 1966.[8] They reported that the misacylated Val-AMP was hydrolyzed by IleRS upon addition of $tRNA^{Ile}$. The editing can take place either after aminoacyl-adenylate (pretransfer editing) and/or aminoacyl-tRNA (posttransfer editing) formation by an aaRS. The post transfer editing in an aaRS was first reported in two separate studies in 1972. Eldred and Schimmel[9] found that IleRS catalyzed the hydrolysis of misacylated Val-$tRNA^{Ile}$, whereas Yarus[10] reported the hydrolysis of misacylated Ile-$tRNA^{Phe}$ by PheRS. Several studies have demonstrated that approximately half of all aaRSs utilize editing activity to maintain translational fidelity.[7]

The ribosome plays a key role in maintaining the translational fidelity by its decoding process. Following a correct match between codon (in mRNA) and anticodon (in tRNA) triplets, the ribosome selects the correct aminoacyl-tRNA as a ternary complex with elongation factor Tu (EF-Tu) and GTP. When the appropriate ternary complex is accommodated in the A-site of the ribosome, GTP hydrolysis as part of a complex with EF-Tu releases the aminoacyl-tRNA.[1] The ternary complex of the aminoacyl-tRNA with EF-Tu–GTP forms a reversible initial binding complex with the ribosome. If there is not a perfect match between the anticodon and codon, the ternary complex dissociates from the ribosome rapidly. A cognate codon-anticodon recognition stabilizes the tRNA–EF-Tu–GTP–ribosome complex which triggers a conformational rearrangement in EF-Tu.

This change in conformation of EF-Tu is accompanied by the hydrolysis of GTP to GDP; the GDP bound EF-Tu loses its affinity for the aminoacyl-tRNA. After the release of the aminoacyl-tRNA by EF-Tu, the 3'-end of the aminoacyl-tRNA moves into the PTC, where it participates in peptide bond formation.[11]

1.2. The ribosome is a ribozyme

The catalytic center of the ribosome resides in the 23S RNA of the 50S ribosomal subunit and is called the peptidyltransferase center (PTC).[3] Enhancement of the rate of peptide bond formation by the ribosome is at least ~10^5-fold as compared to uncatalyzed peptide bond formation between two amino acids.[12] Crick, in 1968, had proposed that the 23S rRNA might contain the catalytic activity of the ribosome and thus introduced the idea of a ribosome being a ribozyme.[13] A ribozyme (**ribo**nucleic acid en**zyme**, also called RNA enzyme or catalytic RNA) is an RNA molecule that catalyzes a chemical reaction in a fashion similar to protein enzymes. Many natural ribozymes catalyze the hydrolysis of one of their own phosphodiester bonds, or those in other RNAs.[14] The first ribozymes were discovered by Sidney Altman[15] and Thomas Cech[16] in separate studies; in 1989, both of them received the Nobel prize in chemistry for this work. Following the discovery of ribozymes, Crick's hypothesis began to gain interest. Noller and co-workers prepared large ribosomal subunit particles after increasingly vigorous deproteinization and showed that the particles retained peptidyltransferase activity.[17] In a separate study, Watanabe and co-workers reported that the *in vitro* synthesized protein free 23S rRNA also demonstrated peptidyltransferase activity.[18] In 2000, Steitz published the 2.3 Å atomic resolution structures of a 50S ribosomal subunit from the archeon *H. marismortui* and its complexes with two aminoacyl-tRNA mimics.[19] In this study, it was established

that there are no ribosomal protein side chain atoms in the vicinity of the peptide bond being formed. They proposed that N3 of A2486 (A2451 in *E. coli*) participates in general acid/base catalysis in the peptidyltransferase reaction.[20] The availability of X-ray crystal structures with improved resolution indicated that N3 of A2451 is not near enough to the α-amino group of the amino acid to enable hydrogen bonding with this functionality.[21] Substitution of the α-amino group of the aminoacyl moiety of the aminoacyl-tRNA with a less reactive hydroxyl group resulted in a reaction rate of depsipeptide bond formation which was pH independent.[22] Subsequently, it was proposed that the 2'-OH group of the sugar moiety of A76 of the peptidyl-tRNA mediates a concerted proton shuttling during peptide bond formation.[2,23] Weinger et al. substituted the 2'-OH group of A76 with H or F and observed a ~10^6 fold reduction in the rate of the peptidyltransferase reaction.[23] In contrast, Sprinzl et al.[24] demonstrated that a peptidyl-tRNA having a 2'-deoxy sugar moiety in A_{76} does not make the peptidyltransferase center inactive; instead the peptidyltransferase reaction still proceeds at a reasonable rate. Therefore, this study suggested that the 2'-OH group of the 3'-adenosine residue of peptidyl-tRNA may not be required for the proton shuttling.

It is now established that the ribosome is a ribozyme which catalyzes the peptidyltransferase reaction, but the role of N3 of A2451and 2'-OH group of the 3'-adenosine residue of peptidyl-tRNA are still not clear.

1.3. Biosynthesis of modified proteins

The study of protein structure and function has been a central focus for the last few decades. Engineering proteins to improve function or create novel functionalities has applications in protein therapeutics,[25] bio-imaging,[26] protein folding and self-assembly,[27]

and organic synthesis.[28] The incorporation of non-canonical amino acids into proteins represents an important tool for protein engineering.[29] The first report of the incorporation of a non-canonical amino acid into protein was published by Cohen and Cowie[30] in 1957, wherein complete substitution of methionine with selenomethionine in *E. coli* auxotrophic for methionine was achieved. Over the past decade, selenomethionine incorporation into proteins has played an important role in protein crystallography.[31]

The major strategy for the site-specific incorporation of unnatural amino acids into proteins has been via the suppression of a nonsense codon in mRNA with a tRNA bearing an unnatural amino acid (Figure 1.2). Hecht and co-workers developed a key methodology for preparing misacylated tRNAs, demonstrating that the chemically synthesized aminoacylated pCpA could be ligated to an abbreviated tRNA using T4 RNA ligase.[32,33] Schultz and co-workers utilized tRNAs bearing unnatural amino acids for the *in vitro* translation of full length proteins.[34] To produce proteins bearing unnatural amino acids at specific sites *in vivo*, orthogonal tRNA/tRNA synthetase pairs for unnatural amino acids were later developed by the Schultz laboratory.[35] For instance, a suppressor $tRNA^{Tyr}$ and a cognate tyrosyl-tRNA synthetase, derived from the archeal bacterium *M. jannasschii*, was modified and introduced into *E. coli* to mediate *in vivo* incorporation of *O*-methyl-L-tyrosine into dihydrofolate reductase (DHFR) in response to a UAG nonsense codon.[35] Such technological advances over the past three decades have revolutionized protein engineering and provided access to proteins with novel functions.[29]

1.4 Dihydrofolate redutase (DHFR)

Dihydrofolate reductase (DHFR) is an essential and ubiquitous enzyme required for normal cellular metabolism in both prokaryotes and eukaryotes. The primary role of DHFR is maintenance of the intracellular pools of 5,6,7,8-tetrahydrofolate (THF) and its derivatives, which are required for the biosynthesis of purines, pyrimidines and several

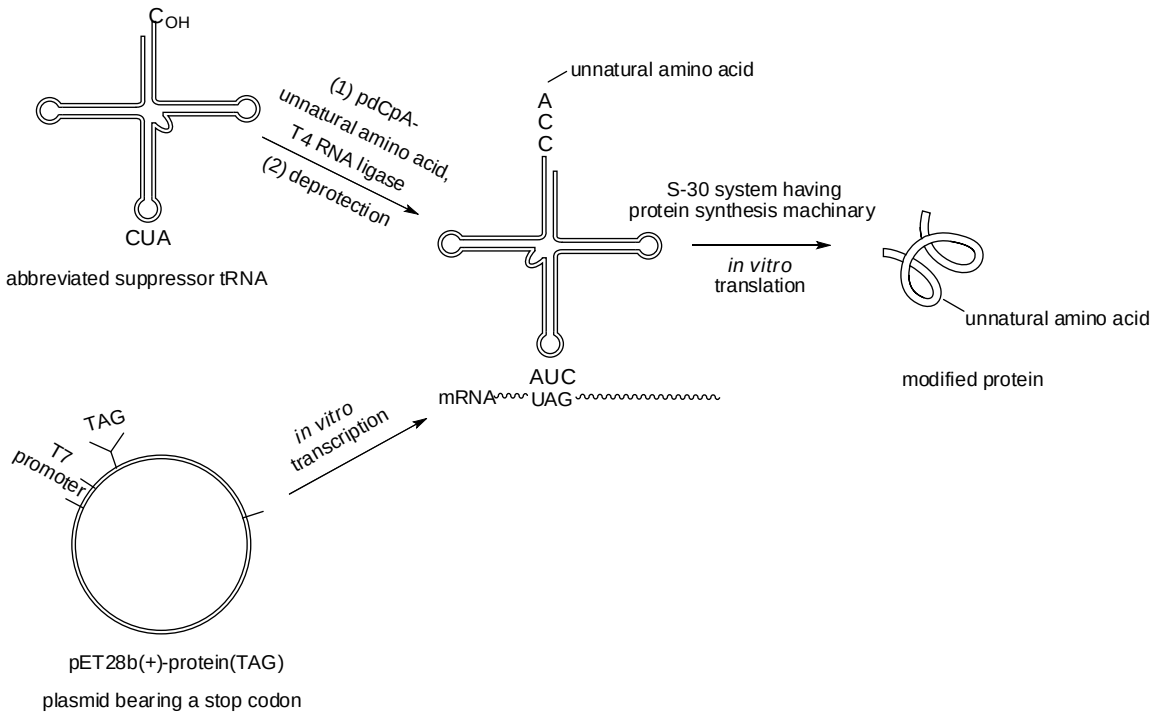
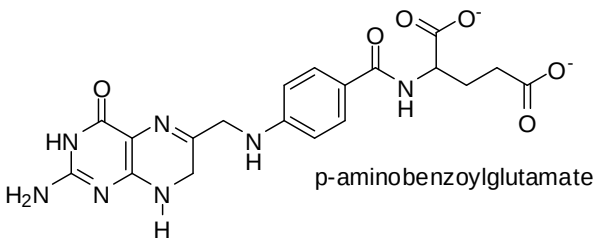


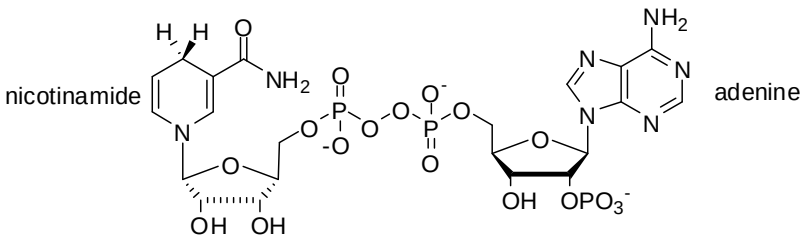
Figure 1.2. Strategy for the incorporation of unnatural amino acids into proteins *in vitro*.[34]

amino acids. The enzyme catalyzes the reduction of 7,8-dihydrofolate (DHF) to THF by transferring the pro-*R* hydride of nicotinamide adenine dinucleotide phosphate (NADPH) cofactor stereospecifically to the C6 atom of the pterin nucleus with simultaneous protonation at N5 (Scheme 1.2).[36,37] DHFR is the only source of THF in both prokaryotes

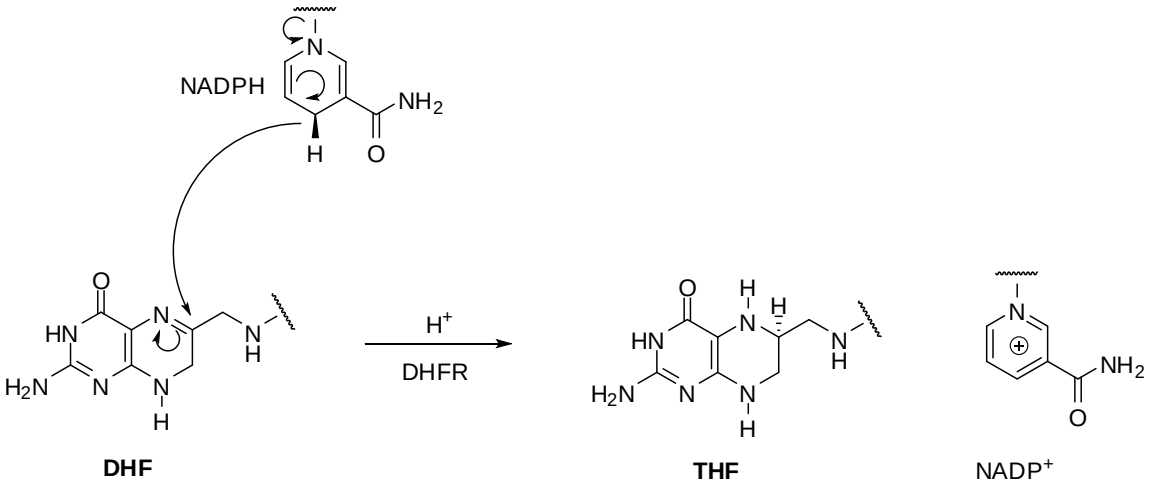
Scheme 1.2. DHFR catalyzed formation of 5,6,7,8-tetrahydrofolate (THF) from 7,8-dihydrofolate (DHF) using nicotinamide adenine dinucleotide phosphate (NADPH) as a cofactor.[37]



DHF



NADPH



DHF **THF** $NADP^+$

and eukaryotes. Several anticancer and antibiotic drugs target DHFR to block the conversion of DHF to THF. Methotrexate is one of the most potent chemotherapeutic drugs and DHFR is the major target of this drug.[38] Another compound, trimethoprim, is an important antibacterial agent as it binds to bacterial DHFR 10^5 times more strongly than to vertebrate DHFRs.[39]

DHFRs from different organisms share low sequence homology but high structural homology. The first crystal structure of a DHFR enzyme was published in 1978.[40] *E. coli* DHFR is an 18 kD protein consisting of eight β-strands and four flanking α-helices. It is divided into two structural subdomains: the adenosine binding subdomain and the major subdomain. In *E. coli* DHFR, the adenosine moiety binds to the smaller adenosine binding domain which is formed by the amino acid residues 38-88. Close to 100 amino acid residues from both the N and C termini make up the major subdomain. This subdomain is dominated by three loops around the active site, which comprise 40% to 50% of the major domain. The loops are defined as 'Met20' (residues 9-24), 'F-G' (residues 116-132) and 'G-H' (residues 142-150). The Met20 loop is present directly above the active site, protecting it from solvent. The F-G and G-H loops play important roles in stabilizing the Met20 loop via a network of hydrogen bonding interactions.

Benkovic and coworkers have reported the complete catalytic scheme for *E. coli* DHFR (Figure 1.3).[41] The most intriguing aspect of the DHFR catalytic mechanism is that after the hydride transfer, product release does not follow directly. Instead, $NADP^+$ is first released followed by rebinding of NADPH, which is the rate-determining step. After binding of NADPH, the product is released, leaving the enzyme primed for the next round of catalysis. There is a synergistic interaction between the substrate and cofactor

binding sites, where the rate of release of $NADP^+$ is increased in the presence of the bound product and the rate of the release of the product is elevated in presence of bound NADPH.

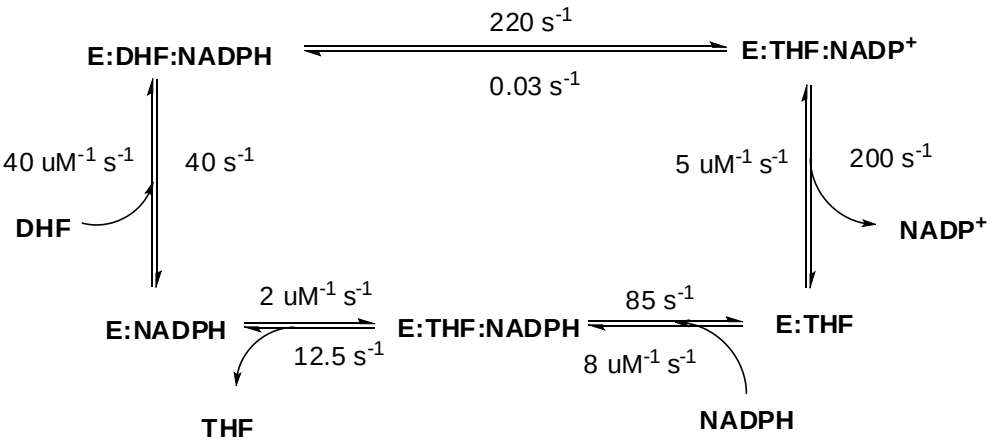


Figure 1.3. The catalytic cycle of *E. coli* DHFR at pH 7.[41]

1.5. Modified ribosomes

Studying ribosome structure and function by mutagenesis at critical nucleotide residues has been carried out for some time, yet studies of ribosome modifications to enable incorporation of amino acids other than α-L-amino acids are non-existent.[2,42] The structural recognition of an aminoacyl-tRNA by the ribosome is crucial for the incorporation of unnatural amino acids into proteins.[43] Puromycin (**1.1**) is an aminonucleoside antibiotic (Figure 1.4), whose structure resembles the 3'-end of the aminoacyl-tRNA, rendering it a putative aminoacyl-tRNA mimic.[44,45] It binds to the A-site of the PTC and takes part in peptide bond formation with the nascent polypeptide attached to the peptidyl-tRNA in the P-site, which leads to premature chain release.[44,45] Puromycin and its derivatives having different amino acid side chains show different

ribosomal inhibition efficiencies.[46,47] This suggests that the ribosomal A-site shows specificity towards binding the different side chains of amino acids, at least at the level of puromycin.

Hecht and co-workers have successfully demonstrated that the modifications at multiple sites in the 23S rRNA can dramatically change the architecture of the ribosome while still preserving essential ribosomal functions.[48-51] The first report of modified ribosomes able to utilize α-D-amino acids was published by Hecht and co-workers in 2003.[48] Modifications in the PTC and helix 89 allowed the modified ribosomes to incorporate α-D-methionine and α-D-phenylalanine into *E. coli* DHFR in moderate yields.[48,49] Following this success, the same group reported the selection of ribosomes having modifications in the PTC of the 23S rRNA using β-puromycin (**1.2**, Figure 1.4).[50] The modified ribosomes were able to incorporate β-alanine into DHFR.[50]

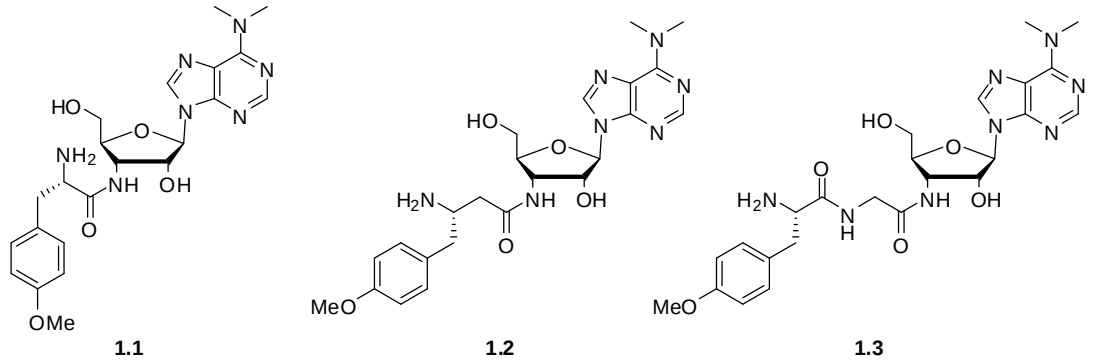


Figure 1.4. Structures of puromycin and its derivatives.[50]

In Chapter 2 are reported five modified ribosomes which were tested for their ability to incorporate five β-amino acids, including β-alanine into DHFR. The modified ribosomes having modifications in two regions of the 23S rRNA (2507-2063 and 2502-

2507 or 2496-2501) were selected using β-puromycin (**1.2**). Similar selection experiments, which were done using the puromycin derivative **1.3** (Figure 1.4), led to the selection of six modified ribosomes enabling the incorporation of a dipeptide and dipeptidomimetic into DHFR (Chapter 3). These modified ribosomes also had modifications in two regions of the 23S rRNA (2507-2063 and 2502-2507).

Chapter 2

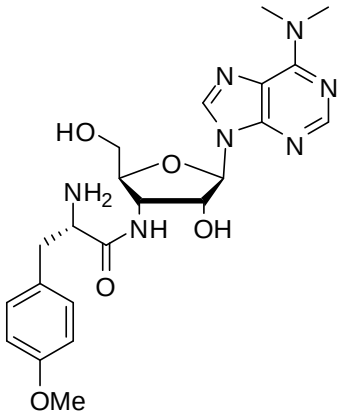
INCORPORATION OF β-AMINO ACIDS INTO *E. COLI* DHFR USING RIBOSOMES HAVING MODIFICATIONS IN THE PEPTIDYLTRANSFERASE CENTER (PTC)

2.1. Introduction

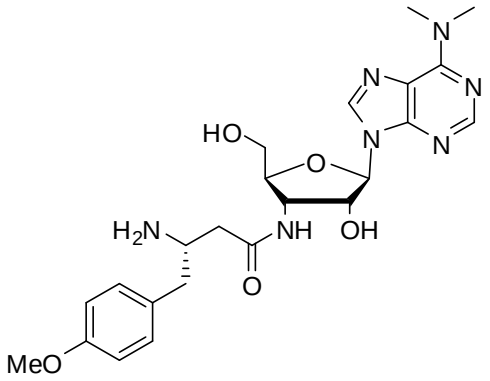
β-Amino acids are not found in ribosomally synthesized proteins or enzymes even though they are present in cells.[52] Synthetic peptide analogues having one or more β-amino acids have been shown to have helical properties, some of which are very similar to the overall structures of peptides formed from α-amino acids.[53-56] Additionally, such peptide analogues having β-amino acids often show enhanced resistance to proteolysis, which make them ideal candidates to be included in the growing list of unnatural amino acids that can be incorporated into proteins.[57-60] The incorporation of β-amino acids into proteins or large peptides can facilitate the creation of artificial protein-like macromolecules with interesting properties. At present, all peptide analogues having β-amino acids are prepared by synthetic or semi-synthetic methods which limits the size of peptide analogues accessible and thus their use. The largest β-mimetic peptide reported to date is only 17 amino acid long.[59] The strong bias of the ribosome against β-amino acids has limited their incorporation into proteins using *in vitro* or *in vivo* protein synthesis methods.[61,62]

As discussed in the previous chapter, puromycin is a small molecule mimic of the 3'-end of aminoacyl-tRNA (Figure 2.1) and is a universal translation inhibitor. It binds at the ribosomal A-site and readily accepts the peptide chain from the peptidyl-tRNA in the P-site, thereby terminating protein synthesis. Analogues of puromycin, in which the *O*-methyl-α-L-tyrosine moiety was replaced by either α-D-amino acids (D-alanine and 4-

methyl-D-phenylalanine) or β-amino acids (β-alanine and 4-methyl-β-L-phenylalanine), were shown to inhibit protein translation weakly as compared to puromycin.[47] This suggests that the ribosome can recognize β-aminoacyl- and α-D-aminoacyl-tRNA mimics.



1.1



1.2

Figure 2.1. Structures of puromycin (**1.1**) and β-puromycin (**1.2**).[50]

In a recent study, Hecht and co-workers demonstrated that multiple mutations in one or more regions of 23S rRNA can result in reorganization of the architecture of the ribosomal peptidyltransferase center in a fashion that facilitates the incorporation of α-D-amino acids[48,49] and β-amino acids[50,51] into protein. Using β-puromycin for selection of the *E. coli* colonies, eight variants of modified ribosomes were selected from a pool of ribosomes randomly mutagenized in two regions of 23S rRNA (2057-2063 and 2502-2507 or 2496-2501) (Figure 2.1).[50] β-Puromycin selection enabled the identification of the ribosomes having architectures suitable for recognizing β-amino acids. S-30 systems prepared from two modified ribosomes were shown to synthesize full size DHFR in

presence of β-alanyl-tRNA$_{CUA}$ and DHFR mRNA with UAG codons.[50] Based on UAG codon suppression to afford full length protein, it was demonstrated that β-alanine was incorporated into protein in moderate yields, but no direct evidence of β-alanine incorporation into DHFR was reported. Wild-type DHFR produced using DHFR mRNA and the modified ribosomes retained good activity, indicating that the ribosomes retained their ability to discriminate individual α-amino acids accurately.

In this thesis, five β-amino acids (Figure 2.2), including β-alanine, were chosen for their incorporation into DHFR. Five different types of modified ribosomes were selected to facilitate their incorporation into DHFR. Two modified DHFRs were selected for a larger scale synthesis and characterization by MALDI mass spectrometry after "in-gel" trypsin digestion.

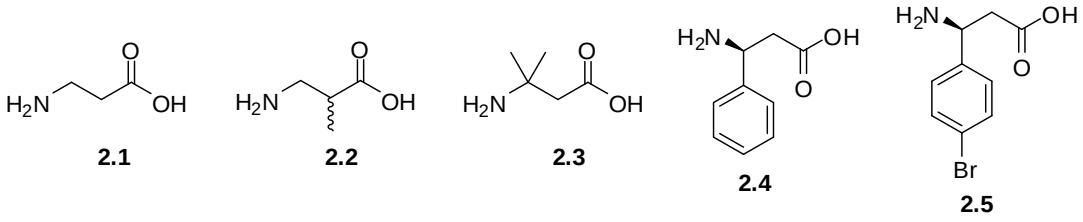


Figure 2.2. Structures of the β-amino acids studied.[51]

2.2. Results

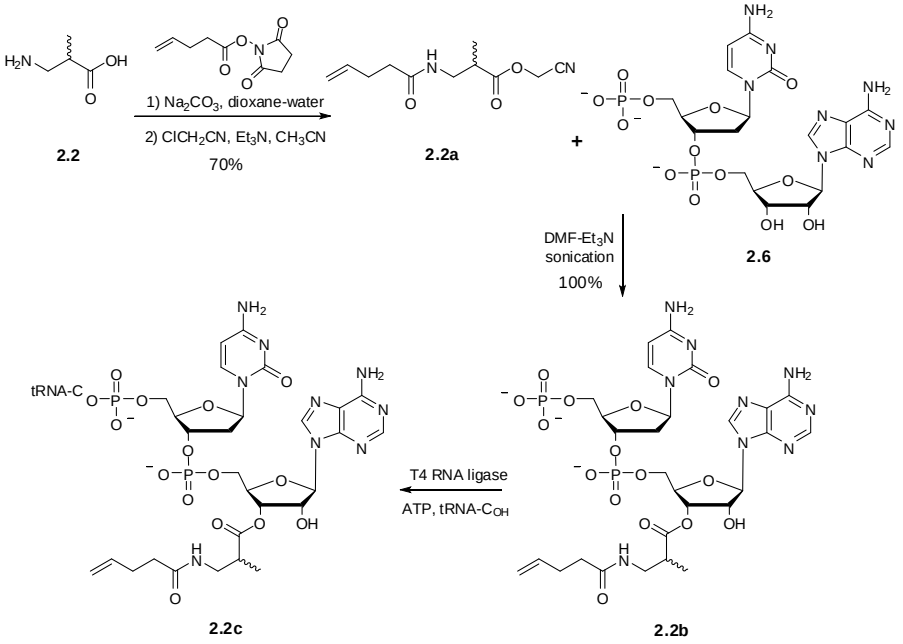
Five β-amino acids were selected either with no side chains (**2.1**) or having different side chains (**2.2**-**2.5**) at the α- or β-positions (Figure 2.2). β-Alanyl-tRNA$_{CUA}$[50] was provided by Dr. Larisa Dedkova and *N*-pentenoyl-α-methyl-β-alanyl-pdCpA[51] was synthesized by Sandipan Roy Chowdhury. Additionally, three *N*-pentenoyl-β-aminoacyl-pdCpA derivatives and four β-aminoacyl-tRNA$_{CUA}$s were prepared. Five different types

of modified ribosomes,[50] selected by Dr. Larisa Dedkova, were used for S-30 preparations.

2.2.1. Preparation of suppressor $tRNA_{CUA}$s activated with β-amino acids

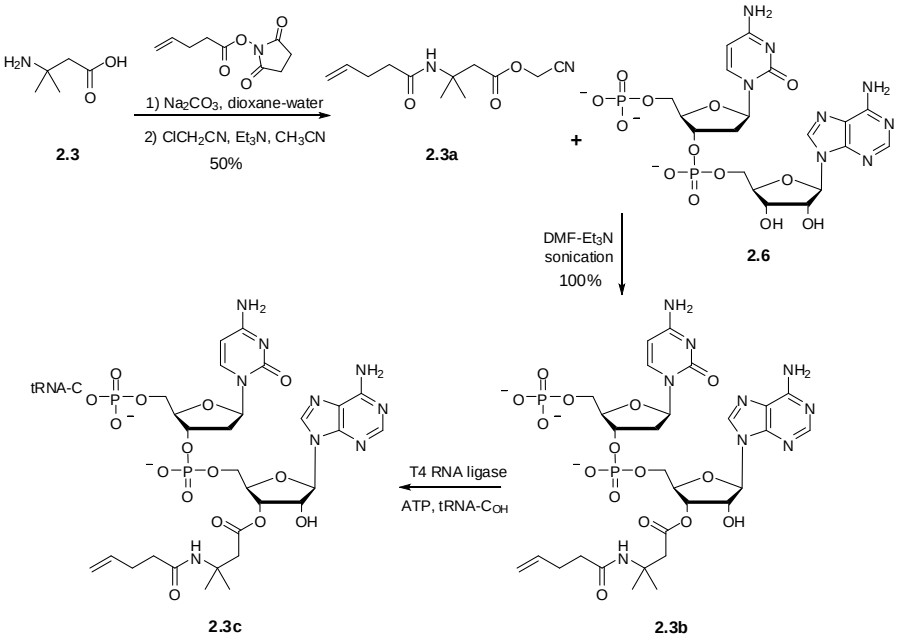
Commercially available amino acids were treated with an *N*-hydroxysuccinimide derivative of pentenoic acid to afford *N*-protected amino acids. The protected amino acids were subsequently treated with chloroacetonitrile in presence of triethylamine, to activate the carboxylic acid moieties as the respective cyanomethyl esters in 38 to 70% overall yields (Schemes 2.1-2.4).[63-66] The pdCpA derivatives of β-amino acids were prepared as previously described.[66] Aminoacylation of pdCpAs (Schemes 2.1-2.4) with the corresponding *N*-pentenoyl β-amino acid cyanomethyl esters was carried out in 9:1 DMF–Et_3N. The reaction mixtures were sonicated for 2 hours in a water bath. The *N*-pentenoyl β-aminoacyl-pdCpA products were purified by reverse phase HPLC using a C_{18} column and were obtained in quantitative yields. Activated pdCpA derivatives were ligated to abbreviated $tRNA_{CUA}$ transcript using T4 RNA ligase and four *N*-pentenoyl-β-aminoacyl-$tRNA_{CUA}$s were prepared (Schemes 2.1-2.4).[50] The *N*-pentenoyl-β-alanyl-$tRNA_{CUA}$ was provided by Dr. Larisa Dedkova. The *N*-pentenoyl protection was removed by treatment with aqueous iodine for five minutes.[63,64] Removal of the protecting group was done immediately prior to use of the misacylated tRNAs in protein synthesis. Each pdCpA derivative was ligated to the abbreviated $tRNA_{CUA}$ with 100% efficiency as shown by 8% polyacrylamide-7 M urea gel electrophoresis at pH 5.0 (Figure 2.3).[67]

Scheme 2.1. Preparation of *N*-pentenoyl-α-methyl-β-alanyl-tRNA$_{CUA}$. (Compounds **2.2a** and **2.2b** were synthesized by Sandipan Roy Chowdhury.)[51]



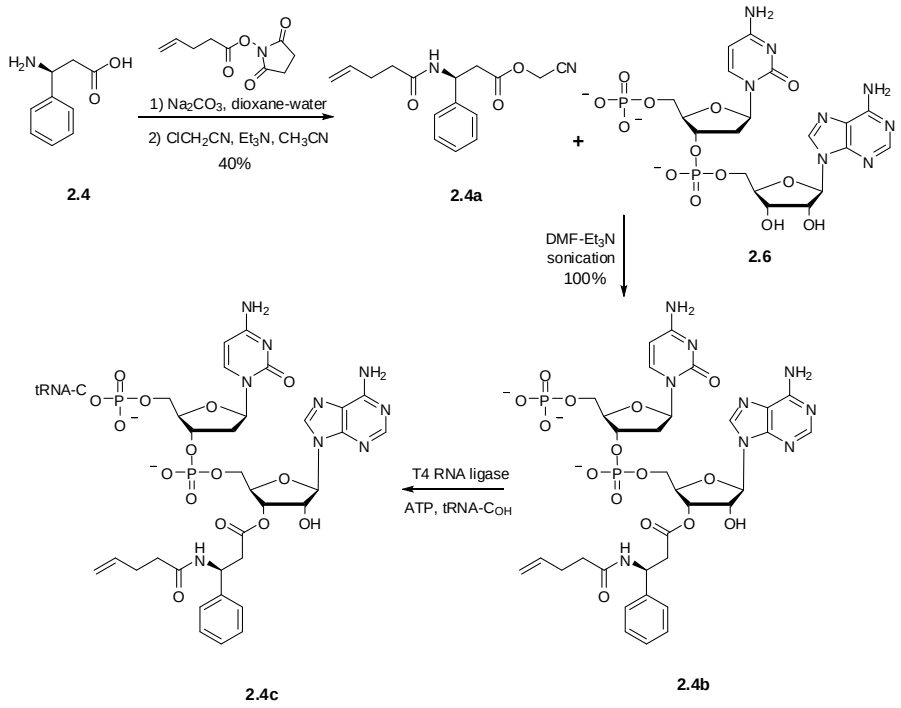
2.2 **2.2a** **2.6** **2.2c** **2.2b**

Scheme 2.2. Preparation of *N*-pentenoyl-β,β-dimethyl-β-alanyl-tRNA$_{CUA}$.[51]

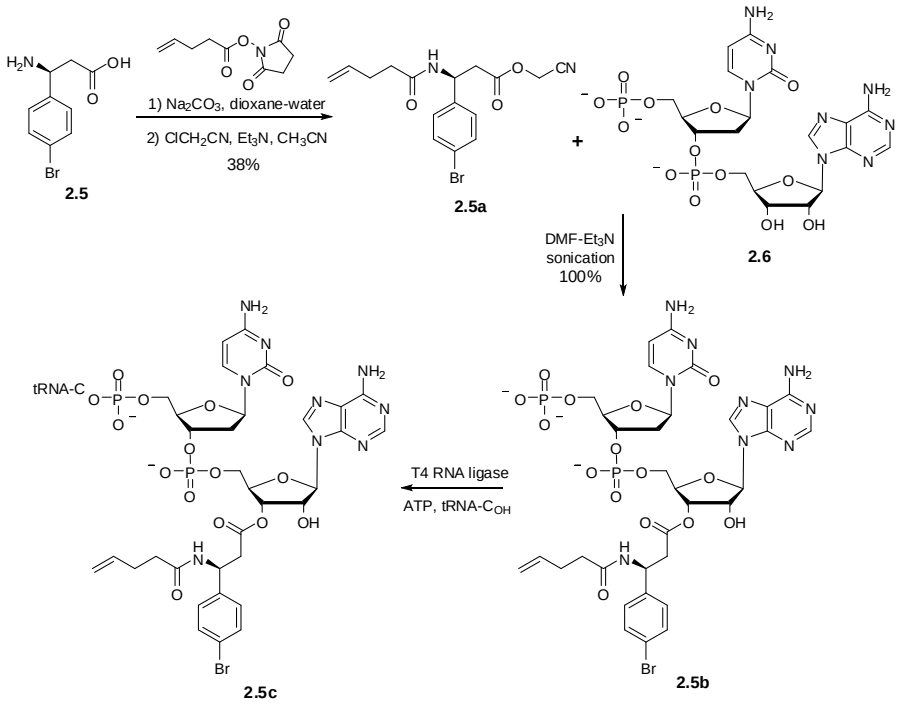


2.3 **2.3a** **2.6** **2.3c** **2.3b**

Scheme 2.3. Preparation of *N*-pentenoyl-β-phenylalanyl-tRNA$_{CUA}$.[51]



Scheme 2.4. Preparation of *N*-pentenoyl-β-(*p*-bromophenyl)alanyl-tRNA$_{CUA}$.[51]



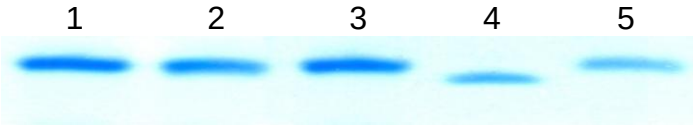


Figure 2.3. Preparation of β-aminoacyl-tRNA$_{CUA}$s. Lane 1, α-methyl-β-alanyl-tRNA$_{CUA}$; lane 2, β,β-dimethyl-β-alanyl-tRNA$_{CUA}$; lane 3, β-phenylalanyl-tRNA$_{CUA}$; lane 4, nonacylated abbreviated tRNA$_{CUA}$; lane 5, β-(*p*-bromophenyl)alanyl-tRNA$_{CUA}$.

2.2.2. *In vitro* translation using β-amino acids with wild-type ribosomes

As discussed previously in Section 2.1, wild-type ribosomes cannot mediate the incorporation of β-amino acids into proteins. Therefore, it seemed logical to verify this experimentally by the use of wild-type *E. coli* ribosomes. The ability of all five β-aminoacyl-tRNA$_{CUA}$s to suppress a UAG codon at position 18 or 49 of DHFR mRNA was compared with α-phenylalanyl-tRNA$_{CUA}$ and α-threonyl-tRNA$_{CUA}$ using the S-30 system prepared from the wild-type ribosome (Figures 2.4 and 2.5).

For position 49 of DHFR, the suppression efficiencies, relative to DHFR synthesized from the wild-type gene, were 40% and 50% for α-phenylalanyl-tRNA$_{CUA}$ and α-threonyl-tRNA$_{CUA}$, respectively. As illustrated in Figure 2.4, the suppression efficiencies for all five β-aminoacyl-tRNAs (0.4%-2.2%) were much lower than for α-phenylalanyl-tRNA$_{CUA}$ and α-threonyl-tRNA$_{CUA}$. Similar results were obtained when position 18 of DHFR was evaluated (Figure 2.5). The incorporation yields of β-amino acids into position 18 of DHFR were under 6% as compared to the wild-type DHFR synthesis. In the same experiment, high levels of full length DHFR were obtained in presence of α-phenylalanyl-tRNA$_{CUA}$ (80%) and α-threonyl-tRNA$_{CUA}$ (90%). These

results demonstrate that the wild-type ribosome cannot efficiently incorporate β-amino acids into proteins.

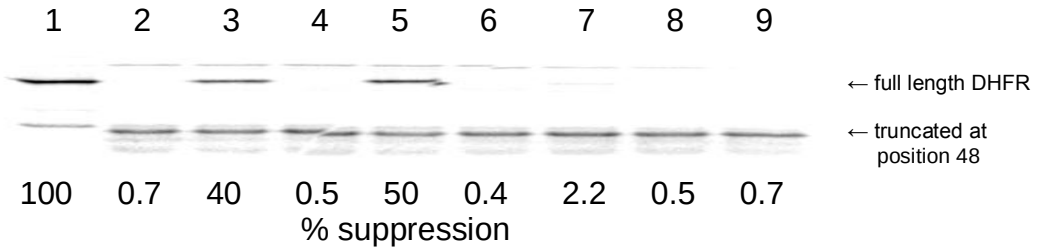


Figure 2.4. Translation of DHFR from wild-type (lane 1) and modified (lanes 2-9) (UAG codon in position 49) mRNA in the presence of different suppressor $tRNA_{CUA}$s. Lane 2, nonacylated $tRNA_{CUA}$; lane 3, $tRNA_{CUA}$ activated with phenylalanine; lane 4, $tRNA_{CUA}$ activated with β-phenylalanine (**2.4**); lane 5, $tRNA_{CUA}$ activated with threonine; lane 6, $tRNA_{CUA}$ activated with β-(*p*-bromophenyl)alanine (**2.5**); lane 7, $tRNA_{CUA}$ activated with β,β-dimethyl-β-alanine (**2.3**); lane 8, $tRNA_{CUA}$ activated with α-methyl-β-alanine (**2.2**); lane 9, $tRNA_{CUA}$ activated with β-alanine (**2.1**). The suppression efficiency relative to wild type is shown below each lane.[51]

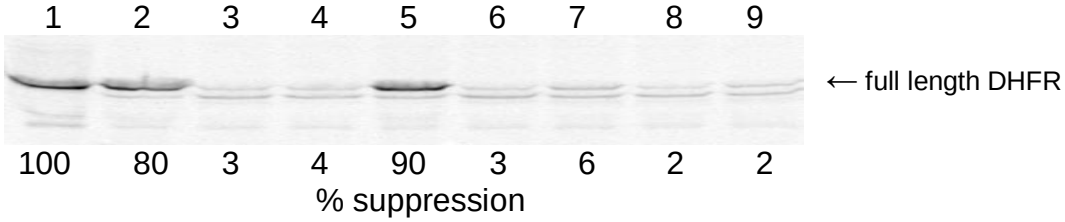


Figure 2.5. Translation of DHFR from wild-type (lane 1) and modified (lanes 2-9) (UAG codon in position 18) mRNA in the presence of different suppressor $tRNA_{CUA}$s. Lane 2,$tRNA_{CUA}$ activated with phenylalanine ; lane 3, nonacylated $tRNA_{CUA}$; lane 4,

$tRNA_{CUA}$ activated with β-phenylalanine (**2.4**); lane 5, $tRNA_{CUA}$ activated with threonine;lane 6, $tRNA_{CUA}$ activated with β-(*p*-bromophenyl)alanine (**2.5**); lane 7, $tRNA_{CUA}$ activated with β,β-dimethyl-β-alanine (**2.3**); lane 8, $tRNA_{CUA}$ activated with α-methyl-β-alanine (**2.2**); lane 9, $tRNA_{CUA}$ activated with β-alanine (**2.1**). The suppression efficiency relative to wild type is shown below each lane.[51]

2.2.3. *In vitro* translation using β-amino acids with modified ribosomes

To study the incorporation of β-amino acids into DHFR with modified ribosomes, five S-30 systems were prepared from *E. coli* colonies harboring different modified ribosomes. Their ability to individually synthesize full length DHFR in presence of β-aminoacyl-$tRNA_{CUA}$ and DHFR mRNA transcripts having a UAG codon was measured. The nucleotide sequence of the ribosomes in the modified regions is compared to the wild-type ribosome in Table 2.1. S-30 systems prepared from the modified ribosomes produced different amounts of wild-type DHFR as compared to the wild-type ribosomal system. Therefore, the suppression efficiencies were expressed relative to the suppression obtained using α-threonyl-tRNA (Table 2.2) or α-phenylalanyl-tRNA (Table 2.3). As a negative control, DHFR synthesis in the presence of nonacylated-$tRNA_{CUA}$ was evaluated. The amounts of DHFR produced by each system were quantified with a phosphoimager, which monitored the incorporation of ^{35}S-methionine into DHFR.

Table 2.2 summarizes the incorporation yields of β-amino acids **2.1**, **2.2** and **2.3** into DHFR at position 10, relative to α-threonine incorporation. Figure 2.6 illustrates the synthesis of full length DHFR in presence of β-aminoacyl-$tRNA_{CUA}$s bearing β-amino

Table 2.1. 23S rRNA sequence modifications in the PTC for the clones used for S-30 preparations.[51]

Clone	Sequence in 23S rRNA of modified ribosome	
	Region 1	Region 2
0403x2	2057AGCGUGA2063	2502UUACCG2507
0403x4	2057AGCGUGA2063	2502AGCCAG2507
040321	2057AGCGUGA2063	2502AGAUAA2507
040329	2057AGCGUGA2063	2502UGGCAG2507
040217	2057AGCGUGA2063	2496AUAGUU2501
Wild type	2057GAAAGAC2063	2496CACCUC2501 2502GAUGUC2507

acids **2.1**, **2.2** and **2.3** using the S-30 system prepared from clone 040329. A racemic mixture of **2.2** was deliberately employed for this study as to see any effect from either of the isomers. The individual β-amino acids were incorporated with slight variations in suppression yields (~2-fold) with the five S-30 preparations having modified ribosomes. Among the three β-amino acids examined for the incorporation into positon 10 of DHFR, the best results were obtained for **2.3** (~11% suppression yield) with S-30 systems prepared from clones 040329 and 040217. Figure 2.7 illustrates the synthesis of full length DHFR in presence of β-aminoacyl-tRNA$_{CUA}$s bearing β-amino acids **2.1** and **2.3** using the S-30 system prepared from clone 040217. All five S-30 preparations produced very low levels of non-specific readthrough of the UAG codon in the absence of any activated suppressor tRNA.

Table 2.3 summarizes the incorporation yields of β-amino acids **2.2**, **2.4** and **2.5** into DHFR at position 18, relative to α-phenylalanine incorporation. These experiments were done in a fashion similar to the experiments described above. The S-30 preparations

from clones 040329 and 0403x4 produced the best results for the three amino acids tested.

Table 2.2. Incorporation of β-amino acids **2.1**, **2.2** and **2.3** into position 10 of *E. coli* DHFR by the use of S-30 systems having different modified ribosomes.[51]

Amino acids	Suppression efficiency in different S-30 systems, having modified ribosomes (%)[a]				
	040329	0403x2	0403x4	040321	040217
Thr	100	100	100	100	100
–	0.5 ± 0.3	2.1 ± 1.4	1.4 ± 0.7	0.9 ± 0.2	2.2 ± 1.0
2.1	8.2 ± 3.0	6.8 ± 3.2	8.1 ± 3.4	5.4 ± 1.3	9.9 ± 3.9
2.2	8.8 ± 1.0	3.8 ± 1.0	5.4 ± 1.9	not tested	not tested
2.3	11[b]	7.9 ± 1.1	7.1 ± 3.6	5.6 ± 3.0	10.9 ± 4.4

[a] Each number represents the average of three independent experiments ± SD.

[b] tested in a single experiment.

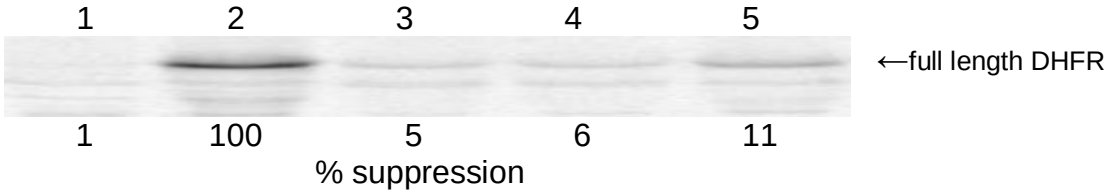


Figure 2.6. Translation of DHFR from a modified DHFR mRNA (UAG codon in position 10) by the use of an S-30 system prepared from clone 040329 in the presence of different suppressor tRNAs. Lane 1, nonacylated $tRNA_{CUA}$; lane 2, threonyl-$tRNA_{CUA}$; lane 3, $tRNA_{CUA}$ activated with α-methyl-β-alanine (**2.2**); lane 4, $tRNA_{CUA}$ activated with β-alanine (**2.1**); lane 5, $tRNA_{CUA}$ activated with β,β-dimethyl-β-alanine (**2.3**).[51]

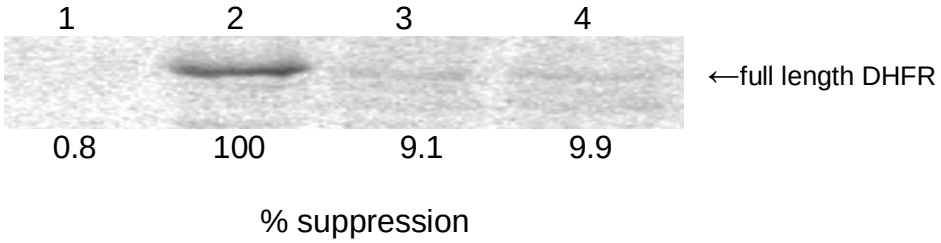


Figure 2.7. Translation of DHFR from a modified DHFR mRNA (UAG codon in position 10) by the use of an S-30 system prepared from clone 040217 in the presence of different suppressor tRNAs. Lane 1, nonacylated $tRNA_{CUA}$; lane 2, threonyl-$tRNA_{CUA}$; lane 3, $tRNA_{CUA}$ activated with β-alanine (**2.1**); lane 4, $tRNA_{CUA}$ activated with β,β-dimethyl-β-alanine (**2.3**).[51]

Table 2.3. Incorporation of β-alanine analogues **2.2**, **2.4** and **2.5** into position 18 of *E. coli* DHFR by the use of S-30 systems having different modifications in the PTC.[51]

Amino acids	**Suppression efficiency in different S-30 systems, having modified ribosomes (%)**[a]				
	040329	0403x2	0403x4	040321	040217
Phe	100	100	100	100	100
–	1.5 ± 0.7	1.2 ± 0.4	2.4 ± 1.0	0.8 ± 0.4	1.0 ± 0.4
2.2	6.0 ± 0.5	1.5 ± 1.0	6.0 ± 1.8	1.3 ± 0.7	0.73 ± 0.2
2.4	18.4 ± 2.7	5.8 ± 1.3	13.2 ± 1.1	6.0 ± 1.8	7.8 ± 0.1
2.5	13.5 ± 1.5	5.3 ± 0.9	15.5 ± 2.0	7.6 ± 2.7	9.2 ± 0.4

[a]Each number represents the average of three independent experiments ± SD.

Figures 2.8 and 2.9 illustrate the incorporation of β-amino acids **2.2**, **2.4** and **2.5** into DHFR at position 18 using the S-30 systems prepared from clones 040329 and 0403x4, respectively. The highest yield of full length DHFR (~18%) was obtained by the combination of β-amino acid **2.4** and the S-30 system prepared from clone 040329. The

incorporation of **2.4** into DHFR gave yields between ~5.8% to ~18%, whereas **2.5** produced full length DHFR in yields ranging from ~5.3% to ~15%.

S-30 systems prepared from clones 0403x2 and 040321 did not perform as well as the other three systems for the incorporation of β-amino acids into positions 10 or 18 of DHFR. For position 10 of DHFR, S-30 systems prepared from clones 0403x2 and 040321 gave best results with β-amino acid **2.3**; ~7.9% and ~5.6% suppression yields, respectively (Table 2.2). In contrast, clone 040321 performed better (~7.6% suppression yield with **2.5**) than clone 0403x2 (~5.8% suppression yield with **2.4**) for position 18 of DHFR. As summarized in Table 2.3, it is evident that the five modified ribosomes did not incorporate **2.2** efficiently into position 18 of DHFR as compared to β-amino acids **2.4** and **2.5**. The suppression yields for β-amino acid **2.2** varied from ~0.7% (for clone 040217) to ~6% (for clone 040329) for position 18. For position 10, S-30 systems prepared from three clones (040329, 0403x2 and 0403x4) were tested for the incorporation of **2.2** into DHFR. The highest amount of full length DHFR having β-amino acid **2.2** was produced by the S-30 system having clone 040329 (~8.8% suppression yield, Table 2.2).

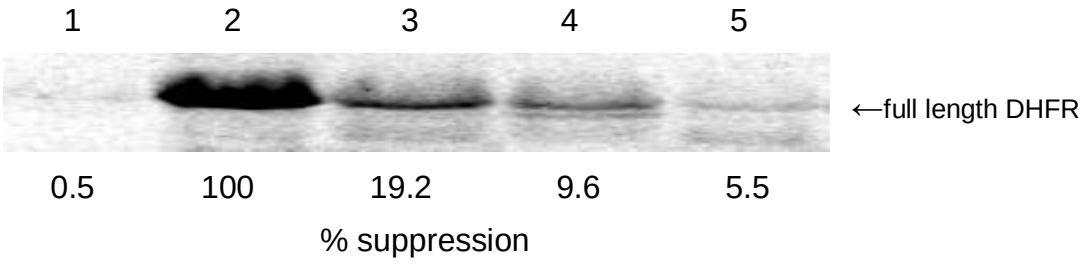


Figure 2.8. Translation of DHFR from a modified DHFR mRNA (UAG codon in position 18) by the use of an S-30 system prepared from clone 040329 in the presence of

different suppressor tRNAs. Lane 1, nonacylated $tRNA_{CUA}$; lane 2, phenylalanyl-$tRNA_{CUA}$; lane 3, $tRNA_{CUA}$ activated with β-phenylalanine (**2.4**); lane 4, $tRNA_{CUA}$ activated with β-(*p*-bromophenyl)-alanine (**2.5**); lane 5, $tRNA_{CUA}$ activated with α-methyl-β-alanine (**2.2**).[51]

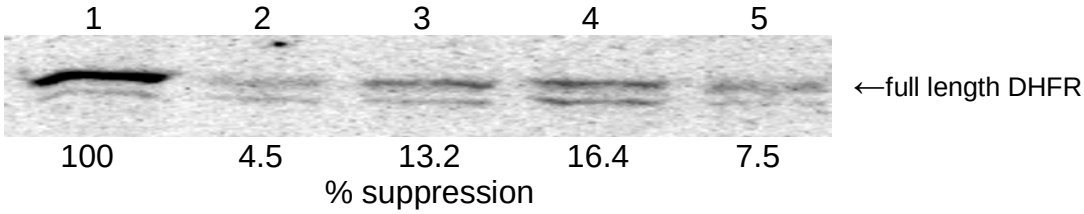


Figure 2.9. Translation of DHFR from a modified DHFR mRNA (UAG codon in position 18) by the use of an S-30 system prepared from clone 0403x4 in the presence of different suppressor tRNAs. Lane 1, phenylalanyl-$tRNA_{CUA}$; lane 2, nonacylated $tRNA_{CUA}$; lane 3, $tRNA_{CUA}$ activated with β-phenylalanine (**2.4**); lane 4, $tRNA_{CUA}$ activated with β-(*p*-bromophenyl)-alanine (**2.5**); lane 5, $tRNA_{CUA}$ activated with α-methyl-β-alanine (**2.2**).[51]

2.2.4. Suppression of the UAG codon in three different positions of DHFR mRNA by β-(*p*-bromophenyl)alanyl-$tRNA_{CUA}$

The ability of β-(*p*-bromophenyl)alanyl-$tRNA_{CUA}$ to suppress the UAG codon in three different mRNA (UAG codon at position 10, 18 or 49) transcripts was studied. As shown in Figure 2.10, β-amino acid **2.5** was incorporated into positions 10, 18 and 49 of DHFR with yields of 11%, 7.4% and 6.3%, respectively. Expression yields for all three experiments were compared to that of phenylalanine. These results indicate that the

modified ribosomes can incorporate β-amino acids at different positions in a protein with similar efficiencies in an *in vitro* translation system.

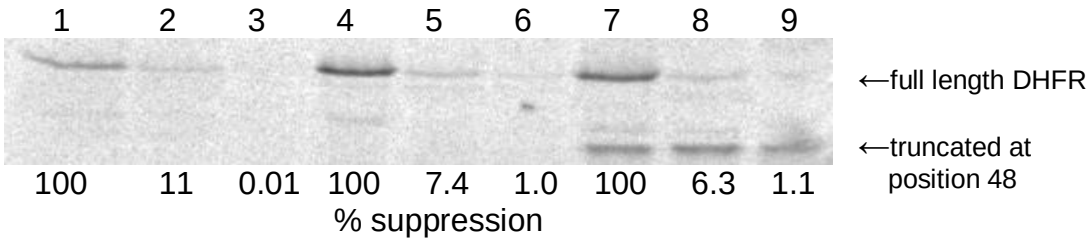


Figure 2.10. Relative suppression efficiencies of β-(*p*-bromophenyl)alanyl-tRNA$_{CUA}$ into three different positions (10, 18 and 49) of DHFR, using an S-30 system prepared from clone 040217. Lanes 1, 4 and 7, L-phenylalanyl-tRNA$_{CUA}$; lanes 2, 5 and 8, β-(*p*-bromophenyl)alanyl-tRNA$_{CUA}$; lanes 3, 6 and 9, unacylated tRNA$_{CUA}$. The suppression efficiency relative to L-phenylalanine is shown below each lane. Lanes 1-3, modification of position 10 in DHFR; lanes 4-6, modification of position 18 in DHFR; lanes 7-9, modification of position 49 in DHFR.[51]

2.2.5. Characterization of β-amino acids introduced into DHFRs by MALDI mass spectrometry analysis of the peptide fragments resulting from "in-gel" trypsin digestion

Two modified DHFRs were synthesized on a larger scale for characterization by MALDI-MS to provide a direct evidence of the incorporation of β-amino acids into DHFR. The plasmid pETD49 and β-(*p*-bromophenyl)alanyl-tRNA$_{CUA}$ were used to obtain DHFR 1 modified at position 49 with β-amino acid **2.5**. Similarly, pETD18 and β,β-dimethyl-β-alanyl-tRNA$_{CUA}$ were used to obtain DHFR 2 modified at position 18 with β-

amino acid **2.3**. Three stages of purification were performed on both the modified DHFRs including Ni-NTA chromatography, DEAE-Sephadex chromatography and SDS-polyacrylamide gel electrophoresis. A previously reported "in-gel" trypsin digestion protocol was used.[68] Samples of the wild-type and modified DHFRs were treated with trypsin to obtain peptide fragments which were subjected to MALDI-MS analysis. Table 2.4 summarizes the calculated and observed (in the MALDI mass spectrum) molecular weights of the peptide fragments obtained from trypsin digestion of the wild-type and

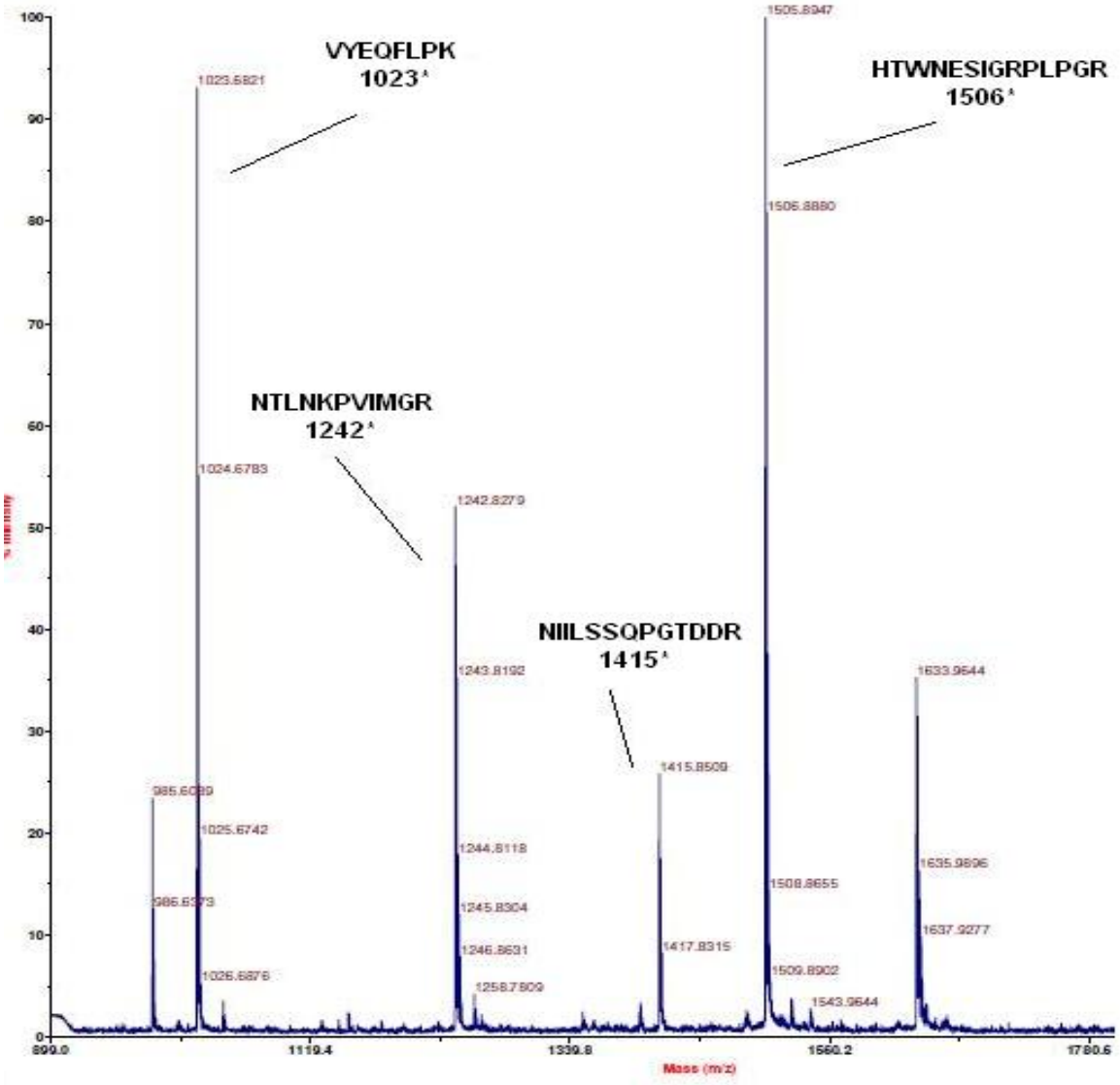
Table 2.4. MALDI-MS analysis of tryptic digests of wild-type and modified DHFR samples.[51]

Position	**Peptide sequence**	**MALDI-MS analysis, molecular mass, Da**					
		Wild-type		Modified DHFR 1		Modified DHFR 2	
		Est	MS	Est	MS	Est	MS
13-32	VIGME**N**AMPWNLPADLAWFK	2304	2304.1	2304	2303.6	**2289**	**2289.9**
34-44	NTLNKPVIMGR	1242	1242.8	1242	1242.6	1242	1242.8
45-57	HTWE**S**IGRPLPGR	1506	1505.8	**1644**	**1645.6**	1506	1505.9
59-71	NIILSSQPGTDDR	1415	1415.8	1415	1415.6	1415	1415.8
72-76	VTWVK	632	632.5	632	634.3	632	632.4
77-98	VDEAIAASGDVPEIMVIGGGR	2028	2029.1	2028	2029.7	2028	2030.7
99-106	VYEQFLPK	1023	1023.7	1023	1023.5	1023	1023.6

Modified DHFR 1: β-alanine analogue **2.5** in position 49; modified DHFR 2: β-alanine analogue **2.3** in position 18. Positions 18 and 49 are indicated in red. The values in bold face reflect the presence of the respective β-amino acid.

two modified DHFRs. For DHFR 1, the tryptic peptide encompassing amino acids 45-57 having β-amino acid **2.5** at position 49, was anticipated to have a molecular mass of 1644 Da. An ion peak was observed at *m/z* 1645.5 displaying the characteristic bromine isotope pattern. This confirmed the presence of the anticipated peptide fragment (Figure 2.11B). Similarly, DHFR 2 produced an ion peak corresponding to *m/z* 2289.9, which corresponds to peptide fragment (amino acids 13-32) having β-amino acid **2.3** at position 18 (Figure 2.12B). MALDI-MS analysis of the tryptic digest of wild-type DHFR did not show ion peaks corresponding to *m/z* 1644 and 2289.9. The anticipated ion peaks for the wild-type peptide fragments encompassing amino acids 45-57 with Ser49 at *m/z* 1506 (Figure 2.11A) or 13-32 with Asn18 at *m/z* 2304 (Figure 2.12A) were present. As listed in Table 2.4, the remaining six peptide fragments obtained from wild-type DHFR matched with those from the modified DHFRs 1 and 2. In replicate experiments, the N-terminal fragment (containing a hexahistidine fusion peptide) and the (large) C-terminal fragment were never observed.

A.


VYEQFLPK
1023*
HTWNESIGRPLPGR
1506*
NTLNKPVIMGR
1242*
NIILSSQPGTDDR
1415*
1023.6821
1024.6783
1025.6742
1026.6876
1242.8279
1243.8192
1244.8118
1245.8304
1246.8631
1258.7809
1415.8509
1417.8315
1505.8947
1506.8880
1508.8655
1509.8902
1543.9644
1633.9644
1635.9896
1637.9277
890.0
1119.4
1339.8
1560.2
1780.6
Mass (m/z)

B.

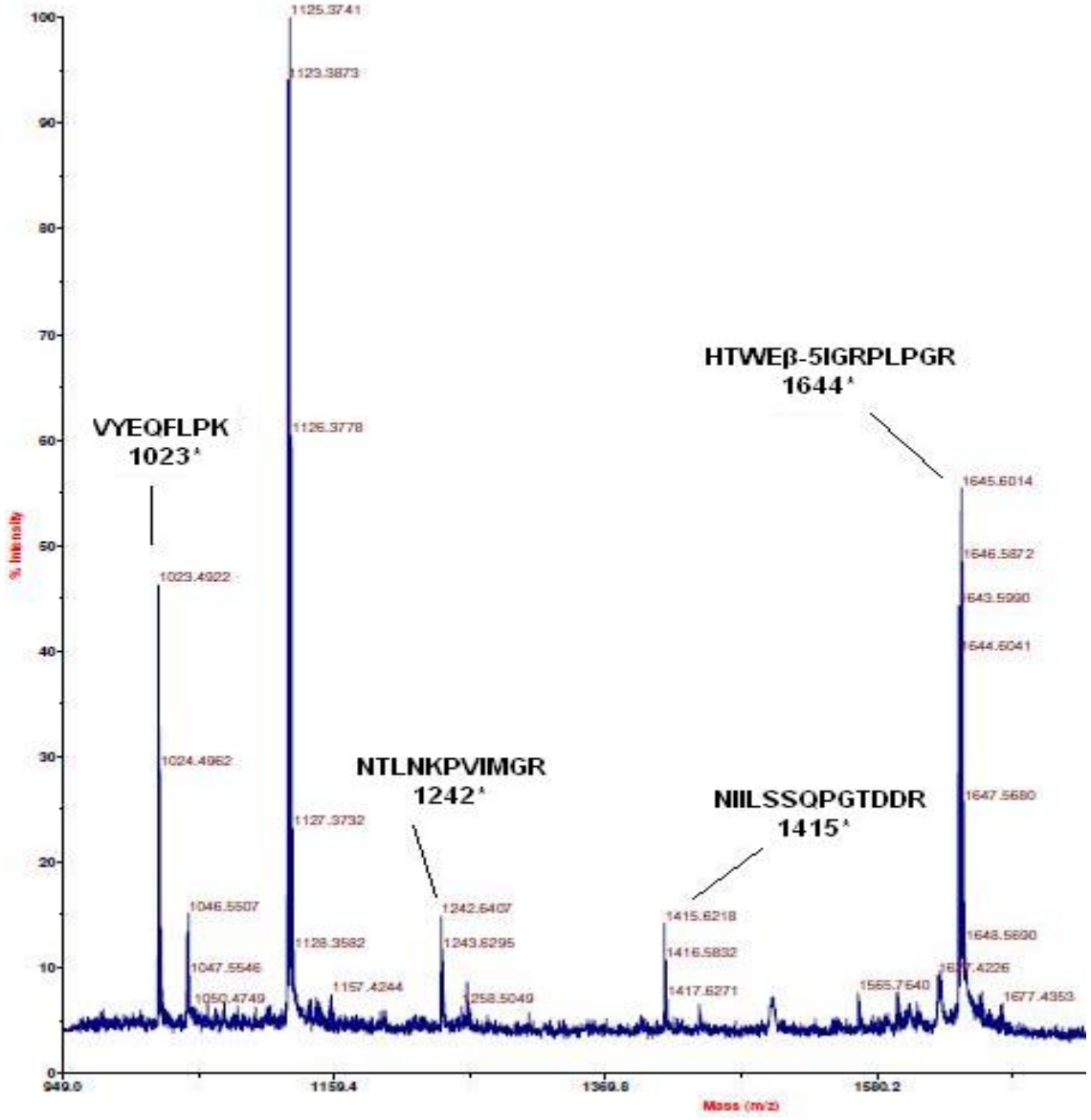
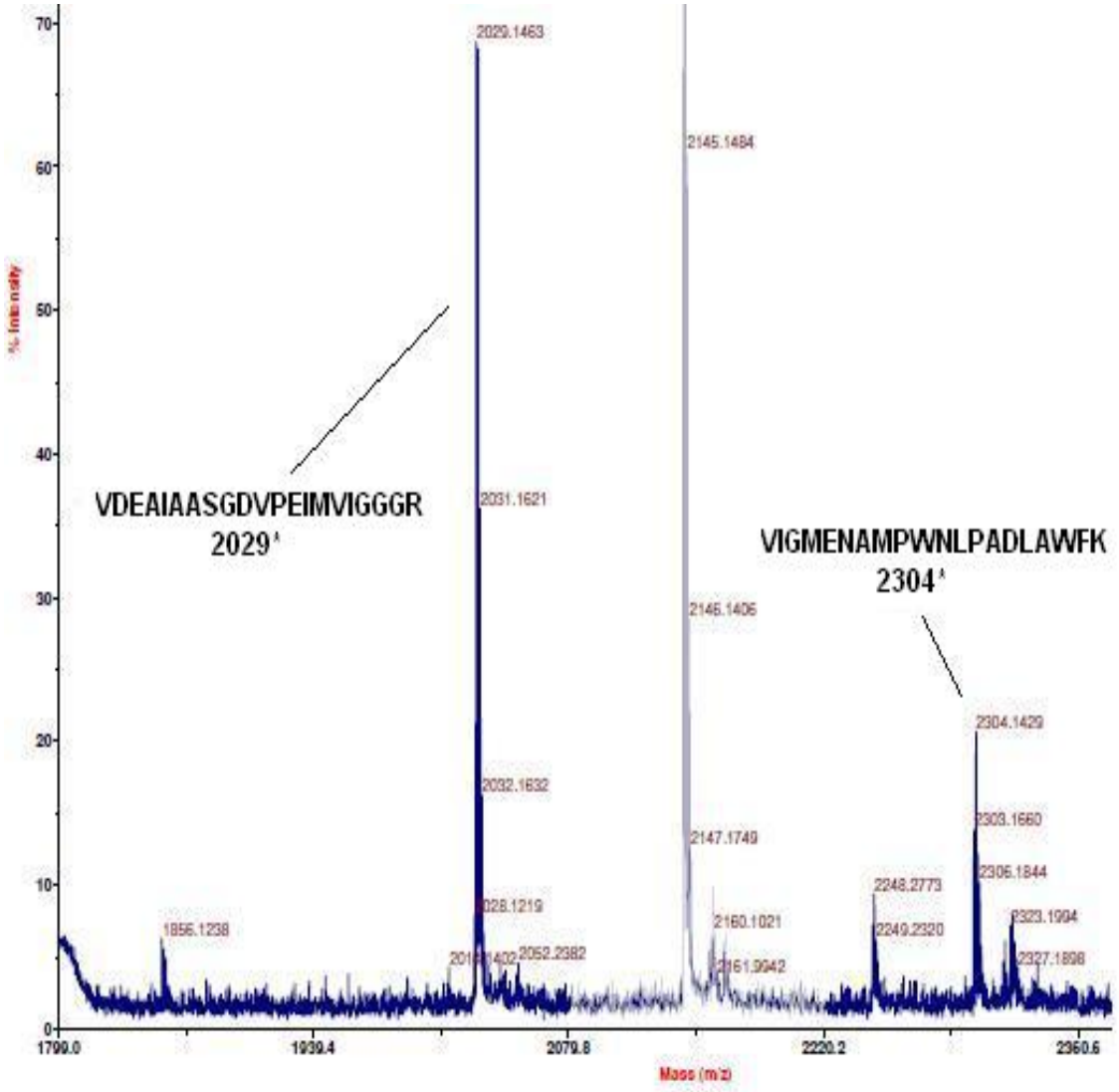


Figure 2.11. MALDI-MS of tryptic fragments of wild-type (**A**) and modified DHFRs (**B**), the latter having β-amino acid **2.5** in position 49. Mass range 1000-1800 Da (* = estimated value in Da).[51]

A.


2029.1463
2145.1484
VDEAIAASGDVPEIMVIGGGR
2029*
2031.1621
VIGMENAMPWNLPADLAWFK
2304*
2145.1406
2304.1429
2032.1632
2303.1660
2147.1749
2306.1844
2248.2773
2323.1994
2160.1021
2249.2320
1856.1238
2052.2382
2161.3942
2327.1898
% Intensity
Mass (m/z)
1799.0
1939.4
2079.8
2220.2
2360.6

B.

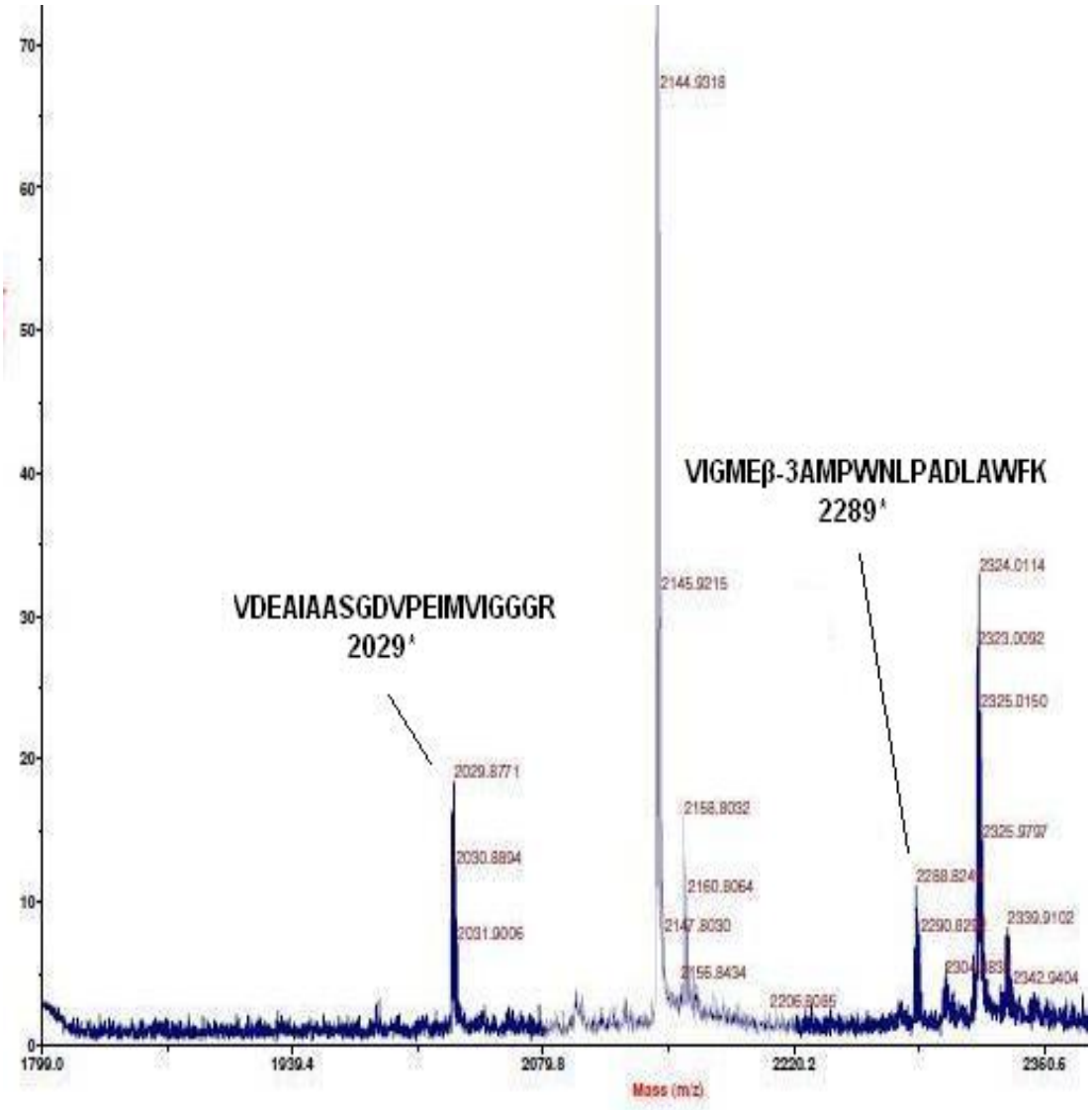


Figure 2.12. MALDI-MS of tryptic fragments of wild-type (**A**) and modified DHFRs (**B**), the latter having β-amino acid **2.3** in position 18. Mass range 1800-2500 Da. (* = estimated value in Da).[51]

2.3. Discussion

In this thesis, four novel β-aminoacyl-tRNA$_{CUA}$s were prepared and the ability of five β-aminoacyl-tRNA$_{CUA}$s to suppress the UAG codon in DHFR mRNA transcripts via S-30 preparations from five different modified ribosomes (Table 2.1) was studied. The preparations of the aminoacylated suppressor tRNAs involved four reactions starting from commercially available β-amino acids. The protection by the amino group as *N*-pentenoyl, followed by the activation of the carboxylic acid moiety as the cyanomethyl ester, afforded the desired products in moderate yields (Schemes 2.1-2.4).

Coupling of the activated β-amino acids with pdCpA (**2.6)** gave the corresponding pdCpA derivatives, activated with β-amino acids at either the 2'- or 3'-position, in quantitative yields after reverse phase HPLC purification (Schemes 2.1-2.4). Subsequently, all four β-aminoacyl-pdCpA derivatives were ligated to a 74-nucleotide abbreviated tRNA$_{CUA}$ by the use of T4 RNA ligase. Full length suppressor tRNA$_{CUA}$s activated with the β-amino acids were obtained in quantitative yields as shown by analysis on a 8% polyacrylamide-7 M urea gel electrophoresis at pH 5.0 (Figure 2.3).

Part of the 23S rRNA (nucleotides 2058 to 2609)[42] forms the peptidyl transferase center of the ribosome. Using β-puromycin for selection of colonies, five variants of modified ribosomes were selected from a random pool of ribosomes having modifications in two regions of 23S rRNA (2057-2063 and 2502-2507 or 2496-2501). The selection experiments were carried out by Dr. Larisa Dedkova.[50] The S-30 preparation protocols followed were as reported earlier.[50] S-30 systems prepared from five modified ribosomal clones were tested for their expression of the wild-type DHFR relative to an S-30 system prepared from the wild-type ribosome. The expression yields

of DHFR from the wild-type gene using S-30 systems having different modified ribosomes ranged from 22% to 77% relative to the S-30 system prepared from the wild-type ribosome. Therefore, the suppression efficiencies of β-aminoacyl-$tRNA_{CUA}$s in producing full length DHFR were compared to α-threonyl-$tRNA_{CUA}$ (Table 2.2) or α-phenylalanyl-$tRNA_{CUA}$ (Table 2.3), also using the same modified ribosomes.

Three β-amino acids (**2.3**, **2.4** and **2.5**, Figure 2.2) were chosen to study the effect of different side chains at the β-position on incorporation yields. Amino acid **2.2** was chosen to study the effect of substitution at α-position. A racemic mixture of **2.2** was deliberately employed for this study to assure that any effect from either of the isomers would be observed. The incorporation yields of β-amino acids into DHFR increased from ~10% to ~19% with increasing hydrophobic character of the side chains. The incorporation of β-(*p*-bromophenyl)alanine into DHFR was slightly higher with three of the five S-30 systems tested. The S-30 preparations from clones 040329 and 0403x2 incorporated β-phenylalanine better than its bromo derivative. In general, β-amino acid **2.2**, having a methyl group at α-position, produced low levels of full length DHFR. The selection procedure utilized β-puromycin in which the side chain of the amino acid moiety is at the β-position. Therefore, it is perhaps not surprising that the selected S-30 systems incorporated β-amino acid **2.2** into DHFR less effectively than they incorporated the other β-amino acids studied (Tables 2.2 and 2.3).

The variants of the ribosomes tested were modified in two regions of 23S rRNA, nucleotides 2057-2063 (region 1) and nucleotides 2502-2507 or 2496-2501 (region 2). All the clones had same sequence in region 1 but a variable sequence in region 2. The wild-type sequence 2057**GAAAGAC**2063 in this region was changed to

2057**AGCGUA**2063 for all the clones. Nucleotide G2061 is the member of the hydrogen bonding network that orients A2451 in the peptidyltransferase center.[20] A G2061U substitution was present in all of the clones and may be crucial for the steric factors responsible for the acceptance of β-aminoacyl-tRNA in the active site. Nucleotide C2063 forms a non conventional base pair with A2450 which is located in the vicinity of A2451.[69] A C2063A substitution was found to be present in all five clones studied, which would most likely disrupt the interaction with A2450. All adenosine nucleotides (A) were substituted by either G or C, changing the hydrogen bond acceptor from a N-atom to an O-atom.

In the second region, there is low sequence homology between the four clones having modifications in region 2502-2507. Clones 0403x4 (2502**AGCCAG**2507) and 040329 (2502**UGGCAG**2507) have good homology in this region. Clones 040321 and 0403x2 have little homology to either of the clones mentioned above. One clone (040217) had modification in a different region (2496-2502) altogether.

The region 2502-2507 is responsible for the stabilization of the 3'-ACC end of aminoacyl-tRNA in the A-site. Nucleotide A_{76} of the aminoacyl-tRNA interacts with a wobble base pair U2506-G2583 in the wild-type ribosome.[20] For clones 0403x4, 040329, 040321 and 0403x2, nucleotide U2506 was substituted by either A or C. Nucleotide A_{76} of the peptidyl-tRNA in the P-site interacts with non Watson-Crick base pair C2501-A2450 which is present in region 2496-2502.[20] Clone 040217 had a C2501U substitution in the second region. Therefore, the mutations in the second region of all the five clones might perturb the interaction of 3'-end of the tRNAs with either the A-site or the P-site.

β-amino acid **2.5** was selected to study the suppression of UAG codon in positions 10, 18 and 49 of the DHFR gene by β-aminoacyl-tRNA$_{CUA}$. S-30 preparations from the modified ribosomal clone 040217 afforded the incorporation of **2.5** into position 10 in 11% yield relative to α-phenylalanine, whereas for positions 18 and 49, suppression yields were 7.4% and 6.3%. This demonstrated that the modified ribosomes can achieve incorporation of β-amino acids into different positions of a protein in reasonable yields.

To verify the incorporation of β-amino acids into DHFR by selected ribosomes, direct evidence was sought. A widely used method, "in-gel" trypsin digestion followed by MALDI-MS analysis[68] of the peptide fragments, was performed on two modified DHFRs. Large scale translation reactions were carried out to synthesize modified DHFRs 1 and 2. DHFR 1 had β-amino acid **2.5** substituted at position 49 and DHFR 2 had β-amino acid **2.3** at position 18. Approximately 1 μg of each modified DHFR was obtained after three stages of purification including Ni-NTA chromatography, DEAE-Sephadex chromatography and SDS-polyacrylamide gel electrophoresis. A tryptic digest of wild-type DHFR was used as a control for MALDI-MS analysis (Table 2.4). In case of the modified DHFR 1, an ion peak at *m/z* 1645.6 showing a characteristic bromine isotope pattern was observed (Figure 2.11B). The estimated molecular mass of the peptide fragment encompassing amino acids 45-57 with β-(*p*-bromophenyl)alanine substituted at position 49 is 1645 Da. Figure 2.11A shows the presence of an ion peak at *m/z* 1506, reflecting the presence of wild-type peptide fragment with serine at position 49. Similarly, MALDI-MS analysis of the tryptic digest from modified DHFR 2 gave an ion peak at *m/z* 2289.9 (Figure 2.12B) and confirmed the incorporation of β-amino acid **2.3** into position 18 of DHFR (estimated mass is 2289 Da). The corresponding wild-type

peptide fragment gave the anticipated ion peak at *m/z* 2304 (Asn18) as seen from Figure 2.12A. For both the modified DHFRs, ion peaks for the other six peptide fragments correlated with wild-type DHFR peptides. The ion peak corresponding to the N-terminal fragment containing a hexahistidine fusion peptide was never seen. This may be a result of aggregation of the peptide caused by histidine residues. The large C-terminal fragment was also not observed for either wild-type or modified DHFRs.

2.4. Experimental

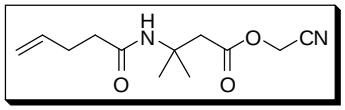
2.4.1. General materials and methods

Reagents and solvents for chemical synthesis were purchased from Aldrich Chemical Co. or Sigma Chemical Co. and were used without further purification. Compounds **2.4** and **2.5** were purchased from Peptech Corporation and were used without further purification. All reactions involving air- or moisture-sensitive reagents or intermediates were performed under argon. Analytical TLC was performed using Silicycle silica gel 60 Å F254 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; bs, broad singlet; d, doublet; q, quartet; m, multiplet. High resolution mass spectra were obtained at the Arizona State University CLAS High Resolution Mass Spectrometry Facility or the Michigan State University Mass Spectrometry Facility.

Tris, acrylamide, bis-acrylamide, urea, ammonium persulfate, *N*,*N*,*N'*,*N'*-tetramethylenediamine (TEMED), dihydrofolic acid, glycerol, ampicillin, pyruvate kinase, lysozyme, erythromycin**,** isopropyl-β-D-thiogalactopyranoside (IPTG), dithiothreitol (DTT) and 2-mercaptoethanol were purchased from Sigma Chemicals (St. Louis, MO). ^{35}S-Methionine (10 μCi/μL) was obtained from Amersham (Pitscataway, NJ). BL-21 (DE-3) competent cells and T4 RNA ligase were purchased from Promega (Madison, WI). Plasmid MaxiKit (Life Science Products, Inc., Frederick, CO) and GenElute^{TM}HP plasmid miniprep kit (Sigma) were used for plasmid purification.

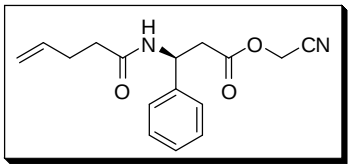
Phosphorimaging analysis was performed using a Molecular Dynamics 400E PhosphoImager equipped with ImageQuant version 3.2 software. Ultraviolet and visible spectral measurements were made using a Perkin-Elmer lambda 20 spectrophotometer.

2.4.2. Synthesis of *N*-4-Pentenoyl Cyanomethyl Ester β-Amino Acids.



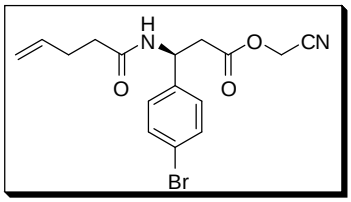
***N*-4-Pentenoyl-*β*,*β*-dimethyl-*β*-alanine Cyanomethyl Ester (2.3a).** To a solution containing 100 mg (0.85 mmol) of **2.3** and 140 mg (1.64 mmol) of $NaHCO_3$ in 4 mL of 1:1 dioxane–H_2O was added 180 mg (0.9 mmol) of 4-pentenoyloxysuccinimide. The reaction mixture was stirred at room temperature for 24 h under argon. The reaction was quenched by the addition of 15 mL of 1 N aq $NaHSO_4$ and the aqueous layer was extracted with two 25-mL portions of ethyl acetate. The combined organic extract was dried over anhydrous $MgSO_4$ and concentrated under diminished pressure. The crude product was then dissolved in 4 mL of acetonitrile. To this solution were added 600 μL (4.2 mmol) of triethylamine and 260 μL (4.2 mmol) of chloroacetonitrile. The reaction

mixture was stirred at room temperature for 24 h at which time 20 mL of ethyl acetate was added. The organic layer was washed with 10 mL of 1 N aq $NaHSO_4$ followed by 10 mL of brine, then dried over anhydrous $MgSO_4$ and concentrated to dryness under diminished pressure, affording a crude residue. The crude product was purified by flash silica gel column chromatography (15 x 1 cm) using 1:1 ethyl acetate–hexanes for elution to obtain **2.3a** as colorless oil: yield 119 mg (50%); R_f 0.5 (1:1 ethyl acetate–hexanes); ^{1}H NMR ($CDCl_3$) δ 1.38 (s, 6H), 2.16-2.20 (m, 2H), 2.29-2.34 (m, 2H), 2.92 (s, 2H), 4.66 (s, 2H), 4.95-5.05 (m, 2H), 5.66 (bs, 1H) and 5.73-5.83 (m, 1H); ^{13}C NMR ($CDCl_3$) δ 27.5, 29.4, 36.3, 42.2, 47.9, 51.9, 114.4, 115.4, 137.0, 169.6 and 172.2; mass spectrum (APCI), *m/z* 239.1398 $(M+H)^+$ ($C_{17}H_{21}NO_3$ requires *m/z* 239.1396).



***N*-4-Pentenoyl-*β*-*S*-phenylalanine Cyanomethyl Ester (2.4a)**. **)**. To a solution containing 100 mg (0.6 mmol) of **2.4** and 105 mg (1.22 mmol) of $NaHCO_3$ in 4 mL of 1:1 dioxane–H_2O was added 135 mg (0.62 mmol) of 4-pentenoyloxysuccinimide. The reaction mixture was stirred at room temperature for 24 h under argon. The reaction was quenched by the addition of 15 mL of 1 N aq $NaHSO_4$ and the aqueous layer was extracted with two 25-mL portions of ethyl acetate. The combined organic extract was dried over anhydrous $MgSO_4$ and concentrated under diminished pressure. The crude product was then dissolved in 4 mL of acetonitrile. To this solution were added 450 µL (3.15 mmol) of triethylamine and 195 µL (3.15 mmol) of chloroacetonitrile. The reaction mixture was stirred at room temperature for 24 h at which time 20 mL of ethyl acetate

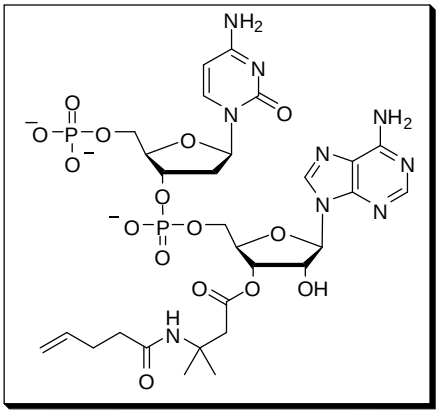
was added. The organic layer was washed with 10 mL of 1 N aq $NaHSO_4$ followed by 10 mL of brine, then dried over anhydrous $MgSO_4$ and concentrated to dryness under diminished pressure, affording a crude residue. The crude product was purified by flash silica gel column chromatography (15 x 1 cm) using 1:1 ethyl acetate–hexanes for elution to obtain **2.4a** as colorless semi-solid: yield 114 mg (40%); R_f 0.31 (1:1 ethyl acetate–hexanes); ^{1}H NMR ($CDCl_3$) δ 2.22-2.36 (m, 4H), 2.82-3.01 (m, 2H), 4.59 (d, 2H, J = 2.4 Hz), 4.95-5.04 (m, 2H), 5.39 (q, 1H, J = 11.6 and 6.4 Hz), 5.71-5.81 (m, 1H), 6.59 (d, 1H, J = 8.0 Hz) and 7.23-7.33 (m, 5H); ^{13}C NMR ($CDCl_3$) δ 29.4, 35.6, 39.6, 48.5, 49.7, 114.2, 115.7, 126.4, 128.0, 128.9, 136.9, 139.8, 169.4 and 171.9; mass spectrum (APCI), *m/z* 287.1516 $(M+H)^+$ ($C_{17}H_{21}NO_3$ requires *m/z* 286.1521).



***N*-4-Pentenoyl-*β*-*S*-(*p*-bromophenyl)alanine Cyanomethyl Ester (2.5a)**. To a solution containing 100 mg (0.41 mmol) of **2.5** and 70.0 mg (0.82 mmol) of $NaHCO_3$ in 4 mL of 1:1 dioxane–H_2O was added 90 mg (0.45 mmol) of 4-pentenoyloxysuccinimide. The reaction mixture was stirred at room temperature for 24 h under argon. The reaction was quenched by the addition of 15 mL of 1 N aq $NaHSO_4$ and the aqueous layer was extracted with two 25-mL portions of ethyl acetate. The combined organic extract was dried over anhydrous $MgSO_4$ and concentrated under diminished pressure. The crude product was then dissolved in 4 mL of acetonitrile. To this solution were added 300 µL (2.10 mmol) of triethylamine and 130 µL (2.10 mmol) of chloroacetonitrile. The reaction mixture was stirred at room temperature for 24 h at which time 20 mL of ethyl acetate

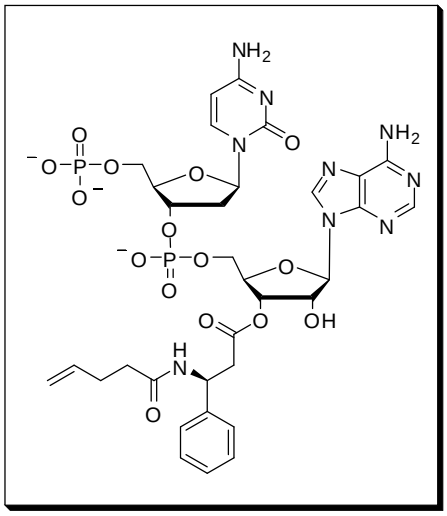
was added. The organic layer was washed with 10 mL of 1 N aq $NaHSO_4$ followed by 10 mL of brine, then dried over anhydrous $MgSO_4$ and concentrated to dryness under diminished pressure, affording a crude residue. The crude product was purified by flash silica gel column chromatography (15 x 1 cm) using 1:1 ethyl acetate–hexanes for elution to obtain **2.5a** as colorless semi-solid: yield 140 mg (38%); R_f 0.33 (1:1 ethyl acetate–hexanes); ^{1}H NMR ($CDCl_3$) δ 2.27-2.40 (m, 4H), 2.85-3.03 (m, 2H), 4.65 (s, 2H), 4.99-5.07 (m, 2H), 5.39 (q, 1H, J = 6.8 Hz), 5.74-5.84 (m, 1H), 6.45 (d, 1H, J = 8.0 Hz), 7.15 (d, 2H, J = 8.0 Hz) and 7.45-7.47 (m, 2H); ^{13}C NMR ($CDCl_3$) δ 29.4, 35.6, 39.2, 48.6, 49.0, 113.9, 115.9, 121.9, 128.1, 132.0, 136.7, 138.8, 169.3 and 171.8; mass spectrum (APCI), *m/z* 365.0631 $(M+H)^+$ ($C_{17}H_{20}NO_3Br$ requires *m/z* 365.0627).

2.4.3. Synthesis of β-aminoacyl-pdCpA derivatives.

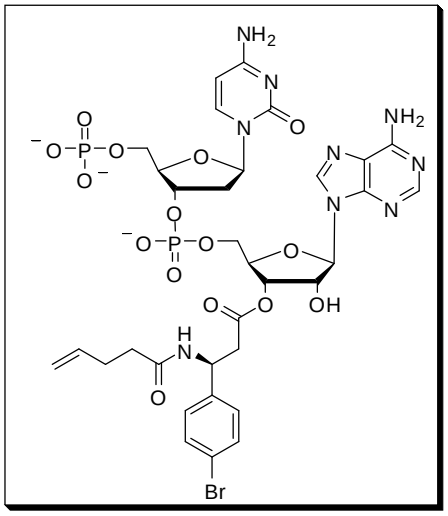


***N*-4-Pentenoyl-*β*,*β*-dimethyl-*β*-alanyl-pdCpA (2.3b).** To a 1.5-mL Eppendorf tube containing 5.3 mg (3.7 µmol) of the tris(tetrabutylammonium) salt of pdCpA was added 10 mg (42 µmol) of **2.3a** dissolved in 100 µL of 9:1 DMF–Et_3N. After 2 h of sonication, the reaction mixture was purified by C_{18} reversed phase HPLC (250 x 10 mm) using a gradient of 1%→65% acetonitrile in 50 mM ammonium acetate, pH 4.5, over a period of 45 min. The fractions eluting at 14.4 and 14.9 min were collected, combined and

lyophilized to afford **2.3b** as a colorless solid: yield 3.6 mg (100%); mass spectrum (ESI), *m/z* 816.2130 $(M-H)^-$ ($C_{29}H_{40}N_9O_{15}P_2$ requires *m/z* 816.2125).



***N*-4-Pentenoyl-*β*-*S*-phenylalanyl-pdCpA (2.4b).** To a 1.5-mL Eppendorf tube containing 5.3 mg (3.7 µmol) of the tris(tetrabutylammonium) salt of pdCpA was added 10 mg (35 µmol) of **2.4a** dissolved in 100 µL of 9:1 DMF–Et_3N. After 2 h of sonication, the reaction mixture was purified by C_{18} reversed phase HPLC (250 x 10 mm) using a gradient of 1%→65% acetonitrile in 50 mM ammonium acetate, pH 4.5, over a period of 45 min. The fractions eluting at 17.4 and 18.5 min were collected, combined and lyophilized to afford **2.4b** as a colorless solid: yield 3.2 mg (100%); mass spectrum (ESI), *m/z* 864.2126 $(M-H)^-$ ($C_{33}H_{40}N_9O_{15}P_2$ requires *m/z* 864.2125).



***N*-4-Pentenoyl-*β*-*S*-(*p*-bromophenyl)alanyl-pdCpA (2.5b)**.To a 1.5-mL Eppendorf tube containing 5.3 mg (3.7 µmol) of the tris(tetrabutylammonium) salt of pdCpA was added 10 mg (27 µmol) of **2.5a** dissolved in 100 µL of 9:1 DMF–Et_3N. After 2 h of sonication, the reaction mixture was purified by C_{18} reversed phase HPLC (250 x 10 mm) using a gradient of 1%→65% acetonitrile in 50 mM ammonium acetate, pH 4.5, over a period of 45 min. The fractions eluting at 20.8 and 22.1 min were collected, combined and lyophilized to afford **2.5b** as colorless solid: yield 3.1 mg (100%); mass spectrum (ESI), *m/z* 942.1234 $(M-H)^-$ ($C_{33}H_{39}N_9O_{15}BrP_2$ requires *m/z* 942.1230).

2.4.4. Preparation of β-aminoacyl-tRNA$_{CUA}$s (2.2c, 2.3c, 2.4c and 2.5c). The activation of suppressor tRNA$_{CUA}$s was carried out as described previously.[50] Briefly, 100-µL reaction mixtures of 100 mM Na Hepes, pH 7.5, contained 1.0 mM ATP, 15 mM $MgCl_2$, 100 µg of suppressor tRNA$_{CUA}$-C_{OH}, 0.5 A_{260} unit of *N*-pentenoyl-protected □-aminoacyl-pdCpA, 15% DMSO, and 100 units of T4 RNA ligase. The reaction mixture was incubated at 37 °C for 1 h, then quenched by the addition of 0.1 vol of 3 M NaOAc, pH 5.2. The N-protected aminoacylated tRNA was precipitated with 3 vol of cold

ethanol. The efficiency of ligation was estimated by 8% polyarylamide–7 M urea gel electrophoresis at pH 5.0 (Figure 2.3, Section 2.2).[52]

The N-pentenoyl-protected aminoacyl-tRNA$_{CUA}$s were deprotected by treatment with 5 mM aqueous I_2 at 25 °C for 15 min. The solution was centrifuged, and the supernatant was adjusted to 0.3 M NaOAc and treated with 3 vol of cold ethanol to precipitate the aminoacylated tRNA. The tRNA pellet was collected by centrifugation, washed with 70% aq EtOH, air dried and dissolved in 50 μL of RNase free water.

2.4.5. Preparation of S-30 extracts from cells having modified ribosomes. Aliquots (5-10 μL) from liquid stocks of *E. coli* BL-21(DE-3) cells, harboring plasmids with a wild-type or modified rrnB gene, were placed on LB agar supplemented with 100 μg/mL of ampicillin and grown at 37 ºC for 16-18 h. One colony was picked from each agar plate and transferred into 3 mL of LB medium supplemented with 100 μg/mL of ampicillin and 0.5 mM IPTG. The cultures were grown at 37 ºC for 3-6 h in a thermostated shaker until OD_{600} ~ 0.15-0.3 was reached (about 3-5 h), diluted with LB medium supplemented with 100 μg/mL ampicillin, 1 mM IPTG and 3 μg/mL of erythromycin (for selectively enhancing the modified ribosome fraction) until OD_{600} 0.01 was reached, and then grown at 37 ºC for 12-18 h. The optimal concentration of the final cultures was OD_{600} 0.5-1.0. Cells were harvested by centrifugation (5000 × g, 4 ºC, 10 min), washed three times with S-30 buffer (1 mM Tris-OAc, pH 8.2, containing 1.4 mM $Mg(OAc)_2$, 6 mM KOAc and 0.1mM DTT) supplemented with β-mercaptoethanol (0.5 mL/L) and once with S-30 buffer having 0.05 mL/L β-mercaptoethanol. The weight of the wet pellet was estimated and 1.27 mL of S-30 buffer was added to suspend each 1 g of cells. The volume of the suspension was measured and used for estimating the amount

of other components. Pre-incubation mixture (0.3 mL) (0.29 M Tris, pH 8.2, containing 9 mM $Mg(OAc)_2$, 13 mM ATP, 84 mM phosphoenol pyruvate, 4.4 mM DTT and 5 μM amino acids mixture), 15 units of pyruvate kinase and 10 μg of lyzozyme were added per 1 mL of cell suspension and the resulting mixture was incubated at 37 ºC for 30 min. The incubation mixture was then frozen at – 80 ºC (~ 30 min), melted (37 ºC, 30 min), and again frozen and melted at room temperature (~ 30 min). Ethylene glycol tetraacetic acid (EGTA) was then added to 2.5 mM final concentration and the cells were incubated at 37 ºC for 30 min and again frozen (– 80 ºC, 30 min). The frozen mixture was centrifuged (15,000 × g, 4 ºC, 1 h) and the supernatant was stored in aliquots at – 80 ºC.

2.4.6. *In vitro* protein translation. Protein translation reactions were carried out in 12-2000 μL of incubation mixture containing 0.2-0.4 μL/μL of S-30 system, 100 ng/μL of plasmid, 35 mM Tris acetate, pH 7.4, 190 mM potassium glutamate, 30 mM ammonium acetate, 2 mM DTT, 0.2 mg/mL total *E. coli* tRNA, 3.5% PEG 6000, 20 μg/mL folinic acid, 20 mM ATP and GTP, 5 mM CTP and UTP, 100 μM amino acids mixture, 0.5 μCi/μL of ^{35}S-methionine and 1 μg/mL rifampicin. In the case of plasmids having a gene with a TAG codon, a suppressor tRNA was added to a concentration of 0.3-0.5 μg/μL (for α-aminoacyl-tRNAs) and 0.6-1.0 μg/μL (for β-aminoacyl-tRNAs). Reactions were carried out at 37 °C for 1 h and terminated by chilling on ice. Aliquots from *in vitro* translation mixtures were analyzed by SDS-PAGE followed by quantification of the radioactive bands by phosphorimager analysis.

2.4.7. Purification of DHFR. Samples of DHFR, prepared during *in vitro* translation, were diluted with 50 mM Na phosphate, pH 8.0, containing 0.3 M NaCl and 0.1 mg/ml

BSA and applied to a 50-μL Ni-NTA agarose column that had been equilibrated with the same buffer. The column was washed with three 500-μL portions of the same buffer and DHFR was eluted by washing with 150 μL of the same buffer also containing 250 mM imidazole.

A 50-μL column of DEAE-Sepharose was equilibrated with 500 μL of 10 mM Na phosphate, pH 7.4. Samples of DHFR purified by Ni-NTA chromatography were diluted 10-fold in the same buffer and applied to the resin. The column was washed with 500 μL of the same buffer. DHFR was eluted with a NaCl step gradient (0.1; 0.2; 0.3; 0.4; 0.5 M) in 10 mM of Na phosphate, pH 7.4. The fractions were analyzed on a 12% polyacrylamide gel, stained using Coomassie R-250, then destained in 20% ethanol–7% acetic acid. The fractions containing DHFR were combined, dialyzed against 50 mM Na phosphate, pH 7.4 and concentrated by ultrafiltration using a YM-10 filter (Amicon Ultra, Millipore Corp, Billerica, MA).

2.4.8. "In gel" trypsin digestion. Samples to be digested in the gel were run in 3-4 lanes of a 12% SDS-polyacrylamide gel, stained with Coomassie R-250 and destained until the background was clear. That area of the gel having the DHFR was cut from the gel and washed with 0.1 M ammonium bicarbonate (1 h, room temperature). The solution was discarded and 0.1-0.2 mL of 0.1 M ammonium bicarbonate and 10-30 μL of 0.045 mM DTT were added. Gel pieces were incubated at 60 ºC for 30 min, cooled to room temperature and incubated at room temperature for 30 min in the dark after the addition of 10-30 μL of 0.1 M iodoacetamide. Gel pieces were washed in 50% acetonitrile–0.1 M ammonium bicarbonate until they became colorless. After discarding the solution, the gel pieces were incubated in 0.1-0.2 mL of acetonitrile (10-20 min at room temperature) and,

after removal of solvent, were re-swelled in 50-100 μL of 25 mM ammonium bicarbonate containing 0.02μg/μL trypsin. After incubation at 37 ºC for 4 h, the supernatant was transferred to a new tube and the peptides were extracted with 60% acetonitrile–0.1% aq TFA (20 min at room temperature). The combined fractions were dried and reconstituted in 10 μL of 60% acetonitrile–0.1% aq TFA for each 1 μg of the protein (measured before trypsin digestion). Approximately, 5 μL of this solution was mixed thoroughly with 2.5 μL of the matrix solution (saturated solution of dihydroxybenzoic acid in 60% acetonitrile–0.1% aq TFA). A 2 μL aliqout from the resulting solution was deposited on one of the wells in a 90-well MALDI-MS plate and dried.This was repeated twice and the plate was subjected to MALDI-MS analysis.

Chapter 3

INCORPORATION OF A DIPEPTIDE AND DIPEPTIDOMIMETIC INTO DHFR USING RIBOSOMES HAVING MODIFICATIONS IN THE PEPTIDYLTRANSFERASE CENTER

3.1. Introduction

Peptidomimetics are compounds that spatially mimic a natural peptide or protein and may retain the ability to interact with a biological target in a manner similar to the natural counterpart. For biological and pharmaceutical studies, peptidomimetics offer interesting advantages over physiologically active peptides. Peptidomimetics can solve problems associated with natural peptides, e.g. stability against proteolysis and poor bioavailability.[70] Other properties, such as receptor selectivity or potency, can be substantially improved with rational design of mimetics.[71] Hence, peptidomimetics have great potential in drug discovery.

3.1.1. β-Turn peptidomimetics

Reversel turns are commonly occurring structural features in globular proteins.[72] β-turns are one of the most important reverse turn motifs and are often located on the surface of proteins, most likely acting as recognition and antigenic sites.[73] In a β-turn, a tight loop is formed when the carbonyl oxygen of one residue (*i*) forms a hydrogen bond with the amide proton of an amino acid (*i*+3) three residues away (Figure 3.1). Bicyclic dipeptide mimetics (Figure 3.2) have been used extensively in peptides to generate β-turn conformations.[74,75] Recently, proline based dipeptidomimetics have been shown to be capable of enhancing biological activities when incorporated into peptides[76] or proteins.[77] Hartley and co-workers synthesized analogues of an anti-HIV protein called RANTES

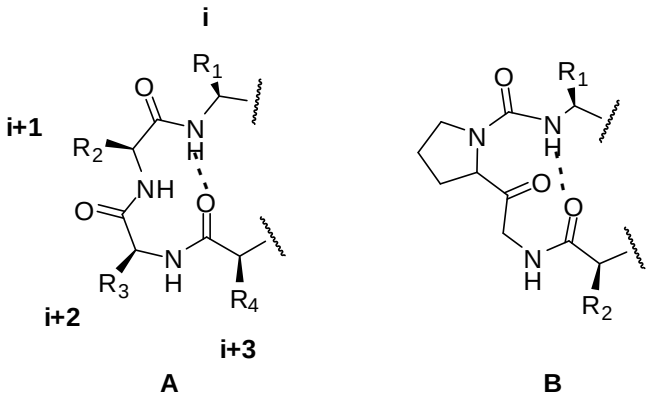


Figure 3.1. General structure of a β-turn (**A**) and a β-turn motif containing pro-gly (**B**) .[1]

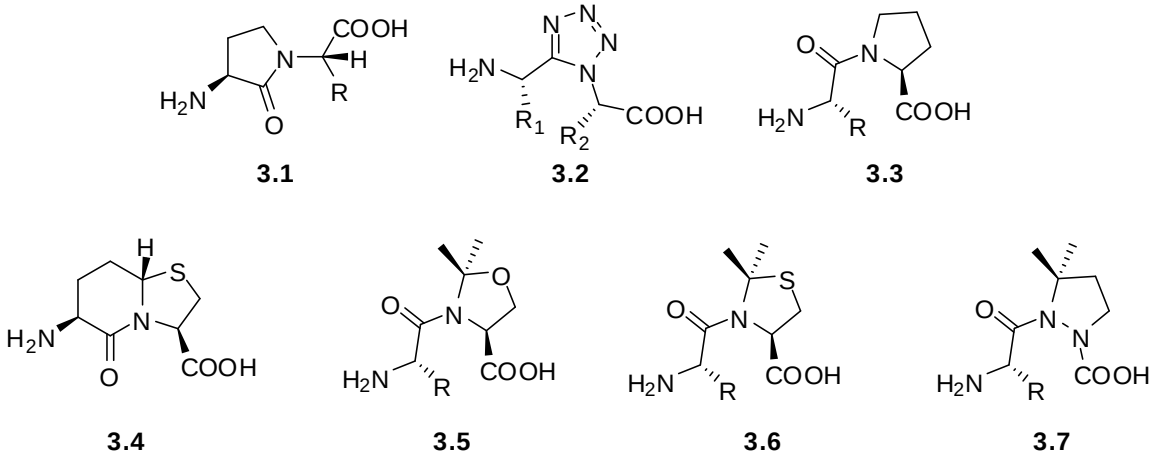


Figure 3.2. Structures of dipeptide β-turn mimetics.[74,75]

bearing the thioproline based dipeptidomimetics shown in Figure 3.3.[77] RANTES is a chemokine ligand of CC-chemokine receptor 5 (CCR5), which inhibits CCR5 dependent HIV entry into target cells.[78] One of the protein analogues, having dipeptide mimic **3.8** (Figure 3.3), showed 50-fold enhancement in potency as compared to AOP–RANTES, the first synthetic analogue of the protein RANTES.[77]

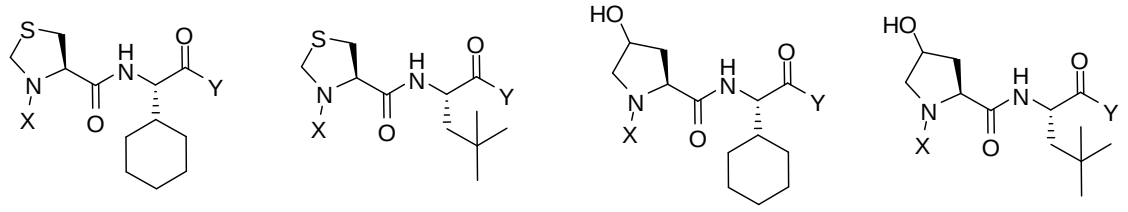


Figure 3.3. Structures of thioproline-based dipeptidomimetics.[77]

3.1.2. Peptide bond isosteres

The backbone of a peptide can be modified by isosteric or isoelectronic exchange of structural elements in the peptide or by the introduction of additional fragments.[79] A few examples of peptide bond isosteres are shown in Figure 3.4. A *p*-cyanophenylalanine (Cnf) containing an abbreviated protein fragment was ligated to a thioamide (**3.12**, Figure 3.4) labeled synthetic peptide to produce a full length protein, α-synuclein, using the native chemical ligation technique.[80] In this study, the Cnf/thioamide FRET pair was studied to determine the N- and C-termini orientations in α-synuclein.

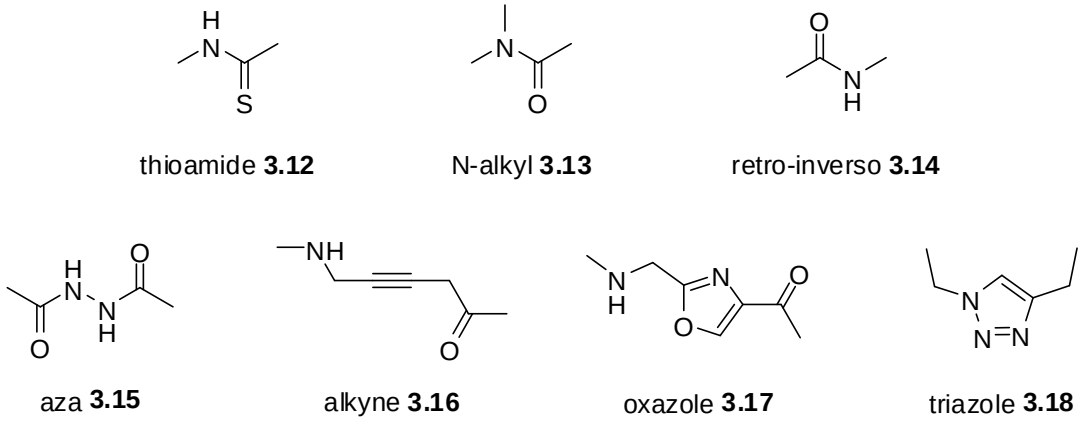


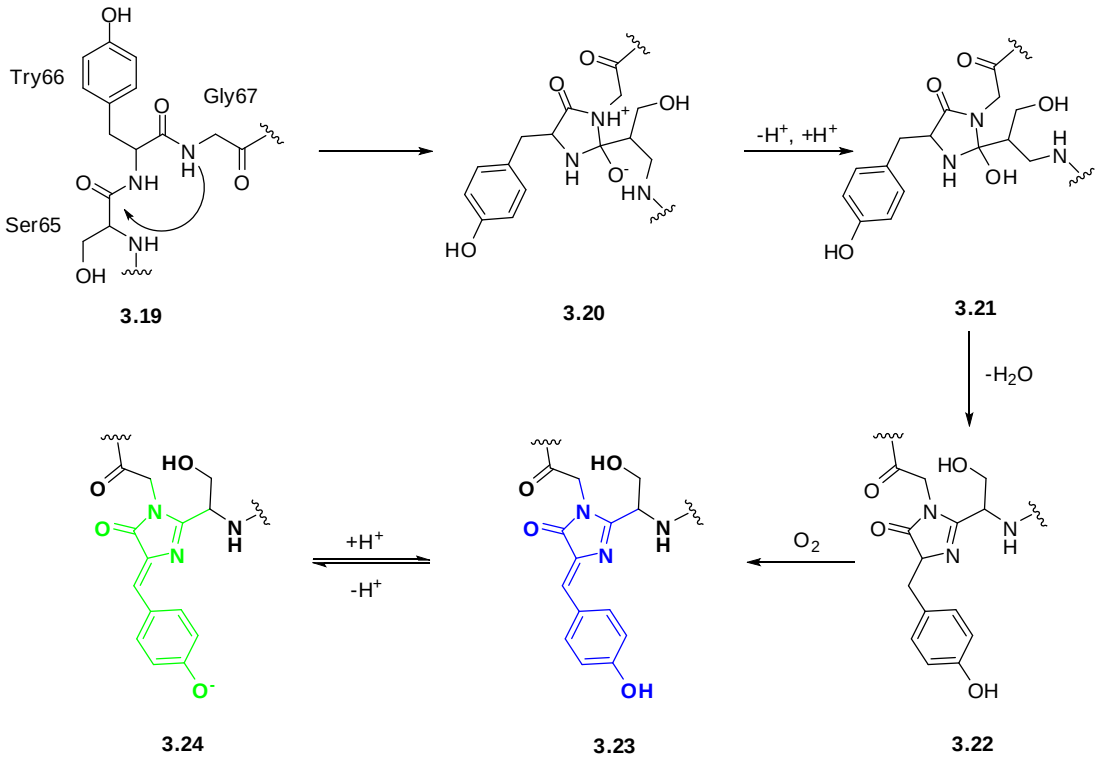
Figure 3.4. Some common peptide backbone modifications.[79]

Peptides exhibiting important biological activities have been modified by incorporating *N*-methyl amino acids (**3.13**, Figure 3.4), which improve the pharmacological properties of these peptides, such as metabolic stability, enhanced potency and bioavailability.[81] A good example of an *N*-methylated peptide is cyclosporine A having seven *N*-methylated amino acids; it is one of the most successful drugs used as immunosuppressant in organ transplantations.[81] The introduction of a triazole group (**3.18**, Figure 3.4) in the backbone of a peptide enabled the formation of β-sheets in small peptides, where a turn was induced by the triazole moiety.[82]

3.1.3. Green fluorescent protein (GFP)

After the discovery of green fluorescent protein (GFP) from the jellyfish *Aequorea victoria* in 1962 by Shimomura et al.,[83] interest in GFP was limited to studies of the luminescence of marine creatures. After its cloning in 1992[84] and its utilization as a marker for gene expression in 1994,[85] it became a widely used tool as a reporter for structures and processes in molecular and cell biology.[86] Amino acid residues in the proteins belonging to the GFP family, fold into a β-barrel structure made up of 11 β-sheets that enclose an internal distorted helix.[87] In *Aequorea victoria* GFP, three amino acid residues of the helix at positions 65-67 undergo a unique posttranslational modification to form the chromophore (Scheme 3.1).[88] The resulting chromophore is located almost at the center of the barrel and therefore is well shielded from solvent contacts.[87] Amino acids Tyr66 and Gly67 are conserved among all natural GFP-like proteins, whereas the amino acid residue at position 65 can vary.[86]

Scheme 3.1. Proposed mechanism for the chromophore formation in GFP.[88]



There has been a steady effort to develop new fluorescent proteins (FP) with improved functions and novel characteristics. Diversity is mainly achieved by site-directed mutagenesis which has produced a broad spectrum of GFP-like proteins.[86] The synthesis and spectral properties of FP chromophores have been studied in detail,[89] yet there is no report of the incorporation of a synthetic FP chromophore or a chromophore mimic into the backbone of proteins using *in vitro* or *in vivo* techniques.

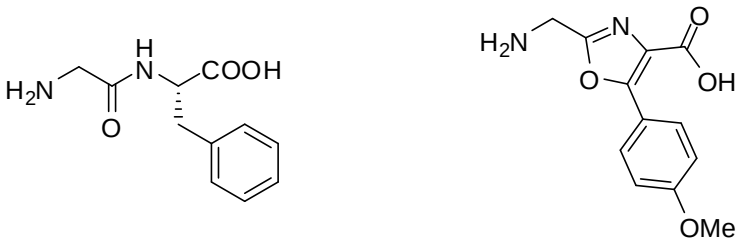
3.1.4. Biosynthesis of proteins having peptidomimetics

Synthetic and semi-synthetic strategies have enabled the incorporation of peptidomimetics into peptides and proteins.[76,77,80] However, the use of such strategies is limited by the size of the polypeptide accessible or control over the site of insertion of

peptidomimetics into a protein. The incorporation of peptidomimetics into proteins using *in vitro* or *in vivo* translation methods has not been reported to date. The architecture of the wild-type ribosome is biased towards the incorporation of α-L-amino acids into proteins.[3,42] Engineering proteins with peptidomimetics, using *in vitro* or *in vivo* methods, has the potential to open new doors for creating interesting biopolymers. Therapeutic proteins having peptidomimetics can enhance the proteolytic and thermal stability of those proteins, as well as increase potency, and improve target specificity and bioavailability. The inability of the wild-type ribosome to incorporate peptidomimetics into proteins, including the ones discussed earlier in this section, is the biggest challenge. Reengineering the architecture of the peptidyltransferase center (PTC) of the ribosome can enable the incorporation of peptidomimetics into proteins. Hecht and co-workers have successfully demonstrated that the suitable mutations in some regions of the 23S rRNA can afford modified ribosomes to facilitate the incorporation of non α-L-amino acids into proteins.[48-51]

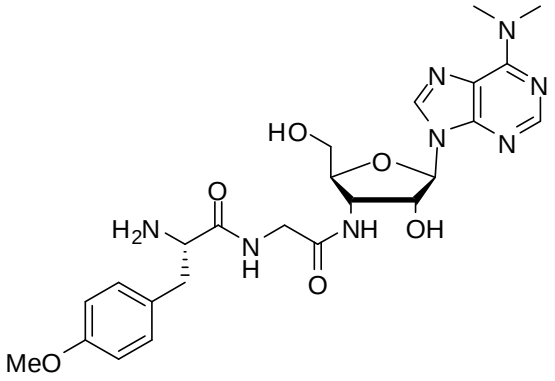
In this thesis, ribosomes having modifications in their 23S rRNA to enable the incorporation of the dipeptide glycylphenylalanine (**3.25**) and dipeptidomimetic **3.26** into DHFR are reported (Figure 3.5). Using puromycin derivative **1.3** (Figure 3.6) for the selection of *E. coli* colonies, six variants of ribosomes having modifications in two regions of the 23S rRNA (2057-2063 and 2502-2507) were selected. The selection experiments were carried out by Dr. Larisa Dedkova in a fashion similar to that reported earlier.[50] The selected modified ribosomes were evaluated for the suppression of the UAG codon at position 10 of DHFR mRNA by glycylphenylalanyl-$tRNA_{CUA}$ (**3.25b**, Scheme 3.2). Dipeptidomimetic **3.26** shown in Figure 3.5 is a fluorescent structural

analogue of the GFP chromophore **3.23** shown in Scheme 3.1. S-30 systems prepared from two different modified ribosomes were used to study the incorporation of **3.26** into position 10 of DHFR and the fluorescence spectra were obtained for the modified DHFR at different protein concentrations.



3.25 **3.26**

Figure 3.5. Structures of glycylphenylalanine (**3.25**) and dipeptidomimetic **3.26**.



1.3

Figure 3.6. Structure of puromycin derivative **1.3**.

3.2. Results

For the selection of ribosomes able to incorporate both α-L-amino acids and dipeptide or dipeptidomimetics, region 2057-2063 in the 23S rRNA was initially altered

to confer erythromycin resistance. Erythromycin resistance allowed the modified ribosomes to grow at the expense of the wild-type ribosomes having little resistance to the antibiotic. Five clones having different sequence in this 23S rRNA region were identified and were further mutagenized in one of the three other regions (2582-2588, 2502-2507 and 2496-2501) of the 23S rRNA. The *E. coli* colonies harboring modified ribosomal clones were grown in the presence of 100 μg/mL of the puromycin derivative **1.3** and plasmids from the colonies showing >50% inhibition of cell growth were isolated and sequenced (Table 3.1). Four clones (010309, 010310, 010326 and 010328, Table 3.1) were identified as mixtures of more than one clone as judged by the automatic DNA sequencing and were transformed again into *E. coli* cells to separate the clones. Clones having modifications in the region 2582-2588, or having the wild-type sequence in the second region were not selected for S-30 preparations (Table 3.1). It was demonstrated previously that modifications in the region 2582-2588 of the 23S rRNA were responsible for low fidelity of translation.[50] Of the remaining clones, nine mediated successful preparation of S-30 systems for *in vitro* translation experiments (Table 3.2). Eight of the clones shown in Table 3.2 were obtained after re-cloning the clone mixtures. Three of the ribosomal clones (010309R9, 010326R6 and 010328R4) had modified ribosomes with the same sequence (2057UGCGUGG2063 and 2502ACGAAG2507), while other two ribosomal clones (010326R1 and 010328R2) also shared same sequence in two modified regions (2057UGCGUGG2063 and 2502CUACAG2507). Therefore, six different variants of modified ribosomes were selected using puromycin derivative **1.3**. The selection experiments were carried out by Dr. Larisa Dedkova.

Table 3.1. Characterization of selected clones sensitive to puromycin derivative **1.3** and erythromycin.

Clones	Purity[a]	Sequence in regions of mutagenesis		Inhibition by	
		Region 1	Region 2	Puromycin derivative **1.3** (%)	Erythro-mycin (MIC, μg/mL)
010120	single	2057TGCGTGG2063	2582CCGAATC2588	65.4	12.5–6.25
010309	mix	2057TGCGTGG2063	2502CTACGG2507*	63.2	12.5–6.25
010310	mix	2057TGCGTGG2063	2502CGCAAT2507*	52.9	12.5–6.25
010322	single	2057TGCGTGG2063	2502ACGAAG2507	77	12.5–6.25
010326	mix	2057TGCGTGG2063	2502CTACAG2507*	50.4	12.5–6.25
010328	mix	2057TGCGTGG2063	2502ACGAAG2507*	77	12.5–6.25
030240	single	2057TTGGTCC2063	2496AGTGGT2501	56	6.25–3.12
030247	single	2057TTGGTCC2063	2496CACCTC2501**	64	6.25–3.12
030248	single	2057TTGGTCC2063	2496CACCTC2501**	61	6.25–3.12
040228	single	2057AGCGTG2063	2496CACCTC2501**	51	6.25–3.12
040322	single	2057AGCGTG2063	2502CTGCTT2507	54	6.25–3.12
040329	single	2057AGCGTG2063	2502TGGCAG2507	54	6.25–3.12
040338	single	2057AGCGTG2063	2502ATCAGG2507	56	6.25–3.12
070307	single	2057ATTCCGG2063	2502GATGTC2507**	77	6.25–3.12
080118	single	2057AGTGAGA2063	2582ATGGGCT2588	66.5	25–12.5
080337	single	2057AGTGAGA2063	2502ATCCGA2507	51	25–12.5
Wild-type		2057GAAAGAC2063	2496CACCTC2501 2502GATGTC2507 2582GATGTC2507	>1000	3.12–1.56

[a] Some colonies harbored more than one ribosomal clones. * Most dominant sequence as estimated by automatic sequencing. ** Same sequence as the wild-type ribosome in that region.

S-30 extracts, prepared from cultures having different modified ribosomes were tested for the incorporation of the dipeptide glycylphenylalanine (**3.25**) into position 10 of DHFR. S-30 systems having two different modified ribosomes were used for the incorporation of the fluorescent dipeptidomimetic **3.26** into DHFR. Glycylphenylalanyl-pdCpA (**3.25a**, Scheme 3.2) was synthesized by Dr. Rakesh Paul and the pdCpA

derivative **3.26a** (Scheme 3.3) activated with dipeptidomimetic **3.26** was prepared by Dr. Maninkandadas M. Madathil.

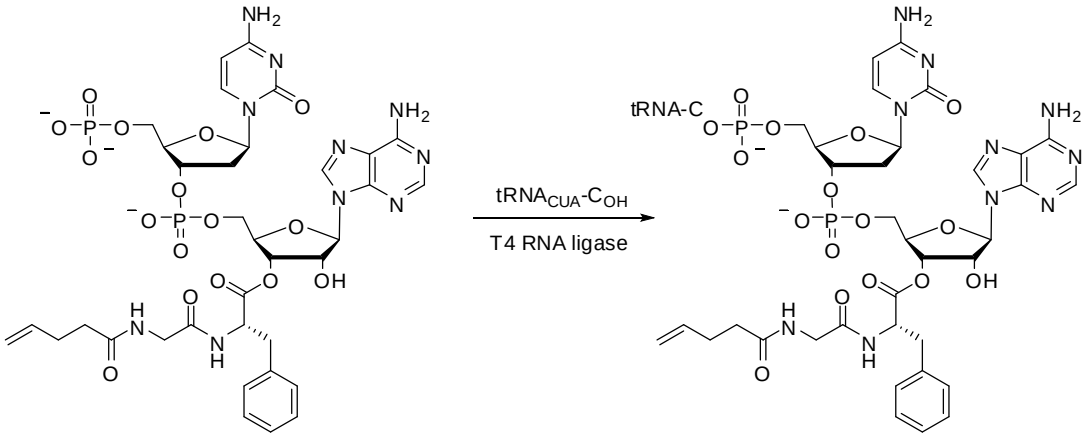
Table 3.2. 23S rRNA sequence modifications in the PTC for the clones used for S-30 preparations.

Clone	Sequence in 23S rRNA of modified ribosomes	
	Region 1 (2057-2063)	Region 2 (2502-2507)
010309R3	UGCGUGG	ACGAAG
010326R6	UGCGUGG	ACGAAG
010328R4	UGCGUGG	ACGAAG
010310R4	UGCGUGG	CGCACG
010326R5	UGCGUGG	CUAUGU
010310R1	UGCGUGG	CGCAAU
010328R2	UGCGUGG	CUACAG
010326R1	UGCGUGG	CUACAG
080337	AGUGAGA	AUCCGA
Wild-type	GAAAGAC	GAUGUC

3.2.1. Preparation of suppressor $tRNA_{CUA}$s activated with glycylphenylalanine (3.25) and dipeptidomimetic 3.26

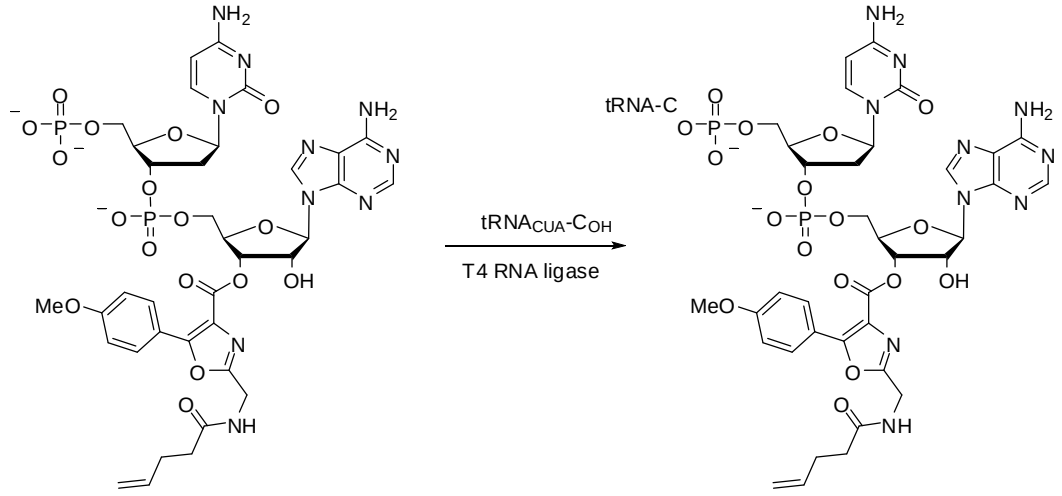
Activated pdCpA derivatives were ligated to an abbreviated 74-nucleotide $tRNA_{CUA}$-C_{OH} transcript using T4 RNA ligase (Schemes 3.2 and 3.3) and the two *N*-pentenoyl-aminoacyl -$tRNA_{CUA}$s (**3.25b** and **3.26b**) were prepared and isolated as described previously.[50] The *N*-pentenoyl protecting groups were removed by treatment with aqueous iodine for five minutes.[63,64] Removal of the protecting group was performed immediately prior to use of the misacylated tRNAs in the protein synthesis reaction. Each pdCpA derivative was ligated to abbreviated $tRNA_{CUA}$-C_{OH} with 100% efficiency as judged by 8% polyacrylamide-7 M urea gel electrophoresis (pH 5.0) analysis (Figure 3.7).[67]

Scheme 3.2. Preparation of glycylphenylalanyl-tRNA$_{CUA}$.



3.25a **3.25b**

Scheme 3.3. Preparation of tRNA$_{CUA}$ activated with the dipeptidomimetic **3.26**.



3.26a **3.26b**

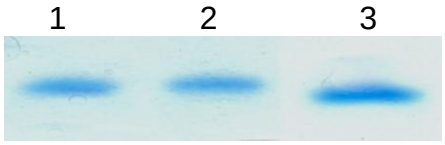


Figure 3.7. Preparation of misacylated tRNAs. Lane 1, glycylphenylalany-$tRNA_{CUA}$; lane 2, $tRNA_{CUA}$ activated with the dipeptidomimetic **3.26**; lane 3, nonacylated abbreviated $tRNA_{CUA}$.

3.2.2. *In vitro* translation using glycylphenylalanine (3.25) with wild-type ribosome

Wild-type ribosomes synthesize polypeptide chains by incorporating one α-L-amino acid at a time via one monoacylated tRNA for each corresponding codon in the mRNA transcript.[42] Ribosome mediated sequential incorporation of two amino acids for one codon using bisacylated tRNAs has not been reported to date. Therefore, the ability of the wild-type ribsome to enable the incorporation of glycylphenylalanine using DHFR mRNA having a UAG codon at position 10 in the presence of glycylphenylalanyl-$tRNA_{CUA}$ was studied. The suppression efficiency of glycylphenylalanyl-$tRNA_{CUA}$ was compared to phenylalanyl-$tRNA_{CUA}$ suppression of the UAG codon using an S-30 preparation having wild-type ribosomes (Figure 3.8). The suppression efficiency was <2% above the background for glycylphenylalanine. This data further proves that during translation, the wild-type ribosome cannot transfer a dipeptide from dipeptidyl-tRNA to a polypeptide chain.[90]

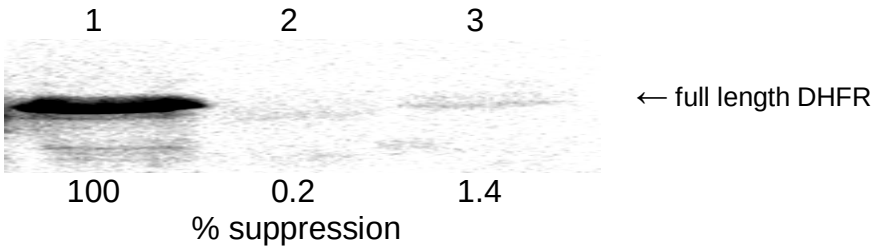


Figure 3.8. Translation of DHFR from wild-type (lane 1) and modified (lanes 2-3) (UAG codon in position 10) mRNA in the presence of different suppressor $tRNA_{CUA}$s. Lane 2, nonacylated $tRNA_{CUA}$; lane 3, $tRNA_{CUA}$ activated with glycylphenylalanine (**3.25**). The suppression efficiency relative to wild type is shown below each lane.

3.2.3. *In vitro* translation using glycylphenylalanine (3.25) and dipeptidomimetic 3.26 with the modified ribosomes

A modified DHFR contruct, having a TAG codon in position 10 (pETDH10 plasmid) was used for the incorporation of glycylphenylalanine (**3.25**) and the dipeptidomimetic **3.26**. Six different S-30 preparations having modified ribosomes produced different levels of full length DHFR from the wild-type gene relative to the S-30 system having the wild-type ribosome. Therefore, the suppression efficiencies were expressed relative to the suppression obtained using phenylalanyl-$tRNA_{CUA}$ (Table 3.3). As a negative control, DHFR synthesis in the presence of nonacylated-$tRNA_{CUA}$ was measured for each experiment. The amounts of DHFR produced were quantified with a phosphoimager, which monitored the incorporation of ^{35}S-methionine into DHFR.

Table 3.3. Incorporation of glycylphenylalanine (**3.25**) and dipeptidomimetic **3.26** into position 10 of *E. coli* DHFR by the use of S-30 systems having different modified ribosomes.

Clones	**Sequence in 23S rRNA of modified ribosomes**		**Suppression efficiency in different S-30 systems, having modified ribosomes (%)**[a]			
	Region 1 (2057-2063)	Region 2 (2502-2507)	Amino acids			
			Phe	-	**3.25**	**3.26**
010309R3 010326R6 010328R4	UGCGUGG	ACGAAG	100	0.8 ± 0.2	18 ± 4.0	15.3 ± 2.5
010310R1	UGCGUGG	CGCAAU	100	2.3 ± 1.0	10.2 ± 0.6	not tested
010310R4	UGCGUGG	CGCACG	100	2.3 ± 1.7	6.2 ± 1.6	not tested
010328R2 010326R1	UGCGUGG	CUACAG	100	0.8 ± 0.4	6.4 ± 1.1	not tested
010326R5	UGCGUGG	CUAUGU	100	1.6 ± 0.8	7.3 ± 2.3	not tested
080337	AGUGAGA	AUCCGA	100	2.2 ± 1.0	15.3 ± 6.1	14 ± 3.0

[a] Each number represents the average of three independent experiments ± SD.

Using glycylphenylalanyl-tRNA$_{CUA}$, S-30 preparations having the modified ribosomes with the sequence 2502ACGAAG2507 (clones 010309R3, 010328R4 and 010326R6) and 2502AUCCGA2507 (clone 080337) in the second region (Table 3.2), produced full length DHFR in ~18% and ~15% yields, respectively. Of these two modified ribosomes, the ribosome with sequence 2502ACGAAG2507 gave low levels of non specific readthrough (~0.8%). Three other modified ribosomes having sequence 2502CGCACG2507 (clones 010310R4, ~6% suppression yield), 2502CUACAG2507 (clones 010328R2 and 010326R1, ~6% suppression yield) and 2502CUAUGU2507 (clone 010326R5, ~7% suppression yield) in the second modified region produced relatively low levels of DHFR, while the modified ribosome with sequence

2502CGCAAU2507 incorporated glycylphenylalanine into DHFR in a moderate yield of ~10%. Figure 3.9 illustrates the incorporation of glycylphenylalanine into position 10 of DHFR using S-30 preparations having different modified ribosomes. The S-30 systems, prepared from the ribosomes having mutations of 2502ACGAAG2507 and 2502AUCCGA2507, were used for the incorporation of the dipeptidomimetic **3.26** into position 10 of DHFR and gave full length DHFR in ~15% and ~14% yields, respectively. Figure 3.10 illustrates the formation of full length DHFR using an S-30 system prepared from ribosomal clone 010328R4 in presence of glycylphenylalanyl-$tRNA_{CUA}$ (**3.25b**) and $tRNA_{CUA}$ (**3.26b)** activated with dipeptidomimetic **3.26**.

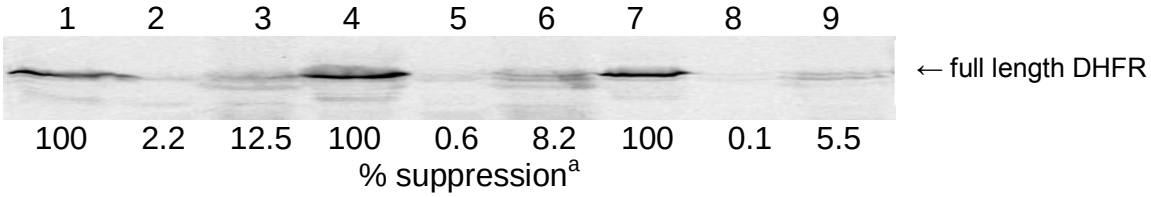


Figure 3.9. Translation of DHFR from a modified DHFR mRNA (UAG codon in position 10) by the use of S-30 systems prepared from clones 080337 (lanes 1-3), 010326R5 (lanes 4-6) and 010310R4 (lanes 7-9) in the presence of different suppressor tRNAs. Lanes 1, 4 and 7, phenylalanyl-$tRNA_{CUA}$; lanes 2, 5 and 8, nonacylated $tRNA_{CUA}$; lanes 3, 6 and 9, glycylphenylalanyl-$tRNA_{CUA}$. [a] The suppression efficiency values are one of the data set used for the average suppression efficiency values in Table 3.2.

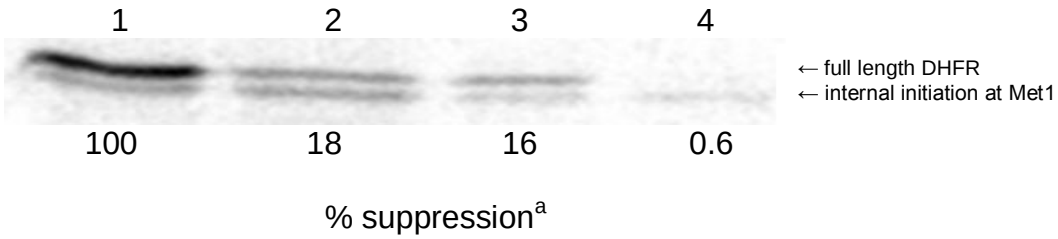


Figure 3.10. Translation of DHFR from a modified DHFR mRNA (UAG codon in position 10) by the use of an S-30 system prepared from clone 010328R4 in the presence of different suppressor tRNAs. Lane 1, phenylalanyl-$tRNA_{CUA}$; lane 2, glycylphenylalanyl-$tRNA_{CUA}$; lane 3, $tRNA_{CUA}$ activated with the dipeptide mimic; lane 4, nonacylated-tRNA. [a] The suppression efficiency values are one of the data set used for the average suppression efficiency values in Table 3.2.

3.2.4. Characterization of the incorporation of glycylphenylalanine (3.25) and dipeptidomimetic 3.26 into DHFR by MALDI-MS analysis of the peptides resulting from "in-gel" trypsin digestion

In order to provide direct evidence for the successful suppression of the UAG codon at position 10 of DHFR mRNA by glycylphenylalanyl-$tRNA_{CUA}$ or $tRNA_{CUA}$ activated with **3.26**, two modified DHFRs putatively bearing glycylphenylalanine (**3.25**) (modified DHFR 1) or dipeptidomimetic **3.26** (modified DHFR 2) were prepared on a larger scale for an "in-gel" trypsin digestion followed by MALDI-MS analysis.[68] As a control, DHFR V10F was also prepared in amounts similar to the two modified DHFRs. The modified DHFRs were purified by the use of Ni-NTA and DEAE-Sephadex chromatography followed by SDS-polyacrylamide gel electrophoresis.[51] Table 3.4 compares the estimated and observed molecular weights in MALDI-MS of the tryptic peptides from DHFR V10F and the modified DHFRs 1 and 2. The tryptic digest of

DHFR V10F gave an ion peak at *m/z* 1321.48 (Table 3.4 and Figure 3.11A) and confirmed the presence of peptide fragment (amino acids 1-12) having phenylalanine at position 10. For DHFR 1, a tryptic fragment encompassing amino acids 1-12 was anticipated to have a molecular mass of 1377 Da. As shown in Table 3.4 and Figure 3.11B, there was an ion peak at *m/z* 1377.33 consistent with the presence of glycylphenylalanine at position 10 of DHFR. As illustrated in Figure 3.11C (modified DHFR 2, Table 3.4), an ion peak at *m/z* 1404.28 generated by the peptide having amino acids 1-12 confirmed the incorporation of dipeptidomimetic **3.26** into position 10 of DHFR (estimated value 1404 Da). The other six tryptic peptides encompassing amino acids 13-106 of DHFR were observed for DHFR V10F and the two modified DHFRs and gave the expected ion peaks in their respective MALDI-MS spectrum. In replicate experiments, the (large) C-terminal fragment was never observed for any of the DHFRs.

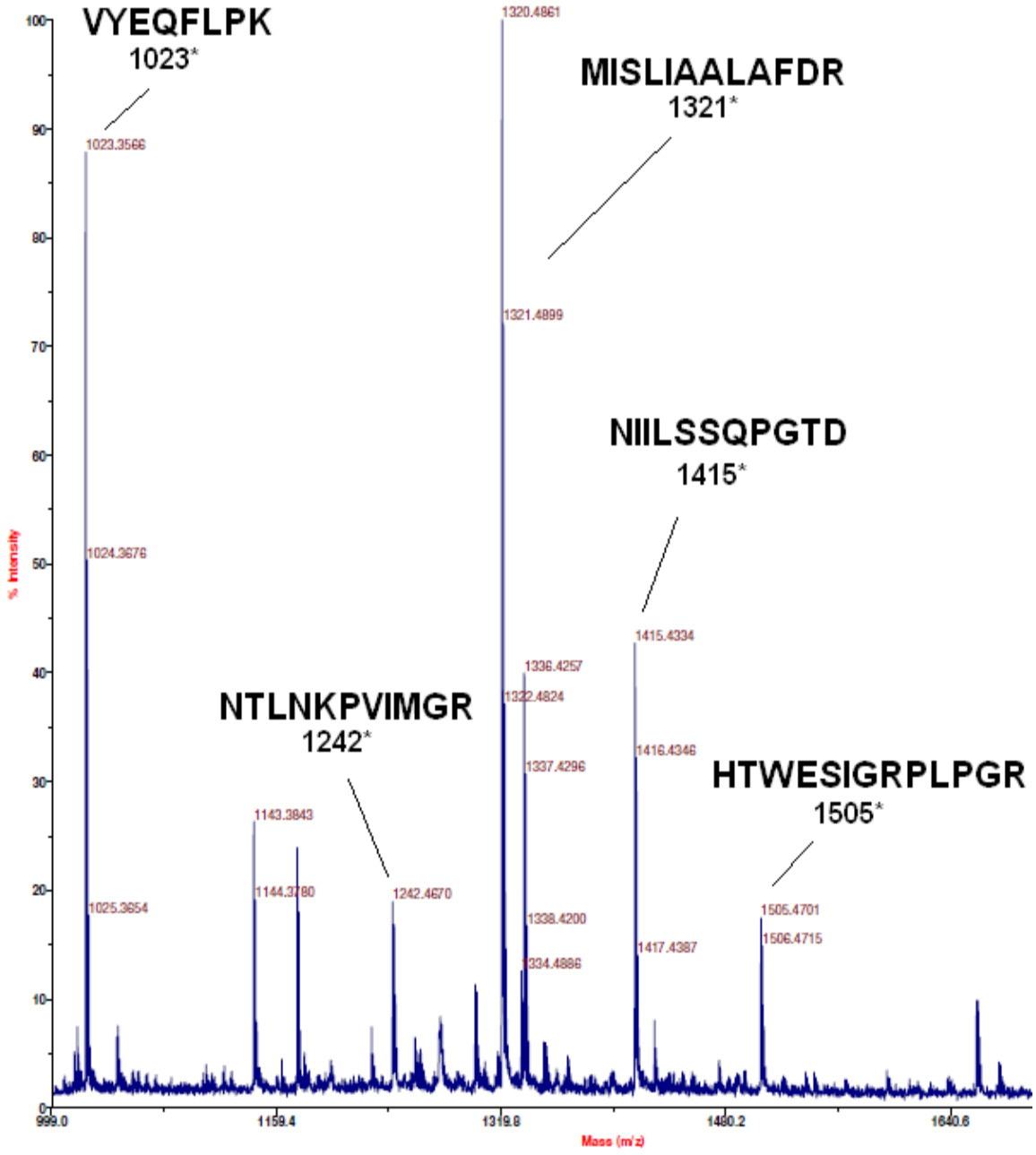
Table 3.4. MALDI-MS analysis of tryptic digests of the wild-type and modified DHFR samples.

Position	**Peptide sequence**	**MALDI-MS analysis, molecular mass, Da**					
		DHFR (V10F)		Modified DHFR 1		Modified DHFR 2	
		Est	MS	Est	MS	Est	MS
-9 - -1	MVHHHHHHR	1272	–	1272	–	1272	–
1-12	MISLIAALA**F**DR	**1321**	**1321.48**	**1377**	**1377.33**	**1404**	**1404.28**
13-32	VIGMENAMPWNLPADLAWFK	2304	2304.43	2304	2303.13	2304	2303.19
34-44	NTLNK PVIMGR	1242	1242.46	1242	1242.29	1242	1242.41
45-57	HTWESIGRPLPGR	1506	1506.47	1506	1506.26	1506	1506.33
59-71	NIILSSQPGTDDR	1415	1415.43	1415	1415.16	1415	1415.08
72-76	VTWVK	632	632.28	632	634.21	632	634.28
77-98	SVDEAIAACGDVPEIMVIGGGR	2159	2159.40	2159	2159.41	2159	2159.12
99-106	VYEQFLPK	1023	1023.35	1023	1023.18	1023	1023.22

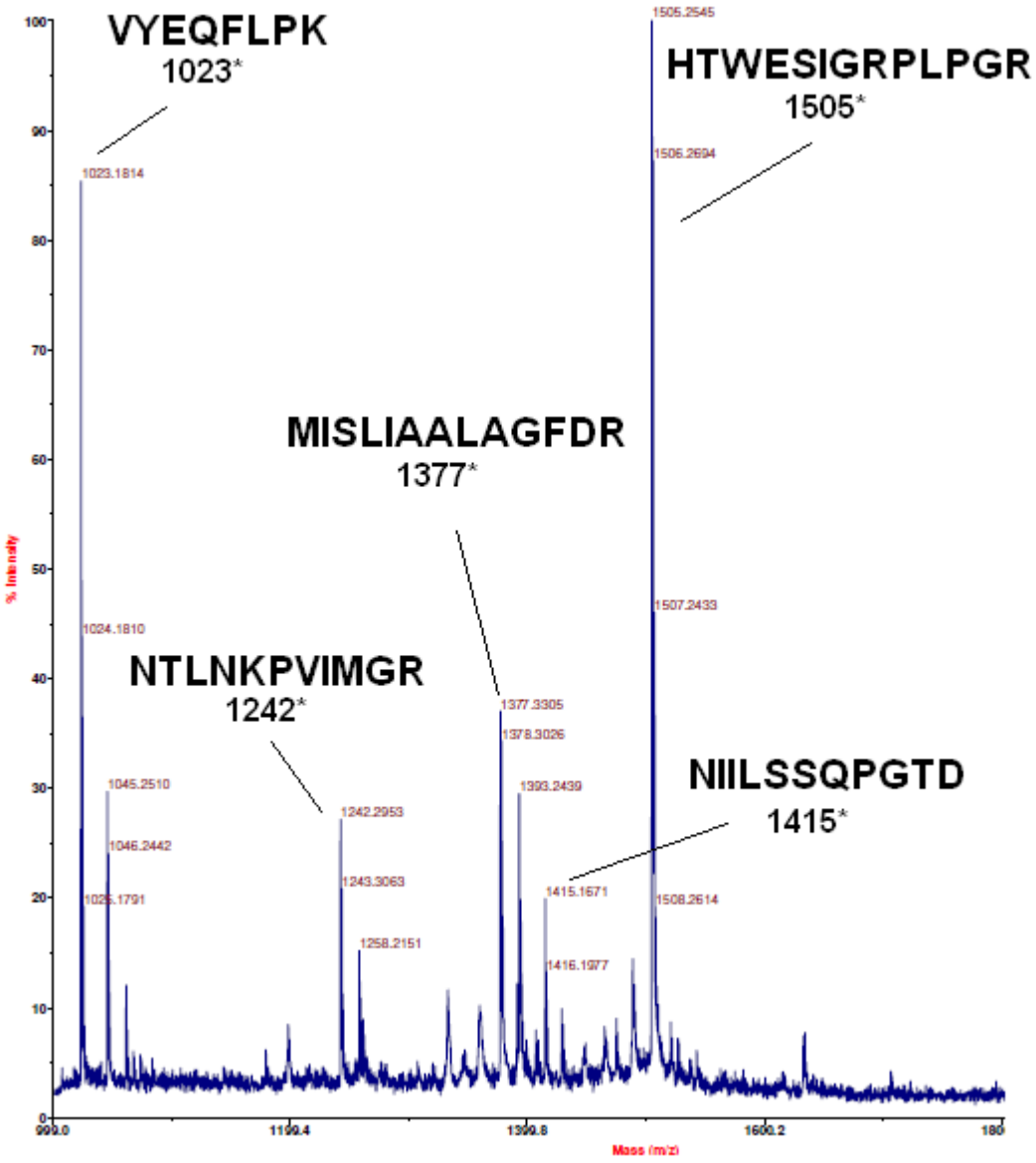
DHFR V10F, phenylalanine at position 10 of DHFR; modified DHFR 1, glycylphenylalanine at position 10 of DHFR; modified DHFR 2, dipeptidomimetic **3.26** atposition 10 of DHFR. The values in boldface reflect the presence of phenylalanine (for

DHFR V10F), glycylphenylalanine (for modified DHFR 1) and dipeptidomimetic (for modified DHFR 2). Position 10 is indicated in red.

A.



B.


VYEQFLPK
1023*
HTWESIGRPLPGR
1505*
MISLIAALAGFDR
1377*
NTLNKPVIMGR
1242*
NIILSSQPGTD
1415*
1023.1814
1024.1810
1045.2510
1046.2442
1025.1791
1242.2953
1243.3063
1258.2151
1377.3305
1378.3026
1393.2439
1415.1671
1416.1977
1505.2545
1506.2694
1507.2433
1508.2614
% Intensity
Mass (m/z)

C.

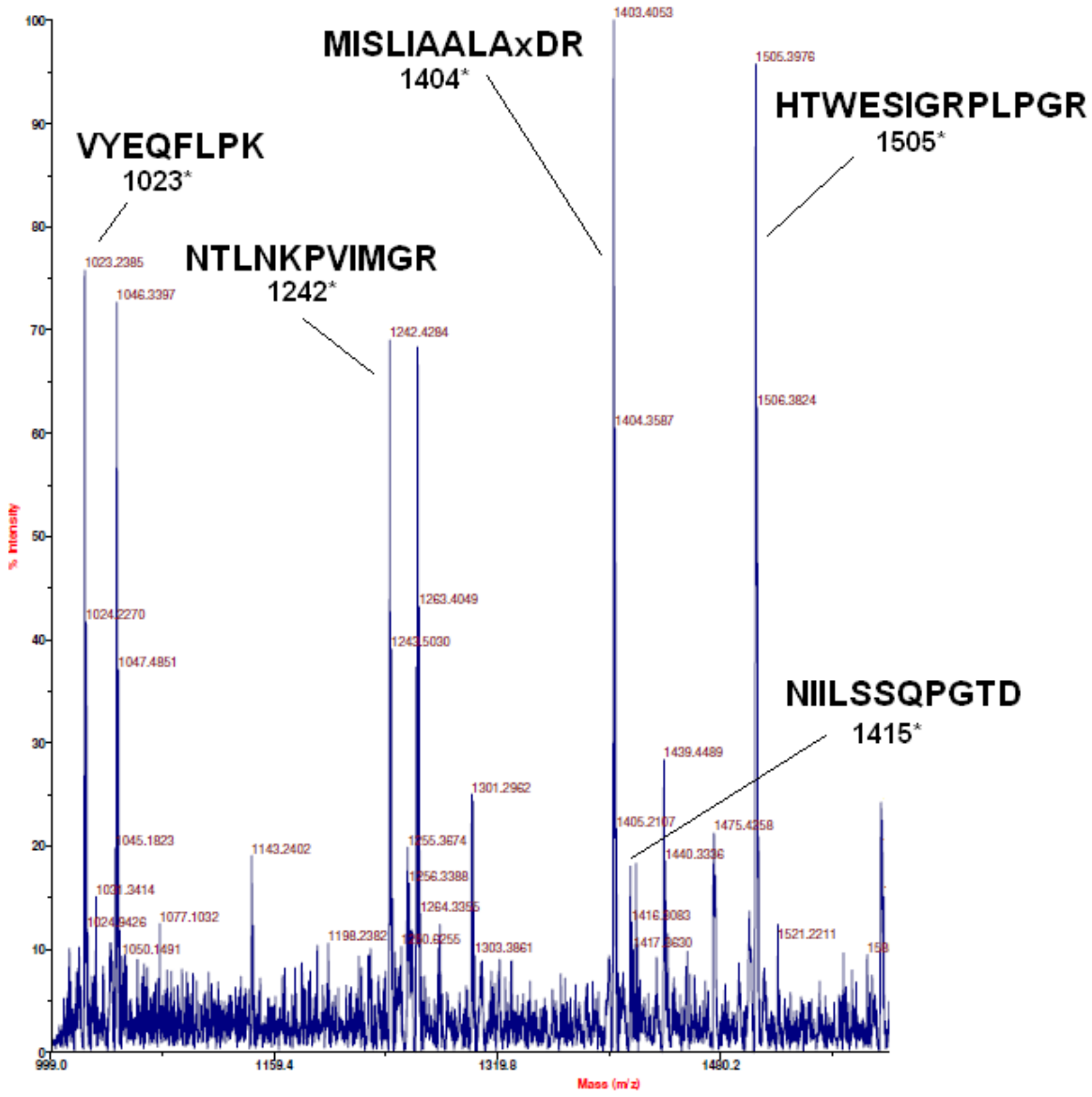


Figure 3.11. MALDI-MS of tryptic fragments of DHFR V10F (**A**); modified DHFR 1 (**B**), having glycylphenylalanine at position 10; modified DHFR 2 (**C**), having dipeptidomimetic **3.26** denoted as 'x' at position 10. Mass range 1000-1800 Da (* = estimated value in Da).

3.2.5. Fluorescence emission spectra of wild-type DHFR and modified DHFR having dipeptidomimetic 3.26 at position 10

Dipeptidomimetic **3.26** is a highly fluorescent structural analogue of the GFP chromophore **3.23** (Scheme 3.1). Figure 3.12 shows the fluorescence emission spectrum of **3.26** in water at four different concentrations. The absorption maximum for **3.26** was at 302 nm, therefore, this wavelength was chosen as excitation wavelength. Dipeptidomimetic **3.26** gave a fluorescence emission maximum at ~403 nm.

The fluorescence emission spectra of wild-type DHFR and modified DHFR having **3.26** at position 10 were compared (Figures 3.13) at 302 nm excitation wavelength. For both the wild-type and modified DHFRs, fluorescence spectra were recorded at three different protein concentrations. The protein concentrations were

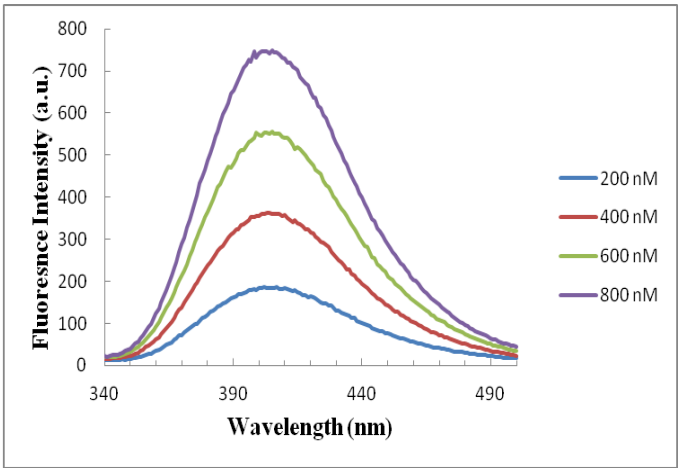
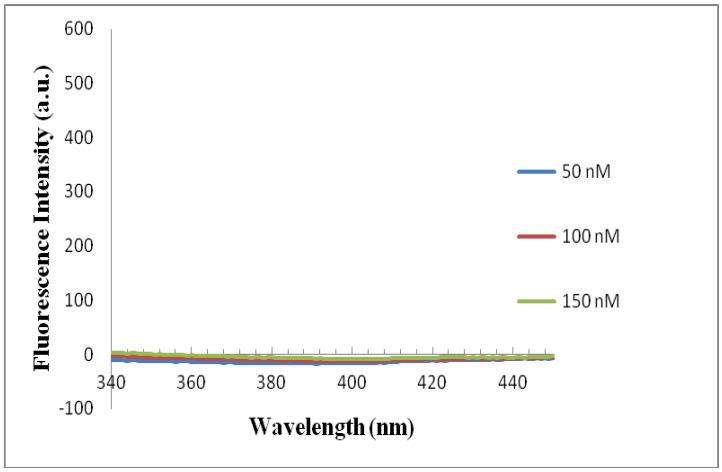


Figure 3.12. Fluorescence emission spectrum of dipeptidomimetic **3.26**. The fluorescence emission was monitored from 340-500 nm following irradiation at 302 nm. The sample concentration range was 200-800 nM in water.

A.



B.

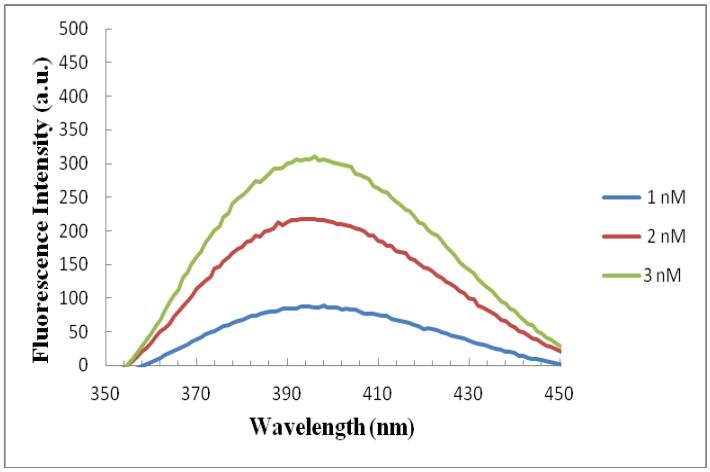


Figure 3.13. Fluorescence emission spectrum of the wild-type (**A**) and modified DHFR containing dipeptidomimetic **3.26** at position 10 (**B**). The fluorescence emission was monitored from 350-450 nm following irradiation at 302 nm. For wild-type DHFR (**A**), the sample concentration range was 50-150 nM and for modified DHFR having **3.26** at position 10 (**B**), the sample concentration range was 1-3 nM.

calculated following a standard BSA assay; the intensities of Coomassie Brialliant Blue staining of wild-type or modified DHFR samples were compared with a BSA standard of known concentration in an SDS-PAGE experiement. When excited at 302 nm, the modified DHFR had a fluorescence emission maximum at ~395 nm (Figure 3.13B) whereas no detectable fluorescence was observed for wild-type DHFR at a protein concentration 50 times greater (Figure 3.13A). Figure 3.14 compares the fluorescence intensitity of modified DHFR having **3.26** at position 10 with that of dipeptidomimetic **3.26**. At 10 nM concentration in water, **3.26** did not show any measurable fluorescence whereas modified DHFR gave an intense fluorescence emission. The incorporation of **3.26** into DHFR enhanced the fluorescence by 100 times; as seen in Figure 3.14, **3.26** gave comparable fluorescence intensity to modified DHFR at 100 times higher concentration (2 nM vs 200 nM).

A.

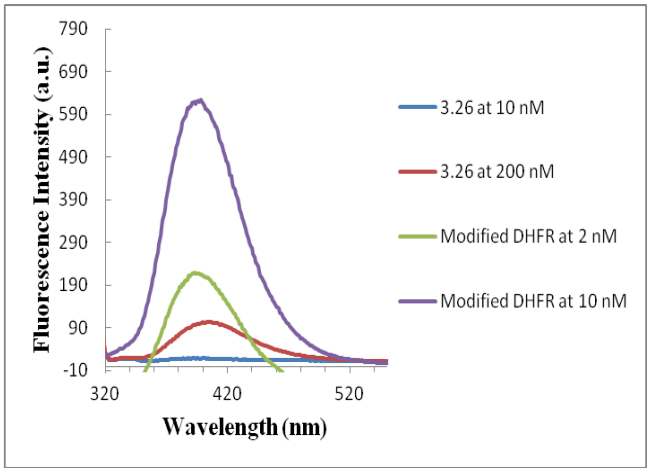


Figure 3.14. Relative fluorescence emission spectra of dipeptidomimetic **3.26**, at concentrations of 10 and 200 nM in water and the modified DHFR having **3.26** at

concentrations of 2 and 10 nM in water.The fluorescence emission was monitored from 320-550 nm following irradiation at 302 nm.

3.3. Discussion

In this study, glycylphenylalanyl-$tRNA_{CUA}$ and $tRNA_{CUA}$ activated with dipeptidomimetic **3.26** were prepared and the ability of these misacylated tRNAs to suppress an UAG codon at position 10 of a DHFR mRNA transcript by using the S-30 preparations having different modified ribosomes was studied (Table 3.2). The glycylphenylalanyl-pdCpA derivative was prepared by Dr. Rakesh Paul, while the pdCpA derivative activated with dipeptidomimetic **3.26** was synthesized by Dr. Maninkandadas M. Madathil. These pdCpA derivatives were ligated to abbreviated $tRNA_{CUA}$ using T4 RNA ligase and the full length $tRNA_{CUA}$s were obtained in quantitative yields as shown by analysis on an 8% polyacrylamide-7 M urea gel at pH 5.0 (Figure 3.7).

Table 3.3 summarizes the suppression yields of the UAG codon at position 10 of DHFR mRNA by the misacylated $tRNA_{CUA}$ activated with glycylphenylalanine (**3.25**) or dipeptidomimetic **3.26**. S-30 preparations having ribosomes with nucleotide sequences 2502ACGAAG2507 and 2502AUCCGA2507 gave the best results for the incorporation of glycylphenylalanine (~18% and ~15% suppression yields, respectively) into position 10 of DHFR. The same two S-30 systems produced full length DHFR in presence of $tRNA_{CUA}$ activated with the dipeptidomimetic **3.26** in ~15% and ~14% yields, respectively. The four other modified ribosomes produced moderate levels (~6% to ~10%

suppression yields) of full length DHFR using glycylphenylalanyl-$tRNA_{CUA}$ and DHFR mRNA having a UAG codon at position 10.

Table 3.2 shows the nucleotide sequence of the wild-type and nine modified ribosomal clones used for S-30 preparations. Clone 080337 had a different sequence in the first modified region (2057AGUGAGA2063), than the other eight ribosomal clones (2057UGCGUGG2063). Three clones, 010309R3, 010326R6 and 010328R4 shared the same nucleotide sequence in the second region (2502ACGAAG2507), while a different set of clones (010326R1 and 010328R2) also had the same sequence in the second region (2502CUACAG2507). Therefore, six different modified ribosomes were tested for the incorporation of glycylphenylalanine (**3.25**) into DHFR.

In the first modified region, the wild type sequence 2057GAAAGAC2063 was modified to either 2507UGCGUGG2063 (for five modified ribosomes) or 2507AGUGAGA2063 (for clone 080337). Interestingly, A2058G, A2060G and A2062G substitutions were present in all six modified ribosomes suggesting the possible importance of a purine nucleotide at these positions. Nucleotide C2063 partcipates in a non-conventional base pair with A2450 which is located in the vicinity of A2451.[69] A C2063G substitution was found to be present in five of the modified ribosomes, whereas a C2063A mutation was present in clone 080337. Nucleotide G2061 is a member of the hydrogen bonding network that orients A2451 in the peptidyltransferase center.[20] Either G2061U or G2061A substitution was present in the selected clones. The mutations at G2061 and C2063 might well play a role in the relaxation of the specificity of the PTC to accommodate a dipeptidyl-tRNA. C2063A and G2061U substitutions were also present in all the ribosomal clones enabling β-amino acid incorporation (Chapter 2).

In the second region (2502-2507), there was lower homology among the modified ribosomes. As discussed in the previous chapter, this region is responsible for stabilization of the 3'-ACC end of aminoacyl-tRNA in the A-site. Nucleotide A_{76} of the aminoacyl-tRNA interacts with wobble base pair U2506-G2583 in the wild-type ribosome.[20] For five of the modified ribosomes, U2506 was substituted by a purine nucleotide G or A, whereas only one modified ribosome had a U2506C substitution. Two modified ribosomes, which gave the best results for the incorporation of both glycylphenylalanine and dipeptidomimetic **3.26**, had G2502A substitutions in the second region. As shown in Chapter 2, clones 0403x4 and 040217 were efficient in incorporating β-amino acids and these two clones also had the same G2502A substitution.

Modified DHFRs having glycylphenylalanine (modified DHFR 1) and dipeptidomimetic **3.26** (modified DHFR 2) at position 10 were prepared on a larger scale for an "in-gel" trypsin digestion followed by MALDI-MS analysis. As a control, DHFR V10F was also prepared in a similar amount as the modified DHFRs. For modified DHFRs having β-amino acids **2.3** and **2.5**, the tryptic peptide, encompassing amino acids -9 to 12, was never observed in MALDI. We speculate that the six-histidine tag attached at the N-terminal of this peptide causes aggregation of the peptide. Therefore, a different DHFR construct having an arginine residue at -1 position was used for the preparation of the modified DHFR having glycylphenylalanine. This enabled hydrolysis at position -1 by trypsin, and removal of the six-histidine tag to generate a peptide with amino acids 1-12.

The tryptic digest of the DHFR V10F gave an ion peak at *m/z* 1321.48 (Table 3.4 and Figure 3.11A) which corresponded well with the estimated molecular mass (1321

Da) for the peptide fragment (amino acids 1-12) having phenylalanine at position 10. In the case of the modified DHFR 1 putatively containing glycylphenylalanine at position 10, a tryptic fragment encompassing amino acids 1-12 was anticipated to have a molecular mass of 1377 Da. As shown in Table 3.4 and Figure 3.11B, there was an ion peak at *m/z* 1377.33 consistent with the presence of the glycylphenylalanine at position 10 of DHFR.

The estimated mass of the peptide fragment (amino acids 1-12) having **3.26** at position 10 was 1404 Da. Figure 3.11C demonstrates the presence of **3.26** in DHFR where an ion peak at 1404.28 is consistent with the estimated value. The other six tryptic peptides encompassing amino acids 13-106 of DHFR were all observed in MALDI-MS spectrum of the wild-type and modified DHFRs. The (large) C-terminal fragment was never observed as was also true in the case of modified DHFRs having β-amino acids discussed in Chapter 2.

The structures of the GFP chromophore before (**3.23**) and after the deprotonation (**3.24**) of the phenolic hydroxyl group are shown in Scheme 3.1. Dipeptidomimetic **3.26** is a highly fluorescent structural analogue of protonated GFP chromophore **3.23** as shown by Figure 3.12. Therefore, the fluorescence emission spectrum of modified DHFR having **3.26** at position 10 was measured and compared with that of wild-type DHFR (Figures 3.13). Figure 3.13 shows the fluorescence emission spectrum following irradiation at 302 nm. Wild-type DHFR did not show any detectable fluorescence at high micro-molar protein concentration (Figures 3.13A). For modified DHFR having **3.26** at position 10, strong fluorescence emission was observered at 50 times lower protein concentration (Figures 3.13B). As evident from Figure 3.14, nearly 100-fold enhancement in

fluorescence intensity was achieved when dipeptidomimetic **3.26** was incorporated into position 10 of DHFR. There was no detectable fluorescence for **3.26** at 10 μM concentration in water. In contrast, modified DHFR having **3.26** was highly fluorescent as compared to **3.26** at 10 μM.

This study demonstrated the unprecedented incorporation of a dipeptide and dipeptidomimetic into a protein *in vitro* using modified ribosomes and mRNA having a UAG codon in presence of respective misacylated suppressor tRNAs. Novel or improved therapeutic proteins can now be prepared by the incorporation of peptidomimetics into such proteins using the modified ribosomes described in this study. Using a similar strategy described in this chapter, modified ribosomes can be selected for the incorporation of structurally varied peptidomimetics into proteins. Reverse turn mimics described in Section 3.1.1 and peptide isosteres in Section 3.1.2 can possibly be incorporated into specific positions in proteins using the modified ribosomes discussed in this study. This can create an important tool in the future to study structure-function relationships in proteins using peptidomimetics. The GFP chromophore **3.23** described in Section 3.1.3 can be incorporated into non-GFPs having similar structure to GFP to prepare artificial GFP-like proteins.

3.4. Experimental

3.4.1. General materials and methods

Tris, acrylamide, bis-acrylamide, urea, ammonium persulfate, *N*,*N*,*N'*,*N'*-tetramethylenediamine (TEMED), dihydrofolic acid, glycerol, ampicillin, pyruvate kinase, lysozyme, erythromycin**,** isopropyl-β-D-thiogalactopyranoside (IPTG), dithiothreitol (DTT) and 2-mercaptoethanol were purchased from Sigma Chemicals (St.

Louis, MO). ^{35}S-Methionine (10 μCi/μL) was obtained from Amersham (Pitscataway, NJ). BL-21 (DE-3) competent cells and T4 RNA ligase were from Promega (Madison, WI). Plasmid MaxiKit (Life Science Products, Inc., Frederick, CO) and GenElute^{TM}HP plasmid miniprep kit (Sigma) were used for plasmid purification.

Phosphorimager analysis was performed using a Molecular Dynamics 400E PhosphoImager equipped with ImageQuant version 3.2 software. Ultraviolet and visible spectral measurements were made using a Perkin-Elmer lambda 20 spectrophotometer. MALDI-TOF mass spectra were obtained at the Arizona State University CLAS High Resolution Mass Spectrometry Facility.

3.4.2. Preparation of aminoacyl-tRNA$_{CUA}$s (3.25b and 3.26b). The activation of suppressor tRNA$_{CUA}$s was carried out as described previously.[50] Briefly, 100-μL reaction mixtures of 100 mM Na Hepes, pH 7.5, contained 1.0 mM ATP, 15 mM $MgCl_2$, 100 μg of abbreviated suppressor tRNA$_{CUA}$-C$_{OH}$, 0.5 A_{260} unit of *N*-pentenoyl-protected □-aminoacyl-pdCpA, 15% DMSO, and 100 units of T4 RNA ligase. The reaction mixtures were incubated at 37 °C for 1 h, then quenched by the addition of 0.1 vol of 3 M NaOAc, pH 5.2. The N-protected aminoacylated tRNA was precipitated with 3 vol of cold ethanol. The efficiency of ligation was estimated by 8% polyarylamide–7 M urea gel electrophoresis (pH 5.0) (Figure 3.6).[67]

The N-pentenoyl-protected aminoacyl-tRNA$_{CUA}$s were deprotected by treatment with 5 mM aqueous I_2 at 25 °C for 15 min. The solution was centrifuged, and the supernatant was adjusted to 0.3 M NaOAc and treated with 3 vol of cold ethanol to precipitate the aminoacylated tRNA. The tRNA pellet was collected by centrifugation, washed with 70% aq EtOH, air dried and the dissolved in 50 μL of RNase free water.

3.4.3. Preparation of S-30 extracts from cells having modified ribosomes. Aliquots (5-10 μL) from liquid stocks of *E. coli* BL-21(DE-3) cells, harboring plasmids with a wild-type or modified rrnB gene, were placed on LB agar supplemented with 100 μg/mL of ampicillin and grown at 37 ºC for 16-18 h. One colony was picked from each agar plate and transferred into 3 mL of LB medium supplemented with 100 μg/mL of ampicillin and 0.5 mM IPTG. The cultures were grown at 37 ºC for 3-6 h in a thermostated shaker until OD_{600} ~ 0.15-0.3 was reached (about 3-5 h), diluted with LB medium supplemented with 100 μg/mL ampicillin, 1 mM IPTG and 3 μg/mL of erythromycin (for selectively enhancing the modified ribosome fraction) until OD_{600} 0.01 was reached, and then grown at 37 ºC for 12-18 h. The optimal concentration of the final cultures was OD_{600} 0.5-1.0. Cells were harvested by centrifugation (5000 × g, 4 ºC, 10 min), washed three times with S-30 buffer (1 mM Tris-OAc, pH 8.2, containing 1.4 mM $Mg(OAc)_2$, 6 mM KOAc and 0.1mM DTT) supplemented with β-mercaptoethanol (0.5 mL/L) and once with S-30 buffer having 0.05 mL/L β-mercaptoethanol. The weight of the wet pellet was estimated and 1.27 mL of S-30 buffer was added to suspend each 1 g of cells. The volume of the suspension was measured and used for estimating the amount of the other components. Pre-incubation mixture (0.3 mL) (0.29 M Tris, pH 8.2, containing 9 mM $Mg(OAc)_2$, 13 mM ATP, 84 mM phosphoenol pyruvate, 4.4 mM DTT and 5 μM amino acids mixture), 15 units of pyruvate kinase and 10 μg of lyzozyme were added per 1 mL of cell suspension and the resulting mixture was incubated at 37 ºC for 30 min. The incubation mixture was then frozen at – 80 ºC (~ 30 min), melted (37 ºC, 30 min), and again frozen and melted at room temperature (~ 30 min). Ethylene glycol tetraacetic acid (EGTA) was then added to 2.5 mM final concentration and the cells were

incubated at 37 ºC for 30 min and again frozen (– 80 ºC, 30 min). The frozen mixture was centrifuged (15,000 × g, 4 ºC, 1 h) and the supernatant was stored in aliquots at – 80 ºC.

3.4.4. *In vitro* protein translation. Protein translation reactions were carried out in 12-2000 μL of incubation mixture containing 0.2-0.4 μL/μL of S-30 system, 100 ng/μL of plasmid, 35 mM Tris acetate, pH 7.4, 190 mM potassium glutamate, 30 mM ammonium acetate, 2 mM DTT, 0.2 mg/mL total *E. coli* tRNA, 3.5% PEG 6000, 20 μg/mL folinic acid, 20 mM ATP and GTP, 5 mM CTP and UTP, 100 μM amino acids mixture, 0.5 μCi/μL of ^{35}S-methionine and 1 μg/mL rifampicin. In the case of plasmids having a gene with a TAG codon, a suppressor tRNA was added to a concentration of 0.3-0.5 μg/μL (for α-aminoacyl-tRNAs) and 0.6-1.0 μg/μL (for β-aminoacyl-tRNAs). Reactions were carried out at 37 °C for 1 h and terminated by chilling on ice. Aliquots from *in vitro* translation mixtures were analyzed by SDS-PAGE followed by quantification of the radioactive bands by phosphorimager analysis.

3.4.5. Purification of DHFR. Samples of DHFR, prepared during *in vitro* translation, were diluted with 50 mM Na phosphate, pH 8.0, containing 0.3 M NaCl and 0.1 mg/ml BSA and applied to a 50-μL Ni-NTA agarose column that had been equilibrated with the same buffer. The column was washed with three 500-μL portions of the same buffer and DHFR was eluted by washing with 150 μL of the same buffer also containing 250 mM imidazole.

A 50-μL column of DEAE-Sepharose was equilibrated with 500 μL of 10 mM Na phosphate, pH 7.4. Samples of DHFR purified by Ni-NTA chromatography were diluted 10-fold in the same buffer and applied to the resin. The column was washed with 500 μL

of the same buffer. DHFR was eluted with a NaCl step gradient (0.1; 0.2; 0.3; 0.4; 0.5 M) in 10 mM Na phosphate, pH 7.4. The fractions were analyzed on a 12% polyacrylamide gel, stained using Coomassie R-250, then destained in 20% ethanol–7% acetic acid. The fractions containing DHFR were combined, dialyzed against 50 mM Na phosphate, pH 7.4, and concentrated by ultrafiltration using a YM-10 filter (Amicon Ultra, Millipore Corp, Billerica, MA).

3.4.6. "In gel" trypsin digestion. Samples to be digested in the gel were run in 3-4 lanes of a 12% SDS-polyacrylamide gel, stained with Coomassie R-250 and destained until the background was clear. That area of the gel having the DHFR was cut from the gel and washed with 0.1 M ammonium bicarbonate (1 h, room temperature). The solution was discarded and 0.1-0.2 mL of 0.1 M ammonium bicarbonate and 10-30 μL of 0.045 mM DTT were added. Gel pieces were incubated at 60 ºC for 30 min, cooled to room temperature and incubated at room temperature for 30 min in the dark after the addition of 10-30 μL of 0.1 M iodoacetamide. Gel pieces were washed in 50% acetonitrile–0.1 M ammonium bicarbonate until they became colorless. After discarding the solution, the gel pieces were incubated in 0.1-0.2 mL of acetonitrile (10-20 min at room temperature) and, after removal of solvent, were re-swelled in 50-100 μL of 25 mM ammonium bicarbonate containing 0.02μg/μL trypsin. After incubation at 37 ºC for 4 h, the supernatant was removed to a new tube and the peptides were extracted with 60% acetonitrile–0.1% aq TFA (20 min at room temperature). The combined fractions were dried and reconstituted in 10 μL of 60% acetonitrile–0.1% aq TFA for each 1 μg of the protein (measured before trypsin digestion). Approximately, 5 μL of this solution was mixed thoroughly with 2.5 μL of the matrix solution (saturated solution of dihydroxybenzoic acid in 60%

acetonitrile–0.1% aq TFA). A 2 μL aliqout from the resulting solution was deposited on one of the wells in a 90-well MALDI-MS plate and dried.This was repeated twice and the plate was subjected to MALDI-MS analysis.

Chapter 4

SYNTHESIS AND BIOLOGICAL EVALUATION OF TOPOPYRONES

4.1. Introduction

The diameter of the nucleus in a typical human cell is approximately 10 µm. Human DNA contains 6 billion base pairs with each base pair about 0.34 nm in length, making the total length of DNA in a human cell approximately 2 m.[1] DNA in the cell nucleus is wrapped around histone proteins having positively charged side chains and is packaged as chromosomes in a space with approximately 10 µm diameter.[1] Processes such as DNA replication, DNA repair, DNA recombination and transcription require unpackaging of the DNA followed by strand separation. The process for generating single-stranded DNA is performed by motor proteins called DNA helicases.[91] During replication or transcription, strand separation generates DNA supercoiling in the flanking regions where two strands are separated by the polymerase–helicase complex. As a result, positive and negative supercoiling in front and behind, respectively, of the replication or transcription sites are generated.[92] Positive supercoiling can stall the replication (or transcription) complex and negative supercoiling can generate abnormal DNA structures.[93] DNA topoisomerases are the class of enzymes which relax supercoiled DNA by generating transient single or double stranded DNA breaks.[1] The first member of the topoisomerase class of enzymes was discovered by James C. Wang in 1971 and named "ω protein", which is *E. coli* topoisomerase I.[94]

DNA topoisomerases are divided into two main categories - type I and type II. The type I enzymes form a transient covalent bond with one DNA strand at a time. By contrast, a pair of strands in a DNA double helix is transiently broken in concert by type

II enzymes (Figure 4.1).[95] Another important difference is that the type II topoisomerases require ATP for the transport of one intact DNA double helix through another, whereas type I enzymes do not require ATP to function.[96,97] All topoisomerases

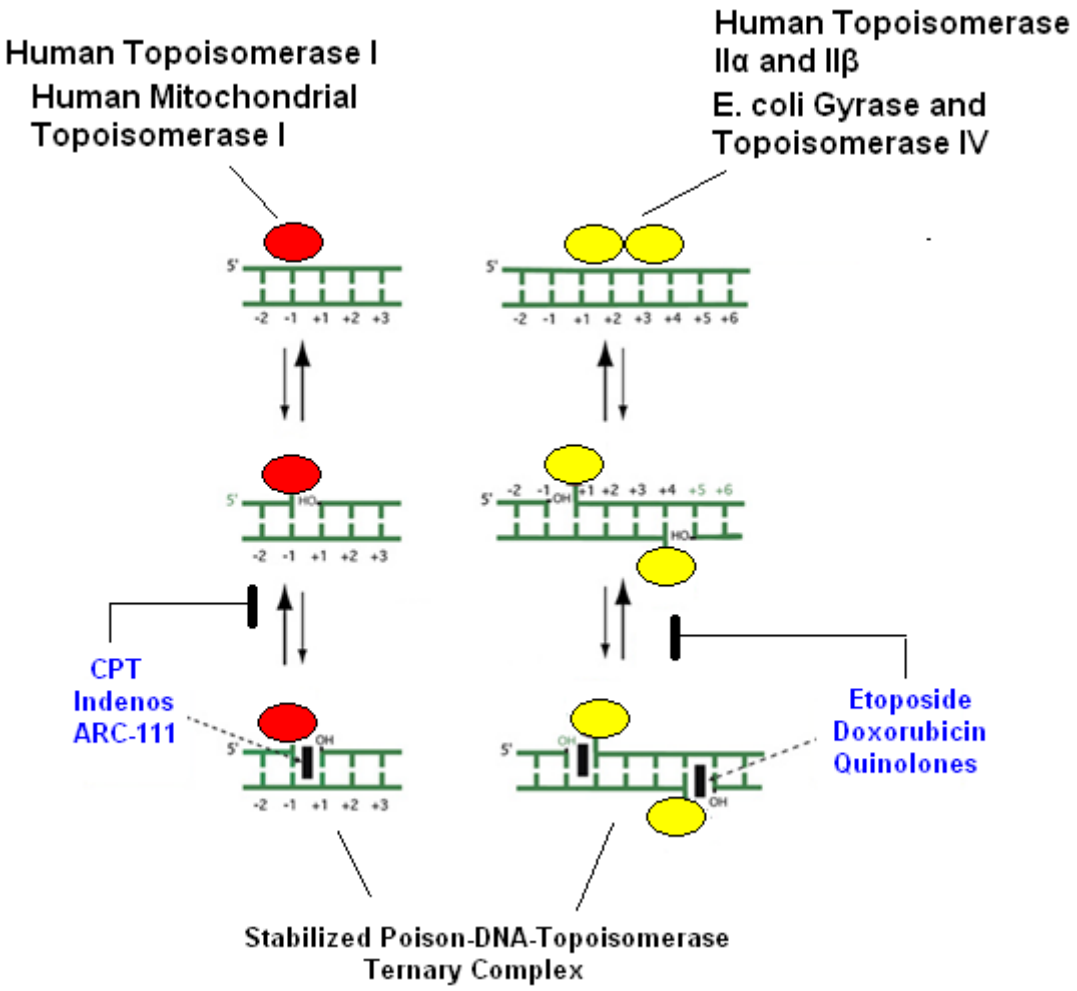


Figure 4.1. Stabilization of DNA–topoisomerase I or DNA–topoisomerase II covalent binary complexes by topoisomerase I or II poisons, respectively.[93]

have a conserved tyrosine residue which is responsible for the catalytic activity. The hydroxyl group of the active site tyrosine moiety mediates a nucleophilic attack on the

DNA phosphodiester bond, yielding a phosphorotyrosine covalent bond and breaking the DNA phosphodiester linkage (Figure 4.2).[93] Rejoining of the DNA strand occurs by another transesterfication reaction, where the hydroxyl group in the sugar moiety attacks the phosphorotyrosine linkage and displaces the tyrosine residue to reform the DNA backbone (Figure 4.2).[93]

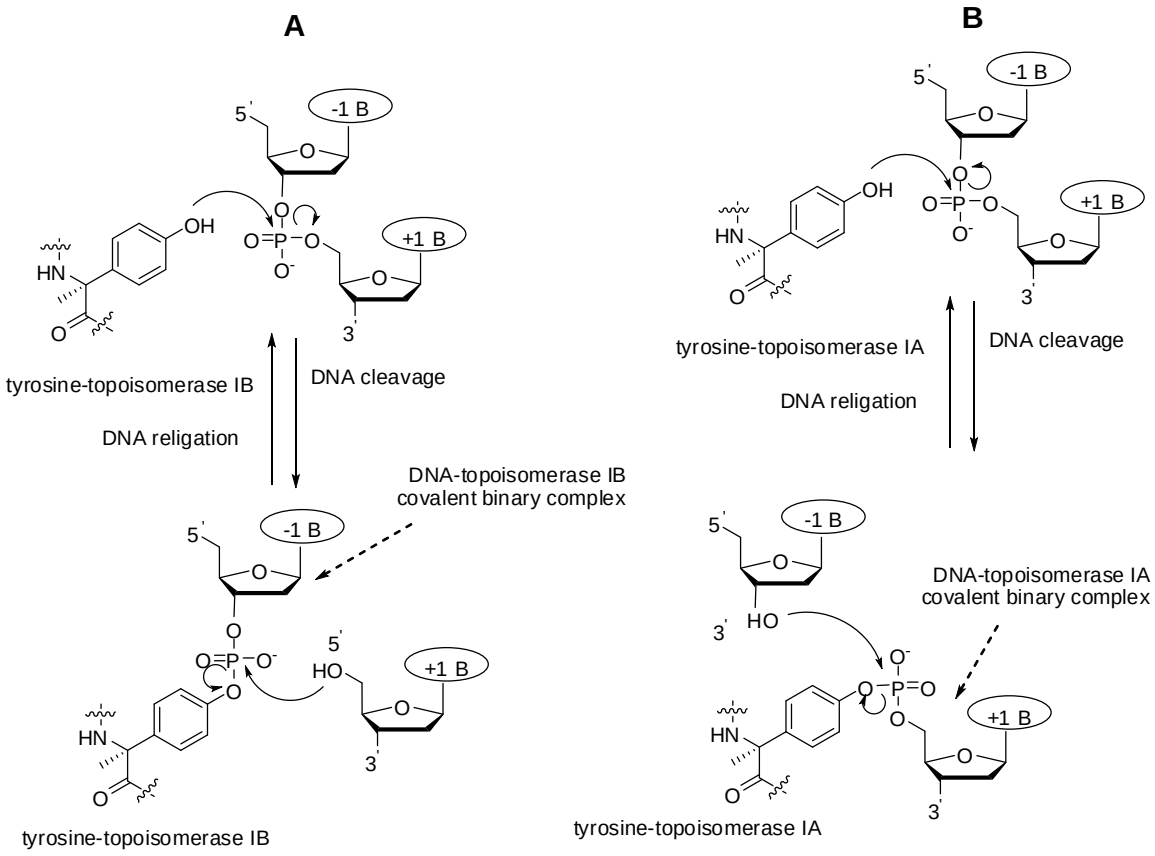


Figure 4.2. Mechanism of DNA relaxation by topoisomerases. (**A**) type IB (**B**) type IA (or type II).[93]

Type I topoisomerases are further divided into two sub-groups, type IA and IB. In general, type IA enzymes generate a covalently bound tyrosyl–DNA intermediate at the

5'-end of the break, whereas type IB enzymes covalently attach to the 3'-end of the break.[98] Type IA topoisomerases can relax only negative supercoils and require Mg^{2+} ions for DNA relaxation activity. Type IB enzymes can relax both negative and positive supercoils but do not require any metal ions to function.[99] There is another mechanistically important difference between type IA and IB enzymes; all enzymes in the type IA subfamily require a single-stranded region within the DNA substrate to function but there is no such requirement for type IB enzymes.[98,100] *E. coli* topo I and topo III, *S. cerevisiae* topo III, human topo IIIα and topo IIIβ are the few enzymes that belong to the class of type IA topoisomerases.[98]

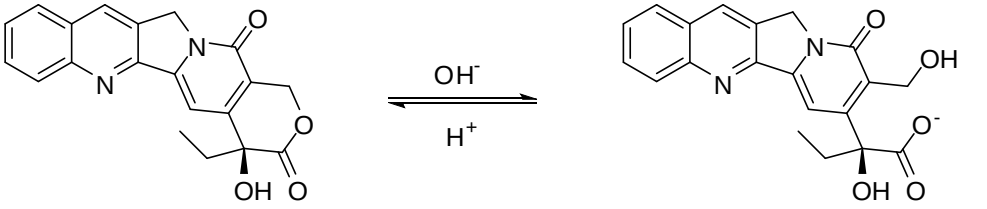
Type II topoisomerases can be divided into two broad classes: type IIA and type IIB. Type IIA topoisomerases include prokaryotic DNA gyrase, prokaryotic topoisomerase IV and eukaryotic topoisomerase II, while type IIB topoisomerases include TopoVI found in the archaeon *Sulfolobus shibatae*.[98] Prokaryotic type II topoisomerases contain two different subunits and are heterotetrameric in structure; on the other hand the eukaryotic topoisomerases are homodimers.[101] This subfamily of enzymes bind to duplex DNA and cleave the opposing strands with a four base stagger.[98] The catalytic mechanism involves covalent attachment of a tyrosine residue in each subunit to the 5'-end of the DNA to form a phosphorotyrosine bond (Figure 4.2).[102] The reactions require Mg^{2+} ions, whereas ATP hydrolysis is required for enzyme turnover and kinetics.[96] The religation reaction is another transesterification reaction where the hydroxyl group of the sugar moiety at the 3'-end of the DNA attacks the phosphorotyrosine bond and releases the covalently bound protein. DNA gyrase is a type

II enzyme found in all bacteria and is the only topoisomerase II enzyme capable of generating negative supercoils in plasmid DNA.[103]

Type I and II DNA topoisomerases have been firmly established as effective molecular targets for many antitumor drugs.[104] There are three type I topoisomerases that have been identified in human cells; topoisomerase I (topo I), topoisomerase IIIα (topo IIIα) and topoisomerase IIIβ (topo IIIβ). Topo I functions as a pivot in DNA replication, RNA transcription and chromosome condensation and segregation.[104] It has been demonstrated that topo I is essential in Drosophila and vertebrates but not in yeast.[105,106] Murine or human cells having no or low levels of topo I were found to have replication defects and exhibited altered gene expression.[107,108] Topo I has transcriptional activities independent of its DNA nicking-ligation activity such as promoter regulation[109] and phosphorylation/activation of splicing factors.[110] Human cells have two topoisomerase II isoenzymes; topoisomerase IIα (topo IIα) and topoisomerase IIβ (topo IIβ).[99] Topo IIα is crucial for cell growth and its levels are maximal in the late S/G2 phase of the cell cycle, where it plays an important role in chromosome condensation and segregation.[111] Despite having similar enzymatic functions *in vitro*, topo IIβ expressions levels are always constant throughout the cell cycle. Human topo IIα is mainly expressed in proliferating cells, and is required for the completion of mitosis.[112] By contrast, topo IIβ is required for normal cell development but is not required for proliferation.[113]

A "classical" topoisomerase I or II poison acts by misaligning the free hydroxyl group of the sugar moiety of DNA in the enyme–DNA covalent binary complex and prevents the reverse transesterfication reaction to religate the DNA (Figure 4.1). In this way topoisomerase poisons stabilize the normally transient enzyme–DNA covalent

intermediate and convert the binary covalent complex into a potential "suicide complex" which leads to preferential cell death in cancer cells.[93] A widely studied group of DNA topoisomerase I poisons are camptothecin (CPT) and its derivatives.[93] CPT was first isolated from *Camptotheca acuminate*, a Chinese tree, in 1966 and is shown to have exclusive topoisomerase I poison activity.[114] In animal studies, CPT showed potent antitumor activity against a broad spectrum of tumors which was accompanied by excessive toxicity during phase I/II trials.[115] CPT is inactive in its ring opened form shown in Figure 4.3. Human topo I was identified as a molecular target of CPT and subsequently, it was established as the sole antitumor target of CPT.[114,116] Two families of non-camptothecin topo I inhibitors have been to the clinical development stage, namely indenoisoquinolines (NSC706744, NSC725766 and NSC724998) and dibenzonaphthyridinones (Topovale, a Genzyme derivative).[93]



Camptothecin (active) Camptothecin carboxylate (inactive)

Figure 4.3. Hydrolysis of camptothecin to its inactive carboxylate form.[93]

Human topo II is a molecular target of many anti-cancer drugs including doxorubicin, mitoxantrone, m-AMSA and etoposide. These topo II poisons stabilize the covalent complex between the enzyme and DNA. Clinically useful topo II poisons are highly sensitive to ATP and the ATP/ADP ratio for their poisoning activity.[115] It has been

shown that the activity of etoposide increased by 50- to 100-fold in presence of ATP.[115] This reflects the need for ATP for type II topoisomerase enzymes to function. The anti-cancer activity of topo II poisons cannot be fully accounted for by drug-induced DNA cleavage. Both DNA replication and RNA transcription have been suggested to be involved in processing the covalent binary complex stabilized by topo II poisons.[117]

Selective inhibition of one type of topoisomerase can result in some negative consequences. Topoisomerase I inhibition can cause a compensatory increase in topoisomerase II levels, which may be associated with drug resistance against topoisomerase I inhibitors/poisons.[118] Simultaneous treatment with topo I poisons like CPT and topo II poisons such etoposide has been shown to be antagonistic.[117,119,120] Etoposide can potentially antagonize CPT by DNA replication inhibition, whereas CPT may decrease the effects of etoposide by inhibiting transcription. Also, combination therapies to target both topo I and topo II have shown significant toxicities, including severe to life threating neutopenia and anemia.[121] Recently, anticancer agents with dual topoisomerase I and II inhibitory activities have generated significant interest. A number of dual topoisomerase I and II inhibitors have been reported[118,121] whereas only saintopin[122] and topopyrones A to D,[123] have been reported as dual topoisomerase I and II poisons to date.

Topopyrones A, B, C and D were first isolated from culture broths of two fungi, *Phoma* sp. BAUA2861 and *Pencillium* sp. BAUA4206 by Kanai et al.[124] in 2000 and were reported to be human topoisomerase I poisons. Topopyrones, structurally similar to both doxorubicin and some pluramycins, are anthraquinone polyphenols having a fused 1,4-pyrone ring in an angular or linear orientation and may or may not have a chlorine

substitution at position 7 (Figure 4.4).[124] The total synthesis of all four topopyrones A to D was reported by Elban and Hecht,[125] followed by the report of topopyrones as dual topoisormerase I and II poisons by Khan et al. in 2008.[123] In the report by Khan et al., each of the topopyrones was able to stabilize enzyme–DNA covalent binary complexes for both topoisomerases I and II, thus showing dual poison activity. A covalent binary complex between 23 bp double stranded DNA, having a strong topoisomerase I cleavage site, and topoisomerase I was shown to stabilize at the same DNA sites as camptothecin. Alternatively, topopyrones A to D were compared with the known topoisomerase II poison, etoposide, for their ability to stabilize topoisomerase IIα–DNA binary complex. At the concentrations tested, all four topopyrones were shown to have potencies similar to etoposide.

With the encouraging biological results for topopyrones A to D and the synthetic route to access all four topopyrones in hand, an SAR study was the next logical step. We initially chose to replace the chlorine substituent with a bromine substituent at position 7 for topopyrones A and B to obtain derivatives **4.8** and **4.11** (Figure 4.4), respectively. The synthetic route followed to obtain the bromo-derivatives was similar as reported earlier for topopyrones A and B.[125] Along with the synthesis of topopyrones A to D, the synthesis of two novel topopyrone derivatives **4.9** and **4.10** was also reported.[125] For this study, these two derivatives were synthesized following a similar strategy as reported earlier.[125] X-ray crystallographic analysis of the binding of other topoisomerase I poisons to the DNA–topoisomerase I covalent binary complex indicates the H-bonding between the hydroxyl groups of the poisons and the binary complex.[127,128] Also, in a docking study done between topopyrones and topopisomerase I, a potential hydrogen bond between

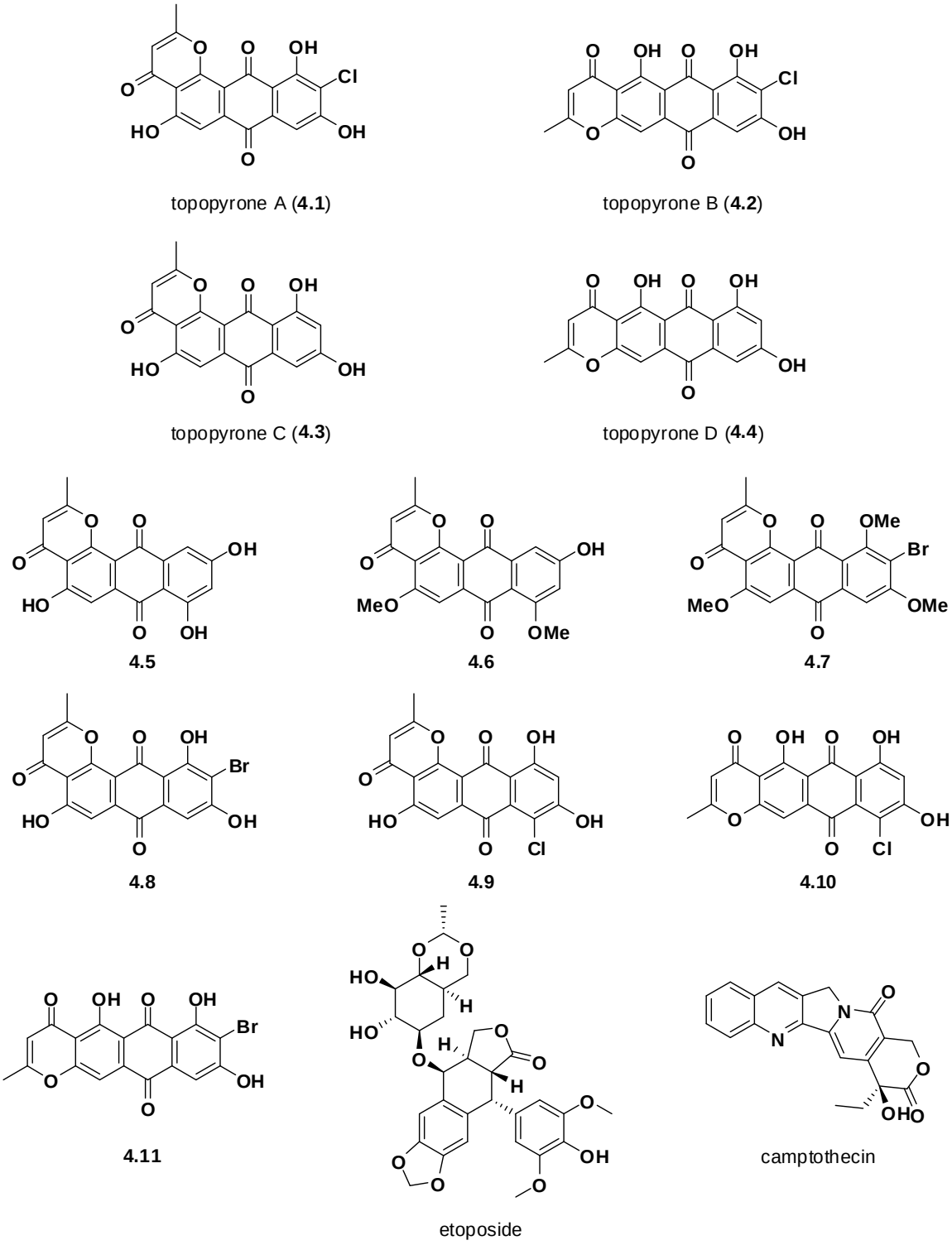


Figure 4.4. Structures of topopyrones A to D, topopyrone derivatives, etoposide and camptothecin.

the C6 OH group and Asn352 was observed for topopyrone C.[129] Therefore, to understand the role of the OH groups, compounds **4.5**, **4.6** and **4.7** were also synthesized.

4.2. Results

The total synthesis of topopyrones A, B, C, and D was reported by Elban and Hecht[125] in 2008 along with the synthesis of two novel topopyrone derivatives (**4.9** and **4.10**). In the present study, a similar synthetic approach was adopted to gain access to two new topopyrone C derivatives (**4.5** and **4.6**), two new topopyrone A derivatives (**4.7** and **4.8**) and one new topopyrone B derivative (**4.11**). Topopyrones **4.9** and **4.10** were obtained in analogy with the procedure reported previously.[125] Highlights of the synthetic route shown here are the Diels–Alder addition reaction (**4.20** and **4.21**, Scheme 4.1), Friedel-Crafts acylation reaction at the ortho position of phenol (**4.16**, Scheme 4.1), the pyrone ring formation using a strong base (**4.18,** Schemes 4.1 and **4.7**, Scheme 4.2), and regioselective mono-bromination (**4.23**, Scheme 4.2).

4.2.1. Synthesis of topopyrones 4.5 and 4.6

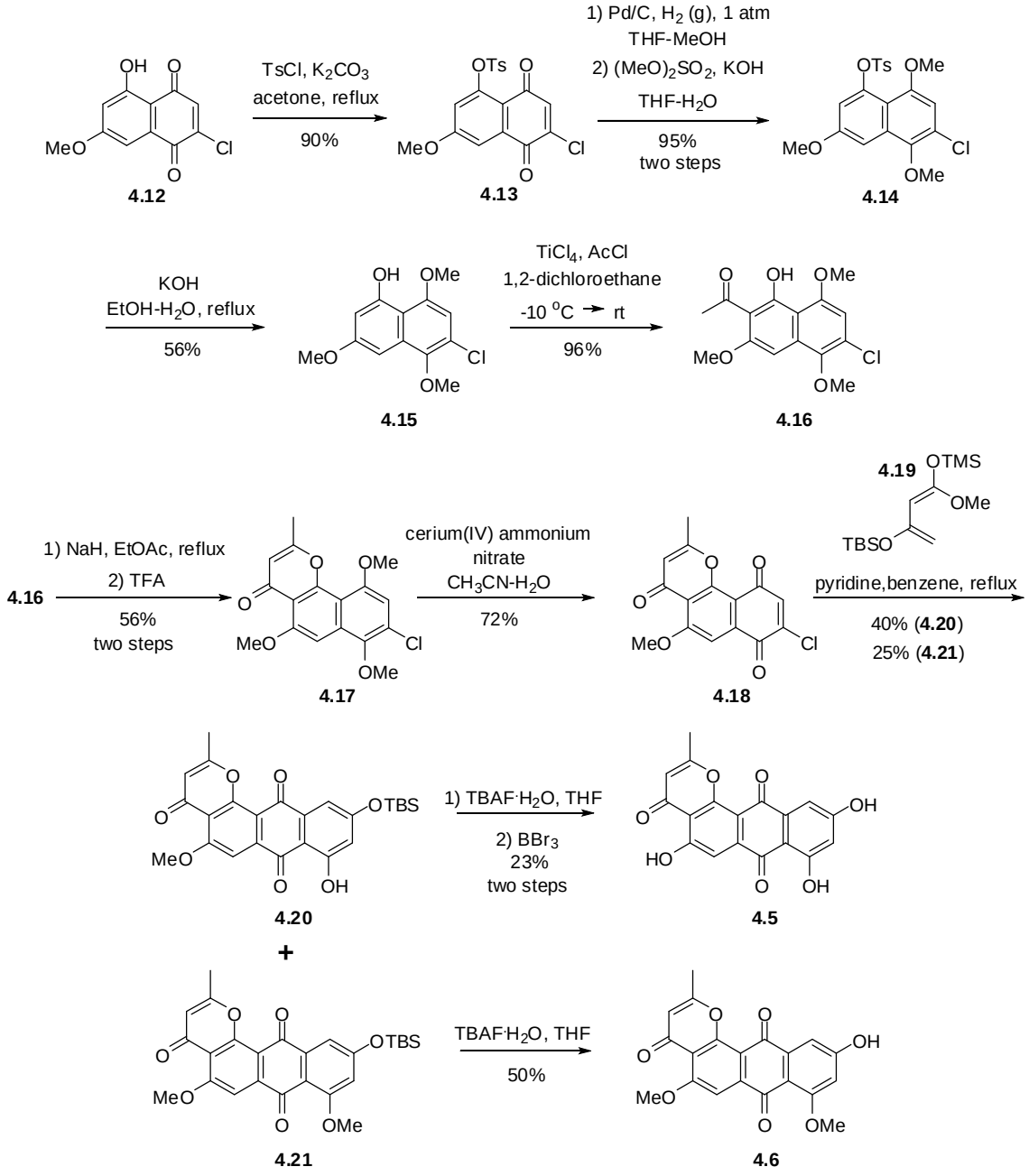
Diels–Alder addition of the previously reported 1,3-dimethoxy-1-trimethylsiloxy-1,3-butadiene[125,130] and commercially available 2,5-dichloro-1,4-benzoquinone in the presence of pyridine gave quinone **4.12**. Protection of the phenol using *p*-toluenesulfonyl chloride (*p*-TsCl) and K_2CO_3 in acetone at reflux afforded **4.13** in 90% yield. Pd/C catalyzed hydrogenation effected reduction of the quinone, which was immediately methylated using dimethyl sulfate and KOH to give **4.14** in 95% overall yield. Compound **4.15** was obtained by removal of the tosylate group from **4.14** (aq KOH at reflux). Although this reaction appeared to proceed cleanly as judged by silica gel TLC analysis, the yield of isolated product was only moderate (56%). An acetyl group was

introduced ortho to the free hydroxyl group using $TiCl_4$ and acetyl chloride in 1,2-dichloroethane at reflux, affording compound **4.16** in 96% yield.[131,132] The pyrone ring was introduced at this stage via base-catalyzed Claisen condensation; compound **4.16** was treated with NaH (60% dispersion in mineral oil) in freshly distilled ethyl acetate at reflux to afford the *β*-diketone. This intermediate was subsequently treated with TFA to obtain **4.17** (56%, two steps). Finally, compound **4.17** was oxidized to quinone **4.18** (72% yield) using cerium(IV) ammonium nitrate in CH_3CN. A second Diels−Alder addition reaction was used to obtain the precursors for topopyrones **4.5** and **4.6**. Thus quinone **4.18** was treated with 1-methoxy-1-trimethylsiloxy-3-tert-butyldimethylsiloxy-1,3-butadiene (**4.19**)[125] in the presence of pyridine to facilitate the aromatization of the intermediate Diels−Alder addition product. Compounds **4.20** and **4.21** were thereby obtained in 40% and 25% yields, respectively. Finally, compound **4.20** was treated with tetra-*n*-butylammonium fluoride hydrate ($TBAF{\cdot}H_2O$) to remove the tertbutyldimethylsilyl protecting group, followed by demethylation using 1 M BBr_3 in CH_2Cl_2. This afforded topopyrone C derivative **4.5** in 23% overall yield. Similarly, topopyrone C derivative **4.6** was obtained in 50% yield by the treatment of **4.21** with $TBAF{\cdot}H_2O$ in THF.

4.2.2. Synthesis of topopyrones 4.7, 4.8 and 4.11

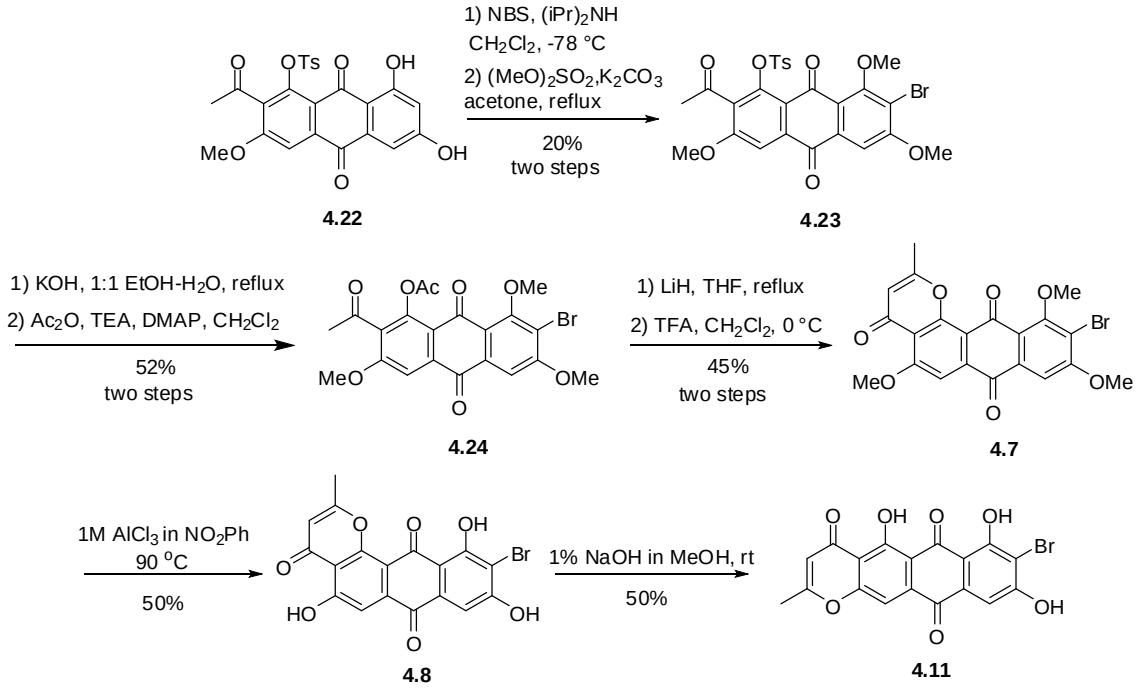
Monobrominated intermediate **4.23** was obtained in 20% overall yield by treating **4.22**[37] with N-bromosuccinimide (Scheme 4.2)[133,134] in the presence of $(iPr)_2NH$ at −78 °C in freshly distilled CH_2Cl_2, followed by methylation of the phenolic hydroxyl groups with dimethyl sulfate and K_2CO_3 in acetone. Removal of the tosylate group, followed by *O*-acetylation of the intermediate phenol afforded **4.24** (52%, two steps). This was achieved

Scheme 4.1. Synthesis of topopyrones **4.5** and **4.6**.[126]



using KOH in EtOH−H_2O at reflux followed by treatment with acetic anhydride in the presence of triethylamine and 4-*N*,*N*-dimethylaminopyridine.[125] Treatment of acetylated intermediate **4.24** with LiH in freshly distilled THF at reflux, followed by treatment with

Scheme 4.2. Synthesis of topopyrones **4.7**, **4.8** and **4.11**.[126]



CF_3COOH at 0 °C (Baker−Venkataraman rearrangement)[135,136] led to pyrone ring formation, affording topopyrone **4.7** in 45% yield for the two steps. Removal of the methyl groups using 1 M $AlCl_3$ in nitrobenzene at 90 °C gave brominated topopyrone A derivative **4.8** in 50% yield.[125] Brominated topopyrone B derivative **4.11** was obtained by the rearrangement of the pyrone ring in **4.8** using 1% NaOH in MeOH (50% yield).[124,125]

4.2.3. Biological evaluation of topopyrone derivatives

Biological evaluation of the topopyrones shown in Figure 4.4 was performed by Dr. Paul A. Zaleski.[126] Compounds **4.5**, **4.6**, **4.7** and **4.8** were analyzed for their ability to stabilize the DNA–topoisomerase I covalent binary complex by polyacrylamide gel electrophoresis (Table 4.1). In this experiment, the amount of DNA cleaved by 1 μM

CPT was arbitrarily assigned a value of 100% and the amount of DNA cleaved by each topopyrone was measured relative to that cleaved by CPT. Topopyrone **4.8** showed the best results of the four topopyrones studied, producing 25% DNA cleavage at 10 μM concentration whereas topopyrone **4.7**, a trimethoxy derivative of topopyrone **4.8**, failed to stabilize the DNA-topoisomerase I covalent binary complex at 1 or 10 μM concentration. Topopyronce C derivatives **4.5** and **4.6**, on the other hand, did not show much difference at 10 μM concentration (Table 4.1) and produced 11% and 10% DNA cleavage, respectively.

In a separate experiment, a nitrocellulose filter binding assay was performed to measure the amounts of DNA–topoisomerase I covalent binary complex stabilized by topopyrones **4.3**, **4.9** and **4.10** at 10 or 20 μM concentrations (Table 4.2). In the absence of any poison, topoisomerase I was covalently bound to 1.4 ± 0.5% of the linearized plasmid DNA substrate; in the presence of 0.5 μM CPT this increased to 39 ± 4.6%. Of the three topopyrones, compound **4.10** employed at 20 μM was as efficient as 0.5 μM CPT in stabilizing the covalent binary complex. Compound **4.9**, having pyrone ring in an angular position, did poorly in stabilizing the covalent binary complex. Only 5-6% DNA covalently bound to topoisomerase I was observed.

Stabilization of the DNA–topoisomerase II covalent binary complex by either etoposide or topopyrone **4.10** was also analyzed by a nitrocellulose filter binding assay. Of the two human type II topoisomerases, isoform IIα was used for this study as it was shown to be more relevant in proliferating cells.[137] In the nitrocellulose filter binding assay, the amount of covalent DNA–topoisomerase II binary complex was relatively high

Table 4.1. Topoisomerase I-mediated DNA cleavage by CPT and topopyrones on a 3′-^{32}P end labeled 23-base pair oligonucleotide substrate.[126]

topoisomerase I poison	% DNA cleavage[a]
DNA alone	1
Topoisomerase I alone	0
1 μM CPT	100
1 μM topopyrone **4.5**	6
10 μM topopyrone **4.5**	11
1 μM topopyrone **4.6**	3
10 μM topopyrone **4.6**	10
1 μM topopyrone **4.7**	1
10 μM topopyrone **4.7**	1
1 μM topopyrone **4.8**	2
10 μM topopyrone **4.8**	25

[a] Relative to CPT; determined by densitometric analysis.

This experiment was carried out by Dr. Paul A. Zaleski.

even in the absence of any poison. Etoposide at 0.5 μM concentration produced 34 ± 4.6% of the covalent binary complex. At 1 and 5 μM concentrations of **4.10**, a similar amount of the DNA was covalently bound to the enzyme.

Topopyrones **4.2**, **4.9** and **4.10** were also analyzed for cleavage site specificity in topopyrone–DNA–topoisomerase I ternary complex. These compounds were shown to stabilize topoisomerase I cleavage at sites similar to CPT along with a new 5'-TG_{72}-3' cleavage site in a 158-bp DNA duplex substrate.[126] Alternatively, topopyrones **4.2** and **4.11** stabilized topoisomerase II cleavage sites in 5'-^{32}P labeled 158 base pair DNA duplex similar to etoposide (5'-CA_{23}-3' and 5'-CG_{74}-3') but with varying efficiencies.

Table 4.2. Nitrocellulose filter binding of CPT or topopyrone stabilized DNA–topoisomerase covalent binary complexes.[126]

topoisomerase poisons	DNA covalently bound to topoisomerase (%)[a]
topoisomerase I	
none	1.4 ± 0.5
0.5 μM CPT	39 ± 4.6
10 μM topopyrone C (**4.3**)	16 ± 4.3
20 μM topopyrone C (**4.3**)	21 ± 7.8
10 μM topopyrone **4.9**	6.1 ± 2.4
20 μM topopyrone **4.9**	5.1 ± 4.4
10 μM topopyrone **4.10**	11 ± 3.6
20 μM topopyrone **4.10**	30 ± 5.2
topoisomerase IIα	
none	23 ± 2.7
0.5 μM etoposide	34 ± 4.6
1 μM topopyrone **4.10**	40 ± 13
5 μM topopyrone **4.10**	34 ± 2.5

[a] Topoisomerase mediated cleavage conditions described in experimental procedures; reaction mixtures contained 6 units of topoisomerase I or 5 units of topoisomerase IIα. Values represent the mean ± SD (n = 3).

This experiment was carried out by Dr. Paul A. Zaleski.

4.3. Discussion

Elban and Hecht[125] reported the total synthesis of topopyrones A to D (Figure 4.4) and two novel topopyrone derivatives (**4.9** and **4.10)**. At the time the topoyrones were discovered, they were reported to be only topoisomerase I poisons.[12436] In contrast, Khan et al.[123] found that all four topopyrones were dual topopisomerase I and II poisons. In this report each of the topopyrones was able to stabilize enzyme–DNA covalent binary complexes for both topoisomerases I and II, thus showing the dual poison activity. Covalent binary complex formation between a 23 bp double-stranded DNA, having a

strong topoisomerase I cleavage site, and topoisomerase I was shown to stabilize the same cleavage sites as camptothecin. On the other hand, topopyrones A to D were compared with known topoisomerase II poison, etoposide, for their ability to stabilize topoisomerase IIα–DNA binary complex. At the concentrations tested, all four topopyrones were shown to have potency similar to etoposide.

In this thesis, the synthesis and biological activities of seven topopyrone derivatives are reported. Initially compounds **4.8** and **4.11** were envisioned where the chlorine substitution at position 7 was changed to a bromine substitution (Figure 4.4). The synthetic route to obtain the bromo-derivatives was similar to that reported earlier for topopyrones A and B.[125] Compounds **4.9** and **4.10** were synthesized following a similar strategy as reported earlier by Elban and Hecht.[125] To understand the role of the hydroxyl groups in stabilization of DNA–topoisomerase I or II covalent binary complexes, compounds **4.5**, **4.6** and **4.7** were also synthesized.

Scheme 4.1 illustrates the synthesis of compounds **4.5** and **4.6**, which began with the protection of the phenol in a previously reported compound, **4.12**. The phenol was protected as a tosylate, where treatment of **4.12** with *p*-toluenesulfonyl chloride and K_2CO_3 in acetone at reflux gave product **4.13** in 90% yield. The next step was to protect the quinone via catalytic reduction followed by methylation. Pd/C catalyzed hydrogenation in 1:1 methanol–THF afforded the reduced quinone, which was immediately derivatized by methylation using dimethyl sulfate and KOH in 1:1 THF–water to give **4.14** in 95% overall yield. The next reaction was the deprotection of the tosylate group in aq KOH at reflux for 4 h. Although this reaction appeared to proceed cleanly as judged by silica gel TLC analysis, the yield of the isolated product was only

moderate (56%). The introduction of the pyrone moiety was achieved during three reaction steps. An acetyl group was introduced ortho to the free hydroxyl group using $TiCl_4$ and acetyl chloride in 1,2-dichoroethane. The acetyl group was first introduced onto the free OH group where $TiCl_4$ mediated Fries rearrangement directed the acetyl group ortho to the phenolic hydroxyl group. The addition of $TiCl_4$ and acetyl chloride was done at -10 ^{0}C and the reaction mixture was stirred at reflux, affording compound **4.16** in 96% yield. The pyrone ring was formed at this stage via base-catalyzed Claisen condensation reaction. The first step was the formation of the *β*-diketone ortho to the free hydroxyl group using NaH (60% dispersion in mineral oil) and freshly distilled ethyl acetate at reflux. The second step was the trifluoroacetic acid mediated ring formation followed by dehydration to afford compound **4.17** (56%, two steps). The oxidation of **4.17** with cerium(IV) ammonium nitrate in acetonitrile gave the quinone **4.18** in 72% yield. The next step was the introduction of the fourth ring which was achieved in two consecutive steps. A regioselective Diels–Alder reaction between quinone **4.18** and Danishefsky's diene 1-methoxy-1-trimethylsiloxy-3-tert-butyldimethylsiloxy-1,3-butadiene[37] (**4.19**) gave the expected *endo* adduct. It was found that the base catalyzed elimination reaction afforded the expected product **4.20** (40% yield) and unexpected product **4.21** (25% yield). Treatment of compound **4.20** with tetra-*n*-butylammonium fluoride hydrate (TBAF·H_2O) to the remove the tertbutyldimethylsilyl protecting group and subsequent treatment with 1 M BBr_3 in CH_2Cl_2 afforded topopyrone C derivative **4.5** in 23% overall yield. Similarly, another topopyrone C derivative (**4.6**) was obtained in 50% yield following the treatment of **4.21** with TABF·H_2O in THF. All attempts at obtaining NMR spectra of compounds **4.5** and **4.6** were met with failure in a variety of

solvent systems. Therefore, **4.5** and **4.6** were converted to 3,5,7-trimethoxy and 7-acetoxy derivatives, respectively. Treatment of **4.5** with dimethyl sulfate and K_2CO_3 in acetone at reflux gave the derivative as an orange solid in 45% yield. Compound **4.6** was treated with acetic anhydride in pyridine to obtain the corresponding acetoxy derivative (68% yield, yellow solid). The NMR data obtained for the trimethoxy or acetoxy derivative was in accordance with the expected structures of compounds **4.5** and **4.6**.

Scheme 4.2 outlines the synthesis of the brominated derivatives **4.7**, **4.8** and **4.11**. Previously reported compound **4.22**[125] was treated with *N*-bromosuccinimide in the presence of diisopropylamine to obtain the monobrominated crude product. The bromination was done at -78 ^oC and the reaction was complete within 5 minutes as judged by silica gel TLC analysis. The crude product was then treated with dimethyl sulfate in presence of K_2CO_3 to obtain the methylated product **4.23** in 20% overall yield. The tosylate protecting group was removed by treatment with KOH in 1:1 ethanol–water at reflux to afford the phenolic intermediate which was subsequently acetylated by using acetic anhydride in presence of triethylamine and 4-*N*,*N*-dimethylaminopyridine to obtain **4.24** (52%, two steps). At this stage, the pyrone ring was introduced following Baker-Venkataraman rearrangement. Treatment of **4.24** with LiH in freshly distilled THF at reflux gave the *β*-diketone intermediate which lead to the formation of pyrone ring by the treatment of trifluoroacetic anhydride at 0 ^oC, affording topopyrone **4.7** in 45% yield for the two steps. This reaction was highly sensitive to moisture; deacylated product was often recovered with no formation of the desired product. A harsh demethylation procedure was applied to obtain topopyrone **4.8**. Treatment of **4.7** with 1 M $AlCl_3$ in nitrobenzene over two days at 90 ^oC gave the brominated topopyrone in 50% yield. Kanai

et al.[124] reported that the angular pyrone ring (topopyrone A) underwent a base catalyzed rearrangement to give thermodynamically stable linear pyrone ring (topopyrone B). A similar method was applied to obtain compound **4.11** in 50% yield, where **4.8** was treated with 1% NaOH in methanol. Again, compounds **4.8** and **4.11** could not be characterized by NMR due to poor solubility in a variety of solvents. Only **4.11** was successfully converted to its trimethoxy derivative by using dimethyl sulfate and K_2CO_3. The NMR characterization of the trimethoxy derivative of **4.11** confirmed the structure of compound **4.11**. All attempts to convert compound **4.8** into its methoxy or acetoxy derivative met with failure; therefore, it was only characterized by high resolution mass spectrometry.

The biological evaluation of the topopyrones was done by Dr. Paul A. Zaleski.[126] The ability of the novel topopyrones **4.5** and **4.6** to stabilize the DNA−topoisomerase I covalent binary complex was investigated using a 23-base pair oligonucleotide substrate that contained a strong topoisomerase I cleavage site having guanosine at the +1 position and thymidine at the −1 position.[138] Both compounds stabilized the DNA–topoisomerase I covalent binary complex with similar efficiencies as analyzed by the polyacrylamide gel electrophoresis. In the densitometric analysis, the percentage of DNA cleavage obtained by 1 μM CPT was arbitrarily assigned a value of 100% (Table 4.1). Topopyrone **4.5** gave 6% and 11% DNA cleavage at 1 and 10 μM concentrations, respectively, which is similar to that obtained with topopyrone **4.6**. On the other hand, considerable variations were seen in DNA cleavage stabilized by topopyrones **4.7** and **4.8** (Table 4.1). Topopyrone **4.8** produced 25% DNA cleavage at 10 μM concentration, whereas topopyrone **4.7** failed to stabilize the DNA–topoisomerase I covalent binary complex at either 1 or 10 μM

concentration. X-ray crystallographic analysis of the binding of other topoisomerase I poisons to the DNA–topoisomerase I covalent binary complex indicates H-bonding between the hydroxyl groups of the poisons and the binary complex. The results shown by **4.7** and **4.8** suggest that one or more hydroxyl groups in topopyrones may be involved in H-bonding and contribute in the stabilization of the ternary complex. A nitrocellulose filter binding assay was performed to measure the activities of topopyrone C (**4.3**), **4.9** and **4.10**. The amounts of the DNA–topoisomerase I covalent binary complexes stabilized in the presence of CPT or the topopyrones were measured by the determination of the amount of the 5'-^{32}P end labeled DNA covalently bound to the enzyme. The protein denaturant sodium dodecyl sulfate (SDS) was used to trap the drug-stabilized DNA–topoisomerase I covalent binary complexes followed by the filtration through a nitrocellulose filter membrane, that specifically binds enzyme, was done. Linearized plasmid pBR322 DNA containing numerous topoisomerase I binding sites, was used as a substrate. The enzyme alone covalently bound 1.4 ± 0.5% DNA, serving as a negative control (Table 4.2). CPT at 0.5 μM concentration bound to 39 ± 4.6% of the DNA which was the positive control. Topopyrone C (**4.3**) bound 16% and 21% of the DNA at 10 and 20 μM concentrations, respectively. Topopyrone **4.10** produced similar levels of stabilized DNA–enzyme binary complex as topopyrone C and clearly gave dose-dependent activity, while topopyrone **4.9** afforded only 5-6% stabilization. Thus, the chlorine substituent at C5 position of topopyrone **4.9** is detrimental to stabilization of the DNA–topoisomerase I covalent binary complex. Topopyrone **4.10**, having the pyrone ring in a linear orientation, restored the activity of the compound as a topoisomerase I poison.

Compound **4.10** produced about the same amount of bound topo I as CPT, albeit at a higher concentration. Therefore, it was an ideal topopyrone derivative for testing activity as a topoisomerase II poison. As already discussed in section 4.1, topoisomerase IIα is relevant in proliferating cells. Therefore, this isoform of the human topoisomerase II enzyme was used for the study. In contrast to topoisomerase I, high levels of ostensibly covalently bound DNA (23 ± 2.7%) were observed without any topoisomerase II poison. In the presence of 0.5 μM etoposide, 34 ± 4.6% of the DNA was covalently bound to topoisomerase II enzyme. Topopyorone **4.10** produced similar levels of stabilized DNA–enzyme binary complex at both 1 and 5 μM concentrations.

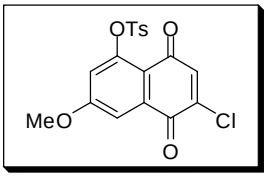
The analysis of the cleavage site specificity of topopyrone-stabilized DNA–topoisomerase I or II covalent binary complexed was also done using either 3'-^{32}P or 5'-^{32}P end labeled 158-base pair DNA duplex substrate. At 1 μM CPT, several DNA cleavage sites at 5'-TG-3' and 5'-TA-3' were observed. The topopyrones were found to stabilize topoisomerase I cleavage at sites similar to CPT but with varying efficiencies. Interestingly, three topopyrones (**4.3**, **4.9** and **4.10**) stabilized a topoisomerase I cleavage site at 5'-GT$_{72}$-3' which was not observed in case of CPT. Topopyrones **4.2** and **4.11** were used to study the cleavage site specificity of stabilized DNA–topoisomerase II covalent binary complexes and were compared with etoposide. Both the topopyrones were found to produce similar DNA cleavage sites as etoposide (5'-CA$_{23}$-3' and 5'-CG$_{74}$-3') with varying efficiencies.

4.4. Experimental

4.4.1. General experimental procedures

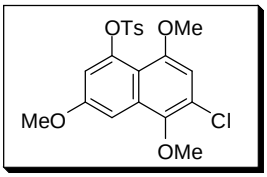
Chemicals and solvents were purchased as reagent grade from Sigma-Aldrich Chemicals and were used without further purification. All the reactions were performed under an argon atmosphere, unless otherwise specified. Thin-layer chromatography (TLC) plates (precoated glass plates with silica gel 60 F254, 0.25 mm thickness) were used for analytical TLC and were visualized by UV irradiation (254 nm). Flash chromatography was carried out using Silicycle 200−400 mesh silica gel. ^{1}H and ^{13}C NMR spectra were obtained using a Varian 400 MHz NMR spectrometer. Chemical shifts (δ) are reported in parts per million (ppm) and are referenced to residual $CHCl_3$ (δ 7.26 ppm for 1H NMR and δ 77.16 for ^{13}C NMR) as the internal strandard. Splitting patterns are designated as s, singlet; d, doublet; m, multiplet. High-resolution mass spectra were obtained at the Michigan State Mass Spectrometry Facility or the Arizona State University CLAS High Resolution Mass Spectrometry Facility.

4.4.2. Synthesis of topopyrone derivatives



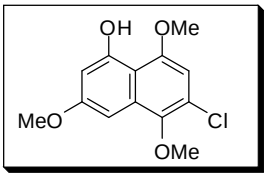
Toluene-4-sulfonic Acid 6-Chloro-3-methoxy-5,8-dioxo-5,8-dihydronaphthalen-1-yl Ester (4.13). To a solution containing 0.41g (1.7 mmol) of **4.12**[43] in 100 mL of acetone at room temperature was added 0.65 g (3.4 mmol) of *p*-toluenesulfonyl chloride (*p*-TsCl) and 0.47 g (3.4 mmol) of K_2CO_3. The reaction mixture was heated at reflux for 6 h, and the cooled solution was filtered through a silica gel plug. The silica gel was washed with

100 mL of ethyl acetate, and the combined filtrate was concentrated under diminished pressure to afford a crude residue. The residue was purified by flash chromatography on a silica gel column (15 × 3 cm). Step gradient elution with 1:1 $CHCl_3$−hexanes → 3:2 $CHCl_3$−hexanes gave **4.13** as a yellow solid: yield 0.60 g (90%); mp 150−152 °C; silica gel TLC R_f 0.47 (2:1 ethyl acetate−hexanes); ^{1}H NMR ($CDCl_3$) δ 2.43 (3H, s), 3.90 (3H, s), 6.95 (1H, s), 7.01 (1H, d, J = 2.8 Hz), 7.33 (2H, d, J = 8.4 Hz), 7.54 (1H, d, J = 2.4 Hz) and 7.84 (2H, d, J = 8 Hz); ^{13}C NMR ($CDCl_3$) δ 21.8, 56.5, 111.7, 115.9, 117.7, 128.8, 129.8, 132.4, 134.5, 137.5, 143.8, 145.9, 149.1, 163.9, 177.2 and 179.0; mass spectrum (APCI), m/z 393.0199 $(M + H)^+$ ($C_{18}H_{14}ClO_6S$ requires m/z 393.0200).

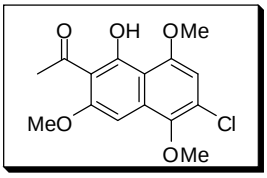


Toluene-4-sulfonic Acid 6-Chloro-3,5,8-trimethoxynaphthalen-1-yl Ester (4.14). To a solution containing 1.1 g (2.9 mmol) of **4.13** in 50 mL of 1:1 MeOH−THF under argon was added ~70 mg of Pd/C. The reaction mixture was purged with H_2 and stirred under H_2 for 2 h at room temperature, at which time it was filtered through a pad of silica gel and washed with 300 mL of ethyl acetate. The combined filtrate was concentrated under diminished pressure, and the resulting residue was dissolved in 100 mL of 1:1 THF−H_2O. To this solution was added 1.5 g (8.7 mmol) of KOH and 2.6 mL (28.6 mmol) of $(MeO)_2SO_2$. The reaction mixture was stirred for 1 h, then quenched with 200 mL of H_2O, and extracted with three 150 mL portions of ethyl acetate. The combined organic phase was dried over anhydrous $MgSO_4$, filtered, and then concentrated under diminished pressure to afford a crude residue. The residue

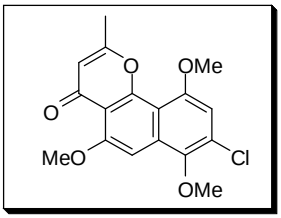
was purified by flash chromatography on a silica gel column (20 × 5 cm). Elution with 1:10 ethyl acetate−hexanes gave **4.14** as a colorless solid: yield 1.1 g (95%); mp 136−138 °C; silica gel TLC R_f 0.61 (2:1 ethyl acetate−hexanes); ^{1}H NMR ($CDCl_3$) δ 2.39 (3H, s), 3.76 (3H, s), 3.83 (3H, s), 3.87 (3H, s), 6.56 (1H, s), 6.59 (1H, d, J = 1.8 Hz), 7.26 (3H, m) and 7.70 (2H, d, J = 6.3 Hz); ^{13}C NMR ($CDCl_3$) δ 21.6, 55.6, 55.8, 60.8, 100.2, 105.6, 112.5, 114.5, 124.4, 128.3, 129.6, 132.3, 133.6, 144.1, 145.2, 146.5, 152.3 and 157.8; mass spectrum (ESI), m/z 423.0676 $(M + H)^+$ ($C_{20}H_{20}ClO_6S$ requires m/z 423.0669).



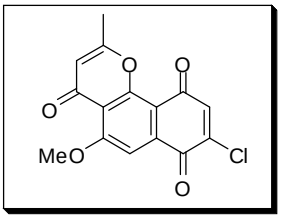
6-Chloro-1-hydroxy-3,5,8-trimethoxynaphthalene (4.15). A solution containing 1.4 g (3.3 mmol) of **4.14** in 100 mL of 1:1 EtOH−H_2O containing 10% (w/v) KOH was heated at reflux for 4 h. EtOH was removed under diminished pressure, and the remaining solution was poured into 6 N HCl (ice cold) and extracted with three 100 mL portions of ethyl acetate. The combined organic extract was dried over anhydrous $MgSO_4$, filtered, and concentrated under diminished pressure to afford a crude oil. The residue was purified by flash chromatography on a silica gel column (12 × 3 cm). Elution with 1:7 ethyl acetate−hexanes gave **4.15** as a colorless solid: yield 0.50 g (56%); mp 115−117 °C; silica gel TLC R_f 0.63 (2:1 ethyl acetate− hexanes); ^{1}H NMR ($CDCl_3$) δ 3.89 (3H, s), 3.91(3H, s), 3.99 (3H, s), 6.55-6.56 (2H, m), 6.92 (1H, d, J = 1.8 Hz) and 9.18 (1H, s); ^{13}C NMR ($CDCl_3$) δ 55.4, 56.4, 60.6, 93.4, 102.2, 103.3, 110.2, 123.1, 132.2, 145.5, 152.7, 156.2 and 160.2; mass spectrum (ESI), m/z 269.0570 $(M + H)^+$ ($C_{13}H_{14}ClO_4$ requires m/z 269.0581).



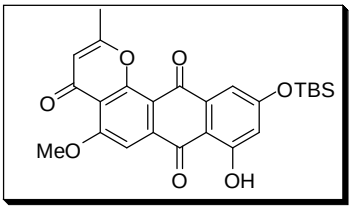
1-(6-Chloro-1-hydroxy-3,5,8-trimethoxynaphthalen-2-yl)- ethan-2-one (4.16). To a solution containing 0.50 g (1.9 mmol) of **4.15** in 20 mL of 1,2-dichloroethane at −10 °C was added 2.1 mL (2.1 mmol) of $TiCl_4$ (1.0 M in CH_2Cl_2). The reaction mixture was stirred for 5 min at −10 °C, and 0.30 mL (3.8 mmol) of acetyl chloride was added. The reaction mixture was heated at reflux for 2 h, then cooled and quenched with 400 mL of 1:1 2 N HCl−saturated aqueous potassium sodium tartrate, filtered through a pad of Celite, and washed with 200 mL of ethyl acetate. The organic phase was separated, and the aqueous phase was extracted with three 300 mL portions of ethyl acetate. The combined organic phase was dried over anhydrous $MgSO_4$, filtered, and then concentrated under diminished pressure to afford a crude residue. The residue was purified by flash chromatography on a silica gel column (15 × 3 cm). Elution with 1:10 ethyl acetate−hexanes gave **4.16** as a yellow solid: yield 0.56 g (96%); mp 128−130 °C; silica gel TLC *Rf* 0.61 (1:2 ethyl acetate−hexanes); ^{1}H NMR ($CDCl_3$) δ 2.69 (3H, s), 3.91 (3H, s), 3.97 (3H, s), 3.98 (3H, s), 6.63 (1H, s) and 6.82 (1H, s); ^{13}C NMR ($CDCl_3$) δ 33.1, 55.5, 56.5, 60.7, 91.0, 105.1, 110.7, 111.3, 127.4, 133.8, 144.3, 155.5, 157.8, 163.7 and 204.0; mass spectrum (ESI), m/z 311.0692 $(M + H)^+$ ($C_{15}H_{16}ClO_5$ requires m/z 311.0686).

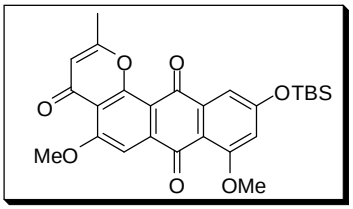


8-Chloro-2-methyl-5,7,10-trimethoxybenzo[h]chromen-4-one (4.17). To a solution containing 0.40 g (1.29 mmol) of **4.16** in 20 mL of freshly distilled ethyl acetate was added 0.57 g (14.1 mmol) of NaH (60% dispersion in mineral oil). The reaction mixture was heated at reflux for 1 h, cooled, and quenched with 100 mL of 2 N HCl (ice cold). The solution was extracted with three 200 mL portions of $CHCl_3$, dried over anhydrous $MgSO_4$, filtered, and concentrated under diminished pressure to afford a crude residue. The residue was dissolved in 15 mL of TFA at 0 °C and stirred for 15 min at 0 °C and then for 30 min at room temperature. The solvent was concentrated under diminished pressure, and the resulting residue was coevaporated with two 50 mL portions of toluene. The resulting residue was purified by flash chromatography on a silica gel column (15 × 2 cm). Step gradient elution with 1:1 ethyl acetate−hexanes → 20% MeOH in $CHCl_3$ gave **4.17** as a colorless solid: yield 0.24 g (56%); mp 182−184 °C; silica gel TLC R_f 0.43 (15% MeOH in 1:1 ethyl acetate− hexanes); ^{1}H NMR ($CDCl_3$) δ 2.42 (s, 3H), 3.95 (s, 3H), 3.97 (s, 3H), 4.06 (s, 3H), 6.26 (s, 1H), 6.76 (s, 1H) and 7.23 (s, 1H); ^{13}C NMR ($CDCl_3$) δ 19.8, 56.3, 56.4, 60.9, 96.4, 106.5, 110.4, 113.5, 113.9, 127.3, 132.7, 143.9, 154.2, 157.2, 157.4, 163.2 and 177.6; mass spectrum (ESI), m/z 335.0675 $(M + H)^+$ ($C_{17}H_{16}ClO_5$ requires m/z 335.0686).



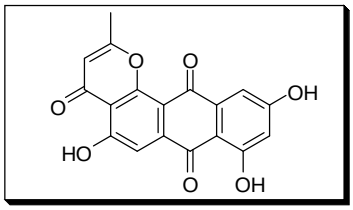
8-Chloro-5-methoxy-2-methylbenzo[h]chromene-4,7,10-trione (4.18). To a solution containing 0.24 g (0.71 mmol) of **4.17** in 50 mL of CH_3CN at 0 °C was added 0.77 g (1.4 mmol) of ceric ammonium nitrate in 8 mL of H_2O. The reaction mixture was warmed to room temperature and stirred for 0.5 h, at which time 150 mL of H_2O was added. The reaction mixture was extracted with four 150 mL portions of $CHCl_3$, and the combined organic phase was dried over anhydrous $MgSO_4$, filtered, and concentrated under diminished pressure to afford a crude residue. The residue was purified by flash chromatography on a silica gel column (15 × 2 cm). Elution with $CHCl_3$ as eluant gave **4.18** as a yellow solid: yield 0.16 g (72%); mp 257−259 °C; silica gel TLC R_f 0.50 (15% MeOH in 1:1 ethyl acetate−hexanes); ^{1}H NMR ($CDCl_3$) δ 2.43 (3H, s), 4.13 (3H, s), 6.18 (1H, s), 7.12 (1H, s) and 7.52 (1H, s); ^{13}C NMR ($CDCl_3$) δ 20.0, 57.4, 104.9, 113.1, 113.3, 117.6, 135.9, 138.2, 143.1, 157.3, 164.2, 165.4, 176.4, 177.6 and 179.3; mass spectrum (ESI), m/z 305.0213 $(M + H)^+$ ($C_{15}H_{10}O_5Cl$ requires m/z 305.0217).





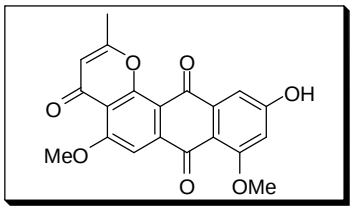
10-(tert-Butyldimethylsilyloxy)-8-hydroxy-5-methoxy-2- methyl-1-oxabenzo[a]anthracene-4,7,12-trione (4.20) and 10-(tert-Butyldimethylsilyloxy)-5,8-dimethoxy-2-methyl-1-oxabenzo[a]anthracene-4,7,12-trione (4.21). To a solution

containing 0.11 g (0.36 mmol) of 18 in 35 mL of dry benzene was added 0.12 mL (1.5 mmol) of pyridine followed by 1.1 g (3.6 mmol) of freshly prepared **4.19**.[37] The reaction mixture was heated at reflux for 18 h, after which the cooled reaction mixture was quenched with 350 mL of 2 N HCl (ice cold) and extracted with three 150 mL portions of $CHCl_3$. The combined organic layer was dried over anhydrous $MgSO_4$, filtered, and concentrated under diminished pressure to afford a crude yellow solid. The residue was purified by flash chromatography on a silica gel column (15 × 2 cm). Elution with 2% MeOH in 1:1 ethyl acetate−hexanes gave **4.20** as an orange solid: yield 67 mg (40%); mp >300 °C dec; silica gel TLC R_f 0.64 (10% MeOH in 1:1 ethyl acetate−hexanes); ^{1}H NMR δ 0.31 (6H, s), 1.00 (9H, s), 2.47 (3H, s), 4.17 (3H, s), 6.21 (1H, d, J = 0.8 Hz), 6.26 (1H, d, J = 2.4 Hz), 7.25 (1H, d, J = 2.4 Hz), 7.70 (1H, s) and 12.40 (1H, s); ^{13}C NMR ($CDCl_3$) δ -4.2, 18.4, 20.2, 25.6, 57.4, 103.8, 110.7, 112.5, 113.3, 113.4, 115.3, 117.6, 136.4, 138.8, 158.2, 164.4, 164.5, 165.2, 165.4, 176.9, 179.3 and 185.1; mass spectrum (ESI), m/z 467.1548 $(M + H)^+$ ($C_{25}H_{27}O_7Si$ requires m/z 467.1526). Further elution with 5% MeOH in 1:1 ethyl acetate−hexanes gave **4.21** as a yellow solid: yield 42 mg (25%); mp >300 °C dec; silica gel TLC R_f 0.5 (10% MeOH in 1:1 ethyl acetate−hexanes); ^{1}H NMR δ 0.32 (6H, s), 1.02 (9H, s), 2.48 (3H, s), 3.99 (3H, s), 4.14 (3H, s), 6.21 (1H, s), 6.68 (1H, d, J = 2.4 Hz), 7.35 (1H, d, J = 2.4 Hz) and 7.71 (1H, s); ^{13}C NMR ($CDCl_3$) δ -4.1, 18.4, 20.2, 25.7, 56.7, 57.2, 104.5, 108.6, 111.3, 113.3, 114.6, 115.3, 116.9, 138.9, 140.6, 157.8, 162.78, 162.84, 164.3, 165.2, 177.2, 179.9 and 180.1; mass spectrum (ESI), m/z 481.1698 $(M + H)^+$ ($C_{26}H_{29}O_7Si$ requires m/z 481.1683).



Topopyrone 4.5. To a solution containing 55 mg (0.12 mmol) of **4.20** in 10 mL of THF at room temperature was added 34 mg (0.13 mmol) of TBAF·H_2O. The reaction mixture was stirred at room temperature for 0.5 h, quenched with 100 mL of 2 N HCl, and extracted with three 50 mL portions of ethyl acetate and two 25 mL portions of $CHCl_3$. The combined organic phase was dried over anhydrous $MgSO_4$, filtered, and concentrated under diminished pressure to afford a crude orange-yellow solid. The residue was dissolved in 10 mL of CH_2Cl_2 at −78 °C, and 1.18 mL (1.18 mmol) of BBr_3 (1.0 M in CH_2Cl_2) was added. The reaction mixture was warmed to room temperature and stirred for 18 h. The reaction was quenched by the addition of 100 mL of 2 N HCl and extracted with three 50 mL portions of ethyl acetate. The combined organic layer was dried over anhydrous $MgSO_4$, filtered, and concentrated under diminished pressure to afford a crude orange-yellow solid. The crude product was purified by flash chromatography on a silica gel column (25 × 2 cm). Elution with 94:5:1 $CHCl_3$−MeOH−AcOH gave **4.5** as an orange solid: yield 9.0 mg (23%); mp >150 °C dec; silica gel TLC R_f 0.51 (94:5:1 $CHCl_3$−MeOH−AcOH); mass spectrum (APCI), m/z 339.0505 $(M + H)^+$ ($C_{18}H_{11}O_7$ requires m/z 339.0503). All attempts at obtaining NMR spectra of **4.5** met with failure in a variety of solvent systems due to insufficient solubility. Accordingly, **4.5** was converted to its 3,5,7-trimethoxy derivative to permit characterization by NMR. The derivative was also characterized by high-resolution mass spectrometry.

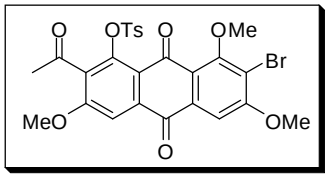
To a solution containing 10.0 mg (0.03 mmol) of **4.5** in 5 mL of acetone was added 0.04 mL (0.30 mmol) of $(MeO)_2SO_2$ and 42 mg (0.30 mmol) of K_2CO_3. The reaction mixture was heated at reflux for 18 h, quenched with 50 mL of 1 N HCl, and extracted with two 20 mL portions of ethyl acetate. The combined organic extract was dried over anhydrous $MgSO_4$, then filtered and concentrated under diminished pressure to afford a crude residue. The crude product was purified by flash chromatography on a silica gel column (15 × 2 cm). Elution with 2% MeOH in $CHCl_3$ gave the 3,5,7-trimethoxy derivative of topopyrone **4.5** as an orange solid: yield 5.0 mg (45%); silica gel TLC R_f 0.48 (2% MeOH in $CHCl_3$); ^{1}H NMR ($CDCl_3$) δ 2.48 (3H, d, J = 0.4 Hz), 3.98 (3H, s), 4.01 (3H, s), 4.14 (3H, s), 6.20 (1H, d, J = 0.4 Hz), 6.82 (1H, d, J = 2.8 Hz), 7.35 (1H, d, J = 2.4 Hz) and 7.58 (1H, s); ^{13}C NMR ($CDCl_3$) δ 20.7, 56.5, 57.2, 57.5, 103.0, 103.5,106.2, 113.5, 117.6, 118.1, 118.5, 136.3, 137.4, 158.1, 162.4, 163.3, 164.5, 165.9, 177.6, 179.6, and 183.5; mass spectrum (ESI), m/z 381.0983 $(M + H)^+$ ($C_{21}H_{17}O_7$ requires m/z 381.0974).



Topopyrone 4.6. To a solution containing 52 mg (0.11 mmol) of **4.21** in 10 mL of THF at room temperature was added 30 mg (0.11 mmol) of TBAF·H_2O. The reaction mixture was stirred for 1 h at room temperature, quenched with 100 mL of 2 N HCl, and extracted with three 50 mL portions of ethyl acetate and two 25 mL portions of $CHCl_3$. The combined organic phase was dried over anhydrous $MgSO_4$, filtered, and then concentrated under diminished pressure to afford a crude orange-yellow solid. The crude

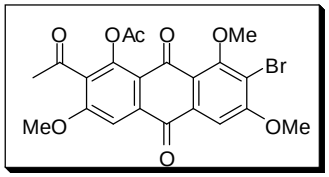
product was purified by flash chromatography on a silica gel column (20 × 2 cm). Elution with 94:5:1 $CHCl_3$−MeOH−AcOH gave **4.6** as a yellow solid: yield 20 mg (50%); mp >180 °C dec; silica gel TLC R_f 0.42 (94:5:1 $CHCl_3$− MeOH−AcOH); mass spectrum (ESI), m/z 367.0807 $(M + H)^+$ ($C_{20}H_{15}O_7$ requires m/z 367.0818). All attempts at obtaining NMR spectra of **4.6** met with failure in a variety of solvent systems due to insufficient solubility. Accordingly, **4.6** was converted to its 7-acetoxy derivative to permit characterization by NMR. The derivative was also characterized by high-resolution mass spectrometry.

A mixture of 12 mg (33 μmol) of **4.6**, 3 mL of acetic anhydride, and 1.5 mL of pyridine was stirred at room temperature for 3 h. The reaction mixture was quenched with 50 mL of 2 N HCl and extracted with two 25 mL portions of ethyl acetate. The combined organic extract was dried over anhydrous $MgSO_4$, filtered, and concentrated under diminished pressure to afford a crude residue. The crude product was purified by flash chromatography on a silica gel column (15 × 2 cm). Elution with 3% MeOH in $CHCl_3$ gave the 7-acetoxy derivative of topopyrone **4.6** as a yellow solid: yield 9.0 mg (68%); silica gel TLC R_f 0.50 (2% MeOH in $CHCl_3$); ^{1}H NMR ($CDCl_3$) δ 2.37 (3H, s), 2.48 (3H, s), 4.04 (3H, s), 4.15 (3H, s), 6.22 (1H, s), 7.07 (1H, d, J = 2.0 Hz), 7.66 (1H, d, J = 2.4 Hz), and 7.69 (1H, s); ^{13}CNMR ($CDCl_3$) δ 19.9, 21.1, 56.8, 57.1, 104.3, 110.4, 113.1, 113.2, 114.4, 116.9, 118.3, 138.4, 139.9, 156.3, 157.7, 161.7, 164.3, 165.1, 168.3, 176.8, 179.0, and 180.2; mass spectrum (ESI), m/z 409.0931 $(M + H)^+$ ($C_{22}H_{17}O_8$ requires m/z 409.0923).



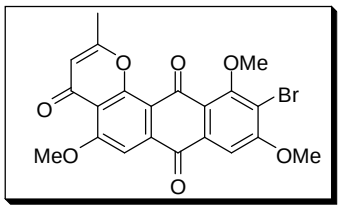
2-Acetyl-7-bromo-3,6,8-trimethoxy-1-tosyloxyanthraquinone (4.23). To a solution containing 0.05 g (0.10 mmol) of **4.22**[11] in 10 mL of CH_2Cl_2 at −78 °C was added 0.02 mL (0.15 mmol) of $(iPr)_2NH$ followed by 0.02 g (0.11 mmol) of *N*-bromosuccinimide in 10 mL of CH_2Cl_2. After the additions, the reaction was quenched by the addition of 20 mL of 1:1 H_2O−CH_2Cl_2. The phases were separated, and the aqueous phase was extracted with three 10 mL portions of CH_2Cl_2. The combined organic extract was dried over anhydrous $MgSO_4$, filtered, and then concentrated under diminished pressure to afford a crude red solid. The residue was purified by flash chromatography on a silica gel column (15 × 2 cm). Elution with 2% MeOH in 1:1 ethyl acetate−hexanes gave the crude monobrominated product **4.23** as a yellow solid. The product obtained was dissolved in 4 mL of acetone, and 0.05 g (0.36 mmol) of K_2CO_3 was added, followed by 0.03 mL (0.40 mmol) of $(MeO)_2SO_2$. The reaction mixture was stirred at reflux for 3 h, and the cooled reaction mixture was quenched with 50 mL of 2 N HCl and then extracted with three 50 mL portions of ethyl acetate. The combined organic extract was dried over anhydrous $MgSO_4$, filtered, and then concentrated under diminished pressure to afford a crude yellow solid. The residue was purified by flash chromatography on a silica gel column (15 × 2 cm). Elution with 1:2 ethyl acetate−hexanes gave **4.23** as a yellow solid: yield 12 mg (20%); silica gel TLC R_f 0.35 (1:1 ethyl acetate−hexanes); ^{1}H NMR ($CDCl_3$) δ 2.46 (3H, s), 2.53 (3H, s), 4.01 (3H, s), 4.02 (3H, s), 4.09 (3H, s), 7.36 (2H, d, J = 8.4 Hz), 7.54 (1H, s), 7.72 (1H, s) and 7.93 (2H, d, J = 8.4 Hz), ^{13}C NMR ($CDCl_3$) 21.8,

31.9, 55.9, 56.66, 56.72, 102.6, 105.6, 107.2, 117.9, 124.5, 129.0, 129.6, 132.3, 132.7, 135.4, 136.0, 143.2, 145.6, 159.3, 161.7, 164.1, 179.2, 182.4 and 198.1; mass spectrum (ESI), m/z 589.0182 $(M + H)^+$ ($C_{26}H_{22}BrO_9S$ requires m/z 589.0168).



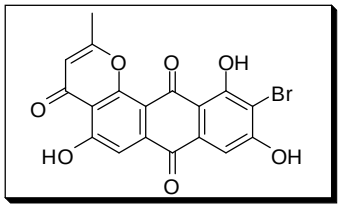
1-Acetoxy-2-acetyl-7-bromo-3,6,8-trimethoxyanthraquinone (4.24). To a solution containing 0.07 g (0.12 mmol) of **4.23** in 15 mL of 1:1 EtOH−H_2O was added 0.06 g (1.0 mmol) of KOH. The reaction mixture was heated at reflux for 4 h. The cooled solution was poured into 2 N HCl and extracted with three 50 mL portions of ethyl acetate. The combined organic extract was dried over anhydrous $MgSO_4$, filtered, and then concentrated under diminished pressure to afford a crude yellow solid. The residue was purified by flash chromatography on a silica gel column (15 × 2 cm). Elution with 1:2 ethyl acetate−hexanes afforded the product having a free hydroxyl group. This material was dissolved in 5 mL of CH_2Cl_2, and 0.03 mL (0.30 mmol) of acetic anhydride was added followed by 0.04 mL (0.30 mmol) of Et_3N and a catalytic amount of 4-*N*,*N*-dimethylaminopyridine. The reaction mixture was stirred at room temperature for 18 h, quenched with 50 mL of 2 N HCl, and extracted with three 50 mL portions of ethyl acetate. The combined organic extract was dried over anhydrous $MgSO_4$, filtered, and then concentrated under diminished pressure to afford a crude yellow solid. The residue was purified by flash chromatography on a silica gel column (15 × 2 cm). Elution with 1:2 ethyl acetate−hexanes gave **4.24** as a yellow solid: yield 30 mg (52%); silica gel TLC R_f 0.29 (1:1 ethyl acetate−hexanes); ^{1}H NMR ($CDCl_3$) δ 2.42 (3H, s), 2.50 (3H, s), 3.91

(3H, s), 4.03 (3H, s), 4.08 (3H, s), 7.55 (1H, s) and 7.66 (1H, s); ^{13}C NMR ($CDCl_3$) δ 21.0, 31.7, 56.6, 57.0, 61.8, 105.2, 106.5, 117.4, 120.2, 121.4, 131.4, 134.1, 135.4, 147.5, 159.4, 159.8, 160.7, 169.1, 178.8, 181.8 and 198.9; mass spectrum (ESI), m/z 477.0162 $(M + H)^+$ ($C_{21}H_{18}BrO_8$ requires m/z 477.0185).

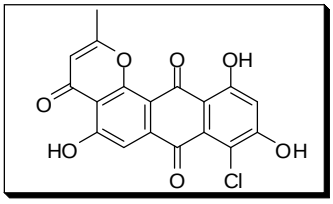


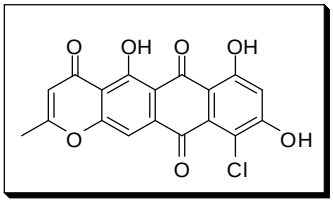
Topopyrone 4.7. To a solution containing 0.04 g (0.08 mmol) of **4.24** in 10 mL of freshly distilled THF was added 0.02 g (2.5 mmol) of LiH. The reaction mixture was heated at reflux for 18 h, and the cooled reaction mixture was quenched carefully with 50 mL of 2 N HCl. The resulting solution was extracted with three 75 mL portions of ethyl acetate, dried over anhydrous $MgSO_4$, filtered, and then concentrated under diminished pressure to afford a crude residue. The residue was dissolved in 5 mL of CF_3COOH. The reaction mixture was stirred at room temperature for 1 h and concentrated under diminished pressure, and the residue so obtained was co-evaporated with two 5 mL portions of toluene. The resulting crude residue was purified by flash chromatography on a silica gel column (7 × 2 cm). Gradient elution with 1:1 ethyl acetate−hexanes → 2% MeOH in 1:1 ethyl acetate−hexanes gave **4.7** as a yellow solid: yield 17 mg (45%); mp >280 °C dec; silica gel TLC R_f 0.21 (1:1 ethyl acetate−hexanes); ^{1}H NMR ($CDCl_3$) δ 2.48 (3H, d, J = 0.4 Hz), 4.04 (3H, s), 4.09 (3H, s), 4.16 (3H, s), 6.21 (1H, d, J = 0.8 Hz), 7.55 (1H, s) and 7.59 (1H, s); ^{13}C NMR ($CDCl_3$) δ 19.9, 56.9, 57.0, 61.8, 103.1, 104.9, 112.9, 116.7, 117.7, 117.9, 122.3, 133.5, 136.8, 157.4, 158.9, 160.3, 163.1, 165.2, 176.8,

178.1, and 182.1; mass spectrum (ESI), m/z 459.0060 $(M + H)^+$ ($C_{21}H_{16}BrO_7$ requires m/z 459.0079).

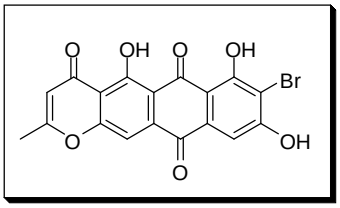


Topopyrone 4.8. A solution containing 14 mg (35 μmol) of **4.7** in 2 mL of 1.0 M $AlCl_3$ in nitrobenzene was heated at 90 °C for 48 h. The reaction was quenched while hot with 20 mL of 6 N HCl and stirred at 90 °C for 1 h. The reaction mixture was extracted with three 50 mL portionsof $CHCl_3$, dried over anhydrous $MgSO_4$, and then concentrated under diminished pressure to afford crude green oil. The crude product was purified by flash chromatography on silica gel column (15 × 2 cm). Step gradient elution with 4:1 ethyl acetate−hexanes → 95:4:1 $CHCl_3$−MeOH−AcOH gave **4.8** as a green solid: yield 6.5 mg (50%); mp >300 °C dec; silica gel TLC R_f 0.30 (10% MeOH in $CHCl_3$); mass spectrum (APCI), m/z 416.9618 $(M + H)^+$ ($C_{18}H_{10}O_7Br$ requires m/z 416.9610). All attempts at obtaining NMR spectra of **4.8** met with failure in a variety of solvent systems due to insufficient solubility.





Topopyrones 4.9 and 4.10. These two compounds were synthesized as described previously.[37]



Topopyrone 4.11. A solution containing 8.0 mg (19 μmol) of **4.8** in 4 mL of 1% NaOH in MeOH was heated at 60 °C for 48 h. The reaction mixture was neutralized carefully with 1 N HCl and concentrated under diminished pressure. The resulting residue was dissolved in 10 mL of H_2O and extracted successively with three 20 mL portions of $CHCl_3$ and two 10 mL portions of ethyl acetate. The combined organic phase was dried over anhydrous $MgSO_4$, filtered, and then concentrated under diminished pressure to afford a crude residue. The residue was purified by flash chromatography on a silica gel column (10 × 2 cm). Elution with 94:5:1 $CHCl_3$−MeOH−AcOH gave **4.11** as a green solid: yield 4.0 mg (50%); mp >300 °C dec; silica gel TLC R_f 0.23 (10% MeOH in $CHCl_3$); mass spectrum (ESI), m/z 416.9613 $(M + H)^+$ ($C_{18}H_{10}BrO_7$ requires m/z 416.9610). All attempts at obtaining NMR spectra of **4.11** met with failure in a variety of solvent systems due to insufficient solubility. Accordingly, **4.11** was converted to its 1,6,8-trimethoxy derivative to permit characterization by NMR. The derivative was also characterized by high-resolution mass spectrometry.

To a solution containing 6.0 mg (14 μmol) of **4.11** in 4 mL of acetone at room temperature was added 0.03 mL (0.23 mmol) of $(MeO)_2SO_2$ and 32 mg (0.23 mmol) of K_2CO_3. The reaction mixture was heated at reflux for 18 h, quenched with 50 mL of 1 N HCl, and extracted with two 20 mL portions of ethyl acetate. The combined organic extract was dried over anhydrous $MgSO_4$, filtered, and concentrated under diminished pressure to afford a crude residue. The crude product was purified by flash

chromatography on a silica gel column (10 × 1 cm). Elution with 2% MeOH in $CHCl_3$ gave the 1,6,8-trimethoxy derivative of topopyrone **4.11** as an orange solid: yield 4.0 mg (66%); silica gel TLC R_f 0.55 (2% MeOH in $CHCl_3$); ^{1}H NMR ($CDCl_3$) δ 2.49 (3H, s), 4.02 (3H, s), 4.10 (3H, s), 4.17 (3H, s), 6.24 (1H, s), 7.66 (1H, s), and 7.71 (1H, s); ^{13}C NMR ($CDCl_3$) δ 20.0, 57.1, 57.3, 61.5, 104.4, 105.9, 113.3, 114.3, 115.9, 119.7, 136.7, 139.6, 157.8, 159.3, 161.7, 164.4, 165.2, 179.19, and 179.21; mass spectrum (ESI), m/z 459.0077 $(M + H)^+$ ($C_{21}H_{16}O_7Br$ requires m/z 459.0079).

.

REFERENCES

1. Karp, G. *Cell and Molecular Biology: Concepts and Experiments;* John Wiley & Sons, Inc.: 2002; .

2. Green, R.; Noller, H. F. *Annu. Rev. Biochem.* **1997**, *66*, 679-716.

3. Schmeing, T. M.; Ramakrishnan, V. *Nature* **2009**, *461*, 1234-1242.

4. Lafontaine, D. L.; Tollervey, D. *Nat. Rev. Mol. Cell Biol.* **2001**, *2*, 514-520.

5. Cech, T. R. *Science* **2000**, *289*, 878-879.

6. Kurland, C. G.; Ehrenberg, M. *Annu. Rev. Biophys. Biophys. Chem.* **1987**, *16*, 291-317.

7. Ling, J.; Reynolds, N.; Ibba, M. *Annu. Rev. Microbiol.* **2009**, *63*, 61-78.

8. Baldwin, A. N.; Berg, P. *J. Biol. Chem.* **1966**, *241*, 839-845.

9. Eldred, E. W.; Schimmel, P. R. *J. Biol. Chem.* **1972**, *247*, 2961-2964.

10. Yarus, M. *Proc. Natl. Acad. Sci. U.S.A.* **1972**, *69*, 1915-1919.

11. Rodnina, M. V.; Wintermeyer, W. *Trends Biochem. Sci.* **2001**, *26*, 124-130.

12. Sievers, A.; Beringer, M.; Rodnina, M. V.; Wolfenden, R. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 7897-7901.

13. Crick, F. *J. Mol. Biol.* **1968**, *38*, 367-379.

14. Doherty, E. A.; Doudna, J. A. *Annu. Rev. Biophys. Biomol. Struct.* **2001**, *30*, 457-475.

15. Guerrier-Takada, C.; Gardiner, K.; Marsh, T.; Pace, N.; Altman, S. *Cell* **1983**, *35*, 849-857.

16. Kruger, K.; Grabowski, P. J.; Zaug, A. J.; Sands, J.; Gottschling, D. E.; Cech, T. R. *Cell* **1982**, *31*, 147-157.

17. Noller, H. F.; Hoffarth, V.; Zimniak, L. *Science* **1992**, *256*, 1416-1419.

18. Nitta, I.; Kamada, Y.; Noda, H.; Ueda, T.; Watanabe, K. *Science* **1998**, *281*, 666-669.

19. Ban, N.; Nissen, P.; Hansen, J.; Moore, P. B.; Steitz, T. A. *Science* **2000**, *289*, 905-920.

20. Nissen, P.; Hansen, J.; Ban, N.; Moore, P. B.; Steitz, T. A. *Science* **2000**, *289*, 920-930.

21. Hansen, J. L.; Schmeing, T. M.; Moore, P. B.; Steitz, T. A. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 11670-11675.

22. Bieling, P.; Beringer, M.; Adio, S.; Rodnina, M. V. *Nat. Struct. Mol. Biol.* **2006**, *13*, 423-428.

23. Weinger, J. S.; Parnell, K. M.; Dorner, S.; Green, R.; Strobel, S. A. *Nat. Struct. Mol. Biol.* **2004**, *11*, 1101-1106.

24. Koch, M.; Huang, Y.; Sprinzl, M. *Angew. Chem. Int. Ed Engl.* **2008**, *47*, 7242-7245.

25. Marshall, S. A.; Lazar, G. A.; Chirino, A. J.; Desjarlais, J. R. *Drug Discov. Today* **2003**, *8*, 212-221.

26. Leblond, F.; Davis, S. C.; Valdes, P. A.; Pogue, B. W. *J. Photochem. Photobiol. B.* **2010**, *98*, 77-94.

27. Lai, Y. T.; Tsai, K. L.; Sawaya, M. R.; Asturias, F. J.; Yeates, T. O. *J. Am. Chem. Soc.* **2013**, *135*, 7738-7743.

28. Sheldon, R. A.; van Pelt, S. *Chem. Soc. Rev.* **2013**, *42*, 6223-6235.

29. Hendrickson, T. L.; de Crecy-Lagard, V.; Schimmel, P. *Annu. Rev. Biochem.* **2004**, *73*, 147-176.

30. Cohen, G. N.; Cowie, D. B. *C. R. Hebd. Seances. Acad. Sci.* **1957**, *244*, 680-683.

31. Hendrickson, W. A.; Horton, J. R.; LeMaster, D. M. *EMBO J.* **1990**, *9*, 1665-1672.

32. Hecht, S. M.; Alford, B. L.; Kuroda, Y.; Kitano, S. *J. Biol. Chem.* **1978**, *253*, 4517-4520.

33. Heckler, T. G.; Chang, L. H.; Zama, Y.; Naka, T.; Chorghade, M. S.; Hecht, S. M.. *Biochemistry* **1984**, *23*, 1468-1473.

34. Noren, C. J.; Anthony-Cahill, S. J.; Griffith, M. C.; Schultz, P. G. *Science* **1989**, *244*, 182-188.

35. Wang, L.; Brock, A.; Herberich, B.; Schultz, P. G. *Science* **2001**, *292*, 498-500.

36. Farnum, M. F.; Magde, D.; Howell, E. E.; Hirai, J. T.; Warren, M. S.; Grimsley, J. K.; Kraut, J. *Biochemistry* **1991**, *30*, 11567-11579.

37. Miller, G. P.; Benkovic, S. J. *Chem. Biol.* **1998**, *5*, R105-R113.

38. Huennekens, F. M. *Protein Sci.* **1996**, *5*, 1201-1208.

39. Hitchings, G. H.; Burchall, J. J. *Adv. Enzymol. Relat. Areas Mol. Biol.* **1965**, *27*, 417-468.

40. Matthews, D. A.; Alden, R. A.; Bolin, J. T.; Filman, D. J.; Freer, S. T.; Hamlin, R.; Hol, W. G.; Kisliuk, R. L.; Pastore, E. J.; Plante, L. T.; Xuong, N.; Kraut, J. *J. Biol. Chem.* **1978**, *253*, 6946-6954.

41. Fierke, C. A.; Johnson, K. A.; Benkovic, S. J. *Biochemistry* **1987**, *26*, 4085-4092.

42. Polacek, N.; Mankin, A. S. *Crit. Rev. Biochem. Mol. Biol.* **2005**, *40*, 285-311.

43. Pape, T.; Wintermeyer, W.; Rodnina, M. V. *EMBO J.* **1998**, *17*, 7490-7497.

44. Nathans, D.; Neidle, A. *Nature* **1963**, *197*, 1076-1077.

45. Pestka, S. *Annu. Rev. Microbiol.* **1971**, *25*, 487-562.

46. Hohsaka, T.; Sato, K.; Sisido, M.; Takai, K.; Yokoyama, S. *FEBS Lett.* **1993**, *335*, 47-50.

47. Starck, S. R.; Qi, X.; Olsen, B. N.; Roberts, R. W. *J. Am. Chem. Soc.* **2003**, *125*, 8090-8091.

48. Dedkova, L. M.; Fahmi, N. E.; Golovine, S. Y.; Hecht, S. M. *J. Am. Chem. Soc.* **2003**, *125*, 6616-6617.

49. Dedkova, L. M.; Fahmi, N. E.; Golovine, S. Y.; Hecht, S. M. *Biochemistry* **2006**, *45*, 15541-15551.

50. Dedkova, L. M.; Fahmi, N. E.; Paul, R.; del Rosario, M.; Zhang, L.; Chen, S.; Feder, G.; Hecht, S. M. *Biochemistry* **2012**, *51*, 401-415.

51. Maini, R.; Nguyen, D. T.; Chen, S.; Dedkova, L. M.; Chowdhury, S. R.; Alcala-Torano, R.; Hecht, S. M. *Bioorg. Med. Chem.* **2013**, *21*, 1088-1096.

52. Porter, E. A.; Weisblum, B.; Gellman, S. H. *J. Am. Chem. Soc.* **2002**, *124*, 7324-7330.

53. Hart, S. A.; Bahadoor, A. B.; Matthews, E. E.; Qiu, X. J.; Schepartz, A. *J. Am. Chem. Soc.* **2003**, *125*, 4022-4023.

54. Cheng, R. P.; DeGrado, W. F. *J. Am. Chem. Soc.* **2001**, *123*, 5162-5163.

55. Frackenpohl, J.; Arvidsson, P. I.; Schreiber, J. V.; Seebach, D. *Chembiochem* **2001**, *2*, 445-455.

56. Appella, D. H.; Christianson, L. A.; Klein, D. A.; Powell, D. R.; Huang, X.; Barchi, J. J.,Jr; Gellman, S. H. *Nature* **1997**, *387*, 381-384.

57. Seebach, D.; Gardiner, J. *Acc. Chem. Res.* **2008**, *41*, 1366-1375.

58. Kritzer, J. A.; Stephens, O. M.; Guarracino, D. A.; Reznik, S. K.; Schepartz, A. *Bioorg. Med. Chem.* **2005**, *13*, 11-16.

59. Porter, E. A.; Wang, X.; Lee, H. S.; Weisblum, B.; Gellman, S. H. *Nature* **2000**, *404*, 565.

60. Cristau, M.; Devin, C.; Oiry, C.; Chaloin, O.; Amblard, M.; Bernad, N.; Heitz, A.; Fehrentz, J. A.; Martinez, J. *J. Med. Chem.* **2000**, *43*, 2356-2361.

61. Heckler, T. G.; Roesser, J. R.; Xu, C.; Chang, P. I.; Hecht, S. M. *Biochemistry* **1988**, *27*, 7254-7262.

62. Roesser, J. R.; Xu, C.; Payne, R. C.; Surratt, C. K.; Hecht, S. M. *Biochemistry* **1989**, *28*, 5185-5195.

63. Lodder, M.; Golovine, S.; Hecht, S. M. *J. Org. Chem.* **1997**, *62*, 778.

64. Lodder, M.; Golovine, S.; Laikhter, A. L.; Karginov, V. A.; Hecht, S. M. *J. Org. Chem.* **1998**, *63*, 794-803.

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

66. Robertson, S. A.; Noren, C. J.; Anthony-Cahill, S. J.; Griffith, M. C.; Schultz, P. G. *Nucleic Acids Res.* **1989**, *17*, 9649-9660.

67. Varshney, U.; Lee, C. P.; RajBhandary, U. L. *J. Biol. Chem.* **1991**, *266*, 24712-24718.

68. Huynh, M. L.; Russell, P.; Walsh, B. *Methods Mol. Biol.* **2009**, *519*, 507-513.

69. Muth, G. W.; Chen, L.; Kosek, A. B.; Strobel, S. A. *RNA* **2001**, *7*, 1403-1415.

70. Renukuntla, J.; Vadlapudi, A. D.; Patel, A.; Boddu, S. H.; Mitra, A. K. *Int. J. Pharm.* **2013**, *447*, 75-93.

71. Lohan, S.; Bisht, G. S. *Mini Rev. Med. Chem.* **2013**, *13*, 1073-1088.

72. Richardson, J. S. *Adv. Protein Chem.* **1981**, *34*, 167-339.

73. Marcelino, A. M.; Gierasch, L. M. *Biopolymers* **2008**, *89*, 380-391.

74. Che, Y.; Marshall, G. R. *Biopolymers* **2006**, *81*, 392-406.

75. Nagai, U.; Sato, K.; Nakamura, R.; Kato, R. *Tetrahedron* **1993**, *49*, 3577-3592.

76. Nachman, R. J.; Kaczmarek, K.; Williams, H. J.; Coast, G. M.; Zabrocki, J. *Biopolymers* **2004**, *75*, 412-419.

77. Hartley, O.; Gaertner, H.; Wilken, J.; Thompson, D.; Fish, R.; Ramos, A.; Pastore, C.; Dufour, B.; Cerini, F.; Melotti, A.; Heveker, N.; Picard, L.; Alizon, M.; Mosier, D.; Kent, S.; Offord, R. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 16460-16465.

78. Berger, E. A.; Murphy, P. M.; Farber, J. M. *Annu. Rev. Immunol.* **1999**, *17*, 657-700.

79. Grauer, A.; Konig, B. *Eur. J. Org. Chem.* **2009**, 5099-5111.

80. Wissner, R. F.; Batjargal, S.; Fadzen, C. M.; Petersson, E. J. *J. Am. Chem. Soc.* **2013**, *135*, 6529-6540.

81. Chatterjee, J.; Gilon, C.; Hoffman, A.; Kessler, H. *Acc. Chem. Res.* **2008**, *41*, 1331-1342.

82. Oh, K.; Guan, Z. *Chem. Commun. (Camb)* **2006**, *29*, 3069-3071.

83. Shimomura, O.; Johnson, F. H.; Saiga, Y. *J. Cell. Comp. Physiol.* **1962**, *59*, 223-239.

84. Prasher, D. C.; Eckenrode, V. K.; Ward, W. W.; Prendergast, F. G.; Cormier, M. J. *Gene* **1992**, *111*, 229-233.

85. Chalfie, M.; Tu, Y.; Euskirchen, G.; Ward, W. W.; Prasher, D. C. *Science* **1994**, *263*, 802-805.

86. Chudakov, D. M.; Matz, M. V.; Lukyanov, S.; Lukyanov, K. A. *Physiol. Rev.* **2010**, *90*, 1103-1163.

87. Ormo, M.; Cubitt, A. B.; Kallio, K.; Gross, L. A.; Tsien, R. Y.; Remington, S. J. *Science* **1996**, *273*, 1392-1395.

88. Zimmer, M. *Chem. Rev.* **2002**, *102*, 759-781.

89. Ivashkin, P. E.; Iampol'skii, I. V.; Luk'ianov, K. A. *Bioorg. Khim.* **2009**, *35*, 726-743.

90. Wang, B.; Zhou, J.; Lodder, M.; Anderson, R. D.,3rd; Hecht, S. M. *J. Biol. Chem.* **2006**, *281*, 13865-13868.

91. Thommes, P.; Hubscher, U. *Chromosoma* **1992**, *101*, 467-473.

92. Liu, L. F.; Wang, J. C. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 7024-7027.

93. Pommier, Y.; Leo, E.; Zhang, H.; Marchand, C. *Chem. Biol.* **2010**, *17*, 421-433.

94. Wang, J. C. *J. Mol. Biol.* **1971**, *55*, 523-533.

95. Wang, J. C. *Nat. Rev. Mol. Cell Biol.* **2002**, *3*, 430-440.

96. Lindsley, J. E.; Wang, J. C. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 10485-10489.

97. Wang, J. C. *Q. Rev. Biophys.* **1998**, *31*, 107-144.

98. Champoux, J. J. *Annu. Rev. Biochem.* **2001**, *70*, 369-413.

99. Wang, J. C. *Annu. Rev. Biochem.* **1996**, *65*, 635-692.

100. Kirkegaard, K.; Wang, J. C. *J. Mol. Biol.* **1985**, *185*, 625-637.

101. Watt, P. M.; Hickson, I. D. *Biochem. J.* **1994**, *303 (Pt 3)*, 681-695.

102. Liu, L. F.; Liu, C. C.; Alberts, B. M. *Cell* **1980**, *19*, 697-707.

103. Gellert, M.; Mizuuchi, K.; O'Dea, M. H.; Nash, H. A. *Proc. Natl. Acad. Sci. U.S.A.* **1976**, *73*, 3872-3876.

104. Liu, L. F. *Annu. Rev. Biochem.* **1989**, *58*, 351-375.

105. Lee, M. P.; Brown, S. D.; Chen, A.; Hsieh, T. S. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 6656-6660.

106. Goto, T.; Wang, J. C. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 7178-7182.

107. Miao, Z. H.; Player, A.; Shankavaram, U.; Wang, Y. H.; Zimonjic, D. B.; Lorenzi, P. L.; Liao, Z. Y.; Liu, H.; Shimura, T.; Zhang, H. L.; Meng, L. H.; Zhang, Y. W.; Kawasaki, E. S.; Popescu, N. C.; Aladjem, M. I.; Goldstein, D. J.; Weinstein, J. N.; Pommier, Y. *Cancer Res.* **2007**, *67*, 8752-8761.

108. Tuduri, S.; Crabbe, L.; Conti, C.; Tourriere, H.; Holtgreve-Grez, H.; Jauch, A.; Pantesco, V.; De Vos, J.; Thomas, A.; Theillet, C.; Pommier, Y.; Tazi, J.; Coquelle, A.; Pasero, P. *Nat. Cell Biol.* **2009**, *11*, 1315-1324.

109. Merino, A.; Madden, K. R.; Lane, W. S.; Champoux, J. J.; Reinberg, D. *Nature* **1993**, *365*, 227-232.

110. Soret, J.; Gabut, M.; Dupon, C.; Kohlhagen, G.; Stevenin, J.; Pommier, Y.; Tazi, J. *Cancer Res.* **2003**, *63*, 8203-8211.

111. Hsiang, Y. H.; Wu, H. Y.; Liu, L. F. *Cancer Res.* **1988**, *48*, 3230-3235.

112. Nitiss, J. L. *Nat. Rev. Cancer.* **2009**, *9*, 327-337.

113. Austin, C. A.; Marsh, K. L. *Bioessays* **1998**, *20*, 215-226.

114. Hsiang, Y. H.; Hertzberg, R.; Hecht, S.; Liu, L. F. *J. Biol. Chem.* **1985**, *260*, 14873-14878.

115. Kupfer, G.; Bodley, A. L.; Liu, L. F. *NCI Monogr.* **1987**, *4*, 37-40.

116. Hsiang, Y. H.; Liu, L. F. *Cancer Res.* **1988**, *48*, 1722-1726.

117. D'Arpa, P.; Beardmore, C.; Liu, L. F. *Cancer Res.* **1990**, *50*, 6919-6924.

118. Van, G. R.; Lendfers, R. R.; Schellens, J.; Bult, A.; Beijnen, J. *J. Oncol. Pharm. Pract.* **2000**, *6*, 92-108.

119. Bertrand, R.; O'Connor, P. M.; Kerrigan, D.; Pommier, Y. *Eur. J. Cancer* **1992**, *28A*, 743-748.

120. Kaufmann, S. H. *Cancer Res.* **1991**, *51*, 1129-1136.

121. Salerno, S.; Da Settimo, F.; Taliani, S.; Simorini, F.; La Motta, C.; Fornaciari, G.; Marini, A. M. *Curr. Med. Chem.* **2010**, *17*, 4270-4290.

122. Yamashita, Y.; Kawada, S.; Fujii, N.; Nakano, H. *Biochemistry* **1991**, *30*, 5838-5845.

123. Khan, Q. A.; Elban, M. A.; Hecht, S. M. *J. Am. Chem. Soc.* **2008**, *130*, 12888-12889.

124. Kanai, Y.; Ishiyama, D.; Senda, H.; Iwatani, W.; Takahashi, H.; Konno, H.; Tokumasu, S.; Kanazawa, S. *J. Antibiot. (Tokyo)* **2000**, *53*, 863-872.

125. Elban, M. A.; Hecht, S. M. *J. Org. Chem.* **2008**, *73*, 785-793.

126. Zaleski, P. A.; Maini, R.; Leiris, S. J.; Elban, M. A.; Hecht, S. M. *J. Nat. Prod.* **2012**, *75*, 577-585.

127. Staker, B. L.; Hjerrild, K.; Feese, M. D.; Behnke, C. A.; Burgin, A. B.,Jr; Stewart, L. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 15387-15392.

128. Staker, B. L.; Feese, M. D.; Cushman, M.; Pommier, Y.; Zembower, D.; Stewart, L.; Burgin, A. B. *J. Med. Chem.* **2005**, *48*, 2336-2345.

129. Elban, M. A. Synthesis of DNA Damaging Agents, University of Virginia, Charlottesville, VA, 2002.

130. Savard, J.; Brassard, P. *Tetrahedron* **1984**, *40*, 3455-3464.

131. Rozenberg, V.; Danilova, T.; Sergeeva, E.; Vorontsov, Z. S.; Lysenko, K.; Belokon, Y. *Eur. J. Org. Chem.* **2000**, *2000*, 3295-3303.

132. Bensari, A.; Zaveri, N. T. *Synthesis* **2003**, *2*, 267-271.

133. Krohn, K.; Bernhard, S.; Florke, U.; Hayat, N. *J. Org. Chem.* **2000**, *65*, 3218-3222.

134. Tietze, L. F.; Gericke, K. M.; Schuberth, I. *Eur. J. Org. Chem.* **2007**, *2007*, 4563-4577.

135. Mahal, H. S.; Venkataraman, K. *J. Chem. Soc.* **1934**, *56*, 1767-1769.

136. Baker, W. *J Chem Soc* **1933**, *55*, 1381-1391.

137. Drake, F. H.; Hofmann, G. A.; Bartus, H. F.; Mattern, M. R.; Crooke, S. T.; Mirabelli, C. K. **1989**, *28*, 8154-8160.

138. Bonven, B. J.; Gocke, E.; Westergaard, O. *Cell* **1985**, *41*, 541-551.