Expanding the scope of templated macromolecular synthesis in vivo: the incorporation of methionine analogues into proteins in vivo by altering the methionyl-tRNA synthetase activity of a bacterial expression host.

Kristi L. Kiick

University of Massachusetts Amherst

Follow this and additional works at: http://scholarworks.umass.edu/dissertations_1

Recommended Citation
http://scholarworks.umass.edu/dissertations_1/1015

This Open Access Dissertation is brought to you for free and open access by the Dissertations and Theses at ScholarWorks@UMass Amherst. It has been accepted for inclusion in Doctoral Dissertations 1896 - February 2014 by an authorized administrator of ScholarWorks@UMass Amherst. For more information, please contact scholarworks@library.umass.edu.
EXPANDING THE SCOPE OF
TEMPLATED MACROMOLECULAR SYNTHESIS IN VIVO:

The Incorporation of Methionine Analogues into Proteins in Vivo by
Altering the Methionyl-tRNA Synthetase Activity of a Bacterial Expression Host

A Dissertation Presented

by

KRISTI L. KIICK

Submitted to the Graduate School of the
University of Massachusetts Amherst in partial fulfillment
of the requirements for the degree of

DOCTOR OF PHILOSOPHY

May 2001

Polymer Science and Engineering
EXPANDING THE SCOPE OF
TEMPLATED MACROMOLECULAR SYNTHESIS IN VIVO:

The Incorporation of Methionine Analogues into Proteins in Vivo by
Altering the Methionyl-tRNA Synthetase Activity of a Bacterial Expression Host

A Dissertation Presented
by
KRISTI L. KIICK

Approved as to style and content by:

David A. Tirrell, Chair

E. Bryan Coughlin, Member

Lila M. Gierasch, Member

Thomas J. McCarthy, Department Head
Polymer Science and Engineering
To my grandparents
ACKNOWLEDGMENTS

There are many people to whom I am indebted for my undertaking, enjoying, and successfully completing this graduate work. Their influence extends over decades and thousands of miles, and for their support, advice, and encouragement I will forever be grateful. First, I would like to thank my advisor, David Tirrell. Working in David's laboratories for the past six and a half years has been an incredible personal and professional journey, and provided opportunities for me to work on truly interesting scientific problems. His graciousness, sense of humor, respect for scientific inquiry, personal strictness, and precise thought and articulation will serve as models for my own behavior, and the gentle guidance he has provided has helped me develop confidence in my capabilities. For all of those things I am very thankful.

I would also like to thank my committee members for their time and helpful suggestions not only throughout the scientific portion of my graduate research, but for their other professional input as well. Bruce Novak was on my committee in its early stages and contributed many helpful suggestions on potential project directions. Bryan Coughlin was kind enough to participate on my committee during the final stages of my research, and his enthusiasm and curiosity about biological polymerizations has been appreciated. Lila Giersch introduced me to biochemistry in 1987, and I was delighted to
have the opportunity to ask her to participate on my doctoral committee after all these years.

For scientific discussions and advice, I am indebted to past and present members of the Tirrell group. A better working environment would be difficult to find, and it is in large part due to these people that my graduate experience was so productive, gratifying, and fun. In the early days, it was Jody Beecher, Vince Conticello, and Jim Thomas. As a student at the University of Massachusetts, I was fortunate enough to have colleagues who have also been good friends: Wendy Naimark, Karen Guhr, Jill Sakata, Scott Kennedy, Kathy DiZio, Jeff Linhardt, Nandita Sharma, David Flanagan, Sarah Jouppi, and Karin Rotem. Their friendship, support, and advice have been truly appreciated. Yi Tang, Marissa Mock, Sarah Heilshorn, Isaac Carrico, Angelika Niemz, Jens Thies, and Kent Kirshenbaum, among others, have made the Caltech part of the journey a pleasure.

On the scientific aspects of my projects, I owe a debt of gratitude to many people. Scott Ross and Pratip Bhattachary helped during the NMR characterization of proteins containing non-natural amino acids. Jan van Hest, Ralf Weberskirch, Helen Blackwell, and Eliana Saxon kindly provided the methionine analogues that we investigated. Rebecca Alexander provided many helpful suggestions and conversations, and having a chance to work with her was an unexpected opportunity to work with an old friend. Heather Maynard and Sheldon Okada were enormously industrious and helpful in our
attempts to modify unsaturated side chains via ruthenium-catalyzed ring-opening
metathesis polymerizations. Nagarajan Vaidehi, Deepshikha Datta, and Jeremy Kua
contributed to our modeling work and the development of computational protocols to
predict methionine analogue/MetRS binding.

On a personal note, I would like to thank some people who had a critical influence
in my life many years ago and are still important in my life today. Richard Person
introduced me to chemistry during my junior year in high school, and it was directly due
to his influence that I was able to discover my interests both in science and in obtaining a
doctoral degree. Burnaby Munson and John Burmeister had an important impact on my
undergraduate experience at the University of Delaware and helped me make important
personal and professional decisions. Mike Johnson, my Master's Thesis advisor at the
University of Georgia, offered me a great start in research, and his continuing support of
my career is appreciated. My dearest old friends Jody Beecher, Janine Sullivan, Jennifer
Lastova, and Rosann Kaylor have offered unwavering support over many years and in
many difficult situations, for which I am extremely grateful. And of course, I am deeply
indebted to my family - Mom, Dad, and Karen - for providing me with the sturdiest of
roots and a desire to contribute to the world in a positive way. I am fortunate to have
such a loving and supportive family.
My final thanks go to my fiance Rick Beyer, who has perhaps contributed the most to making the past few years so enjoyable and satisfying. I am grateful for many things about him, including his intelligence, kindness, sense of humor (which he might not believe), values, and enthusiasm for learning new things. By patiently listening during countless hours of conversation, offering varied perspectives, and constantly encouraging me to express my opinion, he has helped me strengthen my voice. It’s an extremely generous gift, and I will never be able to adequately thank him.
ABSTRACT

EXPANDING THE SCOPE OF TEMPLATED MACROMOLECULAR SYNTHESIS IN VIVO:

The Incorporation of Methionine Analogues into Proteins in Vivo by Altering the Methionyl-tRNA Synthetase Activity of a Bacterial Expression Host

MAY 2001

KRISTI L. KIICK, B.S., UNIVERSITY OF DELAWARE
M.S., UNIVERSITY OF GEORGIA
M.S., UNIVERSITY OF MASSACHUSETTS AMHERST
Ph.D., UNIVERSITY OF MASSACHUSETTS AMHERST

Directed by: Professor David A. Tirrell

The in vivo incorporation of non-natural amino acids is controlled by the aminoacyl-tRNA synthetases (aaRS). The correlation between the incorporation of methionine analogues into proteins in vivo and activation of analogues by methionyl-tRNA synthetase (MetRS) in vitro has therefore been investigated. Activation of methionine analogues 2 - 13 in vitro, assessed via the ATP-PPi exchange reaction, correlates well with the ability of analogues to support protein synthesis in vivo, substantiating the critical role of MetRS in the incorporation of methionine analogues into proteins. Methionine analogues with k_cat/K_m values up to 2000-fold lower than those
for methionine (i.e., 2 (homoallylglycine), 3 (homopropargylglycine), and 9 (norleucine))
can support synthesis of a target protein (DHFR) under standard conditions of protein
expression employing a conventional bacterial host.

Overexpression of MetRS in a bacterial host was investigated as a method to
increase the number of methionine analogues that can be incorporated into proteins in
vivo. Equipping a bacterial host with extra copies of the gene encoding wild-type MetRS
permits incorporation, under certain conditions, of the additional methionine analogues 4
(cis-crotylglycine), 5 (trans-crotylglycine), 7 (2-aminoheptanoic acid), 8 (norvaline), 11
(2-butynylglycine), and 12 (allylglycine). Assessment of the level of replacement of
methionine via amino acid analysis and N-terminal sequencing indicates levels of
replacement of approximately 60% to 98%.

In an effort to broaden further the chemical functionality available for engineering
novel proteins, two azido amino acids, azidoalanine (14) and azidohomoalanine (15)
were investigated for their ability to support PP_i exchange in vitro and protein synthesis
in vivo. While 14 does not support PP_i exchange or protein synthesis, 15 is the most
efficient substrate for MetRS of all the analogues tested to date, with a k_cat/K_m value that
is 400-fold lower than that for methionine. In vivo experiments confirm that this
analogue supports protein biosynthesis, with essentially quantitative replacement of
methionine. DHFR containing 15 can be selectively modified by Staudinger ligation with triarylphosphine reagents.

These results suggest new strategies for incorporation of non-natural amino acids via manipulation of the aaRS activity of a bacterial host and provide new opportunities for the chemical modification of proteins.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>v</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>ix</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>xvi</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>xvii</td>
</tr>
<tr>
<td><strong>CHAPTER</strong></td>
<td></td>
</tr>
<tr>
<td>1. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>1.1 In Vivo Synthesis of Protein Polymers</td>
<td>1</td>
</tr>
<tr>
<td>1.2 Aminoacyl-tRNA Synthetases</td>
<td>6</td>
</tr>
<tr>
<td>1.3 Methionyl-tRNA Synthetase</td>
<td>9</td>
</tr>
<tr>
<td>1.4 Preliminary Investigations of Methionine Analogue Incorporation</td>
<td>17</td>
</tr>
<tr>
<td>1.5 References</td>
<td>25</td>
</tr>
<tr>
<td>2. IN VITRO ACTIVATION OF METHIONINE ANALOGUES BY METHIONYL-tRNA SYNTHETASE</td>
<td>36</td>
</tr>
<tr>
<td>2.1 Introduction and Objectives</td>
<td>36</td>
</tr>
<tr>
<td>2.2 Experimental Section</td>
<td>37</td>
</tr>
<tr>
<td>2.2.1 Analogue Synthesis</td>
<td>37</td>
</tr>
<tr>
<td>2.2.2 Determination of Translational Activity</td>
<td>37</td>
</tr>
<tr>
<td>2.2.2.1 Small Scale Protein Expression</td>
<td>38</td>
</tr>
<tr>
<td>2.2.2.2 Large Scale Protein Expression</td>
<td>39</td>
</tr>
<tr>
<td>2.2.2.3 Protein Purification</td>
<td>40</td>
</tr>
<tr>
<td>2.2.3 MetRS Expression and Purification</td>
<td>41</td>
</tr>
<tr>
<td>2.2.4 Activation of Methionine Analogue in Vitro</td>
<td>43</td>
</tr>
<tr>
<td>2.2.5 Computational Modeling of Methionine Analogues</td>
<td>44</td>
</tr>
</tbody>
</table>
2.3 Results and Discussion

2.3.1 Determination of Translational Activity of Methionine Analogues

2.3.2 MetRS Expression and Purification

2.3.3 Activation of Methionine Analogues by MetRS in Vitro

2.3.4 Determination of Protein Yields

2.3.5 Computational Modeling of Methionine Analogues

2.4 Summary and Conclusions

2.5 References

3. ALTERATION OF THE METHIONYL-tRNA SYNTHETASE ACTIVITY OF A BACTERIAL EXPRESSION HOST

3.1 Introduction and Objectives

3.2 Experimental Section

3.2.1 Genetic Manipulation of a Bacterial Expression Host

3.2.1.1 Construction of Plasmid pUC19-Nhelink

3.2.1.2 Construction of Plasmid pUC19-W305F

3.2.1.3 Construction of Expression Plasmids pQE9-W305F, pQE15-W305F, and pQE15-MRS

3.2.2 Protein Expression

3.2.2.1 Confirmation of MetRS and DHFR Expression

3.2.2.2 Translational Activity of Methionine Analogues in Vivo

3.2.2.3 Determination of Protein Yields

3.2.3 Determination of Intracellular Concentrations of Methionine Analogues
3.2.4 Protein Characterization

3.2.4.1 Amino Acid Analysis and N-terminal Sequencing .................................77
3.2.4.2 NMR Spectroscopy ..............................................................................77

3.2.5 Activation of Methionine Analogues by MetRS in Vitro .........................78
3.2.6 Characterization of GFP Fluorescence ..................................................81

3.3 Results and Discussion ...........................................................................83

3.3.1 Genetic Manipulation of a Bacterial Expression Host ...............................83
3.3.2 Translational Activity of Methionine Analogues in a Modified Bacterial Expression Host .................................................................85
3.3.3 Protein Yields from Modified Bacterial Expression Hosts .......................88
3.3.4 Incorporation of Additional Methionine Analogues into Proteins in Vivo .........................................................................................90
3.3.5 Intracellular Concentrations of Methionine Analogues ............................92
3.3.6 Protein Characterization ........................................................................93

3.3.6.1 Amino Acid Analysis and N-terminal Sequencing .................................93
3.3.6.2 NMR Spectroscopy .............................................................................95

3.3.7 Activation of Methionine Analogues by MetRS in Vitro ............................97
3.3.8 Characterization of GFP Fluorescence ..................................................100

3.4 Summary and Conclusions ...................................................................102
3.5 References ...............................................................................................129

4. IN VIVO INCorporATION OF AZIDOAMINO ACIDS .................................132

4.1 Introduction and Objectives ...................................................................132
4.2 Experimental Section .............................................................................133

4.2.1 Computational Methods ......................................................................133
4.2.2 Analogue Synthesis ...........................................................................133
## LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Kinetic parameters for methionine analogues in the ATP-PP, exchange reaction and protein yields for bacterial cultures supplemented with the analogues</td>
<td>55</td>
</tr>
<tr>
<td>3.1</td>
<td>Intracellular amino acid concentrations as determined by amino acid analysis</td>
<td>104</td>
</tr>
<tr>
<td>3.2</td>
<td>Amino acid analysis results for DHFR containing methionine analogues</td>
<td>105</td>
</tr>
<tr>
<td>3.3</td>
<td>Extent of replacement of methionine by methionine analogues</td>
<td>106</td>
</tr>
<tr>
<td>3.4</td>
<td>Kinetic parameters for activation of methionine analogues by MetRS</td>
<td>107</td>
</tr>
<tr>
<td>3.5</td>
<td>Comparison of catalytic efficiencies for the activation of methionine analogues by MetRS and W305F</td>
<td>108</td>
</tr>
<tr>
<td>4.1</td>
<td>Comparison of kinetic parameters for 14 and 15 with those for methionine, 3, and 9</td>
<td>147</td>
</tr>
<tr>
<td>4.2</td>
<td>Amino acid analysis results for DFHR-Met and DHFR-15</td>
<td>148</td>
</tr>
<tr>
<td>4.3</td>
<td>Mass spectrometric analysis of tryptic fragments of DHFR-15</td>
<td>149</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Schematic representation of the template nature of <em>in vivo</em> protein synthesis</td>
</tr>
<tr>
<td>1.2</td>
<td>Schematic of gene construction and protein synthesis</td>
</tr>
<tr>
<td>1.3</td>
<td>General structure of tRNA</td>
</tr>
<tr>
<td>1.4</td>
<td>The active site of MetRS</td>
</tr>
<tr>
<td>1.5</td>
<td>The structure of MetRS</td>
</tr>
<tr>
<td>2.1</td>
<td>SDS-PAGE analysis of DHFR synthesis by <em>E. coli</em> strain CAG18491/pQE15/pREP4</td>
</tr>
<tr>
<td>2.2</td>
<td>SDS-PAGE analysis of MetRS purification by size-exclusion chromatography</td>
</tr>
<tr>
<td>2.3</td>
<td>Activation of methionine and methionine analogues by MetRS</td>
</tr>
<tr>
<td>2.4</td>
<td>Comparison of kinetic parameters and protein yields for methionine analogues</td>
</tr>
<tr>
<td>2.5</td>
<td>Electron density maps and negative isopotential surfaces for methionine and methionine analogues</td>
</tr>
<tr>
<td>3.1</td>
<td>Genetic strategy for increasing MetRS activity in a bacterial host</td>
</tr>
<tr>
<td>3.2</td>
<td>Schematic of Nhe linker designed to change the identity of the cohesive ends of the gene encoding W305F to <em>Nhe I</em></td>
</tr>
<tr>
<td>3.3</td>
<td>Restriction digestion results for the construction of plasmid pQE9-W305F</td>
</tr>
<tr>
<td>3.4</td>
<td>Restriction digestion results for plasmid pQE15-W305F</td>
</tr>
</tbody>
</table>
3.5 SDS-PAGE analysis of the expression of DHFR and MetRS in CAG18491/pREP4 bacterial hosts .................................................. 113

3.6 Activation rates of methionine by whole cell lysates ........................................ 114

3.7 SDS-PAGE analysis of DHFR expression in bacterial cultures supplemented with methionine analogues ........................................ 115

3.8 Western blot analysis of the DHFR yield from modified bacterial host CAG18491/pQE15-MRS/pREP4 ........................................ 116

3.9 Western blot analysis of DHFR synthesis by bacterial expression hosts CAG18491/pQE15/pREP4 and CAG18491/pQE15-MRS/pREP4 ..... 117

3.10 Yields of purified DHFR obtained from 1-liter cultures supplemented with methionine, 2, 3, or 9 ........................................ 118

3.11 SDS-PAGE analysis of DHFR synthesis in bacterial cultures supplemented with methionine analogues at 500 mg/L .......................... 119

3.12 N-terminal sequencing results for indicating occupancy of the initiator site in DHFR produced in bacterial cultures supplemented with 4 ............ 120

3.13 N-terminal sequencing results for indicating occupancy of the initiator site in DHFR produced in bacterial cultures supplemented with 5 ............ 121

3.14 N-terminal sequencing results for indicating occupancy of the initiator site in DHFR produced in bacterial cultures supplemented with 7 ............ 122

3.15 N-terminal sequencing results for indicating occupancy of the initiator site in DHFR produced in bacterial cultures supplemented with 8 ............ 123

3.16 N-terminal sequencing results for indicating occupancy of the initiator site in DHFR produced in bacterial cultures supplemented with 11 .......... 124

3.17 N-terminal sequencing results for indicating occupancy of the initiator site in DHFR produced in bacterial cultures supplemented with 12 .......... 125
<table>
<thead>
<tr>
<th>Section</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.18</td>
<td>'H NMR spectra (599.69 MHz) of DHFR, 5, and DHFR-5 ..........126</td>
</tr>
<tr>
<td>3.19</td>
<td>Activation of methionine analogues by MetRS .................127</td>
</tr>
<tr>
<td>3.20</td>
<td>Fluorescence spectra for GFP produced from CAG18491/pQE31-GFP/pREP4 cultures supplemented with methionine, 2, 3, or 9 ..................128</td>
</tr>
<tr>
<td>4.1</td>
<td>The Staudinger ligation for the chemical modification of azidoamino acids with a triarylphosphine-FLAG reagent ..................150</td>
</tr>
<tr>
<td>4.2</td>
<td>Electron density maps and negative isopotential surfaces for methionine, 14, and 15 ........................................151</td>
</tr>
<tr>
<td>4.3</td>
<td>Activation of methionine and methionine analogues by MetRS ..........152</td>
</tr>
<tr>
<td>4.4</td>
<td>SDS-PAGE analysis of the translational activity of methionine analogues 14 and 15 ........................................153</td>
</tr>
<tr>
<td>4.5</td>
<td>Western blot analysis of the expression of DHFR in bacterial cultures supplemented with methionine or 15 ..........................154</td>
</tr>
<tr>
<td>4.6</td>
<td>N-terminal sequencing results indicating occupancy of the initiator site in DHFR produced in bacterial cultures supplemented with 15 ..........155</td>
</tr>
<tr>
<td>4.7</td>
<td>Mass spectrometric analysis of tryptic fragments of DHFR-15 ..........156</td>
</tr>
<tr>
<td>4.8</td>
<td>Fluorescence spectra for GFP produced from cultures of bacterial host CAG18491/pQE31-GFP/pREP4 supplemented with methionine or 15 ........................................157</td>
</tr>
<tr>
<td>4.9</td>
<td>Western blot analysis of DHFR produced from cultures of bacterial host CAG18491/pQE15/pREP4 supplemented with methionine or 15 ........................................158</td>
</tr>
<tr>
<td>4.10</td>
<td>Mass spectrometric analysis of tryptic fragments of DHFR-15 ..........159</td>
</tr>
</tbody>
</table>
CHAPTER 1

INTRODUCTION

1.1 In Vivo Synthesis of Protein Polymers

Although many advances in synthetic polymer chemistry have been made over the last several decades to provide the polymer chemist with increasing control over the structure of macromolecules (1-7), none have provided the level of control that is the basis of the catalytic, informational, and signal transduction capabilities of proteins and nucleic acids (8). With increasing scientific and technological interest in producing well-defined architectures and surface chemistries for nanostructure, biomaterials, and biosensors applications, the need to synthesize precisely engineered molecules with the control characteristic of proteins and nucleic acids has gained new importance. We have chosen to focus on the production of protein polymers with novel and useful materials properties owing to the useful structural, biological, and catalytic activities that are mediated by this class of macromolecules.

Protein synthesis in vivo is a template-directed polymerization in which messenger RNA (mRNA) directly encodes cellular DNA information (Figure 1.1). At the ribosome, mRNA is translated by transfer RNA (tRNA) into a corresponding amino
acid sequence. Each tRNA is charged with the appropriate amino acid by a highly selective class of enzymes, the aminoacyl-tRNA synthetases (aaRS). Once charged, the aminoacyl-tRNA is delivered to the ribosome by the elongation factor EF-Tu and accepted at the ribosomal A site. The translationally active amino acid is then covalently attached to the end of the nascent protein chain.

Harnessing the molecular weight and sequence control provided by *in vivo* synthesis should permit control of folding, functional group placement, and self assembly at the angstrom length scale. Indeed, the biosynthetic method shown in Figure 1.2 has been used to produce proteins that exhibit predictable chain-folded lamellar architectures (9-12), unique smectic liquid-crystalline structures with precise layer spacings (13), and controlled reversible gelation (14). It has become of additional interest to expand the novel chemical and physical properties that can be engineered into these classes of protein polymers by the precise placement of non-natural amino acids.

We have focused on *in vivo* incorporation of non-natural amino acids into proteins owing to the synthetic advantages it offers with respect to other methods for analogue incorporation. Introduction of non-natural amino acids can be achieved relatively simply via solid-phase peptide synthesis. While this method circumvents all biosynthetic machinery, the multistep procedure is limited to synthesis of peptides less than or equal to approximately 50 amino acids in length and is therefore not suitable for producing protein materials. Chemical aminoacylation methods, introduced by Hecht and
coworkers (15) and exploited by Schultz, Chamberlin, Dougherty, and others (16-19) provide a powerful method for the site-specific incorporation of non-natural amino acids. But because these methods (except in special cases) require the use of cell-free translation protocols that limit protein yields, they are also unsuitable for production of protein materials. Alteration of the synthetase activities of the cell is also possible through the introduction of heterologous synthetases, but has met with limited success and only permits incorporation of a non-natural amino acid in a site-specific manner (20-25). The simplicity of the in vivo approach, its relatively high synthetic efficiency, and its capacity for multisite substitution make it the method of choice for production of protein materials whenever possible.

While in certain cases it may be desirable to limit incorporation to a single site in a protein chain (e.g., studies of protein folding, structure, and dynamics), multisite substitution also has distinct advantages, as such substitution can cause important changes in protein behavior. For example, incorporation of selenomethionine in place of methionine has long been known to facilitate protein structure determination by x-ray crystallography (26-28). The incorporation of fluorinated functional groups into proteins has imparted to protein films the low surface energy characteristic of fluoropolymers; contact angles of hexadecane on fluorinated protein polymers (70°) are much higher than those on unfluorinated controls (17°) (29). Incorporation of trifluoroleucine in place of leucine also results in increases in the thermal stability of leucine zipper peptides (30);
these results may have important consequences for increasing protein stability, improving protein assembly, or strengthening ligand-receptor interactions. Furthermore, alkene functionality introduced into artificial proteins via dehydroproline can be quantitatively modified via bromination and hydroxylation (31). However, the number of amino acids shown conclusively to exhibit translational activity in vivo is small, and the chemical functionality that has been accessed by this method remains modest.

The incorporation of non-natural amino acids into proteins in vivo requires that the amino acid analogue meet several criteria. First, the analogue must be recognized and transported across the cell membrane into the cell, either by transport mechanisms specific for the natural amino acid or by more general transport machinery (32,33). Once inside the cell, the analogue must also not be degraded. Second, the non-natural amino acid not only must be recognized by the aaRS during translation, but must also be able to form a stable aminoacyl-tRNA that is not subject to the aaRS editing mechanisms that normally prevent the misacylation of tRNA (34). Finally, the non-natural aminoacyl-tRNA must be an efficient substrate for the elongation factor Tu (EF-Tu) and must also be accepted at the ribosomal A site.

Perhaps surprisingly, given the extremely high fidelity of protein biosynthesis, it has been known for decades that a reasonably large number of analogues have been identified which meet these criteria (35-44). For example, we and others have demonstrated the ability of the wild-type translational apparatus to use non-natural amino
acids with fluorinated (29,38), unsaturated (31,45,46), electroactive (47), and other side
chain functions (42,48-50). Strategies that would permit the incorporation of a broader
range of non-natural amino acids in vivo, however, would provide additional
opportunities for protein engineering.

The incorporation of non-natural amino acids into proteins in vivo does not appear
to be limited by transport into the cell or discrimination by EF-Tu and the ribosome.
Transport processes do not appear to be highly specific (51,52) and hundreds of
analogues have been accepted at the ribosome in in vitro translation assays (19,46,53,54).
Therefore, the in vivo incorporation of non-natural amino acids into proteins appears to
be controlled most stringently by the aaRS. Investigations have thus been directed
toward understanding the recognition of amino acid analogues by the aaRS in order to
expand the novel chemical and physical properties that can be engineered into proteins in
vivo.
1.2 Aminoacyl-tRNA Synthetases

Each of the 20 natural amino acids (Appendix A) has a corresponding aaRS that is responsible for catalyzing the covalent attachment of an amino acid to its cognate tRNA in two steps, activation and aminoacylation:

\[
aaRS + aa + ATP \rightleftharpoons [aaRS:aa-AMP] + PP_i
\]

\[
[aaRS:aa-AMP] + tRNA^{aa} \rightarrow aa-tRNA^{aa} + AMP + aaRS
\]

where aaRS is the enzyme, aa is the amino acid, PP_i is pyrophosphate, aaRS:aa-AMP is the aminoacyladenylate complexed with the enzyme, and aa-tRNA^{aa} is the aminoacyl-tRNA. Due to the critical role of these enzymes in the synthesis of cellular proteins that control metabolic processes, they have been the subjects of crystallography and enzymology investigations for decades, and much is known about their structures and mechanisms of action. To date, the crystal structures of 15 of the 20 aaRS have been solved and have provided insight into the detailed mechanisms of enzyme action (55-96).

Although the aaRS are structurally diverse, identification of small stretches of conserved primary sequence, combined with determination of similar three-dimensional structures in many of the aaRS, has led to the division of the 20 aaRS into two distinct groups of 10 enzymes each (97). The class I aaRS (aaRS for Arg, Cys, Gln, Glu, Ile,
Leu, Met, Trp, Tyr, and Val) display a catalytic center which is built around a nucleotide binding domain called the Rossmann fold (98). The Rossmann fold also harbours two signature sequences, HIGH and KMSKS, that are the basis of the initial division of these aaRS into a common class. The conservation of these sequences has been demonstrated to result from the functional role that the residues play in binding of amino acids, ATP, and tRNA to the aaRS (see discussion for MetRS, below). Indeed, aaRS from both prokaryotic and eukaryotic organisms display the signature sequences characteristic of the class I enzymes (99).

The class II aaRS (those for Ala, Asn, Asp, Gly, His, Lys, Phe, Pro, Ser, and Thr), on the other hand, are characterized by a catalytic center that is organized around a seven-stranded antiparallel β-sheet surrounded by α-helices (97,100,101). These secondary structures build a domain which contains three degenerate sequence motifs that are conserved in all the class II aaRS, regardless of the origin of the enzyme. The structural classification of the aaRS classes also appears to be correlated to mechanisms of action. Where the class II aaRS aminoacylate the 3'-OH group of the terminal adenosine of the tRNA, the class I aminoacylate the 2'-OH group (97). The similarities in the sequence, structure, and mechanism within the two classes of enzymes suggest a common ancestor for enzymes belonging to the same class.

Two functions are conducted by the aaRS: activation of an amino acid and aminoacylation of the amino acid's cognate tRNA. Correspondingly, all of the aaRS for
which three-dimensional structures have been determined demonstrate organization into two domains: an active center domain for recognition and activation of the amino acid and a domain specialized in the recognition of tRNA (99). The recognition of tRNA is not a limiting factor in the fidelity of protein biosynthesis, as the tRNA are relatively large molecules with several identity elements. All tRNAs are approximately 76 nucleotide residues in length and assume a typical tRNA cloverleaf structure shown in Figure 1.3, with the amino acid acceptor terminus and the anticodon at opposite ends. Primary targets for recognition of tRNA include the acceptor stem, the three bases of the anticodon, and the D-loop (99). These targets provide for high selectivity of the aaRS for its tRNA; for example, the selectivity of *E. coli* IleRS for tRNA$^{\text{Ile}}$ is $10^7$ against tRNAs for phenylalanine or formylmethionine (34).

The accuracy of amino acid recognition by the active center domain, however, is of prime importance in ensuring the fidelity of amino acid incorporation into proteins, as the amino acids are small molecules often with similar side chain structures. While the discrimination between natural amino acids most certainly involves amino acid residues contained in the signature sequences (class I) and motifs (class II) of the aaRS, often this mechanism for recognizing the amino acid is insufficient to prevent the enzyme from incorrectly forming enzyme-bound noncognate aminoacyl adenylates (misactivation). Indeed, some of the aaRS (AlaRS, IleRS, LeuRS, and ValRS) are very permissive in the
activation of noncognate amino acids, with IleRS and ValRS misactivating 7 and 8 noncognate (but naturally occurring) amino acids, respectively (34).

Despite the frequency with which some amino acids are misactivated, the fidelity of amino acid incorporation into proteins is very high, with error rates (substitution of one natural amino acid for another) on the order of 1 in 10,000 (102). The high fidelity is achieved by editing mechanisms present in some of the aaRS, which occur either by hydrolysis of the misactivated amino acid, or through the deacylation of a mischarged tRNA (34,99,103,104). However, it must be noted that not all aaRS require proofreading for accuracy, and therefore the molecular basis for the specificity of aaRS most often occurs at the first step of amino acid recognition. Because of the importance of the amino acid-aaRS interaction, efforts to manipulate the incorporation of non-natural amino acids into proteins in vivo have focused on the aaRS. In these investigations, such manipulation has focused on methionyl-tRNA synthetase (MetRS).

1.3 Methionyl-tRNA Synthetase

MetRS catalyzes the covalent attachment of methionine to its cognate tRNAs.

Homodimeric native methionyl-tRNA synthetase of E. coli is a Class 1 aaRS comprising two identical subunits of 76 kDa molecular weight (105,106). The formation of the active dimer is thought to be mediated by a C-terminal peptide of the native enzyme; removal of
approximately 120 amino acids from the C-terminus, either by mild trypsinolysis or
genetic engineering, eliminates the ability of the enzyme to dimerize, but produces a fully
active monomer with a single active site (105,107,108). The native dimer is also
observed in T. thermophilus (109), B. stearothermophilus (110), and in eukaryotic
enzymes, although yeast MetRS (both mitochondrial (111) and cytoplasmic (112,113))
lack the C-terminal peptide and are active in the monomeric form in vivo. Yeast and
human MetRS contain unique N-terminal extensions, not found in the prokaryotic
enzymes (99,114,115), that may be involved in import processes in the eukaryotic cell
(99).

Truncated forms of several bacterial MetRS (E. coli, B. stearothermophilus, T.
thermophilus, M. tuberculosis) (96,108-110,116,117) and both yeast (S. cerevisiae) and
human MetRS (117) have been cloned and expressed in E. coli, providing the opportunity
to study the enzymatic mechanisms of a variety of MetRS in vitro via site-directed
mutagenesis. Important structural and functional similarities among the MetRS of
varying origins exist; in fact, MetRS of both prokaryotic (including B.

stearothermophilus and T. thermophilus) and eukaryotic (yeast and human) origin are
capable of recognizing E. coli tRNA^Met (117), despite the presence of only 21-26%
sequence homology (117). The sequences and structures that control the activation of
methionine and aminoacylation of tRNA^Met appear to be well-conserved and are generally
found in both prokaryotic and eukaryotic MetRS. Therefore, although the discussion
below features specific details for the E. coli enzyme, the general features and
decision of action described are likely common to all MetRS.

The class I MetRS contains an active site consisting of a Rossmann nucleotide
binding fold and nearby signature sequences HIGH and KMSKS, and catalyzes the
attachment of methionine to the 2'-hydroxyl group of the terminal adenosine in tRNA^\text{Met}.
The crystal structures of various MetRS confirm that the enzyme is organized into two
subunits: the N-terminal domain which contains the active site, and the C-terminal
domain which provides for recognition of tRNA^\text{Met}. MetRS is a metalloenzyme, the E.
coli MetRS containing one tightly bound zinc ion per protein chain (118-120), which is
thought to aid in the correct folding of the enzyme (121). No direct catalytic role for the
metal ion has been indicated, although maintenance of zinc in the enzyme appears to be
critical for activation and aminoacylation, presumably by correctly orienting the side
chains involved in the methionine binding pocket (116,120,122). The three dimensional
structures solved for the monomeric forms of both E. coli and T. thermophilus MetRS
(96,123-125) illustrate that the zinc-binding region is located in a connective peptide that
makes contacts with both the Rossmann fold and the KMSKS domains. The connective
peptide forms an isolated domain located over the catalytic crevice of the MetRS (124),
consistent with zinc's putative role in correctly orienting side chain residues in the active
site.
The solution of these crystal structures, coupled with site-directed mutagenesis, has provided insight into the possible mechanisms by which MetRS recognizes and activates methionine and then covalently attaches it to tRNA\textsuperscript{Met}. The orientation of bound methionine in the active site of \textit{E. coli} MetRS, described in the recently published crystal structure for MetRS complexed with methionine (125), is shown in Figure 1.4. Residues highlighted in gray are those that line the pocket containing methionine (which can be distinguished by the yellow thioether side chain). The hydrophobic pocket enclosing the side chain of methionine comprises residues Ala12, Leu13, Tyr15, Trp253, Ala256, Pro257, Tyr260, Ile297, His301, and Trp305. Tyr15, Trp253, Pro257, and Tyr260 are in nearest proximity to the side chain, while Trp305 closes the bottom of the hydrophobic cavity (although it is not in direct contact with the terminal methyl group of methionine).

Comparison of the sequences of \textit{E. coli} MetRS and other MetRS shows that Tyr15, Trp253, Ala256, Tyr260, and His301 are strictly conserved, while conservative replacements are observed for Leu13. The other residues are not strictly conserved between species, although Trp305 is always aromatic. These observations are in good agreement with site-directed mutagenesis studies indicating that mutation of Tyr15, His301, or Trp305 to Ala results in significant decreases in the rates of activation and aminoacylation for the enzyme (126-128). Interestingly, mutation of Trp305 to Phe produces an enzyme with full activity toward methionine, consistent with the conservation of the aromaticity of this residue across species. Site-directed mutagenesis
also indicates important roles for Phe197 and Val298, as mutation of either of these residues to Ala also significantly decreases the rates of activation and aminoacylation of methionine by MetRS (116). Although these residues are not found directly in the binding cavity, their mutation must result in a conformational change in the active site that causes a loss of enzyme activity.

The sulfur atom of methionine is anchored in the site by accepting 2 hydrogen bonds (3.4Å), one from the hydroxyl group of Tyr260 and the other from the amide N-H of Leu13. His301 is hydrogen bonded to Tyr260 and shifts toward the methionine ligand upon binding. Residues Asp52 and Arg233 have both been shown to be critical for enzyme activity by site-directed mutagenesis (127). Consistent with the importance of Asp52; the NH$_2$ moiety of methionine makes one hydrogen bond with the Asp52 carboxyl group (125). Arg 233, while not directly in contact with methionine, is linked to the aromatic pocket and to the zinc-binding domain via an extended hydrogen-bonded network involving two different stretches of water molecules (125). The charged side chain of Arg233 is therefore thought to be important in stabilizing the structure of the enzyme in this region via hydrogen bonding.

Comparison of the crystal structure of MetRS complexed with methionine (125) to that of the free enzyme (124), shows that a large number of conformational changes occur in the enzyme upon binding of methionine. Tyr15 and Asn17 undergo 180° rotations toward the ligand, Trp253 shifts by 6Å, His301 shifts by 0.5Å, and Tyr260
shifts slightly. These shifts in position are mediated by concomitant unwinding of the helical peptide 294-304 with the winding of peptide 251-256 into an α-helix; this conformational flexibility at the active site may facilitate binding of methionine analogues with structures and chemical functionality dissimilar to those of methionine.

The enzyme is also capable of binding homocysteine (Hcy, Appendix A), the immediate precursor to methionine in the cell. Incorporation of Hcy into tRNA and protein is prevented by an efficient editing mechanism, which has been demonstrated to operate in E. coli, yeast, and mammalian cells (129-132). The distinct feature of this editing mechanism is that the enzyme-bound Hcy-AMP undergoes intramolecular cyclization to form homocysteine thiolactone (133). The same active site residues that are indicated by site-directed mutagenesis to be important for binding of methionine by MetRS are also important in the discrimination between methionine and homocysteine in editing (128). This suggests that E. coli MetRS may utilize a single active site that partitions an amino acid substrate through either a synthetic pathway or an editing pathway. Studies in which free thiols can be utilized by MetRS for thioester formation with activated methionine (134) suggest that this may occur via binding of the thiol sidechain of homocysteine at a subsite within the active site. Alternatively, a lower affinity of the more polar sidechain of homocysteine for the hydrophobic pocket may leave the sidechain available for cyclization (125). The lack of more general editing by
MetRS is likely to be important for the incorporation of a variety of methionine analogues into proteins \textit{in vivo}.

In addition to binding methionine and discriminating against homocysteine, MetRS must also bind ATP and tRNA; residues indicated to play a role in these binding events have also been deduced on the basis of site-directed mutagenesis, affinity labeling studies, and three-dimensional structures. The binding of ATP involves residues near the methionine binding site, in particular His21 (a well-conserved residue in the HIGH signature sequence), Lys142, Lys335 (in the KMSKS sequence), Tyr 359, and a phylogenetically conserved Tyr residue at position 358 (123,135,136).

Recognition of the tRNA$^{\text{Met}}$ anticodon is indicated to involve two peptidic regions, 391-395 and 451-467, in the C-terminal part of MetRS, specifically Asn452, Trp461, Arg395, and Asn391 (137-140). Not only are these residues highly conserved among MetRS, but they also move upon binding of methionine (125), perhaps indicating a signaling mechanism between the N-terminal and C-terminal domains of the enzyme.

Affinity labeling experiments indicate that the 3' end acceptor arm of the tRNA binds near the KMSKS signature sequence (141-143). Consistent with the fact that the 3' end of the tRNA must react with the anhydride bond of the aminoacyladénylate formed from methionine and ATP, the KMSKS signature sequence is near the ATP binding site (see above). Taken together, these observations suggest a picture of tRNA$^{\text{Met}}$ charging in which ATP and methionine bind in the N-terminal portion of MetRS, with binding of the
tRNA\textsuperscript{Met} anticodon stabilized by the C-terminal portion of the enzyme so that the acceptor arm of the tRNA is aligned with the N-terminal active site containing the enzyme-bound methionyladenylate (Figure 1.5).

Once charged to tRNA\textsuperscript{Met}, methionine is not just incorporated within a protein sequence, but also has the special role of acting as the universal initiator of translation. With these two parallel roles for methionine, all living cells possess two distinct Met-tRNA species that are aminoacylated with methionine by MetRS \textit{in vivo} and \textit{in vitro}; tRNA\textsuperscript{Met} (elongator) and tRNA\textsuperscript{fMet} (initiator) are both recognized by MetRS on the basis of their anticodon, CAU, and are aminoacylated with methionine. Once formed, Met-tRNA\textsuperscript{Met} is used during elongation, but Met-tRNA\textsuperscript{fMet} is formylated by methionyl-tRNA\textsuperscript{fMet} formyltransferase (MTF). The formylation of the Met-tRNA\textsuperscript{fMet} ensures that it will only be accepted at the ribosomal P site, and will not be recognized by the elongation factor EF-Tu or accepted at the ribosomal A site (Figure 1.1). The most important features permitting formylation of tRNA\textsuperscript{fMet} (and not tRNA\textsuperscript{Met}) are the presence of 3 G-C base pairs in the initiator tRNA acceptor stem coupled with a C1A72 base mismatch (144). The ability of the MetRS to recognize methionine analogues and use them efficiently during both initiation and elongation is a key factor in our use of such analogues for protein engineering \textit{in vivo}. That methionine analogues can be efficiently used in both initiation and elongation is suggested by the previously reported \textit{in vivo} incorporation of a several methionine analogues (27,36,42,48,50,145). Even a mutant
Lys-tRNA\(^{\text{lys}}\) can be formylated by MTF and used during initiation (146), suggesting that a variety of methionine analogues may serve efficiently during initiation as well.

In summary, successful incorporation of methionine analogues into proteins \textit{in vivo} requires 1) transport of the analogue into the cell, 2) activation of the analogue by MetRS and aminoacylation of tRNA\(^{\text{Met}}\) with the analogue, 3) lack of editing of the analogue by MetRS, and 4) the ability of the analogue to be appropriately recognized by the elongation factors and ribosome during both initiation and elongation. Previous investigations of the incorporation of methionine analogues into proteins \textit{in vivo} indicate that a variety of methionine analogues of varying chemical function and side chain length meet these requirements (39,40,42,48,145,147). These studies were the basis of the initial investigations described here, which were aimed at expanding the scope of methionine analogues that can be utilized by the translational apparatus for the incorporation of methionine analogues into proteins \textit{in vivo}.

1.4 Preliminary Investigations of Methionine Analogue Incorporation

While in theory any one of the 20 proteinogenic amino acids can be replaced by non-natural amino acids, methionine analogues are of particular interest for investigation. Due to methionine's special function as the initiator amino acid, the ability of a methionine analogue to support protein synthesis indicates that the analogue must also
serve efficiently as the initiator amino acid. The presence of analogues at the N-terminus and at other specific sites along the protein chain not only provides routes for labeling of proteins, but may also permit the production of interesting block copolymer architectures. Investigations of methionine are further motivated by the importance of methionine in mediating protein structure and protein-protein recognition processes (44,148,149); controlled incorporation of methionine analogues may permit purposeful manipulation of these phenomena.

Toward these ends, a series of methionine analogues 2 - 9 (46) (Appendix B), was previously synthesized and investigated for incorporation into proteins in vivo. Previous reports of the in vivo incorporation of selenomethionine (27,48), telluromethionine (42), norleucine (36), trifluoromethionine (50), ethionine (42), and S-nitrosohomocysteine (145) indicate that the MetRS is able to charge at least a small range of analogues to the tRNA<sup>Met</sup>. It therefore seemed possible that at least some of the analogues tested would be utilized by the protein biosynthesis machinery. Indeed, 2 (homoallylglycine) and 3 (homopropargylglycine) demonstrated translational activity in all stages of protein biosynthesis, with extents of substitution of up to 98% (45,46). (The incorporation of 9 (norleucine) had been previously reported (42).) In fact, 3 supported protein biosynthesis at levels similar to that of methionine, as 40 mg of DHFR could be isolated from 1-L cultures supplemented with either methionine or 3. Lower levels of protein synthesis
were observed for cultures supplemented with 2. In contrast, 4 - 8 did not support protein synthesis in the absence of methionine in a conventional bacterial expression host.

Despite the known importance of the aaRS in controlling amino acid incorporation into proteins *in vivo*, there has been a lack of data correlating the kinetics of analogue activation and aminoacylation by MetRS with incorporation into proteins *in vivo*. Therefore, one goal of this dissertation work has been to correlate the activation of methionine analogues by MetRS *in vitro* with their incorporation into proteins *in vivo*. Additionally, methods to expand the range of methionine analogues that can be incorporated into proteins have been investigated; expanding the types of analogues that can be utilized by a bacterial host will permit the production of proteins with novel chemical and physical properties. With the recent publication of the crystal structure of MetRS with bound methionine (125), comparisons of analogue structures with the known, bound structure of methionine may also provide important information regarding the critical features of methionine analogues that control their activation by MetRS.
Figure 1.1 Schematic representation of the template nature of *in vivo* protein synthesis, starting with initiation of translation by the binding of fMet-tRNAfMet at the ribosomal P site. (a) An aminoacyl-tRNA (Arg in this example) is delivered to the ribosomal A site and binds through its anticodon to the mRNA. (b) Peptide bond formation is catalyzed by peptidyl transferase. (c) The ribosome translocates down the mRNA chain, and the next aminoacyl-tRNA (Ala in this example) is delivered to the ribosome.
Figure 1.2  Schematic of gene construction and protein synthesis. An oligonucleotide that encodes the target sequence with potentially desirable materials properties is synthesized via standard solid phase methods and ligated enzymatically into a bacterial cloning vector. The integrity of the insert is verified via restriction analysis and standard sequencing methods. The DNA is multimerized and ligated into a cloning vector, and the desired multimer length is selected after amplification. The target length DNA is ligated into a bacterial expression vector, which then contains the DNA sequence encoding the artificial protein of interest under control of an appropriate promoter. A bacterial host containing this vector is induced to produce the artificial protein polymer, which is purified by appropriate protein purification protocols.
Figure 1.3  General structure of tRNA. The identity elements for recognition by the appropriate aaRS include the anticodon site, the acceptor stem, and the D loop.
Figure 1.4  The active site of MetRS. The thioether of methionine is shown in yellow. Nitrogen is shown in blue, and oxygen in red. The backbone of the protein is shown in green.
The structure of MetRS. The residues forming the binding pocket for methionine are shown in red. Residues that interact with ATP are shown in yellow. Residues indicated by site-directed mutagenesis to direct the binding of tRNA$^{\text{Met}}$ are shown in blue.
1.5 References


(102) Edelmann, P. and Gallant, J.A. Cell 1977, 10, 131-137.


CHAPTER 2
IN VITRO ACTIVATION OF METHIONINE ANALOGUES
BY METHIONYL-tRNA SYNTHETASE

2.1 Introduction and Objectives

Early investigations of the *in vivo* incorporation of methionine analogues (1-3) demonstrated that a variety of methionine analogues can be incorporated into proteins by the biosynthetic machinery. Initial investigations in our laboratories demonstrated that the unsaturated analogues 2 and 3 replace methionine in proteins at levels of substitution as high as 98% (2,3), while analogues 4 - 8 do not support protein synthesis *in vivo* in a conventional bacterial host. Although protein biosynthesis is thought to be controlled most stringently by the aaRS, there had been no systematic investigations of the role of MetRS in controlling the incorporation of methionine analogues into proteins *in vivo*. Therefore, the objectives of these investigations were to characterize the kinetics of activation of a series of methionine analogues 2 - 13 (Appendix B) *in vitro* and to correlate these results with the ability of these analogues to support protein synthesis in the absence of added methionine *in vivo*. 
2.2 Experimental Section

2.2.1 Analogue Synthesis

Each of the analogues 2 - 7 and 11 (Figure 2.1) was prepared by alkylation of diethyl acetamidomalonate with the appropriate tosylate followed by decarboxylation and deprotection of the amine function, as previously described (2-4). Analogues 8, 9, 12, and 13 are available commercially (Sigma-Aldrich, St. Louis, MO). Analogue 10 was a generous gift from Helen Blackwell and was prepared as described by Blackwell et. al (5). Analogues 2 - 7 and 10 - 13 were used as the racemates, and all concentrations given are for the L-isomers of the analogues.

2.2.2 Determination of Translational Activity

Buffers and media were prepared according to standard protocols (6). Due to the preference of MetRS for methionine over analogues, it is necessary that a bacterial expression host that cannot synthesize its own methionine is used during protein expression experiments aimed at the incorporation of methionine analogues. In addition, the medium must be depleted of methionine and supplemented with the analogue of interest. (The medium is supplemented with methionine to permit cell growth, and then
at the time of protein expression, the medium is depleted of methionine and supplemented only with the non-natural amino acid, as described below.)

The E. coli methionine auxotroph CAG18491 (λ, rph-1, metE3079::Tn10) kindly provided by the Yale E. coli Genetic Stock Center, was transformed with plasmids pREP4 and pQE15 (Qiagen), to obtain the expression host CAG18491/pQE15/pREP4.

The plasmid pQE15 encodes the protein murine dihydrofolate reductase (DHFR) under the control of a bacteriophage T5 promoter. The expression plasmid also encodes an N-terminal hexahistidine sequence that permits purification of the target protein by immobilized metal chelate affinity chromatography. Furthermore, DHFR contains eight methionine residues that can be replaced by methionine analogues, and its expression is easily monitored by SDS-polyacrylamide gel electrophoresis (SDS-PAGE) and visualized by Coomassie blue staining.

2.2.2.1 Small Scale Protein Expression

M9AA medium (50 mL) supplemented with 1 mM MgSO₄, 0.2 wt% glucose, 1 mg/L thiamine chloride and the antibiotics ampicillin (200 mg/L) and kanamycin (35 mg/L) was inoculated with 2 mL of an overnight culture of CAG18491/pREP4/pQE15. When the turbidity of the culture reached an optical density at 600 nm (OD₆₀₀) of 0.8, a medium shift was performed to remove methionine from the cell culture. The cells were
sedimented for 10 min at 3030 x g at 4°C, the supernatant was removed, and the cell
pellet was washed twice with 20 mL of 1 x M9 salts. Cells were resuspended in 50 mL
of the M9AA medium described above, without methionine. Test tubes containing 5 mL
aliquots of the resulting culture were prepared, and were supplemented with 300 μL 1
mg/mL either methionine or 2-13. A culture lacking methionine (or any analogue)
served as the negative control. Protein expression was induced by addition of isopropyl-
β-D-thiogalactopyranoside (IPTG, Calbiochem) to a final concentration of 0.4 mM.
Cultures were grown for 4.5 hours, the OD<sub>600</sub> was measured, and 1 mL of each of the
samples was sedimented. After the supernatant was decanted, the cell pellets were
resuspended in a volume of H<sub>2</sub>O (OD<sub>600</sub> x 100 μL) to yield a normalized OD<sub>600</sub> of 10.
Protein expression was monitored by SDS polyacrylamide gel electrophoresis (12% acrylamide running gel, 12 mA, 14 h) and visualized by Coomassie blue staining.

2.2.2.2 Large Scale Protein Expression

Similar procedures were used for preparation and isolation of DHFR from media
supplemented with 1, 2, 3, or 9. M9AA medium (100 mL) supplemented with 1 mM
MgSO<sub>4</sub> 0.2w% glucose, 1 mg/L thiamine chloride and the antibiotics ampicillin (200 mg/L) and kanamycin (35 mg/L) was inoculated with E. coli strain
CAG18491/pQE15/pREP4 and grown overnight at 37°C. This culture was used to inoculate 900 mL M9AA medium supplemented as described. The cells were grown to an OD$_{600}$ of approximately 0.9 and the medium shift was performed as described for the small-scale experiments. Washes were conducted with 500 mL of 1 x M9 salts, and the washed cell pellet was resuspended in 1 L of M9 + 19AA media (- methionine). An aliquot (20 mL) of 1 mg/mL of the analogue was added, followed by IPTG (0.4 mM), and cultures were grown for 4.5 hours. The OD$_{600}$ was measured, and 1 mL-samples were sedimented and resuspended as described for the small-scale cultures. Protein expression was monitored by SDS polyacrylamide gel electrophoresis. The 1-L culture was sedimented (9800 x g, 10 minutes, 4°C), the supernatant removed, and the cell pellet stored at -80°C overnight.

2.2.2.3 Protein Purification

The cell pellet was thawed for 30 min at 37°C, 30 mL of buffer (6 M guanidine-HCl, 0.1 M NaH$_2$PO$_4$, 0.01 M Tris, pH 8) was added, and the mixture was shaken at room temperature for 1 hour. The cell lysate was then sonicated for 60 seconds in 1 second bursts. The cell debris was sedimented (15,300 x g, 30 min, 4°C) and the supernatant was subjected to immobilized metal affinity chromatography (Ni-NTA resin) under denaturing conditions according to the procedure described by Qiagen (7). The
supernatant was loaded onto 10 mL of resin which was then washed with 50 mL of guanidine buffer followed by 25 mL of urea buffer (8 M urea, 0.1 M NaH₂PO₄ and 0.01 M Tris, pH 8). Similar urea buffers were used for three successive 25 mL washes at pH values of 6.3, 5.9 and 4.5, respectively. Target protein was obtained in washes at pH 5.9 and 4.5. These washes were combined and dialyzed (Spectra/Por membrane 1, MWCO = 6-8 kDa) against running distilled water for 4 days, followed by batchwise dialysis against doubly distilled water for one day. The dialysate was lyophilized to yield 35 - 40 mg of modified DHFR for cultures grown on medium supplemented with 3, similar to the yield obtained for DHFR from cultures supplemented with methionine. Proteins produced from cultures (1 L) grown on medium supplemented with 2 yielded 10 mg of DHFR, while supplementation with 9 yielded 20 mg of protein. A control experiment in 2xYT medium afforded 60 mg of DHFR per liter culture.

2.2.3 MetRS Expression and Purification

The fully active, truncated form of the wild-type MetRS was purified from 24-hour cultures of JM101 cells carrying the plasmid pGG3 (8). (The pGG3 plasmid was kindly donated by Professor Hieronim Jakubowski of UMDNJ-New Jersey Medical School, Newark, New Jersey.) The plasmid pGG3 encodes MetRS under control of the E. coli promoter metGpl (Genbank accession number X55791); constitutive expression
of the enzyme in the cellular host provided sufficient levels of enzyme for facile purification and use. The enzyme was purified by size exclusion chromatography as previously described (9). The cell pellet from a 24-hour culture was resuspended in 360 mL of 20 mM Tris-Cl, 0.1 mM EDTA (pH 7.6) and stored at -80°C overnight. The suspension was thawed on ice in a cold room (4°C) and 360 μL of β-mercaptoethanol (β-ME) was added. Streptomycin sulfate (12 g, 3% w/v) was added to the mixture and the solution was stirred at 4°C for two hours. The cell lysate was clarified by centrifugation (25000 x g, 45 minutes, 4°C). Ammonium sulfate (80 g, 35%) was added, the solution was stirred at 4°C overnight, and the precipitate was removed from the solution by centrifugation as above. Additional ammonium sulfate (95g, 70%) was added to the supernate, the solution was stirred at 4°C for 4 hours, and the precipitate was collected by centrifugation as above. The pellet was dissolved in 2 mL 10 mM phosphate buffer, 10 mM β-ME (pH 6.8), and 1 mL of this solution was loaded onto a Sephacryl S-100 HiPrep size exclusion column and eluted isocratically with 10 mM phosphate, 10 mM β-ME buffer at pH 6.8. The fractions containing MetRS (as determined by absorbance measurements at 280 nm) were collected and concentrated in a centicon filter apparatus (10,000 MWCO) via centrifugation. The purified enzyme was stored as 3-μM solutions in 40% glycerol at -20°C until used for enzyme assays. (A truncated mutant MetRS, W305F, was also purified identically from overnight cultures of DH5αF' carrying the
plasmid pBSM547W305F. The plasmid pBSM547W305F was kindly donated by Professor Yves Mechulam of the Ecole Polytechnique, Palaiseau Cedex, France.)

2.2.4 Activation of Methionine Analogue(s) in Vitro

Activation of methionine analogues by wild-type MetRS was assayed via the amino-acid-dependent ATP-PP\(_i\) exchange reaction, also as previously described (9-11). The assay, which measures the \(^{32}\text{P}\)-radiolabeled ATP formed by the enzyme-catalyzed exchange of \(^{32}\text{P}\)-pyrophosphate (PP\(_i\)) into ATP, was conducted in 150 \(\mu\)l of reaction buffer (pH 7.6, 20 mM imidazole, 0.1 mM EDTA, 10 mM \(\beta\)-mercaptoethanol, 7 mM MgCl\(_2\), 2 mM ATP, 0.1 mg/ml BSA, and 2 mM PP\(_i\) (in the form of sodium pyrophosphate (NEN Life Science Products, Inc.) with a specific activity of 0.2-0.5 TBq/mole)). (1 Becquerel (Bq) = 1 disintegration per second (dps); \(3.7 \times 10^{10}\) Bq = 1 Curie (Ci).)

Assays to determine if the methionine analogues 2 - 13 support PP\(_i\) exchange in the presence of MetRS were conducted in solutions 75 nM in enzyme and 5 mM in the L-isomer of the analogue with a reaction time of 20 minutes. Quantitative kinetic parameters for analogues 5 and 11 were obtained with an enzyme concentration of 75 nM and analogue concentrations of 100 \(\mu\)M to 20 mM. Parameters for methionine were obtained by using concentrations ranging from 10 \(\mu\)M to 1 mM. K\(_m\) values for
methionine matched those previously reported (12), although the measured $k_{cat}$ was somewhat lower than the literature value. Aliquots of 20 μl were removed from the reaction mixture at various time points and were quenched in 0.5 ml of a solution comprising 200 mM NaPP₃, 7% w/v HClO₄, and 3% w/v activated charcoal. The charcoal was rinsed twice with 0.5 mL of a 10 mM NaPP₃, 0.5% HClO₄ solution and was then resuspended in 0.5 mL of this solution and counted via liquid scintillation methods. Kinetic constants were calculated by nonlinear regression fit of the data to a Michaelis Menten model.

2.2.5 Computational Modeling of Methionine Analogues

Single-point energy *ab initio* calculations (Hartree-Fock model, 6-31G* basis set) were performed for methionine and for analogues 2, 3, 5, and 11 with fully extended side chains. Electron density maps are shown as surfaces of electron density 0.08 electrons/au³. Isopotential plots are represented as surfaces where the energy of interaction between the amino acid and a point positive charge is equal to -10 kcal/mole. Calculations were performed by using the program MacSpartan (Wavefunction, Inc., Irvine, CA, USA).
2.3 Results and Discussion

2.3.1 Determination of Translational Activity of Methionine Analogues

A bacterial host strain (CAG18491/pQE15/pREP4) suitable for testing the translational activity of methionine analogues 2 - 8 and 10 - 13 was prepared by transformation of E. coli strain CAG18491, a methionine auxotroph, with the repressor plasmid pREP4 and the expression plasmid pQE15. The expression plasmid pQE15 encodes murine DHFR, which contains 8 methionine residues as possible sites of substitution by non-natural amino acids. The translational activity of each analogue was assessed on the basis of its capacity to support synthesis of DHFR in cultures of CAG18491/pQE15/pREP4 that had been depleted of methionine.

The results of the in vivo assays illustrated in Figure 2.1 corroborate our previous reports that 2 and 3 serve effectively as methionine surrogates in bacterial protein synthesis (2,3). Analysis of purified DHFR produced from cultures supplemented with these analogues confirmed incorporation of 2 and 3 at levels up to 98%, also as previously reported. In contrast, analogues 4 - 8 and 10 - 13 do not support measurable levels of protein synthesis in bacterial cultures depleted of methionine. It is highly unlikely that recognition by EF-Tu, recognition by the ribosome, or transport into the cell are the limiting factors for incorporation of these analogues. The ribosome is remarkably
permissive toward amino acid analogues with widely varying chemical functionality, as has been demonstrated by the numerous analogues incorporated into proteins in *in vitro* translation experiments (13-20).

Transport of many analogues into the bacterial cell is indicated by a number of literature reports. Analogue 4 is an antagonist for methionine, inhibiting the growth of *E. coli* cells (21) and 8 replaces leucine in human hemoglobin expressed in *E. coli* (22). Although there is no similar evidence reported for the other analogues, the fact that trifluoromethionine and ethionine are incorporated into proteins expressed in *E. coli* (1,23-25) suggests that neither the trifluoromethyl group nor the longer side chain will inhibit transport of analogues 6 and 7 into *E. coli* cells.

### 2.3.2 MetRS Expression and Purification

The attachment of an amino acid to its cognate tRNA proceeds in two steps, as below,

\[
\text{aaRS} + \text{aa} + \text{ATP} \rightleftharpoons [\text{aaRS:aa~AMP}] + \text{PP}_i
\]

\[
[\text{aaRS:aa~AMP}] + \text{tRNA}^{\text{aa}} \rightarrow \text{aa~tRNA}^{\text{aa}} + \text{AMP} + \text{aaRS}
\]

where aaRS is the enzyme, aa is the amino acid, PP<sub>i</sub> is pyrophosphate, aaRS:aa~AMP is the aminoacyladenylate complexed with the enzyme, and aa~tRNA<sup>aa</sup> is the
aminoacyl-tRNA. Activation, the first step, involves the enzyme-catalyzed formation of an aminoacyl adenylate (aa~AMP) and can be studied by monitoring the rate of exchange of radiolabeled pyrophosphate ($^{32}$P-PP$_i$) into ATP (26,27). Aminoacylation, the second step, can be studied by monitoring the amount of radiolabeled amino acid attached to tRNA in the presence of the enzyme. Because initial recognition of an amino acid by its aaRS is perhaps the most critical step in the incorporation of non-natural amino acids into proteins in vivo, the in vitro activation of methionine analogues by MetRS has been measured and compared to results of studies of in vivo incorporation.

Purified enzyme preparations are necessary to conduct the in vitro assays, and SDS-PAGE analysis was used as an assessment of MetRS purity. The truncated form of MetRS was expressed from 24-hour cultures and purified by size-exclusion methods. The SDS-PAGE analysis of the isolated MetRS (both wild-type and mutant) is shown in Figure 2.2. As indicated in the figure, both the wild-type and mutant forms of MetRS were readily purified from cell lysates by size-exclusion chromatography.

2.3.3 Activation of Methionine Analogues by MetRS in Vitro

The relative rates of activation of methionine and methionine analogues 2-13 by wild-type MetRS were estimated by the ATP-PP$_i$ exchange assay. The results of the in vitro assays shown in Figure 2.3 are consistent with the in vivo results shown in
Figure 2.1, as the analogues that support the highest rates of PP\(_i\) exchange also support protein synthesis in the absence of methionine. Methionine (1) is activated most efficiently by the enzyme, causing exchange of 9 nmoles PP\(_i\) over the time course of the reaction. Analogues 2 and 3 cause exchange of PP\(_i\) at rates similar to that of norleucine (9), while the remaining analogues 4, 6 - 8, and 12 - 13 cause exchange of PP\(_i\) at levels no higher than background (Figure 2.3, lane 14) under these experimental conditions.

Although analogues 5 and 11 effect very slow change of PP\(_i\), the rate of activation is apparently too low to support detectable levels of protein synthesis \textit{in vivo}.

Kinetic parameters were obtained for methionine, 2, 3, 5, 9, and 11 as described in the experimental section; results are shown in Table 2.1. Our measured value of K\(_m\) for methionine matched previously reported values (12), although the value determined for k\(_{cat}\) was slightly lower than that reported. Comparison of the k\(_{cat}\)/K\(_m\) values obtained for methionine (0.54 s\(^{-1}\) μM\(^{-1}\)) and the above analogues show that the analogues are 500-fold to 13825-fold poorer substrates for the enzyme, with catalytic efficiencies in activation decreasing in the following order: Met > 3 > 9 > 2 > 5 > 11. Even 3, which supports the same protein yields as methionine in one-liter cultures, is 500-fold poorer a substrate for the enzyme than methionine. That the loss of catalytic activity is manifest both in increasing values of K\(_m\) and decreasing values of k\(_{cat}\) indicates that the analogues are less efficient in both the binding and catalytic events of the enzyme.
Table 2.1 also demonstrates that methionine analogues that are activated up to 2000-fold more slowly by MetRS than methionine can support protein biosynthesis in a conventional bacterial expression host in the absence of methionine. These results are comparable to those reported previously for the in vitro activation and in vivo incorporation of phenylalanine analogues (28-30); comparisons for other amino acids have been limited by a lack of in vitro activation data. The data suggest that non-natural amino acids can support protein synthesis in vivo even with surprisingly inefficient activation of the amino acid by its aaRS, and that under normal cellular conditions, activation of a natural amino acid by its aaRS is not the limiting factor in protein biosynthesis.

2.3.4 Determination of Protein Yields

The trend in protein yield shown in Table 2.1, coupled with the early observation that the levels of protein synthesis in cultures supplemented with methionine or 3 were significantly greater than those in cultures supplemented with 2 (2,3), led us to probe the correlation of the catalytic efficiency of activation in vitro with protein yield in vivo. This correlation was probed by comparing the kinetic constants for analogue activation by MetRS with the yield of the target protein DHFR obtained from one-liter cultures of the bacterial host CAG18491/pQE15/pREP4.
The relative kinetic constants for analogue activation and the corresponding protein yields are illustrated on the bar graph shown in Figure 2.4. Analogues with the highest $k_{cat}/K_m$ values support the highest levels of protein synthesis; the protein yields scale remarkably well with the $k_{cat}/K_m$ values, at least for the poorer substrates. Analogue 3 supports protein synthesis with yields equivalent to those obtained with methionine (35 mg/L), despite the fact that 3 is a 500-fold poorer substrate for MetRS than methionine. Bacterial cultures supplemented with 9 (1050-fold lower $k_{cat}/K_m$) produce approximately 57% as much DHFR as cultures supplemented with methionine, and cultures supplemented with 2 (1850-fold lower $k_{cat}/K_m$) produce 28% of the control yield of protein. Bacterial cultures supplemented with 5 (4700-fold lower $k_{cat}/K_m$) or 11 (13825-fold lower $k_{cat}/K_m$) do not support measurable levels of protein synthesis in this expression host. These results indicate that the rate of methionine analogue activation in vitro does indeed correlate with protein yields in vivo, and suggest that the kinetics of activation can play a critical role in controlling the rate of protein synthesis in methionine-depleted cultures supplemented with analogues that are poor substrates for MetRS.
Computational Modeling of Methionine Analogues

Computational modeling of methionine analogues provides a basis for comparison of structural and electronic properties of the methionine analogues and potential insight into the analogue features that are important for analogue activation by MetRS. Comparisons of the structural and electronic properties between methionine and the analogues may be useful for better understanding the activation of the analogues, especially given that the crystal structure of MetRS with bound methionine has been solved (31), and the side chain features important in binding are indicated.

Figure 2.5 compares the isopotential surfaces calculated for methionine and for analogues 2, 3, 5, and 11. The representation of the extended side chain conformations is consistent with that observed for methionine in the active site of MetRS (31). That 2 might serve as a substrate for the MetRS is not surprising, given the similar geometries accessible to 1 and 2, the availability of π-electrons near the side-chain terminus of 2, and the known translational activity of 9, the saturated analogue of 2. Indeed, the importance of the thioether of methionine as a hydrogen-bond acceptor (31) suggests that the π-electrons of 2 may serve a similar function. The hydrophobicity of the side chain of 9, coupled with its similar size, likely explains the ability of 9 to be activated by the highly hydrophobic active site of MetRS. The high translational activity observed for 3 (i.e., near-quantitative replacement of methionine without loss of protein yield), was not
anticipated, since the colinearity of the side chain carbons 4 - 6 imposes on 3 a geometry substantially different than that of methionine. However, the electron density associated with the triple bond of 3 is positioned similarly to that of the thioether of the natural substrate 1, despite the differences in side-chain geometry. Additionally, the isopotential surface of 3 (Figure 2.5) extends from the side chain in the same direction and at a similar distance as that for methionine whereas that of 2 does not, which may make 3 a better substrate for MetRS than 2. Furthermore, given the highly aromatic nature of the methionine binding pocket in MetRS, alkynyl C-H/π contacts (32) and the polarizability of the unsaturated side chain may also play a significant role in recognition of 3 by the enzyme.

Figure 2.5 also compares the geometries of 1, 5, and 11. The latter two analogues are not translationally active in a conventional bacterial expression host and are not activated efficiently by MetRS in vitro. Although the geometries of 1 and 5 appear similar in the representation shown, the fixed planarity of the C₄-C₅ double bond may preclude the side-chain conformation required for efficient recognition of 5 by MetRS. The isopotential surface also extends only from the sides of the side chain and not in the same position as that in methionine, which may reduce the ability of 5 to act as a good substrate for MetRS. The linearity of the side chain of 11, coupled with its very different geometry than 1 or 3, must preclude the analogue from achieving a side chain orientation that permits efficient activation by MetRS. That 11 supports even low rates of PP₅.
exchange may be due to the presence of π-electrons in a position on the side chain similar to that of the thioether in methionine. These results suggest that the presence of π electrons and the polarizability of the triple bond play a significant role in the recognition of both 3 and 11 by MetRS, given that the analogues are activated by the enzyme despite their linear geometries.

2.4 Summary and Conclusions

Strategies to optimize the incorporation of non-natural amino acids into proteins in vivo are important for engineering protein materials. Toward these ends, a set of twelve methionine analogues was assayed for translational activity in E. coli and activation by MetRS in vitro. Results of the in vitro assays corroborate the in vivo results, with analogues that support the highest level of PP, exchange in vitro also serving most effectively as methionine surrogates in vivo. Our results indicate that methionine analogues that are up to 2000-fold poorer substrates for MetRS in vitro are still able to support protein synthesis in a conventional bacterial host. Furthermore, quantitative assessment of the kinetics of activation indicates that the catalytic efficiency for a given analogue also correlates well with in vivo yields of the test protein DHFR obtained from cultures supplemented with the analogue, indicating a critical role for MetRS in methionine analogue incorporation into proteins.
The electron density maps and isopotential surfaces for methionine and analogues 2, 3, 5, and 11 indicate the importance of side chain length and electron density in the activation of the analogues by the enzyme. Analogues 2, 3, and 9 all have side chains the same length as that of methionine and all support protein synthesis and PP\textsubscript{i} exchange under the assay conditions. In contrast, none of the analogues 7, 8, 12, or 13 show translational activity or support PP\textsubscript{i} exchange in these in vitro assays. The position of electron density along the side chains of 2 and 3 likely plays a role in their activation and may permit relatively efficient activation of 3 despite the linearity of its side chain. The importance of conformational flexibility was indicated by observations that the internal alkene function of 4 and 5, and the internal alkyne function of 11, prevented their incorporation into test protein under conditions employing a conventional bacterial host. Additionally, the analogues 6 and 10 yielded no evidence of translational activity or PP\textsubscript{i} exchange.

These results suggest that the success or failure of analogue incorporation in vivo is indeed controlled by MetRS. Strategies to incorporate non-natural amino acids into proteins should therefore include assessment of the aaRS activities of the bacterial host.
Table 2.1  Kinetic parameters for methionine analogues in the ATP-PP\textsubscript{i} exchange reaction and protein yields for bacterial cultures supplemented with the analogues.

<table>
<thead>
<tr>
<th>Analogue</th>
<th>$K_m$ (µM)</th>
<th>$k_{cat}$ (s\textsuperscript{-1})</th>
<th>$k_{cat}/K_m$ (s\textsuperscript{-1}µM\textsuperscript{-1})</th>
<th>Protein Yield, mg/L</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>24.3 ± 2</td>
<td>13.3 ± 0.2</td>
<td>5.47 x 10\textsuperscript{-1}</td>
<td>35</td>
</tr>
<tr>
<td>3</td>
<td>2415 ± 170</td>
<td>2.60 ± 0.3</td>
<td>1.08 x 10\textsuperscript{-3}</td>
<td>35</td>
</tr>
<tr>
<td>9</td>
<td>4120 ± 900</td>
<td>2.15 ± 0.6</td>
<td>5.22 x 10\textsuperscript{-4}</td>
<td>20</td>
</tr>
<tr>
<td>2</td>
<td>4555 ± 200</td>
<td>1.35 ± 0.1</td>
<td>2.96 x 10\textsuperscript{-4}</td>
<td>10</td>
</tr>
<tr>
<td>5</td>
<td>15675 ± 250</td>
<td>1.82 ± 0.6</td>
<td>1.16 x 10\textsuperscript{-4}</td>
<td>0</td>
</tr>
<tr>
<td>11</td>
<td>38650 ± 2000</td>
<td>1.51 ± 0.5</td>
<td>3.91 x 10\textsuperscript{-5}</td>
<td>0</td>
</tr>
</tbody>
</table>
Figure 2.1 SDS-PAGE analysis of DHFR synthesis by *E. coli* strain CAG18491/pQE15/pREP4. Cultures were supplemented with nothing (–Met) or with methionine or one of the analogues 2-13, as indicated.
Figure 2.2  SDS-PAGE analysis of MetRS purification by size-exclusion chromatography. (a) Wild-type MetRS (b) Mutant MetRS, W305F. The molecular weight marker at 62 kDa is indicated; both MetRS enzymes have molecular weights of approximately 62.5 kDa. The proteins were visualized by Coomassie Blue staining.
Figure 2.3  Activation of methionine and methionine analogues by MetRS. The amount of PP_i exchanged in 20 minutes, as measured in the ATP-PP_i exchange assay, is shown for methionine (1) and analogues 2 - 13. The background (14) is given for a reaction mixture lacking enzyme and amino acid.
Figure 2.4  Comparison of kinetic parameters and protein yields for methionine analogues. Kinetic parameters $k_{cat}/K_m$, normalized to that for methionine (Table 2.1), are given for the analogues as indicated. The protein yield for 1 L cultures of CAG18491/pQE15/pREP4 supplemented with the analogue is given as a percentage of the yield obtained for cultures supplemented with methionine.
Figure 2.5  Electron density maps (colored surfaces) and negative isopotential surfaces (meshes) for methionine (a) and for analogues 2, 3, 5, and 11 (b - e, respectively). The electron density maps indicate electron-rich (red) and electron-poor (blue) regions of each molecule. For simplicity, the neutral amino acid form is shown; this avoids representations of the highly extended isopotential surfaces of the carboxylate anion of the zwitterion and facilitates comparison of side-chain electronic structure.
2.5 References


CHAPTER 3
ALTERATION OF THE METHIONYL-tRNA SYNTHETASE ACTIVITY OF
A BACTERIAL EXPRESSION HOST

3.1 Introduction and Objectives

The incorporation of methionine analogues in vivo correlates well with the in vitro
activation of the analogues by MetRS, as analogues that support high levels of PP_i
exchange in vitro also support protein synthesis in vivo. Two analogues, 5 and 11, which
support low levels of PP_i exchange in vitro, are unable to support protein biosynthesis in
a conventional bacterial host. Several other methionine analogues (4, 6-8, 10, 12-13)
were not utilized by the bacterial host during protein biosynthesis (and did not support
measurable levels of PP_i exchange in the in vitro assays described in Chapter 2). The
yields of protein obtained from cultures supplemented with a given analogue correlated
well with the catalytic efficiency of the analogue activation by MetRS. Based on these
results, we speculated that manipulation of the MetRS activity of the expression host
might enable additional methionine analogues to serve as substrates during bacterial
protein synthesis.

Two methods were used to investigate this hypothesis. First, overexpression of
the wild-type MetRS in a bacterial host was investigated via the introduction of plasmid
copies of the gene encoding MetRS. This strategy has not been employed previously for incorporating non-natural amino acids into proteins in vivo, but reports of the in vivo misacylation of tRNA substrates by overexpressed aaRS supported the viability of the approach (1-4). Overexpression of a mutant MetRS may also serve as an additional strategy for expanding the substrate range of the translational apparatus. Substitution of Trp305 by Phe (W305F) produces a mutant MetRS with the same enzymatic rates of activation or aminoacylation as the wild-type MetRS (5,6). These results have significance for the in vivo incorporation of methionine analogues, as the substitution of the bulky Trp with the smaller Phe in the active site of MetRS (7) may result in relaxed substrate specificity and permit incorporation of analogues not utilized by the wild-type MetRS. Previous investigations demonstrating the utility of this approach for expanding the substrate range of PheRS to include p-Cl-Phe and p-Br-Phe (8-11) suggest that a similar strategy could be successful for MetRS.

The goals of these investigations were as follows. First, production of bacterial hosts capable of overexpressing wild-type or mutant MetRS was accomplished by insertion of the gene encoding the MetRS or W305F onto a multiple-copy expression plasmid followed by transformation of the plasmid into an appropriate bacterial host. The ability of analogues 2 - 13 to support protein biosynthesis in the modified bacterial hosts was tested, and the effect of overexpression of MetRS on the yields of analogue-containing protein produced by the bacterial host was studied. Differences in the abilities
of the wild-type and mutant MetRS to activate the methionine analogues were assessed by monitoring the *in vitro* ATP-PP\(_i\) exchange activity of both enzymes.

### 3.2 Experimental Section

#### 3.2.1 Genetic Manipulation of a Bacterial Expression Host

Recombinant DNA manipulations were carried out using standard, commercially available reagents (from Sigma, Aldrich, Qiagen, New England Biolabs, Inc.) and standard protocols (12) unless otherwise noted. The expression plasmids pQE9 and pQE15 were obtained from Qiagen. The RGS-His antibody and anti-mouse IgG horseradish peroxidase conjugate used for Western blotting procedures were obtained from Qiagen and Amersham Life Sciences, respectively. The ECL Western blotting detection reagents for developing Western blots were obtained from Amersham Pharmacia Biotech.

Insertion of the gene encoding MetRS into the expression plasmid pQE15 was accomplished using the genetic strategy outlined in Figure 3.1. Simply, a linker was designed to permit the cohesive ends of the MetRS gene to be changed to *Nhe* I. The resulting fragment was then ligated into the unique *Nhe* I site of the expression plasmid pQE9 or pQE15. Specific experimental protocols are delineated below.
3.2.1.1 Construction of Plasmid pUC19-Nhelink

Single-stranded oligonucleotides encoding the Nhe linker sequence shown in Figure 3.2 were supplied by Genosys (The Woodlands, TX) and dissolved in water to yield a final concentration of 1 μg/mL. The oligonucleotides encode a series of restriction enzyme sites that permit the insertion of a DNA fragment at Kpn I and Sac I sites, with subsequent liberation of the gene fragment by digestion with Nhe I. The oligonucleotides were annealed and phosphorylated as follows. To a reaction mixture of 10 μL of each oligonucleotide was added 5 μL of 2M NaCl, 2 mL 1M MgCl$_2$ and 73 μL water to yield a final volume of 100 μL. The solution was vortexed, incubated in boiling water for 5 minutes, and then placed in a styrofoam cooler and allowed to cool to room temperature slowly overnight. The DNA was precipitated by adding 10 μL of 3M sodium acetate (pH 4.8), vortexing, adding 300 μL of cold (-20°C) ethanol, and then incubating the DNA at -80°C for 20 minutes. After centrifugation, the supernatant was removed and the DNA dried under vacuum for 5 minutes. The DNA was then phosphorylated by redissolving the DNA pellet in 44.5 μL water, adding 5 μL of T4 polynucleotide kinase buffer (to bring the final reaction conditions to 70 mM Tris-HCl (pH 7.6), 10 mM MgCl$_2$, 5 mM dithiothreitol), and adding 5 units of T4 polynucleotide kinase. The reaction was incubated at 37°C for 1 hour and immediately mixed with 5.6
μL of 10x agarose gel loading buffer and isolated by agarose gel electrophoresis with a 2% 2w/v agarose gel (visualization with ethidium bromide).

The annealed, phosphorylated, double-stranded DNA was purified by excising the appropriate band (103 bp) from the agarose gel and extracting the DNA from the agarose via the Qiagen QIAquick spin gel extraction protocol (13). The cohesive ends of the linker were designed as Aat II and Sph I, so that the linker could be added easily to the cloning vector pUC19, retaining the necessary unique restriction sites. The plasmid pUC19 (~1 μg) was linearized by digestion for 4 hours at 37°C with 10 units each of Aat II and Sph I enzymes in NEB buffer 4 (50 mM potassium acetate, 20 mM Tris-acetate, 10 mM magnesium acetate, 1 mM dithiothreitol, pH 7.9). All digestions throughout this chapter were performed with a similar amount of DNA and at 37°C for four hours unless otherwise noted. The linearized plasmid was not dephosphorylated, but was isolated by agarose gel electrophoresis as above. The insert and linearized plasmid were ligated in a molar ratio of 5:1 (50:10 ng), with 5 units of T4 DNA ligase in a total volume of 20 mL T4 DNA ligase buffer (50 mM Tris-HCl (pH 7.8), 10 mM MgCl₂, 10 mM dithiothreitol, 1 mM ATP, and 25 μg/L BSA). The ligation was conducted at room temperature for 4 hours and then the ligation mixture was transformed into the E. coli cloning strain DH5αF' according to the following procedure for making and transforming competent
cells. All ligation procedures were done identically as described here unless otherwise noted.

*E. coli* DH5αF' cells were grown to an optical density (OD$_{550}$) of approximately 0.5, placed on ice for 10 minutes, and collected by centrifugation at 2000 x g for 15 minutes at 4°C. The cells were resuspended in 2 mL of TFB1 buffer (10 mM morpholino ethanesulfonic acid (MES) (pH 6.2), 100 mM RbCl$_2$, 10 mM CaCl$_2$·2H$_2$O, 50 mM MnCl$_2$·4H$_2$O) by gentle tilting of the tube. The volume was then adjusted to 16 mL with additional TFB1 buffer. The suspension was left on ice in the cold room for 15 minutes and then the cells were pelleted by centrifugation at 2000 x g for 15 minutes at 4°C. The cells were resuspended in 2 mL TFB2 buffer (10 mM 3-(N-morpholino)propane sulfonic acid (MOPS), 75 mM CaCl$_2$·2H$_2$O, 10 mM RbCl$_2$ and 15% glycerol) and chilled on ice for 15 minutes before being dispensed into 200 μL aliquots which were stored at -80°C.

An aliquot of DH5αF' competent cells (50 μL) was thawed on ice for 20 minutes. To the cells, 5 μL ligation reaction was added, and the mixture was incubated on ice for 45 minutes. The sample was heat-shocked at 42°C for approximately 90 seconds followed by incubation on ice for 5 minutes. The sample was then diluted with 500 μL 2xYT media and incubated on a rotary incubator for 1 hour at 37°C. Cells (200 μL) were spread on agar plates containing 200 μg/mL ampicillin and were grown for 16 hours at 37°C.
Eighteen of the approximately 400 single colonies from these plates were used to inoculate eighteen overnight cultures (5 mL of 2xYT). An aliquot (1.5 mL) of the overnight culture was centrifuged to harvest the cells, and the DNA was isolated from the cells by the spin miniprep procedure described by Qiagen (14). Restriction analysis of the DNA was conducted by digesting the DNA at 37°C for 4 hours with Kpn I and Sac I (10 units of each enzyme in NEB 1 buffer with BSA (10 mM bis tris propane-HCl, 10 mM MgCl₂, 1 mM dithiothreitol, 100 μg/mL BSA, pH 7.0)) to confirm the presence of the Nhelinker insert. Similar digestion with Nhe I for 4 hours at 37°C also confirmed these results (10 units Nhe I in NEB buffer 2 with BSA (50 mM NaCl, 10 mM Tris-HCl, 10 mM MgCl₂, 1 mM dithiothreitol, 100 μg/mL BSA, pH 7.9)). Two samples were submitted for oligonucleotide sequencing analysis (Oligonucleotide Synthesis and Sequencing Facility, California Institute of Technology, Pasadena, CA), which confirmed the sequence (in one of the two samples) of the insert (103 bp) in the plasmid pUC19-Nhelink.

3.2.1.2 Construction of Plasmid pUC19-W305F

The gene encoding MetRS W305F (2449 bp) was removed from plasmid pBSM547W305F (a kind gift from Professor Yves Mechulam of the Ecole Polytechnique, Palaiseau Cedex, France) by restriction enzyme digestion with Kpn I and...
Sac I, employing 10 units of each enzyme in NEB 1 buffer with BSA (10 mM bis tris propane-HCl, 10 mM MgCl₂, 1 mM dithiothreitol, 100 μg/mL BSA, pH 7.0). The plasmid pUC19-Nhelink was linearized by similar digestion with these two enzymes and dephosphorylated by adding calf intestinal alkaline phosphatase (CIP, 1 unit) directly to the digestion mixture and incubating for an additional 30 minutes at 37°C. Both the W305F gene and linearized pUC19-Nhelink plasmid were isolated and purified by agarose gel electrophoresis (1.0% gel with visualization by ethidium bromide) as described above. The purification of the W305F gene required separation of two similarly sized fragments (2449 and 2862 bp), so the fragment was purified by gel electrophoresis twice prior to its use in the ligation mixture. The W305F and linearized pUC19-Nhelink were ligated, and the ligation mixture was transformed into DH5αF' using the procedures described above. Restriction digestion analysis (Nhe I) of the DNA isolated from 9 of the resulting single transformants indicated the presence of insert (2498 bp) in 7 out of 9 colonies tested. One sample was retained as pUC19-W305F and the isolated DNA was stored at -20°C.

3.2.1.3 Construction of Expression Plasmids pQE9-W305F, pQE15-W305F, and pQE15-MRS

The gene encoding W305F was removed from the plasmid pUC19-W305F by digestion with Nhe I (as described above) and isolated and purified by agarose gel...
electrophoresis. The expression plasmids pQE9 and pQE15 (Qiagen, Valencia, CA) were linearized, also by digestion with Nhe I, and similarly purified. The fragment W305F was ligated into each of these expression plasmids separately and the resulting ligation mixtures transformed into cloning strain DH5αF' according to the procedures described above. After amplification and isolation of the plasmid DNA by miniprep procedures (14), digestion with Nhe I was conducted to confirm the presence of the W305F insert (2498 bp) in both of the resulting plasmids pQE9-W305F (5 of 9 colonies) and pQE15-W305F (5 out of 10 colonies). Additional restriction analysis with enzymes Aat II and Nhe I (10 units each, NEB buffer 4) was conducted to confirm the presence of the gene encoding DHFR (920 bp) in the pQE15-W305F plasmid. The orientation of the W305F gene in pQE9-W305F and pQE15-W305F was deduced by restriction analysis with Hind III (10 units, NEB buffer 2).

Transformation of pQE15-W305F into DHFαF' resulted, in one case, in genetic recombination of the mutant gene with the chromosomal copy of the wild-type MetRS gene, yielding plasmid pQE15-MRS. Along with the repressor plasmid pREP4, the plasmids pQE9-W305F, pQE15-W305F, and pQE15-MRS were each transformed into the expression strains CAG18491 or B834(DE3) (Stratagene) (made competent by the same procedures described for DH5αF') and plated onto agar plates containing 200 μg/mL ampicillin and 35 μg/mL kanamycin. Isolation of single colonies yielded the
modified expression hosts CAG18491/pQE9-W305F/ pREP4, CAG18491/pQE15-W305F/pREP4, and CAG18491/pQE15-MRS/pREP4, along with hosts B834(DE3)/pQE9-W305F/ pREP4, B834(DE3)/pQE15-W305F/pREP4, and B834(DE3)/pQE15-MRS/pREP4. Cell stocks from overnight cultures were prepared and stored at -80°C, and DNA stocks were stored at -20°C. Plasmid DNA from all cultures used for protein expression experiments was sequenced to confirm that it encoded the appropriate MetRS.

3.2.2 Protein Expression

3.2.2.1 Confirmation of MetRS and DHFR Expression

Single colonies of the bacterial expression hosts CAG18491/pQE15/pREP4 and either CAG18491/pQE15-W305F/pREP4 or CAG18491/pQE15-MRS/pREP4 were used to inoculate 5 mL 2xYT cultures containing 200 μg/mL ampicillin and 35 μg/mL kanamycin. The cultures were grown to an OD_{600} of 0.8 before inducing expression of the target protein DHFR with 0.4 mM IPTG, and the cultures were permitted to grow for an additional 4.5 hours. Samples (1 mL) of each of the cultures were sedimented and resuspended in distilled H₂O as previously described. The cell lysates from these experiments were analyzed by SDS-PAGE, and the proteins visualized by staining with
Coomassie blue. Similar experiments were conducted for bacterial host

CAG18491/pQE9-W305F/pREP4, except that overnight cultures were grown without
induction of protein expression.

3.2.2.2 Translational Activity of Methionine Analogues in Vivo

The translational activity of the methionine analogues 2 - 13 was tested in both
contventional bacterial hosts equipped with the plasmid pQE15 and modified bacterial
hosts carrying expression plasmids pQE15-MRS or pQE15-W305F. The hosts were
produced by transformation with the appropriate expression plasmid and the repressor
plasmid pREP4 as described above. The small-scale expression experiments in minimal
media were conducted as described in Section 2.2.2, except that the supplementation
level of the analogues was either 20 mg/L, 60 mg/L, or 500 mg/L of the L-isomer. In
experiments to test analogue incorporation, analogue concentrations of 60 mg/L were
used. In experiments designed to test the yield of protein obtained from a conventional
versus a modified bacterial host, supplementation of 20 mg/L was employed. For large-
scale experiments designed to determine protein yields, and for producing proteins for
amino acid analysis and N-terminal sequencing, supplementation of 20 mg/L was
employed for 5, while 500 mg/L was employed for analogues 4, 7, 8, 11, and 12. Protein
expression was monitored by SDS-PAGE, visualized either by Coomassie blue staining
or by Western blotting. The accumulation of the target protein DHFR was taken as evidence for incorporation of the non-natural amino acid. Incorporation was confirmed by purifying the protein by metal chelate affinity chromatography (15,16) as previously described, and analyzing the purified protein by amino acid analysis and N-terminal sequencing, and when possible, NMR spectroscopy.

3.2.2.3 Determination of Protein Yields

A Western blotting procedure designed to assess the relative yields of protein obtained from 5-mL cultures was conducted as previously described (17). Proteins were transferred from an acrylamide gel to a nitrocellulose membrane, and Western blots were developed by treatment with a primary RGS-His antibody (20 µg/mL in 5% w/v dried milk blocking solution), followed by washing and treatment with a secondary anti-mouse IgG conjugated to horseradish peroxidase (50 µg/mL in 5% w/v blocking solution). Chemiluminescence detection of the labeled proteins was afforded by use of ECL detection reagents (Amersham Pharmacia Biotech) and documented on film. Films were exposed for varying times to ensure that band intensity was not saturated. Levels of protein synthesis were estimated by the intensity of the band on the gel, as determined by a Pharmacia Ultrascan XL laser densitometer and analysis by Pharmacia GelScan XL software. Large scale yields were determined by purifying DHFR from 1 L cultures via
metal chelate affinity chromatography under denaturing conditions, as previously described in Chapter 2 (16).

3.2.3 Determination of Intracellular Concentrations of Methionine Analogues

The approximate intracellular concentrations of methionine and methionine analogues 8 or 9 during protein expression were determined by amino acid analysis of cell lysates from 50-mL samples of CAG18491/pQE15-MRS/pREP4 grown on medium containing 19 AA and supplemented with methionine, 8, or 9. Cultures of 250 mL were grown to an OD$_{600}$ of 0.7 on medium supplemented with methionine. After a medium shift, the cells were resuspended in a medium containing 20 mg/L or 500 mg/L of methionine, 8, or 9. Protein expression was induced by the addition of IPTG (0.4 mM), and 50 mL samples were harvested in duplicate when the cultures reached an OD of 1.0 (approximately 2 hours). The cell pellets were washed once with 1 x M9 salts. Free amino acids were extracted as previously described (18). Harvested cells were resuspended in 2.5 mL distilled H$_2$O, and the pH was adjusted to 10 with a few drops of 5N NH$_4$OH. The samples were placed in a boiling water bath for 10 minutes and then cooled to room temperature. The pH was adjusted to 2 with a few drops of 5N HCl, and the precipitate removed by centrifugation. The supernatant was submitted for quantitative amino acid analysis under conditions in which hydrolysis was not employed.
Methionine, 8, and 9 were chosen for these preliminary analyses owing to the fact that these analogues can be detected directly by standard amino acid analysis.

3.2.4 Protein Characterization

3.2.4.1 Amino Acid Analysis and N-terminal Sequencing

The composition of each protein was determined by amino acid analysis performed by the Analytical Chemistry and Peptide/DNA Synthesis Facility at Cornell University, Ithaca, NY. N-terminal sequencing was performed by the Protein/Peptide Microanalytical Laboratory at the California Institute of Technology, Pasadena, CA.

3.2.4.2 NMR Spectroscopy

Proton NMR spectra were recorded using a Varian Inova NMR spectrometer with proton acquisition at 599.69 MHz. Samples were dissolved in D₂O under slightly acidic conditions and spectra were recorded at 25°C overnight. A simple presaturation pulse was used for water suppression. Unambiguous assignment of certain resonances for analogue 5 was confirmed using 1D TOCSY experiments.
1D TOCSY spectra were recorded on a Varian Inova NMR spectrometer with proton acquisition at 599.69 MHz. A 1D TOCSY pulse sequence with gradient-based coherence path rejection and DIPSI-2 mixing sequence (10 kHz bandwidth) was used to establish chemical shift correlations (19). This pulse sequence is designed to determine the identity of protons in a single spin system (in this case, amino acid) using selective irradiation of a single resonance followed by observation of protons to which magnetization is transferred. Selective irradiation of the resonance at 5.35 ppm (unambiguously assigned to olefin protons of 5) was achieved with a 23.67 ms e-SNOB pulse with a 60 Hz bandwidth (20). The selectivity of the pulse is demonstrated in a separate, simple 1D experiment in which the selective pulse is applied alone; no other resonances are observed in the spectrum under these conditions.

3.2.5 Activation of Methionine Analogues by MetRS in Vitro

The activation of methionine analogues 2 - 13 was measured as described in Chapter 2, but in this case, included measurement of the activity for the analogues with the mutant MetRS W305F as well. The kinetic parameters for the analogues determined as described in Section 2.2.4 were determined with W305F using the same concentrations as previously described. Assay conditions for analogues that were not analyzed
previously are given below; measurements were conducted with both the wild-type and mutant MetRS.

For analogue 8, enzyme concentrations of 60 nM and analogue concentrations of 2.5 mM to 35 mM were used in the ATP-PP\textsubscript{i} exchange assay. Analogues 4 and 12 are such poor substrates for the enzyme that only k\textsubscript{cat}/K\textsubscript{m} values could be determined accurately. For 12, enzyme concentrations of 100-150 nM and analogue concentrations of 5 mM to 50 mM were used. The lower solubility of 4 precluded the use of these high analogue concentrations, so the k\textsubscript{cat}/K\textsubscript{m} value was estimated by calculating the slope of the plot of velocity vs. substrate concentration at low substrate concentration. Enzyme concentrations of 150 nM and analogue concentrations of 0.6 mM to 6.6 mM were used. Measurable velocities could not be obtained for analogue 7, owing to the very low solubility of this analogue at the pH of the enzyme assay reaction buffer.

Additionally, ATP-PP\textsubscript{i} exchange assays were conducted on the whole cell lysates from both conventional bacterial expression hosts (CAG18491/pQE15/pREP4 and B834(DE3)/pQE15/pREP4) and modified bacterial hosts (CAG18491/pQE15-MRS/pREP4 and B834(DE3)/pQE15-MRS/pREP4) to confirm that the addition of the MetRS gene to the expression plasmid pQE15 results in an increased level of enzyme activity in the bacterial host. These assays were conducted in a manner similar to that used for analysis of the methionine analogues. M9AA medium was inoculated with the appropriate bacterial host and the culture was grown at 37°C for 8 hours. An aliquot
(3.75 mL) of each culture was centrifuged at 12000 x g to pellet the cells. The cell pellet was washed twice with 1 mL of 1 x M9 salts to remove residual amino acids from the cell pellet. The cell pellets were then resuspended in a volume (OD_{600} x 100 μL) of ATP-PP_{i} exchange assay buffer sufficient to normalize the OD_{600} of the samples to 37.5.

The resuspended cells were lysed by a single freeze-thaw cycle, vortexed, and 30 μL of the cell lysate was used directly in assays with a total volume of 75 μL and a saturating methionine concentration of 750 μM. A control assay in which methionine was excluded was also analyzed for each cell lysate to yield a background level of ATP-PP_{i} exchange that was subtracted from the velocities observed for the cell lysates supplemented with methionine. This procedure corrected for any ATP-PP_{i} exchange that might be catalyzed by the other 19 aaRS and cognate amino acids present in the cell lysate. Using a saturating concentration of methionine permits determination of the maximal velocity, which yields direct comparison of the enzyme concentrations in each cell lysate, assuming a constant value of k_{cat} between the MetRS of the different cultures.

Aliquots (20 μL) were quenched at various time points. Aliquots from the assay mixtures of the conventional bacterial host cell lysate were quenched at 1.5, 4.5, and 9.0 minutes, while those from the modified bacterial host cell lysate were quenched at 60, 90, and 180 seconds. The charcoal was washed twice with the wash buffer, and counted via liquid scintillation methods. Plots of the moles PP_{i} exchanged vs. time permitted the calculation
of the maximal velocity for each sample. Maximal velocities are reported as moles of PPᵢ exchanged per second.

3.2.6 Characterization of GFP Fluorescence

The effect of incorporation of methionine analogues on protein function was investigated by monitoring the fluorescence of the green fluorescent protein (GFP) (21,22) produced from bacterial cultures supplemented with various methionine analogues. The data were kindly provided by Pin Wang at the California Institute of Technology, Pasadena, CA, and the procedures were as follows. The 714 bp fragment encoding GFP (Clonetech, Palo Alto, CA) was amplified via PCR methods to yield a gene fragment with Bam HI/Hind III cohesive ends. This gene fragment was introduced into the multiple cloning site of the expression vector pQE31 (Qiagen, Valencia, CA) to yield the plasmid pQE31-GFP (23). This plasmid was transformed along with pREP4 into the methionine auxotroph CAG18491. The expression plasmid encodes GFP under the control of a bacteriophage T5 promoter, and encodes GFP with an N-terminal hexahistidine sequence for facile purification by metal chelate affinity chromatography. The GFP contains 6 methionine residues, all potential sites for substitution with methionine analogues. Protein expression experiments were conducted as described for the DHFR expression, except that 50 mL cultures of CAG18491/pQE31-GFP/pREP4
were grown on M9 medium supplemented with 500 mg/L of the methionine analogues 2, 3, or 9 (23).

The proteins were purified by metal chelate affinity chromatography under native conditions (24). The cell pellet was resuspended in 2 mL of sonication buffer (50 mM NaH$_2$PO$_4$, 300 mM NaCl, 10 mM imidazole, pH 8.0). After lysis with lysozyme and sonication on ice, the cell debris was pelleted and the supernate loaded onto a nickel spin column equilibrated with 0.6 mL of the sonication buffer. The protein was washed with 2 x 0.6 mL wash buffer (50 mM NaH$_2$PO$_4$, 300 mM NaCl, 20 mM imidazole, pH 8.0) and eluted with 2 x 0.2 mL elution buffer (50 mM NaH$_2$PO$_4$, 300 mM NaCl, 250 mM imidazole, pH 8.0). The protein was stored at 4°C in the elution buffer. The undiluted sample was excited at 396 nm and the intensity of the fluorescence was monitored at 506 nm using a Photon Technology International fluorescence spectrophotometer.

Fluorescence spectra were normalized for protein concentration, which was determined by monitoring the absorption maximum of the sample at 280 nm (23).
3.3 Results and Discussion

3.3.1 Genetic Manipulation of a Bacterial Expression Host

The construction of a plasmid that permits overexpression of MetRS in a bacterial expression host was achieved according to the genetic strategy outlined in Figure 3.1. The gene encoding a mutant MetRS was removed from plasmid pBSM547W305F (5) by treatment with restriction enzymes Sac I and Kpn I. The Sac I/Kpn I fragment (2449 bp) was ligated into the cloning vector pUC19-Nhelink (which was constructed to permit the cohesive ends of the W305F gene to be changed to Nhe I), to yield plasmid pUC19-W305F. A schematic of the linker inserted into pUC19 to form pUC19-Nhelink is shown in Figure 3.2, and its sequence is given in Appendix D. Restriction digestion of the pUC19-W305F with Nhe I confirmed the presence of the W305F fragment (2498 bp) (Figure 3.3). The W305F gene with Nhe I cohesive ends (Appendix E) was ligated into the unique Nhe I site of either the plasmid pQE15 or pQE9 to yield the plasmids pQE15-W305F and pQE9-W305F. Restriction analysis of each of the plasmids with Nhe I confirmed the presence of the gene encoding W305F (2498 bp, Figures 3.3 and 3.4). Restriction digestion of the pQE15-W305F plasmid with Aat II and Nhe I also confirmed the presence of the gene encoding DHFR (920 bp) in the plasmid (Figure 3.4). The orientation of the MetRS genes in the expression plasmids was confirmed by restriction
analysis with Hind III, as shown in Figures 3.3 and 3.4, and allowed determination of the plasmid identity; there are no likely consequences on the expression of the MetRS based on orientation. Both orientations of W305F were observed in the plasmid pQE9-W305F (Figure 3.3), while only the orientation shown in Figure 3.2 (corresponding to orientation 2 in the pQE9-W305F) was observed for pQE15-W305F (Figure 3.4). Transformation of the pQE15-W305F plasmid into a recA positive cell strain resulted in genetic recombination of the mutant MetRS gene with the chromosomal copy of the wild-type MetRS gene, yielding plasmid pQE15-MRS. The plasmid DNA from all cell cultures used for protein expression experiments was sequenced to confirm its identity.

These plasmids were used, along with the repressor plasmid pREP4, to transform the E. coli methionine auxotrophic strains CAG18491 or B834(DE3). Overexpression of MetRS and DHFR was confirmed by SDS-PAGE analysis of cell lysates of 8-hour cultures that had been induced to express DHFR for 4 hours. The presence of the band indicating expression of MetRS is clearly visible along with the band corresponding to DHFR in cultures of CAG18491/pQE15-MRS/pREP4, as shown in Figure 3.5. Similar results confirming the expression of W305F were observed for overnight cultures of CAG18491/pQE15-W305F/pREP4 and CAG18491/pQE9-W305F/pREP4 (no target protein expression could be monitored for the hosts containing pQE9, as this plasmid contains only a multiple cloning site).
The data illustrated in Figure 3.6, obtained from ATP-PP\textsubscript{i} exchange assays of cell lysates of CAG18491/pQE15-MRS/pREP4 and B834(DE3)/pQE15-MRS/pREP4, confirm the results in Figure 3.5 and indicate that the modified bacterial host CAG18491/pQE15-MRS/pREP4 exhibits a $V_{\text{max}}$ value for methionine activation approximately 50-fold higher than that observed for the control host CAG18491/pQE15/pREP4 (Figure 3.6a). A 30-fold increase was observed for the modified host B834(DE3)/pQE15-MRS/pREP4 in comparison to the control host B834(DE3)/pQE15/pREP4 (Figure 3.6b). These values are consistent with the reported copy number of the pQE15 plasmid (approximately 50) (16).

3.3.2 Translational Activity of Methionine Analogues in a Modified Bacterial Expression Host

Bacterial hosts capable of overexpressing MetRS or W305F were produced by transforming the *E. coli* methionine auxotrophs CAG18491 or B834(DE3) with the repressor plasmid pREP4 and an expression plasmid that encodes truncated forms of either the wild-type MetRS or a mutant MetRS W305F, as described in the previous section. These hosts were used to test the translational activity of methionine analogues. The results obtained for both CAG18491 and B834(DE3) cultures were identical; results for CAG18491 are shown and discussed below.
Methionine analogues 2 - 13 were tested for translational activity in both conventional and modified bacterial expression hosts. Cultures of CAG18491/pQE15/pREP4 or CAG18491/pQE15-MRS/pREP4 in M9AA media were grown to an optical density of 0.90, and the cells were sedimented by centrifugation. Cells were washed three times with M9 salts and resuspended to an optical density of 0.90 in M9 test media containing 19 amino acids plus 1) neither methionine nor analogue (negative control); 2) methionine (60 mg/L, positive control); or 3) an analogue of interest (60 mg/L). Expression of DHFR was induced by addition of 0.4 mM isopropyl β-D-thiogalactopyranoside (IPTG), and protein synthesis was monitored after 4.5 hours. Expression of DHFR was monitored by SDS-polyacrylamide gel electrophoresis (SDS-PAGE); accumulation of target protein was taken as evidence for translational activity of the methionine analogue. The results from these experiments are summarized in Figure 3.7.

The target protein was not observed in the negative control culture of CAG18491/pQE15/pREP4, or in cultures of CAG18491/pQE15/pREP4 or CAG18491/pQE15-MRS/pREP4 supplemented with 4, 6, 7, 8, 10, 12, or 13. In contrast, DHFR was detected in both bacterial host cultures supplemented with methionine (1), 2, 3, or 9, as indicated by the appearance of a protein band at the position expected for DHFR in SDS-PAGE.
For the negative control cultures and for cultures supplemented with 5 and 11, however, the behavior of the bacterial hosts differed, as shown in Figure 3.7. The target protein DHFR was not detected in the CAG18491/pQE15/pREP4 cultures supplemented with 5 or 11, while strong induction of DHFR was observed for CAG18491/pQE15-MRS/pREP4 under the same conditions. Even the unsupplemented control culture of CAG18491/pQE15-MRS/pREP4 shows evidence of DHFR synthesis, suggesting that introduction of pQE15-MRS does indeed increase the rate of activation of methionine in the modified host. Results identical to those obtained for CAG18491/pQE15-MRS/pREP4 were observed for bacterial cultures that were equipped with the mutant MetRS W305F (CAG18491/pQE15-W305F/pREP4).

The yields of DHFR obtained from 1-L cultures of CAG18491/pQE15-MRS/pREP4 supplemented with 20 mg/L of 5 were approximately 8 mg/L. Consistent with the fact that 11 is a poorer substrate for MetRS than 5, 60 mg/L supplementation by 11 was required to observe measurable accumulation of target protein. Because supplementation by 11 at a concentration of 500 mg/L improved the yield of protein in 5-mL cultures (Figure 3.8), the yield for 1-L cultures was determined at this level of supplementation and was found to be 35 mg/L. Purified DHFR containing 5 or 11 was analyzed by amino acid analysis, N-terminal sequencing, and, when possible, NMR spectroscopy (see Section 3.3.6).
3.3.3 Protein Yields from Modified Bacterial Expression Hosts

Overexpression of MetRS in a bacterial host has expanded the number of methionine analogues that can be incorporated into proteins *in vivo*. Coupled with the results from Chapter 2 (which indicate that the activation of methionine analogues by MetRS controls the yields of target protein obtained from a conventional bacterial host), these results suggested that protein yields obtained from bacterial cultures supplemented with methionine analogues might be improved by increasing the MetRS activity of the bacterial host. To test this hypothesis, we compared the yields of protein prepared in the conventional bacterial expression host, CAG18491/pQE15/pREP4, to those obtained from a modified host, CAG18491/pQE15-MRS/pREP4.

Protein synthesis was monitored for 5-ml cultures of these hosts supplemented with 20 mg/L of methionine or analogues 2, 3, or 9; Western blot analyses of protein synthesis are shown in Figure 3.9. Although very low levels of protein synthesis were observed for negative control cultures of CAG18491/pQE15-MRS/pREP4, amino acid analyses, N-terminal sequencing, and NMR analyses of proteins produced in cultures of this host supplemented with 5 (a poorer substrate for MetRS than 2, 3, or 9) still show 90-96% replacement of methionine by 5 (25). Thus, the level of protein synthesis shown in Figure 3.9 results from the incorporation of the analogue and is not likely due to incorporation of residual methionine; these results were confirmed by amino acid
analysis. For cultures supplemented with methionine or 3, the modified host, CAG18491/pQE15-MRS/pREP4, does not exhibit higher levels of protein synthesis than the conventional host CAG18491/pQE15/pREP4. Analysis by laser densitometry confirms these results and reveals approximately equal accumulation of target protein for both strains; identical results have been obtained for yields of purified DHFR obtained from large-scale expressions. Activation of the analogue by MetRS therefore does not appear to limit protein synthesis in cultures supplemented with 3.

For cultures supplemented with 2 or 9, however, the modified bacterial host exhibits significantly increased levels of protein synthesis in comparison with the conventional host (Figure 3.9). Laser densitometry analysis indicates that the level of protein synthesis in the modified host is increased approximately 1.5-fold over that in the conventional host for cultures supplemented with 2, and approximately 1.4-fold for cultures supplemented with 9. Activation of these analogues by MetRS appears to limit protein synthesis in the conventional host, such that increasing the MetRS activity of the host is sufficient to restore high levels of protein synthesis. Indeed, as shown in Figure 3.10, the yield of DHFR obtained from large-scale cultures of CAG18491/pQE15-MRS/pREP4 supplemented with 2 or 9 is increased to approximately 35 mg/L (as compared to yields of 10 and 20 mg/L, respectively, from cultures of CAG18491/pQE15/pREP4, (Table 2.1)). The results indicate that simple overexpression of MetRS can improve protein yields for cultures supplemented with methionine.
analogues that are poor substrates for MetRS, and may provide an attractive general method for efficient production of chemically novel protein materials \textit{in vivo}.

3.3.4 \textbf{Incorporation of Additional Methionine Analogues into Proteins \textit{in Vivo}}

Overexpression of MetRS has been indicated not only to expand the range of methionine analogues that can be incorporated into proteins \textit{in vivo}, but also to improve the yields of protein obtained from cultures of conventional bacterial hosts grown in media supplemented with poor substrates of MetRS. The improvement in incorporation and protein yield most likely results from the increase in the rate of analogue activation afforded by overexpression of the enzyme in the bacterial host in cultures supplemented with 20 mg/L of the analogue. For 11, it was observed that raising the supplementation to at least 250 mg/L in cultures of the modified bacterial host also improved protein yield (Figure 3.8), which suggested that the concentration of methionine analogues in the cell could be increased by raising their concentration in the medium. Therefore, the importance of the level of amino acid supplementation on the incorporation of analogues into proteins \textit{in vivo} was probed.

The expression of the target protein DHFR was monitored by SDS-PAGE analysis for both conventional (CAG18491/pQE15/pREP4) and modified bacterial expression hosts (CAG18491/pQE15-MRS/pREP4) with supplementation at 500 mg/L of

90
4, 6, 7, 8, 10, 12, and 13 (3.8 to 4.3 mM). These analogues were chosen owing to their inability to support protein biosynthesis in our previous experiments. The results from the SDS-PAGE analysis of these cultures are shown in Figure 3.11. Negative control cultures did not show any target protein synthesis, while both positive control cultures supported protein synthesis in vivo. For the cultures of the conventional bacterial host (Figure 3.11 (a)), none of the above analogues supported protein synthesis in vivo. For the modified bacterial host CAG18491/pQE15-MRS/pREP4 (Figure 3.11 (b)), however, the accumulation of target protein was observed in cultures supplemented with 500 mg/L of analogues 4, 7, 8, or 12. Increasing the level of amino acid in the medium, coupled with overexpression of the MetRS, is required for these methionine analogues to support protein biosynthesis; their activation appears to be the rate-limiting step in protein biosynthesis. Identical studies were conducted with bacterial cultures of CAG18491/pQE15-W305F/pREP4 to ascertain if overexpression of the mutant MetRS would rescue the translational activity of any additional analogues. No additional analogues were utilized by the bacterial host equipped with extra copies of the mutant MetRS.

Proteins produced by CAG18491/pQE15-MRS/pREP4 cultures (50 mL) supplemented with 500 mg/L of the methionine analogues 4, 7, 8, and 12 were purified and analyzed by amino acid analysis and N-terminal sequencing to confirm the incorporation of these methionine analogues into the target protein DHFR (Section 3.3.6).
Characterization of the ATP-PP$_2$ exchange activity supported by the analogues was also revisited (Section 3.3.7).

3.3.5 **Intracellular Concentrations of Methionine Analogues**

The intracellular concentrations of methionine analogues 8 and 9 were preliminarily investigated to determine to what extent the intracellular concentration of the analogues is increased upon the increase in supplementation. Cultures of CAG18491/pQE15-MRS/pREP4 (250 mL), supplemented with 20 mg/L or 500 mg/L of methionine, 8, or 9, were grown as previously described. Samples (50 mL) were collected in duplicate after two hours of protein expression, and the concentration of free amino acid in the cell lysate was determined by amino acid analysis as described in the experimental section. The choice of 8 and 9 as analogues in these preliminary investigations was based on the ease of their quantitative determination under standard amino acid analysis conditions. As shown in Table 3.1, the concentrations of the analogues in the medium increased by a factor of 5 to 10 upon increasing the supplementation from 20 mg/L to 500 mg/L. The concentrations of the analogues in cultures supplemented with 4.3 mM of 8 or 3.8 mM of 9 were approximately 650 μM and 200 μM, respectively. These values are higher than the reported intracellular concentration of methionine in cellular hosts grown on medium lacking any amino acids.
(~ 10 μM) (18), but are consistent with the concentration of methionine observed in cultures supplemented with 500 mg/L (3.4 mM) methionine (350 μM, Table 3.1). This indicates that their transport is not inhibited relative to that of methionine and may suggest that the analogues are transported into the cell by a similar transport mechanism.

3.3.6 Protein Characterization

3.3.6.1 Amino Acid Analysis and N-terminal Sequencing

Increasing MetRS activity by overexpression of MetRS in a bacterial host is necessary and sufficient to observe translational activity of 5 and 11 under convenient conditions in vivo. Further increasing the rate of analogue activation by concomitantly elevating the concentration of analogue in the medium rescues the translational activity of 4, 7, 8, and 12. Amino acid analysis of proteins produced from cultures supplemented with 500 mg/L of these analogues was conducted to confirm replacement of methionine. As demonstrated in Table 3.2, amino acid analysis shows a decrease in methionine content from the expected value of 3.8 mol%. It is impossible to detect 4, 5, 7, 11, or 12 directly by amino acid analysis, owing to their instability under the analysis conditions. If, however, depletion of methionine is assumed to result from replacement by the analogue, the observed analysis corresponds to overall extents of incorporation for the
analogues ranging from 60% to 98%, as shown in Table 3.3. The amino acid profile given in Table 3.2 also demonstrates that even at high levels of analogue supplementation, the experimentally observed amino acid content is within experimental error (±10% of the expected value) of the theoretical values for DHFR, with the exception of the decrement in methionine content.

Retention of the N-terminal (initiator) methionine in DHFR is expected on the basis of the identity of the penultimate amino acid (26), so N-terminal sequencing provided an additional means of assessing the extent of replacement of methionine by the analogues. Because the analogues are not degraded under the analysis conditions, they can be detected directly, and this analysis provides direct evidence for the presence of the analogues in the target protein. Comparison of chromatograms of the N-terminal residues of DHFR and those for DHFR containing methionine analogues (Figures 3.12, 3.13, 3.14, 3.15, 3.16, and 3.17) demonstrates that the methionine that normally occupies the initiator position of DHFR is replaced with the analogues. The signal corresponding to methionine elutes at approximately 13.0 minutes, while those corresponding to the analogue are listed in the appropriate figure caption. The large peaks that elute at approximately 14.7 minutes correspond to piperidylphenylthiourea (pptu), a product of the analysis resulting from the buffer, and the small peak at approximately 18.5 minutes corresponds to diethylphthalate (diet), an internal standard. These results clearly indicate the incorporation of a variety of methionine analogues at the initiator site of DHFR.
Integration of the peak areas corresponding to methionine and the analogues indicates up to quantitative replacement of methionine by the analogues, with results given for each analogue in Table 3.3. These results are consistent with the amino acid analyses and suggest that analogues 5, 11, 8, 12, 4, and 7 are increasingly poor substrates for MetRS.

3.3.6.2 NMR Spectroscopy

A direct assessment of the extent of incorporation of 5 into DHFR is provided by NMR spectroscopy. (Similar results could not be obtained for the other analogues due to the lack of resonances distinct from those of DHFR.) Comparisons of the 600 MHz proton NMR spectra (Figure 3.18) of DHFR, 5, and DHFR-5 indicate the appearance, in the DHFR-5 spectrum (Figure 3.18c), of the vinylene protons of 5 at δ5.35 (δ-CH) and δ5.60-5.70 (γ-CH). The resonances at δ5.35 and δ5.70 occur at the same chemical shift values as in free 5 and are clearly due to incorporation of 5 into DHFR. That the resonance at 5.60 ppm arises from the γ-CH vinylene proton of 5 is suggested by the fact that the integrated intensity of the resonance at δ5.35 equals the sum of the integrations of the resonances at δ5.60 and δ5.70.

This assignment is confirmed by 1D TOCSY (Total Correlation Spectroscopy) experiments that indicate that the protons at both δ5.60 and δ5.70 are members of the
same spin system (and therefore the same amino acid) as those at δ5.35 (The TOCSY spectra are shown in Appendix C). More importantly, the 1D TOCSY experiments also show that the protons at δ5.35 (and therefore those at δ5.60 and δ5.70) are associated with the spin system of the entire side chain of 5: (¹H-NMR (D₂O): δ 1.6 (d, 3H, J = 6.4 Hz, CH₃-CH=CH-CH₂), 2.50 (t, 2H, J = 5.8, 13 Hz, CH₂-CH=CH-CH₂), 3.70 (t, 1H, J = 5.8 Hz, H₂N-CH-COOH), 5.30-5.43 (m, 1H, Jₑ = 14.7, J = 6.4 Hz, CH₃-CH=CH-CH₂), 5.70 (m, 1H, Jₑ = 14.7, J = 13 Hz, CH₂-CH=CH=CH₂) ppm)

Selective irradiation of the resonance at δ5.35 followed by immediate observation shows only the resonance at δ5.35. Observation after a mixing time of 20 ms, however, shows the protons at δ5.60 and δ5.70, indicating that those protons are members of the same spin system (and therefore the same amino acid residue) as those corresponding to the resonance at δ5.35. The protons of the side chain β and ε carbons are also observed after 20 ms mixing time, at chemical shift values characteristic of the free amino acid (2.5 ppm (β-CH₂) and 1.6 ppm (ε-CH₃)).

Integration of the spectrum suggests that 5 of the 8 methionine positions occupied by 5 are represented by the resonance at δ5.60; these protons must reside in a magnetically-distinct environment from the protons at δ5.70. These results unequivocally demonstrate the translational activity of 5 in the host strain outfitted with:

96
elevated MetRS activity. Integration of the NMR spectrum indicates 90 ± 6% replacement of methionine.

3.3.7 Activation of Methionine Analogues by MetRS in Vitro

Activation of methionine analogues 4, 7, 8, and 12 by MetRS was assayed by the ATP-PP$_i$ exchange assay as previously described. Previous characterization of the exchange supported by these analogues (Figure 2.3) indicated that none of them supported measurable PP$_i$ exchange at 20 minutes at an analogue concentration of 5 mM. In these experiments, analogues at concentrations of 15 mM were incubated with the enzyme for 3.5 hours prior to quenching of a 20 µL aliquot of the assay mixture. As indicated in Figure 3.19, each of the analogues supports levels of PP$_i$ exchange greater than the background under these experimental conditions, confirming the role of MetRS in the incorporation of these analogues into proteins in vivo.

Quantitative assessment of the catalytic efficiency by which MetRS activates these analogues was achieved by determining the kinetic parameters of these analogues as previously described. The analogues are increasingly poor substrates for the enzyme as compared to analogues 5 and 11 (Table 3.4), with $k_{cat}/K_m$ values up to 340,000-fold lower than those observed for methionine. As discussed for 5 and 11 (Chapter 2), the loss of catalytic activity manifest both in increasing values of $K_m$ and decreasing values
of $k_{cat}$ indicates that these analogues are less efficient than methionine in both the binding and catalytic events of the enzyme. These results also confirm that, under normal cellular conditions (i.e., normal expression levels of MetRS and lower amino acid supplementation), activation of analogues 4, 7, 8, and 12 appears to be the rate limiting step in protein biosynthesis and prevents the accumulation of measurable amounts of target protein.

Increasing the MetRS activity of the cellular host and increasing the concentration of analogue in the medium increases the rate of activation of analogues 4, 7, 8, and 12 sufficiently to rescue their translational activity in vivo. Results in Figure 2.4 indicate that the catalytic efficiency of analogue activation can be reduced approximately 2000-fold relative to that of methionine (to a $k_{cat}/K_m$ value of approximately $3.0 \times 10^{-4} \text{ s}^{-1} \text{ M}^{-1}$) prior to the loss of the ability of an analogue to support protein synthesis under a standard set of experimental conditions. The observation of translational activity for 4, 7, 8, and 12 in cellular hosts in which MetRS activity is increased 50-fold and intracellular analogue concentration is likely increased 5-10-fold is consistent with these previous observations, as the rates of activation are effectively raised 250-500 fold under these conditions. Even for the poorest substrate measured for MetRS (12, which has a $k_{cat}/K_m$ value of $1.6 \times 10^{-6} \text{ s}^{-1} \text{ M}^{-1}$), the rate of activation would be effectively increased above the $k_{cat}/K_m$ threshold of $3.0 \times 10^{-4} \text{ s}^{-1} \text{ M}^{-1}$ as a result of the increase in activity and
supplementation, which would rescue the ability of the analogue to support protein synthesis in vivo.

The importance of side chain length in facilitating activation of an analogue by MetRS is again illustrated, as 7, 8, and 12 all have side chains of different length than methionine and are much less efficiently activated by MetRS. Although 4 has the same side chain length as methionine, the conformational restriction placed on the side chain by the cis-conformation of the double bond imposes on 4 a geometry that likely does not bind well at the active site of MetRS. The presence of π-electrons in the side chains of 4 and 12 may also play a role in their activation by MetRS. The lack of observed translational activity for 10 likely results both from the fact that the side chain is longer than that of methionine, and because the electron density is located at the γ-position of the side chain instead of near the δ-position. The shorter side-chain length and linear geometry of 13 may restrict the accessibility of the π electrons of the triple bond and limit binding of this analogue at the MetRS active site. It is surprising, however, that the translational activity of 6 is not rescued under any of the in vivo assay conditions reported here, given the translational activity of both 9 and trifluoromethionine. Perhaps the hydrogen bonding of the thioether sulfur of trifluoromethionine in the active site of MetRS permits accommodation of the electronegative fluorine groups, while the lack of those interactions for the side chain of 6 precludes activation of the analogue by MetRS.
Taken together with the studies in Chapter 2, these results corroborate the importance both of side chain length and of electron density near the δ-position of the analogue side chain in activation of the analogues by MetRS.

Quantitative assessment of the kinetic parameters for activation of the analogues by W305F was also conducted. Consistent with the in vivo results, the mutant MetRS W305F is only a factor of 4 more efficient in activation than the wild-type MetRS, as delineated in Table 3.5. Nevertheless, this small improvement in catalytic efficiency may point to the potential for altering the MetRS activity through directed evolution as a means to further expand the substrate range of the translational apparatus.

3.3.8 Characterization of GFP Fluorescence

The effect on protein function by replacement of methionine by methionine analogues was preliminarily investigated by monitoring the fluorescence of green fluorescent protein (GFP) produced from bacterial cultures supplemented with various methionine analogues. Bacterial cultures CAG18491/pQE31-GFP/pREP4 were obtained by transformation of the E. coli methionine auxotroph CAG18491 with the expression plasmid pQE31-GFP and the repressor plasmid pREP4. The expression plasmid pQE31-GFP encodes GFP with an N-terminal hexahistidine sequence under the control of a bacteriophage T5 promoter. Cultures (50 mL) of the bacterial host were grown in M9
medium lacking methionine and supplemented with analogues 2, 3, or 9. The GFP was purified by metal chelate affinity chromatography under native conditions, and the fluorescence measured from the elution buffer solutions is shown in Figure 3.20.

As can be seen in Figure 3.20, GFP produced from cultures supplemented with 500 mg/L 2, 3, or 9 has an emission spectrum that is similar to that observed for GFP produced from cultures supplemented with the same level of methionine. While the emission maximum is observed at 506 nm in all of the spectra, it is obvious from the normalized spectra shown in Figure 3.20 that the intensity of the emission peak is significantly reduced in the GFP containing the methionine analogues 2, 3, or 9. This suggests that the replacement of methionine does not directly alter the chemical nature of the chromophore, but causes changes in protein structure that reduce the fluorescence of the protein.

Regardless, these results demonstrate that, even at high levels of substitution, the incorporation of non-natural amino acids does not necessarily result in a complete loss of structure or function. There have been a variety of literature reports in which biological activity of proteins is retained after substitution by amino acid analogues (27-32); in fact, selenomethionine supports cell growth (28). Certainly the retention or loss of protein function upon incorporation of non-natural amino acids must be assessed on a case-by-case basis. The effect of analogue incorporation on protein function will depend on the natural amino acid being replaced, the chemical nature of the non-natural amino acid, and
the number, positions, and importance of the natural amino acid in the structure and function of the target protein.

3.4 Summary and Conclusions

The incorporation of 4, 5, 7, 8, 11, and 12 into proteins in vivo constitutes the first example of broadening the amino acid substrate range of the E. coli translational apparatus via overproduction of wild-type MetRS in a bacterial host. (Overexpression of a mutant MetRS W305F provided little additional advantage.) Under an appropriate set of experimental conditions employing overexpression, analogues 5 and 11, which are not incorporated by a conventional bacterial host, replace methionine at levels exceeding 96%. Increasing the analogue supplementation to 500 mg/L in cultures of the modified bacterial host permits incorporation of a larger set of methionine analogues (4, 7, 8, and 12) with replacement levels up to 92%. The poorest of these substrates is approximately 340,000-fold poorer a substrate for MetRS during activation than methionine, which suggests that, under appropriate experimental conditions, much poorer substrates for MetRS than previously imagined can be utilized by the protein biosynthesis machinery. As a result of the versatile chemistry of unsaturated groups (33-39), the incorporation of the unsaturated analogues 4, 5, 11, and 12 may provide new opportunities for macromolecular synthesis through protein engineering. More generally, these results
indicate that this simple strategy—overexpression of wild-type aaRS—may be used to modify proteins by incorporation of non-natural amino acids that are poor substrates for aaRS and that would be essentially inactive in conventional expression hosts.

Our results also demonstrate that increasing the MetRS activity of a bacterial expression host improves the yields of proteins containing methionine analogues that are poor substrates for MetRS. Overexpression of aaRS may therefore also be a general strategy for improving yields of proteins containing other non-natural amino acids, as well as proteins rich in particular natural amino acids. Manipulation of the aaRS activities of a bacterial host therefore has enormous potential for broadening the scope of protein engineering by permitting production of natural and artificial proteins with novel chemical and physical properties.
Table 3.1  Intracellular amino acid concentrations as determined by amino acid analysis.

<table>
<thead>
<tr>
<th>Analogue</th>
<th>Level of supplementation</th>
<th>Intracellular concentration, µM*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>mg/L (µM)</td>
<td></td>
</tr>
<tr>
<td>Met</td>
<td>0</td>
<td>35 ± 5</td>
</tr>
<tr>
<td>Met</td>
<td>20 (135)</td>
<td>58 ± 7</td>
</tr>
<tr>
<td>Met</td>
<td>500 (3400)</td>
<td>350 ± 40</td>
</tr>
<tr>
<td>8</td>
<td>20 (170)</td>
<td>78 ± 8</td>
</tr>
<tr>
<td>8</td>
<td>500 (4300)</td>
<td>600 ± 50</td>
</tr>
<tr>
<td>9</td>
<td>20 (150)</td>
<td>48 ± 10</td>
</tr>
<tr>
<td>9</td>
<td>500 (3800)</td>
<td>220 ± 20</td>
</tr>
</tbody>
</table>

* Average values from duplicate analyses are reported.
Table 3.2  Amino acid analysis results for DHFR containing methionine analogues.

<table>
<thead>
<tr>
<th>Amino acid</th>
<th>Theoretical mole %</th>
<th>DHFR-Met</th>
<th>DHFR-Ana*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ala</td>
<td>2.9</td>
<td>3.6</td>
<td>3.6</td>
</tr>
<tr>
<td>Cys</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Asx</td>
<td>7.1</td>
<td>7.6</td>
<td>7.6</td>
</tr>
<tr>
<td>Glx</td>
<td>11.9</td>
<td>12.3</td>
<td>12.1</td>
</tr>
<tr>
<td>Phe</td>
<td>4.3</td>
<td>4.4</td>
<td>4.6</td>
</tr>
<tr>
<td>Gly</td>
<td>7.6</td>
<td>8.5</td>
<td>8.5</td>
</tr>
<tr>
<td>His</td>
<td>3.8</td>
<td>3.9</td>
<td>3.8</td>
</tr>
<tr>
<td>Ile</td>
<td>6.7</td>
<td>5.6</td>
<td>5.9</td>
</tr>
<tr>
<td>Lys</td>
<td>8.0</td>
<td>8.1</td>
<td>8.2</td>
</tr>
<tr>
<td>Leu</td>
<td>9.5</td>
<td>9.6</td>
<td>9.8</td>
</tr>
<tr>
<td>Met</td>
<td>3.8</td>
<td>3.7</td>
<td>0.2</td>
</tr>
<tr>
<td>Pro</td>
<td>6.2</td>
<td>6.4</td>
<td>6.6</td>
</tr>
<tr>
<td>Arg</td>
<td>5.7</td>
<td>6.0</td>
<td>6.1</td>
</tr>
<tr>
<td>Ser</td>
<td>7.6</td>
<td>7.8</td>
<td>8.0</td>
</tr>
<tr>
<td>Thr</td>
<td>2.9</td>
<td>3.3</td>
<td>3.1</td>
</tr>
<tr>
<td>Val</td>
<td>6.7</td>
<td>6.3</td>
<td>6.4</td>
</tr>
<tr>
<td>Trp</td>
<td>1.9</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Tyr</td>
<td>2.9</td>
<td>2.9</td>
<td>2.9</td>
</tr>
</tbody>
</table>

* Ana = analogue. Average values from duplicate analyses of each of 6 DHFR-Ana samples are reported. (DHFR was produced from cultures supplemented with 4, 5, 7, 8, 11, and 12.)
Table 3.3  Extent of replacement of methionine by methionine analogues.

<table>
<thead>
<tr>
<th>Analogue</th>
<th>Amino Acid Analysis</th>
<th>N-terminal Sequencing</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>92 ± 3</td>
<td>97 ± 2</td>
</tr>
<tr>
<td>11</td>
<td>98 ± 2</td>
<td>97 ± 2</td>
</tr>
<tr>
<td>8</td>
<td>92 ± 3</td>
<td>87 ± 2</td>
</tr>
<tr>
<td>12</td>
<td>84 ± 5</td>
<td>84 ± 2</td>
</tr>
<tr>
<td>4</td>
<td>76 ± 5</td>
<td>79 ± 5</td>
</tr>
<tr>
<td>7</td>
<td>60 ± 5</td>
<td>56 ± 5</td>
</tr>
</tbody>
</table>
Table 3.4  Kinetic parameters for activation of methionine analogues by MetRS.

<table>
<thead>
<tr>
<th>Analogue</th>
<th>( K_m ), ( \mu M )</th>
<th>( k_{cat} ), s(^{-1} )</th>
<th>( k_{cat}/K_m ), s(^{-1} \mu M^{-1} )</th>
<th>Relative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Met</td>
<td>24.3</td>
<td>13.3</td>
<td>0.547</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>2415</td>
<td>2.61</td>
<td>1.08 \times 10^3</td>
<td>1/500</td>
</tr>
<tr>
<td>9</td>
<td>4120</td>
<td>2.15</td>
<td>5.22 \times 10^4</td>
<td>1/1050</td>
</tr>
<tr>
<td>2</td>
<td>4555</td>
<td>1.35</td>
<td>2.96 \times 10^4</td>
<td>1/1850</td>
</tr>
<tr>
<td>5</td>
<td>15675</td>
<td>1.82</td>
<td>1.16 \times 10^4</td>
<td>1/4700</td>
</tr>
<tr>
<td>11</td>
<td>38560</td>
<td>1.51</td>
<td>3.91 \times 10^{-5}</td>
<td>1/14000</td>
</tr>
<tr>
<td>8</td>
<td>14380</td>
<td>0.17</td>
<td>1.18 \times 10^{-5}</td>
<td>1/45600</td>
</tr>
<tr>
<td>4</td>
<td>N/A</td>
<td>N/A</td>
<td>4.31 \times 10^{-6}</td>
<td>1/127200</td>
</tr>
<tr>
<td>12</td>
<td>N/A</td>
<td>N/A</td>
<td>1.58 \times 10^{-6}</td>
<td>1/342000</td>
</tr>
<tr>
<td>7</td>
<td>ND*</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
</tbody>
</table>

* Not determined
Table 3.5  Comparison of catalytic efficiencies for the activation of methionine analogues by MetRS and W305F.

<table>
<thead>
<tr>
<th>Analogue</th>
<th>MetRS $k_{cat}/K_m$, s$^{-1}$ µM$^{-1}$</th>
<th>Relative</th>
<th>W305F $k_{cat}/K_m$, s$^{-1}$ µM$^{-1}$</th>
<th>Relative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Met</td>
<td>0.547</td>
<td>1</td>
<td>1.06</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>$1.08 \times 10^{-3}$</td>
<td>1/500</td>
<td>$4.18 \times 10^{-3}$</td>
<td>1/260</td>
</tr>
<tr>
<td>9</td>
<td>$5.22 \times 10^{-4}$</td>
<td>1/1050</td>
<td>$1.84 \times 10^{-3}$</td>
<td>1/580</td>
</tr>
<tr>
<td>2</td>
<td>$2.96 \times 10^{-4}$</td>
<td>1/1850</td>
<td>$1.27 \times 10^{-3}$</td>
<td>1/850</td>
</tr>
<tr>
<td>5</td>
<td>$1.16 \times 10^{-4}$</td>
<td>1/4700</td>
<td>$4.35 \times 10^{-4}$</td>
<td>1/2400</td>
</tr>
<tr>
<td>11</td>
<td>$3.91 \times 10^{-5}$</td>
<td>1/14000</td>
<td>$1.47 \times 10^{-4}$</td>
<td>1/7200</td>
</tr>
<tr>
<td>8</td>
<td>$1.18 \times 10^{-5}$</td>
<td>1/45600</td>
<td>$3.74 \times 10^{-5}$</td>
<td>1/28400</td>
</tr>
<tr>
<td>4</td>
<td>$4.31 \times 10^{-6}$</td>
<td>1/127200</td>
<td>$2.42 \times 10^{-5}$</td>
<td>1/43800</td>
</tr>
<tr>
<td>12</td>
<td>$1.58 \times 10^{-6}$</td>
<td>1/342000</td>
<td>$1.83 \times 10^{-5}$</td>
<td>1/58000</td>
</tr>
</tbody>
</table>
Figure 3.1  Genetic strategy for increasing the MetRS activity in the bacterial host. (a) The plasmid pUC19 is linearized with restriction digestion by Aat II and Sac I, and the linker insert with these cohesive ends is ligated into the plasmid to yield pUC19-Nhe. (b) The pUC19-Nhe is linearized by digestion with Kpn I and Sac I, and the W305F gene fragment with the same cohesive ends is ligated into the plasmid to yield pUC19-W305F. (c) The W305F gene is removed from pUC19-W305F by digestion with Nhe I and ligated into the unique Nhe I site of pQE15 to yield pQE15-W305F.
Figure 3.2  Schematic of Nhe linker designed to change the identity of the cohesive ends of the gene encoding W305F into *Nhe* I. The W305F fragment can be removed from its original plasmid pBSM547W305F by restriction digestion with *Kpn* I and *Sac* I and inserted in between those sites in this linker. The W305F fragment can then be removed from the linker with restriction digestion with *Nhe* I and ligated into the unique *Nhe* I site in the plasmids pQE9 and pQE15.
Figure 3.3  Restriction digestion results for the construction of plasmid pQE9-W305F.
<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Digest</th>
<th>Expected</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lane 1:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MW markers</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lane 2:</td>
<td>pQE15</td>
<td>uncut</td>
</tr>
<tr>
<td>Lane 3:</td>
<td>pQE15/pREP4</td>
<td>Nhe I</td>
</tr>
<tr>
<td>Lane 4:</td>
<td>pQE15/pREP4</td>
<td>Hind III</td>
</tr>
<tr>
<td>Lane 5:</td>
<td>pQE15-W305F/pREP4</td>
<td>uncut</td>
</tr>
<tr>
<td>Lane 6:</td>
<td>pQE15-W305F/pREP4</td>
<td>Nhe I</td>
</tr>
<tr>
<td>Lane 7:</td>
<td>pQE15-W305F/pREP4</td>
<td>Hind III</td>
</tr>
<tr>
<td>Lane 8:</td>
<td>pQE15/pREP4</td>
<td>Aat II/Nhe I</td>
</tr>
<tr>
<td>Lane 9:</td>
<td>pQE15-W305F/pREP4</td>
<td>Aat II/Nhe I</td>
</tr>
<tr>
<td>Lane 10:</td>
<td>W305F</td>
<td>Nhe I</td>
</tr>
</tbody>
</table>

Figure 3.4  Restriction digestion results for plasmid pQE15-W305F.
Figure 3.5  SDS-PAGE analysis of the expression of DHFR and MetRS in CAG18491/pREP4 bacterial hosts equipped with (a) pQE15 or (b) pQE15-MRS.
Figure 3.6  Activation rates of methionine by whole cell lysates, (a) CAG18491/pREP4 and (b) B834(DE3)/pREP4, equipped with the expression plasmids pQE15 or pQE15-MRS (MRS).
Figure 3.7  SDS-PAGE analysis of DHFR expression in bacterial cultures supplemented with methionine analogues as indicated.  (a) The conventional bacterial host CAG18491/pQE15/pREP4.  (b) The modified bacterial host CAG18491/pQE15-MRS/pREP4.
Figure 3.8  Western blot analysis of DHFR yield from modified bacterial host CAG18491/pQE15-MRS/pREP4 cultures supplemented with 11 (L-isomer) at concentrations of (a) 60 mg/L, (b) 250 mg/L, and (c) 500 mg/L.
Figure 3.9 Western blot analysis of DHFR synthesis by bacterial expression hosts CAG18491/pQE15/pREP4 (pQE15) and CAG18491/pQE15-MRS/pREP4 (MRS). Bacterial cultures were supplemented with methionine, 2, 3, or 9.
Figure 3.10  Yields of purified DHFR obtained from 1-liter cultures supplemented with methionine, 2, 3, or 9. The solid bars represent yields from CAG18491/pQE15/pREP4 cultures, while the striped bars represent yields from CAG18491/pQE15-MRS/pREP4 cultures.
Figure 3.11  SDS-PAGE analysis of DHFR synthesis in bacterial cultures supplemented with methionine analogues at 500 mg/L. Cultures were supplemented with methionine, 4, 6, 7, 8, 10, 12, and 13 (the analogues that did not support protein synthesis at lower levels of supplementation), as indicated. (a) CAG18491/pQE15/pREP4. (b) CAG18491/pQE15-MRS/pREP4.
Figure 3.12  N-terminal sequencing results indicating occupancy of the initiator site in DHFR produced in bacterial cultures supplemented with 4, which elutes at 14.8 minutes. N-terminal residues are shown for (a) DHFR-Met, (b) the free amino acid 4, and (c) DHFR-4, as determined by Edman degradation.
Figure 3.13  N-terminal sequencing results indicating occupancy of the initiator site in DHFR produced in bacterial cultures supplemented with \(5\), which elutes at 16.0 minutes. N-terminal residues are shown for (a) DHFR-Met, (b) the free amino acid \(5\), and (c) DHFR-5, as determined by Edman degradation.
Figure 3.14  N-terminal sequencing results indicating occupancy of the initiator site in DHFR produced in bacterial cultures supplemented with 7, which elutes at 20.8 minutes. N-terminal residues are shown for (a) DHFR-Met, (b) the free amino acid 7, and (c) DHFR-7, as determined by Edman degradation.
Figure 3.15  N-terminal sequencing results indicating occupancy of the initiator site in DHFR produced in bacterial cultures supplemented with 8, which elutes at 14.2 minutes. N-terminal residues are shown for (a) DHFR-Met, (b) the free amino acid 8, and (c) DHFR-8, as determined by Edman degradation.
Figure 3.16  N-terminal sequencing results indicating occupancy of the initiator site in DHFR produced in bacterial cultures supplemented with 11, which elutes at 11.7 minutes. N-terminal residues are shown for (a) DHFR-Met, (b) the free amino acid 11, and (c) DHFR-11, as determined by Edman degradation.
Figure 3.17  N-terminal sequencing results indicating occupancy of the initiator site in DHFR produced in bacterial cultures supplemented with 12, which elutes at 11.6 minutes. N-terminal residues are shown for (a) DHFR-Met, (b) the free amino acid 12, and (c) DHFR-12, as determined by Edman degradation.
$^1$H NMR spectra (599.69 MHz) of DHFR (a), 5 (b), and DHFR-5 (c + d). Samples were dissolved at concentrations of approximately 10 mg/mL in D$_2$O containing 2% [D$_2$]formic acid and spectra were recorded at 25°C overnight.
Figure 3.19  Activation of methionine analogues by MetRS. Assay reactions were supplemented with the methionine analogues indicated, at a concentration of 15 mM. The reaction was quenched at 3.5 hours. The background is given for a reaction mixture lacking both analogue and enzyme.
Figure 3.20  Fluorescence spectra for GFP produced from CAG18491/pQE31-GFP/pREP4 cultures supplemented with methionine, 2, 3, or 9. The spectra were collected at room temperature with an excitation maximum of 396 nm and an emission maximum of 506 nm.
3.5 References


(8) Ibba, M., Kast, P. and Hennecke, H. Biochemistry 1994, 33, 7107-7112.


4.1 Introduction and Objectives

In the previous chapters it has been demonstrated that the translational apparatus of E. coli is remarkably permissive toward methionine analogues under appropriate experimental conditions. In our efforts to further expand the repertoire of selective and mild chemistries that can be used to engineer proteins in vivo, we turned our attention to recently published work indicating that azidosugars can be utilized by the polysaccharide synthesis machinery of the cell (1). Such permissiveness has allowed the engineering of cell surfaces with sialic acid residues bearing azide and other functionality (1-4); the azide can be subsequently derivatized under very mild aqueous conditions at high yields (1). The incorporation of azidoamino acids into proteins in vivo may therefore be a feasible and desirable approach for engineering proteins that can be selectively modified under native conditions.

The chemical modification strategy for proteins decorated with azide functionality is shown in Figure 4.1. The triarylphosphine reagent can selectively deliver a variety of functional moieties (although the FLAG peptide is shown in the figure) via formation of an aza ylide intermediate that rearranges to form an amide bond (1). Based on the
potential of this chemistry for selective chemical modification of proteins, two azido amino acids, 14 and 15 (Appendix B and Figure 4.2), were investigated as potential methionine surrogates in vivo and in activation assays in vitro. Electron density maps and isopotential surfaces for these analogues were generated. Combining the results from in vitro and in vivo studies with models of the electronic and steric features of these side chains may provide additional insight into the primary factors controlling the activation of methionine analogues by MetRS. Chemical modification of the azido residues in the target protein DHFR was also investigated.

4.2 Experimental Section

4.2.1 Computational Methods

The modeling of the electron density maps and isopotential surfaces of analogues 14 and 15 was conducted as described in Section 2.2.5.

4.2.2 Analogue Synthesis

Analogues 14 and 15 were kindly provided by Eliana Saxon, University of California Berkeley, Berkeley, CA, and were synthesized by modification of serine and
homoserine, respectively (5-7). They were characterized by NMR spectroscopy and mass spectrometry and used as provided in assays measuring their translational activity in vivo and their activation by MetRS in vitro.

4.2.3 Activation of Azidoamino Acids by MetRS in Vitro

The activation of analogues 14 and 15 was measured by monitoring the ATP-PP_i exchange supported by the analogues, as described in Section 2.2.4. Assays designed to measure the amount of PP_i exchanged in a 20-minute period were conducted using 75 nM MetRS and an analogue concentration of 5 mM. An additional assay to measure the amount of PP_i exchange at 3.5 hours in a reaction mixture that was 75 nM in enzyme and 15 mM in analogue was also conducted for analogue 14. Quantitative determination of the kinetic parameters for the activation of the analogue 15 by MetRS was also conducted as previously described. Enzyme concentrations of 50 nM and analogue concentrations of 200 µM to 10 mM were used in these assays.

4.2.4 Translational Activity of Azidoamino Acids in Vivo

Experiments to measure the ability of the analogues to support protein synthesis in vivo were conducted as previously described (Section 2.2.2). Small-scale (5 mL)
expressions were conducted for both analogues at analogue concentrations of 20 mg/L and 500 mg/L. Large-scale (1 L) expressions were conducted for analogue 15 with an analogue concentration of 400 mg/L. Proteins were purified via metal chelate affinity chromatography and analyzed by amino acid analysis and N-terminal sequencing as described in Section 3.2.4. Fragments of DHFR-15 were generated by trypsin digestion and were analyzed by matrix-assisted laser desorption ionization mass spectrometry (MALDI-MS) employing α-cyano-4-hydroxycinnamic acid as the matrix. The digestion and MALDI-MS analysis were conducted by Gary Hathaway at the Protein/Peptide Microanalytical Laboratory at the California Institute of Technology, Pasadena, CA.

4.2.5 GFP Fluorescence

A medium-scale (50 mL) expression of GFP (which contains 6 methionine residues) from bacterial cultures CAG18491/pQE31-GFP/pREP 4 was conducted as a preliminary assessment of the impact of the incorporation of 15 on the activity of proteins. The expression protocols, purification procedures, and fluorescence measurements were identical to the procedures described in Chapter 3. The fluorescence spectra of the proteins were measured directly as described in Section 3.2.6 and normalized for protein concentration, which in this case was determined by the Lowry method.
4.2.6 Modification of Azide Residues via the Staudinger Ligation

Protein (DHFR) containing \textbf{15} in place of methionine (DHFR-\textbf{15}) was purified via metal chelate affinity chromatography under denaturing conditions, dialyzed, and lyophilized, as described in Chapter 3 (8). This protein was also analyzed by amino acid analysis and N-terminal sequencing to confirm the incorporation of \textbf{15}. The procedures described below were conducted by Eliana Saxon at the University of California Berkeley, Berkeley, CA. DHFR-\textbf{15} and DHFR-Met (10 \textmu M) were incubated (separately) with the triarylphosphine-FLAG reagent, shown in Figure 4.1, (250 \textmu M) in PBS buffer (137 mM NaCl, 2.7 mM KCl, 4.3 mM Na$_2$HPO$_4$·7H$_2$O, 1.4 mM KH$_2$PO$_4$, pH 7.4) at room temperature for up to 6 hours. The selectivity of the Staudinger ligation was assessed by Western blot analysis with either a primary RGS-His antibody or a primary FLAG antibody. Western blots were developed as previously described. After incubation with the appropriate primary antibody, the blots were washed and treated with a secondary anti-mouse IgG conjugated to HRP and were developed with appropriate chemiluminescence reagents (7).
4.3 Results and Discussion

4.3.1 Computational Modeling of Analogues

The computational studies on azido amino acids are aimed not only at the design of an azido methionine analogue that can replace methionine in vivo, but also at establishing additional evidence for the important structural and electronic features that make an analogue a good substrate for MetRS. The electron density maps and isopotential surfaces for the azido methionine analogues 14 and 15 were calculated and compared with those of methionine, as shown in Figure 4.2.

Our previous studies suggested that the presence of electron density at or near the δ-position of the analogue is one important feature for activation of a methionine analogue by MetRS, as no single side-chain geometry appeared to be required for analogue activation by the enzyme (9-12). These observations are consistent with crystal structure studies demonstrating the importance of hydrogen bonding interactions between the sulfur at the δ-position of the side chain of methionine and the hydroxyl group of Tyr260 of MetRS (13). It is apparent from the representations in Figure 4.2 that 15 not only is a good structural analogue for methionine, but also has electron density at the δ-position of the side chain with an isopotential surface very similar to that of methionine. Analogue 14, on the other hand, is less similar structurally to methionine and has electron
density located at the γ-position and on the opposite side of the extended side chain than that of methionine. Based on these results and by considering the orientation of methionine in the binding pocket of MetRS (Figure 1.4), we hypothesized that 15, and not 14, should serve well as a methionine surrogate in vitro and in vivo.

4.3.2 Activation of Azidoamino Acids by MetRS in Vitro

The results of ATP-PP₃ exchange assays for 14 and 15 support the above hypothesis and indicate that only 15, and not 14, is able to support PP₃ exchange in vitro, as illustrated in Figure 4.3. The in vitro results for all of the methionine analogues are given to facilitate comparison between analogues. As shown in the figure, 15 is indicated to support exchange at a rate similar to that observed for 3, one of the most efficiently activated methionine analogues investigated to date. (Analogue 14 did not support PP₃ exchange at levels above background even in assays that were 15 mM in 14 and were quenched after 3.5 hours.) The quantitative analysis of the kinetic parameters for 15 is consistent with the data in Figure 4.3 and is summarized in Table 4.1. A $k_{cat}/K_m$ value of $1.42 \times 10^{-3} \text{ s}^{-1} \mu\text{M}^{-1}$ was measured for activation of 15 by MetRS in vitro, indicating that it may be a slightly better analogue for activation by MetRS than 3, which has a $k_{cat}/K_m$ value for activation of $1.16 \times 10^{-3} \text{ s}^{-1} \mu\text{M}^{-1}$. The side chains of both 3 and 15 have
electron density near the δ-position of the amino acid, although 15 shares a greater conformational similarity with methionine that may allow it to be activated slightly more efficiently by MetRS than 3.

4.3.3 Translational Activity of Azidoamino Acids in Vivo

Whether or not these in vitro observations correlate with the translational activity of 14 and 15 in vivo was tested in a standard small-scale (5-mL) protein expression assay. The ability of the analogues to support protein synthesis in bacterial cultures supplemented with the analogue and depleted of methionine was monitored by SDS-PAGE analysis. Both conventional (CAG18491/pQE15/pREP4) and modified (CAG18491/pQE15-MRS/pREP4) bacterial expression hosts were investigated. The results in Figure 4.4, shown for the conventional expression host CAG18491/pQE15/pREP4, are representative of both hosts and demonstrate that only analogue 15 supports protein synthesis in vivo. Analogue 14 did not support protein biosynthesis even in a bacterial host equipped with extra copies of the wild-type or mutant MetRS.

Consistent with the relatively high catalytic efficiency with which MetRS activates this analogue (Table 4.1), the synthesis of target protein is observed in cultures of a conventional bacterial host supplemented with 15; overexpression of MetRS is not
required. Also consistent with the relatively high $k_{cat}/K_m$ for this analogue (similar to the $k_{cat}/K_m$ for 3, which supported protein yields identical to those for methionine), the yields of protein obtained from 5-mL cultures supplemented with 20 mg/L 15 are similar to those for cultures supplemented with 20 mg/L methionine, as shown in Figure 4.5. Increasing the concentration of amino acid in the medium to 500 mg/L does not increase the protein yield for cultures supplemented either with 15 or methionine, indicating that protein biosynthesis is not limited by the rate of activation of either of these amino acids under the in vivo assay conditions. Yields from large-scale expressions (1 L) corroborate these observations, with the yield of DHFR obtained from 1-L cultures of CAG18491/pQE15/pREP4 supplemented with 15 matching those obtained from cultures supplemented with methionine or 3 (35 mg/L).

4.3.4 Protein Characterization

Protein obtained from these expressions was purified and analyzed for amino acid content and for the identity of the N-terminal amino acid. The results of the amino acid analysis are listed in Table 4.2, and show that other than the decrement in methionine content, the profile of amino acids is consistent with that expected for DHFR. As 15 is unstable under the analysis conditions, the decrease in methionine content is taken as indirect evidence of replacement of methionine by 15. Based on the decrease of
methionine content from an expected value of 3.8% to an observed value of 0.2%, the amino acid analysis indicates 95 ± 2% replacement of methionine.

As DHFR retains the N-terminal methionine (14), direct measure of the incorporation of the analogue could be obtained from N-terminal sequencing, shown in Figure 4.6. The signal corresponding to methionine elutes at approximately 13.0 minutes while that corresponding to 15 elutes at 11.5 minutes. The large peaks that elute at approximately 14.7 minutes correspond to piperidylphenylthiourea (pptu), a product of the analysis resulting from the buffer, and the small peak at approximately 18.5 minutes corresponds to diethylphthalate (diet), an internal standard. As demonstrated in the figure, the N-terminal methionine of DHFR (Figure 4.6a) is nearly completely replaced by 15 (Figure 4.6b) in DHFR produced from cultures supplemented with 15 (Figure 4.6c). Integration of the peak areas indicates 97 ± 2% replacement of methionine.

Peptide fragments from DHFR-15 were generated by digestion with trypsin and analyzed by MALDI-MS. Results for all observed fragments in which methionine is encoded are delineated in Table 4.3, and a representative mass spectrum for the fragment with the encoded sequence IMQEFESDTFFPEIDLGK is shown in Figure 4.7. The methionine-containing fragment should exhibit a mass of 2146.4, although no such peak is apparent in this data. Instead, a fragment with a mass of 2141.1 is observed, which is consistent with the replacement of the single methionine in the fragment by 15. The complete absence of a peak at the mass expected for the methionine-containing fragment
suggests essentially quantitative replacement of methionine by 15. This is further corroborated by tryptic digest /MALDI-MS analysis of DHFR-Met, in which the fragment at 2146.4 is reliably detected at high intensity. Similar mass spectral data were obtained for three additional tryptic fragments of DHFR-15 (Table 4.3). Again, fragments are observed only at masses consistent with complete replacement of methionine by 15, and not at masses indicating retention of methionine. These results suggest quantitative replacement of methionine by 15 for 6 different positions encoding methionine, consistent with both amino acid analyses and N-terminal sequencing results.

4.3.5 GFP Fluorescence

The impact of this level of azidoamino acid incorporation on protein function was preliminarily investigated by monitoring the fluorescence observed in samples of GFP produced from bacterial cultures supplemented with 15. The GFP contains 6 methionines and has an absorbance maximum at 396 nm and an emission maximum at 506 nm (15). GFP produced from a culture supplemented with 500 mg/L of 15 exhibits an emission spectrum similar to that observed for GFP produced from cultures supplemented with the same level of methionine (Figure 4.7). While the emission maximum appears at 506 nm in both spectra, it is obvious from the normalized spectra shown in Figure 4.8 that the intensity of the emission peak is significantly reduced for the GFP containing the azide
analogue. The fact that fluorescence is observed demonstrates that the quantitative, multisite replacement of methionine by 15 does not result in a complete loss of protein structure or function. The selective modification or covalent capture of biologically active macromolecules therefore seems a feasible application of the incorporation of azidoamino acids into proteins in vivo. As discussed in Chapter 3, it is difficult to make generalizations regarding the impact of analogue incorporation on protein structure and function, as such an impact must be assessed on an individual basis.

4.3.6 Modification of DHFR-15 by the Staudinger Ligation

The ability of DHFR-15 to be selectively modified by the Staudinger ligation (shown in Figure 4.1) has also been preliminarily investigated. DHFR produced from bacterial cultures of CAG18491/pQE15/pREP4 supplemented with 500 mg/L of 15 was purified under denaturing conditions, dialyzed, and lyophilized. Either DHFR-15 or DHFR-Met was incubated with a triarylphosphine-FLAG reagent in PBS buffer for up to 6 hours in order to selectively react the modified DHFR with the triarylphosphine-FLAG (7). Figure 4.9 gives Western blot results for some of these preliminary experiments. The Western blot in Figure 4.9a was developed with an RGS-His antibody, and bands with a molecular weight consistent with that of DHFR are observed in lanes 2-5, demonstrating that both DHFR-15 and DHFR-Met are recognized by the RGS-His
antibody, as expected. The incorporation of 15 at the N-terminal position (near the hexahistidine sequence) does not appear to compromise the binding of the antibody.

Only the DHFR-15 that is incubated with the triarylphosphine-FLAG shows additional bands at higher molecular weights, however, consistent with the modification of multiple sites of DHFR-15. The Western blot in Figure 4.9b, developed with a FLAG antibody, exhibits bands only in lane 2. Coupled with the results in Figure 4.9(a), these data demonstrate that only DHFR-15 is labeled with the FLAG peptide; neither DHFR-15 or DHFR-Met incubated in buffer, nor DHFR-Met incubated with the triarylphosphine-FLAG are detected in the Western analysis employing the FLAG antibody.

These results unequivocally demonstrate the selectivity of the triarylphosphine reagent for the azidohomoalanine in DHFR-15. The presence of multiple bands in lane 2 of Figures 4.9(a) and 4.9(b), however, indicates that the 8 positions containing 15 are not quantitatively modified. Since our characterization results indicate essentially quantitative replacement of methionine by 15, the lack of quantitative modification does not arise from the presence of methionine in the protein. Incomplete reaction of the azide groups with the triarylphosphine reagent may arise from insufficient length of reaction, a lack of accessibility of the azide groups to the triarylphosphine-FLAG, partial hydrolysis of the aza ylide intermediate in the context of the protein, or partial reduction of the azide groups to amine groups during protein biosynthesis and purification.
Reduction during protein biosynthesis and purification is not indicated by N-terminal sequencing, as a single peak with an area corresponding to 97% replacement is observed in the chromatogram of the N-terminal residue of DHFR-15 (Figure 4.6). There is, however, indication of the reduction in the MALDI-MS analysis of tryptic fragments of DHFR-15. Reduction of the azide to the amine results in a loss of 26 mass units, and in the mass spectrum, a small peak corresponding to such a mass loss is observed for each fragment containing 15. A representative mass spectrum is given in Figure 4.10. Based on the N-terminal sequencing and mass spectral data, the degree of conversion to amine prior to chemical modification is not sufficient to account for the incomplete reaction of the azide functionality with the triarylphosphine-FLAG. Methods to quantify the degree of reduction are currently being investigated, along with methods to improve the extent of chemical modification.

4.4 Summary and Conclusions

These studies constitute the first example of the replacement of methionine residues by azidoamino acids in vivo. The capacity of azidoamino acids 14 and 15 to support protein synthesis in vivo correlates well with their activation by MetRS in vitro, as 14, which does not support measurable levels of PP_i exchange, also does not support protein biosynthesis. In contrast, 15, which is the most efficiently activated analogue
investigated to date, supports protein synthesis at yields similar to those supported by methionine, with essentially quantitative levels of replacement. These results provide additional evidence indicating the critical role of MetRS in the incorporation of methionine analogues into proteins \textit{in vivo}, and yield additional information confirming the important side-chain features controlling activation of methionine analogues by MetRS. That 15, and not 14, is activated efficiently by MetRS indicates that the presence of electron density near the \( \delta \)-position of a methionine analogue is an important feature for efficient analogue activation.

Furthermore, the incorporation of azidoamino acids into proteins provides an additional and novel method for the selective modification of proteins. The selectivity of the reaction between triarylphosphine reagents and the azidoamino acid residues in a protein has been demonstrated for the first time. That a protein containing an azidoamino acid might maintain its function has been indicated by experiments demonstrating that GFP produced from bacterial cultures supplemented with 15 retains a significant level of fluorescence. The ligation of phosphines and azides may therefore permit the modification of biologically active macromolecules or the trapping of transiently associate intermediates within the cellular environment. Investigations are underway to demonstrate quantitative modification of azidoamino acids incorporated into proteins \textit{in vivo}. 

146
Table 4.1  Comparison of kinetic parameters for azidoamino acids 14 and 15 with those for methionine analogues 3 and 9.

<table>
<thead>
<tr>
<th>Analogue</th>
<th>$k_{cat}/K_m$, s$^{-1}$ μM$^{-1}$</th>
<th>Relative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Met</td>
<td>$5.47 \times 10^1$</td>
<td>1</td>
</tr>
<tr>
<td>14</td>
<td>no activity</td>
<td>n/a</td>
</tr>
<tr>
<td>15</td>
<td>$1.42 \times 10^3$</td>
<td>1/390</td>
</tr>
<tr>
<td>3</td>
<td>$1.16 \times 10^3$</td>
<td>1/500</td>
</tr>
<tr>
<td>9</td>
<td>$5.22 \times 10^4$</td>
<td>1/1050</td>
</tr>
</tbody>
</table>
Table 4.2 Amino acid analysis results for DHFR-Met and DHFR-15.

<table>
<thead>
<tr>
<th>Amino acid</th>
<th>Theoretical mole %</th>
<th>DHFR-Met</th>
<th>DHFR-15</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ala</td>
<td>2.9</td>
<td>3.6</td>
<td>3.8</td>
</tr>
<tr>
<td>Cys</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Asx</td>
<td>7.1</td>
<td>7.6</td>
<td>8.1</td>
</tr>
<tr>
<td>Glx</td>
<td>11.9</td>
<td>12.3</td>
<td>12.6</td>
</tr>
<tr>
<td>Phe</td>
<td>4.3</td>
<td>4.4</td>
<td>4.3</td>
</tr>
<tr>
<td>Gly</td>
<td>7.6</td>
<td>8.5</td>
<td>8.4</td>
</tr>
<tr>
<td>His</td>
<td>3.8</td>
<td>3.9</td>
<td>3.9</td>
</tr>
<tr>
<td>Ile</td>
<td>6.7</td>
<td>5.6</td>
<td>6.2</td>
</tr>
<tr>
<td>Lys</td>
<td>8.0</td>
<td>8.1</td>
<td>8.5</td>
</tr>
<tr>
<td>Leu</td>
<td>9.5</td>
<td>9.6</td>
<td>9.4</td>
</tr>
<tr>
<td>Met</td>
<td>3.8</td>
<td>3.7</td>
<td>0.2</td>
</tr>
<tr>
<td>Pro</td>
<td>6.2</td>
<td>6.4</td>
<td>6.7</td>
</tr>
<tr>
<td>Arg</td>
<td>5.7</td>
<td>6.0</td>
<td>6.5</td>
</tr>
<tr>
<td>Ser</td>
<td>7.6</td>
<td>7.8</td>
<td>8.6</td>
</tr>
<tr>
<td>Thr</td>
<td>2.9</td>
<td>3.3</td>
<td>3.2</td>
</tr>
<tr>
<td>Val</td>
<td>6.7</td>
<td>6.3</td>
<td>6.7</td>
</tr>
<tr>
<td>Trp</td>
<td>1.9</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Tyr</td>
<td>2.9</td>
<td>2.9</td>
<td>2.7</td>
</tr>
</tbody>
</table>
Table 4.3  Mass spectrometric analysis of tryptic fragments of DHFR-15.

<table>
<thead>
<tr>
<th>Fragment sequence</th>
<th># Methionines</th>
<th>Methionine</th>
<th>Mass 15</th>
<th>Observed</th>
</tr>
</thead>
<tbody>
<tr>
<td>QNLVIMGR</td>
<td>1</td>
<td>930.1</td>
<td>925.0</td>
<td>925.6</td>
</tr>
<tr>
<td>IMQEFESDTFFPEIDLGK</td>
<td>1</td>
<td>2146.4</td>
<td>2141.3</td>
<td>2141.1</td>
</tr>
<tr>
<td>VDMVWIVGGSSVYQEA MNQPGHLR</td>
<td>2</td>
<td>2674.0</td>
<td>2663.8</td>
<td>2663.4</td>
</tr>
<tr>
<td>GSHHHHHHGGSGIMVRP LNSIVAVSQNMGI GK</td>
<td>2</td>
<td>3295.8</td>
<td>3285.6</td>
<td>3284.8</td>
</tr>
</tbody>
</table>
Figure 4.1  The Staudinger ligation for the chemical modification of azido amino acids with a triarylphosphine-FLAG reagent.
Electron density maps (colored surfaces) and negative isopotential surfaces (meshes) for (a) methionine, (b) 14, and (c) 15. The electron density maps indicate electron-rich (red) and electron-poor (blue) regions of each molecule.
Figure 4.3 Activation of methionine and methionine analogues by MetRS. The amount of PP\textsubscript{i} exchanged in 20 minutes is shown for methionine (1) and methionine analogues 2 - 15. The full set of analogues is given to facilitate comparison between analogues 2 - 13 (Chapter 2) and 14 and 15. The background (16) is given for a reaction mixture lacking both enzyme and amino acid.
Figure 4.4 SDS-PAGE analysis of the translational activity for methionine analogues 14 and 15. The SDS-PAGE analysis was conducted on whole cell lysates of CAG18491/pQE15/pREP4 cultures.
Figure 4.5  Western blot analysis of the expression of DHFR in bacterial cultures supplemented with the indicated concentration of (a) methionine or (b) 15.
Figure 4.6  N-terminal sequencing results indicating occupancy of the initiator site in DHFR produced in bacterial cultures supplemented with 15, which elutes at 11.5 minutes. N-terminal residues are shown for (a) DHFR-Met, (b) the free amino acid 15, and (c) DHFR-15, as determined by Edman degradation.
Figure 4.7 Mass spectrometric analysis of a tryptic fragment of DHFR-15. A fragment containing methionine (M), with the sequence shown, would exhibit a mass of 2146.4. Substitution of methionine with 15 (Z) yields a fragment with an expected mass of 2141.3. The relative ratio of methionine-containing fragment vs. 15-containing fragment shown here is observed for three additional fragments.
Figure 4.8  Fluorescence spectra for GFP produced from cultures of bacterial host CAG18491/pQE31-GFP/pREP4 supplemented with methionine or 15. Spectra were collected at room temperature in 50 mM phosphate, 250 mM imidazole buffer at pH 8.0 and were normalized for protein concentration. The spectra were collected at room temperature with an excitation maximum of 396 nm and an emission maximum of 506 nm.
Figure 4.9  Western blot analysis of DHFR produced from cultures of bacterial host CAG18491/pQE15/pREP4 supplemented with methionine or 15 (7).  (a) Western blot analysis with an RGS-His antibody. 1) Molecular weight markers (19.2, 24.7, 36.4 kDa), 2) DHFR-15 incubated with triarylphosphine-FLAG at room temperature in PBS buffer, pH 7.4, 3) DHFR-15 incubated in PBS buffer, 4) DHFR-Met incubated with the triarylphosphine-FLAG, 5) DHFR-Met incubated in PBS buffer. (b) Western blot analysis with a FLAG-antibody. (1 - 5) are as designated for (a).
Figure 4.10  Mass spectrometric analysis of a tryptic fragment of DHFR-15. The peak with a mass of 2141.1 corresponds to a fragment containing 15 (as shown in Figure 4.7). A loss of 26 mass units to yield the peak with a mass of 2115.1 is consistent with the reduction of the azide (15) to an amine.
4.5 References


CHAPTER 5
FUTURE RESEARCH DIRECTIONS

Our studies have demonstrated that overexpression of MetRS in a bacterial host not only increases the number of methionine analogues that can be utilized during protein biosynthesis, but also increases the yields of proteins containing methionine analogues. Efficient incorporation of a variety of unsaturated analogues in place of methionine could be exploited for selective chemical modification via ruthenium-catalyzed (1,2), palladium-catalyzed (3-5), tungsten-catalyzed (6), and other chemistries (7,8). The incorporation of azide functionality permits selective modification with triarylphosphine reagents and may present opportunities for other selective chemical transformations. Furthermore, proteins with high extents of replacement of methionine by unsaturated and azido amino acids may also retain their function, as indicated by preliminary experiments in which GFP containing methionine analogues exhibits significant, albeit reduced, fluorescence. Additional characterization of the modified GFPs (and other proteins, such as DHFR) by circular dichroism and NMR spectroscopies may provide evidence indicating the degree of perturbation of protein folding resulting from the incorporation of methionine analogues. Although these initial studies of modified GFPs indicate that protein fluorescence is reduced, they suggest that protein structure and function might be retained upon analogue incorporation. Incorporation of novel non-natural amino acids
via manipulation of the aaRS activities of a bacterial host therefore has enormous potential for permitting production of natural and artificial proteins with novel chemical and physical properties.

Strategies for manipulating the aaRS activities of a bacterial host include overexpression of wild-type aaRS, site-directed mutagenesis of aaRS, or introduction of heterologous synthetases in the cell (9-11). Overexpression of wild-type MetRS has been demonstrated in our laboratories to permit the incorporation of a variety of methionine analogues, and it is likely that overexpression of other wild-type aaRS will provide similar results. Indeed, overexpression of wild-type LeuRS in a cellular host has rescued the translational activity of hexafluoroleucine, an analogue not utilized by the conventional translational machinery (12). Therefore it seems likely that expanding this simple strategy - overexpression of wild-type aaRS - to include other aaRS should have similar positive results.

Overexpression of a mutant PheRS prepared via site-directed mutagenesis (13) has permitted the in vivo incorporation of 17-Br-Phe, an analogue not normally utilized during protein biosynthesis (14). While overexpression of a mutant MetRS W305F has not afforded advantages over overexpression of wild-type MetRS, the mutant enzyme exhibits slightly higher rates for activation of methionine analogues than the wild-type MetRS. Therefore, mutation of aaRS activities of the cell via directed evolution methods (15,16) may be useful for producing aaRS capable of activating non-natural amino acids.
not activated by wild-type aaRS. The demonstrated promiscuity of the wild-type MetRS, coupled with the increased activity of the mutant MetRS, suggests the potential of evolving MetRS activity toward methionine analogues such as those bearing fluorinated, conjugated, or surfactant-like groups.

In addition to this evolutionary-based strategy for developing new aaRS activity, a rational approach to aaRS design may also be possible through the use of computational methods. Computational methods developed by Goddard and coworkers (17-19) have yielded free energies of binding for phenylalanine analogue/PheRS pairs that correlate well with \textit{in vitro} activation and \textit{in vivo} incorporation. The rational design of methionine analogue/MetRS pairs is a logical extension of this computational work, given the ample kinetic data now available for the activation of methionine analogues. In the long term, these studies would be useful for understanding the basic interactions that govern binding of an amino acid by an aaRS, without the need for synthesis of new analogues or crystallization of an aa/aaRS pair. The prospects are therefore promising for utilizing these approaches to design new aa/aaRS pairs for the purposeful production of chemically and physically novel biopolymers.
5.2 References


APPENDIX A

LIST OF AMINO ACIDS

**ALIPHATIC**
- Glycine (Gly G)
- Alanine (Ala A)
- Valine (Val V)
- Leucine (Leu L)
- Isoleucine (Ile I)

**AROMATIC**
- Phenylalanine (Phe F)
- Tyrosine (Tyr Y)
- Tryptophan (Trp W)

**BASIC**
- Lysine (Lys K)
- Arginine (Arg R)
- Histidine (His H)

**ACIDIC**
- Aspartic Acid (Asp D)
- Glutamic Acid (Glu E)
- Asparagine (Asn N)
- Glutamine (Gin Q)

**HYDROXYL OR SULFUR CONTAINING**
- Serine (Ser S)
- Cysteine (Cys C)
- Threonine (Thr T)
- Methionine (Met M)
- Homocysteine (Hcy)
- Precursor to methionine
APPENDIX B

LIST OF METHIONINE ANALOGUES

Methionine (1)

H₂N\text{S}OH

Homoallylglycine (2)

\text{H₂N-}\text{OH}

Homopropargylglycine (3)

\text{H₂N-}\text{OH}

cis-Crotylglycine (4)

\text{H₂N-}\text{CH=CH-}\text{OH}

trans-Crotylglycine (5)

\text{H₂N-}\text{CH-CH=CH-}\text{OH}

6,6,6 - Trifluoronorleucine (6)

\text{H₂N-}\text{CF}_3

2-Aminoheptanoic acid (7)

\text{H₂N-}\text{CH(CH₃)}-\text{OH}

Norvaline (8)

\text{H₂N-}\text{CH(CH₃)}-\text{OH}

Norleucine (9)

\text{H₂N-}\text{CH(CH₃)}-\text{OH}

o-Allylserine (10)

\text{H₂N-}\text{CH(CH₃)}-\text{OH}

2-Butynylglycine (11)

\text{H₂N-}\text{CH(CH₃)}=\text{CH}-\text{OH}

Allylglycine (12)

\text{H₂N-}\text{CH(CH₃)}-\text{OH}

Propargylglycine (13)

\text{H₂N-C≡CH-}\text{OH}

Azidoalanine (14)

\text{H₂N}-\text{N=N-N}

Azidohomoalanine (15)

\text{H₂N}-\text{N=N-N}
APPENDIX C

TOCSY SPECTRA FOR DHFR-5

(a) Selective irradiation at 5.35 ppm, no mixing time

(b) Selective irradiation at 5.35 ppm, 20 ms mixing time
APPENDIX E

DNA SEQUENCE OF METRS NHE I CASSETTE

Continued next page
APPENDIX E Continued
Continued next page
Appendix E

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>HinP I</td>
<td>Bae I</td>
</tr>
<tr>
<td>Mbo I</td>
<td>Hha I</td>
</tr>
<tr>
<td>Sau3AI</td>
<td>Hpa I</td>
</tr>
<tr>
<td>Msp I</td>
<td>Hha I</td>
</tr>
<tr>
<td>Nsi I</td>
<td>Bst II</td>
</tr>
<tr>
<td>Mbo I</td>
<td>Hpa I</td>
</tr>
<tr>
<td>Sau3AI</td>
<td>Hpa I</td>
</tr>
<tr>
<td>Msp I</td>
<td>Hpa II</td>
</tr>
<tr>
<td>Nsi I</td>
<td>Bst II</td>
</tr>
<tr>
<td>Mbo I</td>
<td>Hpa I</td>
</tr>
<tr>
<td>Sau3AI</td>
<td>Hpa I</td>
</tr>
<tr>
<td>Msp I</td>
<td>Hpa I</td>
</tr>
<tr>
<td>Nsi I</td>
<td>Bst II</td>
</tr>
<tr>
<td>Mbo I</td>
<td>Hpa I</td>
</tr>
<tr>
<td>Sau3AI</td>
<td>Hpa I</td>
</tr>
<tr>
<td>Msp I</td>
<td>Hpa II</td>
</tr>
<tr>
<td>Nsi I</td>
<td>Bst II</td>
</tr>
<tr>
<td>Mbo I</td>
<td>Hpa I</td>
</tr>
<tr>
<td>Sau3AI</td>
<td>Hpa I</td>
</tr>
<tr>
<td>Msp I</td>
<td>Hpa II</td>
</tr>
<tr>
<td>Nsi I</td>
<td>Bst II</td>
</tr>
<tr>
<td>Mbo I</td>
<td>Hpa I</td>
</tr>
<tr>
<td>Sau3AI</td>
<td>Hpa I</td>
</tr>
<tr>
<td>Msp I</td>
<td>Hpa II</td>
</tr>
<tr>
<td>Nsi I</td>
<td>Bst II</td>
</tr>
<tr>
<td>Mbo I</td>
<td>Hpa I</td>
</tr>
<tr>
<td>Sau3AI</td>
<td>Hpa I</td>
</tr>
<tr>
<td>Msp I</td>
<td>Hpa II</td>
</tr>
<tr>
<td>Nsi I</td>
<td>Bst II</td>
</tr>
<tr>
<td>Mbo I</td>
<td>Hpa I</td>
</tr>
<tr>
<td>Sau3AI</td>
<td>Hpa I</td>
</tr>
<tr>
<td>Msp I</td>
<td>Hpa II</td>
</tr>
<tr>
<td>Nsi I</td>
<td>Bst II</td>
</tr>
<tr>
<td>Mbo I</td>
<td>Hpa I</td>
</tr>
<tr>
<td>Sau3AI</td>
<td>Hpa I</td>
</tr>
<tr>
<td>Msp I</td>
<td>Hpa II</td>
</tr>
<tr>
<td>Nsi I</td>
<td>Bst II</td>
</tr>
<tr>
<td>Mbo I</td>
<td>Hpa I</td>
</tr>
<tr>
<td>Sau3AI</td>
<td>Hpa I</td>
</tr>
<tr>
<td>Msp I</td>
<td>Hpa II</td>
</tr>
<tr>
<td>Nsi I</td>
<td>Bst II</td>
</tr>
<tr>
<td>Mbo I</td>
<td>Hpa I</td>
</tr>
<tr>
<td>Sau3AI</td>
<td>Hpa I</td>
</tr>
<tr>
<td>Msp I</td>
<td>Hpa II</td>
</tr>
<tr>
<td>Nsi I</td>
<td>Bst II</td>
</tr>
<tr>
<td>Mbo I</td>
<td>Hpa I</td>
</tr>
<tr>
<td>Sau3AI</td>
<td>Hpa I</td>
</tr>
<tr>
<td>Msp I</td>
<td>Hpa II</td>
</tr>
<tr>
<td>Nsi I</td>
<td>Bst II</td>
</tr>
<tr>
<td>Mbo I</td>
<td>Hpa I</td>
</tr>
<tr>
<td>Sau3AI</td>
<td>Hpa I</td>
</tr>
<tr>
<td>Msp I</td>
<td>Hpa II</td>
</tr>
<tr>
<td>Nsi I</td>
<td>Bst II</td>
</tr>
<tr>
<td>Mbo I</td>
<td>Hpa I</td>
</tr>
<tr>
<td>Sau3AI</td>
<td>Hpa I</td>
</tr>
<tr>
<td>Msp I</td>
<td>Hpa II</td>
</tr>
<tr>
<td>Nsi I</td>
<td>Bst II</td>
</tr>
<tr>
<td>Mbo I</td>
<td>Hpa I</td>
</tr>
<tr>
<td>Sau3AI</td>
<td>Hpa I</td>
</tr>
<tr>
<td>Msp I</td>
<td>Hpa II</td>
</tr>
<tr>
<td>Nsi I</td>
<td>Bst II</td>
</tr>
<tr>
<td>Mbo I</td>
<td>Hpa I</td>
</tr>
<tr>
<td>Sau3AI</td>
<td>Hpa I</td>
</tr>
<tr>
<td>Msp I</td>
<td>Hpa II</td>
</tr>
<tr>
<td>Nsi I</td>
<td>Bst II</td>
</tr>
<tr>
<td>Mbo I</td>
<td>Hpa I</td>
</tr>
<tr>
<td>Sau3AI</td>
<td>Hpa I</td>
</tr>
<tr>
<td>Msp I</td>
<td>Hpa II</td>
</tr>
<tr>
<td>Nsi I</td>
<td>Bst II</td>
</tr>
<tr>
<td>Mbo I</td>
<td>Hpa I</td>
</tr>
<tr>
<td>Sau3AI</td>
<td>Hpa I</td>
</tr>
<tr>
<td>Msp I</td>
<td>Hpa II</td>
</tr>
<tr>
<td>Nsi I</td>
<td>Bst II</td>
</tr>
<tr>
<td>Mbo I</td>
<td>Hpa I</td>
</tr>
<tr>
<td>Sau3AI</td>
<td>Hpa I</td>
</tr>
<tr>
<td>Msp I</td>
<td>Hpa II</td>
</tr>
<tr>
<td>Nsi I</td>
<td>Bst II</td>
</tr>
<tr>
<td>Mbo I</td>
<td>Hpa I</td>
</tr>
<tr>
<td>Sau3AI</td>
<td>Hpa I</td>
</tr>
<tr>
<td>Msp I</td>
<td>Hpa II</td>
</tr>
<tr>
<td>Nsi I</td>
<td>Bst II</td>
</tr>
<tr>
<td>Mbo I</td>
<td>Hpa I</td>
</tr>
<tr>
<td>Sau3AI</td>
<td>Hpa I</td>
</tr>
<tr>
<td>Msp I</td>
<td>Hpa II</td>
</tr>
<tr>
<td>Nsi I</td>
<td>Bst II</td>
</tr>
<tr>
<td>Mbo I</td>
<td>Hpa I</td>
</tr>
<tr>
<td>Sau3AI</td>
<td>Hpa I</td>
</tr>
<tr>
<td>Msp I</td>
<td>Hpa II</td>
</tr>
<tr>
<td>Nsi I</td>
<td>Bst II</td>
</tr>
<tr>
<td>Mbo I</td>
<td>Hpa I</td>
</tr>
<tr>
<td>Sau3AI</td>
<td>Hpa I</td>
</tr>
<tr>
<td>Msp I</td>
<td>Hpa II</td>
</tr>
<tr>
<td>Nsi I</td>
<td>Bst II</td>
</tr>
<tr>
<td>Mbo I</td>
<td>Hpa I</td>
</tr>
<tr>
<td>Sau3AI</td>
<td>Hpa I</td>
</tr>
<tr>
<td>Msp I</td>
<td>Hpa II</td>
</tr>
<tr>
<td>Nsi I</td>
<td>Bst II</td>
</tr>
<tr>
<td>Mbo I</td>
<td>Hpa I</td>
</tr>
<tr>
<td>Sau3AI</td>
<td>Hpa I</td>
</tr>
<tr>
<td>Msp I</td>
<td>Hpa II</td>
</tr>
<tr>
<td>Nsi I</td>
<td>Bst II</td>
</tr>
<tr>
<td>Mbo I</td>
<td>Hpa I</td>
</tr>
<tr>
<td>Sau3AI</td>
<td>Hpa I</td>
</tr>
<tr>
<td>Msp I</td>
<td>Hpa II</td>
</tr>
<tr>
<td>Nsi I</td>
<td>Bst II</td>
</tr>
<tr>
<td>Mbo I</td>
<td>Hpa I</td>
</tr>
<tr>
<td>Sau3AI</td>
<td>Hpa I</td>
</tr>
<tr>
<td>Msp I</td>
<td>Hpa II</td>
</tr>
<tr>
<td>Nsi I</td>
<td>Bst II</td>
</tr>
<tr>
<td>Mbo I</td>
<td>Hpa I</td>
</tr>
<tr>
<td>Sau3AI</td>
<td>Hpa I</td>
</tr>
<tr>
<td>Msp I</td>
<td>Hpa II</td>
</tr>
<tr>
<td>Nsi I</td>
<td>Bst II</td>
</tr>
<tr>
<td>Mbo I</td>
<td>Hpa I</td>
</tr>
<tr>
<td>Sau3AI</td>
<td>Hpa I</td>
</tr>
<tr>
<td>Msp I</td>
<td>Hpa II</td>
</tr>
<tr>
<td>Nsi I</td>
<td>Bst II</td>
</tr>
<tr>
<td>Mbo I</td>
<td>Hpa I</td>
</tr>
<tr>
<td>Sau3AI</td>
<td>Hpa I</td>
</tr>
<tr>
<td>Msp I</td>
<td>Hpa II</td>
</tr>
<tr>
<td>Nsi I</td>
<td>Bst II</td>
</tr>
<tr>
<td>Mbo I</td>
<td>Hpa I</td>
</tr>
<tr>
<td>Sau3AI</td>
<td>Hpa I</td>
</tr>
<tr>
<td>Msp I</td>
<td>Hpa II</td>
</tr>
<tr>
<td>Nsi I</td>
<td>Bst II</td>
</tr>
<tr>
<td>Mbo I</td>
<td>Hpa I</td>
</tr>
<tr>
<td>Sau3AI</td>
<td>Hpa I</td>
</tr>
<tr>
<td>Msp I</td>
<td>Hpa II</td>
</tr>
<tr>
<td>Nsi I</td>
<td>Bst II</td>
</tr>
<tr>
<td>Mbo I</td>
<td>Hpa I</td>
</tr>
<tr>
<td>Sau3AI</td>
<td>Hpa I</td>
</tr>
<tr>
<td>Msp I</td>
<td>Hpa II</td>
</tr>
<tr>
<td>Nsi I</td>
<td>Bst II</td>
</tr>
</tbody>
</table>

Continued next page
APPENDIX F

DNA AND AMINO ACID SEQUENCE OF METRS

1/1
61/21
ATG ACT CAA GTC GCG AAG AAA ATT CTG GTG ACG TGC GCA CTG CCG TAC GCT AAC
GTC TCA ATC CAC CTC GGC CAT ATG CTG GAG CAC ATC CAG GCT GAT GTC TGG GTC
CGT TAC CAG CGA

Met thr gln val ala lys lys ile leu val thr cys ala leu pro tyr ala asn
gly ser ile his leu gly his met leu glu his ile gln ala asp val trp val
arg tyr gln arg

121/41
181/61
ATG CGC GCC CAC GAG GTC AAC TTC ATC TGC GCC GAC GAT GCC CAC GGT ACA CCG
ATC ATG CTG AAA GCT CAG CAG CTT GGT ATC ACC CCG GAG CAG ATG ATT GCC GAA
ATG AGT CAG GAG

met arg gly his glu val asp ile cys ala asp asp ala his gly thr pro
ile met leu lys ala gln gln leu gly ile thr pro glu gln met ile gly glu
met ser gln glu

241/81
301/101
CAT CAG ACT GAT TTC GCA GCC TTT AAC ATC AGC TAT GCC AAC TAC TGC TGG ACG
CAC AGC GAA GAG AAC CGC CAG TGT TCA GAA CTT ATC TAC TCT CGC CGT AAA GAA
AAC GGT TTT ATT

his gln thr asp phe ala gly phe asn ile ser tyr asp asn tyr his ser thr
his ser glu glu asn arg gln leu ser glu leu ile tyr ser arg leu lys glu
asn gly phe ile

361/121
391/131
AAA AAC CGC ACC ATC TCT CAG CTG TAC GAT CCG GAA AAA GGC ATG TTC CTG CCG
GAC CGT TTT GTG AAA GGC ACC TGC CCG AAA TGT AAA TCC CCG GAT CAA TAC GGC
GAT AAC TGC GAA

lys asn arg thr ile ser gln leu tyr asp pro glu lys gly met phe leu pro
asp arg phe val lys gly thr cys pro lys cys lys ser pro asp gln tyr gly
asp asn cys glu

Continued next page
APPENDIX F Continued

481/161
541/181
GTC TGG GGC GCG ACC TAC AGC CCG ACT GAA CTG ATC GAG CCG AAA TCG GTG GTT
TCT TGG GCT AGC CCG GTA ATG CTT GAT TCT GAA CAC TTC TTT GAT CTG CCC
val cys gly ala thr tyr ser pro thr glu leu ile glu pro lys ser val val
ser gly ala thr pro val met arg asp ser glu his phe phe phe asp leu pro
ser phe ser glu

601/201
661/221
ATG TTG CAG GCA TGG ACC CCG AGC GGG TGG CAG GAG CAG GTG GCA AAT AAA
ATG CAG GAG TGG TTT GAA TCT GGC CTG CAA CAG TGG GAT ATC TCC CGC GAC GCC
CCT TAC TTC GGT
met leu gln ala trp thr arg ser gly ala leu gln glu gln val ala asn lys
met gln glu trp phe glu ser gly leu gln gln trp asp ile ser arg asp ala
pro tyr phe gly

721/241
781/261
TTT GAA ATT CCG AAC GCG CCG GGC AAA TAT TTC TAC GTC TGG CTG GAC GCA CCG
ATT GGC TAC ATG GGT TCT TTT AAG AAT CTG TGG GAC AAG CGC GGC GAC AGC GTA
AGC TCC GAT GAA
phe glu ile pro asn ala pro gly lys tyr phe tyr val trp leu asp ala pro
ile gly tyr met gly ser phe lys asn leu cys asp lys arg gly asp ser val
ser phe asp glu

841/281
871/291
901/301
931/311
TAC TGG AAG AAA GAC TCC ACC GCC GAG CTG TAC CAC TTC ATC GGT AAA GAT ATT
GTT TAC TTT CAC AGC CTG TTC TGG CCT GCC ATG CTG GAA GGC AGC AAC TTC CGC
AAG CCG TCC GAT
tyr trp lys lys asp ser thr ala glu leu tyr his phe ile gly lys asp ile
val tyr
phe his ser leu phe trp pro ala met leu glu gly ser asn phe arg lys pro
ser asn

961/321
1021/341
991/331
1051/351
CTG TTT GTT CAT GCC TAT GTG AGC GTG AAG GCC GCA AAG ATG TCC AAG TCT CGC
GCC ACC TTT ATT AAA GCC AGC ACC TGG CTG AAT CAT TTT GAC GCA GAC AGC CTG
CGT TAC TAC TAC
leu phe val his gly tyr val thr val asn gly ala lys met ser lys ser arg
gly thr phe ile lys ala ser thr trp leu asn his phe asp ala asp ser leu
arg tyr tyr tyr

Continued next page
APPENDIX F Continued

1081/361
1141/381
1111/371
ACT GCG AAA CTC TCT TCG CGC ATT GAT GAT ATC GAT CTC AAC CTG GAA GAT TTC
GTT CAG CGT GTG AAT GCC GAT ATC GTT AAC AAA GTG GTT AAC CTG GCC TCC CGT
AAT GCG GCC TTT
thr ala lys leu ser ser arg ile asp asp ile asp leu asn leu glu asp phe
val gln arg val asn ala asp ile val asn lys val val asn leu ala ser arg
asn ala gly phe

1201/401
1261/421
1231/411
ATC AAC AAG CGT TTT GAC GCC GTG CTA AGC GAA CTG GCT GAC CCG CAG TTC
TAC AAA ACC TTC ACT GAT GCC GCT GAA GTG ATT GTG GAA GCC TGG GAA AGC CGT
GAA TTT GTT AAA
il e asn lys arg phe asp gly val leu ala ser gly leu ala asp pro gln leu
ty r lys thr phe thr asp ala ala glu val ile gly glu ala trp glu ser arg
glu phe gly lys

1321/441
1381/461
1351/451
1411/471
GCC GTG CGC GAA ATC ATG GCG CTG GAT CTG GCT GAA AAC CGC TAT GTC GAT GAA
CAG GCT CGT GCA GCA GCT GCT GAA AAA CAG GAA GCC CGC GAT GCC GAC CTG CAG GCA
ATT TGC TCA ATG
ala val arg gly ile met ala leu ala asp leu ala asn arg tyr val asp glu
qln ala pro trp val val ala lys gln glu gly arg asp ala asp leu gln ala
ile cys ser met

1441/481
1501/501
1471/491
1531/511
GGC ATC AAC CTG TTC CGC GTG CTA ATG ACT TAC CTG AAG CGG GTA CTG CCG AAA
CTG ACC GAG CGT GCA GAA GCA TTC CTC AAT AGC GAA CGT ACC TGG GAT GCT ATC
CAG CAA CGG CTG
gly ile asn leu phe arg val leu met thr tyr leu lys pro val leu pro lys
leu thr glu arg ala glu ala phe leu asn thr glu leu thr trp asp gly ile
gln gln pro leu

1561/521
1621/541
1591/531
CTG GGC CAC AAA GTG AAT CGG TTC AAG GCG CTG TAT AAC CGC ATG GAT ATG AGG
CAG GTT GAA GCA CTG GTG GAA GCC TCT AAA TGA
leu gly his lys val asn pro phe lys ala leu tyr asn arg ile asp met arg
gln val glu ala leu val glu ala ser lys OPA


Edelmann, P. and Gallant, J.A. Cell 1977, 10, 131-137.


