BIOCHEMICAL PROPERTIES OF CLASS I LYSYL-tRNA SYNTHETASE

DISSERTATION

A Thesis
Presented in Partial Fulfillment of the Requirements for
The Degree Doctor of Philosophy in the Graduate School of
The Ohio State University

By
Jeffrey David Levengood, B.S.

*****
The Ohio State University
2007

Dissertation Committee:
Professor Michael Ibba, Adviser
Professor Russ Hille
Professor John Reeve
Professor Zucai Suo

Approved by

______________________________
Adviser
The Ohio State Biochemistry Graduate Program
ABSTRACT

The family of aminoacyl-tRNA synthetases (aaRSs) performs the essential cellular function of charging tRNA molecules with their cognate amino acids. This enzyme family can be divided into two unrelated classes with each evolving from different origins. Lysyl-tRNA synthetase (LysRS) is the only synthetase known to have a form in each class. The class I form (LysRS1) is found in most archaea and a few bacteria while the class II form (LysRS2) is found in all eucarya, most bacteria, and a few archaea.

Steady-state kinetics were used to study the mechanisms which LysRS1 employs to recognize its substrates. The binding of both amino acid and tRNA were examined. The binding of lysine was analyzed by studying the ability of several lysine analogues to inhibit the aminoacylation reaction. It was found that the R-group plays a critical role in discrimination rather than the functional groups at the α-carbon. Within the R-group, the size of the chain was a very important point for discrimination.

The binding of tRNA^{Lys} was examined using both LysRS, mutated based on the modeled tRNA^{Lys} bound to LysRS1, and tRNA^{Lys} with mutations in the anticodon.
Analysis of the bound tRNA\(^{\text{Lys}}\) revealed that LysRS is able to specifically recognize the bases at anticodon positions 35 and 36, but not position 34. The experiments with variant enzymes in relation to the model revealed LysRS1 likely uses a variety of bonds to recognize the anticodon. Aromatic residues form non-specific stacking interactions with the bases, charged and polar residues form specific hydrogen bonds with the bases, and other charged residues form non-specific electrostatic interactions with the phosphate backbone of the anticodon stem. There was also evidence that single point mutations in amino acid residues binding the anticodon could have major effects on anticodon recognition and the ability to rescue tRNA mutations.

Comparison of the mechanisms for substrate recognition between LysRS1 and LysRS2 revealed differences in the recognition of substrates that could be the cause for the divergence of the two enzymes. The two LysRSs were shown to have different patterns of inhibition to naturally occurring lysine analogues that had consequences for growth \textit{in vivo}. With the tRNA\(^{\text{Lys}}\) anticodon binding, it was found anticodon binding is much more important for binding with LysRS2 than LysRS1. This could have had consequences with the divergence of other synthetases that must recognize similar anticodons. The presence of LysRS1 may have allowed these other synthetases to put more emphasis on the anticodon in recognition of their cognate tRNA.
Dedicated to my family, who have supported me through everything.
ACKNOWLEDGMENTS

I wish to express my gratitude to my advisor, Dr. Michael Ibba, for his guidance, intellectual support, and endless patience with me as I worked through some of the rough patches of my project. I consider myself lucky to have had an advisor who proved to be such a wealth of knowledge.

I would like to thank my committee members, Dr. Russ Hille, Dr. John Reeve, and Dr. Zucai Suo, for their insight over the years.

I would like to thank Dr. Sandro Ataide and Dr. Herve Roy for their technical support and insight.

I thank Corinne Hausmann and Jiqiang Ling for their support, feedback, and critical review.

I wish to thank the past and present laboratory members of the Ibba lab for providing a supportive and comfortable intellectual environment. I would also like to thank Noah Reynolds and Theresa Rogers for critical review.

This work was supported by grants from the American Heart Association, the National Institutes of Health, and the National Science Foundation.
VITA

September 11, 1978. . . . . . . . . . . . . . . . . . . . Born-Cincinnati, Ohio, USA

2001. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . B.S., Biochemistry, The Ohio State University

2001-present. . . . . . . . . . . . . . . . . . . . . . . . . . OSBP fellow and Graduate Research Associate, The Ohio State University

PUBLICATIONS


FIELDS OF STUDY

Major Field: Biochemistry
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>ii</td>
</tr>
<tr>
<td>Dedication</td>
<td>iv</td>
</tr>
<tr>
<td>Acknowledgements</td>
<td>v</td>
</tr>
<tr>
<td>Vita</td>
<td>vi</td>
</tr>
<tr>
<td>List of Tables</td>
<td>x</td>
</tr>
<tr>
<td>List of Figures</td>
<td>xi</td>
</tr>
<tr>
<td>List of Symbols/Abbreviations</td>
<td>xiii</td>
</tr>
<tr>
<td>Chapters:</td>
<td></td>
</tr>
<tr>
<td>1. Introduction</td>
<td>1</td>
</tr>
<tr>
<td>1.1. The ribosome and protein translation</td>
<td>1</td>
</tr>
<tr>
<td>1.2. Translational fidelity and quality control</td>
<td>2</td>
</tr>
<tr>
<td>1.3. The aminoacyl-tRNA synthetases</td>
<td>5</td>
</tr>
<tr>
<td>1.4. Classes of the aaRSs</td>
<td>6</td>
</tr>
<tr>
<td>1.5. The synthetase aminoacylation reaction mechanism</td>
<td>9</td>
</tr>
<tr>
<td>1.6. Substrate recognition by aaRSs</td>
<td>12</td>
</tr>
<tr>
<td>1.7. tRNA: the adaptor molecule and its recognition by aaRSs</td>
<td>13</td>
</tr>
<tr>
<td>1.8. Amino acid discrimination by the synthetases</td>
<td>17</td>
</tr>
<tr>
<td>1.9. Synthetase universality</td>
<td>18</td>
</tr>
<tr>
<td>1.10. Class I LysRS</td>
<td>21</td>
</tr>
<tr>
<td>1.11. Distribution of LysRS</td>
<td>21</td>
</tr>
</tbody>
</table>


LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Subclasses of the aminoacyl-tRNA synthetases</td>
<td>8</td>
</tr>
<tr>
<td>2.1</td>
<td>Inhibition of LysRS1 by lysine analogues</td>
<td>42</td>
</tr>
<tr>
<td>2.2</td>
<td>Group I and II lysine analogues</td>
<td>48</td>
</tr>
<tr>
<td>3.1</td>
<td>Kinetic data for wild type LysRS1 with mutant tRNAs</td>
<td>69</td>
</tr>
<tr>
<td>3.2</td>
<td>Kinetic data for LysRS1 variant Phe487Ala with mutant tRNAs</td>
<td>71</td>
</tr>
<tr>
<td>3.3</td>
<td>Kinetic data for LysRS1 variant Tyr491Ala with mutant tRNAs</td>
<td>73</td>
</tr>
<tr>
<td>3.4</td>
<td>Kinetic data obtained for LysRS1 variant Tyr491Glu with mutant tRNAs</td>
<td>75</td>
</tr>
<tr>
<td>3.5</td>
<td>Kinetic data for LysRS1 variant Arg502Ala with mutant tRNAs</td>
<td>77</td>
</tr>
<tr>
<td>3.6</td>
<td>Kinetic data for LysRS1 variant Arg502Gln with mutant tRNAs</td>
<td>79</td>
</tr>
<tr>
<td>4.1</td>
<td>Kinetic parameters for lysine analogues</td>
<td>92</td>
</tr>
<tr>
<td>4.2</td>
<td>Kinetic data for wild type LysRS1 and LysRS2 with mutant tRNAs</td>
<td>104</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Translation and the processes involved in quality control</td>
<td>4</td>
</tr>
<tr>
<td>1.2</td>
<td>Schematic of the editing mechanism for ValRS</td>
<td>11</td>
</tr>
<tr>
<td>1.3</td>
<td>The secondary and tertiary structures of tRNA</td>
<td>14</td>
</tr>
<tr>
<td>1.4</td>
<td>3D ribbon model of the tertiary structure of tRNA</td>
<td>15</td>
</tr>
<tr>
<td>1.5</td>
<td>Two step pathway for Gln-tRNA&lt;sub&gt;Gln&lt;/sub&gt; formation</td>
<td>18</td>
</tr>
<tr>
<td>1.6</td>
<td>The two step pathway for the formation of Cys-tRNA&lt;sub&gt;Cys&lt;/sub&gt;</td>
<td>20</td>
</tr>
<tr>
<td>1.7</td>
<td>The distribution of LysRS</td>
<td>22</td>
</tr>
<tr>
<td>1.8</td>
<td>Crystal structure of LysRS2</td>
<td>24</td>
</tr>
<tr>
<td>1.9</td>
<td>Structure of the lysine binding site for &lt;i&gt;E. coli&lt;/i&gt; LysRS2</td>
<td>25</td>
</tr>
<tr>
<td>1.10</td>
<td>Crystal structure of LysRS1 from &lt;i&gt;P. horikoshii&lt;/i&gt;</td>
<td>27</td>
</tr>
<tr>
<td>1.11</td>
<td>The structure of the lysine binding site for LysRS1 from &lt;i&gt;P. horikoshii&lt;/i&gt;</td>
<td>28</td>
</tr>
<tr>
<td>2.1</td>
<td>LysRS1 sequence alignments</td>
<td>38</td>
</tr>
<tr>
<td>2.2</td>
<td>Identity plot of the LysRS1 active site</td>
<td>39</td>
</tr>
<tr>
<td>2.3</td>
<td>Structures of lysine and lysine analogues used in aminoacylation assays</td>
<td>41</td>
</tr>
<tr>
<td>2.4</td>
<td>The amino acid binding pocket of LysRS1 from &lt;i&gt;Pyrococcus horikoshii&lt;/i&gt;</td>
<td>45</td>
</tr>
<tr>
<td>2.5</td>
<td>Structure of S-(2-aminoethyl)-L-cysteine modeled in the active site of &lt;i&gt;P. horikoshii&lt;/i&gt; LysRS1</td>
<td>52</td>
</tr>
</tbody>
</table>
3.1. Model of LysRS1:tRNA^Lys complex.................................................................61
3.2. Structure of GluRS:tRNA^Glu complex...........................................................62
3.3. View of the tRNA^Lys anticodon when bound to LysRS1..............................64
3.4. α-helix cage domain alignments......................................................................66
4.1. Analysis of growth inhibition by AEC using disk assays............................93
4.2. Growth curves of B. subtilis strains.................................................................94
4.3. Diagram of synthetase evolution....................................................................96
4.4. Model for the binding of AEC to LysRS1 (P. horikoshii) and LysRS2 (E. coli) .................................................................98
4.5. Northern analysis of tRNA^Lys charging in B. subtilis.................................101
4.6. Northern analysis of tRNA^Lys from B. burgdorferi charged in B. subtilis ....102
4.7. LysRS2 recognition of the anticodon.............................................................105
Lists of Symbols/Abbreviations

A    adenosine
Å    angstrom
aa   amino acid
AABA  α-aminobutyric acid
aaRS aminoacyl-tRNA synthetase
       (three letter amino acid abbreviation plus suffix RS)
aa-tRNA aminoacyl-tRNA
AEC  S-(2-aminoethyl)-L-cysteine
Ala   alanine
AMP  adenosine monophosphate
Arg   L-arginine
Asp   L-Aspartate
ATP  adenosine triphosphate
BSA  bovine serum albumin
C    cytosine
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Chemical Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>CTP</td>
<td>cytosine triphosphate</td>
</tr>
<tr>
<td>Cys</td>
<td>L-cysteine</td>
</tr>
<tr>
<td>DNA</td>
<td>deoxyribonucleic acid</td>
</tr>
<tr>
<td>DTT</td>
<td>dithiothreitol</td>
</tr>
<tr>
<td>EDTA</td>
<td>ethylenediaminetetraacetic acid</td>
</tr>
<tr>
<td>EF-G</td>
<td>elongation factor G</td>
</tr>
<tr>
<td>EF-Tu</td>
<td>elongation factor Tu</td>
</tr>
<tr>
<td>FPLC</td>
<td>Fast protein liquid chromatography</td>
</tr>
<tr>
<td>g</td>
<td>gram</td>
</tr>
<tr>
<td>G</td>
<td>guanosine</td>
</tr>
<tr>
<td>xg</td>
<td>gravitational constant</td>
</tr>
<tr>
<td>GABA</td>
<td>γ-aminobutyric acid</td>
</tr>
<tr>
<td>Gln</td>
<td>L-glutamine</td>
</tr>
<tr>
<td>Glu</td>
<td>L-glutamate</td>
</tr>
<tr>
<td>Gly</td>
<td>L-glycine</td>
</tr>
<tr>
<td>GMP</td>
<td>guanosine monophosphate</td>
</tr>
<tr>
<td>GTP</td>
<td>guanosine triphosphate</td>
</tr>
<tr>
<td>HEPES</td>
<td>4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid</td>
</tr>
<tr>
<td>His</td>
<td>L-histidine</td>
</tr>
<tr>
<td>IF</td>
<td>initiation factor</td>
</tr>
<tr>
<td>Ile</td>
<td>Isoleucine</td>
</tr>
<tr>
<td>IPTG</td>
<td>Isopropyl β-D-1-thiogalactopyranoside</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>Leu</td>
<td>L-leucine</td>
</tr>
<tr>
<td>Lys</td>
<td>L-lysine</td>
</tr>
<tr>
<td>Lys-eth</td>
<td>L-lysine ethyl ester</td>
</tr>
<tr>
<td>Lys-met</td>
<td>L-lysine methyl ester</td>
</tr>
<tr>
<td>Lys-OH</td>
<td>lysine hydroxamate</td>
</tr>
<tr>
<td>Lys-SA</td>
<td>Lysyl-sulfamoyl adenosine</td>
</tr>
<tr>
<td>m</td>
<td>milli</td>
</tr>
<tr>
<td>M</td>
<td>molar</td>
</tr>
<tr>
<td>mRNA</td>
<td>messenger RNA</td>
</tr>
<tr>
<td>n</td>
<td>nano</td>
</tr>
<tr>
<td>N</td>
<td>nucleotide</td>
</tr>
<tr>
<td>OB</td>
<td>oligonucleotide binding</td>
</tr>
<tr>
<td>OD</td>
<td>optical density</td>
</tr>
<tr>
<td>OH</td>
<td>hydroxyl</td>
</tr>
<tr>
<td>PAGE</td>
<td>polyacrilamide gel electrophoresis</td>
</tr>
<tr>
<td>Phe</td>
<td>L-phenylalanine</td>
</tr>
<tr>
<td>RNA</td>
<td>ribonucleic acid</td>
</tr>
<tr>
<td>RNAsin</td>
<td>RNA nuclease inhibitor</td>
</tr>
<tr>
<td>s</td>
<td>seconds</td>
</tr>
<tr>
<td>SDS</td>
<td>sodium dodecyl sulphate</td>
</tr>
<tr>
<td>Sep</td>
<td>O-phosphoserine</td>
</tr>
<tr>
<td>Ser</td>
<td>L-serine</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Chemical Name</td>
</tr>
<tr>
<td>--------------</td>
<td>------------------------</td>
</tr>
<tr>
<td>T</td>
<td>thymidine</td>
</tr>
<tr>
<td>Thr</td>
<td>L-threonine</td>
</tr>
<tr>
<td>tRNA</td>
<td>transfer RNA</td>
</tr>
<tr>
<td>Trp</td>
<td>L-tryptophan</td>
</tr>
<tr>
<td>Tyr</td>
<td>L-tyrosine</td>
</tr>
<tr>
<td>U</td>
<td>urosine</td>
</tr>
<tr>
<td>UTP</td>
<td>uridine triphosphate</td>
</tr>
<tr>
<td>Val</td>
<td>L-valine</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

Translation, central to all living cells, is the process of decoding nucleic acid sequences into polypeptides. Translation is carried out by several enzymes and macromolecules, most notably aaRS, tRNA, and the ribosome (Laursen et al., 2005). Aminoacyl-tRNA synthetases (aaRSs) are the enzymes responsible for attaching the correct amino acid onto its cognate tRNA. These aa-tRNAs, which serve as the adaptor molecules between the RNA and protein sequence in translation, may then be bound by EF-Tu and delivered to the ribosome as it reads through the mRNA code. The tRNAs contain anticodons designed to bind the codons on the mRNA strands, an action that is mediated by the ribosome. Through these actions the tRNA and ribosome can translate an mRNA sequence into a protein sequence (Crick, 1970).

1.1. The ribosome and protein translation

The ribosome is a macromolecule of immense size that is composed of numerous RNA strands and peptides which form the two major components of
ribosomes. In bacteria these two components are the smaller 30S subunit and the larger 50S subunit (Laursen et al., 2005). The ribosome is the center of protein translation, which can be divided into three phases: initiation, elongation, and termination.

Protein synthesis is initiated in the ribosome with the aid of IF-2. This GTPase binds the 30S ribosome subunit, GTP, and the initiator tRNA fMet-tRNA$^{fMet}$. As the 30S subunit binds the 50S subunit IF-2 places the fMet-tRNA$^{fMet}$ in the ribosomal P site and then dissociates as its GTP is hydrolyzed to GDP (Laursen et al., 2005). Protein elongation then occurs with the aid of elongation factor Tu (EF-Tu) (Rodnina et al., 2001), which binds aa-tRNA in a ternary complex with GTP (Asahara and Uhlenbeck, 2005) and delivers the aa-tRNA to the A site with the hydrolysis of GTP to GDP. After peptide transfer the ribosome goes through translocation, which is promoted by EF-G and the hydrolysis of GTP to GDP (Fraser and Hershey, 2005). The process is repeated until a stop codon is reached on the mRNA. When the termination codon is reached, release factors bind in the A site and release the polypeptide along with the mRNA (Urakov et al., 2006).

1.2. Translational fidelity and quality control

Because of the possibility that the incorporation of one wrong amino acid could render an enzyme inactive, translation must be carried out with great fidelity. Cells employ various mechanisms to ensure each codon is translated accurately (Prætorious-Ibba et al., 2006). These mechanisms result in about 1 in every 10,000
codons being mistranslated, which is lower than the tolerable rate of error for several codons.

Translational fidelity is determined by three processes. First, aaRSs must charge tRNA with its cognate amino acid. In the second process, EF-Tu must recruit the proper aa-tRNA for the codon in the A site of the ribosome (Ibba & Söll, 2001). Finally, the ribosome must match the anticodon to the mRNA codon so translocation can occur and the amino acid can be added to the polypeptide. Figure 1.1 illustrates this process.
Figure 1.1. Translation and the processes involved in quality control. The synthetase must bind the correct amino acid out of a pool of 20 as well as the correct tRNA out of a pool of ~46. It must then bind the two together to form aa-tRNA. The charged tRNA can then be selected by EF-Tu and used in translation. The ribosome must match the anticodon of aa-tRNA to the codon of mRNA. If the process is carried out with fidelity a functional protein is obtained. Adapted from Roy and Ibba, 2006.
While EF-Tu and the ribosome have some ability to screen the tRNAs for proper charging, these activities are limited and insufficient for complete prevention of errors caused by mischarged tRNAs. Because of this limitation it is imperative the aaRSs carry out their respective reactions with a low a level of error. One example of the ability of quality control mechanisms to prevent misincorporation can be seen with IleRS and its discrimination against Leu (Ibba and Söll, 1999). Leu is incorporated into Ile codons at a rate of only 1 in 3000. A large part of this low error rate is attributed to the editing mechanism that IleRS employs (see below).

1.3. The aminoacyl-tRNA synthetases

The aaRSs serve the role of charging tRNA with its cognate amino acid. The synthetases do this through a two step reaction known as aminoacylation.

\[
AA + ATP + AARS \leftrightarrow AARS\cdot AA-AMP + PP_i \quad (1)
\]

\[
AARS\cdot AA-AMP + tRNA \leftrightarrow AARS + AA-tRNA + AMP \quad (2)
\]

In the first step (1) the amino acid is activated with ATP to form an aminoacyl-adenylate intermediate. In the second step (2) the amino acid is transferred to the tRNA, which binds through its α-carboxyl to the 2´ or 3´ hydroxyl of the terminal adenosine (Ibba & Söll, 2000).
1.4. Classes of the aaRSs

In order for aminoacylation to be completed with a high degree of fidelity, it was theorized in Crick’s adaptor hypothesis that an aaRS would be needed for each amino acid (Crick, 1958). The initial isolation and analysis of the enzymes’ structure and sequence showed there were twenty synthetases which could be divided into two classes, class I and class II. Both classes contain ten synthetases, one for each canonical amino acid. These two classes are distinct from each other in both sequence and structure, but perform the same basic function (Ibba & Söll, 2000; Zhang et al., 2006a).

Both synthetase classes follow the two step mechanism in catalyzing aminoacylation, but have different motifs for doing so (Ibba & Söll, 2000; Zhang et al., 2006a). The active site of class I enzymes is composed of a Rossman fold domain, which contains conserved sequences KMSKS and HIGH that stabilize the transition state of amino acid activation. From a mechanistic standpoint, the class I synthetases generally bind to the minor groove of the tRNA, bind ATP in an extended conformation, and attach the amino acid to the 2′-OH of the terminal adenosine. The active site of class II aaRSs is composed of strands of antiparallel β-sheets. It recognizes the major groove of tRNA, binds ATP in a bent conformation, and attaches its amino acid to the 3′-OH of the terminal adenosine. Within the classes there are differences in the way the enzymes recognize the tRNA acceptor stem. These differences have been used to further break down each class into three subclasses, a, b, and c (Ribas de Pouplana and Schimmel, 2001a). This division into subclasses is based on both mechanistic and structural studies. While
the division is primarily based on the binding of tRNA, breakdowns can also be made by the types of amino acids the synthetases bind. For example, subclass Ia synthetases generally charge hydrophobic amino acids (Delagoutte et al., 2000).
<table>
<thead>
<tr>
<th>Class I</th>
<th>Class II</th>
</tr>
</thead>
<tbody>
<tr>
<td>1a</td>
<td>2a</td>
</tr>
<tr>
<td>ValRS (α)</td>
<td>SerRS (α₂)</td>
</tr>
<tr>
<td>LeuRS (α)</td>
<td>ThrRS (α₂)</td>
</tr>
<tr>
<td>IleRS (α)</td>
<td>GlyRS (α₂)</td>
</tr>
<tr>
<td>ArgRS (α)</td>
<td>ProRS (α₂)</td>
</tr>
<tr>
<td>CysRS (α)</td>
<td>HisRS (α₂)</td>
</tr>
<tr>
<td>MetRS (α)</td>
<td></td>
</tr>
<tr>
<td>1b</td>
<td>2b</td>
</tr>
<tr>
<td>GluRS (α)</td>
<td>AspRS (α₂)</td>
</tr>
<tr>
<td>GlnRS (α)</td>
<td>AsnRS (α₂)</td>
</tr>
<tr>
<td>LysRS1 (α)</td>
<td>LysRS2 (α₂)</td>
</tr>
<tr>
<td>1c</td>
<td>2c</td>
</tr>
<tr>
<td>TyrRS (α)</td>
<td>PheRS [(αβ)₂, α]</td>
</tr>
<tr>
<td>TrpRS (α)</td>
<td>AlaRS (α, α₄)</td>
</tr>
<tr>
<td></td>
<td>GlyRS (αβ)</td>
</tr>
<tr>
<td></td>
<td>PylRS¹</td>
</tr>
<tr>
<td></td>
<td>SepRS¹</td>
</tr>
</tbody>
</table>

Table 1.1. Subclasses of the aminoacyl-tRNA synthetases. The oligomer form of each enzyme is in parenthesis. Each synthetase is denoted by its three letter amino acid.

¹The oligomerization of these synthetases is not known.
The conserved structures and sequences of each class are primarily involved in ATP binding. For example, the class I sequences KMSKS and HIGH are amino acid residues that have the function of binding and orienting the ATP so it can react with the $\alpha$-carboxylate of the amino acid (First et al., 1995). This can be seen in GlnRS as the KMSKS is part of a loop which folds over the binding site. The adenine base of ATP fits between the Gly-His of the HIGH motif. The two His residues of this motif interact with the triphosphate both directly and with the aid of water molecules. The second Lys of the KMSKS motif interacts with the phosphates of ATP. These aaRS:tRNA interactions stabilize the transition state at the $\alpha$-phosphate (Arnez and Moras, 1997).

1.5. The synthetase aminoacylation reaction mechanism

The synthetases catalyze the aminoacylation reaction by maximizing their contacts with the transition state (Fersht et al, 1985). In both classes, the reaction occurs in the same manner (Ibba and Söll, 2000). First, the enzymes bind ATP and the amino acid. The $\alpha$-phosphate of the ATP and the $\alpha$-carboxylate of the amino acid must be positioned so that the two can participate in an in-line nucleophilic displacement mechanism. The result of this reaction is an inorganic pyrophosphate that exits the enzyme and an aminoacyl-adenylate that stays bound to the enzyme. The second step of the reaction is the binding of tRNA in a position so that the 2$'$ or 3$'$ hydroxyl of the terminal adenosine can attack as a nucleophile the $\alpha$-carbonyl of the aminoacyl-adenylate. The amino acid and the tRNA are then linked together by an ester bond with the AMP and tRNA dissociating from the enzyme.
Recently, a slight divergence has been observed in the reaction mechanism between the class I and class II enzymes. The divergence is in the rate-determining step for the enzyme (Zhang et al., 2006a). Kinetic studies showed class I enzymes were rate limited by the release of product, aminoacylated tRNA. In contrast, class II enzymes likely have the formation of adenylate as their rate limiting step.

Several synthetases in each class possess an editing reaction as part of their mechanism. This reaction proofreads the charged tRNA. The editing reaction is part of a double sieve mechanism for amino acid discrimination (Beebe et al., 2003). The first sieve is the initial charging reaction in which amino acids larger than the cognate amino acid are discriminated against. The second sieve is the editing reaction where mischarged tRNAs bind in the editing site and are hydrolyzed to give an uncharged tRNA which can then be correctly aminoacylated. The editing site accommodates the aminoacylated end of tRNA if a smaller amino acid is attached. Mischarged amino acids may also be recognized through chemical differences. If the tRNA already has the correct amino acid bound it will be too large to fit in the editing site and the aa-tRNA will dissociate.

The double sieve mechanism can be seen in ValRS (Fukai et al., 2000). In the first step the sieve discriminates against the larger isoleucine (Figure 1.2). Val and Thr are both recognized and charged onto the tRNA. In the editing reaction the mischarged Thr is recognized by the hydroxyl on its side chain. It is accommodated in the editing site and hydrolyzed from the tRNA. The smaller amino acids Cys and α-aminobutyric acid are also charged in the first step, but they fit into the editing site and are hydrolyzed from the tRNA.
Figure 1.2. Schematic of the editing mechanism for ValRS. In the aminoacylation reaction isoleucine cannot be charged because of its larger size. Thr can be accommodated in the active site and bound to tRNA. It then is hydrolyzed as it fits into the editing site due to its hydrophilicity. Val-tRNA\textsuperscript{Val} does not fit and is thus released for protein synthesis. Adapted from Fukai et al., 2000.
Misincorporation due to the lack of editing can have serious effects. Single point mutations can deleteriously affect enzyme activity and lead to pathological diseases. Defects in editing that cause higher rates of misincorporation in protein synthesis have been shown to cause several neurological diseases (Lee et al., 2006; Nangle et al., 2006; Seburn et al., 2006).

A slight exception to the normal aminoacylation reaction mechanism is found in the synthetases ArgRS, GluRS, and GlnRS. In these enzymes the tRNA needs to bind before the amino acid and ATP can react (Francklyn et al., 2002). In all three cases, it was shown that tRNA needs to be bound in order for the enzyme to undergo conformational changes, which create the ATP and amino acid binding sites for formation of the aminoacyl-adenylate intermediate (Sekine et al., 2003; Guigou et al., 2005; Uter and Perona, 2006). After this step, however, no further conformational changes are needed for the transfer step (Zhang et al., 2006a). This induced fit mechanism helps improve recognition of the amino acid (Ataide and Ibba, 2006).

1.6. Substrate recognition by aaRSs

While the catalytic mechanism for synthetases is similar, each enzyme has developed its own method for recognizing its cognate amino acid and tRNA. The tRNA molecules all share the same basic tertiary structure while amino acids only differ in their side chains (Hou, 1997; Ataide and Ibba, 2006). Having to differentiate between similar substrates caused the synthetases to evolve precise mechanisms for recognizing their substrates. To differentiate tRNA, differences in
primary sequence can be utilized. For the amino acids, discrimination has been focused on recognizing differences in the size and charge of the molecules. These mechanisms of substrate recognition and discrimination are described below in more detail.

1.7. tRNA: the adaptor molecule and its recognition by aaRSs

The tRNA molecules are strands generally 76 nucleotides long which adopt a cloverleaf secondary conformation formed by self-complementary base pairing (Schimmel and Ribas de Pouplana, 1995; Ribas de Pouplana and Schimmel, 2001b). The nucleotide strand forms the cloverleaf structure with five arms: the acceptor arm, the anticodon loop, the D arm, the variable loop, and the T loop, all shown below (Figures 1.3 and 1.4).
Figure 1.3. The secondary and tertiary structures of tRNA. The tRNA is shown in its cloverleaf secondary structure (left) with each stem and loop labeled. These four loops fold into a tertiary structure that is shown in a two-dimensional form. Modified from Musier-Forsyth and Schimmel, 1999.
Figure 1.4. 3D ribbon model of the tertiary structure of tRNA. Nucleotides labeled are the anticodon nucleotides (34, 35, 36) and the discriminator base (73). These labeled nucleotides are often used as identity elements for recognition by the synthetases. Modified from Geige et al., 1998.
The two most important structures of the tRNA are the acceptor stem and the anticodon stem-loop (Figure 1.3). The acceptor stem is the section of tRNA where the amino acid is attached. All tRNAs end with the 3’ sequence CCA, with the amino acid being bound to the terminal adenosine after aminoacylation. The anticodon loop contains the anticodon at positions 34-36 of the tRNA strand (Musier-Forsyth and Schimmel, 1999).

The cloverleaf structure folds into an L-shaped tertiary structure containing two domains (Figures 1.3 and 1.4): the acceptor stem and the anticodon stem (Musier-Forsyth and Schimmel, 1999). The acceptor stem is a 12 bp minihelix that includes the 7 bp acceptor arm and is formed by the acceptor arm stacking with the T arm. RNA substrates composed of or resembling this minimal domain are sometimes substrates for the synthetases (Hou et al., 2000). The anticodon stem contains a 10 bp helix with a loop at the end and is formed by the anticodon loop stacking with the D loop (Ribas de Pouplana and Schimmel, 2001b).

The recognition of tRNA by the synthetases is dependent on the position or modification of particular nucleotides that compose a tRNA’s identity set. These can be either positive determinants which enhance aminoacylation, or negative determinants which prevent aminoacylation (Fukunaga et al., 2006). The identity elements on tRNA are generally found in the acceptor and anticodon stems. In a few instances, they are also located on the D arm, T arm, and variable loop (Figure 1.4) (Geige et al., 1998). The most common identity elements used are the anticodon nucleotides and the discriminator base at N73 which precedes the CCA end. By recognizing a combination of these various structural and sequential
elements in tRNA, the synthetases are able to select the correct isoacceptor tRNAs with exceptional fidelity.

1.8. Amino acid discrimination by the synthetases

Due to their smaller size and limited structural diversity, differentiation of amino acids is much more difficult when compared to tRNA recognition (Ataide and Ibba, 2006). Amino acid recognition has an error rate an order of magnitude greater than tRNA (1 in $10^4$ to $10^5$ for amino acid to 1 in $10^6$ for tRNA) (Ahel et al., 2003). While some synthetases can exclude non-cognate amino acids by charge or size, others require more specific interactions with their substrates. For example, SerRS differentiates between Ser and Thr by the precise positioning of the side chain. The enzymes possess a hydrophilic pocket for accommodation of the Ser side chain hydroxyl. Thr is discriminated against by the hydrophilic pocket which excludes the methyl group of the Thr side chain (Bilokapic et al., 2006).

The synthetases not only have to guard against the other canonical amino acids, they must also guard against substrate analogues which can inhibit the formation of properly charged tRNA (Ataide and Ibba, 2006). For example, indolmycin, a derivative of tryptophan, has been shown to be a competitive inhibitor of TrpRS (Wemer et al., 1976). In order to guard against this inhibitor, some organisms contain two copies of TrpRS. One copy is the normal housekeeping gene while the other confers resistance to the inhibitor on the organism. A similar phenomenon has also been seen with the synthetases TyrRS, ThrRS, MetRS, and IleRS (Ataide and Ibba, 2006).
1.9. Synthetase universality

The discovery of synthetase duplication showed this family of enzymes could be fairly diverse. Because of the critical role synthetases play in living organisms, it was assumed they would be universally conserved throughout the living kingdoms (Ibba et al., 2000). However, in the mid-1960s, just a short time after the isolation of the aaRSs, it was discovered that several bacteria do not produce Gln-tRNA\textsuperscript{Gln} through direct aminoacylation (Wilcox and Nirenberg, 1968). In these organisms an alternative two step charging mechanism is used. The alternative reaction is shown below (Figure 1.5).

![Figure 1.5. Two step pathway for Gln-tRNA\textsuperscript{Gln} formation.](image)

Figure 1.5. Two step pathway for Gln-tRNA\textsuperscript{Gln} formation. Pathway for the direct formation of Gln-tRNA\textsuperscript{Gln} (top) and the two step amidotransferase pathway (bottom). Adapted from Schulze et al., 2006.
In the first step of the pathway tRNA\textsuperscript{Gln} is charged with Glu by a non-discriminating GluRS. In a similar manner tRNA\textsuperscript{Asn} is charged with Asp by AspRS. In organisms with this pathway, EF-Tu prevents these mischarged tRNAs from being used in protein synthesis (LaRiviere et al., 2001). An amidotransferase enzyme complex (Glu-Adt) then converts Glu-tRNA\textsuperscript{Gln} and Asp-tRNA\textsuperscript{Asn} to Gln-tRNA\textsuperscript{Gln} and Asn-tRNA\textsuperscript{Asn}, respectively. This transformation occurs by an amide donor, usually glutamine, transferring an ammonium group to the side chain carboxyl of the Glu or Asp (Ruan, et al., 2001). These pathways are widespread and in many organisms provide the only pathway for formation of these aa-tRNAs (Oshikane et al., 2006), or the only method for asparagine synthesis (Bernard et al., 2006).

Other exceptions to synthetase universality have been found since the discovery of the amidotransferase pathway. The genome sequence of the archaean organism Methanococcus jannaschii provides the starkest example of the lack of synthetase universality in the living kingdoms (Bult et al., 1996). When its genome sequence was solved, open reading frames for only 16 of the 20 synthetases were discovered. The synthetases GlnRS and AsnRS are replaced by the amidotransferase pathways mentioned above. CysRS is replaced by a similar two-step pathway (Figure 1.6) (O’Donoghue et al., 2005; Sauerwald et al., 2005).
Figure 1.6. The two step pathway for the formation of Cys-tRNA$^{\text{Cys}}$. The first step has Sep aminoacylated on the tRNA. The second step has the bound Sep converted to Cys. Adapted from O’Donoghue et al., 2005.

In this pathway tRNA$^{\text{Cys}}$ is first aminoacylated with a phosphorylated serine by O-phosphoseryl-tRNA synthetase (SepRS). A sulfur donor is then used by Sep-tRNA:Cys-tRNA synthase (SepCysS) to convert the O-phospho-serine on the tRNA into cysteine.

The final exceptions found involve the 21st and 22nd amino acids of the genetic code, selenocysteine and pyrrolysine. Both amino acids are incorporated at stop codons in mRNA by tRNAs which contain the proper anticodon. Selenocysteine is formed by a two step pathway (Francklyn et al., 2002). SerRS aminoacylates tRNA$^{\text{Sec}}$ with Ser, which is then converted to selenocysteine by Sec synthase enzymes. While selenocysteine is found in all three domains of life, the insertion machinery differs in each domain (Zhang et al., 2006b). Pyrrolysine is directly attached to its tRNA by pyrrolysyl-tRNA synthetase (Polycarpo et al., 2004; Blight et al., 2004). It is extremely rare in nature and so far has been found in only one bacterium and five archaea (Lobanov, et al., 2006).
1.10. Class I LysRS

The other canonical synthetase that was unable to be found was the class II LysRS. The absence of LysRS2 was examined in *Methanococcus maripaludis* as experiments were performed to find how Lys-tRNA$^{\text{Lys}}$ is formed in the organism (Ibba et al., 1997a). Charging with crude extract showed the tRNA was charged directly with lysine, indicating a two step charging mechanism was not used. When the enzyme responsible for the direct charging was isolated it was found to be a class I Lysyl-tRNA synthetase (LysRS1). This made LysRS the first and, to date, only synthetase to be represented in each standard class. As the two classes are not believed to share a common ancestor, this represented a unique example of convergent evolution in the aaRSs.

1.11. Distribution of LysRS

Since the sequence of LysRS1 was discovered, genomes have been scanned to determine the distribution of the LysRS enzymes. These scans have shown that the class II enzyme is in all eukaryotes, most prokaryotes, and a few archaea. The class I enzyme, meanwhile, is found in most archaea and a few bacteria (Jester et al., 2003; Levengood et al., 2004). A few bacterial and archaeal organisms contain both enzymes, such as some archaeal methanogens, and a few members of the *Bacillus* class of bacteria (Figure 1.7) (Ataide et al., 2005).
Figure 1.7. The distribution of LysRS. Organisms harboring LysRS2 are shown in green while organisms with LysRS1 are shown in pink. The phylogenetic tree was constructed from analysis of small-subunit ribosomal RNA (rRNA). The phylogenetic tree was adapted from Barns et al., 1996.
1.12. Class I and class II LysRS

The discovery of the class I LysRS family provided a novel system in which to examine the aminoacylation reaction. As LysRS2 was discovered first, it has been extensively studied and the *E. coli* crystal structure has been solved with bound lysine and the *T. thermophilus* LysRS2 structure with bound tRNA (Figure 1.8) (Onesti et al., 1995; Cusack et al., 1996). The mechanisms of lysine binding in the active site have been studied by biochemical methods (Takita et al., 1996; Ataide et al., 2004). It was found LysRS2 uses an electrostatically charged and acidic active site to bind lysine through a network of hydrogen bonds (Figure 1.9). The identity elements of tRNA that the enzyme recognizes have also been examined (McClain et al., 1990; Tamura et al., 1992; Commans et al., 1998; Francin and Mirande, 2006; Fukunaga et al., 2006). It was found the enzyme uses a network of hydrogen bonds to specifically recognize the bases of the anticodon. It has also been shown that the discriminator base is an important identity element.

From the research mentioned above, LysRS2 was shown to be a class IIb synthetase. It contains an N-terminal β-barrel domain which binds the anticodon of the tRNA (Figure 1.8). The C-terminus is composed of the antiparallel β-sheet motif that serves as the active site for the class II synthetases (Hughes et al., 2006). The binding site for ATP is in the front of the pocket while the lysine binds in the back of the cleft (Figure 1.8). The lysine binds in a negatively charged pocket where it is recognized by several hydrogen bonds. The tRNA\textsubscript{Lys} anticodon is recognized by an OB fold in the β-strands of the N-terminus. All three anticodon nucleotides are recognized by LysRS2 (Skouloubris et al., 2003).
Figure 1.8. Crystal structure of LysRS2. The structure contains bound lysine (blue), ATP (pink), and tRNA (yellow). The structure shown is a docking model made for the *E. coli* enzyme that is consistent with the *T. thermophilus* structure. Adapted from Terada et al., 2002.
Figure 1.9. Structure of the lysine binding site for *E. coli* LysRS2. L-lysine bound in the active site of *E. coli*. The amino acid residues are colored by their conservation in sequence alignments: gold, 100% identity; red, 81-99%; pink, 61-80%; white, 41-60%. For lysine, carbon is white, nitrogen is blue, and oxygen is red.
1.13. Class I LysRS structure

The solving of the crystal structure for LysRS1 from *Pyrococcus horikoshii* has so far given the best points of comparison between LysRS1 and LysRS2 (Terada et al., 2002). The crystal structure revealed the enzyme to be a class Ib synthetase. This is consistent with its tRNA-dependent requirement for aminoacyl-adenylate formation as the other Ib synthetases, GluRS and GlnRS, share this property. It contains an N-terminal Rossmann fold domain that contains the active site. The C-terminal domain is an α-helix cage domain that is involved in recognizing the tRNA anticodon (Figure 1.10.).
Figure 1.10. Crystal structure of LysRS1 from *P. horikoshii*. The substrates lysine (blue), ATP (pink), and tRNA (yellow) are shown bound to the enzyme. On the tRNA the anticodon trinucleotides are shown and labeled as is the terminal adenosine (A76) and A73. Adapted from Terada et al., 2002.
Figure 1.11. The structure of the lysine binding site for LysRS1 from *P. horikoshii*. L-lysine bound in the active site of *P. horikoshii*. The amino acid residues are colored by their conservation in sequence alignments: gold, 100% identity; red, 81-99%; pink, 61-80%; white, 41-60%. For lysine, carbon is white, nitrogen is blue, and oxygen is red.
From the crystal structure, several theories were developed concerning the way the enzyme binds its substrates. These theories are elaborated below (Levengood et al., 2004; Prætorious-Ibba et al., 2006). This dissertation describes the research that was undertaken to biochemically test these theories and examine the way that LysRS1 binds its substrates to catalyze its aminoacylation reaction.

1.14. Lysine binding to LysRS1

The *P. horikoshii* structure with bound L-lysine revealed the lysine binding pocket for the enzyme (Figure 1.11) (Levengood et al., 2004). According to the structure, the enzyme employs both electrostatic and hydrophobic interactions for L-lysine recognition and discrimination. In order to examine the binding and discrimination of L-lysine in LysRS1, steady-state kinetic experiments were performed with L-lysine and lysine analogues. Enzyme and tRNA from *Borrelia burgdorferi*, the causative agent of Lyme disease (Ibba et al., 1997b), were used. The studies were done with analogues to determine how alterations in the substrate affect binding. By examining LysRS1 binding with analogues, a mechanism for the way L-lysine specifically binds in the active site could be developed.

1.15. tRNA binding to LysRS1

The second area of these studies examined the binding of the tRNA anticodon to LysRS1. The crystal structure of LysRS1 showed the enzyme shares a distinctive C-terminus α-helix cage domain with GluRS, which has not been seen in any other aaRS structures. This homology was used to construct a model for tRNA
binding to LysRS1. From this model, several amino acids were identified that were believed to play a role in recognition of the anticodon. Using the *P. horikoshii* enzyme, these amino acids were mutated to create variant enzymes. These variants allowed us to examine the effects these specific residues might have on tRNA$^{\text{Lys}}$ recognition and aminoacylation. Lysine tRNAs from *P. horikoshii* with mutated anticodons were also constructed to test their binding with the variant enzymes. Steady-state kinetics were performed to find the catalytic efficiency of each enzyme with each tRNA.

1.16. The two classes of LysRS

The final section of this report examines the similarities and differences between the class I and II LysRSs. The two LysRSs, originating from different ancestors, evolved to recognize the same amino acid and tRNA. The experiments performed here, when coupled with work done with LysRS2, allowed for analysis of the different mechanisms that the two enzymes use to catalyze the same reaction.

Since the discovery of LysRS1, several hypotheses have been developed for the divergence and distribution of LysRS1 and LysRS2 in nature. One hypothesis was based on the proposal that differences in analogue recognition could have caused the observed distribution, as a few synthetases exist in duplicate in order to confer resistance to certain inhibitors. This occurs in *E. coli* with two forms of LysRS2 (Ataide and Ibba., 2006). One copy is the housekeeping gene while the other copy confers resistance to lysine analogues such as cadaverine. One rationale for studying the binding of lysine analogues to LysRS1 was to examine if there
were differences in the recognition of the analogues between LysRS1 and LysRS2 that could explain the distribution of the two enzymes (Levengood et al., 2004).

Anticodon recognition could also be a factor in the distribution of LysRS1 and LysRS2. Phylogenetic analysis has shown a correlation between the presence of LysRS1 and AsnRS (Söll et al., 2000) as well as LysRS1 and the presence of two GluRS enzymes (Salazaar et al., 2003). These correlations could be the result of differences in anticodon recognition between LysRS1 and LysRS2.

Overall, the studies undertaken provide a framework for the way LysRS1 binds its substrates and catalyzes its reaction. Both the elements of amino acid and tRNA important for recognition were identified. The role of several amino acids in binding the anticodon stem was also determined. Comparison of these mechanisms of LysRS1 with LysRS2 revealed differences that may have influenced the distribution of the two enzymes in nature. The nature of substrate binding for LysRS1 helped fit the enzyme into the overall phylogeny of the synthetase family and showed some of the beneficial results of diversity in the synthetase family.
Chapter 2

The Lysine Binding site of LysRS1

2.1. Introduction

Fidelity in protein translation relies on aminoacylation of tRNA with the correct amino acid and the accurate pairing of mRNA codons and the anticodons of tRNA (Ibba and Söll, 2000). The aaRSs catalyze the charging reaction, which occurs in two steps and involves binding amino acid and activating it with ATP before transferring the amino acid to tRNA. The twenty standard synthetases exist in two classes (class I and II) of ten members each (Arnez and Moras, 1997; Cusack, 1997). The two classes are not believed to share an evolutionary origin. They differ in both sequence and structure, which leads to functional differences in the way the aminoacylation reaction is catalyzed.
Each synthetase has a form in only one class. The exception to this rule is LysRS, which has both a class I and class II form. Despite their difference in origin the two enzymes recognize their substrates in much the same manner. The two enzymes recognize many of the same elements on tRNA, primarily the anticodon and discriminator base (Ibba et al., 1999; Söll et al., 2000). They do differ, however, in the mechanism for lysine activation. LysRS2 forms the aminoacyl-adenylate with just ATP and Lys while LysRS1 also requires bound tRNA^{Lys} (Ibba et al., 1999).

The lysine binding pocket was revealed by the crystal structure of LysRS1 from *P. horikoshii*, which was solved with and without bound L-lysine (Terada et al., 2002). Comparison of this structure with a similar L-lysine bound structure for LysRS2 (Onesti et al., 2000) showed their active site architectures to be fundamentally different while their strategies for recognition of the R-group are quite similar. To further investigate the recognition strategy of LysRS1, the effects of lysine analogues on the aminoacylation reaction *in vitro* were examined.

### 2.2 Materials and Methods

#### 2.2.1. Purification of the class I LysRS

The *B. burgdorferi* encoded LysRS1 was overexpressed in *E. coli* BL21-SI cells. Transformed cells were grown at 37 °C in LB media without salt and supplemented with ampicillin (100 µg/ml) to cell density of OD_{600}=0.5. Expression was induced by addition of NaCl to 0.3 M total concentration. Cells were incubated for another three hours and harvested by centrifugation. The centrifugation and all
subsequent steps were performed at 4 °C. Cells were washed in lysis buffer (50 mM NaH$_2$PO$_4$ pH 8.0, 300 mM NaCl, 5 mM DTT, and 10 mM imidazole) and resuspended in the same buffer with a protease inhibitor cocktail (Roche). Cells were lysed by sonication and cell debris was spun down at 10,000 xg by centrifugation. Qiagen 50% Ni-NTA slurry was added, mixed for ~1 hour, and then applied to an empty PD-10 column with flow through collected. Column was washed first with lysis buffer once, then wash buffer twice (lysis buffer with 20 mM imidazole). Protein was eluted with elution buffer (lysis buffer with 250 mM imidazole) and buffer exchange was performed with PD-10 desalting column (GE Healthcare) in LysRS buffer (50 mM HEPES pH 7.2, 25 mM KCl, 10 MgCl$_2$, 5 mM DTT and 10% glycerol). In LysRS buffer, the enzyme solution was further purified through size separation with a FPLC Superose gel filtration column. Fractions containing LysRS1 were pooled and concentrated through ultrafiltration (Amicon). Concentration was determined using Bradford assay (Bio-Rad) and aliquots were stored at -80 °C. Enzyme was determined to be >90% pure by SDS-PAGE gel, judged by Coomasie blue staining.

2.2.2. Purification of the \textit{B. burgdorferi} tRNA$^{\text{Lys}}$

The pUC19 plasmid containing the gene for \textit{B. burgdorferi} tRNA$^{\text{Lys}}$ was isolated from \textit{E. coli} DH5α (Qiagen Midi kit). The tRNA$^{\text{Lys}}$ gene was cut out of the plasmid using endonuclease \textit{BstnI}. The digested DNA was then used in a transcription reaction with T7 RNA polymerase. The 5 mL transcription reaction was performed in solution of 10 µg digested DNA, 40 mM Tris pH 8.1, 22 mM
MgCl₂, 2 mM spermidine, 5 mM DTT, 50 µg BSA, 600 nM T7 RNA polymerase, 4 mM ATP, 4 mM GTP, 4 mM CTP, 4 mM UTP, 20 mM 5'-GMP, 10 µL pyrophosphatase (Roche, 1 mg/ml), and 0.5 µL RNAsin (Roche). The reaction was incubated at 37 °C for 5 hours, after which it was stopped and the tRNA was isolated using phenol extraction. All subsequent steps were performed at 4 °C unless noted. The aqueous layer was isolated and tRNA was precipitated by the addition of 96% ethanol and 3M NaAC pH 5.3. For ethanol, a volume 2.5 times the sample volume was added while with NaAC, a volume one tenth the sample volume was used. The solution was incubated at -20 °C overnight. The precipitate was spun down by centrifugation and resuspended in column buffer (10 mM HEPES pH 7.3, 1 mM EDTA, 7M Urea). Transcript was heated at 80 °C and cooled on ice to promote denaturation. Transcribed tRNA was further purified by running the sample through a Q-Sepharose FPLC column. tRNA was eluted in column buffer with a gradient of NaCl (0 M-1 M). Fractions containing tRNA were pooled and once again precipitated with ethanol and NaAC, and incubated at -20 °C overnight. Precipitate was spun down by centrifugation, dried, and resuspended in 10 mM HEPES pH 7.2. tRNA was refolded by the addition of MgCl₂ (final concentration of 2 mM), heating at 80 °C for two minutes and cooling to room temperature over two hours. tRNA was quantified by finding the OD₂₆₀. Fractions were then aliquoted and stored at -80 °C.
2.2.3. Purification of T7 RNA Polymerase

T7 RNA polymerase was overexpressed in plasmid pT7-911 transformed into E. coli BL21 cells, grown in M9TB expression media at 37 °C to an OD$_{600}$=0.5-0.6, and expression was induced by the addition of IPTG to give a final concentration of 1 mM. His-tag enzyme purification was then done in a manner identical to LysRS1, stopping after the nickel column step. Enzyme was then dialyzed overnight with storage buffer (10 mM DTT, 20 mM KPO$_4$ pH 7.5, 100 mM NaCl, 50% glycerol, and 1 mM EDTA) and quantitated using the Bradford assay (Bio-Rad). Aliquots were stored at -20 °C.

2.2.4. Aminoacylation Assays

Aminoacylation was performed at 37 °C in 100 mM HEPES, pH 7.2, 25 mM KCl, 10 mM MgCl$_2$, 4 mM DTT, 5 mM ATP, 5 µM tRNA$_{Lys}$, 100 nM LysRS1 enzyme, and [$^{14}$C]L-Lys at concentrations varying between 0.2 and 5 times the $K_M$. 20-µl aliquots were taken every 30 s and spotted onto Whatman 3MM filter disks presoaked in 5% trichloroacetic acid (w/v) containing 0.5% (w/v) [$^{12}$C]-L-Lys. Sample disks were washed three times for 10 min at room temperature in 5% trichloroacetic acid (w/v) containing 0.5% (w/v) [$^{12}$C]-L-Lys, dried at 80 °C, and the level of [$^{14}$C]L-Lys-tRNA$_{Lys}$ quantified by the addition of liquid scintillant (Ultima Gold, Packard Corp.) and scintillation counting. Analogues were added as indicated during determination of $K_I$ values and [$^{14}$C]-L-Lys adjusted so that concentrations varied between 0.2 and 5 times the apparent $K_M$. 

36
2.2.5. \( K_i \) Determination

To determine \( K_i \) values for lysine analogues, at least five different concentrations of analogues were first screened in the aminoacylation reaction under standard conditions with 10 µM of \([^{14}\text{C}]\text{L-Lys}\). Analogue concentrations were then established, at which the initial rate of aminoacylation was decreased by 20–50% when compared with the reaction lacking analogue, and these levels were used for \( K_i \) determinations. Concentrations of analogues used to determine the \( K_i \) were: 150 µM L-lysine hydroxamate; 650 µM S-(2-aminoethyl)-L-cysteine (AEC); 650 µM L-lysinamide; 400 µM L-lysine methyl ester; 250 µM L-lysine ethyl ester; 650 µM DL-5-hydroxylysine; 2.5 mM L-ornithine; 4 mM D-lysine; 200 µM L-cadaverine; 20 nM lysyl-sulfamoyl adenosine; 8 mM L-\( \alpha \)-aminobutyric acid; 2 mM L-\( \gamma \)-aminobutyric acid (GABA); 2.5 mM L-arginine; 10 mM L-glutamic acid. In all cases, stock solutions of inhibitors were maintained at neutral pH. The \( K_i \) values presented represent the average of at least two independent experiments where values deviated by no more than 10% between individual determinations.

2.3. Results

All synthetases catalyze their reaction by the same two step reaction. In the first step amino acid is activated to form an aminoacyl-adenylate intermediate. In the second step the tRNA binds to the complex and has its 3' end esterified with an aminoacyl moiety. The aminoacyl-tRNA is then released from the enzyme. While both LysRS enzymes employ this basic mechanism they have a significant difference in the way they catalyze the first step. The formation of the aminoacyl-
adenylate intermediate requires bound tRNA for LysRS1, but not LysRS2. Because of this difference, we used aminoacylation in the presence of lysine analogues to examine the lysine binding pocket of LysRS1. The binding of these analogues was examined by determining the kinetics of inhibition of steady-state aminoacylation.

2.3.1. L-lysine Binding with LysRS1

The experiments were done using the heterologously produced tRNA and enzyme of *B. burgdorferi*. Sequence alignments showed *B. burgdorferi* LysRS1 contains the same active site residues as *P. horikoshii* LysRS1.

*B. burgdorferi*: 28 **SGITPSGT**VH**IGNFRE** 43  
*P. horikoshii*: 26 **SGITPSGY**VH**GNFRE** 41  
*B. burgdorferi*: 220 **WRIDWP**M**RKY**EK**VDFEPA**GK 240  
*P. horikoshii*: 218 **WRVDWP**M**RWSH**FG**VDFEPA**GK 238  
*B. burgdorferi*: 241 **DHSSGGSFS**DT**SK**NIV**KI**-IFQGSP**PVTFQY** 269  
*P. horikoshii*: 239 **DLVAGSYD**TGEI**KE**VGYKEAP**PLSL**MY 268

**Figure 2.1. LysRS1 sequence alignments.** Active site sequence alignments of *B. burgdorferi* LysRS1 with *P. horikoshii* LysRS1. Residues that compose the lysine binding site are underlined.

To further examine the conservation of the active site residues, sequence alignments were constructed from 44 LysRS1 sequences using ClustalX (Figure 2.1) (Chenna et al., 2003). The data were mapped on to the structure of *P. horikoshii* LysRS1. The identity plot showed a high degree of conservation for the lysine binding site, suggesting that the pattern of recognition for LysRS1 might be conserved (Figure 2.2).
Figure 2.2. Identity plot of the LysRS1 active site. L-lysine in the active site of *P. horikoshii* LysRS. The residues are colored according to their conservation in corresponding sequence alignments: *gold*, 100% identity; *red*, 81–99%; *pink*, 61–80%; *white*, 41–60%. For the substrate lysine, the backbone is shown in white, oxygen in red, and nitrogen in blue.
The properties of L-lysine as a substrate in the aminoacylation reaction were investigated through the use of steady-state kinetics. In these experiments the rate of the aminoacylation reaction was found for several different lysine concentrations. The rates were then plotted against L-lysine concentration and Michaelis-Menten kinetics were employed to obtain a $K_M$ (34 µM) and $k_{cat}$ (0.06 s$^{-1}$) for the binding of L-lysine to LysRS1.

2.3.2. Lysine analogue binding with LysRS1

Steady-state kinetic experiments with the lysine analogues were done in the same manner as for L-lysine. The analogues were tested as inhibitors for their ability to prevent the formation of Lys-tRNA$^{Lys}$. The velocities of the aminoacylation reaction were determined at varying L-lysine concentrations, as before. However, a constant amount of inhibitor was added to each reaction mixture. The inhibitor caused the velocities of the aminoacylation reactions to decrease and give an altered $K_M$, which was then used to calculate a $K_I$ using the equation $K_{M(app)} = K_{M(real)} (1 + [\text{inhibitor}]/K_I)$. The $K_I$ can then be used as a measure of the apparent strength of binding for the analogues (Figure 2.3). The $K_I$ value for each lysine analogue was determined and the values were tabulated. The results are shown in Table 2.1.
Figure 2.3. Structures of lysine and lysine analogues used in aminoacylation assays. Carbon is colored grey, nitrogens are blue, sulfur is yellow, and oxygens are red.
<table>
<thead>
<tr>
<th>Inhibitor</th>
<th>$K_1$ (µM)</th>
<th>$k_{\text{cat}}$ (N)</th>
</tr>
</thead>
<tbody>
<tr>
<td>L-cadaverine</td>
<td>320 ± 45</td>
<td>0.9 ± 0.03</td>
</tr>
<tr>
<td>S-(2-aminoethyl)-L-cysteine</td>
<td>1100 ± 230</td>
<td>1 ± 0.09</td>
</tr>
<tr>
<td>L-lysine ethyl ester</td>
<td>300 ± 45</td>
<td>1.1 ± 0.05</td>
</tr>
<tr>
<td>L-lysine methyl ester</td>
<td>480 ± 100</td>
<td>1 ± 0.07</td>
</tr>
<tr>
<td>DL-Hydroxylysine</td>
<td>1200 ± 140</td>
<td>1.2 ± 0.03</td>
</tr>
<tr>
<td>L-Lysinamidine</td>
<td>2100 ± 450</td>
<td>1 ± 0.07</td>
</tr>
<tr>
<td>L-Ornithine</td>
<td>8800 ± 1320</td>
<td>1.1 ± 0.02</td>
</tr>
<tr>
<td>L-Glutamate</td>
<td>37000 ± 7800</td>
<td>0.9 ± 0.06</td>
</tr>
<tr>
<td>L-Arginine</td>
<td>5100 ± 860</td>
<td>1.1 ± 0.04</td>
</tr>
<tr>
<td>L-α-aminobutyric acid</td>
<td>21000 ± 5300</td>
<td>1.1 ± 0.09</td>
</tr>
<tr>
<td>D-Lysine</td>
<td>6900 ± 2500</td>
<td>1.1 ± 0.1</td>
</tr>
<tr>
<td>L-γ-aminobutyric acid</td>
<td>8040 ± 2200</td>
<td>1 ± 0.07</td>
</tr>
<tr>
<td>Lysyl-sulfamoyl adenosine</td>
<td>0.025 ± 0.004</td>
<td>1 ± 0.05</td>
</tr>
<tr>
<td>L-Lysine hydroxamate</td>
<td>360 ± 72</td>
<td>1.1 ± 0.06</td>
</tr>
</tbody>
</table>

**Table 2.1. Inhibition of LysRS1 by lysine analogues.** Kinetic parameters for the inhibition of steady-state aminoacylation by *B. burgdorferi* LysRS1. $N=k_{\text{cat}}$ determined in the presence of inhibitor relative to $k_{\text{cat}}$ determined in absence of inhibitor.
All the compounds tested acted as competitive inhibitors of the aminoacylation reaction. This is indicated by significant changes for $K_M$, but not $k_{cat}$, when the parameters in the presence of the analogues are compared to the parameters for L-lysine alone (Table 2.1).

The compound with the lowest $K_I$, the strongest inhibitor, was lysyl-sulfamoyl adenosine (Lys-SA), with a $K_I$ of 25 nM. The compound is a non-hydrolyzable analogue of the aminoacyl-adenylate reaction intermediate. The low $K_I$ value indicates that the enzyme is very adept at binding the transition state, as predicted (Fersht et al., 1985). The analogues of L-lysine were much less potent with $K_I$ values ranging from the high µM (300 for L-lysine ethyl ester) to the low mM (21 mM for L-$\alpha$-aminobutyric acid). The least effective inhibitor was the noncognate amino acid glutamate ($K_I$ of 37 mM).

2.4. Discussion

2.4.1. Recognition of L-lysine with LysRS1

The twenty common amino acids are all very similar in structure, only differing in their side chains. Since a large part of the fidelity of the aminoacylation reaction relies on differentiating amino acids by their side chains, each synthetase has evolved a binding pocket for recognizing its cognate amino acid. With LysRS1, the amino acid binding pocket was identified in the structure of the $P.$ horikoshii enzyme. The amino acid binding site for L-lysine was found to be in a pocket that is formed between three strands of the Rossman fold (Terada et al., 2002).
The lysine active site was found to consist of six key amino acids. In this site the lysine is recognized by a combination of hydrophobic interactions and hydrogen and electrostatic bonds. The functional groups of lysine are recognized by the hydrogen and electrostatic interactions while the hydrocarbon side chain is recognized by the hydrophobic interactions.
Figure 2.4. The amino acid binding pocket of LysRS1 from *Pyrococcus horikoshii*. Numbers shown in parentheses are for the *B. burgdorferi* LysRS1. Hydrogen bonds and electrostatic interactions are indicated by dashed lines. Adapted from Wang et al., 2006.
The crystal structure revealed the functional groups of the α-carbon are recognized by four amino acids. The α-amino group forms a hydrogen bond with the main chain carboxyl of Gly27 and an electrostatic interaction with the side chain of Glu41. The oxygens of the α-carboxyl form hydrogen bonds with Thr29 and Trp218.

The side chain is recognized by both hydrophobic and electrostatic interactions. The ε-amino group is primarily recognized by an electrostatic interaction with Glu41. It might also form a hydrogen bond with Tyr268, although the distance between the Tyr hydroxyl and the ε-amino group indicates that this is at best a weak bond (Wang et al., 2006). The hydrocarbon side chain forms hydrophobic interactions with His240 and Trp218. The structure showed these residues are placed very close to the side chain and could create a closed binding pocket.

Some of the amino acids may serve dual roles in the charging reaction. Similar to other synthetases such as GluRS, LysRS1 requires bound tRNA before the aminoacyl-adenylate can be formed. LysRS1 is a Ib synthetase, as are GluRS and GlnRS (Ribas de Pouplana and Schimmel, 2001a; Terada et al., 2002). From comparisons with these two synthetases, the function of some of the amino acid residues can be hypothesized.

In GlnRS, it has been shown that Tyr211 and Phe233 form stacking interactions with the terminal adenosine of the tRNA (Rath et al., 1998; Liu et al., 1998). A conserved Tyr residue serves the same function in GluRS. Sequence alignments revealed Trp218 in LysRS1 is equivalent to the Tyr of GluRS. The
alignments also revealed the LysRS1 His240 is equivalent to Phe233 of GlnRS. A
dual role might also be served by the residues Thr29 and Tyr268. The sequence
alignments showed these two residues might be involved in interacting with the
ATP (Wang et al., 2006).

2.4.2. Analysis of lysine analogue binding with LysRS1

Analogues of L-lysine were examined to determine what aspects of the
molecule are important for binding by LysRS1. The $K_i$ values for inhibition of
steady-state aminoacylation were determined and comparisons were made to
determine the relative ability of each analogue to bind in the active site. As Lys-SA
was an analogue of the transition state, its binding will not be examined here. The
rest of the analogues tested can be divided into two groups. The first group (Group
I) is composed of analogues that differ from L-lysine in the functional groups of the
$\alpha$-carbon. The second group (Group II) is composed of amino acids that differ from
L-lysine in the R-group. GABA does not fit into either of these groups as it differs
in both the R-group and the $\alpha$-carbon, its binding will be analyzed separately. The
sorting of the two groups is shown in Table 2.2.
Table 2.2. Group I and II lysine analogues. $K_I$ values for the inhibition of steady-state aminoacylation of *B. burgdorferi* LysRS1 by lysine analogues.

<table>
<thead>
<tr>
<th>Group I analogues</th>
<th>$K_I$ (µM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>L-lysine ethyl ester</td>
<td>303 ± 45</td>
</tr>
<tr>
<td>L-cadaverine</td>
<td>320 ± 45</td>
</tr>
<tr>
<td>L-lysine hydroxamate</td>
<td>360 ± 720</td>
</tr>
<tr>
<td>L-lysine methyl ester</td>
<td>480 ± 100</td>
</tr>
<tr>
<td>L-Lysinamide</td>
<td>2100 ± 450</td>
</tr>
<tr>
<td>D-Lysine</td>
<td>6900 ± 2500</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Group II analogues</th>
<th>$K_I$ (µM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>S-(2-aminoethyl)-L-cysteine</td>
<td>1100 ± 230</td>
</tr>
<tr>
<td>DL-Hydroxylysine</td>
<td>1200 ± 140</td>
</tr>
<tr>
<td>L-Arginine</td>
<td>5100 ± 860</td>
</tr>
<tr>
<td>L-Ornithine</td>
<td>8800 ± 1320</td>
</tr>
<tr>
<td>L-α-aminobutyric acid</td>
<td>21000 ± 5300</td>
</tr>
<tr>
<td>L-Glutamate</td>
<td>37000 ± 7800</td>
</tr>
</tbody>
</table>

Generalizations can be made by comparing the two groups. From this comparison it is apparent that the enzyme tolerates alterations at the $\alpha$-carbon much better than those in the side chain. This is consistent with the role of the enzyme in its need to discriminate against other amino acids, as they are identical at the $\alpha$-carbon, but differ in the R-groups. Because of this, the mechanism for substrate recognition is focused on the side chain.

### 2.4.3. Group I lysine Analogues

Taking a closer look at the inhibition by each analogue, we can see how each one is discriminated against. The first group of analogues is discriminated against by the failure to form hydrogen bonds with the active site residues Gly29, Thr31, Trp220, and Glu43. All the analogues in group I, except D-Lysine, have alterations
at the α-carboxyl group. These alterations prevent hydrogen bonds from being formed with Thr31 and Trp220.

L-lysine ethyl ester (Lys-eth), L-cadaverine, and L-lysine hydroxamate (Lys-OH) have nearly identical $K_I$ values. Lys-eth has an ethyl group bound to one of the carboxyl oxygens, L-cadaverine is missing the α-carboxyl group, and Lys-OH has a nitrogen bound to an oxygen group in place of one of the carboxyl oxygens. These substitutions likely disrupt the hydrogen bonds the carboxylate forms with Thr 31 and Trp 220 equally since the $K_I$ values are roughly equal.

Steric clashes need to be considered with Lys-eth and Lys-OH as Lys-OH is longer than lysine by an amine while Lys-eth is longer by an ethyl group and L-cadaverine is smaller then lysine as it is missing the α-carboxyl group. Since all three analogues have the same $K_I$ it appears that size of the substrate is not important around the α-carbon. While the binding site is tight around the hydrocarbon side chain, it appears to be fairly open around the α-carboxyl group. This would allow any analogues with additions at the carboxyl to fit in the active site.

Further evidence for the influence of size comes from the results of L-lysine methyl ester (Lys-met). Lys-met is smaller then Lys-eth as it has a methyl group bound to the α-carboxyl instead of an ethyl group. Despite the smaller size, Lys-met has a higher $K_I$ than Lys-eth. This might be because the ethyl group is able to form hydrophobic interactions with the side chain of Trp. Hence, the smaller methyl does not form as strong a hydrophobic interaction and does not bind as well. This data indicates Trp220 might not be very close to the bound lysine in the active
site and plays a limited role in its recognition. This was confirmed in later experiments where the Trp was mutated to Ala. Alterations at this position were shown to have little effect on the binding of lysine (Wang et al., 2006).

The other two analogues in the first group, L-lysinamide and D-lysine, were much poorer inhibitors than the others. L-lysinamide has an amino group in place of one of the carboxyl oxygens. With a $K_i$ value of 2.1 mM the hydrogen bonds are disrupted much more then for the analogues mentioned above. This likely comes from the alteration of charge from negative to positive that comes with the substitution of the amino group for the oxygen. D-Lysine, the D enantiomer of lysine, has the amine and hydrogen around the $\alpha$-carbon switched relative to the position of the R group and the carboxyl. This disrupts the hydrogen bonds the $\alpha$-amino group forms with Gly29 and Glu43. These disruptions are severe as the $K_i$ of D-lysine is much higher than the other analogues in the group.

2.4.4. Group II lysine analogues

As stated previously, the $K_i$ values of the analogues with alterations to the side chain are generally greater than those for the analogues with alterations at the $\alpha$-carbon. In this second group, the analogues with the lowest $K_i$ values were S-(2-aminoethyl)-L-cysteine (AEC) and L-5-hydroxylysine. Their alterations from lysine consisted of bulkier side chains. AEC has its $\gamma$-carbon replaced by a sulfur atom while L-5-hydroxylysine has a hydroxyl bound to the $\delta$-carbon of the side chain. These two analogues are likely discriminated against by His242 (240 P. horikoshii) in the active site (Figure 2.5). From the crystal structure the imidazole
of the His is possibly in direct contact with the hydrocarbon side chain. With the 
two analogues steric clashes are occurring with the substitutions in the side chain. 
The steric clashes that the two analogues cause are possibly equivalent as the $K_i$ 
values are equal.
Figure 2.5 Structure of S-(2-aminoethyl)-L-cysteine modeled in the active site of *P. horikoshii* LysRS1. The L-lysine was modified to present an S instead of the γC, and the resulting structures energy minimized using Swiss-Pdb Viewer v 3.7. The resulting models were visualized in stick and van der Waals surface for active site residues and ball and stick for AEC in MOLMOL v 2k.2. Residues are colored according to their conservation in sequence alignments: gold, 100% identity; red, 81-99%; pink, 61-80%. For the AEC substrate, carbon is in white, nitrogen in blue, oxygen in red, and sulfur in yellow.
L-arginine was another inhibitor used which was larger than lysine. The longer side chain could be discriminated against by steric clashes with Glu43 and Tyr269. The positive charge being in a different position on the side chain might also lead to a weaker electrostatic interaction with Glu43.

In contrast to the previously discussed analogues, L-ornithine has a smaller side chain then lysine. The side chain is one carbon shorter, so the terminal amine is bound to the δ-carbon. This analogue most likely fits in the active site by forming the same bonds as L-lysine with the α-amino and carboxylate groups. This orientation could lead the amino group to occupy the position of the ε-carbon in binding. Examination of the crystal structure showed that this alteration would only increase distances with the functional groups of Glu and Tyr by 0.2-0.3 Å. Given the high $K_i$ of L-ornithine, this increase in distance is either enough to greatly disrupt the electrostatic and hydrogen interactions, or the amine adopts an alternate conformation that places it even farther away from the two amino acid residues. The pressure to adopt a different conformation from the carbon might come from the need to distance itself from the charge of the His side chain.

The analogue L-α-aminobutyric acid (AABA) is also smaller as it is an amino acid that only contains an ethyl group as a side chain. It has no amino group in the side chain to form electrostatic interactions or hydrogen bonds. With a shorter side chain it also must form fewer hydrophobic interactions. These differences give the compound a higher $K_i$ than L-ornithine.

The compound with the highest $K_i$ of all those tested was L-glutamate. The side chain for L-glutamate contains a carboxylate group which gives it a negative
charge. This alteration of charge likely causes L-glutamate to be electrostatically repulsed by the Glu43 of the active site. This repulsion does not completely abolish the ability of glutamate to bind. It is still able to form the bonds with the amino acids around the $\alpha$-amino and $\alpha$-carboxyl groups.

The final compound tested, GABA, is different from L-lysine at both the $\alpha$-carbon and the R-group. It is a three carbon propyl with a carboxylate at one end and an amine at the other. It has the carboxylate group which can form the hydrogen bonds with Thr31 and Trp220. The position of the amino group prevents it from forming hydrogen bonds with Gly29, although the amino group may be able to form an electrostatic interaction with Glu43. This type of bond would be necessary for it to have a lower $K_i$ than AABA (8.0 mM to 21 mM), which could form all the hydrogen bonds around the $\alpha$-carbon.

2.5. Conclusions

Overall, LysRS1 proved very adept at discriminating against the lysine analogues tested. The best inhibitors were only within the range of ten times the $K_M$ for lysine. Experiments which tested the ability of LysRS1 to use the lysine analogues as substrates showed detectable levels of charging for only four of the analogues: AEC, Lys-eth, Lys-met, and L-ornithine (Levgood et al., 2004). This indicates the poor binding of the analogues prevented many of them from being used as substrates. Besides discrimination at the binding pocket, no alternative method for analogue exclusion, such as the conversion of ornithine to ornithine lactam seen in LysRS2 (Jakubowski, 1999), was observed.
Several of the compounds tested are found within cells, and thus need to be discriminated against. As D-lysine is a by product of the synthesis of the L form, strong discrimination against the analogue is needed to ensure it is not misincorporated (Ataide and Ibba, 2006). With LysRS1, arginine, ornithine, and cadaverine are all discriminated against at levels well above concentrations found in nature (Raunio, et al., 1970; Brevet et al., 1995). For glutamate, intracellular concentrations can reach as high as 80 mM in environments of high molarity. The high $K_I$ value for glutamate is thus needed for effective discrimination (Roe, et al., 2004).
Chapter 3

tRNA anticodon binding in LysRS1

3.1. Introduction

In the living kingdoms, the evolution of tRNA has been shaped by two opposing demands (Jackman et al., 2006). All tRNA molecules must be similar enough to be used by the ribosome in protein synthesis. At the same time, they must contain enough variation for the aaRS enzymes to differentiate between them. The tRNAs have adapted to these demands by having identical secondary and tertiary structures for recognition by the enzymes responsible for protein synthesis. For the synthetases, each tRNA has specific identity elements for specific recognition by their respective synthetases.
The identity elements of tRNA\textsubscript{Lys} which LysRS1 recognizes were the first aspect of the enzyme to be studied. \textit{In vitro} kinetics showed the enzyme recognizes the anticodon nucleotides as well as the discriminator base. This correlated well with LysRS2, which recognizes the same elements. However, the anticodon was much more important for tRNA\textsubscript{Lys} recognition by LysRS2 than it was for LysRS1 (Ibba et al., 1999; Söll et al., 2000; Ambrogelly, et al., 2002).

The elements that LysRS1 uses to recognize the anticodon were not well known until the crystal structure was solved. While the crystal structure with tRNA\textsubscript{Lys} could not be solved, the unliganded structure showed the anticodon binding domain was an all α-helix cage domain which is homologous to a similar domain in GluRS (Terada et al., 2002). A crystal structure for the GluRS:tRNA\textsubscript{Glu} complex was solved and this was used as a template to construct a model for the LysRS1:tRNA\textsubscript{Lys} complex. From this model a binding pocket for the anticodon was identified. In order to test the validity of this model steady-state kinetics were used with variant enzymes and mutant tRNAs to examine the binding of tRNA\textsubscript{Lys} to LysRS1.

### 3.2. Materials and Methods

#### 3.2.1. \textit{Pyrococcus horikoshii} LysRS1 purification

Wild type \textit{P. horikoshii} LysRS1 and all variants were overexpressed in BL21 (DE3+ RPC) cells overnight at 37 °C in autoinduction media (Studier and Moffat, 1986). All subsequent steps were carried out at 4 °C unless otherwise noted. Cells
were harvested by centrifugation, washed, and resuspended in column buffer (50 mM HEPES pH 7.2, 25 mM KCl, 10 mM MgCl₂, 5 mM DTT). Cells were lysed with a French pressure cell and cell debris was removed by centrifugation at speed of 10,000 xg. The supernatant was flocculated at 65 °C and flocculent was spun down by centrifugation at 75,000 xg. The resulting solution was fractionated by FPLC using a MonoQ (5/5) column. LysRS1 was eluted using a gradient of 0-500 mM NaCl in column buffer. Fractions containing LysRS1 were pooled and concentrated through ultrafiltration (Amicon). The resulting solution was further purified through FPLC using a Superose 12 gel filtration column. A PD-10 desalting column (GE Healthcare) was used to exchange into LysRS buffer (50 mM HEPES pH 7.2, 25 mM KCl, 10 mM MgCl₂, 5 mM DTT, 10% glycerol). Fractions with LysRS1 were pooled and concentrated by ultrafiltration. The concentration of enzyme was determined by the Bradford assay (Bio-Rad) and aliquots were stored at -80 °C. LysRS1 was judged to be >95% pure by SDS-PAGE gel, judged by Coomassie blue staining.

3.2.2. tRNA_Lys transcription and purification

*P. horikoshii* tRNA_Lys was purified from the plasmid pUC18 in a manner identical to *B. burgdorferi* tRNA_Lys as described in Chapter 2. After determination of the total RNA concentration by UV spectroscopy (260 nm) the plateau of charging was found in the aminoacylation reaction to determine the percentage of active tRNA_Lys.
3.2.3. *In vitro* aminoacylation assays

Aminoacylation was performed at 37 °C in 100 mM HEPES, pH 7.2, 25 mM KCl, 10 mM MgCl₂, 4 mM DTT, 5 mM ATP, 500 nM enzyme for wild type and 1 µM for the variants, 50 µM [¹⁴C]L-Lys, and varying concentrations of tRNA¹⁰⁰Ly⁰ss. 20-µl aliquots were taken every 30 s and spotted onto Whatman 3MM filter disks presoaked in 5% trichloroacetic acid (w/v) containing 0.5% (w/v) [¹²C]-L-Lys. Sample disks were washed three times for 10 min at room temperature in 5% trichloroacetic acid (w/v) containing 0.5% (w/v) [¹²C]-L-Lys, dried at 80 °C, and the level of [¹⁴C]L-Lys-tRNA¹⁰⁰Ly⁰ss quantified by the addition of liquid scintillant (Ultima Gold, Packard Corp.) and scintillation counting. The $k_{cat}/K_M$ values presented represent the average of at least three independent experiments where values deviated by no more than 15% between individual determinations.

3.3. Results

When LysRS1 was discovered, its recognition of tRNA¹⁰⁰Ly⁰ss was the first attribute studied. In early experiments, the identity elements of tRNA¹⁰⁰Ly⁰ss were studied using tRNA¹⁰⁰Ly⁰ss and LysRS1 from the archaeon *M. maripaludis* and the bacterium *B. burgdorferi*. The tRNA recognition was initially studied by determining how class I recognition differed from class II. The anticodon and discriminator base had been shown as identity elements for LysRS2 (Ibba et al., 1999) and were tested as elements for LysRS1. The experiments showed LysRS1 recognizes the tRNA anticodon and discriminator base. However, in a slight differentiation from LysRS2, LysRS1 was found to recognize U35 and U36 of the
anticodon, but not U34. Some species of LysRS1, such as *F. acidarmanus* (Ambrogelly et al., 2002), only used U36 as a recognition element. The anticodon was found to be less significant for LysRS1 when compared to LysRS2 (Ibba et al., 1999; Söll et al., 2000; Ambrogelly, et al., 2002).

In order to explore the basis of this differentiation the binding of the anticodon by LysRS1 was examined using variant enzymes and tRNA\textsubscript{Lys} with mutant anticodons. A model of tRNA\textsubscript{Lys} bound to LysRS1 revealed the way the enzyme might recognize the anticodon. This was used in constructing variant enzymes as residues which are believed to recognize the anticodon were mutated.

### 3.3.1. The LysRS1:tRNA\textsubscript{Lys} complex

When the crystal structure for LysRS1 from *P. horikoshii* was solved, modeling was used to examine the way the enzyme might bind the tRNA substrate. The modeling took advantage of the homology between GluRS and LysRS1. The structure of the GluRS:tRNA\textsubscript{Glu} complex was used as a template to construct a model for the LysRS1:tRNA\textsubscript{Lys} complex. This model focused on recognition of both the acceptor and anticodon stems (Figs. 3.1 and 3.2).
Figure 3.1. **Model of LysRS1:tRNA\textsuperscript{Lys} complex.** The class I Rossman fold is shown in orange while the \(\alpha\)-helix cage domain is shown in red. The other domains are labeled by color: helical insertion is dark blue, CP domain is light blue, SC-fold domain is green, and the \(\alpha\)-helix bundle like domain is pink. L-lysine (blue), ATP (pink), and tRNA (yellow) are shown bound to the enzyme. Zinc ions are shown as dark purple. Some of the residues involved in tRNA recognition in the \(\alpha\)-helix cage are labeled. Adapted from Terada et al. 2002.
Figure 3.2. Structure of GluRS:tRNA$^{\text{Glu}}$ complex. Crystal structure of GluRS with bound tRNA$^{\text{Glu}}$. The Rossman fold (1) is shown in green while the anticodon binding $\alpha$-helix domains (4 and 5) are shown in red and purple. The connective-peptide domain (2) is shown in deep blue and the stem-contact domain (3) is shown in light blue. The tRNA is colored yellow with the anticodon in grey. The N-terminus (N) and C-terminus (C) are labeled. Adapted from Sekine et al., 2001.
For the acceptor stem the discriminator base of A73 was predicted to be recognized by Arg72. This residue is analogous to the Arg47 of GluRS, which binds the discriminator base of tRNA\textsuperscript{Glu}. Further examinations, done by superimposing the structure of LysRS1 over GluRS, revealed the terminal adenosine is stacked by the lysine binding site residues His240 and Trp218. This type of tRNA binding is similar to what was found with GlnRS (Rath et al., 1998; Liu et al., 1998). In GlnRS, it was shown that a Tyr and a Phe residue are used to stack the terminal adenosine. This stacking was needed to form the binding site for the amino acid and is the basis for the tRNA dependency of the amino acid activation step in the aminoacylation reaction.

3.3.2. Anticodon recognition in LysRS1

For the anticodon stem, the conformation of the α-helix cage domain of LysRS1 was altered in order to match the similar structure in the GluRS:tRNA\textsuperscript{Glu} complex. This process gave a possible structure for the α-helix cage when tRNA is bound. The modeled LysRS1:tRNA\textsuperscript{Lys} structure provided a mechanism by which the enzyme may recognize the tRNA anticodon (Fig. 3.3). The tRNA anticodon is recognized by four amino acid residues in the model. Phe487 and Tyr491 stack with U34 and U35, respectively. While these two amino acids recognize the nucleotides, it is not base specific. On the other hand, Lys497 and Arg502 do recognize the nucleotides in a base specific manner. Lys497 forms hydrogen bonds with U35, while Arg502 forms hydrogen bonds with U36.
Figure 3.3. View of the tRNA$_{\text{Lys}}$ anticodon when bound to LysRS1. Shown are the four amino acid residues believed to interact with the anticodon nucleotides. Adapted from Terada et al., 2002.
The nature of recognition from the model can give some insight into the importance of certain anticodon nucleotides. Since U34 is only recognized in a non-specific manner the exact base at this position should not be important. Any mutation at this position will likely be recognized in the same way as U34. This was shown in previous experiments mentioned above. On the other hand, U35 and U36 are recognized specifically. Thus, any alterations in the base at these positions should greatly affect the binding and aminoacylation of tRNA.

3.3.3. Sequence conservation of residues

Analyses of the sequences of LysRS1 from various archaea and bacteria support the importance of the four residues in anticodon recognition (Figure 3.4). Phe487 and Tyr491 were found to be universally conserved. Lys497 is not very well conserved (43% identical). In most cases it is replaced by Gln (43%), which has an amino group that might be able to form the same hydrogen bonds as Lys. In other cases it is replaced by Arg (4%), Glu (2%), or Ala (8%). Arg502 is not universally conserved either (69%). In all cases it is replaced by Lys, which carries an identical charge and might be able to form the same hydrogen bonds as Arg.
P. horikoshii  ISSNREWFSTLYRLFIGKERRGRPLASF
B. burgdorferi  IEPALFKQIPvKILIDKEKGPKLAF
F. acidarmanus  LTPKDFVTLYRFVPIDKDKPRLGYF
M. jannaschii  LNPREDFFQASYYRILLGGKYPKLGF
M. barkeri  VDPRLFKAIYISLLGQSSGPKAGWF
M. maripaludis  LAPKEAVASYKILLGKNYGPKLGSF

Figure 3.4. α-helix cage domain alignments. Sequence alignments of selected LysRS1 species highlighting conservation of residues in the anticodon binding domain. Residues believed to recognize the anticodon are underlined. Alignments were done using MView.

3.3.4. Experimental Overview

From the above model, experiments were designed to test the role of each amino acid in anticodon recognition. Variants of LysRS1 were made where each of the four key amino acids were changed to Ala. This was done to examine the loss of activity upon the loss of the functional group of each amino acid. Two variants were made in which residues were mutated to examine the possibility of new hydrogen bonds being formed with mutant anticodons.

In order to further examine the identity elements of tRNA_{Lys}, molecules with mutated anticodons were used. Some of the mutant anticodons used were selected to test the hypothesis that anticodon selection may have been a cause of divergence of the synthetase enzymes. Others were selected to test the similarities and differences in anticodon recognition between LysRS1 and LysRS2. A total of seven anticodons were used: UUU, CUU, GUU, UCU, UGU, UUG, and UUC.

The wild type anticodons for Lys are UUU and CUU. A correlation between the presence of LysRS1 and a dual GluRS pathway had been found previously.
Recognition of the anticodon by LysRS1 and the two GluRS enzymes might have a role in this correlation. The ability of LysRS1 and GluRS to differentiate at the 36 position could be very important. The anticodons UUC and UUG were used to determine how the enzyme discriminates at this position. A correlation was found between the presence of LysRS1 and AsnRS, indicating another distribution in which the anticodon binding of LysRS1 could be a factor (Söll et al., 2000). The GUU anticodon was examined since it is the Asn anticodon. The results with this anticodon could reveal the role of LysRS1 in AsnRS distribution. The anticodons UCU and UGU were used in these experiments since they had previously been used to investigate discrimination at the 35 position in both LysRS1 and LysRS2 (Commans et al., 1998; Ambrogelly et al., 2002).

The LysRS1:tRNA\textsubscript{Lys} structure showed ways that new interactions between enzyme and tRNA might be formed upon mutation of the anticodon nucleotides. Tyr491 was shown to come close to U34. Mutation of this residue could lead to a new hydrogen bond being formed with a mutant anticodon. The variant Tyr491Glu was made to test if a new hydrogen bond could be made with guanosine at the 34 position. The variant Arg502Gln was also made to test if a new hydrogen bond could be formed with cytidine at the 36 position. Similar variants with GluRS showed this alteration could affect anticodon recognition by the enzyme.

### 3.3.5. LysRS wild type: Differentiation of mutant anticodons

For the anticodons UUU, GUU, and CUU, both the individual $k_{cat}$ and $K_M$ could be determined. The levels of aminoacylation for these three LysRS1:tRNA
pairs are all similar. The parameters for CUU and UUU are not significantly different (Table 3.1), revealing no differentiation in the recognition of C34 when compared to U34. The GUU anticodon was slightly less efficient in aminoacylation in comparison to UUU and CUU, indicating differences in binding for G34.

As mentioned above, the nucleotide at the 34 position is believed to be recognized by a stacking interaction with Phe487. This stacking interaction is not specific for the base of the nucleotide, which explains why there are only slight differences in aminoacylation caused by the mutation at U34. However, the size of the nucleotide does appear to have some importance, as GUU had a lower efficiency than UUU and CUU. This could mean the G at the 34 position is causing steric clashes with the amino acids of the binding site.

The aminoacylation efficiency of each enzyme:tRNA pair was compared to wild type enzyme with wild type tRNA\textsuperscript{Lys}. The loss of aminoacylation efficiency (L) was determined by the formula \( \frac{[k_{\text{cat}}(wt)/K_M(\text{wt})]}{[k_{\text{cat}}(\text{mut})/K_M(\text{mut})]} \). For the CUU and GUU anticodons the L values represented only slight losses in efficiency (Table 3.1). For the tRNA with mutations at positions 35 and 36, tRNA could not be used at saturating concentrations because of the high \( K_M \) compared to practical tRNA concentrations. With \( [S] \ll K_M \), \( k_{\text{cat}}/K_M \) was directly estimated from the equation \( v = k_{\text{cat}}/K_M([E][S]) \).
Table 3.1. Kinetic data for wild type LysRS1 with mutant tRNAs. Steady-state kinetic parameters and comparison of aminoacylation efficiencies for wild type LysRS1 charging with variant tRNAs.

<table>
<thead>
<tr>
<th>tRNA^{\text{Lys}} (N_{34}N_{35}N_{36})</th>
<th>k_{\text{cat}}/K_{M} (\text{s}^{-1} \text{ M}^{-1})</th>
<th>L</th>
</tr>
</thead>
<tbody>
<tr>
<td>UUU</td>
<td>51,000 ± 10,000</td>
<td>1</td>
</tr>
<tr>
<td>CUU</td>
<td>44,000 ± 5,000</td>
<td>1.2</td>
</tr>
<tr>
<td>GUU</td>
<td>29,000 ± 300</td>
<td>1.8</td>
</tr>
<tr>
<td>UCU</td>
<td>1,900 ± 80</td>
<td>27</td>
</tr>
<tr>
<td>UGU</td>
<td>1,700 ± 100</td>
<td>30</td>
</tr>
<tr>
<td>UUC</td>
<td>3,500 ± 200</td>
<td>15</td>
</tr>
<tr>
<td>UUG</td>
<td>2,100 ± 60</td>
<td>24</td>
</tr>
</tbody>
</table>

The L values for the other mutant anticodons were much greater than for CUU and GUU. This indicates alterations at the 35 and 36 positions are not as well tolerated as alterations at the 34 position. These results were predicted as the LysRS1:tRNA^{\text{Lys}} model showed the nucleotides at these positions to be recognized in a base specific manner. While this data does not confirm which residues are being used to identify the base, it does confirm the general recognition pattern which was shown previously.

At positions 35 and 36, it was shown that the C and G mutations behaved similarly at the 35 position while they differed at the 36 position. This indicates that for the 35 position the differences in size between the purine and the pyrimidine did not matter as UCU and UGU had $k_{\text{cat}}/K_{M}$ values that were within the
error of each other. The disruption in hydrogen bonds was equal between the two bases.

For the mutant anticodons UUC and UUG significant differences were found between the two. UUG was discriminated against at a level similar to UCU and UGU while UUC had a better aminoacylation efficiency than these other mutants. The difference seen between C and G could have been caused by two effects. First, the smaller cytidine might be better accommodated by the enzyme than the guanosine. Second, disruption of hydrogen bonds might not be as severe for cytidine as for guanosine.

The better accommodation of C over G at the 36 position could reflect the evolutionary relationship between LysRS1 and GluRS. The tRNA\textsuperscript{Glu} anticodon contains a C at position 36. In the discriminating form of GluRS an Arg residue is used to recognize cytosine and exclude guanosine (Sekine et al., 2001). In non-discriminating GluRS this Arg is replaced by a Gly (Schulze et al., 2006). This alteration removes the ability of the enzyme to discriminate the base at position 36.

### 3.3.6. Phe487Ala: Loss of stacking interactions

The LysRS1:tRNA\textsuperscript{Lys} docking model revealed amino acids that might be responsible for anticodon binding. From this model several LysRS1 variants were constructed. The first variant tested was Phe487Ala. Phe487 is thought to form a stacking interaction with U34. Mutation to Ala would remove this interaction and lead to a loss in aminoacylation efficiency as the enzyme lost the binding energy provided by this residue. The data obtained are shown in Table 3.2.
Table 3.2. Kinetic data for LysRS1 variant Phe487Ala with mutant tRNAs.

Steady-state kinetic parameters and comparison of aminoacylation efficiencies for variant LysRS1 Phe487Ala charging with variant tRNAs.

<table>
<thead>
<tr>
<th>tRNA (N_{34}N_{35}N_{36})</th>
<th>( k_{\text{cat}}/K_M ) (s(^{-1}) M(^{-1}))</th>
<th>L</th>
<th>M</th>
</tr>
</thead>
<tbody>
<tr>
<td>UUU</td>
<td>1,900 ± 60</td>
<td>27</td>
<td>1</td>
</tr>
<tr>
<td>CUU</td>
<td>2,700 ± 100</td>
<td>19</td>
<td>0.7</td>
</tr>
<tr>
<td>GUU</td>
<td>2,200 ± 200</td>
<td>23</td>
<td>0.9</td>
</tr>
<tr>
<td>UCU</td>
<td>990 ± 60</td>
<td>52</td>
<td>1.9</td>
</tr>
<tr>
<td>UGU</td>
<td>870 ± 60</td>
<td>59</td>
<td>2.2</td>
</tr>
<tr>
<td>UUC</td>
<td>1,500 ± 100</td>
<td>34</td>
<td>1.3</td>
</tr>
<tr>
<td>UUG</td>
<td>1,000 ± 60</td>
<td>49</td>
<td>1.8</td>
</tr>
</tbody>
</table>

In evaluating the data for the variant enzymes the \( k_{\text{cat}}/K_M \) for each variant:tRNA pair was compared to the \( k_{\text{cat}}/K_M \) for the wild type enzyme and tRNA to give the L value. In a similar manner the \( k_{\text{cat}}/K_M \) for the variant enzyme with the wild type tRNA was normalized to wild type and each mutant tRNA was then compared to this to give a value termed M.

For the wild type tRNA a large loss in aminoacylation efficiency is seen. Phe487 plays a role in binding the anticodon and the loss of binding energy between this residue and the tRNA causes a loss in catalytic efficiency. Calculations of \( \Delta\Delta G \) showed this loss to be \(~2\) kcal/mol. The other mutations at the 34 position differed more from wild type tRNA than was seen with the wild type enzyme. The \( k_{\text{cat}}/K_M \) values for UUU and GUU fall within the range of each other,
indicating there is no significant difference between these two tRNA mutants. The $k_{\text{cat}}/K_M$ value for CUU is not within the error of UUU or GUU. The difference is not great, but could be meaningful. It is possible the loss of the large Phe perturbs the binding site so that CUU fits better than UUU or GUU.

For the tRNAs with mutations at the 35 position, the losses in aminoacylation efficiency for UCU and UGU are additive for the losses of the variant enzyme and the mutant tRNA. The loss between wild type and mutant tRNA with the variant enzyme indicates U35 is still recognized specifically despite the change in the enzyme. The $k_{\text{cat}}/K_M$ values for the two mutant tRNAs fall within the range of each other, indicating C and G are discriminated against to the same extent.

For the mutant tRNAs with alterations at the 36 position a substantial difference between UUC and UUG was once again seen. The mutant anticodon UUG was discriminated against at a level similar to UCU and UGU. This was also seen with the wild type enzyme. For the UUC mutant the $k_{\text{cat}}/K_M$ value was very close to that of wild type tRNA. Since UUG had a similar level of discrimination as in the wild type enzyme it is unlikely the position of residues in the active site was altered to a level that prevented it from discriminating against UUC. The similar values could indicate LysRS1 is not very effective at discriminating against cytidine at the 36 position.
3.3.7. Tyr491Ala: Loss of stacking interactions

In a manner similar to Phe487, Tyr491 is thought to form stacking interactions with U35. The variant enzyme Tyr491Ala would test if this residue plays a role in binding the anticodon. Data obtained for this variant are shown in Table 3.3.

<table>
<thead>
<tr>
<th>tRNA (N34N35N36)</th>
<th>$k_{cat}/K_M$ (s$^{-1}$ M$^{-1}$)</th>
<th>L</th>
<th>M</th>
</tr>
</thead>
<tbody>
<tr>
<td>UUU</td>
<td>1,600 ± 70</td>
<td>32</td>
<td>1</td>
</tr>
<tr>
<td>CUU</td>
<td>1,700 ± 90</td>
<td>30</td>
<td>0.9</td>
</tr>
<tr>
<td>GUU</td>
<td>1,500 ± 100</td>
<td>34</td>
<td>1.1</td>
</tr>
<tr>
<td>UCU</td>
<td>790 ± 90</td>
<td>65</td>
<td>2.0</td>
</tr>
<tr>
<td>UGU</td>
<td>760 ± 30</td>
<td>67</td>
<td>2.1</td>
</tr>
<tr>
<td>UUC</td>
<td>1,000 ± 70</td>
<td>51</td>
<td>1.6</td>
</tr>
<tr>
<td>UUG</td>
<td>900 ± 5</td>
<td>57</td>
<td>1.8</td>
</tr>
</tbody>
</table>

Table 3.3. Kinetic data for LysRS1 variant Tyr491Ala with mutant tRNAs.
Steady-state kinetic parameters and comparison of aminoacylation efficiencies for variant LysRS1 Tyr491Ala charging with variant tRNAs.

A large loss in efficiency is found for the wild type tRNA with the Tyr491Ala variant, as with Phe487Ala. This results from the loss of binding energy formed from the stacking interaction of Tyr with the anticodon. Calculations of $\Delta\Delta G$ showed this loss of energy to be similar to the loss found for Phe487Ala, around 2.3 kcal/mol. The data support the hypothesis the residue is binding U35. The $k_{cat}/K_M$
values for UUU, GUU, and CUU all fall within the errors of each other. This indicates no differentiation between the bases is taking place at the 34 position.

For the mutant tRNAs with alterations to U35 the results were similar to those obtained previously. The anticodons UCU and UGU had a substantially lower catalytic efficiency than the wild type tRNA. The $k_{cat}/K_M$ values of the two were within the range of each other. This indicates that once again the two mutants are being discriminated against in relation to uridine, but they are discriminated against in a similar manner.

For the tRNAs with alterations to U36, slightly different results were obtained. In contrast to Phe487Ala but in agreement with wild type, the anticodon UUC showed a substantial amount of differentiation from the wild type tRNA. The anticodon UUG was also discriminated against substantially in comparison to wild type tRNA, although the loss of efficiency was not as severe as it was with the UCU and UGU mutants. Instead, the $k_{cat}/K_M$ value was much closer to UUC, which had not been seen before. The results are in reasonable agreement with those of wild type, with substantial differences between wild type and UUC and UUG.

3.3.8. Tyr491Glu: Formation of new hydrogen bonds

While the model of LysRS1 with tRNA$^{Lys}$ showed Tyr491 to be primarily involved in stacking with U35, it showed that it might also come close to U34. As a way to test this placement the Tyr was mutated to Glu to examine if this altered enzyme could form new hydrogen bonds with a guanosine at the 34 position. The data obtained for this enzyme is shown in Table 3.4.
Most of the data obtained for the Tyr491Glu variant is nearly identical to the data for Tyr491Ala. The replacement of Tyr with Glu disrupted the stacking interactions in a manner similar to the Ala variant. A large loss in aminoacylation efficiency was seen with the wild type tRNA as the stacking interactions were lost. This free energy loss was 2.4 kcal/mol. The UUU tRNA had nearly identical $k_{cat}/K_M$ values to GUU and CUU, indicating no differentiation of the bases at the 34 position. The similar results obtained for GUU and wild type disproved the hypothesis that the mutated Glu residue would be in position to form a new hydrogen bond with the GUU anticodon.

<table>
<thead>
<tr>
<th>tRNA ($N_{34}N_{35}N_{36}$)</th>
<th>$k_{cat}/K_M$ (s$^{-1}$ M$^{-1}$)</th>
<th>L</th>
<th>M</th>
</tr>
</thead>
<tbody>
<tr>
<td>UUU</td>
<td>1,500 ± 70</td>
<td>34</td>
<td>1</td>
</tr>
<tr>
<td>CUU</td>
<td>1,400 ± 80</td>
<td>36</td>
<td>1.1</td>
</tr>
<tr>
<td>GUU</td>
<td>1,500 ± 100</td>
<td>34</td>
<td>1</td>
</tr>
<tr>
<td>UCU</td>
<td>940 ± 60</td>
<td>54</td>
<td>1.6</td>
</tr>
<tr>
<td>UGU</td>
<td>870 ± 70</td>
<td>59</td>
<td>1.7</td>
</tr>
<tr>
<td>UUC</td>
<td>1,300 ± 100</td>
<td>39</td>
<td>1.2</td>
</tr>
<tr>
<td>UUG</td>
<td>6,400 ± 400</td>
<td>8</td>
<td>0.2</td>
</tr>
</tbody>
</table>

Table 3.4. Kinetic data obtained for LysRS1 variant Tyr491Glu with mutant tRNAs. Steady-state kinetic parameters and comparison of aminoacylation efficiencies for variant LysRS1 Tyr491Glu charging with variant tRNAs.
The tRNA mutants UCU and UGU once again were less efficient substrates than wild type while their $k_{cat}/K_M$ values were within the error of each other. Specific discrimination was once again taking place at the 35 position.

The most interesting results with the variant enzyme were obtained with UUG. The $k_{cat}/K_M$ value for this anticodon was four times greater than that obtained for wild type tRNA. This indicated new hydrogen bonds were formed with the guanosine in this position and that the Tyr491 residue, instead of fitting between U34 and U35 so it is close to U34, must stack with U35 on the opposite side so it is close to U36. This also shows the variant enzyme could rescue tRNA mutations by recognizing the G at the position.

Despite forming new bonds with UUG, the anticodon UUC was still discriminated against. As with the other variants it did not have as low a level of efficiency as UCU or UGU. Its level of efficiency was close to wild type tRNA, indicating a lack of ability of the enzyme to differentiate cytidine at the 36 position as seen with the Phe487Ala variant.

3.3.9. Lys497Ala: Loss of hydrogen bonds

In contrast to the residues Phe487 and Tyr491, which recognize bases nonspecifically, Lys497 and Arg502 are hypothesized to recognize nucleotides in a base specific manner. This was determined from the LysRS1:tRNA\textsuperscript{Lys} model. Lys497 is thought to form hydrogen bonds with U35 while Arg502 is thought to form hydrogen bonds with U36.
The variant enzyme Lys497Ala was constructed to examine if the loss of the amino acid residue would affect the aminoacylation efficiency. When the enzyme was purified no detectible activity was observed. This may reflect the hydrogen bonds Lys497 forms with the anticodon are essential for the activity of the enzyme. However, the nucleotide to which the residue is binding cannot be confirmed.

3.3.10. Arg502Ala: Loss of hydrogen bonds

From the model the residue Arg502 is thought to form hydrogen bonds with U36, recognizing the nucleotide in a base specific manner. The mutation of this residue to Ala should remove the ability of the enzyme to recognize U36 specifically. The results obtained for this variant are shown in Table 3.5.

<table>
<thead>
<tr>
<th>tRNA (N_{34}N_{35}N_{36})</th>
<th>( k_{cat}/K_M ) (s(^{-1}) M(^{-1}))</th>
<th>L</th>
<th>M</th>
</tr>
</thead>
<tbody>
<tr>
<td>UUU</td>
<td>2,900 ± 200</td>
<td>18</td>
<td>1</td>
</tr>
<tr>
<td>GUU</td>
<td>3,100 ± 100</td>
<td>16</td>
<td>0.9</td>
</tr>
<tr>
<td>CUU</td>
<td>3,200 ± 100</td>
<td>16</td>
<td>0.9</td>
</tr>
<tr>
<td>UCU</td>
<td>1,300 ± 70</td>
<td>39</td>
<td>2.2</td>
</tr>
<tr>
<td>UGU</td>
<td>790 ± 30</td>
<td>65</td>
<td>3.7</td>
</tr>
<tr>
<td>UUC</td>
<td>700 ± 60</td>
<td>73</td>
<td>4.1</td>
</tr>
<tr>
<td>UUG</td>
<td>850 ± 50</td>
<td>60</td>
<td>3.4</td>
</tr>
</tbody>
</table>

Table 3.5. Kinetic data for LysRS1 variant Arg502Ala with mutant tRNAs. Steady-state kinetic parameters and comparison of aminoacylation efficiencies for variant LysRS1 Arg502Ala charging with variant tRNAs.
As seen with the other variants, the wild type tRNA showed a substantial
decrease in aminoacylation efficiency, indicating an involvement of this residue in
tRNA binding. However, the loss seen for wild type tRNA was not as large as that
seen for the other variants that removed stacking interactions. This could indicate
the bonds the Arg are forming are not as important for overall binding as the
stacking interactions. The anticodons mutated at the 34 position had $k_{cat}/K_M$ values
within the error range of UUU, indicating no specific recognition taking place at
this position in this variant.

The mutant anticodons UGU, UUG and UUC all had $k_{cat}/K_M$ values that were
within the error range of each other. There was also a larger difference between the
aminoacylation efficiencies of the mutant tRNAs and the wild type tRNA with the
variant. Also, in contrast to the previous variants, UCU and UGU did not have
identical losses of efficiency. The enzyme was almost twice as effective at
recognizing UCU as UGU. The Arg502Ala variant was able to differentiate
between each of the three bases at the 35 position. This could give some indication
Arg502 is involved in recognition of this base.

The large differences found between charging of tRNA with wild type
anticodon and the mutant anticodons UUC and UUG disproved the hypothesis for
the role of Arg 502. If the residue was responsible for the specific recognition of
U36, the loss of this residue should have removed the ability of the enzyme to
differentiate between anticodons varying at this position. This would mean that
UUC and UUG should have aminoacylation efficiencies similar to wild type. In
contrast to the hypothesis, the values were much higher than the wild type with the
differences between the mutant and wild type being in most cases twice as great as what was seen for the other variants.

### 3.3.11. Arg502Gln: Formation of new hydrogen bonds?

Arg502 was mutated to Gln to examine if new hydrogen bonds could be formed by this mutated residue. The original hypothesis was that Gln would be able to form a new bond with the anticodon UUC. The data for this variant are shown in Table 3.6.

<table>
<thead>
<tr>
<th>tRNA (N34N35N36)</th>
<th>$k_{cat}/K_M$ (s^{-1} M^{-1})</th>
<th>L</th>
<th>M</th>
</tr>
</thead>
<tbody>
<tr>
<td>UUU</td>
<td>570 ± 30</td>
<td>89</td>
<td>1</td>
</tr>
<tr>
<td>GUU</td>
<td>530 ± 6</td>
<td>96</td>
<td>1.1</td>
</tr>
<tr>
<td>CUU</td>
<td>550 ± 30</td>
<td>93</td>
<td>1.0</td>
</tr>
<tr>
<td>UCU</td>
<td>180 ± 10</td>
<td>280</td>
<td>3.2</td>
</tr>
<tr>
<td>UGU</td>
<td>180 ± 0.7</td>
<td>280</td>
<td>3.2</td>
</tr>
<tr>
<td>UUC</td>
<td>73 ± 7</td>
<td>700</td>
<td>7.8</td>
</tr>
<tr>
<td>UUG</td>
<td>78 ± 4</td>
<td>650</td>
<td>7.3</td>
</tr>
</tbody>
</table>

Table 3.6. Kinetic data for LysRS1 variant Arg502Gln with mutant tRNAs. Steady-state kinetic parameters and comparison of aminoacylation efficiencies for variant LysRS1 Arg502Gln charging with variant tRNAs.

For the wild type tRNA with the variant enzyme a very large decrease in aminoacylation efficiency was found. This efficiency was less than any
variant:tRNA pair tested previously. The loss in free energy caused by this variant
with wild type tRNA was 3.0 kcal/mol. The wild type tRNA also had a $k_{cat}/K_M$
value that was within the error range of GUU and CUU. This shows that for the
variant there is still no specific recognition of the base at the 34 position.

For the anticodons UCU and UGU large losses in efficiency between the two
mutants and the wild type were found. This loss was similar to the loss between
UUU and UGU for the Arg502Ala variant. However, unlike the previous variant,
UCU and UGU had nearly identical $k_{cat}/K_M$ values. This indicated this variant was
unable to differentiate between cytidine and guanosine at the 35 position.

The original hypothesis for the creation of the variant was the belief that Gln
would be able to form a new hydrogen bond with cytidine at the 36 position. No
new bond was formed with this nucleotide. A large loss in aminoacylation
efficiency was seen with both UUC and UUG. The $k_{cat}/K_M$ values were within the
error range of each other, so their differences between each other were not
substantial. The loss in efficiency seen between these two mutants and the other
tRNAs was far greater than any differences observed before. The mutation of Arg
to Gln resulted in an enzyme much more sensitive to mutations at the 36 position.

The overall picture of anticodon binding found for *P. horikoshii* LysRS1 is
similar to the results seen with other LysRS1 enzymes. The enzyme recognized
U35 and U36 specifically, but not U34. Some of the data also suggest that at the 36
position the enzyme is more accommodating of cytidine than guanosine mutants.
The anticodon did not appear to be as crucial for recognition as in other synthetases,
as the decreases in aminoacylation efficiency were only about one order of
magnitude for the mutant tRNAs. Other closely related synthetases that use the anticodon for recognition, such as GlnRS, show decreases two orders of magnitude or greater with tRNAs mutated at the anticodon (Uter and Perona, 2004).

3.4. Discussion

Differentiating the similarly structured tRNAs is one of the primary functions of the synthetase enzymes. Early studies on tRNA recognition by the synthetases showed that elements in both the acceptor and anticodon stems were important (Saks et al., 1994). The anticodon nucleotides were generally more important than any elements in the acceptor stem, but this was not universal. For the *E. coli* tRNAs, 17 of the 20 isoacceptors have the anticodon as a vital recognition element. Numerous studies have confirmed that most identity elements are found in the extremities of the tRNA molecule; the acceptor and anticodon stems (Geigé et al., 1998). Early studies of tRNA\textsuperscript{Lys} recognition with LysRS1 revealed the enzyme follows this general pattern (Ibba et al., 1999; Söll et al., 2000; Ambrogelly, et al., 2002).

3.4.1. Elements of tRNA recognized by LysRS1

The elements of tRNA the class I LysRS recognizes were previously studied by enzymatic footprinting and steady-state kinetics (Ambrogelly et al., 2005). Footprinting showed the anticodon itself was protected while the double stranded region of the stem was not. Areas of the D arm and T arm also showed protection.
Kinetic experiments confirmed the anticodon as an important area of recognition, but showed no vitally important elements in the rest of the anticodon stem or the D and T arms. For the acceptor stem, the discriminator base was a minor determinant along with a G2:U71 wobble pair in the double stranded region. The wobble base pair may cause the acceptor stem to adopt a structure which allows the enzyme to bind the minor groove side and direct the CCA end to the active site.

The results of this study correlated well with the results of the earlier studies on tRNA\(^{\text{Lys}}\) identity by LysRS1. These studies had found that LysRS1 recognizes the anticodon and discriminator base (Ibba et al., 1999; Söll et al., 2000; Ambrogelly, et al., 2002). This was similar to the identity elements for LysRS2, which also uses the anticodon and discriminator base. However, the anticodon was much more important for recognition by LysRS2 than LysRS1. Another difference between the two enzymes concerns the G2:U71 wobble pair. Organisms with LysRS1 contained a G2:U71 wobble pair that is absent in the tRNA for organisms with LysRS2. While the presence of this base pair was not needed for recognition by LysRS1, it does serve as an anti-determinant for LysRS2 (Ibba et al., 1999). The anticodon binding in LysRS1 was examined to study the differences between LysRS1 and LysRS2 in anticodon recognition.

3.4.2. Anticodon recognition as a cause of differentiation

Differences in anticodon recognition could also contribute to the phylogenetic distribution of the two classes (Söll et al., 2000). This could have been caused by the similarities of the lysine anticodon to the anticodons for Asn, Asp, Glu, and Gln
and the need for the synthetases to differentiate between the tRNAs with the similar anticodons. As mentioned previously, alternate routes exist for the formation of Gln-tRNA$^{\text{Gln}}$ and Asn-tRNA$^{\text{Asn}}$. These alternate routes present differing requirements for the recognition of tRNA anticodons by the synthetases.

The (U/C)UU anticodon of tRNA$^{\text{Lys}}$ is in the same codon family as the GUC anticodon of tRNA$^{\text{Asp}}$ and the GUU anticodon of tRNA$^{\text{Asn}}$. In organisms where AsnRS is present AspRS, LysRS, and AsnRS differentiate between the various anticodons by recognizing all three nucleotides (Söll et al., 2000). AsnRS is not present in all organisms; when it is absent a non-discriminating AspRS is needed to charge tRNA$^{\text{Asn}}$ with Asp. In these cases, the non-discriminating AspRS only recognizes G34 and U35 (Chuawong and Hendrickson, 2006). This presents a situation where LysRS may only need to differentiate at the 36 position. If this was the case then the anticodon recognition properties of LysRS would correlate with the presence of AsnRS. This hypothesis was substantiated by phylogenetic analyses which showed LysRS1 (which only recognizes U35 and U36, or just U36) correlated strongly with the distribution of AsnRS (Söll et al., 2000).

Discrimination against the anticodons of tRNA$^{\text{Glu}}$ and tRNA$^{\text{Gln}}$ presents a similar situation. The formation of Gln-tRNA$^{\text{Gln}}$ can also occur by a transamidation pathway similar to the one used for Asn-tRNA$^{\text{Asn}}$. A non-discriminating GluRS, which can recognize both Glu and Gln anticodons, can aminoacylate tRNA$^{\text{Gln}}$ with glutamate. The amidotransferase complex can then convert Glu to Gln (Ruan et al., 2001).
3.4.3. The dual GluRS pathway

In divergence from the Asn pathway, GluRS enzymes that only recognize tRNA$^{\text{Gln}}$ were found in several strains of proteobacteria (Skouloubris et al., 2003). These organisms contain two forms of GluRS. The first form is a discriminating GluRS (GluRS1) that only aminoacylates tRNA$^{\text{Glu}}$ with Glu. The second form is a GluRS (GluRS2) that only aminoacylates tRNA$^{\text{Gln}}$ with Glu. This pathway exists primarily in the proteobacteria, with a few exceptions in other organisms such as two members of the Clostridiales (Salazaar et al., 2003).

Experiments to determine the identity elements of tRNA for GluRS demonstrated that in discriminating GluRS the anticodon is important for recognition along with several elements in the D arm (Sekine et al., 1996; Sekine et al., 1999). Mutagenesis with the residue responsible for recognition at the 36 position in a discriminating GluRS revealed tRNA recognition could be altered with this single amino acid change (Sekine et al., 2001). Similar experiments with GluRS2 gave analogous results where the change of a few amino acids showed the ability to alter the specificity of the enzyme with respect to the anticodon (Lee and Hendrickson, 2004).

Phylogenetic analyses of organisms which contain the two forms of GluRS revealed all the $\alpha$-proteobacteria which contain the two GluRS enzymes also contain LysRS1 (Salazaar et al., 2003). In agreement with earlier correlations drawn between the distribution of LysRS1 and AsnRS, there may be functional significance to this distribution. Because of the importance of anticodon recognition for GluRS and GlnRS, and the homology between GluRS and LysRS1,
studies were undertaken to investigate anticodon recognition in LysRS1 to determine the mechanism which it uses to recognize the UUU anticodon.

3.4.4. LysRS1 recognition of the anticodon

This research was conducted to gain some insight into the mechanisms LysRS1 uses to recognize the anticodon. Each variant enzyme showed losses in aminoacylation efficiency when compared to wild type, indicating each residue which was mutated played some role in recognizing elements of the anticodon loop. Examination of the kinetic data in relation to the model revealed some of the ways the enzyme might bind the anticodon loop.

Phe487 was believed to form stacking interactions with U34. Sequence analysis showed the residue to be universally conserved. The variant which replaced this residue with Ala showed substantial losses in aminoacylation efficiency for all tRNAs tested, indicating involvement in anticodon recognition. While the data cannot confirm this residue binds U34, it does lend support to the model.

Similar to Phe487, Tyr491 also was believed to form stacking interactions, with U35. The mutation of this residue to Ala showed losses in aminoacylation efficiency similar to Phe487Ala. While the Ala mutation could not confirm the exact nucleotide it is stacking with, the Tyr491Glu variant did give such insight. The original model showed that while the Tyr would stack with U35 it would also come close to U34. A variant was created where the Tyr was mutated to Glu with the hypothesis that Glu would be able to form hydrogen bonds with a mutant
guanosine at the 34 position of the anticodon. The results for this variant showed the residue could form hydrogen bonds with a guanosine, but at the 36 position instead of the 34 position. This indicated the Tyr fits into the anticodon on the opposite side of U35 from what was shown in the original model.

The unliganded structure of LysRS1 showed the aromatic residues Phe487, Tyr491, and Phe506 are all adjacent to each other in a line. The computer model in combination with the kinetic data indicated Tyr491 could fit between U35 and U36, forming its stacking interactions with U35, while it would be close enough to U36 for the variant Tyr491Glu to form interactions with a G36 mutant. This placement would then indicate Phe487 is forming stacking interactions with U34 and present the possibility of Phe506 forming stacking interactions with U36 on the opposite side of the base from Tyr491. In contrast to Phe487 and Tyr491, Phe506 is not universally conserved. It is a phenylalanine in 78% of the LysRS1 sequences and in all other cases it is Leu.

Finding the mechanism by which LysRS1 specifically recognizes the anticodon bases U35 and U36 was the primary purpose of this study. The LysRS1:tRNA^{Lys} model showed U35 was recognized specifically by hydrogen bonds with Lys497 while U36 was recognized by hydrogen bonds with Arg502. The variant enzyme Lys497Ala proved to be inactive. This would support the model that the residue plays a crucial role in the binding of the anticodon, possibly engaging in the specific recognition of U35.

The model predicted that Arg502 forms specific hydrogen bonds with U36. The kinetic results for the variant enzymes Arg502Ala and Arg502Gln proved this
hypothesize wrong. The model showed the Arg to form hydrogen bonds and specifically bind U36. The mutation of the residue to Ala should have removed the ability of the enzyme to differentiate between nucleotides at the 36 position. What was revealed instead was that the enzyme still specifically recognizes uridine at position 36. The variant Arg502Gln also showed no new hydrogen bonds with the mutant tRNAs, indicating this residue does not occupy a position near the anticodon nucleotides. This indicates the residue is not involved in binding U36.

The variants with alterations of Arg502 still showed losses of efficiency, implicating this residue in binding the anticodon loop. One possibility is that the positively charged Arg is binding the negatively charged phosphate backbone of the anticodon loop. Similar types of binding with positively charged amino acid residues were seen with LysRS2 (Commans et al., 1998). Since the Arg is not specifically recognizing U36, the residue which is the cause of the enzyme’s specific recognition of this nucleotide needs to be identified. Another evaluation of the model showed the residues Glu498 and Arg499 may come close enough to the bound anticodon to form hydrogen bonds with the nucleotides.

The research presented above continued the work on the identification of elements of tRNA interacting with LysRS1 and presented the first biochemical insight into the manner in which the enzyme recognizes the anticodon. It showed LysRS1 uses a network of stacking interactions and hydrogen bonds with the bases to recognize the anticodon. It does not involve as many residues in the binding of the anticodon as LysRS2, however, indicative of the different evolutionary origins of the two forms of LysRS.
Chapter 4

Rationale for two classes of LysRS

4.1. Introduction

The twenty canonical synthetases exist in two classes of ten members each labeled simply class I and class II. They are not believed to share a common origin and differ in both sequence and structure. The structural and sequential differences are reflected in the divergence in the mechanism of catalysis for the two classes. While the two classes use the same basic two step mechanism, they differ in the way they bind their substrates. Class I enzymes bind ATP in an extended conformation and recognize the minor groove side of tRNA on the acceptor stem. Class II enzymes bind ATP in a bent conformation and recognize the major groove of the tRNA.
LysRS is the only synthetase which has both a class I and class II form. Genomic analysis revealed LysRS1 is found in most archaea, a few bacteria, and no eukaryotes. LysRS2 is found in all eukaryotes, most bacteria, and a few archaea. The two enzymes are only found together in a small number of bacterial and archaeal species, such as archaeal methanogens and members of the Bacillus class of bacteria.

Initial comparison of the two LysRS species showed them to be functionally equivalent even though they were not structurally related. The two enzymes were able to recognize the same amino acid and tRNA substrates both in vitro and in vivo. In order to see if the two LysRS forms are truly identical in substrate recognition, the binding of L-lysine and tRNA$_{\text{Lys}}$ to LysRS1 was examined. To analyze the lysine binding pocket of LysRS1, lysine analogues were used to test inhibition of the aminoacylation reaction. To examine the binding of tRNA$_{\text{Lys}}$ at the anticodon, the reaction of variant enzymes with mutant tRNAs was analyzed. Differences in the way the two forms recognize this substrate could be cause for the distribution of the two enzymes in nature.

4.2. Materials and Methods

4.2.1. Strain Construction

To construct B. subtilis strain BCJ157.1 that places lysK on the chromosome under PrpsD expression control, the B. subtilis PrpsD promoter and terminator regions were amplified by using primer pairs RpsF/1–RpsR/1 and RpsF/2-
RpsR/2.2, respectively, ligated, and the fragment was cloned into pDIA5304, yielding plasmid pBCJ164.3. The lysK gene was excised from pBCJ103.7 as an NdeI fragment and inserted into the NdeI site of pBCJ164.3, giving plasmid pBCJ122.3. This PrpsD-lysK expression cassette was then excised on a BamHI fragment and cloned into BamHI-digested pD268, giving plasmid pBCJ137.1, which was then linearized with Scal and integrated into the B. subtilis chromosome by a double cross-over at the amyE locus, giving strain BCJ141.1. The endogenous lysS gene of BCJ141.1 was excised by gene replacement using linearized pBCJ144.3 to give strain BCJ157.1.

4.2.2. In vivo growth inhibition

B. subtilis strains 168 and 157.1 were grown aerobically in LB medium until \( A_{600}=1 \). 1 ml of this culture was then spun down and washed and resuspended in 1 ml of Spizizen’s minimal medium. 250 µL of these cells were then inoculated in 25 ml of Spizizen’s minimal medium supplemented with 2 mM L-lysine or 5 µM of AEC or 400 mM GABA at 37 °C.

4.3. Results

When it was discovered that LysRS has a form in each synthetase class, the distribution of the two forms of the enzyme was examined. In this distribution, LysRS1 is in most archaea and a few bacteria while LyRS2 is in all eucarya, most bacteria, and a few archaea. In order to explain this distribution, the substrate
binding of LysRS1 was examined. It was hypothesized differences in substrate binding could give organisms harboring one form of LysRS a selective advantage over organisms harboring the other form. The binding of both lysine and tRNA was examined *in vivo* and *in vitro*.

### 4.3.1. Divergence in lysine analogue recognition

One of the first differences noted between LysRS1 and LyRS2 was in the method of lysine activation in the first step of the aminoacylation reaction. LysRS1 must bind tRNA in order to activate lysine while LysRS2 can activate lysine without tRNA bound. In other synthetases that require tRNA binding to activate the amino acid it was shown the tRNA induces conformational changes in the amino acid binding pocket. Because of this, the binding pockets of LysRS1 and LysRS2 were further examined by comparing the effects of lysine analogues on the aminoacylation reaction.

The binding of lysine analogues to LysRS1 was described previously (Chapter 1). The data for LysRS1 was compared to similar data for LysRS2 and are shown in Table 4.1. Most of the analogues tested had higher $K_1$ values for LysRS1 when compared to LysRS2. While most of the differences were small, three analogues showed significant differences in binding between LysRS1 and LysRS2. AEC and L-lysinamide had much lower $K_1$s for LysRS2 then LysRS1. On the other hand, GABA had a much lower $K_1$ for LysRS1 then LysRS2.
<table>
<thead>
<tr>
<th>Inhibitor</th>
<th>LysRS1-$K_1$ (µM)</th>
<th>LysRS2-$K_1$ (µM)$^1$</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>L-cadaverine</td>
<td>320 ± 45</td>
<td>260 ± 28</td>
<td>4</td>
</tr>
<tr>
<td>AEC</td>
<td>1140 ± 230</td>
<td>3.9 ± 0.4</td>
<td>290</td>
</tr>
<tr>
<td>L-lysine ethyl ester</td>
<td>303 ± 45</td>
<td>55 ± 6</td>
<td>180</td>
</tr>
<tr>
<td>L-lysine methyl ester</td>
<td>478 ± 100</td>
<td>74 ± 7</td>
<td>6</td>
</tr>
<tr>
<td>DL-Hydroxyllysine</td>
<td>1200 ± 140</td>
<td>500 ± 52</td>
<td>6</td>
</tr>
<tr>
<td>L-Lysinamide</td>
<td>2120 ± 450</td>
<td>17 ± 2</td>
<td>2</td>
</tr>
<tr>
<td>L-Ornithine</td>
<td>8800 ± 1320</td>
<td>6300 ± 600</td>
<td>1</td>
</tr>
<tr>
<td>L-Glutamate</td>
<td>37000 ± 7800</td>
<td>130,000 ± 13,000</td>
<td>1</td>
</tr>
<tr>
<td>L-Arginine</td>
<td>5060 ± 860</td>
<td>64,000 ± 5,000</td>
<td>1</td>
</tr>
<tr>
<td>L-α-aminobutyric acid</td>
<td>21200 ± 5300</td>
<td>14,200 ± 1700</td>
<td>1</td>
</tr>
<tr>
<td>D-Lysine</td>
<td>6900 ± 2500</td>
<td>12,000 ± 1400</td>
<td>1</td>
</tr>
<tr>
<td>L-γ-aminobutyric acid</td>
<td>8040 ± 2200</td>
<td>470,000 ± 51,000</td>
<td>0.02</td>
</tr>
<tr>
<td>Lysyl-sulfamoyl adenosine</td>
<td>0.025 ± 0.004</td>
<td>0.028 ± 0.003</td>
<td>0.08</td>
</tr>
<tr>
<td>L-Lysine hydroxamate</td>
<td>360 ± 72</td>
<td>86 ± 7</td>
<td>0.3</td>
</tr>
</tbody>
</table>

Table 4.1. Kinetic parameters for lysine analogues. Steady-state kinetic data for aminoacylation of tRNA$^{1\text{Lys}}$ by LysRS1 (B. burgdorferi) and LysRS2 (E. coli) in the presence of lysine analogues. $^1$(Data for LysRS2 determined by Sandro Ataide) R=The ratio for an analogue’s $K_1$ value with LysRS1 in comparison to its $K_1$ value with LysRS2 [($K_1$-LysRS1)/($K_1$-LysRS2)].
4.3.2. \textit{In vivo} analogue inhibition

In order to investigate if the \textit{in vitro} differences observed for lysine analogue recognition could cause differences in \textit{in vivo} growth, experiments were performed with several \textit{B. subtilis} strains (Jester et al., 2003). Strain 168 encoded endogenous \textit{B. subtilis} LysRS2, strain 143.3 had the endogenous LysRS2 replaced by LysRS2 from \textit{S. aureus} and strain 157.1 had the endogenous LysRS2 replaced by \textit{B. burgdorferi} LysRS1. These strains were first tested in disk assays where their growth on a plate was tested in the presence of a disk of AEC (Figure 4.1).

![Figure 4.1](image)

\textbf{Figure 4.1. Analysis of growth inhibition by AEC using disk assays.} The LysRS form encoded in each strain is indicated.

The two strains which encode LysRS2, 168 and 143.3, both showed substantial growth inhibition around the disk of AEC (Thialysine). On the other hand, the LysRS1 encoding strain, 157.1, showed minimal growth inhibition around the disk of AEC. While not shown in the Figure 4.1, a strain encoding both classes of LysRS showed growth inhibition in a manner similar to 168 and 143.3.
Similar *in vivo* experiments using liquid culture were done to test possible growth differences in the presence of AEC and GABA. Lysinamide was not used, since concentrations up to 46 mM showed no growth inhibition in either strain 168 or 157.1, indicating the analogue might not have been taken up by the cell or is rapidly destroyed upon passing through the membrane. The analogues AEC and GABA showed different rates of growth inhibition for the two strains (Figure 4.2).

**Figure 4.2. Growth curves of *B. subtilis* strains.** Strains were grown in Spizizen’s minimal medium in the presence of 2 mM lysine, 10 µM AEC, or 400 mM GABA.

In agreement with the disk assays, the growth data showed the strain with LysRS1 was able to grow in an AEC concentration of 10 µM while the strain containing LysRS2 was unable to grow. In contrast, LysRS2 was able to grow in a GABA concentration of 400 mM while LysRS1 was unable to tolerate this GABA concentration.
4.4. Discussion

The synthetases have been found to exist in two unrelated classes, class I and II. They do not share any common sequences or structures, but carry out their reactions in similar manners. They each use the same two step reaction mechanism for aa-tRNA formation, but they bind their substrates in different manners. The foundation of the two classes probably has its origin in the earliest development of coded macromolecule biosynthesis.

Catalysis by macromolecules likely began with the RNA world (Joyce, 2002). Ribozymes catalyzed reactions such as RNA genome replication and RNA synthesis. Eventually, a transition was made to a world of protein catalysis. In this primordial world, protein synthesis would have been catalyzed by ribozymes that used early forms of tRNA as adaptor molecules (Schimmel and Kelley, 1995). The earliest tRNAs were likely minihelices composed of the acceptor helix and TΨC stem. These would have been charged by ribozymes acting as aaRSs with the charged minihelices being used in protein synthesis (Schimmel and de Pouplana, 1995).

The primordial synthetases would have been composed of the basic class domain for binding and aminoacylating the early tRNA minihelix. At some point the early tRNAs added the anticodon stem and the synthetases added domains for recognizing their respective anticodon stems (Figure 4.3) (Schimmel and Ribas de Pouplana, 1995; Ibba et al. 1997c). Because the anticodon domains were added later there is a substantial amount of homology in these domains among all the synthetases (Wolf et al., 1999).
Figure 4.3. Diagram of synthetase evolution. The enzymes were originally the active site tRNA binding domains that are conserved throughout their class. Later, nonconserved domains were joined with the conserved domains to give the full synthetase enzymes. Adapted from Ibba et al., 1997c.
4.4.1. The divergence of the two classes of LysRS

It is thought both classes of LysRS were present at the time of the last common species. As the living kingdoms evolved, divergence occurred that resulted in LysRS2 remaining in all eukarya and most bacteria. For the archaea and a few remaining bacteria, LysRS1 was kept instead of LysRS2 (Saul et al., 2006). Reasons for this distribution of the two enzymes have been hypothesized from the analysis of the reaction mechanisms of LysRS1 and their comparison to LysRS2.

4.4.2. Lysine analogue recognition as a cause of divergence and distribution

Comparison of the crystal structures of LysRS1 and LysRS2 presented the possibility that the lysine binding pocket of LysRS1 is much more compact than the pocket of LysRS2 (Terada et al., 2002). Structural modeling revealed the more compact nature of the LysRS1 binding pocket is responsible for the differences in analogue binding seen with AEC. The more compact nature of the pocket in LysRS1 comes from the manner in which His240 is placed directly against the hydrocarbon side chain of lysine. Figure 4.4 illustrate how this could cause the observed differences between LysRS1 and LysRS2 with respect to the binding of AEC.
Figure 4.4. Model for the binding of AEC to LysRS1 (P. horikoshii) and LysRS2 (E. coli). The model is visualized in ball and stick for the AEC and van der Waals surface for the active site residues. All modeling was done with MOL-MOL version 2k.2.

In AEC a sulfur atom replaces the carbon at the $\gamma$ position. In the model with LysRS1 the sulfur is close enough to His240 to cause steric clashes. In the model for LysRS2, the active site is fairly open at this position and no steric clashes should occur. For GABA, the modeling did not reveal any obvious structural differences that would have led to the differences in binding. It is believed the smaller size of GABA allows it to bind in the more compact LysRS1 pocket. However, its binding in either LysRS class is not strong, as indicated by the $K_i$s for both enzymes being in the millimolar range.
The differences in enzyme inhibition by lysine analogues could have led to the distribution seen in nature if the differences found in vitro had phenotypic effects. In this scenario, organisms encoding one class of the enzyme would be able to grow in the presence of lysine analogue concentrations that would kill organisms encoding the opposite class. Both AEC and GABA are naturally occurring compounds and presented the possibility that the differences in recognition found in vitro could have consequences in nature.

Growth inhibition using both disk assays and growth in liquid media showed AEC causes growth inhibition in organisms containing LysRS2. This growth inhibition occurs because LysRS2 charges tRNA^{Lys} with AEC. The lysine analogue then becomes incorporated into proteins whose function is compromised by the noncognate amino acid. Resistance against this analogue is dependent on having only LysRS1 in the genome. If both classes are present, LysRS2 will still be able to charge tRNA^{Lys} with the analogue and poison protein synthesis. It is believed divergence of the two classes could have occurred if, in the presence of AEC or a similar inhibitor, organisms with only LysRS1 were able to grow while organisms with only LysRS2 or both classes were not.

The same situation could have occurred in the presence of GABA or an analogous compound. Experiments with in vivo growth inhibition revealed organisms encoding LysRS2 could grow in the presence of GABA concentrations that organisms encoding LysRS1 could not. This could have caused divergence and distribution of the two classes of LysRS as organisms with LysRS2 could have
grown in the presence of GABA concentrations that organisms encoding LysRS1 could not tolerate.

4.4.3. tRNA as the cause of LysRS divergence

The recognition of tRNA by each class was examined soon after the discovery of LysRS1 to determine if differences in its recognition could have led to the divergence of the LysRS enzymes (Jester et al., 2003). Binding of tRNA was examined to determine if organisms encoding one class of LysRS would have a selective advantage over the other form in aminoacylating tRNA$^{\text{Lys}}$.

Aminoacylation of tRNA was initially studied in vitro by examining the elements of tRNA that the class I and class II enzymes recognize (Ibba et al., 1999; Söll et al., 2000; Ambrogelly et al., 2002). As stated previously, the class I enzyme was found to recognize the discriminator base as well as the anticodon nucleotides, U36 or U35 and U36. The class II enzyme was shown to be similar in that it recognized the discriminator base along with the anticodon nucleotides, although most class II enzymes examined recognized all three anticodon bases. One element of differentiation between the two classes involved the recognition of a G2:U71 wobble base pair that is found in the acceptor stem of tRNAs from organisms that contain class I enzymes. While the absence of the base pair does not hinder LysRS1 from aminoacylating tRNA, its presence does prevent LysRS2 from doing so.
4.4.4. *In vivo* charging of different tRNA\textsubscript{Lys} species

*In vivo* experiments were carried out to determine if the differences in tRNA recognition found *in vitro* would have phenotypic effects (Jester et al., 2003). *Bacillus subtilis* strains 168, 143.3, and 157.1 were used to find the level of *in vivo* charging of tRNA\textsubscript{Lys} by the various forms of LysRS. The tRNA\textsubscript{Lys} charging levels were found by Northern analysis with the results shown in Figure 4.5.

![Figure 4.5](image)

**Figure 4.5. Northern analysis of tRNA\textsubscript{Lys} charging in *B. subtilis*.** Total RNA was isolated and separated by electrophoresis, transferred to membrane, and probed with a *B. subtilis* tRNA\textsubscript{Lys} specific probe. Adapted from Jester et al., 2003.

The charging of tRNA\textsubscript{Lys} from *Borellia burgdorferi* was also examined *in vivo*. In this system a vector that could express the tRNA was inserted into strains 168 and 157.1. The charging level for the tRNA was then found for each strain by Northern analysis and is shown in Figure 4.6.
Figure 4.6. Northern analysis of tRNA^{Lys} from *B. burgdorferi* charged in *B. subtilis*. RNA was isolated and separated by electrophoresis, transferred to membrane, and probed with *B. burgdorferi* tRNA^{Lys} probe. Data is presented for strains with the vector only (vo) and the vector expressing tRNA^{Lys} (Lys). Adapted from Jester et al., 2003.

The *in vivo* data led to the conclusion that any differences in charging of tRNA^{Lys} by the two classes are not significant. Each class appears able to aminoacylate a substantial amount of tRNA^{Lys} despite differences in sequence. This means that tRNA charging was probably not the cause of the divergence of the two classes.

4.4.5. Differences in anticodon recognition between LysRS1 and LysRS2 as a cause of divergence

While tRNA recognition is not believed to play a major role in the divergence of the LysRSs, it might have contributed to the distribution of LysRS along with the occurrence of AsnRS and the dual GluRS pathway. The role of LysRS in the distribution of other synthetases would have resulted from differences in the recognition of the tRNA anticodon. In LysRS1, the anticodon is recognized by a
C-terminal α-helix cage domain that specifically recognizes U35 and U36 or just U36. LysRS2 recognizes the tRNA anticodon through an N-terminal β-barrel that has the topology of a common oligonucleotide binding (OB) fold (Onesti et al., 2000).

Analysis of anticodon binding in LysRS1 and LysRS2 showed the enzymes recognize different numbers of bases specifically. Also, anticodon binding is much more important in LysRS2 than LysRS1 (Ibba et al., 1999; Söll et al., 2000; Ambrogelly et al., 2002). These differences in anticodon binding could have been a cause in the distribution of the enzymes as a correlation was found in organisms between the presence of AsnRS and LysRS1. In organisms that use a non-discriminating AspRS to aminoacylate tRNA$^{\text{Asp}}$ and tRNA$^{\text{Asn}}$, only the first two positions of the anticodon are recognized. This would allow LysRS to use the 36 position as a primary identity element. Some forms of LysRS1 only use the 36 position of the anticodon as an identity element. A similar correlation was found between the presence of the dual GluRS pathway and the presence of LysRS1. In this pathway GluRS1 aminoacylates tRNA$^{\text{Glu}}$ while GluRS2 aminoacylates tRNA$^{\text{Gln}}$. tRNA differentiation in this pathway is dependent on the binding of the anticodon (Lee and Hendrickson, 2004). The presence of LysRS1, which places less importance on the anticodon, could have increased the importance of the anticodon as an identity element for GluRS.

The results obtained in the studies with *P. horikoshii* LysRS1 confirmed what was found earlier about the binding of LysRS1 with the anticodon. The enzyme does not recognize U34 specifically, but can recognize U35 and U36 specifically.
The anticodon also does not appear to be as significant for LysRS1 as it is for LysRS2 (Table 4.2).

<table>
<thead>
<tr>
<th>tRNA (N&lt;sub&gt;34&lt;/sub&gt;N&lt;sub&gt;35&lt;/sub&gt;N&lt;sub&gt;36&lt;/sub&gt;)</th>
<th>LysRS1-&lt;br/&gt;k&lt;sub&gt;cat&lt;/sub&gt;/K&lt;sub&gt;M&lt;/sub&gt; (s&lt;sup&gt;-1&lt;/sup&gt;/mM)&lt;sup&gt;1&lt;/sup&gt;</th>
<th>LysRS2-&lt;br/&gt;k&lt;sub&gt;cat&lt;/sub&gt;/K&lt;sub&gt;M&lt;/sub&gt; (s&lt;sup&gt;-1&lt;/sup&gt;/mM)&lt;sup&gt;2&lt;/sup&gt;</th>
<th>Relative&lt;br/&gt;k&lt;sub&gt;cat&lt;/sub&gt;/K&lt;sub&gt;M&lt;/sub&gt; drop&lt;sup&gt;2&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>UUU</td>
<td>51 ± 9.7</td>
<td>1000 ± 80</td>
<td>1</td>
</tr>
<tr>
<td>CUU</td>
<td>44 ± 4.9</td>
<td>700 ± 100</td>
<td>1.4</td>
</tr>
<tr>
<td>GUU</td>
<td>29 ± 0.29</td>
<td>8.8 ± 0.8</td>
<td>110</td>
</tr>
<tr>
<td>UCU</td>
<td>1.9 ± 0.082</td>
<td>0.02 ± 0.01</td>
<td>50,000</td>
</tr>
<tr>
<td>UGU</td>
<td>1.7 ± 0.13</td>
<td>&lt;0.01</td>
<td>&gt;100,000</td>
</tr>
<tr>
<td>UUC</td>
<td>3.5 ± 0.22</td>
<td>6.6 ± 0.7</td>
<td>150</td>
</tr>
<tr>
<td>UUG</td>
<td>2.1 ± 0.064</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GUC</td>
<td></td>
<td>0.05 ± 0.01</td>
<td>20,000</td>
</tr>
</tbody>
</table>

**Table 4.2. Kinetic data for wild type LysRS1 and LysRS2 with mutant tRNAs.**<br/><sup>1</sup>Steady-state aminoacylation kinetics for *P. horikohsii* LysRS1. <sup>2</sup>Aminoacylation efficiencies of *E. coli* LysRS2 with tRNAs containing mutant anticodons (Taken from Commans et al., 1998).

The comparison revealed the losses seen with the tRNA mutants are profoundly greater for LysRS2 when compared to LysRS1. The losses in efficiency for LysRS1 were only an order or two of magnitude while the losses for LysRS2 were 2 to 6 orders of magnitude. This indicates a much greater role of the anticodon in tRNA binding for LysRS2.
The differences seen between the two enzymes for the importance of the anticodon can be explained by the structural and biochemical studies examining the binding of the anticodon to the LysRS enzymes. With LysRS2, aromatic amino acid residues such as Phe85 and Phe129 form a hydrophobic pocket that stacks the bases of U34 and U35 (Cusack et al., 1996; Commans et al., 1998). Each anticodon nucleotide is then specifically recognized by hydrogen bonds formed with amino acid residues of the enzyme as shown in Figure 4.7 (Brevet et al., 2003).

![Figure 4.7. LysRS2 recognition of the anticodon.](image)

The specific recognition of the anticodon for tRNA^{Lys} from *E. coli* by its LysRS2 enzyme. Taken from Brevet et al., 2003.

The model for binding of tRNA^{Lys} to LysRS2 shows the anticodon to be specifically recognized by interactions with five amino acid residues, one for the substituted U34 and two each for U35 and U36. This method of recognition differs greatly from what has been found for LysRS1. LysRS1 forms stacking interactions with U34 and U35 in a manner similar to LysRS2. It may also form stacking
interactions with U36. In terms of specific recognition of the bases, LysRS1 uses fewer interactions than LysRS2. LysRS1 appears to use just two amino acid residues in forming hydrogen bonds with the anticodon while LysRS2 uses five. This difference causes the anticodon recognition to be much more crucial for LysRS2 as opposed to LysRS1.

4.4.6. Organisms with both classes of LysRS

While most organisms encode just one form of LysRS, a few organisms have retained both classes. The first organisms found to contain both copies of LysRS was the archaeal methanogen, Methanosarcina barkeri (Krzycki, 2004), in which pyrrolysine was discovered. In organisms that encode this amino acid, free pyrrolysine is synthesized and then used by the synthetase PylRS to charge tRNA\textsuperscript{Pyl}, which has the anticodon for the amber stop codon. This tRNA incorporates pyrrolysine at a specific point in methyltransferase enzymes (Polycarpo et al., 2004; Blight et al., 2004). A secondary route for charging of the tRNA is performed by LysRS1 and LysRS2 together. In this pathway the two enzymes form a ternary complex with the tRNA that results in the tRNA being charged with lysine. This pathway might exist as a way for the stop codon to be read through if pyrrolysine is unavailable for tRNA charging (Polycarpo et al., 2003).

Another example where two LysRS enzymes co-exist can be found in the bacteria Bacillus cereus. In this organism the two enzymes act together to aminoacylate a tRNA that has been labeled tRNA\textsuperscript{Other}, which contains the
anticodon for tryptophan. While the role of this tRNA is still unclear, it may allow Trp codons to be read through when the Trp concentration in the cell is low (Ataide et al., 2005).

4.5. Conclusions

The experiments performed with LysRS1 revealed the mechanisms the enzyme uses to recognize its substrates. Comparisons with the mechanisms LysRS2 uses revealed differences that could be cause for the distribution of the two enzymes. Differences in lysine analogue recognition could have been a cause as organisms encoding LysRS2 were able to grow in the presence of an inhibitor that organisms encoding LysRS1 were unable to tolerate. Also, differences were seen in the recognition of the anticodon for tRNA\(^{\text{Lys}}\) that could also be a factor in the distribution of LysRS as well as other synthetases. Anticodon binding is not as important in LysRS1 as it is LysRS2. In organisms encoding LysRS1, other synthetases that recognize similar anticodons could have made greater use of the anticodon as an identity element and would exhibit greater fidelity in recognizing their cognate tRNAs.
Chapter 5

Conclusions & Perspectives

As the aaRSs provide a very fundamental function in living organisms, it was assumed they would be universally conserved throughout the living kingdoms (Ibba et al., 2000). The discovery of several alternative and novel pathways for tRNA charging revealed diversity does exist within this enzyme family. Instead of using direct charging, several organisms have two step pathways for the formation of Gln-tRNA^{Gln}, Asn-tRNA^{Asn}, and Cys-tRNA^{Cys} (Ruan et al., 2001; O’Donoghue et al., 2005). A two step pathway for Sec-tRNA^{Sec} is found in all three living kingdoms, although the machinery for its aminoacylation differs between each kingdom (Zhang et al., 2006b). In another variation, the rare 22\textsuperscript{nd} amino acid, pyrrolysine, is directly charged onto its tRNA by PylRS (Polycarpo et al., 2004).

A unique exception to synthetase universality is class I LysRS. LysRS had originally been discovered as a class II enzyme in \textit{E. coli}. In the late 1990s, a class I LysRS was found in the archaeon \textit{Methanococcus maripaludis} (Ibba et al., 1997a). This made LysRS the first and so far only synthetase to have a form in
each class. Analysis of genomes revealed LysRS1 to be found in most archaea and a few bacteria, while LysRS2 is found in all eucarya, most bacteria, and a few archaea.

In order to find a cause for this distribution, the mechanisms of LysRS1 substrate binding were examined. The binding of both lysine and tRNA were examined through steady-state kinetics. This was done to make comparisons with LysRS2 to investigate if differences in substrate binding could have caused the distribution of the two enzymes.

The binding of lysine to LysRS1 was examined by using lysine analogues to determine what functional groups of the lysine are important for recognition by LysRS1. Several analogues were tested for their ability to inhibit the aminoacylation reaction. Steady-state kinetics were used to measure the $K_i$ of each analogue to find their relative strength of binding. Comparisons of these values revealed the R-group is a much more important determinant of lysine binding than the functional groups of the $\alpha$-carbon. At the $\alpha$-carbon, the amino group was much more important for binding than the carboxylate. With the R-group, the size of the side chain was substantially important as was the position of the $\varepsilon$-amino group. The side chain size was vital as two amino acid residues in the lysine binding pocket form a closed site around the side chain. These residues form steric clashes with the bulky substitutions of analogues such as DL-hydroxylysine and AEC and prevent them from binding in the active site.

When the crystal structure of LysRS1 from *P. horikoshii* was solved, it was solved with and without bound lysine. While the crystal structure was not solved
with bound tRNA, it was found that the anticodon binding domain is an $\alpha$-helix cage that is homologous to the same domain in GluRS. The crystal structure of the GluRS:tRNA$_{\text{Glu}}$ complex was used as a template to make a model of the LysRS1:tRNA$_{\text{Lys}}$ interaction. This model identified several amino acids that could be involved in anticodon binding. Variant enzymes were then created with mutations of these amino acids to test against tRNA$_{\text{Lys}}$ covariants with mutant anticodons.

Steady-state kinetics of the aminoacylation reaction were used to measure the ability of each variant enzyme to bind each mutant tRNA. The results showed LysRS1 does not specifically recognize the base at position 34 of the anticodon, but does specifically recognize the bases at the 35 and 36 positions. Along with the model, the kinetic data gave some insight into the way the enzyme recognizes the anticodon. Aromatic residues are used to form stacking interactions with the bases of the anticodons. At least two of the nucleotides are recognized by stacking interactions, although it is possible all three are involved. Polar and charged residues are used to specifically recognize the nucleotide bases at positions 35 and 36. The charged residue Arg502 is likely used to make non-specific electrostatic interactions with the negatively charged phosphate backbone of the anticodon stem of the tRNA.

When the kinetic data for substrate binding to LysRS1 was compared to similar data for LysRS2, substantial differences were observed in both the binding of lysine analogues and tRNA. It is believed these differences could have
significant effects in nature that could have led to the distribution of the two enzymes observed today.

A few of the lysine analogues tested showed significant differences in $K_I$ between LysRS1 and LysRS2. The compounds AEC and lysinamide were much more potent inhibitors of LysRS2 than LysRS1 while GABA was a much more potent inhibitor of LysRS1 over LysRS2. The differences for AEC and GABA had phenotypic effects, and in vivo growth experiments emphasized the physiological consequences of the differences in substrate recognition between LysRS1 and LysRS2. These differences could have caused the divergence of the enzymes as organisms with one class of the enzyme would have been able to grow in the presence of inhibitors that organisms with the other LysRS class would not. The patterns of inhibition indicate that inhibitor resistance would only be conferred on organisms harboring a single form of LysRS.

Differences between the two forms of LysRS were also observed for the binding of the anticodon. LysRS1 only specifically recognizes positions 35 and 36 of the anticodon while LysRS2 specifically recognizes all three positions. Also, the anticodon was a much more important identity element for LysRS2 than for LysRS1. It was previously hypothesized that the presence of LysRS1 could have played a role in the divergence of AsnRS and the dual GluRS pathway (Söll et al., 2000; Salazaar et al., 2003).
The discovery of a class I LysRS raised the question why two classes of the enzyme exist in nature. Studies of the mechanisms LysRS1 uses to bind its substrates revealed differences in substrate recognition with LysRS2. Some of these differences could have significance for the divergence of the two enzymes.

Analysis of the binding of lysine analogues through steady-state kinetics revealed the R-group is much more important for discrimination than the functional groups of the α-carbon. The enzyme relies on hydrophobic interactions with the hydrocarbon side chain and electrostatic interactions with the ε-amino group. This method of binding was confirmed in mutagenesis studies of the LysRS1 lysine binding pocket (Wang et al., 2004). Mutation of His242 in the *B. burgdorferi* LysRS1, which is responsible for the hydrophobic interaction, showed a significant decrease in catalytic efficiency. Mutation of Glu43, which forms the electrostatic interaction with the ε-amino group, abolished enzyme activity.

Comparisons of analogue binding between LysRS1 and LysRS2 revealed differences which carried over to growth conditions as organisms harboring one class could grow in the presence of inhibitors organisms harboring the other class could not tolerate. This difference could be a cause of the enzyme divergence. While the naturally occurring AEC could be a good candidate for causing
divergence favoring LysRS1, GABA is not a good candidate for divergence favoring LysRS2. While a substantial difference in inhibition was seen between LysRS1 and LysRS2, the $K_i$ for GABA with LysRS1 was 8.0 mM. This indicates a poor level of binding and suggests the effects of GABA on LysRS1 in nature would not be substantial as free GABA concentrations are not this high. Further research is needed to find a specific inhibitor of LysRS1 with a $K_i$ value several orders of magnitudes lower than GABA. A discovery of a specific inhibitor for LysRS1 could have significance for the development of an antibiotic taking advantage of the differences in inhibition of LysRS1 and LysRS2. Several pathogens such as $B$. burgdorferi (Lyme disease), $T$. pallidum (syphilis), and $B$. hermsii (relapsing fever), contain LysRS1 (Mejlhede et al., 2002). A specific inhibitor of LysRS1 could kill the pathogen while not affecting the activity of the host’s LysRS2.

For the binding of the anticodon, LysRS1 was found to bind the nucleotides through stacking interactions and hydrogen bonds. In comparison to LysRS2, anticodon recognition was not as significant for LysRS1 as LysRS2. This could have consequences for the distribution of LysRS as well as other synthetases. The anticodon for Lys has similarities to the anticodons for Asn, Glu, and Gln. Less importance on the anticodon by LysRS1 could have allowed synthetases that recognize $tRNA^{Asn}$, $tRNA^{Glu}$, and $tRNA^{Gln}$, to put greater importance on recognizing their anticodons. Further studies of the identity elements of each tRNA are needed to determine if this is a plausible cause of synthetase divergence and distribution.
The studies of anticodon binding in LysRS1 were based on a model made of the LysRS1:tRNA<sup>Lys</sup> complex using the structure of GluRS:tRNA<sup>Glu</sup> as a template, as the anticodon binding domains of the two are homologous. The kinetic studies showed some of the hypotheses of the model to be incorrect. For example, Arg502 was believed to form hydrogen bonds with U36 as a homologous Arg in GluRS recognizes C36. Our data showed Arg502 does not perform this function. Further studies of the α-helix cage domain in both LysRS1 and GluRS are needed to determine the level of functional homology between the two.

Recent efforts in the study of synthetases have focused on altering the substrate specificity of the enzymes to expand the genetic code (Wang et al., 2006). LysRS1 from <em>P. horikoshii</em> was engineered to bind the amino acid L-homoglutamine and aminoacylate a tRNA with the four base codon AGGA (Anderson et al., 2004). The findings of these studies could in the future assist in the creation of orthogonal tRNA/synthetase pairs.

The discovery of LysRS1 provided further evidence the synthetases are a diverse family of enzymes. The studies of the way LysRS1 recognizes its substrates and their comparison to LysRS2 showed how amino acid and tRNA recognition have led to the divergence of the aaRS family as a whole. In broader terms, this ability of the aaRSs to adapt to divergent substrates helps to explain how the synthetases facilitate the incorporation into the genetic code of non-canonical amino acids such as Sec and Pyl.


Cusack, S., Yaremchuck, A., and Tukalo, M. (1996) The crystal structures of *T. thermophilus* lysyl-tRNA synthetase complexed with *E. coli* tRNA\textsubscript{Lys} and a *T. thermophilus* tRNA\textsubscript{Lys} transcript: anticodon recognition and conformational changes upon binding of lysyl-adenylate analogue. *EMBO J.* 15:6321-6334.


Jackman, J.E. and Phizicky, E.M. (2006) tRNA$^{\text{His}}$ guanylyltransferase adds G-1 to the 5’ end of tRNA$^{\text{His}}$ by recognition of the anticodon, one of several features shared with tRNA synthetase. *RNA* **12**:1007-1014.


