PROTEINS: Structure, Function, and Bioinformatics 54:693–704 (2004)

Molecular Modeling Study of the Editing Active Site of *Escherichia coli* Leucyl-tRNA Synthetase: Two Amino Acid Binding Sites in the Editing Domain

Keun Woo Lee and James M. Briggs*
Department of Biology and Biochemistry, University of Houston, Houston, Texas

ABSTRACT **Aminoacyl-tRNA synthetases (aaRSs) strictly discriminate their cognate amino acids. Some aaRSs accomplish this via proofreading and editing mechanisms. Mursinna and coworkers recently reported that substituting a highly conserved threonine (T252) with an alanine within the editing domain of *Escherichia coli* leucyl-tRNA synthetase (LeuRS) caused LeuRS to cleave its cognate aminoacylated leucine from $tRNA^{Leu}$ (Mursinna et al., Biochemistry 2001;40:5376–5381). To achieve atomic level insight into the role of T252 in LeuRS and the editing reaction of aaRSs, a series of molecular modeling studies including homology modeling and automated docking simulations were carried out. A 3D structure of *E. coli* LeuRS was constructed via homology modeling using the X-ray structure of *Thermus thermophilus* LeuRS as a template because the *E. coli* LeuRS structure is not available from X-ray or NMR studies. However, both the X-ray *T. thermophilus* and homology-modeled *E. coli* structures were used in our studies. Amino acid binding sites in the proposed editing domain, which is also called the connective polypeptide 1 (CP1) domain, were investigated by automated docking studies. The root mean square deviation (RMSD) for backbone atoms between the X-ray and homology-modeled structures was 1.18 Å overall and 0.60 Å for the editing (CP1) domain. Automated docking studies of a leucine ligand into the editing domain were performed for both structures: homology structure of *E. coli* LeuRS and X-ray structure of *T. thermophilus* LeuRS for comparison. The results of the docking studies suggested that there are two possible amino acid binding sites in the CP1 domain for both proteins. The first site lies near a threonine-rich region that includes the highly conserved T252 residue, which is important for amino acid discrimination. The second site is located in a flexible loop region surrounded by residues E292, A293, M295, A296, and M298. The important T252 residue is at the bottom of the first binding pocket. Proteins 2004;54:693–704.**

Key words: leucyl-tRNA synthetase; amino acid editing; CP1 domain; homology modeling; docking

INTRODUCTION

High fidelity in protein synthesis is essential to all biologic systems and the fidelity is critically dependent on accurate aminoacylation of an amino acid to its correct cognate tRNA.[1–3] The aminoacylation reaction for each of the standard 20 amino acids is controlled by a family of enzymes called the aminoacyl-tRNA synthetases (aaRSs). The overall reaction is achieved by each aaRS in the following two reactions:

$$\mathrm{AA} + \mathrm{ATP} \overset{\text{aaRS}}{\rightleftharpoons} \text{AA-AMP} + \mathrm{PPi}, \quad (1)$$

$$\text{AA-AMP} + \mathrm{tRNA} \overset{\text{aaRS}}{\longrightarrow} \text{tRNA-AA} + \mathrm{AMP}. \quad (2)$$

The first step forms the activated intermediate, an aminoacyl-adenylate (AA-AMP), from the amino acid and adenosine triphosphate (ATP; eq. 1). The aminoacyl moiety is transferred from the adenylate to the 3′-terminal adenosine (3′-A) of tRNA.[1–3]

Some aaRSs have an editing activity that hydrolyzes incorrectly misactivated and/or misaminoacylated amino acids.[4] These include phenylalanyl-, lysyl-, methionyl-, threonyl-, alanyl-, prolyl-, isoleucyl-, valyl-, and leucyl-tRNA synthetases.[5–14] In each case, these aaRSs must discriminate their cognate amino acid from other highly similar or isosteric amino acids. Aminoacylation by the aaRSs have been suggested to function by a "double-sieve" model that relies on discrimination in separate aminoacylation and editing sites.[4,8,13] These two different sieves enable the aaRS to achieve high fidelity by employing varied amino acid specificity strategies. The first activation or "coarse" sieve activates cognate amino acids as well as isosteric or closely related noncognate amino acids that fit into the cognate amino acid binding pocket. The second

Abbreviations: aaRS, amino acid tRNA synthetase; LeuRS, leucyl-tRNA synthetase; IleRS, isoleucyl-tRNA synthetase; ValRS, valyl-tRNA synthetase; CP1, connective polypeptide 1.

*Correspondence to: James M. Briggs, Department of Biology and Biochemistry, University of Houston, Houston, TX 77204-5001.
Email: jbriggs@uh.edu

Received 8 March 2002; Accepted 10 September 2002

"fine" or editing sieve excludes the cognate amino acid but hydrolyzes/edits the misactivated noncognate amino acid and/or mischarged tRNA. For IleRS and ValRS, the separation of the activation and editing sites has been identified.[15–19] The activation site is located within the ATP-binding Rossmann fold that is common to all class I aaRSs[14,20] and the editing site lies in a large inserted domain called connective polypeptide 1 (CP1).[15–19]

The CP1 domain of leucyl-tRNA synthetase (LeuRS) shares extensive homology with the CP1 domain of IleRS and ValRS. LeuRS has been shown to misactivate and edit a series of amino acid derivatives.[14,21,22] Mursinna and coworkers recently reported that a highly conserved threonine residue (T252) within the CP1 domain of *Escherichia coli* LeuRS plays a key role in the editing activity.[14] Substitution of the threonine with an alanine (T252A) altered the editing action of LeuRS such that it hydrolyzed its cognate charged leucine from the $tRNA^{Leu}$. Therefore, the T252A mutant cannot effectively discriminate leucine from noncognate amino acids in its editing active site. The main goal of this article is to attain an atomic-level understanding of this editing-related phenomenon. To achieve it, we carried out a series of molecular modeling studies including homology modeling and automated docking simulations. Because no X-ray or NMR structure is available for the *E. coli* LeuRS, the 3D structure was constructed via homology modeling methods using the X-ray structure of the *Thermus thermophilus* LeuRS[12] as a template. Automated docking experiments were carried out to identify amino acid binding sites in the editing domain of both structures: the homology structure of *E. coli* LeuRS and the X-ray structure of *T. thermophilus* LeuRS. We compared the results with the pocket that was proposed by Mursinna et al.[14] Identification of the editing active site is critically important for novel protein synthesis and drug design. In addition, atomic-level understanding of the mechanism of the editing activity of aaRSs is critical to understand this fundamental process in biochemistry.

METHODS

Homology Modeling

Template structure and sequence alignment

Because currently only one X-ray structure is available for LeuRS, a single X-ray structure was used as the template for homology modeling. The X-ray structure of *T. thermophilus* LeuRS was obtained from Cusack[12] and used as the template structure for building a model of our *E. coli* LeuRS. The sequence of *E. coli* LeuRS was reported by Haertlein et al.[23–25] and obtained from the SWISS-PROT data bank [SYL_ECOLI (P07813)]. The sequence similarity and identity between the two are 73.3 and 45.0%, respectively, according to the ClustalX program. The structure and sequence data were manipulated using the HOMOLOGY and MODELER modules of INSIGHTII, which are well-known homology modeling programs.[26]

Sequence alignment is a central technique in homology modeling. It is used to establish a one-to-one correspondence between the amino acids of the reference protein(s) and those of the unknown protein in the structurally conserved regions. This correspondence is the basis for transferring coordinates from the reference to the model protein. The sequence alignment method we used here is the *segment pair overlap* algorithm developed by Schuler et al. based on local similarity measurements.[27] The main advantage of this method is that it not only aligns the sequences but also identifies blocks of related sequence segments and estimates the probabilities that these high-scoring blocks could be found by chance. Those blocks showing high statistical significance are likely to contain structurally conserved regions.

3D structure generation

3D structures based on the information from the results of the alignment procedure were generated by MODELER. The method used in MODELER is different from those of other general homology programs. The scheme developed by Sali et al. employs probability density functions (PDFs) as the spatial restraints rather than energy.[28–31] The main-chain conformation of a given residue in the model will be described by restraints that depend upon the residue type, the main chain conformation of equivalent residues in the reference proteins, and the local sequence similarity. The PDFs that are used in restraining the model structure are derived from correlations between structural features in a database of families of homologous proteins aligned on the basis of their 3D structures. These functions are used to restrain C_α—C_α distances, main chain N—O distances, main-chain and side-chain dihedral angles, etc. The individual restraints are assembled into a single molecular PDF (MPDF). Each PDF has a similar meaning as the energy terms in a molecular mechanics (MM) force field function. These PDFs were originally constructed from over 400 protein structures in the Protein Data Bank (PDB). They were used, along with the coordinate information from the template protein, to build a final MPDF. The related or reference protein structures are used to derive spatial restraints expressed as PDFs for each of the restrained features of the model. For the aligned residues, all atomic coordinates of the residue are copied from the template protein, according to the restraints. However, for the mismatched residues only the C_α atom coordinates are copied from the template protein while the remaining atomic coordinates are constructed by using internal coordinates derived from a CHARMM (Chemistry at HARvard Macromolecular Mechanics) topology library.[32] All of this coordinate information, as well as the PDFs, are used to build a final MPDF for the model protein.

The 3D protein model is then obtained from an optimization of the newest MPDF. The optimization procedure itself consists of a variable target function method[33] with a *conjugate gradient* minimization scheme. This method, therefore, was designed to find the most *probable* 3D structure of a protein, given its amino acid sequence and its alignment with related structures. This kind of homology modeling scheme is in general based on the assumption that the structure of an unknown protein is similar to known structures of some reference proteins.

Energy minimization

Before doing docking calculations, the homology-modeled structure was energy minimized to account for the effect on the local structure of the mutation of threonine 252 to alanine (T252A). Because we performed a minimization for the T252A mutant protein, we also performed a similar minimization on the WT structure so that both resultant structures were refined in the same way (i.e., through the same minimization procedure and to the same gradient tolerance). Our full LeuRS homology-modeled structure contains 800 residues and has dimensions of ca. 120 × 90 × 80 Å, so a water box for minimization or molecular dynamics should be ca. 20 Å more in each dimension. Because the full protein is so large, our minimization was performed with a 5-Å water layer (4012 water molecules) for the T252A mutant as well as the wild-type (WT) structure. The calculation was performed using the CHARMM program with CHARMM27 parameters[32] and a 12-Å cutoff. First, the *steepest descent* algorithm was used for 1000 steps to remove close van der Waals contacts, followed by the more efficient *Adopted Basis Newton-Raphson* (ABNR) scheme until a tolerance of 0.05 kcal/mol · Å in the gradient was reached.

Automated Docking Simulation

The automated docking of amino acids into the editing site in the CP1 domain of LeuRS was performed with the AUTODOCK3.0 program developed by Olson et al.[34–37] It performs a rapid energy evaluation through precalculated grids of affinity potentials with a variety of search algorithms to find suitable binding positions for the ligand to the given protein. It also allows torsional flexibility of the ligand; however, the protein is required to be rigid in the simulation.

Ligand and protein preparation

In the present work, Leu was docked to three different proteins: WT and T252A of *E. coli* LeuRS and WT of *T. thermophilus* LeuRS for comparison purposes. The l-leucine amino acid structure was built with the BIOPOLYMER module of INSIGHTII; the terminal N and C moieties were charged as NH_3^+ and COO^-.

The WT structure of *T. thermophilus* LeuRS was from the X-ray structure and the WT structure of *E. coli* LeuRS was homology modeled. As described above, the T252A of *E. coli* LeuRS was constructed by virtual mutation from the homology-modeled structure using the BIOPOLYMER module of INSIGHTII. It was then energy minimized with a 5-Å water layer in the same manner as for the WT. The root mean square deviation (RMSD) of backbone atoms between the WT and T252A after minimization was 0.24 Å and the RMSD of backbone atoms between the minimized and starting structures of WT was 0.68 Å.

Computational details

The rotatable torsional angles of ligand molecules are specified by the AUTOTORS program. The center of a 34 × 34 × 34-Å grid was placed near of the conserved T252 residue in the gorge of the CP1 domain of the three different systems. An 85 × 85 × 85-point grid was used with a spacing between each grid point of 0.4 Å. Once the grid size and center were determined, the potential energy at the each grid point was calculated using the AUTOGRID program. After constructing all energy data for all grid points, the main docking simulation followed. The AUTODOCK program computes the interaction energy between the flexible ligand conformer and each grid point via a combined searching algorithm. A translational step of 0.2 Å and a rotational step of 5° were used. The initial starting position and dihedral angles for rotatable angles were selected randomly and finally 50 docking simulations were run to obtain statistically acceptable docked structures of each ligand. Therefore, this docking simulation resulted in 50 docked structures, along with their docking energies per ligand in each protein structure. In these docking studies, the entire protein containing 800 residues was present, although docking was focused on the CP1 domain.

Lamarckian genetic algorithms

AUTODOCK3.0 can use a Lamarckian genetic algorithm (GA) as the search method for the docking simulations.[37] GAs use concepts based on biologic evolution.[38,39] In a molecular docking simulation, the particular arrangement of a ligand can be defined by a set of state variables that describe translation, orientation, and conformation of the ligand. In GAs, each variable corresponds to a gene. The ligand's state corresponds to the *genotype* and the atomic coordinates correspond to the *phenotype*. In molecular docking, the *fitness* is the total interaction energy of the ligand with the protein. After generating the population by distributing the random values to the genotype, random pairs of individuals are mated using *crossover*, in which the new generation inherits genes from either parent. In addition, some offspring undergo random *mutation*.[37] In addition to GA as the global search method, AUTODOCK uses an *adaptive* local search (LS) method that is based on that of Solis and Wets[40] in that it adjusts the step size depending upon the recent history of energy.[37] The *Lamarckian genetic algorithm (LGA)*, as implemented in AUTODOCK, is a hybrid of the GA method (for global search) and adaptive LS (for local search) and has enhanced performance relative to simulated annealing and GA alone.[41]

We used the LGA in our docking simulations with the following settings: a maximum number of 1.5×10^6 energy evaluations, an initial population of 50 randomly placed individuals, a maximum number of 27,000 generations, a mutation rate of 0.02, a crossover rate of 0.80, an elitism value (number of top individuals that automatically survive) of 1, and finally 10 generations for picking the worst individual. For the adaptive local search method, the pseudo-Solis and Wets algorithm was applied using a maximum of 300 iterations per search. The probability of performing a local search on an individual in the population was 0.06 and the maximum number of consecutive successes/failures before changing the local search step size was 4.

```
TT(  1):  MEKYNPHAIEAKWQRFWEEKGPMKAKDLPGGRGKQYVLVMFPYPSGDLH
EC(  1): mqeqyrpeeieskvqlhwdek-rtfevtedeskekyyclsmlpypsgrlh

TT( 50): MGHLKNYTMGDVLARFRRMQGYEVLHPMGWDAFGLPAENAALKFGVHPKD
EC( 50): mghvrnytigdviaryqrmlgknvlqpigwdafglpaegaavknntapap

TT(100): WTYANIRQAKESLRLMGILYDWDREVTTCEPEYYRWNQWIFLKMWEKGLA
EC(100): wtydniaymknqlkmlgfgydwsrelatctpeyyrweqkfftelykkglv

TT(150): YRAKGLVNWCPKCQTVLANEQVVEGRCWRHEDTPVEKRELEQWYLRITAY
EC(150): ykktsavnwcpndqtvlaneqvidgccwrcdtkverkeipqwfikitaya

TT(200): AERLLKDLEGLNWPEKVKAMQRAWIGRSEGAEILFPVEGKEVRIPVFTTR
EC(200): dellndldkldhwpdtvktmqrnwigrsegveitfnvndydntltvyttr

TT(250): PDTLFGATFLVLAPEHPLTLELAAPEKREEVLAYVEAAKRKTEIERQAEG
EC(250): pdtfmgctylavaaghpla--qkaaennpelaafidecrntkvaeaemat

TT(300): REKTGVFLGAYALNPATGERIPIWTADYVLFGYGTGAIMAVPAHDQRDYE
EC(298): mekkgvdtgfkavhpltgeeipvwaanfvlmeygtgavmavpghdqrdye

TT(350): FARKFGLPIKKVIERPGEPLPEPLERAYEEPGIMVNSGPFDGTESEEGKR
EC(348): faskyglnikpvilaadgsepdlsqqaltekgvlfnsgefngldheaafn

TT(400): KVIAWLEEKGLGKGRVTYRLRDWLISRQRYWGTPIPMVHCEACGVVPVPE
EC(398): aiadkltamgvgerkvnyrlrdwgvsrqrywgapipmvtledgtvmptpd

TT(450): EELPVLLPDLKDVEDIRPKGKSPLEAHPEFYETTCPKCGGPAKRDTDTMD
EC(448): dqlpvilpe------dvvmdgitspikadpewakttvngmpalretdtfd

TT(500): TFFDSSWYYLRYTDPHNDRLPFDPEKANAWMPVDQYIGGVEHAVLHLLYS
EC(492): tfmesswyyarytcpqykegmldseaanywlpvdiyiggiehaimhllyf

TT(550): RFFTKFLHDLGMVKVEEPFQGLFTQGMVLAWTDFGPVEVEGSVVRLPEPT
EC(542): rffhklmrdagmvnsdepakqllcqgmvla----------dafyyvgeng

TT(600): RIRLEIPESALSLEDVRKMGAELRPHEDGTLHLWKPAVMSKSKGNGVMVG
EC(582): ernwvspvdaiverdekgrivkakdaaghelvytgmskmsksknngidpq

TT(650): PFVKEQGADIARITILFAAPPENEMVWTEEGVQGAWRFLNRIYRRVAEDR
EC(632): vmverygadtvrlfmmfaspadmtlewqesgveganrflkrvwklvye-h

TT(700): EALLETSGVFQAEALEGKDRELYGKLHETLKKVTEDLEALRFNTAIAALM
EC(681): takgdvaalnvdaltenqkalrrdvhktiakvtddigrrqtfntaiaaim

TT(750): EFLNAL-YEYRKDRPVTPVYRTAIRYYLQMLFPFAPHLAEELWHWFWPDS
EC(731): elmnklakaptdgeqdralmqeallavvrmlnpftphicftlwqelkgeg

TT(799): LFEAGWPELDEKALEK (Total 814)
EC(781): didnapwpvadekamvedstlvvvqvngkvrakitvpvdateeqvrerag

TT     :
EC(831): qehlvakyldgvtvrkviyvpgkllnlvvg (Total 860)
```

Fig. 1. Sequence alignment results. Bold indicates the blocks that show high statistical significance (see the text for details).

RESULTS

Homology Modeling

Sequence alignment

The sequence alignment from the *segment pair overlap* algorithm is shown in Figure 1, where the bold letters represent residues that were used for the blocks, as previously described. The first half (ca. 55%) of the sequence that comprises the conserved Rossmann binding domain in class I aaRRs is almost perfectly aligned. The second half of the protein, which is less conserved, however, contains several regions that were not well aligned. In the figure, the last aligned residue of *E. coli* LeuRS is Val796 and no template structure exists for the remaining 64 residues (from 797–860). Therefore, the first 800 residues, including 3 extra residues, were considered for the structure generation procedure while the remaining 60 residues were discarded because a reliable structure could not be generated for that region.

3D structure generation

Using a comparative protein modeling method, the 3D structure of *E. coli* LeuRS was generated with the MODELER program using multiple alignment information and molecular PDF functions. The final 3D structure generated by our homology modeling method is displayed along with the reference protein in Figure 2. Figure 2(A) shows stereo ribbon diagrams of the X-ray structure of *T. thermophilus* LeuRS,[12] while the final homology-modeled structure of *E. coli* LeuRS is given in Figure 2(B). The RMSD for the backbone atoms of the two entire structures is 1.18 Å; however, the RMSD for the CP1 domains (195 residues of 224–417; green in Fig. 2) is only 0.60 Å. The small RMSD can be interpreted to mean that the two structures share common homology and the generated structure is reasonable. The worst aligned region is the leucyl-specific domain (residues 572–619; purple in the figure). This domain is characteristic of the LeuRS protein and the size and shape of it varies substantially among organisms.[12] This domain may be important for aminoacylation because it is close to one of the key regions of the protein, the KMSKS region.[42,43] However, our main concern in the present study is with the CP1 domain [green in Fig. 2(B)] to investigate the editing reaction. The CP1 domain is globular while the full LeuRS is elongated with ca. 800 residues.

Protein structure validation

To validate our homology-modeled *E. coli* LeuRS structure, a Ramachandran plot was drawn and the structure was analyzed by PROCHECK, a well-known protein structure checking program,[44,45] and PROSTAT from the HOMOLOGY module of INSIGHTII.[26] The phi-psi plot is shown in Figure 3 while the more detailed results are listed in Table I. Figure 3 shows that the structure is reasonable overall because the spot distributions for the homology-modeled structure was similar to the standard X-ray structure.

Table I lists more detailed scores for both structures. The results showed that our modeled structure was competitive with the standard structure. For example, the percentage of the residues in "most favored regions" (the darkest area in the figure) are 92.9% for X-ray *T. thermophilus* LeuRS and 91.8% for homology-modeled *E. coli* LeuRS structure. The overall PROCHECK score for the homology-modeled structure is −0.10. This score indicates that the modeled structure is acceptable because the recommended value is greater than −0.5.[44] In considering only the CP1 domains, which is the focus of this investigation, the PROCHECK score for the modeled structure is −0.06.

Another protein structure analysis program, PROSTAT available in the HOMOLOGY module of INSIGHTII, was used for structure validation.[26] The program measures all bond distances, bond angles, and dihedral angles for the given protein structure and compares them with its own database. It then reports the instances that are out of usual ranges. The results are listed at the bottom of Table I. They show that the two structures are comparable, although the CP1 domains alone showed the best results.

Finally, the 3D homology-modeled structure was profiled using the VERIFY3D program[46] and then compared with the template structure. The 3D-1D scores averaged

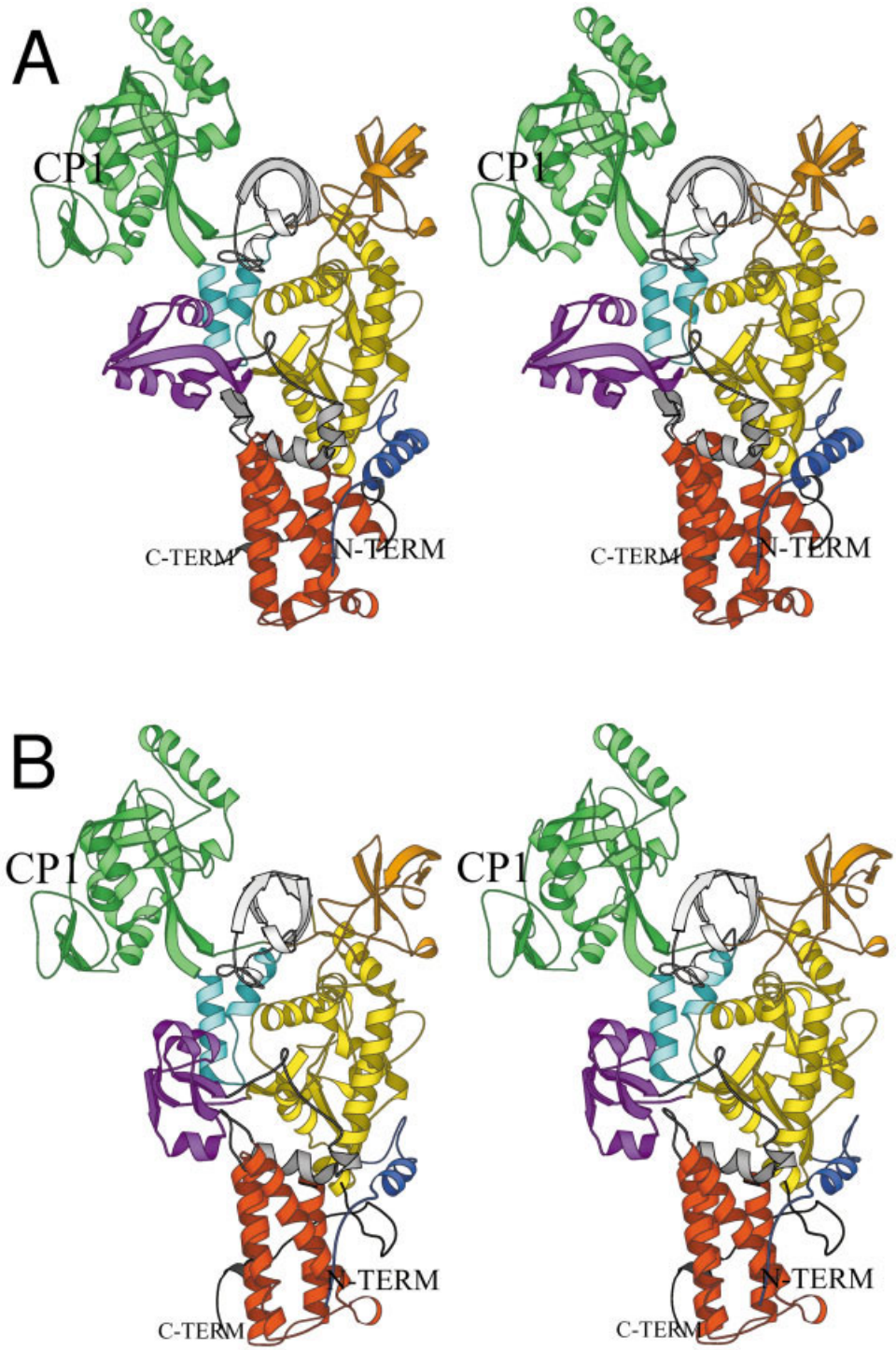

Fig. 2. Stereo ribbon diagrams of the X-ray structure of *T. thermophilus* LeuRS (**A**) and the homology-modeled structure of *E. coli* LeuRS (**B**). The domains are colored as follows: N-terminal extension (blue), catalytic domain (Rossmann fold, yellow), Zn-1 domain (white), helical hairpin insertion (cyan), editing domain (CP1, green), Zn-2 domain (orange), leucyl-specific insertion domain (purple), connecting module (gray), anti-codon binding domain (red), and C-terminal extension (black). The backbone atom RMSD with respect to the X-ray structure for the full modeled structure is 1.18 Å and for the CP1 domain only (green) is 0.60 Å. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

over all residues were 0.35 for our modeled structure and 0.48 for the X-ray structure. The score value, for a well-folded protein, should be above 0.1.[46] A higher score indicates a better 3D structure.

We conclude that our homology-modeled structure is acceptable overall because it does not have any fatal defects. Although the entire structure of the protein will be used in future automated docking studies, our binding site

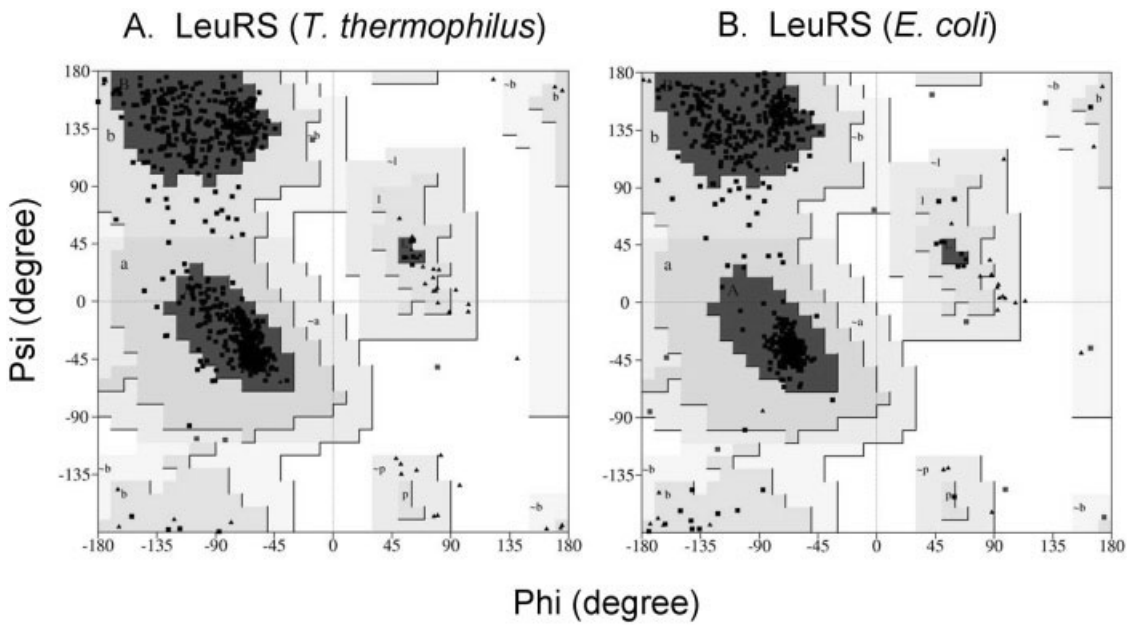

Fig. 3. Ramachandran plot of the X-ray structure of *T. thermophilus* LeuRS (**A**) and the homology-modeled structure of *E. coli* LeuRS (**B**). The different colored areas indicate "disallowed (white)," "generously allowed," "additional allowed." and "most favored (the darkest area)" regions in order of increasing darkness (see Table I).

TABLE I. Results of Protein Structure Check by PROCHECK[44,45] and PROSTAT[26]

	T. thermophilus LeuRS		*E. coli* LeuRS	
	All	CP1 only	All	CP1 only
Residues in most favored regions	645 (92.9%)	149 (94.3%)	614 (91.8%)	149 (93.7%)
Residues in additional allowed regions	45 (6.5%)	8 (5.1%)	46 (6.9%)	9 (5.7%)
Residues in generously allowed regions	3 (0.4%)	1 (0.6%)	7 (1.0%)	1 (0.6%)
Residues in disallowed regions	1 (0.1%)	0 (0.0%)	2 (0.3%)	0 (0.0%)
Number of non-Gly and non-Pro residues	694 (100.0%)	158 (100.0%)	669 (100.0%)	167 (100.0%)
Number of end residues	2	2	2	2
Number of Gly residues (shown as white boxes)	60	20	52	17
Number of Pro residues	58	16	39	8
Total number of residues	814	196	800	194
Overall PROCHECK score[a]	0.36	0.35	−0.10	−0.06
Number of bond distances with significant deviations[b] (by PROSTAT)	0	0	1	0
Number of bond angles with significant deviations[b] (by PROSTAT)	0	0	0	0
Number of dihedral angles with significant deviations[b] (by PROSTAT)	3	1	2	0

[a]Recommended value > -0.50 and investigation is needed for < -1.0.
[b]Number of instances for which the property differs more than 10 standard deviations from the reference value.

search will be focused on the CP1 domain (194 residues from I224 to L417).

Energy minimization

The backbone RMSD between the final minimized WT and the template (*T. thermophilus*) structure is 1.25 Å while that between the starting homology-modeled structure and the template was only 1.18 Å, so the minimization did not appear to cause a substantial distortion in the structure and appears to have accomplished the goal of relieving steric clashes and close contacts. In addition, because we had to minimize the mutant protein (T252A) we also had to minimize the WT protein so that they were at the same level of refinement for use in further studies.

Automated Docking Simulations

Clustering the docked structures

Results of automated docking simulations of a leucine ligand with the CP1 domain of *E. coli* LeuRS are shown in Figure 4: WT [Fig. 4(A)] and T252A mutant [Fig. 4(B)]. Additional docking results of leucine against the WT *T. thermophilus* LeuRS CP1 domain are given in Figure 4(C) for comparison purposes. In the figure, only the 50 C_α atoms of the 50 docked leucines are rendered (as black spheres). As Figure 4 shows, most of the spheres are located in two major regions. For example, 45 of the 50 spheres are gathered in one of these two major regions in the WT LeuRS with leucine [Fig. 4(A) and Table II]. The two regions will be called POCKET 1 and POCKET 2 for

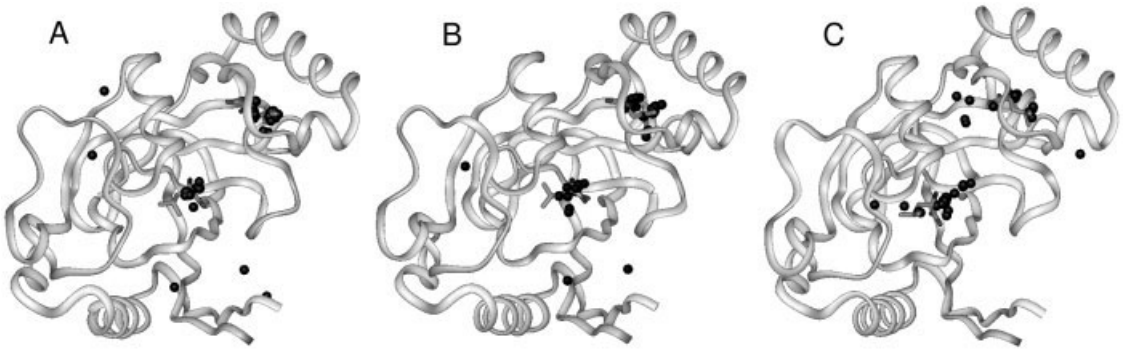

Fig. 4. Automated docking results of the leucine against the CP1 domain. Only the CP1 domain and C_α atoms (black spheres) of 50 docked leucines are visualized for clarity. The results were given for the WT (**A**) and T252A mutant (**B**) of *E. coli* LeuRS and the WT (**C**) of *T. thermophilus* LeuRS for validation and comparison purposes. The lowest-energy docked structures for each pocket are rendered by sticks and shown in both pockets. In each panel, POCKET 1 is in the center and POCKET 2 is in the upper right.

TABLE II. Automated Docking Results of the 50 Leucine Ligands into the CP1 Domain of *E. coli* and *T. thermophilus* LeuRS

	Number of leucines docked into POCKET 1[a]	Number of leucines docked into POCKET 2[a]	Total number of leucines[b]
WT (*E. coli*)	14	31	45
T252A (*E. coli*)	22	24	46
WT (*T. thermophilus*)	22	13	35

[a]The number of the docked ligands in which the C_α atom is located within 4 Å from that of the lowest docked energy structures for each pocket.
[b]The total number of the docked ligands in POCKET 1 or POCKET 2.

convenience. POCKET 1 lies in the central area that contains the conserved threonine-rich editing region [Fig. 4(A)]. POCKET 1 is comprised of residues T247, T248, T252, M336, D342, and D345, while POCKET 2 is surrounded by residues E292, A293, M295, A296, and M298. For each cluster, the lowest-energy docked structure was selected as a sample conformation for each pocket. The representative docked leucine is visualized in the each pocket to show the orientation of the molecule (Fig. 4).

Methionine and several other amino acids were also docked into the CP1 domain and compared to the results from the leucine docking. The overall docking trends are similar to the leucine case (data not shown). The detailed docking results of leucine are summarized in Table II. The docking frequencies of leucine into POCKET 1 or POCKET 2 were determined from 50 runs. The values in the table represent C_α atoms of the docked ligands within 4 Å from that of the lowest-energy docked ligands in each pocket. The table shows that the two pockets are the dominating docking sites. For leucine, 35–46 runs from a total of 50 docked to one of the two pockets for both WT and T252A mutant LeuRS. This strong clustering suggests that both the docked structures and the docking sites are relevant. Therefore, both clustered regions were proposed as amino acid binding sites in the editing domain of the *E. coil* LeuRS. Because the clustering patterns were also similar for the methionine, isoleucine, and valine, we think the pockets are not designed only for the leucine but for other amino acids (results not shown).

Demonstration of the amino acid binding mode

Two sample structures of the docked leucines were compared for the WT and T252A mutant [Figs. 5(A) and 5(B)]. The lowest-energy docked structures are shown for the purposes of illustration of the binding modes. It should be noted that the sampled structures do not mean the real binding mode between the protein and the ligand because in general the docking study has many limitations to extract that kind of conclusion. Four neighboring residues along with the important residue T252 (A252 for the T252A mutant) are shown in the figure. The overall docked conformations for both ligands are similar, although the positions of the side-chains are slightly rotated. The N atom in NH_3^+ of each leucine was hydrogen bonded to the OD (Oδ) atom of the highly conserved D345 and an O atom in COO^- was hydrogen bonded to the OG (Oγ) atom of the T247. To compare these binding patterns with the an experimentally determined structure, an image of the X-ray structure of the complex between valine and the isoleucyl-tRNA synthatase (IleRS) from *T. thermophilus* is shown [Fig. 5(C)].[8,15]* The corresponding four residues are

*The X-ray structures of the IleRS were reported in "protein-only" and "valine-complexed" form by Nureki et al.[15] Because, currently, the coordinates are available only for the protein-only form from the PDB (code: 1ILE), a valine was introduced to the editing site and then manually aligned along the figures in their articles[8,15] only for rough comparison purpose.

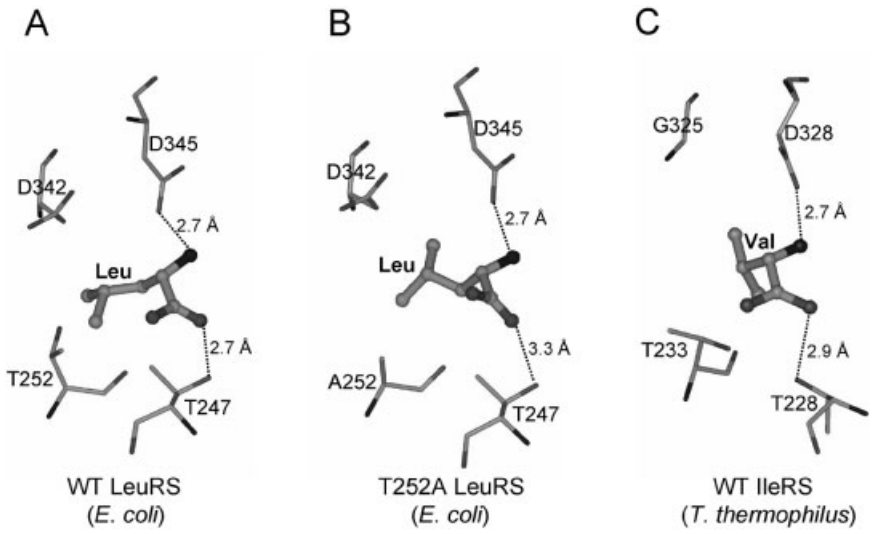

Fig. 5. Binding modes of leucine with the lowest-energy docked leucine structures in POCKET 1 for the WT (**A**) and T252A mutant *E. coli* LeuRS (**B**). Four neighboring residues of the docked leucine, including the highly conserved threonine and aspartate (T252 and D345), are visualized. For the comparison of the amino acid binding pattern to other systems, the editing site of the X-ray structure[8] of *T. thermophilus* IleRS was also given with a manually placed valine in its editing site (**C**).[52]

shown in the figure. Our binding site in *E. coli* LeuRS and the binding mode seem to be reasonable compared to the real amino acid binding structure in *T. thermophilus* IleRS. In particular, the hydrogen bonds of the N atom and O atom of the amino acid were perfectly recreated with the corresponding residues, D328 and T228, which are the key consensus residues of the editing domain in *T. thermophilus* IleRS. This result strongly suggests that our illustration for the amino acid binding mode is meaningful.

Docked energy and predicted free energy

The energies used in the docking studies include intermolecular and intramolecular interaction energies.[37] AUTODOCK reports the final docked energies and estimated free energies for each run according to its energy function for the final docked structures. The predicted free energy includes the intermolecular energy and torsional free energies and an empirically parameterized free energy.[37] In Table III, the docked energies and predicted free energies are reported for the five lowest-energy docked complexes. For WT *E. coli* LeuRS, the docking energies in POCKET 1 and POCKET 2 are similar to one another (average values of −7.61 and −7.42 kcal/mol, respectively) and also the predicted free energy differences are small (−5.61 and −5.67 kcal/mol, respectively). The same is true for WT *T. thermophilus* LeuRS. It was also found that there was no substantial energy difference between the WT and the T252A mutant, that is, the mutation of the conserved threonine (T252) to alanine does not cause a substantial environmental change in the pocket.

It is worth noting that there is a significant difference in average RMSD values for the docked leucines. The RMSD values of the five lowest-energy docked leucines were measured with the lowest-energy docked leucine of each pocket. In *E. coli* LeuRS, the average RMSD values for the T252A mutant are significantly smaller in POCKET 1, 0.18 Å, as compared to other cases (1.04–1.14 Å). Importantly, the mutation of T252 to alanine manipulates the surface of POCKET 1 for the ligand leucine to be more easily bound, as supported by the biochemical evidence reported by Mursinna et al.[14]

Surface rendering for the docked structure

To visualize the amino acid binding pocket of the CP1 domain, the molecular surface of the CP1 domain is shown in Figure 6. In Figure 6(A), two of the lowest-energy docked leucines in each pocket are rendered in a space-filling model and the five residues nearest the docked leucine are yellow. Interestingly, the surface of the protein has a reverse L-shaped channel between the two pockets that has been depicted in blue for clarity. In Figure 6(B), the docked leucines were removed and the whole structure was rotated around the Y-axis to show the relative positions of the highly conserved threonine (T252) along with T247; the two threonines are red [Fig. 6(B)]. The residue to the right is T247 and the one on the left is T252. Interestingly, T252 is located at the bottom of POCKET 1 while T247 is near the entrance. Significantly, this supports the hypothesis that T252 interacts with the side-chain of the incoming amino acid and plays an important role in amino acid discrimination.

Leu-AMP docking to the CP1 domain

There are two different reactions for editing, pre- and posttransfer.[7,17,47,48] In the pretransfer editing pathway, the aaRS directly hydrolyzes AA-AMP to AA + AMP.[7,8] If the amino acid moiety is transferred to the 3′-end of tRNA, the posttransfer editing hydrolyzes the AA-tRNA to produce AA and tRNA. This indicates that the CP1 editing

TABLE III. Docked Energy and Predicted Free Energy for the Five Lowest Docked Energy Structures

		POCKET 1			POCKET 2		
	No.	Docked energy (kcal/mol)	Predicted free energy (kcal/mol)	RMSD (Å)[a]	Docked energy (kcal/mol)	Predicted free energy (kcal/mol)	RMSD (Å)[a]
WT (*E. coli*) + Leu	1	−8.15	−6.13	0.00	−7.51	−5.85	0.00
	2	−7.85	−5.76	1.03	−7.49	−5.58	0.92
	3	−7.38	−5.39	1.14	−7.47	−5.52	1.06
	4	−7.35	−5.41	1.10	−7.37	−5.68	1.33
	5	−7.34	−5.35	1.12	−7.27	−5.71	1.24
Ave.		−7.61	−5.61	1.10	−7.42	−5.67	1.14
T252A (*E. coli*) + Leu	1	−7.43	−5.45	0.00	−7.66	−5.90	0.00
	2	−7.43	−5.45	0.13	−7.21	−5.29	0.98
	3	−7.43	−5.47	0.25	−7.20	−5.27	0.98
	4	−7.42	−5.43	0.18	−7.19	−5.27	1.01
	5	−7.42	−5.44	0.16	−7.18	−5.48	1.19
Ave.		−7.43	−5.45	0.18	−7.29	−5.44	1.04
WT (*T. thermophilus*) + Leu	1	−8.47	−6.61	0.00	−7.93	−6.00	0.00
	2	−8.44	−6.48	0.76	−7.91	−6.07	0.85
	3	−8.40	−6.43	0.83	−7.91	−5.97	0.24
	4	−8.40	−6.46	0.71	−7.87	−5.91	0.32
	5	−8.37	−6.68	0.56	−7.65	−5.89	1.17
Ave.		−8.42	−6.53	0.72	−7.85	−5.96	0.65

[a]RMSD values were measured with the best docked leucine of each pocket.

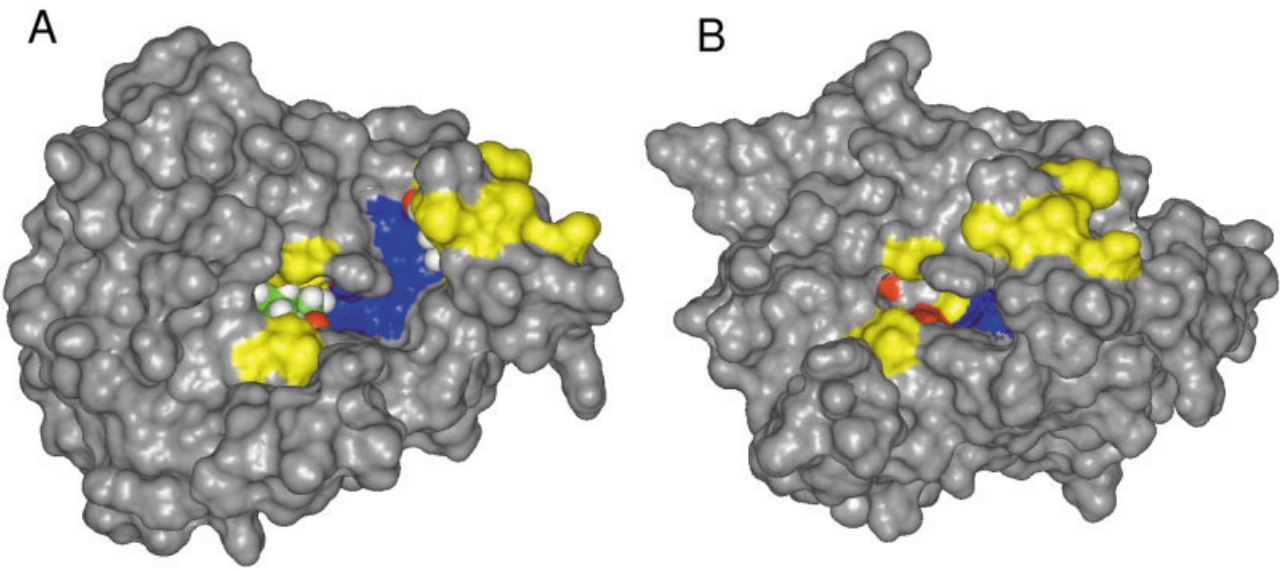

Fig. 6. Two amino acid binding pockets in the editing (CP1) domain of *E. coli* LeuRS. To visualize the two binding pockets, the molecular surface of the protein was rendered using Connolly algorithm and visualized in the bottom view of the structures in Fig. 2(B). Two low-energy docked leucine complexes, one for each pocket, were rendered by CPK in the both pockets (**A**). Yellow shows the closest residues to the docked ligands. The reverse L-shaped channel between the two pockets is blue. To show the relative positions of T247 and T252 inside the pocket, the ligand leucines were removed and the whole structure was slightly rotated around the Y-axis. The two threonines, T247 and T252, are red. The one on the right is T247 and the left is T252. While T247 lies in the entrance, the highly conserved T252 is located at the bottom of the POCKET 1.

domain must bind to both the AA and AMP during pretransfer editing as well as the AA attached to tRNA for the posttransfer editing reaction.

Although only posttransfer editing has been demonstrated for *E. coli* LeuRS,[14,49] it has been revealed that some eukaryotic LeuRSs have pretransfer editing capability.[49] To obtain some clues, in general, about the pretransfer editing capability of LeuRS, proposed binding sites of Leu-AMP in the editing domain were investigated by automated docking simulations with the pre- and posttransfer structures of AA-AMP. *E. coli* and *T. thermophilus* LeuRS were utilized as models although the former has only been demonstrated to have posttransfer editing activity.[14] From an evolutionary perspective, this approach would be valuable because many aaRSs have evolved their editing functions under external pressure. The results showed that there were no acceptable clusterings for Leu-AMP or other AA-AMP substrates (results not given),

that is, the final 50 docked structures were found in different conformations from each other and were also located in different positions, unlike the results for the leucine amino acids. The space between the CP1 domain and the main body of the protein in the current conformation may not be enough to accept AA-AMP because the size of AA-AMP is much larger than the AA itself.

Fukai et al. proposed that there were two different states of aaRSs in the editing reaction: an "open form" and a "closed form."[8] According to this definition, our protein conformation is in the closed form. In the open form, however, the CP1 domain moves away from the main body and may facilitate AA-AMP binding to the editing active site. In an attempt to emulate the open form, the LeuRS main body was removed to avoid the space deficiency problem observed in the docking simulations. Again, no acceptable clustering results were obtained from the docking studies. This shows that there is no acceptable region(s) that can accommodate the entire Leu-AMP structure in the CP1 domain, at least in the present conformation. Alternatively, it may indicate that our emulation of the open form emulation was not enough to be realistic. Finally, interactions may exist between the AMP and the main body of the protein while the AA portion is associated with the CP1 domain.

We also tried manual docking to bind the AA moiety of AA-AMP in one pocket and the AMP part in the other. From the results of automated docking studies, the distance between the two C_α atoms of the two lowest-energy docked leucines in both pockets is about 14 Å and the distance between the two C_α atoms of T248 and A293 is about 20 Å. The longest interatomic distance in a stretched conformation of Leu-AMP is about 15 Å and the distance between the C_α atom of the leucine and the center of the purine ring in AMP approximates 9–10 Å. The results showed that the distance between the both pockets was too far for them to fit together simultaneously (data not shown). A distance adjustment of 4–5 Å between the two pockets is required to achieve complete fitting.

These AA-AMP docking simulation results may suggest that the CP1 domain structures of *E. coli* and *T. thermophilus* LeuRS are not favorable for the pretransfer editing reaction or that the conformations of the CP1 domain of the LeuRSs should be substantially changed to accept Leu-AMP or an AA-AMP into a "pretransfer" editing site.

DISCUSSION

The automated docking simulations using only the amino acid as a ligand revealed that there are two possible amino acid binding pockets in the editing (CP1) domain of *E. coli* LeuRS. POCKET 2 is closer to the aminoacylation site in the main body of LeuRS and POCKET 1 includes T252 and resides near the central region of the CP1 domain. It is possible that one pocket is designed for pretransfer editing and the other is for posttransfer editing as Fukai et al. suggested based on the ValRS crystal structure.[8] Mursinna et al. previously demonstrated that the conserved threonine residue T252, located in POCKET 1, is critical to substrate specificity in posttransfer editing.[14] The equivalent T252A mutation in yeast cytoplasmic LeuRS, which possesses the pretransfer editing activity, did not appear to affect editing activity (T.L. Lincecum and S.A. Martinis, personal communication). This suggests that POCKET 1 may be exclusively designed for the pretransfer editing. To date, there are no experimental results suggesting that POCKET 2 is involved in pretransfer editing.

Alternatively, the mutation experiments for T252 in POCKET 1 and A293 in POCKET 2[14,50,51] reported that the editing activity of *E. coli* LeuRS was affected by mutations in either pocket. Therefore, we can propose that both pockets are related to the posttransfer editing reaction or that both pockets are the posttransfer editing active sites. Possibly, both pockets may work together or allosterically affect the other pocket. This idea is also consistent with the experimental mutation results because some specific changes, such as T252A and A293D, in either pocket affected the editing activity of the protein.[14,51]

Although it has been revealed that mutation of A293 in POCKET 2 affects the aminoacylation and editing activity of *E. coli* LeuRS,[50,51] it is still unclear if POCKET 2 is an editing active site. Chen et al. suggested that the mutation of A293 might cause a substantial change in the editing domain conformation. It can, therefore, be thought that the conformational change, which may be induced by the mutation of A293, would affect the function of POCKET 1 negatively. In this case, the function of POCKET 2 is unclear and requires more experimental and modeling studies to clarify.

In summary, some aaRSs discriminate their cognate amino acids against the others via proofreading and editing mechanisms. It has been reported that substituting a highly conserved threonine (T252) with an alanine within the editing domain of *E. coli* LeuRS caused the protein to hydrolytically cleave its cognate aminoacylated leucine from $tRNA^{Leu}$.[14] To achieve atomic-level insight into the role of T252 in LeuRS and the editing reaction of aaRSs, molecular modeling methods were applied to investigate amino acid binding sites in the editing domain of *E. coli* LeuRS.

First, because the X-ray or NMR structure of *E. coli* LeuRS is not available a 3D structure was constructed using the X-ray structure of *T. thermophilus* LeuRS as a template for homology modeling (Fig. 2). The RMSD for backbone atoms between the two structures was 1.18 Å overall and 0.60 Å for only the editing (CP1) domain. Second, automated docking studies were carried out to elucidate the amino acid binding site(s) in the CP1 domain of both the X-ray *T. thermophilus* and homology-modeled *E. coli* LeuRS structures The results indicate that there are two possible amino acid binding sites in the editing domain. In *E. coli* LeuRS structure, the first is located near the threonine-rich region comprised of residues T247, T248, T252, M336, and D342 (POCKET 1) and the second is in a flexible loop region surrounded by residues E292, A293, M295, A296, and M298 (POCKET 2). It was found that the highly conserved threonine (T252) is located at the bottom of the first binding pocket (POCKET 1).

These automated docking results are supported by the results of the mutation experiments for *E. coli* LeuRS.[14,51]

The T252A mutant lost its discriminating ability for its cognate leucine[14] and the A293D mutation caused the editing activity of LeuRS to decrease.[50] Therefore, based on the results of both computational and mutational analyses, we conclude that both pockets we found are highly correlated to the editing activity of *E. coli* LeuRS.

ACKNOWLEDGMENTS

The authors are grateful to Dr. S.A. Martinis, R.S. Mursinna, T.J. Lincecum Jr., and J. Speidel for helpful discussions and Dr. S.A. Martinis and R.S. Mursinna for careful reading of the article. The authors thank Dr. S. Cusack for providing early release of the crystal structure coordinates for *T. thermophilus* LeuRS and are also grateful to the Institute for Molecular Design at the University of Houston and Accelrys Inc. for making software available. This research is supported by Texas Advanced Research Program (003652-0016-1999) and the National Institutes of Health (GM63789).

REFERENCES

1. Carter CW Jr. Cognition, mechanism, and evolutionary relationships in aminoacyl-tRNA synthetases. Annu Rev Biochem 1993;62: 715–748.
2. Martinis SA, Schimmel P. In: Neidhardt FC, editor. *Escherichia coli* and *Salmonella* cellular and molecular biology. 2nd ed. Washington, DC: ASM Press; 1996. p 887–901
3. Giegé R, Sissler M, Florentz C. Universal rules and idiosyncratic features in tRNA identity. Nucleic Acids Res 1998;26:5017–5035.
4. Fersht AR, Kaethner MM. Mechanism of aminoacylation of tRNA. Proof of the aminoacyl adenylate pathway for the isoleucyl- and tyrosyl-tRNA synthetases from *Escherichia coli* K12. Biochemistry 1976;15:818–823.
5. Eldred EW, Schimmel PR. Rapid deacylation by iso-leucyl transfer ribonucleic acid synthetase of isoleucine-specific transfer ribonucleic acid aminoacylated with valine. J Biol Chem 1972;247:2961–2964.
6. Fersht AR, Kaethner MM. Enzyme hyperspecificity. Rejection of threonine by the valyl-tRNA synthetases by misacylation and hydrolytic editing. Biochemistry 1976;15:3342–3346.
7. Fersht AR. Editing mechanisms in protein synthesis. Rejection of valine by the isoleucyl-tRNA synthetase. Biochemistry 1977;16: 1025–1030.
8. Fukai S, Nureki O, Sekine S, Shimada A, Tao J, Vassylyev DG, Yokoyama S. Structural basis for double-sieve discrimination of L-valine from L-isoleucine and L-threonine by the complex of $tRNA^{Val}$ and valyl-tRNA synthetase. Cell 2000;103:793–803.
9. Freist W, Sternbach H, Cramer F. Threonyl-tRNA synthetase from yeast. Discrimination of 19 amino acids in aminoacylation of tRNA(Thr)-C-C-A and tRNA(Thr)-C-C-A(2′ NH_2). Eur J Biochem 1994;220:745–752.
10. Sankaranarayanan R, Dock-Bregeon AC, Romby P, Caillet J, Springer M, Rees B, Ehresmann C, Ehresmann B, Moras D. The structure of threonyl-tRNA synthetase-$tRNA^{Thr}$ complex enlightens its repressor activity and reveals an essential zinc ion in the active site. Cell 1999;97:371–381.
11. Beuning PJ, Musier-Forsyth K. Species-specific differences in amino acid editing by class II prolyl-tRNA synthetase. J Biol Chem 2001;276:30779–30785.
12. Cusack S, Yaremchuk A, Tukalo M. The 2 Å crystal structure of leucyl-tRNA synthetase and its complex with a leucyl-adenylate analogue. EMBO J 2000;19:2351–2361.
13. Fersht AR. Protein structure: Sieves in sequence. Science 1998;280: 541.
14. Mursinna RS, Lincecum TL, Martinis S. A conserved threonine within *Escherichia coli* leucyl-tRNA synthetase prevents hydrolytic editing of leucy-$tRNA^{Leu}$. Biochemistry 2001;40:5376–5381.
15. Nureki O, Vassylyev DG, Tateno M, Shimada A, Nakama T, Fukai S, Konno M, Hendrickson TL, Schimmel P, Yokoyama S. Enzyme structure with two catalytic sites for double-sieve selection of substrate. Science 1998;280:578–582.
16. Silivan LF, Wang J, Steitz TA. Insights into editing from an Ile-tRNA synthetase structure with $tRNA^{Ile}$ and mupirocin. Science 1999;285:1074–1077.
17. Schmidt E, Schimmel P. Residues in a class I tRNA synthetase which determine selectivity of amino acid recognition in the context of tRNA. Biochemistry 1995;34:11204–11210.
18. Hendrickson TL, Nomanbhoy TK, Schimmel P. Errors from selective disruption of the editing center in a tRNA synthetase. Biochemistry 2000;39:8180–8186.
19. Lin L, Hale SP, Schimmel P. Aminoacylation error correction. Nature 1996;384:33–34.
20. Moras D. Structural and functional relationships between aminoacyl-tRNA synthetases. Trends Biochem Sci 1992;17:159–164.
21. Chen JF, Guo NN, Li T, Wang ED, Wang YL. CP1 domain in *Escherichia coli* leucyl-tRNA synthetase is crucial for its editing function. Biochemistry 2000;39:6726–6732.
22. Martinis SA, Fox GE. Non-standard amino acid recognition by *Escherichia coli* leucyl-tRNA synthetase. Nucleic Acids Symp Ser 1997;36:125–128.
23. Haertlein M, Madern D. Molecular cloning and nucleotide sequence of the gene for *Escherichia coli* leucyl-tRNA synthetase. Nucleic Acids Res 1987;15:10199–10210.
24. Blattner FR, Plunkett G III, Bloch CA, Perna NT, Burland V, Riley M, Collado-Vides J, Glasner JD, Rode CK, Mayhew GF, Gregor J, Davis NW, Kirkpatrick HA, Goeden MA, Rose DJ, Mau B, Shao Y. The complete genome sequence of *Escherichia coli* K-12. Science 1997;277:1453–1474.
25. Oshima T, Aiba H, Baba T, Fujita K, Hayashi K, Honjo A, Ikemoto K, Inada T, Itoh T, Kajihara M, Kanai K, Kashimoto K, Kimura S, Kitagawa M, Makino K, Masuda S, Miki T, Mizobuchi K, Mori H, Motomura K, Nakamura Y, Nashimoto H, Nishio Y, Saito N, Sampei G, Seki Y, Tagami H, Takemoto K, Wada C, Yamamoto Y, Yano M, Horiuchi T. A 718-kb DNA sequence of the *Escherichia coli* K-12 genome corresponding to the 12.7–28.0 min region on the linkage map. DNA Res 1996;3:137–155.
26. InsightII, Version 2000. San Diego: Accelrys Inc. (http://www.accelrys.com); 2001.
27. Schuler GD, Altschul SF, Lipman DJ. A workbench for multiple alignment construction and analysis. Proteins 1991;9:180–190.
28. Sali A, Blundell TL. Comparative protein modeling by satisfaction of spatial restraints. J Mol Biol 1993;234:779–815.
29. Sali A, Matsumoto R, McNeil HP, Karplus M, Stevens RL. Three-dimensional models of four mouse mast cell chymases. Identification of proteoglycan-binding regions and protease-specific antigenic epitopes. J Biol Chem 1993;268:9023–9034.
30. Sali A, Overington JP. Derivation of rules for comparative protein modeling from a database of protein structure alignments. Protein Sci 1994;31:1582–1596.
31. Sali A, Pottertone L, Yuan F, van Vlijmen H, Karplus M. Evaluation of comparative protein modeling by MODELLER. Proteins 1995;23:318–326.
32. Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M. CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. J Comput Chem 1983;4: 187–195.
33. Braun W, Go N. Calculation of protein conformations by proton–proton distance constraints: A new efficient algorithm. J Mol Biol 1985;186:611–626.
34. Goodsell DS, Olson AJ. Automated docking of substrates to proteins by simulated annealing. Proteins 1990;8:195–202.
35. Goodsell DS, Morris GM, Olson AJ. Docking of flexible ligands: Applications of AutoDock. J Mol Recog 1996;9:1–5.
36. Morris GM, Goodsell DS, Huey R, Olson AJ. Distributed automated docking of flexible ligands to proteins: Parallel applications of AutoDock 2.4. J Comp Aided Mol Design 1996;10:293–304.
37. Morris GM, Goodsell DS, Halliday RS, Huey R, Hart WE, Belew RK, Olson AJ. Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. J Comput Chem 1998;19:1639–1662.
38. Holland JH. Adaptation in natural and artificial systems. Ann Arbor, MI: University of Michigan Press; 1975.
39. Četverikov SS. J Exp Biol 1926;2:3.
40. Solis FJ, Wets RJB. Math Oper Res 1981;6:19.
41. Rosin CD, Halliday RS, Hart WE, Belew RK. In: Baeck T, editor. Proceedings of the Seventh International Conference on Genetic Algorithms (ICGA97). San Francisco: Morgan Kauffman; 1997;221–229.
42. Webster T, Tsai H, Kula M, Mackie GA, Schimmel P. Specific sequence homology and three-dimensional structure of an aminoacyl transfer RNA synthetase. Science 1984;226:1315–1317.

43. Brick P, Blow DM. Crystal structure of a deletion mutant of a tyrosyl-tRNA synthetase complexed with tyrosine. J Mol Biol 1987;194:287–297.
44. Laskowski RA, MacArthur MW, Moss DS, Thornton JM. PROCHECK: a program to check the stereochemical quality of protein structures. J Appl Crystallogr 1993;26:283–291.
45. Morris AL, MacArthur MW, Hutchinson EG, Thornton JM. Stereochemical quality of protein structure coordinates. Proteins 1992; 12;345–364.
46. Eisenberg D, Luthy R, Bowie JU. VERIFY3D: assessment of protein models with three-dimensional profiles. Meth Enzymol 1997;277:396–404. (Server Site: http://www.doe-mbi.ucla.edu/Services/Verify_3D/.)
47. Hale SP, Schimmel P. Protein synthesis editing by a DNA aptamer. Proc Natl Acad Sci USA 1996;93:2755–2758.
48. Schmidt E, Schimmel P. Mutational isolation of a sieve for editing in a transfer RNA synthetase. Science 1994;285:1074–1077.
49. Englisch S, Englisch U, von der Haar F, Cramer F. The proofreading of hydroxy analogues of leucine and isoleucine by leucyl-tRNA synthetases from *E. coli* and yeast. Nucleic Acids Res 1986;14: 7529–7539.
50. Tong L, Guo N, Xia X, Wang ED, Wang YL. The peptide bond between E292-A293 of *Escherichia coli* leucyl-tRNA synthetase is essential for its activity. Biochemistry 1999;38:13063–13069.
51. Chen JF, Li T, Wang ED, Wang YL. Effect of alanine-293 replacement on the activity, ATP binding, and editing of *Escherichia coli* leucyl-tRNA synthetase. Biochemistry 2001;40:1144–1149.