Three Common Subunits in Editing Domains of Class Ia tRNA Synthetases

Keun Woo Lee,[†,*] Yong Jung Kwon,[‡] and James M. Briggs[§,*]

[†]*Department of Biochemistry, Division of Applied Life Science, Environmental Biotechnology National Core Research Center (EB-NCRC), Plant Molecular Biology and Biotechnology Research Center (PMBBRC), Gyeongsang National University, Jinju 660-701, Korea. [*]E-mail: kwlee@gnu.ac.kr*

[‡]*Department of Chemical Engineering, Kangwon National University, Chunchon 200-701, Korea*

[§]*Department of Biology and Biochemistry, University of Houston, Houston, TX 77204-5001, USA. [*]E-mail: jbriggs@uh.edu*

Received October 31, 2006

To identify structural or functional common subunit(s) in the CP1 (editing) domains of class Ia tRNA synthetases, five available structures were compared and analyzed. Through the sequence alignments and structural overlapping of the CP1 domains, three conserved regions were identified near the amino acid binding site in the editing domain. Structural overlapping of the three subunits clearly showed the existence of three common structural subunits in all of the five editing RS structures. Based on the established experimental results and our modeling results, it is proposed that subunits 1 and 3 accommodate the incoming amino acid binding, while subunit 2 contributes to the interactions with the adenosine ring of the A76 to stabilize the overall tRNA binding. Since these subunits are critical for the editing reaction, we expect that these key structures should be conserved through the most class Ia editing RSs.

Key Words : Aminoacyl-tRNA synthetase (aaRS), Editing domain, Aminoacylation, CP1 domain

Introduction

Correct aminoacylation of an amino acid to its cognate tRNA is critical for accurate protein synthesis. This very important reaction is controlled by a family of enzymes called aminoacyl-tRNA synthetases (aaRSs).[1-3] Among them, some class Ia RSs including leucyl-, valyl-, and isoleucyl-RSs (LeuRS, ValRS, and IleRS) have developed highly accurate molecular machinery to discriminate their cognate amino acids against structurally similar amino acids.[2,3] The real error rate with which IleRS distinguishes its cognate isoleucine from valine, differing by only one methyl group, was demonstrated to be fewer than 1/3000,[4] although previously Pauling predicted it to be 1/5 based on thermodynamic calculations.[5] This high accuracy is achieved by employing two separate active sites, namely, the activation (aminoacylation) site and the editing (proofreading) site.[6-8] The specificity of the amino acid activation and the editing activity of the editing RSs has been described as a double sieve.[9-11] The activation site (sieve 1) is in the ATP-binding Rossman fold that is common to all class I aaRSs while the editing active site (sieve 2) is located in a large inserted domain called connective polypeptide 1 (CP1).[12-16]

Structural and functional insights into the editing domain of the RSs were provided by available three-dimensional (3D) structures achieved by X-ray crystallography[12-17] and homology modeling.[18] In 1998, the first X-ray structure of the class Ia editing RS was reported by Nureki *et al.*[12] for the *T. thermophilus* IleRS, and then the tRNA complexed *S. aureus* IleRS structure followed from Silvian *et al.*[13] Later, the *T. thermophilus* ValRS structure was revealed by Fukai *et al.*[14,15] and Cusack *et al.*[16] solved *T. thermophilus* LeuRS structures with and without an activation substrate. More recently Fukunaga *et al.*[17,18] determined archaeal *P. horikoshii* LeuRS structure with and without tRNA structure and Lee *et al.*[19] built and refined an *E. coli* LeuRS structure *via* a homology modeling method using the *T. thermophilus* LeuRS X-ray structure[16] as a template. In parallel with the achievements in structural biology, a number of mutagenesis experiments were independently performed on the editing domains of *E. coli* LeuRS and IleRS.[20-24]

Considering *T. thermophilus* and *S. aureus* IleRS, even though they exist in different organisms, they share the same substrate, isoleucine, and the editing domain has the same function, *i.e.* distinguishing Ile from Leu and Val. Like this, although these class Ia type proteins live in different organisms, their substrate (*i.e.* amino acid) structures are universal in the entire life systems. We, therefore, propose that those proteins should have somewhat common structural features in the active sites because they have to bind to the common substrates, amino acids. The goal of this short study was to identify the common structural or functional elements in the editing active site over the class Ia editing RSs.

Methods

3D structure preparation. Currently six full domain 3D structures of the class Ia aaRSs were available for this study by either X-ray crystallography or homology modeling. Five of them are bacterial aaRSs and one belongs to archae. The protein data bank (PDB) codes for the five X-ray crystal structures are 1ILE for *T. thermophilus* IleRS, 1QU3 for *S. aureus* IleRS, 1GAX for *T. thermophilus* ValRS, 1H3N for *T. thermophilus* LeuRS, and 1WKB for *P. horikoshii* LeuRS. Five of the six 3D structures (*i.e.*, all of the bacterial aaRSs) were used for this investigation and the archaeon *P. horikoshii*

LeuRS, was used for validation.

Sequence and structure alignments for finding common subunits. The protein structure manipulations and analyses were performed using INSIGHTII program.[25] and also the five structures were superimposed by overlapping the homologous residues identified by a multiple sequence alignment scheme embedded in the HOMOLOGY module of the INSIGHTII program.[25] First, common amino acid sequences near the editing active sites were investigated to find the common sequence regions. Second, using the obtained common sequence region, many different combinations of structural overlapping were attempted to achieve the best overlap among the structures. The root mean square deviations (RMSD) of the alpha carbons were measured for the available combinations and the average RMSD value was used for comparison.

Structure comparison for pursuing the translocation of CP1 domain. In order to investigate the intriguing question of the reaction mechanism of these proteins, translocation of the CP1 domain resulting from the binding with the tRNA, two structures were compared using our new alignment approach. For the comparison, a pair of RS structures with and without tRNA is required. However, from the five available bacterial aaRS structures, the perfect pair is not yet available. Therefore, two X-ray structures of *T. thermophilus* IleRS and *S. aureus* IleRS were selected for the comparison since they share high levels of structural homology. For the two structures, the sequence similarities of the CP1 domains and the main bodies are *ca*. 68% and 62%, respectively.

Results and Discussion

Structure preparation and sequence alignments. Ribbon diagrams in two different views are shown for the five class Ia tRNA synthetase structures in Figure 1. The five structures show that they have a common, large inserted domain (CP1, colored in blue) with distances between the two active sites, one in the middle of main body and the other in the center of the CP1 domain, of about 30 Å.[12-19]

To identify common structural subunits in the CP1 domain, we focused on the residues located near the editing active site. Sequence alignment of the five CP1 domains showed that there exist three conserved sequences near the amino acid binding sites (Figure 2), which includes the highly conserved threonine rich region (subunit 1) and another established region where the universally conserved aspartic acid exists (subunit 3). The eighteen residues listed in Figure 2 were used for the structural alignments of the three subunits. The RMSD of the alpha carbons were measured for the available combinations and the average RMSD value for the ten measurements was 1.01 Å. The best result was obtained when only the five labeled residues were used for

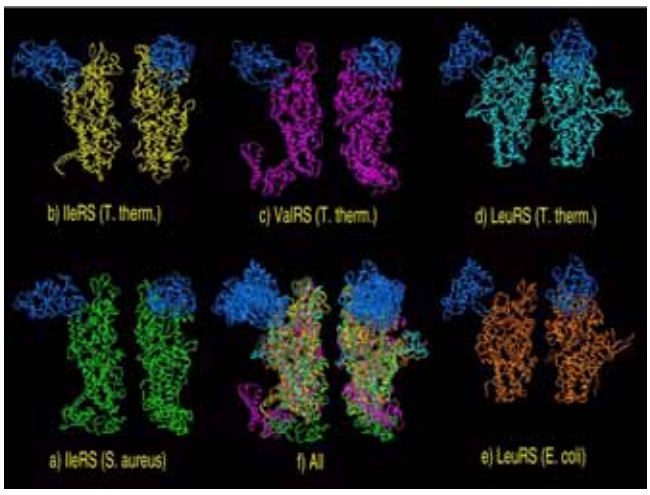

Figure 1. Ribbon diagrams in two different views are shown for the five class Ia tRNA synthetase (RS) structures: four X-ray structures (a-d) and one homology modeled structure (e). All editing domains (CP1) are highlighted in blue for clarity. All five structures were also superimposed (f) by overlapping the homologous residues identified by a multiple sequence alignment scheme embedded in the HOMOLOGY module of the INSIGHTII program.[25]

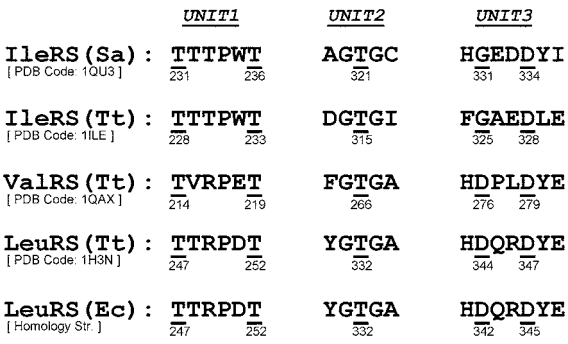

Figure 2. Sequence alignment for the CP1 domains of the five class Ia aaRSs showing the three common conserved sequences near the amino acid binding site. The eighteen residues in three units were used for structural alignment and the five key residues are labeled. Sa, Tt, and Ec in the parentheses represent for *S. aureus*, *T. thermophilus*, and *E. coli*, respectively.

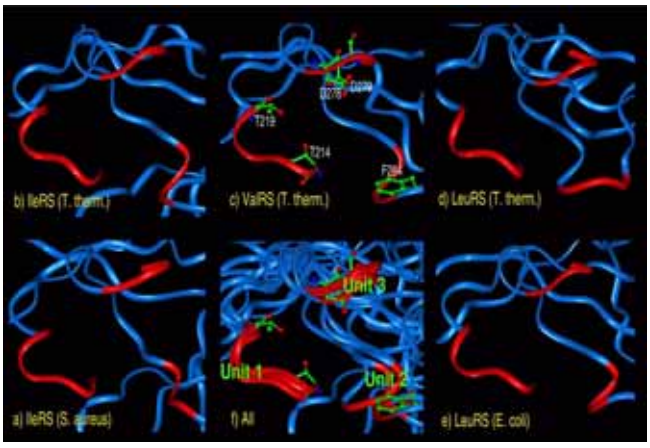

Figure 3. The three structural subunits in the editing domains of the five RSs. Ribbon diagrams of the CP1 domains were built for the four currently available X-ray structures (a-d) and a homology modeled structure (e) for comparison. The three common structural subunits are in red for clarity; all three units were superimposed together (f) using the five residues labeled in Figure 2.

the alignment, resulting in the average RMSD of 0.79 Å.

Three common structural subunits. Using the five residues we determined by sequence alignment and RMSD measurements, the five structures were superimposed (Figure 3). The figure clearly shows that the shapes of the three subunits are very similar and that the structures are well aligned (Figure 3f). In particular, the two established regions, subunits 1 and 3, are almost perfectly aligned through all five structures. For validation purposes, the archaeon *P. horikoshii* LeuRS structure was compared to the five template structures and the results showed that the protein also contains the same subunits with the same shape as the five bacterial structures (data not shown). The role of each subunit will be pursued and discussed in the next section.

In the end of introduction part, we proposed that there might exist some common structural subunits through the class Ia editing tRNA synthetase since their substrate structures were the exactly same. Our alignment result clearly shows that the idea was proved to be true (Figure 3). Based on our observations, we suggest that those three subunits would be conserved in the most class Ia editing RS structures since the three subunits were also found in the archaeal aaRSs as well as the bacterial class Ia aaRSs. Recently we observed the same structural features in the newly reported X-ray structure of *E. coli* LeuRS CP1 domain.[26] Although our idea was proved to make sense though the currently available six structures, we hope that more 3D structures of class Ia aaRSs are reported so that the idea can be confirmed in near future.

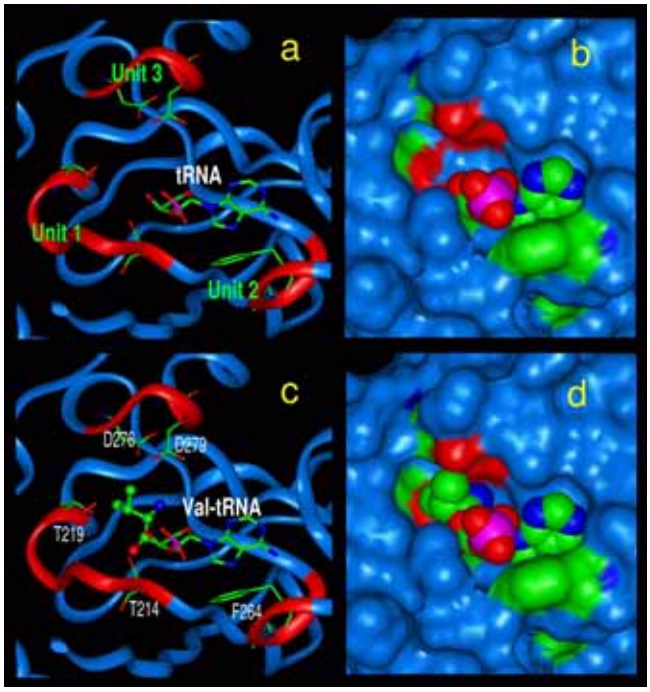

Figure 4. X-ray structure of the editing domain of *T. thermophilus* ValRS complexed with $tRNA^{Val}$. Only A76 from the entire tRNA is shown in the editing pocket for clarity. (a) A76 is connected to C75 (not shown) through the phosphorous atom in pink. The manually introduced valine residue is shown in ball-and-stick (c). The solvent accessible surfaces of the structures in panels a and c were rendered and displayed in panels b and d, respectively, to show the amino acid binding pockets. In panels a and c, the three common structural subunits are in red ribbon, and for all panels, the green, red, blue, and pink color represent carbon, oxygen, nitrogen, and phosphorus atoms, respectively.

Amino acid binding mode in editing site and the role of each subunit. The editing active site of *T. thermophilus* ValRS with a bound tRNA is shown in Figure 4. To demonstrate the amino acid binding mode, a valine residue was manually introduced to the O2' atom of tRNA using the available X-ray structure of the bound valine.[14] The results show that subunit 1 interacts with the side chain of the valine and D279 in subunit 3 forming a salt bridge with the alpha amino group of the valine. Mutagenesis experiments demonstrate that the threonines in subunit 1 are involved in the amino acid discrimination[20-22] and the aspartic acid in subunit 3 (D279, in this case) is essential for the editing activity.[23,24] From these established experimental results and our modeling results, we think that subunits 1 and 3 accommodate the amino acid binding. Subunit 1 interacts with the side chain of the ligand amino acid and plays a key role in amino acid recognition while subunit 3 interacts with the alpha amino group of the ligand and appears to be responsible for anchoring the amino acid by forming a salt bridge between the alpha amino group and the aspartic acid. We also expect that subunit 2 contributes to the interactions with the adenosine ring of the A76 to stabilize overall tRNA binding.

Translocation of the CP1 domain. Translocation of the CP1 domain, resulting from the binding with the tRNA, has been an intriguing question in the reaction mechanism of these proteins. Previously, for the comparison, the main bodies of the aaRS structures were superimposed and then the rotation angles of the CP1 domains were measured by Fukunaga and Yokoyama and they beautifully addressed this issue using the *P. horikoshii* LeuRS structure and available class Ia aaRS structures.[17] But our new alignment approach can provide a slightly different view in studying this interesting conformational transition. For perfect comparison, a pair of RS structures with and without the cognate tRNA is required. From the five available bacterial aaRS structures,

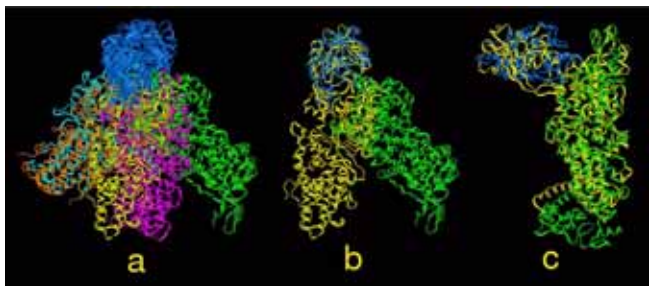

Figure 5. Overlapping of the editing RSs. The five editing RSs are superimposed (a) for the aligning using the three structural subunits in the CP1 domain (blue color). The coloring and structure IDs in panel a are the same in Figure 1. The *T. thermophilus* IleRS (yellow) and *S. aureus* IleRS's (green) are aligned by the three structural subunits (b) and main body (c). In b and c, the CP1 domain of the *T. thermophilus* IleRS is in yellow rather than in blue (a) for clarity.

two were solved with their cognate tRNAs, *T. thermophilus* ValRS (Figure 1c), *S. aureus* IleRS (Figure 1a), and the rest were without their tRNA partners. Unfortunately, the perfect pair is not available. However, the next best pair is that of *T. thermophilus* IleRS (Figure 1b) and *S. aureus* IleRS (Figure 1a) because they share high levels of structural homology. The overall RS structures showed huge distortions in their main bodies after alignment of the three structural subunits (Figure 5a). Although structural alignments achieved by making use of the entire structures (Figures 1f and 5c) did not exhibit any significant translocation movement of the CP1 domain, the newly aligned structures (*i.e. via* the three conserved functional subunits) clearly show the rotation of the CP1 domain with respect to the main bodies of the structures (Figure 5b). The results demonstrate that the CP1 domain may rotate clockwise by *ca.* 40-50 degrees after binding with tRNA.

Conclusion

In order to identify common structural or functional unit(s) in the editing domain of class Ia RSs, the editing active sites of the available RS structures were compared and analyzed. Through sequence alignments and the structural overlaps of the CP1 domain structures, three conserved regions were found near the editing active sites. The structural overlapping of those three subunits clearly showed that there exist three common structural subunits in the editing active sites in the five different CP1 structures. Based on our observation, we suggest the role of the subunits. Subunits 1 and 3 accommodate the amino acid binding. Subunit 1 interacts with the side chain of the ligand amino acid and therefore we think it plays a key role in amino acid recognition. However, since subunit 3 interacts with the alpha amino group of the ligand, it appears to be responsible for anchoring of the amino acid by forming a salt bridge between the alpha amino group and the aspartic acid. Finally, we expect that subunit 2 contributes to the interactions with the adenosine ring of A76 to stabilize overall tRNA binding. Since it seems that the three structural subunits are essential for the editing reaction it is expected that these key structures should be conserved through all class Ia editing RSs. Finally, we showed that our alignment strategy can also provide some clues for the translocation movement study of the CP1 domain caused by the binding with tRNA.

Acknowledgment. This work was also supported by grant from the MOST/KOSEF for the Environmental Biotechnology National Core Research Center (grant #: R15-2003-012-02001-0). This research was also supported by the Texas Advanced Research Program (003652-0016-1999) and the U.S. National Institutes of Health (GM63789).

References

1. Carter Jr., C. W. *Annu. Rev. Biochem.* **1993**, *62*, 715.
2. Martinis, S. A.; Schimmel, P. In *Escherichia coli and Salmonella Cellular and Molecular Biology*, 2nd ed.; Neidhardt, F. C., Ed.; ASM Press: Washington DC, 1996; pp 887.
3. Giegé, R.; Sissler, M.; Florentz, C. *Nucleic Acids Res.* **1998**, *26*, 5017.
4. Loftfield, R. B. *Biochem. J.* **1963**, *89*, 82.
5. Pauling, L. In *Festschrift für Prof. Dr. Arthur Stoll*; Birkhauser Verlag: Basel, 1958; p 597.
6. Eldred, E. W.; Schimmel, P. *J. Biol. Chem.* **1972**, *247*, 2961.
7. Schmidt, E.; Schimmel, P. *Science* **1994**, *285*, 1074.
8. Fersht, A. R. *Biochemistry* **1977**, *16*, 1025.
9. Fersht, A. R. In *Enzyme Structure and Mechanism*; Freeman: San Francisco, 1977; p 283.
10. Fersht, A. R.; Dingwall, C. *Biochemistry* **1979**, *18*, 1238.
11. Fersht, A. R. *Science* **1998**, *280*, 541.
12. Nureki, O.; Vassylyev, D. G.; Tateno, M.; Shimada, A.; Nakama, T.; Fukai, S.; Konno, M.; Hendrickson, T. L.; Schimmel, P.; Yokoyama, S. *Science* **1998**, *280*, 578.
13. Silvian, L. F.; Wang, J.; Steitz, T. A. *Science* **1999**, *285*, 1074.
14. Fukai, S.; Nureki, O.; Sekine, S.; Shimada, A.; Tao, J.; Vassylyev, D. G.; Yokoyama, S. *Cell* **2000**, *103*, 793.
15. Fukai, S.; Nureki, O.; Sekine, S.; Shimada, A.-I.; Vassylyev, D. G.; Yokoyama, S. *RNA* **2003**, *9*, 100.
16. Cusack, S.; Yaremchuk, A.; Tukalo, M. *EMBO J.* **2000**, *19*, 2351.
17. Fukunaga, R.; Yokoyama, S. *J. Mol. Biol.* **2005**, *346*, 57.
18. Fukunaga, R.; Yokoyama, S. *Nat. Struct. Mol. Biol.* **2005**, *12*(10), 915.
19. Lee, K. W.; Briggs, J. M. *PROTEINS: Struct. Funct. Bioinformatics* **2004**, *54*, 693.
20. Hendrickson, T. L.; Nomanbhoy, T. K.; Schimmel, P. *Biochemistry* **2000**, *39*, 8180.
21. Mursinna, R. S.; Lincecum, T. L.; Martinis, S. A. *Biochemistry* **2001**, *40*, 5376.
22. Mursinna, R. S.; Martinis, S. A. *J. Am. Chem. Soc.* **2002**, *124*, 7286.
23. Bishop, A. C.; Nomanbhoy, T. K.; Schimmel, P. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 585.
24. Lincecum Jr., T. L.; Tukalo, M.; Yaremchuk, A.; Mursinna, R. S.; Williams, A. M.; Sproat, B. S.; Van Den Eynde, W.; Link, A.; Van Calenbergh, S.; Grotli, M.; Martinis, S. A.; Cusack, S. *Mol. Cell.* **2003**, *11*, 951.
25. *InsightII*; Accelrys, Inc.: San Diego, (http://www.accelrys.com).
26. Liu, Y.; Liao, J.; Zhu, B.; Wang, E.-D.; Ding, J. *Biochem. J.* **2006**, *394*, 399.