Archaea recruited D-Tyr-tRNA^{Tyr} deacylase for editing in Thr-tRNA synthetase

DANIEL J. RIGDEN

School of Biological Sciences, University of Liverpool, Liverpool L69 7ZB, United Kingdom

ABSTRACT

Aminoacyl-tRNA synthetases (AARSs) are key players in the maintenance of the genetic code through correct pairing of amino acids with their cognate tRNA molecules. To this end, some AARSs, as well as seeking to recognize the correct amino acid during synthesis of aminoacyl-tRNA, enhance specificity through recognition of mischarged aminoacyl-tRNA molecules in a separate editing reaction. Recently, an editing domain, of uncertain provenance, idiosyncratic to some archaeal ThrRSs has been characterized. Here, sequence analyses and molecular modeling are reported that clearly show a relationship of the archaea-specific ThrRS editing domains with D-Tyr-tRNA^{Tyr} deacylases (DTDs). The model enables the identification of the catalytic site and other substrate binding residues, as well as the proposal of a likely catalytic mechanism. Interestingly, typical DTD sequences, common in bacteria and eukaryotes, are entirely absent in archaea, consistent with an evolutionary scheme in which DTD was co-opted to serve as a ThrRS editing domain in archaea soon after their divergence from eukaryotes. A group of present-day archaeobacteria contain a ThrRS obtained from a bacterium by horizontal gene transfer. In some of these cases a vestigial version of the original archaeal ThrRS, of potentially novel function, is maintained.

Keywords: tRNA synthetase; editing; homology modeling; fold recognition; evolution

INTRODUCTION

Aminoacyl-tRNA synthetases (AARSs; for reviews, see Ibba and Söll 2000, 2004; Woese et al. 2000) are responsible for matching amino acids to corresponding tRNAs containing the correct anticodon. As such they play a key role in interpretation of the genetic code, and a premium is placed on their correct matching of amino acid and tRNA. AARSs catalyze the transferase reaction between aminoacyladenylate and tRNA, yielding aminoacyl-tRNA and AMP. In some cases, the chemical characteristics of the amino acid in question are such that sufficient specificity over incorrect amino acids can be achieved solely at the step of covalent attachment of amino acid (delivered in the adenylated form) to the tRNA. Such is the case for cysteine (Fersht and Dingwall 1979; Newberry et al. 2002), whose corresponding AARS contains zinc, which interacts favorably and specifically with the thiol side chain of the amino acid. Nevertheless, in many other cases, amino acids are sufficiently similar in shape, size, and chemistry that selectivity cannot be

achieved in a single step. In these cases, additional editing reactions occur on enzymes with twin catalytic sites. The editing may occur either pre-transfer, with the hydrolysis of misactivated aminoacyladenylate on the tRNA-bound enzyme (Fersht 1977), or post-transfer, where the misaminoacylated tRNA is hydrolyzed (Eldred and Schimmel 1972).

The catalytic synthetase domains of AARSs may be grouped into two classes, I and II, which each contain the enzymes for 10 of the 20 standard amino acids (Eriani et al. 1990; Arnez and Moras 1997). The two classes of enzymes arose independently and have completely different structures, a Rossmann dinucleotide-binding domain in the case of class I (Rould et al. 1989) and an antiparallel β fold in class II (Cusack et al. 1990). Sequence analysis suggests that, in addition to these core catalytic domains, the synthetase sequences contain various extra domains which may be added to the N or C termini or inserted into the catalytic domain (Wolf et al. 1999). Phylogenetic analysis (Wolf et al. 1999; Woese et al. 2000) typically reveals complex and varied patterns of evolution in the different synthetase enzymes, with numerous examples of horizontal gene transfer.

Among the additional domains appended to or inserted into the catalytic domains are those responsible for the editing reactions. In class I, editing is carried out by an insertion into the catalytic domain called CP1 (Schmidt and

Reprint requests to: Daniel J. Rigden, School of Biological Sciences, University of Liverpool, Crown St., Liverpool L69 7ZB, United Kingdom; e-mail: drigden@liv.ac.uk; fax: (+44) 151 795 4414.

Article published online ahead of print. Article and publication date are at <http://www.rnajournal.org/cgi/doi/10.1261/rna.7115404>.

Schimmel 1994). Understanding of editing in class II enzymes has lagged behind, although two types of editing domain have already been characterized, one in ProRS (Wong et al. 2002) and the other, though differently positioned with respect to the catalytic domain, in both AlaRS (Dock-Bregeon et al. 2000) and ThrRS (Beebe et al. 2004). Freestanding editing domains, both those naturally present (Ahel et al. 2003) and those derived artificially from multidomain proteins (Wong et al. 2003), can trans-edit mischarged tRNAs. ThrRSs discriminate effectively against the mischarging with isosteric valine using a bound zinc (Sankaranarayanan et al. 2000), but mischarge with serine at a rate high enough to imply the existence of an editing step (Dock-Bregeon et al. 2000). Interestingly, sequence comparisons show that most archaeal ThrRS sequences do not possess an editing domain of a well-studied type, instead having an alternative, N-terminal domain with no obvious sequence similarity to other proteins (Woese et al. 2000; Beebe et al. 2004). Confirmation that this domain functions as an editing domain has recently been obtained (Beebe et al. 2004). The presence of an idiosyncratic editing domain in archaeal ThrRSs is but one more way in which archaeal aminoacyl-tRNA synthesis differs from that of bacteria and eukaryotes (Praetorius-Ibba and Ibba 2003).

Here we show, by sequence analysis and model building, that archaeal ThrRS editing domains are distantly homologous to D-Tyr-tRNA^{Tyr} deacylases (DTDs), widespread enzymes which function to prevent the misincorporation of D-amino acids into proteins (Soutourina et al. 1999, 2000; Lim et al. 2003). Thus, the archaeal ThrRS editing domain is not a distant homolog of a classical AARS class II editing domain (Dock-Bregeon et al. 2000; Wong et al. 2002; Beebe et al. 2004), but instead represents a novel protein architecture here associated with aminoacyl-tRNA editing. The catalytic site can be located on the model structure, a mechanism proposed, and conserved residues likely involved in binding tRNA substrate located. Patterns of phylogenetic distribution suggest that archaea co-opted the DTD domain for editing purposes soon after their divergence from eukaryotes.

RESULTS AND DISCUSSION

Fold recognition

Simple sequence searches with BLAST and PSI-BLAST did not reveal any more distant homologs outside the archaeobacteria-specific family of ThrRS N-terminal editing domains. Thus doubt remained as to whether they might represent highly divergent examples of known editing domains or, alternatively, a novel protein fold for aminoacyl-tRNA editing (Beebe et al. 2004). The *Methanosarcina mazei* editing domain (residues 1–142) was therefore submitted for fold recognition at the Meta server (Bujnicki et al. 2001). From the alignment of 18 editing domains it was clear that

the C-terminal limit of this domain was at around residue 142, the linker region between this and the transferase domain being highly variable in both length and composition. The results from the various independent methods and their consensus analyses immediately showed a structural relationship between the editing domain and the enzyme D-Tyr-tRNA^{Tyr} deacylase (DTD). Two structures for the latter enzyme are known, from *Escherichia coli* (Ferri-Fioni et al. 2001) and *Hemophilus influenzae* (Lim et al. 2003). These DTD structures were top scoring by the FFAS03 (Rychlewski et al. 2000), 3D-PSSM (Kelley et al. 2000), and Fugue (Shi et al. 2001) methods and third in the Bioinbgu (Fischer 2000) results. All consensus calculations therefore strongly favored the DTD fold for the *M. mazei* editing domain sequence with 3D-Jury (Ginalski et al. 2003), Shotgun on 3 (Fischer 2003), and Pcons2 (Lundstrom et al. 2001) methods producing scores of 70, 38, and 1.55, respectively. Using the results of the latest completed Livebench benchmarking effort (Rychlewski et al. 2003), these results are highly significant. Currently the Shotgun on 3 method best distinguishes between true and false positive fold assignments. With respect to the score achieved by DTD structures for the editing domain, based on the range of difficulty among Livebench submissions, only one in every 16 scores at this level (38 by Shotgun on 3) represents a false positive, the rest being correct fold assignments. Although pairwise sequence identity between representative editing domains and DTDs was very low (7%–18% for the alignment used to produce the models; Fig. 1) a good match was observed between the predicted secondary structure for the editing domains and the observed secondary structure for the DTDs (Fig. 1), supporting the proposed structural correspondence. The exact match between the N-terminal limits of the editing domains and most DTDs was also corroborating evidence since the N-terminal Met residue is buried in the DTD structures. (Some DTD sequences in the database, including the *Azotobacter vinelandii* sequence shown in Fig. 1, have N-terminal extensions but these appear to be the results of incorrect starting codon choice.) Finally, the obvious functional similarity between the two families, both aminoacyl-tRNA deacylases, provided further support for the existence of a distantly homologous relationship leading to structural resemblance.

Model construction

The ability to construct a model that performs well against standard protein structure validation tools is strong evidence for the correctness of a proposed structural correspondence. Modeling of the the *M. mazei* ThrRS editing domain was therefore initiated based on its FFAS03 profile matching alignment with the two template DTD structures. Since both DTDs are dimeric structures with the catalytic sites lying between subunits, model dimers of *M. mazei* ThrRS editing domain were built. There was no inconsis-

FIGURE 1. Sequence alignment of representative archaeal ThrRS editing domains (*above*) and DTDs (*below*). Within each group good conservation is indicated by bold face and absolute conservation by bold italics. Shading marks conservation between the two groups. The numbering of *Methanosarcina mazei* ThrRS is shown *above* the alignment along with its predicted secondary structure (arrows for β -strands and cylinders for α -helices). Numbering and actual secondary structure of *Hemophilus influenzae* DTD (Lim et al. 2003) (PDB code 1j7g) are shown *beneath* the alignment. The figure was made with ALSCRIPT (Barton 1993).

tency with the known quaternary structures of ThrRS which are dimers (Woese et al. 2000). Later evidence independently supported the dimeric structure for the editing domain (see below).

The alignment of editing domains and DTDs implied two large deletions in the former with respect to the known structures. Reassuringly, these both corresponded to extended loops whose loss would have no consequences for core protein fold structure. During the iterative modeling process the main regions of probable inaccuracy, as defined by positive PROSA II profiles or negative VERIFY_3D profiles, were insertions compared to the template structures. Better performing loops in these regions were located during the modeling process and incorporated into later models. A suitable structure for the largest 11-residue insertion was not forthcoming by this standard methodology. Instead, 10 models for the region were more rigorously calculated at the MODLOOP server (Fiser et al. 2000; Fiser and Sali 2003). The best of these, as assessed by profile, stereochemical, and compactness analysis, was included in later models. Outside of the insertions, poor profiles were also obtained at the very C terminus of the model due to the burial from solvent of the Lys119 side chain in initial models. Improved models were obtained after a shift of the editing domain, with respect to the DTDs, of two residues toward the N terminus. After alignment improvement and satisfactory modeling of insertions, three residues remained in disallowed regions of the Ramachandran plot obtained for the model with the best PROSA II score. Two of these

were dealt with by flipping of peptide bonds and the third by local backbone regeneration. In the final round of model generation, one of the 20 models had simultaneously the best PROSA II score (-8.13), the best VERIFY_3D score (104), and the greatest percentage (91.1) of non-Pro, non-Gly residues in the most favored region of the Ramachandran plot and had no disallowed residues. This was taken as the final model.

Model analysis

The interface between the subunits of the final model dimer is strongly hydrophobic, supporting the notion that the archaeobacteria-specific editing domains exist as dimer. No insertions or deletions relative to templates were located at the interface. Further support for the dimeric model comes from mapping of sequence conservation among the full set of editing domains onto the model structure (Fig. 2A). The strongly conserved patch, corresponding to the presumed catalytic site in the DTDs (Ferri-Fioni et al. 2001; Lim et al. 2003), lies between both subunits.

The conservation of this region, combined with the hypothesis of distant homology (rather than structural analogy) between editing domains and DTDs, enables its ready identification as the catalytic site of the editing domains. The catalytic mechanism of DTDs has not yet been the subject of experimental study, although modeling has been employed to propose a plausible hypothesis (Lim et al. 2003). It has been proposed that the side chain of conserved

FIGURE 2. Two views of the final *Methanosarcina mazei* ThrRS model dimer structure in the same orientation. (A) Semitransparent model molecular surface colored by sequence conservation among the archaeal-specific ThrRS editing domains from orange (strongly conserved) to blue (not conserved). (B) Cartoon diagram showing predicted catalytic residues (magenta) and conserved residues predicted to make electrostatic (cyan) or hydrophobic (gray) interactions with substrate tRNA. One subunit is colored according to secondary structure, the other colored uniformly gold. Figures were made with PY-MOL (<http://www.pymol.org>).

Thr80 (numbering of *H. influenzae* DTD; PDB code 1j7g; see Fig. 1) carries out nucleophilic attack on the carbonyl carbon atom of the D-amino acid attached to the tRNA. The amino group of the D-amino acid is proposed to act as general base while an oxyanion hole, comprising the backbone amide nitrogen atoms of Phe79 and Thr80 plus the side chain nitrogen of Gln78, stabilizes the tetrahedral transition states (Lim et al. 2003). Aligned with Thr80 in the editing domains are 16 serines, one threonine, and one alanine (Fig. 1). With the exception of the latter, in the sequence from *Methanopyrus kandleri*, these are suitable for functional substitution of Thr80 as nucleophile. In the *M. mazei* editing domain the presumed nucleophile is therefore Ser84.

Using the described docking of D-Tyr into *H. influenzae* DTD (Lim et al. 2003), L-Ser, the amino acid whose mischarging on tRNA^{Thr} would be corrected by the ThrRS editing domain, was positioned in the editing domain model by an overlay of N, C α , and carbonyl C atoms. In the

DTDs, specificity for D- over L-amino acids was explained through the steric clashes that would occur on binding of the latter, principally with the conserved oxyanion hole residue Gln78 (Lim et al. 2003). This residue is replaced by an absolutely conserved, similarly sized residue, His82 (*M. mazei* ThrRS numbering; Fig. 1), in the editing domains. However, immediately below the His a conserved Leu81 is replaced by a much smaller residue, Ser85, in 16 out of 18 editing domain sequences. Although the present model is not reliable enough for detailed modeling, the resulting cavity could enable a conformational change of His82, allowing the small side chain of the L-Ser component of substrate to bind to the editing domain. After such a change, His82 would be well placed to hydrogen bond to the L-Ser substrate side chain. In the model, an insertion partially disrupts the proposed oxyanion hole of the DTD structure. Nevertheless, the backbone nitrogen of the nucleophilic residue remains suitably positioned to stabilize a tetrahedral intermediate and a new interaction is predicted with the side chain of invariant Tyr78.

As well as the catalytic machinery and the L-Ser binding site, the editing domain must be able to bind the tRNA portion of the substrate. In the DTDs this is predicted to involve a combination of basic residues for binding the phosphate groups of the RNA backbone and hydrophobic groups to interact with exposed bases (Lim et al. 2003). Considering only strongly conserved positions, similar possibilities can be observed in the editing domain model (Fig. 2B). Basic residues which might interact with the tRNA are His7, Lys16, Lys71, Lys119, and His128. Exposed, conserved hydrophobic residues are Leu83, Trp117, Tyr118, Ala120, and Leu130. Note that, as with the DTDs (Lim et al. 2003), both classes of interaction involve both subunits. Two of these residues, His128 and Leu130, are found within the region from 125 to 135 that was replaced wholesale with the sequence AAASAAASAAA, leading to loss of editing activity (Beebe et al. 2004).

Evolutionary considerations

Aminoacyl-tRNA synthetases have complicated evolutionary histories (Wolf et al. 1999; Woese et al. 2000), and archaeal ThrRSs are no exception. Beebe et al. (2004) define two groups of archaea, one with the archaeal type N-terminal editing domain and the other with a typical bacterial/eukaryotic editing domain, along with *Sulfolobus tokodaii*, which is reported to have an N-terminal sequence resembling neither. In fact, searches reveal that *S. tokodaii*, along with *S. solfataricus* and *Aeropyrum pernix*, classed in the archaeal editing domain group, have two sequences with homology to ThrRSs (as already noted; Woese et al. 2000). In each case the smaller homolog has the archaeal-specific domain. The larger homolog has the typical bacterial/eukaryotic editing domain in the two *Sulfolobus* species while the larger of the two *A. pernix* sequences apparently lacks an

N-terminal editing domain. The smaller homologs are only half the size of the larger sequences, through lacking the central transferase catalytic domain. All three organisms with two homologous ThrRS-like sequences are crenarchaeota, although another member of the group, *Pyrobaculum aerophilum* has a single ThrRS with an archaeal-type N-terminal editing domain. Of the present-day archaea with single ThrRS homologs, in *Halobacterium*, *Thermoplasma acidophilum*, *Thermoplasma volcanium*, and *Ferroplasma acidarmanus*, the sequences contain a bacterial/eukaryotic-type editing domain while the archaeal-specific editing domain is found in *M. mazei* (Beebe et al. 2004) and all other archaea not mentioned above.

Interestingly, the phyletic distribution of DTDs seems entirely bacterial/eukaryotic and hence complementary to the archaeal ThrRS-specific editing domains. There is a single DTD-like sequence in the unfinished genome sequence of *Methanosarcina barkeri* (accession NZ_AAAR01000871.1) but it bears 100% identity with a sequence from *Magnetococcus* sp. MC-1 (accession ZP_00043084.1) and therefore presumably represents a bacterial contaminant. One possible evolutionary scenario would therefore be as shown in Figure 3. According to this hypothesis, the common ancestor of DTDs and the archaea-specific editing domain was present before the divergence of the lineages that then gave rise to archaea and eukaryotes.

FIGURE 3. A possible evolutionary scenario that explains present-day distributions of ThrRS and DTD-related domains in bacteria (left), archaeobacteria (center, three different situations labeled with example organisms), and eukaryotes (right). Domains are color coded as follows: red, DTD; green, archaeal-specific ThrRS editing domain (Beebe et al. 2004); magenta, homolog of archaeal-specific ThrRS editing domain, function unknown, in the shorter ThrRS homologous sequences that lack catalytic transferase domains; blue, standard bacterial/eukaryotic class II AARS editing domain (Dock-Bregeon et al. 2000; Beebe et al. 2004). Straight lines represent the remainders of full-length ThrRS sequences with presumed transferase activity. Wavy lines represent the remainders of the shorter ThrRS homologous sequences, lacking transferase domains, that have unknown function. The dotted arrow represents a horizontal gene transfer. Encircled numbers mark key evolutionary events, as explained in the text. Due to its uncertain position in phylogenetic trees (data not shown) *Aeropyrum pernix* is not included in this scheme.

The predecessor enzyme would have been a DTD since that is the present-day activity of the domain in both prokaryotes (Soutourina et al. 1999) and eukaryotes (Soutourina et al. 2000). ThrRSs from bacteria and eukaryotes share the AlaRS/ThrRS class II editing domain (Dock-Bregeon et al. 2000; Beebe et al. 2004) which was presumably also present, therefore, in the earliest ThrRS of the archaeobacterial lineage. After the split giving rise to the progenitors of the archaeal and eukaryotic lineages, acquisition of editing activity would have occurred by the DTD domain and the displacement of the original editing domain occurred (step 1 in Fig. 3). This would have occurred soon after the split since the archaeal-specific editing domain is found in all three archaeal phyla represented in the databases—Euryarchaeota, Crenarchaeota, and Nanoarchaeota. According to the hypothesis, some archaeal organisms would later have acquired a second ThrRS with the bacterial/eukaryotic-type editing domain by horizontal gene transfer (step 2 in Fig. 3), a phenomenon particularly common among AARSs (Wolf et al. 1999; Woese et al. 2000). The catalytic domains of the present-day archaeal ThrRS sequences containing the bacterial/eukaryotic-type editing domains are clearly more closely related to bacterial and eukaryotic enzymes, particularly bacterial enzymes, than they are to the archaeal sequences with archaeal-specific editing domains (Woese et al. 2000; data not shown). This supports the notion of acquisition of a whole second AARS sequence (step 2 in Fig. 3), probably from a bacterium, in contrast to the proposed displacement of an archaeal editing domain by one of the bacterial/eukaryotic type (Beebe et al. 2004). The presence of a second functional ThrRS would have enabled the loss of the central portion of the original archaeal ThrRS along with loss of its transferase activity (step 3 in Fig. 3). The shorter homologs may retain editing activity (Ahel et al. 2003; Wong et al. 2003) but would also thereafter have been free to acquire novel function(s) as observed for many AARS-related sequences (Schimmel and De Pouplana 2000). As mentioned, three extant organisms still have two ThrRS homologs (although the shorter sequences lack the catalytic transferase domain) while others have a single sequence that possesses a bacterial/eukaryotic-type editing domain, these having presumably lost the ancestral archaeal sequence (step 4 in Fig. 3). The scheme shown in Figure 3 is consistent with the hypothesis that AARS editing domains currently found in multidomain contexts may have originated in autonomous editing domains (Ahel et al. 2003).

The widespread presence of DTD proteins across Nature has naturally led to speculation as to how certain groups of organisms in which the enzyme is not present—archaeobacteria, parasitic auxotrophic bacteria, and prototrophic cyanobacteria—cope with the possible dangers of incorporation of D-amino acids into proteins (Soutourina et al. 2000). D-amino acids are produced deliberately in some circumstances—for antibiotic synthesis in bacteria, for example (Friedman 1999)—and through side effects of some en-

zymes, for example, tryptophan synthase (Miles et al. 1986), but their misincorporation into standard proteins is detrimental to the health of the organism (Soutourina et al. 1999, 2000). It is possible that the presence of fewer metabolic pathways in the auxotrophic bacteria may lead to lower intrinsic production of D-amino acids (Soutourina et al. 2000). In archaeobacteria and cyanobacteria it has been proposed that the aminoacyl-tRNA synthetases have evolved improved specificity for L-amino acids over D-amino acids (Soutourina et al. 2000). The distant evolutionary relationship demonstrated here between archaeal-specific editing domains and DTDs raises the additional possibility that D-aminoacyl tRNA deacylase activity may be retained by the editing domains, although not all present-day archaeae have them.

CONCLUSIONS

By sequence analysis and fold recognition, the possibility that archaeal ThrRS-specific editing domains are divergent versions of other AARS editing domains has been ruled out. Instead, they are clearly distant relatives of DTDs and therefore represent a novel protein fold to be associated with AARS editing. Analysis of the final protein model enables the identification of likely catalytic and binding residues and the proposal of a catalytic mechanism, similar to that suggested for DTDs. The recruitment of the DTD domain for its new editing function in ThrRSs seems to have occurred shortly after the divergence of the progenitors of the archaeal and eukaryotic lineages, explaining the lack of typical DTD sequences in present-day archaea. Whether the archaeal ThrRS editing domains retain some DTD activity, or whether archaea have adapted to the possible dangers of incorporation of D-amino acids into proteins in other ways, remains to be determined.

MATERIALS AND METHODS

Sequence analyses

Sequences homologous to *M. mazei* ThrRS were sought in the GenPept and Unfinished Microbial Genome Databases at NCBI using BLAST and PSI-BLAST (Altschul et al. 1997) and the resulting sequence set was aligned using MUSCLE (Edgar 2004). A similar alignment was made for the DTDs. Jalview (Clamp et al. 2004) was used for alignment visualization and for determination of five maximally disparate representatives of each family. ESPRIPT (Gouet et al. 1999) was used for mapping sequence conservation onto structures. The N-terminal editing domain of *M. mazei* ThrRS was analyzed by fold recognition experiments at the Meta server (Bujnicki et al. 2001). This is a portal to the leading fold recognition methods, among them FFAS03 (Rychlewski et al. 2000), 3D-PSSM (Kelley et al. 2000), GenTHREADER (Jones 1999a), Bioinbgu (Fischer 2000), and Fugue (Shi et al. 2001), additionally providing consensus predictions of improved reliability (Lundstrom et al. 2001; Fischer 2003; Ginalski et

al. 2003). These methods utilize sensitive sequence comparisons and, typically, analysis of inferred characteristics such as predicted secondary structure, in order to analyze compatibility of known folds with a given sequence. Secondary structure predictions were made using PSI-PRED (Jones 1999b).

Structural modeling

Based on the fold recognition results, modeling of the N-terminal domain of *M. mazei* ThrRS was carried out using the structures of DTDs from *E. coli* (Ferri-Fioni et al. 2001) (PDB code 1jke) and *H. influenzae* (Lim et al. 2003) (PDB code 1j7g) as templates. Modeling was done with MODELLER 6 (Sali and Blundell 1993). Default regimes of model refinement by energy minimization and simulated annealing were employed. Because of the low sequence similarity between target and template, a rigorous iterative modeling protocol was adopted in which 20 models were constructed and analyzed for each alignment variant. These models were analyzed for packing and solvent exposure characteristics using PROSA II (Sippl 1993) and VERIFY_3D (Luthy et al. 1992), and for stereochemical properties using PROCHECK (Laskowski et al. 1993). Possible misalignments were highlighted as regions that were positive in the PROSA II profiles and/or negative in the VERIFY_3D profiles. Variant alignments were tested for these regions by iterative model building and analysis. When no further improvements could be achieved the model with the highest PROSA II score was taken as the final model. Homodimer models were constructed throughout using restraints to maintain symmetry. Possible conformations for a large, 11-residue insertion in the target, relative to the templates, were obtained from the MODLOOP server (Fiser et al. 2000; Fiser and Sali 2003). The most compact loop structure with good stereochemistry was chosen for inclusion in the final model. MODELLER was also used to calculate percentage sequence identities, using the formula number of identities divided by the length of the shorter sequence (May 2004). Protein structures were superimposed using LSQMAN (Kleywegt 1996) and visualized using O (Jones et al. 1991). Diagrammatic representations of the structures were generated using PyMOL (<http://www.pymol.org>), the dss command of which was used to determine regular secondary structure elements in the final model.

Received June 24, 2004; accepted September 6, 2004.

REFERENCES

- Ahel, I., Korencic, D., Ibba, M., and Söll, D. 2003. Trans-editing of mischarged tRNAs. *Proc. Natl. Acad. Sci.* **100**: 15422–15427.
- Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. 1997. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res.* **25**: 3389–3402.
- Arnez, J.G. and Moras, D. 1997. Structural and functional considerations of the aminoacylation reaction. *Trends Biochem. Sci.* **22**: 211–216.
- Barton, G.J. 1993. ALSCRIPT: A tool to format multiple sequence alignments. *Protein Eng.* **6**: 37–40.
- Beebe, K., Merriman, E., De Pouplana, L.R., and Schimmel, P. 2004. A domain for editing by an archaeobacterial tRNA synthetase. *Proc. Natl. Acad. Sci.* **101**: 5958–5963.
- Bujnicki, J.M., Elofsson, A., Fischer, D., and Rychlewski, L. 2001. Structure prediction meta server. *Bioinformatics* **17**: 750–751.

- Clamp, M., Cuff, J., Searle, S.M., and Barton, G.J. 2004. The Jalview Java alignment editor. *Bioinformatics* **20**: 426–427.
- Cusack, S., Berthet-Colominas, C., Hartlein, M., Nassar, N., and Leberman, R. 1990. A second class of synthetase structure revealed by X-ray analysis of *Escherichia coli* seryl-tRNA synthetase at 2.5 Å. *Nature* **347**: 249–255.
- Dock-Bregeon, A., Sankaranarayanan, R., Romby, P., Caillet, J., Springer, M., Rees, B., Francklyn, C.S., Ehresmann, C., and Moras, D. 2000. Transfer RNA-mediated editing in threonyl-tRNA synthetase. The class II solution to the double discrimination problem. *Cell* **103**: 877–884.
- Edgar, R.C. 2004. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**: 1792–1797.
- Eldred, E.W. and Schimmel, P.R. 1972. Rapid deacylation by isoleucyl transfer ribonucleic acid synthetase of isoleucine-specific transfer ribonucleic acid aminoacylated with valine. *J. Biol. Chem.* **247**: 2961–2964.
- Eriani, G., Delarue, M., Poch, O., Gangloff, J., and Moras, D. 1990. Partition of tRNA synthetases into two classes based on mutually exclusive sets of sequence motifs. *Nature* **347**: 203–206.
- Ferri-Fioni, M.L., Schmitt, E., Soutourina, J., Plateau, P., Mechulam, Y., and Blanquet, S. 2001. Structure of crystalline D-Tyr-tRNA(Tyr) deacylase. A representative of a new class of tRNA-dependent hydrolases. *J. Biol. Chem.* **276**: 47285–47290.
- Fersht, A.R. 1977. Editing mechanisms in protein synthesis. Rejection of valine by the isoleucyl-tRNA synthetase. *Biochemistry* **16**: 1025–1030.
- Fersht, A.R. and Dingwall, C. 1979. Cysteiny-tRNA synthetase from *Escherichia coli* does not need an editing mechanism to reject serine and alanine. High binding energy of small groups in specific molecular interactions. *Biochemistry* **18**: 1245–1249.
- Fischer, D. 2000. Hybrid fold recognition: Combining sequence derived properties with evolutionary information. *Pac. Symp. Biocomput.* 119–130.
- . 2003. 3D-SHOTGUN: A novel, cooperative, fold-recognition meta-predictor. *Proteins* **51**: 434–441.
- Fiser, A. and Sali, A. 2003. ModLoop: Automated modeling of loops in protein structures. *Bioinformatics* **19**: 2500–2501.
- Fiser, A., Do, R.K., and Sali, A. 2000. Modeling of loops in protein structures. *Protein Sci.* **9**: 1753–1773.
- Friedman, M. 1999. Chemistry, nutrition, and microbiology of D-amino acids. *J. Agric. Food Chem.* **47**: 3457–3479.
- Ginalski, K., Elofsson, A., Fischer, D., and Rychlewski, L. 2003. 3D-Jury: A simple approach to improve protein structure predictions. *Bioinformatics* **19**: 1015–1018.
- Gouet, P., Courcelle, E., Stuart, D.I., and Metoz, F. 1999. ESPript: Analysis of multiple sequence alignments in PostScript. *Bioinformatics* **15**: 305–308.
- Ibba, M. and Söll, D. 2000. Aminoacyl-tRNA synthesis. *Annu. Rev. Biochem.* **69**: 617–650.
- . 2004. Aminoacyl-tRNAs: Setting the limits of the genetic code. *Genes & Dev.* **18**: 731–738.
- Jones, D.T. 1999a. GenTHREADER: An efficient and reliable protein fold recognition method for genomic sequences. *J. Mol. Biol.* **287**: 797–815.
- . 1999b. Protein secondary structure prediction based on position-specific scoring matrices. *J. Mol. Biol.* **292**: 195–202.
- Jones, T.A., Zou, J.Y., Cowan, S.W., and Kjeldgaard, M. 1991. Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A* **47**: 110–119.
- Kelley, L.A., MacCallum, R.M., and Sternberg, M.J. 2000. Enhanced genome annotation using structural profiles in the program 3D-PSSM. *J. Mol. Biol.* **299**: 499–520.
- Kleywegt, G.J. 1996. Use of non-crystallographic symmetry in protein structure refinement. *Acta Crystallogr. D* **52**: 842–857.
- Laskowski, R., MacArthur, M., Moss, D., and Thornton, J. 1993. PROCHECK: A program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26**: 283–290.
- Lim, K., Tempczyk, A., Bonander, N., Toedt, J., Howard, A., Eisenstein, E., and Herzberg, O. 2003. A catalytic mechanism for D-Tyr-tRNA^{Tyr} deacylase based on the crystal structure of *Hemophilus influenzae* HI0670. *J. Biol. Chem.* **278**: 13496–13502.
- Lundstrom, J., Rychlewski, L., Bujnicki, J., and Elofsson, A. 2001. Pcons: A neural-network-based consensus predictor that improves fold recognition. *Protein Sci.* **10**: 2354–2362.
- Luthy, R., Bowie, J.U., and Eisenberg, D. 1992. Assessment of protein models with three-dimensional profiles. *Nature* **356**: 83–85.
- May, A.C. 2004. Percent sequence identity; the need to be explicit. *Structure* **12**: 737–738.
- Miles, E.W., Phillips, R.S., Yeh, H.J., and Cohen, L.A. 1986. Isomerization of (3S)-2,3-dihydro-5-fluoro-L-tryptophan and of 5-fluoro-L-tryptophan catalyzed by tryptophan synthase: Studies using fluorine-19 nuclear magnetic resonance and difference spectroscopy. *Biochemistry* **25**: 4240–4249.
- Newberry, K.J., Hou, Y.M., and Perona, J.J. 2002. Structural origins of amino acid selection without editing by cysteinyl-tRNA synthetase. *EMBO J.* **21**: 2778–2787.
- Praetorius-Ibba, M. and Ibba, M. 2003. Aminoacyl-tRNA synthesis in archaea: Different but not unique. *Mol. Microbiol.* **48**: 631–637.
- Rould, M.A., Perona, J.J., Söll, D., and Steitz, T.A. 1989. Structure of *E. coli* glutamyl-tRNA synthetase complexed with tRNA(Gln) and ATP at 2.8 Å resolution. *Science* **246**: 1135–1142.
- Rychlewski, L., Jaroszewski, L., Li, W., and Godzik, A. 2000. Comparison of sequence profiles. Strategies for structural predictions using sequence information. *Protein Sci.* **9**: 232–241.
- Rychlewski, L., Fischer, D., and Elofsson, A. 2003. LiveBench-6: Large-scale automated evaluation of protein structure prediction servers. *Proteins* **53**, Suppl. 6: 542–547.
- Sali, A. and Blundell, T.L. 1993. Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **234**: 779–815.
- Sankaranarayanan, R., Dock-Bregeon, A.C., Rees, B., Bovee, M., Caillet, J., Romby, P., Francklyn, C.S., and Moras, D. 2000. Zinc ion mediated amino acid discrimination by threonyl-tRNA synthetase. *Nat. Struct. Biol.* **7**: 461–465.
- Schimmel, P. and Ribas De Pouplana, L. 2000. Footprints of aminoacyl-tRNA synthetases are everywhere. *Trends Biochem. Sci.* **25**: 207–209.
- Schmidt, E. and Schimmel, P. 1994. Mutational isolation of a sieve for editing in a transfer RNA synthetase. *Science* **264**: 265–267.
- Shi, J., Blundell, T.L., and Mizuguchi, K. 2001. FUGUE: Sequence-structure homology recognition using environment-specific substitution tables and structure-dependent gap penalties. *J. Mol. Biol.* **310**: 243–257.
- Sippl, M.J. 1993. Recognition of errors in three-dimensional structures of proteins. *Proteins* **17**: 355–362.
- Soutourina, J., Plateau, P., Delort, F., Peirottes, A., and Blanquet, S. 1999. Functional characterization of the D-Tyr-tRNA^{Tyr} deacylase from *Escherichia coli*. *J. Biol. Chem.* **274**: 19109–19114.
- Soutourina, J., Blanquet, S., and Plateau, P. 2000. D-tyrosyl-tRNA(Tyr) metabolism in *Saccharomyces cerevisiae*. *J. Biol. Chem.* **275**: 11626–11630.
- Woese, C.R., Olsen, G.J., Ibba, M., and Söll, D. 2000. Aminoacyl-tRNA synthetases, the genetic code, and the evolutionary process. *Microbiol. Mol. Biol. Rev.* **64**: 202–236.
- Wolf, Y.I., Aravind, L., Grishin, N.V., and Koonin, E.V. 1999. Evolution of aminoacyl-tRNA synthetases—Analysis of unique domain architectures and phylogenetic trees reveals a complex history of horizontal gene transfer events. *Genome Res.* **9**: 689–710.
- Wong, F.C., Beuning, P.J., Nagan, M., Shiba, K., and Musier-Forsyth, K. 2002. Functional role of the prokaryotic proline-tRNA synthetase insertion domain in amino acid editing. *Biochemistry* **41**: 7108–7115.
- Wong, F.C., Beuning, P.J., Silvers, C., and Musier-Forsyth, K. 2003. An isolated class II aminoacyl-tRNA synthetase insertion domain is functional in amino acid editing. *J. Biol. Chem.* **278**: 52857–52864.

RNA

A PUBLICATION OF THE RNA SOCIETY

Archaea recruited d-Tyr-tRNA^{Tyr} deacylase for editing in Thr-tRNA synthetase

DANIEL J. RIGDEN

RNA 2004 10: 1845-1851

References

This article cites 49 articles, 14 of which can be accessed free at:
<http://rnajournal.cshlp.org/content/10/12/1845.full.html#ref-list-1>

License

Email Alerting Service

Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).

To subscribe to *RNA* go to:
<http://rnajournal.cshlp.org/subscriptions>