Published online June 23, 2006

3246–3258 *Nucleic Acids Research, 2006, Vol. 34, No. 11*
doi:10.1093/nar/gkl441

Structure of human tryptophanyl-tRNA synthetase in complex with $tRNA^{Trp}$ reveals the molecular basis of tRNA recognition and specificity

Ning Shen[1,2], Litao Guo[1,2], Bei Yang[1,2], Youxin Jin[1] and Jianping Ding[1,*]

[1]State Key Laboratory of Molecular Biology, Institute of Biochemistry and Cell Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences and [2]Graduate School of Chinese Academy of Sciences, 320 Yue-Yang Road, Shanghai 200031, China

Received May 8, 2006; Revised and Accepted June 8, 2006

ABSTRACT

Aminoacyl-tRNA synthetases (aaRSs) are a family of enzymes responsible for the covalent link of amino acids to their cognate tRNAs. The selectivity and species-specificity in the recognitions of both amino acid and tRNA by aaRSs play a vital role in maintaining the fidelity of protein synthesis. We report here the first crystal structure of human tryptophanyl-tRNA synthetase (hTrpRS) in complex with $tRNA^{Trp}$ and Trp which, together with biochemical data, reveals the molecular basis of a novel tRNA binding and recognition mechanism. hTrpRS recognizes the tRNA acceptor arm from the major groove; however, the 3′ end CCA of the tRNA makes a sharp turn to bind at the active site with a deformed conformation. The discriminator base A73 is specifically recognized by an α-helix of the unique N-terminal domain and the anticodon loop by an α-helix insertion of the C-terminal domain. The N-terminal domain appears to be involved in Trp activation, but not essential for tRNA binding and acylation. Structural and sequence comparisons suggest that this novel tRNA binding and recognition mechanism is very likely shared by other archaeal and eukaryotic TrpRSs, but not by bacterial TrpRSs. Our findings provide insights into the molecular basis of tRNA specificity and species-specificity.

INTRODUCTION

In the translation process of protein synthesis, tRNAs recognize a trinucleotide codon on mRNA through a complementary anticodon and deliver the covalently attached amino acids to ribosome where the protein synthesis takes place. Aminoacyl-tRNA synthetases (aaRSs) are the enzymes responsible for the covalent link of amino acids to the 3′ ends of their cognate tRNAs via aminoacylation reaction [see reviews (1,2)]. Since the structure and function of a protein depend critically on its primary structure, or amino acid sequence, errors in protein synthesis usually lead to serious consequences. The fidelity of the translation process during protein synthesis to a large extent depends on the esterification of the correct amino acids to their cognate tRNAs by aaRSs. Therefore, the selectivity and specificity in the recognitions of both amino acid and tRNA by aaRS play a vital role in maintaining the fidelity of protein synthesis.

The aaRS family in general contains 20 enzymes and each enzyme is specific for one of the 20 standard amino acids. The aminoacylation reaction is carried out in two steps: the amino acid is first activated by ATP to form an aminoacyl-AMP, which is then transferred to the 3′ end of the cognate tRNA to form an aminoacyl-tRNA. Although all aaRSs perform the same biological function, they differ substantially in sequence and structure, and thus, are divided into two classes, class I and class II (2–5). Class I aaRSs contain two signature motifs (the HIGH and KMSKS motifs) and share a common catalytic domain with a Rossmann fold (RF) composed of alternative β-strands and α-helices. Class II aaRSs are characterized by three highly conserved sequence motifs (motifs 1, 2 and 3) and their catalytic domain is consisted of a seven-stranded β-sheet flanked by two conserved α-helices. This partition is further correlated with the type of interaction between the enzyme and the tRNA: class I aaRSs usually approach their tRNAs from the minor groove of the acceptor arm, whereas class II aaRSs interact with their tRNAs via the major groove. Moreover, class I enzymes charge their amino acids on the 2′-OH group of the 3′ terminal adenosine, while class II enzymes on the 3′-OH group. Based on similarities in

*To whom correspondence should be addressed. Tel: 86 21 54921619; Fax: 86 21 54921116; Email: jpding@sibs.ac.cn
*Correspondence may also be addressed to Youxin Jin. Tel: 86 21 54921222; Fax: 86 21 54921011; Email: yxjin@sibs.ac.cn

The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors
Protein Data Bank accession codes 2AKE and 2DR2

sequence, structure and biochemical and biophysical properties each class can be further grouped into subclasses. Subclass Ic contains two closely related enzymes, tryptophanyl-tRNA synthetase (TrpRS) and tyrosyl-tRNA synthetase (TyrRS), which share a low sequence identity (about 10–25%) between their eukaryotic and bacterial orthologs and between each other, but a high structure homology in the catalytic domain (6–12). Human TrpRS (hTrpRS) and human TyrRS (hTyrRS) are of particular interest because, compared to their prokaryotic orthologs, hTrpRS has an extra N-terminal domain of about 150 residues and hTyrRS has an extra C-terminal domain of about 170 residues, and the acquisitions of these domains have added additional functions to these enzymes. Alternative splicing or proteolysis of the full-length enzymes yields several short forms that exhibit angiogenic (mini-TyrRS) or antiangiogenic activity (mini-, T1- and T2-hTrpRS), linking protein synthesis to signal transduction pathways (13–15). In addition, biochemical data have shown that the N-terminal domain of hTrpRS plays a critical role in the aminoacylation reaction (13,15).

So far, substantial work has been carried out to elucidate the structures and functions of TrpRSs and TyrRSs, which has revealed insights into the catalytic mechanism of the aminoacylation reaction. The crystal structures of TrpRSs (human and *Bacillus stearothermophilus*) and TyrRSs (human, *B.stearothermophilus* and *Escherichia coli*) in both unliganded form and in complexes with substrates have been determined (6–12,16,17). Recently, the crystal structures of *Thermus thermophilus* TyrRS (tTyrRS) and *Methanococcus jannaschii* TyrRS (mTyrRS) in complexes with their cognate tRNAs were reported, which reveal the binding and recognition mechanisms of $tRNA^{Tyr}$ by bacterial and archaeal TyrRSs (18,19). However, to date, no structure of TrpRS in complex with $tRNA^{Trp}$ has been reported and the recognition mechanism of $tRNA^{Trp}$ by TrpRS especially in eukaryotic organisms is not well understood. Because of the low sequence homology among TrpRSs from different species and between TrpRSs and TyrRSs, there are substantial structural differences even though they share a similar global architecture in the catalytic domain. These differences presumably constitute the basis of the orthogonal tRNA specificities of TrpRSs from different species and the basis of tRNA specificities of TrpRSs and TyrRSs. In these regards, it is of important biological significance to understand how hTrpRS recognizes its cognate substrates (Trp and $tRNA^{Trp}$) and whether the N-terminal domain of hTrpRS plays any functional role in the recognition of its substrates.

We report here the crystal structure of hTrpRS in complex with bovine $tRNA^{Trp}$ and tryptophan (Trp) at 3.0 Å resolution, which is the first TrpRS–$tRNA^{Trp}$ complex structure and the first mammalian aaRS–tRNA complex structure. Analysis of the structure reveals a novel tRNA binding and recognition mechanism by hTrpRS. Similar to TyrRSs, hTrpRS interacts with the tRNA acceptor arm from the major groove which is a typical binding mode of class II aaRSs; however, the 3′ end CCA of the tRNA is bound at the catalytic active site with a deformed conformation, which is a typical binding mode of class I aaRSs. The $tRNA^{Trp}$ acceptor arm is primarily recognized by an α-helix of the N-terminal domain and has no base-specific interaction with the catalytic domain. Recognition of the anticodon is carried out by an α-helix insertion of the C-terminal domain. Sequence and structural comparisons of this tRNA complex with other TrpRSs indicate that this novel tRNA recognition mechanism is very likely shared by other archaeal and eukaryotic TrpRSs, but different from that utilized by bacterial TrpRSs. These diverse recognition mechanisms appear to form the basis of tRNA specificity and species-specificity.

MATERIALS AND METHODS

Expression and purification of hTrpRS and $tRNA^{Trp}$

Expression and purification of the wild-type hTrpRS were carried out as described previously (12). Expression plasmids of the mutant hTrpRSs were generated from the vector containing the wild-type enzyme using the QuikChange Site-Directed Mutagenesis Kit (Stratagene) and the sequences of these genes were verified by DNA sequencing. Mutant hTrpRS proteins were expressed and purified following the procedures used for the wild-type enzyme. Due to the difficulty in over-expressing and purifying large quantity of human $tRNA^{Trp}$ transcript, bovine $tRNA^{Trp}$ transcript was used in the structural studies instead. The bovine $tRNA^{Trp}$ gene was cloned into the pGEM-9zf(−) vector. Expression of the bovine $tRNA^{Trp}$ in *E.coli* and isolation of the total tRNAs were carried out as described previously (20,21).

Crystallization, X-ray diffraction data collection and structure determination

Crystallization of hTrpRS in complex with bovine $tRNA^{Trp}$ was performed by the hanging drop vapor diffusion method at 4°C, using 2 μl of the protein-tRNA solution [about 8 mg/ml of hTrpRS with an equal mole of tRNA, 10 mM ATP, 1 mM Trp, 20 mM K_2HPO_4 (pH 6.8), 10 mM $MgCl_2$, 5 mM 2-mercaptoethanol and 0.5 mM phenylmethylsulfonyl fluoride (PMSF)] and 2 μl of the reservoir solution [2 M $(NH_4)_2SO_4$ and 50 mM HEPES (pH 7.0)]. Crystals of the hTrpRS–$tRNA^{Trp}$ complex grew to approximate dimensions of $0.15 \times 0.15 \times 0.2$ mm^3 in about 1–2 months. Diffraction data were collected from a flash-cooled crystal at 100 K at Cornell High Energy Synchrotron Source F1 beamline. Data processing and scaling were performed using the HKL2000 suite (22). Crystals of the tRNA complex belong to tetragonal space group $P4_322$ with unit cell dimensions of $a = b = 132.3$ Å and $c = 137.0$ Å. The structure of the hTrpRS–$tRNA^{Trp}$ complex was solved by molecular replacement method using the unliganded hTrpRS structure (12) as the search model. In the initial difference Fourier maps there was evident electron density corresponding to the bound tRNA. The initial $tRNA^{Trp}$ model was built based on real-space electron density fitting of the $tRNA^{Asp}$ model in complex with yeast aspartyl-tRNA synthetase (23), which was subjected to further model building and structure refinement. Although both ATP and Trp were present in the crystallization solution, there was only electron density corresponding to Trp at the catalytic active site, but no electron density for the complete activation product Trp-AMP and/or degradation products, (such as AMP or ATP).

Table 1. X-ray diffraction data and structure refinement statistics

Statistics of diffraction data	
Resolution range (Å)[a]	50.0–3.00 (3.05–3.00)
No. of observed reflections	359 396
No. of unique reflections (I/σ > 0)	25 011
Redundancy	14.4 (14.7)
I/σ(I)	36.9 (7.5)
Completeness (%)	100 (100.0)
R_{merge} (%)[b]	13.2 (73.4)
Statistics of refinement and model	
Resolution range (Å)	50.0–3.00 (3.19–3.00)
No. of reflections [Fo ⩾ 0σ(Fo)]	24 957
Completeness (%)	99.8 (99.6)
R-factor (%)[c]	23.1 (27.5)
Free *R*-factor (%)	28.7 (28.4)
No. of protein atoms	3016
No. of tRNA atoms	1602
No. of Trp molecule	1
RMS bond lengths (Å)	0.008
RMS bond angles (°)	1.2
Luzzati atomic positional error (Å)	0.33
Ramachandran plot (%)	
Most favored regions	90.1
Allowed regions	9.6
Generously allowed regions	0.3

[a]Numbers in parentheses refer to the highest resolution shell.
[b]$R_{merge} = ||I_0| - <I>|/<I>$.
[c]*R*-factor $= ||F_o| - |F_c||/|F_o|$.

Therefore, the electron density at the active site was interpreted as the Trp substrate in the current structure. Structure refinement was carried out using the program CNS (24). Model building was facilitated using the program O (25). The statistics of diffraction data and structure refinement are summarized in Table 1.

Aminoacylation activity assay

The aminoacylation activity of both wild-type and mutant hTrpRSs was assayed at 30°C as described previously (26). The reaction buffer consisted of 4 mM ATP, 0.8 mM DTT, 1.0 μCi of [^{3}H]-Trp, 8 mM $MgCl_2$, 80 mM Tris–HCl (pH 7.5) and 0.02 μM bovine $tRNA^{Trp}$ in a total volume of 50 μl. The reaction was stopped after 30 min by quenching on ice and 20 μl aliquots of the reactant were spotted on to Whatmann filter discs (3 mm) for measurement of radioactive counts by scintillation. One unit of the aminoacylation activity was defined as the amount of enzyme needed to charge 1 pmol of tRNA per minute under the assay condition. For all kinetic assays, the concentration of tRNA varied from 0.4 μm to 12.8 μm with the concentration doubled each point and the concentration of Trp was kept at constant of 12 μm. All experiments were repeated at least four times under the same condition. Kinetic parameters k_{cat} and K_M were calculated from Eadie–Hofstee plots. All data were fitted to the Michaelis–Menten equation by non-linear regression.

RESULTS AND DISCUSSION

Overall structure

The crystal structure of hTrpRS in complex with bovine $tRNA^{Trp}$ and Trp was solved using molecular replacement method and the refinement of the final atomic model converged to an *R*-factor of 23.1% and a free *R*-factor of 28.7% to 3.0 Å resolution (Table 1). The asymmetric unit contains one hTrpRS monomer, one $tRNA^{Trp}$ molecule and one Trp molecule with a solvent content of 70%. Figure 1A shows the overall structure of the hTrpRS–$tRNA^{Trp}$–Trp complex. The N-terminal 96 residues of hTrpRS could not be seen in the structure even though the full-length enzyme was used in crystallization. SDS–PAGE analyses of both the dissolved crystals and the crystallization solution showed that the N-terminal part of the enzyme had been hydrolyzed during crystallization and the hTrpRS molecule (residues 97–471) in this complex was equivalent to T2-hTrpRS (Supplementary Figure S1), similar to that in the ligand-free hTrpRS structure (12) (hereafter the nomenclature of the secondary structure of hTrpRS is after Yu *et al*.). Further experiments showed that hydrolysis of the full-length hTrpRS is caused by trace of bacterial serine protease(s) in the protein sample and addition of a protease inhibitor PMSF in the purification and preparation of protein sample prohibits the degradation of the enzyme (data not shown). Bovine $tRNA^{Trp}$ shares very high sequence homology with human $tRNA^{Trp}$ (differing by two bases, U15 versus A15 in the D loop and G57 versus A57 in the variable region) and can be efficiently aminoacylated by hTrpRS (20,27). Therefore, bovine $tRNA^{Trp}$ can be considered as a cognate species for hTrpRS although its post-transcriptional modification pattern could be different because of over-expression in *E.coli*. The anticodon loop of the tRNA was well defined in the electron density maps and the acceptor arm (including the 3′ terminal CCA) was fairly ordered (Figure 1B). However, most of the D loop and the variable region, both of which have no contact with the protein and are exposed to the solvent, were disordered with weak electron density; thus, the model in these regions is tentative. Due to the moderate resolution of the structure, we were unable to identify any post-transcriptional modification on tRNA. Though both Trp and ATP were added in the crystallization solution, only Trp was found to bind at the catalytic active site with well defined electron density; there was no electron density corresponding to the activation product Trp-AMP or degradation products AMP or ATP.

The hTrpRS molecule in this structure consists of three domains: the C-terminal part of the N-terminal domain (residues 97–153, designated as the N-terminal fragment hereafter), the catalytic domain (residues 154–362 and 453–471), and the C-terminal α-helical domain (residues 363–452) (Figure 1A). The N-terminal fragment flanks one side of the catalytic domain and is involved in interaction with the 3′ end of the tRNA acceptor arm. The catalytic domain has a typical RF and contains a connective polypeptide 1 (CP1) insertion. The CP1 (helices α5 and α6) is involved in dimer interface interaction and plays important roles in forming the substrate-binding pocket and binding the tRNA acceptor arm. The C-terminal α-helical domain connects to the catalytic domain via the conserved KMSAS loop and interacts with the tRNA anticodon. Compared to *B.stearothermophilus* TrpRS (bTrpRS), the N-terminal domain of hTrpRS is an extra domain (Figure 1C). In addition, hTrpRS has an 8 residue insertion (384–391) which forms a short α-helix (α11) on the tip of the helical domain.

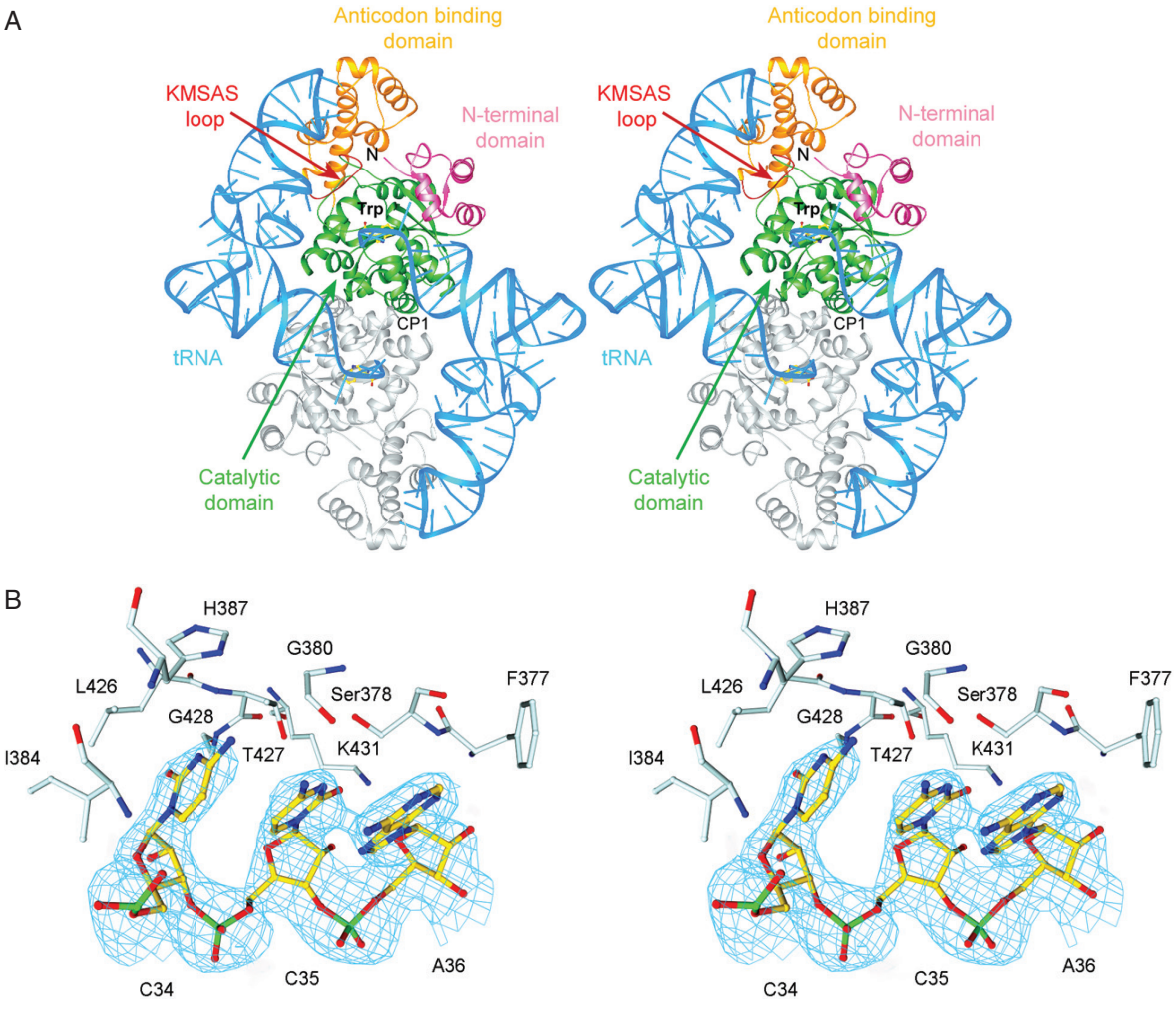
In the ligand-free hTrpRS structure, this α-helix is exposed to the solvent and has a relatively flexible conformation; however, in this structure it has a well ordered conformation and is involved in binding the tRNA anticodon loop. This region is also shown to be crucial for the angiostatic activity of hTrpRS (17).

As expected, hTrpRS forms a homodimer in the tRNA complex with its 2-fold axis coincident with the crystallographic 2-fold symmetry. Two $tRNA^{Trp}$ molecules bind to the hTrpRS homodimer in a symmetrical fashion and each $tRNA^{Trp}$ binds across the two subunits of the homodimer with the acceptor arm binding at the catalytic active site of one subunit and the anticodon loop at the α-helical domain of the other (Figure 1A). This tRNA binding mode is similar to that found in the bacterial and archaeal TyrRS–$tRNA^{Tyr}$ complexes (18,19). The specific recognition of a tRNA by its cognate aaRS is dictated by interactions between the protein and the identity elements of the tRNA, specifically the acceptor arm and the anticodon loop. Indeed, in this TrpRS–$tRNA^{Trp}$ complex, hTrpRS interacts mainly with the acceptor arm and the anticodon loop of the $tRNA^{Trp}$.

Acceptor arm recognition

hTrpRS interacts with the $tRNA^{Trp}$ acceptor arm from the major groove (Figure 2A), which is the characteristic tRNA binding mode of class II aaRSs. This is different from most other class I aaRSs, which approach the acceptor arm of their tRNAs from the minor groove, [such as GlnRS (28,29)], but again similar to that found in the bacterial and archaeal TyrRS–$tRNA^{Tyr}$ complexes (18,19). On the other hand, the binding of the 3′ end of the tRNA at the catalytic active site is more similar to the class I aaRSs rather than the class II aaRSs. The A73 base is twisted away from the tRNA helical axis, and the 3′ end CCA make a sharp turn to intrude into the catalytic active site with a deformed conformation (Figure 2A).

The tRNA acceptor arm has interactions with helix α1′ (100–108) of the N-terminal fragment and helices α6 (259–265) and α9 (321–329) of the catalytic domain (Figure 2 and Supplementary Figure S2). In particular, helix α1′ acts as a platform at the entrance to the catalytic active site to hold the tRNA 3′ end and keep the acceptor arm in a distance from the protein. As a result, the interactions of the first 5 bp of the tRNA with the protein (including Thr259, Asn261, Arg321 and Arg326) involve mainly the sugar-phosphate backbone of the nucleotides and are not base-specific (Supplementary Figure S2). On the other hand, the discriminator base A73 has interactions with three residues (Asp99, Lys102 and Arg106) of helix α1′ (Figure 2B and C). Although, the side-chains of Asp99 and Lys102 were poorly



C

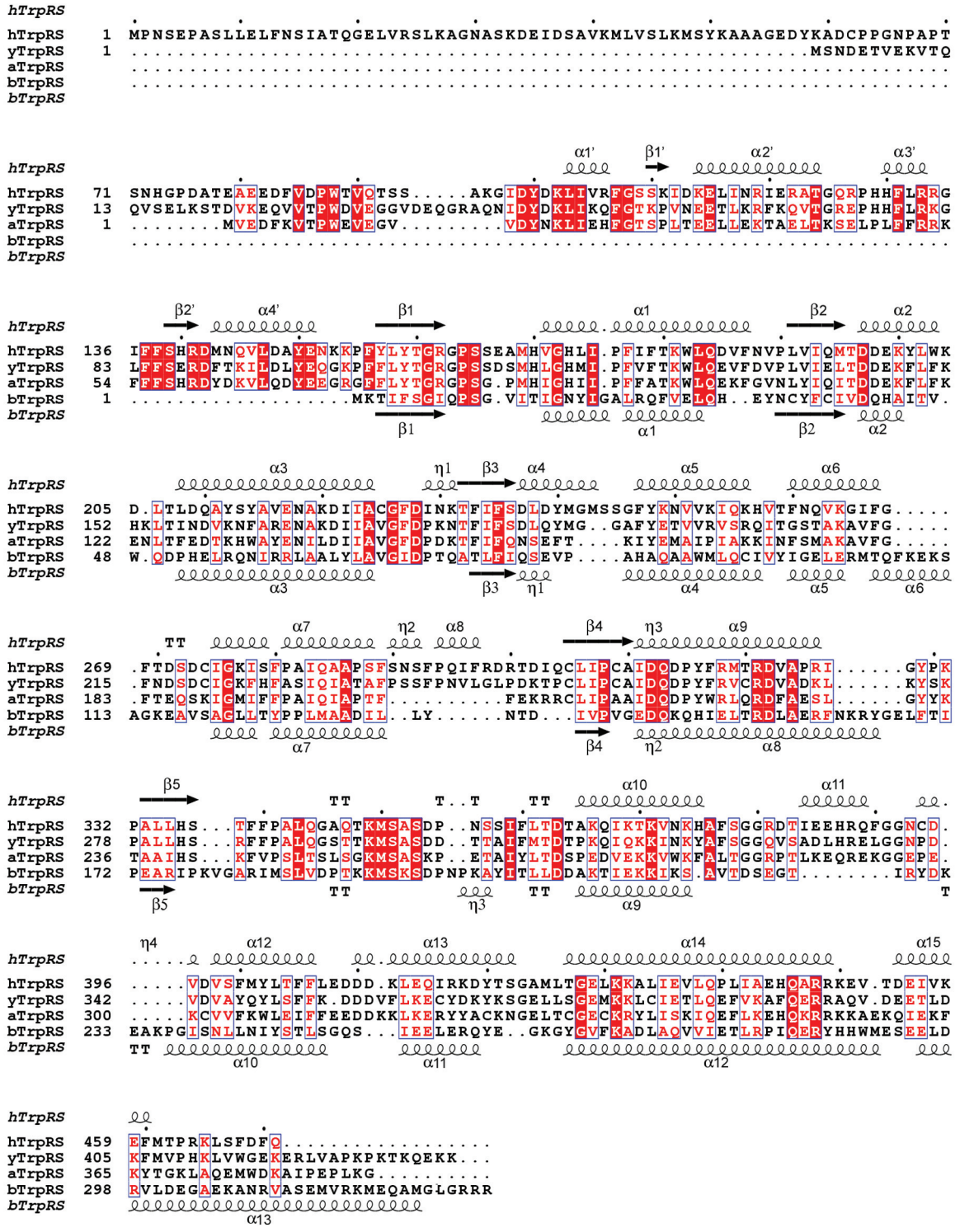


Figure 1. Structure of the hTrpRS–tRNA^{Trp}–Trp complex. (**A**) Overall structure of the dimeric hTrpRS–tRNA^{Trp}–Trp complex. hTrpRS is shown as ribbon with the N-terminal fragment in pink, the catalytic domain in green and the C-terminal domain (also called anticodon binding domain) in gold. For clarity, the second hTrpRS molecule is shown in gray. tRNAs are shown as cyan ribbons and the Trp substrates as ball-and-stick models. (**B**) A representative SIGMAA-weighted $2F_o$–F_c map (1.0 σ contour level) in the anticodon loop region of tRNA^{Trp} at 3.0 Å resolution. The final coordinates of tRNA^{Trp} (in yellow) and hTrpRS (in light blue) are shown as ball-and-stick models. (**C**) Structure-based sequence alignment of hTrpRS with representative TrpRSs from other species. hTrpRS, human TrpRS; yTyrRS, *S.cerevisiae* TrpRS; aTrpRS, *Pyrococcus abyssi* TrpRS and bTrpRS, *B.stearothermophilus* TrpRS. Strictly conserved residues are highlighted in shaded red boxes and conserved are shown in open red boxes. The secondary structure of hTrpRS is placed on top of the alignment [after Yu *et al*. (12)] and the secondary structure of bTrpRS is placed at the bottom of the alignment (9). The N-terminal 96 residues of hTrpRS do not exist in the structure of the hTrpRS–tRNA–Trp complex.

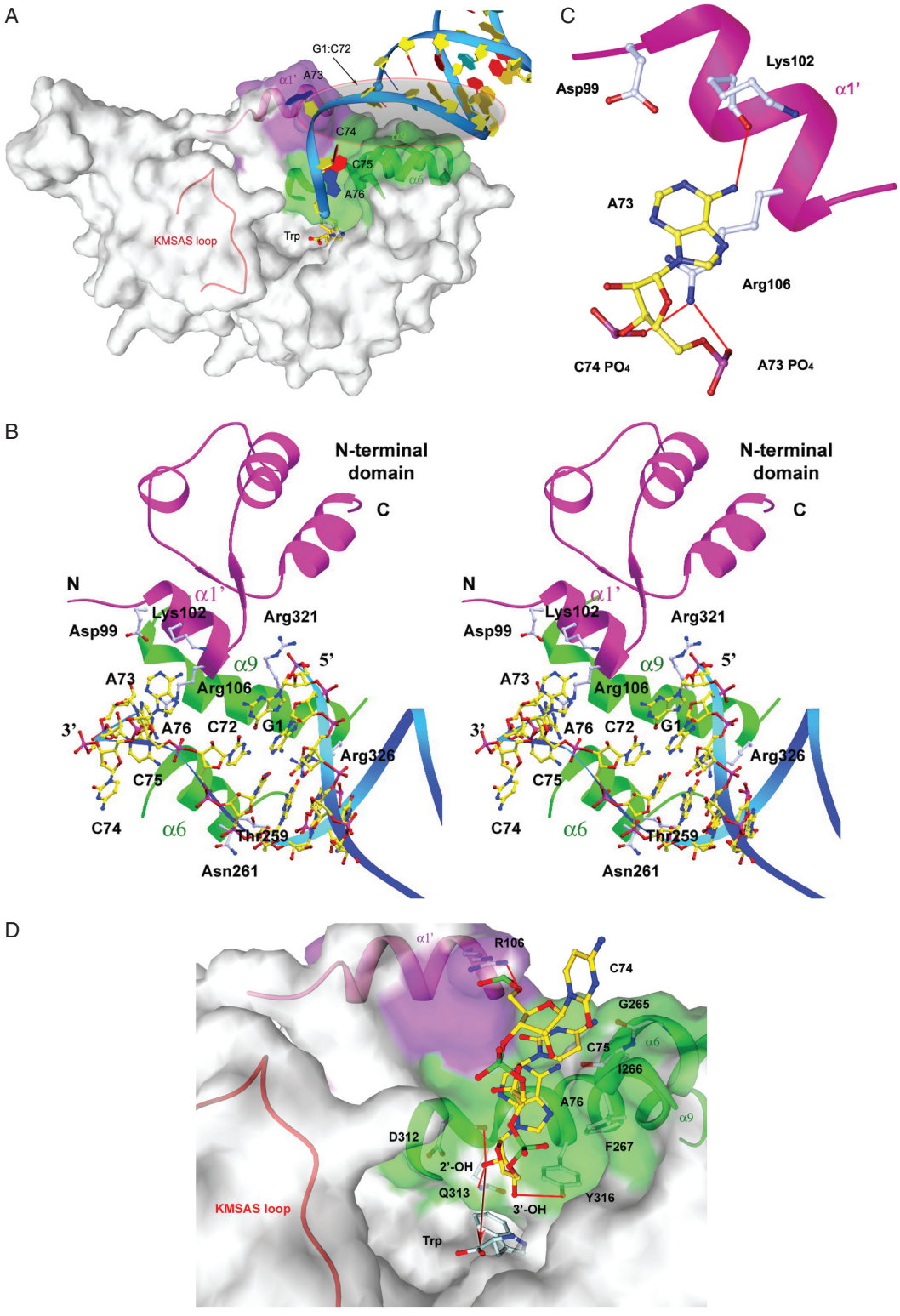


Figure 2. Recognition of the $tRNA^{Trp}$ acceptor arm by hTrpRS. (**A**) Molecular surface of the hTrpRS–$tRNA^{Trp}$–Trp complex showing the interactions between the tRNA acceptor arm and the structural elements of hTrpRS. hTrpRS interacts with the tRNA acceptor arm from the major groove. However, the 3′ end CCA takes a sharp turn to enter into the catalytic active site with a deformed conformation. The bound Trp is shown with a ball-and-stick model. (**B**) A stereoview showing the interactions of the tRNA acceptor arm with the structural elements of hTrpRS. The tRNA is shown as blue ribbons with the nucleotides of the acceptor arm in yellow. The N-terminal fragment of hTrpRS is shown in purple and helices α6 and α9 of the catalytic domain in green. Residues of hTrpRS involved in recognition of the tRNA acceptor arm are shown with side-chains (in light blue). (**C**) Recognition of the discriminator base A73 by residues of helix α1′ of the N-terminal fragment. The hydrogen-bonding interactions are indicated by thin red lines. (**D**) Molecular surface at the catalytic active site showing the interactions of the 3′ end CCA of the tRNA acceptor arm with the surrounding residues of hTrpRS. The 3′ end CCA of the tRNA is in the uncharged form and the 2′-OH group of A76 is positioned about 6 Å away from the α-carbonyl carbon of the Trp (as indicated by arrow).

Table 2. Kinetic data for aminoacylation of $tRNA^{Trp}$ by wild-type and mutant hTrpRS

hTrpRS	K_M (μM)	k_{cat} (s^{-1})	k_{cat}/K_M ($\mu M^{-1}s^{-1}$)	Relative activity[a]
Wild-type	1.21 ± 0.10	1.33 ± 0.08	1.10	1
D99A	3.64 ± 0.53	(7.20 ± 1.75) × 10^{-2}	1.98 × 10^{-2}	−55.6
D99V	0.325 ± 0.091	(1.03 ± 0.29) ×10^{-2}	3.17 × 10^{-2}	−34.7
D99E	1.00 ± 0.13	0.288 ± 0.020	0.288	−3.82
D99K	7.59 ± 1.18	(7.19 ± 0.91) × 10^{-2}	9.47 × 10^{-3}	−116.2
K102A	1.62 ± 0.41	(5.16 ± 0.70) × 10^{-2}	3.19 × 10^{-2}	−34.5
K102I	3.26 ± 0.24	(3.63 ± 0.82) × 10^{-2}	1.11 × 10^{-2}	−99.1
K102D	2.89 ± 0.70	(3.53 ± 0.74) × 10^{-2}	1.22 × 10^{-2}	−90.2
K102R	2.22 ± 0.27	0.282 ± 0.027	0.127	−8.66
R106A	2.51 ± 0.22	(3.00 ± 0.46) × 10^{-2}	1.20 × 10^{-2}	−91.7
R106I	0.239 ± 0.034	(1.71 ± 0.21) × 10^{-2}	7.15 × 10^{-2}	−15.4
R106D	0.549 ± 0.023	(3.68 ± 0.44) × 10^{-2}	6.70 × 10^{-2}	−16.4
R106K	0.361 ± 0.062	0.111 ± 0.014	0.307	−3.58
K431A	2.21 ± 0.11	(2.93 ± 0.19) × 10^{-3}	1.39 × 10^{-3}	−791.4
K431I	1.48 ± 0.19	(5.65 ± 0.29) × 10^{-3}	3.82 × 10^{-3}	−288.0
K431D	0.771 ± 0.039	(1.28 ± 0.03) × 10^{-3}	1.66 × 10^{-3}	−662.7
K431R	1.06 ± 0.12	(4.05 ± 0.17) × 10^{-2}	3.82 × 10^{-2}	−28.8
Q194A	1.63 ± 0.15	(1.16 ± 0.046) × 10^{-2}	7.09 × 10^{-3}	−155.1
Q194L[b]	—	—	—	—
Y159A	11.6 ± 1.0	(7.72 ± 0.17) × 10^{-2}	6.66 × 10^{-3}	−165.2
Y159F	0.364 ± 0.004	(5.66 ± 0.11) × 10^{-2}	0.156	−7.06

[a]The relative activity represents the k_{cat}/K_M value of each mutant relative to that of the wild-type enzyme.
[b]The aminoacylation activity is too low to be detected for this mutant.

defined in the electron density map, their backbones were well resolved. The N6 atom of the A73 base is specifically recognized by the main-chain carbonyl of Lys102 via a weak hydrogen bond. The side-chain carboxylate group of Asp99 is positioned to have a potential hydrogen-bonding interaction with the N1 atom of the A73 base. The side-chain of Arg106 is located beneath the A73 base with well defined electron density and forms stacking interactions with the A73 base. In addition, it forms hydrogen bonds with the phosphate groups of both A73 and C74. Both Asp99 and Lys102 are strictly conserved in archaeal and eukaryotic TrpRSs, while Arg106 can be replaced by Glu, Gln and His (Figure 1C) (20). The importance of these residues in the aminoacylation reaction was investigated by mutagenesis studies (Table 2). Mutation of Asp99 to Glu which has a slightly longer side-chain with a similar property had a minor effect on the aminoacylation activity (−3.82-fold). However, mutation of Asp99 to either Ala or Val, which has a shorter side-chain with a hydrophobic property resulted in a substantial reduction of the activity (−55.6 or −34.7-fold, respectively), and mutation of Asp99 to Lys, which has a longer side-chain with an opposite charge had a severe effect on the activity (−116.2-fold). Similarly, substitution of Lys102 with Arg whose side-chain has a similar property as Lys had a minor effect on the activity (−8.66-fold), but change of Lys102 to Ala, Ile or Asp had a more severe effect on the activity (−34.5-, −99.1- or −90.2-fold, respectively). For Arg106, mutation to Lys with a long side-chain, Ile or Asp with a middle-sized side-chain and Ala with a short side-chain resulted in increasingly severe impairment on the activity (−3.58-, −15.4-, −16.4- and −91.7-fold, respectively). These results indicate that the charge property of the Asp99 and Lys102 side-chains plays a more important role in the aminoacylation presumably through their interactions with the tRNA, while the size and length of the Arg106 side-chain play a more important role. Our structural data are in good agreement with the previous biochemical data that the A73 base of the $tRNA^{Trp}$ acceptor arm is the major identity element and the first and fifth base pairs the minor identity elements recognized by TrpRSs (20,21,27) and the N-terminal domain of hTrpRS plays a critical role in tRNA binding and the aminoacylation reaction (13,15). Based on these structural and biochemical data, we suggest that Lys102, Arg106 and possibly Asp99 are the primary residues responsible for the specific recognition of the discriminator A73 via both hydrogen-bonding and hydrophobic interactions. Nevertheless, since the N-terminal part of the enzyme is absent in this structure, therefore, we cannot completely exclude the possibility that the intrinsically flexible N-terminal region might also play a role in the recognition of the A73 base.

The 3′ end CCA of the tRNA intrude into the catalytic active site with a deformed conformation as observed in other class I aaRSs (28,29) (Figure 2D). The C74 base is flipped towards outside and has no contact with the protein, and its phosphate group is positioned through a hydrogen-bonding interaction with the Nη2 atom of Arg106 (2.8 Å). The C75 base has several hydrophobic contacts with the main chains of Phe267 and Gly268, and its N4 atom forms a hydrogen bond with the main-chain carbonyl of Gly265 (3.3 Å). A76 is located on top of the Trp and has extensive interactions with the surrounding residues. The A76 base is held in the pocket through a hydrogen-bonding interaction between its N6 atom and the main-chain carbonyl of Ile266 (2.4 Å) and hydrophobic interactions with Ile266 and Phe267 of the CP1 helix α6. The A76 ribose is positioned by Asp312 and Gln313 of the conserved AIDQ motif. Compared to its position in the unliganded hTrpRS structure (12), the AIDQ motif is moved closer to the 3′ end of the tRNA and adopts a closed conformation as that seen in the hTrpRS–Trp–AMP complex (10) (see Discussion later). The 2′-OH group forms hydrogen bonds with the main-chain carbonyl of Asp312 (3.1 Å) and the side-chain Nε2 of Gln313 (2.6 Å), and the 3′-OH group forms a hydrogen bond with the side-chain hydroxyl of Tyr316 (3.3 Å). Although both tRNA and Trp are bound at the active site, the 3′ end CCA of the tRNA is in the uncharged form. The ribose 2′-OH group of A76 has no direct interaction with the Trp and is positioned about 6 Å away from the α-carbonyl carbon of the Trp. This complex structure appears to represent a conformational state of the enzyme before the tRNA acylation reaction (see Discussion later).

Anticodon recognition

The anticodon loop of $tRNA^{Trp}$ is recognized by a number of conserved residues from the C-terminal helical domain of hTrpRS (Figure 3A and Supplementary Figure S3). Helices α10 (364–375), α11 (384–390) and α14 (427–450) form a pocket to accommodate the anticodon loop. In particular, the α10–α11 connecting loop wraps around the anticodon triplet and has extensive interactions with the bases. The C34 base is flipped out and sandwiched by the side-chains of ILe384, His387 and Leu426 via hydrophobic interactions.

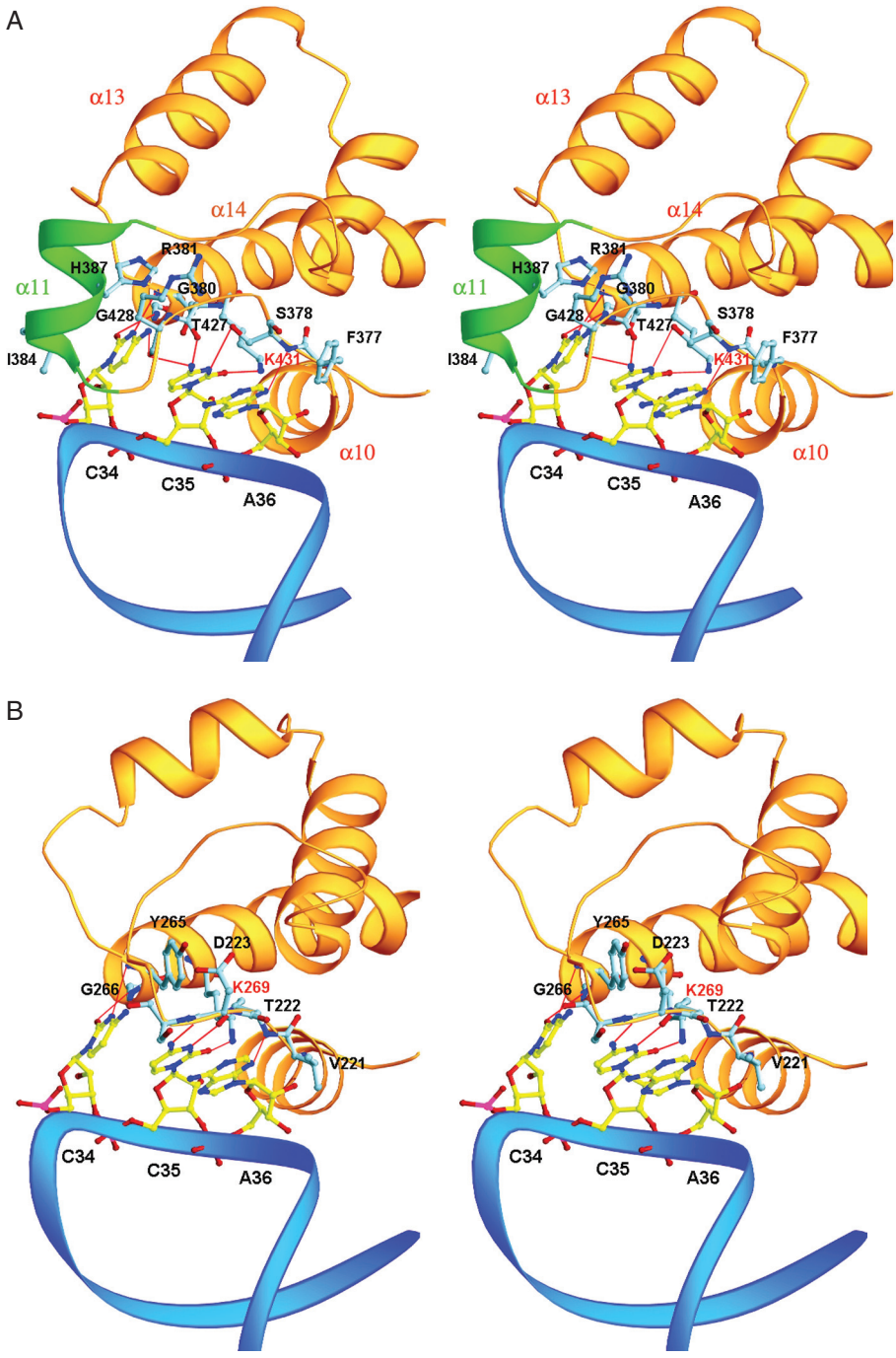


Figure 3. Recognition of the $tRNA^{Trp}$ anticodon by hTrpRS. (**A**) Structure of the tRNA anticodon binding site in the hTrpRS–$tRNA^{Trp}$–Trp complex. The tRNA is shown as blue ribbon with the anticodon CCA in yellow. The small helical domain of hTrpRS is shown in gold and the short helix α11 in green. Residues involved in the recognition of the tRNA anticodon are shown with side-chains (in light blue). The hydrogen-bonding interactions are indicated with thin red lines. (**B**) Structure of the tRNA anticodon binding site in the docking model of the bTrpRS–$tRNA^{Trp}$ complex. Due to the absence of an equivalent helix α11, the C34 base has less interaction with the protein.

O2 of C34 is recognized by the main-chain carbonyl of Gly428 through a hydrogen bond (3.3 Å); N3 of C34 forms a hydrogen bond with the main-chain amide of Thr427 (3.3 Å); and N4 of C34 forms two hydrogen bonds with the main-chain carbonyl of Arg381 (2.8 Å) and the side-chain Nδ1 of His387 (3.5 Å). The C35 base is recognized by Ser378, Gly380, Arg381 and Lys431. O2 and N3 of C35 form two hydrogen bonds with the side-chain amide of Lys431 (3.1 Å) and the hydroxyl of Ser378 (3.1 Å), respectively. N4 of C35 is recognized by the main-chain carbonyl groups of Gly380 and Arg381 through two hydrogen bonds (3.0 and 2.8 Å, respectively). The lack of a side-chain for Gly380 appears to be essential for avoiding steric hindrance with the C34 and C35 bases. The A36 base is positioned

between the C35 base and the Phe377 side-chain, and has less base-specific interactions with the protein than C34 and C35. The N3 atom of A36 forms a hydrogen bond with the main-chain amide of Ser378 (3.3 Å), and N6 of A36 a hydrogen bond with the side-chain Oδ1 of Asp382 (3.3 Å). These results indicate that C34 and C35 of the $tRNA^{Trp}$ anticodon are the major identity elements in its recognition by hTrpRS. This is consistent with the previous biochemical data that anticodon bases C34 and C35 are major, positive identity elements in *Saccharomyces cerevisiae* $tRNA^{Trp}$ (30).

It is noteworthy that Lys431 of hTrpRS is strictly conserved in all TrpRSs (Figure 1C) (20). Biochemical data show that mutation of Lys431 to Ala, Ile or Asp in hTrpRS severely impaired the aminoacylation activity (−791.4-, −288.0- or −662.7-fold, respectively), while change of Lys431 to Arg resulted in a slight decrease of the activity (−28.8-fold) (Table 2). It is likely that Lys431 of hTrpRS or its equivalent in other TrpRSs might be a key residue in the recognition of the second base of the tRNA anticodon.

Recognition of tryptophan

The catalytic active site is located at a deep pocket in the catalytic domain that is exposed to solvent and accessible to substrates. The Trp substrate is bound at the bottom of the pocket and occupies a similar position as the tryptophanyl moiety of Trp-AMP in the hTrpRS–Trp–AMP complex (10) and the bTrpRS–Trp–AMP complex (9). It has extensive interactions with a number of residues from structural elements β1 (Tyr159, Gly161 and Gly163), β2 (Gln194 and Thr196), α2 (Glu199), α7 (Gln284) and α9 (Gln313 and Phe317) (Figure 4A). The indole nitrogen of Trp is specifically recognized via hydrogen-bonding interactions by Tyr159 (with its hydroxyl group, 3.0 Å) and Gln194 (with its Oε1 atom, 2.6 Å). In addition, the side-chain of Tyr159 has a π–π stacking interaction with the Trp indole ring. Both Tyr159 and Gln194 are strictly conserved in eukaryotic TrpRSs (Figure 1C). Biochemical data show that mutation of Tyr159 to Ala severely decreased the aminoacylation activity of hTrpRS (−165.2-fold), whereas change of Tyr159 to Phe resulted in only a minor reduction of the activity (about −7-fold), suggesting that the π–π interaction between Tyr159 and the Trp indole group plays a more important role than the hydrogen-bonding interaction in the recognition and binding of the substrate (Table 2). On the contrary, substitution of Gln194 with Ala severely impaired the activity (−155.1-fold) and substitution of Gln194 with Leu completely abolished the activity, suggesting that the hydrogen-bonding interaction between Gln194 and the Trp indole nitrogen is indispensable in the recognition and binding of the substrate (Table 2). In the bTrpRS–Trp–AMP complex, the indole Nε1 atom of Trp forms a hydrogen bond with the side-chain of Asp132 (not equivalent to Gln194 of hTrpRS) and the indole ring has a π–π stacking interaction with the side-chain of Phe5 (equivalent to Tyr159 of hTrpRS) (9). These results suggest that hTrpRS (and possibly other eukaryotic TrpRSs) uses Tyr159 and Gln194 (or their equivalents) to recognize Trp, while bTrpRS (and possibly other bacterial TrpRSs) uses Phe5 and Asp132 (or their equivalents) instead.

The conformational state of the tRNA complex

Multiple conformations have been observed in the bTrpRS structures (including apo form and its complexes with Trp, ATP and Trp-AMP) representing different intermediate states of the enzyme in the aminoacylation reaction, and conformational changes take place both globally and locally at the catalytic active site during aminoacylation (9,16,31). Previously, crystal structures of hTrpRS in both unliganded form and in complex with Trp-AMP have been reported (10,12,17). Structural comparisons indicate that the unliganded hTrpRS structure assumes an open conformation, which represents the initial state of the enzyme (12) and the Trp-AMP-bound hTrpRS structure adopts a closed conformation, which represents the post-transition product state of the enzyme after Trp activation (10). In the tRNA complex, the overall structure of hTrpRS assumes a closed conformation similar to that found in the Trp–AMP complex (an RMSD of 0.77 Å for 364 Cα atoms), but different from the open conformation seen in the unliganded hTrpRS structure (an RMSD of 1.2 Å for 370 Cα atoms) (Figure 4B). The AIDQ motif also adopts a closed conformation similar to that seen in the Trp–AMP complex and is shifted closer (about 4 Å) to the 3′ end of the tRNA and the Trp compared to its position in the unliganded hTrpRS structure. However, the KMSAS loop assumes a half open conformation between the open conformation seen in the unliganded hTrpRS structure and the closed conformation seen in the Trp–AMP complex. Compared to its position in the Trp–AMP complex, the KMSAS loop in the tRNA complex is moved away from the active site by about 4 Å. The entrance to the substrate-binding pocket is open and accessible by the tRNA acceptor arm. Considering that though both tRNA and Trp are bound at the active site, the 3′ end CCA of the tRNA is in the uncharged form (see Discussion above), therefore, it is very likely that this complex structure represents an intermediate state of the enzyme preceding the formation of the transition state for the tRNA acylation reaction. The enzyme might be trapped in this conformational state due to the lack of aminoacylation activity of T2-hTrpRS. To carry out the acylation reaction, conformational changes have to take place at the active site on both the enzyme (especially the KMSAS loop) and/or the tRNA acceptor arm (especially the 3′ end CCA).

Potential functional role of the N-terminal domain of hTrpRS

In the structure of the hTrpRS–Trp–AMP complex, one hTrpRS molecule contains a partially ordered N-terminal region: residues 7–60 form a helix–turn–helix motif (HTH) and residues 82–92 form a β-hairpin (10) (Figure 5A). Previous biochemical data have shown that mini-hTrpRS (residues 48–471) and T1-hTrpRS (residues 71–471) have comparable aminoacylation activity as the full-length enzyme, whereas T2-hTrpRS (residues 94–471) possesses no aminoacylation activity (13,15). These results suggest that the β-hairpin is required for the aminoacylation reaction, while the HTH motif is not. Analysis of the hTrpRS–Trp–AMP complex reveals that the HTH motif is located beneath the catalytic domain and the β-hairpin is positioned on top of the substrate-binding pocket (Figure 5A). In particular, residues

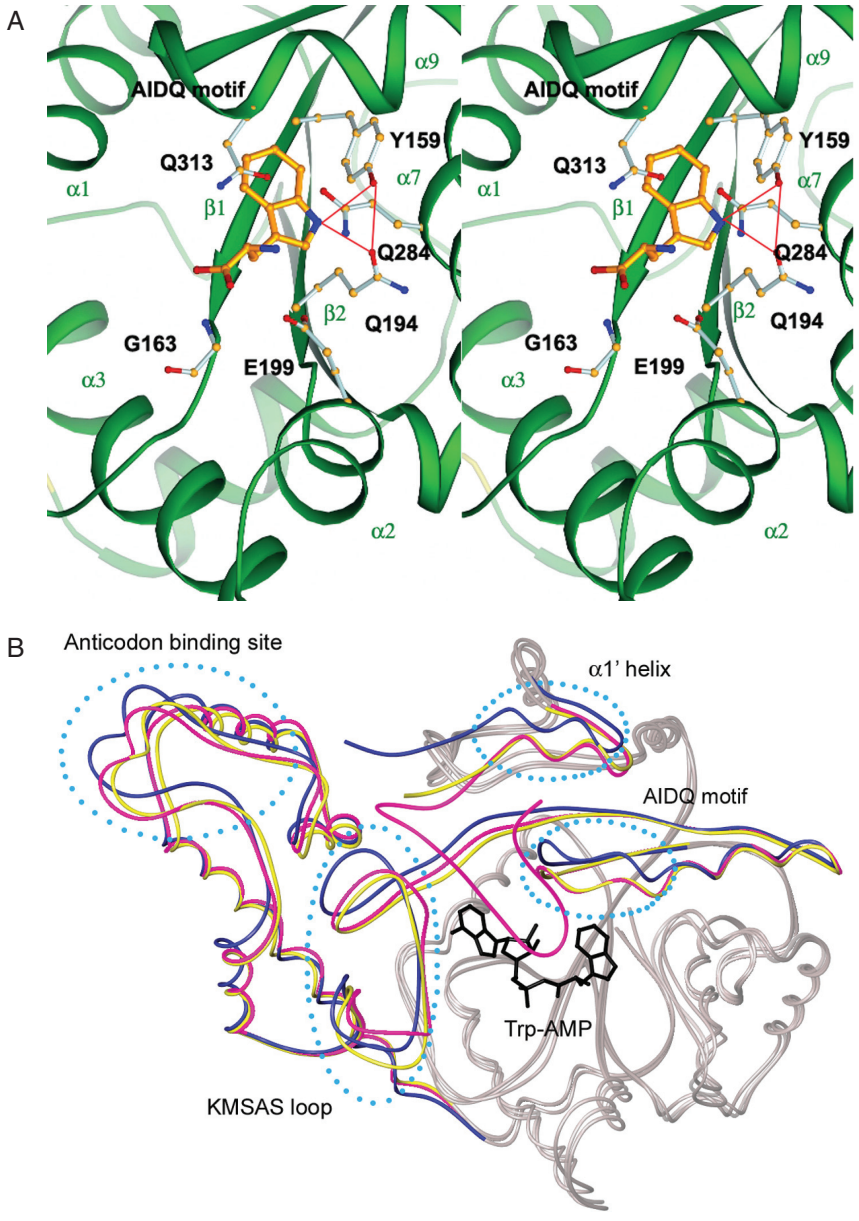


Figure 4. Structure of the catalytic active site of hTrpRS. (**A**) A stereoview showing the interactions between the bound Trp and the surrounding residues of hTrpRS. The Trp is shown in gold and the residues having close interactions with the substrate are shown with side-chains (in gray). The indole N atom of the Trp is specifically recognized by Tyr159 and Gln194 via hydrogen-bonding interactions (indicated by thin red lines). (**B**) Structural comparison of hTrpRS in the hTrpRS–$tRNA^{Trp}$–Trp complex (in yellow), the hTrpRS–Trp–AMP complex (in pink) (10) and the unliganded hTrpRS (in blue) (12). The superposition is based on the catalytic domain (in gray). The regions displaying major conformational changes are indicated with dotted circles. For clarity, only Trp-AMP in the hTrpRS–Trp–AMP complex is shown as a ball-and-stick model.

Pro87 and Trp88 at the turn of the β-hairpin, which are strictly conserved in all archaeal/eukaryotic TrpRSs (Figure 1C) (20), cover on top of the phosphate and sugar moieties of the Trp-AMP. The main-chain amide of Pro87 forms a hydrogen bond with the main-chain carbonyl of Asp312 (3.2 Å) of the AIDQ motif, and the main-chain carbonyl of Trp88 forms a weak hydrogen bond with the side-chain Nζ of Lys349 (3.6 Å) of the KMSAS loop. Both Asp312 and Lys349 are involved in the binding of ATP and removal of the pyrophosphate during the Trp activation. Thus, it seems very likely that the β-hairpin is involved in Trp activation either directly by participating in the hydrolysis of the pyrophosphate and/or indirectly by positioning Asp312 and Lys349 to facilitate the binding and dissociation of the substrate.

On the other hand, in the hTrpRS–tRNA–Trp complex, the N-terminal 96 residues (including both the HTH motif and the β-hairpin) do not exist and the tRNA acceptor arm is bound at the catalytic active site with its 3′ end CCA in the uncharged form. Structural comparison indicates that the β-hairpin in the Trp–AMP complex overlaps with the 3′ end CCA of the tRNA and the disordered region (residues 61–81) linking the β-hairpin and the HTH motif covers on top of the catalytic active site (Figure 5B). Both of them cannot

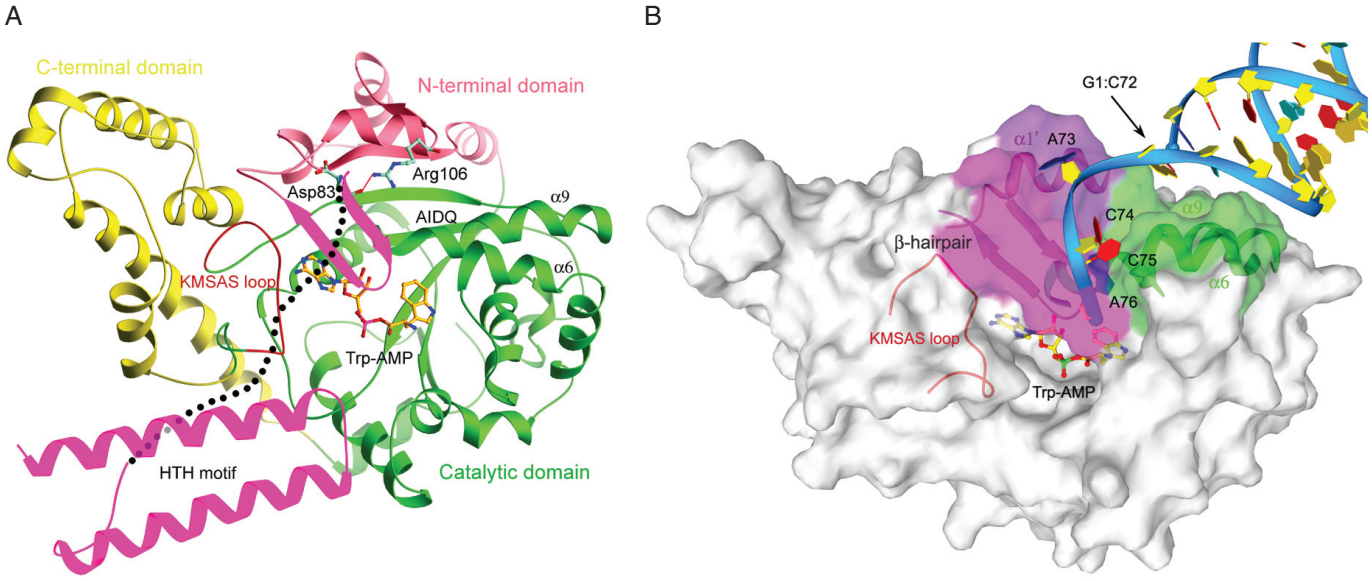


Figure 5. Potential functional role of the N-terminal β-hairpin of hTrpRS. (**A**) Structure of the hTrpRS–Trp–AMP complex (10). The N-terminal HTH motif and β-hairpin are shown in purple and the Trp-AMP as a ball-and-stick model. The disordered region linking the HTH motif and β-hairpin is indicated with a dashed line. (**B**) Molecular surface of the hTrpRS–Trp–AMP complex showing that the β-hairpin covers on top of the catalytic active site and overlaps with the 3′ end CCA of a docked tRNA model based on the hTrpRS–$tRNA^{Trp}$–Trp complex.

take the same positions during tRNA binding and the acylation reaction because of spatial conflicts, implying that the HTH motif and the β-hairpin are not required for tRNA binding and may not be essential for the acylation reaction. Indeed, our preliminary kinetic data show that T2-hTrpRS has no measurable activity in the Trp activation and the overall aminoacylation reaction, but has weak activity in the acylation reaction (X. L. Chen and Y. X. Jin, unpublished data). It is very likely that after the Trp activation, the N-terminal region (including the HTH motif and the β-hairpin) and the KMSAS loop first undergo conformational changes, leading to relocation of the N-terminal region and widening of the entrance to the substrate-binding pocket, and then the tRNA binds to the active site to carry out the acylation reaction. The intrinsic flexibility of the N-terminal region of hTrpRS is well manifested in the crystal structures of hTrpRS. The N-terminal region is only partially ordered in one subunit of the dimeric hTrpRS with bound Trp-AMP, but is disordered in the other subunit (10). In all other hTrpRS structures, most of the N-terminal residues are either hydrolyzed or disordered in both unliganded form (12,17) and complexes with different substrates (including ATP, Trp, Trp analog and Trp-AMP) (N. Shen and J. Ding, unpublished data). Taken together, these results indicate that the N-terminal region of hTrpRS is involved in the Trp activation, but is not required for tRNA binding and the acylation reaction; it has an inherited flexibility and undergoes conformational change during tRNA binding and aminoacylation.

tRNA recognition mechanisms by different TrpRSs

Analysis of the crystal structure of the hTrpRS–$tRNA^{Trp}$–Trp complex has revealed a novel tRNA binding and recognition mechanism by hTrpRS. In addition, this complex structure allows us to perform structural and sequence comparison of hTrpRS with other TrpRSs, which will shed some light into the molecular basis of the tRNA specificity for TrpRSs from different species. Biochemical data have shown that eukaryotic and bacterial TrpRSs cannot efficiently aminoacylate $tRNA^{Trp}$ from other species (20,21,26,27). The specific recognition of $tRNA^{Trp}$ by its cognate TrpRS involves the identity elements on the acceptor arm and the anticodon loop. The N73 base of the $tRNA^{Trp}$ acceptor arm is the major identity element and the first and fifth base pairs the minor identity elements recognized by TrpRSs (20,21,27). Sequence comparison indicates that archaeal and eukaryotic $tRNA^{Trp}$ have conserved A73, G1:C72 and U5:G68 in the acceptor arm, whereas bacterial $tRNA^{Trp}$ have conserved G73, A1:U72 and G5:C68 (32). Mutation of A73G in bovine $tRNA^{Trp}$ resulted in a substantial decrease in the aminoacylation activity of hTrpRS and greatly increases its aminoacylation efficiency by *Bacillus subtilis* TrpRS, and additional mutations (G1A:C72U and U5G:G68C) had only a minor effect (27). Conversely, mutation of G73A in *B.subtilis* $tRNA^{Trp}$ resulted in a drastic increase in its aminoacylation efficiency by hTrpRS and additional mutations (A1G:U72C and G5U:C68G) had a minor positive effect. These biochemical data are in good agreement with our structural data. In the structure of the hTrpRS–tRNA–Trp complex, the $tRNA^{Trp}$ acceptor arm has interactions with primarily the N-terminal fragment of hTrpRS and the A73 base is specifically recognized by Lys102, Arg106 and possibly Asp99 of helix α1′ of the N-terminal fragment via both hydrogen-bonding and hydrophobic interactions, while the first and fifth base pairs of the acceptor arm have no base-specific interactions with the catalytic domain. Change of A73 to G might affect the specific interactions of the A73 base with Lys102, Arg106 and Asp99. In particular, when A73 is replaced by G as in bacterial $tRNA^{Trp}$, the main-chain carbonyl of Lys102 and the O6 atom of the G73 base cannot form hydrogen bond

because both of them are hydrogen bond acceptors, and the N2 group of the G73 base might generate steric hindrance with the side-chain of Asp99. This may explain why eukaryotic TrpRSs cannot efficiently cross-aminoacylate bacterial $tRNA^{Trp}$.

Further sequence comparison of TrpRSs from different species shows that other eukaryotic TrpRSs also contain an N-terminal region similar to that of hTrpRS with varying length, and archaeal TrpRSs contain a shorter N-terminal region equivalent to the N-terminal fragment seen in T2-hTrpRS (Figure 1C) (20). Asp99 and Lys102 of hTrpRS are strictly conserved in both eukaryotic and archaeal TrpRSs, and Arg106 of hTrpRS can be replaced by a polar residue with a large side-chain (His/Glu/Gln). Considering that both the amino acid residues of TrpRS and the nucleotides of the $tRNA^{Trp}$ acceptor arm involved in the recognition of $tRNA^{Trp}$ are strongly conserved in all eukaryotic and archaeal systems, it seems possible that other eukaryotic and archaeal TrpRSs may contain a similar N-terminal domain as that of hTrpRS, and may recognize the $tRNA^{Trp}$ acceptor arm in a similar manner as that by hTrpRS. On the contrary, bacterial TrpRSs do not contain an equivalent N-terminal domain and bacterial tRNAs have conserved G73 and A1:U72 in the acceptor arm, suggesting that bacterial TrpRSs might use a different mechanism to recognize the $tRNA^{Trp}$ acceptor arm. Since bTrpRS has a structure similar to that of TyrRSs, which also lack an N-terminal domain, it is possible that bTrpRS (and possibly other bacterial TrpRSs) might recognize the $tRNA^{Trp}$ acceptor arm through interactions primarily between the structural elements of the catalytic domain and the discriminator G73 and the first base pair of the tRNA in a manner similar to that of TyrRSs (18,19). This suggestion is supported by the biochemical data that two conserved residues Lys149 and Glu153 in the catalytic domain of *B.subtilis* TrpRS (equivalent to Asp314 and Arg318 in helix α9 of hTrpRS) are crucial for the correct recognition of its cognate $tRNA^{Trp}$ (26) and the three G:C base pairs in the *B.subtilis* $tRNA^{Trp}$ acceptor arm are required for its efficient aminoacylation by *B.subtilis* TrpRS (33).

Although the anticodon of $tRNA^{Trp}$ is also a major identity determinant in its recognition by TrpRS (30), all $tRNA^{Trp}$ have strictly conserved CCA as the anticodon (32), suggesting that the anticodon might play a less critical role than the acceptor arm in determining the species-specificity of $tRNA^{Trp}$. Docking experiment of the bTrpRS–$tRNA^{Trp}$ complex shows that the structural elements and residues involved in the recognition of the $tRNA^{Trp}$ anticodon are well conserved, including Val221, Thr222, Asp223, Tyr265, Gly266 and Lys269 of bTrpRS, corresponding to Phe377, Ser378, Gly380, Thr427, Gly428 and Lys431 of hTrpRS (Figure 3B). Similar to that by hTrpRS, the recognition of the $tRNA^{Trp}$ anticodon by bTrpRS appears to be base-specific as well. Nevertheless, there is a marked difference between hTrpRS and bTrpRS: hTrpRS has a short insertion (helix α11) in the small helical domain, which plays a critical role in the recognition of C34 (Figure 3A). Due to the lack of the insertion, the anticodon binding site of bTrpRS differs slightly from that of hTrpRS, resulting in variations of the bTrpRS–tRNA interactions. Most notably, the C34 base has less interactions with the protein and is only poorly recognized by the main-chain carbonyls of Tyr265 and Gly266 (Figure 3B). The C35 base maintains the specific hydrogen-bonding interactions with the main-chain carbonyl of Asp223, the side-chain amide of Lys269, and the hydroxyl of Thr222. The A36 base retains a hydrogen-bonding interaction with the main-chain amide of Thr222. Sequence comparison of the small helical domain of TrpRSs from different species indicates that most eukaryotic and archaeal TrpRSs contain a conserved insertion equivalent to helix α11 of hTrpRS, while bacterial TrpRSs do not (20), suggesting that other eukaryotic and archaeal TrpRSs might recognize both C34 and C35 of the $tRNA^{Trp}$ anticodon in a similar manner as that of hTrpRS, whereas bacterial TrpRSs might recognize mainly C35 of the $tRNA^{Trp}$ anticodon.

Taking together, we propose that in eukaryotes and archaea, A73 of the acceptor arm and the anticodon CCA of $tRNA^{Trp}$ are the major identity elements and the first base pair (G1:C72) is the minor identity element recognized by their cognate TrpRSs; while in bacteria, G73 and the first base pair (A1:U72) of the acceptor arm and the anticodon CCA of $tRNA^{Trp}$ are the major identity elements and the other base pairs of the acceptor arm are the minor identity elements. The discrimination of eukaryotic/archaeal $tRNA^{Trp}$ by their cognate TrpRSs against bacterial $tRNA^{Trp}$ relies primarily on the specific recognition of A73 versus G73, while that of bacterial $tRNA^{Trp}$ relies on the specific recognitions of both G73 versus A73 and A1:U72 versus G1:C72. The differences in the recognition of the $tRNA^{Trp}$ acceptor arm by different TrpRSs form the main basis of the species-specificity, and the differences in the recognition of the $tRNA^{Trp}$ anticodon by the small helical domain of TrpRSs play a secondary role in determining the species-specificity.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

The authors thank the staff members at Beijing Synchrotron Radiation Facility and Cornell High Energy Synchrotron Source for support in diffraction data collection, and other members of our groups for helpful discussion. This work is supported by NSFC grants (30570379 and 30370325) and MOST grants (2004CB720102 and 2002CB512803). Funding to pay the Open Access publication charges for this article was provided by MOST.

Conflict of interest statement. None declared.

REFERENCES

1. O'Donoghue,P. and Luthey-Schulten,Z. (2003) On the evolution of structure in aminoacyl-tRNA synthetases. *Microbiol. Mol. Biol. Rev.*, **67**, 550–573.
2. Woese,C., 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.
3. Arnez,J.G. and Moras,D. (1997) Structural and functional considerations of the aminoacylation reaction. *Trends Biochem. Sci.*, **22**, 211–216.

4. Cusack,S. (1997) Aminoacyl-tRNA synthetases. *Curr. Opin. Struct. Biol.*, **7**, 881–889.
5. 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.
6. Brick,P., Bhat,T.N. and Blow,D.M. (1989) Structure of tyrosyl-tRNA synthetase refined at 2.3 Å resolution: interaction of the enzyme with the tyrosyl adenylate intermediate. *J. Mol. Biol.*, **208**, 83–98.
7. Brown,J.R., Robb,F.T., Weiss,R. and Doolittle,W.F. (1997) Evidence for the early divergence of tryptophanyl- and tyrosyl-tRNA synthetases. *J. Mol. Evol.*, **45**, 9–16.
8. Cusack,S. (1995) Eleven down and nine to go. *Nature Struct. Biol.*, **2**, 824–831.
9. Doublie,S., Bricogne,G., Gilmore,C. and Carter,C.W.,Jr (1995) Tryptophanyl-tRNA synthetase crystal structure reveals an unexpected homology to tyrosyl-tRNA synthetase. *Structure*, **3**, 17–31.
10. Yang,X.L., Otero,F.J., Skene,R., McRee,D.E., Schimmel,P. and de Pouplana,L.R. (2003) Crystal structures that suggest late development of genetic code components for differentiating aromatic side chains. *Proc. Natl Acad. Sci. USA*, **100**, 15376–15380.
11. Yang,X.L., Skene,R.J., McRee,D.E. and Schimmel,P. (2002) Crystal structure of a human aminoacyl-tRNA synthetase cytokine. *Proc. Natl Acad. Sci. USA*, **99**, 15369–15374.
12. Yu,Y., Liu,Y., Shen,N., Xu,X., Xu,F., Jia,J., Jin,Y., Arnold,E. and Ding,J. (2004) Crystal structure of human tryptophanyl-tRNA synthetase catalytic fragment. *J. Biol. Chem.*, **279**, 8378–8388.
13. Otani,A., Slike,B.M., Dorrell,M.I., Hood,J., Kinder,K., Ewalt,K.L., Cheresh,D., Schimmel,P. and Friedlander,M. (2002) A fragment of human TrpRS as a potent antagonist of ocular angiogenesis. *Proc. Natl Acad. Sci. USA*, **99**, 178–183.
14. Wakasugi,K. and Schimmel,P. (1999) Two distinct cytokines released from a human aminoacyl-tRNA synthetases. *Science*, **284**, 147–151.
15. Wakasugi,K., Slike,B.M., Hood,J., Otani,A., Ewalt,K.L., Friedlander,M., Cheresh,D.A. and Schimmel,P. (2002) A human aminoacyl-tRNA synthetase as a regulator of angiogenesis. *Proc. Natl Acad. Sci. USA*, **99**, 173–177.
16. Ilyin,V., Temple,B., Hu,M., Li,G., Yin,Y., Vachette,P. and Carter,C.W.,Jr (2000) 2.9 Å crystal structure of ligand-free tryptophanyl-tRNA synthetase: domain movements fragment the adenine nucleotide binding site. *Protein Sci.*, **9**, 218–231.
17. Kise,Y., Lee,S.W., Park,S.G., Fukai,S., Sengoku,T., Ishii,R., Yokoyama,S., Kim,S. and Nureki,O. (2004) A short peptide insertion crucial for angiostatic activity of human tryptophanyl-tRNA synthetase. *Nature Struct. Mol. Biol.*, **11**, 149–156.
18. Kobayashi,T., Nureki,O., Ishitani,R., Yaremchuk,A., Tukalo,M., Cusack,S., Sakamoto,K. and Yokoyama,S. (2003) Structural basis for orthogonal tRNA specificities of tyrosyl-tRNA synthetases for genetic code expansion. *Nature Struct. Mol. Biol.*, **10**, 425–432.
19. Yaremchuk,A., Kriklivyi,I., Tukalo,M. and Cusack,S. (2002) Class I tyrosyl-tRNA synthetase has a class II mode of tRNA recognition. *EMBO J.*, **21**, 3829–3840.
20. Xu,F., Chen,X., Xin,L., Chen,L., Jin,Y. and Wang,D. (2001) Species-specific differences in the operational RNA code for aminoacylation of $tRNA^{Trp}$. *Nucleic Acids Res.*, **29**, 4125–4133.
21. Xue,H., Shen,W., Giege,R. and Wong,J.T. (1993) Identity elements of $tRNA^{Trp}$. *J. Biol. Chem.*, **268**, 9316–9322.
22. Otwinowski,Z. and Minor,W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Meth. Enzymol.*, **276A**, 307–326.
23. Ruff,M., Krishnaswamy,S., Boeglin,M., Poterszman,A., Mitschler,A., Podjarny,A., Rees,B., Thierry,J.C. and Moras,D. (1991) Class II aminoacyl transfer RNA synthetases: crystal structure of yeast aspartyl-tRNA synthetase complexed with $tRNA^{Asp}$. *Science*, **252**, 1682–1689.
24. Brünger,A.T., Adams,P.D., Clore,G.M., Delano,W.L., Gros,P., Grosse-Kunstleve,R.W., Jiang,J.-S., Kuszewski,J., Nilges,M., Pannu,N.S. *et al.* (1998) Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Cryst.*, **D54**, 905–921.
25. 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 Cryst.*, **A47**, 110–119.
26. Jia,J., Chen,X.L., Guo,L.T., Yu,Y., Ding,J. and Jin,Y. (2004) Residues Lys-149 and Glu-153 switch the aminoacylation of $tRNA^{Trp}$ in *Bacillus subtilis*. *J. Biol. Chem.*, **279**, 41960–41965.
27. Guo,Q., Gong,Q., Tong,K.L., Vestergaard,B., Costa,A., Desgres,J., Wong,M., Grosjean,H., Zhu,G., Wong,J.T. *et al.* (2002) Recognition by tryptophanyl-tRNA synthetases of discriminator base on $tRNA^{Trp}$ from three biological domains. *J. Biol. Chem.*, **277**, 14343–14349.
28. Rould,M.A., Perona,J.J., Söll,D. and Steitz,T.A. (1989) Structure of *E. coli* glutaminyl-tRNA synthetase complexed with $tRNA^{Gln}$ and ATP at 2.8 Å resolution. *Science*, **246**, 1135–1142.
29. Rath,V.L., Silvian,L.F., Beijer,B., Sproat,B.S. and Steitz,T.A. (1998) How glutaminyl-tRNA synthetase selects glutamine. *Structure*, **6**, 439–449.
30. Yesland,K.D. and Johnson,J.D. (1993) Anticodon bases C34 and C35 are major, positive identity elements in *Saccharomyces cerevisiae* $tRNA^{Trp}$. *Nucleic Acids Res.*, **21**, 5079–5084.
31. Retailleau,P., Huang,X., Yin,Y., Hu,M., Weinreb,V., Vachette,P., Vonrhein,C., Bricogne,G., Roversi,P., Ilyin,V. *et al.* (2003) Interconversion of ATP binding and conformational free energies by tryptophanyl-tRNA synthetase: structures of ATP bound to open and closed, pre-transition-state conformations. *J. Mol. Biol.*, **325**, 39–63.
32. Sprinzl,M., Horn,C., Brown,M., Loudovitch,A. and Steinberg,S. (1998) Compilation of tRNA sequences and sequences of tRNA genes. *Nucleic Acids Res.*, **26**, 148–153.
33. Xu,F., Jiang,G., Li,W., He,X., Jin,Y. and Wang,D. (2002) Three G:C base pairs required for the efficient aminoacylation of $tRNA^{Trp}$ by tryptophanyl-tRNA synthetase from *Bacillus subtilis*. *Biochemistry*, **41**, 8087–8092.