The EMBO Journal vol.15 no.22 pp.6321–6334, 1996

The crystal structures of *T.thermophilus* lysyl-tRNA synthetase complexed with *E.coli* $tRNA^{Lys}$ and a *T.thermophilus* $tRNA^{Lys}$ transcript: anticodon recognition and conformational changes upon binding of a lysyl-adenylate analogue

Stephen Cusack[1], Anna Yaremchuk and Michael Tukalo

European Molecular Biology Laboratory, Grenoble Outstation, c/o ILL, 156X, F-38042 Grenoble Cedex 9, France

[1]Corresponding author

The crystal structures of *Thermus thermophilus* lysyl-tRNA synthetase, a class IIb aminoacyl-tRNA synthetase, complexed with *Escherchia coli* $tRNA^{Lys}$(mnm^5s^2UUU) at 2.75 Å resolution and with a *T.thermophilus* $tRNA^{Lys}$(CUU) transcript at 2.9 Å resolution are described. In both complexes only the tRNA anticodon stem–loop is well ordered. The mode of binding of the anticodon stem–loop to the N-terminal β-barrel domain is similar to that previously found for the homologous class IIb aspartyl-tRNA synthetase–$tRNA^{Asp}$ complex except in the region of the wobble base 34 where either mnm^5s^2U or C can be accommodated. The specific recognition of the other anticodon bases, U-35 and U-36, which are both major identity elements in the lysine system, is also described. Additional crystallographic data on a ternary complex with a lysyl-adenylate analogue show that binding of the intermediate induces significant conformational changes in the vicinity of the active site of the enzyme.

Keywords: anticodon/lysyl-tRNA synthetase/modified base/protein–RNA recognition/tRNA

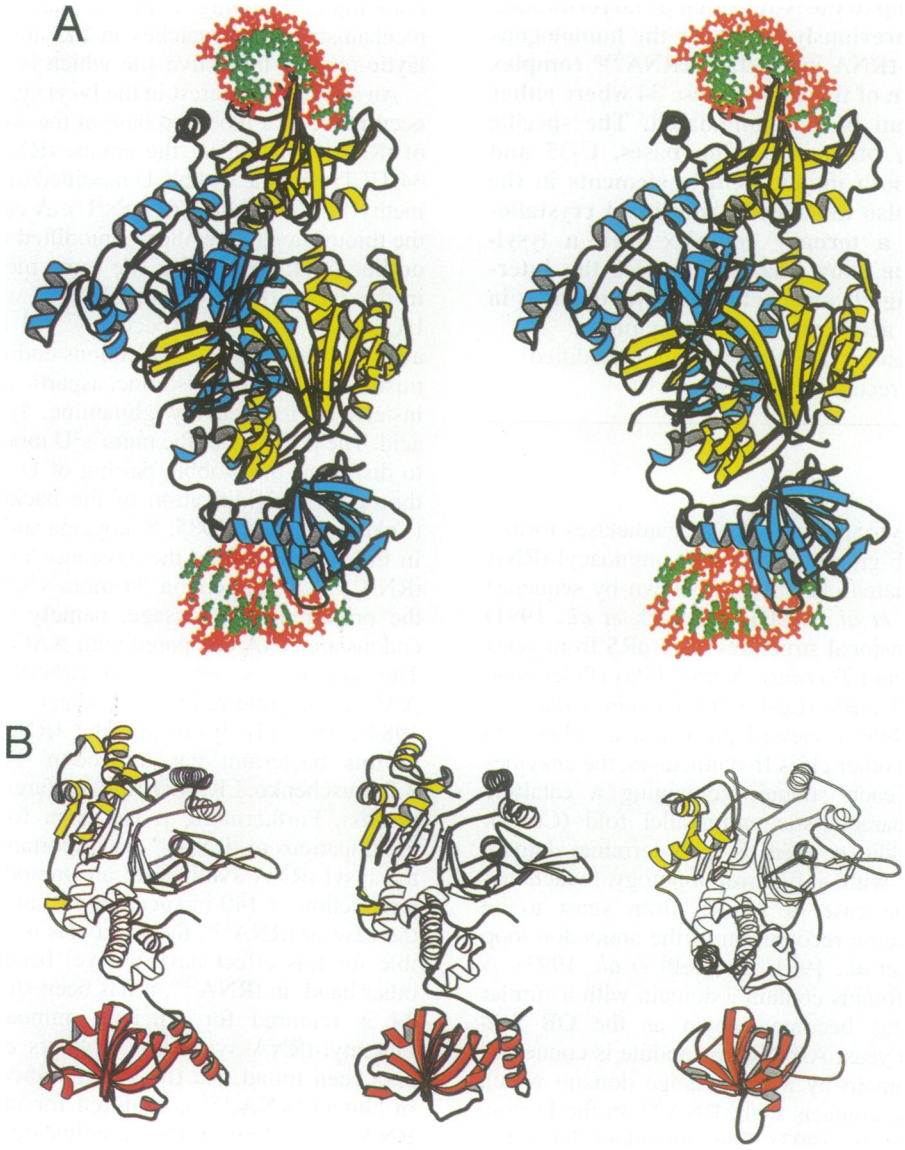
Introduction

Aspartyl-, asparginyl- and lysyl-tRNA synthetases form a closely related sub-group of class II aminoacyl-tRNA synthetases (designated class IIb) as shown by sequence alignments (Eriani *et al.*, 1990a,b; Cusack *et al.*, 1991) and the three-dimensional structures of AspRS from yeast (Ruff *et al.*, 1991) and *Thermus thermophilus* (Poterzman *et al.*, 1994) and LysRS (LysU) from *Escherichia coli* (Onesti *et al.*, 1995; reviewed in Cusack, 1995). In common with most other class II synthetases, the enzymes are homodimers, each subunit containing a catalytic domain with a characteristic antiparallel fold (Cusack *et al.*, 1990). In addition, there is an N-terminal domain of ~130 residues, with a β-barrel topology, which has been shown in the case of AspRS from yeast to be responsible for specific recognition of the anticodon loop of $tRNA^{Asp}$ (Ruff *et al.*, 1991; Cavarelli *et al.*, 1993). A number of other proteins contain a domain with a similar topology which has become known as the OB fold (Murzin, 1993). In yeast AspRS this module is connected to the catalytic domain by a small hinge domain which also makes specific contacts with $tRNA^{Asp}$ in the D-stem region (Cavarelli *et al.*, 1993). The equivalent domain is disordered in the crystal structure of LysRS (LysU) from *E.coli* (Onesti *et al.*, 1995).

For each of the three class IIb synthetases, the anticodon triplet of the cognate tRNA is a major, but not exclusive, identity element (McClain *et al.*, 1990; Pütz *et al.*, 1991; Tamura *et al.*, 1992; Li *et al.*, 1993; reviewed in Saks *et al.*, 1994). On the other hand, the anticodons corresponding to aspartic acid (34-G/QUC, codons GAU, GAC), asparagine (34-G/QUU, codons AAU, AAC) and lysine (34-UUU and 34-CUU, codons AAA, AAG) all contain a central U, and discrimination between them often depends on only one nucleotide. The questions are thus posed, what is the structural basis for discrimination between these anticodons given that they are each recognized by a homologous binding domain? and secondly, by what mechanism do mismatches in the anticodon affect catalytic rates in the active site which is ~60 Å away?

An additional interest in the lysyl system is the common occurrence of a modified base at the wobble position (34) of $tRNA^{Lys}$. In *E.coli*, the unique $tRNA^{Lys}$ has anticodon 34-UUU with the wobble U modified to 5-[(methylamino)-methyl]-2-thiouridine (mnm^5s^2U). A certain proportion of the thiouridine is specifically modified further to 2-selenouridine (Wittwer, 1983). The same modification is found in the *E.coli* $tRNA^{Glu}$ and $tRNA^{Gln}$ with anticodons 34-UUC and 34-UUG respectively. All these tRNAs must avoid wobble pairing with codons ending with U, to avoid misincorporation of histidine, aspartic acid and asparagine instead of respectively glutamine, lysine and glutamic acid. The presence of the mnm^5s^2U modification is thought to disfavour the wobble pairing of U with both G and U through the rigidification of the backbone conformation (Yokoyama *et al.*, 1985; Yokoyama and Nishimura, 1995). In the case of lysine, the existence in *E.coli* of a unique $tRNA^{Lys}$ with anticodon 34-mnm^5s^2UUU correlates with the preferred codon usage, namely that 68% of lysine codons are AAA compared with AAG (Aota *et al.*, 1988). The situation is reversed in *T.thermophilus* which has AAG as the preferred lysine codon (>90%, Kagawa *et al.*, 1984); the only lysine-specific tRNA so far identified in this bacterium has anticodon 34-CUU (M.Tukalo, Z.Petruschenko, I.Kriklivy and A.Yaremchuk, unpublished results). Furthermore, it has been found that in *E.coli*, modifications to $tRNA^{Lys}$ are important for recognition by the lysyl-tRNA synthetase; an unmodified transcript has a reduction of 140 in specificity (Tamura *et al.*, 1992). In the case of $tRNA^{Lys}$, the particular modifications responsible for this effect have not yet been identified. On the other hand, in $tRNA^{Glu}$, it has been shown that mnm^5s^2U-34 is required for efficient aminoacylation by *E.coli* glutamyl-tRNA synthetase (Sylvers *et al.*, 1993). It has also been found that the s^2U modification at position 34 in human $tRNA_3^{Lys}$ is required for interaction with HIV RNA, particularly in the initiation step (Isel *et al.*, 1996).

The structure of the *E.coli* lysyl-tRNA synthetase (LysU) complexed with lysine has been determined by X-ray crystallography (Onesti *et al.*, 1995a) and the solution structure of the N-terminal domain from *E.coli* lysyl-tRNA synthetase (LysS) by NMR (Commans *et al.*, 1995a). We have reported previously the crystallization of the complex of both *E.coli* and *T.thermophilus* $tRNA^{Lys}$ with *T.thermophilus* lysyl-tRNA synthetase (Yaremchuk *et al.*, 1995a). Here we describe the crystal structure of the *T.thermophilus* lysyl-tRNA synthetase complexed with *E.coli* $tRNA^{Lys}$ at 2.75 Å resolution and with a *T.thermophilus* $tRNA^{Lys}$ transcript at 2.9 Å resolution. The structure was determined by molecular replacement using the *E.coli* lysyl-tRNA synthetase (LysU) as search model (Onesti *et al.*, 1995). The crystal structure reveals that the *T.thermophilus* enzyme is, as expected, rather similar to its *E.coli* counterpart (sequence identity 42%) although the relative orientations of the C- and N-terminal domains are different. However, in the crystal form studied, significant electron density for all but the anticodon stem–loop of the tRNA is largely absent, despite biochemical analysis showing that the tRNA is intact (Yaremchuk *et al.*, 1995a) and there being ample room in the crystal for a complete tRNA molecule (Figure 1A). Nevertheless, the electron density of the tRNA anticodon loop is very clear and shows that this region of the tRNA interacts with the N-terminal β-barrel domain of the synthetase in a similar but distinct manner to that observed in the yeast aspartyl system (Ruff *et al.*, 1991; Cavarelli *et al.*, 1993). Details of these interactions, which indicate how specificity for the $tRNA^{Lys}$ anticodon is achieved, are the major subject of the present study. We suggest that, under the conditions of crystallization, a pre-productive complex is formed in which the anticodon is docked correctly to the synthetase, but that the rest of the tRNA is not engaged fully with the synthetase and is consequently disordered. We also describe the low resolution crystal structure of the same complex in a different crystal form obtained by co-crystallization with a lysyl-adenylate analogue. This structure reveals substantial conformational changes in the catalytic domain induced by the binding of the analogue.



Results

Most of the detailed results presented here will be based on the refined model derived from 2.75 Å resolution cryo-crystallographic data of the hetero-complex of LysRSTT with *E.coli* tRNA^{Lys}(mnm^{5}s^{2}UUU). The model has excellent geometry and is refined to an *R*-factor of 18.9% (*R*-free = 22.7%) for all data between 2.75 and 12 Å using a solvent correction. Additional results will be presented on the complex of LysRSTT with an unmodified *T.thermophilus* tRNA^{Lys}(CUU) transcript at 2.9 Å resolution and on 3.8 Å resolution data of the ternary complex of LysRSTT with *E.coli* tRNA^{Lys} and a lysyl-adenylate analogue. Details of the data collections, structure determination and refinement are given in Materials and methods. A general view of the structure determined in this work is shown in Figure 1A.

Comparison of the structure of T.thermophilus lysyl-tRNA synthetase with that of E.coli

The LysRSTT molecule (492 residues/subunit) is structurally very similar to the *E.coli* LysRS (LysU) (Onesti *et al.*, 1995), the primary sequences of the two molecules being 42% homologous (Chen *et al.*, 1994). Both show a strong overall structural resemblance to yeast (Ruff *et al.*, 1991) and *T.thermophilus* aspartyl-tRNA synthetases (Delarue *et al.*, 1994) which are also members of the class IIb subfamily (Figure 1B).

Superposition of 207 Cα out of 339 in LysRSTT and 343 in LysRSECu (of which 124 are identical) in the C-terminal catalytic domain results in an root mean square (r.m.s.) deviation of 0.681 Å (using LSQMAN with cutoff 1.5 Å). However, as seen in Figure 1C, a further rigid-body rotation of 8.75° is then required to superpose the N-terminal anticodon binding domains (86/137 Cα with an r.m.s. deviation of 0.674 Å, 34 identities). The corresponding results for the comparison of LysRSTT and *T.thermophilus* AspRS are 120 Cα (out of 271 in LysRSTT and 361 in AspRSTT, 46 identical) in the C-terminal catalytic domain with an r.m.s. deviation of 0.885 Å. In this case, a further rigid-body rotation of 7° is then required to superpose the N-terminal anticodon binding domains (60/137 Cα with an r.m.s. deviation of 0.642 Å, 14 identities). It has already been pointed out (Onesti *et al.*, 1995) that the anticodon binding domains of *E.coli* LysU and yeast AspRS complexed with tRNA^{Asp} are related by a rigid-body rotation of ~20° and those of yeast and *T.thermophilus* AspRS by 8° (Delarue *et al.*, 1994). The conclusion, therefore, is that all these various aspartyl- and lysyl-tRNA synthetases have significantly different orientations of their N-terminal domains which are, in order from most closed to most open, LysRSU, LysRSTT, AspRSTT and AspRSSC. It is suggestive from the comparison of the two aspartyl enzymes (AspRSTT and AspRSSC) and the two lysyl enzymes (LysRSTT and LysRSECu) that species differences are important, although only two of these structures are complexed with tRNA (AspRSSC and LysRSTT). Since the structure of the free and tRNA complexed enzyme is unknown in all these cases, it is not yet clear whether some of the

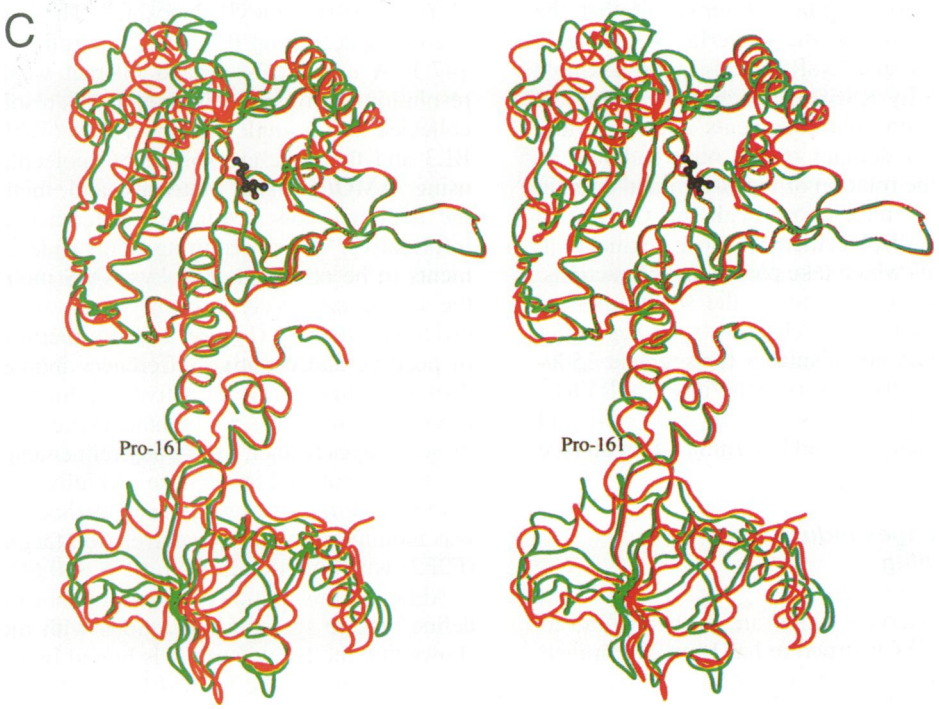


Fig. 1. (**A**) Stereo view of *T.thermophilus* LysRS complexed with a tRNA^{Lys}(CUU) transcript as determined from P42$_1$2 crystals in which the complete dimeric molecule is in the asymmetric unit. The monomers are coloured in cyan and yellow and the visible part of the tRNA anticodon stem–loop is coloured in red (backbone) and green (bases). The insertion domain of the yellow domain has been excluded to permit a view directly into the active site of this subunit. (**B**) Comparison of the three-dimensional structures of LysRSECU (left), LysRSTT (middle) and AspRSSC (right). The conserved core of the class II catalytic domain is shown in white, the N-terminal β-barrel domain characteristic of class IIb in red and the more variable insertion domain in yellow. (**C**) Superposition of Cα traces of LysRSTT (red) and LysRSECU (green) showing rigid-body-like displacements in the position of the insertion domain (top left) and the N-terminal domain (bottom). The lysine substrate is included from the LysRSECU structure (Onesti *et al.*, 1995). The position of Pro161 in LysRSTT at the N-terminal end of the long motif 1 helix is shown (see text).

Table I. Crystallographic data collection statistics for the various $tRNA^{lys}$ complexes of *T.thermophilus* lysyl-tRNA synthetase(LysRSTT)

Crystal contents	LysRSTT *E.coli* $tRNA^{lys}$	LysRSTT *E.coli* $tRNA^{lys}$	LysRSTT *T.thermophilus* $tRNA^{lys}$ transcript	LysRSTT *E.coli* $tRNA^{lys}$ Lys-AMS
Space group	$P42_12$ (#90)	I422 (#97)	$P42_12$ (#90)	$P3_121$ (#152)
Unit cell parameters (Å)	$a = b = 233.0$	$a = b = 234.3$	$a = b = 232.6$	$a = b = 147.3$
	$c = 118.1$	$c = 116.7$	$c = 117.2$	$c = 127.4$
Resolution range (Å)	20–2.75	20–2.75	18–2.9	20–3.9
Temperature (K)	100	100	110	100
Cryoprotectant	30% glycerol	30% glycerol	30% glycerol	30% glycerol
ESRF beamline	BL4 (ID2)	BL4 (ID2)	BL19 (BM14)	BL3 (ID13)
Detector	30 cm Mar	30 cm Mar	30 cm Mar	30 cm Mar
Exposure time/° (s)	40	40	170	200
Total # reflections	208 185	132 767	272 770	51 969
Unique reflections	78 327	40 304	69 586	14 794
Average redundancy	2.7	3.3	3.9	3.5
Completeness (%)	93.1	95.5	97.5	99.1
R_{merge} (highest bin)	0.038 (0.205)	0.046 (0.185)	0.047 (0.220)	0.097 (0.316)

The data in C is for a crystal of the complex of *T.thermophilus* lysyl-tRNA synthetase with a *T.thermophilus* $tRNA^{Lys}$ transcript. All data were integrated with the MOSFLM program (Leslie, 1992)

differences reflect conformational changes associated with tRNA binding.

In the case of LysRSTT and LysRSECu, it is interesting to determine the structural basis for the different relative orientation of the C- and N-terminal domains, this being possible due to the relatively high homology of the two structures. The orientation of the N-terminal domain is determined by inter-domain interactions with the C-terminal domain of both subunits in the dimer (see Figure 1A). For this reason, crystal packing effects upon orientation are likely to be small. It turns out that the detailed interactions making up these interfaces are largely conserved in LysRSTT and LysRSECu, a good example being Arg59t (74u in LysRSECu) which makes identical hydrogen bonds to main chain segments 232t–234t and 448t in the C-terminal domain of the other subunit, in both stuctures. Thus the rotation of the N-terminal domain is not due to different interactions with the C-terminal domains, but due to a displacement of those segments in the C-terminal domains which take part in the interactions. A key element in this appears to be the substitution of Pro161t in place of Asp178u which distorts the beginning of the long motif 1 helix and displaces the segment 150t–166t (containing the strictly conserved motif RYRQRYRL-DLI), which in turn displaces segments 230t–234t and 440t–450t in the symmetry-related C-terminal domain (see Figure 1C).

Conformational changes induced by lysyl-adenylate binding

Some other significant differences between LysRSTT and LysRSECu can be observed which are probably due to the fact that the LysRSECu structure has been determined in the presence of bound lysine. The loop corresponding to 215u–220u (198t–203t) containing Gly217u, whose carbonyl oxygen binds to the α-amino group of the substrate lysine (Figure 10 of Onesti *et al.*, 1995), moves away from the active site in LysRSTT. This in turn pushes the helical hairpin 427t–437t away from the active site, the extremity 427t–434t being disordered in the LysRSTT structure but it is ordered in the LysRSECu structure. The insertion domain (residues 310t–402t) is of a similar topology in LysRSTT and LysRSECu but differs in orientation by a rotation of ~11° (Figure 1C). Furthermore, an extended loop only present in the LysRSECu insertion domain (360u–368u) comes in close proximity to the extremity of the helical hairpin (448u), perhaps helping to stabilize the latter.

Some direct evidence that these differences are induced by lysine binding comes from data on crystals of LysRSTT co-crystallized with *E.coli* $tRNA^{Lys}$ and a non-hydrolysable sulfamoyl analogue of lysyl-adenylate (Lys-AMS), (5′-*O*-[*N*-(L-lysyl)-sulfamoyl]adenosine). These crystals are of trigonal space group $P3_12_1$ with cell dimensions $a = b = 147.3$, Å $c = 127.4$ Å and diffract weakly to ~3.8 Å resolution. A complete data set to this resolution has been collected on a single frozen crystal (Table I) on ESRF BL3 and the structure solved by molecular replacement using AMORE (see Materials and methods). Despite the relatively low resolution, difference maps are very informative, even permitting some side chain displacements to be observed. The Lys-AMS molecule bound in the active site is very clearly seen in strong positive difference density (Figure 2A). Furthermore, the pattern of positive and negative differences indicates substantial displacements in several active site loops, and the entire insertion domain has re-orientated (Figure 2A). To define these changes further, rigid-body refinement using XPLOR has been employed to displace carefully chosen fragments of the structure (see Materials and methods). The procedure was monitored by the free correlation factor on intensities (F2F2) which increased from 0.56 to 0.66.

Although this partially refined model cannot be used to define precise specific interactions with the Lys-AMS, it shows that the lysine moiety is bound in a way consistent with the lysine in the LysRSECu–lysine structure (Figure 10 of Onesti *et al.*, 1995). Furthermore, the AMS moiety is bound in a way consistent with that observed in other class II synthetases (Belrhali *et al.*, 1994, 1995; Poterszman *et al.*, 1994). In particular, the loop 198t–203t (see above) moves into the position observed in the LysRSECu–lysine structure, and there are indications that the hairpin 427t–437t accompanies this movement. Secondly, the motif 2 loop moves into a position in which the discriminatory

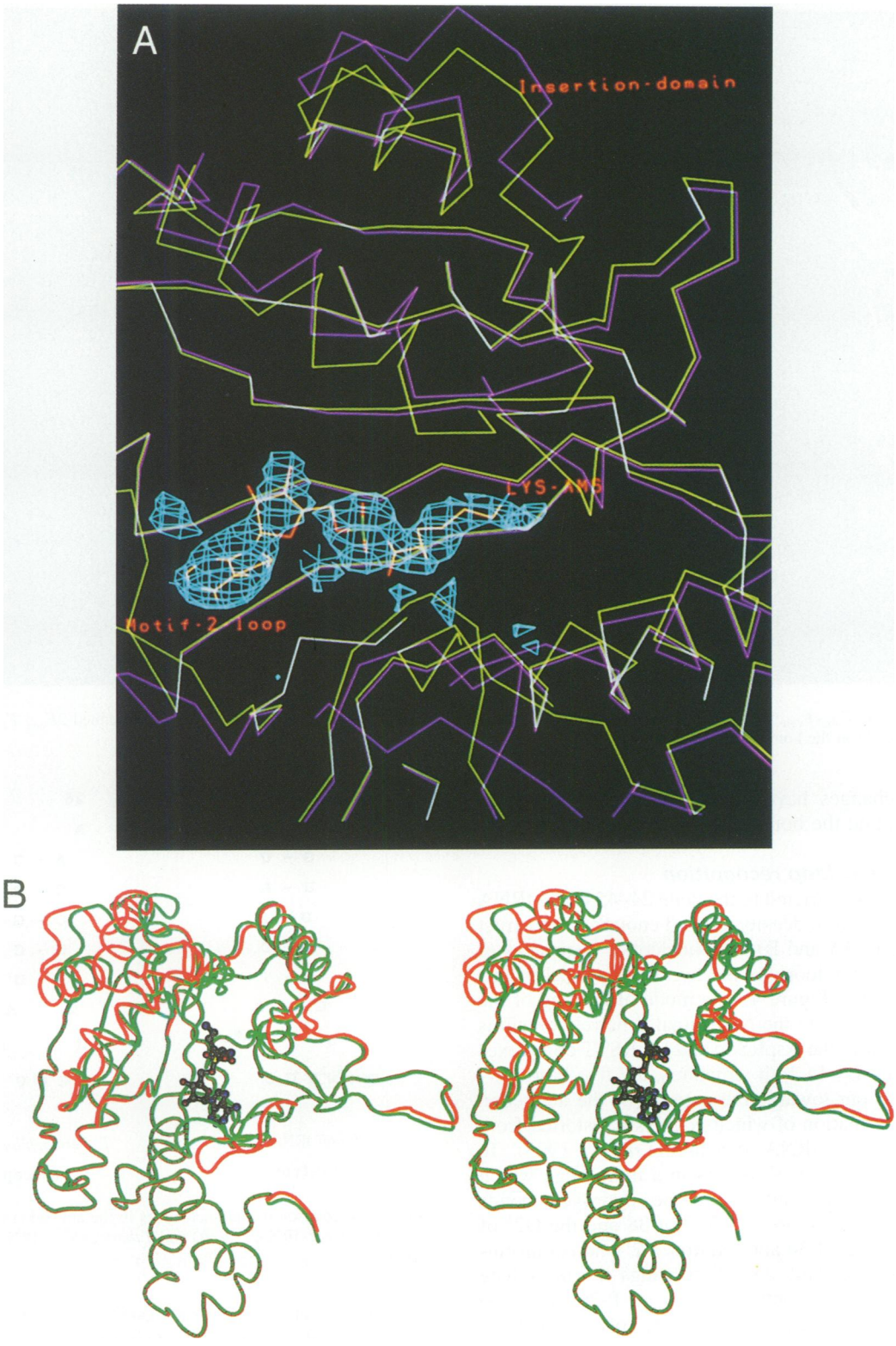


Fig. 2. (**A**) Comparison of Cα trace of the catalytic domain before (purple) and after co-crystallization with the lysyl-adenylate analogue (green). Positive difference electron density for the analogue is shown in blue, contoured at 3 σ. The difference map is calculated after rigid-body refinement (see text) and without ever including the analogue in the model. (**B**) Superposition of the Cα trace of the catalytic domain of LysRSTT without (red) and with (green) the co-crystallized lysyl-adenylate analogue. The orientation is the same as in Figure 3C, and comparison of the two figures shows that the position of the insertion domain and other active site loops in the lysyl-adenylate analogue complex of LysRSTT more closely resembles that of the lysine complex with LysRSECU.

main chain interactions can be made with the adenosine base as observed, for instance, in seryl-tRNA synthetase (Belrhali *et al.*, 1994, 1995). Thirdly, the entire insertion domain (residues 312–368) rotates by ~6° in a direction bringing it closer to the position observed in the LysRSECu–lysine structure (Figure 2B). All these con-

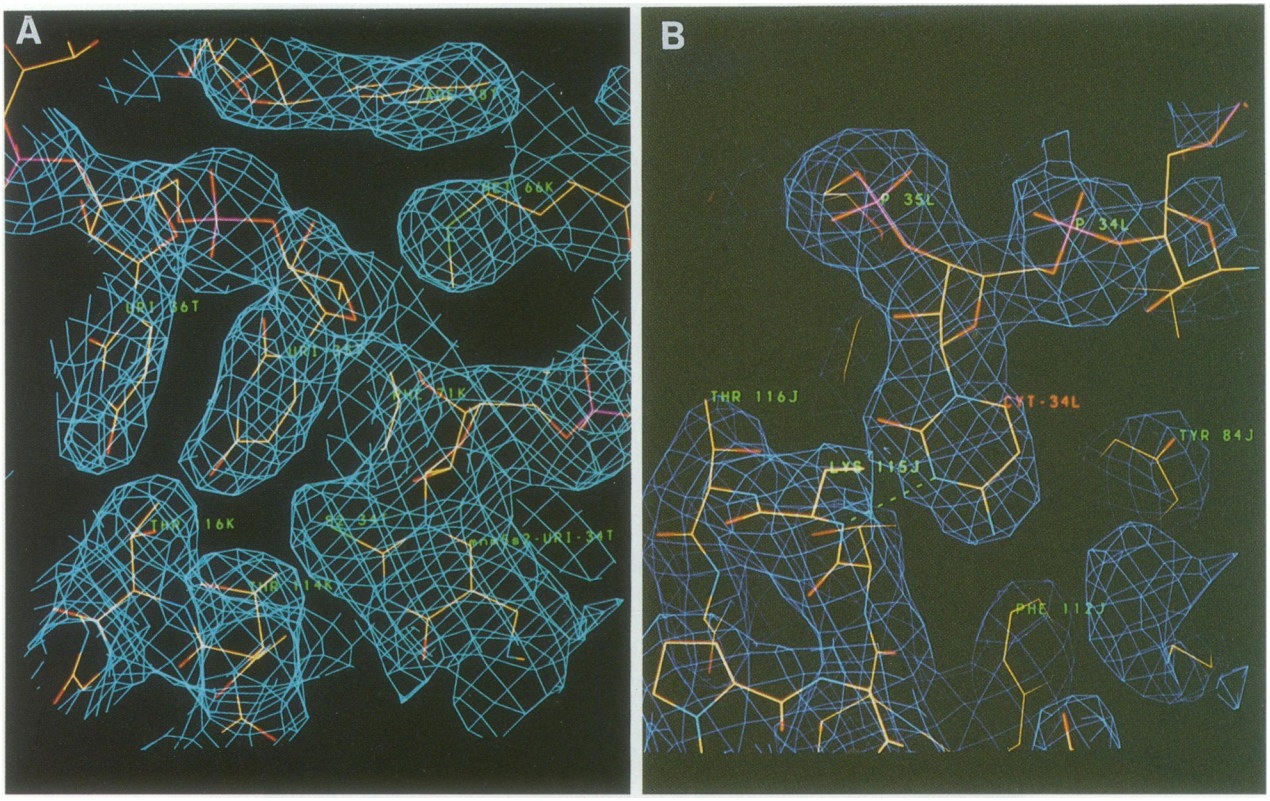


Fig. 3. (**A**) Refined $2F_{obs}$–F_{calc} electron density of anticodon loop bases in the hetero-complex contoured at 1 σ. (**B**) Refined $2F_{obs}$–F_{calc} electron density of base C-34 in the homo-complex contoured at 1 σ.

formational changes have the effect of closing up the active site around the bound Lys-AMS molecule.

Anticodon stem–loop recognition

The discussion is restricted to the zone 24–45 of the tRNA for which the electron density is good enough to construct a model (Figure 3A and B). The nucleotide sequences of the anticodon stem–loop for the two tRNAs used in this study are shown in Figure 4. The mode of binding of the anticodon stem–loop to the N-terminal β-barrel domain is basically similar in the AspRSSc and LysRSTT complexes (Figure 5A and B). In both structures, the five bases 33–37 are splayed out towards the exterior of the anticodon loop, the conformation of which is grossly distorted from that found in free tRNA (Cavarelli *et al.*, 1993). In particular, nucleotide A-37 occurs in a tight bulge which is stabilized by two intra-backbone hydrogen bonds between O2′ of ribose-36 and O1P of 38 and the O2′ of ribose 38 and O2P of 36 and requires a 2′-endo conformation of ribose-37. This distortion brings the phosphate P-39 to a distance of only 7.4 Å from P-26. The two bases 33 and 37 are fully exposed to the solvent and have poor electron density and, in particular, the modification of base A-37 to t^6A is not seen. On the other hand, the three anticodon bases 34–36 and base-38 which interact intimately with the synthetase have very good electron density (Figure 3A and B).

Base-specific interactions

In the LysRSTT complex, base-specific interactions are restricted almost exclusively to the anticodon bases. The tRNA backbone conformation of this region is closely

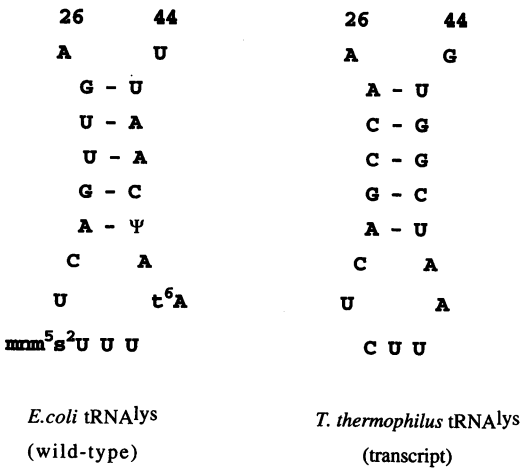


Fig. 4. Comparison of the nucleotides in the anticodon stem–loop of wild-type *E.coli* $tRNA^{Lys}$ (left) and *T.thermophilus* $tRNA^{Lys}$ transcript (right).

similar to that found in the AspRSSc complex, except in the region of nucleotide 34 (Figure 5B).

Base 34. In the majority of species of *T.thermophilus* $tRNA^{Lys}$, base 34 is a cytosine, whereas in *E.coli* $tRNA^{Lys}$, the tRNA in the complex described here, it is a modified uracil, mnm^5s^2U-34. It is not known whether this modified base exists in some minority isoacceptors of *T.thermophilus* $tRNA^{Lys}$ and, therefore, it is possible that the interaction of mnm^5s^2U-34 with LysRSTT does not occur naturally. On the other hand, the residues forming the immediate environment of this base are largely conserved

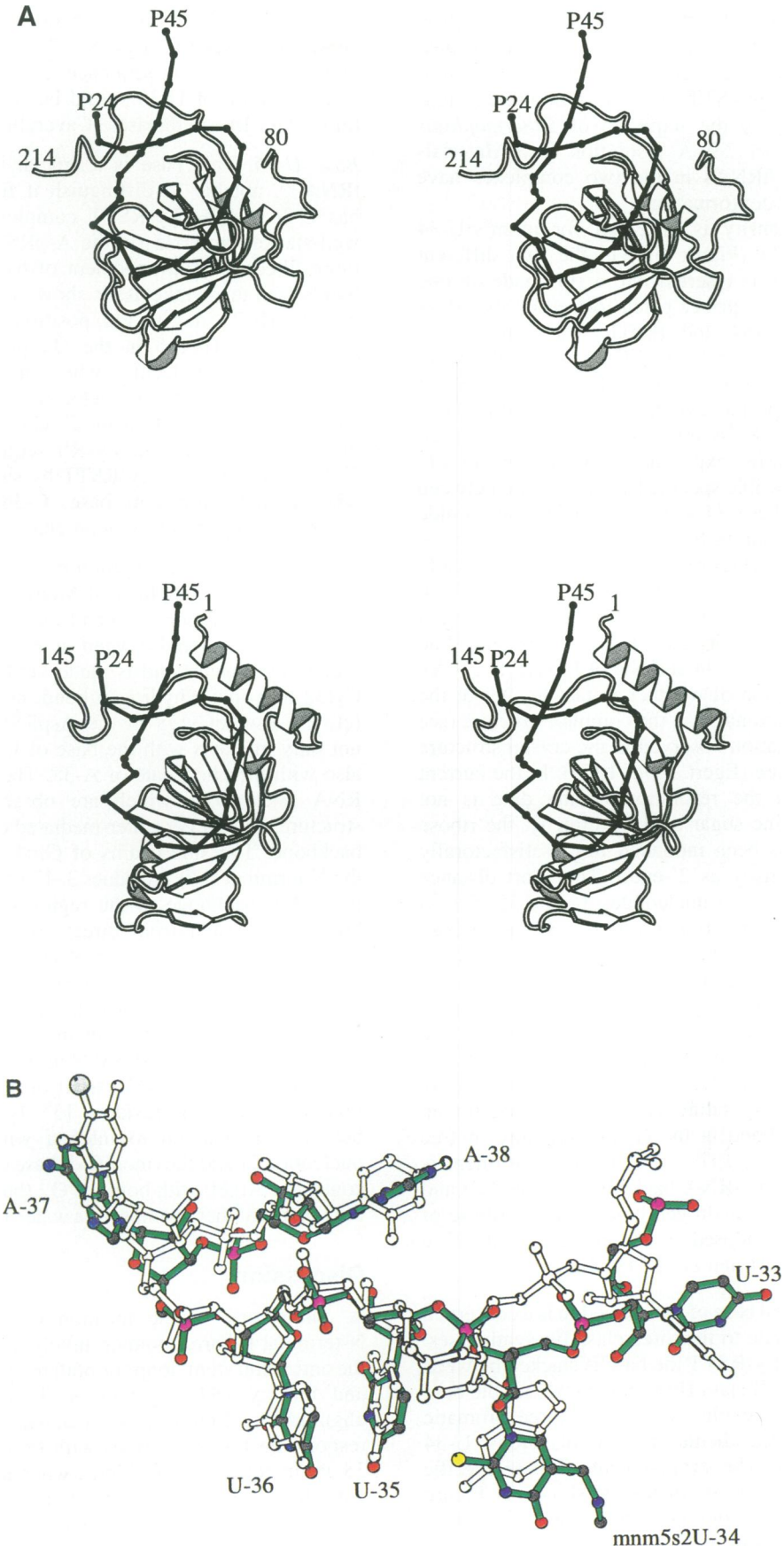


Fig. 5. (**A**) Comparison of the binding of the anticodon stem–loop to the N-terminal β-barrel domain in AspRSSC (top) and LysRSTT (bottom). (**B**) Superposition of the model for bases 33–38 for LysRSTT complex (green) and AspRSSC complex (white). Base identities refer to *E.coli* tRNA^{Lys}.

between LysRSTT and LysRSECu (Figure 7). Furthermore, the relevance of the current structure is confirmed by the determination of the crystal structure of a homocomplex between LysRSTT and an unmodified transcript corresponding to the sequence of *T.thermophilus* $tRNA^{Lys}$ (34-CUU) at 2.9 Å resolution (see Materials and methods). The tRNAs in the two complexes have essentially identical conformations.

Clear electron density is observed for mnm^5s^2U-34 (Figure 3A) or C-34 (Figure 3B) in the two different complexes. The base is inserted into a slot made on one side by a hydrophobic surface patch formed by the edges of U-35, Phe71, Tyr84 and Phe112 (the amino acid residues are conserved in LysRSECu) and on the other side by the aliphatic side chain of Lys115 (Gln131 in LysRSECu). The importance of this cluster of hydrophobic surface residues for tRNA interaction has been deduced previously from NMR experiments (Commans *et al.*, 1995a). The only possible specific hydrogen bond between the synthetase and base 34 is via the main chain amide of Lys115 which is within hydrogen bonding distance of the N3 of the base (Figure 6C). The N3 is normally supposed to be protonated in the case of mnm^5s^2U-34, but not for C-34 where clearly a favourable hydrogen bond interaction occurs in the observed structure (the distance between N3 of C-34 and NH of Lys115 is 2.9 Å). This raises the question of whether in fact the N3 of the mnm^5s^2U-34 is unprotonated in the complex structure (see Discussion), the situation observed in the crystal structure of mnm^5s^2U dihydrate (Egert *et al.*, 1979). In the current structures, although the resolution of the data is not sufficient to determine sugar puckers directly, the ribose of nucleotide 34 has been modelled most satisfactorally into the electron density as 2′-endo. The short distance between the phosphates of nucleotides 34 and 35 (5.3 Å) is one reason why a conformation with a 3′-endo ribose is a much less satisfactory fit to the electron density. However, the poor electron density for nucleotide U-33 suggests flexibility in this region which adds some uncertainty about the exact backbone conformation. In the modelled 2′-endo conformation for the *E.coli* $tRNA^{Lys}$ complex, the distance S2–O2′ is 3.3 Å, which is slightly longer than the corresponding distance observed for an S2–water hydrogen bond in the crystal structure of free mnm^5s^2U (Egert *et al.*, 1979). The mnm^5 substituent is directed away from the tRNA backbone (Figure 6C) and clearly does not form a hydrogen bond with the ribose of U-33, as has been proposed in the context of the free tRNA conformation (Hillen *et al.*, 1978).

Base U-35. The central base of the anticodon is a conserved U in all tRNAs cognate to the three class IIb synthetases. In the complex with LysRSTT the base is stacked between Phe71 (conserved in all class IIb synthetases) and the base U-36 (Figure 6B). The planes of these three aromatic rings are almost perpendicular to that of mnm^5s^2U-34 (Figures 3 and 5B). The arrangement of base-specific hydrogen bonds with the synthetase is shown in Figure 6B and includes Arg64 (conserved in all class IIb synthetases) in the O2 position, Gln82 (conserved in all class IIb synthetases) in the N3 and O4 positions, and Thr114 in the O4 position. All cited residues are absolutely conserved in all known lysyl-tRNA synthetase sequences (except for a conservative substitution of Thr114→Ser in yeast mitochondrial LysRS). This result confirms the hypothesis based on sequence alignments that the mode of recognition of U-35 would be conserved between the three class IIb synthetases (Caverelli *et al.*, 1993).

Base U-36. This base is a crucial identity element for $tRNA^{Lys}$, notably to distinguish it from $tRNA^{Asp}$ which has C-36. In the LysRSTT complex structure, U-36 is well-stacked on U-35 (in the AspRSSc complex, C-36 is more tilted). The arrangement of base-specific hydrogen bonds with the synthetase is shown in Figure 6A and they include His73 in the O2 position, Glu118 in the N3 position and Thr116 in the O4 position. A very well ordered water molecule (which makes three hydrogen bonds with the protein) helps position the key residue which distinguishes U from C, Glu-118 which occurs in the majority of known LysRS sequences. Interestingly, U-36 is identified by LysRSTT by side chain interactions whereas the equivalent base, C-36, is recognized by AspRSSc principally by main chain hydrogen bonds.

Other protein–RNA interactions. The base of A-38 is stacked on the side chain of Met6 (equivalent to Gln121 in AspRSSc) (Figure 3A) and the N6 makes a hydrogen bond to the carbonyl oxygen of Arg65. It is likely that a second hydrogen bond is made via this N6 to the N3 of Cyt32 in a single hydrogen bond, non-standard base pair (cf. the Psu32:Cyt38 in the AspRSSc complex). Arg64 not only interacts with the base of U-36 (see above), but also with the phosphate of A-38. The only other protein–RNA interactions which are observed in the present structure are indirect water-mediated contacts to the tRNA backbone. The side chains of Gln4, Gln7 and Asn11 in the N-terminal helix (residues 3–17 of LysRSTT) approach the tRNA backbone in the region of nucleotides 27–28 but are separated from direct contact by a number of partially ordered water molecules. Interestingly, this N-terminal helix is a distinguishing feature between LysRS and AspRS (Figure 4A) and appears to substitute to some extent for the hinge domain in AspRSSc, which makes backbone contacts to $tRNA^{Asp}$ in the region of nucleotide 27 as well as 10–11 (Cavarelli *et al.*, 1993). The hinge region in LysRSTT (residues 137–145) is poorly ordered but is in a position to interact with the backbone of nucleotide 11 and the (modified) base of A-37. The indirect contact of Arg80 with both the O2 (base) and O2′ (ribose) of U-36 is another example of a water-mediated interaction.

Discussion

A comparison of the location of the residues of the N-terminal β-barrel domain involved in interactions with the anticodon stem–loop for both the yeast aspartyl system and the LysRSTT system is shown in the sequence alignment in Figure 7. It is apparent that the residues responsible for interactions with tRNA bases 35, 36 and 38 are in equivalent positions, whereas those that interact with the wobble base 34 are more variably situated. The alignment also shows, as discussed above, that the most important residues are conserved between LysRSTT and LysRSECu, and indeed with other known LysRS sequences, so that the conclusions about anticodon recognition are probably generally applicable to all LysRS.

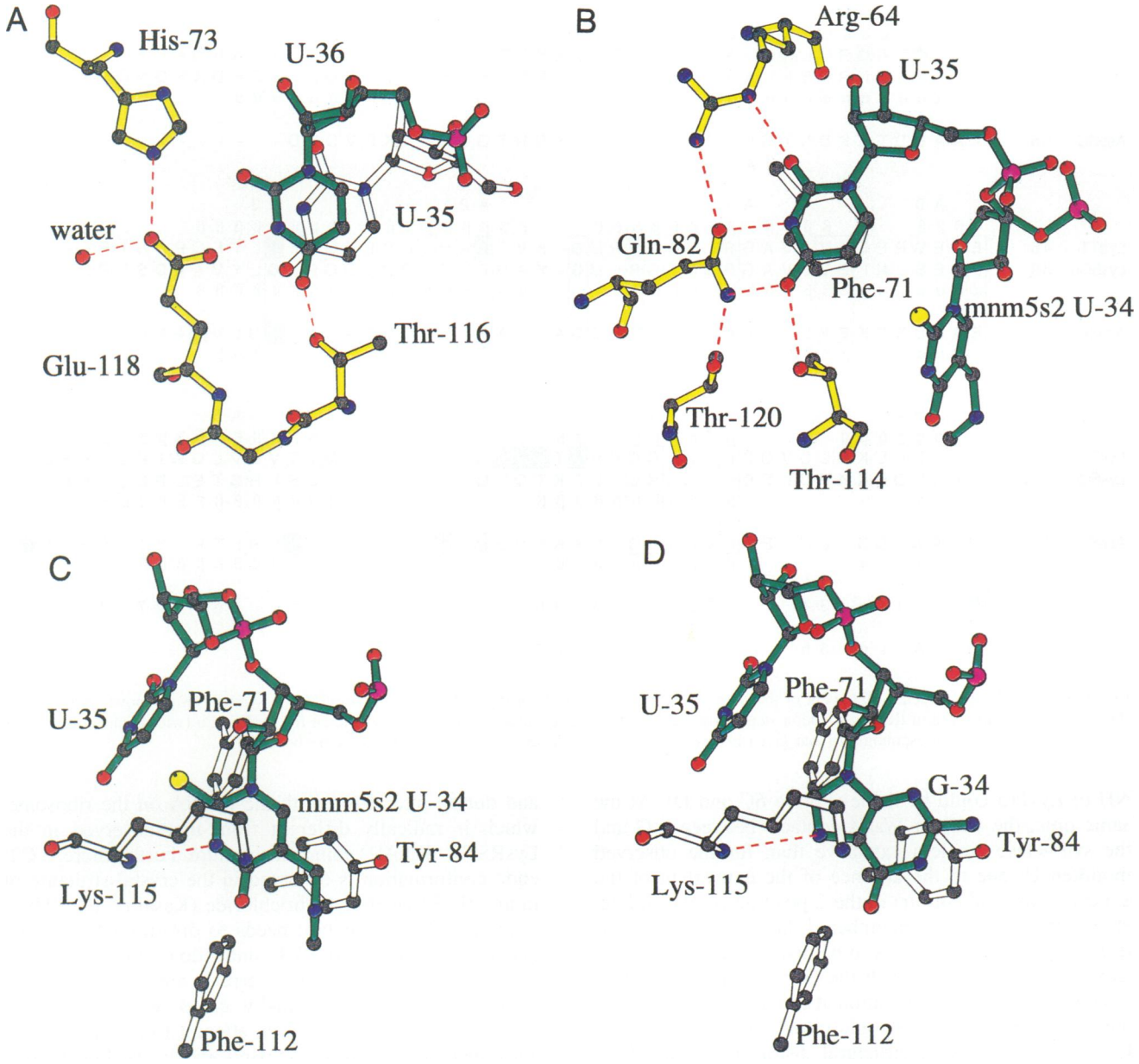


Fig. 6. Detailed interactions of the three anticodon bases of tRNA^{Lys} with LysRSTT. (**A**) U-36, (**B**) U-35, (**C**) mnm5s2U-34. In (**D**) a G-34 has been modelled in place of the mnm5s2U-34 (see text). Red dotted lines indicate putative hydrogen bonds.

The results of McClain *et al.* (1990) and Tamura *et al.* (1992) have identified that the major identity elements of tRNA^{Lys} in the *E.coli* system are the three anticodon bases (particularly U-35) and, to a lesser extent, the discriminator base A-73. Two additional features emerge from the work of Tamura *et al.* (1992), namely that a transcript of *E.coli* tRNA^{Lys} is 140-fold less active than the fully modified wild-type and, secondly, that for essentially all mutant tRNA^{Lys} transcripts tested, only V_{max} changes and not the apparent K_m. Do the structural results presented here explain these aspects of the specificity of LysRS for tRNA^{Lys}? Due to the fact that the tRNA is only partially ordered in the complex structure, we are only able to discuss this question from the point of view of the anticodon. As regards the bases 35 and 36, the specific hydrogen bond interactions with the LysRSTT guarantee that these are uridines (Figure 6A and B). This is in agreement with mutagenesis experiments performed using *E.coli* tRNA^{Lys} transcripts and LysRSEC in which it was found that U35→C or A and U36→C or A abolishes measurable activity and that U36→G leads to a 5-fold decrease in relative V_{max}/K_m without changing the K_m (Tamura *et al.*, 1992). With regard to base 34, the mutagenesis experiments show that the single substitution U34→G leads to a 17-fold decrease in relative V_{max}/K_m, again without changing the K_m (Tamura *et al.*, 1992). From the crystal structure, it is not so obvious how the synthetase discriminates against purines at this position since there is no apparent steric or chemical hindrance to substituting an unmodified G or A into the complex model in this position. In doing this, one places the N7 position of the purine in nearly the same position as that observed for the N3 position of the pyrimidine from which a favourable hydrogen bond interaction with the main chain

Fig. 7. Structure-based sequence alignment of the N-terminal β-barrel domain for three class IIb aminoacyl-tRNA synthetases of known structure showing residues involved in tRNA anticodon stem–loop recognition. The secondary structure is also shown for LysRSECu (with nomenclature of the strands and helices corresponding to that given in Onesti *et al.*, 1995). Absolutely conserved residues are boxed.

NH of Lys115 could be made (Figure 6C and D). At the same time, the van der Waals contacts between a G and the synthetase are less extensive than for the observed modified U, due to the absence of the equivalent of the exocyclic oxygen (sulfur) in the 2-position of pyrimidine. However it should be remembered that, *in vivo*, the most relevant tRNA which needs to be discriminated against is $tRNA^{Asn}$(34-GUU) in which the G is generally modified to Q with its bulky modification at exactly the N7 position mentioned above; LysRS could clearly not accommodate Q-34 without a severe structural distortion (Figure 6D). Thus Q-34 in $tRNA^{Asn}$ is a negative determinant for LysRS which is reinforced by the difference in discriminator base (G in $tRNA^{Asn}$ and A in $tRNA^{Lys}$). This agrees with the conclusions of Li *et al.* (1993) using *in vivo* expressed mutants of $tRNA^{Asn}$ who also showed that $tRNA^{Asn}$(Q→CUU) is almost exclusively charged by LysRS.

In this discussion a further interesting element is the observed C2′ endo ribose pucker for the mnm^5s^2U-34 in the complex structure, which is associated with the locally very different tRNA backbone conformation in $tRNA^{Lys}$ compared with $tRNA^{Asp}$ around position 34 (Figure 4B). This pucker has been shown to be unfavourable compared with the C3′ endo conformation for free mnm^5s^2U nucleotides in solution and in uncomplexed tRNA (Yokoyama *et al.*, 1985). It has been argued that this is due to a steric clash of the bulky thiocarbonyl (s^2) group with the equatorial 2′ OH in the C2′ endo conformation and that this is important for restricting the de-coding ability at the ribosome (Yokoyama *et al.*, 1985, 1995). However, these arguments depend strongly on the presumed stacked conformation of the anticodon nucleotides in free tRNA and during interaction with the mRNA on the ribosome, which is radically different from that observed in the LysRSTT–$tRNA^{Lys}$ complex structure. Furthermore, a C2′ endo conformation is observed in the crystal structure of mnm^5s^2U-34 acetone hydrochloride (Kasai *et al.*, 1979).

Lysyl-tRNA synthetase needs to discriminate strongly against an unmodified UUU anticodon, since this could lead to missense translation errors, and seems to have evolved to function optimally either with a CUU or mnm^5s^2UUU anticodon. In an attempt to understand this and other observations discussed above, we hypothesize that with a C or mnm^5s^2U at position 34, a particular anticodon loop conformation is induced upon initial docking with the synthetase as observed (notably involving the C2′ endo ribose-34). Further, we suppose that this is necessary for correct induced fit interactions of other regions of the tRNA with the synthetase, this effect propagating, like the closing of a zip, to give a productive engagement in the active site. We further hypothesize that part of the role of the s^2 modification is, by virtue of its interaction with the O2′ of the ribose-34, to lower the p*K* of the N3 hydrogen, thus favouring deprotonation. The p*K* of 2-thiouridine is 8.8, compared with 9.2 for uridine, whereas that of mnm^5s^2U is not reported, although the crystal structure of the dihydrate at neutral pH has N3 deprotonated (Egert *et al.*, 1979). This would make the protonation state of the N3 position of mnm^5s^2U-34 resemble that of C-34, both then being able to make favourable hydrogen bonds with the main chain amide of Lys115. In this scenario, neither unmodified U (because it has a protonated N3) nor a pyrimidine (because there is no substituent in the equivalent position to the 2-position

of the purine, Figure 6D) would be capable of binding in such a way as to induce the correct conformation of the anticodon loop backbone, although the binding affinity may be little changed. This might explain why, in these cases, the K_m is unchanged, but V_{max} strongly decreases, since the zip cannot close properly. This hypothesis would predict that a 34-CUU transcript of either *E.coli* tRNA^{Lys} or *T.thermophilus* tRNA^{Lys} should behave much more similarly to wild-type than observed for the 34-UUU transcript. Preliminary biochemical results on the comparative activity of *T.thermophilus* lysyl-tRNA synthetase with wild-type *E.coli* tRNA^{Lys}, de-thiolated *E.coli* tRNA^{Lys} and *E.coli* tRNA^{Lys} or *T.thermophilus* tRNA^{Lys} transcripts with a CUU anticodon are in accordance with these suggestions (M.Tukalo, unpublished results) and lend weight to the idea that the lack of the U-34 modification is an important factor in the poor activity of *E.coli* tRNA^{Lys}(UUU) transcripts. In addition, we expect that a UUU or GUU anticodon should give rise to a different anticodon loop conformation in a crystal structure compared with CUU or mnm^{5}s^{2}UUU. We are in the process of testing these hypotheses by co-crystallizing LysRSTT transcripts with anticodon mutations of *E.coli* and *T.thermophilus* tRNA^{Lys}.

An important additional observation reported here is the significant conformational changes induced upon the binding of the lysine adenylate analogue to LysRSTT, not only, as expected, in the motif 2 loop, but also in the orientation of the insertion domain and other active site loops. These changes bring the active site conformation much closer to that observed in the LysRSECu complex with lysine, strongly suggesting that the closing up of the active site upon substrate binding has functional importance. The observation that LysRSECu only crystallizes in the presence of lysine (Onesti *et al.*, 1995) also suggests that lysine binding induces a different conformation of the synthetase. In the case of adenylate binding to SerRSTT (Belrhali *et al.*, 1994, 1995) and AspRSTT (Poterzman *et al.*, 1994), observed structural changes are much less significant. A recent study of a point mutation (Thr208→Met) in motif 1 in the intersubunit interface of *E.coli* LysRS (LysS) shows that the mutant enzyme, unlike the wild-type, shows highly cooperative binding of lysine to the two subunits (Commans *et al.*, 1995b). This behaviour can be explained by assuming the existence of two states of each subunit, only one of which can bind lysine. Such could be the case for the wild-type, although, perhaps due to the modified intersubunit interface, only in the mutant does an allosteric effect occur. Interestingly, both *E.coli* HisRS (Francklyn *et al.*, 1994) and *T.thermophilus* HisRS (Yaremchuk *et al.*, 1995b) only crystallize in the presence of the substrate histidine, again suggesting that amino acid binding induces a different conformation of the synthetase. Although, in all these cases, the effect of crystal packing forces is unlikely but cannot be ruled out, the conclusion seems to be that class II synthetases differ markedly in the extent to which amino acid binding induces significant structural changes. This raises the question of to what extent tRNA binding and docking are modified by the presence of the amino acid or adenylate in the active site. In the case of *T.thermophilus* seryl-tRNA synthetase, recent crystallographic results on a ternary complex of the enzyme with tRNA^{Ser} and a seryl-adenylate analogue (Cusack *et al.*, 1996) show an ordering of the acceptor stem of the tRNA in the presence of the adenylate compared with the structure without adenylate (Biou *et al.*, 1994). In addition, the extra long motif 2 loop in SerRS is shown to take up two very different conformations depending on whether the tRNA is ordered in the active site or not (Cusack *et al.*, 1996). These results show that the productive docking of tRNA to its cognate synthetase is more subtle than previously thought and may depend on conformational changes induced by binding of the other substrates. In the case of the lysyl-tRNA synthetase complex, we have apparently crystallized an initial docking state of the complex in which the anticodon, but not the acceptor stem, is engaged. Using the analogy introduced in the discussion above, the zip has not been pulled up to the top. The reason for this is not clear and, furthermore, it is puzzling as to why the situation is not improved in the ternary complex with the lysyl-adenylate complex. Further structural studies are required to resolve these outstanding issues.

Materials and methods

Crystals, space group changes, data collection and data quality

Crystals of the *T.thermophilus* lysyl-tRNA synthetase–*E.coli* tRNA^{Lys} hetero-complex were grown as described using wild-type LysRSTT and overexpressed synthetase–*E.coli* tRNA^{Lys} (Yaremchuk *et al.*, 1995a). Data have been collected on a large number of these crystals under various conditions, notably at both room temperature in capillaries and frozen at cryo-temperatures, and with and without soaked or co-crystallized small substrates (lysine, ATP or both). The best quality data at 2.75 Å have been obtained from large crystals of up to 1 mm in length (corresponding to the crystallographic *c*-axis) and of cross-section 0.35×0.35 mm^2, flash frozen to 100 K after transferring in four steps to mother liquor containing 30% glycerol. Images at 0.7° were measured using the extremely intense undulator beam of the high brilliance beamline (ID2/BL4) at the ESRF. Under these conditions, ~10° of data can be collected per crystal position before radiation damage is significant, the length of the crystal permitting a full data set from a single crystal. At room temperature on BL4, only a single high resolution image can be taken before severe radiation damage is visible. Data at 2.75 Å resolution were obtained on BL4 at 100 K on two crystal forms, one native complex crystal which had unit cell dimensions at 100 K of $a = b = 233$ Å, $c = 119$ Å and of space group I422 and a second crystal that was soaked at room temperature for a few minutes in the presence of lysine, ATP and Mn^{2+}, which had very similar cell dimensions but was of space group P42$_1$2 (Table I). In subsequent experiments, it has become apparent that (i) native complex crystals grown above pH 7.0 are of space group I422 at room temperature, whereas those grown at pH 6.8 are P42$_1$2 at room temperature, (ii) soaking or co-crystallization with lysine or lysine and Mg^{2+}-ATP causes a space group change from I422 to P42$_1$2 without significant change in cell dimensions, and (iii) the space group change can sometimes occur upon freezing of a native complex crystal in 30% glycerol. A further observation is that reasonable quality 3.8 Å data can be collected at room temperature on a rotating anode source. This has been done for native and heavy metal-soaked crystals, all in the I422 space group.

Co-crystallization of the ternary complex of LysRSTT, *E.coli* tRNA^{Lys} and the sulfamoyl analogue of lysyl-adenylate (5′-*O*-[*N*-(L-lysyl)-sulfamoyl]adenosine, Lys-AMS) leads to a new trigonal crystal form. The crystallization conditions are 24–26% saturated ammonium sulfate, 50 mM Tris–maleate (pH 7.6), 10 mM $MgCl_2$, 1 mM NaN_3, 4 mg/ml LysRSTT, 2.5 mg/ml *E.coli* tRNA^{Lys} and 325 μM Lys-AMS. The crystals are hexagonal prisms which grow up to 150 μm thick and 600 μm long. These crystals are of trigonal space group P3$_1$21 (#152) with cell dimensions $a = b = 147.3$ Å, $c = 127.4$ Å and diffract weakly to ~3.8 Å resolution. Data have been collected on beamline BL3 at the ESRF on a single frozen crystal, which was dipped rapidly into mother liquor containing 30% glycerol before being frozen (Table I).

Table II. Statistics of the model of *T.thermophilus* LysRS/*E.coli* $tRNA^{lys}$ complex refined in the I422 space-group

LysRSTT residues included (q = 1)	475/505	
LysRSTT residues disordered (q = 0)	138–142, 200–202, 426–436, 485–505	
$tRNA^{Lys}$ residues included	34–38 (occupancy q = 0.7) 24–45 (occupancy q = 0.5)	
Water molecules included	266	
Multiple conformations	Arg175, Glu34, Glu331	
Refinement statistics using all data between	2.75–7 Å	2.75–12 Å (with bulk solvent)
Total reflections used	37 938	39 786
R-free (10%)	24.8	22.8
R-work	20.5	18.9
R.m.s. deviations from ideal geometry		
Bond lengths (Å)	0.011	0.009
Bond angles (°)	1.711	1.653
Average B-factor (Å^2)		
LysRSTT	51.0	
Water	54.7	
tRNA	64.8	

Table III. Progress of the fragment refinement of the ternary complex of *T.thermophilus* LysRS/*E.coli* $tRNA^{lys}$/Lys-AMS as monitored by the free correlation on intensities (F2F2) over the resolution range 4.3–8 Å

	Initial molecular replacement solution	Global rigid-body refinement	Three rigid bodies	Ten fragments	Manual adjustment
Correlation (93%, 8634 refl.)	0.486 (R = 43.7)	0.508	0.523	0.628	0.654 (R = 36.5)
Free correlation (7%, 726 refl.)	0.557 (Rf = 41.0)	0.571	0.579	0.655	0.664 (Rf = 35.5)

As mentioned previously (Yaremchuk *et al.*, 1995a), crystallization of LysRSTT with its cognate wild-type $tRNA^{Lys}$ gives rise to the $P42_12$ form. No good quality high resolution data have been measured on this complex. However, the same crystal form has been obtained using an unmodified transcript of the *T.thermophilus* $tRNA^{Lys}$ (34-CUU). The production of this transcript will be described elsewhere (M.Tukalo, unpublished results). A single crystal at 100 K was measured to 2.9 Å resolution on BM14 at the ESRF (see Table I).

Structure resolution

Using the co-ordinates of the *E.coli* lysyl-tRNA synthetase (LysU, Onesti *et al.*, 1995), the structure of the hetero-complex was determined by molecular replacement using the program AMORE as implemented in the CCP4 package (1994). With the monomer as search model, one clear solution was found with data in the I422 space group with a correlation of 26% for data between 4 and 9 Å. Rigid-body refinement with XPLOR (Brünger, 1992) showed that the N-terminal β-barrel domain was at a different angle to the catalytic domain as compared with the LysU model. Rebuilding of the LysU structure to convert it into that of LysRSTT was straightforward using σ-weighted (Read, 1986) F_{obs}–F_{calc} and $2F_{obs}$–F_{calc} maps and the amino acid sequence of LysRSTT (Chen *et al.*, 1994). Confirmation of the correctness of the model came from difference maps obtained from room temperature, low resolution data on two heavy atom derivatives phased with the model which revealed clear heavy atom binding sites with very reasonable chemistry. A samarium chloride-soaked crystal measured to 3.8 Å resolution gave a single very strong (20 σ) peak adjacent to Glu361 and Glu402 in the enzyme active site. The equivalent position in AspRSSC and AspRSTT is also a samarium binding site and is close to the enzyme magnesium site (Poterszman *et al.*, 1994). A crystal soaked with potassium tetrachloroplatinate measured to 3.5 Å resolution showed four major platinum peaks adjacent to Met260 (19 σ), Met464 and Met401 (11 σ), His215 (7 σ) and His287 (7 σ).

When the LysRSTT model was partly refined it was used as search model for molecular replacement of the $P42_12$ crystal form. Two distinct solutions for the monomer were found and, using the crystal symmetry, could be chosen such as to constitute a non-crystallographic dimer. Similarly, the trigonal form of the ternary complex with Lys-AMS was solved by molecular replacement. The initial AMORE solution after rigid-body fitting gave a correlation of 50.7% (R-factor 42%) for data between 5 and 9 Å.

Crystal packing

The tetragonal crystals of both I422 and $P42_12$ space groups are built up from rings of four dimers with a large central hole of diameter ~53 Å. In the I422 space group, this ring has an exact 422 point group symmetry and is placed on the unit cell origin. The monomer is the asymmetric unit and the molecular dimer axis coincides with the crystallographic 110 axis. Rings related by the I-centering operation contact each other through their N-terminal domains (via an antiparallel arangement of strands A0, see Figure 7) about a crystallographic 2-fold axis parallel to 110 but passing through (0,0, 1/4). There is ample room for complete tRNA molecules associated with each synthetase subunit, as modelled using the yeast aspartyl-tRNA synthetase complex, even though only partially ordered tRNAs are observed. A single tRNA molecule would have no other contact other than to the LysRSTT molecule to which its anticodon is bound. The crystal is thus held together entirely by protein–protein contacts, as was the case for the most highly diffracting form of the seryl-tRNA synthetase–$tRNA^{Ser}$ complex (Biou *et al.*, 1994; Cusack *et al.*, 1996). The transformation from the I422 to $P42_12$ crystal form corresponds to a 1.5° rotation of the ring of four synthetase dimers around the 4-fold axis. The 22 axes of the ring are thus no longer aligned with the crystallographic axes, and become non-crystallographic, the LysRSTT dimer being the asymmetric unit. Instead of being at the origin of a I422 cell, the ring is at a special position (0,0, 1/4) in a $P42_12$ cell. Contacts between the rings in the c-direction are still made as before through the N-terminal domains about a true crystallographic 110 2-fold axis.

Refinement and model quality

Positional refinement, individual B-factor refinement and inclusion of a large number of ordered water molecules was performed using standard XPLOR procedures and monitored by the free R-factor (Brünger, 1992). A simulated annealing run from 2500°C was performed once although the resultant model was not used directly but served as a useful indicator of where to improve the current model. Strong positive difference density could be assigned to the backbone and bases of the anticodon loop of the tRNA with much poorer density for the rest of the anticodon stem and some of the D-stem (nucleotides 21–46 and 8–14). The parameters and topology for the refinement of the protein were from Engh and Huber (1991) and for the tRNA from Parkinson *et al.* (1996). The modified base mnm^5s^2U-34 was refined using the geometry determined from the crystal structure of mnm^5s^2U dihydrate which is observed in

zwitterionic form with the N3 deprotonated and the amino N(51) of the mnm^5 protonated (Egert *et al.*, 1979). The occupancy of the tRNA molecule was set arbitrarily to 0.5–0.7 to avoid unphysically high B-factors. This partial occupancy of the tRNA may be due to poorer binding by tRNAs that are not fully modified, for example at position 34 or to incorporation of uncomplexed synthetase molecules into the crystal lattice. All riboses were constrained to 3′-endo conformation, except those of nucleotides 34 and 37 which were constrained to 2′-endo. A bulk solvent correction was included in the later stages of refinement using SOLRAD = 0.25. This significantly improved visibility of weak parts of the tRNA density. A Ramachandran plot shows all but one of the well-defined residues in the most favoured or additionally allowed regions. The current status of the refined model using the I422 data is given in Table II.

Refinement of the 2.75 Å resolution hetero-complex data in the P42$_1$2 form was restricted to the protein, due to the much poorer quality of the tRNA density. Non-crystallographic symmetry restraints were used except in the limited regions of crystal contacts. Construction of the P42$_1$2 model from the I422 model followed by rigid-body refinement gave an initial *R*-factor of 26%. The quality of the final minimized structure is similar to that of the I422 form with an *R*-factor (*R*-free) of 23.1% (25.0%) for all data between 2.75 and 15 Å resolution (79 157 reflections). The refinement was done with a bulk solvent correction, but no explicit water molecules were added. Only minor differences are seen in the synthetase structure between the I422 and P42$_1$2 models, e.g. in the two crystallographically distinct subunits in the P42$_1$2 cell, His110 (which is near a crystal contact) adopts distinct conformations.

The 2.9 Å data on the homo-complex in space group P42$_1$2 with the *T.thermophilus* tRNA^{Lys}(CUU) transcript reveals much stronger density for the tRNA anticodon stem, which may be due to the fact that there is no heterogeneous modification. Density for the nucleotides 9–12 is also interpretable. The structure shows that the backbone and base conformation of C-34 is identical to that of mnm^{5}s^{2}U-34 in the hetero-complex. This complex has been refined, with solvent correction but with no discrete water molecules, to an *R*-factor (*R*-free) of 23.0% (24.9%) for all data between 2.9 and 15 Å (69 311 reflections) with excellent geometry (r.m.s. bond and angle deviations respectively of 0.008 Å and 1.58°).

Rigid-body refinement of the ternary complex with Lys-AMS was carried out with XPLOR using firstly three rigid bodies (N- and C-terminal domains and tRNA) and then 10 fragments, followed by manual adjustment at fragment boundaries. An atomic model of the lysyl-adenylate analogue was not included in the refinement. The progress of the free correlation factor on intensities (F2F2) is given in Table III for a total of 9360 reflections between 4.3 and 8 Å resolution. In this complex, the tRNA electron density is more continuous than that observed in the original complex but still restricted to the same regions.

Acknowledgements

We gratefully acknowledge the help of the staff of ESRF beamlines BL4 (ID2), BL3 (ID13) and BL19 (BM14) during data collection, in particular Bjarne Rasmussen (EMBL), Michael Wulff (ESRF), Andrew Thompson (EMBL) and Sean McSweeney (EMBL). We also are very grateful to Brian Sproat and Barbro Beijer for synthesis of 5′-*O*-[*N*-(L-lysyl)-sulfamoyl]adenosine and to Sylvia Onesti and Peter Brick for communication of the co-ordinates of LysRSECu prior to publication. Figures 1, 5 and 6 were prepared using MOLSCRIPT (Kraulis, 1991).

References

Aota,S., Gojobori,T., Ishibashi,F., Maruyama,T. and Ikemura,T. (1988) Codon usage from the GenBank Genetic Sequence data. *Nucleic Acids Res.*, **16** (suppl.), 315–401.

Belrhali,H. *et al.* (1994) Crystal structures at 2.5 Å resolution of seryl-tRNA synthetase complexed with two analogues of seryl-adenylate. *Science*, **263**, 1432–1436.

Belrhali,H., Yaremchuk,A., Tukalo,M., Berthet-Colominas,C., Rasmussen,B., Bösecke,P., Diat,O. and Cusack,S. (1995) The structural basis for seryl-adenylate and Ap$_4$A synthesis by seryl-tRNA synthetase. *Structure*, **3**, 341–352.

Biou,V., Yaremchuk,A.D, Tukalo,M. and Cusack,S. (1994) The 2.9 Å crystal structure of *T.thermophilus* seryl-tRNA synthetase complexed with tRNA^{Ser}. *Science*, **263**, 1404–1410.

Brünger,A.T. (1992) *X-PLOR Version 3.1*. Yale University, New Haven, CT.

Cavarelli,J., Rees,B., Ruff,M., Thierry,J.-C. and Moras,D. (1993) Yeast tRNA^{Asp} recognition by its class II aminoacyl-tRNA synthetase. *Nature*, **362**, 181–184.

Cavarelli,J. *et al.* (1994) The active site of yeast aspartyl-tRNA synthetase: structural and functional aspects of the aminoacylation reaction. *EMBO J.*, **13**, 327–337.

Chen,J., Brevet,A., Lapadat-Tapolsky,M., Blanquet,S. and Plateau,P. (1994) Properties of the lysyl-tRNA synthetase gene and product from the extreme thermophile *Thermus thermophilus*. *J. Bacteriol.*, **176**, 2699–2705.

Collaborative Computational Project, Number 4 (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr.*, **D50**, 760–763.

Commans,S., Plateau,P., Blanquet,S. and Dardel,F. (1995a) Solution structure of the anticodon-binding domain of *E.coli* lysyl-tRNA synthetase and studies of its interaction with tRNA^{Lys}. *J. Mol. Biol.*, **253**, 100–113.

Commans,S., Blanquet,S. and Plateau,P. (1995b) A single substitution in the motif 1 of *E.coli* lysyl-tRNA synthetase induces cooperativity toward amino acid binding. *Biochemistry*, **34**, 8180–8189.

Cusack,S. (1995) Eleven down and nine to go. *Nature Struct. Biol.*, **2**, 824–831.

Cusack,S., Berthet-Colominas,C., Härtlein,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.

Cusack,S., Härtlein,M. and Leberman,R. (1991) Sequence, structural and evolutionary relationships between class 2 aminoacyl-tRNA synthetases. *Nucleic Acids Res.*, **19**, 3489–3498.

Cusack,S., Yaremchuk,A. and Tukalo,M. (1996) The crystal structure of the ternary complex of *T.thermophilus* seryl-tRNA synthetase with tRNA^{Ser} and a seryl-adenylate analogue reveals a conformational switch in the active site. *EMBO J.*, **15**, 2834–2842.

Delarue,M., Poterszman,A., Nikonov,S., Garber,M., Moras,D. and Thierry,J.-C. (1994) Crystal structure of a prokaryotic aspartyl-tRNA synthetase. *EMBO J.*, **13**, 3219–3229.

Egert,E., Lindner,H.J., Hillen,W., Gassen,H.G. and Vorbrüggen,H. (1979) Crystal structure and conformation of mnm^{5}s^{2}U dihydrate. *Acta Crystallogr.*, **B35**, 122–125.

Engh,R.A. and Huber,R. (1991) Accurate bond and angle parameters for X-ray protein structure refinement. *Acta Crystallogr.*, **A47**, 392–400.

Eriani,G., Delarue,M., Poch,O., Gangloff,J. and Moras,D. (1990a) Partition of tRNA synthetases into two classes based on mutually exclusive sets of sequences motif. *Nature*, **347**, 203–206.

Eriani,G., Dirheimer,G. and Gangloff,J. (1990b) Aspartyl-tRNA synthetase from *Escherichia coli*: cloning and characterisation of the gene, homologies of its translated amino acid sequence with asparaginyl- and lysyl-tRNA synthetases. *Nucleic Acids Res.*, **18**, 7109–7118.

Francklyn,C., Harris,D. and Moras,D. (1994) Crystallisation of histidyl-tRNA synthetase from *E.coli*. *J. Mol. Biol.*, **241**, 275–277.

Hillen,W., Egert,E., Lindner,H.J. and Gassen,H.G. (1978) Restriction or amplification of wobble recognition. *FEBS Lett.*, **94**, 361–364.

Isel,C., Lanchy,J.-M., Le-Grice,S.F.J., Ehresmann,C., Ehresmann,B. and Marquet,R. (1996) Specific initiation and switch to elongation of human immunodeficiency virus type 1 reverse transcriptase require the post-transcriptional modifications of primer tRNA^{Lys}3. *EMBO J.*, **15**, 917–924.

Kagawa,Y., Nojima,H., Nikiwa,N., Ishizuka,M., Makajima,T., Yasuhara,T., Tanaka,T. and Oshima,T. (1984) High guanine plus cytosine content in the third letter of codons of an extreme thermophile. *J. Biol. Chem.*, **259**, 2956–2960.

Kasai,H., Nishimura,S., Vorbrüggen,H. and Iitaka,Y. (1979) Crystal and molecular structure of the acetonide of 5-methylaminomethyl-2-thiouridine. *FEBS Lett.*, **103**, 270–273.

Kraulis,P.J. (1991) MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.*, **24**, 946–950.

Leslie,A.G.W. (1992) *Joint CCP4 and ESF-EACBM Newsletter on Protein Crystallography*. No. 26, Daresbury Laboratory, Warrington WA4 4AD, UK.

Li,S., Pelka,H. and Schulman,L.H. (1993) The anticodon and discriminator base are important for aminoacylation of *E.coli* tRNA^{Asn}. *J. Biol. Chem.*, **268**, 18335–18339.

McClain,W.H., Foss,K., Jenkins,R.A. and Schneider,J. (1990) Nucleotides that determine *Escherichia coli* tRNA^{Arg} and tRNA^{Lys} acceptor identities revealed by analyses of mutant opal and amber suppressor tRNAs. *Proc. Natl Acad. Sci. USA*, **87**, 9260–9264.

Murzin,A.G. (1993) OB (oligonucleotide/oligosaccharide binding)-fold: common structural and functional solution for non-homologous sequences. *EMBO J.*, **12**, 861–867.

Onesti,S., Miller,A.D. and Brick,P. (1995) The crystal structure of the lysyl-tRNA synthetase (LysU) from *E.coli*. *Structure*, **3**, 163–176.

Parkinson,G., Vojtechovsky,J., Clowney,L., Brünger,A.T. and Berman,H.M. (1996) New parameters for the refinement of nucleic acid-containing structures. *Acta Crystallogr.*, **D52**, 57–64.

Poterszman,A., Delarue,M., Thierry,J.-C. and Moras,D. (1994) Synthesis and recognition of aspartyl-adenylate by *T.thermophilus* aspartyl-tRNA synthetase. *J. Mol. Biol.*, **244**, 158–167.

Pütz,J., Puglisi,J.D., Florentz,C. and Geigé,R. (1991) Identity elements for specific aminoacylation of yeast $tRNA^{Asp}$ by cognate aspartyl-tRNA synthetase. *Science*, **252**, 1696–1699.

Read,R.J. (1986) Improved Fourier coefficients for maps using phases from partial structures with errors. *Acta Crystallogr.*, **A42**, 140–149.

Ruff,M. *et al.* (1991) Class II aminoacyl-tRNA synthetases: crystal structure of yeast aspartyl-tRNA synthetase complexed with $tRNA^{Asp}$. *Science*, **252**, 1682–1689.

Saks,M.E., Sampson,J.R. and Abelson,J.N. (1994) The transfer RNA identity problem: a search for rules. *Science*, **263**, 191–197.

Sylvers,L.A., Rogers,K.C., Shimizu,M., Ohtsuka,E. and Söll,D. (1993) A 2-thio-uridine derivative in the $tRNA^{Glu}$ is a positive determinant for aminoacylation by *E.coli* glutamyl-tRNA synthetase. *Biochemistry*, **32**, 3836–3841.

Tamura,K., Himeno,H., Asahara,H., Hasegawa,T. and Shimizu,M. (1992) *In vitro* study of *E.coli* $tRNA^{Arg}$ and $tRNA^{Lys}$ identity elements. *Nucleic Acids Res.*, **20**, 2335–2339.

Wittwer,A.J. (1983) Specific incorporation of selenium into lysine- and glutamate-accepting tRNAs from *E.coli*. *J. Biol. Chem.*, **258**, 8637–8641.

Yaremchuk,A.D., Kryklivyi,I., Cusack,S. and Tukalo,M. (1995a) Co-crystallization of lysyl-tRNA synthetase from *Thermus thermophilus* with its cognate $tRNA^{Lys}$ and with *E.coli* $tRNA^{Lys}$. *Proteins*, **21**, 261–264.

Yaremchuk,A.D., Cusack,S., Åberg,A., Gudzera,O., Kryklivyi,I. and Tukalo,M. (1995b) Crystallisation of *T.thermophilus* histidyl-tRNA synthetase and its complex with $tRNA^{His}$. *Proteins*, **21**, 426–428.

Yokoyama,S. and Nishimura,S. (1995) Modified nucleosides and codon recognition. In Söll,D. and RajBhandary,U. (eds), *tRNA: Structure, Biosynthesis and Function*. ASM Press, Washington, DC, pp. 207–233.

Yokoyama,S., Watanabe,T., Murao,K., Ishikura,H., Yamaizumi,Z., Nishimura,S. and Miyazawa,T. (1985) Molecular mechanism of codon recognition by tRNA species with modified uridine in the first position of the anticodon. *Proc. Natl Acad. Sci. USA*, **82**, 4905–4909.

Received on May 9, 1996; revised on July 30, 1996