The EMBO Journal vol.15 no.11 pp.2834–2842, 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

Stephen Cusack[1], Anna Yaremchuk and Michael Tukalo

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

[1]Corresponding author

The low temperature crystal structure of the ternary complex of *Thermus thermophilus* seryl-tRNA synthetase with $tRNA^{Ser}$ (GGA) and a non-hydrolysable seryl-adenylate analogue has been refined at 2.7 Å resolution. The analogue is found in both active sites of the synthetase dimer but there is only one tRNA bound across the two subunits. The motif 2 loop of the active site into which the single tRNA enters interacts within the major groove of the acceptor stem. In particular, a novel ring–ring interaction between Phe262 on the extremity of this loop and the edges of bases U68 and C69 explains the conservation of pyrimidine bases at these positions in serine isoaccepting tRNAs. This active site takes on a significantly different ordered conformation from that observed in the other subunit, which lacks tRNA. Upon tRNA binding, a number of active site residues previously found interacting with the ATP or adenylate now switch to participate in tRNA recognition. These results shed further light on the structural dynamics of the overall aminoacylation reaction in class II synthetases by revealing a mechanism which may promote an ordered passage through the activation and transfer steps.

Keywords: class II aminoacyl-tRNA synthetase/protein–RNA recognition/ring–ring interaction/seryl-tRNA synthetase/X-ray crystallography

Introduction

Structural studies of aminoacyl-tRNA synthetases and their substrate complexes have three major interesting aspects. Firstly, until recently, structures of synthetase–tRNA complexes were the only source of detailed structural information on protein–RNA interactions. In particular, crystallographic studies of synthetase–tRNA complexes (Rould *et al.*, 1989; Cavarelli *et al.*, 1993; Biou *et al.*, 1994) are beginning to provide the structural basis for the concept of tRNA identity, i.e. which features of a tRNA are vital to its correct recognition by its cognate synthetase. Secondly, synthetases catalyse a two-step reaction, firstly amino acid activation by means of ATP and secondly amino acid transfer to the 3′-terminal ribose of the cognate tRNA. Exactly how this is acheived with one active site is of general interest for the understanding of more complex enzymes (Perona *et al.*, 1993; Cavarelli *et al.*, 1994; Belrhali *et al.*, 1995) as is the structural basis for amino acid selectivity. Thirdly, aminoacyl-tRNA synthetases are presumably amongst the most ancient of enzymes. Structural studies have and will continue to shed intriguing light on the evolution of amino acid specificity in this family of enzymes and more generally, perhaps, on the origin of the protein biosynthetic machinery (Härtlein and Cusack, 1995). In this last respect, a major discovery was the partition of the 20 aminoacyl-tRNA synthetases into two classes with very different structural characteristics (Cusack *et al.*, 1990; Eriani *et al.*, 1990). Although our knowledge of synthetase structures is advancing rapidly (Cusack, 1995), it still remains of considerable interest to determine which features are common and which are specific to each member of each class. Here we focus on the first two areas of interest and describe crystallographic results on the ternary complex of seryl-tRNA synthetase with $tRNA^{Ser}$ and a seryl-adenylate analogue which shed further light on specific protein–RNA recognition and the mechanism of aminoacylation.

Seryl-tRNA synthetase (SerRS) is a dimeric class II aminoacyl-tRNA synthetase with a number of idiosyncratic features that make it of particular interest. The specific recognition of cognate $tRNA^{Ser}$ by SerRS depends on the mutual interaction of the remarkable α-helical coiled-coil (helical arm) of the synthetase with the long variable arm of the tRNA (Biou *et al.*, 1994 and Figure 1). Both these structural elements are unique to the serine system, other distinctive features being the cross-subunit binding of the tRNA on the synthetase dimer (Biou *et al.*, 1994; Vincent *et al.*, 1995) and the absence of recognition of the $tRNA^{Ser}$ anticodon, a critical identity element in most other synthetase systems (Saks *et al.*, 1994). The importance of the helical arm–long variable arm interaction for efficient and specific aminoacylation has been demonstrated by a variety of biochemical experiments from a number of laboratories in both prokaryotic (Himeno *et al.*, 1991; Normanly *et al.*, 1992; Asahara *et al.*, 1993, 1994) and eukaryotic (Wu and Gross, 1993) systems. As an illustration, we cite the reduction in k_{cat}/K_m for aminoacylation by a factor of 3.5×10^3 when the long variable arm of *Eschericia coli* $tRNA^{Ser}$ is replaced by a short type 1 tRNA loop (Sampson and Saks, 1993). Similarly, truncation of the N-terminal helical arm of *E.coli* SerRS reduces k_{cat}/K_m for aminoacylation by more than four orders of magnitude, without affecting serine activation activity (Borel *et al.*, 1994). Futhermore, the arm-deletion mutant of *E.coli* SerRS is able to misacylate some type 1 tRNAs, albeit at a low level (Borel *et al.*, 1994). This relaxation of specificity suggests firstly, that, in addition to being the major determinant for cognate tRNA, the synthetase helical arm has a role as an anti-determinant for non-cognate tRNAs and secondly, that, in the absence of the helical arm–variable arm interaction, the residual synthetase–

© Oxford University Press

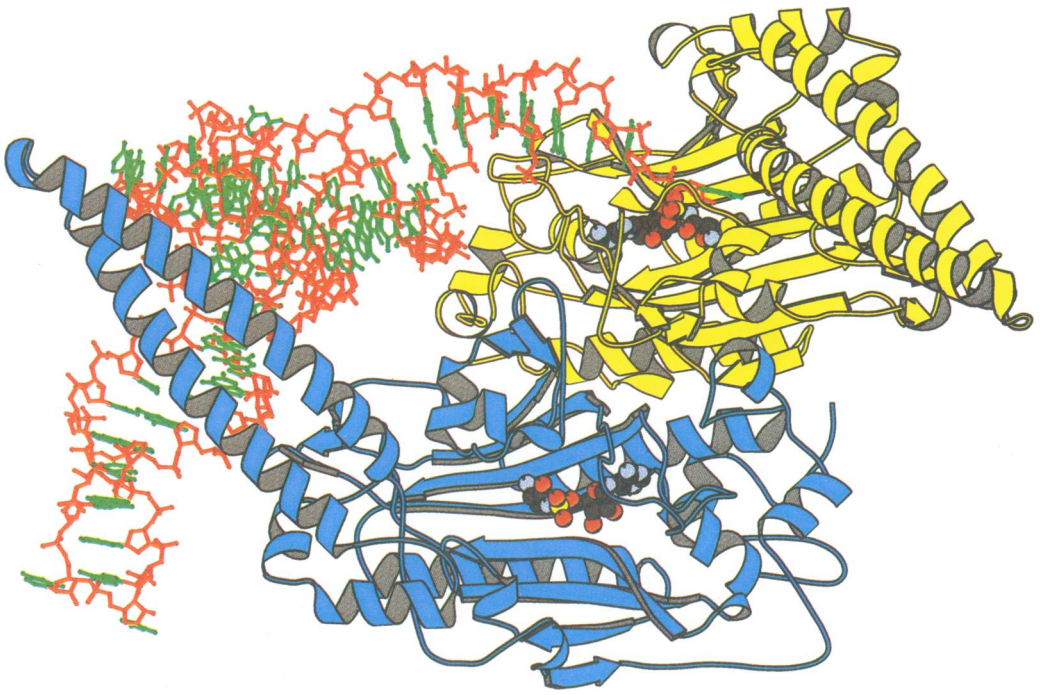

Fig. 1. Overall view of the ternary complex, seryl-tRNA synthetase–tRNA^{Ser} (GGA)–Ser-AMS. The synthetase monomer 1 (residues 1–421) is in yellow and monomer 2 (residues 501–921) in blue. The tRNA backbone is in red and bases in green. The tRNA is viewed looking down the anticodon stem which is not ordered in the crystal structure and not included in the figure. The long variable arm of the tRNA crosses perpendicularly the helical arm of monomer 2 and emerges at the bottom left of the figure. In each active site, the Ser-AMS molecule is represented by spheres. Experimental electron density is poor or lacking for two thirds of the helical arm of monomer 1, the 3′-terminal C75A76 of the tRNA and the extreme end of the tRNA long variable arm. These regions have been modelled manually. Figure prepared using MOLSCRIPT (Kraulis, 1991).

tRNA interactions, notably in the acceptor stem region, are not sufficient to ensure discrimination against non-cognate tRNAs.

The structural basis for the interaction between the synthetase helical arm and the tRNA^{Ser} long variable arm has been visualized in the crystal structure of *Thermus thermophilus* SerRS complexed with its cognate tRNA^{Ser} (GGA) (Biou *et al.*, 1994 and Figure 1). This shows that the synthetase helical arm crosses perpendicularly over the long variable arm of the tRNA, making extensive contacts with the backbone between the second and sixth base pairs of the long variable arm as well specific interactions with the bases of the fourth base pair in the minor groove. In addition, the synthetase arm makes contacts with the tRNA T- and D-loops. This structure also shows that, whereas in the uncomplexed state the synthetase helical arm is flexible, interaction with the tRNA (in particular the long variable arm) twists and orientates the synthetase arm in such a way as to maximize interactions with the tRNA and direct the tRNA acceptor stem into the active site of the other subunit (see Figure 3; Biou *et al.*, 1994).

The importance for specificity of the interactions between the synthetase active site region and the acceptor stem of the tRNA^{Ser} is somewhat less clear. In the original tRNA^{Leu} to tRNA^{Ser} identity swap experiments, which were performed *in vivo* in *E.coli*, Normanly *et al.* (1986, 1992) concluded that the discrimator base (G73) and first three base pairs (G1–C72, G2–C71 and A/U3–U/A70) were important components of the tRNA^{Ser} set of identity elements, all of these elements being fully conserved in all known serine isoacceptors (see Figure 2). On the other hand, an *in vitro* analysis of the recognition of full-length tRNA^{Ser} by *E.coli* SerRS, using T7 transcripts, showed only minor effects of substitution at the discriminator base or of, for instance, substituting an A1–U72 base pair in place of the normal G1–C72 (Asahara *et al.*, 1994). The most significant discrimination was found at position 2–71, where changes from the normal G2–C71 to C2–G71 or U2–G71 were found to be detrimental, with respective V_{max}/K_m values of 0.06 and 0.09 relative to wild-type transcripts. The authors concluded that SerRS does not discriminate strongly the base sequence of the acceptor stem.

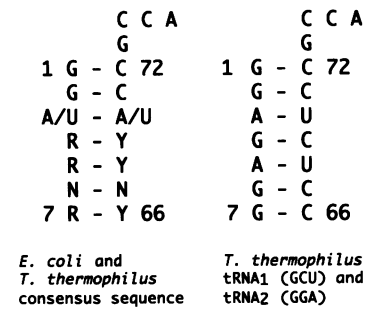

Fig. 2. Sequences of acceptor stems of tRNA^{Ser}. Left, consensus acceptor stem sequence of all known *E.coli* and *T.thermophilus* tRNA^{Ser}. Right, common acceptor stem sequence of *T.thermophilus* tRNA$_1^{Ser}$ (GCU) and tRNA$_2^{Ser}$ (GGA). Abbreviations: R (purine), Y (pyrimidine), N (any nucleotide).

Sequence discrimination of the acceptor stem by *E.coli* SerRS has been re-examined using acceptor stem–TΨC loop minihelices instead of full-length T7 transcripts (Saks

and Sampson, 1996). Mutations were studied in which each of the four canonical base pairs were substituted separately at each of the seven base pairs of the acceptor stem. The results confirm the preference of SerRS for G1–C72, G2–C71 and A/U3–U/A70. In addition, the results reveal for the first time a purine–pyrimidine (R–Y) preference for base pairs 4–69 and 5–68.

The original crystal structure of the complex of *T.thermophilus* SerRS with its cognate $tRNA^{Ser}$ (GGA) showed disorder in the active site region preventing any interpretation of the interactions with the acceptor stem and 3′ end of the tRNA (Biou *et al.*, 1994). Here we present an improved structure in which the interactions of the acceptor stem with the synthetase are now visible, even though bases 74–76 at the extreme 3′ end of the tRNA remain poorly ordered. Data has been remeasured on the same crystal form as that previously described (Yaremchuk *et al.*, 1992; Biou *et al.*, 1994) which contains a single tRNA bound to the SerRS dimer. However, there are two significant differences: firstly, the complex was co-crystallized in the presence of the sulfamoyl analogue of the natural intermediate seryl-adenylate (Ser-AMS, Belrhali *et al.* 1994) and secondly, data collection was performed on frozen crystals at low temperature. The new structure reveals that, upon tRNA binding, the motif 2 loop takes on a completely new conformation compared with that observed when only ATP or Ser-AMS is bound. In this new conformation, certain residues that previously have been observed to interact with ATP or Ser-AMS switch to interact with the tRNA. The interactions of the motif 2 loop of SerRS with the tRNA acceptor stem account, to a large extent, for the biochemical results cited above. The existence of crystal structures of *T.thermophilus* SerRS in the native state (Fujinaga *et al.*, 1993), binary complexes with ATP, Ap_4A, Ser-AMS, Ser-AMP (Belrhali *et al.*, 1994, 1995) and tRNA (Biou *et al.*, 1994) and ternary complexes with tRNA and Ser-AMS, give an unprecedented opportunity to describe the conformational changes associated with substrate binding and the two steps of the catalytic activity.

Results

Cryo-crystallographic data to 2.7 Å resolution were measured on a single frozen crystal of the *T.thermophilus* SerRS complexed with $tRNA^{Ser}$ (GGA) which had been co-crystallized with the seryl-adenylate analogue, Ser-AMS (Belrhali *et al.*, 1994). A new model was refined against this data to an *R*-factor of 18.8% (*R*-free 24.9%) with excellent geometry (Table IV). This model is similar to that previously described in having a single tRNA molecule bound across the two subunits of the synthetase (Biou *et al.*, 1994). The monomer with the tRNA entering the active site has residue numbers 1–421, although residues 36–89 of the helical arm are disordered; the other monomer (residues 501–921) has an ordered helical arm which interacts with the tRNA. There is a major difference from the previous structure, however, namely that the acceptor stem of the tRNA is now clearly visible up to C74 whereas previously the first three base pairs of the acceptor stem were also very poorly ordered. Despite this improvement, the electron density for G73, C74 and G1 is incomplete and that for C75A76-3′ is too poor to permit refinement. The whole acceptor stem has slightly re-orientated from the previous structure, and the motif 2 loop which interacts within the major groove of the acceptor stem is now fully ordered. Both active sites of the SerRS dimer contain the Ser-AMS inhibitor. In general, the cryo-crystallographic data collection results in a better defined structure notably in side chain conformations and the location of ~200 well resolved water molecules mainly associated with the protein and the synthetase–tRNA interface. There are also several instances of multiple conformations, the most striking being that of Trp146 (but not the corresponding residue Trp646 in the other subunit).

The overall structure of the complex is shown in Figure 1, which illustrates the cross-subunit tRNA binding, the importance of the synthetase helical arm–tRNA interactions and the binding of the seryl-adenylate analogue in the active site. Nucleotides C75 and A76 of the tRNA have been manually modelled, but not refined, using the weak electron density observed. The modelled position of the base of A76 is partially stacking on His204. This is a slightly different position and orientation to that found for the second adenosine of Ap_4A as expected (see discussion in Belrhali *et al.*, 1995). Unfortunately, we do not consider this modelling to be sufficiently accurate to discuss further the exact mechanism of the serine transfer to the ribose of A76.

One of the most striking features of the new structure is the existence of two very different ordered conformations of residues 258–267 (758–767) of the motif 2 loop in each synthetase subunit. It should be remembered that both subunits have a Ser-AMS molecule bound, but only one has a tRNA molecule entering the active site. The two conformations are not related by a rigid body movement, and the root mean square (r.m.s.) deviation in the Cα position for the nine residues 269–267 is 5.6 Å. The two conformations are compared in Figure 3. In the subunit which has no tRNA entering the active site, the motif 2 loop is in the A (adenosine)-conformation. This is so designated because the same conformation is observed when either Ser-AMS or ATP (Belrhali *et al.*, 1995) is bound, provided tRNA is absent from the active site. An exactly equivalent ordered A-conformation is also found in the crystal structure of the helical arm-deletion mutant of the *E.coli* SerRS complexed with Ser-AMS (Borel *et al.*, 1994; H.Belrhali, S.Price, C.Berthet-Colominas, M.Härtlein, R.Leberman and S.Cusack, unpublished results). The occurrence of the same conformation in three different crystal forms is strong evidence that this is a functional conformation and not a result of crystal environment although, in the current complex structure, the A-conformation is stabilized by a crystal contact. In the other subunit, the binding of the tRNA acceptor stem induces the motif 2 loop to take up a second conformation which we designate the T-conformation. Finally, in the absence of any substrates, the motif 2 loop is disordered (Fujinaga *et al.*, 1993). The two different ordered conformations are each stabilized by different sets of interactions often involving the same residues. Arg157 makes hydrogen bonds to the main chain carbonyl oxygens of Ser261 and Arg271 in both the A- and T-conformation, but also to that of Lys264 only in the A-conformation. Arg344 hydrogen-bonds to the side chain carboxyl group of Asp265 in the A-conformation and to the main chain

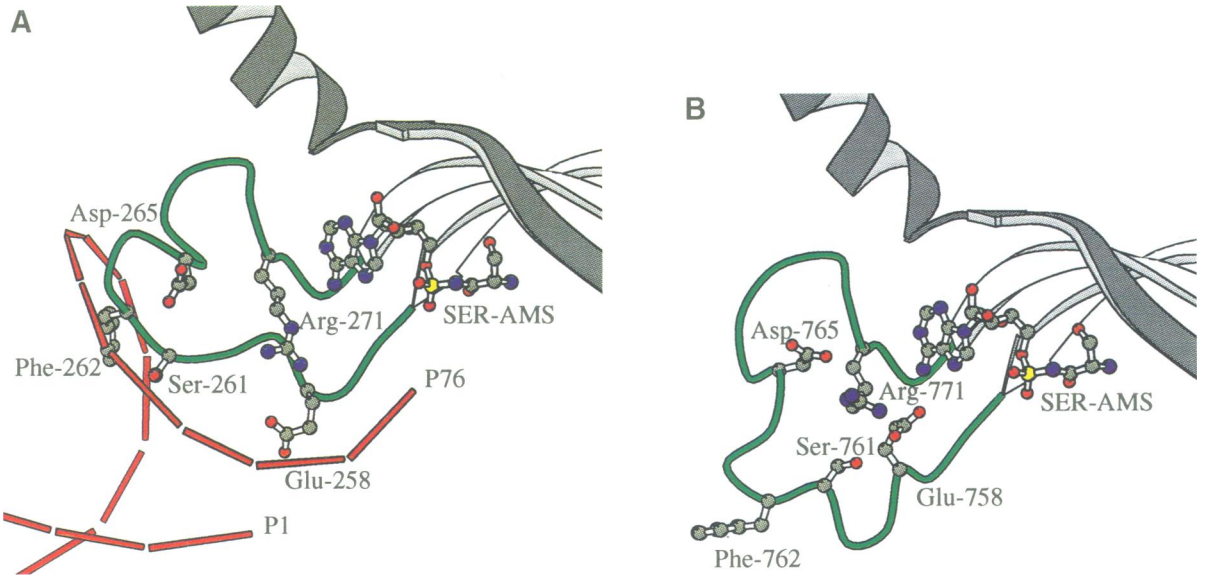

Fig. 3. Comparison of motif 2 loop conformation as observed in (**A**) monomer 1 (T-conformation) and (**B**) monomer 2 (A-conformation) in the ternary complex. Orientations are identical. The side chains of some key residues (see text and Table I) are shown together with the Ser-AMS molecule in each active site and the secondary structure elements corresponding to motif 2 (β-strands white, loop green) and motif 3 (black β-strand followed by α-helix). In (A) the tRNA phosphate positions are linked by red lines. Figure prepared using MOLSCRIPT (Kraulis, 1991).

Table I. Interactions of various residues associated with the motif 2 loop in the T- and A-conformations

Residue monomer 1		T-conformation (monomer 1)		A-conformation (monomer 2)	
Arg157	NE	Arg271	O	Arg771	O
(Arg657)	NH1	Ser261	O	Ser761	O
	NH2			Lys764	O
Glu258	OE2	G73	N2	Ser761	OG
(Glu758)	OE1			ATP, Ser-AMP, Ser-AMS	N6
Ser261	OG	G2	O6	Glu758	OE2
(Ser761)		G1	O6		
Phe262		C69, U68		–	
(Phe762)		(see text)			
Asp265	OD1	–		Arg771	NH2
(Asp765)					
Arg267	NE	C69	O2P	–	
(Arg767)					
Arg271	NH1	C74	N3	ATP	Pγ
(Arg771)	NH2	C74	O2	Asp765	OD1
Arg344	NH2	Asp265	O	Asp765	OD1,OD2
(Arg844)					

Note that monomer 1 (2) residues are numbered 1–421 (501–921) and the motif 2 loop is in the T (A)-conformation.

carbonyl oxygen of the same residue in the T-conformation. Of those residues that interact with substrates, the side chain of Glu258 flips by 180° between the two states (Figure 3), interacting with the *N*6 of adenine (of ATP or Ser-AMS) in the A-conformation and the *N*2 of G73 in the T-conformation. Ser261 helps position Glu258 in the A-conformation but interacts with G2 in the T-conformation. In the A-conformation, Arg271 is positioned by Asp265 and orientated to interact with the γ-phosphate of ATP (Belrhali *et al.*, 1995). In the T-conformation, it forms hydrogen bonds to the *O*2 and *N*3 of C74. All the cited residues of SerRS are highly conserved in all known sequences of the enzyme. These different interactions are summarized in Table I.

Figure 4 compares the conformation of the motif 2 loop relative to the acceptor stem of the tRNA as observed in the crystal structure of the SerRS– tRNA^{Ser} complex and the yeast AspRS–tRNA^{Asp} complex (Ruff *et al.*, 1991). The sequences of the motif 2 loop in various class II synthetases are given in Table II which shows that the loop of SerRS is amongst the longest, being six residues longer than that of yeast AspRS. In the case of yeast AspRS, the motif 2 loop makes extensive base-specific hydrogen bond interactions to the discriminator base (G73) and the first base pair A1–U72 in the tRNA^{Asp} acceptor stem major groove (Cavarelli *et al.*, 1993). In SerRS, residues from the much longer motif 2 loop are able to extend down inside the major groove to the fifth base pair of the acceptor stem.

Two kinds of base interactions in the acceptor stem region are observed. The first is an interaction between Phe262 at the extremity of the motif 2 loop and the two pyrimidine bases U68 and C69 (Figure 5A and B). The Cβ and phenyl ring of Phe262 is in van der Waals contact with the hydrophobic edge (*C*5–*C*6 positions) of the bases,

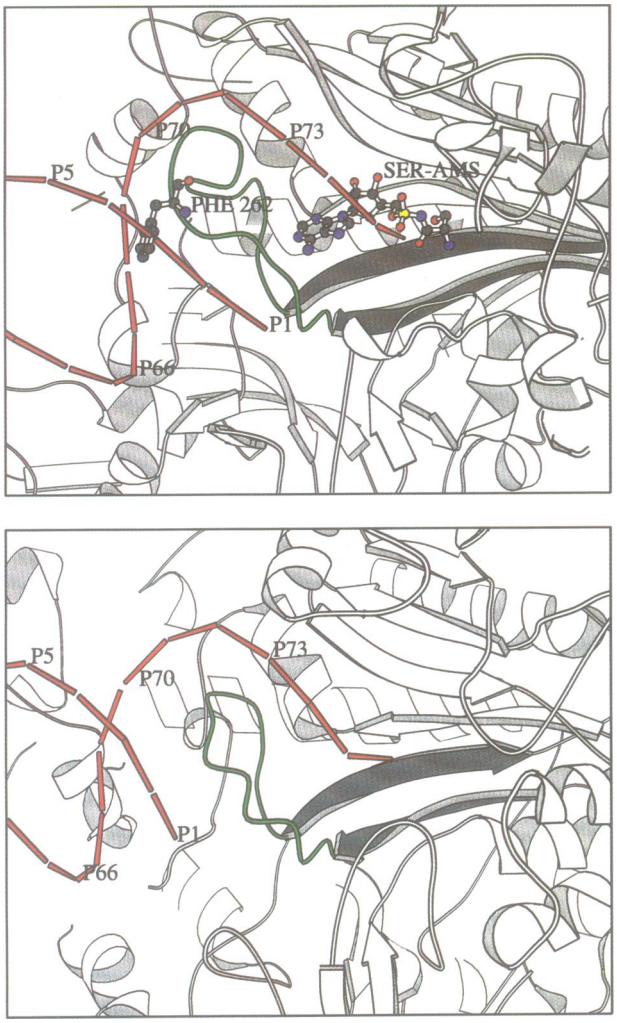

Fig. 4. Comparison of the conformations of the motif 2 loop (green) and tRNA acceptor stem (phosphate trace, red) in the active site region of *T.thermophilus* seryl-tRNA synthetase ternary complex (top) and yeast aspartyl-tRNA synthetase complex (bottom). In the top figure, the side chain of Phe262 is shown extending deep into the major groove of the acceptor stem. Figure prepared using MOLSCRIPT (Kraulis, 1991).

in particular C69. The angle between the phenyl ring and base of C69 is ~60°. This kind of ring–ring interaction was originally described by Burley and Petsko (1985), and a very similar interaction has been observed between Phe565 and the substrate cytidine bound in the intersubunit interface in cytidine deaminase (Betts *et al.*, 1994). This interaction would appear to strongly favour pyrimidine bases in positions 68 and 69 since the ring–ring interaction is stabilizing (Burley and Petsko, 1985), whereas purine bases placed here would position the hydrogen bond acceptor *N*7 in an unfavourable hydrophobic environment. An effect of the interaction of Phe262 is the induction of a significant propeller twist of ~16° on base pairs G4–C69 and A5–U68 (visible in Figure 5A).

The second kind of interaction observed is hydrogen bonding to bases. A hydrogen bond occurs between the main chain carbonyl oxygen of Phe262 and the exocyclic amino group (*N*4) of C71 (although this is somewhat long at 3.3 Å). The hydroxyl oxygen (OG) of Ser261 is

Table II. Comparative sequences of motif 2 loops from a number of class II aminoacyl-tRNA synthetases

		←loop→
SerEC	267	FRSE--AGSYGRDTR--GLIR
SerTT	255	FRSE--AGSFGKDVR--GLMR
SerBS	261	FRSE--AGSAGRDTR--GLIR
SerSC	278	FRRE--AGSHGKDAW--GVFR
SerHS	301	FRQE--VGSHGRDTR--GIFR
SerRS		FRsE--aGSxGrDtr--GΦΦR
ThrEC	362	HRNE----PSGSLH---GLMR
ProEC	139	FRDE----VRPRF----GVMR
GlyHU	277	FRNE----ISPRS----GLIR
AspSC	324	FRAE----NSN------THRH

Abbreviations are SerEC: seryl-tRNA synthetase from *E.coli*; TT: *T.thermophilus*; BS: *Bacillus subtilus*; SC: *Saccharomyces cerevisiae*; HS: human; SerRS: consensus sequence for known seryl-tRNA synthetases (x: any residue, Φ: hydrophobic residue, upper case: absolutely conserved); ThrEC: threonyl-tRNA synthetase from *E.coli*; ProEC: prolyl-tRNA synthetase from *E.coli*; GlyHS: glycyl-tRNA synthetase from human; AspSC: aspartyl-tRNA synthetase from *S.cerevisiae*.

positioned virtually half way between base pairs 2–71 and 1–72 and could be acting as a hydrogen bond donor to *O*6 of either G2 or G1 which are respectively 3.0 and 3.2 Å away (Figure 5C). The OG of Ser261 is also within hydrogen bonding distance of the exocyclic amino groups (*N*4) of both C71 and C72, although the geometry is not favourable. The discriminator base G73 is clearly seen stacking over the centre of the 1–72 base pair and the conserved acidic residue of motif 2, Glu258, is in a position to hydrogen-bond with the *N*2 exocyclic amino group. Arg271 hydrogen-bonds with the *N*3 and *O*2 positions of C74. The interactions of Glu258 and Arg271 with respectively G73 and C74 are similar to the tRNA interactions made by structurally equivalent residues Glu327 and His334 in the yeast AspRS–tRNA^{Asp} complex (Cavarelli *et al.*, 1993).

A list of direct polar interactions between the synthetase and the tRNA acceptor stem is shown in Table III (compare Table 1 of Biou *et al.*, 1994); water-mediated interactions are not listed. These include a variety of hydrogen bonds and electrostatic interactions between the synthetase and tRNA backbone in the acceptor stem region, although they are exclusively to the pyrimidine-rich 3′ strand. Some contact is made with the ribose or phosphate of all nucleotides from 66–71, except 70. It is possible that these one-sided synthetase–backbone interactions, and/or the fact that the acceptor stem is a pure purine (5′)–pyrimidine (3′) helix (Figure 2), induce a particular helical conformation on the acceptor stem that is required for functional interaction with the synthetase. However, in the absence of the structure of the uncomplexed tRNA, or of a complex structure of sufficient resolution to define accurately local helical parameters independently of refinement parameters, we have not examined this further.

Discussion

In the discussion that follows, it should be borne in mind that all quoted biochemical results refer to the *E.coli* seryl system, whereas the crystallographic results pertain to the *T.thermophilus* system. Given the high similarity in structure between the two synthetases (Fujinaga *et al.*,

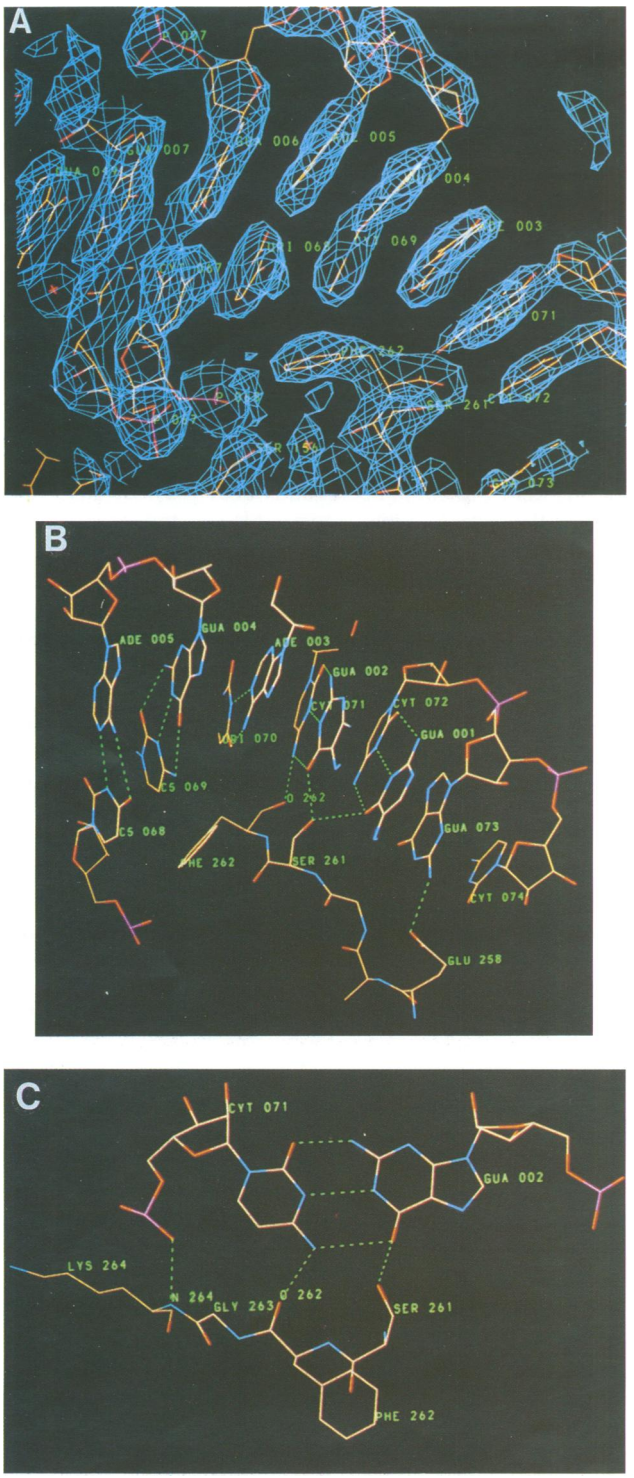

Fig. 5. (**A**) Final $2F_o$–F_c electron density (contour level 1σ) with superimposed model showing the van der Waals contact between Phe262 and the hydrophobic edges of bases U68 and C69. (**B**) General view of the motif 2 loop interacting in the major groove of the acceptor stem of tRNA^{Ser}. (**C**) Interactions of the motif 2 loop with the G2–C71 base pair.

1993), especially in the active site region, the similarity of the acceptor stems of the corresponding tRNAs (Figure 2) and the fact that heterologous aminoacylation in both senses occurs, it is reasonable to assume that both systems behave in a closely similar way. However two differences are worth noting. First, residue Phe262 of the *T.thermophilus* enzyme is equivalent to Tyr274 in the *E.coli* SerRS (Table II). Addition of a hydroxyl group to Phe262 in the *T.thermophilus* complex model shows that the group would point at phosphate 67 of the tRNA but the distance (4.5 Å) would imply that an interaction with the phosphate would have to be water mediated. Secondly, the *E.coli* SerRS has a much longer motif 1 loop than the *T.thermophilus* enzyme (see primary sequence alignments in Fujinaga *et al.*, 1993). This loop forms a β-hairpin and interacts with its symmetry-related partner to form a four-stranded anti-parallel β-sheet about the dimer 2-fold axis (Cusack *et al.*, 1990; Fujinaga *et al.*, 1993). In *T.thermophilus*, the short loop (residues Ala216–Asp220) plays no direct role in substrate binding. In *E.coli*, the loop is long enough (residues Arg221–Asn231) to reach over to the tRNA acceptor stem binding site of the opposite subunit. Using the 2.2 Å resolution model of the arm-deletion mutant of *E.coli* SerRS complexed with Ser-AMS in which the motif 1 loop is ordered (H.Belrhali, S.Price, C.Berthet-Colominas, M.Härtlein, R.Leberman and S.Cusack, unpublished results), it is clear that part of this loop (Asp228–Thr229) could contact the tRNA backbone in the region of nucleotides 65 and 66 and the extremity (Glu224–Glu225) could be in contact with the backbone of nucleotide 1. These putative additonal interactions could alter the exact orientation of the acceptor stem in the *E.coli* complex compared with the *T.thermophilus* complex.

Sequence discrimination of the acceptor stem by *E.coli* SerRS has been re-examined *in vitro* using acceptor stem–TΨC loop minihelices instead of full-length T7 transcripts (Saks and Sampson, 1996). The authors show that these

Table III. Seryl-tRNA synthetase–tRNA^{Ser} polar interactions <3.5 Å distance

Residue	Atom	Nucleotide	Atom	Distance (Å)
Monomer 1				
Ser151	OG	U68	O2P	3.0
Gln152	OE1	C67	O2′	3.0
Ser156	N	C67	O2P	3.1
	OG	U68	O1P	2.5
Glu258	OE2	G73	N2	2.8
Ser261	OG	G1	O6	3.2
	OG	G2	O6	3.0
Phe262	O	C71	N4	3.3
Lys264	N	C71	O1P	3.0
Arg267	NE	C69	O2P	3.0
Arg271	NH1	C74	N3	2.7
	NH2	C74	O2	3.1
Monomer 2				
Lys542	NZ	G47b	O3′	3.2
	NZ	G47c	O2P	3.3
Gln545	OE1	G47a	N2	3.2
	NE2	Cyt47n	O2	2.9
Thr549	OG1	C47n	O2′	2.7
Arg551	NH2	G57	O2P	3.0
Asn552	ND2	C47o	O2P	2.7
	ND2	C47p	O1P	3.3
Ala555	O	G19	N2	2.7
Lys581	NZ	T54	O2P	3.2
Arg588	NH2	G53	O2P	3.2
Arg695	NH2	C66	O3′	2.6
	NH2	C67	O2P	2.6
Arg863	NH1	G46	O2′	3.5
Arg865	NH1	G46	O1P	3.2
	NH2	G46	O2P	2.7

minisubstrates are more sensitive probes of sequence preferences than are full-length transcripts where the contribution of the helical arm–variable arm interactions is overwhelming. Mutations were studied in which each of the four canonical base pairs was substituted separately at each of the seven base pairs of the acceptor stem. The results confirm the preference of SerRS for G1–C72, G2–C71 and A/U3–U/A70, in accordance with previous results (Normanly *et al.*, 1992; Asahara *et al.*, 1993, 1994). The largest effect is at position 2–71 where a substitution to C2–G71 decreases k_{cat}/K_m 80-fold. In addition, the results reveal for the first time a purine–pyrimidine (R–Y) preference for base pairs 4–69 and 5–68. In particular, at position 4–69 there is a decrease in k_{cat}/K_m of factors of 1, 12 and 46 for substitutions from G4–C69 to A–U, U–A or C–G, respectively. This correlates with the fact that all natural *E.coli* serine isoacceptors and the two known *T.thermophilus* $tRNA^{Ser}$s have R–Y in both these positions (Figure 2).

The crystallographic results described above provide the structural basis to understand these observations. The interaction of the motif 2 loop (notably the carbonyl oxygen of Phe262 and the side chain of Ser261) with the 2–71 base pair clearly favours the occurrence of G2–C71, since any other standard choice will reduce the number of possible hydrogen bonds and put other hydrogen bond acceptors in proximity to the carbonyl oxygen of Phe262. The ring–ring hydrophobic interaction between Phe262 and bases C69 and U68 explains the preference in these positions for pyrimidine bases on the 3′ strand, since purine bases placed here would position the hydrogen bond acceptor *N7* in an unfavourable hydrophobic environment. This is the first example of such an interaction in a protein–RNA system, although many examples of parallel stacking interactions between aromatic residues and RNA bases have been noted (Mattaj and Nagai, 1995). Despite these particular examples, perhaps the most significant observation is that there are indeed relatively few base-specific interactions between SerRS and the tRNA acceptor stem. For instance, there are no direct base-specific interactions with the A3–U70 base pair at all, even though this had originally been considered an identity element (Normanly *et al.*, 1986). This structural result is in general agreement with the conclusion that SerRS does not discriminate strongly the base sequence of the acceptor stem (Asahara *et al.*, 1994).

In the case of the *E.coli* system, the weak base-specific preferences shown by SerRS probably reflect the need to avoid mischarging of non-cognate tRNAs, in particular, $tRNA^{Leu}$ and $tRNA^{Thr}$ isoacceptors. In *E.coli*, $tRNA^{Leu}$s invariably have A73 and C2–G71 although, of these, only the A73 is an identity element in the leucyl system (Asahara *et al.*, 1993). $tRNA^{Thr}$ isoacceptors in *E.coli* invariably have G1–C72 as in $tRNA^{Ser}$, but differ from $tRNA^{Ser}$ in having C2–G71, an important identity element in the threonyl system (Hasegawa *et al.*, 1992). Further work needs to be done in the *T.thermophilus* system to determine whether these arguments still apply.

In this work, we show that the motif 2 loop of *T.thermophilus* SerRS can be in two quite different conformations, one in the presence of tRNA (T-conformation), the other in the absence of tRNA but in the presence of either ATP or Ser-AMS (A-conformation). A key residue in the switch between these two conformations is Glu258. In the A-conformation, the side chain of Glu258 interacts with the *N6* of the adenine base of ATP or Ser-AMS, whereas in the T-conformation the main chain and side chain exchange positions (Figure 3). The flip of Glu258 by 180° permits it then to interact with G73 (the 'discriminator' base) of the tRNA. Similarly the side chain of Arg271 alters conformation, binding either the γ-phosphate of ATP or C74 of the tRNA. The occurrence of two glycines in the loop (Gly260 and Gly264) and conserved small residues (A, T or V) in positions 259 and 266 may provide the necessary flexibility and reduced steric hinderance to facilitate the conformational switch. These observations raise the question of what are the dynamic steps in the recognition and catalytic process? In particular, at what stage does the switch from A- to T-conformation of the motif 2 loop occur and what purpose does this serve? We speculate that the variable substrate interactions mediated by the motif 2 loop serve in guiding the active site in an ordered way through the different states needed for substrate recognition, two catalysis steps and product release. The switch from A- to T-conformation is most likely triggered by the successful completion of the serine activation step. The release of the pyrophosphate would free Arg271 to interact with C74 and thus stabilize the tRNA 3′ end conformation permitting the aminoacylation step. The concomitant switch of Glu258 from the adenine *N6* to G73 could also aid the release of AMP after the charging step. Release of AMP would however destabilize the motif 2 loop, whose ordered conformation in either the T- or A-conformation depends on two main chain hydrogen bonds with the adenine base (Belrhali *et al.*, 1994, 1995). This in turn would contribute to the release of the tRNA before the cycle restarts with the re-binding of ATP. It could be that the residual disorder observed in the current structure reflects some 'frustration' in the motif 2 loop as to whether is should adopt the T- or A-conformation when in the presence of both tRNA and the non-hydrolysable intermediate Ser-AMS. In the normal situation, the 3′ end of the tRNA would only transiently be correctly positioned in the presence of the adenylate before the aminoacylation reaction occurred. Further progress in understanding these various states could come from crystal structures of the ternary complex with ATP and tRNA, and the product complex with Ser-$tRNA^{Ser}$ with or without AMP (or AMP analogue). Ideally, however, a probe sensitive to local structure in the active site is required which would permit conformational state and functional state to be correlated as a function of time.

The present results combined with those obtained previously (Biou *et al.*, 1994), provide strong evidence that the functional binding of $tRNA^{Ser}$ to SerRS occurs in at least distinct two steps: firstly the initial recognition and docking which depends largely on the helical arm–long variable arm interaction and secondly, the correct positioning of the 3′ end of the tRNA in the active site. The latter process depends critically on forming the correct T-conformation of the motif 2 loop, which itself appears to be favoured by the presence of the adenylate rather than ATP in the active site.

Finally, this structure has revealed a novel manner in which a protein can discriminate between pyrimidine and purine bases in the major groove of an RNA helix.

This discrimination is achieved through a hydrophobic interaction rather than by hydrogen bonds which are usually associated with base recognition, and may turn out to be a more widespread feature in protein–RNA recognition.

Materials and methods

Form four crystals of the *T.thermophilus* SerRS–tRNA^{Ser} (GGA) complex were grown as described (Yaremchuk *et al.*, 1992) but in the presence of 300 μM {5′-*O*-[*N*-(L-seryl)-sulfamoyl]adenosine} a synthetic analogue of seryl-adenylate (Belrhali *et al.*, 1994). These crystals contain a single tRNA molecule bound to the SerRS dimer in the crystal asymmetric unit. For low temperature data collection, a single crystal was transferred from stabilizing solution containing 34% ammonium sulfate successively to the same solution containing 7.5, 15, 22.5 and 30% glycerol (adjusted so that the ammonium sulfate concentration was constant). Each step took ~1 min. Finally the crystal was picked up by means of a loop made from fine thread and flash frozen *in situ* on the goniometer in a stream of cold nitrogen at 138 (±2) K. Freezing caused a decrease in unit cell volume of 8% compared with room temperature. Data collection was done using a 300 cm diameter Mar Research image-plate system on beamline W32 at LURE (Fourme *et al.*, 1992). The crystal–detector distance was 395 mm and the wavelength 0.9 Å. The dataset comprises 79 images with oscillation range 1° and average exposure 400 s. The crystals diffract beyond 2.7 Å resolution but, for technical reasons, complete data was collected to 2.8 Å and with slightly lower completion to 2.7 Å. No significant radiation damage was observed during the data collection. Images were integrated with the MOSFLM package (Leslie, 1992) and subsequent data processing performed with the CCP4 package (1994). Details of the data collection and processing statistics are given in Table IV.

The starting point for refinement was the complex structure previously described (Biou *et al.*, 1994, PDB entry 1SET). After removal of solvent, this model was repositioned in the new (contracted) unit cell by rigid body refinement using XPLOR (Brünger, 1992). Electron density maps of the type $2mF_o–F_c$ and $mF_o–F_c$ were calculated using SIGMAA (Read, 1986) and examined using FRODO (Jones and Thirrup, 1986). Significant changes in the active site region of the synthetase were immediately visible. Manual alterations were made to the model which included addition of a Ser-AMS molecule in each active site, re-orientation and extension of the tRNA acceptor stem, re-orientation of active site residues and addition of the previously disordered motif 2 loop in one subunit (residues 260–264). A little more of the end of the tRNA long variable arm is also visible (base pair G47d–C47k and C47e) but no more of the anticodon stem. The model was refined by cycles of refinement and manual adjustment using standard XPLOR energy minimization and individual B-factor refinement protocols with parameter files PARHCSDX.PRO for protein (Engh and Huber, 1991) and DNA-RNA.PARAM for nucleic acid (Parkinson *et al.*, 1996). Using this new nucleic acid parameter set, it was found to be no longer necessary to add extra constraints to ensure base planarity. Ribose puckers were defined as either 2′ endo or 3′ endo according to Biou *et al.* (1994) and the corresponding options in the parameter set used.

A large number of very well defined water molecules were observed in both $2mF_o–F_c$ and $mF_o–F_c$ difference maps. Eventually, 197 water molecules (mean B-factor 21.2 Å^2) were included in the model, each selected water molecule being above 4σ in the $mF_o–F_c$ difference map and within hydrogen bonding distance of the protein–tRNA complex or other solvent molecule. A considerable number of more putative water molecules of less significance in the difference map have not been included. A very significant positive difference peak adjacent to the *N*7 position of tRNA base G1 refines satisfactorally as a manganese ion. This position on guanines is known to bind Mn^{2+} as has been observed, for instance, in the crystal structure of the ribozyme (Pley *et al.*, 1994). The initial *R*-factor after rigid body refinement and model adjustment was 0.32. After refinement, the *R*-factors for the working set (37 838 reflections, 91% of all measured data between 2.8 and 8 Å resolution) and test set (3764 reflections, 9% of data) dropped respectively to 0.209 and 0.274. Inclusion of a bulk solvent correction (using solrad = 0.25 in XPLOR) reduced these figures further to *R*-work = 0.188, *R*-free = 0.249 for the same resolution range. This correction improved visibility of some weakly defined regions with high B-factors (e.g. phosphate of G1 near the Mn^{2+} ion). A final refinement cycle was performed with solvent correction using all data between 2.7 and 12 Å resolution (44707 reflections, none free) giving an *R*-factor of 0.199. The final model geometry is excellent (see Table IV).

Table IV. Crystallographic data and refinement statistics

Space group	$P2_12_12_1$ (#19)
Unit cell parameters (Å)	*a* = 121.0
	b = 125.7
	c = 117.5
Resolution range	20–2.7 Å
Temperature	138 K
Total measured reflections	43 406
Number independent reflections	45 115
Average redundancy	3.2
Completeness to 2.7 Å	90.7%
Completeness to 2.8 Å	96.6%
Overall *R*-merge	0.069
R-merge in 2.7–2.8 Å shell	0.225

Crystallographic *R*-factors for all data between 2.8 and 8 Å

	No solvent correction	With solvent correction
R-work (91% of data)	0.209	0.188
R-free (9% of data)	0.274	0.249

Crystallographic *R*-factors for all data from 2.7 Å

	No solvent correction 2.7–8 Å	With solvent correction 2.7–12 Å
R-factor	0.214	0.199

R.m.s. deviations from ideal geometry for fully occupied residues:

	Protein	RNA
Bond lengths (Å)	0.009	0.006
Bond angles (°)	1.536	1.086

Mean B-factors (Å^2)

Monomer 1 main (side)	18.2 (21.0)
Monomer 2 main (side)	15.3 (18.6)
tRNA (all atoms)	39.0
Water	21.2

Acknowledgements

We would like to thank Roger Fourme and Jean-Pierre Benoit for access to beamline W32 at LURE and to Stephanie Monaco for help with the data collection. We also thank Charles Carter,Jr for pointing out the phenylalanine–cytidine interaction in cytidine deaminase. We are indebted to Drs Peggy Saks and Jeff Sampson for access to data before publication and for enthusiastic discussions. We thank Dr Brian Sproat and Babro Beijer for synthesis of Ser-AMS and Dr Michael Härtlein and Christine Vincent for communication of the human SerRS sequence.

References

Asahara,H., Himeno,H., Tamura,K., Hasegawa,T., Watenabe,K. and Shimizu,M. (1993) Recognition of nucleotides of *E.coli* tRNA^{Leu} and its elements facilitating discrimination from tRNA^{Ser} and tRNA^{Tyr}. *J. Mol. Biol.*, **231**, 219–229.

Asahara,H., Himeno,H., Tamura,K., Nameki,N., Hasegawa,T. and Shimizu,M. (1994) *E.coli* seryl-tRNA synthetase recognises tRNA^{Ser} by its characteristic tertiary structure. *J. Mol. Biol.*, **236**, 738–748.

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_4A synthesis by seryl-tRNA synthetase. *Structure*, **3**, 341–352.

Betts,L., Xiang,S., Short,S.A., Wolfenden,R. and Carter,C.W. (1994) Cytidine deaminase. The 2.3 Å crystal structure of an enzyme:transition state analog complex. *J. Mol. Biol.*, **235**, 635–656.

Biou,V., Yaremchuk,A., 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.

Borel,F., Vincent,C., Leberman,R. and Härtlein,M. (1994) Seyl-tRNA synthetase from *E.coli*: implication of its N-terminal domain in aminoacylation activity and specificity. *Nucleic Acids Res.*, **22**, 2963–2969.

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

Burley,S.K. and Petsko,G.A. (1985) Aromatic–aromatic interaction: a mechanism of protein stabilisation. *Science*, **229**, 23–28.

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., Eriani,G., Rees,B., Ruff,M., Boeglin,M., Mitscler,A., Martin,F., Gangloff,J., Thierry,J.-C. and Moras,D. (1994) The active site of yeast aspartyl-tRNA synthetase: structural and functional aspects of the aminoacylation reaction. *EMBO J.*, **13**, 327–337.

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

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.

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. (1990) Partition of tRNA synthetases into two classes based on mutually exclusive sets of sequences motif. *Nature*, **347**, 203–206.

Fourme,R., Dhez,P., Benoit,J.P., Kahn,R., Dubisson,J.M., Besson,P. and Frouin,J. (1992) Bent crystal, bent multilayer optics on a multipole wiggler line for an X-ray diffractometer with an imaging plate detector. *Rev. Sci. Instrum.*, **63**, 982–987.

Fujinaga,M., Berthet-Colominas,C., Yaremchuck,A.D., Tukalo,M.A. and Cusack,S. (1993) Refined crystal structure of the seryl-tRNA synthetase from *Thermus thermophilus* at 2.5 Å resolution. *J. Mol. Biol.*, **234**, 222–233.

Härtlein,M. and Cusack,S. (1995) Structure, function and evolution of seryl-tRNA synthetases: implications for the evolution of aminoacyl-tRNA synthetases and the genetic code. *J. Mol. Evol.*, **40**, 519–530.

Hasegawa,T., Miyano,M., Himeno,H., Sano,Y., Kinura,K. and Shimizu,M. (1992) Identity elements of *E.coli* threonine tRNA. *Biochem. Biophys. Res. Commun.*, **184**, 478–484.

Himeno,H., Hasegawa,T., Ueda,T., Watanabe,K. and Shimizu,M. (1991) Conversion of aminacylation specificity from $tRNA^{Tyr}$ to $tRNA^{Ser}$ *in vitro*. *Nucleic Acids Res.*, **18**, 6815–6819.

Jones,A.T. and Thirrup,S. (1986) Using known structures in protein model building and crystallography. *EMBO J.*, **5**, 819–822.

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, UK.

Mattaj,I.W. and Nagai,K. (1995) Recruiting proteins to the RNA world. *Nature Struct. Biol.*, **2**, 518–522.

Normanly,J., Ogden,R.C., Horvath,S.J. and Abelson,J. (1986) Changing the identity of a transfer RNA. *Nature*, **321**, 213–219.

Normanly,J., Ollick,T., Abelson,J. (1992) Eight base changes are sufficient to convert a leucine-inserting tRNA into a serine-inserting tRNA. *Proc. Natl Acad. Sci. USA*, **89**, 5680–5684.

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.

Perona,J.J., Rould,M.A. and Steitz,T.A. (1993) The structural basis of transfer RNA aminoacylation by *E.coli* glutaminyl-tRNA synthetase. *Biochemistry*, **32**, 8758–8771.

Pley,H.W., Flaherty,K.M. and McKay,D.B. (1994) Three-dimensional structure of a hammerhead ribozyme. *Nature*, **372**, 68–74.

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

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.

Ruff,M., Krishnaswamy,S., Boeglin,M., Poterszman,A., Mitschler,A., Podjarny,A., Rees,B., Thierry,J.C. and Moras,D. (1991) Class II aminoacyl-tRNA synthetases: crystal structure of yeast aspartyl-tRNA synthetase complexed with $tRNA^{Asp}$. *Science*, **252**, 1682–1689.

Saks,M.E. and Sampson,J.R. (1996) Variant minihelix RNAs reveal sequence-specific recognition of the helical $tRNA^{ser}$ acceptor stem by *E.coli* seryl-tRNA synthetase. *EMBO J.*, **15**, 2843–2849.

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

Sampson,J.R. and Saks,M.E. (1993) Contributions of discrete $tRNA^{Ser}$ domains to amino-acylation by *E.coli* seryl-tRNA synthetase: a kinetic analysis using model RNA substrates. *Nucleic Acids Res.*, **21**, 4467–4475.

Vincent,C., Borel,F., Willison,J.C., Leberman,R. and Hartlein,M. (1995) Seryl-tRNA synthetase from *E.coli*: functional evidence for cross-dimer tRNA binding during aminoacylation. *Nucleic Acids Res.*, **23**, 1113–1118.

Wu,X.-Q. and Gross,H.J. (1993) The long extra arms of human $tRNA^{(Ser)Sec}$ and $tRNA^{Ser}$ function as major identity elements for serylation in an orientation-dependednt, but not sequence-specific manner. *Nucleic Acids Res.*, **21**, 5589–5594.

Yaremchuk,A.D., Tukalo,M.A., Krikliviy,I., Malchenko,N., Biou,V., Berthet-Colominas,C. and Cusack,S. (1992) A new crystal form of the complex between seryl-tRNA synthetase and $tRNA^{ser}$ from *Thermus thermophilus* that diffracts to 2.8 Å resolution. *FEBS Lett.*, **310**, 157–161.

Received on December 18, 1995; revised on January 23, 1996