The EMBO Journal Vol.18 No.22 pp.6532–6541, 1999

Synthesis of aspartyl-tRNA^{Asp} in *Escherichia coli*—a snapshot of the second step

S.Eiler, A.-C.Dock-Bregeon, L.Moulinier, J.-C.Thierry and D.Moras[1]

UPR 9004, Laboratoire de Biologie Structurale, Institut de Génétique et de Biologie Moléculaire et Cellulaire, CNRS/INSERM/ULP, 1 rue Laurent Fries, BP 163, 67404 Illkirch Cedex, C.U. de Strasbourg, France

[1]Corresponding author
e-mail: moras@igbmc.u-strasbg.fr

The 2.4 Å crystal structure of the *Escherichia coli* aspartyl-tRNA synthetase (AspRS)–tRNA^{Asp}–aspartyl-adenylate complex shows the two substrates poised for the transfer of the aspartic acid moiety from the adenylate to the 3′-hydroxyl of the terminal adenosine of the tRNA. A general molecular mechanism is proposed for the second step of the aspartylation reaction that accounts for the observed conformational changes, notably in the active site pocket. The stabilization of the transition state is mediated essentially by two amino acids: the class II invariant arginine of motif 2 and the eubacterial-specific Gln231, which in eukaryotes and archaea is replaced by a structurally non-homologous serine. Two archetypal RNA–protein modes of interactions are observed: the anticodon stem–loop, including the wobble base Q, binds to the N-terminal β-barrel domain through direct protein–RNA interactions, while the binding of the acceptor stem involves both direct and water-mediated hydrogen bonds in an original recognition scheme.

Keywords: aminoacylation reaction/aspartyl-tRNA synthetase/crystal structure/Q-base/tRNA

Introduction

The aminoacylation reaction is a key step in the translation of genetic information. Each aminoacyl-tRNA synthetase (aaRS) catalyses the specific attachment of its amino acid substrate to the 3′ extremity of its cognate tRNA through a two-step reaction: the activation of the amino acid substrate by ATP in the presence of magnesium followed by the transfer of the amino acid moiety to the tRNA. This reaction is performed in two topologically distinct types of active sites reflected in the partition of aaRSs based on their sequences and biochemical characteristics (Eriani *et al.*, 1990a). Class I and II aaRSs aminoacylate the 2′ OH and 3′ OH of the 3′-terminal ribose, respectively. Although the aminoacylation reaction proceeds in both classes through an essentially conserved chemical mechanism (Arnez and Moras, 1997), the strategy for the specific recognition of the amino acid and tRNA substrates is unique to each synthetase.

The first step of the aminoacylation reaction, the activation step, proceeds through an in-line nucleophilic displacement mechanism with the formation of a penta-coordinated phosphorus transition state. The binding of the ATP and amino acid necessary for the formation of the energy-rich aminoacyl-adenylate was described for different systems in enzymes of both classes (for a review, see First, 1998). The first high-resolution structure of an aminoacyl-adenylate formed in a crystal of a class II aaRS was observed in the structure of *Thermus thermophilus* aspartyl-tRNA synthetase (Poterszman *et al.*, 1994). However, it is only recently that the various steps of the aspartyl-adenylate formation could be described at very high resolution in the archaeal *Pyrococcus kodakaraensis KOD* AspRS (Schmitt *et al.*, 1998). This structure provided a model also valid for eukaryotic AspRSs based on the high conservation of the catalytic site residues in both worlds.

Our present understanding of the second step of the aminoacylation reaction, the transfer of the amino acid moiety from the adenylate to the tRNA, is essentially based on the study of the glutaminyl system for class I aaRSs (Rould *et al.*, 1989; Perona *et al.*, 1991; Rath *et al.*, 1998) and of yeast AspRS for class II enzymes (Ruff *et al.*, 1991; Cavarelli *et al.*, 1994). In the present study, we describe the crystal structure of the active eubacterial (*Escherichia coli*) class II complex between AspRS, tRNA^{Asp} and aspartyl-adenylate. This structure presents the best available resolution for a comprehensive description of an active form of an aaRS–tRNA complex. It highlights the role of water molecules in the recognition of the substrates and provides a clear picture of the side chains of all amino acid residues involved in the second step of the aspartylation reaction. A molecular mechanism that accounts for the observed conformational changes and provides an explanation for the conservation of all amino acid residues involved can now be proposed. This work reveals the characteristics of eubacterial enzymes as opposed to those from eukaryotes and archaea, and has potential applications for the development of specific inhibitors, i.e. drugs that would inhibit a pathogen synthetase but not its human counterpart (Francklyn *et al.*, 1998).

Results

Overview

A global view of the dimeric complex is shown in Figure 1A; its approximate dimensions are 80 × 115 × 130 Å. A continuous electron density allowed precise model building of the enzyme, tRNA and aspartyl-adenylate with correct geometry. The structure of the *E.coli* AspRS resembles that of the *T.thermophilus* enzyme (Delarue *et al.*, 1994) with a large dimeric interface of 5200 Å^2 (Figure 1A). Each monomer is made up of four modules (Delarue and Moras, 1993) as shown in Figure 1B and C: (i) the N-terminal domain responsible for the tRNA

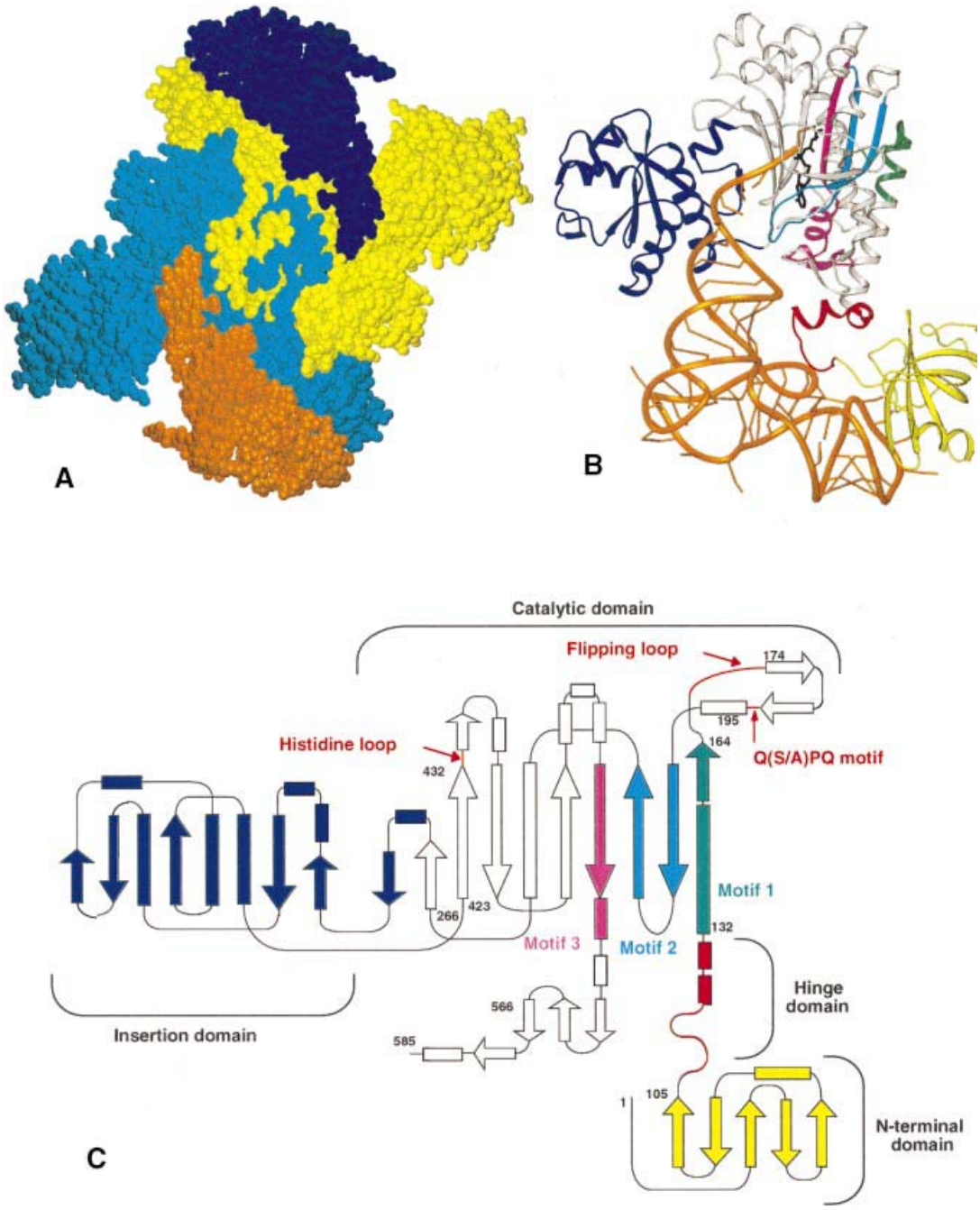

Fig. 1. (**A**) CPK representation of the dimeric *E.coli* AspRS–tRNA^{Asp} complex. The protein subunits are coloured in yellow and cyan with their cognate tRNAs in blue and orange, respectively. (**B**) Ribbon representation of one monomer of the complex showing the domain architecture of AspRS with: (i) the N-terminal domain (residues 1–108) coloured in yellow; (ii) the small hinge module (residues 109–131) in red; (iii) the catalytic domain (residues 132–270 and 422–585) in grey; and (iv) the insertion domain characteristic of eubacterial AspRSs (residues 271–421) in blue. The three signature motifs characteristic of class II aaRSs are shown in green (motif 1), cyan (motif 2) and magenta (motif 3). The tRNA is shown in orange. (**C**) Topology diagram of *E.coli* AspRS. The β-strands are represented as arrows and the helices as rods. Motif 1 is coloured in green, motif 2 in cyan and motif 3 in magenta. The same colour code is used for both (B) and (C). Figures 1–4 were generated using the Program SETOR (Evans, 1998).

anticodon recognition (this module resembles the OB-fold as defined by Murzin, 1993) formed by a five-stranded β-barrel with an α-helix between strands S3 and S4, also interacting with the catalytic domain of the other subunit; (ii) a small hinge domain encompassing a loop of variable sequence and length involved in tRNA recognition followed by a helix–loop–helix motif conserved in all AspRSs; (iii) the catalytic domain comprising the three signature motifs characteristic of class II aaRSs and the C-terminal extension specific to eubacterial AspRSs (Delarue *et al.*, 1994); and (iv) a large insertion domain characteristic of eubacterial AspRS located between motifs 2 and 3 of the catalytic site and composed of a five-stranded antiparallel β-sheet flanked by an α-helix on one side and two helices on the other side. Each domain superposes well with its counterpart in *T.thermophilus* free

enzyme. However, when the superposition is optimized for the catalytic domains, rigid body rotations of 4° and 9° are required to superpose the N-terminal and the eubacterial insertion domains. These conformational changes may be associated with tRNA binding.

Each tRNA molecule interacts only with one monomer and makes contacts with the four domains of the protein (Figure 1B). The three-dimensional structure of $tRNA^{Asp}$ is very similar to that of $tRNA^{Asp}$ from yeast, as observed in the yeast AspRS–$tRNA^{Asp}$ complex. The angle between the two arms of the L-shaped molecule is 93°, significantly smaller than the 110° observed in the free yeast $tRNA^{Asp}$ (Moras *et al.*, 1980; Westhof *et al.*, 1985). The specificity of the recognition of *E.coli* $tRNA^{Asp}$ by AspRS is based upon eight identity elements, three of which are of major importance (Nameki *et al.*, 1992). These are located in the anticodon loop (Q34 and U35) and the acceptor stem (the discriminator base G73) as shown in Figure 2A. The enzyme approaches the $tRNA^{Asp}$ on its variable loop side and binds on the inner side of the L, triggering direct or water-mediated interactions with the four domains of one synthetase subunit. Most interactions with the anticodon loop from C32 to C38 and the single-stranded 3′ GCCA end are base-specific, while those with the double-stranded acceptor stem involve mainly the ribose–phosphate backbone. Water molecules involved in tRNA recognition could clearly be identified and are characterized by comparatively low temperature factors.

Anticodon loop recognition

The mode of binding of the anticodon stem and loop of the tRNA to the N-terminal domain of AspRS is similar to what was described in the previous complexes of subclass IIb synthetases (Cavarelli *et al.*, 1994; Cusack *et al.*, 1996). The similarity is especially marked for U35, a major identity determinant in all tRNA partners of this group of synthetases (AspRS, AsnRS and LysRS). Interactions take place with Phe35 (stacking), Arg28 and Gln46, all three conserved in class IIb enzymes. Hydrophobic residues play a major role in imposing the conformation of the loop, where bases 33–37 unstack and bulge out. In particular, Leu30 intercalates between the non-Watson–Crick loop closing base pair C32–C38 and the ribose–phosphate chain of nucleotides U33, Q34 and U35. The loop conformation is stabilized further by intra-RNA hydrogen bonds between 2′ OH of ribose U35 and O4′ of C38, and between 2′ OH of ribose C38 and O2P of C36. The anticodon bases U35 and C36 sit in a hydrophobic pocket lined by residues Leu33, Phe35 and Phe48 and interact specifically with the enzyme. The base 2-methyl-A37 bulges out like the 1-methyl-G37 in the yeast complex. While the N2 atom of 1-methyl-G37 in the yeast $tRNA^{Asp}$ interacts with the phosphate group of U25, no such interaction is observed in the *E.coli* complex.

Two main differences characterize the *E.coli* complex. First, compared with the yeast $tRNA^{Asp}$, the *E.coli* $tRNA^{Asp}$ has a hypermodified base queuine at position 34. Figure 2B shows four hydrogen bonds made with the purine ring and the ribose. No specific interaction is made with the bulky modification of the Q base. The 4,5-*cis*-dihydroxy-1-cyclopentenyl-3-aminomethyl (DHCPAM) group points away from the tRNA backbone and does not engage in any polar interaction with the protein. This implies that it does not participate in the recognition of this base but could act as an antideterminant, in agreement with mutagenesis experiments (Pütz *et al.*, 1991; Martin *et al.*, 1993). Gln93, which interacts with the N1 and N2 atoms of the purine ring, is also conserved in AsnRSs. The wobble base of $tRNA^{Asn}$ is G or Q as in $tRNA^{Asp}$, suggesting a conserved mode of interaction with their synthetases. However, the hydrogen bonds with Arg76 and Asn82 are specific to the *E.coli* complex. These amino acids belong to a loop located between the β-strands S4 and S5 of the OB-fold, which is also involved in the binding of C36.

Base C36 differentiates the $tRNA^{Asp}$ anticodon triplet from that of $tRNA^{Asn}$, whose anticodon is G/QUU. The N4 position of the base hydrogen-bonds with Asn82, the N3 position with Arg78 and the O2 position with Asn84. These interactions differ from those in the yeast AspRS complex where the contacts are made with the protein backbone (Cavarelli *et al.*, 1994). In both cases, the residues involved belong to the S4–S5 loop whose length varies according to the three worlds of evolution. In some genomes of archaea, the absence of an AsnRS has been noted and Asn-$tRNA^{Asn}$ is synthesized through a trans-amidation reaction from Asp-$tRNA^{Asn}$ (Curnow *et al.*, 1996). In these organisms, the shorter S4–S5 loop allows a C or a U to be accommodated (Schmitt *et al.*, 1998) by removing C-specific interactions.

Acceptor arm recognition

At the corner of the L, the hinge domain of the AspRS comes into contact with the minor groove side of the D stem and C67 on the acceptor stem (Figure 1B). The identity determinant G10 is bound specifically to Asp111. Other interactions involve ribose–phosphate groups of the D stem (Table II). All these interactions contribute to position the tRNA by anchoring the acceptor arm.

The acceptor stem of the tRNA is positioned in a cleft formed by the insertion domain and the catalytic site (Figure 2C). Several loops and helices of AspRS are in contact with the tRNA acceptor stem: a mobile loop called the flipping loop as defined in Schmitt *et al.* (1998) (see Discussion below), the motif 2 loop (a signature peptide of class II aaRSs), a histidine loop, which is a eubacterial AspRS-specific motif located on the N-terminal side of motif 3, the C-terminal extremity and two helices from the insertion domain (Figure 2C). The tRNA arm is fully solvated by a layer of water molecules. It contacts the insertion domain through water-mediated interactions only, while on the catalytic domain side both water-mediated and direct protein–RNA interactions are observed (Table II). Base-specific contacts are formed with G73 and C74. The amino acids involved belong to the motif 2 loop. The discriminator base G73 binds to Asp220 and Arg222, both conserved within all eubacterial AspRSs sequences. Base C74 interacts with Arg225 and its ribose with Glu219, both residues being conserved in all AspRSs. Neither the first base pair G1–C72, nor the second base pair G2–C71, a minor determinant, make any direct base-specific interactions with the protein.

At the level of the tRNA, an interesting conformational change is observed. The acceptor stem contains a wobble pair, G4–U69, which, when changed to A4–U69 or G4–C69, has a small effect on aminoacylation (Nameki *et al.*, 1992). The backbone dihedral angles α and γ of G4 are

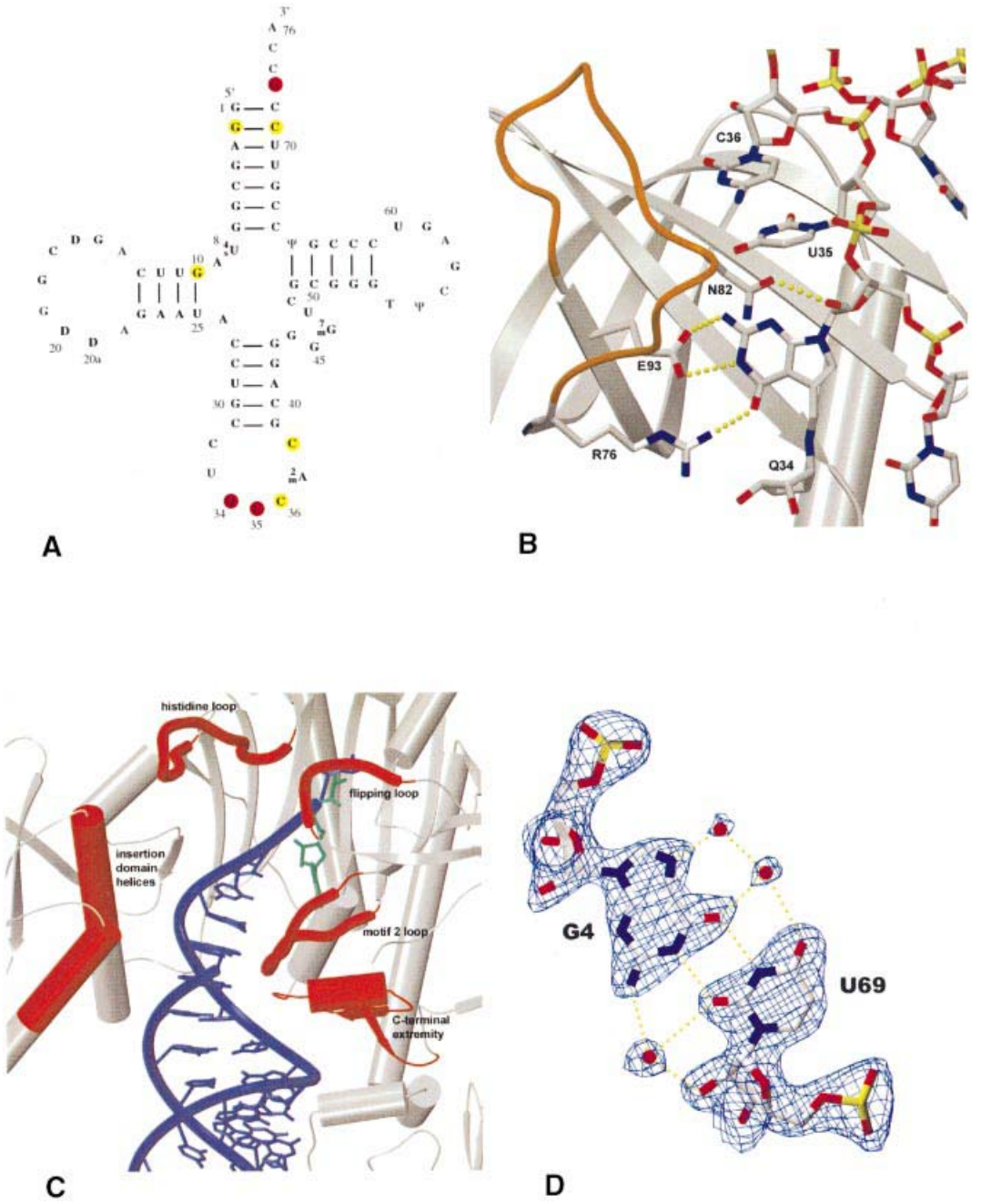

Fig. 2. (**A**) Cloverleaf representation of *E.coli* tRNA^{Asp}. The circles indicate the positions of the identity elements according to Nameki *et al.* (1992). Red circles correspond to major identity elements and yellow circles to minor determinants. (**B**) AspRS–wobble base interaction: hydrogen bonds between the protein and Q34 are shown as yellow dotted lines. Oxygen atoms are represented in red, nitrogen atoms in blue. The N1 and N2 atoms of the purine ring interact with the carboxyl side chain of Glu93, O6 makes a salt bridge with Arg76 and the 2′ OH group hydrogen-bonds to Asn82. The loop between strands S4 and S5, which is missing in archaeal AspRSs, is represented in orange. (**C**) AspRS–tRNA^{Asp} acceptor stem contacts: the interacting loops and helices from the AspRS catalytic domain and the eubacterial insertion domain are shown in red. These include the flipping loop (residues 167–173), the motif 2 loop (residues 216–228), the histidine loop (a eubacterial aspartic acid specific motif located on the N-terminal side of motif 3, residues 436–449), the C-terminal extremity (residues 544–565) and two helices belonging to the insertion domain (residues 335–344 and 399–409). (**D**) Pattern of solvation around the G–U pair. The electron density map shown results from a $2F_{obs} - F_{calc}$ synthesis and is contoured at 1.0 standard deviation.

gauche$^-$/gauche$^+$, a regular RNA conformation, while they are *trans*/*trans* in the free tRNA. This is most likely to be the result of the tRNA binding to the protein. The free functional groups of the purine–pyrimidine G–U mismatch are hydrogen bonded to solvent molecules (Figure 2D) with a pattern of solvation similar to that observed in other G–U-containing structures (Biswas *et al.*, 1997; Ramos and Varani, 1997).

An unexpected observation derived from the present structure concerns the recognition patterns of the acceptor arm of the tRNAs. In yeast, the protein–tRNA interactions are concentrated at the acceptor end of the arm, a feature characteristic of eukaryotes and probably of archaea, as deduced from sequence conservation. In *E.coli* and most likely in all eubacteria, the numerous interactions are located all along the acceptor arm.

Table I. Data collection and refinement statistics for AspRS–tRNA^{Asp}–aspartyl-adenylate complex

Data set	
Resolution (Å)	12–2.4
No. of unique reflections	46 618
Redundancy	5.6
Completeness (%)[a]	98.1 (93.2)
R_{sym} (I) (%)[b]	5.9 (21.4)
Refinement statistics	
R-factor (%)	20.1
No. of reflections in working set	43 082 (90.7%)
Free R-factor (%)	24.9
No. of reflections in test set	3536 (7.4%)
No. of water molecules	514
Average B-factor (Å^2)	28.9
enzyme	24.3
tRNA	42.2
aspartyl-adenylate	17.3
water	29.9
R.m.s. deviation	
bond lengths (Å)	0.006
bond angles (°)	1.324
dihedral (°)	26.731

[a]The values in parentheses correspond to the last shell of resolution (2.46–2.4 Å).

[b]$R_{sym}(I) = \frac{\sum_{hkl}\sum_i |\langle I_{hkl}\rangle - I_{hkl,i}|}{\sum_{hkl}\sum_i |I_{hkl}|}$ where i represents one measure of reflection hkl.

Table II. Direct and water-mediated hydrogen bond interactions in the *E.coli* AspRS–tRNA^{Asp}–aspartyl-adenylate complex between the tRNA and the protein from the same subunit

Acceptor stem			D stem		
C75	*O2.....W.....Nη2*	*Arg537*	G10	N2.....Oδ2	Asp111
C74	N3.....Nη1	Arg225	U11	O2.....Nε2	His114
C74	O2.....Nη2	Arg225	U11	O2′.....Oδ1	Asn116
C74	O2′.....Oε2	Glu219	U12	O2P.....Oγ	Thr117
C74	*O2.....W.....Oε2*	*Glu219*	U12	O2P.....NH	Thr117
C74	*O2.....W.....NH*	*Asp220*	*U12*	*O2′..W.....CO*	*Val115*
C74	*N4.....W.....Nη1*	*Arg225*	U25	O2′.....Oδ2	Asp 111
C74	*O2′.....W.....Nη2*	*Arg174*	*U25*	*O2′.....Oδ1*	*Asn113*
G73	*N1.....Oδ1*	*Asp220*	*A26*	*O2P.W.....Oδ1*	*Asn 65*
G73	O6.....Nη1	Arg222			
G73	N7.....Nη2	Arg222			
G73	*O2P....W.....Nz*	*Lys400*			
G73	*O2P....W.....Oδ1*	*Asp404*	Anticodon stem and loop		
G73	*N2.....W.....Oδ1*	*Asp220*	*C27*	*O2P.W.....Oδ1*	*Asn65*
G73	*N2.....W.....Nη1*	*Arg181*[a]	*C27*	*O1P.W.....Nη1*	*Arg27*
C72	O1P.....Nη2	Arg222	C28	O1P.....Nε	Arg64
C72	*O2′.....W.....CO*	*Ile343*	C32	N4.....CO	Asp29
C72	*O2P....W.....Oη*	*Tyr344*	U33	O5′.....Oγ	Ser32
C72	*N4.....W.....CO*	*Leu221*	Q34	N1.....Oε1	Glu93
C71	O1P.....NH	Ala223	Q34	N2.....Oε2	Glu93
C71	*O2′.....W.....Oη*	*Tyr344*	Q34	O6.....Nη2	Arg76
C71	*O2P....W.....Oδ2*	*Asp224*	Q34	O2′.....Oδ1	Asn82
U70	O2P.....Nη2	Arg549	U35	O2.....Nη1	Arg28
U70	*O4.....W.....Oγ*	*Thr558*	U35	N3.....Oε1	Gln46
U69	O2P.....Nη1	Arg549	U35	O4.....Nη2	Arg78
U69	O1P.....NH	Thr558	C36	N4.....CO	Asn82
G68	O2P.....Oγ	Thr557	C36	N3.....Nη1	Arg78
G68	*O1P....W.....Oγ*	*Thr558*	C36	O2.....Nδ	Asn84
C67	O2′.....Oε1	Glu119	C38	O1P.....Nε	Arg28
C67	*O2.....W.....Oε1*	*Glu119*	C38	N4.....CO	Asp29
G6	*N2.....W.....Oε1*	*Glu119*	*C38*	*O2P.W.....CO*	*Arg27*

[a]Arg181 belongs to the second subunit.

The active site pocket

The aspartyl-adenylate molecule lies at the bottom of the catalytic domain, stretched on the antiparallel β-sheet and held in place by a network of interactions illustrated in Figure 3A. The residues involved belong principally to the class II motifs 2 and 3 and to an AspRS invariant LXQ(S/A)PQXXKQ sequence (residues 190–199). It is remarkable to note that this motif alone allows retrieval of all AspRSs in sequence databases.

A superposition of the catalytic site of this complex with the other known three-dimensional structures of AspRSs shows the strict conservation of the position and conformation of the aspartyl-adenylate. The various aspartyl-adenylates superpose with an r.m.s. difference of <0.7 Å and the ribose adopts a conserved C3′ endo conformation. The ATP and aspartic acid substrates individually have been shown to be superposable on their corresponding counterparts in the aspartyl-adenylate complex (Schmitt *et al.*, 1998). The present study shows that this remains true even in the presence of the tRNA.

The aspartic acid moiety is anchored by several AspRSs invariants: Arg489, Lys198, Gln195, Asp233 and Ser193 (Figure 3A). The water molecule, which bridges the amino group of the aspartyl-adenylate with Asp233 and Ser193, was also observed in the structures of *P.kodakaraensis KOD* and *T.thermophilus* enzymes. The AMP moiety is positioned by class II invariants (Phe229, Arg217 and Arg537) and by AspRSs conserved residues (Gln192, Asp475 and Glu482). Main chain interactions anchor the adenine moiety that is sandwiched between the class II conserved motif 3 arginine (Arg537) and motif 2 phenylalanine (Phe229) (Figure 3A).

Among the residues responsible for the recognition of the aspartyl-adenylate, two are specific for eubacteria. The most important contribution is that of Gln231, which interacts through its Nε with the α-phosphate (O5′) and carbonyl atom of the aspartyl-adenylate (Figure 3B). This interaction, first observed in the *T.thermophilus* AspRS, is not present in archaeal or eukaryotic AspRSs. Instead, in these enzymes, a serine located in a different strand of the catalytic β-sheet (Ser364 in the *P.kodakaraensis KOD*, AspRS and Ser481 in the yeast AspRS) appears to play a similar important role, as shown in Figure 3B. The second is His448, a eubacterial conserved residue with no direct equivalent in eukaryotes or archaea (Figure 3A and B). It interacts specifically with the Oδ of the aspartyl-adenylate and contributes to a water-mediated recognition of the α-phosphate. In the absence of aspartyl-adenylate, the histidine side chain faces the solvent (data not shown). His448 belongs to the H(H/N)λF(T/S) sequence (histidine loop) conserved in all eubacteria, which corresponds to the (R/K)PFYX sequence found in all eukaryotic and archaebacterial AspRSs.

At the tRNA acceptor end, the base of the terminal adenosine sits on top of the aspartyl-adenylate, in a large and open pocket, flanked on one side by a hydrophobic surface (Leu196, Pro450, Phe451 and Phe514) (Figure 4A). The flipping loop together with the conserved AspRS invariant sequence [LXQ(S/A)PQXXKQ] and the histidine loop [H(H/N)λF(T/S)] constitute the other faces of the binding site. The base is fixed in the pocket through interactions between N6 and the side chain hydroxyl of Thr169 and the main chain carbonyl of Pro170, two

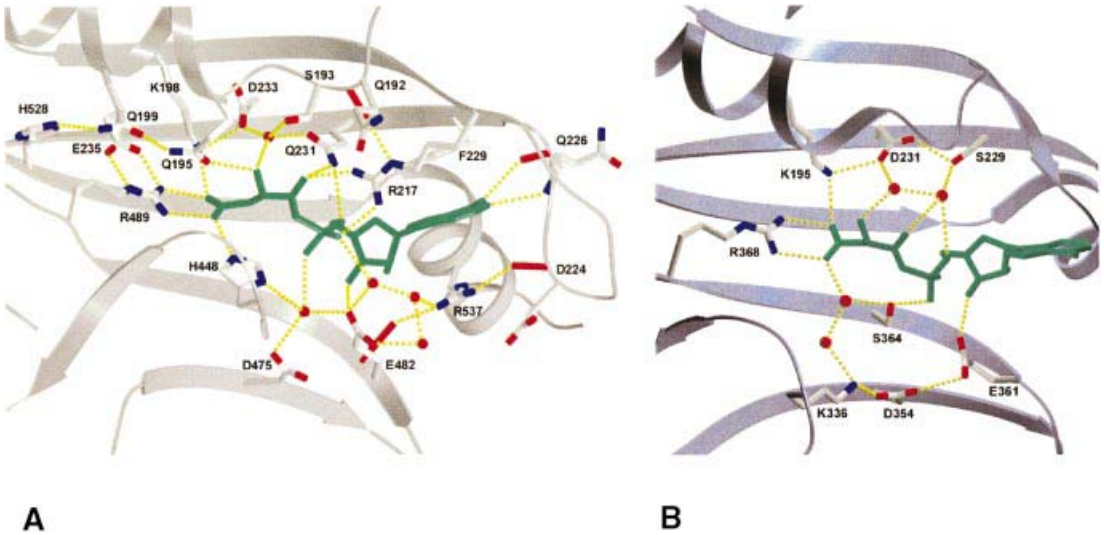

Fig. 3. (**A**) Recognition of the aspartyl-adenylate (in green) by the *E.coli* AspRS. The protein oxygen atoms are represented in red and the nitrogen atoms in blue. The hydrogen bonds between the protein and its substrate are shown as yellow dotted lines. Five water molecules directly responsible for the recognition of the substrate are represented as red spheres. (**B**) Structure of the AspRS–aspartyl-adenylate complex in *P.kodakaraensis KOD*. The stabilization of the intermediate states is mediated essentially by the conserved arginine of motif 2 (not shown in this figure) and a second residue specific to eubacteria (Gln231 in *E.coli*) or to eukaryotes and archaebacteria (Ser364 in *P.kodakaraensis KOD*). The two residues are located on two different strands of the β-sheet and are therefore not structurally homologous. The residue homologous to the eubacterial glutamine is not conserved within eukaryotes and archaebacteria (Ser229 in *P.kodakaraensis KOD* and Gly340 in yeast). His448 in *E.coli* and Lys336 in *P.kodakaraensis KOD* provide another example of the differences associated with each division.

residues of the flipping loop. The ribose is held in place by Ser193, conserved in all AspRSs. A superposition of the active sites of the yeast and *E.coli* complexes shows the identical positioning of A76 in the two structures, with an r.m.s. difference of 0.5 Å.

Discussion

The role of water molecules in tRNA recognition

Water molecules are found to play important roles at protein–nucleic acid interfaces; they fill cavities or surround the regions where protein and nucleic acid atoms in contact are fully buried (Nadassy *et al.*, 1999). The structure of the AspRS–tRNA^{Asp} complex shows an extensive contact surface including direct and solvent-mediated hydrogen bonds. Due to the high quality of the electron density map and the resolution, the water molecules found at the interface between the tRNA and the synthetase (see Table II) have been located reliably. Their distribution varies from one domain to the other and appears to depend on the role of each domain in the recognition process. Two extreme situations are observed. (i) The N-terminal domain anchors the anticodon loop of the tRNA through direct hydrophobic and hydrogen bonding interactions, which contribute to the specific recognition of the two molecules, in agreement with the presence of four out of the eight specificity determinants. No water molecules are seen at this interface. (ii) The acceptor stem binding involves mostly water-mediated interactions. More than 20 water molecules build a shell at the interface between the insertion domain and the tRNA through a network of hydrogen bonds, of which only three are direct protein–water–RNA contacts. The layer of water molecules could favour a dynamic recognition by the formation of a network of non-specific and versatile water-mediated interactions, in contrast to the tryptophan repressor–operator system where water-mediated contacts contribute to specificity (Otwinowski *et al.*, 1988). A hybrid situation is observed at the interface between the catalytic domain of the synthetase and the acceptor stem of the tRNA. The recognition process involves direct and water-mediated interactions (Table II) and proceeds through mutual conformational adjustments of the tRNA and synthetase (the flipping and motif 2 loops, and the C-terminal extremity of the synthetase). These conformational changes can be deduced from the structures of the free *E.coli* enzyme (B.Rees, G.Webster, M.Boeglin, A.C.Dock-Bregeon and D.Moras, in preparation) and of a complex between the *E.coli* tRNA^{Asp} and *T.thermophilus* enzyme (C.Briand, A.Poterszman, G.Webster, J.C.Thierry and D.Moras, in preparation).

The presence of water molecules could be associated with the need for larger adaptability and dynamics during recognition. In contrast, each time a specificity of recognition and/or a precise positioning of the substrates is required (i.e. for the catalysis), the system will tend as much as possible to avoid water-mediated interactions.

The transfer of the aspartic acid moiety to the 3′ OH of A76, the role of the flipping loop

tRNA–aaRS interactions have a unique goal: the precise positioning of the 3′ OH of the ribose of A76 of the cognate tRNA in the catalytic site of the enzyme in order to permit the transfer of the amino acid moiety from the adenylate to tRNA. The location of A76 can vary with the presence and nature of the other substrates (ATP or aspartic acid). Different conformations have been observed in the crystal structure of the heterologous *E.coli* AspRS–yeast tRNA^{Asp} complex (in preparation). A76 gets into its 'functional' location in the active site of the enzyme in a concerted movement with the flipping loop. When comparing all known three-dimensional structures of AspRSs

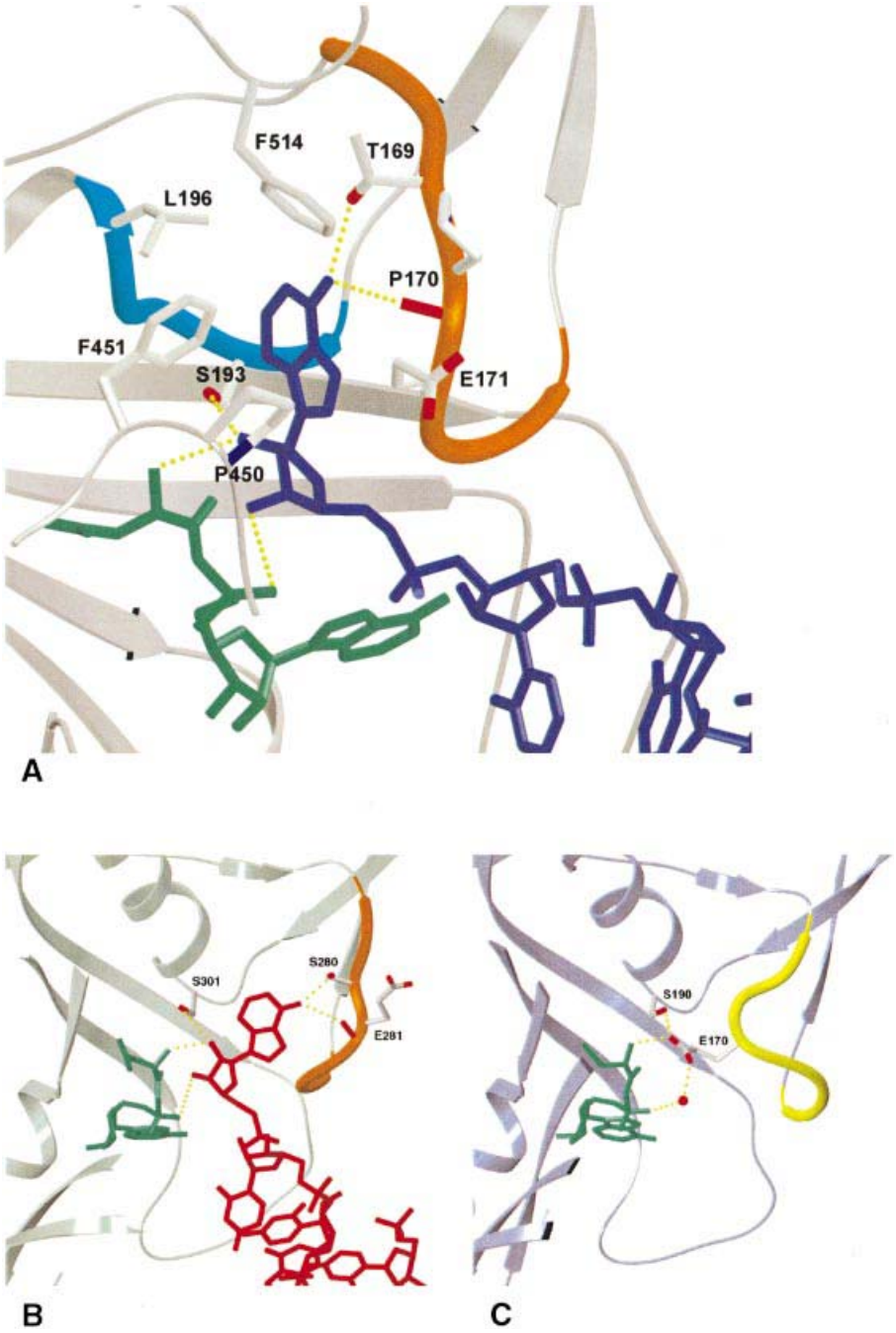

Fig. 4. (**A**) Recognition of the terminal adenosine of $tRNA^{Asp}$ by the catalytic site of AspRS in *E.coli*. The tRNA molecule is shown in purple, the aspartyl-adenylate in green and the enzyme in grey, with the oxygen atoms in red and the nitrogen atoms in blue. The loops involved in the binding of A76 are coloured in orange (flipping loop, residues 167–173) and blue [residues 192–196 of the AspRS invariant LXQ(S/A)PQXXKQ sequence]. The residues of the synthetase forming the hydrophobic surface of the recognition pocket (Pro450, Phe451, Leu196 and Phe514) are also shown. Thr169 and the carbonyl group of Pro170 fix the position of the base, while Ser193 specifically binds to and orients the ribose by interacting with its 2′ OH. The ribose also interacts through its 2′ OH with the amino group of the aspartic acid moiety, while the 3′ OH is in a position favouring both (i) the transfer of its proton to one of the oxygen atoms of the α-phosphate of the AMP moiety and (ii) the nucleophilic attack by the 3′ oxygen atom of A76 on the α-carbonyl carbon of the aspartyl-adenylate. (**B**) Structure of the yeast AspRS–$tRNA^{Asp}$–aspartyl-adenylate complex with the flipping loop coloured in orange in its 'open' conformation. The loop contributes to the positioning of A76 in order to permit the transfer of the aspartic acid to the ribose 3′ oxygen. (**C**) Structure of the *P.kodakaraensis KOD* AspRS–aspartyl-adenylate complex with the flipping loop coloured in yellow in its 'closed' conformation.

complexed with their different substrates, it appears that the 'flipping loop' can adopt an 'open' or a 'closed' conformation (Figure 4). Each of the two conformations can be associated with a specific role depending on the presence of the substrates. The flipping loop is closed when the aspartic acid or the aspartyl-adenylate is bound to the enzyme and when the tRNA is absent. This has been observed in the structures of AspRS from *T.thermophilus* (Poterszman *et al.*, 1994) and *P.kodakaraensis* (Schmitt *et al.*, 1998). The side chain carboxylate of the conserved glutamic acid (Glu171 in *E.coli* and Glu170 in *P.kodakaraensis*) blocks the amino group of

the aspartic acid in its recognition site as shown on Figure 4C. The open conformation of the flipping loop is encountered in two situations. First, it allows the amino acid to get into its recognition site, as observed in the structures of apo AspRSs from *T.thermophilus*, *P.kodakaraensis* and *E.coli*, and in the structure of AspRS from *P.kodakaraensis* complexed with ATP. Secondly, it no longer interacts with the aspartyl-adenylate but is responsible for the accurate recognition and positioning of the terminal adenine of the tRNA through main chain interactions with N6 of the adenine, as seen in the structures of AspRSs complexed with their cognate tRNA^{Asp} in the presence of aspartyl-adenylate (Figure 4A and B).

The flipping loop anchors the base of A76 in a way that allows the formation of three interactions essential for adopting the transfer step conformation (Figure 4A). The 2′ OH of the ribose interacts with both (i) the conserved serine of the LXQ(S/A)PQXXKQ sequence and (ii) the amino group of the aspartyl-adenylate. For the few AspRSs that possess an alanine instead of a serine (three out of 54 known AspRS sequences), the asparagine from the histidine loop can be modelled functionally to replace the serine in order to interact either with the glutamate from the flipping loop or with the ribose of A76. The 3′ OH donates a hydrogen bond to the OPα of the adenylate, which helps to orient the attacking 3′ oxygen and favours the uptake of the hydroxyl proton by the α-phosphate group. No AspRS residue or water molecule is seen to interact with the 3′ OH. This structure differs from that observed in the case of the class I glutaminyl-tRNA synthetase complexed with tRNA^{Gln} and a glutaminyl-adenylate analogue, where the nucleophilicity of the 2′ oxygen is increased through a water-mediated interaction with a glutamic acid residue (Rath *et al.*, 1998).

The conformation of A76 favours the nucleophilic attack of the 3′ OH on the α-carbonyl carbon of the aspartyl-adenylate. Its position and orientation appear to be strictly conserved in the structures of AspRS complexes from yeast and *E.coli* and can be correlated with the conserved conformation and position of the aspartyl-adenylate as observed in all the structures of the AspRSs studied so far.

The stabilization of the transition state is mediated by essentially two amino acid residues (Figure 5), the motif 2 arginine (Arg217) and a glutamine (Gln231) that is conserved only in eubacteria and replaced by a serine in the eukaryotic and archaeal AspRSs (Ser364 in *P.kodakaraensis KOD* and Ser481 in yeast). Gln231, by the polarization contribution of its amido nitrogen, and Arg217 act as electron attractors on the carbonyl oxygen of the aspartyl-adenylate and participate in the delocalization of the negative charges of the α-phosphate. This is the only eubacterial specificity in the mechanism of transfer of the aspartic acid to the tRNA^{Asp}. Two variants of a general aspartylation mechanism can be proposed that take into account these conserved changes. One is characteristic of eukaryotes and archaea and the other of eubacteria (Figure 5).

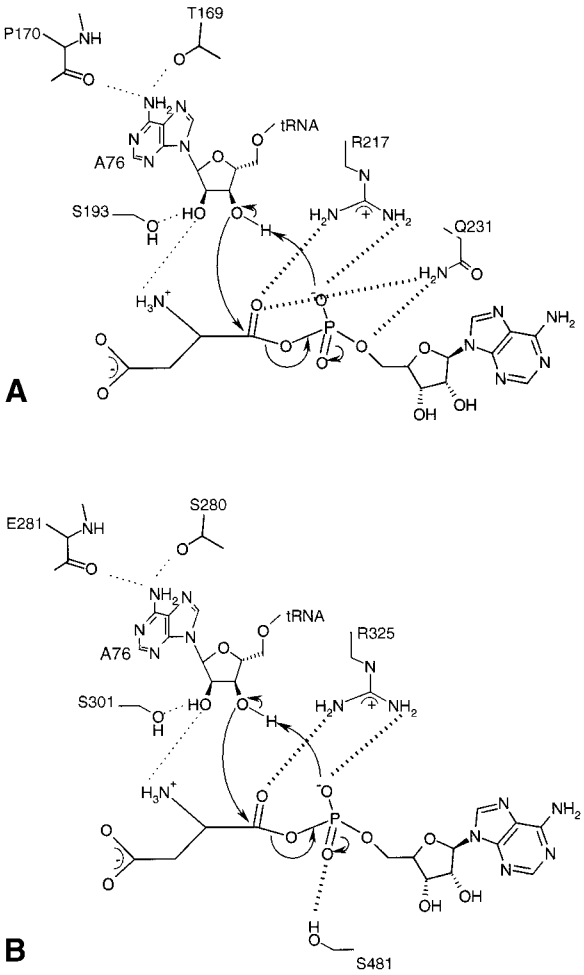

Fig. 5. Schematic drawing of the interactions between the aspartyl-adenylate, the AspRS and the terminal adenine of tRNA^{Asp} in (**A**) *E.coli* and (**B**) yeast with a modelled mechanism for the transfer of the aspartic moiety of the aspartyl-adenylate to the 3′ OH of A76 of the tRNA. Arg217, a residue conserved in all AspRSs, is within electrostatic interaction range of the oxygen atoms from the carbonyl and the phosphate groups of the aspartyl-adenylate. This arginine favours and stabilizes the negative charges on these two oxygen atoms. A glutamine residue conserved in all eubacteria (Glu231 in *E.coli*) or a serine in eukaryotes and archaebacteria (Ser481 in yeast) also contribute to reinforce the role of the arginine residue in the formation of a nucleophilic centre on the carbonyl carbon atom. The transfer of the aspartic acid to the tRNA would proceed through the formation of a tetrahedral intermediate formed during the nucleophilic attack on the carbonyl carbon of the adenylate by the oxygen of the 3′ OH of A76. This 3′ oxygen atom is within interaction range of the electrophilic carbonyl carbon centre and well oriented to allow the transfer of the aspartic acid to the tRNA. The hydroxyl proton is within interaction range of the negatively charged phosphate oxygen to be captured directly by the phosphate group.

Concluding remarks

This study allows comparison of the *E.coli* complex with that of yeast. These are prototypical of eubacteria and eukaryotes or archaea, respectively (sequence analysis shows that eukaryotes and archaea are closely related in the catalytic domain). The binding of the insertion domain, specific to eubacteria, to the tRNA is not sequence specific and all interactions are water mediated. The bulky modification of the *E.coli* hypermodified Q-base is not involved in protein recognition. In contrast to these results, two major differences between the *E.coli* and yeast complex can be pointed out: (i) the mode of recognition of the acceptor arm, which in *E.coli* binds to the enzyme all along the arm backbone, while in yeast the interactions are concentrated at the acceptor end; and (ii) the activation

of the two steps of the reaction uses a glutamine in eubacteria (Gln231 in *E.coli*), replaced by a serine in eukaryotes (Ser481 in yeast) and archaea.

Altogether, the present structure provides an explanation for the conformational changes and the conservation of all amino acid residues involved in the aspartylation reaction. For example, the role of the flipping loop, a major player in the control of the two steps of the aspartylation reaction, is now fully understood. In the absence of tRNA, the serine of the LXQ(S/A)PQXXKQ sequence and the glutamate of the flipping loop are responsible for the stabilization of the aspartyl-adenylate in its recognition site (Figure 4C). When the tRNA is positioned correctly (Figure 4A and B), the glutamate flips towards the solvent and the serine binds the ribose of A76. The 'opening' or 'closing' of the flipping loop is achieved by anchoring its two extremities through direct interactions between the main chain amino group of Thr166 and the carbonyl group of Ala189 at one end, and between Ala173 and Gln192 at the other end. These interactions are conserved in all structures examined so far in the AspRS system.

In addition, the novel description of water in tRNA–synthetase interactions offers new perspectives for RNA recognition. Notably the role of the solvent molecules in the selectivity process needs further investigations.

Materials and methods

Expression and purification

The *E.coli* AspRS gene was cloned in pBR322 and expressed in XL1Blue. The recombinant protein was purified in two chromatographic steps adapted from Eriani *et al*. (1990b) using a Q-Sepharose column and a Toyopearl TSK Butyl column.

Overproduction of $tRNA^{Asp}$ was carried out by growing cultures of *E.coli* cells transformed by the plasmid pTrc99-B-$tDNA^{Asp}$ (Martin *et al.*, 1993). $tRNA^{Asp}$ represents 40% of the total RNA extracted from the cells and was purified in two chromatographic steps, anion exchange (DEAE) followed by hydrophobic interactions chromatography (Eiler *et al.*, 1992).

Crystallization and data collection

Crystallization was done using the hanging drop vapour diffusion method. Drops were equilibrated against a reservoir containing 15% glycerol and 2 M ammonium sulfate in 50 mM Bis-Tris-propane–HCl (BTP–HCl) pH 6.8 buffer at 4°C; they were prepared by mixing 2 µl of protein solution, 2 µl of tRNA solution, 2 µl of substrate solution and 2 µl of reservoir solution. The protein solution contained 10 mg/ml AspRS in 10 mM Tris–HCl buffer pH 7.5. The tRNA solution contained 4.7 mg/ml $tRNA^{Asp}$ (i.e. 2.5 times the molar concentration of the enzyme), 2 mM $MgCl_2$ and 0.25 mM EDTA in 5 mM sodium cacodylate buffer pH 6.5. The substrate solution contained 10 mM ATP, 25 mM aspartic acid and 25 mM $MgSO_4$. Under these conditions, crystals of an average size $500 \times 500 \times 400$ µm^3 were obtained within 2 weeks. They belong to space group $P4_32_12$ with unit cell dimensions $a = b = 101.20$ Å and $c = 231.81$ Å. The dyad axis of the homodimeric complex coincides with a crystallographic 2-fold axis; thus the crystallographic asymmetric unit contains only one monomer of AspRS with a bound aspartyl-adenylate and one tRNA molecule. The adenylate was thus synthesized in the crystallization drop. No transfer of aspartate on the tRNA was observed; this was probably due to the high concentration of ammonium ions and to the slightly acidic pH of the solution (pH ~6.5), lower than the optimal pH for this reaction (pH 7.6). X-ray diffraction data were collected at the synchrotron radiation line D41 ($\lambda = 1.3675$ Å) of the Laboratoire pour l'Utilisation du Rayonnement Electromagnétique (LURE, Orsay, France) using a Mar Research (Hamburg, Germany) imaging plate detector. The data set was collected at 120 K from a single flash-cooled crystal. To impair ice formation, the crystal was transferred into 50 mM BTP–HCl pH 6.8 buffer containing 2.0 M ammonium sulfate and 25% glycerol for 15 s and then plunged into liquid ethane. Raw data (278 frames of 0.25° oscillation with a crystal–image plate distance of 230 mm) were processed with the programs DENZO and SCALEPACK (Otwinowski and Minor, 1996). The merging *R*-value for all measurements was 5.9% (Table I).

Structure determination and refinement

The structure was solved by the molecular replacement method with the program AMoRe (Navaza, 1994) using the yeast AspRS–$tRNA^{Asp}$ complex as the search model (Ruff *et al.*, 1991). After determination of the orientation and position of the monomer in the cell, the yeast enzyme was replaced by the *T.thermophilus* AspRS (Delarue *et al.*, 1994) in order to take into account the large insertion domain present in eubacterial AspRSs. The model was then improved by rigid body refinement and several rounds of model building and simulated annealing with the program CNS (Brünger *et al.*, 1998) using the bulk solvent correction as implemented in CNS. Water molecules were added at positive peaks over 3.5 σ in the $(F_o - F_c)$ maps. In the final stage, cartesian coordinate refinement was followed by individual *B*-factor refinement. The modelling operations were done with the program O (Jones *et al.*, 1991).

The model contains one monomer of AspRS, one tRNA molecule, one aspartyl-adenylate substrate and 514 water molecules. The enzyme is lacking five residues out of 590 at the C-terminal end. The quality of the refined model was assessed using the Biotech Validation Suite for Protein Structures (Wodak *et al.*, 1995). The crystallographic *R*-factor is 20.1% using all reflections between 12 and 2.4 Å with no cut-off (R_{free} = 24.9%). The average *B*-factor for all non-hydrogen atoms is 29.3 Å^2, a value in agreement with the overall *B*-factor determined by Wilson plot on the collected data (B = 34.5 Å^2). The r.m.s. coordinate error is 0.36 Å as estimated by the method of Luzzati (1952). The model shows good geometry and stereochemistry as analysed using the program PROCHECK (Laskowski *et al.*, 1993). All angles but one (φ, ψ) are within the allowed regions of the Ramachandran plot, with 91% in the most favoured regions. Despite our efforts in rebuilding the Glu171 in the electron density, the refinement brings this residue outside the allowed regions of the Ramachandran plot.

Acknowledgements

We thank the staff of LURE, ESRF and DESY for their help during data collection. We also thank Jean Gangloff and Gilbert Eriani from the Unité 'Structure des Macromolécules Biologiques et Mécanismes de Reconnaissance' and our colleagues of the Laboratoire de Biologie Structurale for fruitful discussions, and John G.Arnez of the Albert Einstein College of Medicine for comments on the manuscript. This work was supported by grants from EU project No. BIO4-98-0189, CNRS, INSERM, ULP and Ministère de la Recherche et de la Technologie. Coordinates for the X-ray structure of the *E.coli* AspRS–$tRNA^{Asp}$–aspartyl-adenylate complex have been deposited in the Brookhaven Protein Data Bank under ID code 1C0A.

References

Arnez,J.G. and Moras,D. (1997) Structural and functional considerations of the aminoacylation reaction. *Trends Biochem. Sci.*, **22**, 211–216.

Biswas,R., Wahl,M.C., Ban,C. and Sundaralingam,M. (1997) Crystal structure of an alternating octamer r (GUAUGUA)dC with adjacent G–U wobble pairs. *J. Mol. Biol.*, **267**, 1149–1156.

Brünger,A.T. *et al*. (1998) Crystallography and NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr. D*, **54**, 905–921.

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.

Curnow,A., Ibba,M. and Söll,D. (1996) tRNA dependent asparagine formation. *Nature*, **382**, 589–590.

Cusack,S., Yaremchuk,A. and Tukalo,M. (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. *EMBO J.*, **15**, 6321–6334.

Delarue,M. and Moras,D. (1993) The aminoacyl-tRNA synthetase family: modules at work. *BioEssays*, **15**, 675–687.

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.

Eiler,S., Boeglin,M., Martin,F., Eriani,G., Gangloff,J., Thierry,J.C. and Moras,D. (1992) Crystallization of aspartyl-tRNA synthetase–tRNA (Asp) complex from *Escherichia coli* and first crystallographic results. *J. Mol. Biol.*, **224**, 1171–1173.

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 sequence motifs. *Nature*, **347**, 203–206.

Eriani,G., Dirheimer,G. and Gangloff,J. (1990b) Aspartyl-tRNA synthetase from *E.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.

Evans,S. (1998) SETOR: hardware lighted three-dimensional model representations of macromolecules. *J. Mol. Graph.*, **11**, 134–138.

First,E.A. (1998) Catalysis of tRNA aminoacylation by class I and class II aminoacyl-tRNA synthetases. In M.Sinnott (ed.), *Comprehensive Biological Catalysis*. Academic Press, London, UK, pp. 573–607.

Francklyn,C., Musier-Forsyth,K. and Martinis,S.A. (1998) Aminoacyl-tRNA synthetases in biology and disease: new evidence for structural and functional diversity in an ancient family of enzymes. *RNA*, **3**, 954–960.

Jones,T.A., Zou,J.Y., Cowan,S.W. and Kjeldgaard,M. (1991) Improved methods for the building of protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A*, **47**, 110–119.

Laskowski,R.A., MacArthur,M.W., Moss,D.S. and Thornton,J.M. (1993) PROCHECK: a program to check the stereochemical quality of protein structure. *J. Appl. Crystallogr.*, **26**, 283–291.

Luzzati,P.V. (1952) Traitement statistique des erreurs dans la détermination des structures cristallines. *Acta Crystallogr.*, **5**, 802–810.

Martin,F., Eriani,G., Eiler,S., Moras,D., Dirheimer,G. and Gangloff,J. (1993) Overproduction and purification of native and queuine-lacking *Escherichia coli* tRNA^{Asp}, role of the wobble base in tRNA^{Asp} Acylation. *J. Mol. Biol.*, **234**, 965–974.

Moras,D., Comarmond,M.B., Fischer,J., Weiss,R., Thierry,J.C., Ebel,J.P. and Giegé,R. (1980) Crystal structure of yeast tRNA^{Asp}. *Nature*, **288**, 669–674.

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

Nadassy,K., Wodak,S.J. and Janin,J. (1999) Structural features of protein–nucleic acid recognition sites. *Biochemistry*, **38**, 1999–2017.

Nameki,N., Tamura,K., Himeno,H., Asahara,H., Hasegawa,T. and Shimizu,M. (1992) *Escherichia coli* tRNA^{Asp} recognition mechanism differing from that of the yeast system. *Biochem. Biophys. Res. Commun.*, **189**, 856–862.

Navaza,J. (1994) *AMoRe*: an automated package for molecular replacement. *Acta Crystallogr. A*, **50**, 157–163.

Otwinowski,Z. and Minor,W. (1996) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.*, **276**, 307–326.

Otwinowski,Z., Schevitz,R.W., Zhang,R.-G., Lawson,C.L., Joachimiak,A., Marmorstein,R.Q., Luisi,B.F. and Sigler,P.B. (1988) Crystal structure of *trp* repressor/operator complex at atomic resolution. *Nature*, **335**, 321–329.

Perona,J.J., Rould,M.A., Steitz,T.A., Risler,J.L., Zelwer,C. and Brunie,S. (1991) Structural similarities in glutaminyl- and methionyl-tRNA synthetases suggest a common overall orientation of tRNA binding. *Proc. Natl Acad. Sci. USA*, **88**, 2903–2907.

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

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

Ramos,A. and Varani,G. (1997) Structure of the acceptor stem of *Escherichia coli* tRNA^{Ala}: role of the G3–U70 base pair in synthetase recognition. *Nucleic Acids Res.*, **25**, 2083–2090.

Rath,V.L., Silvian,L.F., Beijer,B., Sproat,B.S. and Steitz,T.A. (1998) How glutaminyl-tRNA synthetase selects glutamine. *Structure*, **6**, 439–449.

Rould,M.A., Perona,J.J., Soll,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 transfer RNA synthetases: crystal structure of yeast aspartyl-tRNA synthetase complexed with tRNA (Asp). *Science*, **252**, 1682–1689.

Schmitt,E., Moulinier,L., Fujiwara,S., Imanaka,T., Thierry,J.C. and Moras,D. (1998) Crystal structure of aspartyl-tRNA synthetase from *Pyrococcus kodakaraensis* KOD: archaeon specificity and catalytic mechanism of adenylate formation. *EMBO J.*, **17**, 5227–5237.

Westhof,E., Dumas,P. and Moras,D. (1985) Crystallographic refinement of yeast aspartic acid transfer RNA. *J. Mol. Biol.*, **184**, 119–145.

Wodak,S.J., Pontius,J., Vaguine,A. and Richelle,J. (1995) Validating protein structures. From consistency checking to quality assessement. In *Proceedings of the CCP4 Daresbury Study Weekend: Making the Most of Your Model*. SERC Daresbury Laboratory, Warrington, UK, pp. 41–51.

Received July 29, 1999; revised and accepted September 28, 1999