The structural basis for seryl-adenylate and Ap$_4$A synthesis by seryl-tRNA synthetase

Hassan Belrhali[1], Anya Yaremchuk[1], Michael Tukalo[1], Carmen Berthet-Colominas[1], Bjarne Rasmussen[1], Peter Bösecke[2], Olivier Diat[2] and Stephen Cusack[1]*

[1]European Molecular Biology Laboratory, Grenoble Outstation, c/o ILL, 156X, F-38042 Grenoble Cedex 9, France and
[2]European Synchrotron Radiation Facility, B.P. 220, F-38043 Grenoble Cedex 9, France

Background: Seryl-tRNA synthetase is a homodimeric class II aminoacyl-tRNA synthetase that specifically charges cognate tRNAs with serine. In the first step of this two-step reaction, Mg·ATP and serine react to form the activated intermediate, seryl-adenylate. The serine is subsequently transferred to the 3'-end of the tRNA. In common with most other aminoacyl-tRNA synthetases, seryl-tRNA synthetase is capable of synthesizing diadenosine tetraphosphate (Ap$_4$A) from the enzyme-bound adenylate intermediate and a second molecule of ATP. Understanding the structural basis for the substrate specificity and the catalytic mechanism of aminoacyl-tRNA synthetases is of considerable general interest because of the fundamental importance of these enzymes to protein biosynthesis in all living cells.

Results: Crystal structures of three complexes of seryl-tRNA synthetase from *Thermus thermophilus* are described. The first complex is of the enzyme with ATP and Mn^{2+}. The ATP is found in an unusual bent conformation, stabilized by interactions with conserved arginines and three manganese ions. The second complex contains seryl-adenylate in the active site, enzymatically produced in the crystal after soaking with ATP, serine and Mn^{2+}. The third complex is between the enzyme, Ap$_4$A and Mn^{2+}. All three structures exhibit a common Mn^{2+} site in which the cation is coordinated by two active-site residues in addition to the α-phosphate group from the bound ligands.
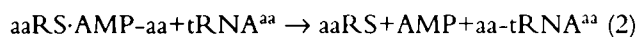
Conclusions: Superposition of these structures allows a common reaction mechanism for seryl-adenylate and Ap$_4$A formation to be proposed. The bent conformation of the ATP and the position of the serine are consistent with nucleophilic attack of the serine carboxyl group on the α-phosphate by an in-line displacement mechanism leading to the release of the inorganic pyrophosphate. A second ATP molecule can bind with its γ-phosphate group in the same position as the β-phosphate of the original ATP. This can attack the seryl-adenylate with the formation of Ap$_4$A by an identical in-line mechanism in the reverse direction. The divalent cation is essential for both reactions and may be directly involved in stabilizing the transition state.

Introduction

Aminoacyl-tRNA synthetases (aaRSs) are a family of enzymes responsible for the specific esterification of an amino acid (aa) to the 3'-end of its cognate tRNA isoacceptor(s) (tRNA^{aa}). This is achieved through a two-step reaction: the enzyme first activates the amino acid using Mg·ATP to form the enzyme bound aminoacyl-adenylate (aa-AMP) intermediate (Eqn. 1; Fig. 1). Then the amino acid is transferred to either the 2' or 3' position of the 3'-terminal ribose of the cognate tRNA (Eqn. 2).

$$aaRS + ATP + aa \overset{Mg^{2+}}{\longleftrightarrow} aaRS \cdot AMP\text{-}aa + PPi \qquad (1)$$

$$aaRS \cdot AMP\text{-}aa + tRNA^{aa} \rightarrow aaRS + AMP + aa\text{-}tRNA^{aa} \qquad (2)$$

Primary sequence analysis and crystallographic structure determination of aaRSs have led to the partition of these enzymes into two classes, each with 10 members [1–3]. Class I synthetases contain two short consensus amino acid motifs, HIGH and KMSKS, that are far apart in the primary sequence but in close proximity in the known three-dimensional structures. Both motifs are implicated

in ATP binding and amino acid activation [4,5]. Class I synthetases have a typical Rossmann dinucleotide-binding fold for the catalytic domain, based on a parallel β-sheet (reviewed in [6]). Class II aaRSs, on the other hand, share three consensus motifs (designated motifs 1, 2 and 3) and a catalytic domain with an antiparallel β-sheet topology. This domain topology has only recently been found in one other protein, biotin synthetase [7], which is also an ATP-binding and activating enzyme. Motif 1 forms part of the conserved inter-subunit interface of homodimeric [8,9] and heterodimeric [10] class II synthetases. Motifs 2 and 3 contain many of the active-site residues important for ATP, amino acid and tRNA acceptor stem recognition [11,12]. On the basis of more extensive sequence homology, class II synthetases have been subdivided into subclass IIa (the enzymes for serine, threonine, proline and histidine) and subclass IIb (the enzymes for aspartic acid, asparagine and lysine) [8].

The partition of aaRSs into two classes is correlated with the fact that class I and class II (except PheRS) enzymes
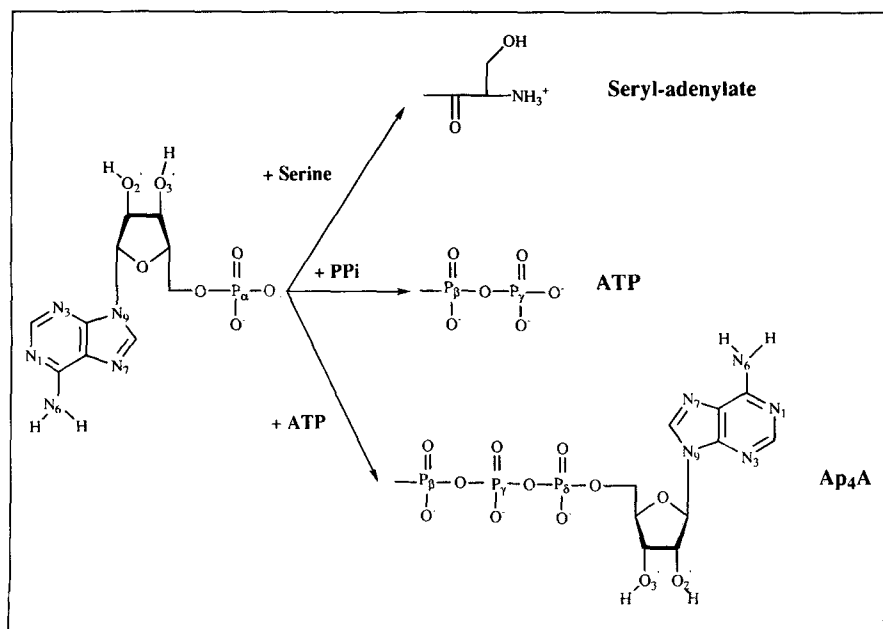
*Corresponding author.

Fig. 1. Structural formulae of the substrates discussed in this paper. The AMP moiety (drawn on the left) is common to each of them.

initially charge the 2'- or 3'- hydroxyl group, respectively, of the terminal tRNA adenosine [1]. A mechanism for amino acid activation and transfer has been proposed for the class I enzymes TyrRS [13] and GlnRS [5]. For class II enzymes, the most extensive structural information exists for SerRS [2,11,14,15] and AspRS [9,16,17]. For the former, the structures of complexes of *Thermus thermophilus* SerRS with two different analogues of seryl-adenylate have been described [11], whereas for the latter, a series of ternary complexes including the cognate tRNA has allowed the proposal of a mechanism for activation and aminoacylation by AspRS [12,17].

In addition to their direct role in protein biosynthesis, aaRSs perform various other biochemical functions. Amongst these is the ability of most aaRSs to synthesize a dinucleotide compound, diadenosine $5',5''-P^1,P^4$-tetraphosphate (AppppA or Ap_4A, see Fig. 1). This was first described by Randerath *et al.* [18] for LysRS from *Escherichia coli* who showed that this enzyme can also synthesize a number of other compounds of the form Ap_4N and Ap_3N, where N can be a variety of nucleosides. Subsequently, many other aaRSs were shown to synthesize Ap_4A *in vitro* and *in vivo* [19,20] (for review, see [21]). For Ap_4A synthesis by an aaRS, the cognate amino acid, ATP and magnesium are absolutely required; in addition, for some synthetases zinc stimulates the reaction [19,21]. The Ap_4A is formed by a reaction analogous to the reverse of amino acid activation with a second ATP molecule, rather than inorganic pyrophosphate, attacking the aminoacyl-adenylate intermediate:

$$\text{aaRS·AMP-aa+ATP} \overset{Mg^{2+}}{\longleftrightarrow} \text{aaRS+Ap}_4\text{A+aa} \qquad (3)$$
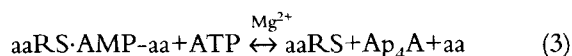
Formation of Ap_4A is favoured by the presence of pyrophosphatase that, by removing the competing pyrophosphate, favours attack of the adenylate by a second ATP molecule. Ap_4A has been found at micromolar

concentrations in a wide variety of prokaryotic and eukaryotic cells. The levels of the compound can fluctuate significantly, for instance in prokaryotic cells subject to oxidative stress. As a result, the compound has been implicated in a number of metabolic processes and stress-related phenomena (for reviews, see [21,22]). Specific enzymes (Ap_4A hydrolases and phosphorylases) exist that can degrade Ap_4A to ATP and/or ADP [21]. This may indicate that the levels of Ap_4A are regulated. On the other hand, it may simply be a mechanism to recover useful metabolites from Ap_4A, which is an inevitable by-product of aaRS activity [21].

SerRS from *E. coli* readily synthesizes Ap_4A in the presence of either magnesium or manganese ions (K Larson, personal communication). In the context of this paper, the interest in studying the Ap_4A synthesis reaction is firstly, to shed light on the mechanism of activation by SerRS, and secondly, to locate the second ATP-binding site. An obvious candidate for the binding site of the adenosine moiety is at, or close to, the site for the terminal adenosine of the tRNA. The binding site for the 3'-end of the tRNA^{Ser} in SerRS is currently unknown because of disorder of this region in the SerRS–tRNA^{Ser} complex structure [15].

In this paper we describe three new crystallographic structures at 2.3–2.6 Å resolution. These are the complexes of SerRS from *T. thermophilus* with ATP, seryl-adenylate (Ser-AMP) and Ap_4A, each in the presence of manganese. The results provide the detailed structural basis to permit a coherent understanding of the enzymatic mechanism of formation of Ser-AMP and Ap_4A and clarify the role played by the divalent cation (manganese or magnesium) in these reactions. In addition, comparisons of the structures of all the known substrate or product complexes of SerRS from *T. thermophilus* illustrate the plasticity of the active site required for completion of the two-step aminoacylation reaction.

Results

Structure of the SerRS–ATP·Mn^{2+} complex

The structure of this complex has been determined from data collected from a single crystal at 2.4 Å resolution. (Statistics of data collection for this and the two other complexes are given in Table 1.) The crystal was soaked in mother liquor containing 20 mM ATP and 20 mM MnSO$_4$ for 30 min before data collection started. A difference map with coefficients $(F_{ATP\cdot Mn}-F_c^{nat})e^{i\phi_c^{nat}}$ revealed extra positive density at a level of up to 10σ in each active site into which an ATP molecule could be fitted, leaving three additional positive peaks unfilled adjacent to the ATP phosphates (Fig. 2a). Calculation of an anomalous Fourier difference map with coefficients $(\Delta F_{ano})_{ATP\cdot Mn}$ $e^{i(\phi_c^{nat}-\pi/2)}$ showed unambiguously that this extra density corresponded to three manganese ions associated with the ATP phosphates (Fig. 2b). A model of the asymmetric unit comprising the synthetase dimer, two ATP molecules, six manganese ions and 229 water molecules was subsequently refined using standard X-PLOR energy-minimization protocols [23] to a crystallographic R-factor of 16.7% and good geometry (see Table 1).

The ATP molecule is in an unusual U-shaped conformation in which the β- and γ-phosphates are bent back into an arginine-rich pocket (comprising Arg271, Arg344 and Arg386) towards the purine ring rather than extending away from it (Figs 2, 3). A similar conformation for ATP

has been observed in the yeast AspRS–tRNA^{Asp}–ATP ternary complex [12]. Fig. 3 shows the hydrogen-bond network and the electrostatic interactions between the enzyme, an ATP molecule and three manganese ions. The interactions with the base and the ribose are as described previously [11]. The manganese site with the highest electron density bridges the α- and β-phosphates and has two ligands from the enzyme, Glu345 and Ser348; this will be referred to as the principal manganese site. Two water molecules complete the octahedral coordination shell of this ion (Fig. 4a). The residue Asp332 is 4–4.5 Å away from this manganese ion and only interacts indirectly with it via one of the coordinating water molecules. Two other manganese ions, at lower occupancy, are found either side of the Pβ–O–Pγ linkage one of them also having Glu345 as a ligand. Two arginines, Arg256 (from motif 2) and Arg386 (from motif 3), stabilize the α- and γ-phosphate groups, respectively (Fig. 3). These arginines are absolutely conserved in class II synthetases, and although disordered in the native enzyme structure [14], are well ordered in the presence of ATP. A third arginine, Arg271 (from motif 2) also stabilizes the γ-phosphate. This residue (which is either an arginine or a histidine in all class II synthetases), as well as Glu258, significantly alter their interactions with ATP depending on the presence or absence of tRNA and other substrates (see Discussion section). Numerous water molecules are found in the
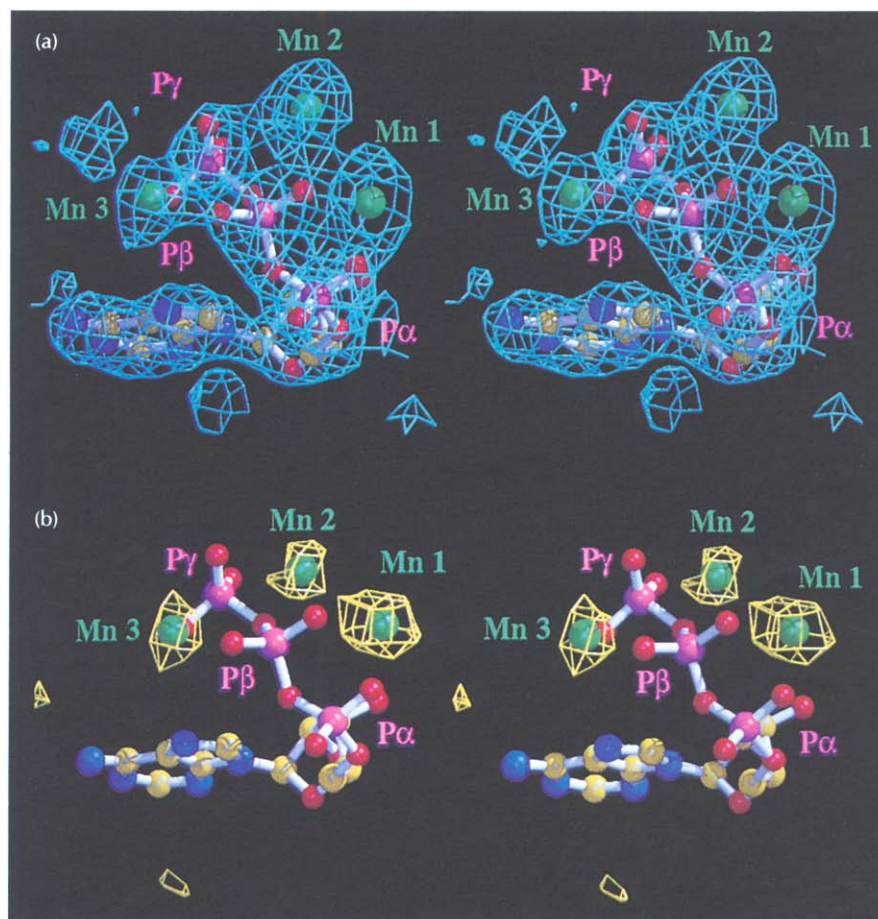

Fig. 2. (a) A stereoview of the σ_a-weighted difference Fourier map contoured at +3σ phased with a model before the inclusion of ATP. A model of an ATP molecule with three manganese ions (green spheres) is superposed. (b) Stereo anomalous difference Fourier map contoured at +3σ calculated with model phases before the inclusion of manganese ions.

Table 1. Data collection and refinement statistics.

Complex	ATP+Mn^{2+}	ATP+Mn^{2+}+serine	Ap$_4$A+Mn^{2+}
Data			
Unit cell			
a (Å)	85.85	85.35	86.49
b (Å)	126.04	125.78	127.12
c (Å)	62.85	62.59	63.26
β (°)	109.20	109.54	109.42
Space group	P2$_1$	P2$_1$	P2$_1$
Experimental site	LURE	LURE	ESRF
Experimental station	W32	W32	BL4
Insertion device type	wiggler	wiggler	undulator
Wavelength (Å)	0.901	0.901	0.906
Total no. of reflections used	153 576	77 551	221 948
Total no. of unique reflections	46 178	28 879	54 994
Completeness (%)	93.7 (81.4)^a	75.6 (73.6)^a	98.6 (93.2)^a
Average redundancy	3.3 (2.0)^a	2.7 (2.4)^a	4.0 (3.6)^a
R$_{merge}$ (%)	3.9 (12.1)^a	4.1 (10.2)^a	4.7 (13.7)^a
No. of reflections I>3σ(I)	86.6 (65.8)^a	88.2 (74.9)^a	89.2 (73.6)^a
Maximum resolution (Å)	2.40	2.60	2.32
Model			
Refinement R-factor (%)	16.7	22.5^b	18.6
Rms deviations of:			
bond lengths (Å)	0.012	–	0.011
bond angles (°)	1.55	–	1.51
dihedrals (°)	24.0	–	23.7
impropers (°)	1.4	–	1.4
No. of water molecules	229	–	186
Average B-factor (Å^2) of:			
water	32.9	–	37.1
enzyme	27.7	–	32.6
nucleotide substrate	28.4	–	51.2
B-factor of Mn^{2+} (Å^2)	34.7/43.9^c	–	34.9/50.2^c

^{a}For the highest resolution shell. ^{b}Because of the lack of complete data, only a rigid-body refinement has been performed. ^{c}Isotropic temperature factor for the principal manganese ion in the first and second active sites.

active site forming a network of hydrogen bonds amongst themselves and with the enzyme. Of particular note are three water molecules found at the positions of the Oγ, NH$_3^+$ and one of the carboxyl oxygen atoms of the substrate serine.

Structure of the SerRS–ATP-Ser·Mn^{2+} complex

A 2.60 Å resolution data set was recorded on a crystal soaked for 30 min in 10 mM serine, 10 mM Mn^{2+} and 10 mM ATP. Because of degradation of the crystal in the X-ray beam, only 75% of the data could be recorded from a single crystal. Nevertheless the data quality is good (Table 1). Calculation of a difference map with phases derived from the native model revealed strong extra positive density into which a Ser–AMP molecule could be fitted. In addition, strong density is observed for a manganese ion at the principal Mn^{2+} site described above for the ATP·Mn^{2+} complex (Fig. 5). This result shows that the enzyme is active in the crystal for the formation of Ser–AMP, as we have previously found for the formation of the adenylate analogue, serine-hydroxamate-AMP [11]. A new element here is that the enzyme can activate serine in the presence of manganese and, furthermore, that the ion is apparently not released after Ser–AMP formation. As the data for this structure are incomplete, the structure has not been refined further. Recently, a new data set has been collected at 2.6 Å resolution for an ATP-Ser·Mg^{2+} complex in 56% saturated sodium citrate rather than 40% ammonium sulphate (see below). The data are 80% complete overall and have an R$_{merge}$ of 6.2% (19.1% in the highest resolution shell). These data show very good difference density for Ser–AMP and have been refined to an R–factor of
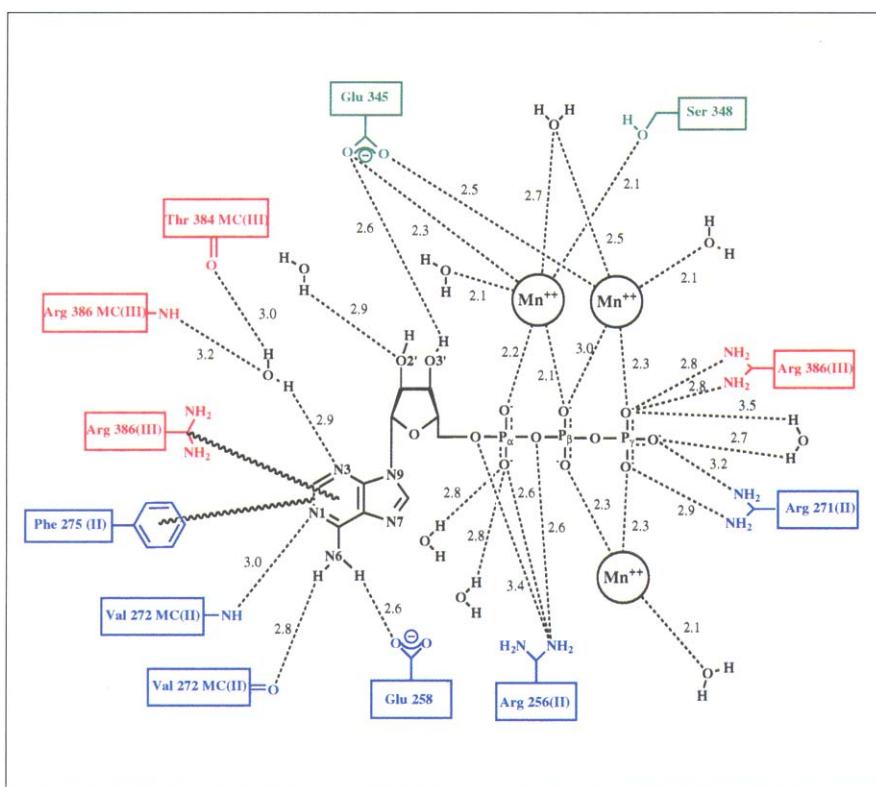

Fig. 3. Diagram showing hydrogen bonds and electrostatic interactions between the enzyme, the ATP molecule and three manganese ions in one of the two active sites of the enzyme: dashed lines represent hydrogen bonds and electrostatic interactions; wavy lines symbolize van der Waals interactions. Residues are colour-coded according to whether they belong to motif 2 (blue), motif 3 (red) or β-strand A4 (green), consistent with Fig. 3 of [11]. Distances (in Å) are between donor and acceptor atoms for the hydrogen-bonding interactions and between the two charged atoms for the electrostatic interactions.
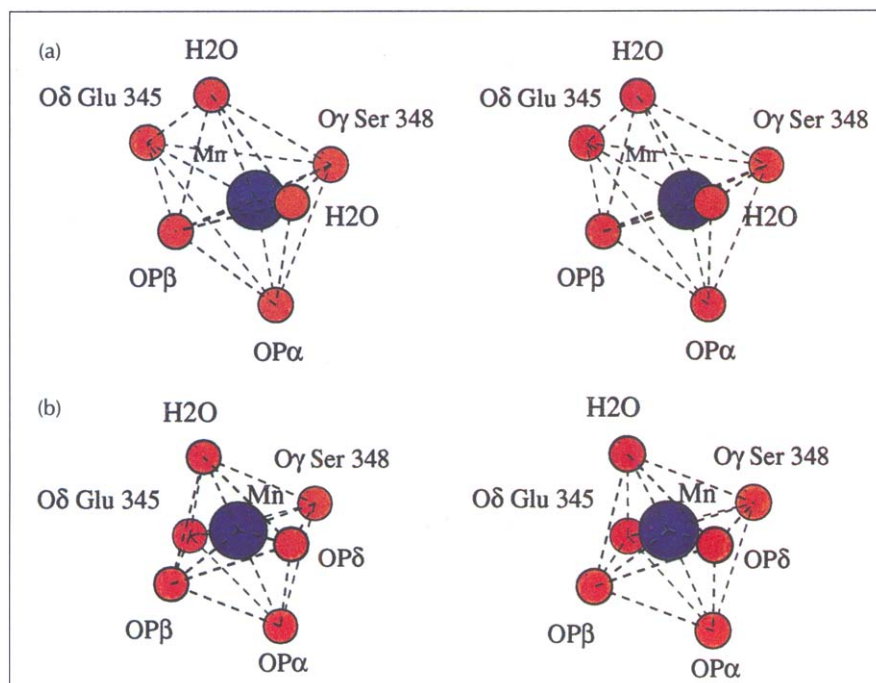
Fig. 4. Stereo diagrams of the octahedral coordination of the principal manganese ion in the complex structure between the enzyme and **(a)** ATP and **(b)** Ap₄A prepared from the final refined atomic coordinates. For distances, see Figs 3,7 respectively. The overall disposition of the atoms is a square-based bipyramid with the divalent cation (blue) at the centre.
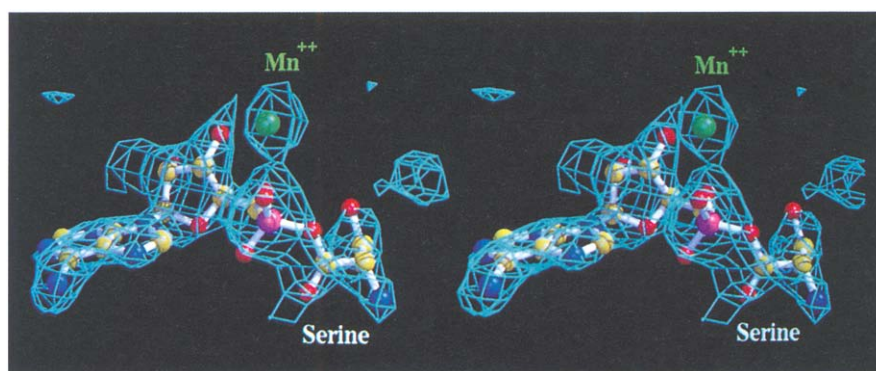

Fig. 5. Stereo σ_a-weighted difference Fourier map contoured at $+3\sigma$ showing the seryl-adenylate intermediate enzymatically formed in the crystal. The map is phased with a model before the inclusion of the ligand. Clear density is observed above the α-phosphate for a manganese ion at the same site as the principal cation in the ATP complex.

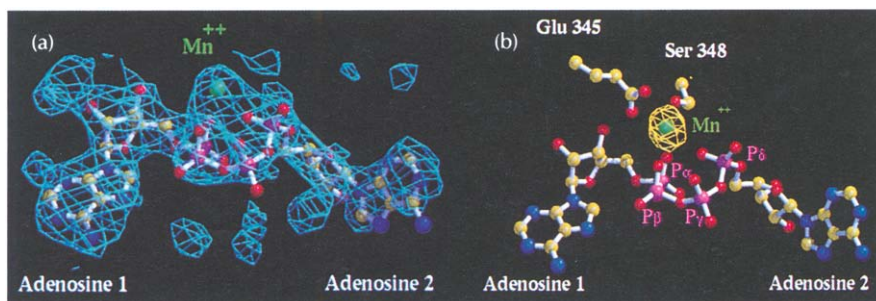**Fig. 6. (a)** σ_a-weighted difference Fourier map contoured at $+3\sigma$ showing density corresponding to an Ap₄A molecule and an associated manganese ion. **(b)** Anomalous difference Fourier map of the Ap₄A·Mn²⁺ complex contoured at $+4\sigma$ showing the presence of the manganese ion. Also shown are the two residues in the protein, Glu345 and Ser348, that coordinate the Mn²⁺ ion.

16.4%. The two structures (ATP-Ser·Mn²⁺ and ATP-Ser·Mg²⁺) both show that the specific interactions made by Ser-AMP are essentially the same as previously described for the two adenylate analogues, serine-hydroxamate-AMP and 5'-O-[N-(L-seryl)-sulfamoyl]-adenosine [11].

Structure of the SerRS–Ap₄A·Mn²⁺ complex

A complete data set at 2.32 Å resolution was recorded from a single crystal soaked for 2 h in 20 mM Ap₄A and 20 mM Mn²⁺ (Table 1). A difference Fourier map revealed extra positive density (Fig. 6a) in both active sites identifiable as an Ap₄A molecule. The four phosphate

groups encircle a manganese ion whose identity was confirmed by its anomalous signal (Fig. 6b). Details of the interactions of the Ap₄A molecule and Mn²⁺ with the protein are shown in Fig. 7. The first adenosine and the Pα and Pβ of Ap₄A are in the same conformation and make the same interactions as described for the corresponding atoms in the ATP complex. The manganese site corresponds to the principal site observed in the ATP and Ser-AMP complexes, with protein ligands Glu345 and Ser348. However, in the Ap₄A complex, the manganese ion is also coordinated by a non-bridging oxygen from the Pδ of Ap₄A but not from Pγ. The octahedral coordination of the manganese is completed by a single
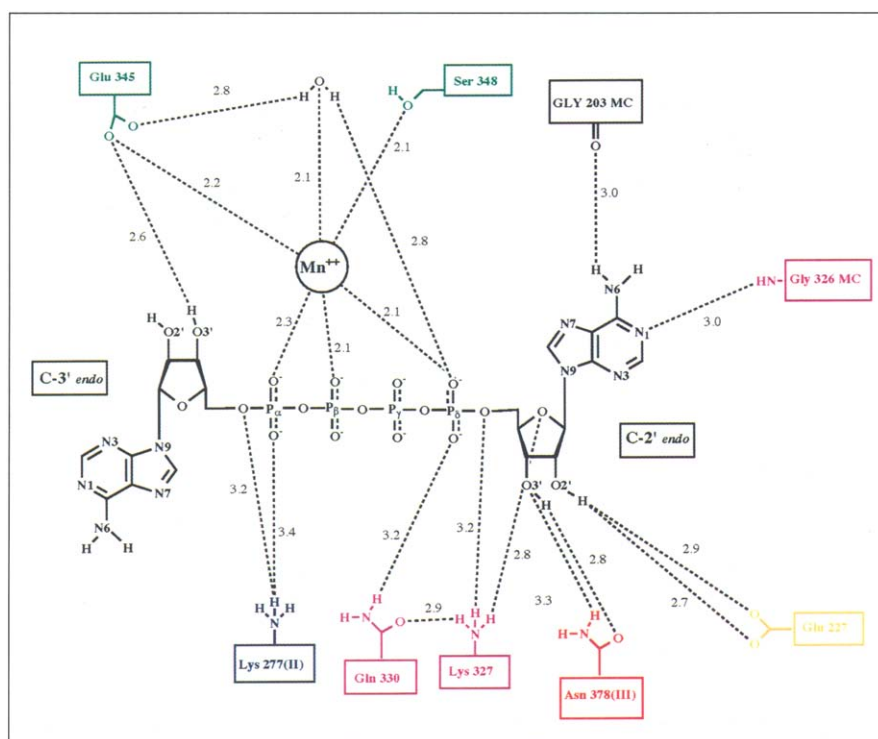
Fig. 7. Diagram showing the hydrogen bonds and the electrostatic interactions between the enzyme, Ap$_4$A molecule and single manganese ion in one of the two active sites of the dimer. Dashed lines represent hydrogen bonds and electrostatic interactions. The residues are colour-coded as in Fig. 3 and [11] with the addition of the TXE loop (yellow) and β-strand A3 (magenta). Distances (in Å) are between donor and acceptor atoms for the hydrogen-bonding interactions and between the two charged atoms for the electrostatic interactions.

water molecule (Fig. 4b). No other manganese sites are apparent. The binding of the Pδ and second adenosine of the Ap$_4$A brings into play new elements of the active site (Fig. 7). Residues Lys327 and Gln330 from β-strand A3 (defined in [11]), which are not, however, conserved in other known SerRS sequences, interact with the Pδ and second ribose. Residues Asn378 (from motif 3 and fully conserved in all known SerRS sequences) and Glu227 (from the TXE loop defined in [11]) also interact with the ribose which is in the C2′-endo conformation, in contrast to the first ribose, which is C3′-endo. As seen in Fig. 6a, the second adenine ring is in weaker electron density and appears to be much less strongly bound than the first adenine (that is, the one in common with ATP). Adenine specificity is conferred by hydrogen bonds between the N6 and N1 positions and the main chain of Gly203 and Gly326 (Fig. 7). The second adenine ring makes no stacking interactions but is in van der Waals contact with an edge of Trp355, and is situated very close to the extreme N terminus of the enzyme at the base of the helical arm. The binding of the second adenine induces a systematic distortion of this region and a disordering of the N-terminal residues 1–3. Substrate complexes in crystals of *T. thermophilus* SerRS generally show significantly higher ligand occupancy in one active site than in the second independent active site. The reason for this asymmetry has not been established. In the case of the Ap$_4$A complex, electron density for the second adenosine is largely absent in the second active site, again indicating the weaker binding of this moiety. The conformation observed for Ap$_4$A is consistent with solution studies on the compound that showed there to be a single high-affinity magnesium-binding site associated with the oligophosphate, and that metal binding leads to destacking of the bases [24].

Discussion

ATP conformation and divalent cation binding site

Above we have described an unusual bent conformation of ATP bound to $T.$ *thermophilus* SerRS that is stabilized by interactions with three manganese ions and with class II conserved arginines from motifs 2 and 3. Previously, we have shown that in the absence of added manganese, the triphosphate of ATP or AMPPCP (an ATP analogue) adopts an extended conformation ([15] and unpublished data). A similar extended conformation has been reported for AMPPCP in the yeast AspRS system [9]. There are two good reasons to believe that the bent conformation is the active conformation of ATP. Firstly, superposition of the active-site structures containing ATP in the extended conformation and Ser-AMP (or analogues) shows that the β-phosphate overlaps the serine-binding site thereby preventing simultaneous binding of serine and ATP in this conformation (data not shown). Secondly, as described below, the bent conformation of the ATP superposed on the conformation determined for Ser-AMP is clearly consistent with the expected in-line displacement mechanism for serine activation (Fig. 8a). Why then, is ATP sometimes observed in the extended conformation? This may be an artifact of the crystallization medium which contains 34–40% saturated ammonium sulphate. The magnesium ion, which is crucial for stabilizing the correct ATP conformation, apparently does not bind at the principal site under these particular crystallization conditions. Furthermore, in the absence of substrates, strong positive difference density is always observed in the arginine-rich pocket formed by Arg271, Arg386 and Arg344 which should accommodate the γ-phosphate. This density can probably be attributed to sulphate ions and/or water molecules, which therefore have to be displaced by ATP. In the absence of bound
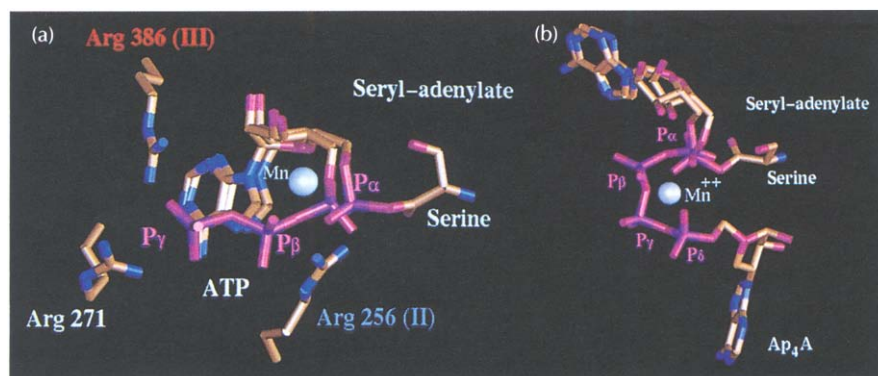
Fig. 8. (a) Superposition of the conformations of the enzyme-bound ATP and seryl-adenylate showing the common divalent cation binding site. **(b)** Superposition of the conformations of the enzyme-bound seryl-adenylate and Ap₄A showing the common divalent cation binding site.

magnesium, which by bridging Pα, Pβ and the enzyme directs the conformation of the triphosphate, displacement of solvent molecules does not occur. In addition, soaking experiments with ATP, serine and even high concentrations of magnesium do not result in Ser-AMP formation in the crystal. Manganese, on the other hand, clearly does bind in the presence of 34–40% saturated ammonium sulphate and, as shown by the second complex structure described above, allows the serine activation reaction to occur in the crystal.

The exact reason for the different behaviour of magnesium and manganese is not known. However, the interpretation we present is supported by recent experiments in which the ammonium sulphate mother liquor was replaced by transferring pre-grown crystals into 56% saturated sodium citrate. Under these conditions, both binding of ATP in the active conformation and serine activation are observed crystallographically in the presence of magnesium. Data at 2.4 Å resolution (99% complete, R_{merge}=4.3%) for the ATP·Mg²⁺ complex show a clear positive difference peak located at the principal manganese site, although no clearly interpretable density equivalent to the other two manganese sites is observed. Furthermore, data at 2.6 Å resolution for an ATP-Ser·Mg²⁺ complex, also in 56% saturated sodium citrate, show that Ser-AMP is formed under these conditions. In this structure, electron density exists which could be a remaining magnesium ion in the principal site, as was the case when manganese was used. These results show that in the presence of ATP, both magnesium and manganese can bind at the same principal site, bridging the α- and β-phosphates and coordinated by Glu345 and Ser348 from the enzyme.

In the yeast AspRS system, a putative magnesium site has been determined from data at 3 Å resolution of a ternary complex between the enzyme, cognate tRNA^Asp and ATP [12]. The magnesium is reported to be bridging the β- and γ-phosphates of ATP and to be directly coordinated by Asp471 and Glu478 [12,17]. Although these two acidic residues are apparently conserved in all class II synthetases (they are equivalent to Asp332 and Glu345, respectively, in SerRS from *T. thermophilus*), the results presented above show conclusively that in *T. thermophilus* SerRS, the principal manganese/magnesium site is

located in a slightly different position from that found in yeast AspRS. The crucial differences between yeast AspRS and *T. thermophilus* SerRS are as follows: firstly, Asp332 only makes a water-mediated interaction with the ion; secondly, the principal ion is between the α- and β-phosphates of the ATP; thirdly, even though a manganese ion is also observed bridging the β- and γ-phosphates, it is still 4.1 Å away from Asp332; and finally, the direct protein ligands are Glu345 and Ser348. Interestingly, Ser348, while conserved in all known SerRSs, is not generally conserved in other class II synthetases. Sequence alignments [8,11] suggest that this residue is functionally conserved as a threonine in several, but not all, class IIa synthetases (e.g. ThrRS and ProRS), but can be glycine, asparagine or even proline in class IIb synthetases (AspRS, AsnRS, LysRS).

The binding of ATP stabilizes several of the residues otherwise disordered in the active site. On the other hand, some of these residues alter their interactions depending on the presence or absence of tRNA in addition to other substrates. For example, in the absence of tRNA^Ser, Glu258 forms a hydrogen bond with the N6 position of the ATP molecule and Arg271 interacts with the γ-phosphate. Furthermore, unlike in the substrate-free enzyme, the flexible motif 2 loop (residues 258–265) is completely ordered (except for some side chains) in the presence of ATP, with Arg157 playing a key role in stabilizing the loop. When cognate tRNA is bound, the motif 2 loop becomes ordered in a quite different conformation (again stabilized by Arg157, but in a different way) that allows it to interact with the major groove of the tRNA acceptor stem (Cusack *et al.*, unpublished data). In so rearranging, both Glu258 and Arg271 break contact with the ATP molecule and move to interact with bases at the 3′-end of the tRNA. This may be a mechanism whereby tRNA binding assists the release of the AMP and pyrophosphate products and thus enhances turnover. Alternatively, it could be a mechanism to ensure that the correct positioning of the tRNA 3′-end occurs only after adenylate formation. In any case, it illustrates the dynamic nature of the active-site interactions during the course of the aminoacylation reaction and also suggests that one should be cautious in assigning definitive roles to certain residues until structural information is available for all steps of the reaction.

Mechanism for Ser-AMP and Ap₄A formation by SerRS

Early studies on the amino acid activation reaction catalyzed by aaRSs strongly suggested that the aminoacyl-adenylate is formed by an in-line displacement mechanism [25,26] (for review, see [5]). The steps in this S_N2-type reaction are nucleophilic attack by the carboxyl group of the amino acid on the ATP α-phosphate, formation of a pentavalent transition state, inversion of stereochemical configuration at the α-phosphate and release of the pyrophosphate leaving group. Such reactions are favoured in several ways: firstly, by arranging the initial conformation of the reactants so that they resemble the geometry of the transition state as closely as possible; secondly, by increasing the partial positive charge on the phosphorus atom; thirdly, by stabilization of the trigonal bipyramidal geometry of the transition state; and, finally, by neutralization of the pyrophosphate leaving group. We will now discuss how SerRS contributes to enhancing the rate of the serine activation reaction.

Fig. 8a shows the superposition of the enzyme-bound ATP and Ser-AMP conformations together with the divalent cation site. Although we do not have a structure with serine alone in the active site, we assume that its initial conformation is very close to that of the serine moiety observed in the Ser-AMP complex. Fig. 8a then clearly shows that the enzyme juxtaposes the reactants in the exact geometric relationship required for the in-line displacement reaction to occur. The specific enzyme–substrate interactions by which this is done were described above (Fig. 3 and [11]). The inversion of the α-phosphate configuration during the reaction is also visible in Fig. 8a; this implies an adjustment of the position of the O5' atom (as discussed in [5]). The probable steps in the activation reaction are shown schematically in Fig. 9. With reference to the other factors mentioned above that contribute to rate enhancement, we can emphasize the three-fold importance of the principal divalent cation site (together with the conserved Arg386 residue from motif 3) in stabilizing the bent conformation of the ATP molecule, helping to polarize the target α-phosphate group for nucleophilic attack, and possibly in directly stabilizing the transition-state geometry. In addition, the divalent cations bridging the β- and γ-phosphates neutralize the leaving pyrophosphate group.

The conserved residue Arg256 (from motif 2) also plays a key role in positioning the α-phosphate and in charge neutralization. Because Arg256 can form hydrogen bonds simultaneously with both the α-phosphate and the amino acid carboxyl group, it is probably important in assembling the reactants and is likely to be responsible for the observed synergistic binding of serine and ATP in the case of SerRS (K Larsen, personal communication). Arg256 is probably also involved in direct stabilization of the transition state. Also of note is residue Lys277 of motif 2, which is conserved in all known SerRS sequences, but not in other class II synthetases. Lys277 is normally observed with its $-C\epsilon-N_zH_3^+$ group bent to
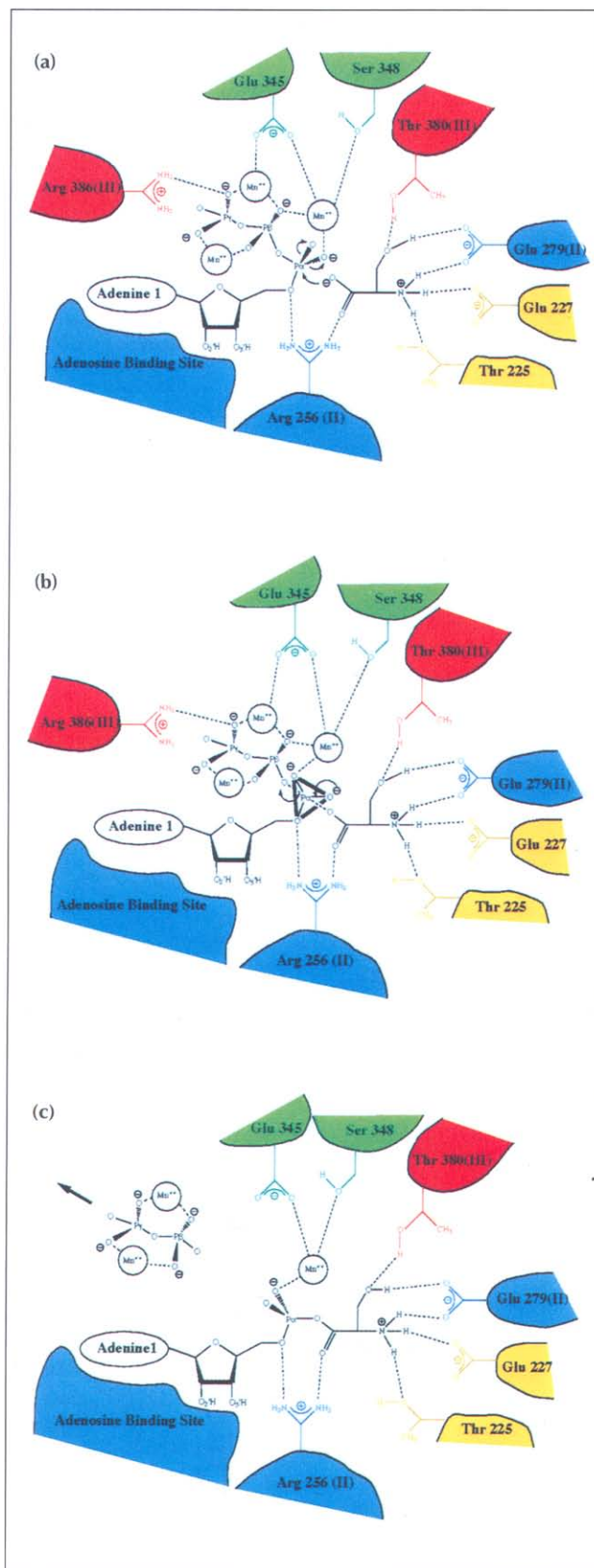

Fig. 9. Schematic diagram for the proposed mechanism of serine activation by SerRS from *T. thermophilus*. Colour coding corresponds to Fig. 3. **(a)** Nucleophilic attack by the serine carboxyl group on enzyme-bound ATP. **(b)** Pentavalent transition state. **(c)** Seryl-adenylate formation and pyrophosphate release.
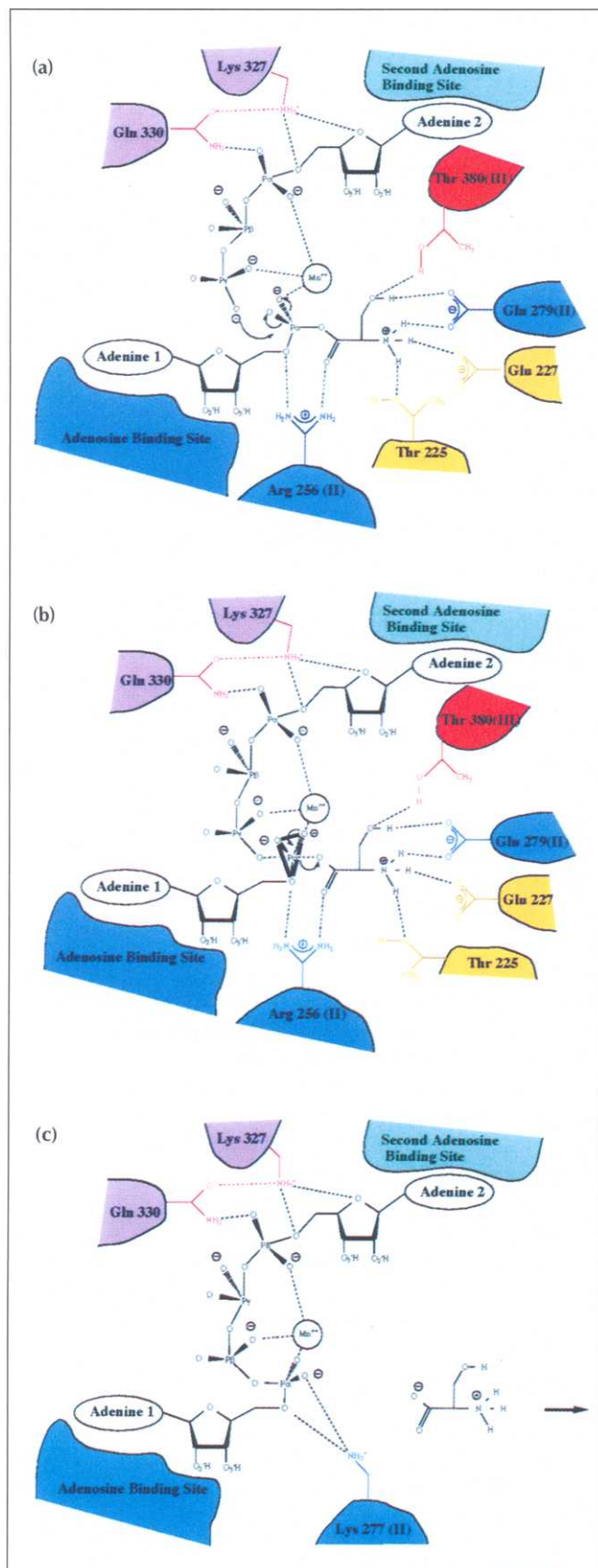
Fig. 10. Schematic diagram for the proposed mechanism of Ap$_4$A synthesis by SerRS. Colour coding corresponds to Fig. 7. **(a)** Nucleophilic attack by a second ATP molecule on enzyme-bound seryl-adenylate. **(b)** Pentavalent transition state. **(c)** Ap$_4$A formation and serine release.

interact with Glu279 (which in turn interacts with the α-amino group of the substrate serine). However, a small movement to straighten the lysine side chain would bring its α-amino group within hydrogen-bonding distance of an α-phosphate oxygen. We hypothesize that this interaction may be important during the initial and transition states of the reaction wherein the two negative charges from the α-phosphate and carboxyl group of the amino acid need to be neutralized and brought close together. This theory might be testable by determining the structure of a complex with a doubly negatively charged transition-state analogue such as pentavalent vanadate.

Basically the same arguments can be now applied to the mechanism of formation of Ap$_4$A by SerRS. As shown in Fig. 8b, superposition of the enzyme-bound Ser-AMP and Ap$_4$A conformations allows one to propose that a second ATP molecule would bind with its γ-phosphate in almost the same position as the β-phosphate of the initial ATP. The steps of the subsequent reaction are shown schematically in Fig. 10. Nucleophilic attack on the α-phosphate of the pre-formed and enzyme-bound Ser-AMP is followed by an in-line displacement reaction in the reverse direction to that illustrated in Fig. 9, giving the products, serine and Ap$_4$A. Once again, the divalent cation at the principal site is a key player in this reaction. The cation permits the initial binding of the second ATP in the correct conformation by binding to its α- and γ-phosphates. Also, by bridging the ATP γ-phosphate and the adenylate phosphate, it could directly stabilize the transition state. Arg256 plays a similar role in this reaction to the one described previously for the mechanism of serine activation. Several studies have shown that Ap$_4$A synthesis is in equilibrium with the reverse reaction in which the synthetase uses Ap$_4$A instead of ATP as a substrate for reaction with serine to form ATP and Ser-AMP. The mechanism for this reverse reaction is easily explained by the current results.

It is interesting to ask whether the second adenosine-binding site is also that of the 3′-terminal adenine 76 (Ade76) of the tRNA. Were this the case, one might expect that upon binding of the second ATP there would be competition between the formation of Ap$_4$A and seryl-3′-ATP, the latter being formed by a reaction analogous to the charging of tRNA (note that class II synthetases charge the 3′-hydroxyl group of the ribose). There is no evidence of seryl-3′-ATP formation in the case of SerRS although the formation of the analogous compound, tryptophanyl-2′-ATP, for the class I synthetase TrpRS, has been reported [27]. Furthermore, the superposition of Ap$_4$A and Ser-AMP (Fig. 8b) shows that the 3′-hydroxyl group of the ribose is directed away from the site of reaction and is ~6 Å from the nearest α-phosphate oxygen position and 5.3 Å from the carboxyl carbon of the serine. On the other hand, for both the GlnRS and AspRS systems, it has been proposed that for aminoacylation to occur, the relevant ribose hydroxyl should be within hydrogen-bonding

distance of an α-phosphate oxygen [5,12]. It is therefore probable that the constraints imposed on binding of ATP or Ap_4A by the oligophosphates forces the adenosine moiety further away from the catalytic site than in the case of the Ade76 of the tRNA. This is consistent with the distortion in the N-terminal region of the protein upon Ap_4A binding referred to above. Following this argument further, the observed specificity of the second nucleoside site for adenosine may be of limited significance because extensive studies on *E. coli* SerRS have shown that it can synthesize Ap_4N equally readily from Ser-AMP and any NTP where N is A, G, U or C (K Larsen, personal communication). Despite these arguments, the observation that the divalent cation remains in place after serine activation and interacts with the δ-phosphate of Ap_4A (which, if not identically placed, must be close to the phosphate of the tRNA Ade76), raises the possibility that the cation is also required for the correct positioning of the 3′-terminal adenosine of the tRNA and hence might also be essential for the transfer reaction.

Comparison with other systems

The reaction mechanism and arrangement of the reactants described here for serine activation is rather similar to that proposed for aspartate activation by AspRS, another class II synthetase [12,17]. However, a key difference, as detailed above, is that in SerRS a major divalent cation site for either magnesium or manganese is found bridging the ATP α- and β-phosphates (α-β site) rather than just the β- and γ-phosphates (β-γ site) as is apparently the case in AspRS. Furthermore, we have shown that the α-β site in SerRS is also occupied in the enzyme complexes with Ser-AMP and Ap_4A. These observations suggest that the divalent cation plays a much more direct catalytic role in both serine activation and Ap_4A synthesis than was previously believed. This is because one of the ligands of the ion is the key reaction centre, the α-phosphate, where bonds are broken or formed. Also, for Ap_4A synthesis, it is difficult to imagine how a divalent cation in an alternative site (for example, the β-γ site) would be much use in stabilizing the binding of the second ATP or the product Ap_4A. The occurrence of a divalent cation site in the α-β site, capable of stabilizing the pentavalent transition state, is reminiscent of the proposed key role of divalent metals ions in several phosphoryl-transfer reactions and ribozymes [28–30].

However, the fact that one of the protein ligands (Ser348) of the divalent cation is not conserved in other synthetases raises the possibility that the location of the cation site is not identical in all class II synthetases. In the ternary complex of a class I aaRS, GlnRS–tRNA^{Gln}–ATP, a putative magnesium site has also been identified bridging the β- and γ-phosphates of the ATP molecule [5]. Further structural studies on other synthetase systems of both classes at sufficient resolution will hopefully clarify the generality of the divalent cation binding site between the α- and β-phosphates of ATP as observed in SerRS from *T. thermophilus*.

Biological implications

As the first step in the overall tRNA aminoacylation reaction, aminoacyl-tRNA synthetases activate the substrate amino acid with Mg·ATP to form the aminoacyl-adenylate. Most synthetases can also synthesize the dinucleotide diadenosine tetraphosphate (Ap_4A) by attack of a second ATP molecule on the enzyme-bound adenylate. Although the cellular role of Ap_4A is uncertain, it has been implicated in a wide variety of metabolic processes [21]. An understanding of the structural basis for the substrate specificity and the catalytic mechanism of aminoacyl-tRNA synthetases is of considerable general interest because of the fundamental importance of these enzymes to protein biosynthesis in all living cells.

As it is now known that the 20 aminoacyl-tRNA synthetases fall into two very distinct structural classes (I and II), it is of particular interest to elucidate the common and distinct features of synthetase activity between the two classes and between the 10 members of each class. By means of a series of crystal structures of binary and ternary complexes of *Thermus thermophilus* seryl-tRNA synthetase with ATP, manganese or magnesium ions, seryl-adenylate (and analogues) and Ap_4A, we have gained a detailed and coherent structure-based description of the mechanism and specificity of Ap_4A and seryl-adenylate formation by a class II aminoacyl-tRNA synthetase.

Both Ap_4A and seryl-adenylate are formed by in-line displacement reactions involving a pentavalent transition state at the α-phosphate position of the adenylate, but with the displacement going in opposite directions. The enzyme active site is designed to hold the substrates (serine and two distinct ATP molecules) in the correct geometrical relationship for these reactions to occur with minimal requirement for additional distortion. A divalent cation, either magnesium or manganese, has been located between the α- and β-phosphates of the first ATP molecule and is also present at the same site in the seryl-adenylate and Ap_4A complexes. Both this cation and the conserved arginine from motif 2 (one of the three conserved motifs of class II aminoacyl-tRNA synthetases), can interact with both the α-phosphate and the carboxyl group of the serine and are suitably positioned to stabilize the pentavalent transition state.

This is the first evidence that the divalent cation plays this role in activation by an aminoacyl-tRNA synthetase, although it is possible that the exact position of the ion-binding site varies amongst the different class II synthetases.

Materials and methods

Crystal preparation and data collection

For the crystallographic experiments described in this paper, crystals were grown in 34% saturated ammonium sulphate (AS), 2% 2-methyl-2,4-pentanediol (MPD, Fluka) [31] from recombinant *T. thermophilus* SerRS expressed in *E. coli* (Tukalo *et al.*, unpublished data). This enzyme has a modified C terminus with sequence P_{419}WHRIRILEDERAS-COOH, instead of the wild-type sequence P_{419}CG-COOH. The C-terminal extension has no effect on enzymatic activity and in electron-density maps is disordered beyond Trp420. Crystals are of space group $P2_1$ with one synthetase dimer per asymmetric unit [14]. Manganese(II)sulphate monohydrate ($MnSO_4 \cdot H_2O$), L-serine, adenosine 5'-triphosphate (ATP) and P^1,P^4-di(adenosine-5')tetraphosphate (Ap₄A) were purchased from MERCK, SERVA, SIGMA and SIGMA, respectively. All solutions for crystal soaking were freshly prepared before use. The crystals, stabilized in 40% AS and 2% MPD, were soaked for relatively short periods to avoid ligand hydrolysis (between 30 min and 2 h) before mounting in glass capillaries for immediate exposure to the X-ray beam. A single crystal was used for each experiment and data were collected at room temperature. Diffraction data for structures 1 (ATP) and 2 (Ser–AMP) were collected on the synchrotron beamline W32 at LURE (Laboratoire Universitaire pour le Rayonnement Electromagnétique, Orsay, France) [32], using a 30 cm Mar Research image-plate scanner. Structure 3 (Ap₄A) diffraction patterns were recorded on a similar detector on the High Brilliance Beamline BL4 [33] at the ESRF (European Synchrotron Radiation Facility, Grenoble, France). All images were integrated using the MOSFLM package [34]. Further data analysis was performed with the CCP4 programme package [35].

Model refinement

The 2.5 Å refined model of SerRS from *T. thermophilus* [14] was used as a starting model. For each set of data, the following general procedure was applied. A rigid-body refinement was initially performed using X-PLOR [23] to account for the slight changes in unit cell (Table 1). σ_a-weighted electron-density maps were then calculated using the CCP4 program SIGMAA [36]. Substrate models were built into the difference density and other modifications made manually with the graphics program FRODO [37]. The complex model was then subjected to a new cycle of rigid-body refinement followed by several hundred steps of Powell energy-minimization and a constrained isotropic individual temperature-factor refinement. The parameter and topology files based on the work of Engh and Huber [38] were used; new topology files were established for ATP, Ap₄A and Ser–AMP with associated parameters adjusted to be compatible with those of the protein. The refined model was then checked visually and further manual reorientation done if necessary. Putative water molecules were then added at positive peaks over 4σ in the $mF_o - F_c$ map that were at suitable distance from the enzyme (between 2.5 Å and 3.5 Å) and in a favourable environment. Results of the refinement are given in Table 1. The manganese ion locations were determined using anomalous difference Fourier maps with coefficients $(\Delta F_{ano})_{ATP \cdot Mn} \, e^{i(\phi_c^{nat}-\pi/2)}$. This technique, originally proposed by Pepinsky and Okaya [39], allowed us to identify unambiguously the extra densities around the phosphate groups as the divalent cation present in the soaking solution rather than as water molecules. At a wavelength of 0.9 Å, manganese has an f'' of about 1.1 electrons.

Structures described in this paper will be deposited in the Brookhaven Protein Data Bank.

Figs 2, 5, 6 were drawn using MINIMAGE [40], MOLSCRIPT [41] and RASTER3D [42]. Fig. 8a, b was drawn using RIBBONS [43].

References

1. Eriani, G., Delarue, M., Poch, O., Gangloff, J. & Moras, D. (1990). Partition of tRNA synthetases into two classes based on mutually exclusive sets of sequence motifs. *Nature* **347**, 203–206.
2. Cusack, S., Berthet-Colominas, C., Härtlein, M., Nassar, N. & 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.
3. Moras, D. (1992). Structural and functional relationships between aminoacyl-tRNA synthetases. *Trends Biochem. Sci.* **17**, 159–164.
4. Mechulam, Y., Dardel, F., Le Corre, D., Blanquet, S. & Fayat, G. (1991). Lysine 355, part of the KMSKS signature sequence, plays a crucial role in the amino acid activation catalysed by the methionyl-tRNA synthetase from *Escherichia coli. J. Mol. Biol.* **217**, 465–475.
5. Perona, J.J., Rould, M.A. & Steitz, T.A. (1993). Structural basis for transfer RNA aminoacylation by *Escherichia coli* glutaminyl-tRNA synthetase. *Biochemistry* **32**, 8758–8771.
6. Delarue, M. & Moras, D. (1993). The aminoacyl-tRNA synthetase family: modules at work. *Bioessays* **15**, 675–687.
7. Artymiuk, P.J., Rice, D.W., Poirrette, A.R. & Willet, P. (1994). A tale of two synthetases. *Nat. Struct. Biol.* **1**, 758–760.
8. Cusack, S., Härtlein, M. & Leberman, R. (1991). Sequence, structure and evolutionary relationships between class 2 aminoacyl-tRNA synthetases. *Nucleic Acids Res.* **19**, 3489–3498.
9. Cavarelli, J., Rees, B., Ruff, M., Thierry, J.-C. & Moras, D. (1993). Yeast tRNA^Asp recognition by its class II aminoacyl-tRNA synthetase. *Nature* **362**, 181–184.
10. Mosyak, L. & Safro, M. (1993). Phenylalanyl-tRNA synthetase from *Thermus thermophilus* has four antiparallel folds of which only two are catalytically functional. *Biochimie* **75**, 1091–1098.
11. Belrhali, H., et al., & Cusack, S. (1994). Crystal structures at 2.5 Å resolution of seryl-tRNA synthetase complexed with two analogs of seryl adenylate. *Science* **263**, 1432–1436.
12. Cavarelli, J., et al., & Moras, D. (1994). The active site of yeast aspartyl-tRNA synthetase: structural and functional aspects of the aminoacylation reaction. *EMBO J.* **13**, 327–337.
13. Fersht, A.R. (1987). Dissection of the structure and activity of the tyrosyl-tRNA synthetase by site-directed mutagenesis. *Biochemistry* **26**, 8031–8037.
14. Fujinaga, M., Berthet-Colominas, C., Yaremchuck, A.D., Tukalo, M.A. & 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.
15. Biou, V., Yaremchuk, A., Tukalo, M. & Cusack, S. (1994). The 2.9 Å crystal structure of *T. thermophilus* seryl-tRNA synthetase complexed with tRNA^Ser. *Science* **263**, 1404–1410.
16. Ruff, M., et al., & Moras, D. (1991). Class II aminoacyl transfer RNA synthetases: crystal structure of yeast aspartyl-tRNA synthetase complexed with tRNA^Asp. *Science* **252**, 1682–1689.
17. Poterszman, A., Delarue, M., Thierry, J.-C. & Moras, D. (1994). Synthesis and recognition of aspartyl-adenylate by *Thermus thermophilus* aspartyl-tRNA synthetase. *J. Mol. Biol.* **244**, 158–167.
18. Randerath, K., Janeway, C.M., Stephenson, M.L. & Zamecnik, P.C. (1966). Isolation and characterization of dinucleotide tetra- and triphosphates formed in the presence of lysyl-tRNA synthetase. *Biochem. Biophys. Res. Commun.* **24**, 98–105.
19. Goerlich, O., Foeckler, R. & Holler, E. (1982). Mechanism of synthesis of adenosine(5')tetraphospho(5')adenosine (AppppA) by aminoacyl-tRNA synthetases. *Eur. J. Biochem.* **126**, 135–142.
20. Brevet, A., Plateau, P., Çirakoglu, B., Pailliez, J.-P. & Blanquet, S. (1982). Zinc-dependent synthesis of 5',5'-diadenosine tetraphosphate by sheep liver lysyl- and phenylalanyl-tRNA synthetases. *J. Biol. Chem.* **257**, 14613–14615.

21. Plateau, P. & Blanquet, S. (1994). Dinucleoside oligophosphates in micro-organisms. *Adv. Microb. Physiol.* **36**, 81–109.
22. Zamecnik, P. (1983). Diadenosine 5',5'''-P¹,P⁴-tetraphosphate (Ap₄A): its role in cellular metabolism. *Anal. Biochem.* **134**, 1–10.
23. Brünger, A.T. (1992) *X-PLOR Version 3.1.* Yale University. New Haven, CT.
24. Holler, E., Holmquist, B., Vallee, B.L., Taneja, K. & Zamecnik, P. (1983). Circular dichroism and ordered structure of bisnucleoside oligophosphates and their Zn²⁺ and Mg²⁺ complexes. *Biochemistry* **22**, 4924–4933.
25. Langdon, S.P. & Lowe G. (1979). The stereochemical course of amino acid activation by methionyl- and tyrosyl-tRNA synthetase. *Nature* **281**, 320–321.
26. Harnett, S.P., Lowe, G. & Tansley, G. (1985). Mechanism of activation of phenylalanine and synthesis of P¹,P⁴-bis(5'-adenosyl) tetraphosphate by yeast phenylalanyl-tRNA synthetase. *Biochemistry* **24**, 2908–2915.
27. Joseph, D.R. & Muench, K.H. (1971). Tryptophanyl transfer ribonucleic acid synthetase of *Escherichia coli.* I. Purification of the enzyme and of tryptophan transfer ribonucleic acid. *J. Biol. Chem.* **246**, 7602–7609.
28. Steitz, T.A. (1993). DNA- and RNA-dependent DNA polymerases. *Curr. Opin. Struct. Biol.* **3**, 31–3.
29. Steitz, T.A. & Steitz, J.A. (1993). A general two-metal-ion mechanism for catalytic RNA. *Proc. Natl. Acad. Sci. USA* **90**, 6498–6502.
30. Yarus, M. (1993). How many catalytic RNAs? Ions and the Cheshire cat conjecture. *FASEB J.* **7**, 31–39.
31. Garber, M.B., Yaremchuk, A.D., Tukalo, M.A., Egorova, S.P., Berthet-Colominas, C. & Leberman, R. (1990). Crystals of seryl-tRNA synthetase from *Thermus thermophilus.* Preliminary crystallographic data. *J. Mol. Biol.* **213**, 631–632.
32. Fourme, R., *et al.,* & 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.
33. Bösecke, P. (1992). The high-flux beamline at the ESRF. *Rev. Sci. Instrum.* **63**, 438–441.
34. Leslie, A.G.W. (1992). *Joint CCP4 and ESF-EACBM Newsletter on Protein Crystallography, No. 26.* SERC Daresbury Laboratory, Warrington, UK.
35. Collaborative Computational Project, Number 4. (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D* **50**, 760–763.
36. Read, R.J. (1986). Improved Fourier coefficients for maps using phases from partial structures with errors. *Acta Crystallogr. A* **42**, 140–149.
37. Jones, T.A. (1978). A graphics model building and refinement system for macromolecules. *J. Appl. Crystallogr.* **11**, 268–272.
38. Engh, R.A. & Huber, R. (1991). Accurate bond and angle parameters for X-ray protein structure refinement. *Acta Crystallogr. A* **47**, 392–400.
39. Pepinsky, R. & Okaya, Y. (1956). Determination of crystal structures by means of anomalously scattered X-rays. *Proc. Natl. Acad. Sci. USA* **42**, 286–292.
40. Arnez, J.G. (1994). MINIMAGE: a program for plotting electron density maps. *J. Appl. Crystallogr.* **27**, 644–653.
41. Kraulis, P.J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* **24**, 946–950.
42. Merritt, E.A. & Murphy, M.E.P. (1994). RASTER3D Version 2.0. A program for photorealistic molecular graphics. *Acta Crystallogr. D* **50**, 869–873.
43. Carson, M. (1992). *RIBBONS, Version 2.0.* University of Alabama at Birmingham, Alabama.