Active Site of Lysyl-tRNA Synthetase: Structural Studies of the Adenylation Reaction†

Gianluigi Desogus, Flavia Todone, Peter Brick, and Silvia Onesti*

Biophysics Section, Blackett Laboratory, Imperial College of Science, Technology and Medicine, London SW7 2BW, U.K.

Received March 24, 2000; Revised Manuscript Received May 18, 2000

ABSTRACT: Aminoacyl-tRNA synthetases play a key role in protein biosynthesis by catalyzing the specific aminoacylation of tRNA. The energy required for the formation of the ester bond between the amino acid carboxylate group and the tRNA acceptor stem is supplied by coupling the reaction to the hydrolysis of ATP. Lysyl-tRNA synthetase from *Escherichia coli* belongs to the family of class II synthetases and carries out a two-step reaction, in which lysine is activated by being attached to the α-phosphate of AMP before being transferred to the cognate tRNA. Crystals of the thermo-inducible *E. coli* lysyl-tRNA synthetase LysU which diffract to 2.1 Å resolution have been used to determine crystal structures of the enzyme in the presence of lysine, the lysyl−adenylate intermediate, and the nonhydrolyzable ATP analogue AMP-PCP. Additional data have been obtained from crystals soaked in a solution containing ATP and Mn^{2+}. The refined crystal structures give "snapshots" of the active site corresponding to key steps in the aminoacylation reaction and provide the structural framework for understanding the mechanism of lysine activation. The active site of LysU is shaped to position the substrates for the nucleophilic attack of the lysine carboxylate on the ATP α-phosphate. No residues are directly involved in catalysis, but a number of highly conserved amino acids and three metal ions coordinate the substrates and stabilize the pentavalent transition state. A loop close to the catalytic pocket, disordered in the lysine-bound structure, becomes ordered upon adenine binding.

Aminoacyl-tRNA synthetases are ubiquitous enzymes that are responsible for the correct charging of transfer RNA (tRNA) with the cognate amino acid. The energy required for the formation of the ester bond between the amino acid and the tRNA acceptor stem is supplied by coupling the reaction to the hydrolysis of ATP to AMP and pyrophosphate. Despite the fact that they differ widely in size, sequence, and oligomeric state, all aminoacyl-tRNA synthetases carry out the same two-step reaction. In the first step, the enzyme binds ATP and the amino acid, and catalyzes the formation of an aminoacyl−adenylate intermediate, in which the carboxylate group of the amino acid is activated by becoming attached to the phosphoryl moiety of AMP. This intermediate is unstable in solution and does not dissociate from the enzyme. In the second step of the reaction, the amino acid moiety is transferred to the 2′- or 3′-ribose hydroxyl group of the terminal adenosine of the tRNA molecule.

Sequence homology and structural studies have revealed that it is possible to divide aminoacyl-tRNA synthetases into two classes, where each class shares sequence motifs and a similar topology for the catalytic domain (*1*). Class I enzymes contain two conserved sequences (HIGH and KMSKS, shown to be involved in ATP binding) and have a catalytic domain which resembles the canonical dinucleotide-binding fold found in many dehydrogenases (*2*−*4*). In contrast, class II enzymes are characterized by three conserved sequence motifs, which have been shown to be involved in dimerization and amino acid activation, and have a catalytic domain built around a large antiparallel β-sheet flanked by long helices (*5*−*7*). This classification generally holds throughout the various kingdoms.

The formation of an enzyme-bound adenylate intermediate is a common mechanism used by both prokaryotic and eukaryotic organisms to activate substrates, especially those having a carboxylate group which can form an acyl−adenylate intermediate. In particular, it is the mechanism used by enzymes which activate amino acid substrates for peptide synthesis. Formation of an acyl−adenylate intermediate is a common step not only to the two unrelated classes of aminoacyl-tRNA synthetases but also to peptide synthetases, which carry out nonribosomal protein synthesis in fungi and bacteria (*8*).

In eukaryotes and most prokaryotes, lysyl-tRNA synthetase (LysRS)[1] belongs to the class II family of aminoacyl-tRNA synthetases, but in some archaea and bacteria, it has recently been shown that the lysyl-tRNA charging activity is associated with a protein containing the conserved motifs characteristic of class I enzymes (*9*, *10*). This is the only example of an aminoacyl-tRNA synthetase activity catalyzed by both classes of enzymes. Within the class II family, a number of

† This work was supported by a Wellcome Trust Grant (Grant 050370). G.D. is a recipient of a Marie Curie Research Training Fellowship. Synchrotron data collection at Hamburg and Trieste was supported by a EU/TMR LSF grant.

* Corresponding author. E-mail: s.onesti@ic.ac.uk. Telephone: 020-75947647. Fax: 020-75890191.

[1] Abbreviations: LysRS, lysyl-tRNA synthetase; AMP-PCP, β,γ-methyleneadenosine 5′-triphosphate; PEG, polyethylene glycol.

10.1021/bi0006722 CCC: $19.00
Published on Web 06/28/2000

subfamilies (called IIa, IIb, and IIc) with more extensive sequence and structural homology can be identified. LysRS from *Escherichia coli* belongs to the IIb subfamily together with AspRS and AsnRS, and shares with these enzymes a conserved topology of the anticodon-binding domain (*11*, *12*).

In *E. coli*, there are two isoforms of LysRS. LysS is expressed under normal growth conditions, while LysU is the product of a normally silent gene, which is overexpressed under certain physiological conditions, such as high temperature, anaerobiosis, low external pH, or the presence of leucine (*13*, *14*). The sequences of the two proteins are 88% identical, and the proteins have similar enzymatic properties. They only appear to differ significantly in their affinity for lysine, with the dissociation constant of the LysU−lysine complex being 8-fold lower than that of the LysS−lysine complex (*15*).

In a previous paper, we reported the crystal structure of *E. coli* LysU in a complex with the amino acid substrate at 2.8 Å (*16*). This structure revealed the overall conformation of the molecule and details of the active site that explain the specificity of the enzyme for the lysine substrate. Subsequently, crystal structures of *Thermus thermophilus* LysRS complexed with the modified *E. coli* $tRNA^{Lys}$ and with the *T. thermophilus* $tRNA^{Lys}$ transcript have been published, together with a low-resolution (3.8 Å) structure of LysRS in a complex with tRNA and a nonhydrolyzable lysyl−adenylate analogue (*17*). In all three structures, only the anticodon of the tRNA molecule is well-ordered.

Here we report high-resolution structures of four distinct complexes of *E. coli* LysU, in the presence of lysine, an ATP analogue, the lysyl−adenylate intermediate, and the unreacted ATP in the presence of Mn^{2+} ions. The substrate ATP reacts with the lysine bound to the active site in the crystal to give the lysyl−adenylate intermediate. In the presence of a nonhydrolyzable ATP analogue, the first step of the reaction does not occur and the resulting structure reveals the conformation of bound ATP prior to catalysis. The positions of the metal binding sites implicated in catalysis have been determined by using $MnCl_2$ instead of $MgCl_2$. The four new crystal structures afford views of the active site at different key stages of the aminoacylation reaction and reveal the structural basis of LysRS specificity and the enzyme mechanism. No major conformational change in the enzyme is observed, but a loop close to the catalytic pocket, disordered in the lysine-bound structure, becomes ordered upon ATP binding.

EXPERIMENTAL PROCEDURES

Crystallization. Recombinant LysU was overexpressed in *E. coli* cells and purified as previously described (*18*). The protein was concentrated to 12 mg/mL in 20 mM Hepes buffer (pH 7.5), 5 mM lysine, and 2 mM β-mercaptoethanol. Initially, orthorhombic crystals were obtained by vapor diffusion in hanging drops from a solution containing 20% PEG 2000, 0.5 M LiCl, and 0.1 M Pipes buffer (pH 6.8). These crystals belong to space group $C222_1$ (cell dimensions of $a = 144.3$ Å, $b = 257.8$ Å, and $c = 182.1$ Å), contain three monomers in the asymmetric unit arranged around a pseudo-3_1 axis, and diffract to 2.1 Å resolution but are very sensitive to radiation damage. Attempts to stabilize the crystals in a cryoprotectant solution, to collect data at low temperatures, were unsuccessful due to crystal cracking. The crystallization protocol was therefore modified to include 17% glycerol as a cryoprotectant, to avoid the handling damage and the osmotic shock upon harvesting. The crystals grown under these conditions have a morphology identical to that of the original orthorhombic crystals, but their diffraction pattern is consistent with space group $P6_122$ with the following cell dimensions: $a = b = 143.1$ Å and $c = 176.1$ Å. These hexagonal crystals contain a single monomer in the asymmetric unit, can be readily frozen, and diffract to 2.1 Å. Since the crystals cracked easily upon handling, substrate binding studies were performed by slowly adding small amounts of a substrate solution directly to the crystallization drop, rather than transferring the crystals into a new harvesting solution.

Data Collection. Diffraction data to 2.1 Å resolution were recorded from a hexagonal crystal of the enzyme−lysine complex flash-frozen to 100 K in a stream of cold nitrogen produced by a Cryostream Cooler (Oxford Cryosystems). The data were collected at DESY (Hamburg, Germany) on the wiggler beamline BW7B (Table 1), using a MarResearch imaging plate system. Diffracted intensities were evaluated and integrated using the program DENZO (*19*).

Data to 2.4 Å resolution were obtained from crystals in the presence of the substrate Mg^{2+} and ATP on the X11 beamline (DESY). A small amount of solution containing ATP and $MgCl_2$ was added to the crystallization drop to yield a final concentration of about 5 mM, and the crystal was frozen directly from the drop. A similar procedure was used to collect data in the presence of the nonhydrolyzable ATP analogue AMP-PCP (β,γ-methyleneadenosine 5′-triphosphate). A solution containing AMP-PCP and $MgCl_2$ was added directly to the drop to give an inhibitor concentration of approximately 5 mM. Diffraction data to 2.4 Å resolution were collected at Elettra (Trieste, Italy), on an 18 cm diameter MarResearch imaging plate system. In a similar manner for data set 4, a solution containing ATP and $MnCl_2$ was added to the drop and a full data set collected using the Synchrotron Radiation Source at Daresbury (Warrington, U.K.). The images were processed and integrated using a version of MOSFLM modified for image plate data (A. G. W. Leslie, personal communication). A summary of the data collection statistics is given in Table 1. All data handling leading to electron density maps was done using the CCP4 program package (*20*).

Structure Solution and Refinement. Despite the apparently dramatic change in space group symmetry (from $C222_1$ to $P6_122$), the presence of glycerol caused only a minor rearrangement of the crystal lattice. A comparison of the cell parameters revealed the relationship between the two lattices, and allowed the original LysU structure from the orthorhombic crystal form to be placed in the hexagonal asymmetric unit by simply translating the model in the *x*-direction by $^1/_2$ the cell dimension.

This initial model was refined against the 2.1 Å data with the program X-PLOR (*21*) using the stereochemical parameters of Engh and Huber (*22*). Low-resolution data to 20 Å were included, and a bulk solvent correction was applied throughout the refinement procedure. A random sample containing 5% of the data was excluded from the refinement, and the agreement between calculated and observed structure

Table 1: Data Collection and Refinement Statistics

		1	2	3	4
soaking conditions	amino acid	5 mM lysine	5 mM lysine	5 mM lysine	5 mM lysine
	ATP		5 mM AMP-PCP	5 mM ATP	5 mM ATP
	metal ion		5 mM $MgCl_2$	5 mM $MgCl_2$	5 mM $MnCl_2$
soaking time		cocrystallized	4 h	12 h	1 h
		Data Collection			
Synchrotron		DESY Hamburg	ELETTRA Trieste	DESY Hamburg	SRS Daresbury
beamline		BW7B	5.2R	X11	9.6
wavelength (Å)		0.86	1.32	0.93	0.87
resolution (Å)		25.0−2.1	10.0−2.4	15.0−2.4	25.0−2.35
no. of measurements		977338	120728	334798	217414
no. of independent reflections		60434	33781	41908	45104
completeness (outer shell) (%)		99.6 (99.6)	84.7 (65.1)	98.2 (94.1)	99.4 (99.4)
R_{merge}[a] (outer shell) (%)		8.5 (26.1)	6.5 (9.9)	8.6 (20.4)	7.2 (18.2)
intensity/σ (outer shell)		8.8 (3.7)	8.1 (6.3)	10.5 (4.1)	15.5 (7.3)
		Crystallographic Refinement			
R_{factor}[b] (outer shell) (%)		19.5 (26.2)	17.6 (25.7)	18.6 (26.7)	18.2 (24.9)
R_{free} (outer shell) (%)		23.3 (30.1)	22.8 (32.9)	23.0 (29.2)	21.3 (24.8)
rmsd for bonds (Å)		0.007	0.006	0.007	0.006
		Modeled Ligands			
		lysine	lysine and AMP-PCP	Lys-AMP and PP_i	lysine, ATP, and Mn

[a] $R_{merge} = \Sigma_h\Sigma_i|I_i(h) - \langle I(h)\rangle|/\Sigma_h\Sigma_i I_i(h)$, where $I_i(h)$ is the ith measurement of reflection h and $\langle I(h)\rangle$ is the weighted mean of all measurements of h. [b] $R_{factor} = \Sigma_h|F_{obs} - F_{calc}|/\Sigma_h F_{obs}$, where F_{obs} and F_{calc} are the observed and calculated structure factors, respectively.

factors for those reflections (R_{free}) was used to monitor the course of the refinement (*23*). The refined model of the LysU−lysine complex includes residues 1−153, 161−268, and 270−502.

All the ATP complexes (data sets 2−4) were refined starting from the high-resolution refined model of the LysU−lysine complex. After rigid-body refinement, difference Fourier maps showed clear density for the nonhydrolyzable ATP analogue (data set 2), the lysyl−adenylate intermediate and the pyrophosphate (data set 3), and the unreacted ATP and Mn^{2+} ions (data set 4). Coordinates for the ligands were included and the structures refined to give models with good stereochemistry and low R_{factor} values. In each case, well-defined electron density for the loop between residues 264 and 271 allowed residue 269 to be included in the model. The refinement statistics are summarized in Table 1. Figures 1−3 were prepared using Raster3D (*24*) and Molscript (*25*, *26*). The coordinates of the four crystal structures have been deposited with the Protein Data Bank. The accession codes for data sets 1−4 in Table 1 are 1E1O, 1E1T, 1E22, and 1E24, respectively.

RESULTS AND DISCUSSION

Structure Solution and Modeled Ligands. Crystals of the LysU−lysine complex obtained in the presence of 17% glycerol display a habit similar to that of the original orthorhombic crystals obtained in the absence of glycerol (*16*, *18*), but belong to a related hexagonal crystal form. Data to 2.1 Å resolution were collected from one crystal (Table 1, data set 1), and the structure was determined by using the original orthorhombic model, translated to account for the different origin of the hexagonal cell. This initial model was refined against the 2.1 Å data to give a high-resolution structure of the enzyme−lysine complex.

Diffraction data to 2.4 Å were collected after soaking hexagonal crystals of the LysU−lysine complex in solutions containing lysine and either ATP or the nonhydrolyzable ATP analogue AMP-PCP, in the presence of magnesium or manganese salts (Table 1). These complexes (data sets 2−4) were refined starting from the 2.1 Å resolution model. After rigid-body refinement, difference Fourier maps showed clear density corresponding to the various substrates, substrate analogues, and metal ions.

The crystals require lysine to grow, and a lysine molecule remains bound to the active site even when the crystals are stored in a lysine-free solution. When the nonhydrolyzable ATP analogue AMP-PCP was used instead of ATP (data set 2), the first step of the reaction did not occur and the resulting structure revealed the conformation of the ATP substrate before catalysis. When crystals were soaked overnight into a solution containing ATP and Mg^{2+} (data set 3), the electron density in the active site clearly showed that the first step of the reaction had occurred within the crystal with the formation of the lysyl−adenylate (Lys−AMP) intermediate. The role of divalent metal ions was studied by soaking crystals in a solution containing ATP and $MnCl_2$ (data set 4). The electron density map showed that in this case the ATP had not reacted with the amino acid lysine, probably due to the shorter soaking time that was used.

The corresponding ligands (lysine and AMP-PCP for data set 2, the lysyl−adenylate intermediate and pyrophosphate for data set 3, and lysine and ATP/Mn^{2+} for data set 4) were built in the enzyme active site, and crystallographic refinement was carried out to produce models with low crystallographic R_{factor} values and good stereochemistry. The refinement statistics and modeled ligands are summarized in Table 1.

Overall Structure. The structure of LysU complexed with the lysine substrate has been previously determined at 2.8 Å resolution (*16*). The enzyme is active as a homodimer, with an extended dimer interface. Each polypeptide chain contains 504 amino acid residues and can be divided into three domains: an N-terminal domain, a catalytic domain, and an insertion α-helical domain. The overall structure of

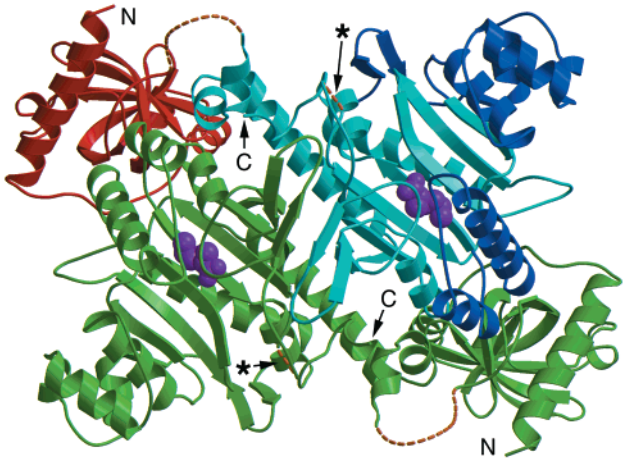

FIGURE 1: Ribbon representation of a dimer of LysU viewed down the molecular 2-fold axis. The structure shown is that of the 2.1 Å complex of LysU with the substrate lysine (shown as a space filling model in violet). One monomer is shown in green, while the other monomer is colored according to the domain structure. The N-terminal anticodon-binding domain is shown in red. Within the larger domain which contains the catalytic site, the conserved core which is common to all class II aminoacyl-tRNA synthetases is shown in light blue, while the secondary structural elements specific to lysyl-tRNA synthetase structure are in dark blue. A dashed orange line has been used to indicate disordered regions of the polypeptide chain, which include the segment of residues 153−160 linking the two domains and residue 269 (indicated by a star) in the motif 2 loop.

a LysU dimer with a lysine molecule bound to each active site is shown in Figure 1.

The N-terminal domain (residues 1−153; shown in red in Figure 1) is involved in binding the tRNA anticodon stem−loop (*17*) and is built around a five-stranded β-barrel with a topology known as OB (oligonucleotide-binding) fold (*27*). A similar fold has been observed in other proteins that bind RNA and single-stranded DNA such as the eukaryotic replication protein A (*28*, *29*) and *Staphylococcus* nuclease (*30*). This module is present in all three members of the IIb subfamily: AspRS (*11*), AsnRS (*12*), and LysRS.

The large catalytic domain (residues 161−504; blue in Figure 1) includes a conserved scaffold shared by all class II synthetases which consists of a large eight-stranded, mostly parallel β-sheet, surrounded by several long helices. The function of the catalytic domain is to bind the amino acid, ATP, and the tRNA acceptor stem and catalyze the two steps of the aminoacylation reaction. The three sequence motifs originally identified in class II synthetases are found in this part of the molecule. Sequence motif 1 (centered around Pro 208) is located at the subunit interface and is not involved in catalysis, but plays a structural role in maintaining the dimer interface. Motifs 2 and 3 (centered around Arg 262 and Arg 480, respectively) contribute conserved residues involved in the binding of ATP and Mg^{2+} ions. A number of secondary structure elements (shown in dark blue in Figure 1) are inserted into the conserved scaffold and are unique to the different class II enzymes.

The model refined against the 2.1 Å data set is not significantly different from that obtained using data to 2.8 Å from the original orthorhombic crystal form (*16*), but provides a more accurate structure for understanding the reaction mechanism at the atomic level. Despite the higher accuracy of the new model, the hydrogen bonding pattern of the lysine substrate does not require revision.

ATP Binding. To examine the position of the ATP substrate before the reaction takes place, diffraction data were collected from crystals grown in the presence of lysine and then soaked in a solution containing the nonhydrolyzable ATP analogue AMP-PCP and $MgCl_2$ (data set 2, Table 1). The corresponding refined model reveals the position and conformation of ATP before the adenylation reaction takes place. What occurs in the crystal is consistent with biochemical data showing that AMP-PCP is an inhibitor of lysyl-tRNA synthetase (unpublished results), despite the fact that the adenylation reaction requires cleavage of the bond between the α- and β-phosphate, and in AMP-PCP the nonhydrolyzable bond is located between the β- and γ-phosphates.

In a subsequent experiment designed to locate the magnesium sites, electron-dense Mn^{2+} ions were used in place of Mg^{2+}. Our biochemical data show that the enzyme activity in the presence of Mn^{2+} ions is comparable to the activity measured in the presence of Mg^{2+} (data not shown). However, when crystals were soaked for 1 h in a solution containing ATP and $MnCl_2$ (data set 4, Table 1), the reaction did not occur and the electron density map could be unambiguously interpreted in terms of an unreacted ATP molecule (Figure 2A). This result is probably due to the short soaking time used in the experiment, but it may also reflect the different behavior of the enzyme under the conditions used for crystallization. A model almost identical to the previous AMP-PCP structure, but including three manganese ions, was refined against these data. Three well-defined electron density peaks (higher than 10 rms) were clearly visible; these peaks are associated with the ATP phosphates, with Mn 1 coordinated by oxygen atoms from the α- and β-phosphate and two carboxylates (Glu 414 and Glu 421), while Mn 2 and Mn 3 bridge the β- and γ-phosphates. Three peaks are present in the AMP-PCP density map, at positions identical to those of the Mn^{2+} sites, and have been interpreted as metal binding sites, although the electron density does not allow us to discriminate between a water molecule and a Mg^{2+} ion.

A difference electron density map calculated for the unreacted ATP is shown in Figure 2A. The substrates of the first step of the aminoacylation reaction (lysine and ATP) are held by a network of hydrogen bond interactions in a deep hydrophilic cleft located on one side of the large antiparallel β-sheet and closed by the surrounding loops of the catalytic domain. Figure 3 shows the hydrogen bonds and electrostatic interactions between the enzyme and the substrates (lysine and ATP) before the first step of the reaction takes place. A schematic diagram of these interaction is shown in Figure 4A. The locations of the Mg^{2+} ions correspond to electron density peaks obtained using diffraction data collected from the crystal soaked in a solution containing $MnCl_2$.

As in all class II aminoacyl-tRNA synthetases, the ATP is found in an unusual bent conformation with the β- and γ-phosphates folded toward the adenine ring into a U-shaped structure (*7*). This places the α-phosphate in an ideal position to attack the carboxylate of the amino acid substrate. We have two different structures (2 and 4) showing the conformation of the bound, unreacted ATP. The corresponding

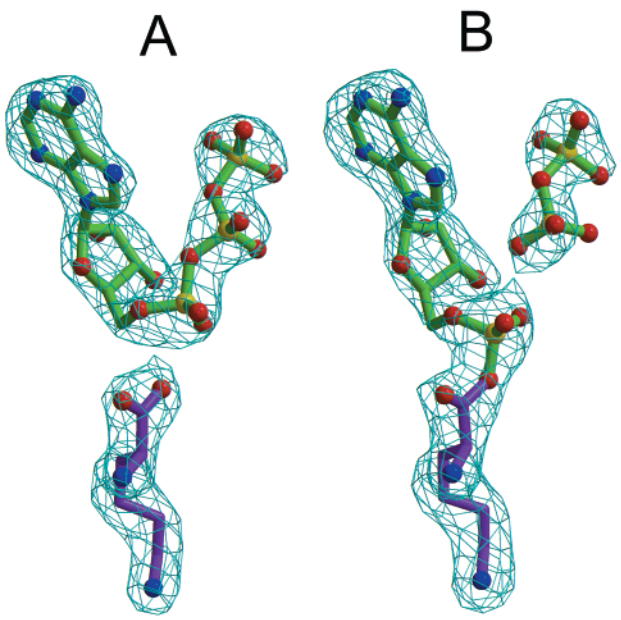

FIGURE 2: (A) Difference electron density map calculated with diffraction data collected from a crystal of the LysU−lysine complex soaked briefly in a solution containing ATP and $MnCl_2$. Under these conditions, the ATP does not react and the map shows the substrate lysine (in violet) and the position and conformation of the ATP (in green) before the first step of the reaction takes place. The map was calculated with the substrates omitted from the model, using all data between 25 and 2.35 Å, and contoured at 4σ. (B) Difference electron density map calculated with diffraction data collected from a crystal of the enzyme−lysine complex soaked overnight in a solution containing ATP and $MgCl_2$. The first step of the reaction occurred within the crystal, leading to the formation of the lysyl−adenylate intermediate. The additional electron density peak can be modeled by a partially occupied pyrophosphate molecule. The map was calculated with the coordinates of the lysyl−adenylate intermediate and pyrophosphate omitted from the model, using all data between 15 and 2.4 Å, and contoured at 3σ.

atomic models are very similar, confirming that the inhibitor AMP-PCP is a good analogue of the unreacted ATP.

Specificity for the ATP adenine base is achieved by hydrogen bond interactions with the purine ring nitrogens. The main chain carbonyl oxygen and nitrogen atoms of residue 271 make hydrogen bonds with both the N1 and 6-amino group of the substrate, while another strong hydrogen bond is found between nitrogen N7 and a well-ordered water molecule. In addition to these polar interactions, the adenine base is sandwiched between the phenyl ring of Phe 274 and the guanidinium moiety of Arg 480, stabilizing the aromatic ring through interactions involving π-electrons (Figure 3). The ATP ribose group is in the C3′-endo conformation and interacts with residue 421, with the 3′-hydroxyl group hydrogen bonding to the glutamate side chain, while the 2′-hydroxyl group is involved in a weaker interaction with the main chain carbonyl oxygen.

The α-phosphate makes a number of strong interactions involving primarily the guanidinium group of Arg 262, the only strictly conserved residue in motif 2, and a tightly bound Mg^{2+} ion (Mg 1). The motif 3 arginine (Arg 480), together with His 270 and two further Mg^{2+} ions (Mg 2 and Mg 3), makes a salt bridge with the γ-phosphate.

Lysyl−Adenylate Binding. When crystals of the LysU−lysine complex were soaked in a solution containing ATP and $MgCl_2$, a lysyl−adenylate intermediate was formed within the crystal. This corresponds to the first step of the reaction catalyzed by LysRS. A difference Fourier map showed well-defined electron density for the adenylate molecule (Figure 2B). An additional electron density peak could be modeled as a partially occupied pyrophosphate molecule. We considered the possibility of the electron density resulting from two overlapping states (the unreacted ATP substrate and the product lysyl−adenylate intermediate), but the electron density corresponding to the lysyl−adenylate intermediate was unambiguous, ruling out the presence of significant amounts of unreacted ATP. The α-phosphate has clearly moved to within binding distance of the α-carbon of the substrate lysine, and the distance between the α- and β-phosphates has increased. The adenosine moiety of the lysyl−adenylate intermediate binds to the synthetase in a manner similar to that of the unreacted ATP. Interactions of the enzyme with the α-phosphate are stronger in the lysyl−adenylate complex than those observed in the ATP complexes. When the electron density map of the lysyl−adenylate complex is compared with the model derived from the $MnCl_2$ soak, two peaks exactly coincide with the positions of the manganese ions. We therefore infer that in the crystal structure of the adenylate complex only two of the magnesium sites (Mg 1 and Mg 3) are retained. The site that is missing (Mg 2) would bridge the oxygen atoms of the β- and γ-phosphates of the partially occupied pyrophosphate molecule.

Aminoacylation Reaction. Superposition of the lysyl−adenylate and ATP/AMP-PCP structures is consistent with an in-line displacement mechanism for lysine activation. The ATP is bent into a U-shaped conformation in such a way as to place the α-phosphate in an ideal position to react with the carboxylate of the amino acid substrate. No enzyme residue participates directly in the reaction, and catalysis is achieved by binding both substrates (lysine and ATP) in the correct conformation. The negative charges on the lysine carboxylate and pyrophosphate moieties are stabilized through basic arginine and histidine residues and through three divalent cations. A subset of these positively charged chemical groups (motif 2 arginine Arg 262 and Mg 1) plays an additional role in assisting the reaction by enhancing the electrophilicity of the α-phosphate and stabilizing the pentavalent transition state. A comparison between the models of the unreacted ATP and the lysyl−adenylate intermediate (Figure 2A,B) clearly illustrates the inversion of configuration at the α-phosphate, caused by the in-line nucleophilic attack.

When the molecular models of the unreacted ATP and the lysyl−adenylate intermediate are compared, most of the enzyme−ligand contacts are unchanged, including contacts between the ATP and the class II invariant residues (Figures 3 and 4A,B). Although there is a shift of nearly 1 Å in the position of the α-phosphate, all the interactions made by the AMP moiety are very similar in the two complexes.

When the nonhydrolyzable ATP analogue AMP-PCP was used rather than ATP (data set 2), the aminoacylation reaction did not occur. Biochemical data (not shown) confirm that the lack of reactivity of AMP-PCP is also observed in solution and is not an artifact of the crystalline state of the enzyme. In AMP-PCP, the oxygen atom bridging the β- and γ-phosphates is substituted with a methylene group, which renders the bond between the β- and γ-phosphate non-

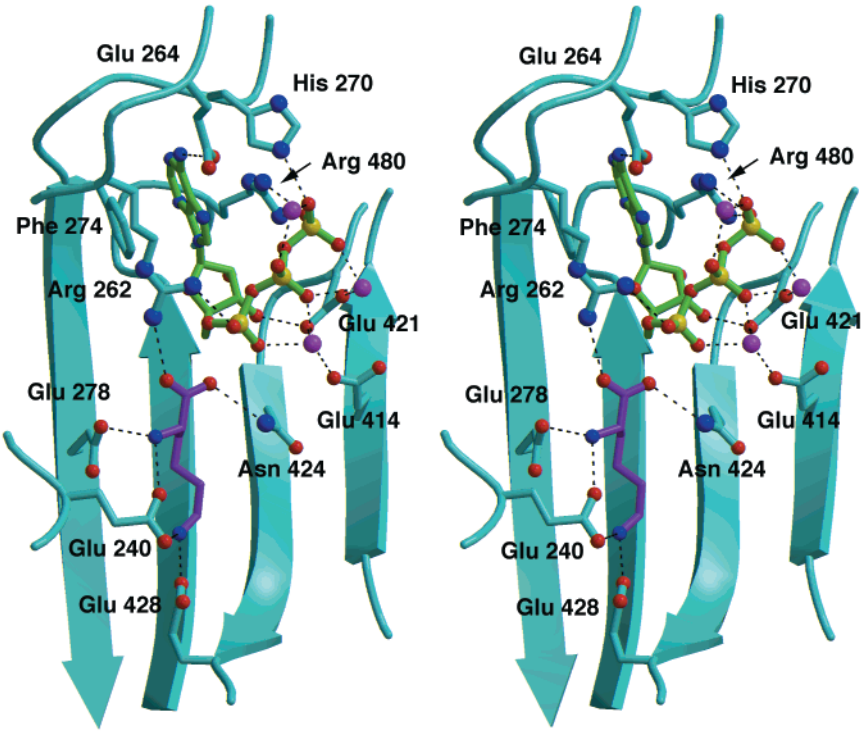

FIGURE 3: Stereoview of the active site of LysU showing the conformation of the substrates lysine (violet) and ATP (green) before the first step of the reaction takes place. The ATP molecule is located on one side of the central β-sheet of the C-terminal domain, with the adenine ring sandwiched between a conserved phenylalanine (Phe 274) and the motif 3 arginine residue (Arg 480). The pyrophosphate moiety is bent toward the adenine so as to place the α-phosphate in the correct position for nucleophilic attack of the lysine carboxylate oxygen. The hydrogen bonding and electrostatic interactions between the substrates and some of the key residues, including the invariant motif 2 arginine (Arg 262), are shown. The three Mg^{2+} ions involved in catalysis are included in magenta.

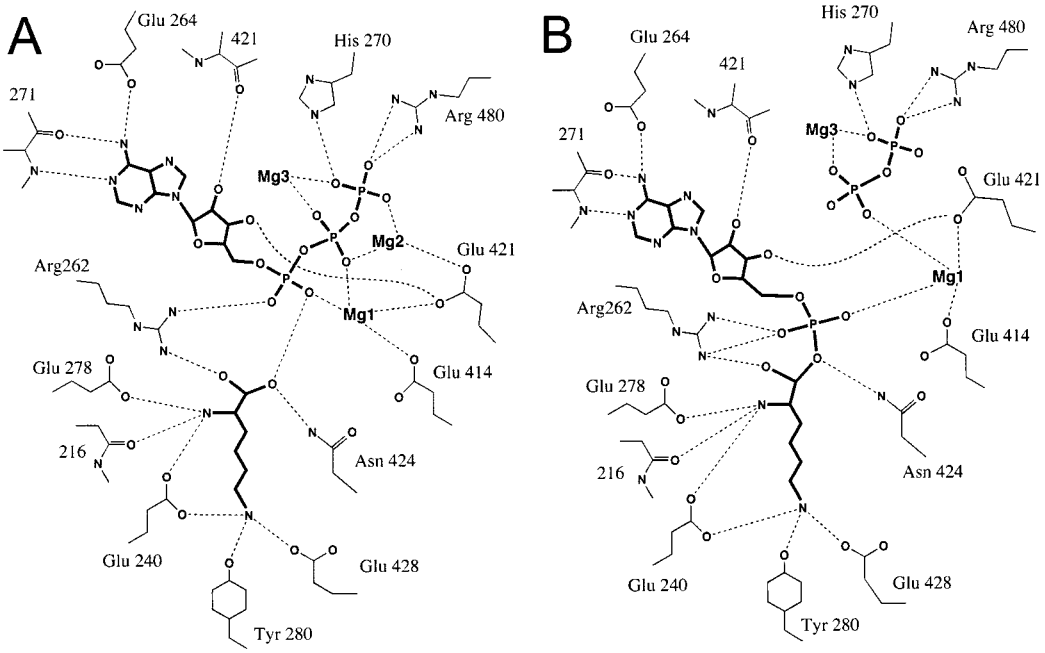

FIGURE 4: Schematic representation of the active site of LysU, showing potential hydrogen bonding interactions (A) in the ternary complex with lysine and ATP and (B) in the complex with the lysyl−adenylate intermediate. The invariant motif 2 Arg 262 plays a key role in the recognition of the lysine carboxylate and the ATP α-phosphate, while the invariant motif 3 Arg 480 binds the γ-phosphate of the ATP. A number of conserved residues in the motif 2 loop (residues 264−271) assume an ordered conformation only upon ATP binding. The positions of the Mg^{2+} sites are indicated.

hydrolyzable. Since ATP hydrolysis by aminoacyl-tRNA synthetases occurs through the cleavage of the bond between the α- and β-phosphates (to give AMP and PP_i), rather than the bond between the β- and γ-phosphates, it is not obvious why AMP-PCP is an inhibitor of the reaction and not a substrate. Most likely, this is due to electronic effects, which make PCP a worse leaving group than PP_i for the nucleophilic attack.

A similar inhibitory effect has been observed for the chemically analogous activity of tyrocidine synthetase 1, in

which both AMP-PCP and AMP-PNP failed to support amino acid-dependent adenylate formation (*31*). In a crystallographic study carried out on the yeast aspartyl-tRNA synthetase (*32*), AMP-PCP was found, in the absence of the amino acid substrate, to bind in an elongated conformation with the γ-phosphate occupying the aspartate specificity pocket. Our results suggest that the inhibition of the adenylation reaction is due to electronic effects rather than nonproductive binding.

In the absence of a nucleotide bound to the active site, the loop between residues 264 and 271 is not well-ordered (Figure 1), despite the fact that it includes some of the most conserved residues in class II synthetases. Most of the amino acid side chains cannot be modeled; residue 269 has not been built, and the position of the main chain for residues 268 and 270 can barely be determined (*16*). When ATP binding occurs, the entire loop acquires a well-ordered conformation and all the side chains show well-defined electron density, with a number of residues (Glu 264, His 270, and the main chain of Asn 271) making direct contact with the adenine moiety or the γ-phosphate (Figures 3 and 4). No other significant structural changes are observed.

Comparison with Other Class II Aminoacyl-tRNA Synthetases. A comparison of the active site of LysU with other class II enzymes reveals a remarkable similarity in the interactions involving ATP (*12*, *32*−*36*). The purine moiety of the nucleotide is invariably sandwiched between the aromatic side chains of topologically equivalent phenylalanine and arginine residues (Figure 3) and held in position by hydrogen bonds with the main chain of the homologues of residue 271 and the side chain of Glu 264 (both belonging to the conserved motif 2 loop, Figure 5). A positive charge supplied by a histidine or arginine residue originating from the same loop (His 270 in LysU) is also well-conserved and is always involved in the stabilization of the pyrophosphate group, together with the invariant motif 3 arginine (Arg 480 in LysU). In some class IIb enzymes, His 270 is preceded in the amino acid sequence by an arginine residue which also interacts with the γ-phosphate (*12*, *35*). LysRS has an identical primary sequence at this position (Arg-His), but the side chain of Arg 269 is slightly displaced and does not make a hydrogen bond with the pyrophosphate moiety. Although there are relatively few contacts between the enzyme and the ribose, the interactions are well-conserved and involve a conserved negatively charged residue that also coordinates the Mg^{2+} ions (Glu 421 in LysU).

Divalent cations such as Mg^{2+} or Mn^{2+} are always implicated in the reactions catalyzed by aminoacyl-tRNA synthetases. Both kinetic (*37*) and structural data (*33*) have shown that the Mg dependence of the aminoacylation reaction differs for class I and class II enzymes, with class I enzymes requiring only one Mg^{2+}, while class II enzymes require three cations. The Mg 1 site corresponds to the strongest metal site observed in all other class II synthetase structures (*12*, *33*, *35*), with the exception being HisRS (*34*, *38*) where the equivalent role is played by a conserved arginine. The two additional metal binding sites bridging the β- and γ-phosphates have been observed in crystal structures of other class II aminoacyl-tRNA synthetases complexed with ATP, such as SerRS (*33*), AspRS (*35*), and AsnRS (*12*), and are consistent with biochemical data (*37*).

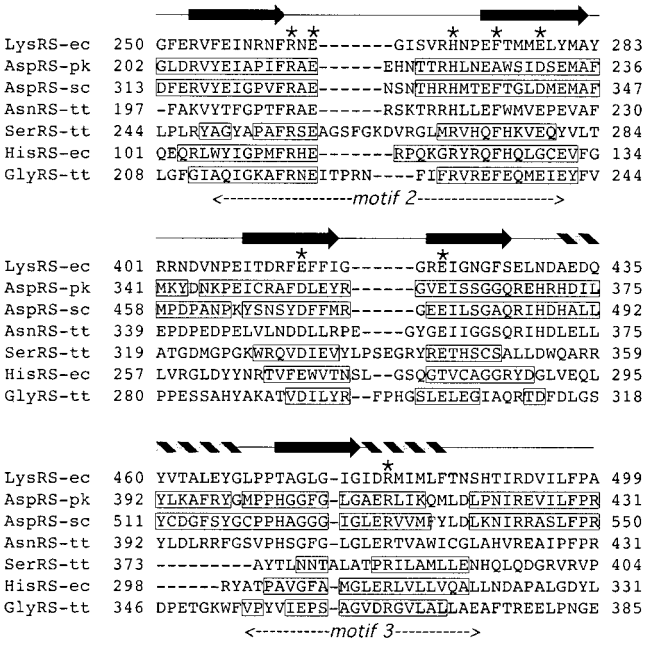

FIGURE 5: Structure-based sequence alignment of selected class II aminoacyl-tRNA synthetases whose three-dimensional structures have been determined in the presence of ATP or the aminoacyl-adenylate intermediate: *E. coli* lysyl-tRNA synthetase LysU (LysRS-ec), *Pyrococcus kodakaraensis* aspartyl-tRNA synthetase [AspRS-pk (*35*)], *Saccharomyces cerevisiae* aspartyl-tRNA synthetase [AspRS-sc (*32*)], *T. thermophilus* asparaginyl-tRNA synthetase [AsnRS-tt (*12*)], *T. thermophilus* seryl-tRNA synthetase [SerRS-tt (*33*)], *E. coli* histidyl-tRNA synthetase [HisRS-ec (*34*)], and *T. thermophilus* glycyl-tRNA synthetase [GlyRS-tt (*36*)]. The alignment is based on structural superposition, with the exception of the AsnRS structure for which the atomic coordinates have not been deposited in the Protein Data Bank. In this instance, the alignment is based on sequence comparisons and the information published in ref *12*. Three regions are shown: motif 2, centered around Arg 262; the segment which includes the negatively charged residues coordinating metal ions (Glu 414 and Glu 421 in LysU); and motif 3, centered around Arg 480. The position of secondary structure elements in LysU is indicated above the sequence. In each sequence, the residues for which the α-carbon atoms are within 2.5 Å of the corresponding position in LysU (after optimal superposition of the catalytic domains) are boxed. Residues which are important for substrate binding are indicated by asterisks.

A more general picture of the structural basis of the aminoacylation reaction can be obtained by extending the comparison to other enzymes unrelated to class II synthetases that carry out the activation of amino acid substrates with ATP. These include not only class I aminoacyl-tRNA synthetases but also the peptide synthetases responsible for nonribosomal peptide synthesis. Peptide synthetases catalyze the ATP-dependent activation of an amino acid as an acyl−adenylate intermediate. The amino acid moiety is subsequently transferred to the thiol group of a phosphopantetheine cofactor which acts as an acceptor, with a role similar to that of the tRNA acceptor stem. This adenylation domain bears no sequence or structural homology to either class of aminoacyl-tRNA synthetases involved in the ribosomal synthesis of polypeptides, despite the fact that the formation of the aminoacyl−adenylate intermediate is chemically analogous in the two systems (*39*).

Both in class I aminoacyl-tRNA synthetases (*40*) and in the adenylation domain of peptide synthetases (*39*), positive charges supplied by the protein (the second lysine of the KMS**K**S motif and a conserved lysine residue in peptide

synthetases) as well as a Mg^{2+} ion stabilize the negative charge on the α-phosphate and are likely to be involved in the stabilization of the pentavalent transition state. Additional positive charges are supplied by the enzyme to stabilize the negatively charged pyrophosphate moiety. Recognition of the adenine base often involves hydrogen bond interaction between the 6-amino group and main chain carbonyl oxygens.

In all these systems, the reaction occurs by a nucleophilic attack of the amino acid carboxylate on the ATP α-phosphate to form an aminoacyl−adenylate intermediate, with the elimination of a pyrophosphate group. The role of the enzyme active site is to ensure the correct position and orientation of the substrates and to counterbalance the negative charges of the carboxylate, the ATP phosphate groups, and the developing charges on the pentavalent transition state.

ACKNOWLEDGMENT

We are grateful to S. Blanquet and P. Plateau (École Polytechnique, Palaiseau, France) for the pXLys5 plasmid and A. Miller (Imperial College, London, U.K.) for samples of purified LysU. We thank the staff at the Synchrotron Radiation Source, Daresbury Laboratory, Elettra, Trieste, and at DESY, Hamburg (particularly Z. Dauter and I. Tews).

REFERENCES

1. Eriani, G., Delarue, M., Poch, O., Gangloff, J., and Moras, D. (1990) *Nature 347*, 203−206.
2. Brick, P., Bhat, T. N., and Blow, D. M. (1989) *J. Mol. Biol. 208*, 83−98.
3. Rould, M. A., Perona, J. J., Söll, D., and Steitz, T. (1989) *Science 246*, 1135−1142.
4. Landes, C., Perona, J. J., Brunie, S., Rould, M. A., Zelwer, C., Steitz, T. A., and Risler, J. L. (1995) *Biochimie 77*, 194−203.
5. Cusack, S., Berthet-Colominas, C., Härtlein, M., Nassar, N., and Leberman, R. (1990) *Nature 347*, 249−255.
6. Cusack, S. (1997) *Curr. Opin. Struct. Biol. 7*, 881−889.
7. Arnez, J. G., and Moras, D. (1997) *Trends Biochem. Sci. 22*, 211−216.
8. Marahiel, M. A., Stachelhaus, T., and Mootz, H. D. (1997) *Chem. Rev. 97*, 2651−2673.
9. Ibba, M., Morgan, S., Curnow, A. W., Pridmore, D. R., Vothknecht, U. C., Gardner, W., Lin, W., Woese, C. R., and Söll, D. (1997) *Science 278*, 1119−1122.
10. Ibba, M., Losey, H. C., Kawarabayasi, Y., Kikuchi, H., Bunjun, S., and Söll, D. (1999) *Proc. Natl. Acad. Sci. U.S.A. 96*, 418−423.
11. Ruff, M., Krishnaswamy, S., Boeglin, M., Poterszman, A., Mitschler, A., Podjarny, A., Rees, B., Thierry, J. C., and Moras, D. (1991) *Science 252*, 1682−1689.
12. Berthet-Colominas, C., Seignovert, L., Härtlein, M., Grotli, M., Cusack, S., and Leberman, R. (1998) *EMBO J. 17*, 2947−2960.
13. Lêvèque, F., Plateau, P., Dessen, P., and Blanquet, S. (1990) *Nucleic Acids Res. 18*, 305−312.
14. Lêvèque, F., Gazeau, M., Fromant, M., Blanquet, S., and Plateau, P. (1991) *J. Bacteriol. 173*, 7903−7910.
15. Brevet, A., Chen, J., Lêvèque, F., Blanquet, S., and Plateau, P. (1995) *J. Biol. Chem. 270*, 14439−14444.
16. Onesti, S., Miller, A. D., and Brick, P. (1995) *Structure 3*, 163−176.
17. Cusack, S., Yaremchuk, A., and Tukalo, M. (1996) *EMBO J. 15*, 6321−6334.
18. Onesti, S., Theoclitou, M. E., Pernilla, E., Wittung, L., Miller, A. D., Plateau, P., Blanquet, S., and Brick, P. (1994) *J. Mol. Biol. 243*, 123−125.
19. Otwinowski, Z. (1993) in *Data collection and processing*, pp 56−62, SERC, Daresbury DL/SC1/R34, Warrington, U.K.
20. Collaborative Computational Project Number 4 (1994) *Acta Crystallogr. D50*, 760−763.
21. Brünger, A. T., Kuriyan, J., and Karplus, M. (1987) *Science 235*, 458−460.
22. Engh, R. A., and Huber, R. (1991) *Acta Crystallogr. A47*, 392−400.
23. Brünger, A. T. (1992) *Nature 355*, 472−474.
24. Merritt, E. A., and Bacon, D. J. (1997) *Methods Enzymol. 227*, 505−524.
25. Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.
26. Esnouf, R. M. (1997) *J. Mol. Graphics 15*, 133−138.
27. Murzin, A. G. (1993) *EMBO J. 12*, 861−867.
28. Bochkarev, A., Bochkareva, E., Frappier, L., and Edwards, A. M. (1999) *EMBO J. 18*, 4498−4504.
29. Bochkarev, A., Pfuetzner, R. A., Edwards, A. M., and Frappier, L. (1997) *Nature 385*, 176−181.
30. Hynes, T. R., and Fox, R. O. (1991) *Proteins 10*, 92−105.
31. Dieckmann, R., Pavela-Vrancic, M., von Dohren, H., and Kleinkauf, H. (1999) *J. Mol. Biol. 288*, 129−140.
32. Cavarelli, J., Eriani, G., Rees, B., Ruff, M., Boeglin, M., Mitschler, A., Martin, F., Gangloff, J., Thierry, J. C., and Moras, D. (1994) *EMBO J. 13*, 327−337.
33. Belrhali, H., Yaremchuk, A., Tukalo, M., Berthet-Colominas, C., Rasmussen, B., Bösecke, P., Diat, O., and Cusack, S. (1995) *Structure 3*, 341−352.
34. Arnez, J. G., Augustine, J. G., Moras, D., and Francklyn, C. S. (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*, 7144−7149.
35. Schmitt, E., Moulinier, L., Fujiwara, S., Imanaka, T., Thierry, J.-C., and Moras, D. (1998) *EMBO J. 17*, 5227−5237.
36. Arnez, J. G., Dock-Bregeon, A. C., and Moras, D. (1999) *J. Mol. Biol. 286*, 1449−1459.
37. Airas, R. K. (1996) *Eur. J. Biochem. 240*, 223−231.
38. Åberg, A., Yaremchuk, A., Tukalo, M., Rasmussen, B., and Cusack, S. (1997) *Biochemistry 36*, 3084−3094.
39. Conti, E., Stachelaus, T., Marahiel, M. A., and Brick, P. (1997) *EMBO J. 16*, 4174−4183.
40. Arnez, J. G., and Steitz, T. A. (1994) *Biochemistry 33*, 7560−7567.

BI0006722