Class II Aminoacyl Transfer RNA Synthetases: Crystal Structure of Yeast Aspartyl-tRNA Synthetase Complexed with $tRNA^{Asp}$

M. RUFF, S. KRISHNASWAMY,* M. BOEGLIN, A. POTERSZMAN, A. MITSCHLER, A. PODJARNY, B. REES, J. C. THIERRY, D. MORAS†

The crystal structure of the binary complex $tRNA^{Asp}$-aspartyl tRNA synthetase from yeast was solved with the use of multiple isomorphous replacement to 3 angstrom resolution. The dimeric synthetase, a member of class II aminoacyl tRNA synthetases (aaRS's) exhibits the characteristic signature motifs conserved in eight aaRS's. These three sequence motifs are contained in the catalytic site domain, built around an antiparallel β sheet, and flanked by three α helices that form the pocket in which adenosine triphosphate (ATP) and the CCA end of tRNA bind. The $tRNA^{Asp}$ molecule approaches the synthetase from the variable loop side. The two major contact areas are with the acceptor end and the anticodon stem and loop. In both sites the protein interacts with the tRNA from the major groove side. The correlation between aaRS class II and the initial site of aminoacylation at 3′-OH can be explained by the structure. The molecular association leads to the following features: (i) the backbone of the GCCA single-stranded portion of the acceptor end exhibits a regular helical conformation; (ii) the loop between residues 320 and 342 in motif 2 interacts with the acceptor stem in the major groove and is in contact with the discriminator base G and the first base pair UA; and (iii) the anticodon loop undergoes a large conformational change in order to bind the protein. The conformation of the tRNA molecule in the complex is dictated more by the interaction with the protein than by its own sequence.

THE FIDELITY OF TRANSLATION OF GENETIC INFORMATION relies upon the accurate aminoacylation of tRNA's by their cognate aminoacyl tRNA synthetases (aaRS's). The 20 or more enzymes present in each living cell perform the task through a two-step reaction (*1*). The catalytic activity involves the specific recognition of three substrates: the amino acid and ATP for the formation of an intermediate adenylate, and the tRNA for the final step of aminoacylation in which the amino acid is attached to the ribose of the terminal adenine. Structural data should aid in understanding the molecular mechanism underlying this highly specific reaction.

From the structural point of view, aaRS's exhibit a wide divergence. Their molecular size, oligomeric state, primary sequence, and three-dimensional (3-D) structure, when known, highlight their diversity. Despite the common catalytic function of these enzymes, most of the available data point toward their idiosyncratic evolution. The first 3-D structures determined at high resolution [*Bacillus stearothermophilus* TyrRS (*2*), *Escherichia coli* MetRS (*3*), and *E. coli* GlnRS (*4*)] underline the existence of one common structural domain similar to the nucleotide binding domain observed in many other molecules, including the kinases and dehydrogenases (*5, 6*). This domain exhibits the characteristic Rossmann fold (α, β structure with a central parallel β sheet flanked by α helices). This correlation led to the hypothesis that all synthetases would have evolved from this common ancestor (*4*). However, sequence analyses of all known aaRS's led Eriani *et al.* (*7*) to discard this hypothesis and to propose the first comprehensive classification of aaRS's into two classes of ten members each. In addition, this study revealed for the first time a correlation between a well-known functional property of aaRS's, their capacity to charge the amino acid at either the 2′ or the 3′ hydroxyl group of the 3′end ribose (*8, 9*), and structural features linked to the class distribution. The crystal structure of *E. coli* SerRS (*10*) supported this classification. In this enzyme the characteristic Rossmann fold is replaced by a domain built around a central antiparallel β sheet.

Class I includes three enzymes with known 3-D structure (MetRS, TyrRS, and GlnRS) and seven others (GluRS, ArgRS, ValRS, IleRS, LeuRS, TrpRS, and CysRS) (*7, 11*), which most certainly share one common 3-D structural domain with the characteristic signatures [HIGH (His-Ile-Gly-His) (*12*) and KMSKS (Lys-Met-Ser-Lys-Ser) (*13, 14*)] associated with the Rossmann fold. Sequence analyses suggest the existence of subgroups with extended homologies (that is, Cys, Met, Val, Leu, and Ile) (*11, 15*). The first crystal structure of a tRNA-synthetase complex in this class of enzymes was obtained for the *E. coli* glutaminyl system (*4*). This structure showed the importance of the shallow minor groove of RNA in the molecular recognition process. A peculiar conformation of the CCA end of $tRNA^{Gln}$ was also noticed.

The ten other enzymes, including SerRS and AspRS, make up for class II synthetases, which share one or more of three structural

The authors are in the Laboratoire de Cristallographie Biologique, Institut de Biologie Moléculaire et Cellulaire du CNRS, Université Louis Pasteur, 15 rue René Descartes, 67084 Strasbourg Cedex, France.

*Present address: Madurai Kamaraj University, Madurai, India.
†To whom correspondence should be addressed.

motifs showing a common sequence signature. Like class I synthetases, this class can be further subdivided in subgroups of extended homologies [that is, Ser, Thr, and Pro (7), or Asp, Asn, and Lys (*16–18*)]. The crystal structure described in this article is that of a complex between a tRNA and a member of this second family of aaRS's. Cocrystals for this system have been available for some time (*19*). Four different crystal forms could be characterized (*20*). A cubic form led to a low-resolution analysis (*21*). Subsequently an orthorhombic form diffracting to better than 2.7 Å resolution led to the present study. The molecular structure of the free tRNA is also known (*22, 23*). Crystals of the free enzyme have also been obtained (*24*).

Structure determination. Yeast AspRS is an α_2 dimer (molecular mass = 2 × 63 kD) that binds two tRNA molecules. Its amino acid sequence (557 residues) was determined by classical peptide analysis (*25*) and from the gene (*26*). Crystals of yeast $tRNA^{Asp}$-AspRS used for this study were grown with the use of the procedure described by Ruff *et al.* (*20*). The synthetase was purified according to a revised version of the published protocol from wild-type yeast cells supplied by FALA (Strasbourg) (*27*). The study has now been pursued with enzyme purified from cloned cells (*26*). The $tRNA^{Asp}$ was purified from bulk tRNA's (Boehringer Mannheim) with the use of countercurrent separation and two chromatographic steps (*28*). Large crystals can now be grown routinely by seeding a solution containing ammonium sulfate at 52 percent saturation. Crystals belong to space group $P2_12_12$ with unit cell parameters a = 210 Å, b = 146 Å, and c = 86 Å and one dimeric molecule per asymmetric unit. The crystal structure was solved by multiple isomorphous replacement (MIR). Three derivatives were used: gold chloride to 6 Å resolution only, mercury (PCMB, *p*-chloromercuribenzoic acid) and samarium ($SmCl_2$) to 3.5 Å resolution. However, the poor phasing power of these derivatives clearly indicates a lack of significant contribution below 4 Å. Data relevant to intensity measurements for native and derivatives are summarized in Table 1. The heavy atom sites, which were located in difference Patterson and difference Fourier maps, were refined first against centric reflections and then against phases calculated from solvent-leveled maps (*29*). This last step improved the mean MIR phase error, as measured against current phases, by 7°. The automatic mask used to define the solvent regions occupied 40 percent of the crystal volume. The solvent-leveled map clearly showed the envelope of the molecular dimer and allowed the positioning of the tRNA molecules (Fig. 1A). Table 2 shows the statistics relevant to MIR phasing. The local twofold axis was determined from the refined heavy atom positions, which are pairwise symmetric about this axis. The orientation of the axis is in agreement with self-rotation functions. This information enabled a further phase improvement in which the MIR phases were combined with phases calculated after averaging the density within the molecular envelope (*30*).

Table 1. Data collection. Diffraction data were collected in the laboratory with a Siemens 2-D area detector for resolution data sets between 40 and 6 Å (essentially native and gold substituted crystals) and at the EMBL synchrotron outstation by DESY in Hamburg with the imaging plate system locally developed by J. Henricks and A. Lentfer and a modified version of the MOSFLM processing system of Imperial College (adapted for image plate by the EMBL Hamburg station) with profile fitting introduced by A. Lesley. Crystals used were plates of 0.4 mm by 0.8 mm in their largest dimensions with a thickness varying between 0.06 mm and 0.2 mm. Systematic errors due to absorption effects were minimized by collecting data at a short wavelength (1.0 Å). Native data were derived from a collection of 310 oscillations from 22 crystals. The high brilliance of the synchrotron X11 beam line allowed native diffraction data to be taken to 2.7 Å resolution.

Crystal	Maximum resolution (Å)	R_{sym}*	Reflections (no.)		$\Delta F/F$†
			Observed	Independent	
Native	2.7	9.0	327,516	70,895	
Hg	3.7	6.5	72,399	20,256	0.191
Au	5.7	7.2	16,068	4,896	0.211
Sm	3.5	7.0	77,909	22,560	0.152

*$R_{sym} = \Sigma|\bar{I}-I|/\Sigma I$, where I is the intensity. †F, structure factor.

Table 2. Location of heavy atoms in derivatives and MIR phasing. The atomic coordinates are expressed, in angstrom units, in a local system: the origin (fractional coordinates x = 0.335, y = 0.377, z = −0.128) is on the noncrystallographic twofold axis, taken as the z axis (direction cosines: −0.088, 0.475, 0.875). The mercury derivative (15 to 4 Å resolution) had 18,976 phased reflections and a phasing power of 2.76. The gold derivative (15 to 6 Å resolution) had 3,468 phased reflections and a phasing power of 1.95. The samarium derivative (15 to 4 Å resolution) had 16,096 phased reflections and a phasing power of 1.64. The average MIR figure of merit for all phased reflections in the resolution range from 15 to 4 Å is 0.61.

Crystal	x	y	z	occ*	x'	y'	z'	occ*
HG1	−21.1	17.9	0.3	4.4	20.8	−18.1	0.1	4.6
HG2	3.9	1.5	9.7	2.9	−3.8	−1.4	9.8	2.2
HG3	−21.2	17.4	−10.1	2.7	21.4	−17.3	−9.8	1.8
AU	−18.4	17.2	−1.8	7.6	17.5	−17.1	−1.4	8.3
SM	9.6	22.7	−3.7	5.1	−9.5	−21.8	−3.1	4.1

*occ = occupancy (arbitrary scale).

Several electron density maps were used for model building of the tRNA and protein with the graphic program FRODO (*31*). Although there is some weak MIR phase information beyond 4 Å resolution, the clearest map was obtained with the use of the MIR phases to 4 Å only, and then extending slowly the resolution from 4 to 3 Å by density modification (Table 3). An example of this 3 Å resolution map is shown in Fig. 1B.

The amino-terminal domain is being refined. It lacks ~100 residues that correspond to the beginning of the sequence. These include the first 70 residues, which can be removed by mild digestion with no apparent change in the enzymatic activity.

The dimeric complex. The molecular complex is made of a dimeric enzyme associated with two tRNA molecules (Fig. 2). The enzyme dimer is roughly diamond shaped, with diagonal dimensions of 110 Å by 90 Å, and rather flat: it is only 40 Å thick at the local twofold axis, but is bent along the main diagonal, which results in a total thickness of ~60 Å. This dimer is the association of two elongated subunits. The tRNA molecules occupy both ends of the main diagonal. The main interaction sites with the enzyme are inside the L shape, as predicted by Rich and Schimmel (*32*), and on the variable loop side (Fig. 3). Large portions of tRNA molecules from the middle of the acceptor stem to the middle of the anticodon stem, including the D and T stems and loops, are accessible to solvent and free to interact with other tRNA's in the crystals. These intermolecular contacts between tRNA molecules account for most of the crystal packing interactions.

Recognition of tRNA. The striking feature of this complex is the mode of interaction of the tRNA with the synthetase. Two major contact areas are clearly visible: (i) the end of the acceptor stem including the single-stranded fragment GCCA and (ii) the end of the anticodon stem and loop. In contrast with the situation in the glutaminyl system, the $tRNA^{Asp}$ approaches the synthetase from the variable-loop side. The two modes of recognition look like mirror images of each other (Fig. 3). As a result, three characteristic features can be noticed. (i) The single-stranded fragment GCCA continues the helical conformation of the stem in a regular way, at least for its backbone part, and dips naturally in the active site of the

synthetase. The terminal adenine base fits in a hydrophobic pocket flanked by the β strands of the active site. (ii) One of the major contacts between synthetase and tRNA involves the central loop of the conserved motif 2 and the end of the acceptor stem on its major groove side. This loop makes close contact with the discriminatory base G and the first base pair UA. The major groove interaction, which extends to the following base pair, is made possible by the larger accessibility of the bases at the end of helices (see below). (iii) A tight contact with strong interactions exists at the anticodon loop level, as shown in the top right corner of Fig. 2B. This loop undergoes an important conformational change in order to present unstacked bases of the anticodon (nucleotides 35 to 38) to the protein. There again, the RNA stem is approached from the major groove side at the end of the helix. Some small interactions involving the minor groove of tRNA occur at the level of anticodon and acceptor stems.

Comparison with the free tRNA. The primary sequence of yeast $tRNA^{Asp}$ is shown on Fig. 4. The superposition of $tRNA^{Asp}$, as seen in the complex with its cognate synthetase, with the molecule in its free state is shown in Fig. 5. (*23*). The superposition indicates a good conservation of the 3-D structure from the middle of the acceptor stem to the middle of the anticodon stem, including the D and T stems and loops where all of the tertiary interactions occur. The conserved part accounts for 54 nucleotides out of 75. In this part of the molecule the fluctuations are small. The three base pairs at the end of the acceptor stem undergo a small but significant change. The conformation of the acceptor strand can not be compared with that of the free tRNA since packing effects dominate in the crystals of free tRNA. The most significant changes occur in the anticodon loop and stem. Deviations start after base pairs G30·U40 and is greatest for residues 34 to 37. The helical conformation is preserved up to residue 33 but then the canonical U-turn conformation is disrupted with a concomitant unstacking of the anticodon bases. All of these changes result in a large movement of the loop. The distance between the two positions of the phosphate groups of residue 35 is close to 20 Å. It can be noticed that the overall pinch of the L-shape involves the most flexible parts of the molecule as detected by *B*-factor analysis of the native structure. Most remarkable is the location of the hinge point in the anticodon stem at base pair 30·40, which is a Watson-Crick base pair for all tRNA's with a few exceptions, such as yeast $tRNA^{Asp}$ where it is a G-U pair. This is also the point where the stems deviate from each other when the structures of yeast $tRNA^{Phe}$ and yeast $tRNA^{Asp}$ are compared.

Comparison with other tRNA's. The molecular structure of yeast $tRNA^{Phe}$ (*33*) is to some extent intermediate between free and complexed $tRNA^{Asp}$ (Fig. 5B). The comparison of these three structures supports the earlier hypothesis that the formation of duplexes of $tRNA^{Asp}$ through anticodon association was responsible for the observed difference in the angular aperture of the limbs between the two free tRNA's (*34*). In these duplexes the flexibility of the stems accommodates most of the conformational changes. The comparisons suggest that the crystal structure of $tRNA^{Phe}$ is the closest view to the structure of a free tRNA, whereas the two conformations of $tRNA^{Asp}$ represent different functional states. The conformational variability corresponds to the mobility of the nucleic acids suggested by the *B*-factor distribution.

Comparison with $tRNA^{Gln}$ as seen in its complexed form enlightens one major difference in the acceptor end part (Fig. 5C). The conformation in $tRNA^{Asp}$ is quite close to that of a regular RNA helix, not only in the stem but also in the single-stranded CCA end, which dips in the active site with a regular helical conformation of its phosphate ribose backbone. In contrast, the first base pair of the CCA stem in $tRNA^{Gln}$ is disrupted and the single-stranded part makes a hairpin turn in order to reach the active site. The discriminator base G73 forms a sequence-specific hydrogen bond with phosphate 72. The question arises whether the type of conformation is dependent on the tRNA sequence at the acceptor end. The large sequence similarity between $tRNA^{Gln}$ and $tRNA^{Asp}$ in this region (Fig. 4) eliminates this hypothesis.

AspRS: A member of class II synthetases. Although the dimer exhibits a rather compact shape each monomer is elongated and composed of at least two distinct domains. The carboxyl-terminal domain forms a central core that includes the catalytic site. It is built around a large antiparallel β sheet surrounded by α helices (Fig. 6). The overall topology of this domain (Fig. 7) is very close to that found in SerRS from *E. coli*. This domain contains the three conserved motifs characteristic of class II synthetases. [When we mention motifs 1, 2, and 3 in the text, we refer to the definition deduced from the sequence analysis (7). The real size of the structurally conserved motifs is obviously larger.] Motif 1 is involved in the dimer interface and motifs 2 and 3, which contribute to the formation of a catalytic cavity, are involved in the recognition of the CCA end of the tRNA and the ATP (Fig. 6). The general organization is the following:

1) Motif 1 contains an α helix (H1 on Fig. 7) followed by a distorted β strand S1 parallel over a few residues to the first strand S4 of the antiparallel β sheet. The helix is in contact with the equivalent helix H1′ of the second monomer.

2) Motif 2 is formed by the two first strands (S4 and S5) of the six-stranded antiparallel sheet. The loop between the two strands interacts with the acceptor stem of tRNA. It is of variable length in the class II family.

3) Motif 3 contains strand S10 and helix H9, which is connected to the carboxyl terminus of the protein.

The dimeric association is strengthened by the interactions within another antiparallel β sheet, namely the four-stranded sheet S3, S2, S2′, and S3′. Even though the corresponding part of the sequence, inserted between motifs 1 and 2, shows no homology within the class II synthetase family, it is interesting that SerRS contains the topological equivalent (*10*). Moreover, the loop between S2 and S3 is in the vicinity of the 5′ end of the second tRNA. This is the only observed contact between one monomer and the tRNA bound to the other. Comparison with the structure of SerRS shows a region with high structural homology (shaded in Fig. 7), which comprises the three motifs plus helices H2 and H3 and strands S2, S3, S7, S8, and S9. The two remaining domains (blank in Fig. 7) with no counterpart in the Ser structure are an insertion of helices H7 and H8 between strands S9 and S10 and an insertion of helix H4, H5, and H6 and strand S6 between helix H3 and strand S7.

Table 3. Density modification and phase extension. Density modification included solvent leveling, high-density and negative-density attenuation, and local averaging. Phase extension was done by slowly increasing the resolution, in steps of $a^* = 1/210$ Å^{-1} in $2(\sin\theta)/\lambda$. During this phase extension procedure, calculated phases were combined with MIR phases for reflections of resolution lower than 4 Å, while the calculated phases were taken for all reflections between 4 and 3 Å.

Procedure	Phased reflections (no.)	Symmetry correlation*	Correlation with 4 Å–MIR map
MIR 4 Å	20,572	0.672	
Density modification at 4 Å	22,409	0.868	0.770
Phase extension at 3 Å	50,286	0.879	0.659

*Average correlation coefficient (weighted by the squared density) between the electron density at points related by the local twofold symmetry within the molecular envelope.

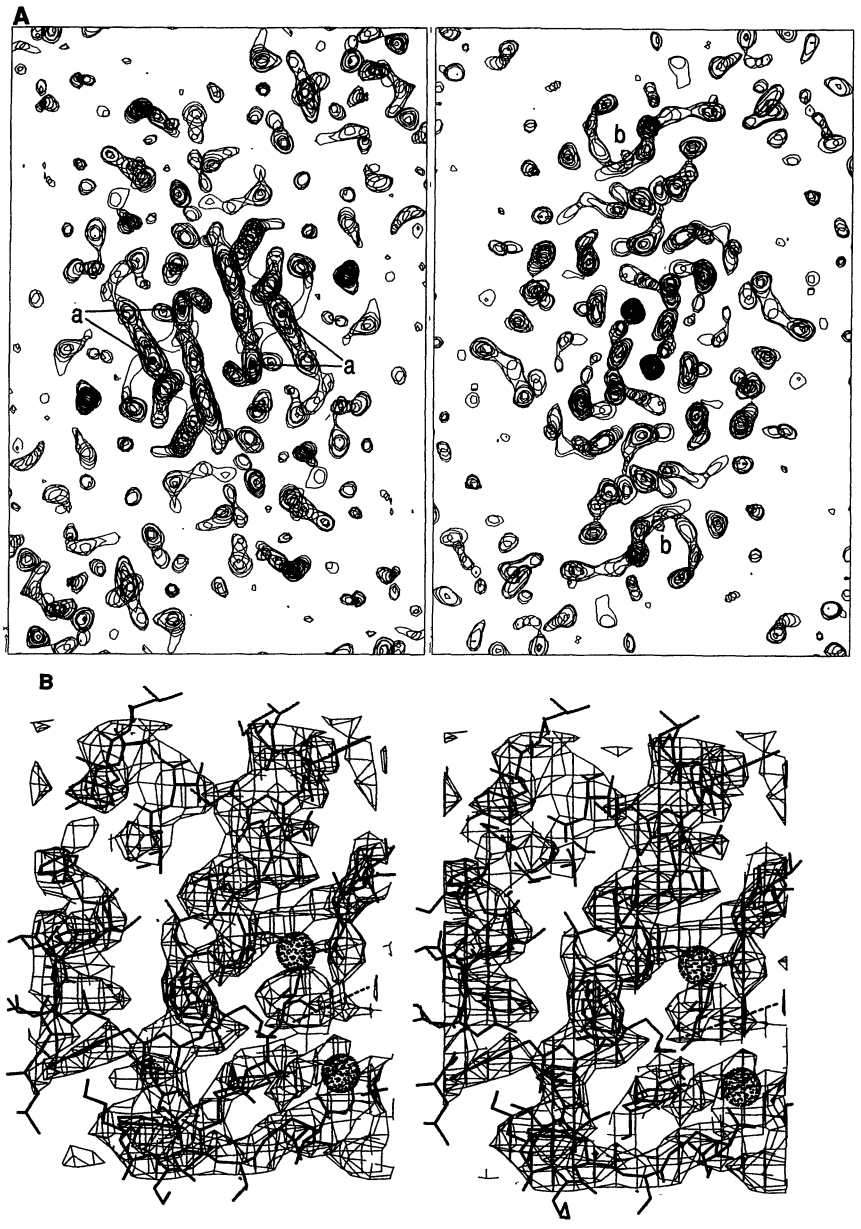


Fig. 1. (**A**) Electron density map at 6 Å resolution. The initial MIR phases were improved by solvent flattening, and the resulting map was averaged around the noncrystallographic molecular twofold axis. Each view is the superposition of 15 sections perpendicular to the molecular axis with a separation of 1 Å between sections. The limits of the molecular dimer are clearly visible. The α helices H1 and H3 (see Fig. 7) are seen in (A), left, part of the anticodon stem of tRNA in (B), right. (**B**) An example of the experimental 3 Å resolution map used for model building (stereoview). The α helix H1 of motif 1 is seen in the middle. The molecular axis (dashed line) is materialized at the lower right. Also shown is the mercury site (and its noncrystallographic symmetry equivalent) near Cys^{255} in the corresponding heavy-atom derivative.

The connection between H8 and S10 of motif 3 occurs between residues 512 and 519, which are two cysteines labeled by the mercury atoms in the corresponding derivative. None of the three cysteines of each polypeptide chain points toward the tRNA molecule or the active site pocket. This observation is in agreement with the solution experiments that did not detect any functional consequence of the labeling of these residues (*35*). This portion of the map clearly indicates the connectivity despite a close contact with the amino-terminal part of the other subunit.

We can now define a general structural domain in place of the Rossmann fold domain of class I enzymes that can address requirements in the catalytic steps, such as positioning of ATP, the amino acid, and the acceptor stem with its single-stranded CCA end. This homologous domain comprises seven strands (six antiparallel and one parallel) and is flanked by three α helices. Even when the SerRS structure also shows an eighth strand placed approximately in the same position as S6, it runs in the opposite direction (parallel instead of antiparallel) and therefore it is not part of the homologous structural domain. The binding of A76 places the ribose in the right orientation for aminoacylation at the 3′ hydroxyl site, in agreement with the functional property associated with this class of enzymes. It is most likely that this can be generalized to all synthetases and that the binding of A76 to either a parallel sheet (class I) or an antiparallel sheet (class II) controls the orientation of the ribose, and as a consequence the primary site of aminoacylation. The variable loop of the conserved motif 2 binds in the major groove of $tRNA^{Asp}$ and makes interactions with the first base pair and the discriminator base. Its variability is most likely associated with a specific character of the recognition at this point.

The amino-terminal part of the sequence forms the anticodon binding domain, which exhibits a barrel-like topology (Fig. 2B). The protein makes a close contact with the ribose phosphate backbone between residues 34 and 38. This domain has no equivalent in any of the four other known synthetases.

Discussion. When the tRNA-synthetase in the yeast aspartyl system is compared with that in the *E. coli* glutaminyl system, the most striking feature is the symmetrical distribution of the molecular recognition patterns (Fig. 4): AspRS and GlnRS recognize opposite sides of the tRNA molecule. The $tRNA^{Asp}$ interacts with the variable-loop side and the major groove of the acceptor stem, whereas the $tRNA^{Gln}$ approaches the D-loop side and the minor groove of the acceptor stem. One obvious consequence of this

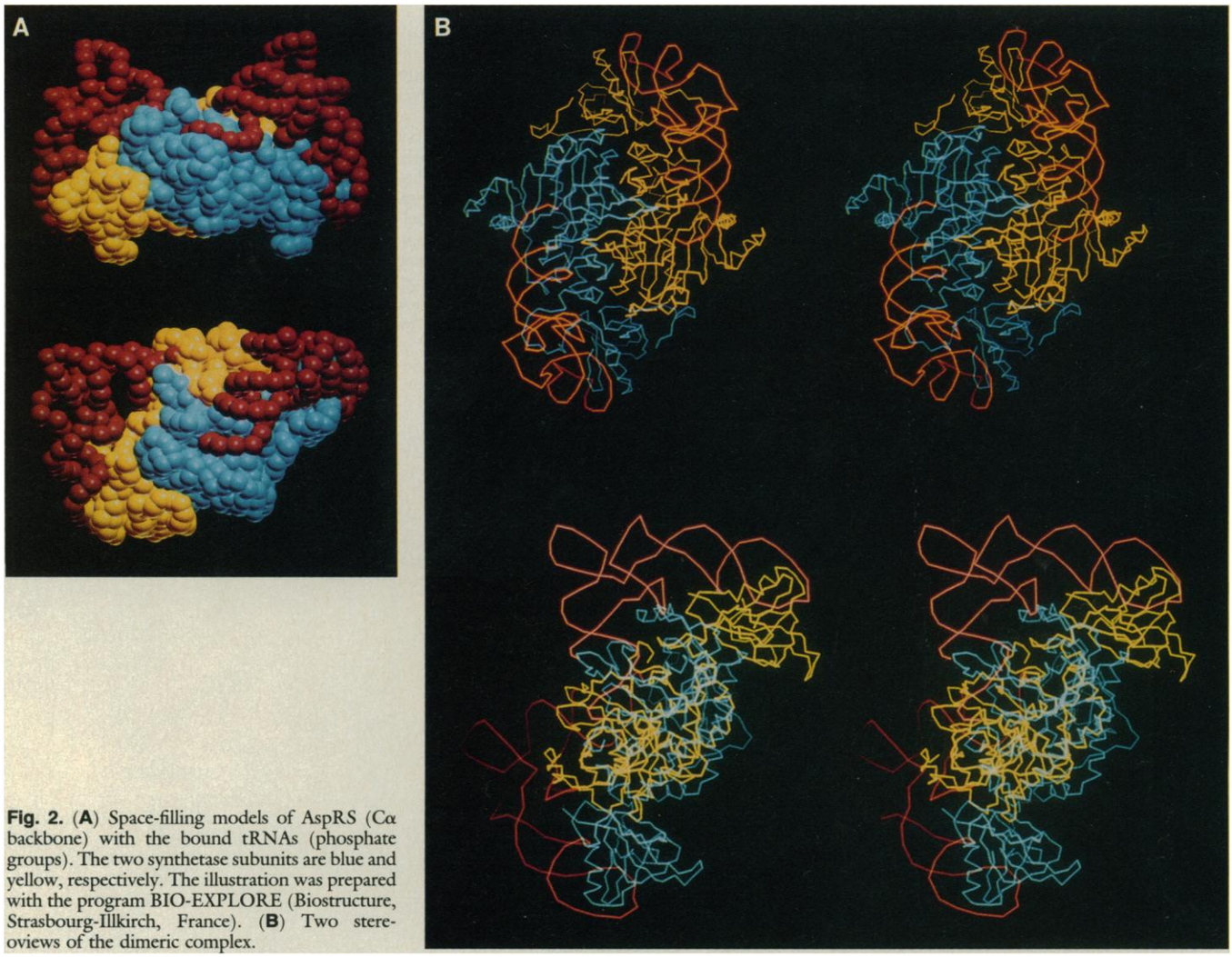


Fig. 2. (**A**) Space-filling models of AspRS (Cα backbone) with the bound tRNAs (phosphate groups). The two synthetase subunits are blue and yellow, respectively. The illustration was prepared with the program BIO-EXPLORE (Biostructure, Strasbourg-Illkirch, France). (**B**) Two stereoviews of the dimeric complex.

difference in the relative position of the tRNA and the synthetase is the conformation of the single-stranded GCCA end. The disruption of the first U·A base pair, the hairpin turn, and the specific intra-RNA interaction with the discriminator G base observed in the glutaminyl system are replaced in the aspartyl system by a regular helical conformation of both the stem and the CCA single strand, which are held in place by the tight RNA-protein contact at the active site cavity (Fig. 6). Since the primary sequences of $tRNA^{Asp}$ and $tRNA^{Gln}$ are similar in this region (Fig. 4), the conformational difference between $tRNA^{Asp}$ and $tRNA^{Gln}$ is most likely due to the interaction with the enzyme and is a clear example of an induced-fit mechanism of binding.

In B-DNA, the major groove is very often implicated in protein recognition. In A-DNA and RNA, however, the width of the narrow and deep major groove (~4 Å) prevents its accessibility to molecules of significant size (those larger than H_2O) and therefore most contacts involve the shallow minor groove. However, in this case the recognition through major groove interactions is made possible by the increased accessibility of approximately four base pairs at the end of each helix. Even when the third base pair (*36*) is an important identity element in some tRNA's, it is not in contact with the protein in the aspartic system. It would be of interest to carry on functional tests on $tRNA^{Asp}$ mutants at this position. In contrast, the first base pair, which does not occur in other yeast tRNA's, is in close interaction with the protein. Despite these observations, changing it to G-C does not affect the aminoacylation properties of the transcript tRNA (*37*). The conformational changes of the anticodon loop of $tRNA^{Asp}$ induced by the complexation enable closer interactions with the protein, essentially from the major-groove side. Calculations of the electrostatic potentials around yeast $tRNA^{Phe}$ and $tRNA^{Asp}$ show the existence of a less

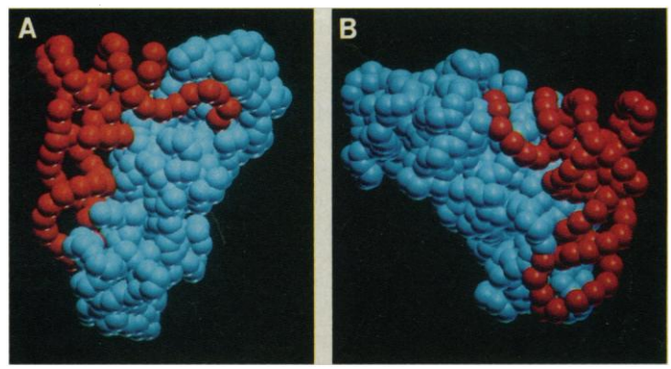
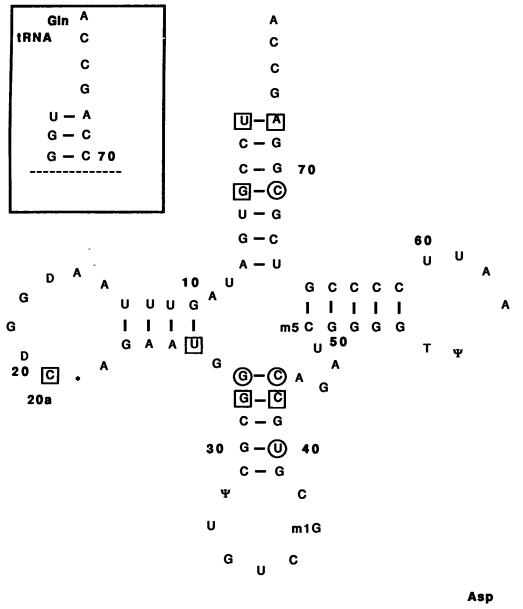


Fig. 3. CPK models of the GlnRS-$tRNA^{Gln}$ complex (**left**), first representative of class I systems, and AspRS-$tRNA^{Asp}$ (**right**) first representative of class II systems. Only phosphate and Cα backbones are shown. This figure emphasizes the opposite sides of approach of the tRNA molecules by the enzymes.



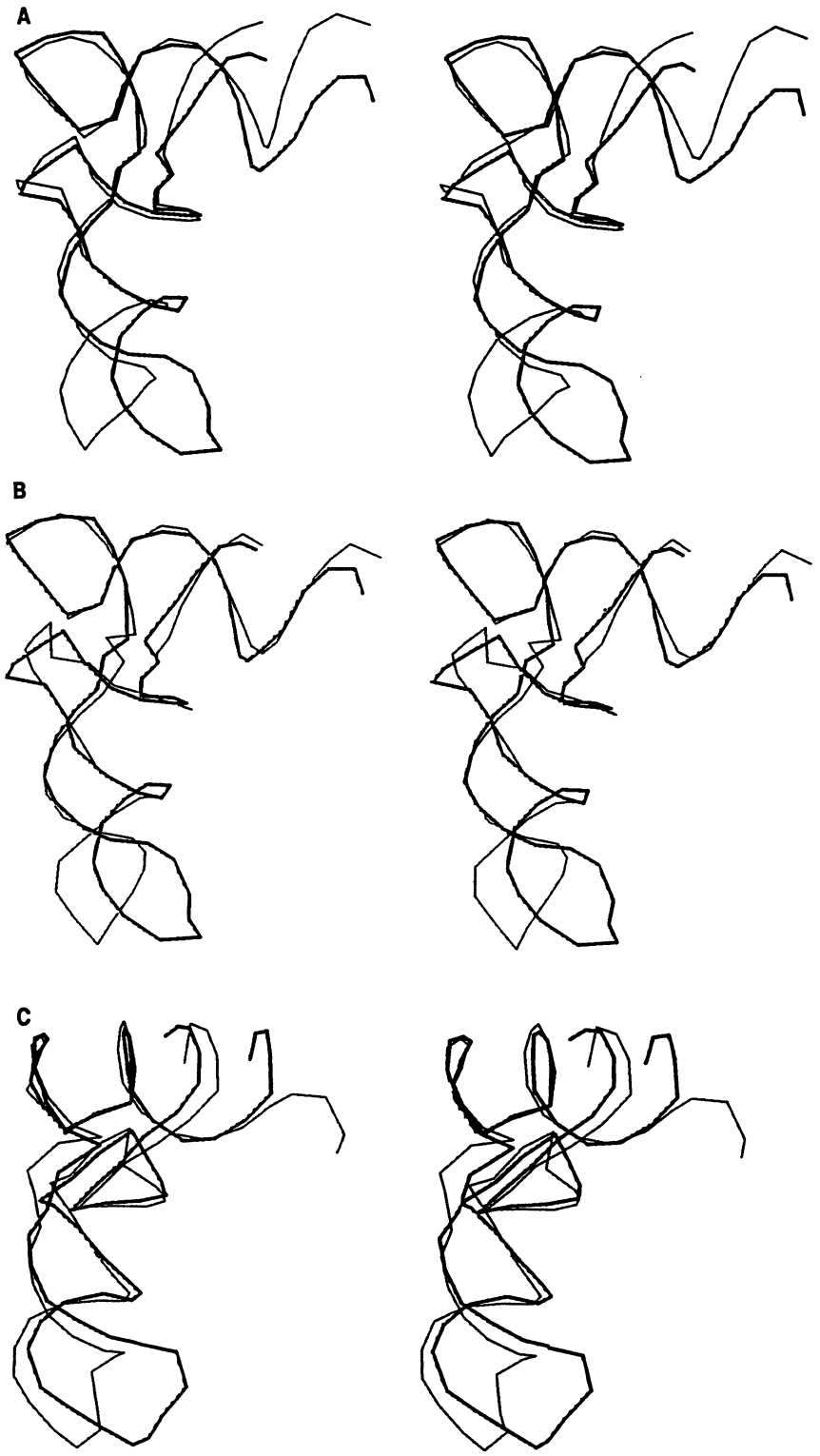


Fig. 4 (left). Nucleotide sequence and secondary structure of yeast $tRNA^{Asp}$ (*45*). The numbering is that of $tRNA^{Phe}$ (in $tRNA^{Asp}$ the variable loop is one nucleotide shorter). Circled residues correspond to nucleotides that are not found more than once at the given position when comparing the sequence to that of the 36 other known yeast tRNA's (*46*). Boxed nucleotides are found less than four times. **Fig. 5 (right).** (**A**) Superposition of the phosphate backbone of uncomplexed yeast $tRNA^{Asp}$ (thin line) on that of $tRNA^{Asp}$ complexed with the synthetase. Thirty phosphate atoms corresponding to the core of the structure (G6 to U12, A23 to A26, and C49 to G67) were used to determine the best superposition by least squares. The root-mean-square (rms) deviation over these 30 atoms was 1.3 Å. The rms deviation for the anticodon loop (G30 to U40) is 12.0 Å. The largest displacement is 20.0 Å for the U35 phosphate. (**B** and **C**) Superposition of the phosphate backbone of complexed yeast $tRNA^{Asp}$ (thick line) with yeast $tRNA^{Phe}$ (B) and with complexed *E. coli* $tRNA^{Gln}$ (C). The same 30 phosphate atoms as in (B) were used for this superposition. The rms deviation for these 30 phosphate atoms is 1.4 Å for (B) and 1.6 Å for (C). View (C) is slightly rotated around the vertical axis in order to show more clearly the major differences in the acceptor end.

negative hole in the potential surface of the anticodon region (*38*). This could direct negatively charged side chains that could then trigger the structural change associated with the binding. Thus all of the conformational changes can be explained as either increasing and optimizing the nucleic-acid protein interactions or favoring the positioning of the CCA end in the active site.

Initial studies of ATP-soaked crystals have allowed the location of the coenzyme to be determined (*39*). The relative position of A76 and ATP in the active site is consistent with the initial site of aminoacylation at the 3′ hydroxyl group of terminal adenine. When comparing with the position of the CCA in the glutaminyl system, it is tempting to generalize these binding modes to all synthetases in class II and class I, respectively. This would explain the observed correlation between the classes and their primary site of aminoacylation (*7*). We note that the classification of synthetases according to their interaction with various ATP analogs (*40, 41*) is in full agreement with the separation in two classes. The biochemical analysis leads to the existence of five different subclasses and suggests further differences in the ATP binding mode. More details are likely to emerge from the refinement.

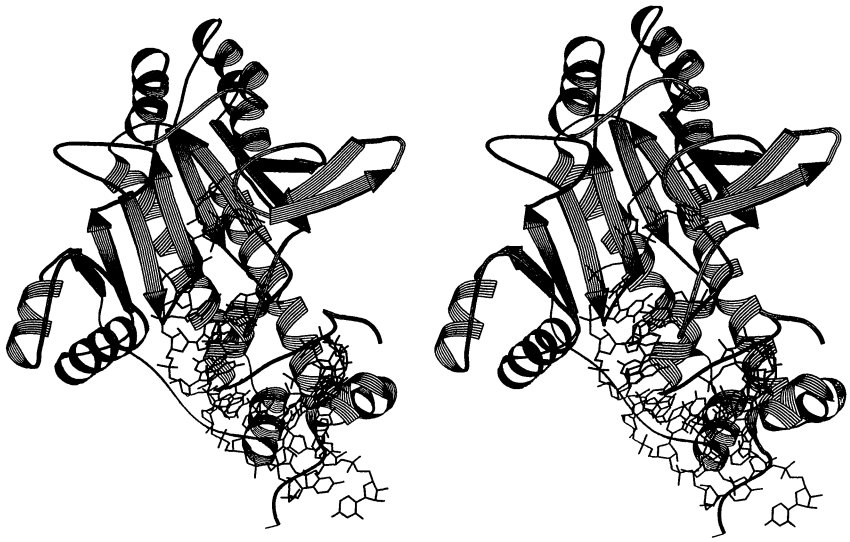


Fig. 6. RIBBON stereoview (*47*) of the carboxyl-terminal domain of one subunit of AspRS (residues 219 to the carboxyl-terminus 557) and the acceptor end (stem and CCA strand) of the interacting tRNA. This view emphasizes the active site cavity built around a six-stranded antiparallel β sheet.

The structural homology between SerRs and AspRS confirms the sequence alignments found earlier and extends them to most of the insertion between motifs 1 and 2 and to part of the insertion domain between motifs 2 and 3. While in SerRS this homologous part corresponds to almost all of the insertion, in AspRS it only represents about one half of the insertion domain. The active site domain with its large antiparallel central sheet is most likely the common scaffold on which class II synthetases are assembled. Its size (~250 amino acids, including the homologous part of the insertion between motifs 2 and 3) is larger than that of the corresponding Rossmann-fold domain, that is common to all class I synthetases (~150 amino acids). This difference could be explained by the fact that in class II enzymes the active site domain functions also include the CCA-end binding. This function is performed by another domain in class I enzyme, as seen in GlnRS, where the CCA end binds to a sheet that is part of a small domain of 100 amino acids inserted between strands of the ATP binding domain. Put together, these two domains have a size similar to that of class II synthetases.

The role of the amino-terminal domain would be to participate in the specific recognition of tRNA through its anticodon end. This is system specific and does not need to bear any homology with other synthetases. The same function is fulfilled by the carboxyl-terminal part in class I synthetases.

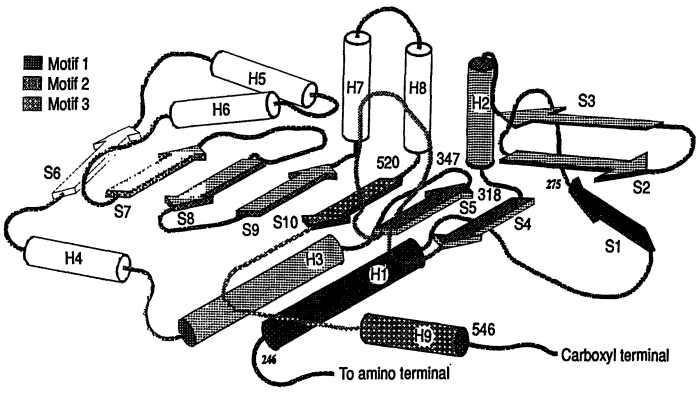


Fig. 7. Topology of the carboxyl-terminal domain of AspRS. The three conserved motifs characteristic of class II synthetases, deduced from sequence alignments (*7*), are shaded as shown: motif 1 (residues 258 to 275) includes α helix H1 and β strand S1; motif 2 (residues 315 to 349) includes β strands S4 and S5; and motif 3 (residues 517 to 549) includes β strand S10 and α helix H8. Comparison with the crystal structure of SerRS (*10*) suggests that most of this carboxyl-terminal domain (all of the shaded areas), with the exception of helices H4 to H8 and strand S6 (shown in blank), is topologically invariant within this class of synthetases.

Among class II synthetases, seven enzymes are α_2 dimers. PheRS is a heterodimer $\alpha_2\beta_2$ in *E. coli*. GlyRS and AlaRS are tetramers in *E. coli*. The dimeric interface involves two main contact areas relating interesting parts of the enzyme: (i) The conserved motif 1 already described and part of motif 2. These motifs are not conserved in the tetrameric enzymes, which constitute a special subgroup. In yeast, AlaRS is a monomeric enzyme (*42*); (ii) The central core domain of one subunit (active site) with the amino-terminal domain of the other subunit. These last contacts suggest potential cooperative recognition upon binding of the tRNA. Whether this point affects the kinetics of this enzyme remains to be seen. The relative positioning of the anticodon binding domain (amino terminal) and the central core are different in SerRS and AspRS. The two amino-terminal domains are on opposite sides of the active site domain. Therefore taking the antiparallel β-sheet plane as a reference, the amino-terminal is above the plane in SerRS and below in AspRS. If we assume that the two synthetases bind the acceptor end in a similar way, as can be inferred from the structural conservation of their active site, the previous observation leads to the conclusion that SerRS does not bind the anticodon loop of the corresponding $tRNA^{Ser}$. On the other hand, a topological study strongly suggests, in agreement with biochemical studies (*43*), that SerRS cannot bind the anticodon loop of the second tRNA in the complex.

Synthetases present a puzzling situation. They can be split in two groups of ten members with two different structural solutions to a common problem. These enzymes seem to have evolved from two original genes. Why two? It could be a requisite for the fidelity of information transfer, as suggested by Rich (*44*). It could also arise from the original binary system in a preexisting RNA world with its intrinsic ambiguity of the attachment at the 3′ end (3′- or 2′-OH). A sequential apparition of the two classes can also be postulated. This hypothesis would imply that class II enzymes appeared first since this class is the only one which contains all of the amino acids essential for protein folding, such as glycine, proline, and alanine. However the presence of most hydrophobic residues in class I and the even distribution of enzymes in the two classes favors a parallel evolution with similar time scales. Answers to these questions and the many more raised by these new structural informations lie at the heart of the elucidation of the origin of the genetic code.

REFERENCES AND NOTES

1. J. Normanly and J. Abelson, *Annu. Rev. Biochem.* **58**, 1029 (1989).
2. P. Brick, T. N. Bhat, D. M. Blow, *J. Mol. Biol.* **208**, 83 (1989).
3. S. Brunie, C. Zelwer, J. L. Risler, *ibid.* **216**, 411 (1990).
4. M. A. Rould, J. J. Perona, D. Söll, T. A. Steitz, *Science* **246**, 1135 (1989).
5. M. G. Rossmann, D. Moras, K. W. Olsen, *Nature* **250**, 194 (1974).
6. J. L. Risler, C. Zelwer, S. Brunie, *ibid.* **292**, 385 (1981).
7. G. Eriani *et al.*, *ibid.* **347**, 203 (1990).
8. T. H. Fraser and A. Rich, *Proc. Natl. Acad. Sci. U.S.A.* **72**, 3044 (1975).
9. F. Von der Haar and F. Cramer, *Biochemistry* **15**, 4131 (1975).
10. S. Cusack *et al.*, *Nature* **347**, 249 (1990).
11. G. Eriani, G. Dirheimer, J. Gangloff, *Nucleic Acids Res.*, in press.
12. T. Webster *et al.*, *Science* **226**, 1315 (1984).
13. C. Hountondji, P. Dessen, S. Blanquet, *Biochimie* **68**, 1071 (1986).
14. C. Hountondji, S. Blanquet, F. Lederer, *Biochemistry* **24**, 1175 (1985).
15. J. P. Heck and D. W. Hatfield, *J. Biol. Chem.* **263**, 868 (1988).
16. J. Anselme and M. Härtlein, *Gene* **84**, 481 (1989).
17. F. Lévêque, P. Plateau, P. Dessen, S. Blanquet, *Nucleic Acids Res.* **18**, 305 (1990).
18. G. Eriani, G. Dirheimer, J. Gangloff, *ibid.*, p. 7109.
19. R. Giégé *et al.*, *C.R. Acad. Sci. Paris D-2* **291**, 393 (1980).
20. M. Ruff *et al.*, *J. Mol. Biol.* **201**, 235 (1988).
21. A. Podjarny *et al.*, *J. Biomol. Struct. Dyn.* **2**, 187 (1987).
22. D. Moras *et al.*, *Nature* **288**, 669 (1980).
23. E. Westhof, P. Dumas, D. Moras, *J. Mol. Biol.* **184**, 119 (1985).
24. A. Dietrich *et al.*, *ibid.* **138**, 129 (1980).
25. I. Amiri *et al.*, *Biochimie* **67**, 607 (1985).
26. M. Sellami *et al.*, *Nucleic Acids Res.* **14**, 1657 (1986).
27. M. Ruff *et al.*, unpublished results.
28. A. C. Dock *et al.*, *Biochimie* **66**, 179 (1984).
29. B. C. Wang, *Methods Enzymol.* **115**, 90 (1985).
30. G. Bricogne, *Acta Crystallogr.* **A32**, 832 (1976).
31. J. A. Jones, *J. Appl. Crystallogr.* **11**, 268 (1978).
32. A. Rich and P. R. Schimmel, *Nucleic Acids Res.* **4**, 1649 (1977).
33. S. H. Kim *et al.*, *Science* **185**, 435 (1974); J. D. Robertus *et al.*, *Nature* **250**, 546 (1974).
34. D. Moras *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **83**, 932 (1986).
35. D. Kern *et al.*, *Eur. J. Biochem.* **193**, 97 (1990).
36. W. H. McClain and K. Foss, *Science* **240**, 793 (1988); Y.-M. Hou and P. Schimmel, *Nature* **333**, 478 (1988).
37. V. Perret *et al.*, *Nature* **344**, 787 (1990).
38. K. A. Sharp, B. Honig, S. C. Harvey, *Biochemistry* **29**, 340 (1990).
39. M. Ruff *et al.*, unpublished results.
40. W. Freist, H. Sternbach, F. Cramer, *Hoppe-Seyler's Z. Physiol. Chem.* **362**, 1247 (1981).
41. W. Freist, *Angew. Chem. Int. Ed. Engl.* **27**, 773 (1988).
42. D. Kern, R. Giégé, J. P. Ebel, *Biochim. Biophys. Acta* **653**, 83 (1981).
43. A. C. Dock, A. Garcia, R. Giégé, D. Moras, *Eur. J. Biochem.* **188**, 283 (1990).
44. A. Rich, in *Horizons in Biochemistry*, M. Kasha and B. Pullman, Eds. (Academic Press, New York, 1962), pp. 103–126.
45. J. Gangloff, G. Keith, J. P. Ebel, G. Dirheimer, *Nature New Biol.* **230**, 726 (1971).
46. H. B. Nicholas, Y.-M. Chen, W. H. McClain, *Cabios* **3**, 53 (1987).
47. J. P. Priestle, *J. Appl. Crystallogr.* **21**, 572 (1988).
48. We thank K. Wilson and the staff of the EMBL outstation at the DESY (Hamburg) for their support during high-resolution data collection, the staff of LURE (Orsay) and CHESS (Cornell University) where preliminary investigations were performed, P. Dumas for help with tRNA modeling, M. Delarue for many helpful discussions, J. Cavarelli for ATP binding analysis, A.-C. Dock for helping with tRNA purification, R. Ripp and M. Bergdoll for support in computing, L. Diebolt for typing the manuscript, and C. Briand for the CPK color figures. We acknowledge many useful contributions from colleagues of our laboratory, a fruitful visit of P. Nordlund, and useful discussions with G. Eriani, J. Gangloff, and D. Kern. We want to express our appreciation to R. Giégé and his group, and J. P. Ebel for their important contribution to the project. Supported by grants from C.N.R.S., Ministère de la Recherche et de la Technologie, and a Human Science Frontier award.

19 February 1991; accepted 5 April 1991