The crystal structure of phenylalanyl-tRNA synthetase from *Thermus thermophilus* complexed with cognate $tRNA^{Phe}$

Yehuda Goldgur[1], Lidia Mosyak[1], Ludmila Reshetnikova[2], Valentina Ankilova[3], Olga Lavrik[3], Svetlana Khodyreva[3] and Mark Safro[1]*

Background: In the translation of the genetic code each aminoacyl-tRNA synthetase (aaRS) must recognize its own (cognate) tRNA and attach the corresponding amino acid to the acceptor end of tRNA, discriminating all the others. The $(\alpha\beta)_2$ phenylalanyl-tRNA synthetase (PheRS) is one of the most complex enzymes in the aaRS family and is characterized by anomalous charging properties. Structurally, the enzyme belongs to class II aaRSs, as its catalytic domain is built around an antiparallel β sheet, but functionally it resembles class I as it aminoacylates the 2′OH of the terminal ribose of tRNA (class II aaRSs aminoacylate the 3′OH). With the availability of the three-dimensional structure of the complex between multisubunit PheRS and $tRNA^{Phe}$, a fuller picture of the specific tRNA–aaRS interactions is beginning to emerge.

Results: The crystal structure of *Thermus thermophilus* PheRS complexed with cognate tRNA has been solved at 3.28 Å resolution. It reveals that one $tRNA^{Phe}$ molecule binds across all four PheRS subunits. The interactions of PheRS with tRNA stabilize the flexible N-terminal part of the α subunit, which appeared to form the enzyme's 11th domain, comprising a coiled-coil structure (helical arm) built up of two long antiparallel α helices. The helical arms are similar to those observed in SerRS and are in the same relative orientation with respect to the catalytic domain. Anticodon recognition upon tRNA binding is performed by the B8 domain, the structure of which is similar to that of the RNA-binding domain (RBD) of the small spliceosomal protein U1A. The *Th. thermophilus* PheRS approaches the anticodon loop from the minor groove side.

Conclusions: The mode of interaction with tRNA explains the absolute necessity for the $(\alpha\beta)_2$ architecture of PheRS. The interactions of $tRNA^{Phe}$ with PheRS and particularly with the coiled-coil domain of the α subunit result in conformational changes in TΨC and D loops seen by comparison with uncomplexed yeast $tRNA^{Phe}$. The $tRNA^{Phe}$ is a newly recognized type of RNA molecule specifically interacting with the RBD fold. In addition, a new type of anticodon-binding domain emerges in the aaRS family. The uniqueness of PheRS in charging 2′OH of tRNA is dictated by the size of its adenine-binding pocket and by the local conformation of the tRNA's CCA end.

Addresses: [1]Department of Structural Biology, Weizmann Institute of Science, 76100, Rehovot, Israel, [2]Institute of Molecular Biology, 117984 Moscow, Russia and [3]Institute of Bioorganic Chemistry, Novosibirsk 690390, Russia.

*Corresponding author.
E-mail: safro@sgms2.weizmann.ac.il

Key words: complex, phenylalanyl-tRNA synthetase, recognition, $tRNA^{Phe}$, X-ray crystallography

Received: 11 September 1996
Revisions requested: 7 October 1996
Revisions received: 24 October 1996
Accepted: 31 October 1996

Electronic identifier: 0969-2126-005-00059

Structure 15 January 1997, 5:59–68

Introduction

Aminoacyl-tRNA synthetases (aaRSs), which belong to the oldest group of proteins, are key enzymes of protein biosynthesis and are of primary importance in the transformation of the genetic information into polypeptide chains [1]. Despite the great diversity of aaRSs in amino acid sequence and quaternary organization, they have been integrated into two classes, each consisting of 10 enzymes [2]. This classification is based on the existence of two different catalytic folds [3] and appears to correlate to significant differences in tRNA recognition and aminoacylation. The crystal structures of class I GlnRS [4,5] and class II AspRS [6] complexed with their cognate tRNAs show highly specific enzyme–anticodon recognition. The property to charge (aminoacylate) the tRNA's terminal ribose on 2′OH by class I aaRSs or on 3′OH by class II was believed to be a consequence of two different modes of acceptor-stem binding [6] and the location of the terminal adenosines of $tRNA^{Gln}$ and $tRNA^{Asp}$ on the opposite sides of the adenylate [7]. Detailed crystallographic results on the class II SerRS–$tRNA^{Ser}$ complex [8] have confirmed that the mode of aaRS–tRNA interactions varies in different systems, even if the aaRSs belong to the same class. It was shown that the SerRS–$tRNA^{Ser}$ interaction does not include recognition of the anticodon of $tRNA^{Ser}$ and in fact the interaction of $tRNA^{Ser}$ with the two very long helices of SerRS is extremely important [8]. The coiled-coil domain, as well as cross-subunit binding of $tRNA^{Ser}$

on the SerRS dimer, was observed in the serine system only. Thus, each tRNA–aaRS complex carries individual and essential information, and promotes an understanding of how this diverse family of enzymes performs the same catalytic function with different although similar substrates, while ensuring high accuracy of tRNA charging.

The structure of the PheRS–tRNA^{Phe} complex is significant for several reasons. Firstly, PheRS from *Th. thermophilus* is an $(\alpha\beta)_2$ enzyme with 350 amino acids in the α subunit and 785 amino acids in the β subunit; its complex with tRNA is the first example of tRNA binding to a tetrameric aaRS. This complex is expected to exhibit a novel pattern of interactions with tRNA and to explain the need for this unusual oligomerization type in the aaRS family. Secondly, PheRS is the only class II enzyme that charges the 2′OH group of the terminal ribose of its cognate tRNA. More information on the complex is obviously needed to resolve this paradox. Thirdly, the crystal structure of PheRS from *Th. thermophilus*, recently determined at 2.9 Å resolution, highlighted 10 structural domains which clustered into four modules: catalytic — CAM (A1–A2, α-subunit), N-terminal — NTM (B1–B5, β-subunit), 'catalytic-like' — CLM (B6–B7, β-subunit) and C-terminal — CTM (B8, β-subunit) [9]. Only CAM could be solely attributed to the domains directly involved in the binding and charging of cognate tRNA. Two further domains, B2 and B8, were considered as alternative candidates for anticodon binding. Having only the native structure at hand, however, the function of other PheRS-specific domains as well as that of the disordered N-terminal part of the α-subunit remained unknown.

We report here the crystal structure of a complex between PheRS and tRNA^{Phe}, both from *Th. thermophilus*, determined at 3.28 Å resolution.

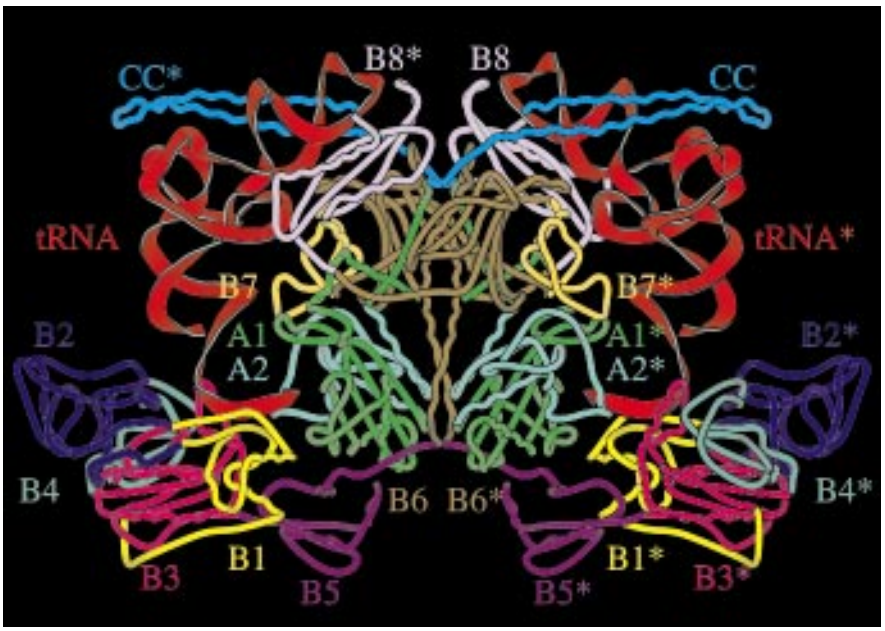
Results

Architecture of the complex

PheRS binds two tRNA molecules. Each tRNA binds across all four subunits of the enzyme (Fig. 1). The total contact area of the tRNA substrate is about 2700 Å^2. This constitutes ~23% of the solvent accessible area of *Th. thermophilus* tRNA^{Phe}, which is somewhat more than in aspartyl and glutaminyl systems. The *Th. thermophilus* tRNA^{Phe} regions that interact with the synthetase are presented at the cloverleaf diagram in Figure 2a. The acceptor end and the stem of the tRNA molecule interact with the active site of the α subunit and with the B1 domain from the same heterodimer. CLM approaches the tRNA from the D-loop side. The remaining protein–tRNA contacts occur with the other heterodimer (its structural elements are marked by an asterisk): B8* interacts with the anticodon stem-loop and the N-terminal domain of the α* subunit approaches the tRNA mainly from the variable loop side. Contrary to our expectations, domain B2, which has the topology of the anticodon-binding domain of AspRS, does not interact with tRNA^{Phe} at all and its function has yet to be determined.

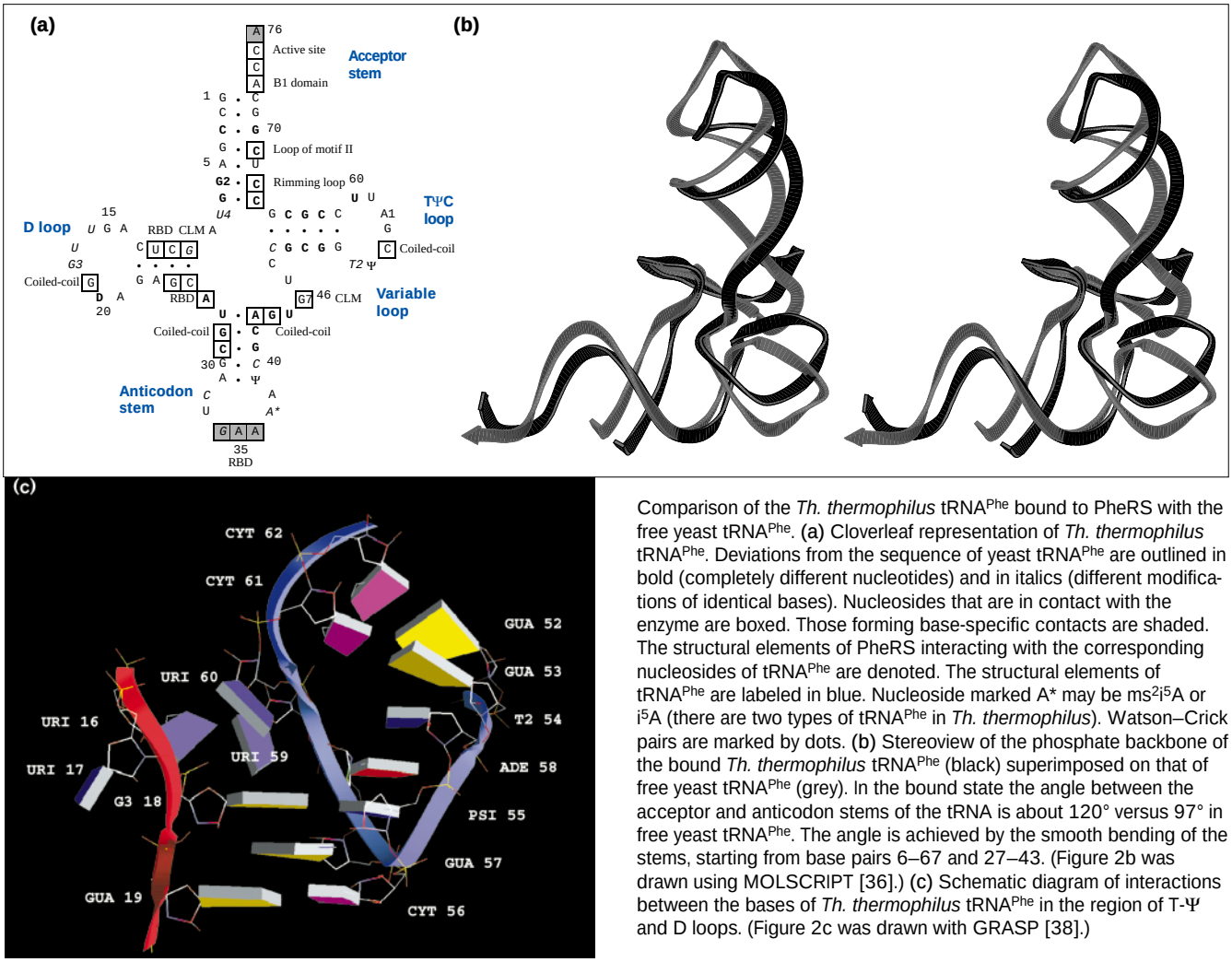
Upon formation of PheRS–tRNA^{Phe} complex, both the enzyme and the substrate undergo conformational changes that conceivably allow better complementarity of interacting surfaces. In PheRS, changes occur only in the close vicinity of the bound tRNA. The most pronounced difference between the two states is the ordering of the

Figure 1



General view of PheRS–tRNA^{Phe} complex structure. The intramolecular twofold axis goes vertically in the plane of the drawing. Domains and their designating inscriptions are the same colour. 'CC' stands for the coiled-coil domain of the α subunit. The structural elements of the second heterodimer are marked with an asterisk. (Figure was drawn using MOLSCRIPT [36].)

Figure 2



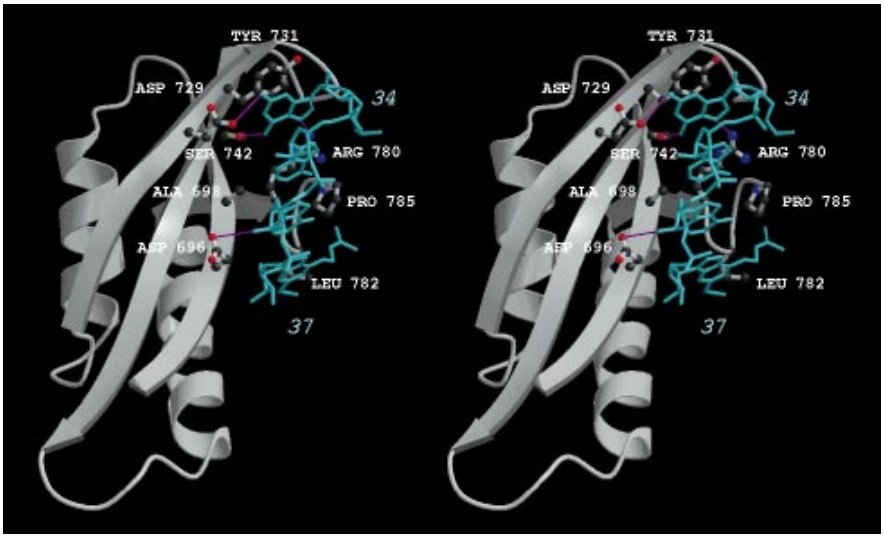
Comparison of the *Th. thermophilus* tRNA^{Phe} bound to PheRS with the free yeast tRNA^{Phe}. **(a)** Cloverleaf representation of *Th. thermophilus* tRNA^{Phe}. Deviations from the sequence of yeast tRNA^{Phe} are outlined in bold (completely different nucleotides) and in italics (different modifications of identical bases). Nucleosides that are in contact with the enzyme are boxed. Those forming base-specific contacts are shaded. The structural elements of PheRS interacting with the corresponding nucleosides of tRNA^{Phe} are denoted. The structural elements of tRNA^{Phe} are labeled in blue. Nucleoside marked A* may be ms^{2}i^{5}A or i^{5}A (there are two types of tRNA^{Phe} in *Th. thermophilus*). Watson–Crick pairs are marked by dots. **(b)** Stereoview of the phosphate backbone of the bound *Th. thermophilus* tRNA^{Phe} (black) superimposed on that of free yeast tRNA^{Phe} (grey). In the bound state the angle between the acceptor and anticodon stems of the tRNA is about 120° versus 97° in free yeast tRNA^{Phe}. The angle is achieved by the smooth bending of the stems, starting from base pairs 6–67 and 27–43. (Figure 2b was drawn using MOLSCRIPT [36].) **(c)** Schematic diagram of interactions between the bases of *Th. thermophilus* tRNA^{Phe} in the region of T-Ψ and D loops. (Figure 2c was drawn with GRASP [38].)

N-terminal region of the α subunit, which appeared to comprise a coiled-coil structure (helical arm). The helices stretch out into the solvent by 65 Å and are apparently flexible in the absence of intermolecular contacts.

Upon binding tRNA, a rearrangement is observed in the synthetase area of residues β780–β785, which belong to part of the fourth strand of the B8 domain (the anticodon-binding domain) in free PheRS. These residues embrace the anticodon loop of tRNA in the complex, forming no regular secondary structure.

For lack of crystallographic data on free tRNA^{Phe} from *Th. thermophilus*, we compared the structure of the bound tRNA with its homologous free yeast tRNA^{Phe} (80% identity, see Fig. 2a) [10,11]. *Th. thermophilus* tRNA^{Phe} displays a standard 'L-shape' conformation, but differs substantially from free yeast tRNA^{Phe} in the stems and loops: the root mean square (rms) deviation for the phosphate atom coordinates in the core (as defined in [8]) is ~2 Å, in contrast to 4.5 Å for the whole molecule (Fig. 2b). The interactions of Ψ55, C56 and G19 with the coiled-coil domain result in essential conformational changes of the TΨC and the D loops of tRNA (Fig. 2c). Further deviations from the structure of free tRNA^{Phe} can be attributed to intermolecular contacts between molecules of tRNA in the crystal, in the area of the crystallographic twofold axis formed by G53, s^{2}T54, Ψ55 and their symmetry mates. Consequently, the D-Ψ stem-loop is bent towards the anticodon stem in such a way that the base pair 19–56, which still stabilizes the interaction between the TΨC and D loops, comes closer to the core. The *trans*-Hoogsteen pair s^{2}T54–m′A58 is broken and instead m′A58 intercalates between s^{2}T54 and Ψ55. The stacking interaction between bases 58–54 and 58–55 presumably compensates for the free energy cost of breaking the hydrogen bond of the 54–58

Figure 3



Stereoview of the anticodon-binding domain of PheRS (grey) with the anticodon loop of $tRNA^{Phe}$ (cyan). Residues in contact with tRNA are depicted as ball-and-stick models. Hydrogen bonds are shown as thin magenta lines. (Figure was drawn using MOLSCRIPT [36] and RASTER3D [37].)

base pair. G18 makes previously observed tertiary interactions with Ψ55, but in contrast to the free yeast $tRNA^{Phe}$ does not intercalate between G57 and m′A58 and stacks only on G57. An important feature of the D loop is the 'bulged in' orientation of the unmodified nucleoside U16. Its tertiary interaction with U59 presumably stabilizes the rearranged architecture of tRNA. The close contacts between the coiled-coil domain of PheRS and the nucleosides of the tRNA variable loop 44–45 cause changes in the conformation of this loop. In particular, bases U45 and A9 are unstacked.

Anticodon recognition

One of the most remarkable features revealed by the structure of the complex is that the specific recognition of the tRNA anticodon is achieved by its interaction with the B8 domain (Fig. 3), which is similar to the RNA-binding domain (RBD) of the U1A spliceosomal protein [12]. In contrast to $tRNA^{Asp}$ and $tRNA^{Gln}$, the enzyme-bound *Th. thermophilus* $tRNA^{Phe}$ keeps the conformation of the anticodon loop relatively similar to that of free yeast tRNA. The main reason for this difference is that AspRS [13] and GlnRS [14] approach the anticodon loop from the major-groove side, and the anticodon bases have to protrude out to form base-specific contacts, whereas in PheRS such contacts exist on the minor groove side of the anticodon loop, maintaining its almost undistorted conformation. The only significant difference is that G34, being coplanar to the base of A35, is unwound outside the interior of the anticodon loop. Consequently, only slight stacking is retained between G34 and A35. Synthetase recognition of G34 is accomplished by means of stacking interaction between G34 and Tyrβ*731, as well as by two base-specific contacts. The first is a hydrogen bond between O6 of G34 and Serβ*742, which belongs to the group of amino acids identified as an RNP1 motif [15], characteristic of various RNA-binding proteins. The second base-specific interaction is between N2 of G34 and Aspβ*729. In addition, the hydrogen bond between N7 of G34 and Argβ*780 may favour purine bases in this position. All four aforementioned amino-acid residues are strictly conserved in the tetrameric species of PheRS. A van der Waals contact of Alaβ*698 with base A35 has an important role because any longer side chain (instead of the conserved Ala) would interfere with the anticodon base. Recognition of A36 can be achieved through van der Waals contact between the Cα atom of Leuβ*697 and C2 atom of this adenine. A hydrogen bond between the sugar O2′ and the Oδ atom of Aspβ*696 probably stabilizes the conformation of tRNA in this area. Formally, residues Aspβ*696, Leuβ*697 and Alaβ*698 interacting with bases A35 and A36 belong to the characteristic RNP2 motif. It is interesting that the binding modes of RNA loops with the spliceosomal protein U1A and with B8 of PheRS appear to be different. When the RBD of U1A is superimposed on B8, the RNA loops are located near the same side of the β sheet, but they are rotated with respect to each other in such a way that the loop directions (5′→3′) are opposite to each other. As determined by mutational analysis of tRNA transcripts, all the anticodon nucleotides (34–36) are involved in the recognition of $tRNA^{Phe}$ by PheRS from different sources, G34 having the most important role [16–19]. The structure of *Th. thermophilus* PheRS-$tRNA^{Phe}$ complex reveals that, among the anticodon nucleotides, G34 makes the largest number of contacts, confirming its vital importance for recognition.

Binding of the CCA end

Partial unwinding of helical conformation of the (acceptor) CCA end of the complexed *Th. thermophilus* $tRNA^{Phe}$ is

largely determined by a network of contacts between the protein and the tRNA sugar–phosphate backbone (Fig. 4). The net of interaction is constructed mostly of three salt bridges between the NH_3^+ terminus of the β subunit and C74 O2P, between Nη1 of Argβ2 and Cyt74 O1P and between Nη2 of Argβ2 and A73 O2P. Hydrogen bonds between C75 O2P and the mainchain nitrogen of Valβ160 as well as between C74 O2′ and Nη of Argβ362 further reinforce this cluster of interactions. In this unwound conformation of the tRNA^{Phe} backbone, which resembles a wide arch, bases C72 and A73 are not strictly stacked on each other, as occurs in yeast tRNA^{Phe}. The base of A73 points away from the CAM and the B1 domain, thus not participating in the contacts with the protein and agreeing with its minor contribution to the recognition set of *Th. thermophilus* [16].

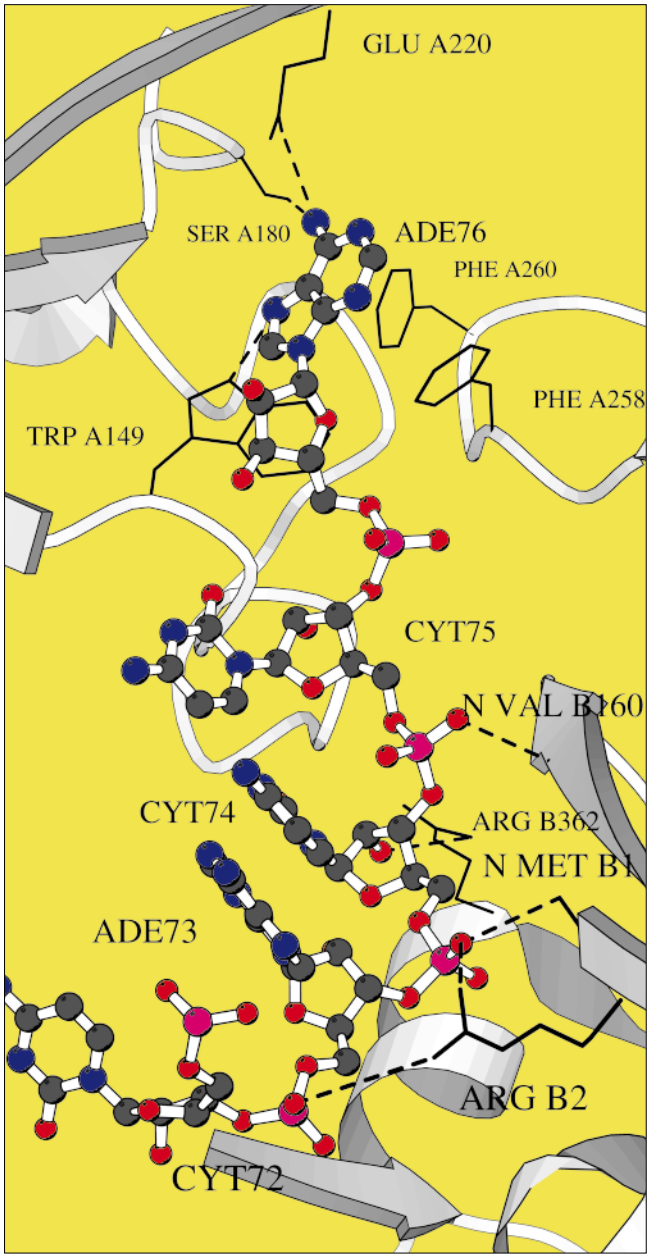
The position of terminal adenosine in the active site cavity is held by three hydrogen bonds: N6 of A76 making contacts with Serα180 and Gluα220; the indole ring of Trpα149 is approximately perpendicular to the base of A76 and makes a hydrogen bond between the Nε1 and N7 of the adenosine. Of these three residues, only Trpα149 is not conserved in the other PheRS amino-acid sequences, but His and Gln at this position in *E. coli* and *Bacillus subtilis* PheRS, respectively, are capable of participating in such an interaction. The angle between the base of A76 and the strictly conserved Pheα258 is about 60°. Such ring–ring interactions are believed to be favourable [20] and may contribute to stabilizing A76. Multiple contacts of the terminal adenosine with PheRS are consistent with the evidence that removing A76 results in a significant increase of K_d of the PheRS–tRNA^{Phe} complex [21].

Comparison with biochemical results

Using affinity labeling and tryptic cleavage, Fasiolo *et al.* [22] found that the N-terminal domain of the small subunit of yeast cytoplasmic PheRS contained important tRNA-binding sites. Affinity labeling of the *E. coli* enzyme showed that tRNA was attached to the large subunit [23]. A drastic reduction of the aminoacylation activity of *Th. thermophilus* PheRS was observed upon removal of the B8 domain (R Kreutzer personal communication). These results, which seem to be in conflict at first glance, are in agreement with our model of the tRNA^{Phe}–PheRS complex. Because the position of the tRNA is fixed by the helical arm and the B8 domain, removing either of these domains may prevent binding.

Footprinting experiments using tRNA^{Phe} transcripts from *Th. thermophilus* with the homologous PheRS [24] indicated remarkable agreement with our model. Indeed the regions of cleavage enhancement are located at the sites where bending of tRNA is likely to occur. Protected regions are found in those fragments of tRNA that are in close contact with the enzyme with one exception, at positions 20–21.

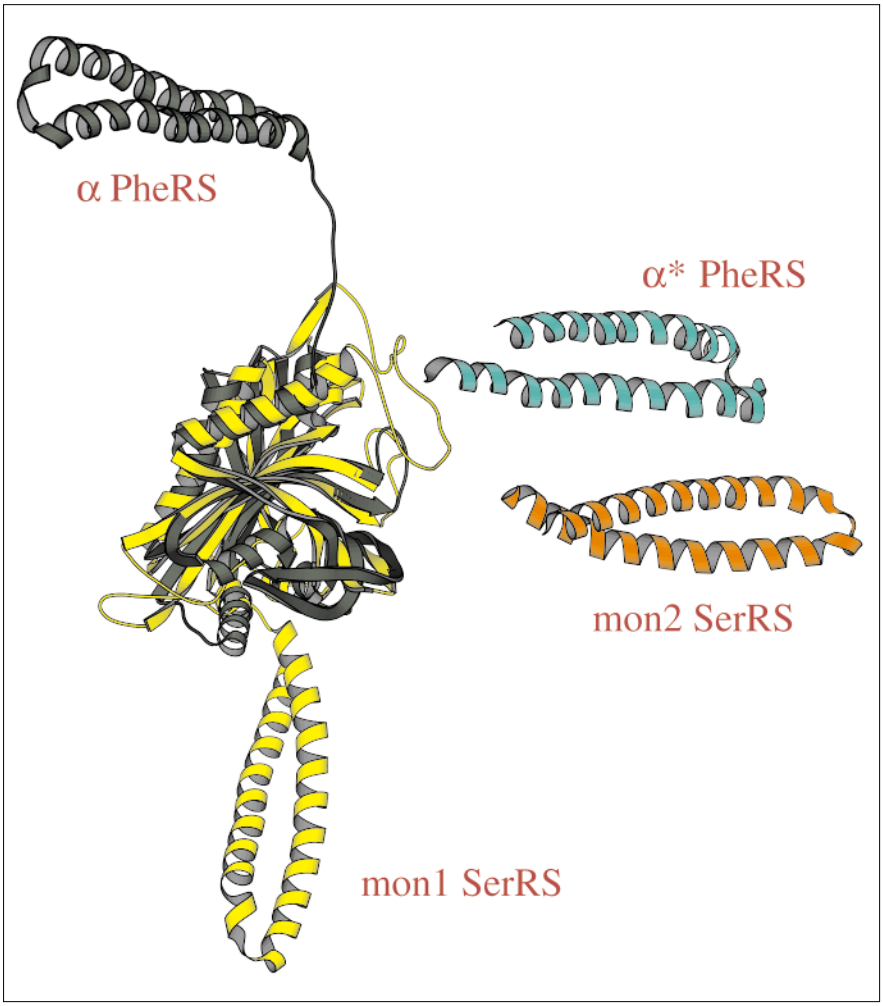
Figure 4



View of the CCA end of tRNA^{Phe} in the active-site cavity of the enzyme. The dashed lines represent the hydrogen bonds. The letters A and B before residue numbers indicate the enzyme subunits. The relatively short distance (~3.8 Å) between the N6 group of A73 (ADE73) and the phosphate group of C72 (CYT72) indicates that their intramolecular interaction may help to stabilize the conformation of the CCA end in a way that resembles tRNA^{Gln} [4]. However, the conformation of tRNA^{Phe} in this region differs from that of tRNA^{Gln}. The letters A and B before residue numbers indicate the enzyme subunits. (Figure was drawn using MOLSCRIPT [36].)

Protection in this area can be explained by the bending of the D loop, which places the base of D20 closer to the phosphate of A21 than it is in yeast tRNA^{Phe}, thus making

Figure 5



The α subunit of PheRS (gray) and a monomer of SerRS (yellow) after the superposition of their catalytic folds. The corresponding helical arms are oppositely directed. However, the orientation of the helical arms from α* subunit of PheRS and the second monomer (mon2) of SerRS with respect to the referenced catalytic folds is similar. (Figure was drawn using MOLSCRIPT [36] and RASTER3D [37].)

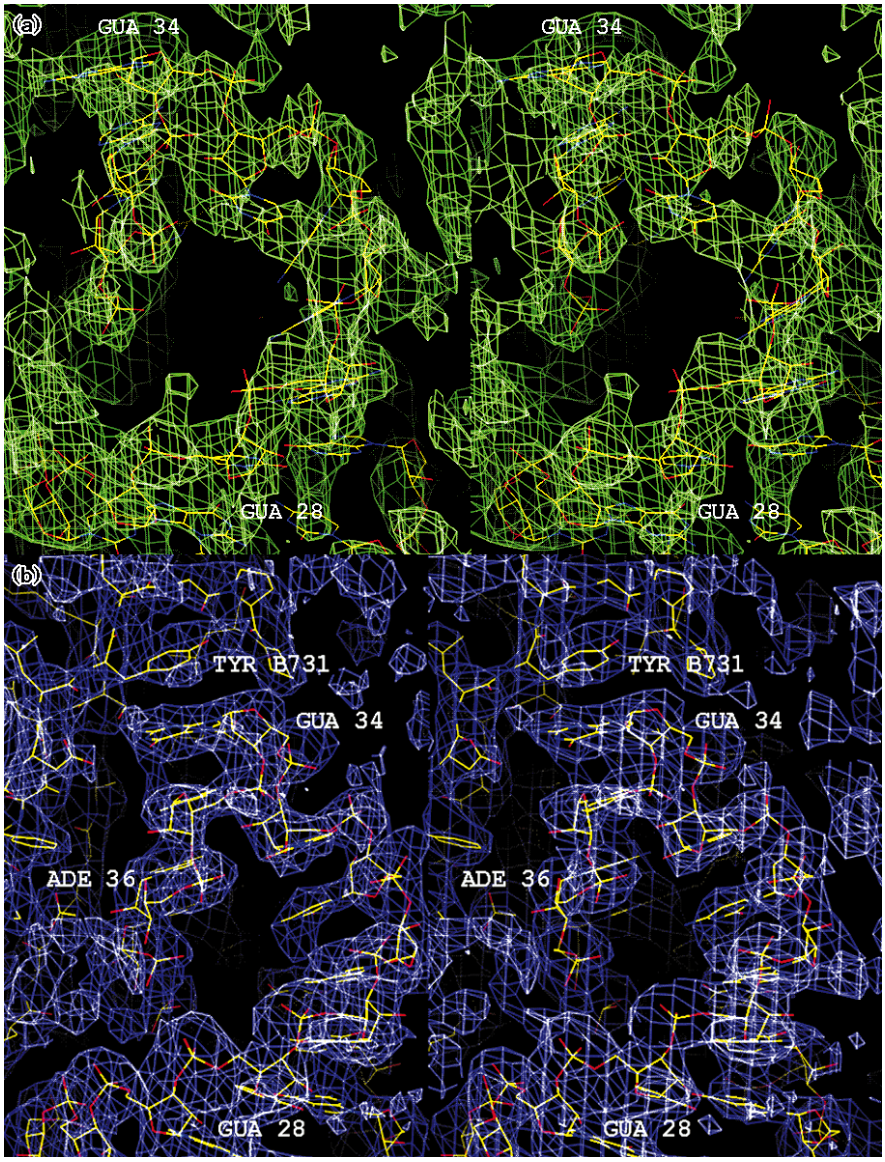
it less accessible. This agrees with the finding [16] that D20 is insignificant to the catalytic efficiency of tRNA aminoacylation and is not involved in specific contacts with *Th. thermophilus* PheRS.

Discussion

Five regions of tRNAPhe form principal contacts with PheRS: anticodon loop and stem; acceptor end and stem; variable loop; D and Ψ loops; and D-stem. There are a few base-specific interactions, but most of the contacts occur between the protein and the backbone of tRNA. On the basis of kinetic experiments for yeast $tRNA^{Ser}$–SerRS and $tRNA^{Phe}$–PheRS systems, it was determined that tRNA binding proceeds in two steps [25]. The initial bimolecular step is rapid and has a broad specificity, whereas the second unimolecular step is related to conformational changes and more precise adjustment/recognition. Non-specific interactions localized at widely spaced regions of tRNA are probably important for recognizing the general shape of the substrate and have the role of ensuring primary specificity. This bimolecular step seems imperative for the specific recognition of the anticodon and the later adjustments in the conformation of the tRNA molecule. As a result, the CCA extremity of the tRNA can be driven into the active site in the correct orientation. Since the anticodon-recognizing domain B8 makes few interactions with other domains of PheRS, it seems unlikely that information can be transferred from it to the active site. The process of stepwise proper orientation of tRNA can be considered as information transduction from anticodon to CCA end, and its correct orientation is sufficient for the aminoacylation to proceed.

Special attention should be given to the role of the helical arm in binding and recognition of the tRNA. Judging from the structure of the $tRNA^{Ser}$–SerRS complex and keeping in mind that the anticodon is not recognized by SerRS, the coiled-coil domain is a major tRNA-recognition element. The existence of the long variable loop in $tRNA^{Ser}$ was believed to be a prerequisite for the helical

Figure 6



Stereoview of the electron-density maps at the region of interaction between the anticodon loop of tRNA and B8. (a) A difference 'omit' map contoured at 1σ. The calculated amplitudes and phases are based on the model of free PheRS that was subjected to slow cool from 4000K using weak harmonic restraints on all atomic positions, as suggested by Kleywegt and Jones [39]. This map is completely unbiased by the tRNA model. (b) The final ($2F_o$–F_c) map calculated at 3.28 Å resolution and contoured at 0.7σ. Stacking of Tyrβ*731 on G34 (GUA34) is clearly seen. Atoms are shown in standard colors. (Figure 6 was drawn with the program O [34].)

arm to perform its function [8]. The phenylalanine system shows that, along with its specific recognition of the anticodon, the coiled-coil domain is also a part of the PheRS structure. This domain plays an important role in PheRS even in the absence of the long variable loop in tRNA^{Phe}. If the catalytic folds of SerRS and PheRS are superimposed, the coiled-coils are directed to the opposite sides (Fig. 5). However, because of the remarkable tetramerization mode of PheRS [26] and the existence of the flexible extended segment (residues α85–α100), the helical arm from the α* subunit is in a position to approach CAM similarly to the way the helical arm from one SerRS monomer approaches the active site domain of the other monomer. Analysis of the contacts between the coiled-coil and tRNA in both systems reveals that the general patterns of interaction have much in common. The two separated areas of contacts are the middle region of the helices, which interacts with a row of paired bases running in a perpendicular direction (along the variable loop in SerRS and the anticodon stem and short variable loop in PheRS), and the end of the helical arm, where a saddle is formed for base pair 19–56. Thus we conclude that the role of the coiled-coil domain consists of recognizing a certain structural pattern characteristic for tRNA (a stem-like structure and an exposed base pair about 15 Å away).

In the context of the relations within the aaRS family, the α subunit is part of an $(\alpha\beta)_2$ enzyme, yet at the same time represents the entire monomer of SerRS. Since CAM is a necessity for all representatives of class II, and a

coiled-coil has been identified only in SerRS and PheRS, we hypothesize that this domain is a required structural component of class II aaRSs with cross-subunit binding of tRNA. Moreover, the presence of the coiled-coil domain in SerRS and PheRS may justify classifying them as a separate group within class II.

Among the nonspecific contacts between PheRS and $tRNA^{Phe}$ those of the acceptor stem are worth mentioning. The phosphate backbone of tRNA in the region of C66 is approached by the rimming loop of CLM (residues β536–β539), but the distances indicate that the interaction is probably water mediated. The rimming loop of the non-catalytic β subunit thus performs the role of the corresponding loop from the second catalytic subunit in AspRS and SerRS, as was predicted [9]. Glnβ207, belonging to the characteristic motif *2* of class II aaRS [*2*], forms a hydrogen bond with OP*2* of C69. Motif *2* is known to participate in ATP/aminoacyl adenylate binding, as well as in interaction with the acceptor stem of tRNA. Although the ATP-binding mode is common for class II, the pattern of interactions with tRNA appeared to be system specific. Contrary to AspRS and SerRS, where residues of motif *2* loop make extended base-specific contacts, in PheRS the role of motif *2* in acceptor stem binding is restricted to forming a nonspecific hydrogen bond. In this respect it is not a surprise that the conformation of the loop does not change upon tRNA binding.

The structure of the PheRS–$tRNA^{Phe}$ complex offers a clearer view of the functional peculiarity of PheRS. Superposition of the catalytic domains of the AspRS–$tRNA^{Asp}$ and SerRS–$tRNA^{Ser}$ complexes on that of PheRS shows that the overall orientation of the tRNA is about the same. From the known structures of AspRS and SerRS complexed with ATP or aminoacyl adenylate analogs [7,27], we can clearly delineate the position of the AMP moiety in the active site of PheRS because it is characterized by interactions with strictly conserved residues and closely similar topology of catalytic domains. Using the common position of AMP as a reference, we can compare the orientation of terminal adenosines of tRNA, which appears to be the only factor that causes the difference in the primary site of aminoacylation. The structure of the complex shows that the terminal adenosine is likely to approach the adenylate moiety from the same side as in the aspartyl system. However, A76 of $tRNA^{Phe}$ penetrates less deeply into the active site cavity and is rotated in such a way that the two sugar moieties are approximately perpendicular. Consequently, the 2′OH of A76 occupies the more preferable position for a nucleophilic attack on aminoacyl adenylate, since it is located between the α phosphate of AMP and 3′OH. To bring the 3′OH into a position favourable for aminoacylation, A76 would have to be displaced significantly deeper into the active-site cavity, which is unlikely considering the size of the binding pocket and the bulkiness of the surrounding sidechains (Pheα258, Pheα260, Trpα149 and Gluα220). The hydrogen bond between Gluα220 and 3′ terminal adenosine partially interferes with the invariant network of interactions of ATP. Thus, ATP binding is likely to be accompanied by repositioning of the terminal adenosine. Nevertheless, whatever alterations occur in the position of A76 in the presence of ATP, the size of the adenine-binding pocket will not permit the displacement required to bring the 3′OH into a position favourable for aminoacylation.

Biological implications

The fidelity of the translation process is governed to a large extent by the recognition of both the amino acid and the tRNA by aminoacyl-tRNA synthetases (aaRS). Much work has been done to understand the structural basis of this specificity. Binary complexes of two class II synthetases and one class I synthetase with tRNA revealed striking differences in the tRNA binding mode between the two classes. Moreover, the mode of aaRS–tRNA specific interactions varies within a class and, as became evident recently, is dictated by global characteristics of the system such as the presence or absence of the long variable loop of tRNA and the domain composition of the aaRS.

The structure of *Th. thermophilus* phenylalanyl-tRNA synthetase (PheRS) complexed with $tRNA^{Phe}$, which has been determined at 3.28 Å resolution, is the first example of tRNA binding to an aaRS molecule with $(\alpha\beta)_2$ subunit organization. The PheRS–$tRNA^{Phe}$ complex is the second system with true cross-subunit tRNA binding (along with the α_2 SerRS). We show that one $tRNA^{Phe}$ molecule interacts with all four subunits of the enzyme. Thus, the structure of native PheRS led us to the conclusion that both α and β monomers are directly involved in the active site formation, whereas the structure of the complex explains why the enzyme has to be a functional $(\alpha\beta)_2$ dimer.

The anticodon loop of the $tRNA^{Phe}$ is specifically recognized by the C-terminal domain of the large β subunit, which is closely similar to the RNA-binding domain of the U1A spliceosomal protein. Unexpectedly, domain B2, which has the topology of the anticodon-binding domain of AspRS and which was considered a strong candidate for performing analogous functions in PheRS, is not involved in interactions with tRNA at all. Remarkably, *Th. thermophilus* PheRS approaches the anticodon loop from the minor groove side. This mode of interaction results in only slightly distorted conformation of anticodon loop as compared with free $tRNA^{Phe}$.

The N-terminal domain of the PheRS α subunit bound with the cognate tRNA manifests a coiled-coil (helical arm) analogous to that in SerRS. Both systems (in spite

of the different subunit organization and the absence of the long variable loop in tRNA^{Phe}) reveal a common pattern of interactions between the coiled-coil and tRNA. This domain functions in the recognition of a stem-like structure and an exposed base pair (19–56) 15 Å away from each other, a structural motif of tRNA. The helical arm is expected to be a necessary structural component of aaRSs with cross-subunit binding of tRNA.

The functional peculiarity of class II PheRS to charge 2′OH of terminal ribose (that is typical for class I aaRSs) is a result of the local orientation of the tRNA CCA end and the size of the A76-binding pocket of the synthetase.

Materials and methods

Crystal preparation and data collection

The crystals of the PheRS–tRNA^{Phe} complex were obtained as described before [28]. They are of the same space group $P3_221$, and were virtually the same unit cell parameters (a = b = 175 Å, c = 140.6 Å) as the crystals of the free enzyme [29]. Data collection was performed using the synchrotron radiation source at Lure (Orsay, France) on a MAR Research image plate detector (completeness = 45.9% and R_{sym} = 10.6% to 3.28 Å resolution) and at the Weizmann Institute of Science (Rehovot, Israel) on a Rigaku RU-300 rotating anode generator with Xentronics area detector (completeness = 70.8% and R_{sym} = 12.3% to 3.54 Å resolution). Data were processed with DENZO [30] and XENGEN [31] packages, respectively. Both data sets were merged with R_{merge} of 15.8% to give an 86.3% full data set of 33 340 reflections in the resolution range 28–3.28 Å (completeness in the highest resolution shell from 3.4 to 3.3 Å is 48.0%, with R_{sym} = 33.8%). The crystals of the complex exhibit high degree of isomorphism with the crystals of native PheRS (R_{iso} = 19%).

Model building and refinement

The model of PheRS was used for phasing the data. Difference Fourier synthesis with experimental amplitudes (F_{comp}–F_{native}) showed density for about half of the tRNA molecule. Two long antiparallel α helices could be fitted into the characteristic density that appeared immediately adjacent to Argα85. Difference maps calculated with PheRS model amplitudes using bulk solvent correction [32] provided additional information (Fig 6a). The partial models of tRNA^{Phe} and of the helical arm were refined together with the parent PheRS model using positional refinement protocol of X-PLOR [33] and bulk solvent correction. Several cycles of refinement alternating with model rebuilding and refitting (using program O [34]) into the maps calculated with Fourier coefficients derived from SIGMAA [35] were performed. Despite the fact that all nucleotides of the tRNA molecule and most of the coiled-coil domain are clearly represented in the maps, attempts to refine them with reasonable temperature factors failed. Therefore, in the final stage, the group occupancy parameters for the nucleotides of tRNA and the amino acids of coiled-coil domain were refined, keeping the B factors constant (80 Å^2). The occupancies appear to be in the range of 0.6–0.9. The weak scattering from tRNA and coiled-coil may be explained by the fact that because the unit cell parameters and the crystallization conditions for the complex and the free enzyme are closely similar, and the crystallization mixture contains both bound and unbound protein at equilibrium concentrations, not all tRNA binding sites in the crystal are occupied. The final model contains 76 nucleotides of tRNA^{Phe} *Th. thermophilus* and 1135 amino-acid residues of the enzyme. The electron density in the region of the first helix and the turn between the two helices of the helical arm (residues α1–α45) is relatively poor because of partial disorder. Therefore, a polyalanine chain represents this region in the model. The crystallographic R factor is 22.1% (R_{free} = 28.7%) for all data. The rms deviations from ideal values for bond lengths and angles are 0.02 Å and 2.3 Å, respectively, for the protein and 0.025 Å and 2.3°, respectively, for the tRNA. φ and ψ angles of 79.1% of protein residues lie in the most favoured regions of the Ramachandran plot. Figure 6b shows a sample of the final ($2F_o$–F_c) electron-density map.

Accession numbers

Coordinates will be deposited in the Brookhaven Protein Data Bank. Whilst these are being processed they may be obtained via electronic mail from safro@sgms2.weizmann.ac.il.

Acknowledgements

We thank A Yonath for reading the manuscript, F Frolow for help during data collection, D Kostrewa for valuable advice. This work was supported by grant from The Israel Academy of Sciences and Humanities and by the Kimmelman Center for Biomolecular Structure and Assembly.

References

1. Meinnel, T., Meshulam, Y. & Blanquet, S. (1995). Aminoacyl-tRNA synthetases: occurrence, structure and function. In *tRNA structure, biosynthesis, function*. (Soll, D. & RajBhandary, U.L., eds.), pp. 251–292, ASM Press, Washington DC, USA.
2. 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.
3. Cusack, S. (1995). Eleven down and nine to go. *Nat. Struct. Biol.* **2**, 824–831.
4. Rould, M., Perona, J., Soll, D. & Steitz, T. (1989). Structure of *E. coli* glutaminyl-tRNA synthetase complexed with tRNA^{Gln} and ATP at 2.8 Å resolution. *Science* **246**, 1135–1142.
5. Perona, J., Rould, M. & Steitz, T. (1993). Structural basis for transfer RNA aminoacylation by *Escherichia coli* glutaminyl-tRNA synthetase. *Biochemistry* **32**, 8758–8771.
6. Ruff, M., *et al*., & Moras, D. (1991). Class II aminoacyl-tRNA synthetases: crystal structure of yeast aspartyl-tRNA synthetase complexed with tRNA^{Asp}. *Science* **252**, 1682–1689.
7. Cavarelli, J., *et al*., & Moras, D. (1994). The active site of yeast aspartyl-tRNA synthetase: structural and functional view of aminoacylation reaction. *EMBO J.* **13**, 327–337.
8. 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.
9. Mosyak, L., Reshetnikova, L., Goldgur, Y., Delarue, M. & Safro, M. (1995). Structure of phenylalanyl-tRNA synthetase from *Thermus thermophilus. Nat. Struct. Biol.* **2**, 537–547.
10. Sussman, J., Holbrook, S., Warrant, R., Church, G. & Kim, S.-H. (1978). Crystal structure of yeast phenylalanine-tRNA. *J. Mol. Biol.* **123**, 607–630.
11. Suddath, F., Quigley, G., McPherson, A., Sneden, D., Kim, S. & Rich, A. (1974). Three-dimensional structure of yeast phenylalanine transfer RNA at 3.0 Å resolution. *Nature* **248**, 20–24.
12. Oubridge, C., Ito, N., Evans, P.R., Teo, C.-H. & Nagai, K. (1994). Crystal structure at 1.92 Å resolution of the RNA-binding domain of the U1A spliceosomal protein complexed with an RNA hairpin. *Nature* **372**, 432–438.
13. Cavarelli, J., Rees, B., Ruff, M., Thierry, J.-C. & Moras, D. (1993). Yeast tRNA^{Asp} recognition by its cognate class II aminoacyl-tRNA synthetase. *Nature* **362**, 181–184.
14. Rould, M., Perona, J. & Steitz, T. (1991). Structural basis of anticodon recognition by glutaminyl-tRNA synthetase. *Nature* **352**, 213–218.
15. Burd, C.G. & Dreyfuss, G. (1994). Conserved structures and diversity of functions of RNA-binding proteins. *Science* **265**, 615–621.
16. Moor, N., Ankilova, V. & Lavrik, O. (1995). Recognition of tRNA^{Phe} by phenylalanyl-tRNA synthetase from *Thermus thermophilus. Eur. J. Biochem.* **234**, 897–902.
17. Pallanck, L., Pak, M. & Schulman, L. (1995). tRNA discrimination in aminoacylation. In *tRNA structure, biosynthesis, function*. (Soll, D. & RajBhandary, U.L., eds), pp. 371–394, ASM Press, Washington DC, USA.
18. Tinkle-Peterson, E. & Uhlenbeck, O. (1992). Determination of recognition nucleotides for *Escherichia coli* phenylalanyl-tRNA synthetase. *Biochemistry* **31**, 10380–10389.
19. Nazarenko, I., Peterson, E., Zakharova, O., Lavrik, O. & Uhlenbeck, O. (1992). Recognition nucleotides for human phenylalanyl-tRNA synthetase. *Nucleic Acids Res.* **20**, 475–478.
20. Burley, S. & Petsko, G. (1985). Aromatic-aromatic interaction: a mechanism of protein structure stabilization. *Science* **229**, 23–28.

21. Moor, N., Repkova, M., Yamkovoy, V. & Lavrik, O. (1994). Alterations at the 3′-CCA end of *Escherichia coli* and *Thermus thermophilus* $tRNA^{Phe}$ do not abolish their acceptor activity. *FEBS Lett.* 351, 241–242.
22. Fasiolo, F., Sanni, A., Potier, S., Ebel, J.-P., Boulanger, Y. (1989). Identification of the major $tRNA^{Phe}$ binding domain in the tetrameric structure of cytoplasmic phenylalanyl-tRNA synthetase from baker's yeast. *FEBS Lett.* 242, 351–356.
23. Khodyreva, S., Moor, N., Ankilova, V. & Lavrik, O. (1985). Phenylalanyl-tRNA synthetase from *E. coli* MRE-600: analysis of the active site distribution on the enzyme subunits by affinity labelling. *Biochem. Biophys. Acta* 830, 206–212.
24. Kreutzer, R., Kern, D., Giege, R. & Rudinger, J. (1995). Footprinting of $tRNA^{Phe}$ transcripts from *Thermus thermophilus* HB8 with the homologous phenylalanyl-tRNA synthetase reveals a novel mode of interaction. *Nucleic Acids Res.* 23, 4598–4602.
25. Krauss, G., Reisner, D., Maass, G. (1976). Mechanism of discrimination between cognate and non-cognate tRNAs by phenylalanyl-tRNA synthetase from yeast. *Eur. J. Biochem.* 68, 81–93.
26. Mosyak, L. & Safro, M. (1993). Phenylalanyl-tRNA synthetase from *T. thermophilus* has four antiparallel folds of which only two are catalytically active. *Biochimie* 75, 1091–1098.
27. Belrhari, 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.
28. Reshetnikova, L., Khodyreva, S., Lavrik, O., Ankilova, V., Frolow, F. & Safro, M. (1993). Crystals of phenylalanyl-tRNA synthetase from *T. thermophilus* HB8 complexed with $tRNA^{Phe}$. *J. Mol. Biol.* 231, 927–929.
29. Chernaya, M., Reshetnikova, L., Korolev, S. & Safro M. (1987). Preliminary crystallographic study of the phenylalanyl-tRNA synthetase from *T. thermophilus* HB8. *J. Mol. Biol.* 198, 555–556.
30. Otwinowski, Z. (1992). Oscillation data reduction program. In *Data collection and processing*. (Sawyer, L., Isaacs, N. & Bailey, S., eds), pp. 56–62, Science and Engineering Research Council, UK.
31. Howard, A., Gilliland, G., Finzel, B., Poulos, T., Olhendorf, D. & Salemme, F. (1987). The use of an imaging proportional counter in macromolecular crystallography. *J. Appl. Cryst.* 20, 383–387.
32. Kostrewa, D. & Winkler, F.K. (1995). Mg^{2+} binding to the active site of *Eco*RV endonuclease: a crystallographic study of complexes with substrate and product DNA at 2.0 Å resolution. *Biochemistry* 34, 683–696.
33. Brunger, A. (1993). X-PLOR version 3.1 manual. Yale University, New Haven, CT.
34. Jones, T.A., Zou, J.-Y., Cowan, S.W. & Kjeldgaard, M. (1991). Improved methods for building protein models in electron-density maps and the location of errors in these maps. *Acta Cryst. A* 47, 110–119.
35. Read, R. (1986). Improved Fourier coefficients for maps using phases from partial structures with errors. *Acta Cryst. A* 42, 140–149.
36. Kraulis, P.J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Cryst.* 24, 946–950.
37. Merrit, E. & Murphy, M. (1994). RASTER3D Version 2.0: a program for photorealistic molecular graphics. *Acta Cryst. D* 50, 869–873.
38. Nicholls, A. (1992). GRASP: graphical Representation and Analysis of Surface Properties. Dept. Biochemistry & Molecular Biophysics, Columbia University, NY, USA.
39. Kleywegt, J. & Jones, T. (1996). Good model-building and refinement practice. *Methods Enzymol.* 277, in press.