A three-dimensional structure model of the complex of glutamyl-tRNA synthetase and its cognate tRNA

Masaru Tateno[a], Osamu Nureki[a], Shun-ichi Sekine[a], Kazuaki Kaneda[b], Mitiko Gō[b], Shigeyuki Yokoyama[a,*]

[a]*Department of Biophysics and Biochemistry, School of Science, University of Tokyo, Tokyo 113, Japan*
[b]*Department of Biology, Faculty of Science, Nagoya University, Nagoya, 464-01, Japan*

Received 18 October 1995

Abstract A docking model of glutamyl-tRNA synthetase (GluRS) and $tRNA^{Glu}$ was constructed, on the basis of the distinguished similarity between the X-ray crystallographic three-dimensional structures of the N-terminal halves of the *Thermus thermophilus* GluRS in the free state and the *Escherichia coli* glutaminyl-tRNA synthetase in a complex with $tRNA^{Gln}$. The modeled structure is energetically favorable and is also well consistent with the results of site-directed mutagenesis studies. The model indicates that the GluRS-specific insertions 2 and 3 fit and bind to the acceptor stem and the D arm, respectively, of the cognate tRNA without affecting other contacts. In particular, insertion 3 strongly interacts with the two D-stem base pairs that are essential for the tRNA·GluRS recognition.

Key words: Aminoacyl-tRNA synthetase; Computer modeling; Molecular mechanics; Molecular superposition

1. Introduction

Aminoacyl-tRNA synthetases (aaRS's) strictly recognize and ligate the specific tRNA('s) and amino acids. The specificity of a tRNA toward an amino acid (the tRNA identity) depends on the recognition of characteristic nucleotides (identity determinants) in the tRNA molecule by aaRS's [1–6]. Based on the sequences and the three-dimensional (3D) structures of the catalytic domains, the aaRS's are divided into two classes, each with ten synthetases [7]. In general, every class I synthetase has catalytic domains containing the classical nucleotide-binding (Rossmann) fold [8].

As for class I, *Escherichia coli* glutaminyl-tRNA synthetase (GlnRS) is the only one for which a crystal structure has been determined for the complex with the cognate tRNA [9,10]. Recently, we have solved the 3D structure of another class I synthetase, glutamyl-tRNA synthetase (GluRS) from *Thermus thermophilus* [11]. While it had been suggested that GluRS is closely related to GlnRS, at the level of their overall primary structures, we found a striking similarity in the 3D structures of the N-terminal halves (catalytic domains) of GluRS and GlnRS [11]. Simply by substituting GluRS for GlnRS in the complex with $tRNA^{Gln}$ (the coordinates of the phosphates only were available in this analysis), we suggested that the $tRNA^{Gln}$ structure fits quite well in the GluRS structure, and furthermore, we found that the specifically inserted secondary structural elements of GluRS are likely to have access to the bound tRNA [11]. Actually, our site-directed mutagenesis analyses demonstrated that some amino acid residues in the insertions are important for the aminoacylation of $tRNA^{Glu}$ by GluRS [11].

However, in the preliminary docking model of GluRS and $tRNA^{Gln}$ (including the coordinates of all the atoms), we detected a steric clash between the two molecules when the energy of the complex was calculated. Therefore, it was still necessary to clarify whether the insertions of GluRS could actually contact $tRNA^{Glu}$ without any energetic discrepancy. In this study, we constructed a much more sophisticated model of the GluRS·$tRNA^{Glu}$ complex, by automatic molecular superposition and computer modeling coupled with molecular mechanics calculation. This in fact allowed us to examine how well the GluRS-specific insertions contact the identity-determining base pairs of $tRNA^{Glu}$.

2. Materials and methods

2.1. Molecular modeling

Molecular mechanics calculations were performed using the united atom force field of AMBER 4.0 [12]. A distance-dependent dielectric function ($\varepsilon = 4r$) was used to approximate the solvent screening effect in the energy calculations. A 10-Å cutoff distance for non-bonded interactions was used. Energy minimizations were carried out by a conjugate gradient method until the gradient of the RMS force reached 0.1 kcal·mol^{-1}·$Å^{-2}$. These calculations were performed on a DEC3000/700 computer. The molecular display and the molecular modeling for the structural modifications were carried out on a Silicon Graphics Indigo^2 workstation using the Insight II molecular graphics program [13].

2.2. Construction of the initial structure of the E. coli $tRNA^{Glu}$ moiety

For the template of the acceptor arm of *Escherichia coli* $tRNA^{Glu}$, the atomic coordinates of *Escherichia coli* GlnRS-bound $tRNA^{Gln}$ [10], a gift from Dr. Steitz to M.G., were used. The coordinates of the anticodon arm of yeast $tRNA^{Phe}$ [14], and those of the D- and T-arms, and the variable loop of yeast $tRNA^{Asp}$ [15] were obtained form the Brookhaven Protein Data Bank [16].

The nucleotide sequence of each part was changed to that of *E. coli* $tRNA^{Glu}$ as follows. First, the coordinates of the phosphate moieties, the ribose moieties, the C4, C8, and N9 atoms of the purine residues, and the N1, C2, and C6 atoms of the pyrimidine residues were taken from the templates. The coordinates of the other atoms were generated using the internal coordinates of AMBER 4.0 by the EDIT module of the program package. Next, for each part, energy minimization was performed with the MINMD module in four steps, first for the base moieties, secondly for the ribose moieties, thirdly for the backbone moieties (the phosphate moieties and the ribose C5′ and O3′ atoms) by the IBELLY option of the MINMD module, and finally for the entire structure.

The D- and T-arms and the variable loop were joined to the acceptor stem, and then to the anticodon arm. G1 was added, since the coordinates of $tRNA^{Gln}$ lack those of residue 1. Next, energy minimization

*Corresponding author. Fax: (81) (3) 5689 5609.

Abbreviations: GluRS, glutamyl-tRNA synthetase; aaRS, aminoacyl-tRNA synthetase; GlnRS, glutaminyl-tRNA synthetase.

SSDI 0014-5793(95)01295-8

was carried out around the phosphate groups at the junctions. Finally, energy minimization was performed by the four-step procedure used for the energy minimization of each part.

2.3. Superposition of the N-terminal halves of GluRS and GlnRS

The 2.5 Å refined crystal structures of GluRS [11] and GlnRS were used for the superposition (the coordinates of GlnRS were a gift from Dr. Steitz to M.G. [10]).

C and FORTRAN programs implementing algorithms by the Rossmann and Argos method [17] for the identification of corresponding atoms were developed (Kaneda et al., in preparation) and combined with the least-squares fitting routine by McLachlan [18]. Using these programs, the coordinates of GluRS were translated and rotated so as to superimpose the Cα atoms of the GluRS N-terminal half (1–322; trial structure) with those of the GlnRS N-terminal half (8–342; reference structure). In this way, GluRS was naturally docked with the modeled $tRNA^{Glu}$, since the acceptor stem of the tRNA was originally derived from that of $tRNA^{Gln}$ bound to GlnRS.

2.4. Molecular mechanics refinement of the docking model of GluRS and $tRNA^{Glu}$

In order to find the overlap of atoms between GluRS and $tRNA^{Glu}$ in the initial docking model, for each nucleotide residue, the surrounding amino acid residues were elucidated and the atomic distances between the nucleotide and the amino acid residues were calculated. When the distances were smaller than the sum of the van der Waals radii of the two atoms, energy minimization of the amino acid residues was carried out locally, in order to reduce the high-energy contact. Next, we performed energy minimization for all of the amino acid residues with at least one atom within 4.0 Å from the $tRNA^{Glu}$ molecule. This energy minimization was carried out further by using a distance 10.0 Å instead of 4.0 Å. In these four processes, the atoms of $tRNA^{Glu}$ were fixed. Then, the energy of the whole $tRNA^{Glu}$ molecule was minimized by the four-step procedure described above (see section 2.2). Finally, the energy of the entire GluRS·$tRNA^{Glu}$ complex was minimized.

3. Results and discussion

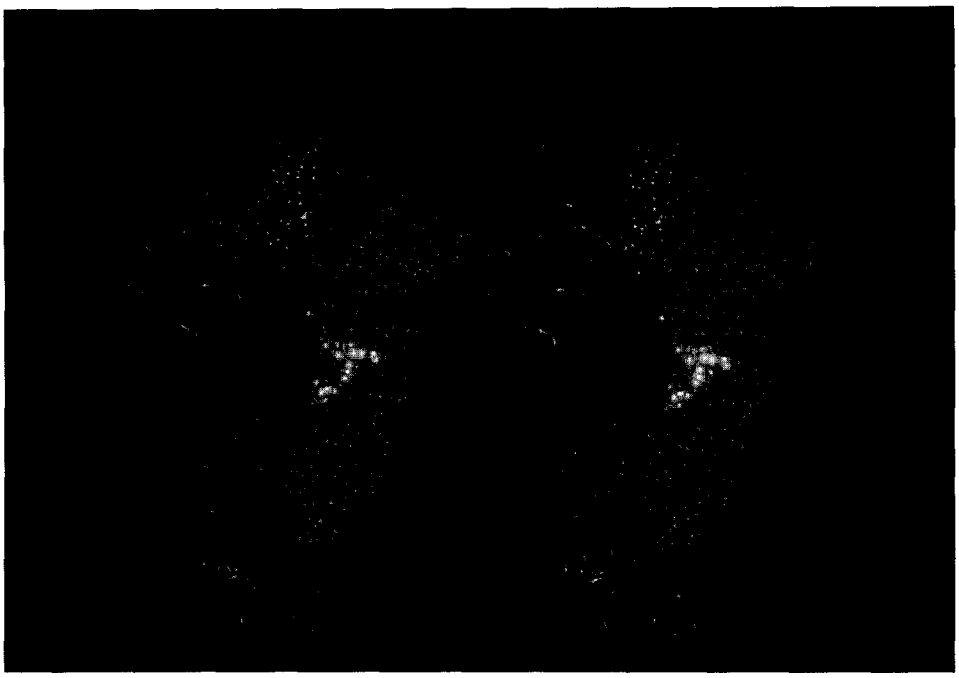
The 3D structures of the N-terminal halves of GluRS and GlnRS are so similar to each other that we could well superimpose them manually [11]. In this study, we have carried out a more precise and quantitative superposition in an automatic manner with the use of a computer program (Table 1). Thereby, GluRS was translated and rotated so as to best-fit the corresponding Cα atoms. Through this analysis, it was found that more than 80% of the Ca atoms in the GluRS N-terminal half could be superimposed with those of the GlnRS N-terminal half. Furthermore, three insertions specific to GluRS were automatically identified; the positions of these insertions in the GluRS sequence correspond well with those of the insertions found by the previous manual superposition, except for a few amino acid residues around the boundaries of insertion 2. Thus, we could definitely determine the inserted sequences specific to GluRS. Similarly, for GlnRS, the specific insertions were unambiguously identified.

Every class I synthetase has the conserved motifs 'KMSK' and 'HIGH'. These motifs were found to be superimposed between the GluRS and GlnRS molecules. In addition, about fifty amino acid residues of the GlnRS N-terminal half contact $tRNA^{Gln}$, and among these amino acid residues, more than forty were found to correspond with those of the GluRS N-terminal half by the superposition. Thus, it was shown that these functional amino acid residues are conserved between the two synthetases. This preservation of the function–structure relationship between the two synthetases naturally led us to consider that the interactions between the acceptor stem of $tRNA^{Glu}$ and the GluRS N-terminal half are also similar to those of the Gln system. The acceptor stem of the $tRNA^{Gln}$ molecule bound to GlnRS is changed to the active form, i.e. the base pair of G1 · U72 is disrupted and the CCA-extremity forms a hair-pin turn toward the catalytic site of GlnRS [9]. Our assumption is that these interactions are essentially conserved in the Glu system. Accordingly, we used the acceptor stem of the $tRNA^{Gln}$ molecule as the template for that of $tRNA^{Glu}$. The sequence of the template was exchanged according to that of $tRNA^{Glu}$, and then the structure was carefully optimized by means of several molecular mechanics steps.

However, for the other parts of $tRNA^{Glu}$, we did not use the coordinates of the GlnRS-bound $tRNA^{Gln}$ as the template in our study, for the following reasons. First, a characteristic feature is found in the variable loop of $tRNA^{Glu}$, i.e. it has four nucleotide residues, in contrast to the five found in most tRNAs, such as *E. coli* $tRNA^{Gln}$ and yeast $tRNA^{Phe}$. The number of nucleotides in the variable loop is expected to influence the 3D structure of the 'augmented D helix', which is composed of the D stem, the variable loop, and their neighboring nucleotide residues [15,19]. In addition, the conformation of the augmented D helix of $tRNA^{Glu}$ could, essentially, be conserved through interactions with GluRS. Consequently, we used the coordinates of yeast $tRNA^{Asp}$ (the free form) as the template for the D- and T-arms and the variable loop of $tRNA^{Glu}$, since $tRNA^{Asp}$ has four nucleotide residues in the variable loop [15]. Second, conformational changes of the anticodon loop of $tRNA^{Gln}$ are induced through the interaction with GlnRS [10]. Accordingly, we used the anticodon arm of yeast $tRNA^{Phe}$, which is considered to be the standard form of the various tRNAs, as the template for that of $tRNA^{Glu}$. We exchanged the sequences of the template according to that of $tRNA^{Glu}$, and then optimized their structures by performing several molecular mechanics steps. The D- and T-arms and the variable loop of $tRNA^{Glu}$ were joined to the acceptor stem, and then the anticodon arm was attached to the D arm and the variable loop. The structures around the junctions were relaxed by molecular mechanics calculation.

In this way, we constructed a starting structure of $tRNA^{Glu}$ to investigate its interactions with GluRS. Thus, we obtained the starting structure of the docking model of GluRS and $tRNA^{Glu}$ (note that the modeled acceptor stem is derived from that of the $tRNA^{Gln}$ bound to GlnRS). In order to optimize the interactions between the two molecules in the model, we examined whether there were high-energy contacts between atoms of the GluRS and $tRNA^{Glu}$ molecules. Although there was no serious steric clash, slight overlaps of atoms between the two molecules were found by the distance calculations. We eliminated the high-energy contacts by performing the molecular mechanics locally around those residues including the overlapped atoms. Finally, we carefully carried out several steps of energy minimization for relaxing the structures of the two molecules, and thus obtained the model of the GluRS·$tRNA^{Glu}$ complex (Fig. 1).

The total energy of the complex was significantly decreased (~150,000 kcal/mol) through the relaxation of the modeled structure. In particular, the van der Waals energy of the complex was decreased and became remarkably low (~−5,000 kcal/mol), indicating that the refinement reduced the overlaps and the high-energy contacts of atoms between the two molecules in the starting structure of the docking model. In order to determine which residues were affected and how largely the structures



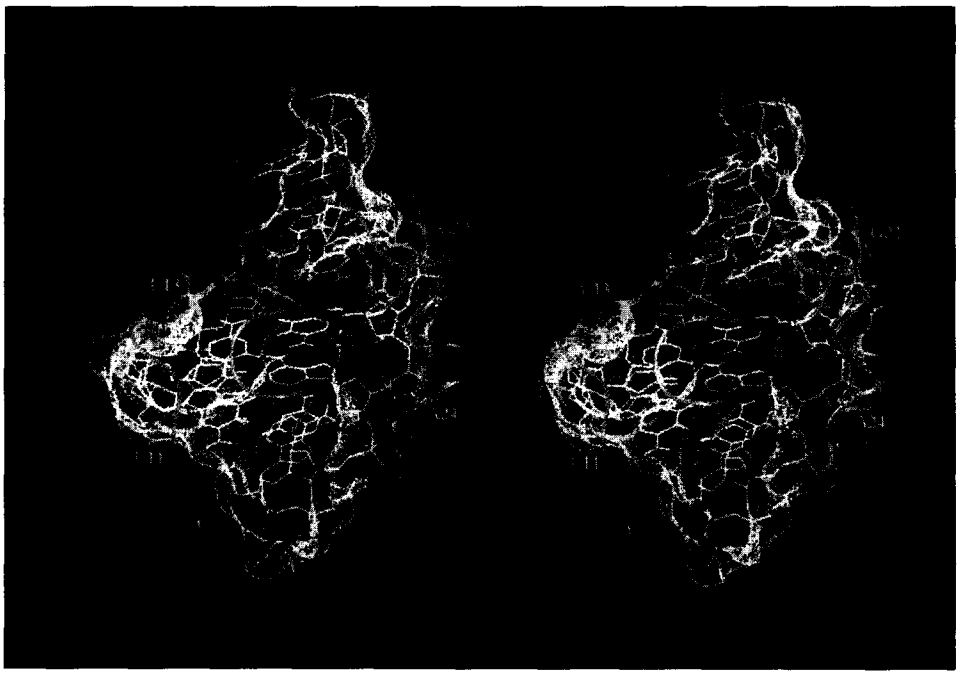


Fig. 1. (stereo view) The model of the GluRS·tRNA^{Glu} complex. GluRS and tRNA^{Glu} are represented by the trace of the Cα atoms (light green) and the CPK model (purple), respectively. Characteristic secondary structural elements specific to GluRS are shown by the CPK model in orange (insertion 2) and red (insertion 3). The molecular graphics picture was produced utilizing the Insight II program [13].

Fig. 3. (stereo view) Interactions between GluRS (represented by the Cα trace) and the determinants (colored in yellow) in the D stem of tRNA^{Glu} (represented by the the wire model on which the solvent-accessible surface is superimposed). Amino acid residues that contact the determinants are represented by ball-and-stick models in red. The molecular graphics picture was produced utilizing the Insight II program [13].

were refined, we calculated the r.m.s.d. before and after the refinement for each residue of GluRS and tRNA^{Glu} (Fig. 2). It was found that the deviations are all within 0.25 Å, and their average values are as small as 0.06 Å and 0.07 Å for GluRS and tRNA^{Glu}, respectively. In addition, the residues with r.m.s.d values larger than the average are not localized, but are scattered throughout the two molecules. These results indicate that only slight structural deviations from the starting conformations for the scattered residues were sufficient to avoid all the atomic overlaps, and the consequence was a remarkable de-

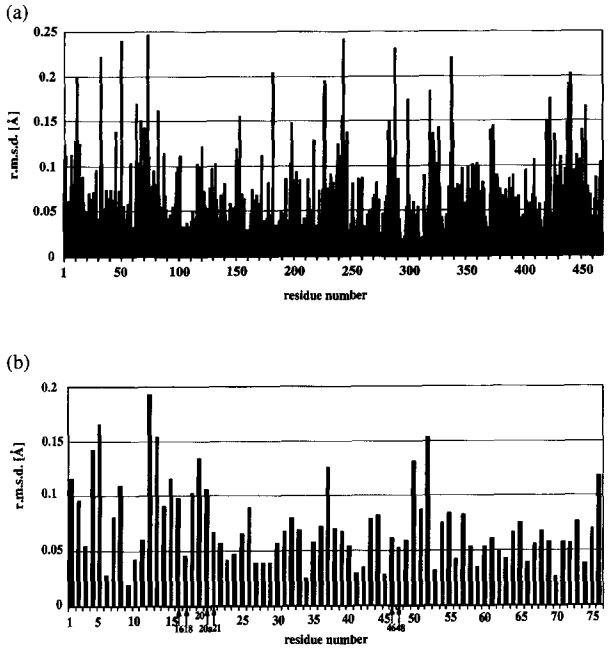


Fig. 2. The r.m.s.d. before and after the refinement in the model building is shown for each residue of GluRS (a) and $tRNA^{Glu}$ (b).

crease in the total energy. Furthermore, the molecular mechanics refinement of the docking model also significantly decreased the electrostatic energy (~3,000 kcal/mol), without sacrificing the other energy terms. Thus, the entire structure of the complex was totally relaxed and equilibrated without any appreciable changes in the conformations of the component molecules and their intermolecular geometry.

To test the plausibility of the present model structure, we examined the specific intermolecular interactions that may be conserved between the Glu and Gln systems. For example, in the X-ray structure of the GlnRS and $tRNA^{Gln}$ complex, the cytidine at position 74 (C74) of the tRNA is looped out and buried in a pocket of GlnRS, and the cytosine base forms a stack with Pro126 [9]. Therefore, it is possible that GluRS recognizes the C74 of $tRNA^{Glu}$. However, in the starting structure of the GluRS·$tRNA^{Glu}$ complex, the looped out C74 nucleotide residue of $tRNA^{Glu}$ only slightly contacted Pro109, which corresponds to Pro126 of GlnRS, according to the superposition. In contrast, in the complex structure after the present molecular mechanics refinement, the cytosine base of C74 is closely and extensively stacked onto Pro109, and therefore, is completely buried in a pocket of GluRS. This indicates that the mechanism of the CCA-extremity recognition is well conserved between GlnRS and GluRS.

We next investigated the characteristic interactions between $tRNA^{Glu}$ and the GluRS-specific insertions that were identified by the superposition. We previously suggested that insertions 2 and 3 directly contact the acceptor stem and the D arm, respectively, of the $tRNA^{Glu}$ molecule [11]. Actually, in the present model structure of the GluRS·$tRNA^{Glu}$ complex, one strand of the antiparallel β-sheet of insertion 2 is entirely bound to the minor groove side of the central five backbone moieties of one strand of the acceptor stem (Fig. 2). Thus, insertion 2 may contribute to the affinity with the tRNA through the van der Waals interaction. On the other hand, insertion 3 fits very well to the minor groove side of the D arm of $tRNA^{Glu}$ (Fig. 2). As the major identity determinants of $tRNA^{Glu}$ (U11·A24 and U13·G22) are located in the D stem [20], it is most likely that insertion 3 directly binds to the determinants and contributes to the tRNA recognition (mentioned later in detail). At the stage of our previous work, using a model of GluRS and $tRNA^{Gln}$ [11], it was impossible to examine whether the insertions specific to GluRS actually bound $tRNA^{Glu}$, since the tRNA molecule in the model was composed of only the phosphates of the $tRNA^{Gln}$ molecule as found in the complex with GlnRS. However, as mentioned above, we succeeded in the construction of an energetically favorable model with only slight conformational changes in the GluRS and $tRNA^{Glu}$ molecules. These indicate that the insertions specific to GluRS (insertions 2 and 3) can fit and bind the specific tRNA without influencing the other intermolecular contacts that are conserved between the Gln and Glu systems.

We performed a more detailed investigation of the interaction between insertion 3 and the D stem of the $tRNA^{Glu}$ molecule in the present model. Residues Ser276 and Glu282 in the insertion, and Ser299 and Lys309, which are spatially close to the insertion, directly contact the strong identity determinants, U11·A24 and U13·G22, in the D stem of $tRNA^{Glu}$ (Fig. 3). Actually, these amino acid residues were indicated to contribute to the aminoacylation activity of GluRS by site-directed mutagenesis [11]. In this context, the model is also consistent with the experimental results.

It should be noted here that the present model is not the final goal of our model building. For example, although the anticodon has some other, major determinants of $tRNA^{Glu}$ [20–23], the GluRS C-terminal domain (domain IV) is too distant from the tRNA molecule to bind in the present model. Therefore, for the anticodon recognition by GluRS, we have proposed a hy-

Table 1.
Corresponding residues revealed by the superposition of the Cα atoms of the GlnRS (reference structure) and GluRS (trial structure) N-terminal halves.

Reference structure (GlnRS)		Trial structure (GluRS)				
Residues	secondary structure	Residues	secondary structure	GluRS-specific insertions	No. of residues	r.m.s.d. [Å]
26- 95	α-B, -C; β-1, -2, -3	1- 70	α-A, -B; β-1, -2, -3	insertion 1	70	1.18
97-133	α-D, E; β-4	80-116	α-C, D; β-4		37	1.54
138-159	α-F	119-140	α-F		22	2.71
164-165	–	141-142	–		2	–
169-181	β-6	143-155	β-5	insertion 2	13	2.65
182-183	–	171-172	–		2	–
185-193	β-7	173-181	β-8		9	1.01
206-263	α-G, -H; β-8, -9, -10	182-239	α-G, -H; β-9, -10, -11		58	2.36
268-273	–	244-249	–		6	0.79
290-314	α-J, -K	250-274	α-I, -J	insertion 3	17	1.19
315-339	α-L	298-322	α-L		25	2.05
Overall correspondence						
Nucleotide binding fold only					180	2.40
All corresponding Cα atoms					261	2.98

pothesis that some conformational change of the interdomain region between domains III and IV allows both of these domains to interact properly with the anticodon arm of $tRNA^{Glu}$ [11]. On the basis of the model constructed in this study, we are now investigating, by molecular dynamics simulation, the large conformational changes of $tRNA^{Glu}$ and GluRS that are necessary for specific recognition of the anticodon and other determinants.

Acknowledgements: We are grateful to Dr. T.A. Steitz for providing M.G. with the coordinates of the GlnRS·$tRNA^{Gln}$ complex. This work was supported by Grants-in-Aid for Science Research on Priority Areas from the Ministry of Education, Science and Culture of Japan (S.Y.).

References

[1] Normanly, J., Ogden, R.C., Horvath, S.J. and Abelson, J. (1986) Nature 321, 213–219.
[2] Yarus, M. (1988) Cell 55, 739–741.
[3] Normanly, J. and Abelson, J. (1989) Annu. Rev. Biochem. 58, 1029–1049.
[4] Schimmel, P. (1989) Biochemistry 28, 2747–2759.
[5] Schulman, L.H. (1991) Progr. Nucleic Acids Res. Mol. Biol. 41, 23–87.
[6] Giegé, R., Puglisi, J.D. and Florentz, C. (1993) Progr. Nucleic Acids Res. Mol. Biol. 45, 129–206.
[7] Eriani, G., Delarue, M., Poch, O., Gangloff, J. and Moras, D. (1990) Nature 347, 203–206.
[8] Rossmann, M.G., Moras, D. and Olsen., K.W. (1974) Nature 250, 194–199.
[9] Rould, M.A., Perona, J.J., Söll, D. and Steitz, T.A. (1989) Science 246, 1135–1142.
[10] Rould, M.A., Perona, J.J. and Steitz, T.A. (1991) Nature 352, 213–218.
[11] Nureki, O., Vassylyev, D.G., Katayanagi, K., Shimizu, T., Sekine, S., Kigawa, T., Miyazawa, T., Yokoyama, S. and Morikawa, K. (1995) Science 267, 1958–1965.
[12] Pealman, D.A., Case, D.A., Caldwell, J.C., Seibel, G.L., Singh, U.C., Weiner, P. and Kollman, P.A. (1991) AMBER 4.0, University of California, San Francisco.
[13] Biosym Technologies (1993) Insight II User Guide, version 2.3.0. San Diego: Biosym Technologies.
[14] Sussman, J.L., Holbrook, S.R., Warrant, R.W., Church, G.M. and Kim, S.H. (1978) J. Mol. Biol. 123, 607–630.
[15] Westhof, E., Dumas, P. and Moras, D. (1985) J. Mol. Biol. 184, 119–145.
[16] Berstein, F.C., Koetzle T.F., Williams, G.J.B., Meyer, E.F., Brice, M.D., Rodgers, J.R., Kennard, O., Shimanouchi, T. and Tasumi, M. (1977) J. Mol. Biol. 112, 535–542.
[17] Rossmann, M.G. and Argos, P. (1975) J. Mol. Biol. 250, 7525–7532.
[18] McLachlan, A.D. (1979) J. Mol. Biol. 128, 49–79.
[19] Ladner, J.E., Jack, A. Robertus, J.D., Brown, R.S., Rhhodes, D., Clark, B.F.C. and Klug, A. (1975) Proc. Natl. Acad. Scie. USA 91, 8670–8674.
[20] Sekine, S., Nureki, O., Sakamoto, K., Niimi, T., Tateno, M., Go, M., Kohno, T., Brisson, A., Lapointe J. and Yokoyama, S., submitted to J. Mol. Biol.
[21] Kern, D. and Lapointe, J. (1979) Biochemistry 18, 5819–5826.
[22] Sylvers, L.A., Rogers, K.C., Shimizu, M., Ohtsuka, E. and Söll, D. (1993) Biochemistry 32, 3836–3841.
[23] Rogers, K.C. and Söll, D. (1993) Biochemistry 32, 14210–14219.