Crystal structure of domain E of *Thermus flavus* 5S rRNA: a helical RNA structure including a hairpin loop

Markus Perbandt[a], Alexis Nolte[a], Siegfried Lorenz[a], Rolf Bald[a], Christian Betzel[b,*], Volker A. Erdmann[a]

[a] *Institut für Biochemie, Freie Universität Berlin, Thielallee 63, 14195 Berlin, Germany*
[b] *Institut für Physiologische Chemie, Universitäts-Krankenhaus Eppendorf, Arbeitsgruppe für Makromolekulare Strukturanalyse, c/o DESY Geb. 22a, Notkestraße 85, 22603 Hamburg, Germany*

Received 25 February 1998

Abstract The synthetic RNA fragment 5′-CUGGGCGG(GCGA)CCGCCUGG (nucleotides in parentheses indicate the loop region) corresponds to the natural sequence of domain E from nucleotides 79–97 of the *Thermus flavus* 5S rRNA including a hairpin loop. The RNA structure determined at 3.0 Å and refined to an *R*-value of 24.1% also represents the first X-ray structure GNRA tetraloop. The loop is in distinctly different conformation from other GNRA tetraloops analyzed by NMR. The conformation of the two molecules in the asymmetric unit is influenced and stabilized by specific intermolecular contacts. The structural features presented here give evidence for the ability of RNA molecules to adapt to specific environments.

Key words: 5S rRNA; Crystal structure; RNA-RNA interaction; Hairpin loop

1. Introduction

The ribosomal 5S RNA is approximately 120 nucleotides long and is an integral part of the large ribosomal subunit. Several parts of the 5S rRNA interact specifically with several ribosomal proteins [1–3]. It is clear that reconstituted 50S ribosomal subunits, lacking the 5S rRNA, are inactive in protein biosynthesis. The function most drastically impaired is that of peptidyltransferase [1,2]. Structural studies will eventually support a more detailed understanding of the precise biological function of the 5S rRNA. Chemical probing as well as sequence alignments have made it possible to build a more general model for the 5S rRNA secondary structure [4], as shown in Fig. 1. Our extensive attempts to crystallize ribosomal 5S RNAs have led to crystals which diffract to about 7.5 Å resolution [5]. So far a native data set has been collected and heavy atom search is currently in preparation. To obtain information at atomic resolution we have turned to the chemical synthesis [6] and crystallization of the various structural domains. The molecular structure of these parts of the 5S rRNA will be the first step to increase our knowledge of how proteins recognize and interact with ribosomal nucleic acids. We have already reported the crystal structure of domain A from *Thermus flavus* [7,8] and preliminary diffraction studies of domain E [9]. Recently Correll et al. [10] have presented the X-ray structure of a 62 nt fragment of *Escherichia coli* 5S RNA mainly in helical conformation, which corresponds, according to Fig. 1, to the sequence 1–120, missing the domains B and C as well as the tetraloop in domain E. At the structural segment where the ribosomal protein L25 is supposed to be bound [11], the major groove of the helix is extremely wide. Domain E, reported here, is of major interest because crosslinks between 5S rRNA and 23S rRNA demonstrated that the location of the hairpin loop in domain E is near the peptidyltransferase center of the ribosome [12]. The hairpin loop of domain E belongs to the highly conserved class of highly stable 5′-GNRA tetraloops occurring in ribosomal RNAs (N is any nucleotide, R is G or A) [13]. Pley et al. [14] have determined the crystal structure of the 5′-GAAA tetraloop occurring in the hammerhead ribozyme and due to their structural relevance three other 5′-GNRA tertraloops have been analyzed so far by NMR techniques [15,16]. However, the exact function of these loops is not clear and surprisingly the 5′-GNRA tetraloop of domain E also shows a different conformation in comparison to the other tetraloops mentioned before. The two molecules in the asymmetric unit are in distinctly different conformations and form slightly distorted RNA helices, including the hairpin loop. The conformation and interaction of the two molecules give some new hints about RNA-RNA interactions as they have been observed for a few other A-RNA fragments [8,10,14,17] and as they might occur for instance in the ribosome or mRNA

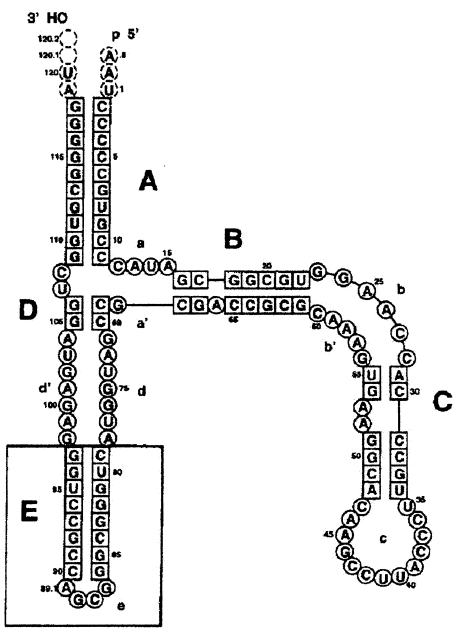

Fig. 1. Secondary structure of *Thermus flavus* 5S rRNA [4]. The individual domains are marked A through E. Domain E is boxed.

*Corresponding author. Fax: (49) (40) 8998-4747.
E-mail: Betzel@unisgi1.desy.de

PII S0014-5793(98)00436-0

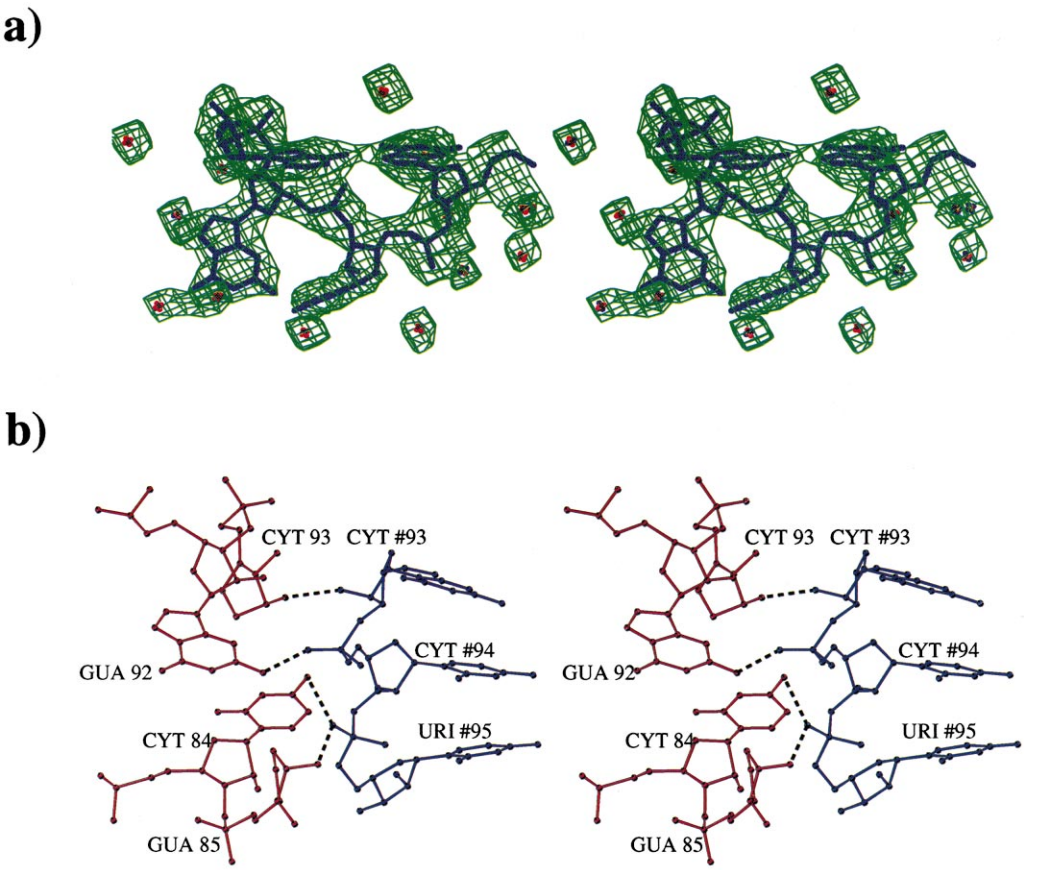

Fig. 2. a: Representative $2F_o-F_c$ electron density map showing the loop region and the surrounding water molecules of molecule B in stereo. b: Stereo plot showing selected H bond contacts of molecule A (in red) to a twofold symmetry related of molecule B (in blue). The intermolecular interaction leads to the loss of the Watson-Crick conformation for guanine 92 and cytosine 84 in molecule A.

structures. The different conformations demonstrate the apparently large reservoir of structural flexibility of RNA molecules.

2. Materials and methods

Domain E of 5S rRNA of *Thermus flavus* was prepared by solid phase chemical synthesis [6] and further purified by reversed phase HPLC. Crystals of the 20-mer suitable for X-ray analysis were obtained by vapor diffusion as reported before [9]. The space group was assigned to $P3_221$ with unit cell parameters of $a=b=42.8$ Å and $c=162.2$ Å. The packing parameter V_M [18,19] is 3.7 Å^3/Da for two molecules in the asymmetric unit. A complete data set to 3.0 Å resolution was collected from one single crystal at 4.0°C using a MAR 300 mm image plate detector at the EMBL beam line X11 at DORIS/DESY. The data set was processed using the program DENZO [20] with a resulting $R_{symm}=7.2\%$. The phase problem was solved by stepwise molecular replacement applying the program AMoRe [21,22] distributed with CCP4 [23]. The 16 nt search model was derived from domain A of 5S rRNA. Several potential solutions were indicated by high correlation factors from the cross-rotation function using data in the resolution range of 8.0–4.0 Å. All potential solutions were orientated with the helical axis parallel to the *c*-axis but differing only by 30° rotation. An extensive and systematic translation search, excluding all positions with major symmetry overlays, yielded a preliminary model with two fragments in the asymmetric unit and a starting *R* value of 45.9% was assigned. This initial model was forwarded to a standard annealing refinement with X-PLOR 3.1 [24] applying a modified RNA geometry parameter library [25] and applying strong restraints according to the limited resolution. The *R* value dropped to 38% for all data between 12.0 and 3.5 Å. At this stage all regions which could not clearly be assigned in the $2F_o-F_c$ and F_o-F_c electron density maps were excluded. Iterative steps of calculating omit density maps, model building and refinement finally revealed two molecules with a hairpin loop. At this stage solvent molecules were introduced during final steps of refinement and the resolution was expanded to 3.0 Å. The solvent molecules were assigned in the

Table 1
Refinement statistics and quality of the model

Space group	$P3_221$
Unit cell dimensions (Å)	
$a=b=$	42.8
$c=$	162.8
Crystal volume/Dalton (V_M)	3.7 Å^3/Da
Resolution range (Å)	15.0–3.0
Number of reflections used	2990
R value (%)	24.1
R_{free} (%)	31.8
Model	
Nucleic acid atoms	858
Water molecules	221
Average *B* value (Å^2)	
of all atoms	30.4
of molecule A	31.1
of molecule B	29.9
of water atoms	29.1
of all phosphates	37.2
of all sugar atoms	35.8
of all bases	23.8
Parameter file	Dna-rna.param [25]
RMS deviations from ideal geometry	
Bond lengths	0.01 Å
Bond angles	1.5°
Dihedral angles	25.1°

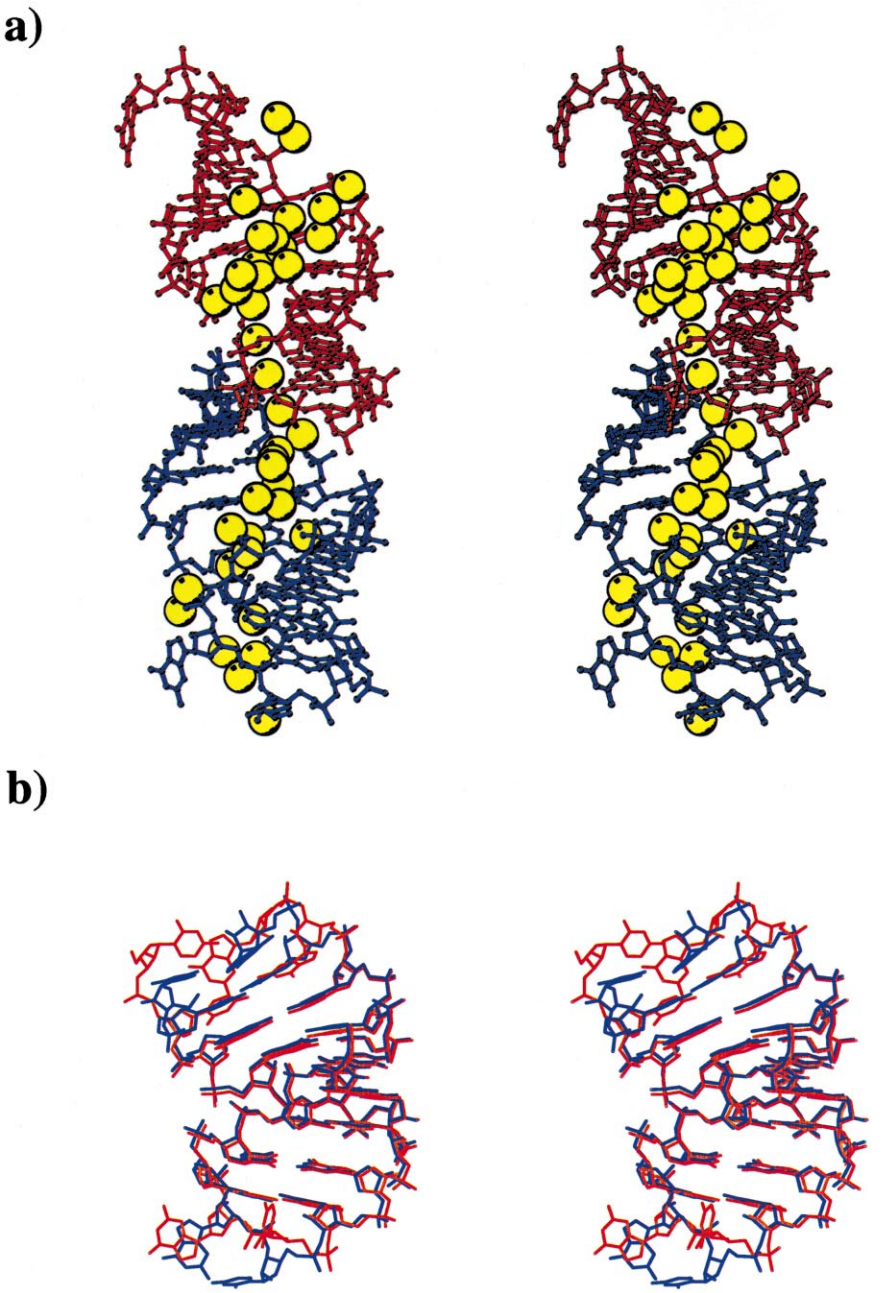

Fig. 3. a: Stereo plot of the two molecules in the asymmetric unit including internal waters of the major groove. The waters are drawn in yellow as a cpk model to indicate the way of space filling the interior of the helices. Molecule A is drawn in red and molecule B in blue. b: Stereo plot of the superimposed molecules: molecule A in red and molecule B in blue. The 'open ends' and the loop regions are due to high intrinsic conformational flexibility of the molecules in different conformations.

F_o-F_c difference electron density maps contoured at 3σ. Waters were deleted if they failed to reappear at 1σ in the $2F_o-F_c$ map after a new refinement cycle. In total 221 solvent molecules were introduced. An example for a representative $2F_o-F_c$ density is given in Fig. 2a. The final refinement using NUCLSQ [27] yielded a conventional R value of 24.1% (R_{free} = 31.8%) including all data up to 3.0 Å. The coordinates of the structure are deposited at the NDB with the entry code URT068.

3. Results and discussion

As a consequence of the slightly distorted helices the averaged twist angles for the two double strands appear to be 30.6° and 30.9° with a high standard deviation as shown in Table 2. The helical part of the synthetic domain E and the corresponding part in the *E. coli* 5S RNA fragment [10] display a root mean square (r.m.s.) deviation of 1.4 Å for phosphate atoms as reference. That proves that both fragments are in the native conformation. Therefore domain E can be classified as a right-handed A-RNA. The final model of domain E consists of two molecules with 858 non-hydrogen atoms in total and 221 solvent molecules. Fig. 3a shows the two molecules including the internal water molecules. The stereochemistry of the final model shows deviations from ideal values comparable to other nucleic acid structures. The quality of

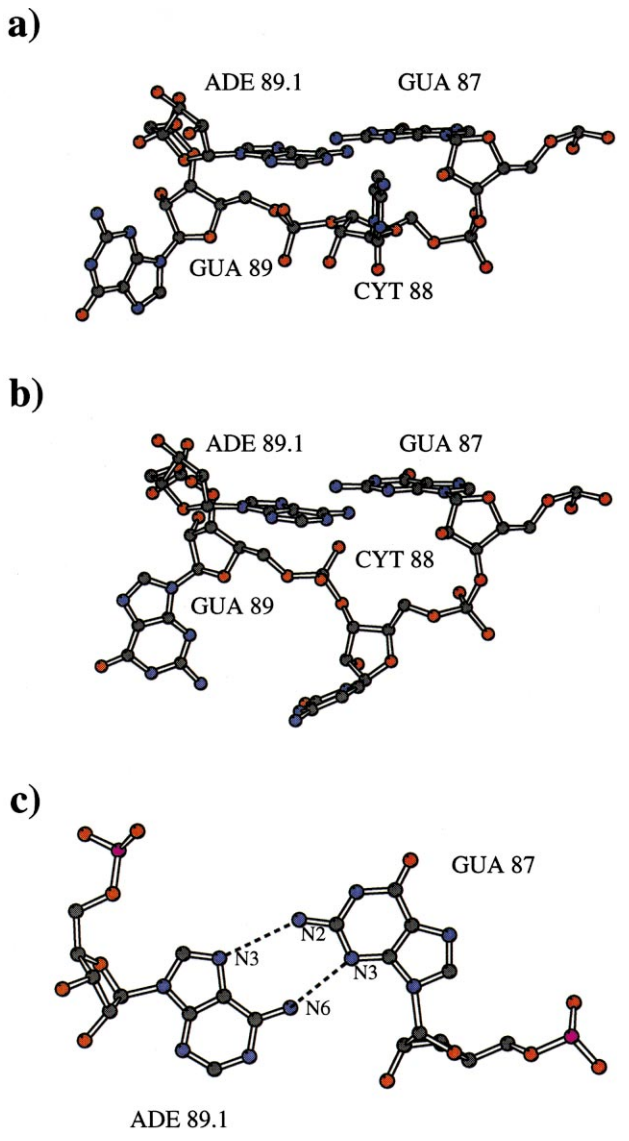

Fig. 4. The terminal bases cytosine 88 and guanine 89 are in different conformations in both molecules. a: The loop region of molecule A. b: The loop region of molecule B. In both molecules the terminal bases stick out towards the solvent. c: In both molecules adenine 89.1 and guanine 87 interact via two H bonds: N2 (G87) to N3 (A89.1) and N3 (G87) to N6 (A89.1). They show an asymmetric hetero-purine formation.

the structure is summarized in Table 1. The mean *B* value is 30.4 Å^2 for all atoms. As found for the crystal structure of domain A of 5S rRNA the highly ordered internal waters are essential for stabilizing the RNA structure and support unusual nucleotide conformations by compensating a lack of conventional H bonds of non-Watson-Crick base pairs [8]. They can be classified as an integral part of the structure [28]. Due to the additional contribution of the 2′-hydroxyl groups, which are absent in DNA molecules, to support H bond networks it is now obvious that solvent is an inherent feature of RNA structures. Biwas et al. [29] have reported that most of the 2′-hydroxyl groups are hydrated or sometimes even hydrogen-bonded to the bases of the same residue. For domain E all molecules are oriented along the long *c*-axis of the unit cell. The molecules stack together and considering the hairpin loop, they form a indefinite layer of antiparallel helices. Both molecules are in 'head to tail' interaction, which means that the 5′ and the 3′ end of one molecule interacts with the loop region of the corresponding molecule (Fig. 3a). Therefore the 'open ends' are in completely different conformations as shown in Fig. 3b. Overall the first two nucleotides for each strand are not in Watson-Crick base pair conformation, the structure is outside the double helical characteristics, the terminal base pairs are 'flipped out'. Similar observations have been also reported for DNA molecules [30].

The most important helical parameters of both molecules are summarized in Table 2 in comparison to other RNA structures. The helical parameters were calculated applying the program NUPARM [31] and indicate features characteristic of alternating sequences. The differences in the average values for the helical twist, rise and *x* displacement of the base pairs and their high standard deviations furthermore illustrate conformational differences of both molecules. The two molecules can be superimposed using all phosphate atoms with a r.m.s. deviation of 1.4 Å between equivalent phosphate atoms. The maximum displacement is 5.5 Å due to several distinctly different inter- and intramolecular contacts. The influence of intermolecular interactions on the structural conformation indicates the flexibility and their ability to adjust and adapt to local geometric restraints.

The conformational differences of both molecules are partially re-stabilized by several direct hydrogen bonds between both RNA helices. The direct contact of the phosphate backbone of molecule A and molecule B initiates a change of the base pairing in molecule A. G83–C93 are in a reversed Watson-Crick formation instead of the normal Watson-Crick observed for molecule B. Furthermore, Fig. 2b illustrates that N-2 of G92 (molecule A) is hydrogen bonded to O-1P of C94 (molecule B) and N-4 of C84 (molecule A) to O-1P of U95 (molecule B). This leads to the loss of the Watson-Crick formation for this particular base pair in molecule A. As a consequence the helical parameters deviate strongly from those expected for a ideal helix formation.

The loop regions as shown in Fig. 4a,b are in different

Table 2
Average values of important helical parameters

	Twist (°)	Rise (Å)	X-Dsp (Å)
Molecule A[a]	30.6 (11.0)	3.4 (0.5)	5.8 (1.9)
Molecule B[a]	30.9 (10.0)	3.2 (0.4)	5.5 (2.7)
A′-RNA[b]	30.0	3.0	4.4
Domain A 5S RNA [8]	33.3	2.4	4.5
$[U(UA)_6]_2$ [32]	33.2	2.8	3.6
tRNA (mono) [33]	33.2	2.5	4.4

[a]All helical parameters are determined by including only the helical part of the structure. Standard deviations are in parentheses.
[b]Fiber model of poly(rI)·poly(rC) A′-RNA [26].

conformations as for example the 5′-GAAA tetraloop in the X-ray structure of the ribozyme [14] and the three other 5′-GNRA tetraloops solved by NMR techniques [15,16]. A representative electron density of the loop regions including the surrounding water molecules is shown in Fig. 2a. However, the electron density for the terminal cytosine 88 and guanine 89 indicates a high intrinsic flexibility. Also for both molecules the guanine 87 and adenine 89.1 show an asymmetric heteropurine formation as described before by Saenger [34] and interact via two unusual H bonds shown in Fig. 4c. The terminal cytosine 88 and guanine 89 are 'sticking out' like 'antennas' towards the solvent and are not stabilized by direct H bonds to other nucleotides. A similar conformation of the terminal nucleotides was also already predicted based on solution data and stereochemical restraints by Westhof et al. for the 5′-GNRA tetraloop [35] of the *Xenopus laevis* oocyte 5S rRNA. In contrast to this computational model the hydrogen bonding scheme of guanine 87 and adenine 89.1 are is different to the H bonds we found for domain E.

These facts illustrate the high structurally and functionally important flexibility of RNA molecules. The orientation of the bases towards the solvent and the fact that the hairpin loop is located near the peptidyltransferase center give us some hints about how the 5S RNA could interact with the peptidyltransferase in the ribosome.

Acknowledgements: We thank Jens-Peter Fürste for helpful discussions and Eric Westhof and Jorge Navaza for help applying the programs NUCLSQ and AMoRe. The project has been supported by the Deutsche Forschungsgemeinschaft (Leibniz Award, SFB 344-D6), the Deutsche Agentur für Raumfahrt-Angelegenheiten and the Fonds der Chemischen Industrie e.V. The technical assistance of Kerstin Brumm is also acknowledged.

References

[1] Erdmann, V.A., Fahnestock, K.H. and Nomura, M. (1971) Proc. Natl. Acad. Sci. USA 68, 2932–2936.
[2] Hartmann, R.K., Vogel, D.W., Walker, R.T. and Erdmann, V.A. (1988) Nucleic Acids Res. 16, 3511–3524.
[3] Horne, J. and Erdmann, V.A. (1972) Mol. Gen. Genet. 119, 337–344.
[4] Specht, T., Wolters, J. and Erdmann, V.A. (1990) Nucleic Acids Res. 18, (Suppl.) 2215–2230.
[5] Lorenz, S., Betzel, Ch., Raderschall, E., Dauter, Z., Wilson, K.S. and Erdmann, V.A. (1991) J. Mol. Biol. 219, 399–402.
[6] Bald, R., Brumm, K., Buchholz, B., Fuerste, J.P., Hartmann, R.K., Jaeschke, A., Kretschmer-Kazemi, F., Lorenz, S., Raderschall, E., Schlegl, J., Specht, T., Zhang, M., Cech, D. and Erdmann, V.A. (1992) in: Structural Tools for the Analysis of Protein-Nucleic Acid Complexes (Lilley, D.J., Heumann, H. and Suck, D., Eds.), pp. 449–466, Birkhäuser Verlag, Basel.
[7] Lorenz, S., Betzel, Ch., Fuerste, J.P., Bald, R., Zhang, M., Raderschall, E., Dauter, Z., Wilson, K.S. and Erdmann, V.A. (1993) Acta Crystallogr. D 49, 418–420.
[8] Betzel, Ch., Lorenz, S., Fuerste, J.P., Bald, R., Zhang, M., Schneider, Th.R., Wilson, K.S. and Erdmann, V.A. (1994) FEBS Lett. 351, 159–164.
[9] Nolte, A., Klußmann, S., Lorenz, S., Bald, R., Betzel, Ch., Dauter, Z., Wilson, K.S., Fuerste, J.P. and Erdmann, V.A. (1995) FEBS Lett. 374, 292–294.
[10] Correll, C.C., Freeborn, B., Moore, P.B. and Steitz, T. (1997) Cell 91, 705–712.
[11] Moore, P.B. (1996) in: Ribosomal RNA (Zimmermann, R.A. and Dahlberg, A.E., Eds.), pp. 199–236, CRC Press, Boca Raton, FL.
[12] Dontsova, O., Tishkov, V., Dokudovskaya, S., Bogdanov, A., Döring, T., Rinke-Appel, J., Thamm, S., Greuer, B. and Brimcombe, R. (1994) Proc. Natl. Acad. Sci. USA 91, 4125–4129.
[13] Woese, C.R., Winkler, S. and Gutell, R.R. (1990) Proc. Natl. Acad. Sci. USA 87, 8467–8471.
[14] Pley, H.W., Flaherty, K.M. and McKay, D.B. (1994) Nature 372, 111–113.
[15] Heus, H. and Pardi, A. (1991) Science 253, 191–193.
[16] Jucker, F.M., Heus, H., Yip, P.F., Moors, H.M. and Pardi, A. (1996) J. Mol. Biol. 264, 968–980.
[17] Schindelin, H., Zhang, M., Bald, R., Fuerste, J.P., Erdmann, V.A. and Heinemann, U. (1995) J. Mol. Biol. 249, 595–603.
[18] Matthews, B.W. (1968) J. Mol. Biol. 33, 491–497.
[19] Feigin, L.A. and Svergun, D.I. (1987) in: Structure Analysis by Small-Angle X-ray and Neutron Scattering, p. 120, Plenum Press, New York.
[20] Otwinowski, Z. (1991) DENZO, Yale University, New Haven, CT.
[21] Navaza, J. and Vernoslova, A. (1995) Acta Crystallogr. A 551, 445–449.
[22] Navaza, J. (1992) Proceedings of the Daresbury Study Weekend on Molecular Replacement.
[23] Collaborating Computing Project No. 4, SERC Daresbury Laboratory, Warrington (1994) Acta Crystallogr. D 50, 760–763.
[24] Brünger, A.T., Karplus, M. and Petsko, G.A. (1989) Acta Crystallogr. A 45, 50–61.
[25] Parkinson, G., Vojtechovsky, J., Clowney, L., Brunger, A.T. and Berman, H.M. (1996) Acta Crystallogr. D 52, 57–64.
[26] Chandrasekaran, R. and Arnott, S. (1989) in: Landolt-Boernstein, New Series, Group VII (Saenger, W., Ed.) Vol. 1b, pp. 31–170, Springer, Berlin.
[27] Westhof, E., Dumas, P. and Moras, D. (1985) J. Mol. Biol. 164, 119–145.
[28] Westhof, E. (1987) Int. J. Biol. Macromol. 9.
[29] Biswas, R. and Sundaralingam, M. (1997) J. Mol. Biol. 270, 511–519.
[30] Nunn, C.M. and Neidle, S. (1996) J. Mol. Biol. 256, 340–351.
[31] Bhattacharyya, D. and Bansal, M. (1990) J. Biomol. Struct. Dynam. 8, 539–572.
[32] Dock-Bregon, A.C., Chevier, B., Podjanry, A., Johnson, J., de Bear, J.S., Gough, G.R., Gilham, P.T. and Moras, D. (1989) J. Mol. Biol. 209, 459–474.
[33] Westhof, E. and Sundaralingham, M. (1986) Biochemistry 25, 4868–4878.
[34] Saenger, W. (1984) Principles of Nucleic Acid Structure, pp. 119–122, Springer-Verlag, New York.
[35] Westhof, E., Romby, P., Romaniuk, P.J., Ebel, J.P., Ehresmann, C. and Ehresmann, B. (1989) J. Mol. Biol. 207, 417–431.