Accelerated Publication

THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 275, No. 11, Issue of March 17, pp. 7439–7442, 2000

Crystal Structure of the Human U4/U6 Small Nuclear Ribonucleoprotein Particle-specific SnuCyp-20, a Nuclear Cyclophilin*

(Received for publication, November 18, 1999, and in revised form, December 14, 1999)

Ulrich Reidt‡§, Klaus Reuter‡, Tilmann Achsel‡, Dierk Ingelfinger§, Reinhard Lührmann‡§, and Ralf Ficner‡¶

From the ‡Institut für Molekularbiologie und Tumorforschung, Universität Marburg, 35037 Marburg, Germany and the §Abteilung für Zelluläre Biochemie, Max-Planck-Institut für Biophysikalische Chemie, 37070 Göttingen, Germany

The cyclophilin SnuCyp-20 is a specific component of the human U4/U6 small nuclear ribonucleoprotein particle involved in the nuclear splicing of pre-mRNA. It stably associates with the U4/U6-60kD and -90kD proteins, the human orthologues of the *Saccharomyces cerevisiae* Prp4 and Prp3 splicing factors. We have determined the crystal structure of SnuCyp-20 at 2.0-Å resolution by molecular replacement. The structure of SnuCyp-20 closely resembles that of human cyclophilin A (hCypA). In particular, the catalytic centers of SnuCyp-20 and hCypA superimpose perfectly, which is reflected by the observed peptidyl-prolyl-*cis/trans*-isomerase activity of SnuCyp-20. The surface properties of both proteins, however, differ significantly. Apart from seven additional amino-terminal residues, the insertion of five amino acids in the loop α1-β3 and of one amino acid in the loop α2-β8 changes the conformations of both loops. The enlarged loop α1-β3 is involved in the formation of a wide cleft with predominantly hydrophobic character. We propose that this enlarged loop is required for the interaction with the U4/U6-60kD protein.

The nuclear splicing of pre-mRNA is catalyzed by the spliceosome, which is a highly dynamic macromolecular assembly formed by the ordered interaction of the snRNPs[1] U1, U2, U4/U6, and U5 and of several non-snRNP proteins (for review see Refs. 1–3). Although the U1 and U2 snRNPs bind to the 5′-splice site and the branch point of the pre-mRNA, respectively, the U4/U6 and U5 snRNPs pre-assemble to the [U4/U6.U5] tri-snRNP. After association of the tri-snRNP with the pre-spliceosome, the mature and active spliceosome is formed in a process that involves extreme rearrangements of the spliceosomal RNA components. Thus, before the first catalytic step of the transesterification reaction, the U4/U6 snRNA duplex is disrupted, resulting in the release of U4 snRNA and the association of U6 snRNA with U2 snRNA and the pre-mRNA (reviewed in Ref. 4). These RNA rearrangements most likely imply significant rearrangements and conformational changes of a number of tri-snRNP proteins. One recently identified protein, which seems a good candidate to catalyze such conformational changes, is the U4/U6 snRNP-specific 20-kDa protein. U4/U6-20kD was shown to be a cyclophilin and therefore termed USA-Cyp (U-snRNP-associated cyclophilin) or SnuCyp-20 (Snurp cyclophilin-20kDa) (5, 6).

Cyclophilins form a structurally well characterized family of closely related proteins that exhibit peptidyl-prolyl *cis-trans* isomerase (PPIase) activity and thereby accelerate the folding of proteins requiring the isomerization of a peptidyl-prolyl bond (reviewed in Ref. 7). However, it remains unclear whether the PPIase activity is the most important cellular function of cyclophilins. There is growing evidence that their main function may be to act as chaperones and that PPIase activity is solely a side effect (8–10). Nevertheless, purified [U4/U6.U5] tri-snRNP particles were shown to be endowed with PPIase activity, which could be inhibited by the immunosuppressive drug cyclosporin A (CsA), consistent with the presence of the cyclophilin SnuCyp-20 (6). SnuCyp-20 could be isolated in the form of a stable RNA-free ternary complex with the U4/U6-snRNP-specific 60- and 90-kDa proteins (5, 6), which represent the human orthologues of the previously characterized yeast splicing factors Prp4 and Prp3 (5, 11). Because SnuCyp-20 does not interact with one of these two proteins via its active site (see "Results and Discussion"), we have solved the crystal structure of the SnuCyp-20 protein to get indications as to how the formation of a complex with these proteins is accomplished.

EXPERIMENTAL PROCEDURES

The coding region of the human *snucyp-20* cDNA was amplified by polymerase chain reaction from an expressed sequence tag (EST) from human liver (GenBank™ accession number T53949) and cloned into the multiple cloning site of expression vector pGEX-4T-2 (Amersham Pharmacia Biotech) resulting in a *GST/snucyp-20* fusion gene. Expression of the fusion gene in *Escherichia coli* strain BL21(DE3) and subsequent affinity purification of the resulting protein using glutathione-Sepharose 4B (Amersham Pharmacia Biotech) was performed according to the protocols of the vendor and yielded an almost pure fusion protein. The GST moiety was cleaved off with 10 units of thrombin/mg of fusion protein and removed by gel filtration using a Superdex 75 (26/60) column (Amersham Pharmacia Biotech). The pure SnuCyp-20 was concentrated to 10 mg ml^{-1} in the buffer used for gel filtration (120 mM NaCl, 2 mM DTT, 20 mM HEPES, pH 7.6) by means of Centriprep-10 concentrators (Amicon). Crystallization of the pure protein was performed at 21 °C in Linbro plates using the hanging drop vapor diffusion technique. A drop of 1.5 μl of protein solution was mixed with an equal volume of reservoir solution and sealed against 1 ml of reservoir solution. The best crystals were obtained with 25% (w/v) PEG6000, 200 mM $MgCl_2$, 100 mM Tris-HCl, pH 8.5. To collect data under cryo conditions, crystals were flash-frozen in a solution containing 25% (w/v) PEG6000, 200 mM $MgCl_2$, 11% (v/v) glycerol, and 100 mM

* This work was supported by grants from the Deutsche Forschungsgemeinschaft (Sonderforschungsbereich (SFB) 286/A11 to R. F. and SFB 379/A6 to R. L.) and the Gottfried Wilhelm Leibnitz Programm (to R. L.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The atomic coordinates and structure factors (*codes 1QOI, R1QOISF, and 2CPL) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ* (*http://www.rcsb.org/*).

¶ To whom correspondence should be addressed. Tel.: 49-6421-2865390; Fax: 49-6421-2867008; E-mail: ficner@imt.uni-marburg.de.

[1] The abbreviations used are: snRNP, small nuclear ribonucleoprotein particle; SnuCyp-20, Snurp cyclophilin-20kDa; hCypA, human cyclophilin A; CsA, cyclosporin A; PPIase, peptidyl-prolyl *cis-trans* isomerase; pre-mRNA, precursors of mRNA; GST, glutathione *S*-transferase; DTT, dithiothreitol; PEG6000, polyethylene glycol 6000 Da.

Tris-HCl (pH 8.5). A complete native data set was collected at a resolution of 2.0 Å. X-ray data were collected on an R-AXIS IV image plate system equipped with a Rigaku RU-300 rotating anode generator operating at 50 kV and 100 mA and focusing mirrors (Molecular Structure Corp.). The crystal-to-detector distance was 120 mm and 1° oscillation images were collected with a 10-min exposure time. Diffraction data were processed using the programs DENZO and SCALEPACK (12). The SnuCyp-20 crystals belong to space group $P2_12_12_1$ (cell constants: a = 47.3 Å, b = 59.9 Å, c = 60.7 Å) and contain one molecule in the asymmetric unit. The structure of SnuCyp-20 was solved by Patterson search methods using the program X-PLOR (13) with the refined crystal structure of hCypA (14) (Protein Data Bank code 2CPL) as a search model. The initial electron density map was significantly improved with the help of the program ARP (15). Model building into the electron density map was done with the program O (16), and the structure was refined using X-PLOR (13). The model was manually improved, and water molecules were built with help of the program ARP (15). The final model contains 173 amino acids and 188 water molecules and has good stereochemistry as evaluated using the program PROCHECK (17). The SnuCyp-20 coordinates have been deposited with the Protein Data Bank and will be released upon publication (Protein Data Bank codes 1QOI and R1QOISF for the structure factor entry).

RESULTS AND DISCUSSION

The human *snucyp-20* gene was overexpressed as a *Schistosoma japonicum* glutathion *S*-transferase (*GST*) fusion gene (18), and the resulting gene product was affinity-purified, yielding almost pure fusion protein. The fusion protein was proteolytically cleaved and the GST moiety subsequently removed by gel filtration. The elution volume of the pure SnuCyp-20 from the gel filtration column corresponded to a monomeric state. Efforts to crystallize the purified SnuCyp-20 resulted in well diffracting crystals using PEG6000 as a precipitant. A complete native data set was collected at a resolution of 2.0 Å, and the structure of SnuCyp-20 was solved by Patterson search methods using hCypA (14) as a search model. The data collection and refinement statistics are summarized in Table I. The structure of SnuCyp-20, a β-barrel of eight antiparallel β-strands and two α-helices sitting on the bottom and top of the barrel (Fig. 1), superimposes well with the structure of hCypA (Fig. 2). The root mean square deviation for 158 common Cα atoms of SnuCyp-20 and hCypA is 0.8 Å. Two short 3_{10} helices are present in SnuCyp-20. The first one lies within the loop connecting strand β6 and β7 and is also found in hCypA, whereas the second one, located in the loop between helix α2 and strand β8, is not present in hCypA. The active center of SnuCyp-20 and hCypA, as well as most residues shown for hCypA to be involved in CsA binding, superimpose almost perfectly, consistent with the fact that SnuCyp-20 ex-

TABLE I
Crystallographic data and refinement statistics

Data collection	
Maximum resolution	2.0 Å
No. of observed reflections	44,280
No. of unique reflections	12,102
Completeness	
Overall	97.8%
Last shell	98.8%
Rsym[a]	
Overall	5.3%
Last shell	11.5%
Refinement statistics	
Resolution range	30−2.0 Å
R-factor,[b] 17.0%	*R*-free, 22.4%
R-factor[b] (last shell), 19.1%	*R*-free (last shell), 22.1%
Deviations from ideal geometry	Bond lengths, 0.009 Å Bond angles, 1.44°

[a] $R_{sym} = \Sigma |I - \langle I \rangle| / \Sigma I$, where I is the observed intensity and $\langle I \rangle$ is the average intensity for multiple measurements.

[b] R-factor = $\Sigma\{F_o - F_c| / \Sigma|F_o|$, where F_o and F_c are the observed and calculated structure factors, respectively. R-free is the cross-validation R-factor calculated for 10% of the reflections omitted in the refinement process.

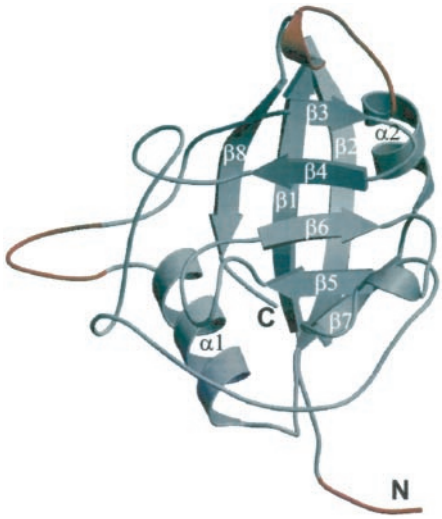

FIG. 1. **Three-dimensional structure of SnuCyp-20.** Residues that represent insertions into the primary structure, with respect to hCypA or whose Cα-atom positions differ significantly from the corresponding amino acids in hCypA are shown in *red*. The Fig. was produced using the programs MOLSCRIPT (33) and RASTER3D (34).

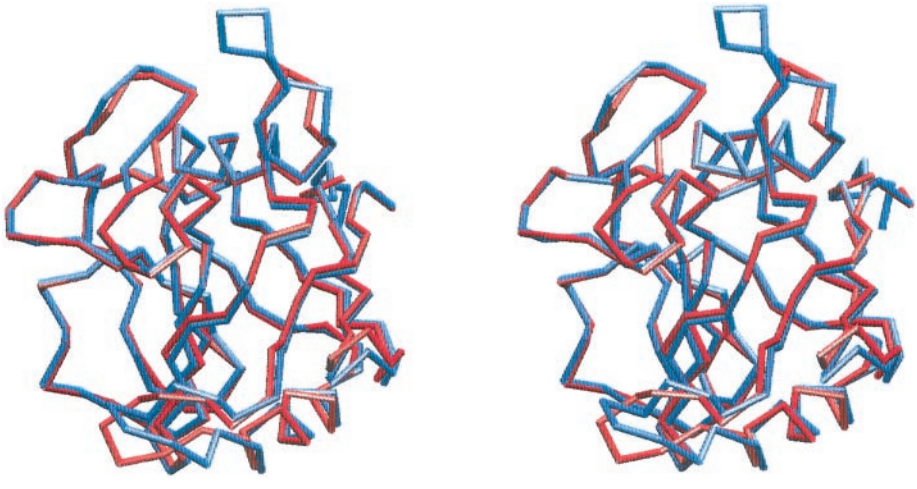

FIG. 2. **Stereo view of the superposition of the Cα-backbones of hCypA and SnuCyp-20.** hCypA is shown in *red* and SnuCyp-20 is *blue*. The superposition was created using the program VMD (35).

hibits PPIase activity, which is inhibited by CsA.[2]

Two cysteine residues of SnuCyp-20, namely Cys^{47} and Cys^{174}, are also present and highly conserved within a family of cyclophilins, whose members contain non-cyclophilin domains and are referred to as divergent cyclophilins (19). The spatial orientation of these cysteines renders the formation of a disulphide bond possible, because the distance between the sulfur atoms of these cysteines amounts to 4.9 Å and can be reduced to 1.7 Å by rotation of the sulfur atoms around the respective Cα-Cβ axes. However, Cys^{47} and Cys^{174} clearly exist in the reduced form in the presented structure of SnuCyp-20 (Fig. 3). Likewise, the crystal structure of a divergent cyclophilin from the nematode parasite *Brugia malayi* reveals the reduced form of the corresponding cysteines (19). This finding is consistent with the fact that both the crystals of both SnuCyp-20 and the *B. malayi* cyclophilin domain were obtained in the presence of DTT.

The main differences in the primary structure of SnuCyp-20 compared with hCypA are an amino terminus elongated by seven amino acids, the insertion of one amino acid in the loop connecting helix α2 with strand β8, and the insertion of five amino acids in the loop connecting helix α1 with strand β3.

Electron density can be allocated to three of the seven residues elongating the amino terminus of SnuCyp-20 compared with hCypA, namely to Asn^{5}, Ser^{6}, and Ser^{7}. The observed conformation of these residues, however, is the result of crystal packing contacts.

The insertion of one residue in loop α2-β8 causes the displacement of residues 158–163 with respect to the corresponding residues in hCypA; this leads to a significantly altered conformation of this loop with respect to hCypA, including the introduction of a short 3_{10} helix (Fig. 1). The respective loops in all other cyclophilins, whose structures have been solved (19–25), resemble however more the loop α2-β8 of SnuCyp-20 than of hCypA, thus rendering it unlikely that the conformation of loop α2-β8 in SnuCyp-20 represents an adaptation to specific requirements for its function in the spliceosome.

The insertion of five amino acids in loop α1-β3 results in the formation of a lobe whose conformation is stabilized by two salt bridges formed by the side chains of Glu^{50} and Arg^{52} or of Asp^{54} and Arg^{92}, respectively. This lobe creates, together with amino

[2] D. Ingelfinger, T. Achsel, U. Reidt, K. Reuter, R. Ficner, and R. Lührmann, unpublished data.

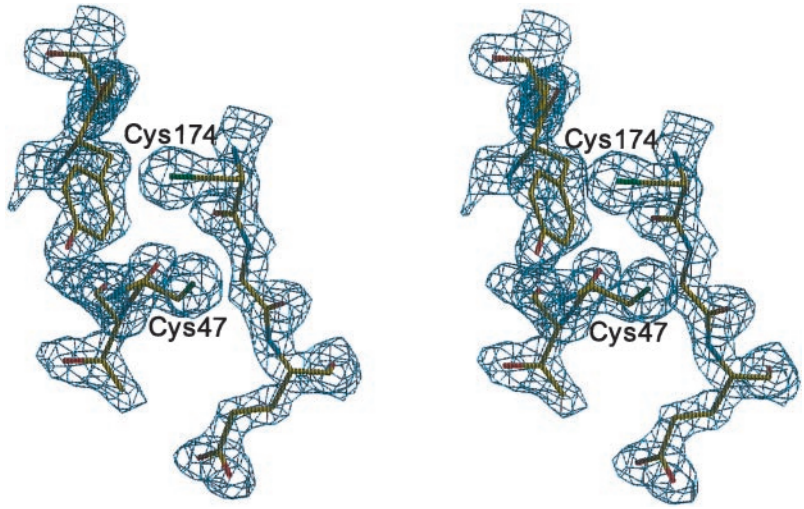

FIG. 3. **Stereo view of residues Cys^{47} and Cys^{174} superimposed with the 2 $F_o - F_c$ electron density map.** The electron density map (contoured at 1.5 σ) clearly shows the reduced form of Cys^{47} and Cys^{174} consistent with the growth of the crystals in the presence of 2 mM DTT.

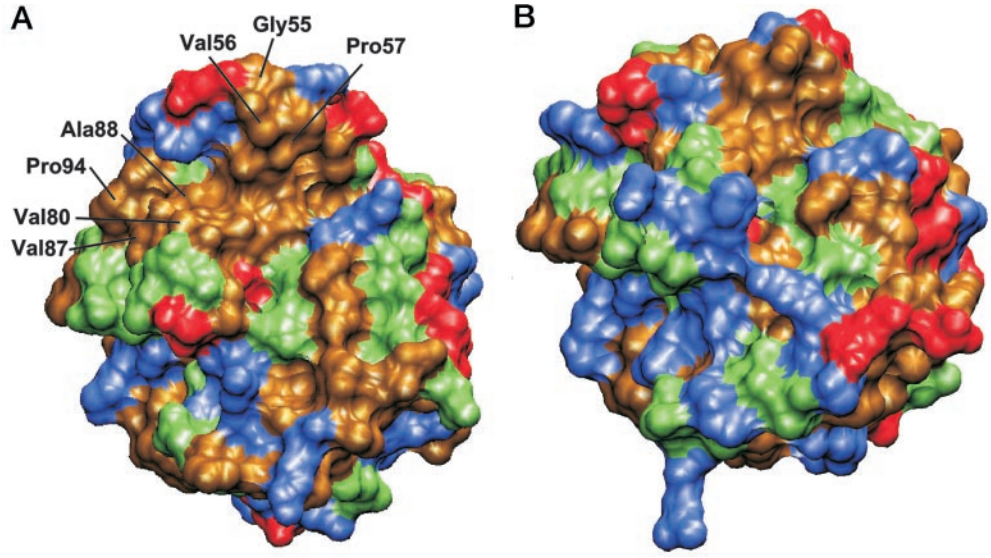

FIG. 4. **Surface representation of SnuCyp-20 (*A*) and hCypA (*B*).** The orientation of both molecules is identical and similar to that shown in Fig. 2. Acidic residues are shown in *red*, basic residues are *blue*, polar residues are *green*, and hydrophobic residues are *ochre*. The enlarged loop α1-β3 of SnuCyp-20 results in the formation of a lobe that creates, together with amino acids located in loop β4-β5, a wide cleft with a predominantly hydrophobic character. Residues contributing to this cleft that are not present in hCypA or whose counterparts are hydrophilic in hCypA are indicated. The surface was generated and displayed with the program VMD (35).

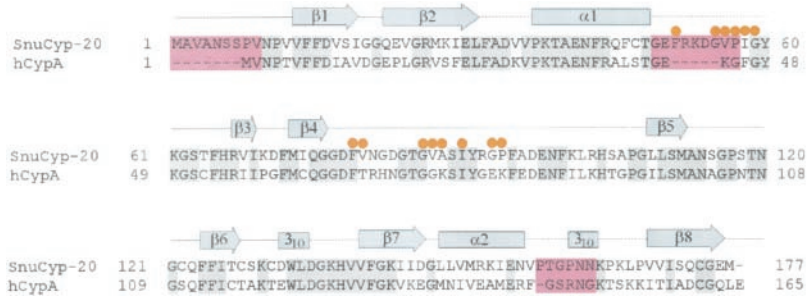

FIG. 5. **Alignment of the primary structures of SnuCyp-20 and hCypA.** The SnuCyp-20 secondary structure was assigned according to PROCHECK (17). Residues that are identical in SnuCyp-20 and hCypA are *shaded* in *gray*. Amino acids that are not present in hCypA or whose Cα positions differ significantly from the corresponding amino acids in hCypA are *shaded* in *magenta*. The *ochre-colored dots* indicate residues involved in the formation of a hydrophobic cleft and proposed to be involved in protein-protein interaction.

acids located in loop β4-β5, a wide cleft with predominantly hydrophobic character (Figs. 4 and 5). This cleft is not present in hCypA or in other cyclophilins of known structure and seems likely to create a protein-protein interaction site. SnuCyp-20 was shown to be part of a stable complex, which further contains the U4/U6-90kD and -60kD proteins (5, 6). Although no direct interaction between SnuCyp-20 and U4/U6-90kD could be demonstrated up to the present, U4/U6-60kD was identified as a direct binding partner of SnuCyp-20. U4/U6-60kD contains seven repeats of the WD40 motif and is therefore expected to fold into a seven-bladed propeller, whose large flat surface forms a platform of protein-protein interaction as shown for the β-subunits of G-proteins by x-ray structure analysis (26–29). This large flat propeller surface of the U4/U6-60kD orthologue Prp4 was shown to interact with the U4/U6-90kD orthologue Prp3 (30) and thus may not be accessible for SnuCyp-20.

In contrast to other known cyclophilin-protein interactions (8, 10) SnuCyp-20 obviously does not bind U4/U6-60kD via its catalytic site. This can be deduced from biochemical data, because hCypA, whose active site is identical to that of SnuCyp-20, is not able to bind U4/U6-60kD. In addition, PPIase activity could be shown for [U4/U6.U5] tri-snRNPs (6), which requires that the SnuCyp-20 catalytic center not be occupied by U4/U6-60kD. Finally, mutations introduced into the active site of SnuCyp-20, which abolished its catalytic activity, did not affect the U4/U6-60kD binding capability.[2] Combining these data, it seems likely that a region outside of the large flat surface of the U4/U6-60kD propeller serves as an anchor site for SnuCyp-20, which may act as a chaperone or PPIase on other sites of the spliceosome, probably to facilitate rearrangements of protein interactions during the splicing process.

The presented structure of SnuCyp-20, which shows a striking similarity to hCypA, reveals loop α1-β3 as well as a number of hydrophobic residues within loop β4-β5 as good candidates for mutational studies to figure out the interactions of this protein with U4/U6-60kD. After the thioredoxin-like protein U5-15kD (31, 32) this paper presents the second three-dimensional structure of an snRNP-associated protein having a close homologue in the cytoplasm, which obviously has been adopted by the spliceosome and adjusted in the course of evolution to the specific needs of the splicing machinery.

Acknowledgments—We are grateful to Gerhard Klebe and Milton Stubbs (Institut für Pharmazeutische Chemie, Universität Marburg), whose x-ray diffractometer we used.

REFERENCES

1. Burge, C. B., Tuschl, T. H., and Sharp, P. A. (1999) In *The RNA World* (Gesteland, R. F., Cech, T., and Atkins, J. F., eds) pp. 525–560, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
2. Krämer, A. (1996) *Annu. Rev. Biochem.* **65,** 367–409
3. Will, C. L., and Lührmann, R. (1997) *Curr. Opin. Cell Biol.* **9,** 320–328
4. Staley, J. P., and Guthrie, C. (1998) *Cell* **92,** 315–326
5. Horowitz, D. S., Kobayashi, R., and Krainer, A. R. (1997) *RNA* **3,** 1374–1387
6. Teigelkamp, S., Achsel, T., Mundt, C., Göthel, S. F., Cronshagen, U., Lane, W. S., Marahiel, M., and Lührmann, R. (1998) *RNA* **4,** 127–141
7. Göthel, S. F., and Marahiel, M. A. (1999) *Cell. Mol. Life Sci.* **55,** 423–436
8. Baker, E. K., Colley, N. J., and Zuker, C. S. (1994) *EMBO J.* **13,** 4886–4895
9. Schreiber, S. L., and Crabtree, G. R. (1992) *Immunol. Today* **13,** 136–142
10. Zhao, Y., Chen, Y., Schutkowski, M., Fischer, G., and Ke, H. (1997) *Structure* **5,** 139–146
11. Lauber, J., Plessel, G., Prehn, S., Will, C. L., Fabrizio, P., Groning, K., Lane, W. S., and Lührmann, R. (1997) *RNA* **3,** 926–941
12. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol.* **276,** 307–326
13. Brünger, A. T. (1993) *X-PLOR*, Version 3.1, Yale University Press, New Haven, CT
14. Ke, H. (1992) *J. Mol. Biol.* **228,** 539–550
15. Lamzin, V. S., and Wilson, K. S. (1997) *Methods Enzymol.* **277,** 269–305
16. Jones, T. A., and Kjeldgaard, M. (1997) *Methods Enzymol.* **277,** 173–208
17. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Cryst.* **26,** 283–291
18. Smith, D. B., and Johnson, K. S. (1988) *Gene* **67,** 31–40
19. Taylor, P., Page, A. P., Kontopidis, G., Husi, H., and Walkinshaw, M. D. (1998) *FEBS Lett.* **425,** 361–366
20. Edwards, K. J., Ollis, D. L., and Dixon, N. E. (1997) *J. Mol. Biol.* **271,** 258–265
21. Konno, M., Ito, M., Hayano, T., and Takahashi, N. (1996) *J. Mol. Biol.* **256,** 897–908
22. Mikol, V., Kallen, J., and Walkinshaw, M. D. (1994) *Proc. Natl. Acad. Sci. U. S. A.* **91,** 5183–5186
23. Clubb, R. T., Ferguson, S. B., Walsh, C. T., and Wagner, G. (1994) *Biochemistry* **33,** 2761–2772
24. Ke, H., Zhao, Y., Luo, F., Weissman, I., and Friedman, J. (1993) *Proc. Natl. Acad. Sci. U. S. A.* **90,** 11850–11854
25. Mikol, V., Ma, D., and Carlow, C. K. (1998) *Protein Sci.* **7,** 1310–1316
26. Lambright, D. G., Sondek, J., Bohm, A., Skiba, N. P., Hamm, H. E., and Sigler, P. B. (1996) *Nature* **379,** 311–319
27. Wall, M. A., Coleman, D. E., Lee, E., Iniguez-Lluhi, J. A., Posner, B. A., Gilman, A. G., and Sprang, S. R. (1995) *Cell* **83,** 1047–1058
28. Sondek, J., Bohm, A., Lambright, D. G., Hamm, H. E., and Sigler, P. B. (1996) *Nature* **379,** 369–374
29. Gaudet, R., Bohm, A., and Sigler, P. B. (1996) *Cell* **87,** 577–588
30. Ayadi, L., Callebaut, I., Saguez, C., Villa, T., Mornon, J. P., and Banroques, J. (1998) *J. Mol. Biol.* **284,** 673–687
31. Reuter, K., and Ficner, R. (1999) *Acta Crystallogr. Sec. D* **55,** 888–890
32. Reuter, K., Nottrott, S., Fabrizio, P., Lührmann, R., and Ficner, R. (1999) *J. Mol. Biol.* **294,** 515–525
33. Kraulis, P. J. (1991) *J. Appl. Crystallogr.* **24,** 946–950
34. Merrit, E. A., and Bacon, D. J. (1997) *Methods Enzymol.* **277,** 505–524
35. Humphrey, W., Dalke, A., and Schulten, K. (1996) *J. Mol. Graphics* **14,** 33–38