Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091

Structure of the pseudouridine synthase RsuA from *Haemophilus influenzae*

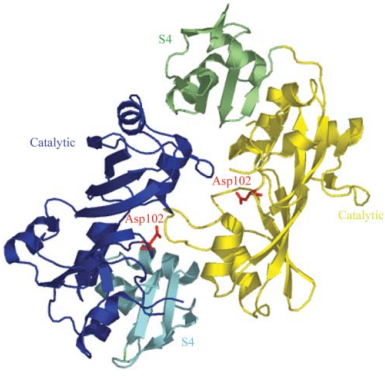
Allan Matte,[a,b] Gordon V. Louie,[c] J. Sivaraman,[d] Miroslaw Cygler[a,b,e] and Stephen K. Burley[c]*

[a]Biotechnology Research Institute, 6100 Royalmount Avenue, Montreal, Quebec H4P 2R2, Canada, [b]Montreal Joint Center for Structural Biology, Montreal, Quebec, Canada, [c]Structural GenomiX Inc., 10505 Roselle Street, San Diego, CA 92121, USA, [d]Department of Biological Sciences, National University of Singapore, 14 Science Drive, 117543 Singapore, Singapore, and [e]Department of Biochemistry, McGill University, Montreal, Quebec, Canada

Correspondence e-mail: mirek.cygler@bri.nrc.ca, sburley@stromix.com

Received 10 December 2004
Accepted 23 February 2005
Online 12 March 2005

PDB Reference: RsuA, 1vio, r1viosf.



The structure of the pseudouridine synthase RsuA from *Haemophilus influenza*, which catalyzes the conversion of uridine to pseudouridine at a single position within 16S ribosomal RNA, has been determined at 1.59 Å resolution and compared with that of *Escherichia coli* RsuA. The *H. influenza* enzyme contains an N-terminal S4-like $\alpha_3\beta_4$ domain followed by a catalytic domain, as observed in the structure of *E. coli* RsuA. Whereas the individual domains of *E. coli* and *H. influenza* RsuA are structurally similar, their relative spatial disposition differs greatly between the two structures. The former displays an extended open conformation with no direct contacts between the domains, while the latter is in a closed conformation with a large interface between the two domains. Domain closure presents several basic and polar residues into a putative RNA-binding cleft. It is proposed that this relative repositioning of the S4 and catalytic domains is used to modulate the shape and size of the rRNA-binding site in RsuA and in other pseudouridine synthases possessing S4 domains.

1. Introduction

Pseudouridine (5-ribosyluracil; ψ) is among the most abundant of RNA modifications (Ofengand, 2002). Uridine is converted to pseudouridine by pseudouridine synthases, enzymes that act on tRNA, rRNA and snoRNA. A number of such enzymes from bacterial sources have been structurally characterized. In *Escherichia coli*, the pseudouridine synthase RsuA (Wrzesinski *et al.*, 1995) catalyzes the formation of the single ψ present at position 516 of 16S rRNA (Bakin *et al.*, 1994). This enzyme possesses a modular organization, with an N-terminal S4 $\alpha_3\beta_4$ domain (Aravind & Koonin, 1999) tethered by a flexible polypeptide linker to the catalytic domain, which contains the catalytic Asp residue (Asp102; Sivaraman *et al.*, 2002).

Amino-acid sequence comparisons have been used to detect S4 domains within various members of the RsuA and RluA ψ-synthase and tyrosyl-tRNA synthetase families, some predicted RNA methyltransferases and a number of proteins of uncharacterized function (Aravind & Koonin, 1999). The S4 designation refers to the presence of this $\alpha_3\beta_4$ domain in the small-subunit ribosomal protein S4, which has been structurally characterized both as an isolated Δ41-residue truncation (Davies *et al.*, 1998, Markus *et al.*, 1998) and as a component of the 30S ribosomal subunit (Schluenzen *et al.*, 2000; Wimberly *et al.*, 2000). Full-length S4 protein (209 residues) consists of three domains: an N-terminal zinc-finger domain, an α_4 domain and the $\alpha_3\beta_4$ domain. Other proteins possessing comparable $\alpha_3\beta_4$ domains include the RNA-binding domain of heat-shock protein 15 (Staker *et al.*, 2000) and the ETS domain found in some winged-helix DNA-binding transcription factors (Davies *et al.*, 1998).

The S4 $\alpha_3\beta_4$ domain is thought to function as an RNA-binding motif (Aravind & Koonin, 1999). Structural studies of a number of RNA-modification enzymes containing the S4 domain have documented that this domain is typically connected to other globular domains by flexible linkers. In two recently published ψ-synthase crystal structures, the S4 domain is either entirely disordered (Del Campo *et al.*, 2004, Mizutani *et al.*, 2004) or was removed to produce diffraction-quality crystals (Corollo *et al.*, 1999; Sivaraman *et al.*, 2004; Mizutani *et al.*, 2004). The structures of tyrosyl-tRNA synthetase from *Bacillus stearothermophilus* and *Staphylococcus aureus* (Brick *et*

al., 1989; Qiu *et al.*, 2001) show the S4 domain in a disordered state, whereas the tyrosyl-tRNA synthetase structure from *Thermus thermophilus* (Yaremchuk *et al.*, 2002) possesses an ordered S4 domain. In the X-ray structures of both *E. coli* RsuA and *T. thermophilus* tyrosyl-tRNA synthetase, the S4 domain adopts multiple ordered conformations relative to the catalytic domain, presumably as a consequence of crystal-packing interactions.

Here, we report the determination of the crystal structure of RsuA from *Haemophilus influenzae*. When compared with the structure of *E. coli* RsuA, *H. influenzae* RsuA is markedly different in the relative positioning of the S4 and catalytic domains. The more compact tertiary structure of the *H. influenzae* protein underscores the relative flexibility of the two domains of RsuA and suggests that spatial mobility of the S4 domain itself may contribute to formation of the RNA-recognition site.

2. Materials and methods

2.1. Cloning, purification and crystallization

The coding sequence for RsuA from *H. influenzae* was cloned from bacterial genomic DNA. A C-terminal hexahistidine-bearing fusion protein was expressed in *E. coli* strain BL21(DE3) from a pET26b-derived expression vector. Full-length RsuA protein was purified by Ni–NTA affinity (Qiagen) and gel-filtration (Pharmacia HiLoad 16/60 Superdex 75) chromatographies and concentrated to 10–30 mg ml^{-1} in a storage buffer containing 10 m*M* HEPES pH 7.5, 150 m*M* NaCl, 10 m*M* methionine, 10%(*v*/*v*) glycerol and 1 m*M* β-mercaptoethanol.

2.2. Crystallization

Crystals of *H. influenzae* RsuA were grown at 277 K by hanging-drop vapour diffusion. A drop containing 1 µl protein in buffer was mixed with 1 µl reservoir solution (2.0 *M* ammonium sulfate, 0.1 *M* sodium acetate pH 4.5) and equilibrated over reservoir solution. The crystals belong to the primitive monoclinic space group $P2_1$ (unit-cell parameters a = 57.2, b = 77.0, c = 62.8 Å, β = 108.6°), with two molecules per asymmetric unit. A cryoprotectant consisting of reservoir solution supplemented with 20%(*v*/*v*) 1,4-butanediol was used to permit direct immersion of the crystals in liquid nitrogen.

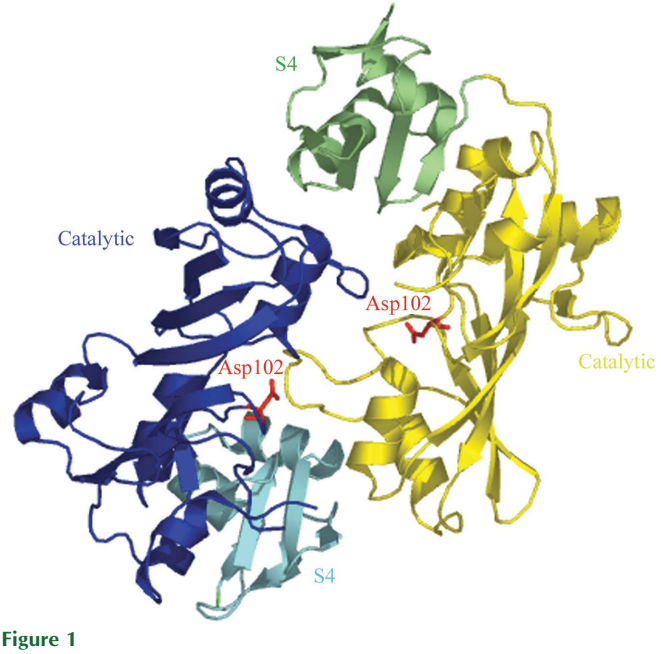


Figure 1
The structure of *H. influenzae* RsuA, showing the head-to-tail orientation of both molecules comprising the asymmetric unit. The two molecules are coloured blue and cyan or yellow and green for the catalytic and $\alpha_3\beta_4$ S4 domains, respectively. The catalytic residue (Asp102, red) is shown in atomic stick figure representation. The figure was prepared using *PyMol* (DeLano, 2002).

Table 1
Data-collection and refinement statistics.

Data collection	
Resolution range† (Å)	22–1.51 (1.59–1.51)
No. of measured reflections	587441
No. of unique reflections	80195
Completeness† (%)	99.6 (97.4)
Redundancy†	7.3 (6.2)
$I/\sigma(I)$†	11.5 (1.8)
R_{sym}†	0.087 (1.109)
Refinement	
Resolution range (Å)	22.1–1.59
No. of reflections	69152
R_{work} (No. of reflections)	0.196 (65715)
R_{free} (No. of reflections)	0.218 (3437)
B factors (Å^2) (No. of atoms)	
Protein atoms	19.7 (3656)
Water molecules	32.0 (241)
Heteroatoms	37.9 (24)
R.m.s.d. bond lengths (Å)	0.004
R.m.s.d. bond angles (°)	1.45
Ramachandran plot	
Most favoured (%)	92.9
Disallowed (%)	0.0

† Values in parentheses are for the highest shell.

2.3. Structure determination and refinement

X-ray diffraction data were collected at the Advanced Photon Source (Argonne National Laboratory) with the SGX-CAT beamline under standard cryogenic conditions at an X-ray wavelength of 0.9794 Å. Data were indexed, integrated and scaled with *MOSFLM* and *SCALA* (*CCP*4; Winn *et al.*, 2002). The structure was determined by molecular replacement with *EPMR* (Kissinger *et al.*, 1999) using the S4 and catalytic domains from the structure of *E. coli* RsuA (PDB code 1ksk) as the search model. Electron-density map interpretation and model building were performed using *XtalView* (McRee, 1999). The program *Crystallography and NMR Explorer* (*CNX*; Accelyrs, San Diego, USA) was employed for structure refinement. The final refinement model has a crystallographic R factor of 0.196 (R_{free} = 0.218) for all reflections at 1.59 Å resolution, with excellent stereochemistry (Badger & Hendle, 2002). Final data-collection and refinement statistics are presented in Table 1.

3. Results and discussion

3.1. Structure description

Our 1.59 Å resolution X-ray structure of *H. influenzae* RsuA revealed the expected overall structural similarity to other pseudouridine synthases, in particular *E. coli* RsuA (Sivaraman *et al.*, 2002), with which it shares 59% identity at the amino-acid sequence level. The closed conformational state of *H. influenzae* RsuA is observed in both molecules comprising the asymmetric unit, which are associated tightly to form a symmetric head-to-tail dimer (Fig. 1). Indeed, this dimeric arrangement, in which the N-terminal domain of one monomer occupies the RNA-binding cleft of the opposing monomer, may serve to stabilize the closed conformation of each monomer. Although this tight packing is suggestive of a stable dimer, several lines of evidence suggest that the observed dimer is not physiologi-

cally relevant. Firstly, very few hydrogen-bonding interactions exist between the two molecules. Secondly, and more importantly, the active-site clefts and catalytic residues (Asp102) are sterically blocked within the dimer and would be inaccessible to the rRNA substrate. A monomeric form of *H. influenzae* RsuA is consistent with those of other rRNA-specific ψ-synthases, including both *E. coli* RsuA (Sivaraman *et al.*, 2002) and RluD (Sivaraman *et al.*, 2004). Finally, there is no evidence of dimer formation in solution as judged by gel-filtration chromatography of *H. influenzae* RsuA (data not shown).

The S4 [residues Ser(−1)–Trp56] and catalytic (residues Gln61–Glu226) domains of *H. influenzae* RsuA (PDB code 1vio) and *E. coli* RsuA (PDB code 1ksk; Sivaraman *et al.*, 2002) differ by only 0.7 and 0.8 Å (calculated root-mean-square deviations for common C^{α}-atomic positions), respectively. However, these two RsuA structures do differ significantly in their relative positioning of the S4 and catalytic domains. In *E. coli* RsuA, the S4 domain extends away from the catalytic domain, with no direct or indirect non-covalent intramolecular interdomain interactions (Sivaraman *et al.*, 2002; Fig. 2). In contrast, in our structure of *H. influenzae* RsuA, a 90° rotation within the linker segment positions the S4 domain much closer to the catalytic domain (Fig. 2) and several side-chain hydrogen bonds are observed between the two domains (Table 2). The specific residues participating in hydrogen bonding are not highly conserved among RsuA family members, but the corresponding residues from a multiple sequence alignment retain the characteristics of being polar or charged. In the closed conformation of *H. influenzae* RsuA, the β-sheet of the S4 domain is essentially in register with the large central β-sheet of the catalytic domain (Fig. 1). Most importantly, as discussed below, this conformational difference substantially alters both the size and shape of the putative rRNA-recognition cleft relative to that observed in *E. coli* RsuA (compare Figs. 2*b* and 2*c*). The overall electrostatic properties of the RsuA RNA-binding clefts are not affected significantly by the conformational differences observed between the *E. coli* and *H. influenzae* enzymes.

3.2. Functional implications

Although participation of the S4 domain in RNA binding is well established, the precise function of this domain in bacterial ψ-synthases has yet to be defined. Only five of the 11 ψ-synthases of *E. coli* possess S4 domains (RsuA, RluB and RluF of the RsuA superfamily, and RluC and RluD of the RluA superfamily; Sivaraman *et al.*, 2002). In RluC, deletion of the S4 domain does not abolish pseudouridine synthase activity; however, the specific activity and specificity of the truncated form of this enzyme relative to its full-length form have not been characterized in detail (Corollo *et al.*, 1999). Sivaraman *et al.* (2002) have suggested that this domain functions primarily as a RNA-recognition module and serves to tether the catalytic module at a specific site within the rRNA in the vicinity of the target uridine. It is also clear, however, that other

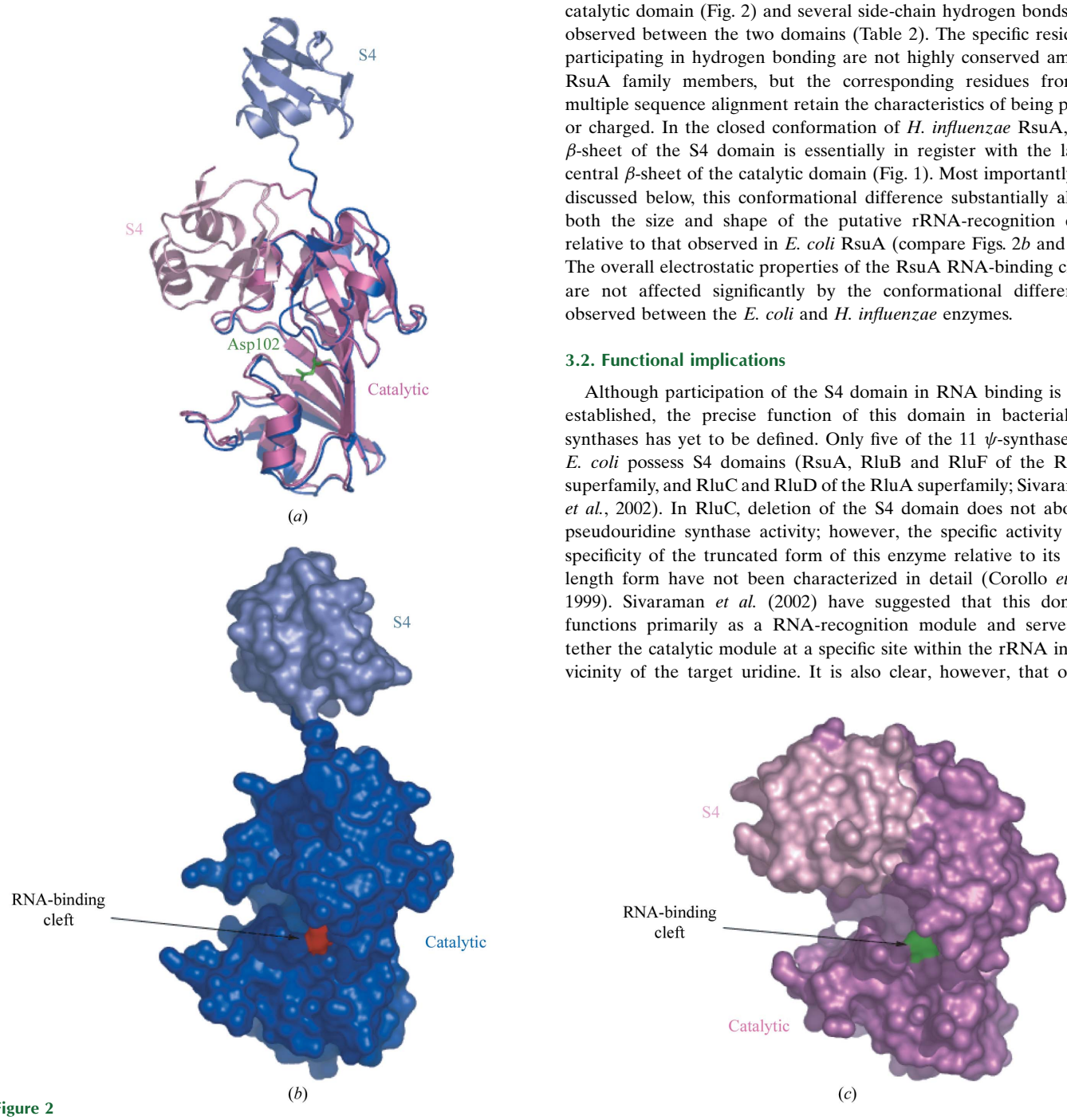


Figure 2
Structural comparisons of *E. coli* and *H. influenzae* RsuA. (*a*) C^{α} superposition of the catalytic domains (residues 60–231) of *E. coli* RsuA (blue) and *H. influenzae* RsuA (magenta). A shift of the $\alpha_3\beta_4$ S4 domains of *E. coli* RsuA (slate) relative to *H. influenzae* RsuA (pink) of about 90° is observed. (*b*) Molecular-surface representation of *E. coli* RsuA and (*c*) *H. influenzae* RsuA using the same orientation as in (*a*). The locations of the catalytic residue (Asp102) in the two enzymes are highlighted in red and green, respectively.

Table 2
Hydrogen-bonding interactions between the N-terminal and catalytic domains of *H. influenzae* RsuA.

N-terminal domain	Catalytic domain	Distance (Å)
Asn10 O	Gln116 N	3.1
Glu50 O	Arg100 NH1	2.8
Glu50 O	Arg100 NH2	2.9
Glu52 OE1	Ser97 N	3.0
Glu52 OE1	Ser97 OG	3.2
Glu52 OE2	Arg100 NE	3.0
Glu52 OE2	Ser97 OG	2.8
Leu53 O	His96 ND1	3.2
Thr55 OG1	Lys94 O	2.9
Trp56 O	Glu58 N	3.3

mechanisms for recognition of the specific uridine targeted for modification almost certainly exist (*e.g.* the shape and size of the RNA-binding cleft adjacent to the catalytic site of each ψ-synthase; Sivaraman *et al.*, 2004).

In productive complexes of both ribosomal S4 protein with 16S rRNA (Wimberly *et al.*, 2000) and *T. thermophilus* tyrosyl-tRNA synthetase with tRNA (Yaremchuk *et al.*, 2002), interaction of the S4 $\alpha_3\beta_4$ domain with RNA involves a region of the S4 domain near α-helices 1 and 2 and the intervening loop. By analogy, residues of the *H. influenzae* RsuA S4 domain responsible for binding the rRNA substrate would include Lys5 and Glu9 (α-helix 1) plus Arg15, Lys20, Arg23 and Gln24 (α-helix 2) and Lys35, Glu50 and Asp51 (Fig. 3). A number of these residues (in particular the basic amino acids Arg2, Lys5, Arg23 and Lys35) are highly conserved within the RsuA family. As a consequence of the proximity of the S4 and catalytic domains

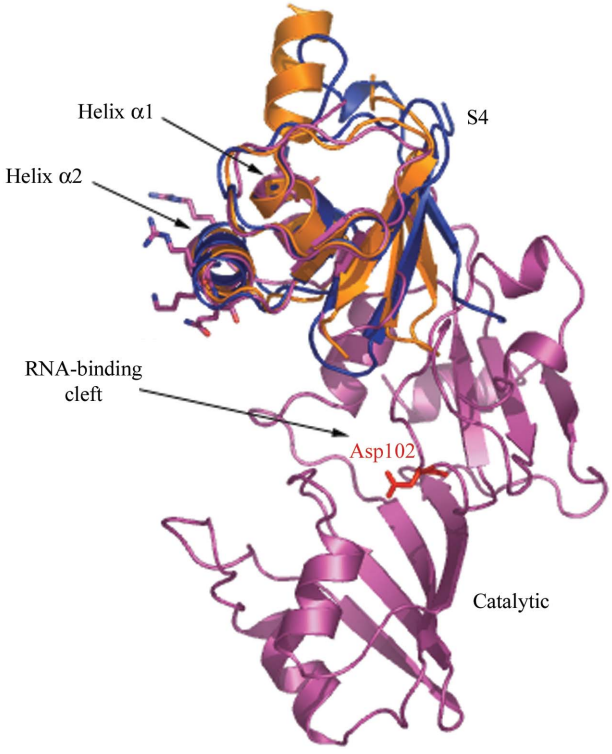


Figure 3
Superposition of $\alpha_3\beta_4$ S4 domains of *H. influenzae* RsuA (magenta; PDB code 1vio), tyrosyl-tRNA synthetase (blue; PDB code 1h3e) and the S4 protein of the 30S ribosome (orange; PDB code 1fjg). The two α-helices, α1 and α2, of the S4 domain involved in RNA binding are indicated. Residues belonging to either α1 (Lys5 and Glu9) or α2 (Arg15, Lys20, Arg23 and Gln24) of RsuA that may contribute directly to RNA-substrate binding are shown.

observed in the closed conformation of *H. influenzae* RsuA, several key residues of the S4 domain, including Arg23, Lys28 and Lys35, occupy part of the rRNA-recognition cleft. Thus, the S4 domain of RsuA may itself influence the shape and chemical character of the rRNA-binding cleft and thereby participate directly in rRNA binding and recognition of nucleic acid features in the vicinity of the modified base. Definitive verification of this predicted role of the S4 domain would require structure determination of RsuA bound to one or more fragments of the 16S-rRNA substrate. A molecular mechanism of action involving the S4 domain modulation of the affinity and specificity of the RNA-recognition site may also be important for other ψ-synthases and RNA-modifying enzymes containing this domain.

Recently, adoption of both closed and open conformations by the catalytic domain itself was recognized in the context of the crystal structure of TruB from *Mycobacterium tuberculosis* (Chaudhuri *et al.*, 2004). In the *M. tuberculosis* TruB structure, a hinge-bending motion of the central β–sheet was observed in the absence of substrate, with the net effect of changing the shape of the catalytic cleft. It has also been documented that RNA binding to ψ-synthases can induce conformational changes in the enzyme (Pan *et al.*, 2003; Phannachet & Huang, 2004).

This research was supported in part by the Canadian Institutes of Health Research (CIHR) through grant (200103GSP-90094-GMX-CFAA-19924) to MC and an Academic Research Grant from the National University of Singapore (JS).

References

Aravind, L. & Koonin, E. V. (1999). *J. Mol. Evol.* **48**, 291–302.
Badger, J. & Hendle, H. (2002). *Acta Cryst.* D**58**, 284–291.
Bakin, A., Kowalak, J., McCloskey, J. A. & Ofengand, J. (1994). *Nucleic Acids Res.* **22**, 3681–2684.
Brick, P., Bhat, T. N. & Blow, D. M. (1989). *J. Mol. Biol.* **208**, 83–98.
Chaudhuri, B. N., Chan, S., Perry, L. J. & Yeates, T. O. (2004). *J. Biol. Chem.* **279**, 24585–24591.
Corollo, D., Blair-Johnson, M., Conard, J., Fiedler, T., Sun, D., Wang, L., Ofengand, J. & Feena, R. (1999). *Acta Cryst.* D**55**, 302–304.
Davies, C., Gerstner, R. B., Draper, D. E., Ramakrishnan, V. & White, S. W. (1998). *EMBO J.* **17**, 4545–4558.
DeLano, W. L. (2002). *The PyMOL Molecular Graphics System*, http://www.pymol.org.
Del Campo, M., Ofengand, J. & Malhotra, A. (2004). *RNA*, **10**, 231–239.
Kissinger, C. R., Gehlhaar, D. K. & Fogel, D. B. (1999). *Acta Cryst.* D**55**, 484–491.
McRee, D. E. (1999). *J. Struct. Biol.* **125**, 156–165.
Markus, M. A., Gerstner, R. B., Draper, D. E. & Torchia, D. A. (1998). *EMBO J.* **17**, 4559–4571.
Mizutani, K., Machida, Y., Unzai, S., Park, S.-Y. & Tame, J. R. H. (2004). *Biochemistry*, **43**, 4454–4463.
Ofengand, J. (2002). *FEBS Lett.* **514**, 17–23.
Pan, H., Agarwalla, S., Moustakas, D. T., Finer-Moore, J. & Stroud, R. M. (2003). *Proc. Natl Acad. Sci. USA*, **100**, 12648–12653.
Phannachet, K. & Huang, R. H. (2004). *Nucleic Acids Res.* **32**, 1422–1429.
Qiu, X., Janson, C. A., Smith, W. W., Green, S. M., McDevitt, P., Johanson, K., Carter, P., Hibbs, M., Lewis, C., Chalker, A., Fosberry, A., Lalonde, J., Berge, J., Brown, P., Houge-Frydrych, C. S. V. & Jarvest, R. L. (2001). *Protein Sci.* **10**, 2008–2016.
Schluenzen, F., Tocilj, A., Zarivach, R., Harms, J., Gluehmann, M., Janell, D., Bashan, A., Bateis, H., Agmon, I., Franceschi, F. & Yonath, A. (2000). *Cell*, **102**, 615–623.
Sivaraman, J., Iannuzzi, P., Cygler, M. & Matte, A. (2004). *J. Mol. Biol.* **335**, 87–101.
Sivaraman, J., Sauvé, V., Larocque, R., Stura, E. A., Schrag, J. D., Cygler, M. & Matte, A. (2002). *Nature Struct. Biol.* **9**, 353–358.
Staker, B. L., Korber, P., Bardwell, J. C. A. & Saper, M. A. (2000). *EMBO J.* **19**,

749–757.
Wimberly, B. T., Brodersen, D. E., Clemons, W. M. Jr, Morgan-Warren, R. J., Carter, A. P., Vornhein, C., Hartsch, T. & Ramakrishnan, V. (2000). *Nature (London)*, **407**, 327–339.
Winn, M. D., Ashton, A. W., Briggs, P. J., Ballard, C. C. & Patel, C. (2002). *Acta Cryst.* D**58**, 1929–1936.
Wrzesinski, J., Bakin, A., Nurse, K., Lane, B. G. & Ofengand, J. (1995). *Biochemistry*, **34**, 8904–8913.
Yaremchuk, A., Kriklivyi, I., Tukalo, M. & Cusack, S. (2002). *EMBO J.* **21**, 3829–3840.