Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Crystallization of Hfq protein: a bacterial gene-expression regulator

Ioulia M. Vassilieva,[a] Alexey D. Nikulin,[a] Udo Blasi,[b] Isabella Moll[b] and Maria B. Garber[a]*

[a]Institute of Protein Research, Russian Academy of Sciences, Pushchino, Russia, and [b]Institute of Microbiology and Genetics, Vienna Biocentre, Austria

Correspondence e-mail: garber@vega.protres.ru

Hfq protein from *Escherichia coli* (EcoHfq) has been overproduced in *E. coli*, purified to homogeneity and crystallized using the hanging-drop vapour-diffusion technique. Crystallization conditions for EcoHfq were found which yielded X-ray quality crystals. Crystals of EcoHfq and of Cd-, Hg- and Se-containing derivatives grew in two months, with unit-cell parameters $a = b = 127.41$, $c = 170.36$ Å. The crystals belong to space group $I4$ and diffract to 2.1 Å resolution. Two hexamers are predicted per asymmetric unit.

Received 21 November 2002
Accepted 26 March 2003

© 2003 International Union of Crystallography
Printed in Denmark – all rights reserved

1. Introduction

Hfq is a highly conserved thermostable bacterial protein (for a review, see Vassilieva & Garber, 2002). It is encoded by the *hfq* gene mapped at 94.8 min on the *E. coli* chromosome (Kajitani & Ishihama, 1991) and located in the *mutL-miaA* superoperon (Tsui & Winkler, 1994). Hfq is an abundant RNA-binding protein that was first identified as a host factor required for plus-strand replication of the Qβ RNA bacteriophage (Franze de Fernandez *et al.*, 1972). In uninfected bacterial cells Hfq appears to work as a regulator of an increasing number of genes, including DNA-repair pathway genes (Tsui *et al.*, 1997). Hfq binds a variety of small regulatory RNAs, known as 'riboregulators', and alters their activity (Sledjeski & Gottesman, 1996; Zhang *et al.*, 1998; Majdalani *et al.*, 1998; Wassarman *et al.*, 2001). The synthesis of more than 30 proteins is affected in the *hfq* null mutant (Muffler *et al.*, 1997). Bacterial Hfq has been shown to exist in a hexameric form in solution (Kajitani *et al.*, 1994; Vassilieva *et al.*, 2002).

Hfq is a virulence factor of many pathogenic bacteria. In *Pseudomonas aeruginosa*, Hfq (PaeHfq) is involved in regulation of virulence-factor synthesis during infection. Hfq protein from *P. aeruginosa* has been shown to functionally replace Hfq protein from *E. coli* (EcoHfq) in terms of its requirement for phage Qβ replication and for *rpoS* expression (Sonnleitner *et al.*, 2002).

Recently, Hfq has been reported to be a bacterial Sm-like protein that mediates RNA–RNA interaction (Zhang *et al.*, 2002; Møller *et al.*, 2002). Sm proteins are known to be RNA-binding proteins that generate complexes that participate in the spliceosome (Will & Lührmann, 2001). Hfq protein facilitates RNA–RNA complex formation by acting as an RNA chaperone. Crystal structures of human Sm proteins, archaeal Sm-like proteins and a bacterial Hfq protein from *Staphylococcus aureus* have recently been determined (Kambach *et al.*, 1999; Törö *et al.*, 2001; Schumacher *et al.*, 2002). Despite sequence homology in the N-terminal region between the *S. aureus* and *E. coli* Hfq proteins, several attempts to model the three-dimensional structure of Hfq from *E. coli* have been made previously (Arluison *et al.*, 2002; Sun *et al.*, 2002). Nevertheless, the structure of EcoHfq has not yet been determined experimentally. There is high interest in the three-dimensional structure of Hfq protein from *E. coli* owing to the fact that the majority of functional experiments have been made using EcoHfq.

It is necessary to understand better the mechanism of Hfq-mediated RNA–RNA complex formation and protein–protein interactions. The crystallization and preliminary X-ray analysis of Hfq protein from *E. coli* reported here are prerequisites for the determination and further analysis of the EcoHfq structure.

2. Materials and methods

2.1. Gene cloning and expression

The *hfq* gene from *E. coli* was cloned into the pET11a-PL expression vector as described in Vassilieva *et al.* (2002). *Nde*I and *Bam*HI restriction sites were used and *E. coli* DH5α strain was transformed by the resultant plasmid. After PCR screening, the plasmid with the correct-sized insert was checked by direct sequencing. *E. coli* strain BL21(DE3) was transformed by the recombinant plasmid and plated on LB agar with 100 µg ml^{-1} ampicillin. 3 ml of LB media containing antibiotics was infected by several colonies and the culture was grown overnight at 310 K. This overnight culture was added to 1 l of LB media containing 100 µg ml^{-1} ampicillin and incubated in a shaker (160 rev min^{-1}) at 310 K

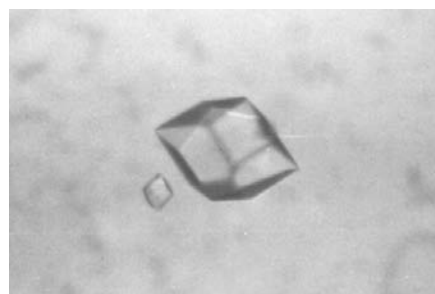
until OD_{590} reached 0.5. At this point, protein production was induced by 0.5 m*M* isopropyl-β-D-thiogalactopyranoside. After 3 h the cells were harvested by centrifugation (8000 rev min^{-1}, Beckman JA-10 rotor, 277 K, 10 min).

E. coli strain B834(DE3) (a methionine auxotroph) was transformed by the recombinant plasmid and grown on a synthetic medium containing selenomethionine (SeMet) in order to produce SeMet-EcoHfq. For details, see Vassilieva *et al.* (2002).

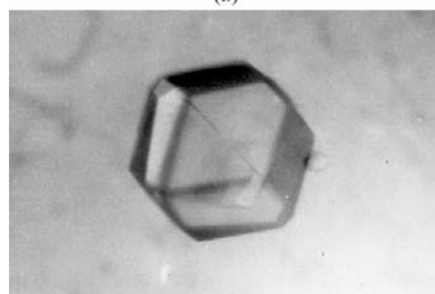
2.2. Protein purification

The purification of EcoHfq was performed as described in Vassilieva *et al.* (2002). 4 g of wet cells was suspended in 20 ml 50 m*M* Tris–HCl pH 8.0, 200 m*M* $MgCl_2$, 1.5 *M* NaCl, 5 m*M* EDTA and disrupted by sonication. Cellular membranes were removed by centrifugation for 20 min at 14 000*g*. Proteins from the host strain were removed by heating at 356 K for 20 min and subsequent centrifugation for 40 min at 14 000*g*. Ammonium sulfate was added to the supernatant to a final concentration of 1.4 *M* and it was loaded onto a butyl-Toyopearl column. After washing the column with solution containing 1.7 *M* ammonium sulfate and 1.5 *M* NaCl, the protein was eluted with a linear gradient of 1.7–0 *M* ammonium sulfate in 1.5 *M* NaCl, 50 m*M* Tris–HCl pH 8.0. The fractions which were homogeneous by UV spectra and Coomassie-blue stained SDS–PAGE analysis were pooled and concentrated to 25 mg ml^{-1}.

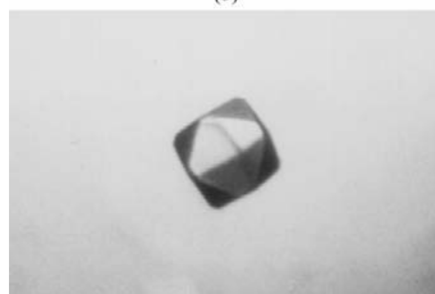
SeMet-containing EcoHfq protein was purified using the same protocol as described above without any modifications.



(*a*)



(*b*)



(*c*)

Figure 1
Crystals of Hfq protein from *E. coli* (EcoHfq): (*a*) a crystal of native protein, (*b*) a heavy-atom derivative crystal containing Se atoms, (*c*) a heavy-atom derivative crystal containing Hg atoms.

Table 1
Statistics of data collection from EcoHfq crystals.

Values given in parentheses correspond to the highest resolution shell.

Parameter	Native EcoHfq	Cd-EcoHfq	Se-EcoHfq	Hg-EcoHfq
Space group	*I*4	*I*4	*I*4	*I*4
Unit-cell parameters (Å)	*a* = *b* = 127.41, *c* = 170.36	*a* = *b* = 128.69, *c* = 171.36	*a* = *b* = 127.21, *c* = 170.34	*a* = *b* = 127.99, *c* = 170.49
Mosaicity	0.203	0.212	0.136	0.281
Beamline	BW7A, DESY	X11, DESY	X11, DESY	X11, DESY
Wavelength (Å)	0.93300	0.81100	0.81100	0.81100
Resolution range (Å)	25–2.42	20–2.1	20–2.2	20–3.3
No. of measured reflections	211548	569167	636089	133107
No. of unique reflections	51255	80266	66630	20578
Completeness (%)	95.2	94.7	95.0	98.6
Multiplicity	4	7	10	6
R_{merge} (%)	4.5 (16.7)	7.5 (29.9)	9.7 (39.7)	17.9 (29.4)
$I/\sigma(I)$	20.5 (6.6)	15.6 (5.4)	9.5 (3.9)	15 (6.4)

2.3. Crystallization

Initial screening for crystallization conditions was performed using commercially purchased sparse-matrix screens (Jancarik & Kim, 1991; Cudney *et al.*, 1994) from Hampton Research (Crystal Screens I and II) with the hanging-drop vapour-diffusion technique. All drops were set up by mixing 2.5 µl of protein dialyzed into 20 m*M* Tris–HCl pH 8.0, 100 m*M* NaCl, 100 m*M* ammonium sulfate with 2.5 µl of screen solution on siliconized cover slides and were equilibrated against 1.0 ml of the same precipitant solution. Crystallization plates were incubated at 294 K. Crystallization conditions were optimized in further experiments to obtain perfect X-ray quality crystals.

EcoHfq formed needles in several Crystal Screen II solutions containing ammonium sulfate and microcrystals in Crystal Screen II solution No. 15 containing 0.5 *M* ammonium sulfate, 1.0 *M* lithium sulfate, 0.1 *M* trisodium citrate pH 5.6. In further experiments, lithium sulfate was substituted for ammonium sulfate at the same concentration and the ratio of protein solution to precipitant solution in the drops was changed to 1:1.5. This optimization of the conditions yielded single crystals of the native EcoHfq protein of 0.37 × 0.25 × 0.25 mm in size that were suitable for diffraction analysis (Fig. 1*a*) and also their Se (Fig. 1*b*), Hg (Fig. 1*c*) and Cd derivatives. Se-derivative crystals were obtained by crystallization of selenomethionine-containing EcoHfq protein. Hg-derivative and Cd-derivative crystals were obtained by crystallizing EcoHfq protein in the presence of equimolar amounts of sodium mersalyl and 3 m*M* $CdCl_2$, respectively. All other conditions were kept constant.

2.4. Data collection

Native, SeMet, Cd- and Hg-derivative data sets of EcoHfq were collected at the EMBL beamlines BW7A and X11, DESY, Hamburg using MAR CCD detectors. All data sets were collected at 100 K using the flash-soaking method with ethylene glycol as a cryoprotectant [the cryosolution contained 15% ethylene glycol, 750 m*M* $(NH_4)_2SO_4$, 50 m*M* trisodium citrate pH 5.6]. The wavelength was 0.93300 Å for native EcoHfq crystals and 0.81100 Å for the heavy-atom derivatives. The rotation step was 1° and the exposure time was 4 min in all cases. All data sets were processed with the *XDS* package (Kabsch, 2001). Details of the experiments are summarized in Table 1.

3. Results and discussion

EcoHfq protein has been highly purified. The most complicated problem causing initial crystallization failure was the presence of RNA impurities in the protein samples. These impurities were eliminated by the inclusion of a hydrophobic chromatography step in the presence of high concentrations of ammonium sulfate and

sodium chloride in the last step of the purification procedure. The absence of RNA in protein samples was monitored by UV spectroscopy (data not shown).

EcoHfq crystallized in the initial experiments. Optimization of the crystallization conditions led to crystals that diffracted to 2.1 Å resolution. The solvent content of the native protein crystals is 52% and the Matthews coefficient V_M is 2.615 $Å^3\,Da^{-1}$ (Matthews, 1968). Two hexamers are predicted per asymmetric unit. Native crystals and their heavy-atom derivatives belong to the same space group *I*4. Statistics of the collection and processing of the data from EcoHfq crystals are presented in Table 1. Structure determination of EcoHfq is in progress.

The crystal structure of a homologous bacterial Hfq protein from *S. aureus* (SauHfq) has recently been published (Schumacher *et al.*, 2002). The crystals belonged to space group $P2_1$ and diffracted X-rays to 1.55 Å resolution. Our data suggest that EcoHfq packs in crystals in a different manner. This difference in crystal packing between EcoHfq and SauHfq molecules arises from their low sequence homology and different lengths (EcoHfq is half as long again in the C-terminus as SauHfq). The C-terminal extension of EcoHfq may be disordered or folded in a special way and may modify the overall structure of the molecule. Determination of the three-dimensional structure of Hfq protein from *E. coli* will provide the best answer to the questions that arise and will explain the features of EcoHfq that have been observed in many functional experiments.

We are grateful to Nina A. Matveeva for very helpful technical assistance. We thank the staff of the EMBL Outstation in Hamburg for assistance during data collection. This work was supported by the Russian Academy of Sciences and the Russian Foundation for Basic Research (grant No. 02-04-48536). The research of MBG was supported in part by an International Research Scholar's award from the Howard Hughes Medical Institute (grant No. 55000308).

References

Arluison, V., Derreumaux, P., Allemand, F., Folichon, M., Hajnsdorf, E. & Regnier, P. (2002). *J. Mol. Biol.* **320**, 705–712.

Cudney, R., Patel, S., Weisgraber, K., Newhouse, Y. & McPherson, A. (1994). *Acta Cryst.* D**50**, 414–423.

Franze de Fernandez, M. T., Hayward, W. S. & August, J. T. (1972). *J. Biol. Chem.* **247**, 824–831.

Jancarik, J. & Kim, S.-H. (1991). *J. Appl. Cryst.* **24**, 409–411.

Kabsch, W. (2001). *International Tables for Crystallography*, Vol. F, edited by M. G. Rossmann & E. Arnold, pp. 730–734. Dordrecht: Kluwer Academic Publishers.

Kajitani, M. & Ishihama, A. (1991). *Nucleic Acids Res.* **19**, 1063–1066.

Kajitani, M., Kato, A., Wada, A., Inokuchi, Y. & Ishihama, A. (1994). *J. Bacteriol.* **176**, 531–534.

Kambach, C., Walke, S., Young, R., Avis, J. M., de La Fortelle, E., Raker, V. A., Lührmann, R., Li, J. & Nagai, K. (1999). *Cell*, **96**, 375–387.

Majdalani, N., Cunning, C., Sledjeski, D., Elliott, T. & Gottesman, S. (1998). *Proc. Natl. Acad. Sci. USA*, **95**, 12462–12467.

Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.

Møller, T., Franch, T., Højrup, P., Keene, D. R., Bächinger, H. P., Brennan, R. G. & Valentin-Hansen, P. (2002). *Mol. Cell*, **9**, 23–30.

Muffler, A., Traulsen, D. D., Fischer, D., Lange, R. & Hengge-Aronis, R. (1997). *J. Bacteriol.* **179**, 297–300.

Schumacher, M., Pearson, R. F., Møller, T., Valentin-Hansen, P. & Brennan, R. G. (2002). *EMBO J.* **21**, 3546–3556.

Sledjeski, D. D. & Gottesman, S. (1996). *EMBO J.* **15**, 3993–4000.

Sonnleitner, L., Moll, I. & Bläsi, U. (2002). *Microbiology*, **148**, 883–891.

Sun, X., Zhulin, I. & Wartell, R. M. (2002). *Nucleic Acids Res.* **30**, 3662–3671.

Törö, I., Thore, S., Mayer, C., Basquin, J., Seraphin, B. & Suck, D. (2001). *EMBO J.* **20**, 2293–2303.

Tsui, H.-C. T., Feng, G. & Winkler, M. E. (1997). *J. Bacteriol.* **179**, 7476–7487.

Tsui, H.-C. T. & Winkler, M. E. (1994). *Biochimie*, **76**, 1168–1177.

Vassilieva, I. M. & Garber, M. B. (2002). *Mol. Biol. (Moscow)*, **36**, 785–791.

Vassilieva, I. M., Rouzanov, M. V., Zelinskaya, N. V., Moll, I., Bläsi, U. & Garber, M. B. (2002). *Biochemistry (Moscow)*, **67**, 1566–1571.

Wassarman, K. M., Repoila, F., Rosenow, C. & Storz, G. (2001). *Genes Dev.* **15**, 1637–1651.

Will, C. L. & Lührmann, R. (2001). *Curr. Opin. Cell. Biol.* **13**, 290–301.

Zhang, A., Altuvia, S., Tiwari, A., Argaman, L., Hengge-Aronis, R. & Storz, G. (1998). *EMBO J.* **17**, 6061–6068.

Zhang, A., Wassarman, K. M., Ortega, J., Steven, A. C. & Storz, G. (2002). *Mol. Cell*, **9**, 11–22.