Biochemistry (Moscow), Vol. 67, No. 11, 2002, pp. 1293-1297. Translated from Biokhimiya, Vol. 67, No. 11, 2002, pp. 1566-1571.

Cloning, Purification, and Crystallization of a Bacterial Gene Expression Regulator—Hfq Protein from *Escherichia coli*

I. M. Vassilieva[1]*, M. V. Rouzanov[1], N. V. Zelinskaya[1], I. Moll[2], U. Bläsi[2], and M. B. Garber[1]

[1]*Institute of Protein Research, Russian Academy of Sciences, Pushchino, Moscow Region, 142290 Russia; fax: (095) 924-0493; E-mail: yuvas@vega.protres.ru*
[2]*Institute of Microbiology and Genetics, Vienna Biocenter, Dr. Bohrgasse 9, 1030 Vienna, Austria; fax: +43-1-4277-9546; E-mail: UDO@gem.univie.ac.at*

Received October 3, 2001
Revision received November 27, 2001

Abstract—Thermostable RNA-binding protein Hfq (also denoted HF1) is a multifunctional expression regulator of many bacterial genes. The regulation takes place both at a translation level (directly) and transcription level (indirectly through the stimulation of bacterial RNA polymerase σ^S-subunit translation). We have cloned and overexpressed the *hfq* gene from *E. coli* and developed a purification procedure for the protein. Using gel filtration and ultracentrifugation techniques it was shown that the obtained Hfq protein is highly homogeneous and well dissolved. It has been crystallized and can be used for structural investigations.

Key words: Hfq, HF1, RNA-binding protein, regulation of bacterial gene expression, *Escherichia coli*

Protein Hfq (also denoted HF1) is a highly conservative (Fig. 1) thermostable protein with molecular weight of 11.2 kD. It was discovered in the late 1970s as a cell protein of *Escherichia coli* involved in the replicative cycle of Qβ bacteriophage *in vitro* [1-3]. Later experiments *in vivo* have shown that Hfq melts the 3′-end of a positive phage RNA strain and provides the interaction of Qβ-replicase with the given RNA region and initiation of negative chain synthesis [4]. In non-infected bacterial cells Hfq protein acts as a multifunctional expression regulator for a large number of genes, including DNA reparation genes [5]. The protein functions both directly (binding with mRNA of the controlled gene and stimulating either its translation or poly(A) synthesis or degradation) and indirectly (stimulating the translation of mRNA of σ^S-subunit of RNA polymerase).

Protein Hfq is a positive expression regulator for the *rpoS* gene, encoding the σ^S (or σ^{38})-subunit of a bacterial RNA polymerase [6, 7]. σ^S-Subunit of RNA polymerase regulates the expression of genes that are induced during osmotic stress, low pH of the medium, shortage of oxygen, heat stress, starvation, or at the transition to the stationary phase of cell growth [8-10]. Furthermore, this subunit is a virulence factor of several pathogenic bacteria species, including *Salmonella typhimurium* [11]. It is assumed that Hfq stimulates the translation of *rpoS* mRNA, altering the secondary structure of a ribosome binding site [6, 12, 13]. OxyS RNA is also involved in the mechanism of *rpoS* gene expression regulation. OxyS binds to Hfq protein *in vivo* and *in vitro*, changes its activity, and hence suppresses the translation of *rpoS* mRNA [14].

Hfq stimulates the degradation of several mRNA in a cell, including *ompA* mRNA (exceptionally stable mRNA encoding the outer membrane protein OmpA) and *rpsO* mRNA encoding the S15 ribosome protein [15]. Hfq competes with 30S ribosomal subunits for binding with 5′-UTR *ompA* RNA [16], participates in the metabolism of poly(A) stimulating the synthesis of poly(A) on mRNA by poly(A)-polymerase I *in vitro* and increasing the length of oligo(A) at mRNA *in vivo* [17]. Hfq interacts with many small regulatory RNA in a cell. It possesses nonspecific DNA-binding activity and along with other proteins is involved in the formation of *E. coli* nucleoid [18]. Since Hfq acts as both DNA- and RNA-binding protein it can be involved in the regulation of transcription and translation coupling in a bacterial cell.

* To whom correspondence should be addressed.

0006-2979/02/6711-1293$27.00 ©2002 MAIK "Nauka/Interperiodica"

The goal of this study was cloning of the *hfq* gene, preparation of *E. coli* Hfq protein superproducer, development of an extraction technique for Hfq protein in preparative quantities, control of homogeneity of the obtained preparation, and preparation of crystals of Hfq protein.

MATERIALS AND METHODS

Preparation of *E. coli* Hfq protein superproducer. The Studier system [19] was employed for the preparation of a superproducer strain. Gene *hfq* was amplified using polymerase chain reaction (PCR); the *E. coli* chromosomal DNA extracted from DH5α strain served as a template as described earlier [20]. We used primers containing recognition sites for *Nde*I and *Bam*HI—forward primer:

5′-GGAAAAGAGCATATGGCTAAGGGG-3′

and reverse primer:

5′-CGCTGGATCCCCGTGTAAAAAAAC-3′.

PCR was performed in 100 μl volume. The reaction mixture contained 10 μl of 10-fold PCR buffer (570 mM Tris, pH 8.8, 166 mM $(NH_4)_2SO_4$, and 15% Tween-20), 2.5 mM $MgCl_2$, mixture of deoxyribonucleotides dNTP (0.2 mM each), primers (100 pM each), 20 ng of chromosomal DNA from *E. coli*, and 5 units of activity of Vent DNA polymerase. The DNA was amplified during 30-35 cycles. Each cycle consisted of three stages: DNA denaturing at 95°C for 30 sec, primer annealing at 57°C for 30 sec, and polymerase reaction at 72°C for 3 min. The results were analyzed by electrophoresis in 1% agarose gel.

PCR product and pET11a-PL vector (being a derivative of pET11a vector and containing four additional restriction sites in the area of the polylinker) were treated with *Nde*I and *Bam*HI restriction endonucleases, and the resulting fragments were purified by electrophoresis and subsequent elution from 1% agarose. DNA ligation was performed in 10 μl. Ligation mixture contained 50 ng of vector and 50 ng of PCR-fragment. *E. coli* BL21(DE3) strain was transformed by the obtained pHfq recombinant plasmid including the *hfq* gene under the control of the T7-promotor. DNA was sequenced (GATC, Germany) using T7 multi-purpose primers.

The synthesis of Hfq protein was induced by the addition of isopropyl-β-D-thiogalactoside (IPTG) into the cell culture to final concentration 0.5 mM. The absorption (A_{600}) of the cultivation medium was 0.8 optical density units per ml. After the addition of an inducer the cells were cultivated for 3 h at 37°C with shaking. The level of Hfq protein production was evaluated using electrophoresis in 15% SDS-PAGE following Laemmli [21]. The gels were treated by Coomassie G-250.

Isolation and purification of the recombinant Hfq protein. To obtain preparative amounts of Hfq protein, cells of the recombinant BL21(DE3)/pHfq strain were spread on agarized LB plates (1% tryptone, 0.5% yeast extract, 1% NaCl, 1.5% agar-agar, pH 7.5), containing ampicillin at the concentration of 100 μg/ml and grown overnight at 37°C. The colonies were placed into liquid LB medium containing ampicillin at the same concentration and cultivated until absorption A_{600} of the medium is 0.8 per ml. The synthesis of Hfq protein was induced by the addition of IPTG until the final concentration of 0.5 mM and the cells were grown for an additional 3 h at 37°C with shaking. The biomass was precipitated by centrifugation for 15 min at 7000*g* and 4°C. The cells were resuspended in 0.1 M Tris-HCl-buffer, pH 8.0, containing 0.2 M $MgCl_2$, 0.8 M NaCl, 5 mM EDTA, 0.1 mM phenylmethylsulfonyl fluoride (PMSF) and then sonicated. The broken cell walls and membranes were removed using low-speed centrifugation for 20 min at 13,000*g* and 4°C. After that the ribosomes were removed from the mixture using high-speed centrifugation at 100,000*g* (4°C) for 20 min. The supernatant was heated at 80°C for 15 min and centrifuged at 13,000*g* and 4°C for 1 h. Sodium-acetate buffer, pH 5.6, was added to the supernatant to the final concentration of 50 mM, dry NaCl until the final concentration of 1 M, and dry $(NH_4)_2SO_4$ until the final concentration of 1.6 M. The protein solution was applied to a butyl-Toyopearl column (10 ml) previously equilibrated with 50 mM of sodium-acetate buffer, pH 5.6, containing 1 M NaCl and 1.6 M $(NH_4)_2SO_4$. Then the column was rinsed with the same solution, and then Hfq protein bound to the carrier was eluted by a gradient of $(NH_4)_2SO_4$ (1.6-0 M) in 100 ml of 50 mM sodium-acetate buffer, pH 5.6, containing 1 M NaCl. The elution speed was 30 ml/h, the fraction volume 2 ml. The absorption (A_{280}) of each fraction was measured. Fractions that did not contain RNA were analyzed by SDS-PAGE. Fractions containing Hfq protein were collected and concentrated using a VivaSpin-10 concentrator (VivaScience, UK) until the absorption (A_{274}) of 25 units per ml.

Crystallization was performed by the hanging drop vapor diffusion technique at 21 and 4°C using commercial kits Crystal Screen and Crystal Screen II (Hampton Research, USA). The protein was transferred into 20 mM Tris-HCl-buffer, pH 8.0, containing 200 mM NaCl by dialysis. The protein concentration in drops was 12-25 mg/ml, drop volume varied in the range of 3-5 μl.

RESULTS AND DISCUSSION

Preparation of an Hfq protein superproducer. The *hfq* gene from *E. coli* was cloned into the pET11a-PL expression vector controlled by the T7-inducible promotor. The resulting pHfq plasmid was used for the transformation of

```
E. coli           : MAKGQSLQDPFLNALRRERVPVSIYLVNGIKLQGQIESFDQFVILLKNTVS  51
P. aeruginosa     : MSKGHSL--PYLNTLRRERVPVSI--VNGIKLQGQIESFDH--ILLKNTVS  45
S. typhimurium    : MAKGQSLQDPFLKPLRRERVPVSIYLVNGIKLQGQIESFDQFVILLKNTVS  51
H. influenzae     : MAKGQSLQDPYLNALRRERIPVSIYLVNGIKLQGQIESFDQFVILLKNTVN  51
Y.enterocolitica: MAKGQSLQDPFLNALRRERVPVSIYLVNGIKLQGQVESFDQFVILLKNTVS  51
Sh. flexneri      : MAKGQSLQDPFLNALRRERVPVSIYLVNGIKLQGQIESFDQFVILLKNTVS  51
P. multocida      : MAKGQSLQDPYLNALRRERIPVSIYLVNGIKLQGQIESFDQFVILLKNTVN  51
Ph. profundum     : MAKGQSLQDPFLNALRRERIPVSIYLVNGIKLQGQIESFDQFVILLKNTVN  51

E. coli           : QMVYKHAISTVVPSRPVSHHSNNAGGGTSSNYHHGSSAQNTSAQQDSEETE 102
P. aeruginosa     : QMVYKHA--TVVPSRPV------LPSGDAEPGNA----------------- 71
S. typhimurium    : QMVYKHAISTVVPSRPVSHHSNNAGGGASNNYHHGSNAQGSTAQQDSEETE 102
H. influenzae     : QMVYKHAISTVVPARSVSHHNNNHHTAPTEAVENVETQAE----------- 91
Y.enterocolitica: QMVYKHAISTVVPSRPVSHHSNNPSGS-TNNYHGSNPSAPQQPQQDSDDAE 101
Sh. flexneri      : QMVYKHAISTV------------------------------------------ 62
P. multocida      : QMVYKHAISTVVPARSVSHH-NNSNNSNQQNYQQEQQTDSNVEKAE----- 96
Ph. profundum     : QMVYKHAISTVI--R----------------------------------- 64
```

Fig. 1. Comparison of Hfq protein sequences from different bacteria. White letters on dark background indicate identical amino acid residues, black letters on gray background correspond to homologous amino acid residues.

E. coli BL21(DE3) cells. The integrity of the *hfq* gene nucleotide sequence in the expression vector was verified by sequencing. The conditions for *hfq* gene expression were optimized. We found an increased level of protein synthesis after the addition of 0.5 mM IPTG into the cultivation medium at the absorption of 0.8 units per ml; after that the cells were cultivated for an additional 3 h at 37°C with shaking. The expression efficiency was verified by the SDS-PAGE technique.

Isolation of the homogeneous soluble Hfq protein preparation from BL21(DE3)/pHfq superproducer. Hfq protein from *E. coli* consists of 102 amino acid residues (Fig. 1) and its molecular weight is 11.2 kD. The first stages of protein purification (removal of membrane fraction, ribosomes, and heat treatment of the resulting extract) yielded preparation of 90% purity (according to SDS-PAGE data). The temperature of heat treatment was chosen experimentally by heating the non-ribosomal extract at various temperatures ranging from 40 to 100°C and evaluating the degree of protein purification by SDS-PAGE (data not presented). When the preparation is heated to 80°C for 15 min most of the *E. coli* proteins denature and form an insoluble precipitate, while almost all Hfq protein remains in the solution. This stage does not change the protein activity as illustrated by toe-printing analysis (U. Bläsi, unpublished data).

The most difficult task proved to be the removal of RNA from the Hfq preparation. Hydrophobic chromatography on a butyl-Toyopearl column in a gradient of ammonium sulfate concentration (1.6-0 M) in the presence of 1 M NaCl allowed the complete purification of

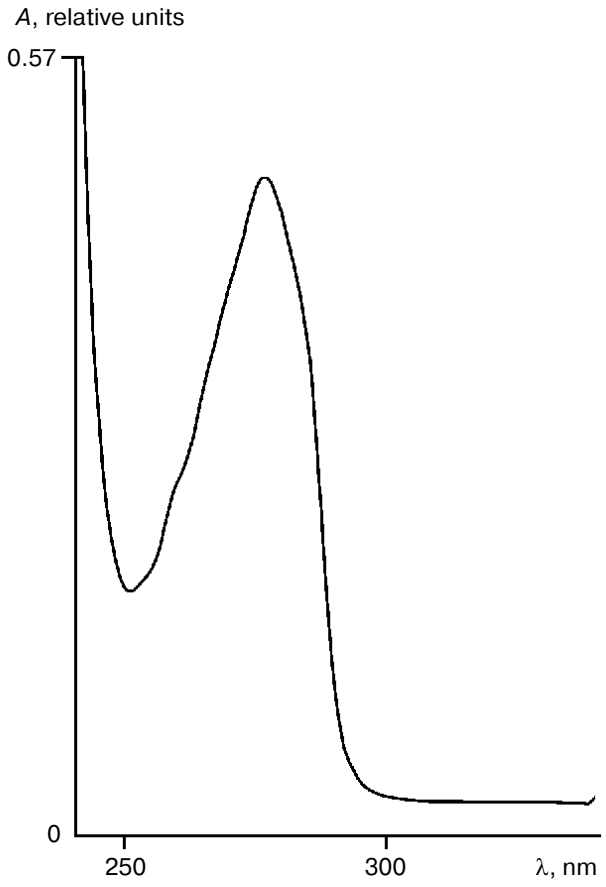

Fig. 2. Absorption spectrum of Hfq protein in the UV-range. Abscissa, the wavelengths (nm); ordinate, the absorption (optical units per 1 ml).

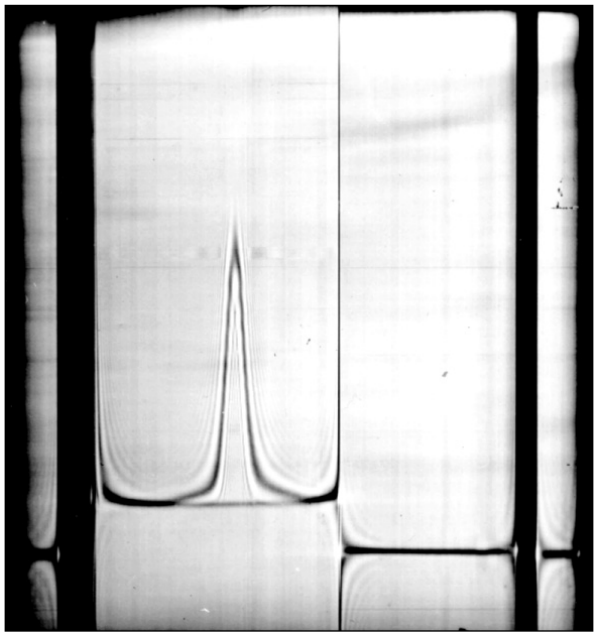

Fig. 3. Sedimentogram of Hfq protein obtained by ultracentrifugation at 20°C.

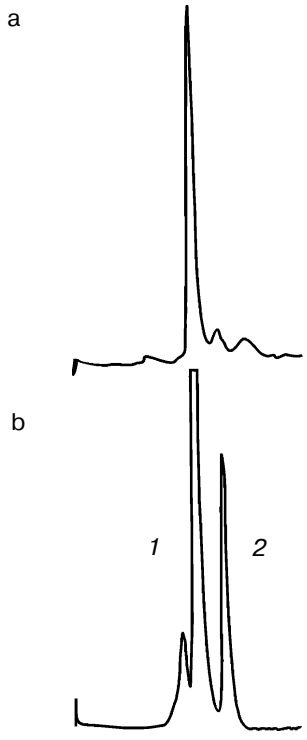

Fig. 4. Gel filtration in HPLC system with molecular weight protein markers on Superose-12B carrier (column length 30 cm, diameter 1 cm, elution speed 0.4 ml/min). a) Elution profile of Hfq protein; b) elution profile of protein markers with a molecular weight of 66 kD (BSA) (*1*), 14.2 kD (α-lactalbumin) (*2*).

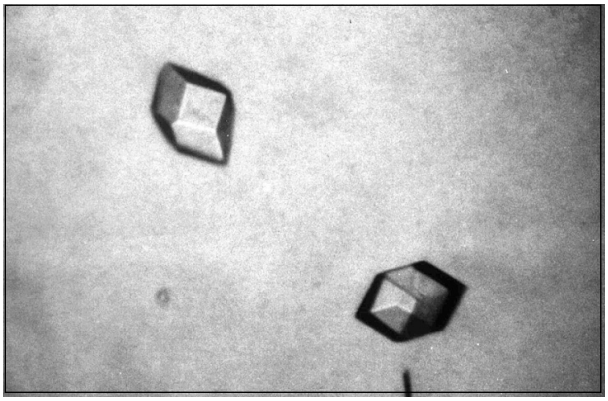

Fig. 5. Crystals of Hfq protein.

the protein from RNA. Hfq protein was strongly bound to the carrier and eluted at $(NH_4)_2SO_4$ concentration of 1 M, while RNA did not bind with the carrier and was eluted during the sample application and column rinsing with the starting buffer. The absence of RNA was confirmed by the absorption spectrum (Fig. 2). Finally, the soluble Hfq preparation was obtained with a purity of 95-98% according to SDS-PAGE data and free from RNA impurities. The developed technique produces 5-7 mg protein from 1 g of the superproducer biomass.

Homogeneity of the preparation is an indispensable condition for the crystallization experiments. Homogeneity of Hfq protein was examined by ultracentrifugation technique and gel filtration in a HPLC system (Figs. 3 and 4). Using the ultracentrifugation technique the sedimentation coefficient of Hfq protein at 20°C was determined to be 4.1 S. The ultracentrifugation and gel-filtration data using HPLC data with molecular markers allow concluding that in the recombinant Hfq protein in solution free from RNA exists as a sufficiently homogeneous hexamer complex in agreement with the literature data for Hfq protein extracted from *E. coli* cells [2].

Crystallization of the Hfq protein. The crystals of Hfq protein reflect X-rays with the limit of resolution of 2.5 Å (Fig. 5). There are also heavy atom derivatives of these crystals containing selenium atoms reflecting X-rays with the limit of resolution of 2.7 Å. The crystals achieve the size of 250 × 200 × 200 μm. Sets of diffraction data have been collected for the crystals and the determination the spatial structure of *E. coli* Hfq protein is now in progress.

This work was financially supported by the Russian Academy of Sciences. The research of M. B. Garber was financed by the Russian Foundation for Basic Research (grant No. 02-04-48536) and partially by The International Research Scholar's Award from the Howard

Hughes Medical Institute. The authors thank N. A. Matveeva, I. A. Kashparov, and A. N. Turkin (Institute of Protein Research, Russian Academy of Sciences, Pushchino) for their technical help in the course of this work.

REFERENCES

1. Franze de Fernandez, M. T., Eoyang, L., and August, J. T. (1968) *Nature*, **219**, 588-590.
2. Franze de Fernandez, M. T., Hayward, W. S., and August, J. T. (1972) *J. Biol. Chem.*, **247**, 824-831.
3. Kajitani, M., and Ishihama, A. (1991) *Nucleic Acids Res.*, **19**, 1063-1066.
4. Schuppli, D., Miranda, G., Tsui, H. C. T., Winkler, M. E., Sogo, J. M., and Weber, H. (1997) *Proc. Natl. Acad. Sci. USA*, **94**, 10239-10242.
5. Tsui, H.-C. T., Feng, G., and Winkler, M. E. (1997) *J. Bacteriol.*, **179**, 7476-7487.
6. Muffler, A., Fischer, D., and Hengge-Aronis, R. (1996) *Genes Dev.*, **10**, 1143-1151.
7. Brown, L., and Elliott, T. (1996) *J. Bacteriol.*, **178**, 3763-3770.
8. Hengge-Aronis, R. (1993) *Cell*, **72**, 165-168.
9. Hengge-Aronis, R. (1996) *Mol. Microbiol.*, **21**, 887-893.
10. Loewen, P. C., and Hengge-Aronis, R. (1994) *Annu. Rev. Microbiol.*, **48**, 53-80.
11. Fang, F. C., Libby, S. J., Buchmeier, N. A., Loewen, P. C., Switala, J., Harwood, J., and Guiney, D. G. (1992) *Proc. Natl. Acad. Sci. USA*, **89**, 11978-11982.
12. Brown, L., and Elliott, T. (1997) *J. Bacteriol.*, **179**, 656-662.
13. Cunning, C., Brown, L., and Elliot, T. (1998) *J. Bacteriol.*, **180**, 4564-4570.
14. Zhang, A., Altuvia, S., Tiwari, A., Argaman, L., Henhhe-Aronis, R., and Storz, G. (1998) *EMBO J.*, **17**, 6061-6068.
15. Vytvytska, O., Jakobsen, J. S., Balcunaite, G., Andersen, J. S., Baccarini, M., and von Gabain, A. (1998) *Proc. Natl. Acad. Sci. USA*, **95**, 14118-14123.
16. Vytvytska, O., Moll, I., Kaberdin, V. R., von Gabain, A., and Blaesi, U. (2000) *Genes Dev.*, **14**, 1109-1118.
17. Hajnsdorf, E., and Regnier, P. (2000) *Proc. Natl. Acad. Sci. USA*, **97**, 1501-1505.
18. Azam, T. A., and Ishihama, A. (1999) *J. Biol. Chem.*, **274**, 33105-33113.
19. Studier, F. W., Rosenberg, A. H., Dunn, J. J., and Dubendorff, J. W. (1990) *Meth. Enzymol.*, **185**, 60-89.
20. Marmur, J. (1961) *J. Mol. Biol.*, **3**, 208-218.
21. Laemmli, U. (1970) *Nature*, **227**, 680-685.