Acta Cryst. (1998). D**54**, 1043–1045

Cloning, overproduction, purification and crystallization of the DNA binding protein HU from the hyperthermophilic eubacterium *Thermotoga maritima*

EVANGELOS CHRISTODOULOU[a] AND CONSTANTINOS E. VORGIAS[b]* *at* [a]*European Molecular Biology Laboratory, c/o DESY, Notkestrasse 85, 22603 Hamburg, Germany, and* [b]*Athens University, Biology Department, Division of Biochemistry and Molecular Biology, Panepistimiopolis-Kouponia, 15701 Athens, Greece. E-mail: cvorgias@biology.db.uoa.gr*

(*Received 14 October 1997; accepted 6 January 1998*)

Abstract

The *humar* gene encoding for the histone-like DNA-binding protein HU from the hyperthermophilic eubacterium *Thermotoga maritima* was efficiently overexpressed in *Escherichia coli* under the T7 promoter. The HU protein was purified using SP-Sepharose ion-exchange and heparin-affinity chromatography and was successfully crystallized in ammonium sulfate. The crystals were grown in the tetragonal form in space group $P4_3$ or $P4_1$ and have unit-cell dimensions $a = b = 46.12$, $c = 77.56$ Å, $\alpha = \beta = \gamma = 90°$. The crystals diffract X-rays to 1.6 Å resolution using synchrotron radiation and are suitable for determination of the HU structure at high resolution.

1. Abbreviations

Amp, ampicillin; EDTA, ethylenediaminetetraacetic acid; *E. coli*, *Escherichia coli*; IPTG, isopropyl-β-D-thiogalactopyranoside; LB, Luria Bertani (medium); NMR, nuclear magnetic resonance; OD_{600}, absorption at 600 nm; PAGE, polyacrylamide gel electrophoresis; PMSF, phenylmethylsulfonyl fluoride; SDS, sodium dodecyl sulfate; aa, amino acids.

2. Introduction

The nucleoid of the prokaryotic cell contains a number of abundant, low-molecular-weight and positively charged proteins classified as histone-like DNA-binding proteins. Among those proteins, HU has been identified as the major protein component of the nucleoid and has attracted considerable attention during the last two decades. In *E. coli*, HU (HUαβ) is the most abundant DNA-binding protein with ~30 000 dimers per cell. It is a heterodimer consisting of homologous α and β subunits each of 90 aa, which are encoded by the *hupB* and *hupA* genes, respectively, and are 70% identical in their sequence. HU appears to be a homodimer in all other bacterial species in which it has been studied (for review see Drlica & Rouviere-Yaniv, 1987; Pettijohn, 1988). Up to now, several functions of HU have been found including: (i) HU binds with little sequence specificity to dsDNA, ssDNA and RNA (Rouvière-Yaniv & Gros, 1975); (ii) DNA–HU complexes condense into nucleosome-like particles and can introduce negative supercoiling into a relaxed circular plasmid DNA in the presence of topoisomerase I (Rouvière-Yaniv *et al.*, 1979; Broyles & Pettijohn, 1986); (iii) the binding of the *lac* repressor as well as the binding of the cAMP receptor protein to the *lac* promoter is facilitated by HU (Flashner & Gralla, 1988); (iv) HU is required for transposition by bacteriophage Mu (Craigie *et al.*, 1985) and (v) HU *in vitro* plays a regulatory role in λ DNA replication (Mensa-Wilmot *et al.*, 1989).

The crystal structure of HU from *Bacillus stearothermophilus* has been solved (Tanaka *et al.*, 1984; White *et al.*, 1989). The solution structure of recombinant HU from *B. stearothermophilus* expressed in *E. coli* (Padas *et al.*, 1992) was also determined using NMR (Vis *et al.*, 1995).

The study of the structural properties responsible for the thermostability of HUs from mesophilic and thermophilic organisms attracted our attention in the past (Wilson *et al.*, 1990). We want to expand our studies to HU protein from extreme thermophilic organisms, such as the hyperthermophilic eubacterium, *Thermotoga maritima* which shows 57.7 and 54.4% identity to HU from *B. stearothermophilus* and *Bacillus globigii*, respectively. We have focused on the determination of the crystal structure of the HU (HUmar) from this organism, as a first step towards the elucidation of the structural properties responsible for the thermal stability at temperatures above 353 K. The overproduction in *E. coli*, purification and crystallization of the HUmar protein, as well as preliminary crystallographic data are described here.

3. Materials and methods

3.1. *Gene cloning*

All cloning procedures were carried out as described in Sambrook *et al.* (1989), or according to the manufacturers instructions. The gene *humar* from *T. maritima* which encodes the HUmar protein, was isolated from chromosomal DNA kindly supplied by Dr Lieb (TU Munchen) using the polymerase chain reaction (PCR). The primers were designed from the primary structure of the *humar* gene which is deposited under the accession number L23541. The *humar* gene was initially cloned into the pcrII vector (InVitrogen) for DNA sequencing and verification, and subsequently subcloned into the expression vector pET-11*a* (Novagen) using the engineered restriction sites *Nde*I and *Bam*HI.

3.2. *Protein overproduction*

The resulting plasmid pET11*a-humar* (Amp resistant) was transformed into the three following *E. coli* host cell lines: BL21(DE3), BL21(DE3)[pLysS] and BL21(DE3)[pLysE] (Studier *et al.*, 1990) for protein expression. The transformations were carried out by the method described by Inone *et al.* (1990). The *E. coli* BL21(DE3) cells have shown the highest HUmar protein production and were used for protein preparation. The cells were grown in LB medium in the presence of 50 μg Amp ml^{-1}, overnight. The next day, the cells were diluted to a final $OD_{600} = 0.6$ and the overexpression of the target *humar* gene was induced with IPTG to a final concentration of 0.5 m*M*. At 0, 3, 6, 12 h post induction, samples were withdrawn and analyzed by the 0.1% SDS15%–

PAGE Laemmli system and stained with Coomassie Blue G-250 (Laemmli, 1970), Fig. 1, lanes 1–4. The highest expression levels of *humar* gene were achieved after 3 h induction as shown in Fig. 1, lane 2.

3.3. *Protein purification*

In a routine protein preparation, a 2 l cell culture induced for 3 h was used. The induced bacteria were collected by low-speed centrifugation and washed once with cold buffer W (50 m*M* Tris–HCl pH 7.5/100 m*M* NaCl/0.1 m*M* PMSF). All further procedures were carried out at 273–277 K, unless otherwise specified. The bacterial paste was resuspended in 5 ml buffer *A* [20 m*M* Tris–HCl pH 8.0/1 m*M* EDTA/0.1 m*M* PMSF/20 m*M* NaCl, 10% glycerol/0.1%(*w*/*v*) Triton X-100] per gram of cell paste. The cells were disrupted by sonication for 15 min. The total extract was clarified by centrifugation in an SS-34 rotor (Sorvall) at 20 000 rev min^{-1} for 20 min. The supernatant was adjusted to 40% saturation in ammonium sulfate. After 30 min of stirring, the non-precipitated material was separated by centrifugation as above. The soluble supernatant was dialysed against buffer *B* (10 m*M* Na-phosphate pH 8.0/1 m*M* EDTA/0.1 m*M* PMSF), overnight. The dialysed HUmar protein fraction was adjusted to 4 *M* urea and applied to a 20 ml SP-Sepharose HP column (Pharmacia), pre-equilibrated against the same buffer in the presence of 4 *M* urea. Bound proteins were eluted with a 400 ml NaCl (0.0–1.0 *M*) linear gradient of buffer *B* in the absence of urea, as a symmetric peak between 350 and 450 m*M* NaCl. The first chromatographic step increased the purity of HUmar protein to about 70%. The peak fractions containing HUmar were combined and directly applied to a 10 ml heparin Sepharose CL-6B column (in 10 m*M* Na-phosphate pH 7.0/1 m*M* EDTA/0.1 m*M* PMSF, 400 m*M* NaCl). The HUmar protein was eluted with a linear 400–1600 m*M* NaCl 100 ml gradient. The final step yielded a very pure HUmar protein, Fig. 1, lane 5. In a routine preparation, we were able to prepare 10 mg of highly purified HUmar per gram wet weight of *E. coli* cells.

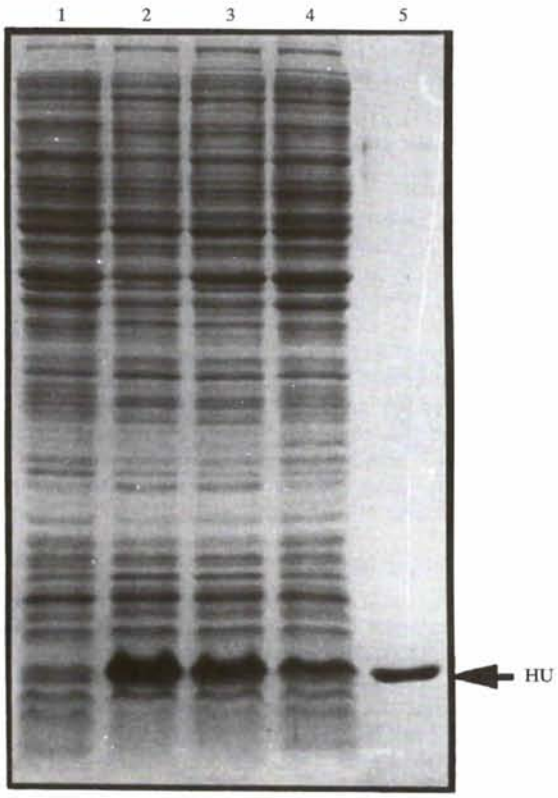

Fig. 1. HUmar overproduction analysed by 0.1% SDS15%–PAGE. Lane 1: BL21(DE3) *E. coli* cells carrying the pET-11*a*-*humar* plasmid used without induction with IPTG; lanes 2, 3, 4 are the same cells 3, 6, 12 h after induction with 0.5 m*M* IPTG, lane 5 is pure protein.

3.4. *Crystallization*

HUmar was screened for crystals using the Hampton Research Crystal Screen I at 277 and 291 K (Jancarick & Kim, 1991). The protein was at a concentration of 10 mg ml^{-1} in 10 m*M* Na-phosphate pH 7.0. All crystals were grown by hanging-drop vapour diffusion, 2–3 µl of protein solution were mixed with an equal volume of precipitating solution and equilibrated against 1.0 ml of precipitating solution (Jancarick & Kim, 1991).

Useful crystals were obtained at pH 4.5 using 80% saturated ammonium sulfate at room temperature. The crystals appeared after 3–5 months and have mostly hexagonal shape and dimensions 0.4 × 0.4 × 0.2 mm, Fig. 2.

3.5. *Preliminary crystallographic analysis*

Diffraction data were collected at room temperature, to 1.6 Å, with synchrotron radiation (λ = 0.8833 Å), from the EMBL BW7B beamline (Van Slifhout & Hermes, 1995) at the DORIS storage ring, DESY, Hamburg, with an imaging-plate scanner (Hendrix & Lentfer, unpublished work). The reflection data were intergrated and scaled using the programs *DENZO* and *SCALEPACK*, respectively (Otwinowski, 1993). The lattice is tetragonal with the space group $P4_1$ or $P4_3$. The

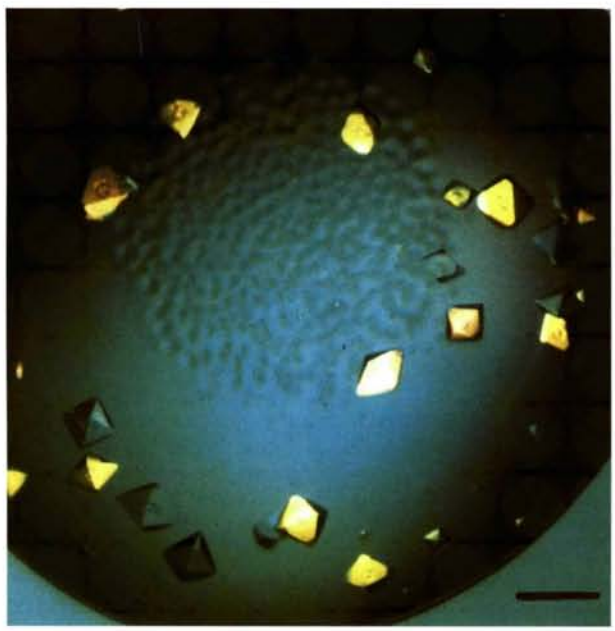

Fig. 2. Crystals of HUmar protein obtained as described in §3. Bar represents 1 mm.

unit-cell dimensions are $a = b = 46.12$, $c = 77.56$ Å, $\alpha = \beta = \gamma = 90^\circ$. The data to 1.6 Å resolution were 99.9% complete and contained a total of 127 048 raw measurements, which were merged to 21 639 unique reflections with an $R_{\rm merge}$ of 0.056. In the high-resolution bin (1.63–1.6 Å) the $R_{\rm merge}$ is 0.594 with $I/\sigma > 2.7$. The molecular mass of recombinant HUmar is 18 kDa, the $V_M = 2.29$ Å^3 Da^{-1}, which is within the range of normal protein crystals (Matthews, 1968), assuming that the asymmetric unit contains one HUmar molecule. The solvent content is 46%(*v*/*v*).

4. Results and discussion

The overproduction of the recombinant histone-like HUmar protein in *E. coli* and the simple and fast protein purification scheme have made crystallization possible. During the first purification step on SP-Sepharose HP column, it was essential to include 4 *M* urea in the column and sample application buffer in order to dissociate DNA fragments which were strongly bound to the protein (data not shown). Under the experimental conditions described in §3 the HUmar protein was tightly bound to the column while DNA fragments were washed out. The elution of the bound HUmar protein was achieved with high salt in the absence of urea without affecting it. Particular care was taken to avoid methionine oxidation. We have found out that methionine oxidation may drive the formation of multimers of HUmar, as resolved by Urea/Triton gel electrophoresis (data not shown).

We have obtained crystals which diffract to a resolution of 1.6 Å. Preliminary X-ray diffraction data are reported for the HUmar crystals and we are in the process of using the diffraction data to solve the structure of the protein.

We thank Dr Lieb (TU Munchen) for providing us the chromosomal DNA of *T. maritima* and Dr Wojciech Rypniewski for helping us with the data collection (EMBL Hamburg Outstation).

References

Broyles, S. & Pettijohn, D. E. (1986). *J. Mol. Biol.* **187**, 47–60.
Craigie, R., Arndndt-Jovin, D. & Mizuuchi, D. (1985). *Proc. Natl Acad. Sci. USA*, **82**, 7570–7574.
Drlica, K. & Rouviere-Yaniv, J. (1987). *Microbiol. Rev.* **51**, 301–319.
Flashner, Y. & Gralla, J. D. (1988). *Cell*, **54**, 713–721.
Inone, H., Nojima, H. & Okayama, H. (1990). *Gene*, **96**, 23–28.
Jancarick, J., & Kim, S.-H. (1991). *J. Appl. Cryst.* **24**, 409–411.
Laemmli, U. K. (1970). *Nature (London)*, **227**, 680–685.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Mensa-Wilmot, K., Carroll, K. & McMacken, R. (1989). *EMBO J.* **8**, 2393–2402.
Otwinowski, Z. (1993). *DENZO: An oscillation data processing program for macromolecular crystallography*. Yale University, New Haven, CT, USA.
Padas, M. P., Wilson, K. S. & Vorgias, C. E. (1992). *Gene*, **117**, 39–44.
Pettijohn, D. E. (1988). *J. Biol. Chem.* **263**, 12793–12796.
Rouvière-Yaniv, J. & Gros, F. (1975). *Proc. Natl Acad. Sci. USA*, **72**, 3428–3432, 1975.
Rouvière-Yaniv, J., Yaniv, M & Germond, J. (1979). *Cell*, **17**, 265–274.
Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular Cloning. A Laboratory Manual*, 2nd ed. New York: Cold Spring Harbor Laboratory Press.
Studier, F. W., Rosenberg, A. H., Dunn, J. J. & Dubendorff, J. W. (1990). *Methods Enzymol.* **185**, 60–89.
Tanaka, I., Appelt, K. Dijk, J., White, S. W. & Wilson, K. S. (1984). *Nature (London)*, **310**, 376–381.
Van Slifhout, R. G. & Hermes, C. (1995). *Rev. Sci. Instrum.* **66**, 1818–1820.
Vis, H., Mariani, M., Vorgias, C. E., Wilson, K. S., Kaptein, R. & Boelens, R. (1995). *J. Mol. Biol.* **254**, 692–703.
White, S., Appelt, K., Wilson, K. S. & Imamoto, J. (1989). *Proteins Struct. Funct. Genet.* **5**, 281–288.
Wilson, K. S., Vorgias, C. E., Tanaka, I., White, S. W. & Kimura, M. (1990). *Protein Eng.* **4**, 11–22.