Expression, purification and crystallization of a protein resulting from the inversion of the amino-acid sequence of a helical bundle

Aikaterini Kefala,^a Dina Kotsifaki,^b Mary Providaki,^b Maria Amprazi^b and Michael Kokkinidis^{a,b*}

Received 7 November 2016

Accepted 19 December 2016

Edited by Z. Dauter, Argonne National Laboratory, USA

Keywords: Rop protein; heptad repeat; protein folding; sequence–structure relationship; hyperthermophilic protein.

Supporting information: this article has supporting information at journals.iucr.org/f

^aDepartment of Biology, University of Crete, Voutes University Campus, PO Box 2208, 70013 Heraklion, Crete, Greece, and ^bInstitute of Molecular Biology and Biotechnology, Foundation for Research and Technology – Hellas, Nikolaou Plastira Street 100, 70013 Heraklion, Crete, Greece. *Correspondence e-mail: kokkinid@imbb.forth.gr

Earlier studies have found that the occurrence of inverse sequence identity in proteins is not indicative of three-dimensional similarity, but rather leads to different folds or unfolded proteins. Short helices, however, frequently keep their conformations when their sequences are inverted. To explore the impact of sequence inversion on long helices, revRM6, with the inverse amino-acid sequence relative to RM6, a highly stable variant of the ColE1 Rop protein, was engineered. RM6 is a highly regular four- α -helical bundle that serves as a model system for protein-folding studies. Here, the crystallization and preliminary crystallographic characterization of revRM6 are reported. The protein was overexpressed in *Escherichia coli*, purified to homogeneity and crystallized. The crystals belonged to space group $P4_12_12$, with unit-cell parameters $a = b = 44.98$, $c = 159.74$ Å, and diffracted to a resolution of 3.45 Å.

1. Introduction

Repressor of primer (Rop) is an RNA-binding protein that is involved in regulation of the copy number of ColE1 plasmids (Polisky, 1988). The 63-residue protein forms two antiparallel α -helices connected by a hairpin turn and assembles itself into a homodimer which forms a highly regular all-antiparallel four- α -helical bundle (PDB entry 1rop; Banner *et al.*, 1987; Eberle *et al.*, 1991). Owing to its structural simplicity, Rop is a widely used model system for analysis of the relationships between the sequence, folding and dynamics of four- α -helix bundles (Glykos *et al.*, 2006; Amprazi *et al.*, 2014). Surprisingly, numerous functional, thermodynamic, kinetic and structural studies of engineered Rop mutants revealed that at the level of sequence–structure relationships the structural simplicity of Rop is deceiving, as the factors which affect its folding are very challenging to interpret (Amprazi *et al.*, 2014). Four- α -helical bundle sequences are characterized by ‘heptad repeats’, a repeating pattern of seven residue types corresponding to the seven topologically distinct positions (a, b, c, d, e, f, g) of the associating α -helices in the bundle (Cohen & Parry, 1990; Gazi *et al.*, 2009). Positions a and d are overwhelmingly hydrophobic as they form the core of the bundle, whereas positions b, c, e, f and g are characterized by amino-acid preferences reflecting geometrical constraints and their solvent exposure (Paliakasis & Kokkinidis, 1992).

The ‘heptad repeats’ pattern of Rop is interrupted in the hairpin-turn region with the insertion of five residues. Deletion of these five residues (the RM6 mutant) produces a continuous pattern of heptad repeats from the N-terminus to the C-terminus. RM6, an inactive variant of Rop, forms a

Table 1

Macromolecule-production information.

Source organism	<i>E. coli</i>
DNA source	DNA synthesized by GenScript, USA
Cloning vector	pET-26b
Expression vector	pET-26b
Expression host	<i>E. coli</i> BL21(DE3)
Complete amino-acid sequence of the construct produced	MLNEGDDGFRALCSRYLEDADHDLSECIDALENLKELLTLTQSRIFRAMNLATKEQKTMHHHHHH

homotetrameric, extremely thermostable helical bundle, with each monomer consisting of a single long α -helix, in contrast to the wild-type Rop structure which forms two antiparallel helices joined by a short hairpin bend (PDB entry 1qx8; Glykos *et al.*, 2006). The reasons for the observed structural plasticity displayed by Rop and RM6 are still poorly understood, despite a wealth of structural and physicochemical data.

In this work, we test the heptad-repeats concept through the complete inversion of the amino-acid sequence of RM6, which gives rise to the revRM6 protein. We expect that the structure of revRM6 will reveal whether the inverted heptad pattern will produce a similar helical bundle to RM6, a behaviour which appears to characterize short helices. It should be noted that in the inverted revRM6 sequence the critical heptad positions *a* and *d* are swapped relative to RM6.

2. Materials and methods

2.1. Macromolecule production

The gene fragment coding for revRM6 was synthesized by GenScript (USA), supplied in the pET-26b (ColE1 plasmids) vector carrying a C-terminal His₆ tag and transformed into *Escherichia coli* strain BL21. A sufficient amount of soluble protein was obtained after induction using the following conditions. The cells were grown in LB medium containing 25 mg ml⁻¹ kanamycin until an OD₆₀₀ of 0.6 was reached. The culture was induced at 301 K with 1 mM IPTG for 18 h. Approximately 3 g of cell paste was resuspended in 50 ml lysis buffer consisting of 25 mM Tris pH 8.0, 300 mM NaCl, 5 mM imidazole, 15 mM β -mercaptoethanol and homogenized.

Figure 1
Crystals of the revRM6 protein.

Table 2

Crystallization.

Method	Hanging-drop vapour diffusion
Plate type	24-well Linbro cell-culture plates
Temperature (K)	291
Protein concentration (mg ml ⁻¹)	14
Buffer composition of protein solution	25 mM Tris pH 8.0, 100 mM NaCl, 15 mM β -mercaptoethanol
Composition of reservoir solution	52%(v/v) MPD, 0.1 M citrate pH 4.3
Volume and ratio of drop	4 μ l, 1:1
Volume of reservoir (ml)	1

After the addition of protease inhibitors (20 mg ml⁻¹ leupeptin, 1 mM PMSF, 150 mg ml⁻¹ benzamidine), the solution was sonicated for 3 min. The precipitate was subsequently removed by centrifugation at 12 000 rev min⁻¹ and 277 K for 45 min. Purification was carried out using His-tag affinity chromatography at 277 K with an 8 ml Ni-NTA column (Qiagen) pre-equilibrated in lysis buffer and initially washed stepwise with 10, 20 and 30 mM imidazole. The imidazole concentration was subsequently increased and the protein started eluting at 200 mM imidazole. Fractions containing the protein were dialyzed against storage buffer consisting of 25 mM Tris pH 8, 100 mM NaCl, 15 mM β -mercaptoethanol and were concentrated to approximately 14 mg ml⁻¹ for subsequent crystallization experiments. Macromolecule-production information is summarized in Table 1.

Further biophysical analysis of the protein was performed using size-exclusion chromatography, SDS-PAGE gels and native gels, which suggested a tetrameric form for revRM6 (Supplementary Figs. S1, S2 and S3). Circular-dichroism analysis was performed using synchrotron radiation on the DISCO beamline at the SOLEIL synchrotron in Paris (Supplementary Fig. S4). The buffer of the protein sample used for these measurements consisted of 25 mM Tris pH 8, 100 mM NaCl, 15 mM β -mercaptoethanol and the protein concentration was up to 14 mg ml⁻¹. The measurements, which were performed at 293 K, revealed that revRM6 is a highly α -helical protein, as is the reverse-sequence protein RM6, which is also α -helical.

2.2. Crystallization

Crystallization conditions for revRM6 were screened using the hanging-drop vapour-diffusion method in 24-well Linbro cell-culture plates. The drops were made up of 2 μ l protein solution mixed with an equal volume of reservoir solution and were equilibrated against 1 ml reservoir solution at 291 K. Initial crystallization screening was performed using commercially available crystallization kits, and the initial crystals were obtained with 40%(v/v) 2-methyl-2,4-pentane-diol (MPD), 0.1 M citrate buffer pH 4.0. Crystal optimization was performed through variation of the above conditions, with the use of various additives from the Silver Bullets additive screen (Hampton Research). The best results were obtained with 52%(v/v) MPD, 0.1 M citrate buffer pH 4.3. The final crystal size was reached in 30 d (Fig. 1). Crystallization information is summarized in Table 2.

2.3. Data collection and processing

X-ray diffraction data were collected from a single crystal using synchrotron radiation on the X06SA beamline at the Swiss Light Source (SLS), Paul Scherrer Institut, Villigen, Switzerland. The crystal was flash-cooled to 100 K in a nitrogen-gas stream. Images corresponding to a rotation of 360° were collected with a 0.2° step per frame (Fig. 2). The diffraction data were recorded on an EIGER 16MX (133 Hz) detector. X-ray diffraction data were indexed, integrated and scaled with *XDS* (Kabsch, 2010) and the *CCP4* program suite (Winn *et al.*, 2011). Data-collection and processing statistics are summarized in Table 3.

3. Results and discussion

The revRM6 crystals exhibited severe diffraction anisotropy, which manifests itself as a directional dependence in diffraction quality; in early crystals the directional variation of diffraction limit ranged from 3 Å along the *c** direction to 7 Å along the *a** and *b** directions. This anisotropy problem, which is generally attributed to a whole-body anisotropic vibration of the unit cell, for example owing to crystal-packing interactions being more uniform in one direction than another, tends to impact the *R* factor (Strong *et al.*, 2006). In an attempt to deal with this issue, we made use of the anisotropy-correction program available at <http://services.mbi.ucla.edu/anisotropy/>. In addition, the optimized crystallization conditions (Table 2) greatly reduced the anisotropy problem. The Matthews coefficient is 3.02 Å³ Da^{−1} and is consistent with a solvent content of 59%. There are two molecules in the asymmetric unit. Calculation of the self-rotation function using *POLARRFN* (Kabsch, 1976) from the *CCP4* package confirmed the presence of local dyads in the *ab* plane (Supplementary Fig. S5).

Further physicochemical analysis of the protein was performed with the collection of circular-dichroism (CD) spectra at the SOLEIL synchrotron, Paris (Supplementary

Table 3

Data collection and processing.

Diffraction source	X06SA beamline, SLS
Wavelength (Å)	1.000
Temperature (K)	100
Detector	EIGER 16MX (133 Hz)
Crystal-to-detector distance (mm)	166
Rotation range per image (°)	0.2
Total rotation range (°)	360
Exposure time per image (s)	0.1
Space group	<i>P</i> ₄ <i>2</i> ₁ <i>2</i>
<i>a</i> , <i>b</i> , <i>c</i> (Å)	44.98, 44.98, 159.74
α , β , γ (°)	90, 90, 90
Mosaicity (°)	1.50
Resolution range (Å)	44.98–3.45 (3.64–3.45)
Total No. of reflections	53899 (6799)
No. of unique reflections	2480 (318)
Completeness (%)	99.1 (95)
Multiplicity	21.7 (21.4)
$\langle I/\sigma(I) \rangle$	19.5 (5.1)
<i>R</i> _{meas}	0.074 (0.492)

Figs. S4 and S6). The CD data revealed a very high (80%) α -helical content. These preliminary results are comparable to those obtained for the RM6 protein, suggesting that despite the sequence inversion the secondary-structure contents of the two proteins are very similar (Supplementary Fig. S6). Structural analysis of revRM6 will reveal if this similarity also extends to the level of tertiary structure.

Acknowledgements

We acknowledge the provision of synchrotron-radiation beamtime at beamline PXI, Swiss Light Source, Paul Scherrer Institut, Villigen, Switzerland and Dr Tomizaki Takashi for assistance. We also acknowledge the SOLEIL synchrotron in Paris for the provision of synchrotron-radiation beamtime at the DISCO beamline and we would like to thank Dr Frank Wien for assistance. This work was supported by the EU FP7 project InnovCrete and the GSRT projects SYN11_1_31 and 12CHN.

References

- Amprazi, M., Kotsifaki, D., Providaki, M., Kapetanidou, E. G., Fellas, G., Kyriazidis, I., Perez, J. & Kokkinidis, M. (2014). *Proc. Natl Acad. Sci. USA*, **111**, 11049–11054.
- Banner, D. W., Kokkinidis, M. & Tsernoglou, D. (1987). *J. Mol. Biol.* **196**, 657–675.
- Cohen, C. & Parry, D. A. D. (1990). *Proteins*, **7**, 1–15.
- Eberle, W., Pastore, A., Sander, C. & Rösch, P. (1991). *J. Biomol. NMR*, **1**, 71–82.
- Gazi, D. A., Charova, S. N., Panopoulos, N. J. & Kokkinidis, M. (2009). *Cell. Microb.* **11**, 719–729.
- Glykos, N. M., Papanikolaou, Y., Vlassi, M., Kotsifaki, D., Cesareni, G. & Kokkinidis, M. (2006). *Biochemistry*, **45**, 10905–10919.
- Kabsch, W. (1976). *Acta Cryst.* **A32**, 922–923.
- Kabsch, W. (2010). *Acta Cryst.* **D66**, 125–132.
- Paliakasis, C. D. & Kokkinidis, M. (1992). *Protein Eng. Des. Sel.* **5**, 739–748.
- Polisky, B. (1988). *Cell*, **55**, 929–932.
- Strong, M., Sawaya, R. M., Wang, S., Phillips, M., Cascio, D. & Eisenberg, D. (2006). *Proc. Natl Acad. Sci. USA*, **103**, 8060–8065.
- Winn, M. D. *et al.* (2011). *Acta Cryst.* **D67**, 235–242.

Figure 2

X-ray diffraction pattern of a revRM6 crystal recorded on the X06SA beamline at the Swiss Light Source.