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

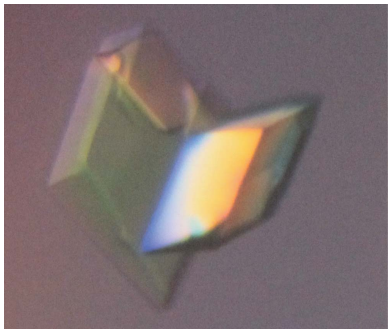
Expression, crystallization and preliminary crystallographic analysis of C-reactive protein from zebrafish

Rong Chen,[a] Jianxun Qi,[b,c] Shugang Yao,[a] Xiaocheng Pan,[a] Feng Gao[d] and Chun Xia[a]*

[a]Department of Microbiology and Immunology, College of Veterinary Medicine, China Agricultural University, Beijing 100094, People's Republic of China, [b]CAS Key Laboratory of Pathogenic Microbiology and Immunology (CASPMI), Institute of Microbiology, Chinese Academy of Sciences, Beijing 100101, People's Republic of China, [c]Beijing Institutes of Life Science, Chinese Academy of Sciences, Beijing 100101, People's Republic of China, and [d]National Laboratory of Biomacromolecules, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, People's Republic of China

Correspondence e-mail: xiachun@cau.edu.cn

Received 8 May 2011
Accepted 14 September 2011



C-reactive protein (CRP) is an acute phase protein that is found in blood, the concentration of which in plasma rises rapidly in response to inflammation. It functions as a pattern-recognition molecule, recognizing dead cells and various pathogenic agents and eliminating them by utilizing the classical complement pathway and activating macrophages. CRP is phylogenetically highly conserved in invertebrates and mammals. To date, information on the *CRP* gene has been reported from numerous species of animals, but little is known about the structure of CRP from species other than humans. In order to solve the structure of CRP from bony fish, the *CRP* gene from zebrafiah (*Danio rerio*) was cloned and expressed in *Escherichia coli*. The zebrafish CRP (*Dare*-CRP) was then purified and crystallized. The crystal diffracted to 2.3 Å resolution and belonged to space group *R*3, with unit-cell parameters $a = b = 114.7$, $c = 61.0$ Å. The Matthews coefficient and solvent content were calculated to be 3.28 Å^3 Da^{-1} and 62.55%, respectively. Determination of the zebrafish CRP structure should be helpful in investigating the evolution of CRPs in the innate immune system.

1. introduction

C-reactive protein (CRP) is an ancient protein that has been studied for about 80 years since its discovery in 1930 in Oswald Avery's laboratory (Tillett & Francis, 1930). CRP is typically comprised of five identical subunits that are held together by noncovalent forces (Shrive *et al.*, 1996). It was originally defined as an acute phase protein as its concentration increases immediately and dramatically in the blood during the early stages of inflammation. Presently, as a member of the classical pentraxins, accumulated studies on CRP have shed light on the role of this molecule in various important disease processes such as cardiovascular, infectious and autoimmune diseases (Agrawal *et al.*, 2009).

CRP genes have been found in almost all animals from the arthropod *Limulus* to humans. It is interesting to note that the primary sequences of different animal CRPs share low identity. In *Limulus* CRP displays obvious polymorphism (Iwaki *et al.*, 1999), but this phenomenon is not found in CRPs from humans or other mammals. In addition to the classical pentamer, the oligomeric form of different CRPs can also vary (Eisenhardt *et al.*, 2009; Shrive *et al.*, 2009). Furthermore, mammalian CRPs possess a number of other important properties. For example, CRP serves as a pattern-recognition molecule that recognizes pathogenic agents and binds to C1q and Fcγ receptors (FcγRs) to eliminate these pathogens (Lu *et al.*, 2008). Although C1q and FcγRs genes have been identified from primal fish to humans, it is still unknown whether fish CRP can bind these two molecules. Here, we report the cloning, expression, purification and crystallization of recombinant CRP (*Dare*-CRP) from zebrafish (*Danio rerio*).

2. Materials and methods

2.1. Cloning and expression

The *CRP* gene was amplified from a *D. rerio* cDNA library using *Ex Taq* DNA polymerase [Takara Biotechnology (Dalian) Co. Ltd]. After the amplification of the target gene with the sense primer

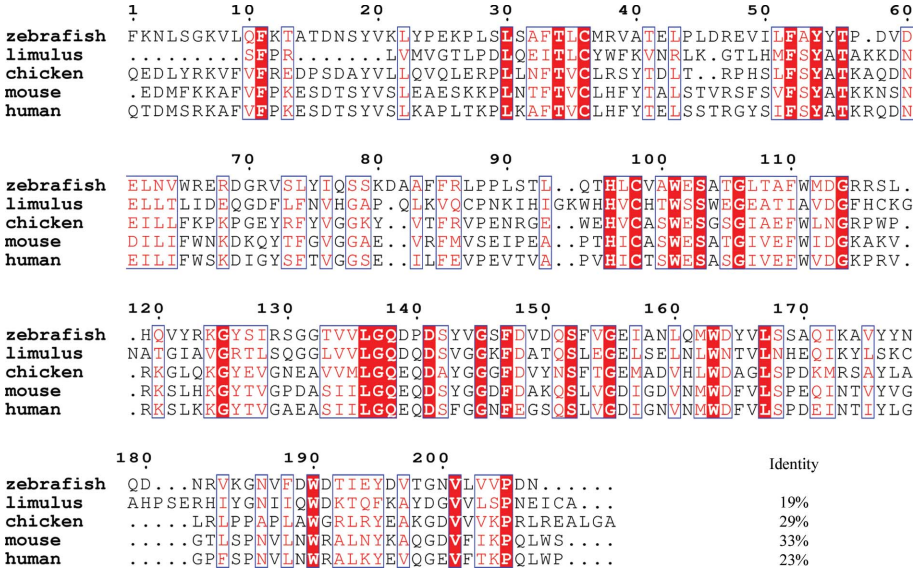


Figure 1
Multiple sequence alignment of CRPs from different species. The sequence alignment was performed using *ClustalW*2. The sequences aligned are as follows: zebrafish (GenBank accession No. JF772178), *Limulus* (GenBank accession No. AAA28270), chicken (GenBank accession No. ABD16281), mouse (GenBank accession No. CAA31928) and human (GenBank accession No. AAA52075). Identical amino acids are shown in white text on a red background, while similar amino acids are shown in red text on a white background. The alignment scores between the *Dare*-CRP sequence and the other sequences are shown at the end of the sequences.

5′-CGGAATTCTTTAAAAATCTGAGCGGTAAAGTG-3′ and the antisense primer 5′-CCGCTCGAGTTATCAGTTATCTGGAACC-ACAAGCA-3′, the PCR product was digested with the restriction enzymes *Nde*I and *Xho*I and then inserted into the pET21a vector (Novagen, Merck KGaA, Darmstadt, Germany). Recombinant *Dare*-CRP protein and its selenomethionine-substituted form were expressed in the form of inclusion bodies in *Escherichia coli* strain BL21 (DE3) and the methionine-auxotrophic *E. coli* strain B834 (DE3) (Blackburn *et al.*, 1999), respectively (Novagen, Merck KGaA, Darmstadt, Germany). When the OD_{600} reached about 0.6, 2 m*M* isopropyl *β*-D-1-thiogalactopyranoside (IPTG) was added to both media. After 8 h, the cells were harvested by centrifugation and lysed by sonication. Pure inclusion bodies were obtained by washing the pellet three times with a solution consisting of 20 m*M* Tris–HCl, 100 m*M* NaCl, 1 m*M* EDTA, 1 m*M* DTT and 0.5% Triton X-100 (Garboczi *et al.*, 1996; Zhang *et al.*, 2010). The inclusion bodies were dissolved to a concentration of 30 mg ml^{-1} in a buffer consisting of 6 *M* guanidine–HCl (Gua–HCl), 50 m*M* Tris–HCl pH 8.0, 10 m*M* EDTA, 100 m*M* NaCl, 10%(*v*/*v*) glycerine and 10 m*M* DTT (Zhang *et al.*, 2010).

2.2. Comparison of the *Dare*-CRP sequence with those of CRPs from other species

Multiple sequence alignment of CRPs from different species was performed using the *ClustalW*2 program (Fig. 1). Identical amino acids are shown as white text on a red background, while similar amino acids are shown as red text on a white background. Alignment scores between the *Dare*-CRP sequence and the other sequences are shown at the ends of the sequences. The sequences aligned are as follows: zebrafish (GenBank accession No. JF772178), *Limulus* (GenBank accession No. AAA28270), chicken (GenBank accession No. ABD16281), mouse (GenBank accession No. CAA31928) and human (GenBank accession No. AAA52075).

2.3. Refolding and purification of the *Dare*-CRP protein

The preparation of native and selenomethionine-substituted *Dare*-CRP protein was essentially carried out as described previously (Garboczi *et al.*, 1996) with modifications in our laboratory (Zhang *et al.*, 2010). Briefly, the *Dare*-CRP inclusion bodies were dissolved in a solution consisting of 6 *M* Gua–HCl, 50 m*M* Tris–HCl pH 8.0. *Dare*-CRP was refolded by the gradual dilution method. After 12 h incubation at 277 K, the soluble *Dare*-CRP protein was concentrated and purified using a Superdex 200 size-exclusion column (GE Healthcare; Zhang *et al.*, 2010).

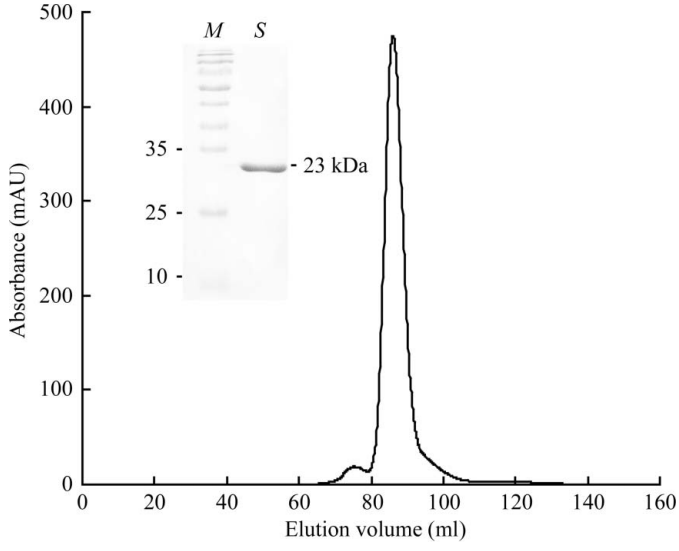


Figure 2
Purification of *Dare*-CRP. The gel-filtration profile of refolded *Dare*-CRP on Superdex G-200 FPLC. The insert shows the reduced SDS–PAGE gel (15%) of the corresponding purified protein; lane *M* contains molecular-weight markers (labelled in kDa) and lane *S* contains the purified protein sample.

2.4. Crystallization

For crystallization, purified *Dare*-CRP protein was dissolved at concentrations of 2.5 and 6 mg ml^{-1} in 50 m*M* NaCl, 20 m*M* Tris–HCl pH 8.0. Screening was carried out with the Index, Crystal Screen, Crystal Screen 2, Crystal Screen Cryo and Crystal Screen 2 Cryo kits (Hampton Research, California, USA; McFerrin & Snell, 2002). The sitting-drop vapour-diffusion method was used for crystal growth at 296 K. Drops were prepared by mixing 1 µl protein solution with 1 µl reservoir solution and were equilibrated against 150 µl of the same reservoir solution. Native and selenomethionine-substituted *Dare*-CRP crystals appeared after 7 d in condition No. 39 [0.085 *M* Na HEPES pH 7.5, 1.7%(*v*/*v*) polyethylene glycol 400, 1.7 *M* ammonium sulfate, 15%(*v*/*v*) glycerol] of Crystal Screen Cryo without any optimization.

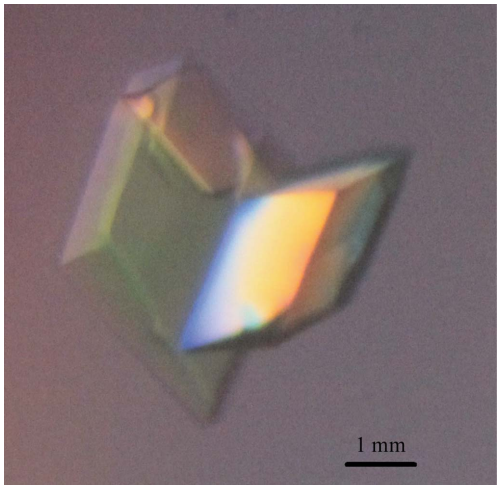


Figure 3
Typical crystals of native *Dare*-CRP.

2.5. Data collection and processing

Diffraction data for native *Dare*-CRP were collected on an R-AXIS IV^{++} image-plate detector using a Rigaku rotating-anode X-ray generator with a radiation wavelength of 1.5418 Å. Data for selenomethionine-substituted *Dare*-CRP were collected on beamline NE3A at the KEK synchrotron facility (Tsukuba, Japan) at a wavelength of 1.0 Å using an ADSC Q270 imaging-plate detector. Both native and selenomethionine-substituted *Dare*-CRP crystals were cryoprotected by adding 33%(*v*/*v*) glycerol. The data were processed and scaled with *HKL*-2000 (Otwinowski & Minor, 1997). Selenomethionine-substituted *Dare*-CRP crystals diffracted to 1.7 Å resolution, while the native crystals diffracted to 2.3 Å resolution. Structure determination will be carried out using the multiple-wavelength anomalous diffraction (MAD) method.

3. Results and discussion

A multiple sequence alignment of CRP amino-acid sequences from GenBank was carried out with *ClustalW*2 (Fig. 1). According to the results of multiple sequence alignment, *Dare*-CRP shares the lowest identity with *Limulus* CRP and resembles mammalian CRPs more closely than avian CRP. Although CRPs exist in invertebrates and vertebrates, they share relatively low similarity (19–33%).

It is often the case that proteins that are not correctly folded will appear in the form of precipitation, aggregation *etc*. Incorrectly folded proteins are unstable and always remain insoluble during the course of refolding. In contrast, correctly refolded protein exists as a monomer, dimer or higher order oligomer, has good solubility and often maintains good stability. The refolding efficiency of *Dare*-CRP inclusion bodies was about 5%. According to the size-exclusion chromatography profile (Fig. 2) and the molecular weight of the *Dare*-CRP monomer (23.5 kDa), soluble *Dare*-CRP exists mainly as a dimer, with a small proportion existing as a higher order oligomer (corresponding to the main peak and the small peak, respectively, in

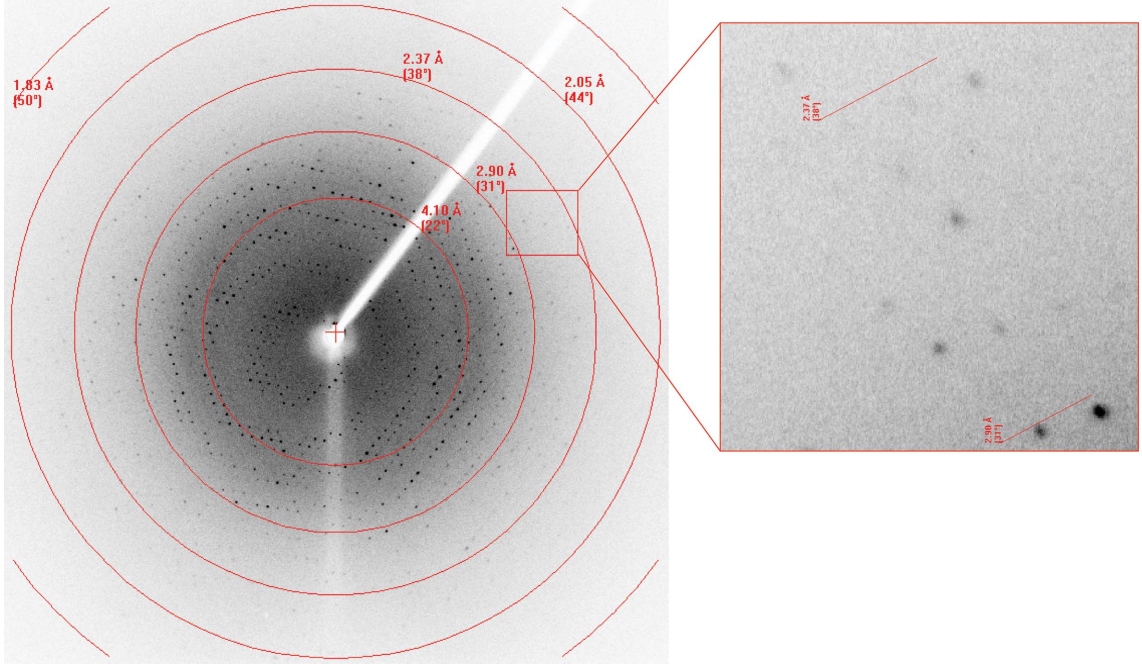


Figure 4
A typical diffraction pattern from a native *Dare*-CRP crystal.

Table 1
X-ray diffraction data and processing statistics of the refined structure.

Values in parentheses are for the highest resolution shell.

	Native	Selenomethionine
Space group	*R*3	*R*3
Unit-cell parameters (Å)	a = 114.7, b = 114.7, c = 61.0	a = 114.7, b = 114.7, c = 60.9
Resolution range (Å)	50–2.3 (2.38–2.30)	50–1.7 (1.76–1.70)
Total No. of reflections	76884	342067
No. of unique reflections	12957	32376
Completeness (%)	99.2 (93.1)	98.1 (93.5)
Average $I/\sigma(I)$	25.4 (5.3)	30.6 (5.2)
$R_{\rm merge}$† (%)	6.9 (25.8)	6.8 (28.7)
Average multiplicity	5.9 (5.3)	10.6 (9.0)

† $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the observed intensity and $\langle I(hkl)\rangle$ is the average intensity from multiple measurements.

Fig. 2). One possible reason for this might be that our *Dare*-CRP was produced in *E. coli*, whereas classical human CRP, which was directly obtained from serum, exists in its native conformation as a pentamer (Shrive *et al.*, 1996). Another possibility is that the dimeric form of *Dare*-CRP is one form of native CRP in zebrafish, as monomeric, hexameric, heptameric and octameric pentraxins have been reported for different species (Eisenhardt *et al.*, 2009; Shrive *et al.*, 2009). The successful production of soluble CRP using a bacterial expression system was first reported by Dortay *et al.* (2011), thus providing a new way of obtaining CRP. The function of this dimeric *Dare*-CRP will be reported in due course.

Optimal *Dare*-CRP crystals appeared after 7 d (Fig. 3). The selenomethionine-subsitituted crystals diffracted to a maximum resolution of 1.7 Å, whereas the native crystals diffracted to 2.3 Å resolution (Fig. 4). Both crystals belonged to space group *R*3. The Matthews coefficient was calculated to be 3.28 Å^3 Da^{-1} (Table 1). There are three methionine residues in the *Dare*-CRP sequence and we intend to determine the structure of *Dare*-CRP using the multiple-wavelength anomalous diffraction (MAD) method. To date, only structures of human CRP (PDB entry 1gnh; Shrive *et al.*, 1996) and of *Limulus* serum amyloid P component (SAP; a highly identical classical pentraxin; PDB entry 3flt; Shrive *et al.*, 2009) have been reported. The monomer structures of human CPR and *Limulus* SAP are both comprised of two main β-sheets, one α-helix and several coils. Human CRP exists in both pentameric and monomeric forms, whereas *Limulus* SAP is reported to form a wider range of oligomers including hexamers, heptamers and octamers. In the absence of any classical pentraxin structures from intermediate species such as bony fish, our results should prove helpful in order to investigate the evolution of pentraxins in the animal kingdom.

This work was supported by grants from the Ministry of Science and Technology of China (863 Program; 2007AA09Z424), the Key National Natural Science Foundation of China (U0631009) and the National Key Basic Research Program of China (973 Program; 2007CB815805). We thank Professor George F. Gao (Institute of Microbiology, Chinese Academy of Sciences) for helpful suggestions. The authors declare no competing financial interests.

References

Agrawal, A., Singh, P. P., Bottazzi, B., Garlanda, C. & Mantovani, A. (2009). *Adv. Exp. Med. Biol.* **653**, 98–116.
Blackburn, N. J., Ralle, M., Gomez, E., Hill, M. G., Pastuszyn, A., Sanders, D. & Fee, J. A. (1999). *Biochemistry*, **38**, 7075–7084.
Dortay, H., Schmöckel, S. M., Fettke, J. & Mueller-Roeber, B. (2011). *Protein Expr. Purif.* **78**, 55–60.
Eisenhardt, S. U., Habersberger, J. & Peter, K. (2009). *Trends Cardiovasc. Med.* **19**, 232–237.
Garboczi, D. N., Ghosh, P., Utz, U., Fan, Q. R., Biddison, W. E. & Wiley, D. C. (1996). *Nature (London)*, **384**, 134–141.
Iwaki, D., Osaki, T., Mizunoe, Y., Wai, S. N., Iwanaga, S. & Kawabata, S. (1999). *Eur. J. Biochem.* **264**, 314–326.
Lu, J., Marnell, L. L., Marjon, K. D., Mold, C., Du Clos, T. W. & Sun, P. D. (2008). *Nature (London)*, **456**, 989–992.
McFerrin, M. B. & Snell, E. H. (2002). *J. Appl. Cryst.* **35**, 538–545.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Shrive, A. K., Burns, I., Chou, H.-T., Stahlberg, H., Armstrong, P. B. & Greenhough, T. J. (2009). *J. Mol. Biol.* **386**, 1240–1254.
Shrive, A. K., Cheetham, G. M., Holden, D., Myles, D. A., Turnell, W. G., Volanakis, J. E., Pepys, M. B., Bloomer, A. C. & Greenhough, T. J. (1996). *Nature Struct. Biol.* **3**, 346–354.
Tillett, W. S. & Francis, T. (1930). *J. Exp. Med.* **52**, 561–571.
Zhang, J., Chen, Y., Gao, F., Chen, W., Qi, J. & Xia, C. (2010). *Acta Cryst.* F**66**, 99–101.