Crystallization and preliminary X-ray diffraction analysis of the two distinct types of zebrafish β_2 -microglobulin

Zhaosan Chen,^a Nianzhi Zhang,^a Shuangshuang Lu,^a Mansoor Tariq,^a Junya Wang^a and Chun Xia^{a,b,*}

Received 12 February 2015

Accepted 21 March 2015

Edited by N. Sträter, University of Leipzig, Germany

Keywords: zebrafish; β_2 -microglobulin; MHC class I.

^aDepartment of Microbiology and Immunology, College of Veterinary Medicine, China Agricultural University, Beijing 100193, People's Republic of China, and ^bKey Laboratory of Animal Epidemiology and Zoonosis, Ministry of Agriculture, Beijing, People's Republic of China. *Correspondence e-mail: xiachun@cau.edu.cn

β_2 -Microglobulin (β_2 m) noncovalently associates with the heavy chain of major histocompatibility complex class I (MHC I) molecules, which bind foreign antigen peptides to control the cytotoxic T lymphocyte (CTL) immune response. In contrast to mammals, there are distinct types of β_2 ms derived from two loci in a number of teleost species. In order to clarify the structures of the β_2 ms, the zebrafish (*Danio rerio*) β_2 ms *Dare- β_2 m-I* and *Dare- β_2 m-II* were expressed in *Escherichia coli*, purified and crystallized, and diffraction data were collected to 1.6 and 1.9 Å resolution, respectively. Both crystals belonged to space group $P2_12_12_1$. The unit-cell parameters were determined to be $a = 38.2$, $b = 50.4$, $c = 50.9$ Å for *Dare- β_2 m-I* and $a = 38.9$, $b = 52.7$, $c = 65.8$ Å for *Dare- β_2 m-II*. Each asymmetric unit was constituted of one molecule, with Matthews coefficients of 2.22 and 3.01 Å³ Da⁻¹ and solvent contents of 45 and 59% for *Dare- β_2 m-I* and *Dare- β_2 m-II*, respectively. These two β_2 m structures will provide relevant information for further studies of the structures of the MHC I complex.

1. Introduction

β_2 -Microglobulin (β_2 m) constitutes the light chain of major histocompatibility complex class I (MHC I) molecules, which play an important role in recognition of the T-cell antigen receptor (TCR) and induction of the cytotoxic T lymphocyte (CTL) immune response (Garboczi *et al.*, 1996). β_2 m noncovalently binds to the heavy chain to stabilize the three-dimensional structure of the MHC I complex (Bjorkman *et al.*, 1987). The mature β_2 m proteins of cartilaginous fish, teleosts, amphibians and mammals are 96, 97, 97 and 99 residues in length, respectively, and have similar structures according to the present data (Hao *et al.*, 2006; Lu *et al.*, 2012). It has been revealed that the structures of β_2 ms are conserved across different species. However, two β_2 m loci that express two distinct types of molecules have been observed in sturgeon, zebrafish, rainbow trout and grass carp, in contrast to mammals, in which β_2 m is encoded by a single locus (Lundqvist *et al.*, 1999; Magor *et al.*, 2004; Hao *et al.*, 2006). This implied that the combinations of β_2 ms and MHC I in teleosts may have their own characteristics.

The first three-dimensional structure of β_2 m to be determined was that of bovine β_2 m in 1985 (Becker & Reeke, 1985). Since then, structures of β_2 ms from mammals and chicken have been solved either in the monomeric form or as the light

Table 1
Macromolecule-production information.

	<i>Dare-β₂m-I</i>	<i>Dare-β₂m-II</i>
Source organism	<i>D. rerio</i>	<i>D. rerio</i>
DNA source	Zebrafish cDNA library	Zebrafish cDNA library
Forward primer	CCAACATATGAAAGTCTCCACTCCGAAAG	CCAACATATGAAAGAAACCCGCGCGAAAGT
Reverse primer	CCGCTCGAGTTACATGTTGGGCTCCCA	CCGCTCGAGTTACATGTCAGGCTCCCATGTGTA
Expression vector	pET-21a(+)	pET-21a(+)
Expression host	<i>E. coli</i> strain BL21 (DE3)	<i>E. coli</i> strain BL21 (DE3)
Complete amino-acid sequence of the construct produced	KVSTPKVHVYSHFPGEYKPNLTICVSSFHPPDISIELLKNQVMSDT-KQTDLAFEKGWQFHLTKSVAFTPEKGEYTCVRHMKETKFSWEPNM	KETPPKVQVYSRNPGEYDKNTLICVSGFHPPDITIELLKNQPIEKT-KQTDLAFEQGWQFHLTKFAAFKPRPGEVYTCQVRHLTETKTYTWEPDM

Table 2
Crystallization conditions.

	<i>Dare-β₂m-I</i>	<i>Dare-β₂m-II</i>
Method	Sitting-drop vapour diffusion	Sitting-drop vapour diffusion
Plate type	48-well VDX plate	48-well VDX plate
Temperature (K)	291	291
Protein concentration (mg ml ⁻¹)	4	7
Buffer composition of protein solution	20 mM Tris-HCl pH 8.0, 50 mM NaCl	20 mM Tris-HCl pH 8.0, 50 mM NaCl
Composition of reservoir solution	0.1 M Tris pH 8.5, 25% (w/v) polyethylene glycol 3350	4.0 M sodium formate
Volume and ratio of drop	2 µl, 1:1 protein solution:reservoir solution	2 µl, 1:1 protein solution:reservoir solution
Volume of reservoir (µl)	200	200

chains of MHC I complexes (Li *et al.*, 2011; Zhang *et al.*, 2012). However, only the structure of the grass carp β_2m type II molecule has been determined from fish (Chen *et al.*, 2010), although an X-ray diffraction data set has been obtained for nurse shark β_2m (Lu *et al.*, 2012). Zebrafish (*Danio rerio*) is regarded as an ideal animal model to study evolution and may help us to reveal the evolution of immune molecules. In order to determine the three-dimensional structures of the two types of zebrafish β_2m s (*Dare-β₂m-I* and *Dare-β₂m-II*), the two genes were cloned and expressed and the proteins were purified and crystallized. Finally, diffraction data were collected from crystals of *Dare-β₂m-I* and *Dare-β₂m-II* to 1.6 and 1.9 Å resolution, respectively. To our knowledge, this is the first attempt to solve the structures of both types of β_2m from the same fish species, and the results will facilitate the solution of the structures of their MHC I molecules.

2. Materials and methods

2.1. Macromolecule production

The *Dare-β₂m-I* and *Dare-β₂m-II* genes were amplified from a zebrafish cDNA library maintained by our research group (Chen *et al.*, 2011) using *ExTaq* DNA polymerase [Takara Biotechnology (Dalian) Co. Ltd]. The coding sequences of the extracellular domains were cloned with two different pairs of primers (Table 1). A unique *NdeI* restriction site was added to the forward primers and a stop codon and a specific *XhoI* recognition site were added to the reverse primers. The PCR products were digested with *NdeI* and *XhoI* at 310 K for 8 h and then separately ligated into the prokaryotic expression vector pET-21a(+) (Novagen, Merck KGaA, Darmstadt, Germany) by T4 DNA ligase. The recombinant plasmids were transformed into *Escherichia coli* strain BL21

(DE3) (Table 1). *Dare-β₂m-I* and *Dare-β₂m-II* were expressed using the same protocol. When the OD₆₀₀ reached about 0.6, a final concentration of 1 mM isopropyl β -D-1-thiogalactopyranoside (IPTG) was used to induce protein expression. After 5 h, the bacteria were harvested by centrifugation at 6000g for 10 min and resuspended in cold phosphate-buffered saline (PBS). In order to obtain pure inclusion bodies, after lysis by sonication, the sample was centrifuged for 10 min at 16 000g and the precipitate was then washed three times with buffer (20 mM Tris-HCl pH 8.0, 100 mM NaCl, 1 mM EDTA, 1 mM DTT) containing 0.5% Triton X-100 and once with the same buffer without Triton X-100. The inclusion bodies were finally dissolved in guanidinium chloride (Gua-HCl) buffer [6 M Gua-HCl, 50 mM Tris-HCl pH 8.0, 100 mM NaCl, 10 mM EDTA, 10 mM DTT, 10% (v/v) glycerine] overnight at a concentration of 30 mg ml⁻¹.

Essentially, the soluble *Dare-β₂m-I* and *Dare-β₂m-II* proteins were prepared separately as described previously, with modifications introduced by our group (Jiang *et al.*, 2014). Each *Dare-β₂m* was gradually dissolved in refolding buffer (100 mM Tris-HCl pH 8.0, 2 mM EDTA, 400 mM L-arginine-HCl, 0.5 mM oxidized glutathione, 5 mM reduced glutathione) using the dilution method. After stirring at 277 K for 12 h, the solution was concentrated and purified by chromatography using a Superdex 200 16/60 HiLoad (GE Healthcare) size-exclusion column followed by Resource-Q (GE Healthcare) anion-exchange chromatography.

2.2. Crystallization

For crystallization, each purified *Dare-β₂m* was concentrated to 4 and 7 mg ml⁻¹ in crystallization buffer (20 mM Tris-HCl pH 8.0, 50 mM NaCl) using Amicon Ultra-15 centrifugal filter devices (Millipore). Screening for crystal-

lization conditions was set up with the Index, Crystal Screen and Crystal Screen 2 kits (Hampton Research, Laguna Hills, USA) at 291 K by the sitting-drop vapour-diffusion method. Drops prepared by mixing protein solution with reservoir solution in a 1:1 ratio were equilibrated against 200 μ l of the same reservoir solution in VDX plates (Hampton Research).

Crystals suitable for data collection were obtained after a week using the optimized conditions summarized in Table 2.

2.3. Data collection and processing

Diffraction data were collected on beamline BL17U at Shanghai Synchrotron Radiation Facility (SSRF; Shanghai,

Figure 1

Purification of refolded *Dare-β₂m-I* and *Dare-β₂m-II* by Superdex 200 16/60 HiLoad gel-filtration and Resource Q anion-exchange chromatography (GE Healthcare). (a) Gel-filtration profile of *Dare-β₂m-I*. Inset, reduced SDS-PAGE gel (15%) of the corresponding purified protein (~11 kDa). Lane M contains molecular-weight markers (labelled in kDa) and lane S contains the purified sample. (b) Result of the further purification of the collected *Dare-β₂m-I* protein by anion-exchange chromatography. The target protein was eluted at an NaCl concentration of 2–8%. Inset, 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 sample. (c) Gel-filtration profile of *Dare-β₂m-II*. Inset, reduced SDS-PAGE gel (15%) of the corresponding purified protein (~11 kDa). Lane M contains molecular-weight markers (labelled in kDa) and lane S contains the purified sample. (d) Result of the further purification of the collected *Dare-β₂m-II* protein by anion-exchange chromatography. The target protein was eluted at an NaCl concentration of 9–14%. Inset, 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 sample.

(a)

(b)

Figure 2

Typical crystals of *Dare-β₂m*s. (a) Diffracting crystal of *Dare-β₂m*-I. (b) Diffracting crystal of *Dare-β₂m*-II.

Table 3

Data-collection statistics.

Values in parentheses are for the outer shell.

	<i>Dare-β₂m</i> -I	<i>Dare-β₂m</i> -II
Diffraction source	BL17U, SSRF	BL17U, SSRF
Wavelength (Å)	0.97916	0.97916
Temperature (K)	100	100
Detector	ADSC 315 CCD	ADSC 315 CCD
Crystal-to-detector distance (mm)	120	200
Rotation range per image (°)	1	1
Total rotation range (°)	200	120
Exposure time per image (s)	1.7	0.8
Space group	<i>P</i> ₂ ₁ <i>2</i> ₁ <i>2</i> ₁	<i>P</i> ₂ ₁ <i>2</i> ₁ <i>2</i> ₁
Unit-cell parameters (Å, °)	<i>a</i> = 38.2, <i>b</i> = 50.4, <i>c</i> = 50.9, α = 90, β = 90, γ = 90	<i>a</i> = 38.9, <i>b</i> = 52.7, <i>c</i> = 65.8, α = 90, β = 90, γ = 90
<i>V_M</i> (Å ³ Da ⁻¹)	2.22	3.01
Solvent content (%)	45	59
Resolution range (Å)	50–1.60 (1.66–1.60)	50–1.90 (1.97–1.90)
Total No. of reflections	100081	52163
No. of unique reflections	13101	11171
Completeness (%)	97.5 (88.8)	100.0 (100.0)
Multiplicity	7.6 (6.5)	4.7 (4.8)
<i>I</i> /σ(<i>I</i>)	29.5 (5.6)	13.8 (3.4)
<i>R</i> _{merge} † (%)	6.4 (31.1)	11.9 (53.1)

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

People's Republic of China) at a wavelength of 0.97916 Å using an ADSC 315 CCD detector. Each crystal was first soaked in reservoir solution supplemented with 15% (v/v) glycerol as a cryoprotectant for about 10 s and then flash-cooled directly in a stream of gaseous nitrogen at 100 K. The collected data were indexed, integrated, corrected for absorption, scaled and merged using *HKL*-2000 (Otwinowski & Minor, 1997).

3. Results and discussion

Soluble *Dare-β₂m*-I and *Dare-β₂m*-II were obtained by *in vitro* refolding of inclusion bodies, which resulted in an efficiency

Figure 3

Diffraction pattern of the *Dare-β₂m*-I crystal. Spots corresponding to diffraction to high resolution are highlighted in the box.


```

Zebrafish type I  KVSTPKVHVYSHFPGEYKPNLTICYVSSFHPDISIELL 40
Zebrafish type II -etp---q---rn---d-d-----g-----t--- 40

Zebrafish type I  KNGQVMSDTKQTDLAFEKGWQFHLTKSVAFTPEKGDEYTC 80
Zebrafish type II ----piek-t-----q-f-----fa--k-rp-ev--- 80

Zebrafish type I  SVRHMKETKKFSWEPNM 97      Identity
Zebrafish type II q---lt---tyt---d- 97      68%

```

Figure 4

Amino-acid sequence alignment of zebrafish β_2 ms. The alignment was performed by DNAMAN. The sequence identity between *Dare- β_2 m-I* and *Dare- β_2 m-II* is given after the sequence alignment. The GenBank accession Nos. of the sequences used are as follows: zebrafish type I, NM_131163; zebrafish type II, NM_213126.

of approximately 10%. Each *Dare- β_2 m* was purified by two chromatographic steps. Firstly, the protein was purified by Superdex 200 16/60 HiLoad size-exclusion chromatography. Both elution profiles showed a main peak at 100 ml (~ 11 kDa; Figs. 1*a* and 1*c*). The collected protein from the target peak was then further purified by Resource-Q anion-exchange chromatography. *Dare- β_2 m-I* eluted at an NaCl concentration of 2–8% (Fig. 1*b*), whereas *Dare- β_2 m-II* eluted at 9–14% NaCl (Fig. 1*d*). Finally, both proteins were obtained with high purity as confirmed by SDS–PAGE, which showed a single pure band (see inset in Fig. 1).

After purification and concentration, crystallization screening was carried out using crystallization kits. The optimized crystallization conditions are summarized in Table 2. Ideal crystals of *Dare- β_2 m-I* and *Dare- β_2 m-II* (Fig. 2) appeared after one week in Index solution No. 45 and Crystal Screen solution No. 33, respectively. The crystals of *Dare- β_2 m-I* and *Dare- β_2 m-II* both belonged to space group $P2_12_12_1$ and diffracted to 1.6 Å (Fig. 3) and 1.9 Å resolution, respectively. Unit-cell parameters were determined as $a = 38.2$, $b = 50.4$, $c = 50.9$ Å for *Dare- β_2 m-I* and $a = 38.9$, $b = 52.7$, $c = 65.8$ Å for *Dare- β_2 m-II*. Selected data statistics are listed in Table 3.

In contrast to the case in mammals, sequence polymorphism of β_2 m genes has been found in fish. In *D. rerio*, two distinct types of β_2 m molecules, *Dare- β_2 m-I* and *Dare- β_2 m-II*, are expressed by two loci. The identity of the two distinct *Dare- β_2 ms* is 68% according to sequence alignment (Fig. 4). Here, crystals of both *Dare- β_2 m-I* and *Dare- β_2 m-II* were grown and

investigated by X-ray diffraction. The results will be helpful in studying the function of the different types of β_2 m, in particular in the formation of the MHC I complexes of zebrafish. Structure determination and refinement using the molecular-replacement method are currently being investigated.

Acknowledgements

The work was supported by the State Key Project of the National Natural Science Foundation of China (grant No. 31230074) and the 973 Project of the National Key Basic Research Program of China (grant No. 2013CB835302). We acknowledge the help from the staff at the Shanghai Synchrotron Radiation Facility of China (SSRF) in the diffraction of crystals.

References

- Becker, J. W. & Reeke, G. N. Jr (1985). *Proc. Natl Acad. Sci. USA*, **82**, 4225–4229.
- Bjorkman, P. J., Saper, M. A., Samraoui, B., Bennett, W. S., Strominger, J. L. & Wiley, D. C. (1987). *Nature (London)*, **329**, 506–512.
- Chen, R., Qi, J., Yao, S., Pan, X., Gao, F. & Xia, C. (2011). *Acta Cryst. F* **67**, 1633–1636.
- Chen, W., Gao, F., Chu, F., Zhang, J., Gao, G. F. & Xia, C. (2010). *J. Biol. Chem.* **285**, 22505–22512.
- Garboczi, D. N., Ghosh, P., Utz, U., Fan, Q. R., Biddison, W. E. & Wiley, D. C. (1996). *Nature (London)*, **384**, 134–141.
- Hao, H.-F., Yang, T.-Y., Yan, R.-Q., Gao, F.-S. & Xia, C. (2006). *Fish Shellfish Immunol.* **20**, 118–123.
- Jiang, B., Liu, Y., Chen, R., Wang, Z., Tariq, M. & Xia, C. (2014). *Acta Cryst. F* **70**, 1072–1075.
- Li, X., Liu, J., Qi, J., Gao, F., Li, Q. R., Li, X., Zhang, N., Xia, C. & Gao, G. F. (2011). *J. Virol.* **85**, 6038–6048.
- Lu, S., Yao, S., Chen, R., Zhang, N., Chen, J., Gao, F. & Xia, C. (2012). *Acta Cryst. F* **68**, 460–463.
- Lundqvist, M. L., Appelkvist, P., Hermesen, T., Pilstrom, L. & Stet, B. J. (1999). *Immunogenetics*, **50**, 79–83.
- Magor, K. E., Shum, B. P. & Parham, P. (2004). *J. Immunol.* **172**, 3635–3643.
- Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
- Zhang, J., Chen, Y., Qi, J., Gao, F., Liu, Y., Liu, J., Zhou, X., Kaufman, J., Xia, C. & Gao, G. F. (2012). *J. Immunol.* **189**, 4478–4487.