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

Crystallization and X-ray diffraction analysis of human CLEC-2

Aleksandra A. Watson and Christopher A. O'Callaghan*

Henry Wellcome Building of Molecular Physiology, University of Oxford, Roosevelt Drive, Oxford OX3 7BN, England

Correspondence e-mail: chrisoc@ccmp.ox.ac.uk

Received 20 October 2005
Accepted 17 November 2005
Online 24 November 2005

The human C-type lectin-like protein CLEC-2 has recently been shown to be expressed on the surface of platelets and to function as a receptor for the snake-venom protein rhodocytin. The C-type lectin-like domain (CTLD) of CLEC-2 was expressed in *Escherichia coli*, refolded and purified. Crystals of this recombinant CLEC-2 were grown by sitting-drop vapour diffusion using polyethylene glycol (PEG) 6000 as a precipitant. After optimization, crystals were grown which diffracted to 2.0 Å using in-house radiation (λ = 1.5418 Å). These crystals belonged to the orthorhombic space group $P2_12_12_1$, with unit-cell parameters a = 35.407, b = 55.143, c = 56.078 Å. The presence of one molecule per asymmetric unit is consistent with a crystal volume per unit weight (V_M) of 1.82 Å^3 Da^{-1} and a solvent content of 32.6%. These results suggest that crystals producing diffraction of this quality will be suitable for the structural determination of human CLEC-2.

1. Introduction

CLEC-2 was identified by *BLAST* homology searches of EST (expressed sequence tag) databases using sequences of known C-type lectin-like molecules (Altschul *et al.*, 1990). A full-length cDNA was then amplified and isolated from human peripheral blood cell cDNA. Colonna *et al.* (2000) used a similar strategy and named the protein CLEC-2. They observed expression in bone marrow cells, liver, myeloid cells (monocytes, dendritic cells and granulocytes) and in several natural killer (NK) cell clones. The *CLEC-2* gene maps to a region containing genes for a number of other C-type lectin-like molecules (Sobanov *et al.*, 2001) and encodes a 32 kDa type 2 transmembrane protein with an extracellular CTLD (Colonna *et al.*, 2000).

The snake *Calloselasma rhodostoma* (Malayan pit viper) produces a potent venom protein, rhodocytin, which causes platelet aggregation (Huang *et al.*, 1995; Shin & Morita, 1998). Rhodocytin is itself a C-type lectin-like protein and is also termed aggretin (Chung *et al.*, 1999). Using rhodocytin affinity purification and mass spectroscopy, CLEC-2 was shown to be a receptor on the surface of platelets for rhodocytin (Suzuki-Inoue *et al.*, 2005). Binding of CLEC-2 by rhodocytin leads to phosphorylation of a tyrosine residue in the cytosolic tail of CLEC-2, which promotes the binding of Syk, further downstream tyrosine phosphorylation events and activation of PLCγ2 (Navdaev *et al.*, 2001). These signalling events result in platelet activation and aggregation. CLEC-2 is the first C-type lectin-like receptor expressed on platelets shown to signal through this pathway. Structural characterization of CLEC-2 will help to elucidate the mechanisms of rhodocytin recognition and binding.

2. Materials and methods

2.1. Cloning and expression

A DNA fragment encoding the extracellular domain of human CLEC-2 from residues 87 to 229 was amplified using the polymerase chain reaction (PCR) and cloned into the T7 vector pGMT7 (O'Callaghan *et al.*, 1998, 2001) using standard cloning techniques. This plasmid (pOC033) then served as a template for PCR amplifi-

cation using the forward and reverse oligonucleotide primers CO257 (5′-TATATTCATATGGGTCATAAATCCAGCCCCTGTGAC) and CO258 (5′-ATAATAGAATTCCTAGCCAGCCTTCCTCTCACAC-ATTAAATAATG). The amplified product was cut with *Nde*I and *Eco*RI and cloned into pGMT7 cut with the same enzymes. The resulting plasmid pOC189 encoded the extracellular domain of human CLEC-2 from residues 96 to 221 (Fig. 1). This construct also incorporates a mutation which changes the predicted unpaired cysteine residue at position 99 to a serine. All PCRs were performed with cloned *Pfu* polymerase (Stratagene, CA, USA). The sequences of all constructs were confirmed using a 3730xl DNA Analyzer (Applied Biosystems, CA, USA).

Plasmids pOC033 and pOC189 were transformed separately into *Escherichia coli* strain BL21 (DE3) pLysS. For expression, cells were grown in Luria–Bertani medium supplemented with 100 µg ml^{-1} ampicillin (Sigma, UK) and protein expression was induced with 0.5 m*M* isopropyl β-D-thiogalactopyranoside (Sigma, UK) in mid-log phase growth (A_{600} = 0.6). Cells were harvested by centrifugation at 4 h post-induction and recombinant protein was expressed as inclusion bodies.

Harvested cells were resuspended in ice-cold phosphate-buffered saline and lyzed by sonication. Inclusion bodies were purified by repeated washes in 0.5% Triton X-100, 50 m*M* Tris pH 8.0, 100 m*M* NaCl, 0.1% sodium azide. Detergent was removed by washing in 50 m*M* Tris pH 8.0, 100 m*M* NaCl. The resulting protein was solubilized overnight in 6 *M* guanidine hydrochloride, 50 m*M* MES pH 6.5, 10 m*M* ethylene diaminetetraacetic acid (EDTA), 2 m*M* dithiothreitol. Protein purity was assessed by sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE) and quantified using the Bio-Rad protein assay (Bio-Rad, Germany).

2.2. Refolding and purification

To minimize aggregation and to maximize the yield of correctly folded protein, various refolding strategies were tested. Parameters that were varied included pH, redox-couple component ratios, arginine concentration and the method of dilution. Optimal refolding for protein expressed by both constructs, as assessed by size-exclusion gel filtration, was achieved by slow dilution of 10–20 mg guanidine-solubilized CLEC-2 protein into 250 ml 1 *M* L-arginine, 200 m*M* Tris pH 8.0, 2 m*M* EDTA, 10 m*M* reduced glutathione, 0.5 m*M* oxidized glutathione, 0.1 *M* phenylmethylsulfonyl fluoride. The mixture was equilibrated at 277 K for 48 h with slow stirring and then concentrated under nitrogen to a volume of 10 ml over a 10 kDa exclusion membrane (Millipore, USA). Refolded protein was analyzed and purified by gel-filtration chromatography in 20 m*M* Tris pH 8.0, 150 m*M* NaCl with a Superdex 200 26/60 column on an Äkta Purifier (GE Healthcare, Amersham Biosciences AB, Sweden). Correctly sized gel-filtration fractions of protein expressed from the pOC033 plasmid were further purified by anion-exchange chromatography. Peak fractions were analyzed by SDS–PAGE, pooled and buffer-exchanged by overnight dialysis at 277 K into 20 m*M* Tris pH 8.0, 20 m*M* NaCl. This protein was loaded onto a 5 ml HiTrap Q XL column (GE Healthcare, Amersham Biosciences AB, Sweden) and eluted with a 1 *M* NaCl gradient. Protein expressed from the pOC189 plasmid crystallized without the need for anion-exchange purification. The purity and molecular weight of all samples entered into crystal trials was assessed by SDS–PAGE and mass spectroscopy, respectively. Liquid-chromatography electrospray ionization mass spectrometry was performed using a reversed-phase C4 column on an Ultima HPLC (Dionex, UK) connected to a Quadrupole Time-of-Flight Micromass spectrometer (Waters, UK).

2.3. Crystallization of CLEC-2

Correctly folded protein was concentrated to 10.0–12.5 mg ml^{-1} in 20 m*M* Tris pH 8.0, 150 m*M* NaCl for initial crystal trials on the basis of precipitation results using the Hampton Pre-Crystallization Test (Hampton Research, CA, USA). Initial crystallization screens were performed using the sitting-drop vapour-diffusion method. A Robbins Hydra (Robbins Scientific, CA, USA) was used to dispense solutions promoting crystal growth into 96-well plates (Greiner, Austria). Sampled screens included Crystal Screen, Crystal Screen 2, Index, Grid Screen PEG 6000, Grid Screen PEG/LiCl, Grid Screen MPD, Grid Screen Ammonium Sulfate and PEG/Ion (Hampton Research, CA, USA) and Wizard I and II (Emerald BioSystems, WA, USA). Crystallization drops were established with 100 nl protein solution and 100 nl reservoir solution using a Cartesian Dispensing Robot and a Cartesian Microsys dispenser (Genomic Solutions, MI, USA), respectively. Plates were stored in a temperature-controlled TAP Homebase storage vault (The Automation Partnership, Royston, England) at 295 and at 277 K (Walter *et al.*, 2005). Crystals

QSELKGTFKGHKCSPCDTNWRYYGDSCYGFFRHNLTWEESKQYCTDMNATLLKIDNRNIVEYIKARTHLIRWV
GLSRQKSNEVWKWEDGSVISENMFEFLEDGKGNMNCAYFHNGKMHPTFCENKHYLMCERKAGMTKVDQLP

Figure 1
Sequences of protein expressed by plasmids pOC033 and pOC189. The truncated protein expressed by plasmid pOC189 is underlined. Predicted paired cysteines are highlighted in the same colour. The red C is a predicted unpaired cysteine in pOC033 mutated to serine in pOC189.

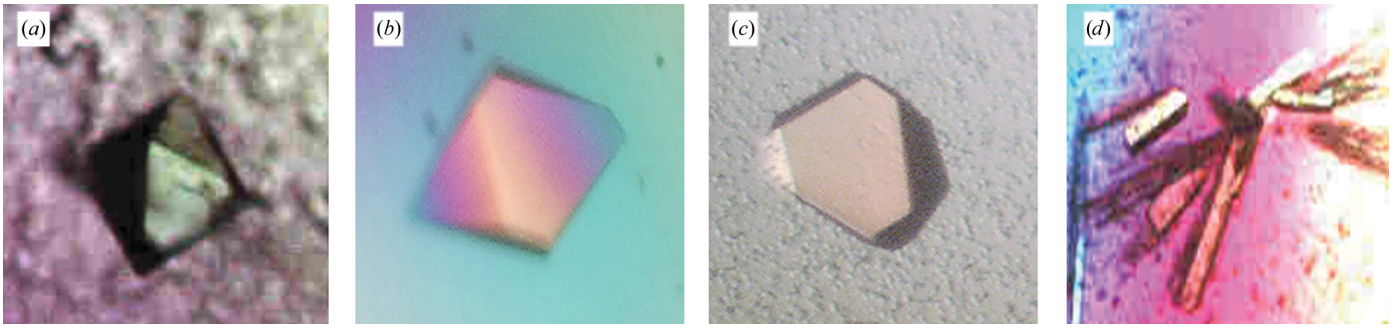

Figure 2
Crystals of the CTLD of CLEC-2. (*a*), (*b*) and (*c*) are large single crystals; (*d*) shows a branching needle crystal form.

Table 1
X-ray diffraction data for human CLEC-2.

Values for the outer resolution shell are shown in parentheses.

Wavelength (Å)	1.5418
Space group	$P2_12_12_1$
Unit-cell parameters	
a (Å)	35.407
b (Å)	55.143
c (Å)	56.078
Matthews coefficient (Å^3 Da^{-1})	1.82
Resolution (Å)	40–2.0 (2.07–2.0)
R_{merge}† (%)	5.9 (13.8)
Completeness (%)	92.0 (94.0)
Average $I/\sigma(I)$	26.2 (12.4)
No. of measured reflections	103691
No. of unique reflections	7838
Redundancy	6.1 (4.4)

† $R_{merge} = \sum|I - \langle I\rangle|/\sum\langle I\rangle$, where I is the intensity of each reflection.

formed reproducibly in 24–48 h at 295 K in a range of conditions. The largest single crystals formed in 0.1 *M* HEPES pH 7.5, 25% PEG 3350 and in 0.1 *M* MES pH 6.5, 20% PEG 6000. Multiple optimization screens were set up around all crystal-forming conditions, both manually and using a RoboGo dispenser (MWG Biotech AG, Germany). Optimization parameters that were varied included buffer composition, molarity and pH, PEG concentration and molecular weight, drop size, protein:reservoir solution ratio and protein age and concentration. Maximum crystal dimensions of 0.65 mm were achieved in 24 m*M* HEPES pH 6.8, 20% PEG 6000. Repeat drops of this condition were dispensed onto polypropylene bridges and contained 2.5 µl reservoir solution and 2.5 µl 13 mg ml^{-1} protein in 20 m*M* Tris pH 8.0, 150 m*M* NaCl.

2.4. Crystallographic studies

Crystals selected for diffraction studies grew in 24 m*M* HEPES pH 6.8, 20% PEG 6000. Crystals grew either as large single crystals in several forms (Figs. 2*a*, 2*b* and 2*c*) or as complex branching needles (Fig. 2*d*). The large single crystals grew to a maximum length of 0.5 mm and diffracted to 2.5 Å. Branching needle crystals were dissected and cleaved with an acupuncture needle prior to mounting. Separated and cleaved needles produced diffraction to a higher resolution (2.0 Å) than the large single variety and were used in all following diffraction studies. Crystals were frozen at 100 K in a nitrogen cryostream. Diffraction was optimal in crystals grown for two weeks. Diffraction to 2.0 Å resolution was achieved using perfluoropolyether oil as a cryoprotectant, but ice rings significantly hindered data processing. The presence of ice rings in the diffraction images was eliminated using 20% glycerol as the cryoprotectant in subsequent studies. X-ray diffraction data were collected in-house using a MAR 345 detector and a Rigaku generator with a copper anode (wavelength = 1.5418 Å) using 1° oscillations. A complete data set was collected from one large two-week-old crystal cleaved from a branching needle complex and mounted in a 0.4–0.7 mm loop. The data were autoindexed and integrated using *DENZO* and scaled using *SCALEPACK* (Otwinowski & Minor, 1997).

3. Results

CLEC-2 was initially expressed using the construct pOC033. However, protein of this length and sequence produced only very small crystals in 0.1 *M* Bis-Tris pH 5.5, 25% PEG 3350, which did not produce any detectable diffraction. A further truncated form of CLEC-2 (pOC189) was generated in which a predicted unpaired cysteine (residue 99) in the construct was changed to a serine to reduce the likelihood of incorrect refolding. Protein expressed and purified from the pOC189 construct proved a better candidate for crystallography in several respects. Expression levels from plasmid pOC189 (approximately 225 mg per litre of culture) were higher than those from pOC033 (less than 100 mg per litre). The relative purity of inclusion bodies from pOC189 was consequently higher and washing with detergent resulted in greater than 95% purity. Much less aggregate was observed in the gel-filtration profile of the CLEC-2 protein encoded by pOC189 compared with that encoded by pOC033. Mass spectroscopy revealed that the protein expressed and purified from the pOC189 plasmid contained three intramolecular disulfide bonds and had a molecular weight of 14 894 Da. This protein crystallized in 24 m*M* HEPES pH 6.8, 20% PEG 6000 as large single crystals or as complex branched needle crystals after two weeks incubation. Large single crystals diffracted to 2.5 Å with space group $P2_12_12_1$ and were not used for data collection. Branching needles had to be dissected with acupuncture needles to facilitate single-crystal mounting and produced 2.0 Å resolution diffraction data when cryoprotected in 20% glycerol. The space group of these crystals was determined by *SCALEPACK* (Otwinowski & Minor, 1997) and *XPREP* (Bruker, Madison, WI, USA) to be $P2_12_12_1$. The crystal volume per unit weight (V_M) was calculated to be 1.82 Å^3 Da^{-1}, with one molecule per asymmetric unit and a solvent content of 32.6% (Table 1; Kantardjieff & Rupp, 2003). CLEC-2 may play a role in platelet function in thrombotic disease. Diffraction data of this quality will provide a good basis for the structural determination of the CTLD of CLEC-2.

We are grateful to Tom Walters, James Brown, Karl Harlos, Charita Christou and Yvonne Jones for helpful discussions. This work was funded by the Medical Research Council.

References

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990). *J. Mol. Biol.* **215**, 403–410.
Chung, C. H, Au, L. C. & Huang, T. F. (1999). *Biochem. Biophys. Res. Commun.* **263**, 723–727.
Colonna, M., Samaridis, J. & Angman, L. (2000). *Eur. J. Immunol.* **30**, 697–704.
Huang, T. F., Liu, C. Z. & Yang, S. H. (1995). *Biochem J.* **309**, 1021–1027.
Kantardjieff, K. A. & Rupp, B. (2003). *Protein Sci.* **12**, 1865–1871.
Navdaev, A., Clemetson, J. M., Polgar, J., Kehrel, B. E., Glauner, M., Magnenat, E., Wells, T. N. & Clemetson, K. J. (2001). *J. Biol. Chem.* **276**, 20882–20889.
O'Callaghan, C. A., Cerwenka, A., Willcox, B. E., Lanier, L. L. & Bjorkman, P. J. (2001). *Immunity*, **15**, 201–211.
O'Callaghan, C. A., Tormo, J., Willcox, B. E., Blundell, C. D., Jakobsen, B. K., Stuart, D. I., McMichael, A. J., Bell, J. I. & Jones, E. Y. (1998). *Protein Sci.* **7**, 1264–1266.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Shin, Y. & Morita, T. (1998). *Biochem. Biophys. Res. Commun.* **245**, 741–745.
Sobanov, Y., Bernreiter, A., Derdak, S., Mechtcheriakova, D., Schweighofer, B., Duchler, M., Kalthoff, F. & Hofer, E. (2001). *Eur. J. Immunol.* **31**, 3493–3503.
Suzuki-Inoue, K., Fuller, G. L., Garcia, A., Eble, J. A., Pohlmann, S., Inoue, O., Gartner, T. K., Hughan, S. C., Pearce, A. C., Laing, G. D., Theakston, R. D., Schweighoffer, E., Zitzmann, N., Morita, T., Tybulewicz, V. L., Ozaki, Y. & Watson, S. P. (2005). In the press.
Walter, T. S., Diprose, J. M., Mayo, C. J., Siebold, C., Pickford, M. G., Carter, L., Sutton, G. C., Berrow, N. S., Brown, J., Berry, I. M., Stewart-Jones, G. B., Grimes, J. M., Stammers, D. K., Esnouf, R. M., Jones, E. Y., Owens, R. J., Stuart, D. I. & Harlos, K. (2005). *Acta Cryst.* D**61**, 651–657.