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

Expression, purification and crystallization of the ecto-enzymatic domain of rat E-NTPDase1 CD39

Xiaotian Zhong,[a]* Madhavan Buddha,[a] Guido Guidotti,[b] Ron Kriz,[a] Will Somers[a] and Lidia Mosyak[a]*

[a]Department of Chemical and Screening Sciences, Wyeth Research, 200 Cambridge Park Drive, Cambridge, MA 02140, USA, and [b]Department of Molecular and Cellular Biology, Harvard University, 7 Divinity Avenue, Cambridge, MA 02138, USA

Correspondence e-mail: xzhong@wyeth.com, lmosyak@wyeth.com

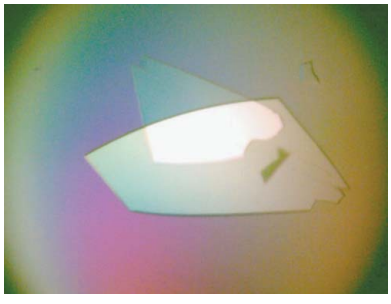
Received 26 August 2008
Accepted 8 October 2008

CD39 is a prototype member of the ecto-nucleoside triphosphate diphosphohydrolase family that hydrolyzes extracellular nucleoside diphosphates and triphosphates in the presence of divalent cations. Here, the expression, purification and crystallization of the ecto-enzymatic domain of rat CD39, sCD39, are described. The 67 kDa secreted soluble glycoprotein was recombinantly overexpressed in a glycosylation mutant CHO line, Lec.3.2.8.1, and purified from conditioned media. Diffraction-quality crystals of sCD39 were produced by the vapor-diffusion method using PEG 3350 and ammonium dihydrogen phosphate as precipitants. The enzyme crystallized in a primitive trigonal form in space group $P3_2$, with unit-cell parameters $a = b = 118.1$, $c = 81.6$ Å and with two sCD39 copies in the asymmetric unit. Several low- to medium-resolution diffraction data sets were collected using an in-house X-ray source. Analysis of the intensity statistics showed that the crystals were invariably merohedrally twinned with a high twin fraction. For initial phasing, a molecular-replacement search was performed against the complete 3.2 Å data set using a maximum-likelihood molecular-replacement method as implemented in *Phaser*. The initial model of the two sCD39 monomers was placed into the $P3_2$ lattice and rigid-body refined and position-minimized with *PHENIX*.

1. Introduction

CD39, originally identified as a B-lymphocyte marker and modulator of homotypic adhesion (Maliszewski *et al.*, 1994), is a prototype member of the ecto-nucleoside triphosphate diphosphohydrolase (E-NTPDase) family (Wang & Guidotti, 1996; Kaczmarek *et al.*, 1996; for a recent review, see Robson *et al.*, 2006). These are ectoenzymes that hydrolyze the terminal phosphoanhydride bonds of extracellular nucleotide diphosphates and triphosphates in the presence of divalent cations and share a set of highly conserved sequences termed apyrase conserved regions (ACRs). ACR1 and ACR4 are similar to the β- and γ-phosphate-binding domains of the actin/HSP70/sugar kinase superfamily. The biological actions of E-NTPDases have many physiological and pathophysiological implications involving the regulated phosphohydrolytic activity on extracellular nucleotides and consequent effects on purinergic signaling (Robson *et al.*, 2006).

The primary structure of CD39 (now referred to as NTPDase1) contains two transmembrane domains near the N- and C-terminal ends as well as a large extracellular enzymatic domain (Maliszewski *et al.*, 1994; Wang & Guidotti, 1996). The eight vertebrate E-NTPDases (NTPDases 1–8) all share a similar membrane arrangement (Bigonnesse *et al.*, 2004), with the exceptions of NTPDases 5 and 6 (Chadwick & Frischauf, 1998). Although the extracellular catalytic domains of NTPDases have a significant degree of sequence identity and share five ACR motifs, they differ in substrate preference, divalent-cation usage and product formation (Robson *et al.*, 2006). NTPDase1 hydrolyzes ATP and ADP similarly, whereas NTPDase2 has a high preference for nucleoside triphosphates. NTPDase3 and NTPDase8 are intermediate between NTPDase1 and NTPDase2, while NTPDases 5 and 6 mainly hydrolyze nucleoside diphosphates. The molecular mechanisms that account for the differences in catalytic properties between these subtypes remain unknown.



Here, we report the successful expression, purification and crystallization of the ecto-enzymatic domain of rat CD39 protein. The recombinant protein consists of a fusion of the murine CD4 signal peptide with amino-acid residues 39–470 of rat CD39 and a C-terminal His_6 tag. Since the protein is heavily and heterogeneously glycosylated, with seven putative N-linked glycosylation sites (Wang *et al.*, 1998; Zhong *et al.*, 2001, 2005), we expressed the protein in a glycosylation-deficient CHO cell line Lec3.2.8.1 that is defective in four steps of Golgi sugar modification (Stanley, 1989). A homogenous recombinant ecto-domain of CD39 with high-mannose glycans such as $Man_5GlcNAc_2$ was produced and purified. Diffraction-quality crystals of this protein were subsequently obtained.

2. Cloning, expression and purification

To express the His_6-tagged soluble ecto-domain of CD39 (sCD39), a polymerase chain reaction (PCR) product was generated with primer XZ235# (5′-ACGCGTCGACGCCGCCACCATGTGCCGAGCCATCTCT-3′, containing a *Sal*I site, a Kozak sequence and the sense sequence of the first six residues of the CD4 signal peptide), primer WZ136# (5′-CCCCCCGGGTTAGTGGTGGTGGTGGTGGTGGGATAATGGCTGTTCAGCTGGGAT-3′, containing an *Xma*I site, an antisense sequence of six histidines and CD39 residues 463–470) and the pcDNA-CD4-sCD39 template (Wang *et al.*, 1998). The resulting PCR product was digested with *Sal*I and *Xma*I and ligated into a glutamine synthase expression vector pSMEG (pWZ1018).

To generate a stable expression cell line for sCD39, Lec3.2.8.1 CHO mutant cells were grown in a humidified incubator with 5% CO_2 at 310 K in alpha medium containing 10% fetal bovine serum (FBS). 24 h before transfection, 1.5×10^6 cells were plated on a 10 cm plate. 15 µg pWZ1018 and 45 µg GeneJuice (Novagen) were mixed in 3.2 ml OPTI-MEM medium (Invitrogen) and incubated at room temperature for 5 min. The mixture was then added to the culture. The medium was replaced with fresh medium after 8 h. 2 d after transfection, the cells were trypsinized, seeded into 10 cm plates at 2×10^6 cells per plate and cultured with Dulbecco's modified Eagle's medium (DMEM)/glutamine-free (SAFC Biosciences) supplemented with GS supplement, 10% dialyzed FBS and 10 µ*M* methionine sulfoximine (MSX). After about three weeks, colonies were formed and transferred into 24-well plates. The CM of individual clones was screened with anti-His4 (Qiagen) and anti-CD39 (Wang *et al.*, 1998). One high-expression clone, L-21, was isolated out of 24 colonies screened.

For large-scale preparations of conditioned media for sCD39, the stable Lec3.2.8.1 line L-21 was grown in DMED/glutamine-free with GS supplement, 10% dialyzed FBS, 10 µ*M* MSX in roller bottles until confluence. After one wash with phosphate-buffered saline, the cells were maintained in the same medium but without FBS and the conditioned media were collected every 3 d for three harvests. Subsequently, the conditioned media were filtered and stored at 193 K.

To purify sCD39, 1 l conditioned media was incubated with 20 ml Ni–NTA beads (Novex) for 2 h at 277 K. The beads were packed onto a column, washed with 50 m*M* Tris pH 8.5, 200 m*M* NaCl, 20 m*M* imidazole pH 8.5 and sCD39 protein was eluted with 50 m*M* Tris pH 8.5, 200 m*M* NaCl, 200 m*M* imidazole pH 8.5. The eluted protein was further purified using size-exclusion chromatography in 50 m*M* Tris pH 8.0, 150 m*M* NaCl; sCD39 protein was purified as a monomer. The protein was homogeneous (Fig. 1), as confirmed by SDS–PAGE and mass spectrometry. Typically, the protein yield of sCD39 was around 1 mg l^{-1}.

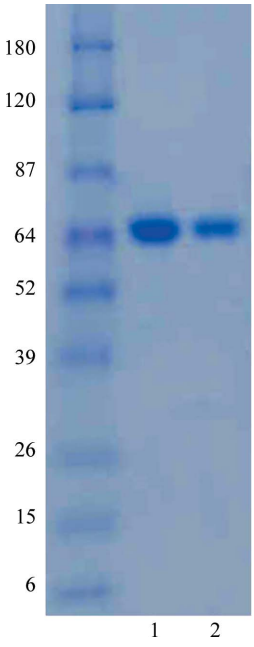


Figure 1
SDS–PAGE analysis of purified sCD39. Lane 1, 2 µg protein; lane 2, 1 µg protein. The corresponding molecular weights (in kDa) for prestained standards (Invitrogen) are shown on the left of the figure.

3. Crystallization, data collection, twinning analysis and molecular replacement

Purified sCD39 protein was concentrated using a Centriprep-10 (Amicon) to 10 mg ml^{-1} in the same buffer, based on the protein absorbance at 280 nm. Crystallization experiments were carried out using the hanging-drop vapor-diffusion method. Equal volumes of protein and reservoir solution (0.2 + 0.2 µl) were mixed and equilibrated against 500 µl reservoir solution. Using the PEG/Ion crystallization screen (Hampton Research) on purified sCD39 led to a number of crystallization hits. The best crystals were grown at 291 K using 20% PEG 3350 and 0.2 *M* ammonium dihydrogen phosphate as precipitating agents. Plate-like crystals formed after 1–2 d and grew to dimensions of 0.2 × 0.2 × 0.1 mm in 3 d (Fig. 2). SDS–PAGE of a crystal that was well washed with the mother liquor confirmed that the molecular weight of the crystallized protein matched that of the intact sCD39 in solution.

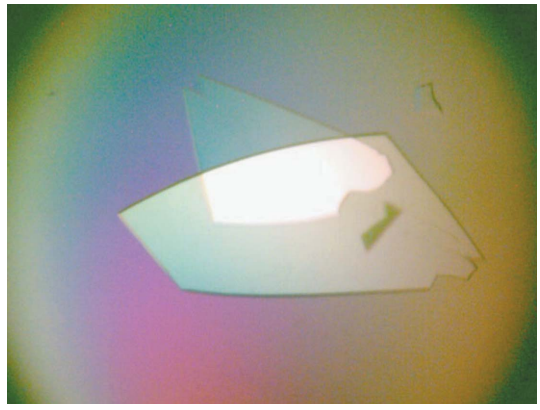


Figure 2
Crystals of sCD39 in polarized light.

Table 1
Data-collection, processing and scaling statistics and twinning analysis.

Space group	$P3_2$
Unit-cell parameters (Å)	$a = b = 118.05$, $c = 81.56$
Resolution range (Å)	20–3.2 (3.31–3.2)
No. of observations	30664
No. of unique reflections	17106
Completeness (%)	92.0 (67.5)
$R_{\rm merge}$†	0.173 (0.334)
Mean $I/\sigma(I)$	5.1 (2.4)
Average redundancy	2 (1.3)
Molecules per ASU	2
Matthews coefficient $V_{\rm M}$ (Å^3 Da^{-1})	2.73
Solvent content (%)	55
Estimated twinning fraction (%)	41.5
Mean $\|L\|$‡	0.36
Mean $\|L^2\|$‡	0.19

† $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where the outer summation is over all unique reflections with multiple observations and the inner summation is over all observations of each reflection. ‡ Mean $|L| = 0.5$ and mean $|L^2| = 0.333$ for untwinned data; mean $|L| = 0.375$ and mean $|L^2| = 0.2$ for a perfect twin.

The crystals were flash-cooled (100 K) prior to data collection in cryoprotection solution consisting of 40% PEG 3350, 0.2 *M* ammonium dihydrogen phosphate, 50 m*M* Tris pH 8.5, 0.2 *M* NaCl. A full 3.2 Å data set was collected in-house from a single crystal using a Saturn92 CCD mounted on an FR-E Cu *Kα* rotating-anode source (Rigaku, Japan). The data were merged and processed with *HKL*-2000 (Otwinowski & Minor, 1997). The crystals belonged to the primitive trigonal space group $P3_1$ or its enantiomorph $P3_2$, with unit-cell parameters $a = b = 118.05$, $c = 81.56$ Å. The presence of two molecules in the asymmetric unit gives a crystal volume per protein weight ($V_{\rm M}$) of 2.73 Å^3 Da^{-1}, with a corresponding solvent content of 55%.

Analysis of the intensity statistics using *phenix.xtriage* (Adams *et al.*, 2002) suggested merohedral twinning with three possible twin laws: $-h, -k, l$; $h, -h-k, -l$; $-k, -h, -l$. Based on analysis of the $R_{\rm obs}$ value, the twin law $-h, -k, l$ was the most likely, with the lowest $R_{\rm obs}$ of 17.8%. Owing to the high degree of twinning (the estimated twin fraction was 41.5%) and to avoid significant errors in the output amplitudes, we decided to use the twinned data directly rather than attempting detwinning. The data-collection, processing, scaling statistics and twinning analysis are given in Table 1.

As in most cases, the presence of twinning did not impede structure solution by molecular replacement. Sequence comparison with structurally characterized members of the NTPDase family revealed that CD39 shares ∼45% identity with NTPDase2 (PDB code 3cj1). The structure of this enzyme was therefore used to construct a search model. The correct solution was found by maximum-likelihood molecular replacement as implemented in the program *Phaser* (McCoy *et al.*, 2005). The correct rotation-function and translation-function peaks were clearly marked by the highest *Z* scores (RFZ = 6.1, TFZ = 9.7) and the highest log-likelihood gain (LLG = 166), producing a solution in space group $P3_2$ with two molecules in the asymmetric unit. This model was then submitted to rigid-body refinement in *PHENIX* (Adams *et al.*, 2002), yielding an $R_{\rm cryst}$ of 45.9% ($R_{\rm free}$ of 46.1%) and after a few rounds of positional minimization an $R_{\rm cryst}$ of 40.4% ($R_{\rm free}$ of 44.4%).

Once the molecular-replacement solution was available, we decided to use a model-based twinning analysis *via* the $R_{\rm obs}$ *versus* $R_{\rm calc}$ plot, which utilizes both observed intensities and intensities derived from the atomic model (Lebedev *et al.*, 2006). The obtained distribution of $R_{\rm obs}$ against $R_{\rm calc}$ (calculated with *phenix.xtriage*) gave a clear indication of the perfect merohedral twin: $R_{\rm obs} = 19\%$ and $R_{\rm calc} = 48.6\%$.

At present, new crystallization trials are being carried out to improve the crystal form, quality and size. Details of structure determination and refinement will be reported elsewhere.

References

Adams, P. D., Grosse-Kunstleve, R. W., Hung, L.-W., Ioerger, T. R., McCoy, A. J., Moriarty, N. W., Read, R. J., Sacchettini, J. C., Sauter, N. K. & Terwilliger, T. C. (2002). *Acta Cryst.* D**58**, 1948–1954.

Bigonnesse, F., Levesque, S. A., Kukulski, F., Lecka, J., Robson, S. C., Fernandes, M. J. & Sevigny, J. (2004). *Biochemistry*, **43**, 5511–5519.

Chadwick, B. P. & Frischauf, A. M. (1998). *Genomics*, **50**, 357–367.

Kaczmarek, E., Koziak, K., Sevigny, J., Siegel, J. B., Anrather, J., Beaudoin, A. R., Bach, F. H. & Robson, S. C. (1996). *J. Biol. Chem.* **271**, 33116–33122.

Lebedev, A. A., Vagin, A. A. & Murshudov, G. N. (2006). *Acta Cryst.* D**62**, 83–95.

Maliszewski, C. R., Delespesse, G. J., Schoenborn, M. A., Armitage, R. J., Fanslow, W. C., Nakajima, T., Baker, E., Sutherland, G. R., Poindexter, K. & Birks, C. (1994). *J. Immunol.* **153**, 3574–3583.

McCoy, A. J., Grosse-Kunstleve, R. W., Storoni, L. C. & Read, R. J. (2005). *Acta Cryst.* D**61**, 458–464.

Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.

Robson, S. C., Sevigny, J. & Zimmermann, H. (2006). *Purinergic Signal.* **2**, 409–430.

Stanley, P. (1989). *Mol. Cell. Biol.* **9**, 377–383.

Wang, T. F. & Guidotti, G. (1996). *J. Biol. Chem.* **271**, 9898–9901.

Wang, T. F., Ou, Y. & Guidotti, G. (1998). *J. Biol. Chem.* **273**, 24814–24821.

Zhong, X., Kriz, R., Kumar, R. & Guidotti, G. (2005). *Biochim. Biophys. Acta*, **1723**, 143–150.

Zhong, X., Malhotra, R., Woodruff, R. & Guidotti, G. (2001). *J. Biol. Chem.* **276**, 41518–41525.