Karen Maree Knapp,^{a,b} Matthias
Zebisch^{a,c} and Norbert Sträter^{a*}

^aInstitute of Bioanalytical Chemistry, Center for
Biotechnology and Biomedicine, University of
Leipzig, Deutscher Platz 5, 04103 Leipzig,
Germany, ^bDepartment of Biochemistry,
University of Otago, 710 Cumberland Street,
Dunedin 9016, New Zealand, and ^cDivision of
Structural Biology, University of Oxford,
Roosevelt Drive, Oxford OX3 7BN, England

Correspondence e-mail:
strater@bbz.uni-leipzig.de

Received 25 September 2012

Accepted 2 November 2012

Crystallization and preliminary X-ray analysis of the open form of human ecto-5'-nucleotidase (CD73)

Eukaryotic ecto-5'-nucleotidase (e5NT) catalyses the hydrolysis of extracellular AMP to adenosine and plays a pivotal role in switching on adenosine signalling via the P1 receptors of the purinergic signalling pathway. With such an important regulatory role, e5NT has become an appealing new drug target, with potential applications in the treatment of inflammation, chronic pain, hypoxia and cancer. In order to gain insight into the structure and function of the eukaryotic e5NT enzymes and to assist in structure-based drug design, the crystal structure of human e5NT has been solved. Recombinant human e5NT comprising four asparagine-to-aspartate surface mutations targeting potential glycosylation sites was refolded from bacterial inclusion bodies. Refolded and purified human e5NT crystallized in space group $P4_332$ and a data set to 1.85 Å resolution was obtained. The structure could be solved by molecular replacement using a polyalanine model generated from *Thermus thermophilus* 5'-nucleotidase (5NT). An anomalous data set revealed the presence of a metal-ion binding site, as well as calcium and chloride ion-binding sites. Structural comparisons with bacterial 5NT homologues showed that the human e5NT crystal structure has an open conformation in which the metal- and substrate-binding sites are distant from each other. Here, the crystallization and preliminary X-ray crystallographic analysis of an open structural conformation of human e5NT are described.

1. Introduction

Ecto-5'-nucleotidase (e5NT; also known as CD73) is a dimeric extracellular glycoprotein that is tethered to the cell membrane by a glycosylphosphatidylinositol (GPI) anchor. Eukaryotic e5NT catalyses the hydrolysis of extracellular AMP to adenosine and regulates signalling by the adenosine-activated P1 receptors. Owing to its pivotal role in purinergic signalling, e5NT has emerged as an appealing new drug target for the treatment of inflammation, chronic pain, hypoxia and cancer (Ohtsuka *et al.*, 2010; Sowa *et al.*, 2010; Thompson *et al.*, 2004; Li *et al.*, 2006; Stagg & Smyth, 2010; Salmi & Jalkanen, 2012; Stagg, 2012; Zhang, 2012). While AMP is thought to be the major physiological substrate, e5NT also catalyses the hydrolysis of a broad range of 5'-nucleoside monophosphates (Sträter, 2006; Zimmermann *et al.*, 2012). Interestingly, ADP and ATP are competitive inhibitors of e5NT.

Based on sequence similarity, eukaryotic e5NT is structurally related to the bacterial 5'-nucleotidase (5NT) enzymes and belongs to a large superfamily of distantly related metallo-phosphoesterases (Koonin, 1994; Knöfel & Sträter, 1999). e5NT is not structurally related to known eukaryotic cytosolic 5NT enzymes (Walldén *et al.*, 2007). Structures of monomeric bacterial 5NT enzymes have revealed that the enzyme comprises two distinct domains, the N- and C-terminal domains, connected by an α -helix. Together, the metal ion-binding site of the N-terminal domain and the substrate-binding site of the C-terminal domain form the active site. The N- and C-terminal domains undergo extensive domain rotations relative to each other and thereby switch between open and closed structural conformations (Garavaglia *et al.*, 2012; Knöfel & Sträter, 2001; Schultz-Heienbrok *et al.*, 2004). Studies on *Escherichia coli* 5NT have

© 2012 International Union of Crystallography
All rights reserved

demonstrated that an $\sim 96^\circ$ domain rotation is necessary for enzyme catalysis, most likely to enable substrate binding and product release from the open form (Knöfel & Sträter, 2001; Schultz-Heienbrok *et al.*, 2004, 2005).

The core catalytic residues and secondary-structure elements of bacterial 5NT enzymes are largely conserved in the eukaryotic e5NT enzymes, and for this reason these enzymes are thought to have similar active-site structures. Interestingly, while e5NT hydrolyses AMP but not ADP or ATP, bacterial 5NT enzymes hydrolyse AMP, ADP and also ATP. Based on structural modelling of the eukaryotic e5NT active site, it is not obvious how the more stringent substrate selectivity of the eukaryotic e5NT enzymes towards AMP (*versus* ADP and ATP) is achieved (Sträter, 2006). Possibly, control of the domain rotation, and thus the distance between the substrate-binding site of the C-terminal domain and the metal ion-binding site of the N-terminal domain, influences activity towards adenine nucleotides that differ in the distance between the adenine moiety and the terminal phosphate group (Sträter, 2006).

A crystal structure of a closed form of human e5NT has recently been obtained for the glycosylated enzyme expressed in HEK293 cells (Heuts *et al.*, 2012). All attempts to obtain a structure with bound ligands or in the open form have so far failed. Here, we describe the crystallization and preliminary X-ray crystallographic analysis of human e5NT in a deglycosylated form prepared *via* bacterial expression in inclusion bodies and refolding. Initial molecular-replacement trials indicate that the enzyme is present in an open conformation. These crystals will help to characterize the domain motion of human e5NT as well as substrate binding to the open form.

2. Materials and methods

2.1. Protein purification

Recombinant human e5NT was overexpressed in *E. coli* in the form of inclusion bodies and was refolded using the method of rapid dilution protein refolding (see below). The recombinant human e5NT protein used in this study consists of the mature human e5NT protein sequence (residues 27–549) and corresponds to the previously described T376A variant (UniProtKB/Swiss-Prot entry P21589/VAR_022091). The signal sequence of human e5NT (residues 1–26) responsible for protein secretion was replaced with an N-terminal purification tag comprising a His tag and an enteropeptidase cleavage sequence (additional N-terminal residues: MAHHHHHHVGTGSN-DDDDKSPDP) to assist in purification. Four asparagine-to-aspartate mutations (N53D, N311D, N333D and N403D) at glycosylation sites were introduced into the e5NT sequence to increase both protein solubility and refolding yields, and two further surface mutations (K478A and R480A) were introduced to enhance the crystallization properties. Details of the procedures developed to produce active human e5NT from bacterial inclusion bodies will be described elsewhere. Briefly, human e5NT was refolded by rapid dilution refolding using a refolding buffer consisting of 0.1 M Tris–HCl pH 8.0, 0.5 M L-arginine, 10% glycerol, 2 mM reduced glutathione, 1 mM oxidized glutathione, 5 mM CaCl₂, 30 μ M ZnCl₂. Renatured human e5NT protein was purified from misfolded species using AMP-agarose affinity chromatography followed by size-exclusion chromatography. The N-terminal purification tag was not removed during protein purification. The enzyme activity has been characterized using a malachite green assay (Zebisch & Sträter, 2007). A specific activity of 577 U mg^{−1} was determined for AMP hydrolysis by the renatured and purified human e5NT protein used in this study, which is higher than the specific activity of 130 U mg^{−1} determined for human e5NT

purified from a human colon adenocarcinoma cell line (Navarro *et al.*, 1998). Also, the K_m value of $6.6 \pm 0.2 \mu$ M determined for AMP hydrolysis is comparable to or lower than the K_m values of 1–50 μ M described for e5NT enzymes from different sources (Zimmermann, 1992).

2.2. Crystallization

Crystallization trials for human e5NT used freshly prepared and filtered (0.22 μ m) protein at a protein concentration of 5.0 mg ml^{−1} in 10 mM Tris–HCl pH 8.0. Immediately prior to crystallization, AMPCP {adenosine [(α,β)-methylene]diphosphate; Jena Bioscience, Germany} was added to give a final concentration of 1 mM in the drop. Initial crystallization screening was carried out using sparse-matrix crystallization screens (Hampton Research, Aliso Viejo, California, USA) and the sitting-drop vapour-diffusion method. Sitting-drop experiments were carried out in 96-well crystallization plates (Greiner Bio-One, Kremsmünster, Austria) and the crystallization drops were dispensed using a Microsys 4000 XL dispensing robot (Cartesian Technologies Inc., Irvine, California, USA). Crystallization drops had a typical volume of 0.4 μ l (0.2 μ l reservoir solution plus 0.2 μ l protein solution) and were equilibrated against 100 μ l reservoir solution. The crystallization plates were sealed and the plates were incubated at 292 K. Once crystallization hits had been identified by sparse-matrix screening, crystallization optimization was carried out using 12-well PVC (polyvinyl chloride) trays (Nelipak, Venray, Netherlands) and the hanging-drop vapour-diffusion method. Systematic grid screens were designed using crystallization drops with a typical volume of 2 μ l (1 μ l reservoir solution plus 1 μ l protein solution); they were equilibrated against 500 μ l reservoir solution and the crystallization plates were incubated at 292 K. The crystals used to collect a high-resolution native data set were grown in 1.0 M lithium chloride, 0.1 M HEPES pH 7.0. These crystals were transferred to a 3.0 M sodium malonate, 0.1 M Tris–HCl pH 8.0 cryoprotectant solution for approximately 10–15 s and were then flash-cooled in liquid nitrogen. The crystals used to collect a lower resolution anomalous data set were grown in 3.5 M ammonium chloride, 0.1 M Tris–HCl pH 8.5. These crystals were transferred to a 20% glycerol, 3.5 M ammonium chloride, 0.1 M Tris–HCl pH 8.5 cryoprotectant solution for approximately 10–15 s and were then flash-cooled in liquid nitrogen.

Crystallization trials were frequently monitored using a Leica MZ16 binocular microscope (Leica Microsystems GmbH, Wetzlar, Germany) and photographs of the crystals were taken with a Canon S40 digital camera (Canon, Amsterdam, Netherlands) mounted to the microscope. Protein crystals were distinguished from salt crystals by UV fluorescence using a DUVI 204 instrument (RiNA GmbH, Berlin, Germany).

2.3. X-ray data collection

X-ray data collection was carried out at 100 K using a wavelength of either 0.91841 Å (native data set) or 1.9 Å (anomalous data set) on beamline 14.2 of the Berlin Synchrotron (BESSY, Berlin, Germany) equipped with a Rayonix MX-225 CCD detector (MAR Research GmbH, Norderstedt, Germany). After collecting two initial diffraction images separated by 90°, the reflections were indexed with *iMOSFLM* (Battye *et al.*, 2011), the mosaicity was estimated and a suitable data-collection strategy was decided upon which would result in the collection of a complete data set.

The native data set was generated by merging two partial data sets collected from a single crystal. The first partial data set was collected with an oscillation angle of 0.5°. In order to improve the separation of

reflections, the oscillation angle was decreased to 0.2° for the second part of the data set. Furthermore, the crystal-to-detector distance was set to 230 mm to provide sufficient separation of reflections. As a result, the resolution is limited to 1.85 Å. The $I/\sigma(I)$ values indicated that the crystal diffracted to somewhat higher resolution.

The X-ray data sets were indexed, integrated, scaled and converted to structure-factor amplitudes using *iMOSFLM* (Battye *et al.*, 2011), *SCALA* and *TRUNCATE* (French & Wilson, 1978) from the *CCP4* suite (Winn *et al.*, 2011). Further data analysis, including the assessment of systematic absences, assignment of space group and calculation of the solvent content, was conducted using software within the *CCP4* suite (Winn *et al.*, 2011; Matthews, 1968). Anomalous data were processed with *SHELX* (Sheldrick, 2008) and anomalous maps were analysed in *Coot* (Emsley *et al.*, 2010). X-ray data-collection parameters, integration and data-reduction statistics are summarized in Table 1.

Molecular replacement was carried out using *MOLREP* (Vagin & Teplyakov, 2010); a polyaniline structural model generated from *Thermus thermophilus* 5NT (PDB entry 2z1a; RIKEN Structural Genomics/Proteomics Initiative, unpublished work), which shares 37% amino-acid sequence identity with human e5NT, was used as an initial search model. Molecular replacement was achieved by searching first for the larger N-terminal domain (residues 28–331) and then for the smaller C-terminal domain (residues 332–531) of the search model. Model building was carried out using the automated model-building program *ARP/wARP* (Langer *et al.*, 2008), and the model was improved by iterative manual rebuilding with *Coot* (Emsley *et al.*, 2010) and refinement with *REFMAC5* (Murshudov *et al.*, 2011).

3. Results and discussion

Following crystallization optimization, cube-shaped crystals of human e5NT were obtained in either 1.0 M lithium chloride, 0.1 M HEPES pH 7.0 or 3.5 M ammonium chloride, 0.1 M Tris–HCl pH 8.5. Crystals grew to dimensions of approximately $100 \times 100 \times 100 \mu\text{m}$ within 1–2 d (Fig. 1).

A native data set was collected from these crystals to 1.85 Å resolution. The X-ray data were indexed in space group *P432*, and scaling of the intensities showed systematic absences along the *h*00 axis (reflection condition $h = 4n$), indicating a screw axis resulting in either space group *P4₁32* or *P4₃32*. The crystals have unit-cell parameters

Table 1

Data-collection statistics for human e5NT.

Values in parentheses are for the highest resolution shell.

	Native data	Anomalous data
X-ray source	BESSY BL14.2	BESSY BL14.2
No. of crystals	1	1
Wavelength (Å)	0.918	1.900
Temperature (K)	100	100
Crystal-to-detector distance (mm)	230	140
Exposure per frame (s)	10	10
Oscillation per frame ($^\circ$)	0.2–0.5	0.5
Total oscillation range ($^\circ$)	85.8	181.0
Space group	<i>P4₃32</i>	<i>P4₃32</i>
Unit-cell parameters (Å)	$a = b = c = 167.6$	$a = b = c = 167.9$
Resolution range (Å)	41.89–1.85 (1.95–1.85)	50.64–2.80 (2.95–2.80)
Total No. of observations	1286375 (104797)	863715 (123100)
Unique reflections	68775 (9864)	20501 (2906)
Completeness (%)	100.0 (100.0)	100.0 (100.0)
Anomalous completeness (%)	100.0 (100.0)	100.0 (100.0)
Mean $I/\sigma(I)$	20.6 (4.1)	24.7 (6.5)
Multiplicity	18.7 (10.6)	42.1 (42.4)
Anomalous multiplicity	9.7 (5.4)	22.8 (22.3)
$R_{\text{merge}}^\dagger$ (%)	10.0 (57.4)	14.8 (88.9)
$R_{\text{p.i.m.}}^\ddagger$ (%)	2.3 (18.3)	3.2 (19.2)
$R_{\text{anom}}^\S$ (%)	N/D ¶	2.1 (10.9)
Wilson <i>B</i> factor (Å ²)	24.0	75.8
Molecules per asymmetric unit	1	1
Matthews coefficient (Å ³ Da ^{−1})	3.2	3.3
Solvent content (%)	62	62

$$\dagger R_{\text{merge}} = \frac{\sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle|}{\sum_{hkl} \sum_i I_i(hkl)}, \quad \ddagger R_{\text{p.i.m.}} = \frac{\sum_{hkl} \{1/[N(hkl) - 1]\}^{1/2} \sum_i |I_i(hkl) - \langle I(hkl) \rangle|}{\sum_{hkl} \sum_i I_i(hkl)}, \quad \S R_{\text{anom}} = \frac{\sum_i |I^+(h) - \langle I^+ \rangle|}{\sum_i (I^+(h) + \langle I^+ \rangle)}, \quad \P \text{ Not determined.}$$

parameters $a = b = c = 167.5 \text{ Å}$; assuming the presence of one molecule per asymmetric unit, the solvent content of the crystals was calculated to be 62% (Matthews, 1968) (Table 1).

Following scaling of the diffraction data and conversion to structure factors, attempts were made to solve the phase problem by molecular replacement. A polyaniline structural model generated from *T. thermophilus* 5NT (PDB entry 2z1a; RIKEN Structural Genomics/Proteomics Initiative, unpublished work), which shares 37% amino-acid identity with human e5NT, was used as the search model. Molecular replacement was achieved with a 2.5 Å resolution data set (which was available before the better data set reported here was obtained) by searching first for the larger N-terminal domain (residues 28–331) and then for the smaller C-terminal domain (residues 332–531) of the polyaniline search model. All possible space groups of the *P432* crystal class were tested and the correct

Figure 1

A crystal of human e5NT grown in 1.0 M lithium chloride, 0.1 M HEPES pH 7.0. Crystals grew to approximately $100 \times 100 \times 100 \mu\text{m}$ in size (a) and showed strong UV fluorescence (b).

Figure 2

Strong anomalous signals observed in the human e5NT crystal structure. (a) A single metal ion bound in the metal ion-binding site. (b) A calcium ion and (c) a chloride ion bound to human e5NT. The $2F_o - F_c$ maps from the partially refined native data set are shown in blue and are contoured at 2.0σ , while the anomalous difference maps are shown in yellow and are contoured at 5.0σ . Figures were prepared in Coot (Emsley *et al.*, 2010).

molecular-replacement solution resulting from space group $P4_32$ was indicated by best scores of 0.223 (the second best score was 0.170) and 0.299 (the next score was 0.233) in the translation function. After partial refinement, the R values dropped to below 30%.

The relative orientations of the N- and C-terminal domains in the correct molecular-replacement solution were found to be significantly different from that observed in the *T. thermophilus* 5NT crystal structure. The *T. thermophilus* 5NT structure represents a closed conformation in which the metal ion-binding site of the N-terminal domain and the substrate-binding site of the C-terminal domain come together to form the active site. In the human e5NT structure the N- and C-terminal domains are oriented such that the metal ion- and substrate-binding sites are very distant from one another; thus, the structure represents an open inactive conformation (data not shown). An extensive $\sim 96^\circ$ domain motion has previously been described for *E. coli* 5NT, for which structures are available in open, closed and intermediate conformations. The domain motion appears to be necessary for enzyme catalysis, most likely to enable substrate binding and product release from the open form (Knöfel & Sträter, 2001; Schultz-Heienbrok *et al.*, 2004, 2005).

It was also possible to collect an anomalous data set to $2.80 \text{ \AA}$ resolution from these crystals (Table 1). The anomalous difference maps phased with an almost complete model showed several strong anomalous signals (Fig. 2).

The strongest anomalous signal (9.5σ) observed was located within the metal ion-binding site of the N-terminal domain and represents a single metal ion bound to the dimetal centre (Fig. 2a). Another strong anomalous signal (6.8σ) was observed and appears to represent a calcium ion from bond distances and coordination with surrounding water molecules and the negative side chain of Asp237 (Fig. 2b). The presence of a calcium ion is not surprising given the high concentration of calcium chloride in the refolding buffer. Furthermore, a strong anomalous signal (6.5σ) was observed and appears to represent a chloride ion coordinated to the backbone N atom of Glu224 (Fig. 2c). Additional weak anomalous signals representing further partially occupied chloride sites were also observed in the anomalous data set. The presence of chloride ions is not surprising given the high concentration of ammonium chloride in the crystallization condition. Anomalous density maps also verified the location of disulfide bridges. In the native data set, strong electron-density signals were also observed for the metal ion, calcium and the chloride ion mentioned above (Fig. 2).

The new crystal form of human e5NT described in this study will allow analysis of the open conformation of the enzyme and of the domain motion between the open and closed forms. It will also enable the structure-based design of inhibitors targeting the open form, including allosteric inhibitors that prevent a closure motion of the enzyme.

The authors thank the Joint Berlin MX Laboratory at the Helmholtz Zentrum Berlin (BESSY II) for beam time and assistance during synchrotron data collection, as well as for travel support. The Deutsche Forschungsgemeinschaft (DFG) is acknowledged for financial support. The authors declare no conflict of interest.

References

- Battye, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. W. (2011). *Acta Cryst.* **D67**, 271–281.
- Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* **D66**, 486–501.
- French, S. & Wilson, K. (1978). *Acta Cryst.* **A34**, 517–525.
- Garavaglia, S., Bruzzzone, S., Cassani, C., Canella, L., Allegrone, G., Sturla, L., Mannino, E., Millo, E., De Flora, A. & Rizzi, M. (2012). *Biochem. J.* **441**, 131–141.
- Heuts, D. P., Weissenborn, M. J., Olkhov, R. V., Shaw, A. M., Levy, C. & Scrutton, N. S. (2012). *ChemBiochem*, **13**, 2384–2391.
- Knöfel, T. & Sträter, N. (1999). *Nature Struct. Biol.* **6**, 448–453.
- Knöfel, T. & Sträter, N. (2001). *J. Mol. Biol.* **309**, 255–266.
- Koonin, E. V. (1994). *Protein Sci.* **3**, 356–358.
- Langer, G., Cohen, S. X., Lamzin, V. S. & Perrakis, A. (2008). *Nature Protoc.* **3**, 1171–1179.
- Li, X., Zhou, T., Zhi, X., Zhao, F., Yin, L. & Zhou, P. (2006). *Microvasc. Res.* **72**, 48–53.
- Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
- Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* **D67**, 355–367.
- Navarro, J. M., Olmo, N., Turnay, J., López-Conejo, M. T. & Lizarbe, M. A. (1998). *Mol. Cell. Biochem.* **187**, 121–131.
- Ohtsuka, T., Changelian, P. S., Boui, D., Noon, K., Harada, H., Lama, V. N. & Pinsky, D. J. (2010). *J. Immunol.* **185**, 1321–1329.
- Salmi, M. & Jalkanen, S. (2012). *Oncimmunology*, **1**, 247–248.
- Schultz-Heienbrok, R., Maier, T. & Sträter, N. (2004). *Protein Sci.* **13**, 1811–1822.
- Schultz-Heienbrok, R., Maier, T. & Sträter, N. (2005). *Biochemistry*, **44**, 2244–2252.
- Sheldrick, G. M. (2008). *Acta Cryst.* **A64**, 112–122.
- Sowa, N. A., Voss, M. K. & Zylka, M. J. (2010). *Mol. Pain*, **6**, 20.
- Stagg, J. (2012). *Oncimmunology*, **1**, 217–218.
- Stagg, J. & Smyth, M. J. (2010). *Oncogene*, **29**, 5346–5358.

- Sträter, N. (2006). *Purinerg. Signal.* **2**, 343–350.
- Thompson, L. F., Eltzschig, H. K., Ibla, J. C., Van De Wiele, C. J., Resta, R., Morote-Garcia, J. C. & Colgan, S. P. (2004). *J. Exp. Med.* **200**, 1395–1405.
- Vagin, A. & Teplyakov, A. (2010). *Acta Cryst. D* **66**, 22–25.
- Walldén, K., Stenmark, P., Nyman, T., Flodin, S., Gräslund, S., Loppnau, P., Bianchi, V. & Nordlund, P. (2007). *J. Biol. Chem.* **282**, 17828–17836.
- Winn, M. D. *et al.* (2011). *Acta Cryst. D* **67**, 235–242.
- Zebisch, M. & Sträter, N. (2007). *Biochemistry*, **46**, 11945–11956.
- Zhang, B. (2012). *Oncoimmunology*, **1**, 67–70.
- Zimmermann, H. (1992). *Biochem. J.* **285**, 345–365.
- Zimmermann, H., Zebisch, M. & Sträter, N. (2012). *Purinerg. Signal.* **8**, 437–502.