Inmaculada Gómez-García,^{a,‡}
 Marchel Stuver,^{b,‡} June Ereño,^a
 Iker Oyenarte,^a María Angeles
 Corral-Rodríguez,^a Dominik
 Müller^b and Luis Alfonso
 Martínez-Cruz^{a*}

^aStructural Biology Unit, CIC bioGUNE,
 Parque Tecnológico de Bizkaia, Edificio 800,
 48160 Derio, Bizkaia, Spain, and

^bDepartment of Pediatric Nephrology, Charité
 Universitätsmedizin, 13353 Berlin, Germany

‡ These authors contributed equally to this
 work.

Correspondence e-mail:
 amartinez@cicbiogune.es

Received 29 May 2012
 Accepted 9 August 2012

Purification, crystallization and preliminary crystallographic analysis of the CBS-domain pair of cyclin M2 (CNNM2)

This work describes the purification and preliminary crystallographic analysis of the CBS-domain pair of the murine CNNM2 magnesium transporter (formerly known as ancient domain protein 2; ACDP2), which consists of a pair of cystathionine β -synthase (CBS) motifs and has 100% sequence identity to its human homologue. CNNM proteins represent the least-studied members of the eight different types of magnesium transporters identified to date in mammals. In humans, the CNNM family is encoded by four genes: *CNNM1–4*. CNNM1 acts as a cytosolic copper chaperone, whereas CNNM2 and CNNM4 have been associated with magnesium handling. Interestingly, mutations in the *CNNM2* gene cause familial dominant hypomagnesaemia (MIM:607803), a rare human disorder characterized by renal and intestinal magnesium (Mg^{2+}) wasting, which may lead to symptoms of Mg^{2+} depletion such as tetany, seizures and cardiac arrhythmias. This manuscript describes the preliminary crystallographic analysis of two different crystal habits of a truncated form of the protein containing its regulatory CBS-domain pair, which has been reported to host the pathological mutation T568I in humans. The crystals belonged to space groups $P2_12_12$ and $I222$ (or $I2_12_12$) and diffracted X-rays to 2.0 and 3.6 Å resolution, respectively, using synchrotron radiation.

1. Introduction

Magnesium, the second most abundant intracellular cation within the human body, is an essential element that is required for the catalytic activity of numerous metalloenzymes. It takes part in oxidative phosphorylation and biochemical reactions involving nucleotides, it modulates muscular function, membrane potential and nerve excitability and can also serve a purely structural role by stabilizing the conformations of certain metal-dependent protein domains and transcriptional regulators (Quamme, 2010). Deficiency of this essential metal has been implicated in many diseases ranging from hypertension, cardiac arrhythmias or asthma to migraines and bone disorders (Agus, 1999; Agus & Agus, 2001; Quamme, 2010). Taking into account the notable physical properties of the Mg^{2+} cation [the unhydrated Mg^{2+} ion has the smallest diameter (0.65 Å) of all biological ions, whereas the fully hydrated cation (5.0 Å) has the largest], together with the particularly restricted octahedral coordination geometry and the relatively slow exchange rate of water in the hydration shell of Mg^{2+} , there are some challenges in explaining how Mg^{2+} moves across cellular membranes. The intracellular Mg^{2+} level is maintained well below the concentration predicted from the transmembrane electrochemical potential, being controlled through regulation of Mg^{2+} uptake, intracellular storage and cellular efflux and ion transport in and out of a variety of organelle compartments by the action of multidomain proteins known as magnesium transporters, which have just begun to be identified at the molecular level (Günther, 1993; Quamme, 1997, 2010; Cefaratti *et al.*, 2000; Romani & Scarpa, 2000; Tashiro *et al.*, 2000; Watanabe *et al.*, 2005; Schweigel *et al.*, 2008). Recently, the crystal structures of the bacterial CorA

© 2012 International Union of Crystallography
 All rights reserved

(Eshaghi *et al.*, 2006; Lunin *et al.*, 2006) and MgtE Mg^{2+} transporters have been reported (Hattori *et al.*, 2007).

In mammals, eight different types of magnesium transporters have been identified to date (Quamme, 2010). Of these, the least studied members are the ancient conserved domain proteins (ACDPs), which were first described by Wang *et al.* (2003). Their name is based on the fact that they all share an 'ancient conserved domain' (ACD) present in a large number of species from bacteria to zebrafish to man, suggesting that the ACDP family may be essential proteins (Wang *et al.*, 2003). In humans, the ACDP family includes four members, ACDP1–4, whereas it appears to be a single-copy gene in lower organisms such as *Caenorhabditis elegans*, yeasts and bacteria (Wang *et al.*, 2004). The ACDP proteins are evolutionarily expressed throughout development and adult tissues, except for ACDP1 which is mainly expressed in the brain, and all show very strong homology to the bacterial CorC and yeast Mam3p proteins, which are involved in magnesium and cobalt efflux and in resistance to manganese toxicity, respectively (Wang *et al.*, 2004; Gibson *et al.*, 1991; Yang *et al.*, 2005). The strong similarity between the ACDP1 and Mam3p proteins was first described by Yang *et al.* (2005) and has more recently been remarked on by Stuver *et al.* (2011). Interestingly, human ACDP2, CorC from *Salmonella typhimurium* and Mam3p from *Saccharomyces cerevisiae* also have significant sequence similarity, which is highest in the region comprising the CBS-domain pair of ACDP2 and CorC. Remarkably, the pathogenic mutation T568I in ACDP2 is located in this region (Stuver *et al.*, 2011). The identity between ACDP2 (residues 440–600) and CorC (residues 60–200) is 37.3% (60.7% similarity), whereas the sequence identity between ACDP2 (residues 440–600) and Mam3p (residues 240–390) is 29.5% (45.6% similarity). The presence of a cyclin box-like motif, the predicted helical architecture (typical of cyclins) and their location in the plasma membrane has led to the suggestion that ACDPs might be involved in cell-cycle regulation (Wang *et al.*, 2003). Accordingly, ACDPs are also known as cyclins M1–4 or CNNM1–4, although these proteins do not appear to have cyclin function *in vivo*. Recent findings have revealed that CNNMs reside at the plasma membrane (Stuver *et al.*, 2011; De Baaij *et al.*, 2012). CNNM1 acts as a cytosolic copper chaperone (Alderton *et al.*, 2007), whereas CNNM2, which was initially identified as an Mg^{2+} transporter (Goytain & Quamme, 2005), has recently been shown to mediate Mg^{2+} -sensitive Na^+ currents and is linked to serum Mg^{2+} levels. Interestingly, mutations in the *CNNM2* gene cause familial dominant hypomagnesaemia (MIM:607803), a rare human disorder characterized by renal and intestinal Mg^{2+} wasting, which may lead to symptoms of Mg^{2+} depletion such as tetany, seizures and cardiac arrhythmias (Wang *et al.*, 2003; Guo *et al.*, 2005; Parry *et al.*, 2009; Polok *et al.*, 2009; Quamme, 2010; Stuver *et al.*, 2011; De Baaij *et al.*, 2012). On the other hand, mutations in *CNNM4* are considered to be the cause of Jalili syndrome, which consists of autosomal recessive cone-rod dystrophy and amelogenesis imperfecta (Wang *et al.*, 2003; Guo *et al.*, 2005; Polok *et al.*, 2009; Quamme, 2010).

Structurally, human CNNM2 (UniProtKB/Swiss-Prot accession No. Q9H8M5), which shares 68% similarity (56% identity) with CNNM4, is a multidomain protein that is predicted to exist in three isoforms that differ in sequence and protein length. The deduced full-length protein (isoform 1) and the exon 6-deleted version are conserved between human and mouse, whereas the third isoform is solely based on one single EST BU189723 from a melanotic melanoma. Isoform 1 of CNNM2 includes 875 residues and is formed of (i) an N-terminal domain with five predicted transmembrane helices (residues 48–68, 251–271, 314–334, 339–359 and 369–389), (ii) a DUF21 domain of unknown function (residues 257–431), (iii) a

cystathionine β -synthase (CBS) motif pair (residues 450–511 and 518–584; Bateman, 1997; Wang *et al.*, 2003) that hosts the pathological mutation T568I in humans and might act as sensor of the intracellular Mg^{2+} concentration (Stuver *et al.*, 2011) and (iv) a cyclic nucleotide monophosphate (cNMP)-binding-like domain (residues 649–791) similar to those usually present in ion channels and cNMP-dependent kinases (Shabb & Corbin, 1992) (Fig. 1). Mouse CNNM2 (UniProtKB/Swiss-Prot accession No. Q3TWN3) shares 97.8% identity with its full-length human counterpart (100% identity in the CBS-domain pair region). It was shown to be upregulated in the case of magnesium deficiency and has been characterized as a nonselective divalent-cation transporter with a preference for Mg^{2+} over Co^{2+} , Mn^{2+} , Sr^{2+} , Ba^{2+} , Cu^{2+} and Fe^{2+} (Goytain & Quamme, 2005). Although no structural data for CNNM2 have been reported to date, its high sequence similarity to its closest homologue, CNNM4, suggests that these transporters might have similar regulatory mechanisms. We recently determined the crystal structure of the CBS-domain pair (also referred as the Bateman domain) of human CNNM4 (to be published elsewhere) and found that it associates to form a disc-shaped dimeric species known as a CBS module. More interestingly, we could demonstrate that Mg^{2+} ions interact with the CBS module and induce a conformational change that reduces its thermal stability.

The CBS domains are considered to be energy-sensing modules which bind adenosine ligands (Scott *et al.*, 2004), metal ions (Hattori *et al.*, 2009) or DNA (Aguado-Llera *et al.*, 2010) with different affinities (Baykov *et al.*, 2011). For instance, CBS domains are involved in gating of the osmoregulatory proteins (Biemans-Oldehinkel *et al.*,

Figure 1

Domain distribution in the human CNNM1–4 metal transporters. DUF21 (Pfam code PF01595; residues 224–414, 251–431, 136–294 and 184–358 in CNNM1, CNNM2, CNNM3 and CNNM4, respectively) is a domain of unknown function usually found in the N-terminus of the proteins adjacent to two intracellular CBS domains. In CNNM2, DUF21 is preceded by one transmembrane (TM) helix (TM1, residues 48–68) and contains another four TM helices including residues 251–271 (TM2), 314–334 (TM3), 339–359 (TM4) and 369–389 (TM5). The cystathionine β -synthase (CBS) pair of CNNM2 includes two CBS motifs (residues 455–505 and 521–577; Pfam code PF00571). In CNNM1, CNNM3 and CNNM4 the corresponding CBS-pair regions include residues 438–561, 323–445 and 377–511, respectively. The cyclic nucleotide monophosphate (cNMP)-binding-like domain (Pfam code PF00027; residues 634–779, 649–805, 521–591 and 575–695 in CNNM1, CNNM2, CNNM3 and CNNM4, respectively) is a motif that is usually present in ion channels and cNMP-dependent kinases. In CNNM1, CNNM2 and CNNM4 the region located between the CBS pair and the cNMP-binding domain (residues 574–614, 589–629 and 516–555, respectively) shares high sequence similarity with the Sec14p motif which is found in *S. cerevisiae* phosphatidylinositol-transfer protein. TM predictions were performed with the TMHMM server v2.0 (<http://www.cbs.dtu.dk/services/TMHMM/>). Known pathological modifications of CNNM2 and CNNM4 are marked for each case. The UniProt database codes for CNNM1, CNNM2, CNNM3 and CNNM4 are Q9NRU3, Q9H8M5, Q8NE01 and Q6P4Q7, respectively.

2006), in transport and binding of Mg^{2+} (Ishitani *et al.*, 2008), in modulation of intracellular trafficking of chloride channels (Carr *et al.*, 2003), in nitrate transport (De Angeli *et al.*, 2009) and as 'internal inhibitors' of pyrophosphatase activity (Jämsen *et al.*, 2010; Tuominen *et al.*, 2010). Interestingly, mutations in the CBS domains of several human proteins have been associated with human diseases (MIM: 607803, MIM:180105 and MIM:600858; Bowne *et al.*, 2002; Scott *et al.*, 2004; Kemp, 2004). With the aim of providing detailed information that will help in understanding the involvement of CNNM2 in the development of familial hypomagnesaemia (MIM:607803) and in unravelling the underlying molecular mechanisms, we initiated crystallographic studies of its individual domains. Here, we describe the results obtained for the purification, crystallization and preliminary analysis of its CBS-domain pair as a first step towards the elucidation of the three-dimensional structure of the whole protein, for which no information is available.

2. Materials and methods

2.1. Cloning, expression and purification of the CBS-domain pair of CNNM2

Various truncated constructs of CNNM2 containing the CBS domains were designed based on alignment with the corresponding region of the human CNNM4 transporter, which has recently been crystallized (Gómez-García *et al.*, 2011). Two final constructs with potential crystallizability, CNNM2_{429–584} (156 residues; MW = 17 849.4; pI = 4.35) and CNNM2_{429–589} (161 residues; MW = 18 443.0; pI = 4.31), were selected after careful analysis of all candidates using the XtalPred protein crystallizability prediction server (<http://ffas.burnham.org/XtalPred/cgi/xtal.pl>; Slabinski *et al.*, 2007). Owing to the exact sequence identity between the murine and human proteins in the CBS-domain region, we decided to use in-house cloned cDNA of murine CNNM2 (Stuiver *et al.*, 2011) as a template to obtain the corresponding constructs, which were amplified by PCR using the forward primer 5'-CACCATGAAGGAAGAGCTGAAC-ATCATC-3' and the reverse primers 5'-TTACTCATCCAAGAT-TTCAGATTTG-3' for CNNM2_{429–584} and 5'-TTAGGTGTACAG-GTCTGTCTCAT-3' for CNNM2_{429–589}. To ensure that none of the resulting protein constructs contained extra amino-acid residues from the plasmid, the forward primer used for directional cloning into pET101/D-TOPO included an ATG start codon immediately before amino acid 429 of CNNM2. On the other hand, the reverse primers placed the TAA stop codons directly after residue 584 or 589 of the CNNM2 sequence. The PCR products were cloned into pET101/D-TOPO plasmids (Invitrogen) and the sequence of the cloned fragments was confirmed by sequencing from the T7 promoter and T7 terminator sequencing points. The plasmids were transformed into *Escherichia coli* strain BL21 Star (DE3) (Invitrogen). 20 ml starter cultures were grown overnight at 310 K and used to inoculate 2 l cultures of Luria-Bertani medium containing 100 mg l⁻¹ ampicillin. The cells were grown at 310 K to an OD₆₀₀ nm of around 0.4. Protein expression was induced by the addition of isopropyl β -D-1-thiogalactopyranoside (IPTG) to a final concentration of 0.5 mM and the cultures were left with shaking for a further 6 h at 310 K. The cells were harvested by centrifugation at 11 000g for 15 min at 277 K, resuspended in lysis buffer (25 mM MES pH 6.0, 1 mM EDTA, 1 mM β -mercaptoethanol, 1 mM benzamidine, 0.1 mM PMSF) and lysed by sonication in a Labsonic P sonicator (Sartorius; 4–5 cycles of 15 s at 90% amplitude with 30 s resting on ice between each cycle to prevent sample overheating). The cell lysate obtained was clarified by ultracentrifugation at 250 000g for 30 min at 277 K and the supernatant

was filtered through a 0.45 μm filter before being applied onto a pre-equilibrated 5 ml HiTrap Q HP column (GE Healthcare) connected to an ÄKTA FPLC system (GE Healthcare) installed within a refrigerated cabinet that maintained the temperature at 277 K during the purification process. The column was then washed with ten column volumes of buffer A (25 mM MES pH 6.0, 1 mM EDTA, 1 mM β -mercaptoethanol) and the bound protein was eluted with buffer B (25 mM MES pH 6.0, 1 mM EDTA, 1 mM β -mercaptoethanol, 1 M NaCl) using a linear gradient over 30 column volumes. The fractions containing the protein of interest (confirmed by SDS-PAGE) were pooled and dialyzed overnight against 300–500 volumes of buffer A. The sample was ultracentrifuged again as described above to remove precipitated protein and filtered through a 0.22 μm filter before being injected onto a 1 ml Mono Q 5/50 GL column (GE Healthcare) pre-equilibrated with buffer A. The column was washed with ten column volumes of buffer A and eluted with a 0–100% gradient of buffer B over 30 column volumes. The fractions of interest were pooled and concentrated using an Amicon Ultra-15 (5000 Da cutoff) centrifugal concentrator (Millipore) to a volume of approximately 2 ml. Subsequently, the concentrated protein was applied onto a HiLoad Superdex 75 16/60 Prep Grade gel-filtration column (GE Healthcare) pre-equilibrated with 25 mM MES pH 6.0, 200 mM NaCl, 1 mM EDTA, 1 mM β -mercaptoethanol and eluted at a flow rate of 0.5 ml min⁻¹. Fractions containing pure protein were pooled and concentrated to around 50 mg ml⁻¹ using an Amicon Ultra-4 5000 Da cutoff concentrator (Millipore). The CNNM2_{429–584} construct was overexpressed and purified with an approximate yield of 10 mg protein per litre of original culture, whereas the CNNM2_{429–589} construct had a much lower yield of approximately 1 mg protein per litre of original culture. Both were judged to be greater than 95% pure by SDS-PAGE analysis. The identity of the proteins was confirmed by mass spectrometry. The amino-acid sequence of CNNM2_{429–584} showed the presence of five nonterminal methionine residues, making it a good candidate for phasing by selenomethionine labelling. Selenomethionine-substituted CNNM2_{429–584} protein was prepared using the methionine-biosynthesis metabolic inhibition method (Doublé, 1997). 20 ml LB culture containing 100 $\mu\text{g ml}^{-1}$ ampicillin was prepared overnight at 310 K. This starter culture was pelleted by centrifugation at 3900g and 277 K for 10 min. This pellet was washed twice with sterile cold M9 minimal medium, resuspended in 20 ml M9 medium pre-warmed to 310 K and used to inoculate 1 l selenomethionine growth medium comprising M9 minimum medium plus sterile-filtered 2 mM MgSO_4 , 0.2% (w/v) glucose, 0.001% (w/v) thiamine, 100 μM CaCl_2 and 0.004% each of all proteinogenic amino acids except glycine, alanine, proline, asparagine, cysteine and methionine (all from Sigma). Cells were grown at 310 K to an OD₆₀₀ of 0.3, at which point 100 mg l⁻¹ threonine, lysine and phenylalanine, 50 mg l⁻¹ leucine, isoleucine and valine and 60 mg l⁻¹ selenomethionine (Acros Organics) were added. The high levels of some amino acids together with the absence of others inhibits the synthesis of methionine in the cells, so that selenomethionine is incorporated instead. The cells were left to grow for a further 30 min before being induced with 1 mM IPTG. The cells were then left to grow overnight at 290 K. The purification of selenomethionine-labelled (SeMet) CNNM2_{429–584} was carried out as described previously for the native form. The concentrations of the purified proteins were determined by UV absorption at 280 nm (OD₂₈₀) using theoretical extinction coefficients computed from their amino-acid sequences ($\epsilon_{280} = 4470$ and 5960 M⁻¹ cm⁻¹ for CNNM2_{429–584} and CNNM2_{429–589}, respectively). Fractions containing the pure proteins were pooled and concentrated using Vivaspin concentrators (5000 Da molecular-weight cutoff) to a final concentration of 50 mg ml⁻¹ for crystallization trials. Pure

proteins were frozen in liquid nitrogen and stored at 193 K. SDS-PAGE (Laemmli, 1970) was used to analyze the protein purity.

2.2. Mass-spectrometric analysis

SDS-PAGE gel bands containing the CNNM2 constructs were subjected to in-gel tryptic digestion according to Shevchenko *et al.* (1996) with minor modifications. The gel piece was swollen in a digestion buffer consisting of 50 mM NH_4HCO_3 and 12.5 ng μl^{-1} trypsin (Roche Diagnostics) in an ice bath. After 30 min, the supernatant was removed and discarded, 20 μl 50 mM NH_4HCO_3 was added to the gel piece and digestion was allowed to proceed at 310 K overnight. Prior to MS analysis, the sample was acidified by adding 5 μl 0.5% TFA. 0.5 μl of the digested sample was directly spotted onto the MALDI target and was then mixed with 0.5 μl α -cyano-4-hydroxycinnamic acid (CHCA) matrix solution [20 μg μl^{-1} in 70:30(v:v) acetonitrile/0.1% TFA]. Peptide mass fingerprinting was performed on a Bruker Autoflex III mass spectrometer (Bruker Daltonics, Bremen, Germany). Positively charged ions were analyzed in reflector mode using delayed extraction. The spectra were obtained by randomly scanning the sample surface. About 600–800 spectra were averaged in order to improve the signal-to-noise ratio. The spectra were externally calibrated; the mass accuracy was <50 p.p.m. when external calibration was performed and typically <20 p.p.m. in the case of internal calibration. Protein identification was performed by searching a nonredundant protein database (NCBI) using the *Mascot* search engine (<http://matrixscience.com>). The following parameters were used for database searches: missed cleavages, 1; allowed modifications, carbamidomethylation of cysteine (complete) and oxidation of methionine (partial).

2.3. Crystallization

Crystallization trials were carried out by the vapour-diffusion technique using a sitting-drop format in 96-well MRC crystallization plates with a variety of commercially available screens (Crystal Screen HT and Index HT from Hampton Research, The JCSG Core Suites I–IV from Qiagen and The JCSG+ Suite and PACTpremier HT-96 from Molecular Dimensions). Screening plates were set up in the high-throughput crystallization facility at CIC bioGUNE and were incubated at a constant temperature of 293 K. Drops consisted

of 100 nl protein solution mixed with 100 nl precipitant solution and the reservoir volume was 50 μl ; the protein concentration was approximately 50 mg ml^{-1} . Improved crystals were subsequently obtained by refining the successful conditions using a hanging-drop format in 24-well VDX plates (Hampton Research) with drops consisting of 0.5 μl protein and 0.5 μl precipitant equilibrated over a reservoir volume of 0.5 ml.

2.4. Preliminary crystallographic analysis

Prior to data collection, the crystals were transferred into crystallization buffer containing 25% sorbitol as a cryoprotectant for few seconds before being flash-cooled by direct immersion into liquid nitrogen at 93 K. Crystals were mounted for X-ray data collection using either CryoLoops (Hampton Research) or MicroMounts loops (MiTeGen). Data sets were collected in-house using a CCD detector mounted on a Microstar-H rotating-anode X-ray generator (Bruker) operated at 60 kV and 100 mA with Helios optics and a copper target ($\text{Cu K}\alpha$; $\lambda = 1.542$ Å) and on beamlines ID14.1 and ID23.1 at the ESRF, Grenoble, France. Diffraction data were processed using *HKL-2000* (Otwinowski & Minor, 1997). Preliminary analysis of the data sets was performed using the *CCP4* program suite (Winn *et al.*, 2011). A plot of the self-rotation function at $\kappa = 180^\circ$ was produced using *MOLREP* (Vagin & Teplyakov, 2010).

3. Results

We have overexpressed and purified two protein constructs, CNNM2_{429–589} (residues 429–589) and CNNM2_{429–584} (residues 429–584), that comprise the intracellular CBS-domain pair of the human/mouse cyclin M2 (CNNM2) protein. The CNNM2_{429–589} construct yielded hexagonal spear-shaped prisms that grew in 240 mM CaCl_2 , 14% PEG 6000, 0.1 M Tris pH 8.0 but diffracted poorly (to ~ 8 Å

Figure 2

Crystals of the CBS-domain pair of CNNM2. (a) Bipyramidal crystals of SeMet CNNM2_{429–584} (space group $I222$ or $I2_12_12_1$). (b) Rhombohedral or square-prismatic crystals of wild-type CNNM2_{429–584} (space group $P2_12_12$). (c) Irregular hexagonal prismatic crystals of CNNM2_{429–589} (which diffracted poorly to ~ 8 Å resolution).

Figure 3

Representative X-ray diffraction image from the $P2_12_12$ CNNM2_{429–584} crystals. The dimensions of the crystals were approximately $0.7 \times 0.7 \times 0.4$ mm and they were exposed for 1 s over a 1° simple rotation range. Resolution circles are included for clarity.

Table 1

Data-processing statistics for CNNM2_{429–584} crystals.

Values in parentheses are for the outer resolution shell.

Data set	CNNM2 _{429–584}	SeMet CNNM2 _{429–584}
Beamline	ID23.1, ESRF	ID23.1, ESRF
Wavelength (Å)	0.9793	0.9793
Total No. of reflections	47643	90175
No. of unique reflections	14016 (663)	6711 (657)
Resolution range (Å)	50–2.0	50–3.6
Space group	<i>P</i> 2 ₁ 2 ₁ 2	<i>I</i> 222 or <i>I</i> 2 ₁ 2 ₁ 2 ₁
Unit-cell parameters (Å, °)	<i>a</i> = 55.789, <i>b</i> = 64.420, <i>c</i> = 53.783, $\alpha = \beta = \gamma = 90$	<i>a</i> = 102.513, <i>b</i> = 104.593, <i>c</i> = 105.694, $\alpha = \beta = \gamma = 90$
Unit-cell volume (Å ³)	193293.5	1133271.5
Mosaicity (°)	1.1	1.3
Completeness (%)	97.0 (99.6)	100 (100)
Multiplicity	3.5 (2.8)	13.5 (12.8)
<i>R</i> _{merge} [†] (%)	9.5 (43.6)	15.1 (54.7)
Mean <i>I</i> / σ (<i>I</i>)	12.0 (1.5)	22.7 (6.9)
Wilson <i>B</i> value (Å ²)	25.2	66.4

[†] $R_{\text{merge}} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$, where $I_i(hkl)$ is the *i*th observation of reflection *hkl* and $\langle I(hkl) \rangle$ is the weighted average intensity for all observations *i* of reflection *hkl*.

resolution; Fig. 2c). In contrast, the CNNM2_{429–584} construct yielded rhombohedral or square-prismatic crystals, occasionally with two bevelled opposing corners, that grew 1–2 d after setup at 293 K from 60 mM NaH₂PO₄, 1.340 M K₂HPO₄ and diffracted X-rays to 2 Å resolution (Figs. 2b and 3, Table 1). Despite the apparent shape heterogeneity of these crystals, preliminary diffraction data showed that both the rhombohedral and the square-prismatic crystals had the same unit-cell parameters and belonged to the same space group (*P*2₁2₁2; Figs. 2b and 3, Table 1). Additionally, we produced SeMet CNNM2_{429–584} protein, which yielded bipyramidal crystals that grew overnight in 325 mM calcium acetate, 0.1 M imidazole pH 8 at 293 K and diffracted X-rays to 3.6 Å resolution (Fig. 2, Table 1).

The presence of one or two molecules within the asymmetric unit gave Matthews coefficients of 2.72 or 1.36 Å³ Da^{−1} with solvent contents of 55 and 10%, respectively, for the *P*2₁2₁2 crystals, indicating that one molecule was the most probable value (Matthews, 1968). The presence of one, two or three molecules within the asymmetric unit gave Matthews coefficients of 7.92, 3.97 or 2.65 Å³ Da^{−1} and solvent contents of 85, 69 or 53%, respectively, for the *I*222/*I*2₁2₁2₁ crystal form. Considering the fragility of the *I*222/*I*2₁2₁2₁ crystals and their limited diffraction power, we estimated that two molecules per asymmetric unit (a dimer of the CBS-domain pair of CNNM2) was the most probable value. In support of this, the corresponding plot of the self-rotation function at $\kappa = 180^\circ$ showed a twofold axis in the *yz* plane at approximately 45° to the *z* axis, suggesting that the two molecules are related by a twofold symmetry axis and probably form a dimer. Such a symmetric dimer is likely to correspond to the most frequently found association of CBS-domain pairs, which is known as the ‘CBS module’ (Lucas *et al.*, 2010; Gómez-García *et al.*, 2010; De Baaij *et al.*, 2012). Data-collection statistics are summarized in Table 1. Crystal structures from both crystal forms of CNNM2_{429–584} are now being determined in our laboratory.

We thank the staff of ESRF beamline ID14.1, and especially Alexander Popov at beamline ID23.1 for his valuable support during synchrotron data collection. We also thank Dr Felix Elortza of the Proteomics Service at CIC bioGUNE for mass-spectrometric analysis and Dr Adriana Rojas for maintenance of the in-house X-ray equipment. This work was supported by grants from Departamento de Educación, Universidades e Investigación del Gobierno Vasco

(PI2010-17), Departamento de Industria, Innovación, Comercio y Turismo del Gobierno Vasco (ETORTEK IE05-147 and IE07-202), Diputación Foral de Bizkaia (Exp. 7/13/08/2006/11 and 7/13/08/2005/14), Ministerio Español de Ciencia e Innovación (MICINN; BFU2010-17857) and the MICINN CONSOLIDER-INGENIO 2010 Program (CSD2008-00005). This study was supported by a grant from the European Community FP7 (EUNEFRON 201590) to DM.

References

- Agüero-Llera, D., Oyenarte, I., Martínez-Cruz, L. A. & Neira, J. L. (2010). *FEBS Lett.* **584**, 4485–4489.
- Agus, M. S. & Agus, Z. S. (2001). *Crit. Care Clin.* **17**, 175–186.
- Agus, Z. S. (1999). *J. Am. Soc. Nephrol.* **10**, 1616–1622.
- Alderton, A., Davies, P., Illman, K. & Brown, D. R. (2007). *J. Neurochem.* **103**, 312–321.
- Baaij, J. H. de, Stuijver, M., Meij, I. C., Lainez, S., Kopplin, K., Venselaar, H., Müller, D., Bindels, R. J. & Hoenderop, J. G. (2012). *J. Biol. Chem.* **287**, 13644–13655.
- Bateman, A. (1997). *Trends Biochem. Sci.* **22**, 12–13.
- Baykov, A. A., Tuominen, H. K. & Lahti, R. (2011). *ACS Chem. Biol.* **6**, 1156–1163.
- Biemans-Oldehinkel, E., Mahmood, N. A. & Poolman, B. (2006). *Proc. Natl Acad. Sci. USA*, **103**, 10624–10629.
- Bowne, S. J., Sullivan, L. S., Blanton, S. H., Cepko, C. L., Blackshaw, S., Birch, D. G., Hughbanks-Wheaton, D., Heckenlively, J. R. & Daiger, S. P. (2002). *Hum. Mol. Genet.* **11**, 559–568.
- Carr, G., Simmons, N. & Sayer, J. (2003). *Biochem. Biophys. Res. Commun.* **310**, 600–605.
- Cefaratti, C., Romani, A. & Scarpa, A. (2000). *J. Biol. Chem.* **275**, 3772–3780.
- De Angeli, A., Moran, O., Wege, S., Filleur, S., Ephritikhine, G., Thomine, S., Barbier-Brygoo, H. & Gambale, F. (2009). *J. Biol. Chem.* **284**, 26526–26532.
- Doublé, S. (1997). *Methods Enzymol.* **276**, 523–530.
- Eshaghi, S., Niegowski, D., Kohl, A., Martínez Molina, D., Lesley, S. A. & Nordlund, P. (2006). *Science*, **313**, 354–357.
- Gibson, M. M., Bagga, D. A., Miller, C. G. & Maguire, M. E. (1991). *Mol. Microbiol.* **5**, 2753–2762.
- Gómez-García, I., Oyenarte, I. & Martínez-Cruz, L. A. (2010). *J. Mol. Biol.* **399**, 53–70.
- Gómez-García, I., Oyenarte, I. & Martínez-Cruz, L. A. (2011). *Acta Cryst.* **F67**, 349–353.
- Goytain, A. & Quamme, G. A. (2005). *Physiol. Genomics*, **22**, 382–389.
- Günther, T. (1993). *Miner. Electrolyte Metab.* **19**, 259–265.
- Guo, D., Ling, J., Wang, M.-H., She, J.-X., Gu, J. & Wang, C.-Y. (2005). *Mol. Pain*, **1**, 15.
- Hattori, M., Iwase, N., Furuya, N., Tanaka, Y., Tsukazaki, T., Ishitani, R., Maguire, M. E., Ito, K., Maturana, A. & Nureki, O. (2009). *EMBO J.* **28**, 3602–3612.
- Hattori, M., Tanaka, Y., Fukai, S., Ishitani, R. & Nureki, O. (2007). *Nature (London)*, **448**, 1072–1075.
- Ishitani, R., Sugita, Y., Dohmae, N., Furuya, N., Hattori, M. & Nureki, O. (2008). *Proc. Natl Acad. Sci. USA*, **105**, 15393–15398.
- Jämsen, J., Baykov, A. A. & Lahti, R. (2010). *Biochemistry*, **49**, 1005–1013.
- Kemp, B. E. (2004). *J. Clin. Invest.* **113**, 182–184.
- Laemmli, U. K. (1970). *Nature (London)*, **227**, 680–685.
- Lucas, M., Encinar, J. A., Arribas, E. A., Oyenarte, I., García, I. G., Kortazar, D., Fernández, J. A., Mato, J. M., Martínez-Chantar, M. L. & Martínez-Cruz, L. A. (2010). *J. Mol. Biol.* **396**, 800–820.
- Lunin, V. V., Dobrovetsky, E., Khutovskaya, G., Zhang, R., Joachimiak, A., Doyle, D. A., Bochkarev, A., Maguire, M. E., Edwards, A. M. & Koth, C. M. (2006). *Nature (London)*, **440**, 833–837.
- Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
- Otwinski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
- Parry, D. A. *et al.* (2009). *Am. J. Hum. Genet.* **84**, 266–273.
- Polok, B., Escher, P., Ambresin, A., Chouery, E., Bolay, S., Meunier, I., Nan, F., Hamel, C., Munier, F. L., Thilo, B., Mégarbané, A. & Schorderet, D. F. (2009). *Am. J. Hum. Genet.* **84**, 259–265.
- Quamme, G. A. (1997). *Kidney Int.* **52**, 1180–1195.
- Quamme, G. A. (2010). *Am. J. Physiol. Cell Physiol.* **298**, C407–C429.
- Romani, A. M. & Scarpa, A. (2000). *Front. Biosci.* **5**, D270–D374.
- Schweigert, M., Kolisek, M., Nikolic, Z. & Kuzinski, J. (2008). *Magn. Res.* **21**, 118–123.

- Scott, J. W., Hawley, S. A., Green, K. A., Anis, M., Stewart, G., Scullion, G. A., Norman, D. G. & Hardie, D. G. (2004). *J. Clin. Invest.* **113**, 274–284.
- Shabb, J. B. & Corbin, J. D. (1992). *J. Biol. Chem.* **267**, 5723–5726.
- Shevchenko, A., Wilm, M., Vorm, O. & Mann, M. (1996). *Anal. Chem.* **68**, 850–858.
- Slabinski, L., Jaroszewski, L., Rychlewski, L., Wilson, I. A., Lesley, S. A. & Godzik, A. (2007). *Bioinformatics*, **23**, 3403–3405.
- Stuiver, M. *et al.* (2011). *Am. J. Hum. Genet.* **88**, 333–343.
- Tashiro, M., Konishi, M., Iwamoto, T., Shigekawa, M. & Kurihara, S. (2000). *Pflugers Arch. Eur. J. Physiol.* **440**, 819–827.
- Tuominen, H., Salminen, A., Oksanen, E., Jämsen, J., Heikkilä, O., Lehtiö, L., Magretova, N. N., Goldman, A., Baykov, A. A. & Lahti, R. (2010). *J. Mol. Biol.* **398**, 400–413.
- Vagin, A. & Teplyakov, A. (2010). *Acta Cryst.* **D66**, 22–25.
- Wang, C.-Y., Shi, J. -D., Yang, P., Kumar, P. G., Li, Q.-Z., Run, Q.-G., Su, Y.-C., Scott, H. S., Kao, K.-J. & She, J.-X. (2003). *Gene*, **306**, 37–44.
- Wang, C.-Y., Yang, P., Shi, J.-D., Purohit, S., Guo, D., An, H., Gu, J.-G., Ling, J., Dong, Z. & She, J.-X. (2004). *BMC Genomics*, **5**, 7.
- Watanabe, M., Konishi, M., Ohkido, I. & Matsufuji, S. (2005). *Am. J. Physiol. Renal Physiol.* **289**, F742–F748.
- Winn, M. D. *et al.* (2011). *Acta Cryst.* **D67**, 235–242.
- Yang, M., Jensen, L. T., Gardner, A. J. & Culotta, V. C. (2005). *Biochem. J.* **386**, 479–487.