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

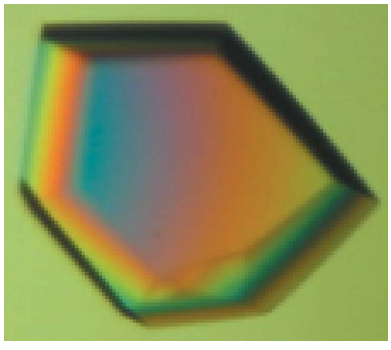
Purification, crystallization and preliminary X-ray diffraction analysis of the CBS-domain pair from the *Methanococcus jannaschii* protein MJ0100

María Lucas,[a] Danel Kortazar,[a] Egoitz Astigarraga,[b] José A. Fernández,[b] Jose M. Mato,[c] María Luz Martínez-Chantar[c] and Luis Alfonso Martínez-Cruz[a]*

[a]Structural Biology Unit, CIC bioGUNE, Parque Tecnológico de Bizkaia, Edificio 800, 48160 Derio, Bizkaia, Spain, [b]Departamento de Química-Física, Universidad del País Vasco UPV-EHU, Lejona, Spain, and [c]Metabolomics Unit, CIC bioGUNE, Parque Tecnológico de Bizkaia, Edificio 801, 48160 Derio, Bizkaia, Spain

Correspondence e-mail: amartinez@cicbiogune.es

Received 3 July 2008
Accepted 1 September 2008



CBS domains are small protein motifs consisting of a three-stranded β-sheet and two α-helices that are present in proteins of all kingdoms of life and in proteins with completely different functions. Several genetic diseases in humans have been associated with mutations in their sequence, which has made them promising targets for rational drug design. The C-terminal domain of the *Methanococcus jannaschii* protein MJ0100 includes a CBS-domain pair and has been overexpressed, purified and crystallized. Crystals of selenomethionine-substituted (SeMet) protein were also grown. The space group of both the native and SeMet crystals was determined to be orthorhombic $P2_12_12_1$, with unit-cell parameters $a = 80.9$, $b = 119.5$, $c = 173.3$ Å. Preliminary analysis of the X-ray data indicated that there were eight molecules per asymmetric unit in both cases.

1. Introduction

CBS domains are small motifs of approximately 60 amino-acid residues that were originally discovered in the enzyme cystathionine-β-synthase (Bateman, 1997). CBS domains are not unique to CBS but can be identified in a wide variety of proteins (Pfam database PF00571; http://www.sanger.ac.uk/Users/agb/CBS/CBS.html), alone or combined with other domains performing different functions. CBS domains are usually associated in tandems, forming a so-called CBS-domain pair or Bateman domain, although some proteins such as 5′-AMP-activated protein kinase (AMPK) contain four CBS domains (Ignoul & Eggermont, 2005). The first CBS-domain structure to be solved was the structure of inosine monophosphate dehydrogenase (IMPDH; Sintchak *et al.*, 1996; Zhang *et al.*, 1999). This structure showed that a single CBS domain consists of a three-stranded β-sheet and two α-helices packed according to the sequence $\beta1$-$\alpha1$-$\beta2$-$\beta3$-$\alpha2$. The crystal structures of CBS domains from many proteins are now available, such as the chloride channels ClC-0 from the ray *Torpedo marmorata* (Meyer & Dutzler, 2006), human ClC-5 (Meyer *et al.*, 2007) and human ClC-Ka (Markovic & Dutzler, 2007), human AMPK (Day *et al.*, 2007), yeast AMPK (Amodeo *et al.*, 2007; Jin *et al.*, 2007; Rudolph *et al.*, 2007; Townley & Shapiro, 2007), truncated rat–human AMPK (Xiao *et al.*, 2007) and the bacterial Mg transporter MgtE from *Thermus thermophilus* (Hattori *et al.*, 2007). Comparison of the CBS domains of the solved structures shows a highly conserved fold despite low sequence similarity. In all these structures, two CBS domains associate to form a compact structure with a cleft between the domains. This cavity has been proven to be the binding site for adenosyl groups in ClC5 and AMPK. The crystal structures of the complex ClC5–ADP/ATP (Meyer *et al.*, 2007) and of the complex AMPK–ATP/AMP/ADP/ZMP (Amodeo *et al.*, 2007; Day *et al.*, 2007; Jin *et al.*, 2007; Townley & Shapiro, 2007; Xiao *et al.*, 2007) raise the question of whether a CBS-domain pair has one or two nucleotide-binding sites. Although there is experimental evidence that the binding of adenosyl groups to the CBS domain of IMPDH (Scott *et al.*, 2004), AMPK (Cheung *et al.*, 2000) and CBS (Finkelstein *et al.*, 1975) has an influence on the catalysis of these proteins, the structural information available does not elucidate how the binding of the

adenosyl compound is transduced to the corresponding catalytic domains. It seems that the physiological functions and binding partners of the Bateman domains may vary considerably between different proteins (Pimkin & Markham, 2008).

Important human pathologies associated with mutations within the CBS domains have been described as homocystinuria (CBS; Shan *et al.*, 2001), autosomal retinitis pigmentosa (IMPDH; Hunter *et al.*, 2002), myotonia (ClC-1; Pusch, 2002), idiopathic epilepsy (ClC-2; Haug *et al.*, 2003), Bartter syndrome (ClC-Kb; Konrad *et al.*, 2000), osteopetrosis (ClC-7; Cleiren *et al.*, 2001) and familiar hypertrophic cardyomiopathy or Wolff–Parkinson–White syndrome (AMPK; Blair *et al.*, 2001), which enhances the chances that these domains will be promising targets for the development of novel drugs.

Some hyperthermophilic microorganisms from archaea, including *Methanococcus jannaschii* (Bult *et al.*, 1996), possess a large number of proteins with CBS domains in their genome. Therefore, they are excellent models for the study and characterization of the binding sites for different types of adenosyl groups in CBS domains, as well as for the study of the specificity of these proteins for their ligands. In this work, we present the cloning, expression, purification and crystallization of the CBS-domain pair of *M. jannaschii* protein MJ0100. The ORF of gene *mj0100* (UniProtKB/Swiss-Prot Q57564) codes for a polypeptide chain of 509 amino acids with a molecular weight of 56 458 Da. Its sequence contains two domains: a DUF39 (Pfam database PF01837) domain (residues 15–320) of so far unknown function found in bacteria and archaea and a CBS-domain pair

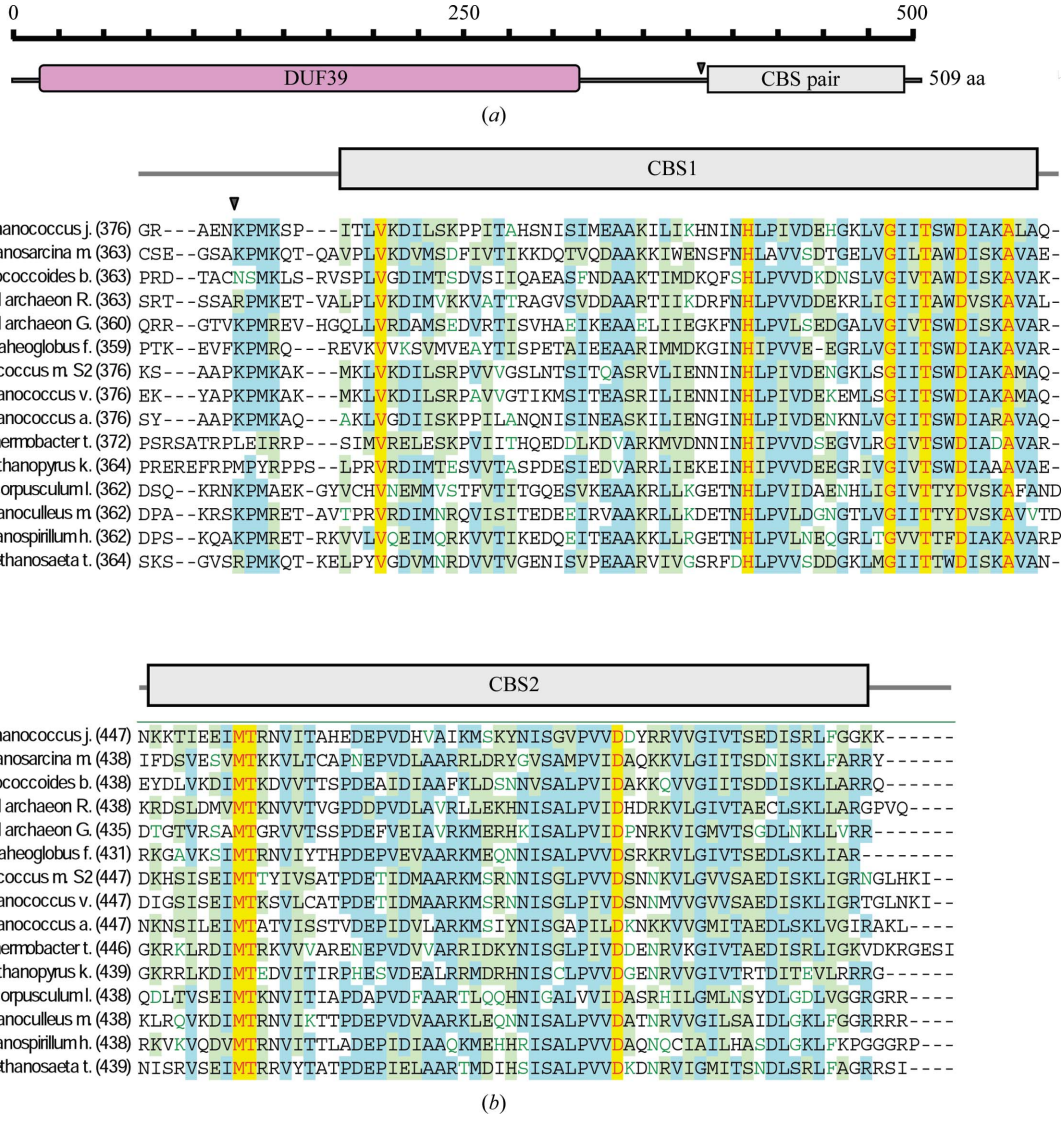


Figure 1
(*a*) Domain organization of protein MJ0100. (*b*) Sequence alignment of the CBS-domain pair of MJ0100 and related proteins. The figure shows a *ClustalW* (Larkin *et al.*, 2007) alignment of the C-terminal part of some of the proteins that contain a combination of DUF39 and CBS domains. The number in brackets indicates the number in the protein sequence of the first amino acid in the alignment. The beginning of the MJ0100c fragment is shown by a triangle. Colour codes: red on yellow background, invariant amino acids; black on blue, strongly conserved; black on green, similar; green on white, weakly similar; black on white, not conserved. The proteins shown and the accession numbers of the proteins in the alignment are MJ0100 from *Methanococcus jannaschii* (NP_247064), MM_0241 from *Methanosarcina mazei* Gö1 (AAM29937), DUF39 from *Methanococcoides burtonii* (YP_565063), RCIX1241 from uncultured archaeon RC-I (YP_685869), DUF39 from uncultured archaeon GZfos26B2 (AAU83016), AF1186 from *Archeoglobus fulgidus* DSM 4304 (NP_070015), MMP1359 from *Methanococcus maripaludis* S2 (NP_988479), DUF39 from *Methanococcus vannielii* SB (YP_001323187), DUF39 from *Methanococcus aeolicus* Nankai-3 (YP_001324988), DUF39 from *Methanothermobacter thermautotophicus* strain Delta H (NP_275992), DUF39 from *Methanopyrus kandlrei* AV19 (NP_614190), Mlab_0451 from *Methanocorpusculum labreanum* Z (YP_001029892), DUF39 from *Methanoculleus marisnigri* Jr1 (YP_001047428), DUF39 from *Methanospirillum hungatei* JF-1 (YP_501633) and DUF39 from *Methanosaeta thermophila* PT (YP_842462).

(residues 392–509) (Fig. 1). There are 15 sequences belonging to archaea with DUF39 + CBS architecture (Fig. 1*b*).

2. Cloning, expression and purification of MJ0100c

An initial attempt to purify full-length MJ0100 protein showed that it was insoluble under the conditions assayed. Therefore, we decided to study only the CBS-domain pair of the protein (MJ0100c; MW = 14 426 Da) consisting of residues 381–509. Plasmid pML1 carries the CBS-domain pair of MJ0100 (residues 381–509) under a T7 promoter. This plasmid was constructed by amplifying from *M. jannaschii* genomic DNA the 397 bp fragment that corresponds to the CBS-domain pair with oligonucleotides MJ0100-381F (CACCATGAAGCCAATGAAGTCACCAATAAC) and MJ0100R (TCATTTTTTCCCTCCGAATAATC). The PCR product was cloned in pET101D plasmid using the Champion pET Directional TOPO Expression Kit (Invitrogen). The DNA sequence was verified by sequencing and then transformed into *Escherichia coli* strain BL21-Codon Plus (Stratagene).

MJ0100c was purified starting from a 2 l BL21-Codon Plus (Stratagene) culture containing plasmid pML1 grown at 310 K with 100 µg ml^{-1} ampicillin and 25 µg ml^{-1} chloramphenicol. Overexpression of MJ0100c was induced by the addition of IPTG to a final concentration of 0.5 m*M* when the optical density of the culture at 600 nm was 0.6. Expression was allowed to proceed for 3 h. Cells were harvested by 10 min centrifugation at 3000*g* and 277 K and stored at 193 K. Cells were thawed at room temperature and resuspended in 80 ml lysis buffer (50 m*M* HEPES pH 7.0, 1 m*M* EDTA, 1 m*M* DTT, 1 m*M* benzamidine, 0.1 m*M* PMSF). Bacterial lysis was carried out by ultrasonication in a Labsonic P instrument (Sartorius). After centrifugation at 145 000*g* for 30 min at 277 K, protein MJ0100c remained mainly as a soluble protein. The resulting supernatant was loaded onto a P11-phosphocellulose (Whatman) column equilibrated with buffer *A*1 (50 m*M* HEPES pH 7.0, 200 m*M* NaCl, 1 m*M* EDTA, 1 m*M* DTT). The column contained 15 ml P11-phosphocellulose previously activated according to the manufacturer's specifications and packed in a K9 column (GE Healthcare). Unbound proteins were washed off with 60 ml buffer *A*1. Elution of the target protein was achieved with buffer *B*1 (50 m*M* HEPES pH 7.0, 750 m*M* NaCl, 1 m*M* EDTA, 1 m*M* DTT). 5 ml fractions were collected. The fractions containing MJ0100c were pooled and dialyzed against buffer *A*2 (50 m*M* HEPES pH 7.0, 50 m*M* NaCl, 1 m*M* EDTA, 1 m*M* DTT). The dialyzed protein was applied onto a 1 ml MonoS column (GE Healthcare) equilibrated with buffer *A*2. After washing with 15 column volumes of buffer *A*2, MJ0100c was eluted with a linear 30 ml NaCl gradient (50–600 m*M* NaCl) using buffer *B*2 (50 m*M* HEPES pH 7.0, 600 m*M* NaCl, 1 m*M* EDTA, 1 m*M* DTT) at a flow rate of 1 ml min^{-1} and collecting 1.0 ml fractions. In a final polishing step, concentrated MJ0100c was applied onto a high-resolution gel-filtration column (Superdex-75 10/300, GE Healthcare) equilibrated in buffer *A*1. The chromatography was performed at a flow rate of 0.5 ml min^{-1}, collecting 1.0 ml fractions. The protein eluted as a single peak with a molecular weight of 26 kDa, which is in concordance with the formation of a dimer (28.8 kDa). Dynamic light-scattering (DLS) analysis confirmed that MJ0100c (at a protein concentration of 1 mg ml^{-1}) is a dimer in solution in the pH range 5–9 (data not shown). The concentration of the purified MJ0100c protein was determined by UV absorption at 280 nm using the theoretical extinction coefficient computed from the amino-acid sequence (ε_{280} = 8250 M^{-1} cm^{-1}; Gill & von Hippel, 1989). The typical yield from a 2 l expression was 20 mg purified protein. The MJ0100c protein was flash-frozen in liquid nitrogen and stored at 193 K. SDS–PAGE (Laemmli, 1970) was used to analyze the protein purity (Fig. 2).

Purification of selenomethionine (SeMet) labelled MJ0100c was performed using *E. coli* B834 (DE3) strain (Novagen; Wood, 1966) and New Minimal Medium supplemented with seleno-L-methionine (SeMet; Budisa *et al.*, 1995). The purification was performed according to the previously described procedure. The yield was slightly lower than that of unlabelled material.

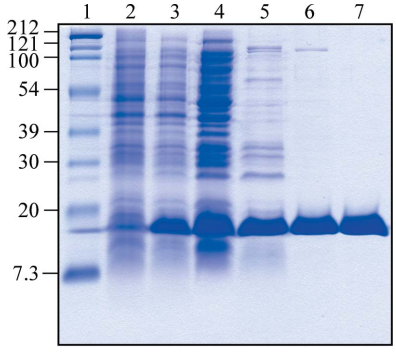


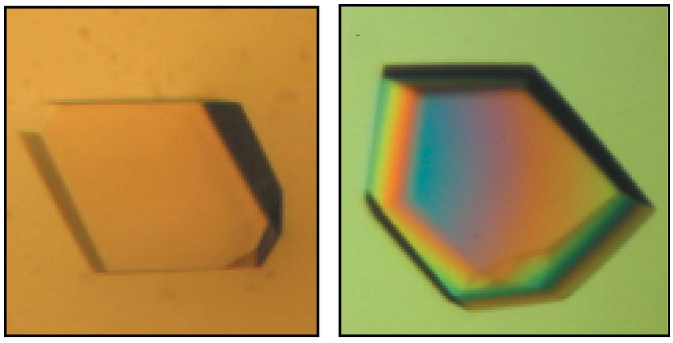
Figure 2
Purification of protein MJ0100c. SDS–PAGE analysis of fractions obtained in a standard purification scheme, as reported in §2. Lane 1, molecular-weight markers (in kDa); lane 2, BL21-Codon Plus (DE3) + pML1 before induction; lane 3, BL21-Codon Plus (DE3) + pML1 after 3 h induction with IPTG; lane 4, soluble lysate after ultracentrifugation; lane 5, 20 µg sample after phosphocellulose chromatography; lane 6, 20 µg sample after MonoS chromatography; lane 7, 20 µg sample after gel-filtration chromatography.

3. Mass-spectrometric analysis

SDS–PAGE gel bands 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 containing 50 m*M* 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 m*M* NH_4HCO_3 was added to the gel piece and digestion was allowed to proceed at 310 K overnight. Prior to mass-spectrometric analysis, the sample was acidified by adding 5 µl 0.5% TFA. 0.5 µl digested sample was directly spotted onto the MALDI target and then mixed with 0.5 µl α-cyano-4-hydroxycinnamic acid matrix solution [20 µg µl^{-1} in acetonitrile, 0.1% TFA, 70:30(*v*:*v*)]. Peptide mass fingerprinting was performed on a Bruker Autoflex III mass spectrometer (Bruker Daltonics). 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. Spectra were externally calibrated, resulting in a mass accuracy of <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 in a nonredundant protein database (NCBI) using the *MASCOT* search engine (http://matrixscience.com). The following parameters were used for database searches: one missed cleavage with allowed modifications carbamidomethylation of cysteine (complete) and oxidation of methionine (partial). Ultimately, MJ0100c expressed in *E. coli* was wild type with no mutations in the amino-acid sequence according to the *M. jannaschii* genome sequence database. Mass spectrometry indicated that the extent of SeMet labelling was >95% (data not shown).

4. Crystallization of MJ0100c

Purified protein was dialyzed against 50 m*M* HEPES pH 7.0 and concentrated to 100–150 mg ml^{-1} using Vivaspin columns (Sartorius) with a molecular-weight cutoff of 5 kDa. Crystallization trials were set up using the hanging-drop vapour-diffusion technique in 24-well VDX plates (Hampton Research) at 293 K. Drops consisting of 1 µl protein solution mixed with 1 µl reservoir solution were equilibrated against 500 µl reservoir solution. Suitable protein concentrations were determined with the native protein using the Pre-Crystallization Screen (PCT, Hampton Research). Initial crystallization conditions were found using Crystal Screen (Hampton Research). Crystals were obtained when the precipitant was 8% PEG 4000 in Tris–HCl pH 8.5 (condition No. 36). These crystallization conditions were optimized with respect to PEG and pH. The best crystallization conditions were 10–15% PEG 6000, 8000 or 10 000 in 100 m*M* Tris–HCl. No differences were observed in the pH range of the Tris–HCl buffer between 8.4 and 9.4. The crystals appeared overnight and grew to maximum dimensions of about 1 × 0.7 × 0.5 mm within 2–3 d (Fig. 3*a*). Crystals which appeared to have two different habits grew in the same crystallization drop, but both had identical unit-cell parameters and space-group symmetry. Crystals of SeMet-substituted MJ0100c (Fig. 3*b*) were obtained using the same crystallization conditions as used for the native crystals.

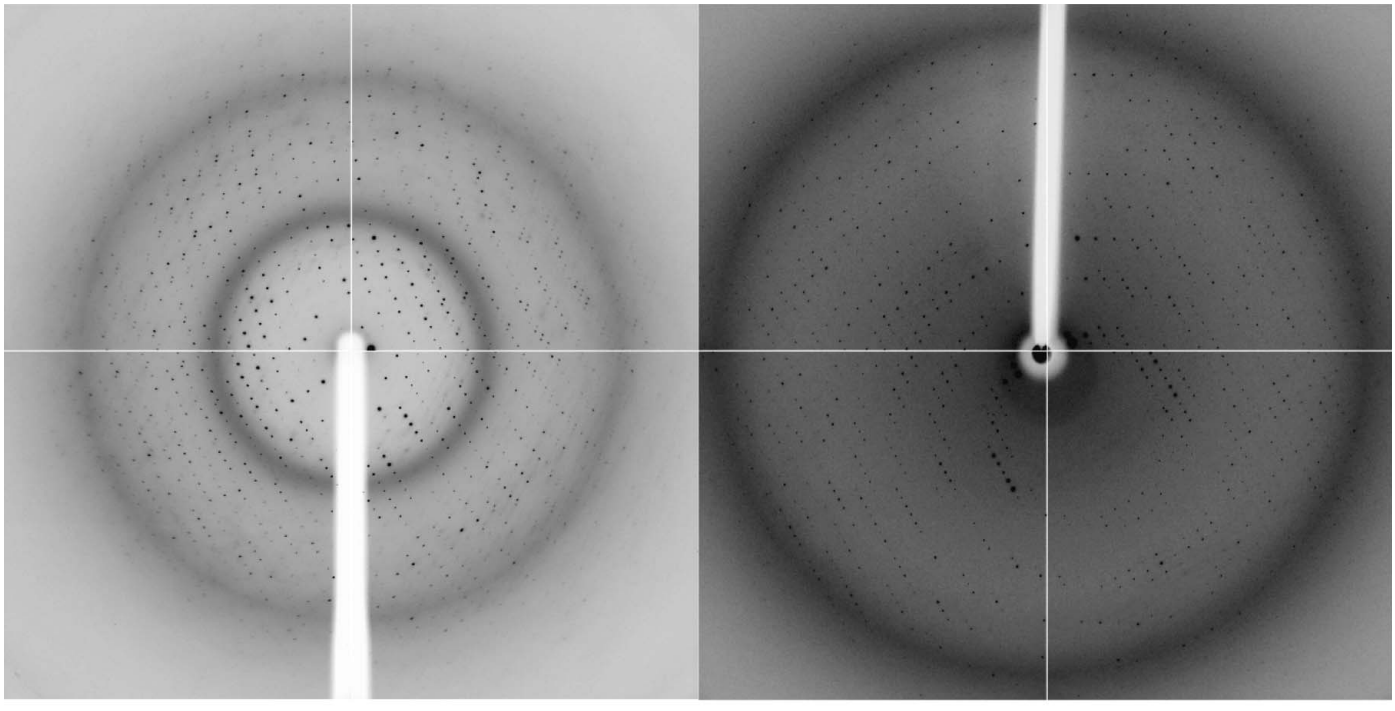


(*a*) (*b*)

Figure 3
(*a*) Crystal of MJ0100c grown in 15% PEG 10 000, 100 m*M* Tris–HCl pH 8.8. (*b*) Crystal of SeMet-labelled MJ0100c grown in 10% PEG 8000, 100 m*M* Tris–HCl pH 8.6, 1 m*M* DTT. Crystal size was 0.3 mm.

5. Preliminary crystallographic analysis

Finding optimal cryogenic conditions to freeze MJ0100c crystals was not straightforward and required significant experimental effort. Owing to their size, the largest crystals grown by the hanging-drop or sitting-drop techniques were not suitable for freezing and usually (in 95% of the trials) fractured just after opening the cover slip of the reservoir or, in the best case, after a brief exposure (1–2 s) to various cryoprotectants including MPD, low-molecular-weight PEGs, sucrose, erythritol, glycerol or ethylene glycol. In order to prevent freezing problems arising from physical tension between the inner and outer regions of the sample, we only froze crystals in the size range 0.1–0.5 mm (in the largest dimension). The best result was obtained by soaking the crystals in crystallization solution containing a final glycerol concentration of 25% and a slight increase (5%) in the precipitant. Unfortunately, this approach only yielded optimal results in 5% of the samples tested. In parallel, crystals were also harvested with stepwise transfer *via* a series of solutions containing increasing concentrations of the cryoprotective agent for varying times with no further success. Dehydration of the crystals by slowly increasing the PEG concentration (in steps of 5%) in the reservoir solution for varying times (from 1 h to one week) prior to flash-cooling in liquid nitrogen and/or to the addition of the cryoprotectants was also tested,



(*a*) (*b*)

Figure 4
Representative X-ray diffraction image from (*a*) native and (*b*) SeMet-labelled MJ0100c crystals. The diffraction ring observed in (*a*) at around 7–8 Å is caused by the presence of silicone oil, which was used as a cryoprotectant in freezing the native crystals. The size of the native and derivatized crystals were 0.4 and 0.3 mm, respectively, and they were exposed for 10 s over a 1° oscillation range. The edge of the detector corresponds to a resolution of 2.13 Å in (*a*) and 3.3 Å in (*b*).

Table 1
Data-processing statistics for native and SeMet-labelled MJ0100c.

Values in parentheses are for the outer resolution shell. No accurate values of the *B* factors could be estimated from the corresponding Wilson plots, since the crystals did not diffract to resolutions lower than 3 Å (Drenth, 2002).

		SeMet-labelled MJ0100c	
	MJ0100c	Peak	Inflection point
Beamline	ID14-2, ESRF	BM16, ESRF	BM16, ESRF
Wavelength (Å)	0.93400	0.97797	0.97908
Temperature (K)	100	100	100
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Unit–cell parameters (Å, °)	a = 80.989, b = 119.527, c = 173.285, $\alpha = \beta = \gamma = 90$	a = 82.395, b = 120.555, c = 175.460, $\alpha = \beta = \gamma = 90$	a = 82.395, b = 120.555, c = 175.460, $\alpha = \beta = \gamma = 90$
Matthews coefficient (Å^3 Da^{-1})	3.7	3.7	3.7
Solvent content (%)	66.5	67.1	67.1
Unit-cell volume (Å^3)	1677463.1	1742866.8	1742866.8
Molecules per ASU	8	8	8
No. of observations	201817	175674	161485
No. of unique reflections	31157	24739	22820
Resolution range (Å)	98–3.1	50–3.4	50–3.5
Mosaicity (°)	0.65	0.51	0.67
Completeness (%)	98.2 (99.1)	99.9 (100)	100 (99.9)
Redundancy	6.6 (5.7)	7.1 (7.2)	7.1 (7.2)
Mean $I/\sigma(I)$	18.43 (5.08)	17.1 (4.3)	17.1 (5.6)
$R_{\rm merge}$† (%)	7.4 (28.5)	9.6 (39.5)	10.1 (36.4)

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

with no further success. In parallel, we tried a different approach to freeze the crystals, consisting of covering the hanging-drop crystallization drops with paraffin oil, Paratone-N or silicon oil DC200 (Fluka) just after opening the cover slip of the crystallization well. This technique yielded optimal results and avoided crystal dehydration and fracture (Pflugrath, 2004). Once the crystallization drop had been covered with oil, the crystals were carefully moved from the inner aqueous solution to the surrounding silicon oil, where they seemed to be stable. Water surrounding the crystal surface was then carefully removed by gently displacing the crystal within the silicon oil with the help of a loop. The crystals could then be flash-cooled by immersion in liquid N_2 prior to data collection. To check whether the cryoprotectants (glycerol and silicone oil) might be responsible for the limited resolution of the data sets, we analyzed the diffraction images at both room temperature and after freezing the crystals. Similar results were obtained, suggesting that the resolution limit might be intrinsic to the crystals despite their apparent external beauty.

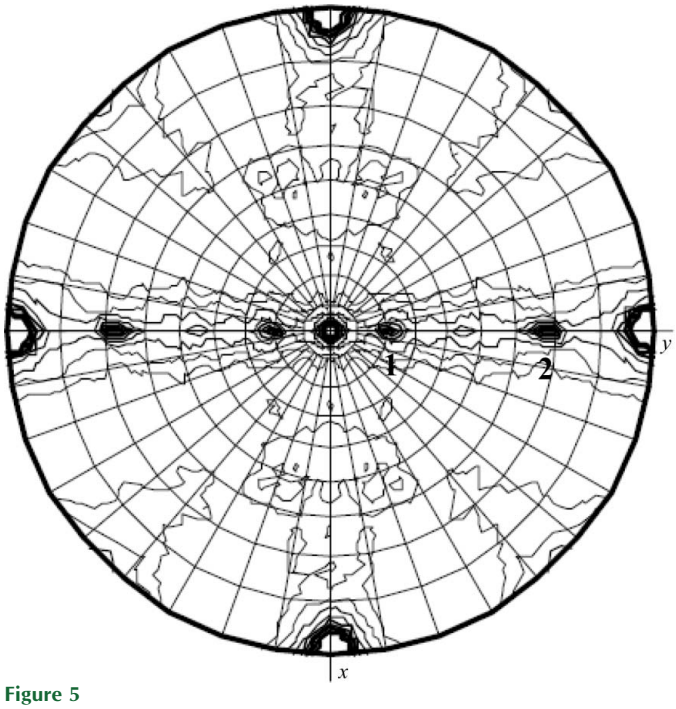


Figure 5
$\kappa = 180°$ section of the self-rotation function of the native MJ0100c data set. CBS-domain pairs usually dimerize in a head-to-head or a head-to-tail manner, both being related by pseudo-twofold symmetry. As shown here, there are two noncrystallographic twofold axis (marked as 1 and 2) in the yz plane that would be consistent with the presence of two dimers (four molecules) in the asymmetric unit. However, the V_M (7.29 Å^3 Da^{-1}) and solvent content (83.14%) values would be too high in this case. Consequently, an additional twofold NCS axis parallel to any of the crystallographic axes a, b or c may be present (although it cannot be distinguished because it overlaps with the crystallographic axis). Consequently, eight molecules (four dimers) in the asymmetric unit may be the most probable value. No additional peaks are observed in the $\kappa = 60°$, 90° or 120° sections of the self-rotation function. This figure was prepared using *MOLREP* (Vagin & Teplyakov, 1998).

Derivative crystals were also very fragile but could be frozen by transferring them for few seconds into a cryosolution containing 27.5% glycerol, 7.5% PEG 6000, 50 m*M* Tris–HCl pH 8.6. Longer soaking periods resulted in crystal breakage. The crystals were then flash-cooled in liquid nitrogen.

Crystals were pre-screened using in-house X8 Proteum equipment (Bruker). A native data set was collected to 3.1 Å resolution at 100 K on beamline ID14-2 at the European Synchrotron Research Facility (ESRF), Grenoble, France from a crystal grown in 15% PEG 10 000, 100 m*M* Tris–HCl pH 8.8 (Fig. 4*a*). The oscillation range was 1.0° and 200 images were collected on an ADSC Q4R CCD detector. Two-wavelength SeMet MAD data were collected to 3.4 Å resolution at 100 K at station BM16 at the ESRF synchrotron (Grenoble, France) from a crystal grown in 10% PEG 8000, 100 m*M* Tris–HCl pH 8.6, 1 m*M* DTT. The oscillation range was 1.0° and 180 images were collected at peak and inflection-point wavelengths on an ADSC Quantum-210 detector (Fig. 4*b*). Data sets were processed using the *HKL*-2000 package (Otwinowski & Minor, 1997).

Native crystals of MJ0100c diffracted to 3.1 Å resolution (Fig. 4*a*) and belonged to space group $P2_12_12_1$, with unit-cell parameters a = 80.9, b = 119.5, c = 173.3 Å. The presence of six, eight or ten molecules within the asymmetric unit give a Matthews coefficient of

4.86, 3.67 and 2.92 Å^3 Da^{-1}, respectively (Matthews, 1968) and a solvent content of 74.7, 66.5 and 57.85%, respectively. After careful analysis of the self-rotation function (Fig. 5) and considering the fragility of MJ0100c crystals and their limited diffraction power, we estimated that eight molecules per asymmetric unit was the most probable value. The crystals of MJ0100c labelled with SeMet diffracted to 3.4 Å resolution (Fig. 4*b*) and belonged to space group $P2_12_12_1$, with unit-cell parameters $a = 82.4$, $b = 120.6$, $c = 175.5$ Å; the asymmetric unit was also likely to contain eight monomers. This is consistent with a Matthews coefficient of 3.7 Å^3 Da^{-1} and a solvent content of 67.1%. The data-collection statistics are summarized in Table 1.

We thank Professor Sung-Hou Kim from University of California at Berkeley for providing us with the genomic DNA from *M. jannaschii* and the staff of ESRF beamlines BM16 and ID14-2 for support during synchrotron data collection. We also thank Beatriz González Callejas from CIC bioGUNE for maintenance of the in-house X-ray equipment. This research was supported by program grants from the Basque Government (ETORTEK IE05-147, IE07-202), Diputación Foral de Bizkaia (Exp. 7/13/08/2006/11 and 7/13/08/2005/14) and the Spanish Ministry of Education (SAF2005-00855), as well as a postdoctoral fellowship from CIC bioGUNE.

References

Amodeo, G. A., Rudolph, M. J. & Tong, L. (2007). *Nature (London)*, **449**, 492–495.
Bateman, A. (1997). *Trends Biochem. Sci.* **22**, 12–13.
Blair, E., Redwood, C., Ashrafian, H., Oliveira, M., Broxholme, J., Kerr, B., Salmon, A., Ostman-Smith, I. & Watkins, H. (2001). *Hum. Mol. Genet.* **10**, 1215–1220.
Budisa, N., Steipe, B., Demange, P., Eckerskorn, C., Kellermann, J. & Huber, R. (1995). *Eur. J. Biochem.* **230**, 788–796.
Bult, C. J. *et al.* (1996). *Science*, **273**, 1058–1073.
Cheung, P. C., Salt, I. P., Davies, S. P., Hardie, D. G. & Carling, D. (2000). *Biochem. J.* **346**, 659–669.
Cleiren, E., Benichou, O., Van, H. E., Gram, J., Bollerslev, J., Singer, F. R., Beaverson, K., Aledo, A., Whyte, M. P., Yoneyama, T., deVernejoul, M. C. & Van, H. W. (2001). *Hum. Mol. Genet.* **10**, 2861–2867.
Day, P., Sharff, A., Parra, L., Cleasby, A., Williams, M., Hörer, S., Nar, H., Redemann, N., Tickle, I. & Yon, J. (2007). *Acta Cryst.* D**63**, 587–596.
Drenth, J. (2002). *Principles of Protein X-ray Crystallography*, 120–122. Berlin: Springer.
Finkelstein, J. D., Kyle, W. E., Martin, J. L. & Pick, A. M. (1975). *Biochem. Biophys. Res. Commun.* **66**, 81–87.
Gill, S. C. & von Hippel, P. H. (1989). *Anal. Biochem.* **182**, 319–326.
Hattori, M., Tanaka, Y., Fukai, S., Ishitani, R. & Nureki, O. (2007). *Nature (London)*, **448**, 1072–1075.
Haug, K. *et al.* (2003). *Nature Genet.* **33**, 527–532.
Hunter, M. S., Ussher, J. M., Browne, S. J., Cariss, M., Jelley, R. & Katz, M. (2002). *J. Psychosom. Obstet. Gynaecol.* **23**, 193–199.
Ignoul, S. & Eggermont, J. (2005). *Am. J. Physiol. Cell Physiol.* **289**, C1369–C1378.
Jin, X., Townley, R. & Shapiro, L. (2007). *Structure*, **15**, 1285–1295.
Konrad, M., Vollmer, M., Lemmink, H. H., van den Heuvel, L. P., Jeck, N., Vargas-Poussou, R., Lakings, A., Ruf, R., Deschenes, G., Antignac, C., Guay-Woodford, L., Knoers, N. V., Seyberth, H. W., Feldmann, D. & Hildebrandt, F. (2000). *J. Am. Soc. Nephrol.* **11**, 1449–1459.
Laemmli, U. K. (1970). *Nature (London)*, **227**, 680–685.
Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J. & Higgins, D. G. (2007). *Bioinformatics*, **23**, 2947–2948.
Markovic, S. & Dutzler, R. (2007). *Structure*, **15**, 715–725.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Meyer, S. & Dutzler, R. (2006). *Structure*, **14**, 299–307.
Meyer, S., Savaresi, S., Forster, I. C. & Dutzler, R. (2007). *Nature Struct. Mol. Biol.* **14**, 60–67.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Pflugrath, J. W. (2004). *Methods*, **34**, 415–423.
Pimkin, M. & Markham, G. D. (2008). *Mol. Microbiol.* **68**, 342–359.
Pusch, M. (2002). *Hum. Mutat.* **19**, 423–434.
Rudolph, M. J., Amodeo, G. A., Iram, S. H., Hong, S. P., Pirino, G., Carlson, M. & Tong, L. (2007). *Structure*, **15**, 65–74.
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.
Shan, X., Dunbrack, R. L. Jr, Christopher, S. A. & Kruger, W. D. (2001). *Hum. Mol. Genet.* **10**, 635–643.
Shevchenko, A., Wilm, M., Vorm, O. & Mann, M. (1996). *Anal Chem.* **68**, 850–858.
Sintchak, M. D., Fleming, M. A., Futer, O., Raybuck, S. A., Chambers, S. P., Caron, P. R., Murcko, M. A. & Wilson, K. P. (1996). *Cell*, **85**, 921–930.
Townley, R. & Shapiro, L. (2007). *Science*, **315**, 1726–1729.
Vagin, A. & Teplyakov, A. (1998). *Acta Cryst.* D**54**, 400–402.
Wood, W. B. (1966). *J. Mol. Biol.* **16**, 118–133.
Xiao, B., Heath, R., Saiu, P., Leiper, F. C., Leone, P., Jing, C., Walker, P. A., Haire, L., Eccleston, J. F., Davis, C. T., Martin, S. R., Carling, D. & Gamblin, S. J. (2007). *Nature (London)*, **449**, 496–500.
Zhang, R., Evans, G., Rotella, F., Westbrook, E., Huberman, E., Joachimiak, A. & Collart, F. R. (1999). *Curr. Med. Chem.* **6**, 537–543.