Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structure of a CBS-domain pair from the regulatory γ1 subunit of human AMPK in complex with AMP and ZMP

Philip Day,[a]* Andrew Sharff,[a] Lina Parra,[a] Anne Cleasby,[a] Mark Williams,[a] Stefan Hörer,[b] Herbert Nar,[b] Norbert Redemann,[b] Ian Tickle[a] and Jeff Yon[a]

[a]Astex Therapeutics Ltd, 436 Cambridge Science Park, Milton Road, Cambridge CB4 0QA, England, and [b]Boehringer Ingelheim Pharma GmbH & Co. KG, 88397 Biberach an der Riss, Germany

Correspondence e-mail: p.day@astex-therapeutics.com

Received 24 January 2007
Accepted 23 February 2007

PDB References: AMP complex, 2uv4, r2uv4sf; ZMP complex, 2uv5, r2uv5sf; R299Q AMP complex, 2uv6, r2uv6sf; R299G AMP complex, 2uv7, r2uv7sf.

AMP-activated kinase (AMPK) is central to sensing energy status in eukaryotic cells *via* binding of AMP and ATP to CBS (cystathionine β-synthase) domains in the regulatory γ subunit. The structure of a CBS-domain pair from human AMPK γ1 in complex with the physiological activator AMP and the pharmacological activator ZMP (AICAR) is presented.

1. Introduction

AMPK plays a key role in regulation of metabolism and is a potential therapeutic target in diabetes, obesity and metabolic syndrome (Kahn *et al.*, 2005; Kemp *et al.*, 2003). It acts as a sensor of the energy status of eukaryotic cells such that an increase in the cellular AMP:ATP ratio results in activation of AMPK, which then acts to inhibit energy-utilizing (anabolic) pathways and activate energy-producing (catabolic) pathways (Hardie *et al.*, 2003). The mechanism of activation of AMPK is complex, involving phosphorylation by an upstream kinase, LKB1 (Hawley *et al.*, 2003; Woods *et al.*, 2003; Shaw *et al.*, 2004), and direct allosteric activation by AMP (Salt *et al.*, 1998). Activation of AMPK plays a key role in the regulation of both lipid and glucose metabolism. Phosphorylation of acetyl-CoA carboxylase (ACC) by AMPK results in a decrease in cellular malonyl-CoA levels, reduced fatty-acid synthesis and an increase in fatty-acid oxidation (Park *et al.*, 2002). Synthesis of triacylglycerol and cholesterol is also inhibited by phosphorylation of glycerol phosphate acyl-transferase and HMG-CoA reductase, respectively, by AMPK (Muoio *et al.*, 1999; Clarke & Hardie, 1990). AMPK activation also results in increased expression and translocation of GLUT-4 receptors to the plasma membrane, resulting in increased cellular glucose uptake (Holmes *et al.*, 1999; Kurth-Kraczek *et al.*, 1999).

AMPK is a heterotrimeric protein composed of a catalytic serine/threonine kinase subunit (α) and two regulatory subunits (β and γ; Hardie & Hawley, 2001; Hardie *et al.*, 1998). Recently, the structures of Snf1, the yeast homologue of the AMPK catalytic α subunit, and the glycogen-binding domain of the β subunit of rat AMPK have been published (Rudolph *et al.*, 2005; Nayak *et al.*, 2006; Polekhina *et al.*, 2005), but there is presently no known structure of any γ subunit. In humans there are multiple isoforms of each subunit (α1, α2, β1, β2, γ1, γ2 and γ3), permitting the formation of 12 different AMPK trimer isoforms, showing different tissue distribution. Assembly of the complete trimer is essential for AMPK

Table 1
Data-collection and refinement statistics.

Values in parentheses are for the highest resolution shell. PCMB, *p*-chloromercuribenzoate; SeMet, selenomethionine.

Crystal	AMP	ZMP	R299Q	R299G	ATP	SeMet	PCMB	K_2PtCl_4
Data collection								
Space group	*C*2	*C*2	*C*2	*C*2	*C*2	*C*2	*C*2	*C*2
Unit-cell parameters								
a (Å)	66.132	66.056	65.995	66.092	66.543	66.567	67.643	66.757
b (Å)	43.757	43.885	43.836	43.775	43.733	43.856	43.111	43.992
c (Å)	55.654	55.798	55.734	55.786	55.648	55.918	55.687	55.620
β (°)	118.835	118.176	118.794	119.040	118.739	118.836	119.232	118.934
Resolution (Å)	35–1.33 (1.38–1.33)	25–1.70 (1.79–1.70)	35–2.00 (2.07–2.00)	35–2.00 (2.07–2.00)	35–1.50 (1.55–1.50)	41–2.30 (2.38–2.30)	25–2.00 (2.07–2.00)	34–2.20 (2.28–2.20)
Wavelength (Å)	0.931	1.5418	1.5418	1.5418	0.931	0.933	0.931	0.931
Unique reflections	31832	15986	9520	9345	22600	6418	9586	7288
Redundancy	6.2 (4.3)	2.5 (2.2)	3.6 (3.4)	3.6 (3.1)	2.5 (2.5)	6.7 (5.1)	3.8 (3.8)	2.7 (1.9)
Completeness (%)	99.0 (90.7)	97.9 (91.1)	99.8 (99.7)	98.0 (93.1)	97.6 (99.2)	98.3 (98.3)	99.3 (100)	93.0 (76.8)
Mean $I/\sigma(I)$	12.5 (3.1)	9.6 (3.2)	11.3 (3.6)	7.9 (2.5)	6.2 (2.1)	15.7 (4.4)	12.2(3.8)	9.3(2.7)
R_{merge}† (%)	6.9 (40.0)	5.7 (23.7)	6.3 (26.2)	8.9 (33.2)	9.6 (39.4)	10.6 (29.6)	6.1 (26.6)	7.6 (27.7)
Refinement								
$R_{working}$‡	17.2	17.0	16.2	17.7				
R_{free}	20.9	23.0	22.9	25.4				
No. of atoms (protein)	1139	1139	1137	1132				
No. of atoms (ligand)	23	22	23	23				
No. of atoms (water)	239	247	187	185				
R.m.s.d. bonds (Å)	0.013	0.012	0.014	0.013				
R.m.s.d. angles (°)	1.4	1.4	1.4	1.3				
Average *B* factor (Å^2)								
Main chain	11.9	17.4	20.0	19.8				
Side chain	16.7	21.9	24.8	24.2				
Solvent	34.8	36.7	36.5	36.8				
Ligand	12.5	18.2	19.2	19.1				

† $R_{merge} = \sum |I_i - \langle I \rangle| / \sum I_i$, where I_i is the observed intensity and I is the mean intensity. ‡ $R_{working} = \sum(|F_o| - |F_c|)/\sum |F_o|$, where F_o are the observed structure-factor amplitudes and F_c are the calculated structure factors.

activity. The β-subunit contains a glycogen-binding region and high cellular glycogen content reduces AMPK activity, suggesting that AMPK could also be part of a glycogen-sensing pathway (Polekhina *et al.*, 2003, 2005; Hudson *et al.*, 2003). The γ-subunit binds AMP and ATP and is thought to be responsible for allosteric activation of the complex (Cheung *et al.*, 2000). The three γ-subunits have variable N-terminal extensions followed by four tandem repeats of a motif known as a cystathionine β-synthase (CBS) domain.

CBS domains are small 50–60 amino-acid domains found in a wide variety of proteins and, although their precise function is poorly understood, they are thought to play a regulatory role in the action of these proteins (Bateman, 1997; Ignoul & Eggermont, 2005). They have been shown to bind adenosyl-containing molecules such as AMP, ATP and *S*-adenosyl-methionine (Scott *et al.*, 2004). The importance of CBS domains is highlighted by the presence of disease-causing mutations in the CBS domains of several proteins, including AMPK, CBS, chloride channels and IMPDH (Milan *et al.*, 2000; Blair *et al.*, 2001; Kluijtmans *et al.*, 1996; Lloyd *et al.*, 1997; Konrad *et al.*, 2000; Pusch, 2002; Bowne *et al.*, 2002; Kennan *et al.*, 2002). CBS domains generally occur in pairs (forming a Bateman module) and it has recently been demonstrated that a pair of CBS domains is required to bind a single adenine nucleotide (Scott *et al.*, 2004). AMPK γ chains are unique in eukaryotes in having two pairs of CBS domains, both of which have been shown to be able to bind a nucleotide. The structure presented here gives the first view of the specific interactions of AMP and ZMP with a CBS-domain pair from a γ subunit.

The incidence of type II diabetes is increasing rapidly, owing in part to the large increase in obesity, the primary contributor to the pathogenesis of diabetes. Many of the current therapeutic approaches to treatment of type II diabetes were developed without knowledge of defined molecular targets. Two of these therapeutics, metformin and rosiglitazone, have subsequently been found to act, at least in part, *via* activation of AMPK (Fryer *et al.*, 2002; Hawley *et al.*, 2002; Saha *et al.*, 2004; Zhou *et al.*, 2001). Furthermore, treatment with AICA-riboside (AICAR), a compound that is taken up by cells and metabolized to ZMP, an AMP mimic and AMPK activator, improved glucose tolerance and lipid profiles in an insulin-resistant rat model (Bergeron *et al.*, 2001; Buhl *et al.*, 2002; Iglesias *et al.*, 2002; Song *et al.*, 2002). Taken together, these data suggest that AMPK may be a suitable target against which to develop drugs for the treatment of type II diabetes. In particular, compounds that increase AMPK activity through binding to the allosteric site on the γ subunit may have therapeutic utility. In order to further elucidate the mechanism of AMP activation, we have solved the crystal structures of a CBS-domain pair from the human γ1 subunit in complex with the physiological activator AMP and with the pharmacological activator ZMP. These structures should assist in structure-based design of compounds binding at this site.

2. Experimental procedures

2.1. Protein production

A construct encoding residues 182–325 with an N-terminal hexahistidine tag was used for structure determination (all numbering is by reference to the SwissProt database sequence P54619). An overnight culture of *Escherichia coli* BL21 (DE3) cells bearing the appropriate AMPK γ1 (CBS3+4) construct was used to inoculate 1 l TB containing 50 µg ml^{-1} kanamycin and grown at 310 K to an OD_{600} of 1.0. Expression was induced by addition of 0.1 m*M* IPTG overnight at 298 K. Harvested cell pellets were stored at 193 K. Pellets were resuspended to 250 ml in lysis buffer (50 m*M* Tris–HCl pH 8, 20% glycerol, 500 mM NaCl, 1.4 m*M* β-mercaptoethanol, 5 m*M* ATP or 0.5 m*M* AMP) containing 10 m*M* imidazole before lysis by sonication. The clarified supernatant was tumbled with Ni–NTA agarose (Qiagen) for at least 1 h at 277 K and the resin was then washed extensively with lysis buffer followed by lysis buffer containing 20 m*M* imidazole. Protein was eluted in lysis buffer containing 200 m*M* imidazole, concentrated and applied onto a Sephacryl S200HR 26/60 gel-filtration column pre-equilibrated in gel-filtration buffer (20 m*M* Tris–HCl pH 8, 20% glycerol, 500 m*M* NaCl, 1.4 m*M* β-mercaptoethanol, 1 m*M* ATP or 0.5 m*M* AMP). Peak fractions were pooled and concentrated to 10 mg ml^{-1}.

2.2. Crystallization and preparation of complexes

For crystallization, purified protein (10 mg ml^{-1}) was exchanged into a buffer comprising 20 m*M* Tris–HCl pH 8.0, 20% glycerol, 1.4 m*M* β-mercaptoethanol and the appropriate nucleotide (0.5 m*M* AMP or 1 m*M* ATP). Crystals were grown in hanging drops of 1 µl protein solution and 1 µl reservoir buffer (0.1 *M* Tris–acetate pH 7.5 and 0.8 *M* trisodium citrate) at 293 K. Crystals appeared overnight and belong to space group *C*2, with unit-cell parameters a = 66.5, b = 43.7, c = 55.6 Å, β = 118.7° and a single molecule in the asymmetric unit. Crystals of selenomethionine-containing protein were grown under the same conditions as the wild-type protein and were isomorphous with the wild-type crystals. Crystals could be soaked and frozen in a solution containing 0.1 *M* Tris–acetate pH 7.5, 0.8 *M* trisodium citrate and 25% glycerol. Heavy-atom derivatives were prepared by soaking crystals for 5 h in 2 m*M* heavy-atom compound. The heavy-atom search produced two useful derivatives: *p*-chloromercuribenzoate (PCMB) and K_2PtCl_4. Complexes with ZMP (5-aminoimidazole-4-carboxamide-1-β-4-ribofuranotide) were prepared by soaking crystals grown in the presence of ATP in 20 m*M* ZMP. The crystals were transferred to 2 µl drops containing 0.1 *M* Tris–acetate pH 7.5, 0.8 *M* trisodium citrate, 25% glycerol plus ZMP for 1–2 d.

2.3. Data collection and structure solution

Data was collected on beamlines ID14EH3 (MAR CCD) or ID14EH2 (ADSC Quantum 4 CCD) at the ESRF or in-house on a Rigaku RU3HR rotating-anode generator (R-AXIS HTC) at 100 K. Data were processed, merged and scaled using *MOSFLM* (Leslie, 1992), *CAD* and *SCALEIT* from the *CCP*4 suite of programs (Collaborative Computational Project, Number 4, 1994). Data-collection statistics are shown in Table 1. The structure of AMPK γ1 CBS3+4 crystallized in the presence of ATP was solved by MIRAS using the selenomethionine protein and two derivatives. A single platinum site was located manually by inspection of Patterson maps. This was input into *SHARP* (de La Fortelle & Bricogne, 1997) and used for phasing and refinement. Difference Fourier methods located one mercury site and two selenium sites. These sites were re-input to *SHARP* and refined and phases were calculated to 2.0 Å. These phases together with the native structure factors were input to the automated model-building programme *ARP*/*wARP* (Perrakis *et al.*, 1993) and after a total of ten rebuilding cycles alternated with five refinement cycles using *REFMAC* (Murshudov *et al.*, 1997) a total of 135 amino acids out of an expected maximum of 158 were built into the map in the correct sequence. Manual rebuilding and refinement gave a total of 151 residues in the final model. This structure was used as a starting point for refinement of the AMP complex at 1.3 Å and for refinement of the ZMP complex at 1.7 Å. In both cases the presence of the ligand in the active site was clearly indicated by the difference density. The ligand was built into the maps, water molecules were added and the complex structures were refined to a final $R_{working}$ and R_{free} of 17.2% and 20.9%, respectively, for the AMP complex and 17.0% and 23%, respectively, for the ZMP complex. The geometry of the structures as illustrated by the Ramachandran plot was good, with the AMP structure having 92.4% and 7.6% of its residues in the most favoured region and additional allowed regions, respectively, and the corresponding values for ZMP complex being 91.7% and 8.3%, respectively. There were no residues in the disallowed region. The statistics for the final structures are shown in Table 1.

3. Results

3.1. Construct design

Previous studies have suggested that expression of full-length and fragments of AMPK γ1 as GST fusions yielded soluble protein (Scott *et al.*, 2004). In our hands, such constructs yielded soluble but highly aggregated material that was unsuitable for structural studies. Full-length His_6-tagged AMPK γ1 was mostly in the insoluble fraction when produced in *E. coli* and the limited soluble material failed to yield crystals. A series of truncations were made based on sequence alignments and published structures of CBS-domain pairs in order to investigate whether a fragment of AMPK γ1 could be expressed in high yield. All constructs that contained the CBS1+2 domain pair either failed to express or proved to be predominantly insoluble. In contrast, a set of constructs based around the CBS3+4 domain pair, or residues 182–331, proved to be soluble following expression at 291 K. N-terminal truncations up to residue 185 and C-terminal truncations back to residue 325 were tolerated, but further truncation rendered the construct insoluble.

Two constructs [His_6-γ1(182–331) and His_6-γ1(182–325)] were progressed for purification. Initial purification trials yielded soluble but aggregated material. Cell lysis in high salt, glycerol and, crucially, ligand (ATP or AMP) facilitated the purification of soluble material that was monodisperse as judged by DLS and that crystallized readily. The majority of crystallization conditions were PEG-based, gave large crystals belonging to space group $P2_1$ and suffered from partial pseudo-merohedral twinning. A single condition containing a high concentration of trisodium citrate gave small triangular crystals belonging to space group $C2$ that diffracted to high resolution (1.33 Å; Table 1).

3.2. Domain structure

For construct His_6-γ1(182–325) residues 182–323 were visible in the electron-density map along with the final histidine of the His_6 tag. For construct His_6-γ1(182–331) residues 324–331 were also visible.

The overall structure shows the expected fold for a CBS-domain pair. The core of a CBS domain consists of a $\beta\alpha\beta\beta\alpha$ structure in which the three β-strands form a sheet with the two helices to one side. The human γ1 chain consists of CBS domains 1, 2, 3 and 4, where domains 1 and 2 form a CBS-domain pair (Bateman module) as do domains 3 and 4. The N-terminus of the single molecule in the asymmetric unit has a short relatively unstructured segment and a short α-helix that precedes another single turn of α-helix (α1 in Fig. 1) that can be considered to be the beginning of the conserved CBS-domain structure. The core of CBS3 consists of $\beta2\alpha2\beta3\beta4\alpha3$ and that of CBS4 $\beta5\alpha6\beta6\beta7\alpha7$ (the numbering of the strands corresponds to that in Fig. 1*a*). CBS3 is formed from a continuous linear sequence consisting of α-helices $\alpha2$, $\alpha3$ and $\alpha4$ and a three-stranded β-sheet (strands $\beta2$, $\beta3$ and $\beta4$), whereas CBS4 is discontinuous, consisting of α-helices $\alpha1$, $\alpha6$ and $\alpha7$ and a four-stranded β-sheet (strands $\beta1$, $\beta5$, $\beta6$ and $\beta7$). CBS3 has an extended helix, $\alpha4$, rather than the short helix and fourth β-strand ($\alpha1$ and $\beta1$) found in CBS4, which together with the presence of $\alpha5$ in the connecting interdomain loop creates asymmetry in the two-domain structure (Fig. 1*a*). The structure presented here is homologous to those of other CBS-domain pairs whose structure has been determined, with the major structural differences being in the interdomain linkers, the loops linking secondary structure and in the N- and C-termini. The main secondary-structural elements overlay well with those of other CBS-domain pairs, with the exception of those of CLC-1 and CLC-5, where there is major rearrangement of the helices equivalent to $\alpha2$ and $\alpha3$ (of AMPK γ1 CBS3) in the first CBS domain (Meyer & Dutzler, 2006; Meyer *et al.*, 2007).

The crystal packing results in the formation of a crystallographic dimer across the twofold symmetry axis (Fig. 1*b*). The extensive dimer interface (1740 $Å^2$) involves exposed hydrophobic residues on helices $\alpha2$ and $\alpha3$ of CBS3 packing against similarly exposed hydrophobic residues on helices $\alpha6$ and $\alpha7$ of CBS4 of the symmetry-related molecule and *vice versa*. Conservation of residues across γ chains of different species is very high, with sequence identity greater than 90% for γ1

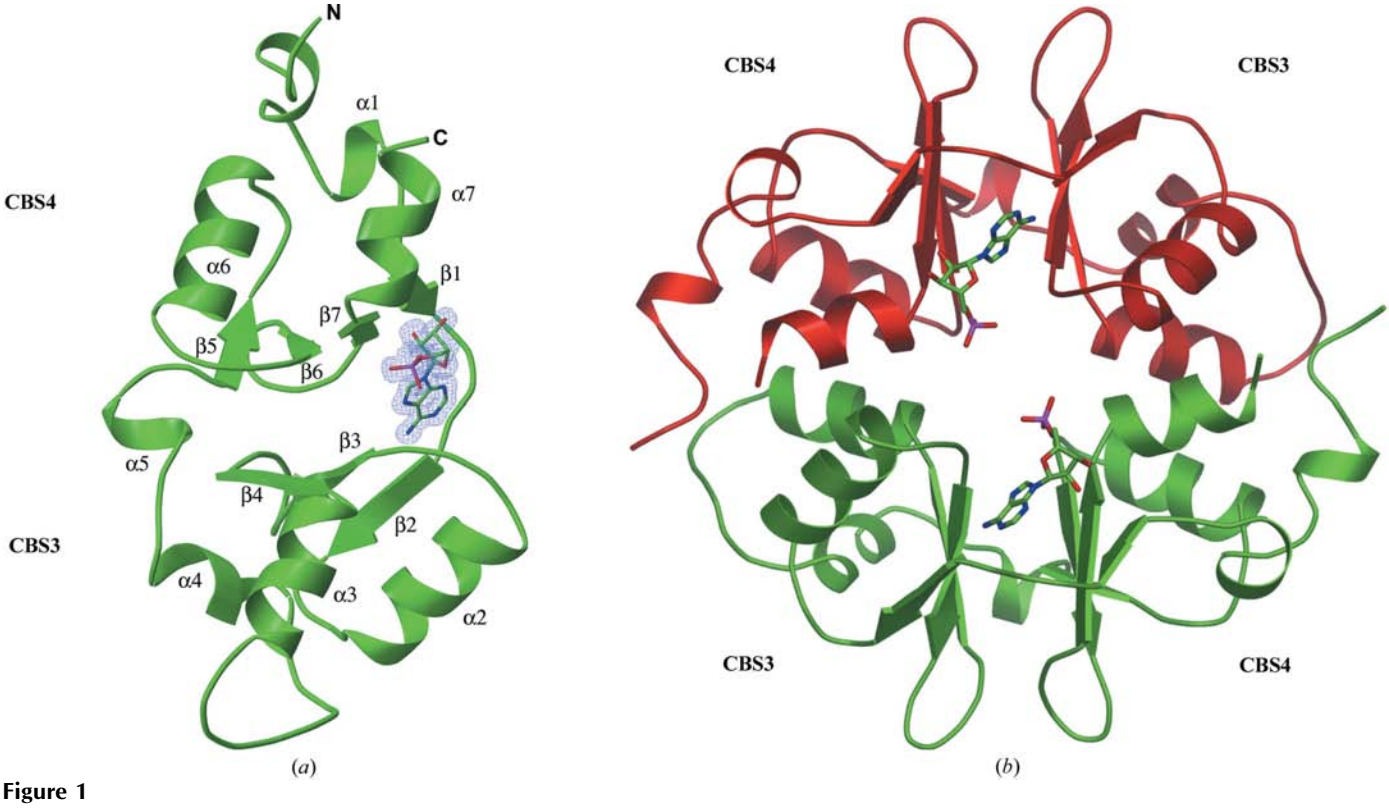

Figure 1
The structure of AMPK γ1 CBS domains 3 and 4. (*a*) Cartoon representation of the structure with bound AMP and $|F_o| - |F_c|$ density for the ligand. (*b*) The crystallographic dimer.

chains from human, pig, cow, mouse and rat. Similarly, sequence identity between known $\gamma1$, $\gamma2$ and $\gamma3$ chains is also high. Conservation of residues across the observed dimer interface is also very high, such that substitutions in AMPK $\gamma2$ and $\gamma3$ are conservative and maintain the pattern of exposed hydrophobic residues. This suggests that the exposed hydrophobic surfaces on the monomer, formed by helices $\alpha2$, $\alpha3$, $\alpha6$ and $\alpha7$, may play a prominent role in forming protein–protein interactions.

Gel filtration and dynamic light scattering (DLS) of purified His_6-$\gamma1$(182–331) was suggestive of a multimeric state in solution, although the full-length γ chain appeared to be monomeric (data not shown). Evidence for higher order oligomerization of the AMPK complex is contradictory (Davies *et al.*, 1994; Neumann *et al.*, 2003). However, it has been shown that the kinase domain of Snf1, the yeast homologue of the α subunit of AMPK, forms dimers that may have physiological relevance and the authors state that the dimerization surfaces of this subunit are conserved in the AMPK family (Nayak *et al.*, 2006; Rudolph *et al.*, 2005). Two other publically available structures of CBS-domain pairs are also dimers [PDB codes 1vr9 (Miller *et al.*, 2004) and 1yav (Kumaran & Swaminathan, unpublished results)]. In both cases the domains pack together with the helices corre-

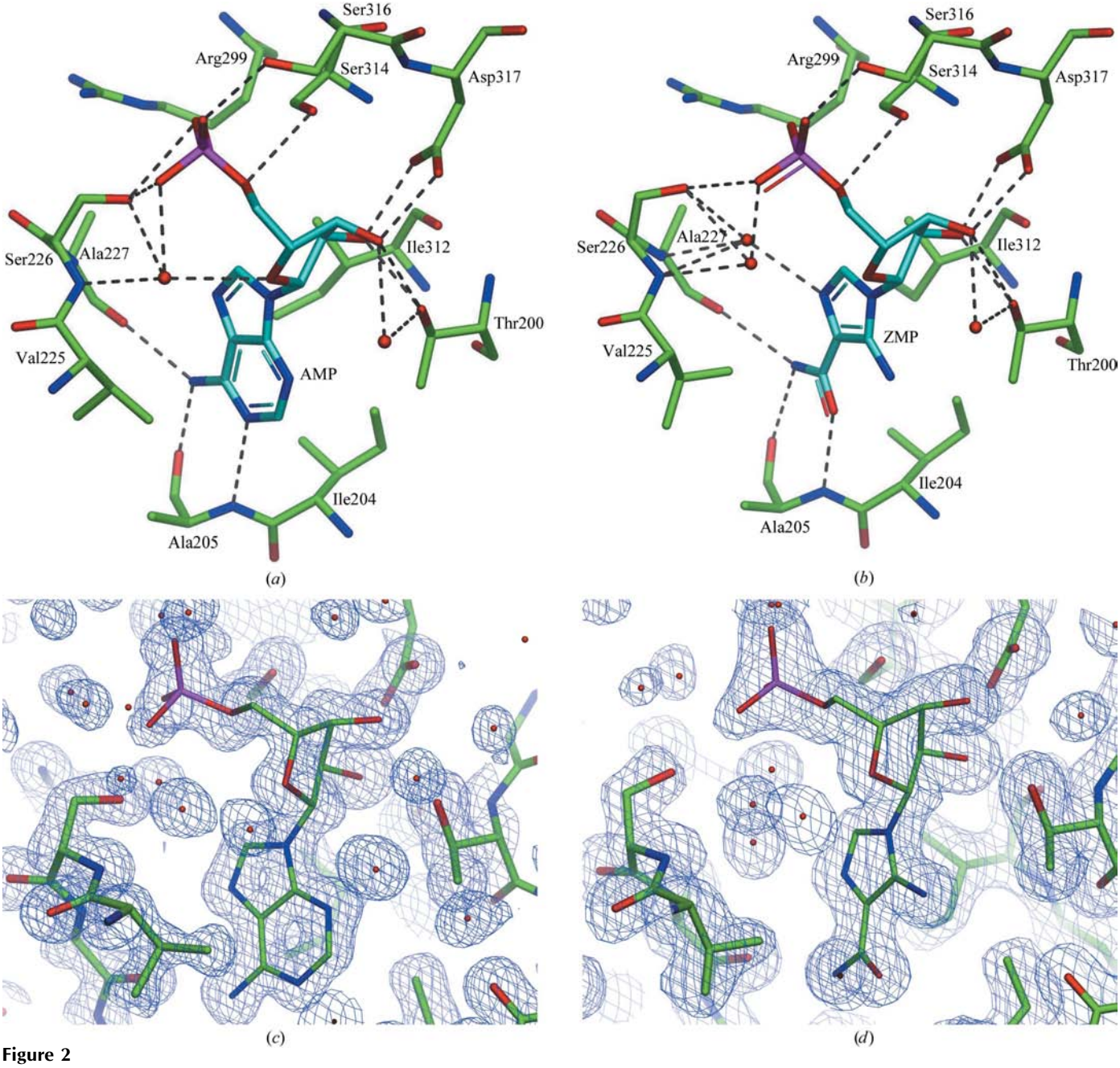

Figure 2
Comparison of AMP and ZMP binding to CBS3+4. The upper panels illustrate protein interactions with AMP (*a*) and ZMP (*b*). Ligands are shown in cyan. The lower panel shows representative $2|F_o| - |F_c|$ density contoured at the 1σ level for the AMP (*c*) and ZMP (*d*) complexes.

sponding to $\alpha2$, $\alpha3$, $\alpha6$ and $\alpha7$ in the interface, but form dimers that differ from those presented here.

3.3. Ligand binding

Difference density could be clearly seen and unambiguously interpreted for both the AMP and ZMP complexes. The difference maps from the crystal grown in the presence of ATP showed clear density for the AMP moiety of the nucleotide; however, the β and γ phosphates could not be located unambiguously in the maps. This may be a consequence of disorder in this part of the molecule, as the phosphates lie very close to the crystallographic symmetry axis.

The nucleotide-binding site sits between adjacent CBS domains and is formed from strands $\beta1$ and $\beta7$ of CBS4 and $\beta2$ and $\beta3$ of CBS3 (Figs. 1*a* and 2*a*). AMP binds in a pocket lined by the side chains of residues Thr200, Ile204, Val225 and Ile312 (Figs. 2*a* and 2*c*). N-6 of the adenine ring is hydrogen bonded to the main-chain carbonyl O atoms of Ala205 and Ala227 and N-1 to the main-chain amide N atom of Ala205. N-3 is hydrogen bonded to a water molecule. The ribose moiety forms several hydrogen bonds to side chains. The 2′-OH hydrogen bonds to Thr200, the 5′-O to Ser314 and the 2′-OH and 3′-OH form charged hydrogen bonds to the side chain of Asp317. The 4′-O hydrogen bonds to the amide N atom of Ser226 *via* a water molecule. The phosphate O atoms hydrogen bond to the hydroxyls of Ser226, Ser314 and Ser316 directly and to the side chain of Arg299, the main-chain carbonyl O atom of residue 298 and the amide N atoms of Ser226 and Ser316 *via* water molecules. In addition to these intramolecular interactions, Lys243 forms a water-mediated interaction with a phosphate O atom of the symmetry-related AMP molecule and a direct hydrogen bond to the side chain of Ser316 of the symmetry-related protein molecule. Both Lys243 and Ser316 are highly conserved across γ chains.

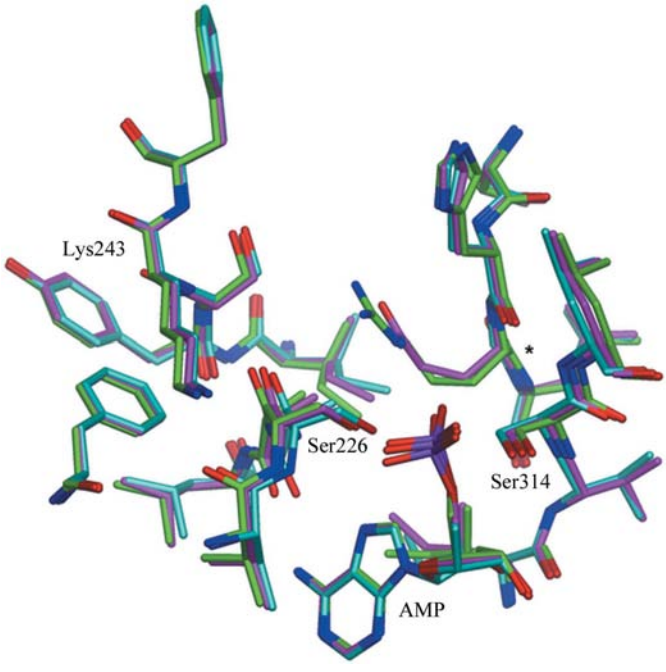

Figure 3
Structure of the AMP-binding site of His_6-$\gamma1$(182–325) superposed with those of the substituted variants Arg299→Gly and Arg299→Gln. The asterisk indicates the position of the C^{α} of residue 299. His_6-$\gamma1$(182–325) is shown in green, His_6-$\gamma1$(182–325, R299Q) is in magenta and His_6-$\gamma1$(182–325, R299G) is in cyan. Water molecules have been omitted for clarity.

The disease-causing substitution Arg531→Gly in AMPK $\gamma2$, the residue equivalent to Arg299 of $\gamma1$, results in a dramatic decrease in affinity for AMP and ATP consistent with the observation that this residue is involved in phosphate binding (Davies *et al.*, 2006; Daniel & Carling, 2002). Similarly, the substitution Arg299→Gln in AMPK $\gamma1$ abolishes the activation of AMPK by AMP (Scott *et al.*, 2004; Burwinkel *et al.*, 2005; Sanders *et al.*, 2007). The structures of His_6-$\gamma1$(182–325, R299G) and His_6-$\gamma1$(182–325, R299Q) were also solved at 2.0 Å resolution and showed that the substitutions had no significant structural effects other than at the site of substitution (Fig. 3), suggesting that the reduction in binding affinity for ATP and AMP was primarily a consequence of the loss of the positive charge on residue 299.

The ZMP-bound structure is very similar to that of the AMP complex (r.m.s.d. = 0.16 Å for all atoms), but there are a few notable changes in the ligand-binding site (Fig. 2*b*). Ile312 adopts an alternative side-chain rotamer and Val225 rotates, resulting in movement of the main chain of residues 225 and 226. This also results in movement of the side-chain hydroxyl of Ser226 by 1.4 Å. The ZMP ring system is bound less (0.4 Å) deeply into the binding site than the adenine ring of AMP and, combined with the main-chain movements, this permits an additional water molecule to bind between the main-chain amide N atom of Ala227 and N-3 of ZMP. These changes were not observed when adenosine was bound (data not shown), suggesting that they may be specific for ZMP and may explain the lower affinity of AMPK for ZMP relative to AMP.

4. Discussion

AMPK $\gamma1$ (CBS3+4) adopts the fold expected for a CBS-domain pair. The hydrophobic face of the small amphipathic N-terminal helix (residues 187–189) packs against the hydrophobic face of the C-terminal helix (residues 315–322). This interaction helps to rationalize the observation that the minimal soluble constructs, containing residues 185–325, keep both of these helices intact. Truncation into these helices would not only expose buried hydrophobic residues but would also disrupt the ligand-binding site which involves a number of residues on the C-terminal helix.

The structure presented suggests that residues 311–317 of CBS4 may form a novel recognition motif (G*h*x[S/T]x[S/T]D, where *x* is any amino acid and *h* is hydrophobic) for the ribose-phosphate moiety. Sequence analysis showed that this motif is conserved within CBS4 of γ chains from many species. Analysis of the full $\gamma1$ sequence revealed the presence of a second motif on CBS1 (residues 84–91), suggesting that nucleotides may bind in a symmetry-related binding pocket of the CBS1+2 pair, consistent with modelling studies (Adams *et al.*, 2004).

Examination of the sequence of the CBS-domain pair from human IMPDH II, which is known to bind ATP (Scott *et al.*, 2004), shows that this motif is present in the first CBS domain, suggesting that it may resemble a CBS1+2-like module (Fig. 4). In the structure of human IMPDH the CBS domains are poorly ordered (Colby *et al.*, 1999); however, they are well defined in the structure of a bacterial IMPDH (Zhang *et al.*, 1999). Superposition of the CBS domains from bacterial IMPDH with the AMP complex described here (Fig. 5) shows that there is good agreement between the binding sites, suggesting that only minimal protein movement in IMPDH would be required for nucleotide binding. Although the CBS-domain pair from IMPDH has motifs on both sides of the putative nucleotide-binding cleft, only one of these could easily accommodate ATP without significant protein movements. This putative binding site aligns with the motif observed for human IMPDH and is thus on the opposite side of the binding cleft with respect to that of the CBS3+4 domain pair of AMPK.

The CBS-domain pair from cystathionine β-synthase has potential motifs on both CBS domains. The substitution of Asp444, a residue in the first CBS domain of cystathionine β-synthase, with Asn causes homocysteineurea owing to reduced affinity for the allosteric activator *S*-adenosylmethionine (Kluijtmans *et al.*, 1996; Scott *et al.*, 2004). Conventionally, residue Asp444 has been considered to be structurally equivalent to Arg70 of AMPK γ1 (Scott *et al.*, 2004; Ignoul & Eggermont, 2005), assuming that all CBS-domain pairs bind nucleotides in the same pocket (Fig. 4). Alternatively, the CBS domains from cystathionine β-synthase can be aligned with the CBS3+4 domain pair (Fig. 4) such that Asp444 aligns with Ser226 of AMPK γ1, a residue that forms a direct hydrogen bond to the phosphate of AMP and would be ideally positioned to interact with both of the positive charges on *S*-adenosylmethionine.

During preparation of this manuscript, the structure of a CBS-domain pair from CLC-5 with bound ATP was published (Meyer *et al.*, 2007). The ligand binds in the same binding pocket that we describe for the AMPK CBS3+4 domain pair (Fig. 6). As in our structure, the adenine moiety is flanked by hydrophobic residues, but is stacked against the side chain of a tyrosine residue (Tyr617) that is equivalent to Val225 of AMPK. N-7 of the adenine ring is hydrogen bonded to the main-chain amide N atom of residue 617 of CLC-5. This interaction is absent in the AMPK γ1–AMP complex, but is equivalent to the interaction seen between N-3 of ZMP and the amide N atom of residue 227. In both structures the hydroxyl groups of the ribose moiety are coordinated by the side chain of the aspartyl residue that is part of the conserved binding motif described above. In addition, there is a further hydrogen bond to Thr200 in AMPK γ1. The binding of the α-phosphate is significantly different in the two structures. The α-phosphate of ATP forms a hydrogen bond to the main-chain amide N atom of Ser618 in CLC-5. The equivalent residue in AMPK γ1, Ser226, is rotated such that it forms a hydrogen bond *via* the side-chain hydroxyl group and the main-chain amide N atom of the same residue forms an indirect hydrogen bond to the α-phosphate *via* a water molecule. The α-phosphate is also coordinated by the side chains of two other serine residues (Ser314 and Ser316) in the AMPK γ1–AMP complex. Thus, although the ligands bind in an equivalent pocket in both structures, the specifics of the protein–ligand interactions are different. Small changes in the position of the bound adenine moiety are propagated through the ligand such that the positions of the

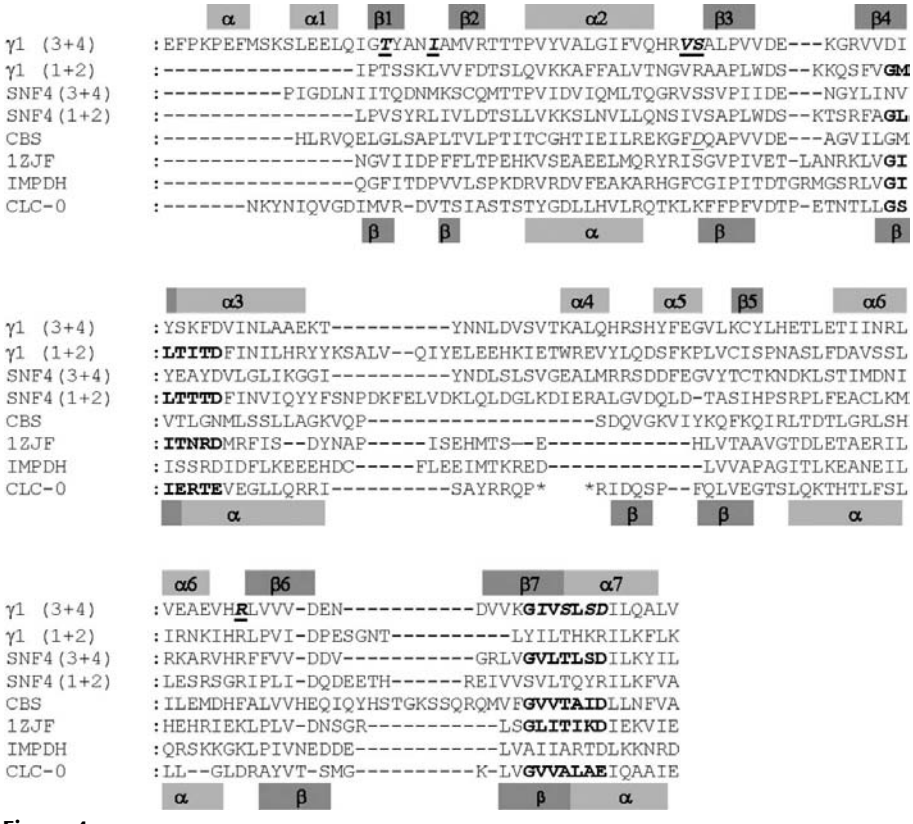

Figure 4
Alignment of AMPK γ1 (UniProt entry P54619; CBS3+4, amino acids 182–323; CBS1+2, amino acids 42–177), SNF4 (*Saccharomyces cerevisiae*, UniProt entry P12904; CBS3+4, amino acids 187–318; CBS1+2, amino acids 35–175), human IMPDH II (UniProt entry P12268, amino acids 112–232), bacterial IMPDH (PDB code 1zfj, UniProt entry Q8K5G1, amino acids 95–211), CBS (UniProt entry P35520, amino acids 416–543) and CLC-0 (*Torpedo marmorata*, GenBank X56758, amino acids 536–606 and 718–769, break in sequence indicated by *). Residues involved in interactions with AMP in AMPK γ1 CBS3+4 are in underlined in bold. Residue 444 of CBS, which is substituted by asparagine in homocysteineurea, is shown in underlined italics. The G*hx*S/T*x*S/TD motif is in bold. Secondary-structural elements derived from the AMPK γ1 (182–325) structure (above) and CLC-0 structure (below) are shown.

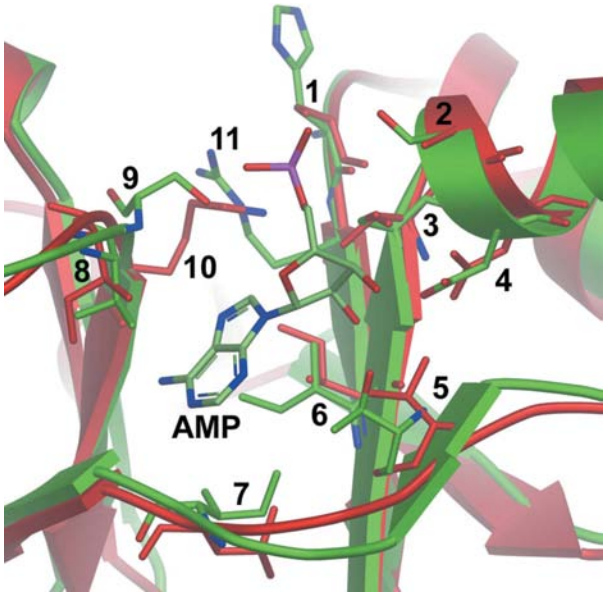

Figure 5
Superposition of AMPK γ CBS3+4 (green) and IMPDH CBS domains (PDB code 1zfj, red). AMP is from the AMPK structure. Highlighted residues are 1, His298 (Ser123); 2, Ser316 (Arg143); 3, Ser314 (Thr141); 4, Asp317 (Asp144); 5, Thr200 (Thr160); 6, Ile312 (Ile139); 7, Ile204 (Leu164); 8, Val225 (Ile185); 9, Ser226 (Glu186); 10, Ala227 (Lys187); 11, Arg299 (Gly124). Residues in parentheses are from IMPDH.

α-phosphate groups are significantly altered. It is perhaps of interest to note that we have solved the structure of AMPK γ1 CBS3+4 in complex with ATP and that in this complex Ser226 rotates such that it could hydrogen bond to either the α- or γ-phosphates of ATP (data not shown). Other than movement of this serine and Arg299, the ATP structure is essentially identical to that of the AMP complex and so we have not presented this structure, as the density for the β- and γ-phosphates of ATP is ambiguous.

The intermolecular interactions seen across the symmetry axis in the crystal structure as described encourage speculation that the observed crystallographic dimer of AMPK γ1 CBS3+4 may be a physiological dimer. This is supported by the considerable surface area buried and by the fact that interactions with the AMP molecule come from both subunits in the crystallographic dimer. However, as the present structure encompasses only a fragment of the AMPK heterotrimer it cannot be ruled out that CBS3+4 might interact with either CBS1+2 or with one of the other subunits in the full complex. The significance of this in terms of the regulatory role of the γ chain and the general regulatory role of CBS domains remains to be determined.

The functional role of CBS domains in the regulation of a variety of proteins is gradually being elucidated and this has been facilitated by the demonstration that many of these

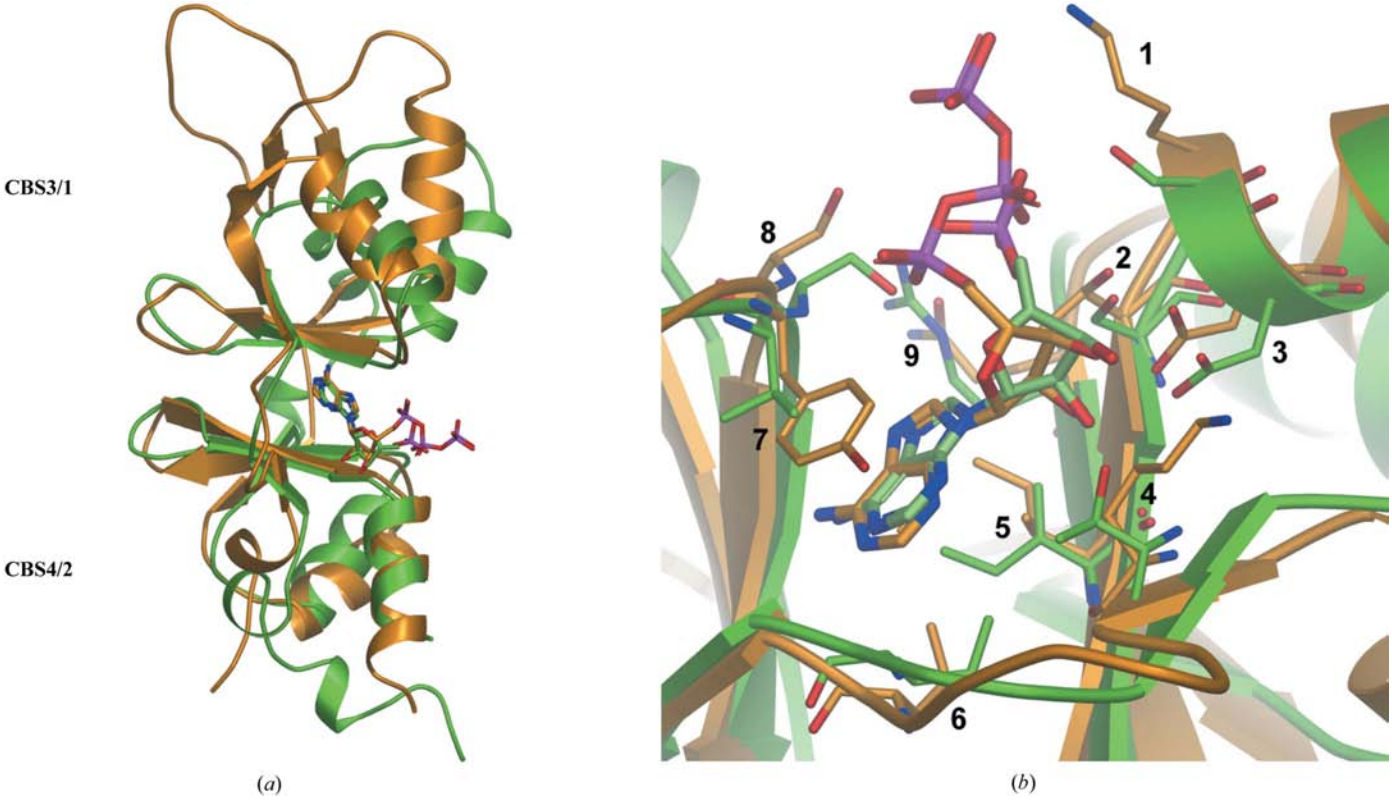

Figure 6
Superposition of AMPK γ CBS3+4 (green) and CLC-5 CBS domains (PDB code 2j9l, orange) with bound AMP and ATP, respectively. (*a*) Superposition of AMPK γ CBS3+4 and CLC-5 CBS1+2 secondary structure. (*b*) Overlay of nucleotide-binding site residues. Highlighted residues are 1, Ser316 (Lys726); 2, Ser314 (Thr724); 3, Asp317 (Asp727); 4, Thr200 (Lys587); 5, Ile312 (Ile722); 6, Ile204 (Leu595); 7, Val225 (Tyr617); 8, Ser226 (Ser618); 9, Arg299 (Gln710). Residues in parentheses are from CLC-5.

domains bind adenine nucleotides. The structures presented here define the residues involved in the binding of adenine nucleotides by the AMPK γ1 CBS3+4 domain pair and will increase the accuracy of models of other ligand-binding CBS pairs including AMPK γ1 CBS1+2. As the pharmacological activator of AMPK, ZMP, has been shown in animal models to induce many of the desirable effects predicted for an AMPK activator, and as the ZMP complex structure has been determined here, it is reasonable to surmise that the present system will afford access to high-resolution structures with other candidate ligands. This should be of great value in structure-based design of agents activating AMPK at this site.

We acknowledge ESRF Grenoble for access to beam time and technical assistance. The authors declare that they have no competing financial interest.

References

Adams, J., Chen, Z., Van Denderen, B. J. W., Morton, C. J., Parker, M. W., Witters, L. A., Stapleton, D. & Kemp, B. E (2004). *Protein Sci.* **13**, 155–165.

Bateman, A. (1997). *Trends Biochem. Sci.* **22**, 12–13.

Bergeron, R., Previs, S. F., Cline, G. W., Perret, P., Russell, R. R. III, Young, L. H. & Shulman, G. I. (2001). *Diabetes*, **50**, 1076–1082.

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.

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.

Buhl, E. S., Jessen, N., Pold, R., Ledet, T., Flyvbjerg, A., Pedersen, S. B., Pedersen, O., Schmitz, O. & Lund, S. (2002). *Diabetes*, **51**, 2199–2206.

Burwinkel, B., Scott, J. W., Buhrer, C., van Landeghem, F. K., Cox, G. F., Wilson, C. J., Hardie, D. G. & Kilimann, M. W. (2005). *Am. J. Hum. Genet.* **76**, 1034–1049.

Cheung, P. C. F., Saly, I. P., Davies, S. P., Hardie, D. G. & Carling, D. (2000). *Biochem. J.* **346**, 659–669.

Clarke, P. R. & Hardie, D. G. (1990). *EMBO J.* **9**, 2439–2446.

Colby, T. D., Vanderveen, K., Strickler, M. D., Markham, G. D. & Goldstein, B. M. (1999). *Proc. Natl Acad. Sci. USA*, **96**, 3531–3536.

Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.

Daniel, T. & Carling, D. (2002). *J. Biol. Chem.* **277**, 51017–51024.

Davies, S. P., Hawley, S. A., Woods, A., Carling, D., Haystead, T. A. & Hardie, D. G. (1994). *Eur. J. Biochem.* **223**, 351–357.

Davies, J. K., Wells, D. J., Liu, K., Whitrow, H. R., Daniel, T., Grignani, R., Lygate, C. A., Schneider, J. E., Noël, G., Watkins, H. & Carling, D. (2006). *Am. J. Physiol. Heart Circ. Physiol.* **290**, H1942–H1951.

Fryer, L. G. D., Parbu-Patel, A. & Carling, D. (2002). *J. Biol. Chem.* **277**, 25226–25232.

Hardie, D. G., Carling, D. & Carlson, M. (1998). *Annu. Rev. Biochem.* **67**, 821–855.

Hardie, D. G. & Hawley, S. A. (2001). *Bioassays*, **23**, 1112–1119.

Hardie, D. G., Scott, J. W., Pan, D. A. & Hudson, E. R. (2003). *FEBS Lett.* **546**, 113–120.

Hawley, S. A., Boudeau, J., Reid, J. L., Mustard, K. J., Udd, L., Makela, T. P., Alessi, D. R. & Hardie, D. G. (2003). *J. Biol.* **2**, 28.

Hawley, S. A., Gadalla, A. E., Olsen, G. S. & Hardie, D. G. (2002). *Diabetes*, **51**, 2420–2425.

Holmes, B. F., Kurth-Kraczek, E. J. & Winder, W. W. (1999). *J. Appl. Physiol.* **87**, 1990–1995.

Hudson, E. R., Pan, D. A., James, J., Lucocq, J. M., Hawley, S. A., Green, K. A., Baba, O., Terashima, T. & Hardie, D. G. (2003). *Curr. Biol.* **13**, 861–866.

Iglesias, M. A., Ye, J. M., Frangioudakis, G., Saha, A. K., Tomas, E., Ruderman, N. B., Cooney, G. J. & Kraegen, E. W. (2002). *Diabetes*, **51**, 2886–2894.

Ignoul, S. & Eggermont, J. (2005). *Am. J. Physiol. Cell Physiol.* **289**, C1369–C1378.

Kahn, B. B., Alquier, T., Carling, D. & Hardie, D. G. (2005). *Cell Metab.* **1**, 15–25.

Kemp, B. E., Stapleton, D., Campbell, D. J., Chen, Z. P., Murthy, S., Walter, M., Gupta, A., Adams, J. J., Katsis, F., van Denderen, B., Jennings, I. G., Iseli, T., Michell, B. J. & Witters, L. A. (2003). *Biochem. Soc Trans.* **31**, 162–168.

Kennan, A., Aherne, A., Palfi, A., Humphries, M., McKee, A., Stitt, A., Simpson, D. A., Demtroder, K., Orntoft, T., Ayuso, C., Kenna, P. F., Farrar, G. J. & Humphries, P. (2002). *Hum. Mol. Genet.* **11**, 547–557.

Kluijtmans, L. A., Boers, G. H., Stevens, E. M., Renier, W. O., Kraus, J. P., Trijbels, F. J., van den Heuvel, L. P. & Blom, H. J. (1996). *J. Clin. Invest.* **98**, 285–289.

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.

Kurth-Kraczek, E. J., Hirshman, M. F., Goodyear, L. J. & Winder, W. W. (1999). *Diabetes*, **48**, 1667–1671.

La Fortelle, E. de & Bricogne, G. (1997). *Methods Enzymol.* **276**, 472–494.

Leslie, A. G. W. (1992). *Jnt CCP4/ESF–EACBM Newsl. Protein Crystallogr.* **26**.

Lloyd, S. E., Gunther, W., Pearce, S. H., Thomson, A., Bianchi, M. L., Bosio, M., Craig, I. W., Fisher, S. E., Scheinman, S. J., Wrong, O., Jentsch, T. J. & Thakker, R. V. (1997). *Hum. Mol. Genet.* **6**, 1233–1239.

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.

Milan, D., Jeon, J. T., Looft, C., Amarger, V., Robic, A., Thelander, M., Rogel-Gaillard, C., Paul, S., Iannuccelli, N., Rask, L., Ronne, H., Lundstrom, K., Reinsch, N., Gellin, J., Kalm, E., Roy, P. L., Chardon, P. & Andersson, L. (2000). *Science*, **288**, 1248–1251.

Miller, M. D. *et al.* (2004). *Proteins*, **57**, 213–217.

Muoio, D. M., Seefeld, K., Witters, L. A. & Coleman, R. A. (1999). *Biochem. J.* **338**, 783–791.

Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.

Nayak, V., Zhao, K., Wyce, A., Schwartz, M. F., Wan-Sheng, L., Berger, S. L. & Marmorstein, R. (2006). *Structure*, **14**, 477–485.

Neumann, D., Woods, A., Carling, D., Wallimann, T. & Schlatter, U. (2003). *Protein Expr. Purif.* **30**, 230–237.

Park, H., Kaushik, V. K., Constant, S., Prentki, M., Przybytkowski, E., Ruderman, N. B. & Saha, A. K. (2002). *J. Biol. Chem.* **277**, 32571–32577.

Perrakis, A., Morris, R. M. & Lamzin, V. S. (1993). *Nature Struct. Biol.* **6**, 458–463.

Polekhina, G., Gupta, A., Michell, B. J., van Denderen, B., Murthy, S., Feil, S. C., Jennings, I. G., Campbell, D. J., Witters, L. A., Parker, M. W., Kemp, B. E. & Stapleton, D. (2003). *Curr. Biol.* **13**, 867–871.

Polekhina, G., Gupta, A., van Denderen, B. J. W., Feil, S. C., Kemp, B. E., Stapleton, D. & Parker, M. W. (2005). *Structure*, **13**, 1453–1462.

Pusch, M. (2002). *Hum. Mutat.* **19**, 423–434.

Rudolph, M. J., Amodeo, G. A., Bai, Y. & Tong, L. (2005). *Biochem. Biophys. Res. Commun.* **337**, 1224–1228.

Saha, A. K., Avilucea, P. R., Ye, J. M., Assifi, M. M., Kraegen, E. W. & Ruderman, N. B. (2004). *Biochem. Biophys. Res. Commun.* **314**, 580–585.

Salt, I. P., Celler, J. W., Hawley, S. A., Prescott, A., Woods, A., Carling, D. & Hardie, D. G. (1998). *Biochem. J.* **334**, 177–187.

Sanders, M. J., Grondin, P. O., Hegarty, B. D., Snowden, M. A. & Carling, D. (2007). In the press.

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.

Shaw, R. J., Kosmatka, M., Bardeesy, N., Hurley, R. L., Witters, L. A., DePinho, R. A. & Cantley, L. C. (2004). *Proc. Natl Acad. Sci. USA*, **101**, 3329–3335.

Song, X. M., Fiedler, M., Galuska, D., Ryder, J. W., Fernstrom, M., Chibalin, A. V., Wallberg-Henriksson, H. & Zierath, J. R. (2002). *Diabetologia*, **45**, 56–65.

Woods, A., Johnstone, S. R., Dickerson, K., Leiper, F. C., Fryer, L. G. D., Neumann, D., Schlattner, U., Walliman, T., Carlson, M. & Carling, D. (2003). *Curr. Biol.* **13**, 2004–2008.

Zhang, R., Evans, G., Rotella, F. J., Westbrook, E. M., Beno, D., Huberman, E., Joachimiak, A. & Collart, F. R. (1999). *Biochemistry*, **38**, 4691–4700.

Zhou, G., Myers, R., Li, Y., Chen, Y., Shen, X., Fenyk-Melody, J., Wu, M., Ventre, J., Doebber, T., Fujii, N., Musi, N., Hirshman, M. F., Goodyear, L. J. & Moller, D. E. (2001). *J. Clin. Invest.* **108**, 1167–1174.