Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structural characterization of the ribonuclease H-like type ASKHA superfamily kinase MK0840 from *Methanopyrus kandleri*

Magdalena Schacherl,[a,b] Sandro Waltersperger[b,c] and Ulrich Baumann[a,b]*

[a]Institute of Biochemistry, University of Cologne, Otto-Fischer-Strasse 12-14, 50674 Cologne, Germany, [b]Department of Chemistry and Biochemistry, University of Bern, Freiestrasse 3, 3012 Bern, Switzerland, and [c]Swiss Light Source, Paul Scherrer Institute, 5232 Villigen, Switzerland

Correspondence e-mail: ubaumann@uni-koeln.de

Received 25 April 2013
Accepted 12 August 2013

PDB References:
MK0840$^{apo\text{-}m}$, 4bg8;
MK0840$^{apo\text{-}o}$, 4bg9;
MK0840–ADP^{open}, 4bga;
MK0840–ADP^{closed}, 4bgb

Murein recycling is a process in which microorganisms recover peptidoglycan-degradation products in order to utilize them in cell wall biosynthesis or basic metabolic pathways. Methanogens such as *Methanopyrus kandleri* contain pseudomurein, which differs from bacterial murein in its composition and branching. Here, four crystal structures of the putative sugar kinase MK0840 from *M. kandleri* in apo and nucleotide-bound states are reported. MK0840 shows high similarity to bacterial anhydro-*N*-acetylmuramic acid kinase, which is involved in murein recycling. The structure shares a common fold with panthothenate kinase and the 2-hydroxyglutaryl-CoA dehydratase component A, both of which are members of the ASKHA (acetate and sugar kinases/Hsc70/actin) superfamily of phosphotransferases. Local conformational changes in the nucleotide-binding site between the apo and holo forms are observed upon nucleotide binding. Further insight is given into domain movements and putative active-site residues are identified.

1. Introduction

Peptidoglycan, or murein, is responsible for the shape and the stability of the eubacterial cell and exhibits various biological activities (Stewart-Tull, 1980; Oken *et al.*, 1981; Atrih & Foster, 1999). It is targeted by small molecules such as β-lactam antibiotics, which block the enzymes responsible for cell-wall synthesis, for example transpeptidases and transglycosylases. Murein consists of two amino sugars, *N*-acetylglucosamine (GlcNAc) and *N*-acetylmuramic acid (MurNAc), connected by a β-(1,4)-glycosidic bond. In most Gram-positive bacteria MurNAc is attached to a four-to-five-residue amino-acid chain containing L-alanine, D-glutamine or D-isoglutamine, L-lysine and D-alanine (de Jonge *et al.*, 1992, 1996). In the archaebacterial kingdom, methanogenic bacteria such as *Methanopyrus kandleri* replace murein with pseudomurein consisting of *N*-acetyl-D-galactosamine (GalNAc) and *N*-acetyl-L-talosaminuronic acid (TalANAc) linked by a β-(1,3)-glycosidic bond and containing only L-amino acids such as L-alanine, L-lysine, L-glutamic acid and L-ornithine (Kurr *et al.*, 1991).

Bacteria break down their murein components to synthesize the cell wall of the next generation (Goodell & Schwarz, 1985). One of the steps involves anhydro-*N*-acetylmuramic acid kinase (AnmK), which converts 1,6-anhydro-*N*-acetylmuramic acid to *N*-acetylmuramic acid 6-phosphate with the simultaneous cleavage of the 1,6-anhydro ring and hydrolysis of ATP (Uehara *et al.*, 2005). As the biosynthetic pathways of pseudomurein and murein differ (Hartmann & König, 1990), no homologous enzymes or similar pathways for pseudomurein recycling have been identified in the genomes of microorganisms to date (Scheffers & Pinho, 2005; Claus & König, 2010). The methanogenic euryarchaeote *M. kandleri*

Table 1
Data-collection and refinement statistics.

Values in parentheses are for the outermost shell.

Structure/data set	Se-SAD peak	MK0840$^{apo\text{-}m}$	MK0840$^{apo\text{-}o}$	MK0840–ADP^{open}	MK0840–ADP^{closed}
Data collection					
Space group	$P2_1$		$P2_12_12$	$P2_1$	$P2_12_12_1$
Unit-cell parameters					
a (Å)	50.51		73.06	73.60	54.79
b (Å)	73.26		116.00	111.70	108.54
c (Å)	116.33		48.38	87.90	117.35
$\alpha = \gamma$ (°)	90		90	90	90
β (°)	95.92		90	89.55	90
No. of monomers per asymmetric unit	2	2	1	2	2
Wavelength (Å)	0.9792	0.9870	1.0000	1.0000	1.0000
Resolution (Å)	61.9–2.40 (2.55–2.40)	29.5–1.96 (2.07–1.96)	61.8–1.90 (2.02–1.90)	73.6–2.60 (2.76–2.60)	45.2–1.34 (1.37–1.34)
No. of observations	137911 (21768)	215899 (31269)	153123 (23725)	335007 (54893)	696252 (42888)
No. of unique reflections	63610 (10158)	60796 (8807)	33085 (5163)	43755 (7132)	154630 (10722)
Multiplicity	2.16 (2.14)	3.55 (3.55)	6.45 (6.40)	6.10 (6.13)	4.50 (4.00)
Completeness (%)	98.4 (97.7)	99.7 (99.3)	99.5 (97.5)	100.0 (100.0)	98.3 (93.2)
R_{merge}† (%)	7.1 (34.9)	6.2 (41.5)	8.6 (71.3)	8.2 (36.7)	3.5 (65.5)
$CC_{1/2}$‡ (%)	99.5 (81.8)	99.7 (89.1)	99.8 (73.1)	99.8 (94.8)	99.9 (86.8)
$\langle I/\sigma(I)\rangle$	10.59 (2.28)	13.30 (2.90)	14.61 (2.17)	21.95 (5.98)	21.12 (2.61)
Refinement statistics					
R_{work}/R_{free}§ (%)		17.7/20.9	17.9/20.9	15.4/20.8	15.7/18.3
No. of non-H protein atoms		4811	2397	9632	5062
No. of water molecules		326	152	222	519
No. of ions/heavy atoms		6 K^+, 2 Cl^-, 8 Se	3 K^+, 1 Cl^-, 4 Se	8 K^+, 3 Mg^{2+}	5 K^+, 4 Mg^{2+}, 1 Ca^{2+}
No. of other molecules		2 ACT	1 ACT	4 ADP, 2 SUC, 4 BGC	2 ADP, 7 GOL
No. of TLS groups/chains		5	5	—	8
Root-mean-square deviations					
Bond lengths (Å)		0.010	0.013	0.006	0.012
Bond angles (°)		1.192	1.375	1.181	1.482
Average B factor (Å^2)					
All protein atoms		35.1	27.5	33.2	25.2
Waters		42.1	31.6	32.5	35.9
Other atoms		35.9	26.6	47.8	26.7
Ramachandran plot¶, residues in (%)					
Most favoured region		98.7	98.0	98.5	99.4
Additionally allowed region		1.3	1.0	1.5	0.6
Disallowed region		0	0	0	0

† $R_{merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the ith measurement of the intensity of the unique reflection hkl and $\langle I(hkl)\rangle$ is the mean over all symmetry-related measurements. ‡ Percentage of correlation between intensities from random half data sets (Karplus & Diederichs, 2012). § $R_{work} = \sum_{hkl} \big||F_{obs}| - |F_{calc}|\big| / \sum_{hkl} |F_{obs}|$. R_{free} was calculated in the same way but using a randomly selected 5% set of reflections that were not used in refinement (Brünger, 1997). ¶ Calculated using *MolProbity* (Chen *et al.*, 2010).

encodes a putative protein (UniProtKB entry Q8TX37; gene MK0840) that is similar to anhydro-*N*-acetylmuramic acid kinase and is annotated as a chaperone distantly related to Hsp70-fold metalloproteases. As archaebacteria lack *N*-acetylmuramic acid, replacing it with *N*-acetyltalosaminuronic acid, it is unlikely that they contain genes encoding the enzymes responsible for *N*-acetylmuramic acid synthesis or degradation; rather, they would encode enzymes for the pathways involved in the metabolism of *N*-acetyltalosaminuronic acid.

There is no structural information available for the enzymes of pseudomurein metabolic pathways. In order to gain a deeper insight into the mechanistic aspects, we have solved the crystal structure of MK0840 in different ligand-bound states and conformations and report the results here.

2. Materials and methods

2.1. Cloning and expression

A synthetic DNA fragment coding for the full-length putative sugar kinase from *M. kandleri* (MK0840, Q8TX37) and optimized for *Escherichia coli* expression (MrGene) was cloned into the pET28a vector (Novagen) with an N-terminal hexahistidine tag using the *Nde*I/*Xho*I restriction sites. Overexpression was carried out in *E. coli* BL21 (DE3) (Novagen) in LB medium overnight at 293 K. Selenomethionine-substituted protein was prepared using the metabolic inhibition method (Van Duyne *et al.*, 1993; Doublié, 1997), yielding approximately the same amount of protein as the wild type. A truncated construct starting from amino acid Leu37 (Leu37–Thr358) was cloned into the pET28a vector and overexpressed as described for the full-length construct.

2.2. Purification

Both full-length and truncated proteins were purified by immobilized metal-affinity chromatography (IMAC) using Ni–NTA Superflow resin (Qiagen) in a buffer consisting of 20 m*M* Tris–HCl pH 7.5, 300 m*M* NaCl with 10–70 m*M* imidazole for the washing step or 250 m*M* imidazole for the elution steps. The hexahistidine tag was cleaved off by thrombin (Sigma–Aldrich) using 1 U thrombin per milligram

of protein during overnight dialysis against IMAC buffer lacking imidazole. The digested protein was again passed over an Ni–NTA column and subsequently further purified by size-exclusion chromatography using a Superdex 200 16/60 column (GE Healthcare) and a buffer consisting of 20 m*M* Tris–HCl pH 7.5, 200 m*M* NaCl, 0.02% sodium azide and 4 m*M* DTT (or 10 m*M* DTT for selenomethionine-substituted protein). Fractions containing dimeric protein were pooled and were concentrated to 6 mg ml^{-1}. The yield of pure protein was 6 mg per litre of culture.

2.3. Crystallization and data collection

Initial crystallization screening was carried out using the sitting-drop vapour-diffusion method at 293 K. As the apo crystals (MK0840^{apo}) obtained from full-length SeMet-substituted MK0840 appeared overnight and were highly intergrown, several optimization procedures were carried out to slow down crystal growth. Finally, a very small number of crystals of diffraction quality could be obtained from sitting-drop experiments with 0.2% low-melting agarose (Hampton Research) and the initial condition 0.5 *M* potassium thiocyanate, 0.1 *M* sodium acetate pH 4.6 at 293 K with poor reproducibility. Small crystals growing at the very rim of the drop were seeded into the same drop by crushing with a thin needle. Only this procedure yielded thin and fragile plates with dimensions of 5 × 50 × 150 µm and a single crystal lattice. For data collection, crystals were cryoprotected with the original mother-liquor solution supplemented with 40% sucrose (water replaced by sucrose) and were flash-cooled in liquid nitrogen. Diffraction data for selenium single-wavelength anomalous dispersion (Se-SAD) were collected under cryogenic conditions on beamline X06DA at the Swiss Light Source (PSI, Villigen, Switzerland) using a MAR225 detector. The wavelength for the Se-SAD peak was 0.9792 Å and that for the remote was 0.9870 Å (data set MK0840$^{apo\text{-}m}$; monoclinic space group). An additional data set from a second MK840^{apo} crystal was collected to give higher resolution data (data set MK0840$^{apo\text{-}o}$; orthorhombic space group). Needle-shaped crystals (MK0840–ADP^{open}) with dimensions of 30 × 30 × 600 µm were obtained from the N-terminally truncated construct (Leu37–Thr358) supplemented with 10 m*M* $MgCl_2$ and 5 m*M* ADP (Sigma–Aldrich) in a condition consisting of 0.2 *M* potassium sodium tartrate, 20% PEG 3350. Crystals were cryoprotected by soaking them in the original mother liquor supplemented with 20% sucrose for 30 s and were flash-cooled in liquid nitrogen. Bar-shaped crystals (MK0840–ADP^{closed}) with dimensions of 40 × 40 × 200 µm were also obtained for the N-terminally truncated construct (Leu37–Thr358) supplemented with 10 m*M* $MgCl_2$, 75 m*M* KCl, 5 m*M* ADP (Sigma–Aldrich) in a cryocondition consisting of 0.14 *M* $CaCl_2$, 0.07 *M* sodium acetate pH 4.6, 10% 2-propanol, 28% glycerol. Native data sets (data sets MK0840–ADP^{open} and MK0840–ADP^{closed}) for molecular replacement were collected from both crystals at a wavelength of 1 Å on beamline X06DA equipped with a MAR225 CCD detector. All data were indexed and processed using the *XDS* software package (Kabsch, 2010). For MK0840–ADP^{closed} a low-resolution pass to 2.77 Å resolution and a high-resolution pass to 1.34 Å resolution were collected from the same crystal and were merged using *XSCALE* (Kabsch, 2010). The merged data set was used for structure determination and refinement. Data-collection and refinement statistics are summarized in Table 1.

2.4. Structure determination and refinement

The MK0840^{apo} structure was solved by Se-SAD (Se-SAD peak data set) using phases computed from eight selenium sites by the *phenix.autosol* routine of the *PHENIX* package (Adams *et al.*, 2010). The initial model was built with the *ARP/wARP* web server (Morris *et al.*, 2002; Langer *et al.*, 2008) and was refined using iterative cycles of *phenix.refine* (Adams *et al.*, 2010). Manual model building was accomplished using *Coot* (Emsley *et al.*, 2010) using the remote data to 1.96 Å resolution (data set MK0840$^{apo\text{-}m}$; space group $P2_1$). For cross-validation, a randomly selected subset (5%) of reflections was set aside in order to monitor the progress of refinement using the free *R* factor (Brünger, 1997). The coordinates were submitted to the *TLS Motion Determination* (*TLSMD*) server (Painter & Merritt, 2006*b*) in order to resolve the optimal number of TLS groups for subsequent use in TLS refinement (Painter & Merritt, 2006*a*) with *phenix.refine*. Data from a second apo crystal (MK0840$^{apo\text{-}o}$; space group $P2_12_12$) grown in the same drop as the $P2_1$ SeMet crystal (MK0840$^{apo\text{-}m}$) could be collected to 1.9 Å resolution. The structure was solved by molecular replacement using the program *Phaser* (McCoy *et al.*, 2007) in the *CCP*4 program suite (Winn *et al.*, 2011) with one monomer of MK0840$^{apo\text{-}m}$ as the search model. For both apo structures, anomalous scattering corrections (f'/f'') for the Se atoms were included during refinement.

The MK0840–ADP^{open} and MK0840–ADP^{closed} structures were determined by molecular replacement with *Phaser* using one monomer of the refined MK0840$^{apo\text{-}m}$ structure as a search model for MK0840–ADP^{open} and one monomer of MK0840–ADP^{open} for MK0840–ADP^{closed}. As one of the two protomers in the MK0840–ADP^{closed} structure differed substantially from the model used, only a partial solution was found. The missing protomer was built by *ARP/wARP*. The nucleotide-bound structures were refined as described for MK0840$^{apo\text{-}m}$, but for MK0840–ADP^{open} TLS refinement did not improve the model as judged by an increase in R_{free} and a larger gap between R_{work} and R_{free}.

The final structures were validated using *MolProbity* (Chen *et al.*, 2010), *WHATCHECK* (Vriend, 1990) and the corresponding tools in *Coot*. Secondary-structure elements were assigned using *DSSP* (Kabsch & Sander, 1983). Structural alignment was performed with *STRAP* (Gille & Frömmel, 2001), sequence alignment was performed using *ClustalW*2 (Larkin *et al.*, 2007) and alignment visualization was performed using *ESPript* (Gouet *et al.*, 1999). Figures were prepared with *PyMOL* (DeLano, 2002) and *LigPlot*$^+$ v.1.4.2 (Laskowski & Swindells, 2011).

2.5. Substrate modelling

Clear densities for ADP and additional ligands bound to the putative sugar-binding pocket were found during refinement of the MK0840–ADP^{open} and MK0840–ADP^{closed} structures. ADP molecules were modelled into the density using *phenix.ligandfit* (Terwilliger *et al.*, 2007). A ligand search was performed using electron-density map correlation with the program *phenix.ligand_identification* (Terwilliger *et al.*, 2006) to identify the molecule occupying the substrate-binding pocket of MK0840–ADP^{open}. The molecules with the best correlation coefficients (CC) were BGC (β-D-glucose; CC = 0.73), AMG (α-methyl-D-galactoside; CC = 0.71) and GAL (β-D-galactose; CC = 0.70). The electron density in the substrate-binding pocket was inspected and β-D-glucose was chosen. The sucrose in MK0840–ADP^{open} and the glycerol molecules in MK0840–ADP^{closed} were manually fitted into the electron density. Simulated-annealing $2F_o - F_c$ OMIT maps were calculated using the *RESOLVE* routine of *phenix.autobuild* (Terwilliger *et al.*, 2008).

2.6. Data deposition

The atomic coordinates and structure factors have been deposited in the Protein Data Bank (Berman *et al.*, 2000; http://www.pdb.org) with PDB codes 4bg8 for MK0840$^{apo\text{-}m}$, 4bg9 for MK0840$^{apo\text{-}o}$, 4bga for MK0840–ADP^{open} and 4bgb for MK0840–ADP^{closed}).

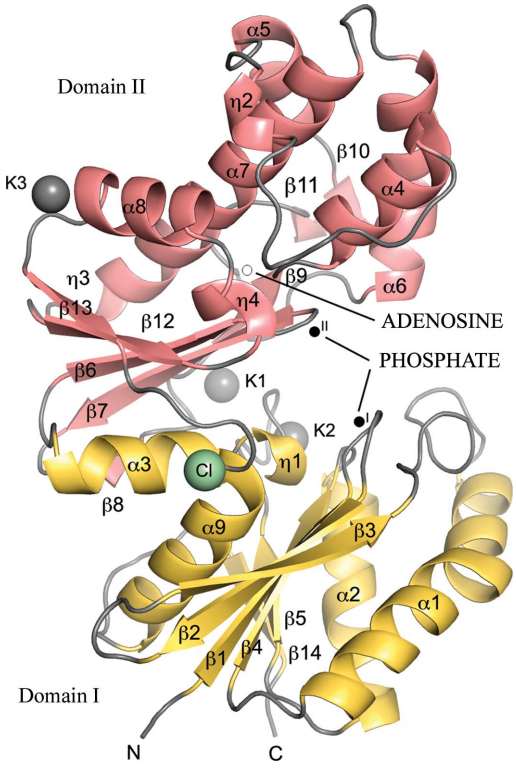


Figure 1
Overall structure of the MK0840^{apo} monomer. The N-terminal domain I is shown in yellow and the C-terminal domain II in light pink. The potassium ions (K1–K3) and the chloride ion (Cl) are shown as spheres. The adenosine and phosphate-binding regions (I and II) are indicated by open and solid circles, respectively.

3. Results and discussion

The following crystal structures were determined in this study: MK0840 without bound nucleotide (termed MK0840^{apo}) in two space groups and ADP-bound open (MK0840–ADP^{open}) and closed (MK0840–ADP^{closed}) forms. The structures were refined at diffraction limits of 2.60–1.34 Å with satisfactory statistics (Table 1).

3.1. Monomer structure

A database search for similar folds resulted in *Pseudomonas aeruginosa* 1,6-anhydro-*N*-acetylmuramic acid kinase (*Pa*AnmK; PDB entry 3qbw; Bacik *et al.*, 2011) and *Acidaminococcus fermentans* 2-hydroxyglutaryl-CoA dehydratase component A (HgdC; PDB entry 1hux; Locher *et al.*, 2001) as top hits as well as an uncharacterized protein from *Shewanella oneidensis* (*So*AnmK; PDB entry 3cqy; Midwest Center for Structural Genomics, unpublished work). Sequence identities for all of these matches were below 20% and the r.m.s. deviations of the superposed structures were 3 Å or greater for equivalent C^{α} atoms. Thus, the overall structure of the *M. kandleri* MK0840 monomer shows a typical ribonuclease H-like type ASKHA superfamily sugar kinase fold (Bork *et al.*, 1992) consisting of two α/β subdomains related to each other by approximate dyad symmetry (Fig. 1). Domain I comprises residues Leu37–Pro159 as well as Gly322–Thr358 and domain II contains residues Asn160–Ala321. The topology of domain I is of the form $\beta\beta\beta\alpha\beta\alpha\beta\alpha$ with the β-strands in the order 3–2–1–4–5 with strand β2 antiparallel to the others, with the very C-terminal strand β14 completing the β-sheet in an antiparallel manner, and helices α1, α2 and η1/α3 (where η represents a 3_{10}-helix) as well as helix α9.

Domain II contains strands β6–β13 in the order 9/8–7–6–12–13, with β7 antiparallel to the others, and β10–β11 as well as helices α4–α8 and η2–η4. The ATP-phosphate binding regions in the kinase fold were located in domain II (L^{45}GNT^{48} and PHOSPHATE I; V^{164}DVGA^{168} and PHOSPHATE II) as well as the adenosine-binding region (S^{296}GGGVKN^{302} and ADENOSINE) on helix η4 (Fig. 1 and Supplementary Fig. S1[1]); pattern identification is according to Bork *et al.* (1992).

MK0840–ADP^{open} binds sucrose molecules originating from the cryoprotectant solution at the contact interface of protomers *A* and *C* as well as that of protomers *B* and *D* (Fig. 3*c*). The residues involved in sucrose coordination are located on strand β3 of both protomers. Several hydrogen-bonded and nonbonded interactions are formed to hydroxyl groups of both the hexose and pentose of the sucrose.

All protomers in both of the MK0840^{apo} structures and the MK0840–ADP^{closed} structure and three of the four protomers in the MK0840–ADP^{open} structure lack electron density for the eight very C-terminal residues in the polypeptide chain. Only for one protomer of MK0840–ADP^{open} could the flexible C-terminus be modelled into the electron-density map owing

[1] Supplementary material has been deposited in the IUCr electronic archive (Reference: GM5025). Services for accessing this material are described at the back of the journal.

to its involvement in crystal contacts. The stretch between Gly351 and Thr358 contains four glycine residues, which probably makes it highly flexible in solution. During refinement and model building we were initially not able to locate the first 36 amino acids of full-length MK0840 in the electron density, indicating that these residues are disordered in the crystal. Secondary-structure and tertiary-structure prediction tools (Yang *et al.*, 2005; Slabinski *et al.*, 2007; Kelley & Sternberg, 2009), on the other hand, predicted an ordered N-terminus with residues Ala19–Thr31 in an α-helical conformation. By closer inspection of the electron density, residues Val23–Thr36 could be modelled into uninterpreted electron density. This part of the structure clearly displays an α-helical conformation, but the remaining part of the N-terminus is still missing. Interestingly, these first 36 residues are absent in other ASKHA family kinases (Supplementary Fig. S1). The function of the first 36 amino acids of MK0840 is unclear, but they seem to be an archaebacterial feature.

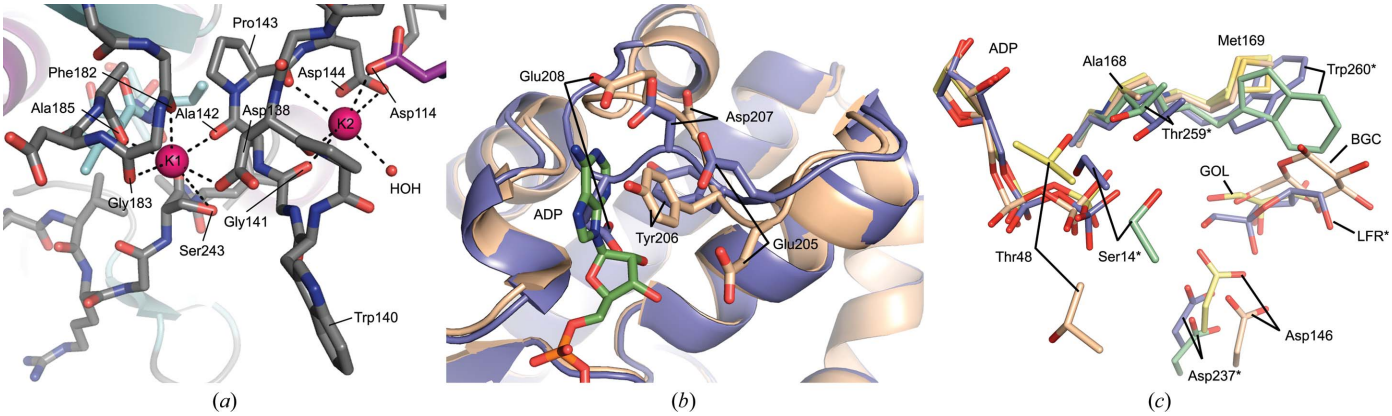
3.2. Ion-binding sites

During refinement, three potential cation-binding sites per monomer were identified in the MK0840^{apo} structure. Based on the type and the arrangement of the coordinating ligands (carbonyl O atoms and carboxylate groups of aspartate or glutamate residues as well as water molecules) and the observed distances (average 2.95 Å), K^+ was selected as the most likely cation. The refined *B* factors of the potassium ions match those of the surrounding protein atoms, indicating full occupancy. Furthermore, for all assigned K atoms anomalous difference signals were detected, which lends further credence to the assignment of these metal ions as potassium. As all crystallization conditions contained potassium salts, the bound K^+ ions most probably originate from the mother liquor.

Two of the three binding sites are present in all crystal forms. They are close to each other, with the ions (K1 and K2) being 7.8 Å apart and 'sandwiched' between three loop structures to the left and right of Trp140 (Fig. 2*a*), a residue that plays a role in dimer formation. The two potassium ions are located near the substrate-binding site, with K2 being approximately 7 Å away from the O2 atom of β-D-glucose in MK0840–ADP^{open} and 6 Å away from the O3 atom of glycerol in MK0840–ADP^{closed}. Strand β8 is interrupted at Pro180 and continues as β9 at Val186. Between them, the polypeptide chain (Asp181–Ala185) performs a turn in order to allow the coordination of the potassium ion K1, a feature that is not found in the structures of other ASKHA members (*e.g. E. coli* RhuK). There is a *cis*-peptide bond between Ala142 and Pro143. Both residues are located in the loop region between strand β5 and helix η1 and are involved in the coordination of potassium ions K1 and K2 (Fig. 2*a*). The *cis*-peptide bond is probably formed to allow bending of the polypeptide chain around K2. Residues Pro143 and Gly141 bind K2 with their backbone O atoms enclosing an angle of 82°. This is only made possible by the *cis* conformation of the Ala–Pro peptide bond. *cis*-Peptide bonds are known to occur in turns and near ion-binding sites (Pal & Chakrabarti, 1999).

In the two MK0840^{apo} structures the third potassium-binding site belongs to the 440 Å^2 contact interface between one monomer and a symmetry-equivalent monomer from another biological dimer (Fig. 3*b*) as determined by the *PISA* server (Krissinel & Henrick, 2007). Contrary to those of K1 and K2, this site appears to be generated by the crystal lattice and therefore may not be occupied in solution. An acetate ion is bound close to K3 which originates from the crystallization solution and interacts with the terminal amino groups N$^{\eta 2}$ of Arg307 and N$^{\eta 1}$ of Arg308 *via* its carboxyl group as well as with Glu311 *via* a hydrogen bond (Fig. 3*b*).

As mentioned above, the observed ions probably originate from the crystallization buffers and it is not known whether potassium is bound to the enzyme in the cellular environment. However, this appears to be likely owing to the full occupancy



(*a*) (*b*) (*c*)

Figure 2
Structural features of MK0840. (*a*) Potassium-binding sites K1 and K2 observed in both the MK0840^{apo} and the MK0840–ADP structures. The polypeptide chain is shown in stick representation. Potassium ions are depicted as pink spheres. Bonds to the potassium ions are shown as dashed black lines. (*b*) Superimposition of the nucleotide-binding region of MK0840^{apo} (blue) and MK0840–ADP^{open} (wheat). The ADP molecule and important side chains are shown in stick representation. (*c*) Superimposition of MK0840–ADP^{open} (wheat), MK0840–ADP^{closed} (yellow) and rhamnulose kinase RhuK in the apo state (green; PDB entry 2cgk) and with bound ADP and fructose (blue; PDB entry 2cgj) in the nucleotide-binding and substrate-binding region. Important side chains are shown in stick representation and asterisks indicates residues of RhuK. The carbohydrates β-L-fructose (LFR), β-D-glucose (BGC) and glycerol (GOL) are shown in stick representation. O atoms are coloured red, N atoms dark blue and S atoms yellow.

together with the general high intracellular potassium ion concentrations. Potassium ions have been reported to be important for the thermostability and activity of proteins in thermophilic archaea (Breitung *et al.*, 1992). It has been shown for proteins from *M. kandleri* that enzyme activity and stability at high temperatures are promoted by high concentrations of potassium salts of cyclic 2,3-diphosphoglycerate, 2,3-bisphosphoglycerate (both at 0.75 *M*) or phosphate (at 1.5 *M*) (Shima *et al.*, 1998; Mamat *et al.*, 2002). This requirement correlates with high intracellular concentrations of cyclic 2,3-diphosphoglycerate (∼1 *M*) in hyperthermophilic *M. kandleri* (Shima *et al.*, 1998).

Additionally to these potassium ions, MK0840^{apo} also binds an anion, which was identified as chloride (Fig. 1). It is coordinated in a pyramidal geometry by Glu324 N, Arg328 N$^{\eta1}$ and a water molecule. The anion–ligand distances are 2.9–3.4 Å and are in the range observed for other chloride-binding sites (Qian *et al.*, 1994; Feller *et al.*, 1996).

3.3. Dimer formation

In solution, full-length and truncated *M. kandleri* MK0840 forms dimers and tetramers in a 3:1 ratio, as observed in size-exclusion chromatography and native PAGE experiments. However, only the dimeric protein produces crystals. In the MK0840$^{apo\text{-}m}$ crystal form belonging to space group $P2_1$, the asymmetric unit contains a dimer with the two monomers related by a twofold noncrystallographic symmetry axis. This dimer is observed in all of the crystal forms reported here and is unusual. Most dimeric ASKHA family members dimerize *via* their C-terminal domains II (Lunin *et al.*, 2004; Mukai *et al.*, 2004; Gorrell *et al.*, 2005; Simanshu *et al.*, 2005; Hong *et al.*, 2006; Nishimasu *et al.*, 2007), burying a surface area of approximately 1400–3100 Å^2 per monomer. In contrast, MK0840 forms a dimer *via* the N-terminal domains I of each monomer (Fig. 3*a* and Supplementary Fig. S4), with a surface area of 1455 Å^2 per monomer in MK0840$^{apo\text{-}m}$ as determined using the *PISA* server (Krissinel & Henrick, 2007). The 40 residues involved from each monomer interact by 13 hydrogen bonds and five salt bridges. A similar dimer formation *via* domains I is only observed in *So*AnmK (PDB entry 3cqy; Midwest Center for Structural Genomics, unpublished work), involving 51 residues and a surface area of 1844 Å^2 per monomer, as well as in glycerol kinase from *Enterococcus casseliflavus* (Yeh *et al.*, 2003), involving 44 residues and an area of 1400 Å^2 per monomer. In the contact region between

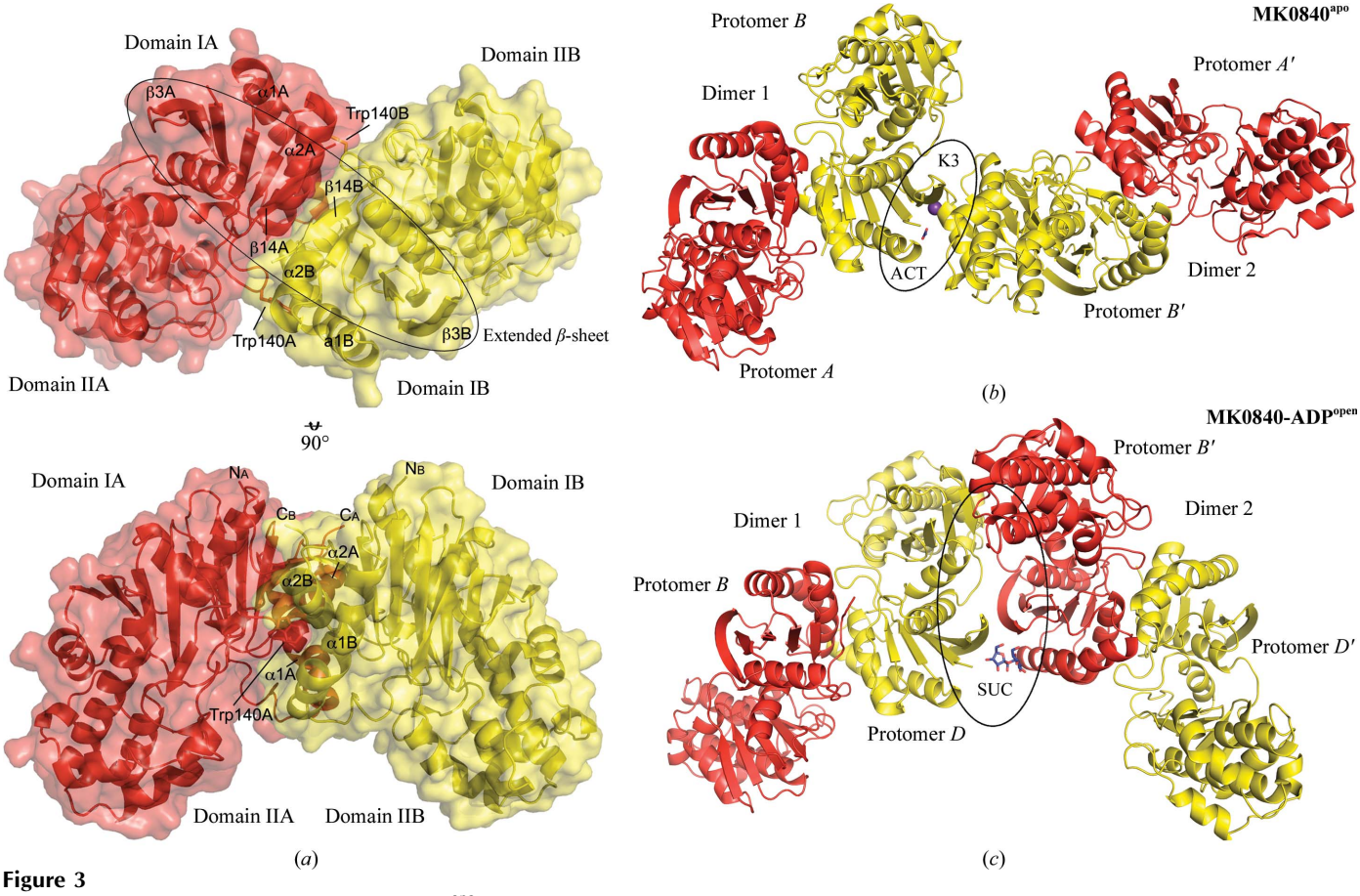


Figure 3
Dimer formation in MK0840. (*a*) Both MK0840^{apo} and MK0840–ADP form a dimer involving both domains I building up an extended β-sheet. Protomers *A* and *B* are coloured red and yellow, respectively. The molecular surface and structural elements involved in dimer formation are also shown. MK0840^{apo} (*b*) and MK0840–ADP^{open} (*c*) build up additional contact interfaces that differ from each other. (*b*) MK0840^{apo} buries an acetate (ACT) molecule and a potassium ion (K3) within the crystallographic contact interface between protomers *B* and *B*′ from two distinct dimers. (*c*) In MK0840–ADP^{open} the crystallographic contact interface between protomers *D* and *B*′ from two distinct dimers buries the sucrose (SUC) molecule.

the two monomers of MK0840 an extended β-interface is formed by the interplay of β-strands 14 (Leu346–Cys348) on both protomers, building up a continuous sheet spanning the two subunits (Fig. 3*a* and Supplementary Fig. S4). Most of the contact area is formed by interactions between residues located in helices α1 and α2. Each monomer provides Trp140, a residue located in a tight turn between strand β5 and helix η1 and which extends into the other monomer, contacting residues Arg88, Ala92, Ile96, Glu126 and Lys127 *via* van der Waals interactions, and Met123 *via* a hydrogen bond.

3.4. Nucleotide-binding site and domain closure

In order to investigate nucleotide binding and potential accompanying conformational rearrangements, we undertook cocrystallization of the N-terminally truncated construct with ADP. Two different crystal forms were obtained under different conditions, termed ADP^{open} and ADP^{closed}. They differ in ligand state, crystal packing and the conformation of the MK0840 molecules, but have the following features in common.

Both MK0840–ADP structures have ADP bound (Fig. 4 and Supplementary Fig. S5). While both apo crystal forms display a less well defined stretch between Glu205 and Lys211 (Fig. 2*b*), indicating flexibility in this part of the structure, this segment becomes more ordered upon nucleotide binding. Several side chains have to move to allow the binding of the adenine moiety of the nucleotide. The largest changes are observed in the χ_1 angles of Glu205, which changes from 87 to -148° (from apo to ADP-bound), and of Tyr206, which flips from -178 to 54°. By adopting the new position in the MK0840–ADP structures, the phenol ring of Tyr206 can engage in π–π stacking interactions with the adenine moiety of ADP. In the apo structures Glu208 occupies the spatial position of the adenosine part of ADP (Fig. 2*b*), whereas in the MK0840–ADP structures the adenine N6 atom is coordinated by O$^{\varepsilon1}$ of this residue (Figs. 2*b*, 4*a* and 4*b*).

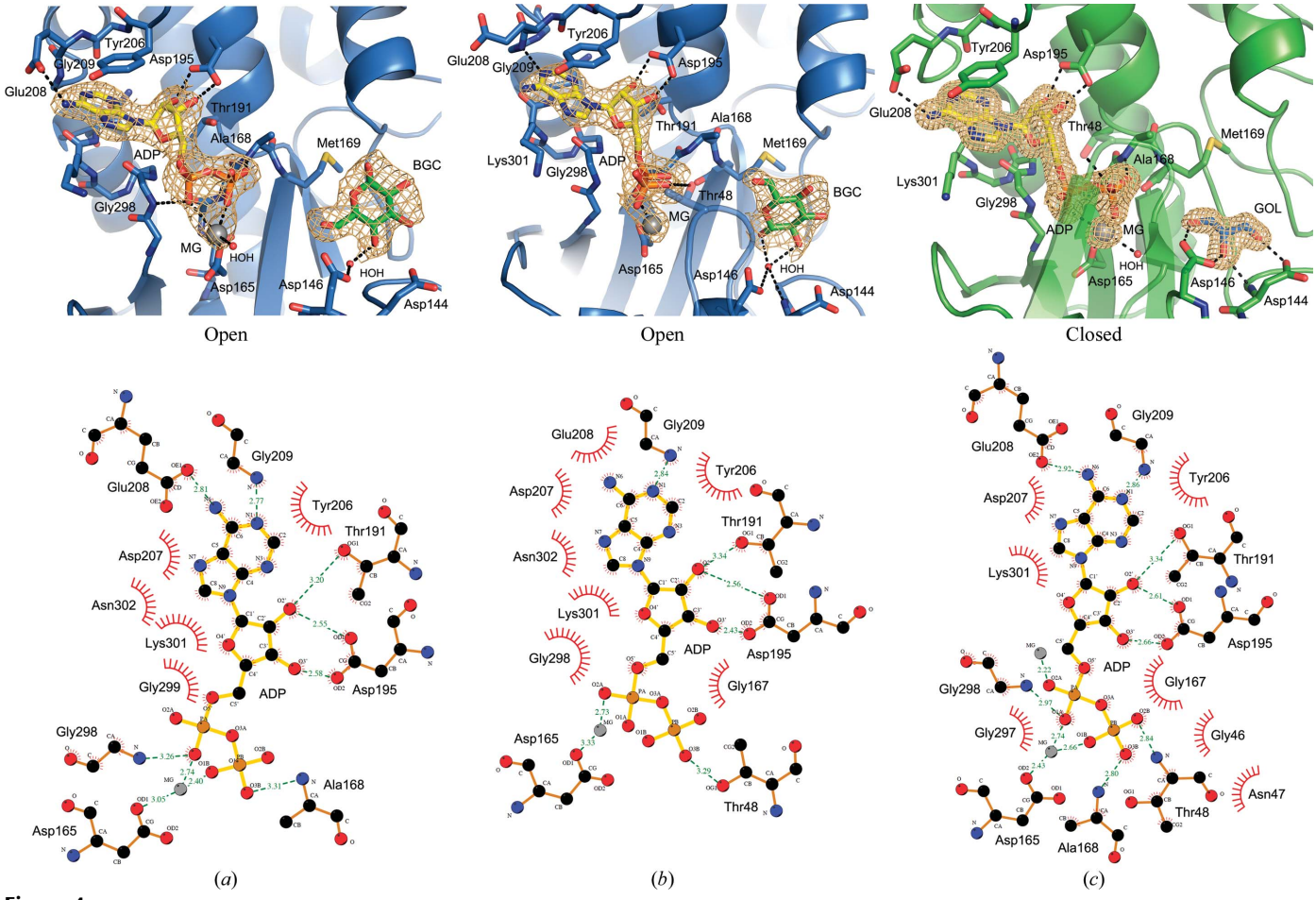


Figure 4
Nucleotide and substrate binding differs in the open and closed forms of MK0840. In protomers *B* and *D* of MK0840–ADP^{open} the β-phosphate of ADP points towards the β-D-glucose (BGC) molecule (*a*), while in protomers *A* and *C* it points towards Thr48 (*b*). The set of residues coordinating the adenine and ribose moiety also differs. In protomers *A* and *C* (*b*) interactions with Ala168 and Glu208 are missing but a new interaction with Thr48 is formed. (*c*) In MK0840–ADP^{closed} glycerol (GOL) is bound. The ADP molecule in (*c*) adopts the same conformation as in (*a*) but interacts simultaneously with Ala168 and Thr48, leading to domain closure. In all six protomers of both MK0840–ADP forms the adenine moiety of ADP is stacked against the phenol ring of Tyr206 and the hydroxyl groups of the ribose are coordinated by Asp195. For better visibility the loop containing Thr48 was removed in (*a*). In (*a*)–(*c*) $2F_o - F_c$ OMIT electron-density maps (light orange) of the ADP molecule, the Mg^{2+} ion, BCG (*a*, *b*) and GOL (*c*) contoured at 1.0σ are shown (upper row). The ligand plots depict interactions of the ADP molecule with protein residues (lower row). Green dashes represent the hydrogen-bonding network; numbers represent distances in Å.

The ADP molecule binds in the cleft of domain II formed by residues located on helix α4 and the loop connecting it to the 3_{10}-helix η2 as well as on helix η4. The ribose ring is in the C3′-*endo* conformation, with the 2′-oxygen and the 3′-oxygen forming hydrogen bonds to the side-chain carboxyl group of Asp195 and van der Waals interactions with the hydroxyl group of Thr191. The adenine ring adopts the *anti* conformation and is sandwiched between the side chains of Tyr206 and Lys301 (Fig. 4 and Supplementary Fig. S5). The adenosine moiety makes additional van der Waals interactions with the backbone of residues in helix η4 as well as water-mediated hydrogen bonds with the backbone amide of Asp207 and the carboxamide side chain of Asn302. The α-phosphoryl group of ADP interacts with the backbone amide of Gly298 and a Mg^{2+} ion, which itself is coordinated by the carboxyl group of Asp165.

The two crystal forms exhibit distinct differences that will be described in the following.

3.4.1. The MK0840–ADP^{open} crystal form. These crystals belonged to the monoclinic space group $P2_1$ and contained four molecules per asymmetric unit. All of them exhibit an open conformation very much resembling the apo forms observed in the non-ligand-bound structures MK0840$^{apo\text{-}o}$ and MK0840$^{apo\text{-}m}$. The r.m.s.d. between protomers *A* of the MK0840$^{apo\text{-}m}$ structure and the MK0840–ADP^{open} structure is 0.65 Å for 313 C^{α} atoms, indicating that ADP binds to the enzyme without causing major changes in domain orientation. It is known for ASKHA family members that domain closure can occur upon substrate or nucleotide binding (Yeh *et al.*, 2003; Grueninger & Schulz, 2006; Hong *et al.*, 2006). For proteins such as Hsc70 or actin, which use MgATP or magnesium polyphosphate to carry out their function, the nucleotides alone are able to induce domain closure (Frieden *et al.*, 1980; Wilbanks & McKay, 1995). However, for kinases which bind MgATP and additionally specific small-molecule substrates, *e.g.* carbohydrates, the substrate alone is sufficient to close the cleft (Buss *et al.*, 2001; Nishimasu *et al.*, 2007). The sole binding of ADP or other nucleotides usually leads to open conformations (Simanshu *et al.*, 2005; Schnick *et al.*, 2009). Both the MK0840^{apo} and the MK0840–ADP^{open} structures resemble the open conformations observed in other ASHKA kinases, *e.g.* *So*AnmK, in which a succinic acid molecule is bound (PDB entry 3cqy; Midwest Center for Structural Genomics, unpublished work), or *E. coli* RhuK, in which ADP and L-fructose were found in the crystal structure (Grueninger & Schulz, 2006).

During refinement, we could observe differences in the conformation and coordination of the ADP molecules in the individual protomers, in which the β-phosphate groups can occupy alternative positions (Figs. 4*a* and 4*b* and Supplementary Figs. S5*a* and S5*b*). Protomers *B* and *D* show torsion angles (C5′—O5′—PA—O3A) of −60 and −28°, respectively, whereas protomers *A* and *C* exhibit angles of 64 and 30°, respectively. These torsion angles lie in the range observed in the structures of other ASKHA family kinases, *e.g.* −50 and −66° for *Pa*AnmK (PDB entry 3qbw; Bacik *et al.*, 2011) and 67° (for both protomers) for acetate kinase (PDB entry 1tuy; Gorrell *et al.*, 2005). Consequently, the β-phosphate (Pβ) groups point in different directions: in protomers *B* and *D* the Pβ is mainly located nearer to the putative substrate-binding site, simultaneously interacting with the backbone amide of Ala168 located in the turn between strands β6 and β7 in domain II (Fig. 4*a* and Supplementary Fig. S5*a*). In protomers *A* and *C* the β-phosphate group points towards Thr48, a residue that is located in the turn connecting strands β1 and β2 in domain I, and interacts with its hydroxyl group (Fig. 4*b* and Supplementary Fig. S5*b*). Superimposition of MK0840–ADP^{open} with other kinases such as rhamnulose kinase (RhuK; Fig. 2*c*), *P. aeruginosa* panthothenate kinase (*Pa*CoaA) and the 2-hydroxyglutaryl-CoA dehydratase component A (HgdC) from *A. fermentans* reveals that Thr48 occupies the same spatial position as serine or threonine residues in these enzymes. These residues are involved in binding to the β-phosphate *via* their backbone amides and/or their side-chain hydroxyls (Ser14 in RhuK, Thr10 in *Pa*CoaA and Ser12-Thr13 in HgdC). They belong to the conserved PHOSPHATE I-

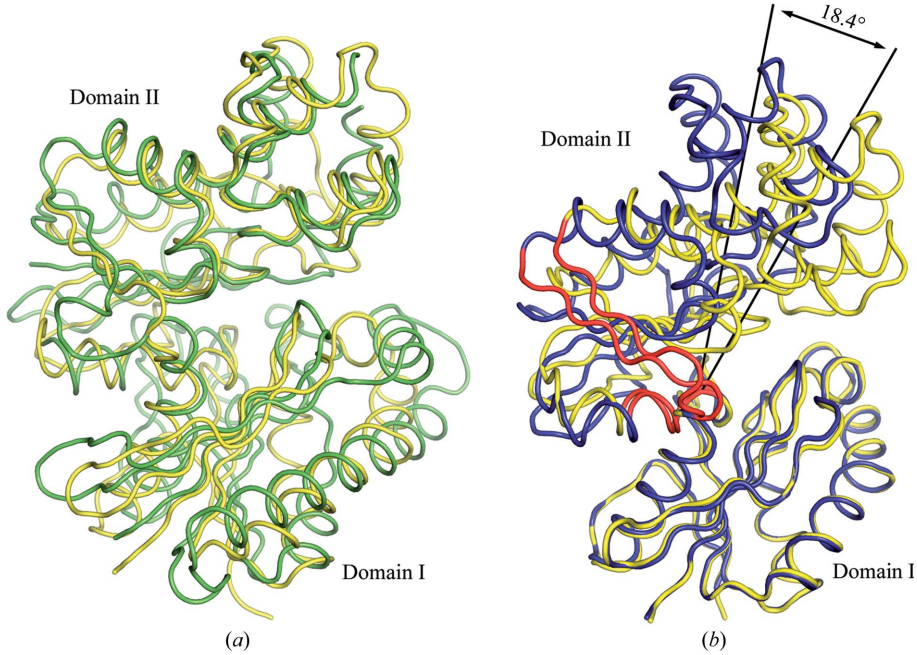


Figure 5
Domain closure in MK0840. (*a*) Superimposition of the closed conformation of MK0840–ADP^{closed} (yellow) with the closed conformation of *Pa*AnmK (green; PDB entry 3qbx). (*b*) Superimposition of the closed conformation of MK0840–ADP^{closed} (yellow) with the open conformation of MK0840–ADP^{closed} (blue). The degree of domain closure is indicated by arrows. The bending region is coloured red. A tube representation of the protein backbone is shown.

binding region in the ASKHA superfamily (Fig. 1 and Supplementary Fig. S1).

3.4.2. The MK0840–ADP^{closed} crystal form. This crystal form belonged to the orthorhombic space group $P2_12_12_1$ and contained two crystallographically independent molecules. Protomer *A* exhibits a conformation very similar to the apo and ADP^{open} crystal forms. The coordination pattern of the bound ADP is similar to that of ADP^{open}, including alternating positions of the β-phosphate group. However, for protomer *B* of MK0840–ADP^{closed} a closed conformation could be assigned similar to the conformation of ADP-bound *Pa*AnmK (Fig. 5*a*). Compared with the open conformation of MK0840^{open}, in the closed conformation domain II makes an 18.4° movement towards domain I (Fig. 5*b*) as calculated by the *DynDom* server (Hayward & Berendsen, 1998). The typical angle of domain movement in ASKHA family kinases lies in the range 10–30°, *e.g.* 14.9° for glycerol kinase [open, PDB entry 2p3r (*E*); closed, PDB entry 2p3r (*A*)], 18.3° for hexokinase [open, PDB entry 2e2n (*A*); closed, PDB entry 2e2q (*A*)] or 29.4° for 3-phosphoglycerate kinase [open, PDB entry 1qpg (*A*); closed, PDB entry 13pk (*A*)], all taken from the *DynDom* database of specific domain movements (Qi & Hayward, 2009). The bending regions in protomer *B* of MK0840–ADP^{closed} were identified as Asp146–Pro151 (helices η1 and α3) and Glu314–Glu324 (strand β13), which are all residues connecting domains I and II. They fall into the conserved motifs CONNECT I and CONNECT II, respectively, as identified by Bork *et al.* (1992), that have been demonstrated to serve as the interdomain hinge in ASKHA proteins. Interestingly, the residue Asp146 that is considered to be the most likely catalytic base in MK0840 (described below) is located at the very edge of the first bending region, indicating interplay of domain closure and catalysis at this point in the structure. Amino acids at various positions in the three-dimensional structure have been identified to play important roles in domain closure in ASKHA kinases. For glycerol kinase, the mutagenesis of a conserved aspartic acid residue from domain I involved in MgATP binding to asparagine or alanine led to a significantly decreased enzymatic activity (Pettigrew *et al.*, 1998). It was speculated that this residue mediates contacts between domains I and II and thus influences domain closure. This Asp corresponds to Gln44 in MK0840. Other residues lying in the CONNECT I/II region and substrate-binding region have been identified to influence domain closure in 3-phosphoglycerate kinase using SAXS, calorimetric and kinetic studies (Szabó *et al.*, 2008*a*,*b*). These residues were found to be important for the structural integrity of the hinge region and the transmission of substrate-induced effects towards the main hinge. These and many other studies point to a cumulative effect of several side-chain interactions in the hinge region and of substrate binding in the mechanism of domain closure.

The interactions of ADP in this closed protomer *B* (Fig. 4*c* and Supplementary Fig. S5*c*) are the total of all interactions observed for both conformers of the ADP molecule in all four crystallographically independent copies of the MK0840–ADP^{open} crystal form (Figs. 4*a* and 4*b* and Supplementary Figs. S5*a* and S5*b*). The distance between Ala168 N and Thr48 O$^{\gamma1}$ in the closed form of MK0840–ADP^{closed} is 4.40 Å, similar to the distance found in the ADP-bound form of *Pa*AnmK (PDB entry 3qbw; 4.21 Å), and is considerably shorter than in the open form (about 8.2 Å). As a result of the domain closure, the α-phosphates and β-phosphates are fixed in space between the two domains I and II by interactions of the phosphate O atoms with Ala168 N and Thr48 O$^{\gamma1}$ (Fig. 4*c* and Supplementary Fig. S5*c*). Thus, one of the two alternating conformations of the nucleotide becomes selected upon domain closure.

3.5. Putative substrate-binding site

In both ADP-bound structures unaccounted-for electron density could be observed proximal to the ADP-binding site. In the ADP^{open} structure this density was interpreted as β-D-glucose (BGC; Figs. 4*a* and 4*b* and Supplementary Figs. S5*a* and S5*b*). Owing to the lower resolution and the probable multiple binding modes, this is a tentative assignment. The glucose could originate from the sucrose solution used as a cryoprotectant. In the ADP^{closed} crystal form glycerol (GOL; Fig. 4*c* and Supplementary Fig. S5*c*) was unequivocally identified as the ligand, originating from the crystallization solution. The ASHKA family comprises glucose and glycerol kinases and these molecules bind in the neighbourhood of the putative substrate-binding site in MK0840 as located in homologous enzymes.

The binding of BGC is essentially accomplished by water-mediated hydrogen-bond interactions of the side chain of Asp146 and the backbone amide N atom of Asp144 with the hydroxyl group of the sugar. Additionally, Met169, Val171 and Val247 form a hydrophobic pocket around the sugar. In contrast to the water-mediated interactions with BGC in MK0840–ADP^{open}, Asp146 and Asp144 interact directly with GOL in MK0840–ADP^{closed}. The carboxylate O atoms of the Asp146 side chain form polar contacts to GOL O1 and O2, and the carboxylate oxygen O$^{\delta1}$ and the backbone amide N atom of Asp144 form contacts to GOL O2 and O3, respectively (Fig. 4*c* and Supplementary Fig. S5*c*). Asp146 occupies the same spatial position as Asp237 in RhuK (Fig. 2*c*; Grueninger & Schulz, 2006) and Asp211 in yeast hexokinase PII (Kuser *et al.*, 2000), where they serve as the general base in the phosphorylation reaction. Met169 superimposes well with Trp260 of RhuK (Fig. 2*c*), which is rotated to fill the empty acceptor-binding site of the apoenzyme and is considered to perform a local induced fit in addition to the general domain closure during substrate binding (Grueninger & Schulz, 2006). A similar function for Met169 can only be surmised, as the substrate preference of MK0840 is as yet unknown and only minor conformational changes are observed for the Met169 side chain in the apo and holo form.

For the following reasons, we believe that neither glucose nor glycerol bind in a substrate-like manner here: sugar kinases usually bind their substrates *via* several side-chain-mediated or water-mediated hydrogen bonds. Here, only one such interaction is formed to BGC. In addition, BGC binding

does not occur in one defined way (Figs. 4*a* and 4*b* and Supplementary Figs. S5*a* and S5*b*) in the four distinct protomers. Furthermore, the enzyme still exhibits an open conformation, whereas binding of substrates usually leads to domain closure. The same holds true for GOL, where despite the involvement of several side-chain and backbone interactions in ligand binding an open conformation is observed in one of the protomers (Fig. 4*c* and Supplementary Fig. S5*c*). Such binding modes have been observed before: for glycerol kinase from *P. falciparum* in the ADP-bound structure (PDB entry 2w41) the cryoprotectant ethylene glycol was found in the substrate-binding site involving several interactions also found in glycerol binding (Simanshu *et al.*, 2005; Schnick *et al.*, 2009) and the enzyme showed an open conformation. Additionally, the binding of ASKHA substrates involves interactions from both domains I and II, as found in the complexes of RhuK with β-L-rhamnulose (PDB entry 2uyt; Grueninger & Schulz, 2007) and β-L-fructose (PDB entry 2cgj; Grueninger & Schulz, 2006) and the complex of *Pa*AnmK with anhMurNAc (PDB entry 3qbx; Bacik *et al.*, 2011). Both BGC and GOL ligands only interact with residues located in domain I, leaving a larger gap between them and domain II, with a distance in the range 4.6–6.5 Å between the nearest atom of the ligand and Met169 S^{δ} in domain II. To corroborate this conclusion, we performed activity assays for 45 putative substrates including GOL and BGC and none were phosphorylated (see below and Supplementary Table S1).

MK0840 is assigned as a member of the anhydro-*N*-acetylmuramic acid kinase family (Pfam PF03702; Finn *et al.*, 2010). As mentioned above, crystal structures of this family are known for *So*AnmK and *Pa*AnmK. The alignment of 291 residues using *DALI* (Holm & Rosenström, 2010) resulted in r.m.s.d.s of 3.9 and 3.8 Å for *So*AnmK and *Pa*AnmK, respectively, for the open confirmation. However, superimposition of $MK0840^{closed}$ with both AnmK homologues (r.m.s.d. of 3.5 Å) shows that the architectures of their substrate-binding sites are quite different (Supplementary Fig. S2). In *Pa*AnmK the general base is reported to be Asp182 (Bacik *et al.*, 2011), which also superposes well with Asp188 of *So*AnmK. These residues occupy the same spatial position as Val187 in MK0840 (Supplementary Fig. S2) and are located on the opposite site of the putative substrate-binding site. Other residues involved in the binding of anhMurNAc, such as Arg129 in *Pa*AnmK and the equivalent Arg134 in *So*AnmK, are replaced by Asp144 in MK0840, a residue involved in glycerol binding in $MK0840–ADP^{closed}$. A global comparison of the MK0840 substrate-binding pocket with ASKHA family kinase structures with highest structural similarity (Z-score > 16, r.m.s.d. < 4.4 Å using *DALI*) showed that one feature of the MK0840 ligand-binding pocket is quite unique. Immediately upstream of Asp144/146 there is an insertion between strand $\beta5$ and helices $\eta1$ and $\alpha3$ (Fig. 2*a*) that harbours Trp140 and the two potassium-binding sites. It is tempting to speculate that it may have an influence on substrate binding or domain closure or even mediate allosteric regulation through Trp140, which interacts with the second protomer of the dimer.

To investigate the substrate preference of MK0840 more closely, we performed kinase assays with 45 different putative substrates (see Supplementary Material and Supplementary Table S1). For some kinases, such as pyruvate kinase M2, two different oligomeric states have been reported, of which only one represents the active form (Eigenbrodt *et al.*, 1992). For this reason, we tested dimeric and tetrameric fractions of MK0840 for activity. None of the compounds was phosphorylated by MK0840. ATPase assays were performed in order to determine whether purified full-length and truncated MK0840 is still in an active conformation or whether it is inactivated by the purification procedure (Supplementary Material). It has been described, for example for *E. coli* glycerol kinase (Pettigrew *et al.*, 1990) and yeast hexokinase (Kaji & Colowick, 1965; DelaFuente *et al.*, 1970), that very low substrate-independent MgATPase activity can occur in the absence of substrate. In the presence of MK0840 we observed a fivefold increase in ATP hydrolysis compared with the background (Supplementary Fig. S3). From all these findings we can conclude that MK0840 is purified in its active form but that its substrate preference remains to be elucidated.

4. Conclusions

The crystal structures of the putative Hsp70 metalloprotease-fold MK0840 from *M. kandleri* reveal an ASHKA family protein that most closely resembles anhydro-*N*-acetylmuramic acid kinase. Like the ASHKA family proteins, it exhibits a bilobal architecture in which a pronounced domain closure can be observed in one crystal form, indicating a functional protein. However, structural differences are observed in the active site of MK0840, especially the lack of conservation of an aspartic acid that is crucial for catalysis in the established anhydro-*N*-acetylmuramic acid kinases. Together with the fact that *M. kandleri*, like other pseudomurein-coated archaea, does not contain anhMurNAc (Hartmann & König, 1990; Kurr *et al.*, 1991; Claus & König, 2010), this indicates that is very unlikely that MK0840 is an anhydro-*N*-acetylmuramic acid kinase. Archaebacterial anhydro-*N*-acetylmuramic acid counterparts (*N*-acetyltalosaminuronic acid and its derivatives) are currently not available and therefore could not be tested as substrates in kinase assays. Further investigations will be necessary in order to unravel the biological function of MK0840.

This work was supported by the Swiss National Science Foundation, the University of Bern and the Berner Hochschulstiftung. The help of all of the beamline scientists at beamline X06DA at the Swiss Light Source, Paul Scherrer Institute, Villigen, Switzerland is highly appreciated.

References

Adams, P. D. *et al.* (2010). *Acta Cryst.* D**66**, 213–221.
Atrih, A. & Foster, S. J. (1999). *Antonie Van Leeuwenhoek*, **75**, 299–307.
Bacik, J.-P., Whitworth, G. E., Stubbs, K. A., Yadav, A. K., Martin, D. R., Bailey-Elkin, B. A., Vocadlo, D. J. & Mark, B. L. (2011). *J. Biol. Chem.* **286**, 12283–12291.

Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N. & Bourne, P. E. (2000). *Nucleic Acids Res.* **28**, 235–242.

Bork, P., Sander, C. & Valencia, A. (1992). *Proc. Natl Acad. Sci. USA*, **89**, 7290–7294.

Breitung, J., Börner, G., Scholz, S., Linder, D., Stetter, K. O. & Thauer, R. K. (1992). *Eur. J. Biochem.* **210**, 971–981.

Brünger, A. T. (1997). *Methods Enzymol.* **277**, 366–396.

Buss, K. A., Cooper, D. R., Ingram-Smith, C., Ferry, J. G., Sanders, D. A. & Hasson, M. S. (2001). *J. Bacteriol.* **183**, 680–686.

Chan, K.-M., Delfert, D. & Junger, K. D. (1986). *Anal. Biochem.* **157**, 375–380.

Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. (2010). *Acta Cryst.* D**66**, 12–21.

Claus, H. & König, H. (2010). *Prokaryotic Cell Wall Compounds: Structure and Biochemistry*, edited by H. König, H. Claus & A. Varma, pp. 231–251. New York: Springer.

DelaFuente, G., Lagunas, R. & Sols, A. (1970). *Eur. J. Biochem.* **16**, 226–233.

DeLano, W. L. (2002). *PyMOL*. http://www.pymol.org.

Doublié, S. (1997). *Methods Enzymol.* **276**, 523–530.

Eigenbrodt, E., Reinacher, M., Scheefers-Borchel, U., Scheefers, H. & Friis, R. (1992). *Crit. Rev. Oncog.* **3**, 91–115.

Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* D**66**, 486–501.

Feller, G., Bussy, O., Houssier, C. & Gerday, C. (1996). *J. Biol. Chem.* **271**, 23836–23841.

Finn, R. D., Mistry, J., Tate, J., Coggill, P., Heger, A., Pollington, J. E., Gavin, O. L., Gunasekaran, P., Ceric, G., Forslund, K., Holm, L., Sonnhammer, E. L., Eddy, S. R. & Bateman, A. (2010). *Nucleic Acids Res.* **38**, D211–D222.

Frieden, C., Lieberman, D. & Gilbert, H. R. (1980). *J. Biol. Chem.* **255**, 8991–8993.

Gille, C. & Frömmel, C. (2001). *Bioinformatics*, **17**, 377–378.

Goodell, E. W. & Schwarz, U. (1985). *J. Bacteriol.* **162**, 391–397.

Gorrell, A., Lawrence, S. H. & Ferry, J. G. (2005). *J. Biol. Chem.* **280**, 10731–10742.

Gouet, P., Courcelle, E., Stuart, D. I. & Métoz, F. (1999). *Bioinformatics*, **15**, 305–308.

Grueninger, D. & Schulz, G. E. (2006). *J. Mol. Biol.* **359**, 787–797.

Grueninger, D. & Schulz, G. E. (2007). *FEBS Lett.* **581**, 3127–3130.

Hartmann, E. & König, H. (1990). *Naturwissenschaften*, **77**, 472–475.

Hayward, S. & Berendsen, H. J. (1998). *Proteins*, **30**, 144–154.

Holm, L. & Rosenström, P. (2010). *Nucleic Acids Res.* **38**, W545–W549.

Hong, B. S., Yun, M. K., Zhang, Y.-M., Chohnan, S., Rock, C. O., White, S. W., Jackowski, S., Park, H.-W. & Leonardi, R. (2006). *Structure*, **14**, 1251–1261.

Jonge, B. L. M. de, Chang, Y.-S., Gage, D. & Tomasz, A. (1992). *J. Biol. Chem.* **267**, 11248–11254.

Jonge, B. L. M. de, Handwerger, S. & Gage, D. (1996). *Antimicrob. Agents Chemother.* **40**, 863–869.

Kabsch, W. (2010). *Acta Cryst.* D**66**, 125–132.

Kabsch, W. & Sander, C. (1983). *Biopolymers*, **22**, 2577–2637.

Kaji, A. & Colowick, S. P. (1965). *J. Biol. Chem.* **240**, 4454–4462.

Karplus, P. A. & Diederichs, K. (2012). *Science*, **336**, 1030–1033.

Kelley, L. A. & Sternberg, M. J. (2009). *Nature Protoc.* **4**, 363–371.

Krissinel, E. & Henrick, K. (2007). *J. Mol. Biol.* **372**, 774–797.

Kurr, M., Huber, R., König, H., Jannasch, H. W., Fricke, H., Trincone, A., Kristjansson, J. K. & Stetter, K. O. (1991). *Arch. Microbiol.* **156**, 239–247.

Kuser, P. R., Krauchenco, S., Antunes, O. A. & Polikarpov, I. (2000). *J. Biol. Chem.* **275**, 20814–20821.

Langer, G., Cohen, S. X., Lamzin, V. S. & Perrakis, A. (2008). *Nature Protoc.* **3**, 1171–1179.

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.

Laskowski, R. A. & Swindells, M. B. (2011). *J. Chem. Inf. Model.* **51**, 2778–2786.

Locher, K. P., Hans, M., Yeh, A. P., Schmid, B., Buckel, W. & Rees, D. C. (2001). *J. Mol. Biol.* **307**, 297–308.

Lunin, V. V., Li, Y., Schrag, J. D., Iannuzzi, P., Cygler, M. & Matte, A. (2004). *J. Bacteriol.* **186**, 6915–6927.

Mamat, B., Roth, A., Grimm, C., Ermler, U., Tziatzios, C., Schubert, D., Thauer, R. K. & Shima, S. (2002). *Protein Sci.* **11**, 2168–2178.

McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.

Morris, R. J., Perrakis, A. & Lamzin, V. S. (2002). *Acta Cryst.* D**58**, 968–975.

Mukai, T., Kawai, S., Mori, S., Mikami, B. & Murata, K. (2004). *J. Biol. Chem.* **279**, 50591–50600.

Nishimasu, H., Fushinobu, S., Shoun, H. & Wakagi, T. (2007). *J. Biol. Chem.* **282**, 9923–9931.

Oken, M. M., Peterson, P. K. & Wilkinson, B. J. (1981). *Infect. Immun.* **31**, 208–213.

Painter, J. & Merritt, E. A. (2006*a*). *Acta Cryst.* D**62**, 439–450.

Painter, J. & Merritt, E. A. (2006*b*). *J. Appl. Cryst.* **39**, 109–111.

Pal, D. & Chakrabarti, P. (1999). *J. Mol. Biol.* **294**, 271–288.

Pettigrew, D. W., Smith, G. B., Thomas, K. P. & Dodds, D. C. (1998). *Arch. Biochem. Biophys.* **349**, 236–245.

Pettigrew, D. W., Yu, G. J. & Liu, Y. (1990). *Biochemistry*, **29**, 8620–8627.

Qi, G. & Hayward, S. (2009). *BMC Struct. Biol.* **9**, 13.

Qian, M., Haser, R., Buisson, G., Duée, E. & Payan, F. (1994). *Biochemistry*, **33**, 6284–6294.

Scheffers, D. J. & Pinho, M. G. (2005). *Microbiol. Mol. Biol. Rev.* **69**, 585–607.

Schnick, C., Polley, S. D., Fivelman, Q. L., Ranford-Cartwright, L. C., Wilkinson, S. R., Brannigan, J. A., Wilkinson, A. J. & Baker, D. A. (2009). *Mol. Microbiol.* **71**, 533–545.

Shima, S., Hérault, D. A., Berkessel, A. & Thauer, R. K. (1998). *Arch. Microbiol.* **170**, 469–472.

Simanshu, D. K., Savithri, H. S. & Murthy, M. R. (2005). *J. Mol. Biol.* **352**, 876–892.

Slabinski, L., Jaroszewski, L., Rychlewski, L., Wilson, I. A., Lesley, S. A. & Godzik, A. (2007). *Bioinformatics*, **23**, 3403–3405.

Stewart-Tull, D. E. (1980). *Annu. Rev. Microbiol.* **34**, 311–340.

Szabó, J., Varga, A., Flachner, B., Konarev, P. V., Svergun, D. I., Závodszky, P. & Vas, M. (2008*a*). *FEBS Lett.* **582**, 1335–1340.

Szabó, J., Varga, A., Flachner, B., Konarev, P. V., Svergun, D. I., Závodszky, P. & Vas, M. (2008*b*). *Biochemistry*, **47**, 6735–6744.

Terwilliger, T. C., Adams, P. D., Moriarty, N. W. & Cohn, J. D. (2007). *Acta Cryst.* D**63**, 101–107.

Terwilliger, T. C., Grosse-Kunstleve, R. W., Afonine, P. V., Moriarty, N. W., Zwart, P. H., Hung, L.-W., Read, R. J. & Adams, P. D. (2008). *Acta Cryst.* D**64**, 61–69.

Terwilliger, T. C., Klei, H., Adams, P. D., Moriarty, N. W. & Cohn, J. D. (2006). *Acta Cryst.* D**62**, 915–922.

Uehara, T., Suefuji, K., Valbuena, N., Meehan, B., Donegan, M. & Park, J. T. (2005). *J. Bacteriol.* **187**, 3643–3649.

Van Duyne, G. D., Standaert, R. F., Karplus, P. A., Schreiber, S. L. & Clardy, J. (1993). *J. Mol. Biol.* **229**, 105–124.

Vriend, G. (1990). *J. Mol. Graph.* **8**, 52–56.

Wilbanks, S. M. & McKay, D. B. (1995). *J. Biol. Chem.* **270**, 2251–2257.

Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.

Yang, Z. R., Thomson, R., McNeil, P. & Esnouf, R. M. (2005). *Bioinformatics*, **21**, 3369–3376.

Yeh, J. I., Charrier, V., Paulo, J., Hou, L., Darbon, E., Claiborne, A., Hol, W. G. J. & Deutscher, J. (2003). *Biochemistry*, **43**, 362–373.