Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structure of the nucleotide-binding domain of a dipeptide ABC transporter reveals a novel iron–sulfur cluster-binding domain

Xiaolu Li,[a,b] Wei Zhuo,[b] Jie Yu,[b] Jingpeng Ge,[b] Jinke Gu,[b] Yue Feng,[b] Maojun Yang,[b] Linfang Wang[a]* and Na Wang[b]*

[a]State Key Laboratory of Medical Molecular Biology, Department of Biochemistry and Molecular Biology, Institute of Basic Medical Sciences, Chinese Academy of Medical Sciences, Peking Union Medical College, Tsinghua University, Beijing 100005, People's Republic of China, and [b]Key Laboratory for Protein Sciences of Ministry of Education, Tsinghua–Peking Center for Life Sciences, School of Life Sciences, Tsinghua University, Beijing 100084, People's Republic of China

Correspondence e-mail: lfwangz@yahoo.com, nawang@tsinghua.edu.cn

Received 15 July 2012
Accepted 1 November 2012

PDB Reference: DppD, 4fwi

Dipeptide permease (Dpp), which belongs to an ABC transport system, imports peptides consisting of two or three L-amino acids from the matrix to the cytoplasm in microbes. Previous studies have indicated that haem competes with dipeptides to bind DppA *in vitro* and *in vivo* and that the Dpp system can also translocate haem. Here, the crystal structure of DppD, the nucleotide-binding domain (NBD) of the ABC-type dipeptide/oligopeptide/nickel-transport system from *Thermoanaerobacter tengcongensis*, bound with ATP, Mg^{2+} and a [4Fe–4S] iron–sulfur cluster is reported. The N-terminal domain of DppD shares a similar structural fold with the NBDs of other ABC transporters. Interestingly, the C-terminal domain of DppD contains a [4Fe–4S] cluster. The UV–visible absorbance spectrum of DppD was consistent with the presence of a [4Fe–4S] cluster. A search with *DALI* revealed that the [4Fe–4S] cluster-binding domain is a novel structural fold. Structural analysis and comparisons with other ABC transporters revealed that this iron–sulfur cluster may act as a mediator in substrate (dipeptide or haem) binding by electron transfer and may regulate the transport process in Dpp ABC transport systems. The crystal structure provides a basis for understanding the properties of ABC transporters and will be helpful in investigating the functions of NBDs in the regulation of ABC transporter activity.

1. Introduction

ATP-binding cassette (ABC) transporters are ubiquitous membrane proteins that use energy from ATP binding and hydrolysis to transport different substrates ranging from small ions to large macromolecules across the membrane bilayer (Dean & Allikmets, 1995). Members of the ABC transporter family are found in organisms from all kingdoms of life (Higgins, 1992). They can function as exporters or importers. Exporters are found in both eukaryotes and prokaryotes, while importers seem to exclusively be present in prokaryotic organisms. ABC transporters comprise the largest protein family in *Escherichia coli*, including 79 proteins from 69 independent functional systems representing 5% of the genome (Linton & Higgins, 1998), whereas only 50 ABC transporters are present in humans (Dean *et al.*, 2001). Members of the seven families of human ABC transporters (Dean & Annilo, 2005) participate in cholesterol and lipid transport (Voloshyna & Reiss, 2011), multidrug resistance (Moitra *et al.*, 2011), antigen presentation, mitochondrial iron homeostasis and ATP-dependent regulation of ion channels (including the cystic fibrosis transmembrane conductance regulator and the sulfonylurea receptors; Higgins, 1992). A

© 2013 International Union of Crystallography
Printed in Singapore – all rights reserved

wide range of disorders, including cystic fibrosis (Wang, 2010), hypercholesterolaemia, atherosclerosis (Voloshyna & Reiss, 2011) and diabetes (Dean *et al.*, 2001; Dean & Annilo, 2005), have been related to mutations in these proteins.

Typical ABC transporters have a characteristic architecture that minimally consists of four domains: two transmembrane domains (TMDs) that are embedded in the membrane bilayer and two ABC domains (or nucleotide-binding domains; NBDs) that are located in the cytoplasm (Doige & Ames, 1993). In prokaryotes these domains are often independent subunits which are assembled into a membrane-bound complex, whereas in eukaryotes these domains are generally fused into a single polypeptide chain. The TMDs have low sequence similarity, but the NBDs contain a set of highly conserved motifs. NBDs contain Walker A (G*xx*G*x*GKS/T, where *x* can be various amino acids) and Walker B (*hhhh*DEP, where *h* is a hydrophobic residue) motifs with consensus sequences for nucleotide binding (Walker *et al.*, 1982) and the ABC 'signature' sequence LSGGQ. The high similarity among the NBDs of all ABC transporters suggest that the mechanism of energy utilization by these systems is conserved. Recent structural studies of the ABC transporter NBDs, complementing the genetic and biochemical data already available in the literature, provide a glimpse at the atomic level of the specific residues involved in ATP binding and hydrolysis (Hung *et al.*, 1998).

E. coli dipeptide/oligopeptide/nickel ABC transporter (DppBCDF) transports peptides consisting of two or three L-amino acids and may also transport haem (Abouhamad & Manson, 1994). Haem, a major iron source, is transported through the outer membrane of Gram-negative bacteria by specific haemoprotein receptors and through the inner membrane by haem-specific periplasmic binding protein-dependent ATP-binding cassette transporters (Ferguson & Deisenhofer, 2004; Wandersman & Delepelaire, 2004; Stojiljkovic & Hantke, 1994). Haem can competitively bind to DppA both *in vitro* and *in vivo* as revealed by biochemical experiments (Létoffé *et al.*, 2006). However, whether the DppD system is a haem transporter remains unknown and requires further investigation. A dipeptide/oligopeptide/nickel ABC transporter is also found in *Thermoanaerobacter tengcongensis*, an anaerobic Gram-negative bacterium (Wang *et al.*, 2004) that utilizes the physical and chemical properties of Fe–S compounds for fundamental metabolism (such as amino-acid metabolism), as might have been present in the earliest forms of life under an anaerobic environment (Sheftel *et al.*, 2010). It has been reported that [4Fe–4S] clusters occur in eukaryotic ABC transporter proteins. However, to date no [4Fe–4S] clusters have been reported in prokaryotic NBDs. Here, we report the crystal structure of DppD, the nucleotide-binding domain of the ABC-type dipeptide/oligopeptide/nickel-transport system from *T. tengcongensis*. This structure is the first report of a native NBD monomer bound with ATP, Mg^{2+} and a [4Fe–4S] cluster. The crystal structure may provide a basis for understanding the properties of ABC transporters and the functions of iron–sulfur clusters in the process of transporting substrates.

2. Materials and methods

2.1. Materials

Enzymes for recombinant DNA technology such as *Pfu* polymerase, T4 DNA ligase, *Nde*I and *Xho*I were purchased from New England Biolabs. PCR buffer and dNTP mixture were also obtained from New England Biolabs. The DNA Quick Purify/Recover Kit and Plasmid Mini Kit were products of Omiga Co. The detergents Triton X-100 and *n*-decyl-β-D-maltopyranoside (DM) were purchased from Affymetrix. Chromatographic columns, such as Ni Sepharose 6 Fast Flow and Superdex 200 HR 10/30 columns, were purchased from GE Healthcare. Crystal screening kits were purchased from Hampton Research and Molecular Dimensions.

2.2. Construction, expression and purification

The *T. tengcongensis DppD* gene, which encodes a mature protein of 326 amino acids, was PCR-amplified using the genomic DNA of *T. tengcongensis* as a template. The forward primer was 5′-CGCCATATGAGCATAATCATAAGAGTTGAGG-3′ and the reverse primer was 5′-CCGCTCGAGTTTTAATGGTCTCCCTTCTTC-3′. The expression construct was generated using a standard PCR-based cloning strategy. In order to allow easy purification, a 6×His tag was added to its C-terminus. The expression construct was subcloned into the prokaryotic expression vector pET21b with *Nde*I and *Xho*I restriction-enzyme sites. The recombinant plasmid was transformed into *E. coli* strain C43 (DE3). Protein expression was induced by the addition of IPTG (0.5 m*M* final concentration) when the OD_{600} of the culture reached approximately 1.2 in LB medium. The cell culture was allowed to grow for a further 12 h at 296 K. The bacterial cells were harvested by centrifugation at 4000 rev min^{-1} for 15 min at 277 K. The cell pellets were resuspended in ice-cold lysis buffer (50 m*M* Tris–HCl pH 8.0, 200 m*M* NaCl, 1 m*M* PMSF) and homogenized by sonication on ice. The cell debris was removed completely by centrifugation of the lysate at 15 000 rev min^{-1} for 15 min at 277 K. The supernatant was collected and subjected to ultracentrifugation at 30000 rev min^{-1} for 1 h. The membrane fraction was harvested and resuspended in buffer *A* (25 m*M* Tris–HCl pH 8.0, 20 m*M* imidazole, 500 m*M* NaCl). After the addition of Triton X-100 to a final concentration of 0.5%(*v*/*v*) and 1 m*M* PMSF, the sample was incubated for 5 h with slow stirring at 277 K. The membrane was removed completely by further ultracentrifugation at 30 000 rev min^{-1} for 30 min at 277 K.

All of the following purification steps were performed in a 277 K cold room. The clear supernatant was loaded onto a Ni Sepharose 6 Fast Flow column pre-equilibrated with buffer *A* containing 0.1% Triton X-100. The contaminant proteins were eluted with wash buffer (buffer *A* containing 0.2% DM). The target protein was eluted with buffer *B* (25 m*M* Tris–HCl pH 8.0, 300 m*M* imidazole, 500 m*M* NaCl, 0.2% DM). The eluate was further purified by gel filtration using a Superdex 200 HR 10/30 column. The buffer for gel filtration consisted of 25 m*M* Tris–HCl pH 7.5, 500 m*M* NaCl, 2.5 m*M* $MgCl_2$, 2.5 m*M* ATP, 5 m*M* DTT, 0.2% DM. The purified protein was concentrated

using an Amicon Ultra centrifugal filter to a concentration of about 15 mg ml^{-1} as measured by the Bradford protein assay (Bradford, 1976). The purified protein was analyzed by SDS–PAGE and LTQ Orbitrap mass spectrometry.

2.3. LTQ Orbitrap mass-spectrometric analysis

The gel strip containing the target protein was removed from the SDS–PAGE gel, cut into small pieces and washed with 100 µl 25 m*M* ammonium bicarbonate pH 8.0 containing 50% acetonitrile for 15 min twice with vortexing. Gel pieces were dehydrated with 100 µl acetonitrile and completely dried with a Speed-Vac before tryptic digestion. The volume of the dried gel was evaluated and three volumes of trypsin (Promega, Madison, Wisconsin, USA) at 12.5 ng ml^{-1} in 25 m*M* NH_4HCO_3 (freshly diluted) were added. The digestion was performed at 303 K overnight. The mixture was sonicated for 10 min and centrifuged. The supernatant was removed and analyzed by high-resolution nano-LC-MS on an Orbitrap (LTQ Orbitrap Velos, Thermo Fisher Scientific).

All data files were processed using *Sequest*. Protein identification was performed by searching against bacterial genomes in the NCBI nonredundant database using the *Mascot* search engine (Matrix Science, London, England). The parameters used were monoisotopic, mass accuracy 0.1 Da, missed cleavages 1.

2.4. Crystallization, data collection and processing

The purified protein sample at a concentration of 15 mg ml^{-1} was centrifuged at 20 000*g* for 45 min to clarify the solution prior to initiating the crystallization trials. Initial screening was performed at 291 K in 48-well plates by the sitting-drop vapour-diffusion method using sparse-matrix screen kits from Hampton Research (PEGRx 1 and PEGRx 2) and Molecular Dimensions (MemStart, MemSys, MemPlus and MemGold) and was followed by a round of refinement of the initial conditions with variation of the precipitants, pH, protein concentration and additives using the hanging-drop vapour-diffusion method in 24-well plates. Typically, 2 µl droplets formed by mixing 1 µl protein solution with 1 µl reservoir solution were placed on siliconized cover slips and equilibrated against 200 µl reservoir solution.

X-ray diffraction data sets were collected from a single native crystal using an ADSC Quantum 315 detector on beamline BL17U at the Shanghai Synchrotron Radiation Facility (SSRF; Shanghai, People's Republic of China). The presence of iron was investigated using anomalous scattering. The reservoir solution supplemented with 12.5%(*v*/*v*) glycerol was used as a cryoprotectant. The crystals were mounted in a suitable nylon-fibre loop, immersed in the cryoprotectant solution for 5–10 s and flash-cooled in a liquid-nitrogen stream at 100 K. The wavelength used was 1.54148 Å, the exposure time was 1 s per frame, the crystal-to-detector distance was 150 mm and the oscillation range was 1° per frame. Highly redundant data sets were collected for the detection of anomalous signal. All intensity data were indexed, integrated and scaled using the *HKL*-2000 programs *DENZO* and *SCALEPACK* (Otwinowski & Minor, 1997).

2.5. Structure determination

The structure of *T. tengcongensis* DppD was determined by the single-wavelength anomalous dispersion (SAD) method using the data collected at the iron peak. In order to determine the iron positions, we collected highly redundant data (data set 2). The iron positions in the crystal were determined using the program *SHELXD* (Sheldrick, 2008). Although four iron sites were expected, only one iron site was found by *SHELXD*. The identified iron site was refined and the initial phases were generated using the SAD experimental phasing module in *Phaser* (McCoy *et al.*, 2007). The phase was then extended to data set 1 using *DM* (Winn *et al.*, 2011). Solvent flattening and histogram matching were performed using *DM*. The initial model was traced automatically using *Buccaneer* (Cowtan, 2006) and was manually rebuilt using *Coot* (Emsley & Cowtan, 2004). The final structure was refined with *PHENIX* (Adams *et al.*, 2002). The *PyMOL* molecular-graphics program (http://www.pymol.org) was used to illustrate the final structure and to produce the figures.

2.6. Spectroscopic analysis

2.6.1. UV–visible absorption spectroscopic analysis. UV–visible absorption spectroscopy of *T. tengcongensis* DppD between 300 and 700 nm was carried out using an Ultraspec 4300 spectrophotometer. All measurements were performed at room temperature using quartz cuvettes with 1 mm path length and a scan step of 1 nm. DppD protein (0.8 mg ml^{-1}) was solubilized in 50 m*M* sodium phosphate buffer pH 7.5. The UV–visible absorption spectrum of BSA at the same concentration was also recorded. A suitable baseline correction was made before recording each spectrum.

2.6.2. Far-UV circular-dichroism (CD) spectroscopic analysis. For far-UV CD spectroscopic analysis, the purified DppD protein was solubilized in 50 m*M* sodium phosphate buffer pH 7.5. To study the effect of ATP hydrolysis on the conformation of DppD, samples of DppD were incubated with 0.5 m*M* ATP and 1 m*M* $MgCl_2$ for 15 min at room temperature before recording each spectrum. To study the effect of reductants on the conformation of DppD, samples of DppD were incubated with 1 m*M* DTT, 0.5 m*M* GSH or 0.5 m*M* NADH for 5 min at 328 K. When recording the spectra, the actual concentration of the protein was 0.1 mg ml^{-1}. Far-UV CD spectroscopy was carried out using an Applied Photophysics Pistar-180 spectropolarimeter with a 1 mm path-length cell and a bandwidth of 3.0 nm. Spectra were recorded from 250 to 190 nm at an interval of 1 nm and were repeated three times. All resultant spectra were obtained by subtraction of the spectrum of the buffer. The results were expressed as the mean residue ellipticity $[\theta]_{MRW}$.

3. Results and discussion

3.1. Expression and purification of the DppD protein

The construct expressing *T. tengcongensis* DppD was successfully generated using a standard PCR-based cloning strategy. The recombinant protein was expressed in *E. coli* strain C43. DppD with a C-terminal 6×His tag was purified using an Ni Sepharose 6 Fast Flow column followed by gel filtration using a Superdex 200 HR 10/30 column.

The purified DppD protein was yellow and its colour became dark yellow when the protein was concentrated, suggesting the presence of a bound ligand molecule in the structure. Previous studies have suggested that the dipeptide/oligopeptide/nickel ABC transporter (DppBCDF) can transport haem (Abouhamad & Manson, 1994); we thus propose that the ligand might be a molecule containing iron. The existence and position of iron was examined in anomalous difference Fourier maps calculated from highly redundant data sets. The anomalous difference Fourier computed with the data collected at the iron peak revealed a bound iron–sulfur cluster in the structure (Fig. 1).

3.2. LTQ Orbitrap mass-spectrometric analysis

The purified DppD protein was analyzed by LTQ Orbitrap mass spectrometry. The resulting peptide mass fingerprint (PMF) of the protein was submitted to NCBInr. Only the ABC-type dipeptide/oligopeptide/nickel-transport system ATPase component (*T. tengcongensis* MB4) was obtained as a result of the search, with a score of 109. The results provided convincing evidence that the purified protein was the ABC-type dipeptide/oligopeptide/nickel-transport system ATPase component from *T. tengcongensis*.

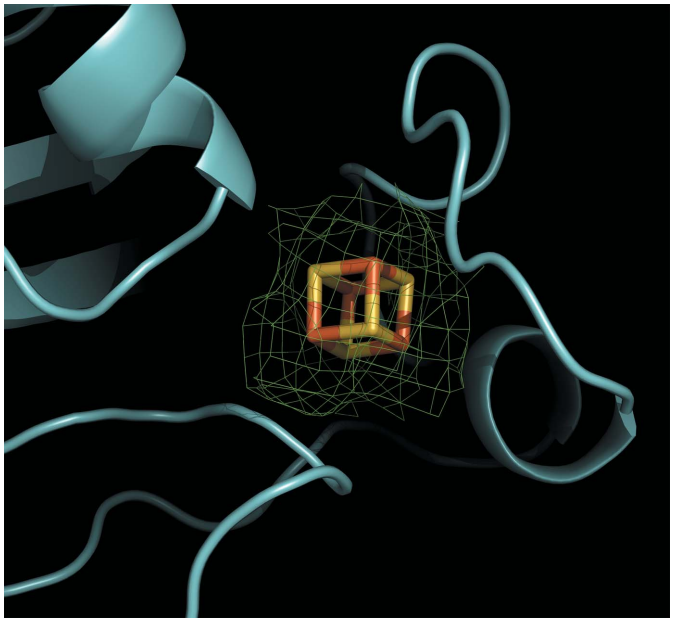


Figure 1
The difference Fourier map of *T. tengcongensis* DppD. The electron-density map for the [4Fe–4S] cluster is coloured green. The [4Fe–4S] cluster is shown as sticks and coloured red for Fe atoms and yellow for S atoms.

Table 1
Data-collection and refinement statistics.

	Data set 1†	Data set 2‡
Data-collection statistics		
Unit-cell parameters		
a (Å)	64.247	64.247
b (Å)	64.247	64.247
c (Å)	320.608	320.608
$\alpha = \beta = \gamma$ (°)	90	90
Space group	$P4_322$	$P4_322$
Wavelength (Å)	1.54148	1.54148
Resolution (Å)	50.00–2.89 (2.99–2.89)	50.00–2.89 (2.99–2.89)
Total No. of reflections	59602	110339
No. of unique reflections	15715	15883
Completeness (%)	97.5 (90.3)	98.2 (96.8)
Average $I/\sigma(I)$	24.57 (2.04)	22.06 (2.23)
R_{merge}§ (%)	8.9 (58.7)	10.2 (73.6)
Multiplicity	3.8 (3.4)	6.9 (6.6)
Wilson B factor (Å^2)	88.9	87.5
Refinement statistics		
No. of reflections used [$\sigma(F) > 0$]	14810	
R_{work}¶ (%)	21.19	
R_{free}¶ (%)	26.79	
R.m.s.d. from ideal bond lengths (Å)	0.015	
R.m.s.d. from ideal bond angles (°)	1.477	
No. of atoms		
Protein atoms	2425	
Ligand atoms	40	
Solvent atoms	0	
B value (Å^2)	88.6	
Ramachandran plot, residues in (%)		
Favoured regions	84.2	
Additional allowed regions	15.8	
Disallowed regions	0	

† For data set 1, all of the .x files were input into *SCALEPACK* for merging and scaling. ‡ For data set 2, the data were scaled separately and output as I^+ and I^- for each reflection for the detection of anomalous signal. § $R_{merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$. ¶ $R_{work} = \sum_{hkl} ||F_{obs}| - |F_{calc}|| / \sum_{hkl} |F_{obs}|$. R_{free} was calculated in the same way as R_{work} but for a test set of reflections that were not included in refinement.

3.3. Crystallization and structure determination

Small crystals were observed after about 2 d from one condition of the PEGRx 2 kit (Hampton Research) containing polyethylene glycol monomethyl ether 5000 as the precipitant. This condition was further optimized and harvestable crystals were obtained using the hanging-drop vapour-diffusion method by mixing 1 µl protein solution with 1 µl reservoir solution [4%(*v*/*v*) 2-propanol, 0.05 *M* CAPSO pH 11.0, 12.25%(*w*/*v*) polyethylene glycol monomethyl ether 5000] and equilibrating the drop against 200 µl reservoir solution. An X-ray diffraction data set was collected from a single crystal. Data-collection statistics and final refinement statistics are summarized in Table 1.

The structure was solved using an X-ray diffraction data set collected from a native crystal. The crystal belonged to space group $P4_322$, with unit-cell parameters $a = b = 64.247$, $c = 320.608$ Å, and there was one molecule in the asymmetric unit. The model was further improved by cycles of manual building and refinement using *Coot* and *PHENIX*. The structure was subsequently refined to a final resolution of

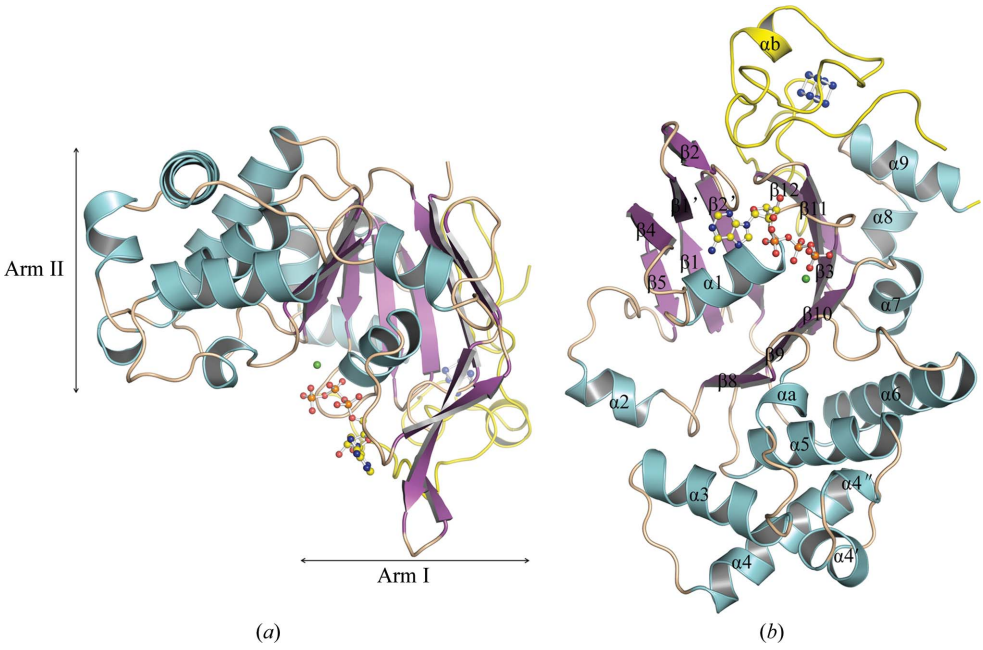


Figure 2
The three-dimensional structure of *T. tengcongensis* DppD. Cartoon representation of *T. tengcongensis* DppD in two views. ATP is coloured by element, Mg^{2+} in green and the iron–sulfur cluster in blue. The C-terminal domain is coloured yellow. (*a*) The L-shaped conformation of DppD with two thick arms (arm I and arm II); the top and the bottom of the protomer face towards the periplasmic and cytoplasmic sides, respectively. α-Helices (cyan) and β-strands (magenta) are differentiated by colour. (*b*) View from the bottom of the structure, as shown in (*a*), showing the ATP-binding pocket and the iron–sulfur cluster located in the C-terminal domain. The secondary-structural elements (α1–α9, αa, αb, β1–β5 and β8–β12) are labelled.

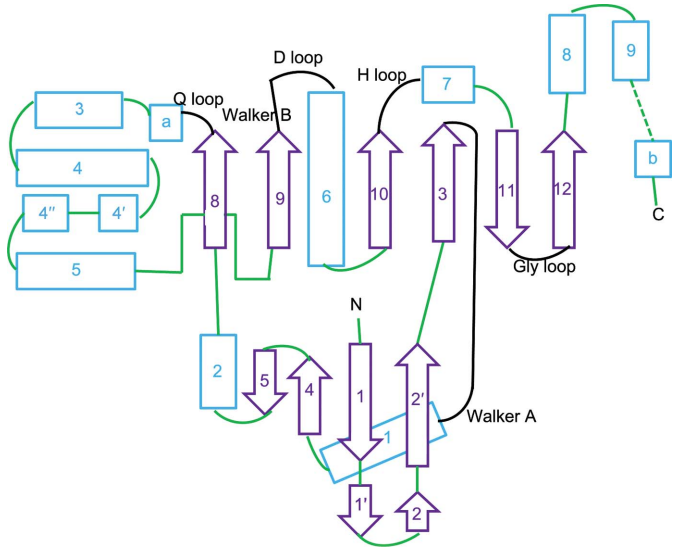


Figure 3
The topology of *T. tengcongensis* DppD. Helices are shown in cyan, β-strands in magenta and loops in green. ATP, Mg^{2+} and the iron–sulfur cluster are omitted for clarity. The conserved motifs are shown in black and labelled. The missing region between α9 and αb (residues 260–269) is represented as a dotted line.

2.89 Å with a working *R* value of 21.19% and a free *R* value of 26.79%. The final model of DppD consisted of residues 2–259 and 270–321. Residues 260–269 were not visible in electron-density maps. Analysis of the Ramachandran plot using *PROCHECK* (Laskowski *et al.*, 1993) showed that 84.2% of the modelled residues were in preferred and allowed regions.

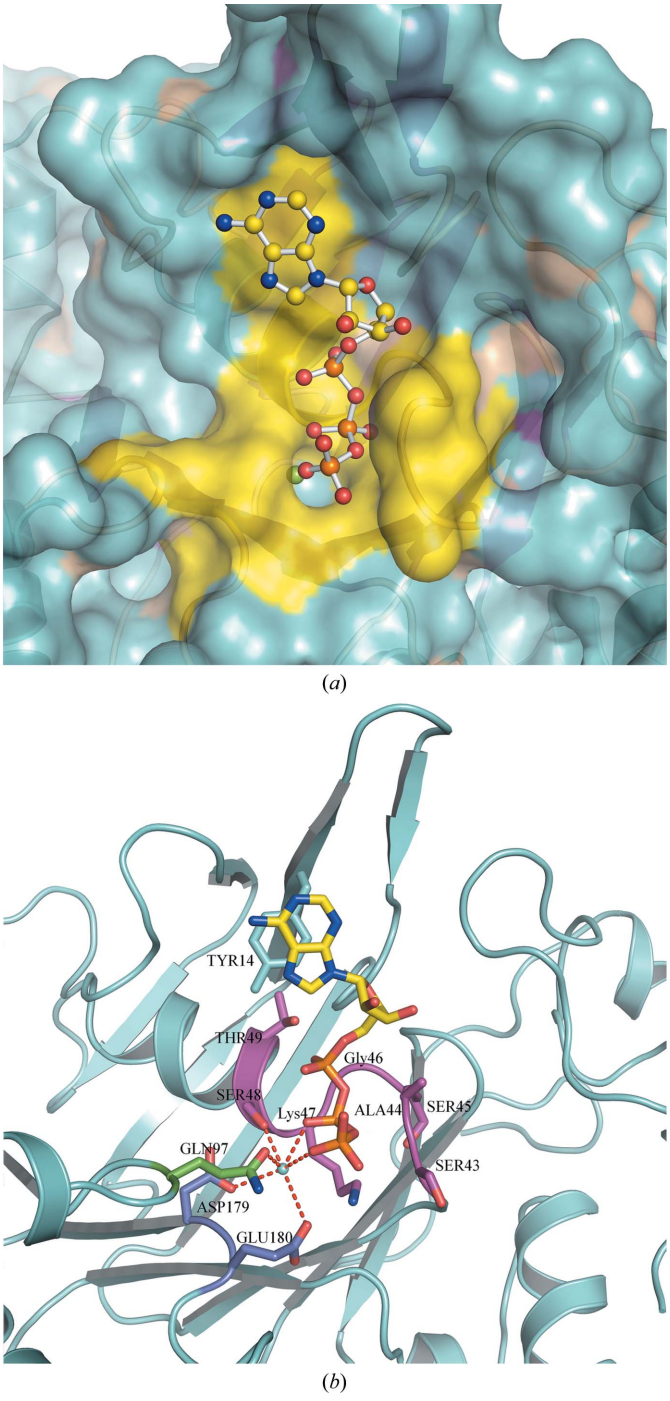
3.4. Structure description

3.4.1. Overall structure. *T. tengcongensis* DppD is composed of two domains: the conserved nucleotide-binding domain (residues 2–259), which is homologous to other NBDs (r.m.s.d. of ∼1.83 Å with MetN), and the C-terminal domain (residues 270–321) bound with a [4Fe–4S] cluster. Residues 260–269 are missing in the structure, probably because this segment, which links the two domains, is flexible. Although the linker is disordered, the C-terminal domain forms a stable loop region bound with a [4Fe–4S] cluster and interacts with the N-terminal domain (Fig. 2). A notable feature of the structure is that the C-terminal domain contains a [4Fe–4S] cluster with four Fe atoms covalently binding to Cys285, Cys291, Cys298 and Cys316. A *DALI* search (Holm & Sander, 1996) revealed no structural fold similar to this domain.

As first observed for HisP, the DppD ABC subunit adopts an L-shaped conformation with a two-arm architecture including 11 α-helices and ten β-sheets. The ATP-binding pocket is near the end of arm I, which consists of seven α-helices and four β-sheets. A six-stranded β-sheet (β3 and β8–β12) spans both arms, with a domain consisting of an α+β-type structure (β1, β2, β4, β5, α1 and α2) on one side (within arm I) and a domain with mostly α-helices (α3–α9, αa and αb) on the other side (arm II) of the sheet (Fig. 2). Arms I and II are connected from β-strand 4 to helix 3 *via* the Q-loop (residues 96–98) and are also connected by a small helix (α-helix a) and the following loop, which is different from other NBDs. The smaller arm II, also termed the signalling domain, is composed of five α-helices and contains the signature motif (residues 155–159; LSGGM), which extends into helix 5 (Holland & Blight, 1999). The whole topology is depicted in Fig. 3. The N-terminus of the structure shows the same ATPase topology and sequence motifs as previously observed in other structures of NBDs from ABC transporters.

3.4.2. ATP-binding pocket. The ATP-binding site observed in the structures of nucleotide-binding proteins such as Ras contains a phosphate-binding loop (P-loop; Saraste *et al.*, 1990). The residues involved in forming the ATP-binding site shape the active-site groove (Fig. 4*a*). ATP interacts with the

residues of the Walker A motif, Walker B motif and Q-loop. Both the Walker A (residues 41–49) and the Walker B (residues 175–180) motifs are located within arm I (Fig. 3). The P-loop (residues 43–49) surrounds the β-phosphate of ATP and forms extensive hydrogen bonds between its main chain and the β-phosphate. Tyr14 provides additional stabilization by hydrophobic interactions between its aromatic ring and the adenine base of ATP. Given the existence of magnesium in the crystallization conditions, it is highly probable that the ions observed in the active site of the structure are magnesium ions. The polar contacts of magnesium with other residues revealed that Asp179 is probably the residue that interacts with the divalent magnesium cation in the process of ATP hydrolysis when substrate is transported (Fig. 4*b*). The interaction of magnesium with the γ-phosphate of ATP can induce ATP hydrolysis when the substrate is transported. In this process, Gln97 or Glu180 could become activated and interact with the γ-phosphate of ATP by forming hydrogen bonds to water (Hung *et al.*, 1998). All of the residues forming hydrogen bonds or stacking with ATP are shown in Fig. 4(*b*).



(*a*)

(*b*)

Figure 4
The arrangement of ATP, Mg^{2+} and key amino-acid residues in the active sites present in *T. tengcongensis* DppD. (*a*) Cartoon representation and surface of *T. tengcongensis* DppD. The residues coloured yellow interact with ATP and Mg^{2+} and form the binding pocket. (*b*) ATP interacts with key amino acids of DppD. The Walker A motif (residues 43–49) is shown in magenta, the Walker B motif (residues 179–180) in blue and the Q-loop (residues 97) in green. The residues that form polar interactions with Mg^{2+} are shown.

3.4.3. C-terminal domain bound with an iron–sulfur cluster. In contrast to the structures of other NBDs such as MetN, the C-terminal domain of *T. tengcongensis* DppD shows a different structural fold (Fig. 5*b*): a V-shaped structure with a [4Fe–4S] cluster embedded in the turn region (Fig. 6*a*). The [4Fe–4S] cluster is coordinated in a tetrahedral geometry by Cys285, Cys291 and Cys316 from the loops between α9 and the C-terminus and by Cys298 from α-helix b (Fig. 6*b*). The residues surrounding the cluster are mostly aromatic amino acids. His248, His288 and His317 make hydrophobic interactions with the [4Fe–4S] cluster. The three histidines may act as ligands for iron when translocating the substrates, as in the [FeFe]-hydrogenase (Peters *et al.*, 1998). Tyr293 interacts with Cys291 and may facilitate substrate binding; thus, it is possibly involved in the substrate-transport process. In order to examine the possible variations in the NBDs of ABC proteins, the sequences and structures of known NBDs were aligned (Fig. 5). It was observed that the residues in the NBDs have very low sequence conservation, and no conserved cysteine residues which can bind to the iron–sulfur cluster were found. *T. tengcongensis* DppD forms a different fold from those of the available NBD structures. The C-terminal [4Fe–4S] cluster could stabilize the structure and we propose that it may function in the regulation of transport of dipeptides or haem.

3.5. Spectroscopic analysis

3.5.1. UV–visible absorbance spectroscopic analysis. In order to verify the existence of the modelled [4Fe–4S] in the crystal structure, the purified DppD protein was subjected to UV–visible absorbance spectroscopic analysis. The absorption spectrum showed a broad peak at ∼400–405 nm (Fig. 7) characteristic of [4Fe–4S] cluster proteins (Alam *et al.*, 2007), whereas the absorption spectrum of BSA showed no peak at ∼400–405 nm. Thus, it was concluded that the freshly purified DppD protein was coordinated with a [4Fe–4S] cluster.

3.5.2. Far-UV CD spectroscopic analysis. To test the conformational changes of the DppD protein upon ATP binding, we recorded far-UV CD spectra in the wavelength range 240–190 nm. Only by incubating the protein with ATP and Mg^{2+} simultaneously could we detect the shift of the

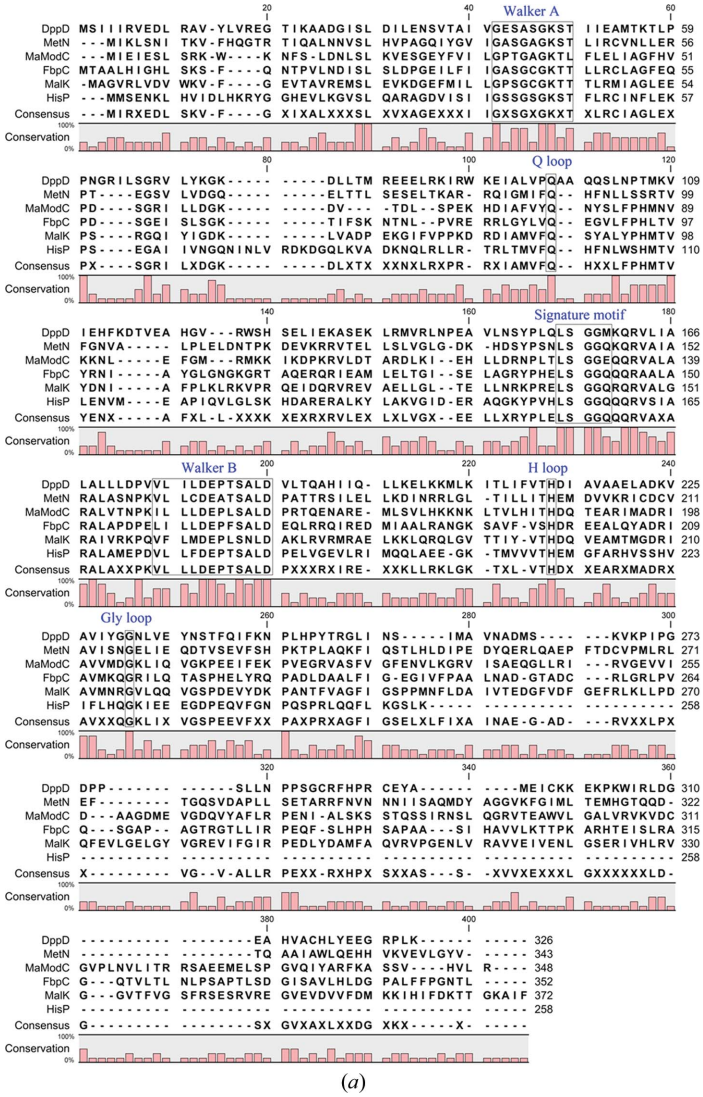
characteristic negative peaks for α-helices (208 and 222 nm; Fig. 8*a*), indicating that the hydrolysis of ATP could induce conformational changes in DppD.

Far-UV CD spectra were also collected to study the functional characteristics of the [4Fe–4S] cluster with different reductants. Incubation of DppD protein with DTT or GSH had little effect on its conformation, whereas the spectrum of the protein changed significantly on incubation with NADH (Fig. 8*b*), indicating that NADH may transfer electrons to the iron–sulfur cluster and cause conformational changes. We propose that the [4Fe–4S] cluster may participate in substrate (dipeptide or haem) binding and activation of the protein by electron transfer when transporting the substrate. However, the biological functions of the [4Fe–4S] cluster and the substrate-binding process still require further detailed investigation.

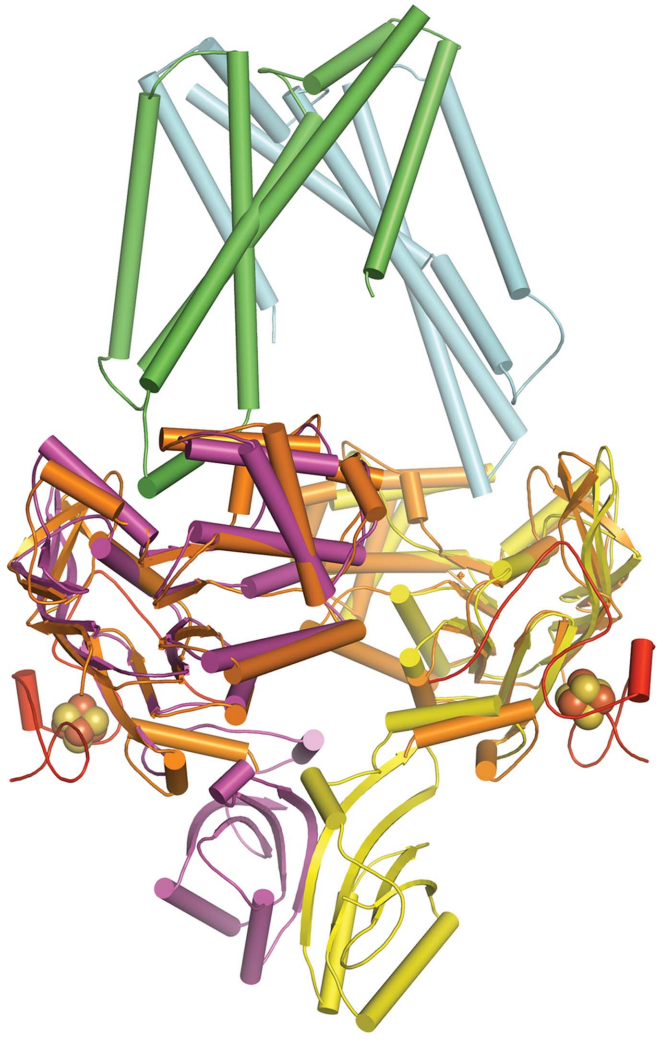
3.6. Comparisons with other NBD structures

Structural analysis of DppD by *BLAST* and *DALI* searches of the Protein Data Bank suggests that the molecule shares the common overall architecture of ABC-transporter NBDs. From sequence and structure alignments (Fig. 5), the N-terminal domain shares conserved motifs with the NBDs of other ABC transporters, but the C-terminal domain shows a notably different fold from those of other NBDs.

The transport of dipeptides or haem requires energy which may become available owing to binding and hydrolysis of ATP



(*a*)



(*b*)

Figure 5
Sequence and structural comparisons of *T. tengcongensis* DppD with other NBDs. (*a*) Sequence alignment of six available NBDs. The consensus sequence is indicated, the residue conservation is shown as a bar plot and the residues framed in black are conserved motifs. The names of the conserved motifs of the NBDs are indicated above the alignment. The alignment was computed by *CLC Sequence Viewer* 6.0.2 using the following sequences retrieved from NCBI and structural templates from the PDB: MetN (NP_414741.1; PDB entry 3dhw; Kadaba *et al.*, 2008), MaModC (NP_615255.1; 3d31; Gerber *et al.*, 2008), FbpC (ZP 04720438.1; 3fvq; Newstead *et al.*, 2009), MalK (AF121946_1; 1g29; Diederichs *et al.*, 2000) and HisP (NP_461293.1; 1b0u; Hung *et al.*, 1998). (*b*) Superposition of *T. tengcongensis* DppD onto the Met ABC transporter MetMN. The helices and strands of DppD are coloured orange. MetM is coloured green and cyan; MetN is coloured magenta and yellow. The [4Fe–4S] cluster is shown in sphere representation. The r.m.s.d. of DppD from MetN is ∼1.83 Å.

by the NBD. All NBDs consist of two constituent domains: an N-terminal domain (also named the RecA-like domain; Ye *et al.*, 2004), which is a catalytic core domain with conserved motifs (Walker A motif, Walker B motif, Q-loop and H-motif), and an α-helical domain. The conformations of the two domains are sensitive to the nucleotide-binding state (Karpowich *et al.*, 2001). It has been proposed that a subregion of the helical domain is related to the function and specific contacts of TMDs and NBDs (Schmitt *et al.*, 2003). In the maltose transporter, it was found that the conserved EAA motif of TMDs interacts with residues in the helical domain (Oldham *et al.*, 2007; Khare *et al.*, 2009). On binding and hydrolysis of ATP, the helical subdomain rotates toward the RecA-like subdomain, moving the signature motif into a favourable position to interact with ATP (Orelle *et al.*, 2010). The far-UV CD spectrum with ATP and Mg^{2+} showed a conformational change of DppD, which was likely to be the result of hydrolysis of ATP. The presence of Mg^{2+} may induce ATP hydrolysis to power the translocation.

A remarkable difference of *T. tengcongensis* DppD from other NBDs is its unique V-shaped C-terminal domain bound with a [4Fe–4S] cluster (Fig. 6*a*). The [4Fe–4S] cluster can covalently bind to the cysteine residues in the C-terminus, which may stabilize the structure of the C-terminal loops. Recent studies of both the methionine-uptake ABC transporter (Gerber *et al.*, 2008) and the molybdate/tungstate ABC transporter (Kadaba *et al.*, 2008) have established that substrates can bind to the extra domains at the C-terminus as allosteric regulators. A superposition of the structure of *T. tengcongensis* DppD with the Met ABC transporter MetN (Fig. 5*b*) revealed a root-mean-square deviation (r.m.s.d.) between equivalent C^{α} atoms in DppD and MetN of 1.83 Å. Interestingly, both of them have additional C-terminal unique domains, while the relative position of the C-terminal domain with respect to the N-terminal domain in the two proteins is different. As noted in *Bacillus subtilis* GPAT, the iron–sulfur cluster is not related to allosteric regulation and does not play a direct role in the enzymatic reaction. It appears to be

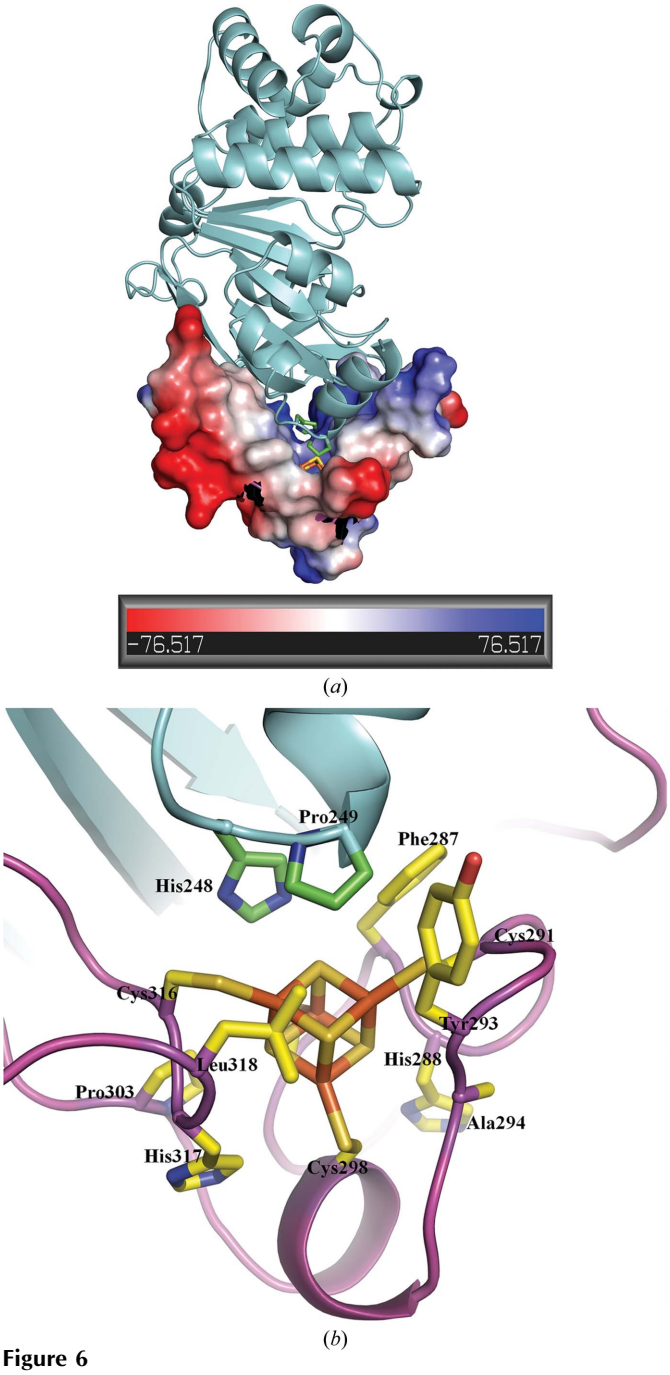


Figure 6
Interaction of the [4Fe–4S] cluster with residues in the C-terminal domain of *T. tengcongensis* DppD. Fe–S centres are shown as sticks; Fe atoms are coloured red and S atoms yellow. (*a*) The surface electrostatic potential of the V-shaped C-terminal domain. The N-terminal [4Fe–4S] cluster is in the turn of the V-shaped C-terminal domain. (*b*) Residues that interact with the [4Fe–4S] cluster in the C-terminal domain. The residues that are likely to be involved in interaction with the [4Fe–4S] cluster are shown in sticks and coloured yellow (from the C-terminal domain) or green (from the N-terminal domain).

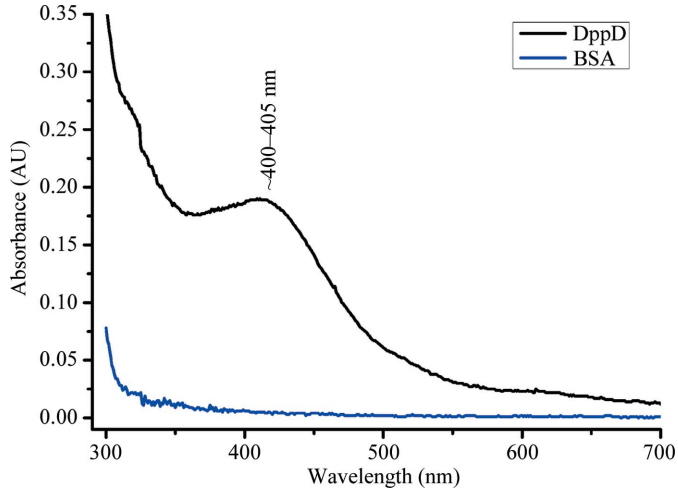


Figure 7
UV–visible absorbance spectroscopic analysis of *T. tengcongensis* DppD. The UV–visible absorption spectrum of freshly purified *T. tengcongensis* DppD shows the presence of a [4Fe–4S] cluster (black line). Purified *T. tengcongensis* DppD (0.8 mg ml^{-1}) and BSA (0.8 mg ml^{-1}) were solubilized in 50 m*M* sodium phosphate buffer pH 7.5. The black line represents the absorption spectrum of DppD. The broad peak (~400–405 nm) indicates the absorption spectrum of the [4Fe–4S] cluster protein. BSA has no absorption peak at this wavelength.

required for both amino-terminal processing and structural stability of the protein (Makaroff *et al.*, 1986; Grandoni *et al.*, 1989). It has also been reported that the ABC-type dipeptide/oligopeptide/nickel-transport system is actually the haem transporter in *E. coli* (Létoffé *et al.*, 2006). From the far-UV CD spectrum of DppD, we could conclude that the [4Fe–4S] cluster may participate in electron transfer. We propose that the [4Fe–4S] cluster can stabilize the protein by binding the substrate, thus stabilizing the transition state of the enzymatic reaction as in aconitase (Jarrett, 2005) in order to regulate substrate transport; thus, the C-terminal domain of DppD does not act as an allosteric regulator. The iron–sulfur cluster probably participates in substrate (dipeptide or haem) binding and activation of the protein by electron transfer during the process of transport. The cysteines, histidines and Tyr293 play important roles in this process. Iron–sulfur clusters are mostly found in eukaryotes and have been shown to play important roles in enzyme catalysis, electron transfer and regulation (Alam *et al.*, 2007). However, how the substrate is transported and the molecular mechanism of the competitive binding of haem and dipeptides require further investigation. Further research on the biological functions of the iron–sulfur cluster in this protein is in progress.

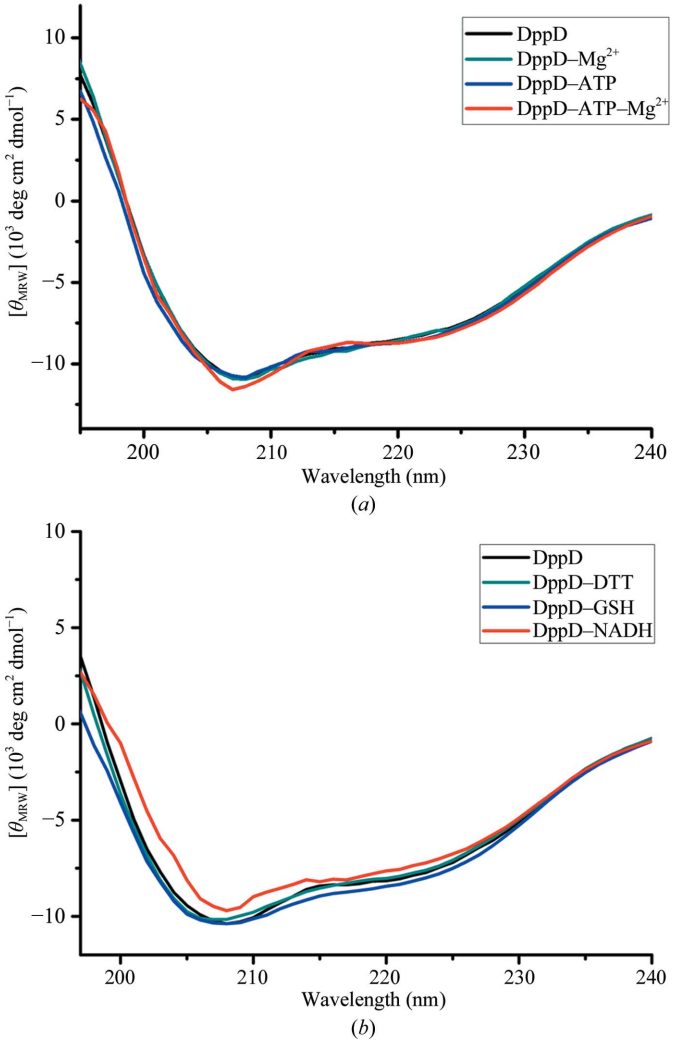


Figure 8
Far-UV circular-dichroism (CD) spectra of *T. tengcongensis* DppD with ATP, Mg^{2+} and various reductants. Purified *T. tengcongensis* DppD (0.1 mg ml^{-1}) was solubilized in 50 m*M* sodium phosphate buffer pH 7.5. (*a*) Far-UV CD spectra of DppD with Mg^{2+} (dark cyan), ATP (blue) and ATP–Mg^{2+} (red). (*b*) Far-UV CD spectra of DppD with DTT (dark cyan), GSH (blue) and NADH (red).

4. Conclusions

We expressed DppD, the nucleotide-binding domain of an ABC transporter, from *T. tengcongensis* in *E. coli* strain C43 and determined its crystal structure bound with ATP, Mg^{2+} and a [4Fe–4S] cluster at 2.89 Å resolution. Structural analysis and comparisons with other ABC transporters revealed that the C-terminus of *T. tengcongensis* DppD forms a novel structural fold and is bound with a [4Fe–4S] cluster. The iron–sulfur cluster may participate in substrate binding to regulate and stabilize the protein in the ABC transport system. The three-dimensional structure of *T. tengcongensis* DppD will help us to further understand the properties of ABC transporters, the functions of NBDs and the regulation of the iron–sulfur cluster. The biological relevance of this iron–sulfur cluster in transporting substrates requires further investigation.

We would like to thank the staff at the SSRF BL17U beamline for their assistance in data collection. This work was supported by the Ministry of Science and Technology (2011CB910501 and 2012CB911101) and the National Natural Science Foundation (31030020 and 31170679). This work was supported by grants from the State Key Laboratory Special fund (2060204) and the National Important Research Plan of China (2012BAI31B08, 2012CB944902, and 2011CB944302) from the Ministry of Science and Technology of China.

References

Abouhamad, W. N. & Manson, M. D. (1994). *Mol. Microbiol.* **14**, 1077–1092.
Adams, P. D., Grosse-Kunstleve, R. W., Hung, L.-W., Ioerger, T. R., McCoy, A. J., Moriarty, N. W., Read, R. J., Sacchettini, J. C., Sauter, N. K. & Terwilliger, T. C. (2002). *Acta Cryst.* D**58**, 1948–1954.
Alam, M. S., Garg, S. K. & Agrawal, P. (2007). *Mol. Microbiol.* **63**, 1414–1431.
Bradford, M. M. (1976). *Anal. Biochem.* **72**, 248–254.
Cowtan, K. (2006). *Acta Cryst.* D**62**, 1002–1011.
Dean, M. & Allikmets, R. (1995). *Curr. Opin. Genet. Dev.* **5**, 779–785.
Dean, M. & Annilo, T. (2005). *Annu. Rev. Genomics Hum. Genet.* **6**, 123–142.
Dean, M., Hamon, Y. & Chimini, G. (2001). *J. Lipid Res.* **42**, 1007–1017.
Diederichs, K., Diez, J., Greller, G., Muller, C., Breed, J., Schnell, C., Vonrhein, C., Boos, W. & Welte, W. (2000). *EMBO J.* **19**, 5951–5961.
Doige, C. A. & Ames, G. F. (1993). *Annu. Rev. Microbiol.* **47**, 291–319.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
Ferguson, A. D. & Deisenhofer, J. (2004). *Cell*, **116**, 15–24.
Gerber, S., Comellas-Bigler, M., Goetz, B. A. & Locher, K. P. (2008). *Science*, **321**, 246–250.
Grandoni, J. A., Switzer, R. L., Makaroff, C. A. & Zalkin, H. (1989). *J. Biol. Chem.* **264**, 6058–6064.
Higgins, C. F. (1992). *Annu. Rev. Cell Biol.* **8**, 67–113.

Holland, I. B. & Blight, M. A. (1999). *J. Mol. Biol.* **293**, 381–399.
Holm, L. & Sander, C. (1996). *Science*, **273**, 595–603.
Hung, L.-W., Wang, I. X., Nikaido, K., Liu, P.-Q., Ames, G. F.-L. & Kim, S.-H. (1998). *Nature (London)*, **396**, 703–707.
Jarrett, J. T. (2005). *Arch. Biochem. Biophys.* **433**, 312–321.
Kadaba, N. S., Kaiser, J. T., Johnson, E., Lee, A. & Rees, D. C. (2008). *Science*, **321**, 250–253.
Karpowich, N., Martsinkevich, O., Millen, L., Yuan, Y.-R., Dai, P. L., MacVey, K., Thomas, P. J. & Hunt, J. F. (2001). *Structure*, **9**, 571–586.
Khare, D., Oldham, M. L., Orelle, C., Davidson, A. L. & Chen, J. (2009). *Mol. Cell*, **33**, 528–536.
Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
Létoffé, S., Delepelaire, P. & Wandersman, C. (2006). *Proc. Natl Acad. Sci. USA*, **103**, 12891–12896.
Linton, K. J. & Higgins, C. F. (1998). *Mol. Microbiol.* **28**, 5–13.
Makaroff, C. A., Paluh, J. L. & Zalkin, H. (1986). *J. Biol. Chem.* **261**, 11416–11423.
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.
Moitra, K., Lou, H. & Dean, M. (2011). *Clin. Pharmacol. Ther.* **89**, 491–502.
Newstead, S., Fowler, P. W., Bilton, P., Carpenter, E. P., Sadler, P. J., Campopiano, D. J., Sansom, M. S. & Iwata, S. (2009). *Structure*, **17**, 1213–1222.
Oldham, M. L., Khare, D., Quiocho, F. A., Davidson, A. L. & Chen, J. (2007). *Nature (London)*, **450**, 515–521.
Orelle, C., Alvarez, F. J., Oldham, M. L., Orelle, A., Wiley, T. E., Chen, J. & Davidson, A. L. (2010). *Proc. Natl Acad. Sci. USA*, **107**, 20293–20298.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Peters, J. W., Lanzilotta, W. N., Lemon, B. J. & Seefeldt, L. C. (1998). *Science*, **282**, 1853–1858.
Saraste, M., Sibbald, P. R. & Wittinghofer, A. (1990). *Trends Biochem. Sci.* **15**, 430–434.
Schmitt, L., Benabdelhak, H., Blight, M. A., Holland, I. B. & Stubbs, M. T. (2003). *J. Mol. Biol.* **330**, 333–342.
Sheftel, A., Stehling, O. & Lill, R. (2010). *Trends Endocrinol. Metab.* **21**, 302–314.
Sheldrick, G. M. (2008). *Acta Cryst.* A**64**, 112–122.
Stojiljkovic, I. & Hantke, K. (1994). *Mol. Microbiol.* **13**, 719–732.
Voloshyna, I. & Reiss, A. B. (2011). *Prog. Lipid Res.* **50**, 213–224
Walker, J. E., Saraste, M., Runswick, M. J. & Gay, N. J. (1982). *EMBO J.* **1**, 945–951.
Wandersman, C. & Delepelaire, P. (2004). *Annu. Rev. Microbiol.* **58**, 611–647.
Wang, J., Xue, Y., Feng, X., Li, X., Wang, H., Li, W., Zhao, C., Cheng, X., Ma, Y., Zhou, P., Yin, J., Bhatnagar, A., Wang, R. & Liu, S. (2004). *Proteomics*, **4**, 136–150.
Wang, G. (2010). *J. Biol. Chem.* **285**, 40438–40447.
Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.
Ye, J., Osborne, A. R., Groll, M. & Rapoport, T. A. (2004). *Biochim. Biophys. Acta*, **1659**, 1–18.