Article

AMP binding stabilizes the KTN domain of the Shewanella denitrificans Kef potassium efflux system

Christos Pliotas, Samuel C Grayer, Silvia Ekkerman, Anthony KN Chan, Jess Healy, Phedra Marius, Wendy Bartlett, Amjad Khan, wilian augusto cortopassi, Shane A Chandler, Tim Rasmussen, Justin LP Benesch, Robert S. Paton, Timothy D. W. Claridge, Samantha Miller, Ian Rylance Booth, James H. Naismith, and Stuart J. Conway

Biochemistry, **Just Accepted Manuscript** • DOI: 10.1021/acs.biochem.7b00300 • Publication Date (Web): 28 Jun 2017

AMP binding stabilizes the KTN domain of the *Shewanella denitrificans* Kef potassium efflux system

Christos Pliotas,[a,†] Samuel C. Grayer,[b,†] Silvia Ekkerman,[c,†] Anthony K. N. Chan,[b,†] Jess Healy,[b,‡] Phedra Marius,[a] Wendy Bartlett,[c] Amjad Khan,[b] Wilian A. Cortopassi,[b] Shane A. Chandler,[d] Tim Rasmussen,[c] Justin L. P. Benesch,[d] Robert S. Paton,[b] Timothy D. W. Claridge,[b] Samantha Miller,[b] Ian R. Booth,[c,*] James H. Naismith[a,e,f,g,*] and Stuart J. Conway[b,h,*]

[a]*Biomedical Sciences Research Complex, University of St Andrews, North Haugh, St Andrews, KY16 9ST, UK*

[b]*Department of Chemistry, Chemistry Research Laboratory, University of Oxford, Mansfield Road, Oxford, OX1 3TA, UK*

[c]*School of Medicine, Medical Sciences and Nutrition, Foresterhill, Aberdeen, AB25 2ZD, UK*

[d]*Department of Chemistry, Physical & Theoretical Chemistry Laboratory, University of Oxford, South Parks Road, Oxford, OX1 3QZ, UK*

[e]*Biotherapy Centre, Sichuan University, Chengdu, China.*

[f]*RCaH, Rutherford Appleton Laboratory, Harwell Oxford, Didcot, OX11 0FA.*

[g]*Division of Structural Biology, University of Oxford, Henry Wellcome Building for Genomic Medicine, Old Road Campus, Roosevelt Drive, Headington, Oxford OX3 7BN.*

[h]*Freiburg Institute for Advanced Studies – FRIAS, Albert-Ludwigs-Universität Freiburg, Albertstr. 19, 79104 Freiburg, Germany.*

†These authors contributed equally to this work.

‡Current address: *Department of Pharmaceutical and Biological Chemistry, UCL School of Pharmacy, University College London, 29/39 Brunswick Square, WC1N, 1AX, UK*

*To whom correspondence should be addressed.

Stuart Conway: stuart.conway@chem.ox.ac.uk; +44(0)1865 285109.

James Naismith: naismith@strubi.ox.ac.uk; +44(0)1235567701.

Ian Booth: i.r.booth@abdn.ac.uk; +44 (0)1224 437396.

Abstract

Ligand binding is one of the most fundamental properties of proteins. Ligand functions fall into three basic types: substrates, regulatory molecules, and co-factors essential to protein stability, reactivity, or enzyme-substrate complex formation. The regulation of potassium ion movement in bacteria is predominantly under the control of regulatory ligands that gate the relevant channels and transporters, which possess subunits or domains that contain Rossmann folds (RFs). Here we demonstrate that AMP is bound to both RFs of the dimeric bacterial Kef potassium efflux system (Kef), where it plays a structural role. We conclude that AMP binds with high affinity ensuring that the site is fully occupied at all times in the cell. Loss of the ability to bind AMP, we demonstrate, causes protein, and likely dimer, instability and consequent loss of function. Regulation of Kef system function is *via* the reversible binding of comparatively low affinity glutathione-based ligands at the interface between the dimer subunits. We propose this interfacial binding site is itself stabilised, at least in part, by AMP binding.

Keywords

Biophysical studies, crystallography, protein dimers, membrane protein, nucleotide

A fundamental property of proteins, by which they express their function in the cell, is the binding of ligands, usually ions or molecules of small mass relative to that of the protein itself. At least three different roles are ascribed to the binding of ligands: firstly, the ligand is a substrate or essential cofactor for an enzyme; secondly, ligand binding may be purely regulatory bringing about changes in protein activity; and thirdly ligands may stabilize a protein fold. The roles are not mutually exclusive and can be combined. The activation of ligand-gated channels is usually brought about by an allosteric transition upon ligand binding at a point distant from the pore; the changes in the concentration of the ligand may reflect the biological state of either the cell or the environment. Regulation of ion flow is critical throughout biology and, for potassium ions (K^+), ligand-gated channels and transporters are central to the modulation of cellular K^+ pools.

In bacteria, both K^+ channels, which play an important role in K^+ influx, and K^+ efflux systems that control the response of the bacterial cell to electrophiles, have K^+ transport and NAD-binding (KTN) domains. These domains may be covalently attached to the pore or may be separately expressed entities that form non-covalent associations with the pore. Some are hybrid systems utilizing a combination of KTN domains that are part of the integral pore-forming subunit together with KTN domains expressed separately from an internal start codon on the same mRNA as the pore protein.[1,2] Another major structural variation is that for the channels (e.g. TrkAH, KtrAB and MthK) octameric rings of KTN domains modulate ion flux, whereas for the Kef systems, dimeric assemblies dominate the known architectures. A conserved feature of KTN domains is a Rossmann fold, a feature known to be

associated with nucleotide binding since its first identification in NAD(H)-binding lactate dehydrogenase.[3]

In *Escherichia coli*, and most other bacteria, multiple transport systems and channels effect control over the K^+ pool, including the Kef systems, which are gated by glutathione (GSH) and its electrophilic conjugates (GSX). In Gram-positive bacteria similar protective systems might exist which employ electrophilic conjugates of other species-specific thiols,[4] for example bacillithiol.[5-7] Whereas the activity of most K^+ transport systems causes modulation of the cytoplasmic pH in the alkaline direction, the Kef systems cause acidification in response to cell-damaging electrophiles.[8,9] The ~600 residue Kef proteins form dimers of an ~380 amino acid membrane domain, which may contain up to 12 trans-membrane spans, although these are poorly defined from a structural perspective. A short hydrophilic linker (20-26 amino acids) connects the membrane domain to two further domains: an ~150 residue KTN domain, and a further, less well-conserved, domain of variable length at the extreme C-terminus of the protein. The KTN domains of separate proteins dimerize and the interface between them contains the GSH binding site. Gating of the K^+ efflux system requires GSH/GSX ligand-mediated communication between the C-terminal domains and the loop containing the ion flow regulating HALESDIEP sequence.[10]

The Kef proteins can be broadly divided into two classes, those including *E. coli* KefC and KefB that require an ancillary protein (KefF[11] and KefG for KefC and KefB, respectively) for full function and those, such as *Shewanella denitrificans*, that do not require an ancillary protein. Gating by GSH/GSX is thought to be almost identical in both protein types. Residues in the predicted GSX binding site of *S.*

denitrificans Kef, identified by sequence alignment and modeling, will likely play similar roles to their *E. coli* KefC counterparts, which were identified by molecular genetic studies and crystallography. The *E. coli* KefC protein has been difficult to study biochemically due to the instability of the KefF-KefC complex, thus our studies have focused on the simpler *S. denitrificans* Kef protein.[12]

The Rossmann folds of the KTN domains in the K^+ uptake systems, TrkAH and KtrAB, have been studied biochemically and shown by crystallography to bind ATP and NADH.[13] These uptake systems have K^+-permeable pores with pseudo-four-fold symmetry to create a pore reminiscent of the classical P-type K^+ channels. Twin pores, arising from separate membrane proteins in the dimer, form associations with an octameric assembly of KTN domains. Binding of ATP and/or NADH modulate the conformation of the octameric rings and regulate the opening of the K^+-conducting pore. ADP and NAD^+ activate the GsuK potassium channel, via its KTN-related RCK domains, whereas Ca^{2+} serves as an allosteric inhibitor.[14] In contrast, the structural basis of regulation of other KTN domain-regulated K^+ channels (e.g. CglK, Kch, MthK) by nucleotides is poorly understood; although reversible gating by divalent cations has been described for MthK, the role of nucleotide binding in the RF is unknown.[15] Similarly, for the GSX-gated Kef systems the role of the bound nucleotide is unclear.

Previously obtained X-ray crystal structures of the KTN domain of trkA from *Methanocaldococcus jannaschii* and the KTN domain of ktrA from *Bacillus subtilis* have electron density consistent with NADH bound in the Rossmann fold.{Roosild:2002tc} Based on these data a homology model was constructed that

had NADH modelled into the Rossmann fold of Kef the *E. coli* KefFC KTN domain.{Roosild:2002tc} In subsequent crystallographic studies on the nucleotide pocket of KTN domain of the *E. coli* KefFC KTN domains,[10] it is suggested that NADH occupies the nucleotide-binding pocket based on the homology model and the presence of this nucleotide in the crystallization liquor, although density only exists that is consistent with a bound AMP molecule. Subsequently, when structures with GSH and GSX were solved AMP was modeled into the GSH structure, but no density consistent with a nucleotide was observed in the GSX structure (density consistent with sulfate ions was observed in the RF of the GSX structure).[16] The uncertainty over the identity of the bound ligand, and the lack of any insight into the role of the bound nucleotide prompted us to re-examine the system in more detail.

Here we report the structural analysis of the KTN domain from *S. denitrificans*. We have established that both the isolated ligand-binding C-terminal domain (*Sd*KefCTD) and full-length integral membrane protein (*Sd*Kef) contain AMP when purified after over-expression in *E. coli*. Differential scanning fluorimetry (DSF, also known as thermal shift) analysis shows that incubation of *Sd*KefCTD with additional AMP results in major stabilizing effects on the protein. NADH although unable to displace AMP from the KTN domain results in some stabilization in DSF studies. In the isolated soluble *Sd*KefCTD KTN domain, introduction of mutations predicted to affect AMP binding led to lower protein expression, consistent with a role for AMP in protein structural integrity. In agreement with this proposal, molecular dynamics simulations indicate reduced stability of the *Sd*KefCTD domain when AMP was removed. In the full-length protein, the same mutations yielded inactive channels. We

propose that AMP is integral to the KTN domain in *Sd*Kef and is required for the stable and functional Kef dimer complex.

Materials and methods

Materials

n-Dodecyl-β-D-maltopyranoside (DDM) was purchased by Anatrace (www.affymetrix.com). Glutathione (reduced) GSH, adenosine monophosphate (AMP), nicotinamide adenine dinucleotide (NAD^+) and reduced nicotinamide adenine dinucleotide (NADH) were ordered from Fisher. Reagents for buffer and other chemicals were purchased from Sigma unless otherwise stated.

Expression and purification of Kef

The KTN construct, denoted as *Sd*KefQCTD, has been characterized previously and contains residues 391-607 of the full length *Sd*Kef protein, including the KTN domain, the carboxy-terminal peripheral domain, the highly charged Q-linker connecting the *Sd*KefQCTD with the transmembrane domains and a peptide corresponding to the regulatory HELEVDIEP loop, plus a C-terminal LEH_6 tag.[12,25] The *Sd*KefQCTD construct was transformed into the *E. coli* strain BL21(DE3) (www.bioline.com). Cells were grown in 500 mL of LB medium at 37 °C to an $OD_{600nm} \approx 0.8$. The cultures were cooled to 25 °C and induced with 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) for 4 h. The cell pellet was resuspended in lysis buffer, 50 mM Tris-HCl buffer, pH 7.8, 300 mM KCl, 40 mM imidazole, 10% glycerol and 1 mM benzamidine. After disruption of the cells with a French press at 18 000 psi, the suspension was centrifuged at 4 000 *g* for 20 min to remove cell debris. The supernatant was then centrifuged at 100 000 *g* for 1 h. The supernatant was then filtered using 0.45 μm diameter filters and was passed through a 25 mL column containing 0.5 mL of nickel-nitrilotriacetic acid (Ni^{2+}-NTA) agarose, at 4 °C. The column was washed with 15 mL of wash buffer, 50 mM Tris-HCl buffer, pH 7.8,

300 mM KCl, 50 mM imidazole, 10% glycerol and 1 mM benzamidine, to remove non-specifically bound proteins and was left at 4 °C overnight. Next morning the elution followed with 10 mL of elution buffer, 50 mM Tris-HCl buffer, pH 7.8, 300 mM KCl, 300 mM imidazole and 0.5 mL fractions were collected. The fractions were analyzed by SDS−PAGE and *UV/vis* absorption spectroscopy and the fractions with the highest protein content from the Ni^{2+}-NTA column were applied to a 120 mL Superose 6 column (General Electrics Healthcare) equilibrated with buffer containing 50 mM Tris-HCl, pH 7.8 and 300 mM KCl. Protein was then eluted at 1 mL/min flow rate. Protein concentration was monitored by absorption at 280 nm. The column was calibrated with Biorad standards. Identity and integrity were confirmed by mass spectrometry (Figure S3).

Alternative protocol

The *E. coli* strain, MJF373,[12] was used to express the *Sd*KefQCTD protein construct, which is encoded in the pTrc*Sd*KefQCTDH_6 plasmid.[12] The expression of the *Sd*KefQCTD protein is inducible by addition of IPTG. For recombinant protein expression, *E. coli* MJF373 host was first transformed with the pTrc*Sd*KefQCTDH_6 plasmid. The resultant transformant was aerobically cultured in the 2× TY medium [16 g/L Oxoid™ Tryptone (Thermo Fisher Scientific Inc.), 10 g/L Oxoid™ Yeast Extract (Thermo Fisher Scientific Inc.), 5 g/L NaCl] at 30 °C with an agitation speed of 180 r.p.m., in the presence of 100 μg/mL of ampicillin (Apollo Scientific Ltd.). When solid medium is required, Bacto™ Agar (BD) was added to a final concentration of 1.5% (*w/v*) in the 2× TY medium. When the bacterial culture reached an optical density of 1.0 at 600 nm IPTG (Apollo Scientific Ltd.), was added to a final concentration of 0.8 mM to induce expression of the recombinant protein. Bacterial

cell pellets were then harvested by centrifugation (Rotor: F10BCI-6x500y, Avanti™ J-25 Centrifuge, Beckman Coulter Inc) at 11305 *g* and 4 °C after 4h of IPTG post-induction, and kept at –80 °C until protein purification. The *Sd*KefQCTD protein was subsequently purified by immobilized metal affinity chromatography (IMAC) and then size-exclusion chromatography (SEC) at 4 °C.

To prepare a sample for protein purification, the frozen cell pellet (7.70 g) was first resuspended in Extraction Buffer. Extraction Buffer was prepared by completely dissolving one tablet of Sigma*FAST*™ Protease Inhibitor Cocktail Tablet, EDTA Free in 100 mL of solution containing 50 mM NaH_2PO_4/Na_2HPO_4, 500 mM NaCl, 10% glycerol, pH 7.4. A hundred milliliter of Extraction Buffer per 20 g of cell mass was used for resuspension. After complete resuspension of the bacterial pellets, the cells were lysed on ice by sonication (50% amplitude, 5 s bursts interrupted by 5 s pauses for 60 cycles; Ultrasonic Processor, Sonics & Materials, Inc.) to release cytosolic proteins. Polyethyleneimine at a final concentration of 0.15% (*v/v*) was added from 5% (*v/v*) stock solution (pH 7.4) to the cell lysates and the mixture was incubated on ice for 15 min to precipitate DNA. Insoluble cell debris and precipitated DNA were removed by centrifugation (Rotor: JA25.50, Avanti™ J-25 Centrifuge, Beckman Coulter Inc.) at 25 000 *g* and 4 °C for 15 min. The resulting supernatant was collected and filtered through 0.45 μm pore size syringe filters (Merck Millipore Corp.). The clarified cell lysate was diluted to 100 mL with Extraction Buffer, and then added with 20 mM imidazole (Sigma-Aldrich Co.) and 10 mM β-mercaptoethanol (Bio-Rad Laboratories, Inc) at final concentrations. This cell lysate preparation was used for the first step of protein purification by IMAC.

For protein purification by IMAC, a HisTrap HP 5 mL column (GE Healthcare) was used to purify hexa-histidine-tagged *Sd*KefQCTD. The affinity purification columns were connected to a computerized ÄKTAFPLC system (GE Healthcare). To equilibrate the column for IMAC, 10 column volumes (CV) of Binding Buffer (50 mM NaH_2PO_4/Na_2HPO_4, 500 mM NaCl, 10% Glycerol, 10 mM β-mercaptoethanol, pH 7.4) were used. The pretreated cell lysates (100 mL; from the above sample preparation procedure) were then loaded into the equilibrated HisTrap HP column. Following the sample loading into the column, Binding Buffer and Elution Buffer (50 mM NaH_2PO_4/Na_2HPO_4, 500 mM NaCl, 10% glycerol, 500 mM imidazole, pH 7.4) were mixed in different ratios to wash out non-specifically bound proteins and elute the protein of interest. Firstly, 10 CV of a step gradient containing 9% Elution Buffer (with 45 mM imidazole) and then a linear gradient ranged from 9% to 30% Elution Buffer (containing up to 150 mM imidazole) over 10 CV were used to wash out contaminating binders. For elution of the polyhistidine-tagged *Sd*KefQCTD protein, 5 CV of 60% Elution Buffer (containing 300 mM imidazole) were used to elute the target in 2 mL fractions.

After the IMAC purification step, *Sd*KefQCTD was further purified by SEC. The IMAC-purified *Sd*KefQCTD protein was first concentrated by using Vivaspin™ sample concentrator (GE Healthcare). The concentrated protein sample (2 mL) was then loaded *via* a 2 mL injection loop into a gel filtration column XK 16/70 (GE Healthcare) packed with 120 mL of Superdex 75 resin (GE Healthcare). This SEC column was pre-equilibrated with 150 mL degassed SEC Buffer (50 mM NaH_2PO_4/Na_2HPO_4, 150 mM NaCl, pH 7.4). The SEC procedure was operated at a constant flow rate (1 mL/min) over 150 mL of a total flow volume and filtrates were

collected in 5-mL fractions (the sample collection was started at 20th mL and stopped at 120th mL). The SEC-purified protein was further concentrated by Vivaspin^{TM} sample concentrator (GE Healthcare) after the purification process. Protein concentrations were determined by a NanoDrop 1000 spectrophotometer (Thermo Fisher Scientific Inc.).

The chemicals, NaH_2PO_4 and Na_2HPO_4, were purchased from Alfa Aesar and BDH Chemicals Ltd., respectively. Glycerol and NaCl were bought from Thermo Fisher Scientific Inc. All other chemicals were purchased from Sigma-Aldrich Co. unless otherwise specified. All of the buffers used in the purification procedures were pre-filtered through 0.2-μm pore size filter papers (Sartorius UK Ltd.) under vacuum to remove insoluble precipitates.

Western blot of membrane and soluble fractions containing full-length *Sd*Kef or *Sd*KefQCTD.

Membrane and soluble protein fractions were prepared from MJF373 cells transformed with either pTrc*Sd*KefH_6 or pTrc*Sd*KefQCTDH_6. Cells were culture overnight in LK medium (10 g/L tryptone, 5g/L yeast extract, 6.4 g/L KCl) with ampicillin (50 μg/mL) and diluted next morning to $O.D._{650}$ = 0.05 into a fresh LK medium as a preculture. Once cells had reached an $O.D._{650}$ of 0.4, they were diluted 10-fold into fresh LK medium and grown again until an $O.D._{650}$ of 0.4, when 0.3mM IPTG was added for induction of expression for 30 min, after which 100 ml cells were harvested by centrifugation, resuspended in PBS containing protease inhibitor cocktail tablet (Roche) and lysed by passage through a French Press at 18 000 psi. Bulk cell debris was removed by centrifugation for 10 minutes at 4 °C 4500 × g and

membrane (pellet) and soluble fractions (supernatant) collected after ultracentrifugation at 90 000 x g, 4 C for 60 min. The pellet was suspended in PBS and Lowry[26] estimation of protein concentration performed on the pellet and supernatant. Samples were separated on 4-12 % Invitrogen™ NuPAGE Bis-Tris gels (ThermoFisher Scientific) run in MES buffer and using SeeBlue Plus 2 Marker. Proteins were transferred onto nitrocellulose membranes and probed for expression using anti-His HRP conjugate antibody (Qiagen). SuperSignal™ West Dura Extended Duration Substrate (ThermoFisher Scientific) was used for ECL detection of bands, exposed to Amersham Hyperfilm™ ECL film (GE Healthcare) developed on an M35 X-OMAT processor.

Expression and purification of the full-length protein *Sd*Kef

The full-length membrane protein, was transformed into the *E. coli* strain BL21(DE3). Cells were grown the same as for *Sd*KefQCTD. The cell pellet was resuspended in buffer 50 mM Tris-HCl buffer, pH 7.8, 300 mM KCl, 1 mM Benzamidine. After disruption of the cells with a French press at 18,000 psi, the suspension was centrifuged at 4 000 *g* for 20 min to remove cell debris. The supernatant was then centrifuged at 100 000 *g* for 1 h. The pellet which contained the cell membrane was solubilized in solubilization buffer 50 mM Tris-HCl buffer, pH 7.8, 300 mM KCl, 1 mM Benzamidine, 1.5% DDM, 10% glycerol, 25 mM imidazole by using a homogenizer and left for an hour gently shaking, at 4 °C. The solubilized sample was subsequently passed through a 25 mL column containing 0.5 mL of nickel-nitrilotriacetic acid (Ni^{2+}-NTA) agarose, at 4 °C. The column was washed with 15 mL of wash buffer; 50 mM Tris-HCl buffer, pH 7.8, 300 mM KCl, 1 mM benzamidine, 0.05% DDM, 35 mM imidazole to remove non-specifically bound

proteins and was left at 4 °C overnight. Next morning the elution followed with 10 mL of elution buffer; 50 mM Tris-HCl buffer, pH 7.8, 300 mM KCl, 0.05% DDM, 300 mM imidazole and 0.5 mL fractions were collected. The fractions were analyzed by SDS−PAGE and UV/vis absorption spectroscopy and the highest fractions were applied to a 120 mL Superose 6 column (General Electrics Healthcare) equilibrated with buffer containing 50 mM Tris-HCl buffer, pH 7.8, 300 mM KCl, 0.05% DDM. Protein was then eluted at 1 mL/min flow rate. Protein concentration was monitored by absorption at 280 nm. The column was calibrated with Biorad standards.

Structural biology

*Sd*KefQCTD was prone to aggregation over extended concentration required for structural studies, but using stirred ultrafiltration cell 8003 with a 30 kDa membrane cut off (www.millipore.com) and a nitrogen stream at 4 °C, prevented this, allowing a concentration of around 20 mg mL^{-1}, in buffer containing 50 mM Tris-HCl, pH 7.8 and 300 mM KCl. Crystal trials were set up by hanging drop on freshly prepared protein samples that have not been previously frozen and involved mixing 1:1 and 2:1 volume of protein solution: precipitant equilibrated against a large volume of precipitant. Crystals grew to full size dimension of (0.2 mm × 0.05 mm × 0.05 mm) in two and a half months at 21 °C. The best crystals (judged by visual inspection) were obtained using 0.2 M sodium malonate pH 7.0 and 20% w/v PEG 3350, as precipitant. Prior to data collection, crystals were transferred into a solution containing 0.2 M sodium malonate pH 7.0 and 40% *w/v* PEG 3350. Data were collected at 100 K on a single crystal, which diffracted to a resolution of 2.92 Å on I24 at Diamond (Oxford, UK). Data were indexed, integrated and merged using MOSFLM / SCALA {Leslie, 1992, Joint CCP4 and ESF-EAMCB newsletter on

protein crystallography, No 26, 1-10} as implemented in CCP4 {CCP4, 1994, Acta Crystallographica Section D, 50, 760-763}. The resolution limits were determined by the data statistics and the Wilson plot. The CCP4 program POINTLESS was used to assign space groups for Kef as $P4_22_12$. The structure was solved using molecular replacement with the program PHASER using the *E. coli* KefC (PDB code 3EYW) as model containing residues 410 - 570 (omitting all water molecules and ligands with non-conserved residues set to alanine). AMP was modeled in both monomers and refinement proceeded by REFMAC5 and manual intervention COOT. Full crystallographic statistics are shown in (Table 1).

Table 1. Statistics of the *Sd*KefQCTD X-ray crystal structure.

Data collection	KefQCTD from *Shewanella denitrificans*
Beamline	Diamond_I24 (21/10/2012)
Wavelength (Å)	0.9686
Resolution (Å)	3.09 (47.6 - 3.09)
Cell constants (Å) α= β =γ=90	a=71.6 b=71.6 c=140.4
Unique reflections	6976 (399)
Mean I/σ	15.3 (1.9)
Completeness (%)	97.6 (85.8)
Multiplicity	6.2 (3.3)
R_{merge}	0.098 (0.651)
Spacegroup	$P4_12_12$
Wilson B-factor (Å^2)	77
Refinement	
R (%)	20.53 (29.8)
R_{free} (%)	26.04 (35.5)
PDB accession code	5NC8
No of atoms	2448
Bond length rmsd	0.01
Bond angle rmsd	1.558
Molprobity (S10) score centile	2.03 (99th centile)
Clashscore, all atoms, score centile	5.43 (100th centile)

Ligand identification

100 µL of 200 µM purified *Sd*KefQCTD in 50 mM Tris-HCl, pH 7.8, 300 mM KCl buffer and full length purified *Sd*Kef in 50 mM Tris-HCl, pH 7.8, 300 mM KCl, 0.05% DDM buffer, were denatured by being subjected to a constant temperature of 95 °C on a bench thermo block for 30 min. Total volume of sample containing the denatured protein was loaded on a vivaspin concentrator with a 3 kDa cut off and centrifuged at full speed on a bench top centrifuge until all liquid had gone through the membrane (30 mins). Filtrate was loaded on a Superdex Peptide 10/300 column (General Electrics Healthcare) with an optimum size separation from 100 to 7000 Da. Prior to sample loading, the column was equilibrated with buffer D. AMP, ATP and NADH of 100 µL volume and 2 mM concentration were individually loaded onto the same column on the same day, under identical conditions to calibrate its behavior. The UV active HPLC peak was subjected to MALDI mass spectrometry, as were the standards.

NMR

NMR experiments were recorded at a ^{1}H frequency of 600 or 700 MHz using a Prodigy BBO probe (600 MHz) or Bruker Avance III spectrometers equipped with a TCI inverse cryoprobe (700 MHz), respectively. The samples were prepared in 5 mm NMR tubes and experiments conducted at listed temperatures. CPMG experiments employed the PROJECT sequence ($90°_x - [\tau - 180°_y - \tau - 90°_y - \tau - 180°_y - \tau]_n$ – acq) as described by Aguilar *et al*. with a total filter time of 96 ms.[27] In all edited ^{1}H experiments water suppression was achieved by presaturation. The pulse tip angle calibration was done for all the samples using the Bruker *pulsecal* routine.

The protein was prepared to a final concentration of 330 μM in deuterated sodium phosphate buffer as follows:

Exchange of protein solution with deuterated sodium phosphate buffer: The purified protein solution was exchanged with deuterated sodium phosphate buffer (50 mM NaH_2PO_4/Na_2HPO_4, 150 mM NaCl, pH 7.4, which were prepared in deuterium oxide) by using PD MiniTrap G-25 columns (GH Healthcare). The exchange was done by the spin protocol, according to the manufacturer's instruction.

The reference spectra of each of the ligands also were run at a concentration of 330 μM in the deuterated sodium phosphate buffer: 50 mM NaH_2PO_4/Na_2HPO_4, 150 mM NaCl, pH 7.4, which were prepared in deuterium oxide.

Denaturation of protein for detection of AMP release: The protein was denatured on a heat block at 80 °C for 3 h. Subsequently, protein precipitate was centrifuged down and the resultant supernatant used directly for NMR studies. The ligands AMP, ADP and NADH were heat treated in the same manner as controls. The experiments were repeated without removing samples from NMR tubes to confirm AMP was not observed from contamination.

Native mass spectrometry

Nanoelectrospray (nESI) MS experiments were carried out on a QToF mass spectrometer (Waters Corp., Wilmslow, UK) with conditions optimised for the transmission of intact noncovalent protein complexes.[18] Sample was buffer-exchanged into 200 mM ammonium acetate and sprayed at a concentration of 7.5 μM

with respect to the dimer. Experiments were conducted in positive polarity with the following instrument settings: capillary voltage 1.4 kV, sample cone 40 V, extraction cone 25 V, backing pressure 3.5 mbar and collision cell pressure 1.75 MPa. The instrument was calibrated using CsI, and data analysis performed using MassLynx (Waters Corp., Wilmslow, UK) software. Protein masses were determined using the three most intense charge states.

Differential scanning fluorimetry

Assays were performed using a Stratagene Mx3005P qPCR (Expt filter set, ex. 492 nm, em. 568 nm). The initial temperature was set to 25 °C (held for 5 minutes), increasing in increments of 1 °C for 55 cycles (held for 1 minute 30 seconds per cycle). Stock solutions of the ligands under examination were prepared to a concentration of 100 mM in buffer containing 50 mM Na-phosphate, 150 mM NaCl, pH 7.4 (with the exception of (−)-adenosine, which was prepared in DMSO). The 100 mM stock solutions were then diluted to a concentration of 10 mM in buffer containing 50 mM Na-phosphate, 150 mM NaCl, pH 7.4. A protein master mix was prepared containing *Sd*KefQCTD (13.3 μM) and Sypro Orange (2.2×, Invitrogen) in buffer containing 50 mM Na-phosphate, 150 mM NaCl, pH 7.4. Ninety-six well plates (Axygen) were prepared using the protein master mix (22.5 μL per well; 12 μM final concentration of protein and 2× final concentration of dye) and the appropriate ligand (2.5 μL per well; 1 mM final concentration). The plate was centrifuged at 1000 rpm for 3 minutes before being run. Controls were performed with dye alone, ligand and dye, and the protein alone. The T_M (melting temperature) was identified by fitting to the Boltzmann equation (Prism 5).[19] The change in unfolding temperature (ΔT_M) was calculated as the shift in T_M relative to the T_M of

the protein + 2.5 μL buffer (DMSO in the case of (−)-adenosine), in the absence of any ligand. A Student's t-test was performed to ensure that the changes were statistically significant.

Expression and purification of Kef mutant construct R416E *Sd*KefQCTD

To express the *Sd*KefQCTD(R416E) the pTrcSdKefQCTDH$_6$-R416E plasmid was first transformed into the *E. coli* strain, MJF373.[12] The expression and purification conditions were the same as those employed in the alternative protocol for purifying the unmodified *Sd*KefQCTD counterpart, except medium LK [10 g/L Oxoid™ Tryptone (Thermo Fisher Scientific Inc.), 5 g/L Oxoid™ Yeast Extract (Thermo Fisher Scientific Inc.), 6.4 g/L KCl (Thermo Fisher Scientific Inc.), 2 g/L glucose (Sigma-Aldrich Co.)] and cell mass of 6.26 g were used instead for protein expression and protein purification, respectively.

Quantification of bound AMP content in proteins by HPLC (R416E)

Analytical HPLC was performed on a PerkinElmer Flexar system with a Binary LC Pump and UV/VIS LC Detector. A Dionex Acclaim® 120 column (C18, 5 μm, 120 Å, 4.6 × 150 mm) was used for analyzing the AMP content in unmodified *Sd*KefQCTD as well as its mutant counterpart, *Sd*KefQCTD(R416E). This HPLC analytic method had a constant flow rate of 1 mL/min and lasted for 20 min per run. It adopted a mobile phase with a mixture of Solvent A (99.9:0.1 : H_2O:Formic acid) and Solvent B (99.9:0.1 : MeCN:Formic acid). This program employed 100% Solvent A at the first minute, then increased the concentration of Solvent B from 0 to 100% over 10 min using a linear gradient and this 100% Solvent B concentration stayed for further 3 min to the 14th min. After that, the method decreased the concentration of Solvent B to 0%

at the 15th min, followed by 100% Solvent A running through the column for the last 5 min of the experiment. The whole HPLC program is summarized in Figure 1.

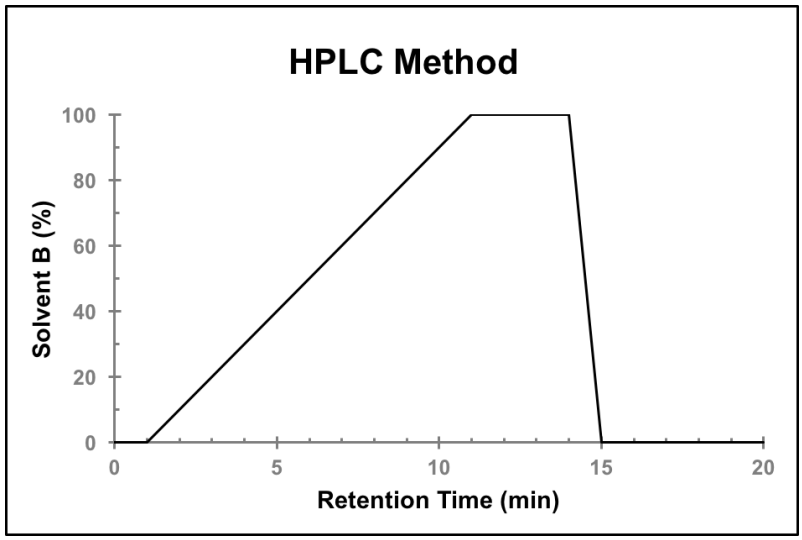


Figure 1. A summary of the HPLC program.

We adopted a similar experimental procedure developed by Chen and co-workers[28] to analyze the bound AMP molecule in proteins. To release the bound AMP, proteins of defined concentrations (10 μL) were first heated at 95 °C in a dry heating block for 5 min, and then subjected to centrifugation at 13 000 rpm for 10 min (MIKRO 20 Centrifuge, Hettich) to pellet down denatured protein precipitates. The resultant pellet was resuspended in 10 μL SEC Buffer (50 mM NaH_2PO_4/Na_2HPO_4, 150 mM NaCl, pH 7.4) and the mixture was then centrifuged at 13 000 rpm for 10 min (MIKRO 20 Centrifuge, Hettich). The resulting supernatants from the above two centrifugation steps were combined (10 + 10 = 20 μL) and used as an injection sample for subsequent HPLC analysis. For the spiking HPLC experiments, supernatants from the first centrifugation step (10 μL) were combined with AMP solutions (10 μL) of defined concentrations.

Computational methods

Molecular dynamics (MD) simulations were performed using the *Sd*KefQCTD model previously built by Healy and co-workers.[1] A total of 4 systems were evaluated: a) *Sd*KefQCTD with gluthathione (GSH) and 2 AMP molecules; b) *Sd*KefQCTD with ESG and 2 AMP molecules; c) *Sd*KefQCTD with GSH and d) *Sd*KefQCTD with ESG. Molecular mechanics parameters for ESG, GSH and AMP were taken from the General AMBER Force Field (GAFF)[2] with AM1-BCC atomic charges. Hydrogen atoms were removed from amino acid residues using the MolProbity Server[2] and added using *tLeap*.[3] Glutamate and aspartate residues were assigned as negatively charged and lysine and arginine as positively charged. Minimization and MD calculations were performed using the AMBER Force Field 12SB within AMBER version 12 with the GPU-accelerated version of PMEMD.[4] Crystallographic waters were not removed, while the protein was further solvated by a box of TIP-3P water molecules.[5] Simulations were carried out in octahedral boxes with an initial volume close to 160 nm^3 containing 3,060 water molecules, adding counter-cations (Na^+) to equilibrate the system. Energy minimization was performed in two steps. Firstly, we used steepest descent followed by conjugate gradients during which the initial position of the small molecule inhibitor and the protein structure obtained after homology modeling were restrained (PR). Secondly, the same minimization methodology was performed without PR. The minimized macromolecule:small molecule(s) complex was then subjected to 500 ps of equilibration and 40 ns of production MD simulation in the NPT ensemble using a Langevin thermostat to simulate a constant temperature at 310 K (τ T = 0.1 ps). Previous work performed by Zou[6], Hong[7], Shiao[8] and Gewert[9] and co-workers showed that nano scale MD could be enough to achieve reasonable protein models obtained by homology modeling.

Isotropic position scaling was used to maintain the pressure of 1 atm (τ p = 2 ps).[10] MD simulation was carried out using 1 fs of integration time and a nonbonding cut-off of 8 Å, with the Shake algorithm[11] turned on to constrain bonds involving hydrogen. A total of 2 000 snapshots were obtained at intervals of 20 ps in producing plots of the geometric variation during the simulation.

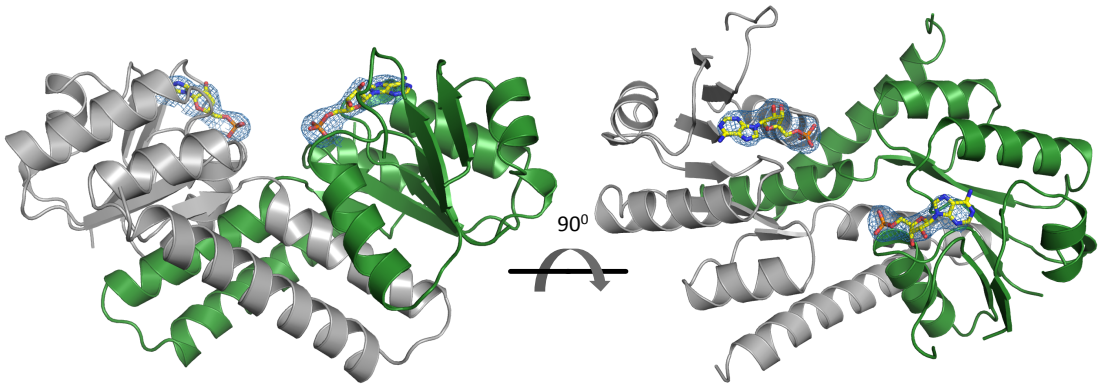
Results

X-Ray crystal structure of the C-terminal domain of Kef from *S. denitrificans* (*Sd*KefQCTD)

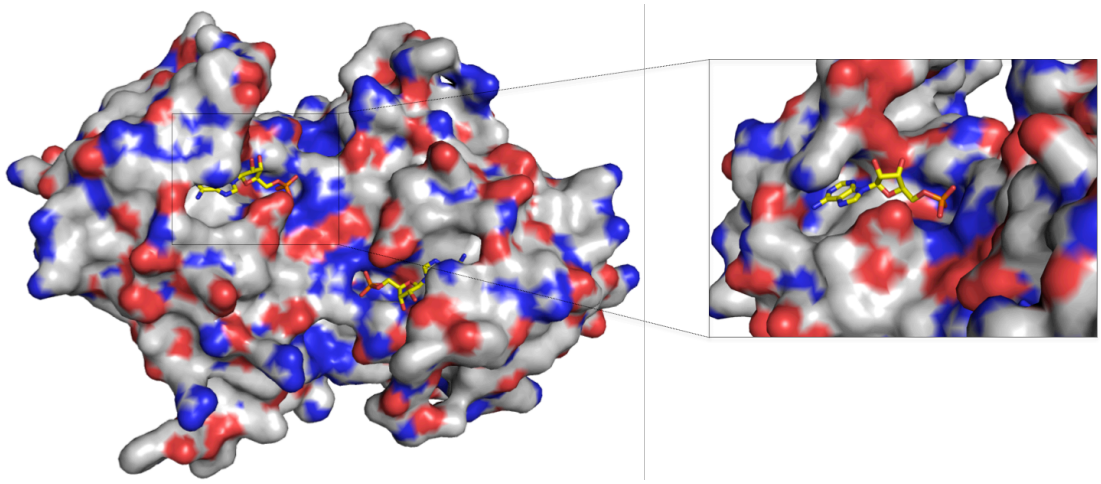
A construct of *S. denitrificans* KTN carboxy-terminal domains (*Sd*KefQCTD), was previously created to optimize stability and solubility.[12] The protein was purified to homogeneity from *E. coli* BL21(DE3) and crystallized. So as not to bias the occupancy of the Rossmann fold, no nucleotide was added before, during or after purification or crystallization of the sample used for all crystallization trials. Therefore, the presence of AMP in the crystal has arisen from nucleotide binding inside the *E. coli* cell and subsequent co-purification. Optimization of initial sparse matrix crystallization conditions yielded a single crystal for which diffraction data were collected to 3.09 Å resolution (Table 1). The structure was solved by molecular replacement using the apo-KefC structure (PDB ID: 3EYW) by removing NAD^+, as the searching model. The asymmetric unit contains two monomers that form the canonical KTN dimer. In line with the PISA[17] prediction, gel filtration (Figure S1), size exclusion chromatography – multi-angle laser scattering (SEC-MALS, Figure S2) analysis, and analytical ultracentrifugation (AUC, Figure S3) experiments indicated that the two Kef monomers form a stable dimer in solution, at the concentrations used. The core fold of the *Sd*KefQCTD domain is essentially identical to that previously described for the *E. coli* protein.[10,16] Each monomer has 6 β-strands (β1-6) arranged in a parallel sheet, which is sandwiched between three α-helices on one face and 1 α-helix on the other. Two C-terminal helices (α5, α6), resolved in the structure, form a helix-turn-helix type arrangement and reach across to the other monomer. The penultimate helix (α5) pairs with a single helix (α1) from the other

monomer and stacks against the β-sheet from the other subunit. The long C-terminal helix (α6) makes contacts with both monomers (Figure 2A).

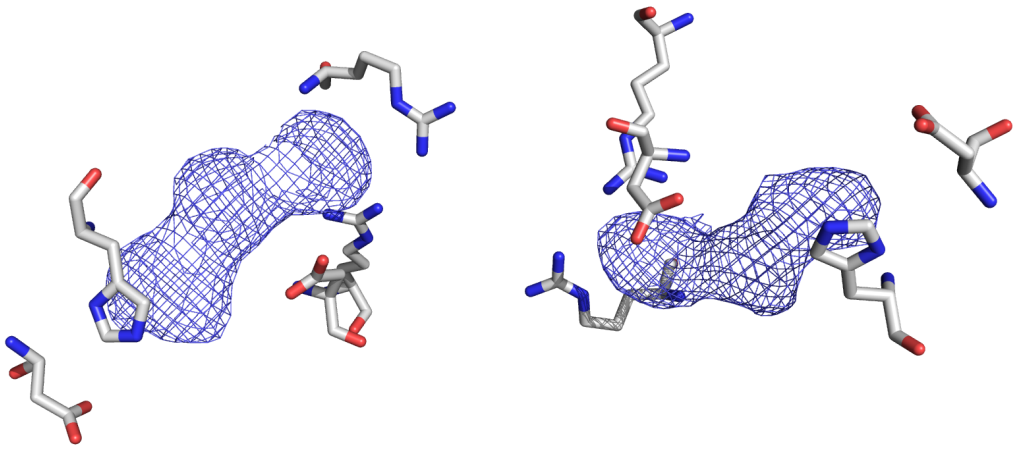
A



B



C



D

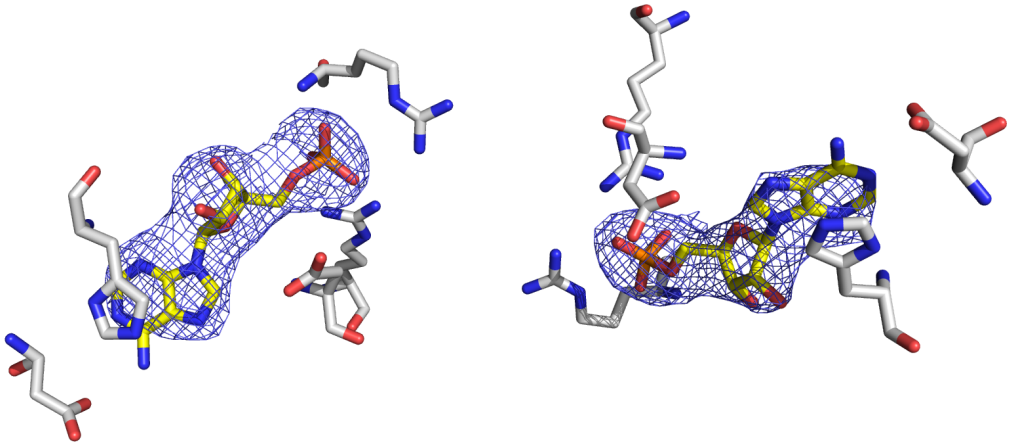


Figure 2. (A) The X-ray crystal structure shows *Sd*KefQCTD as dimeric, with each nucleotide binding pocket occupied by an AMP molecule (PDB code 5NC8). **(B)** View of the Kef protein as looking from the membrane, in surface view. AMP occupies both nucleotide binding pockets and is depicted in a stick representation (carbon = yellow). (**C**) Two views of the F_0-F_c electron density map contoured at 3.0σ. The phases were calculated from a model which had not included AMP. (**D**) The same F_0–F_c map now with the final position of the AMP molecule shown in sticks.

The presence of AMP and key interactions in the KTN domain.

Additional unbiased electron density was clearly visible in both nucleotide-binding pockets (Figure 2A), which was best fitted and refined with an AMP molecule. The nucleotide-binding sites are located within the Rossmann fold of each monomer and the residues that contact AMP come from within one monomer (Figure 2B). However, the phosphate groups are only 12 Å apart and form part of a network of hydrogen bonds that spans the dimer interface where GSH and GSX bind (Figure S4). The side chain of R416 forms salt contacts with the phosphate group of AMP within

the same monomer, (we arbitrarily define as monomer A) and with E524 from the other monomer (monomer B). E524 (B) in turn makes a further salt bridge with R503 (B). R503 (B) binds to the AMP phosphate group in monomer B. The net result can be thought of as an octagon with vertices of α phosphate (monomer A) - R416 (A) – E524 (B) – R503 (B) – phosphate (B) – R416 (B) – E524 (A) – R503 (A) (which links to the phosphate of monomer A) (Figure 3). R416 (A) also interacts with the backbone carbonyl oxygen of I505 (B) from the opposite chain. I505 (B) is located at the entrance to the previously identified GSH-binding pocket (Figure S4). The B-factors observed in this region of the X-ray crystal structure are below the mean for the whole protein indicating the region is well ordered consistent with it being rigid and stable in the crystal form.

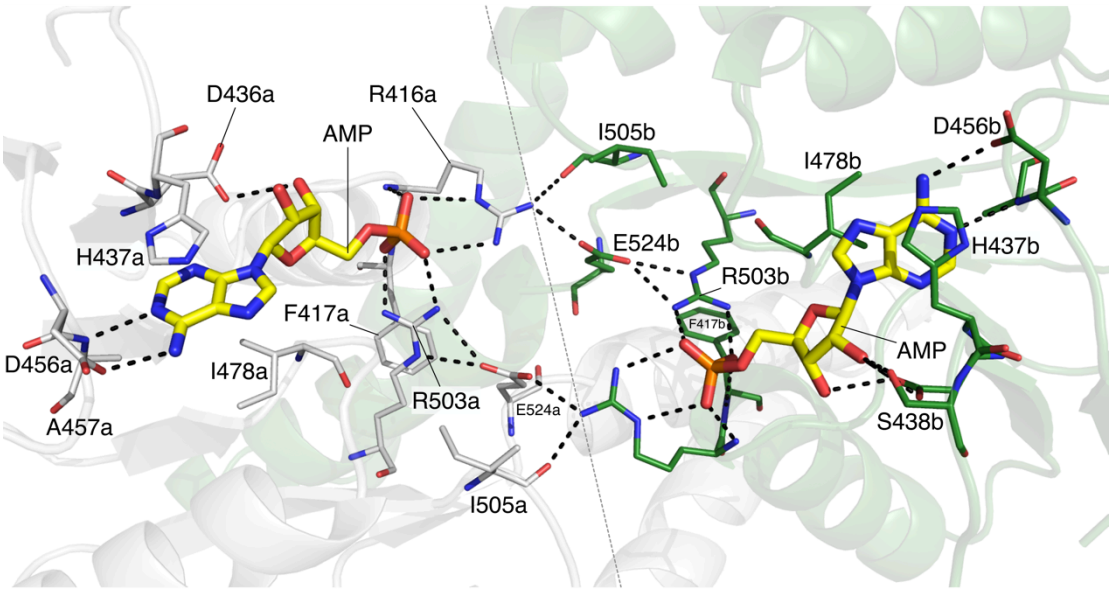


Figure 3. The two AMP binding pockets of *Sd*KefQCTD from *S. denitrificans* (PDB code 5NC8, AMP carbon = yellow). The key residues involved in AMP binding are highlighted as lines, with the residues from chain a shown with carbon = white, and the residues from chain b shown with carbon = green. The gray dotted line indicates the dimer interface.

The adenosine of the AMP ring sits in a pocket formed by I478 and H437 with which it makes π-stacking interactions. The aniline-like amine (N6) of the purine ring forms a hydrogen bond with D456, and the adjacent pyridine-like nitrogen (N1) atom forms a hydrogen bond with the backbone NH of A457 (Figure 4A). This pattern would appear to allow the protein to distinguish between AMP and other nucleotides. O2 and O3 of AMP are located close to D436, and O3 forms a hydrogen bond with this residue. The structure points to a central role for the phosphate group, which forms salt bridges to R416 and R503, and hydrogen bonds to the backbone NH of R416 and F417. Comparison with the X-ray crystal structure of the CTD of KefC from *E. coli*, reported by Roosild *et al.*, reveals that the key residues involved in AMP binding are conserved between these two proteins, and form similar interactions with AMP (Figure 4B).[16] An additional interaction between H434 of *E. coli* KefC is visible, but the density for this residue is not well resolved in the *Sd*KefQCTD X-ray crystal structure.

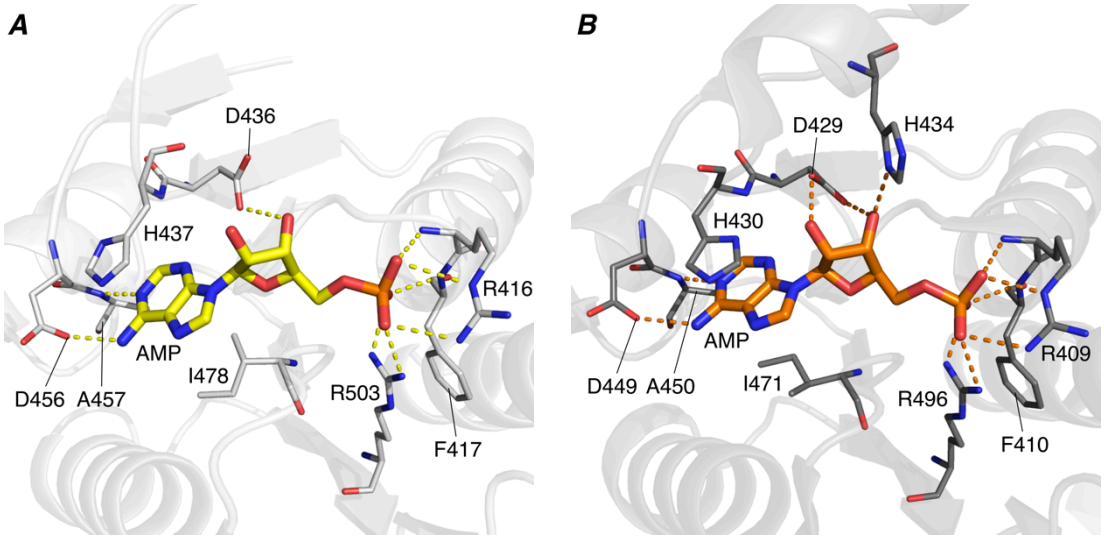


Figure 4. A comparison of the AMP binding pocket of *Sd*KefQCTD (**A**, PDB code 5NC8, AMP carbon = yellow) and the AMP binding pocket of the C-terminal domain of *Ec*KefQCTD from *E. coli* (**B**, PDB code 3L9W, AMP carbon = orange).[16] The key

residues involved in AMP binding are highlighted as lines and are conserved between the two proteins. The equivalent residue of H434 was not well resolved in the *Sd*KefQCTD X-ray crystal structure.

Identification of the bound nucleotide

Crystallography on its own is not a reliable tool to identify bound ligands. Consequently, this technique is unable to distinguish between AMP and NAD^+ with a disordered nicotinamide component. Therefore, we performed HPLC analysis to identify the bound nucleotide. Purified *Sd*KefQCTD protein from the batch used for crystallization trials was denatured and supernatant filtered through a 3 kDa cut-off concentrator. The filtrate was analyzed by HPLC, and a peak with a retention time equal to an authentic sample of AMP (distinct from ATP and NAD^+) was observed (Figure S5A). Further analysis by mass spectrometry shows that the peak has a mass of 348 Da, corresponding to protonated AMP (Figures S5A-C). Quantitation by UV extinction coefficient shows an AMP to protein ratio of at least 0.6 (Figure S5A). Purified full-length integral membrane protein *S. denitrificans* Kef (expressed in *E. coli*) treated in the same way and gave the same result (Figures S5B and C) (n = 2).

^{1}H CPMG NMR analysis

A relaxation-edited ^{1}H NMR experiment was performed on both the native and denatured *Sd*KefQCTD using the Carr-Purcell-Meiboom-Gill (CPMG) based filter. This leads to suppression of the background resonances deriving from the protein and attenuates those of any ligand bound to the macromolecule. The native protein spectrum showed no resonances corresponding to AMP (or NADH or ADP), confirming that any nucleotide present was tightly bound to the protein. However,

when *Sd*KefQCTD was denatured by heating for 3 h at 80 °C, a new set of sharp signals was observed in the spectrum, corresponding to the presence of a free AMP in solution (Figure 5). NADH and ADP were exposed to the same conditions used to denature the protein (heating for 3 h at 80 °C) to formally exclude the possibility that AMP was produced by breakdown of NADH or ADP Neither NADH or ADP led to AMP production (Figures S7 and S8), nor did the spectra of either compound match the molecule detected. AMP was stable to the same heat treatment (Figure S9).

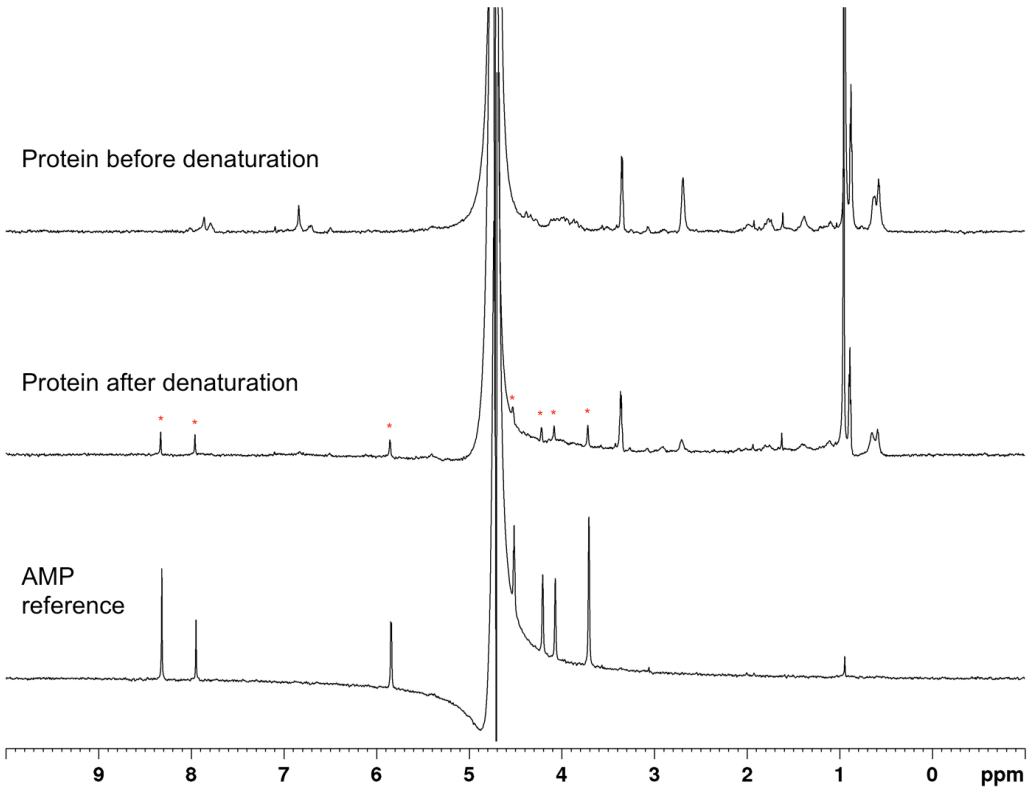


Figure 5. CPMG edited ^{1}H NMR spectra (700 MHz, 278 K) of the native (top) and denatured (middle) *Sd*KefQCTD protein with reference AMP spectrum (bottom). The red stars denote resonances corresponding to AMP that appear in the *Sd*KefQCTD protein spectrum after denaturation, consistent with the release of AMP from *Sd*KefQCTD following denaturation.

Determining the stoichiometry of AMP binding to the SdKefQCTD dimer

Nanoelectrospray MS measurements were then performed to accurately determine the stoichiometry of AMP binding under conditions that preserve noncovalent interactions in the gas phase.[18] The mass spectrum reveals a single charge state series centred around 14+ corresponding to the dimeric protein with two AMP molecules bound (Figure 6A). We did not detect *apo* protein or a dimer with one AMP molecule bound, indicating the protein is homogenously and fully 'loaded' with AMP. Activation, achieved by accelerating the protein into a collision cell,[18] results in peaks corresponding to the mass of a protein dimer with a single AMP molecule bound; and *apo* dimer (Figure 6 and Table S2). This is consistent with the sequential loss of neutral AMP.

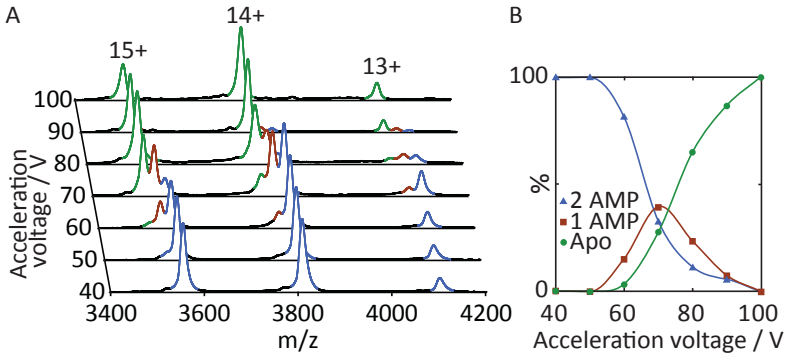


Figure 6. Stoichiometry determination of AMP binding to *Sd*KefQCTD through native mass spectrometry. (A) Mass spectra of *Sd*KefQCTD reveal the sequential removal two AMP with increasing collisional activation. (B) Proportion of bound *Sd*KefQCTD, as a function of collision voltage, averaged across all charge states.

AMP stabilizes SdKefQCTD

Given the regulatory roles played by ATP (and NADH) in TrkAH and KtrAB, we sought to investigate the influence of different nucleotides on the stability of

*Sd*KefQCTD in the presence and absence of the peptide ligands GSH and ESG. We used DSF,[19] in which the protein is heated in the presence of a fluorescent dye (SYPRO orange), and as the protein unfolds an increase in the fluorescence is observed. The point at which the gradient of fluorescence increase is greatest is the measured melting temperature of the protein (T_m). Repeating the experiment in the presence of a ligand that binds to, and hence stabilizes, a folded state of the protein results in an increase in T_m. The change in melting temperature between the free and ligand bound protein states (ΔT_m) roughly correlates to the ligand's affinity for the protein, however, there are some caveats. In particular, larger ΔT_m values are typically observed for more entropically driven (e.g. hydrophobic) binding. Therefore it can be difficult to directly compare ΔT_m values for compounds with very different physicochemical properties.[20]

We hypothesized that stabilization of the Kef T_m by AMP would occur despite the protein being purified with a high percentage of the nucleotide binding sites occupied by AMP. We assumed that as the temperature increases, and the protein unfolds, the off-rate of AMP would also rise, allowing the stabilizing effect of AMP binding to be observed. However, in the presence of the exogenous AMP the equilibrium would be shifted, partially compensating for the increased off rate, resulting in protein stabilization.

Conducting the DSF experiment in the presence of AMP resulted in ΔT_m = +15 °C (Figure 7A), which is consistent with tight binding of AMP to *Sd*KefQCTD. ADP showed ΔT_m = +7 °C, ATP and adenosine had no effect on T_m (Figure 7A) suggesting that much of AMP's affinity for *Sd*KefQCTD derives from the binding of the

phosphate group (Figure 7A). This observation is consistent with the phosphate group-protein interactions that were observed in the *Sd*KefQCTD X-ray crystal structure (Figure 4A), and our computational studies (see below). NADH showed ΔT_m = +3 °C, but NADP and NAD^+ had little effect on T_m.

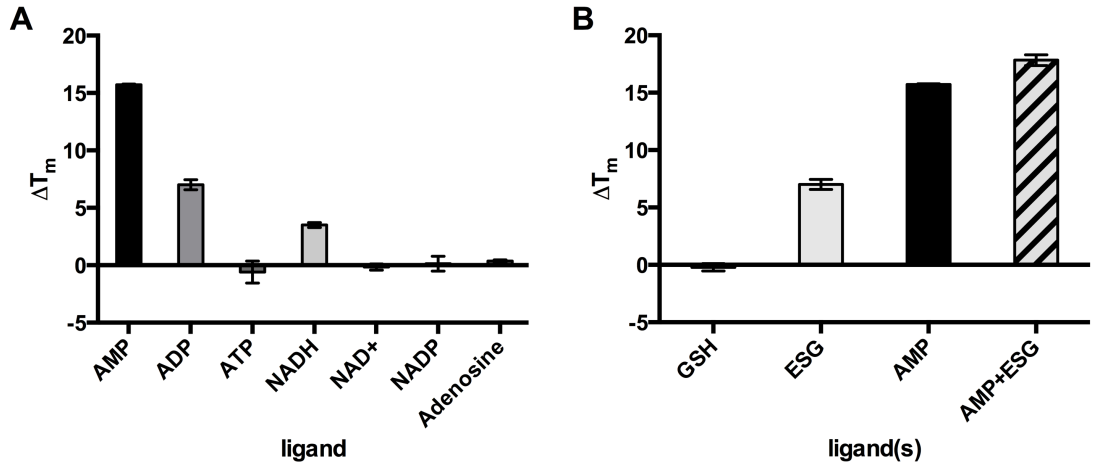


Figure 7. (A) DSF experiments to determine the effect of nucleotides on the stabilization of *Sd*KefQCTD. It is shown that AMP is most effective at stabilizing *Sd*KefQCTD with ΔT_m = +15 °C. **(B)** DSF experiments to determine the effect of both AMP and ESG on the stabilization of *Sd*KefQCTD. Little stabilization is provided by GSH, whereas ESG shows ΔT_m = +7 °C. The stabilization in the presence of both AMP and ESG is +18 °C, consistent separate binding sites for these two ligands.

To compare the binding of the peptide and nucleotide ligands, and to determine whether their binding was competitive, DSF experiments in the presence of two ligands were conducted (Figure 7B and Figure S10). It has previously been shown that GSH has little effect on T_m, which is consistent with its low affinity for *Sd*KefQCTD.[12] ESG, which has a higher affinity for *Sd*KefQCTD, has ΔT_m = +7 °C.

ESG binding was additive with the other ligands (Figure S10), for example the addition of AMP and ESG gives ΔT_m = +18 °C. No release of AMP was observed upon binding of the high affinity ESG analogue, *t*-butyl-*S*-glutathione (Figure S11) using CPMG NMR. These data are consistent with structural data showing distinct nucleotide and peptide binding sites.

Mutagenesis studies to probe the role of AMP when bound to the Rossmann fold of SdKef and SdKefQCTD

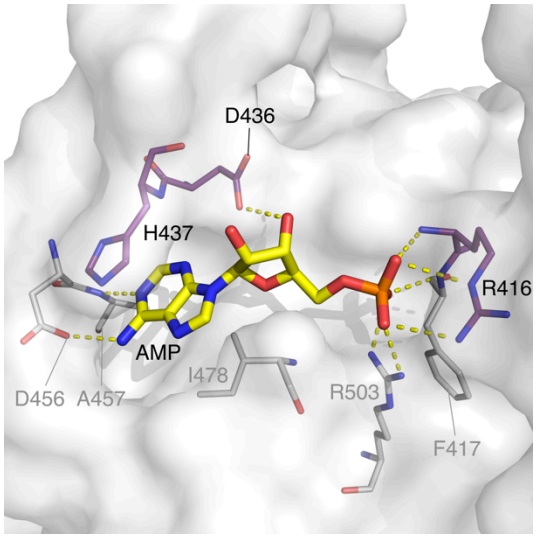


Figure 8. The AMP-binding residues that have been investigated by mutagenesis studies are highlighted with carbon = purple. AMP is shown as a stick representation (carbon = yellow). Image generated using PyMOL and the X-ray crystal structure of *Sd*KefQCTD (PDB code 5NC8).

Mutation of three residues R416, D436 and H437 was carried out in both *Sd*KefQCTD and the full-length protein to test our hypothesis that AMP binding was important for protein stability, and thus function. We selected H437 because it forms the top of the AMP binding pocket and stacks with the purine ring (Figure 8). D436 was identified to form hydrogen bonds with the two hydroxyl groups on the ribose ring of AMP (Figure 8) and R416 because it directly contacts the AMP phosphate

group (Figure 8). Our computational studies (see below) predicted this to be a key interaction and in general interactions between an arginine residue and a phosphate group are typically strong,[21] and would be expected to contribute significantly to the binding interaction of Kef with AMP. The *Sd*KefQCTD mutants were expressed in *E. coli* strain MJF335, which lacks the chromosomally-encoded KefC and KefB proteins, and does not have endogenous GSH biosynthesis due to disruption of the *gshA* locus. This strain permits the presence of GSH to be controlled by medium supplementation. Whole cell-based K^+ efflux assays were performed with the full-length mutant proteins expressed in MJF335 ($KefB^-$, $KefC^-$, $GshA^-$) supplemented during growth with 1 mM GSH (Table S3).

H437

H437 forms the top of the AMP binding pocket, and is observed to be in close proximity (~3.5 Å) to the purine ring of AMP (Figure 8). It is, therefore, plausible that stacking interactions between the two aromatic rings stabilize this interaction. The H437A and M437N mutations remove the stacking interactions and give an indication of how important this residue is to *Sd*Kef affinity for AMP. These mutations, in either *Sd*KefCTD or *Sd*Kef, had only limited effects on protein expression and function. Both H437A and H437N showed expression levels similar to the wild type protein in *Sd*Kef (Figure 9A) and *Sd*KefQCTD (Figure 9B). Consistent with the expression data, full-length proteins carrying either H437A or H437N were active in a K^+ efflux assay although both proteins displayed altered kinetics; H437N exhibited close to wild type activity in its initial K^+ loss rate, but was reduced in final extent, whereas H437A displayed ~50% of wild-type activity in the initial rate of K^+ lost but was similar to wild-type in overall loss. Taken together,

these results indicate that the purine ring of AMP contributes only modestly to its affinity for *Sd*Kef.

D436

D436 forms hydrogen bonds with the two hydroxyl groups on the ribose ring of AMP (Figure 8). D436E would be expected to form similar interactions with AMP. Consistent with this hypothesis, the D436E mutant in both *Sd*Kef and *Sd*KefQCTD exhibited approximately wild-type expression levels (Table 2, and Figure 9). The NEM-elicited K^+ efflux from the *Sd*Kef D436E mutant was not substantially altered from that of the wild-type protein. The uncharged D436N would be expected to form reduced hydrogen bonding interactions with AMP, while D436A would be unable to form hydrogen bonds with AMP. Both D436A and D436N displayed severely reduced levels of protein expression in both *Sd*Kef and SdKefQCTD (Table 2); for *Sd*KefQCTD D436N no protein was detected (Figure 9B). A significant reduction in the K^+ activity of the *Sd*Kef D436A and D436N proteins was found, in line with the reduced expression of these two mutant proteins (Figure 9A). Low expression could indicate lower protein stability resulting from impaired AMP binding.

Table 2. Protein yields and percentage AMP retention of the D436E and D436A mutant.

S. denitrificans Kef mutant	Protein yield (mg / L)	AMP retention (%)
WT-FL	3 ± 1	107.3 ± 21.5 (n=2)
WT - QCTD	9 ± 1	101.5 ± 11.6 (n=2)

D436E - QCTD	8 ± 2	99.7 ± 23.2 (n=2)
D436A -QCTD	2 ± 1	14.7 ± n/a (n=1)

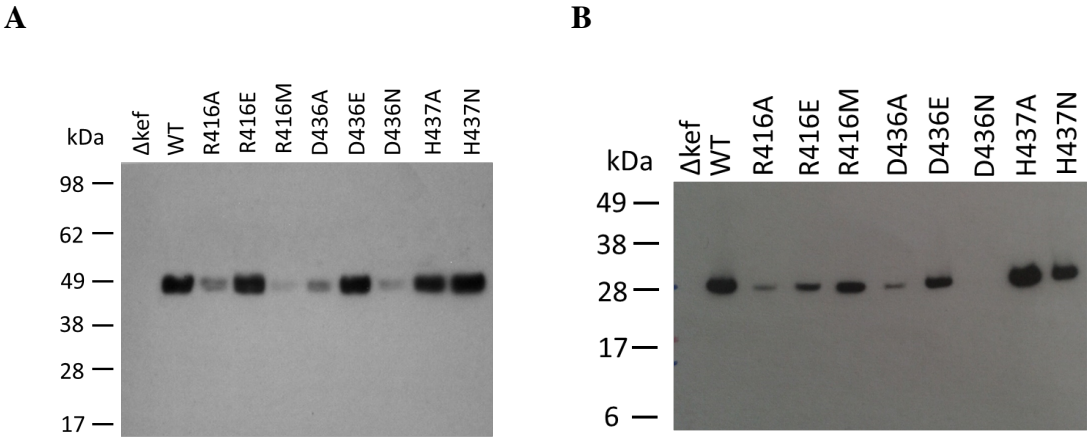


Figure 9. Western blot of the full length *Sd*Kef (A) and the truncated *Sd*KefQCTD (B) mutants. Wild-type (WT) and the mutants in each plasmid were expressed in MJF373 cells and overproduced by induction with 0.3 mM IPTG. Membrane fractions (A) or soluble fractions (B) were isolated, and 15 μg of protein per well separated on SDS-PAGE, transferred to nitrocellulose membrane and an antibody against the C-terminal His_6-tag was used for detection of the proteins. MJF373 alone was used as control (Δkef).

R416

The R416A, R416M, and R416E mutations would be expected to abolish phosphate binding. R416A shows low expression levels for both the full length (*Sd*Kef, Figure 9A) and truncated (*Sd*KefQCTD, Figure 9B) proteins. R416M also showed low expression levels for *Sd*Kef, and reduced expression for *Sd*KefQCTD. These results

are consistent with the mutations reducing AMP binding to the KTN domain. However, while R416E showed very low expression for *Sd*KefQCTD (Figure 9B and Figure S12), the mutated full length *Sd*Kef protein showed wild type expression levels (Figure 9A). This observation implies that this mutant is more stable in the full-length protein. While R416A retained some activity in the K^+ assay, R416M had substantially reduced activity, and an R416E mutant was completely inactive. Consistent with the proposed hypothesis. While AMP might still be accommodated in R416A, steric occlusion of the binding site would be expected from R416M. R416E would likely be the most disruptive to AMP binding, as the negatively charged ɣ-carboxylate of glutamate would repel the phosphate group of AMP. However, it is possible that protein folding occurring better in the full-length protein than the truncated construct. Once folded, it is possible that E416 forms a salt bridge with R503, stabilizing the protein while not binding AMP, explaining the stable but inactive protein.

HPLC analysis of AMP retention by the mutated proteins

To determine whether the variation in protein stability and function correlated to the degree of AMP binding, we used HPLC analysis to assess the amount of AMP that was co-purified with the mutated proteins (Figures 10 and 11). AMP retention by purified *Sd*KefQCTD D436E and D436A proteins was analyzed as above, and it was observed that the D436E protein retained almost the same levels of AMP as the wild type, whereas D436A protein only had ~20% of the normal level (Figure 10). When AMP retention by R416E was analyzed by HPLC, very little AMP could be detected (Figure 11).

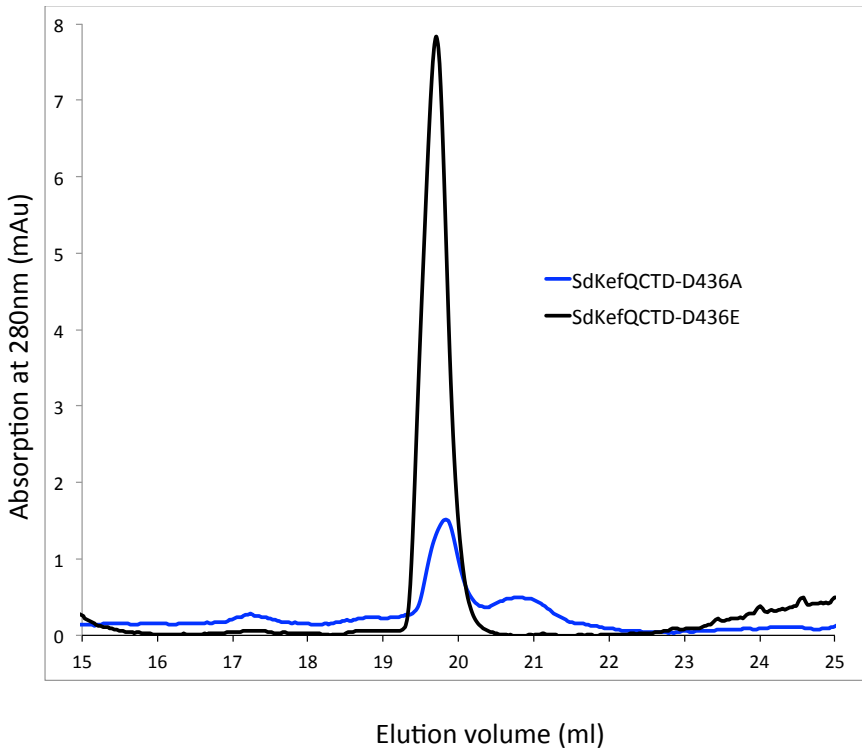


Figure 10. AMP percentage of retention was calculated by comparing the absorption at 280 nm (A280) of the denatured proteins and AMP standards ran the same day, of equal concentration, sample volume and identical buffer as well. A representative example is shown. Mutants D436A and D436E. 100 µL of 200 µM were treated as described in Materials and Methods section, and all resulting single peaks appearing in the gel filtration profiles were subjected to further mass spectrometry analysis for nucleotide identification (Figures S5D-F).

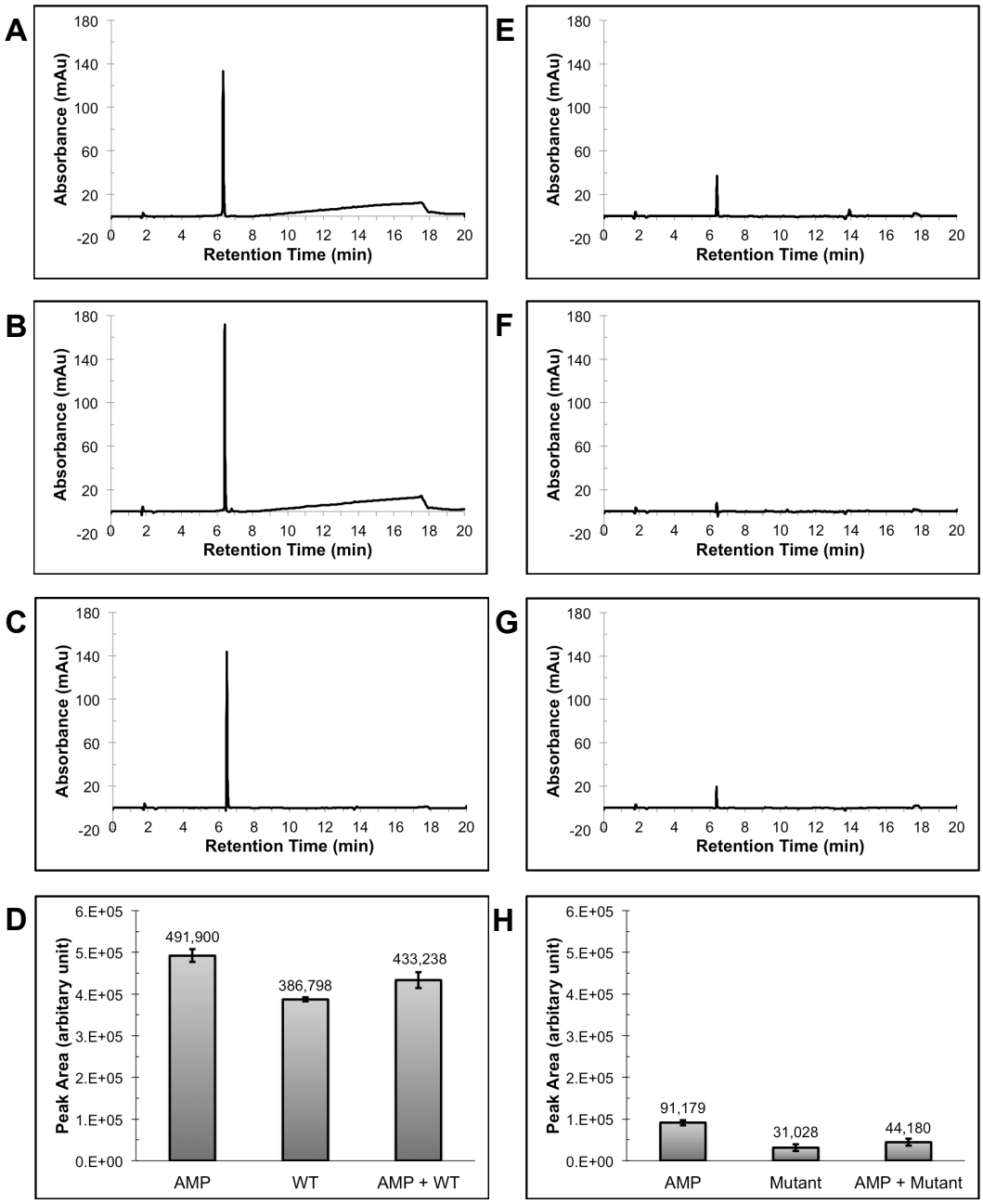


Figure 11. HPLC analysis of AMP released from heat-denatured *Sd*KefQCTD and *Sd*KefQCTD(R416E). All experiments were conducted under the same HPLC conditions, see Materials and Methods section for details. (A) HPLC profile of 50 μM pure adenosine monophosphate (AMP). (B) HPLC profile of 50 μM of denatured wild-type protein (WT) – *Sd*KefQCTD. (C) Spiking experiment containing equal concentration (25 μM each) of pure AMP and denatured WT (i.e. AMP + WT). (D) Quantification of HPLC peak areas from A to C. Data shown are mean ± SD values

from three different experiments (n = 3). (E) HPLC profile of 8.6 μM of pure AMP. (F) HPLC profile of 8.6 μM of denatured mutant protein (Mutant) – *Sd*KefQCTD(R416E). (G) Spiking experiment containing equal concentration (4.3 μM each) of pure AMP and denatured mutant (i.e. AMP + Mutant). (H) Quantification of HPLC peak areas from E to G. Data shown are mean ± SD values from three different experiments (n = 3).

Computation analysis on the effect of ligand binding on SdKefQCTD stability

To investigate computationally the effect of AMP and related ligands on protein stability molecular dynamics (MD) simulations were performed. To include regions of the protein that are disordered in the X-ray crystal structure the *Sd*KefQCTD homology model previously reported by Healy *et al.*,[12] was employed. Four systems were evaluated: a) *Sd*KefQCTD bound to one GSH molecule, and two AMP molecules (one bound to each of the Rossmann folds in the dimer); b) *Sd*KefQCTD bound to one ESG molecule and two AMP molecules bound; c) *Sd*KefQCTD bound to one GSH molecule, and d) *Sd*KefQCTD bound to one ESG bound molecule.

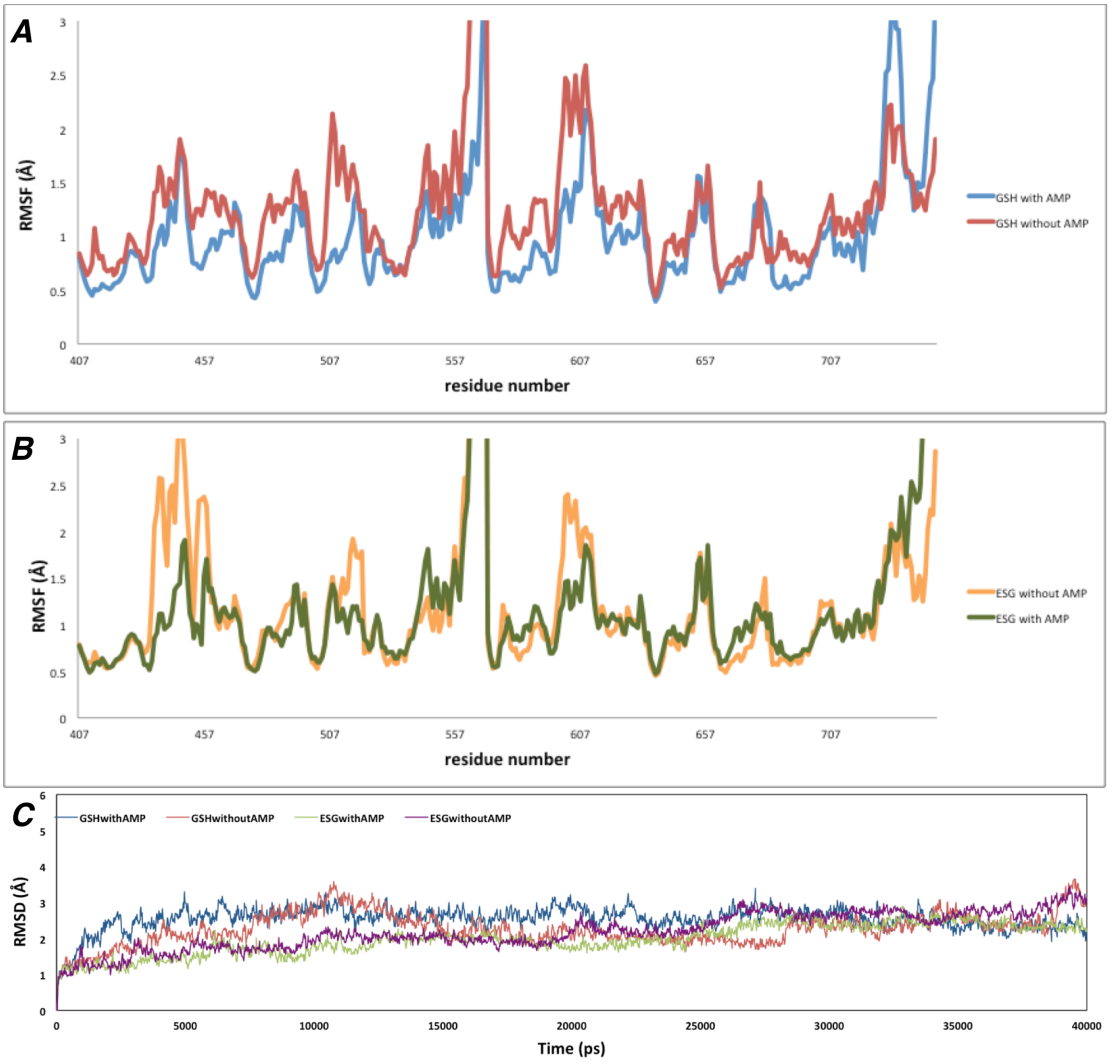


Figure 12. A. Per-residue Root Mean Square Fluctuations (RMSF) of the Cα atoms of a homology model of *Sd*KefQCTD either in the presence of GSH and AMP (blue) or only GSH (red). **B.** Per-residue Root Mean Square Fluctuations (RMSF) of the Cα atoms of a homology model of *Sd*KefQCTD either in the presence of ESG and AMP (green) or only GSH (yellow). **C.** Temporal RMSD (relative to starting structure) for Cα atoms during 40 ns of MD simulation time for *Sd*KefQCTD in complex with GSH (red), GSH and AMP (blue), ESG (purple), or ESG and AMP (green).

Consistent with the experimental data described above, MD simulations of the systems with AMP bound to the Rossmann fold were more stable compared to the

same model without AMP (as quantified by smaller per-residue fluctuations, shown in Figure 12A and B). The *Sd*KefQCTD-AMP-GSH complex did not show large conformational changes during the entire simulation. This model had a stabilized Root Mean Square Deviation (RMSD), relative to the initial point of the MD, of approximately 2.2 Å after the first 5 ns of MD (Figure 12C); the *Sd*KefQCTD-AMP-ESG complex stabilized after 25 ns of MD at approximately 2.3 Å RMSD (Figure 12C). This behavior is to be predicted, since the initial homology model was based on the *Ec*KefQCTD X-ray crystal structure with AMP and GSH both bound (PDB code 3L9W), and so some conformational changes are to be expected when substituting GSH by ESG. In contrast, 40 ns simulation time was insufficient to observe stability for *Sd*KefQCTD-GSH and *Sd*KefQCTD-ESG complexes without AMP bound, which displayed larger geometric fluctuations, as judged by RMSD values greater than 3 Å. These results are qualitatively consistent with AMP being an important factor in stabilizing the *Sd*KefQCTD structure.

The AMP - R416 interactions (electrostatic and hydrogen-bonding) were maintained for 40 ns of MD, with one or two hydrogen bonds observed throughout the simulation (see Figure S13 for further information). Hydrogen bonding interactions at the interface of the protein dimer (R416, E524, R503 and I505) may help to provide structural rigidity. A 40 ns simulation of *Sd*KefQCTD bound only to ESG (Figure 12B) shows high flexibility of the region containing Pro439-Leu445 residues, as a result of unfavorable steric interactions with the succinimide ring of ESG. This effect is not observed in the *Sd*KefQCTD-AMP-ESG complex, due to hydrogen bonds with Asp436 (Figure S14, Figure 12B) and aromatic face-to-face (so-called π-stacking) interactions with His437 (Figure S15). These interactions are maintained during 40 ns

of MD simulation and increase the structural stability of the amino acids in the P439-L445 region.

Discussion

A combination of crystallography, DSF, HPLC analysis, mass spectrometry, and NMR studies provide evidence that the *Sd*KefQCTD domain binds AMP; HPLC analysis suggests the full-length protein also binds AMP. The stoichiometry of the complex, determined using mass spectrometry and HPLC, is 1:1 (one AMP molecule per one *Sd*KefQCTD monomer) which is consistent with that observed by crystallography. Although we were unable to measure a direct binding constant, several pieces of evidence point to this being a high affinity interaction, which we speculate is approximately in the nM range. Firstly, it was only possible to remove AMP by denaturing the protein; secondly, the high thermal stabilization of the *Sd*KefQCTD by additional AMP; thirdly, we were not able to exchange AMP with related nucleotides, and finally the AMP must have been bound in the cytoplasm of *E. coli*, where the concentration of AMP is much lower than other adenosine containing nucleotides – the estimated cellular concentrations of relevant nucleotides in *E. coli* are ATP (9 mM), ADP (0.5 mM), AMP (0.28 mM), NAD^+ (2.6 mM) and NADH (0.08 mM).[22] AMP pools are integrated with those of ADP and ATP *via* the activity of adenylate kinase. The *E. coli* adenylate kinase has a high affinity for AMP (~30 μM) and a high V_{max} (1247 μmol/min/mg protein) ensuring a rapid equilibrium.[23] Thus the growing cell maintains ATP:ADP:AMP in ratios 100:10:2, respectively (stationary phase cells generally have higher AMP concentrations). We did not detect any protein with NAD^+, NADH or ATP bound, which implies that in the cell the domain has binding constant for AMP at least 200-fold tighter than NAD^+, 6-fold

tighter than NADH, 700-fold tighter than ATP, and 35-fold tighter than ADP (these values are predicated upon an estimated upper limit of 5 % for missed detection of such complexes). The observation that the protein appears (within the error of our measurement) to fully occupied with AMP despite extensive purification suggests a very tight binding. It should be noted, however, that the observed K_d values of the KtrAB system for nucleotides falls well below the concentration of cellular pools,[24] and it has been speculated that the change in these pools is accompanied by nucleotide exchanges that are intrinsic to the allosteric transition.

The Kef system is normally in an inactive, GSH-bound, state. When GSH is replaced by an electrophilic adduct of GSH, this triggers channel opening. The formation of such an adduct occurs when the bacterial cell is intoxicated, and the activation of the K^+ efflux system is part of the survival response. In such a scheme, what is the role of AMP? Dimeric Kef protein possesses two GSH-binding sites that are located at the interface between the two protein monomers. These sites are in close proximity to the two nucleotide-binding sites, and in fact complex series of hydrogen bond and salt bridge interactions link AMP, the dimer, and the residues that form the GSH binding site (Figure 3). We hypothesized that the GSH (and GSH-adducts) binding site requires dimer formation, which in turn is dependent on the presence of AMP. Thus, we propose AMP plays a vital role in creating a protein architecture that can respond to the chemical change induced by binding GSH adducts. Specifically, AMP makes a number of contacts with helix 418-437 which extends across the dimer interface and is thought to be involved in the conformational change associated with Kef channel gating.

In support of this model we note that addition of exogenous AMP results in significant stabilization of the protein, which is much larger than typically expected for stabilization of a monomeric protein alone. As *Sd*KefQCTD always co-purifies with the bound nucleotide, we were unable to generate a true *apo* structure to measure the stabilizing effect of ESG on its own, but we noted its stabilization effect was additive to AMP, and NMR analysis confirmed that ESG binding did not displace AMP. Computational modeling was used to generate a structure lacking bound AMP and classical MD simulations showed lower structural stabilities, even when GSH was added. By contrast models with AMP remained stable whether or not GSH was present. Experimental confirmation of the importance of AMP binding for structural integrity was obtained by a panel of mutants selected to disrupt AMP binding. In each case the mutants displayed a decrease in AMP content of purified protein, reduced stability and loss of activity.

Previous crystallographic studies of KTN domains from channels and transporters have identified a variety of bound nucleotides. Usually, these have been ligands added during crystallization rather than being carried through during purification, and in many cases the *apo*-protein is readily isolated. For the octameric assemblies the crystallographic evidence for the role of these specific nucleotides have been supported by biochemical evidence. In contrast, the solved structures of the dimeric complexes, the observed density is consistent with AMP, but crystallography on its own is not a definitive tool for identification of ligands. For the *E. coli* KefC KTN domain structures, density consistent with AMP was observed in two crystal forms (the *apo* protein and the GSH-bound form) whereas a sulfate ion was found in the AMP-binding site in the ESG-bound form. These structures have provided valuable

insight into the nature of the conformational change induced by the electrophilic modification of GSH. A KTN domain structure with AMP has been deposited by a structural genomics consortium (3C85), but no comment is available on what if any ligands were added and the basis of ligand identification. Thus, it seems plausible that the dimeric KTN modules have AMP as a structural moiety, while the octameric proteins have the nucleotides, most probably ATP and NADH, as allosteric modulators of channel activity.

A further difference between the dimeric and octameric assemblies is the recognition of additional ligand, such as GSH, that regulate the activity of the dimeric systems, unlike in the octamers where dissociation of the nucleotide is used to achieve activation. In contrast to the other KTN domains (i.e. Trk, Ktr, and potentially the channels) where the nucleotide may play a major role in gating, AMP in Kef we propose is required to form the structure that binds GSH and its adducts. It follows, therefore, that the affinity for the gating ligands (whether nucleotide or GSH) should be within the dynamic range of actively-metabolizing cells. While a stabilizing ligand should bind sufficiently tightly to ensure that the protein integrity under all metabolic states.

ASSOCIATED CONTENT

Supporting Information

E. coli strains and plasmids used in the study, evidence of *Sd*KefQCTD dimer formation in solution, a comparison of the GSH/GSX-binding site of *Ec*QCTD and the presumed GSH/GSX binding site of *Sd*KefQCTD, additional evidence of AMP binding to *Sd*Kef and *Sd*KefQCTD, and additional computational figures.

Acknowledgements

We would like to thank Dr. Magnus Alphey for help with crystal mounting, and Dr. Sally Shirran for mass spectrometry data analysis and technical assistance.

Funding Information

The project is supported by the Wellcome Trust (WT092552MA, WT100209MA) to IRB, JHN, SM, and SJC, a Biotechnology and Biological Sciences Research Council grant (BB/H017917/1), and a European Union Marie Curie ITN Award (NICHE; 289384) that supported SE. CP would also like to acknowledge additional support by a Tenovus Scotland grant award (T15/41). WAC is supported by a Science Without Borders scholarship. The authors would like to acknowledge that the work presented here made use of the Emerald High Performance Computing facility made available by the Centre for Innovation, formed by the universities of Oxford, Southampton, Bristol, and University College London in partnership with the STFC Rutherford-Appleton Laboratory. RSP acknowledges the use of the EPSRC UK National Service for Computational Chemistry Software (NSCCS) at Imperial College London in carrying out this work (CHEM773). RSP is grateful to NVIDIA for the generous donation of Tesla GPUs as part of the academic partnership scheme. SJC and AC

thank the European Commission for the award of a Marie Curie Fellowship to AC (660156, FLUOROKEF). SJC and SCG thank the EPSRC for studentship support for SCG. SJC thanks St. Hugh's College, Oxford, for research support.

References

(1) Jiang, Y., Pico, A., Cadene, M., Chait, B. T., and MacKinnon, R. (2001) Structure of the RCK Domain from the *E. coli* K+ Channel and Demonstration of Its Presence in the Human BK Channel. *Neuron 29*, 593–601.

(2) Follmann, M., Becker, M., Ochrombel, I., Ott, V., Krämer, R., and Marin, K. (2009) Potassium transport in corynebacterium glutamicum is facilitated by the putative channel protein CglK, which is essential for pH homeostasis and growth at acidic pH. *J. Bacteriol. 191*, 2944–2952.

(3) Adams, M. J., Ford, G. C., Koekoek, R., Lentz, P. J., McPherson, A., Rossmann, M. G., Smiley, I. E., Schevitz, R. W., and Wonacott, A. J. (1970) Structure of Lactate Dehydrogenase at 2.8 Å Resolution. *Nature 227*, 1098–1103.

(4) Loi, V. V., Rossius, M., and Antelmann, H. (2015) Redox regulation by reversible protein S-thiolation in bacteria. *Front. Microbiol. 6*, 187.

(5) Newton, G. L., Rawat, M., La Clair, J. J., Jothivasan, V. K., Budiarto, T., Hamilton, C. J., Claiborne, A., Helmann, J. D., and Fahey, R. C. (2009) Bacillithiol is an antioxidant thiol produced in Bacilli. *Nat. Chem. Biol. 5*, 625–627.

(6) Gaballa, A., Newton, G. L., Antelmann, H., Parsonage, D., Upton, H., Rawat, M., Claiborne, A., Fahey, R. C., and Helmann, J. D. (2010) Biosynthesis and functions of bacillithiol, a major low-molecular-weight thiol in Bacilli. *Proc. Natl. Acad. Sci. USA 107*, 6482–6486.

(7) Helmann, J. D. (2011) Bacillithiol, a new player in bacterial redox homeostasis. *Antioxid. Redox Signal. 15*, 123–133.

(8) Ferguson, G. P., Munro, A. W., Douglas, R. M., McLaggan, D., and Booth, I.

R. (1993) Activation of potassium channels during metabolite detoxification in Escherichia coli. *Mol. Microbiol. 9*, 1297–1303.

(9) Ferguson, G. P., McLaggan, D., and Booth, I. R. (1995) Potassium channel activation by glutathione-S-conjugates in *Escherichia coli*: protection against methylglyoxal is mediated by cytoplasmic acidification. *Mol. Microbiol. 17*, 1025–1033.

(10) Roosild, T. P., Castronovo, S., Miller, S., Li, C., Rasmussen, T., Bartlett, W., Gunasekera, B., Choe, S., and Booth, I. R. (2009) KTN (RCK) Domains Regulate K(+) Channels and Transporters by Controlling the Dimer-Hinge Conformation. *Structure 17*, 893–903.

(11) Lyngberg, L., Healy, J., Bartlett, W., Miller, S., Conway, S. J., Booth, I. R., and Rasmussen, T. (2011) KefF, the Regulatory Subunit of the Potassium Efflux System KefC, Shows Quinone Oxidoreductase Activity. *J. Bacteriol. 193*, 4925–4932.

(12) Healy, J., Ekkerman, S., Pliotas, C., Richard, M., Bartlett, W., Grayer, S. C., Morris, G. M., Miller, S., Booth, I. R., Conway, S. J., and Rasmussen, T. (2014) Understanding the structural requirements for activators of the Kef bacterial potassium efflux system. *Biochemistry 53*, 1982–1992.

(13) Roosild, T. P., Miller, S., Booth, I. R., and Choe, S. (2002) A mechanism of regulating transmembrane potassium flux through a ligand-mediated conformational switch. *Cell 109*, 781–791.

(14) Kong, C., Zeng, W., Ye, S., Chen, L., Sauer, D. B., Lam, Y., Derebe, M. G., Jiang, Y., and Aldrich, R. (2012) Distinct gating mechanisms revealed by the structures of a multi-ligand gated K+ channel. *Elife 1*, e00184.

(15) Li, Y., Berke, I., Chen, L., and Jiang, Y. (2007) Gating and Inward Rectifying

Properties of the MthK K+ Channel with and without the Gating Ring. *J. Gen. Physiol. 129*, 109–120.

(16) Roosild, T. P., Castronovo, S., Healy, J., Miller, S., Pliotas, C., Rasmussen, T., Bartlett, W., Conway, S. J., and Booth, I. R. (2010) Mechanism of Ligand-Gated Potassium Efflux in Bacterial Pathogens. *Proc. Natl. Acad. Sci. USA 107*, 19784–19789.

(17) Krissinel, E., and Henrick, K. (2007) Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol. 372*, 774–797.

(18) Kondrat, F. D. L., Struwe, W. B., and Benesch, J. L. P. (2015) Native mass spectrometry: towards high-throughput structural proteomics. *Methods Mol. Biol. 1261*, 349–371.

(19) Niesen, F. H., Berglund, H., and Vedadi, M. (2007) The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nature Protocols 2*, 2212–2221.

(20) Holdgate, G. A., and Ward, W. (2005) Measurements of binding thermodynamics in drug discovery. *Drug Discov. Today 10*, 1543–1550.

(21) Elliott, T. S., Slowey, A., Ye, Y., and Conway, S. J. (2012) The use of phosphate bioisosteres in medicinal chemistry and chemical biology. *Med. Chem. Commun. 3*, 735–751.

(22) Bennett, B. D., Kimball, E. H., Gao, M., Osterhout, R., Van Dien, S. J., and Rabinowitz, J. D. (2009) Absolute metabolite concentrations and implied enzyme active site occupancy in Escherichia coli. *Nat. Chem. Biol. 5*, 593–599.

(23) Saint Girons, I., Gilles, A. M., Margarita, D., Michelson, S., Monnot, M., Fermandjian, S., Danchin, A., and Barzu, O. (1987) Structural and catalytic characteristics of Escherichia coli adenylate kinase. *J. Biol. Chem. 262*, 622–

629.

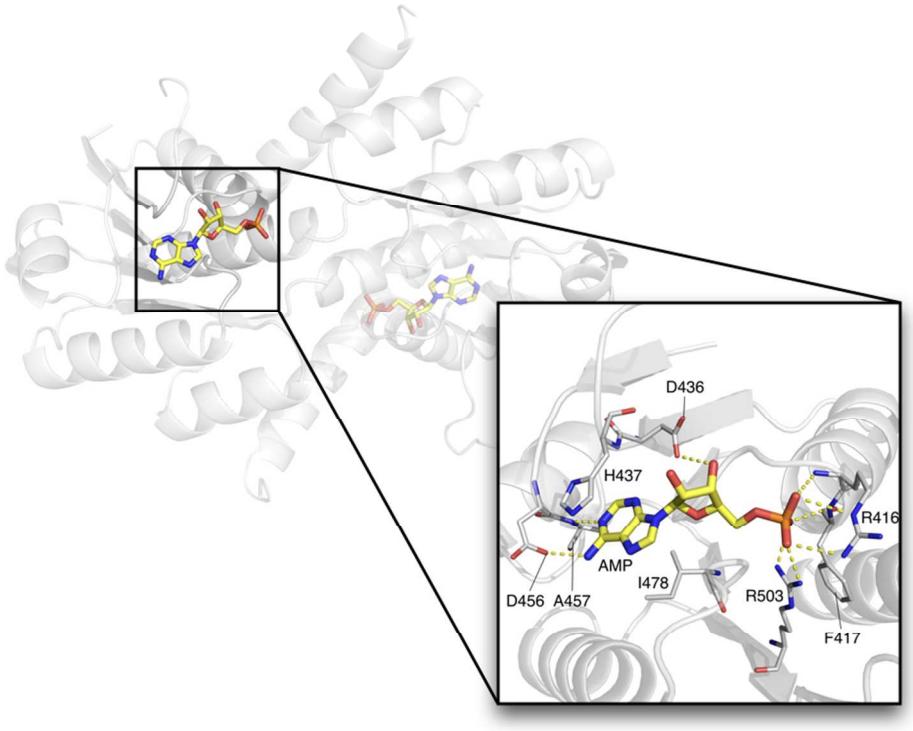
(24) Kröning, N., Willenborg, M., Tholema, N., Hänelt, I., Schmid, R., and Bakker, E. P. (2007) ATP binding to the KTN/RCK subunit KtrA from the K+ -uptake system KtrAB of Vibrio alginolyticus: its role in the formation of the KtrAB complex and its requirement in vivo. *J. Biol. Chem. 282*, 14018–14027.

(25) Healy, J., Rasmussen, T., Miller, S., Booth, I. R., and Conway, S. J. (2016) The photochemical thiol-ene reaction as a versatile method for the synthesis of glutathione S-conjugates targeting the bacterial potassium efflux system Kef. *Org. Chem. Front. 3*, 439–446.

(26) Lowry, O. H., Rosebrough, N. J., Farr, A. L., and Randall, R. J. (1951) Protein Measurement with the Folin Phenol Reagent. *J. Biol. Chem. 193*, 265–275.

(27) Aguilar, J. A., Nilsson, M., Bodenhausen, G., and Morris, G. A. (2012) Spin echo NMR spectra without J modulation. *Chem. Commun. 48*, 811–813.

(28) Chen, L., Wang, J., Zhang, Y.-Y., Yan, S. F., Neumann, D., Schlattner, U., Wang, Z.-X., and Wu, J.-W. (2012) AMP-activated protein kinase undergoes nucleotide-dependent conformational changes. *Nat. Struct. Mol. Biol. 19*, 716–718.



For Table of Contents Use Only

361x270mm (72 x 72 DPI)