

New insights into the domain of unknown function (DUF) of EccC₅, the pivotal ATPase providing the secretion driving force to the ESX-5 secretion system

Fernando Ceballos-Zúñiga,^a Margarita Menéndez^{b,c} and Inmaculada Pérez-Dorado^{a*}

Received 24 January 2024

Accepted 8 May 2024

Edited by B. Kobe, University of Queensland, Australia

Keywords: ESX systems; DUF domain of EccC₅; T7SS; secretion systems; mycobacteria; *Mycobacterium tuberculosis*; ATPase activity; crystallography.

PDB reference: DUF domain of EccC₅ from *Mycobacterium tuberculosis*, 8rin

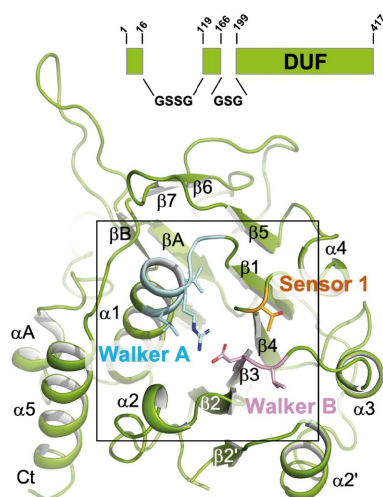
Supporting information: this article has supporting information at journals.iucr.org/d

^aDepartment of Crystallography and Structural Biology, Institute of Physical Chemistry Blas Cabrera, Spanish Research Council, Serrano 119, 28006 Madrid, Spain, ^bDepartment of Structure and Thermodynamics of Macromolecules, Institute of Physical Chemistry Blas Cabrera, Spanish Research Council, Serrano 119, 28006 Madrid, Spain, and ^cCIBER of Respiratory Diseases, ISCIII, Sinesio Delgado 10, 28029 Madrid, Spain. *Correspondence e-mail: jiperez@iqf.csic.es

Type VII secretion (T7S) systems, also referred to as ESAT-6 secretion (ESX) systems, are molecular machines that have gained great attention due to their implications in cell homeostasis and in host–pathogen interactions in mycobacteria. The latter include important human pathogens such as *Mycobacterium tuberculosis* (*Mtb*), the etiological cause of human tuberculosis, which constitutes a pandemic accounting for more than one million deaths every year. The ESX-5 system is exclusively found in slow-growing pathogenic mycobacteria, where it mediates the secretion of a large family of virulence factors: the PE and PPE proteins. The secretion driving force is provided by EccC₅, a multidomain ATPase that operates using four globular cytosolic domains: an N-terminal domain of unknown function (EccC₅^{DUF}) and three FtsK/SpoIIIE ATPase domains. Recent structural and functional studies of ESX-3 and ESX-5 systems have revealed EccC^{DUF} to be an ATPase-like fold domain with potential ATPase activity, the functionality of which is essential for secretion. Here, the crystal structure of the *Mtb*EccC₅^{DUF} domain is reported at 2.05 Å resolution, which reveals a nucleotide-free structure with degenerated *cis*-acting and *trans*-acting elements involved in ATP binding and hydrolysis. This crystallographic study, together with a biophysical assessment of the interaction of *Mtb*EccC₅^{DUF} with ATP/Mg²⁺, supports the absence of ATPase activity proposed for this domain. It is shown that this degeneration is also present in DUF domains from other ESX and ESX-like systems, which are likely to exhibit poor or null ATPase activity. Moreover, based on an *in silico* model of the N-terminal region of *Mtb*EccC₅^{DUF}, it is hypothesized that *Mtb*EccC₅^{DUF} is a degenerated ATPase domain that may have retained the ability to hexamerize. These observations draw attention to DUF domains as structural elements with potential implications in the opening and closure of the membrane pore during the secretion process via their involvement in inter-protomer interactions.

1. Introduction

Molecular machines are sophisticated protein complexes that are ubiquitously present in all cellular organisms (Miller & Enemark, 2016). They convert the chemical energy resulting from the hydrolysis of nucleoside triphosphates (NTPs) into the mechanical force needed in numerous cellular events such as DNA replication, protein degradation, cell motility and protein secretion, amongst others (Miller & Enemark, 2016; Schmidt *et al.*, 2012; Famelis *et al.*, 2023; Crosskey *et al.*, 2020). The members of the superfamily of ATPases associated with diverse cellular activities (AAA+ proteins) are key constituents of many molecular machines, where the mechanical work



is performed at the expense of adenosine triphosphate (ATP) hydrolysis (Miller & Enemark, 2016; Leipe *et al.*, 2002). Type VII secretion systems (T7SSs), also referred to as ESAT-6 secretion (ESX) systems, are AAA+-dependent molecular machines that are found in the *Actinomycetota* phylum and have gained great attention due to their implications in cell homeostasis and host–pathogen interactions in mycobacteria (Famelis *et al.*, 2023; Bunduc *et al.*, 2020; Houben *et al.*, 2014). The latter include devastating pathogenic species such as *Mycobacterium tuberculosis* (*Mtb*), the etiological cause of human tuberculosis (TB), which constitutes a pandemic that is responsible for more than one million deaths every year (World Health Organization, 2022). T7SSs are also found in bacteria belonging to the *Firmicutes* phylum (T7SSb), which include other important pathogens such as *Staphylococcus aureus*, *Listeria monocytogenes*, *Bacillus anthracis* and *B. subtilis* (Zoltner *et al.*, 2016; Mietrach *et al.*, 2020). Of note, T7Sb systems are distantly related to the T7Sa systems found in actinomycetes, thus constituting a divergent type of secretion systems.

Mycobacteria produce up to five ESX secretion systems (T7SSas), named ESX-1 to ESX-5, which consist of large membrane complexes that span the inner bacterial membrane. Among them, ESX-5 is found almost exclusively in slow-growing pathogenic mycobacteria, where it participates in nutrient uptake, intracellular colonization and modulation of the immune response during infection through the secretion of a large family of protein effectors: the PE and PPE proteins (Bunduc *et al.*, 2020; Houben *et al.*, 2014). These roles have important implications in the life cycle and virulence of the pathogens, thus indicating ESX-5 as a potential drug target (Bunduc *et al.*, 2021). Efforts aiming at structural–functional characterization of mycobacterial ESX secretion systems have resulted in structures of ESX-5 from *Mtb* and *M. xenopi* (*Mxp*) and of ESX-3 from *M. smegmatis* (*Msm*), which have illuminated the architecture of the pore complex (Famelis *et al.*, 2019; Poweleit *et al.*, 2019; Bunduc *et al.*, 2021; Beckham *et al.*, 2021; Fig. 1*a*). The latter can be defined as a hexamer of protomers, in which each protomer is formed by four different membrane proteins, EccB:EccC:EccD:EccE, interacting in a 1:1:2:1 stoichiometry. EccC is a pivotal component of the pore complex, providing the secretion driving force via the hydrolysis of ATP (Famelis *et al.*, 2019; Poweleit *et al.*, 2019; Bunduc *et al.*, 2021; Beckham *et al.*, 2021). In the ESX-2 to ESX-5 systems, EccC consists of two N-terminal transmembrane (TM) helices connected to two additional helices (the stalk region) and a domain of unknown function (EccC^{DUF}), which are followed by three FtsK/SpoIIIE AAA+ ATPase domains, here referred to as D1, D2 and D3 (EccC^{D1–D3}; Famelis *et al.*, 2023; Bunduc *et al.*, 2021; Fig. 1*b*). This multi-domain organization is also shared by EccC orthologues (EssC) found in *Firmicutes* (Mietrach *et al.*, 2020), which differ in ESX-1, where EccC is replaced by two functionally equivalent fragments: EccC_{1a}, containing DUF and D1 domains, and EccC_{1b}, containing D2 and D3 domains (Famelis *et al.*, 2023; Houben *et al.*, 2014). EccC/EssC enzymes are expected to operate through a mechanism involving their hexamerization. In line

with this, the cytosolic region of *Mtb*EccC₅ has been observed to transit from an extended/open state to a contracted/closed state in which the enzyme multimerizes to form the cytosolic chamber that is expected to accommodate the effectors to be secreted (Bunduc *et al.*, 2021; Fig. 1*a*).

The D1–D3 domains exhibit the three-layer α – β – α core structure typically found in prokaryotic AAA+ enzymes, which is frequently followed by a C-terminal α -helical lid domain that is missing in EccC/EssC enzymes (Zoltner *et al.*, 2016). The hexameric architecture observed in the ESX-3 and ESX-5 complexes is in line with the functional oligomeric state frequently found in prokaryotic AAA+ enzymes. The latter include the FtsK and SpoIIIE ATPases, which are close relatives of EccC and are reported to function as hexamers, and in which the ATPase sites are located at the interface between two adjacent protomers (Miller & Enemark, 2016; Leipe *et al.*, 2002; Bunduc *et al.*, 2021; Rosenberg *et al.*, 2015). In such cases, both subunits provide the *cis*- and *trans*-acting catalytic elements required to form the active site. The former includes the Walker A motif, the Walker B motif and Sensor 1, which are typically located in the loop connecting β 1 and α 1, the C-terminus of β 3 and the loop connecting β 4 and α 4, respectively. The Walker A motif, also referred to as the P-loop, consists of a glycine-rich sequence (G₁xxG₂xG₃K[S/T] in FtsK homologues) that is involved in stabilization of the phosphate group of the nucleotide via highly conserved lysine and serine/threonine residues that enable the correct orientation of the γ -phosphate required for ATP hydrolysis (Miller & Enemark, 2016; Leipe *et al.*, 2002). The Walker B motif provides two acidic residues (hhhhDE in FtsK homologues, where *h* is any hydrophobic amino acid), the second of which acts as the catalytic base that activates the water involved in nucleophilic attack on the γ -phosphate (Miller & Enemark, 2016; Leipe *et al.*, 2002). Sensor 1 is a key catalytic element consisting of a polar residue, which is typically an asparagine but can also be a serine, threonine or aspartate. It is located between the Walker motifs to assist in the correct orientation of the catalytic glutamate towards the nucleophilic water. *Trans*-acting elements are positively charged amino acids, mostly (but not exclusively) arginine residues, and thus are referred to as arginine fingers or sensors (Miller & Enemark, 2016; Leipe *et al.*, 2002). The latter include the Arg finger, which consists of an arginine (or lysine) residue located at the end of the α 4 helix that orients towards the neighbouring active site to form contacts with the γ -phosphate. Through this interaction, the Arg finger favours the transition state for hydrolysis, thus playing an important role in ATP hydrolysis and, in some cases, in enzyme multimerization (Miller & Enemark, 2016; Ogura *et al.*, 2004). The other two *trans*-acting motifs in AAA+ enzymes are Sensors 2 and 3, which are generally arginines that are involved in sensing and/or stabilizing the ATP/ADP-bound states and promoting conformational changes coupled to these states or nucleotide hydrolysis (Miller & Enemark, 2016; Li *et al.*, 2015).

The first structural and functional studies conducted by Rosenberg and coworkers on EccC from *Thermomonospora curvata* (*Tcr*EccC) showed that the D1 domain is an active

ATPase that exhibits fully conserved Walker A (GxxGxGK[S/T]) and Walker B (hhhhDE) motifs with respect to FtsKs as a close canonical orthologue, including the Arg-finger transacting motif (Rosenberg *et al.*, 2015). This and subsequent studies on EccCs/EssCs from *Mtb* and *S. aureus* showed that the Walker A and B motifs are, however, degenerated in the D2 and D3 domains, which aligns with the nondetectable or poor catalytic activities observed for these domains, which have been proposed to play a regulatory role (Rosenberg *et*

al., 2015; Wang *et al.*, 2020; Zoltner *et al.*, 2016). More recently, structural characterization of the *Msm*ESX3 complex revealed that the N-terminal domain of unknown function (DUF) exhibits an ATPase-like fold, as also observed in the subsequently reported DUF domains of the *Mtb*ESX-5 and *Mxp*ESX-5 complexes (Famelis *et al.*, 2019; Bunduc *et al.*, 2021; Beckham *et al.*, 2021). The authors described the presence of a Walker B motif (hhhhDD³²⁰) in *Msm*EccC₃^{DUF}, which is partially conserved with respect to FtsK orthologs

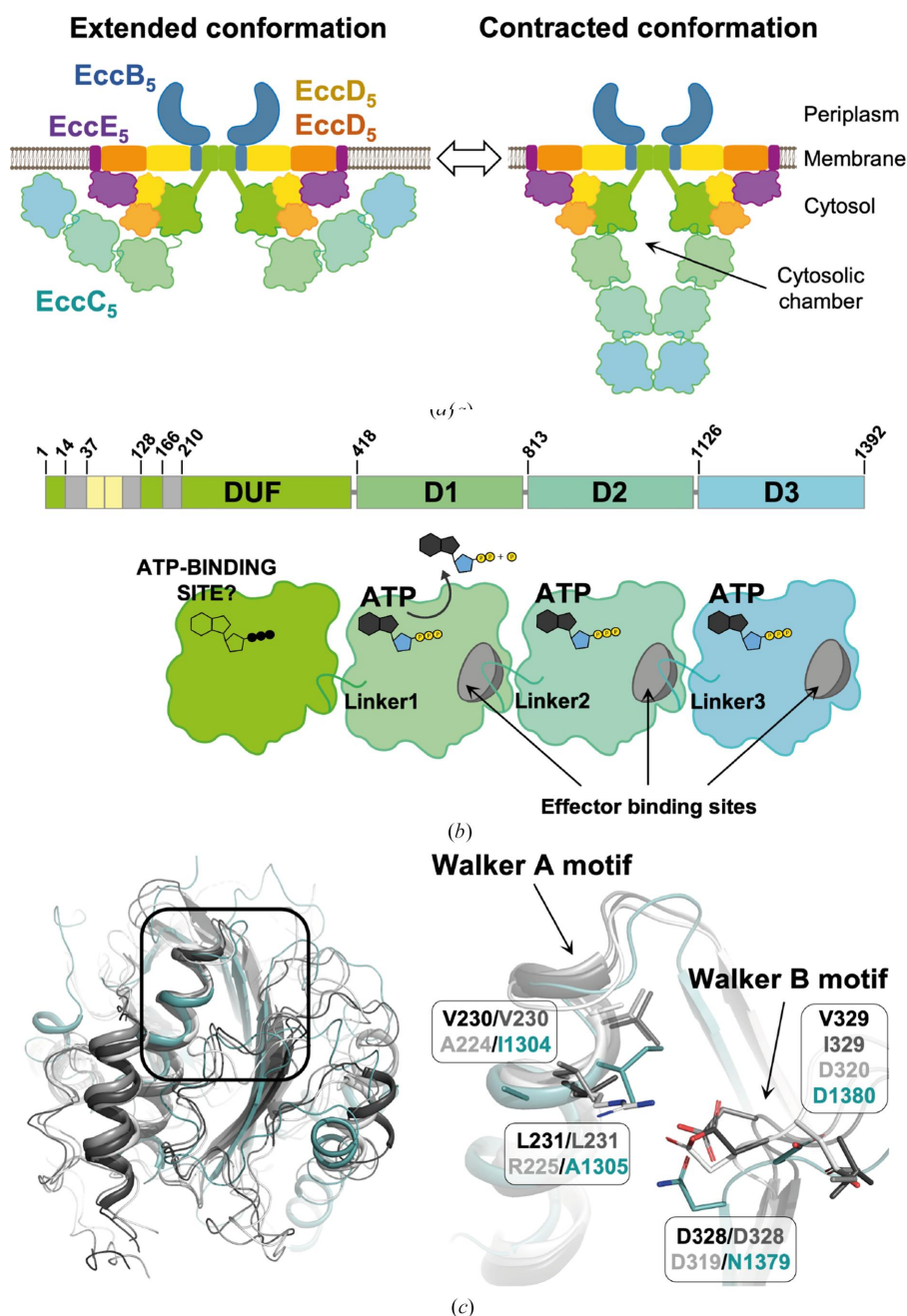


Figure 1
(a) *Mtb*ESX-5 model based on reported cryo-EM structures, showing the proposed conformational alteration of EccC₅ between an open/extended conformation (left) and a close/contracted conformation (right). (b) Domain architecture of EccC₅ consisting of the transmembrane (yellow), stalk (grey) and DUF (green) domains followed by three ATPase domains, D1, D2 and D3 (top), where the DUF and D1–D3 domains are interconnected by connectors referred to as Linker1, Linker2 and Linker3. (c) Superimposition of the *Srs*EssC^{D3} crystal structure (PDB entry 6vt1, blue) with cryo-EM models of *Mtb*EccC₅^{DUF} (PDB entry 7npr, dark grey), *Mxp*EccC₅^{DUF} (PDB entry 7b9s, light grey) and *Msm*EccC₃^{DUF} (PDB entry 6sgx, white) showing the conservation of the overall fold (left) and the similar arrangement of the Walker motifs (right).

(hhhhDE), and how the mutation of each aspartate to alanine obliterates secretion. These observations highlight the essentiality of *MsmEccC₃^{DUF}* in secretion and point to a potential ATPase activity of this domain (Famelis *et al.*, 2019). The structure of *MsmEccC₃^{DUF}* is however devoid of ATP, analogously to the *MtbEccC₅^{DUF}* and *MxpEccC₅^{DUF}* structures, which exhibit valine and isoleucine residues, respectively, that replace the catalytic amino acid (Famelis *et al.*, 2019; Bunduc *et al.*, 2021; Beckham *et al.*, 2021). These observations expose a variable degeneration of the Walker B motif across different *EccC^{DUF}* domains that may impair the ATPase activity to differing extents. In addition, the *MsmEccC₃^{DUF}*, *MtbEccC₅^{DUF}* and *MxpEccC₅^{DUF}* domain structures reveal a noncanonical arrangement of the Walker A motif lacking the typical P-loop structure, which is replaced by an extended α -helical conformation. Interestingly, this arrangement is also observed in the structure of *EssC^{D3}* from *S. aureus* (*SrsEssC^{D3}*; PDB entry 6tv1), which is also devoid of ATP and exhibits an ATP-binding affinity in the low-millimolar range (Mietrach *et al.*, 2020).

The structural and functional information available for T7SSs leaves outstanding questions regarding the molecular basis underlying the ATPase activity proposed for the DUF domain. In this regard, here we report the crystallographic structure of the *MtbEccC₅^{DUF}* domain at 2.05 Å resolution, which provides an unambiguous model showing a nucleotide-free structure with degenerated *cis*-acting and *trans*-acting elements involved in ATP binding and hydrolysis. Our high-resolution structure, together with a biophysical assessment of the interaction of *MtbEccC₅^{DUF}* with ATP/Mg²⁺ *in vitro*, supports the absence of ATPase activity in this domain. These results are in line with an *in silico* analysis carried out in other *EssC* and *EccC* enzymes, which reveals the presence of degenerated DUF domains in other mycobacterial and non-mycobacterial T7S systems that are likely to exhibit null or deficient ATPase activity. These findings suggest that DUF domains play a different role in the secretion process that, based on an *in silico* model of *MtbEccC₅* and the mutagenesis studies reported for *MsmESX3* (Famelis *et al.*, 2019), we propose it may be related to the aperture of the membrane-pore complex during the secretion process.

2. Materials and methods

2.1. Protein expression and purification

A protein construct containing the DUF domain of *EccC₅* from *M. tuberculosis* (*MtbEccC₅^{DUF}*, residues 1–417; UniProt entry P9WNA5) was designed for recombinant production. Residues 17–118 and 167–198 were replaced by GSSG and GSG sequences, respectively, in order to remove the trans-membrane (TM) and stalk regions and enable production as a soluble protein (see Fig. 2 and Supplementary Fig. S1). The construct contained an N-terminal 6×His tag followed by a SUMO tag and a PreScission 3C cleavage site. The DNA coding for the construct cloned into a pET-28a plasmid (between *NcoI* and *XhoI* restriction sites) was purchased from GenScript with codon optimization for *Escherichia coli*

expression (pET-*EccC₅^{DUF}*). Expression of *EccC₅^{DUF}* was performed in *E. coli* BL21 (DE3) Star (Invitrogen) cells using 2×TY medium supplemented with kanamycin (50 µg ml⁻¹). The cells were grown at 37°C to an OD_{600 nm} of 0.8 and then cooled to 16°C for an hour. Expression was then induced at 16°C by the addition of 1 mM isopropyl β -D-1-thiogalactopyranoside for 20 h. The cells were harvested by centrifugation at 5251g for 20 min at 4°C and the cell pellets were stored at –20°C.

The cell pellets were thawed and resuspended in 20 mM Tris–HCl pH 7.5, 500 mM NaCl, 10 mM imidazole (buffer A) supplemented with 5 mM MgCl₂, 1 mM MnCl₂ and 1 µg ml⁻¹ DNase (Sigma). This step was performed at room temperature, while subsequent purification steps were carried out at 4°C. The cells were disrupted by sonication and the clarified extract was loaded into a 1 ml HisTrap HP column (Cytiva) previously equilibrated in buffer A. Protein samples were eluted using an imidazole gradient (from 10 to 500 mM) and then buffer-exchanged to 20 mM Tris–HCl pH 7.5, 150 mM NaCl (buffer B) using a PD-10 desalting column (Cytiva). Samples were diluted to 1 mg ml⁻¹ in buffer B prior to overnight cleavage with 3C protease. The cleaved protein was concentrated using Amicon 10 kDa molecular-weight cutoff concentrators (Millipore) and further purified by size-exclusion chromatography (SEC) using a Superdex 200 (16/60) column (Cytiva) pre-equilibrated with buffer B or 20 mM HEPES pH 7.5, 150 mM NaCl, 5 mM MgCl₂ (buffer C) for crystallographic or ITC experiments, respectively. Pure protein fractions obtained from SEC were pooled, concentrated and flash-frozen in liquid nitrogen for storage at –80°C. Protein sample purity and quantification were assessed by SDS–PAGE and UV–Vis absorbance at 280 nm, respectively.

2.2. Protein crystallization, structure resolution and analysis

All crystallization experiments were carried out using the sitting-drop vapour-diffusion method at 20°C in 96-well MRC plates (Hampton Research) employing an Oryx8 robot (Douglas Instruments). Initial crystals of *EccC₅^{DUF}* were obtained in condition F2 of the Index screen. After optimization, the best crystals grew in 300 nl droplets formed by mixing 150 nl protein solution at 10 mg ml⁻¹ and 150 nl precipitant condition [25% (w/v) PEG MME 2K, 300 mM trimethylamine *N*-oxide, Tris–HCl pH 8.0]. The crystals were flash-cooled in liquid nitrogen without cryoprotection for X-ray data collection, which enabled the measurement of a complete X-ray data set to 2.05 Å resolution at 100 K using synchrotron radiation at ALBA, Barcelona, Spain. The data set was indexed and integrated with *XDS* (Kabsch, 2010) and scaled and reduced using *AIMLESS* (Evans & Murshudov, 2013). The crystals belonged to space group *P1*, with unit-cell parameters $a = 48.30$, $b = 52.24$, $c = 57.44$ Å, $\alpha = 88.39$, $\beta = 84.60$, $\gamma = 80.17^\circ$ (Table 1). Structure solution was carried out using the molecular-replacement method using *Phaser* (McCoy *et al.*, 2007) with the coordinates of the *EccC₅^{DUF}* structure (residues 1–416) solved by cryo-EM (PDB entry 7npt; Bunduc *et al.*, 2021) as the search model. A single and

unambiguous solution was obtained, which consisted of two EccC₅^{DUF} molecules in the asymmetric unit. The initial model was subjected to alternate cycles of model building with *Coot* (Emsley *et al.*, 2010) and refinement with *Phenix* (Liebschner *et al.*, 2019) and *BUSTER* (Bricogne *et al.*, 2017) using NCS restrictions. The electron-density maps allowed us to model both molecules of the asymmetric unit almost completely along with 287 water molecules. Residues 1–16-GSSG-119–123 (in chains *A* and *B*), 275–283 and 311–315 (in chain *A*) and 274–283 (in chain *B*) were not modelled due to poor electron density being observed for these residues. The geometry of the final model was validated using *MolProbity* (Chen *et al.*, 2010). Dihedral angles were analysed with *Coot* (Emsley *et al.*, 2010) and *Mogul* (version 2020.3.0). Figures were generated using *PyMOL* (version 2.0; Schrödinger). The atomic coordinates and structure factors have been deposited in the Protein Data Bank under accession code 8rin.

2.3. Differential scanning fluorimetry

Differential scanning fluorimetry (DSF) assays were carried out to estimate the inflection temperature (T_i) associated with unfolding transitions of the EccC₅^{DUF} construct in the absence and presence of nucleotide/Mg²⁺. The experiments consisted of monitoring the variation in the emission of fluorescence by buried and exposed tryptophan residues at wavelengths of 350 and 330 nm, respectively, using a Tycho NT.6 (NanoTemper). DSF experiments were set up in a final volume of 10 µl buffer *C* containing the EccC₅^{DUF} construct at 2 µM in the absence or presence of ADP–AlF₃ at 4 mM. Mixtures of protein with the ATP analogue were incubated for 15 min before DSF measurements. Graph representations and analysis were performed using the Tycho analysis interface. Three independent runs were used in each case to calculate mean T_i values and the corresponding standard deviations.

2.4. Isothermal titration calorimetry

The ATP-binding and ATPase activities of the EccC₅^{DUF} domain were assessed by isothermal titration calorimetry (ITC) using an ITC-VP instrument from GE Instruments. EccC₅^{DUF} at 61 µM was loaded into the cell at 25°C and was titrated with ATP at 1.3 mM (in the syringe) using 15 injections, firstly of 1 µl, then of 10 µl and subsequently of 20 µl at intervals of 5 min. Both protein and ATP were prepared in buffer *C*. Analysis of the results was carried out and final figures were generated using the *Origin 7* ITC software.

2.5. Primary-, secondary- and tertiary-structure analysis *in silico*

A comparative analysis of the putative ATP/Mg²⁺-binding site across DUF domains and comparison with canonical ATPases was performed by *Clustal Omega* multi-sequence alignment (Sievers *et al.*, 2011) and by structure superimposition in *PyMOL* (version 2.0; Schrödinger) using experimental coordinates deposited in the Protein Data Bank and models predicted by *AlphaFold* (Jumper *et al.*, 2021) available at UniProt. The ATPases used in sequence align-

Table 1

X-ray crystallographic statistics for *MtbEccC₅^{DUF}*.

Values in parentheses are for the highest resolution shell.

Data-collection statistics	
Space group	P1
No. of complexes in asymmetric unit	2
a, b, c (Å)	48.30, 52.24, 57.44
α, β, γ (°)	88.39, 84.60, 80.17
Wavelength (Å)	0.9792
Beamline	BL13-XALOC
Data-collection temperature (K)	100
Resolution range (Å)	57.18–2.05 (2.11–2.05)
$R_{\text{merge}}^{\dagger}$	0.055 (0.354)
$R_{\text{p.i.m.}}^{\ddagger}$	0.035 (0.242)
Total No. of reflections	114509 (6654)
No. of unique reflections	32869 (2122)
Mean $I/\sigma(I)$	9.6 (2.6)
CC _{1/2}	0.998 (0.869)
Completeness (%)	95.1 (77.9)
Multiplicity	3.5 (3.1)
Wilson B factor (Å ²)	29.5
Refinement	
No. of atoms	
Total	4219
Protein	3932
Water	287
R_{work} (%)	0.19
R_{free} (%)	0.22
$\langle B \rangle$ (Å ²)	34.59
R.m.s.d., bond lengths (Å)	0.008
R.m.s.d., angles (°)	0.88

$$\dagger R_{\text{merge}} = \frac{\sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle|}{\sum_{hkl} \sum_i I_i(hkl)} \quad \ddagger R_{\text{p.i.m.}} = \frac{\sum_{hkl} \{1/[N(hkl) - 1]\}^{1/2} \sum_i |I_i(hkl) - \langle I(hkl) \rangle|}{\sum_{hkl} \sum_i I_i(hkl)}$$

ment correspond to FtsK from *Pseudomonas aeruginosa* (*PrgFtsK*; UniProt entry Q9I0M3; PDB entry 2iut; Massey *et al.*, 2006), HerA DNA translocase from *Sulfolobus solfataricus* (*SsHerA*; UniProt entry Q97WG8; PDB entry 4d2i; Rzechorzek *et al.*, 2014), MCM helicase from *Pyrococcus furiosus* (*PfrMCM*; UniProt entry Q8U314; PDB entry 4r7y; Miller *et al.*, 2014), Vps4 from *Metallosphaera sedula* (*MsdVps4*; UniProt entry A4YHC5; PDB entry 4d81; Caillat *et al.*, 2015), EccC from *Thermomonospora curvata* (*TcrEccC*; UniProt entry D1A4G7), *MtbEccCa₁* (UniProt entry P9WNB3), *MtbEccC₂* (UniProt entry O05450), *MtbEccC₃* (UniProt entry P9WNA9), *MtbEccC₄* (UniProt entry P9WNA7), *MtbEccC₅* (UniProt entry P9WNA5; PDB entries 7npt and 7npr; Bunduc *et al.*, 2021), *MxpEccC₅^{DUF}* (UniProt entry I0RZIO; PDB entry 7b9s; Beckham *et al.*, 2021), *MsmEccCa₁* (UniProt entry A0A653FCM4), *MsmEccC₃* (UniProt entry A0QQ40; PDB entry 6sgx; Famelis *et al.*, 2019), *MsmEccC₄* (UniProt entry A0QSN0), EccC from *Geobacillus thermodenitrificans* (*GthEccC*; UniProt entry A4IKE7), EssC from *Staphylococcus aureus* (*SrsEssC*; UniProt entry Q932J9; PDB entry 6tv1), EssC from *Clostridium* sp. (*CspEssC*; UniProt entry A0A3R6STG9), EssC from *Streptococcus oralis* (*SrlEssC*; UniProt entry A0A656Z0M6) and YukB from *Bacillus subtilis* (*BsbYukB*; UniProt entry C0SPA7). Structural analysis of the dihedral angles and their occurrence in the Cambridge Crystallographic Data Centre (CCDC) database was performed using *Mogul* (version 2020.3.0). *AlphaFold2-multimer* (Jumper *et al.*, 2021; Evans *et al.*, 2021) was employed to explore potential protein–protein interactions

of *MtbEccC*₅^{DUF} with the *MtbPE25*–*PPE41* effector complex (UniProt entries I6X486 and Q79FE1, respectively; Abdallah *et al.*, 2006; Daleke *et al.*, 2012), inter-protomer interactions of *MtbEccC*₅^{DUF} with the ubiquitin-like domain of *MtbEccD*₅^{DUF} (UniProt entry P9WNP9) and the potential hexamerization of *MtbEccC*₅^{DUF}. Feasible predicted interfaces were analysed using *jsPISA* (Krissinel, 2015) in order to identify inter-

monomer/domain contacts, which were subsequently inspected and validated in *PyMOL*. *jsPISA* was also used to analyse the predicted interfaces using the interaction radar score, which estimates the likelihood of a biological ensemble based on a statistical analysis of all interfaces found in the PDB. The interaction radar is divided into probability circles (0–100%). The larger the area, the higher the probability of the interface

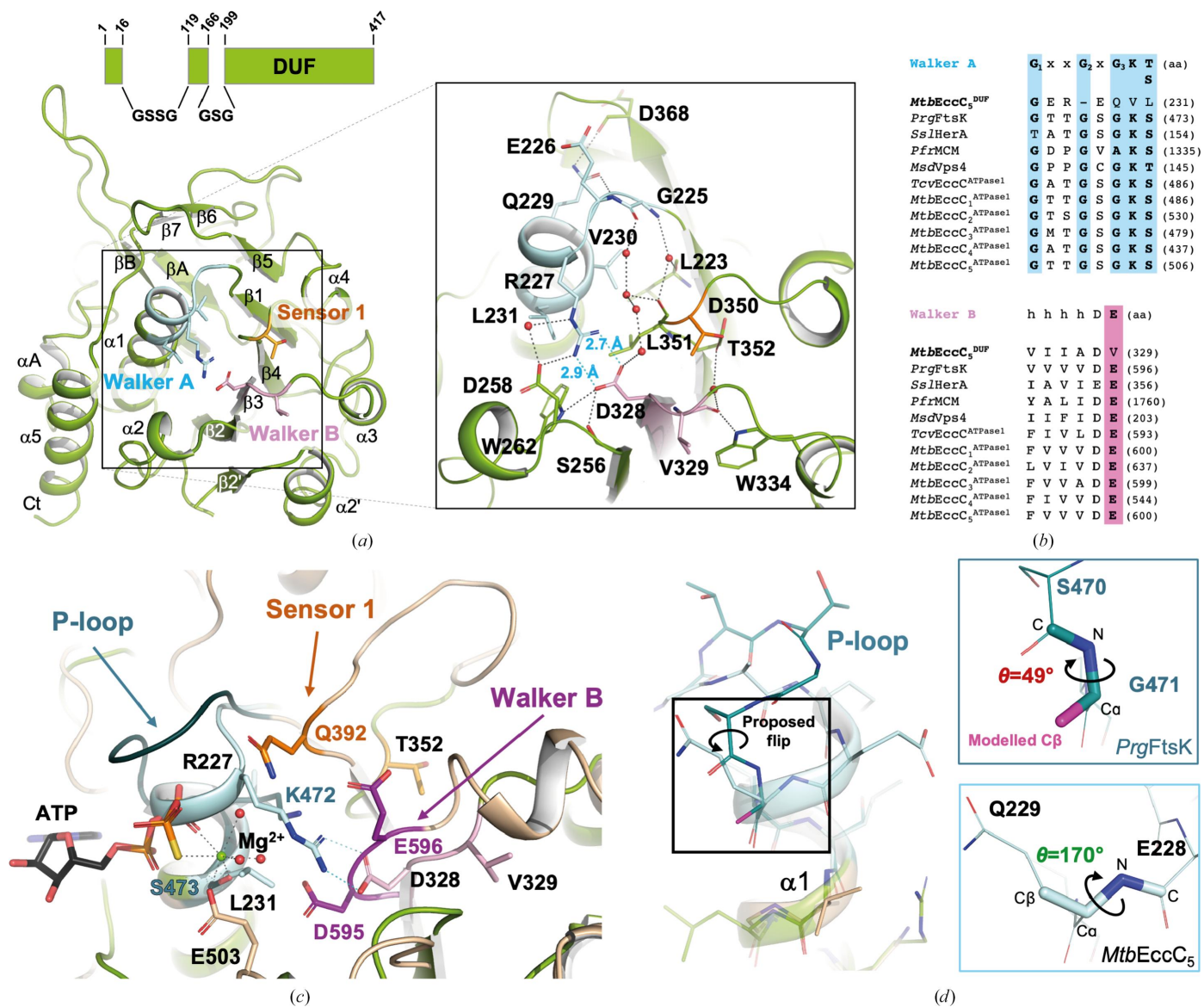


Figure 2
 (a) A general view of the *MtbEccC*₅^{DUF} crystal structure represented as green cartoons is shown (left), as well as details of the putative ATP-binding site (right). Salt-bridge and hydrogen-bond interactions are indicated as dashed lines in blue and black, respectively, as well as the bond distances between Arg227 and Asp328. The Walker A, Walker B and Sensor 1 motifs are coloured blue, pink and orange, respectively, showing key residues (sticks) and water molecules (red spheres). A scheme with details of the regions of *MtbEccC*₅^{DUF} included in the crystallographic construct is shown. (b) Sequence alignment of the Walker A and Walker B motifs of *MtbEccC*₅^{DUF} with the canonical ATPase domains of *PrgFtsK*, *SslHerA*, *PfrMCM*, *MsdVps4* and *MtbEccC*_{1–5}^{D1}, showing the consensus sequence at the top. Key amino acids involved in nucleotide interaction/hydrolysis are highlighted in blue and pink, respectively. (c) Superimposition of the *MtbEccC*₅^{DUF} and *PrgFtsK* (PDB entry 4r7y, beige) structures represented as cartoons. The Walker A, Walker B and Sensor 1 motifs are coloured light blue, pink and orange, respectively, in *MtbEccC*₅^{DUF} and in dark teal, purple and dark orange, respectively, in *PrgFtsK*. In *PrgFtsK*, the magnesium cation is represented as a green sphere, while ATP (in black) and residues key to the interaction with the nucleotide are represented as sticks. The Arg227–Asp328 salt bridge observed in *MtbEccC*₅^{DUF} and the hydrogen bonds involved in ATP/Mg²⁺ stabilization in *PrgFtsK* are represented as dashed lines in blue and black, respectively. (d) Detail of the Walker A motif in *MtbEccC*₅^{DUF} and *PrgFtsK*, showing the favourable dihedral angle C_i–N_{i+1}–C_{i+1}–C_{i+1}^β observed between Glu228 and Gln229 in the former compared with the nonfavourable dihedral angle generated between Ser470 and Gly471 of *PrgFtsK* when the glycine is substituted by an alanine.

having biological significance, and area values fitting or above the 50% probability circle are likely to be part of a biological ensemble (Krissinel, 2015).

3. Results and discussion

3.1. High-resolution structure of the $\text{EccC}_5^{\text{DUF}}$ domain

The overall crystallographic structure of the $\text{MtbEccC}_5^{\text{DUF}}$ domain (Fig. 2a) consists of an α - β - α sandwich that is highly homologous to the α/β fold found in the core of AAA+ domains. The core structure consists of a six-stranded parallel β -sheet ($\beta 5$ - $\beta 1$ - $\beta 4$ - $\beta 3$ - $\beta 2$ - $\beta 2'$) further extended by an additional two-stranded antiparallel sheet ($\beta 6$ - $\beta 7$). This β -sheet core is flanked by two helices ($\alpha 1$ and $\alpha 2$) on one side and four helices ($\alpha 2'A$, $\alpha 2'B$, $\alpha 3$ and $\alpha 4$) on the other side. The C-terminal α -helical lid present in other prokaryotic AAA+ domains is missing in the $\text{MtbEccC}_5^{\text{DUF}}$ structure, which is flanked by two parallel helices (αA and αB) and a two-stranded antiparallel sheet (βA - βB) that connects to the stalk domain. This structure is conserved in both $\text{MtbEccC}_5^{\text{DUF}}$ molecules present in the asymmetric unit (r.m.s.d. of 0.173 Å for 218 C^α atoms). Superimposition of the crystallographic and cryo-EM ($\text{MtbEccC}_5^{\text{CRYO}}$; PDB entry 7npr; Bunduc *et al.*, 2021) models shows overall structure conservation (r.m.s.d. of 0.671 Å for 198 C^α atoms), with the only significant differences in the N-terminal region (residues 1–16 are not defined in the electron-density maps). However, despite the overall conservation of the fold, other differences between our structure and $\text{MtbEccC}_5^{\text{CRYO}}$ are found, as will be described below.

The proposed ATP-binding site should be located at the N-terminus of the $\alpha 1$ helix and the C-terminus of the $\beta 3$ strand, which are expected to contain the P-loop and the Walker B motif, respectively. Sequence analysis of the Walker A motif shows a highly degenerated and one amino acid shorter sequence (GEREQVL231) that is missing the conserved G_2 and G_3 glycines as well as the Lys and Thr/Ser residues that are key for ATP/ Mg^{2+} stabilization (Val30 and Leu231 in $\text{MtbEccC}_5^{\text{DUF}}$; Fig. 2b). Inspection of the $\text{MtbEccC}_5^{\text{DUF}}$ crystal model shows an ATP-free structure in which the degenerated Walker A motif folds as an extended $\alpha 1$ helix instead of the standard P-loop conformation. In this arrangement, $\alpha 1$ would clash with the nucleotide molecule, as revealed by structural superimposition of $\text{MtbEccC}_5^{\text{DUF}}$ with FtsK from *P. aeruginosa* (*PrgFtsK*), a closely homologous canonical ATPase. To understand the structural basis of this extended $\alpha 1$ structure, we investigated the impact of G_2 and G_3 variations in $\text{MtbEccC}_5^{\text{DUF}}$, as invariant flexible features involved in both turns of the P-loop motif. To do so, we modelled the replacement of G_3 by alanine in *PrgFtsK* (G471A) *in silico*, which shows how the newly added C^β gives rise to an unfavoured dihedral angle C_i - N_{i+1} - C_{i+1}^α - C_{i+1}^β ($\theta = 49^\circ$) between Ser470 and Ala471 (Fig. 2d). This is supported by the low occurrence of this torsion angle when searching the CCDC database (Supplementary Fig. S1). This disfavoured conformation is prevented in $\text{MtbEccC}_5^{\text{DUF}}$ by a flip of the Glu228-Gln229 peptide bond, leading to the extended α -helical

structure observed in $\alpha 1$ (Fig. 2d). Analogously, the absence of G_2 should foster an α -helical conformation to avoid unfavourable dihedral angles between the two adjacent non-glycine amino acids. Altogether, this points to variations of G_2 and G_3 as structural factors responsible for the extended $\alpha 1$ helix observed in $\text{MtbEccC}_5^{\text{DUF}}$. Inspection of the Walker B motif also shows conformational deviations in $\text{MtbEccC}_5^{\text{DUF}}$, which include a shorter $\beta 3$ - $\alpha 3$ loop and a hydrophobic residue (Val329) replacing the catalytic base (Fig. 2a). Val329 rests partially buried between the $\beta 3$ - $\alpha 3$ and $\beta 4$ - $\alpha 4$ loops in a conformation that is well stabilized by a hydrogen-bond network with nearby residues. This drags the Walker B motif away from the P-loop by >2 Å compared with *PrgFtsK*, thus providing the room necessary to accommodate the extended $\alpha 1$ helix. This noncanonical arrangement of the Walker motifs is further stabilized by numerous direct and water-mediated interactions. These include a hydrogen bond between the side-chain amide group of Gln229 and the main-chain N atom of Glu226 and the formation of a salt bridge between the side chains of Arg227 and Asp328, which stabilize the extended $\alpha 1$ structure and are well defined in the electron-density maps (Supplementary Fig. S1). Moreover, the short bond distances measured for the salt-bridge interaction indicate that this is a strong contact that, overall, should further challenge the structural rearrangement of this nonproductive conformation observed in $\alpha 1$ for ATP binding. Superimposition of the crystal and cryo-EM structures of MtbEccC_5 shows overall conservation of the noncanonical arrangement observed in the Walker motifs in both models (Supplementary Fig. S1). However, significant differences were found in the structural features that shape the architecture of the degenerated ATP-binding site, supported by the electron-density and cryo-EM maps (Supplementary Fig. S1). These comprise many direct or water-mediated hydrogen bonds that stabilize the Walker and Sensor 1 motifs, including an interaction between Gln229 and Glu226 that is too long to form a hydrogen bond in $\text{MtbEccC}_5^{\text{CRYO}}$ and a water-mediated contact between the side chains of Arg227 and Asp258, which seems to assist the side chain of Arg227 in adopting an optimal orientation to interact with Asp328 and was not observed in the cryo-EM model. Moreover, in contrast to $\text{MtbEccC}_5^{\text{CRYO}}$, the crystal structure of $\text{MtbEccC}_5^{\text{DUF}}$ shows a bidentate salt bridge between the side chains of Arg27 and Asp328 that is well defined in the electron-density maps and contributes to optimize the interactions that stabilize the noncanonical configuration of the Walker A motif (Supplementary Fig. S1). Thus, our structure provides an unambiguous model defining the noncanonical ATP-binding site of MtbEccC_5 at high resolution, which has allowed us to acquire a detailed description and understanding of the key structural features contributing to its degenerated arrangement.

We next analysed the regions expected to contain the *trans*-acting elements in $\text{MtbEccC}_5^{\text{DUF}}$ (Supplementary Fig. S1). No basic residues were observed at the end of $\alpha 4$ as candidates to act as an Arg finger. Only Arg362 was found in the vicinity, and was further down the $\alpha 4$ sequence. However, its quite buried position in the structure makes it a poor candidate to

function as an Arg finger. The helical C-terminal lid, which is habitually found in prokaryotic AAA+ clades and should contain Sensor 2, is missing in *MtbEccC₅^{DUF}*. Besides, inspection of the $\alpha 3$ helix shows no basic residues in this region to form a potential Sensor 3. Taking all these observations together, our structural analysis of *MtbEccC₅^{DUF}* shows the degeneration and absence of *cis*- and *trans*-acting elements, respectively, in *MtbEccC₅^{DUF}*, which would support the lack of ATPase activity in this domain. These observations are in line with our failed attempts to co-crystallize *MtbEccC₅^{DUF}* with ATP/Mg²⁺, with structures devoid of nucleotide and magnesium instead being obtained.

3.2. ATP-binding and ATP-hydrolysis analysis

The structural arrangement of the putative ATP-binding site of *MtbEccC₅^{DUF}* resembles the nucleotide-binding pocket observed in the structure of the *SrsEssC^{D3}* domain (PDB entry 6tv1), which has been reported to bind ATP with an affinity in the low-millimolar range (Mietrach *et al.*, 2020; Fig. 3a). This poor but detectable binding points to potential structural rearrangements of the ATP-binding site of *SrsEssC^{D3}* that enable interaction with the nucleotide. Therefore, we analysed the interaction of *MtbEccC₅^{DUF}* with ATP/Mg²⁺ in order to explore a possible reorganization of the ATP-binding site that might enable nucleotide binding and hydrolysis in solution using two different biophysical approaches.

Firstly, the interaction of *MtbEccC₅^{DUF}* with ADP–AlF₃ (used as an ATP analogue) was assessed by DSF in the presence of magnesium. Despite the high ADP–AlF₃/EccC₅^{DUF} concentration ratio used (~4 mM/2 μ M), the experiments showed that the difference between *T_i* values in the presence (59.13 \pm 0.04°C) and absence (60.46 \pm 0.04°C) of nucleotide was slightly negative (around –1.3°C) (Fig. 3b) and very close to the error limit of the device, thus pointing to an absence of binding. Nucleotide interaction and ATP hydrolysis were next assessed by ITC in the presence of magnesium. The heat released upon the titration of ATP into *MtbEccC₅^{DUF}* was comparable to the thermal effect of ATP dilution in the range of concentrations tested (Fig. 3c). Additionally, no thermal power deflection (proportional to the potential substrate concentration in the cell) was observed after each ATP injection, as would be expected if the nucleotide was being hydrolysed (Menéndez, 2020). Overall, the ITC and DSF results support the lack of ATPase activity of *MtbEccC₅^{DUF}*.

3.3. The nucleotide-binding site in other EccC^{DUF} domains

Given the structural degeneration that was observed in the *MtbEccC₅^{DUF}* structure, and its observed inability to interact with ATP using X-ray crystallography, DSF and ITC, we decided to examine the putative ATP-binding site in other DUF domains. With this purpose, homologous EccC^{DUF} domains and orthologous EssC^{DUF} domains from myco-

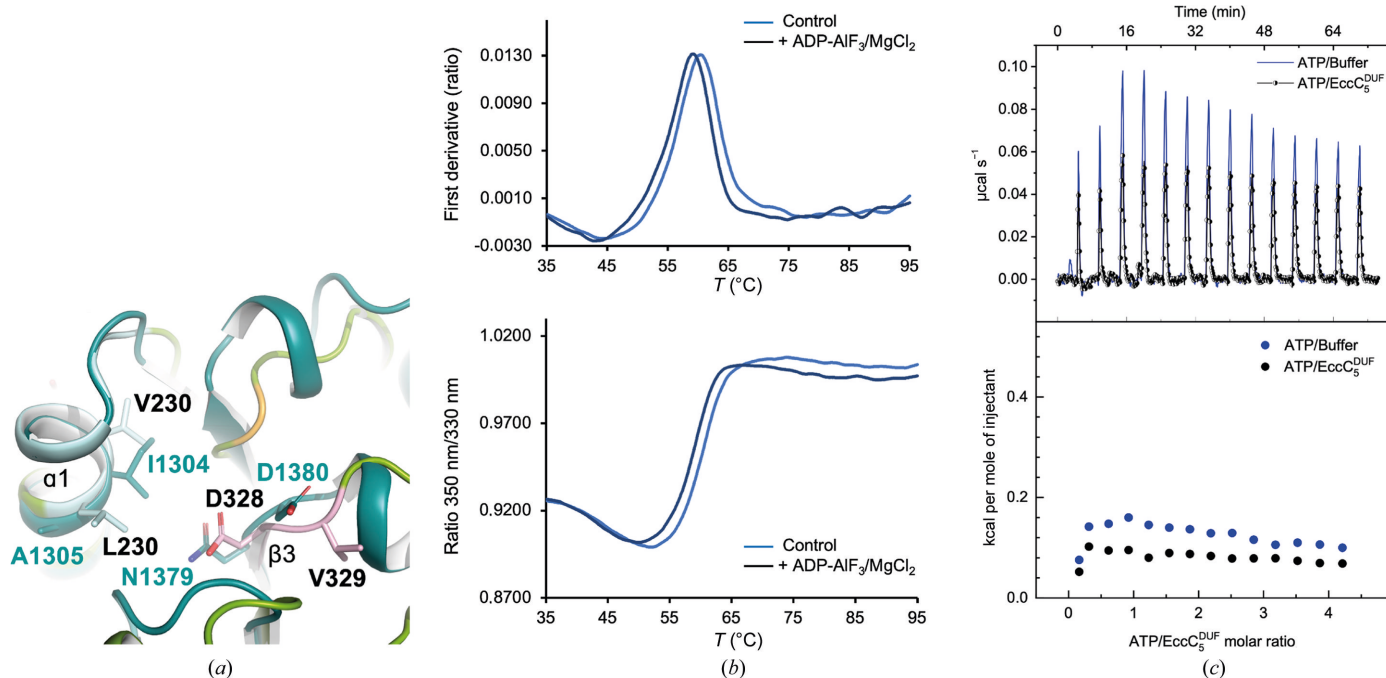


Figure 3 (a) Superimposition of the *MtbEccC₅^{DUF}* (green) and *SrsEssC^{D3}* (PDB entry 6tv1, dark blue) crystal structures showing details of their degenerated ATP-binding sites. Residues occupying key positions for nucleotide/magnesium interaction and hydrolysis are labelled and represented as sticks. The Walker A, Walker B and Sensor 1 motifs of *MtbEccC₅^{DUF}* are coloured following the same colour code as in Fig. 2. (b) Superimposition of *MtbEccC₅^{DUF}* DSF profiles in the absence (control) and presence of ADP–AlF₃ (4 mM) and Mg²⁺ (5 mM). The failure of ADP–AlF₃/Mg²⁺ to stabilize the domain structure against thermal denaturation was consistent with a lack of nucleotide binding. (c) ITC titration of ATP (1.3 mM) into EccC₅^{DUF} (60 μ M) and buffer C (black and blue data, respectively). The upper and lower panels show raw thermograms and normalized heat effect per mole of ATP injected versus ATP:EccC₅^{DUF} molar ratio, respectively.

bacterial and nonmycobacterial species were used in sequence alignment (Fig. 4a), which showed degeneration of the Walker A motif in all of the domains inspected. Analogously to *MtbEccC*₅^{DUF}, these motifs present a sequence that is one amino acid shorter, in which the conserved residues G₂ and G₃ are missing and substituted by a nonglycine amino acid, respectively. They also lack both the lysine and threonine/serine residues that are required for interaction with ATP/Mg²⁺. The DUF domains inspected also exhibit degeneration of their Walker B motifs, in which the catalytic glutamate is replaced by other hydrophobic or polar amino acids (Rosenberg *et al.*, 2015; Zoltner *et al.*, 2016; Wang *et al.*, 2020). Interestingly, analysis of the *MxpEccC*₅^{DUF} and *MsmEccC*₃^{DUF} structures showed that their Walker motifs adopt a similar arrangement to that observed in *MtbEccC*₅^{DUF}, including an extended α 1 helix (Fig. 4a). Analogously to *MtbEccC*₅^{DUF}, the extended α 1 helix can be explained by the amino-acid variations observed at the G₂ and G₃ positions with respect to canonical ATPases, which should elicit a helical conformation compared with the P-loop structure. This structural reorganization was also found in the structural models of *MtbEccC*_{1a}^{DUF}, *MtbEccC*₂^{DUF} and *MtbEccC*₄^{DUF} and in six other DUF domain structures predicted by *AlphaFold*, all of which present the same extended conformation of α 1 (Fig. 4 and Supplementary Fig. S2). Inspection of the Walker motifs in the *MxpEccC*₅^{DUF} and *MsmEccC*₃^{DUF} structures solved

by cryo-EM also reveal an arginine–aspartate salt bridge connecting α 1 to the C-terminus of β 3 (Fig. 4b). The salt bridge observed in *MxpEccC*₅^{DUF} involves an arginine at the N-terminus of α 1, analogously to as in *MtbEccC*₅^{DUF}, while the salt bridge in *MsmEccC*₃^{DUF} is formed by an arginine substituting the canonical threonine/serine of the Walker A motif (Fig. 4a). Of the 15 sequences analysed, five mycobacterial DUF domains exhibit an arginine/lysine and aspartate that could potentially form a salt bridge, which is also predicted in the *AlphaFold* models of *MtbEccC*₂^{DUF} and *MtbEccC*₄^{DUF}, where it is likely to contribute to stabilize this nonproductive arrangement for ATP-Mg²⁺ recognition (Fig. 4b). Thus, our analysis shows that the noncanonical structure observed in the *MtbEccC*₅^{DUF} Walker motifs is present in the cryo-EM structures of *MxpEccC*₅^{DUF} and *MsmEccC*₃^{DUF}, as well as in the predicted models of other DUFs from T7Sa and T7Sb systems. Together, these observations support the idea that the DUF domains are degenerated ATPase domains that exhibit a nonfunctional structure for ATP binding and thus a poor or null ATPase activity.

3.4. Hexameric model of the N-terminal region of *MtbEccC*₅

The lack of ATPase function of *EccC*^{DUF} domains proposed here raises important questions about their role in the secretion process. The relevance of these questions is particularly

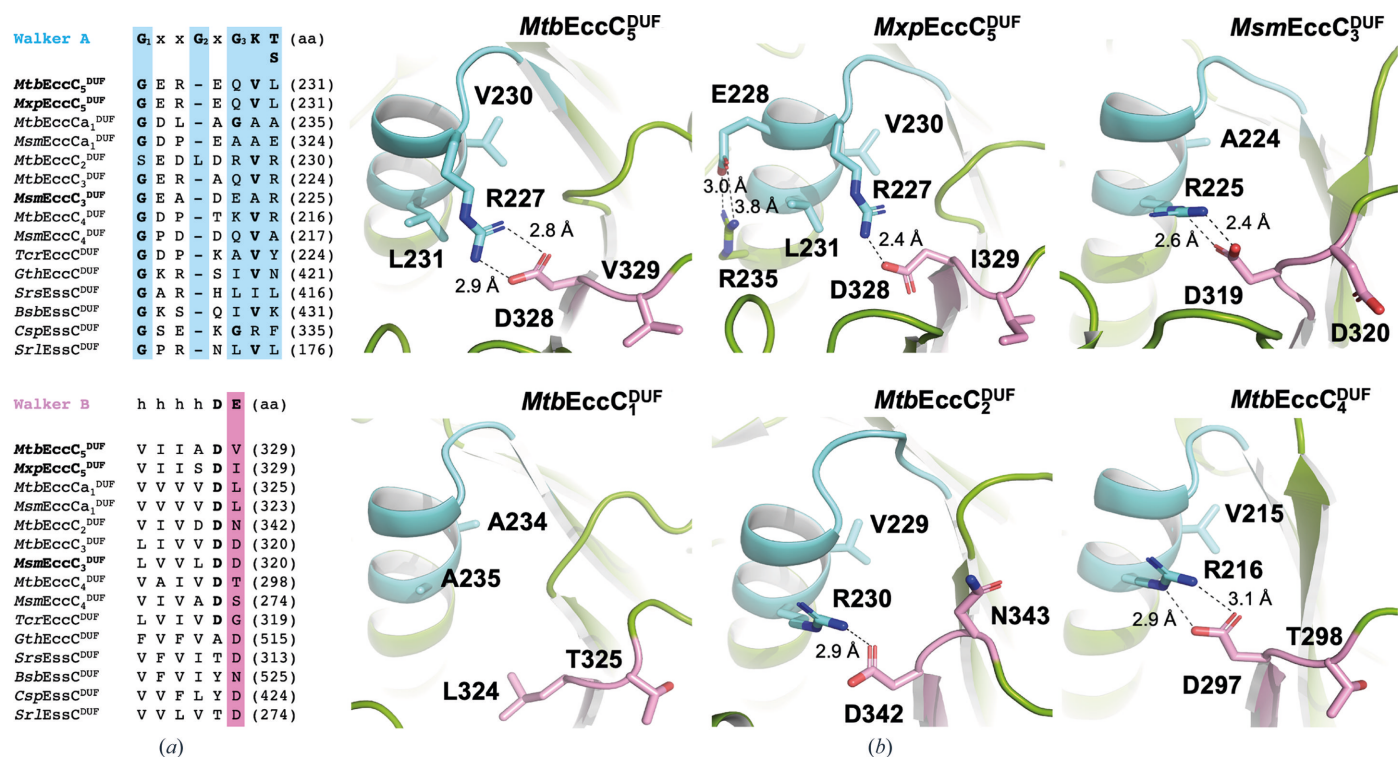


Figure 4 Analysis of a putative ATP/Mg²⁺-binding site in *EccC*^{DUF} domains of ESX-1–ESX-5 systems from *Mtb*, *Msm* and *Mxp*. (a) Sequence alignment of the Walker A and B motifs (coloured blue and pink, respectively). (b) Structural details of the Walker A and B motifs observed in *MtbEccC*₅^{DUF} (crystal structure), *MxpEccC*₅^{DUF} (PDB entry 7b9s), *MsmEccC*₃^{DUF} (PDB entry 6gsx), *MtbEccC*₁^{DUF} (*AlphaFold* model, UniProt entry P9WNB3), *MtbEccC*₂^{DUF} (*AlphaFold* model, UniProt entry O05450) and *MtbEccC*₄^{DUF} (*AlphaFold* model, UniProt entry P9WNA7) represented as cartoons and following the same colour code as in (a). Amino acids substituting key residues for ATP/Mg²⁺ interaction, or involved in hydrogen-bond/salt-bridge interactions, are depicted as sticks.

stressed by site-directed mutagenesis studies conducted on *MsmESX-3*, in which mutations of both of the aspartate residues of the Walker B motif of *MsmEccC₅^{DUF}* to alanines (D319A and D320A) were shown to abrogate secretion (Famelis *et al.*, 2019). This observation leads us to speculate that these mutations might either prevent proper folding of *MsmEccC₅^{DUF}* or affect protein–protein interactions that are key for secretion. Plausible DUF interactors might consist of PE/PPE effector proteins and other cytosolic ESX domains. Considering the second scenario, the ubiquitin-like domains of EccD (ULD or EccD^{ULD}) and DUFs of neighbouring protomers emerge as reasonable inter-protomer interactors, given their proximity in ESX complexes. Given the conservation of the ATPase-like fold of DUF, it is conceivable to expect this domain to have retained the ability to multimerize in the context of DUF–DUF interplay, as has been observed for FtsK homologues and proposed for EccC^{D1–D3} domains (Rosenberg *et al.*, 2015; Miller & Enemark, 2016). This hypothesis aligns with the size-exclusion chromatography profile observed during the production of *MtbEccC₅^{DUF}*, which indicates the presence of monomeric, trimeric and hexameric species (Supplementary Fig. S3). We thus attempted to explore the interaction of *MtbEccC₅^{DUF}* with the *MtbPEP25–PP41* effector complex, as well as *MtbEccC₅^{DUF}–EccC₅^{ULD}* and *MtbEccC₅^{DUF}–EccC₅^{DUF}* inter-protomer interactions, using *AlphaFold*. However, the predicted models showed unreliable and/or unfeasible interfaces (see the supporting information). This could be due to a genuine absence of these interactions, or to limitations in the modelling process arising either from the reliability of the theoretical model itself or the requirement for other domains/components of the cytosolic complex that were missing. Given these results, we decided to explore the potential multimerization of *MtbEccC₅^{DUF}* based on the experimental structures available of AAA+ enzymes in oligomeric states. To do so, we used the crystal coordinates of *PrgFtsK* in the hexameric state (PDB entry 2iuu) as a template to superimpose onto the N-terminal region of *MtbEccC₅*, including the TM, stalk and DUF domains, using the crystal coordinates of *MtbEccC₅^{DUF}* and *MtbEccC₅^{CRYO}* (PDB entry 7npr). This resulted in a hexameric model in which the *MtbEccC₅^{DUF}* domains form a closed ring, with the stalk and TM regions located perpendicular to the DUF domains (Fig. 5a). Analysis of the interface between adjacent DUFs shows that the Walker A, Walker B and Sensor 1 motifs are located near the neighbouring monomer, but they do not participate in inter-domain contacts (Fig. 5b). However, direct interactions via residues in the $\beta 2$ – $\alpha 2$ loop, $\alpha 2$ and $\beta 5$ – $\beta 6$ loop regions were identified. Interestingly, the $\beta 2$ – $\alpha 2$ region is located adjacent to the Walker A and B motifs, with which it forms direct interactions through several residues that include Arg227 and Asp328 (Fig. 5b). Based on these interactions, mutations in Arg227 and Asp328 and/or other amino acids of both Walker motifs might impact on the structure of the $\beta 2$ – $\alpha 2$ region and, by extension, on *MtbEccC₅^{DUF}* multimerization. This would also explain the lack of secretion observed in both Walker B motif mutants, D319A and D320A, of *MsmEccC₃^{DUF}* (Famelis *et al.*, 2019), which correspond to Walker B motif

residues Asp328 and Val329 in *MtbEccC₅^{DUF}*, respectively. Moreover, the residues identified at the DUF–DUF interface differ from those observed in the interaction with the EccC₅^{ULD} domain (Bunduc *et al.*, 2021), which shows the hexameric model of *MtbEccC₅^{DUF}* to be compatible with EccC₅^{DUF}–EccC₅^{ULD} interplay as a key structural element in ESX complexes (Famelis *et al.*, 2019; Bunduc *et al.*, 2021; Beckham *et al.*, 2021). With regard to the TM and stalk regions, these are located perpendicularly and at the edge of the ring formed by *MtbEccC₅^{DUF}* domains, with the TM regions of opposite monomers separated by a distance of ~ 86 Å. This disposition notably contrasts with the EccC₃ and EccC₅ organization observed in the *MsmESX-3*, *MtbESX-5* and *MxpESX-5* structures, in which the TM regions multimerize to form a helical bundle that connects, via the stalk region, to DUF domains positioned away from each other (Famelis *et al.*, 2019; Bunduc *et al.*, 2021; Beckham *et al.*, 2021; Fig. 5c). Of note, the TM helical bundle is located at the centre of the membrane region, where it contributes to a close conformation of the membrane pore. Thus, the aperture of the TM and stalk regions depicted by our model suggests a potential conformational change of the N-terminal region of *MtbEccC₅^{DUF}* that would explain how the membrane pore might open through hexamerization of the DUF domain (Fig. 5c). Given the distance measured across DUF domains in the *MtbEccC₅^{DUF}* model (~ 35 Å), and the dimensions of the periplasmic chamber observed in the *MtbESX-5* complex (Bunduc *et al.*, 2021), the proposed hexamerization could allow a pore opening with a diameter of 35–45 Å, which would be capable of accommodating the secreted PE–PPE heterodimers. Overall, our *MtbEccC₅* model suggests that DUF hexamerization is a plausible event that might play a critical role in the aperture of the membrane pore. This multimerization would be triggered by the interplay of DUFs with the EccC^{D1–D3} ATPase domains, thus linking ATP hydrolysis and effector recognition to the opening of the membrane pore necessary for secretion. Certainly, the inter-protomer interplay proposed here for DUF domains constitutes a hypothesis that requires further studies to be verified, while raising additional outstanding questions that require attention. For instance, how does the coupling of ATPase hydrolysis, effector recognition and multimerization of the D1–D3 ATPase domains occur with the proposed DUF–DUF and/or DUF–ULD interplay and the aperture of the membrane pore?

4. Conclusions

Here, we report the crystallographic structure of the *MtbEccC₅^{DUF}* domain at high resolution, which reveals an ATP/Mg²⁺-free structure with both high degeneracy and the absence of the *cis*- and *trans*-acting motifs required for the binding and hydrolysis of ATP, respectively. Among the most remarkable features, we find the absence of the catalytic glutamate required for nucleophilic attack on ATP and a noncanonical fold of the Walker A motif. Instead of the typical P-loop conformation, the Walker A motif adopts an extended α -helical arrangement that would clash with the nucleotide.

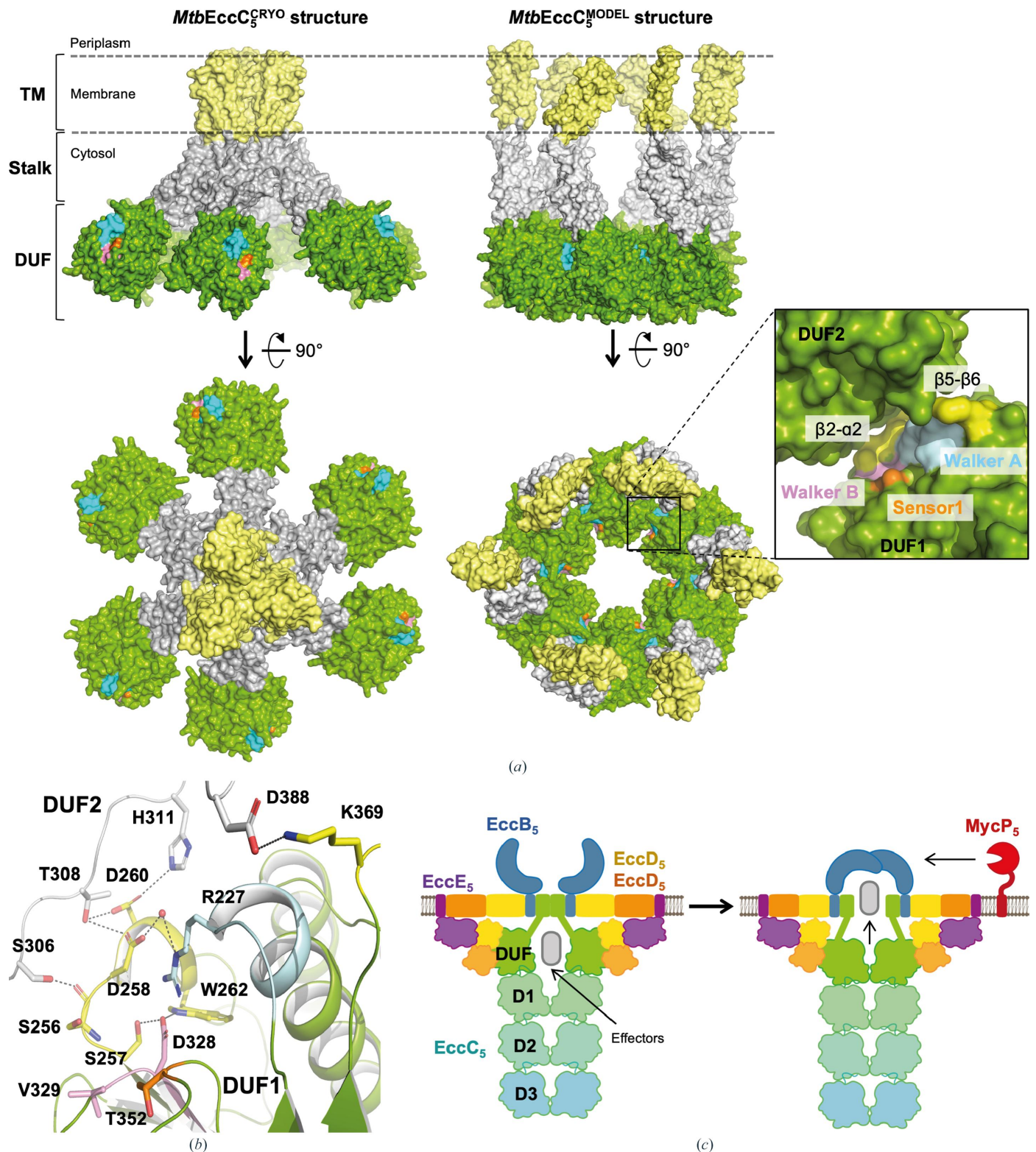


Figure 5
 (a) Structural comparison of the Ecc₅ arrangement observed in the *MtbESX-5* structure (PDB entry 7npr) with our model of the N-terminal region of *MtbEccC5* multimerizing as a hexamer. A surface representation of *MtbEccC5* molecules is shown, with the transmembrane, stalk and DUF domains coloured yellow, grey and green, respectively. The Walker A, Walker B and Sensor 1 motifs are coloured cyan, pink and orange, respectively. An enlarged view showing the molecular interface between adjacent DUF monomers in our model is shown in the box on the right, where identified contact regions are highlighted in yellow. (b) Details of the interface between adjacent DUF1 (grey) and DUF2 [using the colour code used in (a)] molecules are shown, with residues establishing contacts between the monomers and the Walker A and B motifs represented as sticks; their corresponding interactions are indicated as dashed lines. Water molecules are represented as red spheres. (c) Model showing the aperture of the TM region of *MtbEccC5* associated with the proposed multimerization of *MtbEccC5*^{DUF}, in which this domain would transition from a non-multimerized form to a hexameric state to enable opening of the membrane pore.

Although the overall conservation of this arrangement is also observed in *MtbEccC*₅^{CRYO}, our crystal structure provides previously unseen and unambiguously defined interactions that shape the observed noncanonical ATP-binding site. Moreover, our structural analysis of *MtbEccC*₅^{DUF} has led us to identify the structural features that underlie this arrangement. These include the variations of G₂ and G₃ observed in the Walker A motif of *MtbEccC*₅^{DUF}, which introduce geometrical restraints that favour the extended α 1 helical structure instead of the canonical P-loop conformation. We also demonstrated the lack of interaction of *MtbEccC*₅^{DUF} with the nucleotide in solution using DSF and ITC, which aligns with our futile attempts to co-crystallize this domain with ATP/Mg²⁺ and further supports the inability of the DUF domain to bind ATP. Our study also shows that the degenerated ATP-binding site observed in *MtbEccC*₅^{DUF} is present in the experimental and predicted structures of other homologous domains from T7Sa and T7Sb systems, which supports the notion that DUFs are degenerated ATPase domains that are nonfunctional in nucleotide interaction/hydrolysis. DUF domains emerge, however, as pivotal elements in the secretion process, as illustrated by mutagenesis studies of the Walker B motif of *MsmEccC*₃^{DUF} that led to the abrogation of *MsmESX*-3 secretion. Putting all of these observations together, we speculate on the role of DUFs as structural elements involved in inter-protomer interplay, including a possible hexamerization of this domain with potential implications for the opening of the membrane pore during the secretion process. This hypothesis, which requires further studies, opens new important questions such as how ATPase hydrolysis, effector recognition and multimerization of the D1–D3 ATPase domains would couple with the proposed hexamerization of DUF domains and the aperture of the membrane pore.

Acknowledgements

We thank the Synchrotron Radiation Facility ALBA, Barcelona, Spain for the time allocated for X-ray data collection under proposal 2022096995 and the BL13-XALOC beamline staff for their assistance in data collection. We would also like to thank Dr Lourdes Infantes San Mateo and Dr Rafael Molina-Monterrubio from the IQF–CSIC for their input into the structural analysis presented in this work.

Funding information

This work was funded by the Atracción de Talento (Modalidad 1) program from the Regional Government of Madrid, Spain (reference 2019-T1/BMD-14774). M. Menendez was supported by the CIBER of Respiratory Diseases (CIBERES), an initiative of the Spanish Institute of Health Carlos III (ISCIII).

References

Abdallah, A. M., Verboom, T., Hannes, F., Safi, M., Strong, M., Eisenberg, D., Musters, R. J. P., Vandenbroucke-Grauls, C. M. J. E.,

Appelmeik, B. J., Luirink, J. & Bitter, W. (2006). *Mol. Microbiol.* **62**, 667–679.

Beckham, K. S. H., Ritter, C., Chojnowski, G., Ziemianowicz, D. S., Mullapudi, E., Rettel, M., Savitski, M. M., Mortensen, S. A., Kosinski, J. & Wilmanns, M. (2021). *Sci. Adv.* **7**, eabg9923.

Bricogne, G., Blanc, E., Brandl, M., Flensburg, C., Keller, P., Paciorek, W., Roversi, P., Sharff, A., Smart, O. S., Vornrhein, C. & Womack, T. O. (2017). *BUSTER* version 2.11.2. Global Phasing Ltd, Cambridge, United Kingdom.

Bunduc, C. M., Bitter, W. & Houben, E. N. G. (2020). *Annu. Rev. Microbiol.* **74**, 315–335.

Bunduc, C. M., Fahrenkamp, D., Wald, J., Ummels, R., Bitter, W., Houben, E. N. G. & Marlovits, T. C. (2021). *Nature*, **593**, 445–448.

Caillat, C., Macheboeuf, P., Wu, Y., McCarthy, A. A., Boeri-Erba, E., Effantin, G., Göttinger, H. G., Weissenhorn, W. & Renesto, P. (2015). *Nat. Commun.* **6**, 8781.

Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. (2010). *Acta Cryst.* **D66**, 12–21.

Crosskey, T. D., Beckham, K. S. H. & Wilmanns, M. (2020). *Prog. Biophys. Mol. Biol.* **152**, 25–34.

Daleke, M. H., Ummels, R., Bawono, P., Heringa, J., Vandenbroucke-Grauls, C. M. J. E., Luirink, J. & Bitter, W. (2012). *Proc. Natl Acad. Sci. USA*, **109**, 11342–11347.

Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* **D66**, 486–501.

Evans, P. R. & Murshudov, G. N. (2013). *Acta Cryst.* **D69**, 1204–1214.

Evans, R., O'Neill, M., Pritzel, A., Antropova, N., Senior, A., Green, T., Židek, A., Bates, R., Blackwell, S., Yim, J., Ronneberger, O., Bodenstern, S., Zielinski, M., Bridgland, A., Potapenko, A., Cowie, A., Tunyasuvunakool, K., Jain, R., Clancy, E., Kohli, P., Jumper, J. & Hassabis, D. (2021). *bioRxiv*, 2021.10.04.463034.

Famelis, N., Geibel, S. & van Tol, D. (2023). *Biol. Chem.* **404**, 691–702.

Famelis, N., Rivera-Calzada, A., Degliesposti, G., Wingender, M., Mietrach, N., Skehel, J. M., Fernandez-Leiro, R., Böttcher, B., Schlosser, A., Llorca, O. & Geibel, S. (2019). *Nature*, **576**, 321–325.

Houben, E. N. G., Korotkov, K. V. & Bitter, W. (2014). *Biochim. Biophys. Acta*, **1843**, 1707–1716.

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Židek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohli, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., Back, T., Petersen, S., Reiman, D., Clancy, E., Zielinski, M., Steinegger, M., Pacholska, M., Berghammer, T., Bodenstern, S., Silver, D., Vinyals, O., Senior, A. W., Kavukcuoglu, K., Kohli, P. & Hassabis, D. (2021). *Nature*, **596**, 583–589.

Kabsch, W. (2010). *Acta Cryst.* **D66**, 125–132.

Krissinel, E. (2015). *Nucleic Acids Res.* **43**, W314–W319.

Leipe, D. D., Wolf, Y. I., Koonin, E. V. & Aravind, L. (2002). *J. Mol. Biol.* **317**, 41–72.

Liebschner, D., Afonine, P. V., Baker, M. L., Bunkóczi, G., Chen, V. B., Croll, T. I., Hintze, B., Hung, L.-W., Jain, S., McCoy, A. J., Moriarty, N. W., Oeffner, R. D., Poon, B. K., Prisant, M. G., Read, R. J., Richardson, J. S., Richardson, D. C., Sammito, M. D., Sobolev, O. V., Stockwell, D. H., Terwilliger, T. C., Urzhumtsev, A. G., Videau, L. L., Williams, C. J. & Adams, P. D. (2019). *Acta Cryst.* **D75**, 861–877.

Li, N., Zhai, Y., Zhang, Y., Li, W., Yang, M., Lei, J., Tye, B. K. & Gao, N. (2015). *Nature*, **524**, 186–191.

Massey, T. H., Mercogliano, C. P., Yates, J., Sherratt, D. J. & Löwe, J. (2006). *Mol. Cell*, **23**, 457–469.

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.

Menéndez, M. (2020). *Encyclopedia of Life Sciences*, pp. 113–127. Chichester: John Wiley & Sons.

- Mietrach, N., Damián-Aparicio, D., Mielich-Süss, B., Lopez, D. & Geibel, S. (2020). *J. Bacteriol.* **202**, 646–665.
- Miller, J. M., Arachea, B. T., Epling, L. B. & Enemark, E. J. (2014). *eLife*, **3**, e03433.
- Miller, J. M. & Enemark, E. J. (2016). *Archaea*, **2016**, 9294307.
- Ogura, T., Whiteheart, S. W. & Wilkinson, A. J. (2004). *J. Struct. Biol.* **146**, 106–112.
- Poweleit, N., Czudnochowski, N., Nakagawa, R., Trinidad, D., Murphy, K. C., Sasseti, C. & Rosenberg, O. S. (2019). *eLife*, **8**, e52983.
- Rosenberg, O. S., Dovala, D., Li, X., Connolly, L., Bendebury, A., Finer-Moore, J., Holton, J., Cheng, Y., Stroud, R. M. & Cox, J. S. (2015). *Cell*, **161**, 501–512.
- Rzechorzek, N. J., Blackwood, J. K., Bray, S. M., Maman, J. D., Pellegrini, L. & Robinson, N. P. (2014). *Nat. Commun.* **5**, 5506.
- Schmidt, H., Gleave, E. S. & Carter, A. P. (2012). *Nat. Struct. Mol. Biol.* **19**, 492–497.
- Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., Thompson, J. D. & Higgins, D. G. (2011). *Mol. Syst. Biol.* **7**, 539.
- Wang, S., Zhou, K., Yang, X., Zhang, B., Zhao, Y., Xiao, Y., Yang, X., Yang, H., Guddat, L. W., Li, J. & Rao, Z. (2020). *Protein Cell*, **11**, 124–137.
- World Health Organization (2022). *Global Tuberculosis Report 2022*. Geneva: World Health Organization. <https://www.who.int/teams/global-tuberculosis-programme/tb-reports/global-tuberculosis-report-2022>.
- Zoltner, M., Ng, W. M. A. V., Money, J. J., Fyfe, P. K., Kneuper, H., Palmer, T. & Hunter, W. N. (2016). *Biochem. J.* **473**, 1941–1952.