Structure of the mycobacterial ESX-5 type VII secretion system membrane complex by single particle analysis

Katherine S.H. Beckham[1#], Luciano Ciccarelli[2,3,4#], Catalin M. Bunduc[5#], Haydyn D.T. Mertens[1], Roy Ummels[6], Wolfgang Lugmayr[2,3,4], Julia Mayr[2,3,4], Mandy Rettel[7], Mikhail M. Savitski[7], Dmitri I. Svergun[1], Wilbert Bitter[5,6], Matthias Wilmanns[1,4,8], Thomas C. Marlovits[2,3,4,8,9*], Annabel H.A. Parret[1*], Edith N.G. Houben[5*]

1, European Molecular Biology Laboratory, Hamburg Unit, Notkestraße 85, 22607, Hamburg, Germany
2, Research Institute of Molecular Pathology, Dr. Bohr-Gasse 7, 1030 Vienna, Austria
3, Institute of Molecular Biotechnology, Austrian Academy of Sciences, Dr. Bohr-Gasse 3, 1030 Vienna, Austria
4, University Medical Centre Hamburg-Eppendorf, Martinistraße 52, 20246 Hamburg, Germany
5, Vrije Universiteit Amsterdam, De Boelelaan 1108, 1081 HZ Amsterdam, The Netherlands
6, VU University Medical Center, De Boelelaan 1108, 1081 HZ Amsterdam, The Netherlands
7, European Molecular Biology Laboratory, Meyerhofstraße 1, 69117 Heidelberg, Germany
8, Centre for Structural Systems Biology, Notkestraße 85, 22607 Hamburg, Germany
9, Deutsches Elektronen-Synchrotron, Notkestraße 85, 22607 Hamburg, Germany

# Contributed equally to this work

* Corresponding author

SUMMARY

Mycobacteria are characterized by their impermeable outer membrane that is rich in mycolic acids[1]. To transport substrates across this complex cell envelope, mycobacteria rely on type VII (also known as ESX) secretion systems[2]. In *Mycobacterium tuberculosis*, these ESX systems are essential for growth and full virulence and therefore represent an attractive target for anti-tuberculosis drugs[3]. However, the molecular details underlying type VII secretion are largely unknown, due to a lack of structural information. Here, we report the molecular architecture of the ESX-5 membrane complex from *Mycobacterium xenopi* determined at 13 Å resolution by electron microscopy. The four core proteins of the ESX-5 complex ($EccB_5$, $EccC_5$, $EccD_5$ and $EccE_5$) assemble with equimolar stoichiometry into an oligomeric assembly that displays six-fold symmetry. This membrane-associated complex seems to be embedded exclusively in the inner membrane, which indicates that additional components are required to translocate substrates across the mycobacterial outer membrane. Furthermore, the extended cytosolic domains of the EccC ATPase, which interact with secretion effectors, are highly flexible, suggesting an as yet unseen mode of substrate interaction. Comparison of our results with known structures of other bacterial secretion systems demonstrates that the architecture of type VII secretion system is fundamentally different, suggesting an alternative secretion mechanism.

MAIN

During the past decade, it has become apparent that the type VII secretion system (T7SS) is essential for mycobacterial virulence and physiology[4]. The first T7SS to be identified was the 6 kDa early secretory antigenic target (ESAT-6) protein family secretion (ESX)-1 system of the human pathogen *Mycobacterium tuberculosis*[5–7]. Four other paralogous and highly conserved ESX systems have since been described in mycobacteria, named ESX-2 to ESX-5. Each ESX system has a defined role within the bacterial cell that is mediated by a diverse repertoire of secreted effector proteins, with each system exclusively secreting its cognate effectors[8]. Intriguingly, unlike most bacterial secretion systems, such as the type III secretion

system (T3SS), the ESX systems mediate the export of folded substrates across the diderm mycobacterial cell envelope, indicating a novel mode of protein transport.

The ESX-5 system is found only in slow-growing species of mycobacteria, which includes most pathogenic species. In addition to being crucial for nutrient uptake[7], the ESX-5 system secretes two large families of effector proteins – the Pro-Glu (PE) and the Pro-Pro-Glu (PPE) proteins – several of which are required for immune evasion and virulence[9]. The ESX-5 membrane complex that can be stably purified has an estimated size of 1.5 MDa, and comprises the core proteins $EccB_5$, $EccC_5$, $EccD_5$ and $EccE_5$[10]. Whether the four core components are sufficient to span the entire diderm cell envelope of mycobacteria is not clear[1,11]. In addition, the structural and functional roles of the core components within the complex have not been explored, with the exception of $EccC_5$[12]. As a member of the FtsK/SpoIIIE ATPase family, $EccC_5$ is presumed to be the coupling component that mediates effector protein recognition and energizes translocation through the channel[13]. Elucidating the architecture of the ESX-5 core complex is crucial to further our understanding of the mechanism of T7SS-mediated protein transport across the complex mycobacterial cell envelope.

The absence of an endogenous ESX-5 system in the avirulent and fast-growing species *Mycobacterium smegmatis* offers a unique opportunity to investigate the functionality and structure of this system[2]. After screening various mycobacterial species, the ESX-5 system from the mildly thermophilic and slow-growing species *Mycobacterium xenopi* was selected for structural analysis because of its high expression levels and superior stability. To recombinantly express and purify this complex, the *esx-5* locus from *M. xenopi* was cloned into a mycobacterial expression vector and modified to produce $EccC_5$ with a streptavidin tag (Strep-tag II) at its carboxy terminus (Supplementary Fig. 1a). Secretion analysis of the Strep-tagged ESX-5 system in *M. smegmatis* indicated that this system is functional and able to export EsxN, which is encoded by the introduced *esx-5* locus, as well as the heterologously expressed PPE18 from *Mycobacterium marinum*[14] into the culture filtrate (Fig. 1a). Previous data suggested that secreted effector proteins have a potential role in initiating the assembly of purified EccC[12]. We investigated whether substrates are similarly involved in complex assembly by generating an expression construct devoid of the six secreted substrates,

including Esx, PE and PPE proteins, encoded within the *esx-5* locus (Supplementary Fig. 1a). Notably, *M. smegmatis* lacks endogenous ESX-5 substrates and to date, no redundancy in substrate secretion between different ESX systems has been observed. Using blue native polyacrylamide gel electrophoresis (BN-PAGE) followed by immunoblot analysis of isolated membranes, we found that biogenesis of the membrane complex occurs in the absence of these proteins, thus establishing that the ESX-5 complex assembles in a substrate-independent manner, in contrast to what was previously reported for purified EccC[12] (Supplementary Fig. 1b).

To obtain insight into the architecture of the ESX-5 secretion channel, we solubilized membrane preparations with the detergent n-dodecyl β-D-maltoside (DDM) and purified the complex by affinity purification and size-exclusion chromatography (SEC) (Supplementary Fig. 2a). Presence of the four core components $EccB_5$, $EccC_5$, $EccD_5$ and $EccE_5$ was confirmed by SDS-PAGE and mass spectrometry (Fig. 1b). BN-PAGE analysis of the ESX-5 $EccBC_{strep}DE$ complex, hereafter simply referred to as ESX-5 complex, shows a band at approximately 1.5 MDa (Supplementary Fig. 2b) consistent with previous studies[15].

Electron microscopy (EM) was used to visualize the purified complexes. Visual inspection of the electron micrographs of negatively stained DDM-solubilized sample showed the presence of large particles in multiple orientations. We also observed smaller particles, which may represent sub-complexes also seen by BN-PAGE (Supplementary Fig. 2b). To improve sample stability, the DDM in the affinity-purified samples was exchanged to the surfactant amphipol A8-35 prior to SEC purification. This considerably reduces the presence of sub-complexes in the sample and thus improved the quality of the electron microscopic images obtained (Fig. 1c, Supplementary Fig. 2b-d). The inherent ATP hydrolysing activity and hence functional integrity of the purified complex was established using radioactive ATP (γ-^{32}P) (Fig. 1d). In addition, determination of the relative abundance of each component by mass spectrometry indicate that the complex contains all four components in equimolar amounts (Fig. 1e, Supplementary Table 1). An interesting feature observed in the raw images were tail-like extensions, projecting away from the compact globular complex (Fig. 1f). The structure of these approximately 14 nm elongated projections is reminiscent of the electron microscopic 'beads-on-a-string' structure of the *Thermomonospora curvata* EccC[12].

Single-particle analysis and reference-free 2D class averaging of 1,418 particles, followed by an initial 3D reconstruction of the complex without symmetry constraints, revealed a hexameric arrangement around the central axis (Fig. 1g-h, Supplementary Fig. 3a-b, Supplementary Table 2). Further refinement applying six-fold symmetry resulted in a map at 13 Å resolution that shows a compact, globular macromolecular assembly of approximately 28 x 16 nm (Fig. 2a, Supplementary Fig. 3c). Combining our observations on ESX-5 symmetry and our mass spectrometry data leads to a 24-subunit (i.e. 6 x 4) ESX-5 complex with a mass of 1.8 MDa. The dimensions of the complex indicate that the ESX-5 core complex is embedded in the inner membrane, rather than spanning both cell membranes. A central pore extends through the complex, with a diameter of approximately 5 nm (Fig. 2b). Such pore dimensions are sufficient to allow the passage of folded, dimeric T7SS effectors such as the Esx proteins that have an estimated width of 2 nm[16]. The long extensions observed in the electron micrographs of individual single particles are, however, not visible in the 3D reconstruction, probably due to the averaging effects of multiple flexible states.

To confirm that these elongated structures are indeed the integral membrane ATPase, we performed on-the-grid gold labelling of Strep-tagged $EccC_5$ on individual isolated complexes (Fig. 3a). $EccC_5$ comprises an N-terminal double-pass transmembrane domain (TMD), an unclassified domain of unknown function (DUF) and three cytosolic ATPase domains (A1, A2, A3)[4,12]. To investigate the flexibility of the cytosolic segment of $EccC_5$ in solution, a truncated version of $EccC_5$ lacking the TMD ($EccC_5^{\Delta 99}$) was generated. SEC coupled with multi-angle light scattering revealed that $EccC_5^{\Delta 99}$ exists primarily as a monomer in solution (Supplementary Fig. 4a). Small-angle X-ray scattering (SAXS) data collected on monomeric $EccC_5^{\Delta 99}$ yielded parameters consistent with a solution of elongated monomeric particles (Supplementary Fig. 4b, Supplementary Table 3). Furthermore, these data suggest that these elongated particles are flexible (Fig. 3c). To model this conformational flexibility, ensemble optimization methods (EOM) were used in combination with an $EccC_5$ homology model. Three independent model ensembles were tested to probe the origin of flexibility within $EccC_5^{\Delta 99}$ (Supplementary Fig. 4c). Optimized ensembles allowing flexibility only in the DUF domain, while treating the three ATPase domains as a single rigid body (model ensemble 3), accurately described the scattering data (Supplementary Fig. 4c). This analysis indicates that the observed flexibility

largely arising from the DUF domain enables $EccC_5$ to adopt multiple conformations ranging from extended to more compact (Fig. 3d).

Once we identified the flexible cytoplasmic extensions as $EccC_5$, we were able to orientate the EM map in the inner membrane. Next, we aimed to pinpoint the location of the predominantly periplasmic component, $EccB_5$ within the EM map. For this we used the available crystal structure of the soluble domain of $EccB_1$ from *M. tuberculosis*, which shares 32 % identity with *M. xenopi* $EccB_5$[17]. Zhang *et al.* proposed that $EccB_1$ forms a hexamer[18]. Docking of a hexameric model of $EccB_5$ gave a convincing fit to the map placing the majority of $EccB_5$ at the core of the ESX-5 complex (Supplementary Fig. 5a). Alternative docking positions for $EccB_5$ were investigated following segmentation analysis of the EM map[19], revealing a density consistent with the $EccB_5$ homology model (Supplementary Fig. 5b). Both docking approaches position the N-terminal domain of $EccB_5$ at the periplasmic face of the complex forming a collar-like structure around the pore (Fig. 3e).

On the periplasmic side of the complex, additional density was observed that could not be attributed to $EccB_5$ using either model (Fig. 3f). As topology predictions indicate that both $EccC_5$ and $EccD_5$ have minimal mass protruding into the periplasm[15], we hypothesized that the unaccounted density arises from $EccE_5$. This would agree with the previous observation that $EccE_5$ is a peripheral component of the complex, as it is not present in all ESX systems. ESX-1, ESX-2, ESX-3 and ESX-5, but not ESX-4, have EccE homologs that share a relative low sequence identity (19-24% in *M. tuberculosis*) and is particularly sensitive to protease digestion[15]. Bioinformatic predictions of the membrane topology of the different EccE proteins are not consistent, but frequently indicate that the bulk of $EccE_5$ is in the cytosol[20]. To identify the location of $EccE_5$ within the complex, we generated an expression construct in which $EccE_5$ was modified to include a C-terminal hexahistidine-tag. The addition of the tag did not have any effect on complex formation or stability as judged from the appearance of the ESX-$5^{EccE\text{-}His}$ complex visualised by negative staining (Supplementary Fig. 6a-c). However, the functionality of this construct was not tested further. Using gold labelling, we found that $EccE_5$ seems to be located at the periphery of ESX-5 with its C-terminal domain residing in the periplasm (Fig. 3b). This is further supported by the observation that $EccE_5$ forms disulphide

bonds, which are unlikely to occur in the reducing environment of the cytosol (Supplementary Fig. 7).

Because $EccE_5$ is located on the periphery of the complex and this component is not present in all ESX systems, we investigated whether cells lacking the $EccE_5$ component may form a stable sub-complex. Indeed, a complex comprising $EccB_5$, $EccC_5$ and $EccD_5$ could be purified (Supplementary Fig. 8a-c), albeit that expression levels were significantly lower, indicating that complex formation and/or stability is seriously affected. EM analysis of the purified ESX-5 sub-complex (ESX-$5^{\Delta EccE}$) revealed smaller particles than the core ESX-5 complex (Supplementary Fig. 8d-f). Class averages of the ESX-$5^{\Delta EccE}$ similarly show hexameric symmetry around a central pore (Supplementary Fig. 8g). Comparative analysis of the class averages of ESX-5 versus ESX-$5^{\Delta EccE}$ shows the missing density at the periphery of the complex in the latter (Supplementary Fig. 8h-i), which also strengthens the periplasmic localisation of the bulk of $EccE_5$.

Given that most ESX systems are presumably composed of the same set of conserved core components[4], our structural model of the ESX-5 secretion apparatus from *M. xenopi* may serve as a template for mycobacterial T7SSs in general (Fig. 3g). Direct comparison of our T7SS structure with those from other secretion systems reveals that it is fundamentally different both in terms of overall dimensions and symmetry (Fig. 4). Whereas T1SSs, T3SSs, T4SSs and T6SSs span the inner and the outer membrane of Gram-negative bacteria, the overall dimensions of the T7SS molecular assembly confine its role to an inner membrane secretion apparatus[11]. This raises the question as to how the folded substrates are secreted out of the cell after translocation through the T7SS inner membrane complex. Recently, a role of EspC, a specific substrate of the ESX-1 system, in outer membrane translocation of ESX-1 substrates has been proposed, as it localizes to the mycobacterial cell envelope and forms needle-like structures when overexpressed in *Escherichia coli*[21]. As the ESX-5 system lacks an EspC homolog and in contrast to ESX-5, EspC is not essential for growth, we do not foresee that EspC-like needles are involved in ESX-5 mediated transport. Nevertheless, the possibility remains that the outer membrane conduit for the ESX-5 system is formed by another (ESX-5) substrate. A striking feature is the flexibility of $EccC_5$ at the cytosolic face of the ESX-5 complex. Different from the equivalent ATPase in T4SS, which forms two rigid

barrel-like structures[22], $EccC_5$ adopts an elongated, flexible tentacle-like configuration. Moreover, the six-fold symmetry we observed in the ESX-5 complex may provide insights into the overall structural requirements of additional, and as yet unidentified, components involved in subsequent secretion steps. Our data provide novel insights into the mechanism of substrate recognition and translocation through the ESX-5 inner membrane complex. The obtained structure of the ESX-5 complex is a strong basis for future research aimed to identify and characterize interactions required for the complete T7SS assembly such as MycP, required for complex stability[23] as well as the components required to bridge the periplasm and the outer membrane.

METHODS

Molecular biology

Polymerase chain reaction (PCR) was performed using Phusion DNA polymerase (New England Biolabs). For cloning, *Escherichia coli* DH5α was used. The pMV-*eccBC*$_{5mm}$ containing a Strep-tag II coding sequence[23] was modified by PCR amplification to contain the NsiI restriction sites upstream of *eccB*$_5$ and the AflII site between the Strep-tag II sequence and the HindIII restriction site. The resulting plasmid and the previously described pMV plasmid[10] carrying the complete *esx-5* locus of *M. tuberculosis* was further modified by i) replacing the hygromycin resistance cassette by a streptomycin resistance gene; ii) by the introduction of the pAL5000 origin of replication[24] by homologous recombination via a commercial kit (Clontech In-Fusion); iii) and by a deletion of the Hsp60 promoter. The *eccBC*$_{5mx}$ coding region was amplified by PCR from *M. xenopi* RIVM700367 chromosomal DNA to introduce the NsiI and DraI restriction sites. This PCR fragment was cloned into the modified plasmid by replacing the *eccBC*$_{5mm}$ sequence. The MXEN9349-*eccA*$_5$ coding regions were amplified by PCR from *M. xenopi* RIVM700367 chromosomal DNA to introduce the AflII and HindIII restriction sites and ligated in the intermediate pMV plasmid to generate the pMV-ESX-5. The pMV-ESX-5 was further digested with AflII and SnaBI and then religated to result in pMV-ESX-5$^{\Delta subst}$. pMV-ESX-5$^{\Delta EccE}$ resulted from the restriction of pMV-ESX-5 with MluI and HindIII and subsequent religation. Similarly, pMV-ESX-5$^{EccE\text{-}His}$ was derived from pMV-ESX-5 by introducing a hexahistidine-tag sequence downstream of *eccE*$_5$ using the restriction site MluI. The *eccC*$_5$ gene was amplified by PCR to include the cytoplasmic domain (residues 99-1392)

and inserted into the pMyNT vector using SliCE methods[25], generating pMyNT-EccC$_5^{\Delta 99}$.

Secretion analysis

ESX-5 constructs were expressed in *M. smegmatis* mc^{2}155 and functionality of the reconstituted system was analysed by culturing bacteria to mid-logarithmic phase and separating the whole cells from the supernatant by centrifugation and filtration. pSMT3-PE31/PPE18-HA[14], carrying *pe31* and *ppe18*, followed by an HA sequence, from *M. marinum,* was additionally introduced to assess secretion of PPE18 of *M. marinum*. . Subsequently, samples were run on SDS-PAGE for immunoblotting with sera against the HA-tag, EsxN, EccB$_5$ and GroEL2[15]. While anti-EsxN, anti-EccB$_5$ and anti-GroEL2 have been described previously[10], the anti-HA antibody was purchased from Covance/BioLegend (cat. nr. MMS-101P).

Protein overexpression and purification

Equivalent purification protocols were followed for the production of the ESX-5 variants discussed in this study (ESX-5, ESX-5$^{\text{EccE-His}}$ and ESX-5$^{\Delta\text{EccE}}$). *M. smegmatis* mc^{2}155 was transformed with the relevant plasmid and cultured in Middlebrook 7H9 medium (BD Bioscience) supplemented with 0.2% (v/v) glycerol, 10% albumin–dextrose saline (5% (w/v) BSA, 2% (w/v) glucose, 342 mM NaCl), and 0.05% (v/v) Tween-80 at 120 rpm. At an OD_{600} of 1.5, cells were pelleted by centrifugation and suspended in buffer A (20 mM Tris pH 8.0, 300 mM NaCl 10 % (v/v) glycerol) with EDTA-free protease inhibitors (Roche). Cells were lysed by high pressure emulsification and unbroken cells removed by centrifugation (12 min at 10,000*g*). Membrane fractions were isolated from the supernatant by ultracentrifugation (1 h at 98,000*g*). Membranes were resuspended in buffer A and solubilized using 0.25 % DDM for 1 h at 4°C. Insoluble material was removed by ultracentrifugation (30 min at 98,000*g*) and the clarified lysate was incubated with StrepTactin beads (IBA). Following a wash step with buffer A, the complex eluted with buffer A supplemented with 10 mM desthiobiotin (IBA). When exchanging DDM for amphipol A8-35, A8-35 was added in a 5:1 ratio with protein and incubated with BioBeads (Bio-Rad) overnight. The sample was purified further by size exclusion chromatography using a Superose 6 10/300 GL column (GE Healthcare) and analysed by SDS-PAGE and BN-PAGE. Samples were furthermore analysed by immunoblotting using antiserum against EccB$_5$[10].

pMyNT-Ecc$C_5^{\Delta 99}$ was transformed into *M. smegmatis* mc^{2}155 *groEL1ΔC*[26] and cultured in 7H9 medium. Protein expression was induced with 2 % (v/v) acetamide at an OD_{600} of 1. Cell pellets were suspended and lysed as above. Following clarification of the cell lysate by centrifugation (40 min at 30,000*g*), Ecc$C_5^{\Delta 99}$ was purified via immobilized metal affinity chromatography using Ni-NTA agarose (Qiagen). Following the cleavage of the hexahistidine-tag by TEV protease, the protein was concentrated and injected onto a HiLoad Superdex 200 16/600 PG column (GE Healthcare) pre-equilibrated 20 mM Tris pH 8, 200 mM NaCl, 0.25 mM TCEP.

In-gel and in solution digestion

For in-solution processing (replicates 1-3) all reagents were prepared in 50 mM HEPES (pH 8.5). Cysteines were reduced using dithiothreitol (56 °C, 30 minutes, 10 mM). Samples were cooled to 24 °C and alkylated with iodacetamide (room temperature, in the dark, 30 minutes, 10 mM). Subsequently, the samples were prepared for LC-MS/MS using the SP3 protocol[27], digested with trypsin, and the peptides were cleaned up using OASIS HLB µElution Plate (Waters).

For in-gel processing the sample (replicate 4) was separated in a native gel and stained with Coomassie. The bands corresponding to the complex were cut from the gel and subjected to in-gel digestion with trypsin[28]. Subsequently, peptides were extracted from the gel pieces. Samples were sonicated for 15 minutes, centrifuged and the supernatant removed and placed in a clean tube. A solution of 50:50 water: acetonitrile, 1 % formic acid (2 x the volume of the gel pieces) was added and the samples were again sonicated for 15 minutes, centrifuged and the supernatant pooled with the first extract. The pooled supernatants were then dried down with the speed vacuum centrifuge. The samples were dissolved in 10 µL of reconstitution buffer (96:4 water: acetonitrile, 0.1% formic acid and analysed by LC-MS/MS.

Liquid chromatography-mass spectrometry/mass spectrometry

Replicates 1-3 were analysed on a Q Exactive Plus instrument (Thermo Scientific). Peptides were separated using the UltiMate 3000 RSLC nano LC system (Dionex) fitted with a trapping cartridge (µ-Precolumn C18 PepMap 100, 5µm, 300 µm i.d. x 5 mm, 100 Å) and an analytical column (Acclaim PepMap 100 75 µm x 50 cm C_{18}, 3 µm, 100 Å) over a 70-minute gradient and online injected into the mass spectrometer. The resolution of MS scans was set to 70,000.

The filling time was set to a maximum of 100 ms with a target of $3x10^6$ ions. Acquisition was performed in data dependent mode (DDA). The resolution of MS2 scans was set to 17,500, with a fill time of 50 ms and a target of $1x10^5$ ions. HCD normalized collision energy was set to 26.

Replicate 4 was analysed on an Orbitrap Velos instrument (Thermo Scientific). Peptides were separated using the nanoAcquity ultra performance liquid chromatography (UPLC) system (Waters) fitted with a trapping (nanoAcquity Symmetry C18, 5 µm, 180 µm x 20 mm) and an analytical column (nanoAcquity BEH C18, 1.7 µm, 75 µm x 250 mm) over a 35-minute gradient and online injected into the mass spectrometer. Orbitrap Velos mass spectrometer was operated with a data-dependent top 15 method. MS spectra were acquired using a resolution of 30.000 and an ion target of $1. 0 x 10^6$. Collision-induced dissociation (CID) was performed in the linear trap quadrupole (LTQ) using a normalized collision energy of 40%, and an ion target of $3.0 x 10^4$. In order to improve the mass accuracy, a lock mass correction using the ion (m/z 445.12003) was applied.

Data processing

The raw mass spectrometry data were processed with MaxQuant (v1.5.2.8)[29] and searched against a database containing the sequences of proteins of the complex as well as common contaminants. The data were searched with the following modifications: Carbamidomethyl (C) (fixed modification), Acetyl (N-term) and Oxidation (M) (variable modifications). The mass error tolerance for the full scan MS spectra was set to 20 ppm and for the MS/MS spectra to 0.5 Da. A maximum of two missed cleavages was allowed. For protein identification a minimum of two unique peptides with a peptide length of at least seven amino acids and a false discovery rate below 0.01 were required on the peptide and protein level. Quantification was performed using iBAQ values[30] calculated as the sum of the intensities of the identified peptides and divided by the number of observable peptides of a protein. Relative protein concentrations were calculated for each replicate by dividing the iBAQ value of each subunit by the average intensity-based absolute protein quantification (iBAQ) value of the four subunits. The average relative protein concentration was calculated as the average of the relative protein concentration for each subunit over the four replicates.

The standard error of the mean was calculated for each average relative protein concentration (Supplementary Table 3).

Size-exclusion chromatography coupled with multi-angle light scattering

$EccC_5^{\Delta 99}$ was concentrated to 2.5 mg ml^{-1} in 20 mM Tris pH 8.0, 200 mM NaCl, 5 % (v/v) glycerol. Size exclusion chromatography was performed at room temperature using a Superdex 200 10/300 GL column. Separated sample components were analysed in-line with a modular triple detector array (Viscotek TDA 305, Malvern Instruments Ltd., Malvern) to determine right-angle light scattering (RALS), refractive index (RI), and UV-VIS (UV). Data were processed using the Omnisec software (Malvern). The molecular weight of the species eluting from the SEC column (MW_{RALS}) was assessed using correlated concentration measurements derived from baseline-corrected RI in combination with baseline-corrected RALS intensities calibrated against a bovine serum albumin narrow standard (monomeric peak).

Homology modelling and docking

Homology models of the *M. xenopi* $EccB_5$ and $EccC_5$ were generated using MODELER v9.17[17] using the default settings. The homology model of C-terminal part of *M. xenopi* $EccC_5$ (residues 429-1386) was divided into three ATPase domains A1 (residues 431-712), A2 (residues 826-1123) and A3 (residues 1135-1389) based on the *T. curvata* EccC structure (PDB code 4NH0). The hexamer of *the M. xenopi* $EccB_5$ was generated using SymmDock[31] by applying a six-fold symmetry. Models were docked into the EM map or segmented EM volumes using Chimera[32].

Small-angle X-ray scattering

Synchrotron SAXS data were collected at EMBL P12 beamline (DESY, Hamburg) and recorded at 10°C using a PILATUS 1M pixel detector (DECTRIS) at a sample–detector distance of 2.7 m and a wavelength of 1.24 Å (Supplementary Data Table 1). Purified $EccC_5^{\Delta 99}$ was measured in a concentration range from 0.1 to 7 mg ml^{-1} from fractions corresponding to the monomeric species. For the calculation of the forward scattering $I(0)$ and the radius of gyration (R_g), the Guinier approximation implemented in PRIMUS[33] was used. The pair-distance distribution function $P(r)$ was evaluated with GNOM[34], and the

maximum particle dimension (D_{max}) as well as R_g were estimated consecutively. Flexibility was assessed with the Ensemble Optimization Method (EOM) program[35]. The model was divided into rigid bodies corresponding to the ATPase domains A1, A2 and A3. EOM was performed using three independent runs describing the potential flexibility of $EccC_5^{\Delta 99}$. During run 1, A1, A2 and A3 were treated as single independent rigid bodies; run 2 combined A2 and A3 into one rigid body; run 3 fixed A1, A2 and A3 as one rigid body. In all runs the DUF domain was treated as a flexible entity. The fit of the best ensemble of conformations, selected by EOM, was guided by the χ^2 criterion:

$$\chi^2 \frac{1}{K-1} = \sum_{j=1}^{K} \left[\frac{cI(s_j) - Iexp(s_j)}{\sigma(s_j)} \right]^2$$

where $Iexp(s_j)$ is the experimental scattering, K is the number of experimental points, $\sigma(s_j)$ are standard deviations and c is a scaling factor[36].

ATPase assays

ATPase activity was measured by quantifying the hydrolysis of [γ^{32}P]-ATP using thin layer chromatography (TLC). Assays were performed using purified ESX-5 complex in amphipols (100 nM) in 20 mM Tris pH 8.0, 200 mM NaCl, 5 % (v/v) glycerol, 2 mM $MgCl_2$, 1 mM DDT. The reaction was initiated by adding 100 nM [γ^{32}P]-ATP (3000 Ci/mmol) and incubated at 37°C for 1 h to a total reaction volume of 10 µl. The control reaction was conducted in parallel without the addition of ESX-5 complex. The reaction was stopped by the addition of 20 mM EDTA and the reaction products were isolated by phenol:chloroform extraction and loaded onto a PEI-cellulose TLC plate. The TLC plate was run in 0.5 M LiCl, 1 M formic acid and plates were allowed to dry before imaging. Signals were detected following exposure to an image plate (Fujifilm) and analysed using a Phosphor-Imager (Fujifilm). Assays were performed in triplicate.

Disulphide bond analysis

Isolated *M. smegmatis* membranes containing ESX-5 membrane complexes from either *M. xenopi* or *M. tuberculosis* were incubated in buffer A (20 mM Tris pH 8.0, 300 mM NaCl 10 % (v/v) glycerol) including 0 or 10 mM DTT for 10 min at 37°C. Samples were subsequently analyzed by SDS-PAGE and immunoblotting using SDS sample buffer devoid of reducing

agent. The used antisera against $EccB_5$ and $EccE_5$ have been described previously[10].

Electron microscopy and image processing

EM samples were prepared by applying a protein solution (3 µL) on previously glow-discharged carbon-coated copper grids and negatively stained with 2% uranyl acetate. For gold labelling sample screening of the gold-labelled ESX-$5^{EccE\text{-}His}$ complex and the ESX-$5^{\Delta EccE}$, grids were imaged with a Tecnai Morgani electron microscope operating at 80 kV at a defocus of ~2 µm and at a calibrated magnification of 24,194x resulting in a pixel size of 3.72 Å on the specimen. For the single-particle analysis grids were imaged under low-dose conditions with a Tecnai T20 electron microscope operating at 200 kV. Data were collected on a FEI BM EAGLE camera, with a step size of 15 µm, a defocus range of 1.4-5.0 µm and at a nominal magnification of 50,000x, resulting in a pixel size of 2.18 Å on the specimen. For the ESX-5 complex, a total of 65 micrographs were collected and for each of these, a second micrograph at 50° tilt was acquired. A total of 1,418 particles were boxed out manually using 'e2boxer.py' in the EMAN2 software package[37] using a box size of 320 pixels. For the the ESX-$5^{\Delta EccE}$ sample, fifty micrographs were collected using the same conditions and a total of 684 particles were selected using a box size of 180 pixels.

A first step of 2D reference-free class averaging showed the particles to have no preferred orientation on the carbon layer and revealed hexameric classes. To test for symmetry, rotational cross-correlations were calculated from reference-free aligned classes over 360 degrees in 1 degree steps (IMAGIC-5, Images Science Software GmbH). The initial 3D model was obtained using the EMAN2 initial-model-generation program e2initialmodel.py without applying symmetry. This process yielded a six-fold shaped map. A series of independent 3D refinement rounds were performed using different starting 3D references obtained in the previous step of initial model generation and these retained the six-fold symmetry topology and overall EM map. Based on these findings, in the final refinement steps a six-fold symmetry constraint was applied. During the 3D refinement rounds the best self-consistent subset of particles was obtained by keeping the best 70% of the particles from each reference-based class average based on mutual similarity among T7SS particles. The final resolution between two independent reconstructions (gold standard) was estimated at FSC

0.143 within the 'e2refine_easy.py' tool of EMAN2 and was of ~13Å. The final map is CTF amplitude corrected and Wiener filtered based on the FSC curve. Map segmentation was performed using Segger software[19]. Handedness of the map was analysed by tilt-pair validation using the 'e2tiltvalidate.py' program in EMAN2[38]. For the ESX-5$^{\Delta EccE}$ sample, 50 micrographs were collected with a Tecnai Morgani electron microscope operating at 80 kV at a defocus range ~1-4 µm and at a calibrated magnification of 24,194x resulting in a pixel size of 3.72 Å on the specimen. A total of 684 particles were boxed out manually using 'e2boxer.py' with a box size of 180 pixels and subjected to 2D reference-free class averaging.

Gold labelling

Complex components were labelled using 5 nm streptavidin (Nanoncs) for the ESX-5 complex against the EccC$_5$-Strep or Ni-NTA (Nanoprobes) nanogold beads for the ESX-5$^{EccE\text{-}His}$ complex. Purified complexes were incubated for 1 h with varying concentrations of gold beads diluted in SEC buffer, applied directly onto a glow-discharged carbon-coated grid and stained as described above.

Data availability

The electron microscopy data of the ESX-5 membrane complex that support the findings of this study have been deposited in the Electron Microscopy Data Bank under the following accession code EMD-3596. SAXS data are available in the SASBDB (www.sasbdb.org) with accession code SASDBE7.

Correspondence and requests for materials should be addressed to T.M, A.H.A.P and E.N.G.H.

Acknowledgements

This work was generously supported through the Behörde für Wissenschaft, Forschung und Gleichstellung of the city of Hamburg (to T.C.M.) and funded by a VIDI grant from the Netherlands Organization of Scientific Research (NWO) (to E.N.G.H. and C.M.B.). K.S.H.B. was funded by a postdoctoral fellowship from the EMBL Interdisciplinary Postdoc Program (EIPOD) under Marie Curie COFUND actions. We would like to thank the EMBL P12 beamline

and synchrotron facilities at PETRA III (DESY, Hamburg). We are grateful to Mandy Jeske for assistance with the ATPase assays (EMBL, Heidelberg), Shashi Bhushan (Nanyang Technological University, Singapore) for initial EM analysis of ESX-5 particles and Frank Schluenzen for DESY high-performance cluster support. Samples were recorded at the EM facility of the Vienna Biocenter Core Facilities GmbH (VBCF), Austria. We would like to thank Steven J. Ludtke (Baylor College of Medicine) for helpful comments on resolution estimation. The authors are grateful to Arjen Jakobi and Carsten Sachse (EMBL, Heidelberg) for providing valuable feedback on the manuscript.

Author contributions

A.H.A.P., E.N.G.H., W.B., T.C.M. and M.W. supervised and supported the project. K.S.H.B., A.H.A.P., E.N.G.H., C.M.B. and L.C. designed experiments. E.N.G.H., R.U., C.M.B. and K.S.H.B. generated the constructs. K.S.H.B, C.M.B. and E.N.G.H. purified the ESX-5 complexes for EM analysis. K.S.H.B. performed gold labelling and ATPase assays. C.M.B. performed secretion analysis and disulphide bond characterization. L.C. and J.M. performed electron microscopy. L.C. collected and processed the negative stain electron microscopy data together with W.L. K.S.H.B and H.M. collected and analysed SAXS data with the support of D.S. M.M.S. and M.R. conducted the mass-spectrometry analysis. K.S.H.B., A.H.A.P., E.N.G.H., C.M.B, W.B., T.C.M., L.C. and M.W. wrote the paper.

Figure 1

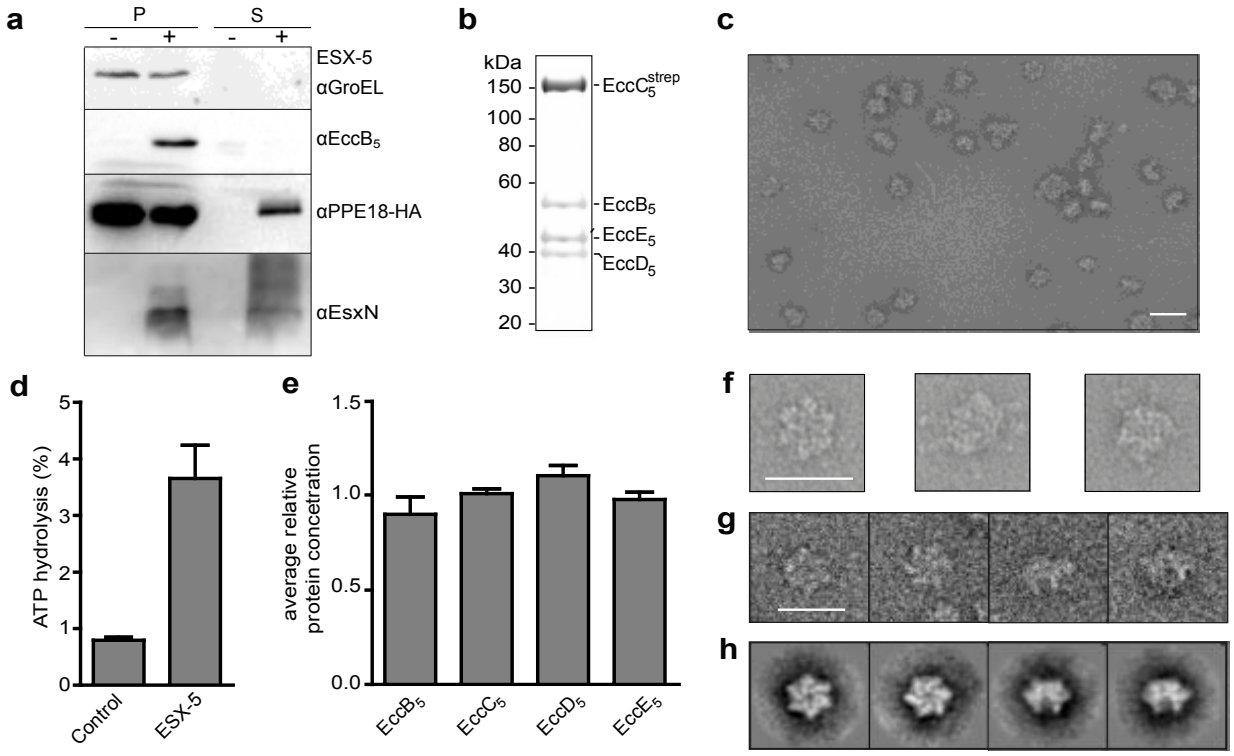

a
P
S
- + - +
ESX-5
αGroEL
αEccB5
αPPE18-HA
αEsxN
b
kDa
150
100
80
60
40
30
20
EccC5strep
EccB5
EccE5
EccD5
c
d
ATP hydrolysis (%)
0
1
2
3
4
5
Control
ESX-5
e
average relative protein concetration
0.0
0.5
1.0
1.5
EccB5
EccC5
EccD5
EccE5
f
g
h

Figure 2

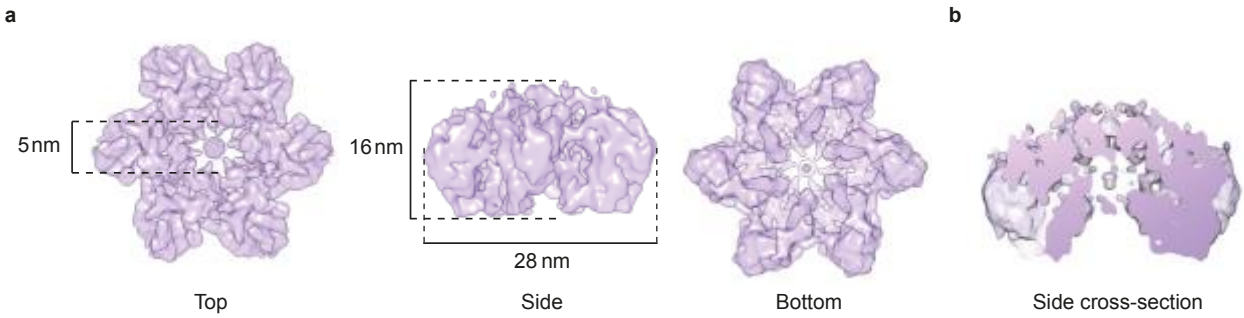


Figure 3

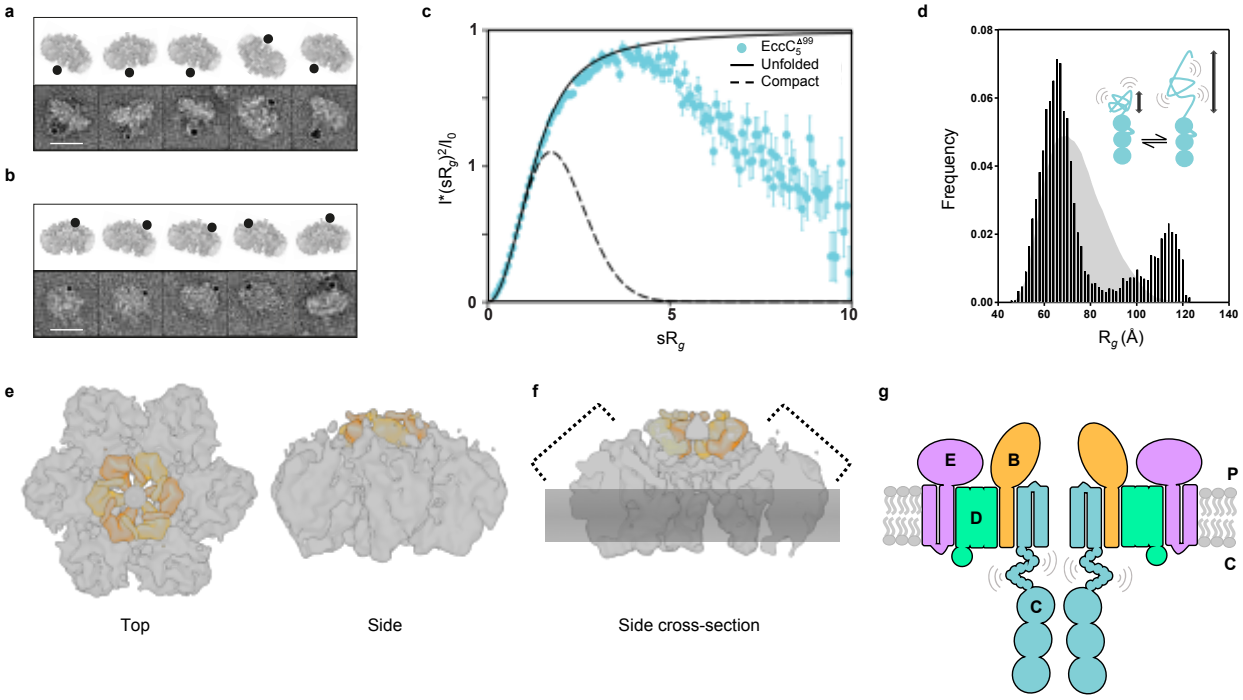
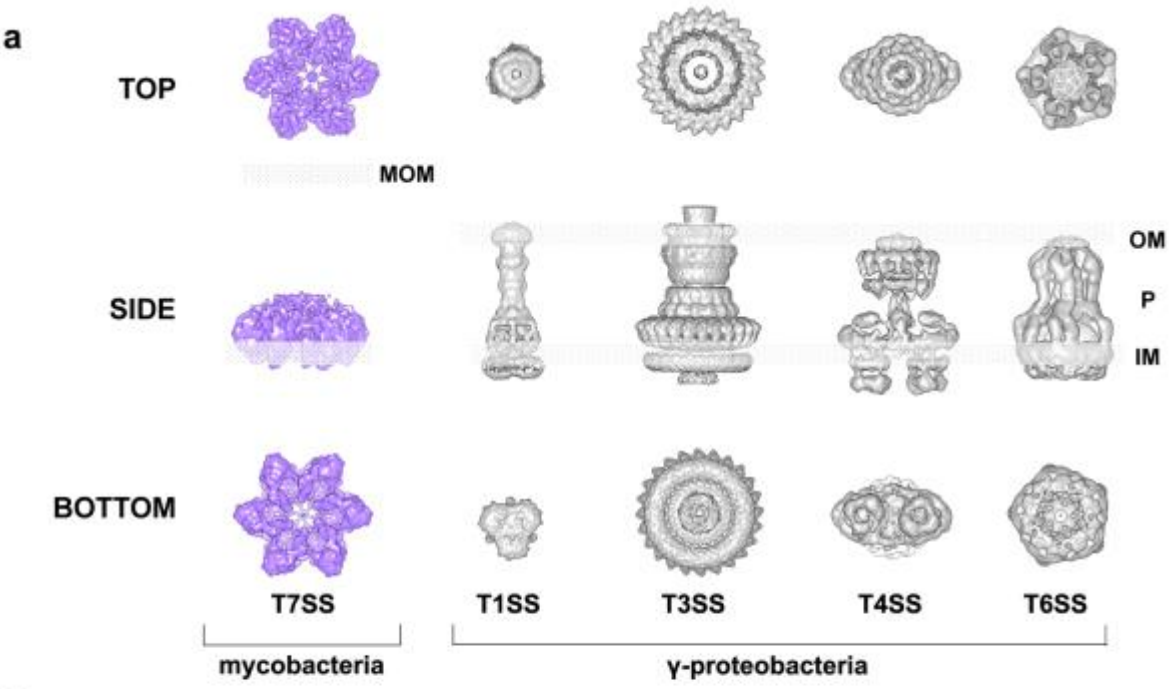

a
b
c
1
1
0
0
5
10
sR_g
I*(sR_g)^2/I_0
EccC_5^Δ99
Unfolded
Compact
d
0.08
0.06
0.04
0.02
0.00
40
60
80
100
120
140
Frequency
R_g (Å)
e
f
g
E
B
D
C
P
C
Top
Side
Side cross-section

Figure 4



b

	T7SS	T1SS	T3SS	T4SS	T6SS
Mass (MDa)	1.8	0.8	3.5	1	1.7
Subunits	4	4	12	12	3
Symmetry	6-fold	3-fold	3-fold	2-fold	5-fold
Methods	NS	Cryo	Cryo	NS	NS
Resolution (Å)	13	16	10	18	12

Figure Legends

Figure 1 | Functional and structural characterization of the *M. xenopi* ESX-5 complex. a, Immunoblot showing substrate secretion is dependent on the presence of reconstituted ESX-5 system (representative of triplicate biological replicates). Anti-$EccB_5$ antibodies were used to confirm ESX-5 complex expression in the cellular fraction; anti-EsxN and anti-HA (the latter to detect PPE18 containing a C-terminal HA-tag) were used to assess secretion; anti-GroEL2 was used as lysis control; (P) cell pellet, (S) culture supernatant. **b,** Coomassie-stained SDS-PAGE gel of purified ESX-5 complex following affinity purification and size exclusion chromatography showing the four conserved components $EccB_5$, $EccC_5^{strep}$, $EccD_5$ and $EccE_5$. Representative gel from at least three independent experiments. **c,** Representative electron micrograph showing ESX-5 particles present in several orientations. Scale bar is 50 nm. **d,** The ATPase activity of the purified ESX-5 complex was measured in three independent experiments. Errors bars show the standard deviation of the mean. T-test analysis indicates a significant increase in ATPase activity compared to the control (p-value ≤ 0.01). **e,** Quantification of relative abundance of the four core subunits of the purified ESX-5 complex. Errors bars show the standard error of the mean derived from four biological replicates. **f,** Representative enlarged images of three individual ESX-5 particles. The first image shows a top view, while the adjacent views highlight a side orientation where the extended filaments are visible. Scale bar is 50 nm. **g**, Four representative views of the ESX-5 complex corresponding to the classes shown in h. Scale bar is 50 nm. **h,** Four typical class averages containing ~30 particles each. The six-fold symmetry is clearly visible for the first two top views from the left.

Figure 2 | 3D reconstruction of the ESX-5 complex. a, Orthogonal views of the reconstructed 3D map (C6) show a 28nm x 16nm structure with a central channel of ~5 nm in diameter. The map is displayed in top, side and bottom view. **b,** Cross-section through the centre of the complex illustrating the inner membrane pore.

Figure 3 | Topology and localization of the individual components of the ESX-5 complex. a, Gold labelling of the C-terminal cytosolic domain of $EccC_5$. The upper panel depicts the orientation of the EM map relative to the particle in the lower panel. Particles shown are representative of labelling experiments performed in triplicate. Scale bar is 20 nm. **b,** Gold

labelling of the C-terminal $EccE_5$ domain. Particles shown are representative of labelling experiments performed in triplicate. Scale bar is 20 nm. **c,** Small angle x-ray scattering analysis of $EccC_5^{\Delta 99}$. Representative normalized Kratky plot showing the scattering intensities of $EccC_5^{\Delta 99}$ in solution (sR_g, versus $sR_g^2*I(q)/I_0$) (cyan). Error bars indicate the standard deviation of the scattering intensity. For reference, typical profiles for a folded, globular particle (dotted line) and an unfolded particle (black line) are included. **d,** Flexibility analysis using EOM shows the R_g distribution for a random pool (shaded grey) of 10,000 structures, where the bimodal population of selected ensembles (black bars) best fit the scattering data. The results suggest that the DUF domain occupies a continuum of populations in solution ranging from compact to rather extended conformations. **e,** Segmentation analysis of the EM volume highlighting the position of $EccB_5$ (orange and yellow, alternating) according to the consensus of both docking approaches within the complex (white) in side and top view**. f**, A side view cross-section of the map shows density on the periplasmic side of the complex that cannot be attributed to $EccB_5$, indicating the possible position of $EccE_5$ (indicated by bracket). The location of the inner membrane is depicted in shadowed grey. **g**, Model of the ESX-5 membrane complex; the periplasmic (P) localization of $EccE_5$ and the flexibility of the DUF region of $EccC_5$ in the cytosol (C) are indicated. The colour codes apply for other panels.

Figure 4 | Unique features of the mycobacterial T7SS in comparison to other bacterial secretion systems. **a,** From left to right: T7SS from mycobacteria (this study, EMD-3596); T1SS from *E. coli*, AcrAB-TolC multi-drug efflux pump (EMD-5915)[39]; T3SS from *Salmonella thyphimurium* (EMD-1875)[40]; T4SS from *E. coli* (EMD-2567)[41]; T6SS from *E. coli* (EMD-2927)[42]. Top, side and bottom views are shown for comparison, demonstrating overall symmetry of the secretion system. The positions on the inner membrane (IM), outer membrane (OM), mycobacterial outer membrane (MOM) and periplasm (P) are indicated. **b,** Summary of the characteristics of each of the displayed secretion systems including the molecular mass (MDa), the number of subunits and overall symmetry. The method used to determine the different models and the achieved resolution is indicated: NS, negative stain electron microscopy; Cryo, cryo-electron microscopy.

References

1. Hoffmann, C., Leis, A., Niederweis, M., Plitzko, J. M. & Engelhardt, H. Disclosure of the mycobacterial outer membrane: cryo-electron tomography and vitreous sections reveal the lipid bilayer structure. *Proc. Natl. Acad. Sci. U. S. A.* **105,** 3963–7 (2008).
2. Houben, E. N. G., Korotkov, K. V. & Bitter, W. Take five — Type VII secretion systems of Mycobacteria. *Biochim. Biophys. Acta - Mol. Cell Res.* **1843,** 1707–1716 (2014).
3. Bitter, W. & Kuijl, C. Targeting bacterial virulence: The coming out of type VII secretion inhibitors. *Cell Host and Microbe* **16,** 430–432 (2014).
4. Gröschel, M. I., Sayes, F., Simeone, R., Majlessi, L. & Brosch, R. ESX secretion systems: mycobacterial evolution to counter host immunity. *Nat. Rev. Microbiol.* (2016). doi:10.1038/nrmicro.2016.131
5. Pym, A. S., Brodin, P., Brosch, R., Huerre, M. & Cole, S. T. Loss of RD1 contributed to the attenuation of the live tuberculosis vaccines Mycobacterium bovis BCG and Mycobacterium microti. *Mol. Microbiol.* **46,** 709–717 (2002).
6. Stanley, S. A., Raghavan, S., Hwang, W. W. & Cox, J. S. Acute infection and macrophage subversion by *Mycobacterium tuberculosis* require a specialized secretion system. *Proc. Natl. Acad. Sci.* **100,** 13001–13006 (2003).
7. Hsu, T. *et al.* The primary mechanism of attenuation of bacillus Calmette-Guerin is a loss of secreted lytic function required for invasion of lung interstitial tissue. *Proc. Natl. Acad. Sci. U. S. A.* **100,** 12420–5 (2003).
8. Ates, L. S. *et al.* Essential Role of the ESX-5 Secretion System in Outer Membrane Permeability of Pathogenic Mycobacteria. *PLOS Genet.* **11,** e1005190 (2015).
9. Fishbein, S., van Wyk, N., Warren, R. M. & Sampson, S. L. Phylogeny to function: PE/PPE protein evolution and impact on Mycobacterium tuberculosis pathogenicity. *Mol. Microbiol.* **96,** 901–916 (2015).
10. Houben, E. N. G. *et al.* Composition of the type VII secretion system membrane complex. *Mol. Microbiol.* **86,** 472–484 (2012).
11. Sani, M. *et al.* Direct visualization by cryo-EM of the mycobacterial capsular layer: a labile structure containing ESX-1-secreted proteins. *PLoS Pathog.* **6,** e1000794 (2010).
12. Rosenberg, O. S. *et al.* Substrates Control Multimerization and Activation of the Multi-Domain ATPase Motor of Type VII Secretion. *Cell* **161,** 501–512 (2015).
13. Daleke, M. H. M., Ummels, R., Bawono, P., Heringa, J. & Vandenbroucke-grauls, C. M. J.

E. General secretion signal for the mycobacterial type VII secretion pathway. *Proc. …* **109,** 11342–11347 (2012).

14. Korotkova, N. *et al.* Structure of the Mycobacterium tuberculosis type VII secretion system chaperone EspG5 in complex with PE25-PPE41 dimer. *Mol. Microbiol.* **94,** 367–382 (2014).
15. Houben, E. N. G. *et al.* Composition of the type VII secretion system membrane complex. *Mol. Microbiol.* **86,** 472–484 (2012).
16. Poulsen, C., Panjikar, S., Holton, S. J., Wilmanns, M. & Song, Y.-H. WXG100 protein superfamily consists of three subfamilies and exhibits an α-helical C-terminal conserved residue pattern. *PLoS One* **9,** e89313 (2014).
17. Webb, B. & Sali, A. Comparative protein structure modeling using MODELLER. *Curr. Protoc. Bioinforma.* **2014,** 5.6.1-5.6.32 (2014).
18. Zhang, X.-L. *et al.* Core component EccB1 of the Mycobacterium tuberculosis type VII secretion system is a periplasmic ATPase. *FASEB J.* **29,** 4804–14 (2015).
19. Pintilie, G. D., Zhang, J., Goddard, T. D., Chiu, W. & Gossard, D. C. Quantitative analysis of cryo-EM density map segmentation by watershed and scale-space filtering, and fitting of structures by alignment to regions. *J. Struct. Biol.* **170,** 427–438 (2010).
20. Bitter, W. *et al.* Systematic Genetic Nomenclature for Type VII Secretion Systems. *PLoS Pathog.* **5,** e1000507 (2009).
21. Lou, Y., Rybniker, J., Sala, C. & Cole, S. T. EspC forms a filamentous structure in the cell envelope of Mycobacterium tuberculosis and impacts ESX-1 secretion. *Mol. Microbiol.* **103,** 1–39 (2016).
22. Peña, A. *et al.* The hexameric structure of a conjugative VirB4 protein ATPase provides new insights for a functional and phylogenetic relationship with DNA translocases. *J. Biol. Chem.* **287,** 39925–39932 (2012).
23. van Winden, V. J. C. *et al.* Mycosins Are Required for the Stabilization of the ESX-1 and ESX-5 Type VII Secretion Membrane Complexes. *MBio* **7,** e01471-16 (2016).
24. Labidi, A., Mardis, E., Roe, B. A. & Wallace, R. J. Cloning and DNA sequence of the Mycobacterium fortuitum var fortuitum plasmid pAL5000. *Plasmid* **27,** 130–140 (1992).
25. Zhang, Y., Werling, U. & Edelmann, W. SLiCE: A novel bacterial cell extract-based DNA cloning method. *Nucleic Acids Res.* **40,** e55 (2012).

26. Noens, E. E. *et al.* Improved mycobacterial protein production using a *Mycobacterium smegmatis groEL1ΔC* expression strain. *BMC Microbiol.* **11,** 27 (2011).
27. Hughes, C. S. *et al.* Ultrasensitive proteome analysis using paramagnetic bead technology. *Mol. Syst. Biol.* **10,** 757 (2014).
28. Savitski, M. M. *et al.* Tracking cancer drugs in living cells by thermal profiling of the proteome. *Science (80-.).* **346,** (2014).
29. Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **26,** 1367–72 (2008).
30. Schwanhäusser, B. *et al.* Global quantification of mammalian gene expression control. *Nature* **473,** 337–342 (2011).
31. Schneidman-Duhovny, D., Inbar, Y., Nussinov, R. & Wolfson, H. J. PatchDock and SymmDock: servers for rigid and symmetric docking. *Nucleic Acids Res.* **33,** W363-7 (2005).
32. Pettersen, E. F. *et al.* UCSF Chimera--a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25,** 1605–12 (2004).
33. Konarev, P. V., Volkov, V. V., Sokolova, A. V., Koch, M. H. J. & Svergun, D. I. PRIMUS: A Windows PC-based system for small-angle scattering data analysis. *J. Appl. Crystallogr.* **36,** 1277–1282 (2003).
34. Svergun, D. I. Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J. Appl. Crystallogr.* **25,** 495–503 (1992).
35. Tria, G., Mertens, H. D. T., Kachala, M. & Svergun, D. I. Advanced ensemble modelling of flexible macromolecules using X-ray solution scattering. *IUCrJ* **2,** 207–217 (2015).
36. Bernado, P., Mylonas, E., Petoukhov, M. V, Blackledge, M. & Svergun, D. I. Structural characterization of flexible proteins using Small-angle X-ray scattering. *J. Am. Chem. Soc.* **129,** 5656–5664 (2007).
37. Tang, G. *et al.* EMAN2: An extensible image processing suite for electron microscopy. *J. Struct. Biol.* **157,** 38–46 (2007).
38. Rosenthal, P. B. & Henderson, R. Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. *J. Mol. Biol.* **333,** 721–745 (2003).
39. Du, D. *et al.* Structure of the AcrAB-TolC multidrug efflux pump. *Nature* **509,** 512–515

(2014).

40. Schraidt, O. & Marlovits, T. C. Three-dimensional model of Salmonella's needle complex at subnanometer resolution. *Science* **331,** 1192–5 (2011).

41. Low, H. H. *et al.* Structure of a type IV secretion system. *Nature* **508,** 550–553 (2014).

42. Durand, E. *et al.* Biogenesis and structure of a type VI secretion membrane core complex. *Nature* **523,** 555–560 (2015).

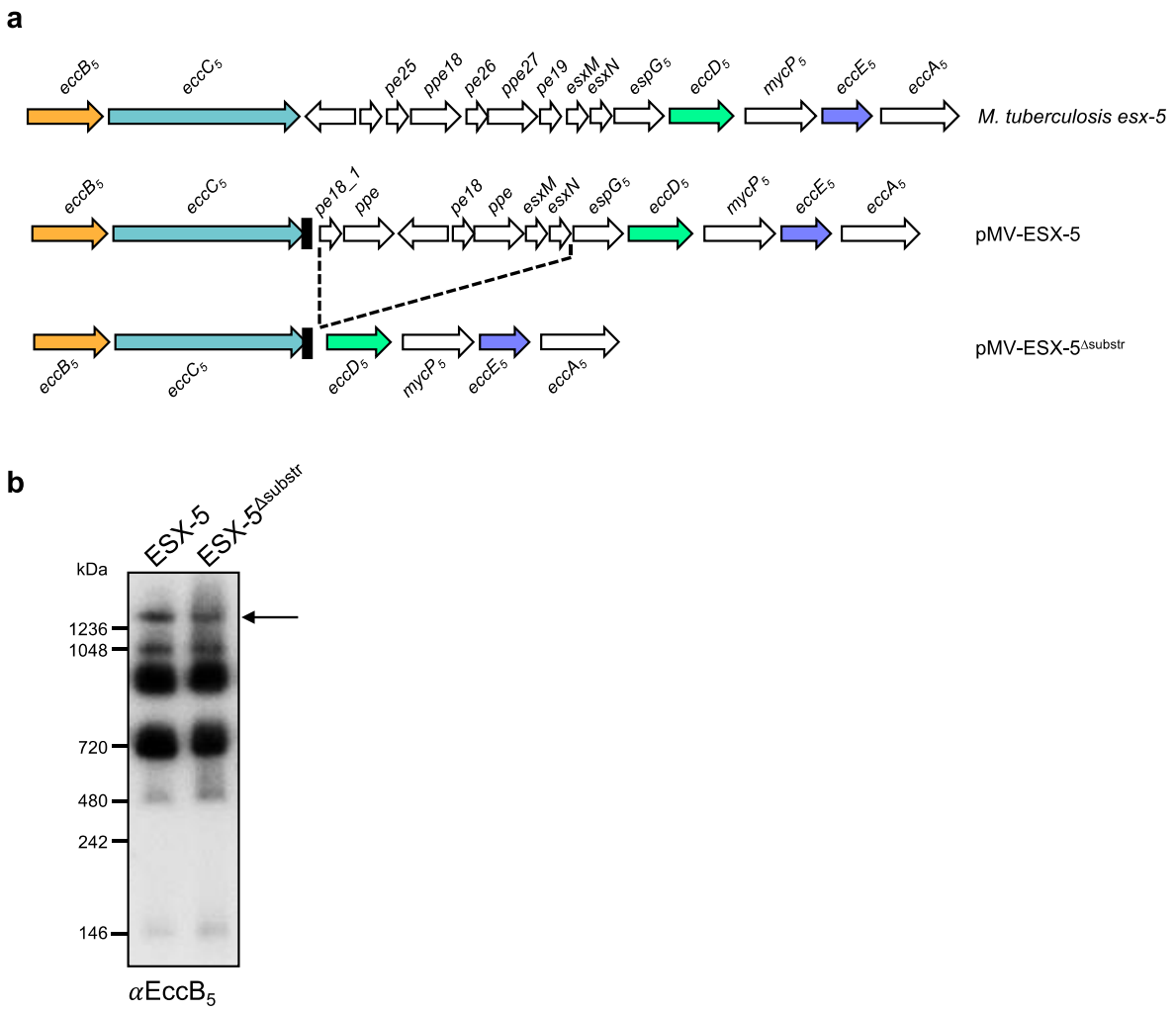
Structure of the mycobacterial ESX-5 type VII secretion system membrane complex by single particle analysis

Katherine S.H. Beckham[1#], Luciano Ciccarelli[2,3,4#], Catalin M. Bunduc[5#], Haydyn D.T. Mertens[1], Roy Ummels[6], Wolfgang Lugmayr[2,3,4], Julia Mayr[2,3,4], Mandy Rettel[7], Mikhail M. Savitski[7], Dmitri I. Svergun[1], Wilbert Bitter[5,6], Matthias Wilmanns[1,4,8], Thomas C. Marlovits[2,3,4,8,9*], Annabel H.A. Parret[1*], Edith N.G. Houben[5*]

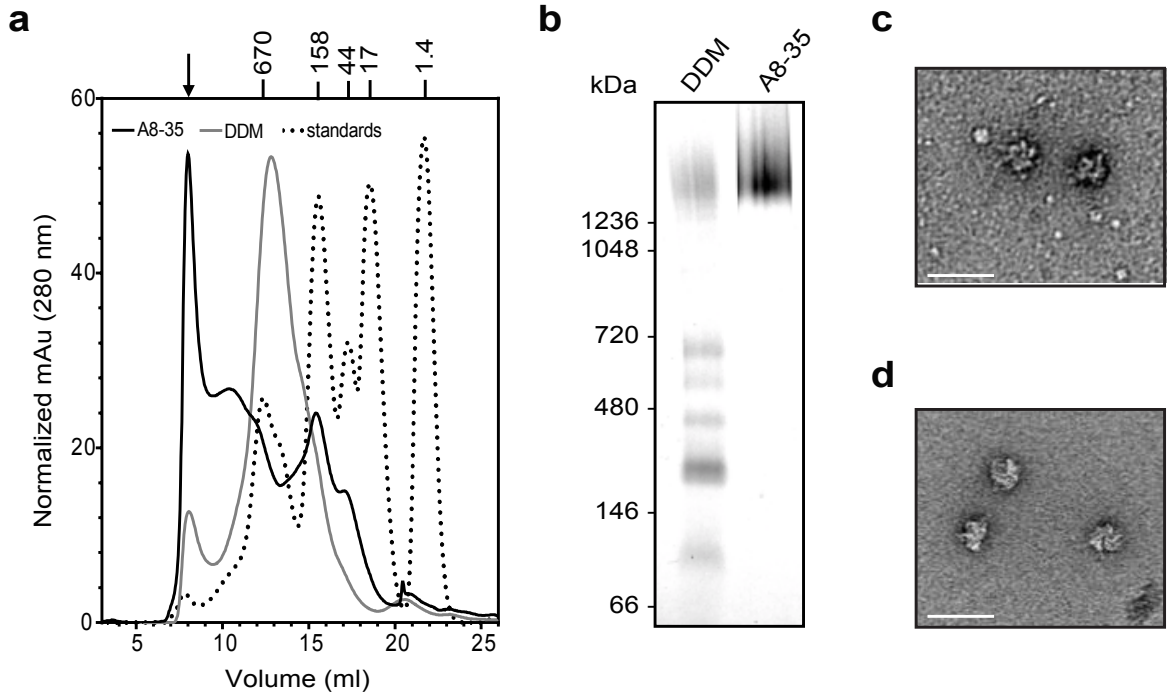
 # equal contribution
** Correspondence: Thomas.Marlovits@imba.oeaw.ac.at, parret@embl-hamburg.de, e.n.g.houben@vu.nl*

[1]European Molecular Biology Laboratory, Hamburg Unit, Notkestraße 85, 22607, Hamburg, Germany
[2]Research Institute of Molecular Pathology, Dr. Bohr-Gasse 7, 1030 Vienna, Austria
[3]Institute of Molecular Biotechnology, Austrian Academy of Sciences, Dr. Bohr-Gasse 3, 1030 Vienna, Austria
[4]University Medical Centre Hamburg-Eppendorf, Martinistraße 52, 20246 Hamburg, Germany
[5]Vrije Universiteit Amsterdam, De Boelelaan 1108, 1081 HZ Amsterdam, The Netherlands
[6]VU University Medical Center, De Boelelaan 1108, 1081 HZ Amsterdam, The Netherlands
[7]European Molecular Biology Laboratory, Meyerhofstraße 1, 69117 Heidelberg, Germany
[8]Centre for Structural Systems Biology, Notkestraße 85, 22607 Hamburg, Germany
[9]Deutsches Elektronen-Synchrotron, Notkestraße 85, 22607 Hamburg, Germany

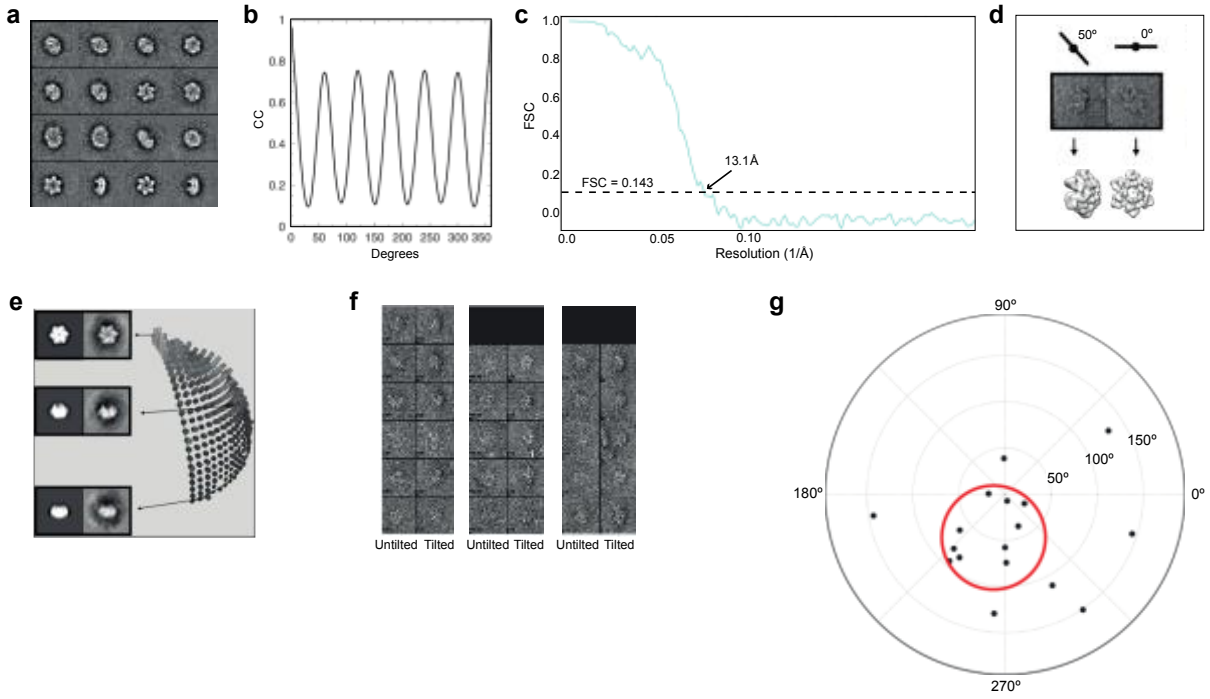
This file contains Supplementary Figure 1-9 and Supplementary Tables 1-3.



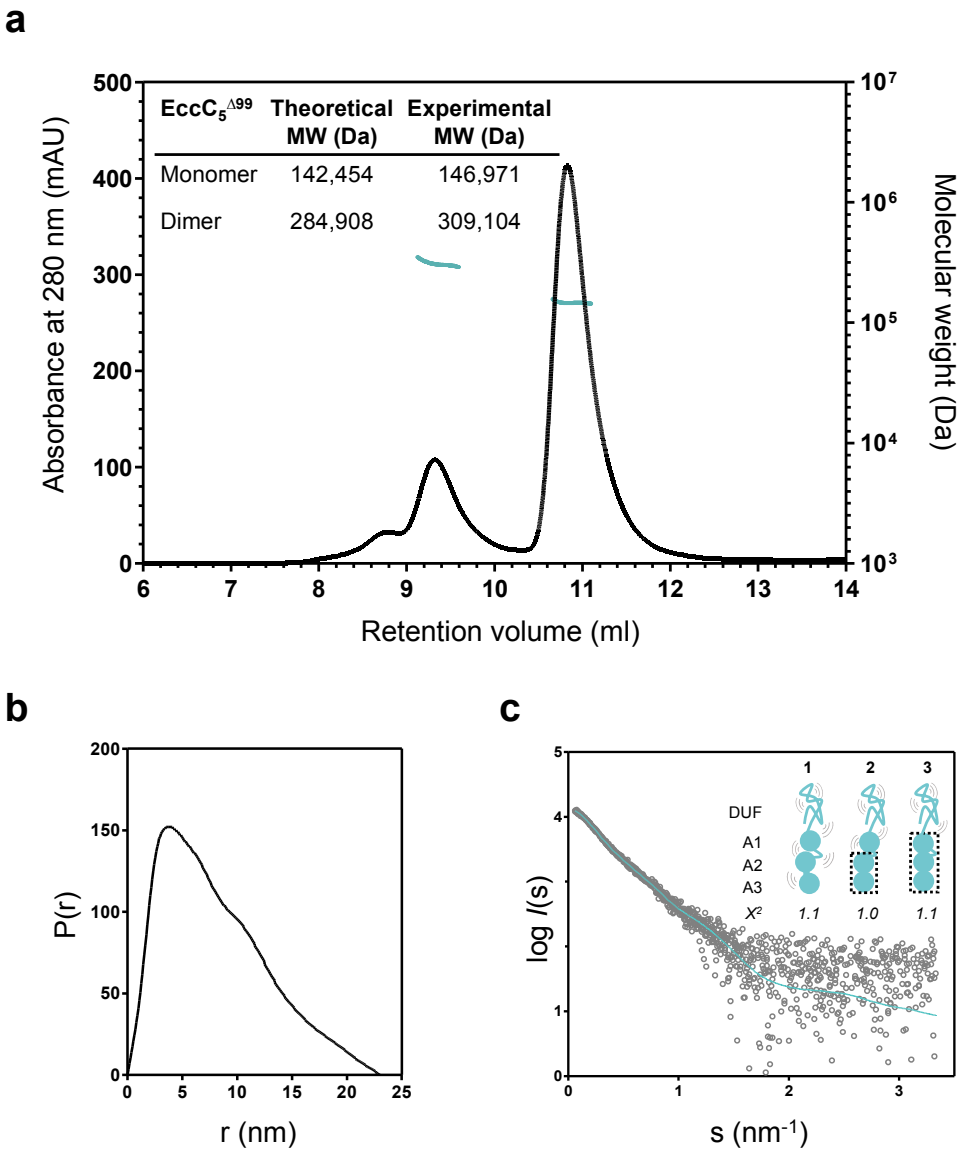
Supplementary Figure 1. Conserved gene architecture and expression of *esx-5* loci. a, Gene organization of the *esx-5* locus in *M. tuberculosis* H37Rv. The homologous *esx-5* locus from *M. xenopi* was expressed in *M. smegmatis* from the pMV-ESX-5 construct. Genes have been coloured according the scheme shown in Fig. 3g. The Strep-tag II on *eccC*$_5$ is shown as a black box. To test the effect of secreted effectors on complex formation, all genes encoding for substrates (two *pe*, two *ppe* and two *esx* genes) were deleted to generate pMV-ESX-5$^{\Delta substr}$ as indicated by the dotted lines. **b**, Immunoblot of detergent-solubilized cell envelope fractions of *M. smegmatis* expressing pMV-ESX-5 and pMV-ESX-5$^{\Delta substr}$ analysed by blue native PAGE. Intact ESX-5 complex is indicated by a black arrow. Representative immunoblot using anti-EccB$_5$ antibodies from at least three independent experiments.



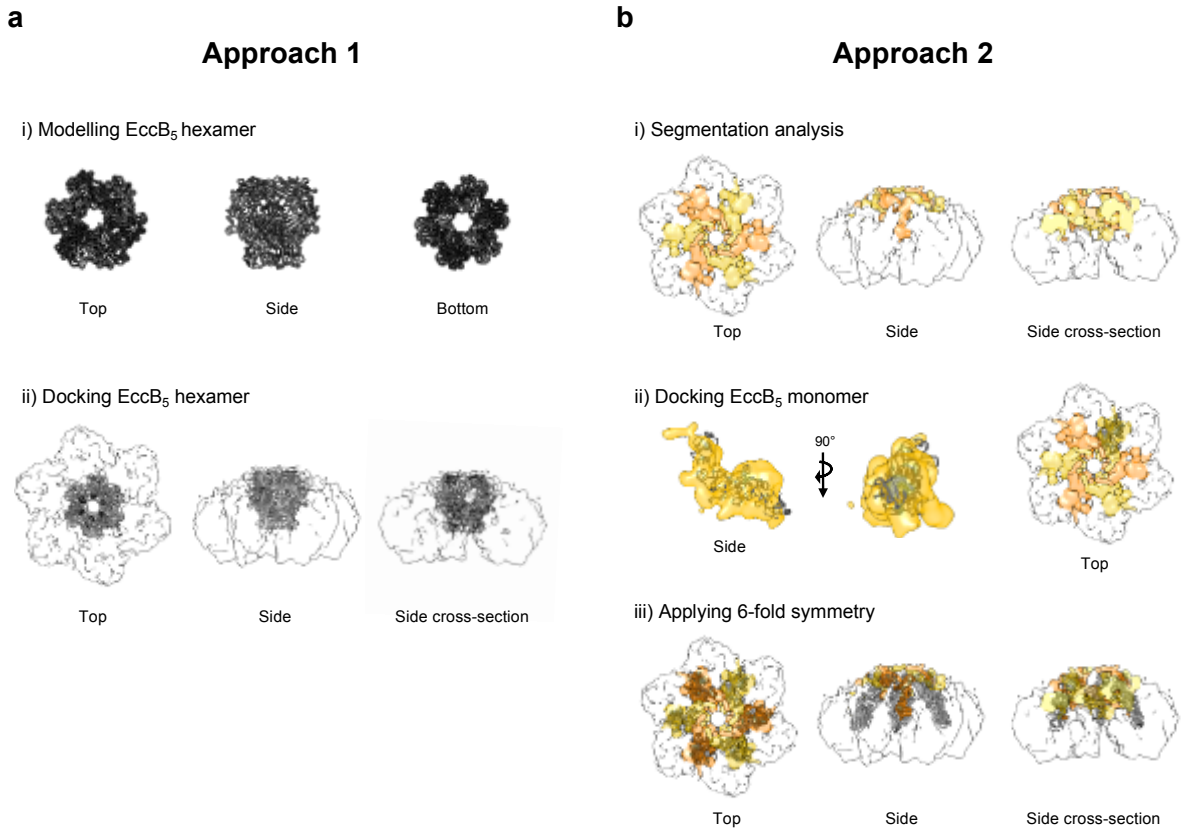
Supplementary Figure 2. Improved stability of the ESX-5 complex in amphipols. a, Analytical size exclusion chromatography analysis of the ESX-5 complex in DDM (grey line) versus amphipol A8-35 (black line). The arrow indicates the elution fraction used for blue native PAGE and EM analysis. The molecular weight of a standard set of protein markers is indicated on top of the corresponding peak (dotted line; in kilodaltons). **b**. Coomassie-stained blue native PAGE gel of ESX-5 complex solubilized in DDM or A8-35**.** Representative gel from at least three independent experiments. **c,** and **d,** Representative negative stain EM showing ESX-5 particles solubilized in DDM and A8-35, respectively. Scale bar is 50 nm.



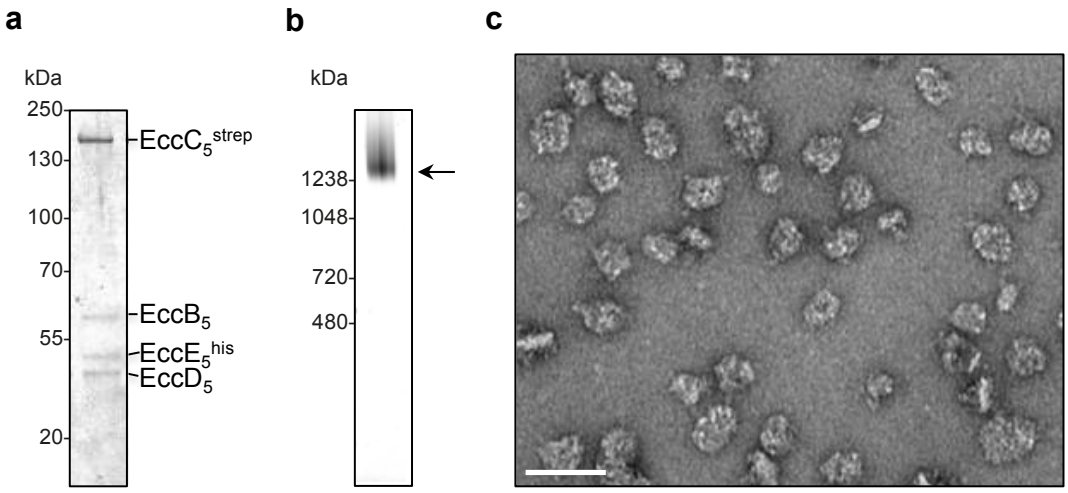
Supplementary Figure 3. Single particle analysis of ESX-5 complexes and map validation. a, Gallery of representative 2D class averages. **b**, Rotational cross-correlation (CC) from a representative 2D class average. **c**, Fourier Shell Correlation curve of the ESX-5 complex reconstruction. The resolution at 0.143 was 13.1Å. **d**, Upper panel shows a representative tilt-pair of ESX-5 particles at 0° and 50° tilting. Below are the corresponding views of the 3D map. **e,** Angular distribution applying C6 symmetry is shown by the graphical display of 'e2eulerxplor.py'. Projections in top, intermediate and side views are shown on the left row adjacent to the corresponding 2D class average. Particle orientation is evenly distributed in the 3D space extending from the equator to the pole, enabling a robust negative stain 3D reconstruction with no missing information. **f,** Representative tilt-pairs collected at 0° and 50° and subjected to tilt-pair analysis. **g,** Tilt-pair parameter plot (TPPP)[1] of tilt-pairs shows a clear clustering around the 50° tilt angle corresponding to the angle that was used to record the second micrograph of the tilt pair.



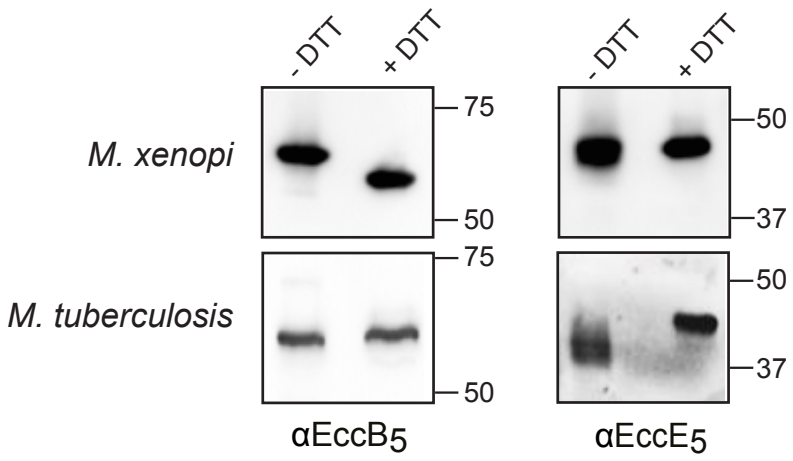
Supplementary Figure 4. Analysis of truncated $EccC_5$ ($EccC_5^{\Delta 99}$) by SEC-MALS and SAXS. a, Molecular weights were calculated from analytical size-exclusion chromatography analysis coupled with static light scattering and plotted across the peaks (blue). The derived molar weights of 147 kDa and 310 kDa match the expected molecular weights of a monomeric and dimeric protein, respectively. **b,** Pair-wise distance distribution (P(r)) of the scattering data. **c,** Fits of model ensemble 3 to the raw SAXS data. Inset shows the schematic overview of modelling approaches used with the Ensemble Optimization Method (EOM). Residue boundaries of $EccC_5$ domains are as follows: DUF, 99-431; A1, 431-712; A2, 826-1123; A3, 1135-1389. Domains boxed by dotted lines were treated as rigid bodies. The chi-squared (χ^2) values of each of the EOM models are indicated.



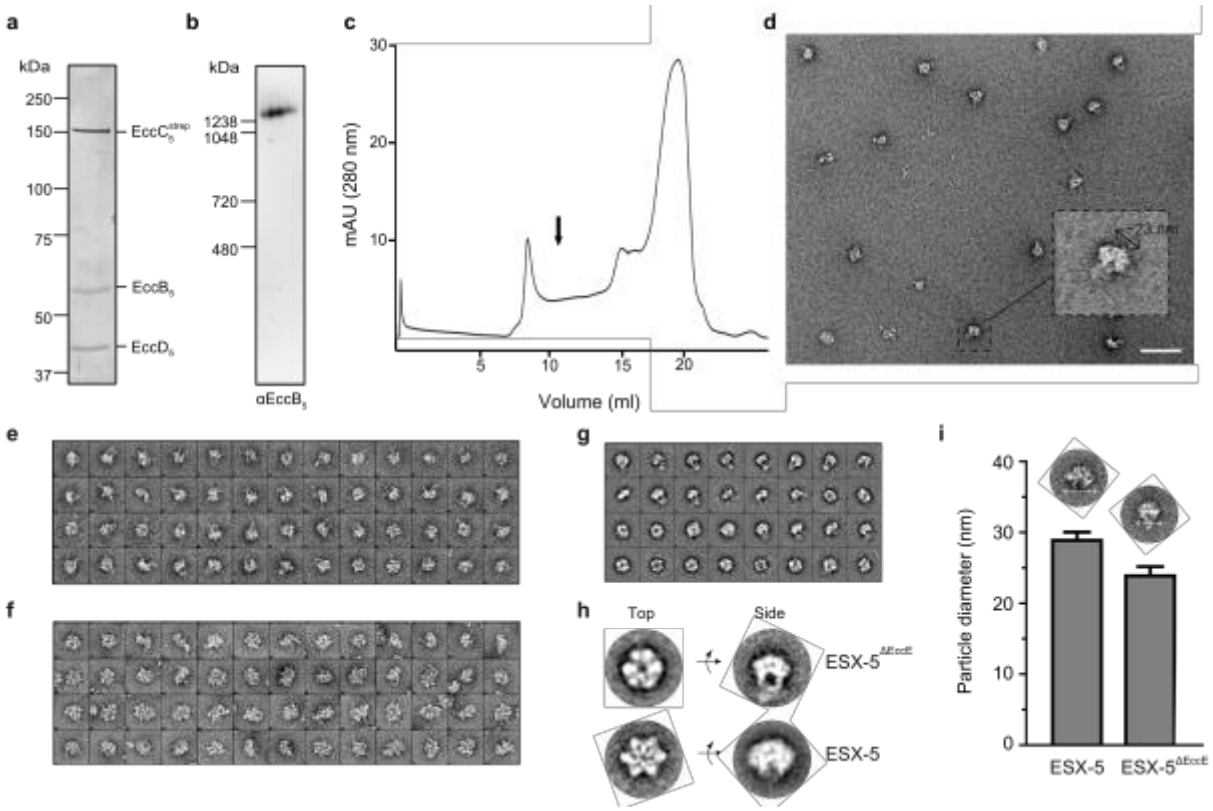
Supplementary Figure 5. Overview outlining $EccB_5$ docking procedures. a, Approach 1 used a hexameric model of $EccB_5$ of *M. xenopi* (i) generated using the method described by Zhang *et al*.[2]. Subsequently, the $EccB_5$ hexamer was docked into the EM map as a rigid body (ii). Top, side and cross section views are indicated. **b,** In approach 2, the EM map was segmented using the program Segger revealing a feature consistent with an $EccB_5$ hexamer (i). Next, individual $EccB_5$ monomers were docked into the segmented volume (ii) and a six-fold symmetry was applied (iii). Both docking approaches place the N-terminus of $EccB_5$ at the apex of the map, but differ in their positioning of the C-terminal region of $EccB_5$ preceding the transmembrane domain.



Supplementary Figure 6. Purification of *M. xenopi* ESX-5$^{EccE\text{-}His}$ complex. **a**, Coomassie-stained SDS-PAGE gel of purified ESX-5 complex following affinity purification using the Strep-tag II on $EccC_5$ and size-exclusion chromatography. Representative gel from at least three independent experiments. **b,** Coomassie-stained blue native PAGE gel of the purified complex as indicated by the black arrow. Representative gel from at least three independent experiments. **c**, Representative negative stain EM of the ESX-5 complex. Visual inspection of the particles suggests that the overall architecture of ESX-5$^{EccE\text{-}His}$ is comparable to the ESX-5 complex, indicating that the addition of a C-terminal His_6-tag on $EccE_5$ has no impact on complex formation or stability. Scale bar is 50 nm.



Supplementary Figure 7. Disulphide bridge formation of $EccE_5$ indicative of its periplasmic location. Immunoblot analysis of purified ESX-5 complexes under reducing (+ DTT) and non-reducing (- DTT) conditions to analyse disulphide bridge formation. This experiment was also performed on the ESX-5 system from *M. tuberculosis* reconstituted in *M. smegmatis* because of the higher number of cysteines in $EccE_5$ from *M. tuberculosis* (6) versus $EccE_5$ from *M. xenopi* (3). As a positive control $EccB_5$, that is known to form a disulfide bridge[3], was also examined. Both $EccB_5$ and $EccE_5$ showed a difference in mobility depending on the presence of DTT. Representative immunoblots using anti-$EccB_5$ and anti- $EccE_5$ antibodies from three independent experiments.



Supplementary Figure 8. Purification of *M. xenopi* ESX-5$^{\Delta EccE}$ complex. a, Coomassie-stained SDS-PAGE gel of purified complex following affinity purification and size-exclusion chromatography, confirming the presence of the three components $EccB_5$, $EccC_5^{strep}$, $EccD_5$ and the absence of $EccE_5$. Representative gel from at least three independent experiments. **b,** Immunoblot of blue native PAGE gel of purified ESX-5$^{\Delta EccE}$ complex indicating a molecular mass of ~ 1,3 MDa. Representative immunoblot using anti-$EccB_5$ antibodies from at least three independent experiments. **c,** Representative size exclusion chromatography profile of ESX-5$^{\Delta EccE}$. The yield of the ESX-5$^{\Delta EccE}$ complex is considerably lower than that of ESX-5. The black arrow indicates the fraction shown in panel d. **d,** Representative electron micrograph showing ESX-5$^{\Delta EccE}$ particles in several orientations. The overall diameter of the particles is ~22-25nm. Scale bar is 50 nm. Representative particles of the ESX-5$^{\Delta EccE}$ (**e**) and the ESX-5 complex (**f**) (box size = 180 x 180 pixels). **g,** Gallery of class averages of 684 ESX-5$^{\Delta EccE}$ particles. **h,** Four class averages of ESX-5 and ESX-5$^{\Delta EccE}$ complexes rotationally aligned are shown at same magnification. **i,** Histogram showing the average diameter for both ESX-5 and ESX-5$^{\Delta EccE}$ complexes. Particles were measured along the horizontal axis, the average diameter of the ESX-5 and ESX-5$^{\Delta EccE}$ complexes is 28.9 ± 1.2 nm and 23.8 ± 1.3, respectively. Error bars indicate the standard deviation of 50 particles. T-test analysis shows a significant difference with a p-value ≤ 0.0001

Supplementary Figure 9
Full blots of results

Figure 1a

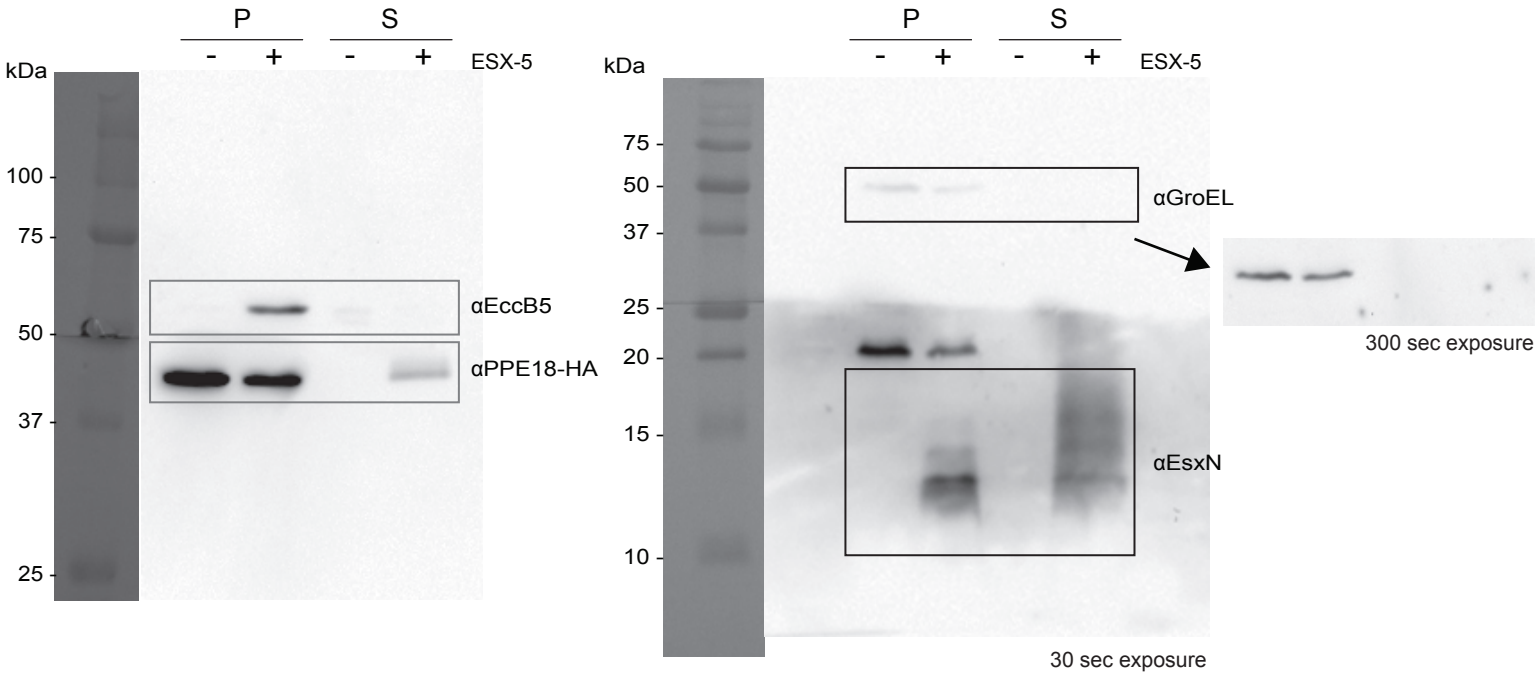
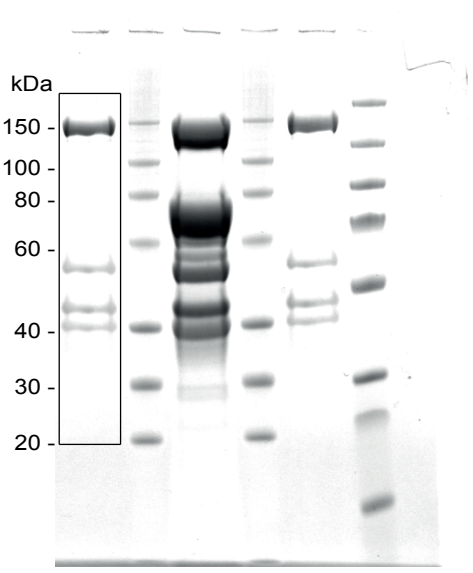
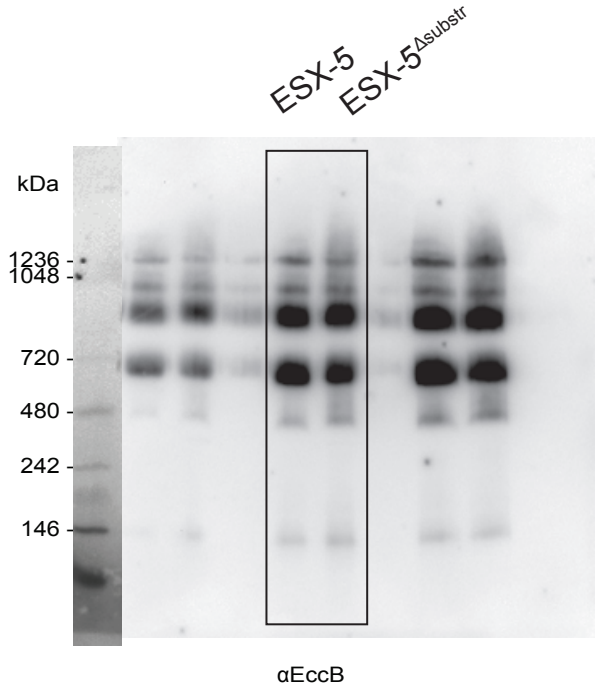


Figure 1b

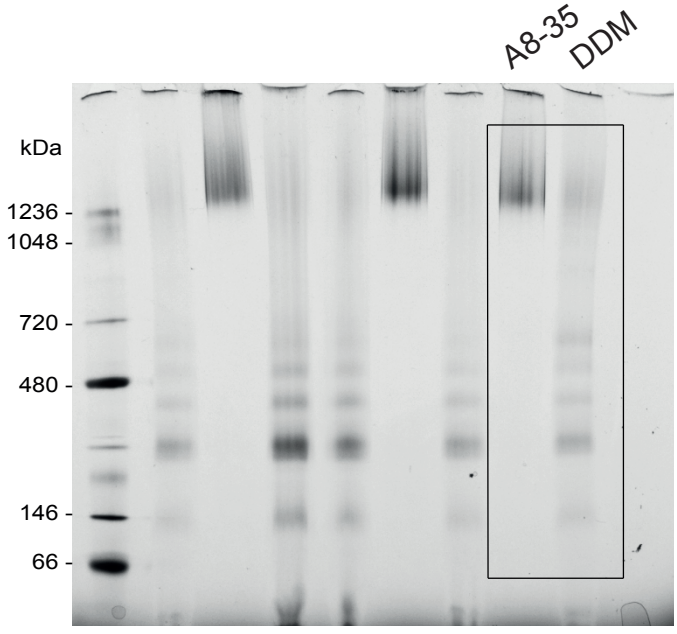


Full blots of results (continued)

Supplementary Fig. 1b

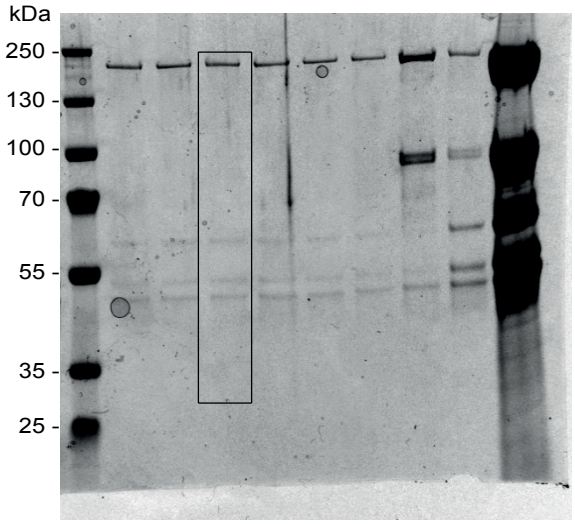


Supplementary Fig. 2b

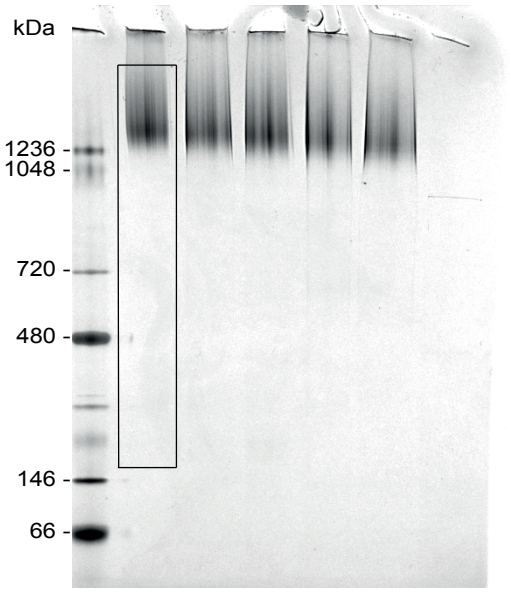


Full blots of results (continued)

Supplementary Fig. 6a



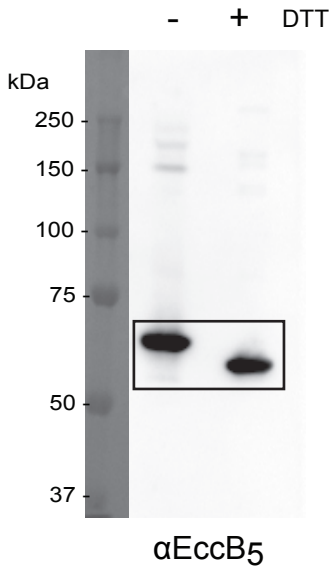
Supplementary Fig. 6b



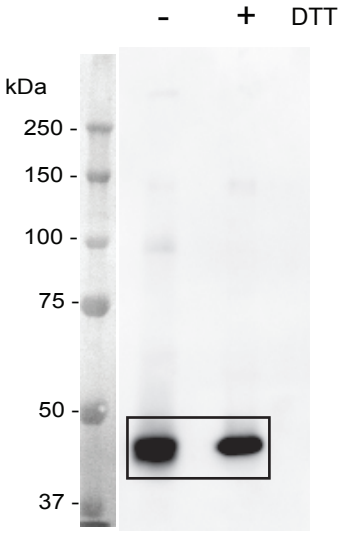
Full blots of results (continued)

Supplementary Fig. 7

M. xenopi

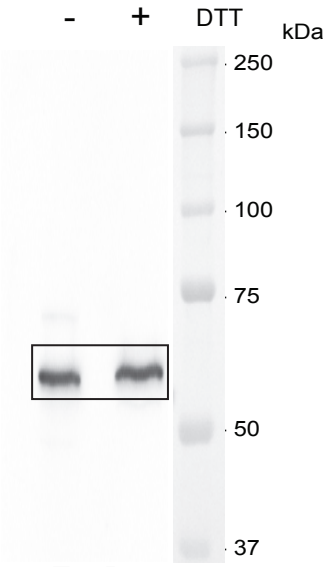


αEccB5

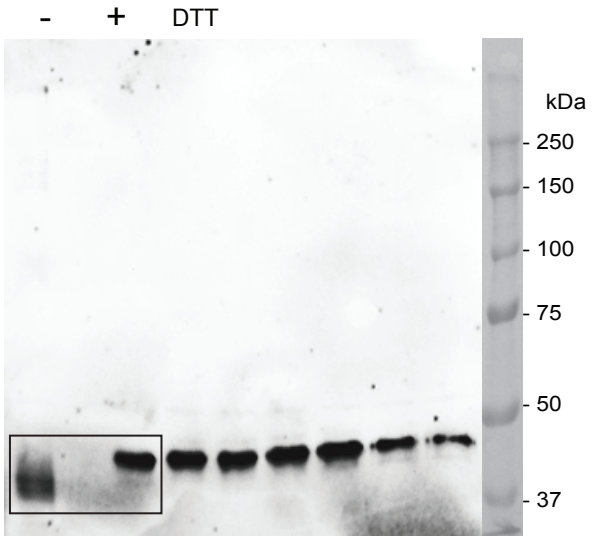


αEccE5

M. tuberculosis



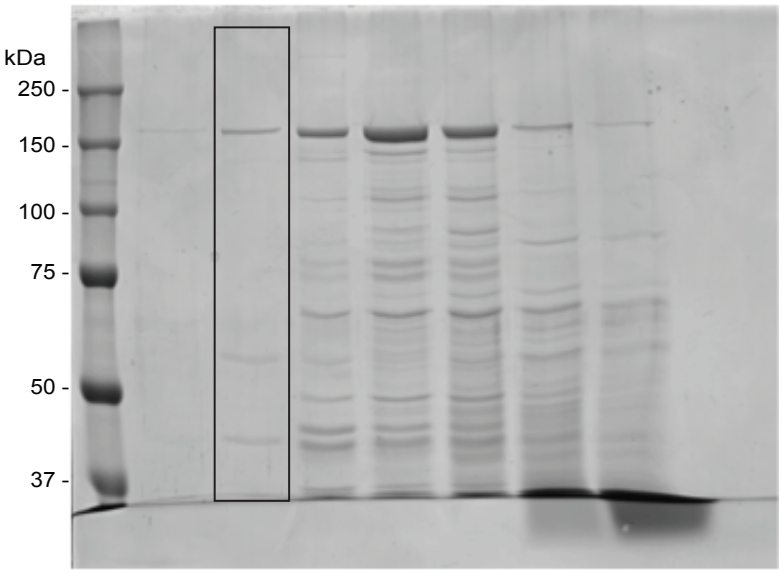
αEccB5



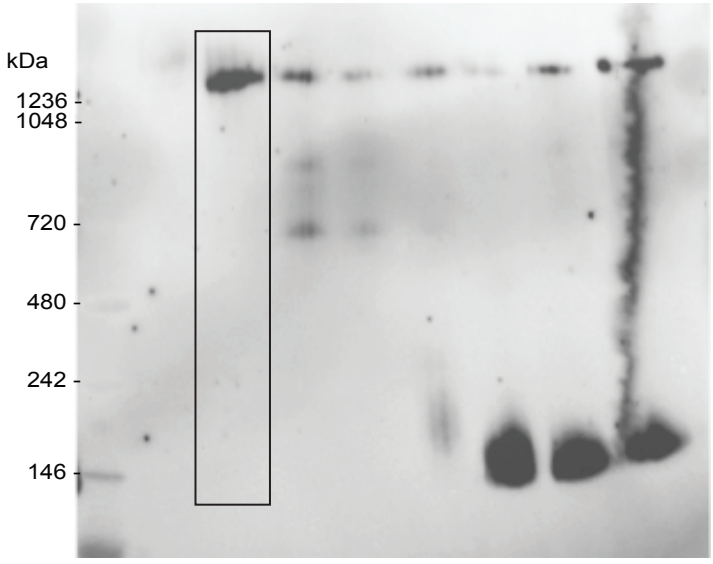
αEccE5

Full blots of results (continued)

Supplementary Fig. 8a



Supplementary Fig. 8b



Supplementary Table 1 | Quantitative mass-spectrometry analysis of ESX-5 to determine subunit stoichiometry. Average relative protein concentrations of the four subunits was determined using iBAQ-based quantitative mass spectrometry and calculated across four biological replicates (1-4). These date indicate a 1:1:1:1 stoichiometry for $EccB_5$:$EccC_5$:$EccD_5$:$EccE_5$.

Proteins	iBAQ				Relative protein concentration				Average relative protein concentration	S.E.	Razor + unique peptides				Sequence coverage (%)				MW (kDa)
	1	2	3	4	1	2	3	4			1	2	3	4	1	2	3	4	
$EccB_5$	3.5E+09	1.3E+10	5.0E+09	2.7E+08	0.8	0.9	0.8	1.2	0.9	0.1	18	22	20	21	48.2	52.0	49.6	42.5	53.2
$EccC_5$	4.5E+09	1.6E+10	6.6E+09	2.2E+08	1.0	1.1	1.0	1.0	1.0	0.0	70	81	77	67	67.2	77.4	71.3	47.7	152.8
$EccD_5$	5.2E+09	1.6E+10	8.0E+09	2.2E+08	1.2	1.1	1.2	1.0	1.1	0.1	13	15	13	8	25.5	26.3	25.5	11.4	53.4
$EccE_5$	4.1E+09	1.5E+10	7.2E+09	2.1E+08	1.0	1.0	1.1	0.9	1.0	0.0	20	23	23	24	65.2	67.0	69.5	57.5	44.4

Supplementary Table 2 | EM data collection and structural analysis of ESX-5.

Micrographs	65
Particles	1,418
Defocus	1.5-5.0 um
Detector	FEI BM Eagle
Pixel size	2.18 Å
Reconstruction method	EMAN2 - *e2initialmodel*
Symmetry	6-fold
Resolution FSC 0.143 gold standard	13.1 Å

Supplementary Table 3 | SAXS data collection and derived parameters for $EccC_5^{\Delta 99}$.

Data collection parameters	
Instrument	EMBL P12 beam line (PETRA-III, DESY, Hamburg)
Beam geometry	0.2 x 0.12 mm^2
Wavelength (Å)	1.24
s range ($Å^{-1}$)[a]	0.01-0.46
Exposure time (s)	1 (20 x 0.05)
Concentration range (mg/mL)	1.0-5.0
Temperature (K)	283
Structural parameters[b]	
I(0) (relative) (from *p(r)*)	0.108 ± 0.001
R_g (Å) (from *p(r)*)	65 ± 1
I(0) (relative) (from Guinier)	0.107 ± 0.001
R_g (Å) (from Guinier)	62 ± 1
D_{max} (Å)	230
Porod volume estimate ($Å^3$)	255,000 ± 30,000
Excluded volume estimate ($Å^3$)	87,000 ± 10,000
Dry volume calculated[c] from sequence ($Å^3$)	172,365
Molecular mass determination	
I(0) (cm^{-1}) BSA (72,000 Da)	0.043 ± 0.001
Molecular mass M_r (Da) (from *I(0)*)	141,000 ± 15,000
Molecular mass M_r (Da) (from Porod volume (V_p/*1.7)*)	159,000 ± 15,000
Molecular mass M_r (Da) (from excluded volume (V_{ex}/*2)*)	173,500 ± 20,000
Calculated monomeric M_r (from sequence)	142,450
Software employed	
Primary data reduction	RADAVER
Data processing	PRIMUS/Qt
Equilibrium analysis	EOM
Computation of model intensities	CRYSOL
3D graphics representations	PyMOL, UCSF Chimera

[a] Momentum transfer $|s| = 4\pi\sin(\theta)/\lambda$

[b] Values reported for 0.4 mg mL^{-1}

[c] Dry volume from: http://www.basic.northwestern.edu/biotools/proteincalc.html

Abbreviations:

M_r : Molecular mass

R_g : Radius of gyration

D_{max} : Maximal particle dimension

V_p : Porod volume

V_{ex} : Particle excluded volume

References

1. Henderson, R. *et al.* Tilt-pair analysis of images from a range of different specimens in single-particle electron cryomicroscopy. *J. Mol. Biol.* **413,** 1028–46 (2011).
2. Zhang, X.-L. *et al.* Core component EccB1 of the Mycobacterium tuberculosis type VII secretion system is a periplasmic ATPase. *FASEB J.* **29,** 4804–14 (2015).
3. Wagner, J. M. *et al.* Structures of EccB1 and EccD1 from the core complex of the mycobacterial ESX-1 type VII secretion system. *BMC Struct. Biol.* **16,** 5 (2016).