Supplementary Material for M.L. Barta *et al*.

Detailed Methodology

Cloning, Purification and Limited Proteolysis of Recombinant Translocator Proteins.

A DNA fragment encoding full-length IpaB (residues 1-580) was amplified from the virulence plasmid of *S. flexneri* via PCR and subcloned into the expression plasmid pT7HMT [1]. Upon verification of its sequence, this vector was co-transformed, along with an analogous fragment encoding for full-length IpgC (residues 1–155) in the pACYC-DUET vector [2], into BL21 (DE3) *E. coli* cells and cultured in kanamycin (50 μg/ml) and chloramphenicol (30 μg/ml)-selective Terrific Broth at 37°C to an $A_{600\ nm}$ of 0.800. Protein co-expression was induced overnight at 18 °C by adding IPTG to 1 mM final concentration. Cells were harvested by centrifugation, resuspended in lysis buffer (20 mM Tris (pH 8.0), 500 mM NaCl, and 10 mM imidazole), and then lysed in a microfluidizer. Soluble translocator/chaperone complex was collected in the supernatant following centrifugation of the cell homogenate and purified on a Ni^{2+}-NTA Sepharose column according to standard protocols [1]. Chromosomal *S.* Typhimurium DNA was used to amplify a DNA fragment encoding full-length SipB (residues 1–593) via PCR, which was also subcloned into pACYC-DUET (Novagen). The co-expression and purification of sequence verified SipB with full-length SicA (residues 1–165) in the pT7HMT vector [1] was accomplished as described above.

Limited proteolysis of purified IpgC/IpaB and SicA/SipB complexes was carried out using subtilisin digestion at 25°C with a protein concentration of 2 mg/ml to generate protease stable translocator fragments. Varying enzyme:substrate ratios (by concentration) were first analyzed in pilot experiments to determine the appropriate ratio for these studies (data not shown); a 1:200 ratio was identified as optimal for IpgC/IpaB while a 1:400 ratio was utilized for SicA/SipB. Over the course of 120 min, 5 μl aliquots were taken at various intervals and boiled immediately in SDS-sample buffer for 10 min. SDS-PAGE was used to separate the samples and various degradation products were characterized by LC-MS/MS as previously described [3].

Using the results of the above experiments as a guide, individual DNA fragments encoding IpaB (residues 28-226) and SipB (30-237) were designed, amplified by PCR, and subcloned into the pT7HMT plasmid. Over-expression and purification of the translocator fragments was carried out as described above, with the exception that recombinant TEV protease was used to digest the fusion affinity tag from each target protein [1]. Final purification for $IpaB^{28.226}$ was achieved by desalting into 20 mM ethanolamine (pH 9.0) prior to performing cation-exchange chromatography over a Resource S column (GE Lifesciences). Purified $IpaB^{28.226}$ was concentrated to 5 mg/ml and exchanged into ddH_2O for further use. Final purification of $SipB^{30.237}$ was achieved by desalting into 20 mM Tris HCl (pH 8.0) prior to performing anion-exchange chromatography over a Resource Q column (GE Biosciences). Purified $SipB^{30.237}$ was concentrated to 10 mg/ml and buffer exchanged into 10 mM Tris (pH 7.1), 50 mM NaCl for further use.

Structure Determination for the IpaB N-terminal Fragment.

Initial crystallization experiments for $IpaB^{28.226}$ failed to yield any samples suitable for X-ray diffraction analysis. However, adventitious proteolysis at the N-terminus of IpaB was observed over the course of several months, and subsequent rescreening using this partly degraded sample identified a condition that yielded single brick-shaped crystals in one day. Following optimization, the final crystal samples were obtained by vapor diffusion of hanging drops at 20 °C where 2.5 µl of protein solution (2.5 mg/ml in water) was mixed with 2.5 µl of reservoir solution (0.6 M sodium formate and 21% (w/v) PEG 2000 MME) and equilibrated over 500 µl of reservoir solution. Crystals were flash cooled in a cryoprotectant solution consisting of reservoir buffer supplemented with 15% (v/v) glycerol.

Screening for phasing atoms was accomplished by Native polyacrylamide gel electrophoresis (PAGE) of $IpaB^{28.226}$ and heavy atom mixtures [4]. Single band gel shifts were considered indicators of changes in protein charge, which potentially arose from heavy atom derivatization. Compounds demonstrating Native PAGE shifts were chosen for further screening via in-drop soaking. Platinum derivatives were generated by adding 1 µl of K_2PtCl_4, K_2PtCl_6 and $K_2Pt(NO_2)_4$ (at concentrations ranging

from 0.5 mM to 20 mM) in reservoir solution to the 5 μl drop, and incubated overnight. Intact crystals were flash cooled as described above.

Monochromatic X-ray diffraction data (λ=1.000 Å) were collected from single native IpaB$^{28.226}$ crystals at -173 °C using beamline 22-BM of the Advanced Photon Source, Argonne National Laboratory. Heavy atom soaked IpaB$^{28.226}$ crystals were exposed to X-rays at Pt peak and remote energies (1.07197 Å and 1.0000 Å; Table 1). Following data collection, individual reflections were indexed, integrated, and scaled using HKL2000 [5]. Experimental phase information was obtained from AutoSol within the PHENIX suite [6] using the anomalous and dispersive differences between diffraction data collected at Pt peak and remote energies from a single K_2PtCl_4-derivatized IpaB$^{28.226}$ crystal. This identified two Pt heavy-atom sites within the asymmetric unit. Iterative rounds of automated chain-tracing, model building, and refinement at 2.5 Å limiting resolution were carried out by AutoBuild [7; 8]. Upon completion, one round of individual coordinate and isotropic atomic-displacement factor refinement was conducted manually, and this refined model was used to calculate both $2F_o$-F_c and F_o-F_c difference maps. These maps were used to iteratively improve the experimental model by manual building in Coot [9; 10], followed by additional coordinate and atom-displacement factor refinement. A final round of modeling and refinement, which included the addition of ordered solvent molecules, was carried out to 2.1 Å limiting resolution using the native data set described in Table 1.

The final model contains two IpaB polypeptides within the asymmetric unit, which corresponds to a Matthews coefficient of 2.28 Å^3/Da and a solvent content of 46%. Electron density corresponding to residues 74-224 of IpaB was modeled for both chains within the asymmetric unit. Additional information and refinement statistics for all three structures is presented in Table 1. The coordinates and refined structure factors have been deposited in the RCSB database [11] under the accession code **3U0C**.

Structure Determination for the SipB N-terminal Fragment.

Crystals of Seleno-L-Methionine-labeled SipB$^{30.237}$ failed to appear upon initial crystallization trials. However, similarly to the crystallization of IpaB$^{28.226}$, adventitious degradation yielded a stable protein core. Crystals of degraded SipB$^{30.237}$ were obtained by vapor diffusion of hanging drops at 20°C.

Specifically, 1.5 µl of protein solution (10 mg/ml in 10 mM Tris (pH 7.1), 50 mM NaCl) was mixed with 1.5 µl of reservoir solution containing 0.15 M potassium bromide and 27% (w/v) PEG 2000 MME, and the drops were equilibrated over 500 µl of reservoir solution. Small, blade-like microcrystals appeared within one day. Diffraction quality crystals were obtained over a one week period following microseeding [12]. Crystals were flash cooled in a cryoprotectant solution consisting of reservoir buffer alone.

X-ray diffraction data were collected at Se remote (λ=0.97243 Å) and peak (λ=0.97934 Å) energies from a single SeMet-labeled SipB$^{30.237}$ crystal at -173 °C using beamline 22-BM of the Advanced Photon Source, Argonne National Laboratory (Table 1). Following data reduction as described above, experimental phase information was obtained by AutoSol [6] from the anomalous and dispersive differences between the two SeMet-labeled SipB$^{30.237}$ data sets; this allowed identification and refinement of twelve Se heavy atom sites within the asymmetric unit. One round of individual coordinate and isotropic atomic-displacement factor refinement were conducted following examination of the initial model built by AutoSol [6], and this refined model was used to calculate both $2F_o$-F_c and F_o-F_c difference maps. These maps were used to iteratively improve the experimental model by manual building in Coot [9; 10], followed by additional coordinate and atom-displacement factor refinement. Side chains were modeled according to the SipB sequence using SeMet sites as a guide. Ordered solvent molecules were added according to the default criteria of *phenix.refine*, and inspected manually using Coot prior to model completion.

Analysis of the asymmetric unit revealed four SipB polypeptides, which corresponds to a Matthews coefficient of 2.53 Å^3/Da and a solvent content of 54%. Electron density corresponding to residues 82-122, 126-174 and 182-226 of SipB was modeled in at least one polypeptide chain. However, poor map quality prevented accurate modeling of two solvent-exposed loop regions (residues 123-125 and 175-181). Additional information and refinement statistics for the structure is presented in Table 1. The coordinates and refined structure factors here have been deposited in the RCSB database [11] under the accession code **3TUL**.

Supplementary References.

1. Geisbrecht, B., Bouyain, S. & Pop, M. (2006). An optimized system for expression and purification of secreted bacterial proteins. *Protein Expr Purif* **46**, 23-32.
2. Barta, M. L., Zhang, L., Picking, W. L. & Geisbrecht, B. V. (2010). Evidence for alternative quaternary structure in a bacterial Type III secretion system chaperone. *BMC Struct Biol* **10**, 21.
3. Kinter, M. & Serman, N. E. (2000). *Protein sequencing and identification using tandem mass spectrometry*, Wiley-Interscience, New York.
4. Boggon, T. J. & Shapiro, L. (2000). Screening for phasing atoms in protein crystallography. *Structure* **8**, R143-9.
5. Otwinowski, Z. a. M., W. (1997). Processing of X-ray Diffraction Data Collected in Oscillation Mode. *Methods in Enzymology* **276**, 307-326.
6. Terwilliger, T. C., Adams, P. D., Read, R. J., McCoy, A. J., Moriarty, N. W., Grosse-Kunstleve, R. W., Afonine, P. V., Zwart, P. H. & Hung, L. W. (2009). Decision-making in structure solution using Bayesian estimates of map quality: the PHENIX AutoSol wizard. *Acta Crystallogr D Biol Crystallogr* **65**, 582-601.
7. Terwilliger, T. C., Grosse-Kunstleve, R. W., Afonine, P. V., Moriarty, N. W., Zwart, P. H., Hung, L. W., Read, R. J. & Adams, P. D. (2008). Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard. *Acta Crystallogr D Biol Crystallogr* **64**, 61-9.
8. Zwart, P. H., Afonine, P. V., Grosse-Kunstleve, R. W., Hung, L. W., Ioerger, T. R., McCoy, A. J., McKee, E., Moriarty, N. W., Read, R. J., Sacchettini, J. C., Sauter, N. K., Storoni, L. C., Terwilliger, T. C. & Adams, P. D. (2008). Automated structure solution with the PHENIX suite. *Methods Mol Biol* **426**, 419-35.
9. Emsley, P. & Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126-32.
10. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**, 486-501.
11. Krissinel, E. & Henrick, K. (2007). Inference of macromolecular assemblies from crystalline state. *J Mol Biol* **372**, 774-97.
12. Oswald, C., Smits, S. H., Bremer, E. & Schmitt, L. (2008). Microseeding - a powerful tool for crystallizing proteins complexed with hydrolyzable substrates. *Int J Mol Sci* **9**, 1131-41.
13. DeLano, W. L. (2002). The PyMOL Molecular Graphics System. **2009**, http://www.pymol.org.
14. Zemla, A. (2003). LGA: A method for finding 3D similarities in protein structures. *Nucleic Acids Res* **31**, 3370-4.
15. Thompson, J., Higgins, D. & Gibson, T. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* **22**, 4673-80.
16. Gouet, P., Courcelle, E., Stuart, D. & Métoz, F. (1999). ESPript: analysis of multiple sequence alignments in PostScript. *Bioinformatics* **15**, 305-8.
17. Jaroszewski, L., Rychlewski, L., Li, Z., Li, W. & Godzik, A. (2005). FFAS03: a server for profile--profile sequence alignments. *Nucleic Acids Res* **33**, W284-8.

Supplementary Figure Legends.

SFig. 1. Limited Subtilisin Digest of Translocator/Chaperone Complexes. *A*, Full length class II chaperone, IpgC, was co-expressed with the full length first translocator, IpaB, from *Shigella flexneri*. Following IMAC purification, the translocator/chaperone complex was treated with subtilisin (1:200 substrate:enzyme ratio) for 120 min, with aliquots taken at various intervals. Samples were separated via SDS-PAGE and their tryptic peptides were identified by LC-MS/MS. The band identified as S3-103-4 was comprised of IpaB residues 28-226. *B*, Full length class II chaperone, SicA, was co-expressed with the full length first translocator, SipB, from *Salmonella* Typhimurium. Following IMAC purification, the translocator/chaperone complex was treated with subtilisin (1:400 substrate:enzyme ratio) for 120 min, with aliquots taken at various intervals. Samples were separated via SDS-PAGE and their tryptic peptides were identified by LC-MS/MS.

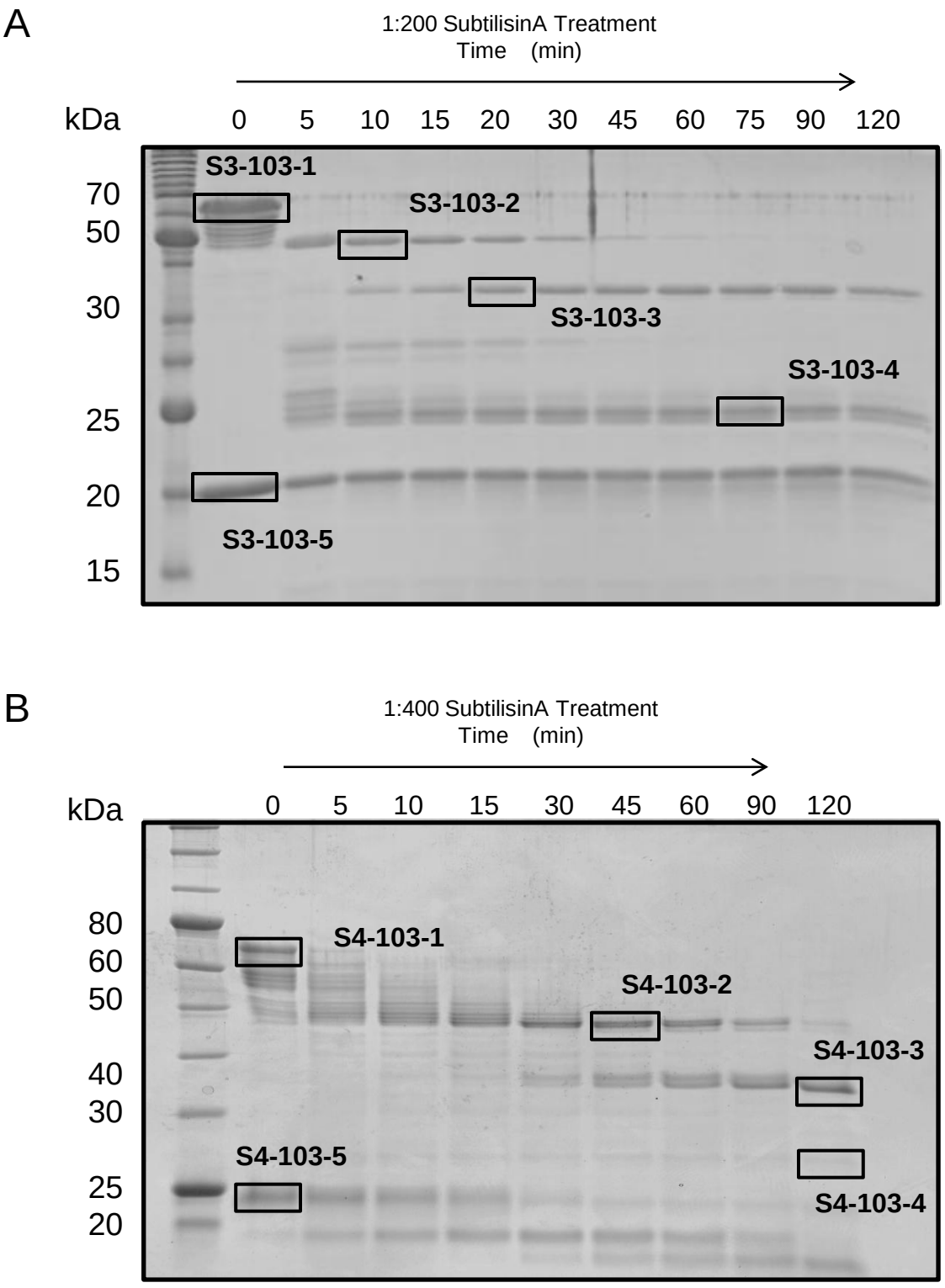
SFig. 2. Refined IpaB and SipB Asymmetric Unit. *A*, Ribbon diagram of IpaB asymmetric unit depicting 2 distinct polypeptides. *B*, Ribbon diagram of SipB asymmetric unit. SipB polypeptides with conformationally rigid α1 are colored cyan and purple. *C*, Structural superposition of chains A (cyan) and B (red) from panel A depicting the 16.8 Å N-terminal movement (measurement from Gly83 carbonyl on each chain). Representations of all structures were generated using PyMol [13].

SFig. 3. Structural Superposition of Type Three Secretion Protein Coiled-coils. *A*, Ribbon diagram of a structural alignment of IpaD (residues 39-322, grey) and IpaB (residues 104-224, purple) with an RMSD of 2.47 Å over 101/121 Cα atoms within 5.0 Å. *B*, Ribbon diagram of a structural alignment of SipD (residues 36-342, grey) and SipB (residues 126-226, cyan) with an RMSD of 2.39 Å over 88/94 Cα atoms within 5.0 Å. Representations of all structures were generated using PyMol [13]. Three-dimensional structures were superimposed using the Local-Global Alignment method (LGA) [14].

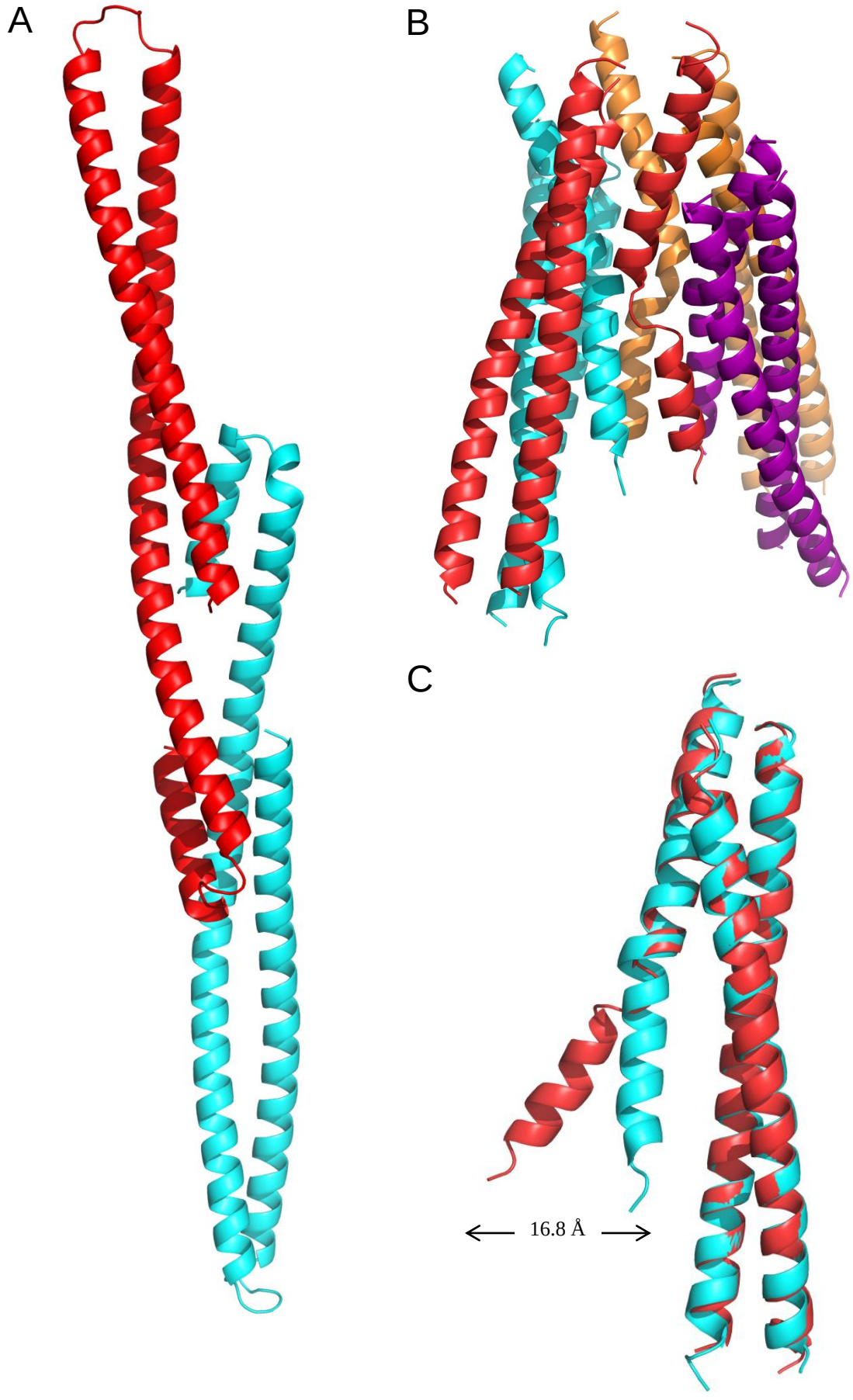
SFig. 4. **Inv-Mxi-Spa T3S Translocator Sequence Alignment.** Sequence alignment of animal pathogen type III secretion first translocators from *S. flexneri*, *B. pseudomallei*, and *S.* Typhimurium colored according to residue conservation (cyan = absolute, purple = similar). Alignment was generated using ClustalW and rendered with ESPRIPT. Numbers above the sequences correspond to *S. flexneri* IpaB. Secondary structure elements of IpaB and SipB are shown above and below the alignment, respectively. Sequence alignments were carried out using CLUSTALW [15] and aligned with secondary structure elements using ESPRIPT [16].

SFig. 5. **Structural Superposition of Colicin Pore-forming Domains.** Ribbon diagrams of the structural superposition of the pore-forming domain from Colicins Ia (green, PDB code: 1CII), E1 (purple, 2I88), N (yellow, 1A87), A (cyan, 1COL) and B (pink, 1RH1). Highlighted in red is the hydrophobic α-helical hairpin responsible for anchoring the pore forming domain in the host lipid bilayer. Crystal structures are rotated 90° downward about the central axis of the hairpin on the right. Representations of all structures were generated using PyMol [13]. Translocators sequences were queried for fold and function homology within the PDB via FFAS [17].

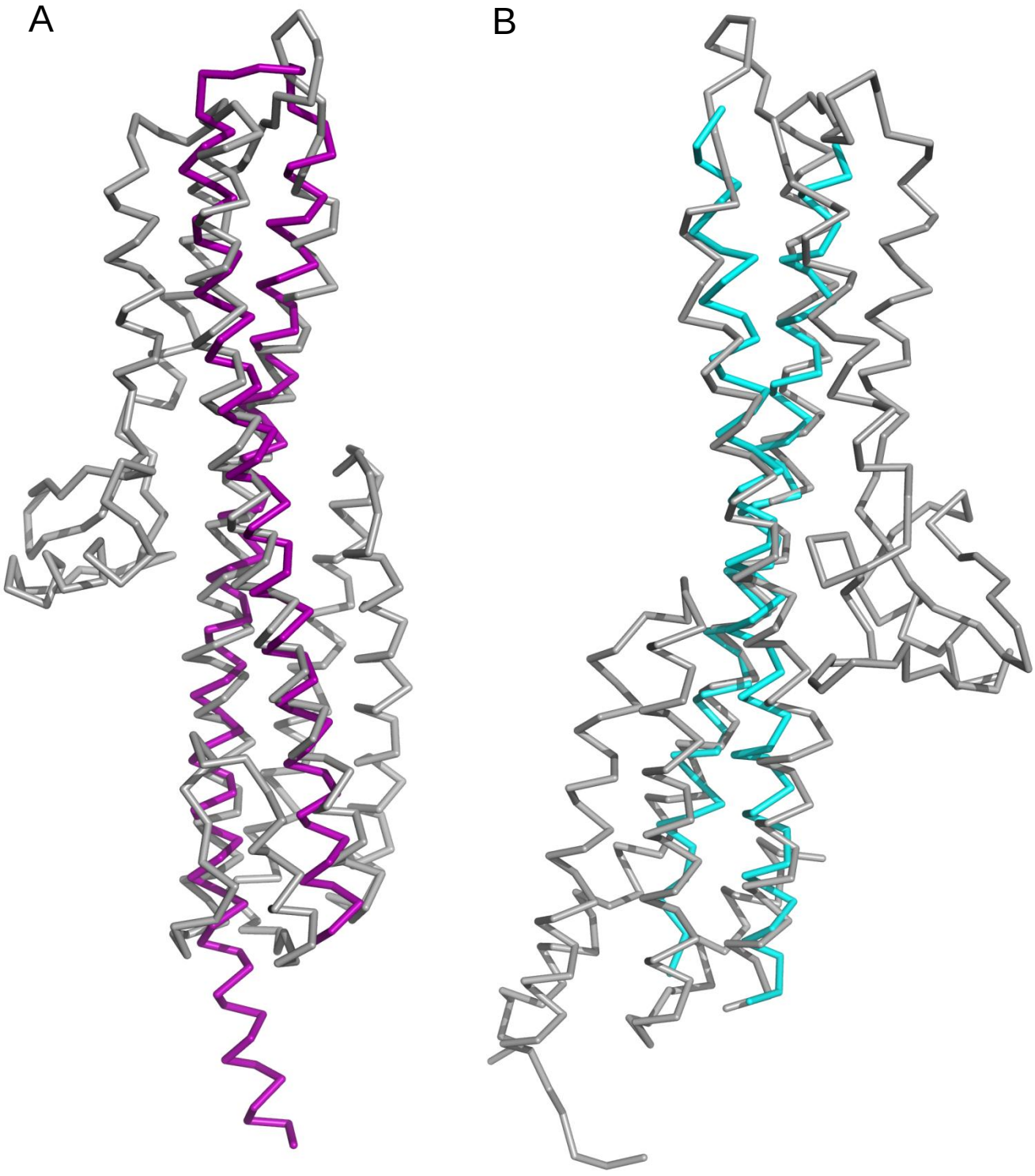
SFig. 1


A
1:200 SubtilisinA Treatment
Time (min)
kDa
0 5 10 15 20 30 45 60 75 90 120
70
50
30
25
20
15
S3-103-1
S3-103-2
S3-103-3
S3-103-4
S3-103-5
B
1:400 SubtilisinA Treatment
Time (min)
kDa
0 5 10 15 30 45 60 90 120
80
60
50
40
30
25
20
S4-103-1
S4-103-2
S4-103-3
S4-103-4
S4-103-5

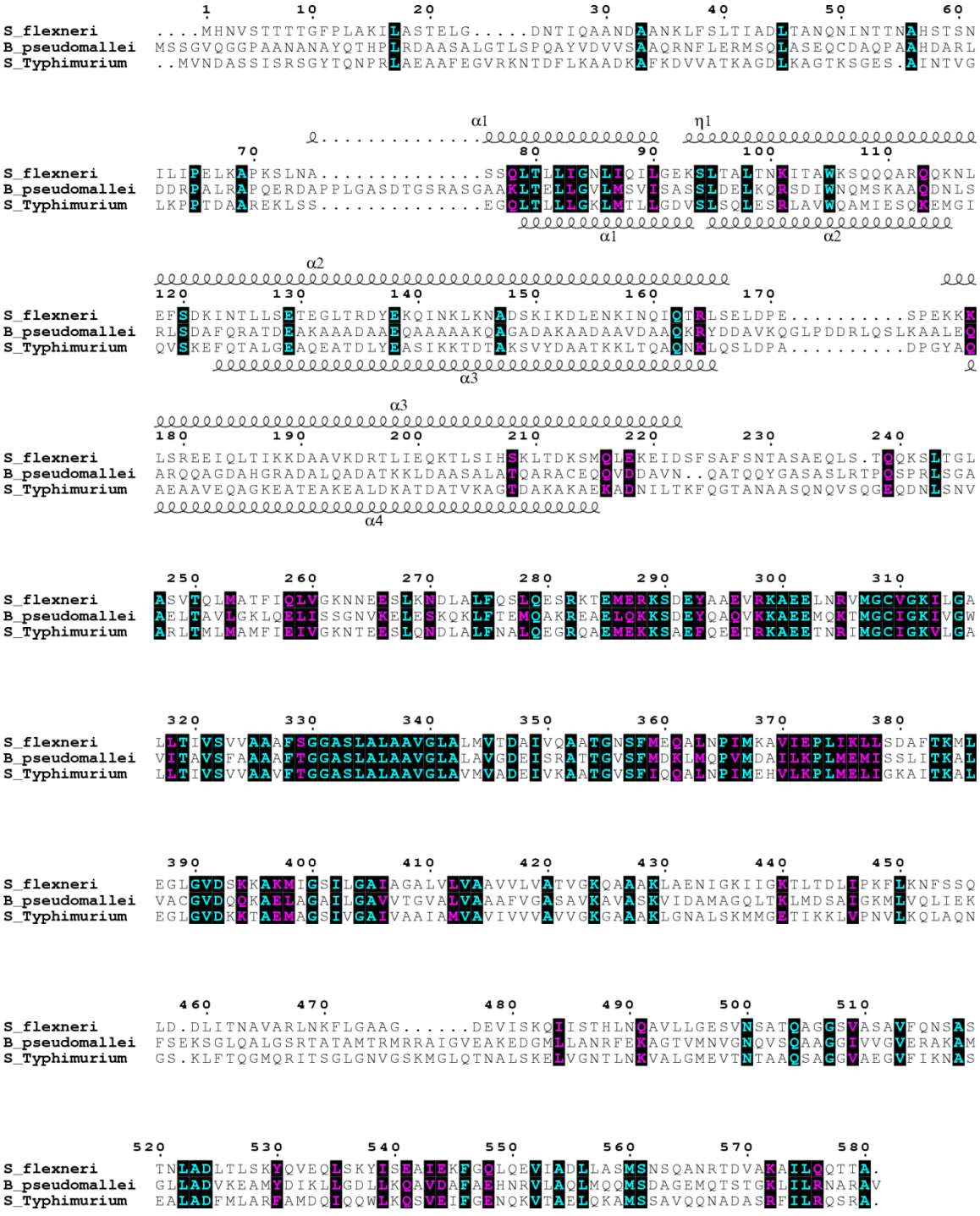
SFig. 2



SFig. 3



SFig. 4



SFig. 5