Directed Assembly of Homopentameric Cholera Toxin B-Subunit Proteins into Higher-Order Structures Using Coiled-Coil Appendages

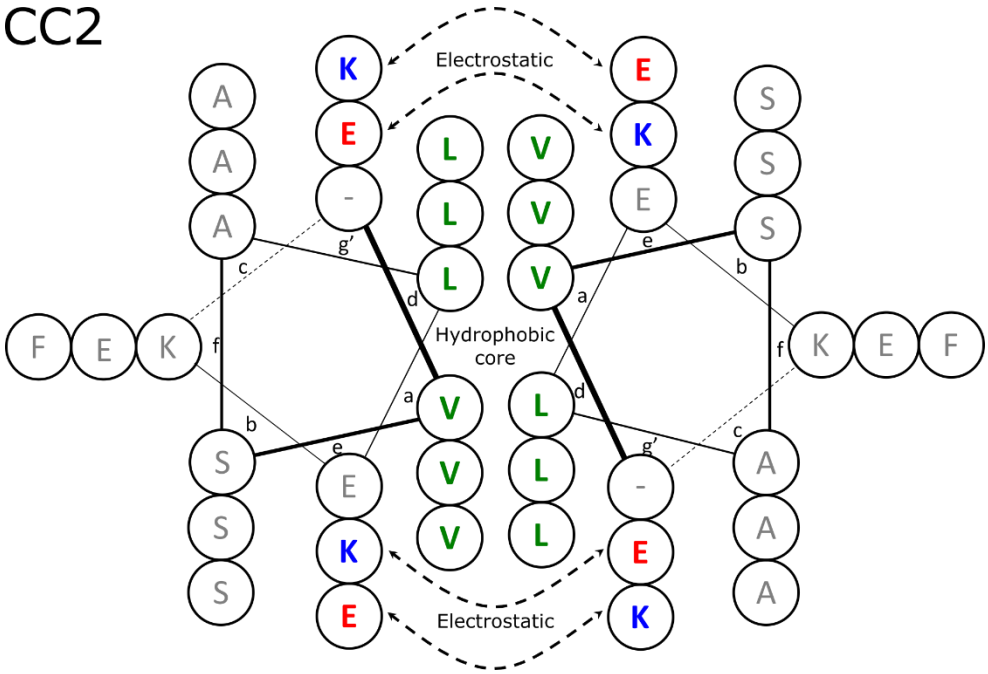
James F. Ross, Gemma C. Wildsmith, Michael Johnson, Daniel L. Hurdiss, Kristian Hollingsworth, Rebecca F. Thompson, Majid Mosayebi, Chi H. Trinh, Emanuele Paci, Arwen R. Pearson, Michael E. Webb and W. Bruce Turnbull

The data associated with this paper are openly available from the University of Leeds Data Repository at DOI: 10.5518/575

Supplementary information

1. General coiled-coil information

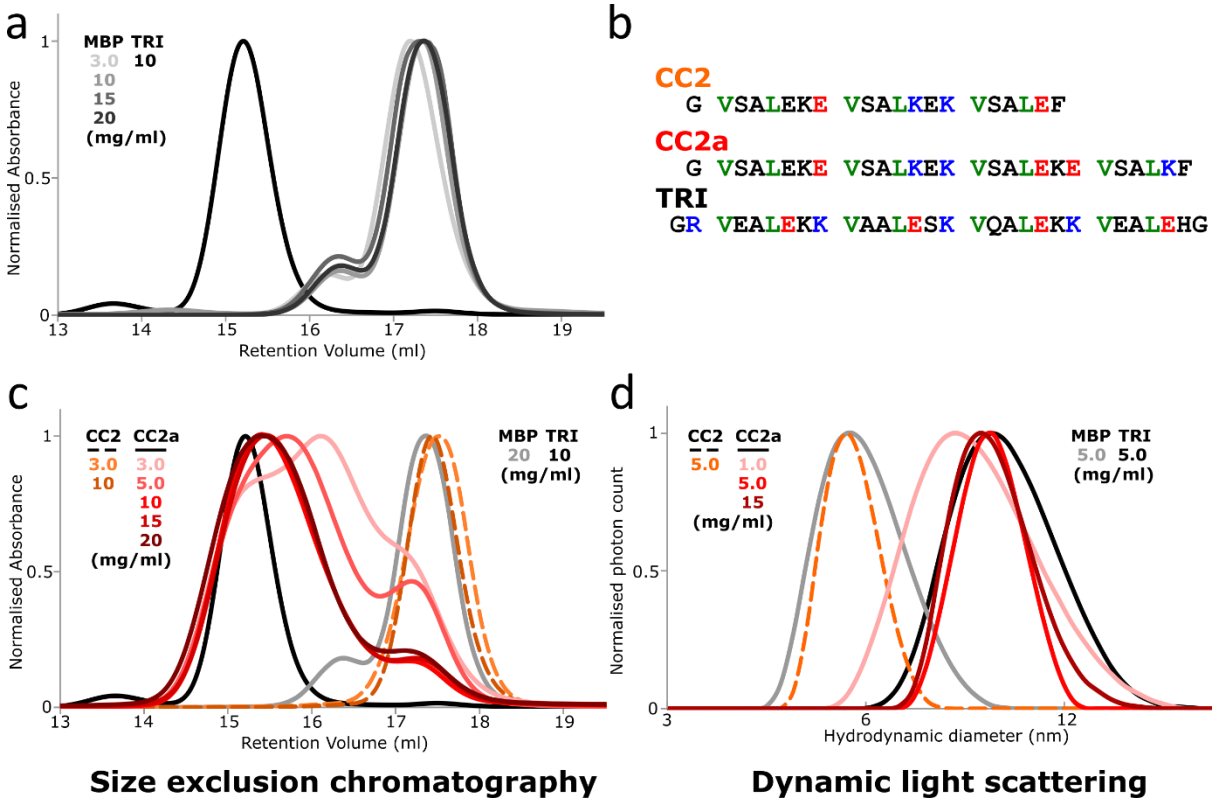
The coiled-coils used in the study were based on the five heptad hetero-dimeric coil-coil described in De Crescenzo et al.[1] To create CC2, a single glycine residue was used as a spacer between the end of the CTB gene and the start of the CC2 sequence. The charged residues at the '*e*' and '*g*' positions were altered to allow for homomeric association of the coils and the '*a*' and '*d*' positions were retained as valine and leucine respectively to promote dimeric association between the coils. As discussed in the main text, this sequence was designed as three-heptad coiled-coil, a four-heptad version of this coil were also designed (SI Fig. S2) where CC2a is the four-heptad version of CC2. Full sequences are given in SI-Fig. S2.



Supplementary Figure S1: Coiled-coil wheel for CC2. Coiled-coil wheel describing the interaction of the homomeric CC2 extension to CTB. The hydrophobic knob residues are highlighted in green and the charged interactions are highlighted in blue and red respective to charge. Non-interacting residues are in grey.

2. Coiled-coil oligomeric state analysis with maltose binding protein

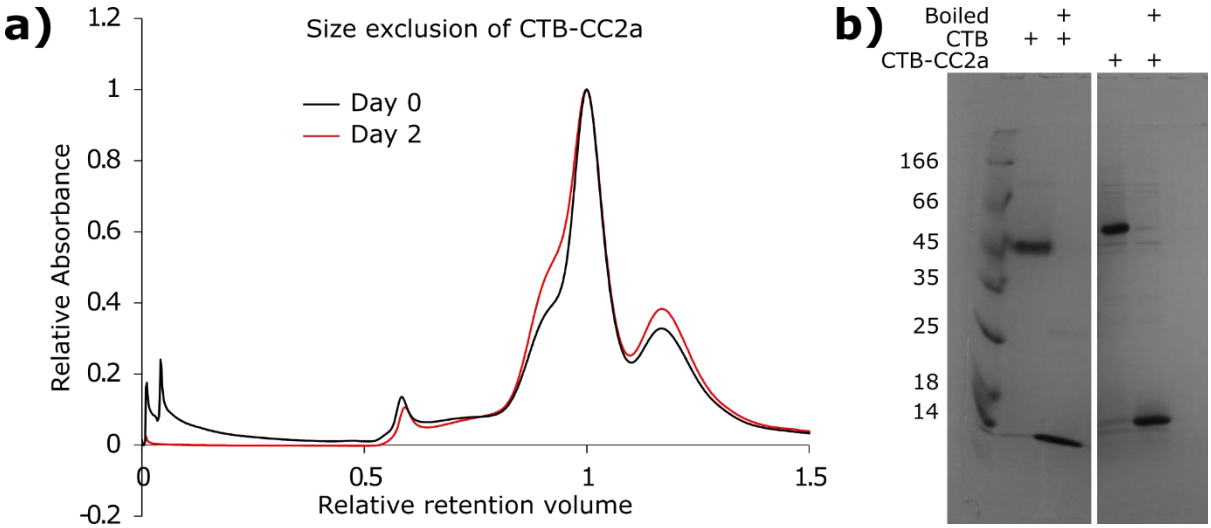
The four-heptad version of coiled-coil CC2a was assessed by SEC and DLS for its propensity to form different oligomeric states in a concentration dependent manner. To observe changes in oligomeric state by SEC we appended the coiled-coil sequence to MBP in order to provide a substantive difference in SEC retention volume and hydrodynamic radius. An additional coiled-coil (TRI) was used which was based on the coiled-coil from PDB 1COI.[2] Throughout SI-Fig. S2 the SEC and DLS for MBP (shades of grey) and MBP-TRI (black) are shown as controls for monomeric and trimerized MBP to compare to the MBP-CC2/a constructs. By SEC CC2a prefers dimeric association at 3 mg/ml, however above this concentration it prefers to associate as a trimer, at all concentrations there are populations of monomers, dimers and trimers. The SEC observations for CC2a were reciprocated by the DLS. CC2 was not seen to associate at the concentrations tested.



Supplementary Figure S2: SEC and DLS investigation of CC2 (dashed lines) with its four-heptad counterpart, CC2a (solid lines). a) MBP and MBP-TRI tested at various concentrations showing monomeric and trimeric retention volumes. **b)** The sequences for CC2, CC2a and TRI. **c and d)** SEC and DLS, respectively, of CC2a, showing a concentration dependent shift from dimer to trimer with increasing concentration.

3. CTB-CC2a experimental characterisation

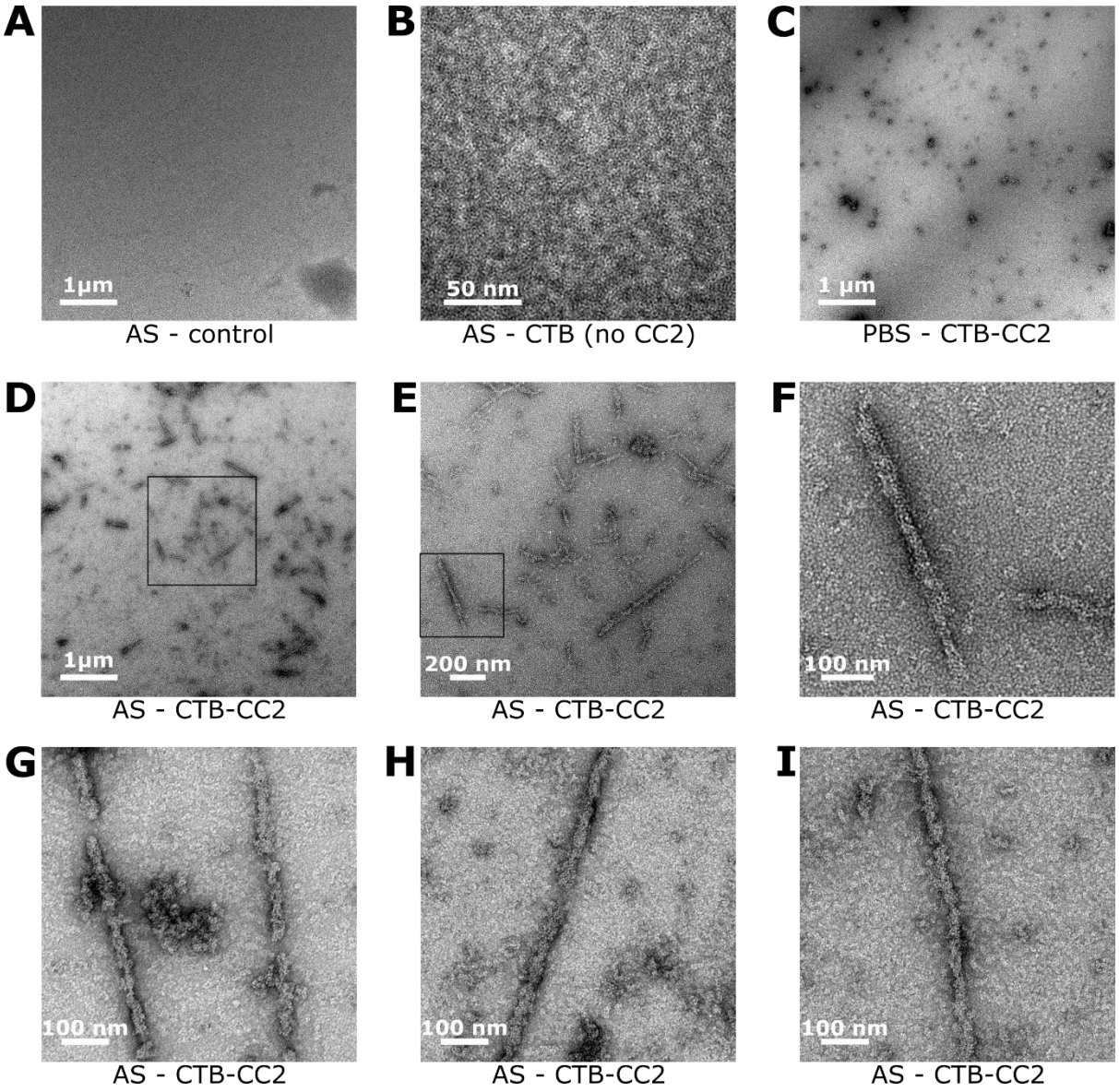
Given the success of the four-heptad coiled-coil to assemble MBP into higher-order structures (SI-Fig. S2), we tested the sequence on CTB. We found that, given two-day incubations of CTB-CC2a, a shoulder on the side of the pentameric peak increased in size. Two CTB pentamers associated via an interaction with five coiled-coils is a species we specifically desired not to encounter in our studies, as this species would likely be more energetically stable than more complicated supra-molecular complexes. Given the existence of this shoulder, we decided to continue with the three-heptad version of the coiled-coil rather than risk poisoning our desired assemblies with dimers of pentamers.



Supplementary Figure S3: Size exclusion chromatography of construct CTB-CC2a, a four-heptad coiled-coil extension. a) Size exclusion chromatography of CTB-CC2a displaying a similar trace as seen with CTB-CC2 (Fig. 2) however, CTB-CC2a shows a noticeable shoulder to the pentameric peak which evidently increases with time, suggestive of the formation of a dimer of pentamers. **b)** an SDS PAGE of the CTB-CC2a construct compared to CTB.

4. TEM Micrographs of CTB-CC2 Tubular Assemblies

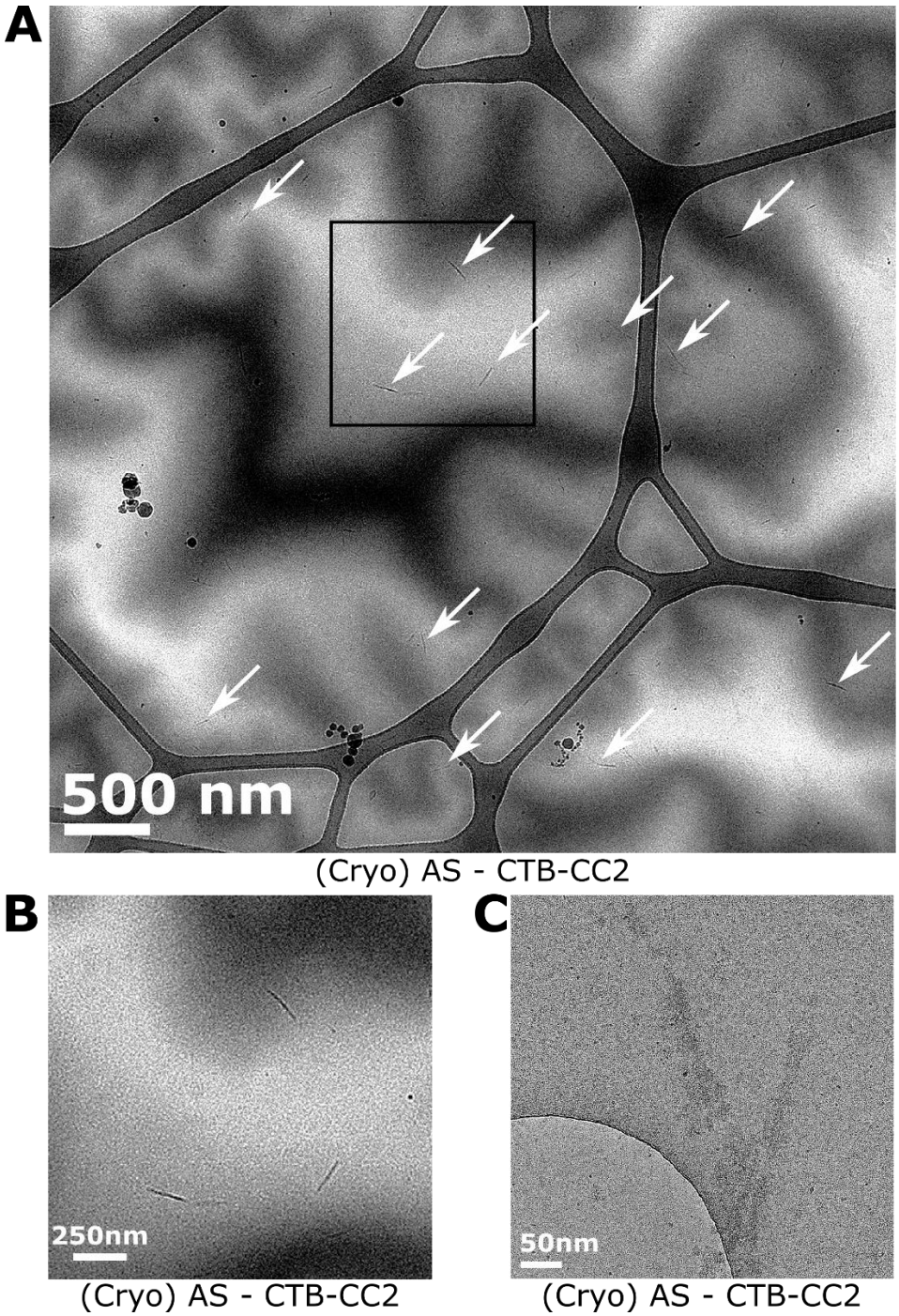
Here we present further TEM images to support those seen in Fig. 2. Controls of the assembly buffer, native CTB in assembly buffer and CTB-CC2 in PBS. We additionally show some 'field of view' micrographs demonstrate the prevalence of the tubular structures as well as some additional extended tubes.



Supplementary Figure S4. Additional micrographs of CTB-CC2 tubular assemblies. Ammonium sulfate (AS) concentrations at 1M. **A)** AS control, no features. **B)** 58 µM CTB (without CC2 fusions) in AS, pentamers evident in the background. **C)** 58 µM CTB-CC2 in PBS, amorphous small aggregates visible. **D-F)** 290 µM CTB-CC2 in AS, step-wise magnification of a representative section of a micrograph. **G-I)** Examples of tubular assemblies from addition sample preparations at 58 µM CTB-CC2. CTB concentrations refer to protomer concentrations (i.e. 5 × pentamer concentration)

5. Cryo-TEM Micrographs of CTB-CC2 Tubular Assemblies

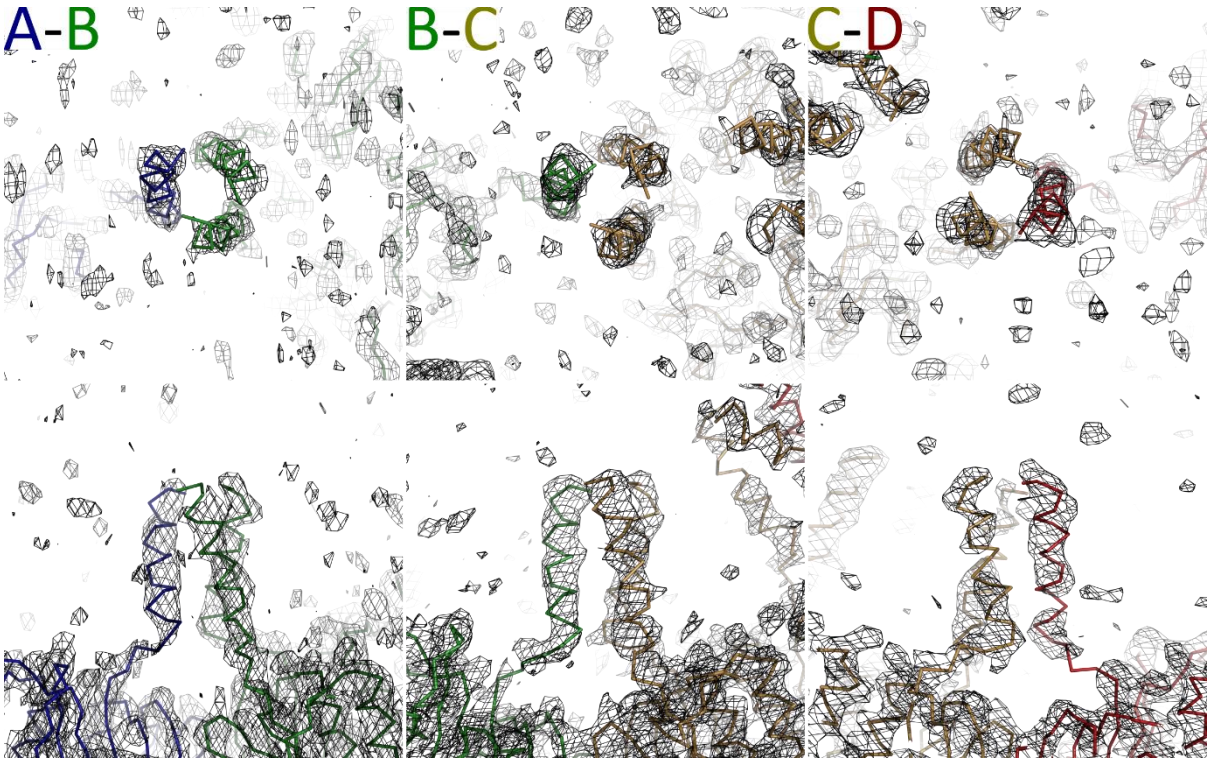
The prevalence of the tubes structures in solution was studied in vitreous ice by cryo-electron microscopy. The protein assemblies have a prevalence for interacting with the carbon support of the microscopy grids (Fig S5c). However, structures of a similar length and diameter to those seen under negative stain TEM, could be observed by cryo-EM when lacey carbon grids with a continuous 2 nm carbon film were used (Fig S5a-b).



Supplementary Figure S5. Cryo-EM micrographs of CTB-CC2 tubular assemblies. Samples prepared from 58 µM CTB-CC2 in 1.25 M ammonium sulfate (AS). **A-B)** Lacey carbon with continuous 2 nm carbon film. The high ammonium sulfate concentration leads to dark areas of the vitreous ice, reducing contrast. **C)** Quantifoil R1.2/1.3 grids – tubular structures show affinity for carbon support.

6. Coiled-coil electron density at 6 Å resolution

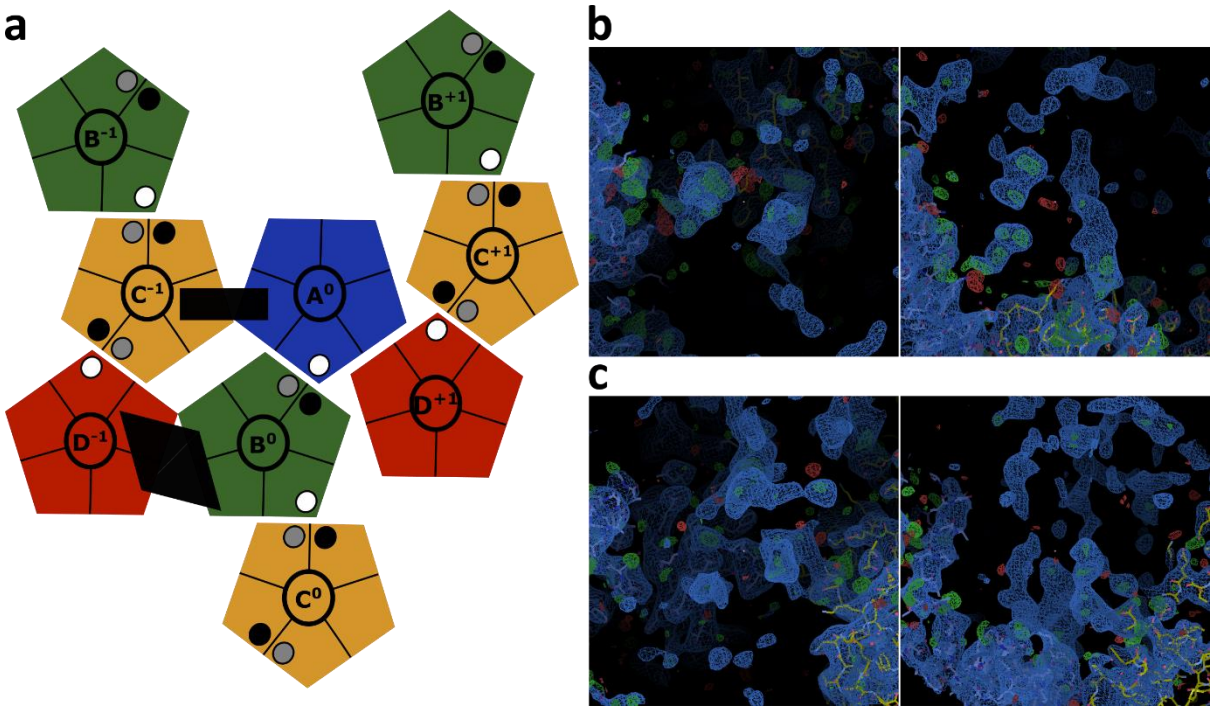
Upon initial molecular replacement of the CTB pentamer (3CHB.pdb) into the CTB-CC2 map, and observation of the tubular configuration of the assembly, it was clear that the coiled-coil structures were present in the density. These were modelled as part of the continuing refinement process. However, we were interested, once good metrics for the fit were achieved, how prominent the coiled-coil density was above the background density of the map. We submitted a refined structure of the asymmetric unit to REFMAC while restricting the resolution to minimum of 6 Å and studied the resultant electron density map, SI-Fig. S6. For each of the coiled-coil sections electron density is clearly visible, however the coiled-coil at the C-D interface is a little less well defined than at A-B and B-C. It should be noted that further potential alpha-helical density was observed, shown in SI-Fig. S7, however this density was not sufficient to reliably model.



Supplementary Figure S6. Observed electron density at 6 Å resolution. Once a final structure of the asymmetric unit had been modelled, a map was created with a resolution of 6 Å. The images above show the density of the three-pronged coiled-coils clearly apparent above the background density in the map.

7. Possible additional coils

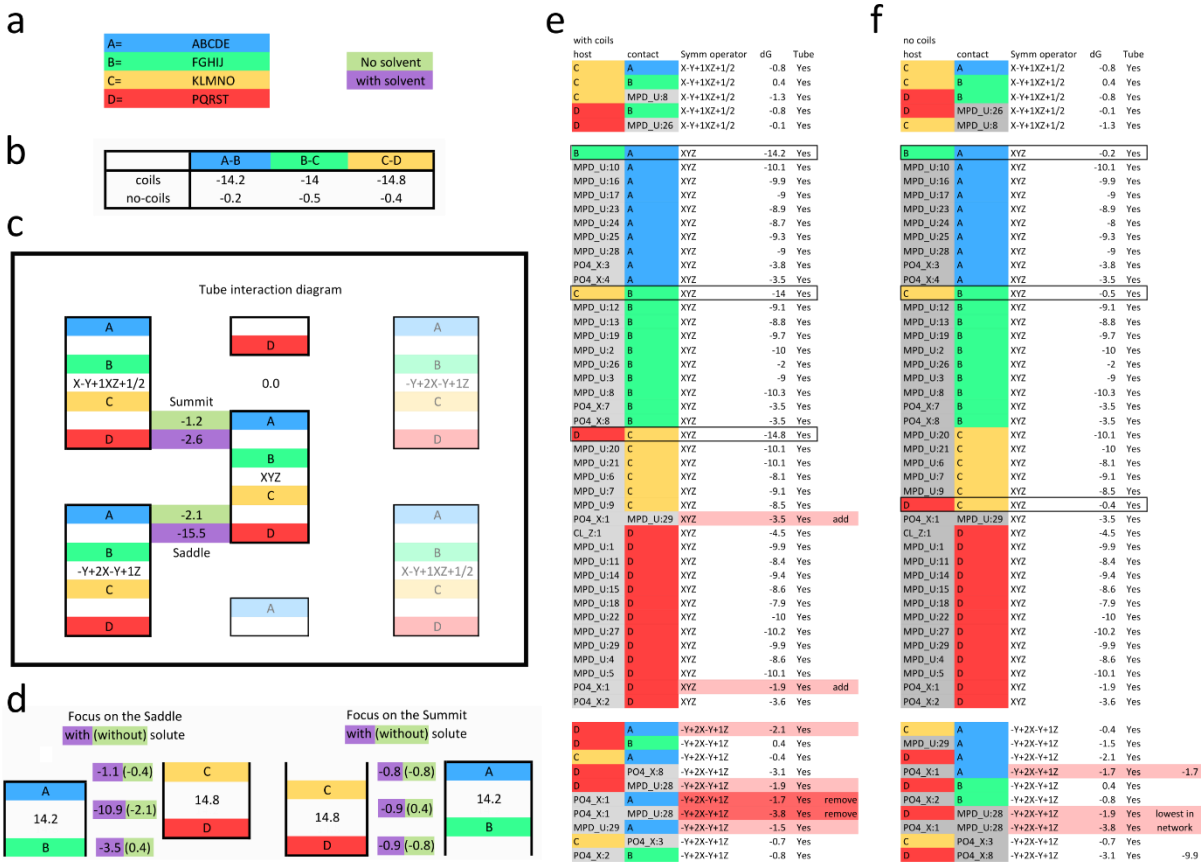
Only nine of a possible twenty helices were modelled in the crystal structure, leaving eleven possible additional helices. The interface between $A^0B^0/C^{+1}D^{+1}$ in the saddle conformation (right-hand side of Fig. S7a) is of interest, given the modelled coiled-coils and the orientation of the pentamers, it is not possible that unmodeled coiled-coils can bridge this interface. For instance, on pentamer C^{+1} there are no free helices in proximity to interact with A^0. Likewise, on B^0 there are no free helices within proximity of D^{+1} to interact. The schematic below may suggest that a possible dimeric coiled-coil may form between a free helix of A^0 and D^{+1}, but due to the -62° interaction angle of the A^0/D^{+1} pentamers, these helices are too distant to interact. However, this is not the case for the $C^{-1}D^{-1}/A^0D^0$ interface (left-hand side of Fig. S7a) which has a number of free helices within interacting distances. The interface between C-1 and A0 has density reminiscent of a dimeric coiled-coil, there is little evidence of a third helix (Fig. S7b). At the interface between D-1 and B0 there is a substantial amount of noise that could represent a number of dimeric or trimeric coiled-coils (Fig. S7c). As discussed in the main text, modelling of these densities was not attempted. Given that these coils are present, it would suggest the proposed assembly would initially form from summit architectures rather than saddle architectures.



Supplementary Figure S7: Potential extra helices. a) From a purely schematic perspective, no coiled-coil interactions can take place across the $A^0B^0/D^{+1}C^{+1}$ interface, as no free helices are present to interact. However, across the $C^{-1}D^{-1}/A^0B^0$ interface there are free helices within a proximity to form coiled-coils. The electron density was inspected at sigma 0.7 and coot map sharpening at 81.0. **b)** Evidence of helical-like density is apparent at the C^{-1}/A^0 interface which may form a dimeric coiled-coil. **c)** Evidence of helical like structures are apparent at the D^{-1}/B^0 interface, although this density is very noisy, it may represent a number of dimeric or trimeric helices in multiple conformations.

8. Subunit interaction analysis with PISA

The interaction free energy ΔG between the components of the crystal structure were estimated with the CCP4 program PISA. Of interest was the interaction between pentameric subunits in the presence or absence of coiled-coils. First the side chains of the coiled-coil residues were rebuilt using the fixbb routine of Rosetta, to represent a more realistic calculation of the interaction energies between coiled-coils. All interactions between pentamers in the absence of coiled-coils (either absent from the density or removed prior to analysis) showed that there was little interaction energy between the pentamers, except for the A-D interaction in the 'saddle' architecture, which is mediated by solute molecules. When a model with the coiled-coils was analysed a ΔG interaction of ~-14 kcal/mol was calculated between coiled-coil associated pentamers.



Supplementary Figure S8: Interaction analysis with PISA. All values are ΔG interaction energy estimates in kcal/mol. **a)** Key: For convenience we label each pentamer A-D, though in the crystal structure, each monomer chain is individually named. **b)** The ΔG interaction between the subunits in the asymmetric unit, either with or without modelled coiled-coils. **c)** The ΔG interaction between asymmetric units in the tube structure, either with (purple) or without (olive) solute molecules. **d)** Focused analysis of the interactions within the saddle and summit architectures, either with (purple) or without (olive) solute molecules. **e and f)** The ΔG interaction energies estimated by PISA either with coiled-coil (**e**) or without coiled-coils (**f**), note that some highlighted values form networked interactions for which a minimum network interaction energy is calculated to represent the interaction energy, as described in detail in Fig. S10.

9. Geometric relationships between pentamers

To help understand the geometric relationship between the pentamers in the tube structures we defined a normal vector for each of the pentamers and calculated the angular relationship between planes perpendicular to these vectors. The planar angles are both positive and negative, but with respect of the tube orientation (a positive angle following the curvature of the tube), it is worth pointing out that the angles for the A^0/D^{+1} planes should be thought of as negative, with respect to the tube curvature. Additionally, it is worth noting that the B^0/C^{-1} planes have angles approaching 0 and 180 suggesting that the planes are almost parallel.

	Coord - 1			Coord - 2		
Subunit	**x**	**y**	**z**	**x**	**y**	**z**
A0	27.85	207.96	36.50	42.75	195.36	25.14
B0	7.19	144.65	40.67	29.70	144.15	38.94
C0	33.73	93.65	77.79	46.24	112.39	79.18
D0	97.10	73.86	81.39	95.88	93.17	93.09
A+1	12.99	128.09	132.51	31.35	134.70	121.15
B+1	57.49	78.55	136.68	69.17	97.80	134.95
C+1	64.24	234.28	77.79	74.21	214.07	79.18
D+1	15.41	189.29	81.39	32.75	180.70	93.09
A-1	78.18	94.50	25.14	74.72	75.29	36.50
B-1	139.87	89.06	40.67	129.05	108.80	38.94
C-1	8.39	172.78	-18.22	30.87	171.30	-16.83
D-1	22.93	107.99	-14.62	39.04	118.71	-2.92

These coordinates place a normal vector along the 5-fold symmetry axis of each pentamer
That is a vector which is perpendicular to the plane of the subunit
these are used to calculate angle between the planes of the subunits

Subunit Interaction	Planar Angle1	Planar Angle2	Dihedral	1nva	2nva	distance
A0 - C+1	40.54	139.46	40.50	100.80	85.30	61.00
A0 - D+1	62.77	117.23	41.20	114.00	114.50	56.20
A0 - C-1	48.49	131.51	7.90	59.60	72.50	67.90
A0 - B0	44.88	135.12	29.60	72.90	73.00	66.70
B0 - D+1	42.86	137.14	42.80	86.10	99.80	61.00
B0 - C-1	8.29	171.71	1.20	85.50	86.30	65.30
B0 - D-1	48.68	131.32	5.80	72.30	59.30	68.20
B0 - C0	58.01	121.99	40.00	68.80	68.30	68.40
C0 - D0	44.73	135.27	31.00	73.50	73.90	66.50

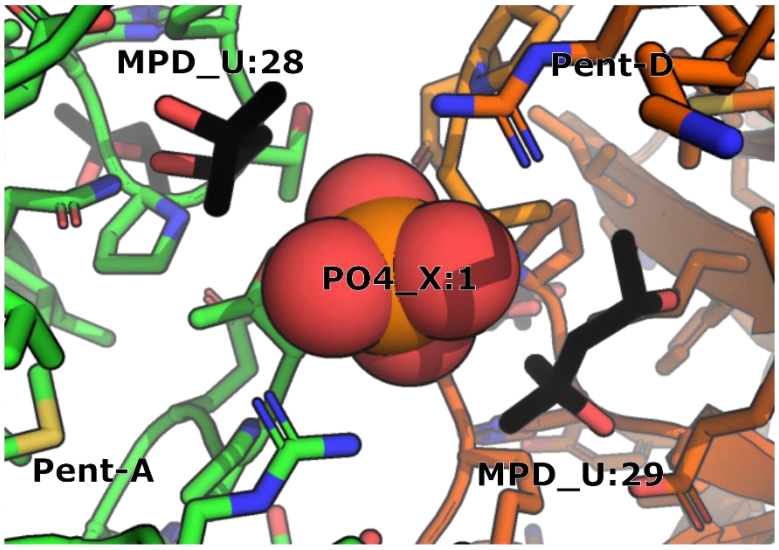
Planar Angle1	Intersecting planes have two angles, each of which are both positive and negative.
Planar Angle2	Intersecting planes have two angles, each of which are both positive and negative.
Dihedral	Twist between subunits
1nva	Angle between 1st subunit normal vector, centre of mass and 2nd subunit centre of mass
2nva	Angle between 2nd subunit normal vector, centre of mass and 1st subunit centre of mass
distance	distance between subunits

Supplementary Figure S9: Analysis of the geometry of the pentamer interactions. Top, a list of coordinates from which normal vectors for subunits of the asymmetric unit and two further interacting asymmetric units, covering the saddle and summit architectures can be calculated as $\vec{n}$=(coord$_2$-coord$_1$)/|coord$_2$-coord$_1$|. Second, a list of the angular intersections of planes perpendicular to these normal vectors between interacting pentamers in the tube. Of note are the angles of intersection for the A^0/D^{+1} pentamers, which can be thought to be negative with respect to the other angles. Also, the B^0/C^{-1} planar intersections reveal the pentamers are almost parallel.

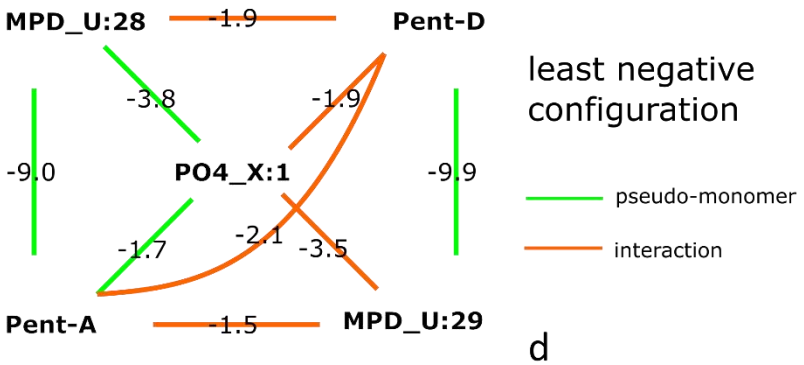
10. Interaction analysis of A0 to D+1 interface

The ΔG interaction energy calculated by PISA between the multi-component interface of the A^0 to D^{+1} interaction of the saddle architecture required further analysis. As an interaction energy is calculated only between two components, to estimate a multi-component interface, two pseudo-monomers must be defined. Here we define the interaction as the interface with the highest (least negative) ΔG, given that one pseudo-monomer contains pentamer A and the other pentamer D, where the two MDP molecules and PO_4 are 'mobile elements'. The least negative grouping of the components puts the PO_4 and MDP_28 with pentamer A and MDP_29 with pentamer D resulting in an interaction energy of -10.9 kcal/mol for the interface.

a



b



c

	Pent-A	MPD_U:28	PO4_X:1	MPD_U:29	Pent-D
Pent-A	0	-9	-1.7	-1.5	-2.1
MPD_U:28		0	-3.8	0	-1.9
PO4_X:1			0	-3.5	-1.9
MPD_U:29				0	-9.9
Pent-D					0

d

Pent-A	Pent-A	Pent-A	Pent-A
MPD_U:28	MPD_U:28	MPD_U:28	-14.3
PO4_X:1	PO4_X:1	-20	
MPD_U:29	-10.9		MPD_U:28
-15.8		PO4_X:1	PO4_X:1
	MPD_U:29	MPD_U:29	MPD_U:29
Pent-D	Pent-D	Pent-D	Pent-D

Supplementary Figure S10: The A^0 to D^{+1} interface in the saddle architecture. a) A close in view of the A^0/D^{+1} interface, A^0 in green, D^{+1} in orange, MDP's in black and PO_4 in spheres. **b)** The interaction energies between all components of the interface, the least negative interface is highlighted in orange lines and the pseudo-monomers with green lines. **c)** The same interaction energies in a chart. **d)** The sum of four possible interactions given different combinations of the mobile elements. The least negative interaction highlighted in yellow.

11. Table of Crystallographic Statistics

PDB Code	6HSV
Beamline	I24, Diamond
Wavelength (Å)	0.97
Space group	*P63*
Unit-cell parameters	
a (Å)	179.2
b (Å)	179.2
c (Å)	192.0
Resolution (Å)	37.28-2.45 (2.51-2.45)
R_{merge} (%)§*	9.2 (69.8)
$R_{p.i.m.}$(%)+*	3.0 (25.6)
CC ½	0.999 (0.941)
Observed reflections	236441 (12331)
Unique reflections	105712 (5468)
Completeness (%)*	88.8 (86.4)
Multiplicity*	10.2 (8.4)
$<I/\sigma(I)>$*	19.0 (3.2)
Refinement	
R_{work} (%)	19.54 (26.50)
R_{free} (%)†	21.37 (28.90)
No. of protein atoms	1722
No. of solvent molecules	584
No. of ligand atoms	246
No. of ion atoms	1
Average *B* factor (Å^2)	
Protein (Å^2)	63.06
Ligands (Å^2)	77.19
Solvent (Å^2)	46.07
Ions (Å^2)	32.6
R.m.s.d., bond lengthsξ (Å)	00187
R.m.s.d., bond anglesξ (°)	1.9536
Ramachandran plot‡	
Most favoured regions (%)	97.13
Outliers (%)	0.23

*Values given in parentheses correspond to those in the outermost shell of the resolution range.

§ $R_{merge} = \sum_{hkl} \sum_i \left| I_i(hkl) - \langle I(hkl) \rangle \right| / \sum_{hkl} \sum I_i(hkl)$

+R_{pim} - precision-indicating (multiplicity-weighted) R_{merge}, relative to I^+ or I^-.

†R_{free} was calculated with 5% of the reflections set aside randomly.

ξBased on the ideal geometry values of Engh & Huber (1991).

‡Ramachandran analysis using the program MolProbity (Chen *et al*., 2010).

12. Nucleotide and amino acid sequences for CTB-CC2 and CTB-CC2a

Generating the CTB-CC genes

Assembly PCR was used with the following primers to generate the final CTB-CC genes

```
CTB-CC-FT CTGTTCAGGCGCATG                    This is the forward terminal primer
CTB-CC-RT CTAAGTTTTCCCTGCAGGTAATTAAATAAGC    This is the reverse terminal primer


CTB-CC2-p1 CTGTTCAGGCGCATGCAACTCCTCAAAATATTACTGATTTGTGCGCAGAATACCACAACAC
CTB-CC2-p2 CGCAGAATACCACAACACACAAATATATACGCTAAATGATAAGATCTTTTCGTATACAGAATCGCTAGCGGGAAAAAGAG
CTB-CC2-p3 GCTAGCGGGAAAAAGAGAGATGGCTATCATTACTTTTAAGAATGGTGCAATTTTTCAAGTAGAGGTACCAGGTAGTC
CTB-CC2-p4 CAAGTAGAGGTACCAGGTAGTCAACATATAGATTCACAAAAAAAAAGCGATTGAAAGGATGAAGGATACCCTGAGGATTGC
CTB-CC2-p5 GATACCCTGAGGATTGCATATCTTACTGAAGCTAAAGTCGAAAAGTTATGTGTATGGAATAATAAA
CTB-CC2-p6 TTATGTGTATGGAATAATAAAACGCCTCATGCGATCGCCGCAATTAGTATGGCAAACTAAGTTTTCCCTGCAGGTAATTCG
CTB-CC2-p7 GCGATCGCCGCAATTAGTATGGCAAACGGTGTGTCGGCCCTGGAAAAAGAGGTATCAGCTTTGAAGGAGAAAGTCTCC
CTB-CC2-p8 GCTTTGAAGGAGAAAGTCTCCGCACTCGAGTTCTAAGTTTTCCCTGCAGG
CTB-CC2a-p8 GCTTTGAAGGAGAAAGTCTCCGCACTCGAGAAAGAGGTATCAGCTTTGAAGTTCTAAGTTTTCCCTGCAGG
```

These assembly PCRs generated the following sequences, which were restriction cloned into pSAB2.2 via SphI and PstI restriction sites (underlined, respectively).

```
>pSAB2.2 (Lac-operon controlled expression plasmid derived from pMAL-p5x following removal of
the malE gene and introduction of an LTIIb leader sequence and El Tor CTB gene)
ATGAGCTTTAAGAAAATTATCAAGGCATTTGTTATCATGGCTGCTTTGGTATCTGTTCAGGCGCATGCAGCTCCTCAAAATATTACTGATTTGT
GCGCAGAATACCACAACACACAAATATATACGCTAAATGATAAGATCTTTTCGTATACAGAATCGCTAGCGGGAAAAAGAGAGATGGCTATCAT
TACTTTTAAGAATGGTGCAATTTTTCAAGTAGAGGTACCAGGTAGTCAACATATAGATTCACAAAAAAAAAGCGATTGAAAGGATGAAGGATACC
CTGAGGATTGCATATCTTACTGAAGCTAAAGTCGAAAAGTTATGTGTATGGAATAATAAAACGCCTCATGCGATCGCCGCAATTAGTATGGCAA
ACTAAGTTTTCCCTGCAG
>CTB-CC2
CTGTTCAGGCGCATGCAACTCCTCAAAATATTACTGATTTGTGCGCAGAATACCACAACACACAAATATATACGCTAAATGATAAGATCTTTTC
GTATACAGAATCGCTAGCGGGAAAAAGAGAGATGGCTATCATTACTTTTAAGAATGGTGCAATTTTTCAAGTAGAGGTACCAGGTAGTCAACAT
ATAGATTCACAAAAAAAAAGCGATTGAAAGGATGAAGGATACCCTGAGGATTGCATATCTTACTGAAGCTAAAGTCGAAAAGTTATGTGTATGGA
ATAATAAAACGCCTCATGCGATCGCCGCAATTAGTATGGCAAACGGTGTGTCGGCCCTGGAAAAAGAGGTATCAGCTTTGAAGGAGAAAGTCTC
CGCACTCGAGTTCTAAGTTTTCCCTGCAGG
>CTB-CC2a
CTGTTCAGGCGCATGCAACTCCTCAAAATATTACTGATTTGTGCGCAGAATACCACAACACACAAATATATACGCTAAATGATAAGATCTTTTC
GTATACAGAATCGCTAGCGGGAAAAAGAGAGATGGCTATCATTACTTTTAAGAATGGTGCAATTTTTCAAGTAGAGGTACCAGGTAGTCAACAT
ATAGATTCACAAAAAAAAAGCGATTGAAAGGATGAAGGATACCCTGAGGATTGCATATCTTACTGAAGCTAAAGTCGAAAAGTTATGTGTATGGA
ATAATAAAACGCCTCATGCGATCGCCGCAATTAGTATGGCAAACGGTGTGTCGGCCCTGGAAAAAGAGGTATCAGCTTTGAAGGAGAAAGTCTC
CGCACTCGAGAAAGAGGTATCAGCTTTGAAGTTCTAAGTTTTCCCTGCAGG
```

Expression of the pSAB2.2_CTB-CC1/2 yielded the following secreted protein sequence.

```
>CTB
TPQNITDLCAEYHNTQIYTLNDKIFSYTESLAGKREMAIITFKNGAIFQVEVPGSQHIDSQKKAIERMKDTLRIAYLTEAKVEKLCVWNNKTPH
AIAAISMAN
>CTB-CC2
TPQNITDLCAEYHNTQIYTLNDKIFSYTESLAGKREMAIITFKNGAIFQVEVPGSQHIDSQKKAIERMKDTLRIAYLTEAKVEKLCVWNNKTPH
AIAAISMANGVSALEKEVSALKEKVSALEF
>CTB-CC2a
TPQNITDLCAEYHNTQIYTLNDKIFSYTESLAGKREMAIITFKNGAIFQVEVPGSQHIDSQKKAIERMKDTLRIAYLTEAKVEKLCVWNNKTPH
AIAAISMANGVSALEKEVSALKEKVSALEKEVSALKF
```

13. Nucleotide and amino acid sequences for MBP-Coiled-coils

Generating the MBP-CC2, MBP-CC2a, MBP-CCTri genes

Assembly PCR was used with the following primers to generate the MBP-CC genes

```
MBP-FT CACATGGGATCCGAGAAC                This is the forward terminal primer
MBP-RT GCTTACCTGCAGAACTTAG               This is the reverse terminal primer

MBP-CC-p1 CACATGGGATCCGAGAACCTGTACTTTCAGGGTAATAACGG
MBP-CC2/2a-p2 GAGACTTTCTCCTTCAAAGCTGATACCTCTTTTTCCAGGGCCGACACACCGTTATTACCCTGAAAGTAC
MBP-CC2-p3 GCTTACCTGCAGAACTTAGAACTCGAGTGCGGAGACTTTCTCCTTCAAAGC
MBP-CC2a-p3 GCTTACCTGCAGAACTTAGAACTTCAAAGCTGATACCTCTTTCTCGAGTGCGGAGACTTTCTCCTTCAAAGC
MBP-CCTri-p2 GACTCTAAAGCGGCGACCTTCTTCTCAAGAGCCTCAACCCTCCCGTTATTACCCTGAAAGTAC
MBP-CCTri-p3 GCTTACCTGCAGAACTTAGCCATGTTCGAGTGCCTCAACCTTCTTCTCCAGGGCCTGAACTTTAGACTCTAAAGCGGCGAC
```

This assembly PCR generated the following sequence, which was restriction cloned into pMAL-c5x via BamHI and PstI restriction sites (underlined).

```
>MBP-CC2
CACATGGGATCCGAGAACCTGTACTTTCAGGGTAATAACGGTGTGTCGGCCCTGGAAAAAGAGGTATCAGCTTTGAAGGAGAAAGTCTCCGCAC
TCGAGTTCTAAGTTCTGCAGGTAAGC
>MBP-CC2a
CACATGGGATCCGAGAACCTGTACTTTCAGGGTAATAACGGTGTGTCGGCCCTGGAAAAAGAGGTATCAGCTTTGAAGGAGAAAGTCTCCGCAC
TCGAGAAAGAGGTATCAGCTTTGAAGTTCTAAGTTCTGCAGGTAAGC
>MBP-CCTri
CACATGGGATCCGAGAACCTGTACTTTCAGGGTAATAACGGGAGGGTTGAGGCTCTTGAGAAGAAGGTCGCCGCTTTAGAGTCTAAAGTTCAGG
CCCTGGAGAAGAAGGTTGAGGCACTCGAACATGGCTAAGTTCTGCAGGTAAGC
```

Expression of the pMAL-c5x_MBP-CC yielded the following cytosolic protein sequence.

```
>MBP-CC2
MFSASALAKIEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHPDKLEEKFPQVAATGDGPDIIFWAHDRFGGYAQSGLLAEITPDKAFQ
DKLYPFTWDAVRYNGKLIAYPIAVEALSLIYNKDLLPNPPKTWEEIPALDKELKAKGKSALMFNLQEPYFTWPLIAADGGYAFKYENGKYDIKD
VGVDNAGAKAGLTFLVDLIKNKHMNADTDYSIAEAAFNKGETAMTINGPWAWSNIDTSKVNYGVTVLPTFKGQPSKPFVGVLSAGINAASPNKE
LAKEFLENYLLTDEGLEAVNKDKPLGAVALKSYEEELVKDPRIAATMENAQKGEIMPNIPQMSAFWYAVRTAVINAASGRQTVDEALKDAQTNS
SSNNNNNNNNNNLGIEGRISHMGHMGSENLYFQGNNGVSALEKEVSALKEKVSALEF
>MBP-CC2a
MFSASALAKIEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHPDKLEEKFPQVAATGDGPDIIFWAHDRFGGYAQSGLLAEITPDKAFQ
DKLYPFTWDAVRYNGKLIAYPIAVEALSLIYNKDLLPNPPKTWEEIPALDKELKAKGKSALMFNLQEPYFTWPLIAADGGYAFKYENGKYDIKD
VGVDNAGAKAGLTFLVDLIKNKHMNADTDYSIAEAAFNKGETAMTINGPWAWSNIDTSKVNYGVTVLPTFKGQPSKPFVGVLSAGINAASPNKE
LAKEFLENYLLTDEGLEAVNKDKPLGAVALKSYEEELVKDPRIAATMENAQKGEIMPNIPQMSAFWYAVRTAVINAASGRQTVDEALKDAQTNS
SSNNNNNNNNNNLGIEGRISHMGHMGSENLYFQGNNGVSALEKEVSALKEKVSALEKEVSALKF
>MBP-CCTri
MFSASALAKIEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHPDKLEEKFPQVAATGDGPDIIFWAHDRFGGYAQSGLLAEITPDKAFQ
DKLYPFTWDAVRYNGKLIAYPIAVEALSLIYNKDLLPNPPKTWEEIPALDKELKAKGKSALMFNLQEPYFTWPLIAADGGYAFKYENGKYDIKD
VGVDNAGAKAGLTFLVDLIKNKHMNADTDYSIAEAAFNKGETAMTINGPWAWSNIDTSKVNYGVTVLPTFKGQPSKPFVGVLSAGINAASPNKE
LAKEFLENYLLTDEGLEAVNKDKPLGAVALKSYEEELVKDPRIAATMENAQKGEIMPNIPQMSAFWYAVRTAVINAASGRQTVDEALKDAQTNS
SSNNNNNNNNNNLGIEGRISHMGHHMGSENLYFQGNNGRVEALEKKVAALESKVQALEKKVEALEHG
```

14. References

(1) De Crescenzo, G.; Litowski, J. R.; Hodges, R. S.; O'Connor-McCourt, M. D. *Biochem.* **2003**, *42*, 1754.
(2) Ogihara, N. L.; Weiss, M. S.; Eisenberg, D.; Degrado, W. F. *Protein Sci.* **1997**, *6*, 80.