Electronic Supplementary Material (ESI) for ChemComm.

Supplementary Material:

A disulfide polymerized protein crystal

Esben M. Quistgaard

Suppelmentary Methods

Crystallization:

Native BAP29 vDED was expressed and purified as described previously, and selenomethionine (SeMet) labeling was carried out as described for BAP31 vDED[1]. For the dimeric structure, crystals were obtained at ~19 °C by hanging drop vapor diffusion. The crystal used for phasing and refinement was ~135 μm wide and 20−40 μm thick, and was grown in 34% 2-methyl-2,4-pentanediol (MPD), 100 mM Na acetate pH 4.6 and 20 mM $CaCl_2$ using SeMet labeled protein purified in 20 mM Tris pH 8, 150 mM NaCl and 2 mM tris(2-carboxyethyl)phosphine (TCEP) and concentrated to ~13 mg/mL. Streak seeding was performed in the drop, which may be significant, as crystals obtained without seeding were generally considerably smaller. The crystal was soaked in reservoir solution before flash freezing in liquid nitrogen. For the tetrameric form, rod-like crystals were readily obtained at 4°C by hanging or sitting drop vapor diffusion when using ammonium sulfate as the precipitant and Na acetate pH 4.6 as buffer. The crystal used for refinement of the high-resolution data set had dimensions of ~800 μm × 70 μm and grew in 1.8 M ammonium sulfate, 100 mM Na acetate pH 4.6 and 5 mM TCEP using protein purified in 20 mM Tris pH 8, 150 mM NaCl and 0.5 mM TCEP and concentrated to ~14 mg/mL. The crystal was cryoprotected before flash freezing by stepwise changing the composition of the drop to 80% v/v saturated $LiSO_4$ and 100 mM Na acetate pH 4.6. The disulfide polymerized tetramer grew from the same protein batch and in the same crystalant except that TCEP was left out. Notably, crystals grown this way were generally smaller and diffracted more poorly than crystals grown in the presence of TCEP. The same cryo-buffer was used as for the other tetramer crystal.

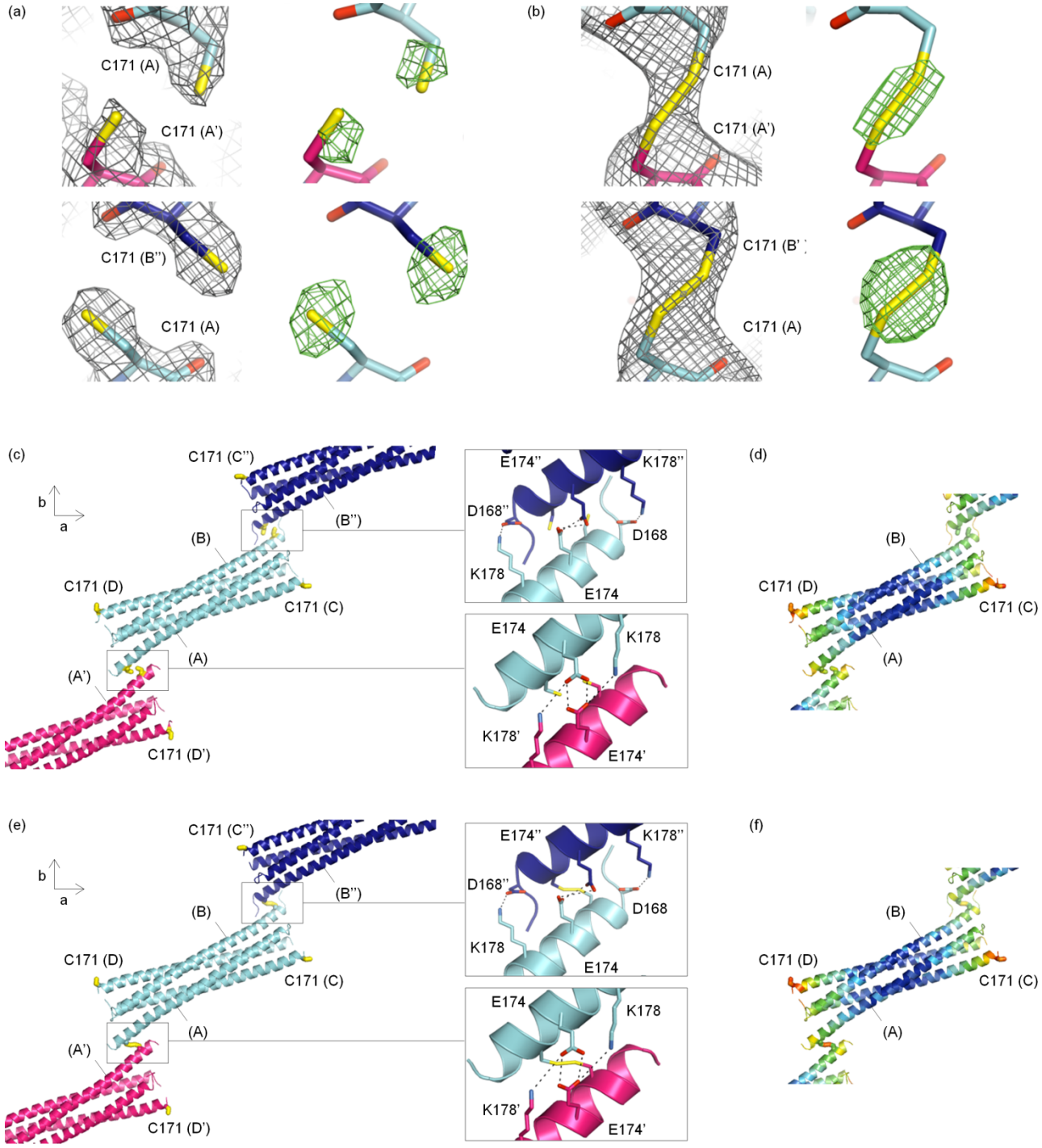
Structure determination and analysis:

All data sets were collected at the Diamond Light Source synchrotron in England (DLS) and processed with XDS and XSCALE[2] (Table 1). SeMet SAD phasing and initial autobuilding were performed using Phenix autosol[3] and the models were then completed by iterative cycles of manual rebuilding in Coot[4] and maximum likelihood refinement in Phenix refine[5] (Table 1). For refinement of the dimeric structure, non-crystallographic symmetry, translation libration screw (TLS) and secondary structure restraints were used. TLS and secondary structure restraints were also used for both tetrameric structures. Finally, reference restraints were employed in the case of the disulfide cross-linked tetramer, using the refined high-resolution tetrameric structure as the reference. All models were validated with Molprobity[6] (Table 1) and structure figures were generated using PyMol[7].

Suppelmentary Figures and Tables

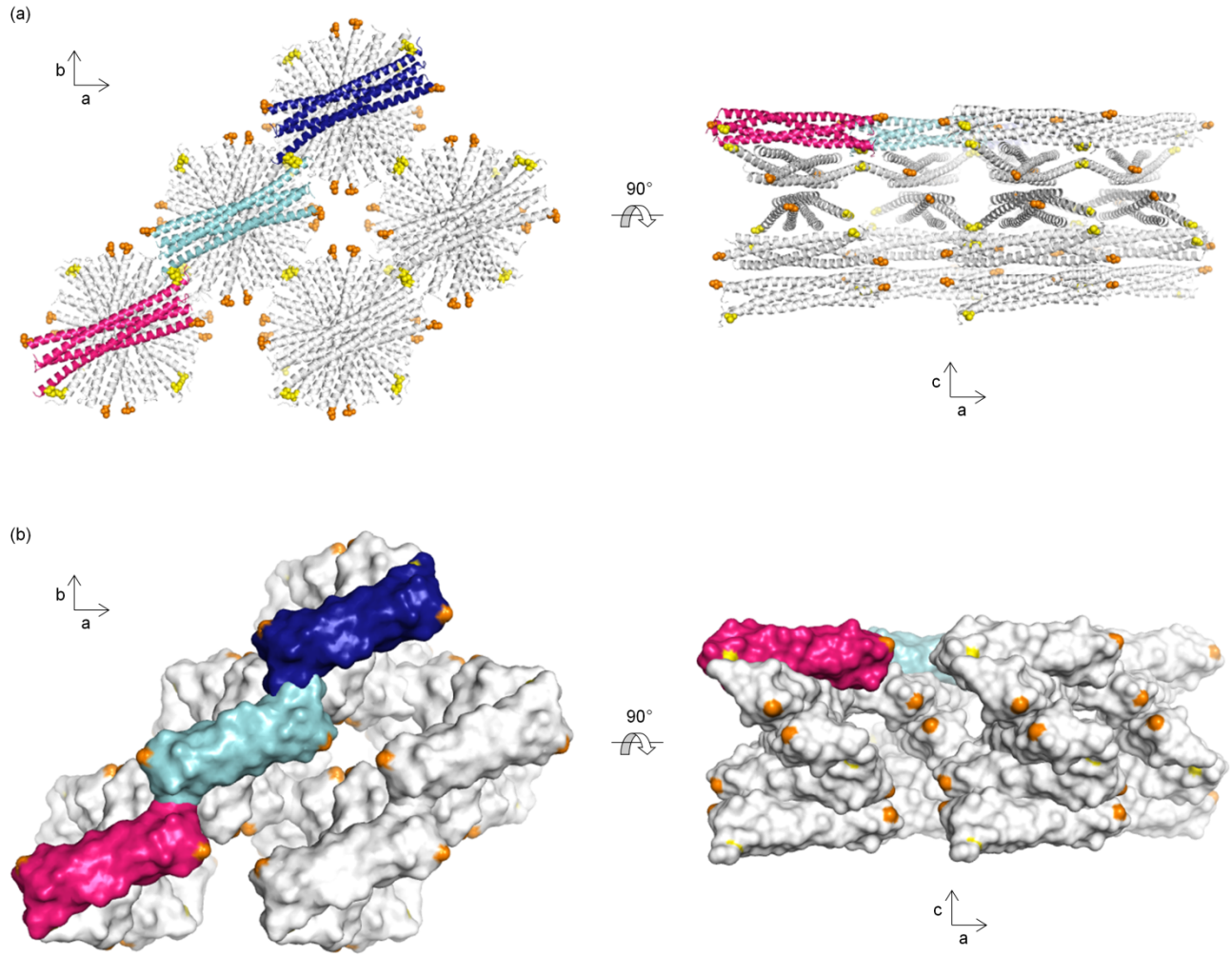
	BAP29 vDED SeMet dimer	BAP29 vDED SeMet tetramer	BAP29 vDED native tetramer, no disulfides	BAP29 vDED native tetramer, with disulfides
Data collection				
Beamline	DLS I02	DLS I02	DLS I02	DLS I04-1
Wavelength (Å)	0.97800	0.97960	0.97949	0.92000
Space group	$P6_5$	$I4_122$	$I4_122$	$I4_122$
Cell dimensions				
a, *b*, *c* (Å)	118.87, 118.87, 30.63	115.54, 115.54, 169.60	115.90, 115.90, 167.20	117.91, 117.91, 168.65
α, β, γ (°)	90, 90, 120	90, 90, 90	90, 90, 90	90, 90, 90
Resolution (Å)	34.31 – 2.50 (2.64 – 2.50)	38.14 – 2.50 (2.57 – 2.50)	29.27 – 2.20 (2.26 – 2.20)	41.69 – 3.20 (3.28 – 3.20)
R_{sym}	0.072 (0.652)	0.070 (0.680)	0.072 (0.568)	0.187 (0.730)
I / σI	17.09 (2.93)	13.02 (1.92)	19.05 (3.87)	12.21 (3.80)
Completeness (%)	99.9 (99.9)	99.6 (98.1)	98.9 (97.1)	96.1 (98.2)
Unique reflections	16720 (2486)	37733 (2784)	28934 (2085)	9752 (712)
Redundancy	6.4 (6.3)	3.9 (3.8)	10.1 (9.7)	13.4 (12.2)
Wilson *B*-factor (Å^2)	49.38	44.43	38.63	52.15
SAD phasing				
No. ordered Se atoms	8	14		
Figure of merit	0.318	0.468		
Refinement				
R_{work} / R_{free}	0.195 / 0.235		0.183 / 0.201	0.199 / 0.215
No. atoms				
Protein	1044		2035	1999
Solvent (water)	31		194	0
Solvent (other)	20		84	50
B-factors				
Protein	73.0		57.8	65.4
Solvent	77.9		69.1	139.0
R.m.s. deviations				
Bond lengths (Å)	0.012		0.008	0.005
Bond angles (°)	1.275		0.976	0.789
Ramachandran plot				
Favored (%)	99.2		100	100
Outliers (%)	0		0	0
Molprobity score				
Clash (percentile)	3.28 (100th)		2.40 (100th)	0.49 (100th)
Overall (percentile)	1.12 (100th)		1.02 (100th)	0.67 (100th)
PDB Accession	4W7Y		4W7Z	4W80

Supplementary Table 1. Data collection and refinement statistics. For the data collection section, the numbers in parentheses refer to the highest resolution shell.



Supplementary Figure 1. Disulfide cross-linking and crystal packing for the tetrameric crystal form. **(a)** C171 in crystal contacts of high-resolution form. Top: C171 of chain A (cyan) and symmetry-related A' (pink) shown with the 2Fo-Fc electron density map at 1 σ (left) and the 1Fo-Fc C171 Sγ omit map at 5 σ (right). Bottom: C171 of chains B (cyan) and symmetry-related B'' (blue). No disulfide bonds are formed. **(b)** C171 in crystal contacts of the disulfide cross-linked form. Shown as for (a). Disulfide cross-links are present here, but due to the modest resolution of the data, the bond geometry cannot be accurately determined and neither can it be ascertained if alternative side chain conformations with minor occupancies

exist. **(c)** Packing of chains A with A’ and B with B’’ in the high-resolution form. Insets show hydrogen bonds and salt bridges. **(d)**. Variation in *B*-factors for the high-resolution form. C171 is colored by *B*-factors and the rest of the protein by Cα *B*-factors. Color scheme is rainbow (blue low, red high). Shown in same orientation as in (c). Note that the lowest *B*-factors are found in the center of the protein and not near the cysteines. **(e)** Packing of chains A with A’ and B with B’’ in the disulfide cross-linked form. Packing is almost identical as in the high-resolution form (c). However, chains A and A’ have moved slightly away from each other and so has chains B and B’’. Consequently, the distance between E174 (A) and E174 (A’) has increased from 2.8 Å to 3.1 Å, E174 (A) and K178 (A’) from 4.1 Å to 4.7 Å, D168 (B) and K178 (B’’) from 2.9 Å to 3.2 Å, E174 Oε1 (B) and E174 Oε1 (B’’) from 2.7 Å to 3.4Å, and E174 Oε2 (B) and E174 Oε2 (B’) from 3.3 Å to 4.2 Å. The weakening of these interactions could be part of the reason for the decreased diffraction potential of the disulfide cross-linked form. **(f)** Variation in *B*-factors for the disulfide cross-linked form. Shown as for the high-resolution form (d). The pattern is almost unaltered by the presence of disulfide bonds, suggesting that these do not appreciably reduce local flexibility in the crystal contacts.



Supplementary Figure 2. Crystal packing of the disulfide cross-linked tetramer. **(a)** The protein assembles into screw-like columns, which are cross-linked to each other by C171 of chains A and B. Protein is grey, except three tetramers, which are colored as in supplementary figure 1. C171 of chains A and B are shown as yellow spheres and the free C171 of chains C and D are orange. Unit cell axes are indicated. **(b)** Same as in (a) but shown in smoothened surface representation to better illustrate the shape of the solvent channels. Screw-shaped columnar solvent channels run along the c-axis all the way through the crystal. They are 30-35 Å wide at the narrowest point. These channels are interconnected by zigzag-like channels, which run parallel to the a- and b-axes and thus perpendicular to the c-axis (left panel). Note that the free C171 of chains C and D are highly exposed and point directly into the solvent channels.