STRUCTURAL BIOLOGY

ISSN 2059-7983

Domain swap in the C-terminal ubiquitin-like domain of human doublecortin

Arne C. Rufer, Eric Kusznir, Dominique Burger, Martine Stihle, Armin Ruf and Markus G. Rudolph*

pRED, Therapeutic Modalities, F. Hoffmann-La Roche, 4070 Basel, Switzerland. *Correspondence e-mail: markus.rudolph@roche.com

Received 5 February 2018
Accepted 23 March 2018

Edited by R. J. Read, University of Cambridge, England

Keywords: domain swap; microtubule-associated protein; doublecortin; analytical ultracentrifugation; lissencephaly; double-cortex syndrome.

PDB reference: C-terminal domain of doublecortin, 6fnz

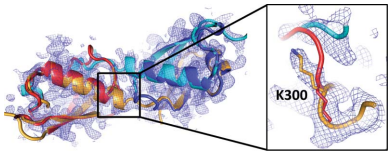
Supporting information: this article has supporting information at journals.iucr.org/d

Doublecortin, a microtubule-associated protein that is only produced during neurogenesis, cooperatively binds to microtubules and stimulates microtubule polymerization and cross-linking by unknown mechanisms. A domain swap is observed in the crystal structure of the C-terminal domain of doublecortin. As determined by analytical ultracentrifugation, an open conformation is also present in solution. At higher concentrations, higher-order oligomers of the domain are formed. The domain swap and additional interfaces observed in the crystal lattice can explain the formation of doublecortin tetramers or multimers, in line with the analytical ultracentrifugation data. Taken together, the domain swap offers a mechanism for the observed cooperative binding of doublecortin to microtubules. Doublecortin-induced cross-linking of microtubules can be explained by the same mechanism. The effect of several mutations leading to lissencephaly and double-cortex syndrome can be traced to the domain swap and the proposed self-association of doublecortin.

1. Introduction

Doublecortin X, or doublecortin for short, is a 40 kDa microtubule-associated protein that is essential for neuronal migration and cortical layering during brain development (Ayanlaja *et al.*, 2017). Binding of doublecortin stabilizes and bundles microtubules, allowing the migration of developing neurons to their designated locations in the brain. Accordingly, mutations in the *DCX* gene that affect doublecortin binding and bundling activities lead to severe neurological diseases including X-linked lissencephaly and double-cortex syndrome. Lissencephaly ('smooth brain') and double-cortex syndrome are rare diseases caused by mutations in, among others, the *DCX* gene (Romero *et al.*, 2018). The impaired neuronal migration during embryonic development results in brains with an abnormally flat and thick cortex. Boys are more severely affected than girls, and clinical manifestations include abnormal muscular tone, seizures and mental retardation to various levels.

At the molecular level, doublecortin contains two ubiquitin-like domains, termed amino-terminal (N-DCX) and carboxy-terminal (C-DCX), which are structurally similar (r.m.s.d. of ~2 Å) but are divergent in sequence (30% identity and 52% similarity over 60 out of 77 residues). In full-length doublecortin, the domains are connected by a linker of ~40 residues. The flexibility of this linker has been inferred from the structures of the individual N-DCX and C-DCX domains. A tryptophan side chain that can bind to N-DCX and thus regulate the distance between the two domains by a



knob-in-hole mechanism is important for the linker (Cierpicki *et al.*, 2006). The tandem arrangement of N-DCX and C-DCX is necessary and sufficient for microtubule binding and bundling. The C-terminal part of doublecortin largely consists of a low-complexity Ser/Pro-rich region that has not been implicated in microtubule binding. Structural information on doublecortin is available for the individual ubiquitin-like domains but not for the full-length protein. The structure of N-DCX has been studied by NMR (PDB entry 1mjd; Kim *et al.*, 2003), X-ray crystallography (PDB entries 2bqq, 5ikc, 5in7, 5io9 and 5ioi; Cierpicki *et al.*, 2006; Burger *et al.*, 2016) and cryo-EM (PDB entries 2xrp and 4atu; Liu *et al.*, 2012; Fourniol *et al.*, 2010), where N-DCX was fitted into electron density on tubulin protofilaments. The only currently available crystal structure of C-DCX was determined in complex with a nanobody (PDB entry 5ip4; Burger *et al.*, 2016), a testimony to the inherent instability of C-DCX. Here, we describe the first crystal structure of free C-DCX, *i.e.* in the absence of a stabilizing nanobody, with an interesting twist that allows the rationalization of some of the biological properties of doublecortin. Our crystallographic and solution studies show that C-DCX may split into an open conformation that can dimerize as a result of a domain swap. Further self-association of C-DCX was observed in solution and the crystal lattice. The experimental evidence is in agreement with previous reports on cooperative interaction with and bundling of microtubules by doublecortin.

2. Materials and methods

2.1. Protein purification, crystallization and data collection

The numbering used here for doublecortin follows that for human doublecortin in UniProt and is shifted by +81 relative to other reports of structures and patient mutations. The C-terminal fragment of human doublecortin encompassing residues 251–351 (C-DCX; molecular mass 10.2 kDa) was cloned using NcoI/NdeI sites into a modified pET-28a(+) vector (Burger *et al.*, 2016) to encode a His_6-SUMO fusion protein cleavable with SUMO protease and thrombin. The plasmid pER2b was transformed into *Escherichia coli* BL21 (DE3) cells and protein production was initiated at an OD_{600} of 0.5 by shifting the culture to 20°C for 20 min, followed by induction with 0.5 m*M* IPTG. The harvested cells were resuspended in 50 m*M* HEPES–NaOH pH 7.5, 300 m*M* NaCl, 2 m*M* TCEP (buffer *A*) and protease inhibitor (Roche complete), and disintegrated using BasicZ cell-disruption equipment at 80 MPa pressure (Constant Systems). The supernatant after centrifugation was passed through a 0.22 µm filter and applied onto an Ni^{2+}–NTA column (HisTrap HP 5 ml, GE Healthcare) equilibrated with buffer *A*. After washing away unbound material, the fusion protein was eluted in a linear 0–300 m*M* imidazole gradient in buffer *A*. Fractions containing His_6-SUMO-C-DCX were pooled and hydrolyzed with SUMO protease. After dialysis at 4°C against 50 m*M* HEPES–NaOH pH 7.5, 100 m*M* NaCl, 10%(*v*/*v*) glycerol, 8 m*M* CHAPS, 2 m*M* TCEP, the hydrolysate was chromatographed again on Ni^{2+}–NTA (as above) equilibrated in buffer *A*. The final step was size-exclusion chromatography of the flowthrough from the Ni^{2+}–NTA column on a Superdex S75 10/300 column (GE Healthcare) equilibrated in 20 m*M* CAPS–NaOH pH 10.5, 100 m*M* NaCl, 5 m*M* TCEP. C-DCX-containing fractions were pooled, concentrated to 5 mg ml^{-1} and stored at −80°C.

Prior to crystallization, C-DCX was concentrated to 20 mg ml^{-1} and 3× the critical micelle concentration (CMC) of the detergent CHAPS was added. Crystals were obtained by mixing 60–70% protein solution with 40–30% reservoir solution consisting of 0.1 *M* HEPES–NaOH pH 7.0, 10% PEG 5000 MME, 5% Tacsimate (100% corresponds to 1.8305 *M* malonic acid, 0.25 *M* ammonium citrate, 0.12 *M* succinic acid, 0.3 *M* DL-malic acid, 0.4 *M* sodium acetate, 0.5 *M* sodium formate, 0.16 *M* ammonium tartrate titrated to pH 7 with NaOH). The calculated pH of the protein/reservoir mixture is ∼7.3. Crystals were cryoprotected with mother liquor containing 25% ethylene glycol and flash-cooled for data collection at 100 K on beamline PXII at the Swiss Light Source (SLS) using a PILATUS 6M detector (440 mm distance, 0.5° oscillation, 0.5 s exposure, 1 Å wavelength, 15% transmission at a flux of 3×10^{11} photons s^{-1}). Data were integrated and scaled in space group *P*422, with unit-cell parameters a = 98.5, c = 114.8 Å, using the *XDS* package (Kabsch, 2010). The likely presence of a fourfold screw axis and of twofold screw axes was established by analysis of the systematically absent reflections [$I/\sigma(I) = 0.7$ for 00l reflections with $l \neq 4n$ and $I/\sigma(I) = 0.7$ for h00 reflections with $h \neq 2n$]. While the native Patterson map was featureless, self-rotation function analysis showed the presence of a twofold ($\kappa = 180°$) noncrystallographic axis at $\omega = 90$, $\varphi = 22.5°$ indicating a dimeric assembly of C-DCX molecules in the crystal (Fig. 1). Assuming 10.2 kDa per monomer, the asymmetric unit may contain 4–6 C-DCX molecules, corresponding to a Matthews coefficient (Matthews, 1968) in the range 3.4–2.3 $Å^3$ Da^{-1} and a solvent content in the range 64–46%, respectively. Data-collection statistics are given in Table 1.

2.2. Phasing and refinement

Molecular replacement using the closely related C-DCX domain extracted from its complex with a nanobody (PDB entry 5ip4) was unsuccessful in either of the space groups ($P4_12_12$ or $P4_32_12$) suggested by the systematic absences. A clear solution of four molecules per asymmetric unit in space group $P4_32_12$ was obtained when the model 5ip4 was trimmed at the termini and surface loops using the *MoRDa* pipeline (Winn *et al.*, 2011). The initial R_{free} for this solution was 39% and the *Q*-factor was 0.77 (Keegan *et al.*, 2011). The next best solution in space group $P4_322$ had a significantly smaller *Q*-factor of 0.48. The electron density after molecular replacement and initial refinement with *REFMAC* (Murshudov *et al.*, 2011) already showed the presence of a domain swap in C-DCX (Fig. 2). The model was rebuilt in *Coot* (Emsley *et al.*, 2010) and refined with *PHENIX* (Adams *et al.*, 2010) using automatically determined TLS domains.

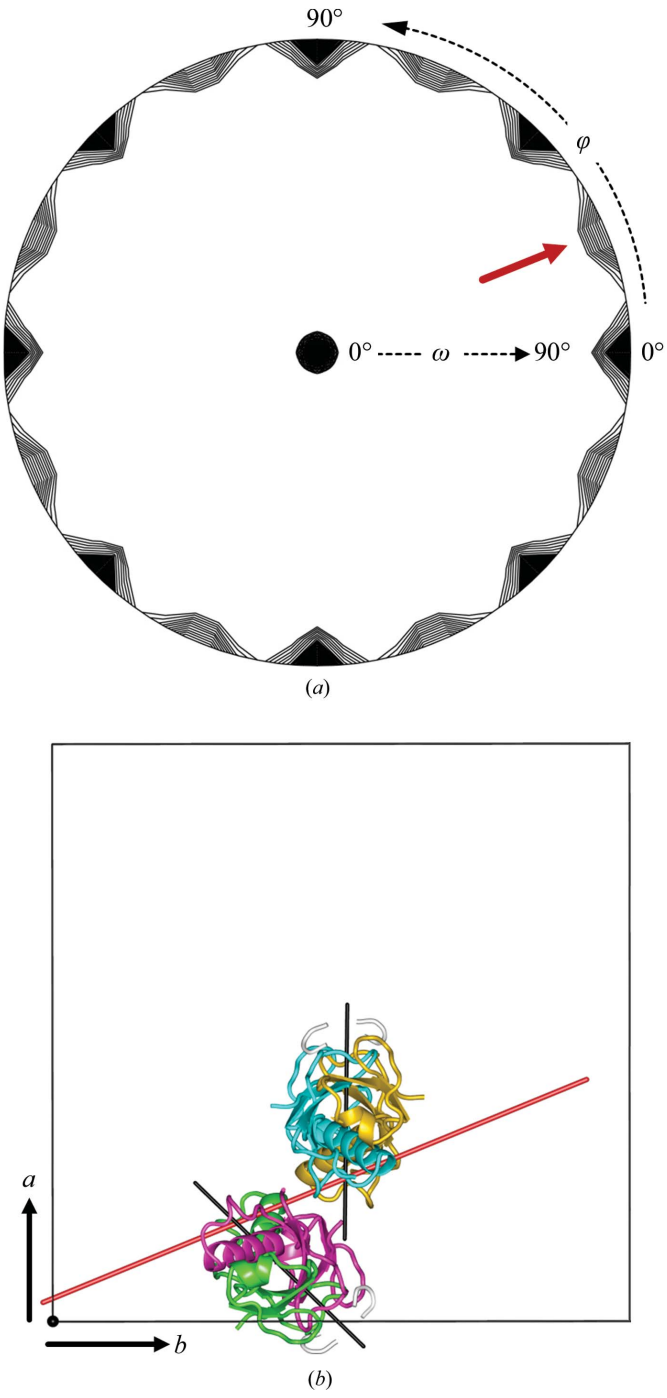


Figure 1
Data analysis and molecular-replacement solution of C-DCX. (*a*) Stereographic projection of the $\kappa = 180°$ section of the self-rotation function. The function was calculated over the full resolution range and a sphere radius of 32 Å using *POLARRFN* (Winn *et al.*, 2011). An additional peak ($\omega = 90$, $\varphi = 22.5°$), 69% of the height compared with the expected positions for 422 symmetry, is marked by a red arrow. (*b*) The four protomers in the asymmetric unit of space group $P4_32_12$, shown as differently coloured ribbons, form two dimers. The final, domain-swapped structures of the protomers are shown. The view is a projection into the *ab* plane of the unit cell. The NCS axes of the dimers are shown as black lines. They make an angle of 45° with and are parallel to the crystallographic axes of this space group. The two dimers are related by another twofold axis (red line), which makes an angle of 22.5° with the *b* axis, explaining the single additional peak in the self-rotation function in (*a*).

Table 1
Data-collection and refinement statistics for C-DCX (PDB entry 6fnz).

Values in parentheses are for the highest resolution shell.

Data collection	
Resolution range (Å)	41.1–2.23 (2.31–2.23)
100% criterion† (Å)	2.23
Space group	$P4_32_12$
Unit-cell parameters (Å)	$a = 98.5$, $c = 114.8$
Total reflections	238822 (21596)
Unique reflections	28139 (2737)
Multiplicity	8.5 (7.9)
Completeness (%)	99.8 (98.8)
R_{sym}‡	0.085 (3.1)
R_{meas}‡	0.091 (3.3)
$CC_{1/2}$‡	0.999 (0.328)
CC*‡	1 (0.703)
Average $I/\sigma(I)$	13.9 (0.68)
Wilson *B* value (Å^2)	60.1
Refinement	
No. of reflections	28101 (2586)
R_{cryst}§	0.208 (0.378)
R_{free}§	0.231 (0.385)
No. of residues	333
No. of waters	42
DPI¶ (Å)	0.36
Phase error†† (°)	29.5
R.m.s.d., bonds (Å)	0.007
R.m.s.d., angles (°)	0.92
Ramachandran plot‡‡ (%)	
Favoured	95.9
Allowed	3.5
Disallowed	0.6
MolProbity score§§	1.65
Clashscore¶¶	2.96
$\langle B \rangle$ (Å^2)	
Protein	80.0
Water	60.9

† The 100% criterion was calculated using *SFTOOLS* (Winn *et al.*, 2011) and represents the resolution in Å of a 100% complete hypothetical data set with the same number of reflections as the measured data. ‡ *R* values and $CC_{1/2}$ are as defined in Diederichs & Karplus (1997) and Karplus & Diederichs (2012), respectively, and were calculated with *PHENIX* (Adams *et al.*, 2010). § $R_{cryst} = \sum_{hkl} ||F_{obs}| - |F_{calc}|| / \sum_{hkl} |F_{obs}|$, where F_{obs} and F_{calc} are the structure-factor amplitudes from the data and the model, respectively. R_{free} is R_{cryst} with 5% of structure factors used as a test set. ¶ Cruickshank diffraction-component precision index based on the *R* value (Blow, 2002). †† The maximum-likelihood-based phase error was calculated with *PHENIX* (Adams *et al.*, 2010). ‡‡ Calculated using *PHENIX* (Adams *et al.*, 2010). §§ The *MolProbity* score should approach the high-resolution limit (Chen *et al.*, 2010). ¶¶ Clashscore is defined as the number of unfavourable all-atom steric overlaps ≥0.4 Å per 1000 atoms (Word *et al.*, 1999).

After completion of the model, residual difference electron density was visible in a surface depression on each C-DCX monomer that spanned a twofold crystallographic axis. The density could not be satisfactorily explained by any crystallization or protein buffer component and was very tentatively modelled as a hexapeptide with sequence PESSEG that might have originated from the bacterial growth medium. Of the side chains, only the glutamate side chains can be clearly identified based on electron density and interactions with nearby arginine side chains. Both C-DCX and the hexapeptide have favourable main-chain torsion angles and the peptide–C-DCX interactions are biophysically sensible (Fig. 3*c*). Since the average *B* values for the peptide (107 Å^2) are higher than those for the C-DCX protein (78 Å^2), the peptide may not be fully occupied. Refinement statistics are collected in Table 1. ESI mass-spectrometric analyses of all C-DCX protein preparations used here in both positive-ion and negative-ion

modes did not detect a smaller (peptide) molecule. Hence, we assume that crystallization led to enrichment of this contaminant.

2.3. Analytical ultracentrifugation

Characterization of C-DCX by analytical ultracentrifugation (AUC) was performed as described previously (Burger *et al.*, 2016), with the following additions to examine self-association: samples of 0.1–31 mg ml^{-1} (7.5–2265 µ*M*) C-DCX in 20 m*M* CAPS–NaOH, 0.1 *M* NaCl, 5 m*M* TCEP pH 10.5 (with density ρ = 1.00503 g ml^{-1} and viscosity η = 1.039 mPa s), where C-DCX is most stable in solution, were loaded into SedVel60K charcoal-filled Epon AE or BE centrepieces (Spin Analytical; 1.2 or 0.3 cm optical path length for concentrations of $\leq$1 mg ml^{-1} or >1 mg ml^{-1}, respectively). Prior to starting AUC runs the rotor with the loaded sample cells was kept for 2 h at 20°C in the centrifuge to ensure complete thermal equilibration of the experimental setup. Samples were analyzed in sedimentation-velocity mode at 60 000 rev min^{-1} and 20°C on a Proteome Lab XLI analytical ultracentrifuge equipped with an An-60Ti rotor

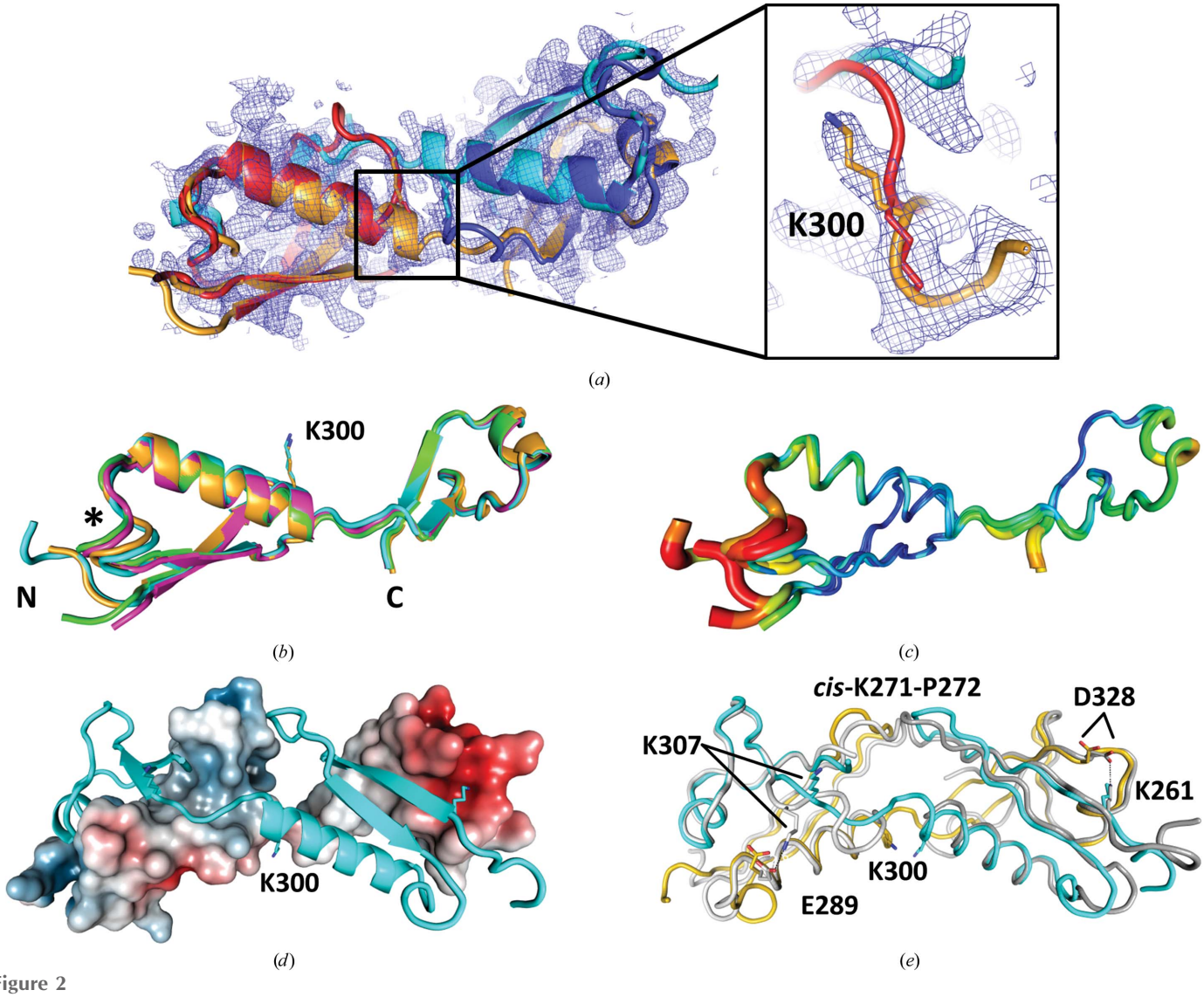


Figure 2
Domain swap in C-DCX. (*a*) Weighted $2F_o - F_c$ electron density contoured at 1 r.m.s.d. directly after molecular replacement, prior to refinement. Two search-model protomers (red and blue) without the domain swap (trimmed PDB entry 5ip4) are shown. Continuous electron density starting from residue Lys300 into the opposing protomer at the lower right of the image indicates that the C-terminal 45 residues (Lys300–Asp344) have been swapped. (*b*) Superposition of the four protomers in the asymmetric unit yields r.m.s.d. values of between 0.5 and 0.9 Å over the entire sequence. Lys300, as the hinge site of the domain swap, is marked for reference. (*c*) The *B*-value putty of the asymmetric unit shows that the site of the domain swap does not display the most elevated *B* values but seems to be rather fixed. Blue and red colours mark the extremes of small and large *B* values. The termini and a surface loop (marked with an asterisk) display the largest variations in *B* values. (*d*) The domain-swapped C-DCX dimer. The surface of one protomer is coloured according to its electrostatic potential and the other is displayed as a ribbon. The view is rotated 90° about the *x* axis relative to (*b*) and (*c*). (*e*) The same dimer as in (*d*) with two charged residue pairs shown as stick models. Two globular C-DCX structures (PDB entry 5ip4) are superimposed and shown as grey ribbons. While in the globular structure the residue pairs Lys261/Asp328 and Glu289/Lys307 form charged hydrogen bonds, these interactions are absent in the domain-swapped dimer.

(Beckman Coulter). Radial scans were monitored by absorbance at 240 nm for 0.1 mg ml^{-1} C-DCX concentration, 275 nm for 1–16 mg ml^{-1} C-DCX concentrations or 300 nm for 31 mg ml^{-1} C-DCX concentration. The partial specific volume of 0.742 ml g^{-1} was calculated from the amino-acid sequence of C-DCX. Raw sedimentation-velocity data were

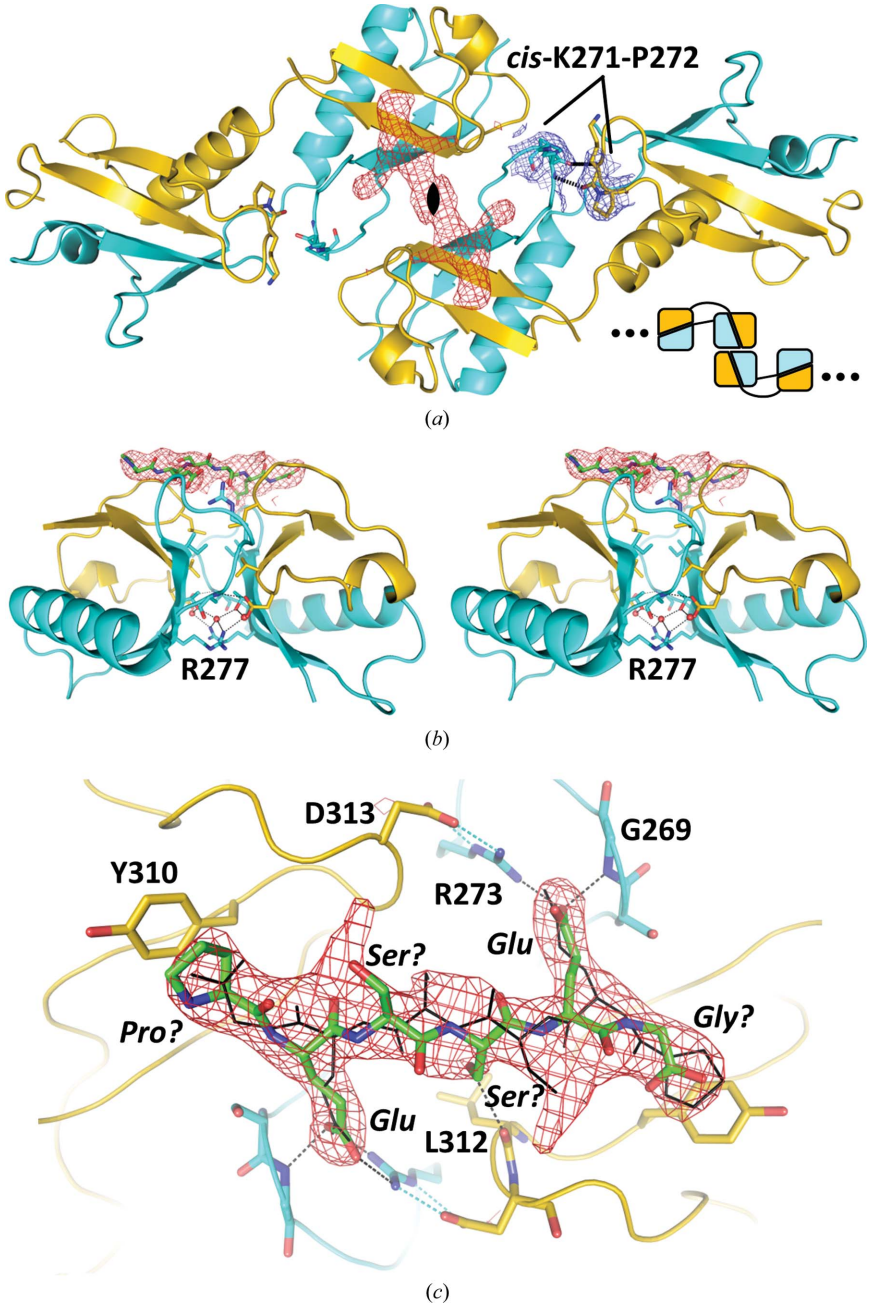


Figure 3
Secondary and tertiary protein–protein interfaces in C-DCX. (*a*) A Lys271-Pro272 *cis*-prolyl peptide bond is present at the secondary interface generated in the domain-swapped dimer. The *cis*-peptide had previously gone unnoticed owing to a lack of clear electron density in the monomeric C-DCX crystal structure (PDB entry 5ip4) but is well defined in this structure ($2F_o - F_c$ density shown as a blue mesh contoured at 1 r.m.s.d.). The backbone amides of this *cis*-peptide form reciprocal hydrogen bonds across the secondary interface (black dashed lines). A large symmetric tertiary interface of ∼2700 Å^2 centred at a twofold axis (marked) connects two domain-swapped dimers into a tetramer. The scheme on the lower right shows that both subunits of each dimer partake in this interface. Difference OMIT electron density contoured at 3 r.m.s.d. indicates a ligand bound at the top of the interface. (*b*) The stereoview of the tertiary interface after rotation by 90° about the horizontal axis shows a small hydrophobic core formed by aliphatic side chains below the difference density. The bottom part of the interface is dominated by polar interactions, including Arg277, which is mutated to histidine in lissencephaly. (*c*) The electron density can be modelled by a peptide. The sequence PESSEG was chosen, but is tentative. The two possible orientations of the peptide are shown as thick green and thin black sticks. The carboxylate groups (Glu and the C-terminus) of the peptide are in hydrogen-bond geometry with the guanidinium side chain of Arg273 and the main-chain amide NH group of Gly269. The proline chosen to model the electron-density stacks on top of the phenolic side chain of Tyr310. In the other orientation, glycine stacks on Tyr310. Amino-acid names are given for the peptide orientation shown in thick sticks.

analyzed with *SEDFIT* (Schuck, 2000) to derive experimental sedimentation coefficients. Sedimentation-coefficient distributions $c(s)$ were plotted and integrated for isotherm construction with *GUSSI* (Brautigam, 2015). The integration interval 0.5–4.0 S covered all protein species, and no aggregates were observed at larger sedimentation coefficients. *SEDFIT* was used to convert experimental sedimentation coefficients to signal-weighted average sedimentation coefficients (s_w), and the coefficients were corrected for buffer density and viscosity ($s_{20,w}$). *SEDPHAT* (Zhao & Schuck, 2015) was used to fit isotherm data. For calculations of theoretical sedimentation coefficients from atomic coordinates, hydrodynamic bead modelling as implemented in *WinHydroPro* (Ortega *et al.*, 2011) was used (mode 1: shell model from atomic level). During data fitting with the isodesmic and monomer–dimer–tetramer (MDT) models as implemented in *SEDPHAT*, $s_{monomer}$ and equilibrium association constants were treated as floating parameters, while s_{dimer} and $s_{tetramer}$ were fixed to the values determined by hydrodynamic modelling using the respective models derived from the C-DCX crystal structure. A second method for the calculation of oligomer sedimentation coefficients starting from that of the monomer uses the hydrodynamic scaling law $s \simeq M^{2/3}$, where $s_{n\text{-mer}} = s_{monomer} \times n^{2/3}$ (Schuck & Zhao, 2017). The fits to the data were plotted with *GraphPad Prism*.

3. Results

3.1. Structure of the domain-swapped C-DCX domain

Previous structure determination of C-DCX in complex with a nanobody (Burger *et al.*, 2016) revealed a small compact domain with a ubiquitin-like fold, *i.e.* a curved, five-stranded mixed β-sheet straddling an α-helix that crosses the whole molecule (PDB entry 5ip4). We obtained crystals of C-DCX purified in alkaline buffer at pH 10.5, where it is most stable, and a Tacsimate condition at pH 7, where the calculated pH of the protein/reservoir mixture is ∼7.3. The diffraction data obtained could not be phased by molecular replacement using the C-DCX domain (PDB entry 5ip4) as the search model in any combination of the possible space groups $P4_32_12$ or $P4_12_12$ and number of molecules to be searched (4–6), suggesting significant structural rearrangements of the C-DCX molecule in this crystal. A solution of four molecules was found using trimmed and modified coordinates of PDB entry 5ip4 as the search model, but the electron density after initial refinement showed a domain swap in C-DCX centred at the hinge-loop sequence 300-KLET-304 (Fig. 2*a*). Lys300 is located at the C-terminus of the central α-helix of the ubiquitin fold. Upon domain swapping, this helix extends by a single turn, thus rigidifying the hinge region and stabilizing the extended form. The extent of the domain swap is substantial in relation to the whole molecule, essentially splitting C-DCX in half and rotating 36 of its 77 residues by 180°, thus duplicating the maximum dimension of the molecule. The split formally increases the surface area of C-DCX by 1690 Å^2, or 31%, from 5370 Å^2 for the globular C-DCX domain to 7080 Å^2 for the extended form, which we term the 'open monomer' (Bennett *et al.*, 1995) as this molecule is also observed in solution (see below). Superposition of the four open monomers reveals root-mean-square distances (r.m.s.d.) of 0.5–0.9 Å, with the largest deviations among the monomers at the N-termini and around residue Lys283, where the maximum C$^\alpha$ distances are ∼5 Å (Fig. 2*b*). These regions also exhibit the highest *B* values (Fig. 2*c*), indicating structural plasticity. Interestingly, the area around the hinge region at Lys300 is rather invariant among the monomers and exhibits low *B* values but no crystal contacts, in line with a rigidifying effect of the helix extension at the hinge.

In the crystal, the surface of the open monomer is masked by another monomer to form a domain-swapped C-DCX dimer (Fig. 2*d*). The dimers in the C-DCX crystal bury a surface area of ∼1500 Å^2 per monomer with a large surface-complementarity coefficient S_c of 0.7 (Lawrence & Colman, 1993). As expected, the closed monomer has similar characteristics for these parts, with 1580 Å^2 buried surface area and S_c = 0.7. These values would argue in favour of the presence of a domain-swapped dimer also in solution, while the open monomer should be unstable, but the latter is not the case (see below). The domain-swapped dimer re-establishes many of the interactions that are present in the globular form of C-DCX, *i.e.* the interface in the domain-swapped dimer is similar to that of a closed monomer, as is typical for domain swaps (Bennett *et al.*, 1995). For instance, the small hydrophobic core of the domain-swapped dimer makes the same interactions in both the extended and globular forms of C-DCX. By contrast, two electrostatic interactions in the globular form between side-chain pairs Lys261/Asp328 and Glu289/Lys307 are absent in the domain-swapped dimer (Fig. 2*e*). As the domain swap requires the opening of these surface-located pairs, a pH-dependent swap mechanism might be envisaged, but this was not tested experimentally in solution as C-DCX is not very stable near its calculated pI (8.2; Burger *et al.*, 2016). The observation of the domain swap using crystallization conditions near the physiological pH would argue against a pH-dependent switch. A secondary interface in the domain-swapped dimer, which is not present in the C-DCX monomer, involves Lys271-Pro272 *cis*-peptide bonds. These are well defined by electron density (Fig. 3*a*) but were not detected in the C-DCX monomer structure, where this region appears to be flexible (Burger *et al.*, 2016). To avoid intermolecular clashes, the loop containing the *cis*-peptide has to adapt in the domain-swapped dimer (Fig. 2*e*). As a result, the main-chain NH group of Lys271 from one monomer forms a (reciprocal) hydrogen bond with the carbonyl group of Pro272 in the opposing monomer (Fig. 3*a*).

A third, larger interface of 2700 Å^2 and S_c = 0.7 is present between protomers in the crystal lattice of C-DCX (Figs. 3*a* and 3*b*). The interface is centred at a crystallographic twofold axis and consists of a small hydrophobic core constructed by Thr264, Ile266, Ala275, Leu312 and Val329. Polar main-chain and side-chain interactions, some of which are water-mediated, are present between Thr264, Arg273, Arg277, Asp313 and Asp328. Together with the domain swap, this interface would

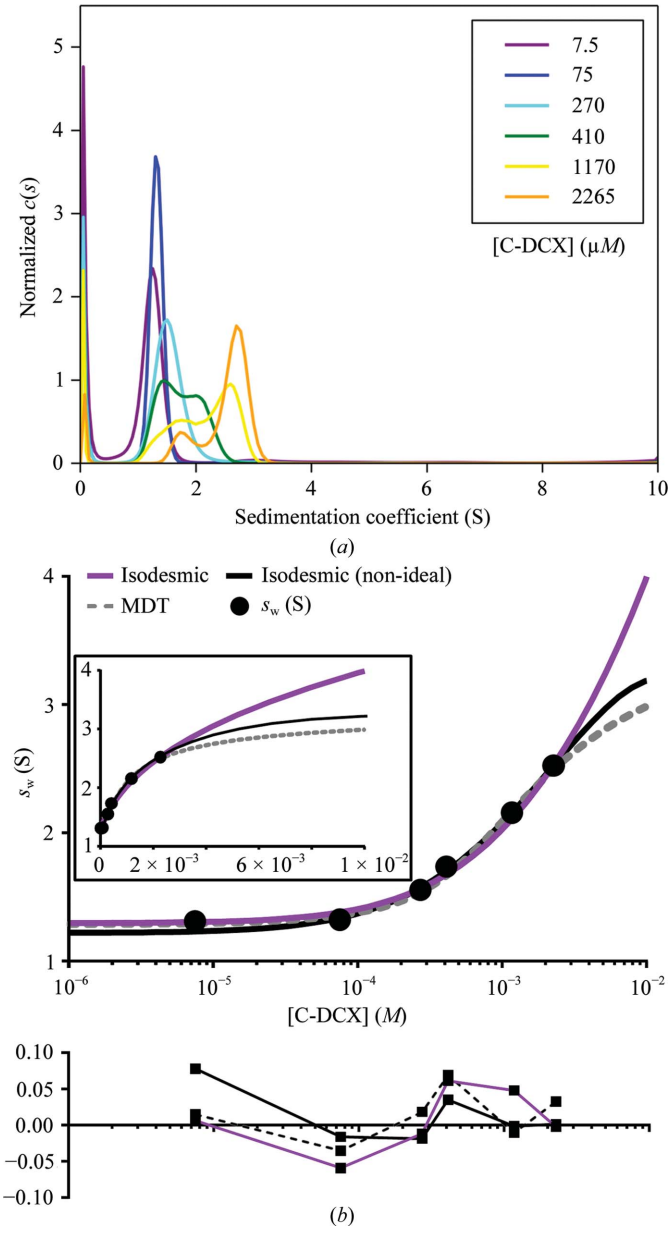


Figure 4
Analytical ultracentrifugation of C-DCX. (*a*) Sedimentation-coefficient distributions $c(s)$ normalized to the total peak area between 0.5 and 4 S. Six sedimentation-velocity experiments were performed at the C-DCX concentrations indicated (micromolar). The higher the concentration, the more prevalent higher-order oligomers are. The concentration-dependent shift of the signal-weighted sedimentation coefficient s_w indicates a fast association equilibrium in solution. No aggregation is observed, as none of the species have sedimentation coefficients of >3.5 S. (*b*) The sedimentation coefficients s_w from the data in (*a*) form an isotherm (black dots). The inset shows the same distribution with a linear scale for the abscissa. Dotted, magenta and black lines represent fits to these data according to the MDT model (reduced $\chi^2 = 4.99 \times 10^5$), the isodesmic model (reduced $\chi^2 = 3.07 \times 10^5$) and the isodesmic model including corrections for non-ideality (reduced $\chi^2 = 0.8 \times 10^5$), respectively. The bottom graph depicts the residuals of the fits to the data. All models deviate <5% from the data. For the isodesmic model the sedimentation coefficient of the dimer was fixed to its theoretical value during the fit, thus limiting oligomer extension to the addition of monomers. For the MDT model, no tetramer formation was allowed by the (improbable) assembly of four monomers into a tetramer and the sedimentation coefficients for the dimer and tetramer were fixed to their theoretical values during the fit.

Table 2
Sedimentation coefficients calculated from coordinates by hydrodynamic modelling.

Sedimentation coefficients were calculated from coordinates by hydrodynamic bead modelling (Ortega *et al.*, 2011). For reference, on the basis of hydrodynamic scaling laws the sedimentation coefficient for a globular dimer of C-DCX would be 2.08 S, assuming a value of 1.31 S for the monomer. A smooth, equivalent sphere of the same partial specific volume and molecular mass as the domain-swapped C-DCX dimer would have a sedimentation coefficient of 2.76 S, the maximum possible for such a (unphysiological) dimer. Generally, values of $s_{dimer}/s_{monomer}$ of 1.6 ± 0.2 are expected for globular proteins (Frigon & Timasheff, 1975; Schuck & Zhao, 2017).

Entity	s_{calc} (S)	$s_{20,w,calc}$ (S)
Monomer (PDB entry 5ip4)	1.48	1.57
Open monomer	1.26	1.33
Domain-swapped dimer	2.22	2.35
Staggered tetramer	3.37	3.56

thus allow tetramers or even higher-order oligomers of C-DCX to form. A staggered tetramer is formed in the crystal by two domain-swapped dimers, which in principle could also lead to chain-type C-DCX oligomers (inset in Fig. 3*a*). The dimer is further stabilized by a small ligand of unknown identity. While none of the crystallization components would fit into this *C*2 symmetric density, it is most satisfactorily explained by a pseudosymmetric hexapeptide, the acidic groups of which could be placed with confidence based on geometric and hydrogen-bonding criteria (Fig. 3*c*). Although partial proteolysis during crystallization is often observed, none of the conceivable hexapeptides from the C-DCX sequence would fit the density. Thorough mass-spectrometric analyses of the C-DCX preparations prior to crystallization did not detect peptides, suggesting that a contaminant peptide of unknown origin is enriched during crystallization. Unfortunately, no C-DCX crystals remained for mass-spectrometric analysis.

3.2. C-DCX forms dimers and tetramers in solution

Given the presence of a dimer and possibly tetramers in the crystal structure, the question arises whether the oligomers also persist in solution and what their shapes might be. This question was addressed using AUC velocity experiments at pH 10.5, where C-DCX is most stable (Burger *et al.*, 2016). Solutions with different C-DCX concentrations were centrifuged at 60 000 rev min^{-1} and the resulting concentration profiles were analyzed in a time-dependent manner *via* protein absorbance measurements (see §2). The lowest concentration that was experimentally accessible was 0.1 mg ml^{-1}, or 7.5 μ*M*, C-DCX. Under these conditions, a symmetric sedimentation-coefficient distribution consistent with a single species of C-DCX is observed (Fig. 4*a*). The species has a signal-weighted sedimentation coefficient of s_w = 1.31 S (68% confidence interval 1.30–1.32 S), which is highly consistent with that of the open C-DCX monomer observed in the crystal structure. Hydrodynamic bead modelling yields s_{calc} = 1.26 S for the open monomer (Fig. 4; Table 2). For comparison, the calculated sedimentation coefficient (s_{calc}) of the compact, non-domain-swapped form of

C-DCX is 1.48 S. Thus, under the experimental conditions and at low protein concentration, C-DCX exists as a single, elongated species in solution with a shape that closely matches that observed in the domain-swapped dimer in the crystal structure. This is despite the fact that a hydrophobic surface patch is exposed to solvent upon opening of the globular monomer (Fig. 2*d*).

At slightly higher concentrations of C-DCX (1 and 3 mg ml^{-1}) a shift of the signal-weighted sedimentation coefficient s_w (integration from 0.5 to 4 S, which included all protein species) towards increasingly higher values is visible. This indicates the formation of a reaction boundary in the sedimentation process owing to concentration-dependent self-association of C-DCX. At concentrations of >3 mg ml^{-1} the distributions separate into two maxima, indicating further association of C-DCX into oligomers (Fig. 4*a*). At the highest C-DCX concentration of 31 mg ml^{-1}, or 2265 µ*M*, that we could attain, the apparent s_w value for the integration interval 0.5–4 S is 2.53 S, with the fast oligomer species migrating at s_w = 2.68 S (Fig. 4*a*). No signal was observed for sedimentation coefficients of >3.5 S, proving that no nonspecific association (aggregation) occurred in any measurement. The observation of gradually increasing s_w values as a function of protein concentration is a hallmark of protein self-association or hetero-association. The increasing s_w values are then a measure of the shift of the association/dissociation equilibrium of monomers into higher-order oligomers. Therefore, the s_w values of the sedimentation-coefficient distributions were re-plotted, and the resulting s_w isotherm was fitted with different monomer–oligomer models (Fig. 4*a*), where the sedimentation coefficients for the monomer and equilibrium dissociation constants were the fit parameters. As the measured apparent sedimentation coefficient is much higher than that expected for a C-DCX dimer (Table 2 and §2) and also changes with concentration, a simple monomer–dimer fast equilibrium model is insufficient to describe the concentration-dependent self-association of C-DCX. Thus, higher-order oligomers must be accounted for.

Two models were applied to describe the measured AUC data: an isodesmic (monomer–multimer) model and an explicit monomer–dimer–tetramer (MDT) model. In the isodesmic model, all equilibrium constants for all association/dissociation steps are assumed to be identical ($K_{d,iso}$) and oligomer extension proceeds stepwise by the addition of monomers. In contrast, the MDT model assumes two sequential steps, monomer–dimer followed by dimer–tetramer formation, governed by distinguishable equilibrium dissociation constants ($K_{d,MD}$ and $K_{d,DT}$, respectively). Based on the residuals of the fits to the data, both models describe the experimental data over the C-DCX concentrations accessible equally well (Fig. 4*b*).

The isodesmic model converged with $s_{20,w,monomer}$ = 1.37 S, with a 68.3% confidence interval (CI) of 1.29–1.45 S, and an equilibrium dissociation constant of $K_{d,iso}$ = 1524 µ*M* (CI of 1462–1614 µ*M*). The MDT model yielded $s_{20,w,monomer}$ = 1.36 S (CI of 0.97–1.60 S) together with $K_{d,MD}$ = 1847 µ*M* (CI of 413–48518 µ*M*) and $K_{d,DT}$ = 227 µ*M* (CI of 1.2–903 µ*M*). Both models return similar $s_{20,w,monomer}$ values close to the expected value for the open monomer (Table 2). The fits for the isodesmic and MDT models deviate at C-DCX concentrations of >2 m*M* (Fig. 4*b*), which were unattainable. Hence, a distinction between the isodesmic and MDT models cannot be made based on the AUC data. However, in the MDT model the confidence intervals of the two K_d values overlap, and thus the latter are statistically indistinguishable. In addition, $K_{d,DT}$ is smaller than $K_{d,MD}$, which is physically implausible in the absence of a ring closure or other higher-order self-interaction during oligomerization (Frigon & Timasheff, 1975). Thus, the most parsimonious model to describe the self-association of C-DCX is the isodesmic model, which accounts for a series of higher-order oligomers. A slight caveat with this view is that isodesmic models usually require identical interaction surfaces, for example in fibre formation, while the domain-swapped C-DCX dimer that self-assembles into the staggered tetramer exhibits two different interfaces (Figs. 2 and 3).

A distinction between the isodesmic and MDT models would require much higher C-DCX concentrations exceeding 31 mg ml^{-1}. These were not attainable and, in addition, such AUC data are expected to suffer from excluded volume effects: non-ideality owing to the reduced average distance of solute molecules approaching their molecular radii would impair the sedimentation process (Erickson, 2009). To address potential hydrodynamic non-ideality, we included a correction in the data analysis as implemented in *SEDPHAT* to account for its possible occurrence at high protein concentrations. The coefficient k_s describes the concentration-dependent reduction of the observed sedimentation velocity s and relates it to the ideal sedimentation coefficient s_0 (extrapolated to 'infinite' dilution) by $s = s_0 \times (1 - k_s \times c)$, where c is the concentration of the solute. Based on the original treatment by Rowe (1977) and a frictional ratio of $f/f_{0(\text{C-DCX},1\,\text{mg ml}^{-1})}$ = 1.28, we estimated k_s = 0.01 ml mg^{-1} for C-DCX. However, fixing k_s at this value results in poor fits to the isotherm for both the isodesmic and MDT models. Adding k_s as a floating parameter to the isotherm analysis with the isodesmic model yields a value of k_s = 0.003 ml mg^{-1}, which translates into a factor of $1 - k_s \times c$ = 0.91 for the correction of the observed sedimentation coefficients at the highest concentration of 31 mg ml^{-1}. For all samples measured at lower concentrations the observed sedimentation coefficients deviate less than 5% from the corrected sedimentation coefficients. With this correction, K_d = 1131 µ*M* (CI of 1006–1412) is calculated and the fitted $s_{20,w,monomer}$ value is 1.29 S (CI of 1.21–1.38). This value is in excellent agreement with s_{calc} = 1.26 S for the open monomer. In the case of the MDT model, k_s converges at a value of negligible magnitude when treated as a floating parameter, and hence $K_{d,MD}$, $K_{d,DT}$ and $s_{20,w,monomer}$ do not change significantly. The self-association of C-DCX is relatively weak, with K_d = 1524 µ*M* (K_d = 1131 µ*M* taking into account hydrodynamic non-ideality). On a mass concentration scale this corresponds to K_d > 10 mg ml^{-1}, and for such weak protein–protein interactions the two parameters k_s and K_d tend to be correlated in the data-fitting procedure (Schuck & Zhao, 2017). Taken together, for the experimentally accessible

Table 3
Lissencephaly- and double cortex-causing mutations mapping to C-DCX.

Numbering follows that of the UniProt entry for human doublecortin and is shifted by +81 relative to other reports of structures and patient mutations.

Mutation	Possible effect on doublecortin structure or function
Arg259Leu	Removes charge near N-terminus, possibly affecting interaction with N-DCX
Pro272Arg	Loss of *cis*-peptide bond removes two hydrogen bonds across secondary interface in the domain-swapped dimer
Arg273Trp	Arg side chain interacts with unknown moiety across tertiary interface in staggered tetramer
Arg277His	Disruption of tertiary interface in staggered tetramer
Asn281Lys	Destabilization of surface-located turn by loss of hydrogen bond; introduces a third positive charge into the turn
Thr284Arg	Surface-located, change of charge
Thr303Ile	Located in the linker region for the domain swap; the two Ile residues clash with each other in the domain-swapped dimer
Gly304Glu	Residue directly following the linker region; reduced flexibility may affect monomer opening
Phe324Leu	Introduces destabilizing cavity in hydrophobic core

concentration range in the case of C-DCX, hydrodynamic non-ideality is not expected to compromise the sedimentation-velocity AUC analysis and, unfortunately, the addition of the hydrodynamic non-ideality parameter k_s did not allow an unequivocal distinction between the isodesmic and MDT models.

In summary, the analytical centrifugation data revealed an open monomer for C-DCX at low protein concentrations that is poised to form a specific domain-swapped dimer. At higher concentrations C-DCX must form higher-order oligomers, likely based on a staggered tetramer. The low overall affinity of the oligomers in either model is consistent with the fast exchange kinetics suggested by the concentration-dependent changes in the s_w values and in the overall forms of the sedimentation-coefficient distributions.

3.3. Structural basis of some lissencephaly mutations

In light of the observed oligomerization of C-DCX, several lissencephaly- and double-cortex-causing mutations may be explained on a molecular level, especially those that would affect microtubule binding and cross-linking (see §3.4). Disease-causing mutations are distributed over the entire *DCX* gene, with nine mutations locating to the C-DCX domain (Table 3). Mapping of these mutations to the C-DCX structure reveals that some surface-located mutations (Arg259Leu, Asn281Lys and Thr284Arg) change the surface entropy and charge but should not have drastic effects on the C-DCX structure. Such mutations may interfere with the electrostatic aspects of microtubule binding. Other mutations (for example Phe324Leu) introduce defects into the hydrophobic core, thus destabilizing doublecortin. More interesting are those mutations that affect the swapping mechanism or that change the secondary and tertiary interfaces of the domain-swapped dimer and the staggered tetramer. Thr303Ile and Gly304Glu are located within or directly next to the linker region involved in the swap. In the domain-swapped dimer, the Ile side chains from each monomer at position 303 sterically clash, thus destabilizing the dimer. By contrast, modelling of Ile at position 303 reveals no detrimental effect on the structure of the globular C-DCX monomer, supporting a biological role for the domain-swapped dimer. While the inherent flexibility of Gly304 may be required for the domain swap, any variation of this side chain, in addition to rigidifying this site, leads to steric clashes with the main-chain carbonyl group, even in the globular monomer. The Gly304Glu mutation will therefore change the structure of C-DCX compared with the wild type. Pro272 forms a *cis*-peptide bond with Lys271, which is abrogated in the Pro272Arg mutation, thus destabilizing the secondary interface in the domain-swapped dimer (Fig. 3*a*). Two more mutations, Arg277His and Arg273Trp, affect the tertiary interface and thus may hinder tetramer formation. Arg277His is located in the polar part of the tertiary interface and this mutation will result in the loss of a few water-mediated interactions (Fig. 3*b*). From a mechanistic standpoint the Arg273Trp mutation is of particular note as the guanidinium side chain of Arg273 forms four hydrogen bonds: two with Asp313 across the tertiary interface and another two with the unknown entity bound to C-DCX. Mutation of Arg273 will not only disrupt the tertiary interface but also may abolish the functional binding of another (microtubule-associated) protein. Generally, the effect of the disease-causing mutations on the structure and function of doublecortin is probably limited, as a complete loss of doublecortin activity should be fatal. In line, several mutant doublecortin proteins retained the ability to stimulate microtubule formation in a turbidity assay, albeit not at the level of wild-type doublecortin, most likely because the cooperativity of doublecortin binding to microtubules was lost (Bechstedt & Brouhard, 2012).

3.4. Cautious model for cooperative DCX–microtubule interaction and bundling

Doublecortin cooperatively binds and bundles microtubules, but the molecular mechanism of these activities is unknown (Bechstedt & Brouhard, 2012). Using the concept of the domain swap, we propose a tentative model for doublecortin–microtubule binding that explains both the cooperativity and bundling activities of doublecortin *via* the same mechanism.

The tubulin $\alpha\beta$-dimer polymerizes into protofilaments, which continue to assemble into sheets or microtubules of 13 or 14 protofilaments. Microtubule *in vitro* pelleting assays have shown that the tandem arrangement of the ubiquitin-like domains connected by their flexible linker (N-DCX_C-DCX, or T-DCX) is necessary and sufficient for co-assembly with microtubules (Taylor *et al.*, 2000; Kim *et al.*, 2003), while the unstructured Ser/Pro-rich C-terminal region of doublecortin is not. In addition, overproduction of T-DCX in cells leads to tubulin polymerization and microtubule bundling (Horesh *et al.*, 1999). Cooperative binding of doublecortin to microtubules has been described, which implies interactions between several doublecortin molecules when bound to tubulin (Bechstedt & Brouhard, 2012). Two low-resolution

(∼8 Å) cryo-EM studies of a complex between doublecortin and microtubules showed electron density at the vertex of four tubulin dimers in the crevice between two protofilaments (Liu *et al.*, 2012; Fourniol *et al.*, 2010). The density observed is large enough to host but a single ubiquitin-like domain, while the rest of doublecortin was disordered. To explain the cooperativity, binding of doublecortin to microtubules *via* one of its ubiquitin-like domains would require intermolecular interactions between several doublecortin molecules mediated by the other ubiquitin-like domain. While one docking study (Fourniol *et al.*, 2010) found N-DCX to fit better than a model of C-DCX, at this resolution both possibilities should be considered. Docking the N-DCX domain into the density would require the C-DCX domains of microtubule-bound doublecortin to interact, and *vice versa*. Superposition of the closed monomeric C-DCX domain (PDB entry 5ip4) onto N-DCX docked to the microtubule places the N-terminus of C-DCX away from the microtubule. In this position, cooperative interactions would require oligomerization *via* N-DCX domains, which however has not yet been reported. In addition, superposition of the domain-swapped C-DCX dimer onto the docked N-DCX domain results in severe clashes of the second C-DCX domain with tubulin. These two arguments make it much more likely that it is indeed N-DCX that contacts tubulin in microtubules and C-DCX protrudes away from them.

In both cryo-EM structures of microtubules complexed with N-DCX, the C-terminus of N-DCX points away from the

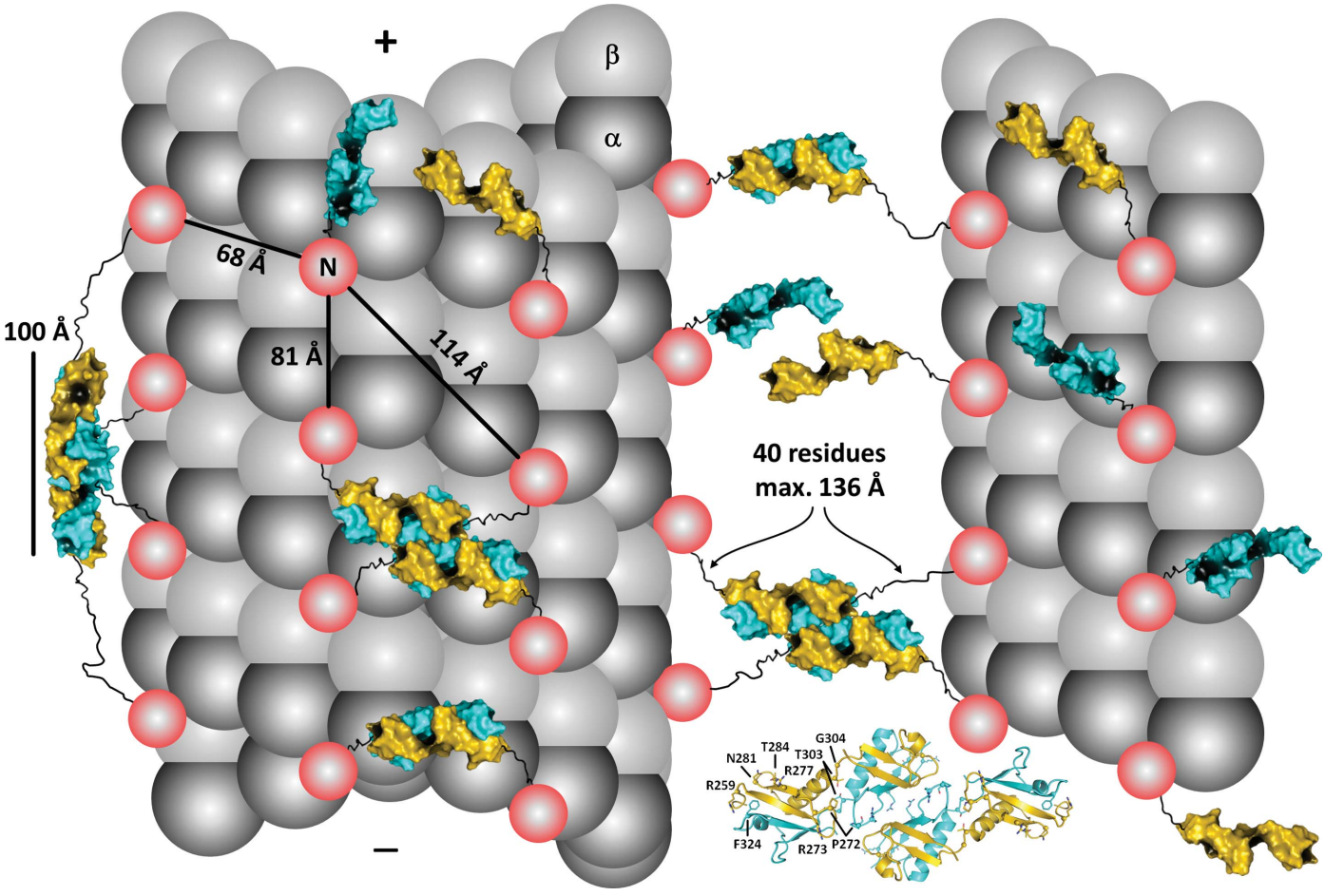


Figure 5
Possible modes for the cooperative binding of doublecortin to microtubules and for microtubule bundling. Schematic microtubules are shown as grey spheres polymerized from α-tubulin and β-tubulin (shades of grey). The exposure of β-subunits at the (+)-end and of α-subunits at the (–)-end of the microtubule determines its polarity. N-DCX (red spheres) binds to the vertex of four $\alpha\beta$-tubulin dimers. Only every other N-DCX binding site around the circumference of the microtubule is occupied in this scheme. Distances between these N-DCX molecules are given, along with the dimension of the staggered C-DCX tetramer (left) and the maximum dimension of the linker region between N-DCX and C-DCX (centre). The linker between N-DCX and C-CDCX is drawn as a black line (not to scale). The microtubule is less densely decorated with doublecortin *in vivo* (estimated at a physiological doublecortin:$\alpha\beta$-tubulin ratio of ∼1:70; Taylor *et al.*, 2000). A few possible orientations of the staggered doublecortin tetramer are shown. Those limited to a single microtubule (left) would explain the cooperative binding of doublecortin, and those cross-linking two microtubules would explain microtubule bundling by doublecortin. In the absence of tertiary interface formation, only a domain-swapped doublecortin dimer may bind and cross-link microtubules, but with lower cooperativity. A few schematic monomers with the open C-DCX conformation, prior to dimer formation, are also indicated. The C-terminal Pro/Ser-rich part of doublecortin is not included in the figure. Note that although the staggered tetramer has *C*2 symmetry, this has to be broken by linker flexibility in order to bundle microtubules with the same polarity. On the lower right, the disease-causing mutations of one protomer of C-DCX are mapped in the context of the staggered tetramer.

microtubule (Liu *et al.*, 2012; Fourniol *et al.*, 2010), which allows the formulation of cooperativity and, by extension, a microtubule cross-linking model for doublecortin (Fig. 5). In this model, doublecortin molecules are bound *via* their N-DCX domains to microtubules, and the flexible linker allows the open form of the C-DCX domain to search for a nearby binding partner. If doublecortin molecules are bound sufficiently close to each other on microtubules, they can dimerize using the domain swap of C-DCX and, if close enough to another doublecortin dimer, form a staggered tetramer. The ∼40-residue linker connecting N-DCX and C-DCX, if fully extended, covers a maximum distance of 136 Å, much more than is required to span the distance from one N-DCX domain to any of its nearest possible neighbours (Fig. 5). The linker length itself is subject to variation. A conformational switching model has been proposed with Trp227 in the linker binding to Arg137 in N-DCX (Cierpicki *et al.*, 2006). This interaction could reversibly shorten the linker, thus modulating the search area of doublecortin for another binding partner. While self-assembly of doublecortin on the same microtubule would explain cooperative binding, microtubule bundling would be the result of self-assembly of doublecortin bound to different fibres. Both the domain-swapped dimer and the staggered tetramer of C-DCX would be able to bundle microtubules, with a higher cooperativity expected for the tetramer. Binding of doublecortin to microtubules saturates at a doublecortin:$\alpha\beta$-tubulin ratio of unity, and at this ratio the rate of tubulin polymerization is largest (Horesh *et al.*, 1999). However, the *in vivo* doublecortin:$\alpha\beta$-tubulin ratio is much smaller and is estimated at 1:70 (Taylor *et al.*, 2000). As our proposed cooperativity and bundling model does not rely on an infinite chain of oligomerized doublecortin, a much smaller *in vivo* ratio is still consistent with the model.

4. Discussion

The exchange of structural elements between monomers to form dimers, cyclic oligomers or linear oligomers is not uncommon in proteins, with at least 60 domain-swapped structures available to date (Liu & Eisenberg, 2002; Rousseau *et al.*, 2012). Although the swapped 'domain' is often just an N- or C-terminal α-helix or β-strand, some proteins exchange larger parts such as immunoglobulin domains in diabodies (Perisic *et al.*, 1994) and triabodies (Pei *et al.*, 1997). Double domain swaps do occur, but are much rarer [for example in catechol-*O*-methyltransferase (Ehler *et al.*, 2014) and streptococcal protein GB1 (Frank *et al.*, 2002)]. Half of the molecule may be swapped in proteins composed of two homologous domains, for example in the eye lens protein βB2-crystallin (97 of 204 residues swapped; Bax *et al.*, 1990) or the cyanovirin-N dimer (48 of 101 residues swapped; Yang *et al.*, 1999). However, even small, usually monomeric single-domain proteins such as the SH3 domains of Eps8 (Radha Kishan *et al.*, 1997) and p47^{phox} (Groemping *et al.*, 2003) and the SH2 domain of Grb2 (Schiering *et al.*, 2000) undergo domain swaps of up to half of their molecular mass. While in this light the domain swap of C-DCX does not appear to be particularly unusual, the presence of a pre-formed extended monomer in solution that has already opened and is thus primed to form the domain-swapped dimer certainly is.

Based on the large Stokes radius derived from gel-permeation studies and on the increase in 8-anilino-1-naphthaline sulfonate (ANS) fluorescence when interacting with C-DCX, a molten globule state was previously proposed for this domain (Kim *et al.*, 2003). ANS fluorescence is quenched by water and increases in nonpolar environments such as the hydrophobic protein interior accessible in a molten globule. In contrast, another study on C-DCX showed cooperative unfolding of the domain at pH 4.5 (T_m = 59°C) and pH 10 (T_m = 68°C, the most stabilizing pH), as well as limited NMR H/D exchange and significant amide-proton dispersion at pH 4. In addition, as revealed by AUC at pH 7.5, the C-DCX monomer was the dominant species (Burger *et al.*, 2016). All of these properties are characteristic of a folded domain. As detected here by AUC, the open C-DCX monomer is present in solution and thus exposes hydrophobic surfaces that are normally buried in the hydrophobic core. Binding of ANS to these surfaces provides an alternative explanation for the increase in quantum yield that was previously observed. Hence, rather than forming a molten globule, C-DCX is more likely to exist as an elongated open monomer in solution that is prone to (self-)associate *via* its hydrophobic surfaces. The question now arises as to why the open C-DCX monomer forms in solution and remains stable without aggregating.

A survey of domain-swapped protein structures (Rousseau *et al.*, 2012) has delineated that shorter, often proline-containing loops are candidates for hinges enabling domain swaps. Shorter loops would stress the protein structure by storing potential energy. Accordingly, hinge loops have been described as driving forces for domain swapping in terms of a loaded spring that releases its potential energy upon generation of the dimer (Schymkowitz *et al.*, 2001). For C-DCX, the role of the short but proline-free linker region KLET would thus be to release energy upon formation of the observed stable open monomer in solution, independent of a subsequent dimerization. The energy that is released must overcompensate both the enthalpic penalty of opening a hydrophobic core and the entropic penalty of water molecules associating with the solvent-exposed hydrophobic surfaces. As the monomer does not expose an exceedingly large hydrophobic area (Fig. 2*d*) it is not very prone to aggregation, at least at elevated pH values. At pH 7.5 and a concentration of 8 µ*M*, which is similar to the lowest protein concentration measured here, monomeric C-DCX had a slightly lower sedimentation coefficient of 1.15 S, while the sedimentation-coefficient distribution also showed larger species (Burger *et al.*, 2016). Compared with our AUC data at alkaline pH, this indicates a larger population of open C-DCX monomers and additionally an even higher propensity for oligomerization at physiological pH. For N-DCX, the equivalent loop region is three residues longer compared with C-DCX and has an unrelated sequence. The longer loop in N-DCX has less potential to store energy, which might explain why no domain swap had hitherto been observed with this domain.

The C-DCX dimer does not form in very dilute solutions but at higher concentrations where an additional small (secondary) interface involving the *cis*-peptide bonds is formed (Fig. 3*a*). Further, this dimer may continue to assemble into larger oligomers using a tertiary interface. The presence of a ligand, possibly but by no means certainly a peptide, that latches the dimers together across the tertiary interface will enhance the stability of the tetramer. Whether a ligand-aided mode of doublecortin oligomerization is of biological significance remains to be determined. However, ligands are frequently found to bind at subunit interfaces, and there is also increasing evidence for biologically relevant ligand-regulated domain swaps: for example, the monomer–dimer equilibrium of glyoxalase I is regulated by glutathione (Saint-Jean *et al.*, 1998) and that of p13suc1, a cell-cycle regulator in fission yeast, is regulated by a phosphopeptide (Schymkowitz *et al.*, 2001).

In a previous study to determine the degree of self-association of full-length doublecortin by AUC, doublecortin was shown to be predominantly monomeric at concentrations below 12.5 µ*M* and to promote microtubule nucleation in its monomeric form (Moores *et al.*, 2006). Furthermore, the tandem arrangement of N-DCX and C-DCX is responsible for enhanced tubulin polymerization, possibly owing to doublecortin acting as a nucleation catalyst (Kim *et al.*, 2003). From these findings, it was concluded that doublecortin would not *a priori* self-associate to serve as a polymerization template for tubulin, but rather that a single doublecortin molecule should facilitate tubulin nucleation by simultaneously interacting with multiple tubulin dimers. Alternatively, molecular crowding of doublecortin on nascent microtubules followed by self-assembly can be envisioned. Doublecortin can be viewed as a 'ligand' for a 'multivalent receptor', the microtubule lattice, with regularly spaced binding sites. Doublecortin was found to cooperatively bind microtubules based on sigmoidal concentration-dependent binding of GFP-labelled doublecortin to microtubules and on decreased dissociation rates at higher degrees of receptor occupancy (Bechstedt & Brouhard, 2012). The structural reason for the cooperativity is not known, but a direct interaction of doublecortin molecules has been discussed (Bechstedt & Brouhard, 2012). Based on the available data, we hypothesize that the formation of the domain-swapped dimer could be a means by which doublecortin cooperatively interacts on the surface of microtubules, thereby exerting its function in the nucleation and stabilization of microtubules. As the open monomer of C-DCX is present in solution, such open doublecortin molecules might also occur on microtubules in search for a nearby binding partner, be it on the same or another fibre. A cross-linking activity of doublecortin for microtubule bundling has previously been suggested (Kim *et al.*, 2003). Both *in vitro* microtubule growth and the formation of domain-swapped protein dimers are favoured under conditions of molecular crowding (Wieczorek *et al.*, 2013; Hernández-Vega *et al.*, 2017; Liu & Eisenberg, 2002). In the crowded environment of the cytosol, protein–protein interactions are favoured compared with diluted *in vitro* solutions where crowding is usually mimicked by polymer co-solutes, and domain swaps are favoured at extremes of pH (Liu & Eisenberg, 2002). Crowding effects are amplified by restricting protein movement to two dimensions, as on the microtubule lattice. The next steps could therefore include the investigation of the possible self-association of doublecortin at higher concentrations, and its interaction with microtubules under conditions of molecular crowding.

Our model of doublecortin–microtubule interaction relies on the flexibility of the linker connecting the N-DCX and C-DCX domains. This linker flexibility would explain why three-quarters of doublecortin remained invisible in the cryo-EM complex structures (Liu *et al.*, 2012; Fourniol *et al.*, 2010): although a stable domain swap plus further assembly into a tetramer are parts of our model that require an ordered C-DCX structure, the flexible linker would lead to averaging out of the electron density for these regions in the cryo-EM complexes. Based on geometric and antibody-inhibition experiments, we previously suggested that the C-DCX domain binds to microtubules (Burger *et al.*, 2016), whereas the model presented here requires N-DCX to bind. The resolution of the cryo-EM structures is not sufficient to determine which of the two ubiquitin-like domains binds to microtubules, and the linker between them is long enough to allow any domain to bind with the other still solvent-exposed. Neither the individual N-DCX nor C-DCX domains alone binds to microtubules, indicating weak affinities (Burger *et al.*, 2016). An anti-N-DCX antibody pelleted with T-DCX and microtubules, while an anti-C-DCX nanobody inhibited co-pelleting of T-DCX with microtubules (Burger *et al.*, 2016). These observations suggested that it was C-DCX that binds to microtubules. However, these observations are also explained if C-DCX cross-linking is required for strong doublecortin binding to microtubules: if a network of doublecortin forms on microtubuli, an anti-N-DCX antibody may recognize an unbound N-DCX domain. By contrast, if an anti-C-DCX nanobody abolishes cross-linking no doublecortin network can form in the first place. The necessity of a network for tight doublecortin binding to microtubules could be tested by mutagenesis of the hinge loop in C-DCX with the aim of inhibiting swapping.

It was found previously that doublecortin can discriminate between microtubules consisting of 13 and 14 protofilaments, with specificity for the former (Bechstedt & Brouhard, 2012). The differences in curvature and N-DCX binding-site distances between these two types of microtubules are subtle and cannot be explained by a model involving a flexible linker. However, a slight change of curvature might already be sufficient to alter the vertex of the four $\alpha\beta$-tubulin dimers that construct the binding site for N-DCX. In line with this, it was found by fluorescence microscopy (Ettinger *et al.*, 2016) that doublecortin preferentially binds to the flat, GDP-bound microtubule lattice but not locally curved microtubules with a lattice consistent with the geometry imposed by GTP-bound tubulin dimers (Alushin *et al.*, 2014). It therefore seems that the N-DCX domain is a microtubule curvature sensor while the C-DCX domain of doublecortin is responsible for cooperative microtubule binding and bundling.

5. Conclusion

Taking the crystallographic and AUC solution data together, two possible scenarios for C-DCX self-association are conceivable. At low concentration, C-DCX either exists predominantly as an open monomer in solution or undergoes a fast equilibrium between closed and open monomers. Two aspects point to the presence of an open monomer: (i) C-DCX has a symmetrical $c(s)$ distribution indicating a single discrete species and (ii) the measured sedimentation coefficient at low concentrations matches the predicted sedimentation coefficient of the open monomer, not the closed monomer. We have no evidence for the co-existence of open and closed monomers under the experimental conditions at high pH values, although a rapid opening and closing of C-DCX, with a preferential population of the open monomer, is likely under physiological conditions. At higher concentrations the domain-swapped dimer is formed, which can crystallize and has a propensity to further self-assemble into a tetramer, possibly of a staggered form. Under physiological conditions and guided by a microtubule template, such low-affinity domain-swapped dimers and staggered tetramers could cooperatively bind to and cross-link microtubules, thus stabilizing the fibre and mediating fibre bundling.

Acknowledgements

We thank the staff of SLS beamline PXII and our colleagues from Expose for support during synchrotron data collection and Johannes Erny for mass-spectrometric analyses of the C-DCX preparations.

References

Adams, P. D. *et al.* (2010). *Acta Cryst.* D**66**, 213–221.
Alushin, G. M., Lander, G. C., Kellogg, E. H., Zhang, R., Baker, D. & Nogales, E. (2014). *Cell*, **157**, 1117–1129.
Ayanlaja, A. A., Xiong, Y., Gao, Y., Ji, G., Tang, C., Abdikani Abdullah, Z. & Gao, D. (2017). *Front. Mol. Neurosci.* **10**, 199.
Bax, B., Lapatto, R., Nalini, V., Driessen, H., Lindley, P. F., Mahadevan, D., Blundell, T. L. & Slingsby, C. (1990). *Nature (London)*, **347**, 776–780.
Bechstedt, S. & Brouhard, G. J. (2012). *Dev. Cell*, **23**, 181–192.
Bennett, M. J., Schlunegger, M. P. & Eisenberg, D. (1995). *Protein Sci.* **4**, 2455–2468.
Blow, D. M. (2002). *Acta Cryst.* D**58**, 792–797.
Brautigam, C. A. (2015). *Methods Enzymol.* **562**, 109–133.
Burger, D. *et al.* (2016). *J. Biol. Chem.* **291**, 16292–16306.
Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. (2010). *Acta Cryst.* D**66**, 12–21.
Cierpicki, T., Kim, M. H., Cooper, D. R., Derewenda, U., Bushweller, J. H. & Derewenda, Z. S. (2006). *Proteins*, **64**, 874–882.
Diederichs, K. & Karplus, P. A. (1997). *Nature Struct. Biol.* **4**, 269–275.
Ehler, A., Benz, J., Schlatter, D. & Rudolph, M. G. (2014). *Acta Cryst.* D**70**, 2163–2174.
Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* D**66**, 486–501.
Erickson, H. P. (2009). *Biol. Proced. Online*, **11**, 32–51.
Ettinger, A., van Haren, J., Ribeiro, S. A. & Wittmann, T. (2016). *Curr. Biol.* **26**, 1549–1555.
Fourniol, F. J., Sindelar, C. V., Amigues, B., Clare, D. K., Thomas, G., Perderiset, M., Francis, F., Houdusse, A. & Moores, C. A. (2010). *J. Cell Biol.* **191**, 463–470.
Frank, M. K., Dyda, F., Dobrodumov, A. & Gronenborn, A. M. (2002). *Nature Struct. Biol.* **9**, 877–885.
Frigon, R. P. & Timasheff, S. N. (1975). *Biochemistry*, **14**, 4559–4566.
Groemping, Y., Lapouge, K., Smerdon, S. J. & Rittinger, K. (2003). *Cell*, **113**, 343–355.
Hernández-Vega, A., Braun, M., Scharrel, L., Jahnel, M., Wegmann, S., Hyman, B. T., Alberti, S., Diez, S. & Hyman, A. A. (2017). *Cell. Rep.* **20**, 2304–2312.
Horesh, D., Sapir, T., Francis, F., Wolf, S. G., Caspi, M., Elbaum, M., Chelly, J. & Reiner, O. (1999). *Hum. Mol. Genet.* **8**, 1599–1610.
Kabsch, W. (2010). *Acta Cryst.* D**66**, 125–132.
Karplus, P. A. & Diederichs, K. (2012). *Science*, **336**, 1030–1033.
Keegan, R. M., Long, F., Fazio, V. J., Winn, M. D., Murshudov, G. N. & Vagin, A. A. (2011). *Acta Cryst.* D**67**, 313–323.
Kim, M. H., Cierpicki, T., Derewenda, U., Krowarsch, D., Feng, Y., Devedjiev, Y., Dauter, Z., Walsh, C. A., Otlewski, J., Bushweller, J. H. & Derewenda, Z. S. (2003). *Nature Struct. Biol.* **10**, 324–333.
Lawrence, M. C. & Colman, P. M. (1993). *J. Mol. Biol.* **234**, 946–950.
Liu, Y. & Eisenberg, D. (2002). *Protein Sci.* **11**, 1285–1299.
Liu, J. S., Schubert, C. R., Fu, X., Fourniol, F. J., Jaiswal, J. K., Houdusse, A., Stultz, C. M., Moores, C. A. & Walsh, C. A. (2012). *Mol. Cell*, **47**, 707–721.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Moores, C. A., Perderiset, M., Kappeler, C., Kain, S., Drummond, D., Perkins, S. J., Chelly, J., Cross, R., Houdusse, A. & Francis, F. (2006). *EMBO J.* **25**, 4448–4457.
Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* D**67**, 355–367.
Ortega, A., Amorós, D. & García de la Torre, J. (2011). *Biophys. J.* **101**, 892–898.
Pei, X. Y., Holliger, P., Murzin, A. G. & Williams, R. L. (1997). *Proc. Natl Acad. Sci. USA*, **94**, 9637–9642.
Perisic, O., Webb, P. A., Holliger, P., Winter, G. & Williams, R. L. (1994). *Structure*, **2**, 1217–1226.
Radha Kishan, K. V., Scita, G., Wong, W. T., Di Fiore, P. P. & Newcomer, M. E. (1997). *Nature Struct. Biol.* **4**, 739–743.
Romero, D. M., Bahi-Buisson, N. & Francis, F. (2018). *Semin. Cell Dev. Biol.* **76**, 33–75.
Rousseau, F., Schymkowitz, J. & Itzhaki, L. S. (2012). *Adv. Exp. Med. Biol.* **747**, 137–152.
Rowe, A. J. (1977). *Biopolymers*, **16**, 2595–2611.
Saint-Jean, A. P., Phillips, K. R., Creighton, D. J. & Stone, M. J. (1998). *Biochemistry*, **37**, 10345–10353.
Schiering, N., Casale, E., Caccia, P., Giordano, P. & Battistini, C. (2000). *Biochemistry*, **39**, 13376–13382.
Schuck, P. (2000). *Biophys. J.* **78**, 1606–1619.
Schuck, P. & Zhao, H. (2017). *Sedimentation Velocity Analytical Ultracentrifugation: Interacting Systems.* Boca Raton: CRC Press.
Schymkowitz, J. W., Rousseau, F., Wilkinson, H. R., Friedler, A. & Itzhaki, L. S. (2001). *Nature Struct. Biol.* **8**, 888–892.
Taylor, K. R., Holzer, A. K., Bazan, J. F., Walsh, C. A. & Gleeson, J. G. (2000). *J. Biol. Chem.* **275**, 34442–34450.
Wieczorek, M., Chaaban, S. & Brouhard, G. J. (2013). *Cell. Mol. Bioeng.* **6**, 383–392.
Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.
Word, J. M., Lovell, S. C., LaBean, T. H., Taylor, H. C., Zalis, M. E., Presley, B. K., Richardson, J. S. & Richardson, D. C. (1999). *J. Mol. Biol.* **285**, 1711–1733.
Yang, F., Bewley, C. A., Louis, J. M., Gustafson, K. R., Boyd, M. R., Gronenborn, A. M., Clore, G. M. & Wlodawer, A. (1999). *J. Mol. Biol.* **288**, 403–412.
Zhao, H. & Schuck, P. (2015). *Acta Cryst.* D**71**, 3–14.