Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Three-dimensional domain swapping and supramolecular protein assembly: insights from the X-ray structure of a dimeric swapped variant of human pancreatic RNase

Andrea Pica,[a]‡ Antonello Merlino,[a,b]‡ Alexander K. Buell,[c] Tuomas P. J. Knowles,[c] Elio Pizzo,[d] Giuseppe D'Alessio,[d] Filomena Sica[a,b] and Lelio Mazzarella[a,b]*

[a]Department of Chemical Sciences, University of Naples 'Federico II', Via Cintia, 80126 Naples, Italy, [b]Institute of Biostructures and Bioimages, CNR, Via Mezzocannone 16, 80134 Naples, Italy, [c]Department of Chemistry, University of Cambridge, Cambridge CB2 1EW, UK, and [d]Department of Structural and Functional Biology, University of Naples 'Federico II', Via Cintia, 80126 Naples, Italy

‡ These two authors contributed equally to this work.

Correspondence e-mail: lelio.mazzarella@unina.it

Received 14 June 2013
Accepted 23 July 2013

PDB Reference: domain-swapped dimer of human pancreatic ribonuclease, 4kxh

The deletion of five residues in the loop connecting the N-terminal helix to the core of monomeric human pancreatic ribonuclease leads to the formation of an enzymatically active domain-swapped dimer (desHP). The crystal structure of desHP reveals the generation of an intriguing fibril-like aggregate of desHP molecules that extends along the **c** crystallographic axis. Dimers are formed by three-dimensional domain swapping. Tetramers are formed by the aggregation of swapped dimers with slightly different quaternary structures. The tetramers interact in such a way as to form an infinite rod-like structure that propagates throughout the crystal. The observed supramolecular assembly captured in the crystal predicts that desHP fibrils could form in solution; this has been confirmed by atomic force microscopy. These results provide new evidence that three-dimensional domain swapping can be a mechanism for the formation of elaborate large assemblies in which the protein, apart from the swapping, retains its original fold.

1. Introduction

Most biological processes at the molecular and cellular levels rely on molecular recognition between proteins and/or nucleic acids. The development of new experimental techniques, as well as the improvement of established techniques, now enables us to obtain deeper insight into biological mechanisms by studying, in atomic detail, the interactions between two or more macromolecules involved in these processes. A striking case of intermolecular interaction is given by protein molecules that can produce large aggregates through a self-association process that eventually become insoluble and precipitate. The formation of aggregates by numerous proteins and peptides unrelated in sequence, structure and function suggests that the propensity to aggregate is an innate property of polypeptides (Dobson, 2004; Rousseau *et al.*, 2006). Proteins occur *in vivo* in many cases at concentrations above their critical concentration for aggregation (Baldwin *et al.*, 2011), but energy barriers against aggregation keep them in a metastable soluble state (Buell *et al.*, 2012).

Understanding the molecular determinants behind biological aggregation has been one of the major goals of structural biologists in recent years. Protein deposits are the hallmark of several severe human diseases, including Alzheimer's, Parkinson's and Huntington's diseases and type 2 diabetes (Dobson, 2001). Amyloidoses, for example, are a

range of clinical disorders caused by extracellular deposition in various tissues of β-sheet-rich elongated, unbranched fibrils known as amyloid fibrils derived from the aggregation of misfolded or partially unstructured peptides and proteins (Pepys, 2006). Sometimes amyloidoses are referred to as conformational diseases in order to emphasize that the aggregation pathway occurs *via* a conformational change which induces the formation of non-native interactions between adjacent protein molecules (Carrell & Lomas, 1997). However, not all reported aggregates and fibrils are enriched in β-sheet structure (Dykes *et al.*, 1978; Lomas & Carrell, 2002) or are associated with deposition diseases (Holmes *et al.*, 1990; Downing & Nogales, 1999).

Three main mechanisms exist by which proteins can specifically self-associate to form fibrils: cross-β-spine aggregation (Nelson *et al.*, 2005), end-to-end stacking (Harrington *et al.*, 1997; Elam *et al.*, 2003) and three-dimensional domain swapping (Bennett *et al.*, 2006; Guo & Eisenberg, 2006). Although cross-β-spine aggregation has drawn the most attention from the scientific community, notable proteins do aggregate by the end-to-end stacking mechanism. The cases of sickle-cell haemoglobin (Harrington *et al.*, 1997) and SOD1 (Elam *et al.*, 2003) are the two best known.

Three-dimensional domain swapping is the process by which two protein molecules reciprocally exchange an identical domain, restoring the same interactions as were present in the monomer (closed interface). The peptide that links the swapped domain to the rest of the chain (the hinge peptide) is often formed by a small number of residues that change their conformation in the swapping process. The resulting dimer has two structural units that are similar to the monomeric species and that are formed by residues from the two protein chains.

Such a dimer is referred to as closed-ended, as there are no unsatisfied (exposed) domains. The open interface is the additional interface not present in the monomer and unique to the dimeric species. Open-ended runaway swapped oligomers can also be formed by consecutive exchange of the swapped domains of three or more monomers. The hypothesis that open-ended runaway swapping is implicated in the formation of protein fibrils has been supported by the observation that human prion protein (Knaus *et al.*, 2001), cystatin C (Janowski *et al.*, 2001) and β_2-microglobulin (Liu *et al.*, 2011), which form fibrils related to diseases, are domain-swapped.

However, even closed-ended swapped dimers have structural features that facilitate aggregation. Indeed, the stability of these dimers is mostly guaranteed by the closed interface that pre-exists in the monomeric form. Moreover, some extra stability is gained upon generation of the open interface, which is usually less important, as the two structural units are often linked by two flexible hinge peptides. This is in contrast to common protein dimers that have acquired the quaternary organization through the development of an extensive open interface. Therefore, domain-swapped dimers are often endowed with a larger freedom in their quaternary assembly (Merlino, Ceruso *et al.*, 2005). This feature may greatly facilitate self-recognition processes and makes these dimers more prone to association.

Here, we present the crystal structure of a variant of human pancreatic ribonuclease (desHP) forced to swap by deleting five residues (16–20) in the loop linking the N-terminal segment (residues 1–15) to the core of the protein (Russo *et al.*, 2000). The N-terminal segment is partially folded as α-helix and embodies some residues that are important for the constitution of the active site (His12 and Asn13). In the variant this helix cannot take over the position occupied in the wild-type enzyme and remains exposed to the solvent, eventually promoting the formation of a stable closed-ended swapped dimer whose specific activity is virtually identical to that of the monomeric wild-type enzyme. The structural characterization by X-ray crystallography reveals that desHP molecules give rise to a densely packed rod-like structure that extends infinitely along one crystallographic axis. On the basis of this arrangement, the formation of fibrils in solution has been predicted and indeed observed.

2. Materials and methods

2.1. Crystallization and data collection

Protein production and purification were performed as described previously (Russo *et al.*, 2000). Crystals of desHP were obtained after one week at 277 K by the hanging-drop vapour-diffusion method using 22%(*w*/*v*) PEG 8000, 0.1 *M* ammonium sulfate as precipitant with a protein concentration of 1.6 mg ml^{-1} in 0.1 *M* sodium citrate buffer pH 6.5, 0.3 *M* NaCl.

X-ray diffraction data were collected at 100 K at the Elettra synchrotron, Trieste, Italy using a 165 mm CCD detector from MAR Research. Diffraction data were processed using the *HKL*-2000 program suite (Otwinowski & Minor, 1997). A summary of the indicators commonly used to estimate the quality of data sets is given in Table 1.

2.2. Structure determination and refinement

Initial phase determination was carried out by the molecular-replacement method using *Phaser* (McCoy *et al.*, 2007). The refined coordinates of des(1–7) HP RNase (PDB entry 1e21; Pous *et al.*, 2001) were used as a search model.

The refinement was carried out with *CNS* v.1.3 (Brünger *et al.*, 1998) and *REFMAC*5 (Murshudov *et al.*, 2011). Several alternating cycles of positional refinement, energy minimization, individual temperature-factor refinement and manual model building were performed. Model rebuilding was carried out using *Coot* (Emsley *et al.*, 2010). Water molecules and sulfate ions were added manually to the model using the same software. *PROCHECK* (Laskowski *et al.*, 1993) was used to analyse the quality of the final structure. The refinement statistics are presented in Table 1. *PyMOL* (Schrödinger; http://www.pymol.org) was used to draw the figures. The final model and structure factors have been deposited in the Protein Data Bank as entry 4kxh.

Table 1
Data-collection and refinement statistics.

Data collection	
Space group	$P2_12_12_1$
Unit-cell parameters (Å)	a = 71.51, b = 74.32, c = 128.53
Resolution range (Å)	20.0–2.70 (2.80–2.70)
Measured reflections	119522
Unique reflections	19442 (1909)
R_{merge} (%)	5.7 (16.1)
Mean $I/\sigma(I)$	28.6 (9.9)
Completeness (%)	100 (100)
Multiplicity	6.1 (6.3)
Dimers per asymmetric unit	2
Refinement statistics	
Resolution limits (Å)	20.0–2.70 (2.80–2.70)
R factor (%)	19.8 (28.0)
R_{free} (%)	24.6 (36.0)
No. of reflections used in refinement	18448 (1806)
No. of reflections used for R_{free} calculation	992 (102)
No. of protein atoms	3836
No. of ions	17
No. of water molecules	211
Average B factors (Å^2)	
Protein, overall	26.7
Solvent atoms	24.3
R.m.s.d., bond lengths (Å)	0.014
R.m.s.d., bond angles (°)	1.7
Ramachandran statistics (%)	
Favoured region	95.8
Allowed region	3.8
Outliers	0.4

2.3. Atomic force microscopy imaging

A 10 µl drop of protein solution (0.6 mg ml^{-1}) was deposited on a freshly cleaved mica sheet (Agar Scientific, Stansted, England), left to dry for about 30 min in air and then rinsed with Milli-Q water to remove salt dissolved in the buffer. The sample was dried with nitrogen. Atomic force microscopy was performed under ambient conditions in intermittent contact mode using a Nanowizard II (JPK Instruments). Mikromasch NSC36 cantilevers with a resonant frequency in the range 75–150 kHz and with nominal spring constants of approximately 1 N m^{-1} were used. Height images were recorded at 2048 × 2048 pixels resolution. The scan rate varied between 0.2 and 1.0 Hz. All image data sets were analysed using *Gwyddion* AFM software (Necas & Klapetek, 2012).

2.4. Fluorescence experiments

The fluorescence spectra of solutions of 1.2 µ*M* Thioflavin T (ThT) were measured with a UV–Vis fluorimeter (constant excitation wavelength 440 nm, emission scanned from 450 to 600 nm). The fluorescence of a solution of 0.6 mg ml^{-1} desHP fibrils in 0.3 *M* NaCl at pH 7.1 and 1.2 µ*M* ThT was then measured.

3. Results

3.1. Overall structure of desHP

desHP was crystallized and diffraction data were collected to a resolution of 2.70 Å. The crystal belongs to space group $P2_12_12_1$. The molecular-replacement approach produced a plausible solution with four copies of the monomeric search model in the asymmetric unit and a very high solvent content (about 60%).

The final model, which includes 3836 protein atoms, 211 water molecules and 13 sulfate ions, was refined to R-factor and R_{free} values of 19.8 and 24.6%, respectively. The statistics of refinement are listed in Table 1. In order to facilitate reading, the residue numbering of wild-type human pancreatic ribonuclease has been adopted; therefore, in the present case Ser15 is covalently linked to Ser21. The monomers are associated as N-terminal swapped dimers with approximate twofold symmetry. In each dimer, residues 1–14 of one chain and residues 24–125 of the partner chain form a structural unit (SU) that closely resembles the native enzyme (PDB entry 1z7x); with respect to the latter, the r.m.s.d. based on C$^\alpha$ atoms of the four SUs is in the range 0.30–0.35 Å.

Each active site is assembled from residues of the two chains and is practically indistinguishable from that of the wild-type enzyme (Fig. 1). The swapping is associated with a four-residue hinge peptide, the conformational features of which control the quaternary assembly of the dimer by determining the relative orientation of the N-terminus with respect to the body of the chain. The two dimers have slightly different quaternary structures; this is shown in Fig. 2(*a*), where the two dimers are drawn after superposition of one SU. The rotation that has to be applied to best superpose the second SU of the two dimers amounts to 23° and is achieved with a modest energy penalty through very small variations of the main-chain torsion angles of the four hinge residues.

In both dimers the position of Asp14 is mostly unperturbed by the swapping: the carboxylate group and the carbonyl O atom are hydrogen-bonded to the side chains of Tyr25 and His48, respectively, of the partner chain; these interactions are typical features of pancreatic-like monomeric ribonucleases. The carboxylate group also makes a further hydrogen bond with the hydroxyl group of Ser21. The polyserine stretch (Ser15, Ser21, Ser22 and Ser23), which constitutes the hinge loop, forms a rather regular threefold helix stabilized by internal hydrogen bonds. In comparison, the swapped dimer (PM8) of a mutated form of the human pancreatic

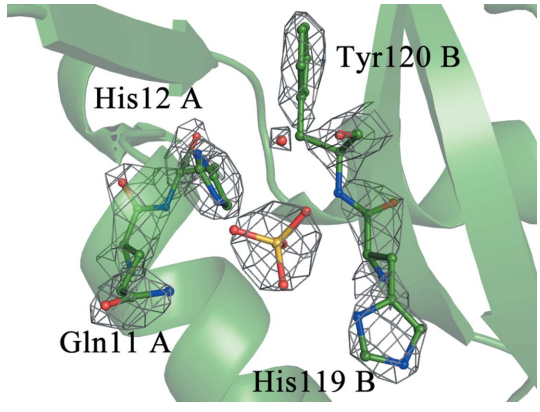

Figure 1
A composite active site generated by residues belonging to subunits A and B of one of the two dimers in the asymmetric unit. The electron-density map is calculated with $2F_o - F_c$ Fourier coefficients and is contoured at the 2.2σ level.

ribonuclease (Canals *et al.*, 2001), shown in Fig. 2(*b*), has a more open structure and the two SUs are further apart with respect to both desHP dimers owing to the longer hinge peptides. In Fig. 2(*c*) the association of the two dimers in the asymmetric unit (formed by the chains labelled *A* and *B* and the chains labelled *C* and *D*) is also shown. Interestingly, the extension of the open surface (about 390 and 255 Å^2 for the two dimers, respectively) is much smaller than that between dimers (see below). It should be noted that both dimers display a pseudo-twofold symmetry, the axis of which is approximately orthogonal and almost intersects the crystallographic screw axis parallel to **c**.

3.2. Supramolecular assembly

The molecular packing displays particularly interesting features. The two dimers are related by an approximate 4_3 axis that is practically coincident with the crystallographic twofold screw axis parallel to **c**. These symmetry operators, together with the twofold symmetry of the dimers, build up rods parallel to the **c** axis with approximate 4_322 symmetry. The inter-dimer association is very strong, burying a surface of about 900 Å^2 (930 and 820 Å^2 for the two crystallographically independent dimer–dimer interfaces) that propagates along the **c** axis, producing tightly packed rods. In contrast, contacts between rods are weak (Fig. 3).

3.3. Sulfate anions

13 sulfate anions have been identified in the electron-density map of the asymmetric unit of the crystal, confirming that this ion plays a fundamental role in the crystallization process. Four sulfate ions are positioned at the active sites, as typically observed in several other members of the pancreatic-like superfamily (Vitagliano *et al.*, 1999; Berisio *et al.*, 1999; Merlino, Mazzarella *et al.*, 2005; Merlino *et al.*, 2009); the remaining anions are located on positive patches of the rod surface. The electrostatic features of the surface calculated with the software *APBS* (Baker *et al.*, 2001) and the bound anions are shown in Fig. 4. In particular, four sulfates are bound at equivalent positions in the four monomers and interact with Arg32 and Arg33. Two more sulfate ions are located at nearly equivalent positions mediating the interaction between rods. The remaining two ions are located at the border of the inter-dimer surfaces. Thus, the presence of the sulfate ions may be important for the stabilization of the rod-like structure characterized by large portions of positively charged surface.

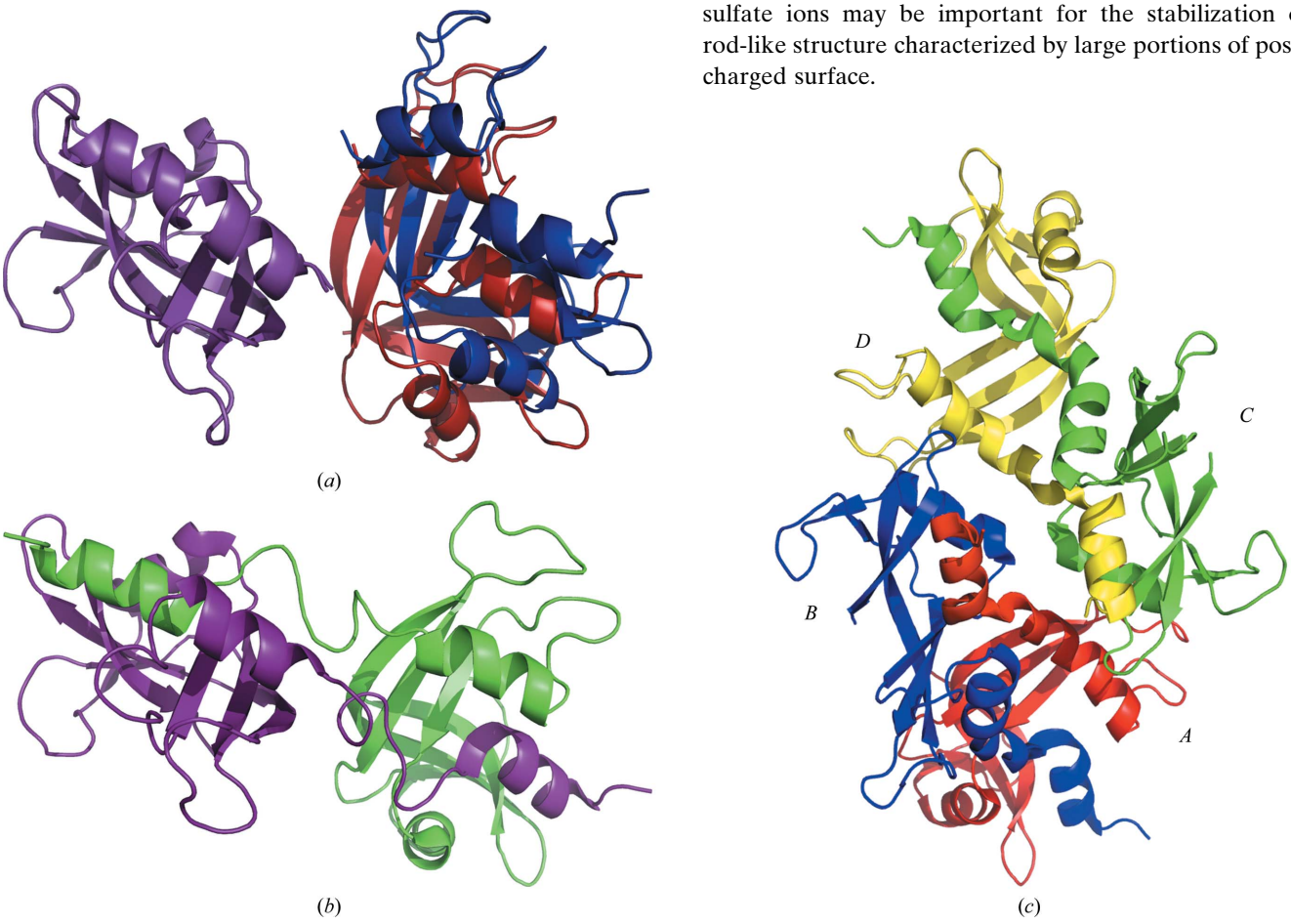

Figure 2
(*a*) View of the two dimers in the asymmetric unit with monomers *A* and *C* (purple) superposed. A further 23° rotation is needed to best superpose the second monomers of the two dimers: *B* (red) and *D* (blue). For clarity, the hinge peptides have not been drawn. (*b*) The structure of a domain-swapped dimeric mutant of human pancreatic ribonuclease (PDB entry 1h8x; Canals *et al.*, 2001) is shown as a cartoon for comparison with the desHP dimers. (*c*) The two dimers in the independent unit are shown as cartoons: chains *A* (red) and *B* (blue) form the first dimer and chains *C* (green) and *D* (yellow) form the second dimer.

3.4. Solution study

In solution desHP displays a strong tendency to form amorphous aggregates, making the growth of crystals suitable for X-ray diffraction difficult. Indeed, at low ionic strength the protein solution is opalescent even at very low concentration (1 mg ml^{-1}). Through an increase in ionic strength (0.3 *M* in NaCl), it was possible to prepare a solution that was more concentrated in protein (up to 2 mg ml^{-1}), from which crystals were obtained. Moreover, a tendency to form fibrils is suggested by the features of the observed crystal packing. Fibrils were observed after incubating a solution of the protein at 0.6 mg ml^{-1} concentration in 0.1 *M* Tris–acetate pH 7.1, 0.2 *M* NaCl both in the presence and in the absence of 0.1 *M* sodium sulfate at room temperature for two weeks. Samples of the solution were deposited on a mica surface, dried and washed with water (see §2). The substrate was then imaged with an atomic force microscope using tapping mode in air. Several AFM images (Fig. 5) clearly showed the presence of long fibrils with a highly homogeneous cross-sectional diameter of about 18 ± 4 Å. In Fig. 6, a histogram of the heights derived from the images is shown. Moreover, the fibrils appear not to be twisted, in contrast to most amyloid fibrils described to date (Fig. 5*b*). We also tested whether these fibrils formed in solution are able to bind ThT and lead to enhanced fluorescence emission at 480 nm upon excitation at 440 nm.

No difference in fluorescence signal was observed between a fresh protein solution and the fibril-containing solution (not shown). This finding suggests that the fibrils are not β-sheet-rich amyloid fibrils and supports the hypothesis that they are structurally related to the rods observed in the crystal.

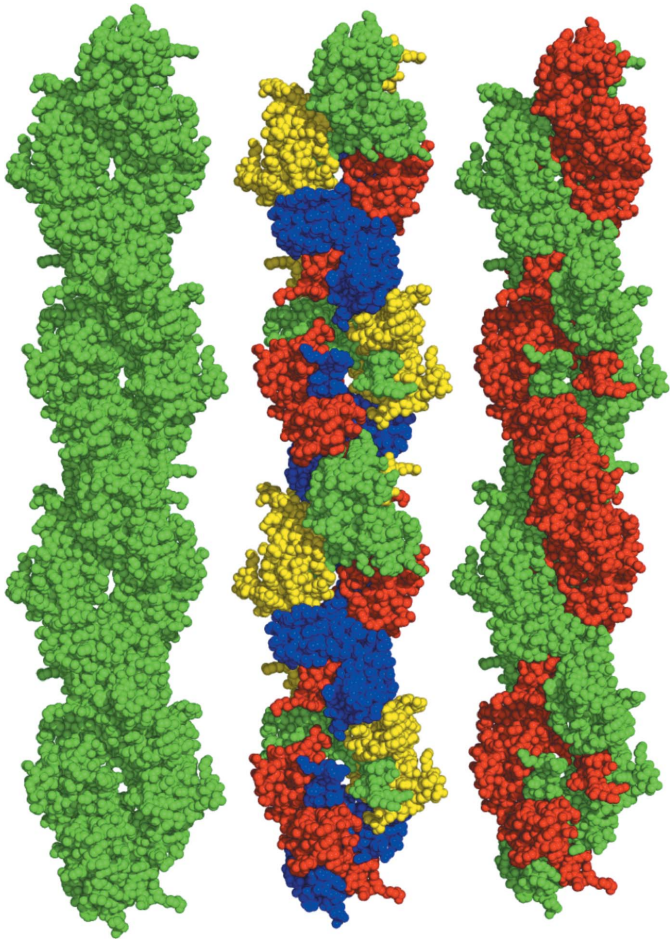

Figure 3
View of the rod structure along the **a** axis. The rod on the left is coloured green. In the central rod, each chain of the two dimers in the asymmetric unit has a different colour according to the colour code used in Fig. 2(*c*). In the third rod, the chains are coloured to highlight the antiparallel double-helical arrangement of the monomers.

4. Discussion

Although proteins can be robust against point mutations, enduring significant numbers of amino-acid substitutions with little change in protein structure or function (Taverna & Goldstein, 2002), it is not uncommon to find cases in which adding/deleting short fragments to/from their sequences partially destabilizes the native fold and favours the formation of dimers or higher order aggregates. Eisenberg and co-workers converted RNase A to a protein able to form fibrils by simply expanding the hinge peptide connecting the core domain of the protein to the exchanging C-terminal β-strand, with an insertion of ten glutamine residues (Sambashivan *et al.*, 2005). The designed amyloid-like fibrils of RNase A contain three-dimensional domain-swapped molecules that are endowed with enzymatic activity. The β-sheet arrangement arises from stacking of the expanded hinge peptides.

Here, we show that the deletion of five residues in the hinge loop of human pancreatic ribonuclease induces the formation of a domain-swapped dimer that leads to the generation of linear aggregates of desHP molecules, as revealed by the crystal packing, in which domain swapping and flexibility of the dimeric assembly are intimately involved. Domain swapping provides the cohesive interaction energy for the formation of dimers through an extensive inter-chain closed interface, whereas the size of the open interface is often small;

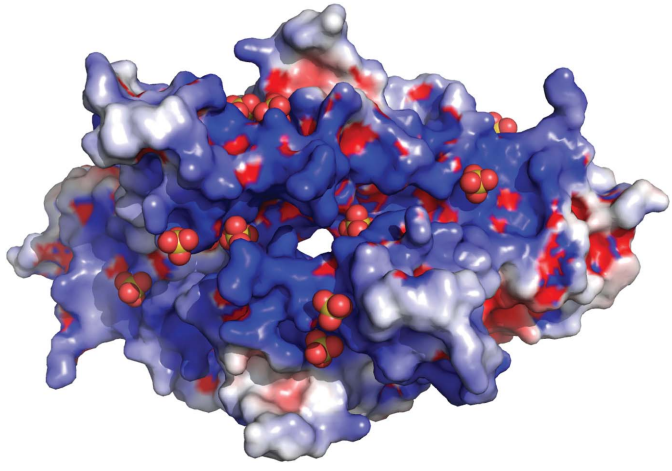

Figure 4
Electrostatic surface of the two dimers in the asymmetric unit. Sulfate ions are shown in ball-and-stick representation.

this feature confers flexibility to the quaternary structure, particularly when the hinge peptide is long, that facilitates auto-recognition of the dimers and favours further aggregation. In this respect, the behaviour of domain-swapped dimers is quite different from that of dimers in which the quaternary structure is held together only by the interactions across an open interface; in this case the stability of the dimer relies on the extension of the interface and this necessarily produces a rather rigid quaternary assembly. In the present case, two dimers with slightly different quaternary structures are present in the crystal: their closed interfaces are very similar and amount to about 1500 Å^2, whereas the open interfaces are small and slightly more different (about 390 and 255 Å^2, respectively). These two dimers alternate on top of each other, forming tightly packed rods which sit on the crystallographic twofold screw axis parallel to **c**. In the rod, the molecules form a left-handed helix with an approximate fourfold symmetry. The small difference in the quaternary assembly of the dimers efficiently increases the complementarity of the two crystallographically independent dimer–dimer contact surfaces, which are 930 and 820 Å^2, respectively. These tightly packed rods run side by side through the crystal with weak lateral contacts. A view of the structure projected down the **c** axis (Fig. 7) shows the almost cylindrical shape of the rods, which have an estimated width of about 25–30 Å.

The desHP crystal structure closely resembles that described for SOD1 (Elam *et al.*, 2003), in which the interfaces within dimers and between dimers along the rod axis are approximately similar and amount to about 650 Å^2. In the present case the interface between dimers is about 50% higher and, most remarkably, the interface that stabilizes the dimeric form through domain swapping amounts to about 2000 Å^2. These data strongly underline the cooperative role between domain swapping and stacking interactions in building up supramolecular structures.

Interestingly, each rod of desHP can be described alternatively as formed by two intertwined antiparallel fourfold helices: one helix is formed by one monomer of the dimers, successively stacked along the rod, say monomers *A*, *C*, *A**, *C** . . . , where the asterisk indicates screw symmetry-related units (see also Fig. 2*c* for the labelling of the chains), and the second helix is built up by monomers *B*, *D*, *B**, *D** . . . of the same dimers. Domain swapping between *A* and *B*, *C* and *D*, *A** and *B** . . . cross-links monomers belonging to different helices and strongly stabilizes the whole structure.

The crystal structure of desHP suggests that the protein could form fibrils in solution and that the rods observed in the

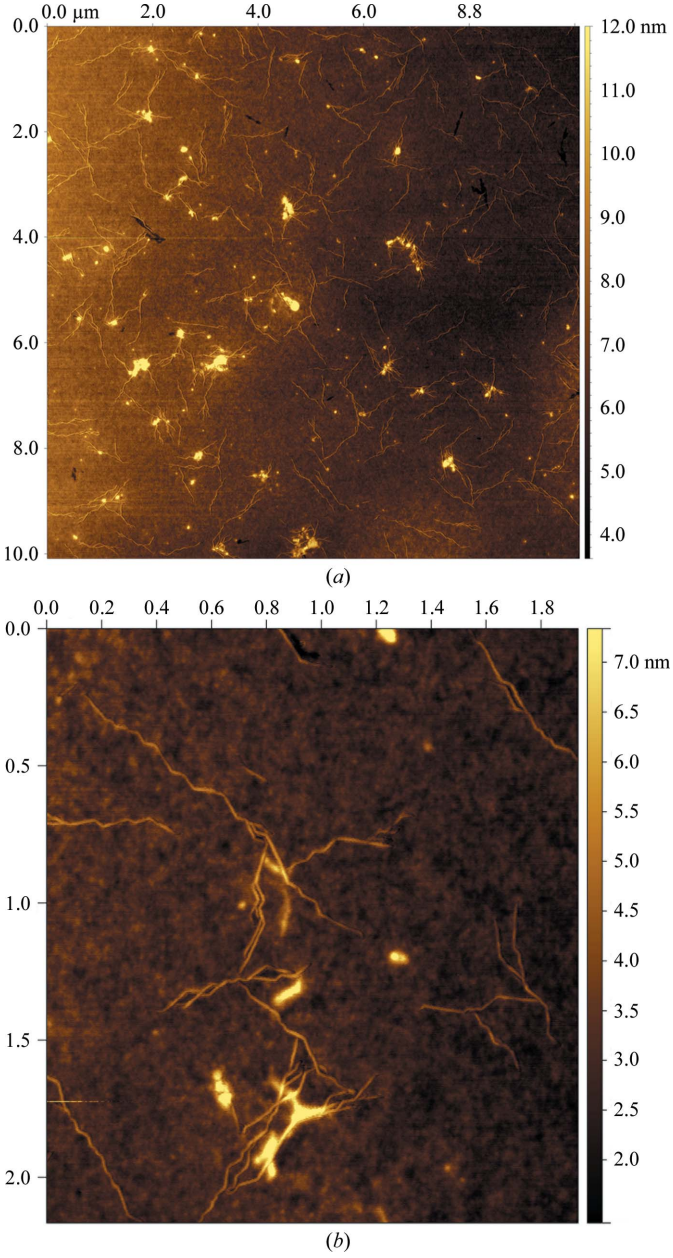

Figure 5
AFM images of desHP fibrils at low (*a*) and high (*b*) magnification.

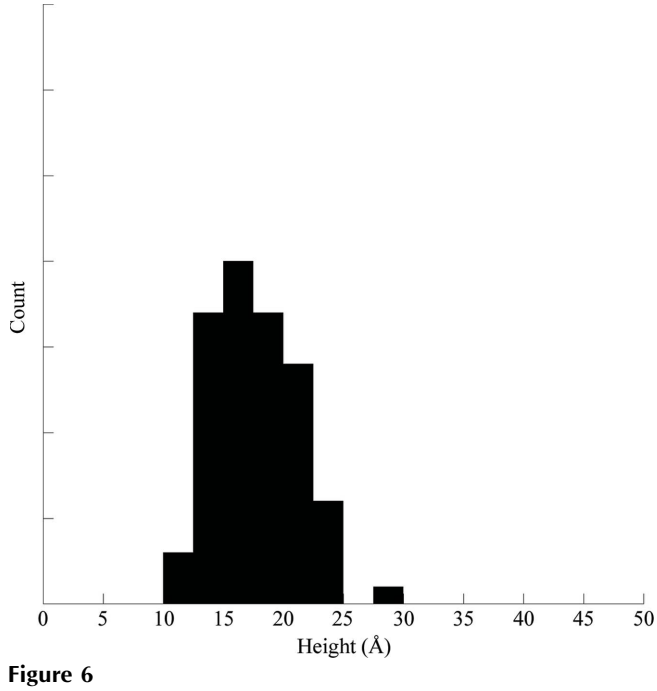

Figure 6
Histogram of desHP fibril heights derived from AFM images.

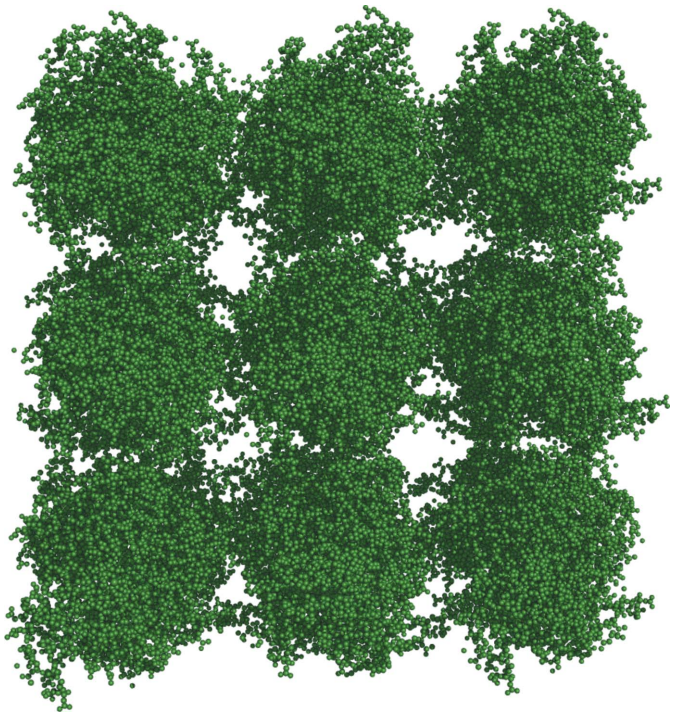

Figure 7
View of the structure along the crystallographic **c** axis.

crystal could be a reasonable model for them. This is enforced by the fact that, as in the case of SOD1, desHP rods are not formed by crystallographic operators only, but their generation involves a noncrystallographic symmetry operation, thus suggesting that the formation of filaments is not simply a necessary consequence of crystallization (Elam *et al.*, 2003). The hypothesis was confirmed by the observation of long unbranched thin fibrils by AFM. Filaments of several micrometres in length were observed in different samples and their thickness was measured (see the histogram of fibril heights in Fig. 6). The diameter measured by AFM (∼20 Å) is smaller than the crystallographic dimension, a result than can be explained by the fact that the AFM images of fibrils were acquired in a dry state on a surface, which can lead to partial denaturation of the structure. The overall effects of these factors are difficult to predict but, taken together, the measurement by AFM of a thinner cross-section than in the crystal is expected and consistent with these various factors.

Based on the AFM images, which show substantial amounts of amorphous protein, it appears that only a small fraction of the protein sample is converted into fibrils, rendering further biophysical examination, for example by CD spectroscopy, difficult. However, the fact that we observe the formation of long unbranched fibrils under quiescent conditions and at low concentrations suggests that these fibrils are indeed the same entities that we found in the crystal state. Moreover, the absence of enhanced ThT fluorescence strongly supports the hypothesis that the fibrils we observe are not cross-β-rich amyloid fibrils but rather fibrils derived from the rod-like assembly observed in the crystal.

To our knowledge, this result represents the first example of a structure-based prediction of protein fibril formation in solution and in turn suggests that these fibrils can be modelled at atomic detail on the basis of the crystal structure.

5. Conclusion

Three-dimensional domain swapping is a mechanism for forming elaborate large assemblies that retain the original fold of the protein, referred to as native-like aggregation. In particular, the interplay between the inter-chain closed interface and the quaternary-structure flexibility, which are characteristic features of a swapped dimer, can play a crucial role in determining self-aggregation processes through a hierarchical organization of the monomers. In particular, starting from the observation of the peculiar supramolecular assembly found in the crystal of desHP, we predicted and indeed observed the formation of long unbranched fibrils. Our results highlight the importance of domain swapping in inducing fibril formation by a mechanism alternative to runaway swapping.

References

Baker, N. A., Sept, D., Joseph, S., Holst, M. J. & McCammon, J. A. (2001). *Proc. Natl Acad. Sci. USA*, **98**, 10037–10041.

Baldwin, A. J., Knowles, T. P., Tartaglia, G. G., Fitzpatrick, A. W., Devlin, G. L., Shammas, S. L., Waudby, C. A., Mossuto, M. F., Meehan, S., Gras, S. L., Christodoulou, J., Anthony-Cahill, S. J., Barker, P. D., Vendruscolo, M. & Dobson, C. M. (2011). *J. Am. Chem. Soc.* **133**, 14160–14163.

Bennett, M. J., Sawaya, M. R. & Eisenberg, D. (2006). *Structure*, **14**, 811–824.

Berisio, R., Lamzin, V. S., Sica, F., Wilson, K. S., Zagari, A. & Mazzarella, L. (1999). *J. Mol. Biol.* **292**, 845–854.

Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.

Buell, A. K., Dhulesia, A., White, D. A., Knowles, T. P., Dobson, C. M. & Welland, M. E. (2012). *Angew. Chem. Int. Ed. Engl.* **51**, 5247–5251.

Canals, A., Pous, J., Guasch, A., Benito, A., Ribó, M., Vilanova, M. & Coll, M. (2001). *Structure*, **9**, 967–976.

Carrell, R. W. & Lomas, D. A. (1997). *Lancet*, **350**, 134–138.

Dobson, C. M. (2001). *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **356**, 133–145.

Dobson, C. M. (2004). *Semin. Cell Dev. Biol.* **15**, 3–16.

Downing, K. H. & Nogales, E. (1999). *Cell Struct. Funct.* **24**, 269–275.

Dykes, G., Crepeau, R. H. & Edelstein, S. J. (1978). *Nature (London)*, **272**, 506–510.

Elam, J. S., Taylor, A. B., Strange, R., Antonyuk, S., Doucette, P. A., Rodriguez, J. A., Hasnain, S. S., Hayward, L. J., Valentine, J. S., Yeates, T. O. & Hart, P. J. (2003). *Nature Struct. Biol.* **10**, 461–467.

Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* D**66**, 486–501.

Guo, Z. & Eisenberg, D. (2006). *Proc. Natl Acad. Sci. USA*, **103**, 8042–8047.

Harrington, D. J., Adachi, K. & Royer, W. E. (1997). *J. Mol. Biol.* **272**, 398–407.

Holmes, K. C., Popp, D., Gebhard, W. & Kabsch, W. (1990). *Nature (London)*, **347**, 44–49.

Janowski, R., Kozak, M., Jankowska, E., Grzonka, Z., Grubb, A., Abrahamson, M. & Jaskolski, M. (2001). *Nature Struct. Biol.* **8**, 316–320.

Knaus, K. J., Morillas, M., Swietnicki, W., Malone, M., Surewicz, W. K. & Yee, V. C. (2001). *Nature Struct. Biol.* **8**, 770–774.

Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
Liu, C., Sawaya, M. R. & Eisenberg, D. (2011). *Nature Struct. Mol. Biol.* **18**, 49–55.
Lomas, D. A. & Carrell, R. W. (2002). *Nature Rev. Genet.* **3**, 759–768.
McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
Merlino, A., Avella, G., Di Gaetano, S., Arciello, A., Piccoli, R., Mazzarella, L. & Sica, F. (2009). *Protein Sci.* **18**, 50–57.
Merlino, A., Ceruso, M. A., Vitagliano, L. & Mazzarella, L. (2005). *Biophys. J.* **88**, 2003–2012.
Merlino, A., Mazzarella, L., Carannante, A., Di Fiore, A., Di Donato, A., Notomista, E. & Sica, F. (2005). *J. Biol. Chem.* **280**, 17953–17960.
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.
Necas, D. & Klapetek, P. (2012). *Cent. Eur. J. Phys.* **10**, 181–188.
Nelson, R., Sawaya, M. R., Balbirnie, M., Madsen, A. Ø., Riekel, C., Grothe, R. & Eisenberg, D. (2005). *Nature (London)*, **435**, 773–778.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Pepys, M. B. (2006). *Annu. Rev. Med.* **57**, 223–241.
Pous, J., Mallorquí-Fernández, G., Peracaula, R., Terzyan, S. S., Futami, J., Tada, H., Yamada, H., Seno, M., de Llorens, R., Gomis-Rüth, F. X. & Coll, M. (2001). *Acta Cryst.* D**57**, 498–505.
Rousseau, F., Schymkowitz, J. & Serrano, L. (2006). *Curr. Opin. Struct. Biol.* **16**, 118–126.
Russo, N., Antignani, A. & D'Alessio, G. (2000). *Biochemistry*, **39**, 3585–3591.
Sambashivan, S., Liu, Y., Sawaya, M. R., Gingery, M. & Eisenberg, D. (2005). *Nature (London)*, **437**, 266–269.
Taverna, D. M. & Goldstein, R. A. (2002). *J. Mol. Biol.* **315**, 479–484.
Vitagliano, L., Adinolfi, S., Sica, F., Merlino, A., Zagari, A. & Mazzarella, L. (1999). *J. Mol. Biol.* **293**, 569–577.