# RSC Advances

## PAPER

# Solution structure of hyperactive type I antifreeze protein†

Luuk L. C. Olijve,[a] Tianjun Sun,[b] Theyencheri Narayanan,[c] Corinne Jud,[d] Peter L. Davies[b] and Ilja K. Voets*[a]

Antifreeze proteins (AFPs) protect freeze-intolerant fish species living in icy polar waters against freeze damage. A 34 kDa dimeric type I antifreeze protein (wfAFP 1h) with unusually high activity in comparison to all other antifreeze proteins in fish was recently discovered in the winter flounder. We have measured the size and shape of this hyperactive AFP by using small angle X-ray scattering. Our results show that wfAFP 1h adopts a long cylindrical shape with a length of 19 $\pm$ 2 nm and a diameter of 2.3 $\pm$ 0.2 nm, which means that wfAFP 1h does not form a fully extended helical dimer in solution. These findings call for a revision of the structural model of wfAFP 1h and the concept of a flat, threonine-rich ice-binding site extending down the length of the protein. Instead, the hyperactive nature of wfAFP 1h may be derived from a unique 3D arrangement of the helices —yet to be resolved— by which it is able to bind to ice surfaces.

## Introduction

The freezing point of seawater ($-1.9$ °C) is more than 1 °C below the equilibrium freezing point of the body fluids of fish surviving in icy polar seawaters. To protect themselves in this undercooled state from rapidly freezing upon contact with ice, these fish produce antifreeze plasma proteins that lower the freezing point by stopping the growth of ice crystals.[1] Antifreeze proteins (AFPs) have specific structural properties that enable them to adsorb to the surface of ice crystals thereby restricting the addition of water molecules to the growing ice surface. The ice surface is forced to grow with a thermodynamically unfavorable curvature, leading to a local depression of the non-equilibrium freezing point, which results in an arrest of further ice growth.[2] Over the years, a wide diversity of AFP structures have been found in many different organisms ranging from fish[3] to insects,[4] bacteria[5] and fungi.[6]

In 2004, Marshall and coworkers reported an unusually potent type I AFP isoform in winter flounder (wfAFP 1h), that was 10 to 100 times more active than the previously discovered type I AFP in the same fish species.[7] Circular dichroism measurements at 4 °C showed that this alanine-rich, 195 residue-long (16.7 kDa) protein is almost entirely $\alpha$-helical.[8] Analytical ultracentrifugation analysis and gel permeation chromatography furthermore indicated that this protein forms a dimer in solution and that it has an extreme asymmetry with an estimated axial ratio of $L/d = 18$.[9] Based on these data, Marshall and coworkers concluded that wfAFP 1h is a helical dimeric rod of 29 nm in length and 1.6 nm in diameter, consisting of two long helices associated side-by-side. A coiled-coil structure was ruled out based on the paucity of hydrophobic amino acid residues (*e.g.*, Leu, Ile, Val), thereby explaining the instability of the protein ($T_{denat} \sim 9$ °C).[8,10] Also, a coiled-coil structure would interfere with regularly positioned threonine residues appearing on a flat ice-binding site along the length of the protein. The complex was therefore modelled as an antiparallel dimer composed of two straight helices with relatively weak interactions between the alanine (Ala) side chains, *i.e.*, as an Ala-zipper, resulting in a closer packing with $L = 27.5$ nm and $d = 1.6$ nm giving an axial ratio of $L/d = 17$.[10] In this model both subunits could simultaneously engage in ice-binding.

To obtain reliable evidence on the exact dimensions and multimeric state of wfAFP 1h, we have performed small angle X-ray scattering (SAXS) experiments on the protein in solution. SAXS is a powerful tool for characterizing the structure of macromolecules in their native environment.[11] Coherent scattering of X-rays by randomly oriented proteins gives information on their size, shape and molar mass. In contrast to X-ray diffraction, SAXS does not require crystalline samples

*a*Laboratory for Macromolecular and Organic Chemistry and Institute for Complex Molecular Systems, Department of Chemical Engineering and Chemistry, Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands. E-mail: I.Voets@tue.nl
*b*Department of Biomedical and Molecular Sciences, Queen's University, Kingston, Ontario, K7L 3N6, Canada
*c*ESRF, 6 rue Jules Horowitz, F-38043, Grenoble, France
*d*Adolphe Merkle Institute, University of Fribourg, Route de l'Ancienne Papeterie, P.O. Box 209, CH-1723 Marly 1, Switzerland
† Electronic supplementary information (ESI) available: See DOI: 10.1039/c3ra22729d

which greatly broadens its application perspective, but reduces the attainable resolution.

The results presented here confirm the dimeric state of wfAFP 1h and provide accurate values for its dimensions. The length of wfAFP 1h as determined from the SAXS experiments is found to be smaller than hypothesized based on the previous model,[10] which implies that the helical dimer is arranged differently and not fully stretched out.

## Materials and methods

### Sample preparation

Recombinant wfAFP 1h was produced as previously described[12] with the following modification. Instead of ice affinity purification, three rounds of ammonium sulfate precipitation were used. This is an effective method for purification because wfAFP 1h precipitates from solution at a relatively low percent saturation of ammonium sulfate (28%) while most *Escherichia coli* proteins remain soluble. The wfAFP 1h pellet from the last precipitation was re-suspended and dialyzed overnight in 50 mM HEPES buffer (pH 6.5) to remove ammonium sulfate. Purity was estimated by SDS-PAGE, and the protein concentration was measured by amino acid analysis. Recombinant wfAFP 1h at 15 mg mL$^{-1}$ was stored at 4 °C and used within 2–3 weeks of its preparation.

### Small angle X-ray scattering

**Data acquisition and data reduction.** The synchrotron radiation X-ray scattering data were collected at the high-brilliance beamline ID02 of the ESRF in Grenoble, France,[13] operating at 12.46 keV. The scattering intensity was measured as a function of momentum transfer vector $q = 4\pi(\sin\theta)/\lambda$, where $\lambda = 0.1$ nm is the radiation wavelength and $2\theta$ is the scattering angle. Two sample-to-detector distances of 1.5 and 3 m were used to cover an angular range of $0.064 < q < 3.85$ nm$^{-1}$. Samples were measured in a polycarbonate (ENKI, KI-Beam) flow through capillary with a diameter of $d = 1.9$ mm kept in a temperature-controlled holder at $T = 5$ °C. The two-dimensional SAXS patterns were normalized to an absolute intensity scale using the calibrated detector response function, known sample-detector distance, and measured incident and transmitted beam intensities.[13] These normalized SAXS patterns were subsequently azimuthally averaged to obtain the one dimensional SAXS profiles. For each sample, 10 frames of 0.3 s were collected and averaged after checking for radiation damage. This corresponds to a total data collection time of 3 s per sample with a reduced flux of about $10^{12}$ photons s$^{-1}$. To obtain the protein scattering curve, the normalized background scattering profile of the buffer and polycarbonate cell was subtracted from the normalized sample scattering profiles. Finally, the absolute calibration of the scattering curves were verified using the known scattering cross-section per unit sample volume, dΣ/dΩ, of water, being $I(0) = 0.01665$ cm$^{-1}$ for $T = 5$ °C.[14]

**Data analysis.** Small angle X-ray scattering is a powerful tool to determine the dimensions and molar mass of proteins and protein complexes directly in solution. The primary require-

ment is an accurate measurement of the $q$-dependence of the scattering intensity of the sample of interest, the buffer, and a calibration standard with known scattering cross-section such as water[14] to compute the differential scattering cross-section per unit sample volume, dΣ/dΩ. This quantity gives direct access to the size, shape, and average molar mass of the protein (complex) under investigation according to

$$\frac{d\Sigma}{d\Omega} = KcM_w S(q)P(q) \tag{1}$$

with the difference in scattering contrast with the solvent, $K$, the weight concentration, $c$, the molecular weight, $M_w$, and interference effects arising from the structure within the sample, *i.e.*, interparticle and intraparticle interference represented by the structure factor, $S(q)$, and the form factor, $P(q)$.

First, the scattering data were analyzed using a Guinier approximation to extract the radius of gyration, $R_g$, and the forward scattering intensity, $I_0$, which is dΣ/dΩ($q \rightarrow 0$). For monodisperse globular proteins, the Guinier approximation is valid for $qR_g \leqslant 1.3$ and $R_g$ and $I_0$ were determined from the slope and y-intercept of the Guinier plot $ln(d\Sigma/d\Omega(q))$ *vs.* $q^2$ using PRIMUS from the ATSAS software package.[15] For a rod-like particle, a modified Guinier approximation can be used in the $q$-range where dΣ/dΩ($q$) $\propto q^{-\alpha}$ with $\alpha \sim 1$ to extract the cross-sectional radius of the cylinder from the determined cross-sectional radius of gyration, with $R_{cs} = \sqrt{2}R_{acs}$.

Using the obtained radius of gyration and the cross-sectional radius from the Guinier approximation, the length of a cylinder can be calculated by

$$R_g^2 = \frac{R_{cs}^2}{2} + \frac{L^2}{12} \tag{2}$$

Subsequently, $I_0 = $ dΣ/dΩ($q \rightarrow 0$) was used to calculate the molar mass of the protein (complex), which is listed in Table 1, according to

$$M_{SAXS} = I_0 \frac{N_{av}}{c(\Delta\rho\bar{v})^2} \tag{3}$$

with the molecular weight $M_{SAXS}$ in g mol$^{-1}$, the forward scattering intensity $I_0$ in cm$^{-1}$, concentration $c$ in g cm$^{-3}$, Avogadro's number $N_{av}$, the scattering length density difference $\Delta\rho$ in cm$^{-2}$ ($\rho_{protein} - \rho_{H_2O}$, where $\rho_{protein} = 1.25 \times 10^{11}$ cm$^{-2}$ and $\rho_{H_2O}$, = $9.44 \times 10^{10}$ cm$^{-2}$)[16] and the partial specific volume of the protein in solution $\bar{v} = 0.7302$ in cm$^3$ g$^{-1}$ calculated using SEDNTERP.[17]

Finally, we have analyzed the scattering profiles in the entire recorded $q$-range with various form factor models using the software package SASfit[18] assuming that wfAFP 1h was measured in an ideal dilute solution where interparticle interactions can be neglected (*i.e.*, $S(q) = 1$). Two form factors were used; one describing a cylindrical object[19] and the other describing a worm-like chain model according to the equations reported by Pedersen and co-workers.[20]

**Table 1** Data-collection and scattering-derived parameters

| | |
|---|---|
| **Data-collection parameters** | |
| Beamline | ID02 (ESRF, Grenoble, France) |
| Operating energy (keV) | 12.46 |
| Flux (photons s$^{-1}$) | $10^{12}$ |
| Wavelength (nm) | 0.1 |
| $q$ range (nm$^{-1}$) | 0.064–3.85 |
| Sample-to-detector distance (m) | 1.5 and 3 |
| Total exposure time (s) | 3 |
| Concentration (mg mL$^{-1}$) | 2–15 |
| $T$ (°C) | 5 |
| Sample holder | Poly carbonate flow-through capillary, $\varnothing$ = 1.9 mm |
| **Structural parameters**$^a$ | |
| $I_0$ (cm$^{-1}$) [Guinier] | 0.055 ± 0.002 |
| $R_g$ (nm) [Guinier] | 5.0 ± 0.09 |
| $I(0)$ (cm$^{-1}$) [$P(r)$] | 0.0557 ± 0.001 |
| $R_g$ (nm) [$P(r)$] | 5.7 ± 0.1 |
| $D_{max}$ (nm)$^b$ | 22 |
| **Molecular mass determination** | |
| Partial specific volume (cm$^3$ g$^{-1}$) | 0.7302 |
| Contrast ($\Delta\rho \times 10^{10}$ cm$^{-2}$) | 3.060 |
| Molecular mass $M_w$ from $I(0)$ (Da) | 34700 ± 1100 |
| Calculated monomeric $M_w$ from sequence (Da) | 16685 |
| **Software employed** | |
| Primary data reduction | SAXSutilities |
| Data processing | PRIMUS |
| *Ab initio* analysis | DAMMIF |
| Validation and averaging | DAMAVER |
| Computation of model intensities | CRYSOL |
| Rigid-body modelling | SASREF |
| Graphical representation | PyMol and SITUS |

$^a$ Reported for 1.9 mg mL$^{-1}$ measurement. $^b$ $D_{max}$ is a model parameter in the $P(r)$ calculation and no variance associated with $D_{max}$ can be calculated.

## Molecular shape reconstruction

To derive information on the solution structure of the protein, the molecular shape was reconstructed using simulated annealing methods. First, the radial distribution function (RDF), $P(r)$, which describes the probable frequency of interatomic vector lengths ($r$) within the scattering particle, was obtained upon indirect inverse Fourier transformation of the scattering data using GNOM.[21] The maximum linear dimension ($D_{max}$) was set to approximately $3 \times R_g$ and adjusted to give the best fit to the experimental data. The RDF was considered to be zero at $r$ = 0 Å and approaches zero at $D_{max}$.

The GNOM output files were used as input for simulated annealing calculations over the range $0.08 < q < 3.8$ nm$^{-1}$ using the online version of DAMMIF.[22] For each scattering curve, 10 independent bead models were generated without predefined shape or symmetry. The 10 different models were superimposed using DAMSEL and DAMSUP. Next, DAMAVER was used to average the aligned models and compute the probability map.[23] Finally, DAMFILT was used to filter the averaged model to give a structure that has high densities on the probability map.

## Results and discussion

### Guinier and form factor analysis

Fig. 1A displays the small angle X-ray scattering profiles obtained at four different concentrations (2–15 mg mL$^{-1}$) at the high-brilliance beamline ID02 of the European Synchrotron Radiation Facility (ESRF, Grenoble, France). The accessible $q$-range covers the full SAXS profile of wfAFP 1h, which extends from the Guinier regime at low $q$-values up to the form factor oscillation at high $q$-values with a slope of $\sim q^{-1}$ typical of cylindrical structures at intermediate $q$-values. A Guinier approximation (Fig. 1B) was used to obtain the radius of gyration $R_g$ and molecular weight $M_w$. The estimated molecular weight of $M_w$ = 34.7 ± 2.5 kDa corresponds to the formation of a dimer in solution. By using a modified Guinier approximation for a rod (Fig. 1C) the cross-sectional radius $R_{cs}$ is obtained. Using these structural parameters, the length of the protein complex was calculated using eqn (2), confirming that wfAFP 1h adopts a highly elongated structure with a length $L$ = 17.6 ± 0.9 nm and a cross-sectional radius of $R_{cs}$ = 1.12 ± 0.01 nm, giving an axial ratio of $L/d$ = 7.9 ± 0.4 (Table 2).

The experimental data are also well described by a rigid cylindrical form factor (Fig. 1A and Table 3) and a worm-like chain model (Fig. S1 and Table S1 of the ESI†). The cylindrical model assumes a rigid cylinder with uniform scattering length density. The structural parameters extracted from fitting the data with a cylindrical form factor are the cross-sectional radius $R_{cs}$ = 1.26 ± 0.02 nm that is close to the $R_{cs}$ obtained from the Guinier analysis, and the length $L$ = 19.7 ± 0.2 nm, which is slightly larger.
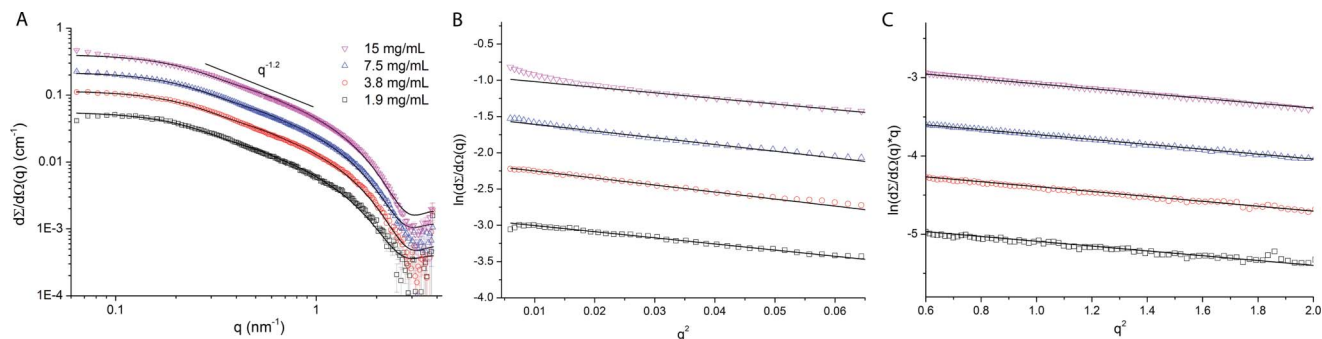
**Fig. 1** (A) Small angle X-ray scattering profiles of 2–15 mg mL$^{-1}$ wfAFP 1h in 50 mM HEPES at pH 6.5 and $T = 5$ °C. The experimental data (symbols) are well described by a cylindrical form factor (solid lines). (B) Fit of the Guinier approximation of the experimental data of wfAFP 1h. The used $q$-ranges are $0.08 \leqslant q \leqslant 0.23$ nm$^{-1}$ (1.9 mg mL$^{-1}$), $0.09 \leqslant q \leqslant 0.21$ nm$^{-1}$ (3.8 mg mL$^{-1}$), $0.13 \leqslant q \leqslant 0.21$ nm$^{-1}$ (7.5 mg mL$^{-1}$) and $0.17 \leqslant q \leqslant 0.23$ nm$^{-1}$ (15 mg mL$^{-1}$). (C) Fit of a modified Guinier approximation for a rod to the experimental data. The used $q$-ranges are $0.65 \leqslant q \leqslant 1.49$ nm$^{-1}$ (1.9 mg mL$^{-1}$), $0.65 \leqslant q \leqslant 1.62$ nm$^{-1}$ (3.8 mg mL$^{-1}$), $0.65 \leqslant q \leqslant 1.57$ nm$^{-1}$ (7.5 mg mL$^{-1}$) and $0.73 \leqslant q \leqslant 1.50$ nm$^{-1}$ (15 mg mL$^{-1}$).

## Structural model of the protein complex

From the results of the Guinier and form factor analyses it is evident that wfAFP 1h has a highly extended shape with a length of $L = 19 \pm 2$ nm and diameter of $d = 2.3 \pm 0.2$ nm. These dimensions imply that wfAFP 1h cannot form a dimer of two fully extended helices, since the latter would correspond to a length of $\sim$27.5 nm. To derive information on the shape of the dimer, we have reconstructed the *ab initio* shape of the protein from simulated annealing calculations using DAMMIF, which features an unlimited and adapting search volume that avoids boundary effects, especially in highly elongated objects.[21] The averaged and filtered volumetric map of wfAFP 1h shows a rod-like shape with a length of 18.5 nm (Fig. 2D), resembling the dimensions found in the Guinier and form factor analysis. Given the dimensions and high helical content of wfAFP 1h, it is evident that the molecular structure of this hyperactive antifreeze protein is unlike any antifreeze protein found so far. The β-helical folds of hyperactive AFPs from insects display a flat β-sheet ice-binding surface with regularly positioned threonine residues.[4] Hence, a likely arrangement of the dimer would be two helices associated side-by-side, enabling a uniform presentation of threonine residues along the length of the protein.[10] Yet, this structure is not consistent with the SAXS results, which means that a different conformation of the helices might confer the hyperactive antifreeze activity.

To get more insight into the 3D arrangement of the helices in the protein complex of wfAFP 1h, we have performed rigid body modelling using SASREF.[25] SASREF uses multiple rigid subunits with known atomic coordinates to perform quaternary structure modelling of a protein complex against the experimental data. Since no crystallographic data of the subdomains in the protein complex of wfAFP 1h is available, we have constructed two long α-helices with a length of 18.5 nm and two short helices to account for the remaining number of residues of the protein complex. These subunits and contact conditions (Fig. 3A) were provided as input for SASREF. The result of the rigid body modelling (Fig. 3C) seems consistent with the results of the more coarse-grained modelling approach and comes very close to the hypothesized structure in which the protein complex is able to display regularly positioned threonine residues on its ice-binding face, enabling a large protein surface area to engage in ice-binding which could explain the hyperactive nature.

However, we need to draw attention to the limitations of the resulting structural model. Firstly, the initial contact conditions influence the SASREF results. Therefore, about six contact conditions per complex were set and various contact conditions were tested (Fig. S3–5 of the ESI†). Reducing the number of contact conditions resulted in one of the subunits losing contact with the complex. Secondly, we propose the structural model in Fig. 3C on the basis of additional information rather than the goodness of fit of the different SASREF models. Thirdly, the subunits are adapted from PDB

**Table 2** Structural parameters obtained from a (modified) Guinier analysis. $R_g$ and $M_w$ are obtained using a Guinier approximation (Fig. 1B), $R_{cs}$ by using a modified Guinier approximation for a rod (Fig. 1C) and $L$ was calculated using eqn (2)

| Conc. (mg mL$^{-1}$) | $R_g$ (nm) | $M_w$ (kDa) | $R_{cs}$ (nm) | $L$ (nm) | $L/d$ |
|---|---|---|---|---|---|
| 1.9 | 5.0 | 34.7 | 1.11 | 17.3 | 7.8 |
| 3.8 | 5.4 | 37.3 | 1.12 | 18.5 | 8.2 |
| 7.5 | 5.3 | 35.4 | 1.12 | 18.1 | 8.1 |
| 15 | 4.8 | 31.4 | 1.11 | 16.4 | 7.4 |
| Average | 5.1 | 34.7 | 1.12 | 17.6 | 7.9 |
| Std | 0.3 | 2.5 | 0.01 | 0.9 | 0.4 |

**Table 3** Structural parameters obtained after fitting the SAXS data with a cylindrical form factor. The radius of gyration is calculated from the obtained $R_{cs}$ and $L$ using eqn (2)

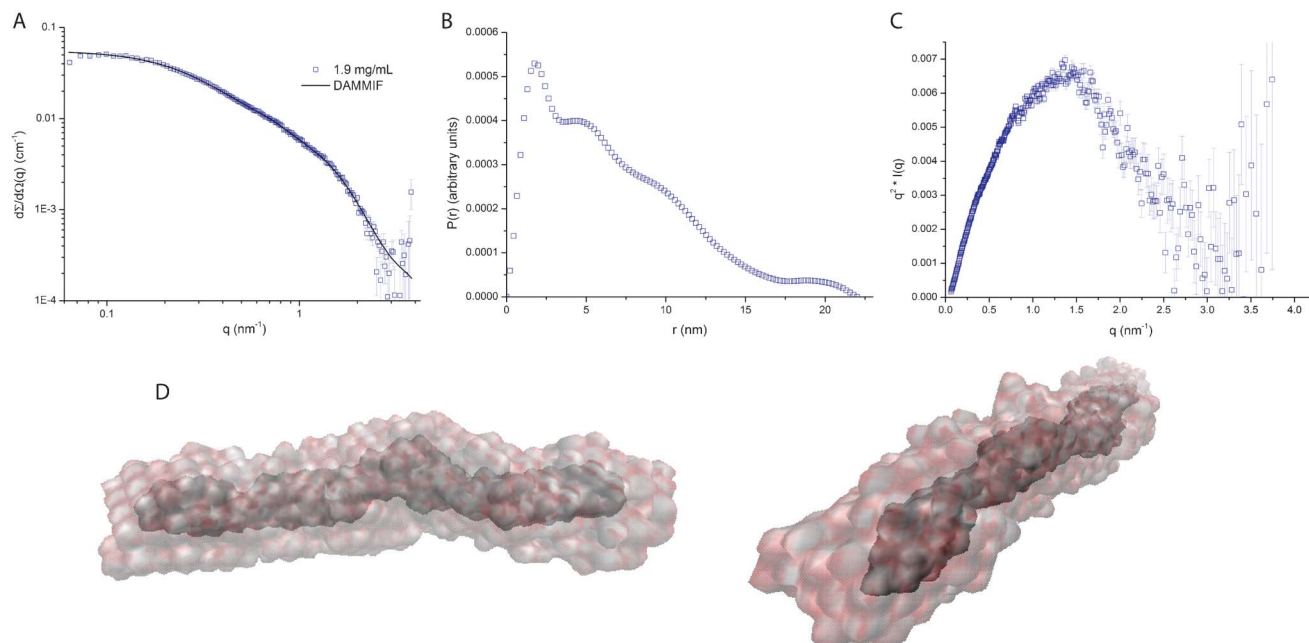| Conc. (mg mL$^{-1}$) | $R_g$ (nm) | $R_{cs}$ (nm) | $L$ (nm) | $L/d$ | $\chi^2$ |
|---|---|---|---|---|---|
| 1.9 | 5.8 | 1.28 | 19.9 | 7.7 | 0.98 |
| 3.8 | 5.7 | 1.23 | 19.5 | 7.9 | 0.85 |
| 7.5 | 5.8 | 1.26 | 19.8 | 7.9 | 0.98 |
| 15 | 5.7 | 1.25 | 19.6 | 7.8 | 0.92 |
| Average | 5.8 | 1.26 | 19.7 | 7.8 | 0.93 |
| Std | 0.1 | 0.02 | 0.2 | 0.1 | 0.06 |

**Fig. 2** (A) Fit of simulated annealing with DAMMIF, (B) radial distribution function (RDF) obtained after inverse Fourier transformation of the experimental data, (C) Kratky plot showing a parabolic shape indicative of a folded state of the protein,[24] and (D) 10 calculations with DAMMIF were performed to produce the average and filtered volumetric map of wfAFP 1h, showing a rod-like shape with a length of 18.5 nm (final $\chi^2$ = 2.57). SITUS was used to create the volumetric maps of the bead models.

entries of other proteins. Hence, the validity of the model needs to be tested with additional biochemical experiments.

### Protein complex has unique molecular structure

Given the extraordinary shape of the protein that is very distinct from any other AFP found so far, we have searched in the Protein Database (PDB) for proteins with similar shape based on four criteria: (1) high α-helicity, (2) dimeric state, (3) number of residues close to 195 as found for wfAFP 1h and (4) length and axial ratio. We have compared the experimental data with theoretical scattering profiles of 10 different crystal structures, and computed the theoretical scattering curves from the electron density profiles of the selected proteins using the known atomic coordinates deposited in the PDB by using CRYSOL.[26] The theoretical scattering curves were then compared to the experimental scattering data (see Fig. S6–7 of the ESI†). Only a handful of the selected α-helix rich dimers of the protein database have a structure that is comparable yet not identical to that of wfAFP 1h. Most of the matching protein structures originate from the cytoskeleton or muscle. Clearly,
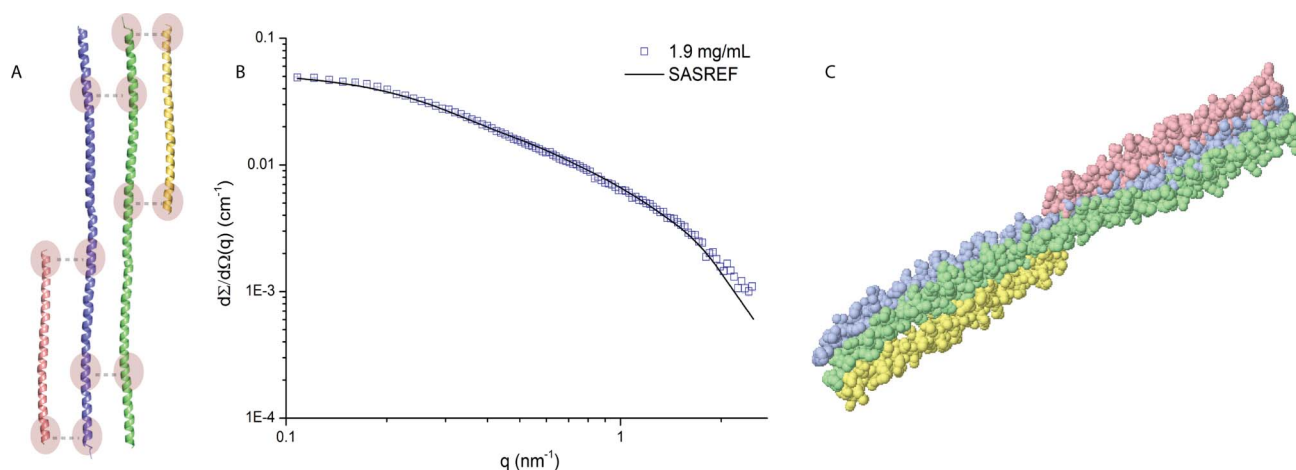


**Fig. 3** (A) Subdomains with provided contact conditions used for the rigid body modelling. The subdomains were constructed from the PDB entry 3VEM. (B) Fit of the model to the experimental data (final $\chi^2$ = 3.79), using P1 symmetry. Data range used for the rigid body modelling 0.08 ⩽ q ⩾ 2.5 nm$^{-1}$. (C) Resulting structural model from the rigid body modelling.

the protein complex adopts a structure with high asymmetry, which is unique for a plasma protein.

## Conclusions

Structural data form the basis for understanding the hyperactive nature of wfAFP 1h. Our SAXS experiments provide the first reliable and independent estimation of the length, diameter and molecular weight of wfAFP 1h in solution. We find that wfAFP 1h self-assembles into a dimer with a highly extended (*i.e.*, non-globular) conformation that is unusual for a plasma protein and mostly found in muscle and cytoskeleton proteins. The protein complex has a rigid cylindrical shape with a length and diameter of $19 \pm 2$ nm and $2.3 \pm 0.2$ nm, respectively. This demonstrates unequivocally that the two $\alpha$-helices in the dimer are not fully extended (*i.e.*, $L = 27.5$ nm) contrary to what was hypothesized previously, irrespective of whether they are arranged in a coiled-coil or parallel conformation. Furthermore, if the helical dimers would be held together *via* an interaction of the alanine side chains, the packing would be too close to match the experimentally observed diameter suggesting that wfAFP 1h adopts a more complex conformation. Further evidence from biochemical methods and/or diffraction studies is necessary to gain more insight into the 3D arrangement of the two monomers in the protein complex and the positioning of regularly placed threonine residues at the ice-binding site.

## Acknowledgements

## References

1 A. L. DeVries and D. E. Wohlschlag, *Science*, 1969, **163**, 1073–1075.

2 J. A. Raymond and A. L. DeVries, *Proc. Natl. Acad. Sci. U. S. A.*, 1977, **74**, 2589–2593.

3 G. L. Fletcher, C. L. Hew and P. L. Davies, *Annu. Rev. Physiol.*, 2001, **63**, 359–390.

4 (*a*) B. L. Pentelute, Z. P. Gates, V. Tereshko, J. L. Dashnau, J. M. Vanderkooi, A. A. Kossiakoff and S. B. Kent, *J. Am. Chem. Soc.*, 2008, **130**, 9695–9701; (*b*) S. P. Greather, M. J. Kuipers, S. M. Gagné, V. K. Walker, Z. Jia, B. D. Sykes and P. L. Davies, *Nature*, 2000, **406**, 325–328.

5 C. P. Garnham, R. L. Campbell and P. L. Davies, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**, 7363–7367.

6 J. H. Lee, A. K. Park, H. Do, K. S. Park, S. H. Moh, Y. M. Chi and H. J. Kim, *J. Biol. Chem.*, 2012, **287**, 11460–11468.

7 C. B. Marshall, G. L. Fletcher and P. L. Davies, *Nature*, 2004, **429**, 153.

8 S. Y. Gauthier, C. B. Marshall, G. L. Fletcher and P. L. Davies, *FEBS J.*, 2005, **272**, 4439–4449.

9 C. B. Marshall, A. Chakrabartty and P. L. Davies, *J. Biol. Chem.*, 2005, **280**, 17920–17929.

10 L. A. Graham, C. B. Marshall, F.-H. Lin, R. L. Campbell and P. L. Davies, *Biochemistry*, 2008, **47**, 2051–2063.

11 (*a*) D. I. Svergun and M. H. J. Koch, *Rep. Prog. Phys.*, 2003, **66**, 1735–1782; (*b*) C. D. Putnam, M. Hammel, G. L. Hura and J. A. Tainer, *Q. Rev. Biophys.*, 2007, **40**, 191–285.

12 F. H. Lin, T. Sun, G. L. Fletcher and P. L. Davies, *Protein Expression Purif.*, 2012, **82**, 75–82.

13 M. Sztucki, E. Di Cola and T. Narayanan, *J. Appl. Crystallogr.*, 2010, **43**, 1479–1487.

14 (*a*) *Small-angle X-ray Scattering*, ed. O. Glatter and O. Kratky, Academic Press, London, 1982; (*b*) E. Mylonas and D. I. Svergun, *J. Appl. Crystallogr.*, 2007, **40**, s245–s249.

15 P. V. Konarev, V. V. Volkov, A. V. Sokolova, M. H. J. Koch and D. I. Svergun, *J. Appl. Crystallogr.*, 2003, **36**, 1277–1282.

16 Online scattering length density calculator of NIST http://www.ncnr.nist.gov/resources/sldcalc.html.

17 D. A. Jacques and J. Trewhella, *Protein Sci.*, 2010, **19**, 642–657.

18 SASfit software package for fitting small angle scattering curves http://kur.web.psi.ch/sans1/SANSSoft/sasfit.html.

19 *Neutron, X-rays and light: scattering methods applied to soft condensed matter*, ed. P. Lindner and T. Zemb, Elsevier, Amsterdam, 2002.

20 J. S. Pedersen and P. Schurtenberger, *Macromolecules*, 1996, **29**, 7602–7612.

21 D. I. Svergun, *J. Appl. Crystallogr.*, 1992, **25**, 495–503.

22 D. Franke and D. I. Svergun, *J. Appl. Crystallogr.*, 2009, **42**, 342–346.

23 V. V. Volkov and D. I. Svergun, *J. Appl. Crystallogr.*, 2003, **36**, 860–864.

24 D. A. Jacques, J.M. Guss, D. I. Svergun and J. Trewhella, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2012, **68**, 620–626.

25 M. V. Petoukhov and D. I. Svergun, *Biophys. J.*, 2005, **89**, 1237–1250.

26 D. I. Svergun, C. Barberato and M. H. J. Koch, *J. Appl. Crystallogr.*, 1995, **28**, 768–773.