doi:10.1016/j.jmb.2006.09.091 J. Mol. Biol. (2007) 365, 411–424

JMB

Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

Insights into Molecular Plasticity of Choline Binding Proteins (Pneumococcal Surface Proteins) by SAXS

Rubén M. Buey[1], Begoña Monterroso[2], Margarita Menéndez[2] Greg Diakun[3], Pablo Chacón[1], Juan Antonio Hermoso[2] and J. Fernando Díaz[1]*

[1]*Centro de Investigaciones Biológicas, Consejo Superior de Investigaciones Científicas Ramiro de Maeztu 9, 28040 Madrid, Spain*

[2]*Instituto de Química-Física Rocasolano, Consejo Superior de Investigaciones Científicas Serrano 119, 28006 Madrid, Spain*

[3]*CCLRC Daresbury Laboratory Warrington WA4 4AD UK*

**Corresponding author*

Phosphocholine moieties decorating the pneumococcal surface are used as a docking station for a family of modular proteins, the so-called choline binding proteins or CBPs. Choline recognition is essential for CBPs function and may also be a determinant for their quaternary structure. There is little knowledge about modular arrangement or oligomeric structures in this family. Therefore, we have used the small angle X-ray scattering (SAXS) technique combined with analytical ultracentrifugation in order to model the three-dimensional envelope of two highly different CBPs: the phage encoded Cpl-1 lysozyme and the pneumococcal phosphorylcholine esterase Pce. Both enzymes have an N-terminal catalytic module and a C-terminal choline-binding module (CBM) that attaches them to the bacterial surface and comprises six and ten sequence repeats in Cpl-1 and Pce, respectively. SAXS experiments have shown an inherent conformational plasticity in Cpl-1 that accounts for the different relative position of these regions in the solution and crystal structures. Dimerization of Cpl-1 upon choline binding has been also visualised for the first time, and monomer–monomer interactions take place through the first CBR where a non-canonical choline binding site has now been identified. This mode of association seems to be independent of the absence or presence of the Cpl-1 catalytic module and reveals that the arrangement of the monomers differs from that previously found in the isolated CBM dimer of pneumococcal LytA amidase. In contrast, Pce displays the same modular disposition in the solution and crystal structures, and remains almost invariant upon choline binding. The present results suggest that protein dimerization and duplication of CBRs may be alternative but not equivalent ways of improving cell wall recognition by CBPs, since they provide different interaction geometries for choline residues present in (lipo)teichoic acids.

Keywords: choline binding proteins; three-dimensional solution structure; protein flexibility; small angle X-ray scattering; analytical ultracentrifugation

Introduction

Streptococcus pneumoniae causes life-threatening diseases such as pneumonia, bacteremia and meningitis worldwide.[1] This human pathogen has a unique nutritional requirement for choline,[2] which is incorporated into the cell wall[3] as covalently linked components of teichoic acids and also of membrane bound lipoteichoic acids.

Choline moieties are specifically recognised by a set of surface proteins, the so-called choline binding proteins (CBPs), attaching them to the bacterial envelope through non-covalent interactions. The CBP

Abbreviations used: CBM(s), choline binding module(s); CBR(s), choline binding repeat(s); C-Cpl-1, isolated choline binding module of Cpl-1 lysozyme; C-LytA, isolated choline binding module of LytA amidase; Cryo-EM, cryo-electromicrospy; R_s, Stokes' radius; SAXS, small angle X-ray scattering.

E-mail address of the corresponding author: fer@cib.csic.es

family includes pneumococcal cell wall murein hydrolases[4] and virulence factors involved in cellular adhesion and colonization[5–11] and also some bacteriophage encoded enzymes.[12] All CBPs are modular proteins composed of at least two clearly separated modules: the functional one, usually located at the N terminus, and the choline binding module (CBM) that selectively targets the enzyme to the cell wall and is made up of homologous repeats of 17–23 residues, called choline binding repeats (CBRs), in tandem disposition. One of the best characterised murein hydrolases encoded by pneumococcal bacteriophages (endolysins) is the Cpl-1 lysozyme encoded by bacteriophage Cp-1.[13]

Choline binding induces dimerization of CBPs with murein hydrolase activity harbouring, as the pneumococcal LytA amidase or Cpl-1 lysozyme, CBMs formed by six repeats and a short C-terminal tail.[14–17] Dimerization may be related to lytic activity regulation, since LytA mutants showing an impaired dimerization display a significant reduction (>90%) in catalytic activity.[15,18] On the other hand, modularity plays an important role in *in vivo* activity of CBPs.[19,20]

Until very recently, only the three-dimensional structures of the choline-bound dimer of C-LytA (the isolated CBM of LytA) and the Cpl-1 monomer[21,22] were solved. Despite the high sequence similarity displayed by the CBMs of both proteins, important structural differences were found between them in the crystal structures. In contrast, the three-dimensional structure of the full-length pneumococcal phosphorylcholine esterase (Pce),[20] a CBP involved in remodelling the distribution of phosphorylcholine moieties on the cell wall,[23,24] shows a CBM made with ten CBRs, the same kind of global fold as C-LytA.

Here, we have used small angle X-ray scattering (SAXS) for modelling the solution three-dimensional structures of Cpl-1 and Pce in their free and choline-bound states, with the final aim of widening the structural knowledge of the CBP family of proteins, including choline-induced conformational changes and flexibility of solution structures. SAXS is a powerful methodology to investigate domain organization in modular proteins or to monitor different conformational states when changes in size or shape are involved. The reliability of the SAXS-derived *ab initio* models has been widely supported by a large number of works (see Koch *et al.*[25] for a review), and our approach has also been validated comparing the hydrodynamic features of the proposed models with those measured by analytical ultracentrifugation under similar experimental conditions. The proposed model for the Cpl-1 dimer formed upon choline binding is the first full-length three-dimensional structure of a choline-bound dimeric lysin. In contrast with the monomer/dimer equilibrium exhibited by Cpl-1 and its isolated CBM (C-Cpl-1) upon choline binding, we show here that Pce is a monomer in the free and in the choline-bound states.

Results

SAXS model for the monomer of Cpl-1

Cpl-1 crystal structure (Figure 1(a)) comprises a catalytic module, consisting of an irregular alpha/beta $(\beta\alpha)_5\beta_3$ barrel, and a CBM arranged in two different structural regions: CI, a left-handed superhelix configuring the functional choline binding sites, and CII, a C terminus beta sheet directly interacting with the catalytic module through a

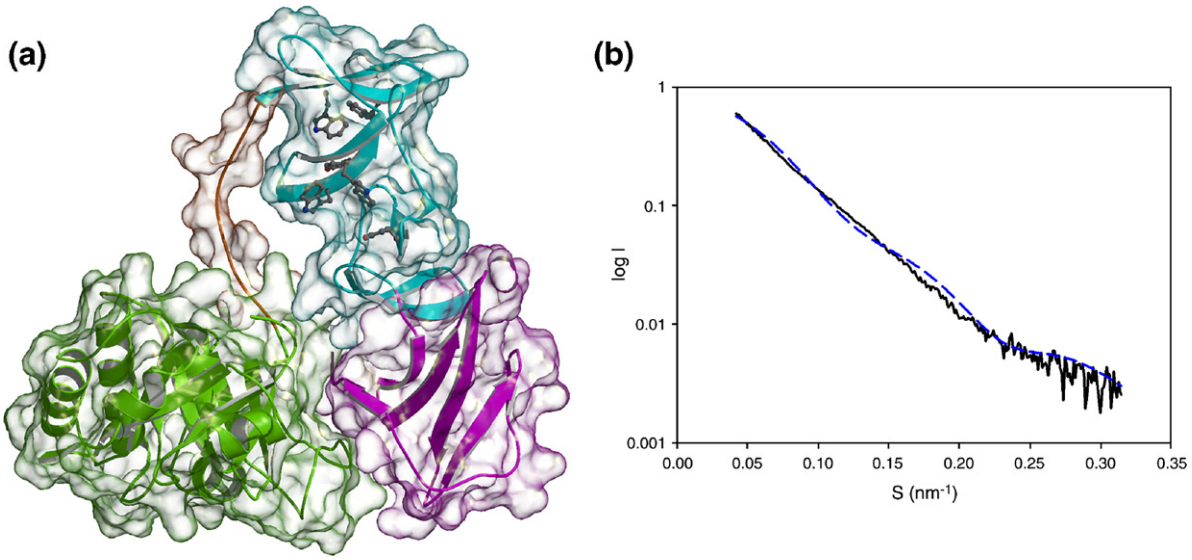

Figure 1. High-resolution crystal structure of Cpl-1 and experimental and theoretical SAXS profiles. (a) Ribbon diagram showing the catalytic module (green), the CI (blue) and CII (purple) structural domains of the CBM, and the connecting linker (orange) of the crystal structure. The molecular surface is also represented. (b) Superimposition of the experimental SAXS profile of Cpl-1 (black continuous line) to the theoretical one, generated by CRYSOL for the crystal structure (blue broken line).

large hydrophobic surface.[22] Both modules are connected by a negatively charged, long extended linker. In solution, and at protein concentrations used for SAXS measurements, Cpl-1 behaves as a monomer ($s_{20,w}$ of 3.17(±0.02) S; Table 1) that self associates into a dimer upon choline binding ($s_{20,w}$ of 4.33(±0.02) S; Table 1). However, only the high-resolution crystal structure of the Cpl-1 monomer is currently available.[22]

The experimental SAXS profile of the Cpl-1 monomer in the absence of choline (continuous black line) is compared in Figure 1(b) with the theoretical SAXS profile of the Protein Data Bank (PDB) structure (broken blue line) calculated by CRYSOL. Discrepancies between them are easily noted (Chi-square = 7.57). The theoretical SAXS profile of the PDB structure shows a prominent shoulder around 0.18 nm^{-1} which becomes wider and shifts towards a lower s value (0.14 nm^{-1}) in the experimental profile. Since this peak probably arises from the mean distance between Cpl-1 modules, its displacement and widening indicates a longer distance between the catalytic and the choline binding modules in the solution structure, suggesting that Cpl-1 modular arrangement in solution might be slightly different from that previously found in the crystal. This behaviour indicates that Cpl-1 structure exhibits some inherent flexibility that allows certain reorientation of its two modules.

To explore this point, Cpl-1 flexibility (i.e. conformational sampling) was computationally studied using CONCOORD,[26] a program that has been shown to be a useful method for sampling protein conformations. It successfully predicts concerted collective atom movements, saving most of the extraordinary computational costs of other methods like classical molecular dynamics simulations.[27–30] Since the CONCOORD method samples the conformational space of the protein fulfilling a set of interatomic distance limits, predicted on the interaction strength of the involved atoms, it is unlikely that it could find a structure largely different from the starting one, as that required to show a change visible by SAXS. However, it is a very powerful tool to analyse the structural flexibility and to find the hinge regions of protein structures.

The output of CONCOORD was further analysed by essential dynamics (ED) to calculate the largest amplitude concerted fluctuations around the Cpl-1 crystal structure, as described in Materials and Methods. They define several rotating/twisting and opening/closing movements of three collective-moving atom groups that coincide with the three domains defined in the crystal structure (Figure 2(a) and (b)). These movements change the relative position of the Cpl-1 modules while the hydrophobic contacts at the interface between the catalytic module and the CII domain are preserved. We used DYNDOM to analyse the collective atom movements described by the three first eigenvectors (derived from the CONCOORD/ED analysis), by comparing the structures of the two extreme projections of the whole trajectory onto each individual eigenvector (see Materials and Methods). The eigenvalues of the three first eigenvectors account for 79% of total Cpl-1 motions, enough to define the "essential" conformational subspace of a protein containing only a few degrees of freedom.[31,32] Collective-moving atom groups predicted by DYNDOM (Table 2) are, again, in close agreement with the three structural regions defined in the crystallographic structure.[22] The predicted bending residues of the inter-domain motions are similar for the three first eigenvectors and comprise the linker residues and those placed at the interface of CI and CII in the CBM (Table 2).

The first eigenvector describes a twist motion between CI and CII regions and a twisting between the former and the catalytic domain (Figure 2(a)). The second eigenvector describes a closure motion between CI and CII and a twisting between the former and the catalytic domain (Figure 2(b)). The third eigenvector (not shown) describes the closure motions between the catalytic domain and every domain (CI and CII) of the CBM.

Table 1. Experimental sedimentation coefficients, maximum distance, radius of gyration and theoretical sedimentation coefficients in the absence and in the presence of 40 mM choline

		SAXS		SV[a]		SAXS envelope[b]			Docked PDB[c]			Crystal PDB[d]		
	v ($ml\ g^{-1}$)	D_{max} (nm)	R_g (nm)	$s_{20,w}$ (S)	R_s (nm)	$s_{20,w}$ (S)	R_s (nm)	R_g (nm)	$s_{20,w}$ (S)	R_s (nm)	R_g (nm)	$s_{20,w}$ (S)	R_s (nm)	R_g (nm)
C-Cpl-1	0.7230	–	–	1.81	2.22	–	–	–	–	–	–	–	–	–
C-Cpl-1/choline	0.7230	9.5	3.17	2.58	3.17	2.32	3.20	3.53	2.46	3.32	3.26	–	–	–
Cpl-1	0.7248	10.0	3.05	3.17	3.01	3.04	3.14	2.97	3.05	3.13	2.68	3.16	3.02	2.42
Cpl-1/choline	0.7248	15.2	4.82	4.33	4.40	4.54	4.20	4.68	4.22	4.52	5.06	4.72	4.04	3.43
												5.15	3.71	2.82
Pce	0.7240	14.3	4.07	3.73	4.54	4.48	3.78	4.01	–	–	–	–	–	–
Pce/choline	0.7240	15.5	4.30	3.64	4.65	4.46	3.80	4.10	–	–	–	–	–	–
Pce_Δ55	0.7240	–	–	4.15	3.73	–	–	–	–	–	–	–		–
PceΔ55/choline	0.7240	–	–	4.15	3.73	–	–	–	3.95	3.93	3.93	3.95	3.93	3.93

[a] Sedimentation velocity data.
[b] Calculated (HYDRO) with bead models from SAXS envelope.
[c] Calculated (HYDROPRO) using coordinates of modified PDB model docked into SAXS-maps.
[d] Calculated (HYDROPRO) using the crystal PDB coordinates. Other parameters used for calculations of $s_{20,w}$: Temperature = 293 K, solvent density = 0.99823 g/cm^3; solvent viscosity = 0.01002 poises.

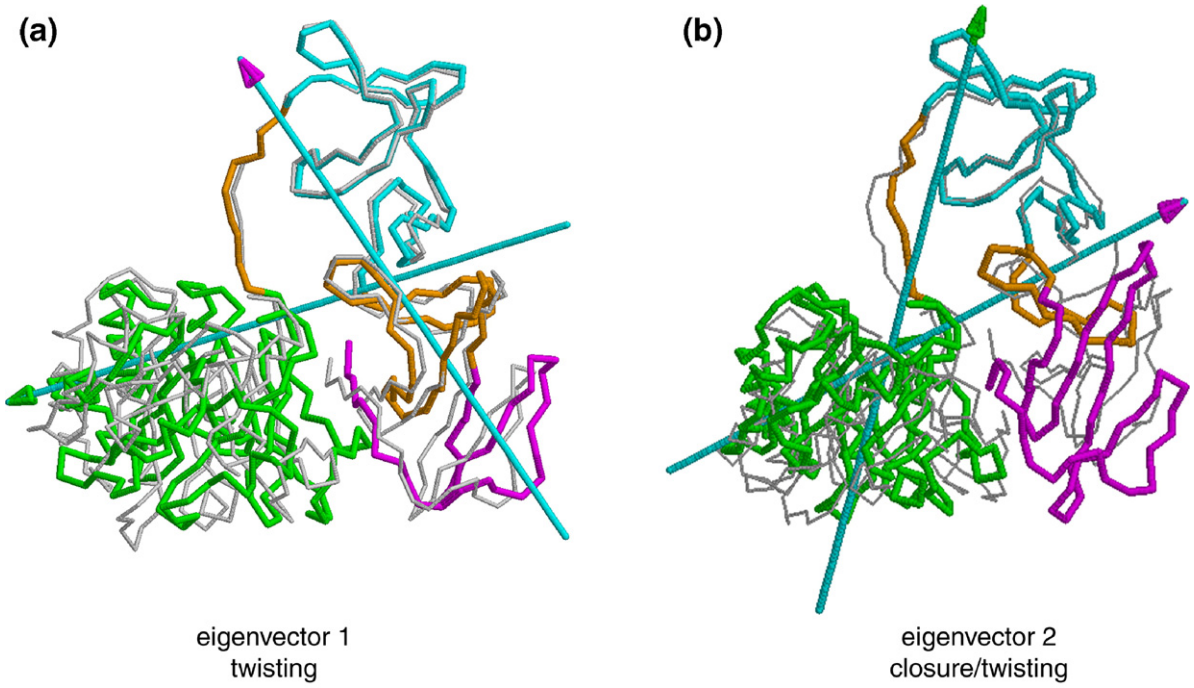

Figure 2. Backbone representation of the motions described by the two first eigenvectors, derived from an essential dynamics analysis of a CONCOORD trajectory for Cpl-1 protein (PDB code: 1h09). The superimposed two extreme structures resulting from the projection of each eigenvector on the CONCOORD trajectory are shown in coloured thick and grey thin traces. The three different moving domains between these two structures predicted by DYNDOM are coloured in green, blue and purple, and the proposed bending residues are coloured in orange. Arrows follow the right-hand-grip rule and are colour-coded with the head indicating the moving domain, and the shaft, the fixed domain.

Docking of the Cpl-1 monomer in the SAXS map

The low-resolution particle shape of Cpl-1 monomer (Figure 3(a)) was reconstructed *ab initio* using the programs DAMMIN[33] and DALAI_GA.[34] Several independent reconstructions were performed starting from random initial approximations to compute the average and the most probable models (see Materials and Methods). Results from both programs are in close agreement. The SAXS-derived low-resolution envelope for Cpl-1 in the absence of choline has a maximum distance of 10.0 nm and a radius of gyration of 3.05 nm (Figure 3(a) and Table 1).

Table 2. Computational flexibility analysis of the Cpl-1 monomer

	Eigenvector 1	Eigenvector 2	Eigenvector 3
Domain 1 (1–188)	3–189	3–191	3–151
Bending residues (domain1-domain2)	188–198 (twist)	185–196 (twist)	149–191 (closure)
Domain 2 (200–281)	190–262	192–260	191–262
Bending residues (domain2-domain3)	262–302 (twist)	260–279 (closure)	262–280 (closure)
Domain 3 (282–336)	300–336	277–336	279–336

Predicted moving domains and bending residues for the three first eigenvectors defining the essential dynamics of Cpl-1 monomer. The eigenvectors are derived from the diagonalization of the covariance matrix of the atomic fluctuations in 500 structures generated by CONCOORD around the crystal structure. Domain and bending residues are predicted by DYNDOM from the two extreme structures resulting from the projection of all the 500 CONCOORD-generated structures onto each eigenvector.

It is impossible to directly fit the crystal structure of the Cpl-1 monomer or any of the CONCOORD generated ones (not very different from the starting crystal structure) into the reconstructed low-resolution envelope generated from SAXS data. Therefore, in order to obtain a first model of the solution structure of Cpl-1 compatible with the low-resolution envelope, we used the semi-automated rigid body modelling software MASSHA.[35] In this approximation the two Cpl-1 modules were considered as completely independent rigid bodies, and the scattering profiles of the different structures resulting from slightly varying their relative disposition were iteratively fitted to the experimental scattering data. The arrangement of Cpl-1 modules in the final model (Figure 3(a), red structure) is similar to that found in the crystal structure but the angle between the catalytic module and the CBM is wider. This reorientation may account for the solution behaviour of Cpl-1, since the theoretical SAXS profile of the MASSHA model (Figure 3(b), broken red line) overlaps the experimental one (Figure 3(b), continuous black line) (Chi-square = 2.10).

Although the hydrophobic interface between the catalytic module and the CII domain of the CBM[22] should not be a hinge area according to CONCOORD simulations, it is largely modified in the model generated by MASSHA. For instance, hydrophobic contacts mediated by A339 are lost and the bifurcated salt-bridge between E163 (catalytic

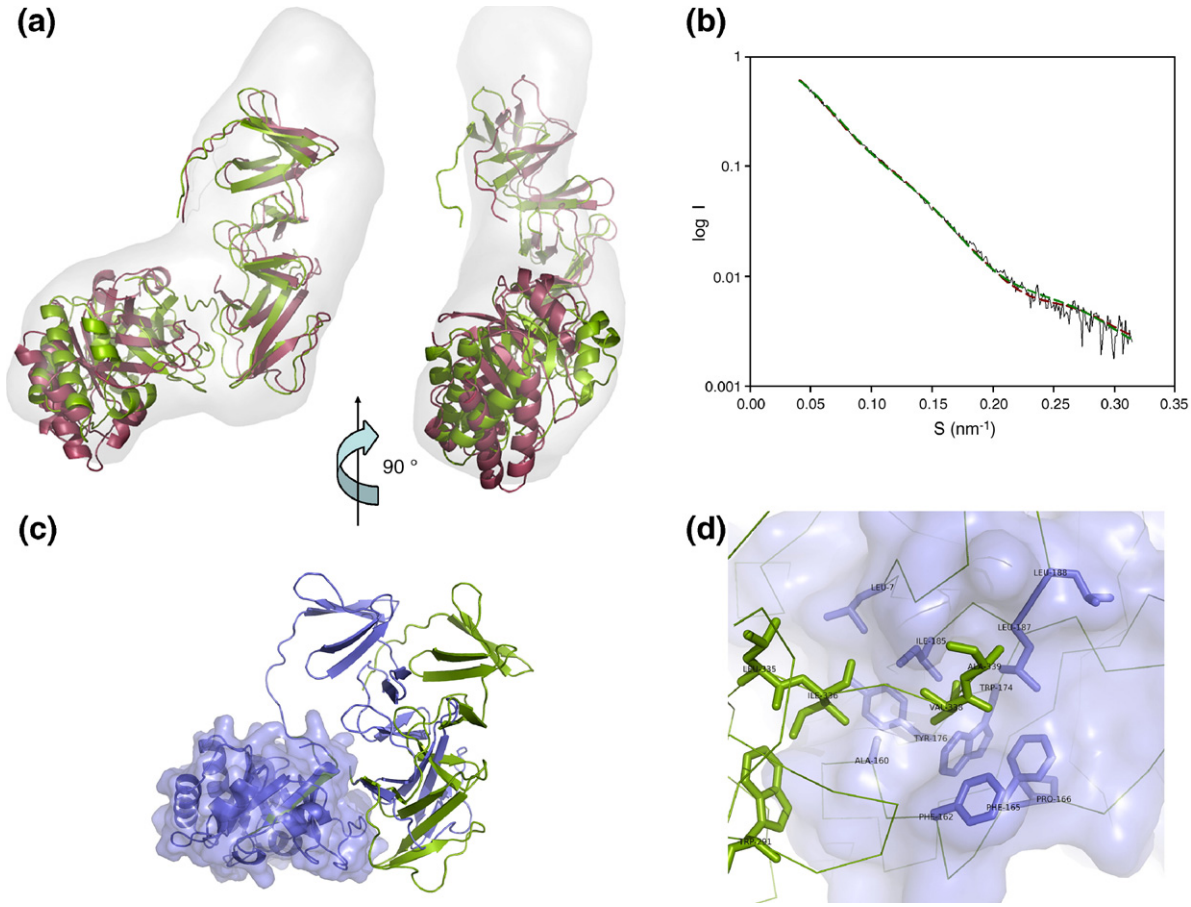

Figure 3. Low-resolution envelope for the monomer of Cpl-1 in solution without choline and SAXS scattering profiles. (a) Best fittings of the proposed solution structures of Cpl-1 into the SAXS-derived low-resolution envelope. The red structure was generated by rigid-body modelling using MASSHA,[35] while the green structure was manually generated as described in the text. (b) Experimental SAXS profile of Cpl-1 monomer (black continuous line) protein in comparison with the theoretical curves generated for the rigid body model (red broken line) and manually generated model (green broken line) using CRYSOL. (c) Backbone/ribbon superposition of the Cpl-1 crystal structure (blue) with the manually generated model (green). Molecular surface of the catalytic module is depicted in blue. (d) Close view of intermodular hydrophobic interactions in the SAXS-based model. Hydrophobic residues of the catalytic barrel (blue) and the CBM (green) joining both modules are drawn in sticks representation. Molecular surface of the catalytic module is coloured in blue.

module) and K308 and K311 (CBM), that reinforces the interactions between modules, do not exist. On the other hand, the flexibility denoted by the essential dynamics analysis of Cpl-1 structures generated by CONCOORD might allow a reorientation of Cpl-1 modules without disrupting the hydrophobic interface. We have, therefore, generated a model where the relative disposition of Cpl-1 modules was varied by manually modifying the torsion angles between the last three residues (337–339) of the CBM (Figure 3(c)) that keep the salt-bridges intact. The linker region (residues 190–199) was considered as a restraint in the maximum distance between modules but it was not modelled (a gap between residues 189–190 was introduced) due to the poor resolution of the SAXS data. The linker stretching arising from realignment of Cpl-1 modules during modelling could be easily overcome since the linker segment is both flexible (*B* factor is twice the crystal average) and long (more than 40 Å in an extended conformation) enough to allow for a certain distance increase between the C terminus of the catalytic module and the N terminus of the CBM. Moreover, its acid nature (six residues out of ten are acidic) helps the linker to adopt an extended, solvated conformation.

The position of the CBM in the model obtained with this approach (Figure 3(a), green structure) is very similar to that previously generated by semi-automated rigid body modelling (Figure 3(a), red structure). As expected, the predicted theoretical scattering curve of this model (Figure 3(b), broken green line) also overlaps to the experimental one (Chi-square = 2.65). This indicates that a minor change of the hydrophobic interface can provide a conformation compatible with the Cpl-1 solution structure, according to SAXS data (Figure 3(d)). It should be noticed that the SAXS envelope is slightly larger than both 3D-models, which might reflect a rather limited degree of sample aggregation.

The theoretical sedimentation coefficients calculated with HYDROPRO/HYDRO for the semi-automated rigid-body-modelling structure ($s_{20,w}$ 3.00 S), the manually generated model ($s_{20,w}$ 3.05 S), the

SAXS low resolution envelope ($s_{20,w}$ 3.04 S) and the crystal structure ($s_{20,w}$ 3.16 S) are very close to the experimental value ($s_{20,w}$ of 3.17(±0.02) S; Table 1). Therefore, hydrodynamic data do not allow distinguishing between the crystal structure and the SAXS-based models for the Cpl-1 monomer.

SAXS model for the dimer of Cpl-1

As described above, in solution, Cpl-1 forms a dimer upon choline binding whose three-dimensional structure remains unknown. Two different models of Cpl-1 dimers might be proposed on the basis of crystal packing for monomeric complexes formed by crystal soaking in choline solutions. In one of them, the monomer/monomer interface is formed by residues from the first β-hairpin (i.e. the beginning of the CI region of the CBM) interacting with the linker residues of a symmetry-related molecule (Figure 4(a)). The second putative crystallographic dimer is formed through the interaction of two symmetry-related CII regions, involving aromatic residues. The CBMs from both monomers would conform an elongated cylinder of approximately 2.0 nm × 10.0 nm forming angles of around 45° with the catalytic barrels, whose catalytic sites face opposite directions (Figure 4(b)). However, the calculated SAXS profiles of both dimers are very different from the experimental one (Figure 4(c)) (Chi-square model 1 = 12.09; Chi-square model 2 = 6.32). These differences are correlated with strong divergences in the particle hydrodynamic parameters such as the radius of gyration or the sedimentation coefficient, whose experimental values (R_g = 4.82 nm and $s_{20,w}$ = 4.33 S) are significantly different from the theoretical estimations yielded by HYDROPRO using the crystallographic coordinates (R_g = 2.82 nm and $s_{20,w}$ = 5.15 S for dimer in Figure 4(a); R_g = 3.43 nm and $s_{20,w}$ = 4.72 S for dimer in Figure 4(b)).

The scattering profile of the Cpl-1 dimer in 40 mM choline shown in Figure 5(a) was *ab initio* modelled (R_g = 4.82 nm; D_{max} = 15.2 nm) and the low-resolution envelope resulting is shown in Figure 5(b). This model has a radius of gyration of 4.68 nm (Table 1) and displays a "pair-of-cherries-like" shape that

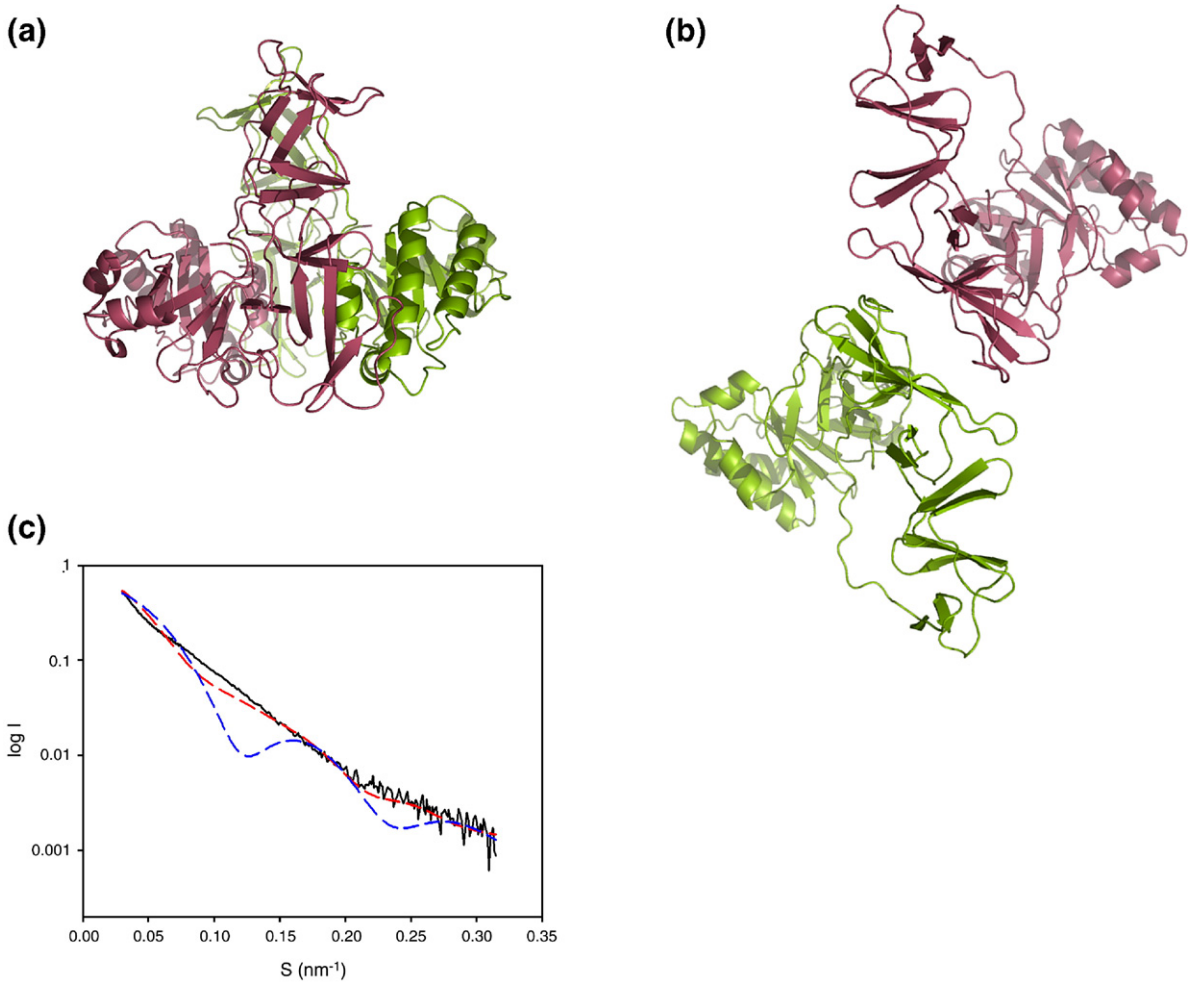

Figure 4. Hypothetical dimers derived from the crystallographic contacts of Cpl-1 monomer and SAXS scattering profiles. (a) and (b) Dimeric models from the crystallographic packing of Cpl-1 monomer. (c) Comparison of the experimental SAXS profile (black continuous line) with the theoretical profiles for the dimers. The red broken line corresponds to the model shown in (a), while the blue broken line corresponds to the model shown in (b).

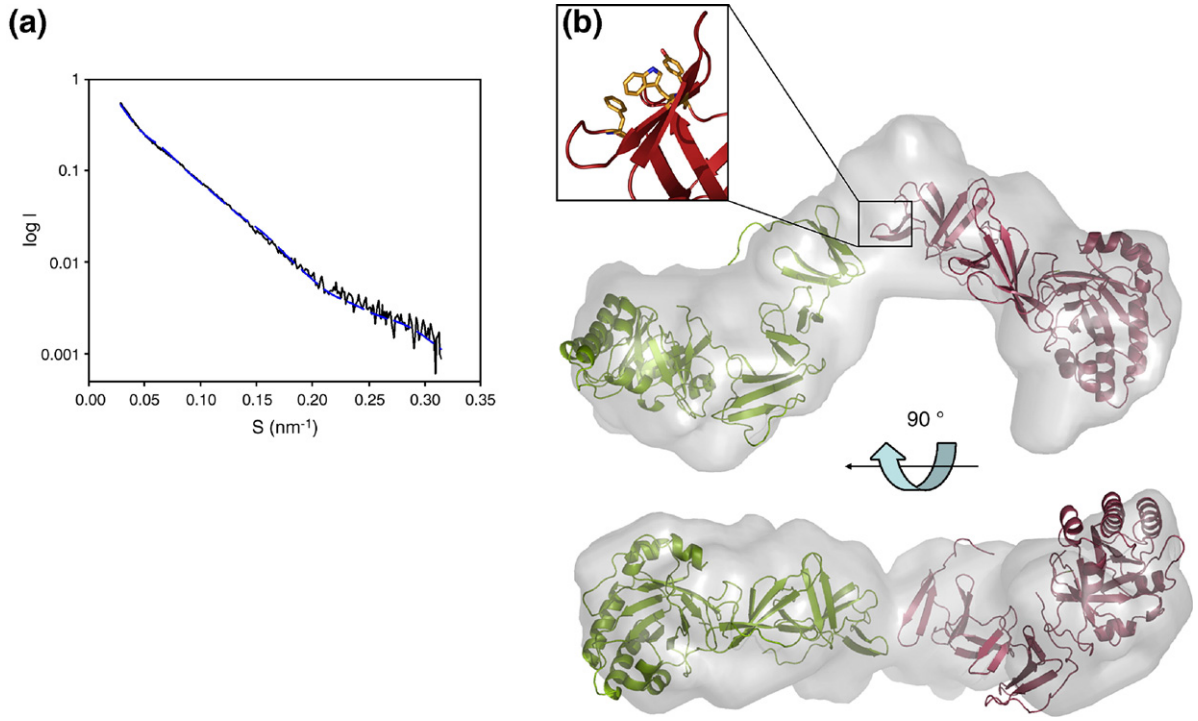

Figure 5. Low-resolution envelope for the dimer of Cpl-1 in the presence of choline and docking of the solution dimeric model. (a) Experimental SAXS profile of Cpl-1 dimer (black continuous line) in comparison with the theoretical one, generated by CRYSOL, for the proposed solution dimeric model (blue broken line). (b) Docking of the SAXS-based model of Cpl-1 dimer in the low-resolution map. Boxed picture: Detail of the dimerization region showing the side-chain of the aromatic residues forming the non-canonical choline binding site.

does not seem to be compatible with the shapes of the two putative crystallographic dimers described above. A high-resolution three-dimensional model of the solution dimer (Figure 5(b)) was built by docking the SAXS-based structure of the monomer (Figure 3(a), green structure) into the low-resolution envelope of the dimer. At the resolution of the SAXS technique, both the monomer and the choline-bound dimer of Cpl-1 would have the same tertiary structure in solution, since the SAXS-based model of the former nicely fits into each of the symmetric arms of the dimer. The resulting model shows that only one mode of association of Cpl-1 monomers would be compatible with the SAXS envelope. The dimerization interface is, therefore, unambiguously identified (Figure 5(b)), and the model's theoretical scattering profile overlaps to the experimental one (Figure 5(a), blue line) (Chi-square = 4.06). Similar results were obtained by semi-automated rigid-body modelling, considering the monomers as rigid particles (not shown). According to the model in Figure 5(b), dimerization would proceed through the N terminus of the CBM, placing the catalytic domains at both extremes of the structure. The dimerization interface would comprise the first β-hairpin of the CBM of each monomer in which three aromatic residues (W210, F218 and Y238) form a previously unnoticed, non-canonical choline-binding site, according to Cpl-1 crystal structure (Figure 5(b), inset). However, at this resolution, we can model neither the interacting residues of the interface nor the possible reorganizations in the hairpins involved. The theoretical sedimentation coefficient and the radius of gyration ($s_{20,w}$ = 4.22 S and R_g = 5.06 nm) of this model are in close agreement with the experimental values (Table 1), further supporting the proposed model.

SAXS model for the dimer of C-Cpl-1

In order to analyse the influence of the intermodular interactions established between the CBM and the catalytic module of Cpl-1 in the former global structure, we also studied the scattering profiles of the isolated CBM of Cpl-1 (C-Cpl-1 protein) in phosphate buffer with 40 mM choline. As in the full-length protein, sedimentation velocity experiments show that C-Cpl-1 behaves as a monomer in solution ($s_{20,w}$ = 1.81 S) that evolves towards a dimer ($s_{20,w}$ = 2.58 S) upon choline binding. When the Stokes' radius of the dimer (R_s = 3.17 nm) and monomer (R_s = 2.22 nm) are compared, it is seen that dimerization proceeds with a significant increase in the particle length, consistent with an end-to-end interaction between monomers.

The scattering profile of this dimer is shown in Figure 6(a) and the SAXS-derived low-resolution envelope in Figure 6(b). The SAXS data yield a radius of gyration of 3.17 nm and a maximum distance of 9.5 nm. The low-resolution envelope has an overall "boomerang-like shape" (Figure 6(b)) that closely resembles the central region in the SAXS model derived for the Cpl-1 dimer, and is consistent with the particle elongation inferred from the increase in

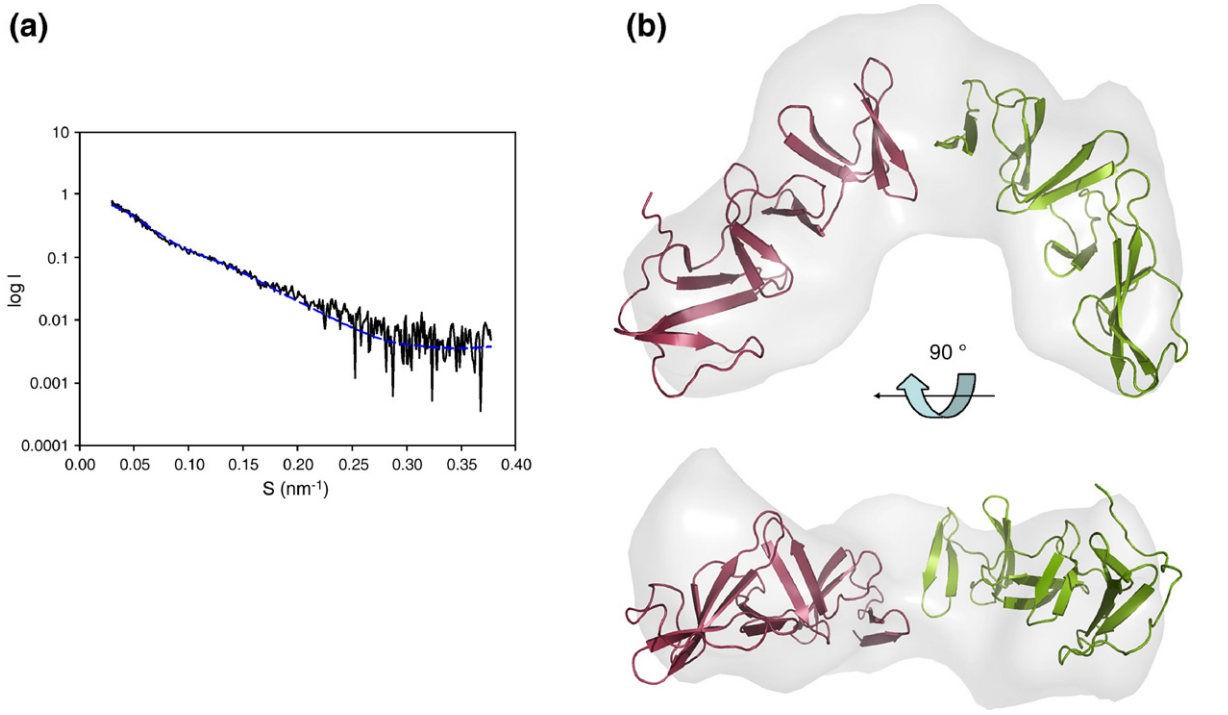

Figure 6. Low-resolution envelope for the solution dimer of C-Cpl-1 in the presence of choline. (a) Experimental SAXS profile of C-Cpl-1 dimer (black continuous line) in comparison with the theoretical profile generated by CRYSOL (blue broken line) for the SAXS-based model extracted from the full-length Cpl-1 dimer shown in Figure 5(b). (b) Docking of the central arms of Cpl-1 dimeric model shown in Figure 5(b) into the SAXS-derived low-resolution envelope of C-Cpl-1.

the Stokes' radius upon C-Cpl-1 dimerization. Indeed, the central body of the full-length dimer in Figure 5(b) fits nicely into the SAXS-calculated envelope of C-Cpl-1 dimer (Figure 6(b)), and its theoretical scattering profile overlaps the experimental one (Figure 6(a), blue line) (Chi-square = 1.88).

The theoretically estimated values of the gyration radius (3.26 nm) and the sedimentation coefficient (2.46 S) of the inferred dimer perfectly agree with the experimental ones (Table 1). Taken together, these results further support that Cpl-1 dimerization proceeds through the first β-hairpin of the CBM and that, at the resolution of our SAXS models, the structure of the CBM is maintained in the absence of the catalytic domain.

SAXS model for Pce

Here, we have also characterised the solution oligomerization state of the pneumococcal phosphorylcholine esterase, Pce, by analytical ultracentrifugation and SAXS experiments. The crystal structure of Pce has been recently published and, as Cpl-1, is a modular enzyme comprising a catalytic module and the CBM, which are joined by a small linker.[20] The N-terminal catalytic module is formed by a single globular domain that folds into an αβ/βα sandwich, while the C-terminal CBM comprises ten CBRs strictly arranged into a left-handed superhelical fold, similar to that observed in C-LytA.[21]

In contrast to Cpl-1, sedimentation velocity experiments evidence that the wild-type enzyme behaves as a monomer in both the free and choline-bound states (Table 1). The X-ray scattering profiles of Pce collected in the absence and in the presence of choline are very similar (Figure 7(a)), although the free form is slightly more compact according to their R_g and D_{max} values (free form: R_g = 4.07 nm and D_{max} = 14.3 nm; choline-bound form: R_g = 4.30 nm and D_{max} = 15.5 nm).

Ab initio modelled low-resolution envelopes of Pce in the presence or absence of choline are very similar, strongly resembling the crystal structure,[20] which is easily docked inside (Figure 7(b)) and their calculated SAXS profile fits the experimental one (Chi-square = 3.51 for Pce-choline). The crystallised form of Pce lacks the last 55 amino acid residues of the C-terminal tail that, according to secondary-structure prediction methods, would be disordered[36] making their precise contribution to the scattering profile and particle hydrodynamics unpredictable. In fact, the SAXS envelope is slightly longer than the crystal structure, which could be due to the contribution of the residues missing in the crystal structure. Nevertheless, there is a reasonable correspondence between the experimental values of R_g and $s_{20,w}$ and those theoretically estimated from Pce Δ55 crystal structure (Table 1).

Discussion

CBPs are modular proteins made of at least two well-defined modules.[37] So far only two complete

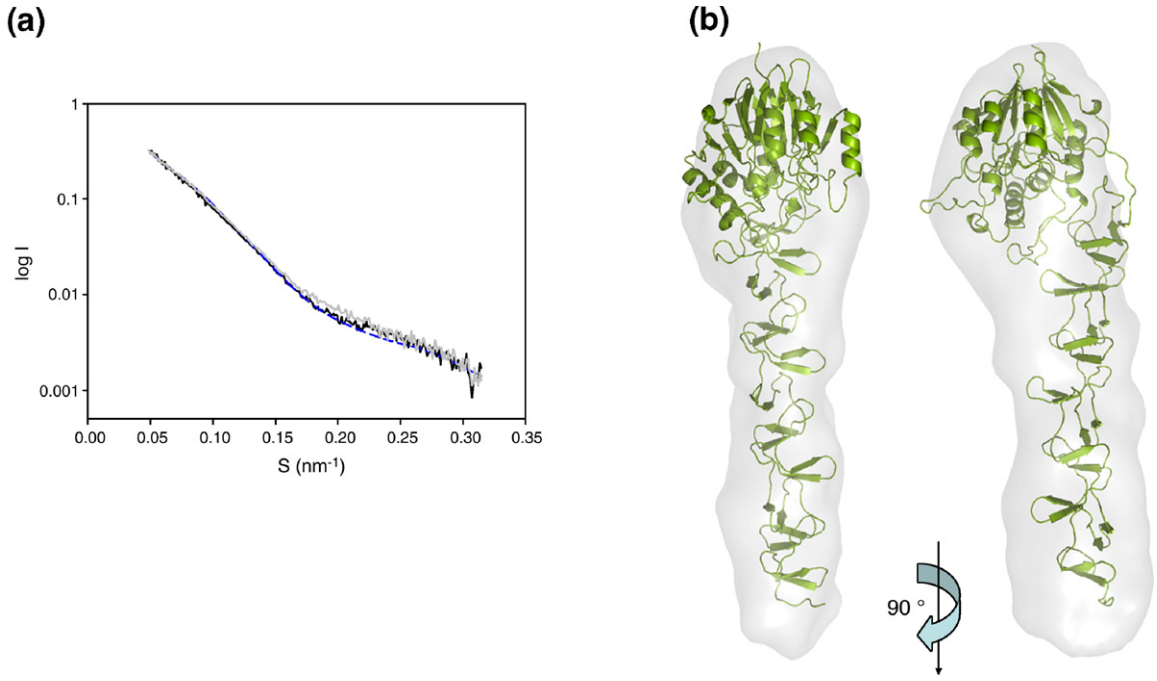

Figure 7. Low-resolution envelope for Pce in solution with choline and docking of the three-dimensional structure derived from crystals grown in the presence of choline analogues. (a) Superimposition of SAXS theoretical profile generated by CRYSOL for the high-resolution crystal structure (blue broken line) to the experimental curves registered in 40 mM choline (black continuous line) or without choline (grey continuous line). (b) Docking of the high-resolution crystal structure of Pce Δ55 into the SAXS low-resolution map obtained in the presence of choline.

CBP structures have been solved at high resolution: Cpl-1[22] and Pce.[20] One the other hand, all efforts to crystallise Pce in the absence of choline analogues or the choline-bound dimer of Cpl-1 lysozyme have been unsuccessful, preventing the study of mechanisms connected with the recognition of choline residues of the cell wall. Since the size of Pce and Cpl-1 dimer (70 and 80 kDa, respectively) are too large for NMR studies and too small for cryo-EM, we have used a small-angle X-ray solution scattering approach, combined with analytical ultracentrifugation experiments and modelling, to analyse choline-mediated changes in the association state of both enzymes.

Solution structure of Cpl-1 lysozyme

Solution scattering data and modelling of Cpl-1 solution structures have provided key insights that complement the atomic details revealed by the protein crystal structure.[22] Firstly, the calculated SAXS profile of the monomer using the crystal structure is remarkably different from the experimental one (Figure 1(b)), indicating an inherent flexibility that allows for a certain variation in the relative disposition of Cpl-1 modules within the particle, also supported by the analysis performed by CONCOORD simulations. Starting from the crystal structure and taking into account the possible hinges and rigid areas in Cpl-1 structure suggested by CONCORD analysis, we have constructed a solution three-dimensional model for the monomeric particle that is consistent with the SAXS low-resolution envelope, and accounts for the experimental scattering profile (Figure 3(b)) and the particle hydrodynamic properties deduced by analytical ultracentrifugation (Table 1). The proposed model shows an elongated structure where Cpl-1 modules form an angle slightly wider than in the crystal structure (Figure 3(a)) and keep the main contacts responsible for the intermodular interactions between the catalytic module (strands β6–β8) and the CII domain of the CBM. It is worth noting the substitution of residues K308, K311 and A339, essential for the stabilization of this interface (see above), by Q285, T290 and K318, respectively, in the pneumococcal LytA amidase. Moreover, the hydrophobic cavity comprised between strands β6 to β8 of Cpl-1 catalytic barrel does not exist in the structure of Cellosyl,[38] a non-modular lysozyme homologous to the Cpl-1 catalytic module. These sequence differences suggest a co-evolution of Cpl-1 modules towards an improvement of intermodular contacts and, therefore, of the lysozyme hydrolytic activity. Indeed, the plasticity of Cpl-1 structure, boosted by the presence of a long and highly hydrophilic linker between the catalytic barrel and the CBM, might facilitate the simultaneous interaction of the enzyme with the glycanic chain to be hydrolysed and the choline residues of (lipo)teichoic acids used as a docking station on the cell wall (essential for activity)[39] and, so, be relevant for the protein function.

Secondly, the first structural characterization of the Cpl-1 dimer stabilized by choline binding is reported. When the SAXS-based model of the

monomer is docked into the SAXS envelope derived for the Cpl-1 dimer, only one mode of association of the two monomers yields a good fitting, unambiguously defining the dimerization interface. Monomer–monomer interactions take place through the first beta hairpin at the N terminus of the CBM of Cpl-1 where a non-canonical choline binding site, previously unnoticed, formed by aromatic residues W210, F218 and Y238, might contribute to stabilize the dimerization interface, increasing at the same time the number of functional sites. In the crystal structure, this choline-binding site was occupied by an aromatic side-chain of a symmetrically related molecule that may impair choline binding. The inspection of the structural fluctuations based in CONCOORD analysis and the SAXS-based models does not suggest that Cpl-1 conformational plasticity might overcome the structural restraints that hamper the occupation of choline canonical sites located between the third and the fifth CBRs in the crystal structure and the disruption of the last one by formation of the CII β-sheet.[22] On the other hand, the low affinity ($K_d \cong 3.6$ mM)[22] of Cpl-1 for choline molecules makes difficult the direct estimation of the binding stoichiometry from solution titration experiments.

Cpl-1 dimerization interface is different from that described for the dimer of C-LytA, the isolated CBM module of LytA, in which dimerization is produced by a tail-to-tail association through the two last β-hairpins from both monomers yielding a boomerang-like homodimer.[21,40] However, it has to be noticed that crystal packing of C-LytA dimers presents contacts through the N-terminal regions, resulting in a dimer of dimers,[21] that resembles those proposed for the monomer association in the Cpl-1 dimer. Indeed, the ability of CBMs for furnishing different dimerization interfaces would also account for the choline-induced self-association of Ejl, the homologous amidase coded by the Ej-1 bacteriophage, into dimers and tetramers.[16]

The C-terminal region of Cpl-1 and C-LytA shows a different global fold in both proteins, and in Cpl-1 strongly interacts with the catalytic module. Therefore, the structural differences found between Cpl-1 and C-LytA may be due to sequence differences in the region comprising the CII domain of Cpl-1, the contacts between CII with the catalytic module or a combination of both. The role played by intermodular contacts in the final global fold of the CII domain was checked by analysing a truncated form of the enzyme (C-Cpl-1) lacking the catalytic module. C-Cpl-1 also dimerises upon choline binding, and the particle dimensions and shape of the SAXS envelope of this dimer are in close concordance with those of the central body in full-length Cpl-1 of the full dimer. In addition, the length of a hypothetical super-helix comprising all the β-hairpins of C-Cpl-1 would be higher than the SAXS envelope. This strongly indicates that formation of the CII β-sheet found in the CBM of the full-length lysozyme does not require the presence of the catalytic module.

The structural differences between the CBMs of Cpl-1 and LytA are probably related to their different functions. LytA is involved in the autolysis of pneumococci at the stationary phase of growth. It allows liberation of toxic substances and cell-wall degradation products into the host tissues, causing inflammation and facilitating invasion.[41,42] Due to its potential suicidal activity, the function of LytA must be strictly regulated to avoid cellular lysis before it is required. In contrast, the role of Cpl-1 endolysin is promoting a fast spread of newly formed virions at the end of the bacteriophage cycle. In addition, although they share the same substrate, i.e. the pneumococcal peptidoglycan, Cpl-1 and LytA also have different bond-cleavage specificities. Therefore, differences exhibited by Cpl-1 and C-LytA dimers suggest that divergent evolution of their CBMs from a common ancestor might have selected the necessary changes to make dimerization possible using different interacting surfaces and, thus, specifically enhancing the catalytic activity of each enzyme. The dramatic decrease in activity found in LytA mutants unable to self-associate upon choline binding[15,18] together with the fact that all pneumococcal lysins harbouring CBMs comprising six repeats and a C-terminal tail form homodimers under choline saturating conditions, point to the relevance of pneumococcal lysins quaternary structure in hydrolytic activity.

Solution structure of Pce protein

In contrast to Cpl-1 protein, the Pce phosphorylcholine esterase behaves in solution as a monomer independently of the presence or absence of choline, as demonstrated by sedimentation and SAXS experiments. Scattering profiles of free and choline-bound Pce are quite similar (Figure 7(a)), being identical in the low angle region, which means identical shapes and dimensions. SAXS-derived models for Pce in the presence and absence of choline were very similar, as expected from the scattering profiles. The shape of the SAXS-derived low-resolution envelopes and the atomic X-ray crystal structure are in close agreement, therefore and in contrast to Cpl-1, Pce has basically identical structures in solution and in the crystal. This finding is in accordance with the very strong interactions found between the CBM and the catalytic module in the crystal structure of Pce.[20] The structural framework of intermodular interactions is determined by three main factors: (i) a short linker on the surface of the catalytic module; (ii) the presence of a very long loop in the catalytic module (residues 36–61) that strongly interacts with the first three repeating units of the choline-binding module, and (iii) two structural Ca^{2+} ions reinforcing the 36–61 loop conformation.[20] This specific disposition of Pce modules seems to be directly related to the orientations of the teichoic acid chain acting as substrate and the active-site cavity. In connection with this, it has been observed that the purified recombinant N-terminal module of Pce retained its full activity

towards small substrates, such as *p*-nitrophenylphosphorylcholine, but becomes almost inactive towards the cell-wall polymeric substrate.[43]

Concluding remarks

In summary, SAXS experiments have revealed great differences in the structural plasticity exhibited by CBPs in solution. Thus, Cpl-1 shows an intrinsic flexibility, allowing certain conformational variability in the relative alignment of the CBM and the catalytic module that preserves the essential intermodular interactions, but could facilitate the recognition of the substrate by the cell-wall-attached enzyme and, therefore, increase the catalytic efficiency. However, in the case of Pce, a very restricted disposition of its modules is found both in solution and in the crystal structure. This fact could be related to its *in vivo* function, specifically releasing only those phosphorylcholine terminal residues relevant for escaping the immune system of the host while choline residues important for normal cell growth and infectivity are preserved.[20]

Dimerization is induced, upon choline binding, in CBPs with amidase or muramidase activity containing CBMs formed by six repeats and a short C-terminal tail. SAXS experiments, here presented, have revealed for the first time the oligomerization state for a complete CBP. Interestingly, dimerization in Cpl-1 is produced through the first choline-binding repeat in which a non-canonical choline binding site has been identified. This alignment seems to be independent of the presence or absence of the catalytic module and reveals a monomer disposition different from that previously found for the dimer of the isolated CBM of the LytA amidase. This finding together with the sequence variations displayed by the CBM of both enzymes could denote an adaptive co-evolution of their constituting modules related to their different functional roles and/or substrate specificity. The absence of choline-induced self-association of Pce suggests that protein dimerization and/or duplication of sequence repeats may be alternative but not necessarily equivalent ways of improving substrate recognition by pneumococcal murein hydrolases. More experiments should be done in order to better understand the relationship between the function displayed by a CBP, the size of its CBM and the modular arrangement within the protein structure.

Materials and Methods

Proteins

Cpl-1 (39,249 Da) lysozyme and its isolated CBM (C-Cpl-1; 18,758 Da) were expressed using, respectively, *Escherichia coli* DH1[pCIP100] and DH1[pCM1] cells. The wild-type Pce phosphorylcholine esterase (69,426 Da) and its truncated PceΔ55 mutant (63,504 Da) were produced from *E. coli* BL21 (DE3) (pRGR12)[23] and *E. coli* BL21 (DE3) (pAPM01),[36] respectively. All proteins were purified by affinity chromatography in DEAE-cellulose following the procedures described.[36,44] The purity of the isolated samples was routinely analysed by SDS-PAGE and the proteins were stored at −20 °C. Protein concentration was determined spectrophotometrically using the following molar absorption coefficients at 280 nm: 97,115 $M^{-1}cm^{-1}$ (Cpl-1), 83,640 $M^{-1}cm^{-1}$ (C-Cpl-1) and 194,020 $M^{-1}cm^{-1}$ (Pce and PceΔ55). Before used, proteins were equilibrated by extensive dialysis against the appropriate buffer. Unless otherwise stated all chemicals (Sigma Chemical Co.) were of the highest available quality.

Analytical ultracentrifugation

Sedimentation velocity experiments were performed by centrifugation of 400 μl samples at 200,000***g*** and 20 °C in an Optima-XLI (An50Ti rotor) analytical ultracentrifuge (Beckman Instruments, Inc.), equipped with an interference detector at protein concentrations used for SAXS ($6 \approx$ mg/ml; 100–350 μM). Radial scans were recorded every 60 s. Sedimentation coefficients (s) were calculated from the ratio of the velocity of the solute boundary (SVEDVERG program[45]). Data correspond in all cases to monodisperse samples and were well fitted by a single ideal species model. The s values were corrected to standard values in water at 20 °C, $s_{20,w}$.[46] Relative hydrodynamic sizes of sedimenting particles were compared using their Stokes' radius (R_s). The effective radius of the equivalent spherical particle was calculated as $R_s = M_w(1-\nu\rho)/6N\nu\eta s_{20,w}$[47] where M_w is the molecular weight, ν is the protein specific volume, ρ the solvent density, N the Avogadro's number and η the solvent viscosity. Solvent densities were determined using a Paar Precision Density Meter DMA 2D (Anton Paar) while viscosities were measured with an Ubbelhode 01 viscosimeter. Protein partial specific volumes were calculated from the amino acid composition using the Sednterp software.[48]

Sedimentation coefficients of SAXS models and crystal structures were calculated by using HYDRO[49] or HYDROPRO.[50] The former was used for bead models obtained as output files of DAMMIN[33] or DALAI_GA,[51] while the latter was used for PDB files.

SAXS measurements

Data collection was performed at the 2.1 station of Daresbury's synchrotron. The camera was set to cover ranges of the scattering vector (defined as reciprocal Bragg spacing, i.e. $2\sin\phi/\lambda$) from approximately 0.01 to 0.33 nm^{-1}. Absolute values of the scattering vector were obtained by reference to the orders of the 67 nm repeat in wet rat tail collagen. The temperature of the samples was set at 4 °C and the X-ray scattering profiles were recorded in time frames ranging from 15 to 60 s.

Data processing was performed using the software package provided by Collaborative Computational Project for Fibre Diffraction and Solution Scattering†. Data were normalized by beam intensity and detector response before being processed. Time frames showing radiation damage were removed before averaging. Radius of gyration (R_g), forward scattering intensity (Io), and the intraparticle distance distribution function $p(r)$ were

† http://www.ccp13.ac.uk/software/software.htm

calculated from the experimental scattering data using the software GNOM.[52] The value of the maximum diameter of the particle, D_{max}, was determined empirically by examining the quality of the fit to the experimental data for a range of D_{max} values. The low-resolution envelopes of the proteins were modelled *ab initio* using programs based on genetic (DALAI_GA[34,51]) and simulated annealing (DAMMIN[33]) algorithms. Results from both programs were very similar, except for C-Cpl-1, for which DAMMIN could not build consistent models. Typically, 10–15 models were generated, superimposed and aligned in pairs with SUPCOMB.[53] The most divergent ones were discarded and the most probable structure was determined from the remaining models. The model with lowest average spatial discrepancy was considered to be the most probable, while those with the highest ones were considered as outliers. The selected aligned structures were then averaged and filtered using DAMAVER and DAMFILT, respectively.[53] The average excluded volume of DAMMAVER was used as cut-off for DAMFILT. The computation of solution scattering from atomic models and the chi-square values between the calculated and experimental SAXS profiles was done using CRYSOL.[54]

The high-resolution crystallographic structures were automatically docked in the low-resolution envelopes obtained from the best averaged results using SITUS.[55] Complementarily, interactive semi-automated rigid body refinement against SAXS data was also performed using MASSHA.[35]

Computer simulations of protein flexibility

Protein flexibility was computationally studied by using the CONCOORD method.[26] This program measures interatomic distances and predicts their strength in order to generate, fulfilling a set of interatomic distance restrictions, random protein structures derived from an experimental structure. Strongly interacting atoms are slightly varied from the observed value, whereas more relaxed limits are allowed to those involved in weaker interactions.

A total of 500 structures were generated by CONCOORD (version 2.0) from the PDB structure 1h09 (Cpl-1). OPLS-UA (united atoms) and CONCOORD default parameters were used for van der Waals and bonded parameters, respectively. The minimum number of distances for each atom was set at 200. The 500 structures were treated as a molecular dynamic trajectory and analysed using the essential dynamic method over the backbone atoms.[32] This method first builds a covariance matrix of the atomic positional deviations, which is further diagonalized to render the eigenvectors. Thus, the total positional fluctuations can be assumed to be built up from the contributions of each individual eigenvector. Only eigenvectors with the largest eigenvalues can describe motions functionally relevant in proteins.[32] Essential dynamic analyses over CONCOORD trajectories were performed with GROMACS v3.2.[56] The predicted essential motions were analysed with the program DYNDOM v1.5[57,58] to visualise conformational changes involving domain-movements in terms of quasi-rigid bodies. DYNDOM is able to identify domains, hinge axis and hinge bending residues in proteins where two conformations are known. Since DYNDOM compares two different conformations of the same protein, we used those arising from the extreme projections of the whole CONCOORD trajectory onto each eigenvector. The window length was adjusted to five residues. The domain window size was set to 10% of the total number of residues in the protein, and the external to internal displacement ratio was set to 1.0.

Acknowledgements

This work was supported by grants BIO2003-01952, BFU2004-00358, BFU2004-01282/BMC, BFU2005-01645 and BFU2006-10288 from the Dirección General de Investigación Científica y Tecnológica (DGICYT) and a grant from BBVA Foundation. The authors thank Dr C. A. Botello, Dr J. L. Sáiz and Dr J. G. Grossmann for technical support in analytical ultracentrifugation and SAXS measurements, Professor J. L. García and Dr P. García for critical reading of the manuscript and help with protein production. We also thank J. I. Garzón-Cañas for help with docking. B. M. and R. M. B. were supported by fellowships from the Spanish Ministry of Education and Science. We also acknowledge the support of the European Community "Research Infrastructure Action" under the FP6 "Structuring the European Research Area" program. Beam time Proposal number 42132.

References

1. Musher, D. M. (2004). A pathogenic categorization of clinical syndromes caused by *Streptococcus pneumoniae*. In *Pneumococcus* (Tuomanen, E. I., Mitchell, T. J., Morrison, D. A. & Spratt, B. G., eds), pp. 211–220, American Society for Microbiology Press, Washington, DC.
2. Rane, L. & Subbarow, Y. (1940). Choline, pantothenic acid and nicotinic acid as essential growth factors for pneumococcus. *J. Biol. Chem.* **134**, 455–456.
3. Tomasz, A. (1967). Choline in the cell wall of a bacterium: a novel type of polymer-linked choline in pneumococcus. *Science*, **157**, 694–697.
4. López, R. & García, E. (2004). Recent trends on the molecular biology of pneumococcal capsules, lytic enzymes, and bacteriophage. *FEMS Microbiol. Rev.* **28**, 553–580.
5. Crain, M. J., Waltman, W. D., 2nd, Turner, J. S., Yother, J., Talkington, D. F., McDaniel, L. S. *et al.* (1990). Pneumococcal surface protein A (PspA) is serologically highly variable and is expressed by all clinically important capsular serotypes of Streptococcus pneumoniae. *Infect. Immun.* **58**, 3293–3299.
6. Rosenow, C., Ryan, P., Weiser, J. N., Johnson, S., Fontan, P., Ortqvist, A. & Masure, H. R. (1997). Contribution of novel choline-binding proteins to adherence, colonization and immunogenicity of *Streptococcus pneumoniae*. *Mol. Microbiol.* **25**, 819–829.
7. Hammerschmidt, S., Talay, S. R., Brandtzaeg, P. & Chhatwal, G. S. (1997). SpsA, a novel pneumococcal surface protein with specific binding to secretory immunoglobulin A and secretory component. *Mol. Microbiol.* **25**, 1113–1124.
8. Brooks-Walter, A., Briles, D. E. & Hollingshead, S. K. (1999). The pspC gene of *Streptococcus pneumoniae* encodes a polymorphic protein, PspC, with elicits cross-reactive antibodies to PspA and provides immunity to pneumococcal bacteraemia. *Infect. Immun.* **67**, 6533–6542.

9. Tu, A. H., Fulgham, R. L., McCroy, M. A., Briles, D. E. & Szalai, A. J. (1999). Pneumococcal surface protein A inhibits complement activation by *Streptococcus pneumoniae*. *Infect. Immun.* **67**, 4720–4724.
10. Gosink, K. K., Mann, E. R., Guglielmo, C., Tuomanen, E. & Masure, R. (2000). Role of novel choline binding proteins in virulence of *Streptococcus pneumoniae*. *Infect. Immun.* **68**, 5690–5695.
11. Mann, B., Orihuela, C., Antikainen, K., Gao, G., Sublett, J., Korhoen, T. K. & Tuomanen, E. (2006). Multifunctional role of choline binding protein G in pneumococcal pathogenesis. *Infect. Immun.* **74**, 821–829.
12. García, E., García, J. L., García, P., Arrarás, P., Sánchez-Puelles, J. M. & López, R. (1988). Molecular evolution of lytic enzymes of *Streptococcus pneumoniae* and its bacteriophages. *Proc. Natl Acad. Sci. USA*, **85**, 914–918.
13. García, P., Martín, A. C. & López, R. (1997). Bacteriophages of *Streptococcus pneumoniae*: a molecular approach. *Microb. Drug Resist.* **3**, 165–176.
14. Usobiaga, P., Medrano, F. J., Gasset, M., García, J. L., Sáiz, J. L., Rivas, G. *et al.* (1996). Structural organization of the major autolysin from *Streptococcus pneumoniae*. *J. Biol. Chem.* **271**, 6832–6838.
15. Varea, J., Sáiz, J. L., López-Zumel, C., Monterroso, B., Medrano, F. J., Arrondo, J. L. R. *et al.* (2000). Do sequence repeats play an equivalent role in the choline-binding module of pneumococcal LytA amidase? *J. Biol. Chem.* **275**, 20496–20501.
16. Sáiz, J. L., López-Zumel, C., Monterroso, B., Varea, J., Arrondo, J. L. R., Iloro, I. *et al.* (2002). Characterization of Ejl, the cell-wall amidase coded by the pneumococcal bacteriophage Ej-1. *Protein Sci.* **11**, 1788–1799.
17. Monterroso, B., López-Zumel, C., García, J. L., Sáiz, J. L., García, P., Campillo, N. E. & Menéndez, M. (2005). Unravelling the structure of the pneumococcal autolytic enzyme. *Biochem. J.* **391**, 41–49.
18. Sánchez-Puelles, J. M., García, J. L., López, R. & García, E. (1987). 3′-end modification of the *Streptococcus pneumoniae* lytA gene: role of the carboxy terminus of the pneumococcal autolysin in the process of enzymatic activation (conversion). *Gene*, **61**, 13–19.
19. Sanz, J. M., Díaz, E. & García, J. L. (1992). Studies on the structure and function of the N-terminal domain of the pneumococcal murein hydrolases. *Mol. Microbiol.* **6**, 921–931.
20. Hermoso, J. A., Lagartera, L., González, A., Stelter, M., García, P., Martínez-Ripoll, M. *et al.* (2005). Insights into pneumococcal pathogenesis from the crystal structure of the modular teichoic acid phosphorylcholine esterase Pce. *Nature Struct. Mol. Biol.* **12**, 533–538.
21. Fernández-Tornero, C., López, R., García, E., Giménez-Gallego, G. & Romero, A. (2001). A novel solenoid fold in the cell wall anchoring domain of the pneumococcal virulence factor LytA. *Nature Struct. Biol.* **8**, 1020–1024.
22. Hermoso, J. A., Monterroso, B., Albert, A., Galán, B., Ahrazem, O., García, P. *et al.* (2003). Structural basis for selective recognition of pneumococcal cell wall by modular endolysin from phage Cp-1. *Structure*, **11**, 1239–1249.
23. de las Rivas, B., García, J. L., López, R. & García, P. (2001). Molecular characterization of the pneumococcal teichoic acid phosphorylcholine esterase. *Microb. Drug Res.* **7**, 213–222.
24. Vollmer, W. & Tomasz, A. (2001). Identification of the teichoic acid phosphorylcholine esterase in *Streptococcus pneumoniae*. *Mol. Microbiol.* **39**, 1610–1622.
25. Koch, M. H., Vachette, P. & Svergun, D. I. (2003). Small-angle scattering: a view on the properties, structures and structural changes of biological macromolecules in solution. *Quart. Rev. Biophys.* **36**, 147–227.
26. de Groot, B. L., van Aalten, D. M., Scheek, R. M., Amadei, A., Vriend, G. & Berendsen, H. J. (1997). Prediction of protein conformational freedom from distance constraints. *Proteins: Struct. Funct. Genet.* **29**, 240–251.
27. de Groot, B. L., Vriend, G. & Berendsen, H. J. (1999). Conformational changes in the chaperonin GroEL: new insights into the allosteric mechanism. *J. Mol. Biol.* **286**, 1241–1249.
28. Jedrzejas, M. J., Mello, L. V., de Groot, B. L. & Li, S. (2002). Mechanism of hyaluronan degradation by *Streptococcus pneumoniae* hyaluronate lyase. Structures of complexes with the substrate. *J. Biol. Chem.* **277**, 28287–28297.
29. Mello, L. V., De Groot, B. L., Li, S. & Jedrzejas, M. J. (2002). Structure and flexibility of *Streptococcus agalactiae* hyaluronate lyase complex with its substrate. Insights into the mechanism of processive degradation of hyaluronan. *J. Biol. Chem.* **277**, 36678–36688.
30. Tai, K. (2004). Conformational sampling for the impatient. *Biophys. Chem.* **107**, 213–220.
31. Rigden, D. J., Lamani, E., Mello, L. V., Littlejohn, J. E. & Jedrzejas, M. J. (2003). Insights into the catalytic mechanism of cofactor-independent phosphoglycerate mutase from X-ray crystallography, simulated dynamics and molecular modeling. *J. Mol. Biol.* **328**, 909–920.
32. Amadei, A., Linssen, A. B. & Berendsen, H. J. (1993). Essential dynamics of proteins. *Proteins: Struct. Funct. Genet.* **17**, 412–425.
33. Svergun, D. I. (1999). Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. *Biophys. J.* **76**, 2879–2886.
34. Chacón, P., Díaz, J. F., Morán, F. & Andreu, J. M. (2000). Reconstruction of protein form with X-ray solution scattering and a genetic algorithm. *J. Mol. Biol.* **299**, 1289–1302.
35. Konarev, P. V., Petoukhov, M. V. & Svergun, D. I. (2001). MASSHA - a graphics system for rigid-body modelling of macromolecular complexes against solution scattering data. *J. Appl. Crystallog.* **34**, 527–532.
36. Lagartera, L., González, A., Stelter, M., García, P., Kahn, R., Menéndez, M. & Hermoso, J. A. (2005). Crystallization and preliminary diffraction studies of the pneumococcal teichoic acid phosphorylcholine esterase. *Pce. Acta Crystallog. sect. F*, **61**, 221–224.
37. Swiatlo, E., McDaniels, L. S. & Briles, D. E. (2004). Choline binding proteins. In *The Pneumococcus* (Tuomanen, E. I., Mitchell, t. J., Morrison, D. A. & Spratt, B. G., eds), pp. 49–60, American Society for Microbiology Press, Washington, DC.
38. Rau, A., Hogg, T., Marquardt, R. & Hilgenfeld, R. (2001). A new lysozyme fold. Crystal structure of the muramidase from *Streptomyces coelicolor* at 1.65 Å resolution. *J. Biol. Chem.* **276**, 31994–31999.
39. Sanz, J. M., Díaz, E. & García, J. L. (1992). Studies on the structure and function of the N-terminal domain of the pneumococcal murein hydrolases. *Mol. Microbiol.* **6**, 921–931.
40. Fernández-Tornero, C., García, E., López, R., Giménez-Gallego, G. & Romero, A. (2002). Two new crystal forms of the choline-binding domain of the major pneumococcal autolysin: insights into the dynamics of the active homodimer. *J. Mol. Biol.* **321**, 163–173.
41. Mitchell, T. J., Alexander, J. E., Morgan, P. J. & Andrew, P. W. (1997). Molecular analyses of virulence

factors of *Streptococcus pneumoniae*. *Soc. Appl. Bacteriol. Symp. Ser.* **26**, 62S–71S.
42. Tuomanen, E. (1999). Molecular and cellular biology of pneumococcal infection. *Curr. Opin. Microbiol.* **2**, 35–39.
43. de las Rivas, B. (2002). Aislamiento y caracterización de nuevas proteínas de union a colina de *Streptococcus pneumoniae*. PhD Thesis, Universidad Complutense de Madrid.
44. Sanz, J. M., López, R. & García, J. L. (1988). Structural requirements of choline derivatives for "conversion" of pneumococcal amidase. *FEBS Letters*, **232**, 308–312.
45. Philo, J. (1997). An improved function for fitting sedimentation velocity data for low-molecular weight solutes. *Biophys. J.* **72**, 435–444.
46. van Holde, K. E. (1985). Sedimentation. In *Physical Biochemistry*. pp. 110–136, Prentice May, Englewood Cliffs, NJ.
47. Waxman, E., Laws, W. R., Laue, T. M., Nemerson, Y. & Ross, J. B. (1993). Human factor VIIa and its complex with soluble tissue factor: evaluation of asymmetry and conformational dynamics by ultracentrifugation and fluorescence anisotropy decay methods. *Biochemistry*, **32**, 3005–3012.
48. Laue, T. M., Shah, B. D., Ridgeway, T. M. & Pelletier, S. L. (1992). Computer-aided interpretation of analytical sedimentation data for proteins. In *Analytical Ultracentrifugation in Biochemistry and Polymer Science* (Harding, S. E., Rowe, A. J. & Horton, J. C., eds), pp. 90–125, Royal Society of Chemistry, Cambridge.
49. García de la Torre, J., Navarro, S., López Martínez, M. C., Díaz, F. G. & López Cascales, J. J. (1994). HYDRO: a computer program for the prediction of hydrodynamic properties of macromolecules. *Biophys. J.* **67**, 530–531.
50. García de la Torre, J., Huertas, M. L. & Carrasco, B. (2000). Calculation of hydrodynamic properties of globular proteins from their atomic-level structure. *Biophys. J.* **78**, 719–730.
51. Chacón, P., Morán, F., Díaz, J. F., Pantos, E. & Andreu, J. M. (1998). Low-resolution structures of proteins in solution retrieved from X-ray scattering with a genetic algorithm. *Biophys. J.* **74**, 2760–2775.
52. Svergun, D. I. (1992). Determination of the regularization parameter in Indirect-Transform Methods usong perceptual criteria. *J. Appl. Crystallog.* **25**, 495–503.
53. Volkov, V. V. & Svergun, D. I. (2003). Uniqueness of ab initio shape determination in small-angle scattering. *J. Appl. Crystallog.* **36**, 860–864.
54. Svergun, D. I., Barberato, C. & Koch, M. H. J. (1995). CRYSOL - a program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. *J. Appl. Crystallog.* **28**, 768–773.
55. Wriggers, W. & Chacón, P. (2001). Using Situs for the registration of protein structures with low-resolution bead models from X-ray solution scattering. *J. Appl. Crystallog.* **34**, 773–776.
56. Lindahl, E., Hess, B. & van der Spoel, D. (2001). GROMACS 3.0: a package for molecular simulation and trajectory analysis. *J. Mol. Model.* **7**, 306–317.
57. Hayward, S. & Berendsen, H. J. (1998). Systematic analysis of domain motions in proteins from conformational change: new results on citrate synthase and T4 lysozyme. *Proteins: Struct. Funct. Genet.* **30**, 144–154.
58. Hayward, S. & Lee, R. A. (2002). Improvements in the analysis of domain motions in proteins from conformational change: DynDom version 1.50. *J. Mol. Graph. Model.* **21**, 181–183.

Edited by M. Moody

(Received 15 June 2006; received in revised form 29 September 2006; accepted 30 September 2006)
Available online 5 October 2006