Biophysical Journal
Article

Biophysical Society

Low-Resolution Structure of Detergent-Solubilized Membrane Proteins from Small-Angle Scattering Data

Alexandros Koutsioubas[1,*]
[1]Jülich Centre for Neutron Science (JCNS) at Heinz Maier-Leibnitz Zentrum (MLZ), Forschungszentrum Jülich, Garching, Germany

ABSTRACT Despite the ever-increasing usage of small-angle scattering as a valuable complementary method in the field of structural biology, applications concerning membrane proteins remain elusive mainly due to experimental challenges and the relative lack of theoretical tools for the treatment of scattering data. This fact adds up to general difficulties encountered also by other established methods (crystallography, NMR) for the study of membrane proteins. Following the general paradigm of ab initio methods for low-resolution restoration of soluble protein structure from small-angle scattering data, we construct a general multiphase model with a set of physical constraints, which, together with an appropriate minimization procedure, gives direct structural information concerning the different components (protein, detergent molecules) of detergent-solubilized membrane protein complexes. Assessment of the method's precision and robustness is evaluated by performing shape restorations from simulated data of a tetrameric α-helical membrane channel (Aquaporin-0) solubilized by n-Dodecyl β-D-Maltoside and from previously published small-angle neutron scattering experimental data of the filamentous hemagglutinin adhesin β-barrel protein transporter solubilized by n-Octyl β-D-glucopyranoside. It is shown that the acquisition of small-angle neutron scattering data at two different solvent contrasts, together with an estimation of detergent aggregation number around the protein, permits the reliable reconstruction of the shape of membrane proteins without the need for any prior structural information.

INTRODUCTION

Membrane proteins perform their biological role that includes signaling, molecular transport, cell-cell recognition, and energy production, by being partially embedded in the cellular membrane. Even though knowledge of their structure is of paramount importance for better understanding a multitude of biological processes, they are heavily underrepresented in the Protein Data Bank (PDB) in comparison to soluble globular proteins, with $<$700 unique membrane structures resolved (http://blanco.biomol.uci.edu/mpstruc/). The severe difficulties in their study by high resolution structural techniques such as crystallography and NMR is, on the one hand, related to the laborious biochemical expression and purification steps needed for the preparation of appropriate samples, and on the other hand, to the need of special strategies for their extraction and solubilization.

In the endeavor to keep membrane proteins folded and active outside their natural membrane environment, different approaches have been developed including the use of detergent micelles (1), amphipols (2), nanodics (3,4), and bicelles (5). In all cases these molecular assemblies attempt to shield the hydrophobic transmembrane surface of the protein, which would otherwise lead to the aggregation and denaturation of the proteins in a solution. The use of detergents (small amphiphilic molecules) above their critical micelle concentration represents the easiest and more widespread method, having the relative drawback of its tendency to destabilize certain proteins, something that can be partially remedied by the development of new milder detergents or by using mixed lipid/detergent micelles (bicelles) that provide a more native environment to the proteins. In turn, short amphiphilic polymers named "amphipols", despite their strong stabilizing properties, tend to be much larger than the protein that they encompass. This disadvantage, together with difficulties in sample preparation and protein reconstitution complexity, is shared by the nanodisc system. Nanodiscs are noncovalent assemblies of a phospholipid bilayer that is surrounded by a genetically engineered lipoprotein dimer and probably represent the best mimic of the membrane environment that can, at the same time, solubilize membrane proteins.

Submitted July 25, 2017, and accepted for publication October 3, 2017.
*Correspondence: a.koutsioumpas@fz-juelich.de
Editor: Jill Trewhella.
https://doi.org/10.1016/j.bpj.2017.10.003

Success in keeping membrane proteins stable and active in solution gives the ability to implement various experimental techniques for the study of membrane protein structure as NMR (6), circular dichroism (7), and small-angle scattering (SAS) (8), whereas for flash-frozen proteins, cryo-electron microscopy represents a quite promising technique with many recent breakthroughs (9). However, in all cases, special attention should be paid for the usually nontrivial subtraction of the contribution of detergents, amphipols, etc., which coexist with the protein in solution. This is especially true in the case of small-angle neutron (SANS) and x-ray (SAXS) scattering, which offer the ability to acquire direct information concerning the low-resolution ($>$1 nm) structure of biological molecules (10,11) for a wide range of solution conditions and molecular sizes.

In a small-angle scattering experiment of detergent-solubilized membrane proteins, the measurable signal may contain contributions from both free detergent micelles and the protein detergent complex (PDC) (12). Even when the scattering intensity belonging only to the PDC can be extracted, calculations regarding scattering invariants or application of structure recovery tools (13,14) that have been developed for soluble proteins do not apply in a direct manner, thus hindering the widespread use of the technique for this class of systems. From the experimental point of view, isolation of the scattering curve corresponding only to the PDC involves the careful subtraction of buffer solution having the same free-micelle concentration as the PDC solution. This can be accomplished either by collecting reference buffer that is eluted before the PDC during a size exclusion chromatography (SEC) step right before the scattering experiment (15), or by performing extensive dialysis of the PDC solution against detergent containing buffer (16). Another proposed method based on singular value decomposition involves the acquisition of scattering data for many different protein-detergent stoichiometries (17). Alternatively, the use of on-line SEC-SAXS has proven as a very efficient way for the proper subtraction of micelle contributions (18,19).

The determination of the scattering curve corresponding to the PDC provides information about the overall size and shape of the system. In previous SAXS and SANS investigations where a high-resolution structure of the protein molecule or a candidate model was available, modeling and molecular simulations (18,20–22) provided a better understanding of detergent organization around the protein molecule. The ability to model adequately the detergent belt around a known protein structure (23) opens the road for experimental investigations, where conformational transitions or interaction of membrane proteins with other molecules need to be resolved (24,25).

A more ambitious approach to the whole problem aims to exploit the big advantage of the SANS technique, i.e., contrast variation, to remove contributions to the scattering from detergent molecules (26–30). In principle, using appropriate D_2O/H_2O ratios as solvent and the full or partial deuteration of detergent molecules, one may obtain the same scattering length density (sld) for the solvent and the detergents, thus rendering the detergent belt around the protein and free micelles invisible to neutrons (31). However, due to the chemically heterogeneous nature of the hydrophilic and hydrophobic part of detergents, the elimination of their scattering contribution over the entire q-range is difficult to achieve, and has been demonstrated only for special cases of detergent molecules (32). Such combinations of detergent and solvent with homogeneous matching, provided that they are able to solubilize the protein under study, permit easy sample preparation and the acquisition of scattering curves that correspond only to the protein. This same approach has been shown to be feasible also for the nanodisc system (33).

Because the requirement for simultaneous membrane protein solubilization and detergent contrast matching in a SANS experiment is quite strict and is expected to apply in limited cases, it would be interesting to explore different approaches for the determination of membrane protein structure from small-angle scattering data. Recently, Kynde et al. (34) developed a sophisticated hybrid ab initio approach with the aim to recover the low-resolution structure of a protein embedded in a nanodisc system from a combination of SANS and SAXS data. It was reported that for the case of bacteriorhodopsin embedded in a nanodisc, the amount of experimental information was not enough for this approach to be successful. The same method was shown to give better results for human cytochrome P450 anchored in a nanodisc environment (35).

The goal of this work is to explore the potential of an ab initio strategy for the general case of detergent-solubilized membrane proteins. In this context, we develop a multiphase coarse-grained bead model with a set of appropriate constraints and a minimization algorithm where, with no prior knowledge of the shape of the protein, we attempt to recover the low-resolution structure of the entire PDC. The required input to the algorithm consists of a set of scattering curves, the molecular characteristics of the detergent, and its aggregation number around the protein, which can be independently estimated by other experimental techniques (e.g., analytical ultracentrifugation or refractometry). We explicitly show that a single scattering curve contains insufficient amount of information for obtaining a unique and physically meaningful solution. On the contrary, for a set of at least two SANS curves at different solvent contrasts, the concurrent fit of the data leads to stable solutions where the obtained molecular envelopes are representative of the structure of each PDC component. The efficiency of the presented methodology is evaluated using two different sets of data: 1) scattering curves of a tetrameric α-helical membrane channel (Aquaporin-0) solubilized by n-Dodecyl β-D-Maltoside (dDM), calculated from all-atom molecular dynamics trajectories of the system; and 2) SANS

experimental curves of the filamentous hemagglutinin adhesin (Fhac) β-barrel protein transporter solubilized by n-Octyl β-D-glucopyranoside (OG_D), taken from the work of Gabel et al. (22).

METHODS

Model and minimization procedure

Following the methodology that was initially proposed by Chacón et al. (13), further improved and popularized by Svergun and co-worker (14,36) for the low-resolution restoration of biomolecule shape from small-angle scattering data, and also based on previous work (20,23,37,38), we have developed an algorithm where a set of different types of beads participate in a simulated annealing procedure that aims to arrange the beads in a way that is compatible with the experimentally measured scattering curves and also with the general anticipated organization of detergent molecules around a membrane protein.

Small-angle scattering curves of membrane protein/detergent complexes contain three different contrast contributions from the protein, detergent heads, and tails. In this respect, the constructed model consists of three types of beads (dummy atoms) corresponding to the membrane protein, detergent hydrocarbon tails, and detergent hydrophilic heads, with different contrasts $\delta\rho_M$, $\delta\rho_T$, and $\delta\rho_H$ relative to the solvent. These beads of packing radius r are placed on a close-packed face-centered cubic (fcc) lattice, having a maximum of 12 nearest neighbors. The fcc lattice constant L is equal to $4r/\sqrt{2}$. The search space available for the placement of beads is not bounded.

The scattering curve $I(q)$ versus q (scattered intensity versus wavevector transfer) of each bead configuration at each solvent contrast is calculated using the Debye equation (39):

$$I(q) = \sum_{i=1}^{N}\sum_{j=1}^{N} f_i(q) f_j(q) \frac{\sin qr_{ij}}{qr_{ij}}, \quad (1)$$

where N is the total number of beads, $f_i(q)$ is the form factor of each bead given by $V_i\delta\rho_i$ with $V_i = L^3/4$ being the effective bead volume, and r_{ij} is the distance between the ith and jth bead.

The target of the simulated annealing procedure is to obtain a compact configuration that reproduces the (M) experimentally measured scattering curves for a given number of detergent molecules (N_{det}) around the protein. The algorithm proceeds by selecting random positions of the fcc lattice and by performing trial additions, removals, and changes of the type of beads in an effort to obtain an overall bead configuration (X) that minimizes the score function:

$$f(X) = \sum_{l=1}^{M} R_l\left(I_{\exp,l}, X\right)^2 + \sum_{k=1}^{4} \alpha_k P_k(X). \quad (2)$$

The main contribution in $f(X)$ comes from the sum for each different contrast (l) of the q^2-weighted discrepancies $R(I_l,X)^2$ between the scaled intensity $I_l(X)$ (scaling factor is calculated according to Eq. 16 in (40)) corresponding to the configuration X and the experimental scattering intensity $I_{exp,l}$, given by

$$R_l(I_l, X)^2 = \frac{\sum_q \left[I_l(X,q)q^2 - I_{\exp,l}(q)q^2\right]^2}{\sum_q \left[I_{\exp,l}(q)q^2\right]^2}. \quad (3)$$

Because small-angle scattering curves are decaying rapidly as a function of q, the choice of q^2-weighting ensures that experimental points at intermediate-q (where most of the shape information is contained) will contribute substantially to the score function (41). Note that especially in the case of SANS curves where resolution effects have to be considered, $I_{exp,l}$ represents the desmeared curve calculated according to the iterative method proposed by Lake (42), taking into account the beam divergence $\Delta\theta$ and wavelength spread $\Delta\lambda/\lambda$.

Additional terms in the score function are related to various penalty weights α_k and penalties P_k having to do with protein's compactness P_1, detergent belt compactness P_2, detergent aggregation number P_3, and overall radius of gyration ($R_{g,i}$) penalty P_4. Penalties P_1 and P_2 are calculated as described by Svergun (14) for the ensemble of protein and tail beads, respectively. Penalty $P_3(X)$ ensures that the number of head and tail beads corresponds to the given detergent aggregation number N_{det}, and is defined as

$$P_3(X) = \frac{\left(N_{det} - N_t L^3/4V_t\right)^2}{N_{det}^2} + \frac{\left(N_{det} - N_h L^3/4V_h\right)^2}{N_{det}^2}, \quad (4)$$

where N_h and N_t represent the number of head and tail beads in the configuration X, and V_h,V_t is the known molecular volume of detergent heads and tails, respectively.

Penalty $P_4(X)$ is given by

$$P_4(X) = \sum_{l=1}^{M} \frac{\left(R_{g,l} - R_{gIFT,l}\right)^2}{R_{gIFT,l}^2}, \quad (5)$$

where $R_{g,l}$ is the radius of gyration of the configuration X for the contrast l, and $R_{gIFT,l}$ is the corresponding radius of gyration that is estimated from the indirect Fourier transform of the experimental curve (see also next subsection). Because the search space is not bounded, this penalty prohibits large deviations of the size of the complex from the expected values, especially in the beginning of the minimization procedure.

To guide the algorithm toward physically meaningful arrangements, it is essential to impose an additional set of constraints related to the anticipated general properties of detergent organization around the hydrophobic transmembrane surface of the proteins. These are as follows: 1) prohibition of any hydrophobic bead contacts with empty lattice positions that represent the solvent; 2) hydrophobic detergent tail beads must be separated from the solvent by a single layer of hydrophilic head beads; 3) thickness of the hydrophobic tail bead core along the z axis is bounded and cannot exceed the limit $2L_{hb\text{-}max}$; 4) tail beads should always stay connected and also encircle the z axis; and 5) protein beads cannot be disconnected.

The first two constraints reflect the requirement that hydrophobic tails should avoid contact with water. The third and fourth constraint is implemented to promote doughnut-like detergent belt structures around the protein transmembrane hydrophobic surface. Note that the z axis is assumed to coincide with the approximate axis of the protein's transmembrane part. Bounding the thickness of the detergent structure (along the z axis) is related to the properties of both detergent and proteins. For compact lipid packings, we expect that there is a maximum detergent hydrocarbon tail extension ($L_{hb\text{-}max}$) given by the Tanford formula (43) $L_{hb\text{-}max} = 1.5 + 1.265\ n_c$ Å, where n_c is the number of carbons of the detergent tail. However, for more disordered detergent arrangements, a more relaxed limit can be used by arguing that the maximum thickness of the tail structure ($2L_{hb\text{-}max}$) should be roughly equal to the typical hydrocarbon core thickness of cell membranes ($\approx$40 Å) where the transmembrane part of proteins is embedded and performs its function. Finally, the last constraint is simply implemented for keeping the protein beads interconnected, as it is expected for a protein molecule.

From the definition of the model, it can be seen that the thickness of the hydrophilic detergent head layer depends on the selected size of the beads, and is equal to two times the packing radius, $2r$. However, the real thickness of this layer t, as can be determined by fitting small-angle scattering data of detergent micelles in solution (44), might be different than $2r$. Because the

contribution of this thin layer to the scattering depends primarily on the product of layer thickness and layer contrast (45), the effective contrast of the head beads $\rho_H{}'$ is adjusted so that $2r\delta\rho_H{}' = t\delta\rho_H$. Additionally, a modified molecular volume $V_h{}'$ is used in the second term of Eq. 4, following the relation $V_h{}' = (2r/t)V_h$.

A spherical configuration of tail beads, covered with a single layer of head beads, together with an embedded string of protein beads, along the z axis, is constructed. This configuration respects all the constraints described above and serves as the initial state for the minimization of the score function $f(X)$. The simulated annealing minimization involves the definition of an effective annealing temperature T and proceeds as proposed in (14,36). We begin at a high temperature where the majority of trials is accepted. Then random trials producing modified bead configurations X' (that always respect the defined constraints) are attempted and accepted if the score function is reduced or if $f(X') > f(X)$ with a probability $e^{[f(X) - f(X'))/T}$. For the results presented in this work, an annealing schedule of $T' = 0.90T$ was used, and $100N$ trials at each temperature are attempted. Note that because only one bead is modified in each trial, the recalculation of only N cross terms of Eq. 1 are needed, thus avoiding the computationally intensive recalculation of the whole double sum (37). The annealing procedure is stopped when no further reduction of the score function is observed. The following values of the penalty weights were used: $a_1 = 0.1$, $a_2 = 0.1$, $a_3 = 1$, and $a_4 = 1$, which are expected to provide in the general case a balance between compactness and flexibility in the obtained PDC shapes. The used bead packing radius during the runs was in the range 2.5–3.0 Å. In all reconstructions, scattering data up to $q = 0.2$ Å^{-1} are considered, because in this q-range the scattering contribution from the shape of the PDC dominates (14). For smaller protein molecules where a higher q-range needs to be considered and internal inhomogeneities from each phase might contribute to the scattered intensity, an appropriate constant has to be subtracted from the scattering data, so that shape scattering curves are used with a q^{-4} behavior at high q.

Scattering data and initial treatment

Two different sets of scattering data were considered for validating the efficiency and robustness of the presented algorithm. The first set concerns theoretically calculated SAXS and SANS curves from all-atom molecular dynamics (MD) trajectories of n-Dodecyl β-D-Maltoside/Aquaporin-0 (dDM/AQ0) complexes, whereas the second set consists of experimental SANS curves of OG_D-solubilized Fhac protein transporter in 90% D_2O and 42% D_2O, taken from the elegant work of Gabel et al. (22).

The AQ0 dataset was calculated as follows. In previous computational works (19,20), it was shown that all-atom MD simulations of dDM/AQ0 complexes can reliably reproduce the SEC-SAXS experimental data (18) of this system. Following the exact same methodology as in Koutsioubas et al. (20) and by placing 270 dDM molecules around the backbone-restrained structure of AQ0 (PDB: 2B6P (46)), we have sampled, after the proper equilibration of the system, the MD trajectory over a time frame of 5 ns. Using the programs CRYSOL v.2.8.2 (40) and CRYSON v.2.7 (45) (applying default parameters including a 3 Å, 20% contrast hydration layer), we calculated the corresponding SAXS and SANS curves at various solvent contrasts (100% D_2O, 42% D_2O, and 0% D_2O). Averaging of scattering curves corresponding to the various time frames (over a simulation time of 5 ns) yielded the final scattering curves that were further smeared (in the neutron case), by assuming a beam divergence of $\Delta\theta = 0.002$ rad and $\Delta\lambda/\lambda = 0.1$, which are typical for a neutron instrument.

All simulated and experimental scattering curves were treated using the program GNOM v.4.6 (47), which performs an indirect transform of the data and among other parameters, provides as output a regularized scattering curve I_{exp} and an estimation of the protein/detergent complex radius of gyration R_{gIFT} that are, in turn, used as input for our simulated annealing algorithm.

Postprocessing of molecular envelopes

For each shape reconstruction, 10 independent runs were executed. Using the programs DAMAVER (48), we aligned the obtained protein molecular envelopes, compared all models with each other, and found the most probable one. Additionally, the mean normalized spatial discrepancy (NSD) is calculated, which represents a measure of the similarity between the obtained envelopes. For ideally superimposed objects, NSD tends to 0, although it exceeds 1 if the objects are systematically different. Additionally, the SD of the NSD is indicative of the overall stability of the solution. The most probable envelope is in turn aligned with the crystallographic structure of the protein using the program SUPCOMB (49). The structural similarity between the ab initio shape and the crystallographic structure is quantified as follows: using the program PDB2SAX (SITUS package) (50), the all-atom structure of the protein is mapped to a close-packed sphere model on a hexagonal lattice, with the same lattice parameter as the ab initio bead model. Subsequently, the NSD_{cryst} between the ab initio shape and the mapped crystallographic structure is calculated using SUPCOMB. Visualization of the PDC envelopes was performed with the program PyMOL (51).

RESULTS

First example: dDM-solubilized Aquaporin-0 channel

The generation of simulated scattering data concerning a well-characterized system such as AQ0, gave us the ability to explore different combinations of data input to our algorithm and also understand the required amount of experimental information that is needed for obtaining stable and realistic shape reconstructions. In this respect, ab initio runs from a single SAXS curve and from two or three SANS curves at different solvent contrasts were executed for the 270 dDM/AQ0 system. Additionally, for each of these input data combinations, two sets of runs are performed—one with no assumed symmetry and another one with imposed fourfold axial symmetry around the z axis. The used sld or electron densities ρ for each component of the system and detergent molecular volumes V are given in Table 1. Furthermore, the thickness of the detergent hydrophilic heads t is taken to be equal to 6 Å, as found by

TABLE 1 Slds at Different Contrasts, Electron Densities, and Molecular Volumes of Protein, Detergent Parts, and Solvent Used during the Ab Initio Runs for the AQ0/dDM System

Bead Phase	$\rho(e^-/\text{Å}^3)$	$sld_{100\%D_2O}$ (10^{-6} Å^{-2})	$sld_{42\%D_2O}$ (10^{-6} Å^{-2})	$sld_{0\%D_2O}$ (10^{-6} Å^{-2})	V (Å^3)
dDM detergent heads	0.52	3.93	2.72	1.85	350
dDM detergent tails	0.28	−0.41	−0.41	−0.41	340
AQ0 protein	0.44	2.90	2.26	1.80	—
Solvent	0.33	6.40	2.36	−0.56	—

Molecular volume (V) and sld values for the detergents were taken from the work of Breyton et al. (31). Electron densities (ρ) were taken from the work of Oliver et al. (44). Note that owing to the H/D exchange occurring in solvents containing a certain ratio of D_2O (55), the sld of a protein is adjusted accordingly for each solvent contrast. The related calculations are performed using the contrast calculator of the program SASSIE (56).

core-shell model fits of experimental free-micelle scattering curves (44).

In all ab initio runs, the algorithm was able to fit the MD generated input curves. In Fig. 1, PDC shapes for the SAXS and two curve SANS (100% D_2O and 0% D_2O) reconstructions are depicted, whereas in Table 2 we give a summary of the mean NSD values and their corresponding SD for each set of shape reconstructions. Also, the calculated NSD_{cryst} between the most probable shape and the AQ0 tetramer crystallographic structure is presented. It is found that for a single SAXS curve as input without symmetry, the obtained NSD values are quite high, indicating shape variability in the final solution. By comparing the most probable shape with the crystallographic structure, we observe systematic deviations. Only when we impose a fourfold symmetry, NSD_{cryst} tends close to 1, but still the obtained shape possesses clear differences from the AQ0 tetramer structure. We also note that the detergent belt around the protein is not quite compact as it is expected from the MD trajectories. The same conclusions are reached in the case where a single SANS curve is used as input.

In contrast, when two or three SANS curves at different solvent contrast are used as input, then the ab initio run is much more constrained and the algorithm is guided toward stable solutions (NSD < 1) whereas, as can be seen in Fig. 2, the resemblance of ab initio shapes with the crystallographic structure is evident ($NSD_{cryst} \approx 0.7$) and all major structural features of the tetramer are well reproduced. The application of symmetry restrictions improves solution stability, and leads to marginally better agreement with the crystallographic structure. We also note that the inclusion of a third solvent contrast (42% D_2O) or of the SAXS curve as input does not lead to lowered NSD and NSD_{cryst} values compared to the two-curve SANS runs, while identical final shapes are obtained. Additionally, from a visual inspection of Fig. 1 we may notice that, especially for the two SANS curves run with fourfold symmetry, the detergent belt is highly compact and its size and shape identical to the one of the MD trajectories.

All above results were obtained by assuming a fixed detergent aggregation number $N_{det} = 270$. But because in real cases this quantity has to be estimated experimentally and might be prone to error, we have tested the stability of the ab initio reconstructions against small variations of the parameter N_{det}. We have found that a 10% variation of N_{det} leads to small degradation of the obtained NSD_{cryst} values, whereas visual inspection of the protein envelopes shows only minor differences compared to the runs with $N_{det} = 270$. In more detail: for $N_{det} = 240$, $NSD_{cryst} = 0.83$; and for $N_{det} = 300$, $NSD_{cryst} = 0.72$.

Second example: OG_D-solubilized Fhac protein transporter

The second application of the algorithm concerned an actual set of SANS experimental data (Fig. 3) from the

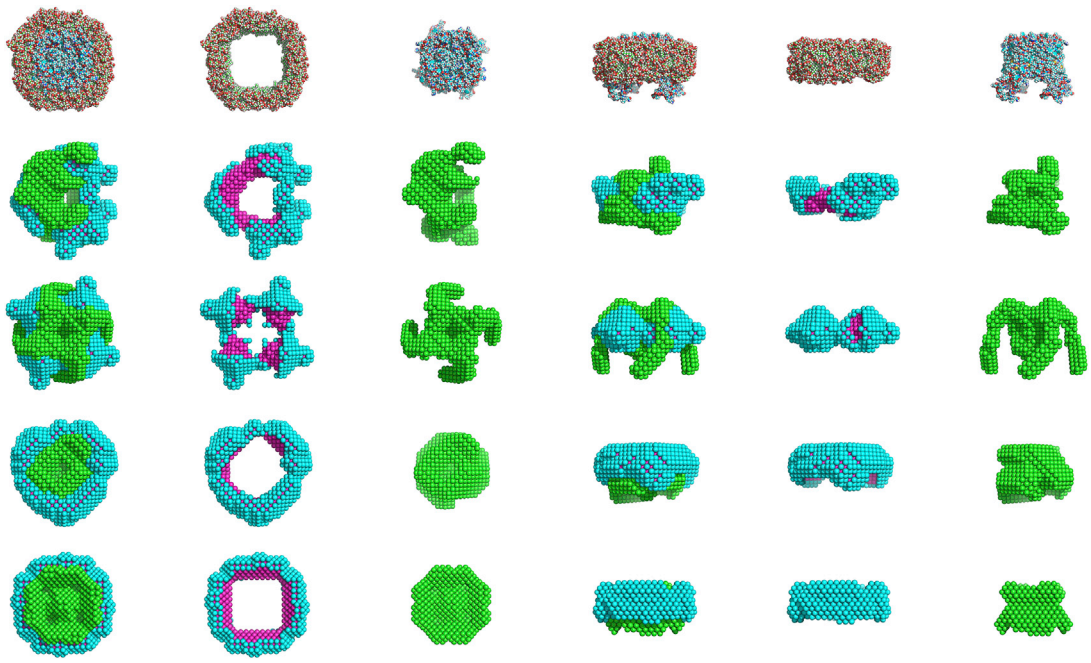

FIGURE 1 Ab initio shape reconstructions of the 270 dDM/AQ0 tetramer. (*First row*) Given here is the MD structure. (*Second row*) Given here is a single SAXS curve with no imposed symmetry. (*Third row*) Given here is a single SAXS curve with imposed fourfold axial symmetry. (*Fourth row*) Given here are two SANS curves with no imposed symmetry. (*Fifth row*) Given here are two SANS curves with imposed fourfold axial symmetry. The first three columns show the top view of the different PDC components (entire PDC, detergent belt, and protein, respectively), whereas the last three columns show the side view. (The bead colors correspond to *green*, protein; *cyan*, detergent hydrophilic heads; and *magenta*, detergent hydrophobic tails.) To see this figure in color, go online.

TABLE 2 Summary of Mean NSD, its SD, and the NSD_{cryst} Values for Ab Initio Runs of the 270 dDM/AQ0 System with Different Numbers of Curve Inputs and Imposed Symmetries

Ab Initio Run	NSD	σNSD	NSD_{cryst}
Single SAXS, no symmetry	1.082	0.056	1.64
Single SAXS, fourfold symmetry	1.035	0.161	1.10
Two SANS, no symmetry	0.668	0.077	0.73
Two SANS, fourfold symmetry	0.457	0.048	0.70
Three SANS, no symmetry	0.715	0.041	0.75
Three SANS, fourfold symmetry	0.454	0.018	0.73

OG_D-solubilized outer-membrane β-barrel transporter of the *Bordetella pertussis* Fhac (22), which consists of a 16-stranded β-barrel that is occluded by an N-terminal α-helix and an extracellular loop and a periplasmic module composed of two aligned polypeptide-transport-associated (POTRA) domains (52). The used detergent had its hydrophobic tail part fully deuterated, thus presenting high contrast in relation to the protein. The two SANS measurements were performed at 90% D_2O and 42% D_2O solvent contrast, which represents the matching point for the overall detergent molecule and protein, respectively. As reported, by the use of a dialysis step against solvents with the same detergent concentrations, the free-micelle contribution was subtracted and the SANS curves belonging only to the PDC were obtained. Additionally, for this particular system, the detergent aggregation number around the protein N_{det} was determined by analytical ultracentrifugation and was found to be equal to 142 $\pm$ 10 OG_D molecules.

For the ab initio runs using both SANS contrasts, the applied sld for each component of the system and OG_D molecular volumes V are given in Table 3. N_{det} is fixed at the value 140 and the thickness of the detergent hydrophilic heads t is taken to be equal to 3.2 Å (44). The resulting fit of the experimental results is presented in Fig. 3 and the corresponding PDC shapes are presented in Fig. 4. The obtained NSD values of the ab initio runs are equal to 0.657 $\pm$ 0.046, suggesting a stable set of solutions, whereas from the comparison with the crystallographic structure we get $NSD_{cryst} = 0.75$. Inspection of the crystallographic structure aligned with the most probable shape (Fig. 5) reveals that the ab initio envelope is representative of the overall molecular shape and that the β-barrel, extramembrane periplasmic POTRA domains, and their relative orientation, are nicely reproduced. It is interesting to note that although the detergent belt encompasses the transmembrane part of the protein, its shape is somewhat disordered and not highly compact, which exposes a small part of the protein β-barrel surface to the solvent. This result will be further discussed below in relation to previous MD modeling investigations (22).

As in the case of the AQ0 tetramer, we have checked the effect of detergent aggregation variation on the final solution, where we again find that for N_{det} in the range 125–155, the determined envelopes are very similar and NSD_{cryst} values remain almost constant (for $N_{det} = 125$, we find $N_{cryst} = 0.77$ and for $N_{det} = 155$, $N_{cryst} = 0.75$).

DISCUSSION

In general, the different chemical nature of detergent heads and tails makes it difficult to contrast-match their contribution over the entire q-range in a small-angle scattering experiment. That means that to directly access protein structural information from SAXS or SANS measurements where detergents bound to the protein contribute to the scattering, we need to rely on modeling. We note that since their introduction almost two decades ago, ab initio structure recovery algorithms for soluble globular proteins have become increasingly popular in the literature, assuming a complementary role in integrative structural studies of proteins and nucleic acids. What makes ab initio approaches quite attractive, despite their low resolution, is that no prior structural information is required concerning the biomolecule under study.

The main goal of this work is to develop an ab initio shape recovery method for detergent-solubilized membrane proteins without the need for special contrast-matched detergents, so that sample preparation can be more flexible and focused on the proper solubilization of the protein in a folded and active state. Following the paradigm of ab initio methods, we constructed a four-phase bead model (protein,

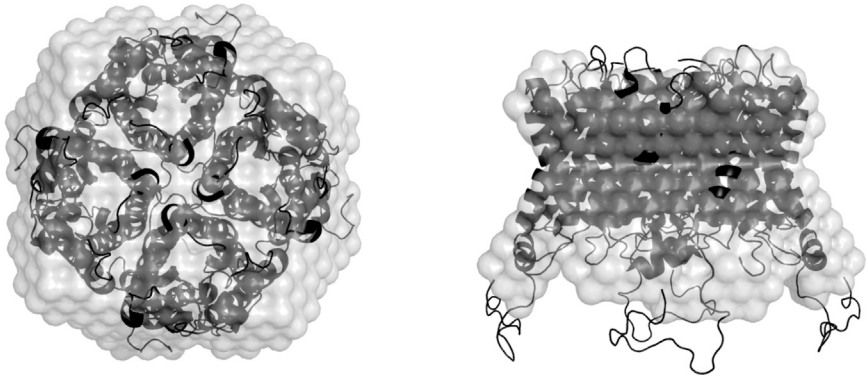

FIGURE 2 Comparison of most probable ab initio envelope (*shaded*) with the crystallographic structure of the AQ0 tetramer PDB: 2B6P (46) (*solid*). The reconstruction belongs to the ab initio run performed with two SANS curves as input and with imposed fourfold symmetry.

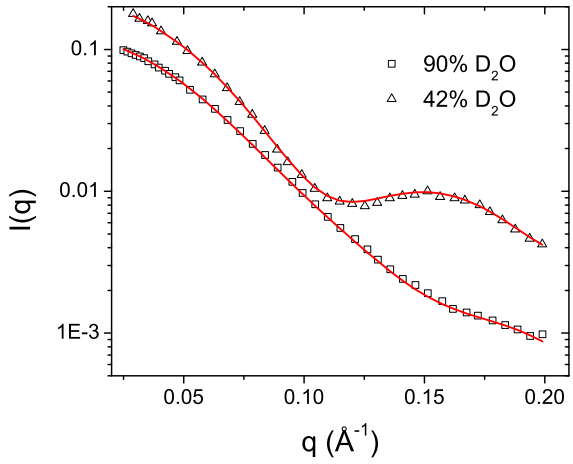

FIGURE 3 Fit of the two Fhac/OG_D SANS experimental curves by the ab initio algorithm. To see this figure in color, go online.

detergent heads, detergent tails, and solvent) together with a set of geometric constraints that reflect the anticipated structure of the detergent belt around the protein, which should resemble a core-shell doughnut-like structure. The importance of applying constraints may be also illustrated by performing ab initio reconstructions using the multiphase program MONSA (14), where no detergent belt related constraints are implemented and only bead connectivity and compactness is considered. Related results for the Fhac/OG_D system correspond to nonphysical final shapes, whereas reconstructions for the dDM/AQ0 system using three SANS curves as input (Fig. S1) reveal that in the final models, phases of similar contrast (protein, detergent heads) cannot be easily differentiated and the obtained protein shape possesses systematic differences from the crystallographic structure.

Ideally, an additional bead phase corresponding to the hydration layer that surrounds the entire PDC has to be included in our model (53), but because only a marginal improvement of the obtained shapes is expected (38) and due to the increased complexity of such an approach, we opted not to implement it in this algorithm. We stress that no particular geometry (for example, a torus) is assumed for the detergent belt. Two extra restraints that are imposed are related 1) to the maximum thickness of the belt that cannot exceed the typical cellular membrane thickness, and 2) to the total number of detergent molecules that are bound to the protein (aggregation number, N_{det}). By fixing N_{det} and considering the molecular volume of detergent heads and tails, we essentially fix the corresponding number of beads in the final solution.

TABLE 3 Slds at Different Contrasts and Molecular Volumes of Protein, Detergent Parts, and Solvent Used during the Ab Initio Runs for the Fhac/OG_D System

Bead Phase	$sld_{90\%D_2O}$ (10^{-6} A^{-2})	$sld_{42\%D_2O}$ (10^{-6} A^{-2})	V (Å^3)
OG_D heads	3.79	2.72	190
OG_D tails	7.28	7.28	230
Fhac protein	3.07	2.43	—
Solvent	5.70	2.36	—

Molecular volume (V) and sld values for the detergents were taken from the work of Breyton et al. (31). Protein sld is adjusted for each solvent contrast (56).

All the geometric constraints together with the requirement for model compactness and the discrepancy between the calculated and input scattering curves contribute to the defined score function $f(X)$. Successful minimization of the score function by simulated annealing leads to a bead arrangement that respects all the described constraints and also fits the small-angle scattering curves. The final bead arrangement represents the molecular envelope for each of the PDC components. The major finding of the minimizations of the MD-generated dDM/AQ0 scattering data is that a single curve, even with appropriate axial symmetry applied, cannot lead to a stable solution; and for obtaining molecular envelopes that are representative of the PDC shape, we need to provide at least two SANS curves at different solvent contrasts as input (Fig. 1). Addition of a third contrast does not improve the shape reconstructions. Enforcement of fourfold axial symmetry, as it is expected for the AQ0 tetramer, improves the agreement with the known crystallographic structure.

In an effort to better understand these results, we can think of the PDC system as having two principal contributions to the scattering—one coming from the protein and a second from the hydrophobic part of the detergent belt. The hydrophilic head layer that covers the belt is indeed contributing to the scattering, but by design, this structural feature is inherently considered in our model, because hydrophobic beads are separated from the solvent by a layer of hydrophilic beads that reproduce the scattering contribution of detergent head layers as determined in measurements of detergent micelles (31,44). Consequently, it is not surprising that two scattering curves at different contrast can lead to stable solution and to an objective shape for both the protein and the detergent belt. Here we note that the term "two different contrasts" refers to well-separated contrasts (>30% D_2O content). Additional tests have revealed that shape reconstruction quality is improved when one of the two contrasts is close to the matching point of the hydrophobic part of detergent micelles. During experimental planning and in the case of two SANS measurements, one should consider acquiring scattering data either close to the overall matching points of detergent tails and protein or at the most distant possible contrasts (0% and 100% D_2O). In the latter case, one of the two measurements will always be close to the matching point of detergent tails for both deuterated and hydrogenated hydrocarbon chains.

Having established that two contrast-separated SANS curves provide enough information for the ab initio reconstruction, we tested the algorithm using previously

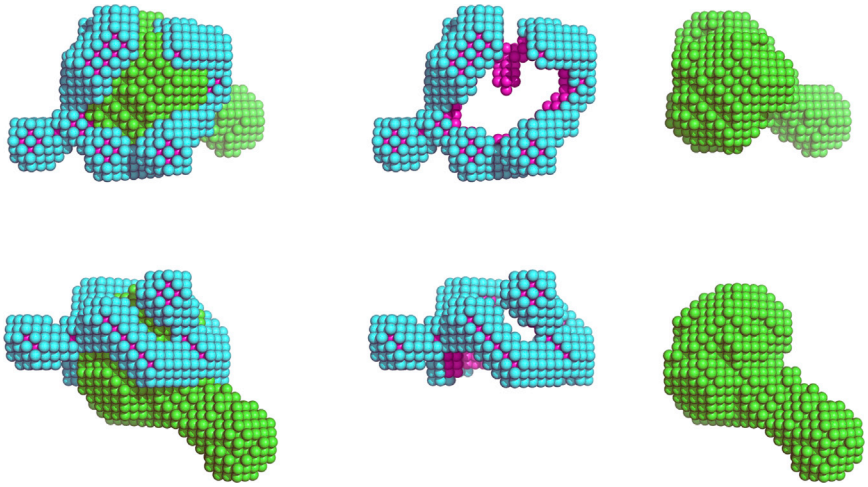

FIGURE 4 Most probable ab initio shape reconstruction of the 140 OG_D/Fhac PDC. (*First row*) Given here is the top view. (*Second row*) Given here is the side view. In each column, different PDC components (entire PDC, detergent belt, and protein, respectively) are depicted. (The bead colors correspond to *green*, protein; *cyan*, detergent hydrophilic heads; and *magenta*, detergent hydrophobic tails.) To see this figure in color, go online.

published SANS data of the Fhac transporter in two different solvent contrasts. The final obtained molecular shape of the protein is highly similar to the crystallographic structure (Fig. 5), with the β-barrel and POTRA domains easily distinguishable. It is interesting that the resulting detergent belt seems somewhat disordered and does not cover the entire β-barrel surface (Fig. 4). Inspection of PDC shapes belonging to different ab initio runs (Fig. S2) reveals detergent belt shapes with gaps and, in some cases, protrusions. The same topological variability was present in the all-atom MD models that gave the best fits to the scattering data in the work of Gabel et al. (22), which indicates, on the one hand, the relative ambiguity of small-angle data-derived models and, on the other hand, the dynamic nature of the PDC. In this respect, we note that a compact detergent belt is not a universal characteristic of solubilized membrane proteins, something that has to be carefully considered in modeling approaches.

Given that the number of protein-bound detergents N_{det} is an important input parameter in our method and because its estimation by experimental methods like analytical ultracentrifugation (22), refractometry, or absolute scale small-angle scattering (18) may contain systematic or statistical errors, it was relevant to investigate the stability of the shape reconstructions to small variations of N_{det}. Our findings suggest that an error of the order of $\pm 10\%$ is not appreciably affecting the shape reconstructions. This observation gives confidence that the typical uncertainty that characterizes the aforementioned experimental techniques should not affect the accuracy of our method.

The design of the PDC bead model was performed by having mainly in-focus systems of integral membrane proteins solubilized by detergents. However, our effort to not heavily constrain the detergent belt shape gives the ability to straightforwardly extend the applicability of this method to bicelle/protein systems (5). Provided that the contrast and structural parameters of empty bicelles are known from previous investigations, the bicelle belt can be modeled in the same manner as in the case of detergents. Furthermore, the method in its current formulation also seems appropriate for the study of peripheral membrane proteins that have only a small fraction of their amino-acid sequence anchored inside the detergent micelle.

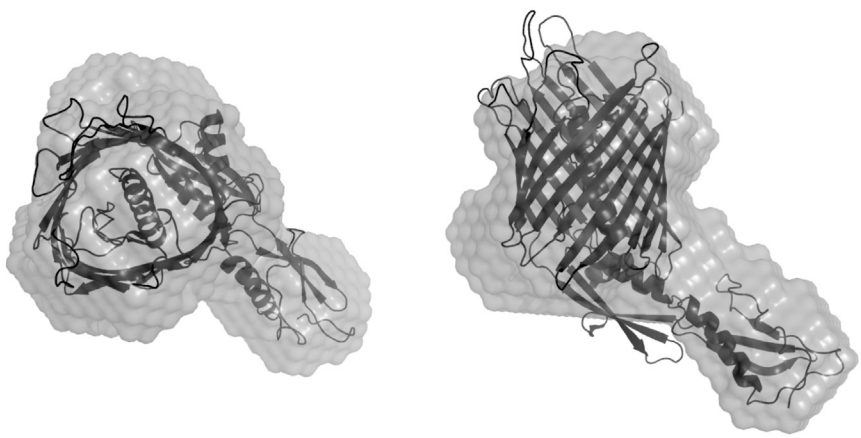

FIGURE 5 Comparison of most probable ab initio envelope (*light shading*) with the crystallographic structure of the Fhac transporter PDB: 2QDZ (52) (*solid*). (*Left*) Given here is the top view. (*Right*) Given here is the side view.

From the experimental and sample preparation point of view, it is absolutely essential to properly subtract free-micelle contributions from the scattering data to obtain curves corresponding solely to the PDC. Extensive dialysis against detergent-containing solutions is a viable method, but the in-situ combination of SEC and SAXS has proved to be a very efficient technique for this task (18). In this context and because our methodology relies on multicontrast SANS, the recent development of an in situ SEC-SANS platform at the Institute Laue-Langevin (54) presents special interest in relation to the acquisition of high-quality scattering data from detergent-solubilized membrane proteins. Such setups coupled with upcoming high flux diffractometers at future neutron sources like the European Spallation Source, may bring down the requirements for protein sample amounts, thus rendering the method as a medium-throughput option for membrane protein structure determination.

The executable code of the presented ab initio algorithm (program DANVILLE) for all major software platforms is freely available for academic use from the author upon request. A typical shape reconstruction with no imposed symmetry takes 5–20 h on the single core of modern processors.

CONCLUSIONS

In summary, we have presented a method for the determination of the low-resolution structure of detergent-solubilized membrane proteins from SANS datasets. Measurements at two different solvent contrasts together with an independent estimation of the detergent aggregation number around the protein appear to provide reliable shape reconstructions. The generality of the model permits the treatment of data from detergent- or bicelle-solubilized integral or peripheral membrane proteins where their membrane-embedded part is surrounded by a detergent belt of arbitrary shape. Given the extreme difficulties in the study of membrane protein structure, this work offers a powerful tool that allows the general use of small-angle scattering for detailed studies of membrane proteins in solution without the need for any prior structural information. Related ab initio approaches for soluble globular proteins have proved increasingly popular in the literature in recent years. In this respect, we expect that the proposed methodology coupled with the preparation of high-quality samples and the implementation of adapted experimental protocols will allow the investigation of challenging systems and will also complement current advances in other fronts like cryo-electron microscopy.

SUPPORTING MATERIAL

Two figures are available at http://www.biophysj.org/biophysj/supplemental/S0006-3495(17)31083-4.

AUTHOR CONTRIBUTIONS

A.K. designed and performed research, wrote scientific software, analyzed data, and wrote the manuscript.

ACKNOWLEDGMENTS

The author thanks Dr. Javier Pérez for his introduction into the general problem of the use of small-angle scattering for the study of membrane protein structure.

REFERENCES

1. Seddon, A. M., P. Curnow, and P. J. Booth. 2004. Membrane proteins, lipids and detergents: not just a soap opera. *Biochim. Biophys. Acta.* 1666:105–117.
2. Tribet, C., R. Audebert, and J.-L. Popot. 1996. Amphipols: polymers that keep membrane proteins soluble in aqueous solutions. *Proc. Natl. Acad. Sci. USA.* 93:15047–15050.
3. Bayburt, T. H., Y. V. Grinkova, and S. G. Sligar. 2002. Self-assembly of discoidal phospholipid bilayer nanoparticles with membrane scaffold proteins. *Nano Lett.* 2:853–856.
4. Knowles, T. J., R. Finka, …, M. Overduin. 2009. Membrane proteins solubilized intact in lipid containing nanoparticles bounded by styrene maleic acid copolymer. *J. Am. Chem. Soc.* 131:7484–7485.
5. Dos Santos Morais, R., O. Delalande, …, S. Combet. 2017. Contrast-matched isotropic bicelles: a versatile tool to specifically probe the solution structure of peripheral membrane proteins using SANS. *Langmuir.* 33:6572–6580.
6. Kang, C., and Q. Li. 2011. Solution NMR study of integral membrane proteins. *Curr. Opin. Chem. Biol.* 15:560–569.
7. Miles, A. J., and B. A. Wallace. 2016. Circular dichroism spectroscopy of membrane proteins. *Chem. Soc. Rev.* 45:4859–4872.
8. Le Roy, A., K. Wang, …, C. Ebel. 2015. AUC and small-angle scattering for membrane proteins. *Methods Enzymol.* 562:257–286.
9. Rawson, S., S. Davies, …, S. P. Muench. 2016. The changing landscape of membrane protein structural biology through developments in electron microscopy. *Mol. Membr. Biol.* 33:12–22.
10. Koch, M. H., P. Vachette, and D. I. Svergun. 2003. Small-angle scattering: a view on the properties, structures and structural changes of biological macromolecules in solution. *Q. Rev. Biophys.* 36:147–227.
11. Jacques, D. A., and J. Trewhella. 2010. Small-angle scattering for structural biology—expanding the frontier while avoiding the pitfalls. *Protein Sci.* 19:642–657.
12. Lipfert, J., and S. Doniach. 2007. Small-angle x-ray scattering from RNA, proteins, and protein complexes. *Annu. Rev. Biophys. Biomol. Struct.* 36:307–327.
13. Chacón, P., F. Morán, …, J. M. Andreu. 1998. Low-resolution structures of proteins in solution retrieved from x-ray scattering with a genetic algorithm. *Biophys. J.* 74:2760–2775.
14. Svergun, D. I. 1999. Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. *Biophys. J.* 76:2879–2886.
15. Cooley, R. B., J. P. O'Donnell, and H. Sondermann. 2016. Coincidence detection and bi-directional transmembrane signaling control a bacterial second messenger receptor. *eLife.* 5:e21848.
16. Zimmer, J., D. A. Doyle, and J. G. Grossmann. 2006. Structural characterization and pH-induced conformational transition of full-length KcsA. *Biophys. J.* 90:1752–1766.
17. Lipfert, J., L. Columbus, …, S. Doniach. 2007. Analysis of small-angle x-ray scattering data of protein–detergent complexes by singular value decomposition. *J. Appl. Cryst.* 40:s235–s239.

18. Berthaud, A., J. Manzi, …, S. Mangenot. 2012. Modeling detergent organization around aquaporin-0 using small-angle x-ray scattering. *J. Am. Chem. Soc.* 134:10080–10088.
19. Chen, P. C., and J. S. Hub. 2015. Structural properties of protein-detergent complexes from SAXS and MD simulations. *J. Phys. Chem. Lett.* 6:5116–5121.
20. Koutsioubas, A., A. Berthaud, …, J. Pérez. 2013. Ab initio and all-atom modeling of detergent organization around Aquaporin-0 based on SAXS data. *J. Phys. Chem. B.* 117:13588–13594.
21. Døvling Kaspersen, J., C. Moestrup Jessen, …, J. S. Pedersen. 2014. Low-resolution structures of OmpA•DDM protein-detergent complexes. *ChemBioChem.* 15:2113–2124.
22. Gabel, F., M. F. Lensink, …, C. Ebel. 2014. Probing the conformation of FhaC with small-angle neutron scattering and molecular modeling. *Biophys. J.* 107:185–196.
23. Pérez, J., and A. Koutsioubas. 2015. Memprot: a program to model the detergent corona around a membrane protein based on SEC-SAXS data. *Acta Crystallogr. D Biol. Crystallogr.* 71:86–93.
24. Gao, Y., G. Westfield, …, S. Ramachandran. 2017. Isolation and structure-function characterization of a signaling-active rhodopsin-G protein complex. *J. Biol. Chem.* 292:14280–14289.
25. Wojtowicz, H., A. Prochnicka-Chalufour, …, N. Izadi-Pruneyre. 2016. Structural basis of the signalling through a bacterial membrane receptor HasR deciphered by an integrative approach. *Biochem. J.* 473: 2239–2248.
26. Clifton, L. A., C. L. Johnson, …, J. H. Lakey. 2012. Low resolution structure and dynamics of a colicin-receptor complex determined by neutron scattering. *J. Biol. Chem.* 287:337–346.
27. Nogales, A., C. García, …, J. González-Rodríguez. 2010. Three-dimensional model of human platelet integrin αIIb β3 in solution obtained by small angle neutron scattering. *J. Biol. Chem.* 285:1023–1031.
28. Johs, A., M. Hammel, …, R. Prassl. 2006. Modular structure of solubilized human apolipoprotein B-100. Low resolution model revealed by small angle neutron scattering. *J. Biol. Chem.* 281:19732–19739.
29. Cardoso, M. B., D. Smolensky, …, H. O'Neill. 2009. Insight into the structure of light-harvesting complex II and its stabilization in detergent solution. *J. Phys. Chem. B.* 113:16377–16383.
30. Tang, K.-H., V. S. Urban, …, R. E. Blankenship. 2010. SANS investigation of the photosynthetic machinery of *Chloroflexus aurantiacus*. *Biophys. J.* 99:2398–2407.
31. Breyton, C., F. Gabel, …, C. Ebel. 2013. Small angle neutron scattering for the study of solubilised membrane proteins. *Eur. Phys. J. E Soft Matter.* 36:71.
32. Compton, E. L. R., E. Karinou, …, A. Javelle. 2011. Low resolution structure of a bacterial SLC26 transporter reveals dimeric stoichiometry and mobile intracellular domains. *J. Biol. Chem.* 286:27058–27067.
33. Maric, S., N. Skar-Gislinge, …, L. Arleth. 2014. Stealth carriers for low-resolution structure determination of membrane proteins in solution. *Acta Crystallogr. D Biol. Crystallogr.* 70:317–328.
34. Kynde, S. A. R., N. Skar-Gislinge, …, L. Arleth. 2014. Small-angle scattering gives direct structural information about a membrane protein inside a lipid environment. *Acta Crystallogr. D Biol. Crystallogr.* 70:371–383.
35. Skar-Gislinge, N., S. A. R. Kynde, …, L. Arleth. 2015. Small-angle scattering determination of the shape and localization of human cytochrome P450 embedded in a phospholipid nanodisc environment. *Acta Crystallogr. D Biol. Crystallogr.* 71:2412–2421.
36. Franke, D., and D. I. Svergun. 2009. *DAMMIF*, a program for rapid ab initio shape determination in small-angle scattering. *J. Appl. Cryst.* 42:342–346.
37. Koutsioubas, A., and J. Perez. 2013. Incorporation of a hydration layer in the dummy atom ab initio structural modeling of biological macromolecules. *J. Appl. Cryst.* 46:1884–1888.
38. Koutsioubas, A., S. Jaksch, and J. Pérez. 2016. DENFERT version 2: extension of ab initio structural modelling of hydrated biomolecules to the case of small-angle neutron scattering data. *J. Appl. Cryst.* 49:690–695.
39. Debye, P. 1915. Zerstreuung von Röntgenstrahlen. *Ann. Phys.* 351: 809–823.
40. Svergun, D., C. Barberato, and M. H. J. Koch. 1995. *CRYSOL*—a program to evaluate x-ray solution scattering of biological macromolecules from atomic coordinates. *J. Appl. Cryst.* 28:768–773.
41. Sokolova, A. V., V. V. Volkov, and D. I. Svergun. 2003. Prototype of a database for rapid protein classification based on solution scattering data. *J. Appl. Cryst.* 36:865–868.
42. Lake, J. 1967. An iterative method of slit-correcting small angle x-ray data. *Acta Crystallogr.* 23:191–194.
43. Tanford, C. 1980. The Hydrophobic Effect: Formation of Micelles and Biological Membranes, 2nd Ed. J. Wiley, Hoboken, NJ.
44. Oliver, R. C., J. Lipfert, …, L. Columbus. 2013. Dependence of micelle size and shape on detergent alkyl chain length and head group. *PLoS One.* 8:e62488.
45. Svergun, D. I., S. Richard, …, G. Zaccai. 1998. Protein hydration in solution: experimental observation by x-ray and neutron scattering. *Proc. Natl. Acad. Sci. USA.* 95:2267–2272.
46. Gonen, T., Y. Cheng, …, T. Walz. 2005. Lipid-protein interactions in double-layered two-dimensional AQP0 crystals. *Nature.* 438:633–638.
47. Svergun, D. 1992. Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J. Appl. Cryst.* 25:495–503.
48. Volkov, V. V., and D. I. Svergun. 2003. Uniqueness of ab initio shape determination in small-angle scattering. *J. Appl. Cryst.* 36:860–864.
49. Kozin, M. B., and D. I. Svergun. 2001. Automated matching of high- and low-resolution structural models. *J. Appl. Cryst.* 34:33–41.
50. Wriggers, W., and P. Chacón. 2001. Using *Situs* for the registration of protein structures with low-resolution bead models from x-ray solution scattering. *J. Appl. Cryst.* 34:773–776.
51. Schrödinger, LLC. 2015. The PyMOL Molecular Graphics System, V. 1.8. https://pymol.org/2/.
52. Clantin, B., A.-S. Delattre, …, V. Villeret. 2007. Structure of the membrane protein FhaC: a member of the Omp85-TpsB transporter superfamily. *Science.* 317:957–961.
53. Rambo, R. P., and J. A. Tainer. 2013. Super-resolution in solution x-ray scattering and its applications to structural systems biology. *Annu. Rev. Biophys.* 42:415–441.
54. Jordan, A., M. Jacques, …, A. Martel. 2016. SEC-SANS: size exclusion chromatography combined *in situ* with small-angle neutron scattering. *J. Appl. Cryst.* 49:2015–2020.
55. Jacrot, B. 1976. The study of biological structures by neutron scattering from solution. *Rep. Prog. Phys.* 39:911–953.
56. Sarachan, K. L., J. E. Curtis, and S. Krueger. 2013. Small-angle scattering contrast calculator for protein and nucleic acid complexes in solution. *J. Appl. Cryst.* 46:1889–1893.

Biophysical Journal, Volume *113*

Supplemental Information

Low-Resolution Structure of Detergent-Solubilized Membrane Proteins from Small-Angle Scattering Data

Alexandros Koutsioubas

Supporting Information:
Low Resolution Structure Determination of Detergent Solubilised Membrane Proteins from Small Angle Scattering Data.

Alexandros Koutsioubas*
Jülich Centre for Neutron Science (JCNS) at Heinz Maier-Leibnitz Zentrum (MLZ), Forschungszentrum Jülich GmbH, Lichtenbergstr. 1, 85748 Garching, Germany

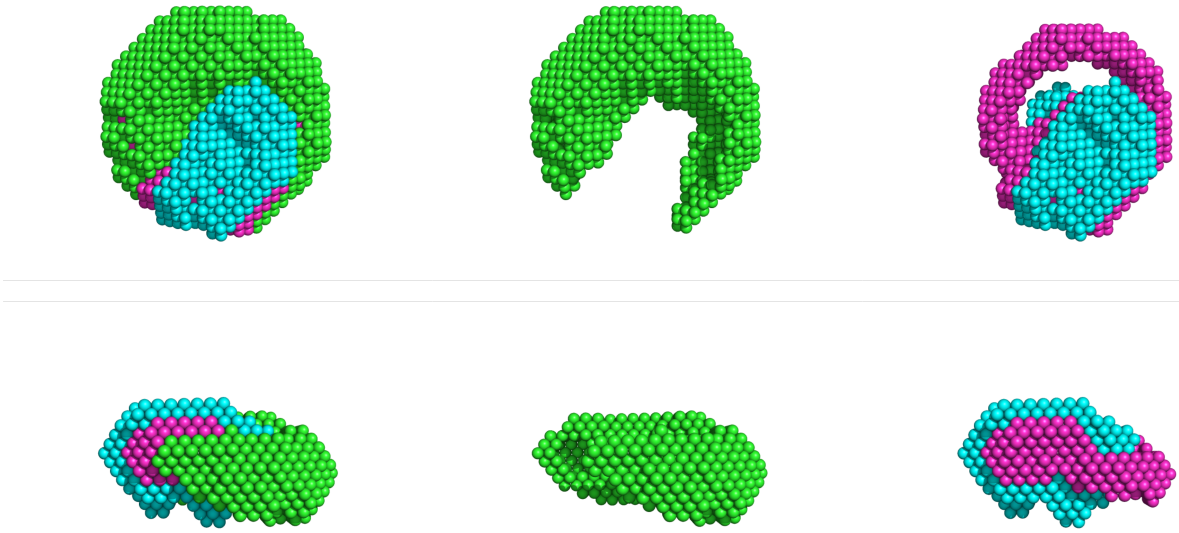

FIG. S1. Ab-initio shape reconstructions of the 270 dDM/AQ0 tetramer using the program MONSA [1], 3 SANS curves as input (0%, 42% and 100% D_2O) and default parameters concerning bead compactness and connectivity, volume fraction and center penalty. Bead colours correspond to: green-protein, cyan - detergent hydrophilic heads and magenta - detergent hydrophobic tails. The volumes of each different phase are provided as input. This is equivalent to providing the detergent aggregation number and head/tail molecular volumes, however the protein volume is an extra parameter that has to be provided to MONSA. The used scattering length densities are the same as those given in Table 1 of the manuscript. Note that due to the similar contrast of detergent heads and protein in all MONSA reconstructions the protein phase beads are positioned erroneously around the detergent tail phase while the detergent head beads form a core close to the centre of the structure. Even considering this core (cyan beads) as the structure representing the protein, we find systematic structural differences ($NSD_{cryst} > 1.0$) comparing to the Aquaporin-0 crystallographic structure. Similar MONSA reconstructions using the two OG_D/Fhac experimental SANS curves as input led always to non-physical final shapes for both the protein and the detergent belt.

* a.koutsioumpas@fz-juelich.de

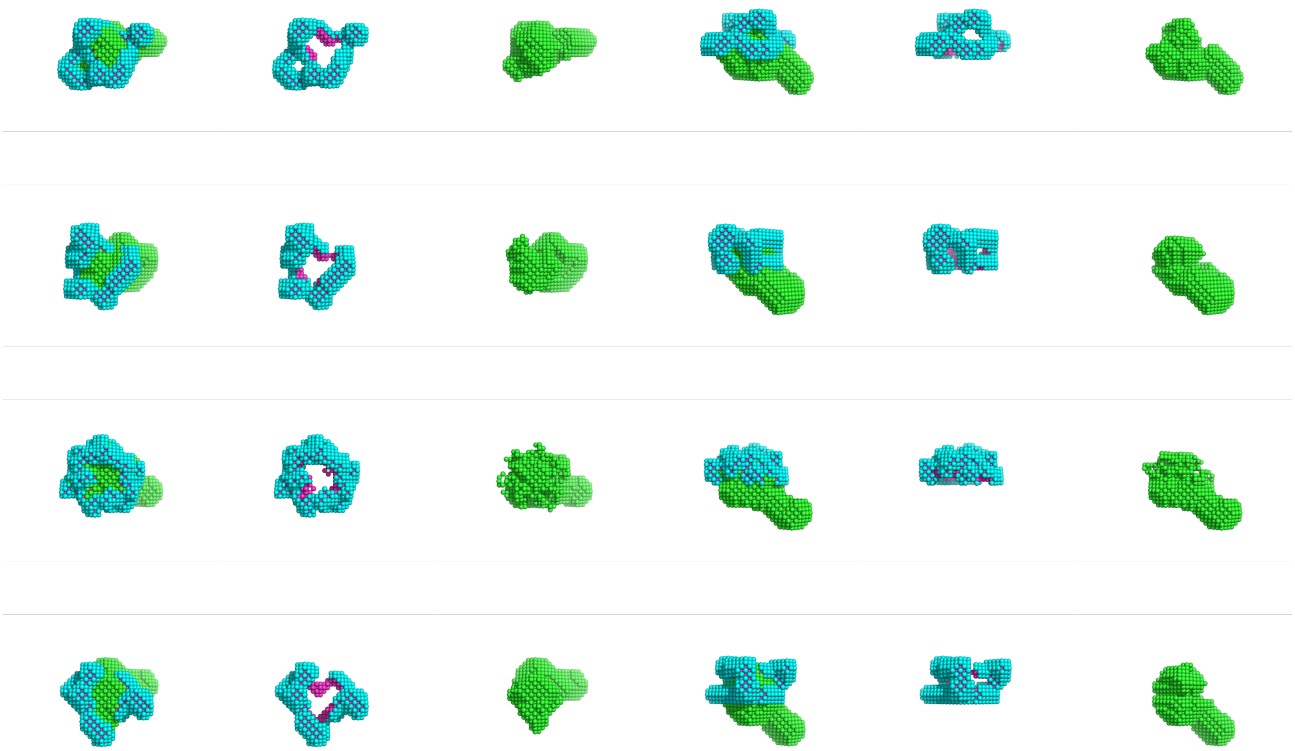

FIG. S2. Results of different ab-initio shape reconstructions of the 140 OG_D/Fhac PDC as the one presented in the manuscript's fig. 4. Each line represents a different ab-initio run. In the first three columns the top view of the different PDC components (entire PDC, detergent belt and protein respectively) are depicted, while in the last three columns the side view. The bead colours correspond to: green - protein, cyan - detergent hydrophilic heads, magenta - detergent hydrophobic tails.

[1] Svergun, D. I., 1999. Restoring Low Resolution Structure of Biological Macromolecules from Solution Scattering Using Simulated Annealing. *Biophysical journal* 76:2879–2886.