Ab Initio and All-Atom Modeling of Detergent Organization around Aquaporin-0 Based on SAXS Data

ARTICLE *in* THE JOURNAL OF PHYSICAL CHEMISTRY B · OCTOBER 2013

Impact Factor: 3.3 · DOI: 10.1021/jp407688x · Source: PubMed

4 AUTHORS:

Alexandros G Koutsioubas
Forschungszentrum Jülich
35 PUBLICATIONS **228** CITATIONS
SEE PROFILE

Alice Berthaud
Institut Curie
3 PUBLICATIONS **28** CITATIONS
SEE PROFILE

Stéphanie Mangenot
Institut Curie
22 PUBLICATIONS **517** CITATIONS
SEE PROFILE

Javier Perez
SOLEIL synchrotron
78 PUBLICATIONS **1,397** CITATIONS
SEE PROFILE

THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

pubs.acs.org/JPCB

Ab Initio and All-Atom Modeling of Detergent Organization around Aquaporin-0 Based on SAXS Data

Alexandros Koutsioubas,[†,§] Alice Berthaud,[†,‡] Stéphanie Mangenot,[‡] and Javier Pérez*[,†]

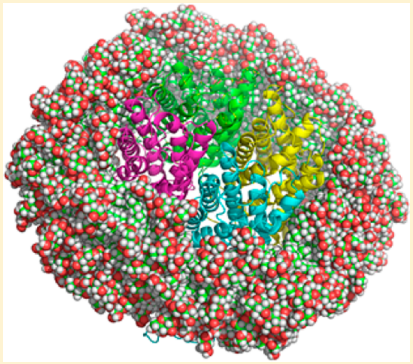
[†]Synchrotron Soleil, Beamline SWING, Saint Aubin BP48, F-91192 Gif sur Yvette Cedex, France

[‡]Institut Curie, Centre de Recherche, CNRS UMR168, Université Pierre et Marie Curie, F-75248 Paris Cedex, France

[§]Jülich Centre for Neutron Science JCNS, Forschungszentrum Jülich GmbH, Outstation at MLZ, Lichtenbergstrasse 1, 85747 Garching, Germany

Supporting Information

ABSTRACT: A necessary initial step for the application of small angle X-ray scattering (SAXS) as an analytical probe for structural investigations of membrane proteins in solution is the precise knowledge of the structure of spontaneously formed detergent assemblies around the protein. Following our recent article (Berthaud et al. *J. Am. Chem. Soc.* **2012**, *134*, 10080−10088) on the study of the n-dodecyl β-D-maltopyranoside (dDM) corona surrounding Aquaporin-0 tetramers in solution, we aimed at the development of more elaborate models, exploiting the information content of the scattering data. Two additional approaches are developed here for the fit of SAXS experimental data, one based on a generalized ab initio algorithm for the construction of a coarse-grained representation of the detergent assemblies, and a second based on atomistic molecular dynamics. Accordingly, we are able to fit the SAXS experimental data and obtain a better insight concerning the structure of the detergent corona around the hydrophobic part of the Aquaporin-0 surface. The present analysis scheme represents an additional step toward future conformational studies of transmembrane proteins in solution.



INTRODUCTION

Membrane proteins perform their biological role in the cellular membrane, and they are involved in a wide range of cellular functions. The difficulty to study their structure is reflected in the low, although rapidly increasing, number of deposited entries in the Protein Data Bank.[1] This fact is partially related to the need of membrane protein solubilization by a membrane-mimetic detergent or lipid assembly around their transmembrane surface.[2] Quite often this prerequisite may hinder crystallization, which is necessary for X-ray crystallographic studies, while may also complicate NMR studies due to protein-detergent interactions.[3]

Small angle X-ray scattering (SAXS) is a well-established low resolution technique, adapted for the study of ternary and quaternary conformational changes of soluble proteins under physiological conditions.[4] However the applicability of the technique in the case of membrane proteins has been also hindered by the necessary presence of detergent molecules in solution that permit membrane protein solubilization, leading to extra contributions to the measurable signal that are difficult to separate from the protein signal. These unwanted contributions are related to the formation of detergent free micelles in solution and also to the corona-like self-assembled detergent structure that covers the hydrophobic surface of the protein.

Several approaches for the separation of protein/detergent complex and micellar signal have been proposed, including electron density matching and collection of scattering data at different protein-detergent stoichiometries.[5] However, these methods suffer from inevitable perturbations of the system and from the need of multiple protein/detergent samples respectively. On the other hand, the coupling of the SAXS technique with size-exclusion high-pressure liquid chromatography (SE-HPLC) in a Synchrotron source,[6] permits the size separation of different molecular assemblies in solution and the concurrent acquisition of unbiased SAXS curves ($I(q)$ vs q). Through this same methodology, high-quality scattering data concerning the free micelles and the pure protein/detergent belt complexes can be acquired.

Despite this very important experimental achievement, there is still the need of a method in order to decouple the protein/detergent contributions from the detergent-protein complex. Due to their different chemical composition, detergent heads and tails, have quite different electron densities that are both different from the typical mean electron density of a protein molecule. So the relative complexity of the protein/detergent complex means that the usual analysis of SAXS data, based on the calculation of scattering invariants or the use of "ab initio" modeling,[7−9] is not applicable in a straightforward manner. Even in the case of the complementary to SAXS small angle

Received: August 1, 2013
Revised: September 24, 2013
Published: September 27, 2013

dx.doi.org/10.1021/jp407688x | *J. Phys. Chem. B* 2013, 117, 13588−13594

neutron scattering (SANS) technique, where the contrast of different species in solution can be systematically varied, adequate contrast matching of the two different detergent parts is very difficult to achieve.[10]

In a recent article,[11] we have shown that SE-HPLC coupled SAXS measurements from dDM solubilized Aquaporin-0 tetramers can be modeled satisfactory by an elliptic toroid coarse-grained bead model of the detergent belt around the hydrophobic surface of the protein. This work with a membrane protein of known structure represents a first step toward a detailed description of the detergent belt, opening a possible way for future conformational studies of membrane proteins in solution. The fits with this geometrical model gave parameters concerning detergent organization (overall thickness of the detergent layer, extension of hydrophobic/hydrophilic regions) that are in agreement with previous experimental studies of detergent micelles[12] and with independent measurements from refractometry coupled UV absorption spectroscopy.[11]

In an effort to overcome the restrictions that are imposed by the geometrical representation of the detergent corona, we propose here a "model-free" ab initio coarse-grained fitting-algorithm that takes into account the different electron densities and general physical properties of the system. Application of this algorithm provides high quality fits of the SAXS data and a more objective estimation of the detergent corona shape. At a second step, by executing molecular dynamics (MD) simulations together with a special methodology for the initial equilibration of the system, we are able to obtain a detailed all-atom model that is consistent with the experimental results for Aquaporin-0/dDM system to an extent never reached until now.

In light of these findings, we improve the interpretation of our previous modeling approach and also attempt to draw conclusions concerning the effects of protein/detergent complexation, and the potential general applicability of the proposed methodology in membrane protein structural studies.

METHODS

Ab-Initio Algorithm. Following the methodology that was initially proposed by Chacon,[7] further improved by Svergun[8,13,14] and implemented in the popular programs DAMMIN, DAMMIF, and MONSA for the low-resolution restoration of biological structures from small-angle scattering data, we have developed an algorithm where a set of different types of beads that are placed around the membrane protein, participate in a simulated annealing procedure that aims to arrange the beads in a way that is compatible with the experimentally measured scattering. Comparing to the mentioned ab initio approaches, here the model is hybrid, because the calculation of the scattering of the system contains contributions from the all-atom structure of the protein (PDB code 2B6P[15] with added residues 1 and 36) and from a coarse-grained description of the surrounding detergent corona.

Two types of beads (voxels) with different electron densities are introduced ($\rho_{\text{hydrophilic}} \cong 0.52e/\text{Å}^3$) and $\rho_{\text{hydrophilic}} \cong 0.27e/\text{Å}^3$), populating regions of detergent hydrophilic heads and hydrophobic tails. These beads are placed in a close-packed face-centered cubic (fcc) lattice of spacing L, and they are not allowed to overlap with atoms of the known protein structure. In this lattice, each bead may have up to 12 nearest-neighbors.

The initial configuration is prepared by the construction of a hollow thin cylinder of hydrophobic beads around the approximate trasmembrane part of the protein (with no protein-bead overlaps). The outer layer of this cylinder is then covered by a layer of hydrophilic beads, so that no contact between hydrophobic beads and empty lattice points (representing the solvent) are present in the initial model (see also Supporting Information 1 (SI1)). Initial cylinder configurations of radius in the range 60 ± 10 Å and thickness in the range 30 ± 10 Å, always led to the convergence of the simulated annealing algorithm toward similar final solutions.

As it is well documented[16,17] the hydration layer around the protein surface, is expected to contribute substantially to the measured scattering. In order to partially take into account this extra contribution, the bound water at the solvent exposed protein surface is represented by a third type of beads that correspond to a hydration layer of r_h = 3 Å thickness and electron density contrast $\delta\rho_b = 0.03e/\text{Å}^3$.[18] Initial placement of hydration beads is performed by finding the most distant non-hydrogen atom (of atomic radius r_{atom}) of the protein in each direction around its center of mass, and by placing a hydration bead on the nearest fcc lattice point that resides on this direction and that also has a distance less than the sum $r_{\text{atom}} + r_w$, where r_w = 1.5 Å is the radius of the water molecule.

These hydration layer beads cover the entire protein surface creating a monolayer between the protein surface and the "solvent". Since the contribution of hydration to the scattering depends primarily on the product of hydration layer thickness and layer contrast, the electron density of the hydration beads in each simulated annealing run is adjusted so that the product of layer thickness and contrast is equal to $3{\cdot}0.03e/\text{Å}^2$. Hydration beads also participate in the simulated annealing procedure, so that they may give their place to hydrophobic or hydrophilic beads that touch the protein surface. Ideally, a hydration layer around the surface of the detergent corona should also be modeled. However inclusion of such a feature would mean that a single bead modification trial may affect multiple neighboring beads, which would largely increase the complexity of the algorithm at the present step.

The goal of the simulated annealing procedure is to find interconnected and compact arrangements of beads that are also characterized by a scattering curve of minimum discrepancy comparing to the experimental curve. In order to ensure the convergence of the algorithm toward physically meaningful structures, it was found necessary to impose an additional rule related to the general properties of detergent assemblies, by which any contact between a hydrophobic bead and a hydration bead or an empty lattice position that represents the solvent is prohibited. So since there are no "solvent exposed" hydrophobic beads in the starting configuration, trials that expose hydrophobic beads are immediately canceled.

The scattering curve $I(q)$ versus q for the initial configuration is calculated using the Debye equation[19]

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

where $f_i(q)$ is the atomic form factor or the bead form factor, K is the sum of the protein atomic groups (taking into account covalently bound hydrogens) plus the number of beads of all types (hydrophilic, hydrophobic, hydration), and r_{ij} is the distance between the ith and jth atomic group/bead. The atomic form factor is calculated as $f_i(q) = f_{i,\text{vacuum}}(q) - \rho_0 g_i(q)$, where $f_{i,\text{vacuum}}(q)$ is the atomic form factor in vacuum and $g_i(q)$

the form factor of a Gaussian sphere that represents the excluded solvent volume.[18] The hydrophobic/hydrophilic beads form factor is that of a sphere of electron density contrast $\delta\rho$, and with a volume given by $V_{bead} = L^3/4$.

During the simulated annealing procedure, random positions of the fcc lattice are selected, and trial additions/removals or changes of the type of beads are attempted in an effort to obtain a bead configuration X that minimizes a score function

$$f(X) = \chi^2 + \alpha_1 P_1(X) + \alpha_2 P_2(X) \quad (2)$$

which contains contributions from the discrepancy χ^2 between the intensity of the X configuration and the experimental scattering intensity and also from the models' looseness $P_1(X)$ and connectivity $P_2(X)$.[8,14] The acceptance of trials is dependent on the annealing temperature T, which decreases during the annealing procedure. Since only one bead is modified in each trial, the recalculation of only K cross terms of eq 1 are needed, thus avoiding the computationally intensive recalculation of the whole double sum.

A final compact and interconnected bead configuration is characterized by penalty values $P_1(X) \approx 0.02$ and $P_2(X) = 0$. Accordingly, the values of parameters $(\alpha_1,\alpha_2 \cong 10)$ in eq 2 were determined by test runs, so that their final contribution in the $f(X)$ is of the order of 10%. The annealing procedure proceeds as suggested in ref 3, with 10^5 trials at each annealing temperature, and a cooling step given by $T' = 0.9T$. The algorithm is executed till no improvement in the score function is observed.

Molecular Dynamics Simulations. All-atom modeling of the detergent assemblies around Aquaporin-0 involved a special pre-equilibration procedure for the preparation of initial protein/detergent corona configurations and subsequent 10 ns molecular dynamics simulations of the system. To a large extent, for the setup of the system and the simulation details, we follow the general methodology from the work of Bond, Sansom et al.[20−22] It is important to note that the SAXS pattern is never used as a constraint in the MD simulation.

The initial "micelle-like toroid" structure of the protein/detergent system was prepared by adding N vacuum equilibrated dDM molecules around the approximate transmembrane surface of the Aquaporin-0 tetramer, with their molecular axes radiating from the mean perimeter of the transmembrane protein part, and with the maltose groups situated in the outer surface of the structure. The minimal intramolecular distance was set equal to 2 Å in order to avoid atomic overlaps.

Subsequently, the protein/detergent complex was subjected to nonequilibrium MD runs, where two geometrical parameters are used in order to shape the dDM corona around the backbone-restrained protein, through the application of a mild external force. In more detail, an algorithm periodically checks the number of dDM molecules with maltose heads having a distance greater than h from the plane that crosses the transmembrane part of the protein and/or having a distance larger than d from the protein surface. On each atom of these dDM molecules, an external constant acceleration starting at 1 Å/ps^2 is imposed with a direction either toward the pore axis of the protein or toward the plane that crosses the protein pore axis. This procedure is iteratively applied every 10 ps more than 400 times, by also progressively decreasing the magnitude of the imposed acceleration, until finally no acceleration is applied. The two geometrical parameters (h,d) help to define the shape of the detergent corona and also help the production of more regular configurations. The values of the two parameters are systematically varied in an effort to obtain compact corona structures that are characterized by scattering curves similar to the experimentally acquired one. For the present system the used values of the parameters were $h = 21 \pm 1$ Å and $d = 23 \pm 1$ Å. Details on the MD runs that were performed using GROMACS[23−25] are presented in SI2.

RESULTS

In our previous work concerning the structure of the detergent corona around Aquaporin-0,[11] the model is based on a geometrical representation of an elliptic two component torus with four independent parameters (a, b, e, and t), where be, b/e, and a are the major and minor semi axis and height of the hydrophobic core, and t is the thickness of the hydrophilic outer layer. This approach proved successful in the fit of SAXS experimental data, giving detergent aggregation numbers around the protein that agree with independent estimates from refractometry and SAXS forward scattering. In addition, the obtained parameters a,t are fully consistent with expected values from experimental investigations of detergent micelles. However, the necessary ellipticity that is introduced in the model has no direct physical foundation.

In an effort to propose a model with physically meaningful constraints, we have developed an ab initio algorithm for the coarse-grained description of the detergent corona. Instead of using a predefined geometric model for the construction of the two-component detergent belt, we rely on the versatility of simulated annealing for the arbitrary placement of "hydrophobic" and "hydrophilic" beads around the crystallographic structure of Aquaporin-0, in order to reproduce the SAXS results. The only physically relevant restriction concerns the prohibition of "hydrophobic" beads being in contact with solvent. This approach is essentially "model-free" and may capture detergent corona structure in a more general manner. In this sense the hybrid model represents a step forward in a more objective treatment of the system.

The electron densities used for the different parts of the system were chosen to be close to the values found for the geometric coarse grained model, that are also in agreement with values previously reported in the literature for dDM micelles.[12] We have found that the algorithm could not reach low χ^2 if large deviations from these values are imposed. A large number of simulated annealing runs with different number of starting configurations and various bead sizes, always converged toward similar final models characterized by Normalized Spatial Discrepancy (NSD) values[26] in the range 0.45−0.55, thus revealing the stability of the final solution. We were able to obtain high quality fits of the SAXS data with $\chi^2 \cong 1.0$ (Figure 1), while final hybrid models (Figure 2) were characterized by a very high similarity with the all-atom structures that are discussed in the following. The estimated number of detergent molecules in the corona was found equal to 340 ± 10.

Going beyond the coarse-grained description of the detergent corona, we attempted to make a direct comparison with the structure of the protein/detergent complex as it is obtained through MD simulations. For this purpose, 10 ns MD runs of preformed Aquaporin/dDM complexes were executed with a variable number of detergent molecules ($N = 220 - 320$). Since the number of molecules in the MD simulations is fixed, multiple simulations with variable number of detergents are needed in order to test different detergent aggregation numbers around the protein. Furthemore, effects associated

dx.doi.org/10.1021/jp407688x | J. Phys. Chem. B 2013, 117, 13588−13594

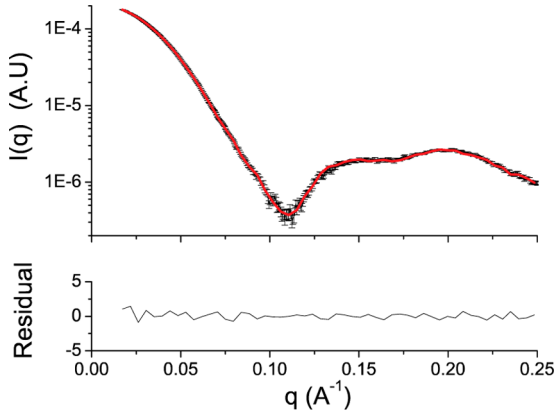


Figure 1. Best fit with the hybrid model and standard deviation normalized residuals.

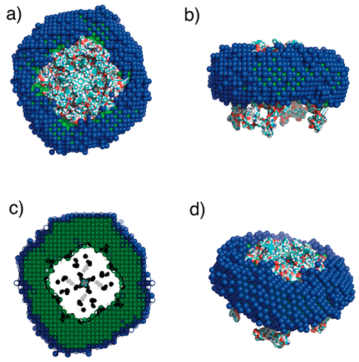


Figure 2. Hybrid ab initio model of the protein/detergent complex. Blue and green beads represent the hydrophilic and hydrophobic parts of the detergent corona, respectively. Note that the hydration beads around the solvated protein surface are omitted for clarity. (a) Top view; (b) side view; (c) cross-section perpendicular to the pore axis; (d) tilted view.

with the possible structural deformation of the protein inside the detergent environment are tested by keeping the protein backbone structure either constrained or not.

Our early observations showed that when the MD starting configurations were far from what is expected for the overall size and shape of the detergent corona, the relaxation of the system may easily reach very large simulation times (over 20 ns). Inspection of related system snapshots revealed non-compact arrangement of detergent molecules, and large discrepancies with the experimentally acquired scattering data especially in the $q > 0.1$ Å^{-1} range (see Figure S1).

For these reasons, special care was taken for the preparation of initial equilibrated configurations for the MD runs. In order to force the protein/detergent complex to more compact structures with overall shape closer to what is expected from our previous models, we chose to impose a nonequilibrium external force on the dDM molecules. Two parameters h and d that represent the maximal distance of an atom belonging to a dDM molecule from the transmembrane plane are used. During MD equilibration runs, at different number of dDM molecules N, the nonequilibrium force tends to bring dDM molecules within the spatial limits that are imposed by the values of h,d.

After this pre-equilibration step, we have observed that the radius of gyration of the complex is quite stable during 10 ns of MD runs (see Figure S2a), meaning that the system is close to an equilibrium state. Comparison of calculated scattering curves at different simulation times also revealed the structural stability of the complex (see Figure S3). In the case of nonrestrained protein backbone runs, the root-mean-squared deviation (RMSD) from the crystallographic structure of the protein reached plateau-values of about 4−5 Å within 3−4 ns, thus providing an indication of a mild swelling of Aquaporin-0 within the complex. This phenomenon of structural drift in the micellar environment has also been observed in previous MD studies, like for OmpA membrane protein solubilized by DPC molecules.[20] Unlike in ref 20, extended drift of the protein at higher times is not observed (see Figure S2b).

Overall, we note that the special pre-equilibration procedure together with the use of dDM adapted force field parameters from the work of Abel el al.[27] results in protein/corona structures with scattering curves that are characterized by fair agreement with the experimentally acquired data (Figure 3), far

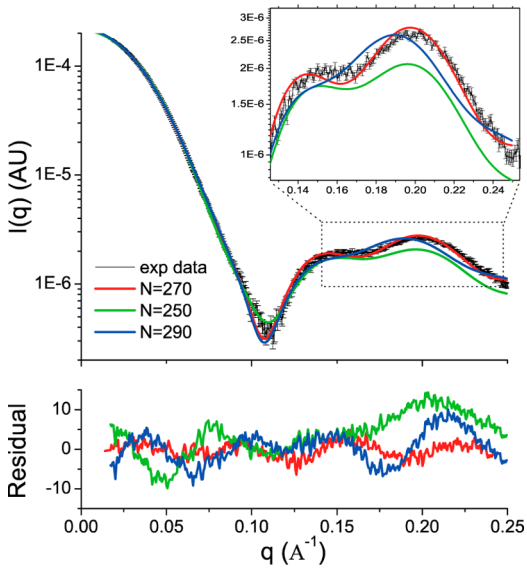


Figure 3. Comparison of SAXS experimental results with time averaged scattering curves from the MD last 5 ns. MD simulated Aquaporin-0/dDM system (backbone restrained protein) with N = 250, 270, 290. The standard deviation normalized residuals between the experimental and simulated curves are plotted at the bottom. (inset) magnification of the $q = 0.12 - 0.25$ Å^{-1} region.

better actually than without pre-equilibration and extremely satisfying considering that the SAXS curve was not used as a "driving potential". The R_g values of the simulations for (N = 250−290), tend to agree with the experimentally found value $R_{g,exp}$ = 44 Å, within ±1 Å. Fine details are best captured (χ = 2.0 ± 0.2) by the model with 270 dDM molecules, especially for $q > 0.1$ Å^{-1}, that is, the q-range where the "overall fold" of the structure contributes the most. By leaving the protein backbone free during N = 270 MD runs, we found slightly higher discrepancies with the experimental curve ($\chi^2 \cong 2.5$; see Figure S4). So for the rest of the discussion, we would refer to the obtained structure (see Figure 4) from the backbone-restrained N = 270 MD simulation, as the best-fit all-atom model.

DISCUSSION

Compared to the all-atom model, the ab initio algorithm certainly reaches lower χ^2 values. This is a direct consequence of the lower number of constraints in the hybrid model.

dx.doi.org/10.1021/jp407688x | J. Phys. Chem. B 2013, 117, 13588−13594

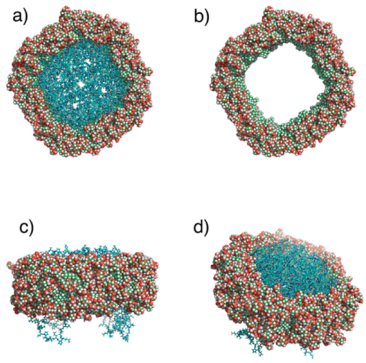


Figure 4. Several views of the $N = 270$ protein-backbone restrained all-atom protein/detergent complex. (a) Top view; (b) top view without protein atoms; (c) side view; (d) tilted view.

Table 1. Summary of Structural Parameters and Goodness of Fit for the Three Different Modeling Approaches

model	R_g (Å)	D_{max} (Å)	N	thickness (Å)	χ^2
geometric[11]	44.0	120	271 ± 30	41	1.4
hybrid	44.1	128	340 ± 31	45	1.0
all atom	44.2	124	270	42	2.0

Despite its small shortcomings in the reproduction of the experimental data (mainly in the region near the curve minimum), the all-atom model may still be regarded as a solid foundation for future all-atom modelization attempts, taking particular account of the extremely bad experimental agreement obtained with a system not using our specific pre-equilibration, and the fact that the SAXS curve was not used as a driving force in the molecular dynamics. Consequently, in the following subsections we will focus on several structural aspects of detergent organization around Aquaporin-0, as revealed by the hybrid and all-atom best-fit models.

General Detergent Corona Shape and Radial Electron Density Distributions. Two structural parameters that can be evaluated without specific modeling from the scattering curve, are the radius of gyration R_g and the maximum diameter of the studied structure D_{max}. Both current modeling approaches give identical R_g values ($R_{g,MD}$ = 44.2 Å, $R_{g,hybrid}$ = 44.1 Å) that agree with the experimental one R_g = 44.1 Å as calculated through Guinier analysis corresponding to a reduced range of $0.6 < qR_g < 1.3$. Using the program GNOM,[28] we may calculate the pair-distance distribution function through an indirect Fourier transform of the experimental scattering curve, and obtain the D_{max}, which is found equal to 124 Å. By direct geometric evaluation we find for the all-atom model exactly the same value and a slightly higher estimation for the hybrid model $D_{max(ab\ initio)}$ = 128 Å. It should be noted that in our ab initio algorithm, Dmax is not an a priori fixed parameter. The bead model is left free to adjust its size during the simulated annealing procedure.

Regarding the shape of the detergent corona, we notice that the all-atom model follows the overall shape of the transmembrane surface, something that is intuitively expected. So the corona surface tends to adopt the square-like shape of the Aquaporin-0 tetramer (see Figure 4). The same trend is observed in the case of the hybrid model with a somewhat more rounded periphery, a feature that possibly contributes to the better fits obtained with this approach. Both current models possess a natural roughness on the interfaces between different parts of the structure (Figure 2 and Figure 4), something that could not be present in our previous geometric models.[11] Indeed for a detergent corona around a protein we may expect that the system may be characterized by a certain degree of heterogeneity, dynamic polydispesity and conformational fluctuations that in our experiments would be averaged in the final SAXS curves.[11] The introduction of ellipticity in the geometric model partially compensates for these effects, but it can not represent a real feature of detergent shape.

We note that comparison of the overall corona thickness parallel to the protein pore axis between all models gives 41 Å for the geometric elliptic toroid model, ≅ 42 Å for the all-atom MD model, and ≅ 45 Å for the hybrid model. As it will be discussed in the following, the higher value that is found for the hybrid model is most probably related to the absence of a hydration layer around the corona surface.

Further insight can be gained by calculation of the radial (relative to the pore axis) electron density profiles. Due to the special geometry of the protein/detergent complex, we have chosen to perform the radial density calculations only for atoms (or beads) that reside less than 5 Å apart from the transmembrane plane. In Figure 5 we plot the electron density

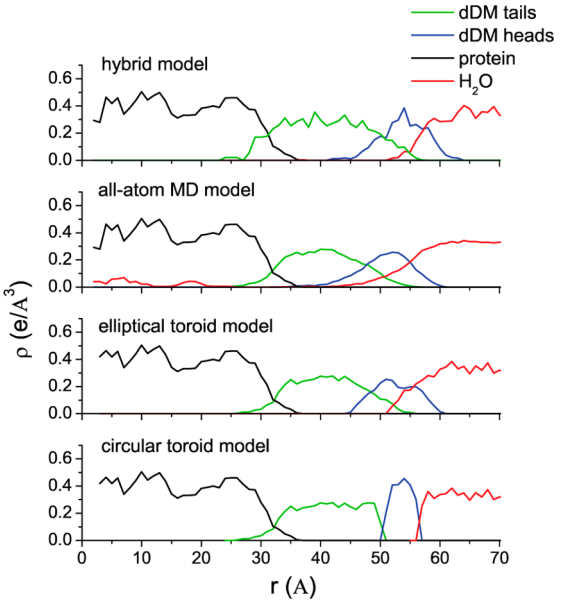


Figure 5. Hybrid, all-atom, elliptic, and circular geometrical model electron density. Radial electron density plots for each component of the system as a function of distance r from the Aquaporin-0 channel axis and for $|Z| < 5$ Å. The parameters for the two geometric models were the same as for the best-fit models in ref 11. The electron density of the solvent for the hybrid and geometrical model is evaluated by counting the empty lattice sites.

contributions from the different parts of the system. Since the studied structure does not possess a circular symmetry, the radial overlap between different parts of the corona can be mostly attributed to the squared shape of the detergent belt. The plots for the hybrid, all-atom and elliptic geometrical models give similar results, with similar overlaps between different parts of the complex and especially for detergent tails and heads atoms. This offers clear indications for the need of breaking the circular symmetry, either by the introduction of ellipticity in the geometrical model or by the more realistic current hybrid and all-atom models. Interestingly, the width of

dx.doi.org/10.1021/jp407688x | J. Phys. Chem. B 2013, 117, 13588−13594

the region where hydrophobic tails dominate is equal to about 16−18 Å, which is close to the theoretical maximum (16.7 Å) for the extension of an alkyl chain with 12 carbons.

Detergent Aggregation Number. The aggregation number in a protein/detergent complex is a quantity that should depend both on the nature of the detergent and also on the hydrophobic surface area of the protein. Experimentally, the number of detergent molecules that self-assemble around Aquaporin-0, was previously estimated by two independent methods, using either the ratio of the forward scattering to the protein concentration $I(0)/C$, or through concurrent refractometry measurements.[11] Both methods gave estimations in the range 200−250 dDM molecules. The fit of the SAXS data with the elliptic toroid model gives an estimation of 271 ± 30 molecules. Interestingly, the all-atom MD model that best described the SAXS data, has exactly the same number of dDM molecules as predicted by the geometric coarse-grained model. However, the hybrid model tends to give a higher estimation (340 ± 31), something that should be attributed to the fact that in the ab initio method the hydration layer around the detergent corona is not included, whereas it was included in the previous geometrical approach. This is the same effect as in the case of the program DAMMIN[8] where the absence of a hydration layer around the bead structure tends to swell the final configuration. This swelling concerns both parts of the detergent corona since the annealing algorithm has to keep a certain ratio between their thicknesses, so that the characteristic scattering peaks at $q > 0.1$ Å^{-1} are reproduced by the final fitting curve. It is found that by also excluding from the model the hydration layer around the protein surface, even larger deviations of the detergent aggregation number (+5%) are obtained.

Concomitantly, all structural parameters found for the hybrid model, tend to have slightly higher apparent values comparing to the geometric and all-atom models, something that is also reflected in the radial electron density plot (Figure 5), where a shift of the distributions toward higher radial distances is present. Note that the estimation of the number of dDM molecules for the hybrid model is performed by simply counting the number of electrons in the two parts of the corona independently and dividing them by the number of corresponding electrons of heads and tails in the actual dDM molecule. The agreement between the two values, estimated from the number of "hydrophobic" and "hydrophilic" beads is an indication of the validity of the model. In the present case, the hybrid method gave estimations that were the same within ±5%.

Another quantity of interest that is also related to structural aspects of the protein/detergent complex, is the mean surface area occupied by each molecule. By numerically estimating the outer surface of the detergent corona from our MD model, we find a surface per dDM molecule equal to 73 Å^2. Then by consideration of the aggregation number and geometry of free dDM micelles as found by fitting SAXS data with a core−shell ellipsoid of revolution model,[12] we may estimate the surface per detergent molecule to be equal to 79.5 Å^2. Comparison of these two values suggests that the organization of detergents around the transmembrane surface is characterized by about the same order as in the micellar environment. This is also supported by the calculation of the occupied volume per molecule for the micelle and protein/detergent complex. We found essentially identical volumes of 731 Å^3 and 727 Å^3 per molecule for our all-atom model and micelle structure, respectively.

CONCLUSIONS

Structural studies that require protein solubilzation by detergents should greatly benefit from a precise description of the detergent assembly. Here we have shown that the detergent corona of dDM molecules around α-helical Aquaporin-0 tetramers can be modeled to a certain detail based on the fitting of high-quality SE-HPLC coupled SAXS data by two different approaches.

Our ab initio coarse-grained approach with a minimal number of restrictions converged toward corona parameters that explain the artificial nature of the incorporation of ellipticity in the geometric model. The use of all-atom molecular dynamics provides a realistic description of interatomic interactions while accounting for the SAXS data, thus providing a convincing molecular description of the protein/detergent complex. The obtained structural parameters and detergent aggregation numbers agree to a large extent with the estimations of the elliptical toroid model.

It is important to note that the general structural features and parameters that were deduced from the analysis should apply at least for cases where a membrane protein is solubilized by dDM or similar nonionic detergent molecules, provided that detergent self-assembly takes place only around the transmembrane surface.[29] Based on the present approach, further work with different types of detergent molecules and proteins of various transmembrane part morphologies can be envisaged in order to derive general rules regarding detergent organization around membrane proteins.

The lack of a hydration layer around the corona in the hybrid model and the observed discrepancies in the comparison of experimental and MD generated scattering curves leave open room for further future improvements of both modeling approaches. The simulated annealing algorithm for the hybrid model could be adapted so that during the minimization of the score function the coverage of the hydrophilic corona surface by hydration beads can be taken into account. The MD model could include a more sophisticated initial equilibration procedure targeted at the preparation of less compact and more rounded initial structures. In any case, the hybrid and the MD models provide a more realistic description of the corona, which both validate a posteriori our previous geometrical approach[11] and offer new perspectives for model developments.

Applications taking advantage of the present modeling scheme may include the validation of candidate membrane protein structures from crystallography and homology modeling techniques, or the monitoring of conformational changes in solution associated with changes of the extra-membrane part of the protein. Such studies may provide unique information concerning the function of membrane channels. Finally, we note that as in the case of soluble proteins, SAXS from detergent solubilized membrane proteins will serve as a complementary technique to NMR and crystallography. We believe that the present study will further stimulate experimental and computational work toward this direction.

ASSOCIATED CONTENT

Supporting Information

Details on the implementation of the ab initio algorithm and MD simulations; Additional results concerning system equilibration and stability during MD simulations. This information is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*E-mail: javier.perez@synchrotron-soleil.fr.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The authors thank Dr. Dominique Durand, Dr. Konrad Hinsen, and Dr. Stéphane Abel for fruitful discussions and comments on this work, and Dr. Patrice Vachette for constant interest and support.

REFERENCES

(1) White, S. H. Biophysical Dissection of Membrane Proteins. *Nature* **2009**, *459*, 344−346.
(2) White, S. H.; Wimley, W. C. Membrane Protein Folding and Stability: Physical Principles. *Annu. Rev. Biophys. Biomol. Struct.* **1999**, *28*, 319−365.
(3) Tamm, L. K.; Liang, B. NMR of Membrane Proteins in Solution. *Prog. Nucl. Magn. Reson. Spectrosc.* **2006**, *48*, 201−210.
(4) Koch, M. H. J.; Vachette, P.; Svergun, D. I. Small-Angle Scattering: A View on the Properties, Structures and Structural Changes of Biological Macromolecules in Solution. *Q. Rev. Biophys.* **2003**, *36*, 147−227.
(5) Lipfert, J.; Doniach, S. Small-Angle X-ray Scattering from RNA, Proteins, and Protein Complexes. *Annu. Rev. Biophys. Biomol. Struct.* **2007**, *36*, 307−327.
(6) Pérez, J.; Nishino, Y. Advances in X-ray scattering: from solution SAXS to achievements with coherent beams. *Curr. Opin. Struct. Biol.* **2012**, *22*, 670−678.
(7) Chacon, P.; Morn, F.; DÍaz, J. F.; Pantos, E.; Andreu, J. M. Low-Resolution Structures of Proteins in Solution Retrieved from X-ray Scattering with a Genetic Algorithm. *Biophys. J.* **1998**, *74*, 2760−2775.
(8) Svergun, D. I. Restoring Low-Resolution Structure of Biological Macromolecules from Solution Scattering Using Simulated Annealing. *Biophys. J.* **1999**, *76*, 2879−2886.
(9) Svergun, D. I.; Petoukhov, M. V.; Koch, M. H. Determination of Domain Structure of Proteins from X-ray Solution Scattering. *Biophys. J.* **2001**, *80*, 2946−2953.
(10) Zimmer, J.; Doyle, D. A.; Grossmann, J. G. Structural Characterization and pH-Induced Conformational Transition of Full-Length KCSA. *Biophys.l J.* **2006**, *90*, 1752−1766.
(11) Berthaud, A.; Manzi, J.; Pérez, J.; Mangenot, S. Modeling Detergent Organization around Aquaporin-0 Using Small-Angle X-ray Scattering. *J. Am. Chem. Soc.* **2012**, *134*, 10080−10088.
(12) Lipfert, J.; Columbus, L.; Chu, V. B.; Lesley, S. A.; Doniach, S. Size and Shape of Detergent Micelles Determined by Small-Angle X-ray Scattering. *J. Phys. Chem. B* **2007**, *111*, 12427−12438.
(13) Svergun, D. I.; Nierhaus, K. H. A Map of Protein-RNA Distribution in the 70 s *Escherichia coli* Ribosome. *J. Biol. Chem.* **2000**, *275*, 14432−14439.
(14) Franke, D.; Svergun, D. I. DAMMIF, a Program for Rapid Ab-Initio Shape Determination in Small-Angle Scattering. *J. Appl. Crystallogr.* **2009**, *42*, 342−346.
(15) Gonen, T.; Cheng, Y.; Sliz, P.; Hiroaki, Y.; Fujiyoshi, Y.; Harrison, S. C.; Walz, T. Lipid-Protein Interactions in Double-Layered Two-Dimensional AQP0 Crystals. *Nature* **2005**, *438*, 633−638.
(16) Svergun, D. I.; Richard, S.; Koch, M. H. J.; Sayers, Z.; Kuprin, S.; Zaccai, G. Protein Hydration in Solution: Experimental Observation by X-ray and Neutron Scattering. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 2267−2272.
(17) Park, S.; Bardhan, J. P.; Roux, B.; Makowski, L. Simulated X-ray Scattering of Protein Solutions Using Explicit-Solvent Models. *J. Chem. Phys.* **2009**, *130*, 134114 1−8.
(18) Svergun, D.; Barberato, C.; Koch, M. H. J. CRYSOL − A Program to Evaluate X-ray Solution Scattering of Biological Macromolecules from Atomic Coordinates. *J. Appl. Crystallogr.* **1995**, *28*, 768−773.
(19) Debye, P. Zerstreuung von Röntgenstrahlen. *Ann. Phys.* **1915**, *351*, 809−823.
(20) Bond, P. J.; Sansom, M. S. Membrane Protein Dynamics versus Environment: Simulations of OMPA in a Micelle and in a Bilayer. *J. Mol. Biol.* **2003**, *329*, 1035−1053.
(21) Bond, P. J.; Cuthbertson, J. M.; Deol, S. S.; Sansom, M. S. P. MD Simulations of Spontaneous Membrane Protein/Detergent Micelle Formation. *J. Am. Chem. Soc.* **2004**, *126*, 15948−15949.
(22) Psachoulia, E.; Bond, P. J.; Sansom, M. S. P. MD Simulations of Mistic: Conformational Stability in Detergent Micelles and Water. *Biochemistry* **2006**, *45*, 9053−9058.
(23) Berendsen, H. J. C.; van der Spoel, D.; van Drunen, R. Gromacs: A Message-Passing Parallel Molecular Dynamics Implementation. *Comput. Phys. Commun.* **1995**, *91*, 43−56.
(24) Lindahl, E.; Hess, B.; van der Spoel, D. Gromacs 3.0: A Package for Molecular Simulation and Trajectory Analysis. *J. Mol. Model.* **2001**, *7*, 306−317.
(25) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. Gromacs 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J. Chem. Theory Comp.* **2008**, *4*, 435−447.
(26) Kozin, M. B.; Svergun, D. I. Automated Matching of High- and Low-Resolution Structural Models. *J. Appl. Crystallogr.* **2001**, *34*, 33−41.
(27) Abel, S.; Dupradeau, F. Y.; Raman, E. P.; MacKerell, A. D.; Marchi, M. Molecular Simulations of Dodecyl-β-Maltoside Micelles in Water: Influence of the Headgroup Conformation and Force Field Parameters. *J. Phys. Chem. B* **2011**, *115*, 487−499.
(28) Svergun, D. I. Determination of the Regularization Parameter in Indirect-Transform Methods Using Perceptual Criteria. *J. Appl. Crystallogr.* **1992**, *25*, 495−503.
(29) Andersen, K. K.; Westh, P.; Otzen, D. E. Global Study of Myoglobin Surfactant Interactions. *Langmuir* **2007**, *24*, 399−407.