Eur Biophys J (2008) 37:131–141
DOI 10.1007/s00249-007-0185-8

ORIGINAL PAPER

Outer membrane proteins: comparing X-ray and NMR structures by MD simulations in lipid bilayers

Katherine Cox · Peter J. Bond · Alessandro Grottesi · Marc Baaden · Mark S. P. Sansom

Received: 24 January 2006 / Revised: 7 May 2007 / Accepted: 8 May 2007 / Published online: 6 June 2007

Abstract The structures of three bacterial outer membrane proteins (OmpA, OmpX and PagP) have been determined by *both* X-ray diffraction and NMR. We have used multiple (7×15 ns) MD simulations to compare the conformational dynamics resulting from the X-ray versus the NMR structures, each protein being simulated in a lipid (DMPC) bilayer. Conformational drift was assessed via calculation of the root mean square deviation as a function of time. On this basis the 'quality' of the starting structure seems mainly to influence the simulation stability of the transmembrane β-barrel domain. Root mean square fluctuations were used to compare simulation mobility as a function of residue number. The resultant residue mobility profiles were qualitatively similar for the corresponding X-ray and NMR structure-based simulations. However, all three proteins were generally more mobile in the NMR-based than in the X-ray simulations. Principal components analysis was used to identify the dominant motions within each simulation. The first two eigenvectors (which account for >50% of the protein motion) reveal that such motions are concentrated in the extracellular loops and, in the case of PagP, in the N-terminal α-helix. Residue profiles of the magnitude of motions corresponding to the first two eigenvectors are similar for the corresponding X-ray and NMR simulations, but the directions of these motions correlate poorly reflecting incomplete sampling on a ~10 ns timescale.

Keywords Membrane protein · Outer membrane · Simulation · Molecular dynamics · Lipid bilayer

K. Cox · P. J. Bond · A. Grottesi · M. S. P. Sansom (✉)
Department of Biochemistry, University of Oxford,
South Parks Road, Oxford, OX1 3QU, UK
e-mail: mark.sansom@bioch.ox.ac.uk

M. Baaden
Laboratoire de Biochimie Théorique, CNRS UPR 9080,
Institut de Biologie Physico-Chimique,
13, rue Pierre et Marie Curie, 75005 Paris, France

Present Address:
A. Grottesi
CASPUR, Supercomputing Consortium for University
and Research, Via dei Tizii 6, 00185 Rome, Italy

Introduction

Bacterial outer membrane proteins (OMPs) share a common basic architecture, with the transmembrane (TM) domain formed by an anti-parallel β-barrel. OMPs constitute about half of the mass of the bacterial outer membrane. The structure of 26 OMPs have been solved by X-ray diffraction (Buchanan 1999; Koebnik et al. 2000; Schulz 2000), of which three have also been determined by NMR (Arora et al. 2001; Fernandez et al. 2001; Hwang et al. 2002). OMPs provide an excellent opportunity for using molecular dynamics (MD) simulation studies to explore the conformational dynamics of a family of structurally related membrane proteins. This should allow us to define common dynamic properties alongside functionally important differences between individual OMPs (Bond et al. 2002; Domene et al. 2003). Thus, starting from an average structure (in time and space) as determined by protein crystallography (PX) or NMR the dynamic behaviour of a single protein molecule can be simulated in a membrane (Fig. 1) (Baaden and Sansom 2004) or micellar (Böckmann and Caflisch 2005; Bond and Sansom 2003) environment in order to help explore the relationship between its structure and biological function.

The majority of OMP structures have been determined by PX. However, recent developments in NMR methods

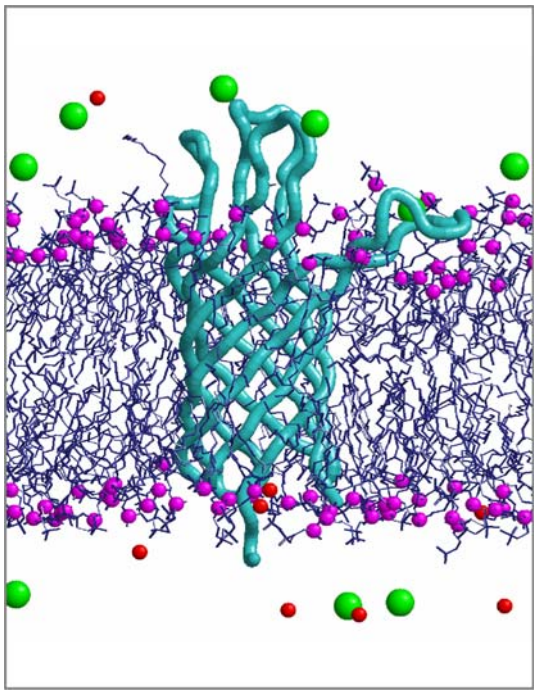

Fig. 1 Simulation setup for an OMP in a DMPC bilayer, illustrated for $OmpA^{NMR}$. The protein backbone is shown as a *tube* (*cyan*). The bilayer lipids are in *blue*, with the headgroup phosphorus atoms as *small purple spheres*. Na^+ ions are shown as *red spheres*; Cl^- ions as *green spheres*; water molecules are omitted for clarity

(Arora et al. 2001; Arora and Tamm 2001) have enabled the structure determination of three OMPs in detergent micelle environments. These are: OmpA (Arora et al. 2001); OmpX (Fernandez et al. 2001) (which has been refined with additional constraints derived from selective protonation of Val, Leu and Ile residues; Fernandez et al. 2004); and the OM enzyme PagP (the structure of which has been determined in two different detergents) (Hwang et al. 2002). The structures of these proteins have also been determined using PX (Ahn et al. 2004; Pautsch and Schulz 2000; Vogt and Schulz 1999).

In general structures solved by NMR are of lower accuracy and precision than high resolution PX structures (Spronk et al. 2002), although it remains difficult to define the 'quality' of an NMR structure (Snyder et al. 2005). Also, the experimental difficulties inherent in studying membrane proteins mean that NMR structures of the latter are generally based on fewer restraints per residue than for a water soluble protein of comparable size. A comparison of the behaviour of NMR and PX structures of water soluble proteins in MD simulations suggested that NMR-derived structures on average exhibit a higher degree of internal strain than PX structures (Fan and Mark 2003). Given the increasing importance of NMR for structure determination of membrane proteins, it is of interest to explore the influence of the origin of the starting structure (NMR versus PX) on the conformational dynamics of OMPs as seen in MD simulations in lipid bilayers. To this end, we have performed multiple 15 ns MD simulations starting from both PX and NMR structures of OmpA, OmpX and PagP. The results of these studies enable us to explore the extent to which the main features of MD simulations are robust to changes in the 'quality' of the starting model for the (membrane) protein.

Methods

Protein structures

Three OMPs for which both PX and NMR structures were available were used as the basis of the simulations: OmpA, OmpX and PagP (Table 1; Fig. 2). For OmpA the 2.5 Å resolution PX structure (1BXW) (Pautsch and Schulz 1998) was used as this had a complete coordinate set for the extracellular loops. With respect to the NMR structures, we manually selected one model from each ensemble, which was judged to be most compatible with a lipid bilayer environment. In particular, models were selected such that, e.g. loops were not in orientations that would bring them in contact with the hydrophobic core of the bilayer. For OmpX two simulations were performed: one using an earlier structure (1ORM), and the other using a more recent structure (1Q9F), which was derived from substantially more restraints per residue (see Table 1). Residues missing from the PX structure of PagP (i.e. residues 38–47 of loop L1) were built using Modeller v4 (http://www.salilab.org/modeller/) (Fiser et al. 2000; Sali and Blundell 1993).

The ionisation state of the sidechains in each structure was based on pK_A calculations, performed using the program WhatIf (Vriend 1990) combined with locally written code. The local code enabled us to take the complex environment of the ionisable sidechains and the low dielectric of the lipid bilayer slab into account in these calculations. In the case of OmpA and OmpX, the calculations were performed using the PX structures and the results were applied to both the NMR and PX starting models. For PagP (as the NMR structure was published first) the pK_A calculations were performed on the NMR structure and applied to the PX structure.

Setup of simulation systems

Each protein was embedded in a pre-equilibrated DMPC Bilayer (Fig. 1). To orient the protein relative to the hydrophobic core of the bilayer the tilt and translation of the protein relative to a 30 Å thick hydrophobic slab was optimised, by minimisation of the membrane–protein interaction energy as calculated using empirical hydrophobicities

Table 1 OMP structures used in simulations

Molecule	PDB[a]	X-ray: resolution (Å)	NMR: restraints per residue	Reference
$OmpA^{PX}$	1BXW	2.5		Pautsch and Schulz (1998)
$OmpA^{NMR}$	1G90::1		2.0	Arora et al. (2001)
$OmpX^{PX}$	1QJ8	1.9		Vogt and Schulz (1999)
$OmpX^{NMR\text{-}LOW}$	1ORM::10		1.6	Fernandez et al. (2001)
$OmpX^{NMR}$	1Q9F::18		5.5	Fernandez et al. (2004)
$PagP^{PX}$	1THQ	1.9		Ahn et al. (2004)
$PagP^{NMR}$	1MM5::10		3.5	Hwang et al. (2002)

[a] For the NMR structures the PDB code is followed by the model number in the NMR ensemble, i.e. the entry in the table is PDB::model

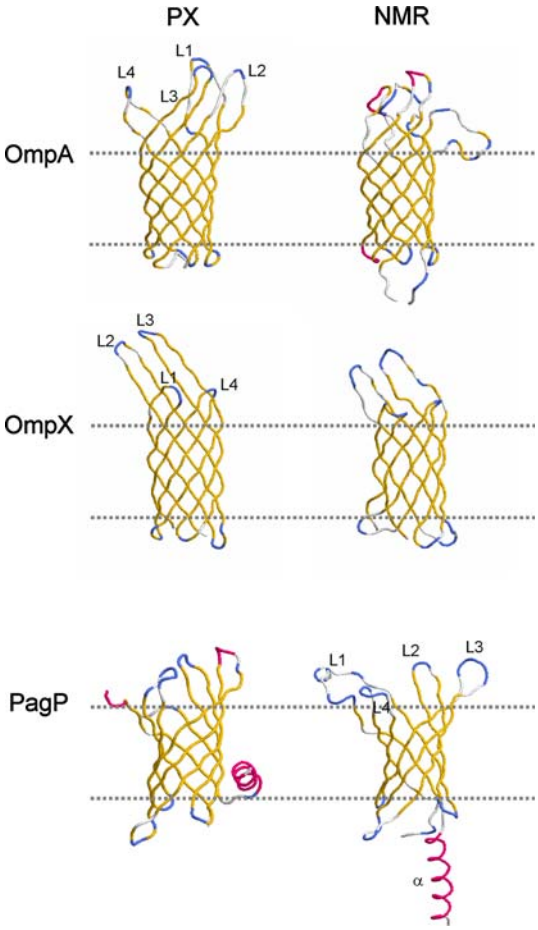

Fig. 2 Comparison of starting structures used in the simulations (also see Table 1) with the backbone shown as *tubes* coloured according to secondary structure (*yellow*, *β*-strand; *pink*, α-helix; *blue*, turn). The *horizontal broken lines* indicate the approximate location of the bilayer/water interfaces when these structures are docked in a DMPC bilayer

(Roseman 1988). Each protein was subsequently embedded in a DMPC bilayer as described in (Faraldo-Gómez et al. 2002). After insertion of the protein, water and ions equivalent to ~0.1 M NaCl were added by simple process of superimposition. The system was energy minimised and then underwent 0.5 ns of protein restrained dynamics to allow lipid relaxation. This was followed by 15 ns of production MD.

Simulation protocols

The $OmpA^{NMR}$, OmpX and PagP simulations were performed using GROMACS v3.14 (Lindahl et al. 2001) (see also www.gromacs.org); v2 was used for $OmpA^{PX}$ (Bond and Sansom 2003). The GROMOS96 force field (van Gunsteren et al. 1996) was used for all simulations. Energy minimisations used up to 500 steps of steepest descents. During protein restrained runs, harmonic position restraints with force constants of 10 kJ mol^{-1} $Å^{-2}$ per atom were applied to all non-hydrogen atoms. Electrostatic energies were calculated using particle mesh Ewald (Darden et al. 1993) with a 10 Å cut-off for the real space calculations. A cut-off of 10 Å was also used for van der Waals interactions. All the simulations were performed at constant temperature, pressure and number of particles. The temperatures of the lipid, solvent (i.e. water and ions) and proteins were coupled separately. The Berendsen temperature coupling algorithm (Berendsen et al. 1984) was applied at 310 K, i.e. well above the main phase transition temperature of DMPC $T_M = 297$ K, with coupling constant $\tau_T = 0.1$ ps, and the system pressure was isotropically coupled using a Berendsen barostat (Berendsen et al. 1981) at 1 bar with coupling constant $\tau_P = 1$ ps. The timestep for integration was 2 fs, and coordinates and velocities were saved every 5 ps. The LINCS algorithm was used throughout to restrain bond lengths (Hess et al. 1997). Trajectories have been deposited in the BioSimGRID database (www.biosimgrid.org).

Secondary structure analysis used DSSP (Kabsch and Sander 1983). VMD (Humphrey et al. 1996) and Rasmol (Sayle and Milner-White 1995) were used for visualisation. Convergence analysis (Bond and Sansom 2003; Faraldo-Gómez et al. 2004) was carried out on the final 10 ns of each simulation: average mean square fluctuation (MSF) values were calculated for time windows of 0.1, 0.25, 0.5, 1, 2, 4 and 8 ns. In each case the MSF values are evaluated separately for TM and non-TM residues. Principal component

analysis of the covariance matrix of the positional fluctuations of the Cα atoms was performed as described by (Amadei et al. 1993; Garcia 1992). This matrix was generated from the final 10 ns of each trajectory.

Results

Starting structures

Visual inspection of the starting structures (Fig. 2) reveals differences between the PX and NMR structures, mainly in the conformations of the extracellular loops (and of the N-terminal α-helix in the case of PagP), but also in some aspects of the local secondary structure of the TM barrels. These differences reflect the different experimental conditions and techniques used during structure determination. In particular, they are likely to reflect differences in the environment of the flexible loops experienced by the proteins in detergent micelles (for the NMR experiments) and in crystal lattices (for the X-ray studies) (see, e.g. Bond et al. 2006, for a simulation perspective of OmpA loop mobility within a crystal). The starting structures may also be compared by calculation of the root mean squared deviations (RMSDs) of the equivalent Cα atoms in the PX and NMR structures (Table 2). It can be seen that in each case the RMSDs are significantly higher for the residues outside the TM region (i.e. the loops and turns) than for the β-barrel residues within the TM region. Comparing the TM residue RMSDs between structures, they are higher for OmpA (a lower resolution PX structure), and for the comparison of $OmpX^{NMR\text{-}LOW}$ (which has a small number of restraints per residue) with $OmpX^{PX}$. More generally, if the structures are ordered according to a common 'quality'/resolution scale, one obtains starting from the highest quality structure: $OmpX^{PX} \approx PagP^{PX} > OmpA^{PX} > OmpX^{NMR} > PagP^{NMR} > OmpA^{NMR} > OmpX^{NMR\text{-}LOW}$. The higher quality structures have a lower RMSD between their TM residues (OmpX and PagP), whereas the lower quality structures yield a higher RMSD for those residues (OmpA and $OmpX^{NMR\text{-}LOW}$). However, this does not remain true for the non-TM residues where no direct correlation with the quality of the starting structure can be seen.

Each structure was energy minimised prior to the simulations. Comparing the potential energies (per residue) of the resultant minimised structures (Table 3) does not reveal any correlation between the origin of the structure (PX versus NMR) and the potential energy of the protein following minimisation. Note that the potential energies per residue cited are therefore after minimisation to a *local* energy minimum nearest to the published structure. Thus, the analysis demonstrated the absence of a simple correlation between the experimental method used and the corresponding force-field energy for the (locally) minimised structure. It is interesting to note that NMR structures can be of slightly lower energy, even though they are derived from relatively few restraints per residue.

Table 2 RMSDs between NMR and PX structures

Molecule	RMSD[a] (Å)		
	All residues	TM residues	Non-TM residues
OmpA	7.4	3.7	10.0
$OmpX^{NMR\text{-}LOW}$	5.1	3.5	6.2
OmpX	3.2	1.9	4.5
PagP	12.0	1.6	14.8[b]

[a] Cα RMSD between NMR starting structure used in simulation and PX starting structure, fitted onto the PX structure

[b] Contains the L1 loop, nine missing residues of which were built via homology modelling

Table 3 Summary of simulations

Simulation	Potential energy of protein[a] (kJ mol^{-1})	Potential energy per residue[a] (kJ mol^{-1})	Final[b] Cα RMSD for TM residues (Å)	Final[b] Cα RMSD for non-TM residues (Å)
$OmpA^{PX}$	-0.52×10^4	−30.4	1.34 ± 0.04	3.57 ± 0.25
$OmpA^{NMR}$	-1.36×10^4	−79.5	3.08 ± 0.86	8.00 ± 0.44
$OmpX^{PX}$	-1.28×10^4	−86.5	0.87 ± 0.53	4.40 ± 0.26
$OmpX^{NMR\text{-}LOW}$	-0.51×10^4	−34.5	2.90 ± 0.07	4.38 ± 0.18
$OmpX^{NMR}$	-0.94×10^4	−64.0	2.00 ± 0.80	4.90 ± 0.25
$PagP^{PX}$	-1.02×10^4	−62.6	1.21 ± 0.07	4.24 ± 0.24
$PagP^{NMR}$	-0.83×10^4	−51.0	1.32 ± 0.05	7.51 ± 0.27

Each simulation was of 15 ns duration. The number of lipid (DMPC) molecules varied between 105 and 113 molecules. The number of waters ranged from ~4,800 to ~5,600, and the total numbers of atoms from ~21,000 to ~24,000

[a] Calculated for the local minimum closest to the experimental starting structure after 100 steps of energy minimisation of the protein

[b] Calculated for the last 5 ns of each simulation

Structural drift

The overall drift of a membrane protein structure from its initial conformation during a simulation provides a crude measure of the 'stability' of a structure in its membrane environment. Structural drifts for each simulation can be quantified by calculating the RMSD with respect to time of the Cα atoms from the starting structure. For this purpose we divided the structure into TM, and non-TM regions (Fig. 3; Table 3).

For each simulation the RMSDs are, as expected, lower for the TM region than for the non-TM region. Comparing the two OmpA simulations, the RMSDs are significantly higher when starting from the NMR structure, both for the TM (i.e. β-barrel) and non-TM (i.e. loop and turn) regions. In both OmpA simulations the main initial rise in RMSD is complete after ~9 ns.

For the OmpX simulations, as the 'quality' of the structure decreases the RMSD of the TM region increases, i.e. $OmpX^{NMR\text{-}LOW} > OmpX^{NMR} > OmpX^{PX}$. For the non-TM regions the equivalent RMSDs are all in the range 4–5 Å. Thus it would appear that the 'quality' of the structure mainly influences the simulation stability of the TM β-barrel. In both $OmpX^{PX}$ and $OmpX^{NMR}$ simulations, the initially rising RMSD reaches a plateau at ~2 ns. However for the $OmpX^{NMR\text{-}LOW}$ simulation (data not shown), the TM region RMSD levels off but the non-TM RMSD continues to drift.

For all of the simulations, Procheck (Laskowski et al. 1993) was used to assess the stereochemical quality of the protein structure before and after the simulation. For the PX structure-based simulations, there was some drift of β-barrel residues from the most favoured region to the additionally allowed regions. For the simulations based on NMR structures, the corresponding drift was less, possibly as a result of a poorer stereochemical quality of the starting structures. This may help to explain the higher TM region RMSD for the NMR-based simulations, as it has been shown that there is a correlation between backbone stereochemical quality and final RMSD in membrane protein simulations (Law et al. 2005).

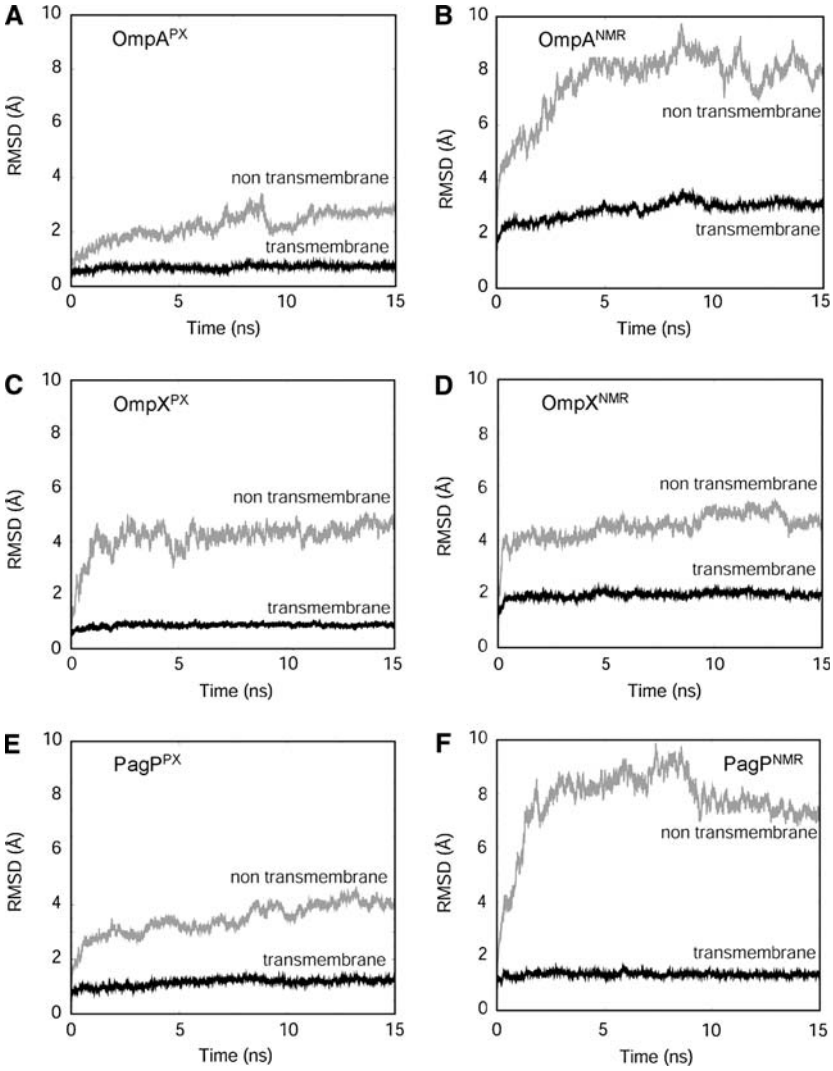

Fig. 3 Root mean square deviation (*RMSD*) of the Cα atoms versus time for the **a** $OmpA^{PX}$; **b** $OmpA^{NMR}$; **C** $OmpX^{PX}$; **d** $OmpX^{NMR}$, **e** $PagP^{PX}$ and **f** $PagP^{NMR}$ simulations. In each case the structures from the simulation were fitted to the initial structure relative to which the RMSD was evaluated. The *black* and *grey* lines give the RMSDs for the transmembrane and non-transmembrane residues, respectively

For all three proteins, the non-TM RMSDs are high, even for the PX structures. Such elevated RMSDs may reflect the release of restraints formed by crystal contacts. For example, the large extracellular 'flag' of $OmpX^{PX}$ forms a major crystal contact between the L3 loops of adjacent monomers, and in $PagP^{PX}$ the L4 loop is stabilised by a contact to turn T2. When comparing simulations of globular proteins starting from both NMR and PX structures (Fan and Mark 2003) it was noted that for those PX structures that gave a core backbone RMSD >3 Å this was mainly due to the release of crystal packing forces.

Cα RMSD matrices were calculated to enable a pairwise comparison of each structure within a PX trajectory with each structure within the corresponding NMR trajectory. Such analysis (data not shown) did not provide any support for convergence of the PX and NMR structures over the course of the simulations. This may reflect the incomplete sampling of membrane protein conformational dynamics on a ~10 ns timescale (Faraldo-Gómez et al. 2004). To check this, the RMSD of structures along each NMR simulation was calculated against the final structure from the corresponding PX simulation. Again, these results did not suggest convergence of the NMR and PX simulations.

Convergence analysis

A more quantitative comparison of relative protein mobility in the different simulations, and for the TM versus non-TM regions of each protein, was carried out by means of block analysis of the MSFs of Cα atoms about the structure at the beginning of the time frame to be analysed. For the last 10 ns of each simulation, MSF averages were calculated for time windows of 0.1, 0.25, 0.5, 1.0, 2.0, 4.0 and 8.0 ns for both the TM and non-TM regions of the protein and plotted on log–log scales (Fig. 4).

From the non-zero slopes of the lines we conclude that the MSFs have not converged in any of the simulations. This is a general property of the current timescales (tens of nanoseconds) accessible by MD simulations of membrane protein systems (Faraldo-Gómez et al. 2004). However, these graphs reveal that *trends* in relative magnitudes of MSFs are preserved over the analysis time windows, suggesting that extrapolation to longer timescales associated with biological function is not unreasonable. Hence our estimates of the relative magnitudes of fluctuations of, e.g. TM and non-TM regions are probably correct, even if the absolute values are underestimated due to failure to fully capture longer timescale motions. In all cases, at ~10 ns the non-TM regions show on average ~4× higher conformational fluctuations (in terms of Cα MSF) than do the TM regions.

Comparing the MSFs between PX and NMR structure simulations reveals a more complex pattern. For the OmpA TM regions the PX and NMR simulation MSFs are approximately equal. However, for the OmpX and PagP simulations it can be seen that the MSFs are consistently greater

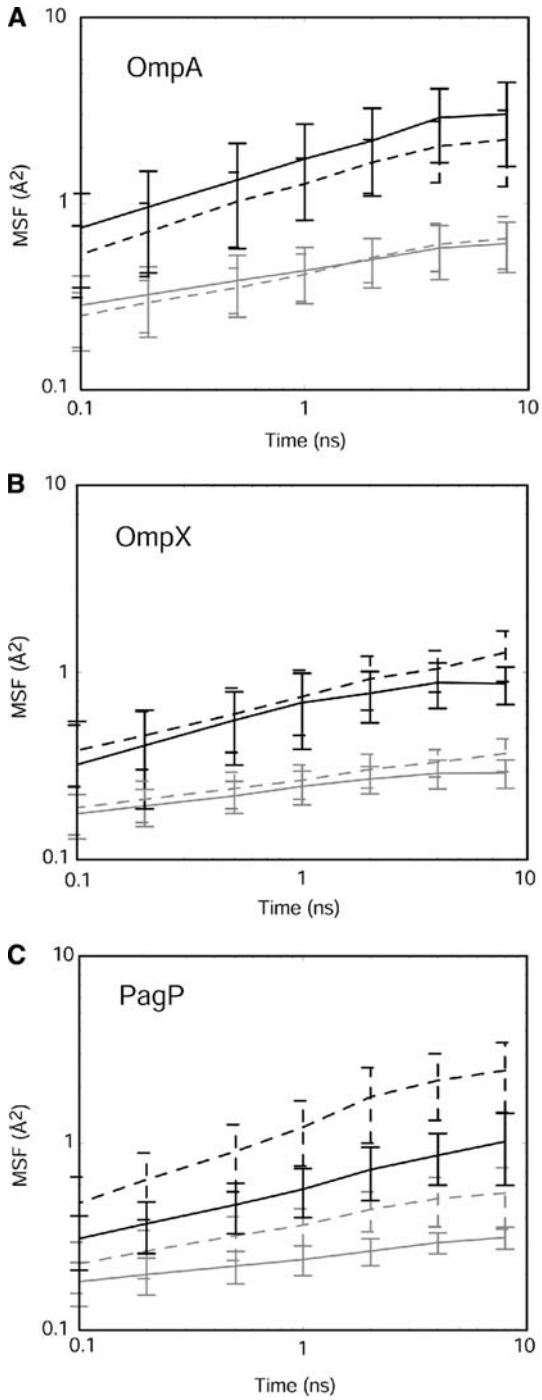

Fig. 4 Block analysis of mean square fluctuations (*MSFs*) calculated for Cα atoms for the **a** OmpA, **b** OmpX and **c** PagP simulations. The *grey lines* show the MSFs for TM structure elements, the *black lines* the MSFs for non-TM structure elements. The *solid lines* represent the PX structures and the *broken lines* the NMR structures. MSFs are also averaged over all windows of the same length; *error bars* represent standard deviations about the window-average

for NMR simulations than for the PX simulations. Focusing on the non-TM regions, OmpA has a higher MSF for the PX than NMR simulation whereas for OmpX and PagP the reverse is true. This may reflect the more poorly defined conformations of the extracellular loops in the OmpA PX structure, as has been suggested previously (Faraldo-Gómez et al. 2004). Indeed, coordinates were missing for several loop residues from the higher resolution structure of OmpA (1QJP; Pautsch and Schulz 2000).

Fluctuations versus residue

To further characterise the pattern of flexibility versus rigidity within the proteins, and to compare the PX versus NMR simulations, we have measured the Cα root mean square fluctuation (RMSF) as a function of residue number for each simulation (Fig. 5). Comparing NMR- and PX-based simulations, it is evident that all three proteins are generally more mobile in the NMR-based simulations than in the PX simulations. Comparing the OmpX NMR simulations it can also be seen that as the quality of the structure (i.e. number of restraints per residue) improves the flexibility of the protein in the membrane decreases (data not shown). The only exceptions to this are loop L1 of OmpA and loop L3 of OmpX (part of the loop L2 and L3 'flag'—see below), where the PX-based simulations yield higher mobilities. The third periplasmic turn (*ca.* residue 119) of OmpX also shows an elevated mobility in the NMR-based simulation (possibly resulting from correlations between the motions of turn 3 and the L2/L3 flag). In the case of PagP there are considerable differences between the PX and NMR simulations as far as the RMSF profiles are concerned. In both simulations, the RMSF is high for the N-terminal α-helix, but particularly so in the case of the $PagP^{NMR}$ simulation. This difference may reflect the initial orientation of the α-helix. (In the PX structure this is close to the preferred location, i.e. at the membrane/water interface, whereas in the NMR structure used in the simulation the initial orientation is directed away from the membrane.) Furthermore, for loops L3 and L4 the fluctuations are considerably greater for $PagP^{NMR}$ than for $PagP^{PX}$. This may reflect the rather more 'closed' nature of the loops in the PX structure. Interestingly, closure of the loops may play an important role in forming the active site of PagP (K. Cox and M. S. P. Sansom, ms. in preparation).

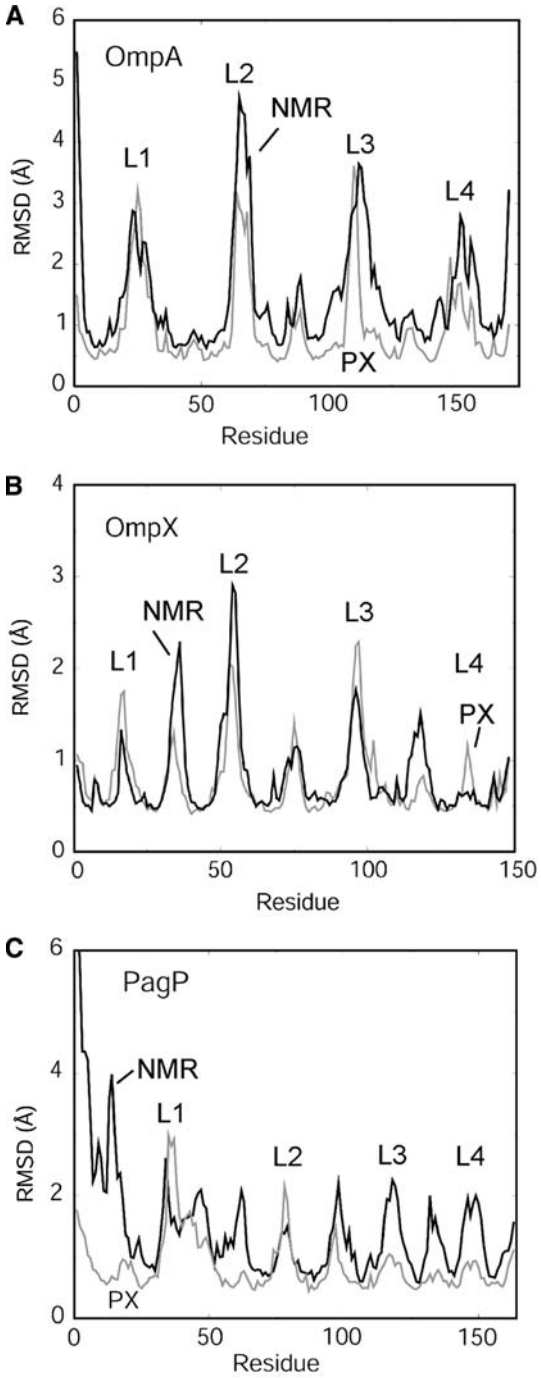

Fig. 5 Cα atom root mean square fluctuations (*RMSFs*) versus residue number. RMSFs were calculated from the final 10 ns of each simulation: **a** OmpA; **b** OmpX and **c** PagP. The *grey lines* correspond to simulations based on the NMR structures; the *black lines* to simulations based on the PX structures

Secondary structure

For each simulation the secondary structure versus time has been examined (results not shown) and used to calculate the predominant secondary structure (i.e. the secondary structure elements present for at least 75% of each simulation; Fig. 6). For OmpA the secondary structure elements did not fluctuate significantly with respect to time. From the predominant secondary structure analysis it is clear that the only major differences are at the ends of the β-strands. Thus, on average the length of strands of $OmpA^{NMR}$ is shorter for β3, β4 and β6 than in $OmpA^{PX}$. For OmpX, a

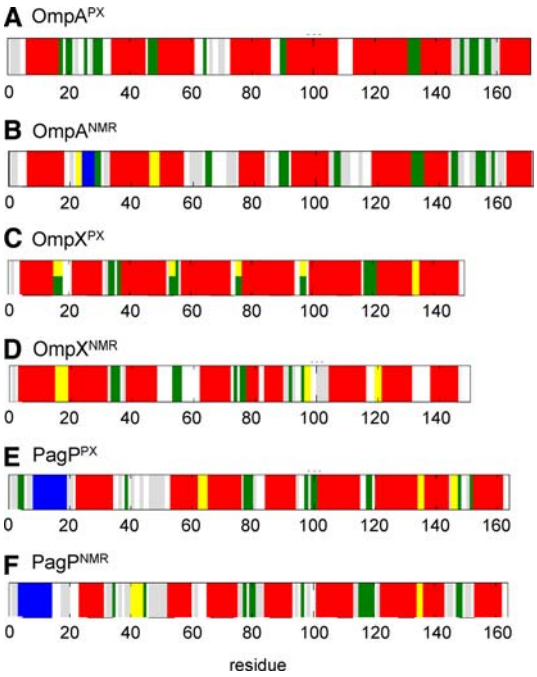

Fig. 6 Dominant secondary structure (i.e. secondary structure, as determined using DSSP, maintained for >75% of the simulation, calculated over the whole duration of each 15 ns simulation) for: **a** $OmpA^{PX}$; **b** $OmpA^{NMR}$; **c** $OmpX^{PX}$; **d** $OmpX^{NMR}$, **e** $PagP^{PX}$ and **f** $PagP^{NMR}$. *Red*, β-sheet; *grey*, coil; *green*, bend; *yellow*, turn; and *blue*, α-helix

similar pattern is seen, with shorter strands for, e.g. β3, β4, β5 and β6 in the $OmpX^{NMR}$ simulation. For PagP the difference between the two simulations is less pronounced. Interestingly, given the difference in loop mobility noted above, the secondary structure of L4 is less defined in the $PagP^{NMR}$ than in the $PagP^{PX}$ simulation.

Principal components analysis

Principal components analysis has been employed to identify the dominant motions in each simulation. The eigenvalue profiles for the seven simulations (Table 4) reveal that

Table 4 Percentage of motions represented by the eigenvectors

Simulation	Percentage motion of eigenvector		
	Eigenvectors 1 and 2	Eigenvectors 1–5	Eigenvectors 1–10
$OmpA^{PX}$	55	71	80
$OmpA^{NMR}$	55	75	83
$OmpX^{PX}$	37	54	66
$OmpX^{NMR\text{-}LOW}$	50	66	76
$OmpX^{NMR}$	52	69	78
$PagP^{PX}$	50	61	72
$PagP^{NMR}$	65	78	85

in all cases except one the first two eigenvectors captured ≥50% of the motion sampled during the simulation. However, when the subspace overlap of the eigenvectors was calculated for the covariance matrices from corresponding PX and NMR simulations (data not shown), for all the pairwise covariance matrix comparisons a low dot product was observed, indicating that the essential motions of the two systems were different. This is presumed to reflect the relative lack of convergence of the simulations (see above).

We have compared visually the nature of the first two principal motions (eigenvectors) for each structure and simulation (Fig. 7). For both OmpA and OmpX these motions are concentrated in the extracellular loops. The first two principal components for both OmpA simulations show the presence of a concerted motion between loops L1, L2 and L3. From analysis of the displacement of the Cα carbons of the OmpA simulations along the first three eigenvectors, it can be seen that the amplitude of the fluctuations is similar (data not shown).

For OmpX, the principal motions are also concentrated in the loops, and especially in the L2/L3 flag. However, there are significant differences between the PX- and NMR-based simulations, as can be seen from Fig. 7. In particular, for eigenvectors 1 and 3 of $OmpX^{PX}$ the loop motion dominates, but for eigenvector 2 there is a correlated motion between the flag loops (L2 and L3) and the first β-hairpin (i.e. β1–L1–β2). These fluctuations are amplified in the $OmpX^{NMR}$ simulations.

For the PagP simulations, comparison of the principal motions revealed a more complex behaviour. For example, a concerted motion can be detected between the N-terminal α-helix and the L1 extracellular loop in the $PagP^{NMR}$ simulation, whereas the same motion is virtually absent from the $PagP^{PX}$ simulation. These differences in sampled motions presumably reflect the rather different orientations of the α-helix in the PX and NMR structures.

In general, there appears to be a correlation in the relative magnitude of the principal motions between the corresponding PX and NMR based simulations. However, a more detailed residue-by-residue comparison of the angles between the displacements corresponding to the first eigenvectors (data not shown) revealed somewhat imperfect correlations, suggesting that the two simulations have sampled different regions of subspace. Again, this probably results from incomplete sampling during 15 ns simulations.

Pores

A number of measurements of ionic currents using OmpA and OmpX reconstituted into planar lipid bilayers have suggested that these proteins are capable of forming ion permeable pores (Dupont et al. 2004). (There have not been any suggestions of pore formation by PagP.) Previous simulation

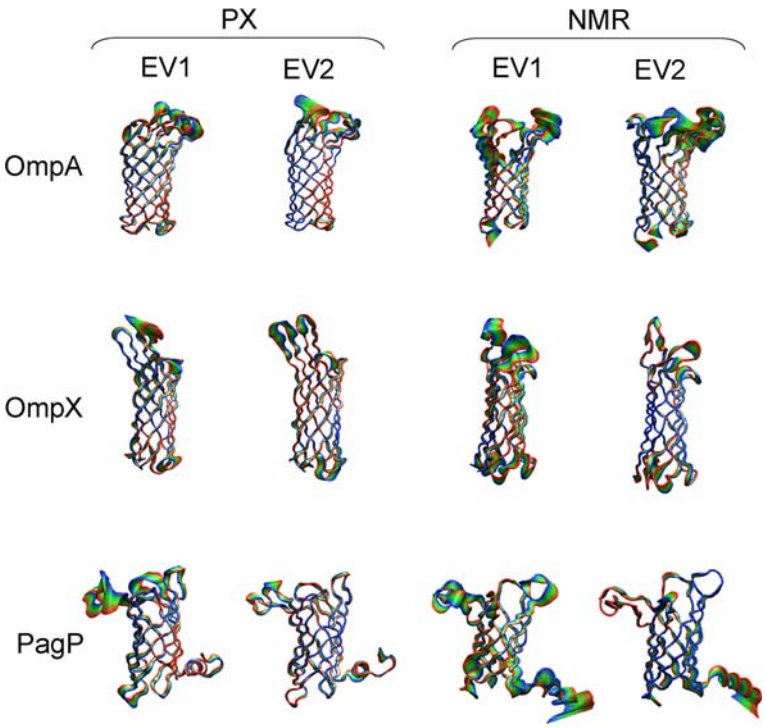

Fig. 7 Superimposed Cα traces corresponding to projections along the first (EV1) or second (EV2) eigenvector from each simulation. Twenty Cα trace structures are superimposed and coloured from *red* to *blue*

studies of OmpA, in both lipid bilayer and detergent micelle environments (Bond et al. 2002; Bond and Sansom 2003) suggest that small changes in the packing of sidechains within the β-barrel core could result in opening of a transbilayer pore. We therefore examined whether OmpX showed any evidence of pore formation. None of the simulations of OmpX revealed any pore activity. For the $OmpX^{PX}$ simulation analysis of the trajectories of selected water molecules present within the TM barrel revealed three possible constrictions preventing the complete penetration of water. Interestingly, one of these constrictions (Fig. 8) is formed by two amino acids, F90 and S130, which were found to have two alternative conformations in the crystal structure. This suggested that fluctuations in sidechain conformation might be capable of 'gating' an OmpX pore. However, no such pore opening event was observed in the current OmpX simulations. We analysed the water occupancy in the OmpX pore over the last 5 ns of the simulation. The ten most solvent exposed pore-lining residues are E128, T4, T66, S42, D124, R29, Q114, T44, K27 and Y62, listed in decreasing order. Thus, there are at least some polar residues with potential for pore-lining seen within the simulations. However, our results suggest that longer simulations might be needed to fully address biologically relevant questions of pore formation by OmpX.

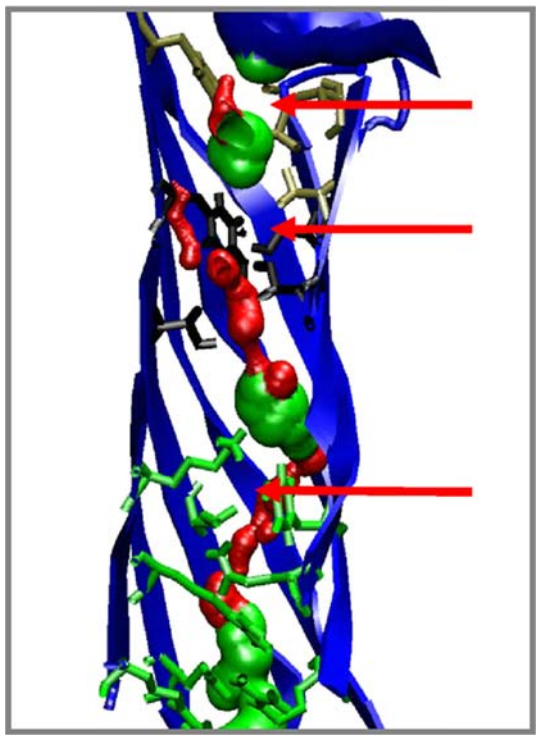

Fig. 8 Snapshot from the $OmpX^{PX}$ simulation, showing the inner surface of the OmpX 'pore' as calculated using HOLE (Smart et al. 1993, 1996). It can be seen that the 'pore' is constricted (*red surface*) in several places as indicated by the *horizontal arrows*

Discussion

These studies suggest that there are differences in the conformational dynamics between OMP simulations based on PX versus on NMR structures. At present it is difficult to dissect out those differences which are a consequence of the various environments (crystal versus detergent micelle) used in the structure determinations (Bond and Sansom 2003), those differences which arise from the variable quality of the starting structures, and thus differences due to incomplete sampling on a ~10 ns timescale. However, common patterns between NMR and PX structure simulations *do* emerge, providing a strong indication of those simulation features that are robust to changes in starting structure.

It is important to recognise that simulations based on NMR structures of OMPs do yield meaningful data, especially in the context of the emergence of further membrane protein structures based on NMR techniques upon which simulation studies may be based. Whilst we note that a recent simulation study of water soluble proteins has suggested that a significant number of NMR structures may be less stable than their crystallographic counterparts (Fan and Mark 2003), for the OMPs any such differences in stability of the TM regions (i.e. β-barrels) are relatively low. It is also important to note that NMR methods for membrane protein structure are advancing (Arora and Tamm 2001; Fernández and Wüthrich 2003) and as higher quality NMR structures become available (Fernandez et al. 2004) the reliability of subsequent simulations also improves. However, the number of restraints per residue is still low for membrane protein structures compared to their water soluble counterparts. For example, even the 'best' NMR structure of an OMP (OmpX and 1Q9F) has only ~6 restraints per residue. In contrast, for many water soluble proteins NMR structures with $\geq$20 restraints per residue (comparable to a 2.0–2.5 Å resolution PX structure) are available.

In terms of mobility, the overall patterns observed are similar to those seen in previous simulations (Baaden et al. 2003; Baaden and Sansom 2004; Bond and Sansom 2003; Faraldo-Gómez et al. 2003; Tieleman and Berendsen 1998), i.e. extracellular loop residues are more mobile than TM β-barrel residues. In general the mobility is a little higher for NMR than for PX structures. Of course, one must remember that the current simulations are in a DMPC membrane, i.e. under in vitro conditions. In vivo the outer leaflet of the OM is composed of lipopolysaccharide, the conformational dynamics of which (Lins and Straatsma 2001; Shroll and Straatsma 2002, 2003) are likely to modulate the behaviour of the protein.

To this extent, we can see that PX and NMR structures yield similar information on OMP flexibility. However, at the level of more detailed analysis of the conformational spaces occupied by the principal motions, there is limited overlap between the two classes of simulation. This suggests an upper bound on the level of interpretation of simulations of OMPs of duration 10–20 ns. To address more generic aspects of pore *function* and of protein–lipid interactions for these systems, one probably needs to extend the simulations by an order of magnitude.

These simulations have probed some similarities and some differences of the conformational dynamics based on NMR versus PX structures of OMPs. A number of other simulations have explored the conformational dynamics of OMPs in bilayer, detergent micelle (Böckmann and Caflisch 2005; Bond et al. 2004; Bond and Sansom 2003) and crystal (P. J. Bond and M. S. P. Sansom, ms. in preparation) environments. There remains a need to extend such studies to a wider range of membrane proteins in order to more fully understand the influence of both environment and structure quality on the dynamics of membrane proteins, and in order to relate differences in conformation and dynamics (e.g. in PagP; Hwang et al. 2004) to membrane protein function. In particular, the structures of a number of simpler α-helical membrane proteins have now been determined by NMR, including the glycophorin A TM helix dimer (MacKenzie et al. 1997), Mistic (Roosild et al. 2005) and MerF (Howell et al. 2005). Thus an understanding of the influence of the quality of α-helical membrane protein structures in simulations should also soon be possible.

Acknowledgements KC was a BBSRC research student. Research in MSPS's laboratory is funded by the BBSRC, EPSRC and the Wellcome Trust. Our thanks to all of our colleagues, especially Sundeep Deol and Charles Blundell.

References

Ahn VE, Lo EI, Engel CK, Chen L, Hwang PM, Kay LE, Bishop RE, Privé GG (2004) A hydrocarbon ruler measures palmitate in the enzymatic acylation of endotoxin. EMBO J 23:2931–2941

Amadei A, Linssen ABM, Berendsen HJC (1993) Essential dynamics of proteins. Proteins: Struc Funct Genet 17:412–425

Arora A, Abildgaard F, Bushweller JH, Tamm LK (2001) Structure of outer membrane protein A transmembrane domain by NMR spectroscopy. Nature Struct Biol 8:334–338

Arora A, Tamm LK (2001) Biophysical approaches to membrane protein structure determination. Curr Opin Struct Biol 11:540–547

Baaden M, Meier C, Sansom MSP (2003) A molecular dynamics investigation of mono- and dimeric states of the outer membrane enzyme OMPLA. J Mol Biol 331:177–189

Baaden M, Sansom MSP (2004) OmpT: molecular dynamics simulations of an outer membrane enzyme. Biophys J 87:2942–2953

Berendsen HJC, Postma JPM, van Gunsteren WF, DiNola A, Haak JR (1984) Molecular dynamics with coupling to an external bath. J Chem Phys 81:3684–3690

Berendsen HJC, Postma JPM, van Gunsteren WF, Hermans J (1981) Intermolecular forces. Reidel, Dordrecht

Böckmann RA, Caflisch A (2005) Spontaneous formation of detergent micelles around the outer membrane protein OmpX. Biophys J 86:3191–3204

Bond PJ, Cuthbertson JM, Deol SD, Sansom MSP (2004) MD simulations of spontaneous membrane protein/detergent micelle formation. J Am Chem Soc 126:15948–15949

Bond PJ, Faraldo-Gómez JD, Deol SS, Sansom MSP (2006) Membrane protein dynamics and detergent interactions within a crystal: a simulation study of OmpA. Proc Natl Acad Sci USA 103:9518–9523

Bond PJ, Faraldo-Gómez JD, Sansom MSP (2002) OmpA—a pore or not a pore? Simulation and modelling studies. Biophys J 83:763–775

Bond PJ, Sansom MSP (2003) Membrane protein dynamics vs. environment: simulations of OmpA in a micelle and in a bilayer. J Mol Biol 329:1035–1053

Buchanan SK (1999) b-Barrel proteins from bacterial outer membranes: structure, function and refolding. Curr Opin Struct Biol 9:455–461

Darden T, York D, Pedersen L (1993) Particle mesh Ewald—an N.log(N) method for Ewald sums in large systems. J Chem Phys 98:10089–10092

Domene C, Bond P, Sansom MSP (2003) Membrane protein simulation: ion channels and bacterial outer membrane proteins. Adv Prot Chem 66:159–193

Dupont M, Dé E, Chollet R, Chevalier J, Pagès JM (2004) *Enterobacter aerogenes* OmpX, a cation-selective channel *mar*- and osmoregulated. FEBS Lett 569:27–30

Fan H, Mark AE (2003) Relative stability of protein structures determined by X-ray crystallography or NMR spectroscopy: a molecular dynamics simulation study. Proteins: Struct Funct Bioinf 53:111–120

Faraldo-Gómez JD, Forrest LR, Baaden M, Bond PJ, Domene C, Patargias G, Cuthbertson J, Sansom MSP (2004) Conformational sampling and dynamics of membrane proteins from 10-nanosecond computer simulations. Proteins: Struct Funct Bioinf 57:783–791

Faraldo-Gómez JD, Smith GR, Sansom MSP (2002) Setup and optimisation of membrane protein simulations. Eur Biophys J 31:217–227

Faraldo-Gómez JD, Smith GR, Sansom MSP (2003) Molecular dynamics simulations of the bacterial outer membrane protein FhuA: a comparative study of the ferrichrome-free and bound states. Biophys J 85:1–15

Fernandez C, Hilty C, Bonjour S, Adeishvili K, Pervushin K, Wüthrich K (2001) Solution NMR studies of the integral membrane proteins OmpX and OmpA from *Escherichia coli*. FEBS Lett 504:173–178

Fernandez C, Hilty C, Wider G, Güntert P, Wüthrich K (2004) NMR structure of the integral membrane protein OmpX. J Mol Biol 336:1211–1221

Fernández C, Wüthrich K (2003) NMR solution structure determination of membrane proteins reconstituted in detergent micelles. FEBS Lett 555:144–150

Fiser A, Kinh Gian Do R, Sali A (2000) Modeling of loops in protein structures. Prot Sci 9:1753–1773

Garcia AE (1992) Large-amplitude nonlinear motions in proteins. Phys Rev Lett 68:2696–2699

Hess B, Bekker H, Berendsen HJC, Fraaije JGEM (1997) LINCS: a linear constraint solver for molecular simulations. J Comp Chem 18:1463–1472

Howell SC, Mesleh MF, Opella SJ (2005) NMR structure determination of a membrane protein with two transmembrane helices in micelles: MerF of the bacterial mercury detoxification system. Biochemistry 44:5196–5206

Humphrey W, Dalke A, Schulten K (1996) VMD—visual molecular dynamics. J Molec Graph 14:33–38

Hwang PM, Bishop RE, Kay LE (2004) The integral membrane enzyme PagP alternates between two dynamically distinct states. Proc Natl Acad Sci USA 101:9618–9623

Hwang PM, Choy WY, Lo EI, Chen L, Forman-Kay JD, Raetz CRH, Privé GG, Bishop RE, Kay LE (2002) Solution structure and dynamics of the outer membrane enzyme PagP by NMR. Proc Natl Acad Sci USA 99:13560–13565

Kabsch W, Sander C (1983) Dictionary of protein secondary structure: pattern-recognition of hydrogen-bonded and geometrical features. Biopolymers 22:2577–2637

Koebnik R, Locher KP, Van Gelder P (2000) Structure and function of bacterial outer membrane proteins: barrels in a nutshell. Mol Microbiol 37:239–253

Laskowski RA, Macarthur MW, Moss DS, Thornton JM (1993) Procheck—a program to check the stereochemical quality of protein structures. J Appl Cryst 26:283–291

Law RJ, Capener C, Baaden M, Bond PJ, Campbell J, Patargias G, Arinaminpathy Y, Sansom MSP (2005) Membrane protein structure quality in molecular dynamics simulation. J Mol Graph Mod 24:157–165

Lindahl E, Hess B, van der Spoel D (2001) GROMACS 3.0: a package for molecular simulation and trajectory analysis. J Molec Model 7:306–317

Lins RD, Straatsma TP (2001) Computer simulation of the rough lipopolysaccharide membrane of *Pseudomonas aeruginosa*. Biophys J 81:1037–1046

MacKenzie KR, Prestegard JH, Engelman DM (1997) A transmembrane helix dimer: structure and implications. Science 276:131–133

Pautsch A, Schulz GE (1998) Structure of the outer membrane protein A transmembrane domain. Nature Struct Biol 5:1013–1017

Pautsch A, Schulz GE (2000) High-resolution structure of the OmpA membrane domain. J Mol Biol 298:273–282

Roosild TP, Greenwald J, Vega M, Castronovo S, Riek R, Choe S (2005) NMR structure of Mistic, a membrane-integrating protein for membrane protein expression. Science 307:1317–1321

Roseman MA (1988) Hydrophilicity of polar amino acid side-chains is markedly reduced by flanking peptide bonds. J Mol Biol 200:513–522

Sali A, Blundell TL (1993) Comparative protein modeling by satisfaction of spatial restraints. J Mol Biol 234:779–815

Sayle RA, Milner-White EJ (1995) RasMol: biomolecular graphics for all. Trends Biochem Sci 20:374–376

Schulz GE (2000) b-Barrel membrane proteins. Curr Opin Struct Biol 10:443–447

Shroll RM, Straatsma TP (2002) Molecular structure of the outer bacterial membrane of Pseudomonas aeruginosa via classical simulation. Biopolymers 65:395–407

Shroll RM, Straatsma TP (2003) Molecular basis for microbial adhesion to geochemical surfaces: computer simulation of Pseudomonas aeruginosa adhesion to goethite. Biophys J 84:1765–1772

Smart OS, Goodfellow JM, Wallace BA (1993) The pore dimensions of gramicidin A. Biophys J 65:2455–2460

Smart OS, Neduvelil JG, Wang X, Wallace BA, Sansom MSP (1996) Hole: a program for the analysis of the pore dimensions of ion channel structural models. J Mol Graph 14:354–360

Snyder DA, Bhattacharya A, Huang YJ, Montelione GT (2005) Assessing precision and accuracy of protein structures derived from NMR data. Proteins: Struc Funct Bioinf 59:655–661

Spronk C, Linge JP, Hilbers CW, Vuister GW (2002) Improving the quality of protein structures derived by NMR spectroscopy. J Biomol NMR 22:281–289

Tieleman DP, Berendsen HJC (1998) A molecular dynamics study of the pores formed by *Escherichia coli* OmpF porin in a fully hydrated palmitoyloleoylphosphatidylcholine bilayer. Biophys J 74:2786–2801

van Gunsteren WF, Kruger P, Billeter SR, Mark AE, Eising AA, Scott WRP, Huneberger PH, Tironi IG (1996) Biomolecular simulation: the GROMOS96 manual and user guide. Biomos & Hochschulverlag AG an der ETH Zurich, Groningen, Zurich

Vogt J, Schulz GE (1999) The structure of the outer membrane protein OmpX from *Escherichia coli* reveals possible mechanisms of virulence. Structure Fold Des 7:1301–1309

Vriend G (1990) WhatIf—a molecular modeling and drug design program. J Mol Graph 8:52–56