

Membrane protein structure quality in molecular dynamics simulation

Richard J. Law^{a,*}, Charlotte Capener^a, Marc Baaden^b, Peter J. Bond^a, Jeff Campbell^a,
George Patargias^a, Yalini Arinaminpathy^a, Mark S.P. Sansom^a

^a *Laboratory of Molecular Biophysics, Department of Biochemistry, The University of Oxford, The Rex Richards Building,
South Parks Road, Oxford OX1 3QU, UK*

^b *Institut de Biologie Physico-Chimique, CNRS-UPR9080, 13 rue Pierre et Marie Curie, F-75005 Paris, France*

Accepted 4 May 2005

Available online 15 August 2005

Abstract

Our goal was to assess the relationship between membrane protein quality, output from protein quality checkers and output from molecular dynamics (MD) simulations. Membrane transport proteins are essential for a wide range of cellular processes. Structural features of integral membrane proteins are still under-explored due to experimental limitations in structure determination. Computational techniques can be used to exploit biochemical and medium resolution structural data, as well as sequence homology to known structures, and enable us to explore the structure–function relationships in several transmembrane proteins. The quality of the models produced is vitally important to obtain reliable predictions. An examination of the relationship between model stability in molecular dynamics (MD) simulations derived from RMSD (root mean squared deviation) and structure quality assessment from various protein quality checkers was undertaken. The results were compared to membrane protein structures, solved at various resolution, by either X-ray or electron diffraction techniques. The checking programs could predict the potential success of MD in making functional conclusions. MD stability was shown to be a good indicator for the quality of structures. The quality was also shown to be dependent on the resolution at which the structures were determined.

© 2005 Elsevier Inc. All rights reserved.

Keywords: Protein structure; Homology modeling; Simulation; Protein quality checking; X-ray crystallography; Electron microscopy

1. Introduction

It is theoretically possible for every protein structure solved to be at atomic resolution, without the need for model refinement, in this case the quality would not need to be checked [1]. But this is not the case. For membrane proteins, in particular, the situation is worse due to the added difficulties with solving their structures [2]. This explains why only structures for 82 different membrane proteins (including homologous proteins from different organisms) (http://blanco.biomol.uci.edu/Membrane_Proteins_xtal.html) have been solved so far (<1% of the estimated total number of membrane proteins) even though they represent

about 30% of most genomes. However, this number is rapidly increasing with the release of several recent membrane protein structures with diverse biological function [3–11]. Homology modeling is a process that attempts to redress the balance somewhat. One caveat is that significant uncertainty remains concerning the atomic coordinates of both homology models and medium resolution structures and therefore it is important that we have methods to check their validity. The accuracy of a structure has direct implications on the elucidation of structure–function relationships and protein quality checkers can avoid drawing invalid mechanistic conclusions from bad starting structures. There is a need to assess the quality of a homology model, as well as asking how well it compares to experimentally determined structures of various resolutions. The large number of models for aquaporins and potassium channels is one area where model and experimental structure can readily be compared, and will be

* Corresponding author at: Chemistry and Biochemistry Department, 9500 Gilman Drive, University of California, San Diego, La Jolla, CA 92093-0365, USA. Tel.: +1 858 349 6369; fax: +1 858 349 6369.

E-mail address: rlaw@mccammon.ucsd.edu (R.J. Law).

studied in some detail in this work. The study also includes several other membrane proteins, namely OmpA, OmpLA (bacterial outer membrane proteins), GluR0 (Glutamate Receptor), and BtuCD (ABC transporter protein), to broaden the comparison.

There are potential errors and biases in the quality checking methods that reflect the way in which protein structures are refined and the way in which homology models are constructed. Refinement of a crystal structure, against less than atomic resolution (i.e. 1 Å or better) data, uses a library of stereochemical information to optimize the model. Homology modeling programs such as Modeller and What-If use similar libraries to produce homology models, as do protein quality checking programs such as WhatCheck [12] and Procheck [13]. It is thus important to ensure that one is not simply examining the refinement process but really assessing the genuine quality of the structure. This implies the use of different libraries and/or datasets for structure determination and quality assessment. On the other hand, if one library is used to refine a structure, another one to check it and then a force-field is employed in molecular dynamics, which results are the most reliable and to what extent are differences dependent on these representations? One way to objectively check structures would be to use a database made from very high quality structures, determined via direct methods, at atomic resolution. Yet currently such a database, especially for membrane proteins, cannot be compiled, as there are not enough structures available.

This makes modeling heavily dependent on the quality of the library and force-field data used. Most refinement algorithms use a database of small molecule and peptide structures for the geometric attributes, such as Engh & Huber '91—constructed from the CSD (Cambridge Structural Database-1983). Structures in the PDB (Protein Data Bank) [14] are then used for more complex conformational restraints and serve as libraries for non-stereochemical factors such as atom volume, atom–atom contacts, or hydrogen bonding networks. Comparison to a separate set of high-resolution structures provides an independent assessment of these libraries. But this methodology does not always come up with correct models, as there are many structures at lower resolution that contain significant errors [15]. This leads to a dilemma, as without an a priori fundamental understanding of the exact atomic resolution driving forces of protein folding, the determined structures are the only reliable experimental information we have, both in terms of constructing models and in terms of designing programs to check their quality.

It has been shown that the checkers examined (Procheck, Prove [16], and WhatCheck) agree very well for a test set of twelve soluble protein structures at atomic resolution, used for comparison [17]. One can then create a reference library of structural attributes that were not applied during the refinement and determine how these attributes are related to structure quality, for subsequent use by a protein quality checker. Frequent updating of libraries used by the checkers

is required in order to take the constantly increasing number of high-resolution structures into account thus ensuring an accurate calibration of the quality assessment. In the context of this paper it seems especially desirable to eventually increase the number of membrane proteins represented.

The protein quality checkers used in this study were chosen according to several criteria. Firstly, the checker must be able to assess the quality and hence biological relevance of a given model. Secondly, it should be possible to simplify the checker results and to extract the most relevant information for the assessment of homology models. And lastly, as molecular dynamics is often subsequently used to draw biological/physiological conclusions about the structure–function relationship of a given protein, we wanted to relate checker results to the stability of the protein in MD simulation. Are static model checking results consistent with the stability of an MD simulation of the same model? How does MD affect the checked properties, for instance in a short 2 ns simulation? Is it possible, if expensive, to use MD to check model quality and can it distinguish between two very similar models where other methods cannot? To address these questions, we examined the following protein quality checking methods: phi/psi (Φ/Ψ) angle check in Procheck, WhatCheck New Packing Quality analysis (directional contact analysis), WhatCheck Chi1/Chi2 (χ_1/χ_2) check, and Prove atom volume score. Procheck measures how well the Φ/Ψ fit to ideal values in the Ramachandran plot. Packing analysis scores the contact distances between atoms in residues adjacent in the structure. χ_1/χ_2 analysis tests the fit of the side-chain angles to ideal values. Prove checks whether the “volume” of atoms, based on a Voronoi cell description of each buried atom, is correct, by its contacts and distances to its neighbouring atoms. All these checkers use a force-field/dataset based on already determined structures.

2. Methods

2.1. System preparation

Homology modeling used Modeller v4.0 [18] (<http://www.salilab.org/modeller/modeller.html>) with the 2.2 Å resolution structure of GlpF (1FX8) [19] as a template for the aquaporins (structures and models 2–8 in Table 1), and the 2.8 Å KcsA structure [20] for the potassium channel models (models 10–15 in Table 1). Four-fold symmetry was imposed when modeling the tetramer structures. Sequence alignments for input to modeling were generated using ClustalW in Jalview (<http://www.ebi.ac.uk/jalview/index.html>). Prior to running simulations, side-chain ionisation states were adjusted to match the results of pK_a calculations, performed as described in [21], using UHBD [22] to calculate free energy differences between ionised and unionised side-chains of the protein. Crystal structures were used to set up the BtuCD, OmpA [23] and OmpLA [24] MD

Table 1

The table presents results for the selected protein quality checks, including three qualitative ranges as defined by the documentation for each program

#	Source	Protein	PBD#	Resolution	Procheck Φ/Ψ (%)	WhatCheck packing	WhatCheck χ_1/χ_2	Prove atom volume	MD C α RMSD (2 ns)
1	X-ray	GlpF	1FX8	1.8	90	0.52	0.39	-0.09	1.9
2	X-ray	bAQP1	1J4N	2.2	82	-1.80	-2.15	-0.14	2.1
3	EM	hAQP1	1IH5	3.7	71	-5.17	-4.92	0.15	4.5
4	EM	hAQP1	- ^a	3.7	70	-4.88	-4.93	0.13	6.1
5	EM	hAQP1	1FQY	3.8	69	-4.56	-3.58	0.16	5.4
6	Homology	hAQP1	1H6I	–	80	-2.48	-4.37	0.44	3.4
7	Homology	hAQP1	–	–	80	-2.84	-2.06	-0.39	3.1
8	Homology	AQPZ	–	–	82	-2.43	-2.71	-0.46	3.3
9	X-ray	KcsA	1K4C	1.8	97	-0.43	1.23	0.10	1.5
10	Homology	Kir 6.2-A	–	–	76	-4.32	-4.02	-0.05	4.0
11	Homology	Kir 6.2-B	–	–	71	-3.09	-3.99	-0.51	4.8
12	Homology	Shaker	–	–	84	-0.52	-3.25	-0.17	3.0
13	Homology	Twk-5CW	–	–	73	-3.48	-3.89	-0.89	4.5
14	Homology	Twk-5ACW	–	–	69	-3.89	-3.86	-0.69	4.5
15	Homology	GluR0	–	–	85	-1.44	-3.15	-0.35	3.1
16	X-ray	BtuCD	1L7V	3.2	85	-0.68	-1.10	0.03	3.1
17	X-ray	OmpA	1BXW	2.5	92	-2.82	-1.63	0.35	2.5
18	X-ray	OmpLA	1QD6	2.1	86	-1.10	-0.44	0.10	2.1

Bold indicates good, normal text, medium, and italic, a bad structure. This is compared to the drift of each structure in a 2 ns molecular dynamics simulation and the experimental resolution of the structure. The homology models do not have resolutions.

^a A refined version of the 1IH5 hAQP1 (Ren et al., 2001).

runs [25–29]. Structures, models and selected snapshots from the trajectories were checked and compared using Procheck, Prove and WhatCheck. The average z -scores for the whole protein, and fragments, were recorded. Correlation coefficients between z -scores, root mean square deviation (RMSD) from the starting structure in MD and resolution of the experimental data were calculated from a line of regression (best fit). For analysis, the absolute values of the Prove scores were taken, as the highest quality z -score here is actually centred at zero.

For the comparative simulations on aquaporins and potassium channels each model/structure was embedded in a membrane mimetic octane slab [30,31] and the system was then solvated with SPC [32,33] water. Counter-ions were added to yield an electroneutral system that was energy minimized prior to starting simulations.

2.2. Simulations

All simulations were performed with GROMACS using a Nose–Hoover thermostat, Berendsen barostat, and NPT and periodic boundary conditions. A constant pressure of 1 bar was applied independently in all three directions, using a coupling constant of $\tau_p = 1.0$ ps. Water, octane and protein were coupled separately to a temperature bath at 300 K using a coupling constant $\tau_T = 0.1$ ps [34]. Long-range interactions were dealt with using a twin-range cut-off: 10 Å for van der Waals interactions and 17 Å for electrostatic interactions. The timestep was 2 fs, using LINCS [35] to constrain bond lengths, and the force-field was based on GROMOS 96 [32]. Each system underwent 100 steps of steepest-gradient minimisation following the set-up. Prior to the production run a 500 ps equilibration simulation was

conducted, for each system, with the position of the non-hydrogen protein atoms restrained with a force constant of $1000 \text{ kJ mol}^{-1} \text{ nm}^{-2}$. This was to enable the water and octane to pack properly around the protein.

3. Results

3.1. Protein quality versus MD

The results presented here are a simplified view of the output from the checking programs. Eighteen starting models, on which simulations were performed in a solvated octane slab, were used in the study. The initial structures (described in more detail in Section 2) were three EM structures, nine homology models, and six X-ray structures. These are summarized in Table 1. z -Scores from the checker programs are reported here. The X-ray structures are the most stable in simulation, which confirms the observations made recently by Fan and Mark for a set of 5 ns globular protein MD simulations starting from X-ray and NMR structural data [36]. The homology models are intermediate, and the EM structures show the highest fluctuations. These fluctuations observed in MD were compared to results from various protein quality checking programs. The analyzed parameters were clustered into three qualitative bands (based on the documentation for each method) in order to obtain an overall impression of quality for each structure.

By plotting the C α RMSD at the end of an MD simulation against the various protein quality checker results one can observe a correlation between quality of the static structure and dynamic stability in a short 2 ns simulation (Fig. 1). If the absolute value of the correlation coefficient (r) is ≥ 0.8

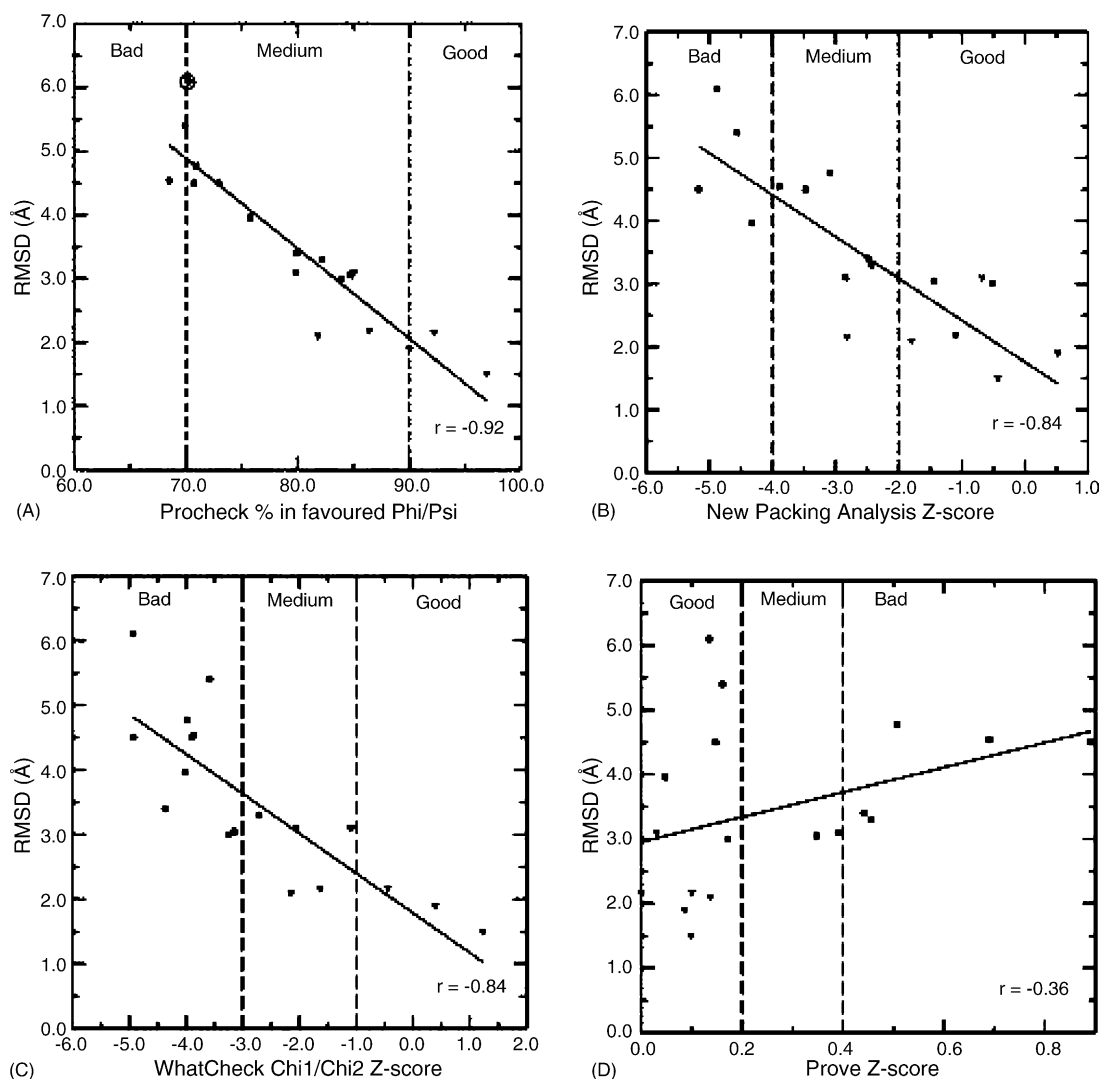


Fig. 1. This graph illustrates the correlation between RMSD in an MD simulation starting from a given structure and (A) Procheck percentages for Φ/Ψ angles lying within the favoured regions of the Ramachandran plot, (B) the results of the new version of the atomic directional contact analysis from WhatCheck, (C) the results of a χ_1/χ_2 angle check and (D) Prove atom volume analysis results. For each graph, the line is a regression line calculated for the compared data. The correlation coefficient for this data is then indicated in each plot. The different point symbols indicate the source of the structure: X-ray (▼), homology model (■), or EM structure (●).

then the two sets of data are strongly correlated. If $r = 0.7$ then $r^2 = 50\%$ (correlation determination factor), which means that less than 50% of the data is explained by the correlation line. Therefore if the correlation coefficient is less than 0.7 then the two sets of data cannot be considered to have a statistically significant correlation. Procheck Φ/Ψ , WhatCheck packing analysis, and WhatCheck χ_1/χ_2 are well correlated with the RMSD seen in the 2 ns simulation. Prove atom volume, and bond angle/length checks (results not shown) were not well related to simulation performance.

The homology models show the same level of correlation as the whole data set for the examined structures. Due to their differing overall quality, the homology models do cluster in a slightly different region of the plot than the EM and X-ray structures. The X-ray structures are generally better quality, both in terms of the static checkers and the

stability in simulation, probably as a result of the improved resolution, relative to the EM structures. This agrees well with Fan et al.'s results [36].

There is a noticeable outlier, highlighted on the Procheck Φ/Ψ versus RMSD plot. It is the Ren et al. [37] refined hAQP1 structure (structure no. 4 in Table 1). The refinement did not increase the protein's stability in simulation but in fact reduced it. In this case interactive modeling of the structure prevented Procheck from distinguishing between good and bad models, but packing analysis and MD simulation do assess the quality as being bad [38].

It is interesting to look in more detail at the regions of the structures giving rise to high RMSD values, and determine whether the stability of these fragments can be predicted as reliably as the overall stability. For this purpose the $C\alpha$

RMSD and Procheck results for transmembrane (TM) segments were compared to those of the loop regions. In each case the values for the TM segments indicate a higher stability than for the loops. Concerning the Procheck results this is hardly surprising given that the Ramachandran plot on which the check is based relies on the satisfaction of ideal secondary structure Φ/Ψ bond angles and loops often contain non-ideal, non-secondary-structure bond angles. Loops/turns are the most mobile sections of membrane proteins, and their increased mobility in simulations has been frequently observed and has been abundantly reported in the literature. For the whole structures, the homology models and X-ray/EM structures show a similar correlation between Φ/Ψ score and RMSD. Yet, for the homology models, both for helical TM segments, and even more so the loop regions, the situation is markedly different. The correlation coefficient, r , is actually positive for the modeled TM segments of these models, indicating an inverse relationship. This seems to suggest that a better protein quality score for the transmembrane helices results in a greater RMSD, and thus increased flexibility of the structure. This seems, at first, counter-intuitive and may indicate a flaw in part of the homology modeling process. For instance, all of the K-channel homology models (six of the nine homology models) were produced using ideal helical restraints on the TM sections, which may lead to exaggerated strain in the overall structure and thus increased flexibility to accommodate the different segments. The strain imposed by the compliance with these restraints may be explained by unsatisfactory packing between the helices, neglected inter-helix interactions and distortion of the loop. The aquaporin homology models were constructed without using TM helical restraints and subsequently the relationship between the Φ/Ψ scores and RMSD is more similar to that of the experimentally determined structures. Perhaps another desirable improvement would be better loop modeling as it might make these regions more stable in simulation and more favourably assessed by the checker programs (all of the models were produced with Modeller 4.0. A loop-modeling module exists in Modeller v.6.1). Yet the loop mobility may be indicative of an intrinsic property of these regions.

For the experimentally solved structures we observe a greater correlation of Procheck scores with RMSD for the loops compared to the TM segments. This may indicate that accurate determination of loop structure is a more sensible probe for the accuracy of the whole model than the assessment of the TM segments. As this feature is highly resolution dependent, loops are generally of better quality in the X-ray structures than in the EM structures, as long as their conformational freedom is not affected by crystal packing. This does not hold for the OmpA structure where the loops had to be modeled separately as they were not well resolved in the X-ray data due to their mobility. Interestingly, the quality of the loops can also be independently assessed by analysis of the convergence of an MD simulation [39].

As expected, the drift of the structures in MD is not correlated with the bond angle or bond length ($r = 0.24$ and 0.1 , respectively) analyses (results not shown). This is expected as standard values are used for these parameters during refinement, whether the method used is X-ray diffraction, EM, or homology modeling and similar bond parameters are again used in the force-field terms of a subsequent MD simulation. The only difference might arise when a position restraint based refinement program, such as Shelxl [40] is used for refinement of atomic resolution X-ray structures [17]. No membrane proteins have yet been solved at a high enough resolution that this kind of minimal refinement has been possible.

There are currently two versions of the WhatCheck packing analysis that had a very similar correlation to MD RMSD, hence only the results for the newer version are presented. New packing analysis looks with greater precision at the structure and is meant for X-ray structures at better than 2.0 \AA resolution (Personal Communication, G. Vriend). No marked difference is seen between the distribution of the points for the homology models, or the experimental structures, for this quality check.

The χ_1/χ_2 torsion angle checking method implemented in WhatCheck is well correlated with the MD drift from the starting structure. In this method the arrangement of the residue side-chains is analyzed. Optimization of such positions is not considered paramount in homology modeling or for the refinement of medium resolution structures where the side-chains may not be well defined by the electron density. This is especially true for 2D EM structures where an anisotropic distribution of the resolution is observed, and the resolution perpendicular to the membrane is worse than in the plane of the membrane. This anisotropy renders the assignment of residues to the density difficult and can lead to mis-assignments. The effect is very similar to the mis-threading that can occur in homology models due to incorrect sequence alignment. Both mis-assignments and mis-threading can cause side-chain structure disruption due to incorrect packing of the side-chains. Given the large number of homology models and EM structures (and structures at a wide range of resolutions) we examined this factor in detail and showed that it is indeed very important for the quality and stability of a protein structure.

The Prove results were not consistent with the RMSD seen in simulation. Prove appears to assess too many structures as being “good”—in particular the EM Aquaporin structures. But it does still distinguish between the best and worst homology models. Prove was one of the checker analyses that looked at an attribute not intentionally refined in the structures and therefore could possibly make an independent assessment of the quality. Many of the methods, in terms of studying membrane proteins, may suffer from the lack of high-resolution data in the PDB. This is particularly true for the Prove method, which is critically dependent on the number of available membrane protein structures [16]. It

is also possible that the atom volume measured by Prove may be indirectly refined in the same way as simpler stereochemical factors.

Looking at the details of the results for the checkers (Table 1) there are few cases where all methods agree on the ‘quality’ of a given structure. This suggests using more than one checker for quality assessment. Focusing on overall consistency and correlation with MD data, a recommended set of checkers for membrane proteins would be a combination of Procheck, directional contact (packing) analysis and χ_1/χ_2 analysis from WhatCheck. Since simulation is well correlated with several of these methods, it seems that MD is indeed sensitive to protein quality and does in some cases provide more detailed results. For instance the simulations could distinguish between various aquaporins and the Kir 6.2 models where the checker results were much more inconclusive. The consistency of the results may at first seem surprising, given the incomplete sampling in such a short 2 ns MD simulation. However, the major rearrangements of badly packed protein fragments and side-chain conformations occur in the first few hundred picoseconds of a simulation and are represented by a jump in RMSD observed at the beginning of each simulation. Differences in force-fields and between the modeling and simulation package certainly contribute to these changes but only represent a small part compared to the RMSD jump due to a bad starting structure. This also indicates that MD should be able to finely distinguish between “bad” structures that reach high RMSD values, but will be more insensitive to “good” structures. An additional factor, perhaps, is the different environment in the simulation (membrane mimetic) and in the experimental structure (crystal), but this environment change is not consistent enough that conclusions about its effect can be made.

3.2. Homology model refinement using quality checkers?

All checker programs provide not only an overall assessment of a structure but also indicate the contribution of each residue, and thus allow us to identify the bad parts in a structure or the potential errors in the underlying sequence alignment. Valuable insight into the quality of the different parts of a homology model can be gained by investigating the contributions of each residue to the quality assessment of the structure.

Stretches of contiguous sequence where the packing is particularly bad can indicate areas of misalignment in homology models. Such information can be used to refine the underlying sequence alignment, and hence allow us to improve the homology models for the lower quality structures. The residue contributions can also help to distinguish between two possible alignments, as was the case for the two Kir 6.2 models, A and B. Indeed, model B contains a stretch with a particularly abnormal packing environment whereas Model A does not. This indicates that

the Model A alignment is the correct one and is corroborated by MD simulations, which distinguish between the models in the same way [41]. On the other hand, the packing analysis actually gives Model B an overall better score so, on the whole, this issue appears somewhat inconclusive.

The Shaker, GluR0, hAQP1 and hAQPZ homology models do not appear to contain such stretches of four or more residues with abnormal packing suggesting that their alignments are correct.

Torsion angle checks might provide another means to discriminate between less reliable homology models, as it has been reported that errors in sequence alignment lead to distorted χ_1/χ_2 angles. We tested this hypothesis for the case of the Twk-5 models, where our aim was to reveal stretches of residues with particularly bad χ_1/χ_2 angles. This would have pinpointed areas of misalignment and thus provide direct guidance in refinement of the model. Unfortunately it was not possible to distinguish between the two Twk-5 models, differing by their subunit arrangement (clockwise or counter-clockwise). The models had a bad overall χ_1/χ_2 z-score, but neither was good enough to reveal an improved quality. This is probably due to the uncertainties in the alignment, especially for the second pore-domain. χ_1/χ_2 z-score analysis agrees with Prove and indicates model 14 being of slightly better quality than model 13. Procheck and packing analysis favour model 13.

3.3. Protein quality as a function of resolution

Can ‘good’ resolution be defined by protein quality checks? The protein quality checker results were well correlated with the resolution of the starting structures obtained from EM or X-ray diffraction. A ‘good’ resolution threshold for membrane proteins can be set between 2.2 and 2.6 Å based on the correlation of the checker results with resolution (graph not shown). This is based on the lower limit set for a high quality structure defined by each quality checker program. The 2.2 Å threshold is not surprising as these methods were parameterized with structures of better than 2.2 Å resolution. Yet that does not imply that it is not a valid result for this reason, as higher resolution structures are, by definition, more correct [15].

3.4. Reasonable RMSD from good protein quality

One may also use the checker results to relate resolution and protein quality to possible artifacts in an MD simulation. This allows the definition of an acceptable C α RMSD for membrane proteins in 2 ns of carefully set-up simulations in an octane slab. Resolution versus C α RMSD for X-ray/EM structures (Fig. 2) is probably not well represented by a simple regression line, although the calculated correlation coefficient of 0.89 is high. The standard deviation of the points around the regression line is high and a *t*-test of the statistical significance only just falls within 95% confidence limits. One caveat is that RMSD values depend on the size of

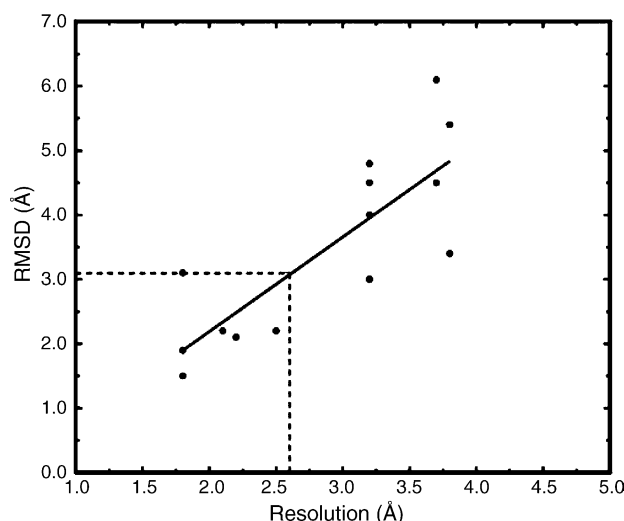


Fig. 2. Correlation of the resolution of EM/X-ray structures with the deviation in MD simulation. The dashed line demonstrates the calculation of the “good deviation” in an MD simulation, based on the protein quality checkers and the resolution.

the protein which can lead to some dispersion of the data points. But more importantly, the spread of the data points in this graph clearly suggests that the level of correlation is

actually dependent on the window of resolution investigated. That is, the lower the resolution the less confident we can be about predicting how much deviation we will see in a particular structure, in simulation, based purely on that value.

The resolution was also compared to RMSD for just the TM segments but their correlation was much lower than for the whole structure. One would expect this as the resolution is calculated for the whole structure, and thus biased by the less reliably characterized parts like loops, turns and mobile stretches. This is true for both α -helical and β -barrel structures. In the β -barrel structures (OmpA and OmpLA) the TM β -barrel tends to be very stable, yet the loops are very mobile. High resolutions are seen for these structures, the barrels are very stable in simulation, yet the loops were seen to be extremely mobile, dominating the RMSD for that experiment.

It is also possible to use RMSD and checker results to extrapolate the RMSD deviation of a ‘good’ protein structure in MD simulation (Fig. 2). Hence the upper limit for deviation of a protein structure in 2 ns of MD, as indirectly defined by the protein quality checkers, using the graphs in Fig. 2, is 2.9 Å, for all the C α atoms of the structure. Of course, using these results, an estimate of an

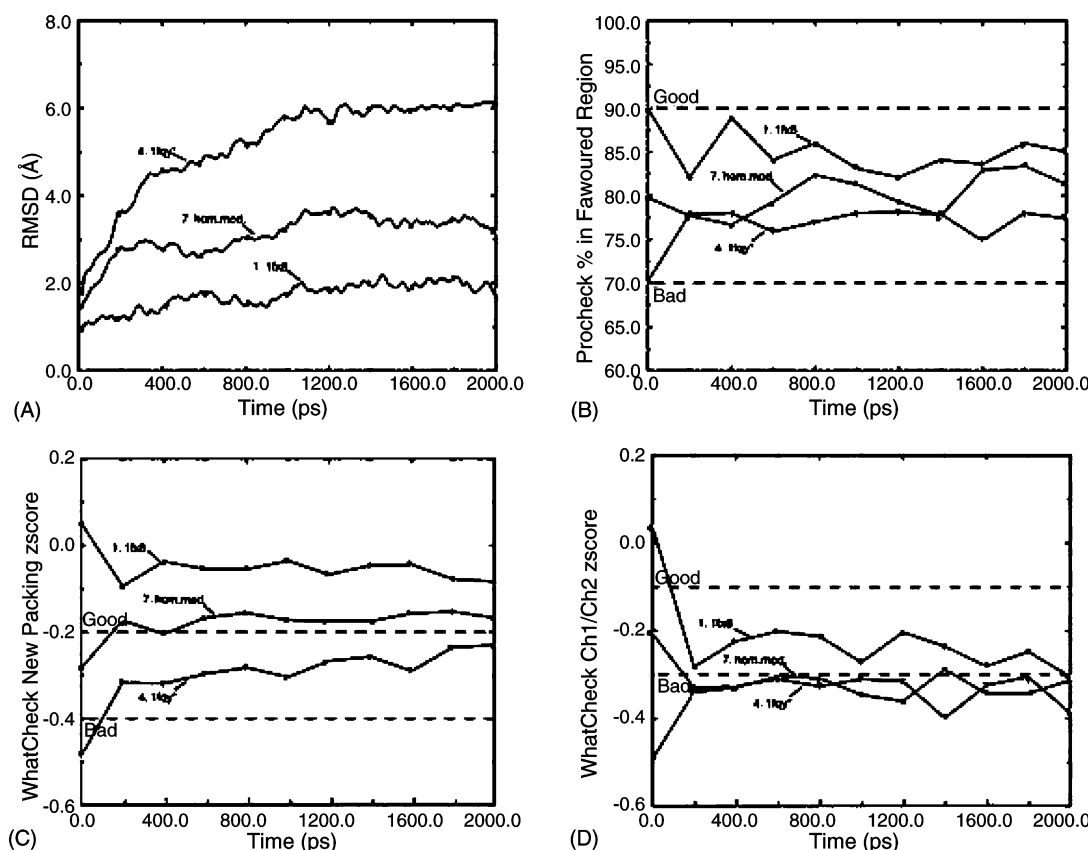


Fig. 3. These plots demonstrate the effect of MD simulation on the three protein checker programs that best correlate with the drift observed in MD from initial structure: (A) shows the RMSD of the three structures chosen for comparison—GlpF (1.1fx8) (—▼—), Aqp1 homology model (7.hom.mod.) (—■—), Aqp1-11h5-ref. Model (4.11h5*) (—●—); (B) shows the effect on Procheck Φ/Ψ scores; (C) plots the effect on WhatCheck packing quality, and (D) shows the result on the WhatCheck χ_1/χ_2 scores.

acceptable RMSD can be taken directly from each checking program and does not require the extra step of defining a good resolution. The three checking programs that are well correlated with MD RMSD all give an estimate of acceptable C α RMSD at around 3.1 Å from the starting structure.

3.5. MD simulation effects on protein quality

In MD simulation bond angles and lengths will change quickly, mostly in minimization, simply due to differences between the modeling/refinement force-field and the MD force-field. But what is the effect on the packing quality and general protein quality? Three structures were chosen to demonstrate the influence of MD on the checker scores: one of the best structures (Glpf, model no. 1), one of the worst structures (1ih5-ref, model no. 4) and an intermediate one (Aqp1 homology model, model no. 7) (Fig. 3). Contrary to the correlation of MD with the quality checkers, these results suggest that the differences in force-field do affect the chosen tests as we observe a degree of convergence of the checked property with simulation time for each case. The χ_1/χ_2 scores mostly drop below the threshold for a bad structure. The Procheck scores tend to converge around an intermediate value. The packing in the worst structure is improved, but the packing is slightly worsened in the best structure. The packing of the intermediate structure, the homology model, is improved. The latter suggests that MD may have a role in the improvement of medium resolution (~ 4 Å) structures and homology models. And indeed, as well as restricted molecular dynamics being used in some solved structure refinement programs [42], it has also been used in the completion/refinement of homology models [43].

4. Conclusions and discussion

Protein Quality, as assessed by several checking programs is well correlated with stability in MD simulation, as well as resolution of solved structures. It is insufficient to check only the aspects of the protein structure which are very likely to be correct due to the employed homology modeling or structure refinement method; i.e. local factors such as bond-angles and lengths. It is instead necessary to check parameters related to the stability of the overall structure like those assessed in the packing quality and χ_1/χ_2 checks. Procheck Φ/Ψ angle check also gives very good results in terms of predicting accuracy in MD simulation. It is prudent to use more than one program in the evaluation process, as the results of a single test can sometimes be inconclusive for medium resolution or medium quality models.

It appears that MD is capable of differentiating, in terms of quality and stability, between two models that are very similar, even when the checker analysis is inconclusive. For instance, in this study, the two Kir homology models, the two Twk-5 homology models, and the various Aquaporin

models were compared. These models differ from each other only slightly. In two of these three cases, MD was capable of distinguishing between the models and correlated with several of the protein quality checking programs, in each case. The two Kir 6.2 models differ by their alignment to KcsA and MD seems capable of discerning the most stable model, which is the one based on a better alignment. The alignment mostly determines the quality of a homology model and changes therein most radically change the quality of the resulting structure.

Two possible criticisms of this study are the size and somewhat unrepresentative nature of the sample set. That is, only 18 structures were used and most were either potassium channel or aquaporin family structures. The dataset was selected for its source diversity. That is, both α -helix and β -barrel, as well as X-ray, EM, and homology models, from bacterial and mammalian cells, and from a range of resolutions. The results of this study seem to be independent of these differences and can be applied to any of these systems. Any lack of representative sampling is symptomatic mainly of the deficiency in membrane structures available for study in the PDB (although this situation is beginning to improve [44]). Homology modeling can help to bridge this gap but only for similar proteins for which structures already exist. Homology models can, as seen in studies with aquaporin, sometimes provide more reliable information than medium resolution (~ 4 Å) EM structures.

It is a philosophical point perhaps, but might a protein structure require a degree of ‘instability’ in order to be active? It has been shown that small deviations from ideal structural attributes may be at regions of functional significance in the structure [16]. It is then not an easy task to distinguish between deviations that are ‘allowed’ and those that are due to experimental error or deficiencies in the simulation protocol or force-field.

A parallel investigation of the relationship between protein quality and structural drift in a lipid simulation would be interesting, although longer timescales are required for a correlation to be seen. At the moment though, not enough simulations have been conducted to make such a comparison. If such an investigation could be carried out, it may be possible to determine the degree to which large scale motions are altered by a membrane environment. An investigation of the temperature environment might also be informative since the checker programs are based mainly on low temperature structures yet the simulations are run at room temperature.

The influence of the quality of the template on the quality of the homology model has not been directly investigated. The overwhelming evidence points to the crucial role of the pairwise alignment between the template and target sequence. There are only a few protein structures available where this could be checked, but potassium-channels (KcsA; 3.2 and 1.8 Å), Aquaporins (Aqp1; various resolutions), and perhaps Bacteriorhodopsin [45] (various resolutions) may provide reasonable test sets.

Now that more than one member of these membrane protein families has a solved structure, one could also begin to compare the quality of the homology modeling programs for use with membrane proteins. This has been done for soluble proteins [46]. It could also be used to check the ability of the sequence alignment algorithms for membrane proteins against direct structural alignments—the only truly ‘correct’ alignment.

Acknowledgements

RJL was supported by a MRC studentship; research in MSPS’s laboratory is supported by grants from the Wellcome Trust, the BBSRC and the EPSRC. MB thanks the EU for a grant (Contract No. QLK2-CT-2000-51210). Our thanks also to Gary Ren and Alok Mitra for providing two of the EM models used in this study. Thank you also to R.W.W. Hooft, G. Vriend and P. Keller for helping to construct and maintain the Biotech server (<http://biotech.ebi.ac.uk:8400/>) on which these investigations were carried out.

References

- [1] C. Branden, T. Jones, *Nature* 343 (1990) 687–689.
- [2] J. Rosenbusch, *J. Struct. Biol.* 136 (2001) 144–157.
- [3] C. Oomen, P. Van Ulsen, P. Van Gelder, M. Feijen, J. Tommassen, P. Gros, *EMBO J.* 23 (2004) 1257–1266.
- [4] D. Chimento, A. Mohanty, R. Kadner, M. Wiener, *Acta Crystallogr. Sect. D—Biol. Crystallogr.* 59 (2003) 509–511.
- [5] Y. Jiang, V. Ruta, J. Chen, A. Lee, R. MacKinnon, *Nature* 423 (2003) 42–48.
- [6] A. Kuo, et al. *Science* 300 (2003) 1922–1926.
- [7] A. Miyazawa, Y. Fujiyoshi, N. Unwin, *Nature* 423 (2003) 949–955.
- [8] R. Dutzler, E.B. Campbell, R. MacKinnon, *Science* 300 (2003) 108–112.
- [9] J. Abramson, I. Smirnova, V. Kasho, G. Verner, H.R. Kaback, S. Iwata, *Science* 301 (2003) 610–615.
- [10] W.Y. Huang, M.J. Lemieux, J. Song, M. Auer, D.N. Wang, *Science* 301 (2003) 616–620.
- [11] V. Yankovskaya, et al. *Science* (2003) 299.
- [12] G. Vriend, C. Sander, *J. Appl. Crystallogr.* 26 (1992) 47–60.
- [13] R.A. Laskowski, M.W. Macarthur, D.S. Moss, J.M. Thornton, *J. Appl. Crystallogr.* 26 (1993) 283–291.
- [14] H. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, P.E. Bourne, *Nucl. Acids Res.* (2000) 28.
- [15] R.W.W. Hooft, G. Vriend, C. Sander, E.E. Abola, *Nature* 381 (1996) 272.
- [16] J. Pontius, J. Richelle, S. Wodak, *J. Mol. Biol.* 264 (1996) 121–136.
- [17] G.G. Wilson, A. Karlin, *Neuron* 20 (1998) 1269–1281.
- [18] A. Sali, T.L. Blundell, *J. Mol. Biol.* 234 (1993) 779–815.
- [19] D. Fu, A. Libson, L.J.W. Miercke, C. Weitzman, P. Nollert, J. Krucinski, R.M. Stroud, *Science* 290 (2000) 481–486.
- [20] D.A. Doyle, J.M. Cabral, R.A. Pfuetzner, A. Kuo, J.M. Gulbis, S.L. Cohen, B.T. Cahit, R. MacKinnon, *Science* 280 (1998) 69–77.
- [21] C. Adcock, G.R. Smith, M.S.P. Sansom, *Biophys. J.* 75 (1998) 1211–1222.
- [22] M.E. Davis, J.D. Madura, B.A. Luty, J.A. McCammon, *Comput. Phys. Commun.* 62 (1991) 187–197.
- [23] A. Pautsch, G.E. Schulz, *Nature Struct. Biol.* 5 (1998) 1013–1017.
- [24] H. Snijder, I. Ubarretxena-Belandia, M. Blaauw, K. Kalk, H. Verheij, M.R. Egmond, N. Dekker, B.W. Dijkstra, *Nature* 401 (1999) 717–724.
- [25] P. Bond, J. Faraldo-Gómez, M.S.P. Sansom, *Biophys. J.* 83 (2002) 763–775.
- [26] P. Bond, M.S.P. Sansom, *J. Mol. Biol.* 329 (2003) 1035–1053.
- [27] C. Domene, P.J. Bond, S.S. Deol, M.S.P. Sansom, *JACS* 125 (2003) 14966–14967.
- [28] O. Beckstein, P.C. Biggin, P. Bond, J.N. Bright, C. Domene, A. Grottesi, J. Holyoake, M.S.P. Sansom, *FEBS Lett.* 555 (2003) 85–90.
- [29] M. Baaden, C. Meier, M.S.P. Sansom, *J. Mol. Biol.* 331 (2003) 177–189.
- [30] D.P. Tieleman, H.J.C. Berendsen, M.S.P. Sansom, *Biophys. J.* 80 (2001) 331–346.
- [31] C.E. Capener, M.S.P. Sansom, *J. Phys. Chem. B* 106 (2002) 4543–4551.
- [32] J. Hermans, H.J.C. Berendsen, W.F. van Gunsteren, J.P.M. Postma, *Biopolymers* 23 (1984) 1513–1518.
- [33] W.F. van Gunsteren, P. Kruger, S.R. Billeter, A.E. Mark, A.A. Eising, W.R.P. Scott, P.H. Huneberger, I.G. Tironi, *Biomos & Hochschulverlag AG an der ETH Zurich, Groningen & Zurich*, 1996.
- [34] H.J.C. Berendsen, J.P.M. Postma, W.F. van Gunsteren, A. DiNola, J.R. Haak, *J. Chem. Phys.* 81 (1984) 3684–3690.
- [35] B. Hess, H. Bekker, H.J.C. Berendsen, J.G.E.M. Fraaije, *J. Comput. Chem.* 18 (1997) 1463–1472.
- [36] A. Fan, A.E. Mark, *Proteins: Struct., Funct. Genet.* 52 (2003) 111–120.
- [37] G. Ren, V. Redy, A. Cheng, A. Mitra, *PNAS* 98 (2001) 1398–1403.
- [38] B. de Groot, A. Engel, H. Grubmuller, *FEBS Lett.* 504 (2001) 206–211.
- [39] J. Faraldo-Gómez, G.R. Smith, M.S.P. Sansom, *Biophys. J.* 85 (2003) 1406–1420.
- [40] Sheldrick, Schneider, *Meth. Enzymol.* 277 (1997) 319–343.
- [41] C. Capener, M. Sansom, *J. Phys. Chem. B* 106 (2002) 4543–4551.
- [42] R.B. Sutton, D. Rasshauer, R. Jahn, A.T. Brunger, *Nature* 395 (1998) 347–353.
- [43] J. Flohil, G. Vriend, H. Berendsen, *Proteins: Struct., Funct. Genet.* 48 (2002) 593–604.
- [44] T. Noguchi, Y. Akiyama, *Nucl. Acid Res.* 31 (2003) 492–493.
- [45] H. Luecke, B. Schobert, J.K. Lanyi, et al. *Science* 293 (2001) 1499–1503.
- [46] A. Marchler-Bauer, A.R. Panchenko, N. Ariel, S.H. Bryant, *Proteins: Struct., Funct. Genet.* 48 (2002) 439–446.