Volume 327, number 3, 256–260 FEBS 12755 August 1993

Structure and dynamics of bacteriorhodopsin

Comparison of simulation and experiment

M. Ferrand[a,b], G. Zaccai[a,c], M. Nina[b], J.C. Smith[b], C. Etchebest[d] and B. Roux[e,f]

[a]*Institut Laue Langevin, BP 156X, 38042 Grenoble Cedex, France,* [b]*Service de Biophysique des Protéines et des Membranes, Département de Biologie Cellulaire et Moléculaire, C.E.-Saclay, 91191 Gif-sur-Yvette, France,* [c]*Laboratoire de Biophysique Moléculaire, Institut de Biologie Structurale, 41 Avenue des Martyrs, 38027 Grenoble Cedex 1, France,* [d]*Institut de Biologie Physico-Chimique, 13 rue Pierre et Marie Curie, 75005 Paris, France,* [e]*Laboratoire d'Ingéniérie et d'Etudes des Protéines, C.E.-Saclay, 91191 Gif-sur-Yvette, France* and [f]*Département de Physique, Université de Montréal, C.P. 6128 succursale A, Montréal H3C 3J7, Canada*

Received 27 May 1993

Global features of the structure and dynamics of bacteriorhodopsin are investigated using molecular modelling, dynamical simulations and neutron scattering experiments. The simulations are performed on a model system consisting of one protein molecule plus intrinsic water molecules. The simulation-derived structure is compared with neutron diffraction data on the location of water and with the available electron microscopy structure of highest resolution. The simulated water geometry is in good accord with the neutron data. The protein structure deviates slightly but significantly from the experiment. The low-frequency vibrational frequency distribution of a low-hydration purple membrane is derived from inelastic neutron scattering data and compared with the corresponding simulation-derived quantity.

Bacteriorhodopsin; Molecular dynamics; Neutron scattering

1. INTRODUCTION

Bacteriorhodopsin (BR), a 26 kDa retinal binding protein, is a light-activated proton pump organised with lipids on a highly ordered two-dimensional lattice in the purple membrane patches of *Halobacterium halobium*. The elucidation of its three-dimensional structure at relatively high resolution using electron microscopy [1] has given impetus to theoretical efforts directed at an atomic detail understanding of the function of the molecule. However, considerable uncertainty remains on some important aspects of the structure, particularly on the side-chain and water positions [1]. As a result, stable molecular dynamics simulations in which the molecule is equilibrated and does not deviate significantly from the experimental structure are difficult to perform at present. The quality of simulation-based models for the BR photocycle and proton translocation mechanism [2] will depend on the accuracy of the structure and dynamics obtained in the simulations.

In the present work we have constructed a structural model for the BR monomer starting from the electron microscopy coordinates with added loops and water molecules. The model structure is subjected to a molecular dynamics simulation of 100 ps. We examine the accuracy of the simulation at a resolution similar to that achieved experimentally: a detailed analysis of the high-resolution local structure is not made as this contains considerable uncertainties in both the experiments and the simulation. The intrinsic water molecule positions are compared in projection with the exchangeable hydrogen contrast density seen in neutron diffraction studies [3] and found to be in good agreement. Deviations of the coarse-grained simulated structure from that derived from the electron microscopy experiment are described. We also examined large-amplitude low-frequency collective motions in the protein. These are accessible using inelastic neutron scattering spectroscopy [4]. Due to their delocalised nature these vibrations are likely to be relatively insensitive to detailed structural aspects. We have derived the vibrational frequency distribution, $G(\omega)$ from neutron experiments and compared this with the corresponding simulation-derived quantity. Approximate agreement is seen in the overall form of $G(\omega)$. The simulations are used to investigate the vibrational contributions from the different helices.

Correspondence address: J.C. Smith, Service de Biophysique des Protéines et des Membranes, Département de Biologie Cellulaire et Moléculaire, C.E.-Saclay, 91191 Gif-sur-Yvette, France.

2. MATERIALS AND METHODS

2.1. *Modelling and simulation* The starting coordinates for the modelling procedure were those deposed in the Protein Data Bank [1]. The modelling and molecular dynamics were performed with the program CHARMM [5] using the version 21 force field.

The helix backbone geometry was taken from the experimental coordinates [1]. The modelling procedure involved positioning the helix side chains, placing water molecules and adding loops onto the interhelix regions. The side-chain positioning was performed with the Sparse Matrix Driven (SMD) method [6]. In this method combinations of discrete rotameric states are sampled. The rotamer combination with the lowest potential energy is chosen for the model. The SMD method has been shown to position side chains of highly refined proteins with an RMS deviation of 1.69 Å [6]. No detailed analysis is made here of the side-chain positions obtained; the SMD algorithm was used primarily to obtain a helix packing geometry of low energy for the simulations.

The interhelix loops were poorly defined in the electron microscopy analysis [1], and they are too long to be determined using modelling techniques alone. They were included in the simulations so as to approximately model their effect on the helix structure and dynamics. The partial changes on the loop atoms were divided by two to approximately model the electrostatic shielding by the solvent surrounding the membrane.

With a model for the protein atoms constructed, water molecules were positioned in the remaining unoccupied volume using a grid search method. The volume enclosed by the whole protein was divided up into 5.4×10^4 cells of side 0.5 Å. Each grid node was examined and if no protein atom was within 1.8 Å it was considered as a possible water site.

Molecular dynamics simulations were performed on the complete model system using an integration timestep of 2 fs. The nonpolar hydrogens were considered together with the heavy atoms to which they were attached in the 'extended atom' approximation [5]. Partly due to the relatively low resolution of the experimental structure, and probably also to the fact that the environment of protein was neglected, the starting model is of high free energy. The method of simulated annealing, in which the temperature of the system is increased and then decreased, was used to bring the system into a relatively stable region of phase space, associated with a low free energy. It was necessary to constrain those aspects of the geometry considered reliable during the annealing and the subsequent early phase of 300K equilibration. Harmonic constraints were applied to the helix α carbons, to the helix backbone dihedrals and to the intrahelical hydrogen bond distances. The simulated annealing consisted of a 5 ps heating phase to 800K followed by a reduction of the temperature of 50K per ps until 300K. After the simulated annealing the 300K structure was energy minimised. The energy minimised structure was heated to 300K during 5 ps and equilibrated for ~ 35 ps. During the equilibration the constraints were progressively lessened until, at the end of the equilibration no constraints remained.

The production dynamics was performed without velocity rescaling for 65 ps. The resulting kinetic energy fluctuation was 1.3%. The neutron-weighted vibrational density of states, $G(\omega)$ was derived from the dynamical trajectories by numerical Fourier transformation of the incoherent cross-section-weighted velocity autocorrelation function as described previously [7].

2.2. *Neutron experiments*

The neutron scattering experiments on purple membrane have been described [8]. In the present analysis $G(\omega)$ was derived from the inelastic scattering intensities by extrapolation towards $q = 0$ (where q is the wavevector transfer between the neutrons and the sample), following the method described in [9]. Data used in the present work were collected from a low-hydration purple membrane sample (0.02 gD_2O/gBR). The low-hydration sample contains less quasielastic scattering, facilitating the extraction of $G(\omega)$ from the data.

3. RESULTS

The production phase of the molecular dynamics simulation presents a structural and dynamical model that explores a stable region of configurational space. However, due partly to the uncertainties in the experimental geometry from which the simulation starting structure was derived, and partly to the approximations present in the simulation methodology and potential function, the region of configurational space sampled in the simulation is likely to differ from that sampled by BR in the purple membrane. To examine the accuracy of the global structure of the molecule obtained by simulation we compared the structure averaged over the production dynamics with the neutron diffraction data and the electron microscopy structure.

An estimation of 11 ± 2 exchangeable protons (i.e. 4 to 7 water molecules) in the channel of the protein was made in the neutron diffraction work [3]. Our initial

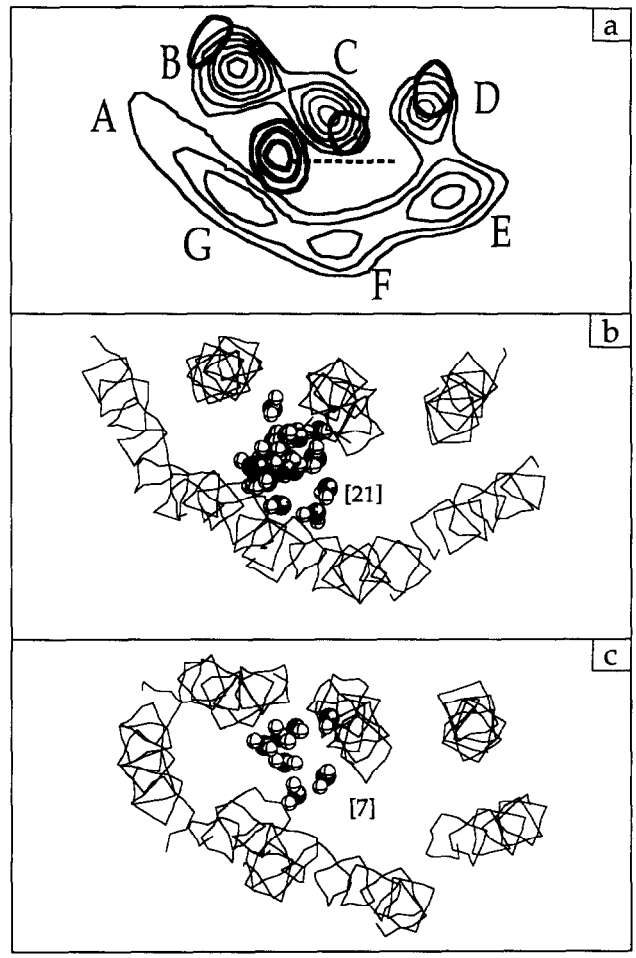

Fig. 1. Membrane-plane projections. (a) Two-dimensional difference Fourier map obtained by neutron diffraction [3] between oriented purple membranes at 15% relative humidity in ^2H_2O and H_2O. Bold contour lines represent 90%, 70% and 50% between-zero density level and the maximum positive level. The finer lines correspond to density areas of native BR. Position of retinal is indicated by a dotted line in the centre of helix bundle. (b) Helix backbone tracing obtained from the experimental electron microscopy structure with the 21 water molecule positions resulting from the grid search. (c) Helix backbone tracing of average structure from molecular dynamics simulation with the 7 water molecules remaining stable after equilibration.

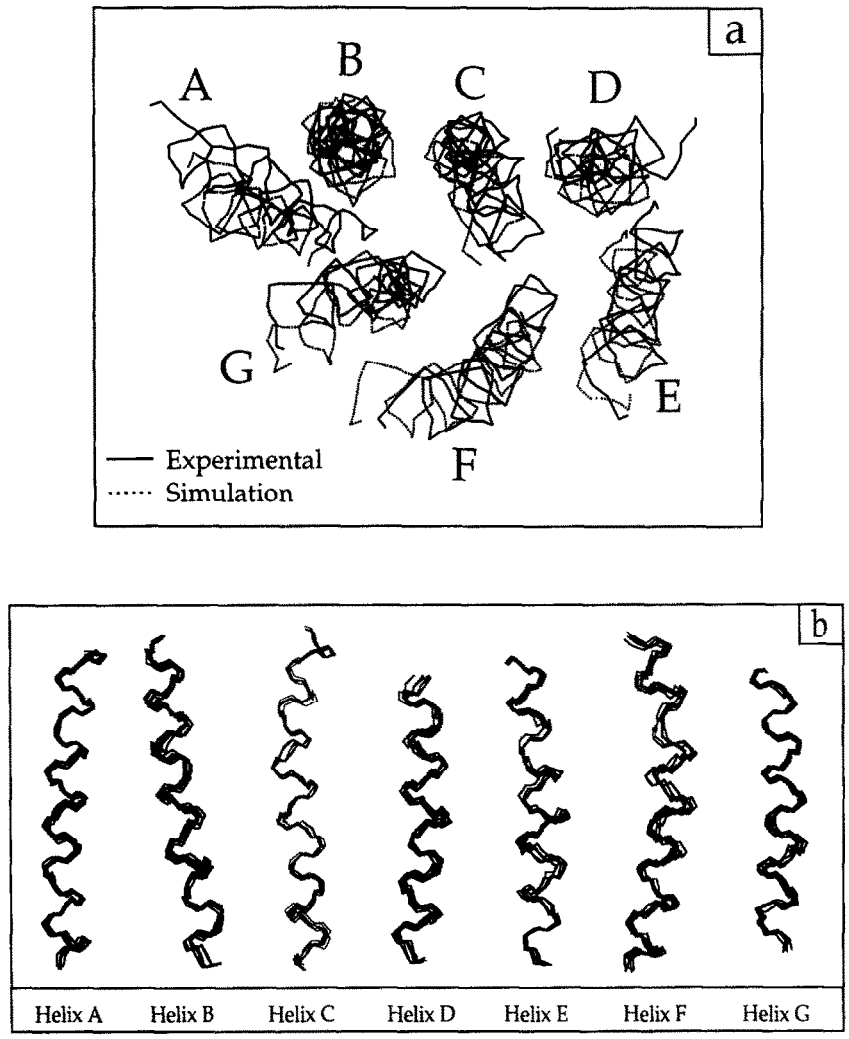

Fig. 2. (a) Projection in the membrane plane of the helix backbone of BR for the experimentally-derived structure (solid line) and for the average production dynamics structure (dotted line). (b) Superposed 10 ps-separated frames of the different helices obtained from the production dynamics phase.

water modelling, based on the unoccupied volume in the protein after side-chain positioning, resulted in the placing of 21 internal water molecules. The molecules were all located in the channel limited by helices B, C and G, 11 on the extracellular side and 10 on the cytoplasmic side. During the equilibration stages it became apparent that only 7 of the 21 water molecules were confined in the channel. The remaining 14 molecules were at the surface and exhibited correspondingly large fluctuations. Since our interest was confined to those water molecules playing a primary structural role, within the channel, the 14 surface molecules were removed. Two of the remaining water molecules were located in the cytoplasmic side of the channel between the Schiff base and ASP 96 and five were located in the extracellular side.

The low-resolution features of the water channel can be usefully compared with experiment. Fig. 1 shows a projection in the plane of the membrane of the neutron scattering contrast density [3] together with the water positions after the grid search and averaged over the simulation. The projection obtained from the grid search method is in excellent agreement with the experiment, although there are too many water molecules. The simulation-derived projection is in reasonable accord, a slight shift of the water molecules towards helix B being apparent.

The protein structure can be examined by comparing the average production dynamics structure with that derived by electron microscopy. The RMS deviation of the helix backbone (N, C_α and C) atoms of the average simulation structure from that derived experimentally was 2.15 Å. This deviation is significantly higher than the typical RMS deviations of 1.0–1.5 Å that are obtained in molecular dynamics simulations starting from high-resolution structures of water-soluble globular

proteins. In the present simulation the RMS deviation after equilibration and before production was 2.26 Å. The RMS deviation between the initial and final structures of the production dynamics is 0.52 Å. Thus, the majority of the deviation of the structure occurs during the heating and equilibration procedures of the simulation, as the system searches for a stable region of configurational space, of low free energy.

Fig. 2a shows a drawing of the average structure superposed on the electron microscopy structure in the projection parallel to the membrane surface. It can be seen that the global helix packing geometry is largely conserved. Some tilting of the helices was observed in the simulation, such that they became more perpendicular to the membrane surface. During the simulation the protein became slightly more spherical due to shifts of helices A, D and E. Fig. 2b shows superposed helix snapshots (every 10 ps) during the simulation. It can be seen that the internal helix structures are well preserved.

Fig. 3 shows the RMS fluctuations from their mean positions of the α-carbon atoms along the chain. The fluctuations are reduced in the helical regions (0.48 Å) compared with the loops (0.71 Å). A characteristic property of the helix fluctuations is that they increase at the ends of the helices. This has been analysed in myoglobin simulations where it was found to be due to rigid-body rotational motions of the helices perpendicular to their axes [10].

Further analysis and comparison with the experiment is possible by examining the frequency dependence of the vibrational motions in the molecule. Fig. 4 compares the experimentally-derived and simulation-derived $G(\omega)$s in the frequency range 0–200 cm^{-1}. In this range the largest amplitude, collective vibrations of the protein are expected to occur. Both simulation and experiment show a similar overall form of $G(\omega)$. The absence of intensity at higher frequencies in the simulation-derived plot might be due in part to the use of the extended-atom approximation in the simulation; this reduces the number of vibrations at higher frequencies [11]. At very low frequencies (< 15 cm^{-1}) the experimental $G(\omega)$ approximates to a Debye-like form ($\alpha\omega^2$) indicating the presence of large-amplitude vibrations describable by continuum dynamics. This behaviour is apparent over a less extended range (0–5 cm^{-1}) in the simulation-derived plot, due possibly to the absence of an explicit environment in the simulations. At higher frequencies the simulation-derived plot is clearly more structured than that observed experimentally. This is partly an effect of instrumental resolution broadening that is not present in the simulation but may also be due to an absence of environmental frictional damping effects on the simulated system. Fig. 5 shows $G(\omega)$ at frequencies < 20 cm^{-1} for each helix separately. All helices present a peak at 7 cm^{-1}; at higher frequencies differences exist. Helices B and C, implicated in retinal binding and proton transfer, have very similar frequency distributions.

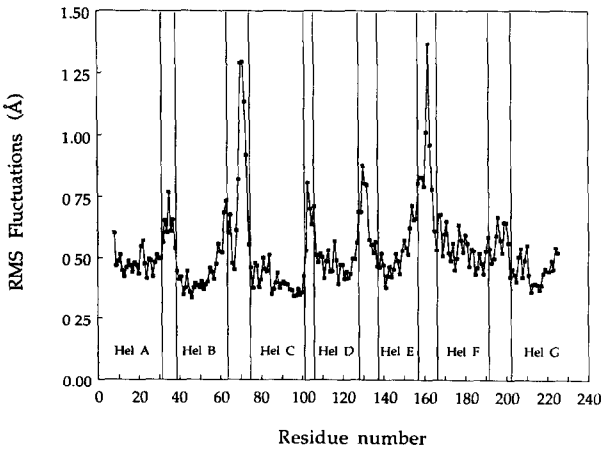

Fig. 3. RMS α-carbon fluctuations along the polypeptide chain during the production dynamics phase.

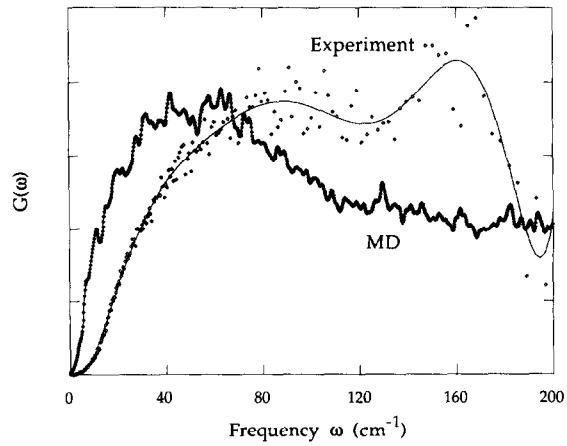

Fig. 4. Vibrational density-of-states $G(\omega)$ derived from neutron scattering experiments on purple membranes at a hydration of 0.02 g^2H_2O/gBR. Measurements were performed on a time-of-flight spectrometer (IN6) at the Institut Laue Langevin, Grenoble. The same quantity calculated from the molecular dynamics simulation is presented.

4. CONCLUSIONS

A complete understanding of the proton pump mechanism in BR will require an accurate and detailed structural and dynamical description of the protein. The present results give an indication as to the degree of convergence of the simulations and experiments. Progress on the experimental side can be expected, for example from further refinement of the electron microscopy data and from neutron diffraction on specifically deuterated samples [12]. Improved simulations will be made possible by the presently rapidly increasing available computer power. In the near future this will allow the simulation of the protein with explicit environment.

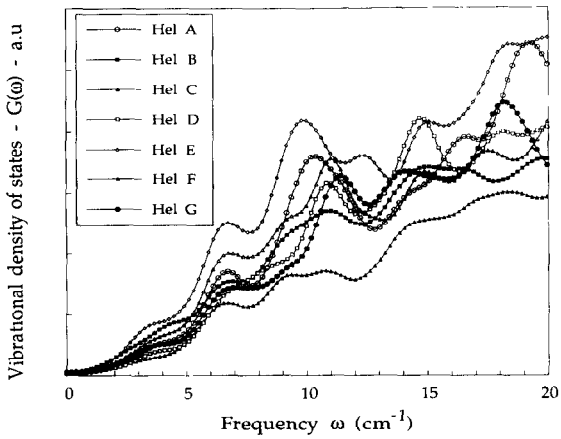

Fig. 5. $G(\omega)$ derived from the molecular dynamics simulation for individual transmembrane helices.

Improvements in the potential energy surface for the retinal and retinal–water interactions are also being made [13,14].

Inelastic neutron scattering experiments have shown that BR function is strongly correlated with the presence of unharmonic motions in the purple membrane [8]. These occur only in highly hydrated samples above 230K. The continued refinement of dynamical simulation models will be an important means of interpreting the experiments and will eventually provide a direct link between dynamics and function in BR. A large conformational change involving the tilting of two or three helices occurs in the photocycle of BR [15–17]. The time scale of this event may be much longer than that directly accessible to molecular dynamics simulation. However, low-frequency vibrations and/or diffusive motions in the protein must be characterised to enable a description of the dynamics of the proton transfer reactions.

Acknowledgements. We acknowledge Dr. G.R. Kneller for useful advice and for programs to calculate the vibrational density of states from the simulations. We acknowledge useful discussions with Drs. J.L. Popot and R. Lavery.

REFERENCES

[1] Henderson, R., Baldwin, J.M., Ceska, T.A., Zemlin, F., Beckmann, E. and Downing, K.H. (1990) J. Mol. Biol. 213, 899–929.
[2] Zhou, F., Windemuth, A. and Schulten, K. (1993) Biochemistry 32, 2291–2306.
[3] Papadopoulos, G., Dencher, N.A., Zaccaï, G. and Büldt, G.J. (1990) J. Mol. Biol. 214, 15–19.
[4] Smith, J.C. (1991) Quart. Rev. Biophys. 24 (3), 227–291.
[5] Brooks, B., Bruccoleri, R., Olafson, B., States, D., Swaminathan, S. and Karplus, M. (1983) J. Comp. Chem. 4, 187–217.
[6] Tuffery, P., Etchebest, C., Hazout, S. and Lavery, R. (1991) J. Biomol. Struct. Dyn. 8, 1267–1289.
[7] Kneller, G.R., Doster, W., Settles, M., Cusack, S. and Smith, J.C. (1992) J. Chem. Phys. 97 (12), 8864–8879.
[8] Ferrand, M., Dianoux, A.J., Petry, W. and Zaccaï, G. (1993) Proc. Natl. Acad. Sci. USA (in press).
[9] Cusack, S. and Doster, W. (1990) Biophys. J. 58, 243–251.
[10] Furois-Corbin, S., Smith, J.C. and Kneller, G.R. (1993) Proteins: Struct. Funct. Genetics 16, 141–154.
[11] Smith, J.C., Cusack, S., Tidor, B. and Karplus, M. (1990) J. Chem. Phys. 93, 2974–2991.
[12] Samatey, F., Popot, J.L., Etchebest, C. and Zaccaï, G. (1992) in: Structure and Function of Retinal Proteins (Rigaud, J.L., Ed.) Colloque INSERM 221, pp. 9–12.
[13] Nina, M., Roux, B. and Smith, J.C. (1992) in: Structure and Function of Retinal Proteins (Rigaud, J.L., Ed.) Colloque INSERM 221, pp. 17–20.
[14] Nina, M., Smith, J.C. and Roux, B. (1993) J. Mol. Struct. (THEOCHEM), in press.
[15] Dencher, N.A., Dresselhaus, D., Zaccaï, G. and Büldt, G. (1989) Proc. Natl. Acad. Sci. USA 85, 6358–6361.
[16] Koch, M.H.J., Dencher, N.A., Oesterhelt, D., Plöhn, H.J., Rapp, G. and Büldt, G. (1991) EMBO J. 10, 521–526.
[17] Subramanian, S., Gerstein, M., Oesterhelt, D. and Henderson, R. (1993) EMBO J. 12, 1–8.