Molecular dynamics simulation of Pf1 Coat Protein

Douglas J. Tobias, Michael L. Klein, and Stanley J. Opella
Department of Chemistry, University of Pennsylvania, Philadelphia, Pennsylvania 19104-6323

ABSTRACT The results of molecular dynamics simulations of Pf1 coat protein are described and compared to experimental NMR data on both the membrane bound and structural forms of this viral coat protein. Molecular dynamics simulations of the 46 residue coat protein and related model sequences were performed according to a simple protocol. The simulations were initiated with the polypeptides in a completely uniform α helical conformation in a dielectric continuum ($\epsilon = 2$) and the motions of individual residues were followed as a function of time by monitoring the angular fluctuations of amide NH bond vectors. The simulations of Pf1 coat protein were able to identify the same mobile and structured segments found in experimental NMR studies of the membrane bound form of the protein (Shon, K.-J., Y. Kim, L.A. Colnago, and S.J. Opella. 1991. *Science (Wash. DC)*. 252:1303–1305). Significantly, in addition to mobile amino and carboxyl terminal regions, a mobile internal loop was found that connects the rigid hydrophobic and amphipathic helices in the protein. NMR experiments show that this mobile loop is present in both the viral and membrane bound forms of the protein and that it plays a role in viral assembly (Nambudripad, R., W. Stark, S.J. Opella, and L. Makowski. 1991. *Science (Wash. DC)* 252:1305–1308). The results of simulations of several alanine based 46 residue polypeptides with some of the charged residues present in the Pf1 coat protein sequence suggest that interactions between the Asp 14 and Asp 18 sidechains and the peptide backbone are responsible for the formation of the mobile loop. The agreement between the results of the calculations presented here and the previously reported NMR experiments suggest that molecular dynamics simulations might be useful in the prediction of the secondary structure and dynamics of individual residues in membrane and structural proteins with predominantly α helical secondary structure.

INTRODUCTION

The membrane bound form of the coat protein of the filamentous bacteriophage Pf1, like other membrane proteins, has both rigidly structured and mobile segments in its polypeptide backbone (Shon et al., 1991). One of the mobile segments is also present in the structural form of the coat protein in virus particles (Nambudripad et al., 1991). Since the structures of many membrane proteins are dominated by long hydrophobic helices spanning the bilayer and shorter amphipathic helices in the plane of the bilayer (Deisenhofer et al., 1985; Rees et al., 1989; Henderson et al., 1990), the mobile connecting loops and terminal regions appear to be key elements of their overall architecture (Cross and Opella, 1980; Keniry et al., 1984; Leo et al., 1987; Bowers and Oldfield, 1988). The juxtaposition of mobile and structured regions of both membrane associated and structural proteins means that experimental and theoretical investigations of the dynamics of non-globular proteins, in general, are likely to be highly informative as part of structure determinations and other investigations.

Classical molecular dynamics (MD) simulations, which consist of numerically solving Newton's equations of motion for all of the atoms in the system, have profoundly influenced our understanding of the structure, dynamics, and functions of globular proteins (Brooks et al., 1988). However, membrane associated and structural proteins are not nearly as well understood as globular proteins, largely because the experimental and theoretical methods of structural biology were developed with globular proteins in mind, and these methods are not directly applicable to other classes of proteins. Therefore, even qualitative information from credible MD simulations is likely to be valuable, since so few experimental results are available that describe the structure and dynamics of non-globular proteins. While the methods for MD simulations of globular proteins *in vacuo* and in aqueous solution are now fairly well established (Brooks et al., 1985), the development of techniques for simulating the components of biological membranes are still in their infancy (Egberts and Berendsen, 1988; Edholm and Jahnig, 1988; DeLoof et al., 1991; Berkowitz and Raghavan, 1991; Sanders et al., 1991).

The membrane bound form of the coat protein of the filamentous bacteriophage Pf1 provides a tractable model system for molecular dynamics simulations of a membrane protein. High resolution solid-state NMR and multidimensional solution NMR studies of this small 46 residue protein (Rowitch and Perham, 1987) in the model membrane environments of phospholipid bilayers and detergent micelles show it to be surprisingly complex with five distinct structural regions (Shon et al., 1991). The secondary structure consists of a long hydrophobic helix (residues 19 to 42) that spans the bilayer and a relatively short amphipathic helix (residues 6 to 13) in the plane of the bilayer. The amino terminus (residues 1 to 5), the carboxyl terminus (residues 43 to 46), and residues 14 to 18 connecting the two helices are mobile. All of these features are characteristic of membrane bound proteins, identifying Pf1 coat protein as a typical membrane protein in terms of both its structure and dynamics. Pf1 coat protein also exists as the major structural element of virus particles and this form of the protein has been analyzed by solid-state NMR spectroscopy (Cross et al., 1983; Tsang and Opella, 1986; Shon et al.,

Address correspondence to Stanley Opella.

0006-3495/93/03/670/06 $2.00

1991), neutron diffraction (Nambudripad et al., 1991), and x-ray diffraction (Makowski et al., 1980). The MD simulations give insight into the mechanism of interconversion of these two forms of the protein.

METHODS

The dynamics of several 46 residue polypeptides with an initial structure of a single continuous α helix were simulated. One of the polypeptides had the amino acid sequence of Pf1 coat protein (Rowitch and Perham, 1987). Several independent (i.e., different initial conditions) 500–1,800 ps long MD simulations of the Pf1 coat protein sequence in a dielectric continuum were performed. The dielectric constant $\epsilon = 2$ was chosen because it is approximately the value in the hydrocarbon region of a bilayer membrane (Ashcroft et al., 1981) and because it slightly reduces electrostatic (e.g., hydrogen bonds, charge-charge, and charge-dipole) interactions. 500–1,000 ps simulations of several 46 residue alanine-rich peptides were performed for comparison. Polyalanine was simulated as a control as well as to investigate the roles of electrostatic interactions. In addition, simulations were performed on the following peptides: polyalanine with all the charged residues of the Pf1 coat protein (Asp 4, Glu 9, Asp 14, Asp 18, Lys 20, Arg 44, and Lys 45), the same peptide with Asp 14 and Asp 18 replaced by alanines, polyalanine with a lysine at residue 20, and polyalanine with aspartic acid at residues 14 and 18. Simulations of the latter peptide were also carried out with $\epsilon = 10$ and 20 to investigate the effects of dielectric screening on the results.

All of the simulations were started from α helical structures (with the backbone dihedral angles, $\phi = -60°$, $\psi = -40°$). The equations of motion were integrated using the Verlet algorithm (1967) with a timestep of 0.001 ps and the CHARMM potential energy function (Brooks et al., 1983) and PARAM19 peptide parameters (Reiher, 1985). No experimental restraints were used in the calculations. The velocities were occasionally reassigned to maintain temperatures of ~25°C. The nonbonded energies and forces, which were smoothly truncated at 8 Å with shifting functions (Brooks et al., 1983), were processed using a list-based algorithm (Verlet, 1967), and the list was updated every ten timesteps. The SHAKE algorithm (Ryckaert et al., 1977) was used to fix the lengths of NH and OH bonds throughout the simulations. The independent simulations of the Pf1 coat protein sequence were started from the same configurations but with different initial velocities. All of the simulations were performed using the CHARMM program (Brooks et al., 1983) on a Convex C210 computer.

RESULTS AND DISCUSSION

In all of the Pf1 coat protein simulations, the starting structure of a continuous α helix evolved into five distinct regions in remarkable qualitative agreement with NMR experimental results. Two rigid helices are separated by a mobile segment and both ends of the protein are mobile. In contrast, the α helix in polyalanine remained intact throughout the simulation. These findings are illustrated by the structures shown in Fig. 1.

Fig. 2 presents plots of the angular fluctuations of N-H bond vectors at representative residues of Pf1 coat protein during the initial 200 ps of a 500-ps simulation. These bond vectors reflect the motions of individual residues and can be directly compared to the experimental results from relaxation measurements, since ^{15}N nuclear spin relaxation results from fluctuations of the $^{1}H/^{15}N$ heteronuclear dipole-dipole interactions described by this same vector. They should also be closely related to the motional averaging effects observed with the chemical shift anisotropy at each amide group, since of the principal elements of the ^{15}N amide chemical shift tensor are aligned approximately along this vector. The plot in Fig. 2 *d* for Ala 36, which is representative of the residues in the hydrophobic helix, shows only small oscillations. All of the residues in polyalanine give plots very similar to that in Fig. 2 *d*. The plot in Fig. 2 *b* for Thr 13 in the amphipathic helix shows that this residue experiences oscillations with only slightly larger amplitude than those for residues in the hydrophobic helix of Pf1 coat protein and any of the residues of polyalanine; it undergoes a slow drift, apparently tracking the change in orientation of the amphipathic helix with respect to the hydrophobic helix. The plots in Fig. 2 *a* (Val 2) and Fig. 2 *e* (Lys 45) show the large amplitude angular fluctuations that occur in the penultimate residues of the polypeptide, indicating that the ends of the peptide are quite mobile. Fig. 2 *c* is the plot for Gly 17 in the loop region connecting the two helices; while its fluctuations are not as dramatic as those at the ends of the peptide, they are substantially larger than those for residues in either of the helices. It is significant that the angular drift for this residue does not appear to follow the reorientation of the amphipathic helix.

Fig. 3 compares the time evolution of the angle between the two helices in Pf1 coat protein from an 1,800-ps simulation to that for the two halves of polyalanine. There is a rapid increase of the angle in the first 1,000 ps of the simulation of the coat protein, followed by a slow rise. The apparent "sawtooth" structure in the curve suggests that there is a faster (100 ps) timescale inter-helical motion superimposed on the slow, overall inter-helical movement. Molecular graphics analysis revealed this faster motion to be due to the coupling between the amphipathic helix motion and a bending mode in the hydrophobic helix. This coupling appears to be transmitted via strong interactions between the Asp 18 sidechain and the hydrophobic helix backbone and between the Lys 20 sidechain and the carbonyl groups of Ile 12 and Thr 13 in the amphipathic helix. Similar behavior was exhibited in other simulations of the coat protein, although the timescale for the initial approach to the plateau value and the plateau value itself differed somewhat. The angle between the top and bottom halves of polyalanine shows that the helix was undergoing only small amplitude bending motions, remaining essentially straight on the 1,000-ps timescale.

The overall inter-helical movement is an example of a functionally important inter-domain motion in a protein, analogous to the hinge-bending motion in lysozyme (McCammon et al., 1976), resulting from a flexible loop or hinge composed of a few residues, connecting two structured, rigid domains. In lysozyme the inter-domain motion appears to be involved in the closing (opening) of the active site cleft upon binding (releasing) the substrate. In Pf1 coat protein the inter-domain

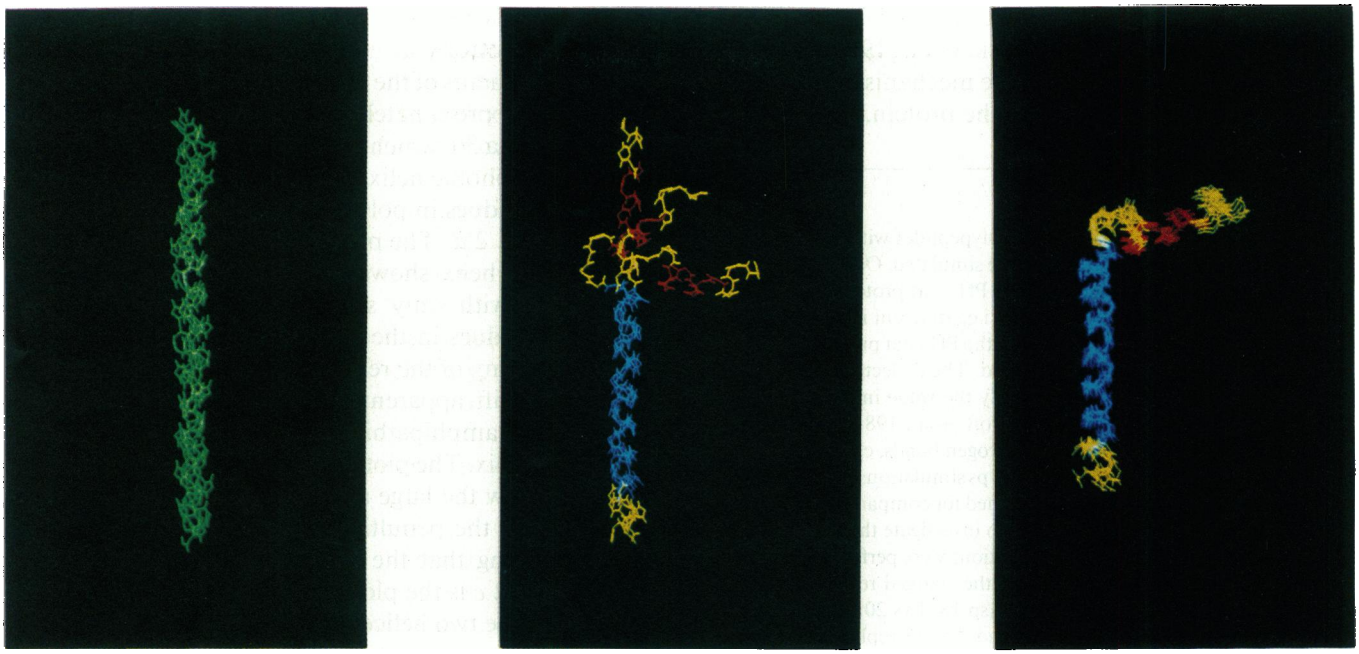
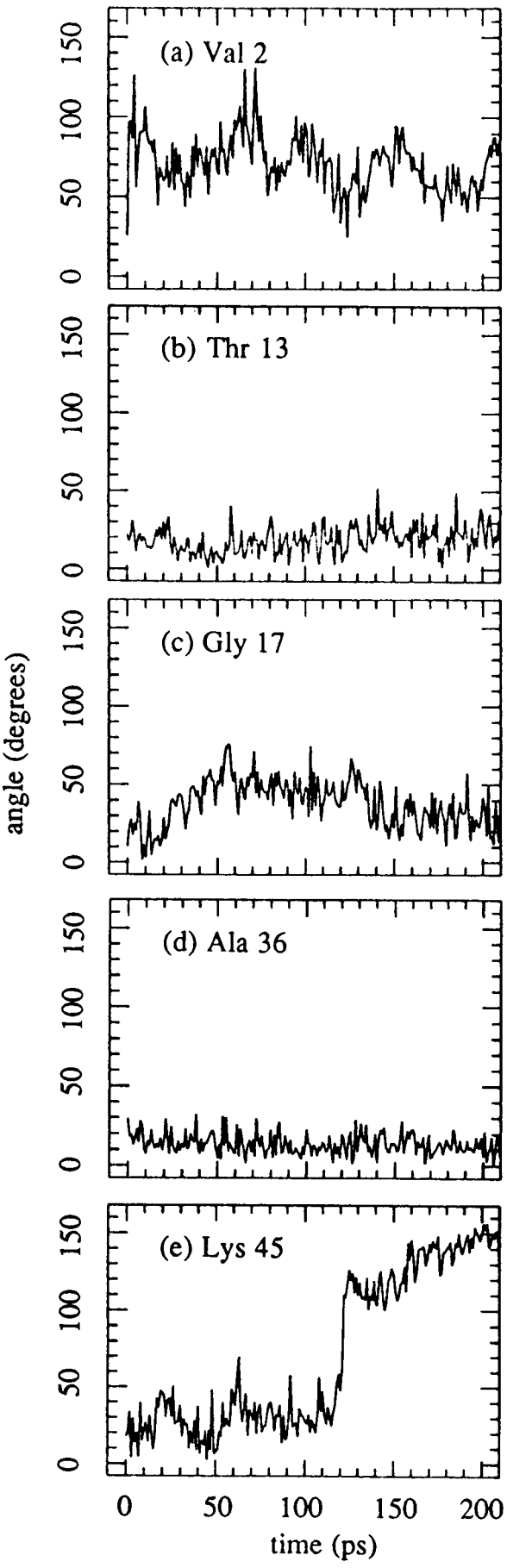


FIGURE 1 Backbone atoms of 46 residue polypeptides at various times during the molecular dynamics simulations. (*Left*) Structures from a simulation of the polyalanine peptide at 0, 450, and 900 ps. (*Middle*) Structures from a simulation of the Pf1 coat protein at 0, 450, and 900 ps. (*Right*) Six structures from a simulation of the Pf1 coat protein at 10-ps intervals starting at 900 ps. The coloring in the middle and right figures denotes the distinct regions identified by NMR experiments (Shon et al., 1991); "rigid" hydrophobic (*blue*) and amphipathic (*red*) helices, and mobile ends and connecting loops (*yellow*). The structures were drawn using the MolX program (Sneddon, 1991).

motion has an important role in the viral assembly process where there is a transition from the membrane bound form of the protein with an inter-helical angle of approximately 90° to the form in the viral coat where the angle is ~15° (Nambudripad et al., 1991).

We performed 500-ps simulations of several initially α helical alanine-rich peptides in order to investigate the possible role of electrostatic interactions in the formation of the mobile loop between the two helices in our model of the Pf1 coat protein. The first was simulation of a 46 residue polyalanine peptide ("I") with seven of the alanine residues replaced by the corresponding charged residues of the Pf1 coat protein sequence (Asp 4, Glu 9, Asp 14, Asp 18, Lys 20, Arg 44, and Lys 45). During the course of the simulation of polypeptide I, the terminal regions of the polypeptide were mobile and a mobile internal loop formed while the remainder of the polypeptide remained helical with an inter-helical angle of ~90°, as in the simulations of wild-type Pf1 coat protein. Polypeptide II differed from polypeptide I in having Asp 14 and Asp 18 replaced by alanine residues. After 500 ps of simulation polypeptide II displayed a slight kink between the helices, with an inter-helical angle of ~10°, in the region of the mobile loop in the native Pf1 coat protein, but a mobile loop did not develop. The kink appeared to be caused by a strong interaction between the sidechain of Lys 20 and the peptide backbone. To test this hypothesis a 46 residue polyalanine peptide ("III") with only Ala 20 replaced by Lys 20 was simulated and a kink almost identical to that observed in polypeptide II was formed. This suggested that Lys 20 could cause the formation of a slight kink, but Asp 14 and Asp 18 were necessary for the formation of the mobile loop. This conclusion was supported by simulation of a 46 residue polyalanine peptide ("IV"), with Ala 14 and Ala 18 replaced by Asp 14 and Asp 18, in which a mobile loop did form in the correct location with the remainder of the peptide remaining helical. Molecular graphics analysis suggested that the loop formation was initiated by insertion of the Asp sidechains into hydrogen bonds of the initial helix. A simulation of a polyalanine peptide with only Ala 14 replaced by Asp 14 ("V") indicates that a single Asp is capable of disrupting a helix at one hydrogen bond, but can not cause loop formation. These findings suggest that two nearby Asp are responsible for formation of the mobile loop, regardless of whether or not they are in phase, i.e., at the i and $i + 4$ positions in the sequence. To examine the effects of dielectric screening on the ability of Asp 14 and Asp 18 to cause loop formation two additional simulations of polypeptide IV were performed with higher dielectric constants (ϵ = 10, 20). A mobile loop was formed in the simulation with ϵ = 10 but not in the simulation with ϵ = 20.

Taken together, the results of the various simulations of polypeptides I–V show that electrostatic interactions between the sidechains of Asp 14 and Asp 18 and the peptide backbone are sufficient to cause mobile loop formation in Pf1 coat protein, but only if the local dielectric constant is low (i.e., ϵ = 2 − 10). We cannot say that



these interactions are necessary for kink formation because our model does not contain explicit solvent or lipid molecules. Since the mobile loop is present in both the membrane bound and viral coat forms (Shon et al., 1991), we can rule out the role of lipid molecules for mobile loop formation, although water molecules could accomplish the same task as the Asp sidechains. With regard to the role of the charged residues in disrupting α helices in membrane bound proteins, it is worth pointing out that the analysis of the simulation of peptide I suggested that interactions of the Asp 4 but not the Glu 9 sidechain contribute to the mobility of the NH_2-terminal region.

The results for Pf1 coat protein show that vacuum MD simulations can provide information directly relevant to the structure and dynamics of both membrane bound and structural proteins. The simulations correctly identified the locations of the structured helical regions, mobile internal loop, and mobile termini in the protein. Furthermore, the relative mobility of the residues in the helical and loop regions were reproduced qualitatively in the simulations. Previously reported short (100 ps) simulations of membrane associated polypeptides of similar size displayed fluctuations that resulted in temporary tilts and bends of the helical backbone structure (Edholm and Jahnig, 1988; Sanders et al., 1991). The agreement between the results of our calculations and NMR experiments suggests that a simple MD simulation protocol might be useful in the prediction of the secondary structure and dynamics of membrane proteins.

There are some quantitative features of the protein that apparently cannot be reproduced by simulations in a dielectric continuum. In particular, some of the independent simulations did not reproduce the expected relative orientations of the two helices in the Pf1 coat protein. In four independent simulations, the inter-helical angle (Fig. 3) exhibited variable behavior: the angle converged to $\sim 50°$ after ~ 500 ps in two 600-ps simulations and approached 100–110° by the end of 500 and 1,800-ps simulations. Furthermore, in all of the simulations, the amphipathic helix was generally oriented so that the polar and nonpolar faces were perpendicular to the plane containing the two helices, while we expect that the two faces are actually in the plane, with the nonpolar face buried in the membrane and the polar face sticking into the solvent above the membrane surface. It is likely that we need to simulate the protein in an amphipathic environment (bilayer membrane with aqueous solvent) to reproduce these features. The presence of the polar aqueous solvent above the membrane, the polar headgroups at the membrane surface, and the nonpolar chains in the membrane will help to correctly orient the

FIGURE 2 Time evolution of the angle between the axis of the hydrophobic helix (residues 22-43) and the NH bond vector of (*a*) Val 2, (*b*) Thr 13, (*c*) Gly 17, (*d*) Ala 36, and (*e*) Lys 45, from a simulation of the Pf1 coat protein.

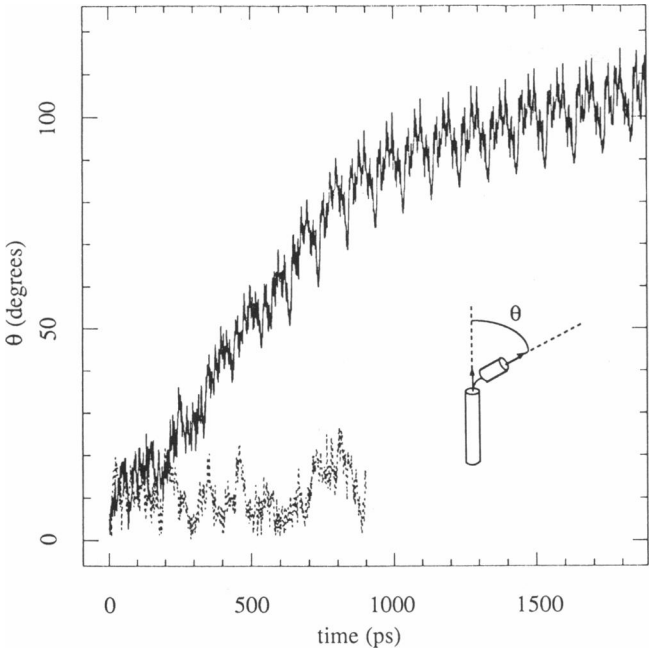


FIGURE 3 Time evolution of the interhelical angle from an 1,800-ps simulation of the Pf1 coat protein (*solid curve*) and a 46 residue alanine polypeptide (*dashed curve*). We defined the interhelical angle as the angle between the axes of two helices, one containing residues 7–14 and the other containing residues 22–42 (these residue ranges span most of the amphipathic and hydrophobic helices, respectively, in the Pf1 coat protein). The helix axis is taken as the unit vector along the axis of a cylinder fit to the α carbon atoms of the specified residues.

amphipathic helix at the membrane surface. With the amphipathic helix in the proper orientation, the sidechains on the nonpolar face can participate in favorable van der Waals interactions with the lipid hydrocarbon chains, while the sidechains on the polar face can be favorably solvated by the polar region of the membrane and/or water. Finally, the polar lipid headgroups will not only help stabilize the membrane/water interface, but will also help anchor the protein in the membrane by interacting with the backbone and with charged protein sidechains located near the membrane surface.

CONCLUSIONS

Molecular dynamics simulations of the 46 residue Pf1 coat protein in a dielectric continuum ($\epsilon = 2$), starting with an α helical structure, were able to identify the same mobile and structured segments found in experimental NMR studies of the membrane bound form of the protein. The simulations correctly identified the locations of the rigid hydrophobic and amphipathic helices as well as the mobile terminal residues and mobile rigid loop connecting the two helices. Simulations of several alanine-rich peptides with some or all of the charged residues of the Pf1 coat protein suggest that interactions between Asp 14 and Asp 18 sidechains and the peptide backbone are responsible for the formation of the mobile loop. This physical picture should be viewed with some caution, however, since the interactions of the protein with the solvent may also cause loop formation. Comparisons of simulations and experimental results are needed before any rules for predicting the disruptive effects of charged sidechains on a helices can be proposed. However, our simulations certainly do suggest that local interactions of specific residues are responsible for determining many aspects of the structure and dynamics of Pf1 coat protein. Therefore, the simple MD simulation protocol employed in this study may prove useful in describing structural and dynamic features of other membrane proteins. Indeed, preliminary results from our laboratories, to be reported elsewhere, reveal good agreement between simulations and NMR experiments on the structure and dynamics of the coat protein and procoat protein from fd and M13 bacteriophages.

It is a pleasure to acknowledge the many valuable direct and indirect contributions of Harden McConnell to the research described in this paper. Chief among them are his role as mentor in the graduate education of one of the authors (S. Opella) and the formulation of biophysical research on membranes, especially focussing attention on the importance of the dynamics of the molecular constituents.

This research is supported by grants from the National Institutes of Health to M. Klein (R01 GM40712), S. Opella (R01, A120770; R01 GM29754; R01 GM24266), and a postdoctoral fellowship to D. Tobias (F32 GM14463).

Received for publication 18 September and in final form 12 November 1992.

REFERENCES

Ashcroft, R. G., H. J. L. Coster, and J. R. Smith. 1981. The molecular organisation of bimolecular lipid membranes. *Biochim. Biophys. Acta.* 643:191–204.

Berkowitz, M. L., and K. Raghavan. 1991. Computer simulation of a water/membrane interface. *Langmuir.* 7:1042–1044.

Bowers, J. L., and E. Oldfield. 1988. Quantitative carbon 13 nuclear magnetic resonance spectroscopy study of mobile residues in bacteriorhodopsin. *Biochemistry.* 27:5156–5161.

Brooks, B. R., R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus. 1983. CHARMM: a program for macromolecular energy, minimization and dynamics calculations. *J. Comput. Chem.* 4:187–217.

Brooks III, C. L., M. Karplus, and B. M. Pettitt. 1988. Proteins: a theoretical perspective of dynamics, structure, and thermodynamics. *Adv. Chem. Phys.* 71:1–259.

Cross, T. A., and S. J. Opella. 1980. Structural properties of fd coat protein in sodium dodecyl sulfate micelles. *Biochem. Biophys. Res. Commun.* 92:478–484.

Cross, T. A., P. Tsang, and S. J. Opella. 1983. Comparison of protein and DNA structure in fd and Pf1 Bacteriophages. *Biochemistry.* 22:721–726.

Deisenhofer, J., O. Epp, K. Miki, R. Huber, and H. Michel. 1985. Structure of the protein subunits in the photosynthetic reaction center of rhodopseudomonas viridis at 3A resolution. *Nature (Lond.).* 318:618–624.

DeLoof, H., S. C. Harvey, J. P. Segrest, and R. W. Pastor. 1991. Mean

field stochastic boundary molecular dynamics simulation of a phospholipid in a membrane. *Biochemistry.* 30:2099–2113.

Egberts, E., and H. J. C. Berendsen. 1988. Molecular dynamics simulation of a smectic liquid crystal with atomic detail. *J. Chem. Phys.* 89:3718–3732.

Henderson, R., J. M. Baldwin, T. A. Ceska, N. Zemlin, E. Beckmann, and K. H. Downing. 1990. Model for the structure of bateriorhodopsin based on high-resolution electron cryomicroscopy. *J. Mol. Biol.* 213:899–929.

Keniry, M. A., H. S. Gutowsky, and E. Oldfield. 1984. Surface dynamics of the integral membrane protein bacteriorhodopsin. *Nature (Lond.).* 307:383–386.

Makowski, L., D. L. D. Caspar, and D. A. Marvin. 1980. Filamentous bacteriophage Pf1 structure determined at 7A resolution by refinement of models for the α-helical subunit. *J. Mol. Biol.* 140:149–181.

McCammon, J. A., B. R. Gelin, M. Karplus, and P. G. Wolynes. 1976. The hinge-bending mode in lysozyme. *Nature (Lond.).* 262:325–326.

Nambudripad, R., W. Stark, S. J. Opella, and L. Makowski. 1991. Membrane mediated assembly of filamentous bacteriophage Pf1. *Science (Wash. DC).* 252:1305–1308.

Rees, D. C., H. Komiya, T. O. Yeates, J. P. Allen, and G. Feher. 1989. The photosynthetic reaction center as a model for membrane proteins. *Annu. Rev. Biochem.* 58:607–633.

Reiher, W. E. 1985. Ph.D. Thesis, Harvard University, Cambridge, MA.

Rowitch, D. H., and R. N. Perham. 1987. Cloning and expression of the filamentous bacteriophage Pf1 major coat protein gene in Escherichia coli. *J. Mol. Biol.* 195:873–884.

Ryckaert, J. P., G. Ciccotti, and H. J. C. Berendsen. 1977. Numerical integration of Cartesian equations of motion of a system with contraints-molecular dynamics of n-alkanes. *J. Comput. Phys.* 23:327–341.

Shon, K. J., Y. Kim, L. A. Colnago, and S. J. Opella. 1991. NMR studies of the structure and dynamics of membrane bound bacteriophage Pf1 coat protein. *Science (Wash DC).* 252:1303–1305.

Sneddon, S. F. 1991. *MolX Series of Molecular Graphics Programs.* Pfizer Central Research, Groton, CT.

Tsang, P., and S. J. Opella. 1986. Pf1 virus particle dynamics. *Biopolymers.* 25:1859–1864.

Verlet, L. 1967. Computer "experiments" on classical fluids. I. Thermodynamical properties of Lennard-Jones molecules. *Phys. Rev.* 159:98–103.