Structure *18*

Supplemental Information

PagP Crystallized from SDS/Cosolvent Reveals the Route for Phospholipid Access to the Hydrocarbon Ruler

Jose Antonio Cuesta-Seijo, Chris Neale, M. Adil Khan, Joel Moktar, Christopher Tran, Russell E Bishop, Régis Pomès, and Gilbert G. Privé

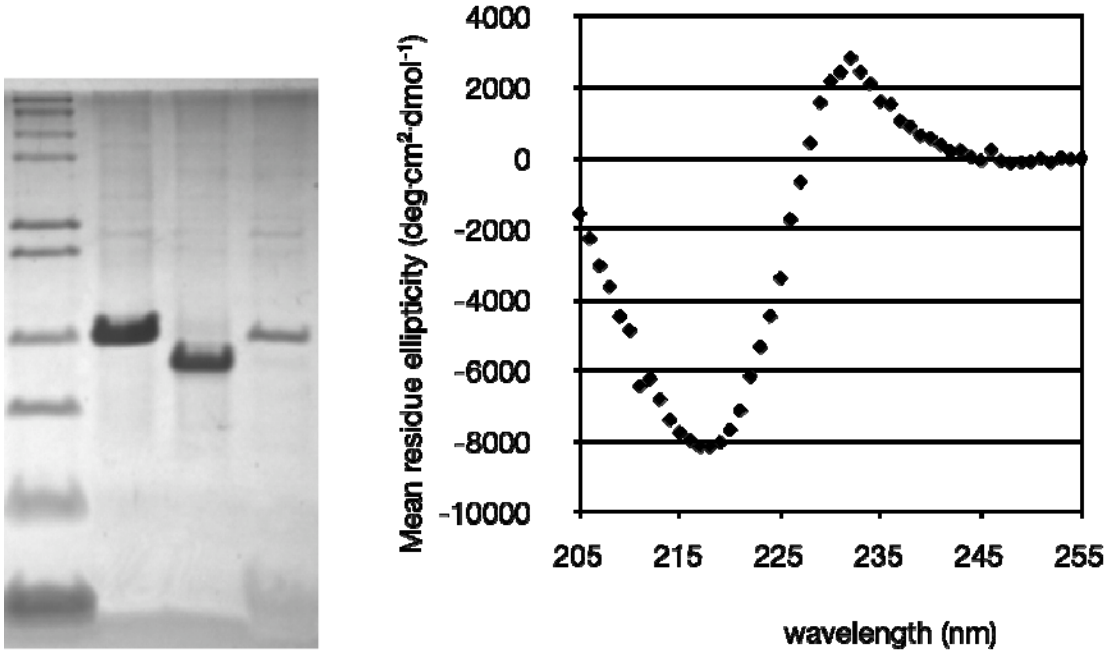

Figure S1, related to Figure 2. Refolding of PagP in SDS/MPD

Left. 18% SDS-PAGE of the protein crystallization stock. From left to right, Mark 12 molecular weight standards, 1 μL of the crystallization protein stock, 5 μg of PagP unfolded in SDS alone and 2 μg of PagP folded in LDAO. The gel was stained with Simply Blue (Invtrogen).

Right. CD spectrum of the sample prior to the concentration step (see Methods). The maximum at 232 nm is indicative of the correct formation of the hydrocarbon ruler, the minimum at 218 nm of the β–sheet structure.

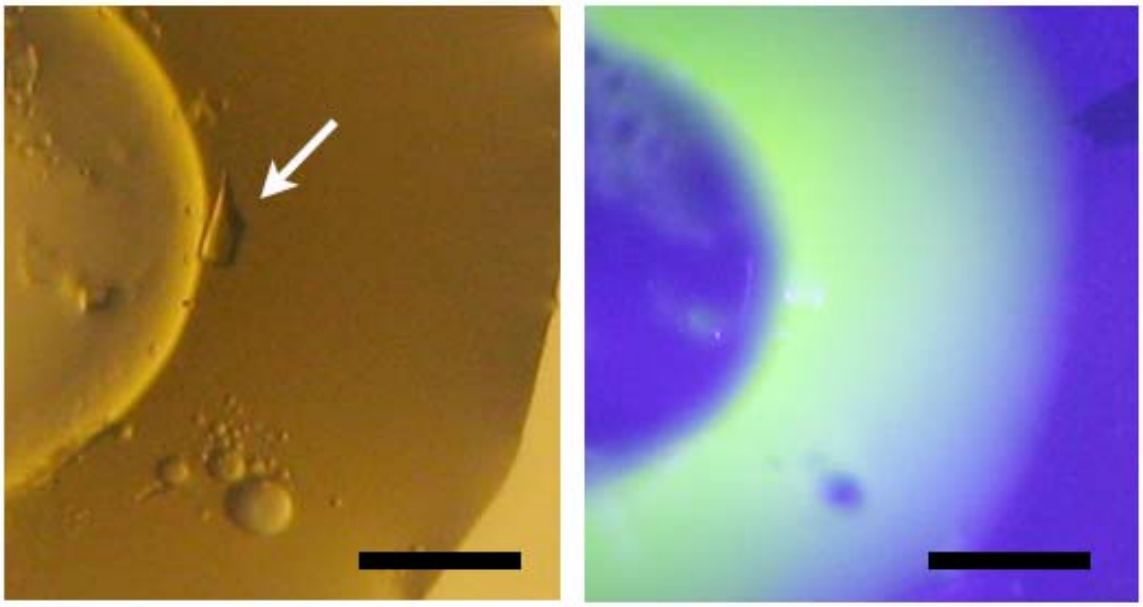

Figure S2, related to Figure 3. PagP crystals grow from a SDS-rich phase
Left: Phase separation and crystallization of PagP from SDS/MPD. The white arrow indicates a PagP crystal. Right: The same droplet after the addition of the fluorescent lipid NBD-DPPC and viewed under UV light. The lipid partitions into the phase that contains the crystal, as seen by the green fluorescence. The scale bar represents 100 μm.

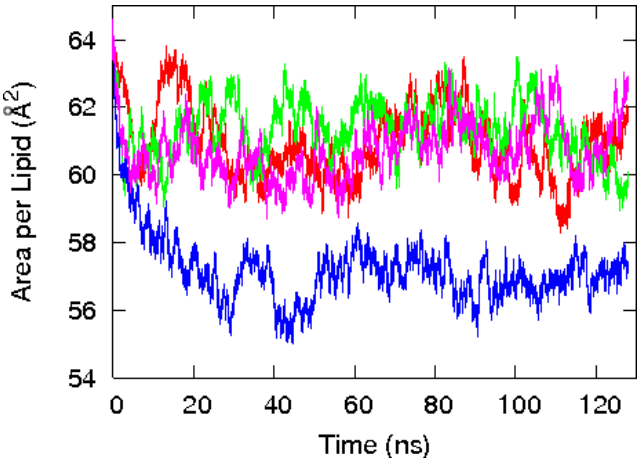

Figure S3, related to Figure 5.
The area per lipid time series at 323K for simulations employing the forcefield parameters of (red) standard ffgmx, (green) ffgmx converted to ε and δ, (blue) unmodified ε and δ using OPLSAA scaling rules, (purple) the half-ε double-pairlist method for combination with OPLSAA.

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

Molecular Dynamics Simulations

Simulation conditions. Simulations were conducted with version 3.3.1 of the GROMACS simulation package (Lindahl et al., 2001), modified to apply a biasing force of constant magnitude. The Berger parameters (Berger et al., 1997) for POPC (Tieleman et al., 1998) and the OPLSAA parameters (Kaminski et al., 2001) for PagP were combined using the half-ε double-pairlist method (Chakrabarti et al., 2010). The water model was TIP4P (Jorgensen et al., 1983). Periodic boundary conditions were enforced via a rectangular unit cell. Lennard–Jones interactions were evaluated using a group-based twin-range cutoff (van Gunsteren and Berendsen, 1990) calculated every step for separation distances less than 0.9 nm and every ten steps for distances between 0.9 and 1.4 nm, when the nonbonded list was updated. Coulomb interactions were calculated using the smooth particle-mesh Ewald method (Darden et al., 1993; Essmann et al., 1995) with a real-space cutoff of 0.9 nm and a Fourier grid spacing of 0.12 nm. Simulation in the *NpT* ensemble was achieved by isotropic coupling to a Berendsen barostat (Berendsen et al., 1984) at 1 bar with a coupling constant of 4 ps and separate coupling of the solute and the solvent to Berendsen thermostats (Berendsen et al., 1984) at 310 K with coupling constants of 0.1 ps. Bonds involving hydrogen were constrained with SETTLE (Miyamoto and Kollman, 1992) and LINCS (Hess et al., 1997) for solvent and solute, respectively. The integration time step was 2 fs.

POPC bilayer construction. A POPC molecule was constructed based on the 1,2-dilauryl-*sn*-glycero-3-phosphoethanolamine (DLPE) crystal structure of Elder et al. (Elder et al., 1977) by replacing amide hydrogens with methyls and extending the acyl chains in the *trans* conformation. A POPC bilayer with 160 lipids per leaflet measuring 7.77 nm by 7.96 nm in the bilayer plane was then constructed based on the DLPE crystal

symmetry (Elder et al., 1977). The rectangular unit cell was set to 20 nm normal to the plane of the bilayer and solvated with 29,953 TIP4P water molecules. The crystal symmetry was broken using a protocol derived from Takaoka *et al.* (Takaoka et al., 2000). However, we found that a preliminary low temperature simulation was required to prevent leaflet separation during the 510 K pulse, and that a large excess of water normal to the bilayer was required to accommodate undulations that were induced when, at 510 K, the area per lipid increased more quickly than could be accommodated in the bilayer plane by pressure coupling. Our protocol was 0-1 ns: 310 K, 1-1.1 ns: gradual heating to 510 K, 1.1-1.17 ns: gradual cooling to 360 K, 1.17-5.17 ns: 360 K, 5.17-40 ns: 323 K. At this point, the system dimensions were 10.4 nm by 9.8 nm in the bilayer plane and 12 nm along the bilayer normal. The undulations had reduced significantly by this point and we reduced the unit cell to 8 nm along the bilayer normal, excluding a large number of water molecules. This smaller system was simulated for an additional 50 ns at 310 K. A time-dependent analysis of the order parameters of the acyl chains indicated that the bulk properties of the bilayer had converged after 25 ns (not shown). The coordinates of this bilayer at 25 ns were used for the study of PagP.

*PagP all-atom modelling***.** The available coordinates for PagP, including crystal waters within 5 Å, were taken from the crystal structure presented in this article. We then constructed 1000 candidate models of L1 loop residues 38-45 as random coil using the program Loopy (Xiang et al., 2002), keeping the two loop conformations of lowest colony energy. To both of these structures, we modelled the missing sidechains for residues 4, 35, 36, 38-45, 46, 47, 146, 147, and 148 with SCWRL 3.0 (Canutescu et al., 2003), during which we found it necessary to manually direct Y46 $\chi 1$ to -120˚ to avoid atomic overlap. A single molecule of dodecyl phosphate (DP) was placed in the PagP binding pocket as a dodecyl sulphate analog. Parameters for DP were constructed based on those of dodecyl phosphocholine (DPC) (Tieleman et al., 2000) by removing the choline group and adding a partial charge of 0.1 to each of the four oxygens.

Composite system creation. Two copies of PagP having different L1 loop conformations were overlaid on the POPC bilayer with the principle axis of each β-barrel oriented 15° to the bilayer normal, bringing the L3 loop toward the bilayer, in order to align the exposed hydrophobic/hydrophilic surfaces with the hydrophobicity profile of the bilayer as suggested by Ahn *et al.* (Ahn et al., 2004). To reduce any anisotropic effects that the mobile loop regions of the protein may inflict upon the relative surface tension of each leaflet, the proteins were inserted anti-parallel to one another. Any phospholipid within 0.2 Å was removed, additionally removing a minimal number of phospholipids while ensuring that each leaflet contained the same amount. A surface representation of each protein was then constructed using MSMS (Sanner et al., 1996) and PagP-shaped holes were made in the bilayer according to the protocol of Faraldo-Gomez *et al.* (Faraldo-Gomez et al., 2002) in 3 segments of 20 ps each while applying a position restraint along the bilayer normal with a force constant of 1000 kJ mol^{-1} nm^{-2} on phosphorous atoms of all lipids. The strength of the hole-making force was 10, 100, and 500 kJ mol^{-1} nm^{-2} in the first, second and third segments, respectively. The composite system, composed of a solvated POPC bilayer, two protein molecules and associated DP and crystal waters, was neutralized with 8 Na^{+} ions and simulated for 50 ns at 310 K, with the position of protein and DP heavy atoms restrained during the first 5 ns. The rate at which the backbone RMSD to the crystal structure increased was significantly attenuated by 15 ns (not shown), and hence the final 30 ns of this simulation was used to generate conformations with which to initiate non-equilibrium steered MD simulations. Starting configurations for steered insertion of the *sn-1* chain of a POPC molecule into the PagP binding pocket and steered extraction of a DP molecule from the PagP binding

pocket were taken at 2 ns and 200 ps intervals, respectively. Prior to the initiation of steered insertion simulations, both DP molecules and two Na^+ ions were removed from the system.

Steered MD simulations. Steered insertion simulations employed a constant magnitude biasing force of the form

$$F = \frac{-K}{\left| x^{ligand} - x^{ref} \right|}(x^{ligand} - x^{ref}) \qquad (1)$$

where x_{ligand} represents the current position of the centre of mass of the distal 13 carbons in the *sn-1* chain of the selected phospholipid and x_{ref} represents the reference position calculated as

$$x^{ref} = x^{PagP} + x_0^{DP} + x_0^{PagP} \qquad (2)$$

in which the current position of the centre of mass of the PagP β-barrel, x^{PagP}, is offset by the vector from the initial center of mass of the PagP β-barrel, x_0^{PagP}, to the initial centre of mass of the DP molecule that was removed from the binding pocket, x_0^{DP}. Only biasing force components in the bilayer plane were applied. The constant K was set to 500 kJ mol^{-1} nm^2 such that the targeted *sn-1* chain is attracted to the PagP binding pocket. For steered extraction simulations, x^{ligand} represents the current position of the centre of mass of DP and the constant K was set to -500 kJ mol^{-1} nm^2 such that DP is repelled from its initial position.

The 100-ps length of the simulations is sufficient to observe productive insertions, yet short enough to allow a large number of repeats (1507 in this work) for the sake of statistical significance. In addition, the applied force should be large enough to promote a significant number of productive insertions yet small enough to ensure a significant number of unsuccessful attempts; furthermore, the latter should not perturb the protein structure significantly. With this in mind, we conducted preliminary simulations with three different force constants of 100, 500, and 1000 $kJ \cdot mol^{-1} \cdot nm^{-2}$ (results not shown), based on which we selected 500 $kJ \cdot mol^{-1} \cdot nm^{-2}$.

Half-ε double-pairlist combination rules for Berger lipids and OPLSAA proteins. During MD simulation, non-bonded interactions between atoms that are separated by exactly three bonds (the so-called 1-4 interactions) are often scaled by some factor from their full strength. This scaling factor differs amongst forcefields and thus the simultaneous use of multiple forcefields requires special attention to ensure that the 1-4 interactions of all molecules are properly scaled. Indeed, combination of the Berger lipid parameters and the OPLSAA protein parameters within the molecular dynamics simulation package GROMACS is complicated by the inability to specify unique parameters for Coulombic 1-4 interactions. It is this deficiency that led Tieleman *et al.* (Tieleman et al., 2006) to reparametrize the dihedral angle energy functions of the Berger lipids such that the nonbonded component of the 1-4 interactions is accounted for by the new dihedral potentials. While the published reparametrization of dihedral parameters appears to be a valid solution, one would like to avoid reparametrizing every lipid and detergent for which parameters have already been generated using this lipid forcefield. To this end, we apply a simple method that can be used within GROMACS to combine the above-mentioned forcefields while properly scaling the 1-4 nonbonded interactions of each that is at once simpler and more inclusive. This method makes use of the fact that (i) it is possible to specify a unique value for the LJ component of the 1-4 interactions, and

(ii) the scaling factor for the Coulombic component of the 1-4 interactions of one forcefield is an integer multiple of the scaling factor used in the other forcefield (1.0 for Berger lipids and 0.5 for OPLSAA). Specifically, in the GROMACS topology file, the ε values of the 1-4 LJ parameters of the lipids are multiplied by an additional factor of 0.5 in the pairtypes section and the list of 1-4 interactions in the pairs section is duplicated. The regular OPLSAA combination rules are then applied. In this way, the LJ and Coulombic 1-4 interactions are both cut in half and then included twice for the lipids, yielding properly scaled 1-4 interactions for both Berger lipids and OPLSAA protein. This method is referred to as the half-ε double-pairlist method (Chakrabarti et al., 2010).

Supplemental Figure S3 shows the area per lipid obtained from simulating a bilayer composed of 1,2-dipalmitoyl-*sn*-glycero-3-phosphocholine (DPPC), containing 64 lipids per leaflet solvated by 28.5 SPC (Berendsen et al., 1981) waters per lipid, using a variety of methods to treat the lipid forcefield. Initial DPPC parameters were taken from Tieleman et al. (Tieleman and Berendsen, 1996). The area per lipid derived from the direct use of lipid.itp and ffgmx, a forcefield that does not arbitrarily scale 1-4 interactions, is traced by the red line. Similar results are found when the functional form of the LJ variables are converted from C6/C12 to δ/ε while leaving 1-4 interactions as intended (blue line) and when applying the half-ε double-pairlist method under the OPLSAA scaling rules (purple line). However, the area per lipid is significantly reduced when the unmodified δ/ε formulation of LJ interactions is used while scaling lipid 1-4 interactions according to the rules of OPLSAA (blue line). In this set of conditions, the Coulombic 1-4 interactions of lipids are erroneously assigned only half of their full strength.

SUPPLEMENTAL REFERENCES

Ahn, V.E., Lo, E.I., Engel, C.K., Chen, L., Hwang, P.M., Kay, L.E., Bishop, R.E., and Privé, G.G. (2004). A hydrocarbon ruler measures palmitate in the enzymatic acylation of endotoxin. EMBO J *23*, 2931-2941.

Berendsen, H., Postma, J., van Gunsteren, W., DiNola, A., and Haak, J. (1984). Molecular dynamics with coupling to an external bath. The Journal of Chemical Physics *81*, 3684.

Berendsen, H., Postma, J., van Gunsteren, W., and Hermans, J. (1981). Interaction models for water in relation to protein hydration. In Intermolecular Forces, B. Pullman, ed. (Dordrecht, The Netherlands: Reidel), pp. 331-342.

Berger, O., Edholm, O., and Jahnig, F. (1997). Molecular dynamics simulations of a fluid bilayer of dipalmitoylphosphatidylcholine at full hydration, constant pressure, and constant temperature. Biophys J *72*, 2002-2013.

Canutescu, A.A., Shelenkov, A.A., and Dunbrack, R.L., Jr. (2003). A graph-theory algorithm for rapid protein side-chain prediction. Protein Sci *12*, 2001-2014.

Chakrabarti, N., Neale, C., Payandeh, J., Pai, E.F., and Pomes, R. (2010). An iris-like mechanism of pore dilation in the CorA magnesium transport system. Biophys J *98*, 784-792.

Darden, T., York, D., and Pedersen, L. (1993). Particle mesh Ewald: An N+log(N) method for Ewald sums in large systems. Journal of Chemical Physics *98*, 10089-10092.

Elder, M., Hitchcock, P., Mason, R., and Shipley, G. (1977). A Refinement Analysis of the Crystallography of the Phospholipid, 1, 2-Dilauroyl-DL-Phosphatidylethanolamine, and Some Remarks on Lipid-Lipid and Lipid-Protein Interactions. Proceedings of the

Royal Society of London. Series A, Mathematical and Physical Sciences (1934-1990) *354*, 157-170.
Essmann, U., Perera, L., Berkowitz, M., Darden, T., Lee, H., and Pedersen, L. (1995). A smooth particle mesh Ewald method. The Journal of Chemical Physics *103*, 8577.
Faraldo-Gomez, J.D., Smith, G.R., and Sansom, M.S. (2002). Setting up and optimization of membrane protein simulations. Eur Biophys J *31*, 217-227.
Hess, B., Bekker, H., Berendsen, H., and Fraaije, J. (1997). J. Comput. Chem. *18*, 1463-1472.
Jorgensen, W., Chandrasekhar, J., Madura, J., Impey, R., and Klein, M. (1983). Comparison of simple potential functions for simulating liquid water. The Journal of Chemical Physics *79*, 926.
Kaminski, G., Friesner, R., Tirado-Rives, J., and Jorgensen, W. (2001). Evaluation and Reparametrization of the OPLS-AA Force Field for Proteins via Comparison with Accurate Quantum Chemical Calculations on Peptides. Journal of Physical Chemistry B *105*, 6474-6487.
Lindahl, E., Hess, B., and van der Spoel, D. (2001). GROMACS 3.0: a package for molecular simulation and trajectory analysis. Journal of Molecular Modeling *7*, 306-317.
Miyamoto, S., and Kollman, P. (1992). Settle: An analytical version of the SHAKE and RATTLE algorithm for rigid water models. Journal of Computational Chemistry *13*, 952-962.
Sanner, M.F., Olson, A.J., and Spehner, J.C. (1996). Reduced surface: an efficient way to compute molecular surfaces. Biopolymers *38*, 305-320.
Takaoka, Y., Pasenkiewicz-Gierula, M., Miyagawa, H., Kitamura, K., Tamura, Y., and Kusumi, A. (2000). Molecular Dynamics Generation of Nonarbitrary Membrane Models Reveals Lipid Orientational Correlations. Biophysical Journal *79*, 3118-3138.
Tieleman, D., and Berendsen, H. (1996). Molecular dynamics simulations of a fully hydrated dipalmitoylphosphatidylcholine bilayer with different macroscopic boundary conditions and parameters. The Journal of Chemical Physics *105*, 4871.
Tieleman, D., MacCallum, J., Ash, W., Kandt, C., Xu, Z., and Monticelli, L. (2006). Membrane protein simulations with a united-atom lipid and all-atom protein model: lipid-protein interactions, side chain transfer free energies and model proteins. JOURNAL OF PHYSICS CONDENSED MATTER *18*, 1221.
Tieleman, D., van der Spoel, D., and Berendsen, H. (2000). Molecular Dynamics Simulations of Dodecylphosphocholine Micelles at Three Different Aggregate Sizes: Micellar Structure and Chain Relaxation. JOURNAL OF PHYSICAL CHEMISTRY A *104*, 6380-6388.
Tieleman, D.P., Forrest, L.R., Sansom, M.S., and Berendsen, H.J. (1998). Lipid properties and the orientation of aromatic residues in OmpF, influenza M2, and alamethicin systems: molecular dynamics simulations. Biochemistry *37*, 17554-17561.
van Gunsteren, W., and Berendsen, H. (1990). Computer Simulation of Molecular Dynamics: Methodology, Applications, and Perspectives in Chemistry. Angewandte Chemie International Edition in English *29*, 992-1023.
Xiang, Z., Soto, C.S., and Honig, B. (2002). Evaluating conformational free energies: the colony energy and its application to the problem of loop prediction. Proc Natl Acad Sci U S A *99*, 7432-7437.