

Classification: Biological Sciences/Biophysics and Computational Biology

Clustering and Dynamics of Crowded Proteins near Membranes and Their Influence on Membrane Bending

Grzegorz Nawrocki¹, Wonpil Im², Yuji Sugita^{3,4,5}, Michael Feig^{1,5*}

¹Department of Biochemistry and Molecular Biology, Michigan State University, East Lansing, Michigan 48824, USA

²Department of Biological Sciences and Bioengineering Program, Lehigh University, Bethlehem, PA 18015, USA

³RIKEN Cluster for Pioneering Research, Wako, Saitama 351-0198, Japan

⁴RIKEN Center for Computational Research, Kobe, Hyogo 640-0047, Japan

⁵RIKEN Center for Biosystems Dynamics Research, Kobe, Hyogo 640-0047, Japan

*Corresponding author.

603 Wilson Road, Room 218 BCH

East Lansing, MI 48824, USA

E-mail: mfeiglab@gmail.com

Phone: 517-432-7439

RUNNING TITLE

Protein Crowding near Membranes

ABSTRACT

Atomistic molecular dynamics simulations of concentrated protein solutions in the presence of a phospholipid bilayer are presented to gain insights into the dynamics and interactions at the cytosol-membrane interface. The main finding is that proteins that are not known to specifically interact with membranes are preferentially excluded from the membrane leaving a depletion zone near the membrane surface. As a consequence, effective protein concentrations increase leading to increased protein contacts and clustering, whereas protein diffusion becomes faster near the membrane for proteins that do occasionally enter the depletion zone. Since protein-membrane contacts are infrequent and short-lived in this study, the structure of the lipid bilayer remains largely unaffected by the crowded protein solution, but when proteins do contact lipid head groups, small but statistically significant local membrane curvature is induced on average.

KEYWORDS

Molecular dynamics; diffusion; crowding; membrane curvature; phospholipids

STATEMENT OF SIGNIFICANCE

Interactions between crowded cytosols and membrane surfaces are unavoidable inside cells. This all-atom simulation study suggests that non-specific protein-membrane interactions create a water-rich protein depletion zone between the membrane and the crowded environment, leading to an increased propensity of proteins to aggregate in bulk but also allow for accelerated diffusion on the surface of the membrane when proteins come closer to the surface occasionally. The simulation results furthermore provide evidence of a non-specific mechanism for protein-induced membrane curvature formation as a result of crowding near the membrane.

\body

INTRODUCTION

Proteins and nucleic acids have to function under highly crowded conditions inside cells (1). An unresolved question is how such environments impact biomolecular structure and dynamics compared to the *in vitro* non-crowded conditions in most experimental and computational studies (2, 3). Earlier work has described the volume exclusion effect of crowding (4), but more recent studies emphasize the role of weak, non-specific interactions between biomolecules in the cell (5-11). There is increasing evidence that protein-protein interactions in highly concentrated environments can potentially destabilize native folds contrary to what the volume-exclusion effect predicts (12-16). Transient molecular cluster formation between biomolecules has emerged as the primary determinant of reduced diffusion in crowded cellular environments (10, 11, 17-19).

In addition to studies of biomolecular crowding in cytoplasmic environments, crowding inside or near membranes has also been examined (20-24). Membrane surfaces are ubiquitous not just at the boundaries of cells but also as part of lipid vesicles and cellular organelles such as the endoplasmic reticulum (ER). Protein crowding within the membrane slows down dynamics as in the cytoplasm (23), but the effects are more complicated due to domain formation and confinement by cytoskeletal elements (25). Moreover, as predicted by Saffman-Delbrück theory (26), diffusion of proteins inside the membrane is only weakly dependent on particle size R (23). In contrast, the Stokes-Einstein model for isotropic solvent describes an $1/R$ dependence. Consequently, molecular association and clustering is not expected to strongly impact diffusion within the membrane.

Less is known about the interface between membrane surfaces and crowded cytoplasmic environments, especially when non-membrane binding proteins are involved. Previous studies

have found a role of protein crowding in inducing membrane curvature (27). This effect was attributed to the anisotropic pressure that is generated by proteins moving laterally on a membrane surface. More recent work argues that this effect is much less significant compared to membrane curvature induced by hydrophobic insertion of peripherally associated membrane proteins (28, 29). Moreover, it seems that a high fraction of the membrane surface needs to be covered by proteins and an asymmetric distribution of crowding between membrane leaflets is needed to realize significant overall curvature (28). It remains unclear how high concentrations of proteins may modulate other membrane properties. There is also little insight into how the properties of cytoplasmic proteins may be affected by the presence of a membrane surface under crowded conditions. The structural and dynamic properties of proteins in concentrated solutions could be altered in the presence of a membrane. It may be expected, for example, that proteins forced to interact with a membrane due to crowding experience reduced diffusion and are subject to destabilization when surrounded partially by a non-aqueous environment.

To examine these questions in molecular detail, we present 10 μ s-scale atomistic molecular dynamics (MD) simulations of mixtures of proteins in the presence and absence of a membrane. The simulations suggest that the presence of the membrane increases protein clustering and allows proteins to diffuse faster on the membrane surface than in the crowded milieu. However, the proteins affect the membrane properties only to a small extent and the membrane has little effect on protein stability.

RESULTS AND DISCUSSION

We carried out simulations of concentrated mixtures of villin, protein G, and ubiquitin in the absence and presence of a lipid bilayer composed of POPC, sphingomyelin, and cholesterol (**Fig.**

1), initially with NAMD on sub- μ s time scales followed by 10 μ s on Anton2 for which results are reported here (*SI Appendix*, Table S1). The simulations follow well-tested protocols for describing protein-membrane interactions that have resulted in excellent agreement between simulation and experiment in many previous studies (30-33). To avoid the overestimation of protein-protein interactions, we have applied a recently introduced force field modification (10).

Stability of proteins

Most of the proteins remained stable near their native structure, but at least one villin in each of the simulations and two copies of protein G in the simulation at 10% without the membrane deviated significantly from the experimental reference structures after several microseconds (*SI Appendix*, Figs S1-S3). Increased RMSD was correlated with increases in the radius of gyration (*SI Appendix*, Figs S4-S6) and indicates unfolding. In previous simulations of similar protein solutions we partial unfolding was also observed, especially for villin (10, 13). Villin has marginal stability around 4 kcal/mol to full unfolding and 2-3 kcal/mol to partially unfolded states with the force field used here (10). Experimental stability is estimated to be 2-3 kcal/mol at 298K with unfolding rates around 100 μ s or less (34). Some unfolding during 10- μ s may thus be expected. However, the main conclusions are not significantly affected when only the first 2 μ s are analyzed to avoid partially unfolded structures (*SI Appendix*, Figs S15, S32, and S34).

The volume exclusion effect of crowding is expected to increase the stability of compact native states (35), but a destabilization of the native state due to protein-protein interactions upon crowding has also been proposed (13). To analyze whether protein structures may be altered at higher concentrations and whether the presence of the membrane affects stability, average RMSD and radii of gyration for each type of protein were calculated as a function of concentration (*SI Appendix*, Table S2). Small differences may be consistent with native state destabilization (such

as a slight increase in average RMSD for folded villin with an RMSD below 2.5 Å) or with native state compaction due to volume exclusion (such as a slight decrease in the radius of gyration for ubiquitin) when comparing between the most (30%) and least (5%) concentrated systems. However, p-values of 0.2 or larger suggest only weak significance of these observations. The presence of the membrane also does not significantly affect protein stability (*SI Appendix*, Table S3).

Similar conclusions are found when comparing root mean square fluctuations (RMSF) (*SI Appendix*, Figs S7-S9). For most residues, different concentrations and the presence of the membrane have only a small effect, but the loop in ubiquitin around residue 35 fluctuates significantly less at higher protein concentrations. A decreased RMSF is consistent with the smaller radius of gyration in ubiquitin upon crowding.

Protein contacts and clustering

At the high concentrations considered here, interactions between proteins are unavoidable. Indeed, we find extensive contacts between proteins at all concentrations (*SI Appendix*, Fig. S10). In the absence of the membrane, about 5-10% of the theoretical maximum contacts are formed at any time largely independent of concentration. When contacts are analyzed between the same type of protein, we find that contacts between ubiquitin are more likely than between villin or protein G (*SI Appendix*, Table S4). This finding generally holds even after normalizing the number of contacts by the surface area of spheres with volumes equivalent to the proteins (*SI Appendix*, Table S4) and can be understood based on the differences in net charges. Ubiquitin is neutral and villin and protein G are positively and negatively charged, respectively. Overall, interactions between different proteins are non-specific without a strong bias towards specific protein-protein interfaces. However, there are some preferences for involving certain residues in protein-protein contacts

(**Fig. 2** and *SI Appendix*, Figs S11-S13) that vary only slightly depending on what the other protein partner is. Interestingly, the RMSF is elevated for many of the residues involved in forming contacts (*SI Appendix*, Figs S7-S9).

When the membrane is introduced, the number of contacts per protein increases with the most significant change at 5% and 30% (based on p-values of 0.03 (5%), 0.55 (10%), and 0.07 (30%); *SI Appendix*, Fig. S10 and Table S4). The trend is less clear for contacts between proteins of the same type (*SI Appendix*, Table S4). This contrasts with a geometrically expected decrease in contacts due to the membrane since contacts along z between proteins above and below the membrane are prevented.

We further analyzed the formation of clusters based on the protein contacts. Cluster size distributions based on C α -C α contacts are shown in **Fig. 1B** and are very similar results to contacts based on heavy atom distances (*SI Appendix*, Fig. S14). A simple hard-sphere model without attraction results in significantly smaller cluster sizes (*SI Appendix*, Fig. S16).

At 5%, we find a decaying cluster size distribution indicative of transient cluster formation similar to what we described previously for concentrated villin solutions below the solubility limit (10). At 30%, most proteins are found in a single large cluster consistent with a phase change to an aggregated form (36). The presence of the membrane generally shifts the cluster size distribution to larger cluster sizes. This can be seen most clearly from a reduction in monomers and other small clusters. This effect is greater at 5% and 30% consistent with a greater increase in contacts (Table S4). We note that differences at the largest cluster sizes reflect in part different numbers of protein copies in non-membrane and membrane systems (Table S1).

The increased clustering in the presence of the membrane indicates a decrease in solubility that could be observable macroscopically. Here, the protein solutions are sandwiched between two

membrane surfaces (when considering the periodic images along z) with a gap of about 150 Å. Such conditions may be found in the ER or Golgi apparatus. Although our systems are only a simple approximation of these complex biological environments, our results are consistent with previous observations of condensation of enzymes in the rough ER (37) and the aggregation and sorting of secretory proteins in the ER and Golgi apparatus (38, 39). While specific mechanisms involving membrane-bound receptors likely play a role as well (40), this study proposes a generic mechanism for selective protein aggregation in the ER based on close membrane surfaces.

Protein-membrane interactions

None of the proteins studied here are expected to bind the membrane peripherally or via insertion. However, crowding may be expected to force proteins to make close contacts with the membrane surface. Density profiles of proteins relative to the membrane show that this is not the case (**Fig. 3** and *SI Appendix*, Fig. S17). Instead, the proteins are preferentially excluded from the membrane surface, even at 30%. Next to the membrane surface, a water and ion layer of about 10-15 Å is formed where protein concentrations are low. This finding is independent of the scaling of protein-water interactions (*SI Appendix*, Fig. S18). For comparison, we simulated hard spheres with radii equivalent to the proteins in the presence of a hard surface (*SI Appendix*, Fig. S18). The hard spheres pack more closely to the surface compared to the proteins where density increases slowly up to 70 Å away from the membrane center. Our results also do not agree with the increased membrane interactions reported from a simple hydrodynamic model (41). This indicates clearly that the proteins are thermodynamically excluded from the membrane. This means that protein-membrane binding affinities are low and that solvation of the membrane lipid head groups with water and ions is preferred over membrane-protein interactions. This finding is consistent with previous analyses showing weak protein-membrane binding unless anionic lipids are present (42-

44). Moreover, examples of strongly binding peripheral membrane proteins typically involve hydrophobic anchors that are partially inserted into the lipid bilayer to interact with the lipid acyl chains (45) and/or π -cation interactions between aromatic residues and choline headgroups (30, 44). The proteins studied here do not have a large fraction of aromatic residues. surface-exposed hydrophobic elements, or a lipidation modification such as myristoylation suitable for partial membrane insertion.

Protein exclusion from the membrane surface results in an increased protein concentration away from the membrane. This is consistent with increased protein contacts and clustering in the presence of the membrane. However, all types of proteins occasionally come into contact with the lipid headgroups (*SI Appendix*, Fig. S19). The interactions are slightly more frequent for villin and protein G than for ubiquitin (*SI Appendix*, Fig. S19). This is expected from stronger electrostatic interactions between the charged proteins and the zwitterionic lipid head group. Protein residues that are preferred in protein-membrane interactions vary by protein (**Fig. 2**) and only partially overlap with the residues involved in protein-protein contacts (*SI Appendix*, Figs S11-S13). The residues that are most likely to be involved in protein-membrane contacts are tyrosine, which is known to interact strongly with choline head groups (30, 44), hydrophobic alkanes, which interact favorably with the lipid acyl chains, and polar residues, which can form hydrogen bonds with the lipid head groups (*SI Appendix*, Table S5).

An analysis of contact residence times (based on protein-lipid heavy atom distances within 5 Å) revealed a typical contact time of around 2 ns for protein G and ubiquitin; villin remained bound slightly longer, i.e. 4-6 ns (*SI Appendix*, Table S6). A much longer time component of around 1 μ s is attributed to unbinding, reinsertion into the main protein cluster, and later rebinding.

The predictions from the simulations could be tested experimentally by employing X-ray or neutron scattering to study the density variations. Protein-membrane interactions could be probed via site-directed spin labels in combination with paramagnetic resonance or other fluorescence techniques. We are not aware of such experiments for comparable systems.

Protein diffusion

Retarded diffusion of proteins upon crowding is well known. Recent work (10, 17) (19, 46) has suggested that this is largely a result of transient cluster formation. Here, we analyzed whether the presence of the membrane affects diffusive properties. Translational and rotational diffusion coefficients in bulk solutions without the membrane bilayer match previously reported values for villin (10). *SI Appendix* provides averages in Tables S7-S8, mean-squared displacement (MSD) in Figs S20-S22, and correlation functions in Figs S23-S25. The TIP3P water model used here underestimates solvent viscosity about threefold and reported diffusion rates are thus three times faster than in experiment. Diffusion coefficients for the ubiquitin are retarded more than for villin or protein G, as expected based on size. There is only a moderate decrease in translational diffusion between short (<1 ns) and longer (>10 ns) time scales for 5% and 10%, consistent with extensive cluster formation that results in a lack of transiently varying diffusion rates (10). At 30%, diffusion on longer time scales is retarded more significantly compared to shorter times, reflecting cage effects (*SI Appendix*, Table S7).

In the presence of the membrane, diffusion rates vary depending on the location of a protein with respect to the membrane (**Fig. 3**). The discussion here primarily focuses on motion parallel to the membrane. Diffusion perpendicular to the membrane (*SI Appendix*, Fig. S27) shows similar trends. Diffusion is slowest at the farthest point from the membrane and increases towards the membrane surface up until 35 Å from the membrane center. When proteins come into direct contact with the

lipids, sharply reduced diffusion is observed in most cases. Translational and rotational diffusion are affected similarly, but translational diffusion is accelerated more strongly near the membrane than rotational diffusion at 5% and 10%. The opposite trend is found at 30%. At 5%, rotational diffusion appears to slow down significantly near 40 Å, but the translational diffusion is affected less. This may indicate that proteins near the membrane are subject to orientational restraints due to preferential involvement of certain protein residues in membrane interactions (**Fig. 2**), but since this observation is only made at 5%, an artefact due to limited sampling is also possible. Increased diffusion near the membrane is most pronounced for long-time diffusion (>10 ns) (*SI Appendix*, Figs S26-S27). This finding is largely independent of whether the initial, mid- or endpoint of a diffusion interval is used to assign the distance from the membrane (*SI Appendix*, Fig. S28).

The translational diffusion of proteins parallel to the membrane in a confined membrane system can be estimated from bulk 3D diffusion in a non-membrane system (47). The long-time x-y diffusion far away from the membrane is lower than or equal to the bulk-based estimate (using Eq. 4 in *SI Appendix*), but diffusion near the membrane surface is significantly faster (*SI Appendix*, Fig. S29). Faster diffusion near the membrane can be understood from the protein concentration gradient in the presence of the membrane where there are fewer obstacles in the depletion zone near the membrane surface, but the reduced protein concentration also means that proteins are less likely to experience the faster diffusion near the membrane surface. However, there is still a net effect of accelerated diffusion parallel to the membrane when the probability of finding a protein close to the membrane is considered (*SI Appendix*, Fig. S30). Villin and protein G benefit most when they move from the bulk to about 50-60 Å from the membrane center, and the acceleration is greater at 30% than at 5 and 10% with an increase of up to almost 90% for protein G. The increase in diffusion described here is in disagreement with the retarded diffusion found near the

cell wall based on a simplified model that emphasizes hydrodynamic effects but neglects the details of protein-protein and protein-lipid interactions (41).

Rotational diffusion is not expected to be affected strongly in the presence of confinement when the membrane surfaces are separated by more than several times the size of the proteins. Therefore, we compared the rotational diffusion rates directly between the non-membrane and membrane systems (*SI Appendix*, Fig. S31). Generally, we find a similar conclusion of rotational diffusion far away from the membrane being slower than or equal to bulk diffusion, while diffusion near the membrane surface exceeds bulk diffusion values at 10% and 30% volume fractions. At 5% volume fraction, only villin surpasses bulk diffusion rates at around 47 Å from the membrane center, whereas protein G and ubiquitin remain below bulk diffusion rates for all distances from the membrane center. This can be understood from a greater sensitivity of rotational diffusion to contact formation (10) and the greater increase in the number of protein contacts at 5% in the presence of the membrane relative to the non-membrane systems (*SI Appendix*, Fig. S10), as well as apparently longer-lasting protein-membrane contacts.

The overall picture that is emerging from the above analysis is that proteins near a membrane surface may diffuse faster than in the crowded bulk solution due to the protein depletion zone that is formed by non-membrane interacting proteins in the vicinity of a membrane bilayer. The faster diffusion is most evident in translational diffusion over longer time scales and could suggest a mechanism for circumventing the challenge of slow transport of biomolecules within the crowded cellular milieu and to reach membrane-embedded receptors and transporters. One could test this idea experimentally by comparing diffusion via nuclear magnetic resonance or fluorescence recovery after photobleaching in a very membrane-rich environment of a cell, such as the ER, with other parts of the cell or by studying comparable *in vitro* systems.

Membrane properties

Lipid order parameters as a function of protein concentration (*SI Appendix*, Fig. S33) were found to be very similar values to previous results for pure POPC-sphingomyelin-cholesterol mixtures with the same force field that was used here (48). Protein concentration had virtually no effect, indicating that the internal structure and dynamics of the lipid bilayer and the liquid-ordered state expected for the lipid composition in this study are not affected by the presence of the proteins.

Previous studies have suggested that crowding may introduce membrane bending (27). Membrane deformations as a result of protein-membrane contacts were examined by averaging the distance of lipid phosphate atoms from the membrane center for phosphates near the closest protein-membrane contact point as a function of the protein-membrane distance (**Fig. 4**). The average phosphate distance from the membrane center is decreased by as much as 0.2 Å as proteins approach the membrane to within 3-10 Å of heavy-atom distances. This suggests membrane indentation upon either direct contact or via indirect interactions mediated e.g. by water or longer-range electrostatics. This effect is most pronounced for villin, which also has the longest protein-membrane contact residence times. The weakest effect is observed for ubiquitin, for which the membrane is only slightly indented upon contact at 5% and 10%, while the membrane is slightly wider when ubiquitin touches the membrane at 30% (**Fig. 4**). For very short protein-membrane contacts (<3 Å), the membrane appears to become distorted more strongly with indentations and extrusions varying by protein and concentration, but there are high statistical uncertainties.

In the absence of any specific interactions it is expected that proteins are most likely to bump into membrane surfaces that extrude furthest into the aqueous solvent. In principle, lipid head groups could also be pulled away from the membrane due to electrostatic attraction. However, the opposite finding of an indented membrane upon protein contact suggests a specific mechanism for inducing

membrane curvature. We did not observe any overall net membrane bending, presumably because of periodic boundary conditions and an equal distribution of proteins on either side of the membrane. However, our results predict that net bending would arise if proteins are unequally distributed on either side of the membrane in terms of concentration and/or composition. An indentation of 0.2 Å over a 30 Å diameter disk (the diameter of the spherical region within which phosphates around the protein-membrane contact point were analyzed) is equivalent to the curvature on the surface of a sphere with a 560 Å radius (**Fig. 4**). Such spheres are in the range of lipid vesicle sizes suggesting that crowding could stabilize such vesicles. While the effect may appear small, we find that at about 25% of the time there is either a villin or protein G within 10 Å of a lipid (the maximum distance at which we see an effect on membrane curvature). This translates into one protein per 700 nm² or about 56 proteins bound to the membrane of an entire vesicle. This could provide significant overall stabilization when the contributions of all proteins are summed up. The stabilization of vesicles or formation of curved membranes could be tested experimentally in the presence of proteins at high concentration.

CONCLUSIONS

In this study, we are reporting atomistic simulations of concentrated protein solutions near a neutral phospholipid bilayer. The model proteins considered here, villin, protein G, and ubiquitin, are not known to interact specifically with phospholipid membranes and we found that even at the highest concentrations of 30% volume fraction, the proteins are preferentially excluded from the membrane surface. This finding has two major consequences: 1) The proteins effectively experience a higher concentration as they occupy a smaller volume in the presence of the membrane. This leads to increased contacts and increased clustering; 2) When proteins enter the

depletion zone of 10-15 Å near the membrane surface, they can diffuse significantly faster than in the crowded environment, especially over longer (>10 ns) time scales. Additional insights from this work are that the membrane structure remains largely unperturbed in the presence of the crowding proteins, which may be expected since the proteins largely avoid the membrane. However, when proteins do contact the membrane, they appear to be able to induce local curvature that could support lipid vesicles. The conclusions are experimentally testable hypotheses that we hope will stimulate new studies of the interaction between crowded cellular environments and membrane surfaces either *in vitro* or *in vivo*.

A major limitation of this work is the relatively small size of the systems dictated by the available computer resources and future work will aim at extending the spatial and temporal scales via coarse-grained modeling. Another limitation is that proteins that are expected to interact with the membrane were not included and that no integral membrane proteins were present in the phospholipid bilayer. In both cases, we would expect that cytoplasmic proteins may interact more extensively with the membrane surface. This could mitigate the membrane-induced increase in clustering described here and alter the diffusive characteristics of proteins near the membrane. Extending our current work to such more complicated systems will be another aim of future work.

METHODS

Systems: Concentrated solutions of proteins with and without a lipid bilayer were constructed (**Fig. 1** and Table S1). All systems contained equal numbers of three types of proteins: the chicken villin head piece (HP-36; “villin”), the B1 domain of streptococcal protein G (“protein G”), and human ubiquitin (“ubiquitin”). The proteins were chosen because of their small size, variation in charge (villin: +2e, protein G: -4e, ubiquitin: 0), and variation in secondary structures. Phospholipid bilayers consisted of equal numbers of 1-palmitoyl-2-oleyl-phosphatidylcholine (POPC),

sphingomyelin, and cholesterol to mimic the typical composition of animal cell membranes. Crowded systems were prepared at three protein concentrations (5%, 10%, and 30%) based on the total volume of the proteins relative to the aqueous solvent. In terms of weight, the concentrations were about 40, 80, and 250 g/L. All systems were solvated in explicit water and K^+/Cl^- were added to neutralize the systems and achieve excess KCl concentrations of about 150 mM (*SI Appendix*, Table S1) to reflect typical physiological conditions (49).

MD simulations: The initial systems were equilibrated before commencing production MD simulations. The equilibration of the systems was performed using NAMD (version 2.10) (50) and the CHARMM-GUI protocol (51, 52). After initial setup and equilibration (see *SI Appendix*), production simulations of the 10% and 30% systems were then carried out without any restraints using NAMD (version 2.10) (50) for time periods of 300-800 ns (*SI Appendix*, Table S1). Subsequently, simulations were extended for 10 μ s using the special-purpose Anton2 hardware (53). Additional simulations at 5%, with and without a membrane were also carried out on Anton2 over 10 μ s each. All systems were simulated under periodic boundaries in the NPT ensemble. In the production simulations, temperature was set to 310 K and the pressure to 1.01325 bar. See *SI Appendix* for further details.

DATA AVAILABILITY

All data discussed in the paper will be made available to readers.

ACKNOWLEDGEMENTS

Funding was provided by the National Science Foundation grants MCB 1330560, MCB 1817307, and MCB 1810695, by the National Institute of Health (NIGMS) grant R35 GM126948, by RIKEN pioneering research projects on “dynamic structural biology”, “glycolipidologue”, and a grant for the RIKEN center for biosystems dynamics research (BDR). Computer time was used on XSEDE resources (TG-MCB090003) and on the Anton2 special-purpose supercomputer at the Pittsburgh Supercomputing Center (PSCA18053P).

REFERENCES

1. R. J. Ellis, Macromolecular Crowding: an Important but Neglected Aspect of the Intracellular Environment. *Curr Opin Struct Biol* **11**, 114-119 (2001).
2. J. Danielsson, M. Oliveberg, Comparing Protein Behaviour *in vitro* and *in vivo*, What Does the Data Really Tell Us? *Curr Opin Struct Biol* **42**, 129-135 (2017).
3. G. Rivas, A. P. Minton, Macromolecular Crowding In Vitro, In Vivo, and In Between. *Trends Biochem Sci* **41**, 970-981 (2016).
4. S. B. Zimmerman, A. P. Minton, Macromolecular Crowding - Biochemical, Biophysical, and Physiological Consequences. *Annu Rev Biophys Biomol Struct* **22**, 27-65 (1993).
5. M. Feig, I. Yu, P.-h. Wang, G. Nawrocki, Y. Sugita, Crowding in Cellular Environments at an Atomistic Level from Computer Simulations. *J Phys Chem B* **121**, 8009-8025 (2017).
6. I. Yu *et al.*, Biomolecular Interactions Modulate Macromolecular Structure and Dynamics in Atomistic Model of a Bacterial Cytoplasm. *eLife* **5**, e19274 (2016).
7. W. B. Monteith, R. D. Cohen, A. E. Smith, E. Guzman-Cisneros, G. J. Pielak, Quinary Structure Modulates Protein Stability in Cells. *Proc Natl Acad Sci USA* **112**, 1739-1742 (2015).
8. S. Majumder *et al.*, Probing Protein Quinary Interactions by In-Cell Nuclear Magnetic Resonance Spectroscopy. *Biochemistry* **54**, 2727-2738 (2015).
9. M. Roos *et al.*, Coupling and Decoupling of Rotational and Translational Diffusion of Proteins under Crowding Conditions. *J Am Chem Soc* **138**, 10365-10372 (2016).
10. G. Nawrocki, P.-h. Wang, I. Yu, Y. Sugita, M. Feig, Slow-Down in Crowded Protein Solutions Correlates with Transient Oligomer Formation. *J Phys Chem B* **121**, 11072-11084 (2017).
11. F. Etoc *et al.*, Non-specific interactions govern cytosolic diffusion of nanosized objects in mammalian cells. *Nat Mater* **17**, 740-746 (2018).
12. K. Inomata *et al.*, High-Resolution Multi-Dimensional NMR Spectroscopy of Proteins in Human Cells. *Nature* **458**, 106-111 (2009).
13. R. Harada, N. Tochio, T. Kigawa, Y. Sugita, M. Feig, Reduced Native State Stability in Crowded Cellular Environment Due to Protein-Protein Interactions. *J Am Chem Soc* **135**, 3696-3701 (2013).
14. M. Feig, Y. Sugita, Variable Interactions between Protein Crowders and Biomolecular Solutes are Important in Understanding Cellular Crowding. *J Phys Chem B* **116**, 599-605 (2012).
15. A. C. Miklos, M. Sarkar, Y. Wang, G. J. Pielak, Protein Crowding Tunes Protein Stability. *J Am Chem Soc* **133**, 7116-7120 (2011).
16. Q. H. Wang, A. Zhuravleva, L. M. Gierasch, Exploring Weak, Transient Protein-Protein Interactions in Crowded In vivo Environments by In-Cell Nuclear Magnetic Resonance Spectroscopy. *Biochemistry* **50**, 9225-9236 (2011).
17. G. Nawrocki, A. Karaboga, Y. Sugita, M. Feig, Effect of Protein-Protein Interactions and Solvent Viscosity on the Rotational Diffusion of Proteins in Crowded Environments. *Phys Chem Chem Phys* **21**, 876-883 (2019).
18. C. Beck *et al.*, Nanosecond Tracer Diffusion as a Probe of the Solution Structure and Molecular Mobility of Protein Assemblies: The Case of Ovalbumin. *J Phys Chem B* **122**, 8343-8350 (2018).

19. S. von Bülow, M. Siggel, M. Linke, G. Hummer, Dynamic cluster formation determines viscosity and diffusion in dense protein solutions. *Proc Natl Acad Sci USA* 10.1073/pnas.1817564116, 201817564 (2019).
20. H. Kirchhoff, S. Haferkamp, J. F. Allen, D. B. Epstein, C. W. Mullineaux, Protein diffusion and macromolecular crowding in thylakoid membranes. *Plant Physiol* **146**, 1571-1578 (2008).
21. H.-X. Zhou, Crowding effects of membrane proteins. *J Phys Chem B* **113**, 7995-8005 (2009).
22. M. Lindén, P. Sens, R. Phillips, Entropic tension in crowded membranes. *Plos Comp Biol* **8**, e1002431 (2012).
23. S. Ramadurai *et al.*, Lateral diffusion of membrane proteins. *J Am Chem Soc* **131**, 12650-12656 (2009).
24. M. Javanainen, H. Martinez-Seara, R. Metzler, I. Vattulainen, Diffusion of Integral Membrane Proteins in Protein-Rich Membranes. *J Phys Chem Lett* **8**, 4308-4313 (2017).
25. J. A. Dix, A. S. Verkman, Crowding Effects on Diffusion in Solutions and Cells. *Annu Rev Biophys* **37**, 247-263 (2008).
26. P. Saffman, M. Delbrück, Brownian motion in biological membranes. *Proc Natl Acad Sci USA* **72**, 3111-3113 (1975).
27. J. C. Stachowiak *et al.*, Membrane bending by protein–protein crowding. *Nat Cell Biol* **14**, 944 (2012).
28. M. M. Kozlov *et al.*, Mechanisms shaping cell membranes. *Curr Op Cell Biol* **29**, 53-60 (2014).
29. Z. Chen, E. Atefi, T. Baumgart, Membrane Shape Instability Induced by Protein Crowding. *Biophys J* **111**, 1823-1826 (2016).
30. A.-S. Schillinger, C. Grauffel, H. M. Khan, Ø. Halskau, N. Reuter, Two homologous neutrophil serine proteases bind to POPC vesicles with different affinities: When aromatic amino acids matter. *BBA Biomemb* **1838**, 3191-3202 (2014).
31. M. P. Muller *et al.*, Characterization of Lipid–Protein Interactions and Lipid-Mediated Modulation of Membrane Protein Function through Molecular Simulation. *Chem Rev* **119**, 6086-6161 (2019).
32. V. Monje-Galvan, J. B. Klauda, Peripheral membrane proteins: Tying the knot between experiment and computation. *BBA Biomemb* **1858**, 1584-1593 (2016).
33. N. Schwierz, S. Krysiak, T. Hugel, M. Zacharias, Mechanism of Reversible Peptide–Bilayer Attachment: Combined Simulation and Experimental Single-Molecule Study. *Langmuir* **32**, 810-821 (2016).
34. M. Buscaglia, J. Kubelka, W. A. Eaton, J. Hofrichter, Determination of Ultrafast Protein Folding Rates from Loop Formation Dynamics. *J Mol Biol* **347**, 657-664 (2005).
35. H.-X. Zhou, G. Rivas, A. P. Minton, Macromolecular Crowding and Confinement: Biochemical, Biophysical, and Potential Physiological Consequences. *Annu Rev Biophys* **37**, 375-397 (2008).
36. D. Karandur, K.-Y. Wong, B. M. Pettitt, Solubility and Aggregation of Glys in Water. *J Phys Chem B* **118**, 9565-9572 (2014).
37. J. Tooze, H. F. Kern, S. D. Fuller, K. E. Howell, Condensation-sorting events in the rough endoplasmic reticulum of exocrine pancreatic cells. *J Cell Biol* **109**, 35 (1989).
38. E. Chanat, W. B. Huttner, Milieu-induced, selective aggregation of regulated secretory proteins in the trans-Golgi network. *J Cell Biol* **115**, 1505 (1991).

39. M. Mizuno, S. J. Singer, A soluble secretory protein is first concentrated in the endoplasmic reticulum before transfer to the Golgi apparatus. *Proc Natl Acad Sci USA* **90**, 5732 (1993).
40. J. Dancourt, C. Barlowe, Protein Sorting Receptors in the Early Secretory Pathway. *Annu Rev Biochem* **79**, 777-802 (2010).
41. E. Chow, J. Skolnick, Effects of confinement on models of intracellular macromolecular dynamics. *Proc Natl Acad Sci USA* **112**, 14846-14851 (2015).
42. B. Rogaski, J. B. Klauda, Membrane-Binding Mechanism of a Peripheral Membrane Protein through Microsecond Molecular Dynamics Simulations. *J Mol Biol* **423**, 847-861 (2012).
43. T. Heimburg, D. Marsh, "Thermodynamics of the Interaction of Proteins with Lipid Membranes" in *Biological Membranes: A Molecular Perspective from Computation and Experiment*, K. M. Merz, B. Roux, Eds. (Birkhäuser Boston, Boston, MA, 1996), 10.1007/978-1-4684-8580-6_13, pp. 405-462.
44. Hanif M. Khan *et al.*, A Role for Weak Electrostatic Interactions in Peripheral Membrane Protein Binding. *Biophys J* **110**, 1367-1378 (2016).
45. M. B. Sankaram, D. Marsh, "Chapter 6 Protein-lipid interactions with peripheral membrane proteins" in *New Comprehensive Biochemistry*, A. Watts, Ed. (Elsevier, 1993), vol. 25, pp. 127-162.
46. M. Rothe *et al.*, Transient binding accounts for apparent violation of the generalized Stokes-Einstein relation in crowded protein solutions. *Phys Chem Chem Phys* **18**, 18006-18014 (2016).
47. P. Simonnin, B. Noetinger, C. Nieto-Draghi, V. Marry, B. Rotenberg, Diffusion under confinement: hydrodynamic finite-size effects in simulation. *J Chem Theory Comput* **13**, 2881-2889 (2017).
48. I. Bera, J. B. Klauda, Molecular Simulations of Mixed Lipid Bilayers with Sphingomyelin, Glycerophospholipids, and Cholesterol. *J Phys Chem B* **121**, 5197-5208 (2017).
49. R. B. Martin (2004) Bioinorganic Chemistry. in *Encyclopedia of Molecular Cell Biology and Molecular Medicine*, ed R. A. Meyers (Wiley-VCH).
50. J. C. Phillips *et al.*, Scalable Molecular Dynamics with NAMD. *J Comput Chem* **26**, 1781-1802 (2005).
51. S. Jo, T. Kim, V. G. Iyer, W. Im, CHARMM-GUI: A Web-Based Graphical User Interface for CHARMM. *J Comput Chem* **29**, 1859-1865 (2008).
52. S. Jo, J. B. Lim, J. B. Klauda, W. Im, CHARMM-GUI Membrane Builder for mixed bilayers and its application to yeast membranes. *Biophys J* **97**, 50-58 (2009).
53. D. E. Shaw *et al.* (2014) Anton 2: Raising the Bar for Performance and Programmability in a Special-Purpose Molecular Dynamics Supercomputer. in *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis* (IEEE Press, New Orleans, Louisiana), pp 41-53.

FIGURE LEGENDS

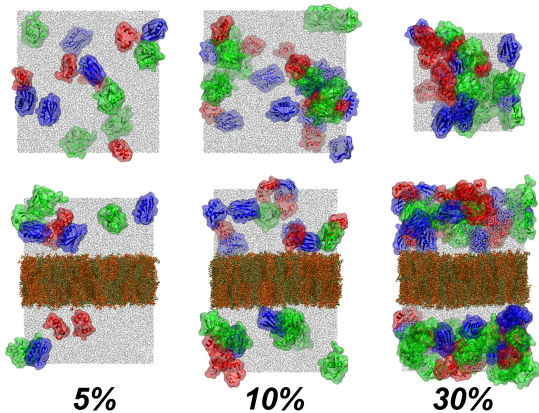
Figure 1 Systems and Protein Clustering. (A) Overview of simulated systems without (top) and with (bottom) phospholipid bilayers at protein volume fractions of 5, 10, and 30% with villin (red), protein G (blue), ubiquitin (green), POPC (tan), sphingomyelin (orange), cholesterol (dark brown), and water (grey); (B) Cluster size distributions between all proteins at 5% (purple), 10% (light blue), and 30% (tan) in the absence (solid lines) and presence (dashed lines) of the membrane based on protein contacts with minimum C α -C α distances <7 Å. A cluster size of 1 corresponds to monomers.

Figure 2 Protein Interactions. Normalized preferences for residue-residue interactions projected onto the molecular structures of villin (+2e), protein G (-4e), and ubiquitin (neutral) from simulation at 30% volume fraction in the presence of the membrane. The coloring in the leftmost column reflects residue types (polar – green, basic – blue, acidic – red, hydrophobic – white). Subsequent coloring reflects preferences for interactions with the membrane (yellow), villin (red), protein G (blue), or ubiquitin (green). More saturated colors indicate stronger preferences. The N- and C-termini are shown as blue and red spheres, respectively.

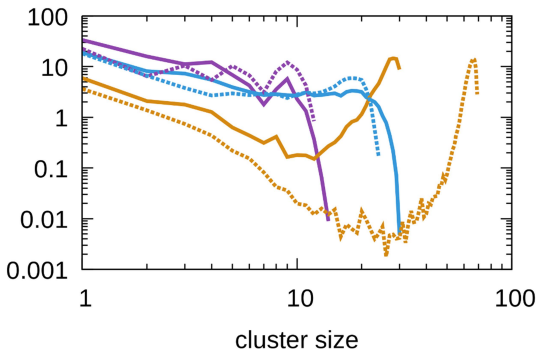
Figure 3 Density and Diffusion near the Membrane. Heavy-atom density distributions of molecular components (top), translational diffusion constants parallel to the membrane (center), and rotational diffusion constants (bottom) as a function of the distance from the membrane center along the membrane normal for systems with proteins at 5% (purple), 10% (light blue), and 30% (tan). Densities are shown for proteins (solid lines), lipids (long dashes), and water molecules (short dashes). Translational and rotational diffusion constants were assigned to the center of mass of a given protein at the beginning of the intervals for which mean-square displacements (MSD) and rotational correlation functions were obtained. Diffusion on shorter time scales, along the

membrane normal, and based on mid- or endpoints of diffusion time intervals is shown in *SI Appendix*, Figs S26-S28. Translational diffusion was estimated from MSD vs. time during 10-100 ns. Statistical errors for density distributions are less than 1%. The lipid bilayer projected at scale and a grey sphere at the size of protein G at the point of closest membrane contact are shown for perspective.

Figure 4 Crowding-Induced Membrane Deformations. (top): Membrane distortion as a function of protein interactions at 5% (purple), 10% (light blue), and 30% (tan). Protein-membrane distances are defined based on minimum heavy-atom distances between proteins and lipids. Membrane distortions are characterized by average phosphate distances from the membrane center for phosphate atoms within a 15 Å radius from the lipid atom in closest contact with the protein. The average phosphate distance to the center irrespective of any protein contact is indicated as a grey line. (bottom, left): Snapshot of a curved membrane (grey with purple phosphates) when in contact with villin (red) from the simulation at 30%. Other proteins are shown in blue; (bottom, right): projection of a sphere with a radius compatible with the induced curvature (green).

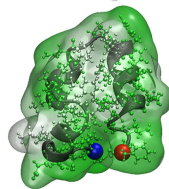
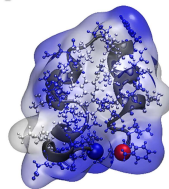
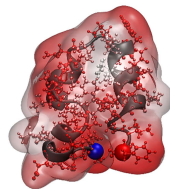
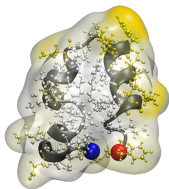
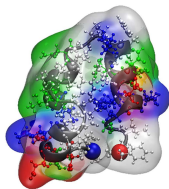
A**B**

percentage

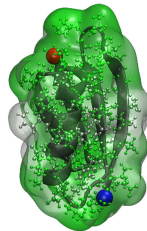
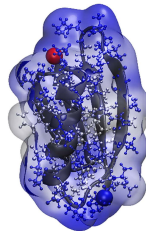
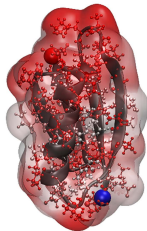
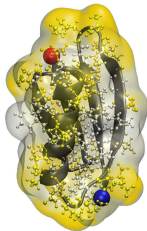
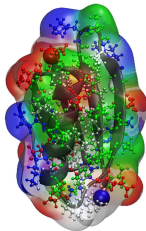


→ membrane → villin → protein G → ubiquitin

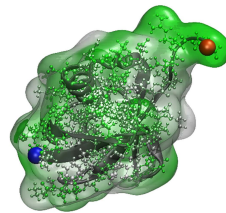
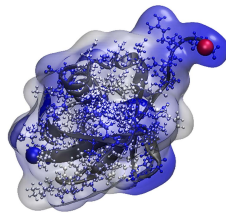
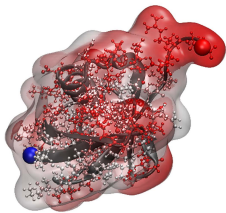
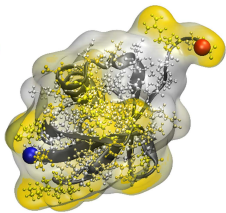
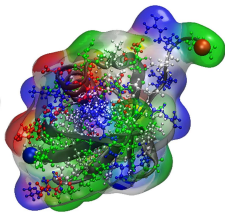
villin

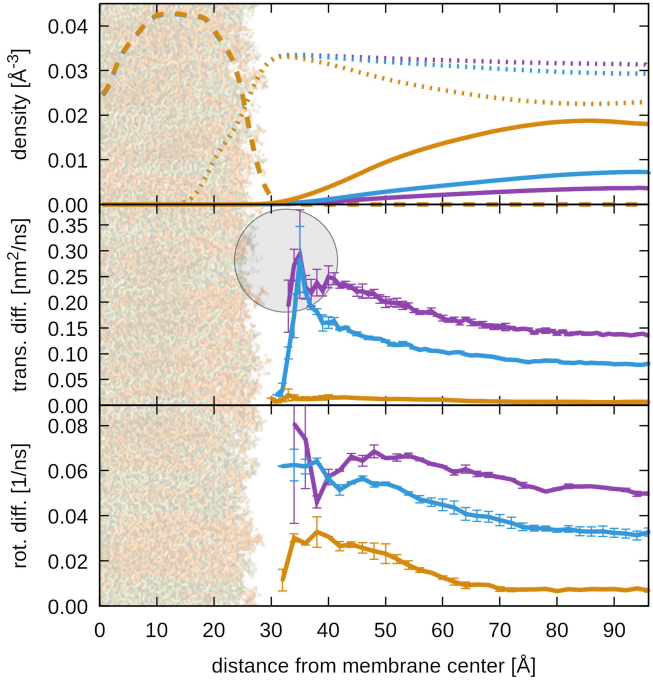


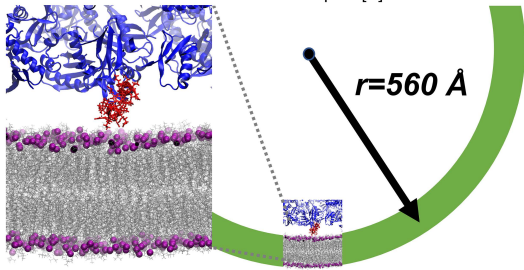
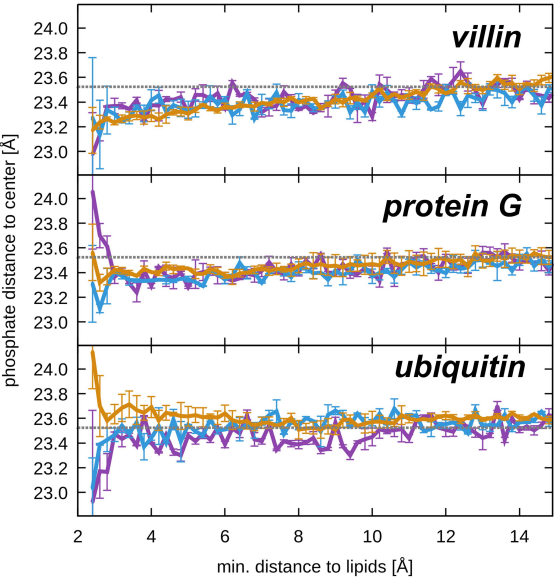
protein G



ubiquitin







Supplementary Information for

Clustering and Dynamics of Crowded Proteins near Membranes and Their Influence on Membrane Bending

Grzegorz Nawrocki¹, Wonpil Im², Yuji Sugita^{3,4,5}, Michael Feig^{1,5*}

¹Department of Biochemistry and Molecular Biology, Michigan State University, East Lansing, Michigan 48824, USA

²Department of Biological Sciences and Bioengineering Program, Lehigh University, Bethlehem, PA 18015, USA

³Cluster for Pioneering Research, RIKEN, Wako, Saitama 351-0198, Japan

⁴Center for Computational Research, RIKEN, Kobe, Hyogo 640-0047, Japan

⁵Center for Biosystems Dynamics Research, RIKEN, Kobe, Hyogo 640-0047, Japan

*Corresponding author.

Michael Feig

603 Wilson Road, Room 218 BCH

East Lansing, MI 48824, USA

E-mail: mfeiglab@gmail.com

Phone: 517-432-7439

This PDF file includes:

Supplementary methods

Figs. S1 to S34

Tables S1 to S8

References for SI reference citations

Supplementary Methods

System Setup. The systems were assembled using the CHARMM-GUI framework and Membrane Builder (1-3). Initial protein structures were taken from the experimental structures deposited in the Protein Data Bank: 1VII for villin (4), 3GB1 for protein G (5), and 1UBQ for ubiquitin (6). The orientations of the proteins were randomized before assembling the crowded systems by randomly packing proteins inside the cubic (non-membrane) or rectangular (membrane) system volumes. The number of proteins and system sizes were adjusted to limit the total number of atoms to less than 400,000 atoms so that sufficiently long simulations could be carried out on the Anton2 hardware. Therefore, the number of protein copies was decreased from 10% to 5% volume fraction instead of increasing the system size (Table S1). Furthermore, the membrane systems were constructed under the constraint of keeping the number of lipids (and therefore the x-y box dimension) constant for all protein concentrations. Snapshots of the initial systems are shown in Figure 1A.

Setup of NAMD and Anton2 simulations. The initial step involved four cycles of minimization with 50 steps each (250 steps for membrane systems) using the steepest descent and adopted basis Newton-Raphson methods. Minimization was followed by simulations at 300 K (303.15 K for membrane systems) under restraints on heavy atoms (with a force constant of 1.0 kcal/mol/Å² for backbone atoms and 0.1 kcal/mol/Å² for non-backbone atoms). In the membrane systems, water molecules were also restrained from entering the hydrophobic core of the lipid bilayer and lipids were restrained to remain oriented with the head groups near the water phase and the tails inside the hydrophobic core. For non-membrane systems, 10 ps were initially simulated in the NVT ensemble with a 1-fs time step, followed by 20 ps with a 2-fs time step, and another 20 ps in the NPT ensemble. The equilibration of the membrane systems involved longer simulations of two 250 ps runs in the NVT ensemble with a 1-fs time step followed by another 250 ps in the NPT ensemble before the time step was increased to 2 fs and another three simulations were carried out over 700 ps with decreasing restraints on lipids and water molecules.

Production runs on Anton2 were begun after 5000 steps of minimization and 500 ps of equilibration without restraints using NAMD. The 5% simulations were started from the systems at 10% after the initial equilibration after replacing half of the proteins with water molecules and adjusting ion concentrations accordingly.

Additional details for NAMD and Anton simulations. Proteins were described by a modified version of the CHARMM36 force field (7), where protein-water Lennard-Jones interactions were increased by a factor of 1.09 to avoid aggregation artefacts as introduced previously (8). Lipid interactions were described by the CHARMM36 lipid force field (9), and explicit water was modeled with the CHARMM version of the TIP3P water model (10). Protein-lipid interactions were not altered based on a previous study that found good agreement with experiment for protein-membrane interactions with the (unmodified) CHARMM36 protein and lipid parameters (11). Initial ion parameters (12) were modified via NBFIX based on osmotic pressure corrections (13). Energies were matched between Anton2 and the Desmond software and between Desmond and CHARMM to ensure correct implementation of the NBFIX modifications and enhanced water scaling.

Changes of the box size were controlled isotropically in the x and y dimensions in the membrane systems and for all dimensions in the non-membrane simulations. A Berendsen thermostat and barostat (14) were used in the NAMD simulations with a thermal coupling constant of $\tau=1$ ps and an isothermal compressibility $\beta=0.0000457$ bar⁻¹. A time step of 2 fs was used in the NAMD simulations in combination with SHAKE applied to all bonds involving hydrogen atoms. In the Anton2 simulations, integration was carried out via the “multigrator” algorithm (15) with a 2.5 ps time step applied to bonded and near-range non-bonded interactions.

Far-range interactions were evaluated every third time step based on a RESPA scheme. All bonds involving hydrogen atoms were constrained with M-SHAKE (16). Pressure was controlled according to a Martyna-Tobias-Klein (MTK) barostat (17) with an interval length of 1.2 ps and temperature was maintained via a Nose-Hoover thermostat (18) with an interval length of 60 fs. A relaxation time of $\tau=0.041667$ ps was used for baro- and thermostats.

In NAMD simulations, Lennard-Jones interactions were truncated with a switching function that was effective from 10 to 12 Å. Particle-mesh Ewald summation was used to estimate long-range electrostatic interactions with a grid spacing of 1 Å and a spline interpolation order of 6. In Anton2 simulations, Lennard-Jones interactions were truncated at 9 Å and the u-series version of Gaussian-split Ewald summation was used to calculate long-range Electrostatic interactions (19). The direct space cutoff was optimized for accuracy and speed based on system size and varied between 9.12 to 12.65 Å.

Analysis. The results presented here were extracted from the Anton2 simulations. The first 100 ns of each trajectory were omitted from analysis as equilibration based on variations in contact formation at the beginning of the trajectories (see Figure S10). Statistical uncertainties were estimated from variations in the reported results between different protein copies (where possible) or from block averaging along the trajectory.

Protein contacts were determined based on two different criteria to facilitate comparisons with previous work: 1) Closest C α -C α distances <7 Å (8) and 2) closest heavy atom distances <2.7 Å (an additional criterion introduced when analyzing the detailed effect of protein contacts on the rotational diffusion of villin (20)). The parameters were chosen so that cluster-size distributions extracted from all-atom simulations based on protein contacts were similar independent of the chosen criterion. Clusters were determined from contacts based on any protein being in contact with at least one other protein in the cluster.

Translational and rotational diffusion coefficients were evaluated as described previously (8). To describe translational diffusion, the mean square displacement (MSD) was calculated as a function of time from $\text{MSD}(\tau) = \langle (\mathbf{r}(\mathbf{t}_0 + \tau) - \mathbf{r}(\mathbf{t}_0))^2 \rangle_{t_0}$ based on the center of mass of a given protein after unwrapping coordinates due to periodic boundary conditions, $\mathbf{r}(t)$, and averaged over different initial times t_0 along the trajectory. From the slope of the linear fits over a given time interval $\Delta\tau$, $s(\Delta\tau)$, translational diffusion coefficients were then obtained according to the Einstein relationship: $\mathbf{D}_t = \mathbf{s}(\Delta\tau)/(2N\Delta\tau)$ where N is the dimension ('3' for 3D diffusion, '2' for diffusion in the x-y plane, and '1' for diffusion along the membrane normal z). In order to analyze transient and anomalous behavior of D_t , we considered time intervals $\Delta\tau$ of 0 – 1 ns, 1 ns – 10 ns, and 10 ns – 100 ns. Translational diffusion coefficients were estimated separately for each protein and as a function of distance from the membrane based only on the position of the protein at time t_0 , irrespective of where it may diffuse afterwards.

Estimates of translational diffusion are subject to finite-size effects in the presence of periodic boundary conditions. To obtain infinite-size values, we added the correction term given in Eq. 1 (21):

$$D_{t,PBCcorr} = \frac{k_B T}{6\pi\eta L} \left(\xi - \frac{4\pi R_h^2}{3L^2} \right) \quad (1)$$

where $\xi = 2.837$, k_B is the Boltzmann constant, T the temperature of the system, L is the length of the cubic simulation box, η is the shear viscosity of the solvent, and R_h is the hydrodynamic radius of a given molecule, estimated with HYDROPRO (22), as $R_h(\text{villin}) = 9$ Å, $R_h(\text{protein G}) = 11$ Å, and $R_h(\text{ubiquitin}) = 12$ Å.

For crowded systems, the viscosity was further adjusted from the viscosity of pure solvent, η_w . Instead of a simple hard-sphere based estimate used earlier (8, 23) for how the effective viscosity varies with the protein volume fraction, ϕ , we used here the expanded formalism in Eq. 2 that was introduced recently (24):

$$\eta = \eta_w (1 + 2.5\phi + b\phi^2) \quad (2)$$

where the parameter b captures effective increases in viscosity not just due to volume exclusion but also as a result of increased clustering (8, 24). Von Bülow et al. determined different values of b for different proteins depending on their propensity to form clusters (24), but because similar values were found for villin, protein G, and ubiquitin, the three proteins simulated here as well in a mixture, we used an average value of 58.2.

In the presence of the membrane, there are additional confinement effects to consider. Following the analysis by Simonnin et al. (25), the finite size correction for translational diffusion parallel to the membrane in a fluid that is infinite in x-y directions but constrained in the z direction by a membrane is:

$$D_{t,\parallel,PBCcorr} = \frac{k_B T}{\eta} \left(\frac{3 \ln(1+\sqrt{2})}{4\pi L} - \frac{3H}{40L^2} \right) \quad (3)$$

where L is the box size in x and y directions, and H is the width of the fluid slab that is calculated as the box length in the z direction minus the width of the membrane layer ($2 \times 27 \text{ \AA} = 54 \text{ \AA}$). Other parameters were set as in Eq. 1.

The diffusion parallel to the membrane in a slab of width H for infinite box dimensions in the x-y directions that is obtained after correction based on Eq. 3 can then be compared to bulk diffusion in the absence of a constraining slab (i.e., $H=\infty$) according to Eq. 4 (25):

$$D_{\parallel}(H, L = \infty) = D_{bulk} \left(1 + \frac{9RH}{8H} \ln \left(\frac{RH}{H} \right) \right) \quad (4)$$

Rotational diffusion coefficients were estimated following the method introduced by Wong and Case (26), where randomly distributed unit vectors are rotated along with the protein and a correlation function is obtained based on the rotation of the vectors. The correlation functions up to 100 ns (up to 24 ns for z-dependent analysis of rotational diffusion) were then fitted with

double-exponential functions $C_O(t) = S_R^2 e^{-\frac{t}{\tau_{Rs}}} + (1 - S_R^2) e^{-\frac{t}{\tau_{Rf}}}$ to obtain slow and fast correlation times, τ_{Rf} and τ_{Rs} , weighted by S_R^2 . An overall relaxation time τ was determined according to $\tau = \left(\frac{S_R^2}{\tau_{Rs}} + \frac{1-S_R^2}{\tau_{Rf}} \right)^{-1}$ and the rotational diffusion coefficient D_r was calculated as

$D_r = 1/6\tau$. Because rotational correlation functions converge slowly for a single protein, we averaged correlation functions from multiple proteins before fitting the exponential functions. Statistical uncertainties were estimated from comparing results obtained with different subsets of proteins. We did not correct the rotational diffusion estimates for periodic boundary conditions (27) because there is little change for the large systems studied here.

The MMTSB Tool Set (28), analysis functions in CHARMM (29), and custom-written programs in C/C++ and perl were used for all of the analysis. VMD (30) and gnuplot (<http://www.gnuplot.info>) were used for visualization and plotting. Gnuplot was also used for fitting linear and exponential functions for the determination of diffusion coefficients.

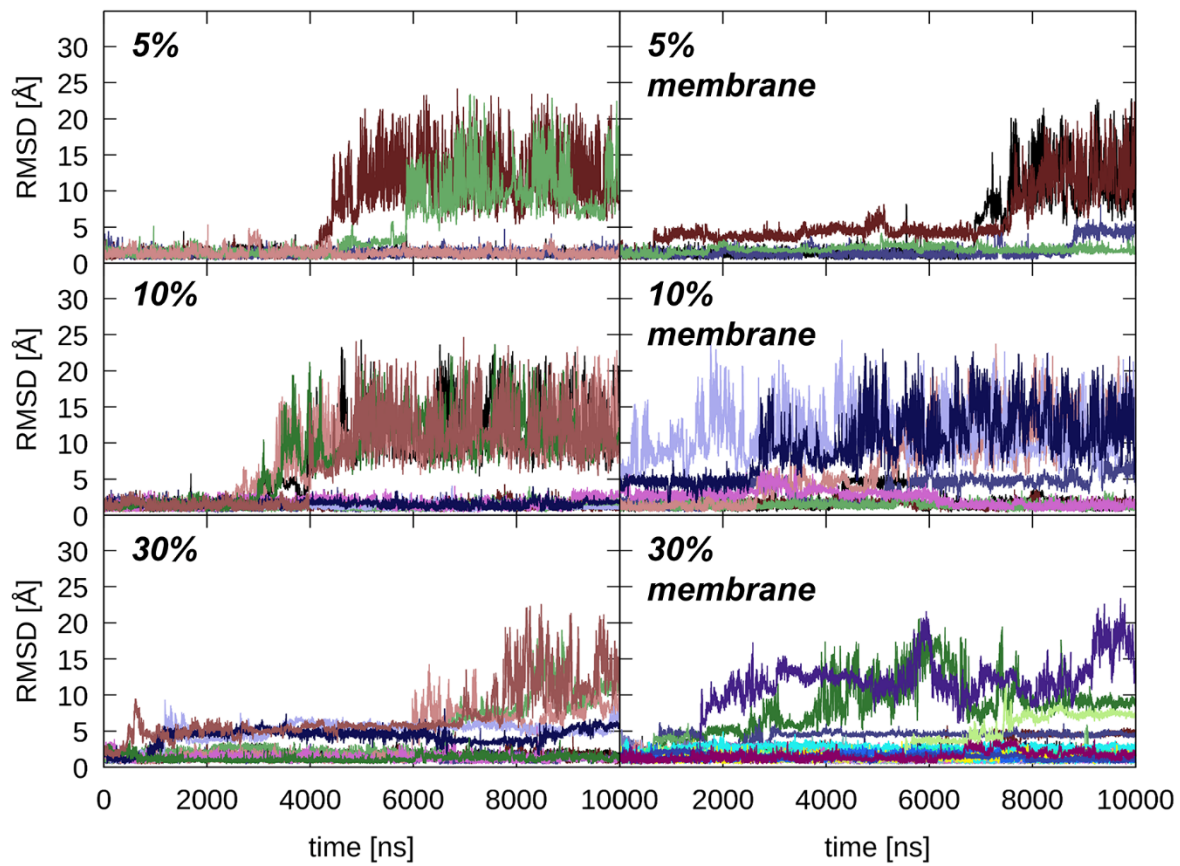


Figure S1: Root mean square deviation (RMSD) of villin molecules for C α atoms after optimal structural superposition compared to the experimental reference structure (PDB code: 1WY3 (31)) as a function of simulation time in the Anton2 simulations. Different colors indicate different molecules.

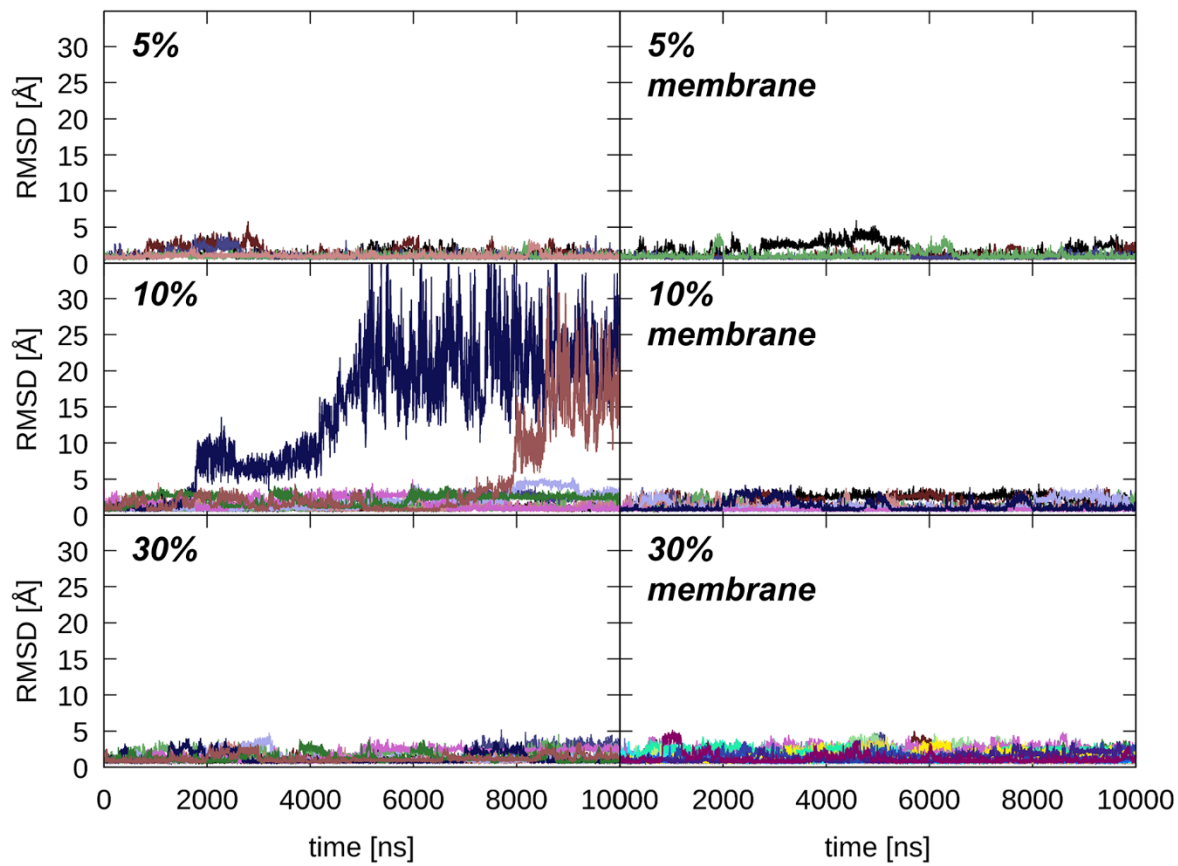


Figure S2: RMSD of protein G molecules for C α atoms after optimal structural superposition compared to the experimental reference structure (PDB code: 3GB1 (5)) as a function of simulation time in the Anton2 simulations. Different colors indicate different molecules.

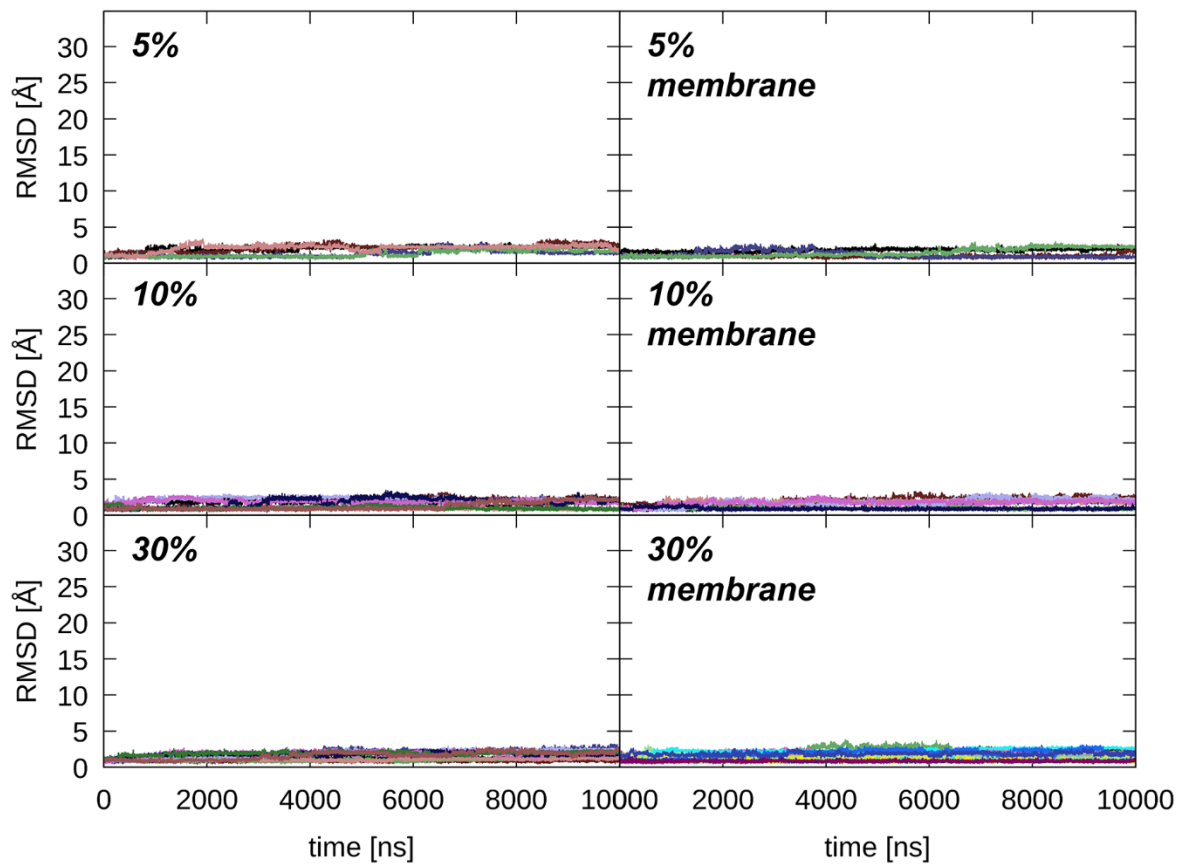


Figure S3: RMSD of ubiquitin molecules for C α atoms after optimal structural superposition compared to the experimental reference structure (PDB code: 1UBQ (6)) as a function of simulation time in the Anton2 simulations. Only residues 1-72 were considered to exclude the flexible C-terminus. Different colors indicate different molecules.

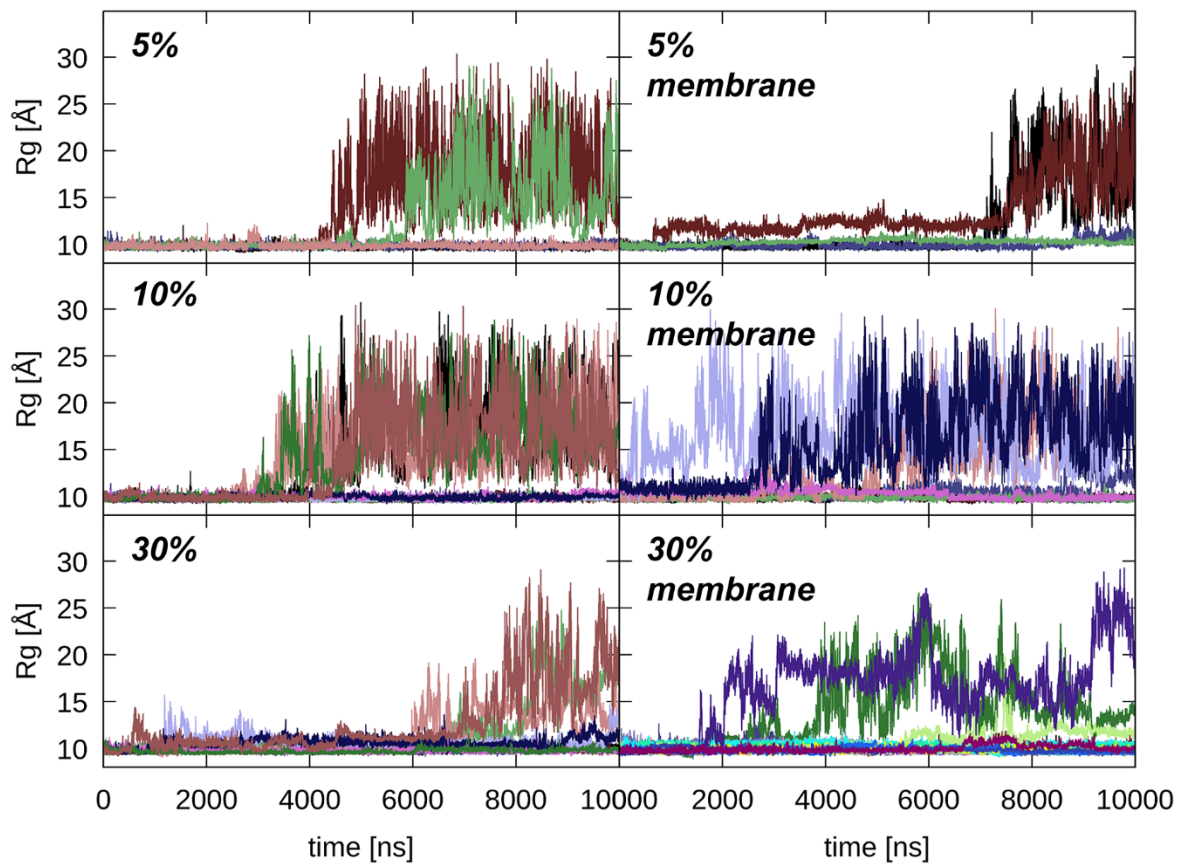


Figure S4: Radius of gyration R_g of villin molecules based on C α atoms as a function of simulation time in the Anton2 simulations. Different colors indicate different molecules.

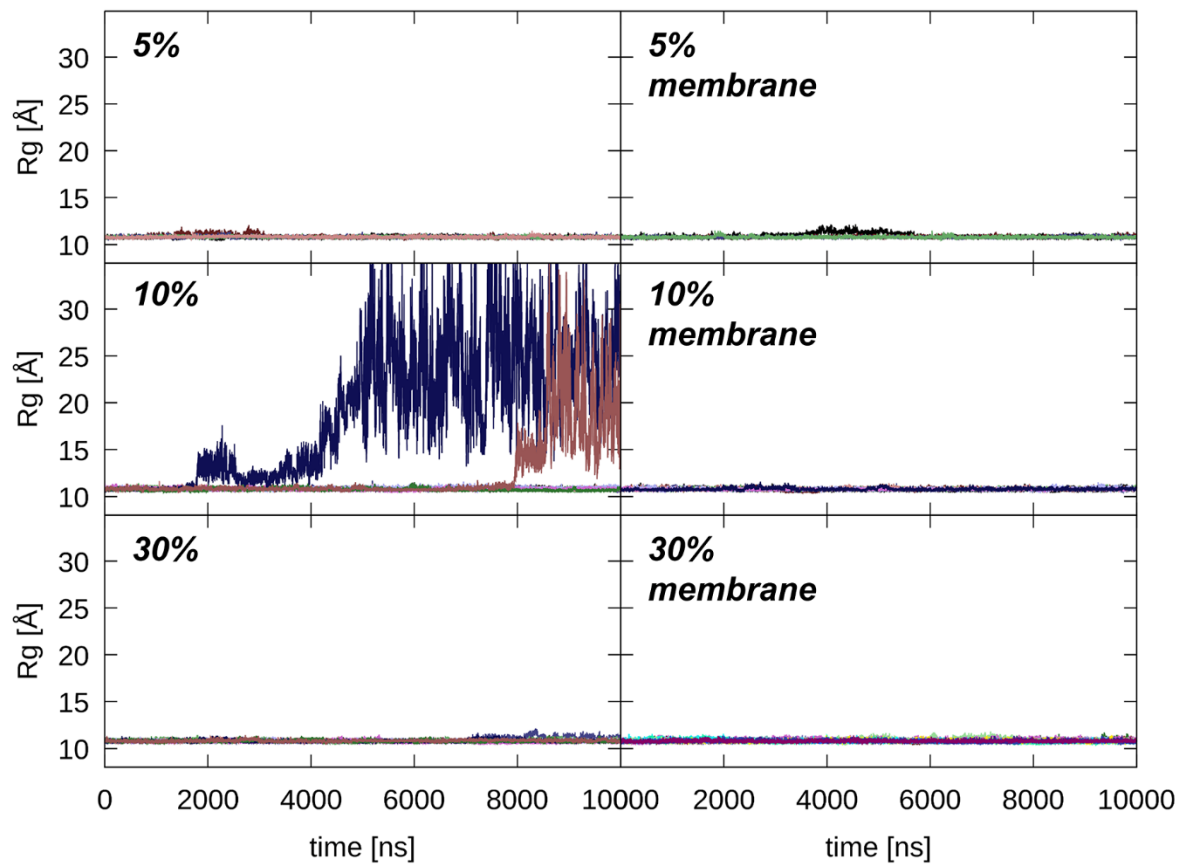


Figure S5: R_g of protein G molecules based on C α atoms as a function of simulation time in the Anton2 simulations. Different colors indicate different molecules.

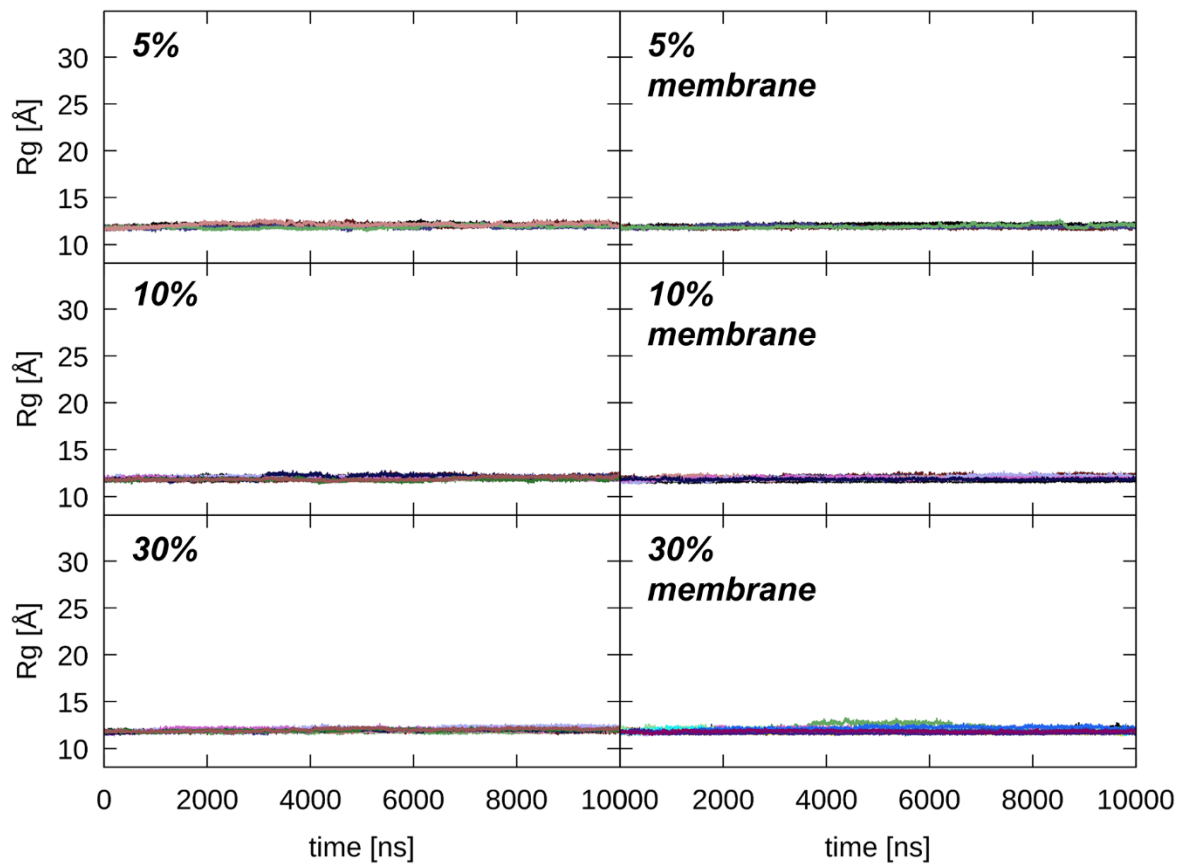


Figure S6: R_g of ubiquitin molecules based on $C\alpha$ atoms as a function of simulation time in the Anton2 simulations. Different colors indicate different molecules.

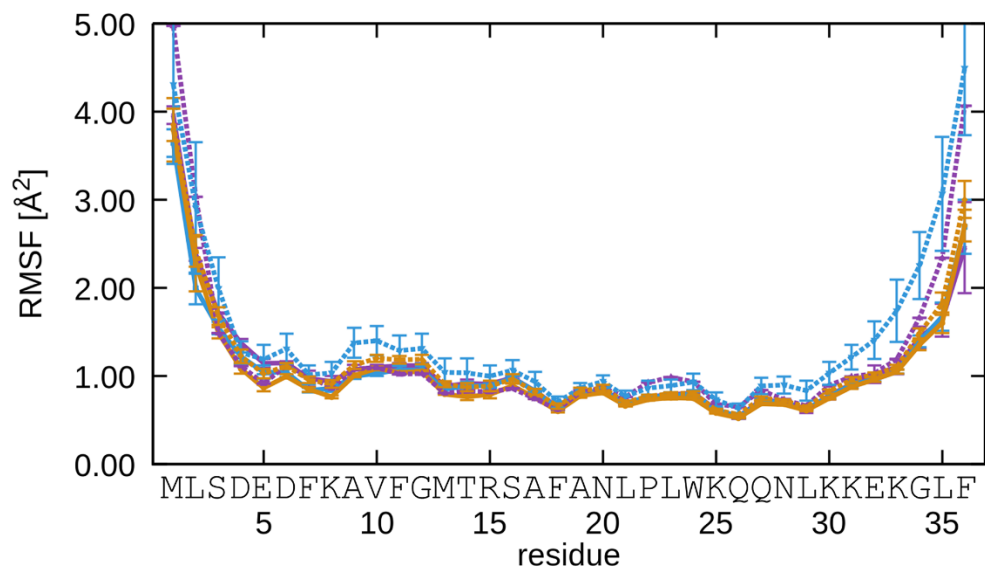


Figure S7: Root mean square fluctuations (RMSF) for $C\alpha$ atoms in villin molecules at 5% (purple), 10% (light blue), and 30% (tan) protein concentration in the absence (solid lines) and presence of the membrane (dashed lines). Error bars reflect uncertainties obtained from variations between different protein molecules in a given system. Only protein molecules where the $C\alpha$ RMSD at the end of the trajectory was below 2.5 Å were included.

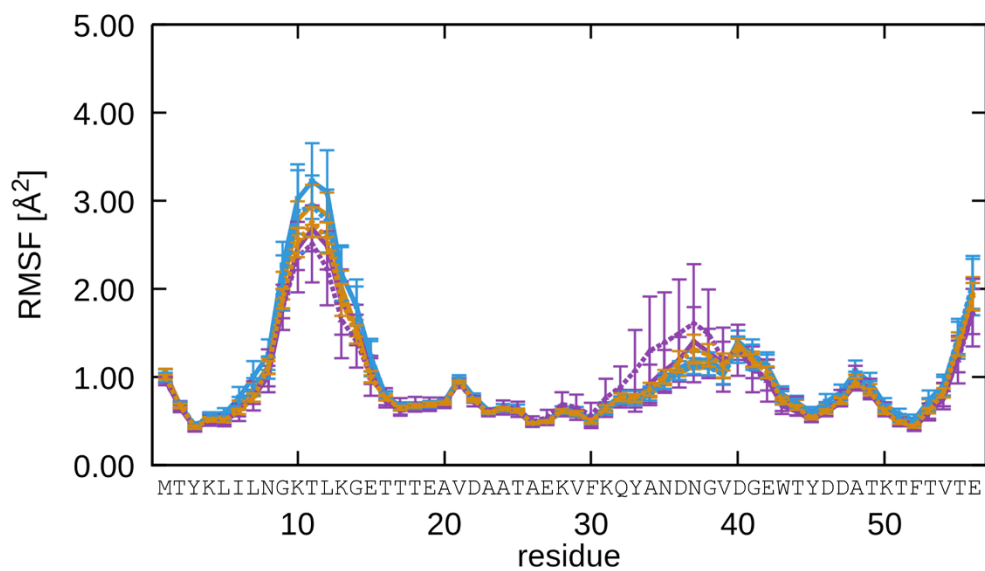


Figure S8: Root mean square fluctuations (RMSF) for C α atoms in protein G molecules at 5% (purple), 10% (light blue), and 30% (tan) protein concentration in the absence (solid lines) and presence of the membrane (dashed lines) as in Figure S7

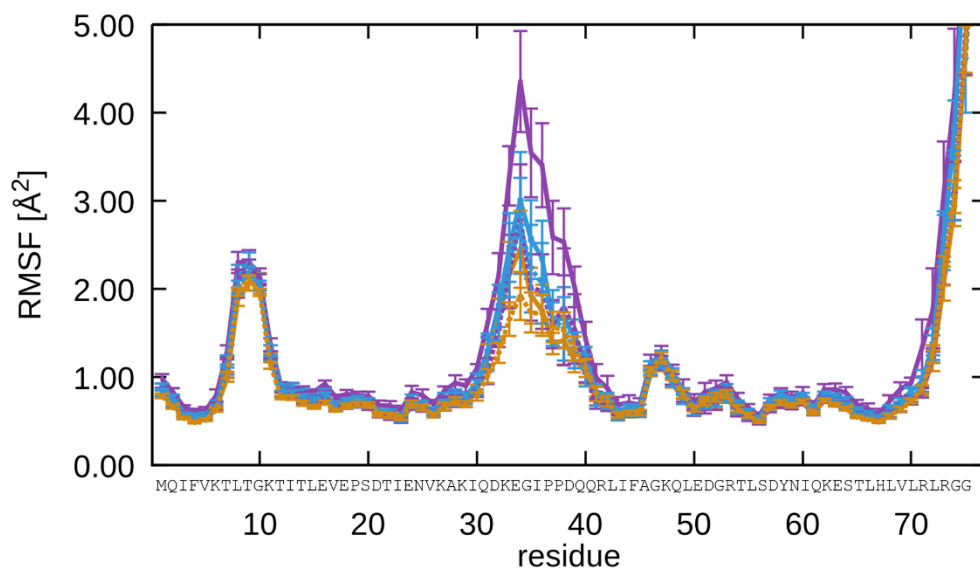


Figure S9: RMSF for C α atoms in ubiquitin molecules at 5% (purple), 10% (light blue), and 30% (tan) protein concentration in the absence (solid lines) and presence of the membrane (dashed lines) as in Figure S7.

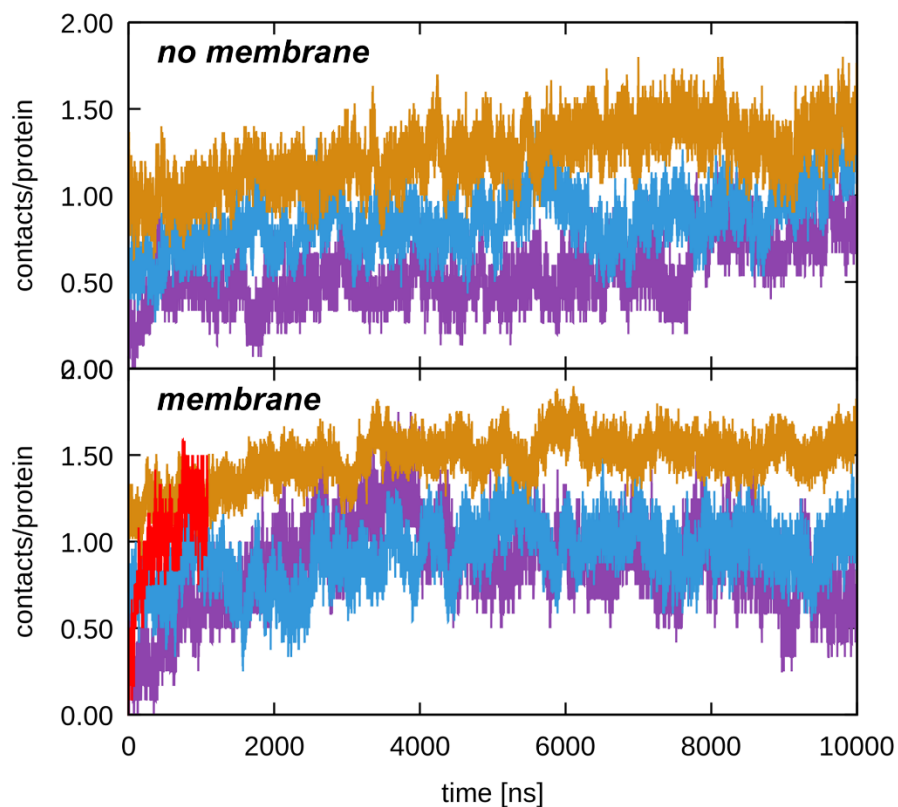


Figure S10: Protein contacts normalized by number of proteins as a function of simulation time at 5% (purple), 10% (light blue), and 30% (tan) protein concentration in the absence (top) and presence of the membrane (bottom). Results from a simulation at 5% with the original CHARMM c36 force field without scaling protein-water interactions are shown in red. Contacts were defined as minimum $C\alpha$ distances of less than 7 Å.

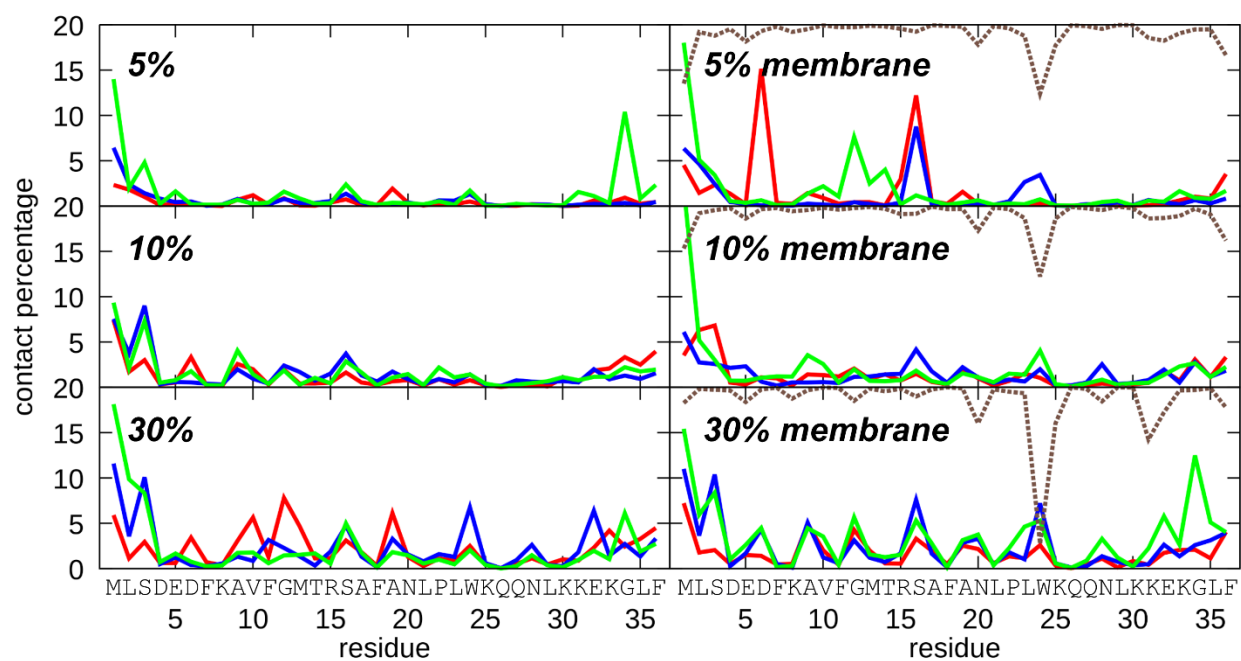


Figure S11: Percentage of residues (x10) involved in minimum heavy-atom distances below 5 Å between villin and other villin (red), protein G (blue), ubiquitin (green), or the membrane (black dashed line, shown as 20-10*percentage).

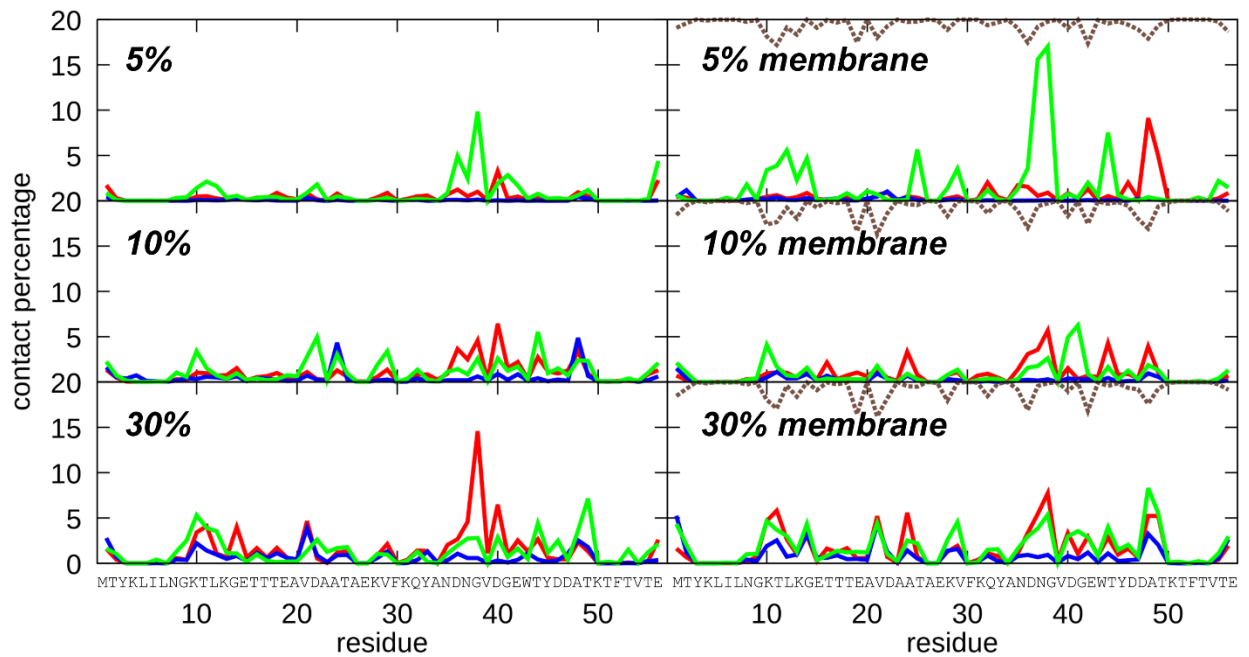


Figure S12: Percentage of residues (x10) involved in minimum heavy-atom distances below 5 Å between protein G and villin (red), other protein G (blue), ubiquitin (green), or the membrane (black dashed line, shown as 20-10*percentage).

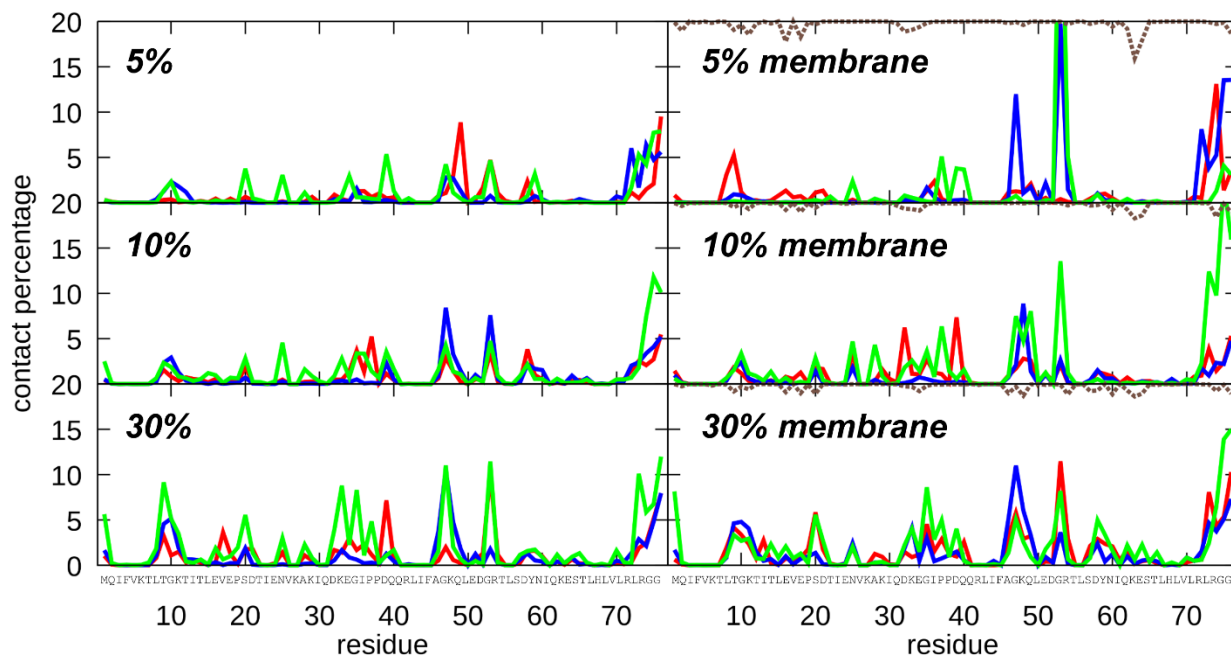


Figure S13: Percentage of residues (x10) involved in minimum heavy-atom distances below 5 Å between ubiquitin and villin (red), protein G (blue), other ubiquitin (green), or the membrane (black dashed line, shown as 20-10*percentage).

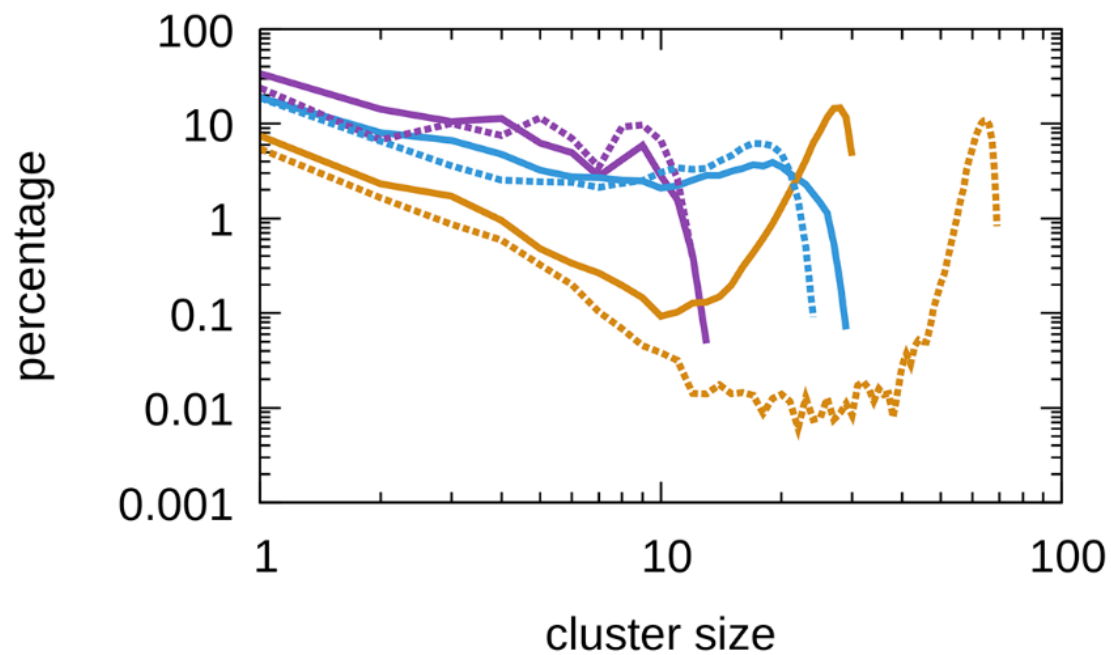


Figure S14: Cluster size distributions between all proteins at 5% (purple), 10% (light blue), and 30% (tan) in the absence (solid lines) and presence (dashed lines) of the lipid bilayer based on protein contacts defined as minimum heavy atom distances of less than 2.7 Å.

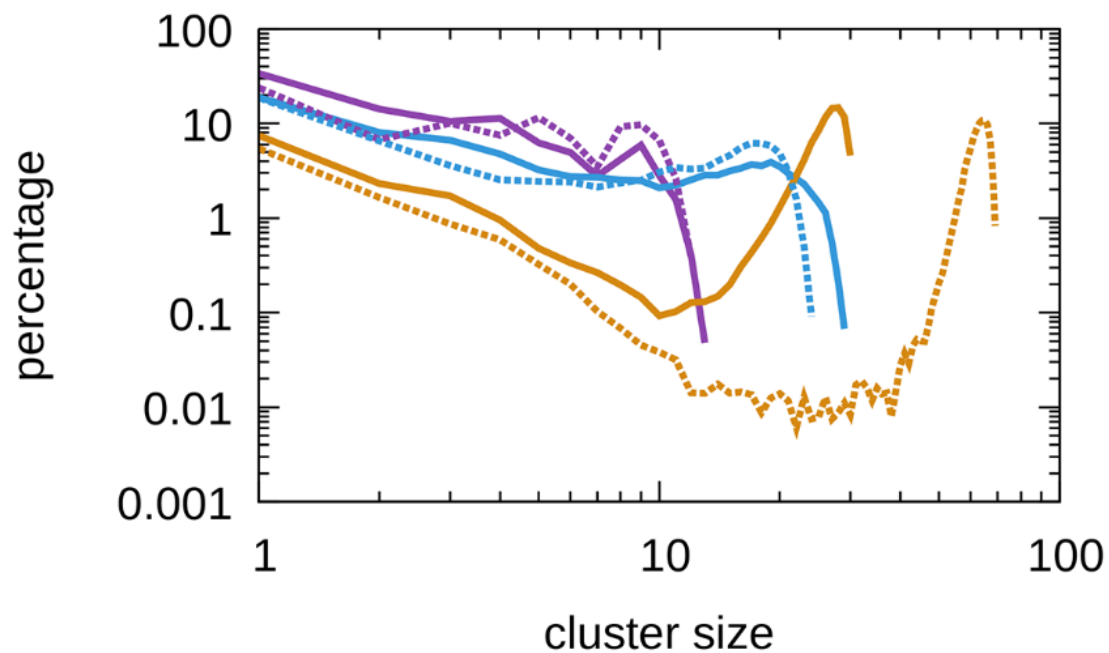


Figure S15: Cluster size distributions between all proteins at 5% (purple), 10% (light blue), and 30% (tan) in the absence (solid lines) and presence (dashed lines) of the lipid bilayer based on protein contacts defined as minimum $C\alpha$ - $C\alpha$ atom distances of less than 7 Å as in Figure 1B but using only data from the first 2 μ s in order to exclude frames with partially unfolded proteins.

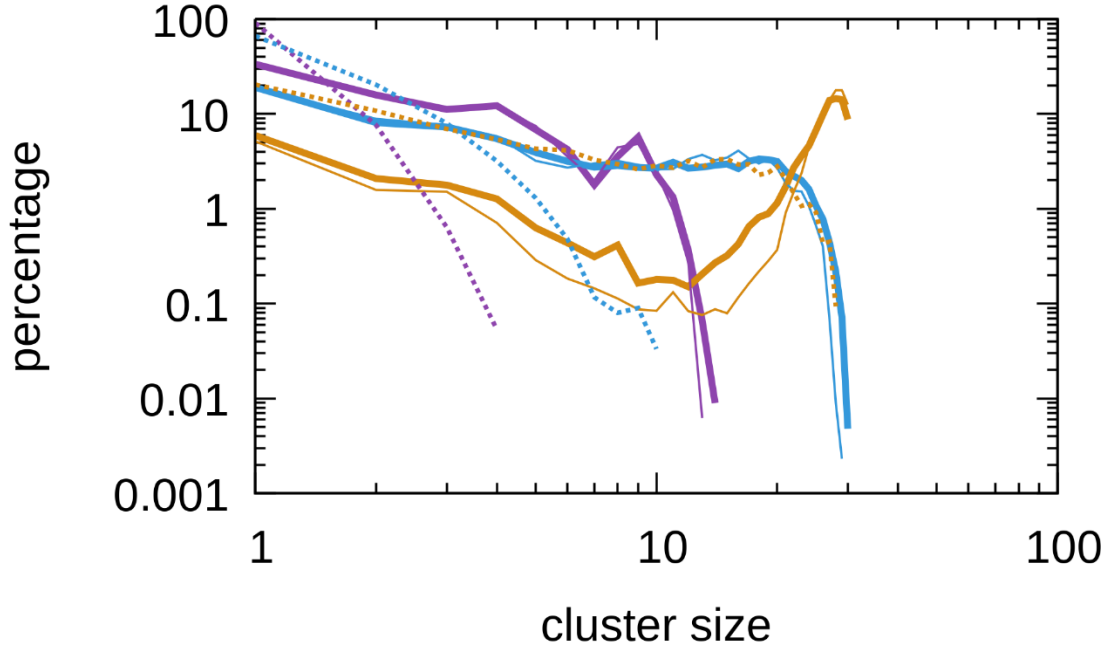


Figure S16: Cluster size distributions between all proteins in the absence of a membrane surface at 5% (purple), 10% (light blue), and 30% (tan) based on all-atom simulations using a C α -C α distance criterion of 7 Å (thick solid lines), based on all-atom simulations using an equivalent center-of-mass based criterion (thin solid lines), and based on coarse-grained (CG) simulations using the same center-of-mass based criterion (dashed lines). According to the center-of-mass criterion two proteins A and B were considered in contact when the distance between their centers was less than $(\sigma_A + \sigma_B)/2 + 7\text{Å}$ with values for σ given below. The CG simulations involved spherical models with 10 particles for each protein type at box volumes that were adjusted to result in the same volume fractions as in the atomistic simulations. The interaction potential for the CG model consisted of a purely repulsive short-range Lennard-Jones type potential as in the work by Mani et al. (32): $U_{ij}(r_{ij}) = 4\epsilon \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{100}$. The value of ϵ was set to 4 kJ/mol for protein-protein interactions. The size-dependent parameter σ_{ij} was calculated as $\sigma_{ij} = (\sigma_i + \sigma_j)/2$ with σ_i determined from the diameter of a sphere with a volume equivalent to the molecular volume of molecule i . Specifically, we set $\sigma_{\text{villin}}=18.13\text{ Å}$, $\sigma_{\text{proteinG}}=20.67\text{ Å}$, and $\sigma_{\text{ubiquitin}}=23.34\text{ Å}$. The CG simulations were run with OpenMM for 1 μs at 298 K using a Langevin thermostat.

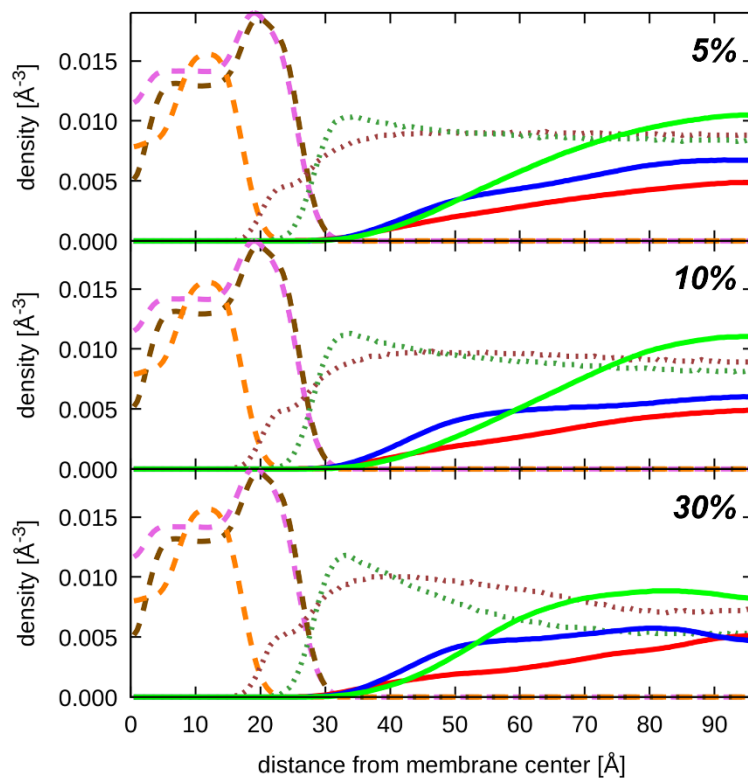


Figure S17: Heavy-atom density distributions along the membrane normal for different proteins (solid lines, villin – red, protein G – blue, ubiquitin – green), individual lipid types (long dashed lines; POPC – dark brown, sphingomyelin - purple, cholesterol – orange), and ions (short dashed lines, Na^+ - red, Cl^- - green). Ion concentrations are shown at 100x of the actual densities. Protein concentrations are shown at 3x for 10% protein concentration and 6x for 5% concentration to facilitate comparisons.

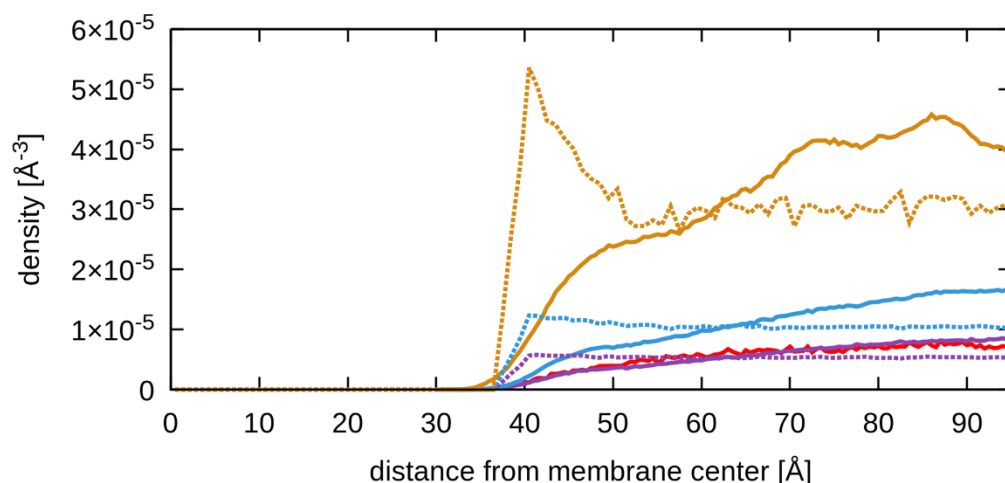


Figure S18: Density distribution of all protein center of mass positions at 5% (purple), 10% (light blue), and 30% (tan) from all-atom (solid lines) and hard-sphere (dashed lines) simulations. Results from a simulation at 5% with the original CHARMM c36 force field of the protein-membrane system without scaling protein-water interactions are shown in red. The hard-sphere results were obtained from coarse-grained (CG) simulations using spherical models with 10 particles for each protein type at box volumes as described in the caption for Figure S16. Interactions with the membrane were also repulsive using the same potential form based on the distance from the membrane plane along the membrane normal. The membrane plane was set at 2.8 nm from the center of the membrane. The value of ϵ was set to 4 kJ/mol for protein-membrane interactions. For interactions with the membrane, the value of σ for different proteins was used directly. The increase in density near the membrane with the CG model, especially at the higher concentrations, has been observed before and is due to anisotropic collisions with other molecules at a planar surface (33, 34).

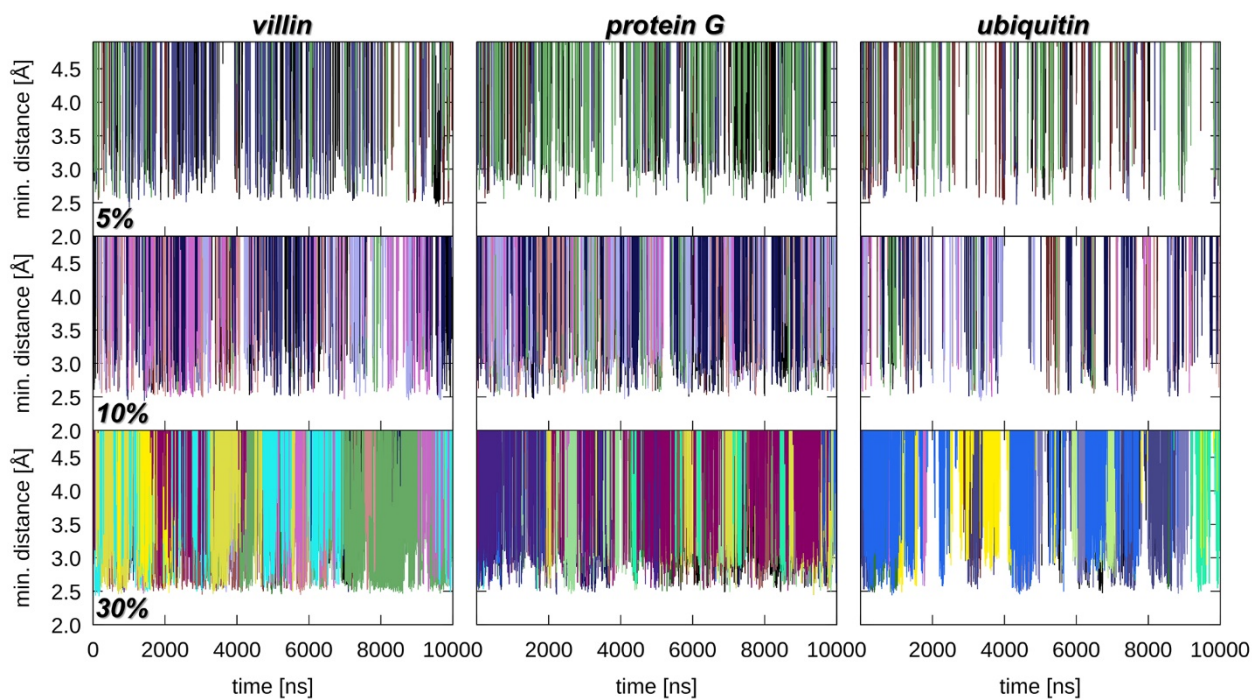


Figure S19: Minimum heavy-atom distances between individual proteins and membrane lipids as a function of simulation time. Different colors indicate different proteins.

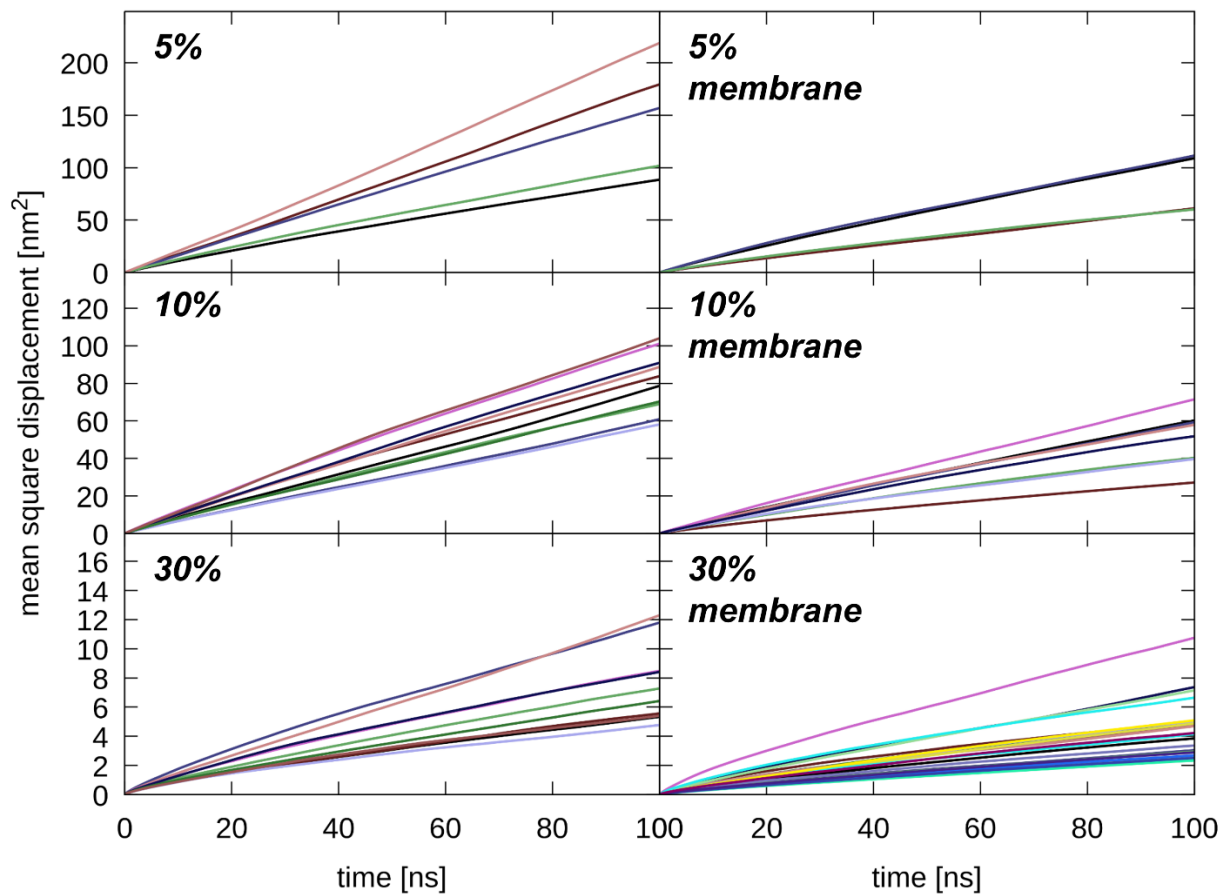


Figure S20: Mean-square displacement (MSD) curves for individual villin molecules (shown with different colors) in the absence (left) and presence (right) of a membrane bilayer at total protein concentrations of 5%, 10%, and 30%.

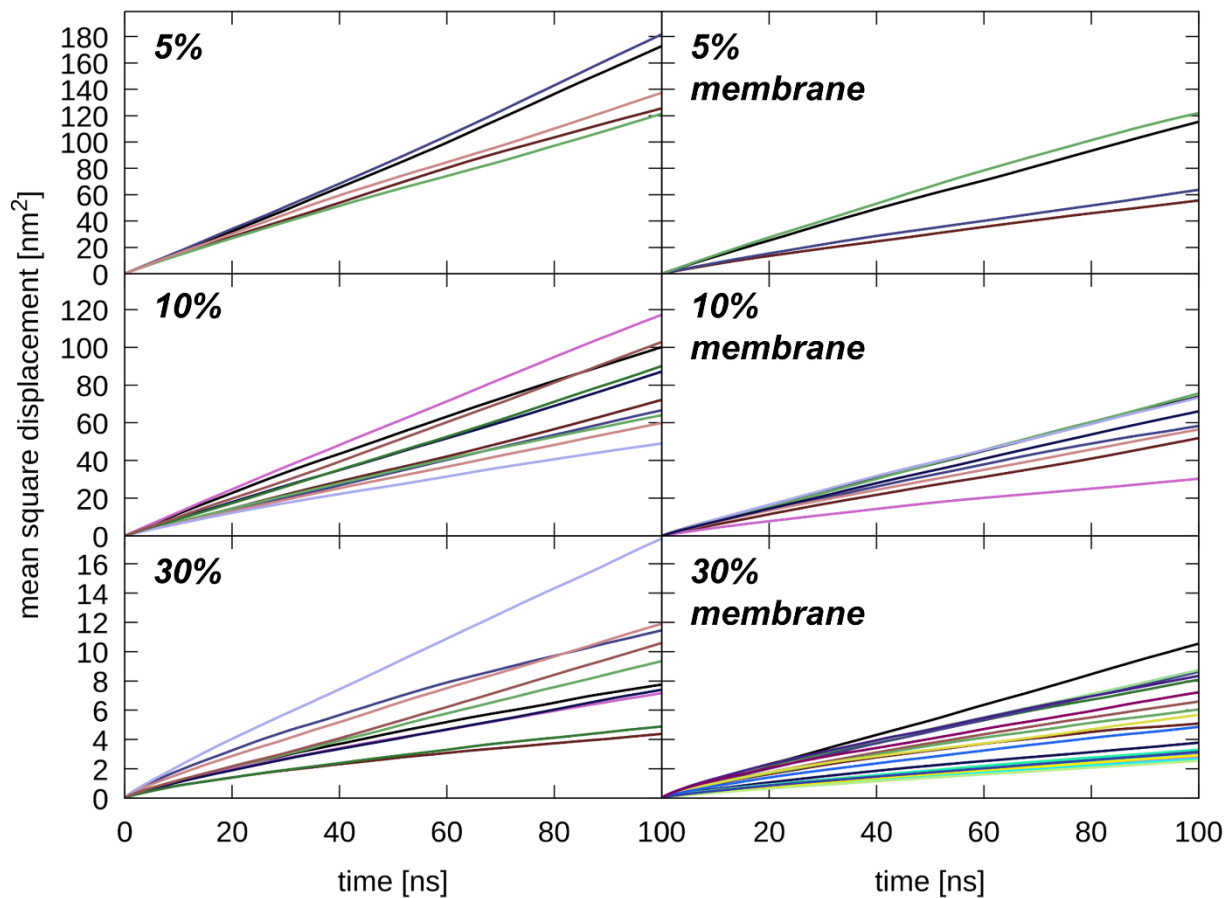


Figure S21: MSD curves for individual protein G molecules (shown with different colors) in the absence (left) and presence (right) of a membrane bilayer at total protein concentrations of 5%, 10%, and 30%.

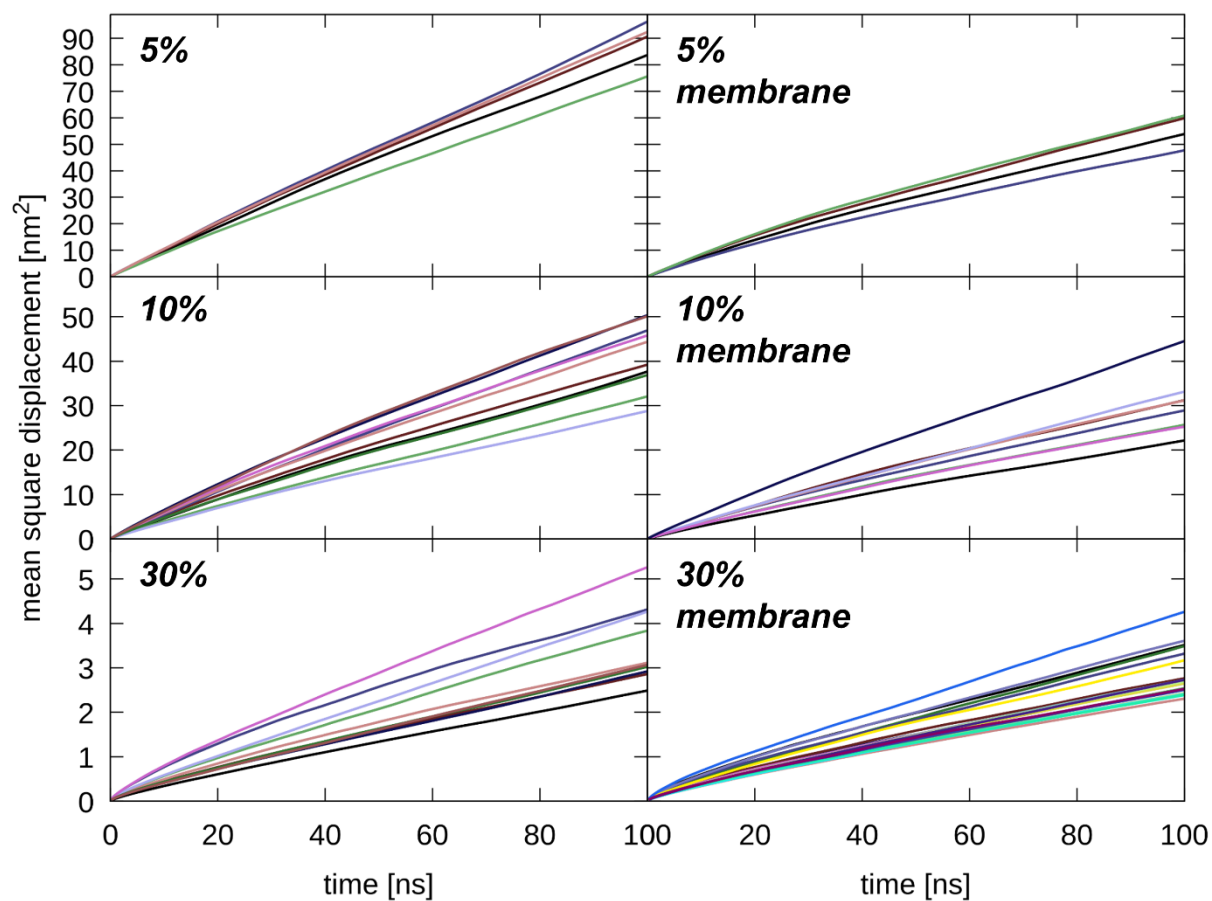


Figure S22: MSD curves for individual ubiquitin molecules (shown with different colors) in the absence (left) and presence (right) of a membrane bilayer at total protein concentrations of 5%, 10%, and 30%.

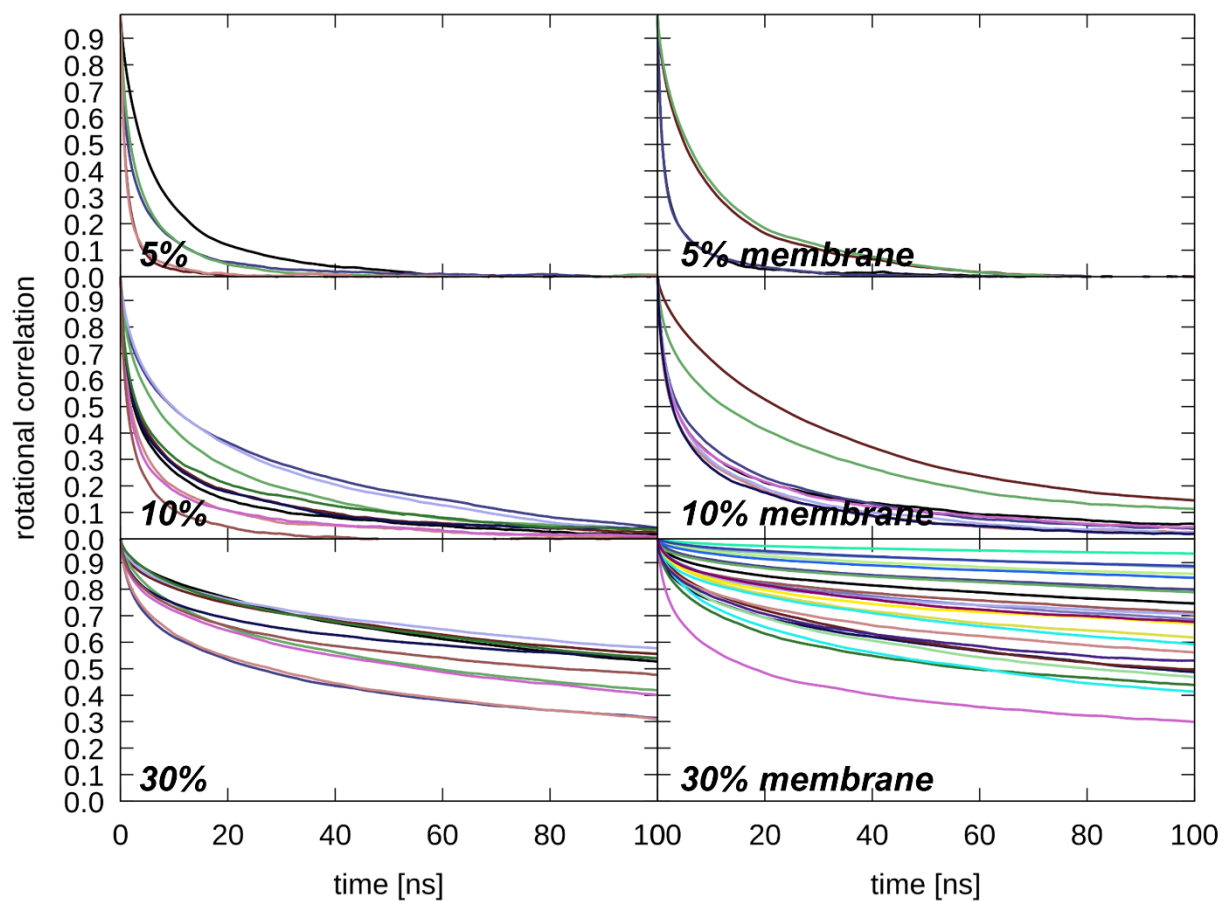


Figure S23: Rotational correlation functions for individual villin molecules (shown with different colors) in the absence (left) and presence (right) of a membrane bilayer at total protein concentrations of 5%, 10%, and 30%.

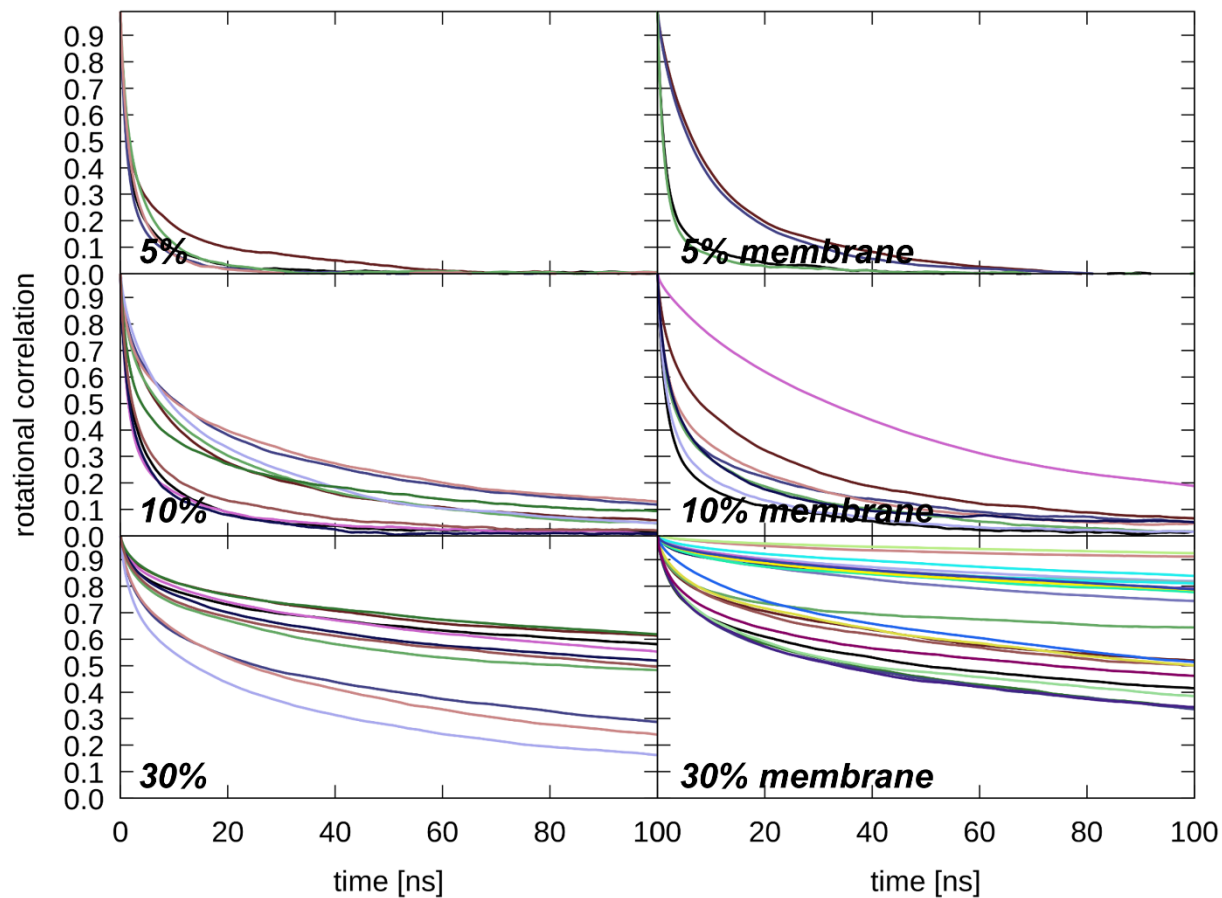


Figure S24: Rotational correlation functions for individual protein G molecules (shown with different colors) in the absence (left) and presence (right) of a membrane bilayer at total protein concentrations of 5%, 10%, and 30%.

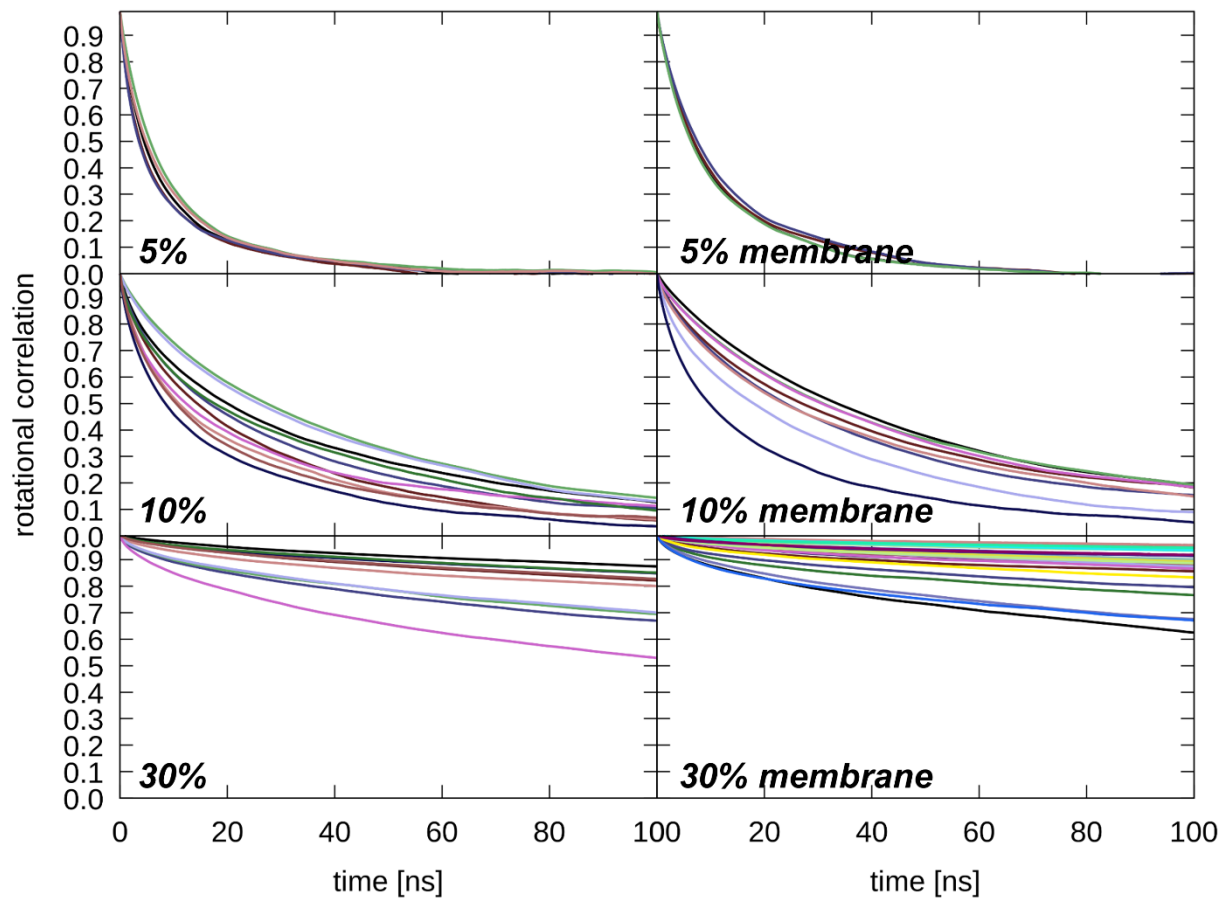


Figure S25: Rotational correlation functions for individual ubiquitin molecules (shown with different colors) in the absence (left) and presence (right) of a membrane bilayer at total protein concentrations of 5%, 10%, and 30%.

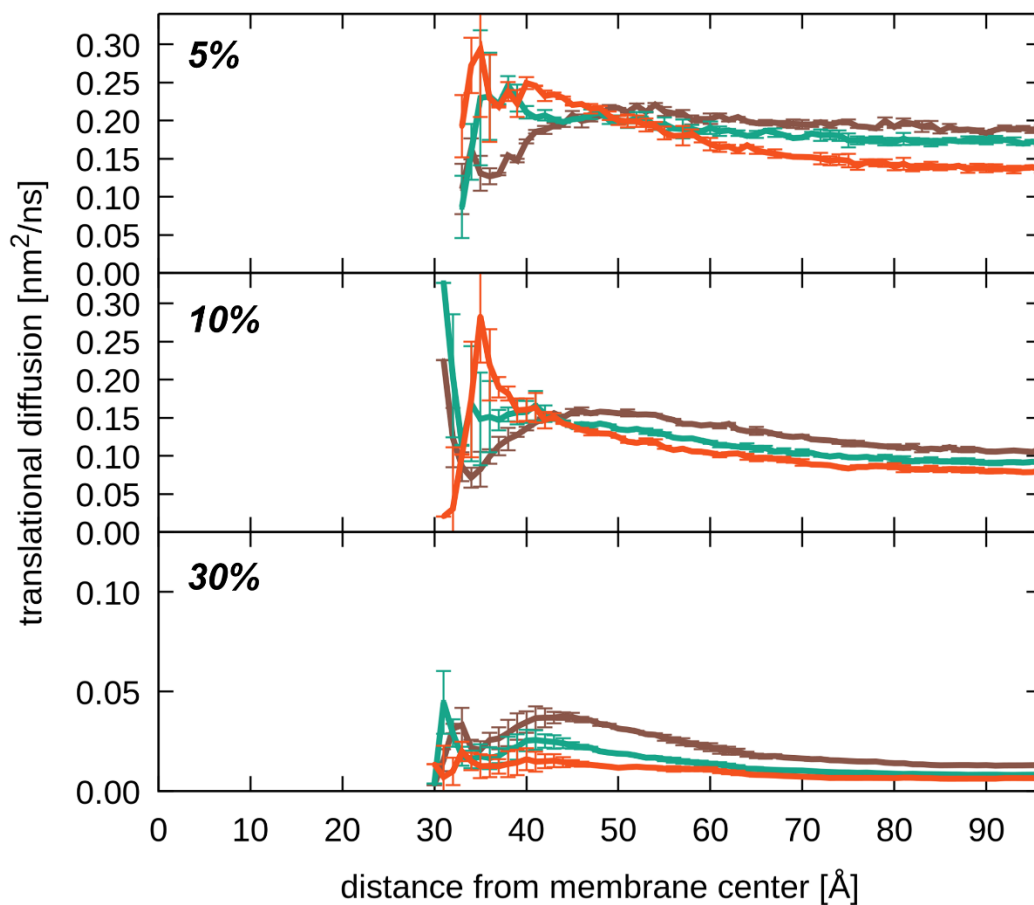


Figure S26: Translational diffusion parallel to the membrane (in the x-y plane) as a function of the distance from the membrane center extracted from mean-square displacement curves at different time scales: 0-1 ns (dark brown), 1-10 ns (green), 10-100 ns (red) for protein concentrations of 5%, 10%, and 30%. MSD curves were generated by combining data from all proteins and the results are not corrected for PBC artefacts.

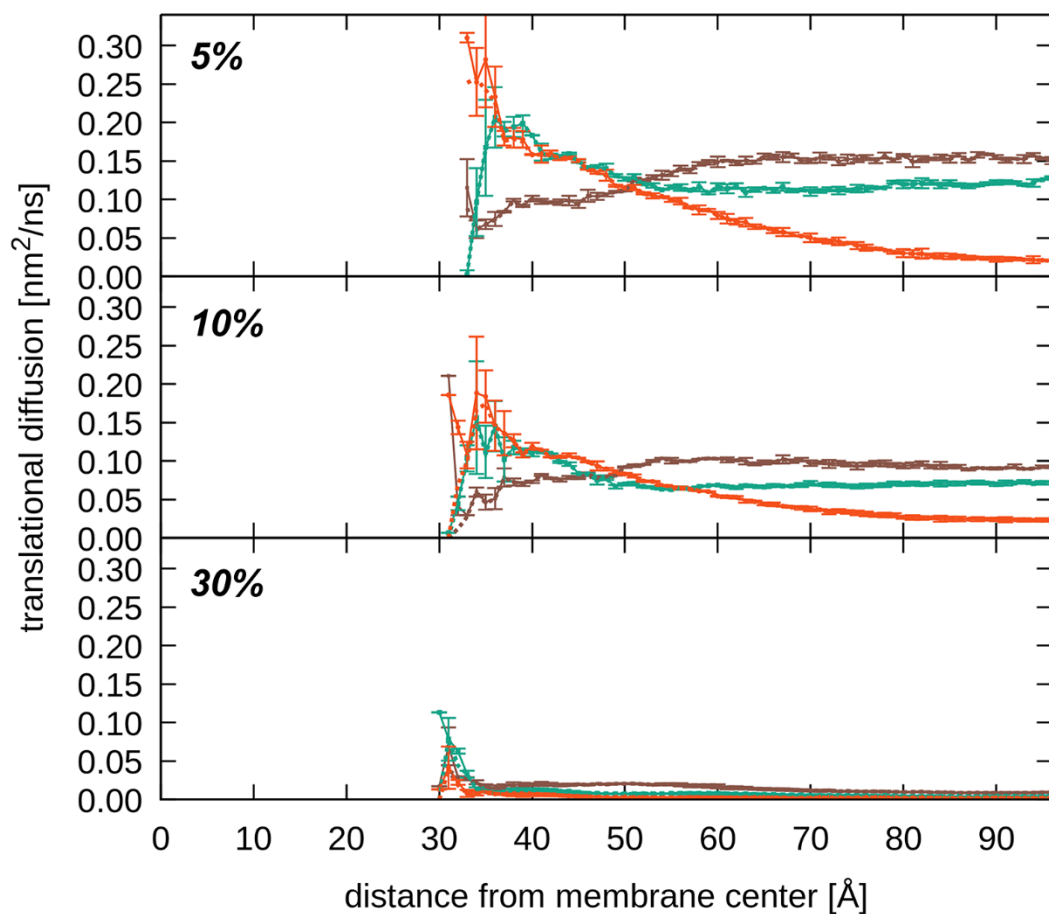


Figure S27: Translational diffusion perpendicular to the membrane (along the z-axis) as a function of the distance from the membrane center extracted from MSD curves at different time scales: 0-1 ns (dark brown), 1-10 ns (green), 10-100 ns (red) for protein concentrations of 5%, 10%, and 30%. MSD curves were generated by combining data from all proteins and the results are not corrected for PBC artefacts. Dashed lines indicate results from a modified analysis where potential drift along z due to a potential gradient was subtracted according to $\langle (dz - \langle dz \rangle)^2 \rangle$.

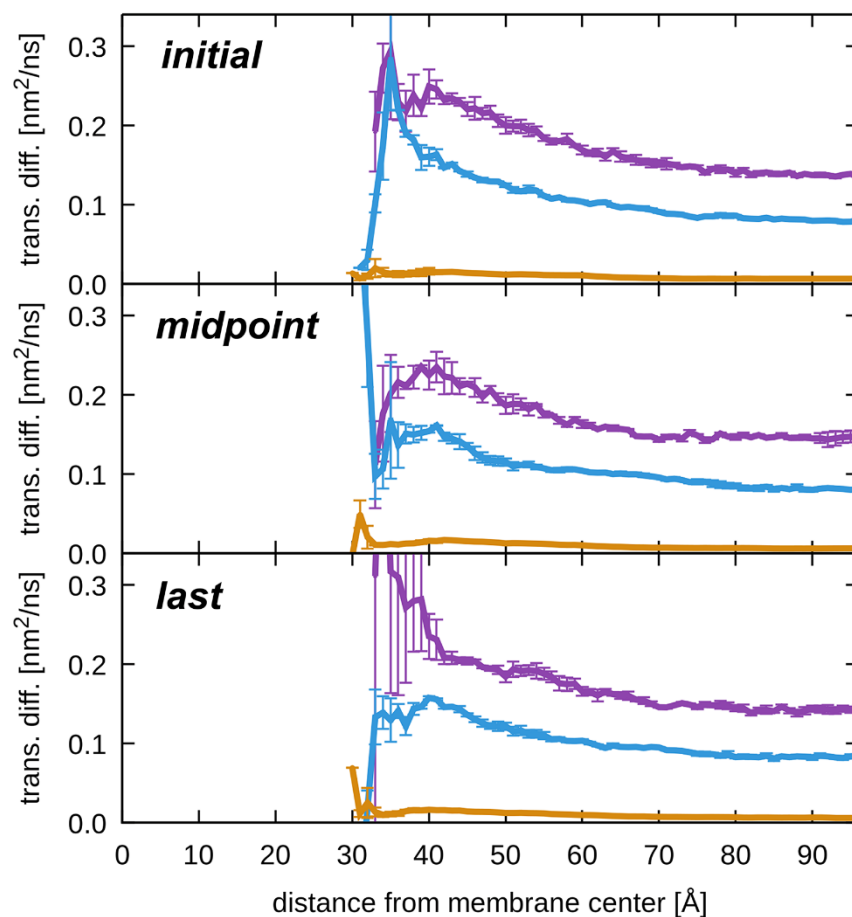


Figure S28: Translational diffusion constants parallel to the membrane (in the x-y plane) as a function of the distance from the membrane center extracted from MSD curves at 10-100 ns for all proteins at total protein concentrations of 5% (purple), 10% (light blue), and 30% (tan). Top, center, and bottom panels compare the choice of the initial, mid, or last point of the time interval over which diffusion is measured in assigning the distance from the membrane center.

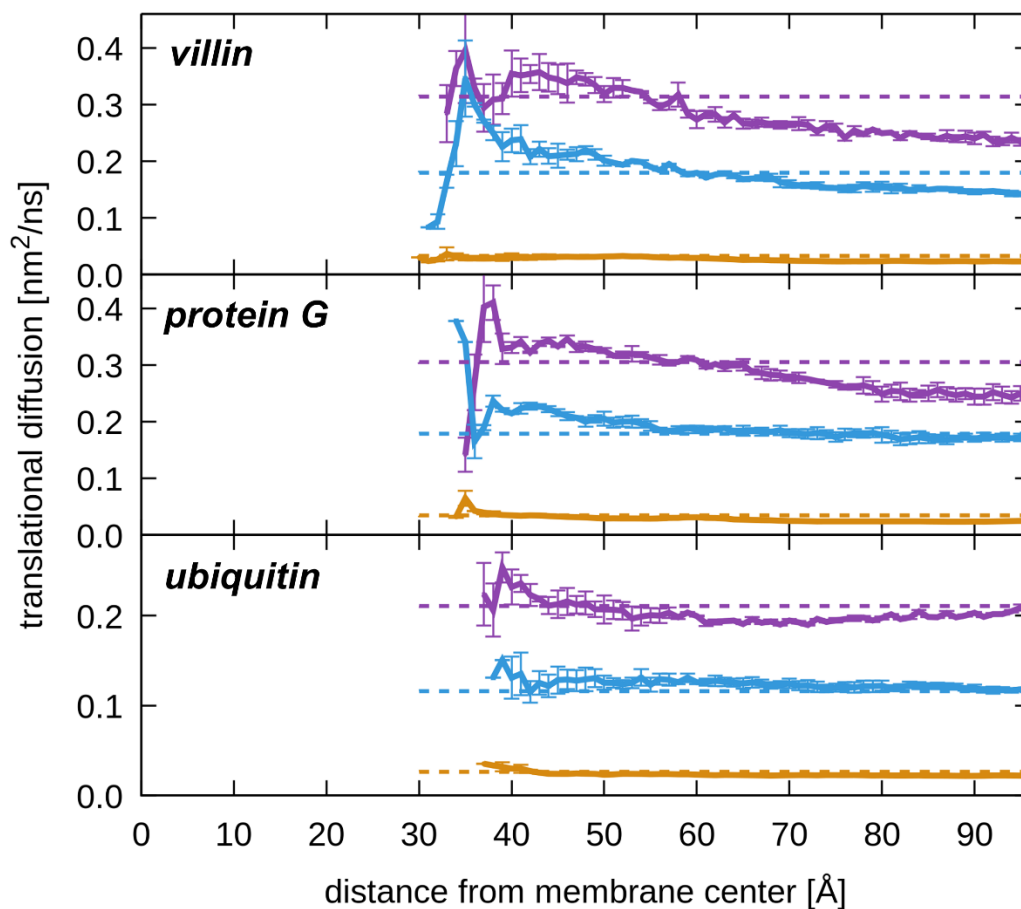


Figure S29: Translational diffusion constants parallel to the membrane (in the x-y plane) as a function of the distance from the membrane center extracted from MSD curves at 10-100 ns for only villin, protein G, or ubiquitin molecules at total protein concentrations of 5% (purple), 10% (light blue), and 30% (tan). Diffusion values were corrected for PBC artefacts according to Eq. 3 with viscosity estimated by Eq. 2. The dashed lines indicate the predicted diffusion rates parallel to the membrane according to Eq. 4 based on PBC-corrected diffusion for different proteins in non-membrane systems at the same concentration (see data in Table S7).

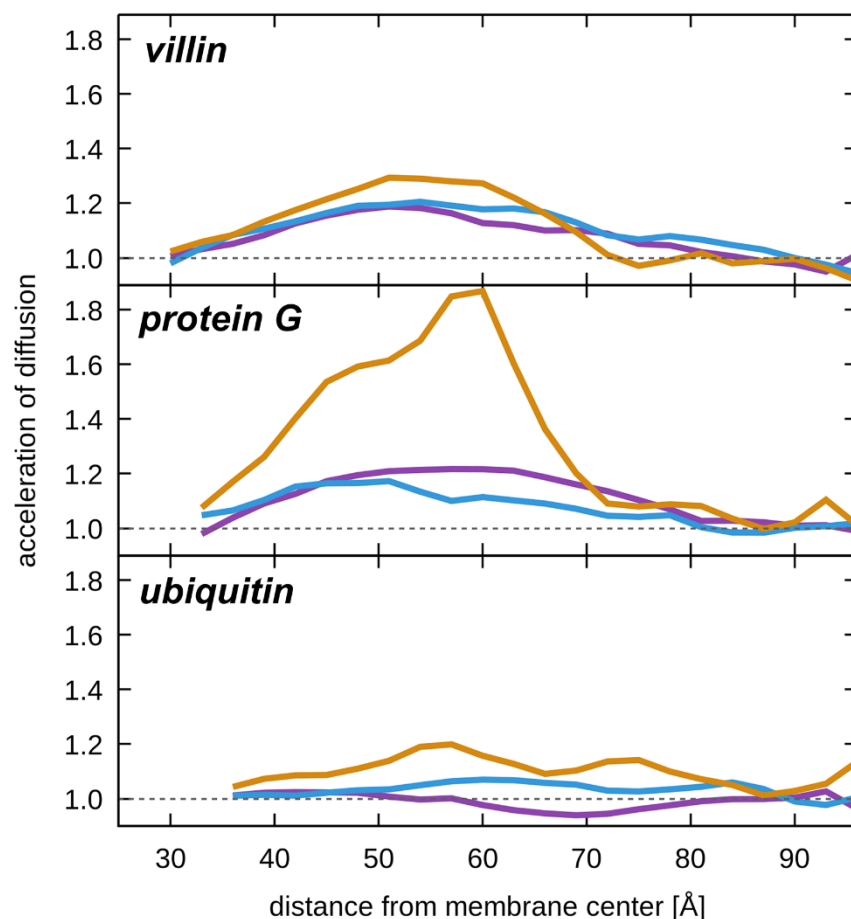


Figure S30: Population-weighted acceleration of translational diffusion parallel to the membrane (in the x-y plane) as a function of the distance from the membrane center extracted from MSD curves at 10-100 ns for only villin, protein G, or ubiquitin molecules at total protein concentrations of 5% (purple), 10% (light blue), and 30% (tan). Acceleration was calculated as the ratio of relaxation times $\tau_{\text{bulk}}/\tau_z$ where $\tau_{\text{bulk}}=1/D_{\text{bulk}}$ and $\tau_z=(1-p)/D_z + p/D_{\text{bulk}}$ and $p=\rho_z/\rho_{\text{bulk}}$ is the probability of a protein to be found at a given distance from the membrane center based on the z-dependent density profiles shown in Figure S18 and the z-dependent diffusion profiles shown in Figure S29.

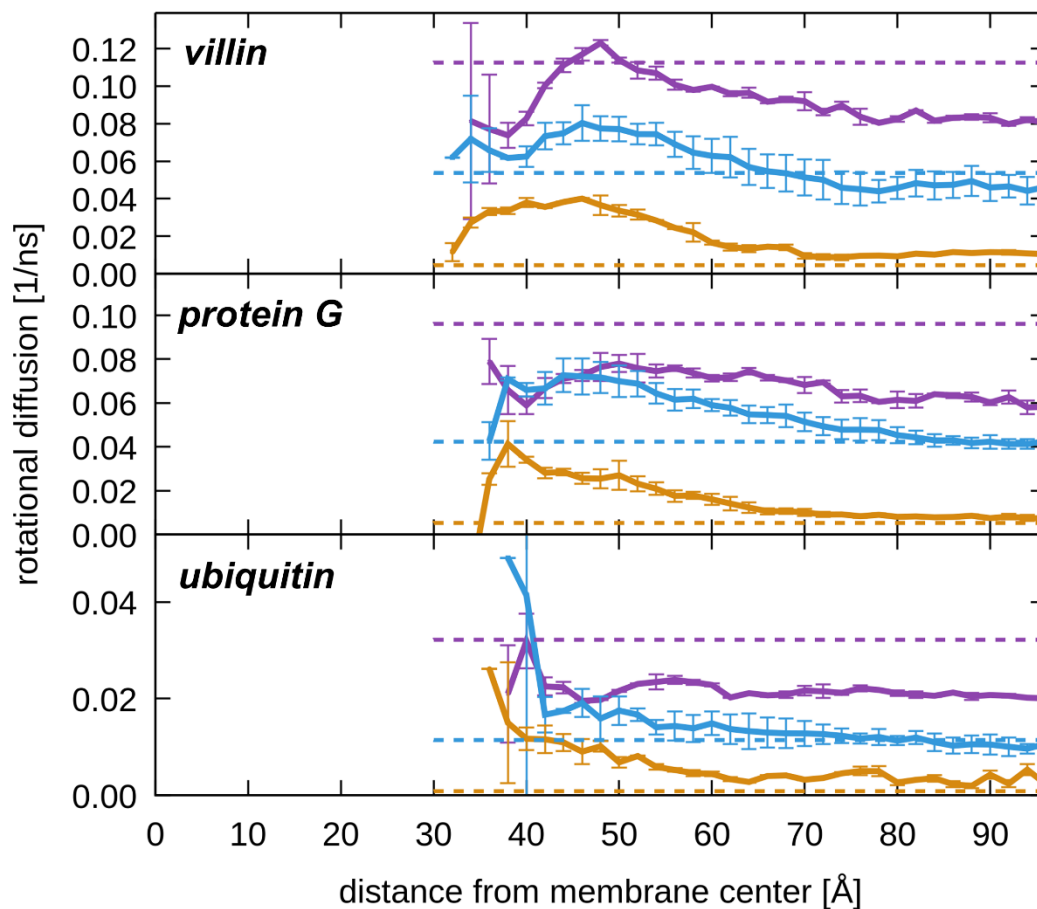


Figure S31: Rotational diffusion constants as a function of the distance from the membrane center for only villin, protein G, or ubiquitin molecules at total protein concentrations of 5% (purple), 10% (light blue), and 30% (tan). The dashed lines indicate rotational diffusion for different proteins in non-membrane systems at the same concentration (see data in Table S8).

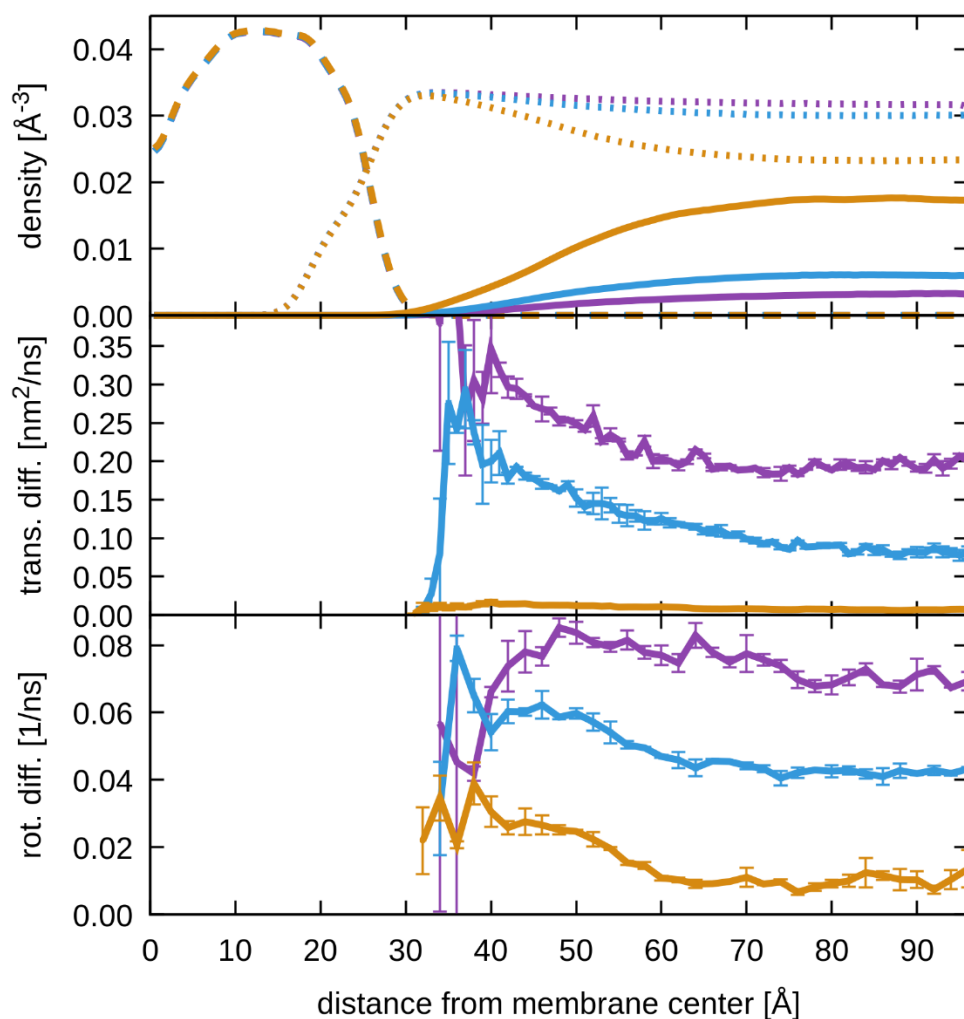


Figure S32: Heavy-atom density distributions of molecular components (top), translational diffusion constants parallel to the membrane (center), and rotational diffusion constants (bottom) as a function of the distance from the membrane center along the membrane normal as in Figure 3 but based on only the first 2 μ s before some proteins unfold. Results are shown for systems with proteins at 5% (purple), 10% (light blue), and 30% (tan). Densities are shown for proteins (solid lines), lipids (long dashes), and water molecules (short dashes). Translational and rotational diffusion constants were assigned to the center of mass of a given protein at the beginning of the intervals for which mean-square displacements (MSD) and rotational correlation functions were obtained. Translational diffusion was estimated from MSD vs. time during 10-100 ns. Statistical errors for density distributions are less than 1%.

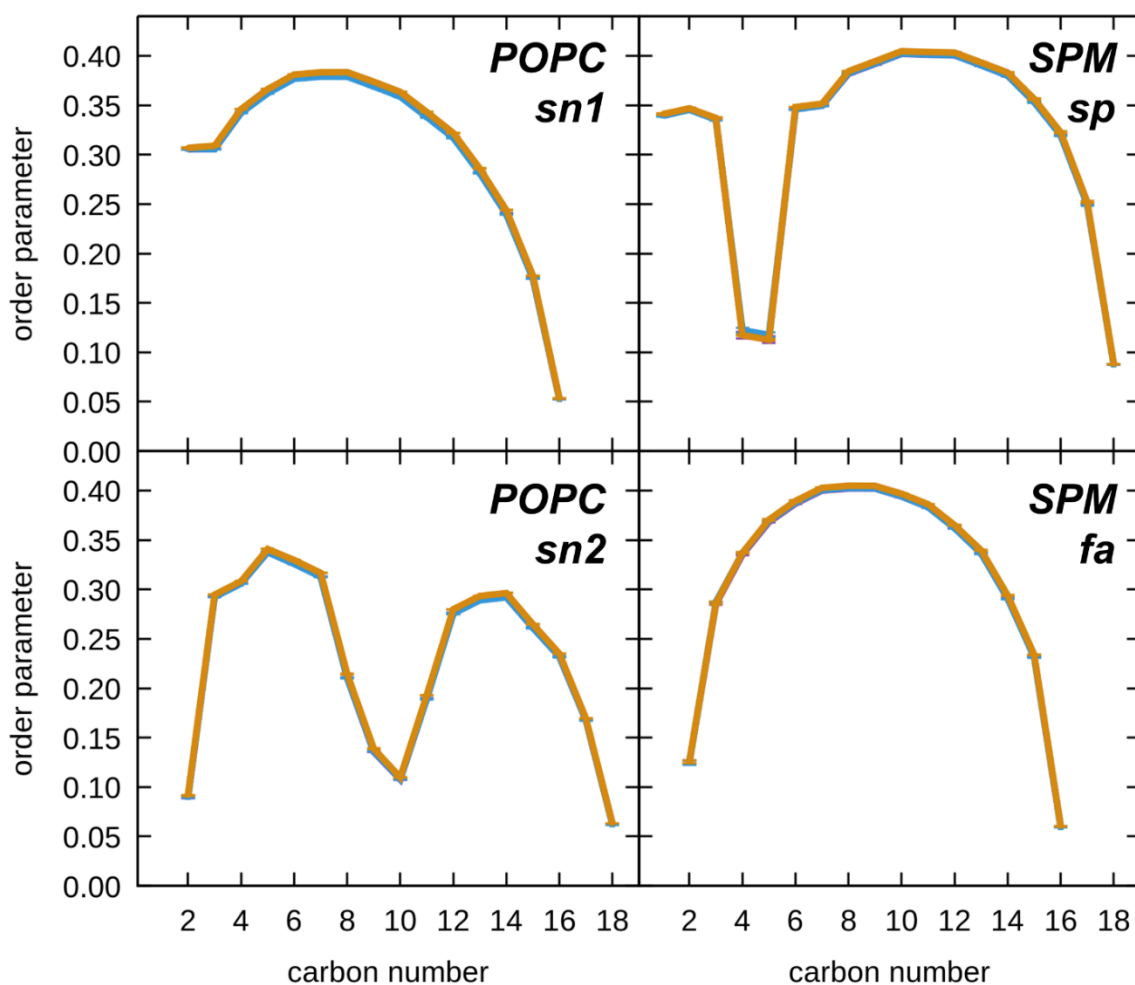


Figure S33: Lipid acyl chain order parameters for POPC (sn1 and sn2 chains) and sphingomyelin (sphingosine (sp) and fatty acid (fa) chains) with protein concentrations of 5% (purple), 10% (light blue), and 30% (tan). Order parameters were calculated as $S = |\langle (3\cos^2\theta - 1)/2 \rangle|$, where θ is the time-dependent angle of a given C-H bond vectors along the acyl chains relative to the membrane normal.

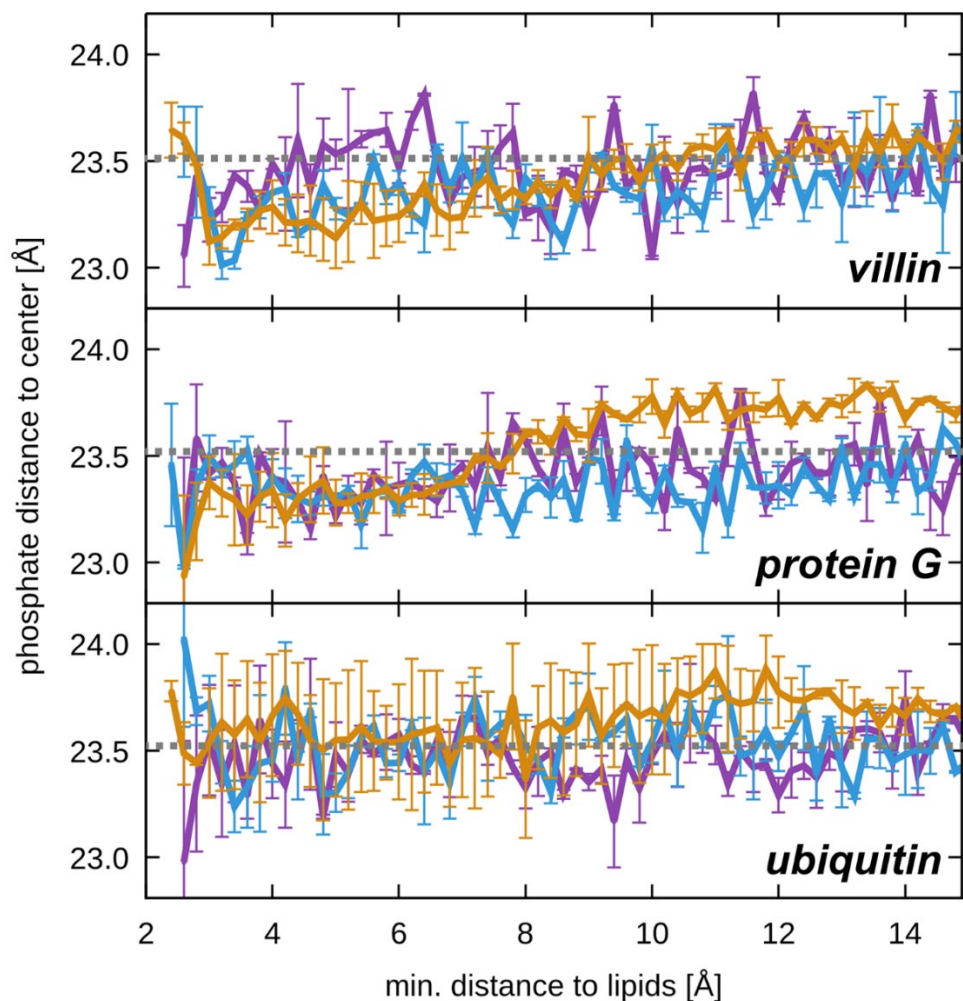


Figure S34: Membrane distortion as a function of protein interactions at protein concentrations of 5% (purple), 10% (light blue), and 30% (tan) as in Figure 4 but based on only the first 2 us of each trajectory before some proteins unfold. Protein-membrane distances are defined based on minimum heavy-atom distances between proteins and lipids. Membrane distortions are characterized by average phosphate distances from the membrane center for phosphate atoms within a 15 Å radius from the lipid atom in closest contact with the protein. The average phosphate distance to the center irrespective of any protein contact is indicated as a grey line.

Table S1. Simulated all-atom systems.

System	%vol	#proteins ¹	#lipids ²	#K ⁺ /Cl ⁻	#atoms	Box x-y [<i>Å</i>]	Box z [<i>Å</i>]	NAMD time [<i>μs</i>]	Anton2 time [<i>μs</i>]	
a5	5	15 (5,5,5)		0	306/296	364721	153.32	153.32	0	10
a10	10	30 (10,10,10)		0	326/306	371813	154.06	154.06	0.5	10
a30	30	30 (10,10,10)		0	103/83	124602	106.31	106.31	0.8	10
a5m	5	12 (4,4,4)	828 (276,276,276)	232/224	363360	134.21	193.71	0		10
a10m	10	24 (8,8,8)	828 (276,276,276)	248/232	367575	134.18	195.34	0.3		10
a30m	30	69 (23,23,23)	828 (276,276,276)	227/181	369657	133.74	194.83	0.7		10

¹number of villin, protein G, and ubiquitin molecules given in parentheses; ²number of POPC, sphingomyelin, and cholesterol molecules given in parentheses.

Table S2. Protein stability vs. concentration.

	%vol	Villin			Protein G			Ubiquitin		
		avg. ¹	err. ²	p ³	avg. ¹	err. ²	p ³	avg. ¹	err. ²	p ³
RMSD [Å]	5	3.36	0.81		1.14	0.12		1.56	0.13	
	10	4.41	0.87		2.17	0.65		1.48	0.11	
	30	2.79	0.41	0.59	1.29	0.06	0.39	1.44	0.08	0.53
RMSD (<2.5 Å) ⁴ [Å]	5	1.31	0.08		1.04	0.07		1.53	0.13	
	10	1.32	0.11		1.20	0.07		1.45	0.11	
	30	1.42	0.05	0.35	1.18	0.05	0.24	1.42	0.08	0.54
R_g [Å]	5	11.26	0.596		10.80	0.017		11.96	0.048	
	10	11.82	0.624		11.25	0.396		11.93	0.032	
	30	10.53	0.231	0.37	10.80	0.009	0.90	11.90	0.022	0.37
R_g (<2.5 Å) ⁴ [Å]	5	9.84	0.059		10.79	0.006		11.95	0.047	
	10	9.21	0.616		10.78	0.005		11.92	0.031	
	30	9.85	0.037	0.88	10.79	0.007	0.73	11.90	0.020	0.37

¹averages over Anton2 simulations with and without membrane; ²standard errors obtained via error propagation formula from trajectory errors; ³p-values for null hypothesis of values at 5% and 30% not being significantly different; ⁴considering only folded snapshots with an RMSD of less than 2.5 Å.

Table S3. Protein stability vs. membrane presence.

	mem.	Villin			Protein G			Ubiquitin		
		avg. ¹	err. ²	p ³	avg. ¹	err. ²	p ³	avg. ¹	err. ²	p ³
RMSD	no	3.45	<i>0.591</i>		1.82	<i>0.431</i>		1.59	<i>0.094</i>	
[Å]	yes	3.59	<i>0.599</i>	0.88	1.25	<i>0.093</i>	0.32	1.39	<i>0.090</i>	0.27
RMSD (<2.5 Å)⁴	no	1.34	<i>0.033</i>		1.15	<i>0.041</i>		1.56	<i>0.088</i>	
[Å]	yes	1.36	<i>0.088</i>	0.81	1.13	<i>0.058</i>	0.73	1.37	<i>0.087</i>	0.28
R_g	no	11.13	<i>0.400</i>		11.11	<i>0.264</i>		11.95	<i>0.026</i>	
[Å]	yes	11.28	<i>0.440</i>	0.83	10.79	<i>0.016</i>	0.35	11.91	<i>0.032</i>	0.39
R_g (<2.5 Å)⁴	no	9.81	<i>0.026</i>		10.79	<i>0.005</i>		11.94	<i>0.025</i>	
[Å]	yes	9.46	<i>0.412</i>	0.48	10.78	<i>0.005</i>	0.56	11.90	<i>0.031</i>	0.41

¹averages over Anton2 simulations at all three concentrations; ²standard errors obtained via error propagation formula from trajectory errors; ³p-values for null hypothesis of values without and with membrane not being significantly different; ⁴considering only folded snapshots with an RMSD of less than 2.5 Å.

Table S4. Protein contacts by protein.

System	%vol	Villin			Protein G			Ubiquitin			All	
		$f_{\text{cont.}}^1$	err.^2	f/A^3	$f_{\text{cont.}}^1$	err.^2	f/A^3	$f_{\text{cont.}}^1$	err.^2	f/A^3	$f_{\text{cont.}}^1$	err.^2
a5	5	0.08	0.04	0.77	0.02	0.002	0.15	0.36	0.16	2.10	0.55	0.07
a10	10	0.24	0.05	2.32	0.13	0.05	0.97	0.48	0.08	2.80	0.86	0.06
a30	30	0.39	0.09	3.78	0.18	0.03	1.34	0.76	0.06	4.44	1.24	0.08
a5m	5	0.27	0.03	2.61	0.04	0.01	0.30	0.39	0.05	2.28	0.86	0.06
a10m	10	0.25	0.04	2.42	0.07	0.01	0.52	0.74	0.10	4.32	0.92	0.07
a30m	30	0.30	0.02	2.91	0.24	0.03	1.79	0.75	0.08	4.38	1.48	0.06

¹fraction of contacts calculated as the number of contacts between the same type of proteins normalized by the number of proteins; contacts were defined as minimum C α distances of less than 7 Å; ²standard errors based on variation between different sets of molecules; ³fraction of contacts divided by the surface areas of a sphere with an equivalent volume to the respective protein structures ($A_{\text{villin}}=10.33 \text{ nm}^2$, $A_{\text{proteinG}}=13.42 \text{ nm}^2$, $A_{\text{ubiquitin}}=17.11 \text{ nm}^2$) multiplied by 100.

Table S5. Amino acid preferences in protein-membrane contacts¹.

Amino acid	5%	err.	10%	err.²	30%	err.²
Alkanes ³	19.1	11.7	11.8	1.0	18.2	2.0
Polar ⁴	16.6	1.1	15.1	2.9	22.6	7.2
Acidic ⁵	6.0	0.1	6.2	0.4	5.7	1.6
Basic ⁶	4.0	0.01	4.9	1.5	5.9	2.3
Sulfur ⁷	4.6	1.2	3.9	1.4	2.3	0.2
Phe	3.1	0.7	2.7	0.6	2.6	0.1
Tyr	16.6	2.3	12.4	2.3	18.2	10.2
Trp	6.5	0.6	9.6	1.4	10.8	4.3
Gly	1.7	0.1	1.7	0.3	2.8	0.6
Pro	1.9	0.1	2.6	1.1	3.1	0.6

¹Percentages of closest protein-lipid heavy atom contacts within 5 Å as a function of amino acid type, e.g. in 6.5% of all instances where Trp is the closest amino acid in a protein to the membrane at 5% concentration; ²Statistical uncertainties from comparing subsets of proteins;

³Ala, Val, Leu, Ile; ⁴Asn, Gln, His, Ser, Thr; ⁵Asp, Glu; ⁶Arg, Lys.

Table S6. Protein-membrane contact residence times.

System	%vol	Protein	τ_1 [ns]	err.	τ_2 [ns]	err	a	err.
a5m	5	villin	5.28	1.10	1585	727	0.95	0.02
		protein G	1.51	0.41	1843	650	0.97	0.001
		ubiquitin	3.87	0.95	945	408	0.95	0.003
		all	3.20	0.18	1144	11.5	0.94	0.008
a10m	10	villin	3.93	0.42	1622	1302	0.91	0.02
		protein G	2.22	0.13	1296	118	0.93	0.004
		ubiquitin	1.65	0.04	75	12.3	0.81	0.03
		all	3.09	0.37	829	420	0.92	0.01
a30m	30	villin	5.63	1.05	684	15.4	0.78	0.03
		protein G	2.28	0.41	622	278	0.83	0.03
		ubiquitin	1.82	0.47	489	31.1	0.79	0.02
		all	2.97	0.29	589	78.6	0.80	0.02

Residence times from fitting contact survival functions to a double exponential expression with parameters a , τ_1 , and τ_2 : $a \cdot \exp(-t/\tau_1) + (1-a) \cdot \exp(-t/\tau_2)$. Protein-membrane contacts were defined based on protein-lipid heavy-atom distances within 5 Å.

Table S7. Translational diffusion rates for proteins in non-membrane systems.

System	%vol	Protein	D_t	err.	D_t	err.	D_t	err.
			<i>0-1 ns</i> [nm ² /ns]		<i>1-10 ns</i> [nm ² /ns]		<i>10-100 ns</i> [nm ² /ns]	
a5	5	villin	0.36	0.03	0.35	0.03	0.34	0.04
		protein G	0.36	0.01	0.34	0.01	0.33	0.02
		ubiquitin	0.27	0.007	0.25	0.005	0.23	0.006
a10	10	Villin	0.23	0.01	0.21	0.01	0.19	0.01
		protein G	0.22	0.01	0.21	0.01	0.20	0.01
		ubiquitin	0.16	0.005	0.15	0.005	0.13	0.004
a30	30	villin	0.051	0.002	0.041	0.002	0.035	0.001
		protein G	0.055	0.002	0.043	0.002	0.038	0.002
		ubiquitin	0.037	0.001	0.031	0.001	0.029	0.0004

Diffusion from mean-square center of mass displacement curves with correction for periodic boundary artefacts according to Eq. 1 with viscosity estimated by Eq. 2 using $\eta_w=0.35$ cP (for CHARMM TIP3P with Ewald summation) (35).

Table S8. Rotational diffusion rates for proteins in non-membrane systems.

System	%vol	Protein	D_r [1/ns]	err.	τ_f [ns]	err	τ_s [ns]	err.	S_R^2	err.
a5	5	villin	0.112	0.027	1.36	0.42	8.79	2.01	0.38	0.05
		protein G	0.096	0.007	1.21	0.07	9.68	2.51	0.37	0.04
		ubiquitin	0.032	0.003	3.66	0.60	17.59	1.49	0.42	0.04
a10	10	villin	0.054	0.008	2.22	0.22	30.50	3.13	0.40	0.04
		protein G	0.042	0.008	3.31	0.50	43.61	7.51	0.39	0.03
		ubiquitin	0.011	0.001	9.07	1.28	58.25	4.33	0.54	0.03
a30	30	villin	0.005	0.001	21.49	3.34	547.5	64.8	0.55	0.03
		protein G	0.005	0.001	21.62	2.81	708.1	124.7	0.54	0.03
		ubiquitin	0.0008	0.0002	72.47	10.8	1502.9	276.8	0.77	0.02

Diffusion rates were extracted from fitting rotational correlation functions to double exponentials: $S_R^2 \exp(-t/\tau_s) + (1 - S_R^2) \exp(-t/\tau_f)$.

References

1. S. Jo, T. Kim, V. G. Iyer, W. Im, CHARMM-GUI: A Web-Based Graphical User Interface for CHARMM. *J Comput Chem* **29**, 1859-1865 (2008).
2. S. Jo, J. B. Lim, J. B. Klauda, W. Im, CHARMM-GUI Membrane Builder for mixed bilayers and its application to yeast membranes. *Biophys J* **97**, 50-58 (2009).
3. E. L. Wu *et al.*, CHARMM-GUI Membrane Builder toward realistic biological membrane simulations. *J Comput Chem* **35**, 1997-2004 (2014).
4. C. J. McKnight, P. T. Matsudaira, P. S. Kim, NMR Structure of the 35-Residue Villin Headpiece Subdomain. *Nat Struct Biol* **4**, 180-184 (1997).
5. A. M. Gronenborn *et al.*, A Novel, Highly Stable Fold of the Immunoglobulin Binding Domain of Streptococcal Protein G. *Science* **253**, 657-661 (1991).
6. S. Vijaykumar, C. E. Bugg, W. J. Cook, Structure of Ubiquitin Refined at 1.8 Å Resolution. *J Mol Biol* **194**, 531-544 (1987).
7. R. B. Best *et al.*, Optimization of the Additive CHARMM All-Atom Protein Force Field Targeting Improved Sampling of the Backbone ϕ , ψ and Side-Chain χ_1 and χ_2 Dihedral Angles. *J Chem Theory Comput* **8**, 3257-3273 (2012).
8. G. Nawrocki, P.-h. Wang, I. Yu, Y. Sugita, M. Feig, Slow-Down in Crowded Protein Solutions Correlates with Transient Oligomer Formation. *J Phys Chem B* **121**, 11072-11084 (2017).
9. J. B. Klauda *et al.*, Update of the CHARMM All-Atom Additive Force Field for Lipids: Validation on Six Lipid Types. *J Phys Chem B* **114**, 7830-7843 (2010).
10. A. D. MacKerell, Jr. *et al.*, All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *J Phys Chem B* **102**, 3586-3616 (1998).
11. A. Sandoval-Perez, K. Pluhackova, R. A. Böckmann, Critical Comparison of Biomembrane Force Fields: Protein–Lipid Interactions at the Membrane Interface. *J Chem Theory Comput* **13**, 2310-2321 (2017).
12. D. Beglov, B. Roux, Finite representation of an infinite bulk system: solvent boundary potential for computer simulations. *J Chem Phys* **100**, 9050-9063 (1994).
13. Y. Luo, B. Roux, Simulation of Osmotic Pressure in Concentrated Aqueous Salt Solutions. *J Phys Chem Lett* **1**, 183-189 (2010).
14. H. J. C. Berendsen, J. P. M. Postma, W. F. v. Gunsteren, A. DiNola, J. R. Haak, Molecular dynamics with coupling to an external bath. *J Chem Phys* **81**, 3684-3690 (1984).
15. R. A. Lippert *et al.*, Accurate and efficient integration for molecular dynamics simulations at constant temperature and pressure. *J Chem Phys* **139**, 164106 (2013).
16. V. Kräutler, W. F. Van Gunsteren, P. H. Hünenberger, A fast SHAKE algorithm to solve distance constraint equations for small molecules in molecular dynamics simulations. *J Comput Chem* **22**, 501-508 (2001).
17. G. J. Martyna, D. J. Tobias, M. L. Klein, Constant pressure molecular dynamics algorithms. *J Chem Phys* **101**, 4177-4189 (1994).
18. S. Nose, A molecular dynamics method for simulations in the canonical ensemble. *Mol Phys* **52**, 255-268 (1984).

19. Y. Shan, J. L. Klepeis, M. P. Eastwood, R. O. Dror, D. E. Shaw, Gaussian split Ewald: A fast Ewald mesh method for molecular simulation. *J Chem Phys* **122**, 054101 (2005).
20. G. Nawrocki, A. Karaboga, Y. Sugita, M. Feig, Effect of Protein–Protein Interactions and Solvent Viscosity on the Rotational Diffusion of Proteins in Crowded Environments. *Phys Chem Chem Phys* **21**, 876-883 (2019).
21. I. C. Yeh, G. Hummer, System-Size Dependence of Diffusion Coefficients and Viscosities from Molecular Dynamics Simulations with Periodic Boundary Conditions. *J Phys Chem B* **108**, 15873-15879 (2004).
22. A. Ortega, D. Amorós, J. García de la Torre, Prediction of Hydrodynamic and Other Solution Properties of Rigid Proteins from Atomic- and Residue-Level Models. *Biophys J* **101**, 892-898 (2011).
23. C. Tanford, *Physical Chemistry of Macromolecules* (John Wiley & Sons, New York, 1961).
24. S. von Bülow, M. Siggel, M. Linke, G. Hummer, Dynamic cluster formation determines viscosity and diffusion in dense protein solutions. *Proc Natl Acad Sci USA* 10.1073/pnas.1817564116, 201817564 (2019).
25. P. Simonnin, B. Noetinger, C. Nieto-Draghi, V. Marry, B. Rotenberg, Diffusion under confinement: hydrodynamic finite-size effects in simulation. *J Chem Theory Comput* **13**, 2881-2889 (2017).
26. V. Wong, D. A. Case, Evaluating Rotational Diffusion from Protein MD Simulations. *J Phys Chem B* **112**, 6013-6024 (2008).
27. M. Linke, J. Köfinger, G. Hummer, Rotational Diffusion Depends on Box Size in Molecular Dynamics Simulations. *J Phys Chem Lett* **9**, 2874-2878 (2018).
28. M. Feig, J. Karanicolas, C. L. Brooks, MMTSB Tool Set: Enhanced Sampling and Multiscale Modeling Methods for Applications in Structural Biology. *J Mol Graph Modell* **22**, 377-395 (2004).
29. B. R. Brooks *et al.*, CHARMM: The Biomolecular Simulation Program. *J Comput Chem* **30**, 1545-1614 (2009).
30. W. Humphrey, A. Dalke, K. Schulten, VMD: Visual Molecular Dynamics. *J Mol Graph* **14**, 33-38 (1996).
31. T. K. Chiu *et al.*, High-resolution x-ray crystal structures of the villin headpiece subdomain, an ultrafast folding protein. *Proc Natl Acad Sci USA* **102**, 7517-7522 (2005).
32. E. Mani, W. Lechner, W. K. Kegel, P. G. Bolhuis, Equilibrium and Non-Equilibrium Cluster Phases in Colloids with Competing Interactions. *Soft Matter* **10**, 4479-4486 (2014).
33. D. Deb *et al.*, Hard sphere fluids at a soft repulsive wall: A comparative study using Monte Carlo and density functional methods. *J Chem Phys* **134**, 214706 (2011).
34. R. Sibug-Aga, B. B. Laird, Structure of a soft-sphere fluid at a soft repulsive wall: A comparison of weighted density-functional theories. *Phys Rev E* **69**, 051502 (2004).
35. S. E. Feller, R. W. Pastor, A. Rojnuckarin, S. Bogusz, B. R. Brooks, Effect of Electrostatic Force Truncation on Interfacial and Transport Properties of Water. *J Phys Chem* **100**, 17011-17020 (1996).

