PNAS RESEARCH ARTICLE | BIOPHYSICS AND COMPUTATIONAL BIOLOGY

Formation of extramembrane β-strands controls dimerization of transmembrane helices in amyloid precursor protein C99

George A. Pantelopulos[a,b], Daisuke Matsuoka[c], James M. Hutchison[d], Charles R. Sanders[d,e,f], Yuji Sugita[c], John E. Straub[a,1], and D. Thirumalai[g]

Edited by Ken Dill, Stony Brook University, Stony Brook, NY; received July 20, 2022; accepted November 3, 2022

The 99-residue C-terminal domain of amyloid precursor protein (APP-C99), precursor to amyloid beta (Aβ), is a transmembrane (TM) protein containing intrinsically disordered N- and C-terminal extramembrane domains. Using molecular dynamics (MD) simulations, we show that the structural ensemble of the C99 monomer is best described in terms of thousands of states. The C99 monomer has a propensity to form β-strand in the C-terminal extramembrane domain, which explains the slow spin relaxation times observed in paramagnetic probe NMR experiments. Surprisingly, homodimerization of C99 not only narrows the conformational ensemble from thousands to a few states through the formation of metastable β-strands in extramembrane domains but also stabilizes extramembrane α-helices. The extramembrane domain structure is observed to dramatically impact the homodimerization motif, resulting in the modification of TM domain conformations. Our study provides an atomic-level structural basis for communication between the extramembrane domains of the C99 protein and TM homodimer formation. This finding could serve as a general model for understanding the influence of disordered extramembrane domains on TM protein structure.

amyloid | lipid bilayer | intrinsically disordered | protein–protein interaction | molecular dynamics

Significance

Structural characterization of intrinsically disordered protein domains in transmembrane (TM) proteins is challenging due to the inherent conformational diversity. Using all-atom molecular dynamics simulations, we provide a detailed structural characterization of the conformational ensemble of full-length APP-C99 protein in monomeric and homodimeric states. The C-terminal extramembrane domain of C99 monomer forms transient β-strands that stabilize metastable α-helices, conjectured to play a role in Aβ oligomerization and cytosolic signaling. Most importantly, formation of transient β-strands influences TM homodimer formation, constituting long-range communication between the extramembrane and transmembrane domains. These observations emphasize the necessity to model extramembrane domains in TM protein complexes used to regulate cell signaling.

The amyloid β (Aβ) protein is believed to play a central role in the etiology of Alzheimer's disease (AD) (1, 2). Aβ is the product of the sequential cleavage of the amyloid precursor protein (APP), a single-pass transmembrane (TM) protein, by secretases. Initial cleavage by β-secretase produces the 99-amino acid C-terminal fragment, known as APP-C99 or simply C99, which is subsequently cleaved within its TM segment by γ-secretase to produce Aβ (3, 4). This last step in the biogenesis of Aβ protein lacks fidelity, resulting in a distribution of isoforms of Aβ, including the predominant $Aβ_{40}$ and the more amyloidogenic $Aβ_{42}$ (5). It is known that familial AD mutations as well as membrane lipid composition can influence the $Aβ_{40}$:$Aβ_{42}$ ratio and the development of AD (6–11).

Given the importance of this final step in the biogenesis of Aβ, there has been an intense effort to characterize the structure of the enzyme γ-secretase and its C99 substrate using experiment (12, 13) and simulation (14–16). Both APP (17) and C99 (18) are known to exist as both monomers and as homodimers. Experimental studies of full-length monomeric C99 have provided a detailed characterization of the TM domain (TMD) as well as insights into the nature of the flanking juxta-membrane (JM) domains.

Congeners of C99, composed of residues 16 to 55 ($C99_{16–55}$) or shorter sequences, are known to form homodimers sometimes defined by the TM domain glycine zipper $G_{29}xxxG_{33}xxxG_{37}$ (19). This sequence motif is known to play a role in stabilizing homodimers of TM helices such as glycophorin A (20). The propensity for different dimer motifs, "Gly-in," "Gly-side," and "Gly-out," defined by the Crick angle between C99 homodimers along both Gly33 alpha carbons, have been shown to be dependent on whether lipids are organized into micelles or bilayers, and are dependent on lipid composition (21–24). The JM domain, $K_{16}LVFFAED_{23}$, has been demonstrated to form an α-helix both in certain familial AD mutants and in lower pH environments characteristic of endosomes (25–29). While studies have been devoted to the impact of membrane lipid composition on the structure of the TM domain of monomers and dimers of C99 congeners, the impact of membrane composition and extramembrane domains on the structure of full-length C99 monomer and dimer remains largely unknown.

Solution phase NMR experiments have been successful in partially characterizing the structure of extramembrane domains through the use of detergent micelles and bicelles to mimic lipid bilayer environments. These experiments faced a variety of limitations which have gradually been overcome with time. Initial experiments were performed using 14:0 lyso PG (LMPG) micelles with varying concentrations of cholesterol (28) and

The authors declare no competing interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence may be addressed. Email: straub@bu.edu.

This article contains supporting information online at http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2212207119/-/DCSupplemental.

Published December 20, 2022.

a cholesterol analogue (27), which introduce a substantial amount of curvature to the hydrophobic surface experienced by C99. Later experiments which used lipid/DHPC (lipid/detergent) ratios (q) of 0.5 which were expected to form an oblate, disk-shaped bicelle to serve as a bilayer-like environment, found little difference in C99 paramagnetic probe response relative to LMPG micelle conditions as a function of lipid identity (29). Caldwell et al. recently characterized 07:0 PC (DHPC) bicelles using small-angle X ray and neutron scattering, fluorescence anisotropy, and MD simulation, finding bicelles of $q \leq 0.5$ to be spheroidal (micellear) in shape, due to mixing of DHPC with other lipids (30). In light of this issue with DHPC bicelles, recent solution phase NMR experiments have been performed using bicelles in which the detergent component is n-dodecyl-β-melibioside (DDMB). This bicelle medium has been characterized by small-angle X ray and neutron scattering and cryoelectron microscopy and were observed to oblate even at q = 0.33. C99 can be successfully solubilized in these bicelles in a variety of lipid compositions (31), allowing for characterization of C99 in both liquid ordered (L_o) and liquid disordered (L_d) lipid phases (32).

Recent studies of C99 at q = 0.33 in 4:2:1 14:0 PC (DMPC):eSM:Chol (14:0 PC lipid, egg sphingomyelin, and cholesterol) L_o phase DDMB bicelles (sphingomyelin and cholesterol-rich bicelles: "SCOR bicelles") at 318.15 K and 4.5 pH and at dilute, 300 to 400 μM concentration have suggested a few key results regarding domains beyond the JM region. 1) Extramembrane residues are generally random coil, but chemical shifts and relative reductions of 15N-amide T_2 spin relaxation times suggest α-helical propensity and surface association in the C-terminal residues of C99 (residues 91-99), as in previous experiments with LMPG and in DHPC-DMPC bicelles. 2) Cytosolic C-terminal domain residues 63 to 73 exhibit no amide proton exchange with water. Moreover, paramagnetic probe experiments employing the lipid-like hydrophobic probe 16-DOXYL-stearic acid (16-DSA) and the hydrophilic probe, gadopentetic acid (Gd-DTPA), have demonstrated that residues 63 to 73 are little affected by either of these probes. Surprisingly, 15N-amide T_2 spin relaxation times for these sites exhibit no reduction relative to other random coil residues in the extramembrane domains, implying that this domain is obfuscated from the bilayer or solution due to involvement in a complicated dynamic ensemble of intraprotein interactions within the C-terminal extramembrane domain. 3) N-terminal residues 2 to 6 exhibit some inhibition of amide proton exchange and reduced accessibility by both paramagnetic probes, also implying that a dynamic ensemble of intraprotein interactions occurs involving the N-terminal extramembrane domain, albeit possibly less complicated than for the 63 to 73 segment.

There is some evidence suggesting that extramembrane domains of C99 in these bicelles form transient secondary structures in the monomeric state. The homodimeric state would be expected to form a more stable secondary structure, though the homodimeric state conformational ensemble remains poorly characterized due to the significant propensity for C99 to also populate monomeric and homotrimeric states (31). The N-terminal extramembrane domain of C99 contains much of the Aβ sequence and can stabilize oligomers and fibrils by forming a variety of β-strands (2, 33). It is also known to modulate production of Aβ length via interaction with γ-secretase (14). The C-terminal extramembrane cytosolic domain includes phosphorylation sites and engages in many cytoplasmic interactions, many likely to involve induced secondary structure (34, 35). Our previous simulations of the full-length C99 monomer in implicit solvent bilayers observed ensembles of metastable states characterized by unique α-helical and β-strand structures dependent on membrane thickness (34). β-strands observed in the extramembrane domains during these simulations suggested that the C99 N terminus might be available as a seed for Aβ aggregation on the membrane surface, believed to be a potential mechanism for Aβ toxicity (36), and that the C99 N and C termini might serve as interfaces for stabilizing the C99 homodimer and other protein–protein interactions. However, the implicit solvent model used in our past work is expected to produce a more collapsed state and higher α-helix propensity (37), so more accurate explicit solvent simulations employing force fields more appropriate to modeling intrinsically disordered proteins are necessary to rigorously characterize the conformational ensemble (38).

Here, we report generalized Replica Exchange with Solute Tempering (gREST) molecular dynamics (MD) simulations (39) of the C99 monomer and homodimer in 16:0-18:1 PC (POPC) lipid bilayers. gREST is a Hamiltonian replica exchange method that substantially enhances the sampling of protein structures by scaling down the potential energies of bonded and nonbonded interactions within a predefined "solute" set of atoms and the nonbonded interactions of the solute with the rest of the system (the "solvent") as if increasing the temperature of only the solute subset of the system. A novelty of the formulation and implementation of gREST is the option to selectively scale only some types of bonded and nonbonded interactions. In this work, we define the gREST solute as the C99 protein and POPC, scaling bonds, angles, dihedrals, CMAP terms, and Lennard–Jones interactions down with effective temperature but leaving electrostatic interactions unperturbed to preserve membrane–water and membrane-protein interactions at all effective temperatures. We use POPC lipid bilayers as a simple model of a L_d phase bilayer to which wild-type C99 is evidenced to strongly associate with $P(L_d) \approx 0.86 - 0.9$ (40), and is often used as a reductive model of the bulk plasma membrane,

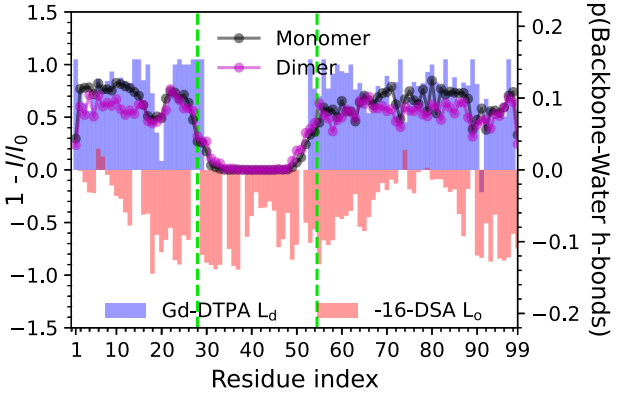

Fig. 1. Left y-axis and bar plots: I/I_0 is the ratio of full-length C99 1H, 15N backbone amide NMR signals in the presence of either hydrophilic (Gd-DTPA) or hydrophobic (16-DSA) paramagnetic probes over peak intensities observed in probe-free conditions. Right y-axis and scatter plots: The observed likelihood of C99 monomer and homodimer backbone amide or carbonyl hydrogen bond formation with water in gREST simulations at an effective temperature of T = 310 K. Green dashed lines indicate bilayer surface position. The L_d phase Gd-DTPA NMR experiment was performed with 100 μM uniformly 15N-enriched monomeric C99 in 10 wt% in solution q = 0.33 POPC/DDMB bicelles with 1 mM Gd-DTPA. L_o phase Gd-DTPA and 16-DSA experiments performed with 200 to 400 μM uniformly enriched partially multimeric C99 in 5 to 10 wt% in q = 0.33 4:2:1 DMPC:eSM:Chol/DDMB bicelles and in the presence of either 6 to 9 mol% 16-DSA (relative to the total moles of lipid) or 1 mM Gd-DTPA in solution (31).

where 90% of APP cleavage occurs (41). We also report NMR paramagnetic probe experiments of C99 in POPC lipid DDMB bicelles following the same protocol as in recent work (31).

Cluster analysis is used to characterize the remarkable diversity of structures in the monomer ensemble and the relative simplicity of the homodimer ensemble. The role of secondary structure formation and interactions in the extramembrane domains in simplifying the homodimer ensemble, stabilizing homodimer formation, and modulating the populations of the three principal TM homodimer structures is examined. Overall, these simulations provide evidence for the role of the C99 extramembrane domains in stabilizing a structurally heterogeneous C99 homodimer, nature of the C99 monomer conformational ensemble, and generally demonstrate the role of extramembrane domains in controlling TM domain protein structures. This study provides a detailed characterization of the structural ensemble of the full-length C99 monomer and homodimer, providing insights into the role of extramembrane regions in modulating TM domain structure.

Results

gREST simulations of C99 monomer were initiated from 16 conformational states identified in 30 Å hydrophobic thickness implicit bilayer REMD simulations (34). Simulations of C99 dimer were prepared from two initial structures of Gly-in and

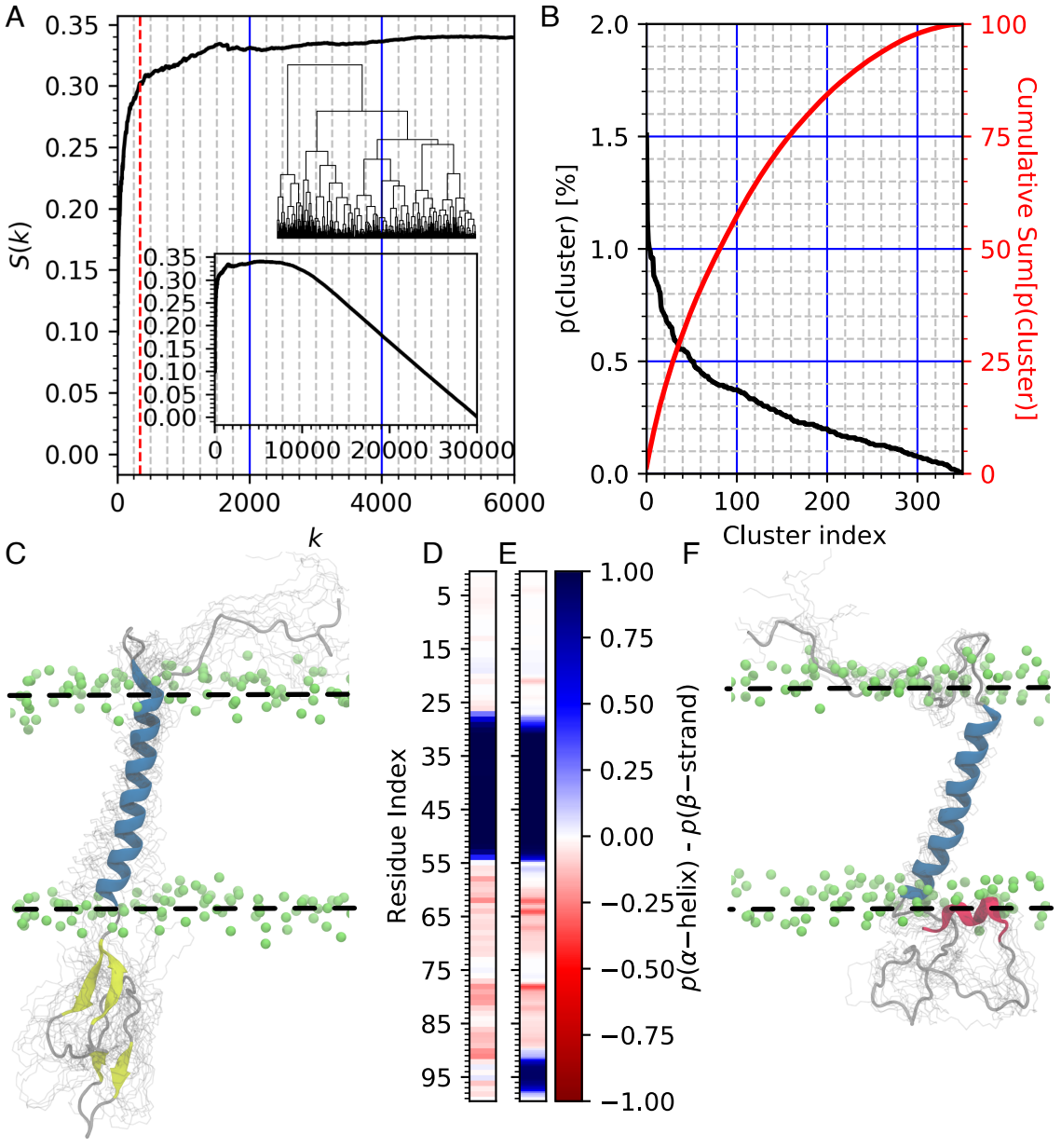

Fig. 2. Analysis of the full-length C99 monomer in hydrated L_d phase bilayers at 310 K. (*A*) Silhouette scores for Ward's minimum variance hierarchical clustering (inset dendrogram) quantify the quality of assignments of configurations to *k* number of clusters. The dashed red line is in the elbow of S(k) at $k = 350$. (*B*) Percent of conformational ensemble for each cluster, p(cluster) (black), ranked in order from largest to smallest, and the cumulative sum of p(cluster) (red). (*C*) First largest cluster, β-sheets in yellow. TMD in blue and POPC phosphorous in green. (*D*) Ensemble-averaged secondary structure of monomer ensemble assigned using STRIDE showing random coil propensities in the N-terminal extramembrane domain and β-strand propensities in the C-terminal extramembrane domain also suggested by NMR probe solvent accessibility in Fig. 1 (31). (*E*) Secondary structure of conformations in which the C-terminal helix is present. (*F*) 5th largest cluster, C-terminal helix in red.

Gly-side TM dimers (24) with random coil constructions of the extramembrane domains which were simulated for 200 ns, allowing for collapse in the radius of gyration of both extramembrane domains. The convergence of these gREST simulations was evaluated using an ergodic measure-like parameter of the observed probability of each trajectory to sample each temperature (*SI Appendix*, Eq. S1), (42–44) to select a subset of sampled conformations to treat as the equilibrium ensemble (*SI Appendix*, Fig. S2 to S4).

Accessibility of C99 to Hydrophobic and Hydrophilic Environments in NMR and MD. The accessibility of backbone amide C99 sites to either the hydrophilic probe Gd-DTPA or the hydrophobic probe 16-DSA was measured experimentally as $1 - I/I_0$, the ratio of NMR peak intensities at low C99 concentrations from probe solutions divided by the corresponding intensities observed in probe-free solutions. Measurements of $1-I/I_0$ of C99 at a 100 μM concentration with Gd-DTPA in POPC (L_d phase, reported here) 10% wt DDMB bicelles are hypothesized to be mostly (70 to 80%) monomeric and match well with the likelihood of protein backbone hydrogen bonds with water (Fig. 1) and height of Cα atoms above the membrane surface (*SI Appendix*, Fig. S1) observed in both C99 monomer and homodimer simulations. The notable measurement of hydrophobic and hydrophilic probe inaccessibility to C99 residues 63 to 73 observed at 200 to 400 μM concentration in SCOR (L_o phase, reported previously) (31) 5 to 10% wt DDMB bicelles, which also feature coexistence of C99 monomers, homodimers, and homotrimers, is not present in these L_d phase POPC systems, in which residues 63 to 73 are solvent exposed.

The simulations indicate that C99 homodimer JM domain residues 16 to 23 do not reinsert to the membrane, as the JM domain does in the monomer, but maintain similar solvent accessibility. Additionally, the extreme C-terminal domain of the homodimer does not exhibit the tendency to associate as an amphipathic helix with the model membrane surface as observed for the monomer in MD simulation and in NMR experiments. This abolition of association of the JM domains and extreme C terminus with the membrane surface, as previously observed for the C99 monomer (34), is a general indication of a change in membrane surface interactions induced by homodimerization.

The C99 Monomer Is Intrinsically Disordered and Features C-terminal β-strands. Using the RMSD of pair distances of N- and C-terminal Cα as a criterion of the distance between conformations of C99 (Eq. **1**), we clustered the C99 monomer and homodimer in POPC bilayers using Ward's minimum variance hierarchical clustering (Eq. **2**) (45) and computed silhouette scores (Eq. **3**), (46) S, for k number of clusters to determine the most appropriate number of clusters and cluster assignments to describe C99 conformations. Silhouette scores for the C99 monomer plateau on the order of thousands of states and decrease linearly to zero on the order of tens of thousands of states (Fig. 2*A*). As such, monomeric C99 is similar to monomeric Aβ in that the protein conformational ensemble has an enormous number of states (47). For the sake of partitioning the ensemble into some nonideal partitioning of C99 monomer conformational states, we chose to make a clustering of 350 states, which is an apparent elbow in $S(k)$. With this clustering, we find the population of C99 monomer conformational states to be a mix of exponential and linear decay as a function of cluster size (Fig. 2*B*).

The C99 monomer expresses a significant but widely varying β-strand character throughout the extramembrane domains (Fig. 2 *C* and *D*), as can be seen in the first 100 largest clusters in the $k = 350$ clustering (*SI Appendix*, Fig. S5). The JM domain residues 18 to 23 were found to participate in α-helices and β-strands at 4.17 and 5.15% propensities in the ensemble. The N-terminal residues 2 to 6 were found to participate in α-helices and β-strands at a mere 0.67 and 4.62% propensity, suggesting that the obfuscation of residues 2 to 6, 18 to 23 from paramagnetic

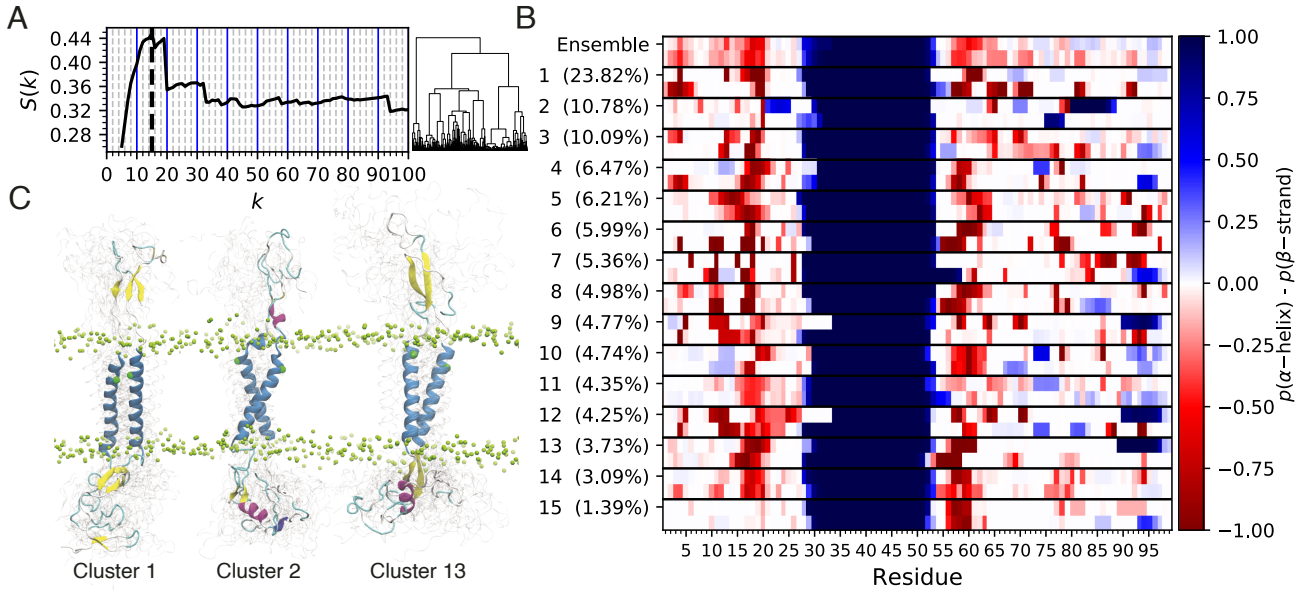

Fig. 3. Analysis of the full-length C99 homodimer in hydrated L_d phase bilayers at 310 K. (*A*) Silhouette scores for Ward's minimum variance hierarchical clustering (inset dendrogram) for partitioning of C99 homodimer conformations into k number of clusters. The dashed black line identifies the optimal number of clusters at $k = 15$. (*B*) Ensemble average and population-ranked cluster secondary structures assigned using STRIDE show many unique β-strand motifs in the N- and C-terminal extramembrane domains. (*C*) Representative conformations of 1st, 2nd, and 13th largest clusters. TMD in blue, C-terminal helix in red, and POPC phosphorous and Gly33 Cα in green.

probes in both POPC and SCOR bicelles may result from these transient structures.

The C-terminal residues 93 to 97 were found to participate in α-helices at 13.94% propensity in the ensemble. The β-strand propensities observed throughout the C99 monomer C-terminal residues also persist in the subset of conformations where residues 93 to 97 are α-helical. Even in the presence of the α-helix (Fig. 2 *E* and *F*), β-strand propensities in the C-terminal extramembrane domain are present. The presence and rapid exchange of these β-strands seems a likely explanation for the random coil secondary structure observed throughout extramembrane domains despite the reduction in spin relaxation rates. Previous NMR characterization of the 47 C-terminal APP residues in aqueous solution identified residues $T_{72}PEE_{75}$ and $N_{88}PTY_{91}$ as exhibiting a propensity for forming reverse turns (48). We observe that these subdomains similarly exhibit a propensity for β-strands in our simulation of the monomeric state of membrane-bound C99.

The C99 Homodimer Exhibits Metastable States. Silhouette scores of C99 homodimer clustering show that 15 clusters is the optimal number to represent the conformational ensemble (Fig. 3*A*). The secondary structure of these clusters is characterized by unique β-strands formed within and between N- and C-terminal extramembrane domains (Fig. 3*B*). The JM domain only partially manifests in the 2nd largest cluster. The C-terminal α-helix manifests in the 9th, 10th, 12th, and 13th largest clusters. However, neither the C-terminal nor JM domain α-helices observed in the homodimer associate with the membrane surface, as they do in the monomer (Fig. 1). The formation of N-terminal domain β-strands and the conversion of C-terminal domain β-strands to stabilization as a consequence of interprotein interactions are particularly important features introduced by homodimerization. As visualized in Fig. 3*C*, the conformation of extramembrane domains is found to impact the structure of the C99 homodimer TM interface.

C99 Changes Upon Homodimerization. Aside from a dramatic reduction in the number of states, there are other notable ensemble-averaged conformational changes upon C99 homodimerization. All secondary structure motifs in the C99 monomer result from intraprotein backbone hydrogen bonds (Fig. 4*A*). However, specific intraprotein interactions in the dimer extramembrane domains are observed (Fig. 4*B*). Despite this change, the accessibility of the C99 backbone to water decreases only slightly upon homodimerization, and the TM domain remains protected from water (Fig. 1). Interprotein interactions introduced upon homodimerization introduce β-strands to the N-terminal extramembrane domain involving residues 2 to 6 and 10 to 20 (Fig. 4*C*). Also observed in the dimer is a β-strand involving the Lys anchor of residues 53 to 54 and residues 57 to 61 near the membrane surface, much like the β-strand introduced in the structure of C83 complexes with γ-secretase, as determined via cryo-EM (49). C83 is the C-terminal TM domains of APP released from longer forms of APP by α-secretase cleavage. For the C99 dimer, the β-strand propensities are slightly reduced elsewhere in the C-terminal extramembrane domain. The introduction of β-strands in the N-terminal domain Aβ sequence and the relative increase in α-helical propensity in the C-terminal extramembrane domains are both representative of the predicted secondary structure propensities in putative Aβ oligomers and fibrils and C99 C-terminal domain binding motifs, respectively.

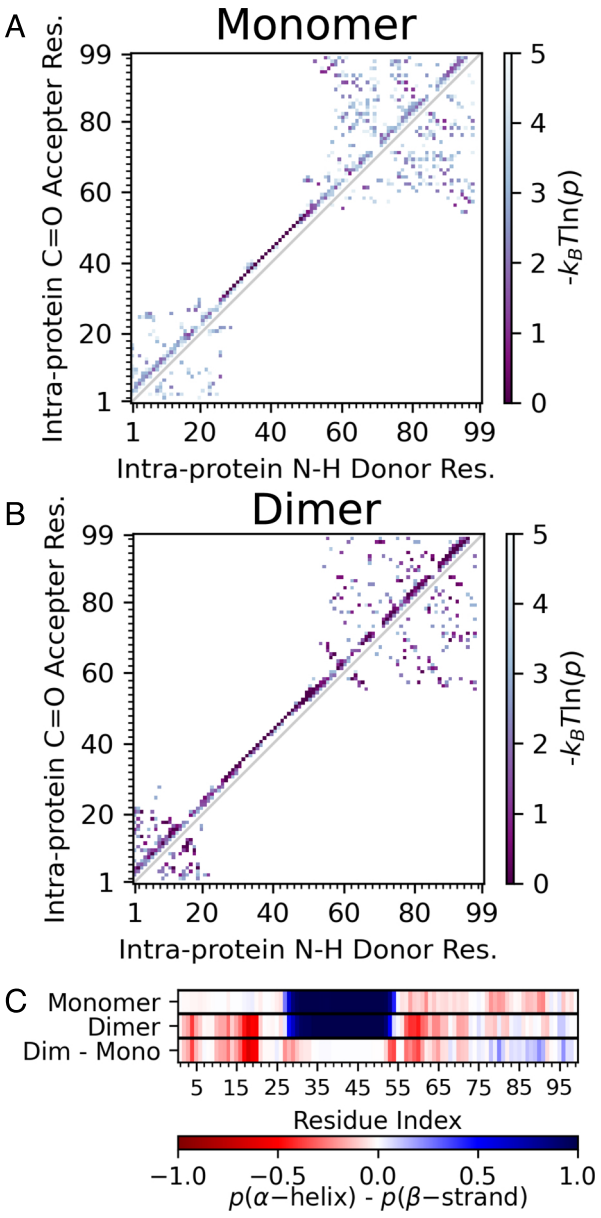

Fig. 4. Differences in ensemble-averaged structure of monomeric and homodimeric C99. Intraprotein backbone N-H···O=C hydrogen bonds for monomer (*A*) and dimer (*B*) demonstrate formation of more stable intraprotein hydrogen bonds in extramembrane domains. (*C*) Ensemble average secondary structures of monomeric C99, homodimeric C99, and the change in secondary structure upon homodimerization show introduction of N-terminal β-strands, C-terminal juxtamembrane β-strands, and C-terminal extramembrane α-helices upon homodimerization.

Intramembrane Homodimer Motifs Depend on Extramembrane Domain States. The TM tilt relative to the bilayer normal (θ) and the Gly33–Gly34 hinge angle (κ), previously demonstrated to be sensitive to membrane thickness and curvature (23, 24, 34), are unimodal and change only slightly upon homodimerization, decreasing in κ and increasing in θ as the TM domains interact (Fig. 5 *A* and *B*). Despite this, the dimerization interface experiences significant changes due to the extramembrane domain.

We evaluate the homodimer superhelix handedness and find 60% and 40% of conformations adopt left- and right-handed superhelices (Fig. 5*C*). We also evaluate the superhelix handedness $\phi_{\text{Gly29Gly37}}$, xy-plane Crick angles $\Psi_{\text{Crick}}^{\text{Gly33}}$, $\Psi_{\text{Crick}}^{\text{Ile45}}$,

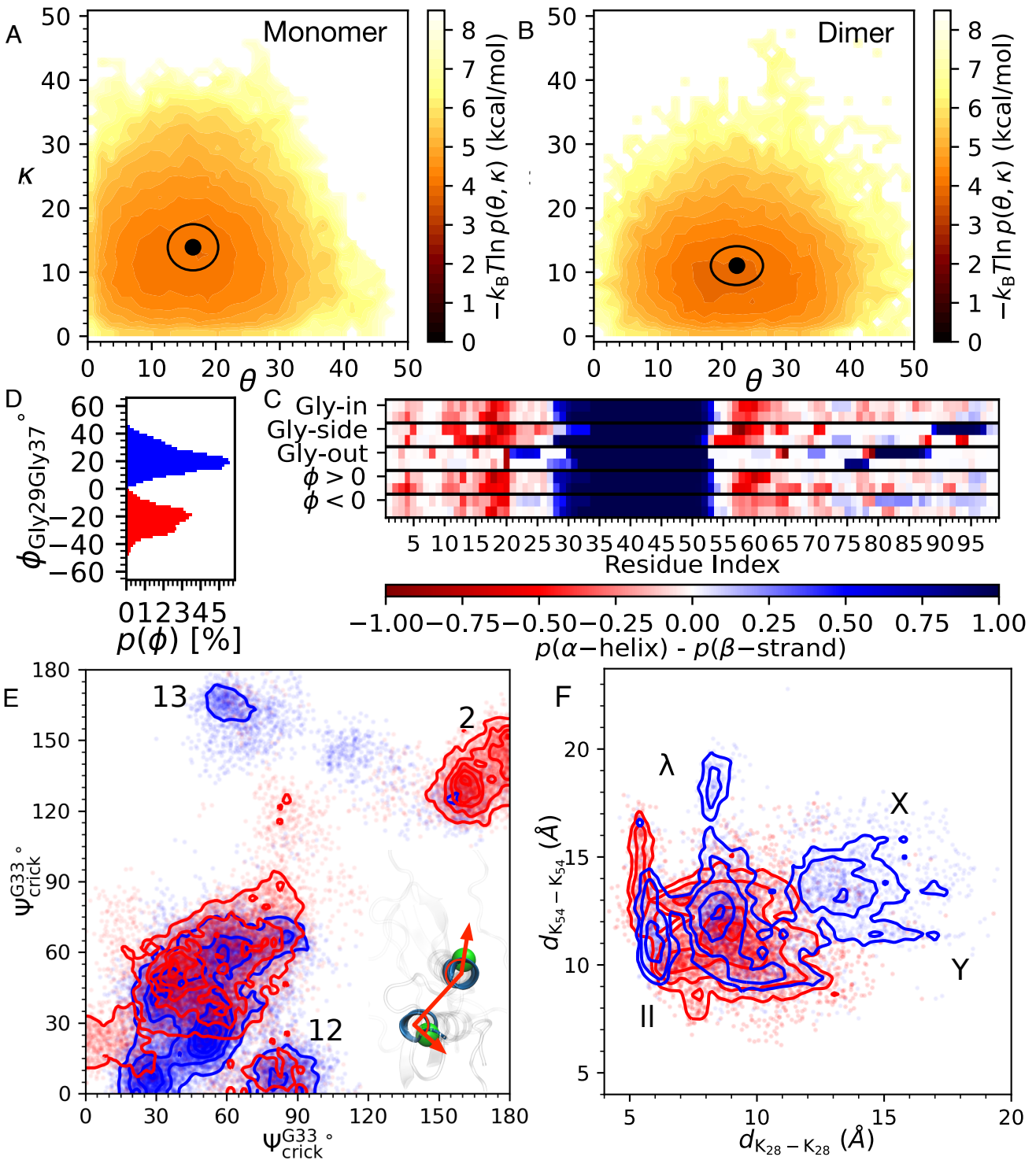

Fig. 5. Monomeric (*A*) and homodimeric (*B*) C99 TM domain tilt angle θ and Gly33-Gly34 hinge angle κ demonstrate a slight stiffening in the Gly-Gly hinge and a broader TM tilt upon homodimerization. (*C*) Dihedral of Gly29–Gly33–Gly33–Gly29 in the full-length homodimer, defining a left- ($\phi > 0$) and right-handed ($\phi < 0$) superhelix. (*D*) Secondary structure of C99 for Gly-side (12th and 13th-largest clusters), Gly-out (2nd-largest cluster), and Gly-out (the remaining clusters), and secondary structure for left- and right-handed superhelices. (*E*) Gly33–Gly33 Crick angles in the xy-plane, colored by conformations belonging to left- and right-handed superhelices. Inset homodimer conformation from 13th-largest cluster shows a Gly-side $\Psi \sim (90°, 150°)$ Crick angle configuration. (*F*) Lys28–Lys28 and Lys54–Lys54 Cα pair distances. Only left-handed superhelices (blue) access X- and Y-shaped homodimer motifs.

and the pair distances of the alpha carbons of Lys28–Lys28 and Lys28–Lys28 (*SI Appendix*, Figs. S6 to S10). The secondary structure of the C99 homodimer adopting Gly-in, Gly-side, and Gly-out and left- and right-handed superhelical structures reveals that the orientation of the glycine zipper, often used to define the C99 homodimer, is influenced by extramembrane domain structures and that the handedness of the TM superhelix is not strongly influenced by extramembrane domains (Fig. 5*D*).

We see that the extramembrane domains determine the relative rotation of the C99 TMDs and partially determine features of the

superhelix, such as these interhelix distances and the superhelix handedness. The Gly-out structure is uniquely accessed by left-handed superhelices and is only described by the unique, 2nd-largest cluster (Fig. 5*E*), which features α-helices in residues 21 to 24, 75 to 78, and 79 to 88. The Gly-side structure is described by both the 12 and 13th-largest clusters, which feature the C-helix (residue 89 to 99 α-helix). The relative alignment of the homodimer is defined using the pair distances for the Lys28–Lys28 and Lys28–Lys28 alpha carbons, which sit in the lipid bilayer surface and define ||, X, Y, and λ-shaped TM homodimer motifs previously observed to depend on surface curvature and lipid composition (23, 24). We observe that particular alignment shapes are not exclusively accessible to a particular cluster (*SI Appendix*, Fig. S7) but do find that the expanded Y and X are accessible only to left-handed superhelices (Fig. 5*F*).

Homodimer Motifs Are Not Primarily Controlled by Glycine Zipper Formation. Mutations of residues Gly29 and Gly33 have been demonstrated to reduce the amyloidogenic $A\beta_{42}/A\beta_{40}$ ratio (50). Familial AD-associated mutations in residues 42 to 46, 48, 52, and 53 have been demonstrated to reduce APP dimerization and enhance the proamyloidogenic $A\beta_{42}/A\beta_{40}$ ratio (8, 51–54). It is believed that the motif driving formation of the C99 homodimer is defined by Cα-H$\cdots$O=C Gly–Gly hydrogen bonds in the glycine zipper as this stabilizes the homodimerization of GxxxG-containing TM polypeptides that lack extramembrane domains (55–57), including the excised C99 TM domain (19, 54). However, evaluation of the interprotein backbone hydrogen bond propensities observed in the C99 homodimer reveals a surprising lack of propensity for the glycine zipper to form the Cα-H$\cdots$O=C Gly–Gly hydrogen bond. Only the 1st largest cluster, 24% of the ensemble, features Gly–Gly hydrogen bonds between Gly29–Gly29, Gly29–Gly33, and Gly33–Gly33. Rather, it is extramembrane residues that prominently form interprotein backbone–backbone hydrogen bonds and that are the primary determinant of homodimer conformations (Fig. 6).

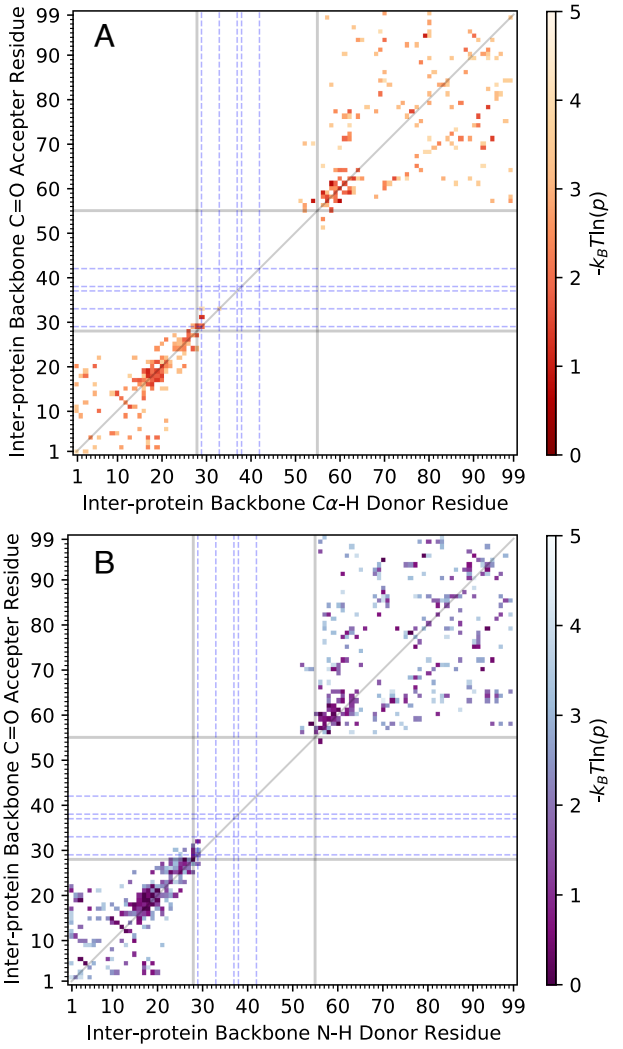

Fig. 6. Interprotein backbone hydrogen bonds within 2.5 Å hydrogen distance and ≥ 150 angle between the C99 carbonyl oxygen acceptor and (*A*) Cα–hydrogen donor and (*B*) amide hydrogen donor. Solid gray lines specify the beginning of the TM domain from residues 28 to 55. Dashed blue lines identify TM glycine residues 29, 33, 37, and 38. The lack of Cα–hydrogen bonding in the TM domain involving the glycine zipper and frequently observed N- and C-terminal hydrogen bonds suggests that these extramembrane domain interactions stabilize the dimer.

Discussion

We have performed replica exchange simulations of the full-length C99 monomer and homodimer in neutral pH conditions and liquid disordered POPC lipid bilayers at 310 K. Through these extensive MD simulations, we have determined that the C99 monomer extramembrane domains are intrinsically disordered and that the C-terminal residues generally form thousands of unique β-strand configurations that occlude the C-terminal residues from interaction with the hydrophobic and hydrophilic probes used in NMR experiments by Hutchison et al. at dilute C99 concentrations (Figs. 1 and 2) (31). Finding our monomeric simulations to independently recapitulate the same conclusions as NMR characterizations, we propose that homodimeric simulations performed with the same simulation parameters would produce a reasonable conformational ensemble of the C99 homodimer.

We found the C99 homodimer to be stabilized by 15 conformational states, principally defined by unique β-strands formed between C99 proteins (Fig. 3). These states are stabilized through interprotein interactions, which in turn allow for the formation of more stable α-helices (Fig. 4). Of these states, we found that the Gly-side and Gly-out homodimer TM helix configurations were associated with clusters featuring the C-terminal helix (residues 91 to 99) and unique helices in the N and C termini (21 to 25, 75 to 79, 81 to 88), respectively (Fig. 5). Cα-H$\cdots$O=C hydrogen bonds, previously demonstrated to stabilize homodimers formed by the excised C99 TM polypeptide and to generally stabilize TM polypeptides featuring glycine zippers, were not found to be present in the C99 homodimer with the exception of the 1st-largest cluster (*SI Appendix*, Figs. S11 and S12), though an abundance of extramembrane domain backbone hydrogen bonds form (Fig. 6) defining the unique secondary structures which characterize the 15 conformational states.

Insight was provided into the role of cooperative effects in homodimer structure formation. Right-handed superhelices were found to produce only parallel-packed TM helices, with a preference for Gly-out conformations. Left-handed superhelices were found to promote crossed-angle TM helices, with a preference for Gly-side conformations (Fig. 5). Taken together, these observations demonstrate that extramembrane domain structure influences TM domain conformational states. This suggests that accurate characterization of TM domain structure in the C99 homodimer requires inclusion of extramembrane domains, a consideration that may be important for all TM proteins exhibiting homodimer formation.

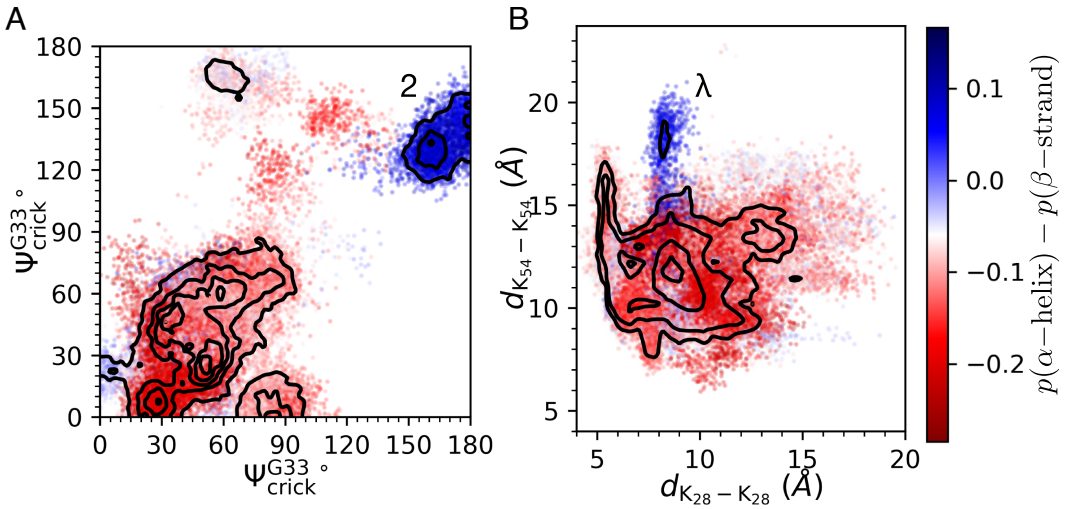

Fig. 7. Homodimer extramembrane residues 1 to 28 and 55 to 99 average secondary structure propensity as a function of (*A*) Gly33–Gly33 Crick angles and (*B*) Lys28–Lys28, Lys54–Lys54 Cα pair distances demonstrate the unique dimer interfaces induced by promotion of α helices in extramembrane domains, particularly formation of the juxtamembrane helix in the 2nd-largest cluster.

We rarely observe formation of an N-terminal juxtamembrane helix in our simulations of C99 in our model membrane at neutral pH. We observe that when the juxtamembrane helix is present, the C99 homodimer tends to adopt a Gly-out conformation with reduced extramembrane β-strand propensity (Fig. 7). The adoption of the higher free energy Gly-out homodimer structure (55–57) or a reduction in extramembrane secondary structure may act to destabilize the homodimer. It has been observed that low pH or the presence of FAD mutations D22Q and E23N enhances formation of α helices in the juxtamembrane region (25, 26). As homodimer formation is evidenced to protect C99 from cleavage by γ-secretase (18), if FAD mutants D22Q or E23N act to destabilize homodimer formation, they would stand to increase Aβ production.

The first model of the C99 homodimer TM domain was presented by Münter et al. (50) Since that time, further progress toward characterization of the structure of the full-length C99 homodimer, and in particular, the extramembrane domain regions, has proved challenging. Using all-atom models and well-parameterized gREST simulations, this study provides insights into the nature of both the TM and extramembrane domains, including how extramembrane structure modulates TM structure and the dimer interface. While there are many structures of TM homodimer proteins, most have been determined using congener proteins (TM domain-only fragments) lacking most or all of the flanking extramembrane domains. The demonstrated impact of secondary structure in the extramembrane domain regions upon TM domain structure provides insight into the role of extramembrane domains in affecting TM domain structures in general and on the complexity of the C99 homodimer in particular. The computational protocol employed in this study provides one effective means of providing a more complete understanding of the role of extramembrane domains in modulating the structure and function of TM protein homodimers.

Materials and Methods

All simulations were performed using MD software GENESIS ver 1 (SPDYN) (58, 59) following simulation methodology similar to that of Matsuoka et al. (60) Software and methods for used initial structure preparation, simulation parameters (24, 25, 34, 61–63), convergence of simulations to equilibrium (42–44), and cluster scoring (45, 46) are described in *SI Appendix*. Experimental details for C99 characterization in POPC lipid DDMB bicelles are described as previously reported (31) and in *SI Appendix*.

Hierarchical Clustering. The RMSD of Cα pair distances of two nonintersecting sets of N- and C-terminal extramembrane residues, 1 to 28 and 53 to 99, between each pair of conformations of the monomer or dimer is used as the clustering distance metric, expressed as

$$\text{dRMSD}(i,j) = \sqrt{\frac{1}{N_{C\alpha\text{pairs}}} \sum_{k}^{N_{C\alpha\text{pairs}}} (\vec{x}_i(d_k) - \vec{x}_j(d_k))^2}, \quad [1]$$

where $\vec{x}_i(d_k)$ ($\vec{x}_j(d_k)$) is the *kth* Cα pair distance belonging to the *ith* (*jth*) conformation. Each conformation is agglomeratively clustered using Ward's method (45), in which the cluster size and distance between the cluster centers of all pairs of clusters at each iteration are evaluated to minimize the deviation in dRMSD in the next set of clusters defined as

$$\Delta(A,B) = \sum_{i\in A\cup B} \text{dRMSD}(\vec{x}_i, \vec{x}_{A\cup B}) - \sum_{i\in A} \text{dRMSD}(\vec{x}_i, \vec{x}_A) - \sum_{i\in B} \text{dRMSD}(\vec{x}_i, \vec{x}_B) = \frac{n_A n_B}{n_A + n_B} \text{dRMSD}(\vec{x}_A, \vec{x}_B), \quad [2]$$

where $\Delta(A,B)$ is the cluster–cluster distance between two clusters indexed by A and B, $\vec{x}_A$ ($\vec{x}_B$) is the conformation with minimum distance to all other conformations (the cluster center) in cluster A (B), $\vec{x}_{A\cup B}$ is the cluster center of the putative merging of cluster A with B, and n_A (n_B) is the number of conformations belonging to cluster A (B). This approach to clustering extramembrane domain conformations is similar to that employed by Baul et al. (64) to cluster a variety of intrinsically disordered proteins and by Chakraborty et al. to characterize the Aβ monomer conformational ensemble in solution (47).

Rather than use a somewhat arbitrary cutoff in the dendrogram to define clusters, we instead use a scoring approach. The optimal number

of clusters for the ensemble is evaluated using the Silhouette scoring metric (46)

$$a(i) = \frac{1}{n_A - 1} \sum_{i,j \in A}^{n_A} \mathrm{dRMSD}(\vec{x}_i, \vec{x}_j),$$

$$b(i) = \frac{1}{n_B} \sum_{i \in A, j \in B}^{n_B} \mathrm{dRMSD}(\vec{x}_i, \vec{x}_j), \quad [3]$$

$$S(k) = \frac{1}{N} \sum_{i}^{N} \frac{b(i) - a(i)}{\max\{a(i), b(i)\}},$$

where $b(i)$ is computed for the cluster B, which is the nearest cluster to conformation i outside of cluster A to which i belongs, $\max\{a(i), b(i)\}$ is the greater of $a(i)$ or $b(i)$, N is the total number of conformations, and $S(k)$ is the silhouette score for a partitioning of the system to k number of clusters. $S(k)$ is maximized when the number of clusters k best-partitions data into well-separated clusters. Silhouette scores have been used for optimizing numbers of clusters in the past even when clustering on kinetic transitions rather than conformations, such as in the construction of Markov state models by Nedialkova et al. (65).

Data, Materials, and Software Availability. Simulation conformations at equilibrium, conformational cluster assignments, simulation input parameters, and NMR paramagnetic probe peak intensities. data have been deposited in Mendeley (https://doi.org/10.17632/j6rtdskrvc.1). Previously published data were used for this work [Hutchison, J. M. et al. Bicelles Rich in both Sphingolipids and Cholesterol and Their Use in Studies of Membrane Proteins. J. Am. Chem. Soc. 142, 12715-12729 (2020). DOI: 10.1021/jacs.0c04669].

ACKNOWLEDGMENTS. G.A.P. was supported as a RIKEN short-term international program associate and by R01 GM107703. J.E.S. and D.T. were supported by R01 GM107703. J.M.H. was supported by NIH T32 CA00958229 and F31 AG061984. C.R.S. was supported by RF1 AG056147. Y.S. and D.M. were supported by RIKEN Pioneering Projects "Integrated Lipidology," "Glycolipidologue Initiative," and "Dynamic Structure Biology." The RIKEN K-computer (ra000009) and Institute for Molecular Science Molecular Simulator were used in the production of gREST simulations. Data availability: C99 conformations sampled at equilibrium at 310 K in monomeric and homodimeric states, homodimer cluster assignments, GENESIS gREST simulation input parameters, and NMR paramagnetic probe data are available on the Mendeley database G.A.P.; D.M.; J.M.H.; C.R.S.; Y.S.; J.E.S.; D.T. (2022), "Data from: Formation of beta-strands in the extramembrane controls dimerization of transmembrane helices in amyloid precursor protein C99 intrinsically disordered domains," Mendeley Data, V1, DOI: 10.17632/j6rtdskrvc.1.

Author affiliations: [a]Department of Chemistry, Boston University, Boston, MA 02215; [b]Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20814; [c]Theoretical Molecular Science Laboratory, RIKEN Cluster for Pioneering Research, Saitama 351-0198, Japan; [d]Department of Biochemistry, Vanderbilt University, Nashville, TN 37240; [e]Center for Structural Biology, Vanderbilt University, Nashville, TN 37240; [f]Department of Medicine, Vanderbilt University School of Medicine, Nashville, TN 37240; and [g]Department of Chemistry, University of Texas, Austin, TX 78712

Author contributions: G.A.P., Y.S., J.E.S., and D.T. designed research; G.A.P., D.M., and J.M.H. performed research; G.A.P. analyzed data; and G.A.P., C.R.S., Y.S., J.E.S., and D.T. wrote the paper.

1. D. J. Selkoe, J. Hardy, The amyloid hypothesis of Alzheimer's disease at 25 years. *EMBO Mol. Med.* **8**, 595–608 (2016).
2. P. H. Nguyen *et al.*, Amyloid oligomers: A joint experimental/computational perspective on Alzheimer's disease, Parkinson's disease, type II diabetes, and amyotrophic lateral sclerosis. *Chem. Rev.* **121**, 2545–2647 (2021).
3. M. Takami *et al.*, γ-Secretase: Successive tripeptide and tetrapeptide release from the transmembrane domain of β-carboxyl terminal fragment. *J. Neurosci.* **29**, 13042–13052 (2009).
4. I. Coburger *et al.*, Analysis of the overall structure of the multi-domain amyloid precursor protein (APP). *PLoS ONE* **8**, 1–12 (2013).
5. T. Tomita, Molecular mechanism of intramembrane proteolysis by γ-secretase. *J. Biochem.* **156**, 195–201 (2014).
6. S. Kumar-Singh *et al.*, Mean age-of-onset of familial Alzheimer disease caused by presenilin mutations correlates with both increased Aβ42 and decreased Aβ40. *Hum. Mutat.* **27**, 686–695 (2006).
7. E. Winkler *et al.*, Generation of Alzheimer disease-associated amyloid beta 42/43 peptide by gamma-secretase can be inhibited directly by modulation of membrane thickness. *J. Biol. Chem.* **287**, 21326–21334 (2012).
8. T. H. Xu *et al.*, Alzheimer's disease-associated mutations increase amyloid precursor protein resistance to γ-secretase cleavage and the Aβ42/Aβ40 ratio. *Cell Discov.* **2**, 16026 (2016).
9. N. Tang, K. P. Kepp, Aβ42/Aβ40 ratios of presenilin 1 mutations correlate with clinical onset of Alzheimer's disease. *J. Alzheimer's Dis.* **66**, 939–945 (2018).
10. P. Osenkowski, W. Ye, R. Wang, M. S. Wolfe, D. J. Selkoe, Direct and potent regulation of γ-secretase by its lipid microenvironment. *J. Biol. Chem.* **283**, 22529–22540 (2008).
11. D. Petit *et al.*, Aβ profiles generated by Alzheimer's disease causing PSEN1 variants determine the pathogenicity of the mutation and predict age at disease onset. *Mol. Psychiatry* **27**, 2821–2832 (2022).
12. M. Okochi *et al.*, γ-secretase modulators and presenilin 1 mutants act differently on presenilin/γ-secretase function to cleave Aβ42 and Aβ43. *Cell Rep.* **3**, 42–51 (2013).
13. M. Szaruga *et al.*, Alzheimer's-causing mutations shift Aβ length by destabilizing γ-secretase-Aβn interactions. *Cell* **170**, 443–456.e14 (2017).
14. D. Petit *et al.*, Extracellular interface between APP and Nicastrin regulates Aβ length and response to γ' secretase modulators. *EMBO J.* **38**, 2953–2966 (2019).
15. M. Silber, M. Hitzenberger, M. Zacharias, C. Muhle-Goll, Altered Hinge Conformations in APP Transmembrane Helix Mutants May Affect Enzyme-Substrate Interactions of γ-Secretase. *ACS Chem. Neurosci.* **11**, 4426–4433 (2020).
16. B. Dehury, N. Tang, R. Mehra, T. L. Blundell, K. P. Kepp, Side-by-side comparison of Notch- And C83 binding to γ-secretase in a complete membrane model at physiological temperature. *RSC Adv.* **10**, 31215–31232 (2020).
17. S. Eggert, B. Midthune, B. Cottrell, E. H. Koo, Induced dimerization of the amyloid precursor protein leads to decreased amyloid-β protein production. *J. Biol. Chem.* **284**, 28943–28952 (2009).
18. E. Winkler, A. Julius, H. Steiner, D. Langosch, Homodimerization protects the amyloid precursor protein C99 fragment from cleavage by gamma-secretase. *Biochem.* **54**, 6149–6152 (2015).
19. K. D. Nadezhdin, O. V. Bocharova, E. V. Bocharov, A. S. Arseniev, Dimeric structure of transmembrane domain of amyloid precursor protein in micellar environment. *FEBS Lett.* **586**, 1687–1692 (2012).
20. D. T. Moore, B. W. Berger, W. F. DeGrado, Protein-protein interactions in the membrane: Sequence, structural, and biological motifs. *Structure* **16**, 991–1001 (2008).
21. N. Miyashita, J. E. Straub, D. Thirumalai, Y. Sugita, Transmembrane structures of amyloid precursor protein dimer predicted by replica-exchange molecular dynamics simulations. *J. Am. Chem. Soc.* **131**, 3438–3439 (2009).
22. N. Miyashita, J. E. Straub, D. Thirumalai, Structures of β-amyloid peptide 1–40, 1–42, and 1–55–the 672–726 fragment of APP–in a membrane environment with implications for interactions with γ-secretase. *J. Am. Chem. Soc.* **131**, 17843–17852 (2009).
23. L. Dominguez, L. Foster, S. C. Meredith, J. E. Straub, D. Thirumalai, Structural heterogeneity in transmembrane amyloid precursor protein homodimer is a consequence of environmental selection. *J. Am. Chem. Soc.* **136**, 9619–9626 (2014).
24. L. Dominguez, L. Foster, J. E. Straub, D. Thirumalai, Impact of membrane lipid composition on the structure and stability of the transmembrane domain of amyloid precursor protein. *Proc. Natl. Acad. Sci. U.S.A.* **113**, E5281–E5287 (2016).
25. A. Panahi, A. Bandara, G. A. Pantelopulos, L. Dominguez, J. E. Straub, Specific binding of cholesterol to C99 domain of amyloid precursor protein depends critically on charge state of protein. *J. Phys. Chem. Lett.* **7**, 3535–3541 (2016).
26. G. A. Pantelopulos, A. Panahi, J. E. Straub, Impact of cholesterol concentration and lipid phase on structure and fluctuation of amyloid precursor protein. *J. Phys. Chem. B* **124**, 10173–10185 (2020).
27. A. J. Beel *et al.*, Structural studies of the transmembrane C-terminal domain of the amyloid precursor protein (APP): Does APP function as a cholesterol sensor? *Biochem.* **47**, 9428–9446 (2008).
28. P. J. Barrett *et al.*, The amyloid precursor protein has a flexible transmembrane domain and binds cholesterol. *Science* **336**, 1168–1171 (2012).
29. Y. Song, K. F. Mittendorf, Z. Lu, C. R. Sanders, Impact of bilayer lipid composition on the structure and topology of the transmembrane amyloid precursor C99 protein. *J. Am. Chem. Soc.* **136**, 4093–4096 (2014).
30. T. A. Caldwell *et al.*, Low-q bicelles are mixed micelles. *J. Phys. Chem. Lett.* **9**, 4469–4473 (2018).
31. J. M. Hutchison *et al.*, Bicelles rich in both sphingolipids and cholesterol and their use in studies of membrane proteins. *J. Am. Chem. Soc.* **142**, 12715–12729 (2020).
32. G. A. Pantelopulos, J. E. Straub, Regimes of complex lipid bilayer phases induced by cholesterol concentration in MD simulation. *Biophys. J.* **115**, 2167–2178 (2018).
33. T. Kakeshpour *et al.*, A lowly populated, transient β-sheet structure in monomeric Aβ1-42 identified by multinuclear NMR of chemical denaturation. *Biophys. Chem.* **270**, 106531 (2021).
34. G. A. Pantelopulos, J. E. Straub, D. Thirumalai, Y. Sugita, Structure of APP-C99 1–99 and implications for role of extra-membrane domains in function and oligomerization. *Biochim. Biophys. Acta (BBA) - Biomembr.* **1860**, 1698–1708 (2018).
35. M. Yuksel, O. Tacal, Trafficking and proteolytic processing of amyloid precursor protein and secretases in Alzheimer's disease development: An up-to-date review. *Eur. J. Pharmacol.* **856**, 172415 (2019).
36. E. J. Fernandez-Perez, C. Peters, L. G. Aguayo, Membrane damage induced by amyloid beta and a potential link with neuroinflammation. *Curr. Pharm. Des.* **22**, 1295–1304 (2016).

37. D. R. Roe, A. Okur, L. Wickstrom, V. Hornak, C. Simmerling, Secondary structure bias in generalized born solvent models: Comparison of conformational ensembles and free energy of solvent polarization from explicit and implicit solvation. *J. Phys. Chem. B* **111**, 1846–1857 (2007).
38. J. Huang *et al*., CHARMM36m: An improved force field for folded and intrinsically disordered proteins. *Nat. Methods* **14**, 71–73 (2016).
39. M. Kamiya, Y. Sugita, Flexible selection of the solute region in replica exchange with solute tempering: Application to protein-folding simulations. *J. Chem. Phys.* **149**, 072304 (2018).
40. R. Capone *et al*., The C99 domain of the amyloid precursor protein resides in the disordered membrane phase. *J. Biol. Chem.* **296**, 100652 (2021).
41. C. Venugopal, C. Demos, K. Jagannatha Rao, M. Pappolla, K. Sambamurti, Beta-Secretase: Structure, function, and evolution. *CNS Neurol. Disord. - Drug Targets.* **7**, 278–294 (2008).
42. D. Thirumalai, R. D. Mountain, T. R. Kirkpatrick, Ergodic behavior in supercooled liquids and in glasses. *Phys. Rev. A* **39**, 3563–3574 (1989).
43. D. Thirumalai, R. D. Mountain, Ergodic convergence properties of supercooled liquids and glasses. *Phys. Rev. A* **42**, 4574–4587 (1990).
44. J. E. Straub, D. Thirumalai, Exploring the energy landscape in proteins. *Proc. Natl. Acad. Sci. U.S.A.* **90**, 809–813 (1993).
45. J. H. W. Jr., Hierarchical grouping to optimize an objective function. *J. Am. Stat. Ass.* **58**, 236–244 (1963).
46. P. J. Rousseeuw, Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *J. Comput. Appl. Math.* **20**, 53–65 (1987).
47. D. Chakraborty, J. E. Straub, D. Thirumalai, Differences in the free energies between the excited states of Aβ40 and Aβ42 monomers encode their aggregation propensities. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 19926–19937 (2020).
48. C. D. Kroenke, D. Ziemnicka-Kotula, J. Xu, L. Kotula, A. G. Palmer, Solution conformations of a peptide containing the cytoplasmic domain sequence of the β amyloid precursor protein. *Biochem.* **36**, 8145–8152 (1997).
49. R. Zhou *et al*., Recognition of the amyloid precursor protein by human γ-secretase. *Science* **363**, eaaw0930 (2019).
50. L. M. Münter *et al*., GxxxG motifs within the amyloid precursor protein transmembrane sequence are critical for the etiology of Aβ42. *EMBO J.* **26**, 1702–1712 (2007).
51. P. M. Gorman *et al*., Dimerization of the transmembrane domain of amyloid precursor proteins and familial Alzheimer's disease mutants. *BMC Neurosci.* **9**, 1–11 (2008).
52. S. Weggen, D. Beher, Molecular consequences of amyloid precursor protein and presenilin mutations causing autosomal-dominant Alzheimer's disease. *Alzheimer's Res. Therapy* **4**, 9 (2012).
53. M. Dimitrov *et al*., Alzheimer's disease mutations in APP but not γ-secretase modulators affect epsilon-cleavage-dependent AICD production. *Nat. Commun.* **4**, 2246 (2013).
54. Y. Yan *et al*., Dimerization of the transmembrane domain of amyloid precursor protein is determined by residues around the γ-secretase cleavage sites. *J. Biol. Chem.* **292**, 15826–15837 (2017).
55. D. Langosch, B. Brosig, H. Kolmar, H. J. Fritz, Dimerisation of the glycophorin A transmembrane segment in membranes probed with the ToxR transcription activator. *J. Mol. Biol.* **263**, 525–530 (1996).
56. B. K. Mueller, S. Subramaniam, A. Senes, A frequent, GxxxG-mediated, transmembrane association motif is optimized for the formation of interhelical Cα-H hydrogen bonds. *Proc. Natl. Acad. Sci. U.S.A.* **111**, (2014).
57. S. M. Anderson, B. K. Mueller, E. J. Lange, A. Senes, Combination of Cα-H hydrogen bonds and van der Waals packing modulates the stability of GxxxG-mediated dimers in membranes. *J. Am. Chem. Soc.* **139**, 15774–15783 (2017).
58. J. Jung *et al*., GENESIS: A hybrid-parallel and multi-scale molecular dynamics simulator with enhanced sampling algorithms for biomolecular and cellular simulations. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **5**, 310–323 (2015).
59. C. Kobayashi *et al*., GENESIS 1.1: A hybrid-parallel molecular dynamics simulator with enhanced sampling algorithms on multiple computational platforms. *J. Comput. Chem.* **38**, 2193–2206 (2017).
60. D. Matsuoka, M. Kamiya, T. Sato, Y. Sugita, Role of the N-Terminal transmembrane helix contacts in the activation of FGFR3. *J. Comput. Chem.* **41**, 561–572 (2020).
61. E. L. Wu *et al*., CHARMM-GUI Membrane Builder toward realistic biological membrane simulations. *J. Comput. Chem.* **35**, 1997–2004 (2014).
62. 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).
63. B. R. Brooks *et al*., CHARMM: The biomolecular simulation program. *J. Comput. Chem.* **30**, 1545–1614 (2009).
64. U. Baul, D. Chakraborty, M. L. Mugnai, J. E. Straub, D. Thirumalai, Sequence effects on size, shape, and structural heterogeneity in intrinsically disordered proteins. *J. Phys. Chem. B* **123**, 3462–3474 (2019).
65. L. V. Nedialkova, M. A. Amat, I. G. Kevrekidis, G. Hummer, Diffusion maps, clustering and fuzzy Markov modeling in peptide folding transitions. *J. Chem. Phys.* **141**, 114102 (2014).