ACCOUNTS
of chemical research

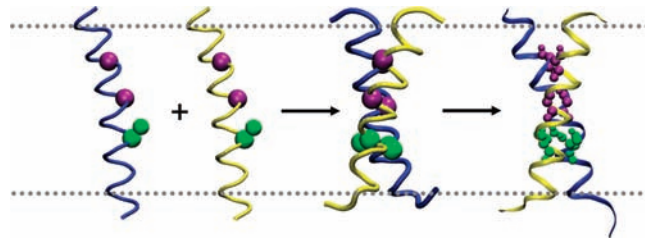
Molecular Dynamics Simulations of the Dimerization of Transmembrane α-Helices

EMI PSACHOULIA, DAVID P. MARSHALL, AND
MARK S. P. SANSOM*

Department of Biochemistry and Oxford Centre for Integrative Systems Biology, University of Oxford, South Parks Road, Oxford OX1 3QU, U.K.

RECEIVED ON JULY 20, 2009

CONSPECTUS



Membrane proteins account for nearly a quarter of all genes, but their structure and function remain incompletely understood. Most membrane proteins have transmembrane (TM) domains made up of bundles of hydrophobic α-helices. The lateral association of TM helices within the lipid bilayer is a key stage in the folding of membrane proteins. It may also play a role in signaling across cell membranes. Dimerization of TM helices is a simple example of such lateral association.

Molecular dynamics (MD) simulations have been used for over a decade to study membrane proteins in a lipid bilayer environment. However, direct atomistic (AT) MD simulation of self-assembly of a TM helix bundle remains challenging. AT-MD may be complemented by coarse-grained (CG) simulations, in which small numbers of atoms are grouped together into particles. In this Account, we demonstrate how CG-MD may be used to simulate formation of dimers of TM helices. We also show how a serial combination of CG and AT simulation provides a multiscale approach for generating and refining models of TM helix dimers.

The glycophorin A (GpA) TM helix dimer represents a paradigm for helix−helix packing, mediated by a GxxxG sequence motif. It is well characterized experimentally and so is a good test case for evaluating computational methods. CG-MD simulations in which two separate TM helices are inserted in a lipid bilayer result in spontaneous formation of a right-handed GpA dimer, in agreement with NMR structures. CG-MD models were evaluated via comparison with data on destabilizing mutants of GpA. Such mutants increased the conformational flexibility and the dissociation constants of helix dimers. GpA dimers have been used to evaluate a multiscale approach: A CG model is converted to an AT model, which is used as the basis of an AT-MD simulation. Comparison of three AT-MD simulations of GpA, one starting from a CG model and two starting from NMR structures, leads to convergence to a common refined structure for the dimer.

CG-MD self-assembly has also been used to model dimerization of the TM domain of the syndecan-2 receptor protein. This TM helix contains a GxxxG motif, which mediates right-handed helix packing comparable to that of the GxxxG motif in GpA. The multiscale approach has been applied to a more complex system, the heterodimeric αIIb/β3 integrin TM helix dimer. In CG-MD, both right-handed and left-handed structures were formed. Subsequent AT-MD simulations showed that the right-handed structure was more stable, yielding a dimer in which the GxxxG motif of the αIIb TM helix packed against a hydrophobic surface of the β3 helix in a manner comparable to that observed in two recent NMR studies.

This work demonstrates that the multiscale simulation approach can be used to model simple membrane proteins. The method may be applied to more complex proteins, such as the influenza M2 channel protein. Future refinements, such as extending the multiscale approach to a wider range of scales (from CG through QM/MM simulations, for example), will expand the range of applications and the accuracy of the resultant models.

Published on the Web 12/17/2009 www.pubs.acs.org/acr
10.1021/ar900211k

1. Introduction

Membrane proteins play many roles in the biology of cells, including transport, signaling and cell−cell interactions. Reflecting this, ca. 25% of genes encode membrane proteins.[1] The majority of membrane proteins have folds based upon bundles of transmembrane (TM) α-helices. Membrane protein folding may be described in terms of two stages: translocon-mediated insertion, followed by lateral association of TM α-helices.[2,3]

The lateral association (packing) of helices within a membrane has been studied using a number of biochemical, biophysical, and computational approaches.[3−5] Helix dimerization provides a simple model of lateral association in folding. It also is important in mechanisms of signaling across membranes by membrane-bound receptors, which have ectodomains that bind extracellular ligands and are often linked to intracellular signaling domains via single TM helices.

Several computational methods enable modeling and simulation of membrane proteins.[6−8] Such methods may provide mechanistic insights or predictive capability concerning membrane proteins and the packing of their TM helices. In particular, molecular dynamics (MD) simulations provide a useful approach for modeling membrane proteins. MD simulations of membrane proteins may be conveniently divided into those that employ a full atomistic (AT) representation and those that employ a simplified, coarse-grained (CG) model of the protein and of the bilayer lipids. AT-MD simulations are accurate but computationally demanding, whereas the more approximate nature of CG-MD simulations allow longer simulation times (>1 μs) and larger systems (>10^5 atoms) to be explored with modest computational resources.

In this Account, we review some of our recent studies of CG-MD and related methods to model TM helix dimerization. In particular, we describe studies of the TM helix dimer of glycophorin A (GpA). We also provide a brief account of recent extensions of these studies to dimerization of related TM helices involved in receptor-mediated cell signaling.

2. Computational Approaches

2.1. Coarse-Grained and Atomistic Models for Transmembrane α-Helices. A limitation of AT-MD simulations is that extended (>1 μs) simulations of membrane systems require substantial computer resources, thus limiting their applications. Coarse-grained (CG) simulations address the timescale and system size issues via simplification of the representation of the component molecules of biomolecular systems.[9−13] For example, in one CG approach (using the MARTINI coarse-grained forcefield[14,15]) that has been applied to simple membranes,[10] atoms are grouped together to form particles, each particle corresponding to approximately four non-H atoms. For example, a dipalmitoyl phosphatidylcholine (DPPC) phospholipid molecule (which contains 50 atoms) can be modeled by eight hydrophobic particles (four each for the two C_{16} tails), two mixed polar/apolar particles (for the glycerol group), a negatively charged particle (representing the phosphate group), and a positively charged particle (for the choline group; Figure 1A). Such CG models yield 2−3 orders of magnitude increased speed of simulations. Furthermore the CG model yields overall peptide/bilayer system configurations consistent with, for example, solid-state NMR data[16] and may be used to assemble more complex protein/bilayer systems.[17]

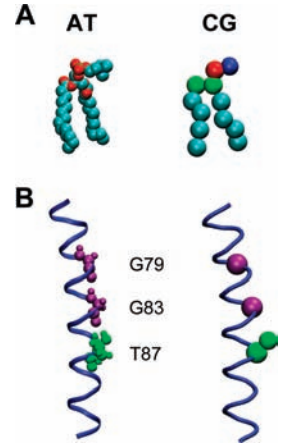


FIGURE 1. Atomistic (AT) and coarse-grained (CG) representations of (A) a phospholipid lipid (DPPC) and (B) a transmembrane (TM) α-helix. In the CG model of DPPC, the particle types are color coded as type C in cyan, type N in green, and type Q negatively charged and positively charged in red and blue, respectively. In panel B, the TM α-helix of glycophorin A (GpA) is shown in AT and CG format for key residues (Gly79, Gly83, and Thr87) of the dimerization interface.

An example of the relationship between an AT model and a CG model of a TM helix is shown for GpA (Figures 1B and 2). GpA is a simple membrane protein from red blood cells, which has been the subject of numerous biophysical and biochemical studies.[2] The structure of the GpA dimer from a solution/micelle NMR study (PDB 1AFO)[18] provided an initial structure of a GpA TM helix *monomer* for use in subsequent simulations. The AT structure was converted to a corresponding CG model[16,19,20] using a modified version of Marrink's original CG forcefield.[10] A CG peptide model generated from the corresponding AT structure consists of a chain of backbone particles (one per residue) with attached side chain particles.[10] Four CG particle types (all with an effective diameter of 4.7 Å) are distinguished: "polar" (P); "mixed polar/apolar" (N); "hydrophobic apolar" (C); and "charged" (Q), along with fur-

GpA (73-95)
ITLIIFGVMAGVIGTILLISYGI
Syn2 (145-169)
VLAAVIAGGIVGFLFAIFLILLLVY
Integrins
αIIb (967-997)
WWVLVGVLGGLLLLTILVLAMWKVGFFKRNR
β3 (694-727)
LVVLLSVMGAILLIGLAALLIWKLLITIHDRKEF

FIGURE 2. Comparison of the sequences of the TM domains of glycophorin A (GpA), syndecans 2 (Syn2), and integrins (αIIb and β3). In each case, the GxxxG or equivalent dimer interface motif is highlighted in red.

ther subtypes for the N and Q particles that allow fine-tuning of Lennard-Jones interactions to reflect hydrogen-bonding capacities. Nonbonded interactions between these particles are described by a Lennard-Jones (LJ) potential with five levels of interaction. Hydrogen-bonding subtypes for N and Q particles modulate this LJ interaction with other particles. Charged (Q) particles also interact via a screened Coulombic potential. Other than for glycine, either one or two particles are used to model each amino acid side chain. The geometry of individual amino acids and of peptide links between adjacent amino acids are modeled via appropriate distance and angle restraints.[20] In addition, the secondary structure of each TM α-helix is modeled via restraining the distances between backbone particles of residues i and $i + 4$ to mimic H-bonds in the corresponding AT structure.

For multiscale simulations,[21] a selected structure from a CG-MD simulation was converted back to an AT model using an in-house protocol.[22] To convert a CG lipid to an AT model, energy-minimized atomistic lipid fragments were aligned to the CG particles. The AT lipids were then energy-minimized. To convert a peptide, AT side chains were aligned with and substituted for the CG particles used to model the side chain. Pulchra[23] was used to grow the protein backbone from the Cα trace. The resultant AT peptide structure was energy-minimized.

2.2. Coarse-Grained MD Simulations. CG-MD simulations were performed using GROMACS 3.0 (www.gromacs.org).[24] Lennard-Jones interactions were shifted to zero between 9 and 12 Å, and electrostatic interactions were shifted to zero between 0 and 12 Å, with a relative dielectric constant of 20. All simulations were performed at constant temperature, pressure, and number of particles. The temperature of the protein, lipid, and solvent were each coupled separately using the Berendsen algorithm[25] at 323 K with $\tau_T = 1$ ps. The system pressure was semi-isotropically coupled in the x/y and z directions using the Berendsen algorithm at 1 bar with $\tau_P = 1$ ps and

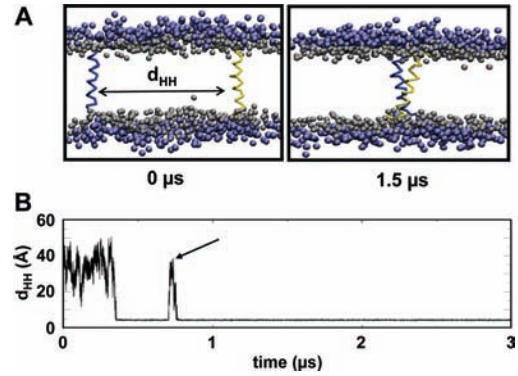


FIGURE 3. Coarse-grained GpA TM helix dimerization simulation. (A) The initial system configuration (0 μs) consists of two helices (blue and yellow) inserted in a DPPC bilayer in a parallel orientation with an interhelix separation of $d_{HH} \approx 55$ Å. The choline and phosphate (blue) and glycerol (gray) backbone particles of the DPPC molecules are shown. The snapshot at 1.5 μs illustrates the stable TM helix dimer. (B) Interhelix distance (d_{HH}) as a function of time for a GpA dimerization simulation. The arrow indicates a transient (for <0.1 μs) dissociation event of the stable helix dimer.

a compressibility of 5×10^{-6} bar^{-1}. The time step for integration was 20 fs (for GpA and integrins) or 40 fs (for Syn2). Coordinates were saved for subsequent analysis every 200 or 400 ps, respectively. VMD[26] was used for visualization.

For all helix dimerization simulations, two α-helices were inserted into a preformed dipalmitoyl phosphatidylcholine (DPPC) bilayer (containing 245 lipids) such that they were separated by an interhelix distance $d_{HH} \approx 55$ Å (Figure 3). Each system was solvated with ~2200 CG water particles and appropriate counterions. The energy of the system was minimized and followed by a 3 μs MD simulation.

2.3. Atomistic Simulations. GROMACS was also used for all AT simulations, with the GROMOS96 united atom force field[27] and the SPC water model.[28] Long-range electrostatics were calculated using the particle mesh Ewald method[29] with a real-space cutoff of 10 Å. For the van der Waals interactions, a cutoff of 10 Å was used. The simulations were performed at a temperature of 323 K using a Berendsen thermostat with $\tau_T = 0.1$ ps. A constant pressure of 1 bar was maintained with an isotropic coupling constant $\tau_P = 1.0$ ps and compressibility $= 4.5 \times 10^{-5}$ bar^{-1}. The integration time step was 2 fs. The LINCS method[30] was used to constrain bond lengths. Coordinates were saved every 5 ps for analysis.

3. Results

3.1. Coarse-Grained Simulations of Glycophorin Helix Dimerization. To model the dimerization of GpA, two CG helices were inserted in a parallel orientation relative to one another in a preformed DPPC bilayer (Figure 3). Seven simu-

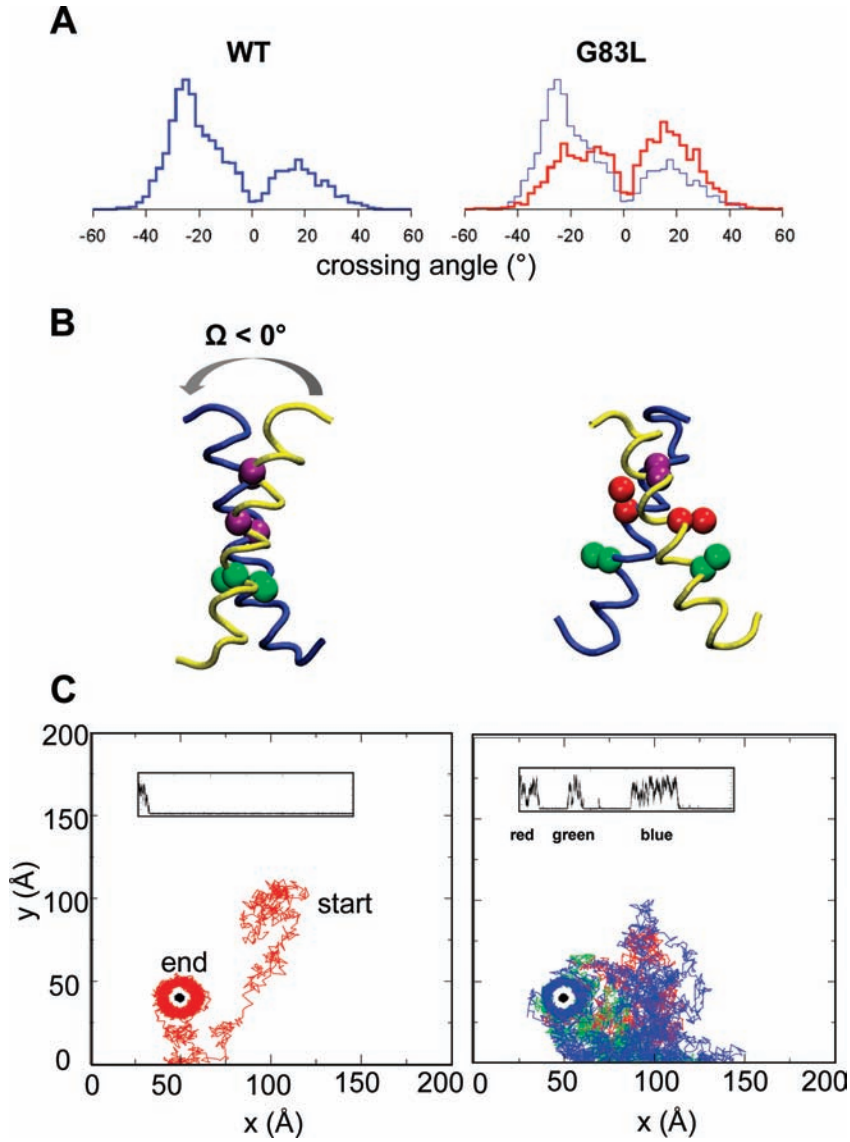


FIGURE 4. Glycophorin A TM helix dimerization: wild type and mutant (G83L) compared. (A) Helix crossing angle distributions for the WT (blue) and the G83L (red) mutant. Helix crossing angles (Ω) were then evaluated from merged dimer trajectories corresponding in each case to an ensemble of 7 × 3 μs simulations. Note that a positive crossing angle corresponds to left-handed helix packing and a negative crossing angle to right-handed packing. (B) Representative structures of the WT and G83L mutant GpA TM helix dimers. Purple spheres correspond to glycine, green to threonine, and red to the G83L leucine. (C) Example motions of one GpA helix relative to the other in the reference frame of the lipid bilayer. The positions of the centers of mass of a WT simulation and a G83L simulation are shown. The positions of the fixed helices are drawn in black. The coordinates (*x*,*y*) of the mobile helix are drawn in red for the wild-type and in red, green, and blue for the G83L mutant to distinguish the dissociation events. The inset figures show the corresponding interhelix vs time plots (see Figure 3B for details).

lations of duration 3 μs were run both for the GpA wild-type (WT) helices and for several mutant helices.[31] Here we will focus on a comparison of the GpA-WT with a mutant, G83L, which has been shown to *destabilize* the GpA TM helix dimer. During each simulation, the two helices diffused randomly in the bilayer, encountered one another, and formed a dimer (Figure 3B). If the initial separation was increased from 55 to ~110 Å, a similar dimerization was seen, but the initial delay was increased from <0.5 to >1 μs.

The GpA-WT helices formed a long-lasting dimer, examination of which revealed that the helices were packed in a *right-handed* fashion (i.e., with a negative value of the helix-crossing angle, Ω; Figure 4). The Ω distribution across all GpA-WT simulations (total simulation time 21 μs) is bimodal, with the major mode (72%) having a crossing angle of $\Omega = -25°$.[31] This Ω distribution compares well with the right-handed packing of the GpA helices observed in both the solution/micelle NMR ($\Omega = -43°$)[18] and the solid-state/bilayer

NMR ($\Omega \approx -35°$)[32] structures. Analysis of the helix/helix contacts in the right-handed GpA model from CG-MD showed that the Gly79, Gly83, and Thr87 residues formed the helix/helix interface (Figure 3B), in good agreement with NMR[18,32] and mutagenesis data.[33−36] A number of test cases (e.g., a simple model $KL_{19}K$ α-helix and the ζζ TM helix dimer[37]) have shown that in the absence of a GxxxG motif, right-handed packing is not observed.

The diffusion of the two GpA helices within a lipid bilayer may be examined by plotting the position of the center of mass of one helix relative to the other. Such analysis reveals a two-dimensional random walk followed by helix−helix encounter and subsequent stable dimer formation (Figure 4C). We have also explored simulations with four GpA helices. In these simulations, right-handed helix dimers formed, in addition to some trimers (also right-handed) and loose tetrameric aggregates.

Comparable helix/helix association CG-MD simulations have been performed for both nondisruptive and disruptive mutations. For the former (e.g., I85F), the resultant dimers were indistinguishable in terms of their helix packing and stability from the WT GpA dimer. However, with the disruptive mutants, both the dimer structure and stability were perturbed. Thus for the GpA-G83L mutant, the helices formed a dimer but subsequently dissociated once or twice in four out of the seven simulations (each of duration 3 μs).[31] This can be seen via analysis of the position of the center of mass of one G83L helix relative to the other (Figure 4C). In general, for a number of disruptive mutant helices, the initial two-dimensional random walk and helix−helix encounter is followed by a dynamic process of dimer dissociation and reassociation, on a $\sim$1 μs time scale.

Comparison of the Ω distribution for the disruptive G83L mutant (while in a dimer) reveals loss of the preferred right-handed helix packing. Indeed, the distribution becomes markedly bimodal suggesting that the disruptive mutant "softens" the interhelix interface. The tight dimerization interface observed in the GpA-WT is lost upon mutation of one of the key glycine residues that provide stabilizing direct interactions as well as allow the close approach of other nearby residue side chains within the interface such as Thr87. In general, simulations of three disruptive mutants revealed dimers that were less stable than those for the WT and two nondisruptive mutants. From the dimer−monomer equilibria observed in the simulations, values of $\Delta\Delta G_{\text{DIMERIZATION}}$ were estimated, which were in agreement with experimental data.[38] Direct calculation of $\Delta G_{\text{DIMERIZATION}}$ of the WT GpA was not undertaken, for which umbrella sampling or related approaches would be required.[39]

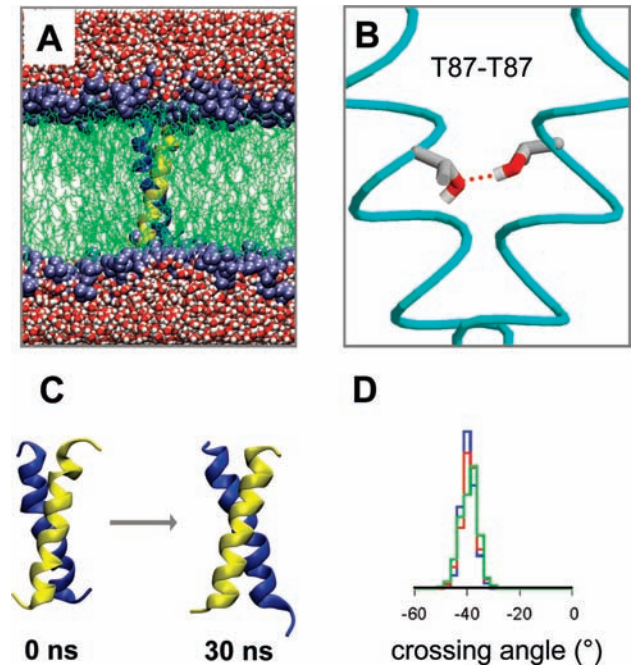


FIGURE 5. Atomistic simulations of the GpA dimer. (A) Simulation snapshot from an AT-MD simulation starting from a GpA dimer model converted from CG. The choline and phosphate groups of the bilayer are in blue, the remainder of the lipid molecules in green, and the water molecules in red/white. (B) Converged structure from an AT-MD simulation showing the H-bond between the Thr87 side chains of opposite GpA monomers. (C) Comparison of the GpA dimer at the start (0 ns) and end (30 ns) of an AT-MD simulation initiated from the CG-MD generated model of the dimer. (D) Comparison of the helix crossing angle distributions from three AT-MD simulations, starting from the micelle/NMR (blue), bilayer/NMR (red), and CG-MD model (green) structures.

3.2. A Multiscale Approach: From CG to AT. Overall, our studies of GpA suggest that CG-MD is sufficiently sensitive to detect disruptions of the helix/helix interface introduced by mutation and for GpA-WT to yield a structural model comparable to those derived from NMR data. This encourages us to refine CG-MD-generated models via atomistic simulations, allowing, for example, local distortions of the TM α-helices, thus providing a multiscale approach to predict the packing of TM helices.

To this end, a representative (i.e., right-handed, with a crossing angle close to the mode of the distribution) GpA dimer from the ensemble of CG simulations was converted to an AT model and used as the starting point for a 30 ns AT-MD simulation (Figure 5). Visualization and comparison of the structures at the start and end of the AT-MD simulation indicated a shift of crossing angle from $\Omega \approx -25°$ to $\Omega \approx -40°$, the latter value giving an improved agreement with the NMR

structures. However, the helix−helix interface of the AT dimer was broadly the same as that in the CG model, namely, corresponding to the LIxxGVxxGVxxT motif. In contrast, AT-MD simulations starting from a left-handed CG model for GpA resulted in a much looser interface that changed substantially over the course of the 30 ns simulation and did not show interactions between the GxxxG motifs of adjacent helices.

The AT-MD simulation starting from the CG model may be compared with two previously published AT simulations of a GpA-WT dimer, with initial dimers from the solution/micelle NMR structure and the solid-state/bilayer NMR structure.[40] All three AT-MD simulations converge to the same model, with a modal crossing angle of $\Omega \approx -40°$. Moreover, in all three simulations, the dimer is stabilized by packing interactions between the glycines of the GxxxG motif and also by interhelix side chain H-bonds of the Thr87 residues across the interface. Thus, in all three atomistic simulations the GpA-WT dimer converged to a common structure. This suggests that a multiscale approach to helix dimerization may be able to provide refined model structures of accuracy comparable to that of NMR experimental studies.

3.3. TM Helices and Transmembrane Signaling. As discussed above, GpA TM helices contain a motif (LIxxGVxxGVxxT) that plays a key role in homodimerization.[4,36] Such GxxxG motifs may play a more general role in TM helix dimerization.[41] Related motifs may also be found at helix−helix interfaces, in which one or both glycines of GxxxG are substituted by alanine or serine,[36] again leading to a right-handed crossing angle for the dimer.[2,36] It is therefore of some interest to use CG-MD simulations to explore the role of such motifs in, for example, TM helix dimerization for receptors that signal across the bilayer and for which helix association is thought to play a key functional role. Two examples will be described here: the syndecans and the integrins (Figure 2).

3.4. The Syndecans: A GxxxG Dimerization Motif in a Signaling Protein. The syndecan family of receptors participate in a number of cellular processes, for example, in aiding integrins in focal adhesion development and in wound healing. All syndecans have an N-terminal extracellular domain (containing heparan sulfate and chondroitin sulfate binding sites), a single TM helix (∼25 residues), and a short C-terminal cytoplasmic domain.[42] The cytoplasmic domain contains interaction sites for a number of intracellular proteins. Syndecan-2 (Syn2) is the predominant syndecan expressed during embryonic development and is involved in regulating TGF-β signaling.[42] Syn2 may form homodimers.[43] Its TM helix contains a conserved dimerization motif similar to that of GpA, containing GxxxG but lacking a threonine side chain (Figure 2). As yet, there are no structural data for the Syn2 TM domain. Thus, it is of interest to apply the CG-MD methodology evaluated using GpA to this biomedically relevant protein.

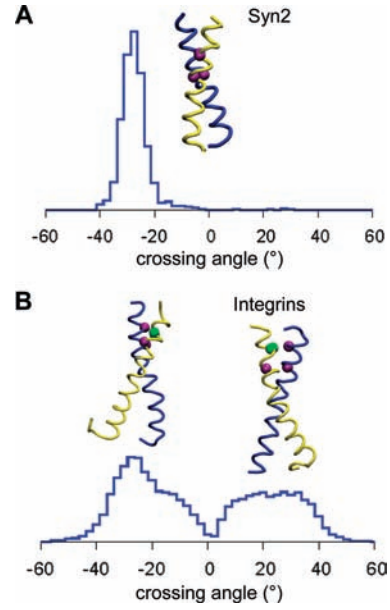


FIGURE 6. Helix crossing angles and selected TM helix dimer structures from CG-MD simulations, with the glycines in purple: (A) Syn2; (B) integrin αIIb/β3, showing LH and RH helix dimers. The glycines of αIIb (blue Cα trace) are in purple and the serine and alanine (of the SxxxA motif) of β3 (yellow Cα trace) are in green and purple, respectively.

The initial system for the Syn2 simulations was similar to that for GpA, that is, with two Syn2 TM helices inserted 55 Å apart, and five 3 μs simulations were performed. As for GpA-WT, Syn2 helices dimerized within the first ∼0.7 μs, subsequently remaining as a dimer for most of each simulation. In general, Syn2 behaved similarly to GpA, with the main interactions being between the glycines of the GxxxG motif (Figure 6A). However, the crossing angle distribution of Syn2 is a unimodal distribution with a modal $\Omega = -30°$, that is, a right-handed packing orientation. Thus it would seem that homodimerization mediated by a simple GxxxG motif yields a helix packing mode the same as that of GpA, even though Syn2 lacks the threonine side chain present in GpA.

3.5. The Integrins: The αIIb/β3 Helix Dimer. The multiscale approach employed to model the GpA dimer may also be used to explore a more complex example such as the integrin heterodimer formed by the α-helical TM domains of integrins. The integrins are a major class of cell adhesion receptors involved in many cellular processes, for example, cell migration, adhesion, differentiation, and proliferation.[44] Integrins are heterodimeric type I transmembrane (TM) proteins of nonco-

valently associated α and β subunits. Each α and β subunit consists of a large (∼700−1000 residues) N-terminal extracellular domain, a single TM helix, and a short (∼20−50 residues) C-terminal cytoplasmic tail. The TM helices of both α and β subunits consist of ∼23 mostly hydrophobic residues and are terminated by either a tryptophan/lysine pair or a single lysine residue at the intracellular (C-terminal) side.[45]

In mammals, 19 α and 9 β subunits have been identified, and different combinations of these α and β subunits produce 24 different heterodimers that each bind to a specific ligand, cell surface, and extracellular matrix. As an example, we will explore the αIIb/β3 dimer. The TM domain of αIIb contains a GxxxG motif, while β3 contains a GxxxG-like motif, namely, SxxxA (Figure 2). These two motifs are expected to be important for dimerization, alongside other residues close to the membrane/cytoplasmic interface. Thus, simulations were performed with longer TM helices for αIIb and β3 (Figure 2), in agreement with recent NMR studies.[46,47]

As for GpA and Syn2, CG models of the two TM helices (αIIb and β3) were inserted in a DPPC bilayer with the helices separated by a distance of ∼55 Å, and 10 3 μs simulations were performed. In 9 out of 10 simulations, the αIIbβ3 dimer was formed within the first 0.5 μs and for the remaining simulation within 1 μs. In all simulations, once formed the dimer did not subsequently dissociate over the course of the remainder of the simulation.

The helix crossing angle distribution for the integrin dimer is bimodal (Figure 6B), with a slight bias (56% vs 44%) in favor of right-handed dimers. To further explore the packing of integrin TM helices, two models of the dimer, one left-handed ($\Omega = +28°$) and one right-handed ($\Omega = -25°$), were extracted from the CG-MD trajectory and used to initiate AT-MD simulations.

Visualization of the structures at the start and end of the atomistic αIIbβ3-LH simulation (data not shown) suggests that the left-handed dimer is relatively unstable, with the two helices not remaining closely packed during the simulation. In contrast, visualization of the structures at the start and end of the αIIbβ3-RH simulation (Figure 7) suggests that the right-handed dimer is relatively stable. This difference is confirmed by analysis of the interhelix distances of the αIIb and β3 TM helices in each simulation as a function of time (Figure 7A). Thus, for αIIbβ3-RH, the interhelix distance remains at ∼8 Å throughout the simulation, whereas for αIIbβ3-LH, the interhelix distance doubles over the course of 30 ns. Examination of crossing angle distributions suggests that a tight right-handed interface ($\Omega = -25° \pm 3°$) is maintained, while there is more flexibility of packing in the LH ($\Omega = +20°$ to $+50°$) dimer.

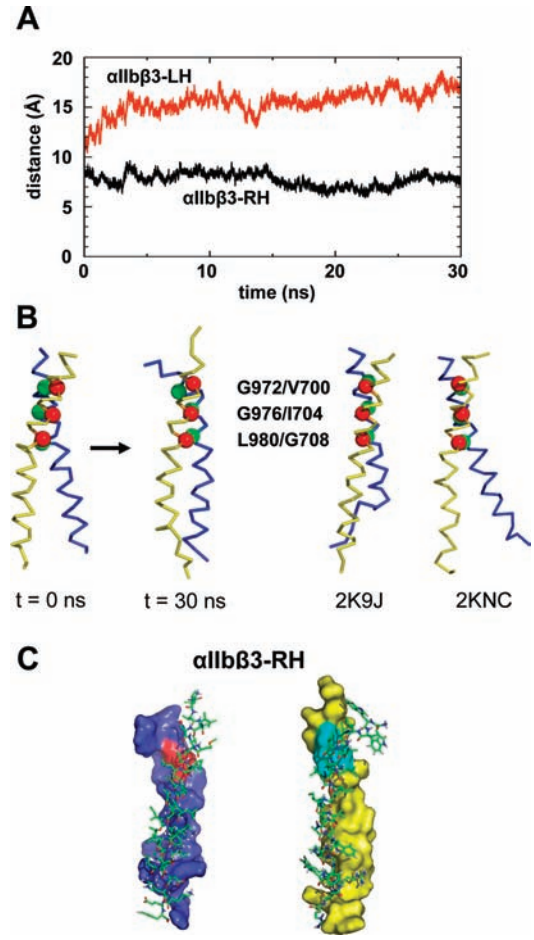


FIGURE 7. AT-MD simulations of integrin aIIb/β3 TM helix dimers, starting from models generated by CG-MD. (A) Progress of the two simulations (black = RH; red = LH), monitored as the interhelix separation distance as a function of simulation time. (B) Snapshots at the start (t = 0 ns) and end (t = 30 ns) of the simulation of the RH dimer model, compared with NMR structures (2K9J[46] and 2KNC[47]). The αIIb chain is in blue and the β3 chain in yellow. The Cα atoms of the contact residues (Gly972, Gly976, and Leu980 of αIIb in green; Val700, Ile704, and Gly708 of β3 in red) are shown as spheres. (C) Two views of the 30 ns AT-MD structure of the aIIb/β3 RH helix dimer. On the left, the surface of the αIIb helix is shown in blue, with Gly972, Gly975, and Gly976 in red. On the right, the surface of the β3 helix is in yellow with residues Val700, Ala703, and Ile704 in cyan.

Analysis of the interacting residues (over the last 5 ns) between the two helices in each of the dimers showed that in αIIbβ3-RH there were interactions along the whole length of both helices (i.e., αIIb and β3). Two main regions of interactions were (a) the G^{972}xxxG^{976} motif and nearby residues of αIIb interacting with the hydrophobic patch on β3 close to its GxxxG-like motif, namely, residues Val700, Ala703, and Ile704, and (b) the cytoplasmic, membrane-proximal residues

987−997 of αIIb with the cytoplasmic, membrane proximal residues 715−727 of β3. The interactions of the residues located in the cytoplasmic region appear to modulate the interactions of residues located in the GxxxG region. This is of particular interest because earlier combined experimental and computational studies of the αIIbβ3 TM helix dimer[48] had already implicated a similar interaction surface between the G^{972}xxxG^{976} region of the helices, but not of the more polar residues toward the C-termini. Similarly, two recent NMR structures of a αIIb/β3 TM helix dimer (one in DHPC/POPC bicelles[46] and one in a membrane-mimetic solvent system[47]) both revealed a right-handed packing ($\Omega = -26°$ and $-30°$, respectively, compared with $\Omega = -25°$ for the simulation) of the N-terminal segments of the two helices, with close approach of G^{972}xxxG^{976} of the αIIb subunit to V^{700}xxxI^{704} of β3. In the NMR structures, the C-termini of the two helices are further apart than in the AT-MD model structure. It will therefore be of interest to further explore the details and possible functional significance of the predicted interactions closer to the C-termini of the helices, especially as mutagenesis studies have implicated the more C-terminal region of β3 in helix/helix interactions[45] and in the context of proposed changes in helix packing modes upon integrin activation.[49]

4. Future Directions

These studies indicate a future direction for MD simulations of lateral association and oligomerization of TM helices. In particular, the methods outlined in this Account may be applied to a wider range of TM helices and association motifs. One example of interest is provided by ErbB receptor homo- and heterodimers.[50] The ErbB TM helices interact via complex GxxxG and related sequence motifs and may exhibit multiple modes of interaction reflecting the activation status of a receptor. Another direction will be to extend studies to larger assemblies of TM helices. For example, preliminary studies have shown that CG-MD simulations may also be used to self-assemble the tetrameric TM helix bundle of the influenza A M2 protein channel.[51]

From a methodological perspective, there are two immediate challenges. The first is to extend the multiscale approach to embrace a wider range of scales, from CG through, for example, QM/MM simulations, thus enabling more detailed analysis of the role of H-bonds in stabilizing TM helix packing.[52,53] It will also be important to develop high-throughput (HT) methods of multiscale MD simulation setup and analysis in order to enable a wider range of biological systems to be explored.

This work was supported by grants from the BBSRC and the Wellcome Trust. We thank our colleagues for their interest in this work, especially Iain Campbell, Kia Balali-Mood, Beatrice Nikolaidi, and Philip Fowler.

BIOGRAPHICAL INFORMATION

Emi Psachoulia joined the University of Oxford as a Wellcome Trust funded graduate student in 2004 and was awarded a D.Phil. in computational biochemistry in 2008. She is interested in the interactions of membrane-associated proteins involved in cellular signaling using multiscale MD simulations.

David Marshall completed his M.Bioch. from the University of Oxford in 2009. He is currently a graduate student in the Department of Biochemistry funded by the BBSRC.

Mark Sansom has a D.Phil. in Molecular Biophysics from the University of Oxford (1983). He is currently Director of the Structural Bioinformatics and Computational Biochemistry Unit in the Department of Biochemistry at the University of Oxford. His research interests are focused on computer modeling and simulation to understand the structure and function of membrane proteins.

FOOTNOTES

*To whom correspondence should be addressed. E-mail: mark.sansom@bioch.ox.ac.uk. Phone: +44-1865-613306. Fax: +44-1865-613238.

REFERENCES

1 Nilsson, J.; Persson, B.; von Heijne, G. Comparative analysis of amino acid distributions in integral membrane proteins from 107 genomes. *Proteins: Struct., Funct., Bioinf.* **2005**, *60*, 606–616.

2 Popot, J. L.; Engelman, D. M. Helical membrane protein folding, stability, and evolution. *Annu. Rev. Biochem.* **2000**, *69*, 881–922.

3 Bowie, J. U. Solving the membrane protein folding problem. *Nature* **2005**, *438*, 581–589.

4 Senes, A.; Engel, D. E.; DeGrado, W. F. Folding of helical membrane proteins: The role of polar, GxxxG-like and proline motifs. *Curr. Opin. Struct. Biol.* **2004**, *14*, 465–479.

5 Langosch, D.; Arkin, I. T. Interaction and conformational dynamics of membrane-spanning protein helices. *Protein Sci.* **2009**, *18*, 1343–1358.

6 Ash, W. L.; Zlomislic, M. R.; Oloo, E. O.; Tieleman, D. P. Computer simulations of membrane proteins. *Biochim. Biophys. Acta* **2004**, *1666*, 158–189.

7 Lindahl, E.; Sansom, M. S. P. Membrane proteins: Molecular dynamics simulations. *Curr. Opin. Struct. Biol.* **2008**, *18*, 425–431.

8 Khalili-Araghi, F.; Gumbart, J.; Wen, P. C.; Sotomayor, M.; Tajkhorshid, E.; Schulten, K. Molecular dynamics simulations of membrane channels and transporters. *Curr. Opin. Struct. Biol.* **2009**, *19*, 128–137.

9 Shelley, J. C.; Shelley, M. Y.; Reeder, R. C.; Bandyopadhyay, S.; Klein, M. L. A coarse grain model for phospholipid simulations. *J. Phys. Chem. B* **2001**, *105*, 4464–4470.

10 Marrink, S. J.; de Vries, A. H.; Mark, A. E. Coarse grained model for semiquantitative lipid simulations. *J. Phys. Chem. B* **2004**, *108*, 750–760.

11 Murtola, T.; Falck, E.; Patra, M.; Karttunen, M.; Vattulainen, I. Coarse-grained model for phospholipid/cholesterol bilayer. *J. Chem. Phys.* **2004**, *121*, 9156–9165.

12 Nielsen, S. O.; Lopez, C. F.; Srinivas, G.; Klein, M. L. Coarse grain models and the computer simulation of soft materials. *J. Phys.: Condens. Matter* **2004**, *16*, R481–R512.

13 Stevens, M. J. Coarse-grained simulations of lipid bilayers. *J. Chem. Phys.* **2004**, *121*, 11942–11948.

14 Marrink, S. J.; Risselada, J.; Yefimov, S.; Tieleman, D. P.; de Vries, A. H. The MARTINI forcefield: Coarse grained model for biomolecular simulations. *J. Phys. Chem. B* **2007**, *111*, 7812–7824.

15 Monticelli, L.; Kandasamy, S. K.; Periole, X.; Larson, R. G.; Tieleman, D. P.; Marrink, S. J. The MARTINI coarse grained force field: Extension to proteins. *J. Chem. Theory Comput.* **2008**, *4*, 819–834.

16 Bond, P. J.; Holyoake, J.; Ivetac, A.; Khalid, S.; Sansom, M. S. P. Coarse-grained molecular dynamics simulations of membrane proteins and peptides. *J. Struct. Biol.* **2007**, *157*, 593–605.

17 Scott, K. A.; Bond, P. J.; Ivetac, A.; Chetwynd, A. P.; Khalid, S.; Sansom, M. S. P. Coarse-grained MD simulations of membrane protein-bilayer self-assembly. *Structure* **2008**, *16*, 621–630.

18 MacKenzie, K. R.; Prestegard, J. H.; Engelman, D. M. A transmembrane helix dimer: Structure and implications. *Science* **1997**, *276*, 131–133.

19 Bond, P. J.; Sansom, M. S. P. Insertion and assembly of membrane proteins via simulation. *J. Am. Chem. Soc.* **2006**, *128*, 2697–2704.

20 Bond, P. J.; Wee, C. L.; Sansom, M. S. P. Coarse-grained molecular dynamics simulations of the energetics of helix insertion into a lipid bilayer. *Biochemistry* **2008**, *47*, 11321–11331.

21 Ayton, G. A.; Noid, W. G.; Voth, G. A. Multiscale modeling of biomolecular systems: In serial and in parallel. *Curr. Opin. Struct. Biol.* **2007**, *17*, 192–198.

22 Stansfeld, P. J.; Hopkinson, R.; Ashcroft, F. M.; Sansom, M. S. P. PIP_2 binding site in Kir channels: Definition by multiscale biomolecular simulations. *Biochemistry* **2009**, *48*, 10926–10933.

23 Rotkiewicz, P.; Skolnick, J. Fast procedure for reconstruction of full-atom protein models from reduced representations. *J. Comput. Chem.* **2008**, *29*, 1460–1465.

24 Lindahl, E.; Hess, B.; van der Spoel, D. GROMACS 3.0: A package for molecular simulation and trajectory analysis. *J. Mol. Model.* **2001**, *7*, 306–317.

25 Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **1984**, *81*, 3684–3690.

26 Humphrey, W.; Dalke, A.; Schulten, K. VMD - Visual Molecular Dynamics. *J. Mol. Graphics* **1996**, *14*, 33–38.

27 van Gunsteren, W. F.; Kruger, P.; Billeter, S. R.; Mark, A. E.; Eising, A. A.; Scott, W. R. P.; Huneberger, P. H.; Tironi, I. G. *Biomolecular Simulation: The GROMOS96 Manual and User Guide*; Biomos & Hochschulverlag AG an der ETH Zurich: Groningen and Zurich, 1996.

28 Hermans, J.; Berendsen, H. J. C.; van Gunsteren, W. F.; Postma, J. P. M. A consistent empirical potential for water-protein interactions. *Biopolymers* **1984**, *23*, 1513–1518.

29 Darden, T.; York, D.; Pedersen, L. Particle mesh Ewald - an N.log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **1993**, *98*, 10089–10092.

30 Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. LINCS: A linear constraint solver for molecular simulations. *J. Comput. Chem.* **1997**, *18*, 1463–1472.

31 Psachoulia, E.; Bond, P. J.; Fowler, P. W.; Sansom, M. S. P. Helix−helix interactions in membrane proteins: Coarse-grained simulations of glycophorin A helix dimerization. *Biochemistry* **2008**, *47*, 10503–10512.

32 Smith, S. O.; Song, D.; Shekar, S.; Groesbeek, M.; Ziliox, M.; Aimoto, S. Structure of the transmembrane dimer interface of glycophorin A in membrane bilayers. *Biochemistry* **2001**, *40*, 6553–6558.

33 Lemmon, M. A.; Flanagan, J. M.; Treutlein, H. R.; Zhang, J.; Engelman, D. M. Sequence specificity in the dimerisation of transmembrane α-helices. *Biochemistry* **1992**, *31*, 12719–12725.

34 Treutlein, H. R.; Lemmon, M. A.; Engelman, D. M.; Brunger, A. T. The glycophorin A transmembrane domain dimer: Sequence-specific propensity for a right-handed supercoil of helices. *Biochemistry* **1992**, *31*, 12726–12733.

35 Lemmon, M. A.; Treutlein, H. R.; Adams, P. D.; Brunger, A. T.; Engelman, D. M. A dimerisation motif for transmembrane α helices. *Nat. Struct. Biol.* **1994**, *1*, 157–163.

36 Russ, W. P.; Engelman, D. M. The GxxxG motif: A framework for transmembrane helix-helix association. *J. Mol. Biol.* **2000**, *296*, 911–919.

37 Call, M. E.; Schnell, J. R.; Xu, C. Q.; Lutz, R. A.; Chou, J. J.; Wucherpfennig, K. W. The structure of the ζζ transmembrane dimer reveals features essential for its assembly with the T cell receptor. *Cell* **2006**, *127*, 355–368.

38 Doura, A. K.; Fleming, K. G. Complex interactions at the helix-helix interface stabilize the glycophorin A transmembrane dimer. *J. Mol. Biol.* **2004**, *343*, 1487–1497.

39 Hénin, J.; Pohorille, A.; Chipot, C. Insights into the recognition and association of transmembrane α-helices. The free energy of α-helix dimerization in glycophorin A. *J. Am. Chem. Soc.* **2005**, *127*, 8478–8484.

40 Cuthbertson, J. M.; Bond, P. J.; Sansom, M. S. P. Transmembrane helix−helix interactions: Comparative simulations of the glycophorin A dimer. *Biochemistry* **2006**, *45*, 14298–14310.

41 Senes, A.; Gerstein, M.; Engelman, D. M. Statistical analysis of amino acid patterns in transmembrane helices: The GxxxG motif occurs frequently and in association with beta-branched residues at neighboring positions. *J. Mol. Biol.* **2000**, *296*, 921–936.

42 Tkachenko, E.; Rhodes, J. M.; Simons, M. Syndecans: New kids on the signaling block. *Circ. Res.* **2005**, *96*, 488–500.

43 Dews, I. C.; MacKenzie, K. R. Transmembrane domains of the syndecan family of growth factor coreceptors display a hierarchy of homotypic and heterotypic interactions. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 20782–20787.

44 Hynes, R. O. Integrins: Bidirectional, allosteric signaling machines. *Cell* **2002**, *110*, 673–687.

45 Partridge, A. W.; Liu, S.; Kim, S.; Bowie, J. U.; Ginsberg, M. H. Transmembrane domain helix packing stabilizes integrin αIIbβ3 in the low affinity state. *J. Biol. Chem.* **2005**, *280*, 7294–7300.

46 Lau, T. L.; Kim, C.; Ginsberg, M. H.; Ulmer, T. S. The structure of the integrin αIIbβ3 transmembrane complex explains integrin transmembrane signalling. *EMBO J.* **2009**, *28*, 1351–1361.

47 Yang, J.; Ma, Y. Q.; Page, R. C.; Misra, S.; Plow, E. F.; Qin, J. Structure of an integrin αIIbβ3 transmembrane-cytoplasmic heterocomplex provides insight into integrin activation. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 17729–17734.

48 Li, W.; Metcalf, D. G.; Gorelik, R.; Li, R. H.; Mitra, N.; Nanda, V.; Law, P. B.; Lear, J. D.; DeGrado, W. F.; Bennett, J. S. A push-pull mechanism for regulating integrin function. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 1424–1429.

49 Wegener, K. L.; Campbell, I. D. Transmembrane and cytoplasmic domains in integrin activation and protein-protein interactions (Review). *Mol. Membr. Biol.* **2008**, *25*, 376–387.

50 Landau, M.; Ben-Tal, N. Dynamic equilibrium between multiple active and inactive conformations explains regulation and oncogenic mutations in ErbB receptors. *Biochim. Biophys. Acta* **2008**, *1785*, 12–31.

51 Carpenter, T.; Bond, P. J.; Khalid, S.; Sansom, M. S. P. Self-assembly of a simple membrane protein: coarse-grained molecular dynamics simulations of the influenza M2 channel. *Biophys. J.* **2008**, *95*, 3790–3801.

52 Senes, A.; Ubarretxena-Belandia, I.; Engelman, D. M. The Cα—H···O hydrogen bond: A determinant of stability and specificity in transmembrane helix interactions. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 9056–9061.

53 Park, H.; Yoon, J.; Seok, C. Strength of C_α—H···O═C hydrogen bonds in transmembrane proteins. *J. Phys. Chem. B* **2008**, *112*, 1041–1048.