Predicting Transmembrane (TM) Domain Dimer Structures using Martini 3

Amita R. Sahoo[1], Paulo C. T. Souza[2], Zhiyuan Meng[1+] and Matthias Buck[1]*

[1]*Department of Physiology and Biophysics, Case Western Reserve University, School of Medicine, 10900 Euclid Avenue, Cleveland, Ohio 44106, U. S. A.*

[2]*Molecular Microbiology and Structural Biochemistry (MMSB, UMR 5086), CNRS & University of Lyon, 7 Passage du Vercors, 69007, Lyon, France*

+ current address: *Biophysics Graduate Program, The Ohio State University, Columbus, Ohio 43210, U. S. A.*

* To whom correspondence may be addressed: matthias.buck@case.edu

Abstract

Determination of the structure and dynamics of transmembrane (TM) domains of single-transmembrane receptors is key to understanding their mechanism of signal transduction across the plasma membrane. Although many studies have been performed on isolated soluble extra- and intracellular receptor domains in aqueous solutions, limited knowledge exists on the lipid embedded TM region. In this study, we predict the assembly of alternate configurations of receptor TM domain dimers using the Martini 3 force field for coarse-grain (CG) molecular dynamic simulations. This recent version of Martini has new bead types and sizes, which allows more accurate predictions of molecular packing and interactions compared to the previous version. Our results with Martini 3 simulations show overall good agreement with *ab initio* predictions using PREDDIMER and with available NMR derived structures for TM helix dimers. Understanding and predicting the association TM domains may help us to better understand the signalling mechanism of TM receptors, in turn providing the opportunity for development of new pharmaceuticals, some of which are peptide based.

Introduction

Membrane proteins account for 20-30 % of all proteins identified in the genomes of prokaryotes and eukaryotes.[1] Compared with the multi-pass membrane proteins, single-pass transmembrane receptors are the most abundant and functionally diverse category of membrane proteins.[2] These proteins are often highly flexible near the membrane and very difficult to characterize structurally. Signal transduction across the plasma membrane typically involves receptor dimerization. The TM domains can contribute to the stability of full-length receptor dimers and hence, help maintain configurations which are either competent for signalling or inactive. Specific sets of interhelix contacts between the TM domain, set up distinct "on"/ "off" states through change in the orientation, if not oligomerization of the TM domains. Within a biological membrane, individual TM helices usually interact to form one or only a few thermodynamically stable structures. Properties of the lipid bilayer, such as the thickness of the membrane, the nature of lipid tails and headgroups are a main contributing factor to the stability of TM domain configurations. Solution NMR has been the main tool for the determination of TM helical dimer structures using detergent micelles (typically DPC), or more realistic membrane mimics, such as bicelles, (a mixture of DMPC lipid and DHPC detergent). However, the NMR data are typically analysed in a way to determine the structure only for one particular configuration of a dimer promoted by such environment.[3]

To date the structures of several TM helix homo- or heterodimers have been solved.[4,5,14,6–13] Analyzing these structures and how they assemble by studying the dynamics of the simple TM helix dimers helps us to understand the type and mode of interactions between individual TM helices, at the level of amino acids and/or amino acid motifs.[15] In this study, we focussed on a computational characterization of the structural dynamics of the TM domains of a total of 11 receptors.[6-16] All the receptors that we study are receptor tyrosine kinases (RTK) with an exception of Bnip3, which is a bcl2-family (non-kinase receptor) and GpA, Glycophorin A, a model system for TM dimers. Although *ab initio* predictions have been made for these receptors[16,17] and further structural refinements using μs level all-atom molecular dynamics (MD) simulations have been carried out for some,[17,18] the configurational dynamics in the TM region remain less explored. All-atom simulations to study TM helix association have been employed previously but since the movement of lipids and TM helices in lipids is relatively slow, these methods demand extensive computational resources that are difficult to

access.[19] Significant computational speed-ups (typically 100-fold over all-atom MD) can be achieved by carrying out coarse grained (CG) simulations using a recently developed Martini 3 force field.[20] This new version has improved packing and interaction balance of all the non-bonded interaction terms, allowing more realistic modelling of biomolecular systems, including protein-protein interactions examined here. Using Martini 3, we gained insights into how these 11 TM peptides associate/dimerize in the CG simulations in a DMPC lipid bilayer, and also examined their configurational dynamics and validate the dimer structures with reference to existing solution NMR structures[4,5,14,6–13] and compare to *ab initio* PREDDIMER dimer predictions[16]. Further, we describe our results in the context of available experimental data.

Results and Discussion

The transmembrane (TM) regions for the 11 receptors were modeled as ideal α-helices. Because the regions immediately outside the TM segment may affect its structure and stability, we extended the N-terminus and C-terminus of the TM region by 8-9 residues and 10 residues, respectively, as extended conformation. For each protein, we modeled two identical TM peptides and placed them in parallel to the membrane normal, in the center of the DMPC bilayer with an interhelical separation of 50Å from each other. Each simulation was run in quadruplicate for 4 µs each. Therefore, for each receptor, we ran a total time of 16 µs and for all the 11 receptors, the total sampling was 176 µs. Via diffusion in the membrane, the TM peptides came closer and soon interacted with each other forming the TM dimer as shown in fig 1. To follow the association of the TM peptides, we monitored the distance between the center of mass of the helix monomers for all the 11 receptors, as shown in Fig S1. The TM peptides for all the 11 receptors dimerize in the DMPC bilayer and appeared stable throughout the simulation.

To simplify the analysis of the systems, all 4 repeat simulations for each TM dimer system were concatenated and total populations of the TM dimers are calculated considering the inter-helical angle and distance as shown in the 2D plot (Fig S2). Each 2D plot has one global minimum (highly dense population) and 2/3 local minima. Therefore, we selected the 3 most populated clusters for each TM protein and up to 5 representative structures from each cluster were then compared with the NMR structure for further analysis. The best ones, in terms of crossing angle and root mean square deviation (RMSD), are reported in Table1. Details of the methods are given in the Supporting Information. Because even the CG simulations are still time consuming and require some expertise for set-up and analysis, we also carried out *ab initio* prediction the TM homo/heterodimer configurations for all the 11 receptors from the amino acid sequences using the PREDDIMER server. This approach is based on the alignment of the peptide's surfaces, with the possible dimers ranked by a parameter, Fscor, which consider the best packing and polar/hydrophobic complementarity between peptides and exposed dimer surface to lipids.[16,17] Most important here is that PREDDIMER also gives several models (i.e. alternate configurations) which can be compared with the structures from the CG simulations (Table S3). We also made the comparison because PREDDIMER had been used before for predicting the TM dimer structures of these receptors.[16–18]

The comparison of the RMSD of the CG simulated TM dimers with the NMR structures (Fig 2, Table 1) shows that the CG structures range from nearly similar (1.4 Å for ErbB1-B2 dimer) to more distant (4.8 Å for PDGFRb). The same is the case for the *ab initio* PREDDIMER models (Table S3) where the RMSD with the NMR structure also ranges from 1.7 Å (for GpA) to very distant 9 Å (for ErbB1/2). Interestingly, most of the CG simulated TM

dimers (10 out of 11) share similar values of crossing angle with that of the NMR structures (Table 1). The exception is PDGFRb, where the CG TM dimer configurations are right-handed (a negative value of crossing angle) compared to left-handed (a positive value of crossing angle) NMR dimer structures. Furthermore, the PREDDIMER models share similarities with both CG simulated and the NMR derived structures (Table 1 and S3). As shown in Fig S3, the best fit structure (explained in the Fig legend) from the 3rd cluster of PDGFRb from simulation is right-handed dimer with a larger crossing angle of -55° and shares close similarity (RMSD 3.2 Å) with the 3rd predicted model. By contrast the NMR structure for this protein's TM domain is left-handed dimer with crossing angle of 23° which shares similarity (RMSD 2.3Å) with the 2nd predicted model. As the *ab initio* PREDDIMER program[16] predicts different possible arrangements between the TM peptides, our CG simulation also provides similar information represented by the best fit configurations from the three most populated clusters. (Table 1 and S3). We observed transitions between the clusters for all the receptors (Fig S1, S2 and S4). Moreover, the analysis of the overall TM crossing angle distribution (Fig S2 and S4) for all the 11 TM dimers shows the existence of both the right- and left-handed configurations. By contrast, the NMR structures of each TM dimer only provide one configuration, resolved at particular pH and in a suitable bicelle or micelle environment. Therefore, as shown in Table 1 and S3, the TM models obtained from the CG simulations go hand in hand with *ab initio* PREDDIMER predicted models and share similarity with the NMR structures in terms of RMSD, Fscor and inter-helical crossing angle (Fig S3). However, there are also examples of less than perfect agreement in case of Bnip3, EphA2, FGFR3 and PDGFRb. Possible origins of these differences are discussed below.

A switching between the interaction interface of the TM regions results in different helix orientations in the membrane environment which needs to be coordinated with the proximity as well as orientation with both the extracellular and the intracellular domains. To measure the change in TM orientation during the simulations, we calculated the 2D distribution of helix dimerization as a function of inter helical distance and the inter helical crossing angle (Fig S2) and also plotted the contact map for the obtained TM dimers (Fig S5). As expected, based on population density, both left-handed or right-handed configurations of the TM dimer are seen in some of the cases, most clearly for EphA1 (but several configurations- even though some are relatively close in crossing angle are seen for most). The dimeric interface of the TM peptides from our study shows a high level of similarity with the NMR structures (Fig S5). All the TM dimers except ErbB3, FGFR3 and PDGFRb contain a canonical extended GxxxG or SxxxG motif, starting from the middle of the sequence to the C-termini of the TM domains (Table S2). This motif allows for potential stabilization by Cα-H hydrogen bonding in the typical GAS_{right} dimer motif.[21,22] In fact, EphA1, EphA2, ErbB1, ErbB1/2, ErbB3, ErbB4, FGFR3 and GpA showed an equal population of both right-handed and left-handed TM dimer configurations. Structural studies on the isolated TM helices of EphA1, EphA2, ErbB2, ErbB3, and the heterodimer ErbB1/ ErbB2 have also shown that the TM helices dimerize in two different ways: either right-handed or left-handed helical dimers.[5,6,8–10]. By contrast TM dimers of Bnip3, ErbB2 and PDGFRb are mostly populated as the right-handed configurations (Fig S2). In general, some ambiguity remains about what governs the arrangement in either of the two configurations and whether both configurations are relevant for activation/may represent different activation states. Beyond this, it is not yet clear for most of the systems how the interaction of the soluble extracellular and intracellular domains either supports or hinders the formation of thermodynamically favoured TM dimers. Several studies also suggested that the lipid environment may control the mode of dimerization.[14,17,23] As a further examination of the convergence of the dimer binding modes, we calculated the free energy of binding for EphA1,

EphA2, ErbB2 and for the well-studied helix dimer of GpA (for details, see Fig S6, Fig S7 and additional discussion in the supporting material).

Caution needs to be exercised when comparing *ab initio* predictions, MD simulation results or even structures derived from experiments. First, the comparison needs to be performed for the same TM peptide or protein construct, as well as for the same membrane composition. Second, some of the NMR structures were determined in DPC detergent, rather than in bicelles, but there is no systematic difference among this set of examples. However, even bicelles are not the idealized DMPC-composed planar bilayer in the middle, with DHPC detergent at the bicelle-disk edge, as some studies are showing a peptide dependent mixing of the DHPC and DMPC molecules. Third, the NMR derived structures are biased by restraints between the two helices which are for the most part symmetric. By contrast, in the simulations a sliding of helices relative to one another is observed (most noticeably for EphA2 and ErbB2). It is worth noticing that *ab initio* predictions as PREDDIMER cannot consider variations in membrane thickness, lipid composition or salt concentration in the aqueous solution, while MD simulations with Martini 3 can naturally include these environmental effects. Even differences in pH can be mimicked by defining different charged states for acidic/basic groups, or by using the Titratable Martini approach.[24] In addition, trimers or oligomer can be accurately studied[25] with Martini 3 or even interactions with protein complexes. Additional simulations and experimental studies are needed to delineate the possible effects of several such variables on the TM dimer structural configurations and their stability. In conclusion, we have demonstrated with an established test set that Martini 3 CG simulations can reveal both the strength of TM dimers, as well as the tendency to form alternate structures, reflecting different specific residue-residue (motif) interactions, between the TM peptides in an overall reliable manner.

Methods

Data Sets and Modeling of the TM peptides

The NMR structures of 11 TM dimers were extracted from the PDB database, using the first structure of the ensemble. All the TM dimer structures were determined either in DPC micelles or DMPC/DHPC bicelles with the pH value ranges from 4.5 to 6.8 (Table S1). The TM sequences for all the 11 receptors were extracted from the UniProt database. All the TM peptides of these 11 receptors were modeled as an ideal α-helix using PyMOL 2.4, which was maintained by a DSSP algorithm (which determines the secondary structure of the peptide backbone) during the simulations, as is standard for regular secondary structure in Martini. We then added extra 8-9 residues at N-terminal and 10 residues at the C-terminal of the TM models as an extended conformation (Table S2) as these extra- and intracellular membrane proximal regions are known to provide better stability for some systems in the lipid bilayer.

Coarse-grain molecular dynamics simulation

In order to characterize the dimerization of TMs, we built 11 TM peptide systems with the monomers placed 50 Å apart from each other (Fig 1). For this, atomistic (AT) modeled systems of all the 11 TMs were converted to coarse-grained (CG) representation using the *martinize2.py* workflow module of the MARTINI 3 force field[20] (see https://github.com/marrink-lab/vermouth-martinize) considering the secondary structure DSSP assignment.[26] CG simulations were performed using GROMACS version 2016.5.[27] The *insane.py* script[28] was used for setting up of the DMPC bilayer (typically 306 lipids and 4870 CG water molecules)

around the peptides in a cubic box with dimensions of 100×100×100 $Å^3$. The pH of the systems was considered neutral. All the simulations were run in presence of regular MARTINI water and neutralised and brought up to 0.15M NaCl. The systems were equilibrated for 500 ps. The long-range electrostatic interactions were used with a reaction type field having a cutoff value of 11 Å.[29] We used potential-shift-verlet for the Lennard-Jones interactions with a value of 11 Å for the cutoff scheme and the V-rescale thermostat with a reference temperature of 320 K in combination with a Berendsen barostat with a coupling constant of 1.0 ps, compressibility of 3.0×10^{-4} bar $^{-1}$, and a reference pressure of 1 bar was used. The integration time step was 20 fs. All the simulations were run in quadruplicate for 4 µs. For further analysis and comparison, the extracted CG structures were then converted to all atomistic (AA) representation using the Backward tool[30]of Martini for further analysis.

Data Analysis

The PREDDIMER webserver was used both to predict models for the 11 TM dimers *ab initio*, but also to analyse the NMR and CG MD derived structures based on the Fscor and helix crossing angle.[31] Analysis of the trajectories was carried out using tools in GROMACS.[27] All the analysis have been carried out considering only the TM region of the peptides. The contact maps between the helices (i.e. hydrophobic TM regions only) were calculated with a distance cut off 5.5 Å for all the backbone and side-chain atoms. Interhelix distances are calculated between the center of masses of the TM, but not of the N- and C-terminal regions. 2D plots were plotted in Origin2020b using the inter-helical distance and the inter-helical angle between the TM helices, as x- and y-axes respectively. Each plot shows the presence of one global minimum and up to 2 local minima and we therefore divided into three populations. From each of these three populations, 5 representative configurations were extracted and then further compared with the experimentally derived NMR structures. The best one, in terms of RMSD and crossing angles, was then reported for the 3 minima (Table 1 and Figure 2). The comparison with the NMR structure is done by calculation of backbone RMSD considering only the TM regions of the receptors.

Umbrella Sampling (US) simulations

We used the umbrella sampling method with the center of mass (COM) distance between the TM peptides as the reaction coordinate and calculated the free energy of TM association of these receptor's TM domains. We calculated the dimerization free energy profiles for the TM peptides of GpA, EphA1, EphA2 and ErbB2, as examples using a standard umbrella sampling (US) protocol. The best fit structures obtained from the three most populated clusters (according to 2D plot in Fig S2) from the CG simulation were chosen as the starting point for the US simulations. Unlike the CG simulation above, the peptide length was reduced by removing the N- and C-terminal residues and we have considered only the transmembrane regions for all the US simulations. The TM dimers were embedded in the lipid bilayer consisting of typically 328 DMPC molecules, 5200 water beads with a salt concentration of ~0.15 NaCl using the same *insane.py*.[28] The center of mass distance between the TM domains was used as the reaction coordinate (ξ). US simulations were set up with 19 replicas spaced between the reaction coordinate of 0 to 3.6 nm with a spacing of 0.2 nm. The spring constant of 1000 kJ mol^{-1} nm^{-2} was employed for the TM domains. Each US window was simulated for 0.5 µs with a time step of 20 fs. The gmx wham[32] tool was used for calculating the potential mean force (PMF) profiles, estimating the free energy of TM domain dimerization. All the simulations were performed in GROMACS 2016.5. In order to calculate an average free energy

and standard deviation between the US simulations begun with the three different conformers, the lowest energy minima values for each were weighted by the CG populations, derived from the three population clusters (listed in the legend of the 2D plot shown in Fig S2).

References:

1. Wallin, E. & von Heijne, G. Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms. *Protein Sci.* **7**, 1029–1038 (1998).

2. Arkin, I. T. & Brunger, A. T. Statistical analysis of predicted transmembrane alpha-helices. *Biochim. Biophys. Acta* **1429**, 113–128 (1998).

3. Polyansky, A. A., Volynsky, P. E. & Efremov, R. G. Structural, dynamic, and functional aspects of helix association in membranes: a computational view. *Adv. Protein Chem. Struct. Biol.* **83**, 129–161 (2011).

4. Bocharov, E. V *et al.* Unique dimeric structure of BNip3 transmembrane domain suggests membrane permeabilization as a cell death trigger. *J. Biol. Chem.* **282**, 16256–16266 (2007).

5. Bocharov, E. V *et al.* Spatial structure and pH-dependent conformational diversity of dimeric transmembrane domain of the receptor tyrosine kinase EphA1. *J. Biol. Chem.* **283**, 29385–29395 (2008).

6. Bocharov, E. V *et al.* Left-handed dimer of EphA2 transmembrane domain: Helix packing diversity among receptor tyrosine kinases. *Biophys. J.* **98**, 881–889 (2010).

7. Endres, N. F. *et al.* Conformational coupling across the plasma membrane in activation of the EGF receptor. *Cell* **152**, 543–556 (2013).

8. Mineev, K. S. *et al.* Spatial structure of the transmembrane domain heterodimer of ErbB1 and ErbB2 receptor tyrosine kinases. *J. Mol. Biol.* **400**, 231–243 (2010).

9. Bocharov, E. V *et al.* Spatial structure of the dimeric transmembrane domain of the growth factor receptor ErbB2 presumably corresponding to the receptor active state. *J. Biol. Chem.* **283**, 6950–6956 (2008).

10. Mineev, K. S. *et al.* Spatial structure and dimer--monomer equilibrium of the ErbB3 transmembrane domain in DPC micelles. *Biochim. Biophys. Acta* **1808**, 2081–2088 (2011).

11. Bocharov, E. V, Mineev, K. S., Goncharuk, M. V & Arseniev, A. S. Structural and thermodynamic insight into the process of 'weak' dimerization of the ErbB4 transmembrane domain by solution NMR. *Biochim. Biophys. Acta* **1818**, 2158–2170 (2012).

12. Bocharov, E. V *et al.* Structure of FGFR3 transmembrane domain dimer: implications for signaling and human pathologies. *Structure* **21**, 2087–2093 (2013).

13. MacKenzie, K. R., Prestegard, J. H. & Engelman, D. M. A transmembrane helix dimer: structure and implications. *Science* **276**, 131–133 (1997).

14. Muhle-Goll, C. *et al.* Hydrophobic matching controls the tilt and stability of the dimeric platelet-derived growth factor receptor (PDGFR) β transmembrane segment. *J. Biol. Chem.* **287**, 26178–26186 (2012).

15. Sahoo, A. R. & Buck, M. Structural and Functional Insights into the Transmembrane Domain Association of Eph Receptors. *Int. J. Mol. Sci.* **22**, (2021).

16. Polyansky, A. A. *et al.* PREDDIMER: a web server for prediction of transmembrane helical dimers. *Bioinformatics* **30**, 889–890 (2014).

17. Polyansky, A. A., Volynsky, P. E. & Efremov, R. G. Multistate Organization of Transmembrane Helical Protein Dimers Governed by the Host Membrane. *J. Am. Chem. Soc.* **134**, 14390–14400 (2012).

18. Zhang, L., Sodt, A. J., Venable, R. M., Pastor, R. W. & Buck, M. Prediction, refinement, and persistency of transmembrane helix dimers in lipid bilayers using implicit and explicit solvent/lipid representations: microsecond molecular dynamics simulations of ErbB1/B2 and EphA1. *Proteins* **81**, 365–376 (2013).

19. Chavent, M., Duncan, A. L. & Sansom, M. S. Molecular dynamics simulations of membrane proteins and their interactions: from nanoscale to mesoscale. *Curr. Opin. Struct. Biol.* **40**, 8–16 (2016).

20. Souza, P. C. T. *et al.* Martini 3: a general purpose force field for coarse-grained molecular dynamics. *Nat. Methods* **18**, 382–388 (2021).

21. Cymer, F., Veerappan, A. & Schneider, D. Transmembrane helix–helix interactions are modulated by the sequence context and by lipid bilayer properties. *Biochim. Biophys. Acta - Biomembr.* **1818**, 963–973 (2012).

22. Anderson, S. M., Mueller, B. K., Lange, E. J. & Senes, A. 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).

23. Gopal, S. M., Pawar, A. B., Wassenaar, T. A. & Sengupta, D. Lipid-dependent conformational landscape of the ErbB2 growth factor receptor dimers. *Chem. Phys. Lipids* **230**, 104911 (2020).

24. Grünewald, F. *et al.* Titratable Martini model for constant pH simulations. *J. Chem. Phys.* **153**, 24118 (2020).

25. Westerfield, J. M. *et al.* Conformational clamping by a membrane ligand activates the EphA2 receptor. *J. Mol. Biol.* **433** (18), 167144 (2021).

26. Kabsch, W. & Sander, C. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **22**, 2577–2637 (1983).

27. Abraham, M. J. *et al.* GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1–2**, 19–25 (2015).

28. Wassenaar, T. A., Ingólfsson, H. I., Böckmann, R. A., Tieleman, D. P. & Marrink, S. J. Computational Lipidomics with insane: A Versatile Tool for Generating Custom Membranes for Molecular Simulations. *J. Chem. Theory Comput.* **11**, 2144–2155 (2015).

29. de Jong, D. H., Baoukina, S., Ingólfsson, H. I. & Marrink, S. J. Martini straight: Boosting performance using a shorter cutoff and GPUs. *Comput. Phys. Commun.* **199**, 1–7 (2016).

30. Wassenaar, T. A., Pluhackova, K., Böckmann, R. A., Marrink, S. J. & Tieleman, D. P. Going Backward: A Flexible Geometric Approach to Reverse Transformation from

Coarse Grained to Atomistic Models. *J. Chem. Theory Comput.* **10**, 676–690 (2014).

31. Polyansky, A. A. *et al.* PREDDIMER: a web server for prediction of transmembrane helical dimers. *Bioinformatics* **30**, 889–890 (2014).

32. Hub, J. S., de Groot, B. L. & van der Spoel, D. g_wham—A Free Weighted Histogram Analysis Implementation Including Robust Error and Autocorrelation Estimates. *J. Chem. Theory Comput.* **6**, 3713–3720 (2010).

Acknowledgements: This work is supported by a NIH R01 grant from the National Eye Institute R01EY029169 and previous grants from NIGMS (R01GM073071 and R01GM092851) to the Buck lab.

Author Contributions: ARS generated the protein coarse-grained models and performed the MD simulations and analysed the data. ZM carried out early simulations with the beta release of Martini 3. PCTS provided critical data and discussion. ARS and MB co-wrote the paper. All authors read and approved the final version of the manuscript and supporting information.

Competing Interests: Authors declare no competing interests.

Supporting Information Available: Materials and methods, additional results, supporting tables and figures are included in the supporting information

Table 1: Comparison of CG simulated TM dimers with the NMR structures. Central conformers of the three most populated clusters from the CG simulation are considered for calculating the mean and SD of the crossing angle values. Similar/ near crossing angle (X) values (± 20°) and RMSD ≤ 4.5 Å between the CG and NMR are marked as red. All the receptors except PDGFRb show similarity in the crossing angle values and are within the RMSD range of 4.5 Å.

TM Dimers	CG simulation						NMR
	X (deg)			RMSD (Å) from NMR			X (deg)
	1st Cluster	2nd Cluster	3rd Cluster	1st Cluster	2nd Cluster	3rd Cluster	
1-Bnip3	-60 ± 10	-87 ± 4	-89 ± 5	4.1 ± 0.5	5.2 ± 0.3	5.4 ± 0.3	-41
2-EphA1	-47 ± 5	49 ± 6	-15 ± 10	3.5 ± 0.2	6.5 ± 0.2	4.8 ± 0.1	-52
3-EphA2	-11 ± 3	29 ± 4	33 ± 8	5.1 ± 0.7	5.2 ± 0.5	4.6 ± 0.1	17
4-ErbB1	-30 ± 2	-8 ± 6	-20 ± 8	4.2 ± 0.3	4.3 ± 0.3	4.6 ± 0.3	-42
5-ErbB1/2	-16 ± 4	-55 ± 5	-26 ± 5	4.0 ± 0.2	3.6 ± 0.3	2.4 ± 0.9	-51
6-ErbB2	-60 ± 4	-18 ± 5	-88 ± 8	2.7 ± 0.9	4.1 ± 0.1	3.6 ± 0.6	-43
7-ErbB3	40 ± 11	-8 ± 4	-41 ± 9	3.9 ± 0.1	4.5 ± 0.2	5.1 ± 0.3	30
8-ErbB4	-30 ± 4	-56 ± 4	-10 ± 6	3.9 ± 0.4	4.1 ± 0.2	5.4 ± 0.3	-41
9-FGFR3	-18 ± 8	-50 ± 3	34 ± 5	5.6 ± 0.1	5.2 ± 0.2	5.6 ± 1.2	35
10-GpA	-19 ± 5	0 ± 4	14 ± 3	3.0 ± 0.2	4.6 ± 0.2	3.4 ± 0.3	-39
11-PDGFRb	-84 ± 3	-132 ± 7	-61 ± 8	5.5 ± 0.4	6.9 ± 0.1	5.3 ± 0.5	23

FIGURE LEGENDS

Figure 1. Association of TM dimers in the DMPC bilayer. TM peptides are initially placed 50 Å apart from each other and then inserted into DPMC bilayer. TM peptides interact quickly and the final 4 µs dimer conformation is shown here. TM region of the peptides are shown as magenta and red. N- and C-terminal regions are shown as yellow.

Figure 2. Superposition of the 11 solution NMR TM dimer structures (in yellow) with best fit CG simulated structures (vars. colors). All structures are shown in cartoon/mainchain representation. Best fit structure represents the one (among the structures from all the clusters) that has lowest backbone RMSD from the NMR structure. Only the TM region is considered for backbone RMSD calculation.

Figure 1.

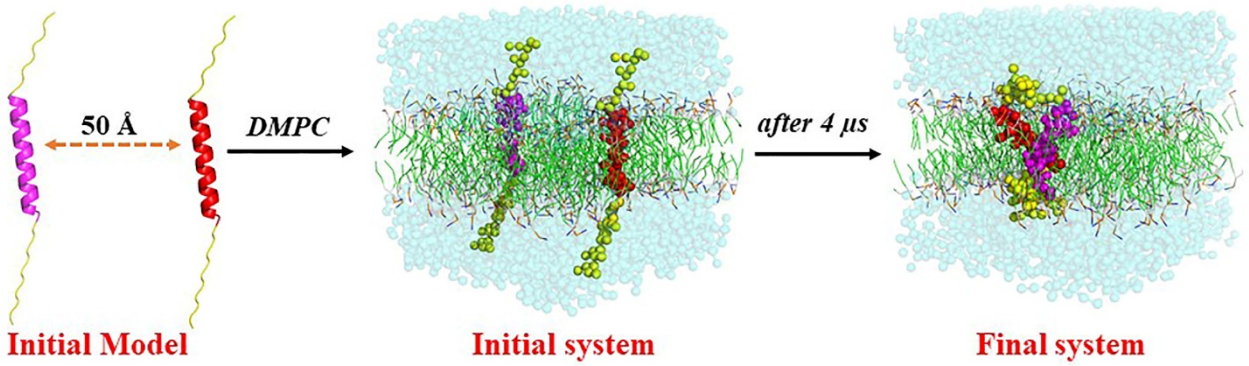
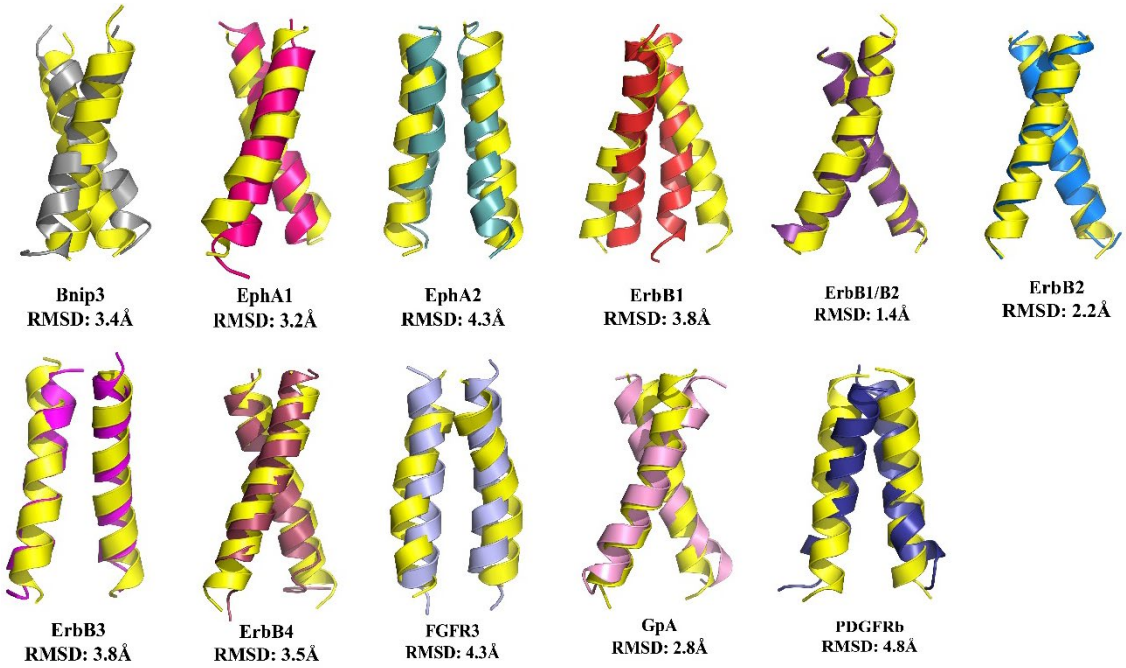


Figure 2.


Bnip3
RMSD: 3.4Å
EphA1
RMSD: 3.2Å
EphA2
RMSD: 4.3Å
ErbB1
RMSD: 3.8Å
ErbB1/B2
RMSD: 1.4Å
ErbB2
RMSD: 2.2Å
ErbB3
RMSD: 3.8Å
ErbB4
RMSD: 3.5Å
FGFR3
RMSD: 4.3Å
GpA
RMSD: 2.8Å
PDGFRb
RMSD: 4.8Å