Architecture and membrane interactions of the EGF receptor

Supplemental Information

Anton Arkhipov, Yibing Shan, Rahul Das, Nicholas F. Endres, Michael P. Eastwood,

David E. Wemmer, John Kuriyan, and David E. Shaw

Supplemental Figure Legends

Figure S1, related to Figure 3. Simulations of the extracellular dimers of EGFR. EGFR subunits shown in red and blue, EGF in orange. The force fields (AMBER ff99SB-ILDN or CHARMM22*; see Experimental Procedures) used in the simulations are labeled. Simulations using the same setup and force field are each assigned an identification number as shown. Simulations "AMBER (1)" for 2-ligand, 1-ligand, and ligand-free dimers are those shown in Figure 3 of the main text. The length of a simulation and the distance (d_{CC}) between the C-termini of the two extracellular domains averaged over the last 3 μs is shown. (A) The 2-ligand extracellular dimer, in which the C-termini largely remain close to each other. (B) The 1-ligand extracellular dimer, in which the C-termini are separated. Two snapshots are shown for the AMBER simulation of the 1-ligand dimer (AMBER (2)). In this simulation, a large conformational change in domain IV similar to that seen in all other 1-ligand dimer simulations was observed at around t = 1 μs. Later in the simulation, the dimer switched back to the initial conformation in which the C-termini of the extracellular domains are close to each other. (C) The ligand-free extracellular dimers, in which the C-termini are also separated. (D) Residue–residue contacts in simulations of the ligand-free extracellular dimers (residues starting from position 300 are shown, corresponding to the C-terminus of domain II and domains III and IV).

A contact is marked if two residues are within 5 Å of each other for a total of at least 10% of the simulation time. Two groups of asymmetric contacts are marked with green and magenta frames and illustrated on the right using a structure from one of the simulations. (E) Distribution of the distance d_{CC} in simulations of the extracellular dimers using the AMBER force field (left) and using the CHARMM force field (right). The distributions observed in simulations of the complete active and inactive EGFR dimers (Figure 6) are also shown. (F) Root-mean-square deviation (RMSD) of the dimer of EGFR extracellular modules from the crystal structure of the 2-ligand dimer (PDB entry 3NJP). The distributions are obtained by aggregating the RMSDs from simulations of the same dimer type (2-ligand, 1-ligand, or ligand-free). (G) "Staggered" and "flush" conformations of the extracellular dimers, as observed in PDB entries 1IVO and 3NJP for the former and 1MOX for the latter shown from the side and the top. For the latter view, the yellow subunits of the two crystal structures are superposed to show the difference between the two structures in terms of the relative orientations of the two subunits. The orientation can be characterized by the angle θ formed by the Cα atoms of Ile190 and Pro204 of one subunit and Pro204 of the other. Distributions of θ observed in our simulations (right) show fluctuations between the staggered and flush conformations. Two slightly different θ values are obtained for each crystal structure, because the structures are not exactly symmetric; the space between these values is depicted as colored rectangles on the plots.

Figure S2, related to Figure 4. Simulations of transmembrane dimers of EGFR mutants and Her2. The distance between the dimerization interface (d_{int}) shown as a function of simulation time. (A) Effect of mutations of the GxxxG-like motif to the stability of the N-terminal transmembrane dimer. The transmembrane dimers with the quadruple mutation (T624I/G625I/G628I/A629I) dissociate readily in two simulations, whereas those with single mutations such as T624L, G625L, G628L, and A629L (Lu et al., 2010) are stable on the time scale of ~20 μs. (B) As shown, the Her2 N-terminal dimer dissociated and reformed. A similar event was observed for the EGFR N-terminal transmembrane dimer (Figure 4D, between 190 and 200 μs).

Figure S3, related to Figure 5. Simulations of the TM–JM-A constructs. (A) The number of satisfied experimental NOEs as a function of simulation time. The plots and images on the extreme left illustrate simulations of the left-handed transmembrane dimer assembled with the JM-A dimer. As shown, the left-handed transmembrane conformation was not stable in the simulation, and the number of satisfied NOEs decreased significantly in the course of the simulations. The black and red plots correspond to two independent simulations; in both simulations, the starting structure satisfied all 13 NOEs. Conformations at the beginning and end of one simulation (black on the plot) are shown. The remaining plots correspond to the simulations of TM–JM-A constructs in DMPC and POPC/POPS membranes, as shown in Figure 5 and in panel (E) of this figure, as well as simulations of the active dimers of complete EGFR and TM–JM–KD construct (Figures 6 and 7). (B, C) Interactions between the Arg/Lys side chains of the JM-A segment with the lipid head groups. The JM-A dimer is illustrated in (B) and a membrane-embedded monomer in (C). Electrostatic interactions are highlighted with green lines. The three hydrophobic residues in the JM-A segments (Leu655, Leu658, and Leu659) are shown in orange. (D) The C-terminal transmembrane dimer stabilized by the membrane-embedded JM-A segments. The plot shows the distance between the C-terminal GxxxG-like motifs of the transmembrane helices. (E) The JM-A dimer stable when connected to the N-terminal transmembrane dimer and destabilized when connected to the C-terminal transmembrane dimer. Satisfaction of the inter-chain juxtamembrane NOEs, as measured experimentally, is used here as an indicator of dimer stability. The simulation with the N-terminal transmembrane dimer is illustrated on the left (another similar simulation is illustrated in Figure 5A) and two simulations with the C-terminal transmembrane dimer in the center and on the right. In the center, the juxtamembrane dimer breaks completely long before the end of the ~6-μs simulation. (F) Helicity of the JM-A segments (see Figure 5C) in an active dimer based on the data from a simulation of an active dimer of the TM–JM–KD construct.

Figure S4, related to Figure 7. Effect of EGFR interaction with the anionic lipids. (A) Details of the interaction of the inactive kinase dimer with anionic lipids. (B) The fraction of

POPS among all the lipids in contact with the JM–KD portion. The dashed lines are used to indicate the fraction of POPS among all the lipids in the intracellular leaflet in a simulation. Note that all POPS lipids are located in the intracellular leaflet, and thus the fraction of POPS in this leaflet is twice its fraction in the membrane. (C) Clustering of anionic lipids (red) toward EGFR kinase in a simulation. The shaded areas are areas of the membrane inner leaflet in direct contact with EGFR. (D) Detachment of the kinase domain from the inner membrane. (E) The linearity of EGF-stimulated EGFR activity (blue) suggesting that EGFR is fully dimerized and that its slope corresponds to the activity of an EGFR dimer. The dependence of EGF-free activity on EGFR density at the cell surface (the companion paper) is better fit by a decreasing dissociation constant of EGFR active dimer (red) than an invariable one (brown), suggesting an autoinhibition mechanism that breaks down at high EGFR density. Details of the fitting are included in the Supplemental Experimental Procedures.

Supplemental Movie Legends

Movie S1, related to Figure 3. The model of the inactive EGFR extracellular dimer. The simulated conformational change of the extracellular dimer upon removal of a bound ligand from the ligand-bound (active) extracellular dimer produces the conformation of an inactive extracellular dimer in which the two C-termini of the extracellular modules are separated by a significant distance.

Movie S2, related to Figure 6. The model of the near-complete inactive EGFR dimer in simulation.

Movie S3, related to Figure 6. The model of the near-complete active EGFR dimer in simulation.

Supplemental Structure Files Legends

Structure S1, related to Figure 6. Coordinates of the structural model of intact EGFR monomer.

Structure S2, related to Figure 6. Coordinates of the structural model of the intact EGFR inactive dimer.

Structure S3, related to Figure 6. Coordinates of the structural model of the intact EGFR active dimer.

Supplemental Experimental Procedures

Simulation protocols

A typical simulation system in this study consists proteins, lipids, and water molecules. The AMBER ff99SB-ILDN (Cornell et al., 1995; Hornak et al., 2006; Lindorff-Larsen et al., 2010) force field, combined with the ff99SB* backbone correction (Best and Hummer, 2009), was used for proteins; the CHARMM C36 force field was used for lipids (Klauda et al., 2010); and the water model was TIP3P (Jorgensen et al., 1983). As indicated in Results, for a number of systems studied additional simulations were also performed using the CHARMM22* force field (MacKerell et al., 1998; Piana et al., 2011) for proteins. The simulated systems were solvated in water with 0.15 M NaCl. Residue protonation states corresponded to pH 7.

The protein backbone atoms were restrained to their initial positions using a harmonic potential with a force constant of 1 kcal mol^{-1} $Å^{-2}$) for at least 5 ns as an equilibration step; restraints were subsequently removed. Simulations were mostly performed in the NPT ensemble with T = 310 K, P = 1 bar, and Berendsen's coupling scheme (Berendsen et al., 1984) with one temperature group (NVT was used for those post-equilibration simulations that did not involve the membrane). Water molecules and all bond lengths to hydrogen atoms were constrained using M-SHAKE (Krautler et al., 2001). Van der Waals and short-range electrostatic interactions were cut off at 13.5 Å for simulations with <100,000 atoms and at 10.5 Å otherwise. Long-range electrostatic interactions were calculated using the *k*-space Gaussian Split Ewald method (Shan et al., 2005) with a 32 × 32 × 32 or 64 × 64 × 64 mesh. The simulation time step was 1 fs for the equilibration stage and 2.5 fs for production simulations; the r-RESPA integration method (Tuckerman et al., 1992) was used, with long-range electrostatics evaluated every 5 fs.

The model membrane consisted of neutral POPC lipids, with 15% (molar) POPC randomly replaced by negatively charged POPS lipids. This POPS fraction was chosen to mimic the abundance of anionic lipids in the mammalian plasma membrane (Zachowski, 1993; van Meer et al., 2008), in which ~10% of lipids are PS species, and few more per cent are other anionic species, such as phosphoinositides, the total content of anionic lipids thus being up to 15%. The POPS lipids were introduced only in the intracellular leaflet, which is where the anionic lipids are almost exclusively found in cell membranes. The lipid content thus was 30% POPS and 70% POPC in the intracellular and 100% POPC in the extracellular leaflets. The POPS fraction specified in the text always refers to the overall fraction, and, unless specified otherwise, it is 15%. As indicated in Results, a number of simulations were repeated with a smaller POPS content (0, 2.5, and 7.5% of the total lipid molecules), to investigate the effect of the membrane charge. In a number of simulations of TM–JM-A construct a pure DMPC bilayer was used to mimic the conditions of the NMR experiments (Endres et al., 2012). Modeling and preparation for simulations, as well as analysis and visualization, were performed using VMD (Humphrey et al., 1996).

Simulations

To check the robustness of the observed coupling between ligand binding and domain IV conformation in EGFR dimers, an observation from our simulations using the AMBER ff99SB-ILDN force field (see Experimental Protocols), we repeated the simulations using another force field (CHARMM22*). In the CHARMM simulations (Figure S1), the 1-ligand and ligand-free extracellular dimers arrived at a conformation with widely separated C-termini similar to that reached in the AMBER simulations. As indicated by the histograms of d_{CC} and the angle of "stagger" (Figure S1), while the extracellular dimers in the CHARMM simulations tended to be more conformationally variable than in the AMBER simulations, the overall trends, such as the transition from "staggered" to "flush" conformations and the transition to a larger d_{CC} upon removal of the bound ligands, were robustly produced by both force fields. Contact maps of the residues of the two domain IVs after their bending in the simulations of the ligand-free dimer

(Figure S1D) may be useful in the design of crosslinking experiments to further validate the conformation of the ligand-free extracellular domains.

Summary of simulations

Details of the simulations performed for the study reported here are summarized in the table below. Most simulations involved the membrane, with the exception of simulations of the extracellular modules. When applicable, multiple simulations of the same category are placed in one row. Unless specified otherwise, the content of POPS in the membrane is 15% and the force field used for protein is AMBER ff99SB-ILDN (see Experimental Procedures in the main text). These simulations of systems of relatively small size (~34,000–48,000 atoms) were performed by a 512-node Anton supercomputer; systems totaling ~141,000–272,000 atoms were simulated on a 1024-node Anton; larger systems (beyond 272,000 atoms) were simulated on a 2048-node Anton. Most simulations were ended when the system was found by visual inspection to settle in a conformation. Several simulations of the transmembrane helices were extended to a simulation time of ~100–200 μs to test the stability of the transmembrane dimers or to directly simulate the dimerization of the helices.

Description	**Number of atoms (rounded to 1,000)**	**Duration (μs)**	**Membrane size (Å^2)**	**Illustration**
Simulations of the extracellular domains				
Tethered extracellular monomer	241,000; 251,000	3.6; 2.7	NA	Figure 2
Extended extracellular monomer with EGF	394,000	5.5		Figure 2
Extended extracellular monomer without EGF	395,000	5.4		Figure 2
The 2-ligand extracellular dimer	272,000–391,000	12.8; 6.2		Figure 3, S1
The 2-ligand extracellular dimer (CHARMM22*)	271,000	10.2; 10.5; 8.7		Figure S1
The 1-ligand extracellular dimer	272,000	14.1; 10.7		Figure 3, S1
The 1-ligand extracellular dimer (CHARMM22*)	271,000	15.0; 10.5; 10.8		Figure S1
The ligand-free extracellular dimer	272,000	11.0		Figure 3, S1
The ligand-free extracellular dimer	271,000	16.6; 7.5		Figure S1

(CHARMM22*)				
Simulations of the transmembrane segments				
The N-terminal transmembrane dimer	38,000	100.2; 51.2	68 × 68	Figure 4
The C-terminal transmembrane dimer	35,000–38,000	30.4; 14.5; 10.7; 33.4	68 × 68	Figure 4
The N-terminal transmembrane dimer (I640E)	36,000	38.6	68 × 68	Figure 4
The N-terminal transmembrane dimer (T624L, G625L, G628L, A629L)	36,000	19.2; 18.5; 18.5; 18.1	68 × 68	Figure 4
The N-terminal transmembrane dimer (T624I/G625I/G628I/A629I)	36,000	5.2; 1.3	68 × 68	Figure 4
The C-terminal transmembrane dimer (I640E)	36,000	19.0; 33.9	68 × 68	Figure 4
Self-assembly simulation with 4 transmembrane helices	37,000	201.2	68 × 68	Figure 4
Self-assembly simulation with 9 transmembrane helices	37,000	103.1	68 × 68	Figure 4
The HER2 N-terminal transmembrane dimer	33,000	134.6	68 × 68	Figure S2
The HER2 C-terminal transmembrane dimer	34,000	120.8	68 × 68	Figure S2
Simulations of the TM–JM-A construct				
The N-terminal transmembrane dimer and the JM-A dimer	45,000	125.0; 40.3	70 × 70	Figure 5, S3
The N-terminal transmembrane dimer and the JM-A dimer (in DMPC bilayer)	46,000	19.0; 14.5, 14.3; 103.0	70 × 70	Figure 5
The left-handed transmembrane dimer and the JM-A dimer (in DMPC bilayer)	46,000	5.1; 10.1	70 × 70	Figure S3
The C-terminal transmembrane dimer and the JM-A dimer	39,000	30.4; 6.4	70 × 70	Figure 5, S3
The TM–JM-A monomer	48,000	18.8; 10.3	68 × 68	Figure 5
Simulations of the TM–JM–KD construct				
The TM–JM–KD, monomer	121,000	2.7	94 × 94	Figure 7
The TM–JM–KD, active dimer	244,000	21.7	138 × 138	Figure 7
The TM–JM–KD, inactive dimer	229,000	26.7; 11.5	138 × 138	Figure 7, S4
The TM–JM–KD, inactive dimer (7.5% POPS)	204,000	10.2	138 × 138	Figure 7
The TM–JM–KD, inactive dimer (2.5% POPS)	246,000	16.6	138 × 138	Figure 7

The TM–JM–KD, inactive dimer (without POPS)	204,000–263,000	8.7	138 × 138	Figure S4
Simulations of the monomer and dimers of the near-complete EGFR				
Monomer	229,000	4.7	116 × 116	Figure 6
Inactive dimer	422,000	4.1; 1.5; 2.1	148 × 148	Figure 6
Active dimer	554,000	4.7	150 × 150	Figure 6

Simulations of the transmembrane segments

An EGFR transmembrane segment in our simulations includes residues 615–647, which correspond to residues 624–656 in HER2. The self-assembly simulations for EGFR transmembrane helices employed either four or nine helices that were initially oriented perpendicular to the membrane. The helices readily adopted an angle of ~20° with respect to the normal vector of the membrane in the simulations. The transmembrane helices were initially placed according to a 2 × 2 or 3 × 3 uniform grid.

In simulations of the I640E mutants, the glutamates were protonated as they are expected to be in the membrane interior (Bocharov et al., 2008; Smith et al., 1996). To test the protonation state of the glutamate, we also performed a simulation in which the residue was deprotonated. In this simulation, the glutamate side chain established strong interactions with charged lipid head groups and was exposed to water. As a result, the transmembrane helices were distorted near the mutated residue (kinks developed) and they adopted a smaller angle with the membrane normal compared to the 20° angle for a wild-type helix.

The d_{int} plotted in figures is defined as distance between the centers of mass of the heavy backbone atoms of the Gly or Ala residues of the GxxxG-like motifs of each transmembrane helix.

Simulations of the TM–JM-A and TM–JM–KD constructs

Simulations of the TM–JM-A and TM–JM–KD constructs included residues 615–663 and residues 615–995, respectively. Initial structures of the TM–JM-A dimer included an antiparallel dimer of the JM-A α helices, as suggested by NMR studies (Jura et al., 2009). A double mutation (M626L and M644I) in the transmembrane helix was used for all but one TM–JM-A simulation in the DMPC membrane in order to be consistent with the NMR experiment (see the companion manuscript). (In the simulation illustrated in "DMPC 1" in Figure 5, the wild-type sequence was used.) We here did not observe significant a difference between the simulations of the double mutant and the wild type.

For the simulations of the TM–JM–KD inactive dimer, the TM–JM-A portion was taken from a conformation near the end of a simulation of a TM–JM-A monomer. The kinase dimer was taken from PDB entry 3GT8 (Jura et al., 2009). For the simulation of the active dimer, the TM–JM-A conformation was taken from the end of a TM–JM-A dimer simulation and the JM-A–KD conformation was taken from the asymmetric active dimer structure in PDB entry 2GS6 (Zhang et al., 2006).

Simulations of the extracellular modules

The dimer simulations of the extracellular domains (residues 1–614) were started from the structures of PDB entry 1IVO (Ogiso et al., 2002). The simulations of the extracellular domains in the tethered conformation were started from PDB entry 1NQL (Ferguson et al., 2003).

At the time of the simulation, most of domain IV was not resolved in the two available crystal structures of the extracellular dimer (Ogiso et al., 2002; Garrett et al., 2002). We thus modeled the conformation of the domain IVs by incorporating the conformation of domain III and IV from the tethered structure. More recently, domain IV was resolved in the context of the extracellular dimer (PDB entry 3NJP; Lu et al., 2010). The new structure is highly consistent

with the model we constructed. (The Cα RMSD between the structure and our model, computed for the whole EGFR extracellular dimer, is 1.2 Å).

The 1-ligand extracellular dimer was simulated starting from the same structure that was used for the 2-ligand dimer; the only difference was that either one or the other bound EGF molecule was removed from the structure before simulation. The simulations of the ligand-free extracellular dimers were started from the end conformation of one such simulation of the 1-ligand extracellular dimers (the remaining EGF molecule was removed before the simulations).

Fitting to EGFR activity-density measurements

The slope of the EGF-stimulated (EGFR + EGF) activity linear to EGFR density, which is $a \cong 24\ \mu m^2$ from linear fitting, corresponds to the activity of EGFR active dimer. Assuming the activity of a ligand-free active dimer is the same as that of an EGF-bound dimer, ligand-free activity then is $a\mathrm{D}(\mathbf{s}, k_d)$, where D denotes the density of active EGFR, k_d is the disassociation constant of EGFR active dimer without EGF, $\mathbf{s}$ is the surface density of EGFR, and $\mathrm{D} = (k_d + 4\mathbf{s} - (k_d^2 + 8s \cdot k_d)^{1/2})/8$. Fitting $24\mathrm{D}(\mathbf{s}, k_d)$ to the EGFR–EGF data produces $k_d = 1480\ \mu m^{-2}$ and a poor match with the raw data (Figure S4E). An improved fitting can be produced with the assumption of a k_d decreasing with density, which implies that the active dimers are favored at high EGFR density besides the effect of higher concentration. In this fitting, $k_d \Rightarrow k_0/\exp(\mathbf{s}^2/\mathbf{s}_0^2)$ was devised to reflect a decreasing dissociation constant. The fitting yielded $k_0 = 3301\ \mu m^{-2}$, $\mathbf{s}_0 = 2150\ \mu m^{-2}$ (Figure S4E). Note that $1\ \mu m^{-2}$ in surface density is approximately equivalent to 0.14 μM in concentration.

Supplemental References

Berendsen, H.J.C., Postma, J.P.M., DiNola, A., and Haak, J.R. (1984). Molecular dynamics with coupling to an external bath. J. Chem. Phys. *81*, 3684–3690.

Best, R.B., and Hummer, G. (2009). Optimized molecular dynamics force fields applied to the helix-coil transition of polypeptides. J. Phys. Chem. B *113*, 9004–15.

Cornell, W.D., Cieplak, P., Bayly, C.I., Gould, I.R., Merz, K.M. Jr., Ferguson, D.M. Spellmeyer, D.C., Fox, T., Caldwell, J.W., and Kollman, P.A. (1995). A second generation force field for the simulation of proteins, nucleic acids and organic molecules. J. Am. Chem. Soc. *117*, 5179–5197.

Jorgensen, W.L., Chandrasekhar, J., Madura, J.D., Impey, R.W., and Klein, M.L. (1983) Comparison of simple potential functions for simulating liquid water. J. Chem. Phys. *79*, 926–935.

Hornak, V., Abel, R., Okur, A., Strockbine, B., Roitberg, A., Simmerling, C. (2006). Comparison of multiple Amber force fields and development of improved protein backbone parameters. Proteins: Struct., Funct., Bioinf. *65*, 712–725.

Humphrey, W., Dalke, A., and Schulten, K. (1996). VMD: visual molecular dynamics. J. Mol. Graph. *14*, 33–38.

Klauda, J.B., Venable, R.M., Freites, J.A., O'Connor, J.W., Mondragon-Ramirez, C., Vorobyov, I., Tobias, D.J., MacKerell, A.D., and Pastor, R.W. (2010). Update of the CHARMM all-atom additive force field for lipids: validation on six lipid types. J. Phys. Chem. B *114*, 7830–7843.

Krautler, V., van Gunsteren, W.F., and Hunenberger, P.H. (2001). A fast SHAKE algorithm to solve distance constraint equations for small molecules in molecular dynamics simulations. J. Comput. Chem. *22*, 501–508.

Lindorff-Larsen, K., Piana, S., Palmo, K., Maragakis, P., Klepeis, J.L., Dror, R.O., and Shaw, D.E. (2010). Improved side-chain torsion potentials for the AMBER ff99SB protein force field. Proteins: Struct., Funct., Bioinf. *78*, 1950–1958.

MacKerell, A.D., Bashford, D., Bellott, M., Dunbrack, R.L., Evaseck, J.D., Field, M.J., Fischer, S., Gao, J., Guo, H., Ha, S., et al. (1998). All-atom empirical potential for molecular modeling and dynamics studies of proteins. J. Phys. Chem., B *102*, 3586–3617.

Piana, S., Lindorff-Larsen, K., and Shaw, D.E. (2011). How robust are protein folding simulations with respect to force field parameterization? Biophys. J. *100*, L47–L49.

Shan, Y., Klepeis, J.L., Eastwood, M.P., Dror, R.O., and Shaw, D.E. (2005). Gaussian split Ewald: a fast Ewald mesh method for molecular simulation. J. Chem. Phys. *122***,** 1–13.

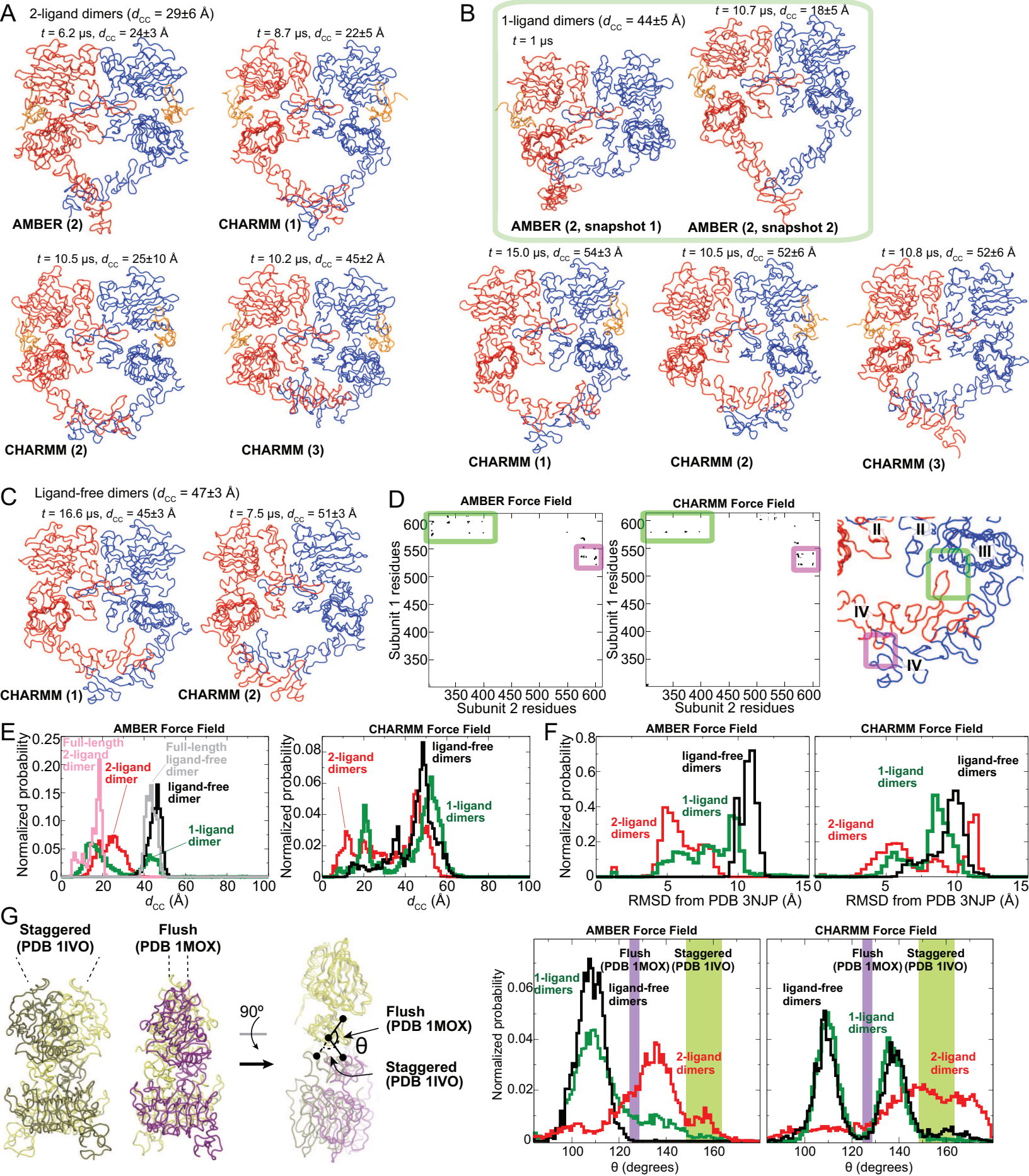
Smith, S.O., Smith, C.S., and Bormann, B.J. (1996). Strong hydrogen bonding interactions involving a buried glutamic acid in the transmembrane sequence of the neu/erbB-2 receptor. Nat. Struct. Biol. *3*, 252–258.

van Meer, G., Voelker, D.R., and Feigenson, G.W. (2008). Membrane lipids: where they are and how they behave. Nat. Rev. Mol. Cell Biol *9*, 112–124.

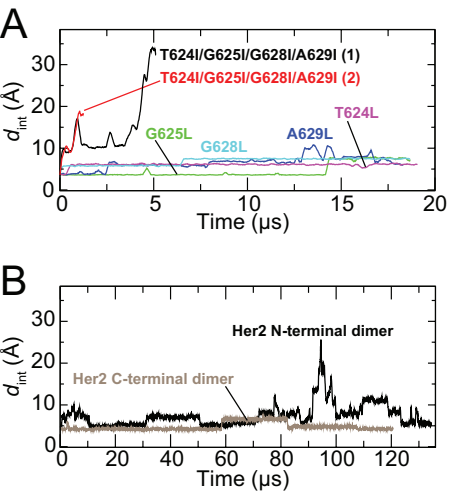
Tuckerman, M., Berne, B.J., and Martyna, G.J. (1992). Reversible multiple time scale molecular dynamics. J. Chem. Phys. *97*, 1990–2001.

Zachowski, A. (1993). Phospholipids in animal eukaryotic membranes: transverse asymmetry and movement. Biochem. J. *294*, 1–14.

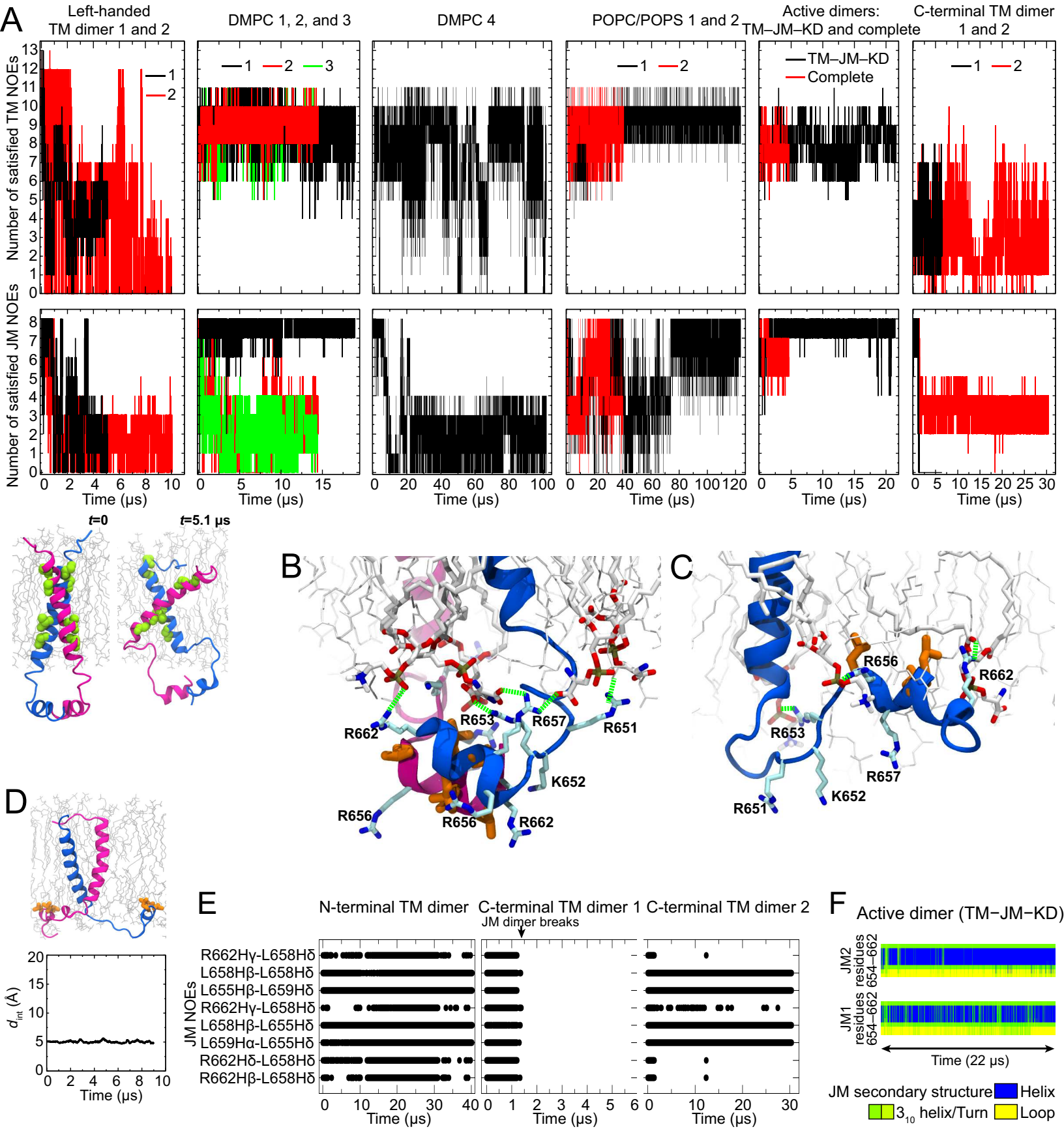
Supplemental Figure S1



Supplemental Figure S2



Supplemental Figure S3



Supplemental Figure S4