Supporting Information for:

Dimerization of the EphA1 Receptor Tyrosine Kinase Transmembrane Domain: Insights into the Mechanism of Receptor Activation

Matthieu Chavent, Alan P. Chetwynd, Phillip J. Stansfeld, Mark S. P. Sansom

Text: Details of methods of analysis of convergence of the PMF profile calculations:

Convergence and PMF profile calculations: We first divided our data from umbrella simulation into 750 bins, each 10 ns wide. We then calculated the PMF profile. This allows calculating the variance $\sigma^2(x)$ for the whole simulation with a reasonable approximation. We then grouped the data into bins of different width n creating, for each width, a dataset x_n with a variance $\sigma^2(x_n)$. We then calculated the statistical inefficiency $s(n)$:

$$s(n) = \frac{n\sigma^2(x_n)}{\sigma^2(x)}$$

If the data converge, the statistical inefficiency tends to reach a plateau. The value of the plateau corresponds to the correlation time. We performed this calculation exploring the simulation dataset backward, beginning by the end using the approach of Yang *et al.* [54]. Fig. S1A shows that the last 10% of the simulation the $s(n)$ reached a plateau. This allows defining the time for which the simulation reached a convergence: around 90% of the simulation and the correlation time for this part: c.a. 75 ns. Thus, for the last 10% of the simulation, we can consider that bins with at least a width of 75 ns divide this simulation in independent datasets. We preferred to choose slightly larger bins with a width of 100 ns. We produced 8 independent PMF profiles based on these data (see Fig. S1B). These profiles present different shapes which suggest that the datasets are really independent. We then averaged these curves and calculated an estimate of the statistical error to create Fig. 1A.

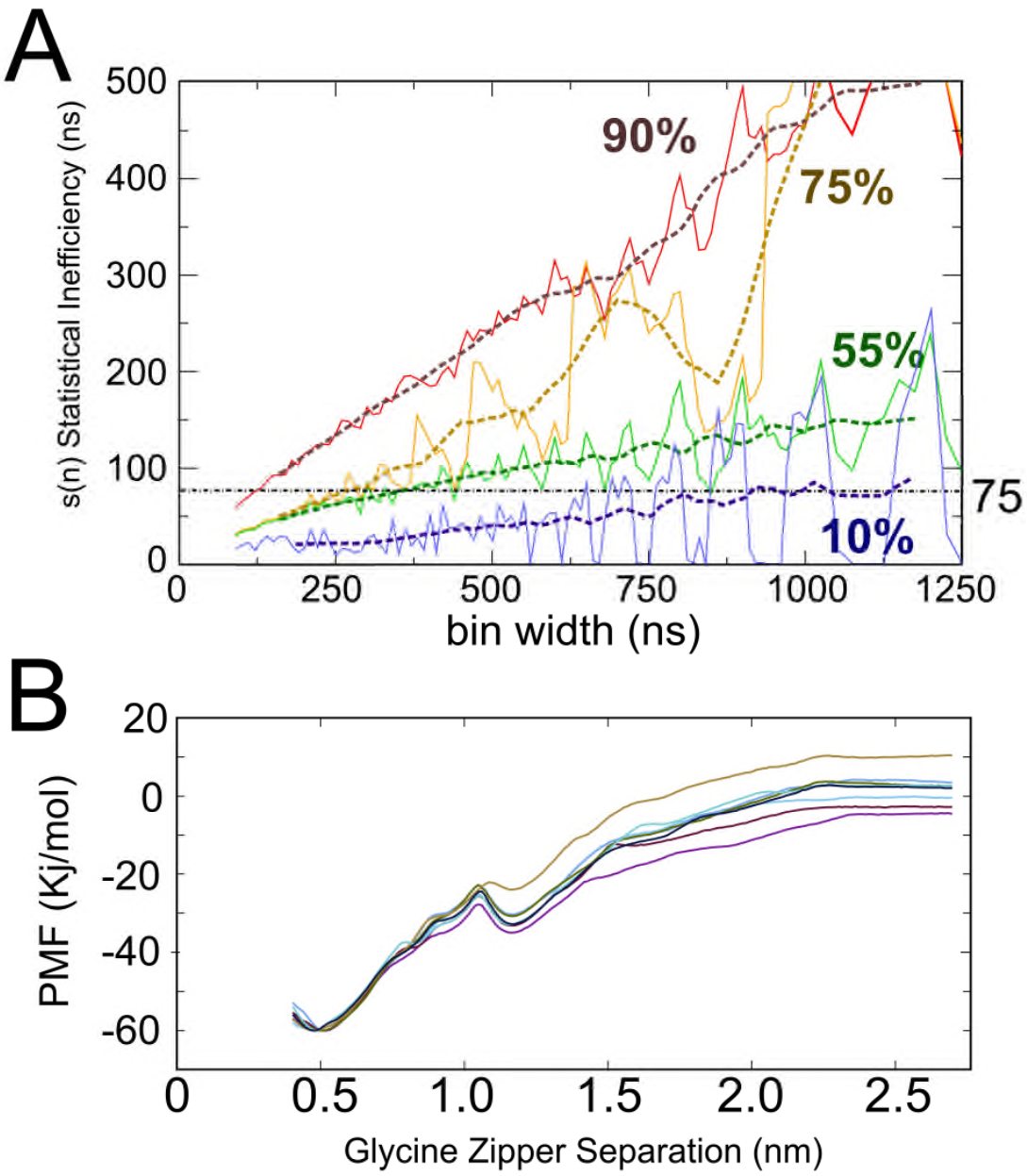

Figure S1:

(A) Statistical inefficiency, *s(n)*, calculated for various degrees of penetration (10% to 90%) into the dataset (see Supporting Information Text above). Dashed lines are averaged values of plain lines. We estimate that the last 10% of the simulation is converged and that the correlation time is around 75 ns.). We used this last 10% to create 8 independent PMF profiles (B). These profiles were used to create Fig. 1A.

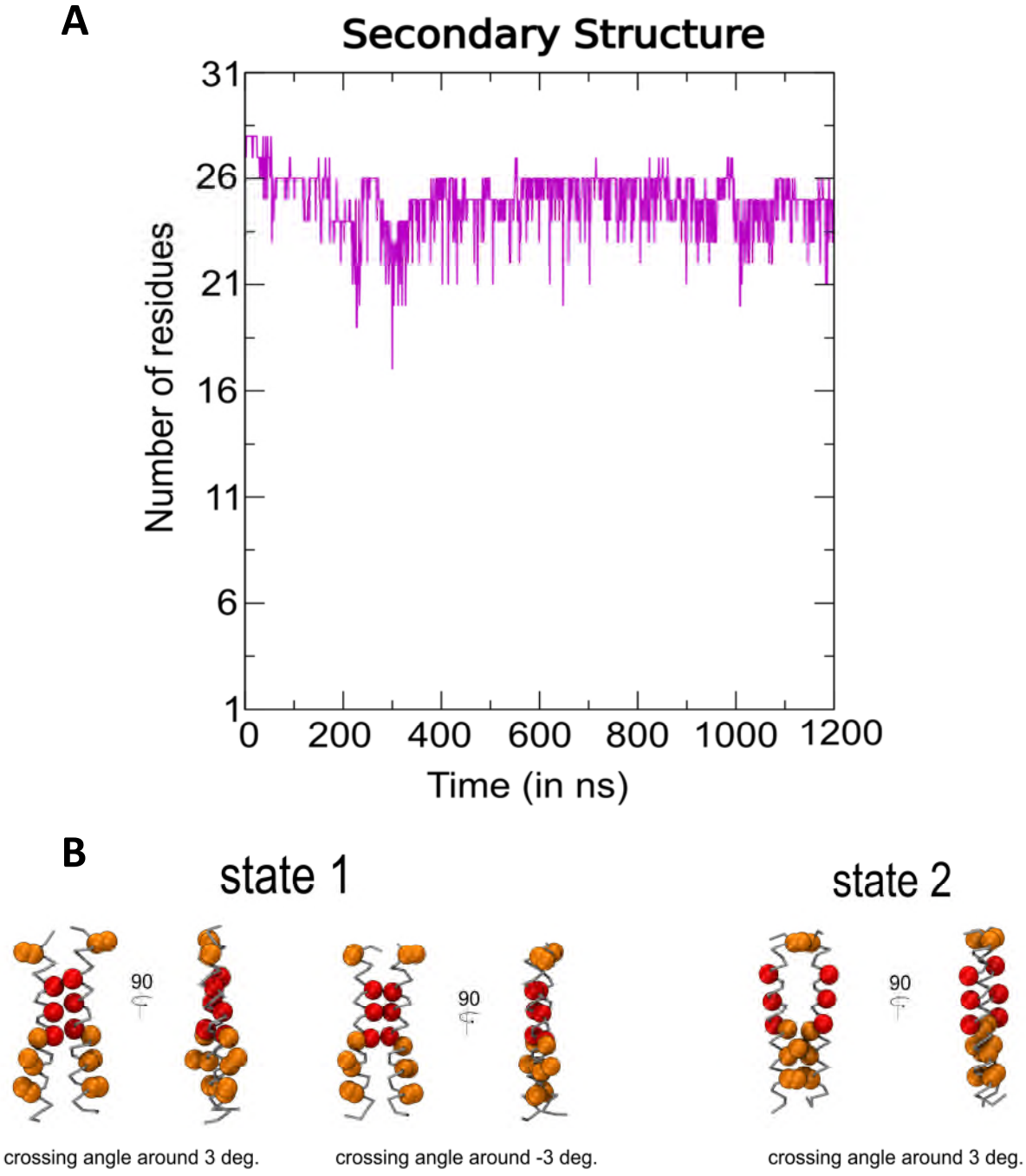

Figure S2:

(A) Secondary structure conservation in function of the time. This refers to the extended (1.2 µs) simulation of a single EphA2 TM helix in a DPPC bilayer discussed in the main text. (B) Representative structures. Two representative structures for state 1 (for a crossing angle around -3° and +3°) and one representative structure for state 2 (for a crossing angle around +3°). Spheres depict residues involved in the two states. In red, residues forming the glycine zipper, in orange additional residues involved in the dimer interaction for state 2.

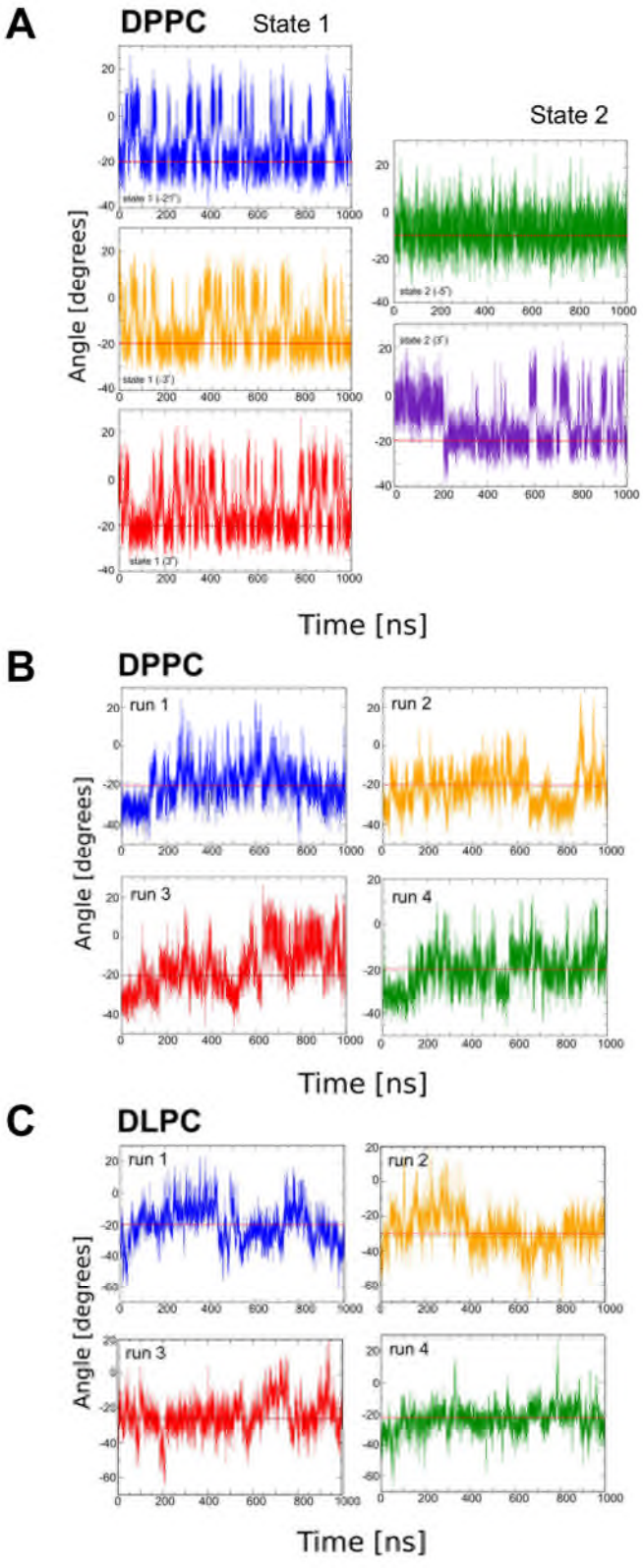

Figure S3:

Helix crossing angle as a function of time, based on simulations starting from PMF derived-structures (**A**), or from based on simulations starting from the NMR structure in a DPPC (**B**) or DLPC (**C**) bilayer.

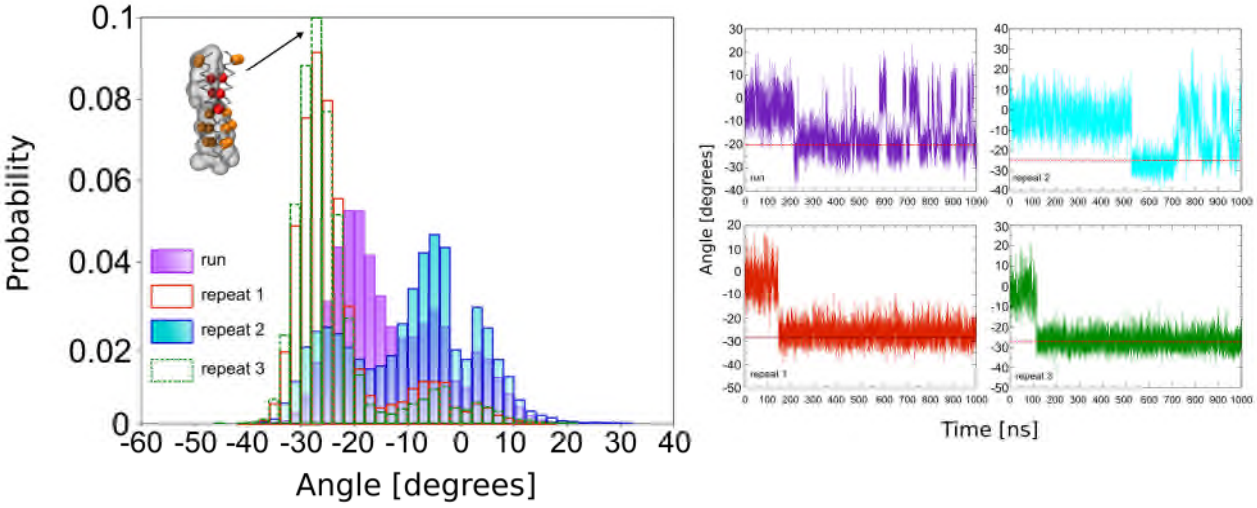

Figure S4:

Helix crossing angle distribution based on the 1 µs trajectories starting with a structure in state 2 with a crossing angle of ~3°. On the right, helix crossing angle in function of the time. In red, residues forming the glycine zipper, in orange additional residues involved in the dimer interaction for state 2.

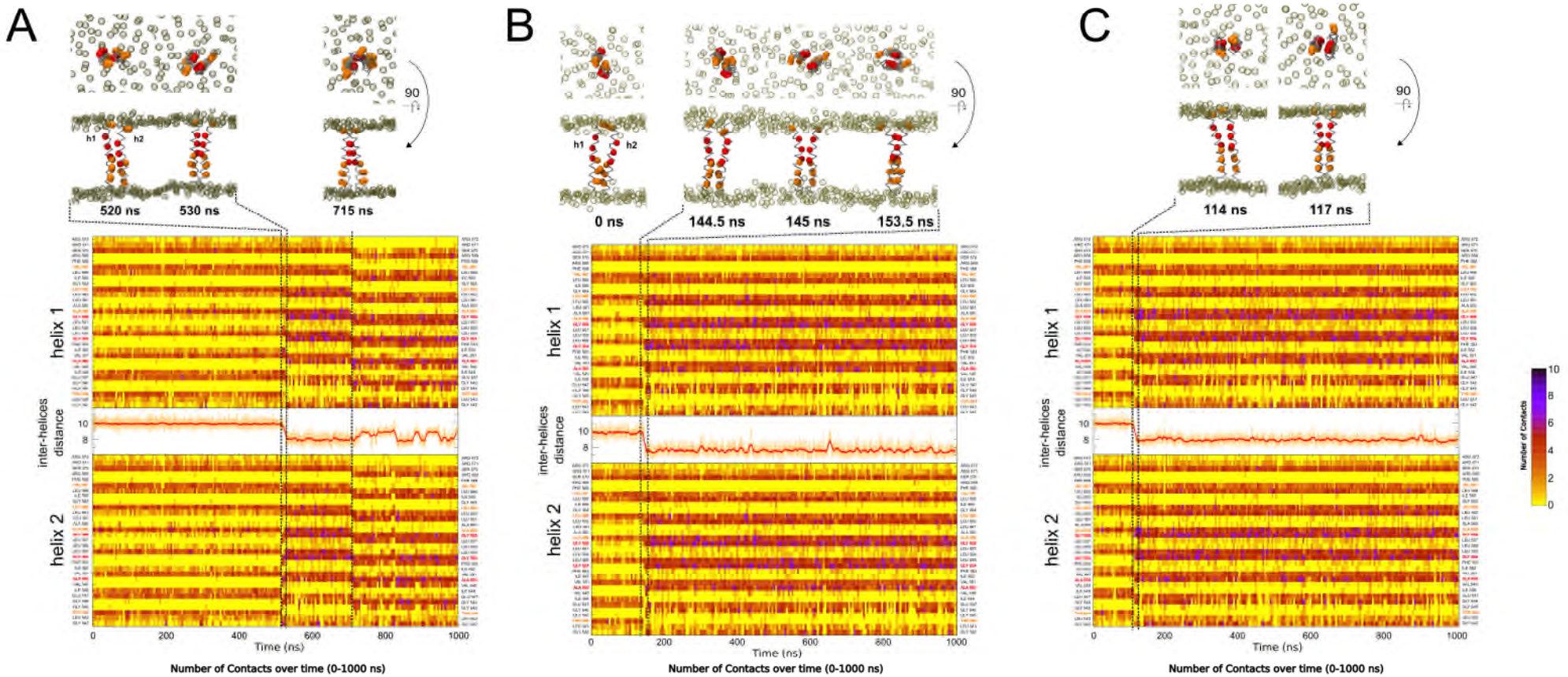

Figure S5:

The number of contacts at the dimer interface as a function of time for repeat simulations presented in Fig. S4. In the graphs in the centre, the red curves show the averaged inter-helix distances (in Å) with the raw values in orange. At the top, snapshots are shown of the dimers at slected points along each trajectory. Red spheres depict residues forming the glycine zipper and orange spheres represent residues involved in the dimer interaction for state. (**A**) corresponds to repeat 2, (**B**) corresponds to repeat 1, and (**C**) corresponds to repeat 3.

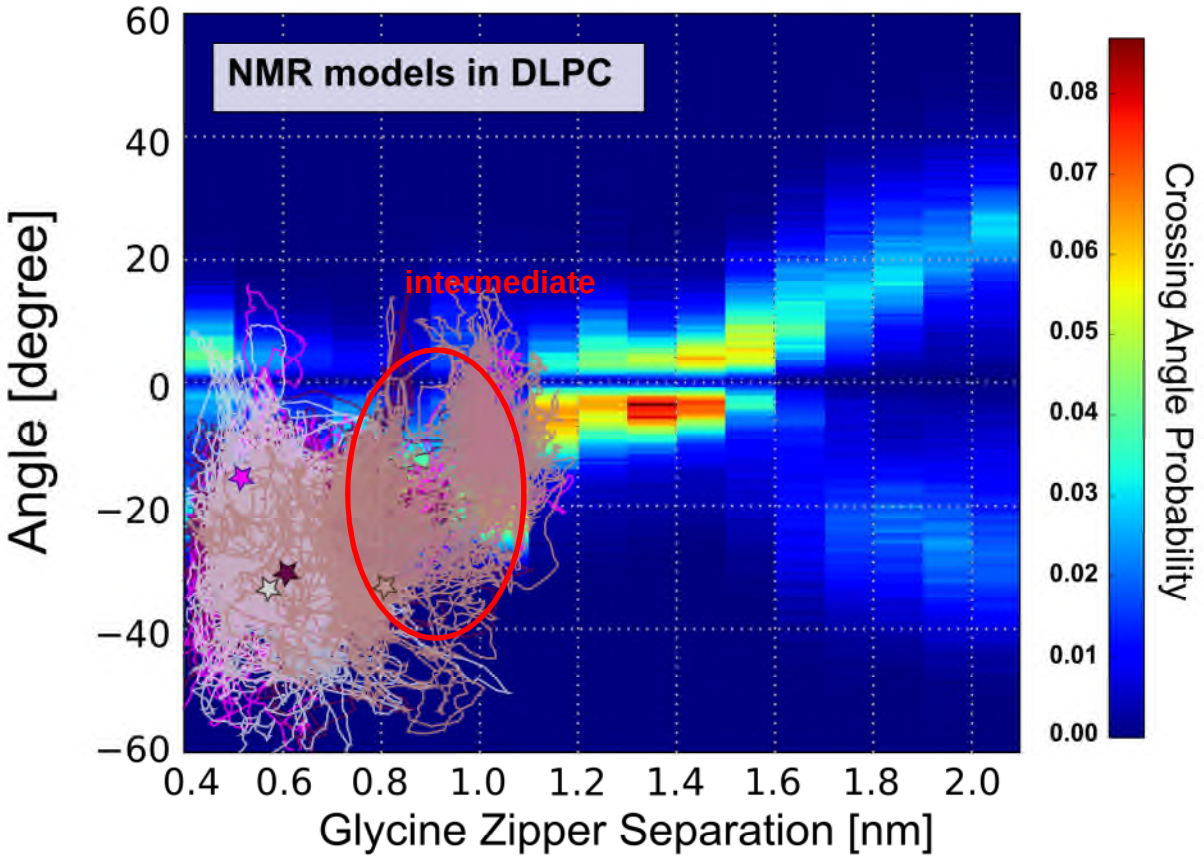

Figure S6:

Trajectories of the four CG simulations based on the NMR structure in a DLPC bilayer superimposed on the CG conformational landscape presented in Figure 6A. In plotting these trajectories we sample averaged positions (based on 10 steps) every 0.5 ns. Stars depict the final configuration of the dimer for each trajectory. Trajectory colours correspond to run numbers presented in Fig. S3: light brown: run 1, gray: run 2, mauve: run3, and dark purple: run 4. The red ellipse indicates the area corresponding to the intermediate state (Fig. 5).

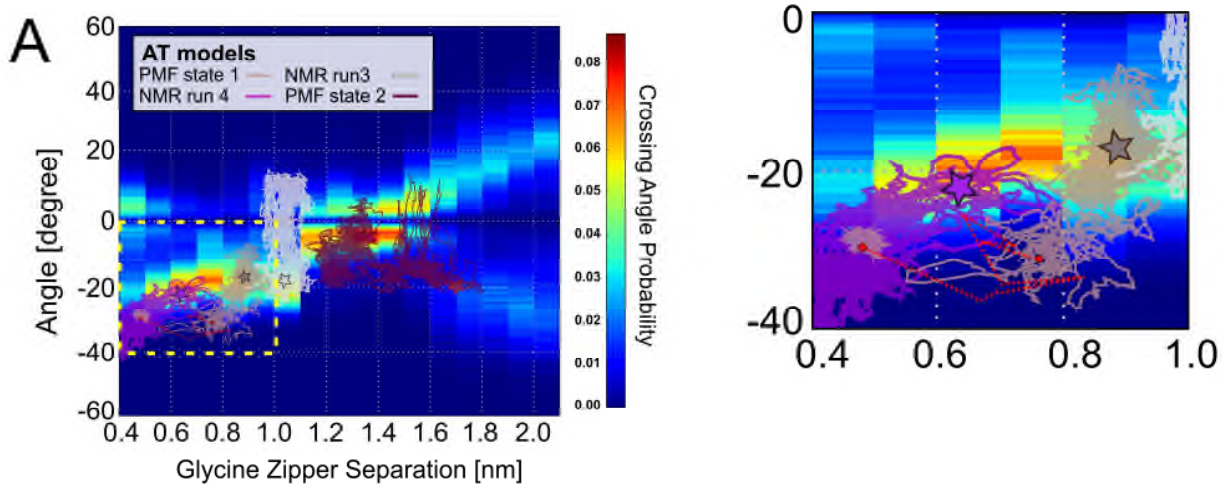

Figure S7:

Trajectories for the four AT simulations superimposed on the CG conformational landscape presented in Fig. 5A In plotting these trajectories we sample averaged positions (based on 10 steps) every 0.2 ns. The red dashed line displays the equilibration trajectory.

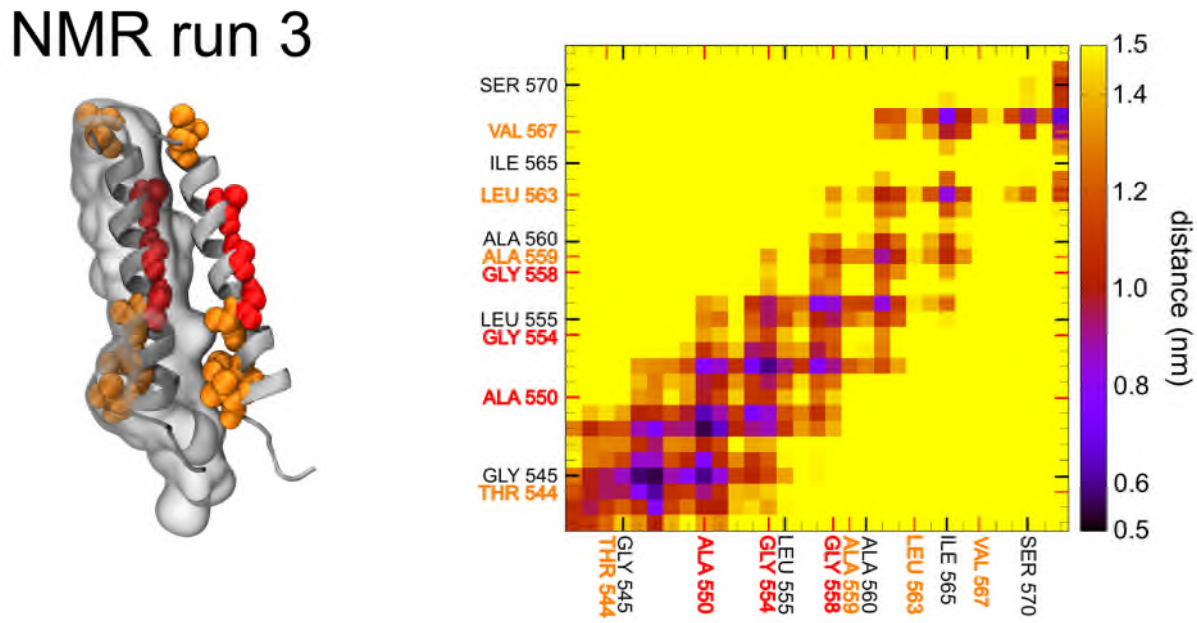

Figure S8:

Contact maps and most representative structures of the helix dimer based on the last 10 ns for AT simulations starting with NMR structure. For one helix, the surface is represented as a transparent envelope. The main residues involved for State 1 and for State 2 are depicted in red and orange respectively, as previously described.

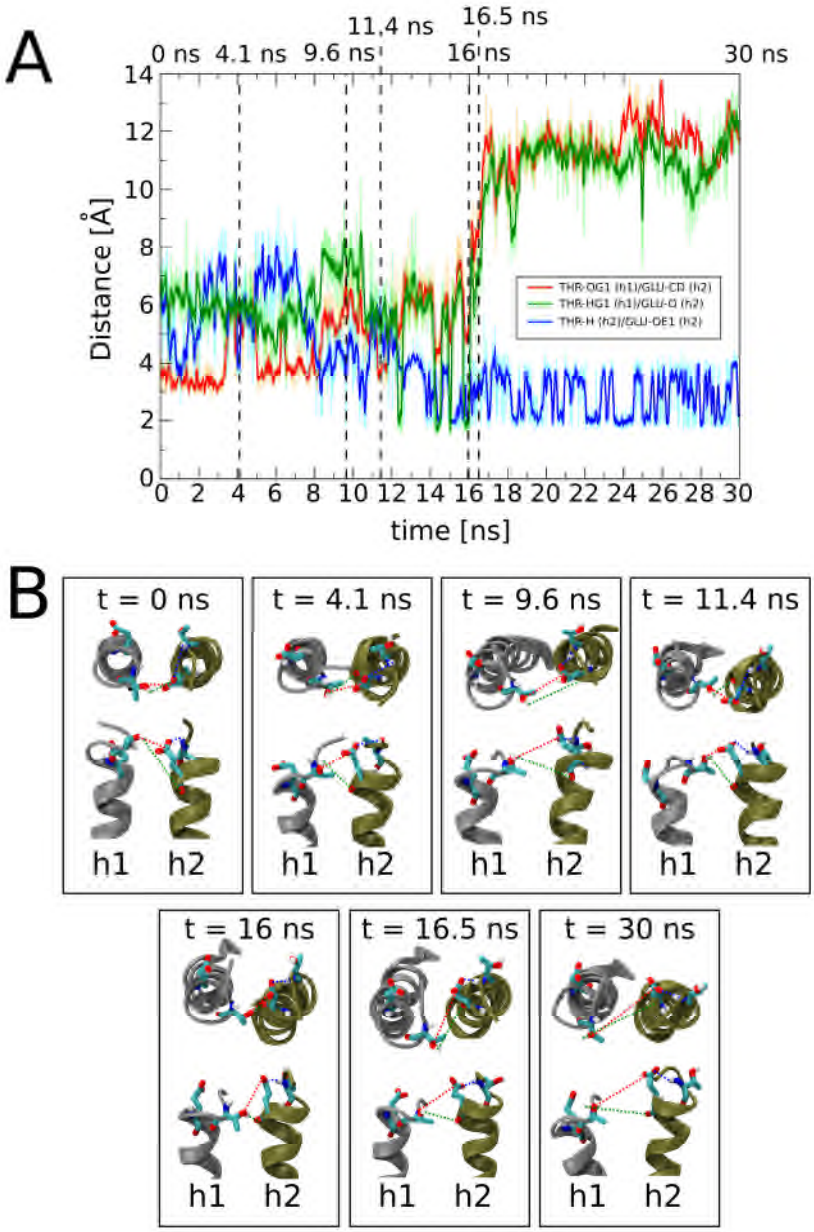

Figure S9:

(A) Diagram visualizing the evolution of inter-helix and intra-helix distances between E^{547} and T^{544} around the partial dissociation of the dimer (t ~16 ns) for *State 2* AT simulation. (B) The corresponding interactions visualized at the dimer interface.

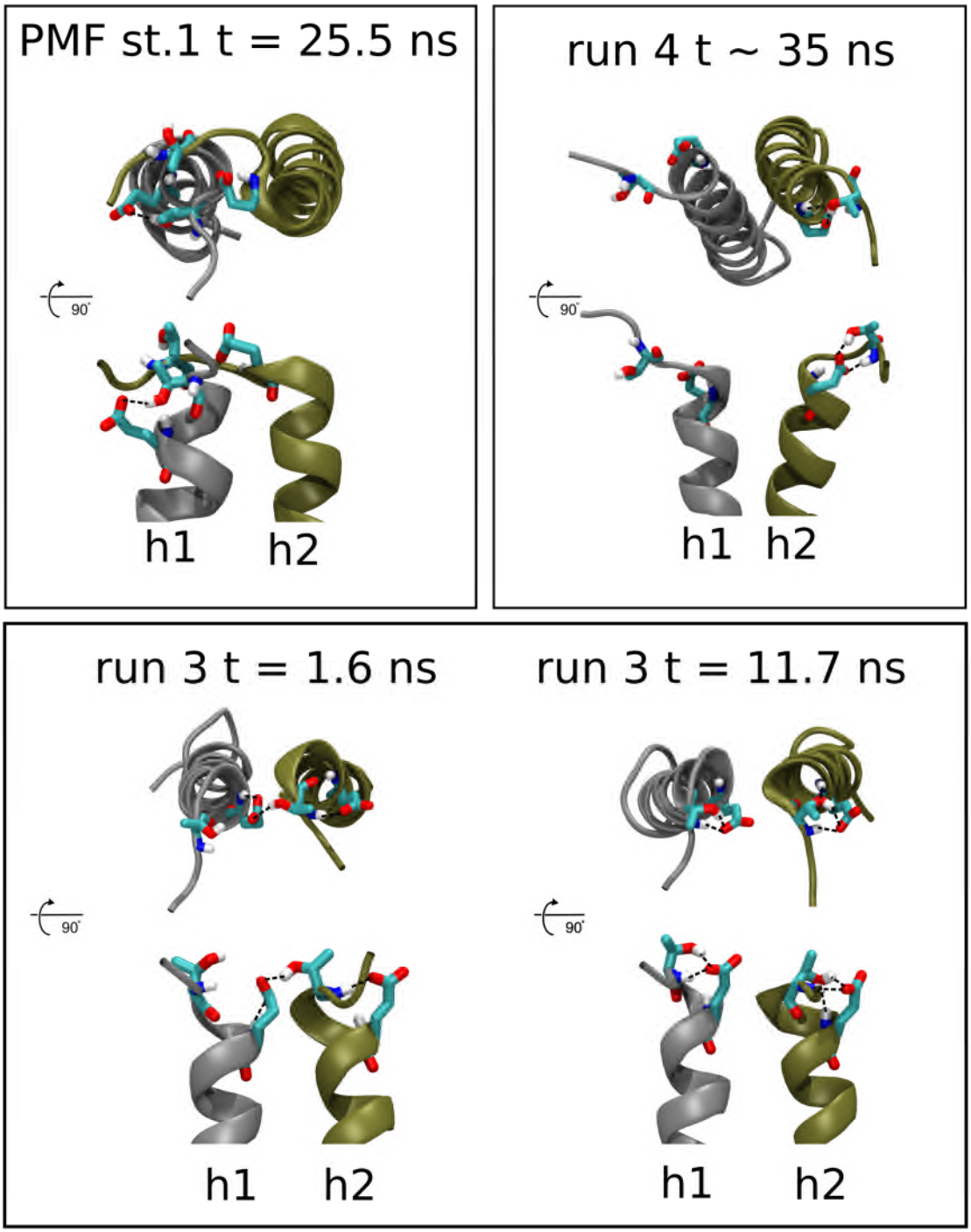

Figure S10:

Visualization of H-bonds from the different AT simulations. See also Table S1.

Table S1: Residue pairs for hydrogen bonds created during AT simulations.

helix 1	helix 2	helix 1 intra	helix 2 intra	time associated inter	time associated intra h1	time associated intra h2
PMF state 1 (h-bonds: 63 inter, 11 intra helix 1, 6 intra helix 2)*						
Gly 546 - Main	**Thr 544 - Main**	#	**Glu 547 - Main**	44.22 %	#	2.29 %
Gly 545 - Main	**Glu 547 - Side**	#	#	42.60 %	#	#
Glu 547 - Main	Leu 543 - Main	#	#	42.34 %	#	#
Arg 569 - Main	Arg 569 - Side	#	#	40.64 %	#	#
Glu 547 - Side	Leu 543 - Main	**Thr 544 - Side**	#	37.16 %	79.66 %	#
Thr 544 - Main	#	Ile 548 - Main	#	#	65.80 %	#
PMF state 2 (h-bonds: 42 inter, 7 intra helix 1, 7 intra helix 2)*						
Ser 570 - Side	Arg 571 - Side	#	#	15.79 %	#	#
Thr 544 - Side	**Glu 547 - Side**	#	**Thr 544 - Main**	8.91 %	#	52.27%
Leu 566 - Main	Arg 571 - Side	#	#	8.42 %	#	#
Arg 569 - Main	Arg 571 - Side	#	#	4.71 %	#	#
Val 567 - Main	Arg 571 - Side	#	#	2.76 %	#	#
Arg 572 - Side	Arg 571 - Side	#	#	2.56 %	#	#
Gly 542 - Main	**Glu 547 - Side**	#	**Thr 544 - Main**	2.35 %	#	52.27%
Thr 544 - Main	**Glu 547 - Side**	Val 549 - Main	**Thr 544 - Main**	2.20 %	54.65 %	52.27%
Thr 544 - Main	#	Ile 548 - Main	#	#	3.54 %	#
#	**Thr 544 - Main**	#	Val 549 - Main	#	#	66.67 %
#	**Thr 544 - Main**	#	Ile 548 - Main	#	#	49.17 %
NMR run 4 (h-bonds: 30 inter, 8 intra helix 1, 8 intra helix 2)*						
Arg 569 - Main	Arg 569 - Side	#	#	18.64 %	#	#
Ser 570 - Side	Arg 569 - Side	#	#	11.50 %	#	#
Phe 568 - Main	Arg 569 - Side	#	#	10.74 %	#	#
Ile 565 - Main	Arg 569 - Side	#	#	7.68 %	#	#
Phe 553 - Side	Ala 550 - Main	#	#	3.78 %	#	#
Ser 570 - Main	Arg 569 - Side	#	#	2.76 %	#	#
Phe 553 - Side	Val 551 - Main	#	#	2.42 %	#	#
#	**Thr 544 - Side**	#	**Glu 547 - Side**	#	#	88.36 %
#	**Thr 544 - Main**	#	**Glu 547 - Side**	#	#	77.68 %
#	**Thr 544 - Main**	#	Ile 548 - Main	#	#	29.51 %
#	**Thr 544 - Main**	#	Gly 546 - Main	#	#	2.52 %
NMR run 3 (h-bonds: 45 inter, 7 intra helix 1, 9 intra helix 2)*						
Glu 547 - Side	Gly 545 - Main	***Thr 544 - Main***		27.93 %	48.53 %	#
Glu 547 - Side	***Thr 544 - Side***	***Thr 544 - Side***	***Glu 547 - Side***	1.72 %	93.4 %	86.62 %
#	***Thr 544 - Main***	#	***Glu 547 - Side***	#	#	75.16 %
#	***Thr 544 - Main***	#	Ile 548 - Main	#	#	72.97 %
#	***Thr 544 - Side***	#	***Glu 547 - Main***	#	#	9.00 %
#	***Thr 544 - Main***	#	***Glu 547 - Main***	#	#	3.10 %
Thr 544 - Main	#	Ile 548 - Main	#	#	77.66 %	#

* In this table, inter-helix H-bonds are presented, as well as intra-helix H-bonds formed by Thr 544 with other residues. Only H-bonds with a life time more than 2% of the whole trajectory time are included in this table. H-bonds were identified using a cut-off distance of 3.5 Å and angle of 30° via the VMD plugin Hbonds (http://www.ks.uiuc.edu/Research/vmd/plugins/hbonds/)

Supporting Information - Movies

Movie S1:

Movie illustrating the concerted rotations needed to pass from *State 2* to the intermediary structure in which the glycine zipper in addition to the C-terminal segments of the helices are involved. Movie based on trajectory for repeat 2 presented in Fig. S4B. A high resolution version is available at: http://www.matthieuchavent.com/videos/EphA1_dimer_sup_movie1.mp4

Movie S2:

Movie illustrating the decoupled rotations to pass from *State 2* to *State 1*. Movie based on trajectory presented in Fig. 4. A high resolution version is available at: http://www.matthieuchavent.com/videos/EphA1_dimer_sup_movie2.mp4.