Supplementary Information for:

Structure-Based Simulations Reveal Concerted Dynamics of GPCR Activation

Nicholas Leioatts, Pooja Suresh, Tod D. Romo, & Alan Grossfield*

*alan_grossfield@urmc.rochester.edu

Department of Biochemistry & Biophysics, University of Rochester Medical Center, Rochester, NY, 14642, USA

S1 Supplemental Methods

S1.1 Drawing bonds to represent contact space

Bonds were drawn on protein structures to represent the unique contacts that changed during transition simulations (see Figures 7, S7, S8, and S9). The bonds were only drawn for the subset of contacts that varied most within a given principal component (specifically, those accounting for the first 50% of the variance), and the bond thickness was proportional to that contact's contribution to the principal component.

S1.2 Defining Non-monotonic Transitions

In Figure S6 we map the non-monotonic or "backtracking" contacts back onto the protein structures. In this analysis a contact was considered to backtrack only if it made a backward transition of 0.2 or more (defined using Equation 2). Many more contacts made these transitions in β_2AR than in rhodopsin; this may be the result of more topological frustration in the former protein.

S1.3 Differences in contact-space PCA for comparing proteins

For the β_2AR to rhodopsin comparison in Section 3.4.1 one change was made to the transition contacts matrix for the contact-based PCA (no change was needed for the Cartesian PCA). Specifically, we used matrices of all possible residue pairs within the transmembrane region (Table ii) to define the transition matrix (not just the unique contacts in a different state). This was simply because the set of unique contacts for β_2AR and rhodopsin differ (see Table Si). This created a much larger matrix, so only 30 randomly chosen samples were used.

Supplementary Tables

Protein	Contacts Broken	Contacts Formed	Total Contacts
β_2AR (all contacts)	99	89	188
β_2AR (activation crucial)	20	22	42
β_2AR (deactivation crucial)	21	21	42
rhodopsin (all contacts)	132	108	240
rhodopsin (activation crucial)	21	23	44
rhodopsin (deactivation crucial)	27	22	49

Table Si: Total number of contacts formed and broken during transition, and the number of contacts contributing the most (up to 50%) to the transition of β_2AR and rhodopsin.

First dataset	Second dataset	P-value Activation	P-value Deactivation
cartesian results			
cross	β_2AR	$2.273\text{x}10^{-87}$	$1.325\text{x}10^{-92}$
cross	rhodopsin	$5.014\text{x}10^{-182}$	$2.191\text{x}10^{-222}$
β_2AR	rhodopsin	$1.105\text{x}10^{-68}$	$1.494\text{x}10^{-58}$
contact-space results			
cross	β_2AR	$2.683\text{x}10^{-36}$	$8.479\text{x}10^{-34}$
cross	rhodopsin	$7.225\text{x}10^{-57}$	$4.286\text{x}10^{-53}$
β_2AR	rhodopsin	$3.035\text{x}10^{-38}$	$1.880\text{x}10^{-30}$

Table Sii: P-values between paths taken by β_2AR and rhodopsin. These two-tailed p-values test the hypothesis that the ensemble of paths taken by β_2AR is statistically different from those taken by rhodopsin. The data here compares two columns using the results plotted in Figure 8.

First dataset	Second dataset	P-value Rhodopsin	P-value β_2AR
cartesian results			
cross	activation	$1.202\text{x}10^{-176}$	$1.201\text{x}10^{-96}$
cross	deactivation	$7.276\text{x}10^{-175}$	$1.404\text{x}10^{-97}$
activation	deactivation	$7.441\text{x}10^{-2}$	$1.149\text{x}10^{-1}$
contact-space results			
cross	activation	$1.789\text{x}10^{-117}$	$5.615\text{x}10^{-117}$
cross	deactivation	$5.250\text{x}10^{-135}$	$1.473\text{x}10^{-132}$
activation	deactivation	$4.055\text{x}10^{-3}$	$1.436\text{x}10^{-29}$

Table Siii: P-values between activation and deactivation paths. These two-tailed p-values test the hypothesis that the ensemble of paths taken during activation is statistically different from those taken during deactivation. The data here compares two columns using the results plotted in Figure 9.

Supplementary Movies

Supplementary Movie S1 This movie shows the average displacement of simulations along the first two principal components (also called right singular vectors). Proteins are shown first activating, then deactivating. The transmembrane domain is shown from the cytoplasmic side.

Supplementary Movie S2 This video shows the time-progression of contacts in example trajectories for each system. Proteins are shown from the cytoplasmic side in yellow. The unique contacts are shown as sticks between c_α atoms. Cyan bonds are those present only in the initial structure. Purple bonds are those present only in the ending structure. These bonds are only drawn when the contacts are formed (see 2.2.2 for details). Note that the presence of unique end-state contacts at the beginning of each movie is due to formation of these contacts during our equilibration period.

Supplementary Figures

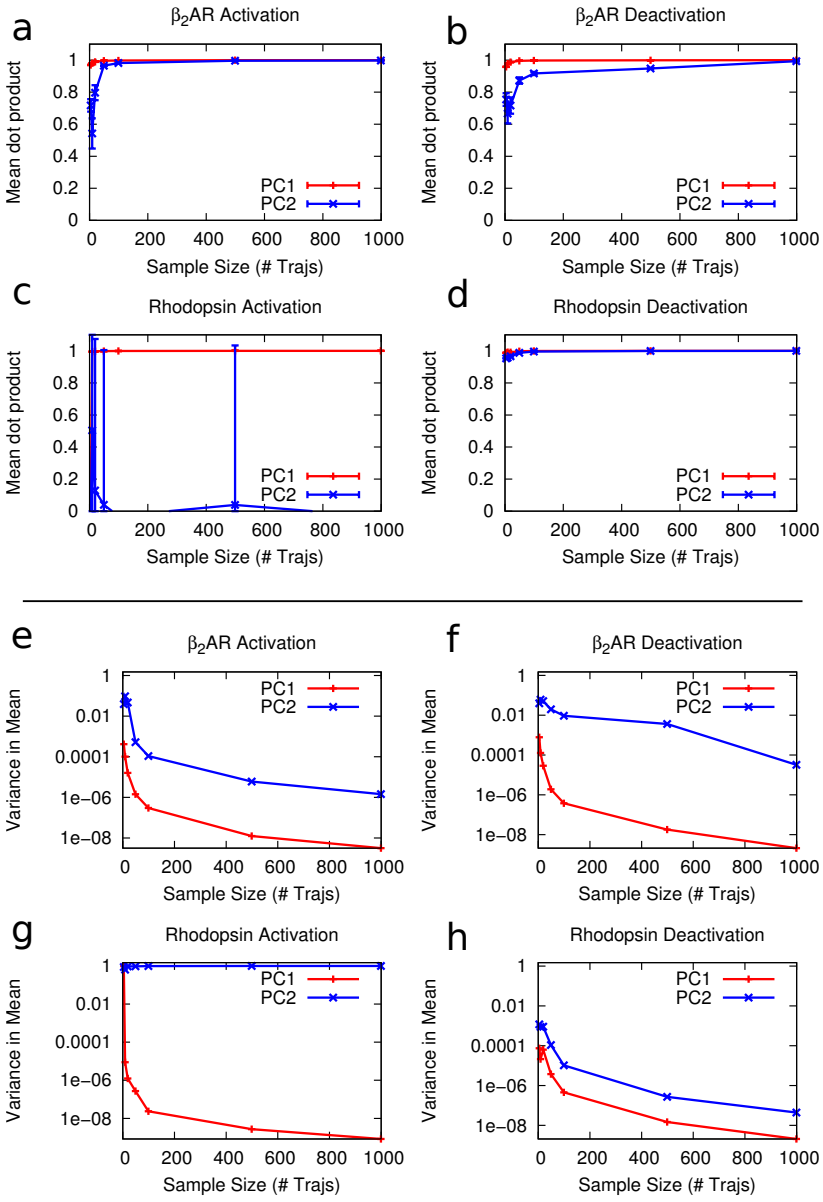

Figure S1: Convergence of the first two principal components in cartesian space. (a-d) Mean absolute dot product of the first two principal components. The mean dot product (y-axis) was calculated from pairs of modes taken from our bootstrapped data set. This was plotted as a function of the number of trajectories included in the bootstrap ensemble (x-axis). Calculations were between PC1 eigenvector pairs (blue lines) or PC2 eigenvector pairs (red lines) from two independent bootstrap samples. The mean was taken over 30 independent bootstrap runs. Error bars show variance in the mean. (e-h) The variances in PC1 and PC2 (error from panels a-d) are replotted on a log scale (y-axis) as a function of ensemble size (x-axis).

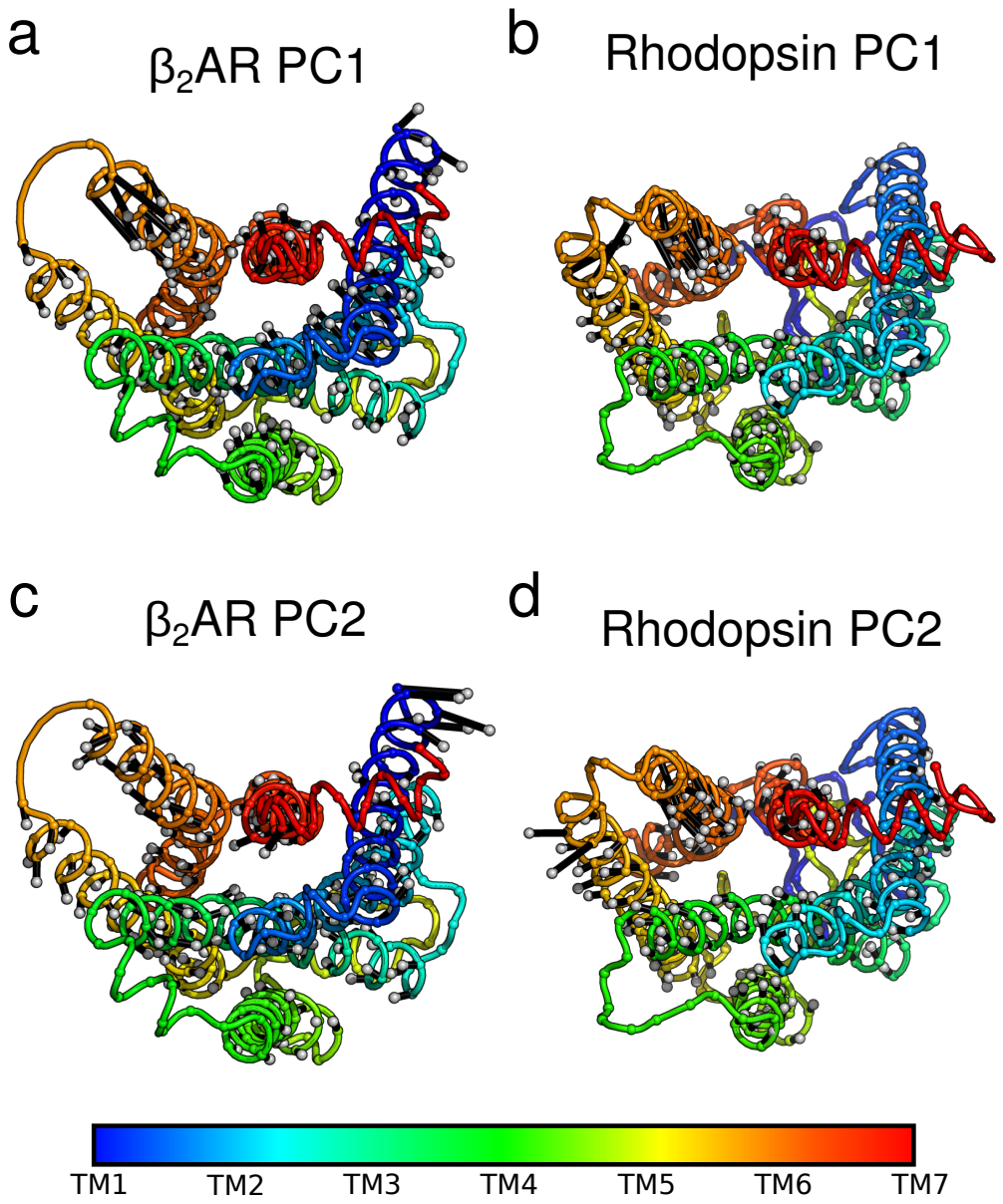

Figure S2: Direction of the most concerted motions in deactivation. Principal components (Section 3.2.1) are plotted as black arrows extending from c_α's on cartoons of the active protein. Arrow length is proportional to each residue's contribution to the principal component. Data is shown for (a) PC1 of β_2AR, (b) PC1 of rhodopsin, (c) PC2 of β_2AR, and (d) PC2 of rhodopsin.

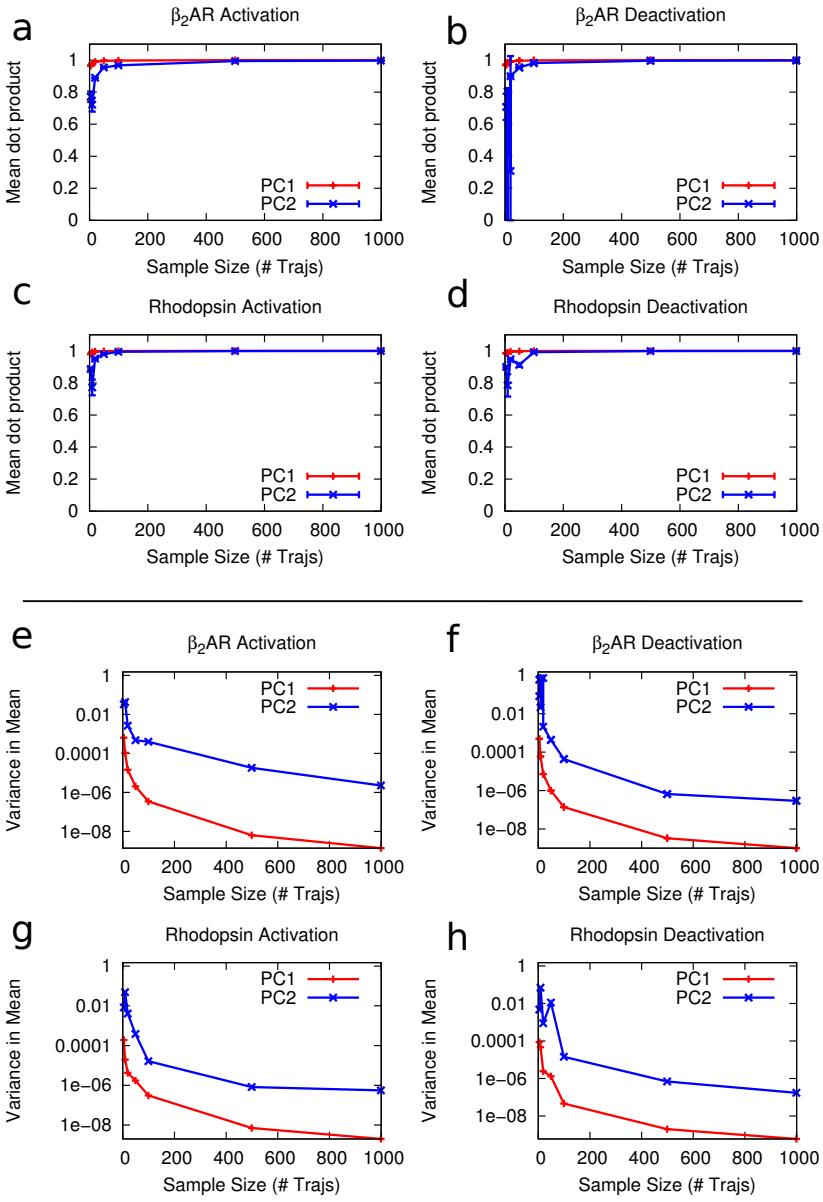

Figure S3: Convergence of the first two principal components in contact space. (a-d) Mean dot absolute product of the first two principal components in contact space. The mean dot product (y-axis) was calculated from pairs of modes taken from our bootstrapped data set. This was plotted as a function of the number of trajectories included in the bootstrap ensemble (x-axis). Calculations were between PC1 eigenvector pairs (blue lines) or PC2 eigenvector pairs (red lines) from two independent bootstrap samples. The mean was taken over 30 independent bootstrap runs. Error bars show variance in the mean. (e-h) The variance in PC1 and PC2 (error from panels a-d) is replotted on a log scale (y-axis) as a function of ensemble size (x-axis).

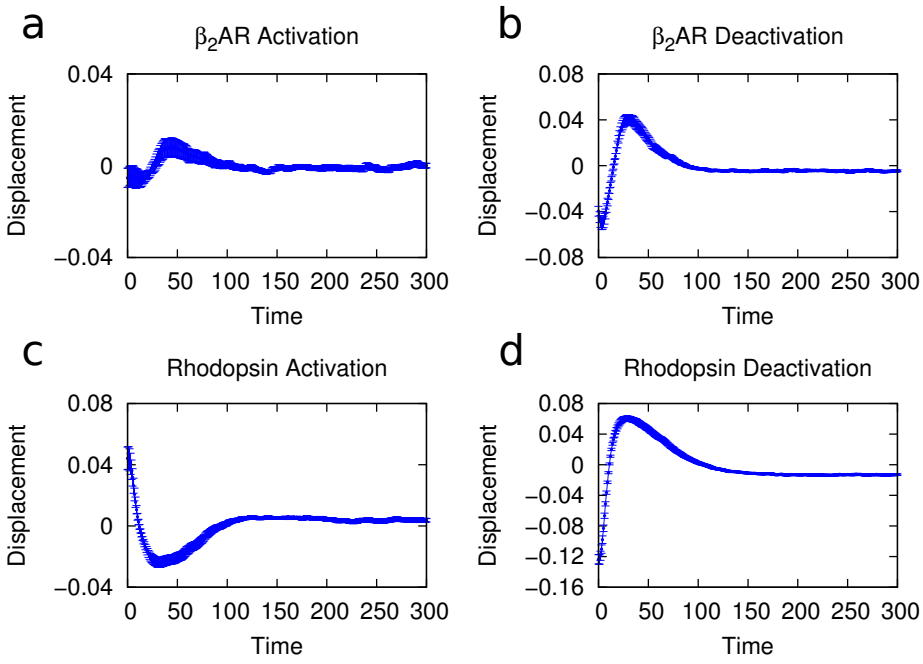

Figure S4: Displacement of PC2 in contact-space. The average projection of each simulation onto PC2 is shown over time, plotted for each ensemble. The error bars represent the standard deviation from the mean displacement across the entire 1,000-trajectory ensemble. Data shown is from the contact-based PCA (Section 3.3.1). Data is shown for (a) β_2AR activation, (b) β_2AR deactivation, (c) rhodopsin activation, and (d) rhodopsin deactivation.

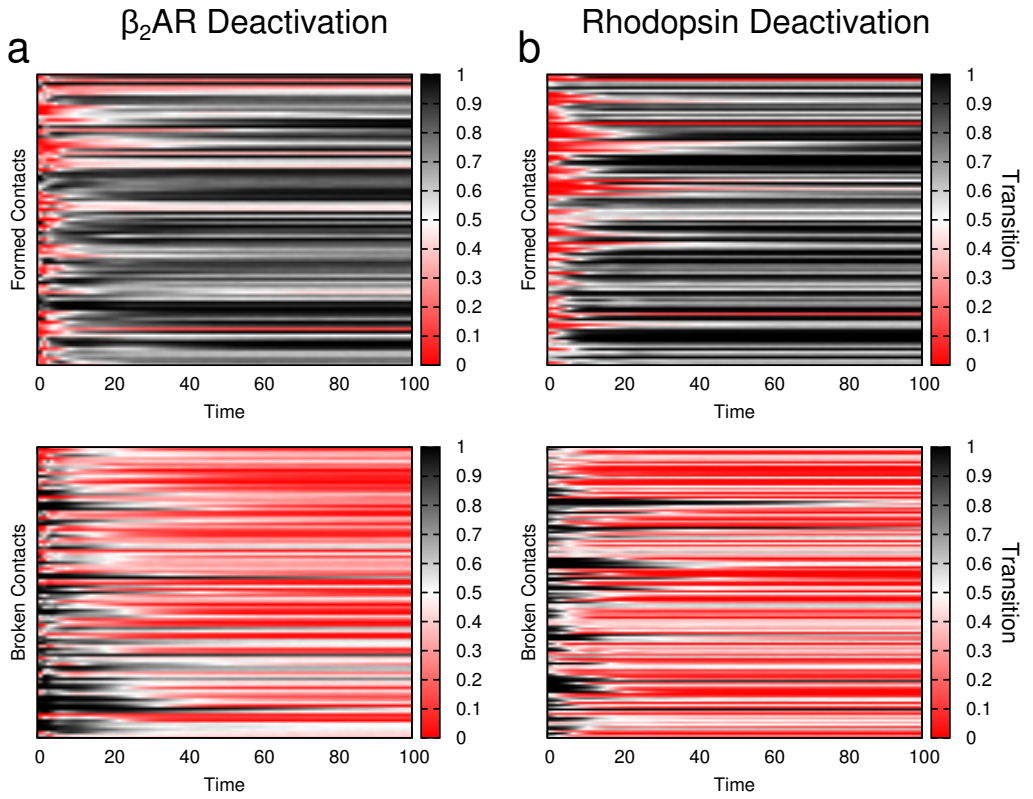

Figure S5: Deactivation contact transition matrices. The average transition value (Equation 2), indicated by color, is plotted for each unique contact (y-axis). The first 100 frames (x-axis) are shown for (a) β_2AR deactivation and (b) rhodopsin deactivation.

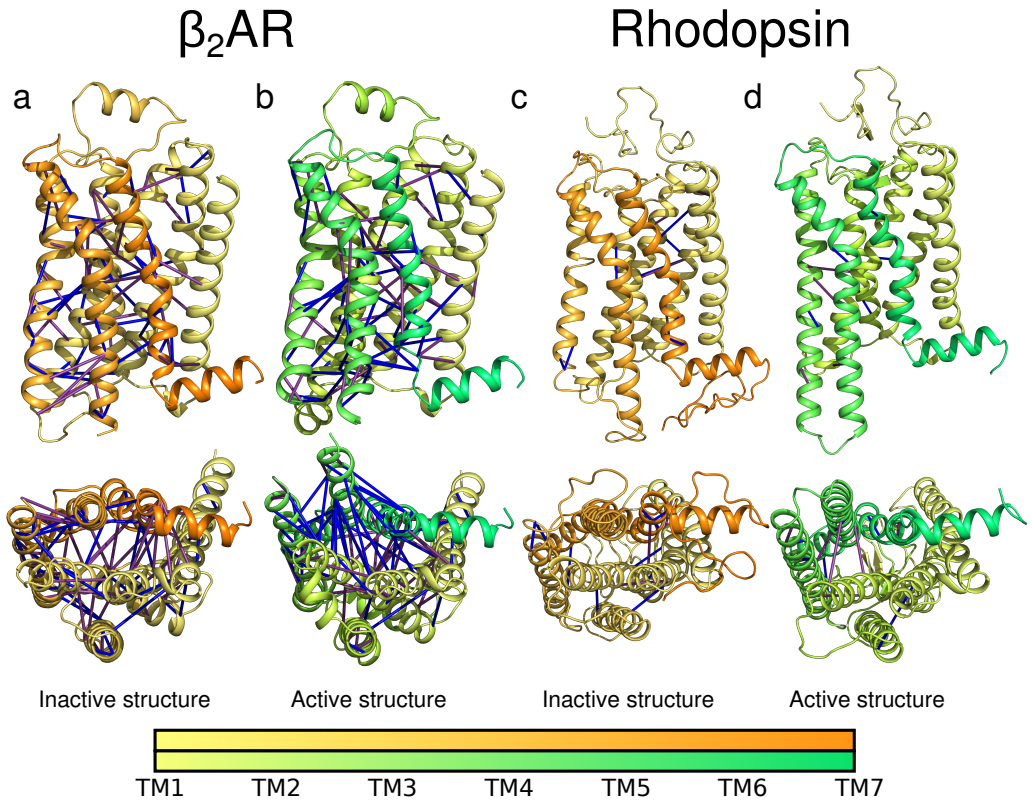

Figure S6: Backtracking contacts. The contacts that made "backtracking" transitions are mapped onto the protein structures. Here, backtracking contacts were defined as those that made at least an average 0.2 transition back toward their starting structure using the transition value in Equation 2. Bonds broken during a deactivation are shown in *blue*, those formed are shown in *purple*. Cartoons of inactive protein are colored *yellow to orange*, the active proteins are colored *yellow to green*. Data is shown for (a) β_2AR activation, (b) β_2AR deactivation, (c) rhodopsin activation, and (d) rhodopsin deactivation.

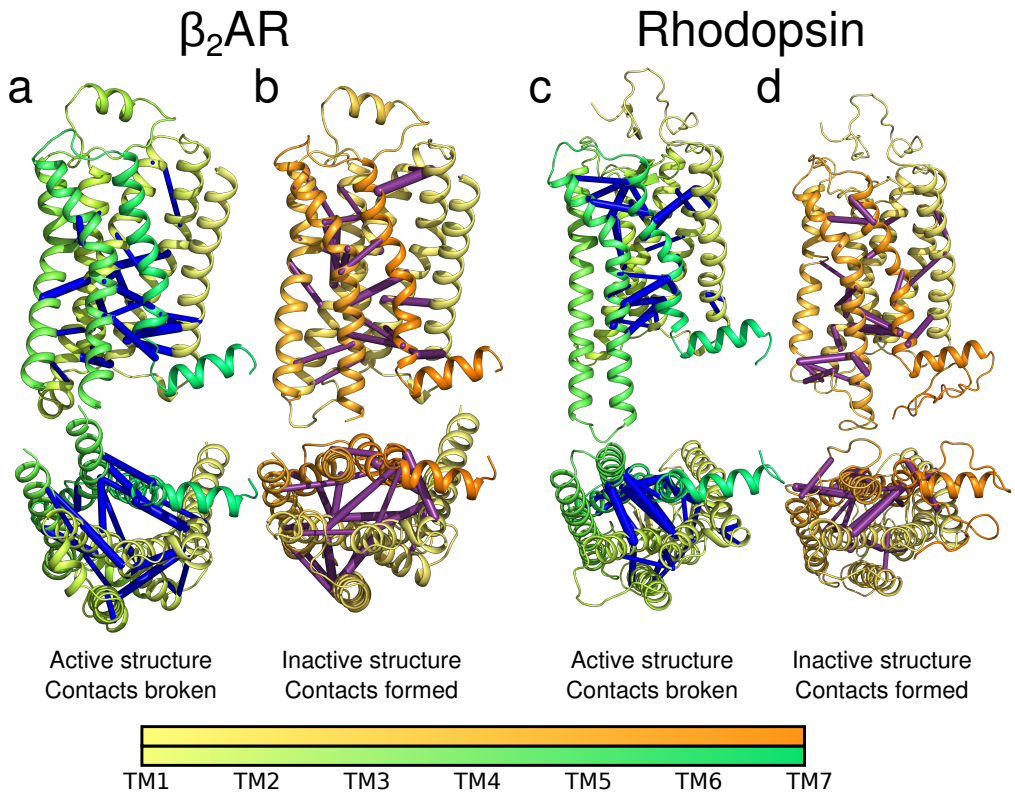

Figure S7: Contacts characterizing PC1 deactivation transitions. Those contacts whose change contributed the most to PC1 (accounting for 50% of PC1 displacement, see Section S1.1) are mapped onto cartoons of β_2AR and rhodopsin as bonds between c_α's . The bond thickness is proportional to the contribution made to the principal component. Bonds broken during a deactivation are shown in *blue*, those formed are shown in *purple*. Cartoons of inactive protein are colored *yellow to orange*, the active proteins are colored *yellow to green*. Panels (a) and (b) show data for β_2AR deactivation, panels (c) and (d) show data for rhodopsin deactivation.

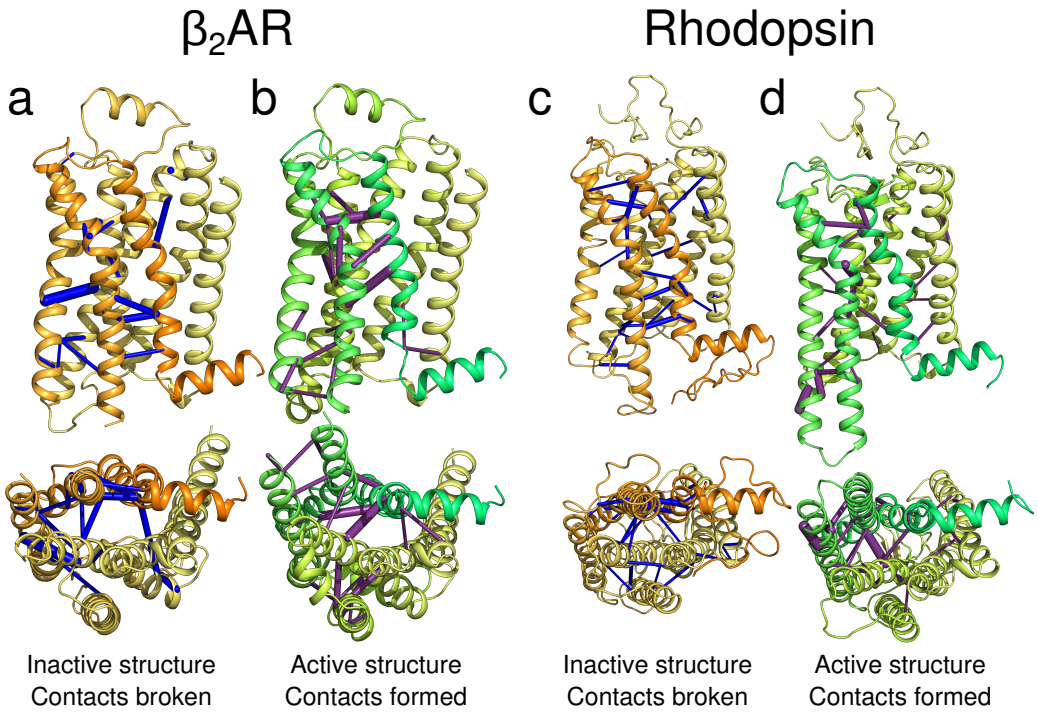

Figure S8: PC2 contacts characterizing activation. Mapping similar to Figure 7, but for those contacts crucial to *PC2*. Panels (a) and (b) show data for β_2AR activation. Panels (c) and (d) show data for rhodopsin activation.

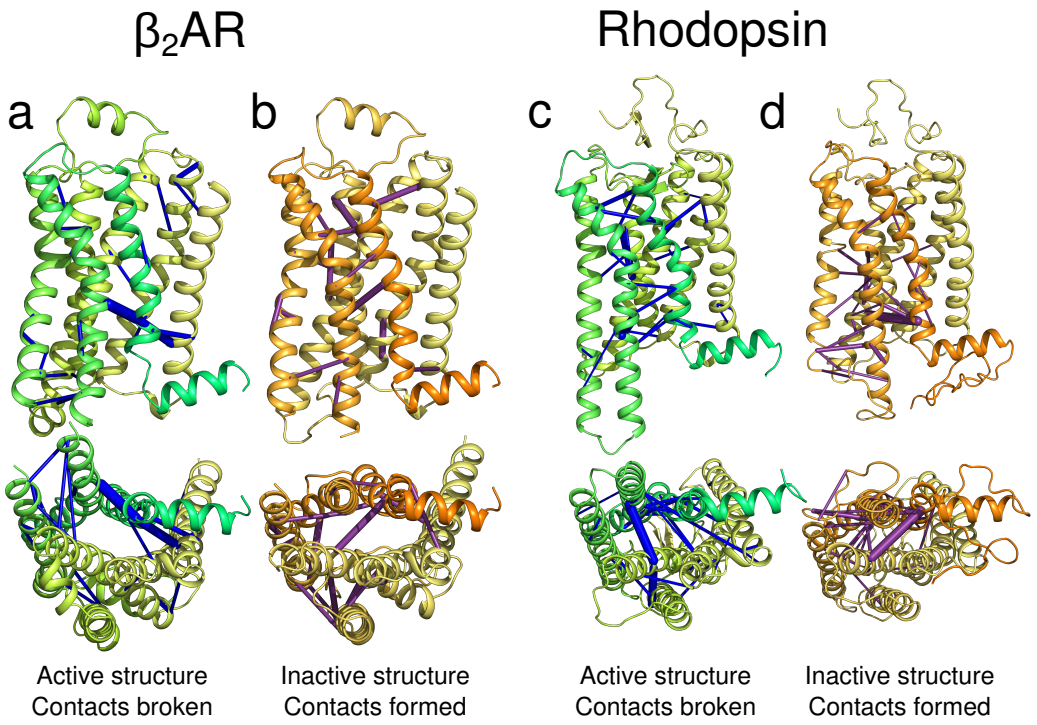

Figure S9: PC2 contacts characterizing deactivation. Mapping similar to Figure S7, but for those contacts crucial to *PC2*. Panels (a) and (b) show data for β_2AR simulations. Panels (c) and (d) show data for rhodopsin simulations.