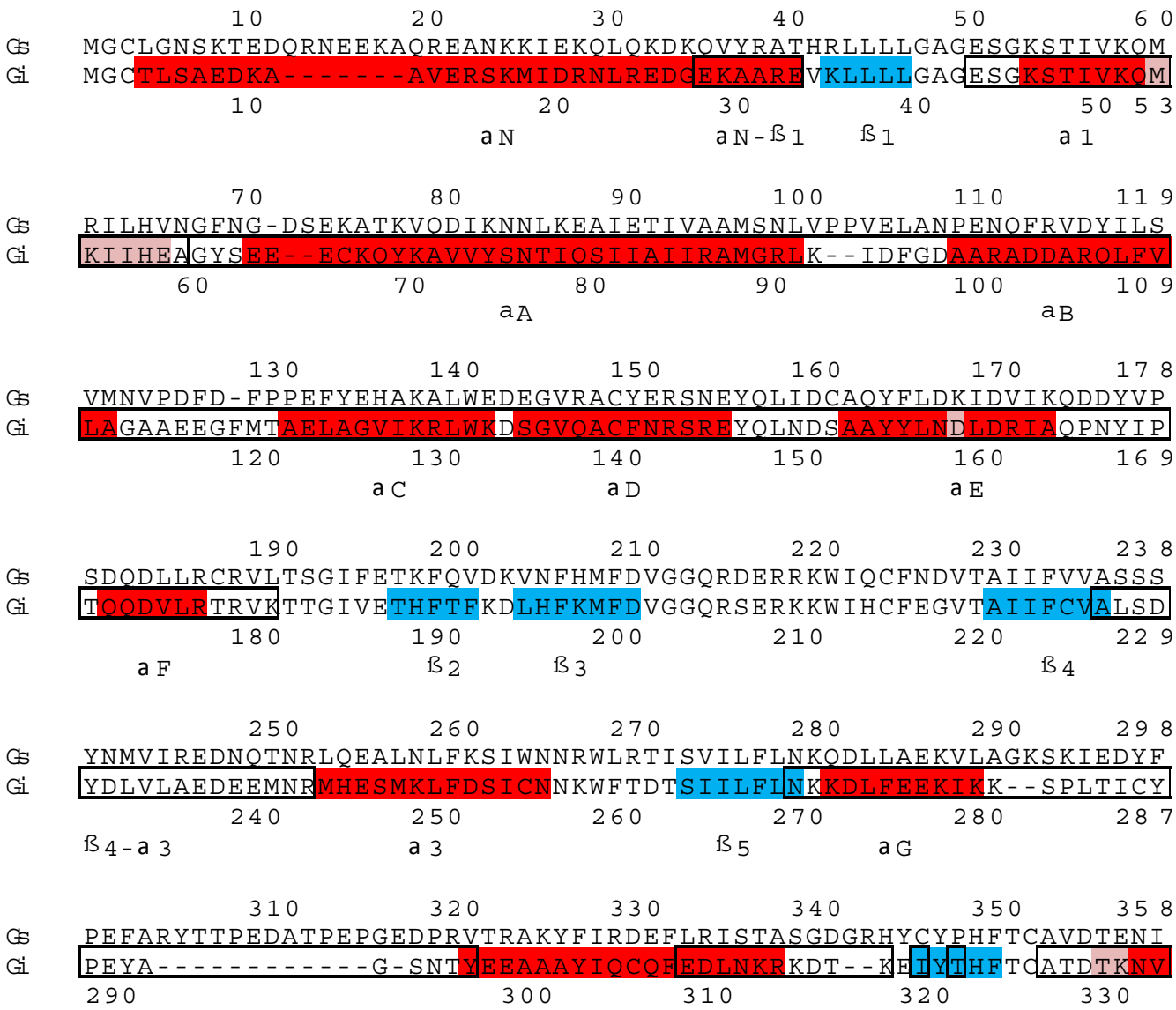
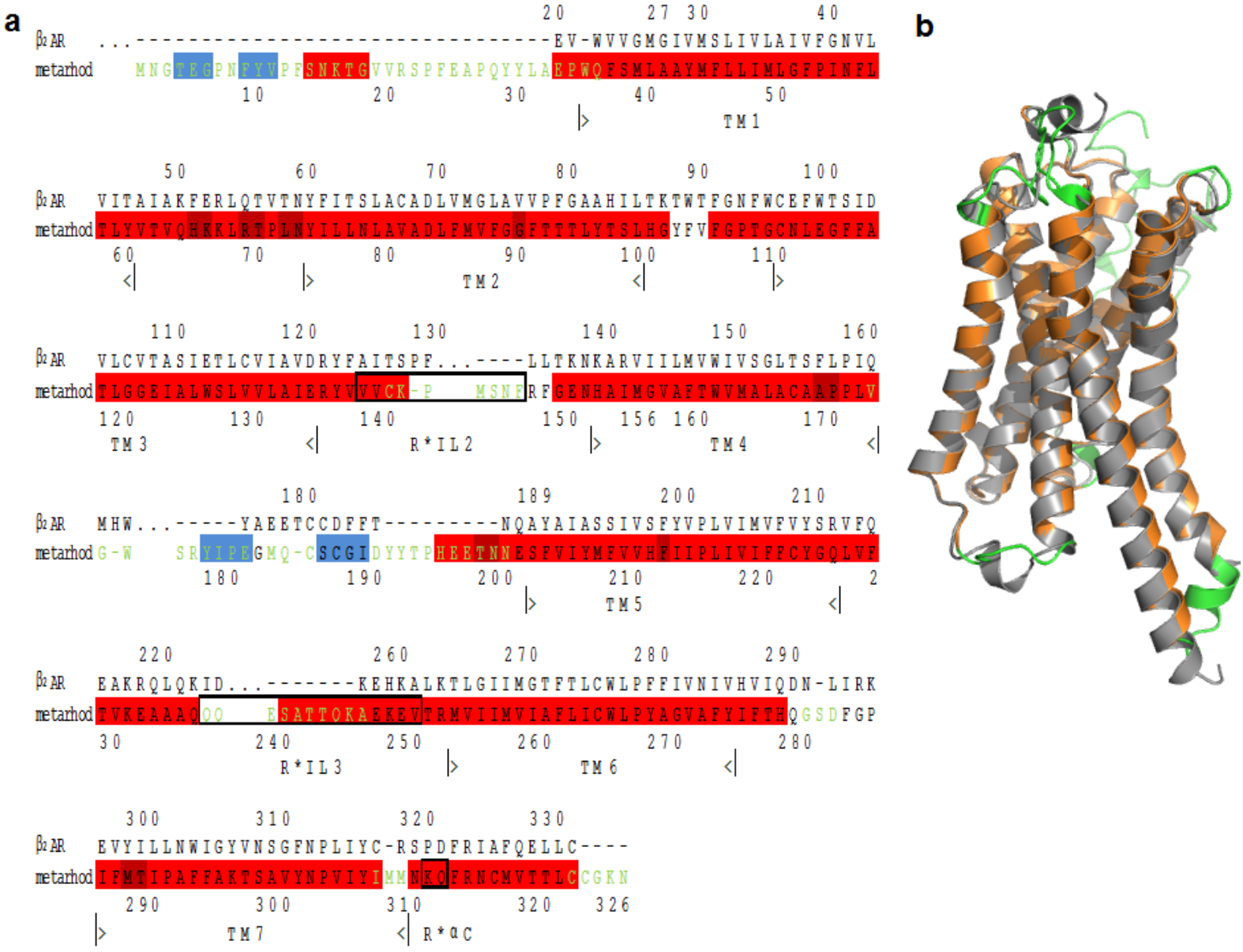
Energetic analysis of the R*-G complex links the α5 helix to GDP release and domain opening

Nathan S. Alexander, Anita M. Preininger, Ali I. Kaya, Richard A. Stein, Heidi E. Hamm, and Jens Meiler

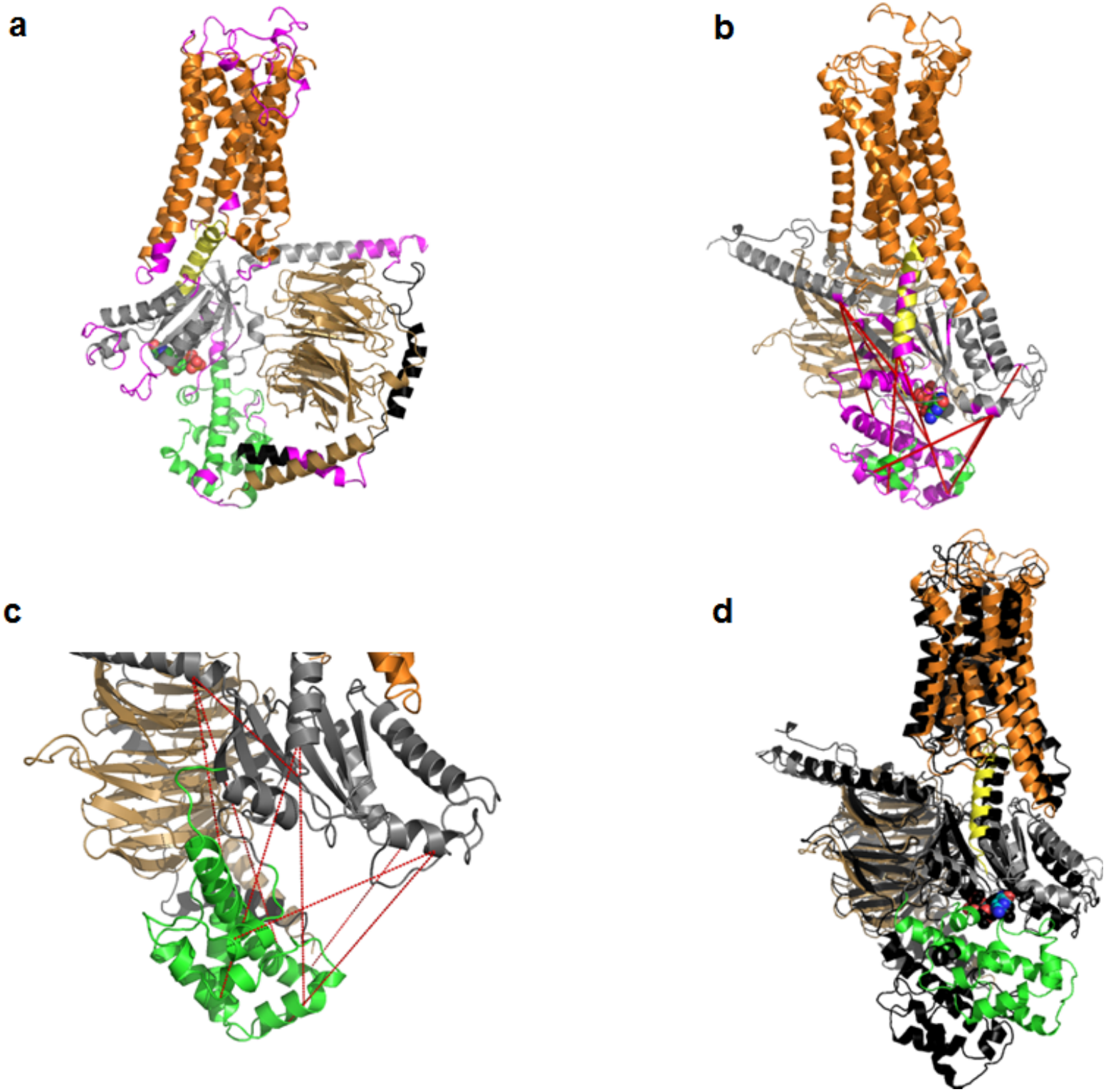
Supplemental Figures



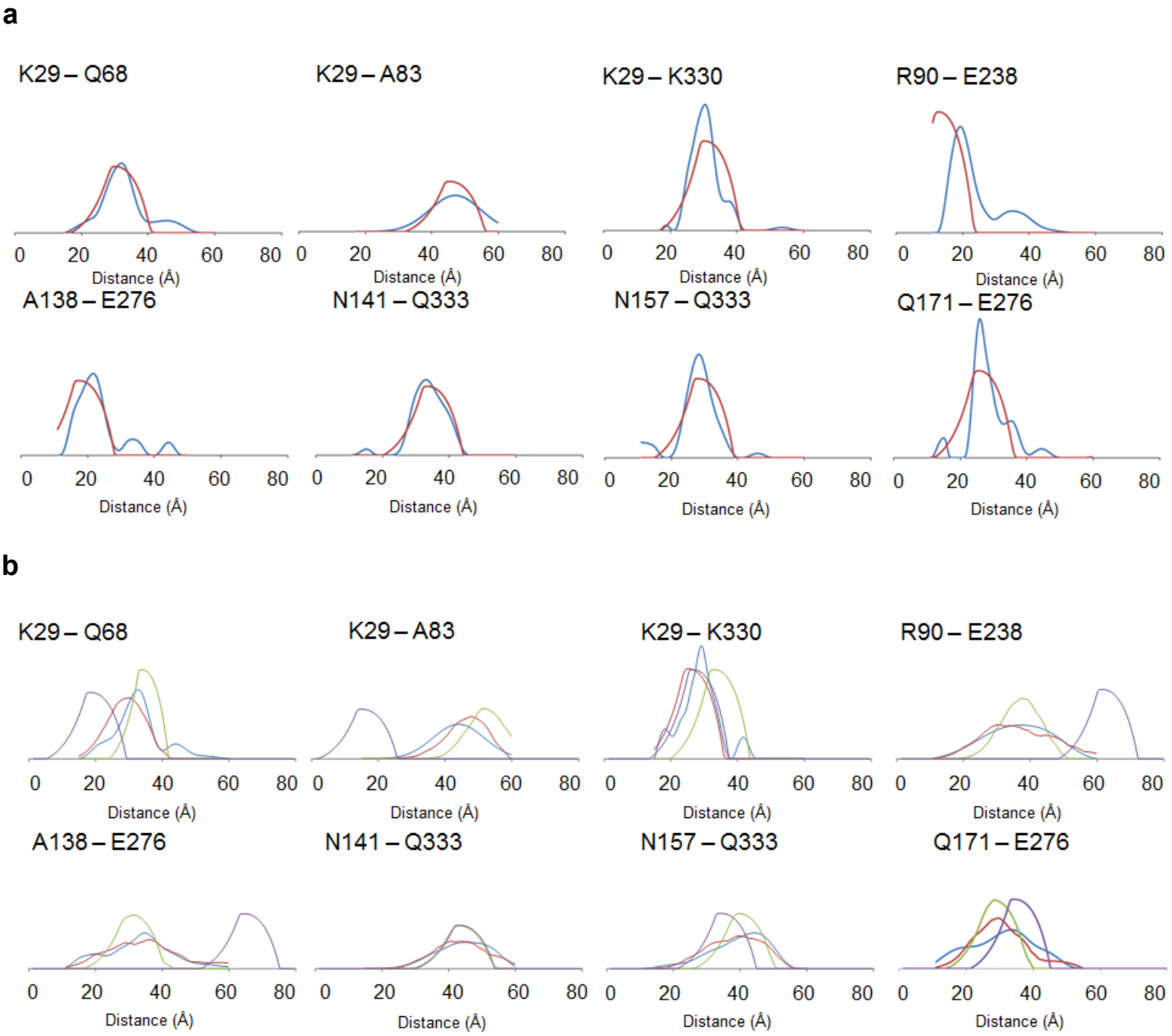
Supplemental Figure 1 Alignment of $G_{\alpha i}$ and $G_{\alpha s}$ sequences used for construction of the initial comparative model of the R*-$G\alpha_i$ complex. Alignment is from[1]. α-Helices are labeled in red. β-strands are labeled in blue. All secondary structure elements are labeled according to convention in the manuscript. Residue ranges critical for the energetic analysis are framed and named according to convention in the manuscript.



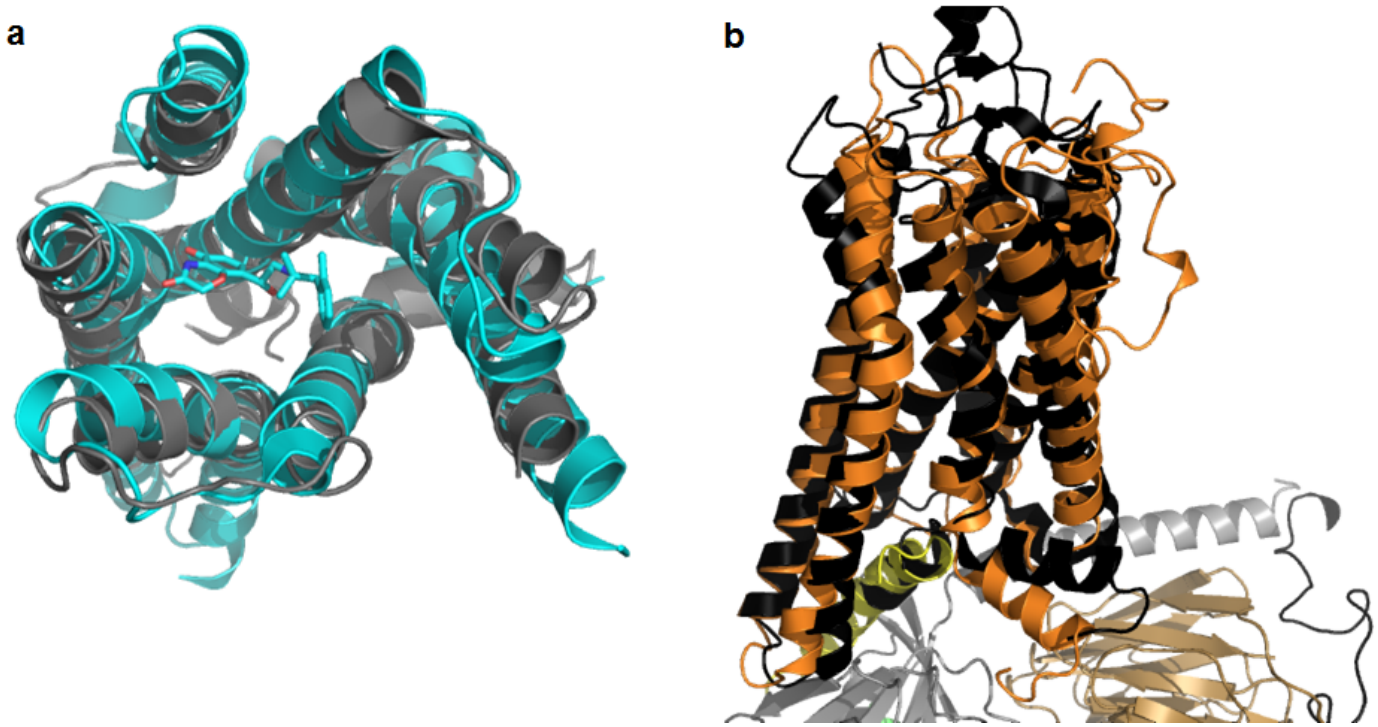
Supplemental Figure 2 Receptor comparative model construction. **(a)** Alignment of B2-adrenergic receptor with metarhodopsin as performed by MUSTANG[2]. Amino acids groups for the energetic analysis are outlined in bold and named according to seqence location. Strands are highlighted in blue. Helices are highlighted in red; dark red is used to separate consecutive helices. Residues with green text were rebuilt using the Rosetta loop building protocol. **(b)** Superimposition of the receptor comparative model (orange/green) overlayed on the template structure B2-adrenergic receptor (grey). Residues rebuilt using the Rosetta loop building protocol are colored in green.



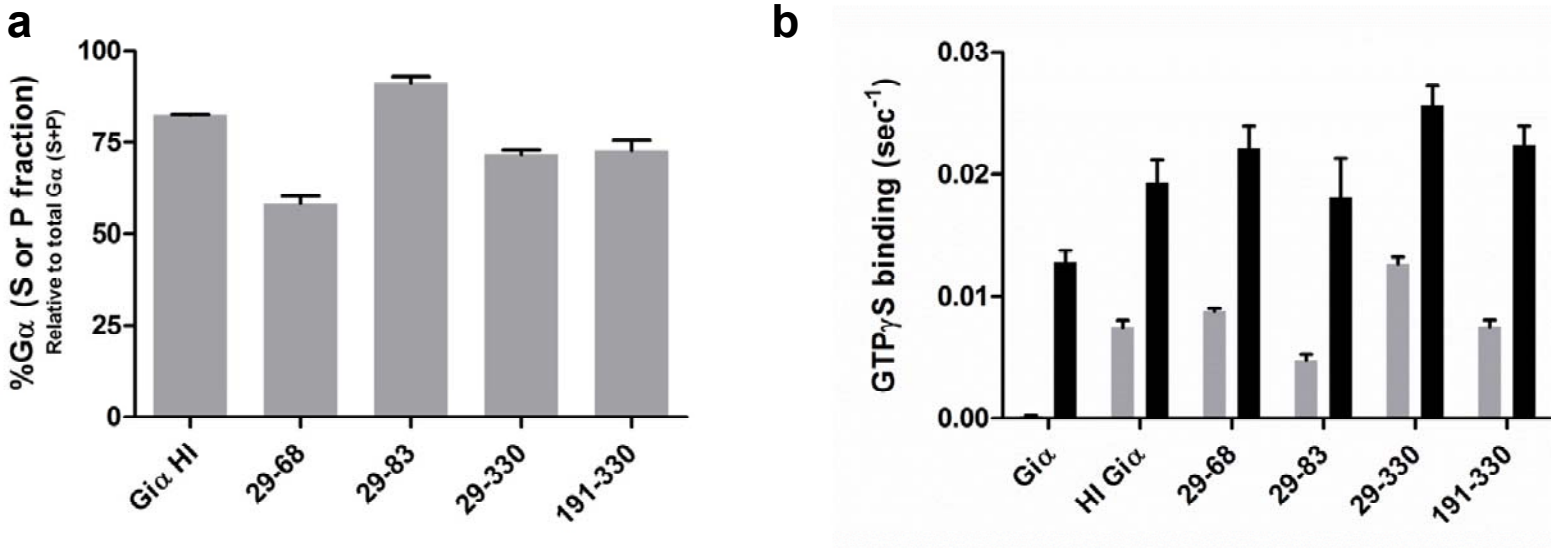
Supplemental Figure 3 Residues of the model construction basis. **(a)** Residues rebuilt using the Rosetta loop building protocol are shown in magenta coloring. **(b)** Residues with experimental data are shown in magenta coloring. Regions not rebuilt using the Rosetta loop building protocol are structured as from the template 3sn6, with the exception of the helical domain (green) which undergoes rigid body movements. In (b) DEER distances measured are shown as red lines. **(c)** Distances measured in Gα by DEER. **(d)** In black is the model before relaxation, which is compared to the colored model after relaxation. The rmsd between them is 4.1Å, excluding the helical domain residues 57-180. This shows that the relaxation protocol is a small perturbation protocol.



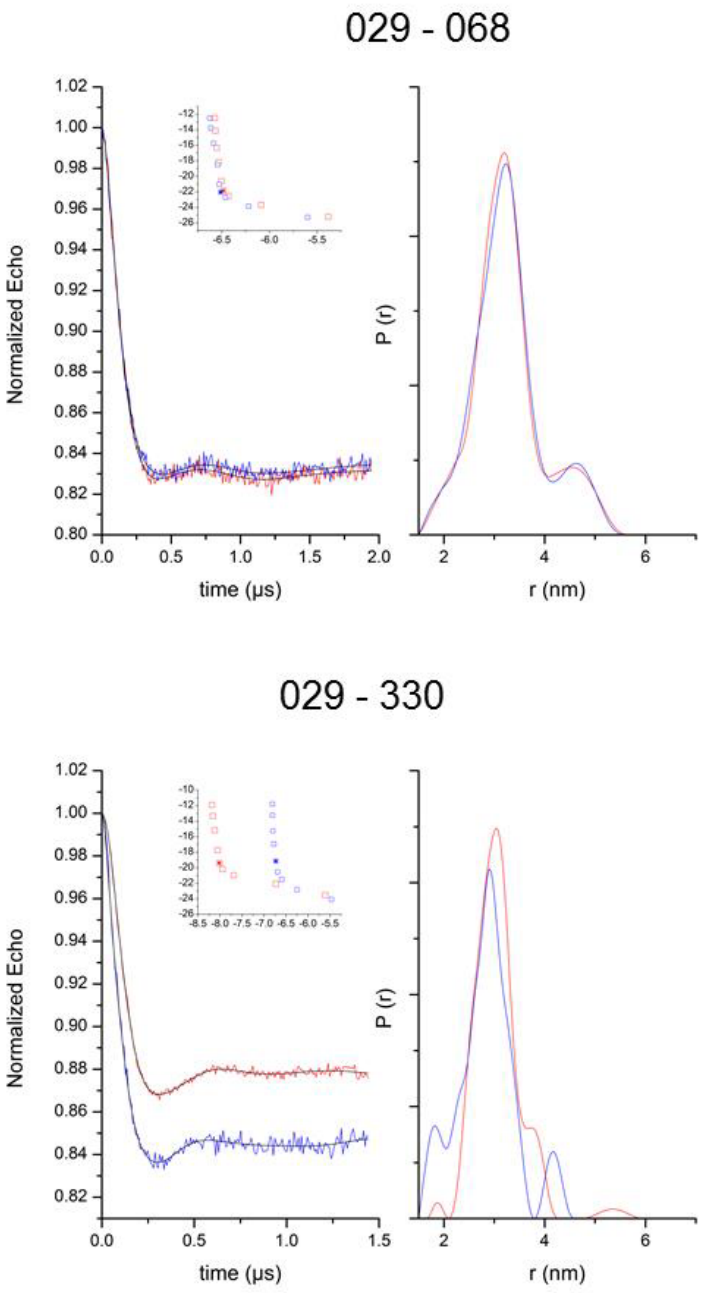
Supplemental Figure 4 DEER distributions compared to model and experimental structures. **(a)** Comparison of the experimental distance distribution as observed in EPR DEER measurements (blue) with the predicted distribution computed from the ensemble model of receptor unbound $G\alpha_i$ (red). **(b)** Comparison of the experimental distance distribution as observed in EPR DEER measurements (blue) with the predicted distribution computed from the ensemble model of the R*-G_i complex (red). In green we show the distance distribution of our previous model which reproduces average distance accurately but not the distance distribution[3]. In magenta we display the distance distribution based on the crystal structure[4].

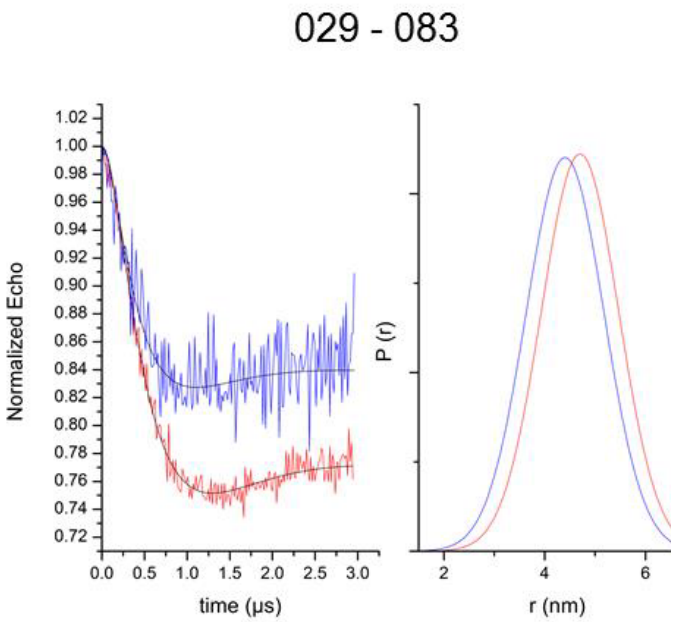


Supplemental Figure 5 Comparison of receptor structures. **(a)** Comparison of the activated rhodopsin structure (3DQB, grey) with the β_2AR (cyan). The ligand does not significantly perturb the structure. **(b)** Comparison of model receptor (salmon) with activated rhodopsin structure (3DQB, black).



Supplemental Figure 6 Characterization experiments of double mutants. **(a)** Binding of doubly spin-labeled mutant G proteins to rhodopsin in disc membranes. **(b)** Basal (grey) and receptor (black) catalyzed nucleotide exchange rates for the doubly spin-labeled mutant α-subunit. Bars represent the mean of a minimum of three independent experiments, and error bars show standard error of the mean.





Supplemental Figure 7 DEER echo decays and the corresponding distance distribution fits.

Supplemental Tables

Supplemental Table 1 Agreement of unified model with EPR DEER distance measurements

Mutants	Receptor Unbound Distance (Å) DEER	Unified Model μ ±σ	Receptor Bound Distance (Å) DEER	Unified Model μ ±σ	β_2AR-Gs [a]	Source
29-68	31.3	28.7 ±0.6	32.0	26.8 ± 2.6	17.3	new data
29-83	48.7	44.0 ±0.4	45.4	44.1 ±3.9	13.7	new data
29-330	29.7	29.1 ±0.3	28.6	24.0 ±0.3	25.4	new data
90-238	19.2	11.5 ±0.4	38.0	34.5 ±9.4	60.4 [b]	[3]
138-276	20.0	16.1 ±0.1	34.0	32.8 ±10.4	63.2	[3]
141-333	33.0	32.2 ±0.2	46.0	41.4 ±6.6	41.8	[3]
157-333	28.0	26.9 ±0.2	45.0	36.0 ±6.6	32.9	[3]
171-276	26.0	24.2 ±0.5	34.0	28.1 ±6.5	33.2	[3]

[a] β_2 adrenergic receptor-Gs complex structure from reference[4]

[b] Corresponding residue is not resolved in reference[4]. Residue 240 in $G_{\alpha i}$ is used (263 in G_s)

Supplemental Table 2 Agreement [a] of unified model with EPR DEER distance distributions

Mutants	Receptor Unbound	Receptor Bound	β_2AR-Gs[b]	Previous Ensemble	Source
29-68	0.06	0.11	0.37	0.17	new data
29-83	0.05	0.08	0.65	0.24	new data
29-330	0.05	0.08	0.04	0.19	new data
90-238	0.23	0.02	0.53[4]	0.08	[3]
138-276	0.12	0.03	0.60	0.11	[3]
141-333	0.02	0.03	0.07	0.07	[3]
157-333	0.04	0.04	0.16	0.08	[3]
171-276	0.07	0.05	0.15	0.11	[3]
Sum	0.65	0.45	2.57	1.06	

[a] Calculated as the cumulative EUCLIDIAN distance[5], where a perfect fit would give a score of 0.0. For each distance within a model, the cone-model probability distribution [6] is applied to relate the Cβ-Cβ distance into a distribution of probable spin label distances. This can then be directly compared to the measured EPR distance probability distribution

[b] β_2 adrenergic receptor-G_s complex structure from reference[4]

[c] Structural ensemble from reference[7]

Supplemental Table 3 Agreement of unified model with changes in accessibility observed by EPR CW measurements[a]

entity	amino acid	CW EPR	Δ exposure		Z-score
α1	V050	-1^{8}	-1.0	±0.5	-0.3
αF	Q171	1^{8}	4.1	±1.5	1.4
switch1	V179	1^{8}	-3.4	±1.4	-1.1
switch1	K180	2^{8}	-3.1	±2.2	-1.1
switch1	T182	0^{8}	2.8	±1.4	1.0
switch1	I184	0^{8}	2.8	±1.8	0.9
switch1	E186	2^{8}	-1.5	±0.5	-0.5
switch1	T187	0^{8}	-1.1	±0.4	-0.4
GTPase	F191	2^{9}	-1.7	±0.3	-0.6
GTPase	L194	-1^{9}	-0.7	±0.3	-0.3
switch2	S206	1^{10}	9.3	±0.8	3.2
switch2	K209	0^{10}	0.0	±0.1	-0.0
switch2	W211	0^{10}	0.4	±0.3	0.1
switch2	C214	1^{10}	0.0	±0.2	0.0
switch2	G217	-1^{10}	1.9	±0.7	0.6
αG	L273	2^{9}	0.6	±0.2	0.2
α4	A300	-1^{10}	1.1	±0.1	0.4
GTPase	E318	-2^{9}	-6.3	±0.2	-2.2
GTPase	Y320	-2^{9}	-3.5	±0.1	-1.2
α4-β6	T321	1^{9}	-2.2	±0.2	-0.7
β6-α5	K330	-2^{9}	-2.0	±0.4	-0.7
β6-α5	N331	-2^{9}	-1.4	±0.2	-0.5
α5	Q333	0^{9}	1.0	±0.1	0.3
α5	F334	-1^{9}	-2.7	±0.1	-0.9
α5	T340	0^{9}	1.3	±0.2	0.4
α5	V342	-2^{9}	-2.4	±0.3	-0.8
α5	I344	-2^{9}	-3.1	±0.3	-1.1
α5	K349	-2^{9}	-2.5	±0.1	-0.9

[a] Categorization was done manually using the difference in peak width between receptor bound and receptor unbound CW-EPR measurements. An increasing width going from receptor unbound to the receptor bound state indicates a decrease in mobility, while a decreasing spectral width indicates an increase in mobility[11]. Here, we use mobility as an indicator of exposure.

Exposure changes going from the receptor unbound to receptor bound state fall into five categories: 2 = large increase; 1 = small increase; 0 = neutral; -1 small decrease; -2 = large decrease.

Supplemental Table 4 Agreement of unified model with changes in dynamics observed H/D exchange measurements

entity	amino acid	H/D exchange [a]	Δ exposure μ ±σ		Z-score
GTPase	A033-Q038	2	-0.4	±0.4	-0.3
α1	A041-Q052	2	0.8	±1.3	0.6
α1/helical	M053-I081	1	0.2	±1.9	0.2
helical	I082-A087	0	-0.4	±0.6	-0.3
helical	F095-A099	1	0.0	±0.5	-0.0
helical	A101-A104	1	0.2	±0.2	0.1
helical	Q106-L110	0	0.4	±1.0	0.3
helical	A111-A114	0	0.9	±2.5	0.7
helical	E116-M119	0	-1.4	±2.4	-1.1
helical	E122-L123	0	1.0	±0.2	0.8
helical	V126-G135	0	0.3	±0.6	0.2
helical	I127-F140	0	0.1	±0.5	0.1
helical	L148-A153	2	0.4	±0.9	0.3
helical	N157-I168	0	-0.5	±1.2	-0.3
helical	T170-L175	1	2.2	±1.9	1.7
helical/switch1	T177-V185	2	0.3	±3.7	0.2
GTPase	H188-F196	1	-1.0	±0.8	-0.7
switch2	F199-C214	1	0.6	±2.5	0.5
GTPase	G217-A220	0	-0.1	±1.3	-0.1
β4-α3	A226-L232	2	3.6	±2.6	2.8
β4-α3/ GTPase	A235-L249	1	0.6	±1.1	0.4
GTPase	S252-W258	0	-0.5	±0.7	-0.4
GTPase	F259-S263	0	-0.9	±0.8	-0.7
GTPase	I264-L268	0	-0.4	±0.7	-0.3
αG	N269-L273	2	0.9	±0.5	0.7
αG	F274-I278	1	0.3	±0.8	0.2
αG	E289-Y290	0	-1.0	±0.2	-0.8
αG	A291-N294	0	-1.1	±2.5	-0.9
GTPase	T295-Y302	0	-0.3	±0.9	-0.2
GTPase	E297-Y302	0	-0.3	±0.8	-0.2
GTPase	Q304-E308	0	-1.5	±1.3	-1.2
α4-b5	L310-Y320	1	-3.3	±2.9	-2.5
GTPase	H322-A326	1	1.8	±2.4	1.4
β6-α5	D328-Q333	2	1.0	±2.2	0.7
α5	F334-T340	0	-1.3	±1.7	-1.0
α5	D341-F354	-2	-3.6	±2.4	-2.8

[a] Data from [12]

Supplemental Table 5 Interaction energies across selected interfaces in free and R* bound Gαi

free $G_{αi}$

entity	amino acid	GDP | $G_{αi}$ interface energy in REU		(Z-score)
GDP		5.1	±0.2	(21)
α1	S044	1.0	±0.0	(27)
α1	S047	0.8	±0.0	(24)
α1	T048	0.9	±0.1	(19)
helical	Y154	0.8	±0.0	(25)
β6- α5	T327	0.5	±0.0	(21)
α1	cumulative	3.1		
helical	cumulative	0.8		
β6- α5	cumulative	0.9		
αG	cumulative	1.0		

entity	amino acid	helical domain | $G_{αi}$ interface energy in REU		(Z-score)
helical	E065	0.9	±0.1	(8)
helical	R090	0.5	±0.4	(1)
helical	R144	0.8	±0.1	(15)
helical	Q147	1.2	±0.1	(9)
helical	D150	1.5	±0.2	(8)
helical	Y154	0.8	±0.0	(21)
helical	Q171	0.5	±0.2	(3)
helical	L175	1.3	±0.1	(16)
helical	R178	1.0	±0.1	(15)
helical	cumulative	10.1		
α1	E043	1.0	±0.1	(16)
α1	T048	0.7	±0.1	(14)
α1	K051	0.8	±0.1	(9)
α1	K054	1.4	±0.1	(19)
α1	I055	0.7	±0.2	(4)
β4- α3	V233	0.5	±0.0	(15)
β4- α3	E238	0.6	±0.4	(1)
αG	K270	0.8	±0.0	(21)
αG	K277	0.5	±0.2	(2)
GDP		2.0	±0.1	(18)
α1	cumulative	5.5		
β4- α3	cumulative	2.5		
αG	cumulative	1.8		

entity	amino acid	α5 | $G_{αi}$ interface energy in REU		(Z-score)
β6-α5	A326	2.4	±0.1	(39)
β6-α5	T327	1.1	±0.1	(13)
β6-α5	D328	0.4	±0.0	(39)
β6-α5	T329	0.8	±0.0	(40)
β6-α5	K330	0.0	±0.0	(3)
α5	N331	0.2	±0.1	(2)
α5	V332	0.9	±0.0	(27)
α5	Q333	0.5	±0.0	(32)
α5	F334	0.0	±0.0	(0)
α5	V335	1.1	±0.1	(9)
α5	F336	2.2	±0.1	(34)
α5	D337	0.2	±0.0	(13)
α5	A338	0.2	±0.0	(14)
α5	V339	1.0	±0.1	(21)
α5	T340	0.5	±0.0	(17)
α5	D341	0.0	±0.0	(2)

R*-$G_{αi}$ complex

entity	amino acid	R* | $G_{αi}$ interface energy in REU		(Z-score)
R* IL2	V138	0.6	±0.0	(21)
R* IL2	V139	0.8	±0.0	(18)
R* IL2	K141	0.6	±0.0	(15)
R* IL2	F146	2.0	±0.1	(25)
R* IL3	Q237	1.8	±0.1	(14)
R* IL3	S240	1.3	±0.0	(50)
R* IL3	T242	1.4	±0.1	(25)
R* IL3	E249	1.0	±0.0	(23)
R* IL3	V250	0.6	±0.1	(14)
R* αC	K311	1.0	±0.2	(5)
R* αC	Q312	1.4	±0.2	(8)
R*	cumulative	17.2		
αN- β1	R032	1.4	±0.1	(21)
α4- β6	E308	2.1	±0.0	(49)
α4- β6	D315	0.6	±0.2	(4)
α4- β6	K317	0.8	±0.1	(9)
α4- β6	T321	1.7	±0.1	(14)
α5	I344	0.9	±0.1	(19)
α5	N347	0.8	±0.0	(21)
α5	L348	0.8	±0.1	(15)
α5	D350	1.8	±0.1	(18)
α5	C351	0.6	±0.0	(19)
α5	L353	1.4	±0.1	(22)
α5	F354	0.7	±0.1	(13)
αN- β1	cumulative	2.2		
α4- β6	cumulative	5.6		
α5	cumulative	8.2		

entity	amino acid	α5 | R*-$G_{αi}$ interface energy in REU		(Z-score)
β6-α5	A326	-0.2	±0.1	(5)
β6-α5	T327	0.0	±0.1	0
β6-α5	D328	0.4	±0.3	(2)
β6-α5	T329	0.4	±0.2	(2)
β6-α5	K330	0.7	±0.3	(3)
α5	N331	1.7	±0.1	(20)
α5	V332	1.8	±0.1	(13)
α5	Q333	0.5	±0.0	(19)
α5	F334	1.0	±0.1	(11)
α5	V335	1.6	±0.1	(32)
α5	F336	1.9	±0.0	(46)
α5	D337	0.1	±0.0	(6)
α5	A338	1.2	±0.0	(35)
α5	V339	1.4	±0.0	(41)
α5	T340	0.2	±0.0	(6)
α5	D341	0.5	±0.1	(5)

α5	V342	0.6	±0.1	(7)	α5	V342	0.5	±0.0	(12)
α5	I343	1.1	±0.0	(42)	α5	I343	0.2	±0.0	(8)
α5	I344	0.3	±0.2	(2)	α5	I344	0.9	±0.1	(19)
α5	K345	0.0	±0.0	(0)	α5	K345	0.1	±0.2	(1)
α5	N346	0.1	±0.1	(1)	α5	N346	0.1	±0.0	(8)
α5	N347	0.0	±0.0	(2)	α5	N347	0.8	±0.0	(21)
α5	L348	0.0	±0.0	n.d.	α5	L348	0.8	±0.1	(15)
α5	K349	0.1	±0.1	(1)	α5	K349	0.7	±0.5	(1)
α5	D350	0.2	±0.0	(4)	α5	D350	1.8	±0.1	(18)
α5	C351	0.0	±0.0	n.d.	α5	C351	0.6	±0.0	(19)
α5	G352	0.0	±0.0	n.d.	α5	G352	0.5	±0.0	(28)
α5	L353	0.0	±0.0	n.d.	α5	L353	1.4	±0.1	(22)
α5	F354	0.0	±0.0	n.d.	α5	F354	1.3	±0.1	(23)
β6-α5	cumulative	4.7			β6-α5	cumulative	1.2		
α5	cumulative	9.2			α5	cumulative	21.1		
α1	T048	0.8	±0.0	(23)	α1	T048	0.0	±0.1	(0)
α1	Q052	1.5	±0.0	(39)	α1	Q052	1.0	±0.1	(16)
α1	M053	0.5	±0.1	(10)	α1	M053	0.6	±0.0	(24)
α1	I056	1.1	±0.0	(32)	α1	I056	0.0	±0.0	n.d.
GTPase	F191	0.9	±0.0	(21)	GTPase	F191	1.8	±0.1	(14)
GTPase	K192	0.2	±0.1	(3)	GTPase	K192	0.8	±0.0	(20)
GTPase	L194	0.2	±0.1	(3)	GTPase	L194	0.5	±0.1	(8)
GTPase	F196	0.8	±0.1	(16)	GTPase	F196	1.0	±0.0	(26)
GTPase	I265	0.6	±0.0	(28)	GTPase	I265	0.8	±0.1	(14)
GTPase	F267	0.6	±0.1	(7)	GTPase	F267	0.9	±0.1	(18)
GTPase	E318	0.0	±0.0	(0)	GTPase	E318	0.8	±0.3	(2)
GTPase	Y320	0.6	±0.1	(10)	GTPase	Y320	1.2	±0.2	(6)
GTPase	H322	0.7	±0.1	(9)	GTPase	H322	0.8	±0.1	(12)
					R* IL2	V138	0.6	±0.0	(21)
					R* IL2	V139	0.8	±0.0	(18)
					R* IL2	K141	0.6	±0.0	(15)
					R* IL3	E249	0.7	±0.1	(15)
					R* IL3	V250	0.6	±0.1	(14)
					R* aC	K311	0.7	±0.1	(8)
					R* aC	Q312	1.4	±0.2	(9)
GDP		1.4	±0.1	(11)	GDP		0.2	±0.0	(5)
α1	cumulative	5.0			α1	cumulative	2.2		
GTPase	cumulative	6.4			GTPase	cumulative	10.3		
R*	cumulative	0.0			R*	cumulative	8.6		

Supplemental Table 6 Comparison of inter-residue distances[a] within the β_2 adrenergic receptor-Gs complex structure and model structures at the R* | $G_{\alpha i}$ interface

G_α	R* 138	R* 139	R* 249	R* 250	R* 311	
32	-0.8	-1.3	4.5	2.1	0.2	0.3
308	0.1	0.9	4.2	1.8	6.1	0.1
317	-1.9	0.4	8.2	4.8	3.1	0.1
321	1.0	1.5	4.3	2.4	5.1	0.1
344	1.0	3.5	7.5	6.4	3.4	0.1
347	4.3	3.7	4.4	3.6	2.1	0.0
348	-4.6	-1.2	8.7	6.3	5.3	0.0
350	2.0	-0.3	-0.7	-2.6	5.3	0.1
	0.1	0.1	0.1	0.1	0.2	**Models' Distances Std. Dev.[b]**

[a] values are crystal structure distance minus the average model distance, therefore blue indicates larger model distances, red indicates larger crystal structure distance

[b] for model structures, the average of the standard deviation of distances between the given residue and the residues of the opposing interface, indicating the amount of residue-residue interaction variation within the ensemble of models

Supplemental Table 7 Comparison of inter-residue distances[a] within the β_2 adrenergic receptor-Gs complex structure and model structures at the α5 | R*-$G_{\alpha i}$ interface

α5	R* 048	R* 194	R* 196	R* 265	R* 267	R* 318	R* 320	R* 322	R* 138	R* 139	R* 249	R* 250	R* 311	
328	3.7	3.5	3.5	0.6	0.0	-1.3	-0.5	-1.4	1.9	0.9	2.8	1.0	4.3	0.1
329	-3.3	3.9	1.7	1.6	-0.2	1.5	2.6	1.9	4.9	4.9	7.6	6.2	6.4	0.1
330	1.4	6.0	5.4	1.8	-0.2	0.6	1.0	-0.9	4.6	3.3	5.3	3.5	6.6	0.1
331	5.2	0.4	2.0	5.1	6.1	3.5	4.2	4.2	1.4	0.9	6.8	3.7	6.9	0.1
332	-0.3	5.4	4.8	-1.7	-4.5	-1.8	-2.2	-4.7	2.7	1.4	3.4	1.5	3.9	0.1
333	-1.1	-0.6	-3.3	-0.2	-0.3	0.3	1.6	1.9	1.6	1.9	6.2	4.4	3.8	0.1
334	4.1	-1.8	-0.4	4.9	5.5	3.2	5.0	5.9	0.3	0.2	6.6	3.6	6.4	0.1
335	2.9	4.5	5.9	1.4	-0.7	-0.1	-1.0	-3.6	2.2	0.5	3.7	1.0	5.7	0.1
336	-2.0	4.5	0.0	-4.2	-4.7	-2.2	-2.3	-3.1	3.1	2.4	4.4	3.0	3.1	0.1
337	0.4	-3.4	-4.4	0.7	0.9	1.4	3.0	2.9	0.9	1.6	7.0	5.0	4.8	0.1
338	4.0	-0.7	1.3	5.1	3.5	3.2	4.6	2.6	0.0	-0.8	5.5	2.1	6.9	0.1
339	0.8	5.4	4.4	-3.0	-3.3	-2.4	-4.7	-4.6	3.0	1.0	3.0	0.7	5.0	0.1
340	-1.4	0.6	-3.6	-3.8	-3.0	-1.9	-1.5	-1.2	2.9	3.0	5.5	4.2	2.9	0.1
341	1.9	-4.1	-3.0	1.8	1.8	3.2	3.8	2.7	-0.7	0.0	7.5	4.8	6.2	0.1
342	2.9	1.9	2.8	1.0	0.2	0.6	-1.1	-1.4	0.9	-1.3	3.2	-0.4	6.8	0.1
343	-1.3	3.3	0.0	-5.1	-4.5	-4.9	-4.7	-3.8	4.8	2.9	3.2	1.8	3.3	0.1
344	-0.4	-3.1	-4.6	-1.9	-1.2	0.0	0.4	0.6	1.0	3.5	7.5	6.4	3.4	0.1
345	2.5	-2.8	-1.4	1.5	1.5	3.9	2.0	1.4	-2.6	-3.2	7.0	2.8	7.8	0.1
346	1.6	2.4	2.2	-1.4	-1.6	-2.8	-2.9	-2.5	1.9	-1.1	0.8	-2.3	6.0	0.1
347	-1.4	-0.3	-2.7	-3.8	-3.4	-3.2	-2.7	-2.1	4.3	3.7	4.4	3.6	2.1	0.1
348	1.5	-3.3	-3.1	0.3	0.9	2.4	1.8	1.9	-4.6	-1.2	8.7	6.3	5.3	0.1
349	2.8	-1.1	0.3	0.8	0.8	2.2	0.2	0.1	-3.1	-4.9	3.3	-2.3	9.5	0.1
350	-0.2	1.0	0.1	-3.1	-3.2	-3.8	-3.7	-3.3	2.0	-0.3	-0.7	-2.6	5.3	0.1
	0.2	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.2	0.1	0.2	**Models' Distances Std. Dev.[b]**

[a] values are crystal structure distance minus the average model distance, therefore blue indicates larger model distances, red indicates larger crystal structure distance

[b] for model structures, the average of the standard deviation of distances between the given residue and the residues of the opposing interface, indicating the amount of residue-residue distance variation within the ensemble of models

Supplemental Table 8 Average and S.E.M. (in parentheses) of nucleotide exchange rates (x10^{-3} sec^{-1}) for basal and stimulated states

	Gαi	M53C	F196C	E308C	F336C
basal	0.636	1.713	1.547	0.632	10.694
	(0.077)	(0.109)	(0.067)	(0.027)	(0.292)
stimulated	11.278	8.115	6.116	4.358	4.453
	(1.220)	(0.604)	(1.086)	(0.167)	(0.660)

Supplemental Table 9 Crystal contacts in 3SN6 as determined by COOT [13]			
Helical Domain	PRO 122	GLN 337	Receptor
Helical Domain	Lys 151	SER 352	GTPase Domain
Helical Domain	GLU 164	ASN 31	Nanobody
Helical Domain	ARG 165	ASN 268	Beta Subunit
Helical Domain	ASP 180	SER 64	Nanobody

Supplemental Note

Receptor unbound model of G_i

A comparative model of $G\alpha_i\beta\gamma$ was constructed based on the PDB coordinates 1GOT[3, 14]. The β sequence was set to bovine $G\beta_1$ (UniProt ID 62871), and the Gγ sequence was set to $G\gamma_1$ (UniProt ID 02698). The model of the receptor unbound state was then subjected to 100 independent relaxation trajectories that iterate between backbone perturbation, fast side chain optimization using a rotamer library[15], and all atom gradient minimization in ROSETTA full-atom force field[16]. The command line used for relaxation was:

```
relax.linuxgccrelease -database rosetta-3.3/rosetta_database/ -
in:file:s start_structure.pdb -out:file:fullatom -out:prefix m_16_  -
run::constant_seed -run::jran $seed_number -nstruct 2 -
in:file:fullatom -use_input_sc -relax:fast -in:file:extra_res_fa
GDP_1gp2.params
```

The ten models with lowest ROSETTA energy form the conformational ensemble representing Gαiβγ in the receptor unbound state. GDP was present throughout all steps of the protocol.

Receptor bound model R-G_i*

We focus on the receptor – G protein interface and the helical domain of Gα, which correspond to two regions of the β_2AR-G_s crystal structure exhibiting weak electron density. It is because of the weak electron density and the resulting placement of the helical domain in an orientation unanticipated by our previous modeling efforts that the helical domain of Gα was one point of focus. Our model seeks to add detail to these regions deriving a model that can be interrogated.

The allosteric mechanism by which this causes GDP release is unknown and of critical importance for understanding the GPCR-G protein system. It is therefore unfortunate that this region also exhibits weak electron density. However, the crystal structure is extremely suitable for our modeling efforts. The crystal structure is used as a template to build a homology model of the rhodopsin-transducin complex. During model construction, only the backbone coordinates of the template crystal structure are used as a starting point. This means that any errors in the placement of crystal structure side chains will not be propagated into our model. The accuracy of Rosetta's scoring functions is the key to mitigating the effects of errors and weak electron density in the crystal structure on our modeling efforts. It has been shown that the scoring functions of Rosetta are able to accurately identify the native structure of a given protein sequence to an accuracy which resolves the proper orientation of the side chains [16].

The weak electron density could not only result in incorrect side chain orientations in the crystal structure but also incorrect positioning of backbone coordinates. Further, it is our objective to model R*-Gi complex which will deviate from the β2AR-Gs complex in important details. Therefore, we use the backbone coordinates of the crystal structure as the starting point for our homology model. However, the Rosetta structural refinement protocol utilized in our manuscript has been shown to be able to drive a protein into the native structure for its sequence, when the starting structure is within 3-4Å of the native structure. We, therefore, do not require that the template structure provides the exactly correct structure for its sequence, or the sequence for which we are building our homology model. We require only that the template structure provides starting coordinates near to our target sequence's native structure.

The crystal structure of the β2AR-Gs protein complex (PDB 3SN6[17]) is used as the template for constructing a comparative model for the rhodopsin bound state of $G\alpha_i\beta\gamma$. The $G\alpha_i$ sequence of *Rattus norvegicus* (UniProt ID 10824) is used and threaded onto the G_s sequence according to the sequence alignment of[1]. The BioChemical Library (BCL) was used to perform the threading according to the command line:

```
bcl.exe TemplateModel -alignment sequences/ratgai1_3sn6A_reallyA.pir -
inputformat pir -start_model 3sn6A.pdb -prefix chain_a_thread_ -
write_zero_coordinates -print_problem_αAs -message_level Debug
```

Next, the helical domain of Gα is placed in the average position derived from previous EPR DEER data[3]. This was accomplished by superimposing the nucleotide binding domains of the threaded $G\alpha_i$ structure and the bound state model from[3] using the "super" command in PYMOL[18]. The helical domain of the threaded $G\alpha_i$ was then moved into the same relative position as the previously published model using the "super" command in PYMOL.

Missing regions of $G\alpha_i$ were reconstructed using the ROSETTA loop building protocol. The command line used was:

```
loopmodel.linuxgccrelease -fa_input –database rosetta-
3.2/rosetta_database/ -loops::input_pdb
chain_a_thread_threaded_3sn6A.pdb -out:prefix decoys/chain_a_01_ -
nstruct 1 -loops::loop_file chain_a.loops -loops::frag_sizes 9 3 1 -
loops::frag_files αArgai109_05.200_v1_3 αArgai103_05.200_v1_3 none -
loops::build_initial -loops::remodel quick_ccd -loops::refine no -
loops::relax no -out::file::fullatom
```

The regions that were built included residues 1-12, 52-66, 88-97, 116-120, 178-185, 227-241, 280-283, 286-295, and 314-318.

The receptor sequence was aligned using structure-structure alignment of 3SN6 with the structure of metarhodopsin 3PQR[19] using MUSTANG[20]. The metarhodopsin sequence was then threaded onto the 3SN6 structure using the BCL command line given above for $G\alpha_i$. Missing residues were added using the command line below:

```
loopmodel.linuxgccrelease -loops::input_pdb
receptor_threaded_3sn6O.pdb -out:prefix receptor_O_ -nstruct 1 -
file:spanfile 3pqrO.spanfile -fa_input –database rosetta-
3.2/rosetta_database/ -loops::loop_file 3pqrO.loops -loops::frag_sizes
9 3 1 -loops::frag_files aα3pqrO09_05.200_v1_3 aα3pqrO03_05.200_v1_3
none -loops::build_initial -loops::remodel quick_ccd -loops::refine no
-loops::relax no
```

The residues that were modeled during loop building included residues 1-33, 34-36, 140-146, 173-181, 183-185, 190-200, 237-246, 280-282, 307-309, and 322-326. The contents of the span file used are given below, where spanning definitions are taken from the PDB-TM:

```
TM region definition for LeuT from PDB-TM 3pqr
7 326
antiparallel
n2c
   40   65    40   65
   69   96    69   96
```

```
  113  137   113  137
  149  172   149  172
  203  228   203  228
  250  274   250  274
  290  311   390  311
```

The sequence of Gβ and Gγ were modeled as $G\beta_1$ (Uniprot ID 02693) and $G\gamma_1$ (UniProt ID 02693). For $G\beta_1$, this entailed only changing the first residue in the 3SN6 structure to a methionine. For $G\gamma_1$, a blast alignment was performed between the $G\gamma_2$ sequence (UniProt ID 63212) of 3SN6. Threading and loop building were performed as described above. The residues that were built to complete the sequence included 1-5, 15-28, and 65-74.

After loop construction, the model was relaxed in ROSETTA 46 times. To accommodate the receptor, the relaxation utilized ROSETTA's full atom membrane potential[21]. The relax command line used was:

```
relax.linuxgccrelease -/rosetta-3.3/rosetta_database/ -in:file:s
bound_gdp.pdb -out:file:fullatom -out:prefix m_0_  -run::constant_seed
-run::jran $seed_number -nstruct 1 -in:file:fullatom -use_input_sc -
relax:fast -relax:membrane -score:weights membrane_highres.wts -
membrane:normal_cycles 100 -membrane:normal_mag 15 -
membrane:center_mag 2 -file:spanfile 3pqr_renum.span -
constraints::cst_fa_file bound.cst -constraints::cst_fa_weight 12 -
constraints::epr_distance -in:file:extra_res_fa GDP_1gp2.params
```

The contents of the restraint file were:

```
AtomPair CB  90 CB 238 SPLINE EPR_DISTANCE 38   1.0 0.5
AtomPair CB 157 CB 333 SPLINE EPR_DISTANCE 45   1.0 0.5
AtomPair CB 171 CB 276 SPLINE EPR_DISTANCE 34   1.0 0.5
AtomPair CB 141 CB 333 SPLINE EPR_DISTANCE 46   1.0 0.5
AtomPair CB 138 CB 276 SPLINE EPR_DISTANCE 34   1.0 0.5
```

The contents of the span defining TM regions was derived from the PDB-TM:

```
TM region definition for 3pqr
7 1095
antiparallel
n2c
  808 833 808 833
  837 864 837 864
  881 905 881 905
  917 940 917 940
  971 996 971 996
  1018 1042 1018 1042
  1058 1079 1058 1079
```

The model with lowest ROSETTA energy was used as the starting point for further modeling of the R*-G_i complex.

Helical domain position sampling

The BioChemicalLibrary (BCL) was used to perform rigid body conformational sampling of the helical domain in Gα relative to the nucleotide binding domain. EPR distance restraints will be used for filtering models in a later step and therefore are not used here in order to sample the largest conformational space. The BCL command line used for sampling was:

```
bcl.exe Fold -protocols Default Dock -mutate_protocols Default Dock -
score_protocols Default Dock -prefix m_356 -nmodels 2 -native
m_12_bound_gdp_0001.pdb -start_model m_12_bound_gdp_0001.pdb -
use_native_pool -mc_number_iterations 750 500 -min_sse_size 0 0 0 -
αAclass AABackBone -mc_temperature_fraction 0.5 0.2 -quality RMSD -
domain_specify helical_bound_23.domain -random_seed $seed_number
```

The maximum translation and rotation of the helical domain allowed to occur in a single Monte Carlo step was 10.0Å and 60º, respectively. The residues 58-62 and 178-185 linking the helical domain to the nucleotide binding domain were removed to allow free movement of the helical domain. In addition, to ensure the helical domain could again be reconnected to the nucleotide binding domain, the loop closure tolerance was set to zero.

Models were filtered to remove those with clashes and 743 remained after filtering. Each of the models underwent four independent ROSETTA loop building trajectories to reconnect the helical and nucleotide binding domains of Gα. Residues 57-63 and 177-186 were built using the following command line:

```
loopmodel.linuxgccrelease -loops::input_pdb bcl.pdb -out:prefix m_0 -
nstruct 4 -loops::loop_file chain_a.loops -run::constant_seed -
run::jran $seed_number -fa_input –database rosetta-
3.2/rosetta_database/ -loops::frag_sizes 9 3 1 -loops::frag_files
αArgai109_05.200_v1_3 αArgai103_05.200_v1_3 none -loops::build_initial
-loops::remodel quick_ccd -loops::refine no -loops::relax no -
out::file::fullatom
```

After modeling the missing residues, each model was relaxed in ROSETTA. The command line for relaxation includes:

```
relax.linuxgccrelease –database rosetta-3.3/rosetta_database/ -
in:file:s m_0S_0004.pdb -out:file:fullatom -out:prefix m_0_ -
run::constant_seed -run::jran $seed_number -nstruct 1 -
in:file:fullatom -relax:fast -relax:membrane -score:weights
membrane_highres.wts -membrane:normal_cycles 100 -membrane:normal_mag
15 -membrane:center_mag 2 -file:spanfile 3pqr_renum.span -
in:file:extra_res_fa GDP_1gp2.params -constrain_relax_to_start_coords
```

The best relaxed model according to ROSETTA score for each model was determined. These were then used for finding subsets of models which agree with experimentally measured distance probability distributions.

An ensemble of helical domain positions consistent with EPR distance restraints

The BCL was used to perform the Monte Carlo optimization of finding a set of models which reproduce the EPR measured distance probability distributions. The double mutant EPR measurements that were used included 90-238, 138-276, 141-333, 157-333, and 171-76. Each distribution was trimmed to a maximum of under 60Å. The command line for optimization was:

```
bcl.exe FitEPRDistribution -exp_hist_list epr_distributions_trimmed.ls
-exp_hist_data_columns 0 1 -model_list relaxed_best_renum_full.ls 0 -
num_fits 1000 -start_size_range 5 20 -prefix fit_01/ -
terminate_criteria 0.1 2500 -use_pdbid_numbering -message_level
Standard
```

The protocol for the minimization is as follows. An initial ensemble with between 5-20 models is randomly selected from the pool of all models. Next, a maximum of 2500 Monte Carlo steps occurs, where a step includes the removal, addition, or swapping of a model in the current ensemble of models. Changes to the ensemble are scored and accepted or rejected based upon the cumulative EUCLIDIAN distance. For each distance distribution, the cumulative EUCLIDIAN distance is calculated between the experimental distribution and the distribution derived from the current ensemble of models. The total score of the current ensemble is the sum of the cumulative EUCLIDIAN distances. This procedure was performed in 1000 independent trajectories. The ensemble with the smallest total cumulative EUCLIDIAN distance score is used as the representative model of the R*-G_i complex.

Comparing EPR CW observed changes in exposure with models of free G_i and R-G_i*

Using CW-EPR, the mobility of a side chain can be measured, and it is assumed that the mobility of the side chain indicates its exposure. Buried residues will have many neighbors and be immobile, while exposed side chains will have fewer neighbors and be more mobile. This correlation has been demonstrated experimentally but is imperfect[22]. To avoid over-interpretation of the experimental data, we manually classified the experimental measurements into five categories depending on the magnitude and sign of exposure change: strong decrease (-2) to strong increase (+2) (Supplemental Table 2).

For our models of free G_i and R*-G_i we calculated solvent accessibility in the ensembles using a neighbor count measure that is optimized to correlate with relative solvent accessible surface area (rSASA)[23]. The accessibility change between the unbound and bound states predicted by the model was calculated. This was accomplished by taking all residues and calculating their neighbor count in the receptor bound (RB) and unbound (RU) and states[23]. The change in neighbor count between the two states is summarized for the ensemble of models by the average and standard deviation (Supplemental Table 2 and Supplemental Table 3). The average change in neighbor count for an individual residue is calculated as $\overline{\Delta NC} = \frac{\sum_i \sum_j \left(NC_i^{RB} - NC_j^{RU}\right)}{N_{RB} * N_{RU}}$, where NC_i^{RB} and NC_j^{RU} are the receptor bound and receptor unbound neighbor counts of a residue of interest in models i and j, respectively. N_{RB} and N_{RU} are the number of models in the receptor bound and receptor unbound ensembles, respectively. As expected, model and experiment agree in the direction of change for the majority (61%) of the cases yielding a moderate correlation coefficient of 0.33.

To plot differences between predicted and experimentally observed exposure changes, the $\overline{\Delta NC}$ for each residue is converted to a z-score relative to the other residues of interest as $Z_{\overline{\Delta NC}} = \frac{\overline{\Delta NC}}{\sigma_{\overline{\Delta NC}}}$, where $\sigma_{\overline{\Delta NC}}$ is the standard deviation of all $\overline{\Delta NC}$. The z-score indicates the number of standard deviations of a given $\overline{\Delta NC}$. In Figure 3, the difference between experimental classification and model z-scores are plotted.

The command line used for analysis of mobility measurements is given below:

```
bcl-all-static.exe AnalyzeRestraintAgreement -pdb_list
bound_ensemble.ls -analysis_prefix analysis -analysis_type_enumerated
```

```
'AccessibilityChange
(filename_postfix=.AnalyzeAccessibilityChange,start_ensemble_filename=
unbound_ensemble.ls,experimental_data_filename=data_formatted.cst,mean
_min_cutoff=0,mean_max_cutoff=9999,zscore_min_cutoff=0,zscore_max_cuto
ff=9999,pymol_output_filename="accessibility.pml",ensemble_representat
ive_index=0,ensemble_representative_from_start_ensemble=1,gradient_min
=-4,gradient_max=4,direct_relation=0)'
```

Comparing H/D exchange observed changes in exposure with models of free G_i and R-G_i*

Using H/D exchange experiments the solvent accessibility of amino acids can be measured. Buried residues will be difficult to access and experience little H/D exchange as compared to solvent exposed ones[24]. The protein is digested and the ratio of H/D exchange is determined for the resulting peptides using mass spectrometry. To avoid over-interpretation of the experimental data, we manually classified the experimental measurements into five categories depending on the magnitude and sign of exposure change: strong decrease (-2) to strong increase (+2) Supplemental Table 2).

Per residue exposure changes were computed as discussed above. The exposure change for stretches of amino acids was predicted by averaging $\overline{\Delta NC}$ over the respective residues. As expected, model and experiment agree in the direction of change for the majority (72%) of the cases yielding a moderate correlation coefficient of 0.56.

The command line used for analysis of H/D exchange measurements is given below:

```
bcl-all-static.exe AnalyzeRestraintAgreement -pdb_list
bound_ensemble.ls -analysis_prefix hd_exchange -
analysis_type_enumerated 'AccessibilityChange
(filename_postfix=.AnalyzeAccessibilityChange,start_ensemble_filename=
unbound_ensemble.ls,experimental_data_filename=data_uniq.cst.bak,mean_
min_cutoff=0,mean_max_cutoff=9999,zscore_min_cutoff=0,zscore_max_cutof
f=9999,pymol_output_filename="accessibility_hdexchange.pml",ensemble_r
epresentative_index=0,ensemble_representative_from_start_ensemble=1,gr
adient_min=-4,gradient_max=4,direct_relation=0)'
```

Energetic analysis of key interfaces in the complex

The Rosetta ΔΔG protocol[25] is used to calculate the interaction energies of four interfaces in the Gαi and R*-$G\alpha_i$ complex. These interfaces included: $G\alpha_i$-helical domain|$G\alpha_i$-GTPase domain interface, GDP|$G\alpha_i$ domain interface, R*|$G\alpha_i$ domain interface, and C-terminal helix α5|$G\alpha_i$ domain.

The $G\alpha_i$-helical domain|$G\alpha_i$-GTPase domain interface involves residues in Gαi and is most relevant for the receptor unbound state of $G\alpha_i$. The $G\alpha_i$-helical domain is considered to be residues 61-180. The $G\alpha_i$-GTPase domain is considered to be residues 1-60 and 181-354. The protocol keeps $G\alpha_i$-GTPase and GDP in place while moving the Gαi-helical domain away to calculate the change in energy.

The GDP|$G\alpha_i$ domain interface ΔΔG calculation includes all residues of Gαi and the bound GDP molecule and is most relevant for the receptor unbound state of $G\alpha_i$. The GDP is moved away from its binding pocket to calculate the change in energy.

The R*|Gα$_i$ domain interface involves interactions between the receptor and Gα$_i$. The protocol calculates the energy differences at the interface of Gα$_i$ and the receptor between when the receptor is bound and unbound.

The C-terminal helix α5|R*-Gα$_i$ domain interface examines interactions of the C-terminal α5 helix of Gα$_i$. The C-terminal helix is considered to be residues 325-354. The energetic of this interface can be calculated in the receptor unbound Gαi state and the R*-Gα$_i$ state. During the ΔΔG calculation the C-terminal helix of Gα$_i$ is moved away from its native position. This protocol gives information about the key interactions α5 makes with Gα$_i$ when not receptor bound, and how those interactions change as a result of receptor binding. It also gives key interactions of α5 with the receptor.

For each interface, the individual residues were examined that contribute to the total ΔΔG. The receptor unbound and bound states are represented as ensembles of structures, so the ΔΔG for a residue is the average ΔΔG seen for that residue over the ensemble. A residue is considered to be important to the interface if the mean ΔΔG is more than 0.5 REU. In addition, the standard deviation of ΔΔG over the ensemble must be less than half the mean ΔΔG (i.e. Z-score greater than 2); this ensures that the ΔΔG is consistently seen across the ensemble.

The difference in the energy values between the unbound and bound states for a given interface indicates changes in key interactions. Calculations were conducted over the ensembles for the receptor unbound and bound states to compute mean ΔΔG and standard deviations. Only statistically significant contributors (Z-score larger than 2) that are large (|ΔΔG| > 0.5 REU) were considered for further analysis.

A sample command line for calculating ΔΔG is given below.

```
rosetta_scripts.static.linuxgccrelease -database rosetta_database/ -
in:file:extra_res_fa gdp.params -parser:protocol interface.xml -
in:file:l pdbs.ls -out:prefix ddgs/ -score:weights
membrane_highres.wts -membrane:normal_cycles 100 -membrane:normal_mag
15 -membrane:center_mag 2 -file:spanfile 3pqr_renum.span
```

The protocol file includes the following options:

```
<ROSETTASCRIPTS>
        <MOVERS>
                <ddG name=ddgy jump=5 symmetry=0 per_residue_ddg=1
repack=0 scorefxn=score12/>
        </MOVERS>
        <PROTOCOLS>
                <Add mover_name=ddgy/>
        </PROTOCOLS>
</ROSETTASCRIPTS>
```

G protein activation pathway by energetic analysis

Recent successes in modeling membrane proteins[26,27,28] accurately suggest that computational methods are becoming capable of adding atomic detail in regions where experimental information is not at high resolution. Here we present a model of receptor-mediated changes in a G protein that lead to its activation. The unified model combines experimental data from orthogonal methods such as CW-EPR and DEER, steady-state fluorescence, cryo-EM, and H/D exchange experiments. New insights include: The structural dynamics of the helical domain is

defined through a realistic ensemble. The model takes into account receptor-mediated translocation of the C terminus, which is communicated both through the GTPase domain and across the helical domain to lead to domain opening and GDP release. Furthermore, a quantitative energetic analysis details the residues which are predicted to be important in the inactive heterotrimer as well as receptor-bound states. This model helps reconcile the crystal structure with the available experimental data.

The activated receptor R* interacts with $G\alpha_i$ through three major pathways: the high affinity C terminus of α5 (~1/2 of energetic contribution), the interaction of R* IL 3 with α4-β6 loop (~1/3 of the energetic contribution) and through the interaction of R* IL 2 with the αN-β1 loop (~1/6 of the energetic contribution). The consequences of receptor binding the high-affinity C terminus of α5 include a 5.7Å and a 63º rotation of α5 (Supplemental Movie 8).

The interaction of α5 with the GTPase β-sheet shifts from one energetic minimum (6.4 REU) to a second, deeper energetic minimum (10.3 REU) upon receptor activation, breaking interactions and forming new ones. This conformational change is captured by crystallography[4] and was previously deduced from CW-EPR mobility studies[10]. The interaction between I343 and β2, β3 is destabilized. This residue is part of a hydrophobic cluster with F191 and F196, an interaction that is weakened in the R* state. F191 and F196 as well as the outmost tips of the β-sheet at K192, E318, and Y320 rearrange and engage in improved, tighter contacts with α5.

The energetic stabilization of the C terminus of α5 and its interaction with the β-sheet drive the conformational reorientation of this helix. The rotation and shift of α5 causes residues A326-T329, to lose contacts with the rest of G_i. Specifically, the interaction energy of this region decreases from 4.7 REU to 1.0 REU, which brings energetic changes induced by receptor binding from the C terminus of α5 towards the GDP binding site. Residues T329-V332 of the α5 helix unwind, loosening attractive interactions with GDP (1.4 REU to 0.2 REU). More importantly, an exquisitely strong interaction of α5 with α1 is also weakened (5.0 REU to 2.2 REU) leading to its structural destabilization. This is also indicated by the absence of crystallographic coordinates for the C terminus of α1 in the crystal structure of the β_2AR-$G\alpha_s\beta\gamma$ protein complex[17]. Note that α1 links to the helical domain so that weakening of its interaction with α5 ultimately contributes to the release of the helical domain as discussed below.

In addition to α1, the helical domain is anchored through a network of polar interactions between the αA helix, αD-αE loop, αF-β2 loop on the helical domain and the β5-αG-α4 loop and β4-α3 loop on the GTPase domain (4.3 REU). These interactions need to be broken for the helical domain to be released. We hypothesize that two mechanisms contribute to this event: firstly, unwinding of residues T329-V332 in α5 lengthens the β6-α5 loop leading to loss of its interaction with the guanine ring of GDP. The loop adopts a different conformation and requires extra space, thus exerting force on αG. Secondly, we hypothesize that the interaction of R* IL 3 with α4-β6 loop is propagated to αG and possibly the β4-α3 loop. Our models show an 11º rotation of αG that shifts the C terminus of αG accompanied by conformational changes in the respective loop regions connecting αG. Our unified model also shows an attractive interaction between R* IL 2 with the αN-β1 loop. This signal could be propagated via the GTPase domain β-sheet to α5, α1, or the switch regions – a process difficult to track given the small amplitude of this interaction.

Once released, the helical domain samples a wide but well-defined space which is somewhat less dramatic than that observed in the crystal structure of the β_2AR-$G\alpha_s\beta\gamma$ protein complex[17] but consistent with cryo-EM studies[29]. All of these studies are qualitatively consistent in terms of the flexibility of the helical domain. Since crystallization represents the lowest energy structure, it may be just one possible conformation among many, influenced by both crystal contacts as well as stabilizing proteins required to obtain this structure. Indeed, based on the β_2AR-G_s

structure [4], the helical domain makes crystal contacts with the helical domain, receptor, GTPase domain, nanobody, and beta subunit (Supplemental Table 9). Ideally, resolution of structures of other receptor-G protein complexes without accessory proteins should clarify the importance of this conformation.

Supplemental Movies

Supplemental Movie 1 Modeled interaction of activated rhodopsin (R*) with G_i. The receptor starts in the non-activated conformation (red) and then moves into the activated state (orange). Upon receptor activation, the α5 C-terminal helix of $G\alpha_i$ (yellow, blue) binds to R*. The helical domain (green) opens away $G\alpha_i$-GTPase (grey) and GDP is released (spheres). Gβ is shown in brown; Gγ shown in black.

Supplemental Movie 2 Agreement of unified model with single particle EM class averages, relating to Figure 3A.

Supplemental Movie 3 Energetics of helical domain|$G\alpha_i$ interface in free $G\alpha_i$, relating to Figure 4A.

Supplemental Movie 4 Energetics of the GDP|$G\alpha_i$ interface in free $G\alpha_i$, relating to Figure 4B.

Supplemental Movie 5 Energetics of R*|$G\alpha_i$ interface in the R*-$G\alpha_i$ complex, relating to Figure 4C.

Supplemental Movie 6 Energetics of the interface between α5|$G\alpha_i$-GTPase in the basal state, relating to Figure 5A.

Supplemental Movie 7 Energetics of the interface between α5|$G\alpha_i$-GTPase in the R*-$G\alpha_i$ complex, relating to Figure 5B.

Supplemental Movie 8 Rotation of the α5 C-terminal helix of $G\alpha_i$ upon receptor binding. As α5 binds to activated rhodopsin (R*), α5 undergoes a 63º rotation $G\alpha_i$-GTPase. Asparagine 341 is shown as sticks for reference.

Supplemental Data

Supplemental Data 1 Ten model structures representing the receptor unbound model of $G\alpha_i\beta\gamma$ and corresponding summary PDB validation reports.

Supplemental Data 2 Nine model structures representing the R*-Gi complex and corresponding summary PDB validation reports.

Bibliography

1. Johnston, C. A.; Siderovski, D. P., Receptor-mediated activation of heterotrimeric G-proteins: Current structural insights. *Mol. Pharmacol.* **2007,** *72* (2), 219-230.
2. Konagurthu, A. S.; Whisstock, J. C.; Stuckey, P. J.; Lesk, A. M., MUSTANG: A multiple structural alignment algorithm. *Proteins: Structure, Function, and Bioinformatics* **2006,** *64* (3), 559-574.
3. Van Eps, N.; Preininger, A. M.; Alexander, N.; Kaya, A. I.; Meier, S.; Meiler, J.; Hamm, H. E.; Hubbell, W. L., Interaction of a G protein with an activated receptor opens the interdomain interface in the alpha subunit. *Proceedings of the National Academy of Sciences of the United States of America* **2011,** *108* (23), 9420-4.
4. Rasmussen, S. G. F.; DeVree, B. T.; Zou, Y.; Kruse, A. C.; Chung, K. Y.; Kobilka, T. S.; Thian, F. S.; Chae, P. S.; Pardon, E.; Calinski, D.; Mathiesen, J. M.; Shah, S. T. A.; Lyons, J. A.; Caffrey, M.; Gellman, S. H.; Steyaert, J.; Skiniotis, G.; Weis, W. I.; Sunahara, R. K.; Kobilka, B. K., Crystal structure of the [bgr]2 adrenergic receptor-Gs protein complex. *Nature* **2011,** *advance online publication*.
5. Kamarainen, J. K.; Kyrki, V.; Ilonen, J.; Kalviainen, H., Improving similarity measures of histograms using smoothing projections. *Pattern Recognit. Lett.* **2003,** *24* (12), 2009-2019.
6. Hirst, S. J.; Alexander, N.; McHaourab, H. S.; Meiler, J., RosettaEPR: an integrated tool for protein structure determination from sparse EPR data. *J Struct Biol* **2011,** *173* (3), 506-14.
7. Van Eps, N.; Preininger, A. M.; Alexander, N.; Kaya, A. I.; Meier, S.; Meiler, J.; Hamm, H. E.; Hubbell, W. L., Interaction of a G protein with an activated receptor opens the interdomain interface in the alpha subunit. *Proceedings of the National Academy of Sciences*.
8. Oldham, W. M.; Van Eps, N.; Preininger, A. M.; Hubbell, W. L.; Hamm, H. E., Mapping allosteric connections from the receptor to the nucleotide-binding pocket of heterotrimeric G proteins. *Proceedings of the National Academy of Sciences of the United States of America* **2007,** *104* (19), 7927-32.
9. Oldham, W. M.; Van Eps, N.; Preininger, A. M.; Hubbell, W. L.; Hamm, H. E., Mechanism of the receptor-catalyzed activation of heterotrimeric G proteins. *Nature structural & molecular biology* **2006,** *13* (9), 772-7.
10. Van Eps, N.; Oldham, W. M.; Hamm, H. E.; Hubbell, W. L., Structural and dynamical changes in an alpha-subunit of a heterotrimeric G protein along the activation pathway. *Proc Natl Acad Sci U S A* **2006,** *103* (44), 16194-9.
11. McHaourab, H. S.; Lietzow, M. A.; Hideg, K.; Hubbell, W. L., Motion of spin-labeled side chains in T4 lysozyme, correlation with protein structure and dynamics. *Biochemistry* **1996,** *35* (24), 7692-7704.
12. Chung, K. Y.; Rasmussen, S. G.; Liu, T.; Li, S.; DeVree, B. T.; Chae, P. S.; Calinski, D.; Kobilka, B. K.; Woods, V. L., Jr.; Sunahara, R. K., Conformational changes in the G protein Gs induced by the beta2 adrenergic receptor. *Nature* **2011,** *477* (7366), 611-5.
13. Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K., Features and development of Coot. *Acta Crystallographica Section D* **2010,** *66* (4), 486-501.
14. Lambright, D. G.; Sondek, J.; Bohm, A.; Skiba, N. P.; Hamm, H. E.; Sigler, P. B., The 2.0 A crystal structure of a heterotrimeric G protein. *Nature* **1996,** *379* (6563), 311-9.
15. Bower, M. J.; Cohen, F. E.; Dunbrack, R. L., Jr., Prediction of protein side-chain rotamers from a backbone-dependent rotamer library: a new homology modeling tool. *Journal of molecular biology* **1997,** *267* (5), 1268-82.
16. Bradley, P.; Misura, K. M.; Baker, D., Toward high-resolution de novo structure prediction for small proteins. *Science* **2005,** *309* (5742), 1868-71.
17. Rasmussen, S. G.; DeVree, B. T.; Zou, Y.; Kruse, A. C.; Chung, K. Y.; Kobilka, T. S.; Thian, F. S.; Chae, P. S.; Pardon, E.; Calinski, D.; Mathiesen, J. M.; Shah, S. T.; Lyons, J. A.; Caffrey, M.; Gellman, S. H.;

Steyaert, J.; Skiniotis, G.; Weis, W. I.; Sunahara, R. K.; Kobilka, B. K., Crystal structure of the beta2 adrenergic receptor-Gs protein complex. *Nature* **2011,** *477* (7366), 549-55.
18. PyMOL *The PyMOL Molecular Graphics System, Version 1.2r1*, Schrödinger, LLC.
19. Choe, H. W.; Kim, Y. J.; Park, J. H.; Morizumi, T.; Pai, E. F.; Krauss, N.; Hofmann, K. P.; Scheerer, P.; Ernst, O. P., Crystal structure of metarhodopsin II. *Nature* **2011,** *471* (7340), 651-5.
20. Konagurthu, A. S.; Whisstock, J. C.; Stuckey, P. J.; Lesk, A. M., MUSTANG: a multiple structural alignment algorithm. *Proteins* **2006,** *64* (3), 559-74.
21. (a) Barth, P.; Wallner, B.; Baker, D., Prediction of membrane protein structures with complex topologies using limited constraints. *Proceedings of the National Academy of Sciences of the United States of America* **2009,** *106* (5), 1409-14; (b) Barth, P.; Schonbrun, J.; Baker, D., Toward high-resolution prediction and design of transmembrane helical protein structures. *Proceedings of the National Academy of Sciences of the United States of America* **2007,** *104* (40), 15682-7.
22. McHaourab, H. S.; Lietzow, M. A.; Hideg, K.; Hubbell, W. L., Motion of spin-labeled side chains in T4 lysozyme. Correlation with protein structure and dynamics. *Biochemistry* **1996,** *35* (24), 7692-704.
23. Durham, E.; Dorr, B.; Woetzel, N.; Staritzbichler, R.; Meiler, J., Solvent accessible surface area approximations for rapid and accurate protein structure prediction. *J Mol Model* **2009**.
24. Konermann, L.; Tong, X.; Pan, Y., Protein structure and dynamics studied by mass spectrometry: H/D exchange, hydroxyl radical labeling, and related approaches. *J. Mass Spectrom.* **2008,** *43* (8), 1021-1036.
25. Kortemme, T.; Kim, D. E.; Baker, D., Computational alanine scanning of protein-protein interfaces. *Sci STKE* **2004,** *2004* (219), pl2.
26. Nugent, T.; Jones, D. T., Accurate de novo structure prediction of large transmembrane protein domains using fragment-assembly and correlated mutation analysis. *Proceedings of the National Academy of Sciences of the United States of America* **2012,** *109* (24), E1540-E1547.
27. Barth, P.; Schonbrun, J.; Baker, D., Toward high-resolution prediction and design of transmembrane helical protein structures. *Proceedings of the National Academy of Sciences of the United States of America* **2007,** *104* (40), 15682-15687.
28. Barth, P.; Wallner, B.; Baker, D., Prediction of membrane protein structures with complex topologies using limited constraints. *Proceedings of the National Academy of Sciences* **2009,** *106* (5), 1409-1414.
29. Westfield, G. H.; Rasmussen, S. G.; Su, M.; Dutta, S.; DeVree, B. T.; Chung, K. Y.; Calinski, D.; Velez-Ruiz, G.; Oleskie, A. N.; Pardon, E.; Chae, P. S.; Liu, T.; Li, S.; Woods, V. L., Jr.; Steyaert, J.; Kobilka, B. K.; Sunahara, R. K.; Skiniotis, G., Structural flexibility of the G alpha s alpha-helical domain in the beta2-adrenoceptor Gs complex. *Proceedings of the National Academy of Sciences of the United States of America* **2011,** *108* (38), 16086-91.