

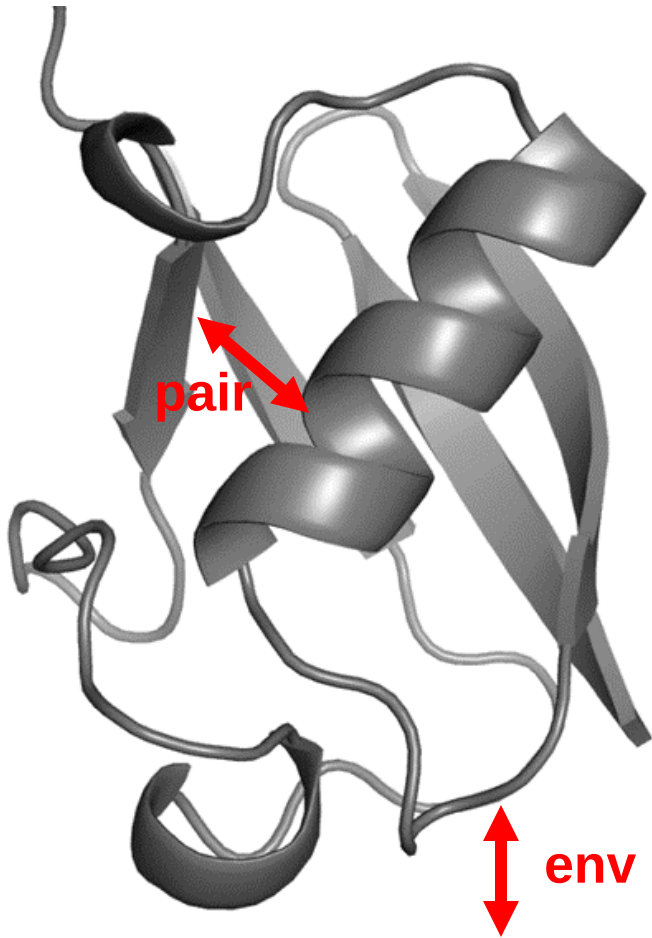
An Introduction to RosettaMembrane

Jens Meiler



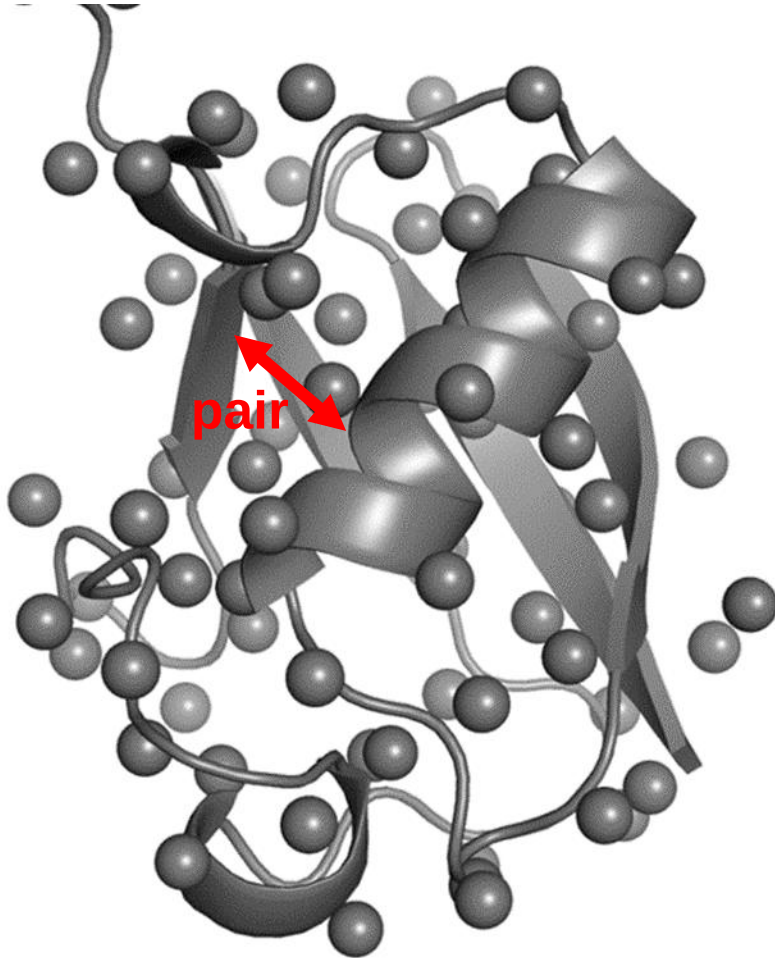
Center for Structural Biology and Institute of Chemical Biology
Departments of Chemistry, Pharmacology, and Biomedical Informatics

Scoring Function Terms



- Explicit: Interaction within the protein
 - Covalent interactions: disulfides
 - Electrostatic interactions
 - Hydrogen Bonds
 - Van der Waals interactions
- Implicit: Interactions with environment
 - Solvation
 - Radius of Gyration

Low Resolution Scoring Terms: Explicit Pairwise Interactions



Residue pair
interactions
(electrostatics,
disulfides)

Strand pairing
(hydrogen
bonding)

Strand
arrangement
into sheets

Helix-strand
packing

Steric repulsion

$$\sum_i \sum_{j>i} -\ln \left[\frac{P(aa_i, aa_j | s_{ij} d_{ij})}{P(aa_i | s_{ij} d_{ij}) P(aa_j | s_{ij} d_{ij})} \right]$$

$$\text{Scheme A : } SS_{\phi, \theta} + SS_{hb} + SS_d$$

$$\text{Scheme B : } SS_{\phi, \theta} + SS_{hb} + SS_{d\sigma}$$

where

$$SS_{\phi, \theta} = \sum_m \sum_{n>m} -\ln [P(\phi_{mn}, \theta_{mn} | d_{mn}, sp_{mn}, s_{mn})]$$

$$SS_{hb} = \sum_m \sum_{n>m} -\ln [P(hb_{mn} | d_{mn}, s_{mn})]$$

$$SS_d = \sum_m \sum_{n>m} -\ln [P(d_{mn} | s_{mn})]$$

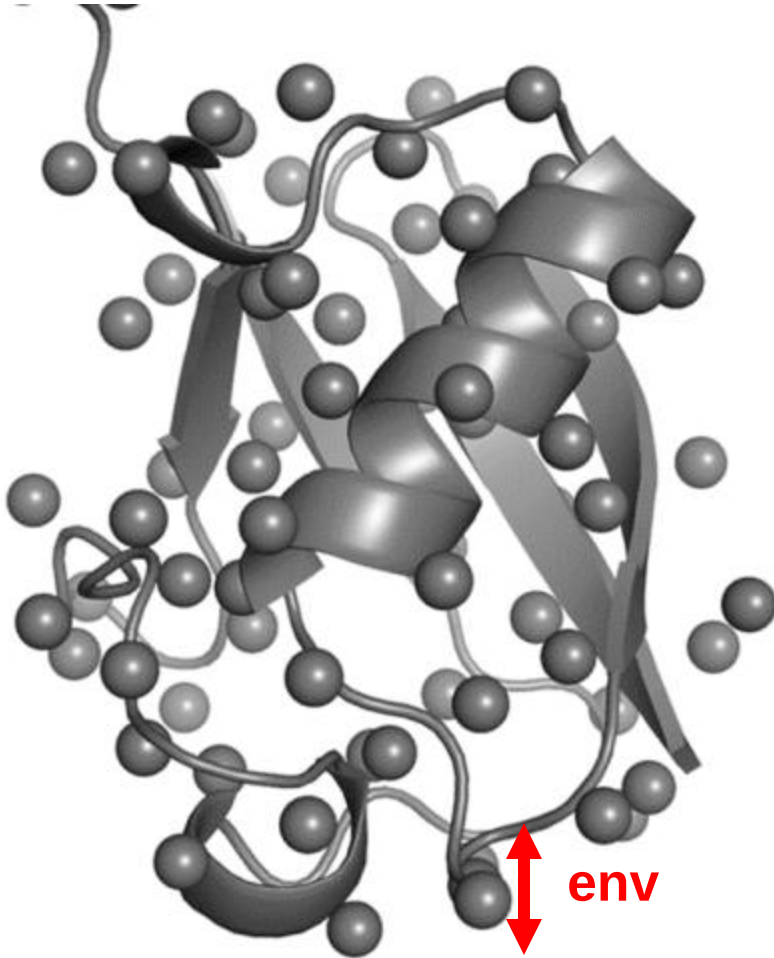
$$SS_{d\sigma} = \sum_m \sum_{n>m} -\ln [P(d_{mn} \sigma_{mn} | \rho_m, \rho_n)]$$

$$- \ln [P(n_{\text{sheets}} n_{\text{lonestrands}} | n_{\text{strands}})]$$

$$\sum_m \sum_n -\ln [P(\phi_{mn}, \psi_{mn} | sp_{mn} d_{mn})]$$

$$\sum_i \sum_{j>i} \frac{(r_{ij}^2 - d_{ij}^2)^2}{r_{ij}}; d_{ij} < r_{ij}$$

High Resolution Scoring Terms: Implicit Interactions with Solvent



Residue
environment
(solvation)

$$\sum_i -\ln [P(aa_i|nb_i)]$$

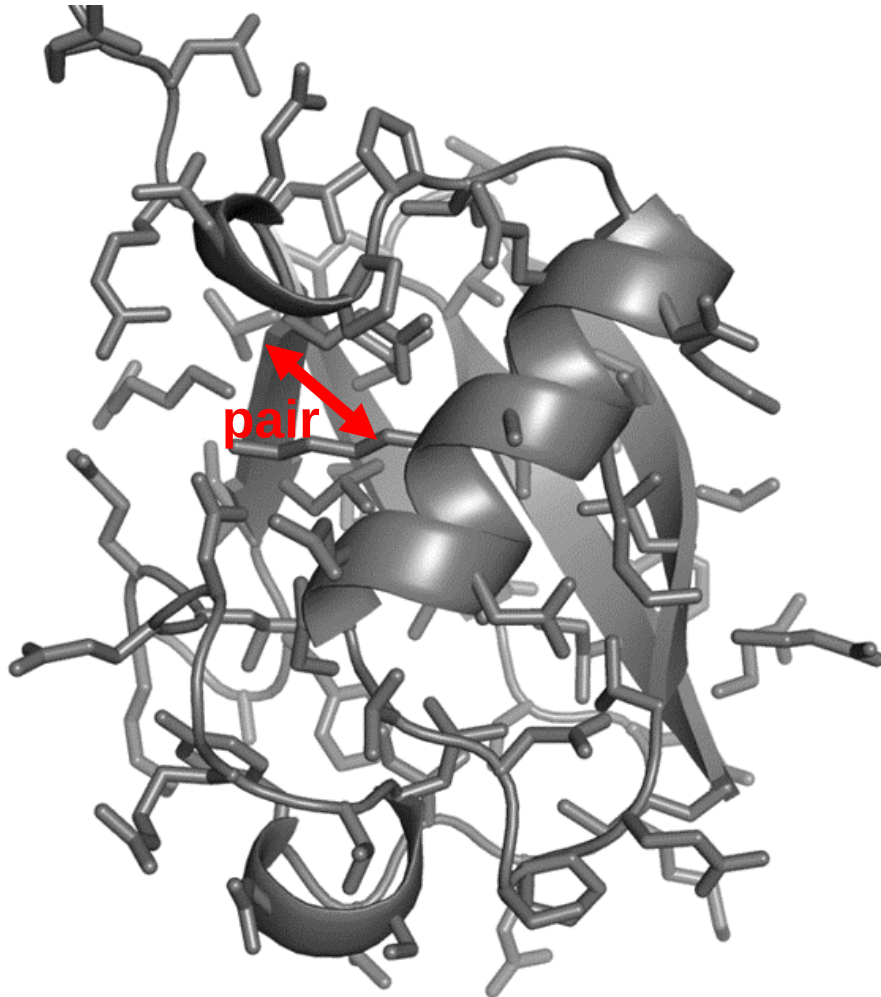
Radius of
gyration (vdw
attraction;
solvation)

$$\sqrt{\langle d_{ij}^2 \rangle}$$

C β density
(solvation;
correction
for excluded
volume effect
introduced by
simulation)

$$\sum_i \sum_{sh} -\ln \left[\frac{P_{\text{compact}}(nb_{i,sh})}{P_{\text{random}}(nb_{i,sh})} \right]$$

High Resolution Scoring Terms: Explicit Pairwise Interactions



Residue pair
interactions
(electrostatics,
disulfides)

$$\sum_i \sum_{j>i} -\ln \left[\frac{P(aa_i, aa_j|d_{ij})}{P(aa_i|d_{ij})P(aa_j|d_{ij})} \right]$$

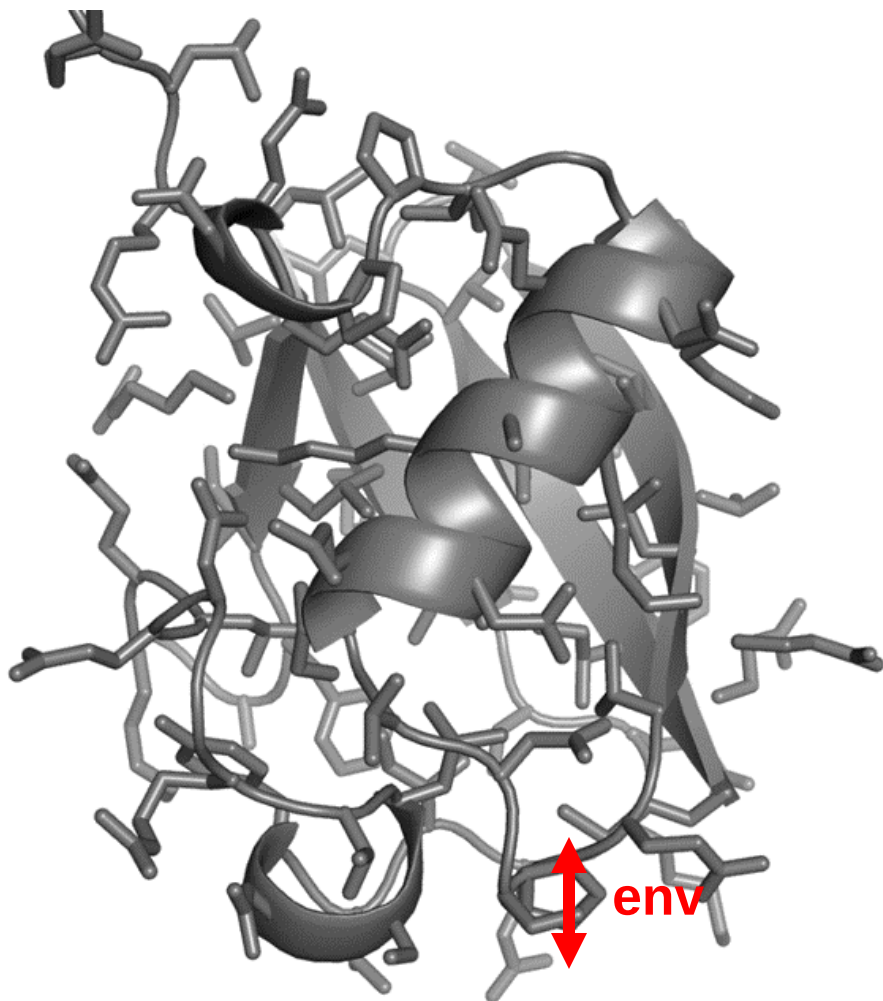
Hydrogen
bonding

$$\sum_i \sum_j (-\ln [P(d_{ij}|h_jss_{ij})] \\ -\ln [P(\cos \theta_{ij}|d_{ij}h_jss_{ij})] \\ -\ln [P(\cos \psi_{ij}|d_{ij}h_jss_{ij})])$$

Lennard-Jones
interactions

$$\sum_i \sum_{j>i} \begin{cases} \left[\left(\frac{r_{ij}}{d_{ij}} \right)^{12} - 2 \left(\frac{r_{ij}}{d_{ij}} \right)^6 \right] e_{ij}, & \text{if } \frac{d_{ij}}{r_{ij}} > 0.6 \\ -8759.2 \left(\frac{d_{ij}}{r_{ij}} \right) + 5672.0 \right] e_{ij}, & \text{else} \end{cases}$$

High Resolution Scoring Terms: Implicit Interactions with Solvent



Solvation

$$\sum_i \left[\Delta G_i^{\text{ref}} - \sum_j \left(\frac{2\Delta G_i^{\text{free}}}{4\pi^{3/2}\lambda_i r_{ij}^2} e^{-d_{ij}^2} V_j + \frac{2\Delta G_i^{\text{free}}}{4\pi^{3/2}\lambda_j r_{ij}^2} e^{-d_{ij}^2} V_i \right) \right]$$

Ramachandran torsion preferences

$$\sum_i -\ln [P(\phi_i, \psi_i | \text{aa}_i, \text{ss}_i)]$$

Rotamer self-energy

$$\sum_i -\ln \left[\frac{P(\text{rot}_i | \phi_i, \psi_i) P(\text{aa}_i | \phi_i, \psi_i)}{P(\text{aa}_i)} \right]$$

Unfolded state reference energy

$$\sum_{\text{aa}} n_{\text{aa}}$$

Implicit Solvation Potentials

- Implicit Solvation Models can be:
 - knowledge-based (RosettaMembrane low-res)
 - physics-based (IMM1 in CHARMM)

- SASA model

- $$\Delta G_{\text{solv}} = \sum_i \sigma_i \text{ASA}_i$$

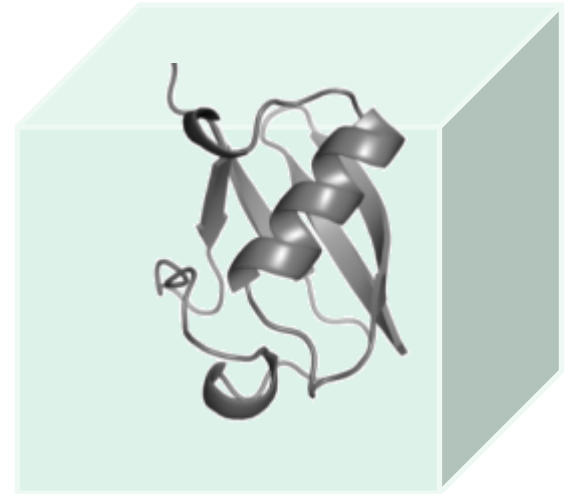
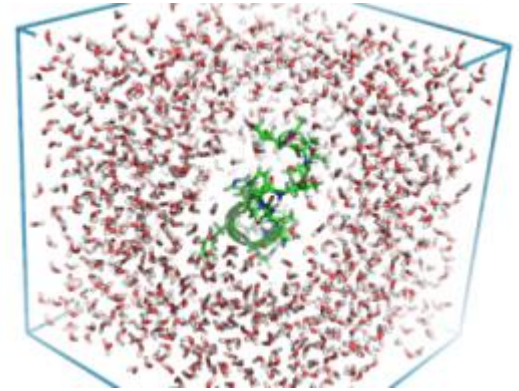
- Poisson-Boltzmann

- $$\vec{\nabla} \cdot [\epsilon(\vec{r}) \vec{\nabla} \Psi(\vec{r})] = -4\pi \rho^f(\vec{r}) - 4\pi \sum_i c_i^\infty z_i q \lambda(\vec{r}) e^{\frac{-z_i q \Psi(\vec{r})}{kT}}$$

- Generalized Born Approximation

- $$G_s = \frac{1}{8\pi} \left(\frac{1}{\epsilon_0} - \frac{1}{\epsilon} \right) \sum_{i,j} \frac{q_i q_j}{f_{GB}}$$

- $$f_{GB} = \sqrt{r_{ij}^2 + a_{ij}^2} e^{-D}$$



The Lazaridis – Karplus Model EEF1: A Model of Desolvation



- EEF1: Gaussian-shaped solvent exclusion

$$\Delta G_i^{solv} = \Delta G_i^{ref} - \sum_j \int_{V_j} f_i(r) dr$$

- unaccounted effects:
 - hydrophobic effect
 - viscosity of the solvent
 - H-bonds with solvent
- limitations:
 - choice of solvent is important
 - models are only tested on small proteins!!!
 - ionization of charged groups neglected

T. Lazaridis and M. Karplus; "Effective energy function for proteins in solution"; *Proteins*; **1999**; Vol. 35 (2): p. 133-152.

Layer and Burial Distinction for each Amino Acid

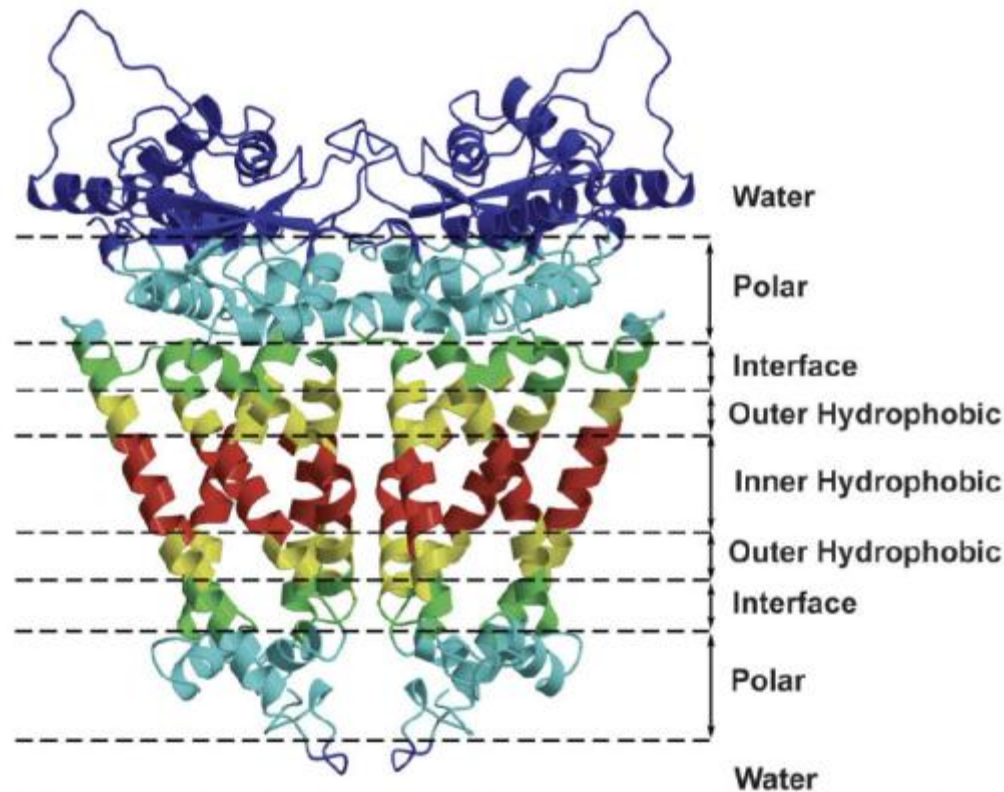
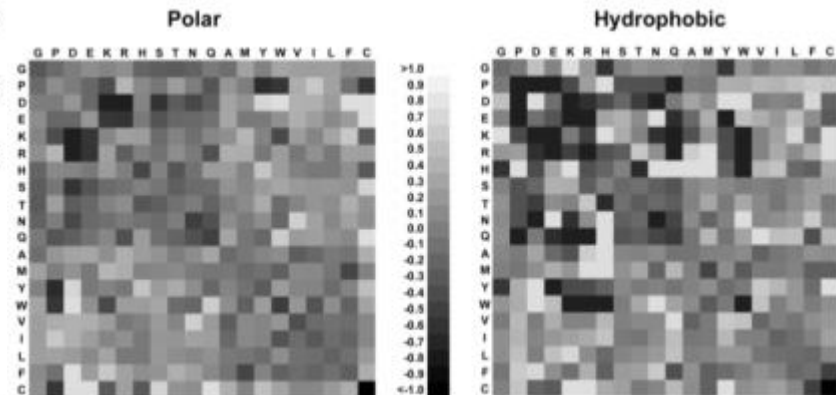


Fig. 1. Membrane layer definition. The lowest energy embedding of the fumarate reductase complex structure (PDB 1QLA) found using the method discussed in the text is colored as follows: "water-exposed" layer—dark blue; "polar" layer—light green; "interface" layer—green; "outer hydrophobic" layer—yellow; and "inner hydrophobic" layer—red.

$$E_{total} = E_{env} + E_{pair} + E_{clash} + E_{density} + E_{strand}$$

$$E_{env} = \sum_i -\ln\left(\frac{P(aa_i | L, B)}{P(aa_i)}\right)$$

$$E_{pair} = \sum_i \sum_{i < j} -\ln\left(\frac{P(aa_i, aa_j | d_{ij}, L)}{P(aa_i | d_{ij}, L) * P(aa_j | d_{ij}, L)}\right)$$

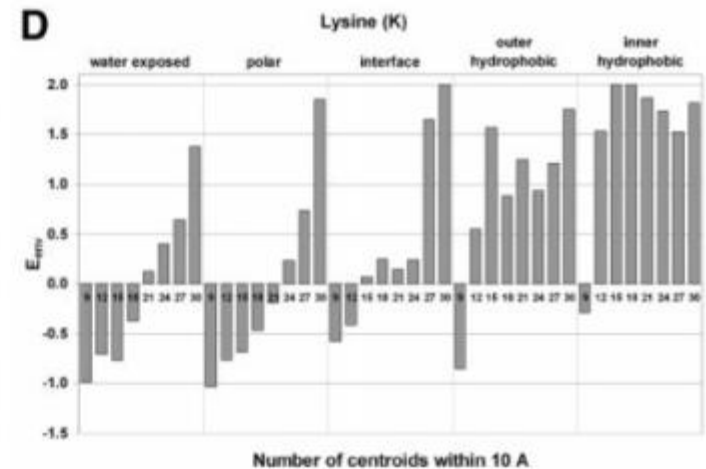
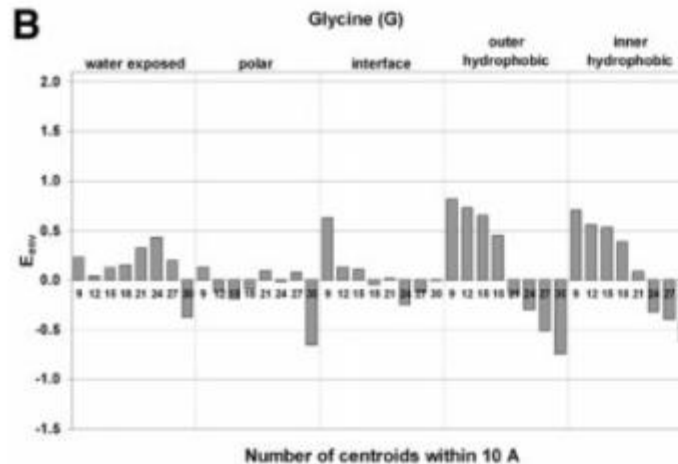
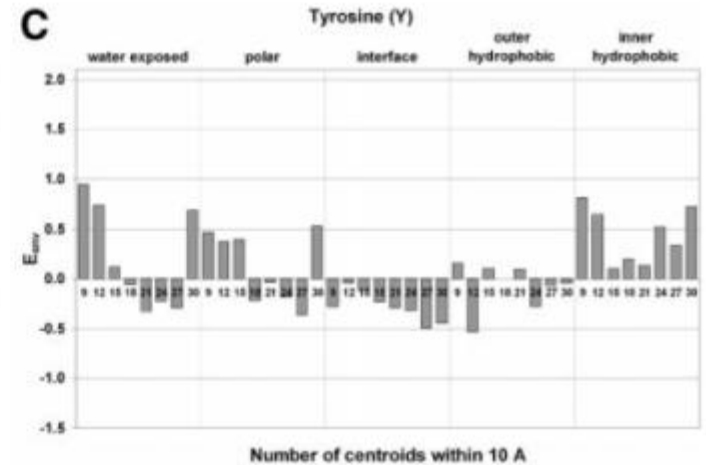
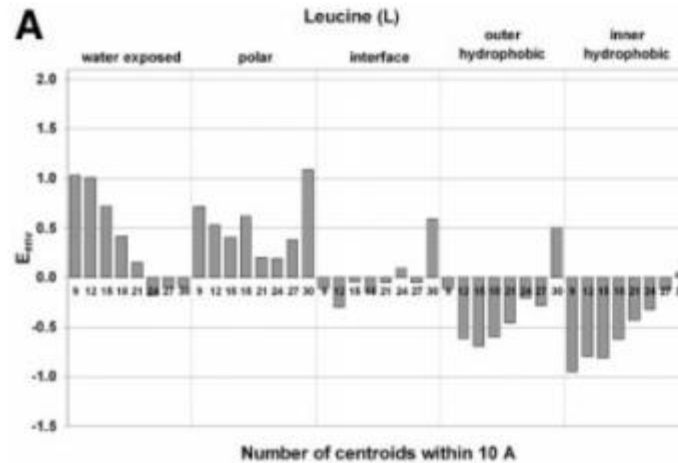


V. Yarov-Yarovoy, J. Schonbrun and D. Baker; "Multipass membrane protein structure prediction using Rosetta"; *Proteins*; **2006**; Vol. 62 (4): p. 1010-25.

Environment Potential for Selected Amino Acids



$$E_{env} = \sum_i -\ln\left(\frac{P(aa_i | L, B)}{P(aa_i)}\right)$$



Residue Density Profiles

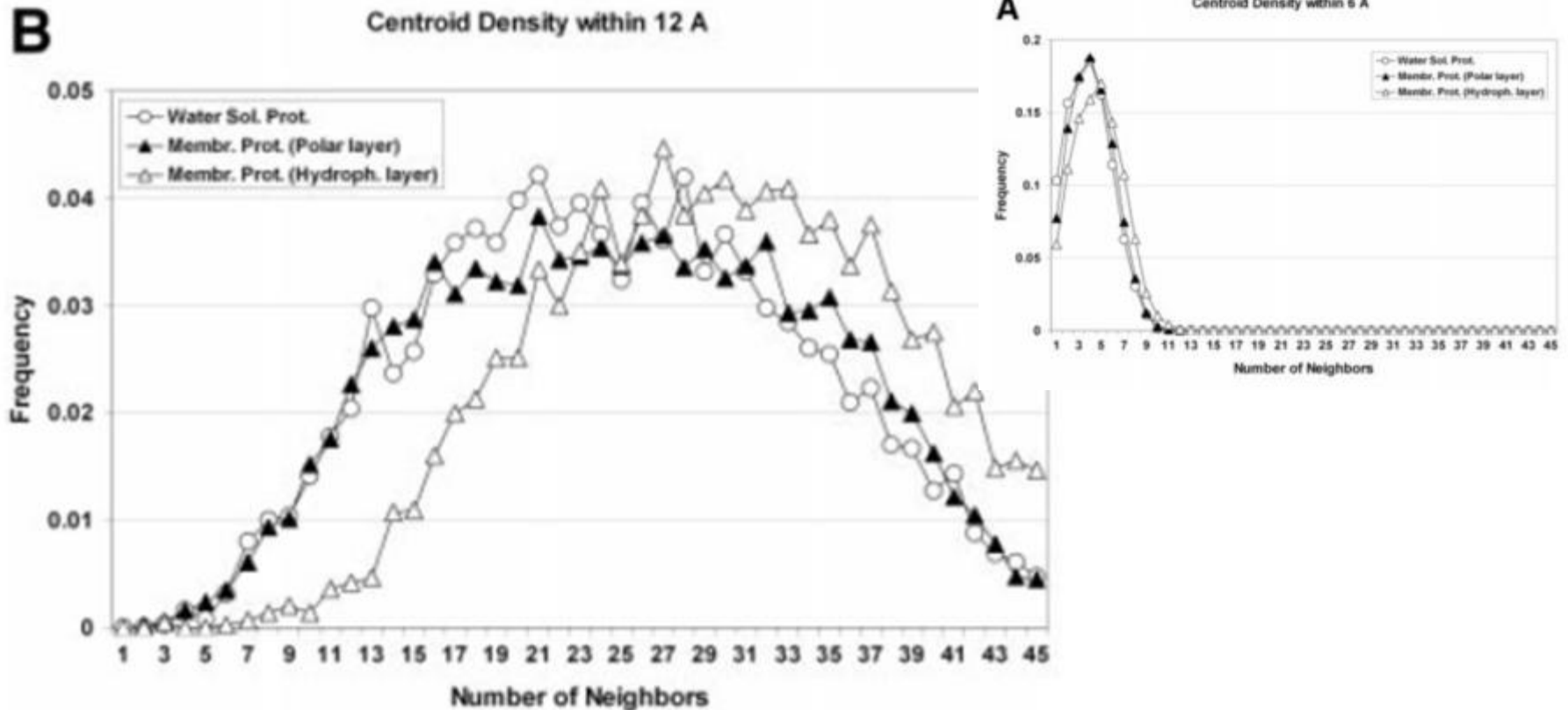


Fig. 3. Residue density profiles observed in the hydrophobic and polar layers of the membrane and α -helical type water-soluble proteins.

High Resolution Scoring Terms: Adapted from Solution Potential

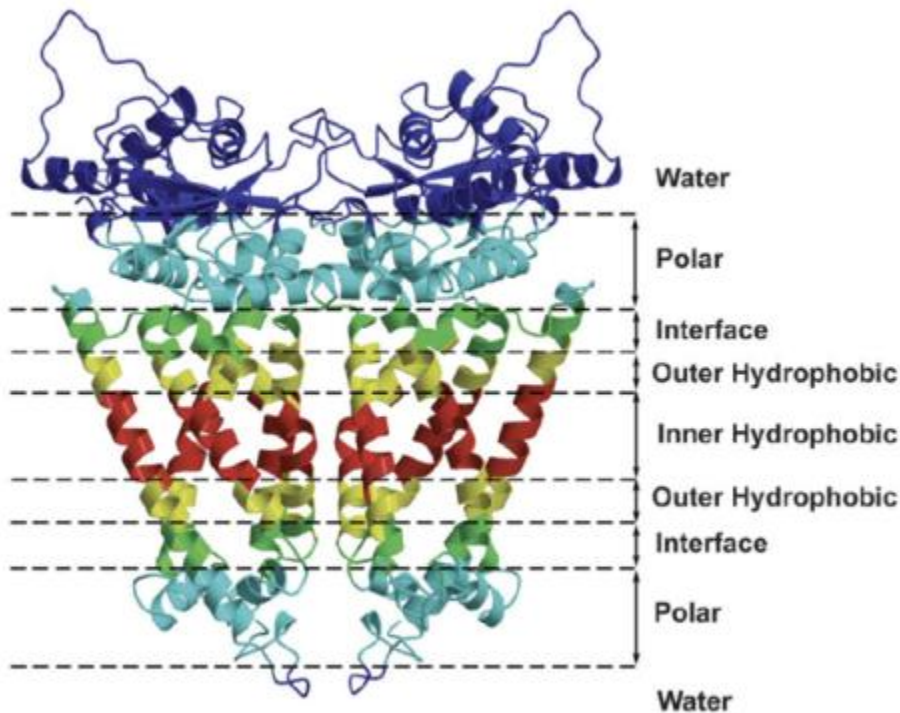


Fig. 1. Membrane layer definition. The lowest energy embedding of the fumarate reductase complex structure (PDB 1QLA) found using the method discussed in the text is colored as follows: "water-exposed" layer—dark blue; "polar" layer—light green; "interface" layer—green; "outer hydrophobic" layer—yellow; and "inner hydrophobic" layer—red.

H-bonding weight Increases in Strength
with Membrane Depth

Residue pair
interactions
(electrostatics,
disulfides)

$$\sum_i \sum_{j>i} -\ln \left[\frac{P(aa_i, aa_j|d_{ij})}{P(aa_i|d_{ij})P(aa_j|d_{ij})} \right]$$

Hydrogen
bonding

$$\sum_i \sum_j (-\ln [P(d_{ij}|h_j ss_{ij})] \\ -\ln [P(\cos \theta_{ij}|d_{ij} h_j ss_{ij})] \\ -\ln [P(\cos \psi_{ij}|d_{ij} h_j ss_{ij})])$$

Lennard-Jones
interactions

$$\sum_i \sum_{j>i} \left\{ \left[\left(\frac{r_{ij}}{d_{ij}} \right)^{12} - 2 \left(\frac{r_{ij}}{d_{ij}} \right)^6 \right] e_{ij}, \quad \text{if } \frac{d_{ij}}{r_{ij}} > 0.6 \right. \\ \left. -8759.2 \left(\frac{d_{ij}}{r_{ij}} \right) + 5672.0 \right] e_{ij}, \quad \text{else}$$

Solvation

$$\sum_i \left[\Delta G_i^{\text{ref}} - \sum_j \left(\frac{2\Delta G_i^{\text{free}}}{4\pi^{3/2} \lambda_j r_{ij}^2} e^{-d_{ij}^2} V_j \right. \right. \\ \left. \left. + \frac{2\Delta G_i^{\text{free}}}{4\pi^{3/2} \lambda_j r_{ij}^2} e^{-d_{ij}^2} V_i \right) \right]$$

Ramachandran torsion
preferences

$$\sum_i -\ln [P(\phi_i, \psi_i|aa_i, ss_i)]$$

Rotamer
self-energy

$$\sum_i -\ln \left[\frac{P(\text{rot}_i|\phi_i, \psi_i)P(aa_i|\phi_i, \psi_i)}{P(aa_i)} \right]$$

Unfolded state
reference energy

$$\sum_{aa} n_{aa}$$



IMM1 – Implicit Membrane Model 1

- extension of the EEF1 energy function (Gaussian-shaped solvent exclusion)
- IMM1 includes:
 - solvation parameters for non-polar phase derived from experimental data for transfer from water to cyclohexane
 - reference solvation free energy dependent on z-coord
 - distance-dependent dielectric model accounts for reduced electrostatic interactions within the membrane
 - adjustment of side-chain pair-potential to account for potentials of mean force
- electrostatic model is adjusted to match experimental observations for polyAla, polyLeu, glycophorin A

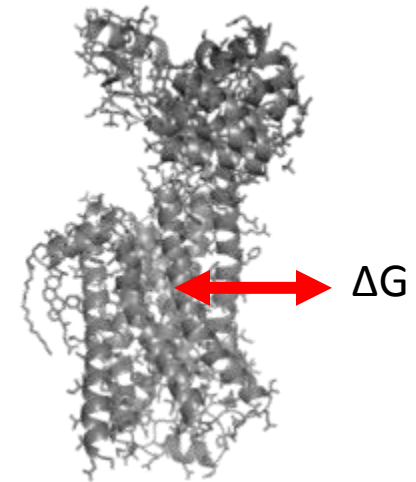
T. Lazaridis; "Effective energy function for proteins in lipid membranes"; *Proteins*; **2003**; Vol. 52 (2): p. 176-92.

Solvation is Sum of Desolvation and Membrane Embedding



- Desolvation term:

$$E_{\text{solv}} = - \sum_i^{\text{natom}} \sum_{j>i}^{\text{natom}} \left\{ \frac{2\Delta G_i^{\text{free}}}{4\pi\sqrt{\pi}\lambda_i r_{ij}^2} \exp(-d_{ij}^2) V_j + \frac{2\Delta G_j^{\text{free}}}{4\pi\sqrt{\pi}\lambda_j r_{ij}^2} \exp(-d_{ji}^2) V_i \right\}$$



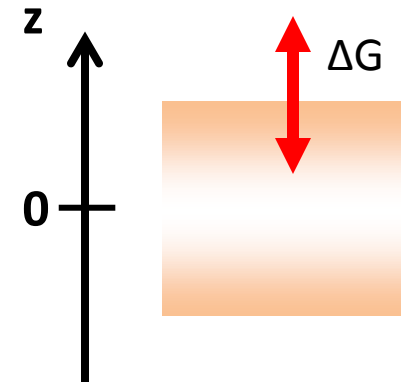
- Embedding term:

$$E_{\text{mbenv}} = \sum_i^{\text{natom}} \Delta G_i^{\text{ref}}(z')$$

$$\Delta G_i^{\text{free}}(z') = f(z') \Delta G_i^{\text{free,water}} + (1 - f(z')) \Delta G_i^{\text{free,hex}}$$

$$\Delta G_i^{\text{ref}}(z') = (1 - f(z')) * (\Delta G_i^{\text{ref,hex}} - \Delta G_i^{\text{ref,water}})$$

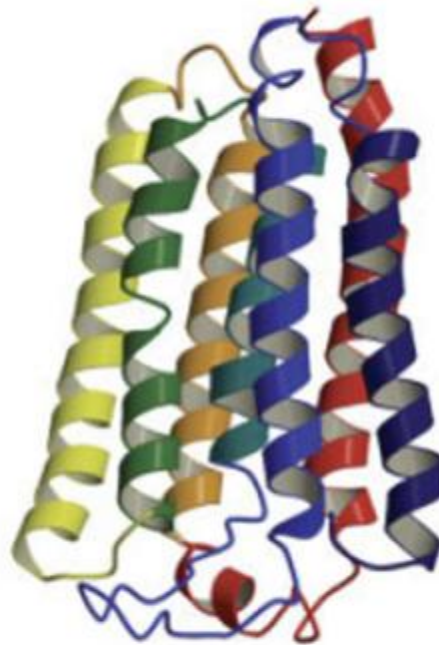
$$f(z') = \frac{z'^n}{1 + z'^n} \quad z' = |z|/(T/2)$$



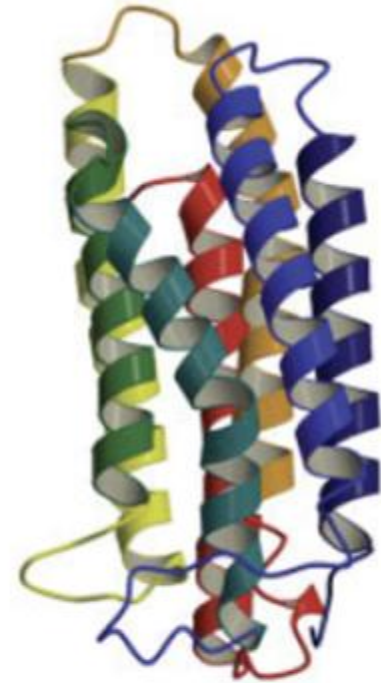
7TM helix bacteriorhodopsin



Native BRD7



Lowest RMSD

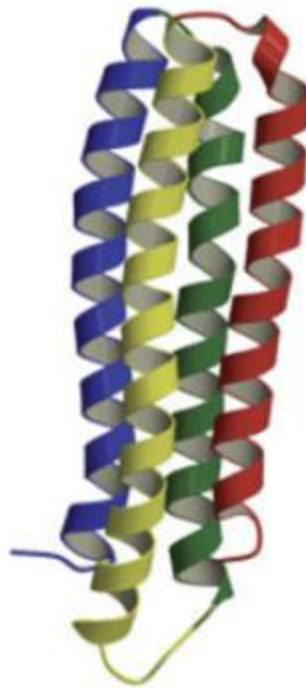


Top5 Cluster Center

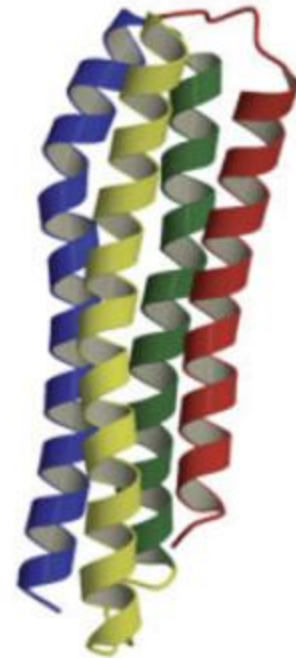
4 TM helix subdomain of V-type Na-ATPase (VATP)



Native VATP

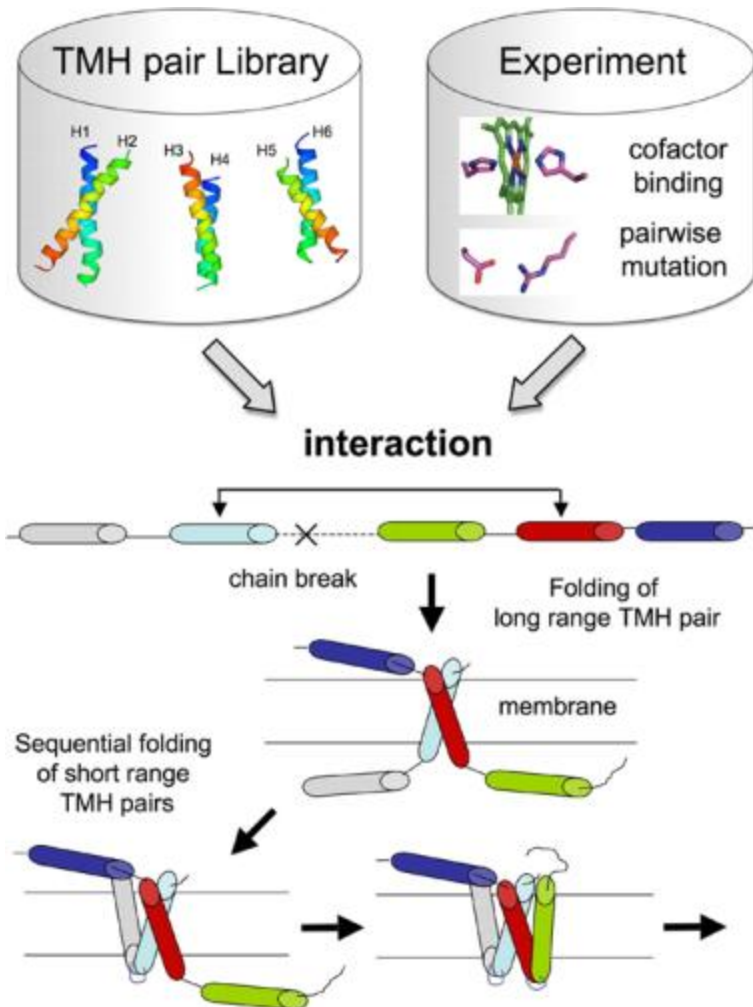


Lowest RMSD



Top5 Cluster Center

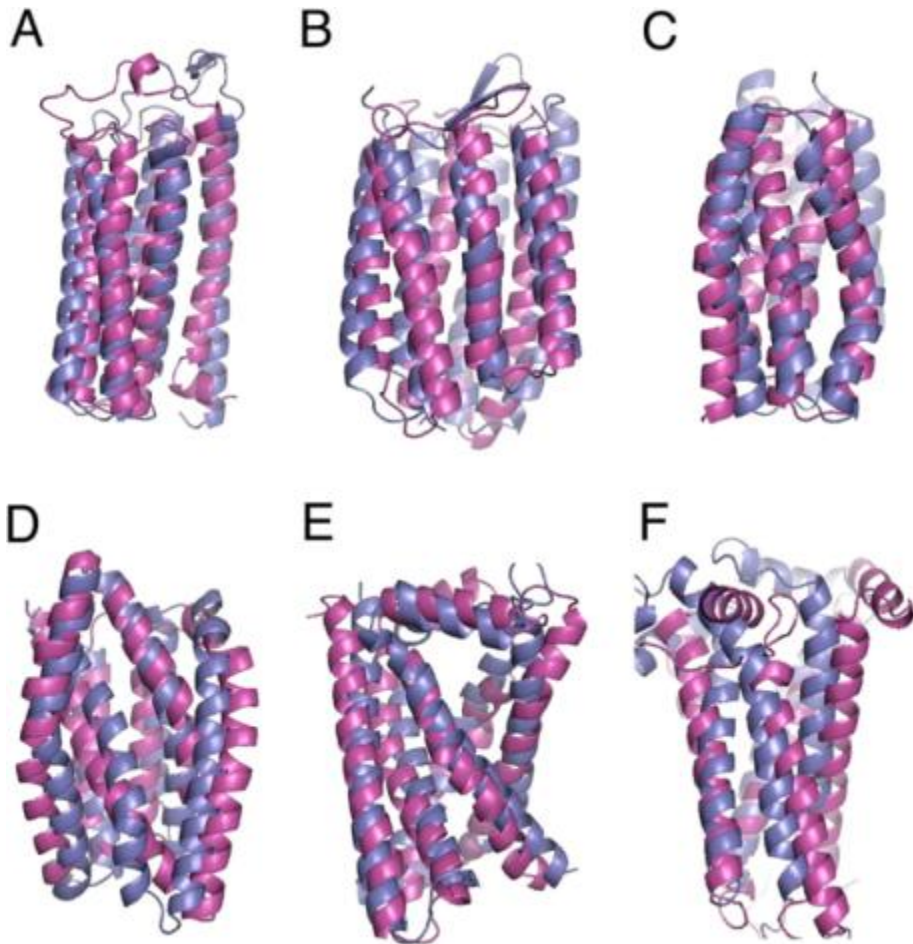
Prediction of membrane protein structures using limited constraints



Ab initio folding protocol with long-range interactions. Interactions can be predicted from sequence information using a database of TMH pairs of known structure or can be inferred from experiments. Once an interaction is selected, the two helices connected through space by that interaction are inserted and folded in the membrane. Adjacent individual TMHs are then randomly selected and folded in the membrane by Monte-Carlo fragment insertion sampling. After all TMHs are assembled in the membrane, the initial chain break is closed.

P. Barth, B. Wallner and D. Baker; "Prediction of membrane protein structures with complex topologies using limited constraints"; *Proc Natl Acad Sci U S A*; **2009**; Vol. 106 (5): p. 1409-14.

Prediction of membrane protein structures using limited constraints

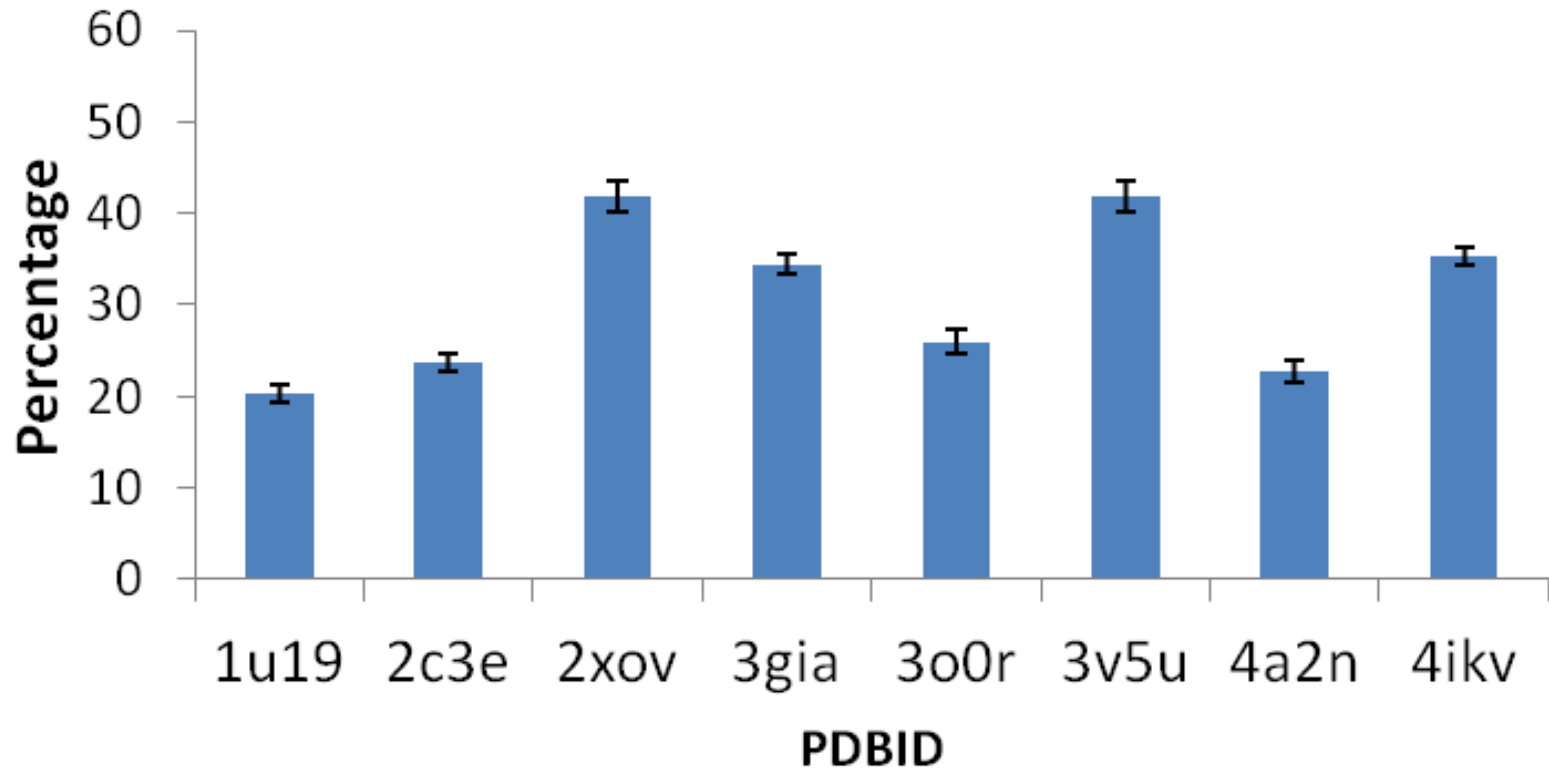


Prediction of membrane protein structures. Superposition between the most accurate models of the 5 lowest all-atom energy models (magenta) and X-ray structure of: chain A of cytochrome c (A), Bacteriorhodopsin (B), chain H of fumarate reductase (E), and chain D of cytochrome bc1 (F). The cluster size was used as an initial filter for the selection of the models.

Benchmarking Full Design of MPs



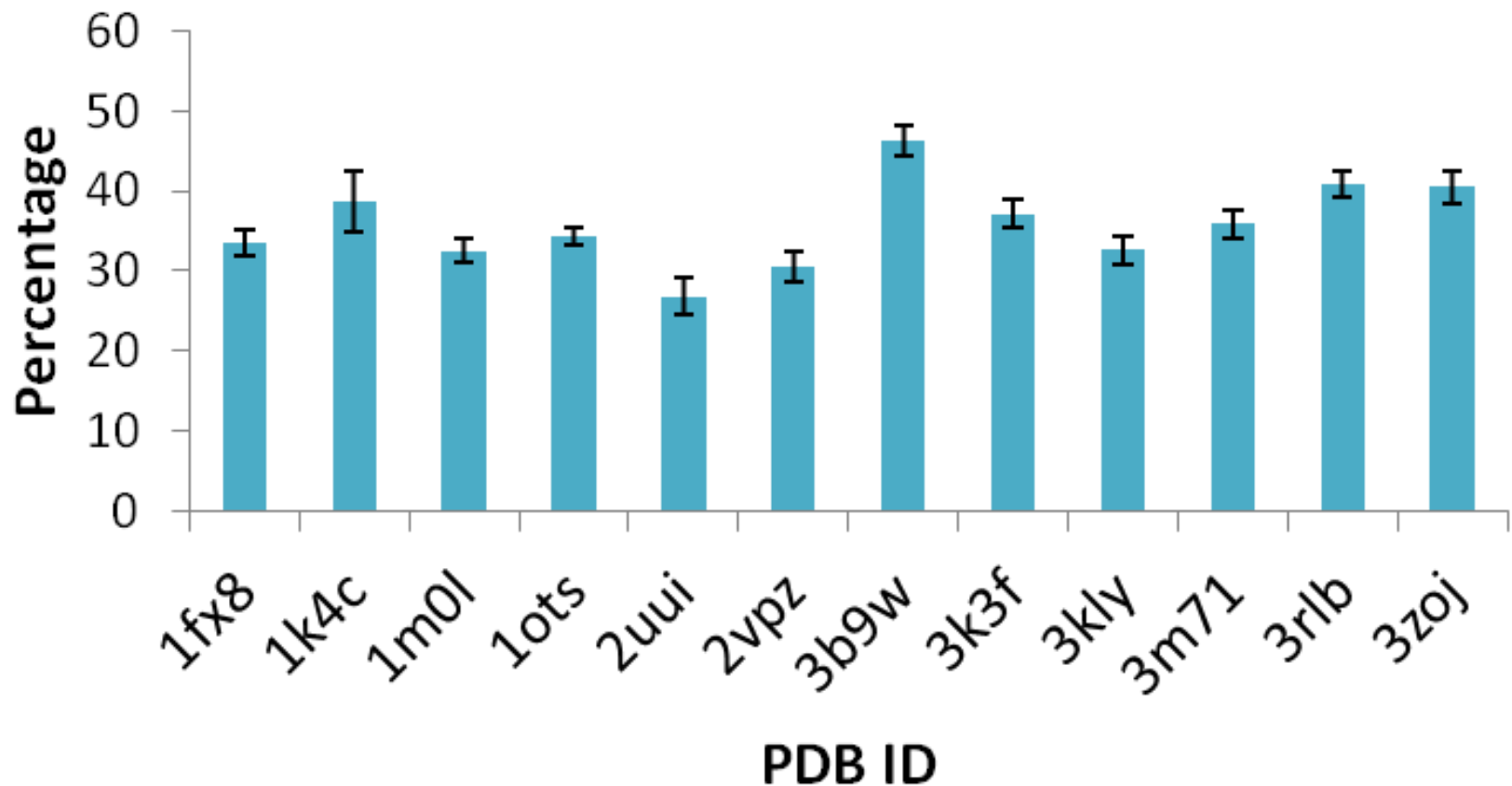
Sequence Recovery in Monomers



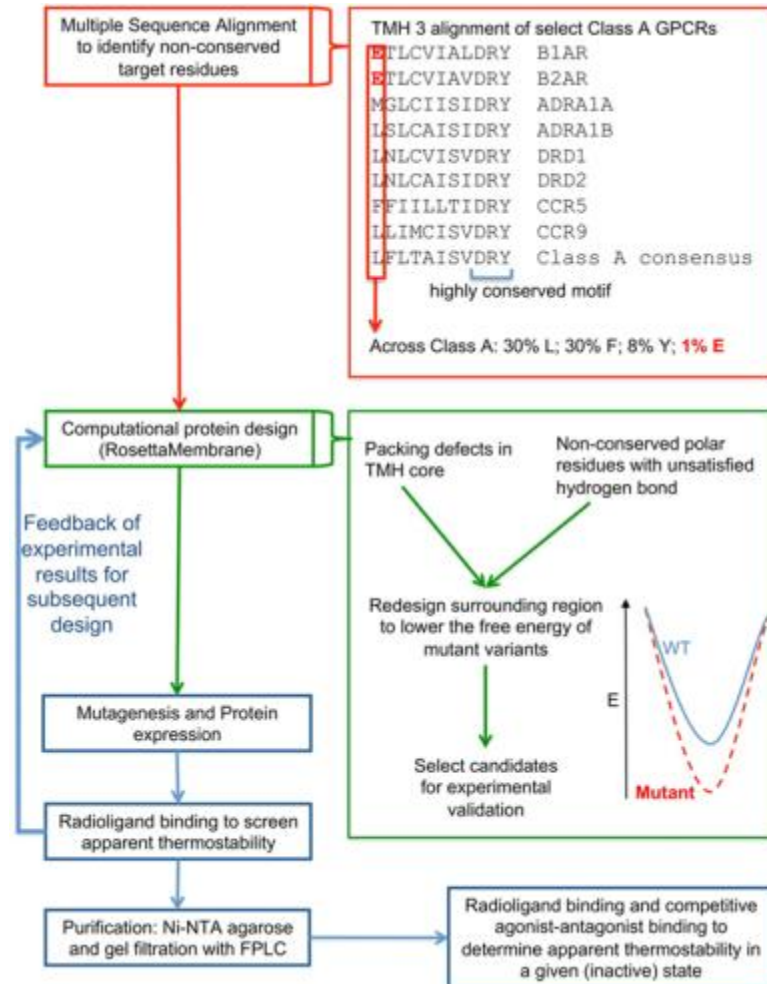
Benchmarking Full Design of MPs



Sequence Recovery in Multimers



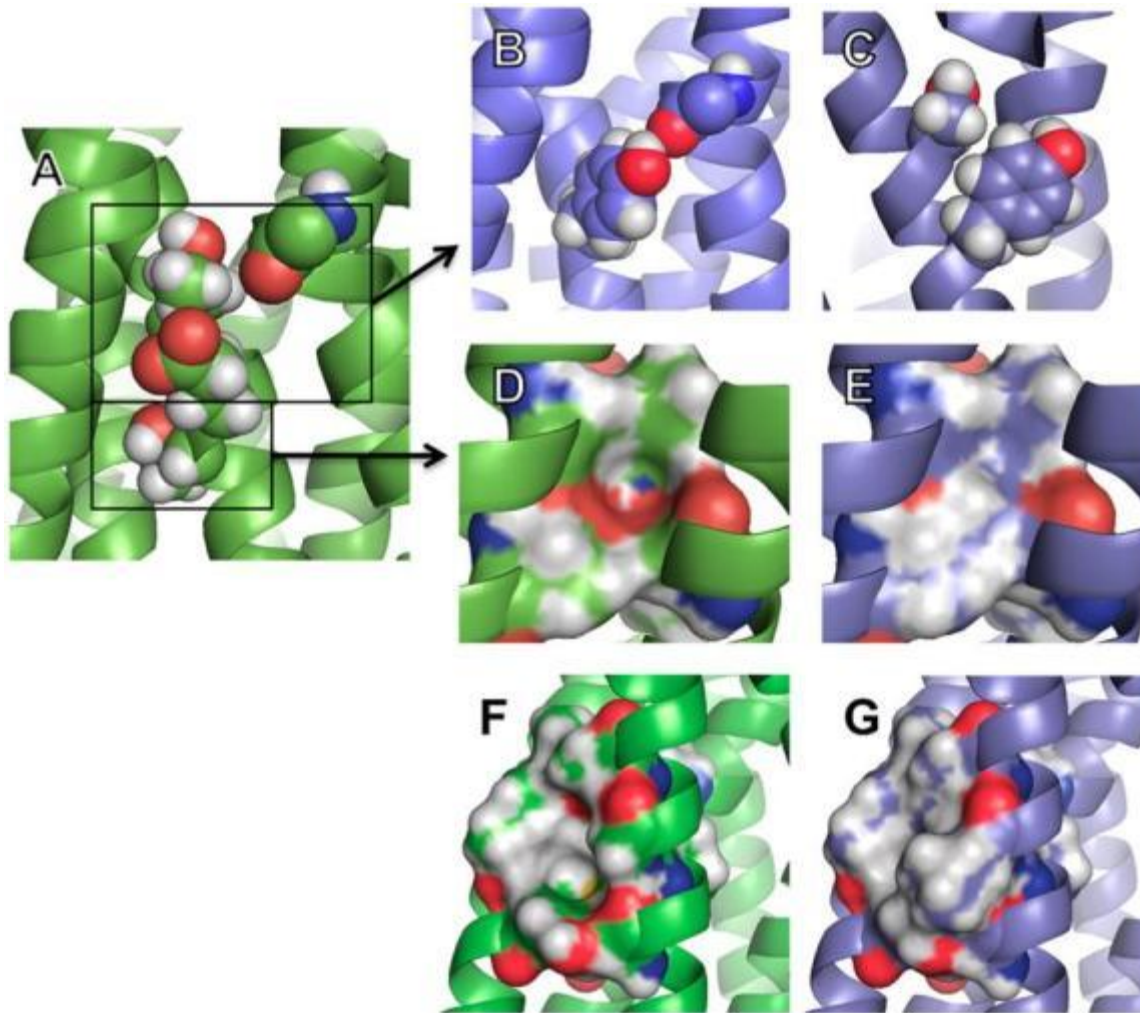
GPCR Thermostabilization



Integrated computational/experimental approach to design stabilized membrane proteins. (*Top*, red) Multiple sequence alignments (MSA) across class A GPCRs are analyzed to identify potential target sites for stabilization (i.e., nonconserved residues). *Top Inset*: MSA of Class A GPCRs transmembrane helix 3 (TM3) exhibit a highly conserved motif (DRY) and a nonconserved polar residue specific to beta-adrenergic receptors (E3.41). (*Middle*, green) Structural analysis is first performed to (i) select unsatisfied among nonconserved polar residues (i.e., nonstabilized by hydrogen bond or electrostatic interactions) and (ii) identify local regions with packing defects. RosettaMembrane is then used to redesign these target regions to identify amino-acid substitutions minimizing the energy of the inactive B1AR structure (middle inset). Designs are subsequently refined and filtered for false positives. (*Bottom*, blue) Rapid experimental validation using analytical radioligand binding assays determines the apparent melting temperature (T_m) of selected receptor variants. The results are fed back into the design process for subsequent rounds of computational design combining and accommodating locally stabilized regions. The variants resulting in the highest apparent conformational stability are purified to accurately measure their apparent T_m and to assess their conformational stabilization using agonist-antagonist competition binding assays.

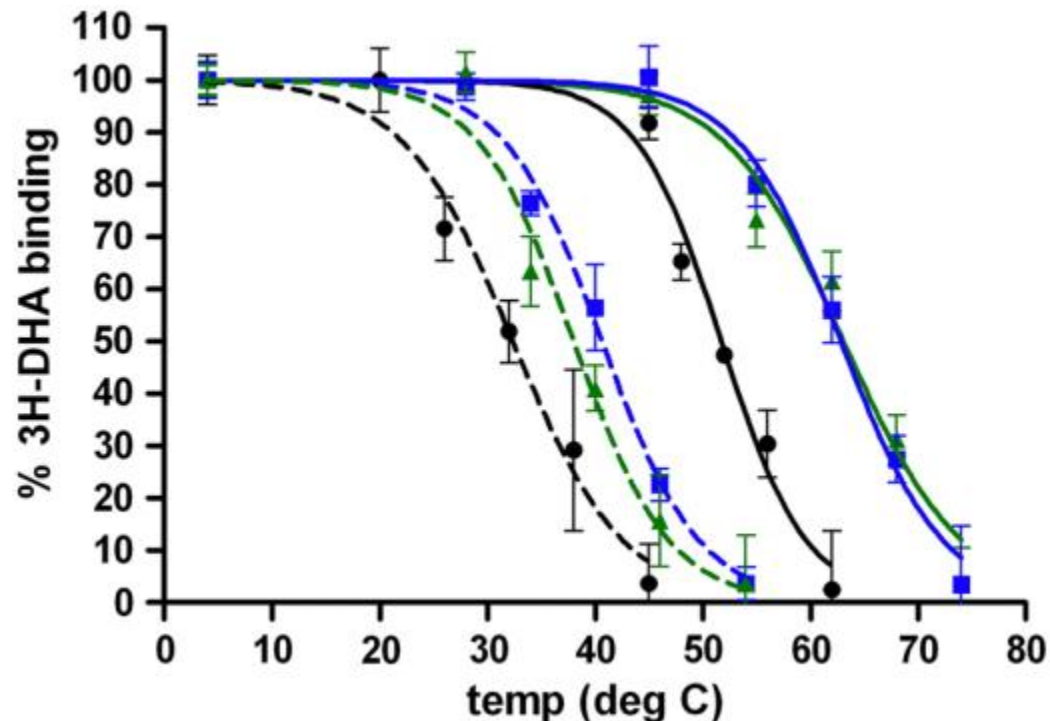
K. Y. Chen, F. Zhou, B. G. Fryszczyn and P. Barth; "Naturally evolved G protein-coupled receptors adopt metastable conformations"; *Proc Natl Acad Sci U S A*; **2012**; Vol. 109 (33): p. 13284-9.

GPCR Thermostabilization



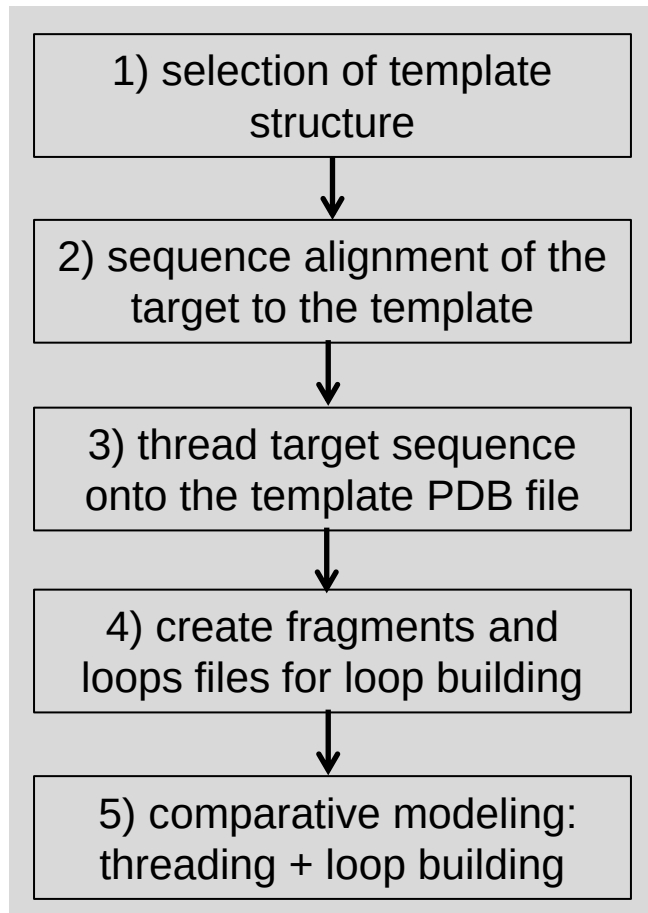
Conformational stabilization of B1AR by optimization of polar and van der Waals interhelical interactions. WT and designed structures are colored in green and blue, respectively. (A–E) Unsatisfied polar residues and packing defects at TM2/4/5 interface. (A) E3.41 and T3.37 are nonconserved and unsatisfied polar residues in B1AR WT. (B) Y3.41 improves the packing, eliminates an unsatisfied polar group, and creates a close to optimal hydrogen bond with a backbone unsatisfied carbonyl group on TM4. (C) S3.37, unlike T3.37, accommodates Y3.41 conformation. (D) Surface representation of T3.42 and associated packing defect. (E) I3.42 improves the packing and eliminates an unsatisfied polar group. (F and G) Packing defects at TM5/6 interface. (F) G6.42 and T6.45 create a packing defect between TM5 and TM6 in B1AR WT. (G) The redesigned stabilized region consists of two substitutions: L6.42 and I6.45, which optimize the VDW interactions between TM5 and TM6.

GPCR Thermostabilization

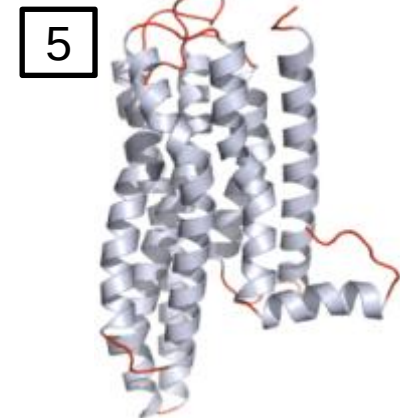
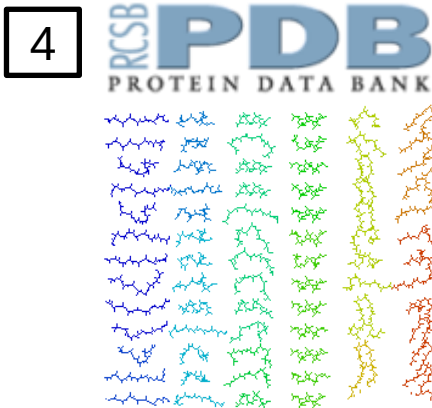
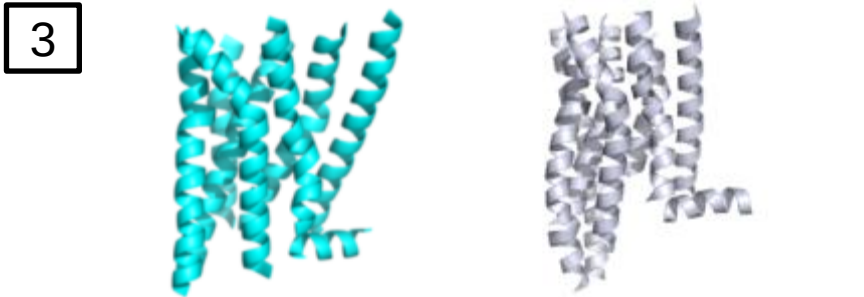


Hyperstabilized B1AR variants. Melting curves of B1AR variants were obtained by measuring tritiated antagonist (3H-DHA) binding to detergent solubilized receptor incubated at increasing temperatures. When designed in WT B1AR (dotted lines), 5 m (blue square) and 7 m (green triangle) variants exhibited T_m increases of 8 and 6 °C over B1AR-WT (black circle), respectively. When designed in the M23 background (solid lines), 5 m (blue square), and 7 m (green triangle) variants both exhibited T_m increases of 11 °C over B1AR-M23 (black circle, solid line) and 31 °C over B1AR-WT (black circle, dotted line). Error bars show the standard errors $N = 4$.

Comparative modeling & docking of GPCRs

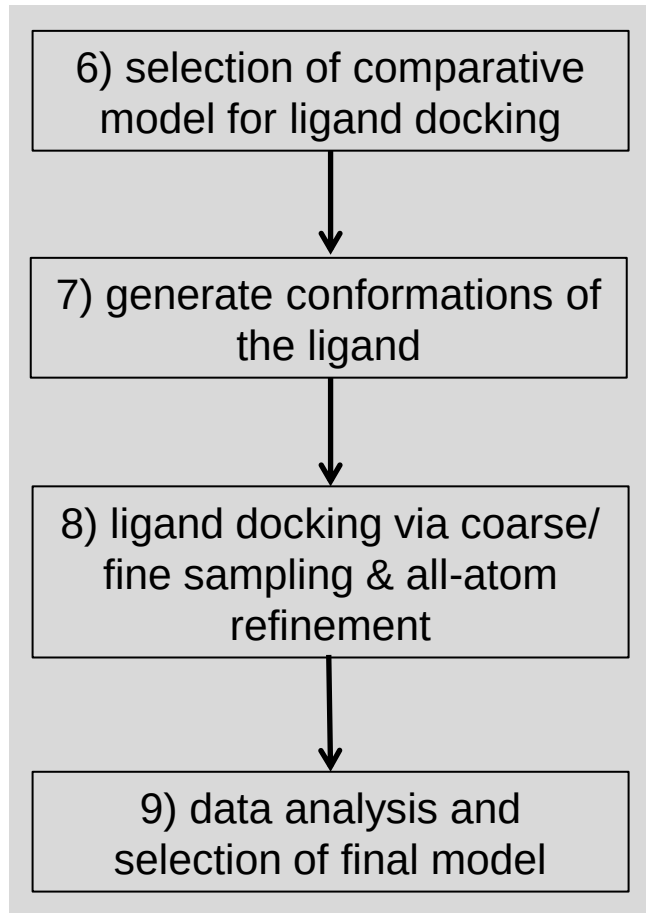


2 target 1 MCFSVSLSATVALGCMFVPKVY
template 1 KEVYILLNWIGYVNSGFNPLIT

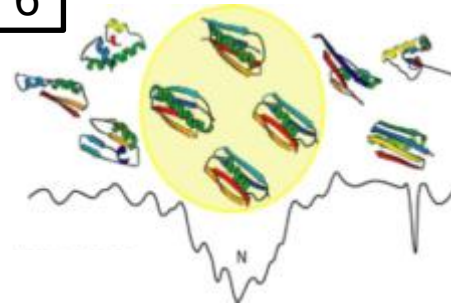


Combs, S.A., et al (2013) Nat Protoc, 1(7), 1277-98.
Nguyen, E.D., et al (2013) PLOS ONE 8(7):e67302

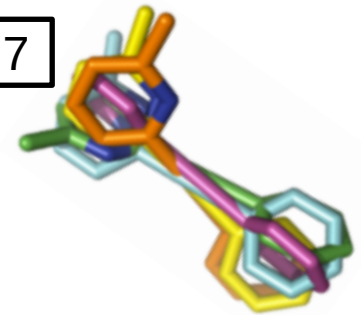
Comparative modeling & docking of GPCRs



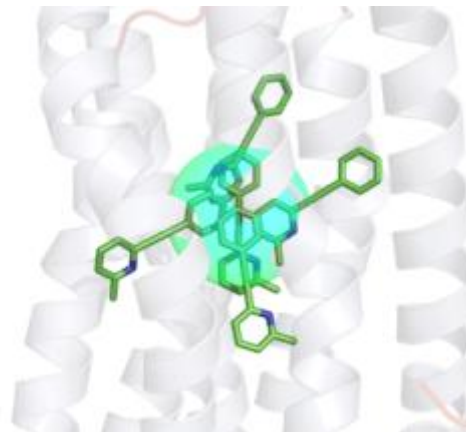
6



7



8



9

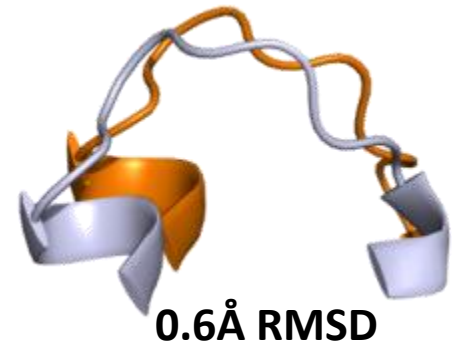
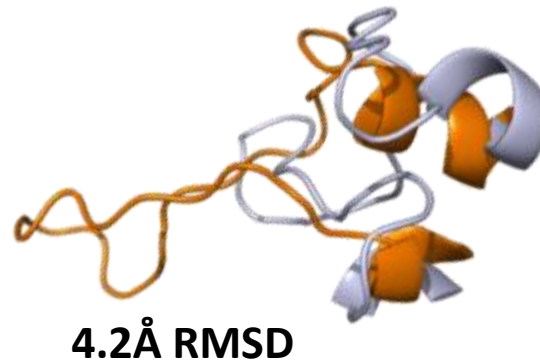


Combs, S.A., et al (2013) Nat Protoc, 1(7), 1277-98.
Nguyen, E.D., et al (2013) PLOS ONE 8(7):e67302

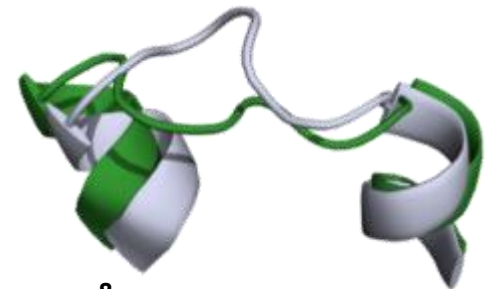
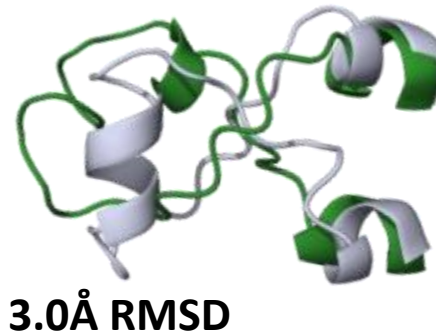
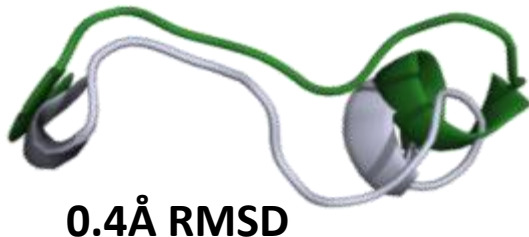
Rosetta recovers native loop structures in comparative models



Rhodopsin



β_1 Adrenergic

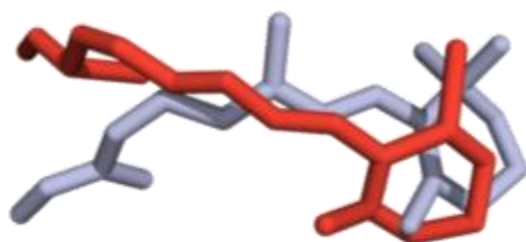


EC1 (7-8 residues)

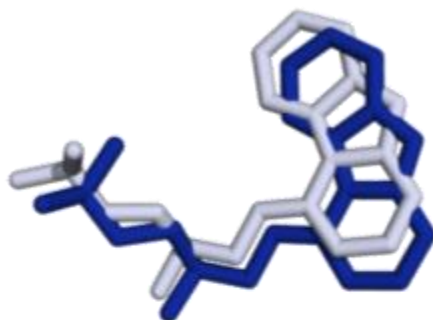
EC2 (17-34 residues)

EC3 (8-10 residues)

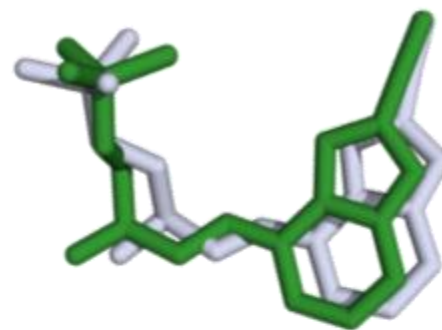
RosettaLigand docking recovers ligand binding poses



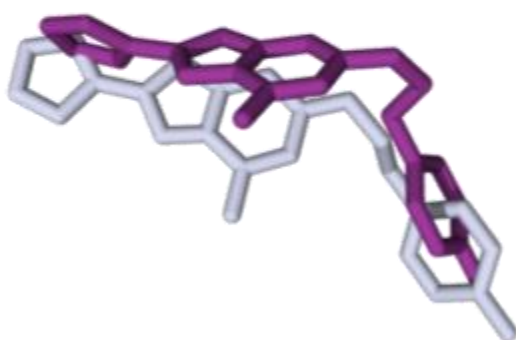
Retinal docked into
rhodopsin (3.7Å RMSD)



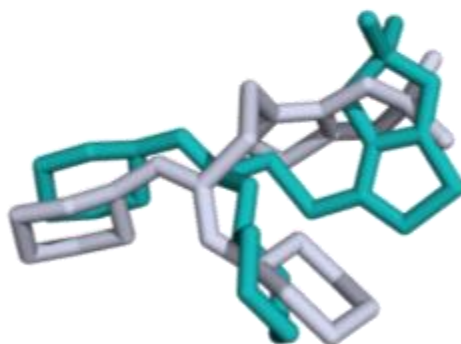
Carazolol docked into β_2 AR
(1.7Å RMSD)



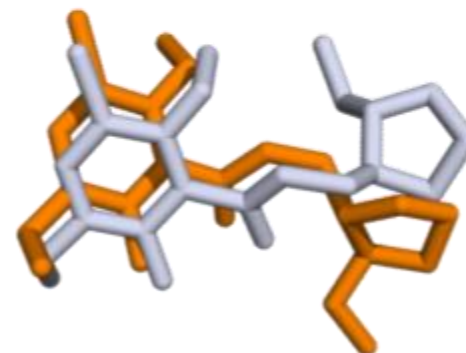
Cyanopindolol docked
into β_1 AR (1.9Å RMSD)



ZM241385 docked into A2A
adenosine (2.5Å RMSD)

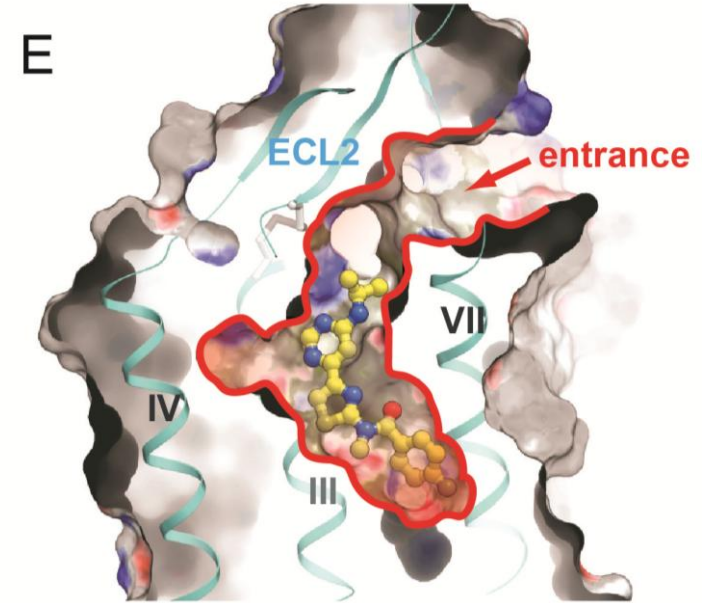
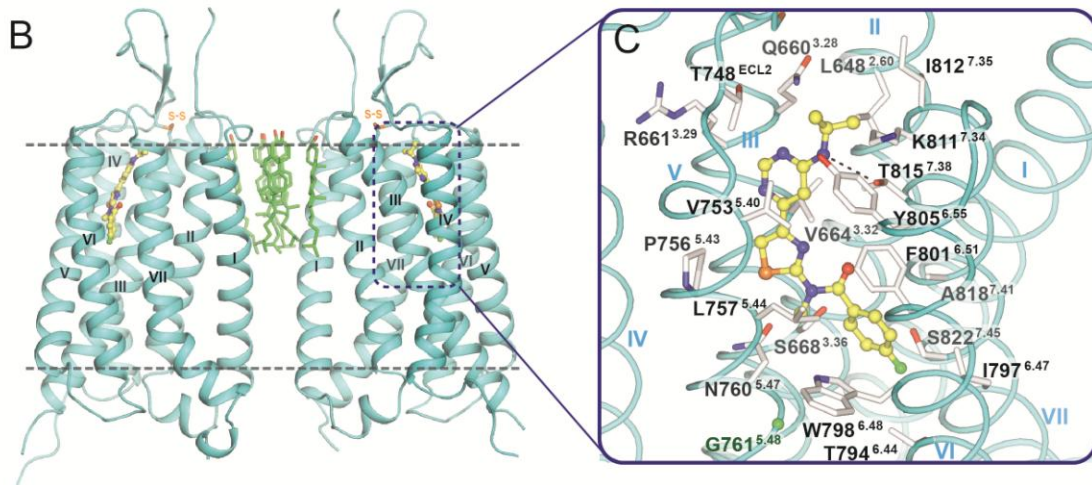
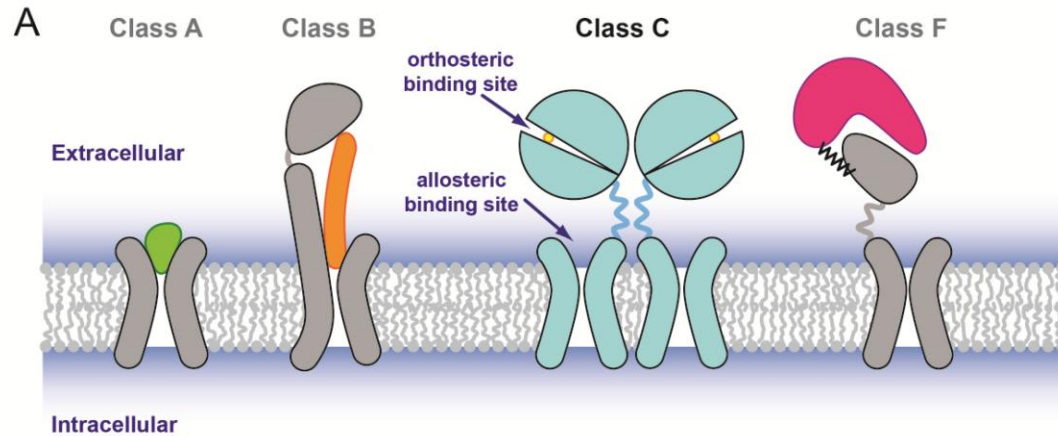


IT1t docked into CXCR4
(3.7Å RMSD)



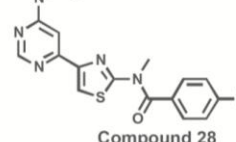
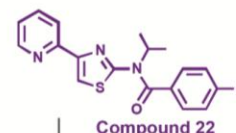
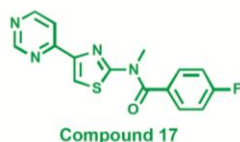
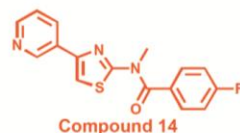
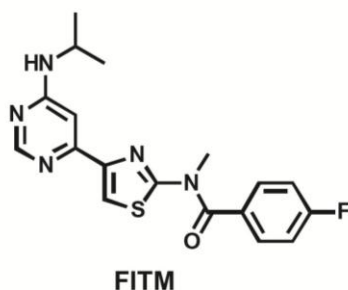
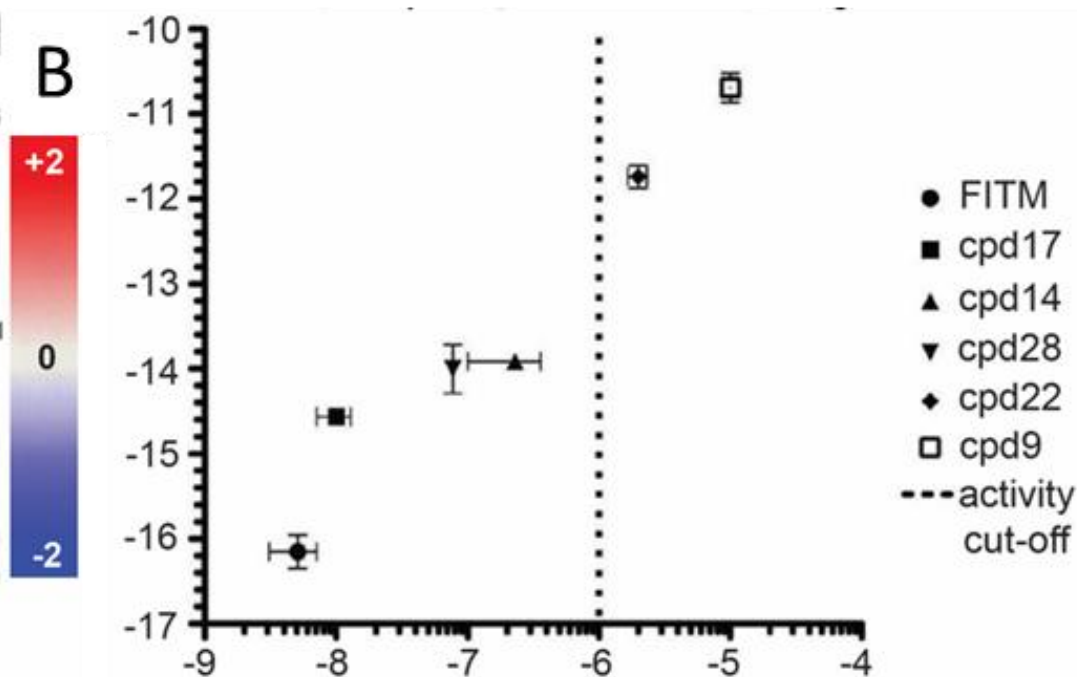
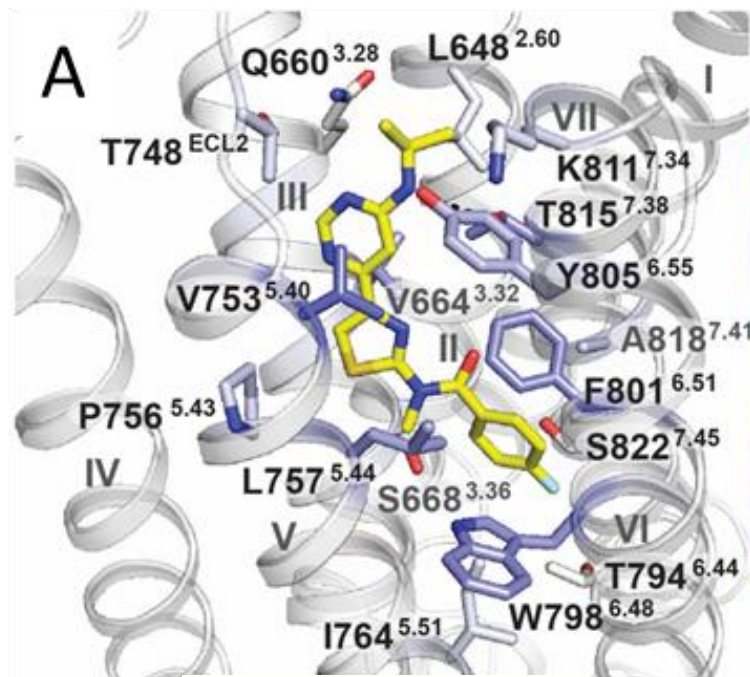
Eticlopride docked into
dopamine D3 (2.6Å RMSD)

Overall Structure of the mGlu1 TMD and Entrance to the Ligand Site



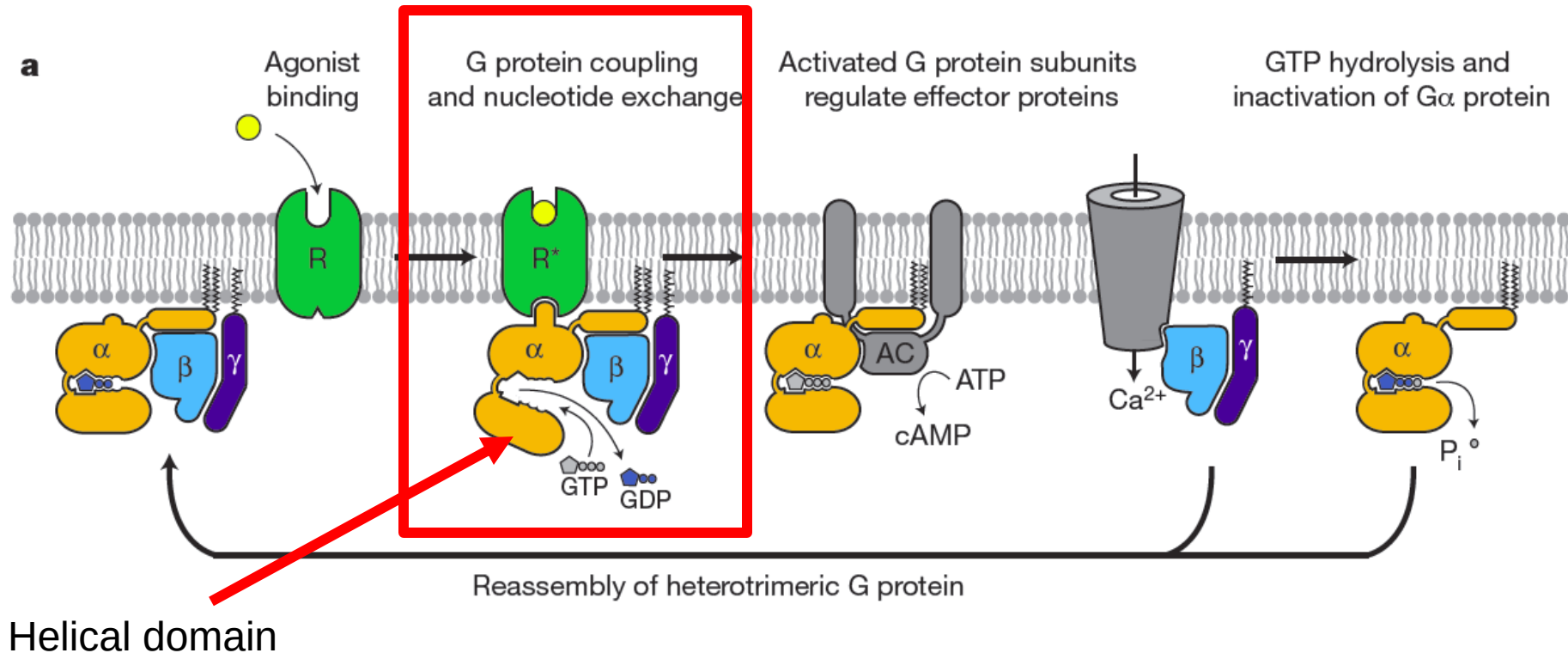
H. Wu, et al.; "Structure of a class C GPCR metabotropic glutamate receptor 1 bound to an allosteric modulator"; *Science*; accepted

Structure of a Class C GPCR in complex with an Allosteric Modulator



H. Wu, et al.; "Structure of a class C GPCR metabotropic glutamate receptor 1 bound to an allosteric modulator"; *Science*; accepted

G protein cycle: Formation of the G protein | R* complex

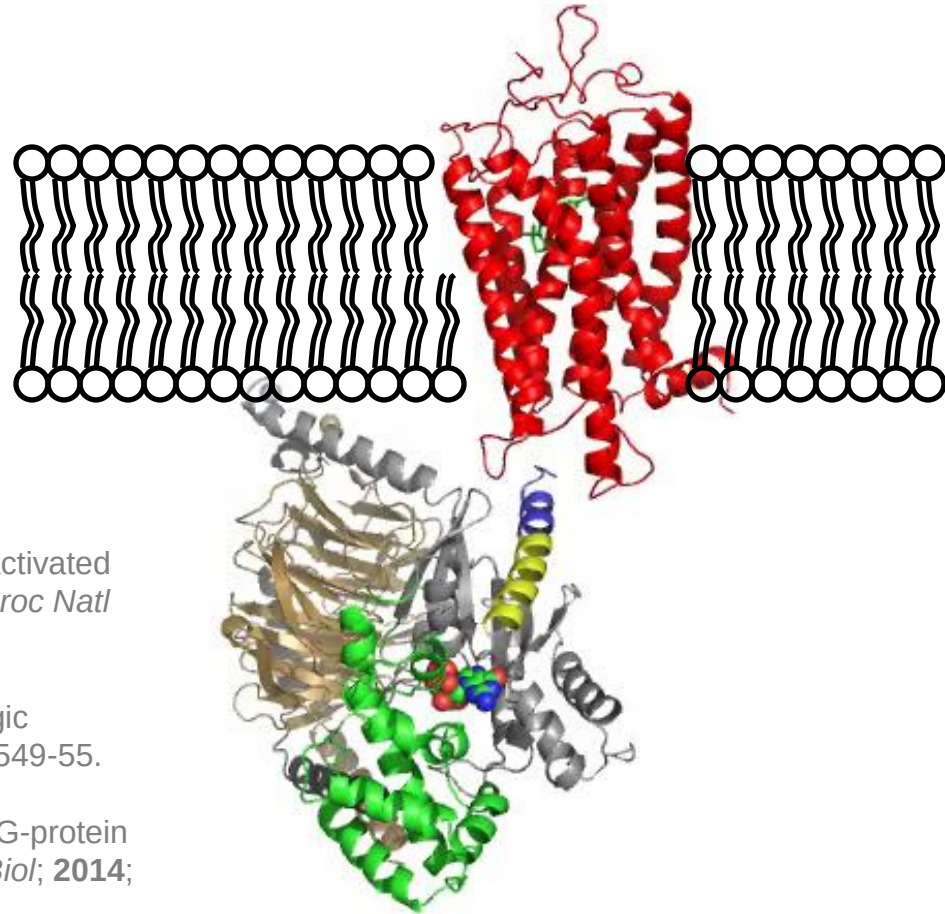


Rasmussen, S. G., et. al.; "Crystal structure of the beta(2) adrenergic receptor-Gs protein complex"; *Nature*; **2011**; Vol. p.

GPCR Signaling Involves Substantial Conformational Changes



- Light induces a configurational change of Retinal
- A conformational change of the receptor is induced
- Leading to G-Protein binding and GDP release

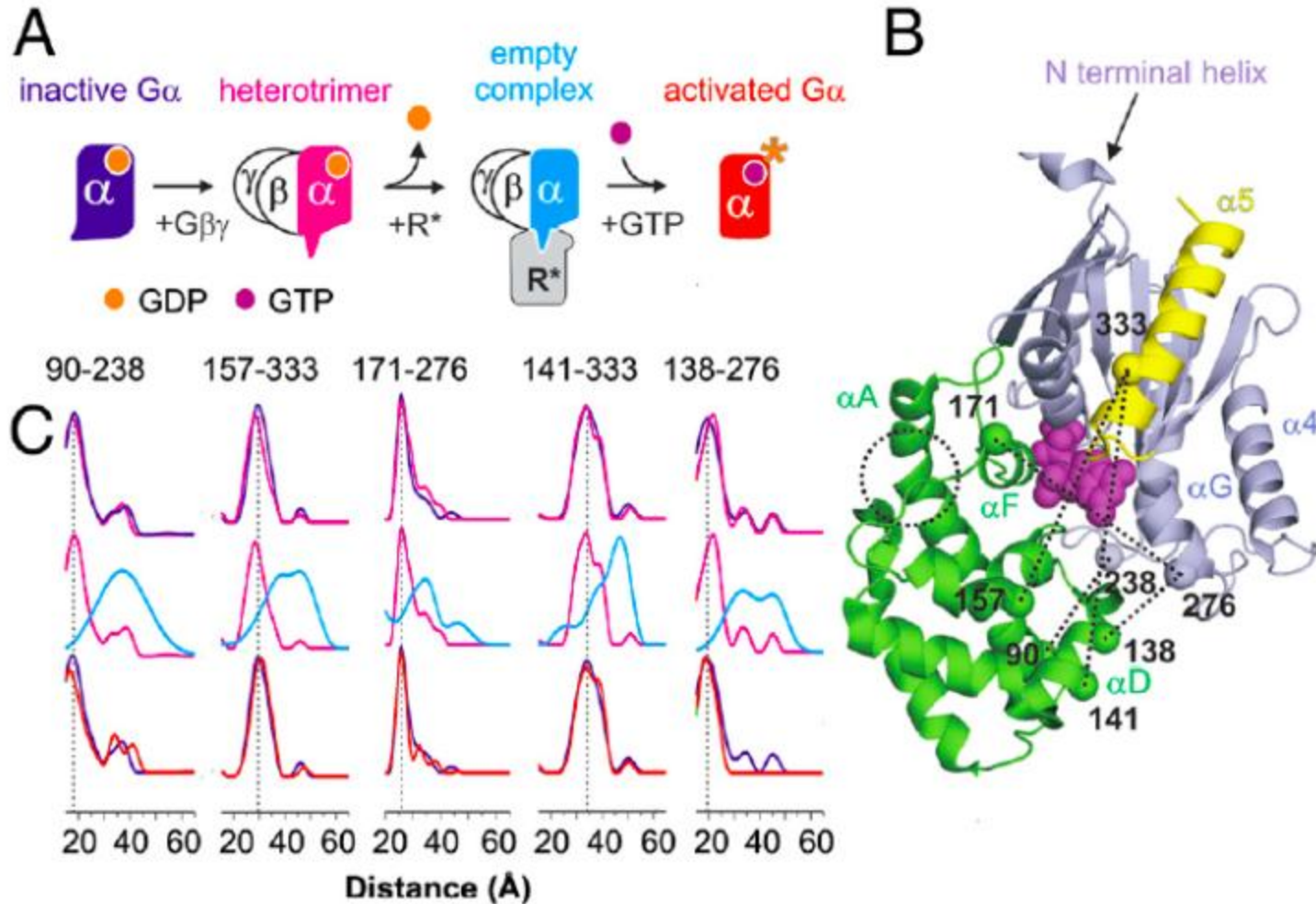


N. Van Eps, **J. Meiler**, et al.; "Interaction of a G protein with an activated receptor opens the interdomain interface in the alpha subunit"; *Proc Natl Acad Sci U S A*; **2011**; Vol. 108 (23): p. 9420-4.

S. G. Rasmussen, et al.; "Crystal structure of the beta2 adrenergic receptor-Gs protein complex"; *Nature*; **2011**; Vol. 477 (7366): p. 549-55.

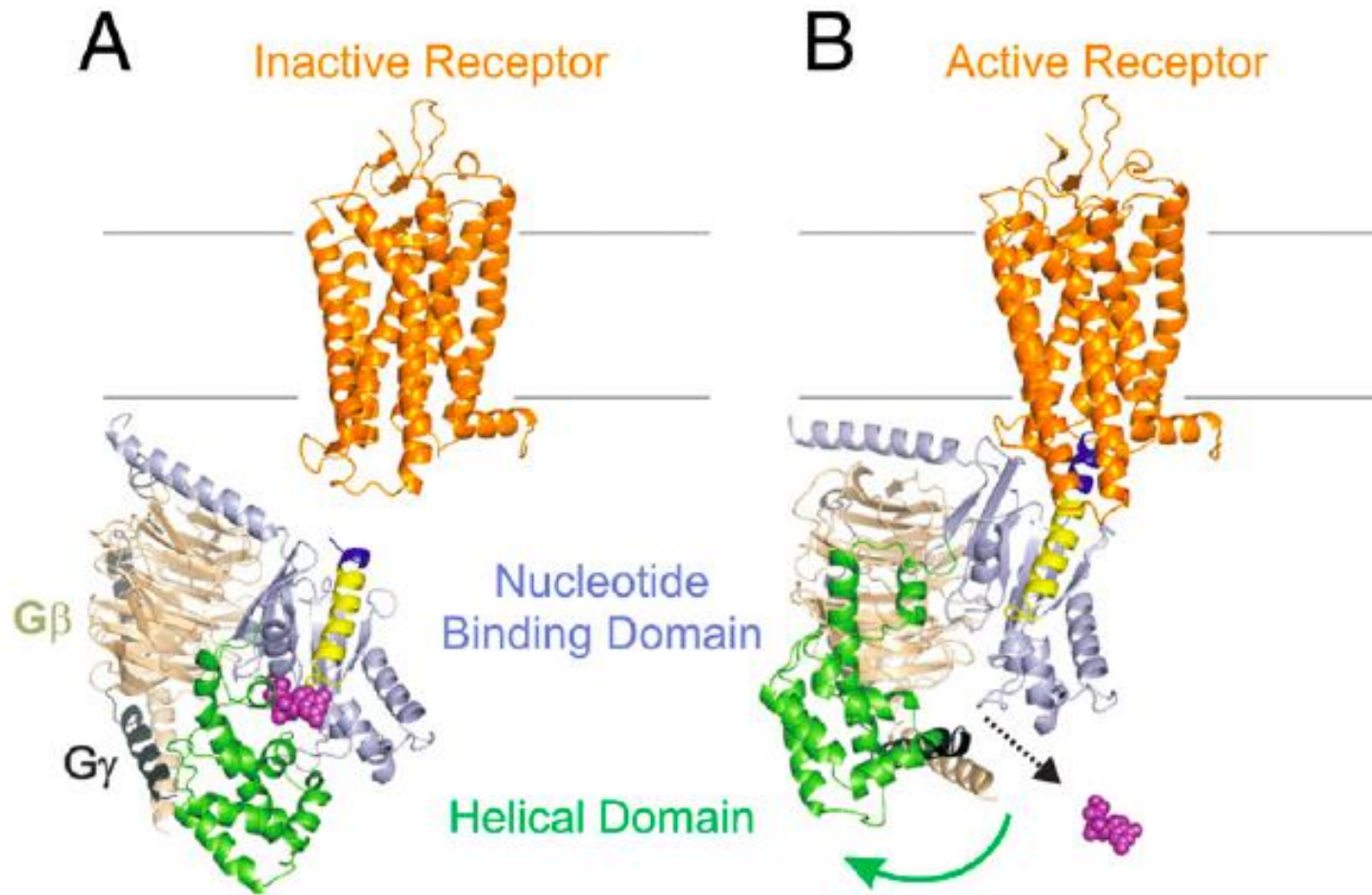
N. S. Alexander, **J. Meiler**; "Energetic analysis of the rhodopsin-G-protein complex links the alpha5 helix to GDP release"; *Nat Struct Mol Biol*; **2014**; Vol. 21 (1): p. 56-63.

Receptor activation of G proteins leads to domain separation



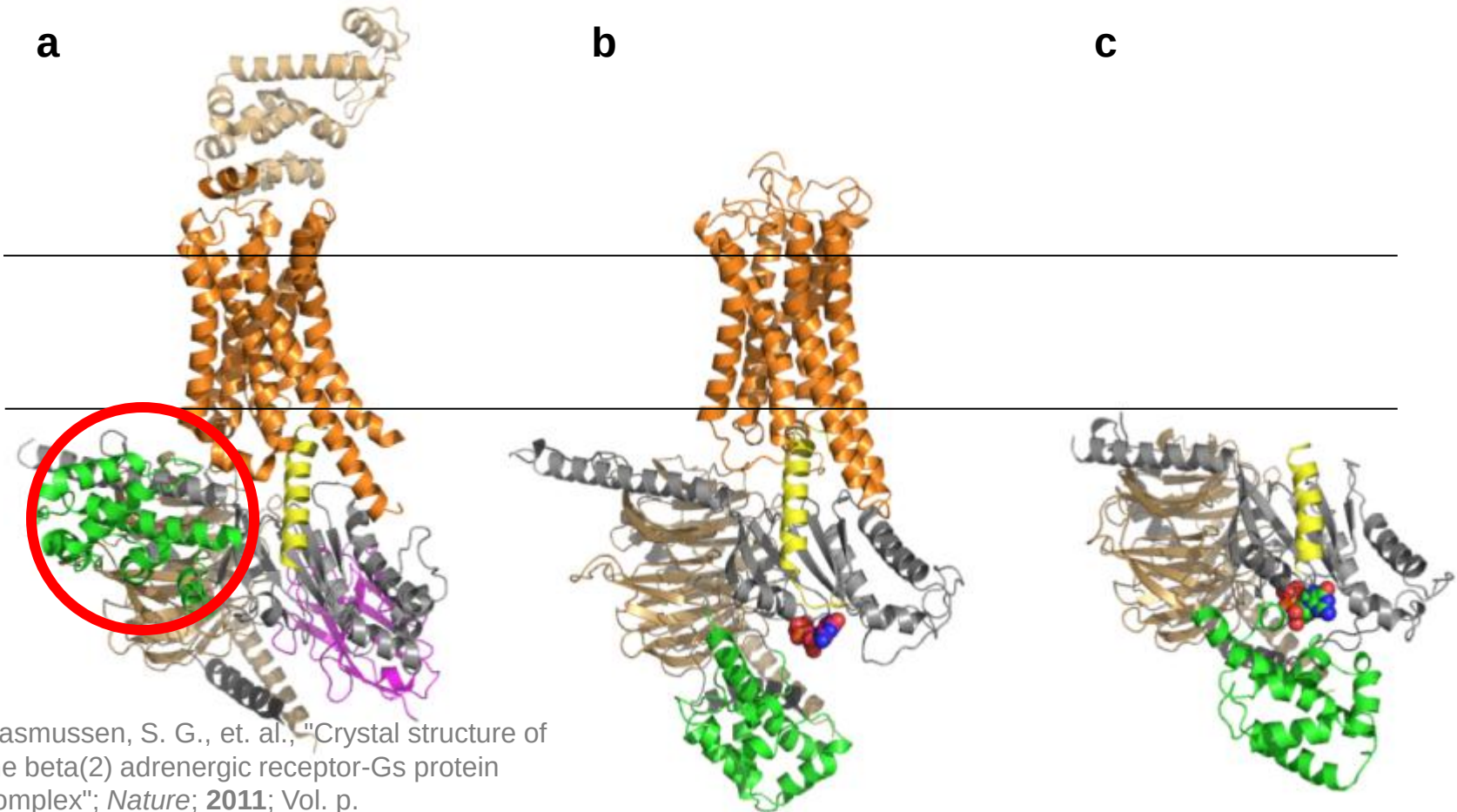
Van Eps, N., et. al.;
 "Interaction of a G
 protein with an
 activated receptor
 opens the interdomain
 interface in the alpha
 subunit"; *Proc Natl
 Acad Sci U S A*; **2011**;
 Vol. 108 (23): p.
 9420-4.

A model showing the opening of the inter-domain

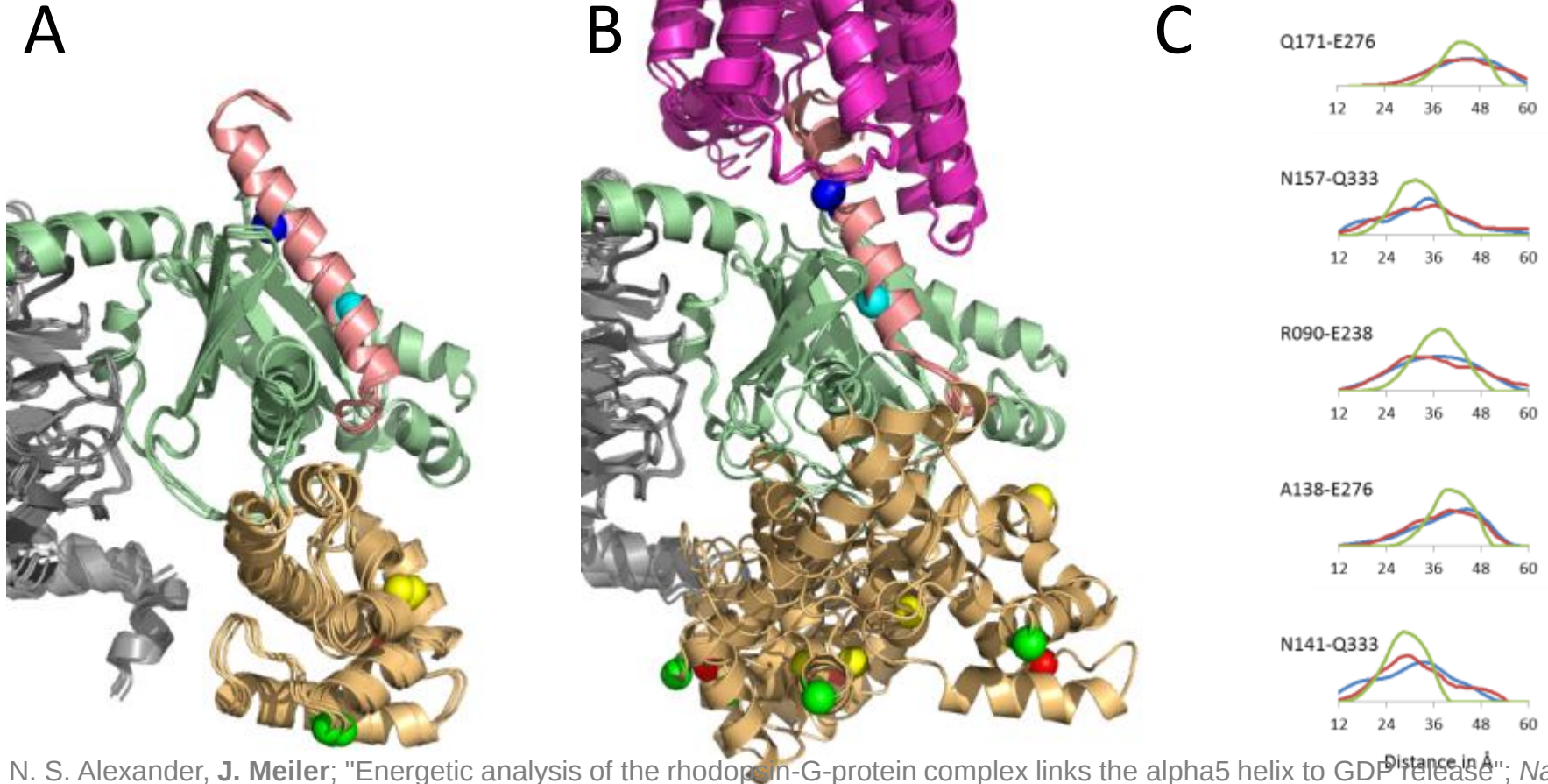


Van Eps, N., et. al.;
"Interaction of a G
protein with an
activated receptor
opens the interdomain
interface in the alpha
subunit"; *Proc Natl
Acad Sci U S A*; **2011**;
Vol. 108 (23): p.
9420-4.

Overall structure of the b2AR|Gs complex omitting crystallization aids

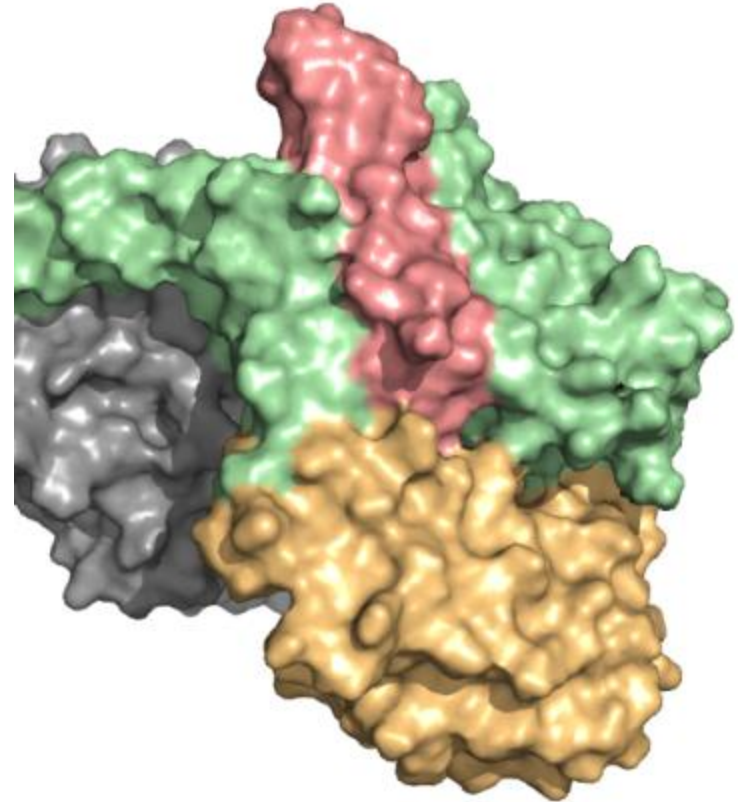
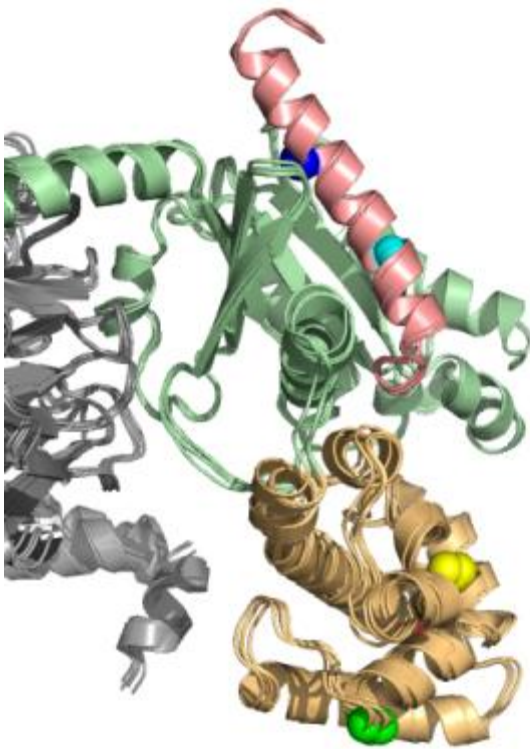


Placement of helical domain and rotation of $\alpha 5$ as observed by EPR



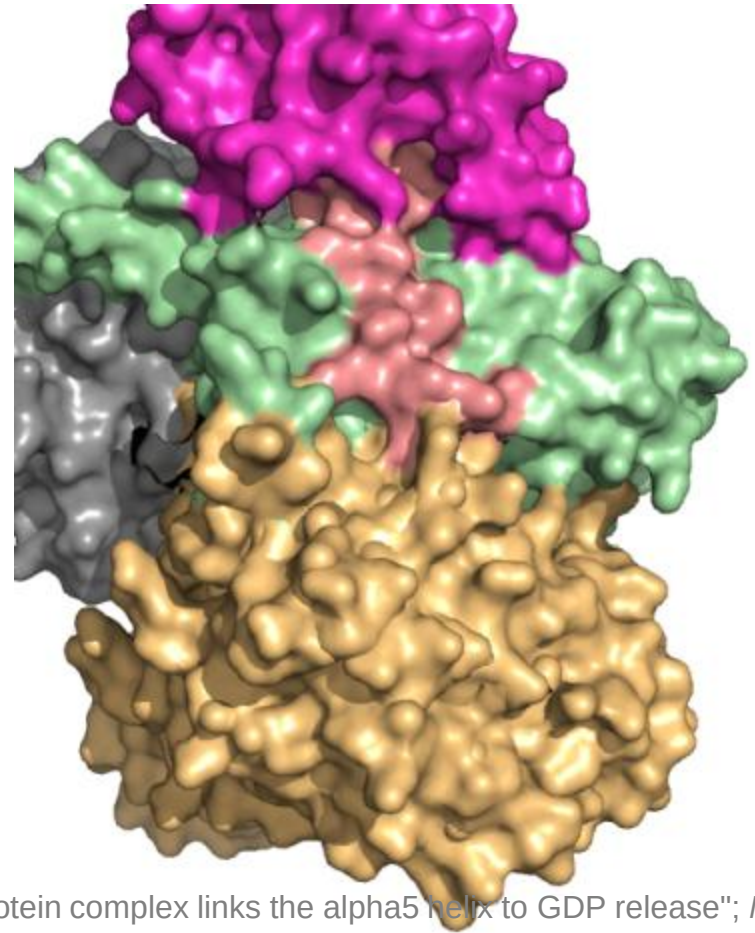
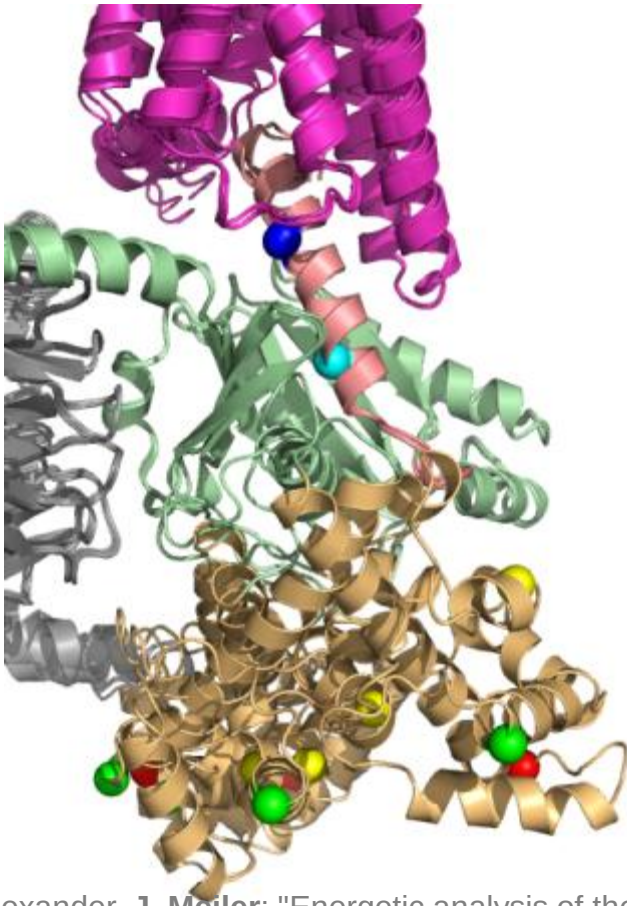
N. S. Alexander, **J. Meiler**; "Energetic analysis of the rhodopsin-G-protein complex links the $\alpha 5$ helix to GDP release"; *Nat Struct Mol Biol*; **2014**; Vol. 21 (1): p. 56-63.

Space occupied by helical domain in the basal state



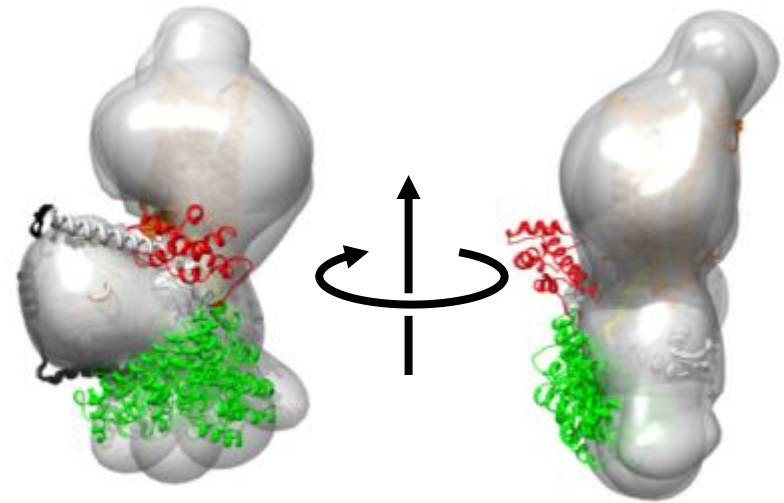
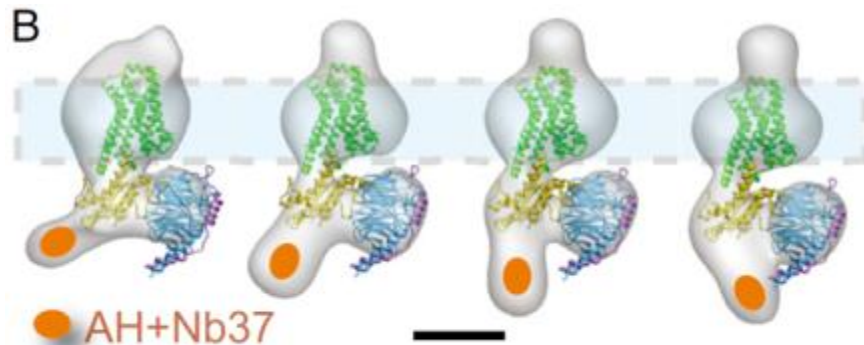
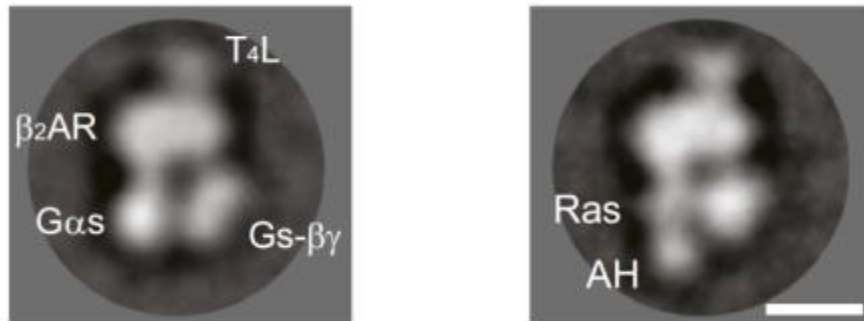
N. S. Alexander, **J. Meiler**; "Energetic analysis of the rhodopsin-G-protein complex links the alpha5 helix to GDP release"; *Nat Struct Mol Biol*; **2014**; Vol. 21 (1): p. 56-63.

Space occupied by helical domain in Gai/R* complex



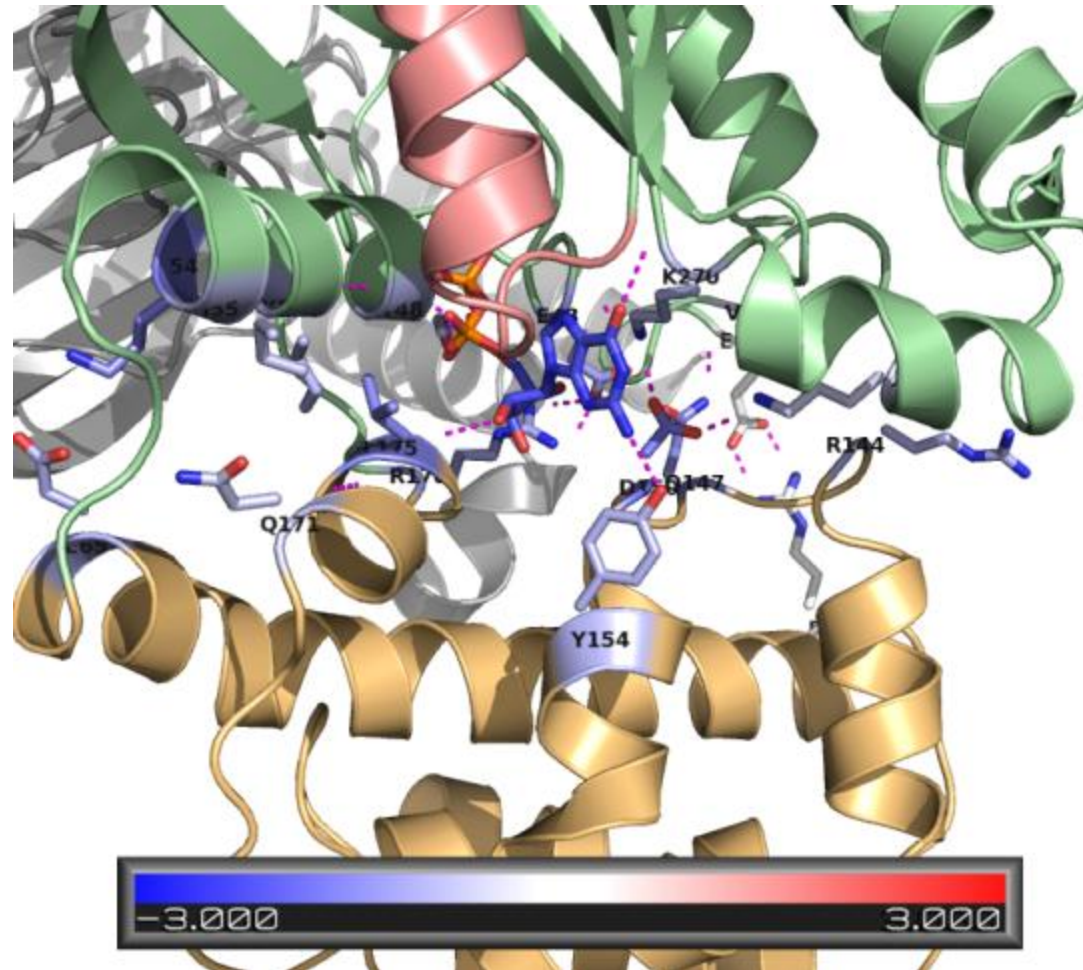
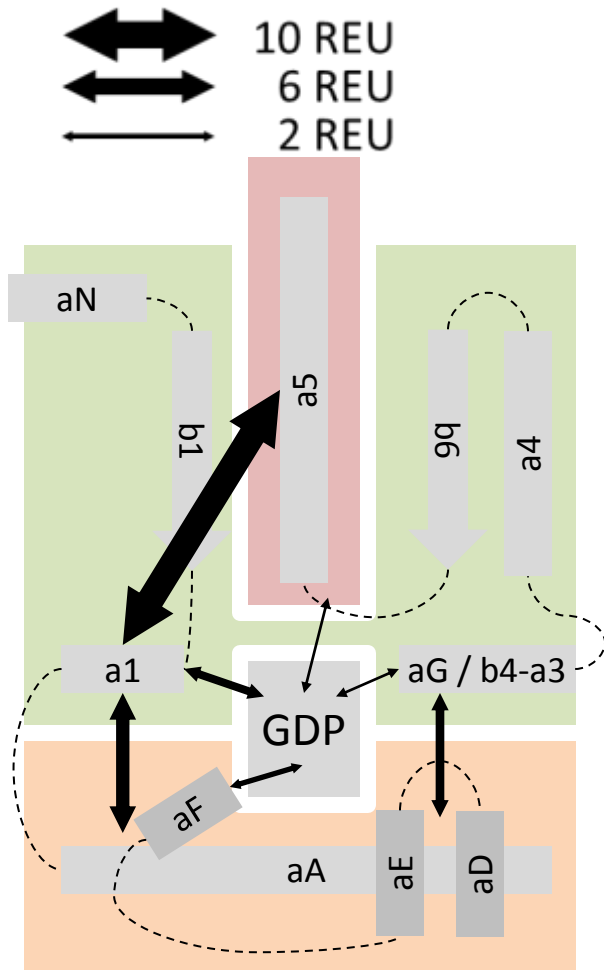
N. S. Alexander, **J. Meiler**; "Energetic analysis of the rhodopsin-G-protein complex links the alpha5 helix to GDP release"; *Nat Struct Mol Biol*; **2014**; Vol. 21 (1): p. 56-63.

Single particle EM reconstructions of the $\beta 2AR|Gs$ complex

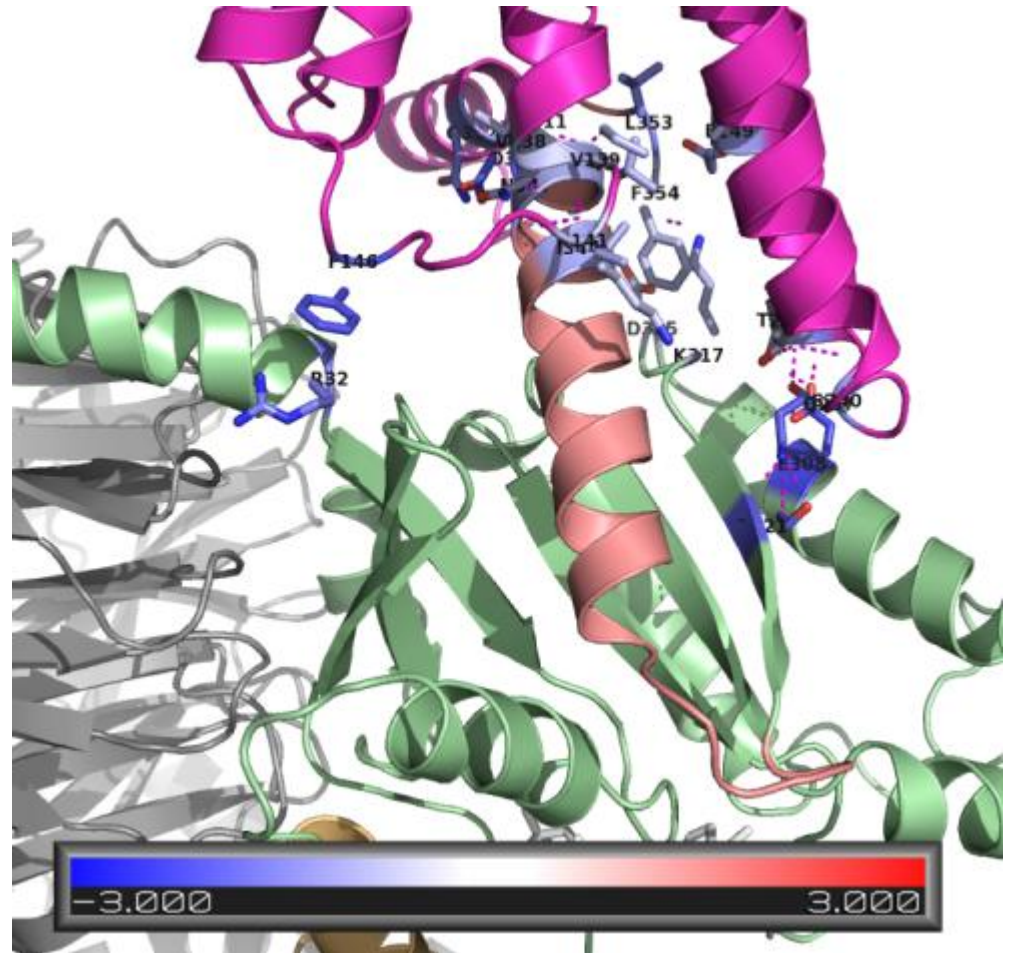
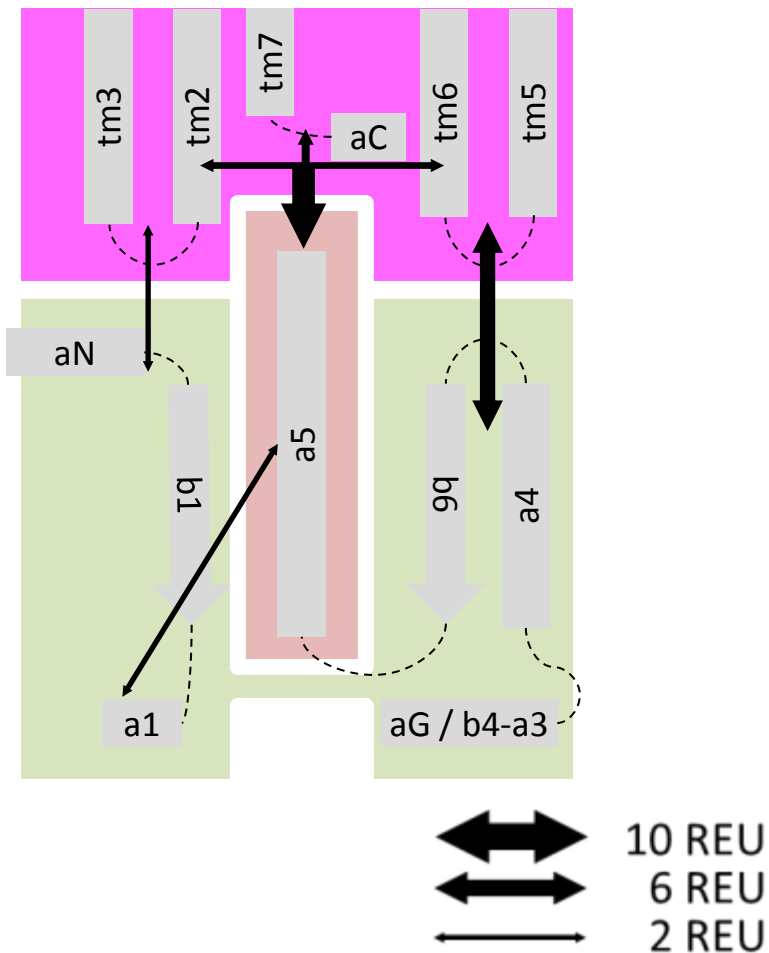


Westfield, G. H.; "Structural flexibility of the G alpha s alpha-helical domain in the beta2-adrenoceptor Gs complex"; *Proc Natl Acad Sci U S A*; **2011**; Vol. 108 (38): p. 16086-91.

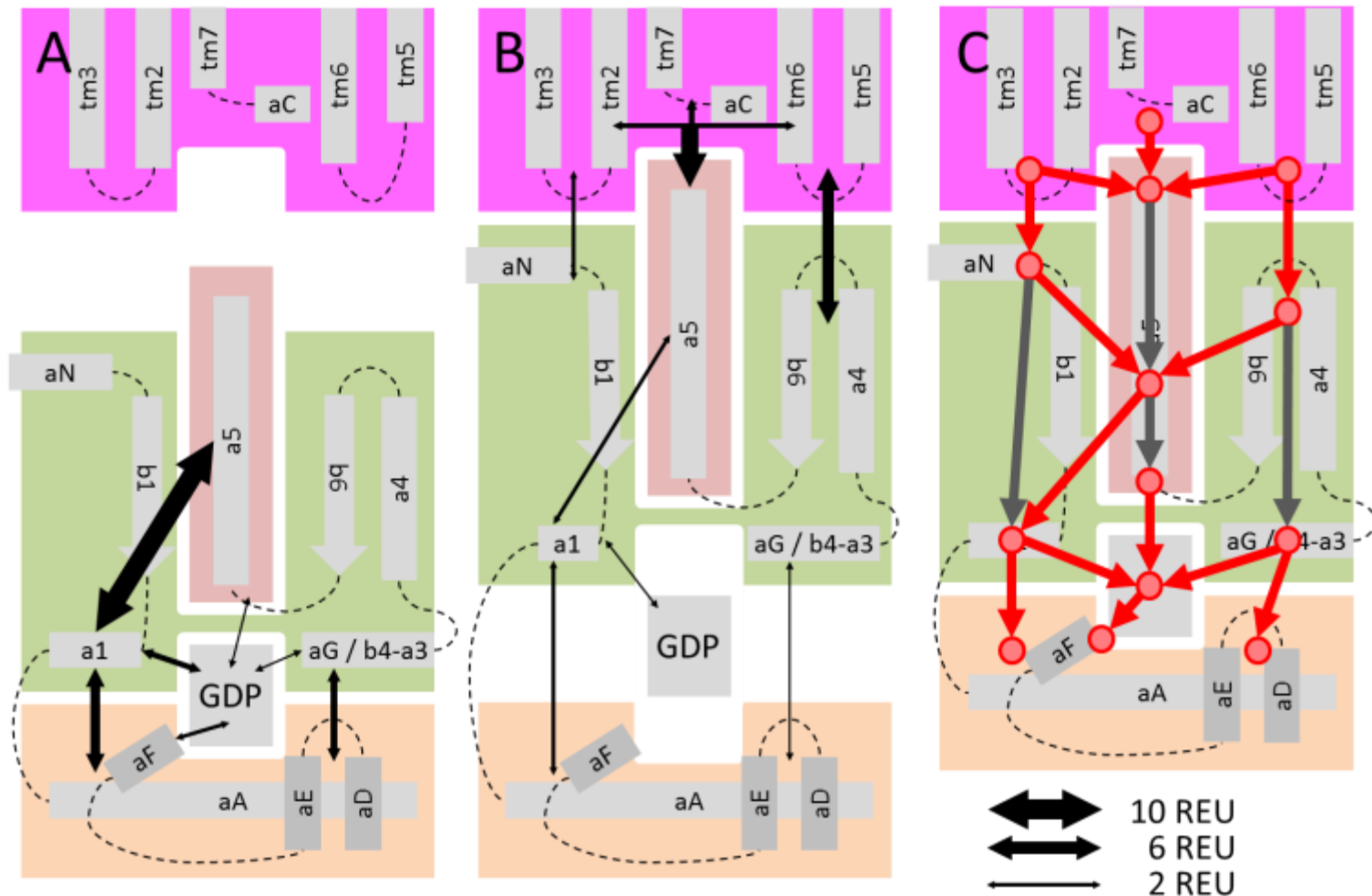
Energetics of helical domain | Gai interface in free Gai.



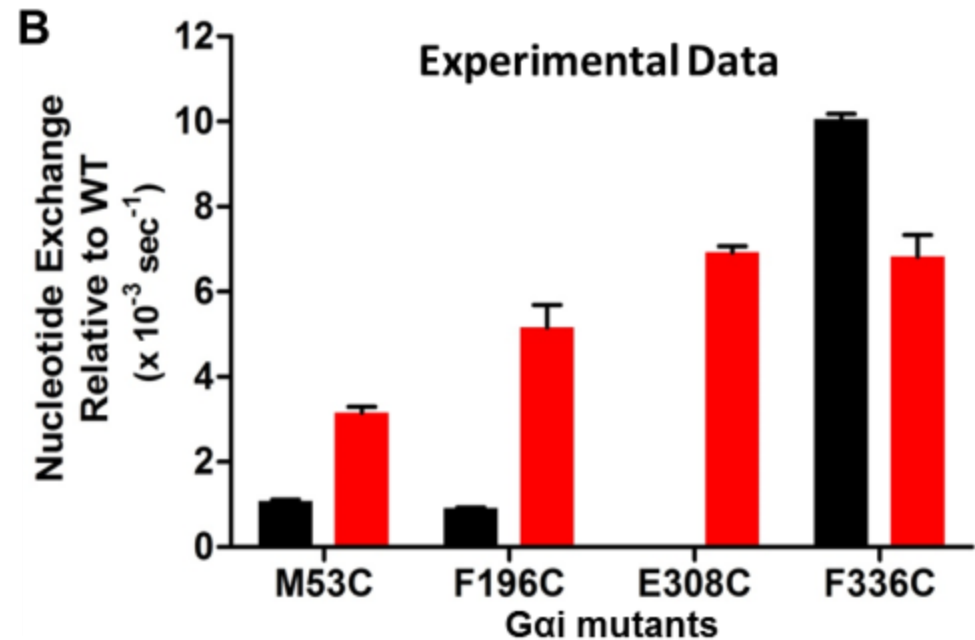
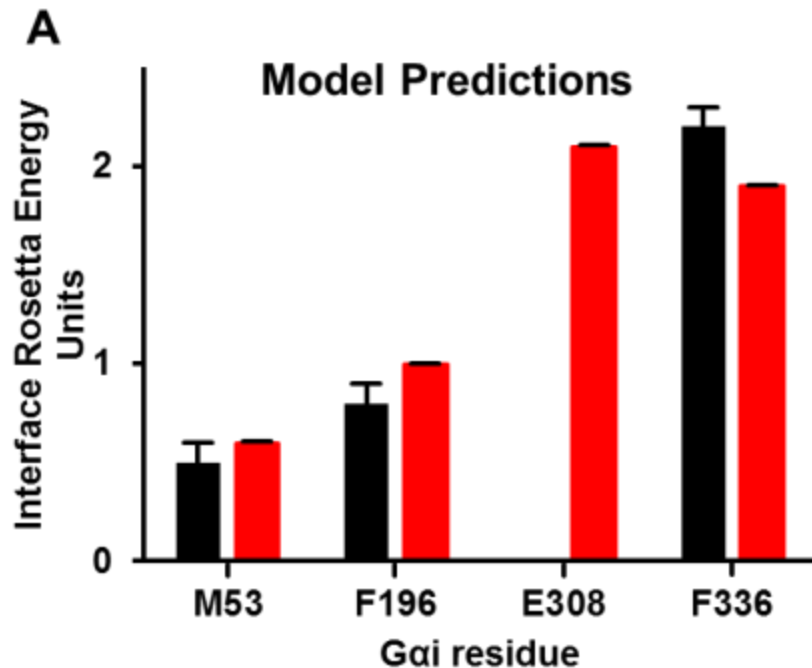
Energetics of R* | Gai interface in the R*-Gai complex



Energetic basis of signal transduction during Gai interaction with receptor R*



Basal and receptor-mediated nucleotide exchange rates match model



N. S. Alexander, **J. Meiler**; "Energetic analysis of the rhodopsin-G-protein complex links the alpha5 helix to GDP release"; *Nat Struct Mol Biol*; **2014**; Vol. 21 (1): p. 56-63.