# Centrality of Weak Interhelical H-bonds in Membrane Protein Functional Assembly and Conformational Gating

**Ilan Samish[1], Eran Goldberg[1], Oksana Kerner[1] & Avigdor Scherz[1,2]**

[1]**Department of Plant Sciences, Weizmann Institute of Science, 76100 Rehovot, Israel.**
[2]To whom correspondence should be addressed.

Received line

## ABSTRACT

Our analysis demonstrates that backbone-mediated interhelical hydrogen-bonds cluster laterally in the conserved core of transmembrane helical proteins. Each residue's propensity to bear these interactions is in correlation with the residue's packing-value scale; giving biophysical meaning to this phenomenological scale. Residues participating in such an intersubunit, structurally conserved H-bond in reaction centers of photosystem II were combinatorially mutated and characterized *in silico* and *in vivo* suggesting that H-bond reversible association regulates protein-gated electron transfer. Similar motifs may be involved in folding and conformational flexibility of other membrane proteins. Hence, these findings provide new parameters for structure and function prediction.
**Contact**: ilan.samish@weizmann.ac.il

## INTRODUCTION

Membrane protein conformational gating plays key roles in protein systems (Spencer and Rees 2002). Although functionally distinct in many catalytic cycles, the underlying mechanisms remain enigmatic (Warshel and Parson 2001). In order to flip between two alternative conformations, weak interactions must rearrange. Such interactions may include weak (as defined by Desiraju (Desiraju and Steiner 1999)) and strong interhelical H-bonds as well as packing interactions (Richards 1974). Examples to weak bonds we studied include: (a) A bifurcated H-bond in which a backbone carbonyl accepts an inter- and intra- helical H-bond. (b) An interhelical bond donated by the acidic $C_\alpha$ atom (Senes, Ubarretxena-Belandia et al. 2001). Amplified by the low dielectric milieu of the membrane, weak interactions in membrane proteins have a stronger effect in comparison to their soluble counterparts. More generally, similar interactions are involved in membrane protein functional assembly (DeGrado, Gratkowski et al. 2003) playing a part in the helix-assembly step within the 3-stage model for membrane protein folding (Engelman, Chen et al. 2003).

Following statistical characterization of interhelical H-bonds in a non-redundant set of structurally known helical proteins, we concentrated on photosynthetic reaction centers (PRCs) as a model system – the 1st and most varied group of structurally-known helical membrane proteins. Due to 3.5 billion years of evolution, these protein complexes have undergone considerable sequence degeneracy, but have structurally maintained important motifs required for function. Consequently, functionally important and structurally conserved motifs can be datamined followed by discrete analysis of the common motif in each available structure. Further, the study of membrane proteins in their natural membrane environment imposes technical difficulties due to the lack of 'reporting agents' that can teach us about molecular level changes. Hence, the multi-step electron transfer (ET) in PRCs can serve as a molecular reporter to changes in different microenvironments of the protein. Last, photosystem II PRC in cyanobacteria is a well-suited system for genetic manipulation and *in vivo* biophysical characterization. Thus, combinatorial mutagenesis, combined with temperature-dependent biophysical analysis and *in silico* mutant characterization enables one to utilize PRCs as a model system to study the role/fitness of different residues/chemical moieties in the functional assembly and conformational gating of membrane proteins.

Our study integrates bioinformatic tools in a feedback loop – guiding the genetic manipulation strategy and explaining the resulting biophysical findings. Generalization of key findings in the model system is presented by searching for the phenomenon in sequence and structure space and vice versa. Finally, the recent accumulation of helical membrane protein structures enabled us to conduct statistical datamining of the phenomenon suggested by our model system in a non-redundant dataset of all helical membrane proteins. Cumulatively, our results suggest new roles for weak interhelical H-bonds in the membrane milieu as well as biophysical meaning to interactions previously regarded phenomenologically as 'packing'.

## RESULTS & DISCUSSION

A structural core of all PRCs was computed based on the multiple structural alignment and core alignment algorithm. 200 out of 700 locations of $C_\alpha$-atoms were structurally conserved in all PRCs. A cluster of high-packing amino acids in the central core of all complexes was found. This cluster was analyzed in theoretical and crystal structures of type II (i.e. bacterial and photosystem II) PRCs. Conserved

intersubunit H-bonds were detected in all structures suggesting a role in protein functionality.

*In vivo* combinatorial mutagenesis of residues found in this conserved region was performed in photosystem II to resolve the significance of weak interhelical H-bond, as identified by computer modeling, in controlling different protein conformations required for protein-gated ET. Rotamer-based *in silico* mutagenesis conducted on this conserved region confirmed the need for such a bond in order to carry the conformational gating. Next, the ability of each mutant to conduct gating and the local stability were estimated using two complimentary temperature-dependent methods: single-turnover double-modulated fluorescence and thermoluminescence. Thermodynamic parameters related with multiple conformations and folding were resolved and used to establish a physical model for gating.

In order to understand the general role of interhelical H-bonds and packing motifs, statistical characterization of these bonds was conducted using a non-redundant dataset of helical membrane proteins. Findings include: (a) backbone-mediated interhelical H-bonds are laterally distributed in transmembrane domains; (b) residues bearing these bond are evolutionary conserved; and (c) the propensity of residues to participates in such bonds is in correlation with the amino-acids packing value scale (Eilers, Shekar et al. 2000). Cumulatively, these findings provided biophysical meaning to this previously published phenomenological characterization.

Hence, our study underscores the significance of a balance between different interhelical forces in: (a) keeping the protein exterior hydrophobic during helix insertion into the membrane; (b) supporting electrostatic interactions in the protein interior; and (c) maintain conformational flexibility required for protein function. Weak interhelical interactions have evolutionary evolved in support of these constraints maintaining an intact, yet functionally flexible protein. This biophysical understanding provides new parameters that can be utilized for structure and function prediction.

## MATERIALS & METHODS

Data Set: Photosynthetic reaction center structures analyzed include (**PDB** (Berman, Westbrook et al. 2000) code): 1prc, 1m3x, 1aij, 1aig, 1eys, 1jb0, 1izl, 1dop (theoretical model), 2 theoretical models by Xiong *et al.* (Xiong, Subramaniam et al. 1996; Xiong, Subramaniam et al. 1998) and a photosystem II reaction center structure in publication at 3.2 Å resolution (Loll B., Biesiadka J. & Saenger W., personal communication). H-bonds were statistically characterized in the following structures (resolution 2.8 Å or better): 1c3w, 1eul, 1ezv, 1fx8, 1j95, 1l9h, 1ots, 1qla, 1occ 1jb0, 1prc.

In order to define the transmembrane segment for each structure the **Swiss-Prot** (Boeckmann, Bairoch et al. 2003) annotation was used, except for 1ezv, where 1bgy entry of **MPtopo** (Jayasinghe, Hristova et al. 2001) was used, with slight modifications based on structural alignment.

Datamining: Virtual mutagenesis by rotamer library was conducted using **SCWRL** (Bower, Cohen et al. 1997). PDB files were manually edited and described as a single subunit to adhere with the software requirements. Resulting structures were subject to energy minimization prior to H-bond detection.

In order to detect H-bonds, hydrogen atoms were added to each structure using a slightly-modified version of the program **REDUCE** (Word, Lovell et al. 1999), adding hydrogens also to the carboxyl group of Asp and Glu – groups that tend to be in their uncharged, i.e. protonated, form in the low dielectric environment of the membrane (Popot and Engelman 2000). REDUCE was executed with the '-BUILD' flag to optimize hydrogen atoms positions and Asn, Gln and His orientations. An in-house C++ program, **HBD** (Goldberg E., *et al.* unpublished), has been implemented to identify H-bonds, using the **GAMB++** library (HJ Wolfson, Tel-Aviv University). The geometrical criteria for H-bonds were taken from Senes *et al*.(Senes, Ubarretxena-Belandia et al. 2001).

Sequence conservation scores were calculated using **Rate4Site** (Pupko, Bell et al. 2002), implementing the maximum likelihood principle. Sequence conservation of the high-packing motif in the D1 subunit of photosystem II was conducted by searching sequence space with **PSI-BLAST** (Altschul, Madden et al. 1997), using both the original D1 sequence and the representative sequence (cobbler) from the **BLOCKS** database (Henikoff, Greene et al. 2000) – block IPB000484C. In addition, a Blast search (using **TBLASTN**) was carried out against the full genome sequences of plants and cyanobacteria that have not yet been parsed into gene databases.

Structure conservation within PRCs was calculates using multiple structural alignment and core detection (**MUSTA)** (Leibowitz, Nussinov et al. 2001). In order to avoid the bias demonstrated when using a single template, each group of size $n$ was computed $n$ times, each time with a different protein serving as a template. A search for intersubunit H-bonds in the conserved residues found was conducted using **Swiss-PDB-Viewer** (Guex and Peitsch 1997).

Internal cavities were detected using **Volbl** (Liang, Edelsbrunner et al. 1998; Liang, Edelsbrunner et al. 1998) with a probe radius of 1Å. Solvent accessible surface (SAS) was calculated for every residue using **MOLMOL** (Koradi, Billeter et al. 1996). As buried residues are known to be more conserved than lipid-facing residues, the conservation of residues participating in backbone-mediated interhelical H-bonds was conducted relative to the conservation of other buried residues (SAS < 7%).

## REFERENCES


Altschul, S. F., Madden, T. L. et al. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nuc. Acid Res. **25**(17): 3389-402.

Berman, H. M., Westbrook, J. et al. (2000). The Protein Data Bank. Nuc. Acids Res. **28**(1): 235-242.

Boeckmann, B., Bairoch, A. et al. (2003). The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003. Nuc. Acid Res. **31**(1): 365-370.

Bower, M. J., Cohen, F. E. et al. (1997). Prediction of protein side-chain rotamers from a backbone-dependent rotamer library: a new homology modeling tool. J. Mol. Biol. **267**(5): 1268-82.

DeGrado, W. F., Gratkowski, H. et al. (2003). How do helix-helix interactions help determine the folds of membrane proteins? Perspectives from the study of homo-oligomeric helical bundles. Prot. Sci. **12**: 647-665.

Desiraju, G. R. and Steiner T. (1999). The weak hydrogen bond in structural chemistry and biology, Oxford university press.

Eilers, M., Shekar, S. et al. (2000). Internal packing of helical membrane proteins. Proc. Natl. Acad. Sci. U S A **97**(11): 5796-5801.

Engelman, D. M., Chen, Y. et al. (2003). Membrane protein folding: beyond the two stage model. FEBS Lett. **555**(1): 122-5.

Guex, N. and Peitsch M. (1997). Swiss-model and Swiss-PDB Viewer: An environment for comparative protein modeling. Electrophoresis **18**: 2714-2723.

Henikoff, J. G., Greene, E. A. et al. (2000). Increased coverage of protein families with the blocks database servers. Nucleic Acids Res. **28**: 228-30.

Jayasinghe, S., Hristova, K. et al. (2001). MPtopo: A database of membrane protein topology. Prot. Sci. **10**(2): 455-458.

Kleinfeld, D., Okamura, M. Y. et al. (1984). Electron-transfer kinetics in photosynthetic reaction centers cooled to cryogenic temperatures in the charge-separated state: evidence for light-induced structural changes. Biochemistry **23**: 5780-6.

Koradi, R., Billeter, M. et al. (1996). MOLMOL: a program for display and analysis of macromolecular structures. J. Mol. Graph. **14**(1): 51-5, 29-32.

Leibowitz, N., Nussinov, R. et al. (2001). MUSTA--a general, efficient, automated method for multiple structure alignment and detection of common motifs: application to proteins. J. Comput. Biol. **8**(2): 93-121.

Liang, J., Edelsbrunner, H. et al. (1998). Analytical shape computation of macromolecules: II. Inaccessible cavities in proteins. Proteins **33**(1): 18-29.

Liang, J., Edelsbrunner, H. et al. (1998). Anatomy of protein pockets and cavities: measurement of binding site geometry and implications for ligand design. Prot. Sci. **7**(9): 1884-97.

Popot, J.-L. and Engelman D. M. (2000). Helical membrane protein folding, stability, and evolution. Annual Reviews in Biochemistry **69**: 881-922.

Pupko, T., Bell, R. E. et al. (2002). Rate4Site: an algorithmic tool for the identification of functional regions in proteins by surface mapping of evolutionary determinants within their homologues. Bioinformatics **18 Suppl 1**: S71-7.

Richards, F. M. (1974). The interpretation of protein structures: total volume, group volume distributions and packing density. J. Mol. Biol. **82**(1): 1-14.

Senes, A., Ubarretxena-Belandia, I. et al. (2001). The Calpha ---H...O hydrogen bond: a determinant of stability and specificity in transmembrane helix interactions. Proc. Natl. Acad. Sci. U S A **98**(16): 9056-61.

Spencer, R. H. and Rees D. C. (2002). The alpha-helix and the organization and gating of channels. Annu. Rev. Biophys. Biomol. Struct. **31**: 207-33.

Stowell, M., McPhillips, T. et al. (1997). Light-induced structural changes in photosynthetic reaction center: implications for mechanism of electron-proton transfer. Science **276**(5313): 812-816.

Warshel, A. and Parson W. W. (2001). Dynamics of biochemical and biophysical reactions: insight from computer simulations. Q. Rev. Biophys. **34**(4): 563-679.

Word, J. M., Lovell, S. C. et al. (1999). Asparagine and glutamine: Using hydrogen atom contacts in the choice of side-chain amide orientation. J. Mol. Biol. **285**(4): 1735-1747.

Xiong, J., Subramaniam, S. et al. (1996). Modeling of the D1/D2 Proteins and Cofactors of the Photosystem-II Reaction-Center - Implications For Herbicide and Bicarbonate Binding. Prot. Sci. **5**(10): 2054-2073.

Xiong, J., Subramaniam, S. et al. (1998). A knowledge-based three dimensional model of the Photosystem II reaction center of *Chlamidomonas reinhardtii.* Photosyn. Res. **56**: 229-254.