UNIVERSITY OF CALIFORNIA

Los Angeles

# Hydrogen Bond Shaping of Membrane Protein Structure

A dissertation submitted in satisfaction of the requirements

for the degree Doctoral of Philosophy in Chemistry

by

Zheng Cao

2013

ABSTRACT OF THE DISSERTATION

Hydrogen Bond Shaping of Membrane Protein Structure

by

Zheng Cao

Doctoral of Philosophy in Chemistry

University of California, Los Angeles, 2013

Professor James U. Bowie, Chair

The intricate functions of membrane proteins would not be possible without bends or breaks that are remarkably common in transmembrane helices. The frequent distortions are nevertheless surprising because backbone hydrogen bonds should be strong in an apolar environment, potentially rigidifying helices. It is therefore mysterious how distortions can be generated by evolutions. Through my studies on bacteriorhodopsin and $Ca^{2+}$-ATPase, I found that helix distortions are facilitated by shifting hydrogen bonding partners. My results explained how evolution has been able to liberally exploit transmembrane helix bending by for optimizing membrane protein structure, function and dynamics.

It has recently been found that the stability of bacteriorhodopsin assessed through unfolding did not achieve equilibrium. Here I made a new equilibrium test by measuring

the transition of folded bacteriorhodopsin to unfolded bacterioopsin. My result suggests that the energetic effects of most mutations that have been studied before may be focused on the folded state.

To further investigate how hydrogen bonds may shape transmembrane helices, I sought a non-perturbing method to measure strengths of both backbone and side-chain hydrogen bond strengths, because backbone hydrogen bonds cannot be probed by mutation and mutagenesis to remove side-chains may not mimic the breaking of side-chain hydrogen bonds. I therefore decided to employ the equilibrium hydrogen / deuterium fractionation factors ($\phi$) for hydrogen bonds. I used model compounds to study the relationship between $\phi$-value and the hydrogen bond free energy. By applying this relationship, I found that hydrogen bonding stabilizes enzyme intermediate state by up to ~ 4 kcal/mol more than the resting state and marginally favor soluble-protein unfolding.

To measure hydrogen bond strengths in a membrane protein, I focused on the isolated voltage-sensor domain of the voltage-dependent potassium-selective channel. By utilizing its 2D NMR spectrum, I obtained $\phi$-values and thus free energies for ~ 70 % of the backbone hydrogen bonds in this membrane protein. I found that it has similar backbone hydrogen bond strength to water-soluble proteins on average. Moreover, I found that the flexible (dynamic) point of a transmembrane helix in this protein can be predicted from the backbone hydrogen bond strengths in that helix.

The dissertation of Zheng Cao is approved.

David S. Eisenberg

Feng Guo

James U. Bowie, Committee Chair

University of California, Los Angeles

2013

TABLE OF CONTENTS

PAGE

LIST OF TABLES

LIST OF FIGURES

# ACKNOWLEDGMENTS

The progress in research that I have made in the past six years are owned to the great efforts of my PI, Dr. James U. Bowie. My achievements would not be realized without his patience, support and instructions. I also wish to thank Dr. David Eisenberg, Dr. William Gelbart and Dr. Feng Guo for serving on my committee. Besides, I will never forget the great help given by Dr. Duilio Cascio, Dr. Michael Sawaya, Dr. Jane Strouse, Dr. Robert Taylor and Dr. Robert Peterson. Without their help with technical issues in X-ray crystallography and NMR, I would have got involved in a lot troubles in my research. In addition, I would like to thank Dr. Chiwook Park and Dr. Jonathan Schlebach from Purdue University for collaborating in solving the problem in equilibrium unfolding of bacteriorhodopsin.

Last, I personally would like to acknowledge the present and past members of Bowie lab for help, support and guidance.

VITA

| | |
|---|---|
| 2007 | B. S.<br>Beijing (Peking) University, Beijing, China |
| 2007-present | Graduate Student<br>University of California, Los Angeles |
| 2008 | Teaching Assistant<br>University of California, Los Angeles |
| 2008-present | Research Assistant<br>University of California, Los Angeles |

**Awards**

| | |
|---|---|
| 2012 | Dissertation Year Fellowship<br>University of California, Los Angeles |
| 2006 | Beijing University President's Research Fellowship for Undergraduate Students<br>Beijing (Peking) University, Beijing, China |
| 2006 | Outstanding Student Scholarship<br>Beijing (Peking) University, Beijing, China |
| 2006 | First Prize at the 14th Beijing Challenge-Cup Competition for University Students in Scientific and Technological Research<br>Beijing, China |
| 2004 | First Prize at the 21st Beijing Physics Contest for University Students<br>Beijing, China |
| 2003 | Silver Medal at the 16th China National Chemistry Olympiad for High-School Students<br>Wuhan, Hubei Province, China |

**Publications**

1. Cao Z, Bowie JU (2013) Hydrogen bonds in a membrane protein are not stronger than those in water-soluble proteins. *Manuscript in preparation*.

2. Cao Z, Bowie JU (2013) An energetic scale for equilibrium H/D fractionation factors illuminates hydrogen bond free energies in proteins. *Manuscript in preparation*.

3. Cao Z, Bowie JU (2012) Shifting hydrogen bonds may produce flexible transmembrane helices. *Proc Natl Acad Sci* 109:8121–8126.

4. Cao Z, Schlebach J, Park C, Bowie JU (2012) Thermodynamic stability of bacteriorhodopsin mutants measured relative to the bacterioopsin unfolded state. *Biochimica Et Biophysica Acta* 1818:1049-1054.

5. Schlebach JP, Cao Z, Bowie JU, Park C. (2012) Revisiting the folding kinetics of bacteriorhodopsin. *Protein Sci* 21:97-106.

6. Hong H, Blois TM, Cao Z, Bowie JU. (2010) Method to measure strong protein-protein interactions in lipid bilayers using a steric trap. *Proc Natl Acad Sci USA* 107:19802-19807.

7. Yang Z, Cao Z, Sun H, Li Y**.** (2008) Composite films based on aligned carbon nanotube arrays and a poly(*N*-isopropyl acrylamide) hydrogel. *Adv Mater* 20:2201-2205.

**Presentations**

1. Cao Z, Bowie JU. (Aug 2 – 6, 2010) Interactions stabilizing proline-induced transmembrane helices. **(Poster)** *FASEB Summer Research Conference "Molecular Biophysics of Cellular Membranes"*. Saxtons River, Vermont, USA.

# CHAPTER 1

# INTRODUCTION

## 1.1. FLEXIBILITY OF TRANSMEMBRANE HELICES

A striking feature of transmembrane helices is the many obvious distortions from ideality. Overall, about half of all transmembrane helices contain clear bends and some contain multiple bends (1–3). The presence of so many transmembrane helix distortions in membrane proteins was quite surprising. In theory, backbone hydrogen bonds should be much stronger in the apolar membrane environment, which would argue that transmembrane helices should be straight and more rigid (4). However, the straight for a helix introduces an evolutionary problem because it limits the placement of functional side chains (1) and would also restrict movements of helix that may be necessary for function and / or folding (5–7). Besides, kinking of a transmembrane helix has also been found to help recruit water into functional sites (8). There are good reasons why evolution wants to bend transmembrane helices, but, if transmembrane helices are hard to bend, how evolution makes them bent. Or perhaps there is another possibility, which is transmembrane helices are actually easy to bend. In other words, they are flexible. If this is true, then I asked why they are flexible. Previously, researchers have focused on the effect of deleting and placing proline residues in the structure of transmembrane helices (1, 9). The reason why proline is studied is that usually proline cannot be accommodated

in the middle of an α-helix in water-soluble proteins because the side chain removes the backbone hydrogen bond and creates a steric clash with the residue at *i*-4. Thus, proline introduces a distortion in the helix. In fact proline is found at the center of about 20 % of all helix kinks (3, 10). In bR, three of the kinked transmembrane helices have a proline residue around the kink center (1). Therefore, Yohannan *et al*. assumed that proline was creating these kinks and thus they tested this idea by mutating each of them to alanine. Surprisingly, these changes led to no appreciable change in the kinks according to the crystal structures (1). This suggests that at least something else is also responsible for distorting the helix. However, since the P→A mutant helix has the potential to change to the straightened form, it is very possible that the P→A brings a strain that is felt by the bent helix. In my dissertation work, I found out the strain in energetics and the interactions that are stabilizing the bent transmembrane helix in bR. In addition, I also investigated into the conformational changes and shifting of backbone hydrogen-bond patterns among the transmembrane helices of the intermediate states of sarcoplasmic $Ca^{2+}$-ATPase in its functional cycle. From the studies in bR and $Ca^{2+}$-ATPase, I answered the problem whether transmembrane helices are easy to bend (11) as described in Chapter 2.

## 1.2. THE REVERSE UNFOLDING OF BACTERIORHODOPSIN

Understanding protein folding is important in the field of protein science. So far, a profound progress have been made in studying water-soluble protein folding. Although membrane protein folding is also interesting and significant, a crude search for "protein

folding" and for "membrane protein folding" on PubMed yielded 38,633 papers for the former but only 166 for the latter. This difference is due to the difficulty in techniques of studying folding of membrane proteins.

A requirement for studying protein folding, if I want to study the thermodynamics of this process, is to know under which condition folding/unfolding equilibrium of the protein can be reached. For water-soluble proteins this is relatively easy to accomplish by adding denaturants or changing temperature or pressure (12, 13), but for membrane proteins most of those methods will result in irreversible unfolding or aggregation. Previously, Faham *et al*. developed an SDS-driven unfolding method for studying unfolding of one of the helical membrane proteins, bacteriorhodopsin (bR). In their study, SDS was added to bR which was originally solubilized in DMPC/CHAPSO micelles to unfold the bR from the native state, $bR_f$, to the unfolded state in which the chromophore retinal (RET) still bound to Lys216, $bR_u$ (14). Although Faham *et al*. claimed that unfolding $bR_f$ to $bO_u$ is reversible (14), and Booth and co-workers have shown that the main kinetic phases are consistent with this unfolding equilibrium model (15), it has recently turned out that this unfolding process is actually not reversible (16, 17) because RET hydrolysis is spontaneous (18) at a rate faster than the folding in the apparent transition zone of the unfolding curve (17). This means that there should be a mixture of $bR_f$, $bR_u$, unfolded bacterio-opsin ($bO_u$) and free RET in their samples within the apparent transition zone of the unfolding curve and no equilibrium was reached among those species. Therefore, there may be a big error in the unfolding free energy derived from the previous unfolding curves. To solve this problem, I studied the equilibrium of unfolding $bR_f$ to $bO_u$, which is different from the unfolded state that was previously assumed. In my

work as described in Chapter 3 of this dissertation, I proved that unfolding under the new condition is reversible and re-visited the energetic contributions by most of the side-chains in Helix B of bR using this method (16).

## 1.3. HYDROGEN BOND STRENGTH AND EQUILIBRIUM HYDROGEN / DEUTERIUM FRACTIONATION FACTORS

The energetic contribution of hydrogen bonds to protein structure was controversial in the past. At the beginning, it was assumed that backbone hydrogen bonds should be stronger in the fully buried folding core of water-soluble proteins and in membrane proteins than those highly exposed to water. The reason is that the environments for membrane proteins, micelles and lipid bilayers (4, 19), and for the fully buried folding core of water-soluble proteins (20) are more hydrophobic environment than the surface of water-soluble proteins. Another reason is that due to high exposure to water molecule, backbone groups on surface of water-soluble proteins form alternative hydrogen bonds with water, which further impair the strength of intra-molecular hydrogen bonds (4). This assumption is supported by computations and experiments using model compounds (4). However, my research in the membrane proteins bR and $Ca^{2+}$-ATPase indicates that shifting of transmembrane helices is not a big deal (11), which suggests that backbone hydrogen bonds in membrane proteins may not be so strong as expected before. In fact, experiments by employing mutagenesis also suggest that there is almost no difference among the side-chain-side-chain hydrogen bond strengths in membrane proteins, in the fully buried folding core of water-soluble proteins

and on the surface of water-soluble proteins on average (4). However, mutagenesis may not be a reliable way to study side-chain-side-chain hydrogen bond strengths because it may introduce factors which complicate the hydrogen bonding patters in a protein (4, 21). Moreover, classical mutagenesis methods cannot be applicable to backbone atoms. Although Kelly and co-workers have developed a special way to make amide-to-ester mutations (22–24), such method cannot be widely used for all proteins. Besides, all the mutagenesis experiments have to be coupled with reverse unfolding of the protein in order to obtain the unfolding free energy. The requirement of reverse unfolding also limits the range of proteins in which intra-molecular hydrogen bonds need to be measured. Thus, I asked whether the strength of hydrogen bonds formed by backbone amide groups and side chains could be measured experimentally by not mutating any protein residue or unfolding the protein. To answer this question, I focused on the equilibrium hydrogen/deuterium fractionation factors of hydrogen bonds: $\phi$ = ([D]/[H])$_{protein}$/([D]/[H])$_{solvent}$ (25–38). H/D fractionation factors have been measured in several water soluble proteins (26–30, 33–37), but these results could only be interpreted qualitatively at the time since they were not put on an energetic scale. In my study, I measured the relationship between the $\phi$-value and $\Delta G^{HB}$ by using a series of weak acids and weak bases and apolar, aprotic solvents to mimic the formation of weak hydrogen bonds in proteins and obtained a reasonable result for the scale factor (SF) which relates $\phi$-value and $\Delta G^{HB}$ by SF = $\partial(\Delta G^{HB})/\partial(RT\ln\phi)$. By applying the scale factor I determined and using the measured $\phi$-values for several water-soluble proteins (26–30, 33–37), I analyzed the relative free energy of hydrogen bonds formed by backbone groups and side chains in those proteins and its relationship with secondary structures where the hydrogen

bonds are and charges that the hydrogen bonds carry. This work is described in Chapter 5.

## 1.4. MEASUING HYDROGEN BOND STRENGTH IN A MEMBRANE PROTEIN VIA EQUILIBRIUM HYDROGEN / DEUTERIUM FRACTIONATION FACTORS

Last, to obtain the distribution of hydrogen bond strengths in a membrane protein and answer the question whether membrane proteins have stronger hydrogen bonds than water-soluble proteins on average, I investigated in hydrogen bond strengths in the voltage-sensing domain (VSD) of the voltage-dependent potassium-selective channel from *Aeropyrum pernix* ($K_V$AP). Previously, both the crystal and solution-NMR structures of $K_V$AP-VSD were determined in good agreement with each other (39, 40). NOEs from the protein residues to the lipids 1,2-diheptanoyl-sn-glycerol-3-phosphocholine (D7PC), which form micelles solubilizing the protein in the experiment for determining its solution-NMR structure, indicate that the micelle well mimics the chemical environment of a phospholipid bilayer (40). This suggests that the solution-NMR structure, as well as the crystal structure, is quite similar to the native protein structure in cell membrane. Thus, I measured the $\phi$-values of the exchangeable NH groups in this membrane protein by utilizing its well assigned $^{1}H$-$^{15}N$ 2D-NMR spectrum (40). By applying the scale factor that I determined as described in Chapter 4, I obtained the strengths of 70 % of the backbone hydrogen bonds in this membrane protein and compared them with those in water-soluble proteins. In addition, I also analyzed the

distribution of weak and strong backbone hydrogen bonds and discussed its correlation with the head groups of the lipids 1,2-diheptanoyl-sn-glycerol-3-phosphocholine (D7PC), which form the micelles that solubilize this membrane protein. It is the first time strengths for the vast majority of backbone hydrogen bonds in a membrane protein can be measured experimentally. This work is described in Chapter 5.

## 1.5. REFERENCES


1. Yohannan S, Faham S, Yang D, Whitelegge JP, Bowie JU (2004) The evolution of transmembrane helix kinks and the structural diversity of G protein-coupled receptors. *Proc Natl Acad Sci U S A* 101:959–963.

2. Cordes FS, Bright JN, Sansom MSP (2002) Proline-induced distortions of transmembrane helices. *J Mol Biol* 323:951–960.

3. Hall SE, Roberts K, Vaidehi N (2009) Position of helical kinks in membrane protein crystal structures and the accuracy of computational prediction. *J Mol Graph Model* 27:944–50.

4. Bowie JU (2011) Membrane protein folding: how important are hydrogen bonds? *Curr Opin Struct Biol* 21:42–49.

5. Bright JN, Shrivastava IH, Cordes FS, Sansom MSP (2002) Conformational dynamics of helix S6 from Shaker potassium channel: simulation studies. *Biopolymers* 64:303–313.

6. Shi L et al. (2002) Beta2 adrenergic receptor activation. Modulation of the proline kink in transmembrane 6 by a rotamer toggle switch. *J Biol Chem* 277:40989–96.

7. Wigley WC et al. (2002) A protein sequence that can encode native structure by disfavoring alternate conformations. *Nat Struct Biol* 9:381–8.

8. Miyano M, Ago H, Saino H, Hori T, Ida K (2010) Internally bridging water molecule in transmembrane alpha-helical kink. *Curr Opin Struct Biol* 20:456–463.

9. Yohannan S et al. (2004) Proline substitutions are not easily accommodated in a membrane protein. *J Mol Biol* 341:1–6.

10. Meruelo AD, Samish I, Bowie JU (2011) TMKink: A method to predict transmembrane helix kinks. *Protein Sci Publ Protein Soc* 20:1256–1264.

11. Cao Z, Bowie JU (2012) Shifting hydrogen bonds may produce flexible transmembrane helices. *Proc Natl Acad Sci* 109:8121–8126.

12. Brockwell DJ, Smith DA, Radford SE (2000) Protein folding mechanisms: new methods and emerging ideas. *Curr Opin Struct Biol* 10:16–25.

13. Travaglini-Allocatelli C, Ivarsson Y, Jemth P, Gianni S (2009) Folding and stability of globular proteins and implications for function. *Fold Bind Protein-Nucl Acid Interactions* 19:3–7.

14. Faham S et al. (2004) Side-chain contributions to membrane protein structure and stability. *J Mol Biol* 335:297–305.

15. Curnow P, Booth PJ (2007) Combined kinetic and thermodynamic analysis of alpha-helical membrane protein unfolding. *Proc Natl Acad Sci U S A* 104:18970–18975.

16. Cao Z, Schlebach JP, Park C, Bowie JU (2012) Thermodynamic stability of bacteriorhodopsin mutants measured relative to the bacterioopsin unfolded state. *Biochim Biophys Acta* 1818:1049–1054.

17. Schlebach JP, Cao Z, Bowie JU, Park C (2012) Revisiting the folding kinetics of bacteriorhodopsin. *Protein Sci Publ Protein Soc* 21:97–106.

18. Cooper A, Dixon SF, Nutley MA, Robb JL (1987) Mechanism of retinal Schiff base formation and hydrolysis in relation to visual pigment photolysis and regeneration: resonance Raman spectroscopy of a tetrahedral carbinolamine intermediate and oxygen-18 labeling of retinal at the metarhodopsin stage in photoreceptor membranes. *J Am Chem Soc* 109:7254–7263.

19. Bowie JU (2005) Solving the membrane protein folding problem. *Nature* 438:581–589.

20. Hass MAS, Ringkjøbing Jensen M, Led JJ (2008) Probing electric fields in proteins in solution by NMR spectroscopy. *Proteins Struct Funct Bioinforma* 72:333–343.

21. Hammen PK, Klevit RE, Martin Scholtz J, William Anderson J, Bruce Waygood E (1995) Investigation of a side-chain-side-chain hydrogen bond by mutagenesis, thermodynamics, and NMR spectroscopy. *Protein Sci* 4:936–944.

22. Fu Y, Gao J, Bieschke J, Dendle MA, Kelly JW (2006) Amide-to-E-olefin versus amide-to-ester backbone H-bond perturbations: Evaluating the O-O repulsion for extracting H-bond energies. *J Am Chem Soc* 128:15948–15949.

23. Gao J, Kelly JW (2008) Toward quantification of protein backbone-backbone hydrogen bonding energies: An energetic analysis of an amide-to-ester mutation in an alpha-helix within a protein. *Protein Sci Publ Protein Soc* 17:1096–1101.

24. Deechongkit S, Dawson PE, Kelly JW (2004) Toward assessing the position-dependent contributions of backbone hydrogen bonding to beta-sheet folding thermodynamics employing amide-to-ester perturbations. *J Am Chem Soc* 126:16762–16771.

25. Hibbert F, Emsley J (1991) in *Advances in Physical Organic Chemistry*, ed D. Bethell (Academic Press), pp 255–379. Available at: http://www.sciencedirect.com/science/article/pii/S0065316008600477.

26. Shi Z, Krantz BA, Kallenbach N, Sosnick TR (2002) Contribution of Hydrogen Bonding to Protein Stability Estimated from Isotope Effects†. *Biochemistry (Mosc)* 41:2120–2129.

27. Loh SN, Markley JL (1994) Hydrogen Bonding in Proteins As Studied by Amide Hydrogen D/H Fractionation Factors: Application to Staphylococcal Nuclease. *Biochemistry (Mosc)* 33:1029–1036.

28. Bowers PM, Klevit RE (1996) Hydrogen bonding and equilibrium isotope enrichment in histidine-containing proteins. *Nat Struct Mol Biol* 3:522–531.

29. LiWang AC, Bax A (1996) Equilibrium Protium/Deuterium Fractionation of Backbone Amides in U−13C/15N Labeled Human Ubiquitin by Triple Resonance NMR. *J Am Chem Soc* 118:12864–12865.

30. Khare D, Alexander P, Orban J (1999) Hydrogen bonding and equilibrium protium-deuterium fractionation factors in the immunoglobulin G binding domain of protein G. *Biochemistry (Mosc)* 38:3918–3925.

31. Kreevoy MM, Liang TM (2012) Structures and isotopic fractionation factors of complexes, A1HA2-. *J Am Chem Soc* 102:3315–3322.

32. Kreevoy MM, Liang T-M, Chang K-C (2012) Structures and isotopic fractionation factors of complexes AHA-1. *J Am Chem Soc* 99:5207–5209.

33. Halkides CJ, Wu YQ, Murray CJ (1996) A Low-Barrier Hydrogen Bond in Subtilisin: 1H and 15N NMR Studies with Peptidyl Trifluoromethyl Ketones†. *Biochemistry (Mosc)* 35:15941–15948.

34. Harris TK, Abeygunawardana C, Mildvan AS (1997) NMR Studies of the Role of Hydrogen Bonding in the Mechanism of Triosephosphate Isomerase†. *Biochemistry (Mosc)* 36:14661–14675.

35. Lin J, Westler WM, Cleland WW, Markley JL, Frey PA (1998) Fractionation factors and activation energies for exchange of the low barrier hydrogen bonding

proton in peptidyl trifluoromethyl ketone complexes of chymotrypsin. *Proc Natl Acad Sci* 95:14664–14668.

36. Takeda M, Jee J, Terauchi T, Kainosho M (2010) Detection of the Sulfhydryl Groups in Proteins with Slow Hydrogen Exchange Rates and Determination of Their Proton/Deuteron Fractionation Factors Using the Deuterium-Induced Effects on the 13Cβ NMR Signals. *J Am Chem Soc* 132:6254–6260.

37. Markley JL, Westler WM (1996) Protonation-State Dependence of Hydrogen Bond Strengths and Exchange Rates in a Serine Protease Catalytic Triad: Bovine Chymotrypsinogen A†. *Biochemistry (Mosc)* 35:11092–11097.

38. Jarret RM, Saunders M (1985) A new method for obtaining isotopic fractionation data at multiple sites in rapidly exchanging systems. *J Am Chem Soc* 107:2648–2654.

39. Jiang Y et al. (2003) X-ray structure of a voltage-dependent K+ channel. *Nature* 423:33–41.

40. Butterwick JA, MacKinnon R (2010) Solution structure and phospholipid interactions of the isolated voltage-sensor domain from KvAP. *J Mol Biol* 403:591–606.

# CHAPTER 2

# SHIFTING HYDROGEN BONDS MAY PRODUCE FLEXIBLE TRANSMEMBRANE HELICES.

## 2.1. INTRODUCTION

There are many advantages of transmembrane helix kinks for membrane protein structure and function. They create weak points for movement during catalytic cycles (1, 2), they enable the precise positioning of key side chains (3), they can help recruit water to functional sites (4) and they can prevent off-pathway folding (5). It is nevertheless surprising that distortions are so much more common in transmembrane than soluble protein helices (6, 7), because helices are more stable in the apolar membrane environment (8–11) where backbone hydrogen bonds are stronger (12, 13). It has therefore remained mysterious how distortions can possibly be generated by the evolutionary currency of random point mutations. While significant structural fluctuations have been seen in molecular dynamics simulations of isolated transmembrane helices (6, 14–16), there have been no experimental measurements of transmembrane helix deformability.

To learn how hard it is to deform a transmembrane helix, I attempted to engineer an alternative helix conformation into a membrane protein so that the energetic differences could be assessed. I focused on the kink at Pro50 in Helix B of bacteriorhodopsin because it is relatively distant from the retinal chromophore, a useful probe of folding. The helix contains a relatively modest bend ranging from 11-19 °, depending on how it is measured. Proline can play a significant role in creating helix distortions because it is incompatible with a regular helix (6, 7, 17, 18). Nevertheless, a P50A mutation does not remove the kink (19), indicating that other residues participate in bending the helix.

I reasoned that the P50A substitution might introduce an energetic cost to bend the helix because a more regular hydrogen bonding pattern could form if the helix straightened. The putative cost must be overcome by favorable side chain interactions in the bent helix in order to maintain the distortion. It might therefore be possible to break the bend by removing a second key side chain, because there would then be insufficient stabilization to counteract the cost of bending.

Indeed by making a second mutation, T46A, I were able to free the helix to adopt several distinct, straighter conformations. Much to surprise, however, the straightened helices include non-canonical $i \rightarrow i$+3 hydrogen bonds. I estimate that the energetic differences between these conformations are quite modest, which suggests that shifting of hydrogen bonds within transmembrane helices could provide a mechanism for imparting considerable flexibility into transmembrane helices. Consistent with this view, I find that significant conformational changes can occur in transmembrane helices during functional

cycles and that these structural changes are accommodated by frequent shifting to alternative backbone hydrogen bond patterns.

## 2.2. RESUTLS AND DISCUSSION

It is speculated that in the hydrophobic membrane environment, transmembrane helix structure can be maintained in the unfolded state (10, 20). In the case of unfolding of bacteriorhodopsin in micelles, approximately 65% of the helix structure remains intact (11, 21) and recent distance measurements by EPR throughout the protein are consistent with an unfolded state in SDS consisting of mostly helical structure with frayed ends (22). In the case of Helix B specifically, where P50 resides, circular dichroism spectra of a B helix peptide indicates that roughly 19 of the residues remain helical in SDS, whereas an NMR structure of a fragment of bR from residues 1-71 finds that the region from 39-62 within helix B remains helical (23). Thus, the current evidence supports a model in which the B helix loses tertiary contacts and frays at the ends in SDS (22), but the hydrophobic center of the helix surrounding P50 maintains its helical character.

As illustrated in Fig. 2.1*A*, Helix B will remain at least somewhat distorted in the unfolded state of bacteriorhodopin because the proline blocks canonical helix formation. In the P50A mutant, however, the helix is free to adopt a canonical helix in the unfolded state, but it must become bent in the folded structure. In P50A, the helix bend is not a simple breakage of a single hydrogen bond at residue 50, as the hydrogen bond actually remains intact in the P50A structure (19). Rather the bend is created by subtle alterations

involving many backbone atoms and side chains. Thus, to the extent that there is an energetic penalty for bending, it is paid for in subtle ways throughout the helix.

If there is indeed a higher helix bending cost in P50A relative to wild-type that is compensated by other side chains in the helix upon folding, I would expect to see energetic coupling effects between the P50A mutant and other side chains in the Helix B. **I** therefore compared the effects on protein stability of side-chain substitutions in Helix B in both the wild type and P50A backgrounds.

**The P50A mutation produces long-range energetic coupling in Helix B**. A plot of the unfolding free energy contributions of 14 side-chains in the P50A versus the wild-type background, shown in Fig. 2.1*B*, reveals long-range energetic coupling with remarkable uniformity. In particular, the points fall roughly on a straight line with a slope of 0.75 ± 0.05. These results indicate that energetic contribution of a side chain in the P50A background is reduced by ~ 25 % compared with that in the wild-type background. As a control, I performed the same experiment in a different mutant background, M56A, targeting a residue not obviously involved in helix bending. As shown in Fig. 2.1*B* the slope of the line is 1.02 ± 0.02, reflecting no energetic coupling between M56A and other residues in the same helix. Thus, the P50A mutation has long range energetic consequences throughout the helix.

Why are the coupling effects so uniform? In particular, regardless of the energetic importance of the side chain, its contribution in P50A is reduced by ~ 25 %? As outlined in the *Supporting Information Appendix*, this can be explained in terms of a

highly simplified model in which I envision the energetic contributions of each of the side chains is represented by springs of various strengths. I can then envision that the P50A mutation generates a new "straightening force" that is resisted by all the springs. The new counteracting force will be distributed among the springs according to their strengths. By analogy, one can imagine a right handed man holding on to a chin-up bar. If a new weight is now attached to his belt, the additional force will be distributed more toward his stronger right arm than his left. These new forces will ultimately play out in the reduced energetic contribution of each side chain as I observed. I recognize that the situation in the protein is much more complex than this conceptual model, but I believe the basic principle may be operating here.

**Freeing the helix to adopt a new conformation.** The presence of possible helix bending cost in the P50A mutant suggests that it might be possible to free the helix to adopt an alternative conformation by the removal of a side chain that is important in bending the helix. The largest contributor to stability in Helix B is T46, and possibly Y57. A T46A mutation lowers stability by 3.0 kcal/mol. A Y57F mutation lowers stability by ~ 3.9 kcal/mol for the first unfolding transition but its total contribution to the stability of the protein is difficult to assess because of the presence of a low signal for a second transition (see *Supporting Information Appendix* for details). I therefore decided to test whether mutations at T46 and Y57 in the P50A mutant background would free the helix to adopt an alternative conformation by determining structures of the single and double mutants. The structure of P50A was determined previously (19) and is very similar to the wild-type structure with only small changes near the site of the mutation.

For both Y57F and the double mutant, P50A/Y57F, I observed only modest changes in the structure of the protein (*Supporting Information Appendix: Fig. S2.4*). Thus, Y57 does not appear to be a critical residue driving the kink. T46 is different, however. While the backbone structure of the T46A single mutant is very similar to the wild-type protein, the double mutant shows a significant straightening of Helix B (Fig. 2.2*A*; *Supplementary files: Movie S2.1*). T46 is hydrogen bonded to D96 and forms one of the strongest interhelical hydrogen bonds I measured in prior work (24). The straightening of the helix in the P50A/T46A double mutant suggests that this strong interhelical hydrogen bond helps to effectively pin the helix in a bent conformation.

**Non-canonical hydrogen-bond patterns in the new helix conformations.** In the P50A/T46A double mutant, Helix B is noticeably straightened (Fig. 2.2*A*; *Supplementary files: Movie S2.1*), but the structural change cannot be described by a simple straightening of Helix B into a more regular helix. Instead, Helix B adopts a new yet still distorted structure, with remarkable alterations at the atomic level. Fig. 2.2*B* plots the backbone O-N distances for residues *i* to *i*+4 throughout Helix B. For the single mutants P50A and T46A, there are only small differences in the hydrogen bond distances compared to the wild-type protein for both molecules in the crystal asymmetric unit. On the other hand, there are large changes for the P50A/T46A double mutant compared to the wild-type structure as well as between the two molecules in the asymmetric unit of the crystal. Most notably, the canonical $i \rightarrow i+4$ $\alpha$-helical hydrogen bonds are stretched or broken for residues 38, 39, 40 and 42 in molecule A and for residues 38, 39, 41 and 46 in molecule B. As illustrated in Fig. 2.2*C*, the broken hydrogen bond at residue 42 in molecule A is

readily explained by a new ordered water molecule that bridges the side chain of D96 and the carbonyl oxygen of F42. However, as shown in Fig. 2.2*D* the backbone hydrogen bond reforms in molecule B and the water molecule disappears. While this leaves the D96 side chain without an observable hydrogen bond partner, the D96 side chain appears to form a stabilizing electrostatic interaction with the aromatic ring of F219 (25). But how can so many of the other hydrogen bonds break without severe energetic consequences?

Further examination reveals that with one exception, the backbone hydrogen bonds have not disappeared, but have shifted to make new, or improved, $i \rightarrow i+3$ hydrogen bonds (Fig. 2.2 *C* and *D*). Segments of $3_{10}$ helices and π-helices have been observed in transmembrane helices (26–29) and the conversion of $\alpha$-helices to $3_{10}$ helices is thought to be a relatively low energy transition, particularly in an apolar environment (30). But the hydrogen bonding patterns in the P50A/T46A mutant are not consistent with conversion to a regular $3_{10}$ helix and by the criterion used in DSSP (31), the structure remains $\alpha$-helical. Thus, the new conformation is not a canonical $3_{10}$ or $\alpha$-helix.

**Non-canonical hydrogen bonds are parts of a continuum of helix conformations.** Hildebrand *et al.* made the interesting observation that backbone hydrogen bonds in membrane proteins more commonly show bifurcation between the $i \rightarrow i+3$ and $i \rightarrow i+4$ types (32). Moreover, To assess how commonly the $i \rightarrow i+3$ hydrogen bonds found in transmembrane helices are parts of non-canonical helices as opposed to regular $3_{10}$ helix segments, I analyzed a database of transmembrane helices from 41 non-homologous,

high-resolution membrane protein structures described previously (33). Of the 743 $i \rightarrow i+3$ backbone hydrogen bonds identified, only 11% were in contiguous patterns composed of 3 or more $i \rightarrow i+3$ hydrogen bonds, leaving 89% that cannot be part of a $3_{10}$ helix. These results suggest that hydrogen bond shifts are part of a continuum of helical conformations rather than wholesale conversions to new helix types.

**Small energetic cost of helix distortions.** The fact that the P50A/T46A mutant helix adopts a somewhat distorted structure over a more ideal $\alpha$-helix suggests that the hydrogen bond shifts are relatively low energy transitions. To obtain an actual experimental estimate of the energetic consequences of the engineered helix change I engineered in bR, I constructed the thermodynamic cycle illustrated in Fig. 2.3. In essence, I can measure the free energy contributions of the P50A and T46A mutations with and without helix bending. The result suggests that bending the helix only costs ~ 0.6 kcal/mol in the context of the protein. Moreover, the fact that I see multiple helix conformations in the P50A/T46A double mutant indicates that transmembrane helices are relatively flexible when freed from other constraints. Indeed, if bending the helix costs only 0.6 kcal/mol, the conformation seen in the wild-type protein could be explored by thermal motions.

**Rampant hydrogen-bond shifting seen in conformational changes.** If alternative hydrogen bonding patterns provide low energy pathways for helix bending that are accessible by thermal motions, I might also expect to see hydrogen bond shifts during the

conformational changes that occur in catalytic cycles. I therefore examined the backbone hydrogen bonding patterns that occur in the sarcoplasmic $Ca^{2+}$-ATPase, for which there are many structures representing distinct stages in the transport cycle (34). As shown in Fig. 2.4 *A* - *E*, the backbone hydrogen bonding patterns in the transmembrane helices change significantly during pumping. These changes in hydrogen bonding patterns translate into clear conformational changes in the helices. Fig. 2.4*F* and *Supplementary files: Movie S2.2* show a particular helix straightens and kinks during pumping with extensive hydrogen bond shifting. These hydrogen bond shifts are not conversions to alternative regular helices, but rather isolated shifts to accommodate local structural deformations.

In summary, my results indicate that transmembrane helices are quite flexible. Thus, the introduction of kinks appears to be well within the realm of simple evolutionary steps and helix distortions can be readily accessed during conformational changes. I propose that flexibility may at least in part be explained by backbone hydrogen bonding donors and acceptors shifting to different partners. While the introduction of a Pro residue in a helix is one dramatic way to distort helices, my results indicate that the structure of transmembrane helices can be altered readily by a single point mutation because the energy cost for helix bending is not extremely high. In the absence of P50, Helix B in bR can be shifted to distinct conformations by a single mutation (T46A). I had earlier proposed that most kinks are introduced by proline mutations which can then be removed once the kink is fixed in the structure (3). This hypothesis was supported by my finding that at the vast majority of kinks I found prolines in homologous proteins, suggesting an

evolutionary connection. More recent work with much larger databases suggests that these so-called vestigial prolines are less frequent than I had observed (6, 7, 33). Moreover, I found that the introduction of Pro into a transmembrane helix is generally very deleterious and is likely to happen only infrequently (17). Another possibility is that prolines are accommodated favorably after the kink structure is generated by other mutational pathways, which appear to be well within the realm of simple evolutionary steps. My results indicate that membrane protein structure is much more malleable than I might have imagined; a feature that has apparently been essential for optimizing membrane protein structure and function.

## 2.3. MATERIALS AND METHODS

**Preparation of mutant bR proteins.** All mutant bR proteins were prepared as described previously (19, 35).

**$bR_f$-to-$bO_u$ unfolding assays.** The $bR_f$-to-$bO_u$ unfolding assays of wild-type and mutant bR were performed with the addition of excess all-*trans* retinal added as described (36). Fitting equations and fitting parameters are given in Supplementary Information.

**X-ray crystallography.** The mutant bR proteins were crystallized by vapor diffusion using the bicelle method (37, 38). Purple membrane in water at >10 mg/ml was mixed with a 40% (w/v) 2.8:1 DMPC/CHAPSO bicellar solution at a volume ratio of 4:1 and equilibrated on ice for more than 1 hour in dark. The drops contained 2 μL of the protein/bi-

celle solution mixed with 0.75 μL of well solution composed of ~ 3 M sodium phosphate (pH 3.6 - 4.0), 3.5 % triethylene glycerol and 0.03 M 1,6-hexanediol. After grown in dark for 2 - 14 days at 37 °C, crystals were washed in the well solution, dipped into a liquid perfluoropolyether cryoprotectant and frozen in liquid nitrogen. Diffraction data were collected at the Advanced Photon Source beamline 24-ID-C and were integrated and scaled by using the DENZO/SCALEPACK program package (39). Structures were phased by molecular replacement with PHASER (40) using the wild-type bR model 1PY6. T46A and P50A/T46A were refined at resolutions of 2.47 Å and 2.37 Å using CNS (41) and COOT (42) software with the twinning operator ($-h, -k, h + l$) applied to the reciprocal lattice at twin fractions of 0.31 and 0.50, respectively. 5% of the reflections, which were the same as those omitted in the refinement of the wild-type structure 1PY6, were selected to calculate $R_{free}$ for both of the mutants. Composite omit maps for T46A and P50A/T46A were computed by using detwinned reflection data (see *Supporting Information Appendix: Fig.S3.3A and B*). T46A was detwinned as described (43), but for detwinning the highly twinned P50A/T46A crystal I used the method of Redinbo and Yeates (44), employing coordinates from the wild-type structure (1PY6) with residues 46 - 50 deleted. Y57F and P50A/Y57F were refined at resolutions of 2.06 Å and 2.40 Å using REFMAC (40) and COOT (42) software, respectively. 5% of the reflections, which were the same as those omitted in the refinement of the wild-type structure 1XJI, were selected to calculate $R_{free}$ for both of the mutants. Data collection and refinement statistics details are given in the Supplementary Information.

**Backbone hydrogen bonds.** Backbone hydrogen bonds of T46A and P50A/T46A mutant bR were identified using HBPLUS (45) and COOT (42).

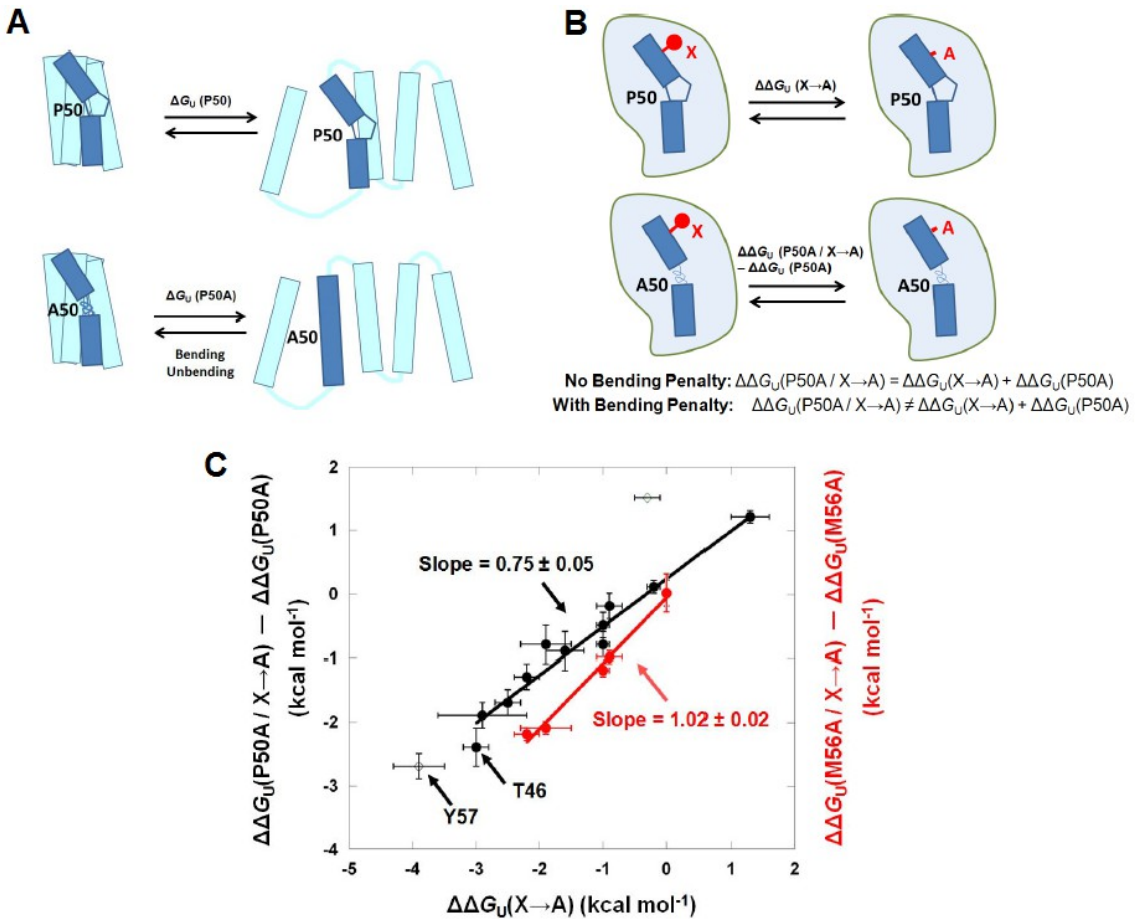


**Fig. 2.1.** Detecting energetic cost of bending Helix B in bR. (*A*) A cartoon depicting the equilibrium between folded and unfolded bacteriorhodopsin molecules. Upon unfolding, Helix B containing P50 is expected to remain at least somewhat distorted (or more flexible) than the same helix containing a P50A mutation. In the case of P50A, the helix is free to adopt a helical structure in the unfolded state, and needs to become bent for the protein to fold properly. (*B*) The experimental scheme. Helix B contains a kink centered around P50. The P50A mutation does not destroy the bend, but I envision a bending cost is imposed on the helix (relative to the unfolded state not depicted here), represented by the spring. To detect possible bending energy I made side chain substitutions to alanine in either the wild-type background or the P50A background. A generic side chain X is represented by the red ball and stick and the mutation to Ala is represented by a stick. I then compare unfolding free energy differences to obtain $\Delta\Delta G_U$(X→A) (the change in unfolding for the single X→A mutations in the wild-type, P50, background compared to the wild-type) and $\Delta\Delta G_U$(P50A/X→A) (the change in unfolding free energies for the

double mutations of X→A in the P50A background compared to the wild-type). If there is no coupling between the sites, the side chain substitution should have the same effect in both the wild-type and P50A backgrounds. A difference indicates an energetic penalty. (*C*) The observed unfolding free energy differences. The changes in unfolding free energy for various side chain, X, substitutions to alanine in either the P50A background (black) or M56A background (red) are plotted against the same substitutions in the wild-type background. The only exception is for the side chain Y57, which was changed to an F because the A substitution was too destabilizing. The black line represents a least squares fit to all the black points. The red line is a least squares fit to all the red points. The error bars reflect the standard deviation of two or more independent measurements and the fitting uncertainty from each experiment. Two mutants, Y57F and V49A, produced anomalous unfolding curves which were analyzed as described in the *Supporting Information Appendix*. These results are shown in open diamonds and were not included in the curve fits.

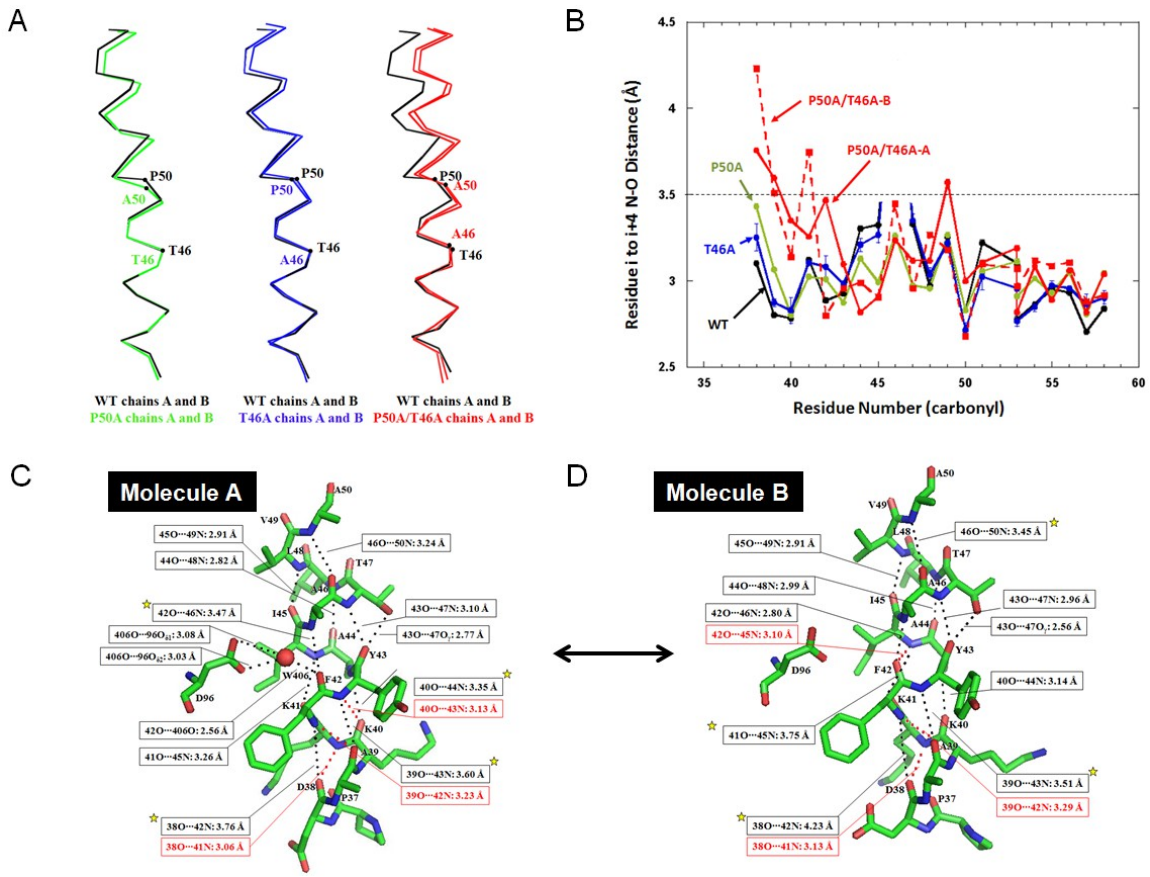


**Fig. 2.2.** Changes in Helix B structure imparted by mutations in bR. (*A*) Superposition of Helix B structures for wild-type bacteriorhodopsin and variants. $C_\alpha$ traces are shown. The wild type structure is shown in black, P50A is shown in green, T46A is shown in blue and P50A/T46A is shown in red. For each structure, both molecules A and B in the unit cell are shown, but are indistinguishable for all but the P50A/T46A double mutant. Only the residues in the N-half of the kinked helix (residues 39 to 46) were employed for the superposition to highlight the change in bending. (*B*) A plot of the backbone O-N distances between residues *i* and *i*+4 throughout helix B for bacteriorhodopsin variants. The wild-type is shown in black, the T46A mutant in blue, the P50A mutant in green and the P50A/T46A double mutant in red. For all the variants but P50A/T46A, the average distances are shown for the two chains in the asymmetric unit of the crystal, because the structures were so similar. The error bars reflect the standard deviation of the distances between molecules A and B. Because the A and B molecules are quite different in the P50A/T46A double mutant, the distances for the two molecules are shown separately. (*C*

and *D*) Detailed hydrogen bonding patterns in the structure P50A/T46A seen for molecule A (panel *C*) and molecule B (panel *D*) in the asymmetric unit of the crystal structure. Backbone O-N distances between residues *i* and *i*+4 for a canonical α-helix and other side chain hydrogen bonds are shown in black. Backbone O-N distances between residues *i* and *i*+3 are shown in red. Residues with severely lengthened *i*→*i*+4 O-N distances are highlighted with the yellow stars.

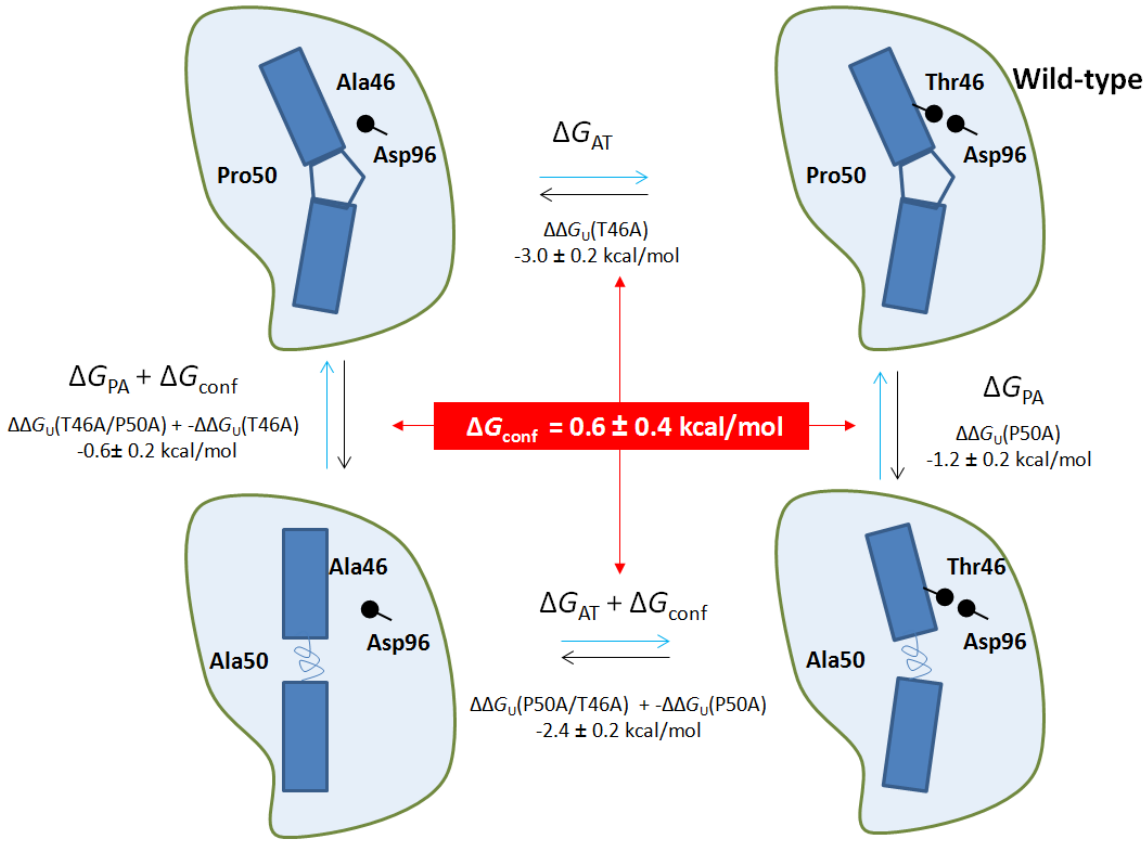


**Fig. 2.3.** Measuring the energetic difference between helix conformations in bR. I have measured the contribution of two mutations, P50A and T46A, to stability with and without the inclusion of a conformational change in the helix. $\Delta G_{AT}$ and $\Delta G_{PA}$ are the free energy differences between the wild-type and the T46A mutant or P50A mutant, respectively. $\Delta G_{conf}$ is the free energy difference between two helix conformation states. The blue arrows indicate the arbitrarily chosen direction of the reactions. The free energy differences obtained from the unfolding free energies are indicated for each reaction. By comparing the horizontal and vertical reactions I can extract $\Delta G_{conf}$, highlighted in red.

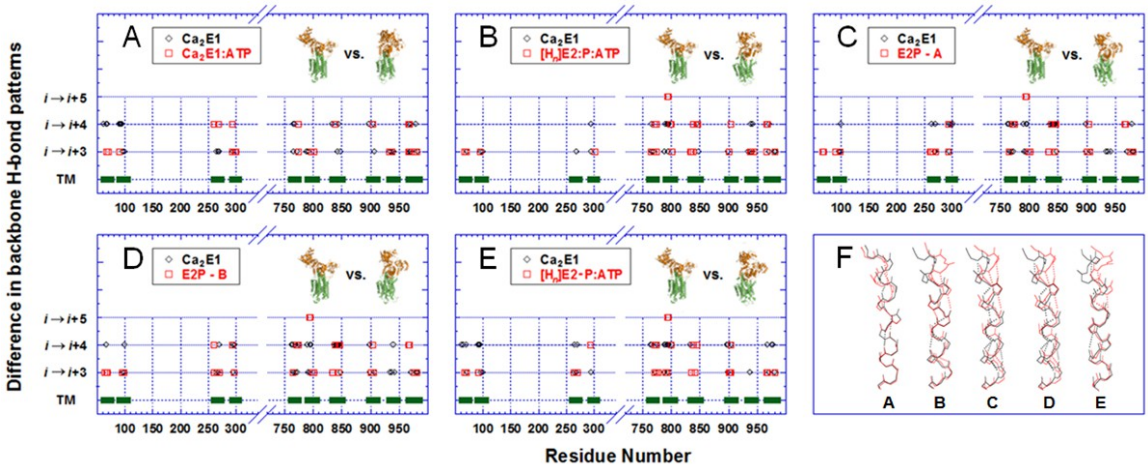


**Fig. 2.4.** Helix flexing by hydrogen bond shifting during conformational changes in $Ca^{2+}$-ATPase. (*A* - *E*) The backbone hydrogen-bond patterns in the TM region of (*A*) $Ca_2E1$:ATP (1T5S), (*B*)[$H_n$]E2P:P:ATP (1WPG), (*C*) E2P molecule A (2ZBE), (*D*)E2P molecule B (2ZBE) and (*E*) [$H_n$]E2-P:ATP (2ZBG) were compared with the backbone hydrogen-bond patterns in the TM region of the ligand-free form, $Ca_2E1$ (1SU4). Backbone hydrogen bonds which are found in the reference model, $Ca_2E1$, but not in the comparing model are labeled with open black diamonds. Backbone hydrogen bonds which are found in the comparing model but not in the reference model are labeled with open red squares. The TM regions are labeled with solid green rectangles. The 5 structures used all had resolutions of 2.7 Å or better. (*F*) An example of the resulting helix alterations for the TM helix spanning residues 762 to 780. The Cα atoms of residues 762 - 770 in the N-half of the helix from each comparing model are aligned with those from the reference model. The hydrogen bonds shown and the color scheme are the same as in panels (*A* - *E*).

## 2.4. SUPPORTING INFORMATION APPENDIX

**Methods for fitting of abnormal SDS-induced $bR_f$-to-$bO_u$ unfolding curves.** The I52A mutant unfolding curves showed two transitions. For this mutant the unfolding free energy, $\Delta G_U$, for each transition was assumed to have a linear relationship with the mole fraction concentration of SDS, $X_{SDS}$:

$$\Delta G_{U,1} = m_1 (X_{SDS} - C_{m,1}),$$

and

$$\Delta G_{U,2} = m_2 (X_{SDS} - C_{m,2}).$$

The total $\Delta G_U$ corresponding to the free energy change of unfolding $bR_f$ to $bO_u$ is the sum of $\Delta G_{U,1}$ and $\Delta G_{U,2}$. In addition, the absorptivities at 560 nm for both the folded and intermediate states are assumed to have a linear response to $X_{SDS}$:

$$\varepsilon_1 = a_1 X_{SDS} + b_1,$$

and

$$\varepsilon_2 = a_2 X_{SDS} + b_2.$$

Thus, the unfolding curves of I52A with double transitions were fit according to the following equation:

$$\mathrm{Abs}_{560} = \frac{(a_1 \cdot X_{SDS} + b_1) \cdot \exp[m_1 \cdot (X_{SDS} - C_{m,1})/(RT)] + (a_2 \cdot X_{SDS} + b_2)}{1 + \exp[m_1 \cdot (X_{SDS} - C_{m,1})/(RT)] + \exp[-m_2 \cdot (X_{SDS} - C_{m,2})/(RT)]},$$

where $RT$ is 0.592 kcal $mol^{-1}$ at room temperature (22.5 ± 0.5 °C). Parameters $a_1$, $a_2$, $b_1$, $b_2$, $m_1$, $m_2$, $C_{m,1}$ and $C_{m,2}$ were fit using Kaleighdagraph 4.1.

For the Y57F and P50A/Y57F mutants, which also have double transitions in their unfolding curves, only the first transition was fit because the two transitions are well separated but the signal for the second transition is too low to fit (Fig. S2.1*A*). The fitting equation is described as following:

$$\mathrm{Abs}_{560} = \frac{(a_1 \cdot X_{\mathrm{SDS}} + b_1) \cdot \exp[m_1 \cdot (X_{\mathrm{SDS}} - C_{\mathrm{m},1})/(RT)] + (a_2 \cdot X_{\mathrm{SDS}} + b_2)}{1 + \exp[m_1 \cdot (X_{\mathrm{SDS}} - C_{\mathrm{m},1})/(RT)]}.$$

Parameters $a_1$, $a_2$, $b_1$, $b_2$, $m_1$ and $C_{\mathrm{m},1}$ were fit using Kaleighdagraph4.1.

For the V49A and P50A/V49A mutants, their absorbance peak shifts by ~ 25 nm in the transition region (Fig. S2.1*B*) as a result of isomerization of the bound RET (46, 47). The unfolding curves were also fit using the single-transition model. However, energetic change corresponding to the RET isomerization is hidden behind the apparent unfolding free energy.

**Table S2.1.** Thermodynamic parameters of SDS-induced unfolding of bR variants.

| bR variants | $m$ (kcal mol$^{-1}$) [a] | $C_m$ [b] | $\Delta\Delta G_U$ (kcal mol$^{-1}$) [c] |
|---|---|---|---|
| WT | – 46.1 ± 5.5 | 0.609 ± 0.004 | 0.0 ± 0.0 |
| K41A | – 51.5 ± 3.3 | 0.590 ± 0.001 | – 1.0 ± 0.2 |
| F42A | – 45.3 ± 2.6 | 0.554 ± 0.004 | – 2.5 ± 0.2 |
| Y43A | – 41.7 ± 3.7 | 0.566 ± 0.003 | – 1.9 ± 0.4 |
| I45A | – 43.3 ± 3.0 | 0.558 ± 0.005 | – 2.2 ± 0.2 |
| T46A | – 42.7 ± 1.8 | 0.540 ± 0.002 | – 3.0 ± 0.2 |
| T47A | – 41.7 ± 4.7 | 0.585 ± 0.004 | – 1.0 ± 0.1 |
| V49A | – 48.1 ± 2.7 | 0.605 ± 0.002 | – 0.3 ± 0.2 |
| P50A | – 45.1 ± 3.1 | 0.584 ± 0.004 | – 1.2 ± 0.2 |
| I52A | | | |
| 1st transition | – 51.2 ± 3.4 | 0.555 ± 0.004 | – 2.8 ± 0.3 |
| 2nd transition | – 41.6 ± 16.3 | 0.609 ± 0.009 | – 0.1 ± 0.3 |
| Total | | | – 2.9 ± 0.6 |
| F54A | – 59.9 ± 7.1 | 0.597 ± 0.002 | – 0.8 ± 0.2 |
| M56A | – 42.3 ± 4.1 | 0.641 ± 0.003 | 1.3 ± 0.3 |
| Y57F | | | |
| 1st transition | – 39.6 ± 3.3 | 0.511 ± 0.004 | – 3.9 ± 0.4 |
| S59A | – 54.8 ± 3.4 | 0.606 ± 0.001 | – 0.2 ± 0.1 |
| M60A | – 43.3 ± 3.4 | 0.575 ± 0.004 | – 1.6 ± 0.3 |
| P50A/K41A | – 40.8 ± 3.7 | 0.571 ± 0.002 | – 1.7 ± 0.2 |
| P50A/F42A | – 41.5 ± 3.1 | 0.534 ± 0.003 | – 2.9 ± 0.2 |
| P50A/Y43A | – 45.3 ± 2.6 | 0.569 ± 0.001 | – 1.9 ± 0.2 |
| P50A/I45A | – 38.4 ± 3.0 | 0.543 ± 0.004 | – 2.5 ± 0.2 |
| P50A/T46A | – 42.2 ± 2.9 | 0.524 ± 0.003 | – 3.6 ± 0.3 |
| P50A/T47A | – 43.8 ± 4.7 | 0.563 ± 0.003 | – 2.0 ± 0.2 |
| P50A/V49A | – 47.0 ± 6.2 | 0.615 ± 0.003 | 0.3 ± 0.1 |
| P50A/I52A | – 48.7 ± 3.1 | 0.544 ± 0.003 | – 3.1 ± 0.2 |
| P50A/F54A | – 47.2 ± 6.6 | 0.582 ± 0.002 | – 1.4 ± 0.2 |
| P50A/M56A | – 47.7 ± 5.0 | 0.611 ± 0.003 | 0.0 ± 0.2 |
| P50A/Y57F | | | |
| 1st transition | – 34.3 ± 3.1 | 0.498 ± 0.003 | – 3.9 ± 0.2 |
| P50A/S59A | – 40.9 ± 2.4 | 0.589 ± 0.001 | – 1.1 ± 0.2 |
| P50A/M60A | – 40.8 ± 2.7 | 0.554 ± 0.001 | – 2.2 ± 0.3 |
| M56A/Y43A | – 45.4 ± 2.6 | 0.593 ± 0.002 | – 0.8 ± 0.2 |
| M56A/I45A | – 51.4 ± 6.0 | 0.592 ± 0.003 | – 0.9 ± 0.2 |
| M56A/T47A | – 46.3 ± 2.2 | 0.610 ± 0.001 | 0.1 ± 0.1 |
| M56A/F54A | – 37.7 ± 2.3 | 0.619 ± 0.003 | 0.4 ± 0.2 |

[a] Dependence of unfolding free energy on SDS mole fraction concentration.

[b] SDS mole fraction concentration at the midpoint of a transition.

[c] Change in stability compared with wild-type at the wild-type $C_m$.

**Table S2.2.** X-ray data collection and refinement statistics.

| bR mutants | T46A | P50A/T46A | Y57F | P50A/Y57F |
|---|---|---|---|---|
| **Data collection** | | | | |
| Wavelength (Å) | | | | |
| Space Group | $P\,2_1$ | $P\,2_1$ | $C\,2\,2\,2_1$ | $C\,2\,2\,2_1$ |
| Cell dimensions | | | | |
| *a*, *b*, *c* (Å) | 45.35, 108.67, 56.18 | 45.78, 116.46, 57.08 | 46.08, 104.01, 129.74 | 45.28, 103.06, 128.88 |
| *α*, *β*, *γ* (°) | 90, 113.56, 90 | 90, 112.55, 90 | 90, 90, 90 | 90, 90, 90 |
| Resolution (Å) | 90.00 – 2.47 (2.56 – 2.47) | 90.00 – 2.37 (2.45 – 2.37) | 100.00 – 2.06 (2.13 – 2.06) | 100.00 – 2.40 (2.49 – 2.40) |
| $R_{sym}$ | 0.141 (0.492) | 0.177 (0.425) | 0.140 (0.466) | 0.140 (0.457) |
| $I/\sigma I$ | 8.46 (2.60) | 5.68 (2.09) | 12.40 (3.94) | 7.37 (2.03) |
| Completeness (%) | 99.7 (99.8) | 97.4 (96.2) | 98.8 (97.7) | 95.9 (98.8) |
| Redundancy | 3.5 (3.5) | 3.2 (3.0) | 6.3 (6.2) | 2.8 (2.8) |
| **Refinement** [a] | | | | |
| Resolution (Å) | 51.50 – 2.47 | 39.74 – 2.37 | 40.07 – 2.06 | 47.85 – 2.40 |
| No. Reflections | 17877 | 21867 | 19416 | 11633 |
| $R_{work}$ / $R_{free}$ | 0.209 / 0.249 | 0.214 / 0.239 | 0.198 / 0.211 | 0.203 / 0.246 |
| No. Atoms | | | | |
| Protein | 3,510 | 3,506 | 1,755 | 1,753 |
| Ligand/ion | 52 | 52 | 66 | 87 |
| Water | 11 | 7 | 10 | 9 |
| *B*-factors (Å$^2$) | | | | |
| Protein | 49.3 | 36.7 | 29.0 | 28.9 |
| Ligand/ion | 47.7 | 31.6 | 56.3 | 61.6 |
| Water | 38.2 | 24.5 | 37.0 | 39.3 |
| r.m.s. Deviation | | | | |
| Bond Length (Å) | 0.015 | 0.013 | 0.0255 | 0.023 |
| Bond Angle (°) | 1.747 | 2.016 | 1.952 | 2.489 |
| Ramachandran Stats (%) | | | | |
| Core | 96.6 | 96.2 | 97.9 | 99.0 |
| Allowed Region | 3.4 | 3.8 | 2.1 | 1.0 |
| Gen. Allowed Region | 0 | 0 | 0 | 0 |
| Disallowed Region | 0 | 0 | 0 | 0 |

[a] Twinning operator (*-h*, *-k*, *h+l*) was applied to the reciprocal lattices of T46A and P50A/T46A.

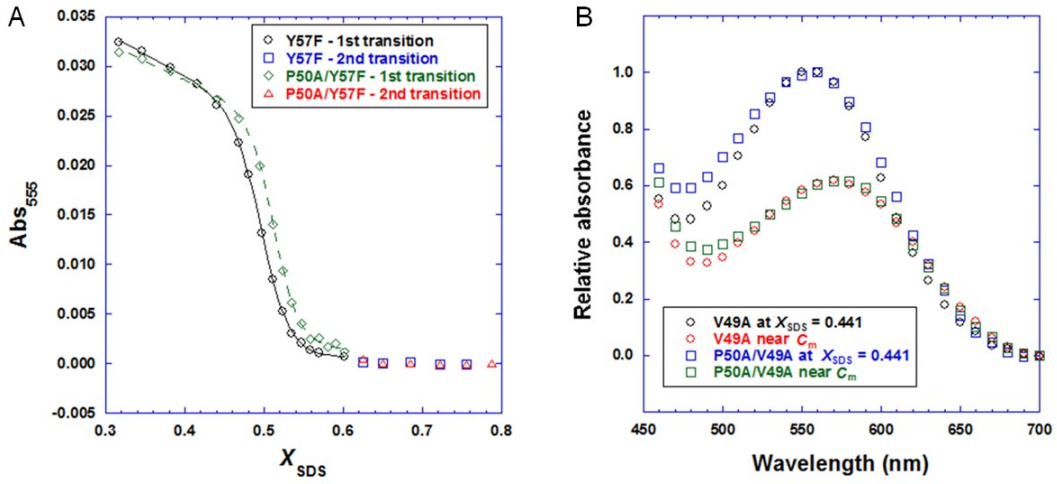


**Fig. S2.1.** Unfolding complexities in four bR mutants. (*A*) Separate double transitions are observed in the unfolding curves of both Y57F and P50A/Y57F. Signals for the second transition are not high enough for fitting the whole curve using a double-transition model. The solid curves reflect the fitting of the first transition only. (*B*) A large shift of absorbance spectra was observed during the unfolding of V49A and P50A/V49A. Compared with the peak at $X_{SDS} = 0.441$ in the sub-denaturant region, the peak at a point near $C_m$ shifts by ~ 25 nm for each of the two mutants.

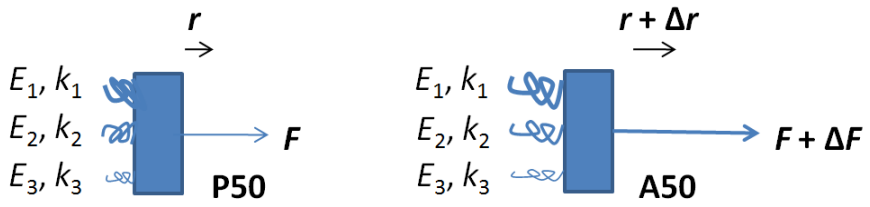


**Fig. S2.2.** A conceptual model to explain the finding that the P50A mutant uniformly reduces the energetic contribution of distant side chains by a constant fraction. I imagine that I can represent each side chain by a spring. For simplicity, I will assume that the springs are attached to a single rigid body that represents half of Helix B. In the figure, only 3 springs (side chains) are shown. Different side chains have varying energetic contributions, which I can model with different spring strengths with elastic constants $k_1$, $k_2$, $k_3$, …, $k_X$ (different-sized springs in the figure). I will call the energy cost for breaking the springs starting from zero stretch $E_1$, $E_2$, $E_3$, …, $E_X$. In the folded wild type protein, the spring is at some equilibrium position, stretched by $r$ away from zero-stretch position. In P50A, I envision that the bent conformation has a higher energy so that there is effectively a force that is pulling it away from the equilibrium position seen for the wild type. The displacement of the whole rigid body changes from $r$ to $r + \Delta r$ and the force it feels changes from $F$ to $F + \Delta F$ in the presence of the P50A mutation. Note that these displacements may be too small to see at the resolution of my structures, although some structural perturbations are visible. In the wild-type protein the springs are stretched by $r$, away from the zero-stretch position. Thus the energetic change due to the breaking of an arbitrary spring X (corresponding to the energetic contribution of the residue in the wild-type background), is

$$\Delta\Delta G_U(\text{single mutant}) = E_X - (1/2)\, k_X r^2. \qquad \text{(Eq. S2.1)}$$

In other words, the energetic contribution of the side chain spring is reduced according to how far it is stretched from its optimal position. I assume that $E_X$ is proportional to $k_X$ by a coefficient $A$ which is only dependent on the atom-atom distance at the zero-stretch position, *i.e.*

$$E_X = A\, k_X,$$

then Eq. S2.1 can be re-written as

$$\Delta\Delta G_U(\text{single mutant}) = A\, k_X - (1/2)\, k_X r^2 = [A - (1/2)\, r^2]\, k_X. \qquad \text{(Eq. S2.2)}$$

Suppose that the displacement of the whole rigid body changes by $\Delta r$ when P50A mutation takes place. Similar to Eqs. S2.1 and S2.2, the energetic change due to the breaking of an arbitrary spring X, represented by the energetic contribution of the corresponding residue in the P50A background, is

$\Delta\Delta G_U(\text{double mutant}) - \Delta\Delta G_U(\text{P50A})$

$$= E_X - (1/2)\, k_X (r + \Delta r)^2 = A\, k_X - (1/2)\, k_X (r + \Delta r)^2 = [A - (1/2)\, (r + \Delta r)^2]\, k_X.$$

Then the ratio between the energetic contributions of an arbitrary residue in the P50A and wild-type backgrounds will be a constant number irrelevant to the strength of the interaction in which this residue is involved, $k_X$, as shown by the following equation:

$$\frac{\Delta\Delta G_U(\text{double mutant}) - \Delta\Delta G_U(\text{P50A})}{\Delta\Delta G_U(\text{single mutant})} = \frac{[A - \frac{1}{2}(r + \Delta r)^2]}{[A - \frac{1}{2} r^2]}. \qquad \text{(Eq. S2.3)}$$

If I re-arrange Eq. S2.3, I can get the linear dependence of $\Delta\Delta G_U$(double mutant) on $\Delta\Delta G_U$(single mutant) with a slope less than 1 and a *y*-intercept equal to $\Delta\Delta G_U$(P50A) as follows:

$$\Delta\Delta G_U(\text{double mutant}) = \frac{[A - \frac{1}{2}(r + \Delta r)^2]}{[A - \frac{1}{2} r^2]} \cdot \Delta\Delta G_U(\text{single mutant}) + \Delta\Delta G_U(\text{P50A}).$$

Thus, this simple conceptual model explains the main features seen in Fig. 2.1*B*.

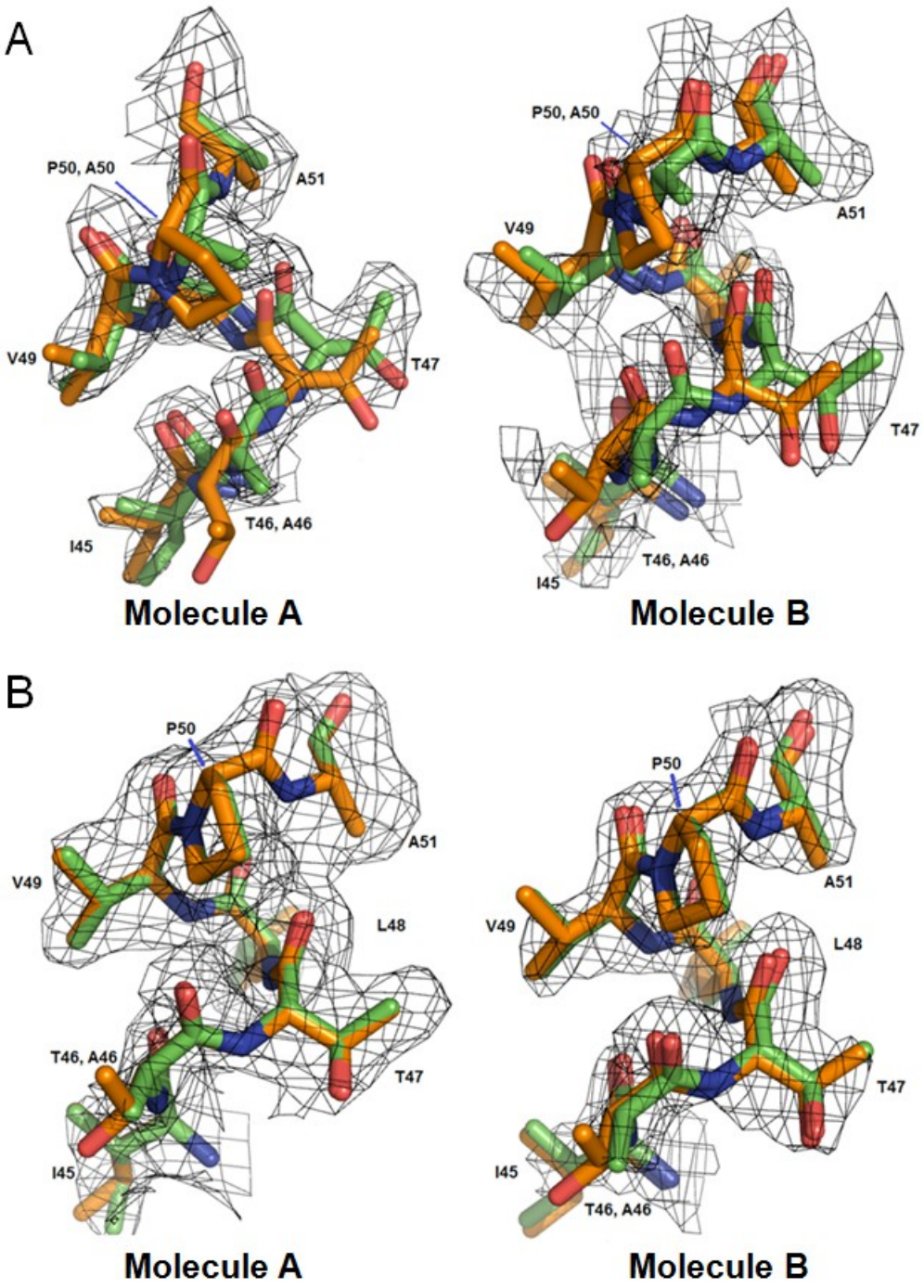


**Fig. S2.3.** Composite omit maps of bR mutants. (*A*) P50A/T46A composite omit maps of the central region of Helix B. (*B*) T46A composite omit maps of the central region of Helix B. For both panels, the P50A/T46A or T46A structure is colored in dark green and

the wild-type structure as a reference model is colored in orange. The P50A/T46A and T46A structures are aligned with the wild-type structure by using the Cα atoms of all the residues. Residue types and numbers are labeled beside each side chain.

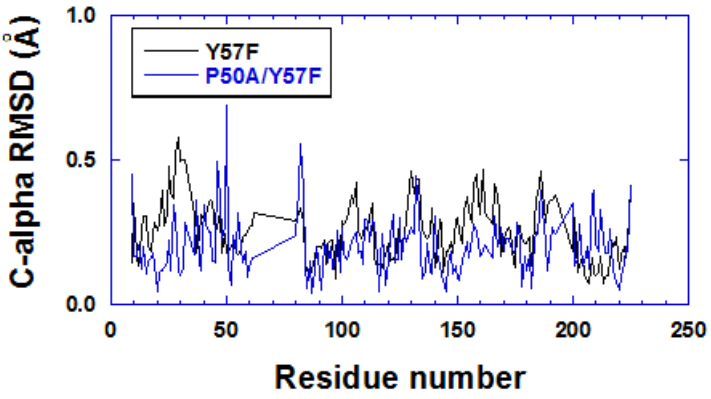


**Fig. S2.4.** Comparison of Y57F and P50A/Y57F with wild-type bR. The plot shows the root mean square deviation (RMSD) of the Cα coordinates of the transmembrane helices after alignment with the wild-type structure (1XJI) using Cα atoms from all residues. Only modest structural differences (RMSD < 0.7 Å) are observed throughout the helical region.

## 2.5. REFERENCES


1. Bright JN, Shrivastava IH, Cordes FS, Sansom MSP (2002) Conformational dynamics of helix S6 from Shaker potassium channel: simulation studies. *Biopolymers* 64:303-313.

2. Shi L et al. (2002) Beta2 adrenergic receptor activation. Modulation of the proline kink in transmembrane 6 by a rotamer toggle switch. *J Biol Chem* 277:40989-96.

3. Yohannan S, Faham S, Yang D, Whitelegge JP, Bowie JU (2004) The evolution of transmembrane helix kinks and the structural diversity of G protein-coupled receptors. *Proc Natl Acad Sci U S A* 101:959-63.

4. Miyano M, Ago H, Saino H, Hori T, Ida K (2010) Internally bridging water molecule in transmembrane alpha-helical kink. *Curr Opin Struct Biol* 20:456-463.

5. Wigley WC et al. (2002) A protein sequence that can encode native structure by disfavoring alternate conformations. *Nat Struct Biol* 9:381-8.

6. Hall SE, Roberts K, Vaidehi N (2009) Position of helical kinks in membrane protein crystal structures and the accuracy of computational prediction. *J Mol Graph Model* 27:944-50.

7. Langelaan DN, Wieczorek M, Blouin C, Rainey JK (2010) Improved helix and kink characterization in membrane proteins allows evaluation of kink sequence predictors. *J Chem Inf Model* 50:2213-2220.

8. Lau FW, Bowie JU (1997) A method for assessing the stability of a membrane protein. *Biochemistry* 36:5884-5892.

9. Haltia T, Freire E (1995) Forces and factors that contribute to the structural stability of membrane proteins. *Biochim Biophys Acta* 1241:295-322.

10. Engelman DM et al. (2003) Membrane protein folding: beyond the two stage model. *FEBS Lett* 555:122-125.

11. Riley ML, Wallace BA, Flitsch SL, Booth PJ (1997) Slow alpha helix formation during folding of a membrane protein. *Biochemistry* 36:192-196.

12. Bowie JU (2011) Membrane protein folding: how important are hydrogen bonds? *Current Opinion in Structural Biology* 21:42-49.

13. White SH, Wimley WC (1999) Membrane protein folding and stability: physical principles. *Annu Rev Biophys Biomol Struct* 28:319-365.

14. Bright JN, Shrivastava IH, Cordes FS, Sansom MSP (2002) Conformational dynamics of helix S6 from Shaker potassium channel: simulation studies. *Biopolymers* 64:303-313.

15. D'Rozario RSG, Sansom MSP (2008) Helix dynamics in a membrane transport protein: comparative simulations of the glycerol-3-phosphate transporter and its constituent helices. *Mol Membr Biol* 25:571-583.

16. Quint S et al. (2010) Residue-specific side-chain packing determines the backbone dynamics of transmembrane model helices. *Biophys J* 99:2541-2549.

17. Yohannan S et al. (2004) Proline substitutions are not easily accommodated in a membrane protein. *J Mol Biol* 341:1-6.

18. Cordes FS, Bright JN, Sansom MSP (2002) Proline-induced distortions of transmembrane helices. *J Mol Biol* 323:951-960.

19. Faham S et al. (2004) Side-chain contributions to membrane protein structure and stability. *J Mol Biol* 335:297-305.

20. Bowie JU (2005) Solving the membrane protein folding problem. *Nature* 438:581-589.

21. London E, Khorana HG (1982) Denaturation and renaturation of bacteriorhodopsin in detergents and lipid-detergent mixtures. *J Biol Chem* 257:7003-7011.

22. Krishnamani V, Hegde BG, Langen R, Lanyi JK (2012) Secondary and Tertiary Structure of Bacteriorhodopsin in the SDS Denatured State. *Biochemistry*. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22242919 [Accessed January 18, 2012].

23. Pervushin KV, Orekhov VYu, Popov AI, Musina LYu, Arseniev AS (1994) Three-dimensional structure of (1-71)bacterioopsin solubilized in methanol/chloroform and SDS micelles determined by 15N-1H heteronuclear NMR spectroscopy. *Eur J Biochem* 219:571-583.

24. Joh NH et al. (2008) Modest stabilization by most hydrogen-bonded side-chain interactions in membrane proteins. *Nature* 453:1266-1270.

25. Philip V et al. (2011) A survey of aspartate-phenylalanine and glutamate-phenylalanine interactions in the protein data bank: searching for anion-π pairs. *Biochemistry* 50:2939-2950.

26. Riek RP, Rigoutsos I, Novotny J, Graham RM (2001) Non-[alpha]-helical elements modulate polytopic membrane protein architecture. *Journal of Molecular Biology* 306:349-362.

27. Choe S, Grabe M (2009) Conformational dynamics of the inner pore helix of voltage-gated potassium channels. *J Chem Phys* 130:215103.

28. Bjelkmar P, Niemelä PS, Vattulainen I, Lindahl E (2009) Conformational changes and slow dynamics through microsecond polarized atomistic molecular simulation of an integral Kv1.2 ion channel. *PLoS Comput Biol* 5:e1000289.

29. Khalili-Araghi F et al. (2010) Calculation of the gating charge for the Kv1.2 voltage-activated potassium channel. *Biophys J* 98:2189-2198.

30. Smythe ML, Huston SE, Marshall GR (1995) The Molten Helix: Effects of Solvation on the .alpha.- to 310-Helical Transition. *J Am Chem Soc* 117:5445-5452.

31. Kabsch W, Sander C (1983) Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* 22:2577-2637.

32. Hildebrand PW, Preissner R, Frömmel C (2004) Structural features of transmembrane helices. *FEBS Lett* 559:145-151.

33. Meruelo AD, Samish I, Bowie JU (2011) TMKink: A method to predict transmembrane helix kinks. *Protein Sci* 20:1256-1264.

34. Møller JV, Olesen C, Winther A-ML, Nissen P (2010) The sarcoplasmic Ca2+-ATPase: design of a perfect chemi-osmotic pump. *Q Rev Biophys* 43:501-566.

35. Oesterhelt D, Stoeckenius W (1974) Isolation of the cell membrane of Halobacterium halobium and its fractionation into red and purple membrane. *Meth Enzymol* 31:667-678.

36. Cao Z, Schlebach J, Park C, Bowie JU (2011) Thermodynamic stability of bacteriorhodopsin mutants measured relative to the bacterioopsin unfolded state. *Biochimica Et Biophysica Acta*. Available at: http://www.ncbi.nlm.nih.gov/pubmed/21880269 [Accessed November 17, 2011].

37. Faham S, Bowie JU (2002) Bicelle crystallization: a new method for crystallizing membrane proteins yields a monomeric bacteriorhodopsin structure. *J Mol Biol* 316:1-6.

38. Faham S et al. (2005) Crystallization of bacteriorhodopsin from bicelle formulations at room temperature. *Protein Sci* 14:836-840.

39. Otwinowski Z, Minor W (1997) in *Macromolecular Crystallography Part A* (Academic Press), pp 307-326. Available at: http://www.sciencedirect.com/science/article/pii/S007668799776066X [Accessed July 20, 2011].

40. The CCP4 suite: programs for protein crystallography (1994) *Acta Crystallogr D Biol Crystallogr* 50:760-763.

41. Brunger AT (2007) Version 1.2 of the Crystallography and NMR system. *Nat Protoc* 2:2728-2733.

42. Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* 66:486-501.

43. Yeates TO (1997) Detecting and overcoming crystal twinning. *Meth Enzymol* 276:344-358.

44. Redinbo MR, Yeates TO (1993) Structure determination of plastocyanin from a specimen with a hemihedral twinning fraction of one-half. *Acta Crystallogr D Biol Crystallogr* 49:375-380.

45. McDonald IK, Thornton JM (1994) Satisfying hydrogen bonding potential in proteins. *J Mol Biol* 238:777-793.

46. Fischer U, Oesterhelt D (1979) Chromophore equilibria in bacteriorhodopsin. *Biophys J* 28:211-230.

47. Chen D-liang, Wang G-yu, Xu B, Hu K-sheng (2002) All-trans to 13-cis retinal isomerization in light-adapted bacteriorhodopsin at acidic pH. *J Photochem Photobiol B, Biol* 66:188-194.

# CHAPTER 3

# THERMODYNAMIC STABILITY OF BACTERIORODOPSIN MUTANTS MEASURED RELATIVE TO THE BACTERIOOPSIN UNFOLDED STATE

## 3.1. INTRODUCTION

Measuring protein stability requires a method to drive the folding equilibrium in favor of the unfolded state. For water-soluble proteins this is typically accomplished using urea, Gdn·HCl or temperature. However, thermal unfolding has not been found to be reversible for any helical membrane proteins (1–6). Urea and Gdn·HCl have been particularly effective for beta-barrel membrane proteins (7) and have been found to reversibly unfold a few large helical membrane protein (7, 8). An alternative to urea and Gdn·HCl is to use a denaturing detergent (9, 10). SDS has the advantage that the protein is maintained in a micelle environment, which leads to maintenance of considerable secondary structure and may therefore be a better mimic of the more structured unfolded state in a bilayer (9, 11–14).

Bacteriorhodopsin can be completely refolded from an SDS unfolded state (11, 15, 16). Titration of bR with SDS from a DMPC/CHAPSO or DMPC/CHAPS solution, leads to $bR_f$-to-$bR_u$ unfolding curves that are reasonably stable over the course of an hour (see below). Moreover, kinetic constants in DMPC/CHAPS conditions indicate that unfolding to $bR_u$ should reach equilibrium in a matter of minutes (17–19). The unfolding curves are also well modeled by a two-state equilibrium and the extrapolated kinetic constants of the unfolding and refolding reac-

tions are consistent with extracted equilibrium constants within the transition zones (17, 18). Thus, I have assumed that the unfolding transitions reflect a folding equilibrium, in spite of the fact that the most rigorous test of equilibrium is not possible, *i.e.*, overlap of the unfolding and refolding curves. I have now come to the conclusion, however, that true equilibrium cannot be achieved under the DMPC/CHAPSO conditions I have used in the past (16 mM DMPC, 6 mM CHAPSO) (20–25), because RET hydrolysis exceeds the rate of refolding near the midpoint of the observed unfolding transition (see below).

Instead of measuring the $bR_f$-to-$bR_u$ transition that is accessed from unfolding at short time scales, it is also possible to measure the $bR_f$-to-$bO_u$ transition if the reaction is allowed to proceed for many days. The advantage of the $bR_f$-to-$bO_u$ transition is that it can be rigorously shown that the reaction is at equilibrium in the transition zones because the folding and unfolding curves are essentially the same (26). To measure the error associated with the equilibrium assumption for the $bR_f$-to-$bR_u$ transition, I compared the effects of mutations on both reactions. Fortuitously, the results are remarkably similar for both measurements.

## 3.2. RESULTS

The $bR_f$-to-$bR_u$ reaction is complicated by the slow hydrolysis of RET, which occurs from the fully unfolded protein with a half life of around 12 min (17). Because loss of RET is slow until there is a significant fraction of unfolded protein, the unfolding curves are relatively stable over a modest time scale. Figure 2.1A shows the change in unfolding curves over time. Minimal change is seen over the first 30 min. These results, combined with measurements of unfolding rates under similar conditions (17, 18), which indicate that unfolding should be

achieved in a matter of minutes, suggest that the $bR_f$-to-$bR_u$ unfolding curves can be used to extract equilibrium constants over a useable time scale. Nevertheless, due to RET hydrolysis during refolding, the refolding curves do not superimpose with the unfolding curves (Figure 2.1B). At low $X_{SDS}$, where refolding is rapid, the protein refolds to near completion, but in the transition zones where refolding is slow, RET hydrolysis becomes a significant factor, complicating the interpretation.

I therefore sought a further test of the equilibrium assumption. A simplified reaction scheme for unfolding followed by essentially irreversible RET hydrolysis is given by:

$$bR_f \underset{k_f}{\overset{k_u}{\rightleftharpoons}} bR_u \xrightarrow{k_{hyd}} bO_u + RET$$

where $k_u$ is the unfolding rate constant, $k_f$ is the refolding rate constant and $k_{hyd}$ is the RET hydrolysis rate constant. For equilibrium between $bR_f$ and $bR_u$ to be achieved, $k_f$ must be greater than $k_{hyd}$. To test whether this is true in the transition zone, I rapidly unfolded the protein to $bR_u$ at high SDS concentration ($X_{SDS}$ = 0.78), then jumped back an SDS concentration at the midpoint of the observed transition ($X_{SDS}$ = 0.67) and monitored both refolding at 560 nm and hydrolysis at 390 nm. As shown in Figure 2.1C, I essentially saw no refolding before hydrolysis was complete. Thus, under the solution conditions I have used for measuring stability, equilibrium cannot be established near the midpoint of the transition and the extent of the observed unfolding must be limited by the unfolding rate.

I therefore measured the unfolding free energy differences for a variety of mutants using an alternative thermodynamic stability measurement. If the unfolding reactions are left to incubate, a $bR_f$-to-$bO_u$ equilibrium is attained as described by Chen and Gouaux (26): $bR_f \leftrightarrow bO_u$ + RET. In my hands it took 9 days to reach equilibrium at room temperature as judged by the

nearly superimposable folding and unfolding curves monitored by absorbance at 560 nm and fluorescence at 335 nm (Figure 2.2A and B). By adding excess RET in the samples, I shortened the equilibration time to 4 days and the folding and unfolding curves monitored by absorbance at 560 nm and fluorescence at 335 nm (Figure 2.2C and D) were more superimposable. The SDS concentration at the midpoint of the transition zone, $C_m$, was much higher for the $bR_f$-to-$bR_u$ reaction (17, 18) than for the $bR_f$-to-$bO_u$ reaction, which is consistent with the spontaneous hydrolysis of the RET Schiff base in SDS solutions at neutral pH (27). In addition, the only peaks observed in the absorbance spectra throughout the transition were at 560 nm, corresponding to $bR_f$, and at 390 nm, corresponding to free RET (Figure 2.3). No indications of a contribution from $bR_u$ at 440 nm were evident, which indicates that the product of the unfolding experiment was bO and free RET but not $bR_u$ (15, 17, 26). The unfolding curves derived from $Abs_{560}$ and $Flu_{335}$ were essentially the same (Figure 2.4), and the absorbance spectra exhibited an isosbestic point (Figure 2.3), suggesting that the unfolding of $bR_f$ to $bO_u$ can be described by a two-state model. As the $bR_f$-to-$bO_u$ reaction should release RET, I expected that the addition of free RET would drive the equilibrium to higher SDS concentrations. Indeed, this is what I observed (Figure 2.4), and is consistent with a thermodynamic equilibrium process. The addition of free RET has two advantages: (1) the concentration of RET becomes a constant, simplifying analysis of the unfolding curves and (2) the time it took to reach equilibrium decreases from 9 days to 4 days.

The $\Delta\Delta G_u$ values obtained using the $bR_f$-to-$bO_u$ transition and $\Delta\Delta G_u^{app}$ values using the $bR_f$-to-$bR_u$ transition were measured for 13 mutants and are listed in Table 1. I compare measurements at the center of the transition zones for the wild-type protein to minimize extrapolations outside of the observable range of fraction unfolded ($X_{SDS}$ = 0.572 for the $bR_f$-to-

$bO_u$ transition without added RET, $X_{SDS} = 0.611$ for the $bR_f$-to-$bO_u$ transition in the presence of 11.2 μM RET, and $X_{SDS} = 0.673$ for the $bR_f$-to-$bR_u$ transition). For 12 of the 13 mutants, the $\Delta\Delta G_u^{app}$ and $\Delta\Delta G_u$ values measured using the two transitions $bR_f$-to-$bR_u$ and $bR_f$-to-$bO_u$, respectively, were generally within the experimental error. The correlation between the different measures of $\Delta\Delta G_u$ for the 12 mutants is shown in Figure 2.5, illustrating the close correspondence between the two methods. An ideal correlation would have an intercept at 0 kcal/mol and a slope of 1.0. For the fitted line, the intercept is (-0.01 ± 0.12) kcal/mol and the slope is 0.98 ± 0.07.

### 3.3. DISCUSSION

While the $bR_f$-to-$bR_u$ unfolding transition apparently reaches a steady state rapidly, RET hydrolysis and slow refolding precludes establishment of a true equilibrium under the conditions I have used. In contrast, the $bR_f$-to-$bO_u$ equilibrium is stable. Nevertheless, when I compare the effects of 12 of the 13 mutations on the two reactions, very similar results are obtained for most mutations. The origin of the fortuitous congruence of the two measures is unclear. One possibility is that mutations largely alter $k_u$, but not $k_f$. This could happen if the mutations have large effects on the free energy of the native state, but minimal effects on the transition state or unfolded state free energies. The effects of one mutant, P50A, were very different when measured by the two methods. It was found to be significantly stabilizing in the $bR_f$-to-$bR_u$ transition, but destabilizing in the $bR_f$-to-$bO_u$ transition. P50 is in the center of kinked helix in the folded bR structure, so it is perhaps not surprising that there might be long-range effects imparted by this mutation that differentially alter unfolding transition state.

Booth and co-workers have found excellent correspondence between equilibrium and kinetic measurements (17, 18). Their conditions are slightly different than the ones I have used, however, employing CHAPS instead of CHAPSO and the CHAPS is used at a higher concentration. I have observed that increasing CHAPSO concentrations from 6 to 16 mM increases refolding rates (unpublished result), which could then lead to equilibrium for the $bR_f$-to-$bR_u$ transition. Nevertheless, the $bR_f$-to-$bO_u$ transition is the more reliable way to measure the effects of mutations on thermodynamic stability.

### 3.4. MATERIALS AND METHODS

**Preparation of materials.** All bR variants were prepared as described (20, 28, 29). 1,2-dimyristoyl-sn-glycerol-3-phosphocholine (DMPC) was purchased from Avanti Polar Lipids. 3((3-cholamidopropyl)dimethylammonio)-2-hydroxy-1-propanesulphonate (CHAPSO) was purchased from Anatrace. Sodium dodecyl sulfate (SDS) and All-*trans* retinal (RET) were purchased from Sigma-Aldrich.

**$bR_f$-to-$bR_u$ unfolding and $bR_u$-to-$bR_f$ refolding assays.** The $bR_f$-to-$bR_u$ unfolding assays were performed as described by Faham S. *et al.* (20). Purple membrane was dissolved in a 2500 μL solution containing 15 mM DMPC, 6 mM CHAPSO and 10 mM sodium phosphate [pH 6.3]. The final concentration of the bR protein was in a range of 3 – 7 μM. After equilibration in dark for 1 hour, the solution was titrated by using an SDS titrant at room temperature in a 1-cm cu-vette stirred with a magnetic stir bar. The SDS titrant contained 20 % (w/v) SDS, 15 mM DMPC, 6 mM CHAPSO and 10 mM sodium phosphate [pH 6.3]. 10 μL of the SDS titrant was added every 3 min during the titration.

To test the effect of RET hydrolysis on the time evolution of the $bR_f$-to-$bR_u$ unfolding curve, a series of 200 μL solutions containing 3.7 μM wild-type bR, 15 mM DMPC, 6 mM CHAPSO, 10 mM sodium phosphate [pH 6.3] and different amounts of SDS varying from 5 to 71 mM were made at the same time and then loaded on a 96-well micro-plate (Thermo Scientific). The absorbance of each solution at 560 nm was monitored using a SpectraMax M5 plate reader (Molecular Devices).

To examine the $bR_u$-to-$bR_f$ refolding reactions, purple membrane was dissolved to make a final solution containing 37 μM bR, 15 mM DMPC, 6 mM CHAPSO, 10 mM sodium phosphate [pH 6.3] and 66.5 mM SDS. 3 min later, the absorbance at 560 nm disappeared and absorbance at 440 nm reached the maximal value. 20 μL of the unfolded bR solution was then mixed with a series of 180 μL solutions. The final solutions contained 3.7 μM wild-type bR, 15 mM DMPC, 6 mM CHAPSO, 10 mM sodium phosphate [pH 6.3] and different amounts of SDS varying from 5.2 to 66.5 mM.

**$bR_f$-to-$bO_u$ unfolding and $bO_u$-to-$bR_f$ refolding assays.** The $bR_f$-to-$bO_u$ unfolding and $bO_u$-to-$bR_f$ refolding assays were performed similarly to Chen & Gouaux (26). The main difference from their experiments was that I equilibrated each sample for much longer times.

For the unfolding assays, a stock solution was prepared by dissolving purple membrane in 20.625 mM DMPC, 22 mM CHAPSO and 13.75 mM sodium phosphate [pH 6.3]. For experiments with added RET, 15.4 μM all-*trans* RET was included in the stock solution. 160 μL aliquots of the stock solution were combined with 60 μL SDS solutions at various concentrations. The final solutions contained 1.5 - 3 μM bR, 15 mM DMPC, 16 mM CHAPSO, 10 mM sodium phosphate [pH 6.3], 0 or 11.2 μM all-*trans* RET and SDS varying from 8 to 138 mM.

For the refolding assays, bR was first unfolded in a buffer containing 15 mM DMPC, 16 mM CHAPSO, 10 mM sodium phosphate [pH 6.3] and 104 mM SDS. After equilibration for 1 h in dark, the absorbance at 560 nm disappeared and the absorbance at 390 nm reached the maximal value. Then, the unfolded bR solution was mixed with a series of DMPC/CHAPSO/SDS/RET/sodium phosphate [pH 6.3] solutions with varying SDS concentrations. The final solutions contained 1.5 - 3 μM bR, 15 mM DMPC, 16 mM CHAPSO, 10 mM sodium phosphate [pH 6.3], 0 or 11.2 μM all-trans RET and different amounts of SDS varying from 13 to 79 mM.

After the samples for unfolding and refolding assays were prepared, all the samples without added RET were equilibrated in dark at room temperature for 212 hours (~ 9 days) and those with added RET were equilibrated in dark at room temperature for 96 hours (~ 4 days). 200 μL of each unfolding and refolding sample was loaded on a 96-well UV-star micro-plate (Greiner Bio-One). Absorbance at 560 nm and fluorescence emission at 335 nm (excitation at 290 nm) of each solution were measured by SpectraMax M5 plate reader (Molecular Devices).

For unfolding experiments without added RET, the unfolding curves can be described by

$$\mathrm{Abs}_{560} = \mathrm{Abs}^{\mathrm{o}}_{560} \cdot F_{\mathrm{f}}\,, \qquad \text{(Eq. 3.1)}$$

where $\mathrm{Abs}^{\mathrm{o}}_{560}$ is the absorbance of the subdenaturant line and its extension over the experimental $X_{\mathrm{SDS}}$ range, which represents the theoretical absorbance if all the bR is folded and is assumed to be linearly dependent on $X_{\mathrm{SDS}}$:

$$\mathrm{Abs}^{\mathrm{o}}_{560} = a\,X_{\mathrm{SDS}} + b, \qquad \text{(Eq. 3.2)}$$

and $F_{\mathrm{f}}$ is the fraction of folded state in each bR variant, *i.e.* $F_{\mathrm{f}} = [\mathrm{bR_f}]/c$, where $c$ is the total concentration of each bR variant in the unit of μM. Since the unfolding free energy calculated for a standard state of 1 μM is defined as

$$\Delta G_U = - RT \ln([bO_u]\cdot[RET]/[bR_f]) = - RT \ln[c(1- F_f)^2/F_f],$$

where $RT$ is 0.592 kcal/mol, $F_f$ can be written in the expression of $\Delta G_U$ and $c$:

$$F_f = \frac{2 + \frac{\exp[-\Delta G_U /(RT)]}{c} - \sqrt{\frac{\exp[-2\Delta G_U /(RT)]}{c^2} + \frac{4\exp[-\Delta G_U /(RT)]}{c}}}{2}, \quad \text{(Eq. 3.3)}$$

If I assumed that $\Delta G_U$ has a linear relationship with $X_{SDS}$:

$$\Delta G_U = m\,(X_{SDS} - C_m) - RT \ln(0.5c),$$

Then Eq. 3.3 can be re-written as

$F_f =$

$$\frac{2 + \exp\frac{-[m\cdot(X_{SDS} - C_m) + \ln 2]}{RT} - \sqrt{\exp\frac{-2\cdot[m\cdot(X_{SDS} - C_m) + \ln 2]}{RT} + 4\cdot\exp\frac{-[m\cdot(X_{SDS} - C_m) + \ln 2]}{RT}}}{2}$$

(Eq. 3.4)

By plugging Eqs. 3.2 and 3.4 into Eq. 3.1, I derived the equation of $Abs_{560}$ in the expression of $X_{SDS}$ and used this equation to fit each unfolding curve without added RET:

$Abs_{560} = (a\,X_{SDS} + b)\cdot$

$$\frac{2 + \exp\frac{-[m\cdot(X_{SDS} - C_m) + \ln 2]}{RT} - \sqrt{\exp\frac{-2\cdot[m\cdot(X_{SDS} - C_m) + \ln 2]}{RT} + 4\cdot\exp\frac{-[m\cdot(X_{SDS} - C_m) + \ln 2]}{RT}}}{2}$$

(Eq. 3.5)

Parameters $a$, $b$, $m$ and $C_m$ were fit using Kaleighdagraph4.1. The unfolding free energies for wild-type and mutant were compared by subtracting $\Delta G_U^{WT}$ from $\Delta G_U^{mutant}$ at the midpoint of the wild-type transition, $X_{SDS} = C_m^{WT} = 0.572$.

For unfolding experiments with added RET, the RET concentration was held constant, so the RET concentration was combined with the equilibrium constant, giving the unfolding reac-

tion had a pseudo 1:1 stoichiometry, *i.e.* $\Delta G_U = -RT \ln([bO_u]/[bR_f]) = -RT \ln[(1-F_f)/F_f])$. $F_f$ can be written as

$$F_f = \frac{1}{1+\exp[-\Delta G_U/(RT)]}, \quad \text{(Eq. 3.6)}$$

If I assumed that $\Delta G_U$ has a linear relationship with $X_{SDS}$:

$$\Delta G_U = m\,(X_{SDS} - C_m),$$

Then $F_f$ can be re-written as

$$F_f = \frac{1}{1+\exp[-m\cdot(X_{SDS}-C_m)/(RT)]}. \quad \text{(Eq. 3.7)}$$

By plugging Eqs. 3.2 and 3.7 into Eq. 3.1, I derived the equation of $Abs_{560}$ in the expression of $X_{SDS}$ and used this equation to fit each unfolding curve with added RET:

$$Abs_{560} = \frac{a\cdot X_{SDS}+b}{1+\exp[-m\cdot(X_{SDS}-C_m)/(RT)]},$$

Parameters *a*, *b*, *m* and $C_m$ were fit using Kaleighdagraph4.1. The unfolding free energies for wild-type and mutant were compared by subtracting $\Delta G_U^{WT}$ from $\Delta G_U^{mutant}$ at the midpoint of the wild-type transition, $X_{SDS} = C_m^{WT} = 0.611$.

**Table 3.1.** $\Delta\Delta G_U$ of bR variants tested from the $bR_f$-to-$bO_u$ and $bR_f$-to-$bR_u$ reactions at certain $X_{SDS}$.

| Single Mutants | $\Delta\Delta G_U$[a] ($bR_f$-to-$bO_u$ without added RET) | $\Delta\Delta G_U$[a] ($bR_f$-to-$bO_u$ with 11.2 μM RET) | $\Delta\Delta G_U$[a] ($bR_f$-to-$bR_u$) | | |
|---|---|---|---|---|---|
| $X_{SDS}$ | 0.572[b] | 0.611[b] | 0.572 | 0.611 | 0.673[b] |
| K41A | - | -1.2 ± 0.3 | -1.4 ± 0.3 | -1.4 ± 0.3 | -1.4 ± 0.2 |
| F42A | -2.6 ± 0.5 | -2.2 ± 0.3 | -2.1 ± 0.3 | -2.1 ± 0.2 | -2.0 ± 0.2 |
| Y43A | -1.9 ± 0.3 | -2.1 ± 0.3 | -1.9 ± 0.4 | -1.9 ± 0.3 | -1.6 ± 0.2 |
| I45A | -2.2 ± 0.4 | -2.1 ± 0.2 | -2.1 ± 0.3 | -2.1 ± 0.3 | -1.9 ± 0.3 |
| T46A | - | -2.7 ± 0.4 | -2.3 ± 0.3 | -2.3 ± 0.3 | -2.2 ± 0.3 |
| T47A | -1.1 ± 0.3 | -0.9 ± 0.2 | -1.6 ± 0.3 | -1.4 ± 0.2 | -1.1 ± 0.2 |
| V49A | -0.3 ± 0.4 | -0.3 ± 0.3 | -0.7 ± 0.2 | -0.6 ± 0.2 | -0.3 ± 0.2 |
| I52A | -2.0 ± 0.3 | -1.6 ± 0.2 | -2.0 ± 0.2 | -1.9 ± 0.2 | -1.5 ± 0.1 |
| F54A | - | -0.9 ± 0.3 | -0.5 ± 0.2 | -0.4 ± 0.2 | -0.4 ± 0.1 |
| M56A | - | 1.6 ± 0.3 | 2.0 ± 0.2 | 1.9 ± 0.2 | 1.6 ± 0.1 |
| S59A | - | -0.1 ± 0.2 | -0.1 ± 0.2 | -0.1 ± 0.2 | -0.1± 0.1 |
| M60A | - | -1.0 ± 0.4 | -1.1 ± 0.3 | -1.1 ± 0.3 | -1.0 ± 0.2 |
| P50A | -1.1 ± 0.3 | -1.0 ± 0.2 | 0.7 ± 0.4 | 0.7 ± 0.3 | 0.6 ± 0.1 |

[a] Unit of $\Delta\Delta G_U$: kcal/mol.
[b] $C_m$ of the wild-type bR for each reaction.

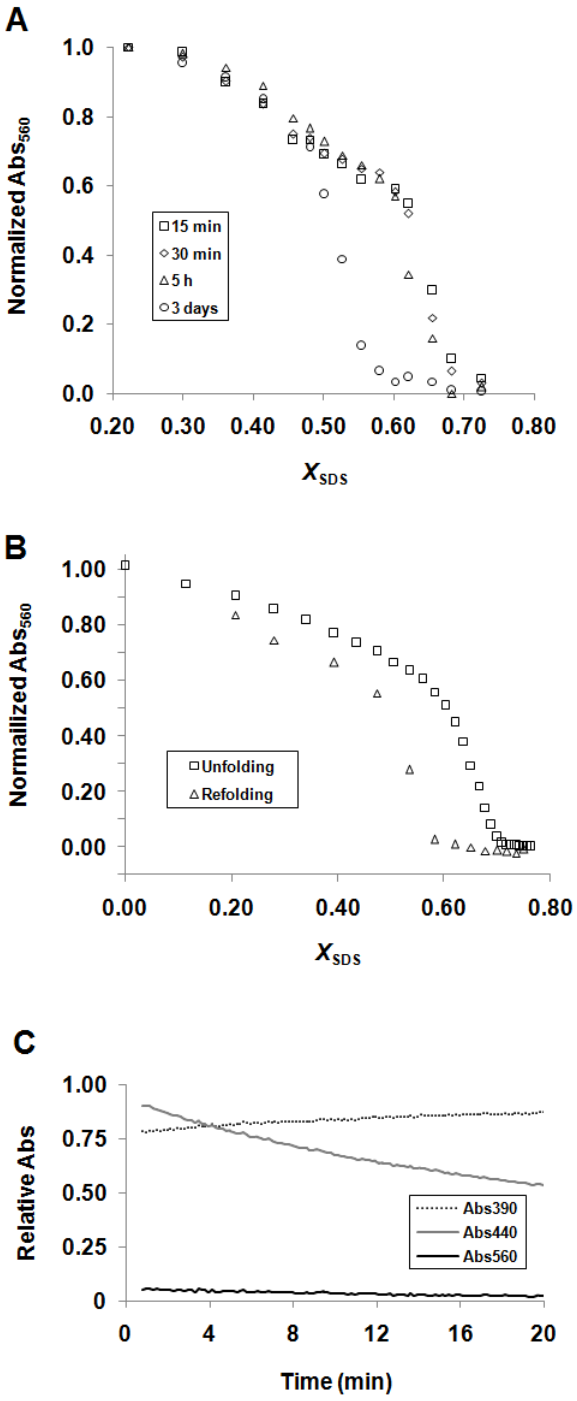


**Figure 3.1.** Effect of RET hydrolysis on the SDS-induced unfolding of wild-type bR. (*A*) Change of the wild-type $bR_f$-to-$bR_u$ unfolding curve in time. (*B*) Plot of normailized absorbance at 560 nm against the SDS mole fraction concentration, $X_{SDS}$, for the wild-type $bR_f$-to-$bR_u$ unfolding and $bR_u$-to-$bR_f$ refolding experiments. (*C*) Change of absorbances at 390, 440 and 560

nm in time when the wild-type protein was refolded from $bR_u$ at the apparent $C_m$ of the $bR_f$-to-$bR_u$ transition ($X_{SDS}$ = 0.673).

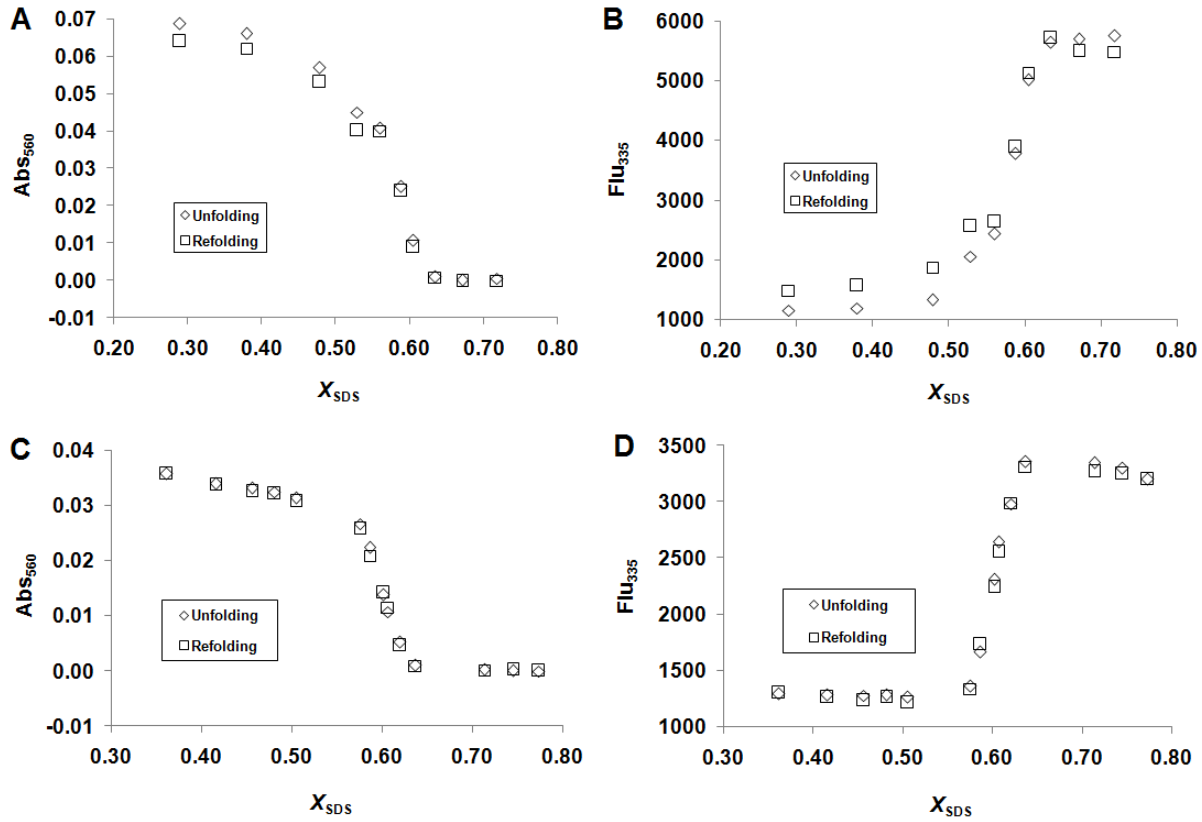


**Figure 3.2.** Equilibrium unfolding ($bR_f$-to-$bO_u$) and refolding ($bO_u$-to-$bR_f$) of wild-type protein. (*A*) The unfolding and refolding samples without added RET monitored by absrobance at 560 nm after incubated for ~ 9 days. (*B*) The unfolding and refolding samples without added RET monitored by fluorescence at 335 nm after incubated for ~ 9 days. (*C*) The unfolding and refolding samples with added RET monitored by absrobance at 560 nm after incubated for ~ 4 days. (*D*) The unfolding and refolding samples with added RET monitored by fluorescence at 335 nm after incubation for ~ 4 days.

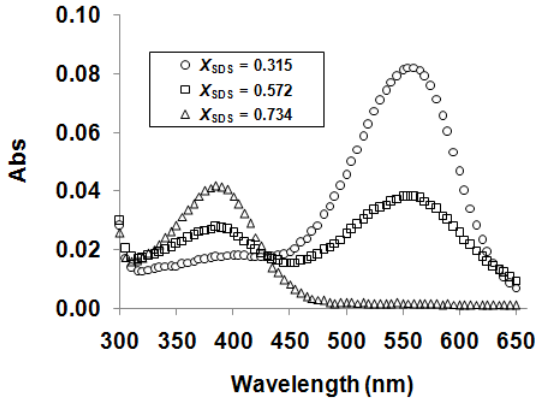


**Figure 3.3.** Absorbance spectra of the wild-type bR as a function of SDS concentration. The spectra were taken at $X_{SDS}$ = 0.315, 0.572 and 0.734, where the protein was completely folded, 50 % unfolded and completely unfolded, respectively.

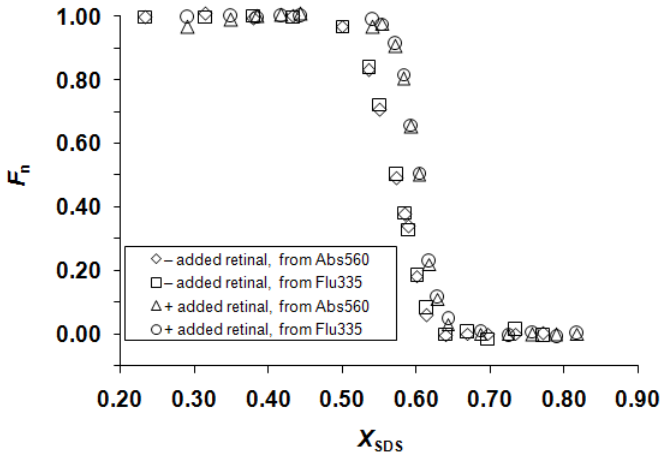


**Figure 3.4.** Wild-type $bR_f$-to-$bO_u$ unfolding curve. Plot of the fraction of the native-state protein, $F_n$, derived by monitoring absorbance and fluorescence for the reactions without and with added RET. Superposition of the curves is consistent with a two-state model for SDS-induced unfolding of $bR_f$ to $bO_u$. The transition shifts to higher SDS concentration in the presence of added RET, consistent with an equilibrium reaction that releases RET. Under both conditions with and without added RET, 3 μM of wild-type protein was included in the samples.

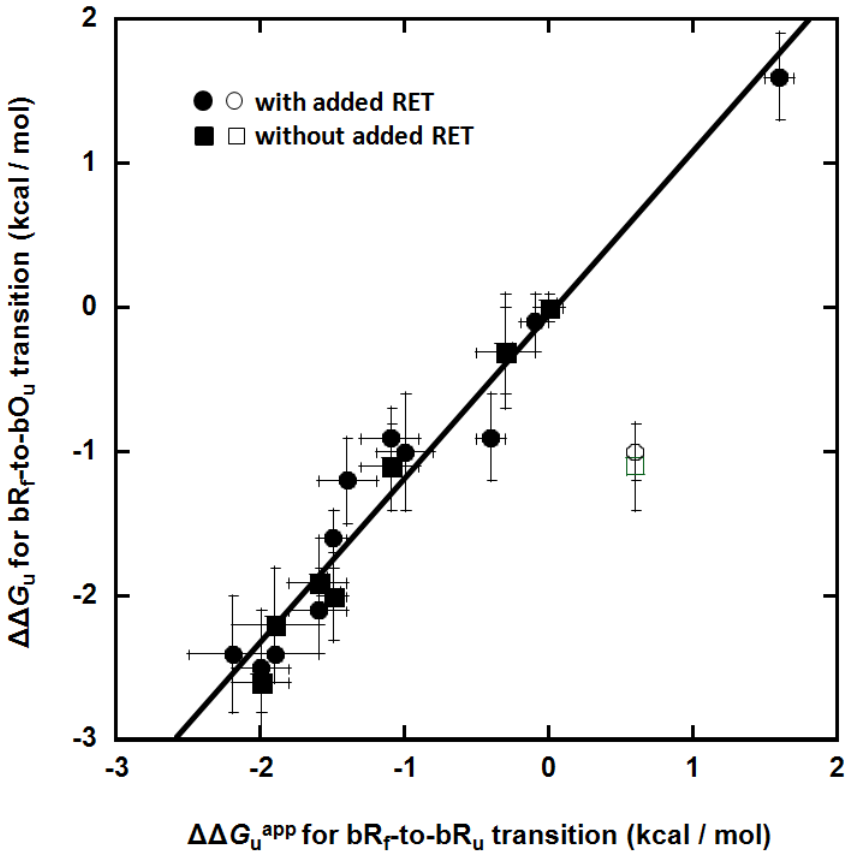


**Figure 3.5.** Correlation between $\Delta\Delta G_u^{app}$ values measured for the $bR_f$-to-$bR_u$ transition and the $\Delta\Delta G_u$ values measured for the $bR_f$-to-$bO_u$ transition. Square symbols refer to values measured without added RET and circle symbols refer to values measured with added RET. The open symbols refer to values for the P50A mutation, which is an outlier. The least squares fit line through all the filled symbols has a slope of 0.98 ± 0.07 and an intercept of - 0.01 ± 0.12 kcal/mol. The fitting correlation coefficient, *R*, is 0.96.

## 3.5. REFERENCES


1. Morin PE, Diggs D, Freire E (1990) Thermal stability of membrane-reconstituted yeast cytochrome c oxidase. *Biochemistry* 29:781–788.

2. Thompson LK, Miller AF, Buser CA, de Paula JC, Brudvig GW (1989) Characterization of the multiple forms of cytochrome b559 in photosystem II. *Biochemistry* 28:8048–8056.

3. Maneri LR, Low PS (1988) Structural stability of the erythrocyte anion transporter, band 3, in different lipid environments. A differential scanning calorimetric study. *J Biol Chem* 263:16170–16178.

4. Chen GQ, Gouaux E (1997) Reduction of membrane protein hydrophobicity by site-directed mutagenesis: introduction of multiple polar residues in helix D of bacteriorhodopsin. *Protein Eng* 10:1061–1066.

5. Kahn TW, Sturtevant JM, Engelman DM (1992) Thermodynamic measurements of the contributions of helix-connecting loops and of retinal to the stability of bacteriorhodopsin. *Biochemistry* 31:8829–8839.

6. Haltia T, Freire E (1995) Forces and factors that contribute to the structural stability of membrane proteins. *Biochim Biophys Acta* 1241:295–322.

7. Roman EA, Argüello JM, González Flecha FL (2010) Reversible unfolding of a thermophilic membrane protein in phospholipid/detergent mixed micelles. *J Mol Biol* 397:550–559.

8. Findlay HE, Rutherford NG, Henderson PJF, Booth PJ (2010) Unfolding free energy of a two-domain transmembrane sugar transport protein. *Proc Natl Acad Sci USA* 107:18451–18456.

9. Lau FW, Bowie JU (1997) A method for assessing the stability of a membrane protein. *Biochemistry* 36:5884–5892.

10. Renthal R (2006) An Unfolding Story of Helical Transmembrane Proteins. *Biochemistry* 45:14559–14566.

11. Huang KS, Bayley H, Liao MJ, London E, Khorana HG (1981) Refolding of an integral membrane protein. Denaturation, renaturation, and reconstitution of intact bacteriorhodopsin and two proteolytic fragments. *J Biol Chem* 256:3802–3809.

12. Jirgensons B (1982) Factors determining the reconstructive denaturation of proteins in sodium dodecyl sulfate solutions. Further circular dichroism studies on structural reorganization of seven proteins. *J Protein Chem* 1:71–84.

13. Jirgensons B, Capetillo S (1970) Effect of sodium dodecyl sulfate on circular dichroism of some nonhelical proteins. *Biochim Biophys Acta* 214:1–5.

14. Snel MM, Kaptein R, de Kruijff B (1991) Interaction of apocytochrome c and derived polypeptide fragments with sodium dodecyl sulfate micelles monitored by photochemically induced dynamic nuclear polarization 1H NMR and fluorescence spectroscopy. *Biochemistry* 30:3387–3395.

15. London E, Khorana HG (1982) Denaturation and renaturation of bacteriorhodopsin in detergents and lipid-detergent mixtures. *J Biol Chem* 257:7003–7011.

16. Marti T (1998) Refolding of bacteriorhodopsin from expressed polypeptide fragments. *J Biol Chem* 273:9312–9322.

17. Curnow P, Booth PJ (2007) Combined kinetic and thermodynamic analysis of alpha-helical membrane protein unfolding. *Proc Natl Acad Sci USA* 104:18970–18975.

18. Curnow P, Booth PJ (2009) The transition state for integral membrane protein folding. *Proc Natl Acad Sci USA* 106:773–778.

19. Booth PJ et al. (1995) Intermediates in the folding of the membrane protein bacteriorhodopsin. *Nat Struct Biol* 2:139–143.

20. Faham S et al. (2004) Side-chain contributions to membrane protein structure and stability. *J Mol Biol* 335:297–305.

21. Yohannan S et al. (2004) A C alpha-H...O hydrogen bond in a membrane protein is not stabilizing. *J Am Chem Soc* 126:2284–2285.

22. Yohannan S, Faham S, Yang D, Whitelegge JP, Bowie JU (2004) The evolution of transmembrane helix kinks and the structural diversity of G protein-coupled receptors. *Proc Natl Acad Sci U S A* 101:959–63.

23. Yohannan S et al. (2004) Proline substitutions are not easily accommodated in a membrane protein. *J Mol Biol* 341:1–6.

24. Joh NH et al. (2008) Modest stabilization by most hydrogen-bonded side-chain interactions in membrane proteins. *Nature* 453:1266–1270.

25. Joh NH, Oberai A, Yang D, Whitelegge JP, Bowie JU (2009) Similar energetic contributions of packing in the core of membrane and water-soluble proteins. *J Am Chem Soc* 131:10846–10847.

26. Chen GQ, Gouaux E (1999) Probing the folding and unfolding of wild-type and mutant forms of bacteriorhodopsin in micellar solutions: evaluation of reversible unfolding conditions. *Biochemistry* 38:15380–15387.

27. Cooper A, Dixon SF, Nutley MA, Robb JL (1987) Mechanism of retinal Schiff base formation and hydrolysis in relation to visual pigment photolysis and regeneration: resonance Raman spectroscopy of a tetrahedral carbinolamine intermediate and oxygen-18 labeling of retinal at the metarhodopsin stage in photoreceptor membranes. *Journal of the American Chemical Society* 109:7254–7263.

28. Cline SW, Doolittle WF (1987) Efficient transfection of the archaebacterium Halobacterium halobium. *J Bacteriol* 169:1341–1344.

29. Oesterhelt D, Stoeckenius W (1974) Isolation of the cell membrane of Halobacterium halobium and its fractionation into red and purple membrane. *Meth Enzymol* 31:667–678.

# CHAPTER 4

## AN ENERGETIC SCALE FOR EQUILIBRIUM H/D FRACTIONATION FACTORS ILLUMINATES HYDROGEN BOND FREE ENERGIES IN PROTEINS

### 4.1. INTRODUCTION

The role of hydrogen bonds in protein folding has been extensively studied, yet their contribution to stability remains controversial (1). The crux of the problem is the contribution of hydrogen bonding to solvent in the unfolded state relative to the folded state: are hydrogen bonds in the folded protein stronger than solvent hydrogen bonding in the unfolded state?

Model compound studies of backbone hydrogen bond mimics suggest that backbone hydrogen bonds are not stabilizing in water (2, 3). Chemical mutagenesis of the backbone suggested that the backbone hydrogen bonds can be stable (4, 5), but stability depends strongly on the context of the hydrogen bond (6). In other words, some backbone hydrogen bonds may be stabilizing and others are not. A similar picture emerges from mutagenesis of side-chain hydrogen bonds. Mutagenesis studies of hydrogen bonded side chains indicate that the interactions can be stabilizing (7–18). Indeed, a comprehensive analysis by Pace and co-workers argues that the net contribution of side-chain hydrogen bonds to folded proteins is on the order of 1 kcal/mol on average, but there is a wide variation in this net contribution (14). Undoubtedly part of the

variability arises from true differences in hydrogen bond strength while part is experimental error due to the complexity of combined mutagenesis and unfolding experiments.

Another experimental way to measure hydrogen bond strength, that is essentially non-perturbing, is by the use of equilibrium H/D fractionation factors. For a hydrogen bond between a weak acid, AH, and a weak base, B, the fractionation factor, $\phi$, is defined as the equilibrium constant for the following reaction (19–34):

$$AH{:}B + \tfrac{1}{2}\, D_2O \rightleftharpoons AD{:}B + \tfrac{1}{2}\, H_2O,$$

$$\phi = ([D]/[H])_{solute}/([D]/[H])_{water}.$$

For weak hydrogen bonds typically observed in proteins, the $\phi$-value increases as the strength of the hydrogen bond decreases (22–34). In other words, weaker hydrogen bonds accumulate more deuterium than stronger hydrogen bonds (22–34).

Loh and Markley were the first to employ NMR to evaluate the contribution of large numbers of $\phi$-values within a protein structure, in their investigation of staphylococcal nuclease H124L (25). Since then, similar approaches were used to measure $\phi$-values in enzyme:substrate complexes (29–32), in DNA base pairs (23), and in protein structures (25–28). In DNA base pairs, the $\phi$-values fall within a narrow range from 0.83 to 1.10 (23). In proteins, the $\phi$-values are more diverse, ranging from 0.28 to 1.42(25–33) (ignoring an extreme outlier with a $\phi$-value of 2.0 (32)).

Several trends have emerged from studies on proteins: [1] $\alpha$-helical hydrogen bonds are stronger on average than $\beta$-sheet hydrogen bonds (25–28). [2] Side-chain hydrogen bonds tend to be stronger than backbone hydrogen bonds (25, 26, 28). [3] Charge-stabilized hydrogen bonds are stronger than neutral hydrogen bonds (25, 26, 28,

34). In particular, the $\phi$-values for neutral hydrogen bonds measured in experiments fall between 0.52 and 1.42 (25–33), whereas charged-stabilized hydrogen bonds have been seen with a $\phi$-value as low as 0.28 (25). [4] Cooperative networks can strengthen hydrogen bonds (25, 26, 28, 34). [5] Equilibrium H/D fractionation factors in proteins have no correlation with H/D exchange rates (25). [6] Low-barrier (single-potential-well) hydrogen bonds to substrates in enzyme active sites can have $\phi$-values as low as 0.32 (29–32).

A major missing component of these earlier studies on protein hydrogen bonds was an energetic scale. In particular, how much stronger is a hydrogen bond with a $\phi$-value of 0.52 relative to one with a $\phi$-value of 1.42? Shi *et al*. attempted to answer this question with a model compound study (24). They measured two distinct equilibrium H/D fractionation factors and the corresponding free energies, $\Delta G^{HB}$. From these two points and the assumption that $\Delta G^{HB}$ is linear with $RT\ln\phi$, they obtained the scale factor, $SF = \partial(\Delta G^{HB})/\partial(RT\ln\phi)$, which relates any measured $\phi$ to $\Delta G^{HB}$. The scale factor obtained was - 74 ± 27 (24). Although the approach is important conceptually, we believe this scale factor cannot be correct. From the known $\phi$-values for neutral hydrogen bonds, this scale factor implies that they range over 60 kcal/mol in free energy. This is nearly an order of magnitude larger than the maximum expected enthalpy of a neutral hydrogen bond in a vacuum (35–37). Moreover, the linearity assumption is lacking experimental validation.

In this work we revisit and extend the idea of Shi *et al.* (24). By measuring free energies of hydrogen bonding and fractionation factors for a series of 18 weak acid and weak base pairs, we obtained a more extensive experimental delineation of the

relationship between $\phi$ and $\Delta G^{HB}$ values. We confirm the expected linear relationship (24) between $\Delta G^{HB}$ and $RT\ln\phi$, and obtain a scale factor of - 7.0 ± 0.7. The new scale factor is more consistent with what we know about hydrogen bonds. With a reliable $\phi$-value we can place prior work on an energetic scale. Moreover, we provide a new analysis of the overall contribution of hydrogen bonding to protein folding. We find that hydrogen bonding is slightly more stabilizing the unfolded state of proteins relative to the folded state. Our scale factor provides a simple, general and non-perturbing method for measuring hydrogen bond strengths in proteins and other systems.

## 4.2. RESULTS AND DISCUSSION

**Approach to measuring the scale factor.** To determine the relationship between $\Delta G^{HB}$ and $\phi$-values, we employed a series of weak acid base pairs listed in Fig. 4.1*A* and measured their $\phi$-values and the corresponding $\Delta G^{HB}$-values. To obtain $\phi$-values relative to water, we employed the three equilibria shown in Fig. 4.1*B*.

We first measured the fractionation factors ($\phi_1$) relative to the hydrogen-bond complex formed by triphenylmethanol ($Ph_3COH$) (Reaction 1). $Ph_3{}^{13}COH$ and the weak acid of interest were dissolved at a low concentration in the cognate base. Since the base acts as the solvent, the dissolved weak acids will be completely hydrogen bonded to the solvent base. The ratio of deuterium to protium in the solution was controlled by adding deuterated and protonated methanol in various ratios. The $\phi_1$-values of hydrogen-bonded weak acids were measured relative to the hydrogen-bonded $Ph_3{}^{13}COH$ by $^{13}$C-NMR as illustrated in *Supporting Information Appendix*: Fig. S4.1 *B* and *C*.

Once the $\phi_1$-values were obtained, they were converted to $\phi$-values relative to water. This was done by measuring the fractionation factors ($\phi_2$) of the hydrogen bonded $Ph_3COH$ dissolved in the corresponding base relative to solid $Ph_3COH$ (Reaction 2), followed by applying the fractionation factor ($\phi_3$= 1.10) for solid $Ph_3COH$ relative to water obtained from Kreevoy and Liang(20) (Reaction 3). As shown in Fig. 4.1*B*, the sum of the three reactions provides the desired $\phi$-value relative to water and is simply the product of three measured equilibrium constants. The $\phi_1$-values are listed in *Supporting Information Appendix*: Table S4.1. The $\phi_2$-values for triethylamine, *p*-dioxane and di(*n*-propyl) ketone, were measured to be 1.01 ± 0.02, 1.24 ± 0.07 and 1.14 ± 0.06, respectively. The final $\phi$-values of the hydrogen-bonded complexes with reference to water are also listed in *Supporting Information Appendix*: Table S4.1.

The free energy of hydrogen bonding was determined by measuring the dissociation constants of the various weak acid:base pairs in an aprotic solvent as illustrated in *Supporting Information Appendix*: Fig. S4.1*D*. We used solvents that closely matched the dielectric constant of the base that was employed to measure the $\phi_1$-values so that the hydrogen bond strengths would be matched in both solutions (Reaction 1). Dissociation curves were obtained by varying the concentrations of bases and monitoring chemical shifts of the proton involved in hydrogen bonding by $^1$H-NMR as an indication of the relative fractions associated and dissociated. The equilibrium dissociation constants, *K*, for the hydrogen-bond complexes and their $\Delta G^{HB}$ values are listed in *Supporting Information Appendix*: Table S4.1.

**The scale factor relating the $\Delta G^{HB}$ and $\phi$-values.** A plot of $\Delta G^{HB}$ against $RT\ln\phi$ from 18 hydrogen-bond complexes formed by our model weak acids and weak bases is shown in Fig. 4.1*C*. As expected (24), the points fall on a line with a linear fitting correlation coefficient of 0.93. The slope yields the scale factor of - 7.0 ± 0.7. The data used in the scale factor determination includes six different acids and three different bases, suggesting that the empirical relationship between $\Delta G^{HB}$ and $RT\ln\phi$ values we report here is robust and largely independent of the type of hydrogen donors and acceptors. The $\phi$-values of the hydrogen-bond complexes in this plot ranges from 0.65 to 1.54, which covers most of the $\phi$-values observed for protein hydrogen bonds previously (25–33). Some protein hydrogen bonds have a $\phi$-value lower than 0.65 (25–33), but given the linear relationship between $\Delta G^{HB}$ and $RT\ln\phi$ values those protein hydrogen bonds are within a short linear extrapolation of the plot in Fig. 4.1*C*. Thus, we can use the scale factor reported here to compare the strengths of any two intra-molecular hydrogen bonds in proteins as long as their $\phi$-values with reference to water are available.

**Hydrogen bond strength in proteins and enzymes.** The highest $\phi$-value measured in proteins is 1.42 for the backbone NH group of G55 in the unligated from of staphylococcal nuclease H124L. Although the amide of G55 appears to be hydrogen-bonded to the backbone carbonyl groups of E52, it must be an exceedingly weak interaction. We therefore set a $\phi$-value of 1.42 as the reference, *i.e.*, a free energy of 0. Using a reference $\phi$-value of 1.42 and our scale factor, we can obtain the relative free energies, $\Delta\Delta G^{HB}$, of hydrogen bonds.

For hydrogen bonds formed by backbone amide groups, the average values of $\Delta\Delta G^{HB}$ are 2.2 ± 1.2 kcal/mol in staphylococcal nuclease H124L (25), 1.9 ± 0.7 kcal/mol in histidine-containing proteins (26), 1.0 ± 0.2 kcal/mol in human ubiquitin (27) and 1.1 ± 0.3 kcal/mol in the two immunoglobulin G binding domains of protein G (28) (The uncertainties in free energies here and in the following are standard deviations). The backbone hydrogen bond strengths span 7.0, 3.5, 0.7 and 1.8 kcal/mol in the four proteins, respectively (Fig. 4.2). While a 7 kcal/mol range in staphylococcal nuclease H124L appears extreme, it is due to a backbone hydrogen bond to a charged side chain (see below). If we exclude backbone hydrogen bonds made to charged residues the overall range for neutral backbone hydrogen bonds decreases to 4.2 kcal/mol with reference to the weakest neutral backbone hydrogen bond (for L108 in the unbound staphylococcal nuclease H124L with the lowest neutral $\phi$-value of 0.52) (25). This value is well within the theoretical maximal enthalpy of a neutral hydrogen bond, 6.6 kcal/mol (35). Thus, the scale factor we obtain appears physically reasonable.

The $\phi$-value data for side chain hydrogen bonds is more sparse and biased toward residues involved in enzyme catalysis. To our knowledge the $\phi$-values of 20 side chains are known (excluding one extreme outlier which has an abnormally high $\phi$-value of 2.0) (26, 28–33). They include Ser, Tyr, Cys and His residues involved in side-chain:backbone, side-chain:side-chain or side-chain:ligand hydrogen bonds. As shown in Fig. 4.2, we find that the $\Delta\Delta G^{HB}$ for the side-chain hydrogen bonds have an average value of 3.9 ± 1.6 kcal/mol and span a range of 5.7 kcal/mol. Among the 20 side-chain hydrogen bonds, 18 are charged stabilized, which is likely to be an important factor in their strength (see below), and 16 are located at active sites of enzymes.

Of the 16 active site hydrogen bonds with known $\phi$-values, 13 form in intermediate states during catalytic cycles and are thought to stabilize those intermediates (29–32). We can now quantify the degree of stabilization. For example, His64 $H_{\delta 1}$ in subtilisin Carlsberg, which hydrogen bonds to the side chain of Asp32 (29), has a $\phi$-value of 1.2 in the ground state, while in the intermediate state its $\phi$-value decreases at least to 0.85. From our scale factor, this change corresponds to an improvement in hydrogen-bond free energy of over 1.5 kcal/mol. The His57 $H_{\delta 1}$ and $H_{\varepsilon 2}$ protons in bovine chymotrypsinogen A are involved in hydrogen bonds to the side changes of Asp102 and Ser195, respectively (32), and have $\phi$-values of 1.4 and 0.54 in the ground state, while in the intermediate state of the catalytic cycle their $\phi$-values change to 0.4 and 0.69, respectively. This corresponds to an overall improvement in hydrogen bond free energy of 4.2 kcal/mol, stabilizing the intermediate.

Charge-stabilized hydrogen bonds are expected to be stronger than neutral hydrogen bonds (25, 26, 28, 34) and this is revealed in the measured $\phi$-values. The average $\phi$-value for neutral hydrogen bonds at room temperature is $0.96 \pm 0.20$, and for hydrogen bonds with nominally charged acceptors, it is $0.60 \pm 0.29$(25–33). Based on our scale factor, the energetic difference between these two kinds of hydrogen bonds is 1.9 kcal/mol on average.

As noted previously, $\alpha$-helical backbone NH groups have an average $\phi$-value 6 - 14 % less than $\beta$-sheet backbone NH groups in the same proteins(25–28) at room temperature (see Fig. 4.3*A*). This corresponds to a difference of 0.3 - 0.6 kcal/mol higher stability in $\alpha$-helices relative to $\beta$-sheets on average (Fig. 4.3*B*). Our experimental value corresponds well with the results from theoretical calculations indicating that the

enthalpy cost to rupture a backbone N-H···O=C hydrogen bond is 0.35 kcal/mol higher in $\alpha$-helices than in $\beta$-hairpins in $H_2O$ (38).

**Contribution of hydrogen bonding to protein unfolding free energy.** Our scale factor allows us to determine the relative energetic contribution of hydrogen bonds to the folded state of a protein, but what is their contribution to stabilizing the folded state relative to the unfolded state? To answer this question we would require a complete picture of the change in hydrogen bonding upon unfolding between the protein and the solvent and also the perturbation of the solvent caused by the protein hydrogen bonds. While addressing this question for individual hydrogen bonds is not straightforward, we believe we can now answer this question for the overall net contribution of all hydrogen bonds to folding.

We consider the $\phi$-value of all the hydrogen bonds in a complex composed of the protein and water molecules whose structure is closely affected by the unfolding and folding of the protein. As shown in Fig. 4.4*A*, we can imagine a protium / deuterium exchange reaction between protein:water complexes in the unfolded and folded states. The equilibrium constant for the reaction in Fig. 4.4*A* is simply the ratio of the $\phi$-values for these two states, $\phi^{\text{U-Prot}}/\phi^{\text{F-Prot}}$. This reaction is equivalent to the sum of the two reactions described in Fig. 4.4*B*, which represents the folding of the protonated protein in $H_2O$ and the unfolding of the deuterated protein in $D_2O$, respectively. Thus, the equilibrium constant of the reaction shown in Fig. 4.4*A* will be equal to the product of the equilibrium constants of the two reactions shown in Fig. 4.4*B*:

$$\phi^{\text{U-Prot}}/\phi^{\text{F-Prot}} = K_{\text{u}}^{\text{D-Prot:(D2O)n}}/K_{\text{u}}^{\text{H-Prot:(H2O)n}}.$$

Thus, if we know the unfolding equilibrium constants or unfolding free energy of a protein in the protonated form in $H_2O$ and in the deuterated form in $D_2O$ at room temperature, we will be able to calculate the contribution made by the changes of all the hydrogen bonding interactions to the protein unfolding free energy, by using the following equation:

$$\Delta\Delta G^{HB} = SF \cdot RT\ln(\phi^{U\text{-}Prot}/\phi^{F\text{-}Prot}) = SF \cdot RT\ln(K_u^{D\text{-}Prot:(D2O)n}/K_u^{H\text{-}Prot:(H2O)n}).$$

The unfolding free energies of three proteins, rat CD2 (39), ribonuclease (40) and monomeric wild-type lambda repressor ($\lambda^{6\text{-}85}$) (10) have been measured in both $H_2O$ and $D_2O$. The deuterated proteins in $D_2O$ are more stable than the protonated proteins in $H_2O$ by 1.2 kcal/mol for rat CD2, 1.0 kcal/mol for ribonuclease and for 0.55 kcal/mol for $\lambda^{6\text{-}85}$. This means the $RT\ln(K_u^{D\text{-}Prot:(D2O)n}/K_u^{H\text{-}Prot:(H2O)n})$ and thus the $RT\ln(\phi^{U\text{-}Prot}/\phi^{F\text{-}Prot})$ values for these three proteins at room temperature are simply - 1.2, - 1.0 and - 0.55 kcal/mol, respectively. In other words, the $\phi$-value for the unfolded protein is lower than the folded protein, implying stronger hydrogen bonding in the unfolded state. By applying the scale factor we obtain the change in free energy due to the hydrogen bonding interactions upon unfolding. For rat CD2, ribonuclease and $\lambda^{6\text{-}85}$, the hydrogen bonds favor the unfolded state by 8.3, 7.0 and 3.8 kcal/mol overall and by 0.09, 0.06 and 0.05 kcal/mol per residue, respectively, which are small compared to other interactions that drive protein folding such as van der waals forces, electrostatic interactions and hydrophobic interactions (41). This result suggests that the unfolding of all the three proteins is marginally favored by changes in the hydrogen bonding interactions, which contrasts with the analysis of Pace and co-workers based on side chain mutagenesis experiments (14). We believe this contradiction can be easily reconciled if side chain hydrogen bonds are generally more

stable than the average backbone hydrogen bond. For the hydrogen bonds with known $\phi$-values, this seems to be true. Moreover, many side chain hydrogen bonds are charge stabilized and most are of the O-H···O type which may be stronger than N-H···O hydrogen bonds(42) because, we think, the O-H···O hydrogen bonds carry more partial charges than N-H···O hydrogen bonds. Thus, an analysis based on side chain hydrogen bonding may overestimate their overall contribution(28).

The strong hydrogen bond free energies in proteins that we report suggests that hydrogen bonds can play an important role in specifying the structure of a protein. They will also play an important role in defining protein movements. For example, many strong side-chain hydrogen bonds were found at active sites in intermediate conformations of catalytic enzymes (29–32). We find that the hydrogen bonds can stabilize the intermediate states by 1.5 to 4.2 kcal/mol more than they stabilize the ground states, thereby facilitating the formation of intermediate states of enzymes during the catalytic pathway. It seems likely that proteins can define modes of motion by modulating hydrogen bond strengths. Thus the wide range of backbone hydrogen bond strengths may be an evolved feature of proteins that deserves further scrutiny. Our scale factor now provides way to quantitatively measure these contributions.

## 4.3. MATERIALS AND METHODS

**Preparation of materials.** *p*-(Trifluomethyl)phenol (*p*-$CF_3C_6H_4OH$), phenol ($C_6H_5OH$), 2,2,2-trifluoroethanol ($CF_3CH_2OH$), triphenylmethanol ($Ph_3COH$), $^{13}C$(OH)-labeled triphenylmethanol ($Ph_3{}^{13}COH$), pyrrole ($C_4H_4NH$), *p*-dioxane ($C_4H_8O_2$) and di(*n*-propyl)

ketone (($n$-$C_3H_7$)$_2$C=O) were purchased from Sigma-Aldrich. Methanol-O-$d$ ($CH_3OD$), cyclohexane-$d_{12}$ ($C_6D_{12}$), toluene-$d_8$ ($C_7D_8$), $o$-dichlorobenzene-$d_4$ ($C_4D_4Cl_2$) and chloroform-$d$ ($CDCl_3$) were purchased from Cambridge Isotope Laboratories, Inc.. Among these materials, the chemicals in solid state, $p$-(trifluomethyl)phenol and phenol, were dried in a vacuum oven at room temperature over night before use and triphenylmethanol and $^{13}$C(OH)-labeled triphenylmethanol were used directly because they were already dry enough. All chemicals in liquid state except cyclohexane-$d_{12}$, because it was already dry enough, were dried by mixing with 4 Å molecular sieves purchased from Sigmal-Aldrich for 6 - 18 hr before use. Triethylamine ($Et_3N$) and methanol ($CH_3OH$), which were stored under dry argon and were pre-dried by using a home-made aluminum column, were a gift from Neil Garg Lab at UCLA.

Triphenylmethanol-O-$d$ ($Ph_3COD$) was made by mixing ~ 2 g triphenylmethanol ($Ph_3COH$), 1 mL acetonitrile ($CH_3CN$) and 2 mL deuterium oxide ($D_2O$) for 2 days, followed by drying the mixture in a vacuum oven at room temperature for one day. The deuteration level in triphenylmethanol-O-$d$ was verified by dissolving the dry powder of triphenylmethanol-O-$d$ in dry chloroform-$d$ at a concentration of more than 0.5 M and recording the $^1$H-NMR spectrum in a BRUKER AV300 spectrometer. Almost no signal for the hydroxyl proton was detected compared to the peak for aromatic protons, suggesting nearly full deuteration at the hydroxyl group. Acetonitrile and deuterium oxide were purchased from Fisher Scientific and Cambridge Isotope Laboratories, Inc., respectively.

**Measurement of $\phi$-values for hydrogen-bond complexes formed between weak acids and weak bases**. In aqueous solution, weak acids and weak bases cannot form hydrogen bonds with each other, but instead are hydrogen-bonding to water and are partially neutralized by one another. Therefore, we did not measure the $\phi$-values of hydrogen-bond complexes formed between each weak acid and weak base (see Fig. 4.1*A*) directly in water, but we measured them at room temperature in an indirect way by converting the $\phi$-value to a water reference in three steps (see Fig. 4.1*B*), where $\phi_1$, $\phi_2$, and $\phi_3$ are the equilibrium constants for the three steps, respectively.

The $\phi_1$-values were measured using a method similar to Jarret and Saunders (19). In our experiment, 0.10 M *p*-(trifluomethyl)phenol (*p*-$CF_3C_6H_4OH$), 0.10 M phenol ($C_6H_5OH$), 0.10 M 2,2,2-trifluoroethanol ($CF_3CH_2OH$), 0.60 M methanol-OH/OD ($CH_3OH$/OD), 0.005M $^{13}$C-labeled $Ph_3{}^{13}COH$ and 0.10 M pyrrole ($C_4H_4NH$) were added together to solvent which was composed of the pure base, triethylamine ($Et_3N$), *p*-dioxane ($C_4H_8O_2$) or di(*n*-propyl) ketone ((*n*-$C_3H_7)_2C$=O), and a small amount of cyclohexane-$d_{12}$ ($C_6D_{12}$) at volume ratios of 9 / 1, 9 / 1, and 5 / 1, respectively. Except for the $^{13}$C-labeled $Ph_3{}^{13}COH$, the natural abundance of $^{13}$C was utilized for the weak acids. In each of the three solutions, all the weak acids formed hydrogen bonds with the same base. The small amount of cyclohexane-$d_{12}$ ($\varepsilon$ = 2.0) (43) was used to lock the magnetic field and keep the dielectric constants of the mixtures (44) with each base, triethylamine ($\varepsilon$ = 2.3) (43), *p*-dioxane ($\varepsilon$ = 2.3) (43) and di(*n*-propyl) ketone ($\varepsilon$ = 12.5) (43), almost the same as the dielectric constants of toluene ($\varepsilon$ = 2.3) (43), toluene ($\varepsilon$ = 2.3) (43) and*o*-dichlorobenzene ($\varepsilon$ = 9.9) (43), respectively, because we measured the strengths of the hydrogen bonds in these solvents. The ratio of the XH / XD group (X = O or N) for each

acid was varied in a series samples by changing the ratio of OH/OD in the added methanol-OH/OD.

By monitoring the signal for the carbon atoms directly linked to the XH / XD group (X = O or N), *i.e.* the $\alpha$-C, from the $^{13}$C-NMR spectra, we measured the $\phi_1$-values of hydrogen bond complexes formed between weak acids and each of the three bases with reference to the $\phi$-value of the hydrogen bond complex formed between the $^{13}$C-labeled $Ph_3{}^{13}COH$ and the same base at exchange equilibrium (~ 3 hr after mixing).

As illustrated in *Supporting Information Appendix*: Fig. S4.1*A*, two types of $^{13}$C-NMR spectra for $\alpha$-C atoms were observed depending on the rate of H/D exchange. In the case of slow exchange, two separate peaks for $\alpha$-C were seen, each having the same chemical shift as the pure C-XH or C-XD (X = O or N) state. Thus, for slow exchange we obtained the XH / XD ratio from the corresponding peak areas, *A*. The hydrogen bond complexes for $^{13}$C-labeled $Ph_3{}^{13}COH$ with all the three bases followed the slow-exchange regime. For the hydrogen bond complexes formed by other acids and the same base which also follow this slow-exchange regime, we plotted the ratios of *A*(AD:B) / *A*(AH:B) against the ratios of *A*($Ph_3{}^{13}COD$:B) / *A*($Ph_3{}^{13}COH$:B) for the solutions with different total amounts of protons and deuterons. Since

$$\phi_1 = \frac{A(\mathrm{AD{:}B}) \,/\, A(\mathrm{AH{:}B})}{A(\mathrm{Ph_3{}^{13}COD{:}B}) \,/\, A(\mathrm{Ph_3{}^{13}COH{:}B})}, \qquad \text{(Eq. 4.1)}$$

their $\phi_1$-values were simply read from the slope of the least-square linear-line fitting of this plot, forcing the line to pass the origin (see *Supporting Information Appendix*: Fig. S4.1*B*). In the case of fast exchange, a merged peak for the $\alpha$-C was seen, whose chemical shift, $\delta$, is dependent on fraction of deuteron or proton within the hydrogen bond complex according to

$$[\mathrm{AD{:}B}] / [\mathrm{AH{:}B}] = [\delta(\mathrm{AH{:}B}) - \delta(\mathrm{AL{:}B})] / [\delta(\mathrm{AL{:}B}) - \delta(\mathrm{AD{:}B})], \quad \text{(Eq. 4.2)}$$

where L represents the mixed state of labile protons and deuteron. The chemical shift of the pure protonated state, $\delta$(AH:B), can be measured by adding the protonated $Ch_3OH$ only, but the chemical shift $[\mathrm{AD{:}B}] / [\mathrm{AH{:}B}] = \frac{1}{\frac{\delta(\mathrm{AH{:}B}) - \delta(\mathrm{AD{:}B})}{\delta(\mathrm{AH{:}B}) - \delta(\mathrm{AL{:}B})} - 1}$. of deuterated state, $\delta$(AD:B), cannot be determined directly. Consequently, their $\phi_1$-values were obtained in a different way. We first convert Eq. 4.2 into the following form,

(Eq. 4.3)

If we define $\Delta = \delta(\mathrm{AH{:}B}) - \delta(\mathrm{AD{:}B})$ in Eq. 4.3 and then plug Eq. 4.3 into Eq. 4.1, we obtain

$$\phi_1 = \frac{A(\mathrm{Ph_3{}^{13}COH{:}B}) / A(\mathrm{Ph_3{}^{13}COD{:}B})}{\frac{\Delta}{\delta(\mathrm{AH{:}B}) - \delta(\mathrm{AL{:}B})} - 1}. \quad \text{(Eq. 4.4)}$$

Eq. 4.4 can be further re-written as

$$\frac{1}{\delta(\mathrm{AH{:}B}) - \delta(\mathrm{AL{:}B})} = \frac{1}{\phi_1 \cdot \Delta} \cdot \left(\frac{A(\mathrm{Ph_3{}^{13}COH{:}B})}{A(\mathrm{Ph_3{}^{13}COD{:}B})} + \phi_1\right).$$

Thus, for the hydrogen bond complexes in the fast-exchange regime, their $\frac{1}{\delta(\mathrm{AH{:}B}) - \delta(\mathrm{AL{:}B})}$ values were plotted against the ratios of $A(Ph_3{}^{13}COH{:}B)$ / $A(Ph_3{}^{13}COD{:}B)$. The $\phi_1$-value is the absolute value of the *x*-intercept from the linear least-square fitting of this plot (see *Supporting Information Appendix*: Fig. S4.1*C*). The uncertainty of the $\phi_1$-value for each hydrogen bond complex is the sum of the standard deviation from three experiments and the average error from the data fitting for each experiment, *i.e.* STD + ($d_{exp1}$ + $d_{exp2}$ + $d_{exp3}$) / 3, where STD stands for the standard deviations from the three experiments and $d_{exp1}$, $d_{exp2}$ and $d_{exp3}$ are the errors in the $\phi_1$-value from fitting for the three experiments, respectively.

We measured $\phi_2$-values for the hydrogen bond complex formed between $Ph_3COH$ and each base as follows. 0.30 g $Ph_3COH$ and 0.30 g $Ph_3COD$ were added to 1.6 mL of the solvents triethylamine / cyclohexane-$d_{12}$ ($V/V$ = 9/1), *p*-dioxane / cyclohexane-$d_{12}$ ($V/V$ = 9/1) and di(*n*-propyl) ketone / cyclohexane-$d_{12}$ ($V/V$ = 5/1), respectively, and the mixtures were stirred in dark at room temperature for 2 - 4 days to allow for equilibrium to be established. The ratio of [$Ph_3COD$:B] / [$Ph_3COH$:B] in each solution, *i.e.* the supernatant of each solid/liquid mixture containing ~ 0.3 M $Ph_3COH/D$, were determined by measuring the ratio between the peak areas of $^{13}$C-OD and $^{13}$C-OH from the $^{13}$C-NMR spectrum by utilizing its natural $^{13}$C abundance. To determine the $Ph_3COH/D$ in the solid state, each solid-liquid mixture was filtered under vacuum and the solid was dried for 1 hr under vacuum at room temperature. The dried samples were then dissolved in deuterated chloroform to a concentration of ~ 0.5 - 0.7 M and the ratio of [$Ph_3COD$] / [$Ph_3COH$] was determined from the $^{13}$C-NMR spectrum. The uncertainty in $\phi_2$-value for each solid-liquid mixture is the standard deviation from three experiments.

The $\phi_3$-value was previously determined to be 1.10 ± 0.06 by Kreevoy and Liang(20).

By multiplying $\phi_1$-, $\phi_2$- and $\phi_3$-values together, we obtain the $\phi$-value of each hydrogen bond complex with reference to water (see Fig. 4.1*B*). The standard deviation of the final $\phi$-value for each hydrogen bond complex, $\delta\phi$, was derived from the uncertainties of $\phi_1$- and $\phi_2$-values according to the following equation,

$$\delta\phi/\phi = \delta\phi_1/\phi_1 + \delta\phi_2/\phi_2,$$

where $\delta\phi_1$ and $\delta\phi_2$ are the uncertainties in $\phi_1$- and $\phi_2$-values, respectively. As the uncertainty of $\phi_3$-values will be applied to each hydrogen bond complex in the same way,

it will have no effect on the relative uncertainties between any two hydrogen bonds, so it was not considered for determining the uncertainty in $\phi$-value. The uncertainty in $RT\ln\phi$ values was determined as $RT(\delta\phi/\phi)$.

All the $^{13}$C-NMR measurements were performed on a BRUKER AV500 spectrometer equipped with a cyroprobe optimized for $^{13}$C sensitivity.

**Measurement of strengths of hydrogen-bond complexes formed between weak acids and weak bases.** The strength of the hydrogen bond, $\Delta G^{HB}$, formed by pairs of weak acids (AH) and weak bases (B) was measured as the dissociation free energy of the hydrogen-bond complex (AH:B) in apolar, aprotic organic solvents. The weak acids we used were *p*-(trifluomethyl)phenol, phenol, 2,2,2-trifluoroethanol, methanol-OH, triphenylmethanol and pyrrole. The weak bases we used were triethylamine, *p*-dioxane and di(*n*-propyl) ketone. For hydrogen bonds formed between each of the weak acids and triethylamine or *p*-dioxane, a small amount of the weak acid was dissolved in toluene-$d_8$ to make the concentration of the acid of $c_o$ = 0.0300 M and then the solution was divided into seven 600 uL aliquots. Next, various amounts of triethylamine or *p*-dioxane were added to the seven samples to make the final concentrations of the bases ranging from 0 to 0.5 M and 0 to 1.0 M, respectively. For hydrogen bonds formed between each of the weak acids and di(*n*-propyl) ketone, a small amount of the weak acid was dissolved in *o*-dichlorobenzene-$d_4$ to make the concentration of the acid of $c_o$ = 0.0300 M and then the solution was divided into seven 600 uL aliquots. Next, di(*n*-propyl) ketone at different volumes were added to the seven samples to make the final concentrations of the base ranging from 0 to 1.0 M.

$^1$H-NMR spectra were acquired for each sample at room temperature soon after they were prepared using BRUKER spectrometers AV300, DRX500, AV500 or AV600. Since the exchange between labile protons from the hydrogen-bonded complex and from the free acid is very fast, a $AH{:}B \rightleftharpoons AH + B$ merged peak representing the mixed state was always observed for the labile proton. The chemical shift of the labile proton was plotted against the volume of added base to obtain a binding isotherm. The equilibrium dissociation constant, $K$, of the reaction was determined by least-square fitting as described by Fielding (45) (see *Supporting Information Appendix*: Fig. S4.1*D*). A standard state of 1 M was used for $K$. The uncertainty of the $K$-value for each hydrogen bond complex was determined by comparing the discrepancy between two experiments. $\Delta G^{HB}$ value for each hydrogen bond complex was determined as $-RT\ln(NK)$, where $R$ is the gas constant, $N$ is the number of potential hydrogen acceptors per molecule of base, which is 2, 4 and 1 for di(*n*-propyl) ketone, *p*-dioxane and triethylamine, respectively. The reason the number $N$ is included in the conversion of $K$ into $\Delta G^{HB}$ is that the concentration of hydrogen acceptors from the bases is actually $N$ times higher than the concentration of each base. The uncertainty in $\Delta G^{HB}$ was determined as $RT(\delta K/K)$, where $\delta K$ is the uncertainty in $K$.

**Determination of the scale factor**. The strength of each hydrogen bond formed between the model weak acids and weak bases expressed in the free energy change of breaking the hydrogen bond, $\Delta G^{HB}$, was plotted against the value of $RT\ln\phi$ of each at room temperature, where $R$ is the ideal gas constant. A least-square fitting was made from the plot to determine the scale factor, which is simply the slope (see Fig. 4.1*C*).

**Applying the new scale factor to study the difference in strength of protein hydrogen bonds.** The $\phi$-values for hydrogen bonds in several proteins have been measured (25–33), but those experiments were performed at different temperatures ranging from room temperature to 45 $^{o}$C. Thus, we need to convert their results into $\phi$-values at room temperature before we apply the scale factor because our scale factor is determined at room temperature. As pointed out by Kreevoy and Liang (20), the $\phi$-value of a hydrogen bond complex with reference to water is dependent on the zero-point energies, *i.e.* the lowest (or ground-state) eigenvalues of energy for the vibrational potential functions, of hydrogen bonds formed between the weak acid and weak base and formed between water molecules in the protonated and deuterated forms. Their relationship can be written as the following form (20),

$$RT\ln\phi = hc[(\mathrm{ZPE(AH{:}B)} - \mathrm{ZPE(AD{:}B)} - \mathrm{ZPE(H_2O)} + \mathrm{ZPE(D_2O)}], \quad \text{(Eq. 4.7)}$$

where $h$ is the Plank constant, $c$ is the speed of light in vacuum, and ZPE stands for zero-point energy in the unit of $cm^{-1}$. Since ZPE is not dependent on temperature especially when the change in temperature is very small, the right-hand side of Eq. 4.7, and therefore the $RT\ln\phi$ value, can be considered as a constant for each hydrogen bond complex. Thus, the previously measured $\phi$-values for hydrogen bonds at the experimental temperature, $T_o$, in the unit of Kelvin can be converted into the $\phi$-values at room temperature as $\exp\{[T_o/(298\ \mathrm{K})] \cdot \ln\phi_{To}\}$. Thus, by applying the scale factor we report here to the $\phi$-values of exchangeable hydrogen atoms in proteins at room temperature, we can calculate the difference between the free energies of any two protein hydrogen bonds.

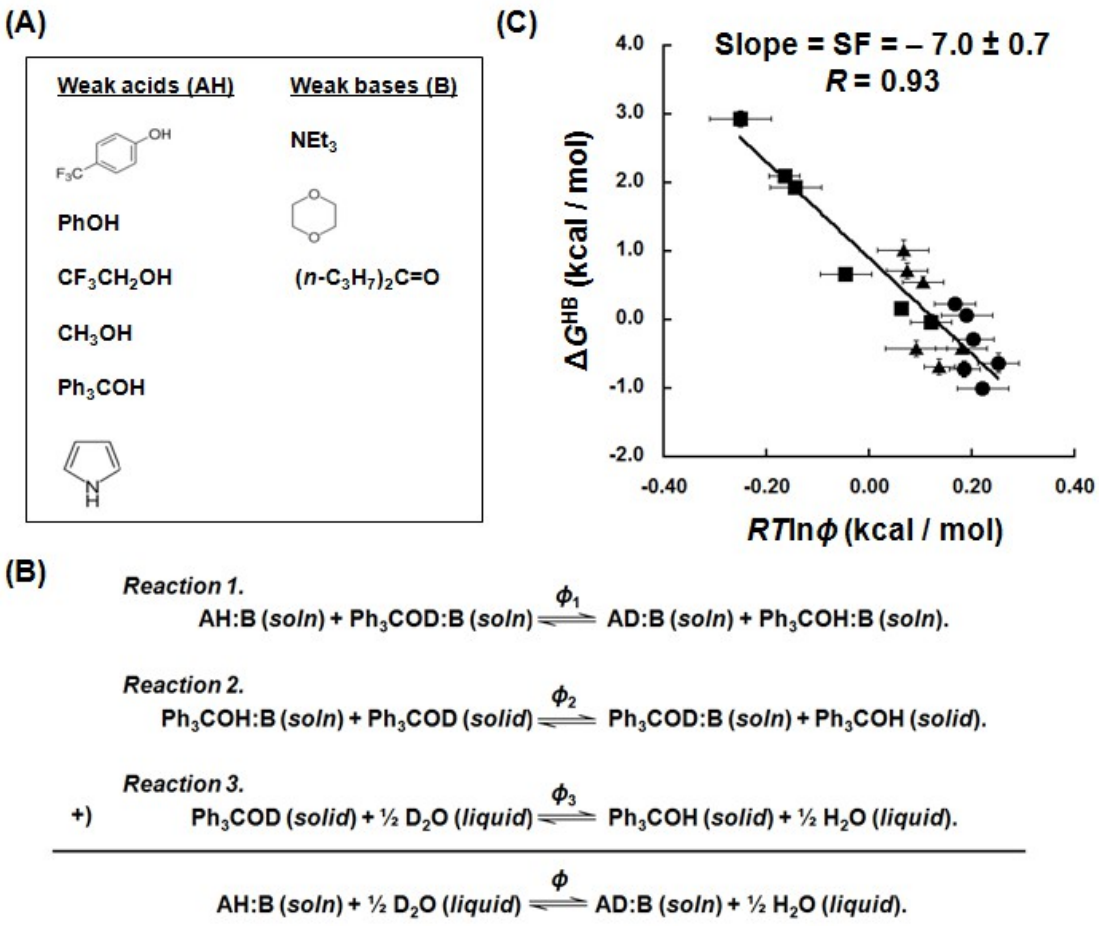


**Fig. 4.1.** Determination of the scale factor. (*A*) The 6 weak acids and 3 weak bases used to form model hydrogen-bonded complexes. (*B*) The three equilibria used to determine the $\phi$-value relative to water. (*C*) Plot of $\Delta G^{HB}$ against $RT\ln\phi$ values at room temperature from 18 model hydrogen-bond complexes to determine the scale factor. The hydrogen-bond complexes formed by triethylamine in toluene, *p*-dioxane in toluene and di(*n*-propyl) ketone in *o*-dichlorobenzene were labeled using square, round and triangle symbols, respectively.

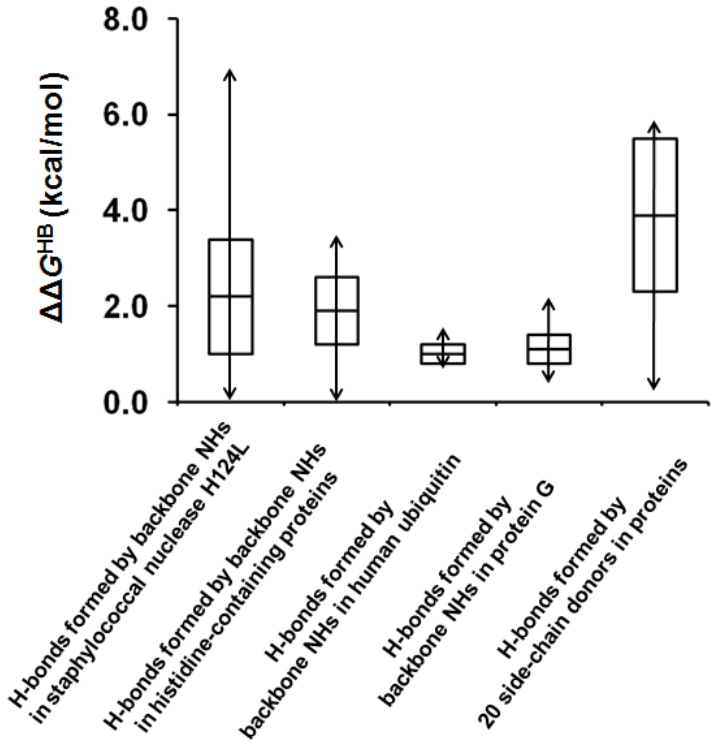


**Fig. 4.2.** Distribution of protein hydrogen bond strengths. The average $\Delta\Delta G^{HB}$ values, for backbone NH groups and side-chain donors are shown as a bar in the middle of each box. The standard deviations, determined by considering both the uncertainty in the scale factor and the standard deviations of the $\phi$-values, are indicated as the distance between the middle bar and the upper or lower edge of each box. The ranges of $\Delta\Delta G^{HB}$ are indicated by the double-headed arrows across each box.

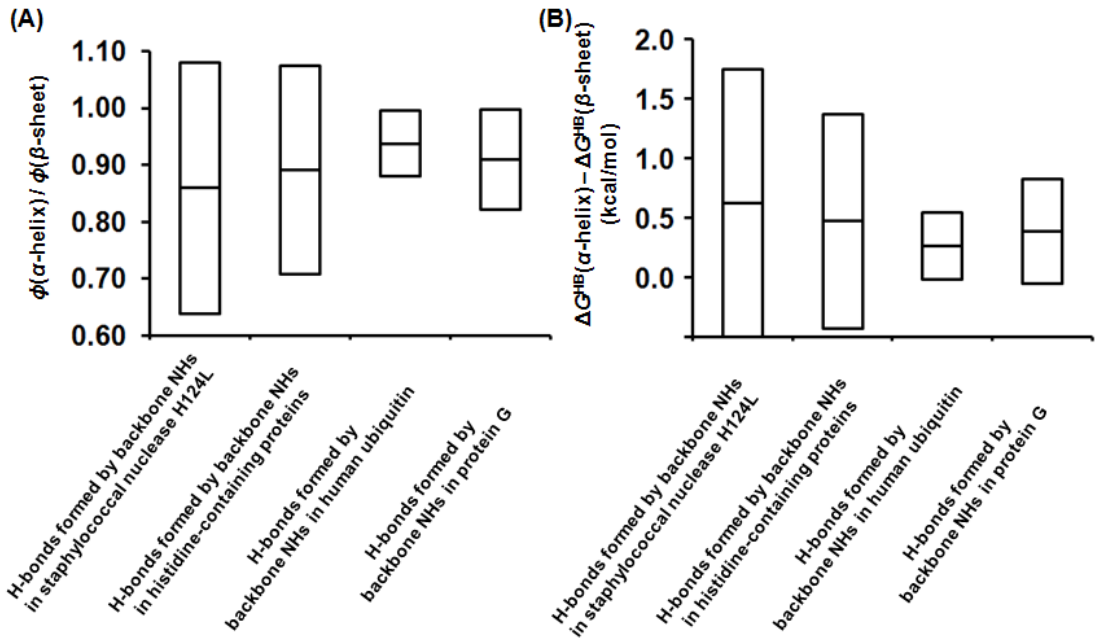


**Fig. 4.3.** Comparison between backbone hydrogen bond strengths in α-helices and β-sheets. (*A*) The ratios between $\phi$-values for $\alpha$-helices and $\beta$-sheets, $\phi(\alpha\text{-helix}) / \phi(\beta\text{-sheet})$ in various proteins. The average values are shown as a bar in the middle of each box and the standard deviations are indicated as the distance between the middle bar and the upper or lower edge of each box. (*B*) The differences between $\Delta G^{HB}$-values for $\alpha$-helices and for $\beta$-sheets, $\Delta G^{HB}(\alpha\text{-helix}) - \Delta G^{HB}(\beta\text{-sheet})$, in various proteins. The average values are shown as a bar in the middle of each box. The standard deviations, determined by considering both the uncertainty in the scale factor and the standard deviations of the $\phi$-values, are indicated as the distance between the middle bar and the upper or lower edge of each box.

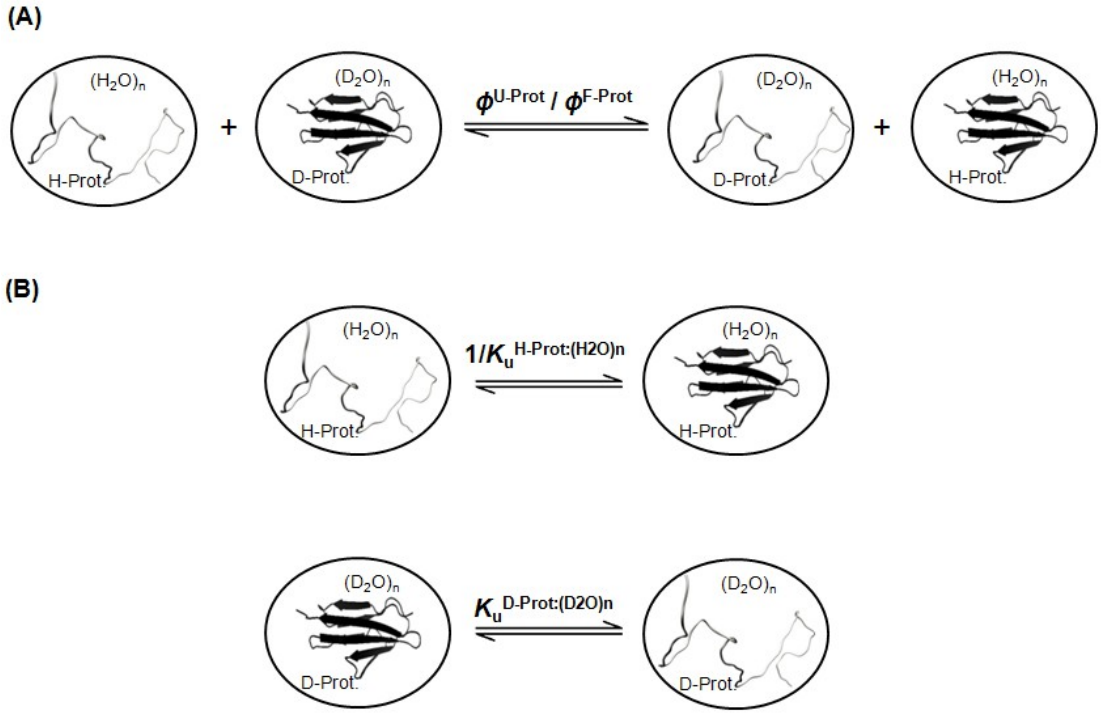


**Fig. 4.4.** Measuring the net hydrogen bond contribution to folding using fractionation factors and the scale factor. (*A*) A hypothetical protium/deuterium exchange reaction between folded and unfolded protein:water complexes. (*B*) Equilibrium folding and unfolding reactions of the protonated and deuterated protein:water complexes, respectively. $(H_2O)_n$ and $(D_2O)_n$ represent the water molecules whose structures are closely affected by the folding and unfolding of the protein.

## 4.4. SUPPORTING INFORMATION APPENDIX

**Table S4.1.** Thermodynamic parameters for model hydrogen-bond complexes.

| Acid | Base | $\phi_1$ | $\phi$ | $RT\ln\phi$ (kcal/mol) | $K$ (M) | $\Delta G^{HB}$ (kcal/mol) |
|---|---|---|---|---|---|---|
| *p*-$CF_3C_6H_4OH$ | $Et_3N$ | 0.59 ± 0.05 | 0.65 ± 0.07 | -0.250 ± 0.062 | 0.0076 ± 0.0015 | 2.93 ± 0.12 |
| PhOH | $Et_3N$ | 0.68 ± 0.02 | 0.76 ± 0.04 | -0.165 ± 0.029 | 0.030 ± 0.002 | 2.10 ± 0.04 |
| $CF_3CH_2OH$ | $Et_3N$ | 0.71 ± 0.04 | 0.79 ± 0.06 | -0.144 ± 0.046 | 0.040 ± 0.004 | 1.93 ± 0.06 |
| $CH_3OH$ | $Et_3N$ | 0.83 ± 0.05 | 0.93 ± 0.07 | -0.046 ± 0.049 | 0.33 ± 0.02 | 0.66 ± 0.04 |
| $Ph_3COH$ | $Et_3N$ | 1.00 ± 0.00 | 1.11 ± 0.02 | 0.063 ± 0.012 | 0.77 ± 0.09 | 0.16 ± 0.07 |
| Pyrrole | $Et_3N$ | 1.10 ± 0.05 | 1.22 ± 0.08 | 0.120 ± 0.041 | 1.1 ± 0.2 | -0.03 ± 0.11 |
| *p*-$CF_3C_6H_4OH$ | *p*-dioxane | 0.97 ± 0.03 | 1.32 ± 0.12 | 0.166 ± 0.054 | 0.17 ± 0.02 | 0.24 ± 0.07 |
| PhOH | *p*-dioxane | 1.01 ± 0.02 | 1.37 ± 0.10 | 0.190 ± 0.045 | 0.22 ± 0.02 | 0.07 ± 0.05 |
| $CF_3CH_2OH$ | *p*-dioxane | 1.03 ± 0.01 | 1.40 ± 0.09 | 0.203 ± 0.040 | 0.40 ± 0.04 | -0.28 ± 0.06 |
| $CH_3OH$ | *p*-dioxane | 1.06 ± 0.03 | 1.45 ± 0.12 | 0.221 ± 0.051 | 1.3 ± 0.2 | -1.00 ± 0.09 |
| $Ph_3COH$ | *p*-dioxane | 1.00 ± 0.00 | 1.36 ± 0.08 | 0.185 ± 0.041 | 0.82 ± 0.17 | -0.72 ± 0.10 |
| Pyrrole | *p*-dioxane | 1.12 ± 0.02 | 1.52 ± 0.11 | 0.252 ± 0.045 | 0.71 ± 0.17 | -0.63 ± 0.14 |
| *p*-$CF_3C_6H_4OH$ | $(n\text{-Pr})_2C{=}O$ | 0.89 ± 0.02 | 1.12 ± 0.08 | 0.066 ± 0.045 | 0.09 ± 0.02 | 1.02 ± 0.11 |
| PhOH | $(n\text{-Pr})_2C{=}O$ | 0.90 ± 0.02 | 1.13 ± 0.08 | 0.073 ± 0.045 | 0.15 ± 0.03 | 0.72 ± 0.10 |
| $CF_3CH_2OH$ | $(n\text{-Pr})_2C{=}O$ | 0.95 ± 0.02 | 1.19 ± 0.09 | 0.105 ± 0.044 | 0.20 ± 0.03 | 0.55 ± 0.09 |
| $CH_3OH$ | $(n\text{-Pr})_2C{=}O$ | 0.93 ± 0.05 | 1.17 ± 0.12 | 0.092 ± 0.064 | 1.0 ± 0.2 | -0.42 ± 0.12 |
| $Ph_3COH$ | $(n\text{-Pr})_2C{=}O$ | 1.00 ± 0.00 | 1.25 ± 0.07 | 0.136 ± 0.032 | 1.6 ± 0.3 | -0.68 ± 0.12 |
| Pyrrole | $(n\text{-Pr})_2C{=}O$ | 1.10 ± 0.04 | 1.35 ± 0.12 | 0.193 ± 0.051 | 1.0 ± 0.1 | -0.42 ± 0.07 |

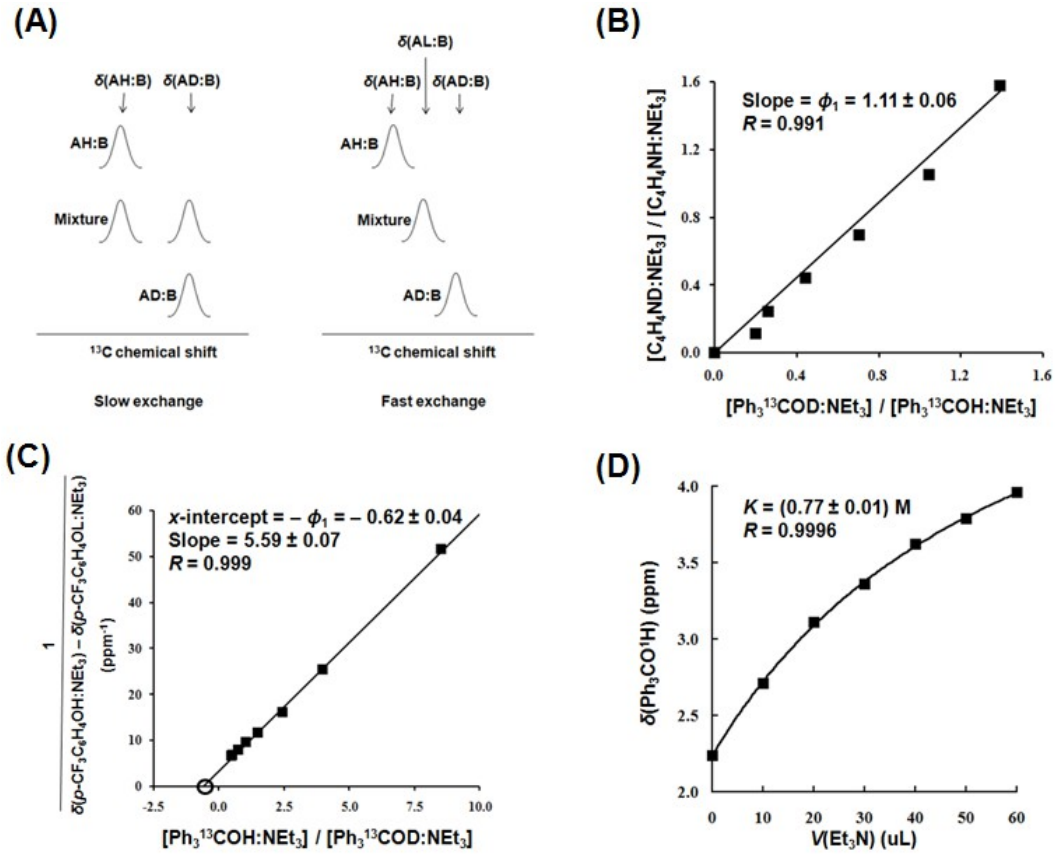


**Fig. S4.1.** Determination of the $\phi_1$-values and the dissociation equilibrium constant of the hydrogen-bond complexes. (*A*) $^{13}$C-NMR peaks for the carbon atoms directly connected to XH groups (X = O or N) are sketched for hydrogen-bond complexes under slow and fast exchanges between the protonated and deuterated states. (*B*) The plot of pyrrole:triethylamine ($C_4H_4NH$:$NEt_3$) is shown as an example for determining the $\phi_1$-values of hydrogen-bond complexes in the show-exchange regime. (*C*) The plot of *p*-trifluoromethylphenol:triethylamine (*p*-$CF_3C_6H_4OH$:$NEt_3$) is shown as an example for determining the $\phi_1$-values of hydrogen-bond complexes in the fast-exchange regime. The position of *x*-intercept was labeled by a circle. (*D*) The chemical shift of the labile proton in the mixture containing the free triphenylmethanol and the hydrogen bond complex triphenylmethanol:triethylamine ($Ph_3COH$:$NEt_3$) is plotted against the volume of added

base as an example for determining the dissociation equilibrium constant of hydrogen-bond complexes.

## 4.5. REFERENCES


1. Bowie JU (2011) Membrane protein folding: how important are hydrogen bonds? *Curr Opin Struct Biol* 21:42–49.

2. Klotz IM, Franzen JS (1962) Hydrogen Bonds between Model Peptide Groups in Solution. *J Am Chem Soc* 84:3461–3466.

3. Eberhardt ES, Raines RT (1994) Amide-Amide and Amide-Water Hydrogen Bonds: Implications for Protein Folding and Stability. *J Am Chem Soc* 116:2149–2150.

4. Fu Y, Gao J, Bieschke J, Dendle MA, Kelly JW (2006) Amide-to-E-olefin versus amide-to-ester backbone H-bond perturbations: Evaluating the O-O repulsion for extracting H-bond energies. *J Am Chem Soc* 128:15948–15949.

5. Gao J, Kelly JW (2008) Toward quantification of protein backbone-backbone hydrogen bonding energies: An energetic analysis of an amide-to-ester mutation in an alpha-helix within a protein. *Protein Sci Publ Protein Soc* 17:1096–1101.

6. Gao J, Bosco DA, Powers ET, Kelly JW (2009) Localized thermodynamic coupling between hydrogen bonding and microenvironment polarity substantially stabilizes proteins. *Nat Struct Mol Biol* 16:684–690.

7. Fersht AR, Matouschek A, Serrano L (1992) The folding of an enzyme: I. Theory of protein engineering analysis of stability and pathway of protein folding. *J Mol Biol* 224:771–782.

8. Marqusee S, Sauer RT (1994) Contributions of a hydrogen bond/salt bridge network to the stability of secondary and tertiary structure in λ repressor. *Protein Sci* 3:2217–2225.

9. Myers JK, Pace CN (1996) Hydrogen bonding stabilizes globular proteins. *Biophys J* 71:2033–2039.

10. Myers JK, Oas TG (1999) Contribution of a buried hydrogen bond to lambda repressor folding kinetics. *Biochemistry (Mosc)* 38:6761–6768.

11. Fernández-Recio J, Romero A, Sancho J (1999) Energetics of a hydrogen bond (charged and neutral) and of a cation-π interaction in apoflavodoxin. *J Mol Biol* 290:319–330.

12. Takano K et al. (1999) Contribution of Intra- and Intermolecular Hydrogen Bonds to the Conformational Stability of Human Lysozyme†,‡. *Biochemistry (Mosc)* 38:12698–12708.

13. Albeck S, Unger R, Schreiber G (2000) Evaluation of direct and cooperative contributions towards the strength of buried hydrogen bonds and salt bridges. *J Mol Biol* 298:503–520.

14. Takano K, Scholtz JM, Sacchettini JC, Pace CN (2003) The Contribution of Polar Group Burial to Protein Stability Is Strongly Context-dependent. *J Biol Chem* 278:31790–31795.

15. Jang DS et al. (2004) Structural double-mutant cycle analysis of a hydrogen bond network in ketosteroid isomerase from Pseudomonas putida biotype B. *Biochem J* 382:967.

16. Campos LA, Cuesta-López S, López-Llano J, Falo F, Sancho J (2005) A Double-Deletion Method to Quantifying Incremental Binding Energies in Proteins from Experiment: Example of a Destabilizing Hydrogen Bonding Pair. *Biophys J* 88:1311–1321.

17. Joh NH et al. (2008) Modest stabilization by most hydrogen-bonded side-chain interactions in membrane proteins. *Nature* 453:1266–1270.

18. Hong H, Blois TM, Cao Z, Bowie JU (2010) Method to measure strong protein-protein interactions in lipid bilayers using a steric trap. *Proc Natl Acad Sci U S A* 107:19802–19807.

19. Jarret RM, Saunders M (1985) A new method for obtaining isotopic fractionation data at multiple sites in rapidly exchanging systems. *J Am Chem Soc* 107:2648–2654.

20. Kreevoy MM, Liang TM (2012) Structures and isotopic fractionation factors of complexes, A1HA2-. *J Am Chem Soc* 102:3315–3322.

21. Kreevoy MM, Liang T-M, Chang K-C (2012) Structures and isotopic fractionation factors of complexes AHA-1. *J Am Chem Soc* 99:5207–5209.

22. Hibbert F, Emsley J (1991) in *Advances in Physical Organic Chemistry*, ed D. Bethell (Academic Press), pp 255–379. Available at: http://www.sciencedirect.com/science/article/pii/S0065316008600477.

23. Vakonakis I, Salazar M, Kang M, Dunbar KR, LiWang AC (2003) Deuterium isotope effects and fractionation factors of hydrogen-bonded A:T base pairs of DNA. *J Biomol NMR* 25:105–112.

24. Shi Z, Krantz BA, Kallenbach N, Sosnick TR (2002) Contribution of hydrogen bonding to protein stability estimated from isotope effects. *Biochemistry (Mosc)* 41:2120–2129.

25. Loh SN, Markley JL (1994) Hydrogen Bonding in Proteins As Studied by Amide Hydrogen D/H Fractionation Factors: Application to Staphylococcal Nuclease. *Biochemistry (Mosc)* 33:1029–1036.

26. Bowers PM, Klevit RE (1996) Hydrogen bonding and equilibrium isotope enrichment in histidine-containing proteins. *Nat Struct Mol Biol* 3:522–531.

27. LiWang AC, Bax A (1996) Equilibrium Protium/Deuterium Fractionation of Backbone Amides in U−13C/15N Labeled Human Ubiquitin by Triple Resonance NMR. *J Am Chem Soc* 118:12864–12865.

28. Khare D, Alexander P, Orban J (1999) Hydrogen bonding and equilibrium protium-deuterium fractionation factors in the immunoglobulin G binding domain of protein G. *Biochemistry (Mosc)* 38:3918–3925.

29. Halkides CJ, Wu YQ, Murray CJ (1996) A Low-Barrier Hydrogen Bond in Subtilisin: 1H and 15N NMR Studies with Peptidyl Trifluoromethyl Ketones†. *Biochemistry (Mosc)* 35:15941–15948.

30. Harris TK, Abeygunawardana C, Mildvan AS (1997) NMR Studies of the Role of Hydrogen Bonding in the Mechanism of Triosephosphate Isomerase†. *Biochemistry (Mosc)* 36:14661–14675.

31. Lin J, Westler WM, Cleland WW, Markley JL, Frey PA (1998) Fractionation factors and activation energies for exchange of the low barrier hydrogen bonding proton in peptidyl trifluoromethyl ketone complexes of chymotrypsin. *Proc Natl Acad Sci* 95:14664–14668.

32. Markley JL, Westler WM (1996) Protonation-State Dependence of Hydrogen Bond Strengths and Exchange Rates in a Serine Protease Catalytic Triad: Bovine Chymotrypsinogen A†. *Biochemistry (Mosc)* 35:11092–11097.

33. Takeda M, Jee J, Terauchi T, Kainosho M (2010) Detection of the Sulfhydryl Groups in Proteins with Slow Hydrogen Exchange Rates and Determination of Their Proton/Deuteron Fractionation Factors Using the Deuterium-Induced Effects on the 13Cβ NMR Signals. *J Am Chem Soc* 132:6254–6260.

34. Edison AS, Weinhold F, Markley JL (2012) Theoretical Studies of Protium/Deuterium Fractionation Factors and Cooperative Hydrogen Bonding in Peptides. *J Am Chem Soc* 117:9619–9624.

35. Mitchell JBO, Price SL (1990) The nature of the N □ H…□O□C hydrogen bond: An intermolecular perturbation theory study of the formamide/formaldehyde complex. *J Comput Chem* 11:1217–1233.

36. Ben-Tal N et al. (1997) Free Energy of Amide Hydrogen Bond Formation in Vacuum, in Water, and in Liquid Alkane Solution. *J Phys Chem B* 101:450–457.

37. Rose GD, Wolfenden R (1993) Hydrogen Bonding, Hydrophobicity, Packing, and Protein Folding. *Annu Rev Biophys Biomol Struct* 22:381–415.

38. Sheu S-Y, Schlag EW, Selzle HL, Yang D-Y (2008) Molecular dynamics of hydrogen bonds in protein-D2O: the solvent isotope effect. *J Phys Chem A* 112:797–802.

39. Parker MJ, Clarke AR (1997) Amide Backbone and Water-Related H/D Isotope Effects on the Dynamics of a Protein Folding Reaction. *Biochemistry (Mosc)* 36:5786–5794.

40. Hermans Jr. J, Scheraga HA (1959) The thermally induced configurational change of ribonuclease in H2O and D2O. *Biochim Biophys Acta* 36:534–535.

41. Brändén C-I, Tooze J (1999) *Introduction to Protein Structure* (Garland Pub.).

42. Schulz GE, Schirmer RH (1979) (Springer-Verlag, New York), pp 33 – 36.

43. Maryott AA, Smith ER (1951) *Table of Dielectric Constants of Pure Liquids* (National Bureau of Standards Circular 514) Available at: http://oai.dtic.mil/oai/oai?verb=getRecord&metadataPrefix=html&identifier=ADA278956 [Accessed September 14, 2013].

44. Jouyban A, Soltanpour S, Chan H-K (2004) A simple relationship between dielectric constant of mixed solvents with solvent composition and temperature. *Int J Pharm* 269:353–360.

45. Fielding L (2000) Determination of Association Constants (Ka) from Solution NMR Data. *Tetrahedron* 56:6151–6170.

# CHAPTER 5

# HYDROGEN BONDS IN A MEMBRANE PROTEIN ARE NOT STRONGER THAN THOSE IN WATER-SOLUBLE PROTEINS.

## 5.1. INTRODUCTION

Hydrogen bonds have long been expected to be more important for membrane proteins than soluble proteins because the hydrocarbon core of the bilayer is a much less polarizable environment than water (1). Moreover, water molecules compete with hydrogen bonds in the protein. Since water-soluble proteins are highly exposed to water, they are expected to have more opportunities to form alternative hydrogen bonds with water than membrane proteins (1). This idea is supported by classic experiments examining dimerization of the model compound N-methylacetamide, which is used to mimic the formation of protein hydrogen bonds. Dimerization was found to be 4 kcal/mol stronger more favorable in carbon tetrachloride than in water (2). In addition, calculations performed by Ben-Tal *et al*. suggest that the energy of an amide hydrogen bond is reduced from 6.6 kcal/mol in vacuum to 1.3 kcal/mol in water (3). However, experimental studies analyzing the effects mutations on protein unfolding free energy have found that hydrogen bond strengths between side chains are comparable to those in water soluble proteins. Even fully buried and highly exposed hydrogen bonds in water-soluble proteins are similar in strength (1).

It is unclear whether these similarities extend to backbone hydrogen bonds, however. I have previously shown that considerable conformational flexibility is possible in transmembrane helices without breaking hydrogen bonds because the hydrogen bonding connections can readily shift between *i*+4 and *i*+3 positions (4). To compare backbone hydrogen bond strengths in membrane proteins with those in water-soluble proteins and avoid the problems summarized by Bowie (1) in measuring side-chain hydrogen-bond strengths in membrane and water-soluble proteins by mutagenesis, I performed the analysis using the equilibrium hydrogen / deuterium fractionation factors.

The equilibrium fractionation factor, $\phi$, between proton (H) and deuteron (D) in a hydrogen-bonded species is defined as $\phi = ([D]/[H])_{protein}/([D]/[H])_{solvent}$. It is the equilibrium constant of the reaction shown in Scheme 5.1.

$$(N\text{-}H)_{protein} + (O\text{-}D)_{Water} \leftrightarrow (N\text{-}D)_{protein} + (O\text{-}H)_{Water}$$

**Scheme 5.1.** Proton/Deuteron exchange between NH/D groups in proteins and OH/D groups in water molecules.

The $\phi$-value of labile hydrogen atoms involved in hydrogen bonds is correlated with the free energy, $\Delta G^{HB}$, of the hydrogen bonds (see Chapter 4). Previously, I have determined the scale factor (SF), which relates these two parameters via $SF = \partial(\Delta G^{HB}) / \partial(RT\ln\phi)$, to be 7.5 ± 0.7, and have used this scale factor to relate $\phi$-values of hydrogen bonds with their free energies (see Chapter 4). Therefore, if I can obtain $\phi$-values for hydrogen bonds in a membrane protein I can obtain the hydrogen bond strengths.

I focused on the voltage-sensing domain (VSD) of the voltage-dependent potassium-selective channel from *Aeropyrum pernix* ($K_V$AP), because its NMR spectrum had been previously assigned, allowing me to measure specific fractionation factors for many hydrogen bonds. VSDs are one of the main molecular detectors for electrical potentials across biological lipid bilayers. The crystal structures of $K_V$ proteins show that the VSDs consist of four transmembrane helices, S1–S4. These four transmembrane helices directly interact with lipid molecules in the membrane and both the structure and function of VSDs have been found to be highly depend on the chemical and physical properties of the lipids (5–8). The Mackinnon lab has determined both the crystal and solution-NMR structures of $K_V$AP-VSD, and they are in good agreement with each other (9, 10). NOEs from the protein residues to the lipids 1,2-diheptanoyl-sn-glycerol-3-phosphocholine (D7PC) used to solubilizing the protein for the solution-NMR experiments indicate that the micelle well mimics the chemical environment of a phospholipid bilayer (10), suggesting that the observed structures are quite similar to the native protein structure in cell membrane.

Here I report my experimental results of $\phi$-values for ~ 70 % of backbone hydrogen bonds and a few side-chain hydrogen bonds in $K_V$AP-VSD by utilizing its well assigned $^{1}H$-$^{15}N$ HSQC spectrum (10) and my analysis on the strengths of those hydrogen bonds after converting the $\phi$-values into the $\Delta G^{HB}$'s by applying the scale factor I determined previously (see Chapter 4).

## 5.2. **RESULTS AND DISCUSSION**

**Backbone hydrogen bond strengths in $K_V$AP-VSD.** By using the $^{1}H$-$^{15}N$ SOFAST-HMQC spectrum, I measured the $\phi$-values for backbone amide groups of ~ 70 % of the non-proline residues and three side chains, N17, Q49 and W70, of the membrane protein $K_V$AP-VSD by fitting Eq. 5.1 as shown in Fig. 5.1. The measured $\phi$-value for the backbone amide groups of this protein ranges from 0.49, for T90, to 1.44, for R151, as shown in Table 5.1 with an average value of 0.97 and a standard deviation of 0.23. Calculated from these $\phi$-values, the free energy of the hydrogen bonds formed by the backbone amide groups with reference to the strength of hydrogen bond formed by the backbone amide group of R151, $\Delta\Delta G^{HB}$, are found to range from 5.0 to 0.0 kcal/mol as shown in Table 5.1 with an average value of 1.9 kcal/mol and a standard deviation of 1.2 kcal/mol.

I can compare the average $\phi$ and $\Delta\Delta G^{HB}$ values for the backbone amide groups in $K_V$AP-VSD and in the four water-soluble proteins mentioned in Chapter 4. I find that this membrane protein has stronger hydrogen bonds on average than human ubiquitin (11) and the two immunoglobulin G binding domains of protein G (12) and has a weaker than the histidine-containing proteins (13) and staphyloccal nuclease H124L (14). This means that the average $\phi$ and $\Delta\Delta G^{HB}$ values for exchangeable NH groups in $K_V$AP-VSD is within the range of average $\phi$ (11–14) and $\Delta\Delta G^{HB}$ values for water-soluble proteins (see Chapter 4). In other words, this membrane protein does not have stronger backbone hydrogen bonds than water-soluble proteins on average. My result here supports the previous results from mutagenesis experiments indicating that side-chain hydrogen bond strengths in membrane proteins and water-soluble proteins are quite similar to each other

(1). Moreover, the distribution of the number of backbone amide groups of this membrane protein in different $\phi$-value ranges has a bell shape as shown in Fig. 5.2, which is also similar to the $\phi$-value distribution for the four water-soluble proteins (11–14) discussed in Chapter 4.

As found previously (12–22), if the hydrogen bond acceptor carries a full negative charge, the exchangeable hydrogen atom from the donor will have a low $\phi$-value. In $K_V$AP-VSD, I also find that the backbone amide group of E53, which is in the loop region and hydrogen-bonds to the negatively charged carboxylic group of its side chain, has a $\phi$-value of 0.60.

Unlike the histidine-containing proteins (13), the backbone amide groups with low $\phi$-values do not cluster together in the folding core region of this membrane protein. Instead, Fig. 5.3 shows the locations of the 14 strongest hydrogen bonds formed by the backbone amide groups with $\phi$-values lower than 0.75 and the 15 weakest ones with $\phi$-values higher than 1.25. The low $\phi$-values tend to be mixed with the high ones in this membrane protein and no folding core can be found based on the positions of the residues with low $\phi$-values. The poor separation between the residues forming strong and weak backbone hydrogen bonds suggests that the TM helices in this membrane protein are very flexible and have a high potential to make conformational change to form more canonical helices, if we assume a canonical helix should have almost identical backbone hydrogen bond strengths. Indeed, there is a lot of mixing between residues forming strong and weak hydrogen bonds in Helices S1 and S3, but there is almost no such mixing in Helix S4, which corresponds well with the fact that Helices S1 and S3 is bent and has breaks, respectively, but Helix S4 is almost straight.

**Flexibility in TM Helix S4 predicted by backbone hydrogen bond strength.** TM helix S4 has the most number of backbone hydrogen bonds whose $\phi$-values have been measured. Thus, an investigation of the relationship between backbone hydrogen bond strength and potential conformational change in this helix is allowed. TM helix in the isolated $K_V$AP-VSD, which only contains helices S1 - S4, has a distinct conformation from the full length version, which contains helices S1- S6. In the isolated $K_V$AP-VSD, S4 is an intact straight helix spanning residues from 120 to 151. However, In the full-length $K_V$AP, S4 is broken at residues 134 - 137 and the C-branch of the helix merges with S5 (9). Fig. 5.4*A* shows residues 122 - 147 in cartoon from the two versions. In Fig 5.4 *B* and *C*, the distances between backbone N atom at residue *i* and backbone O atom at residue *i* - 4 in the two versions of $K_V$AP and strengths of hydrogen bonds formed by backbone NH groups in the isolated $K_V$AP-VSD are plotted against the residue numbers, respectively. Obviously, around residues 134 - 137, where the helix breaks in the full-length form, the hydrogen bonds formed by backbone NH groups have the lowest strength in the isolated form. The means that weak backbone hydrogen bonds in a TM helix is an indicator of helix flexibility.

**Side-chain hydrogen bond strengths in $K_V$AP-VSD.** Interestingly, the side chains tend to have more extreme $\phi$-values. The side-chain amide group of N17 has a $\phi$-value of around 0.60 for both the labile protons and the side-chain amide group of Q49 has a similar result. As shown in Fig. 5.3, these two side chains are in the N-terminal and loop regions, respectively, both exposed to the lipid head groups. Although I cannot find their

hydrogen bond acceptors in the protein structure, it is very likely that N17 and Q49 side chains form hydrogen bonds with the negatively charged phosphate groups from the lipids because these two side chains are both exposed to the lipid head groups (see Fig. 5.3). However, the side-chain NH group of W70 has a $\phi$-value of 1.57, which is the highest $\phi$-value for all the NH groups in this protein. As shown in Fig. 5.3, this side chain is near the center of the bilayer, exposed to the exterior and no probable hydrogen-bond acceptor can be found.

## 5.3. MATERIALS AND METHODS

**Preparation of samples.** The $^{15}$N-labeled $K_V$AP-VSD protein (from Residues 5 to 151) was expressed and purified as described by Butterwick and Mackinnon (10). The purified protein was divided into seven aliquots, each containing ~ 0.28 mM protein, 5.3 mM D7PC, 20 mM KCl and 20 mM HEPES/$Na^+$ (pH 7.00). After repeated dilutions and concentrations using a Amicon Centrifugal Filter Unit with a MWCO of 30 kD purchased from EMD Milipore, the final mole fractions of deuterated water in the seven samples were changed to 0.050, 0.200, 0.350, 0.475, 0.590, 0.700 and 0.940, respectively. For example, to make a final fraction of deuterated water 0.475, I first added a certain amount of buffer which has the same concentrations of D7PC, KCl and HEPES/$Na^+$ (pH/D 7.00) and a fraction of deuterated water of 0.97 to the purified protein sample to make the fraction of deuterated water roughly 0.475. Because the volume of protein samples cannot be measured correctly due to presence of lipids, there is an uncertainty of less than 10 % in the fraction of deuterated water by this step. So, I concentrated the sample to a

protein concentration of ~ 0.28 mM again, followed by adding roughly the same volume of buffer which has the same concentrations of D7PC, KCl and HEPES/$Na^+$ (pH/D 7.00) and a fraction of deuterated water of exactly 0.475 to the sample. This step was repeated 10 times until the final fraction of deuterated water in the sample was exactly 0.475. Finally, the sample was concentrated to a protein concentration of ~ 0.28 mM. The accurate final concentration of the protein in each sample was determined from the UV-absorption spectrum at 280 nm in order to make sure that the final concentrations of the protein did not differ much among all the samples.

**NMR spectroscopy.** $^1$H-$^{15}$N SOFAST-HMQC spectra decoupled during acquisition, (23, 24) were recorded at 45 $^o$C on a Bruker Bosch-800MHz spectrometer equipped with a cryoprobe and were processed using TOPSPIN 3.1. The SOFAST-HMQC NMR technique was applied to only excite the amide protons, whose chemical shifts were between 6.6 and 11.0 ppm, in my protein samples. Since alkyl protons from the protein, whose chemical shifts are outside this range of chemical shift, were left unperturbed, the amide proton $T_1$'s were greatly shortened, allowing the use of a short interscan delay as described by Schanda *et al*. (23, 24). Moreover, as LiWang and Ad Bax pointed out (11), water presaturation and/or the use of an interscan delay shorter than water $T_1$ (~ 2 - 3 sec) (12) attenuates resonances of protons in rapid exchange with water and this attenuation factor is roughly proportional to the $H_2O/D_2O$ ratio, which will result in a decrease in the apparent $\phi$-value. In addition, through the NOE, the intensities of other protons in their vicinity will also be attenuated in the same way and thus also have decreased apparent $\phi$-values. However, in my $^1$H-$^{15}$N SOFAST-HMQC experiment, since the water resonance

($\delta$ = 4.71 ppm) is outside the excited proton band, the water protons were also left unperturbed, which leads to perfect water suppression and no amide groups with attenuated signals due to fast exchange with water.

During the data acquisition, spectral widths of 13.6 and 32.5 ppm were used in the $^{1}H$ and $^{15}N$ dimensions, respectively, with 216 $t_1$ increments of 2 K complex points and 128 transients per increment.

Before the NMR spectra were recorded, each sample was incubated in a 45 $^{o}C$ incubator for 10 hr before the acquisition of their NMR spectrum. The incubation time was found to be sufficient to achieve equilibrium by using the sample with a $D_2O$ fraction of 0.940. A series of $^{1}H$-$^{15}N$ SOFAST-HMQC spectra were recorded for this sample at different times after the $D_2O$ fraction was changed to 0.940. After 10 hr, most of the peaks disappeared, leaving a few peaks which represent the non-exchangeable amide groups as described by Butterwick and Mackinnon (10) still present with constant intensities in time.

As the $D_2O/H_2O$ ratio increases in the solvent, the amide proton $T_1$'s can get longer by ~ 2.8 fold (25). I therefore determined the appropriate interscan delays for each sample. A series of $^{1}H$-$^{15}N$ SOFAST-HMQC spectra were collected for the protein sample with a $D_2O$ fraction of 0.050 at interscan delays of 0.10, 0.20, 0.57, 1.05, 1.52 and 2.00 sec, respectively, and for the protein sample with a $D_2O$ fraction of 0.700 at interscan delays of 0.10, 1.52, 2.00 and 3.00 sec, respectively. For the former sample, no peak volume changes were seen at interscan delays more than 1.05 sec, and for the latter one, no peak volume changes were seen at interscan delays more than 2.00 sec. Therefore, a 1.05 sec interscan delay was used for the sample with a $D_2O$ fraction of

0.050 and a 2.00 sec interscan delay was used for samples with $D_2O$ fractions of 0.200, 0.350, 0.475, 0.590 and 0.700, respectively. To provide a reference for peak volume comparisons, I employed a 10 mM N-acetylglycine sample with a $^{15}N/^{14}N$ ratio of 1/9 in pure $H_2O$ which was sealed in a capillary tube made from Type 1, Class A borosilicate glass with an outer diameter of 1.7 mm purchased from New Era Enterprises, Inc.. The reference sample was coaxially inserted into a high-quality NMR tube designed for 800 NMR spectrometers and with an outer diameter of 5 mm purchased from Wilmad-Lab Glass and transferred from sample to sample.

The data were processed using Guassian functions in both dimensions with zero-filling to give 2K x 2K matrices. A fifth-order polynomial baseline correction was used in both the $^{1}H$ and $^{15}N$ dimensions. Peak volumes were obtained by directly integrating each cross peak. For partially overlapped peaks, the *dcon* command in TOPSPIN 3.1 was used to mathematically separate them and the simulated single peaks were integrated.

**Determination of the fractionation factors.** Fractionation factors were determined from the decoupled $^{1}H$-$^{15}N$ HMQC spectra using methods similar to those described previously (12, 14, 13). In Cross-peak volumes were normalized internally using the sum of the peak volumes of the reference, N-acetylglycine and residues V66, W70 and A74 which have non-exchangeable amide protons and well separated cross peaks. For every cross peak, the inverse of the normalized peak volume, $y$, was plotted against the ratio of deuterated to protonated water, $x$, and the plot was least-square fit to a line according to the following equation(12, 14)

$$1/y = (\phi/y_{max}) \cdot (x + 1/\phi),$$

where $y_{max}$ is the cross-peak volume of the exchangeable sites of the protein in the 100 % protonated state and was fit together with $\phi$-value. The $\phi$-value at each exchangeable site was obtained from the $x$-intercept of the fit line, $-1/\phi$. The software KaleidaGraph 4.1 was used for all the fittings.

The distribution of $\phi$-values for backbone amide groups in histogram was obtained by classifying all the residues into 11 groups from $\phi = 0.40$ to $\phi = 1.50$ with each group spanning a $\phi$-value range of 0.10, *i.e.* $0.40 \leq \phi < 0.50$, $0.50 \leq \phi < 0.60$, ... , and $1.40 \leq \phi < 1.50$.

The free energies of the hydrogen bonds formed by each exchangeable site relative to the weakest backbone hydrogen bond found in this protein, $\Delta\Delta G^{HB}$'s, at room temperature were calculated by first converting the $\phi$-values measured at 45 °C into the fractionation factors at room temperature, $\phi_{298K}$'s, as described (see Chapter 4) and then using the scale factor that I determined previously (see Chapter 4) and the highest $\phi_{298K}$-value that I found for the backbone amide groups of this protein.

**Table 5.1.** Fractionation factors of NH groups and relative strengths of hydrogen bonds formed by the NH groups in $K_V$AP-VSD.

| Residues [a] | $\phi$ [b] | $\Delta\Delta G^{HB}$ [c] | Residues [a] | $\phi$ [b] | $\Delta\Delta G^{HB}$ [c] |
|---|---|---|---|---|---|
| L11 | 0.82 | 2.6 | A77 | - [f] | - [f] |
| G12 | 0.97 | 1.9 | Y78 | 0.94 | 2.0 |
| G13 | 1.20 | 0.9 | K79 | 0.73 | 3.2 |
| R14 | 0.94 | 2.0 | S80 | 0.98 | 1.8 |
| V15 | 1.01 | 1.7 | G81 | 1.01 | 1.7 |
| R16 | 0.82 | 2.6 | D82 | 0.59 | 4.2 |
| N17 | 0.70 | 3.4 | A84 | 1.16 | 1.0 |
| N17$\delta_1$ [d] | 0.60 | 4.1 | G85 | 1.04 | 1.5 |
| N17$\delta_2$ [d] | 0.62 | 3.9 | Y86 | 1.02 | 1.6 |
| I18 | 1.07 | 1.4 | V87 | 0.93 | 2.0 |
| G19 | 0.80 | 2.8 | K88 | 1.09 | 1.3 |
| D20 | 0.90 | 2.2 | K89 | 1.09 | 1.3 |
| V21 | 0.77 | 2.9 | T90 | 0.49 | 5.1 |
| M22 | 1.01 | 1.6 | L91 | 1.17 | 1.0 |
| E23 | 1.26 | 0.6 | Y92 | 1.42 | 0.1 |
| H24 | 0.83 | 2.6 | E93 | 1.11 | 1.2 |
| V27 | 1.31 | 0.4 | I94 | 0.78 | 2.9 |
| E28 | 0.81 | 2.7 | A96 | 1.21 | 0.8 |
| L29/E45 [e] | 1.20 | 0.9 | L97 | 0.52 | 4.8 |
| G30 | 0.85 | 2.5 | V98 | 0.97 | 1.8 |
| V31 | 0.74 | 3.1 | G101 | 1.43 | 0.0 |
| S32 | 0.59 | 4.1 | L102 | 1.39 | 0.2 |
| Y33 | 1.27 | 0.6 | L103 | 1.16 | 1.0 |
| A34 | 0.71 | 3.3 | A104 | 0.80 | 2.7 |
| A35 | - [f] | - [f] | L105 | 0.64 | 3.8 |
| L37 | - [f] | - [f] | I106 | 1.39 | 0.2 |
| V39 | 1.38 | 0.2 | G108 | 1.11 | 1.2 |
| I40 | - [f] | - [f] | H109 | 1.17 | 1.0 |
| V41 | - [f] | - [f] | L110 | 1.28 | 0.6 |
| V43 | - [f] | - [f] | A111 | 0.93 | 2.0 |
| V44 | 1.31 | 0.4 | G112 | 0.97 | 1.9 |
| Y46 | 1.02 | 1.6 | L113 | 0.89 | 2.3 |
| T47 | 0.59 | 4.2 | L115 | 1.26 | 0.6 |
| M48 | 0.94 | 2.0 | F116 | - [f] | - [f] |
| Q49 | 0.97 | 1.8 | L118 | - [f] | - [f] |
| Q49$\varepsilon_1$ [d] | 0.63 | 3.9 | V119 | 1.11 | 1.2 |

| | | | | | |
|---|---|---|---|---|---|
| Q49ε$_2$ [d] | 0.60 | 4.1 | R120 | 1.14 | 1.1 |
| L50 | 1.41 | 0.1 | L121/L128 [e] | 1.17 | 1.0 |
| S51 | 0.87 | 2.4 | L122 | 0.99 | 1.8 |
| G52 | 1.38 | 0.2 | R123 | 0.69 | 3.5 |
| E53 | 0.60 | 4.1 | F124 | - [f] | - [f] |
| Y54 | 0.69 | 3.4 | L125 | 0.87 | 2.4 |
| L55 | 0.98 | 1.8 | R126 | 0.92 | 2.1 |
| V56/D146 [e] | 0.68 | 3.5 | I127 | - [f] | - [f] |
| R57 | 1.07 | 1.4 | I130 | - [f] | - [f] |
| L58 | 1.00 | 1.7 | I131 | 1.17 | 1.0 |
| Y59 | 0.85 | 2.5 | S132 | 0.85 | 2.4 |
| L60 | 1.25 | 0.6 | R133 | - [f] | - [f] |
| V61 | - [f] | - [f] | G134 | 1.19 | 0.9 |
| D62 | - [f] | - [f] | S135 | 0.98 | 1.8 |
| L63 | - [f] | - [f] | K136 | - [f] | - [f] |
| I64 | - [f] | - [f] | F137 | - [f] | - [f] |
| L65 | - [f] | - [f] | L138 | 0.92 | 2.1 |
| V66 | - [f] | - [f] | S139 | 0.57 | 4.3 |
| I67 | - [f] | - [f] | A140 | 1.10 | 1.3 |
| I68 | - [f] | - [f] | I141 | 0.95 | 2.0 |
| L69 | - [f] | - [f] | A142 | 0.78 | 2.9 |
| W70 | - [f] | - [f] | D143 | 0.83 | 2.6 |
| W70ε [d] | 1.57 | -0.4 | A144 | 0.57 | 4.3 |
| A71 | - [f] | - [f] | A145 | 0.76 | 3.0 |
| Y73 | - [f] | - [f] | K147 | 0.86 | 2.4 |
| A74 | - [f] | - [f] | L148 | 0.90 | 2.2 |
| Y75 | - [f] | - [f] | V149 | 1.09 | 1.3 |
| R76 | - [f] | - [f] | R151 | 1.44 | 0.0 |

[a] For backbone amide groups unless noted.

[b] All the ϕ-values have an uncertainty of < 10 % from the fitting using Eq. 5.1.

[c] All the $\Delta\Delta G^{HB}$ values have an uncertainty of < 25 % determined from the uncertainties of ϕ-values and the scale factor.

[d] For side-chain NH groups.

[e] Due to 100 % overlapping, the two peaks for Residues X and Y cannot be separated but their total peak volumes can be fit using Eq. 5.1 to get an apparent ϕ-value. If the

apparent ϕ-value is very low or very high, at least one of the residues, either X or Y with a larger peak volume, has a very low or very high ϕ-value, respectively, but the ϕ-value of the other one that has a smaller peak volume cannot be determined because the volume ratio of the two overlapped peaks may be very large.

[f] Data unavailable either because the NH groups are not exchangeable or their $^{1}H$-$^{15}N$ HMQC peaks are seriously overlapped with others.

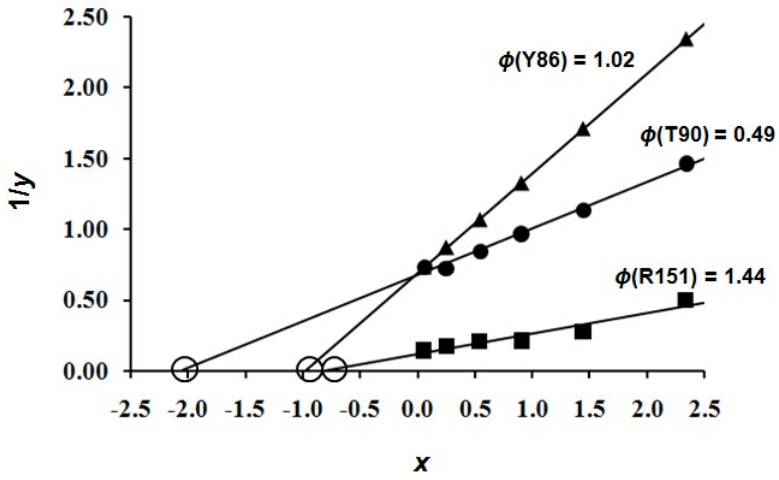


**Fig. 5.1**. Determining $\phi$-values of exchangeable NH groups in $K_V$AP-VSD. Plots of the fitting for the $\phi$-value of three backbone amide groups, T90, R151 and Y86, which have the lowest, the highest and a middle-level $\phi$-values, as an example. The $x$-intercepts of three fit lines are circled. The fitting correlation coefficients, $R$, for T90, R151 and Y86, are 0.995, 0.979 and 1.000, respectively.

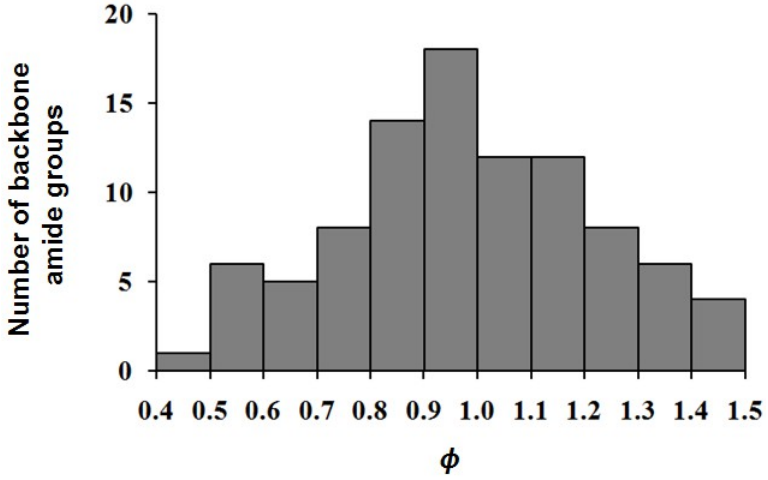


**Fig. 5.2.** Distribution of the number of $K_V$AP-VSD backbone amide groups in different $\phi$-value ranges.

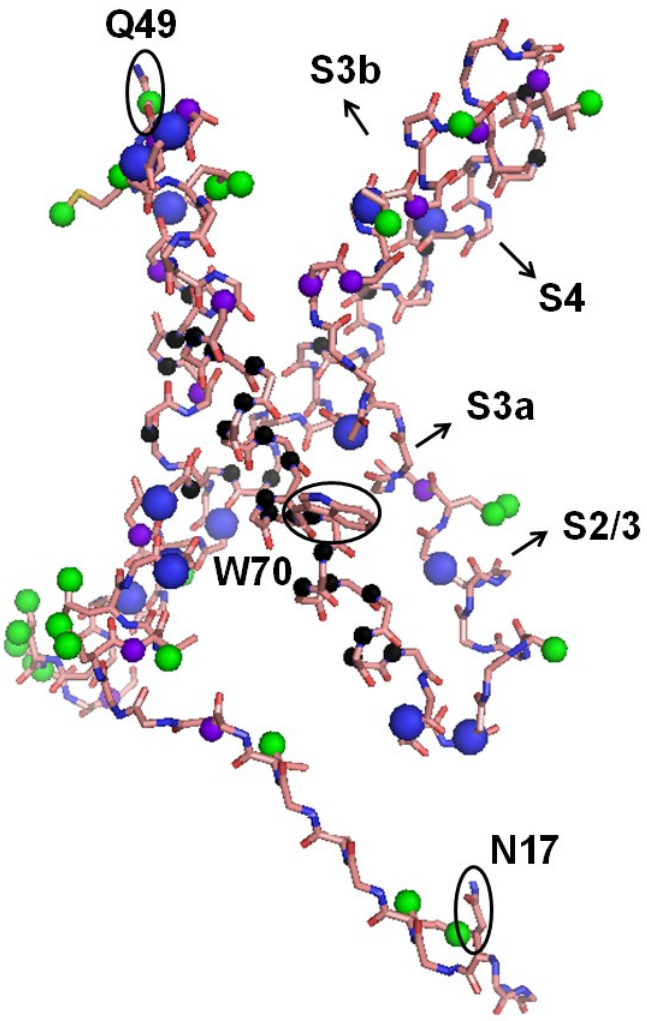


**Fig. 5.3.** Positions of different types of amide groups. The backbone atoms in the KVAP-VSD structure (PDB code: 1ORS) are shown in stick with the carbon, oxygen and nitrogen atoms colored in orange, red and blue or purple, respectively. The backbone amide groups with $\phi$-values lower than 0.75, higher than 1.25 and unavailable are labeled with large blue, small purple and small black spheres, respectively, at the position of the nitrogen atoms. The side-chain carbon atoms whose NOEs to lipid head groups have been detected by Butterwick and Mackinnon (data acquired through private email) are labeled with green spheres and their whole side chains are shown in stick. The side chains of N17 and Q49 are both shown in stick and are highlighted with ellipses.

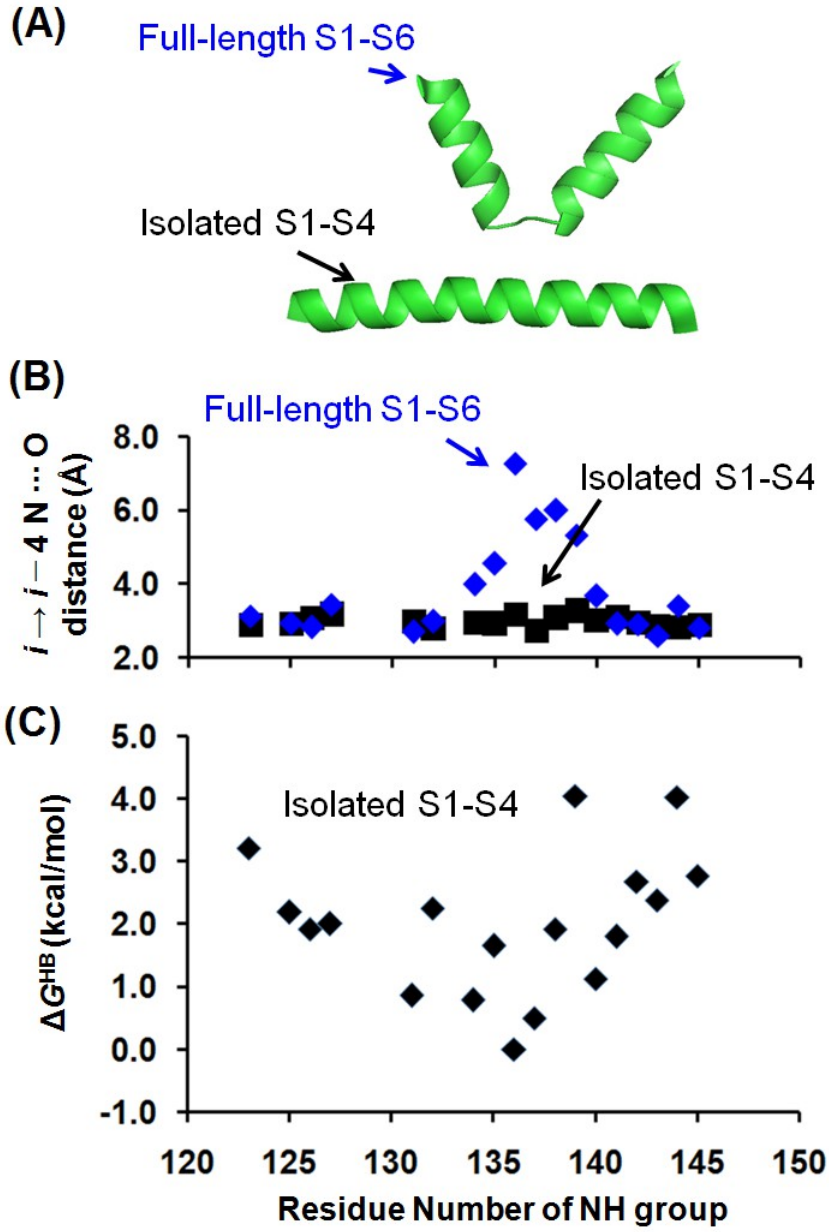


**Fig. 5.4.** Flexibility in TM helix S4 predicted by backbone hydrogen bond strength. (*A*) Residues 122 - 147 in the isolated and full-length $K_VAP$ are shown in cartoon. (*B*) the distances between backbone N atom at residue *i* and backbone O atom at residue *i* - 4 in the two versions of $K_VAP$ are plotted against the residue numbers. (*C*) The strengths of

hydrogen bonds formed by backbone NH groups in isolated $K_V$AP-VSD are plotted against the residue numbers.

## 5.4. REFERENCES


1. Bowie JU (2011) Membrane protein folding: how important are hydrogen bonds? *Curr Opin Struct Biol* 21:42–49.

2. Klotz IM, Farnham SB (1968) Stability of an amide-hydrogen bond in an apolar environment. *Biochemistry (Mosc)* 7:3879–3882.

3. Ben-Tal N et al. (1997) Free Energy of Amide Hydrogen Bond Formation in Vacuum, in Water, and in Liquid Alkane Solution. *J Phys Chem B* 101:450–457.

4. Cao Z, Bowie JU (2012) Shifting hydrogen bonds may produce flexible transmembrane helices. *Proc Natl Acad Sci* 109:8121–8126.

5. Ramu Y, Xu Y, Lu Z (2006) Enzymatic activation of voltage-gated potassium channels. *Nature* 442:696–699.

6. Schmidt D, MacKinnon R (2008) Voltage-dependent K+ channel gating and voltage sensor toxin sensitivity depend on the mechanical state of the lipid membrane. *Proc Natl Acad Sci* 105:19276–19281.

7. Schmidt D, Jiang Q-X, MacKinnon R (2006) Phospholipids and the origin of cationic gating charges in voltage sensors. *Nature* 444:775–779.

8. Xu Y, Ramu Y, Lu Z (2008) Removal of phospho-head groups of membrane lipids immobilizes voltage sensors of K+ channels. *Nature* 451:826–829.

9. Jiang Y et al. (2003) X-ray structure of a voltage-dependent K+ channel. *Nature* 423:33–41.

10. Butterwick JA, MacKinnon R (2010) Solution structure and phospholipid interactions of the isolated voltage-sensor domain from KvAP. *J Mol Biol* 403:591–606.

11. LiWang AC, Bax A (1996) Equilibrium Protium/Deuterium Fractionation of Backbone Amides in U−13C/15N Labeled Human Ubiquitin by Triple Resonance NMR. *J Am Chem Soc* 118:12864–12865.

12. Khare D, Alexander P, Orban J (1999) Hydrogen bonding and equilibrium protium-deuterium fractionation factors in the immunoglobulin G binding domain of protein G. *Biochemistry (Mosc)* 38:3918–3925.

13. Bowers PM, Klevit RE (1996) Hydrogen bonding and equilibrium isotope enrichment in histidine-containing proteins. *Nat Struct Mol Biol* 3:522–531.

14. Loh SN, Markley JL (1994) Hydrogen Bonding in Proteins As Studied by Amide Hydrogen D/H Fractionation Factors: Application to Staphylococcal Nuclease. *Biochemistry (Mosc)* 33:1029–1036.

15. Kreevoy MM, Liang TM (2012) Structures and isotopic fractionation factors of complexes, A1HA2-. *J Am Chem Soc* 102:3315–3322.

16. Kreevoy MM, Liang T-M, Chang K-C (2012) Structures and isotopic fractionation factors of complexes AHA-1. *J Am Chem Soc* 99:5207–5209.

17. Halkides CJ, Wu YQ, Murray CJ (1996) A Low-Barrier Hydrogen Bond in Subtilisin: 1H and 15N NMR Studies with Peptidyl Trifluoromethyl Ketones†. *Biochemistry (Mosc)* 35:15941–15948.

18. Harris TK, Abeygunawardana C, Mildvan AS (1997) NMR Studies of the Role of Hydrogen Bonding in the Mechanism of Triosephosphate Isomerase†. *Biochemistry (Mosc)* 36:14661–14675.

19. Lin J, Westler WM, Cleland WW, Markley JL, Frey PA (1998) Fractionation factors and activation energies for exchange of the low barrier hydrogen bonding proton in peptidyl trifluoromethyl ketone complexes of chymotrypsin. *Proc Natl Acad Sci* 95:14664–14668.

20. Takeda M, Jee J, Terauchi T, Kainosho M (2010) Detection of the Sulfhydryl Groups in Proteins with Slow Hydrogen Exchange Rates and Determination of Their Proton/Deuteron Fractionation Factors Using the Deuterium-Induced Effects on the 13Cβ NMR Signals. *J Am Chem Soc* 132:6254–6260.

21. Markley JL, Westler WM (1996) Protonation-State Dependence of Hydrogen Bond Strengths and Exchange Rates in a Serine Protease Catalytic Triad: Bovine Chymotrypsinogen A†. *Biochemistry (Mosc)* 35:11092–11097.

22. Edison AS, Weinhold F, Markley JL (2012) Theoretical Studies of Protium/Deuterium Fractionation Factors and Cooperative Hydrogen Bonding in Peptides. *J Am Chem Soc* 117:9619–9624.

23. Schanda P, Brutscher B (2005) Very Fast Two-Dimensional NMR Spectroscopy for Real-Time Investigation of Dynamic Events in Proteins on the Time Scale of Seconds. *J Am Chem Soc* 127:8014–8015.

24. Schanda P, Kupče Ē, Brutscher B (2005) SOFAST-HMQC Experiments for Recording Two-dimensional Deteronuclear Correlation Spectra of Proteins within a Few Seconds. *J Biomol NMR* 33:199–211.

25. Markus MA, Dayie KT, Matsudaira P, Wagner G (1994) Effect of deuteration on the amide proton relaxation rates in proteins. Heteronuclear NMR experiments on villin 14T. *J Magn Reson B* 105:192–195.

# CHAPTER 6

## SUMMERY

My study described in Chapter 2 on transmembrane helix flexibility has indicated that transmembrane helices are quite flexible. The introduction of kinks appears to be well within the realm of simple evolutionary steps and helix distortions can be readily accessed during conformational changes. I have proposed that flexibility may at least in part be explained by backbone hydrogen bonding donors and acceptors shifting to different partners. While the introduction of a Pro residue in a helix is one dramatic way to distort helices, our results have indicated that the structure of transmembrane helices can be altered readily by a single point mutation because the energy cost for helix bending is not extremely high. In the absence of P50, Helix B in bR can be shifted to a distinct conformation ("straight" but still non-canonical helix) by a single mutation (T46A). This conformational shift only cost ~ 0.6 kcal/mol in free energy. My results suggest that membrane protein structure is much more malleable than people might have imagined; a feature that has apparently been essential for optimizing membrane protein structure and function.

My study described in Chapter 3 on the equilibrium of unfolding $bR_f$ to $bO_u$ has proved that unfolding under the new condition is reversible. I showed that unfolding free energy in the transition zones is linearly dependent on SDS concentration using the new method, which allows simple short-range extrapolation of the unfolding free energies and thus allows comparison between unfolding free energies of mutant and wild-type bR

proteins at the same SDS concentration. I found that the energetic contributions by most of the side-chains in Helix B of bR obtained using this new method are similar to the results obtained using the old method, although by using the old method the equilibrium of bR unfolding was not reached.

My study described in Chapter 4 has provided a reliable way to convert $\phi$-values of protein hydrogen bonds into a relative energetic scale by SF = $\partial\Delta G^{HB}$ / ($\partial RT\ln\phi$) = -7.0 ± 0.7. By employing our method, it is possible to study the relative free energies of a large number of protein hydrogen bonds at the same time without unfolding the protein or employing mutagenesis, as long as the NMR spectrum has been assigned. This is extremely useful for studying the strengths of backbone hydrogen bonds and side-chain hydrogen bonds which are involved in a complex hydrogen bond network. By making the conversion from the previously measured $\phi$-values of hydrogen bonds in water-soluble proteins, I have found that the strengths of backbone hydrogen bonds in those proteins span a large range in free energy (~ 7 kcal/mol) and their standard deviations seem to be correlated with the sizes of the proteins. I have also shown that backbone hydrogen bonds in $\alpha$-helices are ~ 0.5 kcal/mol stronger than those in $\beta$-sheets on average. Moreover, I have shown that the appearance of a charge in the protein hydrogen bond acceptor greatly increases the strength of the protein hydrogen bond by ~ 2 kal/mol. During the catalytic cycle of enzymes, hydrogen bonding interactions at the catalytic center can stabilize the intermediate state by up to ~ 4 kcal/mol more than the resting state. Last, I have shown that the overall hydrogen bonding interactions in a protein solution marginally favors soluble-protein unfolding by < 10 kcal/mol.

My study described in Chapter 5 has shown that the membrane protein $K_V$AP-VSD does not have stronger backbone hydrogen bonds than water-soluble proteins on average. Another interesting finding from our experiment is that all the strongest and weakest hydrogen bonds formed by backbone amide groups are mixed with each other. Both of these two findings support the previous conclusion that side-chain hydrogen bond strengths measured via mutagenesis in membrane proteins, in the fully buried folding core of water-soluble proteins and on the surface of water-soluble proteins are quite similar to each other. I have suggested that the poor separation between residues forming strong and weak backbone hydrogen bonds, especially in Helices S1 and S3, is a sign that TM helices are flexible in this membrane protein. More importantly, I have found that residues forming weak backbone hydrogen bonds at the center of TM Helix S4 have a high potential to have conformational change.

In sum, under the instruction of Dr. James U. Bowie at UCLA, I have made a great progress in research on proteins, especially membrane proteins, which further our understanding of protein energetics and structures. Much work still needs to be done to fully reveal the difference between membrane and water-soluble proteins. I believe our studies open a new door for future research on this area.