The Contribution of C_α−H···O Hydrogen Bonds to Membrane Protein Stability Depends on the Position of the Amide†

Madhusoodanan Mottamal and Themis Lazaridis*

Department of Chemistry, City College of New York / CUNY, 138th Street and Convent Avenue, New York, New York 10031

Received September 8, 2004; Revised Manuscript Received November 8, 2004

ABSTRACT: Structural analyses of membrane proteins reveal a large number of C_α−H···O contacts between transmembrane helices, presumed to be hydrogen bonds. Recent experiments produced conflicting results for the contribution of such hydrogen bonds to membrane protein stability. An FTIR study estimated an energy of −0.88 kcal/mol for the G79-C_α−H···I76-O hydrogen bond in glycophorin A, whereas a mutagenesis study showed that the A51-C_α−H···T24-O_γ hydrogen bond does not stabilize bacteriorhodopsin. Here, we reconcile these results using molecular mechanics calculations and an implicit membrane model (IMM1). With explicit hydrogen atoms, the potential energy of the G79-C_α−H···I76-O interaction in GpA ranges from −0.54 to −0.9 kcal/mol and its contribution to stability (effective energy) from −0.49 to −0.83 kcal/mol, depending on the structural model used. The average values of these quantities in GpA-like motifs are similar. In bR, the corresponding numbers for the A51-C_α−H···T24-O_γ interaction are +0.15 and +0.32 kcal/mol. The difference results from the different arrangement of the interacting groups and specifically the position of the acceptor with respect to the C_α and N atoms. This conclusion likely applies to soluble proteins as well.

Membrane proteins account for about 30% of all proteins but less than 0.5% of known structures. Hence, theoretical methods for predicting membrane protein structure from sequence could be of great value. Understanding and quantifying the forces that determine membrane protein structure is essential in such efforts.

The folding of α-helical membrane proteins is envisioned to occur in two stages: in the first stage, α-helices are formed and inserted into the membrane, driven by the hydrophobic effect. In the second stage the helices associate with each other to form the final folded structure (*1*, *2*). The driving forces for the second stage are most likely van der Waals packing and polar interactions, such as hydrogen bonds, the latter being especially strong in the low polarity of the membrane interior. However, strongly polar residues are relatively infrequent in TM[1] helices and may, in fact, lead to nonspecific ("promiscuous") association (*3*−*5*). A weaker interaction proposed to be more "gentle" and specific is C−H···O hydrogen bonds (*6*). Surveys of high-resolution protein structures reveal widespread occurrence of close C−H···O contacts (*6*−*11*), the majority of which involve hydrogens bonded to α-carbons. Quantum mechanical calculations on model systems in the gas phase gave a C_α−H···O hydrogen bond energy of 2.5−3.0 kcal/mol (*12*−*14*), almost half the strength of a conventional hydrogen bond.

The simplest membrane protein where interhelical C_α−H···O hydrogen bonds are evident is the transmembrane helix dimer of glycophorin A (GpA). Its NMR structure in detergent and lipid bilayers shows that a GXXXG motif forms the binding interface (*15*, *16*). The close apposition of the helices, enabled by the small side chain of the Gly residues, allows the formation of C_α−H···O hydrogen bonds. GXXXG motifs have also been found in polytopic membrane proteins, such as glycerol facilitator (*17*), calcium ATPase (*18*) and photosystem I (*19*). More than 50% of the GXXXG-containing α-helices have been found to participate in helix-helix interactions, suggesting that the GXXXG motif is a determinant of helix association in membranes (*20*). C_α−H···O contacts are also observed outside GXXXG motifs. For example, six C_α−H···O contacts in bacteriorhodopsin (bR) exhibit H···O distance <2.7 Å (*6*).

Recent experiments gave conflicting results for the contribution of these hydrogen bonds to membrane protein stability. Arbely and Arkin (*21*) used an empirical correlation between vibrational frequency shifts and interaction energy and estimated an energy of −0.88 kcal/mol for the G79-C_α−H···I76-O hydrogen bond in GpA. Yohannan et al. (*22*) estimated the contribution of the A51-C_α−H···T24-O_γ hydrogen bond to the stability of bR by mutating Thr24 to Ala, Val, and Ser. They found that the mutants were marginally less stable or even more stable than wild type, suggesting that this C_α−H···O bond does not stabilize the protein.

The aim of this work is to reconcile the above results by calculating the energetics of C_α−H···O hydrogen bonds

† Financial support was provided by the National Science Foundation (MCB-0316667) and by NIH (2S06GM008168−23). Computer resources were partially provided by an RCMI grant from NIH (SGI2RR003060).

* Corresponding author. E-mail: tlazaridis@ccny.cuny.edu. Telephone: 212-650-8364. Fax: 212-650-6107.

[1] Abbreviations: GpA, Glycophorin A; bR, bacteriorhodopsin; IMM1, Implicit membrane model 1; IMM1-P22, implicit membrane model 1-parameter22; EEF1, effective energy function 1; ABNR, adopted basis Newton Raphson; TM, transmembrane.

10.1021/bi048065s CCC: $30.25
Published on Web 01/12/2005

using a molecular mechanics energy function that takes into account the membrane environment implicitly (IMM1). The "strength" of these bonds is quantified as the interaction energy between the neutral groups containing the hydrogen bond donor and acceptor. In addition to GpA and bR, we performed calculations on other membrane proteins containing GpA-like motifs (glycerol facilitator and calcium ATPase).

METHODS

The implicit membrane model IMM1 implemented in the program CHARMM was used to model the lipid bilayer (*23*). It is an extension of the effective energy function EEF1, which has been successfully used to model proteins in solution (*24*, *25*). The solvation free energy of a protein in EEF1 is treated as the sum of group contributions. The solvation free energy of a group is equal to a reference value, obtained from experimental data for small model compounds in water, minus the amount of solvation lost due to exclusion of solvent by other protein atoms. In IMM1, the same idea is used for proteins inside a model membrane, which lies parallel to the xy plane with its center at $z = 0$. The solvation free energy parameters in the nonpolar core of the membrane were derived from experimental data for small model compounds in cyclohexane. The solvation parameters of all atoms are dependent on the vertical coordinate, z or $z' = |z|/(T/2)$, where T is the thickness of the nonpolar core of the membrane

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

and ΔG_i represents the solvation free energy of atom i. The transition from the membrane interior to water is accomplished by the function $f(z')$

$$f(z') = \frac{z'^n}{1 + z'^n} \quad (2)$$

where n controls the steepness of the transition ($n = 10$). In addition, in IMM1 a modified dielectric screening function is used to account for the strengthening of electrostatic interaction in the membrane that is compatible with the distance dependent dielectric used in EEF1. This function has the following form

$$\epsilon = r^{(a+(1-a)\sqrt{f_i f_j})} \quad (3)$$

where f_i and f_j are obtained from eq 2; a value of 0.85 for "a" was found to give reasonable results for various systems studied earlier (*23*). The hydrophobic thickness of a membrane depends on the lipid and usually ranges between 20 and 30 Å. In this study, a thickness of 26 Å was used, as in previous work.

The IMM1 model is based on the charmm 19 force field which treats the nonpolar hydrogen atoms implicitly. For example, a C_α−H group is treated as a united atom. To estimate the energy of C_α−H···O contacts, explicit representation of the hydrogen atom may be important. Thus a model called IMM1-P22, based on the all-atom charmm 22 force field (*26*), is also used in our calculations. Both energy functions treat hydrogen bonds as a purely electrostatic interaction.

For GpA, we used two initial structures: one was obtained by solution NMR in detergents (pdb code: 1AFO) (*15*); and the other by solid-state NMR in lipid bilayers (*16*, *27*). Solution NMR produced 20 models in most of which the distance between G79-C_α and I76-O is greater than the ideal C_α···O hydrogen bonding distance (≤3.8 Å) (*8*). Hence, we focused mainly on the 19th model, where all C_α···O contacts that are thought to make hydrogen bonds are <3.8 Å. However, we also obtained the average C_α···O hydrogen bonding energy over all 20 structures. Amino acid residues 73−95 were included in the calculations. The initial structures were oriented with their principal axis perpendicular to the membrane surface and the energetics of these structures in the IMM1 membrane model was obtained after 300 steps of adopted basis Newton Raphson (ABNR) energy minimization. The X-ray structure of bR (pdb code: 1PY6) was again oriented with its principal axis perpendicular to the membrane. This makes all helices more or less perpendicular to the membrane. The protein was translated along the z-axis such that a maximum of hydrophobic residues reside in the hydrophobic membrane interior. Then the system was minimized for 300 steps by the ABNR method. The same protocol was used for glycerol facilitator (1fx8.pdb) and calcium ATPase (1eul.pdb).

A reasonable definition for the energy of a hydrogen bond is the interaction energy between the neutral groups that contain the H donor and the acceptor (*28*). In the case of the CHARMM force field, such neutral groups are NHC_α and CO for the backbone and $C_\beta H_\beta O_\gamma H_\gamma$ for the Thr side chain. In the EEF1 and IMM1 energy functions, the solvation free energy is pairwise decomposable. The interaction energy reported by CHARMM includes the desolvation cost resulting from this interaction; i.e., it is an effective interaction making a thermodynamic contribution to stability in a solvent environment. We also report the potential energy of these interactions, i.e., the van der Waals and electrostatic components only.

RESULTS

It has been proposed that the C_α−H···O hydrogen bonds between residues Gly79:Ile76, Val80:Gly79, Gly83:Val80, and Val84:Thr87 stabilize the GpA dimer (*6*). The C_α···O distances in the energy-minimized structures (Table 1) are shorter than 3.8 Å, except for the G79-C_α···O.I79 interaction. Thus, these C_α−H···O hydrogen bonds are structurally possible. The effective interactions between the NHC_α and CO groups range from −0.08 to −0.42 kcal/mol, except for V84-C_α···T87-O_γ which gives a positive effective interaction energy in the solid-state structure. This is so because in this structure the Thr hydroxyl makes a hydrogen bond with the backbone CO of V84 and thus points its hydrogen toward the C_α−H of V84, leading to an unfavorable electrostatic interaction with the C_α group of V84. In the solution structure Thr87 makes a hydrogen bond with the CO of Gly83 in the same helix.

The interactions in Table 1 may be underestimated due to lack of explicit representation of H_α atoms in IMM1. Table 2 shows the interhelical distances and interaction energies using a membrane model with explicit H_α atoms (IMM1-P22). Here again, with one exception (G79-C_α···I76-O), all distances are shorter than 3.8 Å. The C_α and O group

Table 1: Interhelical Distances between C_α and O or O_γ Atoms in the Energy-Minimized GpA Structures and the Interaction Energy between the NHC_α and CO or $C_\beta O_\gamma H_\gamma$ Groups in IMM1[a]

	solution NMR structure (model 19)		solid-state NMR structure	
donor:acceptor	$R_{C_\alpha\cdots O}$	total energy (vdw, ele, solv)	$R_{C_\alpha\cdots O}$	total energy (vdw, ele, solv)
G79-C_α:I76-O	4.00	−0.34 (−0.27, −0.14, +0.07)	4.03	−0.34 (−0.26, −0.14, +0.05)
	3.72	−0.40 (−0.30, −0.18, +0.09)	4.03	−0.33 (−0.26, −0.14, +0.07)
V80-C_α:G79-O	3.33	−0.13 (−0.20, −0.10, +0.17)	3.37	−0.15 (−0.20, −0.11, +0.16)
	3.36	−0.08 (−0.13, −0.11, +0.17)	3.36	−0.14 (−0.18, −0.12, +0.16)
G83-C_α:V80-O	3.51	−0.41 (−0.33, −0.18, +0.10)	3.68	−0.42 (−0.34, −0.17, +0.10)
	3.48	−0.41 (−0.34, −0.17, +0.09)	3.69	−0.41 (−0.34, −0.17, +0.09)
V84-C_α:T87-O_γ	3.52	−0.20 (−0.17, −0.14, +0.11)	3.46	+0.25 (+0.05, +0.08, +0.12)
	3.62	−0.25 (−0.21, −0.14, +0.11)	3.36	+0.33 (+0.14, +0.06, +0.13)
av effective energy		−0.28 ± 0.13		−0.30 ± 0.12[b]
av potential energy		−0.39 ± 0.10		−0.40 ± 0.10[b]

[a] All energies are in kcal/mol and distances are in Å. The values in parentheses are van der Waals, electrostatic, and desolvation contributions, in this order. Two values are given in each category, one for each helix. The C_α group includes amide N, amide H, and the α carbon. The O group includes the carbonyl carbon and oxygen atoms. The O_γ group includes the γ oxygen, the γ hydrogen, and the β carbon. [b] The V84-C_α···T87-O_γ contact is omitted from this average.

Table 2: Interhelical Distances between C_α and O or O_γ Atoms in the Energy-Minimized GpA Structures and the Interaction Energy between the NHC_α and CO or $C_\beta H_\beta O_\gamma H_\gamma$ Groups in IMM1-P22[a]

	solution NMR structure (model 19)		solid-state NMR structure	
donor:acceptor	$R_{C_\alpha\cdots O}$	total energy (vdw, ele, solv)	$R_{C_\alpha\cdots O}$	total energy (vdw, ele, solv)
G79-C_α:I76-O	3.70	−0.90 (−0.35, −0.64, +0.09)	4.10	−0.64 (−0.26, −0.44, +0.06)
	3.92	−0.75 (−0.31, −0.51, +0.07)	4.08	−0.65 (−0.27, −0.45, +0.07)
V80-C_α: G79-O	3.35	−0.85 (−0.11, −0.92, +0.18)	3.39	−0.88 (−0.17, −0.88, +0.17)
	3.37	−0.79 (−0.08, −0.89, +0.18)	3.39	−0.88 (−0.16, −0.89, +0.17)
G83-C_α: V80-O	3.66	−0.92 (−0.39, −0.63, +0.10)	3.74	−0.90 (−0.37, −0.62, +0.09)
	3.71	−0.92 (−0.38, −0.65, +0.10)	3.76	−0.88 (−0.37, −0.60, +0.09)
V84-C_α:T87-O_γ	3.56	−0.61 (+0.04, −0.77, +0.11)	3.46	+0.58 (+0.13, +0.34, +0.12)
	3.66	−0.68 (−0.11, −0.67, +0.10)	3.42	+0.63 (+0.25, +0.25, +0.13)
av effective energy		−0.80 ± 0.12		−0.81 ± 0.12[b]
av potential energy		−0.92 ± 0.12		−0.91 ± 0.16[b]

[a] All energies are in kcal/mol and distances are in Å. The values in parentheses are van der Waals, electrostatic and desolvation contributions, in this order. Two values are given in each category, one for each helix. The C_α group includes amide N, amide H, α carbon, and α hydrogen. The O group includes the carbonyl carbon and oxygen atoms. The O_γ group includes the γ oxygen, γ hydrogen, β carbon, and β hydrogen. [b] The V84-C_α···T87-O_γ contact is omitted from this average.

Table 3: Interhelical Distances between C_α and O or O_γ Atoms in the Energy-Minimized GpA Structures, and the Interaction Energy between the NHC_α and CO or $C_\beta O_\gamma H_\gamma$ or $C_\beta H_\beta O_\gamma H_\gamma$ Groups in All Solution NMR Models[a]

	all 20 solution NMR structures (IMM1 model)		all 20 solution NMR structures (IMM1-P22 model)	
donor:acceptor	$R_{C_\alpha\cdots O}$	total energy (vdw, ele, solv)	$R_{C_\alpha\cdots O}$	total energy (vdw, ele, solv)
G79-C_α:I76-O	4.41 ± 0.26	−0.26 (−0.19, −0.12, +0.05) ± 0.06	4.51 ± 0.34	−0.49 (−0.18, −0.35, +0.05) ± 0.18
	4.42 ± 0.29	−0.25 (−0.19, −0.11, +0.05) ± 0.05	4.49 ± 0.30	−0.48 (−0.19, −0.35, +0.05) ± 0.14
V80-C_α:G79-O	3.44 ± 0.09	−0.23 (−0.25, −0.12, +0.15) ± 0.05	3.50 ± 0.15	−0.93 (−0.22, −0.86, +0.15) ± 0.10
	3.46 ± 0.08	−0.21 (−0.21, −0.15, +0.15) ± 0.07	3.50 ± 0.14	−0.93 (−0.20, −0.88, +0.15) ± 0.10
G83-C_α:V80-O	3.75 ± 0.15	−0.41 (−0.33, −0.17, +0.09) ± 0.03	3.79 ± 0.19	−0.92 (−0.35, −0.66, +0.09) ± 0.13
	3.75 ± 0.15	−0.40 (−0.33, −0.17, +0.09) ± 0.03	3.78 ± 0.21	−0.92 (−0.34, −0.66, +0.09) ± 0.14
V84-C_α:T87-O_γ	3.58 ± 0.10	−0.21 (−0.17, −0.15, +0.11) ± 0.03	3.56 ± 0.05	−0.57 (+0.03, −0.72, +0.11) ± 0.05
	3.55 ± 0.09	−0.22 (−0.18, −0.15, +0.11) ± 0.05	3.58 ± 0.06	−0.58 (+0.00, −0.69, +0.11) ± 0.07
av effective energy		−0.27 ± 0.13		−0.73 ± 0.12
av potential energy		−0.38 ± 0.10		−0.83 ± 0.10

[a] All energies are in kcal/mol and distances are in Å. The values in parentheses are van der Waals, electrostatic, and desolvation contributions, in this order. Two values are given in each category, one for each helix. The atoms that are included in the C_α, O, and O_γ groups are the same as in Table 1 and Table 2 for IMM1 and IMM1-P22, respectively.

interaction energies are significantly stronger in the IMM1-P22 model. The effective energies range from −0.61 to −0.92 kcal/mol, with an average of −0.80 kcal/mol.

Solution NMR has produced 20 models for GpA in detergent consistent with the NOE constraints. Model 19 exhibits the shortest distance for the G79-C_α···I76-O hydrogen bond and is likely better than the others. However, to avoid any bias, we repeated the calculations on all 20 models. The average (G79)C_α···O(I76) distance in all 20 models is ∼4.5 Å. Table 3 shows the average C_α···O distance and the average energy of each hydrogen bonding pair in the 20 NMR models. As expected, the average G79-C_α··· I76-O energy over all models is smaller in magnitude for both IMM1 and IMM1-p22 functions. However, it is still substantial. The average effective energy of all C_α−H···O hydrogen bonds in the IMM1 model is −0.27 kcal/mol,

Table 4: Interaction Energy between the C_α and O Groups in the Energy-Minimized GpA-Like[b] Motif Structures[a]

donor:acceptor	IMM1 model total (vdw, ele, solv)	IMM1-P22 model total (vdw, ele, solv)	vacuum total (vdw, ele)
	Glycerol Facilitator (1fx8.pdb): Helix 1−Helix 4		
F15-C_α:S92-O	−0.22 (−0.27, −0.07, +0.12)	−0.92 (−0.30, −0.73, +0.12)	−1.36 (−0.25, −1.11)
G19-C_α:G96-O	−0.39 (−0.45, −0.12, +0.18)	−1.00 (−0.52, −0.70, +0.21)	−1.69 (−0.47, −1.22)
Q93-C_α:E14-O	−0.19 (−0.18, −0.06, +0.05)	−0.42 (−0.20, −0.27, +0.06)	−0.99 (−0.28, −0.71)
Q93-C_α:T18-O_γ	−0.15 (−0.14, −0.13, +0.12)	−0.34 (−0.10, −0.37, +0.14)	−0.75 (−0.05, −0.69)
G96-C_α:F15-O	−0.39 (−0.29, −0.20, +0.10)	−0.98 (−0.33, −0.77, +0.11)	−1.18 (−0.24, −0.94)
	Glycerol Facilitator (1fx8.pdb): Helix 5−Helix 8		
P240-C_α:T156-O_γ	−0.05 (−0.09, −0.14, +0.18)	−0.03 (−0.10, −0.11, +0.18)	−0.23 (+0.11, −0.34)
G243-C_α:M153-O	−0.23 (−0.13, −0.22, +0.12)	−0.84 (−0.08, −0.90, +0.14)	−0.87 (+0.20, −1.07)
A244-C_α:T156-O	−0.19 (−0.17, −0.08, +0.06)	−0.28 (−0.19, −0.16, +0.07)	−0.73 (−0.21, −0.52)
A157-C_α:G243-O	+0.05 (−0.01, −0.11, +0.16)	−0.73 (−0.03, −0.88, +0.18)	−1.11 (+0.22, −1.33)
G247-C_α:A157-O	−0.30 (−0.24, −0.17, +0.12)	−0.84 (−0.38, −0.58, +0.12)	−1.12 (−0.30, −0.82)
G161-C_α:G247-O	−0.39 (−0.44, −0.04, +0.09)	−0.69 (−0.44, −0.35, +0.10)	−1.27 (−0.51, −0.76)
	Calcium ATPase (1eul.pdb): Helix 5−Helix 7		
R762-C_α:Y837-O_η	−0.29 (−0.36, −0.05, +0.12)	0.05 (−0.39, +0.33, +0.11)	−0.58 (−0.38, −0.20)
S766-C_α:Y837-O	−0.22 (−0.20, −0.08, +0.06)	−0.53 (−0.23, −0.37, +0.07)	−0.97 (−0.25, −0.72)
G770-C_α:G841-O	−0.31 (−0.30, −0.14, +0.13)	−0.92 (−0.41, −0.65, +0.13)	−0.78 (−0.43, −0.35)
C774-C_α:G845-O	−0.18 (−0.23, −0.08, +0.12)	−0.35 (−0.24, −0.24, +0.13)	−0.87 (−0.27, −0.60)
G841-C_α:S766-O	−0.30 (−0.31, −0.12, +0.12)	−0.81 (−0.30, −0.65, +0.14)	−0.64 (−0.40, −0.23)
G845-C_α:G770-O	−0.37 (−0.40, −0.08, +0.11)	−0.93 (−0.46, −0.59, +0.12)	−1.57 (−0.41, −1.16)
av effective energy	−0.24 ± 0.12	−0.62 ± 0.34	
av potential energy	−0.36 ± 0.12	−0.75 ± 0.34	−0.98 ± 0.37

[a] All energies are in kcal/mol. The values in parentheses are van der Waals, electrostatic, and desolvation contributions, respectively. The atoms that are included in the C_α, O, and O_γ groups are the same as in Table 1 and Table 2 for IMM1 and IMM1-P22, respectively. The atoms that are included in the O_η group are the η oxygen, η hydrogen and ζ carbon. [b] A right-handed interhelical contact where at least one of the interacting helices has a GXXXG or SXXXG motif at the contact point (6).

which is identical to the average effective energy for model 19 (Table 1). In the IMM1-P22 model, except for G79-C_α···I76-O, all interaction energies are very close to that of model 19.

The experimental value −0.88 kcal/mol obtained by Arbely and Arkin (*21*), as well as the quantum mechanical calculations (*12, 13*), correspond to a potential energy, rather than an effective energy. Thus, the desolvation contribution should not be included when comparing to this experiment. With IMM1-P22 the potential energies range from −0.72 to −1.03, with an average of −0.91 kcal/mol for all pairs listed in Table 2, excluding the V84-C_α···T87-O_γ interaction in the solid-state structure. For Gly79-C_α−H···O.Ile76, the hydrogen bond examined by Arbely and Arkin (*21*), the potential energy is −0.71 kcal/mol in the solid-state structure, −0.90 kcal/mol in model 19, and −0.54 kcal/mol average in all models of the solution NMR structure. These values are in good agreement with the FTIR experiment.

The interaction energies between C_α and O_γ, or O_η groups in other GpA-like motifs were calculated with IMM1 and IMM1-P22 and also in vacuum (Table 4). We considered the interactions between helices 1 and 4, helices 5 and 8 from the X-ray structure of glycerol facilitator, and helices 5 and 7 from the X-ray structure of calcium ATPase. In all cases, the effective interaction energies are negative with favorable electrostatic contributions, except for A157-C_α···G243-O in IMM1 and R762-C_α···Y837-O_η in IMM1-P22. The effective interaction energies range from +0.05 to −0.39 kcal/mol for IMM1, +0.05 to −1.00 kcal/mol for IMM1-P22 and −0.23 to −1.69 kcal/mol in vacuum. The average effective interaction energy of all hydrogen bonds in Table 4 is −0.62 kcal/mol for IMM1-P22. The average potential energy (excluding desolvation) of all the hydrogen bonds in IMM1-P22 is −0.75 kcal/mol. The average potential energy in vacuum is −0.98 kcal/mol. It is worth noting that C_α−H···O hydrogen bonds with Glycine as donor tend to have larger energy than the rest. The average interaction energy of C_α−H···O hydrogen bonds in Table 4 that have hydrogen donors from glycine is −0.88 kcal/mol, but the same for hydrogen donors from other residues is −0.39 kcal/mol. Presumably glycine, because of its small size, can allow better access of electronegative atoms to one of its two H^α atoms.

Interaction energies for putative C_α−H···O hydrogen bonds in bR were also calculated (Table 5). These interactions can be positive or negative. With all three energy functions, the effective interaction energy for A51-C_α···T24-O_γ (the hydrogen bond eliminated in the mutagenesis experiment of Bowie and co-workers (*22*)), is positive, mainly due to large, positive electrostatic energy. The same is true for the D212-C_α···Y185-O_η interaction. In some cases the interaction energy is favorable due to van der Waals energy contribution, unlike GpA where most of the favorable interaction energy arises from electrostatics in IMM1-P22. The average effective interaction energy of the six contacts is positive in IMM1 and slightly negative in IMM1-P22. These calculations show that some C_α−H···O interactions can be stabilizing and others destabilizing.

To investigate the origin of the difference in C_α−H···O hydrogen bond energy between GpA and bR, Figure 1 shows the configuration of hydrogen bonding residues G83 and V80 in the solid-state structure of GpA; and T24 and A51 in the crystal structure of bR. One can see that the NHC_α group is oriented differently with respect to the O in the two proteins. In GpA the O is closer to the C_α (3.63 Å) than to the neighboring backbone N (4.42 Å), but in bR, the O_γ of Thr24 is closer to the N (3.26 Å) than to the C_α (3.48 Å). In

Table 5: Distances and Interaction Energies between C_α and O_γ or O_η Atoms in the Energy-Minimized bR Structure[a]

	IMMI model		IMMI-P22 model		vacuum	
donor:acceptor	$R_{C_\alpha\cdots O}$	total (vdw, ele, solv)	$R_{C_\alpha\cdots O}$	total (vdw, ele, solv)	$R_{C_\alpha\cdots O}$	total (vdw, ele, solv)
A51-C_α:T24-O_γ	3.56	+0.15 (−0.41, +0.38, +0.17)	3.53	+0.32 (−0.23, +0.38, +0.17)	3.60	+0.11 (−0.30, +0.41, +0.0)
W86-C_α:T89-O_γ	3.19	+0.51 (+0.25, +0.09, +0.17)	3.46	−0.11 (+0.17, −0.43, +0.14)	3.38	−1.22 (+0.30, −1.53, +0.0)
M118-C_α:S141-O_γ	3.98	−0.17 (−0.26, −0.00, +0.09)	3.88	−0.19 (−0.27, −0.01, +0.10)	3.85	−0.55 (−0.28, −0.27, +0.0)
E166-C_α:S169-O_γ	3.38	+0.58 (−0.13, +0.24, +0.47)	3.59	−0.37 (−0.23, −0.58, +0.44)	3.56	−1.80 (−0.26, −1.54, +0.0)
M209-C_α:Y57-O_η	3.78	−0.24 (−0.39, +0.02, +0.13)	3.62	−0.27 (−0.35, −0.04, +0.13)	3.49	−0.52 (−0.28, −0.24, +0.0)
D212-C_α:Y185-O_η	3.42	+0.05 (−0.36, +0.25, +0.16)	3.51	0.39 (−0.25, +0.50, +0.15)	3.29	+1.52 (+0.32, +1.20, +0.0)
av effective energy		+0.15 ± 0.34		−0.04 ± 0.32		
av potential energy		−0.05 ± 0.26		−0.23 ± 0.38		−0.41 ± 0.89

[a] All energies are in kcal/mol and distances are in Å. These calculations are based on 1PY6.pdb. Hydrogens were built with HBUILD and the energy was minimized for 300 ABNR steps using IMM1, IMM1-P22 or vacuum. The values in parentheses are van der Waals, electrostatic, and desolvation contributions, in this order. The atoms included in the C_α, O_γ, and O_η are the same as in Table 4.

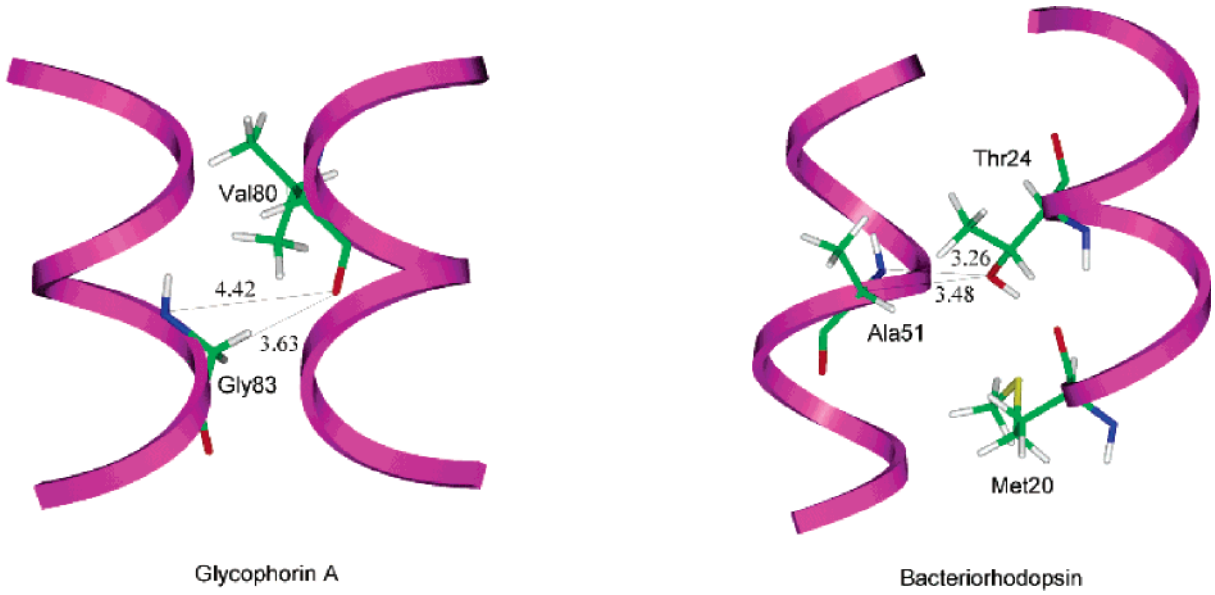

FIGURE 1: C_α−H···O hydrogen bonds and distances between C_α···O or O_γ and N···O or O_γ in the solution NMR structure of Glycophorin A(F78−I85) and the crystal structure of Bacteriorhodopsin (M20-L28 and T47-T55). Val80 and Gly83 in GpA and Thr24, Ala51 and Met20 in bR, are shown in sticks. Nitrogen, carbon, oxygen and hydrogen atoms are shown in blue, green, red and gray, respectively.

addition, the C_α−H···O_γ angle deviates significantly from linearity. This nonideal T24-A51 interaction is dictated by the constraints of protein structure. Since the partial charge on N is negative, the repulsion between N and O dominates over the attraction between O and C_α in the bR T24-A51 interaction. This observation extends to the other hydrogen bonding pairs studied above. Average distances between the O···N and O···C_α atoms for all the hydrogen bonding pairs in the GpA structures (Table 1) are ∼4.3 and ∼3.5 Å, respectively. For the M118-C_α···S141-O_γ and M209-C_α···O_η-Y57 pairs of bR, whose interaction energy is negative, the O_γ or O_η is closer to the C_α than to N. For the D212-C_α···Y185-O_η pair, whose interaction energy is positive, the opposite is true. In general, positive effective energy hydrogen bonding pairs have the O···N distance shorter than the O···C_α, whereas negative effective energy hydrogen bonding pairs have the opposite.

In GpA most of the hydrogen bond acceptors are carbonyl oxygen atoms, whereas in bR all the h-bond acceptors are hydroxyl oxygen atoms. Thus, the difference in strength between the GpA and bR hydrogen bonds may, to some extent, be due to the different chemical properties of CO and OH. To clarify this issue we "mutated" the −$C_\beta H_\beta$−$O_\gamma H_\gamma$ of threonine, C_β−$O_\gamma H_\gamma$ of serine, and C_ζ−$O_\eta H_\eta$ of tyrosine to −CO, preserving the geometry around the C atom. Table 6 shows the effective interaction energy between the C_α donor group and −OH or −CO acceptor groups in bR after 300 steps of ABNR minimization in the IMM1-P22 model. Upon substitution of OH by CO, the energy increases or decreases depending on the position of the H with respect to the C_α and N. If the H is closer to C_α, its elimination makes the interaction more negative or less positive. The opposite is true if H is closer to N. In conclusion, the C_α−H···O interaction is mainly determined by the position of the acceptor O atom but is also modulated by the position of the H atom in the case of OH.

To examine whether the calculated C_α···O hydrogen bond energy depends strongly on its position in the membrane we estimated the energy of C_α···O hydrogen bonds while moving the protein along the *z*-axis in both directions. Figure 2 shows the energy of the four C_α···O hydrogen bonds in GpA as a function of the *z* coordinate of the center of mass of the hydrogen bond donor and acceptor group atoms. Arrows in the figure indicate the original position of these bonds in the membrane. One can notice that inside the membrane, especially within ±10 Å from the center, the

Table 6: Interaction Energy between the Hydrogen Bond Donor C_α and the Hydrogen Bond Acceptor −OH or −CO[b] Groups in bR Using the IMM1-P22 Model[a]

donor:acceptor	−OH acceptor total (vdw, ele, solv)	−CO[b] acceptor total (vdw, ele, solv)
A51-C_α:T24-$O_\gamma H_\gamma$/CO[b]	+0.32 (−0.23, +0.38, +0.17)	+0.47 (−0.32, +0.61, +0.18)
W86-C_α:T89-$O_\gamma H_\gamma$/CO[b]	−0.11 (+0.17, −0.43, +0.14)	−0.32 (+0.12, −0.61, +0.16)
M118-C_α:S141-$O_\gamma H_\gamma$/CO[b]	−0.19 (−0.27, −0.01, +0.10)	−0.35 (−0.20, −0.24, +0.09)
E166-C_α:S169-$O_\gamma H_\gamma$/CO[b]	−0.37 (−0.23, −0.58, +0.44)	−0.43 (−0.08, −0.76, +0.42)
M209-C_α:Y57-$O_\eta H_\eta$/CO[b]	−0.27 (−0.35, −0.04, +0.13)	−0.19 (−0.32, +0.00, +0.12)
D212-C_α:Y185-$O_\eta H_\eta$/CO[b]	+0.39 (−0.25, +0.50, +0.15)	+0.05 (−0.34, +0.24, +0.16)

[a] All energies are in kcal/mol. The values in parentheses are van der Waals, electrostatic, and desolvation contributions, in this order. The atoms included in the C_α, O, O_γ, and O_η are the same as in Table 4. [b] CO from the mutant structures, where the C−OH was replaced by C═O.

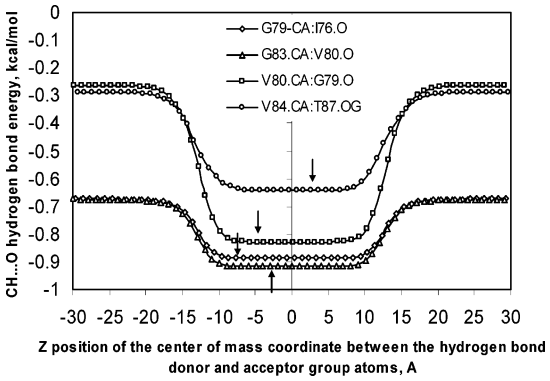

FIGURE 2: C_α−H···O hydrogen bond energy as a function of the position of the hydrogen bond along the *z*-axis of the membrane. Arrows indicate the original position of the hydrogen bonds in the membrane along the *z*-axis.

hydrogen bond energy remains the same. As the position of hydrogen bond approaches the boundary of the membrane, its strength decreases until it reaches another constant value corresponding to pure implicit water. Therefore, our estimated hydrogen bond energy is not sensitive to small displacements along the *z* axis.

To test the dynamic stability of these hydrogen bonds in GpA and bR we performed 1 ns Nose-Hoover molecular dynamics simulations using the IMM1-P22 model. We used the X-ray structure for bR and the solid-state NMR structure for GpA. To include the effect of polar residues outside the membrane, we considered a 29-residue fragment of GpA. The final structures after 1 ns simulation were again energy minimized for 300 steps by the ABNR method. In the case of GpA, except for V84-C_α···T87-O_γ, the interaction energies of all the C_α−H···O contacts after 1ns dynamics and minimization were found to remain close to that of the initial energy minimized structure (averages over the two helices are −0.60 vs −0.65, −0.87 vs −0.88, −0.70 vs −0.89 and −0.54 vs +0.61 kcal/mol for G79-C_α:I76-O, V80-C_α:G79-O, G83-C_α:V80-O, and V84-C_α:T87-O_γ interactions, respectively). During the dynamics, T87 switched from an interhelical hydrogen bond with the CO of V84 to an intrahelical hydrogen bond with the CO of G83. This reversed the sign of the V84-C_α···T87-O_γ interaction. In the case of bR, the interaction energies of C_α−H···O contacts in A51-T24, W86-T89, M118-S141, E166-S169, M209-Y57, and D212-Y185 after the simulation are +0.21, −0.78, −0.14, −0.18, −0.49, and +0.01 kcal/mol, respectively. The W86-C_α···T89-O_γ and M209-C_α···Y57-O_η contacts remain within hydrogen bonding distance, whereas the C_α−O distance for the other contacts increases.

CONCLUSIONS

Interaction energy calculations between the neutral groups that contain the hydrogen donor and acceptor were used in this work to estimate the strength of C_α−H···O hydrogen bonds. The united-atom membrane model (IMM1) gave an average effective energy of ∼−0.29 kcal/mol for a C_α−H···O hydrogen bond in GpA (Table 1). The explicit hydrogen membrane model IMM1-P22, which is more realistic, gave an average effective energy of −0.77 kcal/mol and an average potential energy of −0.87 kcal/mol (Tables 2 and 3). In general, C_α−H···O hydrogen bonds in GpA-like motifs were found to stabilize the protein with an average effective energy of −0.62 kcal/mol and an average potential energy of −0.75 kcal/mol. These values are similar to a recent experimental estimate of −0.88 kcal/mol for a similar interaction in GpA (*21*). The sum of all the effective energies of C_α−H···O hydrogen bonds in all 20 solution NMR structures with IMM1-P22 is −5.82 kcal/mol. Thus, the six weak C_α−H···O hydrogen bonds in GpA make a substantial contribution to the stability of the dimer, although not the largest one (the average total effective interaction energy between the two helices over the 20 models is −24.93 kcal/mol; −34.49 kcal/mol van der Waals, −1.39 kcal/mol electrostatic, and +10.94 kcal/mol desolvation).

Both the FTIR experiment and our calculations give values much lower than the ab initio estimates for the energy of C_α−H···O bonds (2.5−3.0 kcal/mol). One of the reasons could be that the quantum mechanical calculations are done in the gas phase. However, we found a maximum energy of 1.8 kcal/mol for this bond in vacuum. The lack of explicit polarization in our molecular mechanics energy functions could lead to some underestimation of the interaction energy. Second, the protein environment is completely absent in the quantum calculations, and its effect on the stability of C_α−H···O bond is completely neglected. The small model compounds used in the quantum mechanical calculations do not experience the constraints that similar groups have in the context of a large helical fragment and have more freedom to rearrange so as to optimize their interaction. For example, in the quantum mechanical calculations (*12, 13*), the C···O distances in the optimized geometries of the model molecules used range from 3.29 to 3.33 Å. However, in the NMR structures of GpA, the average C_α···O distance is 3.6 Å.

This study showed that in bR the hydrogen bond between the C_α−H of Ala51 and O_γ of Thr24 does not stabilize the protein. This is consistent with the experimental observation by Yohannan et al. (*22*) that replacement of Thr24 by Ala or Ser stabilizes the protein. The origin of the difference in

stability of C_α−H···O hydrogen bonds in GpA and bR is the position of the O atom with respect to the C_α and N atom of its binding partner. In GpA the O is closer to C_α than to N, whereas the opposite is true in bR. Though this specific hydrogen bond in bR is not stabilizing, other C_α−H···O hydrogen bonds were found to show some weak stabilizing capacity. The position of the hydrogen of the hydroxyl group with respect to C_α and N also influences the strength of the interaction.

ACKNOWLEDGMENT

We thank Iban Ubarretxena-Belandia for valuable discussions. Jin-Ming Zhang performed some initial interaction energy calculations on GpA.

REFERENCES

1. Popot, J. L., and Engelman, D. M. (1990) Membrane protein folding and oligomerization: the two-stage model, *Biochemistry 29*, 4031−4037.
2. Popot, J. L., and Engelman, D. M. (2000) Helical membrane protein folding, stability, and evolution, *Annu. Rev. Biochem 69*, 881−922.
3. Choma, C., Gratkowski, H., Lear, J. D., and DeGrado, W. F. (2000) Asparagine-mediated self-association of a model transmembrane helix, *Nat. Struct. Biol. 7*, 161−166.
4. Zhou, F. X., Cocco, M. J., Russ, W. P., Brunger, A. T., and Engelman, D. M. (2000) Interhelical hydrogen bonding drive strong interactions in membrane proteins, *Nat. Struct. Biol. 7*, 154−160.
5. Zhou, F. X., Merianos, H. J., Brunger, A. T., and Engelman, D. M. (2001) Polar residues drive association of polyleucine transmembrane helices, *Proc. Natl. Acad. Sci. U.S.A. 98*, 2250−2255.
6. Senes, A., Ubarretxena-Belandia, I., and Engelman, D. M. (2001) The C_α−H···O hydrogen bond: A determinant of stability and specificity in transmembrane helix interactions, *Proc. Natl. Acad. Sci. U.S.A. 98*, 9056−9061.
7. Derewenda, Z. S., Derewenda, U., and Kobos, P. M. (1994) (His)-C^ϵ−H···O═C< Hydrogen bond in the active sites of Serine Hydrolases, *J. Mol. Biol. 241*, 83−93.
8. Derewenda, Z. S., Lee, L., and Derewenda, U. (1995) The occurrence of C−H···O hydrogen bonds in proteins, *J. Mol. Biol. 252*, 248−262.
9. Bella, J., and Berman, H. M. (1996) Crystallographic evidence for C_α−H···O═C hydrogen bonds in collagen triple helix, *J. Mol. Biol. 264*, 734−742.
10. Fabiola, G. F., Krishnaswamy, S., Nagarajan, V., and Pattabhi, V. (1997) C−H···O Hydrogen bonds in β-sheets, *Acta Crystallog. Sect. D 53*, 316−320.
11. Taylor, R., and Kennard, O. (1982) Crystallographic evidence for the existence of C−H···O, C−H···N and C−H···Cl hydrogen bonds, *J. Am. Chem. Soc. 104*, 5063−5070.
12. Vargas, R., Garza, J., Dixon, D. A., and Hay, B. P. (2000) How strong is the C^α−H···O═C hydrogen bond?, *J. Am. Chem. Soc. 122*, 4750−4755.
13. Scheiner, S., Kar, T., and Gu, Y. (2001) Strength of the C^αH···O hydrogen bond of amino acid residues, *J. Biol. Chem. 276*, 9832−9837.
14. Gu, Y., Kar, T., and Scheiner, S. (1999) Fundamental properties of the CH···O interaction: Is it a true hydrogen bond?, *J. Am. Chem. Soc. 121*, 9411−9422.
15. MacKenzie, K. R., Prestegard, J. H., and Engelman, D. M. (1997) A transmembrane helix dimer: Structure and implications, *Science 276*, 131−133.
16. Smith, S. O., Smith, C., Shekar, S., Peersen, O., Ziliox, M., and Aimoto, S. (2002) Transmembrane interactions in the activation of the Neu receptor tyrosine kinase, *Biochemistry 41*, 9321−9332.
17. Fu, D., Libson, A., Miercke, L. J., Weitzman, C., Nollert, P., Krucinski, J., and Stroud, R. M. (2000) Structure of a glycerol conducting channel and the basis for its selectivity, *Science 290*, 481−486.
18. Toyoshima, C., Nakasako, M., Nomura, H., and Ogawa, H. (2000) Crystal structure of the calcium pump of sarcoplasmic reticulum at 2.6A resolution, *Nature 405*, 647−655.
19. Jordan, P., Fromme, P., Witt, H. T., Klukas, O., Saenger, W., and Krauss, N. (2001) Three-dimensional structure of cyanobacterial photosystem I at 2.5A resolution, *Nature 411*, 909−917.
20. Kleiger, G., Grothe, R., Mallick, P., and Eisenberg, D. (2002) GXXXG and AXXXA: Common α-helical interaction motifs in proteins, particularly in extremophiles, *Biochemistry 41*, 5990−5997.
21. Arbely, E., and Arkin, I. T. (2004) Experimental measurement of the strength of a C^α−H···O bond in a lipid bilayer, *J. Am. Chem. Soc. 126*, 5362−5363.
22. Yohannan, S., Faham, S., Yang, D., Grosfeld, D., Chamberlain, A. K., and Bowie, J. U. (2004) A C^α−H···O hydrogen bond in a membrane protein is not stabilizing, *J. Am. Chem. Soc. 126*, 2284−2285.
23. Lazaridis, T. (2003) Effective energy function for proteins in lipid membranes, *Proteins 52*, 176−192.
24. Lazaridis, T., and Karplus, M. (1997) "New view" of protein folding reconciled with the old through multiple unfolding simulations, *Science 278*, 1928−1931.
25. Lazaridis, T., and Karplus, M. (1999) Effective energy function for proteins in solution, *Proteins 35*, 133−152.
26. MacKerell, J., A. D., Bashford, D., Bellott, M., Dunbrack, J., R. L., Evanseck, J. D., Field, M. J., Fischer, S., Gao, J., Guo, H., Ha, S., Joseph-McCarthy, D., Kuchnir, L., Kuczera, K., Lau, F. T. K., Mattos, C., Michnick, S., Ngo, T. N., D. T., Prodhom, B., Reiher, I., W. E, Roux, B., Schlenkrich, M., Smith, J. C., Stote, R., Straub, J., Watanable, M., Wiorkiewicz-Kuczera, J., Yin, D., and Karplus, M. (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins, *J. Phys. Chem. B 102*, 3586−3616.
27. Smith, S. O., Eilers, M., Song, D., Crocker, E., Ying, W. W., Groesbeek, M., Metz, G., Ziliox, M., and Aimoto, S. (2002) Implications of threonine hydrogen bonding in the glycophorin A transmembrane helix dimer, *Biophys. J. 82*, 2476−2486.
28. Lazaridis, T., Archontis, G., and Karplus, M. (1995) The enthalpic contribution to protein stability: insights from atom-based calculations and statistical mechanics, *Adv. Prot. Chem. 47*, 231−306.

BI048065S