doi:10.1016/j.jmb.2004.09.011 J. Mol. Biol. (2004) 343, 1487–1497

JMB

Available online at www.sciencedirect.com

SCIENCE DIRECT®

ELSEVIER

Complex Interactions at the Helix–Helix Interface Stabilize the Glycophorin A Transmembrane Dimer

Abigail K. Doura and Karen G. Fleming*

T.C. Jenkins Department of Biophysics, Johns Hopkins University, 3400 North Charles Street, Baltimore, MD 21218 USA

**Corresponding author*

To explore the residue interactions in the glycophorin A dimerization motif, an alanine scan double mutant analysis at the helix–helix interface was carried out. These data reveal a combination of additive and coupled effects. The majority of the double mutants are found to be equally or slightly more stable than would be predicted by the sum of the energetic cost of the single-point mutants. The proximity of the mutated sites is not related to the presence of coupling between those sites. Previous studies reveal that a single face of the glycophorin A monomer contains a specific glycine-containing motif (GxxxG) that is thought to be a driving force for the association of transmembrane helices. Double mutant cycles suggest that the relationship of the GxxxG motif to the remainder of the helix–helix interface is complex. Sequences containing mutations that abolish the GxxxG motif retain an ability to dimerize, while a sequence containing a GxxxG motif appears unable to form dimers. The energetic effects of weakly coupled and additive double mutants can be explained by changes in van der Waals interactions at the dimer interface. These results emphasize the fact that the sequence context of the dimer interface modulates the strength of the glycophorin A GxxxG-mediated transmembrane dimerization reaction.

Keywords: glycophorin A; dimerization motif; helix–helix interface; transmembrane helices; GxxxG

Introduction

The specificity of protein–protein interactions is critical to the establishment of native protein structure and to the assembly of protein complexes. A common method to study the interactions that specify native protein folds is site-specific mutagenesis.[1] The interpretation of structural effects of mutation may be simplified in membrane proteins. As compared to soluble proteins, membrane proteins exist in a constrained environment in which the association of helices can be considered independently from the formation of secondary structure elements.[2] Thus, an advantage to the study of helical membrane protein folding is that the unfolded state may be considered a stable helix monomer, whereas the unfolded state in soluble proteins explores a greater conformational space.[3,4] By comparing the structural and thermodynamic effects of mutants in soluble and membrane proteins, it may be possible to elucidate whether similar principles determine the native fold in such different environments.

The human erythrocyte protein glycophorin A (GpA) contains a single transmembrane domain that associates to form a symmetric homodimer. Multiple thermodynamic studies using qualitative and quantitative techniques have been carried out to probe the GpATM dimerization reaction.[5–7] These experiments resulted in a prediction of a sequence motif involved in dimerization,[8] which was found at the dimer interface when the NMR structure was solved.[9] Initial mutational analysis suggested that the glycine residues within the glycophorin motif were the most critical residues for dimerization.[10] Genetic and statistical screens lead to the hypothesis that the interactions in the glycine motif (GxxxG) are the primary force driving the helices to associate and that the presence of this motif could mediate strong association in other transmembrane alpha helices.[11,12] In contrast, large-scale mutagenesis has shown that an intact GxxxG

Abbreviations used: GpATM, glycophorin A transmembrane domain; AUC, analytical ultracentrifugation; WT, wild-type.

E-mail address of the corresponding author: karen.fleming@jhu.edu

motif is not necessary for dimerization of the GpATM.[13] These results suggest that the adjacent amino acid residues at a GxxxG dimer interface can be a determinant for the strength of dimerization. Current simple models of helical protein interaction suggest that both packing moments[14] and the presence of small side-chains[15] can be predictors of a helix–helix interface. However, previous thermodynamic studies show that the ability of helices to dimerize is governed by complex principles and involves detailed interactions at the helix–helix interface.[13]

Site-specific mutagenesis has been used extensively in biochemistry to determine the effect of a residue on the structure and function of a protein. The effect of a single-point mutant may be attributed to many interactions between the mutated site and the remaining sequence. Double-mutant cycles were first used to probe the interaction between two residues more directly.[1] Using this method, one can determine whether an interaction between sites exists and quantify the strength of that interaction. Double-mutant cycles have been used in several soluble proteins[16] to probe cooperativity in folding[17] and enzymatic activity.[18] In simple cases, the presence of coupling can be explained by the proximity of the mutated residues.[19] However, coupling can be due to long-range electrostatic interactions[20] established through both direct and indirect pathways.[21] While there is no direct correlation between the distance between mutated sites in the native structure and the energy of coupling in soluble proteins, coupling is more likely to occur when residues are less than 12 Å apart.[22] In addition, some non-additive interactions are believed to be caused by interactions in the denatured state.[23,24] The ensemble of conformations available to an unfolded protein in solution is much greater than that available in a membrane environment. The decreased ambiguity in the denatured state of membrane proteins may lessen the effect of a hydrophobic mutation on the monomeric transmembrane helix as compared to a soluble denatured helix. The intrinsic contribution of the monomer stability to the free energy of association could remain the same in a mutated transmembrane domain. Therefore, in membrane proteins the interpretation of the effect of mutation on the free energy of association can be considered as changes in inter-helical interactions only. The intention of this study is to employ double-mutant cycles to investigate the role of coupling at the helix–helix interface in stabilizing the transmembrane protein glycophorin A.

Results

To better understand the role of coupling in specifying the native fold of a membrane protein, a double-mutant analysis of the GpATM was carried out. The study comprises double alanine mutations at the dimer interface. A total of 21 double alanine mutants were created and analyzed using sedimentation equilibrium analytical ultracentrifugation (AUC). Table 1 shows both the experimentally determined changes in the free energy of association as well as those calculated assuming additivity for each double mutant. Experimentally, all mutations are destabilizing with respect to the wild-type (WT) with the exception of Leu75Ala-Val84Ala. By comparing the free energy of association if additive to the experimentally determined free energy of association, it is possible to quantify the energy of interaction between the mutated sites. This value is referred to as the free energy of coupling (ΔG^{o}_{coup}). Although the vast majority of mutations are destabilizing with respect to the WT, the combination of double mutations can have a stabilizing effect as compared to the sum of single-point mutant losses in free energy. The most stabilizing ΔG^{o}_{Coup} is -2.2 kcal mol^{-1} (1 cal = 4.184 J) for the mutant Leu75Ala-Val84Ala (i.e. the double mutant is more stable than the combination of the single-point mutants). The most destabilizing ΔG^{o}_{Coup} is $+2.6$ kcal mol^{-1} observed for the mutant Leu75Ala-Thr87Ala (i.e. the double mutant is less stable than the combination of the single-point mutants). Ten of the 21 mutants (48%) have coupling free energies that are stabilizing. Eight of the 21 mutants (38%) exhibit an additive ΔG^{o}_{Coup}. The remaining three (14%) have a destabilizing ΔG^{o}_{Coup}.

Most double mutant proteins still dimerize

All GpATM double mutants except one retain the ability to dimerize, and none of the double mutant proteins associates more strongly than the WT protein (Figure 1; Table 1). The most destabilizing

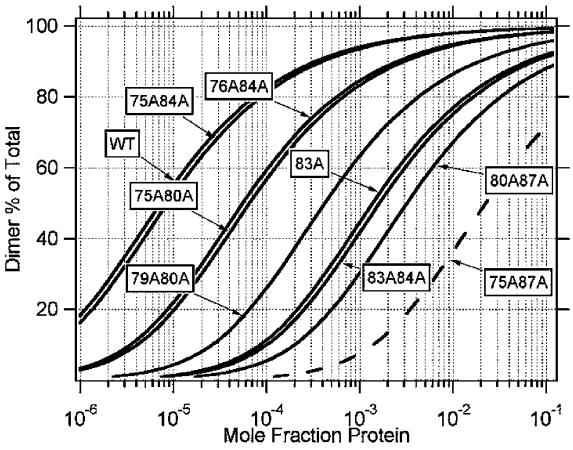

Figure 1. Relative dimer population of selected GpATM mutants. The fraction dimer is plotted *versus* the mole fraction protein for selected mutants. The populations are calculated data, based on the standard-state free-energy values determined by sedimentation equilibrium analytical ultracentrifugation. For reference, the WT and Gly83Ala single-point mutant distributions are shown. The broken line shows the maximum possible dimer distribution for the Leu75Ala-Thr87Ala double mutant.

Table 1. The sequence for each double mutant is shown in column 2. The motif for dimerization is colored in red and sites of substitution are colored in blue

Mutant	Sequence	Additive $\Delta\Delta G^o_{MUT}$ (a)	Experimental $\Delta\Delta G^o_{MUT}$ (b)	ΔG^o_{coup} (c)	Threshold (d)
WT	ITLIIFGVMAGVIGTILLISTGI		0.0 ± 0.1		
75A76A	ITAAIFGVMAGVIGTILLISTGI	3.1 ± 0.2	2.4 ± 0.1	-0.7	± 0.2
75A79A	ITAIIFAVMAGVIGTILLISTGI	3.0 ± 0.2	2.3 ± 0.1	-0.7	± 0.3
75A80A	ITAIIFGAMAGVIGTILLISTGI	1.7 ± 0.2	1.3 ± 0.3	-0.4	± 0.3
75A83A	ITAIIFGVMAAVIGTILLISTGI	4.4 ± 0.2	3.6 ± 0.2	-0.8	± 0.3
75A84A	ITAIIFGVMAGAIGTILLISTGI	2.3 ± 0.2	0.1 ± 0.1	-2.2	± 0.3
75A87A	ITAIIFGVMAGVIGAILLISTGI	2.2 ± 0.2	≥ 4.8	≥ 2.6	
76A79A	ITLAIFAVMAGVIGTILLISTGI	3.5 ± 0.2	2.9 ± 0.2	-0.6	± 0.3
76A80A	ITLAIFGAMAGVIGTILLISTGI	2.2 ± 0.2	2.5 ± 0.3	0.3	± 0.4
76A83A	ITLAIFGVMAAVIGTILLISTGI	4.9 ± 0.2	3.4 ± 0.2	-1.6	± 0.3
76A84A	ITLAIFGVMAGAIGTILLISTGI	2.8 ± 0.2	1.2 ± 0.1	-1.6	± 0.3
76A87A	ITLAIFGVMAGVIGAILLISTGI	2.7 ± 0.2	2.8 ± 0.2	0.1	± 0.2
79A80A	ITLIIFAAMAGVIGTILLISTGI	2.1 ± 0.3	2.4 ± 0.1	0.3	± 0.3
79A83A	ITLIIFAVMAAVIGTILLISTGI	4.8 ± 0.3	3.6 ± 0.3	-1.2	± 0.4
79A84A	ITLIIFAVMAGAIGTILLISTGI	2.7 ± 0.3	3.0 ± 0.2	0.3	± 0.4
79A87A	ITLIIFAVMAGVIGAILLISTGI	2.6 ± 0.2	2.8 ± 0.2	0.2	± 0.3
80A83A	ITLIIFGAMAAVIGTILLISTGI	3.5 ± 0.3	3.5 ± 0.1	0.0	± 0.3
80A84A	ITLIIFGAMAGAIGTILLISTGI	1.4 ± 0.3	2.0 ± 0.2	0.6	± 0.4
80A87A	ITLIIFGAMAGVIGAILLISTGI	1.3 ± 0.2	3.7 ± 0.2	2.4	± 0.3
83A84A	ITLIIFGVMAAAIGTILLISTGI	4.1 ± 0.3	3.2 ± 0.1	-0.9	± 0.3
83A87A	ITLIIFGVMAAVIGAILLISTGI	4.0 ± 0.2	3.3 ± 0.1	-0.7	± 0.3
84A87A	ITLIIFGVMAGAIGAILLISTGI	1.9 ± 0.2	1.0 ± 0.1	-0.9	± 0.3
79L83L	ITLIIFLVMALVIGTILLISTGI	5.8 ± 0.3	4.1 ± 0.2	-1.8	± 0.4

[a] The $\Delta\Delta G^o_{MUT}$ of the double mutant predicted by the sum of the for the single-point mutants.
[b] The $\Delta\Delta G^o_{MUT}$ for the double mutant determined experimentally by sedimentation equilibrium analytical ultracentrifugation.
[c] Calculated using equation (3).
[d] Threshold for additivity is calculated using equation (7).

mutant that associates under experimental conditions, Val80Ala-Thr87Ala, is still 30% dimeric at a protein:detergent mole fraction of 10^{-3}. The double mutant Gly79Ala-Gly83Ala, which contains the two most destabilizing single-point mutations, retains an ability to dimerize even though this double mutant does not contain the GxxxG dimerization motif. The free energy of association for the mutant Leu75Ala-Thr87Ala is inaccessible by our methods, because of an inability to access adequate concentrations of dimer to confirm the equilibrium constant. From the highest observable concentration of protein, the upper limit for the ΔG^o_x is calculated to be ≥ -2.3 kcal mol^{-1}. The mutant Leu75Ala-Thr87Ala is therefore destabilized by at least an additional 0.7 kcal mol^{-1} compared to the most destabilizing mutants for which an accurate ΔG^o_x can be obtained.

The free energy of coupling is stabilizing in most cases

The free energy of coupling can be determined by comparing the free energy of association for the double-mutant protein to the sum of the free energies for the corresponding single-point mutant proteins (Figure 2). A free energy of coupling equal to zero denotes an additive interaction. The majority of mutations exhibit a stabilizing free energy of coupling. Those with the strongest stabilizing free energy of coupling are mutations that occur when the primary site mutated is Leu75 or Ile76. The residue pair Leu75 and Val84 exhibits the most stabilizing interaction. The ΔG^o_{Coup} between Leu75 and Val84 is great enough to result in a WT-like propensity to dimerize in the double mutant (Figure 1). Residue Ile76 exhibits stabilizing coupled effects when doubly mutated with residues Gly83 or Val84. These data demonstrate that in many cases the double substitutions to alanine in the GpATM allow the recovery of favorable interactions, providing for greater overall stability as compared to the sequences that contain single point mutations to alanine.

The pattern of coupling is complex

In soluble proteins, the simplest explanation for thermodynamic coupling is a van der Waals interaction between mutated residues.[19] This simple pattern of coupling can be observed by comparing the proximity of mutated residues to the presence of coupling. Interestingly, in the GpATM there appears to be no relationship between the proximity of mutated residues and their coupling, and the coupling of residues in the GpATM appears to follow no simple pattern. This is visualized in Figure 3, where each structure shows the pattern of coupling between the first substitution and all other secondary substitutions created in the double-mutant proteins. Each schematic represents the coupling for six double mutants for which the

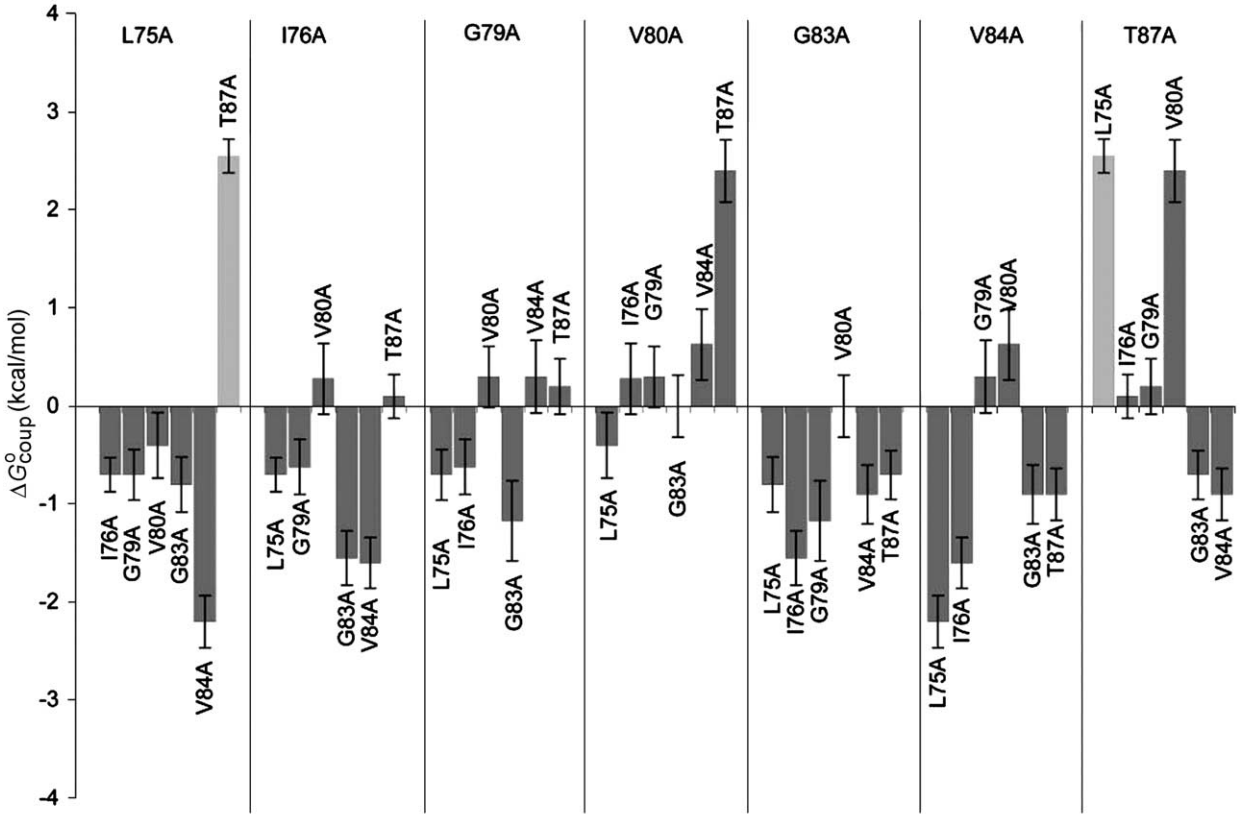

Figure 2. The free energy of coupling for alanine double mutants. Each bar represents the free energy of coupling for a double mutant. The first mutation in each double mutant is noted in the header and the second is the label on the bar. The error bars reflect the standard deviation of each measurement. A free energy of zero indicates additivity. The gray bar shows an estimated minimum for the Leu75Ala-Thr87Ala double mutant. Note that each double mutant is shown twice.

primary substitution is indicated in the text below the model. Notably, residues in van der Waals contact do not exhibit a destabilizing coupling in any of the double mutants created. In contrast, GpATM residues within van der Waals contact in the WT structure have a free energy of coupling that is either additive or slightly stabilizing.

The pattern of coupling for each site along the interface is different from the pattern for all other sites. Leu75 exhibits a coupling profile similar to that of Ile76 with the exception of the strongly destabilizing coupling at Thr87. Val80 exhibits a particularly unique coupling profile. It is additive with the residues to the N-terminal side, yet is coupled in a destabilizing fashion to the C terminus. Val84 (the valine residue adjacent to the second glycine residue in the GxxxG motif) exhibits a coupling profile different from that of Val80. Gly79 and Gly83 exhibit very different coupling profiles; for instance, Gly79 is not coupled to as many sites as Gly83. This distinction may be due to the fact that the helices are packed more tightly at Gly83 than they are at Gly79. The difference between the coupling at the N and C termini with Val80 and the distinct profiles of Gly79 and Gly83 indicate a lack of symmetry in the nature of interactions at the distal ends of the helix and within the GxxxG motif.

The role of sequence motifs as a driving force for dimerization

Alanine single-point mutations created at Gly79 and Gly83 have been shown to maintain a propensity to dimerize. When a double mutant to alanine is created at the GxxxG motif, the GpATM still dimerizes and the ΔG^{o}_{Coup} is stabilizing. It has been proposed that an AxxxA motif can drive dimerization,[12,25] and an AxxxA motif is created by alanine substitution of the glycine residues in the GxxxG motif. The result that the ΔG^{o}_{Coup} is stabilizing for the Gly79Ala-Gly83Ala mutant may support the hypothesis that an AxxxA motif also drives dimerization. This may occur because alanine has a small side-chain that can facilitate a close approach of the helices to maintain strong van der Waals interactions. To address this question, we created a double mutant to leucine at Gly79 and Gly83. The introduction of leucine is expected to create a large steric clash that should not allow backbone interactions at positions 79 and 83. Single-point substitutions to leucine at positions 79 and 83 have been shown to dimerize with a weaker affinity than single-point substitutions to alanine at those sites.[13] Figure 4(B) shows that the data obtained for the Gly79Leu-Gly83Leu mutant is best fit to a monomer–dimer model, demonstrating that a sequence

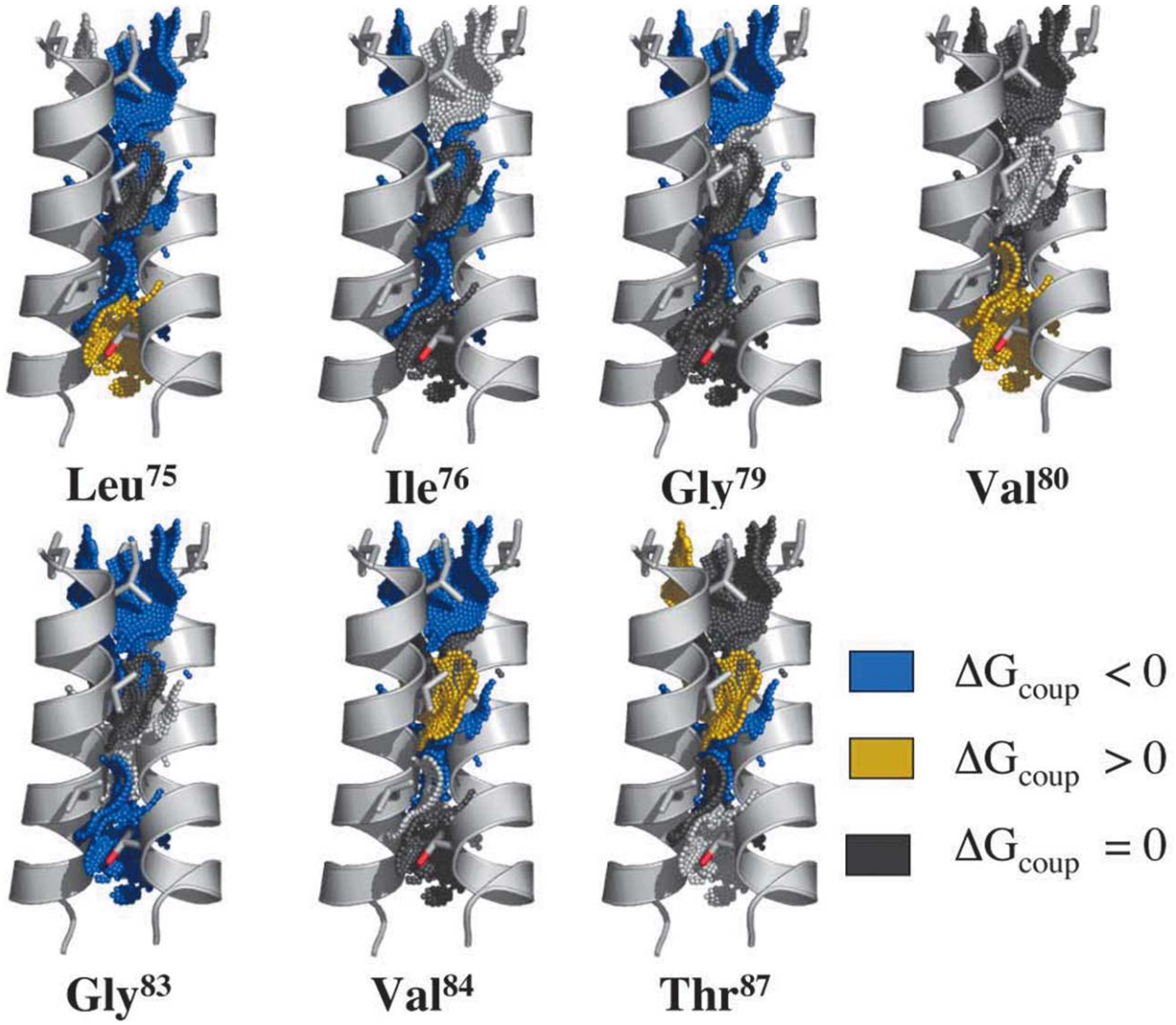

Figure 3. The thermodynamic coupling of sites along the helix–helix interface. Each GpATM schematic structure represents the type of coupling between the indicated primary site and each secondary site along the dimerization interface. These couplings are projected onto the WT structure to compare the pattern of coupling for each site. The dot surface represents the molecular contact surface area of each residue at the dimer interface in the WT structure. The primary site of substitution is colored in light gray and indicated in the text below the schematic. A destabilizing ($\Delta G^{o}_{Coup} > 0$) free energy of coupling is represented by a coloring of the dot surface for the secondary mutations in yellow. A stabilizing free energy of coupling ($\Delta G^{o}_{Coup} < 0$) is represented by a coloring of the dot surface for the secondary mutations in blue. An additive free energy of coupling ($\Delta G^{o}_{Coup} = 0$) is represented by a coloring of the dot surface for the secondary mutations in dark gray.

replacing a GxxxG motif with leucine residues also has the ability to dimerize in the GpATM ($\Delta\Delta G^{o}_{MUT} = +4.1$ kcal mol^{-1}). Notably, the stability of this mutant dimer is much greater than that of the Leu75Ala-Thr87Ala mutant (Figure 4(A)) which still contains the GxxxG dimerization motif.

The energetics of additive and weakly coupled mutants can be predicted by structure-based calculations

Structure-based parameterization using computational models has been used to explain the relationship between structural changes and energetic changes for glycophorin A single-point mutants.[7,13] To determine whether this relationship applies to double mutants, a structure-based calculation was carried out using the WT GpATM NMR structure as the basis for double mutant models. The coefficients determined in the single-point mutant parameterization were used to predict the $\Delta\Delta G^{o}_{MUT}$ for a double mutant-based computational model. Of the 21 double mutants, the association free energies of 15 can be predicted using structure-based calculations (Figure 5) ($R = 0.89$, $p = 5.364 \times 10^{-5}$). All 15 are either additive or coupled weakly. The free energies of association for those double mutants that are coupled strongly are not predicted well by the structure-based calculation. To test whether the predictive value of the parameterization could be improved, an additional parameterization was carried out that included both double-point

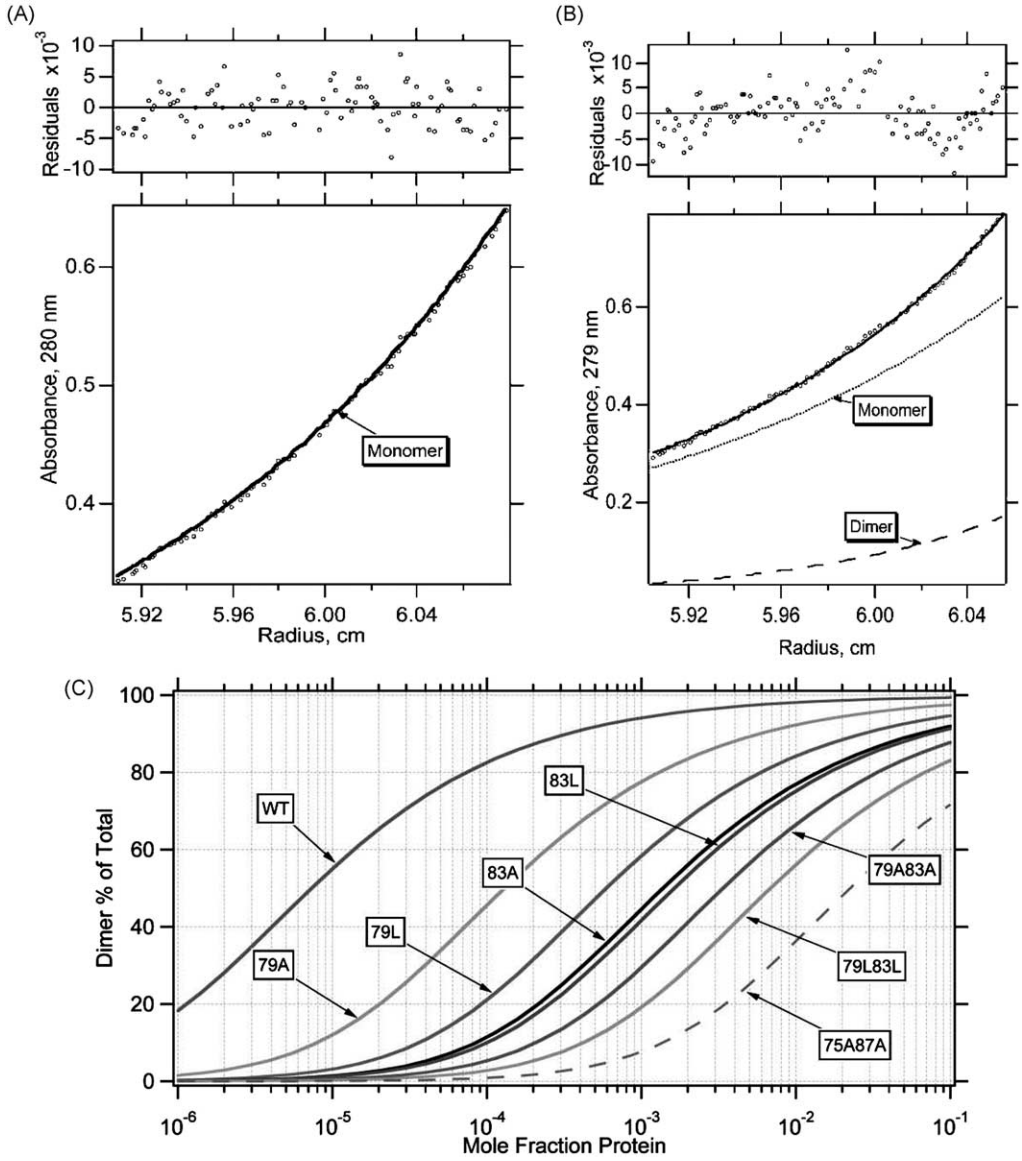

Figure 4. A GxxxG motif is neither necessary nor sufficient for dimerization. (A) and (B) are representative sedimentation equilibrium data for selected mutants. A shows a purely monomeric population for the double mutant Leu75Ala-Thr87Ala and the residuals of the single-species fit of AUC data. (B) shows the monomer dimer fit and the residuals for AUC data of the double mutant Gly79Leu-Gly83Leu. (C) shows the relative populations of the single and double mutants affecting the GxxxG motif as a function of mole fraction protein. These distributions are calculated from the experimental standard-state free-energy values.

and single-point mutant data. There was no significant change in the values of the coefficients or in the correlation of predicted and experimental changes in free energy.

Structurally, coupling may be due to global structural changes in the dimer, an accumulation of small changes at the level of side-chain conformation at the dimer interface, or structural changes in the unfolded state. Under the assumption that small conservative mutations do not affect the monomeric state of a transmembrane helix it can be inferred that coupling in the GpATM is due to either global or local structural changes in the dimer. The modeling protocol used in this study allows for changes in side-chain conformation at the dimer interface but does not allow backbone rearrangements or other large global changes in structure. The inability to use these models to predict strong coupling indicates that the most likely explanation for strong coupling in the GpATM is global structural rearrangements in the dimer structure.

Discussion

GxxxG is neither necessary nor sufficient for dimerization in the GpATM

Double-mutant cycles allow the exploration of the sequence context of the glycophorin A motif in more detail. Using this method, we were able to

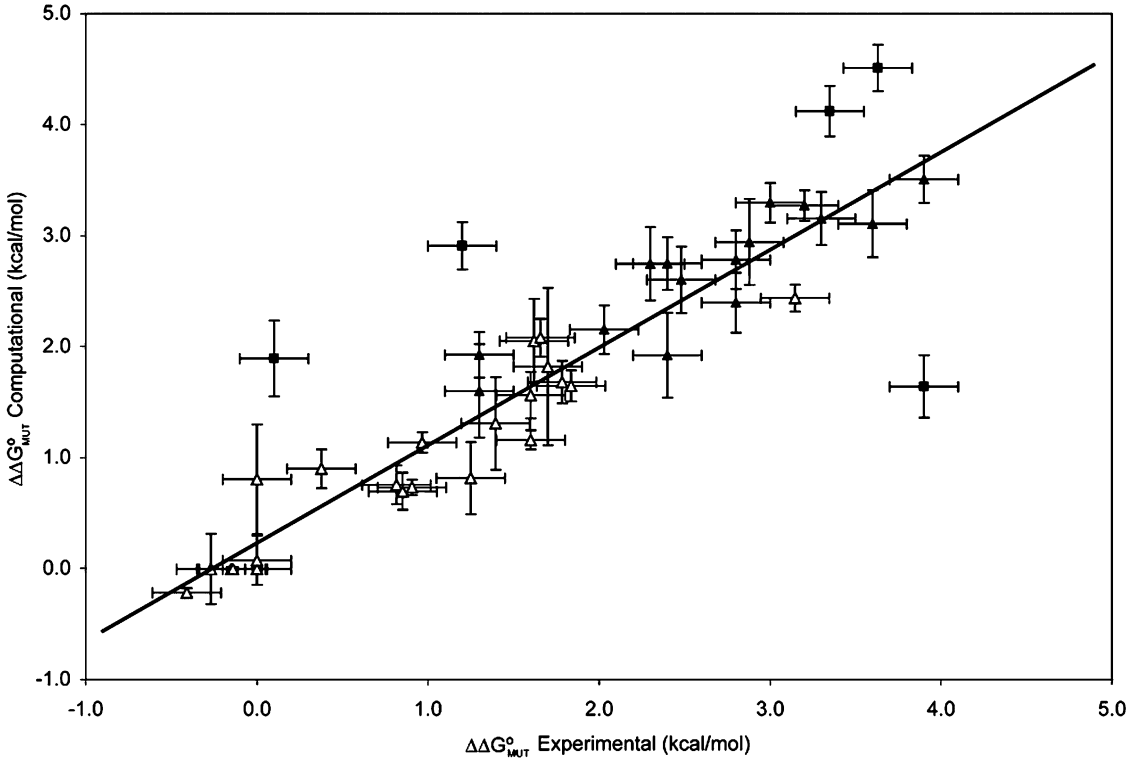

Figure 5. The free energy of association for double mutants can be predicted using structure-based calculations. A comparison between the free energies determined through analysis of structural parameters and those determined experimentally is shown. Single-point mutant free energies are indicated by open triangles. Weakly coupled and additive double-point mutant free energies are indicated by filled triangles. Strongly coupled double-mutant free energies are indicated by filled squares. The trend line indicates the correlation between filled and open triangle data ($R^2 = 0.9209$, slope = 0.88). The *X*-value error bars are the experimental standard deviation. The *Y*-value error bars are based on the conformational heterogeneity in the NMR structure.

determine the degree of coupling at the dimer interface to reveal how sites interact to specify a transmembrane dimer. Several previous studies attempt to develop simple models for helix–helix interactions based on statistics[12] and genetic screens.[11] These studies have postulated that the interactions at the GxxxG motif provide a driving force for helix–helix interactions in a membrane environment. Single-point mutants at Gly79 and Gly83 lead to large disruptions in the propensity of dimerization in the GpATM.[8,10,13,26] Therefore, interactions at these glycine residues have been implicated as the driving force for dimerization. However, our double-mutant analysis has shown that many permutations in the GpATM sequence still dimerize with significant affinity. The single exception in this study is Leu75Ala-Thr87Ala, which is monomeric under accessible experimental conditions. Therefore, the presence of a GxxxG motif is not sufficient for dimerization. Alanine single-point substitutions created at Gly79 and Gly83 have been shown to maintain a propensity to dimerize. When a double mutant to alanine is created at the GxxxG motif there maintains a propensity to dimerize, and the ΔG^{o}_{coup} is stabilizing. The introduction of leucine at these sites should create a large steric clash and would not be expected to allow backbone interactions at positions 79 and 83. Surprisingly, the mutant Gly79Leu-Gly83Leu dimerizes with significant affinity. These data indicate that a GxxxG-like motif is not necessary for dimerization in the GpATM. Because the sedimentation equilibrium method allows a direct determination of mass, our experiments confirm that this LxxxL sequence still forms a defined oligomer (dimer) that must be specified by the remaining amino acid sequence.

Long-range coupling specifies native interactions in the GpATM

Thermodynamic coupling of residues can be explained easily in situations where the coupling of residues is correlated to the proximity of residues in the secondary and tertiary structure of the protein. Mutations at sites that are in direct contact can affect the total stability of the protein by changing multiple components. The change in free energy can be due to changes in a particular site's inherent contribution as well as alterations in the interactions between that site and other sites in the protein. Previous studies in soluble proteins have shown that residues outside of van der Waals radii are largely energetically independent.[16] Since glycophorin A is a very well packed and compact structure, it is an unexpected result that most sites

are not coupled strongly, although they are proximal. Indeed, the greatest degree of destabilizing coupling is between residue 75 and 87, sites that are 18 Å apart on distal ends of the helices. The greatest degree of stabilizing coupling is between sites 75 and 84, which are 14 Å apart. Therefore, the residues with the greatest degree of both positive and negative coupling are those most distant in the GpA interface. Recent studies have shown that the long-range coupling between Leu75 and Thr87 are necessary to initiate high-affinity dimerization in the membrane protein bacteriophage M13 major coat protein (MCP), a GxxxG containing low-affinity dimer.[27] In contrast to soluble protein studies, the nature of the coupling along the interface of the GpATM appears unique, due to the predominance of long-range coupling stabilizing the native fold.

The structures of helical membrane proteins and soluble helix bundles have distinct differences. The helices in membrane proteins tend to be longer and have packing angle preferences different from those in soluble proteins.[28,29] The greater degree of long-range coupling in a transmembrane protein as compared to a soluble protein may be due to lesser constraints at the ends of the helices, which allows for greater conformational rearrangements as a function of mutation. A possible model for the structural effects that result in long-range coupling in a transmembrane dimer is visualized in Figure 6. When the crossing angle is changed, there could be an effect on the interactions on the opposing end through a pivot about the crossing point of the helices. This combination of effects could lead to both positive and negative coupling between the distal ends of the helices. The function of the sequence context at the distal ends in a GxxxG-mediated association could be to stabilize a specific crossing angle that allows a close approach of the helices. The presence of glycine residues at the crossing point may maximize favorable inter-monomer interactions. Therefore, long-range coupling could be essential for the formation of the native-like structure in a GxxxG protein.

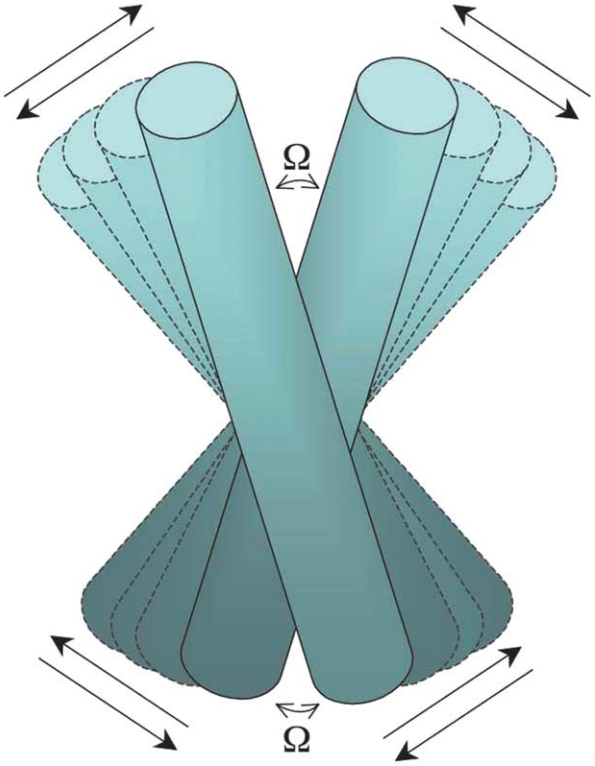

Figure 6. Long-range coupling can be mediated through a change in the helix–helix crossing angle. The schematic illustrates a possible mechanism of long-range coupling in a symmetric dimer. Changes in interactions at a distal end of the helix could have a significant effect on the opposing end mediated by a change in the crossing angle (Ω) through a movement about the pivot point.

Structural rearrangements are necessary to explain strongly coupled mutants

Structure-based parameterization has been used to probe the relationship between structure and energetics in many soluble proteins[30,31] and in the membrane protein bacteriorhodopsin.[32] Recently, a set of coefficients was generated in a parameterization using structural models and energetic data for GpATM single-point mutants.[13] Using these coefficients, it is possible to calculate an accurate $\Delta\Delta G^{o}_{MUT}$ for most GpATM double mutants. In this study, a structure-based parameterization using both single and double mutants was carried out. Neither the coefficients nor the correlation to the experimental free energies change significantly when double-mutant models are included in the parameterization. Therefore, using a limited set of data, an accurate correlation between calculated and experimental free energies can be determined.

The ability to predict the energetics of association for most double mutants using models based on the WT NMR structure indicates that modeling these mutations does not require global rearrangements in the structure of the GpATM. Although many double mutants are weakly stabilizing, their energetics can still be explained using structural modeling that allows structural rearrangements only at the level of side-chain conformation. Since most mutations can be explained using simple models that deviate only slightly from the NMR structure, the GpATM dimer structure appears to be relatively insensitive to mutation, even though these mutations can cause the energetics of association to vary greatly. Additionally, the parameters used to calculate the free energies are based only on changes at the helix–helix interface, validating the assumption that these mutations affect only the dimer stability. Although the molecular modeling in this study is simple, it is still possible to predict the stability of many sequence permutations. In contrast, the free energy of association for strongly coupled mutants cannot be predicted on the basis of this simple model. We hypothesize that it is

necessary to introduce global structural rearrangements to explain the interactions that stabilize strongly coupled mutants. This is true for double mutants that have both positive and negative free energies of coupling. The strength of a study comparing experimentally determined free energies to those calculated on the basis of simple models is that it may be possible to predict which mutations are inducing global structural changes. Conversely, without experimentally determined free energies it is not possible to distinguish the probability that a structural model is valid. Therefore, the relationship between structure and energetics determined by structure-based parameterization can serve as a method to compare more comprehensive models for a particular GpA mutant to predict possible structural rearrangements induced by mutation.

Conclusions

The affinity of the glycophorin A transmembrane dimer can be affected greatly by the presence of mutations at the helix–helix interface. Double point substitutions to alanine exhibit a range of effects. Although all double mutants are destabilized with respect to the WT, few double mutants are coupled in a destabilizing fashion. Most double mutants are coupled only weakly or they are additive. Double-mutant cycles reveal that long-range coupling is essential for the native structure and association of the GpATM. Structure-based parameterization reveals that mutations at sites that are coupled strongly are likely to induce global structural changes. Moreover, we find that a GxxxG motif is neither necessary nor sufficient for dimerization to occur. The interactions along the dimer interface appear complex and are modulated greatly by sequence context.

Materials and Methods

Mutagenesis and protein purification

Double mutants were generated using the pET11A-SNGpA99 construct[5] as a template for site-directed mutagenesis using the Quikchange kit (Stratagene, LaJolla CA) with the appropriate primers. All mutant SNGpA fusion proteins were purified using extractions in the detergent Thesit (Fluka, Switzerland) as described.[5] Immediately before sedimentation equilibrium analysis, the SNGpA fusion protein of interest was exchanged into the desired detergent (C_8E_5, Sigma-Aldrich) by ion-exchange chromatography as described.[7]

Sedimentation equilibrium

Experiments were carried out at 25 °C in a Beckman XL-A analytical ultracentrifuge. The wavelength of absorbance chosen in each experiment was dependent on the ability to observe an adequate dimer population to determine an accurate equilibrium constant. Mutants with a greater propensity for dimerization were observed at 230 nm. For mutants with low propensity to dimerize, the wavelength of absorbance chosen for the experiment was 280 nm, which allows a higher concentration of protein in the cell. Experiments carried out at both wavelengths were shown to have consistent standard-state free energies of association (data not shown). A minimum of nine data sets were used in a global fitting of the data using MacNonlin.[33] The data used in analysis consisted of three significantly different initial protein concentrations run at three or four significantly different speeds. The monomeric sigma was calculated from the amino acid composition using SEDNTERP[34] and held constant during global fitting. Each free energy was measured independently a minimum of three times to determine an average and standard deviation.

A standard-state free-energy value (ΔG_x^o) is calculated by assuming an ideal dilute solution.[35] By extrapolating to the standard state, it is possible to directly compare experiments conducted at different concentrations of detergent. This facilitates the determination of an accurate free-energy value by adjustment of the experimental conditions to populate both monomeric and dimeric species. The equation below is used to calculate the standard state free energy of association:[35]

$$\Delta G_x^o = -RT \ln(K_{app}[\text{micellarDet}]_w) \quad (1)$$

Calculation of the additivity threshold

The $\Delta\Delta G_{MUT}^o$ due to mutation was determined by taking the difference between the average standard-state free energy of association for the WT and mutant protein as shown in the following equation:

$$\Delta\Delta G_{MUT}^o = \Delta G_{x,MUT}^o - \Delta G_{x,WT}^o \quad (2)$$

The free energy of coupling was determined by comparing the free energy perturbation for a double mutant to the sum of the previously determined free energy perturbations for the corresponding single-point mutants as follows:

$$\Delta G_{Coup}^o = \Delta\Delta G_{MUT1MUT2}^o - \Delta\Delta G_{MUT1}^o - \Delta\Delta G_{MUT2}^o \quad (3)$$

A $\Delta G_{Coup}^o = 0 \pm$ threshold indicates an additive mutation, where the threshold represents the limiting value for additivity based on propagation of the experimental standard deviations. If the ΔG_{Coup}^o is outside the threshold, the double mutant is considered coupled. The standard deviation of the $\Delta\Delta G_{MUT}^o$ for the single-point mutants was determined previously and is given as σ_{MUT}.[26] The standard deviation in the $\Delta\Delta G_{MUT}^o$ of double mutants, $\sigma_{\Delta\Delta G_{experimental}}$ (equation (4)), is determined using the standard deviations obtained from multiple centrifugation experiments of each variant as follows:

$$\sigma_{\Delta\Delta G_{experimental}} = \sqrt{\sigma_{\Delta G_{WT}}^2 + \sigma_{\Delta G_{MUT1MUT2}}^2} \quad (4)$$

Since each $\Delta\Delta G_{MUT}^o$ value contains the $\Delta G_{x,WT}^o$ term the equation for coupling can be expressed as follows:

$$\Delta G_{Coup}^o = (\Delta G_{x,MUT1MUT2}^o - \Delta G_{x,WT}^o) - (\Delta G_{x,MUT1}^o - \Delta G_{x,WT}^o) - (\Delta G_{x,MUT2}^o - \Delta G_{x,WT}^o) \quad (5)$$

Two $\Delta G_{x,WT}^o$ terms can be canceled out, and the equation can be simplified by rearrangement of terms into:

$$\Delta G_{Coup}^o = \Delta G_{x,MUT1MUT2}^o - \Delta G_{x,MUT1}^o - \Delta G_{x,MUT2}^o + \Delta G_{x,WT}^o \quad (6)$$

The threshold for the ΔG_{Coup}^o is therefore calculated by the square-root of the sum of the standard deviation of each

component as follows:

$$\text{threshold} = \sqrt{\sigma^2_{\Delta G_{WT}} + \sigma^2_{\Delta G_{MUT1}} + \sigma^2_{\Delta G_{MUT2}} + \sigma^2_{\Delta G_{MUT1MUT2}}} \quad (7)$$

Computational modeling and structure-based parameterization

Computational modeling and structure-based parameterization were carried out as described.[13] The coordinates (1AFO) for the WT glycophorin A NMR structure were used as a basis to model the mutant structures.[9] Amino acid substitutions were made in the 20 NMR structures using WHAT-IF[36] and mild steric clashes were relieved by using the *deball* function. Structural parameters were calculated using the occluded surface algorithm version 7.2.2.[37] The occluded surface algorithm calculates three values used in the parameterization: favorable interchain occluded surface (FOS), unfavorable interchain occluded surface (UOS), and exposed surface (ES). Since the $\Delta\Delta G^{o}_{MUT}$ is the basis for the parameterization, each WT value is subtracted from each mutant value to give ΔFOS and ΔUOS values. The exposed surface area is used as the basis for a calculation of side-chain conformational entropy by comparing the conformational freedom of a residue in an extended state to the conformational freedom in monomer and dimer models considering the maximum possible conformational entropy for that side-chain.[30,38] The difference between the monomer and dimer side-chain conformational entropy values results in the ΔS_{sc} for a mutant model. Structure-based parameterization is than carried out using the following parameters: ΔFOS, the change in inter-monomer favorable packing interactions; ΔUOS, the change in inter-monomer unfavorable packing interactions; and $T\Delta\Delta S_{sc}$, the change in side-chain conformational entropy. Each parameter is the difference of the average value for the 20 mutant models and the corresponding value for the 20 WT NMR structures. The parameterization was done by simultaneous fitting of the single-point mutant values to the following equation by floating the coefficients in equation (8):

$$\Delta\Delta G^{o}_{Calc} = \sigma\Delta FOS + \alpha\Delta UOS + (-T\Delta\Delta S_{sc}) \equiv \Delta\Delta G^{o}_{MUT} \quad (8)$$

The previously determined values for σ and α (using only the single-point mutant models) were used to predict the $\Delta\Delta G^{o}_{Calc}$ for the double mutants. These values are $\sigma = -0.039$ and $\alpha = 6.44 \times 10^{-2}$. No significant change in the coefficients or in the correlation between the $\Delta\Delta G^{o}_{MUT}$ and the $\Delta\Delta G^{o}_{Calc}$ is observed when double-mutant models are included in the parameterization of σ and α.

Acknowledgements

This work was supported by a grant from the NIH (GM57534) and by a Career Award from the Department of Defense (DAMD17-02-1-0427).

References

1. Ackers, G. K. & Smith, F. R. (1985). Effects of site-specific amino acid modification on protein interactions and biological function. *Annu. Rev. Biochem.* **54**, 597–629.
2. Popot, J. L. & Engelman, D. M. (1990). Membrane protein folding and oligomerization: the two-stage model. *Biochemistry*, **29**, 4031–4037.
3. Creamer, T. P., Srinivasan, R. & Rose, G. D. (1995). Modeling unfolded states of peptides and proteins. *Biochemistry*, **34**, 16245–16250.
4. Shortle, D. (1996). The denatured state (the other half of the folding equation) and its role in protein stability. *FASEB J.* **10**, 27–34.
5. Lemmon, M. A., Flanagan, J. M., Hunt, J. F., Adair, B. D., Bormann, B. J., Dempsey, C. E. & Engelman, D. M. (1992). Glycophorin A dimerization is driven by specific interactions between transmembrane alpha-helices. *J. Biol. Chem.* **267**, 7683–7689.
6. Russ, W. P. & Engelman, D. M. (1999). TOXCAT: a measure of transmembrane helix association in a biological membrane. *Proc. Natl Acad. Sci. USA*, **96**, 863–868.
7. Fleming, K. G., Ackerman, A. L. & Engelman, D. M. (1997). The effect of point mutations on the free energy of transmembrane alpha-helix dimerization. *J. Mol. Biol.* **272**, 266–275.
8. Lemmon, M. A., Flanagan, J. M., Treutlein, H. R., Zhang, J. & Engelman, D. M. (1992). Sequence specificity in the dimerization of transmembrane alpha-helices. *Biochemistry*, **31**, 12719–12725.
9. MacKenzie, K. R., Prestegard, J. H. & Engelman, D. M. (1997). A transmembrane helix dimer: structure and implications. *Science*, **276**, 131–133.
10. Langosch, D., Brosig, B., Kolmar, H. & Fritz, H. J. (1996). Dimerisation of the glycophorin A transmembrane segment in membranes probed with the ToxR transcription activator. *J. Mol. Biol.* **263**, 525–530.
11. Russ, W. P. & Engelman, D. M. (2000). The GxxxG motif: a framework for transmembrane helix–helix association. *J. Mol. Biol.* **296**, 911–919.
12. Senes, A., Gerstein, M. & Engelman, D. M. (2000). Statistical analysis of amino acid patterns in transmembrane helices: the GxxxG motif occurs frequently and in association with beta-branched residues at neighboring positions. *J. Mol. Biol.* **296**, 921–936.
13. Doura, A. K., Kobus, F. J., Dubrovsky, L., Hibbard, E. & Fleming, K. G. (2004). Sequence context modulates the stability of a GxxxG mediated transmembrane helix–helix dimer. *J. Mol. Biol.* **341**, 991–998.
14. Liu, W., Eilers, M., Patel, A. B. & Smith, S. O. (2004). Helix packing moments reveal diversity and conservation in membrane protein structure. *J. Mol. Biol.* **337**, 713–729.
15. Jiang, S. & Vakser, I. A. (2004). Shorter side chains optimize helix–helix packing. *Protein Sci.* **13**, 1426–1429.
16. Chen, J. & Stites, W. E. (2001). Energetics of side chain packing in staphylococcal nuclease assessed by systematic double mutant cycles. *Biochemistry*, **40**, 14004–14011.
17. Horovitz, A. & Fersht, A. R. (1992). Co-operative interactions during protein folding. *J. Mol. Biol.* **224**, 733–740.
18. Mildvan, A. S., Weber, D. J. & Kuliopulos, A. (1992). Quantitative interpretations of double mutations of enzymes. *Arch. Biochem. Biophys.* **294**, 327–340.
19. Wells, J. (1990). Additivity of mutational effects in proteins. *Biochemistry*, **29**, 8509–8517.
20. Perry, K. M., Onuffer, J. J., Gittelman, M. S., Barmat, L. & Matthews, C. R. (1989). Long-range electrostatic

interactions can influence the folding, stability, and cooperativity of dihydrofolate reductase. *Biochemistry*, **28**, 7961–7968.
21. LiCata, V. J., Dalessio, P. M. & Ackers, G. K. (1993). Single-site modifications of half-ligated hemoglobin reveal autonomous dimer cooperativity within a quaternary T tetramer. *Proteins: Struct. Funct. Genet.* **17**, 279–296.
22. Fodor, A. A. & Aldrich, R. W. (2004). On evolutionary conservation of thermodynamic coupling in proteins. *J. Biol. Chem.* **279**, 19046–19050.
23. Green, S. M. & Shortle, D. (1993). Patterns of nonadditivity between pairs of stability mutations in staphylococcal nuclease. *Biochemistry*, **32**, 10131–10139.
24. Green, S. M., Meeker, A. K. & Shortle, D. (1992). Contributions of the polar, uncharged amino acids to the stability of staphylococcal nuclease: evidence for mutational effects on the free energy of the denatured state. *Biochemistry*, **31**, 5717–5728.
25. Kleiger, G., Grothe, R., Mallick, P. & Eisenberg, D. (2002). GXXXG and AXXXA common alpha-helical interaction motifs in proteins, particularly in extremophiles. *Biochemistry*, **41**, 5990–5997.
26. Fleming, K. G. & Engelman, D. M. (2001). Specificity in transmembrane helix–helix interactions can define a hierarchy of stability for sequence variants. *Proc. Natl Acad. Sci. USA*, **98**, 14340–14344.
27. Melnyk, R. A., Kim, S., Curran, A. R., Engelman, D. M., Bowie, J. U. & Deber, C. M. (2004). The affinity of GXXXG motifs in transmembrane helix–helix interactions is modulated by long-range communication. *J. Biol. Chem.* **279**, 16591–16597.
28. Bowie, J. U. (1997). Helix packing in membrane proteins. *J. Mol. Biol.* **272**, 780–789.
29. Bowie, J. U. (1997). Helix packing angle preferences. *Nature Struct. Biol.* **4**, 915–917.
30. Baker, B. M. & Murphy, K. P. (1998). Prediction of binding energetics from structure using empirical parameterization. *Methods Enzymol.* **295**, 294–315.
31. Murphy, K. P. (1999). Predicting binding energetics from structure: looking beyond DeltaG degrees. *Med. Res. Rev.* **19**, 333–339.
32. Faham, S., Yang, D., Bare, E., Yohannan, S., Whitelegge, J. P. & Bowie, J. U. (2004). Side-chain contributions to membrane protein structure and stability. *J. Mol. Biol.* **335**, 297–305.
33. Johnson, M. L., Correia, J. J., Yphantis, D. A. & Halvorson, H. R. (1981). Analysis of data from the analytical ultracentrifuge by nonlinear least-squares techniques. *Biophys. J.* **36**, 575–588.
34. Laue, T. M., Shah, B., Ridgeway, T. M. & Pelletier, S. L. (1992). Computer-aided interpretation of analytical sedimentation data for proteins. In *Analytical Ultracentrifugation in Biochemistry and Polymer Science* (Harding, S. E., Rowe, A. J. & Horton, J. C., eds), pp. 90–125, Royal Society of Chemistry, Cambridge.
35. Fleming, K. G. (2002). Standardizing the free energy change of transmembrane helix–helix interactions. *J. Mol. Biol.* **323**, 563–571.
36. Vriend, G. (1990). WHAT IF: a molecular modeling and drug design program. *J. Mol. Graph.* **8**, 52–56 see also page 29.
37. Pattabiraman, N., Ward, K. B. & Fleming, P. J. (1995). Occluded molecular surface: analysis of protein packing. *J. Mol. Recogn.* **8**, 334–344.
38. Lee, K. H., Xie, D., Freire, E. & Amzel, L. M. (1994). Estimation of changes in side chain configurational entropy in binding and folding: general methods and application to helix formation. *Proteins: Struct. Funct. Genet.* **20**, 68–84.

Edited by G. von Heijne

(Received 2 July 2004; received in revised form 31 August 2004; accepted 8 September 2004)