Computational Methodology for Estimating Changes in Free Energies of Biomolecular Association upon Mutation. The Importance of Bound Water in Dimer−Tetramer Assembly for β37 Mutant Hemoglobins†

James C. Burnett, Glen E. Kellogg,* and Donald J. Abraham*

Department of Medicinal Chemistry, School of Pharmacy, Institute for Structural Biology and Drug Discovery, Virginia Commonwealth University, Richmond, Virginia 23298-0133

Received July 26, 1999

ABSTRACT: The computational modeling program HINT (Hydropathic INTeractions), an empirical hydropathic force field that includes hydrogen bonding, Coulombic, and hydrophobic terms, was used to model the free energy of dimer−tetramer association in a series of deoxy hemoglobin β37 double mutants. Five of the analyzed mutants (β37W → Y, β37W → A, β37W → G, β37W → E, and β37W → R) have been solved crystallographically and characterized thermodynamically and subsequently made a good test set for the calibration of our method as a tool for free energy prediction. Initial free energy estimates for these mutants were conducted without the inclusion of crystallographically conserved water molecules and systematically underestimated the experimentally calculated loss in free energy observed for each mutant dimer−tetramer association. However, the inclusion of crystallographic waters, interacting at the dimer−dimer interface of each mutant, resulted in HINT free energy estimates that were more accurate with respect to experimental data. To evaluate the ability of our method to predict free energies for de novo protein models, the same β37 mutants were computationally generated from native deoxy hemoglobin and similarly analyzed. Our theoretical models were sufficiently robust to accurately predict free energy changes in a localized region around the mutated residue. However, our method did not possess the capacity to generate the long-range secondary structural effects observed in crystallographically solved mutant structures. Final method analysis involved the computational generation of structurally and/or thermodynamically uncharacterized β37 deoxy hemoglobin mutants. HINT analysis of these structures revealed that free energy predictions for dimer−tetramer association in these models agreed well with previously observed energy predictions for structurally and thermodynamically characterized β37 deoxy hemoglobin mutants.

The formation of two or more biomolecules into a stable assembly is a complex process characterized by an intricate balance of entropic and enthalpic forces (*1*−*5*). The net combination of these terms is the free energy of binding, $\Delta G_{association}$, which is key to the understanding of biomolecular processes. A major component of the free energy of association results from a wide variety of noncovalent interactions that include hydrogen bonding, ion pairing, and hydrophobic coupling (*6*−*8*). Moreover, the relative contributions of each of these interactions to the overall binding are highly variable with respect to the biological system of interest (*9*). Some complexes obtain their binding strength from a few, energetically favorable interactions, while others achieve tight binding from many interactions distributed over the entire binding interface. Additionally, structurally conserved water molecules play a key role in both binding stabilization and binding site recognition (*10*−*20*). Understanding the role of such varied and complex determinants of protein−protein associations will ultimately aid in providing a general methodology for the rational design and optimization of therapeutic agents engineered to modify distinct biomolecular systems.

Recent advances in the field of X-ray crystallography have facilitated the study of numerous biologically important systems. To this end, a variety of computational chemistry methods, utilizing a wide range of complex static and dynamic molecular force field methods, have been used in an attempt to examine the strength of biomolecular atom−atom interactions in X-ray crystal structures (*10*, *21*−*28*). However, few of these methods have attempted to quantitatively translate these interactions into binding energy estimates. In particular, a study by Covell and Wallqvist (*26*) utilized a method based on weighted atomic pairwise surface burial to accurately model the thermodynamics of several antibody−enzyme complexes, while Tawa, Topol, Burt, and Erickson (*28*) have used a model based on molecular mechanics, dielectric continuum solvation, and exposed surface area to successfully predict the binding free energies of peptidic inhibitors of HIV protease mutants. However, while these computer-aided methods have demonstrated degrees of success in predicting biomolecular free energies, none to date have attempted to directly account for the hydrophobic binding force.

† This research was supported in part by the National Institutes of Health Grant 5RO1HL32793-15.

* To whom correspondence should be addressed. (G.E.K.) Telephone: +1 804 828 6452. Fax: +1 804 827 3664. E-mail: glen.kellogg@vcu.edu. (D.J.A.) Telephone: +1 804 828 8183. Fax: +1 804 827 3664. Email: donald.abraham@vcu.edu.

10.1021/bi991724u CCC: $19.00
Published on Web 01/28/2000

The computational chemistry program HINT (Hydropathic INTeractions) was developed for the global analysis of all possible noncovalent biomolecular atom−atom interactions, including hydrogen bonding, Coulombic, acid−base, and hydrophobic forces. HINT differs from most other models that use force fields derived from Newtonian physics to calculate atom−atom interaction strengths. Instead, HINT employs thermodynamic, atom-based, hydropathy values derived from the measurement of solvent partitioning (octanol/water) of organic molecules to quantitatively "score" interactions. Importantly, the solvent partitioning effect, represented by hydrophobicity (log $P_{octanol/water}$), was shown as early as the 1960s (*29*, *30*) to be an important term in quantitative structure−activity relationship correlations, describing small molecule binding to proteins.

Because HINT uses the same experimentally derived scoring algorithm for all interaction types, calculated scores arise from the empirical (hydropathic) properties of the interacting atoms and are not the product of interaction-specific mathematical functions. This results in a unified description of all noncovalent interactions in the biological environment. Several studies utilizing HINT have shown that these interaction scores are related empirically to binding affinity and/or free energy for a variety of biological systems (*31*−*33*). In effect, each atom−atom interaction score is a self-contained free energy estimate of that interaction. Thus, the basis of the HINT model is that these atom−atom ΔG estimates can be summed to calculate a total ΔG estimate for the entire biomolecular interaction. This is fundamentally different from molecular mechanics approaches that assume energy component additivity by summing total contributions from different types of energy terms (*34*).

In a previous study that utilized high-resolution crystal structures, HINT convincingly predicted free energy changes associated with native and mutant hemoglobin (Hb)[1] dimer−tetramer assembly (*35*). This analysis identified and quantitated key noncovalent atom−atom interactions, which stabilized the Hb dimer−dimer interface during the assembly of two $\alpha\beta$ dimers into the $\alpha_2\beta_2$ tetramer. Interestingly, it was observed that not only electrostatic but also numerous hydrophobic contacts were responsible for overall stability of the general Hb tetramer (*35*). While this previous analysis was accurate in predicting changes in free energy, it focused exclusively on residue−residue interface interactions and did not elaborate on the contribution of structurally conserved water molecules within examined three-dimensional structures.

Recently, the high-resolution deoxy Hb crystal structures of mutants in which the Trp37β residue position was replaced by Tyr, Ala, Gly, Glu, and Arg were reported (*36*, *37*). The wealth of experimental data on this set of mutant structures makes it an attractive system to use in the further calibration of the HINT method. Distinct crystal structures of each of the mutants allows not only for the examination of the direct effects of individual mutant residue interactions on the stability of the tetrameric complexes but also facilitates the examination of far reaching structural changes associated with each mutant. Furthermore, each crystal structure includes its own set of structurally conserved water molecules, making possible a detailed and self-consistent examination of the contribution of water to the stability of the Hb dimer−dimer interface. Finally, experimental measurements of the change in free energy associated with dimer−tetramer assembly is available for all of these β37 mutant crystal structures (*38*, *39*).

The research in this paper focuses on translating distinct differences in quantitated atom−atom interactions into free energy predictions for a very specific set of structural mutations. This unique opportunity, to investigate several mutant crystal structures of the same residue position, facilitates the examination of the thermodynamic impact of each mutant residue on different structural levels. Thus, analysis of first level thermodynamic effects encompasses the observed free energy change resulting from atom−atom interactions made solely by the mutated amino acid residue. Second level thermodynamic examination involves free energy changes resulting from structural alterations in regions not directly adjacent to the residue mutation site. Finally, examination of third level thermodynamic changes includes the effects of mutations on the overall energetic contribution of water molecules to dimer−dimer interface stability.

In general, HINT interaction scores for dimer−tetramer assembly—in mutant β37 hemoglobins—correlate well with corresponding experimental data and calibrate this analysis as a method for free energy prediction. The inclusion of crystallographically important waters in the analysis of dimer−tetramer assembly indicates that the binding energy contributed by these small molecules cannot be ignored and that hydration effects are an important link for fully understanding the nature of protein−protein interactions. Finally, through computer-assisted modeling of both structurally and thermodynamically characterized and uncharacterized β37 mutants, this study indicates that our computer-aided method for free energy predictions is sufficiently robust to apply to computationally derived biomolecular systems.

METHODS

Structural Data. All structural data in this study was obtained from the Brookhaven Protein Data Bank. All examined native and β37 mutant crystal structures were solved by Kavanaugh et al. in two separate studies and were thus sequestered into separate crystal sets (*36*, *37*). Both studies included a native crystal structure that was used as a comparative reference for corresponding β37 mutant structures in this research (identified by the year in which it was reported). Thus, the 1998 crystal set is composed of native deoxy Hb structure β37W^{1998} (Brookhaven coordinate file 1A0U; Kavanaugh, Weydert, Rogers, and Arnone, 1998 (*36*)) and four corresponding double mutant structures: β37Y, β37A, β37G, and β37E (Brookhaven coordinate files 1A0V, 1A0W, 1A0X, and 1A0Y, respectively; Kavanaugh et al., 1998 (*36*)). For all structures in the 1998 crystal set, protein overexpression in *Escherichia coli* necessitated that the

[1] Abbreviations: Hb, hemoglobin; rms, root-mean-square; $\Delta\Delta G$, difference between native and mutant association free energies; β37, general referral to hemoglobin crystal structures examined in this study; β37W^{1998}, native hemoglobin crystal structure from ref *36*; β37W^{1992}, native hemoglobin crystal structure from ref *37*; β37Y, hemoglobin crystal structure with Trp37β → Tyr mutations; β37A, hemoglobin crystal structure with Trp37β → Ala mutations; β37G, hemoglobin crystal structure with Trp37β → Gly mutations; β37E, hemoglobin crystal structure with Trp37β → Glu mutations; β37R, hemoglobin crystal structure with Trp37β → Arg mutations; β1M, recombinant hemoglobin with Val1β → Met mutations.

β-subunit NH_2 terminal valine be replaced by a methionine (*40*, *41*). However, it is important to note that previous studies have documented in detail the lack of any significant structural or functional consequences of the β1M mutation (*41*). In this study, the β1M mutation does not impact on dimer−dimer interface scores. The 1992 crystal set is composed of native deoxy Hb structure β37W^{1992} (Brookhaven coordinate file 1HBB; Kavanaugh, Rogers, Case, and Arnone, 1992 (*37*)) and a corresponding β37R double mutant (Brookhaven coordinate file 1HBA; Kavanaugh et al., 1992 (*37*)).

Molecular Models and Energy Minimizations. Hydrogens were modeled to all Hb structures at pH = 7.0, using InsightII version 97.2 of Biosym Molecular Simulations (*42*). The N- and C-termini were modeled as NH_3^+ and COO^-, respectively. Hydroxyl protons on Tyr, Thr, and Ser residues were manually checked to verify that hydrogen bond geometries were present. The side chain amido Nδ2 and Oδ1 atoms of Asn97α were rotated by 180° about the Cβ−Cγ axis to effect a hydrogen bond with the side chain carboxylate of Asp99β at dimer−dimer interfaces for all examined structures. Each Hb thus prepared was subjected to 100 cycles of hydrogen-only steepest descents minimization with Discover version 2.9 as incorporated into InsightII. Afterwards, the hydrogens of crystallographic waters at all hemoglobin dimer−dimer interfaces were manually optimized with respect to surrounding amino acid residues and adjacent waters, and all crystallographic water molecules were further subjected to conjugate gradients minimization until convergence to an average derivative of 0.01 kcal mol^{-1} Å^{-1}. This described minimization procedure adequately removed bad contacts involving the modeled hydrogen atoms for all structures. Note that no heavy atom position was modified by this procedure, and errors (crystallographic or otherwise) existing in the PDB files were not corrected, except as noted above.

All computationally generated double mutant β37 models were created with the Biopolymer module as incorporated in InsightII version 97.2. All de novo mutants were generated from the β37W^{1998} crystal structure. Additionally, computational models for mutants β37S and β37T were also generated from the β37A crystal structure, and mutant β37F was also modeled from the β37Y crystal structure. Generation of mutant models involved replacing the 37β position of the starting crystal structure with the desired mutant residue, and then manually optimizing it with respect to surrounding amino acid residues and crystallographic waters. Next, the program GRID (version 17 of Molecular Discovery Ltd.) (*43*) was used to access mutant residue effects on the presence or absence of water molecules within a 5 × 5 × 5 Å box centered on the mutant residue. An initial GRID analysis was performed on the starting crystal structure β37W^{1998} to assess the accuracy of the program. Results from these analyses indicated that GRID accurately predicted the positions of crystallographic waters to within approximately 1.0 Å of their coordinate positions within a 5 × 5 × 5 Å box centered on the native Trp37β residue. Subsequent GRID analysis of all de novo mutant models indicated no mutant-induced displacement of preexisting crystallographic waters. GRID did, however, indicate the insertion of a new water molecule bridging between the Asp94α carboxylate and hydroxyl moieties of residues Ser37β and Thr37β in the β37S and β37T mutants generated from both the β37W^{1998} and β37A crystal structures.

Structural optimization of de novo β37 mutants involved the minimization of a 5 Å zone (including waters suggested by GRID) centered on each mutant 37β residue. Initial optimization involved 100 iterations of steepest descents minimization (to remove bad contacts), followed by conjugate gradients minimization until convergence to an average derivative of 0.01 kcal mol^{-1} Å^{-1}.

Determination of Interface Waters. The ideology behind this study was to treat crystallographic waters at Hb dimer−dimer interfaces as ligands that stabilized formation of the Hb tetramer. Thus, our analysis was concerned with waters that interacted with both dimers of the Hb tetramer—bridging the interface and engaging in facilitative interactions with both. For example, in the 1998 crystal set structures (β37W^{1998}, β37Y, β37A, β37G, and β37E), waters at the dimer−dimer interface were chosen based on the criteria that each water molecule must make at least one significant interaction (hydrogen bonding or acid/base interaction with score >50 HINT units) with each assembling dimer. Using this protocol, 23, 21, 22, 20, and 22 significant interface waters were identified for structures β37W^{1998}, β37Y, β37A, β37G, and β37E, respectively. Comparison of the separate interdimer waters for each 1998 crystal set structure resulted in the organization of a comparative hydration "superset" that included all identified interface waters from each examined structure. In this way, HINT total interaction scores for superset waters from each structure were quantitated and compared. Note that if a particular mutant did not contain one of the waters from the hydration superset, then this water's HINT score was equal to zero. Using the described protocol, the 1998 crystal set hydration superset contained 27 waters, while the 1992 crystal set hydration superset contained 26 waters.

Analysis of Hydropathic Interactions. The role of hydropathy in dimer−tetramer association was analyzed using the computational program HINT. HINT was devised to simply and intuitively interpret biological structure in terms of experimentally derived parameters and structural constants that are readily available, either from empirical or calculational sources. Using the HINT model specific interactions between Hb subunits were described as a double sum over the atoms within each component:

$$B = \sum_{j=1}^{\text{atoms}} \sum_{i=1} b_{ij} = \sum\sum (S_i a_i S_j a_j R_{ij} T_{ij} + r_{ij}) \qquad (1)$$

where S is the solvent-accessible surface area, a is the hydrophobic atom constant, T is a descriptor function, and R and r are functions of the distance between atoms i and j. From this empirically derived equation, a binding score is calculated where b_{ij} describes the specific interaction between atoms i and j, and B describes the total interaction between the two species.

The hydrophobic atom constants (a_{ij}) are derived by reduction of the fragment constants for the water/octanol partition coefficient (*44*, *45*). Positive-signed atom (fragment) constants indicate hydrophobic atoms (fragments), while negative-signed constants indicate polar or hydrophilic atoms (fragments). Partition coefficients (sum of hydrophobic atom

constants) for small molecules calculated by HINT are similar to values calculated by other methods (*45*). Solvent-assessible surface area (S_i) is a constant describing the shape and accessiblity of the atom and its tendency for interaction. Buried atoms have a smaller *S* and are less involved in interactions.

The descriptor function, T_{ij}, differentiates among three possibilities for polar−polar interactions (acid−acid, hydrogen bonding/acid−base, and base−base) in order to maintain the convention that favorable interactions have positive scores. Each InsightII CVFF atom type is assigned descriptor variables to represent its hydrogen bonding acceptor/donor character, Bronsted acid/base character, and Lewis acid/base character. These are used by T_{ij} to calculate a value of +1, −1, or 0 for each atom−atom interaction. The HINT analysis considers the units of T_{ij} to be Å^{-4}, so that *B* and b_{ij} have no units.

The functional form of the range dependence is described by two terms (R_{ij} and r_{ij}). The former scales the hydrophobic atom constant/solvent-accessible surface area product with distance, while the second is independent of hydropathy and responds only to distance variations. For this work, R_{ij} has been set to the simple exponential, e^{-r} where *r* is the distance between the interacting atoms in Å (*46*). The 6-12 Lennard-Jones function,

$$r_{ij} = Ae_{ij}\,[(\upsilon dw/r)^{-6} - 2(\upsilon dw/r)^{-12}] \qquad (2)$$

where e_{ij} is the van der Waals parameter, and *A* is a scaling factor balancing the contributions of hydropathic van der Waals forces, was used for r_{ij}. For this study, $A = 50$ kcal mol^{-1}.

HINT Calculation Details. This study was performed with HINT v.2.26I (eduSoft LC, Ashland, VA), integrated into insightII 97.2, using adjustable HINT parameters as reported previously. The atom potential types used by HINT v.2.26I are based on the CVFF force field. The fragment values and log *P* calculation method of Hansch and Leo (*45*) was modified and adapted to the CVFF atom primitive set.

The minimized Hbs were apportioned to individual molecular subunits (α_1, β_1, α_2, β_2), and HINT parameters for each subunit and interfacial hydration superset were calculated using the HINT dictionary partitioning algorithm. Total interaction scores for $\alpha_1\alpha_2$, $\alpha_1\beta_2$, and $\alpha_2\beta_1$ inter-dimer interfaces, and $\alpha_2\beta_2$ tetramer-interface waters were calculated for each uniquely modeled and structurally optimized Hb.

ΔΔG Calculation Details. HINT interaction estimates are based on thermodynamically determined atomic hydropathy parameters and, as such, the HINT interaction scores should provide an approximation of free energies for biomolecular contacts (associations). Thus, HINT total dimer−dimer interface scores were converted to kcal mol^{-1} estimates in the following manner: the total free energy for the assembly of two $\alpha\beta$ dimers for native deoxy Hb at pH 7.0 and 21.5 °C, $\Delta G_{\text{dimer-tetramer}}$ ($\Delta G_{\text{D-T}}$), has been determined experimentally as −14.3 kcal mol^{-1} (± 0.2) (*39*). This value is divided into the HINT total interface score, without the inclusion of structural waters, for native deoxy Hb dimer−dimer interface scores for β37W^{1998} and β37W^{1992} to generate conversion factors of 528 and 554 HINT score units per kcal, respectively, with the following caveat: the $\Delta G_{\text{D-T}}$ value includes both the energy of formation of the noncovalent bonds that comprise the dimer−dimer interface and the statistical solvent rearrangement that accompanies bond formation.

Using the applicable HINT score units per kcal conversion factor (crystal set specific) for each Hb structure examined, the total binding energy for each Hb was calculated as the sum of binding energies resulting from the total dimer−dimer interface score between associating $\alpha\beta$ dimers without water, $\Delta G_{\text{D-T,}}$ and the binding energy resulting from binding interactions between interface waters and the associated tetramer, $\Delta G_{\text{tetramer-water}}$ ($\Delta G_{\text{T-W}}$):

$$\Delta G_{\text{calculated}} = \Delta G_{\text{D-T}} + \Delta G_{\text{T-W}} \qquad (3)$$

As indicated initially, this formalization treats the crystallographically examined water molecules as ligands that stabilize the interaction between associated dimers.

On the basis of the predicted $\Delta G_{\text{calculated}}$ for each examined native and mutant Hb structure, ΔΔ*G* values of differences in free energy between native and corresponding mutant Hbs were calculated:

$$\Delta\Delta G_{\text{calculated}} = (\Delta G_{\text{native[D-T]}} - \Delta G_{\text{mutant[D-T]}}) + (\Delta G_{\text{native[T-W]}} - \Delta G_{\text{mutant[T-W]}}) \qquad (4)$$

RESULTS AND DISCUSSION

Hemoglobin is an allosteric tetrameric protein, containing two subunit types (α and β), which readily dissociate in dilute solutions into two dimers, $\alpha_1\beta_1$ and $\alpha_2\beta_2$. The assembly of these dimers into a tetramer results in four subunit interface interactions ($\alpha1-\alpha2$, $\beta1-\beta2$, $\alpha1-\beta2$, and $\alpha2-\beta1$). Thus, this protein has a 2-fold symmetry axis, and residue−residue interactions between the $\alpha_1\beta_2$ and $\alpha_2\beta_1$ subunits are identical. The Trp37β residue plays a key role in stabilizing both the deoxy and oxy Hb tetrameric complexes through residue interactions at the $\alpha_1\beta_2$ and $\alpha_2\beta_1$ interfaces. Specifically, the indole NH of Trp37β forms a strong hydrogen bond with the Asp94α side chain carboxylate, while the size and position of the entire indole aids in sterically maintaining the composure of the dimer−dimer interface in both deoxy and oxy Hb structures (*47*). Interestingly, the strong hydrogen bond between the Trp37β and Asp94α residues is one of two hydrogen bonds that are maintained during the allosteric transition from the deoxy to the oxy form of Hb (*35*) and thus is a key pivot point for movement within the Hb tetramer.

While it is clear that the Trp37β residue is important for deoxy Hb tetramer stability, the magnitude of thermodynamic destabilization produced by mutations at this residue site is surprising; the average loss of free energy accompanying dimer−tetramer association for thermodynamically characterized β37 mutants is approximately 5.0 kcal mol^{-1} (*38*, *39*). On average, this observed loss in free energy is much larger than the destabilization observed for nearly all deoxy Hb residue mutations that have been characterized previously (*39*). Thus, while the HINT analysis in this study focuses on the calibration of our method in the prediction of biomolecular free energies, this research has also been undertaken to elucidate the chemical reasons for the large thermodynamic cost associated with deoxy Hb 37β mutations.

Table 1: Summary of HINT Subunit Interface Scores and $\Delta\Delta G$ Estimates for β37 Mutants without the Inclusion of Crystallographic Water Molecules

hemoglobin	$\alpha_1\beta_2$	$\alpha_2\beta_1$	$\alpha_1\alpha_2$	$\beta_1\beta_2$	interface total	HINT $\Delta\Delta G$[a]	experimental $\Delta\Delta G$[b]
β37W[1998]	2282	1627	3699	0	7608		
β37Y	2616 (2249[c])	1534	2834	0	6993	1.2 (1.9[c])	5.3
β37Y[d]	2252 (1885[c])	1436	2834	0	6522	2.2 (2.9[c])	5.3
β37A	1725	1025	3366	0	6116	2.8	6.0
β37G	1619	920	2839	0	5378	4.2	6.1
β37E	1468	764	3223	0	5455	4.1	9.2
β37W[1992]	2373	1752	3796	0	7921		
β37R	2973	2074	3804	0	8851	−1.7	1.0

[a] HINT $\Delta\Delta G$ predictions calculated in kcal mol^{-1}. [b] Experimental $\Delta\Delta G$ calculated in kcal mol^{-1} (*38*, *39*). [c] Normalized for subunit interface $\alpha_1\beta_2$ residue−residue interaction α40K−β146H. [d] HINT scores and $\Delta\Delta G$ predicitions for β37Y as a tyrosinate.

The analyzed deoxy Hb mutant crystal structures were solved in two separate studies and are thus sequestered into two separate crystal sets in this analysis. Structures of the 1998 mutant crystal set were solved by Kavanaugh et al. (*36*) and include mutants β37Y, β37A, β37G, β37E, and a corresponding native structure, β37W[1998]. The 1992 mutant crystal set was solved by Kavanaugh et al. (*37*) and contains two structures, a native hemoglobin, β37W[1992], which appears to be nearly isomorphous with the β37W[1998] structure; and a corresponding β37R mutant.

HINT analysis indicates that the observed thermodynamic destabilization produced upon mutation to Tyr37β, Ala37β, Gly37β, and Glu37β results from: the gradual loss of 37β residue interdimer interactions with alpha subunit residues (and in the case of the Ala37β and Glu37β mutations, new unfavorable interdimer interactions); long-range perturbations to the alpha subunit dipeptide carboxy terminus (*36*); dimer−dimer interface water contacts that are, in general, less thermodynamically stable than waters in the native structure. In contrast, the Arg37β mutation stabilizes the hemoglobin tetramer through new favorable interactions between the side chain guanidinium and the Asp94α side chain carboxylate. However, examination of this mutant structure's interface waters indicates that these small molecules contribute less thermodynamic stability to the dimer−dimer interface than corresponding native deoxy Hb β37W[1992] interface waters.

In general, mutations at the 37β site adhere to a logical order of effects that one would intuitively predict for the residue replacements that were examined. However, two of the mutant structures, β37Y and β37R, produce results that impact generally on the modeling of protein thermodynamics and test the applicability of the HINT analysis as a predictive model for protein thermodynamics. Furthermore, the analysis of distinct mutant Hb interface waters indicates the critical nature of these small but powerful molecules for modeling and understanding the energetics of biomolecular systems. Finally, using our method to predict free energy changes associated with computationally generated β37 mutant models illuminates error estimates involved in the HINT method and indicates that this computer-aided method is sufficiently robust for use in predictions of free energy changes in structurally uncharacterized biomolecular systems.

HINT Interface Interactions without Inclusion of Structural Water. Native and mutant hemoglobin dimer−dimer interfaces were initially analyzed without the inclusion of crystallographic water molecules for all examined structures. Summation of subunit interactions, $\alpha_1\alpha_2$, $\alpha_1\beta_2$, and $\alpha_2\beta_1$, results in a total dimer−dimer HINT score for examined structures (Table 1). As observed in Table 1, the $\beta_1\beta_2$ interface gives a HINT total interaction score of zero for all deoxy Hb structures and consequently does not contribute to the total HINT interface scores. Moreover, $\alpha_1\beta_2$ and $\alpha_2\beta_1$ interactions, though symmetrically related, are not equal as the tetrameric 2-fold axis does not superimpose upon the crystallographic 2-fold axis in deoxy Hb. Thus, refined nonsymmetry averaged electron density maps permit residues to vary from the molecular 2-fold axis, and $\alpha_1\beta_2$ and $\alpha_2\beta_1$ interfaces give different HINT total interaction scores (*48*).

For the two different native deoxy Hb structures examined, β37W[1998] and β37W[1992], estimates for thermodynamic scaling factors are 528 and 554 HINT score units per kcal mol^{-1}, respectively. This difference suggests a precision in the HINT calibration, ca. 5%, which is similar in magnitude to the experimental error of the solution thermochemistry experiments (*38*, *39*). Subsequently, differences between each mutant β37 HINT total interface score versus its respective native deoxy Hb HINT total interface score are translated into $\Delta\Delta G$ predictions as described in eqs 3 and 4.

The Hb $\alpha_1\beta_2$, $\alpha_2\beta_1$, and $\alpha_1\alpha_2$ interfaces are composed of a large number of noncovalent residue−residue and residue−water interactions (Figure 1). In the native hemoglobin structure, Trp37β is a central residue in the flexible joint region of both the $\alpha_1\beta_2$ and $\alpha_2\beta_1$ interfaces, and while it does not interact significantly with crystallographic waters, it does make several favorable hydrophobic contacts with residues Arg92α, Pro95α, and Tyr140α and engages in an important stabilizing electrostatic interaction with the side chain carboxylate of residue Asp94α (Figures 1 and 2) (*35*). Atom−atom contacts between residues Trp37β and Asp94α are dominated by the formation of a strong hydrogen bond and, to a lesser degree, acid/base interactions, between the Trp37β indole NH group and the side chain carboxylate of Asp94α (combined $\alpha_1\beta_2$ and $\alpha_2\beta_1$ 37β HINT residue scores = 800). As the Trp37β residue is replaced with a Tyr, Ala, Gly, and Glu mutation, the gradual loss of both favorable hydrophobic interactions and a strong hydrogen bond with the Asp94α residue results in a gradual decrease in $\alpha_1\beta_2$ and $\alpha_2\beta_1$ interface scores (Table 1), and also corresponds with the gradual increasing separation of the dimer−dimer interface, as observed crystallographically (*36*). RMS deviations for these mutants, when superimposed on the corresponding native deoxy Hb (β37W[1998]) reflect this increasing structural change with values of 0.28, 0.30, 0.41, and 0.45 for heavy atom positions of mutant crystal structures β37Y,

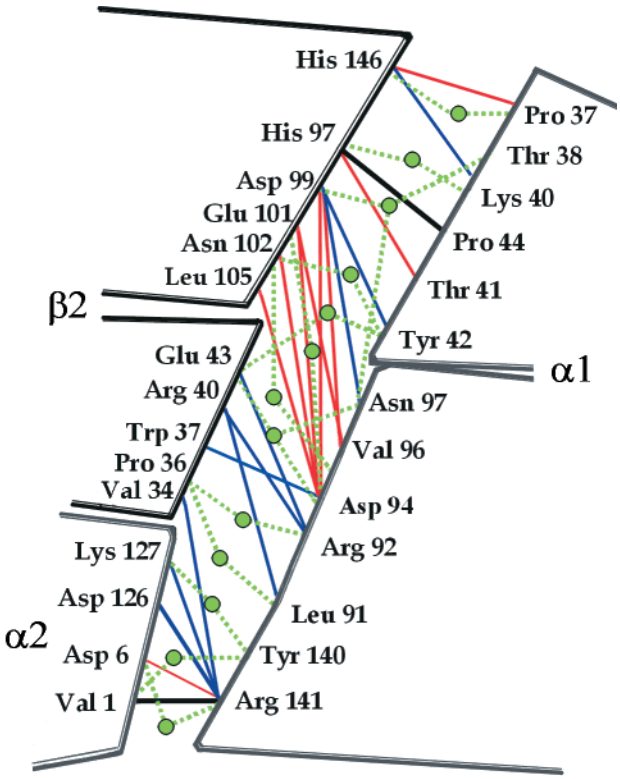

FIGURE 1: After Abraham, Kellogg, Holt, and Ackers (1997), schematic of the deoxy hemoglobin $\alpha_1\beta_2$ and $\alpha_1\alpha_2$ interfaces with residue−residue and residue−water interactions determined by HINT. Although each contact is typically composed of both positive and negative interactions, the dominance of negative interactions is indicted by red, and net positive interactions by blue and black. Important interdimer waters are indicated by green circles, and water−residue hydrogen bonds are indicated by dashed green lines.

β37A, β37G, and β37E, respectively. Stereo diagrams of the environments of examined native and mutant β37 crystal structures are shown in Figure 2.

The Tyr37β mutation results in the least thermodynamic perturbation to the $\alpha_1\beta_2$ and $\alpha_2\beta_1$ interfaces, as the hydroxyl group of the side chain phenol maintains weak acid/base interactions with the Asp94α side chain carboxylate and the side chain methylene maintains favorable hydrophobic interactions with residue Arg92α (combined $\alpha_1\beta_2$ and $\alpha_2\beta_1$ HINT residue scores = 392) (Figures 2 and 3). However, for the β37Y mutant structure, the total HINT $\alpha_1\beta_2$ interface score is greater than the same total HINT interface score for the native structure (Table 1) and is the result of an unexpected increase in the ionic interaction between interface residues Lys40α and the His146β (with respect to the same interaction in the native deoxy hemoglobin $\alpha_1\beta_2$ interface, a thermodynamic cost of −0.7 kcal mol^{-1} is incurred for the β37Y $\alpha_1\beta_2$ interface). This interaction occurs approximately 22 Å away from the 37β residue position, and closer examination indicates that it may be an artifact of the refined three-dimensional structure and should not be a direct result of the Tyr37β mutation. However, for the purposes of this investigation, this interaction is not modified, as one of the goals of this study is the direct utilization of high-resolution crystallographic data. Mutations Ala37β and Gly37β both result in elimination of all bonding interactions with the Asp94α side chain carboxylate. The Ala37β mutation is comparatively less stable, engaging in hydrophobic-polar interactions with residues Arg92α and Tyr140α (combined $\alpha_1\beta_2$ and $\alpha_2\beta_1$ HINT residue scores = −80) (Figure 3), while the Gly37β mutation engages in fewer unfavorable interactions (combined $\alpha_1\beta_2$ and $\alpha_2\beta_1$ HINT residue scores = −5). The Glu37β mutation produces the largest thermodynamic destabilization to the $\alpha_1\beta_2$ and $\alpha_2\beta_1$ interfaces, due to unfavorable atom−atom interactions between the Glu37β and Asp94α side chain carboxylates (combined $\alpha_1\beta_2$ and $\alpha_2\beta_1$ HINT residue scores = −457). Mutation Arg37β results in new hydrogen bonding and acid−base interactions between the mutant residue side chain guanidinium and the Asp94α side chain carboxylate (in agreement with a previous HINT analysis of this mutant) (*35*). For this mutant, the HINT score for the combined $\alpha_1\beta_2$ and $\alpha_2\beta_1$ Arg37β residue score is 1969 vs a combined $\alpha_1\beta_2$ and $\alpha_2\beta_1$ total HINT residue score of 770 for the corresponding native residue in the β37W^{1992} structure.

All β37 mutant structures, except the β37R structure, show a general decrease in $\alpha_1\alpha_2$ intersubunit HINT scores due to long-range mutational effects, which translate into increased disorder at the alpha subunit carboxy terminal dipeptide (*36*). Mutant β37G displays the greatest decrease in the $\alpha_1\alpha_2$ total HINT interface score. For this mutation, the Val1α amino terminus interaction with the Arg141α carboxy terminus is substantially weakened with respect to the same interaction in the native β37W^{1998} structure, and the overall decrease in this interface strength is equal to a loss of 860 HINT score units (1.6 kcal mol^{-1}). Similarly, mutant structures β37Y, β37A, and β37E all display varying degrees of decrease in $\alpha_1\alpha_2$ HINT interface scores for the same reason as indicated for the β37G mutant (Table 1). The β37R mutant $\alpha_1\alpha_2$ interface is virtually unchanged (HINT $\alpha_1\alpha_2$ total interface score = 3804) when compared with its corresponding native structure (β37W^{1992}) ($\alpha_1\alpha_2$ total HINT interface score = 3790).

HINT ΔΔ*G* predictions for all β37 mutants, without the inclusion of structural waters, clearly indicate that our ΔΔ*G* predictions systematically underestimate experimentally determined ΔΔ*G* values by an average of 3.5 kcal mol^{-1}. This trend is surprising, as previous HINT examination of several mutant deoxy Hb crystal structures, in which water was not explicitly included, resulted in ΔΔ*G* predictions that were considerably more accurate (*35*). From this analysis, it is clear that another factor must be contributing significantly to the free energy changes (ΔΔ*G*) associated with these mutations. Consequently, differences in the binding energy of structurally conserved water molecules are examined for each mutant and native deoxy Hb crystal structure.

Solvent Contributions to Dimer−Dimer Interface Stability. Previous studies have indicated that inclusion of crystallographic waters is necessary for realistic estimations of free energy associations in biomolecular systems (*26, 27, 49*−*51*). Covell and Wallquist (*26*) have clearly shown that failure to account for crystallographic water molecules in the D1.3-HE and D2.1-E5.2 antibody−enzyme interfaces for native and mutant structures resulted in underestimates of changes in free energy of binding for all mutant structures examined. Bhat, Bentley, Boulot, Greene, Tello, Dall'Acqua, Souchon, Schwarz, Mariuzza, and Poljak (*49*) have indicated that bound water molecules play a key role in mediating intermolecular interactions during biomolecular associations.

For all crystal structures examined in this study, structurally conserved water molecules are observed to line each

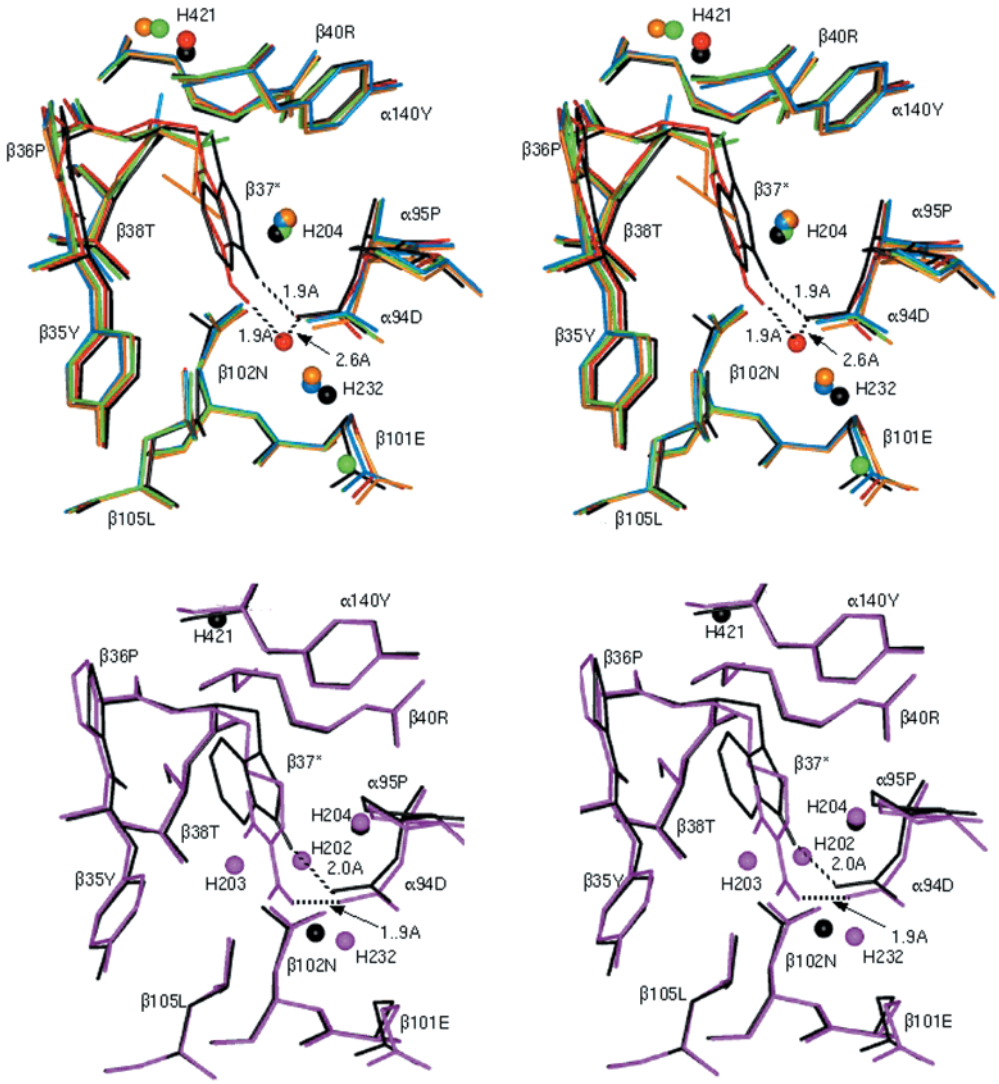

FIGURE 2: Superimposed stereo diagrams showing the environment of residue β37 in the $\alpha_1\beta_2$ interfaces of the 1998 crystal set (top): β37W[1998](black), β37Y (red), β37A (green), β37G (blue), and β37E (orange). The 1992 crystal set (bottom): β37W[1992] (black) and β37R (purple). Hydrogen bonds are indicated by dashed lines, and bound water molecules are indicated by filled circles.

deoxy Hb dimer−dimer interface. As seen in Figure 1, these waters are observed to mediate numerous residue−residue interactions and consequently should be important for predicting the free energy of dimer−tetramer association. Thus, for the two separate mutant crystal sets, crystallographically conserved hydration supersets are examined for atom−atom interactions at dimer−dimer interfaces. $\Delta\Delta G$ contributions of water molecules are determined by dividing the total HINT interaction score for each hydration superset by the predetermined number of HINT score units per kcal (as estimated from native deoxy Hb structures β37W[1998] and β37W[1992], vide supra). For mutants of the 1998 crystal set, a total of 27 interfacial waters are determined to make significant hydrogen bonding contacts at the dimer−dimer interface. HINT interaction totals for all hydration superset interactions with their respective Hb tetramers are shown in Table 2. On the basis of these total interaction scores, the average number of HINT score units for a single water molecule in the native Hb β37W[1998] and corresponding mutants β37Y, β37A, β37G, and β37E is 130, 138, 98, 110, and 86, respectively. For the β37Y mutation, a new water-mediated hydrogen bond between the mutant residue tyrosyl OH and residue Asp94α results in a slight increase in the total HINT score for this superset (Figures 2 and 3). The HINT total water contribution for the β37Y mutant is greater than the native Hb (β37W[1998]) total HINT score by 336 HINT score units (−0.7 kcal mol^{-1}). Lower HINT total interaction scores for interface waters of the β37G, β37A, and β37E mutant crystal structures are attributed to weaker overall hydrogen bonding (Table 2). This weaker bonding results from the far reaching structural consequences of these mutations, which give rise to a widening of the dimer−dimer interface (*36*), as reflected in dimer−dimer interface scores as calculated without structural waters (Table 1) and produce a weakening of all atom−atom interaction scores for these hemoglobin water supersets (Table 2). Moreover, the β37A HINT total interaction score for interface waters is lower still, as this mutant structure is missing two water molecules included in the hydration superset.

In both the native crystal structure β37W[1992] and its corresponding β37R mutant, a total interface superset of 26 waters is examined. The mutant hydration superset is missing three structural waters that are identified as important in the corresponding native deoxy Hb hydration network, but at

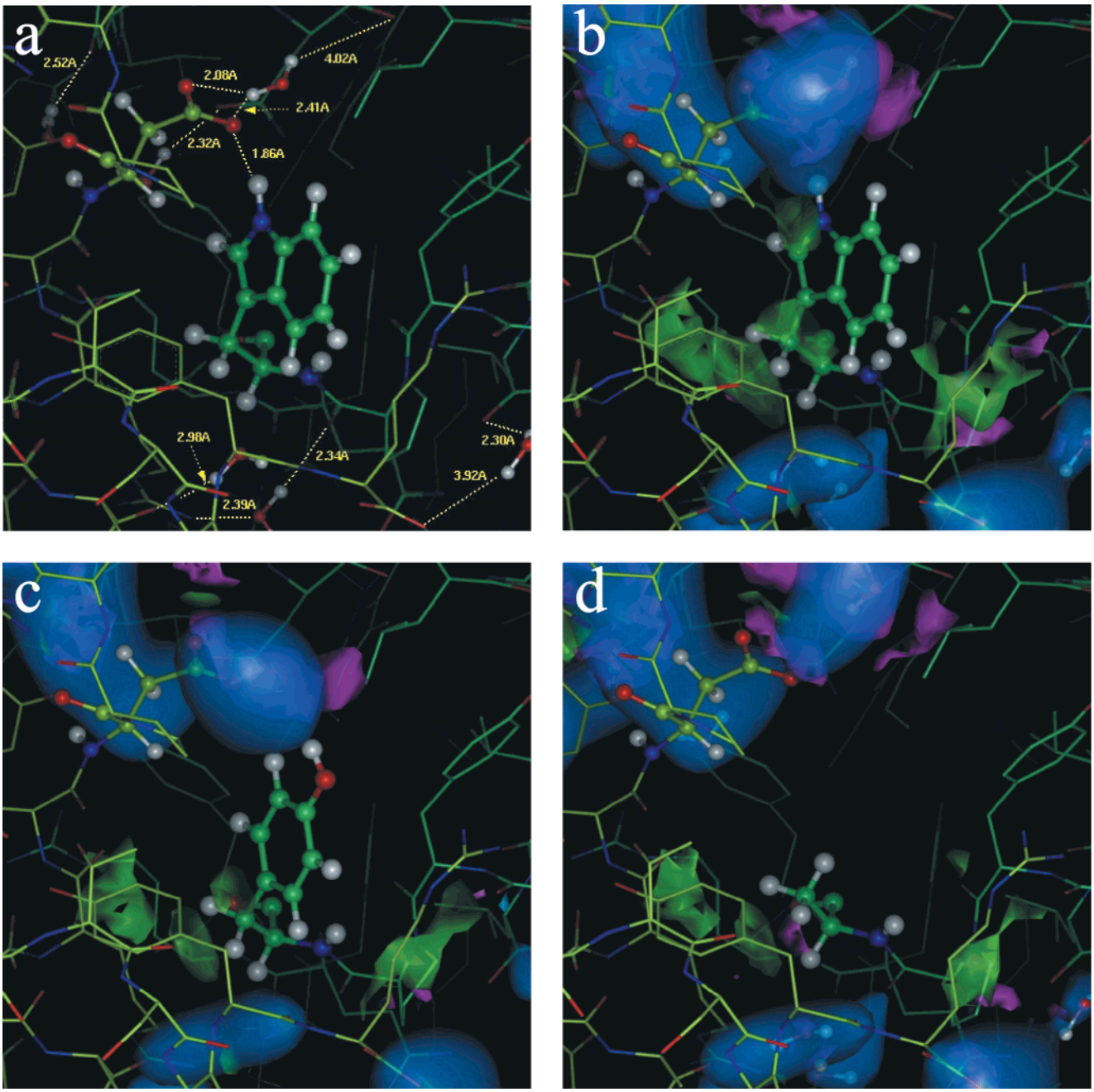

FIGURE 3: Molecular models of the localized region around native residue Trp37β and mutant residues Tyr37β and Ala37β in the $\alpha_1\beta_2$ interface. The carbon atoms of the α1 subunit are yellow-green and the carbon atoms of the β2 subunit are blue-green. Residue 37β is depicted with large ball-and-stick, while its main intersubunit contact, residue Asp94α, and water molecules are depicted in smaller ball-and-stick. (a) For comparison, the intersubunit region in the vicinity of native residue Trp37β is displayed without contours: intersubunit hydrogen bonds are labeled in yellow. Subfigures (b), (c), and (d) display HINT maps summing the residue−residue and residue−water interactions in a localized region around native residue Trp37β and mutants Tyr37β and Ala37β, respectively. HINT interaction contour maps display, visually, the quality and magnitude of the intersubunit contacts. The contour surfaces are color-coded by interaction type and the relative volume of the enclosed contour surface can be correlated with the relative magnitude of the interaction. Blue surfaces represent favorable polar−polar contacts, which are generally hydrogen bonds; green surfaces represent regions where there are favorable hydrophobic−hydrophobic contacts; magenta surfaces indicate unfavorable hydrophobic-polar contacts. (b) Intersubunit contacts by and surrounding native residue Trp37β. (c) Intersubunit contacts by and surrounding mutant residue Tyr37β. The tyrosine maintains a water mediated hydrogen bond with residue Asp94α. (d) Intersubunit contacts by and surrounding mutant residue Ala37β. All favorable bonding interactions with residue Asp94α are lost, while hydrophobic-polar interactions with Tyr140α have evolved.

the same time, the mutant hydration superset possesses three water molecules that are not in the native deoxy Hb hydration superset. HINT total atom−atom interaction scores show that the β37R mutant waters engage in more hydrogen bonding interactions overall (the Arg37β side chain guanidinium forms new hydrogen bonding interactions with surrounding water molecules) but at the same time this mutant residue moiety also engages in numerous unfavorable acid−acid interactions with surrounding waters. The net effect of these unfavorable interactions is that the native β37W[1992] water superset is, as a whole, more stable than the mutant hydration superset.

For mutant structures of both crystal sets, differences in hydration superset HINT total scores were converted to ΔΔ*G*

Table 2: Summary of HINT Total Interface Scores for Interactions between Hb Tetramers and Hydration Supersets, and HINT $\Delta\Delta G$ Estimates vs Experimental $\Delta\Delta G$

hemoglobin	total HINT interface score for hydration supersets[a]	HINT $\Delta\Delta G$ (kcal mol^{-1}) hydration supersets	interface w/o water	total	experimental $\Delta\Delta G$[b]
β37W^{1998}	3516				
β37Y	3862	−0.7	1.2 (1.9[c])	0.5 (1.2[c])	5.3
β37Y[d]	3530	0.7	2.1 (2.9[c])	2.8 (3.6[c])	5.3
β37A	2753	1.5	2.8	4.3	6.0
β37G	3076	0.8	4.2	5.0	6.1
β37E	2419	2.1	4.1	6.2	9.2
β37W^{1992}	3108				
β37R	1700	2.5	−1.7	0.8	1.0

[a] Sum of all observed hydration superset HINT interactions, including hydrogen bonding, acid−base, acid−acid, base−base, and hydrophobic−polar. [b] Experimental $\Delta\Delta G$ calculated in kcal mol^{-1} (*38*, *39*). [c] normalized for subunit interface $\alpha_1\beta_2$ residue−residue interaction α40K−β146H. [d] HINT scores and $\Delta\Delta G$ predicitions for β37Y as a tyrosinate.

predictions with respect to native structures (Table 2). Examination of the effects of these waters indicates that they had a significant effect on observed dimer−dimer stability. When changes in free energy associated with interface waters are added to free energy changes observed for dimer−dimer assembly not including structural waters, the resulting HINT $\Delta\Delta G$ predictions correlate more closely with experimental thermodynamic results, averaging within 2.0 kcal mol^{-1} of corresponding experimental $\Delta\Delta G$ predictions (Table 2). Subsequently, final HINT $\Delta\Delta G$ predictions for mutants β37A and β37G differ from experimentally determined $\Delta\Delta G$ estimates by −1.7 and −1.1 kcal mol^{-1}, respectively, while the final HINT $\Delta\Delta G$ prediction for mutant β37E differs from the experimental $\Delta\Delta G$ estimate by −3.0 kcal mol^{-1}. The mutant β37Y HINT $\Delta\Delta G$ prediction produces the least accurate free energy prediction, differing from experimental $\Delta\Delta G$ results by −4.8 kcal mol^{-1} (−4.1 kcal mol^{-1} with normalized α40K−β146H interaction). Finally, the β37R mutant HINT $\Delta\Delta G$ prediction differs from the experimental $\Delta\Delta G$ estimate by −0.2 kcal mol^{-1}.

The β37R mutant, also known as mutant Rothschild, is of particular interest in this study, as both current and previous HINT analyses of this crystal structure, without including structural waters, indicate a mutant residue related stabilization to the tetramer, in direct conflict with experimental measurements which indicate a thermodynamic destabilization of the dimer−dimer interface. This destabilization results in an increased oxygen affinity mutant due to an increased level of dissociated $\alpha\beta$ dimers (*39*). Interestingly, our analysis of the binding energy of dimer−dimer interface waters for this mutant shows that they contribute less free energy to the tetramer overall, compared to corresponding native Hb waters, suggesting that the binding energy of these waters accounts for 100% of the experimentally observed thermodynamic destabilization associated with this mutant.

Free Energy Estimate for a Single Water Molecule. Inclusion of water molecules in the HINT analysis of native and mutant hemoglobins clearly indicates that they are important for more accurate thermodynamic predictions. Furthermore, while the basic HINT equation implicitly includes the thermodynamic effects of bulk water on any examined biomolecular system, our results show that structurally conserved water molecules, in three-dimensional structures, transcend the general bulk water effect. Thus, these small molecules, in effect, act as multiple ligands that additively contribute a significant amount of binding energy during biomolecular associations.

Examination of dimer−dimer interface waters from the two separate deoxy Hb native structures provides the basis for an estimate of the average energy contribution of a single water molecule. For each structure, an average of 26.5 water molecules are determined to contribute to the free energy process during dimer−dimer association, with the average contribution of a single water molecule quantitating as 130 HINT score units for the β37W^{1998} structure and 135 HINT score units for the β37W^{1992} structure. From these two estimates, our method indicates that the average free energy contribution of a single interface water molecule is approximately −0.3 kcal mol^{-1}.

Previous studies have estimated the change in free energy contributed by a single water molecule to range from −0.2 to −2.0 kcal mol^{-1} (*27*). A more recent investigation by Hisler, Gomez, and Freire (*50*) estimated an enthalpic contribution of −0.5 kcal mol^{-1} for a single water molecule participating in the association of the D1.3-HE antibody-hen egg white lysozyme complex. Our energy estimate for a single water (−0.3 kcal mol^{-1}) molecule is approximately 60% of this estimate and may reflect the fact that our analysis not only considers the favorable enthalpic contributions of each water molecule but also includes the unfavorable hydrophobic-polar interactions that are ubiquitously incurred during the binding of these small molecules to a biomolecular interface, and which are certainly at least partially entropic in origin. When our estimate for the average free energy of a single water molecule is calculated with only enthalpic binding forces, the average free energy increases to −0.8 kcal mol^{-1}, in closer agreement with previous calculations.

Closer Examination of Mutant β37Y. Experimental thermodynamic analysis indicates that this mutation results in 5.3 kcal mol^{-1} thermodynamic destabilization compared to dimer−tetramer assembly for native deoxy Hb. In contrast, the thermodynamic destabilization predicted by HINT analysis of this mutant is much more conservative, indicating a destabilization of only 1.2 kcal mol^{-1} when examined without water. However, as indicated earlier, it should be noted that the $\alpha_1\beta_2$ interface for this mutant crystal structure gives a higher than expected total HINT interface score and is the result of a stronger than usual interaction between interface residues Lys40α and His146β. When this interaction is normalized, with respect to the same observed interaction

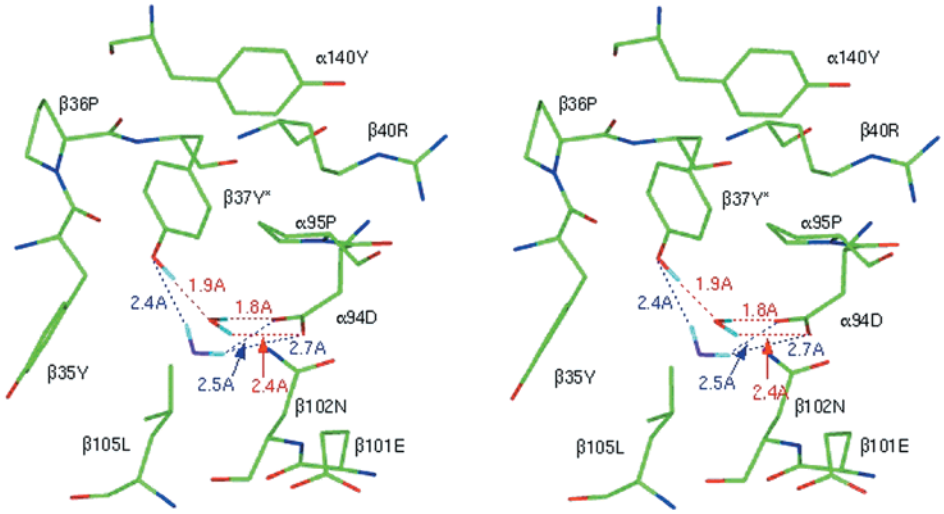

FIGURE 4: Stereo diagram showing the environment of superimposed Tyr37β tyrosine and tyrosinate residues in the $\alpha_1\beta_2$ interface. Carbons are green, protein oxygens are red, nitrogens are dark blue, and hydrogens are light blue. Hydrogen bonds are indicated by dashed lines. A key hydrogen bond donating water for the tyrosine mutant is indicated with a red oxygen, while the same water interacting with the tyrosinate mutant is indicated by a purple oxygen.

in the native structure β37W[1998], the predicted destabilization for this mutant is 1.9 kcal mol^{-1}.

Analysis of the effects of interdimer waters on the thermodynamic stability of this tetramer indicates that these waters thermodynamically stabilize this mutation, relative to the native deoxy Hb structure, by −0.7 kcal mol^{-1}. Detailed analysis of β37Y interface water interactions indicates that the Tyr37β OH forms new hydrogen bonds with an adjacent water (Figures 3 and 4). This bonding is not observed between the native Hb Trp37β indole NH and an equivalent water molecule due to steric constraints and a strong hydrogen bond with the carboxylate of residue Asp94α (Figure 2). Thus, with the ΔΔ*G* prediction for interfacial waters included, the total ΔΔ*G* prediction for this mutant is only 0.5 kcal mol^{-1} using our method (1.2 kcal mol^{-1} with normalized subunit interface $\alpha_1\beta_2$ residue−residue Lys40α−His146β interaction). This prediction grossly underestimates the change in experimental free energy during dimer−tetramer assembly for this mutant. However, our results do seem to corroborate oxygen-binding studies, which suggest that unlike other β37 mutants, β37Y is comparatively similar to native deoxy hemoglobin, displaying less tetramer dissociation and cooperative oxygen binding, which is of approximately the same magnitude as native Hb tetramers (*38*).

To further investigate the lack of thermodynamic instability predicted for this mutation, the Tyr37β residue of the β37Y crystal structure was computationally modeled as a tyrosinate. The effect of placing an ionized tyrosine in this position is extensive with respect to both $\alpha_1\beta_2$ and $\alpha_2\beta_1$ interface contacts, as well as dimer−dimer interface water stability. Examination of the tyrosinate mutation without water indicates that it would thermodynamically destabilize the dimer−dimer interface by a total of 2.2 kcal mol^{-1} (Table 1). Furthermore, the effect of this ionized mutation on interface waters would result in a further loss of 0.7 kcal mol^{-1}. In this instance, the ionized mutant residue effects on water shows that the protons of the water molecule bridging between the tyrosinate and the Asp94α side chain carboxylate would be rotated by 90° to facilitate an interaction with the tyrosinate (Figure 4). This rotation not only would lead to an overall decrease in hydrogen bonding between this water and the Asp94α carboxylate but also would result in increased unfavorable interactions between this water and surrounding protein residues. Thus, increases in unfavorable interactions, directly attributed to the tyrosinate, would result in a total ΔΔ*G* of 2.9 kcal mol^{-1} for this ionized species. When this thermodynamic destabilization is calculated with an $\alpha_1\beta_2$ interface score that has been normalized for the previously described Lys40α−His146β interaction, the ΔΔ*G* prediction for this mutant would be 3.6 kcal mol^{-1}, in significantly better agreement with experimental data. While the scope of this research cannot conclusively determine that the Tyr37β mutant residue exists as an ionized species, resonance Raman studies on human hemoglobin and horse and sperm whale myoglobins have indicated that the pK_a values of tyrosine residues are highly dependent on their local protein environment (*52*−*54*). Specifically, it has been shown that tyrosine residues located in polar and solvent accessible regions of these proteins have lower pK_a values than tyrosine residues buried in hydrophobic/solvent inaccessible protein pockets (*52*). The Tyr37β residue, in the deoxy Hb tetramer, is located in a highly polar region of the dimer−dimer interface, surrounded by residues Glu101β, Arg40β, Asp94α, and Asp99α and is a solvent-exposed residue on the αβ dimers prior to and after tetrameric assembly. Thus, it may be postulated that the pK_a of this mutant residue is considerably lower than commonly observed, and that it may exist as an ionized species during the process of thermodynamic characterization under aqueous conditions. However, based solely on the mutant β37Y crystal structure used in this study, and the neutral pH under which it was crystallographically solved, there is no absolute proof to corroborate that this residue exists as an ionized species.

Computationally Derived β37W Mutant Hemoglobins. Error Estimates for Model-based Free Energy Predictions. To evaluate our ability to accurately generate and thermodynamically evaluate de novo protein models, crystallographically and thermodynamically characterized structures

Table 3: Summary of HINT ΔΔ*G* Estimates for Computationally Derived β37 Mutants

hemoglobin[a,b,c]	HINT ΔΔ*G* w/o water[d]	HINT water ΔΔ*G*[d]	total HINT ΔΔ*G*[d]	experimental ΔΔ*G*[e]
β37W[1998]				
β37W-Y	1.3	−0.2	1.1	5.3
β37W-A	2.9	0.9	3.8	6.0
β37W-G	2.7	0.7	3.4	6.1
β37W-E	4.4	−0.1	4.3	9.2
β37W-R	−2.3	0.5	−1.8	1.0
β37W-S	1.1	1.2	2.3	5.5
β37W-S	1.8	1.1	2.9	5.5
β37W-F	1.1	−0.5	0.6	
β37Y-F	1.6 (2.3[f])	0.6	2.2 (2.9[f])	
β37W-T	1.7	1.3	3.0	
β37A-T	2.9	2.0	4.9	

[a] β37W-, mutants computationally generated from native deoxy Hb β37W[1998]. [b] β37A-, mutants computationally generated from mutant deoxy Hb β37A. [c] β37Y-, mutants computationally generated from mutant deoxy Hb β37Y. [d] HINT ΔΔ*G* predictions calculated in kcal mol^{-1}. [e] Experimental ΔΔ*G* calculated in kcal mol^{-1} (*38*, *39*). [f] Normalized for subunit interface $\alpha_1\beta_2$ residue−residue interaction α40K−β146H.

in the text of this work are computationally modeled from the native deoxy Hb structure β37W[1998]. Furthermore, crystallographically unsolved deoxy Hb mutants β37S (*55*), β37T (*47*), and β37F (*56*) are modeled from native deoxy Hb structure β37W[1998] as well as from structurally similar mutant β37A and β37Y crystal structures. Thermodynamic predictions for these species are based on combined intersubunit HINT interface scores not including structural waters and HINT interaction scores from interface waters within a 5 Å hydration shell around the computationally modeled mutant residues.

In general, HINT ΔΔ*G* predictions are very comparable for modeled mutants (Table 3) with respect to HINT ΔΔ*G* predictions for corresponding crystallized mutants (Table 1 vs Table 3). However, the de novo computationally modeled β37 mutants failed to reproduce the long-range structural effects observed during examination of crystallographic mutant structures (largely at the $\alpha_1\alpha_2$ interface). The omission of these effects, which clearly contribute to the thermodynamic destabilization observed in the corresponding mutant crystal structures, results in an average error of approximately 3.0 kcal mol^{-1} for HINT predictions based on computationally generated models (vs 2.0 kcal mol^{-1} for HINT predictions based on experimentally solved crystal structures).

Interface waters within a 5 Å hydration shell around the 37β position of computational mutant models may be thought of as a microcosm of the total interface waters examined in the analysis of the individual mutant crystal structures and provide a good approximation of water contributions to the thermodynamic stability/instability resulting from modeled mutants. For mutants with solved crystal structures, the most substantial deviations between a computer generated mutant model water ΔΔ*G* prediction and a crystal structure hydration superset ΔΔ*G* prediction occur for mutants β37W-E and β37W-R. The observed deviation for the β37E mutant is a result of the movement of a bridging water between the Asp94α and Glu37β carboxylate moieties; the observed deviation for the β37R mutant is most likely a result of the large changes associated with the interfacial water lattice of this mutant (as revealed by its crystallographic structure) (*37*), which are not observed when de novo modeling this mutant from the native Hb crystal structure.

HINT ΔΔ*G* predictions for mutant models of β37S, β37F, and β37T are also presented in Table 3. The effects of these mutations at $\alpha_1\beta_2$ and $\alpha_2\beta_1$ interfaces conform to the general mutant effects observed during examination of β37 mutant crystal structures, with each analyzed synthetic mutant structure displaying decreased total dimer−dimer interface energies as a result of both lost favorable interactions with the Asp94α carboxylate and increased unfavorable interactions with surrounding residues. We would be interested to compare our thermodynamic predictions to measured ΔΔ*G* values for the β37F and β37T mutants when these data become available.

CONCLUSIONS

This study indicates that the HINT method, a molecular model based on experimental thermodynamic data, may be used as a highly descriptive tool for the prediction of binding free energies in biomolecular systems. We have focused on translating quantitated atom−atom interactions in Hb dimer−dimer interfaces into predictions of mutant-induced thermodynamic change, showing that direct mutant residue interactions, mutant-induced long-range structural changes, and differences in the stability of bound water molecules all contribute to free energy changes observed for each mutant structure.

Other unique conclusions include the following: (1) Thermodynamic estimates of dimer−tetramer association free energy, generated from the detailed analysis of high resolution β37 mutant Hb crystal structures, correlated well with corresponding experimental thermodynamic measurements taken by Kiger et al. (*38*) and Turner et al. (*39*), which indicated that mutations of the Trp37β position result in large thermodynamic destabilizations to Hb tetramer stability. (2) Hydropathic analysis of the binding interactions of important crystallographic water molecules at mutant β37 dimer−dimer interfaces indicated that bound waters may account for up to 100% of observed free energy changes, and on average accounted for approximately 15% of the total estimated free energy change. These results were in good agreement with findings from a previous analyses on protein−protein association energies conducted by Covell and Wallqvist (*26*). (3) Computationally generated β37 mutant models accurately reproduce the localized structural and thermodynamic mutant effects observed in corresponding crystallographically solved β37 mutants, but fail to reproduce the long-range secondary structural changes that were observed in the same crystallographic structures. Thus, computer-generated mutant models do not generate thermodynamic predictions as accurately as unique mutant crystal structures and incur an average error of approximately 3.0 kcal mol^{-1} that is likely due to long-range, structural effects. However, most computer-generated models produced hydration effects that agreed qualitatively with the same effects observed in corresponding mutant crystal structures.

ACKNOWLEDGMENT

We wish to thank Paolo Botti for his help in the development of our computational methods and Dr. Neel

Scarsdale and Dr. Jason Rife for helpful discussions regarding the computational preparation of mutant models. We also wish to acknowledge the support of VCU and the Institute for Structural Biology and Drug Discovery.

REFERENCES

1. Frank, H., and Evans, M. (1945) *J. Chem. Phys. 13*, 507−532.
2. Kauzmann, W. (1959) *Adv. Protein Chem. 14*, 1−63.
3. Janin, J., and Chothia, C. (1990) *J. Biol. Chem. 265*, 16027−16030.
4. Perutz, M. (1992) in *Protein Structure*: *New Approaches to Disease and Therapy* pp 41−76, W. H. Freeman and Company, New York.
5. Buckle, A. M., Schreiber, G., and Fersht, A. R. (1994) *Biochemistry 33*, 8878−8889.
6. Novotony, J., Bruccoleri, R. E., and Saul, F. A.(1989) *Biochemistry 28*, 4735−4749.
7. Cunningham, B. C., and Wells, J. A. (1993) *J. Mol. Biol. 234*, 554−563.
8. Tulip, W. R., Harley, U. R., Webster, R. G., and Novotony, J. (1994) *Biochemistry 33*, 7986−7997.
9. Clackson, T., and Wells, J. A. (1995) *Science 267*, 383−386.
10. Dall'Acqua, W., Goldman, E., Eisenstein, E., and Mariuzza, R. A. (1996) *Biochemistry 35*, 9667−9676.
11. Colombo, M. F., Rau, D. C., and Parsegian, V. A. (1992) *Science 256*, 655−659.
12. Timasheff, S. N. (1992) *Biochemistry 31*, 9857−9864.
13. Mariuzza, R. A., and Poljak, R. J. (1993) *Curr. Opin. Immunol. 5*, 50−55.
14. Timasheff, S. N. (1993) *Annu. Rev. Biophys. Struct. 22*, 67−97.
15. Karplus, P. A., and Paerman, C. (1994) *Curr. Opin. Struct. Biol. 4*, 770−776.
16. Lavoie, T. A., and Carey, J. (1994) *Nucleic Acids Mol. Biol. 8*, 184−196.
17. Braden, B. C., and Poljak, R. J. (1995) *FASEB J. 9*, 9−16.
18. Poornima, C. S., and Dean, P. M. (1995) *J. Comput.-Aided Mol. Des. 9*, 500−512.
19. Goldbaum, F. A., Schwartz, F. P., Eisenstein, E., Cauerhff, A., Mariuzza, R. A., and Poljak, R. J. (1996) *J. Mol. Recognit. 9*, 6−12.
20. Wang, H., and Ben-Naim, A. (1996) *J. Med. Chem. 39*, 1531−1539.
21. Horton, N., and Lewis, M. (1992) *Protein Sci. 1*, 169−181.
22. Rao, B. G., Tilton, R. F., and Covell, D. G. (1995) *Protein Sci. 4*, 1881−1903.
23. Ajay, and Murcko, M. A. (1995) *J. Med. Chem. 38*, 4953-4967.
24. Wallqvist, A., Jernigan, R. L., and Covell, D. G. (1995) *Protein Sci. 4*, 1881−1903.
25. Babine, R. E., and Bender, S. L. (1997) *Chem. Rev. 97*, 1359.
26. Covell, D. G., and Wallqvist, A. (1997) *J. Mol. Biol. 269*, 281−297.
27. Holdgate, G. A., Tunnicliffe, A., Ward, W. H. J., Weston, S. A., Rosenbrock, G., Barth, P. T., Taylor, I. W. F., Pauptit, R. A., and Timms, D. (1997) *Biochemistry 36*, 9663−9673.
28. Tawa, G. J., Topol, I. A., Burt, S. K., and Erickson, J. W. (1998) *J. Am. Chem. Soc. 120*, 8856−8863.
29. Hansch, C., Steward, A. R., Anderson, S. M., and Bentley, D. (1968) *J. Med. Chem. 11*, 1−11.
30. Helmer, F., Kiehs, K., Hansch, C. (1968) *Biochemistry 7*, 2858−2863.
31. Kellogg, G. E., Joshi, G. S., and Abraham, D. J. (1992) *Med. Chem. Res. 1*, 444−453.
32. Gussio, R., Pattairaman, N., Zaharevitz, D. W., Topol, I. A., Rice, W. G., Schaeffer, C. A., Erickson, J. W., and Burt, S. K. (1996) *J. Med. Chem. 39*, 1645−1650.
33. Kellogg, G. E., Scarsdale, J. N., and Fornari, F. A. Jr. (1998) *Nucleic Acids Res. 26*, 4721−4732.
34. Dill, K. A. (1997) *J. Biol. Chem. 272*, 701−704.
35. Abraham, D. J., Kellogg, G. E., Holt, J. M., and Ackers, G. K. (1997) *J. Mol. Biol. 272*, 613−632.
36. Kavanaugh, J. S., Weydert, J. S., and Ackers, G. K. (1998) *Biochemistry 37*, 4358−4373.
37. Kavanaugh, J. S., Rogers, P. H., Case, D. A., and Arnone, A. (1992) *Biochemistry 31*, 4111−4121.
38. Kiger, L., Klinger, A. L., Kwiatkowski, L. D., Young, A. D., Doyle, M. L., Holt, J. M., Noble, R. W., and Ackers, G. K. (1998) *Biochemistry 37*, 4336−4345.
39. Turner, G. J., Galacteros, F., Doyle, M. L., Hedlund, B., Pettigrew, D. W., Turner, B. W., Smith, F. R., Moo-Penn, W., Rucknagel, D. L., and Ackers, G. K. (1992) *Proteins*: *Struct., Funct., Genet. 14*, 333−350.
40. Doyle, M. L., Lew, G., De Young, A., Kwaitkowski, L., Wierzba, A., Noble, R. W., and Ackers, G. K. (1992) *Biochemistry 31*, 8629−8639.
41. Kavanaugh, J. S., Rogers, P. H., and Arnone, A. (1992) *Biochemistry 31*, 8640−8647.
42. *Insight II. Version 97.2* Biosym Technologies, San Diego, California.
43. GRID V.17, Molecular Discovery Ltd., West Way House, Elms Parade, Oxford, 1998.
44. Abraham, D. J., and Leo, A. J. (1987) *Proteins*: *Struct., Funct., Genet. 2*, 130−152.
45. Hansch, C., and Leo, A. J. (1979) in *Substituent Constants for Correlation Analysis in Chemistry and Biology*. John Wiley and Sons, Inc., New York.
46. Isrealachvili, J., and Pashley, R. (1982) *Nature 300*, 341−342.
47. Vallone, B., Bellelli, A., Miele, A. E., Brunori, M., and Fermi, G. (1996) *J. Biol. Chem. 271*, 12472−12480.
48. Fermi, G., Perutz, M. F., Shaanan, B., and Fermi, R. (1984) *J. Mol. Biol 175*, 159−174.
49. Bhat, T. N., Bentley, G. A., Boulot, G., Greene, M. I., Tello, D., Dall'Acqua, W., Souchon, S., Schwartz, F. P., Mariuzza, R. A., and Poljak, R. J. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 1089−1093.
50. Hisler, V. J., Gomez, J., and Freire, E. (1996) *Proteins*: *Struct., Funct., Genet. 26*, 123−133.
51. Royer, W. E. Jr., Pardanani, A., Gibson, Q. H., Peterson, E. S., and Friedman, J. M. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 14526−14531.
52. Asher, S. A., Larkin, P. S., and Teroka, J. (1991) *Biochemistry 30*, 5944−5954.
53. Rodgers, K. R., Su, C., Subramaniam, S., and Spiro, T. G. (1992) *J. Am. Chem. Soc. 114*, 3697−3709.
54. Huang, S., Peterson, E. S., Ho, C., and Freedman, J. M. (1996) *Biochemistry 36*, 6197−6206.
55. Yamaoka, K. (1971) *Blood 38*, 730−738.
56. Ishimori, K., Imai, K., Miyazaki, G., Kitagawa, T., Wada, Y., Morimoto, H., and Morishima, I. (1992) *Biochemistry 31*, 3256−3264.

BI991724U