α-Helix Stability and the Native State of Myoglobin†

Laura Lin, Rachel J. Pinker, and Neville R. Kallenbach*

Department of Chemistry, New York University, New York, New York 10003

Received July 15, 1993; Revised Manuscript Received August 31, 1993®

ABSTRACT: Native proteins fold to form structures that contain secondary-structure regular patterns in the peptide backbone, such as α-helix, β-structure, and turns with high frequency. The role of this secondary structure in stabilizing the native folded state is presently unclear. Alanine substitutions at helical sites in myoglobin show no correlation with the helical propensity of the side chains involved. In an effort to demonstrate a relationship between the effect of a side chain on stabilizing secondary structure and the native structure, we have carried out site-directed changes in the sequence of the helical protein sperm whale myoglobin. Fully buried hydrophobic side chains were exchanged for similar side chains at sites corresponding to midhelical positions in the native state. The results show a positive correlation between the α-helix-forming ability of the substituted side chain and the stability of the mutant proteins, when differences between the size of the side chains are taken into account. If, in addition, each type of amino acid substitution is averaged over different sites, the helix propensities of the amino acids account for much of the residual variation. This implies that the stability of the native state of a protein is coupled to that of secondary structural elements in the structure. In magnitude, the net contribution of propensity differences is smaller than hydrophobic effects, but not negligible in terms of the net free energy of unfolding.

Secondary structure is a common feature in known protein structures (Fasman, 1989). However, the role of secondary structure in formation of the native state of proteins and its contribution to the stability of the native state have been difficult to establish. On the one hand, hydrophobic interactions and packing constraints dominate the thermodynamics of folding a chain into its native state (Murphy et al., 1990), suggesting that secondary structure may play a minor role in stabilizing native proteins (Dill, 1990). On the other hand, there is mounting evidence that proteins can assume molten globule states in which nativelike secondary structure forms, in the absence of tight packing of many side chains (Kuwajima, 1989; Jeng & Englander, 1991; Ptitsyn, 1992; Barrick & Baldwin, 1993). The fact that protein fragments or peptides can acquire helical structure in solution is consistent with the idea that secondary structure can arise independently of packing constraints. Experiments on short α-helical peptides (O'Neil & Degrado, 1990; Lyu et al., 1990; Padmanabhan et al., 1990) reveal significant differences in helix formation among chemically similar side chains: for example, the nonpolar side chains, Ala, Leu, and Met, are helix forming while Ile and Val are weak. Quantitative scales of the stabilizing effect of these side chains in the middle of an α–helix have been derived (O'Neil & Degrado, 1990; Lyu et al., 1990; Gans et al., 1991), specifying the Zimm–Bragg *s* values for helix propagation (Zimm & Bragg, 1959) or the intrinsic α-helical propensities, for each side chain.

The role of secondary structure in stabilizing the native state of proteins has been addressed experimentally in a number of studies of amino acid substitutions in model proteins (Hecht et al., 1986; Horovitz et al., 1992; Blaber et al., 1993). A difficulty in these efforts is that the α-helices in globular proteins tend to be short, 12 residues on average, placing many sites of interest near the N- or C-terminus of a helix (Presta & Rose, 1988; Richardson & Richardson, 1988). In such cases, the scale of helical propensities is position dependent because of "capping" interactions at the ends of a helix (Serrano & Fersht, 1989; Lyu et al., 1992). A highly helical protein such as myoglobin (Phillips et al., 1990), with five long helical regions, offers a favorable system for analysis of the role of secondary structure. We initially set out to test the relationship between α-helix stability and that of native myoglobin by preparing a series of alanine substitutions at helical sites on the surface of the molecule (Pinker et al., 1993). The results do not reveal a correlation with differences in midhelix propensities (*s* values; O'Neil & Degrado, 1990; Gans et al., 1991) between alanine and the side chains substituted. In most cases, substituting Ala for a helical side chain on the surface actually destabilizes the native protein. The reason appears to be that differences in the buried surface area of the protein before and after alanine substitution contribute more to the stability of the mutant proteins than differences in helix propensity. Quantitative scales of midhelix propensity indicate that Ala differs from Gly by less than 1 kcal/mol (O'Neil & Degrado, 1990; Gans et al., 1991; Horovitz et al., 1992; Blaber et al., 1993). An analysis of α–helix propensities of side chains at two sites in T4 lysozyme reveals a free energy difference of about this magnitude, although there is a correlation with the hydrophobicity of side chains (Blaber et al., 1993), consistent with the idea that differences in buried surface area influence the contribution of surface side chains to the stability of helices in proteins (Pinker et al., 1993).

This led us to consider a strategy in which the extent of exposure of each site chosen for substitution is controlled more precisely. This is most directly achieved by substituting side chains at internal sites that have zero exposure in the native state (see Table I). Ala substitutions for bulkier side chains at internal sites are generally destabilizing (Shortle et al., 1990; Eriksson et al., 1992) since they can perturb packing; we used these as controls, focusing on substitutions among bulky Leu, Met, Ile, and Val side chains at fully buried helical sites in myoglobin. When intrinsic differences in the size of the side chains are taken into account, the stability of the

† This work was supported by Grant GM 40746 from the NIH and a Patricia Roberts Harris Fellowship to R.J.P.

* Address correspondence to this author.

® Abstract published in *Advance ACS Abstracts,* November 1, 1993.

0006-2960/93/0432-12638$04.00/0

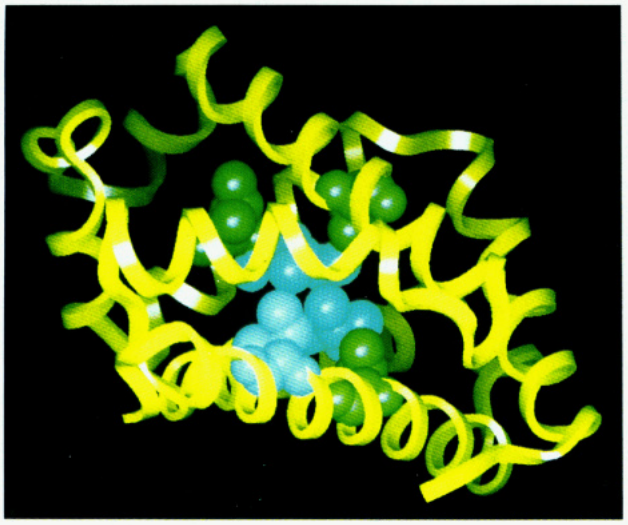

FIGURE 1: Identity of mutant sites in the core of sperm whale myoglobin. Coordinates are taken from 1MBW (Phillips et al., 1990) in the Brookhaven Protein Databank. The heme group is not shown for clarity. Color codes: blue for isoleucine, green for leucine.

mutant proteins now shows a positive correlation with the tendency of these amino acids to stabilize α-helical structure (Figure 2A). If, in addition, the data for each type of substitution are averaged over different sites in the protein, then the helix propensity correlates well with the stability of the mutant proteins (Figure 2B). This correlation suggests that stabilizing secondary structure indeed stabilizes the native state of myoglobin and hence that the stability of secondary structure in proteins is coupled to that of the native fold.

MATERIALS AND METHODS

Mutant Selection. Crystal structures of sperm whale myoglobin (1MBO, 1MBC, 1MBD, 1MBW, 4MBN, 5MBN) in the Brookhaven Protein Databank (Bernstein et al., 1977) were examined using an IRIS work station and the molecular graphics package Insight II. Three Leu sites were chosen for alteration to Ile, Met, Val, and Ala, and three Ile sites were chosen for alteration to Leu, Met, Val, and Ala (Figure 1). Each of these is a midhelical site, more than four residues removed from the N- or C-termini of a helix, to avoid helix capping effects mentioned which affect the propensity of side chains near the ends of α-helices in both peptides and proteins (Presta & Rose, 1988; Richardson & Richardson, 1988; Serrano & Fersht, 1989; Lyu et al., 1992). Solvent accessibility calculations using the algorithm of Lee and Richards (Richards, 1977; Lesser & Ross, 1990) indicate that each site is fully buried in the native myoglobin molecule, without making direct contact with the heme. An unpublished program written by George Rose was used to evaluate the depth of burial of the selected side chains in the protein, by peeling successive layers of atoms that are exposed on the outer surface. All the sites selected are deeply buried. We used Leu 49 and Ile 99 as controls since they are deeply buried but not located in α-helices.

Site-Directed Mutagenesis and Protein Purification. Mutagenesis of the synthetic gene for sperm whale myoglobin from plasmid pMb413 (Springer & Sligar, 1987) was carried out using the procedure of Kunkel (1985). The gene was cloned into a phage M13 vector. Oligonucleotides (10 ng) containing the codons for leucine (CTG), isoleucine (ATC), methionine (ATG), valine (GTT), or alanine (GCT) were annealed with a uridine-containing phage template. The desired mutations were verified by DNA dideoxy sequencing (Sanger et al., 1977). Mutant genes were subcloned into the vector pUC19 and transformed into *Escherichia coli* strain TB-1. Proteins were isolated as described by Springer and Sligar (1987). The purity of the mutant proteins was monitored by SDS–gel electrophoresis; all the proteins used were at least 95% homogeneous by this assay. The heme absorbance and the CD spectrum of each mutant protein have been monitored and found to be equivalent to those of the native protein (data not shown). Thus neither the heme environment nor the helix content of the mutant proteins differs from those of native unsubstituted myoglobin.

Thermal Denaturation. Thermal unfolding experiments on all the proteins were carried out in 40 mM glycine–KOH buffer at pH 11.0, to avoid precipitation of the unfolded protein (Acampora & Hermans, 1967), in the presence of 0.5 mM KCN as ligand. Samples were dialyzed extensively against the above buffer for 24–36 h. Protein concentrations were 0.1–0.05 mg/mL for the melting curves and around 1 mg/mL for DSC experiments. Heme absorbance in the visible (Soret band) and circular dichroism at 222 nm were monitored as a function of temperature. Temperature was increased at a uniform rate of 1 °C/min. Visible melting profiles were obtained using a thermoelectrically controlled Perkin-Elmer 552 spectrophotometer interfaced to a PC-XT computer for acquisition and analysis of experimental data. The thermal unfolding curves were analyzed by means of a two-state transition model to give the temperature of melting, T_m, and enthalpy of unfolding, ΔH, at T_m (Acampora & Hermans, 1967). Differential scanning calorimetry was performed on the wild-type protein and selected mutants to calibrate the thermodynamic parameters (Privalov & Khechinashvili, 1974; Privalov et al., 1986; Nozaki & Tanford, 1971). For a mutant with a transition ΔT_m from the wild-type protein, with midpoint temperature $T_{m,wt}$, $\Delta\Delta G°$ is calculated using the equation

$$\Delta\Delta G°(\text{kcal/mol}) = -\Delta T_m[\Delta H_{wt} + \Delta C_p \Delta T_m]/T_{m,wt}$$

where ΔC_p = 2.8 kcal deg^{-1} mol^{-1} for myoglobin (Privalov et al., 1986) and ΔH_{wt} = 118 kcal/mol from DSC measurements in this laboratory. This equation assumes that the enthalpy of the mutant protein corresponds to that of the wild type at the unfolding temperature of the mutant protein, $T_{wt} + \Delta T_m$, corrected for the effect of ΔC_p. Thermal transitions at pH 11 are reversible on slowly cooling, with no evidence for decomposition of the protein (Acampora & Hermans, 1967).

Thermal Denaturation in the Presence of Urea. Addition of a nonreactive solvent such as urea or guanidinium chloride cooperatively unfolds proteins (Tanford, 1970; Schellman, 1978). The effect can be described in terms of preferential binding of the solvent to unfolded states or to differential solvation of the unfolded protein. In practice, the isothermal unfolding of a very stable molecule such as myoglobin by urea occurs at high urea concentrations (Puett, 1973), making it difficult to estimate the free energy of the molecule in the absence of denaturant. The temperature of unfolding as a function of urea concentration is easier to measure for several concentrations of urea and can be interpreted in terms of Schellman's linear free energy model to reconstruct the isotherms if these are needed. From thermodynamic arguments, Schellman shows that

$$\partial(T_m^{-1})/\partial C_3 = (R/\Delta H)\Delta b_{23}$$

where T_m is the unfolding temperature of the protein in the presence of a molar concentration C_3 of urea, ΔH is the enthalpy of unfolding, Δb_{23} is the differential free energy of solvation of the protein due to urea, and R is the gas constant.

Table I: Effect of Internal Side-Chain Substitutions on the Stability of Sperm Whale Myoglobin

mutants	positions	ΔT_m (°C)[a]	$\Delta\Delta G°$ (kcal/mol)[b]	$m' \times 10^5$ (mol^{-1} K^{-1})[c]
I28L	B-8	−1.7	0.55	2.30
I28M		−1.7	0.56	2.60
I28V		0.1	−0.04	2.28
I28A		−4.4	1.33	2.82
I111L	G-11	−2.0	0.64	2.53
I111M		−3.7	1.14	2.50
I111V		*d*		
I111A		−6.4	1.84	2.99
I142L	H-18	1.8	−0.63	2.36
I142M		2.6	−0.93	2.25
I142V		−0.3	0.12	2.34
I142A		−3.5	1.10	2.67
L29I	B-9	−3.6	1.12	2.36
L29M		0.3	−0.12	2.29
L29V		−6.0	1.72	2.67
L29A		−8.9	2.37	3.40
L69I	E-11	−0.1	0.02	2.14
L69M		0.0	0.00	2.41
L69V		−0.3	0.10	2.39
L69A		−3.8	1.18	2.71
L135I	H-11	−5.2	1.54	2.59
L135M		−2.5	0.79	2.71
L135V		−8.3	2.25	2.99
L135A		*d*		
L49I	CD-7	−2.5	0.80	*e*
I99L	FG-5	*d*		

[a] $\Delta T_m = T_{m,mutant} - T_{m,wildtype}$ is the average of triplicate determinations. [b] $\Delta\Delta G° = \Delta T_m/T_m[\Delta H_{wt} + \Delta C_p \Delta T_m]$ (kcal/mol), where T_m = 76.5 °C is the transition temperature of the wild-type myoglobin, ΔH_{wt} = 118 kcal mol^{-1}, and ΔC_p = 2.8 kcal deg^{-1} mol^{-1}. [c] See Figure 5 and text. Wild-type myoglobin gives a slope of 2.38×10^{-5} under the same condition. [d] The mutant proteins at these sites cannot be purified by normal procedure. [e] Not done.

With the approximation that the differential free energy of solvation is linear in the urea concentration itself

$$\Delta b_{23} = \Delta b_{23}{}^0 C_3$$

a plot of T_m^{-1} vs C_3 would be predicted to be linear, as we find, provided ΔH varies slowly with C_3 (see Figure 5 and Mayr and Schmid (1993)]. The slope of the line would be

$$m' = R\Delta b_{23}{}^0/\Delta H$$

and values of m' are given in Table I. Plots of T_m^{-1} using ln [a] as variable, where [a] is the urea activity, turn out to be curved (data not shown). From linear plots of T_m^{-1} vs C_3, the isotherm for unfolding a protein at 25 °C by urea can be reconstructed, assuming that ΔH at the T_m corresponding to a given urea concentration is controlled by the value of ΔC_p alone [see Santoro & Bolen (1992)]. The enthalpy change ΔH includes three contributions: $\Delta H = \Delta H° + h^U - h^N$, where $\Delta H°$ is the enthalpy of unfolding of the protein in the absence of urea and h^U and h^N are the excess enthalpies of the unfolded and folded states of the protein in the presence of urea.

RESULTS AND DISCUSSION

The effect of mutations at internal sites on the thermal stability of myoglobin is summarized in Table I. Only two out of 22 internal mutations significantly stabilized the protein, both in cases where the helical propensity would be expected to increase. In the case of one nonhelical site, the effect is opposite. I99L proves to be too unstable to allow purification

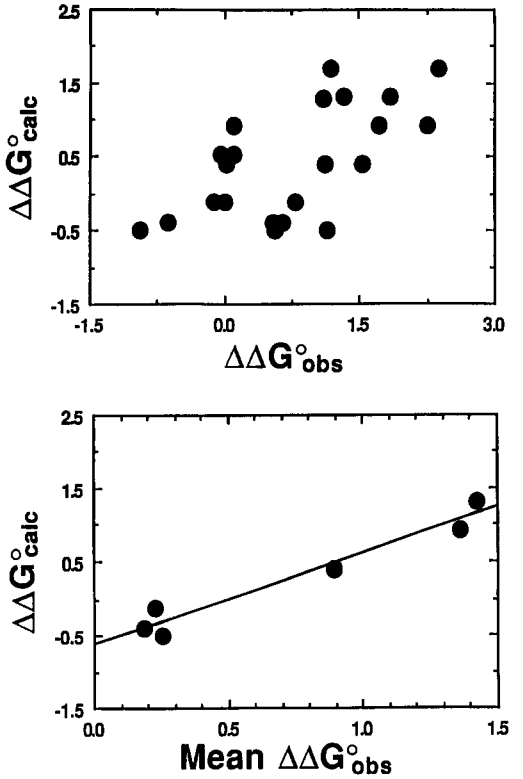

FIGURE 2: (A, top) Observed free energy change $\Delta\Delta G°_{obs}$ plotted against $\Delta\Delta G°_{calc}$, the calculated free energy change. $\Delta\Delta G°_{calc}$ is the sum of two components. One is the free energy change due to the surface area differences between side chains upon mutation (Lesser & Rose, 1990), multiplied by 22 cal/Å^2 (Chothia, 1976) to convert into an energy term. The second is the helical propensity difference between the wild-type amino acid and the mutant (Gans et al., 1991). (B, bottom) Mean free energy change of each type of mutation plotted against $\Delta\Delta G°_{calc}$. The line shown is the best-fit linear regression line ($\Delta\Delta G°_{calc} = -0.62 + 1.2\Delta\Delta G°_{obs}$; $R^2 = 0.95$). Isoleucine to valine and leucine to alanine mutations are omitted from this plot, because I111V and L135A are both too unstable to be purified and the average of the remaining two is not statistically equivalent to the others.

by the normal procedure (Springer & Sligar, 1987), even though the mutation corresponds to an increase in helix propensity. In the other, L49I decreases the helix propensity, and gives a protein with a T_m 2.5 °C lower than wild type. There is no trend in substitutions at nonhelical internal sites. These two sites are the only two possible hydrophobic sites we found in the interior of the myoglobin located on the nonhelical position and without direct contact with the heme group.

At each helical site, the least stable proteins result from substituting Ala for a bulkier side chain, regardless of helix propensity, if we assume L135A is less stable than L135V and does not survive isolation. Changes from bulky side chains to Ala can disrupt packing or even create cavities in the core (Eriksson et al., 1992) and hence should be destabilizing. Since the substitutions described here are all in the hydrophobic core of the protein, which is at least partly disrupted on unfolding, the difference in hydrophobicity accompanying each amino acid change has to be taken into account (Pinker et al., 1993). By converting the change in side-chain surface area upon mutation, estimated as the difference in total side-chain surface area between the side chains (Lesser & Rose, 1990), to a free energy scale using Chothia's factor of 22 cal/Å^2 (1976), we obtain the result shown in Figure 2A. The free energy difference between the wild-type and mutant proteins, $\Delta\Delta G°_{obs}$, shows an overall positive correlation with $\Delta\Delta G°_{calc}$, the sum of midhelix propensity differences evaluated from

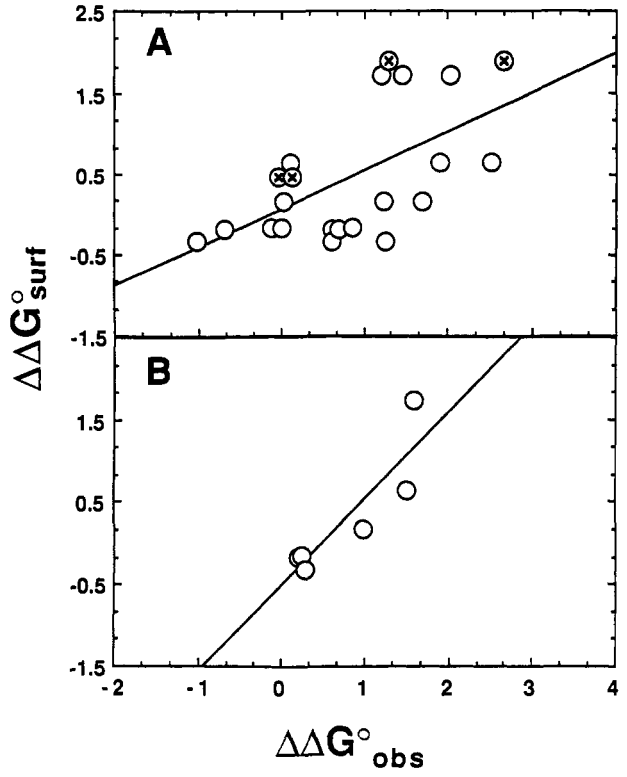

FIGURE 3: $\Delta\Delta G°_{obs}$ plotted against $\Delta\Delta G°_{surf}$ without taking into account the helix propensity scale. Panel A: unaveraged data; the best fitting line has $R^2 = 0.35$. Panel B is the site-averaged data. The points with different symbols are not included in this graph (see Figure 2B legend). The equation of the best fitting straight line for these data is $\Delta\Delta G°_{calc} = -0.53 + 1.1(\text{mean } \Delta\Delta G°_{obs})$, with $R^2 = 0.77$.

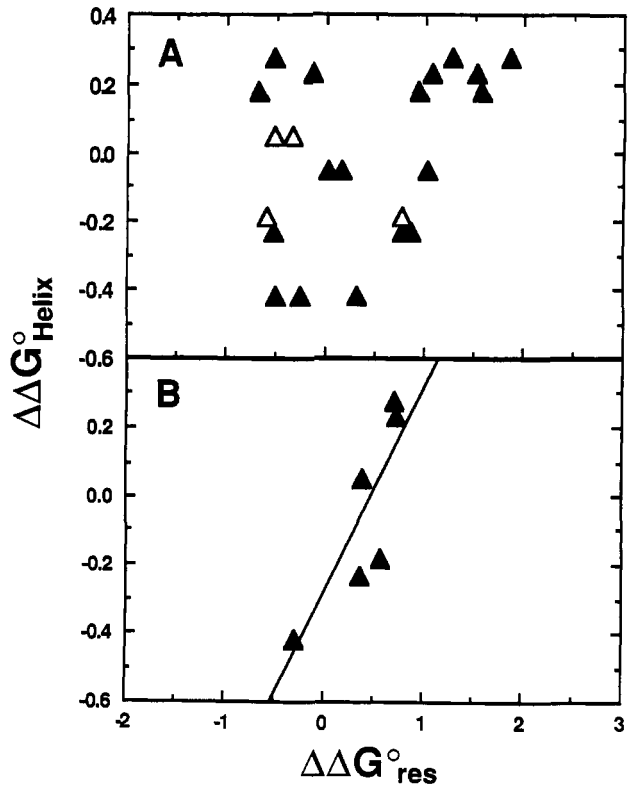

FIGURE 4: $\Delta\Delta G°_{res}$, the residual free energy change after the free energy contributionis subtracted from the side-chain surface area ($\Delta\Delta G°_{obs} - \Delta\Delta G°_{surf}$), plotted against $\Delta\Delta G°_{helix}$, the helical propensity scale (Gans et al., 1991). Again, panel A is the raw data without site averaging (no trend can be seen in this case), while in panel B, after site averaging, the correlation ($\Delta\Delta G°_{helix} = -0.29 + 0.60\Delta\Delta G°_{res}$, $R^2 = 0.68$) indicates that, despite the major effect coming from hydrophobicity, the secondary structure effect accounts for a considerable fraction of the remaining stability change.

model peptides (O'Neil & Degrado, 1990; Gans et al., 1991) and the change in hydrophobicity calculated from residue area as above. The data show considerable scatter, which is not unexpected, since each internal site represents a distinct microenvironment of the protein (Green et al., 1992). The scatter can be reduced substantially (Figure 2B) if we average the free energy change for each *type* of substitution over the different sites. This procedure assumes that effects at a given site are local and improves the correlation notably. It is interesting that the correlation in Figure 2B holds even if the side-chain substitutions that involve an appreciable change in surface area are discarded. For example, removing the Ala and Val substitutions from this plot, we find that the remaining hydrophobic differences among Leu, Ile, and Met still account for most of the variation in free energy of the mutant proteins.

How important is helix propensity relative to side-chain hydrophobicity? If we examine the raw data, a modest correlation still can be seen (Figure 3A) between side-chain hydrophobicity differences and the stability of the mutant proteins (correlation coefficient $R^2 = 0.35$). Averaging local effects at different sites produces a stronger correlation (see Figure 3B):

$$\Delta\Delta G°_{surf} = -0.5 + 1.1\Delta\Delta G_{obs}$$

with a correlation coefficient $R^2 = 0.77$. However, when we compare Figure 2B ($R^2 = 0.95$) with Figure 3B, we see that including differences in helix propensity among the side chains significantly improves the correlation. To pursue this further, in Figure 4, $\Delta\Delta G°_{res}$, which is the residual experimental free energy that is *not* accounted for by the surface area contribution, $\Delta\Delta G°_{res} = \Delta\Delta G°_{obs} - \Delta\Delta G°_{surf}$, is plotted against the difference in helix propensities between the side chains substituted. In panel A, the data are not averaged, and no trend can be discerned. As seen in panel B, a positive correlation results when sites are averaged, with $R^2 = 0.68$. This result is in marked contrast to what we have observed in the case of external Ala substitutions in myoglobin (Pinker et al., 1993). In the case of the fully internal sites, the magnitude of the helix propensity differences is sufficient to account for the residual free energy against a background of all other contributions other than the hydrophobicity. Use of hydrophobic free energies derived from transfer experiments in which model compounds are partitioned between water and nonaqueous solvents weakens the correlations we find [see Lesser & Rose (1990)].

Figure 5 illustrates the dependence on urea concentration of the T_m for the wild-type and several mutant myoglobins. Inspection of the values of the slopes in Table I indicates that the largest value corresponds to X → A differences. Thus size difference between the side chains seems to influence the urea dependence of T_m, as it does the T_m values in the absence of urea. Since all the sites we are discussing are chosen to be fully buried in the native protein, these differences can reasonably be attributed to changes in accessibility or binding of urea by the unfolded protein, as Green et al. (1992) have argued, if we assume a two-state transition model. The actual situation might be more complex, however, if populations of intermediate states are involved.

CONCLUSION

The above results lead us to conclude that the role of the α helix propensity of a side chain in the interior of myoglobin

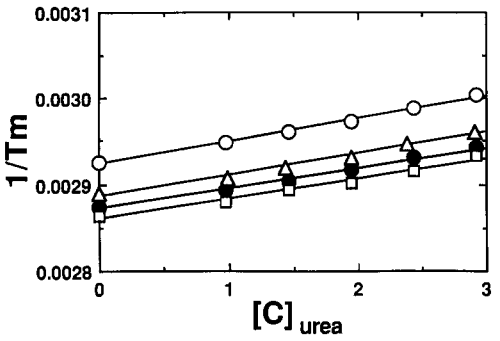

FIGURE 5: Dependence of T_m^{-1} on urea concentration. Symbols: wild type (●), L29V (○), I142L (□), and I111L (△). Urea concentrations were determined by refractive index measurements (Pace, 1986). All melting curves were done in the same buffer system as described under Materials and Methods with the addition of urea. The slopes, m', are listed in Table I.

is not negligible, so that the secondary-structure stabilization contributes directly to that of the native state. It seems important to qualify this conclusion with care. In substitutions among chemically similar alkyl side chains at internal helical sites in myoglobin, differences in hydrophobicity between the side chains involved still account for about twice the variation that can be attributed to differences in helix propensity *per se* (see Figure 4). Thus while helix propensities can be seen to play a role, they do not represent the dominant effect in stabilizing the native folded state of the protein. The relatively small difference in area (or hydrophobicity) between Leu and Ile, for example, still contributes more to the free energy of folding on average than the difference in their s values. The contribution of α-helix to the free energy of folding the native protein can be supposed to play a determinative role in acquiring the native fold in any case. Secondary structure seems to be present in the earliest detectable folding intermediates in a protein such as cytochrome c (Englander & Mayne, 1992) and hence might prove crucial in the pathway of folding in terms of framework models (Kim & Baldwin, 1982). Pace's group (Pace, 1992) has recently revised upward the estimated equilibrium thermodynamic contribution of H-bonding to the native structure of a protein, concluding that the total effect of H-bonds is not insignificant relative to the hydrophobic component. As pointed out by Creamer & Rose (1992), relative helix propensities based on Gly, for example, do not include the H-bonding contribution to helix formation. Helix propensity differences alone thus might underestimate the stabilizing contribution to folding from helical structure.

In attempting to understand the extent of scatter seen in Figure 2A, several quite different contributions of local environment might play a role. First, according to Shortle et al. (1990), substitutions at internal sites in staphylococcal nuclease vary in their effect on stability depending on the local density of Cα carbons in the vicinity of the site. While side chains in the interior of proteins are mobile, as implied in models which treat the interior as an apolar fluid (Baldwin, 1986; Murphy et al., 1990), mutational studies implicate localized "internal packing energy" differences (Kellis et al., 1988; Lim & Sauer, 1991; Sandberg & Terwilliger, 1991). Thus local rigid body contributions or density differences might account for the large variation in effect of single mutations from site to site seen in Figures 2A, 3A, and 4A. We avoided sites close to the heme group in order to remove heme contacts as a source of such rigid body effects. However, the heme may have a general impact on local dynamics or strain of side chains in the core of myoglobin. A second and more

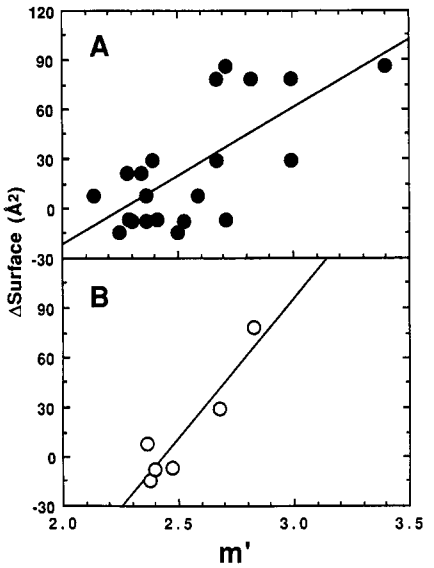

FIGURE 6: Calculated changes in surface area upon mutation (Lesser & Rose, 1990), Δsurface (Å²), plotted against m', the slopes of the graphs in Figure 5. Panel A: The straight line fit gives a correlation coefficient $R^2 = 0.5$. Note the scatter in the slopes for the same type of mutation at different sites in the protein, which suggests the differential binding of urea to the unfolded states of mutant proteins. Panel B: Effect of site to site averaging on the data in panel A. A similar reduction in scatter is seen as in Figure 2; $R^2 = 0.86$.

problematical factor is the residual structure that is seen in the unfolded states of proteins under a variety of denatured conditions (Shortle & Meeker, 1986; Neri et al., 1992). This can contribute to site–site differences in unpredictable ways.

It is worth noting that a plot of the difference in surface area between the side chains substituted against the value of the slope, m', shows a scatter similar to that in Figure 2A (Figure 6). In the case of these slopes, one might expect the extent of exposure to urea of side chains in the unfolded state to be more uniform, reducing the dispersion. This is not observed, and the result of averaging the substitutions over different sites produces a comparable correlation to what we find for $\Delta\Delta G°$ values in Figure 2B. Thus the exposure to urea in the unfolded myoglobin reveals a site to site variation that is consistent with the presence of residual structure in that state or some local effect that prevents uniform solvation of side chains by urea. The parallel between Figure 2 and Figure 6 might suggest that local effects in the native state do not account for the scatter in Figure 2A, if the m' values indeed reflect differences in unfolded states of the protein. This depends however on the role of intermediate state(s) in the unfolding process.

In principle, structural rearrangements in the native state might accompany some, if not all, of the substitutions carried out, perturbing the free energy of the native state. Even if all these could be determined in detail, it might prove difficult to interpret their thermodynamic contributions quantitatively. Depending on how one interprets m' differences, it may be reasonable to attribute important effects to intermediates or unfolded states of the protein. In either case, it seems that the simplifying procedure of averaging several sites together in order to try to isolate the effect of a particular substitution improves the scatter in the resulting data. This is decisively so in the case of Figure 4.

ACKNOWLEDGMENT

George Rose and Oleg Ptitsyn provided essential assistance and discussions during this work. Drs. B. Springer and S.

Sligar kindly provided us with their cloned myoglobin gene and guidance in expressing and isolating myoglobins. Dr. Luis Marky gave expert advice in DSC measurements.

REFERENCES

Acampora, G., & Hermans, J., Jr. (1967) *J. Am. Chem. Soc. 89*, 1543–1552.

Baldwin, R. L. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 8069–8072.

Barrick, D., & Baldwin, R. L. (1993) *Protein Sci. 2*, 869–876.

Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyerr, E. F., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T., &Tasumi, M. (1977) *J. Mol. Biol. 112*, 535–542.

Blaber, M., Zhang, X.-j., & Matthews, B. W. (1993) *Science 260*, 1637–1640.

Chothia, C. (1976) *J. Mol. Biol. 105*, 1–14.

Creamer, T. P., & Rose, G. D. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 5937–5941.

Dill, K. A. (1990) *Biochemistry 29*, 7133–7155.

Englander, S. W., & Mayne, L. (1992) *Annu. Rev. Biophys. Biomol. Struct. 21*, 243–265.

Eriksson, A. E., Baase, W. A., Zhang, X.-J., Heinz, D. W., Blaber, M., Baldwin, E. P., & Matthews, B. W. (1992) *Science 255*, 178–183.

Fasman, G. D. (1989) in *Prediction of Protein Structure and the Principles of Protein Conformation* (Fasman, G. D., Ed.) pp 193–316, Plenum Press, New York.

Gans, P. J., Lyu, P. C., Manning, M. C., Woody, R. W., & Kallenbach, N.R. (1991) *Biopolymers 31*, 1605–1614.

Green, S. M., Meeker, A. K., & Shortle, D. (1992) *Biochemistry 31*, 5717–5728.

Hecht, M. H., Sturtevant, J. M., & Sauer, R. T. (1986) *Proteins: Struct., Funct., Genet. 1*, 43–46.

Horovitz, A., Matthews, J. M., & Fersht, A. R. (1992) *J. Mol. Biol. 227*, 560–568.

Jeng, M.-F., & Englander, S. W. (1991) *J. Mol. Biol. 221*, 1045–1061.

Kellis, J. T., Nyberg, K., Sali, D., & Fersht, A. F. (1988) *Nature 333*, 784–786

Kim, P. S., & Baldwin, R. L. (1990) *Annu. Rev. Biochem. 59*, 631–660.

Kunkel, T. A. (1985) *Proc. Natl. Acad. Sci. USA 82*, 488–492.

Kuwajima, K. (1989) *Proteins: Struct., Funct., Genet. 6*, 87–103.

Lesser, G. L., & Rose, G. D. (1990) *Proteins: Struct., Funct., and Genet. 6*, 6–13.

Lim, W. A., & Sauer, R. T. (1991) *J. Mol. Biol. 219*, 359–376.

Lyu, P. C., Liff, M. I., Marky, L. A., & Kallenbach, N. R. (1990) *Science 250*, 669–6738.

Lyu, P. C., Zhou, H. X., Jelveh, N., Wemmer, D. E., & Kallenbach, N. R. (1992) *J. Am. Chem. Soc. 114*, 6560–6562.

Mayr, L. M., & Schmid, F. X. (1993) *Biochemistry 32*, 7994–7998.

Murphy, K. P., Privalov, P. L., & Gill, S. J. (1990) *Science 247*, 559–561.

Neri, D., Billeter, M., Wider, G., & Wuthrich, K. (1992) *Science 257*, 1559–1563.

Nozaki, Y., & Tanford, C. (1971) *J. Biol. Chem. 246*, 2211–2217.

O'Neil, K. T., & Degrado, W. F. (1990) *Science 250*, 646–651.

Pace, C. N. (1986) *Methods Enzymol. 131*, 266–280.

Padmanabhan, S., Marqusee, S., Ridgeway, T., Laue, T. M., & Baldwin, R.L. (1990) *Nature 344*, 268–270.

Phillips, G. N., Arduini, R. M., Springer, B. A., & Sligar, S. G. (1990) *Proteins: Struct., Funct., Genet. 7*, 358–365.

Pinker, R. J., Lin, L., Rose, G. D., & Kallenbach, N. R. (1993) *Protein Sci. 2*, 1099–1106.

Presta, L. G., & Rose, G. D. (1988) *Science 240*, 1632–1641.

Privalov, P. L., & Khechinashvili, N. N. (1974) *J. Mol. Biol. 86*, 665–684.

Privalov, P. L., Griko, Y. V., Venyaminov, S. Y., & Kutyshenko, V. P. (1986) *J. Mol. Biol. 190*, 487–498.

Ptitsyn, O. B. (1992) The molten globule state, in *Protein Folding* (Creighton, T. E., Ed.) pp 243–299, W. H. Freeman, New York.

Puett, D. (1973) *J. Biol. Chem. 248*, 4623–4634.

Richards, F. M. (1977) *Annu. Rev. Biophys. Bioeng. 6*, 151–175.

Richardson, J. S., & Richardson, D. C. (1988) *Science 240*, 1648–1652.

Sandberg, W. S., & Terwilliger, T. C. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 1706–1710.

Sanger, F., Nicklen, S., & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463–5467.

Santoro, M. M., & Bolen, D. W. (1992) *Biochemistry 31*, 4901–4907.

Schellman, J. A. (1978) *Biopolymers 17*, 1305–1322.

Serrano, L., & Fersht, A. R. (1989) *Nature 342*, 296–299.

Serrano, L., Kellis, J. T., Jr., Cann, P., Matouschek, A., & Fersht, A. R. (1992) *J. Mol. Biol. 224*, 783–804.

Shirley, B. A., Stanssens, P., Hahn, U., & Pace, N. C. (1992) *Biochemistry 31*, 725–732.

Shortle, D., & Meeker, A. K. (1986) *Proteins: Struct., Funct., Genet. 1*, 81–89.

Shortle, D., Stites, W. E., & Meeker, A. K. (1990) *Biochemistry 29*, 8033–8041.

Springer, B. A., & Sligar, S. G. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 8961–8965.

Tanford, C. (1970) *Adv. Protein Chem. 24*, 2–95.

Zimm, B. H., & Bragg, W. K. (1959) *J. Chem. Phys. 31*, 526–535.