Structural Consequences of Heme Removal: Molecular Dynamics Simulations of Rat and Bovine Apocytochrome b_5[†]

Elizabeth M. Storch and Valerie Daggett*

Department of Medicinal Chemistry, Box 357610, University of Washington, Seattle, Washington 98195-7610

Received March 11, 1996; Revised Manuscript Received May 17, 1996[⊗]

ABSTRACT: Molecular dynamics simulations of rat and bovine apocytochrome b_5 were performed to investigate the structural and dynamical consequences of heme removal. A crystal structure is available for the bovine holoprotein, while experimental studies of apocytochrome b_5 have focused on the rat protein. The rat and bovine proteins are 93% homologous by sequence, and the sequence differences (six residues) appear to have no effect on the structure of the native holoprotein, as seen by the correlation of a bovine simulation with rat holocytochrome b_5 experimental data (Storch & Daggett, 1995). There was a marked effect, however, on the structure and dynamics of the apo form. The bovine apocytochrome b_5 simulation displayed subtle inconsistencies when compared to the experimental results on the rat apoprotein. Therefore, the rat protein was constructed from the bovine crystal structure coordinates. The MD simulation of the rat apoprotein displayed greater deviations from the crystal structure, yet it was in much closer agreement to the experimental data for the apoprotein. Additionally, the six variant residues fall in the regions where the bovine protein deviated from experiment. The two hydrophobic cores of the rat protein behaved very differently. Core 2 was well maintained, retained native-like structure, and is in good agreement with NMR data (Moore & Lecomte, 1990). Conversely, core 1, which is normally constrained by the prosthetic heme group, exhibited conformational heterogeneity, increased mobility, and some loss of secondary structure. Thus, the model of rat apocytochrome b_5 complements past studies by providing structural information about core 1 that has proved difficult to obtain by experiment. The bovine simulation serves as a prediction, since little to no experimental data exist for this form of the apoprotein.

An understanding of protein folding hinges upon elucidation of the progression of events occurring along the folding/unfolding pathway. Many experimental techniques have been applied to this problem; however, the cooperative nature of folding generally results in only minute amounts of partially folded intermediates at equilibrium. Exceptions to this behavior have been found in so-called molten globules, but these are generally found under denaturing conditions. Investigation of the nature of partially unfolded protein conformations under physiological or native conditions has been possible for a few hemoproteins destabilized through removal of their heme groups: apomyoglobin (Breslow et al., 1965; Harrison & Blout, 1965; Lecomte & Cocco, 1990; Hughson et al., 1990, 1991; Brooks, 1992; Fink, 1995), apocytochrome b_{562} (Feng & Sligar, 1991; Feng et al., 1991, 1994; Laidig & Daggett, 1996a), and apocytochrome b_5 (Huntley & Strittmatter, 1972; Tajima et al., 1976; Moore & Lecomte, 1990, 1993; Moore et al., 1991; Falzone et al., 1996).

In addition to being a possible protein folding intermediate under native, *in vitro* conditions, apocytochrome b_5 appears to be populated *in vivo*. Apocytochrome b_5 is synthesized on the endoplasmic reticulum (Gonzalez & Casper, 1980; Krieter & Shires, 1980), and an excess of the protein exists in microsomes (Hara & Minakami, 1970; Hara et al., 1970; Negishi & Omura, 1970; Shawver et al., 1984). Later, insertion of the iron and porphyrin moiety into the protein occurs in the mitochondria (Jones & Jones, 1969). Therefore, it may be necessary for apocytochrome b_5 to maintain a quasistable conformation following synthesis and intracellular transport prior to heme insertion and subsequent participation in electron transfer reactions with a variety of partners (Lederer, 1994).

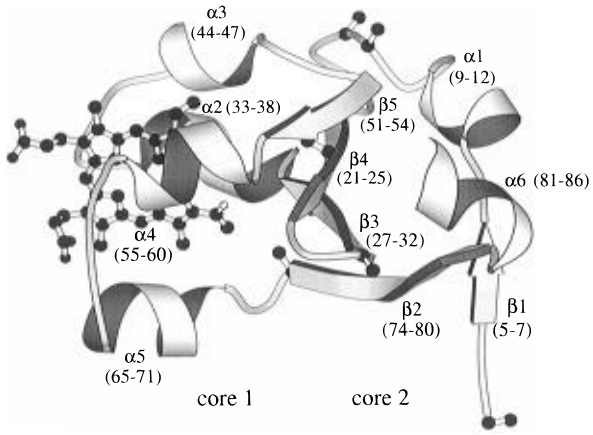

FIGURE 1: Ribbon diagram of the crystal structure of bovine cytochrome b_5 including the prosthetic heme group (Mathews et al., 1972). The heme group was not included in the apoprotein simulations. The two cores and the residue numbers comprising secondary structure elements are indicated. The C_β atoms for the six variant residues are displayed. This figure was made using MOLSCRIPT (Kraulis, 1991).

Cytochrome b_5 consists of six α-helices, five strands making up a β-sheet, and various β-turns. There are two hydrophobic cores in the molecule, one of which binds the heme group (core 1, Figure 1), while the other is important in maintaining the structural integrity of the protein (core 2,

[†] Supported by the National Institutes of Health (Grant GM 50789).
* To whom correspondence should be addressed.
[⊗] Abstract published in *Advance ACS Abstracts,* August 15, 1996.

S0006-2960(96)00598-3 CCC: $12.00

Figure 1). When the heme group is extracted under native conditions, the protein expands and is less stable than the holoprotein (Huntley & Strittmatter, 1972; Tajima et al., 1976; Pfeil, 1993). Circular dichroism (CD) studies suggest that there is an increase in disorder and less repetitive secondary structure in the apo form (Huntley & Strittmatter, 1972).

Using one- and two-dimensional NMR spectroscopy, Moore and Lecomte (1990) found that there is a stable hydrophobic cluster in apocytochrome b_5, in which Trp 22 figures prominently; this cluster is a key element of the second core. In addition, four of the five β-strands and at least two of the six α-helices are present in the apo form (Moore & Lecomte, 1993). The rest of the structure appears to fluctuate between different conformations (Moore et al., 1991). In fact, even structured portions of the protein are fluctuating more than in the holo form as measured by hydrogen exchange and NMR (Moore & Lecomte, 1990, 1993; Lecomte & Moore, 1991). More recent NMR studies utilizing ^{15}N-labeled protein support these earlier findings and provide further structural information (Falzone et al., 1996). This intermediate displays many of the hallmarks of the molten globule state except that there are well-populated tertiary interactions. Moore and Lecomte (1993) have proposed a model in light of these results in which one of the cores folds rapidly (core 2), while core 1 requires the heme to fold properly.

Although structural data for apocytochrome b_5 exist, precise characterization of all portions of the molecule has been difficult because of its dynamic nature. Computer simulations provide an avenue for elucidating the atomic details of structural transitions and the detailed dynamic behavior of proteins. However, the information obtained from a simulation must be compared to experiment to ensure relevancy.

Cytochrome b_5 is a good choice for molecular dynamics (MD) simulation studies. It is small, containing 98 amino acid residues. The partially folded state occurs under native conditions, so we do not have to include denaturant or use high temperature to destabilize the native state. Most importantly, there is a wealth of structural information for both holo- and apocytochrome b_5, and the necessary control simulation of holocytochrome b_5 is well behaved (Storch & Daggett, 1995). Even though these structural data exist, there is still much to be learned from the simulations. For example, what do the unstructured portions of the molecule look like? How much conformational sampling does the secondary structure experience, since it is present but fluctuating more than in the holo form? Are nonnative interactions important in stabilizing the folding intermediate? We can address these questions with simulations and provide a detailed dynamic, molecular model of apocytochrome b_5 that goes beyond, but is complementary to, what can be garnered using experimental methods.

To perform MD simulation studies of apocytochrome b_5, one needs a starting structure. The only available crystal structure of cytochrome b_5 is of the bovine form, while the structural experimental work described above was performed on the rat protein. The sequence homology of the water-soluble fragment (residues 3−87) is 93% in the rat and bovine forms (Ozols et al., 1976); thus, structural differences between the two should be minor. This does indeed appear to be true for holocytochrome b_5, as the six variant residues seem to have little to no effect based on the fact that our MD simulation of bovine holoprotein is in good agreement with both bovine and rat experimental results (Storch & Daggett, 1995). However, the MD simulation of bovine apocytochrome b_5 presented here exhibited slight structural deviations from the experimental data for the rat apoprotein. The discrepancy was localized to three regions of the protein, and four of the six species variant residues are located within these regions. This led us to further investigate the structural and dynamical consequences of "mutating" these six residues. When the starting structure was modified to the rat sequence, the agreement between experiment and the MD results improved significantly, allowing us to propose a detailed model for the rat apocytochrome b_5 folding intermediate. In addition, these simulations demonstrate how mutations can lead to structural and dynamical changes in partially unfolded conformations while having no apparent effect on the folded state.

METHODS

The starting conformation of apocytochrome b_5 was the crystal structure of bovine cytochrome b_5 with the prosthetic heme group removed (Brookhaven Protein Data Bank, 3b5c; Mathews et al., 1972). The crystal structure contains residues 3−87; residues 1 and 2 and 88−93 were not visible in the electron density maps and were therefore omitted in the simulation. For the rat protein simulation, six modifications were made to the bovine crystal structure coordinates by placing the new side chains in approximately the same orientation. The following mutations were made: Ala 3 → Asp, Asn 16 → Lys, Asn 17 → Asp, Leu 23 → Val, Tyr 27 → His, and Phe 74 → Tyr (bovine → rat) (Ozols, 1989). The positions of these variant residues are indicated in Figure 1.

The potential energy function (Levitt, 1983, 1989; Levitt et al., 1995; Laidig & Daggett, 1996b) and associated protocols (Daggett & Levitt, 1992) are described in detail elsewhere. Energy minimization and the MD simulations were performed using the program ENCAD (Levitt, 1990). The simulations were performed at pH 2 for 200 ps and continued at pH 6.9 for 1400 ps. This procedure was followed because heme removal occurs experimentally by lowering the pH to 2 (Teale, 1959), which protonates the coordinating histidine residues (His 39 and 63) and subsequently releases the heme. Experiments are then resumed at a physiologically relevant pH of 6.9. The water density was set to the experimental value (0.997 g/mL) at 298 K (Kell, 1967) by adjusting the volume of the box prior to beginning the simulation. The MD simulation of rat protein contained 2901 water molecules at pH 2 and 3084 water molecules at pH 6.9. The bovine protein simulation contained 2771 water molecules at pH 2 and 3096 water molecules at pH 6.9.

Beginning with the low pH conditions (Glu, Asp, and His protonated), the systems were subjected to a variety of preparatory steps as described elsewhere (Storch & Daggett, 1995). MD was then performed for 1×10^5 steps utilizing a 2 fs time step, resulting in the first 200 ps of the trajectory at 298 K. The starting structure for this portion of the simulation was used as the starting point for the simulation at pH 6.9. All Glu and Asp residues and His 26, His 27 (in the case of the rat simulation; the bovine form contains Tyr

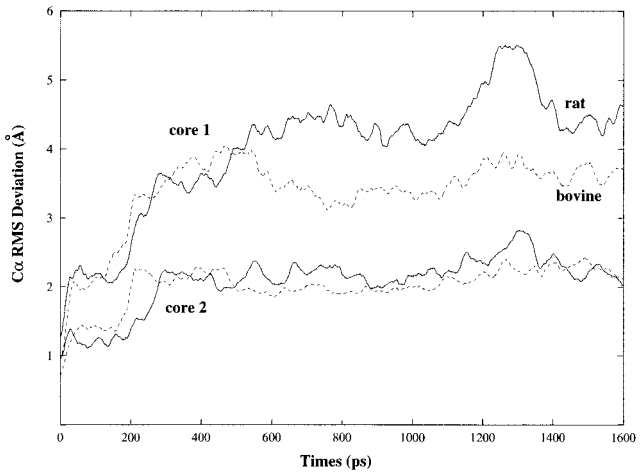

FIGURE 2: C_α RMS deviation from the crystal structure as a function of time. The rat simulation is depicted by the solid line and the bovine simulation by the dashed line. The two upper lines represent the deviation of core 1 residues (33−73), and the two lower lines represent core 2 residues (3−32 and 74−83). For simplicity, we define the cores as being inclusive in sequence, instead of marking each individual hydrophobic group pointing into each core. The last three residues (84−87) were not monitored. The proteins were modeled at pH 2 for the initial 200 ps and at pH 6.9 thereafter.

at this position), and His 80 were deprotonated. His 15, His 39, and His 63 and all Lys and Arg residues were protonated and positively charged. The experimentally determined pK_a values for the His residues were used to determine the ionization states (Altman et al., 1989; Moore et al., 1991). Again, the proteins were solvated and prepared for MD. The MD trajectory was then continued for 1400 ps at 298 K, and structures were saved every 0.2 ps for analysis (7000 structures).

RESULTS

Rat Apocytochrome b_5: Tertiary Structure. The starting structure for the simulation of rat apocytochrome b_5 was the modified crystal structure of bovine cytochrome b_5 (Mathews et al., 1972); the six variant residues which were "mutated" are shown in Figure 1. Cytochrome b_5 contains two distinct regions: core 1 (residues 33−73) and core 2 (residues 3−32 and 74−87) (Figure 1). The C_α root-mean-square (RMS) deviation from the crystal structure of each core was monitored separately (Figure 2), and a drastic difference between the cores is evident. Within the first 300 ps of the simulation, the deviation from the crystal conformation of core 1 climbed to 3.7 Å. These residues fluctuated greatly with time and reached values greater than 5.5 Å. Conversely, core 2 deviated much less from the crystal structure (∼2 Å) and experienced only minor fluctuations.

In order to localize the motion to more distinct areas of the protein, the C_α RMS deviation was plotted as a function of residue number and time (Figure 3). The deviation was greatest in core 1, primarily for residues 33−76 and the C-terminus. The highest deviations, shown in black, occurred predominantly between residues 39−52 and 65−72. These residues incorporate the secondary structure elements α3 and the preceding loop, α5, and the loop prior to β2 (Figure 1). In the latter part of the simulation, residues 52−60 (β5, α4, and preceding loop) also showed large deviations. α6 accounted for the large degree of deviation in core 2, as it had few tethers to the rest of the protein. Throughout the

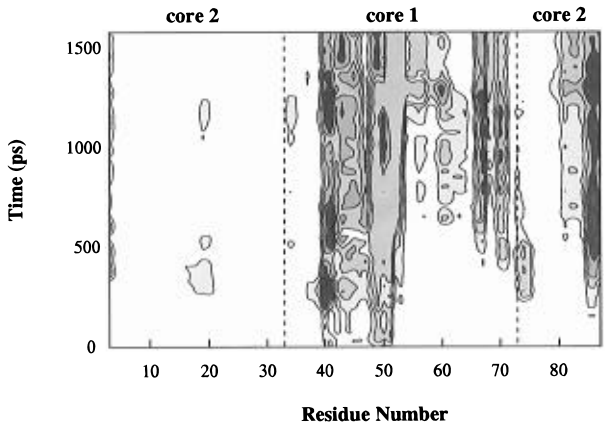

FIGURE 3: Contour plots of the C_α RMS deviation from the crystal structure per residue number as a function of time. Deviations are coded as follows: black, RMS > 5 Å; gray, 4 Å < RMS < 5 Å; light gray, 3 Å < RMS < 4 Å; white, RMS < 3 Å. The core boundaries are denoted by dashed lines.

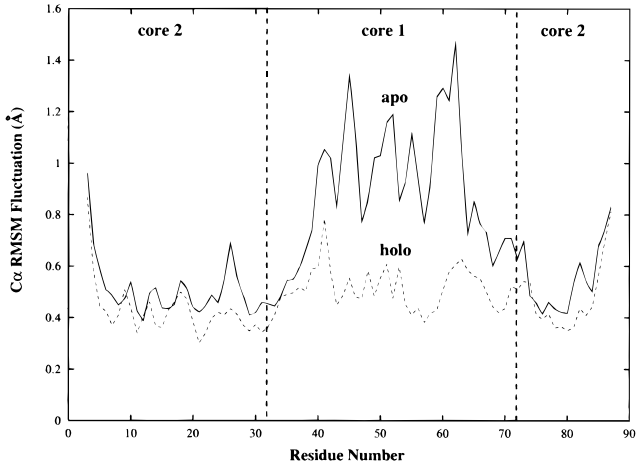

FIGURE 4: Plots of the C_α RMS fluctuation per residue number. The values were averaged over the last 200 ps of the simulations. The apo rat simulation is given by the solid line, and the holo bovine run is depicted by the dashed line.

simulation, more dramatic deviation was observed in core 1 (helices α2−α5) than core 2.

Even though core 1 deviated considerably from the crystal structure, the deviation does not necessarily reflect heightened mobility. For example, a conformation may be adopted that is different from the crystal structure, but it could be quite immobile once it has stabilized in the new orientation. A more direct way to determine the mobility of a state is by calculating the C_α RMS fluctuation about the mean structure as in Figure 4. There was a dramatic increase in the mobility of core 1 in the apo form compared to its heme-bound state, although at certain time points the mobility of residues 40−50 was high in the control simulation as well. Thus, not only did the placement of core 1 in apo b_5 change compared to the crystal structure but after doing so it remained very mobile.

A series of snapshots at various time points illustrate the areas of high mobility (Figure 5). The backbone displacement was highest in core 1 (residues 33−73). The helices in this core (α2−α5) moved out and away from the center of the heme-binding pocket and became conformationally heterogeneous. β5, which borders the heme pocket, became displaced from the β-sheet and disrupted the connectivity between cores 1 and 2 (this strand moved ∼12 Å from its

starting position). This is particularly evident when considering the solvent accessibility of β4 residues (residues 21−25), all of which are sequestered in the holoprotein. The solvent accessibility of these residues increased 7−54 Å^2, with an average increase of 26 Å^2. In contrast to core 1, the backbone placement was strikingly similar to the crystal structure for most of core 2 with the possible exception of the N- and C-termini which fluctuated a great deal, and at times making even closer contacts with the protein core than in the crystal structure.

Additional tertiary contacts were monitored throughout the simulation and compared to the experimental NMR data. Potential nuclear Overhauser effect connectivities (NOEs) were evaluated by calculating an average-weighted distance ($\langle r^{-6}\rangle^{-1/6}$) between interacting hydrogens during the simulation. All of the hydrogen-hydrogen distances corresponding to the experimentally observed NOEs (Moore & Lecomte, 1993) in hydrophobic core 2 [see Figure 4 of Moore and Lecomte (1990)] remained below 5 Å, which is considered a reasonable cutoff for an NOE (Wüthrich, 1986). These interactions were experimentally observed for both rat apo- and holocytochrome b_5 (Moore & Lecomte, 1990). Falzone et al. (1996) have found that the amide hydrogens of core 1 exchange with solvent, while many of the residues in core 2 are protected. In the simulation, the protected residues in core 2 participated in hydrogen bonds and would therefore be unlikely to exchange with solvent.

Pfeil (1993) has reported second derivative heat capacity spectra of apo- and holocytochrome b_5, indicating that the Trp and Tyr residues are in similar environments in the two forms of the protein, while the environment of the Phe residues changes upon heme removal. However, in comparison to our MD simulation, it must be noted that Pfeil's work focuses on the rabbit protein. The Phe, Tyr, and Trp residues are identical in the rat and rabbit sequences with the exception of Phe 74, which is a Tyr in rat and a Phe in the bovine form (Ozols, 1989). In the MD simulations of apo- and holocytochrome b_5, the solvent accessibility of the Tyr and Trp residues was low (the apoprotein showed an increase of 4−13 Å^2, where the residues are not exposed in holo), with the exception of Tyr 74 (with a 35 Å^2 increase from holo). The solvent exposure of the Phe residues increased in the apoprotein MD simulation (20−44 Å^2), as suggested by experiment. Phe 35 and 58 and Tyr 74 interact with the heme and are sequestered from solvent in the holoprotein.

Secondary Structure. The boundaries of the secondary structure were determined via (ϕ,ψ) dihedral angles using a previously described method (Daggett & Levitt, 1992; Daggett et al., 1991) that requires that at least three residues in succession fit the dihedral angle criterion. The overall amounts of α-helical and β-sheet structure in apocytochrome b_5 are in good agreement with the CD results of Huntley and Strittmatter (1972) (Table 1); however, it must be noted that CD is not reliable for definitive quantitative determination of secondary structure, as evidenced by the disparity in the results from two different research groups (Table 1).

When compared with the crystal structure, the helical content for α2 increased (Table 1); this phenomenon was also observed in the control simulation of holocytochrome b_5 (Storch & Daggett, 1995). α3 showed low helix content throughout the simulation. After the first 500 ps, the helix content of α5 dropped considerably, and it was structurally labile for the remainder of the simulation. α6 decreased significantly in helical content compared to the crystal structure. In contrast, α1 and α4 retained high helix contents, even though α4 was extremely mobile (residues 55−60; Figures 4 and 5). α2 was the most stable of the helices in core 1, in terms of internal structure and its dynamic behavior (Figure 5).

Table 1: Overall Percentage of Secondary Structure for Segments of Apocytochrome b_5 in the Holo Crystal Structure and during Molecular Dynamics[a]

secondary structure[b]	crystal structure	MD, rat apo b_5	experiment, rat apo b_5	MD, bovine apo b_5
α1 (9−12)	100	99 (1)		97 (14)
α2 (33−38)	33	94 (10)		78 (32)
α3 (44−47)	100	75 (9)		94 (15)
α4 (55−60)	100	98 (1)		97 (10)
α5 (65−71)	100	62 (18)		97 (8)
α6 (81−86)	83	16 (16)		26 (38)
overall α[c]	33	28	25−30,[d] 21[e]	31
β1 (5−7)	100	92 (8)		25 (44)
β2 (74−80)	57	8 (6)		11 (19)
β3 (27−32)	67	66 (2)		65 (6)
β4 (21−25)	100	98 (2)		90 (19)
β5 (51−54)	0	0		0
overall β	19	14	10−15,[d] 44[e]	12

[a] α-Helical or β-structure is defined as having at least three consecutive residues with the appropriate (ϕ,ψ) dihedral angles (Daggett & Levitt, 1992). Secondary structure contents are given in percentage units, with the standard deviation in parentheses. The values are averages over the entire simulations (200−1600 ps). [b] The elements of secondary structure (residue numbers in parentheses) are shown in Figure 1. [c] The overall fraction of secondary structure is for the entire protein over time divided by the total number of residues. [d] These data were obtained from CD experiments (Huntley & Strittmatter, 1972). [e] These data were obtained from CD experiments (Pfeil, 1993).

Two of the five strands in the β-sheet (β2 and β5) were structurally heterogeneous in the simulation. β2 (residues 74−80) and β5 (residues 51−54) are defined as irregular in the Brookhaven PDB file (Mathews et al., 1972). This irregularity is reflected in the low β-content of these segments in the crystal structure (Table 1). β2 further decreased almost immediately and remained low for the duration of the simulation. β5, on the other hand, never adopted β-structure according to the dihedral angle definition used. The remainder of the β-strands (β1, β3, and β4) were well maintained during the simulation. Even though the β-content of β3 was low, it is comparable to that of the crystal structure.

The intramolecular main-chain hydrogen bonds were monitored as another probe of secondary structure content. A hydrogen bond was defined as having a distance of ≤2.6 Å between carbonyl oxygen and amide hydrogen atoms and a hydrogen-bonding angle (N−H···O═C) within 45° of linearity. The values for the $i \rightarrow i + 4$ hydrogen bonds diagnostic of helical structure within the helices of the crystal structure (with added hydrogen atoms) were all ≤2.6 Å, with the exception of two hydrogen bonds within α2 (between residues 33−37 and 34−38) and α6 (residues 81−85 and 82−86). During the simulation, α2 formed the two helical hydrogen bonds not observed in the crystal structure. α6 also formed one of the missing hydrogen bonds (residues 81−85) in the first part of the simulation (0−750 ps) but lost it after 750 ps, reaching distances of ∼3.7 Å. The $i \rightarrow i + 4$ distances of α2 (residues 34−38) exceeded ∼3 Å intermittently throughout the simulation (900−1200 ps) as did those of α3 (residues 44−48, 500−1300 ps). One of

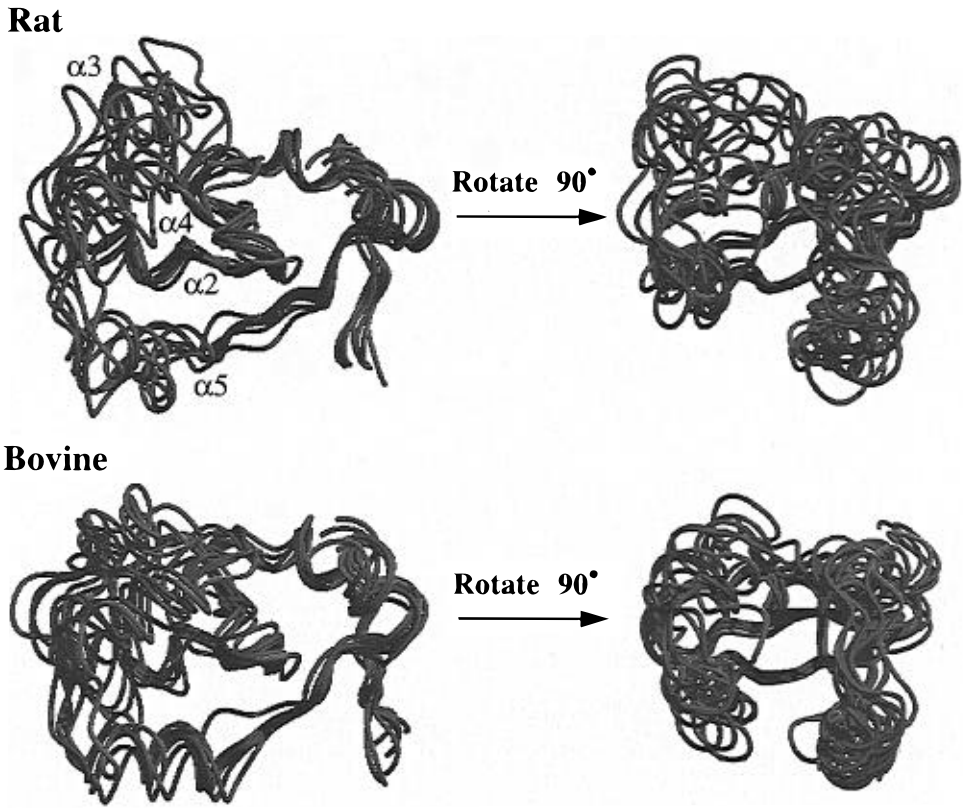

FIGURE 5: Main-chain traces from various time points throughout the rat and bovine simulations (350, 550, 750, 950, 1150, and 1350 ps). The crystal structure is shown in cyan. The figures on the left are in the same orientation as in Figure 1. The figures on the right provide a view into the heme binding pocket.

the internal $i \rightarrow i + 4$ distances of α5 (residues 66−70) began increasing at 550 ps, and the hydrogen bond was lost for the duration of the simulation. All other helical main-chain hydrogen bonds remained intact.

Interresidue hydrogen−hydrogen distances were also monitored and compared with the available NOEs. Moore and Lecomte (1993) found that α1 contains many of the NOEs diagnostic of helical structure. There is some evidence for helical NOEs in α2 and α3, but the intensities are lower than in the holo form. There is one probable helical NOE in α4. Residues in α5 could not be assigned. There is some evidence for helical NOEs in α6, but a lack of resolution prevented full assignment.

In the simulation, all of the $i \rightarrow i + 3$ connectivities expected for α-helices remained below 4 Å in α1−α4 throughout the simulation. The $i \rightarrow i + 3$ NOEs for α5 and α6, however, fluctuated greatly. In the C-terminal region of α5 (residues 66−69, 67−70 and 68−71), the distances were between 5 and 6 Å. The corresponding distances in the crystal structure were ≤3.9 Å. α6 has three $i \rightarrow i + 3$ NOEs within the defined region, residues 81−86. For the first 500 ps, all of these remained mostly below 4 Å. For the remainder of the simulation, the NOEs between residues 82−85 and 83−86 were ∼4.8 Å. Overall, α5 experienced the most loss of structure of the α-helices as determined by calculated NOEs.

A network of experimentally observed NOEs exist in the β-sheet of holocytochrome b_5 [Guiles et al., 1990; see Figure 6 Moore and Lecomte (1993)] and are retained in apocytochrome b_5 with the exception of the interactions between $\beta4/\beta5$ (Moore & Lecomte, 1993). All of these β-sheet NOEs were observed throughout the entire MD simulation, except for two NOEs between $\beta2/\beta3$ and the three between $\beta4/\beta5$ (Guiles et al., 1990). As mentioned previously, $\beta2$ (residues 74−80) is irregular in the crystal structure. Furthermore, crystallographic studies indicate that there is a bend or kink in $\beta2$ and that Ile 75, which is located in the middle of the strand, does not participate in the β-structure (Mathews et al., 1979). In addition, the hydrogen bond between the amide hydrogen of residue 77 and the carbonyl oxygen of residue 29 is long in the crystal structure, 3.8 Å. The NOEs corresponding to this region exceeded 5 Å in the simulation for the first 300−500 ps; the mean values for the hydrogen−hydrogen interactions between CαH30−NH76 and CαH30−CαH75 were 5 and 5.6 Å, respectively. After 500 ps, however, the distances dropped to well below 5 Å. The loss of the contacts between $\beta4/\beta5$ occurred first at the N-terminus of $\beta5$ with residue 51. Eventually, distances for the $\beta4/\beta5$ region reached ∼15 Å.

Bovine Apocytochrome b_5: Tertiary Structure. The starting conformation for the simulation of bovine apocytochrome b_5 was the bovine crystal structure without the heme group. The C_α RMS deviation of the bovine apo simulation from the crystal structure is depicted in Figure 2. A difference of approximately 1.5 Å was observed between cores 1 and 2, relative to the crystal structure. Interestingly, the rat form had a higher C_α RMS deviation for core 1.

Residues 31−80 (not shown) displayed C_α RMS deviations of ≥3 Å. This region includes the C-terminal portion of core 1, extending into $\beta2$. In the latter part of the simulation (1200−1600 ps), residues 16−18 exhibited increased deviations, as well. This behavior differed from that of the rat form, where deviations were localized almost entirely to core 1 (Figure 3). As in the rat simulation, core 2 deviated much

less from the crystal structure than core 1 (Figure 5). Core 1, however, was much less mobile than that of rat b_5. In addition, the $\beta4/\beta5$ connection was never completely lost, as was observed for rat b_5 (Moore & Lecomte, 1993). Moreover, the connectivities between $\beta1/\beta2$ (CαH6−NH78, NH7−NH78) and $\beta2/\beta3$ (CαH30−CαH75, CαH30−NH76) were disrupted, unlike the rat b_5 simulation. The regions that differ between the two simulations contain four of the six variant residues (Figure 1).

Secondary Structure. The secondary structure elements of the bovine protein were analyzed via the methodology described above. All helices maintained high helical contents throughout the simulation, with the exception of $\alpha2$ and $\alpha6$ (Table 1). For the first 500 ps, the helical content of $\alpha2$ remained below 42%, after which it rose to ≥88% and remained high for the duration of the simulation. Large fluctuations were observed for $\alpha6$, as was seen in the rat simulation. In contrast to the rat simulation, however, the overall α-helical content remained somewhat higher at 31%.

The β-structure, which comprises much of core 2, experienced much greater fluctuations. The β-content of $\beta1$ dropped to 0% by 900 ps, where it remained for the duration of the simulation. $\beta2$ and $\beta5$ had very low β-content throughout the simulation. $\beta3$ and $\beta4$, the central strands in the β-sheet, maintained relatively high β-contents. Note that although $\beta3$ had only 65% β-structure, it is the same as that observed in the crystal structure. $\beta4$ maintained high β-content (≥94%) until 1100 ps when it dropped to 70−76%.

In addition, $i \rightarrow i + 3$ α-helical hydrogen−hydrogen NOE distances were measured, and almost all of them remained in the detectable range (≤5 Å) throughout the simulation. The distances of $\alpha2$ became slightly extended (≥4 Å) from 200 to 1100 ps, and those of $\alpha6$ reached higher values (≥4 Å) between 200−400 and 800−1600 ps. In contrast to the rat simulation, none of the contacts corresponding to NOEs in the helices exceeded 5 Å.

Many of the experimentally observed β-sheet NOE connectivities for core 2 of rat apocytochrome b_5 were lost for the bovine apoprotein (Guiles et al., 1990). The corresponding values for the $\beta1/\beta2$ contacts, CαH6−NH78 and NH7−NH78, were 5.4 and 6.9 Å, respectively. The $\beta2/\beta3$ NOEs for CαH75−CαH30 and CαH76−CαH30 were 6.2 and 5.2 Å. The $\beta4/\beta5$ contacts CαH23−NH51, CαH23−NH52 and NH24−CαH53, which were lost in both the rat simulation and experiments, were 5.8, 4.2, and 3.8 Å, respectively, in the bovine simulation. Additionally, the NOE contacts between residues 22 and 76 of the hydrophobic core were lost after 500 ps. Interestingly, all of these NOEs were retained in the MD simulation of bovine holocytochrome b_5 (Storch & Daggett, 1995).

Structural Effects of Variant Residues. The local structural effects of the variant residues were investigated in order to elucidate the observed global differences between the MD simulations of rat and bovine apo b_5. The variant residues are given as changes from the bovine to the rat sequence. Two of the six residues, Tyr 27 → His and Phe 74 → Tyr, showed no obvious stabilizing or destabilizing interactions in the two simulations.

One of the local effects of the Ala 3 → Asp substitution was a salt bridge that formed between Asp 3 and Lys 5 in the rat simulation. This interaction is not present in the crystal structure and formed during the first 50 ps of the rat simulation at pH 6.9 and was maintained for the rest of the simulation. The corresponding residues in the bovine simulation, Ala 3 and Lys 5, were unable to interact in such a way, which resulted in increased conformational sampling in $\beta1$.

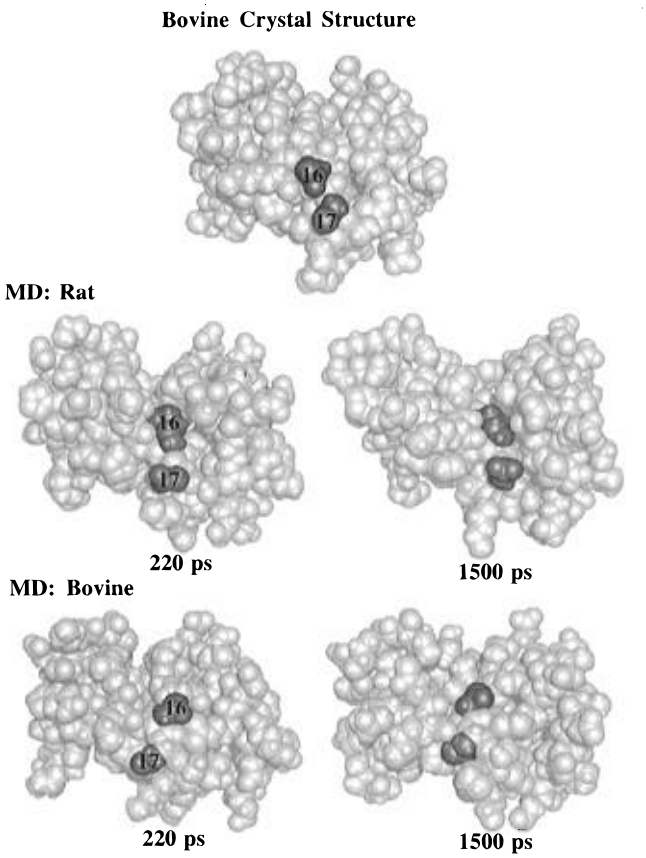

FIGURE 6: Space-filling representations of rat and bovine apocytochrome b_5. The variant residues in the rat (Lys 16 and Asp 17) and bovine (Asn 16 and 17) proteins are shown in gray. The crystal structure is shown for reference.

The Asn 16 → Lys and Asn 17 → Asp substitutions are illustrated in Figure 6. Residue 16 is located in the loop between $\alpha1$ and $\beta4$. In the rat simulation, the charged amine of Lys 16 burrowed between the two hydrophobic cores by 420 ps and remained in this position for the rest of the simulation. Although Asn 16 showed some inclination toward such displacement in the bovine simulation, the structural effects were not as dramatic. Thus, Lys 16 contributed to the disruption of interactions between $\beta4$ and $\beta5$, while the neutral Asn 16 did not. Furthermore, Asn 16 formed a stabilizing hydrogen bond with the carbonyl of Gly 51 in the bovine simulation. In both proteins, residue 17 remained on the surface completely solvent accessible and did not participate in obvious stabilizing or destabilizing interactions.

The structural effects of the Leu 23 → Val substitution resulted in sustained van der Waals contacts between Leu 23, Ala 50, and Ala 54 in the bovine simulation, which constituted the prominent stabilizing interactions between $\beta4$ and $\beta5$. This effect was not seen in the rat simulation, and the $\beta4/\beta5$ sheet became disrupted.

DISCUSSION

Initially, we intended to simulate the bovine form of apocytochrome b_5 and compare to the experimental data on

rat apocytochrome b_5, as the available crystal structure is of the bovine form. The sequence identity between the rat and bovine sequences (for the water-soluble fragment of the protein used for simulation and experiment) is 93%, which corresponds to six amino acid substitutions. The high degree of homology suggests that the two forms should be structurally similar, if not identical. Furthermore, a control simulation of bovine holocytochrome b_5 (Storch & Daggett, 1995) is well behaved as determined by comparison to structural experimental data (Veitch et al., 1988; Guiles et al., 1990; Moore & Lecomte, 1990; Moore et al., 1991). However, during the simulation of bovine apocytochrome b_5 presented here, we observed progressively more deviations from what is expected on the basis of experiments on the rat protein. The disagreement was localized to specific areas of the protein located near four of the variant residues.

On the basis of this observation, six residue substitutions were made to the bovine protein, and we subsequently simulated "rat apocytochrome b_5". The rat simulation was in closer agreement to experiment. In this discussion, we first compare our results with the data on the rat protein and then present findings that illustrate the subtle differences in structure and dynamics between the rat and bovine simulations.

Rat Apocytochrome b_5. Previous studies indicate that apocytochrome b_5 retains limited native-like structure upon removal of the prosthetic heme group. The structure of the apoprotein is less compact, experiences an increase in disorder, and changes in secondary structure elements (Huntley & Strittmatter, 1972; Tajima et al., 1976). A variety of NOEs remain upon removal of the heme, and four of the five β-strands ($\beta1-\beta4$) are configured as in the holoprotein by 2-D NMR studies (Moore & Lecomte, 1993). Lack of resolution and weak NOE patterns have hindered the assignment of all helical NOEs, and Moore and Lecomte concluded that the two terminal helices ($\alpha1$ and $\alpha6$) are present, as in the holoprotein, with conformational multiplicity occurring in the others. Furthermore, characterization of core 2 has provided details regarding a stable unit in the partially unfolded form (Moore & Lecomte, 1990, 1993).

As indicated in Figures 2−4, the MD simulation of rat apocytochrome b_5 is in agreement with the general observation of core stability. That is, the residues of core 1 both deviated substantially from the crystal structure and experienced a conformational flexibility not observed in core 2. Numerous side-chain interactions lead to the damping of mobility in core 2 (Figure 4), as observed by Moore and Lecomte (1990).

Our studies indicate that loss of structure was localized to particular regions of the protein. Of the heme-binding helices, $\alpha3$ and $\alpha5$ experienced the greatest internal fluctuations in structure, and both of these, in addition to $\alpha4$, became displaced from their orientations in the crystal structure (Figures 2 and 5). $\alpha5$ also became quite disrupted at its N-terminus as the simulation progressed. Characteristic α-helical NOEs were lost for $\alpha5$, and it displayed a marked movement away from the $\beta3/\beta4$ loop. It should be reiterated that although experiment detects fluctuations in core 1 of the protein (Moore & Lecomte, 1993), the ability to resolve the exact magnitude and precise placement of the motion and structural disruptions is limited. Our simulations provide a structural model in which secondary structure disruption to the α-helices is limited, with $\alpha3$, $\alpha4$, and $\alpha5$ being most affected in accord with experiment (Falzone et al., 1996), although all of the core 1 α-helices were highly mobile (Figures 2, 4, and 5).

Core 2 contains primarily the β-sheet and α-helices 1 and 6. The simulation revealed that the majority of the structural stability in this unit lies within $\beta1$, $\beta3$, and $\beta4$. The C-terminal tail of $\alpha6$ exhibited multiple conformations. Structure was progressively lost in this helix; however, the retention of the experimental NOE connectivities between NH7 and both NH78 and NH80 ($\beta1$ and $\alpha6$, respectively) (Moore & Lecomte, 1993) throughout the simulation reflected the fact that tertiary structure in this region remained intact.

In contrast to the stability of the majority of core 2, conformational multiplicity of $\beta5$ and the nearby α-helical regions ($\alpha3$ and $\alpha4$) of core 1 was observed. The NOEs lost between $\beta4$ and $\beta5$ lead to disruptions in the interface between cores 1 and 2, thus allowing core 1 to become less ordered. The attachment of $\beta5$ to the β-sheet is nonidentifiable in the apoprotein 2-D NMR experiments (Moore & Lecomte, 1993). Comparison of solvent accessibility of the residues in apocytochrome b_5 and holocytochrome b_5 revealed that residues in $\beta4$ (21−25) gained varying amounts of exposure, as has also been seen experimentally (Moore & Lecomte, 1993).

Bovine Apocytochrome b_5 and Structural Effects of the Variant Residues. The amount of experimental information for bovine apocytochrome b_5 is practically nonexistent (Huntley & Strittmatter, 1972). Therefore, we compared MD simulations of bovine and rat apocytochrome b_5 to explore their structural differences and similarities. Core 1 of the bovine form possessed a greater deal of conformational flexibility relative to core 2, as was also observed in the rat apo simulation (Figures 2 and 5). The increased main-chain mobility, resulting from heme removal, extended through residue 80, located in core 2. In contrast, the mobility present in the rat simulation was confined to core 1 (excluding the C-terminus), ending at residue 75 (Figure 3).

Although a "loosening" of core 1 occurred relative to the starting conformation, the overall secondary structure content of the helices was higher than in the rat form, with $\alpha2$ being the most disrupted. Experimentally, a structural disruption at the interface between cores 1 and 2 is observed for the rat protein (Moore & Lecomte, 1993). Two of the three variant residues located in this region are near the interface of cores 1 and 2; Asn 16 → Lys, Asn 17 → Asp, and Leu 23 → Val (bovine → rat). In both proteins, residue 16 moved in toward the interface (Figure 6). Lys 16 contributed to the disruption of the hydrophobic interactions between cores 1 and 2 in the rat simulation, while Asn 16 in the bovine simulation did not achieve the same effects and instead formed some stabilizing hydrogen bonds with the loop between $\alpha3$ and $\beta5$. In both cases, residue 17 was not critical to the structure and was completely accessible to solvent. An important stabilizing interaction at the interface of the two cores in the bovine simulation was a localized hydrophobic cluster involving Leu 23, Ala 50, and Ala 54. In the rat protein Val 23 interacted intermittently with Ala 54 but not Ala 50. This interaction and the lack of disruption by Lys 16 as seen in the rat simulation, were the prominent factors leading to a more stable core 1 for bovine apo b_5.

The secondary structure content of core 2 was lower overall in the bovine simulation. In contrast to the rat form,

β1 had low β-content and β4 displayed high deviations. Furthermore, the conformational deviations of β1 and β2 had an effect on tertiary contacts and the overall stability of the sheet. These two regions contain three variant residues: Ala 3 → Asp, Tyr 27 → His, and Phe 74 → Tyr (bovine → rat). The Ala 3 → Asp variant in the rat simulation showed drastic improvements in connectivity between the β1/β2 sheet. In addition, a salt bridge formed between Asp 3 and Lys 5, which stabilized the N-terminus of β1, forming a more rigid "appendage" that remained in contact with the main portion of the protein. The lack of the rigidity in this region of the bovine protein allowed for more conformational sampling of the N-terminus. The added constraint in the rat simulation was also reflected in the secondary structure content of β1, which was 92% in the rat form versus 25% in bovine (Table 1). Although the last two variant residues (27 and 74) are located in the region where the rat simulation shows greater agreement with experiment, there were no obvious local stabilizing interactions that occurred as a consequence of the substitutions. It is more likely that the subsequent stabilization between β2 and β3 was due to the indirect effects of other substitutions.

CONCLUSIONS

Experimental studies suggest that core 1 of rat apocytochrome b_5 is well maintained during the beginning stages of the unfolding process, while core 2 is conformationally heterogeneous and highly dynamic (Moore & Lecomte, 1990, 1993; Moore et al., 1991; Falzone et al., 1996). Our simulations revealed a similar pattern of structural and dynamic variation between the two cores. Core 2 remained compact and intact while core 1 experienced some loss of tertiary and secondary structure. A decrease in β-structure was observed at the interface of the cores, and portions of the α-helical structure in core 1 were disrupted as the entire core became dynamic and conformationally heterogeneous. We would suggest that α-helices in core 1 are basically intact, with the possible exception of α5, which becomes quite disrupted. These structural characteristics are consistent with NMR experiments of rat apo b_5 (Falzone et al., 1996), although we may be overpredicting α-helical content for α4. Even though the helical content of helices in core 1 is high in the simulation, they become highly dynamic and may confound experiments designed to monitor their structural properties. For example, the difficulty in resolving the NMR cross peaks in core 1 (Moore & Lecomte, 1993) may be due to the dynamic behavior of the apoprotein as opposed to extensive loss of helical structure, as has been observed for apocytochrome b_{562} (Feng et al., 1994; Laidig & Daggett, 1996a).

Initially our studies were performed on bovine apocytochrome b_5. Unfortunately, while displaying the same general properties as the rat simulation, many of the NOEs predicted from the MD-generated model did not accurately reproduce some of the more precise structural NMR information available. Namely, the bovine simulation showed too much structure in some areas of the protein and too little in others. Therefore, we constructed and proceeded with a model of the rat form. Although the sequence changes appear to be innocuous when viewed in the crystal structure, they have a marked effect on the structure and dynamics of the apoprotein. Interestingly, this structural variance only occurs once the heme group is removed, since the MD simulation of bovine holocytochrome b_5 correlates very well with the rat holocytochrome b_5 experimental data (Storch & Daggett, 1995).

ACKNOWLEDGMENT

We thank Dr. Juliette Lecomte for interesting discussions, for providing many helpful comments on the manuscript, and for sharing unpublished data. All molecular graphics images (except Figure 1) were produced using the MidasPlus program from the Computer Graphics Laboratory, University of California, San Francisco (supported by National Institutes of Health Grant RR-01081) (Ferrin et al., 1988; Huang et al., 1991).

REFERENCES

Altman, J., Lipka, J. J., Kuntz, I. D., & Waskell, L. (1989) *Biochemistry 28*, 7516−7523.
Breslow, E., Beychok, S., Hardman, K. D., & Gurd, F. R. N. (1965) *J. Biol. Chem. 240*, 304−309.
Brooks, C. L., III (1992) *J. Mol. Biol. 227*, 375−80.
Daggett, V., & Levitt, M. (1992) *J. Mol. Biol. 223*, 1121−1138.
Daggett, V., Kollman, P. A., & Kuntz, I. D (1991) *Biopolymers 31*, 1115−1134.
Falzone, C. J., Mayer, M. R., Whiteman, E. L., Moore, C. D., & Lecomte, J. T. J. (1996) *Biochemistry* (in press).
Feng, Y. Q., & Sligar, S. G. (1991) *Biochemistry 30*, 10150−10155.
Feng, Y. Q., Wand, A. J., & Sligar, S. G. (1991) *Biochemistry 30*, 7711−7717.
Feng, Y. Q., Sligar, S. G., & Wand, A. J. (1994) *Nat. Struct. Biol. 1*, 30−35.
Ferrin, T. E., Huang, C. C., Jarvis, L. E., & Langridge, R. (1988) *J. Mol. Graphics 6*, 13−27.
Fink, A. L. (1995) *Subcell. Biochem. 24*, 27−53.
Gonzalez, F. J., & Kasper, C. B. (1980) *Biochemistry 19*, 1790−1796.
Guiles, R. D., Altman, J., Kuntz, I. D., & Waskell, L. (1990) *Biochemistry 29*, 1276−1289.
Hara, T., & Minakami, S. (1970) *J. Biochem. (Tokyo) 67*, 741−743.
Hara, T., Tanaka, S., & Minakami, S. (1970) *J. Biochem. (Tokyo) 68*, 805−810.
Harrison, S. C., & Blout, E. R. (1965) *J. Biol. Chem. 240*, 299−303.
Huang, C. C., Pettersen, E. F., Klein, T. E., Ferrin, T. E., & Langridge, R. (1991) *J. Mol. Graphics 9*, 230−236.
Hughson, F. M., Wright, P. E., & Baldwin, R. L. (1990) *Science 249*, 1544−1548.
Hughson, F. M., Barrick, D., & Baldwin, R. L. (1991) *Biochemistry 30*, 4113−4118.
Huntley, T. E., & Strittmatter, P. (1972) *J. Biol. Chem. 247*, 4641−4647.
Jones, M. S., & Jones, O. T. (1969) *Biochem. J. 113*, 507−514.
Kell, G. S. (1967) *J. Chem. Eng. Data 12*, 66−69.
Kraulis, P. (1991) *J. Appl. Crystallogr. 24*, 946−950.
Krieter, P. A., & Shires, T. K. (1980) *Biochem. Biophys. Res. Commun. 94*, 606−611.
Laidig, K. E., & Daggett, V. (1996a) *Folding Des. 1*, 335−346.
Laidig, K. E., & Daggett, V. (1996b) *J. Phys. Chem. 100*, 5616−5619.
Lecomte, J. T. J., & Cocco, M. J. (1990) *Biochemistry 29*, 11057−11067.
Lecomte, J. T. J., & Moore, C. D. (1991) *J. Am. Chem. Soc. 113*, 9663−9665.
Lederer, F. (1994) *Biochimie 76*, 674−692.
Levitt, M. (1983) *J. Mol. Biol. 168*, 595−617.
Levitt, M. (1989) *J. Mol. Biol. 181*, 423−447.
Levitt, M. (1990) *ENCAD−Energy Calculation and Dynamics*, Molecular Applications Group, Stanford, CA.
Levitt, M., Hirshberg, M., Sharon, R., & Daggett, V. (1995) *Comput. Phys. Commun. 91*, 215−231.
Mathews, F. S., Argos, P., & Levine, M. (1972) *Cold Spring Harbor Symp. 36*, 387−395.

Mathews, F. S., Czerwinski, E. W., & Argos, P. (1979) The X-ray crystallographic structure of calf liver cytochrome b_5. In *The Porphyrins*, Vol. VII, pp 107−147, Academic Press, New York.
Moore, C. D., & Lecomte, J. T. J. (1990) *Biochemistry 29*, 1984−1989.
Moore, C. D., & Lecomte, J. T. J. (1993) *Biochemistry 32*, 199−207.
Moore, C. D., Ousaima, N. A., & Lecomte, J. T. J. (1991) *Biochemistry 16*, 4975−4978.
Negishi, M., & Omura, T. (1970) *J. Biochem. (Tokyo) 67*, 745−747.
Ozols, J. (1989) *Biochim. Biophys. Acta 997*, 121−130.
Ozols, J., Gerard, C., & Nobrega, F. G. (1976) *J. Biol. Chem. 251*, 6767−6774.
Pfeil, W. (1993) *Protein Sci. 2*, 1497−1501.
Shawver, L. K., Seidel, S. L., Krieter, P. A., & Shires, T. K. (1984) *Biochem. J. 217*, 623−632.
Storch, E. M., & Daggett, V. (1995) *Biochemistry 34*, 9682−9693.
Tajima, S., Enomoto, K., & Sato, R. (1976) *Arch. Biochem. Biophys. 172*, 90−97.
Teale, F. W. J. (1959) *Biochim. Biophys. Acta 35*, 543.
Veitch, N. C., Concar, D. W., Williams, R. J., & Whitford, D. (1988) *FEBS Lett. 238*, 49−55.
Wüthrich, K. (1986) Nuclear Overhauser effects, $^1H-^1H$ distances, and molecular dynamics, in *NMR of Proteins and Nucleic Acids*, 1986, pp 111−113, John Wiley & Sons, Inc., New York.

BI960598G