THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

pubs.acs.org/JPCB

Open Access on 01/23/2015

Filling Up the Heme Pocket Stabilizes Apomyoglobin and Speeds Up Its Folding

J. S. Goodman,[†] S.-H. Chao,[‡] T. V. Pogorelov,*[,†,§] and M. Gruebele*[,†,‡,∥]

[†]Department of Chemistry, University of Illinois, Urbana, Illinois 61801, United States

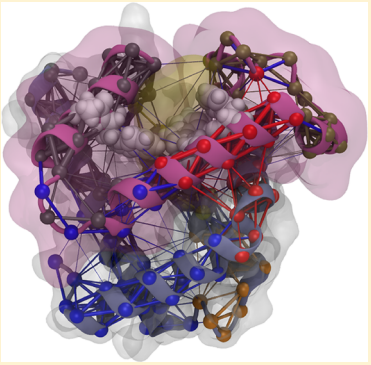
[‡]Department of Physics, Center for the Physics of Living Cells, University of Illinois, Urbana, Illinois 61801, United States

[§]School of Chemical Sciences, University of Illinois, Urbana, Illinois 61801, United States

[∥]Center for Biophysics and Computational Biology, University of Illinois, Urbana, Illinois 61801, United States

Supporting Information

ABSTRACT: Wild type apomyoglobin folds in at least two steps: the ABGH core rapidly, followed much later by the heme-binding CDEF core. We hypothesize that the evolved heme-binding function of the CDEF core frustrates its folding: it has a smaller contact order and is no more complex topologically than ABGH, and thus, it should be able to fold faster. Therefore, filling up the empty heme cavity of apomyoglobin with larger, hydrophobic side chains should significantly stabilize the protein and increase its folding rate. Molecular dynamics simulations allowed us to design four different mutants with bulkier side chains that increase the native bias of the CDEF region. *In vitro* thermal denaturation shows that the mutations increase folding stability and bring the protein closer to two-state behavior, as judged by the difference of fluorescence- and circular dichroism-detected protein stability. Millisecond stopped flow measurements of the mutants exhibit refolding kinetics that are over 4 times faster than the wild type's. We propose that myoglobin-like proteins not evolved to bind heme are equally stable, and find an example. Our results illustrate how evolution for function can force proteins to adapt frustrated folding mechanisms, despite having simple topologies.



INTRODUCTION

Myoglobin is the prototype of the globin fold,[1] with eight helices A−H arranged in two cores, ABGH and CDEF. The latter contains a non-covalently bound heme group. When the heme group is removed, the resulting apomyoglobin molecule (apoMb) still folds into a native-like structure,[2] with some loss of secondary structure particularly in the F helix.[3]

ApoMb is one of the first proteins for whose refolding detailed structural predictions were made.[4] In the collision-diffusion model, secondary structure elements fluctuate in and out of existence, and when they diffuse together, tertiary contacts form. In the denatured state, the AB and GH fragments of the ABGH domain are separated by a long linker made up from the central CDEF helices, which themselves are directly connected and have a much smaller contact order. Thus, it is not surprising that the model predicted the CDEF core would form first, followed by AB and GH docking onto it later.

When refolding experiments were carried out, the exact opposite turned out to occur. Stopped flow experiments monitored by circular dichroim and amide proton protection showed that the ABGH core of apoMb was first to form, with CDEF following later in a separate step.[5] Indeed, the ABGH core formed so fast that its folding kinetics could not be resolved until laser T-jump refolding experiments revealed that it folds in ∼10 μs when monitored by fluorescence of apoMb's two tryptophan residues, both contained in the A helix.[6,7]

The discrepancy is easily rationalized in terms of the missing heme group, a very large hydrophobic molecule that sits at the very core of the CDEF helices. Without the prosthetic group, the side chains in CDEF are missing most of their hydrophobic contacts, and no significant packing occurs in the native structure (Figure 1). Thus, CDEF cannot form a stable hydrophobic core, key for a stable tertiary structure. On the other hand, ABGH is a well-packed core.[3,8] The linker connecting AB and GH is about 70 residues long, but it has been shown by loop contact formation rate measurements that such contacts can occur on a 100 ns time scale.[9] Fast folding experiments have revealed a changing tryptophan environment on that time scale, followed by native-like fluorescence in 5−20 μs depending on the protein mutant.[7] Thus, the higher helix propensity and more stable hydrophobic core of ABGH win out over the smaller contact order of CDEF.

CDEF is the functional core of apoMb, and it has been noted that amino acid residues present for function (here: to pack around the oxygen-carrying heme group) often frustrate

Special Issue: William C. Swope Festschrift

Received: December 19, 2013
Revised: January 21, 2014
Published: January 23, 2014

   dx.doi.org/10.1021/jp412459z | J. Phys. Chem. B 2014, 118, 6511−6518

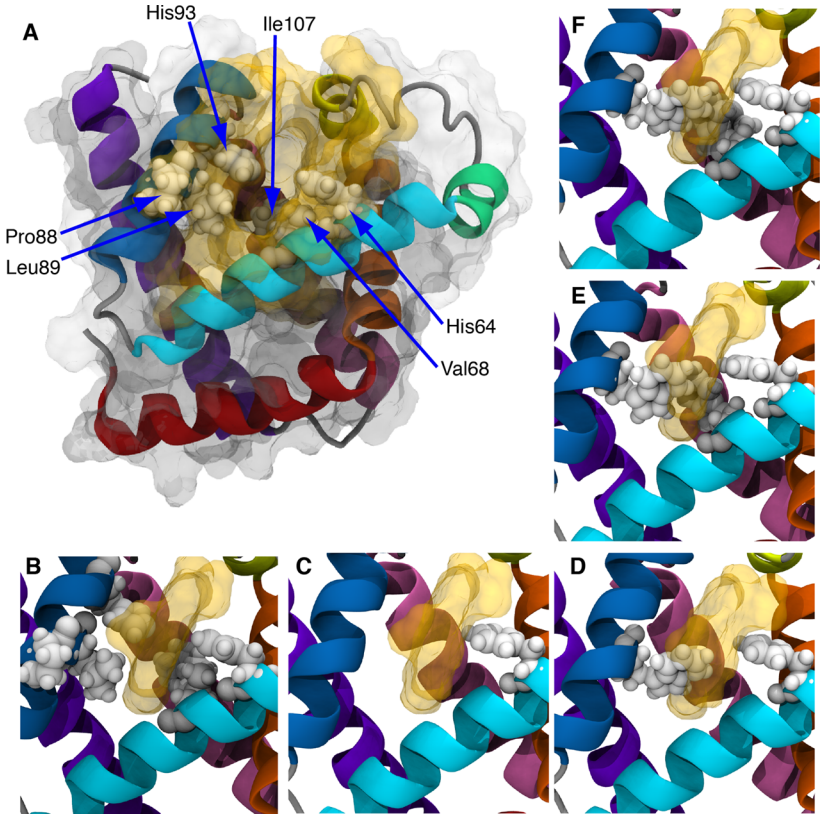


Figure 1. Apomyoglobin (A and B) and its mutants: apoMb2 (C), apoMb3 (D), apoMb4 (E), and apoMb5 (F). Helices are shown in the cartoon representation and colored as follows: A (red), B (orange), C (yellow), D (green), E (cyan), F (blue), G (dark pink), and H (magenta). Residues lining the heme group pocket are shown as a transparent surface (yellow in A). Parts A and B highlight all mutation sites as well as the heme-binding His 93 residue. Parts C–F highlight mutations as white van der Waals surfaces, excluding Pro88Ala for clarity.

folding.[10,11] Folding is favored by highly stable and therefore rigid secondary and tertiary structures with large hydrophobic cores, whereas function often requires metallo-centers or polar/charged residues in a flexible environment to adaptively bind to small ligands or other biomacromolecules. Thus, apoMb is a three-state folder with a folding mechanism at least as complex as U → ABGH → ABCDEFGH = N.[12]

This naturally leads to the question whether that empty core can be stabilized so the folding of CDEF is more on par with ABGH. In principle, it should be possible even to speed up CDEF beyond ABGH due to the former's smaller contact order. In practice, it may not be possible to insert sufficiently many large natural amino acid side chains to reach two-state folding as was done for RNaseH,[13] but more globally cooperative folding should be possible. Such stabilization and speed-up would illustrate that the three- (or more) state mechanism of apoMb is a consequence of evolution for heme-binding and oxygen carrying function, not a necessity of the globin fold.

We examined the packing of CDEF computationally and designed core mutations expected to speed up folding. Such replacements include replacement of a proline known already to disrupt the F helix, substitution of a needless heme-binding histidine by phenylalanine, and various new mutations of heme-lining side chains illustrated in yellow at the top of Figure 1A. We designed a sequence of mutants going from two to five substitutions in and near the heme pocket.

We tested the sequence of mutants by thermal and guanidine hydrochloride (GuHCl) titrations to determine protein stability and by stopped flow experiments to monitor changes in the speed of refolding. We find as a general trend that apoMb becomes progressively more stable as the heme pocket is filled with hydrophobic side chains and folds progressively faster as detected by tryptophan fluorescence. The fastest time we achieve is <13 ms, not yet as fast as the formation of the ABGH core in wild type horse apomyoglobin but considerably closer. This observation suggests that a natural hemeless myoglobin analogue could perhaps fold via an apparent two-state mechanism.

METHODS

Protein Systems. All of our mutants are based on horse heart apomyoglobin.[3] Table 1 summarizes the sequence of mutants that were designed on the basis of MD simulations,

Table 1. Nomenclature of apoMb Mutants and Melting Temperatures Derived from the Global Model Fits (1 σ Uncertainty Including Global Parameter Correlations <0.8 °C)

name	mutations	T_m (CD_{222nm})	T_m (mean λ_{fl})
ApoMbWT	wild type[a]	65.5 °C[14]	55.0 °C[15]
ApoMb2	H64F; P88A	85.7 °C	52.6 °C
ApoMb3	H64F; P88A; L89W	84.1 °C	78.6 °C
ApoMb4	H64F; P88A; L89W; V68W	83.6 °C	82.8 °C
ApoMb5	H64F; P88A; L89W; V68W; I107M	81.9 °C	78.6 °C

[a]Our remeasurements yield a lower value of 58 °C by CD, using the two-state model in the Methods section.

dx.doi.org/10.1021/jp412459z | J. Phys. Chem. B 2014, 118, 6511–6518

and expressed for experimental studies. The sites of mutation and arrangement of substituted side chains are shown in Figure 1.

Molecular Dynamics Simulations. The coordinates of wild type protein were obtained from a crystal structure of horse heart myoglobin (PDB ID: 3LR7, resolution of 1.60 Å[16]) with the coordinates of the heme prosthetic group excluded. Mutants were prepared using the *Mutator Plugin* of Visual Molecular Dynamics (VMD) software.[17] Each system was minimized for 1000 time steps, and equilibrium MD was performed for 20 ns using NAMD2 software[18] with the CHARMM27 force field,[19] the TIP3P model of water,[20] and the CMAP corrections[21] for proteins in the NPT ensemble. Langevin dynamics with a 0.5 ps^{-1} damping coefficient and the Langevin piston Nosé−Hoover procedure were used to maintain temperature and pressure to constant values (300 K and 1 atm).[22,23] The particle mesh Ewald (PME) method was used to calculate long-range electrostatic forces without truncation, using a grid density of 1 $Å^{-3}$.[24] The van der Waals interaction cutoff value was 12 Å, while the integration time steps for bonded, nonbonded, and PME calculations were 2, 1, and 2 fs, respectively. Trajectory geometry analyses (RMSD, RMSF, and R_{gyr}) were performed using Tcl scripts within VMD.

Void Volume Calculations. To analyze empty volume (voids) within the heme cavity, 20 ns simulations were used and data was extracted one structure per nanosecond. van der Waals radii for the atoms were assigned from the CHARMM27 force field using VMD. Voids were identified using the McVol software.[25] McVol uses a Monte Carlo approach in combination with a fine grid to classify points that are parts of the protein, solvent, or voids (cavity or cleft). Protein surfaces were defined with a probe radius of 1.1 Å. The surface of each atom was represented by a maximum of 2500 points. The cavity grid spacing was set to 0.5 Å and the minimum cavity volume to 7 $Å^3$. The heme pocket was defined by the residues within 5 Å of heme in the wild type structure (PDB ID 3LR7). The void volume was calculated by counting the number of probes (volume of 5.6 $Å^3$) within 4 Å of C_α atoms in the heme cavity.

Network Analysis. The dynamical networks were analyzed using the NetworkView plugin in VMD.[17,26] The last 10 ns of the simulation data were used with a time step of 20 ps. Briefly, a network is constructed of nodes (C_α) and connecting edges. An edge connects two nodes if they are in contact (4.5 Å), excluding nearest neighbors. The weight of the edge is the probability of information transfer between nodes, measured by correlation in the dynamics between the two nodes.[26] Community analysis of the networks was also performed using the Girvan−Newman algorithm[27] as implemented in the NetworkView plugin.[17,26]

Protein Expression and Purification. A DNA fragment coding for apoMb2 was inserted between the BamHI and NdeI restriction sites of plasmid p-ET15b (Genscript, Piscataway, NJ). Subsequent mutant vectors were synthesized by site-directed mutagenesis (Strategene, La Jolia, CA) of apoMb2 and were amplified in BL21 DE3 cells at 37 °C with 100 mg/L ampicillin. After reaching an OD_{600} value of 0.6−0.8, the cells were induced with 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) at 20 °C for 12 h. The cells turned red upon expression and were sonicated in lysis buffer (300 mM NaCl, 50 mM Na_2HPO_4, and 10 mM imidazole, pH 8). The mutants were purified by Ni-NTA chromatography, washing with phosphate buffer (300 mM NaCl, 50 mM Na_2HPO_4, pH 8) at different imidazole concentrations starting from 20 to 500 mM. Heme molecules were extracted using 2-butanone at 4 °C,[28] and the N-terminal histadine tags were removed by thrombin cleavage (EMD Millipore, Billerica, MA). The mutants were dialyzed twice in 10 mM sodium acetate, pH 6, for 6 h at 4 °C prior to taking measurements.

The N-terminal domain of protein RsbR was similarly expressed and purified (without the need for heme extraction). The sequence was residues 1−143 of the protein from PDB structure 1OR4.[29]

Thermal and Chemical Denaturation. Circular dichroism (CD) spectra for all mutants were obtained using a Jasco-715 spectropolarimeter (Jasco Inc., Easton, MD) at a scan rate of 100 nm/min in 1 nm intervals. All GuHCl-containing samples were collected from 220 to 250 nm and 200 to 250 nm in 0 M GuHCl samples. All thermal melts at a constant GuHCl concentration were done using a Peltier temperature controller covering from 20 to 92 °C in 3° increments. Room temperature GuHCl melts were measured by hand-mixing appropriate denaturant volumes with the protein solution. ApoMb2 was measured at 30 μM, while all other samples were measured at 15 μM using a 1.0 mm path length cuvette.

Tryptophan fluorescence spectra for each mutant was collected from 290 to 450 nm at a scan rate of 600 nm/min in 1 nm intervals using a Cary Eclipse fluorimeter (Agilent Technologies, Santa Clara, CA). All samples were measured at 10 μM with a 280 nm excitation wavelength and a photomultiplier voltage ranging between 580 and 650 V. All thermal melts were done with a Peltier temperature controller within the same temperature range as the CD measurements.

Stopped Flow Refolding. The folding kinetics for each mutant was measured using a custom-built stopped flow apparatus (Unisoku Co. Ltd., Osaka, Japan) with a 1:6 mixing ratio. The dead time of the instrument was ≤20 ms. Refolding was observed by mixing one part of 2 M GuHCl denatured samples with six parts of 0 M GuHCl buffer in a 1.0 mm path length sapphire cell containing a 50 μM window, resulting in a 0.3 M final GuHCl concentration. Fluorescence relaxation was monitored by exciting with a titanium-sapphire laser (KMLabs Inc., Boulder, CO) that was frequency tripled to 280 nm with a third harmonic generator. Fluorescence was guided out of the stopped flow apparatus with an optical fiber, passed through a B370 band-pass filter (Hoya Corp., Santa Clara, CA), detected by a photomultiplier tube (R7400U-03, Hamamatsu Corp., Bridgewater, NJ), and collected by an oscilloscope with 2.5 GHz bandwidth and a 10 GS/s sampling frequency (DPO7254, Tektronix Inc., Beaverton, OR). The oscilloscope was set to record 2 μs data windows every 10 ms, triggered by a function generator (DG4102, Rigol Technologies Inc., Oakwood Village, OH). The total length of data collection was 3.1 s, and contains 312 data points from averaging the intensity of each 2 μs window using Matlab (Mathworks Inc., Natick, MA). Mixes were taken at one or two concentrations for each mutant. Sample concentrations were 200 μM for apoMbWT, 200 μM for apoMb2, 133 and 190 μM for apoMb3, 100 and 200 μM for apoMb4, and 100 μM for apoMb5.

Data Analysis. Melting curves for fitting were generated by plotting the mean residue elipticity (MRE) values at 222 nm for each CD spectrum and the mean wavelength of each tryptophan emission spectrum in Igor Pro (Wavemetrics Inc., Lake Oswego, OR). We used two models. Individual denaturation traces could be fitted by an apparent two-state

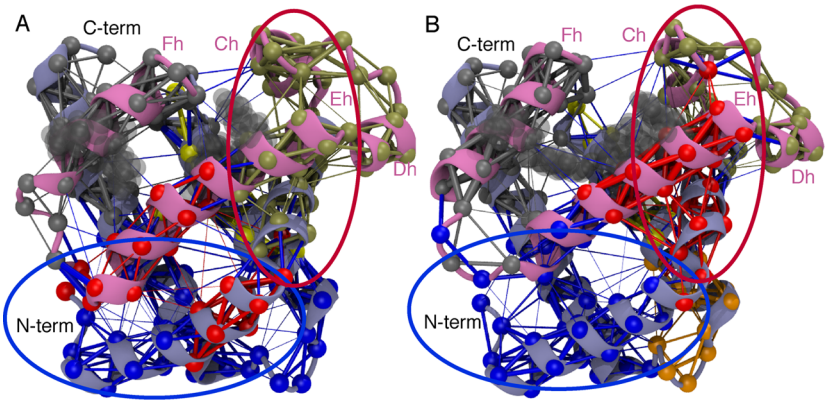


Figure 2. Network analysis of apoMbWT (A) and apoMb5 (B). ABGH (light blue) and CDEF (pink) cores are shown in the cartoon representation. Network nodes and edges are colored by community (blue, red, orange, yellow, tan, and gray). Ovals (red and blue) highlight the interbundle regions of network stabilization in apoMb5. CDEF helices are marked (Ch−Fh).

model with the free energy difference of native minus denatured state given by

$$\Delta G = g_{\mathrm{T}}^{(1)}(T - T_{\mathrm{m}})$$

$$\Delta G = g_{\mathrm{Gu}}^{(1)}(C - C_{\mathrm{m}})$$

depending on whether the titration was thermal or denaturant. T_{m} is the midpoint temperature, or melting temperature, of the thermal titration. C_{m} is the midpoint concentration of the denaturant titration in molar units. This first order Taylor expansion neglects heat capacity effects but was sufficient to fit all individual traces, which do not contain reliable information on the heat capacity of folding. In addition, multiple thermal and guanidine hydrochloride denaturation traces were fitted simultaneously by a "global" model for the free energy difference

$$T_0 = T_0^{(0)} + T_0^{(1)}C$$

$$\Delta G = g^{(0)} + g_{\mathrm{Gu}}^{(1)}C + g_{\mathrm{T}}^{(2)}(T - T_0)^2 + g_{\mathrm{T,Gu}}^{(3)}(T - T_0)^2C$$

This quadratic model allows for cold[30] and heat denaturation. $g^{(0)}$ is the free energy difference between native minus denatured state at the temperature T_0 of maximal stability of the protein. A linear guanidine hydrochloride dependence of T_0 was included in the fit, as well as a linear dependence of ΔG on denaturant concentration ($g_{\mathrm{Gu}}^{(1)}$ is the "*m*-value"[31]). In both models, the equilibrium constant was then calculated as

$$K_{\mathrm{eq}} = \mathrm{e}^{-\Delta G/RT}$$

The native state N and unfolded state U were allowed a linear baseline, to account for temperature dependence of the CD and fluorescence spectra in the folded and denatured states:

$$S_{\mathrm{N}} = a_{\mathrm{N}} + b_{\mathrm{Gu}}^{\mathrm{N}}[\mathrm{GuHCl}] + b_{\mathrm{T}}^{\mathrm{N}}(T - T_i)$$

$$S_{\mathrm{U}} = a_{\mathrm{U}} + b_{\mathrm{Gu}}^{\mathrm{U}}[\mathrm{GuHCl}] + b_{\mathrm{T}}^{\mathrm{U}}(T - T_i)$$

where $T_i = T_{\mathrm{m}}$ or T_0, depending on the model used. The final signal was then evaluated as

$$S = S_{\mathrm{N}}\frac{K_{\mathrm{eq}}}{1 + K_{\mathrm{eq}}} + S_{\mathrm{U}}\frac{1}{1 + K_{\mathrm{eq}}}$$

Kinetic traces for apoMbWT, apoMb2, apoMb4, and apoMb5 were fitted to a single-exponential function starting at $t = 0$ where mixing was initiated. In two traces, an additional slow exponential accounted either for bleaching of the tryptophan residues or a possible slow phase. Data at <20 ms was not fitted due to the instrument dead time.

RESULTS

Molecular-Dynamics-Based Design of apoMb Mutants. To investigate the role mutations play in filling the heme pocket, apoMb wild type (apoMbWT) and successive variants with two to five mutations (Table 1) were modeled using equilibrium molecular dynamics simulations with explicit water. We started out with two mutations already known from the literature to contribute to greater stability, secondary structure content, and faster folding of CDEF.[32,33] Mutant apoMb2 (H64F, P88A) adds a bulky Phe 64 residue in the E helix in place of the now useless heme binding histidine and an F-helix-stabilizing Ala 88 side chain instead of the helix-destabilizing proline residue. Additional mutations added bulky side chains that in our simulations pointed toward the void occupied by heme in the holoprotein. The triple mutant apoMb3 (H64F, P88A, L89W) has Trp89 on the F helix pointing toward the empty heme pocket. The quadruple mutant apoMb4 (H64F, V68W, P88A, L89W) has Trp mutations in both helices E and F. Finally, the quintuple mutant adds a Met107 mutation on the G helix. All proteins were stable for the 20 ns of each simulation, as evident from the radius of gyration and the root-mean-square deviation (RMSD) values (Figure S1A,B, Supporting Information). Root-mean-square fluctuation (RMSF) values revealed the dynamic nature of the side chain motions in the E−F, G−H, B−C, and F−G loop regions of the mutants (Figure S2, Supporting Information). The mutants with the largest number of bulky side chains added generally lead to smaller fluctuations in the B−C and F−G loop regions.

Voids were detected near the C_α atoms of the residues lining the heme group pocket, and their volumes were estimated (Figure S3, Supporting Information). While voids remained dynamic, addition of the bulky point mutations led to reduction of the void volume, making the heme group packed tighter. Of the total heme void volume (638 Å^3), apoMb2 filled 5%, apoMb3 10%, apoMb4 16%, and apoMb5 21%. On the basis of computation, we expected that even apoMb5 could retain some heme-binding ability. This was confirmed by protein

dx.doi.org/10.1021/jp412459z | J. Phys. Chem. B 2014, 118, 6511−6518

expression: even apoMb4 and apoMb5 still bind some heme and have to be purified to remove heme, as described in the Methods section.

A network analysis looking at correlated motion of side chains (shown for apoMbWT and apoMb5 in Figure 2) revealed increased size of the detected side chain community located in the N-terminus of the proteins (A- and H-helices) and in the community connecting the helices C, D, and E surrounding the heme cavity (Figure 2, blue and red ovals). The increased size of the communities reflects stabilization of the CDEF core of the variants. Additionally, the stabilization of the F-helix of apoMb5 was improved, as evident from the increased thickness of the edges. This agrees with the reduction in the RMSF values for F−G and B−C loops (see the Supporting Information).

Thermodynamics: Filling up the Heme Cavity Increases Stability and Two-State Character. To monitor the change in stability upon mutation, thermal denaturation was measured from ∼20 to ∼90 °C for each mutant at several (0−1.5 M) GuHCl concentrations, and a GuHCl melt was measured from 0 to ca. 4.5 M GuHCl at room temperature. Denaturation was detected by both CD and tryptophan fluorescence spectroscopy. Mutant apoMb2 has tryptophan probes only in the A helix (W7 and W14), whereas mutants 3−5 have one or two additional tryptophans probing the heme pocket environment.

CD spectra of all apoMb mutants without denaturant were characteristic of a helix bundle (Figure 3A). The apoMb2 ellipticity at room temperature was comparable to that of apoMbWT (MRE ∼ −16 000 deg cm^2 dmol^{-1} [14]). Mutants 3−5 had about 20% more helix content, suggesting that L89W plays a significant role in increasing helical content, whereas

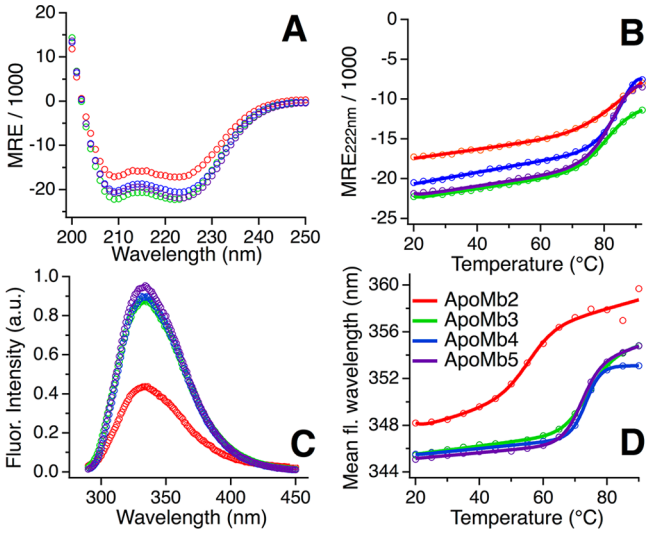


Figure 3. Thermodynamic behavior of the apoMb mutants. (A) CD spectra in 10 mM sodium acetate buffer at 22 °C and pH 6. MRE = mean residue ellipticity in deg cm^2 dmol^{-1}. (B) CD-detected thermal denaturation measurements were performed under the same buffer conditions. (C) Emission spectra were taken at 22 °C in 10 mM sodium acetate buffer at pH 6. (D) Fluorescence-detected thermal denaturation under the same buffer conditions. The mean fluorescence wavelength is shown. The solid curves in parts B and D are individual fits to an apparent two-state model, with parameters shown in the Supporting Information (Tables S1 and S2).

successive addition of V68W and I107M leads to only small additional increases in propensity.

Because apoMb folds in at least two steps, it was important to detect changes by two different probes.[34] Upon thermal denaturation detected by circular dischroism at 222 nm (Figure 3B), all mutants unfolded at about 20 °C higher temperature than apoMbWT (65.5 °C in ref 14 and reproduced here). When detected by mean fluorescence wavelength shift (Figure 3D), apoMb2 was comparable to the wild type, but apoMb3−5 melted at a 20 °C higher temperature similar to that detected by CD. Thus, apoMb2 is clearly at least a three-state folder, like the wild type, whereas mutants 3−5 are closer to two-state folders by the CD/mean fluorescence criterion.

We confirmed this by a global thermodynamic fit (see the Methods section) to obtain the most consistent melting temperatures. For each mutant, all CD-detected melts were fitted simultaneously, and all fluorescence-detected melts were fitted simultaneously, as shown in Figure 4 for our most stable mutant apoMb4 and in the Supporting Information for the other mutants (Figures S4−S6). A perfect fit could be obtained by overparameterizing the model, but we preferred the more robust fit obtained by the very simple model discussed in the Methods section, holding several parameters fixed (see the Supporting Information). The melting temperatures determined by CD at 222 nm and mean fluorescence wavelength shift are shown in Table 1 and confirm the fits to the 0 M GuHCl data in Figure 3: By CD, all mutants have a similar stability, much higher than that of the wild type; by fluorescence, apoMb2 resembles the wild type, whereas 3−5 have a much higher melting temperature very close to the CD melting temperature. Thus, when tryptophans are added in the CDEF core, protein stability probed by fluorescence approaches the average stability of secondary structure, and the protein moves from three-state behavior (apoMbWT and apoMb2) toward two-state behavior (apoMb3, apoMb4, and apoMb5). Indeed, the CD/fluorescence global melting temperatures of apoMb4 are both remarkably close to one another and actually higher than the reported value for heme-bound myoglobin (T_m ∼ 80 °C).[35] However, addition of GuHCl restores clear three-state behavior (Figure 4A).

Kinetics: Filling up the Heme Cavity Decreases the Refolding Time. We performed stopped flow measurements on apoMbWT and the four mutants. Unfolded protein solutions in 2 M GuHCl were mixed with 10 mM sodium acetate, pH 6, in order to promote refolding, which is represented by the change in fluorescence intensity (Figure 5), from tryptophans in helix A (all proteins) and also from tryptophans in the CDEF core (apoMb3, apoMb4, and apoMb5). Previous studies showed that the folding of apoMbWT initiates with the ABGH helix bundle in only a few microseconds[6,7] but does not reach its full native conformation until a relatively longer period of time (∼0.5 s by CD).[5] Due to the instrument dead time (∼20 ms), the microsecond phase was not resolved. The refolding of apoMbWT was observed in 71 ms when detected by fluorescence intensity, attributed to the folding of the heme-binding CDEF core. This rate is somewhat faster than CD-detected refolding,[5] yet another indication of wild type three-state behavior. The refolding phase of double mutant apoMb2 was one-half that of apoMbWT, and that of apoMb4 and apoMb5 was at least 4 times as fast, although the rate of apoMb5 could not be resolved due to instrument response.

dx.doi.org/10.1021/jp412459z | J. Phys. Chem. B 2014, 118, 6511−6518

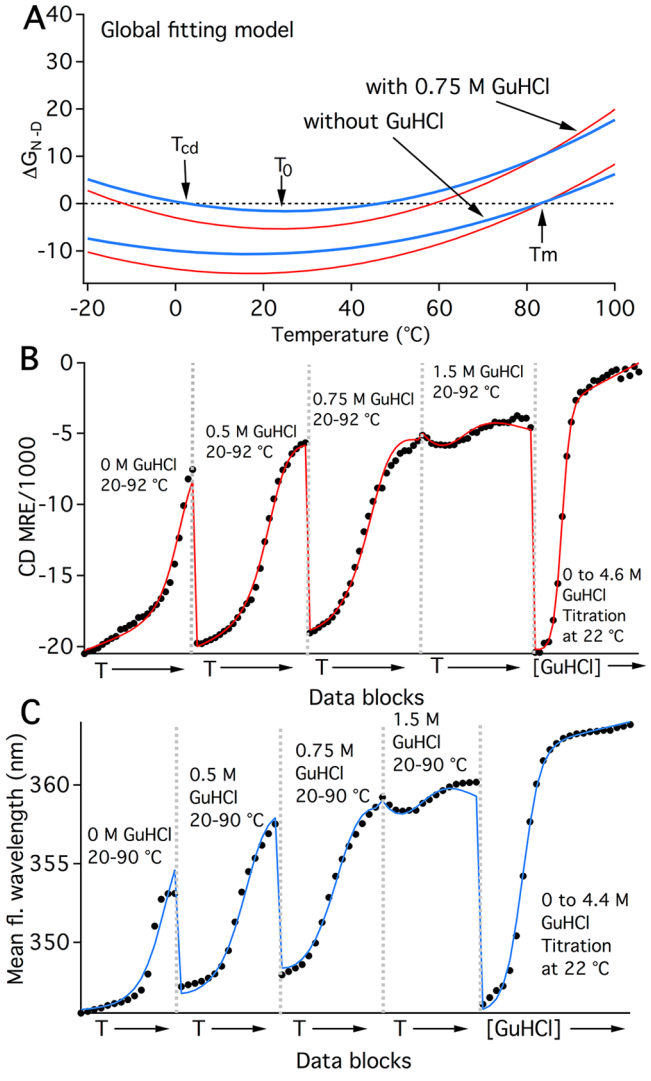


Figure 4. Global fit of the CD (red) and fluorescence (blue) of apoMb4 (experimental data in black). (A) Free energies of the global model in the Methods section as a function of temperature in 0 and 0.75 M GuHCl. The cold denaturation temperature T_{cd}, the temperature of lowest free energy T_0, and the heat denaturation melting temperature T_m are illustrated. (B) Five CD data blocks (4 thermal titrations at increasing GuHCl concentration, 1 GuHCl titration at room temperature) and global fit. (C) Five fluorescence data blocks and fit.

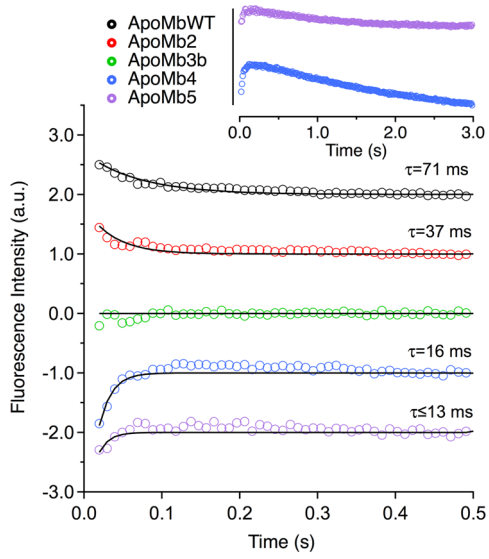


Figure 5. Short time and long time (inset) kinetics of apoMbWT and mutants 2–5. The short-time fit is a single exponential function. Traces are offset from 0 for clarity, a.u. = arbitrary units. The long traces in the inset are fitted to $\tau \approx 1.2$ s (purple) and 3 s (blue).

ApoMb refolding speeds up as the CDEF core is stabilized and the molecule approaches two-state thermodynamic behavior.

The amplitude of the refolding phase switched from decreasing to increasing when more tryptophans were added as probes in the CDEF region. This is due to the opposite behavior of the tryptophans in the A helix and the new tryptophans added in the CDEF core: the former are quenched upon refolding, whereas the latter show increased fluorescence intensity upon refolding. In apoMb3, the two effects cancel and no net phase is observed. As discussed previously for ubiquitin, apomyoglobin, and phosphoglycerate kinase, tryptophan residues that are not well packed (such as W89 and W68 are likely to be) show hyperfluorescence in the native state.[36] The transition from a decreasing phase to an increasing phase can be accounted for by hyperfluorescence from the partially confined tryptophan side chains in the heme pocket: 0 in apoMbWT and apoMb2, 1 in apoMb3, and 2 in apoMb4 and apoMb5.

Finally, we note that tryptophan undergoes bleaching during the latter part of the 3 s stopped flow measurement. We subtracted a bleaching baseline obtained by mixing protein in 2 M GuHCl buffer with aqueous 2 M GuHCl buffer to suppress the refolding reaction (Figure S7 in the Supporting Information). This subtraction yielded flat kinetic traces from 0 to 3 s for apoMbWT, apoMb2, and apoMb3, but a slow phase was left over for apoMb4 and apoMb5 (Figure 5, inset). This may represent genuine slow structural relaxation of the tryptophan residues in the heme pocket or merely indicate that bleaching proceeds at different rates for tryptophans in the heme pocket in the presence (control) or absence (dilution into 0 M GuHCl buffer) of denaturant.

DISCUSSION

Apomyoglobin is a prototypical three-state (or more) folder,[12] whose wild type has a large separation of folding time scales for the ABGH core and CDEF core.[5,6] It is a clear example of function frustrating folding. The contact order of the CDEF helices is smaller than that of the ABGH helices, their fold topology is similar, and yet they assemble more slowly. Hence, a model that assumes roughly equally good tertiary contact formation must fail in this case, predicting faster folding of the CDEF core.[4]

This begs the question whether improved packing of the empty heme pocket with large side chains can stabilize the CDEF core, speed up its folding, and bring apomyoglobin closer to two-state folding. The void volume left by the heme group (638 Å^3) is too large to be completely filled by the four side chains we enlarged; they fill at most 21% of it. Nonetheless, we can answer the question in the affirmative. The most stable of our mutants, the fourth-generation construct apoMb4, has identical melting temperatures when measured by circular dichroism and fluorescence in 0 M GuHCl, indicating more two-state-like folding. Moreover, its unique measured melting temperature is greater than that of

dx.doi.org/10.1021/jp412459z | J. Phys. Chem. B 2014, 118, 6511–6518

myoglobin. Its folding kinetics has sped up considerably over the wild type, although not yet into the microsecond range. Finally, MD simulations show that the motion of side chains surrounding the heme pocket has become more correlated, indicating rudimentary packing. Additional probes may yet reveal that apoMb4 still differs significantly from two-state folding even in 0 M GuHCl. The only indication we currently have is a difference in the free energy profiles in Figure 4A away from the melting temperature.

For other proteins, three-state folding is less robust than in apomyoglobin. In RNaseH, a single isoleucine to aspartate mutation was sufficient to restore two-state folding.[13] The mechanism for restoration was quite different there: the mutation destabilizes the intermediate, so only the unfolded state is populated and converts directly to the native state. Here, a subdomain of a protein is stabilized until it can fold in synchrony with another already stable subdomain.

On the basis of these observations, we predict that myoglobin-like proteins without a prosthetic group could turn out to be very stable fast folders. One such protein is the N-terminal domain of RsbR, a regulator of stress response genes in *B. subtilis*[29] whose structure is shown in Figure 6. This protein has the same topology as myoglobin, but helix B is directly connected to helix E, with the small heme-binding helices C and D gone. We expressed this protein and determined its T_m by thermal denaturation, which reaches a denatured baseline in 4 M GuHCl but not in 0 M GuHCl below 90 °C (Figure 6). On the basis of the two-state model described in the Methods section, and using the 4 M GuHCl data as a denatured baseline, we estimate the N-RsbR melting temperature at 87 ± 2 °C, even higher than myoglobin or our fourth generation construct apoMb4. It will be interesting to explore the folding dynamics of such hemeless myoglobins in the future, to see if their folding can approach the "speed limit", measured by downhill folding[37] and single molecule[38] experiments to lie at about 1 μs for small proteins.

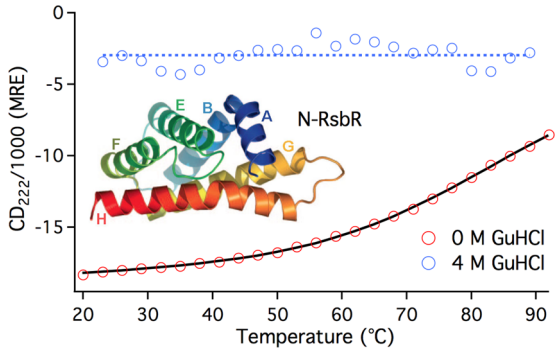


Figure 6. Thermal denaturation of the N-terminal domain of protein RsbR, a hemeless globin. The blue curve is the denatured state CD baseline and the black curve is an apparent two-state fit to the CD-detected thermal denaturation trace, using the blue denatured baseline as the upper limit.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

Description of additional measurements with text and figures. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Authors

*Phone: 001-217-333-1624. E-mail: mgruebel@illinois.edu.
*E-mail: pogorelo@illinois.edu.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

J.S.G. thanks Dr. Maxim B. Prigozhin for his guidance and mentorship. T.V.P. would like to thank Dr. Yekaterina (Katka) Golubeva and Michael Hallock for stimulating discussions. Financial support was provided by the National Institutes of Health grant R02 GM093318-4. All simulations were submitted using the Triton cluster at the SCS Computer Center.

■ REFERENCES

(1) Kendrew, J. C.; Bodo, G.; Dintzis, H. M.; Parrish, R. G.; Wyckoff, H.; Phillips, D. C. A Three-Dimensional Model of the Myoglobin Molecule Obtained by X-Ray Analysis. *Nature* **1958**, *181*, 662−666.
(2) Kirby, E. P.; Steiner, R. F. The Tryptophan Microenvironments in Apomyoglobin. *J. Biol. Chem.* **1970**, *245*, 6300−6306.
(3) Cocco, M. J.; Lecomte, J. T. J. Characterization of Hydrophobic Cores in Apomyoglobin: A Proton NMR Spectroscopy Study. *Biochemistry* **1990**, *29*, 11067−11072.
(4) Bashford, D.; Cohen, F. E.; Karplus, M.; Kuntz, I. D.; Weaver, D. L. Diffusion-Collision Model for the Folding Kinetics of Myoglobin. *Proteins* **1988**, *4*, 211−227.
(5) Jennings, P. A.; Wright, P. E. Formation of a Molten Globule Intermediate Early in the Kinetic Folding Pathway of Apomyoglobin. *Science* **1993**, *262*, 892−896.
(6) Ballew, R. M.; Sabelko, J.; Gruebele, M. Direct Observation of Fast Protein Folding: The Initial Collapse of Apomyoglobin. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 5759−5764.
(7) Ballew, R. M.; Sabelko, J.; Gruebele, M. Observation of Distinct Nanosecond and Microsecond Protein Folding Events. *Nat. Struct. Biol.* **1996**, *3*, 923−926.
(8) Rischel, C.; Thyberg, P.; Rigler, R.; Poulsen, F. M. Time-Resolved Fluorescence Studies of the Molten Globule State of Apomyoglobin. *J. Mol. Biol.* **1996**, *257*, 877−885.
(9) Bieri, O.; Wirz, J.; Hellrung, B.; Schutkowski, M.; Drewello, M.; Kiefhaber, T. The Speed Limit of Protein Folding Measure by Triplet-Triplet Energy Transfer. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 9597−9601.
(10) Gruebele, M. Protein Folding: The Free Energy Surface. *Curr. Opin. Struct. Biol.* **2002**, *12*, 161−168.
(11) Gruebele, M. Downhill Protein Folding: Evolution Meets Physics. *C. R. Biol.* **2005**, *328*, 701−712.
(12) Barrick, D.; Baldwin, R. Three-State Analysis of Sperm Whale Apomyoglobin Folding. *Biochemistry* **1993**, *32*, 3790−3796.
(13) Spudich, G. M.; Miller, E. J.; Marqusee, S. Destabilization of the Escherichia Coli Rnase H Kinetic Intermediate: Switching between a Two-State and Three-State Folding Mechanism. *J. Mol. Biol.* **2004**, *335*, 609−618.
(14) Sabelko, J.; Ervin, J.; Gruebele, M. The Cold Denatured Ensemble of Apomyoglobin: Implications for the Early Steps of Folding. *J. Phys. Chem. B* **1998**, *102*, 1806−1819.
(15) Gilmanshin, R.; Dyer, R. B.; Callender, R. H. Structural Heterogeneity of the Various Forms of Apomyoglobin: Implications for Protein Folding. *Protein Sci.* **1997**, *6*, 2134−2142.
(16) Yi, J.; Orville, A. M.; Skinner, J. M.; Skinner, M. J.; Richter-Addo, G. B. Synchrotron X-Ray-Induced Photoreduction of Ferric Myoglobin Nitrite Crystals Gives the Ferrous Derivative with Retention of the O-Bonded Nitrite Ligand. *Biochemistry* **2010**, *49*, 5969−5971.
(17) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual Molecular Dynamics. *J. Mol. Graphics* **1996**, *14* (33−38), 27−38.

(18) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K. Scalable Molecular Dynamics with NAMD. *J. Comput. Chem.* **2005**, *26*, 1781−1802.
(19) Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, I.; Mackerell, A. D., Jr. Charmm General Force Field: A Force Field for Drug-Like Molecules Compatible with the Charmm All-Atom Additive Biological Force Fields. *J. Comput. Chem.* **2010**, *31*, 671−690.
(20) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926−935.
(21) Mackerell, A. D., Jr.; Feig, M.; Brooks, C. L. Extending the Treatment of Backbone Energetics in Protein Force Fields: Limitations of Gas-Phase Quantum Mechanics in Reproducing Protein Conformational Distributions in Molecular Dynamics Simulations. *J. Comput. Chem.* **2004**, *25*, 1400−1415.
(22) Martyna, G. J.; Tobias, D. J.; Klein, M. L. Constant Pressure Molecular Dynamics Algorithms. *J. Chem. Phys.* **1994**, *101*, 4177−4189.
(23) Feller, S. E.; Zhang, Y.; Pastor, R. W.; Brooks, B. R. Constant Pressure Molecular Dynamics Simulation: The Langevin Piston Method. *J. Chem. Phys.* **1995**, *103*, 4613−4621.
(24) Darden, T.; York, D.; Pedersen, L. Particle Mesh Ewald: An N Log(N) Method for Ewald Sums in Large Systems. *J. Chem. Phys.* **1993**, *98*, 10089−10092.
(25) Till, M. S.; Ullmann, G. M. McVol - a Program for Calculating Protein Volumes and Identifying Cavities by a Monte Carlo Algorithm. *J. Mol. Model.* **2010**, *16*, 419−429.
(26) Sethi, A.; Eargle, J.; Black, A. A.; Luthey-Schulten, Z. Dynamical Networks in Trna:Protein Complexes. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 6620−6625.
(27) Girvan, M.; Newman, M. E. Community Structure in Social and Biological Networks. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 7821−7826.
(28) Teale, F. W. Cleavage of the Haem-Protein Link by Acid Methylethylketone. *Biochim. Biophys. Acta* **1959**, *35*, 543.
(29) Murray, J. W.; Delumeau, O.; Lewis, R. J. Structure of a Nonheme Globin in Environmental Stress Signaling. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 17320−17325.
(30) Nishii, I.; Kataoka, M.; Tokunaga, F.; Goto, Y. Cold Denaturation of the Molten Globule States of Apomyoglobin and a Profile for Protein Folding. *Biochemistry* **1994**, *33*, 4903−4909.
(31) Tanford, C. Protein Denaturation. *Adv. Protein Chem.* **1970**, *24*, 1−95.
(32) Picotti, P.; Marabotti, A.; Negro, A.; Musi, V.; Spolaore, B.; Zambonin, M.; Fontana, A. Modulation of the Structural Integrity of Helix F in Apomyoglobin by Single Amino Acid Replacements. *Protein Sci.* **2004**, *13*, 1572−1585.
(33) Garcia, C.; Nishimura, C.; Cavagnero, S.; Dyson, H. J.; Wright, P. E. Changes in the Apomyoglobin Folding Pathway Caused by Mutation of the Distal Histidine Residue. *Biochemistry* **2000**, *39*, 11227−11237.
(34) Chiba, K.; Ikai, A.; Kawamurakonishi, Y.; Kihara, H. Kinetic Study on Myoglobin Refolding Monitored by Five Optical Probe Stopped-Flow Methods. *Proteins* **1994**, *19*, 110−119.
(35) Wan, L.; Twitchett, M. B.; Eltis, L. D.; Mauk, A. G.; Smith, M. In Vitro Evolution of Horse Heart Myoglobin to Increase Peroxidase Activity. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 12825−12831.
(36) Ervin, J.; Larios, E.; Osvath, S.; Schulten, K.; Gruebele, M. What Causes Hyperfluorescence: Folding Intermediates or Conformationally Flexible Native States? *Biophys. J.* **2002**, *83*, 473−483.
(37) Yang, W. Y.; Gruebele, M. Folding at the Speed Limit. *Nature* **2003**, *423*, 193−197.
(38) Chung, H. S.; McHale, K.; Louis, J. M.; Eaton, W. A. Single-Molecule Fluorescence Experiments Determine Protein Folding Transition Path Times. *Science* **2012**, *335*, 981−984.