Functional Cavities in Proteins: A General Method for Proximal Ligand Substitution in Myoglobin

Gia D. DePillis,† Sean M. Decatur,† Doug Barrick,‡ and Steven G. Boxer*,†

Department of Chemistry, Stanford University Stanford, California 94305-5080

Institute for Molecular Biology University of Oregon, Eugene, Oregon 97403

Received December 9, 1993

The amino acid ligands to the heme iron provide an important distinction between different classes of heme proteins. For example, the peroxidases and the oxygen carrier proteins myoglobin (Mb) and hemoglobin bear a proximal His, electron-transfer proteins such as cytochrome *c* have His, Met, and Lys ligands, and oxidases such as cytochrome P450 and catalase contain Cys and Tyr ligands, respectively. In contrast to chemical model systems, where a large range of ligands has been used to modulate heme function,[1] replacement of the proximal ligand in Mb has been limited in scope to the natural amino acids Cys[2] and Tyr.[2,3] In order to understand the vast range of reactivities among heme proteins, it would be ideal to possess the flexibility to vary the proximal ligand, while retaining the structure, recognition, specificity, and ease of manipulation provided by the protein matrix. We report a simple method to achieve this goal for myoglobin, with broader implications for protein engineering in general.[4]

When the proximal His residue 93 in sperm whale Mb is changed to Gly by site-directed mutagenesis and the protein is expressed in *Escherichia coli*[5] using a medium containing imidazole (Im), a protein is isolated which has a molecular weight and electronic absorption spectrum similar to those of wild-type (WT) Mb.[6] The high-resolution X-ray crystal structure of this protein, H93G(Im), illustrated schematically in Figure 1, shows that Im occupies the cavity created by conversion of His93 to Gly and is coordinated to the heme iron.[7] As shown in Figure 2, replacement of proximal His93 by Im leads to significant changes in the rebinding kinetics for CO, bound on the distal side, following flash photolysis.[8] These results demonstrate that the small changes in the Fe–N bond distance and the orientation of the Im plane relative to the heme plane seen in the X-ray structure[7] lead to significant changes in function.[9]

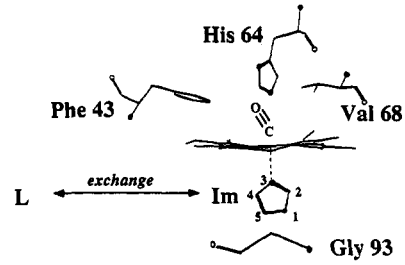

Figure 1. Schematic diagram of heme pocket residues in carbonmonoxymyoglobin mutant H93G(Im),[6] illustrating the replacement of proximal His93 with free imidazole and further exchange with other small ligands L.

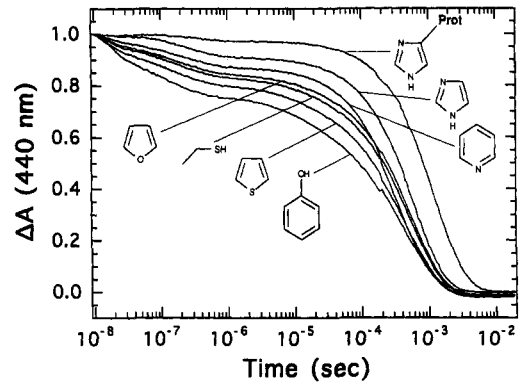

Figure 2. Ligand rebinding kinetics following flash photolysis of CO from H93G(L) (22 °C; 1 atm CO; aqueous buffer; see ref 8 for experimental details). The rebinding curves have been normalized to the same change in absorption at the earliest time (about 10–12 ns following the laser flash[17]).

We have discovered that a wide range of small organic ligands can replace Im in the proximal binding site simply by addition of an excess of the ligand to a solution of ferric H93G(Im). For example, relatively conservative substitutions, such as pyridine and methyl-substituted imidazoles, lead to relatively minor absorption spectral changes (Table 1) while producing readily measured differences in the ligand binding kinetics (Figure 2) and NMR spectra (Figure 3).[10] Ligand exchange is completely reversible, as addition of Im to the other ligand-substituted proteins yields back the spectrum of the Im-ligated protein (data not shown). The hyperfine-shifted ^{1}H NMR spectra of the paramagnetic metcyano complexes of heme proteins and models are exquisitely sensitive to proximal ligand orientation.[11] The pattern of the hyperfine-shifted heme methyl resonances of H93G(Im)-CN is different from that of WT MbCN (Figure 3A and B), consistent with the different orientation of the Im and His ligands observed in their respective crystal structures.[12] Addition of excess deuterated Im (Im-d_3) results in the loss of a peak at

† Stanford University.

‡ University of Oregon.

(1) Ji, L.-N.; Liu, M.; Hsieh, A.-K.; Hor, T. S. A. *J. Mol. Catal.* **1991**, *70*, 247–257. James, B. R. In *The Porphyrins*; Dolphin, D., Ed.; Academic Press: New York, 1978; Vol. V, pp 205–302.

(2) Adachi, S.; Nagano, S.; Ishimori, K.; Watanabe, Y.; Morishima, I.; Egawa, T.; Kitagawa, T.; Makino, R. *Biochemistry* **1993**, *32*, 241–252. Adachi, S.; Nagano, S.; Watanabe, Y.; Ishimori, K.; Morishima, I. *Biochem. Biophys. Res. Commun.* **1991**, *180*, 138–144.

(3) Egeberg, K. D.; Springer, B. A.; Martinis, S. A.; Sligar, S. G.; Morikis, D.; Champion, P. M. *Biochemistry* **1990**, *29*, 9783–9791.

(4) Alternative methods include protein semisynthesis, which has been applied to cytochrome *c* (Wuttke, D. S.; Gray, H. B.; Fisher, S. L.; Imperiali, B. *J. Am. Chem. Soc.* **1993**, *115*, 8455–8456. Wallace, C. J. A.; Clark-Lewis, I. *J. Biol. Chem.* **1992**, *267*, 3852–3861), and chemical aminoacylation of suppressor tRNA complementary to a nonsense codon introduced at the desired site, followed by *in vitro* protein expression (Noren, C. J.; Anthony-Cahill, S. J.; Griffith, M. C.; Schultz, P. G. *Science* **1989**, *244*, 182–188). The strategy of substituting amino acid side chains with exogenous molecules has been applied to, e.g., aspartate amino transferase (Toney, M. D.; Kirsch, J. F. *Science* **1989**, *243*, 1485–1488); subtilisin (Carter, P.; Abrahmsen, L.; Wells, J. A. *Biochemistry* **1991**, *30*, 6142–6148); trypsin (Perona, J. J.; Hedstrom, L.; Wagner, R. L.; Rutter, W. J.; Craik, C. S.; Fletterick, R. J. *Biochemistry* **1994**, *33*, 3252–3259); rhodopsin (Zhukovsjy, E. A.; Robinson, P. R.; Oprian, D. D. *Science* **1991**, *251*, 558–560); and azurin (den Blaauwen, T.; Canters, G. W. *J. Am. Chem. Soc.* **1993**, *115*, 1121–1129).

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

(6) Barrick, D. *Biochemistry* **1994**, *33*, 6546–6554.

(7) Relevant differences in the structure of H93G(Im) compared to WT reported in ref 6 include (1) Fe–N bond length 0.3 Å shorter; (2) imidazole plane rotated 36° relative to the His93 plane; (3) nonliganded Im N atom 0.2 Å closer to the hydroxyl group of Ser92, adopting a favorable hydrogen-bonding geometry; (4) 0.6 Å increase in the distance between Fe and the distal H_2O; and (5) Fe 0.2 Å closer to the plane of the heme.

(8) For experimental methods, see, e.g.: Balasubramanian, S.; Lambright, D. G.; Boxer, S. G. *Biochemistry* **1993**, *32*, 2202–2212.

(9) Measurements of k_{off} for CO by the classical method of NO substitution are complicated by interesting secondary charges associated with the NO complex. These changes are under investigation and will be described in detail elsewhere.

(10) In all cases, CO is found to bind reversibly to the reduced form of H93G(L), exactly as for WT and H93G(Im), and yields similar absorption spectra irrespective of L. O_2 rebinding has not yet been measured, as the complexes are very sensitive to autooxidation.

(11) (a) Traylor, T. G.; Berzinis, A. P. *J. Am. Chem. Soc.* **1980**, *102*, 2844–2846. (b) Yamamoto, Y.; Nanai, N.; Chujo, R.; Suzuki, T. *FEBS Lett.* **1990**, *264*, 113–116.

0002-7863/94/1516-6981$04.50/0

Table 1. Electronic Absorption Spectra of Proteins

protein	soret (nm) (bandwidth)[a]	relative ϵ_{Soret}[b]	visible (nm)
WT Mb[c]	408 (N)		502, 632
H93G(imidazole)	409 (N)	1.00	503, 629
H93G(pyridine)	408 (N)	0.92	497, 630
catalase[d]	405 (B)		620
H93Y Mb[e]	402 (B)		480, 598
H93G(phenol)	403 (B)	0.68	495, 605
$P450_{cam}$[e]	391 (VB)		510, 600
H93C human Mb[f]	391 (VB)		509, 629
H93G(ethanethiol)	393 (VB)	0.54	505, 626
H93G(thiophene)	393 (VB)	0.59	505, 622
H93G(furan)	408 (N)	0.89	497, 629

[a] Width at half-height: N, narrow (~30 nm); B, broad (>40 nm); VB, very broad (>50 nm). [b] Measured relative to H93G(Im); all H93G(L) at same concentration. [c] Both human[2] and sperm whale.[3] [d] Osbahr, A. J.; Eichhorn, G. L. *J. Biol. Chem.* **1962**, *237*, 1820–1824. [e] High-spin ferric, camphor-bound.[15] [f] Reference 3.

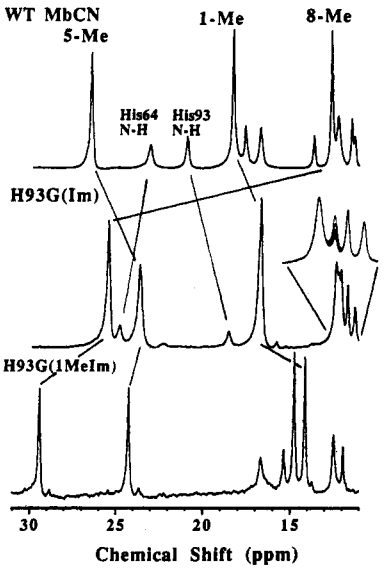

Figure 3. Comparison of ^1H NMR spectra of WT MbCN and H93G-(Im)CN, taken in 90% H_2O buffer, pH 9.0, 30 °C, and H93G(1-MeIm)-CN, taken in D_2O buffer, pH 7.0 (500 MHz). Assignments for WT are from the work of La Mar and co-workers; assignments for H93G(L) are based on 1D NOE, 2D NOESY, and COSY spectra, and isotopic labeling of the heme and the imidazole as in ref 11. The inset above the H93G-(Im) spectrum shows the disappearance over a period of 50 min of an Im proton upon addition of an excess of Im-d_3 to H93G(Im) in D_2O buffer, pH 7.0. The small extra peak at 12.0 ppm in the parent spectrum taken in H_2O buffer, pH 9.0, is an exchangeable amide backbone proton.

12.1 ppm (inset, Figure 3B), unambiguously demonstrating that the proximal Im ligand is reversibly exchanged with exogenous ligands.[13] Substitution with other ligands yields unique and characteristic spectra (e.g., 1-methylimidazole, Figure 3C). We have also obtained the ^1H NMR spectra of the diamagnetic Fe^{II}-CO complexes of H93G(L), because the chemical shifts of the Val68 protons of MbCO are sensitive to structural changes on the distal side.[14] The chemical shifts of the Val68$C_\gamma H_3$ protons of all H93G(L) are essentially the same as WT (data not shown). Thus, the diamagnetic and paramagnetic NMR data together indicate that ligand exchange perturbs the proximal but not the distal side of the heme pocket. Finally, addition of the same ligands to solutions of the WT protein produces none of the changes reported above for H93G.[15] These findings support the contention that exogenous ligands replace Im specifically in the proximal cavity of H93G.

This ligand exchange strategy proves to be remarkably general and can be used to make more radical changes in the proximal residue. We have initially focused on replacing Im with phenol and ethanethiol, as they are the analogs of the amino acids Tyr and Cys, which have been introduced into Mb by conventional mutagenesis of His93.[2,3] Replacement of Im in H93G(Im) by these ligands leads to very large changes in the absorption spectrum. As shown in Table 1, the spectrum of H93G(phenol) is quite similar to those of H93Y and catalase, while the spectrum of H93G(ethanethiol) is very similar to that of H93C. The CO rebinding kinetics of H93G(ethanethiol) and H93G(phenol) are quite different from those of H93G(Im) (Figure 2). The CO recombination kinetics vary widely with different exogenous ligands,[10] thus the possibility that the distal His64 is serving as the fifth ligand, with CO bound on the proximal side, is ruled out.[16] As shown in Table 1 and Figure 2, we have also replaced Im with other ligands, such as furan and thiophene, thereby introducing entirely unnatural functionality into the cavity created by removal of His93 and producing large effects on ligand binding kinetics.

Replacement of the proximal His with exogenous ligands generates a rich variety of novel proteins. With this exchange method it is possible to systematically alter the size, chemical, and isotopic properties of the proximal ligand, using methodology which is much simpler than the chemical or biosynthetic introduction of unnatural amino acids.[4] In combination with other distal- and proximal-side changes produced by conventional site-directed mutagenesis, it should be possible to discover the precise mechanisms by which the protein modulates heme function. Extensions of this ligand exchange strategy are in progress to examine the role of the proximal ligand in determining cooperativity in Hb and in modulating the binding of chlorophylls in photosynthetic proteins and to probe electron transfer pathways in proteins.

Acknowledgment. We thank Professor Steven G. Sligar for his gift of the sperm whale Mb expression system. G.D.P. is supported in part by an NIH Postdoctoral Fellowship; S.M.D. is the recipient of an NSF Predoctoral Fellowship. This work was supported in part by a grant from the National Institutes of Health.

(12) NMR assignments of hyperfine shifted resonances followed the methods described in: Emerson, S. D.; La Mar, G. N. *Biochemistry* **1990**, *29*, 1545–1556. Yamamoto et al.[11b] have correlated the pattern of heme methyl shifts with His imidazole orientation. The assignments in Figure 2B are consistent with the rotation of the imidazole plane observed in the crystal structure.[7]

(13) Exchange of Im with Im-d_3 in the metcyano complex occurs with a half-life of about 9 min (Decatur et al., to be published).

(14) The chemical shift of the Val68$C_\gamma H_3$ protons is sensitive to the geometry of the distal pocket. For example, modification of the distal His of WT SW Mb with the tetrazole moiety leads to a shift in the Val68$C_\gamma H_3$ resonance indicative of conformational changes at His64 (Adachi, S.; Morishima, I. *Biochemistry* **1992**, *31*, 8613–8618). Even larger shifts are observed in the NMR spectrum of the ferrous-Im complex of WT Mb (Decatur et al., unpublished observations). The insensitivity of this chemical shift in H93G-(L) makes it very unlikely that added ligands bind on the distal side or that His64 serves as the "proximal" ligand.

(15) Im and simple thiols have been shown to bind weakly to the distal site of SW WT Mb (Sono, M.; Andersson, L. A.; Dawson, J. H. *J. Biol. Chem.* **1982**, *257*, 8308–8320) but only at a much higher concentration of ligand than employed here.

(16) The Soret band maximum in H93C MbCO is found at 420 nm, while the maximum of the corresponding complex in cysteine-ligated P450 is 450 nm.[2] This led the authors to conclude that distal H64 replaces the proximal Cys ligand as the fifth ligand, with CO bound in the proximal pocket. However, their results are consistent with model studies of CO complexes in which fifth ligands, including Im, ethers, thiols, and thioethers, all yield 420 nm Soret maxima. The Soret band shifts to 450 nm only when a thiol ligand is deprotonated (Collman, J. P.; Sorrell, T. N. *J. Am. Chem. Soc.* **1975**, *97*, 4133–4134).

(17) We assume that as for WT Mb, no geminate recombination occurs on the subnanosecond time scale for CO (see, e.g.: Cornelius, P. A.; Steele, A. W.; Chernoff, D. A.; Hochstrasser, R. M. *Proc. Natl. Acad. Sci. U.S.A.* **1981**, *78*, 7526–7529). This assumption is being tested by picosecond time-resolved measurements.