J|A|C|S
ARTICLES
Published on Web 00/00/0000

An Electrostatic Model for the Frequency Shifts in the Carbonmonoxy Stretching Band of Myoglobin: Correlation of Hydrogen Bonding and the Stark Tuning Rate

Stefan Franzen*

Contribution from the Department of Chemistry, North Carolina State University, Raleigh, North Carolina 27695

Received December 7, 2001

Abstract: The effect of internal and applied external electric fields on the vibrational stretching frequency for bound CO (ν_{CO}) in myoglobin mutants was studied using density functional theory. Geometry optimization and frequency calculations were carried out for an imidazole−iron−porphine−carbonmonoxy adduct with various small molecule hydrogen-bonding groups. Over 70 vibrational frequency calculations of different model geometries and hydrogen-bonding groups were compared to derive overall trends in the C−O stretching frequency (ν_{CO}) in terms of the C−O bond length and Mulliken charge. Simple linear functions were derived to predict the Stark tuning rate using an approach analogous to the vibronic theory of activation.[1] Potential energy calculations show that the strongest interaction occurs for C−H or N−H hydrogen bonding nearly perpendicular to the Fe−C−O bond axis. The calculated frequencies are compared to the structural data available from 18 myoglobin crystal structures, supporting the hypothesis that the vast majority of hydrogen-bonding interactions with CO occur from the side, rather than the end, of the bound CO ligand. The ν_{CO} frequency shifts agree well with experimental frequency shifts for multiple bands, known as A states, and site-directed mutations in the distal pocket of myoglobin. The model calculations quantitatively explain electrostatic effects in terms of specific hydrogen-bonding interactions with bound CO in heme proteins.

Introduction

The electrostatic environment is a key determinant of both binding affinity and reactivity of proteins. A central issue in the current understanding of electrostatic effects is the distinction between specific short-range hydrogen-bonding interactions, on one hand, and long-range interactions from internal fields, on the other. Internal fields arise from a sum of smaller interactions at a distance. Given the long-range nature of Coulombic interactions, internal fields that arise from specific protein geometries may be significant. Enzymatic catalysis and molecular recognition involve highly specific electrostatic interactions that include salt bridges and hydrogen bonds. To address the relative magnitude of these effects, an experimental reporter such as carbon monoxide bound to heme in myoglobin provides a useful probe of the protein electrostatic environment.[1−22] The stretching frequency of the CO molecule, ν_{CO}, serves as a probe that can be studied by structural and spectroscopic methods to provide correlations of electrostatic environment and function.[4,19,23] The iron−carbon stretching frequency, ν_{Fe-CO}, is correlated with ν_{CO} and provides complementary information on the structure of the distal pocket (Figure 1). In wild-type

* To whom correspondence should be addressed. Phone: (919)-515-8915. Fax: (919)-515-8909. E-mail: Stefan_Franzen@ncsu.edu.

(1) Stavrov, S. S.; Decusar, I. P.; Bersuker, I. B. *New J. Chem.* **1993**, *17*, 71−76.
(2) Balasubramanian, S.; Lambright, D. G.; Boxer, S. G. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 4718−4722.
(3) Decatur, S. M.; Boxer, S. G. *Biochem. Biophys. Res. Commun.* **1995**, *212*, 159−164.
(4) Park, E. S.; Andrews, S. S.; Hu, R. B.; Boxer, S. G. *J. Phys. Chem. B* **1999**, *103*, 9813−9817.
(5) Park, E. S.; Thomas, M. R.; Boxer, S. G. *J. Am. Chem. Soc.* **2000**, *122*, 12297−12303.
(6) Varadarajan, R.; Lambright, D. G.; Boxer, S. G. *Biochemistry* **1989**, *28*, 3771−3781.
(7) Varadarajan, R.; Zewert, T. E.; Gray, H. B.; Boxer, S. G. *Science* **1989**, *243*, 69−72.
(8) Kim, S.; Deinum, G.; Gardner, M. T.; Marletta, M. A.; Babcock, G. T. *J. Am. Chem. Soc.* **1996**, *118*, 8769−8770.
(9) Lloyd, E.; Burk, D. L.; Ferrer, J. C.; Maurus, R.; Doran, J.; Carey, P. R.; Brayer, G. D.; Mauk, A. G. *Biochemistry* **1996**, *35*, 11901−11912.
(10) Kiger, L.; Stetzkowski-Marden, F.; Poyart, C.; Marden, M. *Eur. J. Biochem.* **1995**, *228*, 665−668.
(11) Anderton, C. L.; Hester, R. E.; Moore, J. N. *Biochim. Biophys. Acta* **1997**, *1338*, 107−120.
(12) Cameron, A. D.; Smerdon, S. J.; Wilkinson, A. J.; Habash, J.; Helliwell, J. R.; Li, T. S.; Olson, J. S. *Biochemistry* **1993**, *32*, 13061−13070.
(13) Ivanov, D.; Sage, J. T.; Keim, M.; Powell, J. R.; Asher, S. A.; Champion, P. M. *J. Am. Chem. Soc.* **1994**, *116*, 4139−4140.
(14) Sage, J. T.; Morikis, D.; Champion, P. M. *Biochemistry* **1991**, *30*, 1227−1237.
(15) Unno, M.; Christian, J. F.; Olson, J. S.; Sage, J. T.; Champion, P. M. *J. Am. Chem. Soc.* **1998**, *120*, 2670−2671.
(16) Zhu, L. Y.; Sage, J. T.; Champion, P. M. *Biochemistry* **1993**, *32*, 11181−11185.
(17) Christian, J. F.; Unno, M.; Sage, J. T.; Champion, P. M.; Chien, E.; Sligar, S. G. *Biochemistry* **1997**, *36*, 11198−11204.
(18) Oldfield, E.; Guo, K.; Augspurger, J. D.; Dykstra, C. E. *J. Am. Chem. Soc.* **1991**, *113*, 7537−7541.
(19) Phillips, J. G. N.; Teodoro, M. L.; Li, T.; Smith, B.; Olson, J. S. *J. Phys. Chem. B* **1999**, *103*, 8817−8829.
(20) Kushkuley, B.; Stavrov, S. *Biophys. J.* **1996**, *70*, 1214−1229.
(21) Kushkuley, B.; Stavrov, S. S. *Biophys. J.* **1997**, *72*, 899−912.
(22) Olson, J. S.; Phillips, G. N., Jr. *J. Biol. Inorg. Chem.* **1997**, *2*, 544−552.
(23) Ansari, A.; Berendzen, J.; Braunstein, D.; Cowen, B. R.; Frauenfelder, H.; Hong, M. K.; Iben, I. E. T.; Johnson, J. B.; Ormos, P.; Sauke, T. B.; Scholl, R.; Schulte, A.; Steinbach, P. J.; Vittitow, J.; Young, R. D. *Biophys. Chem.* **1987**, *26*, 337−355.

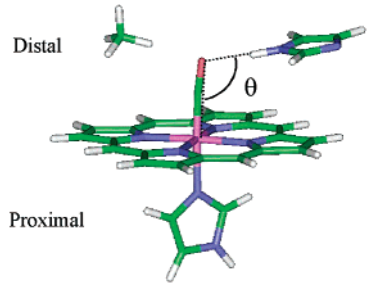

Figure 1. Depiction of the iron−imidazole−carbonmonoxy porphine adduct studied. In this figure, two hydrogen-bonding groups are shown as an example. The CH_4 and C_3H_4N (Im) hydrogen-bonding distances are 2.0 and 1.92 Å, respectively. The angle θ is defined as the Fe−O···H angle.

sperm whale myoglobin (SWMb), there are at least three observed ν_{CO} and ν_{Fe-CO} bands in the pH range from 4.5 to 7.0. The origin of these bands is still a matter of some interest because they determine the relationship of conformational substates of myoglobin with functional states of the protein.[23−29] Frequency shifts of ν_{CO} in mutants that do not contain a distal histidine[19,22,30] show that the electrostatic effect on bound CO is not exclusively due to the tautomers of histidine.[18] In addition to the effects on the distal side, there are also proximal effects on ν_{CO} that arise from changes in axial ligation trans to bound CO. Proximal effects on ν_{CO} have been studied in the H93G proximal cavity mutant and other proximal mutants.[31−33]

The environment of bound CO in myoglobin is an important test case for understanding protein electrostatic effects. A large number of myoglobin mutants have been structurally characterized by X-ray crystallography in the carbonomonoxy (MbCO) form.[34−41] Table 1 lists all atoms that are less than 4.5 Å from the oxygen of CO in each of the X-ray structures of wild-type and mutant MbCO. In each case, there are several close contacts between protein side chains and bound CO. Table 1 suggests

Table 1. Listing of Closest Atoms in Myoglobin Mutant Crystal Structures[a]

mutant	*R* (Å)	Fe–O···X	mutant	*R* (Å)	Fe–O···X
2MGK	wild type[35]		1MLF	V68A[36]	
H64-Ne	3.07	103.68	H64-Ne	2.74	102.27
V68-C	3.23	98.94	F43-C	3.86	99.77
F43-C	3.56	108.81	A68-C	3.70	104.15
1MNK	H64V,V68T (pig)[38]		1MLJ	V68F[36]	
T68-O	2.84	136.00	H64-Nδ	2.67	106.5
V64-C	4.28	141.14	F43-C	3.98	96.49
F43-C	4.38	110.12	F68-C	4.13	96.82
1MOC	H64T[35]		1MCY	H64Q,L29F[39]	
V68-C	3.03	98.98	V68-C	2.94	100.0
F43-C	3.43	113.57	F43-C	3.57	107.6
T64-C	3.63	131.96	Q64-N	3.27	100.0
2MGC	H64L[35]		F29-C	3.28	151.5
V68-C	3.21	125.49	1MLM	V68I[36]	
V68-C	4.53	101.24	H64-Ne	2.96	103.03
L64-C	3.93	158.17	I68-C	3.53	90.64
F43-C	3.65	106.32	I68-Ca	3.27	101.52
2MGA	H64G[35]		F43-C	3.51	130.12
V68-C	2.78	100.19	1MTI	F46L[40]	
V68-C	4.69	98.45	H64-Ne	2.63	130.96
F43-C	3.53	103.99	F43-C	4.15	113.75
H2O	2.94	108.00	V68-C	3.14	127.03
H2O	2.87	98.12	2SPL	L29F[41]	
1MLU	H64G,V68A[88]		H64-Ne	3.02	117.69
A68-C	3.59	100.76	F29-C	3.18	158.09
F43-C	3.80	102.7	F43-C	3.74	102.24
H2O	2.94	100.0	V68-C	3.46	96.14
1MYM	F46V[49]		1VXC	pH 5.0 wild type[37]	
V68-C	2.83	103.07	H64-Ne	3.16	105.3
V68-C	4.97	94.21	I F43-C	3.48	116.92
F43-C	3.48	108.05	V68-C	3.82	103.57
1YCA	V68T[12]		1VXF	pH 6.0 wild type[37]	
V68-C	3.22	96.24	H64-Ne	3.29	99.81
H64-Ne	2.90	95.85	F43-C	3.64	107.9
F43-C	3.52	108.36	V68-C	3.16	103.2
2MGF	H64Q[35]		1MBC	wild type[89]	
V68-C	2.88	100.97	H64-Ne	2.66	107.98
Q64-N	3.05	111.71	F43-C	3.31	115.92
F43-C	3.59	110.23	V68-C	3.32	92.76

[a] The distances and Fe−O···X angles are reported on the basis of the X-ray crystal structure data.

that the electrostatic effect of, at most, a few amino acid residues is likely to play a major role in the observed frequency shifts of ν_{Fe-CO} and ν_{CO} bands in mutants of myoglobin. In some cases, a single amino acid dominates the interaction because of its proximity to CO. For example, the available evidence indicates that the conformations of the closest residue in wild-type myoglobin, H64, play a dominant role in the hydrogen-bonding environment of CO.[42−44] The distribution of Fe−O···L angles, defined in Figure 1 as the angle, θ, is dominated by angles of 110° or less. Out of 60 amino acids listed in Table 1 that contain hydrogen atoms within a radius of 4.5 Å, only 15 have the defined angle $\theta > 110°$, and only nine have $\theta > 120°$, suggesting that there is a preferred geometry of interaction. These considerations indicate that specific hydrogen-bonding interactions are important and, moreover, that there is an orientational preference for hydrogen bonding to CO.

There are four ν_{CO} stretching frequencies observed by infrared spectroscopy in many forms of wild-type myoglobin and

(24) Rejto, P. A.; Freer, S. T. *Prog. Biophys. Mol. Biol.* **1996**, *66*, 167−196.
(25) Miller, L. M.; Patel, M.; Chance, M. R. *J. Am. Chem. Soc.* **1996**, *118*, 4511−4517.
(26) Schulze, B. G.; Grubmuller, H.; Evanseck, J. D. *J. Am. Chem. Soc.* **2000**, *122*, 8700−8711.
(27) Schott, J.; Dreybrodt, W.; Schweitzer-Stenner, R. *Biophys. J.* **2001**, *81*, 1624−1631.
(28) Sage, J. T.; Morikis, D.; Champion, P. M. *Biochemistry* **1991**, *30*, 1237−1247.
(29) Steinbach, P. J.; Ansari, A.; Berendzen, J.; Braunstein, D.; Chu, K.; Cowen, B. R.; Ehrenstein, D.; Frauenfelder, H.; Johnson, J. B.; Lamb, D. C.; Luck, S.; Mourant, J. R.; Nienhaus, G. U.; Ormos, P.; Philipp, R.; Xie, A.; Young, R. D. *Biochemistry* **1991**, *30*, 3988−4001.
(30) Balasubramanian, S.; Lambright, D. G.; Marden, M. C.; Boxer, S. G. *Biochemistry* **1993**, *32*, 2.
(31) Decatur, S.; DePillis, G. D.; Boxer, S. G. *Biochemistry* **1996**, *35*, 3925−3932.
(32) Peterson, E.; Chien, E.; Sligar, S.; Friedman, J. *Biochemistry* **1998**, *37*, 12301−12319.
(33) Abadan, Y.; Chien, E. Y. T.; Chu, K.; Eng, C. D.; Nienhaus, U.; Sligar, S. G. *Biophys. J.* **1995**, *68*, 2495−2502.
(34) Andrews, B. K.; Romo, T.; Clarage, J. B.; Pettitt, B. M.; Phillips, G. N. *Structure Folding and Design* **1998**, *6*, 587−594.
(35) Quillin, M. L.; Arduini, R. M.; Olson, J. S.; Phillips, G. N. *J. Mol. Biol.* **1993**, *234*, 140−155.
(36) Quillin, M. L.; Li, T.; Olson, J. S.; Phillips, G. N.; Dou, Y.; Ikeda-Saito, M.; Regan, R.; Carlson, M.; Gibson, Q. H.; Li, H.; Elber, R. *J. Mol. Biol.* **1995**, *245*, 416−436.
(37) Yang, F.; Phillips, G. N., Jr. *J. Mol. Biol.* **1996**, *256*, 762−774.
(38) Smerdon, S. J.; Krzywda, S.; Brzozowski, A. M.; Davies, G. J.; Wilkinson, A. J.; Brancaccio, A.; Cutruzzola, F.; Allocatelli, C. T.; Brunori, M.; Li, T.; et al. *Biochemistry* **1995**, *34*, 8715−8725.
(39) Zhao, X.; Vyas, K.; Nguyen, B. D.; Rajarathnam, K.; Mar, G. N. L.; Li, T.; Phillips, G. N., Jr.; Eich, R. F.; Olson, J. S.; Ling, J.; Bocian, D. F. *J. Biol. Chem.* **1995**, *270*, 20763−20774.
(40) Lai, H. H.; Li, T.; Lyons, D. S.; Phillips, G. N., Jr.; Olson, J. S.; Gibson, Q. H. *Proteins* **1995**, *22*, 322.
(41) Carver, T. E.; Brantley, R. E., Jr.; Singleton, E. W.; Arduini, R. M.; Quillin, M. L.; Phillips, G. N., Jr.; Olson, J. S. *J. Biol. Chem.* **1992**, *267*, 14443−14450.
(42) Caughey, W. S.; Shimada, H.; Miles, G. C.; Tucker, M. P. *Proc. Natl. Acad. Sci. U.S.A.* **1981**, *78*, 2903−2907.
(43) Li, X. Y.; Spiro, T. G. *J. Am. Chem. Soc.* **1988**, *110*, 6024−6033.
(44) Morikis, D.; Champion, P. M.; Springer, B. A.; Sligar, S. G. *Biochemistry* **1989**, *28*, 4791−4800.

BATCH: ja11a34 USER: jld69 DIV: @xyv04/data1/CLS_pj/GRP_ja/JOB_i44/DIV_ja017708d DATE: September 17, 2002

hemoglobin.[45] These bands have been designated as A_0, A_1, A_2, and A_3 bands, and it has been suggested that their frequencies depend on different conformational states of myoglobin, hemoglobin, and other heme proteins.[23,46−48] However, studies of myoglobin mutants in the H64, V68, F43, F46, and L29 positions often exhibit a single ν_{CO} frequency.[49] In particular, among the H64 mutants, only H64A and H64G show evidence for multiple ν_{CO} bands. Disordered H_2O in the distal pocket has been hypothesized to lead to a broad set of IR bands.[19] The A_1 and A_3 conformers are observed in wild-type Mb at pH 7.0 at $\nu_{CO} \approx 1945$ cm^{-1} and a less intense IR band at $\nu_{CO} \approx 1932$ cm^{-1}, respectively. The A_3 conformer is present in a number of different species of myoglobin (MbCO) and hemoglobin (HbCO), but is much smaller in adducts of the H93G mutant of SWMb. [31] The IR band corresponding to the A_0 state at $\nu_{CO} \approx 1967$ cm^{-1} is observed to grow in intensity as the pH is lowered to pH 4.5. The A_0 state, also known as the "open" state, has a lower binding affinity and a more rapid CO recombination rate.[28,50] The X-ray crystal structure of a low pH form of myoglobin corresponds to the structure of the A_0 substate.[37] The X-ray crystal structure and single-crystal Raman data[51] of the A_0 conformer are consistent with a protonated H64 that has moved to the surface of the protein so that it is exposed to solvent, which increases the volume of the distal pocket in the A_0 substate. Direct experimental evidence for specific hydrogen-bonding interactions in the A_1 substate has been obtained from deuteration shifts of the ν_{Fe-CO} Raman band.[15,52] The origin of A states of myoglobin categorized by the ν_{CO} and ν_{Fe-CO} stretching frequency of observed IR bands is still in need of an explanation that ties together structure and function.

The reversible binding of oxygen essential for the function of myoglobin must occur in competition with endogenous CO. The origin of the discrimination against CO binding has been considered in terms of both an electrostatic[13,53] and a steric role.[54] A molecular understanding of the electrostatic effect requires determination of the relative strength of hydrogen bonding as opposed to other internal fields.[3,4,15] Hydrogen bonding by the distal histidine may favor the binding of O_2 over that of CO.[15,53,55] In the hydrogen-bonding model, the open state has lower O_2 binding affinity and may represent a physiological trigger for O_2 release.[14,28] Because the ν_{CO} frequency is correlated with the open (A_0) and closed (A_1) conformers, it reports on the difference in electrostatic environment that has functional consequences for diatomic ligands other than CO.[25] The vibrational Stark effect on ν_{CO} demonstrates that frequency shifts can arise from an internal field. The vibrational Stark effect arises from an electric field interaction with a change in the CO dipole moment, $\Delta\mu_{CO} = \mu_{CO}$(excited) $- \mu_{CO}$(ground), known as the Stark tuning rate.[4,5,56−58] Internal or matrix electric field effects can arise due to specific molecular interactions such as hydrogen bonding of a group CO···H−X, where X = C, N, or O. The concerted motion of several groups can also be viewed as a steric mechanism because the motion is coupled to α-helices that shift in position upon photolysis.[54] The goal of the present theoretical study is to quantitatively relate these two phenomena. The ansatz of this study is that both the Stark tuning rate and the effect of hydrogen bonding should be connected to structure through the change in the CO bond length.

The present study investigates the effect of an applied electric field and of hydrogen bonding to explain the ν_{CO} frequency shifts of CO in mutants and conformational substates of myoglobin. The vibrational frequency of ν_{Fe-CO} and ν_{CO} has been calculated for a range of different hydrogen-bond partners CH_4, NH_3, H_2O, H^+, imidazole, imidazolium, pyrrole, and combinations of these at angles from 180° to 100° and over a range of distances. Density functional theory (DFT) was used to determine the frequencies. DFT is an appropriate method for the calculation of frequencies and potential energy surfaces.[59−66] Comparisons have been made between two functionals, the BLYP functional and the GGA functional, as was done in a previous study of proximal effects on the ν_{Fe-CO} and ν_{CO} frequencies.[67] The calculated geometries and Mulliken charges for over 70 structures are compared to obtain a correlation plot that represents the vibronic effect of hydrogen bonding with the Fe−C−O oscillator. The approach is conceptually related to the vibronic theory of activation (VTA),[1] although Mulliken charge is used for the correlation instead of the population of specific molecular orbitals.[20] This study shows the consistency of the hydrogen-bonding hypothesis[28,68] with vibrational Stark effect measurements,[4] solutions of the Poisson−Boltzmann equation,[19] and the VTA approach[20] to obtain the magnitude of electrostatic effects.

Methods

The optimized ground-state geometries were obtained using both the generalized gradient approximation (GGA) of Perdew and Wang[69] and the BLYP functional[70,71] as implemented in DMol3 (Molecular

(45) Shimada, H.; Caughey, W. S. *J. Biol. Chem.* **1982**, *257*, 11893−11900.
(46) Potter, W. T.; Hazzard, J. H.; Choc, M. G.; Tucker, M. P.; Caughey, W. S. *Biochemistry* **1990**, *29*, 6283−6295.
(47) Caughey, W. S.; Alben, J. O.; McCoy, S.; Boyer, S. H.; Carache, S.; Hathaway, P. *Biochemistry* **1969**, *8*, 59−62.
(48) Frauenfelder, H.; Park, F.; Young, R. D. *Annu. Rev. Biophys. Chem.* **1988**, *17*, 451−479.
(49) Li, T. S.; Quillin, M. L.; Phillips, G. N.; Olson, J. S. *Biochemistry* **1994**, *33*, 1433−1446.
(50) Tian, W. D.; Sage, J. T.; Champion, P. M.; Chien, E.; Sligar, S. G. *Biochemistry* **1996**, *35*, 3487−3502.
(51) Zhu, L.; Sage, J. T.; Rigos, A. A.; Morikis, D.; Champion, P. M. *J. Mol. Biol.* **1992**, *224*, 207−215.
(52) Ramsden, J.; Spiro, T. G. *Biochemistry* **1989**, *28*, 3125−3128.
(53) Sigfridsson, E.; Ryde, U. *J. Biol. Inorg. Chem.* **1999**, *4*, 99−110.
(54) Kachlova, G. S.; Popov, A. N.; Bartunik, H. D. *Science* **1999**, *284*, 473−476.
(55) Springer, B. A.; Sligar, S. G.; Olson, J. S.; Phillips, G. N. *Chem. Rev.* **1994**, *94*, 699−714.
(56) Hush, N. S.; Reimers, J. R. *J. Phys. Chem.* **1995**, *99*, 15798−15805.
(57) Reimers, J. R.; Hush, N. S. *J. Phys. Chem. A* **1999**, *103*, 10580−10587.
(58) Andrews, S. S.; Boxer, S. G. *J. Phys. Chem.* **2002**, in press.
(59) Groot, M. J. d.; Havenith, R. W. A.; Vinkers, H. M.; Zwaans, R.; Vermeulen, N. P. E.; Lenthe, J. H. v. *J. Comput.-Aided Mol. Des.* **1998**, *12*, 183−193.
(60) Wirstam, M.; Blomberg, M. R. A.; Siegbahn, P. E. M. *J. Am. Chem. Soc.* **1999**, *121*, 10178−10185.
(61) Kuramochi, H.; Noodleman, L.; Case, D. A. *J. Am. Chem. Soc.* **1997**, *119*, 11442−11451.
(62) Vogel, K. M.; Kozlowski, P. M.; Zgierski, M. Z.; Spiro, T. G. *J. Am. Chem. Soc.* **1999**, *121*, 9915−9921.
(63) Vogel, K. M.; Kozlowski, P. M.; Zgierski, M. Z.; Spiro, T. G. *Inorg. Chim. Acta* **2000**, *297*, 11−17.
(64) Harvey, J. N. *J. Am. Chem. Soc.* **2000**, *122*, 12401−12402.
(65) McMahon, B. H.; Stojkovic, B. P.; Hay, P. J.; Martin, R. L.; Garcia, A. E. *J. Chem. Phys.* **2000**, *113*, 6831−6850.
(66) Wasileski, S. A.; Koper, M. T. M.; Weaver, M. J. *J. Phys. Chem. B* **2001**, *105*, 3518−3530.
(67) Franzen, S. *J. Am. Chem. Soc.* **2001**, *123*, 12578−12589.
(68) Tian, W. D.; Sage, J. T.; Champion, P. M. *J. Mol. Biol.* **1993**, *233*, 155−166.
(69) Perdew, J. P.; Chevary, J. A.; Vosko, S. H.; Jackson, K. A.; Pederson, M. R.; Singh, D. J.; Fiolhais, C. *Phys. Rev. B* **1992**, *46*, 6671−6687.
(70) Becke, A. D. *J. Chem. Phys.* **1997**, *107*, 8554−8560.
(71) Lee, C. L.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785−789.

Simulations Inc.).[72,73] All calculations were carried out on the SGI/Cray Origin 2000 and IBM SP supercomputers at the North Carolina Supercomputer Center (NCSC). A numerically tabulated numerical basis set of double-ζ plus polarization (DNP) quality was employed as described in the Supporting Information. The calculations included geometry optimization, potential energy surface calculation, and vibrational frequency calculation. Figure 1 shows an example of an imidazole−iron−porphine−CO adduct. Geometry optimizations of this structure were carried out with two (O, H), three (Fe, O, H), or four (Fe, O, H, X where X = O, N, or C) fixed atoms. The reference point for determination of the relative energy of the hydrogen bond was a calculation with the hydrogen-bonding group placed at a distance of 10 Å.

The procedure for the vibrational frequency calculation is given in the following. The geometry optimizations were carried out until the energy difference was less than 10^{-6} au on subsequent iterations. Following geometry optimization, the Hessian matrix, **V**, was constructed by finite differences of analytic gradients. The finite differencing proceeds from atom to atom. An energy calculation is carried out for displacements in the *x*, the *y*, and the *z* directions of each atom in the molecule. **V** has dimension $3N \times 3N$, where *N* is the number of atoms in the molecule. The Cartesian second derivatives in energy (force constants), which are the elements of **V**, are listed in order of x_j, y_j, and z_j for each atom *j* successively. The vibrational frequencies are obtained by matrix diagonalization of the resulting Hessian matrix in mass-weighted Cartesian coordinates. Mass weighting is applied by the transformation $\mathbf{H} = \mathbf{M}^{-1/2}\mathbf{V}\mathbf{M}^{-1/2}$, where $\mathbf{M}^{-1/2}$ is a diagonal matrix of dimension 3*N* x 3*N*, where the diagonal elements are $1/\sqrt{m_j}$ for $j =$ 1 to 3*N*. The elements of $\mathbf{M}^{-1/2}$ are also in order of x_j, y_j, and z_j for each atom *j* successively. The matrix equation for this procedure is $\mathbf{S}^{-1}\mathbf{H}\mathbf{S} = \Lambda$, where **S** is the matrix of eigenvector coefficients, and Λ is the diagonal matrix of eigenvalues for *j* modes λ_j for $j = 1$ to $3N$.[74] The dimensions of **S** are 3*N* rows by *N* columns. The Cartesian matrix diagonalization procedure gives six eigenvalues (for a nonlinear molecule) that correspond to translations and rotations of the center of mass of the molecule and $3N - 6$ eigenvalues that correspond to vibrational normal modes. The vibrational frequencies, ν in units of wavenumbers (cm^{-1}), are obtained from the eigenvalues by $\lambda_j = 4\pi^2c^2\nu^2$. The infrared intensities are calculated using the finite difference procedure to estimate the difference dipole moment derivative $\partial\mu/\partial Q_j$ along a given normal mode, where the normal coordinates for each of the *j* modes correspond to each of three consecutive rows of **S** (x_j, y_j, z_j, etc.). The transition intensity is obtained within the harmonic approximation from the square of the vibrational transition moment.[75] AVS (Advanced Visual System) and insightII (Molecular Simulations) programs were used for visualization of the results.

Results

The effect of various chemical groups that hydrogen bond to heme-CO was considered systematically as a function of distance and angle. The geometry, electronic structure, and frequencies of the iron−carbonyl-bonded system were compared for four different distal groups, H_2O (as a model for serine and threonine), NH_3 (as a model for unprotonated lysine), CH_4 (as a model for leucine, isoleucine, and valine), and imidazole (as a model for histidine). These groups span most of the chemical interactions of biologically relevant amino acids in heme proteins. The potential energy of interaction and optimal hydrogen-bond length for each of the groups are both given in Figure 2. The magnitude of the hydrogen-bond interaction energy is smallest for CH_4, and increases for NH_3 and H_2O. This increase is not surprising because the hydrogen-bonding interaction considered here, CO···H−X, is expected to increase in order of increasing electronegativity for X = C, N, O. The imidazole (Im) π-system must contribute significantly to the interaction because the stabilization energy is larger and the optimal CO···H−N distance is significantly shorter for Im than for NH_3. The interaction energies calculated are comparable to those obtained in a previous study,[53] −17.5, −14.5, and −8.2 kJ/mol for Im, H_2O, and NH_3, respectively, as compared to values of −18.3, −14.4, and −7.5 kJ/mol as shown in Figure 2 (see also Supporting Information).

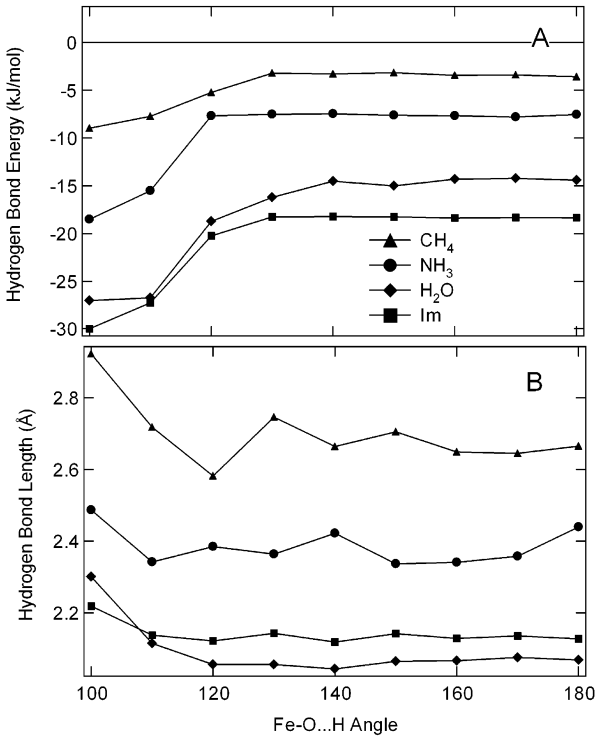

Figure 2. Hydrogen-bond energy and bond length for four hydrogen-bonding groups to Im−porphine−CO as a function of the Fe−O···H bond angle, θ. (A) The hydrogen-bond energy is shown as calculated using a potential energy surface relative to a reference energy calculated at a distance of 10 Å for each hydrogen-bonding group. (B) The optimum hydrogen-bonding distance determined from the potential energy surface (PES). The PES for CH_4 is relatively shallow, and thus there is a larger variability in the minimum position than for the other hydrogen-bonding groups.

The interaction energy and optimal hydrogen-bond lengths given in Figure 2 reflect the change in orbital interaction of the groups with the CO σ-orbitals, which is maximal at $\theta = 180°$, and π-orbitals, which is maximal at $\theta = 100°$. The hydrogen bond with CO increases in strength by a factor of 2 for the polar groups, and by nearly a factor of 3 for CH_4, as the angle θ is varied from 180° to 100° (see Figure 1). The optimal nonbonding distances for CH_4, NH_3, H_2O, and Im increase by 0.23, 0.12, 0.25, and 0.1 Å, respectively, as the angle θ is varied from 180° to 100° (Figure 2).

Geometry. The Fe−C and C−O bond lengths obtained from the GGA DFT calculation are d(Fe−C) ≈ 1.77 Å and d(C−O) ≈ 1.16 Å, respectively, as shown in Figure 3 and tabulated in the Supporting Information. These values are comparable to the values of other studies. For example, the Fe−C bond length was calculated to be d(Fe−C) ≈ 1.74 Å[53] and d(Fe−C) ≈ 1.77

(72) Delley, B. *J. Chem. Phys.* **1990**, *92*, 508−517.
(73) Delley, B. *J. Chem. Phys.* **2000**, *113*, 7756−7764.
(74) Wilson, E. B., Jr.; Decius, J. C.; Cross, P. C. *Molecular Vibrations*; Dover: New York, 1955.
(75) McHale. *Molecular Spectroscopy*; Prentice Hall: Upper Saddle River, NJ, 1999.

BATCH: ja11a34 USER: jld69 DIV: @xyv04/data1/CLS_pj/GRP_ja/JOB_i44/DIV_ja017708d DATE: September 17, 2002

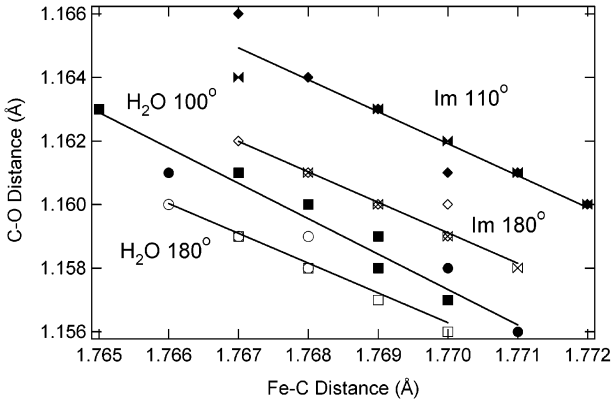

Figure 3. Dependence of the Fe−C and C−O bond lengths on the distance and angle of hydrogen-bonding groups. The effects of both H_2O and imidazole (Im) are shown for two angles each. The series of points represent distances of 2.0, 2.2, 2.4, 2.6, 2.8, and 3.0 Å from left to right across the figure. Results of both calculations are shown on the same plot. Open and filled circles represent the BLYP calculation for H_2O at 180° and 100°, respectively. Open and filled squares represent the GGA calculation for H_2O at 180° and 100°, respectively. Open and filled diamonds represent the BLYP calculation for Im at 180° and 110°, respectively. Open and filled bowties represent the GGA calculation for Im at 180° and 110°, respectively.

Å[76] for nonporphine model systems. Porphine−Fe−CO calculations similar to those performed here gave values of d(Fe−C) ≈ 1.72 Å[77] and d(Fe−C) ≈ 1.78 Å[63] by DFT methods. The DFT calculated C−O bond length is d(C−O) ≈ 1.17 Å in a number of other studies.[53,63,76,77] Hydrogen bonding by imidazole (CO···H−Im) results in calculated bond length changes of −0.019 and +0.006 Å, respectively, for d(Fe−C) and d(C−O).[76] These compare well with values of −0.017 and + 0.006 Å obtained in the present study. DFT calculations are also in reasonable agreement with the most recent high-resolution X-ray crystal structure that gives the Fe−C and C−O bond lengths as d(Fe−C) ≈ 1.73 Å and d(C−O) ≈ 1.12 Å, respectively.[54] A discrepancy in the C−O bond length is the major difference between the DFT calculations[53,63,76,77] and the X-ray structural data.[54,78−80] The calculated geometries are in closer accord with the Fe $d_\pi \rightarrow$ CO π^* back-bonding interaction model that predicts longer bond lengths for iron-bound CO than for free CO commensurate with the trends in the frequency data.[78,81]

To calculate the effects of hydrogen bonding, geometric constraints were imposed on the geometry optimization procedure. The angle dependence for various hydrogen-bonding groups was determined at a fixed hydrogen-bond distance of 1.92 Å with fixed positions for atoms Fe, O, and H. Thus, the carbon in Fe−C−O···H was free to vary as were all remaining atoms. Details are given in the Supporting Information.

In all of the calculations, the iron−pyrrole-nitrogen (Fe−Np) bond lengths changed little with hydrogen-bonding distance or geometry. In the more than 70 calculated structures considered, the total change in the Fe−Np distance was less than 0.0016 Å for calculations using the GGA functional. The Fe−Np distance is nearly constant for the hydrogen-bonding groups H_2O, NH_3, and CH_4. On the other hand, the changes in the Fe−Np distance are nearly 3 times as large for changes in proximal ligation. A recent study of hydrogen bonding to the Nδ−H hydrogen found significant core size effects (i.e., effect on Fe−Np) as the axial imidazole bonding distance, that is, the Fe−Nϵ bond length, decreased.[67] Fe−Nϵ bond lengths are relatively constant for all of the structures studied here except for a decrease of 0.05−0.1 Å in Fe−Nϵ bond length for all hydrogen-bonding groups for 110 < θ < 140° (results not shown).

The effect of hydrogen-bond distance on the Fe−C−O geometry is shown in Figure 3 for H_2O and Im interaction with bound CO. The range of CO···H distances studied was 2.0−3.0 Å. Figure 3 shows an inverse relationship between the Fe−C and C−O bond lengths as a function of the hydrogen-bond distance. The effect of both H_2O and Im on the C−O bond length is larger for smaller Fe−O···H angles. In summary, based on the GGA calculation, a side hydrogen-bonding interaction results in a 3-fold greater increase in C−O bond length than an end interaction. Numerical values are given in the Supporting Information.

Mulliken Charge. The magnitude of charge displacement from the iron to CO increases as θ approaches 100° for all of the distal hydrogen-bonding groups studied. The Fe → CO charge displacement was estimated by the sum of the C and O Mulliken charges, $q_C + q_O$. Because the total Mulliken charge $q_C + q_O$ is zero for free CO, the magnitude of this sum is a measure of the net charge displacement from the iron−porphine complex onto bound CO. For the hydrogen-bonding ligands studied here, there is between −0.02 and −0.11 units of negative charge displaced onto CO. The Mulliken charge sum $q_C + q_O$ grows more negative by > −0.01 charge unit for H_2O, NH_3, CH_4, and Im as the angle decreases down to an angle of 140°. For Im and CH_4, the largest negative value of Mulliken charge is reached at 100°. For H_2O and NH_3, the charge displacement onto CO is greatest for $\theta \approx 140°$.

Frequencies. The ν_{Fe-CO} and ν_{CO} stretching frequencies range from 522 to 541 cm^{-1} and 1955−2043 cm^{-1}, respectively, for the GGA DFT calculation as shown in Figure 4. The range represents the angle and distance dependence of the principal four hydrogen-bonding groups studied, H_2O, NH_3, CH_4, and Im, in proximity to bound CO. The values of ν_{Fe-CO} and ν_{CO} for a non-hydrogen-bonded CO are 498 and 2013 cm^{-1}, respectively, using the GGA DFT method.[67] The values are comparable to the ν_{Fe-CO} and ν_{CO} stretching frequency calculated by other DFT methods. For example, ν_{Fe-CO} and ν_{CO} in a porphine model system are shifted from 586 and 1907 cm^{-1} to 573 and 1927 cm^{-1}, respectively, for an interaction with the lone pair of NH_3 as compared to CO alone.[53] The shift of ν_{Fe-CO} to lower values and of ν_{CO} to higher values is consistent with lone pair interactions as shown in Table 2. A second comparison with other DFT calculations is provided by the shift of ν_{CO} from 1967 cm^{-1} for Fe−CO to 1935 cm^{-1} for ν_{CO} with a hydrogen bond to imidazole (Fe−CO···Im).[76] Using the BLYP functional, the current method gives 464 and 1995 cm^{-1} for ν_{Fe-CO} and ν_{CO}, respectively, without any hydrogen bonding and 515 and 1955 cm^{-1} for a hydrogen bond to imidazole.[67] Tabulated numerical results that correspond to Figure 4 are given in the Supporting Information.

For all of the hydrogen-bonding groups, the ν_{Fe-CO} and ν_{CO} stretching vibrations increase in parallel with increasing angle θ and inversely with increasing distance (Figure 4). The ν_{CO}

(76) Cui, Q.; Karplus, M. *J. Chem. Phys.* **2000**, *112*, 1133−1149.
(77) Rovira, C.; Kunc, K.; Hutter, J.; Ballone, P.; Parrinello, M. *J. Phys. Chem. A* **1997**, *101*, 8914−8925.
(78) Ray, G. B.; Li, X.-Y.; Ibers, J. A.; Sessler, J. L.; Spiro, T. G. *J. Am. Chem. Soc.* **1994**, *116*, 162−176.
(79) Momenteau, M.; Reed, C. A. *Chem. Rev.* **1994**, *94*, 659−698.
(80) Peng, S. M.; Ibers, J. A. *J. Am. Chem. Soc.* **1976**, *98*, 8032−8036.
(81) Spiro, T. G.; Smulevich, G.; Su, C. *Biochemistry* **1990**, *29*, 4497−4508.

BATCH: ja11a34 USER: jld69 DIV: @xyv04/data1/CLS_pj/GRP_ja/JOB_i44/DIV_ja017708d DATE: September 17, 2002

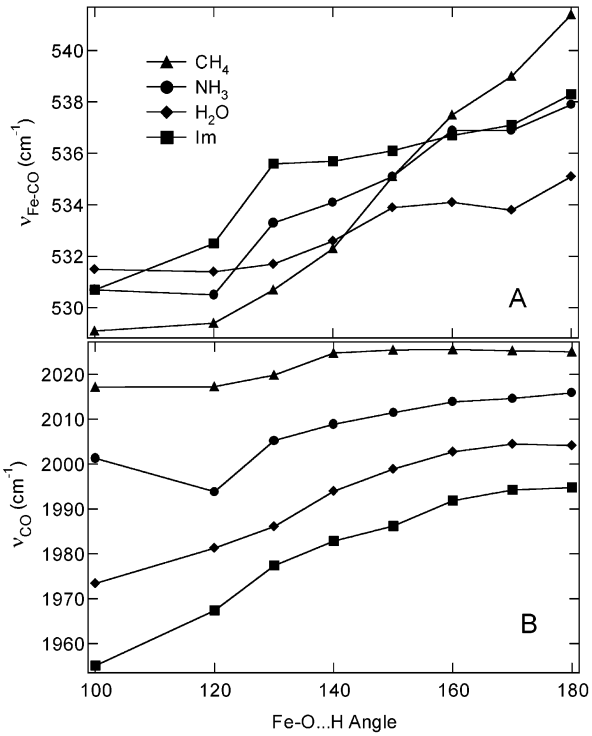

Figure 4. Model hydrogen-bonding GGA calculations as a function of angle at a fixed CO···H distance of 1.916 Å. (A) The ν_{Fe-CO} frequencies are given for the hydrogen-bonding molecules indicated in the legend in order of the strength of their interaction with CO. (B) The ν_{CO} frequencies are given as a function of the angle of the hydrogen-bonding group with respect to the Fe−C−O axis.

frequency decreases in the series $CH_4 > NH_3 > H_2O >$ Im as can be seen by comparing the frequencies in Figure 4B. This decrease in ν_{CO} is expected on the basis of the order of the hydrogen-bond strength Im > H_2O > NH_3 > CH_4 (Figure 2). The effect of hydrogen bonding on the ν_{Fe-CO} vibration is less pronounced. However, it is generally the case that stronger hydrogen-bonding groups such as H_2O (see Figure 2) have lower ν_{Fe-CO} frequencies. The values for imidazole show an interesting deviation from this trend. Although imidazole has a stronger interaction with CO than the nonaromatic groups H_2O, NH_3, and CH_4, the ν_{Fe-CO} frequency of imidazole is closest to that of NH_3. The interaction of the π system of imidazole presumably contributes to the interaction energy with bound CO, but does not affect the ν_{Fe-CO} frequency as greatly. On the other hand, the ν_{Fe-CO} and ν_{CO} frequencies show a π-back-bonding correlation as a function of the distance of the hydrogen-bonding group at a fixed angle. The ν_{Fe-CO} and ν_{CO} frequencies given in Figure 4 for two different angles, θ, of the hydrogen-bonding group Fe−O···H show a negative correlation with the hydrogen-bond distance.

Table 2 provides frequency data for a number of additional hydrogen-bonding groups including imidazolium and the lone pairs of H_2O and Im. Table 2 shows that imidazole protonation leads to a significant frequency lowering of ν_{CO}. Imidazolium hydrogen bonding at a distance of 1.916 Å leads to a lengthening of the C−O bond to >1.17 Å for two different θ angles. For θ = 110°, the Fe−C bond is lengthened to 1.771 Å, leading to a much lower ν_{Fe-CO} frequency than that in neutral imidazole. The pyrrole model is a test of the effect of substitution of the Nδ nitrogen by carbon. Table 2 shows that the frequencies of both ν_{Fe-CO} and ν_{CO} are increased for a pyrrole model at θ = 110°, indicative of significantly weaker hydrogen bonding for C−H than for Nϵ−H of imidazole.

Table 2. Comparison of Bond Lengths and Frequencies of a Number of Hydrogen-Bonding Environments around CO Calculated Using the GGA DFT Method[a]

	Fe–Nϵ	Fe–Np	Fe–C	C–O	ν_{Fe-CO}	ν_{CO}	intensity
CO···H−ImH$^+$							
110	2.054	2.022	1.771	1.170	517.8	1923.6	696
180	2.051	2.020	1.754	1.175	547.3	1914.5	1315
CO···H−Im							
110°	2.048	2.020	1.763	1.166	529.0	1958.5	848
180°	2.056	2.019	1.764	1.163	538.3	1994.8	1050
CO···H−Py							
110°	2.047	2.020	1.766	1.167	502.0	1993.2	653
180°	2.055	2.020	1.766	1.163	530.4	2028.3	892
CO···LP−Im							
100°	2.047	2.019	1.796	1.154	491.6	2023.2	586
180°	2.055	2.019	1.766	1.163	538.3	1996.5	679
CO···OH_2							
100°	2.053	2.019	1.776	1.151	519.1	2074.2	558
180°	2.056	2.019	1.773	1.153	523.0	2061.1	624
CO···H_2O							
100°	2.054	2.019	1.764	1.164	531.5	1973.4	756
180°	2.055	2.019	1.765	1.161	535.1	2004.2	945
CO···H_4C							
100°	2.056	2.019	1.769	1.159	529.1	2017.2	691
180°	2.056	2.019	1.769	1.158	541.4	2025.1	796
CO···H_3N							
100°	2.055	2.019	1.767	1.161	530.7	2001.3	696
180°	2.056	2.019	1.767	1.159	537.9	2015.9	844
angle 100°							
no H-bond	2.074	2.020	1.786	1.157	498.1	2013.4	631
Im(1.92), OH_2(3.1)	2.054	2.020	1.771	1.163	511.1	1953.5	743
Im(1.92), H_2O(2.2)	2.054	2.020	1.755	1.169		1951.8	797
Im(1.92), CH_4(2.0)	2.047	2.020	1.762	1.167	535.4	1939.1	847
CH_4(1.92), CH_4(2.0)	2.051	2.019	1.767	1.160	529.7	1993.6	800
CH_4(1.92), H_2O(2.0)	2.052	2.019	1.762	1.166	535.0	1957.7	864
H_2O(1.92), CH_4(2.0)	2.052	2.019	1.761	1.166	535.0	1956.0	1011
ImH^+(2.2), CH_4(2.0)	2.051	2.022	1.769	1.173	518.4	1896.7	757
ImH^+(3.0), CH_4(2.0)	2.055	2.021	1.778	1.166	508.8	1945.0	671
ImH^+(3.4), CH_4(2.0)	2.052	2.021	1.780	1.164	508.0	1959.9	652
ImH^+(3.8), CH_4(2.0)	2.052	2.021	1.782	1.1625	508.0	1971.0	641

[a] All calculations that involve a single hydrogen-bonding group have a CO···H distance at a fixed value of 1.91 Å. The groups studied are imidazolium (ImH^+), imidazole (Im) and pyrrole (Py), imidazole Ne-tautomer (LP−Im), water with the lone pair directed at CO (OH_2), water (H_2O), methane (CH_4), ammonia (NH_3), free CO, and combinations of the above groups.

It has been proposed that lone pair interactions increase ν_{CO} by virtue of a polarization of CO opposite to that in hydrogen-bonding interactions.[18,19] As an example of such a lone pair interaction, Table 2 gives the results of a vibrational frequency calculation for the Nϵ-tautomer of imidazole. In this tautomer, the hydrogen bond is eliminated, and the Nδ lone pair is the only interaction with CO. The consequence is that ν_{CO} shifts from ~1959 to ~2023 cm^{-1} (Table 2). For the interaction of a water lone pair, CO···OH_2 results in a ν_{CO} shift from ~1973 to ~2074 cm^{-1}. For both H_2O and Im lone pair interactions, the increase in ν_{CO} frequency is accompanied by a downshift of ν_{Fe-CO}.

It is apparent from the X-ray crystal structure data given in Table 1 that specific interactions of groups in the distal pocket with bound CO involve multiple interactions. Calculations were carried out for doubly hydrogen-bonded species as well (Table 2). Particular emphasis is placed on double hydrogen-bonding geometries that include CH_4, where methane serves as a model for aliphatic amino acids. These geometries account for the mutants of myoglobin that involve substitution of aliphatic amino acids in the distal pocket. Examination of Table 1 indicates that the majority of close contacts involve C−H hydrogen bonding as one of a number of interactions with bound CO. Table 2 shows that the addition of a CH_4 hydrogen bond at a distance of 2.0 Å to a first hydrogen bond from CH_4, H_2O, Im, or ImH^+ results in a frequency lowering of ν_{CO} by 24, 17, 19, or 23 cm^{-1}, respectively, relative to the singly hydrogen-bound species. The hydrogen bonding of water in the distal pocket can have a profound effect. For a single CH_4 group at a distance of 1.92 Å, $\nu_{CO} \approx 2017$ cm^{-1}. This frequency shifts down by ~60 cm^{-1} to a value of $\nu_{CO} \approx 1957$ cm^{-1} when an additional H_2O is hydrogen bonded at a distance of 2.0 Å opposite CH_4. A series of double hydrogen-bond geometries consisting of ImH^+ and CH_4 were calculated as models of the "open" or A_0 state of myoglobin[28,68] that has a protonated histidine. As a result of protonation, the histidine swings out toward the surface of the globin and is located further from CO than neutral imidazole.[37] Because the interaction of imidazolium with bound CO is quite strong, it is reasonable to consider the effect of long-range interactions in this case. As shown in Table 2, the combination of ImH^+ and CH_4 gives rise to relatively low frequencies (and long C−O bond lengths) even when the hydrogen of ImH^+ is as far as 3.8 Å from the oxygen of bound CO.

A Vibronic Approach. The vibronic theory of activation (VTA) is based on the idea that the bonding of bound ligands is altered by charge transfer from a metal center that leads to a change in orbital occupancy.[82] The change in orbital occupancy can be calibrated by calculation of molecule properties (e.g., bond length and frequency) for a small molecule along with its cation and anion. The calculation for the isolated ligand serves to establish the correlation of change in bond length and frequency as charge is injected or withdrawn from the molecule. The vibronic theory has been applied to CO using semiempirical methods.[20,21] The difference in the present approach is that a change in Mulliken charge is taken as an indicator of the strength of the Fe−CO interaction instead of the occupancy of the 5σ and $2\pi^*$ orbitals. While the correlation with occupancy of the 5σ and $2\pi^*$ orbitals is of interest, the correlation that emerges in DFT calculations is complicated by the fact that these orbitals are involved in a number of MOs in iron porphine. The current approach employs C−O bond lengths (Figure 3), frequencies (Figure 4), and Mulliken charges given in Table 2 to study CO polarization.

A linear correlation for bond length versus frequency for four species, neutral CO, CO cation, CO anion, and ^{1}CO excited state, is shown in Figure 5. A linear correlation with frequency is the same as Badger's rule,[83] because a plot of the bond length versus force constant similar to Figure 5 is still linear. The bond length−frequency correlation obtained from over 70 vibrational frequency calculations in Table 2 and in the Supporting Information is shown in Figure 5. A fit of these calculated values to a line yields the following relation:

$$\nu_{CO} = 2168 - 8677(R_{CO} - R_{CO}^0)\ cm^{-1} \qquad (1)$$

where R_{CO} is obtained from a geometry optimization, and R_{CO}^0 is the equilibrium bond length of neutral CO, $R_{CO}^0 = 1.141$ Å. Figure 5 shows that the slope of the line fit to the calculated metalloporphine CO frequencies is larger than the slope for the unbound CO neutral, cation, anion, and singlet excited state. For this reason, the free CO values are not used for quantitative applications and are shown in Figure 5 only for comparison. The bond length−frequency correlation given in eq 1 can be used to estimate the CO frequency after a geometry optimization of a particular hydrogen-bonding group or groups.

The correlation of the frequency with the Mulliken charge is shown in Figures 6 and 7. The sum of the Mulliken charges on the carbon q_C and oxygen q_O atoms, $q_C + q_O$, gives an estimate of the relative amount of charge transfer from the iron−imidazole porphine moiety. The charge displacement obtained from a sum of Mulliken charges ranges from $q_C + q_O = -0.02$ to −0.1. Because the VTA approach is based on the idea that charge transfer between metal and bound ligand is proportional

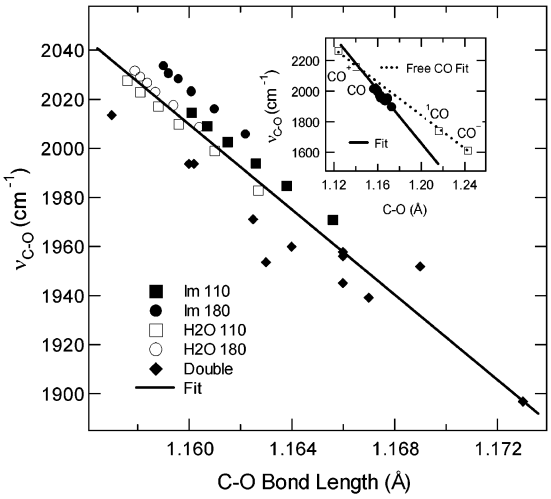

Figure 5. Distance−frequency correlation obtained for a number of data points obtained from the frequency calculations performed. The calculations used to obtain the correlation include the distance dependence for Im and H_2O hydrogen-bonding groups and the double hydrogen-bonding groups. The dashed line in the inset is the frequency correlation for free CO along with its cation (CO^+), anion (CO^-), and excited state (^{1}CO) calculated using the same GGA DFT basis set used for the porphine model calculations. The fit to the solid line is given by eq 1.

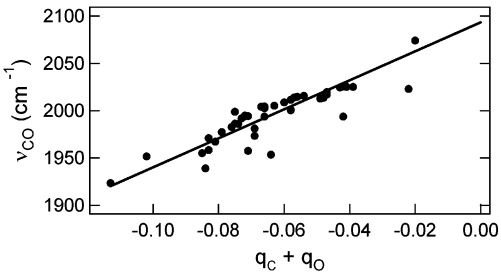

Figure 6. The charge displacement−frequency correlation shown for a number of the calculated frequencies. The sum of the Mulliken charges on C and O of the CO ligand is shown as $q_C + q_O$. The linear correlation obtained is given in eq 2.

(82) Stavrov, S. S. *Biophys. J.* **1993**, *65*, 1942−1950.
(83) Badger, R. M. *J. Chem. Phys.* **1934**, *2*, 128−131.

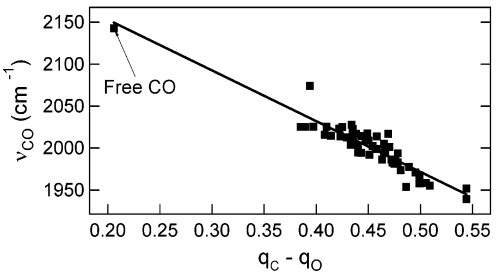

Figure 7. A plot of the ν_{CO} frequency as a function of the charge displacement within bound CO. The calculated frequencies shown represent all of the data in the Supporting Information, except those involving the charged residue imidazolium (ImH^+).

to the strength of the interaction, this Mulliken method has the same significance, but loses the distinction between σ- and π-effects in the original VTA theory. [1] The line in Figure 6 is given by

$$\nu_{CO} = \{2094 + 1535(q_C + q_O)\}\ \text{cm}^{-1} \qquad (2)$$

There are several outlying points that are well below the line. These are the species Im OH_2 at $q_C + q_O = -0.064$, $\nu_{CO} = 1953.5\ \text{cm}^{-1}$, the N$\epsilon$ tautomer of Im at $q_C + q_O = -0.022$, $\nu_{CO} = 2023.2\ \text{cm}^{-1}$, and the dimethane hydrogen-bonded species at $q_C + q_O = -0.042$, $\nu_{CO} = 1993.9\ \text{cm}^{-1}$, and other double hydrogen-bonded species. These deviations appear to lie on a second line shifted 30 cm^{-1} to lower energy. If all of the calculated frequencies in this study are included, the slope is 1433 cm^{-1}/Mulliken charge, and the intercept is the same.

A third correlation is observed for the polarization of the CO bond. The interaction of Fe with CO removes electron density from the carbon by interaction with the 5σ orbital and increases electron density on oxygen by π-back-bonding to the 2π* orbital. Of course, π-back-bonding tends to replenish some of the charge on oxygen. Nonetheless, the difference in the Mulliken charges on carbon and oxygen is a measure of this polarization that does not require detailed examination of the occupancies of the 5σ and 2π* orbitals in a number of different MOs where they interact with metal, porphine, and trans ligand orbitals. The correlation for the polarization $\Delta q_{CO} = q_C - q_O$ with frequency is excellent as shown in Figure 7. All of the calculations for neutral hydrogen-bonding ligands that were carried out in this paper are included in Figure 7. The most appropriate fit is to a linear function that allows the frequency to approach the free CO value as Δq_{CO} approaches zero.

$$\nu_{CO} = \{2274 - 605\Delta q_{CO}\}\ \text{cm}^{-1} \qquad (3)$$

The polarization Δq_{CO} is defined so that it will be a positive number representing the extent of charge displacement within bound CO.

Calculation of the Stark Tuning Rate. The effect of applied electric field on the C−O bond length and Mulliken charge was calculated by geometry of optimization of the imidazole−iron−porphine−CO system at four values of an applied electric field perpendicular to the plane of the porphine. The simplified VTA approach described above provides a means to use geometry-optimized structures to correlate distance or Mulliken charge presented in Table 3 with ν_{CO} frequency. The application of VTA is shown in Table 3 for four values of the applied electric field spanning ~±1 MV/cm. Using C−O bond distances and Mulliken charges, the ν_{CO} frequency was calculated using eqs 1 and 2. The calculated Stark tuning rates using eqs 1 and 2 are 2.6 and 2.25 cm^{-1}/(MV/cm), respectively. These values compare extremely well with the experimental value of 2.4 cm^{-1}/(MV/cm).[4] The Stark tuning rate for free CO is calculated to be 0.5 cm^{-1}/(MV/cm), which is also in reasonable agreement with experiment.[84] The bond-length frequency correlation is validated by this comparison.

Discussion

The DFT calculations provide a set of frequencies and intensities that span conformational and chemical interactions relevant to the CO bands of heme proteins. Hydrogen bonding in the distal pocket emerges as a dominant factor that gives rise to an electrostatic effect on the ν_{Fe-CO} and ν_{CO} frequencies. The A states of wild-type Mb and the effects of mutation can be categorized on the basis of the calculated frequency shifts in ν_{CO} presented in Table 2 and further studied using the VTA approach developed on the basis of these calculations. The origin of open (A_0) and closed (A_1) states of wild-type myoglobin can be sought in the correlation of the X-ray structural data with model calculations of hydrogen bonding that compare imidazole−iron−porphine−CO in the vicinity of imidazolium and imidazole, respectively. The effects of mutations can also be calculated using the VTA approach on the basis of the geometry of hydrogen-bonding groups that surround the CO tabulated in Table 1.

The Hydrogen-Bonding Model for ν_{CO} Frequency Shifts. The hypothesis of the current study is that specific hydrogen-bonding environments are responsible for the observed electrostatic effects on ν_{CO}. The calibration of the calculated effect using an applied electric field, the Stark tuning rate, in terms of a change in C−O bond length presents a complementary view of the origin of the vibrational Stark effect. Current theories of the vibrational Stark effect are based on the anharmonicity of the potential energy surface[56−58,85,86] and a bond polarization term.[58] The current approach is based on the field-induced perturbation of the geometry. Anharmonic terms have been shown to be consistently too small to account for the observed Stark tuning rate.[58] According to the DFT calculations, the small change in bond length of 3×10^{-4} Å/(MV/cm) (Table 3) is the primary effect giving a calculated Stark tuning rate of 2.25−2.5 cm^{-1}/(MV/cm) using eq 1. Given a structural Stark shift of 3×10^{-4} Å/(MV/cm), an internal field acting at a distance within the protein the magnitude of the field would require 10^7 V/cm to give the frequency shift between the A_0 and A_1 substate of $\Delta\nu_{CO} \approx 22\ \text{cm}^{-1}$.

The π-Back-Bonding Correlation. The π-back-bonding correlation observed in many heme proteins consists of an inverse correlation between the bands, ν_{Fe-CO} and ν_{CO}.[78,81] The π-back-bonding correlation for distal hydrogen-bonding groups holds only if the Fe−Np bond length is nearly constant in a series of structures. Equatorial π-back-bonding competes with axial π-back-bonding and complicates the correlation as shown in a previous study of proximal effects.[67] In the present study, the change in Fe−Np bond length is less than $\pm 1.6 \times 10^{-3}$ Å. The resulting small change in the core size has little effect on

(84) Park, E. S.; Boxer, S. G. *J. Am. Chem. Soc.* **2002**, submitted.
(85) Lambert, D. K. *J. Chem. Phys.* **1988**, *89*, 3847−3860.
(86) Bishop, D. M. *J. Chem. Phys.* **1992**, *98*, 3179−3184.

BATCH: ja11a34 USER: jld69 DIV: @xyv04/data1/CLS_pj/GRP_ja/JOB_i44/DIV_ja017708d DATE: September 17, 2002

Table 3. Effect of an Applied Electric Field on Bond Length and Mulliken Charge Using the GGA Functional[a]

MV/cm	Fe–C	C–O	q_C	q_O	$q_C + q_O$	$q_C - q_O$	ν_{CO} eq 1	ν_{CO} eq 2
−1	1.7864	1.1569	0.190	−0.237	−0.047	0.427	2021.4	2021.8
−0.5	1.7865	1.1572	0.190	−0.238	−0.048	0.428	2018.7	2020.3
0	1.7861	1.1572	0.190	−0.239	−0.049	0.429	2018.7	2018.9
0.5	1.7861	1.1573	0.190	−0.239	−0.049	0.429	2017.9	2018.8
1	1.7858	1.1575	0.190	−0.240	−0.050	0.430	2016.2	2017.3

[a] Columns 7 and 8 are calculated values of the CO stretching frequency based on the VTA correlations in the text. The equations used are eqs 1 and 2, respectively. Geometry optimizations were carried out at applied electric fields in the range of $\pm 2 \times 10^{-4}$ au. Using the conversion for applied electric field, 10^{-4} au = 0.5142 MV/cm, to a good approximation the range is ~±1 MV/cm as reported in the table.

the π-overlap of the ring with the iron dπ orbitals. As a consequence, the ν_{Fe-CO} and ν_{CO} frequencies shown in Figure 4 are inversely proportional to the hydrogen-bond distance for either H_2O or Im. However, the calculations shown in Figure 4 for hydrogen bonding on the side of CO ($\theta = 100-110°$) are closer to the slope of the experimental correlation than those for end-on hydrogen bonding ($\theta = 180°$) consistent with the structural data in Table 1. The structural trend of bonding to CO from the side, that is, at Fe−O···H angles of <120°,[34−41] is also consistent with the π-back-bonding correlation in vibrational spectroscopy.[63,78,81]

The orbital symmetry helps to explain why the π-back-bonding effect is so much larger in the case of side-on hydrogen bonding than for end hydrogen bonding. The interaction of the hydrogen-bond group with the σ- and π-bonding orbitals of bound CO will depend greatly on the angle. At 180°, a pure σ-type hydrogen bond has no overlap with the π^* orbital on CO. As the angle decreases, there is an increase in the interaction of the hydrogen-bonding group with the π^* orbital. End and side hydrogen-bonding interactions show π-back-bonding correlations of a different nature. As discussed elsewhere, σ-bonding interactions that involve the Fe−C bond will also affect the π-bonding in the Fe−C−O subsystem.[31,67] An effect that tends to shorten the σ-bond (e.g., lessen σ-bonding repulsion to the imidazole trans to CO) will also increase the π-overlap of the dπ orbitals. Thus, an increase in Fe−C π bonding is expected, as is a concurrent increase in electron density in the $2\pi^*$ orbital of CO leading to a decrease of the CO bond order. By contrast, for hydrogen bonding from the side, interaction with the $2\pi^*$ orbitals of CO leads to a direct π-back-bonding effect that reduces the frequency of ν_{CO} by lengthening the CO bond. As either the O−H group of H_2O or the Nδ−H group of Im moves into the proximity of CO, the π-bonding interaction of this group tends to shorten the Fe−C bond leading to an increase in ν_{Fe-CO} frequency (see Figure 3).

Comparison of chemically different groups is often difficult because many factors change simultaneously when one atom is substituted for another. Nonetheless, there is a consistent trend for the hydrogen-bonding groups studied. The interaction of the hydrogen-bonding group depends on the electronegativity of the element X in the X−H donor. Figure 4 shows that there are two regimes. Near 100°, ν_{Fe-CO} and ν_{CO} have an inverse relationship, and near 180°, ν_{Fe-CO} and ν_{CO} both increase in parallel in the order $CH_4 > NH_3 > H_2O$. The parallel change in ν_{Fe-CO} and ν_{CO} frequency in this series arises from a change from π-type bonding at angles near 90° to σ-type bonding at angles near 180°. The crossover occurs near 140°.

The Correlation of Structure and Spectroscopy in A_0 and A_1 States. The multiple CO stretching bands observed in many myoglobins and hemoglobins can be explained by comparison of imidazole and imidazolium hydrogen bonding calculated by the DFT methods in this study. On the basis of the wild-type myoglobin X-ray crystal structure (2MGK), the Im−H distance of H64 is <2.1 Å, and there are simultaneous hydrogen bonds from V64 at 2.2 Å and F43 at 2.6 Å.[35] The model calculation for the two hydrogen-bond groups, CH_4 and Im (Table 2, row 30), has a calculated frequency of $\nu_{CO} = 1939$ cm^{-1}. The experimental values for the largest (A_1) substate of myoglobin are $\nu_{Fe-CO} \approx 507$ cm^{-1} and $\nu_{CO} \approx 1943$ cm^{-1}. The A_0 state is represented by X-ray crystal structures 1VXA, 1VXC, and 1VXF at pH 4.0, 5.0, and 6.0, respectively.[37] In these structures, H64 has swung out of the distal pocket to varying degrees and is nearer to the surface of the globin.[37] The pK_a of H64 is 6.0, which suggests that H64 is in the imidazolium form in the low pH structures.[68] The distance of the imidazolium ImH^+ from the CO is greater than that of Im in wild-type Mb.[37] The geometries in these structures were modeled in DFT calculations given in Table 2 (rows 33−36) for a fixed CH_4 distance representing V68 and a variable ImH^+ distance. The calculations in Table 2 show that the ν_{CO} frequency varies from 1897 to 1971 cm^{-1} as the ImH^+ hydrogen-bonding distance increases from 2.0 to 3.8 Å. The experimental frequencies for the A_0 substate are $\nu_{Fe-CO} \approx 493$ cm^{-1} and $\nu_{CO} \approx 1967$ cm^{-1}. The calculated ν_{Fe-CO} frequencies corresponding to A_0 and A_1 substates are 508 and 535 cm^{-1}, in agreement with the experimental trend that A_1 is larger than A_0. The DFT calculations are consistent with a weak interaction of imidazolium with bound CO. The best agreement is found for a CO··· H distance of 3.4 Å, which is long enough to no longer be considered a hydrogen bond. The corresponding 4.4 Å distance between Nδ of ImH^+ and the oxygen of CO is approximately 1.2 Å larger than that observed in the pH 5.0 X-ray structure. It is possible that solvation effects change the structure slightly as the pH is lowered or that experimental data are affected by the equilibrium between the protonated and unprotonated forms. In summary, the model gives a reasonable representation of the type of ν_{CO} frequency shift that is expected for a structural change associated with A_0 and A_1 conformers.

The Effect of Mutation on the ν_{CO} Frequency. The results of the VTA correlation (eqs 1−3) can be used to address specific hydrogen-bonding geometries in mutants of myoglobin. A test of the VTA model is carried out by geometry optimization of each of the structures considered as a model for the mutant in question. The geometries and Mulliken charges and ν_{CO} frequencies for three examples are given in Table 4. Mutations in the H64 position have a profound effect on the ν_{CO} stretching frequency. The calculations suggest that a combination of CH_4 and H_2O hydrogen bonding gives rise to a frequency in the range observed, that is, $\nu_{CO} \approx 1957$ cm^{-1} (Table 2, rows 31 and 32). This value of ν_{CO} is in the correct range for H_2O hydrogen

Table 4. Models for Myoglobin Mutants Based on the Vibronic Theory of Activation[a]

mutant/model	C–O	q_C	q_O	$q_C + q_O$	$q_C - q_O$	ν_{CO} (eq 1)	ν_{CO} (eq 2)
H64L CH_4, CH_4, CH_4	1.163	0.239	−0.282	−0.043	0.521	1967.5	1967.9
H64L/V68T CH_4, CH_4, OH_2	1.157	0.220	−0.242	−0.022	0.462	2017.9	2000.2
H64Q CH_4, CH_3CONH_2	1.164	0.225	−0.288	−0.063	0.513	1963.2	1967.3

[a] The distances and Mulliken charges obtained from a geometry optimization are used to calculate the ν_{CO} frequency for comparison with experiment.

bonding in H64A and H64G mutants, assuming V68 is also involved in hydrogen bonding to CO. Such a hypothesis has been advanced to explain the very broad distribution of ν_{CO} frequencies observed in the FTIR spectra of those mutants.[19] The substitution of an aliphatic residue in H64L results in an observed frequency of $\nu_{CO} \approx 1968$ cm^{-1}, which corresponds to a frequency shift of $\Delta\nu_{CO} \approx +23$ cm^{-1} relative to wild type.[19] As indicated in Table 1, the X-ray structural data do not indicate any single group that appears to replace H64 giving a strong hydrogen-bonding interaction with CO. For mutants such as H64L, it is possible that hydrogen bonding to multiple aliphatic groups gives rise to the observed frequency shift. A model with three CH_4 groups surrounding CO was geometry-optimized as a model for H64L (Table 4). As shown in Table 4, the model consisting of three hydrogen-bonding groups is in the correct range. A second example, the H64Q mutant, was modeled using methyl amide CH_3CONH_2 in the place of imidazole (Table 4). A geometry optimization alone was carried out for CH_4 and $CH_3(C{=}O)NH_2$ as a model for glutamine in place of the imidazole model for histidine. As shown in Table 4, using the VTA approach, good agreement with experiment is achieved for the model of the H64Q mutant. As a third example, the H64L/V68T double mutant gives a large shift up in ν_{CO} frequency. This interaction too was modeled by application of the VTA method using two CH_4 and one OH_2 interaction with bound CO, and reasonable agreement is found (Table 4).

Connection to Vibronic Theory. The vibronic theory of activation (VTA) can be related to eq 2.[87] According to the vibronic theory, the force constant K_0 is shifted by a term F^2/Δ, where F is the vibronic constant, $F = \langle 0|\partial V/\partial Q|1\rangle$, and Δ is the energy gap between the ground state $\langle 0|$ and the vibronic state $|1\rangle$. The effective force constant for the ground state is $K_0 - F^2/\Delta$.[87] The reduction in force constant leads to a softening of the mode as a result of vibronic coupling. The vibronic term, F, is equal to the sum of the changes in occupation numbers Δp_i times the orbital vibronic constant f_i.

$$F = \sum_i \Delta p_i f_i \tag{4}$$

The unperturbed frequency is $\nu_0 = \sqrt{K_0/\mu}$ (in cm^{-1}) within the harmonic approximation, and the shifted frequency, ν, is

$$\nu = \nu_0 \sqrt{1 - \sum_i \frac{\Delta p_i^2 f_i^2}{\nu_0^2 \Delta}} \tag{5}$$

We equate the change in the sum of the square of the occupation numbers with the absolute value of the change in Mulliken charge $\Delta p_i^2 = |q_C + q_O|$. The vibronic model becomes

$$\nu = \nu_0 \sqrt{1 - \frac{|q_C + q_O| f_{CO}^2}{\nu_0^2 \Delta}} \tag{6}$$

where f_{CO} is the vibronic matrix element for an iron CO charge transfer state. This model is equivalent to the empirical fit to a line (eq 2) because an expansion of eq 2 yields

$$\nu = \nu_0 \left\{ 1 - \frac{f_{CO}^2}{2\nu_0^2 \Delta} |q_C + q_O| \right\} \tag{7}$$

and the slope is $f_{CO}^2/2\nu_0\Delta$. Alternatively, eq 6 can be used to fit the data in Figure 6, which yields a value of $\nu_0 = 2091$ cm^{-1} and $f_{CO}^2/\nu_0^2\Delta = 1.4$. Using the parameters from the fit of the values in Figure 6 in eq 7, the frequency and slope in a linear fit are estimated to be 2091 and 1464 cm^{-1} as compared to 2093 and 1535 cm^{-1} obtained using eq 2. The value of χ^2 for the fits differs by only 1%.

Conclusion

The effect of hydrogen bonding on the vibrational frequency of CO in iron−porphine adducts has been compared to the effect of an applied electric field, the vibrational Stark effect.[4,57] The magnitude of the applied electric field is consistent with a correlation of frequency with bond length and Mulliken charge on bound CO obtained from a large number of DFT calculations that include many different hydrogen-bonding environments. The calculations have been interpreted with a VTA approach that unites the experimentally observed vibrational Stark effect and observed hydrogen-bonding shifts on ν_{CO}. A change in the C−O bond length occurs both for hydrogen bonding and in the presence of an applied electric field. A quantitative interpretation of the dependence of the bond length on an externally applied field provides a structural interpretation for the Stark tuning rate. The calculations are consistent with the observation that hydrogen bonding occurs mainly from the side at Fe−O···H angles of $\theta = 100$−$120°$. The π-back-bonding correlation is maintained for changes in distance at a fixed angle θ and to a certain extent even for chemical changes in the hydrogen-bonding group for the same geometry. The calculations provide further evidence that specific hydrogen-bonding interactions dominate the electrostatic effects observed in the vibrational spectra of heme proteins. In particular, the dominant effect on the ν_{CO} frequency in myoglobin arises from specific hydrogen-bonding effects.

Acknowledgment. S.F. acknowledges support through NSF grant MCB-9874895 and a grant from the North Carolina

(87) Bersuker, I. B.; Stavrov, S. S. *Coord. Chem. Rev.* **1988**, *88*, 1−68.
(88) Carlson, M. L.; Regan, R.; Elber, R.; Li, H.; Phillips, G. N.; Olson, J. S.; Gibson, Q. H. *Biochemistry* **1994**, *33*, 10597−10606.
(89) Kuriyan, J.; Wilz, S.; Karplus, M.; Petsko, G. A. *J. Mol. Biol.* **1986**, *192*, 133−154.

BATCH: ja11a34 USER: jld69 DIV: @xyv04/data1/CLS_pj/GRP_ja/JOB_i44/DIV_ja017708d DATE: September 17, 2002

Supercomputer Center (NCSC). Thanks is given to Lee Bartolotti of the NCSC for many helpful suggestions.

Supporting Information Available: Information on the basis set. Detail information on energies, frequencies, and Mulliken charge as a function of hydrogen-bond geometry (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

JA017708D