Journal of Molecular Modeling (2020) 26:281
https://doi.org/10.1007/s00894-020-04519-w

ORIGINAL PAPER

Critical assessment of the FeC and CO bond strength in carboxymyoglobin: a QM/MM local vibrational mode study

Marek Freindorf[1] · Elfi Kraka[1]

Received: 8 May 2020 / Accepted: 26 August 2020

Abstract

The interplay between FeC and CO bonding in carboxymyoglobin (MbCO) and the role of potential hydrogen bonding between the CO moiety and the side chains of the surrounding protein amino acids have been the subject of numerous experimental and theoretical studies. In this work, we present a quantitative measure for the intrinsic FeC and CO bond strength in MbCO, as well as for CO···H bonding, based on the local vibrational mode analysis, originally developed by Konkoli and Cremer. We investigated a gas phase model, two models of the wild-type protein, and 17 protein mutations that change the distal polarity of the heme pocket, as well as two protein mutations of the heme porphyrin ring. Based on local mode force constants, we could quantify for the first time the suggested inverse relationship between the CO and FeC bond strength, the strength of CO···H bonding, and how it weakens the CO bond. Combined with the natural orbital analysis, we could also confirm the key role of π back donation between Fe and the CO moiety in determining the FeC bond strength. We further clarified that CO and FeC normal modes couple with other protein motions in the protein environment. Therefore, normal mode frequencies/force constants are not suited as bond strength descriptors and instead their local mode counterparts should be used. Our comprehensive results provide new guidelines for the fine-tuning of existing and the design of MbCO models with specific FeC, CO, and CO···H bond strengths.

Keywords Vibrational spectroscopy · Myoglobin · Local vibrational theory · Local mode force constants · DFT · QM/MM

Introduction

The metallo-enzyme myoglobin (Mb) is responsible for storage of oxygen in vertebrates and plays an important role in many physiological functions of the heart and skeletal muscles [1–4]. The Mb active site involves a heme group with iron in an oxidized (Fe^{3+}, ferric) or reduced (Fe^{2+}, ferrous) form. As shown in Fig. 1, the heme group is fixed in the Mb active site via Fe-N bonding with the side chain of H93 (*proximal histidine*), while the opposite site of the heme pocket is occupied by H64 (*distal histidine*), leaving enough space for small ligands such as CO, $NO^{\cdot}$, and O_2 to enter and bind to the iron. These small molecular ligands play a key role in many biological processes via transferring signals for conformational changes of the Mb protein, resulting in diverse enzymatic activity [5]. One has monitored the interaction between the small ligands coordinated to the iron of the heme group and the active site pocket of Mb by vibrational spectroscopy, in particular by resonance Raman techniques [6] and time-resolved resonance Raman spectroscopy [7]. CO has been considered a useful vibrational spectroscopy probe because the CO absorption band can easily be identified in the spectrum. It has been assumed that the CO vibration does not strongly couple with other vibrational modes, and in this way directly reflects the interaction with Fe and the heme pocket, an important assumption, which has also laid the basis for using CO as a frequent *vibrational Stark effect* (VSE)

This paper belongs to Topical Collection XX - Brazilian Symposium of Theoretical Chemistry (SBQT2019)

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00894-020-04519-w) contains supplementary material, which is available to authorized users.

✉ Elfi Kraka
ekraka@smu.edu

Marek Freindorf
mfreindorf@smu.edu

[1] Computational and Theoretical Chemistry Group (CATCO), Department of Chemistry, Southern Methodist University, 3215 Daniel Avenue, Dallas, TX 75275-0314, USA

Published online: 24 September 2020

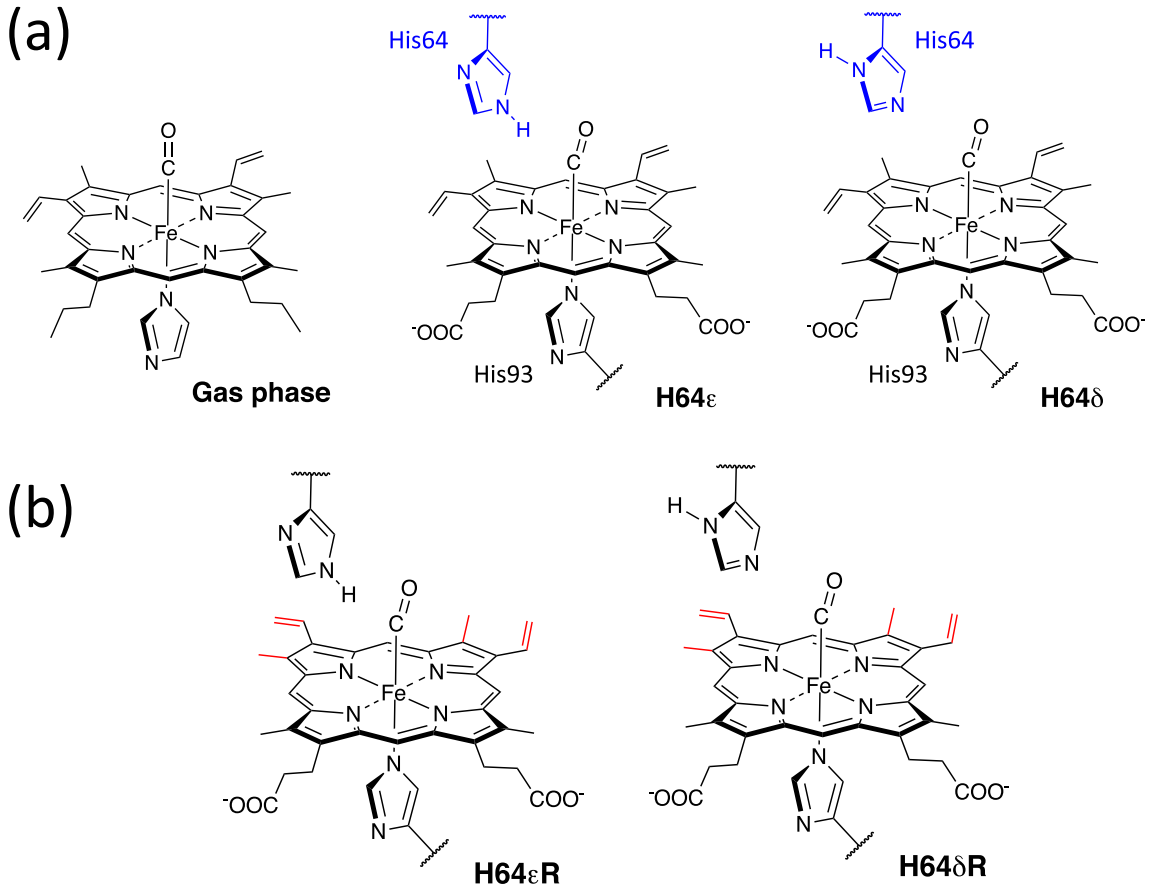

Fig. 1 **a** MbCO active site model in the gas phase (**Gas**), QM part of MbCO **H64ϵ** and **H64δ**; the distal His64 is shown in blue. **b** QM part of MbCO **H64ϵR** and **H64δR**; the different position of the heme vinyl groups in the R forms is marked in red. For a detailed description, see text. Snapshots of the active site pocket for all MbCO systems investigated in this work are collected in the electronic supplementary material, in Figs. S1–S4

probe [8–14]. Caused by the popularity of the CO ligand, carboxymyoglobin (MbCO) has been the subject of intense experimental [15–36] and theoretical [37–52] studies since the first Mb X-ray structure published in 1960 [2].

The IR absorbance spectrum of sperm whale MbCO [28] (PDB entry: 1A6G) shows four maxima, which were originally assigned to the CO stretching vibration of different conformational states, labeled as $A_0 - A_3$; A_0 (1965 cm^{-1}), A_1 (1947 cm^{-1}), A_2 (1942 cm^{-1}), and A_3 (1932 cm^{-1}) [53–55]. At a neutral or alkaline pH, the most dominant band was found at 1945 cm^{-1} being identified as a combination of the two A_1 and A_2 conformers [53, 54]. It has been suggested that the observed differences in the CO absorptions originate from changes in the electrostatic interaction between the CO coordinated to Fe and the heme pocket, which adapts a somewhat different electronic structure for each of the different conformational states [20, 21, 29]. The A_0 state was assigned to a structure where the protonated imidazole side chain of H64 is moved from the distal heme pocket into the surrounding water solvent. The A_1 and A_3 states were characterized by a neutral imidazole side chain inside the heme pocket. Exchange kinetics between the A_1 and A_3 states suggested that they originate from two slightly different conformations of H64 in the distal pocket, where A_1 corresponds to a weaker interaction between H64 and CO, while A_3 corresponds to a stronger interaction [17, 18, 32, 33, 54, 56]. The situation has been complicated by the fact that the distal histidine H64 can form two tautomers (ϵ and δ) leading to a different interaction with the CO molecule coordinated to iron, as shown in Fig. 1. So far, the relationship between the A conformational states and the H64 tautomer remains controversial, mainly because of the fact that it is still unclear which tautomer of H64 is more stable [17, 18, 20, 23, 26, 28, 29, 37–39, 42, 51].

Other studies [40, 43, 52, 57] pointed out that there is obviously an inverse correlation between the FeC and CO stretching frequencies observed in MbCO and its mutations, suggesting that strengthening of the FeC bond leads to weakening of the CO bond. It has been argued that this inverse correlation originates from π back donation from the occupied

d_π orbitals of iron into the vacant CO π^* orbitals. The resulting electron transfer into the CO π^* orbitals weakens the CO bond and strengthens the FeC bond [52, 58, 59]. However, to quantify this important reverse relation for the wild-type Mb and its protein mutations, one needs a quantitative measure of the intrinsic strength of a chemical bond. The local mode theory, originally developed by Konkoli and Cremer [60–64], provides this highly sought-after bond strength measure based on vibrational spectroscopy, which was applied in this work.

It has been known for a long time that the Mb environment lowers the binding affinity of CO to the heme group compared to O_2, whereas in solution CO binds to the free heme stronger than O_2 [65]. One possible explanation was based on different electrostatic interactions between the ligand and the distal site of the heme pocket [66], resulting in a stronger hydrogen bond between **H64ϵ** and O_2 compared to the CO ligand. Some theoretical studies argued that the difference in the hydrogen bond energies of H···O_2 and H···OC leads to the discrimination between the O_2 and CO ligands by the protein [39, 44]. These findings are in line with neutron scattering data suggesting the existence of a H···O_2 bond in the oxymyoglobin (MbO_2) complex [67], whereas no H···OC hydrogen bond was observed in the MbCO complex [68]. So far, the potential role of hydrogen bonding as a stabilizing factor of **H64ϵ** over **H64δ** has remained unclear.

To shed light into the currently open questions in a systematic way, our study was primarily aimed at the following:

1. To derive intrinsic FeC and CO bond strength measures based on local vibrational force constants and related bond strength orders BSO *n*.
2. To quantify the suggested inverse relationship between FeC and CO bond strength based on local vibrational force constants.
3. To further explore the postulated relationship between π back donation and FeC bond strength.
4. To clarify the role of hydrogen bonding on the CO bond strength.
5. To assess to what extent the FeC and CO vibrational modes in these proteins are localized, which determines their suitability as bond strength measures.

As a test set, we chose 20 different MbCO systems; the MbCO active site model in the gas phase (**Gas**), the MbCO wild-type ϵ and δ tautomers (see Fig. 1a), complemented with 17 protein mutations involving the replacement of distal His64 and/or other distal amino acids with amino acids that change the distal polarity, and two protein mutations involving a reversed orientation of the vinyl groups in the heme porphyrin ring (Fig. 1b). The selected protein mutations were based on experimental spectroscopy data being available in the literature [17, 21, 22, 25, 33, 35, 69]. The nomenclature of these mutations was adapted from the literature: **H64A**, **H64G**, **H64I**, **H64L**, **H64M**, **H64Q**, and **H64V** are MbCO protein mutations where the distal histidine H64 is mutated; **F43W**, **F43V**, **F46V**, **L29F**, **V68F**, and **V68N** are MbCO protein mutations where other amino acids of the distal pocket are exchanged while keeping H64 in the ϵ tautomeric form [18, 29, 41]; **H64QL29F** and **H64LV68F** are MbCO double protein mutations involving mutations of H64; and **H64ϵR** and **H64δR** are MbCO mutations with a reversed orientation of the vinyl groups [35]. Table S1 of the electronic supplementary material provides further details about the protein mutations. While in this work we focused predominantly on distal modifications, in future work, we will investigate the influence of proximal modifications as well.

As a suitable measure of the intrinsic bond strength, local mode force constants were derived from the local vibrational mode theory originally developed by Konkoli and Cremer [60–64]. A short introduction into the local mode theory is given in the following computational section. Local mode force constants were complemented with atomic charges of the Fe, C, O, and N_{H93} atoms, FeCO bond angles, the population of the CO π^* molecular orbitals, and electron density features such as the energy density at the FeC and CO bond critical points determining the covalent bond character.

Computational methods

Local vibrational mode theory

The normal vibrational modes of a molecule are generally delocalized over parts or even over the whole molecule [70–72]. Therefore, normal mode vibrational frequencies or the corresponding normal mode force constants are ineligible as bond strength measures [73–75]. A molecular vibration can be described via the Euler-Lagrange equations:

$$L\left(\mathbf{x},\dot{\mathbf{x}}\right) = T\left(\dot{\mathbf{x}}\right) - V(\mathbf{x}) \tag{1}$$

$$= \frac{1}{2}\dot{\mathbf{x}}^{\dagger}\mathbf{M}\dot{\mathbf{x}} - \frac{1}{2}\mathbf{x}^{\dagger}\mathbf{F}^{x}\mathbf{x} \tag{2}$$

where the Lagrangian $L\left(\mathbf{x},\dot{\mathbf{x}}\right)$ is defined as the difference of the kinetic energy $T\left(\dot{\mathbf{x}}\right)$ and the potential energy $V(\mathbf{x})$ of the vibrational mode.

Vector $\mathbf{x}$ contains the *3N* Cartesian displacement coordinates of the molecule being composed of *N* atoms describing the deviation of the atoms from their equilibrium position during the vibration, and $\dot{\mathbf{x}}$ denotes differentiation with respect to time. $\mathbf{F}^x$ is the force constant matrix (Hessian) in Cartesian coordinates $\mathbf{x}$. $\mathbf{M}$ is the diagonal mass matrix containing the

atomic masses. Both square matrices have the dimension $(3N \times 3N)$.

The Lagrangian can be also expressed in terms of $N_{vib} = (3N - K)$ internal coordinates $\mathbf{q}$ (K number of translations and rotations; 6 for nonlinear and 5 for linear molecules) [70]:

$$L\left(\mathbf{q}, \dot{\mathbf{q}}\right) = T\left(\dot{\mathbf{q}}\right) - V(\mathbf{q}) \tag{3}$$

$$= \frac{1}{2}\dot{\mathbf{q}}^{\dagger}\mathbf{G}^{-1}\dot{\mathbf{q}} - \frac{1}{2}\mathbf{q}^{\dagger}\mathbf{F}^{q}\mathbf{q} \tag{4}$$

$\mathbf{F}^q$ is the force constant matrix in internal coordinates $\mathbf{q}$; $\mathbf{G}$ is the Wilson $\mathbf{G}$ matrix, also called *inverse kinetic energy* matrix [70]. Both $\mathbf{F}^q$ and $\mathbf{G}$ are square matrices with the dimension $N_{vib} \times N_{vib}$.

The relationship between internal and Cartesian coordinates is provided by the Wilson $\mathbf{B}$ matrix [70], a rectangular $(N_{vib} \times 3N)$ matrix containing the first derivatives of the internal coordinates $q_n (n = 1, 2, 3 \ldots N_{vib})$ with respect to the Cartesian coordinates $x_i (i = 1, 2, 3 \ldots 3N)$:

$$\mathbf{B}_n = \frac{\delta q_n(\boldsymbol{x})}{\delta x_i} \tag{5}$$

The $\mathbf{B}$ matrix also connects the $\mathbf{M}$ and $\mathbf{G}$ matrices via the following relationship [70, 76]:

$$\mathbf{G} = \mathbf{B}\mathbf{M}^{-1}\mathbf{B}^{\dagger} \tag{6}$$

Equation (1) reveals that there are two different coupling mechanisms between the vibrational modes, *mass coupling* due to the off-diagonal elements of the Wilson $\mathbf{G}$ matrix, reflecting pairwise kinetic coupling between the internal coordinates and *electronic coupling* due to the off-diagonal elements of the $\mathbf{F}^q$ matrix [70].

The *electronic coupling* can be eliminated via the Wilson GF-matrix formalism applied to Eq. (1) [70–72, 76], i.e., solving the vibrational secular equation:

$$\mathbf{F}^{x}\mathbf{L} = \mathbf{M}\mathbf{L}\boldsymbol{\Lambda} \tag{7}$$

$\boldsymbol{\Lambda}$ is a diagonal matrix with the eigenvalues λ_μ, which leads to the N_{vib} (harmonic) vibrational frequencies ω_μ according to $\lambda_\mu = 4\pi^2 c^2 \omega_\mu^2$, and $\mathbf{L}$ collects the vibrational eigenvectors $\mathbf{l}_\mu$ in its columns, with the following properties:

$$\mathbf{L}^{\dagger}\mathbf{F}^{x}\mathbf{L} = \mathbf{F}^{\mathbf{Q}} = \mathbf{K} \tag{8}$$

$$\mathbf{L}^{\dagger}\mathbf{M}\mathbf{L} = \mathbf{M}^{R} \tag{9}$$

Equations (8) and (9) define the diagonal normal force constant matrix $\mathbf{F}^{\mathbf{Q}} = \mathbf{K}$ and the reduced mass matrix $\mathbf{M}^R$ (with elements m_μ^R), respectively. $\mathbf{Q}$ is a vector that collects the N_{vib} normal coordinates [77]. In essence, diagonalization of the force constant matrix $\mathbf{F}^x$, i.e., transforming to normal coordinates $\mathbf{Q}$ [78–80], eliminates the *electronic coupling* [70] a procedure which is routinely performed in standard quantum chemistry packages during a harmonic vibrational frequency calculation.

However, the transformation to normal coordinates $\mathbf{Q}$ does not eliminate the *mass-coupling*. Konkoli and Cremer eliminated the remaining mass coupling via mass-decoupled Euler-Lagrange equation [60–64]:

$$p_i = \frac{\delta L\left(\mathbf{q}, \dot{\mathbf{q}}\right)}{\delta \dot{q}_i} \neq 0, \quad \dot{p}_i = \frac{\delta V(\mathbf{q})}{\delta q_i} \neq 0, \tag{10}$$

$$p_j = \frac{\delta L\left(\mathbf{q}, \dot{\mathbf{q}}\right)}{\delta \dot{q}_j} = 0, \quad \dot{p}_j = \frac{\delta V(\mathbf{q})}{\delta q_j} = 0, \forall j \neq i$$

which defines a one-dimensional subspace in the full vibrational space for each internal coordinate q_i, i.e., only the masses of the atomic fragment involved in the vibration described by internal coordinate q_i are non-zero. This led for the first time to mass-decoupled local vibrational modes $\mathbf{a}_i$ directly derived from normal vibrational modes $\mathbf{d}_\mu$ and the force constant matrix $\mathbf{K}$:

$$\mathbf{a}_i = \frac{\mathbf{K}^{-1}\mathbf{d}_i^{\dagger}}{\mathbf{d}_i\,\mathbf{K}^{-1}\,\mathbf{d}_i^{\dagger}} \tag{11}$$

To each local mode $\mathbf{a}_i$, a corresponding local mode frequency ω_i^a, local mode mass G_{ii}^a, and a local force constant k_i^a can be defined [60–63], The local mode frequency ω_i^a is defined by:

$$\left(\omega_i^a\right)^2 = \frac{1}{4\pi^2 c^2}\, k_i^a\, G_{ii}^a \tag{12}$$

and the corresponding local mode force constant k_i^a by:

$$k_i^a = \mathbf{a}_i^{\dagger}\,\mathbf{K}\,\mathbf{a}_i = \left(\mathbf{d}_i\,\mathbf{K}^{-1}\,\mathbf{d}_i^{\dagger}\right)^{-1} \tag{13}$$

As shown by Zou and Cremer [81], local mode stretching force constants can be directly associated with the intrinsic strength of a chemical bond and/or weak chemical interaction. Based on their landmark paper, local vibrational stretching force constants have been successfully applied to assess the strength of strong and weak covalent bonds [81–88], to investigate weak chemical interactions including hydrogen, halogen, pnicogen, chalcogen and tetrel bonding [83, 89–109], to derive a new description of metal-ligand bonding, [96, 100, 110–113] and [114] and a new aromaticity index [110, 115–117]. There exists a 1:1 relationship between a complete set of non-redundant local modes and the normal modes via an *adiabatic connection scheme*, allowing a smooth transition from local to normal modes [117]. This forms the basis of the characterization of the normal mode (CNM) procedure [63]. CNM decomposes each normal mode into local mode

contributions offering a new, comprehensive way of analyzing IR and Raman spectra, which we successfully applied to assess the usefulness of vibrational Stark effect probes [118]. Recently, the local mode theory was extended for the description of chemical bonding in periodic systems [119, 120].

When comparing a large set of k^a values, the use of a relative bond strength order (BSO) n is convenient. Both are connected according to the generalized Badger rule derived by Cremer, Kraka, and co-workers [73, 74], via the following power relationship:

$$BSO\ n = A(k^a)^B \tag{14}$$

The constants A and B are fully determined by two reference compounds with known BSO n and k^a values and the requirement that for a zero-force constant BSO n becomes 0. In this work, we used as references for the FeC bonds the CuC bond in $CuCH_3$ as a bond close to a single bond and the NiC bond in $NiCH_2$ as a bond close to a double bond (details are given in the supplementary material). To quantify the single- or double-bond character, the Mayer bond orders [121–123] for these molecules were calculated to be $n(Mayer, CuC) = 1.0048$ and $n(Mayer, NiC) = 1.8937$, which corresponds to a ratio of 1.00:1.8846. The corresponding constants A and B are 0.248 and 1.184. For the CO bonds, we used the CO single bond in CH_3OH with BSO $n = 1$ and the CO double bond in CH_2O with BSO $n = 2$ [93], leading to A and B values of 0.309 and 0.704, respectively. For the hydrogen bonds, we used the HF bond of FH with BSO $n = 1$ and the HF bond in the $[H\cdots F\cdots H]^-$ anion with BSO $n = 0.5$ as references [93], leading to A and B values of 0.451 and 0.348, respectively.

We determined the covalent character of the FeC, CO, and the hydrogen bonds via the Cremer-Kraka criterion [124, 125] for covalent bonding which is based on the local energy density $H(\mathbf{r})$:

$$H(\mathbf{r}) = G(\mathbf{r}) + V(\mathbf{r}) \tag{15}$$

where $G(\mathbf{r})$ is kinetic energy density (positive, destabilizing) and $V(\mathbf{r})$ is potential energy density (negative, stabilizing). Taken at the bond critical point $\mathbf{r}_b$ of the electron density $\rho(\mathbf{r})$ [126, 127] between two bonded atoms, $H(\mathbf{r}_b) < 0$ indicates a covalent bond, while $H(\mathbf{r}_b) > 0$ indicates an electrostatic interaction.

Computational details

All gas phase calculations were performed with the PBE0 functional [128] and Pople's 6-31G(d,p) basis set [129]. The QM/MM calculations were performed at the PBE0/6-31G(d,p)/AMBER level of theory using the ONIOM methodology with electronic embedding [130] and the AMBER force field for the MM part [131]. The protein geometries were based on the X-ray structure of sperm whale MbCO [28] (PDB entry: 1A6G). The QM part was composed of all atoms of the heme group and side chain atoms of both the distal H64 and the proximal H93 histidine (average number of QM atoms, 100). In the case of **H64A**, **H64G**, **H64I**, **H64L**, **H64M**, **H64Q**, **H64V**, **H64QL29F**, and **H64LV68F** with a replacement of H64, the QM part included the side chain atoms of the corresponding replacement amino acid. All remaining protein atoms were included in the MM part (average number of MM atoms, 2960). Additional technical details of the QM/MM calculations are provided in the electronic supplementary material. For both gas phase and all protein molecules, full geometry optimizations were performed followed by harmonic frequency calculations and subsequently by local mode analyses. All vibrational frequency calculations were completed without imaginary frequencies, confirming all systems as minima on the potential energy surface. The DFT calculations in the gas phase and in the protein were carried out with Gaussian16 [132]. The local mode analysis was performed with COLOGNE16 [133]. Natural bond orbital (NBO) charges and CO π^* populations were calculated with the program NBO 6 [134]. The energy densities at bond critical points were calculated using the AIMALL program [135].

Results and discussion

In Table 1, calculated FeC and CO distances d, local mode force constants k^a and corresponding bond strength orders BSO n, FeCO bond angles, and energy densities $H(\mathbf{r}_b)$, as well as NBO charges for Fe, C, O, N_{H93}, and CO π^* populations for the 20 compounds investigated in this work are collected. In Table 3, calculated local mode stretching frequencies ω^a(FeC) and ω^a(CO) and the experimental normal mode frequencies ω_{exp} (being identified by isotope substitutions as FeC and CO stretching frequencies) for 14 complexes with available experimental data are shown. Table 2 contains hydrogen bond (HB) distances d, local mode force constants k^a(HB) and corresponding BSO n values, NBO charges of O and H atoms, and energy densities $H(\mathbf{r}_b)$ for the hydrogen bonds formed between the CO oxygen and the side chain His64ϵ hydrogen of the MbCO wild type and those of the corresponding mutations. Figures S1–S4 of the electronic supplementary material provide active site snapshots of all QM/MM optimized protein structures.

In the following, we will focus first on the local mode discussion of the FeC and CO bonds, followed by the assessment of experimental normal mode frequencies as bond strength descriptors. Then, we will discuss hydrogen bonding.

Table 1 Calculated FeC and CO distances d, local mode force constants k^a, corresponding bond strength orders BSO n, FeCO bond angles α, and energy densities $H(\mathbf{r}_b)$; NBO charges for Fe, C, O, N_{H93}, and CO π^* populations. Gas phase calculations at the PBE0/6-31G(d,p) and protein calculations at the PBE0/6-31G(d,p)/AMBER levels of theory

Model	d (Å)		k^a (mDyn/Å)		BSO n		(α°)	$H(\mathbf{r}_b)$ (Hartree/Å^3)		NBO charge (e)				π^* (e)
	FeC	CO	FeC	CO	FeC	CO	FeCO	FeC	CO	Fe	C	O	N_{H93}	CO
Gas phase	1.760	1.148	2.654	18.209	0.786	2.388	180.00	− 0.3617	− 5.2273	0.666	0.600	− 0.464	− 0.491	0.3940
H64ϵ	1.728	1.156	3.023	16.977	0.918	2.273	173.76	− 0.4326	− 5.1024	0.678	0.599	− 0.518	− 0.513	0.4665
H64δ	1.759	1.147	2.580	18.176	0.761	2.384	171.78	− 0.3833	− 5.2414	0.691	0.601	− 0.454	− 0.508	0.3838
H64A	1.744	1.150	2.824	17.926	0.846	2.361	173.42	− 0.4049	− 5.1996	0.676	0.598	− 0.475	− 0.506	0.4173
H64G	1.746	1.150	2.773	17.915	0.828	2.360	171.63	− 0.4056	− 5.1949	0.679	0.596	− 0.477	− 0.509	0.4181
H64I	1.750	1.150	2.748	17.870	0.820	2.356	171.19	− 0.3995	− 5.1902	0.679	0.593	− 0.479	− 0.508	0.4181
H64L	1.745	1.150	2.807	17.914	0.840	2.360	174.14	− 0.4015	− 5.2016	0.680	0.596	− 0.476	− 0.510	0.4185
H64M	1.749	1.150	2.769	17.888	0.827	2.358	173.27	− 0.3975	− 5.1989	0.678	0.594	− 0.470	− 0.510	0.4117
H64Q	1.735	1.154	2.931	17.288	0.885	2.302	173.64	− 0.4224	− 5.1422	0.679	0.595	− 0.498	− 0.513	0.4489
H64V	1.749	1.150	2.783	17.888	0.832	2.358	173.48	− 0.3955	− 5.1922	0.677	0.595	− 0.478	− 0.510	0.4170
F43W	1.746	1.151	2.797	17.807	0.837	2.350	170.57	− 0.4103	− 5.1848	0.676	0.585	− 0.478	− 0.509	0.4247
F43V	1.734	1.156	2.934	17.012	0.886	2.276	172.00	− 0.4278	− 5.1078	0.683	0.596	− 0.512	− 0.512	0.4595
F46V	1.722	1.156	3.126	17.064	0.955	2.281	176.14	− 0.4420	− 5.0984	0.668	0.616	− 0.529	− 0.509	0.4758
L29F	1.722	1.156	3.178	17.109	0.973	2.285	172.77	− 0.4488	− 5.0990	0.671	0.610	− 0.531	− 0.511	0.4700
V68F	1.730	1.156	3.003	16.947	0.910	2.270	178.01	− 0.4211	− 5.1017	0.674	0.598	− 0.517	− 0.512	0.4702
V68N	1.721	1.159	3.131	16.608	0.956	2.238	177.59	− 0.4386	− 5.0626	0.672	0.596	− 0.549	− 0.513	0.5025
H64QL29F	1.730	1.154	3.050	17.372	0.927	2.310	171.78	− 0.4346	− 5.1395	0.676	0.601	− 0.507	− 0.512	0.4484
H64LV68F	1.750	1.150	2.770	17.860	0.827	2.355	178.85	− 0.3833	− 5.1949	0.676	0.596	− 0.476	− 0.510	0.4158
H64ϵR	1.723	1.157	3.147	16.974	0.962	2.272	179.17	− 0.4366	− 5.0943	0.674	0.608	− 0.523	− 0.511	0.4932
H64δR	1.742	1.150	2.864	17.871	0.861	2.356	178.63	− 0.4015	− 5.1976	0.676	0.600	− 0.475	− 0.506	0.4176
$Fe(CO)_5$(ax)	1.791	1.144	3.006	18.731	0.911	2.436	180.00	− 0.3274	− 5.2885	− 0.842	0.614	− 0.416	–	0.4188
$Fe(CO)_5$(eq)	1.777	1.148	3.025	18.164	0.918	2.383	180.00	− 0.4053	− 5.2165	− 0.842	0.583	− 0.434	–	0.4433
CO	–	1.135	–	20.264	–	2.574	–	–	− 5.2227	–	0.508	− 0.508	–	0.0

Table 2 Bond lengths d, local force constants k^a, bond strength orders BSO n, NBO charges of O and H atoms, energy densities $H(\mathbf{r}_b)$ of the NH···OC hydrogen bond between CO and the imidazole hydrogen of distal His64 of **H64ϵ**, and its mutations with possible hydrogen bonding (PBE0/6-31G(d,p)/AMBER level of theory). For comparison, the hydrogen bonds of the water dimer and the ammonia-formaldehyde dimer are included (gas phase; PBE0/6-31G(d,p) level of theory)

Model	d	k^a	BSO n	NBO charge		$H(\mathbf{r}_b)$
	(Å)	(mDyn/Å)		O/N(e)	H(e)	(Hartree/Å^3)
H64ϵ	2.078	0.109	0.209	− 0.518	0.464	0.0013
H64Q	2.444	0.039	0.146	− 0.498	0.431	0.0047
F43V	2.125	0.098	0.201	− 0.512	0.464	0.0020
F46V	2.082	0.072	0.181	− 0.529	0.470	0.0040
L29F	1.980	0.192	0.254	− 0.531	0.468	0.0034
V68F	2.130	0.080	0.187	− 0.517	0.464	0.0027
V68N	2.197	0.050	0.159	− 0.549	0.464	0.0047
H64QL29F	2.180	0.114	0.212	− 0.507	0.434	0.0020
H64ϵR	2.013	0.145	0.230	− 0.523	0.467	0.0007
H_2O···HOH	1.906	0.214	0.264	− 0.960	0.486	− 0.0088
H_2NH···OCH_2	2.224	0.079	0.187	− 0.527	0.397	− 0.0013

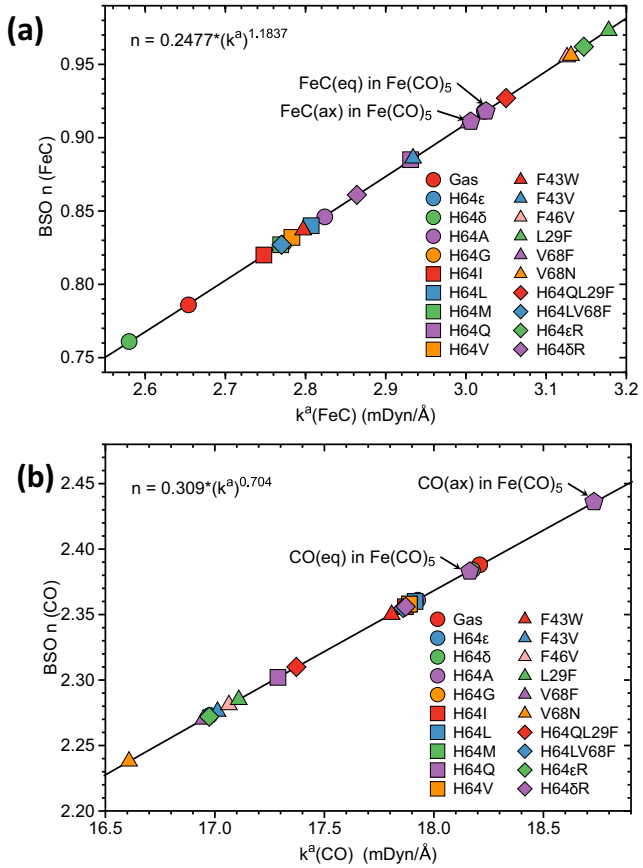

Fig. 2 **a** Bond strength order BSO *n* for FeC bonds calculated from local mode force constants k^a(FeC) via Eq. (14). **b** Bond strength order BSO *n* for CO bonds calculated from local mode force constants k^a(CO)via Eq. (14). Gas phase calculations at the PBE0/6-31G(d,p) and protein calculations at the PBE0/6-31G(d,p)/AMBER levels of theory

Analysis of FeC and CO bonds

In Fig. 2, the BSO *n* values for FeC and CO bonds are shown, calculated from the corresponding local mode force constants k^a via Eq. (14). For comparison, the data of the corresponding FeC and CO bonds of $Fe(CO)_5$ (D_{3h} symmetry) are added. The BSO *n* values of the FeC bonds range from 0.761 to 0.973 (see Fig. 2a and Table 1). Weakest FeC bonds are found for wild-type **H64*δ*** (BSO $n = 0.761$) and the gas phase model (BSO $n = 0.786$), whereas the strongest FeC bond is encountered for **L29F** (BSO $n = 0.973$), followed by **H64*ε*R** (BSO $n = 0.962$), **V68N** (BSO $n = 0.956$), and **F46V** (BSO $n = 0.955$).

The averaged FeC bond strength in $Fe(CO)_5$ (BSO $n = 0.915$) is close to that in the wild-type **H64*ε*** (BSO $n = 0.918$).

As revealed from the data in Table 1 and from Fig. 2, modification of the porphyrin frame as realized in **H64*ε*R** versus **H64*ε*** and **H64*δ*R** versus **H64*δ*** with reversed peripheral vinyl groups has a larger impact on FeC bond strength. Resonance Raman studies observed substantial changes in the low-frequency region for these modifications; however, their impact on the FeC vibration could not be quantified [35]. For both **H64*ε*R** and **H64*δ*R**, the FeC bond strength increases compared to the native orientation (**H64*ε*R**: BSO $n = 0.962$, **H64*ε***: BSO $n = 0.918$; **H64*δ*R**: BSO $n = 0.861$, **H64*δ***: BSO $n = 0.761$). However, there is no effect on the FeC bond strength for the same modification of the porphyrin frame in the gas phase model (see Table S2 of the electronic supplementary material). This strongly suggests that the increase of the FeC bond strength observed in **H64*ε*R** and **H64*δ*R** is mediated via the protein environment.

The BSO *n* values of the CO bonds range from 2.238 to 2.388 (see Fig. 2b and Table 1). The weakest CO bond is found for **V68N** (BSO $n = 2.238$) and the compounds with the strongest CO bonds are the **Gas** model (BSO $n = 2.388$) and wild-type **H64*δ*** (BSO $n = 2.384$), which are both weaker than the CO bonds in $Fe(CO)_5$ (BSO $n = 2.410$, average value). In comparison, the BSO *n* value of the CO molecule is 2.574 (see Table 1). This value should not be confused with the conceptual molecular orbital description of the CO bond as a triple bond, resulting from formally subtracting the number of occupied antibonding MOs from the number of occupied bonding MOs leading for CO to a value of 3. The conceptual description of the carbon monoxide bond does not consider the magnitude of overlap and the bond polarity in any quantitative sense and therefore is of limited value. A bond strength measure based on vibrational spectroscopy is more relevant [84]. The chemical bonding of CO in transition metal complexes, usually described in terms of the Dewar-Chatt-Duncanson model [136–139], suggests a synergetic interplay between donation from the 7σ homo of CO into an empty σ metal orbital and a back donation from an occupied d(π) orbital into the empty π^* orbital of CO leading to a weakening of the CO bond strength, observed via red shift of the CO vibrational stretching frequency [140–143]. A prominent example is Tolman's electronic parameter (TEP) [144, 145] describing the transition metal ligand strength in metal-carbonyl complexes indirectly via a redshift of the A_1 symmetric CO stretching frequency, which recently has been replaced by the metal ligand electronic parameter (MLEP) of Cremer, Kraka, and co-workers, a direct metal ligand bond strength measure being based on local vibration modes [111–113].

Two points are noteworthy; (i) the two strongest CO bonds belong to compounds which cannot form a hydrogen bond with the distal protein residue, whereas the opposite holds for the compounds with the weakest CO bonds, i.e., H···OC hydrogen bonding is one factor weakening the CO bond; (ii) comparing Fig. 2a and b it also becomes already apparent that stronger FeC bonds are related to weaker CO bonds and vice versa.

In Fig. 3a, local mode force constants k^a(FeC) are compared with the FeC distances and in Fig. 3b local mode force constants k^a(CO) with the corresponding CO distances. In both cases, we find a significant correlation (FeC bonds: R^2

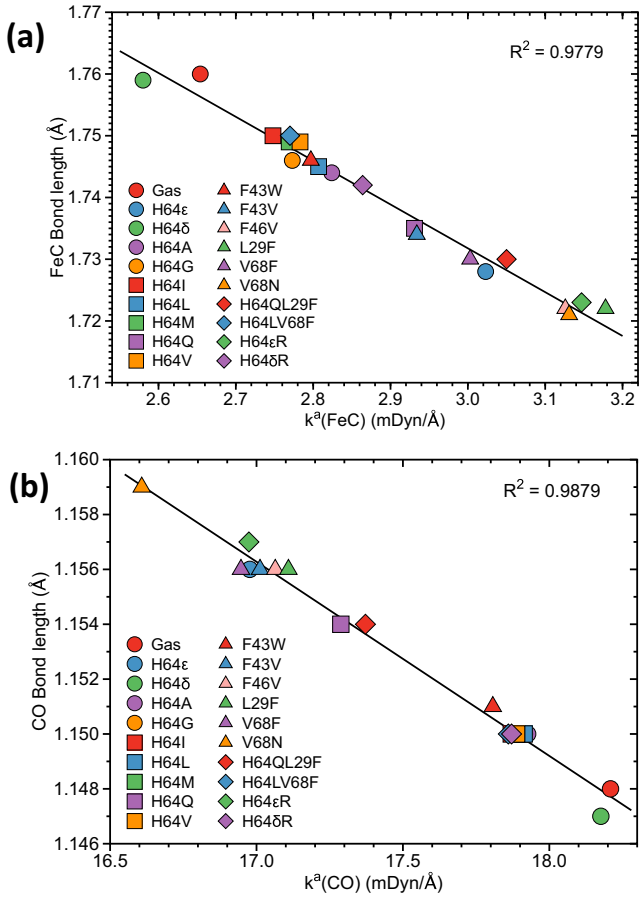

Fig. 3 **a** FeC bond lengths versus local mode force constants k^a(FeC). **b** CO bond lengths versus local mode force constants k^a(CO). Gas phase calculations at the PBE0/6-31G(d,p) and protein calculations at the PBE0/6-31G(d,p)/AMBER levels of theory

= 0.9779, CO bonds: $R^2 = 0.9879$) confirming that the generalized Badger rule of Kraka, Larsson, and Cremer [74] applies to the 20 compounds investigated in this work. The generalized Badger rule extends Badger's original rule [146] predicting an inverse power relationship between the bond length and the stretching force constant for diatomic molecules to polyatomic compounds by replacing normal mode force constants with their local mode counterparts. A caveat is appropriate; it is not always true that the stronger bond is the shorter bond [82, 147]; however, for our series with only marginal bond length changes (FeC bonds, 0.039 Å (2%); CO bonds, 0.012 Å (1%)) reflecting the space confinement enforced by the Mb pocket, it is valid. As revealed by the data in Table 1, the longest and weakest FeC bonds are found for the gas phase model **Gas** ($d = 1.760$ Å) and wild-type **H64*δ*** ($d = 1.759$ Å), whereas **V68N** ($d = 1.721$ Å), **F46V** ($d = 1.722$ Å), **L29F** ($d = 1.722$ Å), and **H64εR** ($d = 1.723$ Å) possess the shortest and strongest FeC bonds, belonging to the class of compounds which can form a hydrogen bond with the distal protein residue. The axial FeC bond in $Fe(CO)_5$ (**$Fe(CO)_5$(ax)**, $d = 1.791$ Å) is 0.031 Å longer than the gas phase value, suggesting that not only space confinement but also electronic effects influence the FeC bond distance. Figure 3b corroborates an inverse relationship between FeC and CO bonds. The longest and weakest CO bond is found for **V68N** ($d = 1.159$ Å) and the shortest and strongest for the gas phase model **Gas** ($d = 1.148$ Å) and wild-type **H64*δ*** ($d = 1.147$ Å), which is in the same range of the CO bonds in $Fe(CO)_5$ (**$Fe(CO)_5$(ax)**, $d = 1.144$ Å; **$Fe(CO)_5$(eq)**, $d = 1.148$ Å), however 0.012 Å longer than the CO bond in carbon monoxide (see Table 1).

In order to connect the description of bonding via the local force constant, a potential energy-related property with an electron density-related property was correlated with local force constants k^a with energy densities $H(\mathbf{r}_b)$. Figure 4a shows the correlation between the energy density $H(\mathbf{r}_b)$(FeC) and the local force constant k^a(FeC) and Fig. 4b the corresponding correlation for the CO bonds. As shown in Fig. 4a, stronger FeC bonds trend toward having more covalent character than their weaker counterparts. However, there is some scattering, as is reflected by the R^2 value of 0.8796. The largest covalent character is found for **L29F** ($H(\mathbf{r}_b) = -0.4488$ Hartree/A^3), **F46V** ($H(\mathbf{r}_b) = -0.4420$ Hartree/A^3),

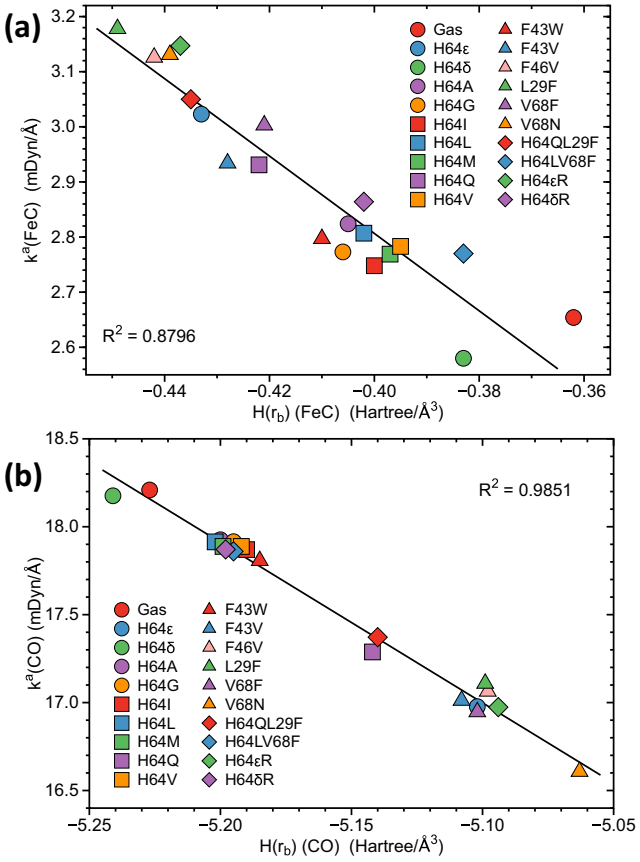

Fig. 4 **a** Energy density $H(\mathbf{r}_b)$(FeC) versus k^a(FeC). **b** Energy density $H(\mathbf{r}_b)$(CO) versus k^a(CO). Gas phase calculations at the PBE0/6-31G(d,p) and protein calculations at the PBE0/6-31G(d,p)/AMBER levels of theory

V68N ($H(\mathbf{r}_b) = -0.4386$ Hartree/A^3), and **H64ϵR** ($H(\mathbf{r}_b) = -0.4366$ Hartree/A^3). This is completely in line with the BSO *n* values (see Fig. 2) and the FeC distances (see Fig. 3). The weaker FeC bonds show some irregularities as depicted in Fig. 4a. In contrast, we find a significant correlation (R^2 = 0.9851) for the CO bonds; stronger CO bonds are related to smaller $H(\mathbf{r}_b)$ values, i.e., the covalent character of the CO bonds increases with increasing bond strength. We found the smallest $H(\mathbf{r}_b)$(CO) values for the gas phase model **Gas** (−5.2273 Hartree/A^3) and wild-type **H64δ** (−5.2414 Hartree/A^3) (see Table 1). These values are in good agreement with $H(\mathbf{r}_b)$ of carbon monoxide (−5.2227 Hartree/A^3) and the $H(\mathbf{r}_b)$ values of the two CO bonds in $Fe(CO)_5$ (axial CO bonds: −5.2885 Hartree/A^3; equatorial CO bonds: −5.2165 Hartree/A^3). It seems that in particular for FeC bonds additional factors determine the bond strength, which are not reflected by the $H(\mathbf{r}_b)$ values focusing just on the bond critical point $\mathbf{r}_b$, such as the charge of carbon monoxide oxygen and the CO π^* back donation, which will be discussed in the following.

In Fig. 5a, the CO π^* populations reflecting π back donation from the occupied d_π orbitals of iron into the vacant CO π^* orbitals are correlated with the local mode force constants k^a(CO) and in Fig. 5b the NBO charges on the carbon monoxide oxygen are related to the local mode k^a(CO) force constants. As expected, we find the lowest CO π^* populations for the wild-type **H64δ** (CO π^* = 0.3838 *e*) and the gas phase model **Gas** (CO π^* = 0.3940 *e*) being identified via their BSO *n* values as the compounds with the strongest CO bonds among the 20 compounds investigated in this work. The largest CO π^* population is encountered for **V68N** (CO π^* = 0.5025 *e*) followed by **H64ϵR** (CO π^* = 0.4932 *e*) belonging to the compounds with the weakest CO bonds. In comparison, π back donation in $Fe(CO)_5$ (**$Fe(CO)_5$(ax)**, CO π^* = 0.4188 *e*; **$Fe(CO)_5$(eq)**, CO π^* = 0.4433 *e*) falls into the medium range. It is noteworthy that all compounds with larger CO π^* population and as such weaker CO bonds belong to the class of compounds forming a hydrogen bond between the distal histidine and the carbon monoxide oxygen atom. This important finding will be further discussed below. Figure 5b looks like a mirror image of Fig. 5a, i.e., CO π^* population correlates with the charge on O, which is shown in Figure S5 of the electronic supplementary material; $R^2 = 0.9604$. The largest NBO charge on O is found for **V68N** (NBO(O) charge = −0.549 *e*) and the smallest for **H64δ** (NBO(O) charge = −0.454 *e*) and gas phase model **Gas** (NBO(O) charge = −0.464 *e*). For comparison, the oxygen charge of carbon monoxide (NBO(O) charge = −0.508 *e*) lies in between. Again, all compounds with larger oxygen charges belong to the class of compounds which form hydrogen bonds.

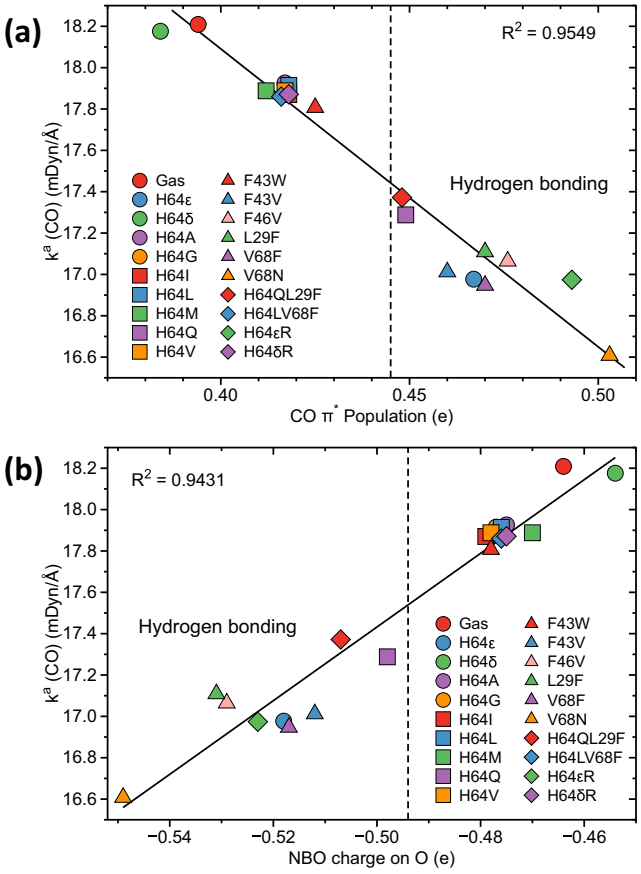

Fig. 5 **a** Local mode force constants k^a(CO) versus CO π^* population. **b** Local mode force constants k^a(CO) versus NBO charge of the O atom of CO. Gas phase calculations at the PBE0/6-31G(d,p) and protein calculations at the PBE0/6-31G(d,p)/AMBER levels of theory

We did not find a correlation between the NBO charges on the carbon monoxide carbon and the local mode force constants k^a(FeC); see Fig. S6 of the electronic supplementary material. On the other hand, we observed an interesting correlation between the CO π^* population and local mode k^a(FeC) force constants, as shown in Fig. 6a, and a similar correlation between NBO(O) charges and local k^a(FeC) force constants, as shown in Fig. S7 of the electronic supplementary material. Weaker FeC bonds are related to smaller CO π back donation, i.e., stronger CO bonds and vice versa. Closing the circle, Fig. 6b shows a direct comparison of local k^a(CO) and k^a(FeC) quantifying for the first time the suggested inverse relationship [40, 52, 57–59] between the CO and FeC bond strengths in these compounds on the basis of local mode force constants as quantitative bond strength measure.

Hydrogen bonding

In the following, we will address two so far open questions: (i) how distal hydrogen bonding is influenced by the protein environment and (ii) what role it plays for the modulation of the CO and FeC bond strengths. In Fig. 7a, hydrogen bond

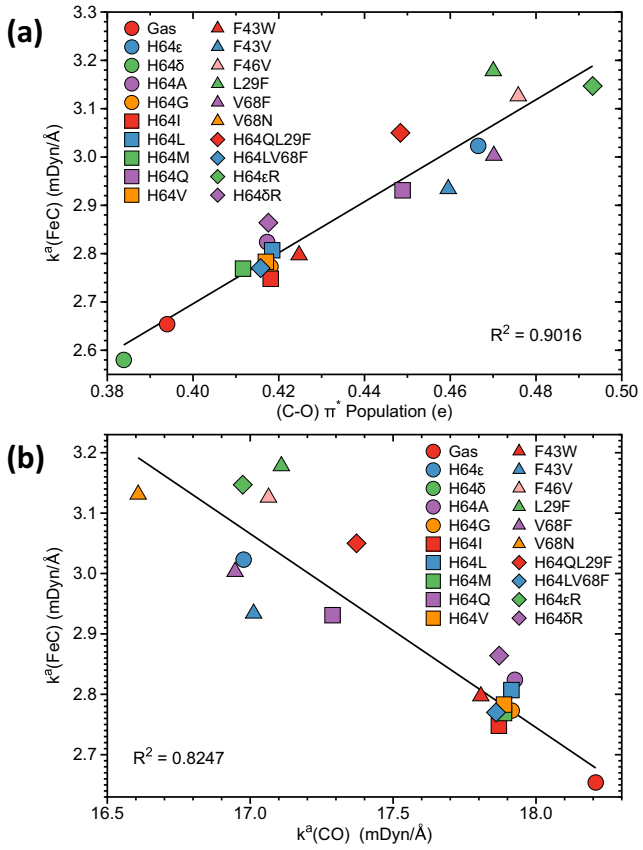

Fig. 6 **a** Local mode force constants k^a(FeC) versus CO π^* population. **b** Local mode force constants k^a(FeC) versus local mode force constants k^a(CO). Gas phase calculations at the PBE0/6-31G(d,p) and protein calculations at the PBE0/6-31G(d,p)/AMBER levels of theory

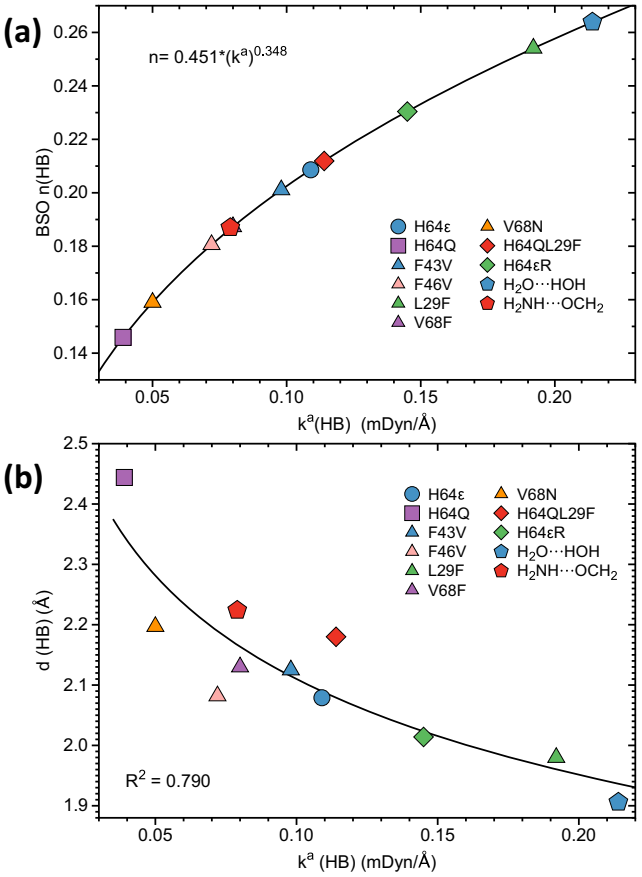

Fig. 7 **a** Bond strength order BSO n for the hydrogen bonds calculated from the local mode force constants k^a(HB) via Eq. (14). **b** Hydrogen bond distances d(HB) versus k^a(HB). Gas phase calculations at the PBE0/6-31G(d,p) and protein calculations at the PBE0/6-31G(d,p)/AMBER levels of theory

strength orders BSO n(HB) calculated from local mode force constants k^a(HB) are shown and in Fig. 7b hydrogen bond distances d are related with the local mode force constants k^a(HB). In Fig. 8a, the covalent character of the hydrogen bond is shown as reflected by the $H(\mathbf{r}_b)$(HB) values and in Fig. 8b local mode force constants k^a(HB) are related to the charge on the CO oxygen atom. Table 2 summarizes the corresponding hydrogen bond data. As reference, the water dimer and the ammonia-formaldehyde dimer are included.

As demonstrated by our results, local mode force constants k^a(HB) and related BSO n(HB) are sensitive antennas monitoring even small changes in hydrogen bonding resulting from a different electrostatic and steric environment in the modified heme pocket of the mutants. Although all hydrogen bonds investigated in this work are of the same type, i.e., NH···OC bonds formed between CO and the His64 ϵ tautomer or its replacement with another amino acid (see Table S2 of the electronic supplementary material), they are of different strengths with BSO n values ranging from 0.146 (**H64Q**) to 0.254 (**L29F**), which almost reaches the hydrogen bond strength of the water dimer (NBO $n = 0.264$) while the hydrogen bond of the ammonia-formaldehyde dimer lies in the middle range with an NBO n value of 0.187. The hydrogen bond distances stretch from 1.980 Å for **L29F** to 2.444 Å for **H64Q**. As in the case of the CO and FeC bonds, the Badger rule applies to the hydrogen bonds (although some scattering, $R^2 = 0.790$), i.e., the shortest hydrogen bond is also the strongest whereas the longest is the weakest, as shown in Fig. 7b.

In the **V68N** mutation, Val68 of wild-type **H64ϵ** is replaced with asparagine. Replacing the non-polar Val68 side chain with a polar form amide group enhances the electrostatic field in the heme pocket. The O atom of CO gets attracted by the amide hydrogens too, enlarging its distance to the His64 hydrogen from 2.078 Å in **H64ϵ** to 2.197 Å, substantially weakening the hydrogen bond. In **H64Q,** His64 is replaced by glutamine, which is a more drastic change. Instead of a hydrogen bond via the imidazole side chain, the hydrogen bond is now formed via a hydrogen of the amide side chain, leading to the weakest (BSO $n = 0.146$) and longest hydrogen bond ($d = 2.444$ Å) of our series. Noteworthy is that the strongest hydrogen bond is not found for a His64 replacement but for replacing the distal Leu29 with phenylalanine in wild-type **H64ϵ** distal heme pocket, i.e., replacing the isopropyl side chain with a phenyl group, as realized in **L29F**. The phenyl

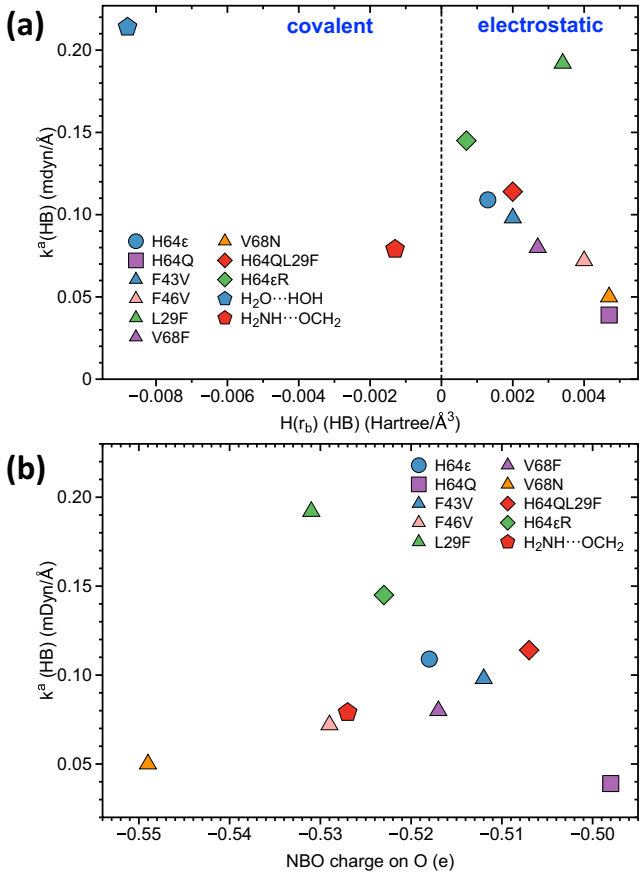

Fig. 8 **a** Energy density $H(\mathbf{r}_b)$(HB) versus k^a(HB). **b** k^a(HB) versus NBO(O). Gas phase calculations at the PBE0/6-31G(d,p) and protein calculations at the PBE0/6-31G(d,p)/AMBER levels of theory

π cloud pushes the O atom closer to His64, leading to the shortest H···O distance of 1.980 Å and strongest hydrogen bond. It is also interesting to note that the second strongest hydrogen bond (BSO $n = 0.230$) is found for **H64εR** with a reversion of the peripheral vinyl groups of the porphyrin framework. This result clarifies for the first time that in addition to effecting overall heme properties, such as the reduction potential, this reversion also influences the binding of exogenous ligands, such as CO, as has been questioned [35, 52]. Combining the replacement of the distal Leu29 with a phenylalanine and the replacement of His64 with glutamine as realized in **H64QL29F** leads to a medium hydrogen bond strength of 0.212, obviously balancing the two extremes.

In Fig. 8a, the energy density $H(\mathbf{r}_b)$(HB) is compared with the local force constant k^a(HB). According to the Cremer-Kraka criterion [120, 125] for covalent bonding, all hydrogen bonds investigated in this work are of electrostatic character. There is the trend that weaker hydrogen bonds have an increase in electrostatic character, in line with our studies on other hydrogen bonded systems [89]. According to its k^a value, the outlier **L29F** should have covalent hydrogen bond character in contrast to all other compounds investigated in this work. It is interesting to note that **L29F** has been originally identified as A_3 conformer with a somewhat different electron density distribution [25]. $H(\mathbf{r}_b)$ as a local electron density measure focusing on the bond critical point $\mathbf{r}_b$ cannot pick up these more regional differences.

As evidenced by Fig. 5, mutations with distal hydrogen bonding possess stronger CO π back donation and as such weaker CO and stronger FeC bonds. Therefore, the question arises if hydrogen bonding leads to enhanced CO charge polarization and stabilization of the negative charge on the CO oxygen which goes in line with stronger CO π back bonding. However, as revealed by Fig. 8b, this is not the case; there is no correlation between k^a(HB) and the NBO charge on oxygen. The same holds for CO π back donation. As shown in Fig. S8 of the electronic supplementary material, there is only no direct correlation between the hydrogen bond strength and the CO π^* population. Therefore, the conclusion is that the key role of distal hydrogen bonding is structural rather than electronic; hydrogen bonding keeps the CO ligand at a Fe distance with optimum orbital overlap between the Fe d_π and the CO π^* orbitals, which is a prerequisite for π back donation. There is no significant correlation between hydrogen bonding and FeCO bending angle as reflected by the FeCO bond angles in Table 1 and Fig. S9 of the electronic supplementary material. In line with experimental studies reporting only small FeCO angular distortions in the crystal structures of MbCO mutants in comparison with NO and O_2 ligands [52], we find angle variations from 180° of less than 10°, which are too small to allow a separation of mutants with and without hydrogen bonding.

Assessment of experimental frequencies as FeC and CO bond strength descriptors

The discussion of the FeC bond and CO bond strength based on resonance Raman spectroscopic data has drawn upon normal vibrational FeC and CO vibrational frequencies [17, 22, 25, 35, 69]. As discussed above, normal vibrational modes are generally delocalized; therefore, the identification of a particular normal mode as FeC or CO stretching (in most studies performed via isotope substitution [52, 148]) seems to be questionable, in particular for a complex system with more than 100 atoms and normal mode vectors with a dimension of almost 300. The local vibrational mode analysis does not only provide local mode stretching force constants and frequencies as useful bond strength measures, it can also serve as an effective tool to analyze the character of normal vibrational modes.

For smaller systems, in particular gas phase molecules, this can be accomplished via the CNM procedure [63, 118]. In Fig. 9, the decomposition of the CO, FeC, and axial FeN normal modes into local mode contributions is shown in the form of a pie diagram for iron-porphyrin imidazole, a simpler

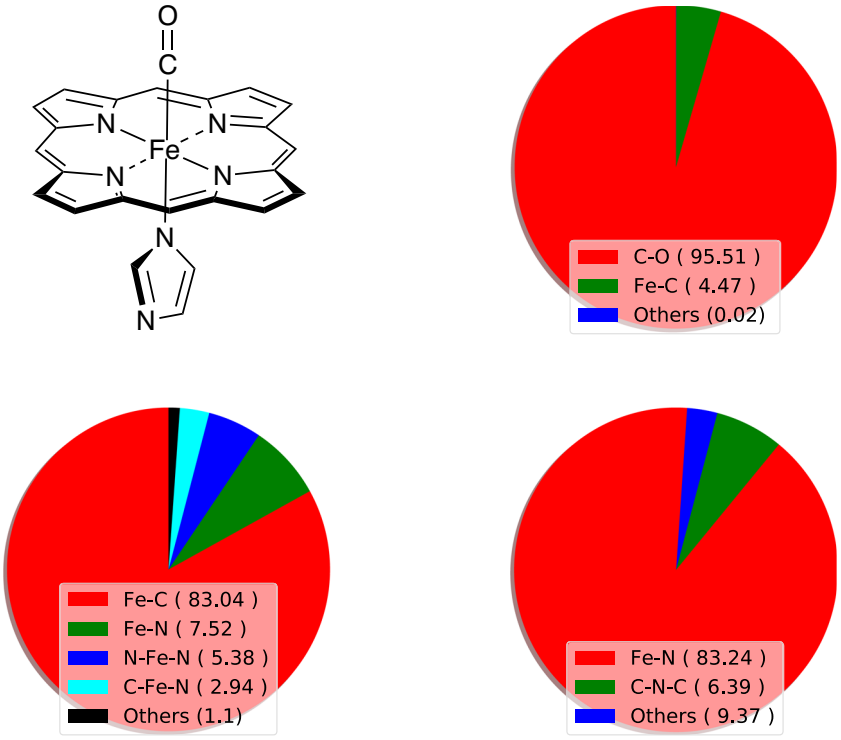

Fig. 9 Decomposition of the CO, FeC, and axial FeN normal modes into local mode contribution, shown in the form of a pie chart for a simplified gas phase model. Contributions are given as percentages. Calculated at the PBE0/6-31G(d,p) level of theory

mimic of the gas phase model, which was recently suggested as potential vibrational Stark effect probe [118]. The complete CNM of all normal modes is shown in Fig. S11 of the electronic supplementary material in the form of a bar diagram. The CO normal mode is localized by more than 95% with a small contribution of less than 5% of the FeC mode. The FeC normal mode is more delocalized (83% FeC contribution), resulting from mode coupling of 7.5% with the axial FeN stretching mode, 5.4% coupling with $N_{eq}FeN_{ax}$, and about 3% coupling with equatorial CFeN bending modes. The axial FeN modes show a similar delocalization (83% FeN contribution), with a coupling of 6.4% from the imidazole CNC bonding mode.

For protein systems, the CNM procedure is no longer doable; however, we can compare the normal FeC and CO vibrational stretching frequencies with their local mode counterparts. If the normal modes would be free of mode-mode coupling, both sets of frequencies should correlate. Therefore, we tested for all complexes with published experimental FeC and/or CO normal mode frequencies to what extent these frequencies accord with the corresponding local mode frequencies. The results are summarized in Table 3 and Fig. 10.

As revealed by the data in Table 3 and Fig. 10, there is some weak correlation between FeC normal and local modes. Smallest normal and local frequencies are found for **H64I** and **H64V** with weaker FeC bonds (see Fig. 2a), while **H64ϵR** with a strong FeC bond is characterized by both large normal and local mode frequencies. Figure 2a also shows some outliers. This demonstrates that (i) the assignment of a protein normal vibration mode with many degrees of freedoms via isotope substitution is not free from ambiguities, and (ii) that

Table 3 Calculated local mode frequencies ω^a and experimental normal mode frequencies ω_{exp} of FeC and CO bonds for those complexes with available experimental data. Calculated frequencies at the PBE0/6-31G(d,p)/AMBER level of theory

Model	FeC			CO		
	ω^a (cm^{-1})	ω_{exp} (cm^{-1})	Ref. (exp)	ω^a (cm^{-1})	ω_{exp} (cm^{-1})	Ref. (exp)
H64ϵ	721	508	[25]	2050	1944	[69]
		509	[69]		1945	[25]
					1946	[17]
H64A	696	504	[25]	2106	1958	[25]
H64G	690	506	[17]	2106	1944	[17]
H64I	687	505	[22]	2103	–	–
H64L	694	506	[22]	2106	–	–
H64M	690	506	[17]	2104	1947	[17]
H64Q	709	507	[69]	2069	1944	[69]
H64V	691	504	[22]	2104	–	–
F43W	693	503	[25]	2099	1955	[25]
F46V	733	504	[25]	2055	1956	[25]
L29F	739	526	[25]	2058	1932	[25]
V68N	733	508	[25]	2028	1944	[25]
H64QL29F	724	513	[69]	2074	1938	[69]
H64ϵR	735	511	[35]	2050	–	–

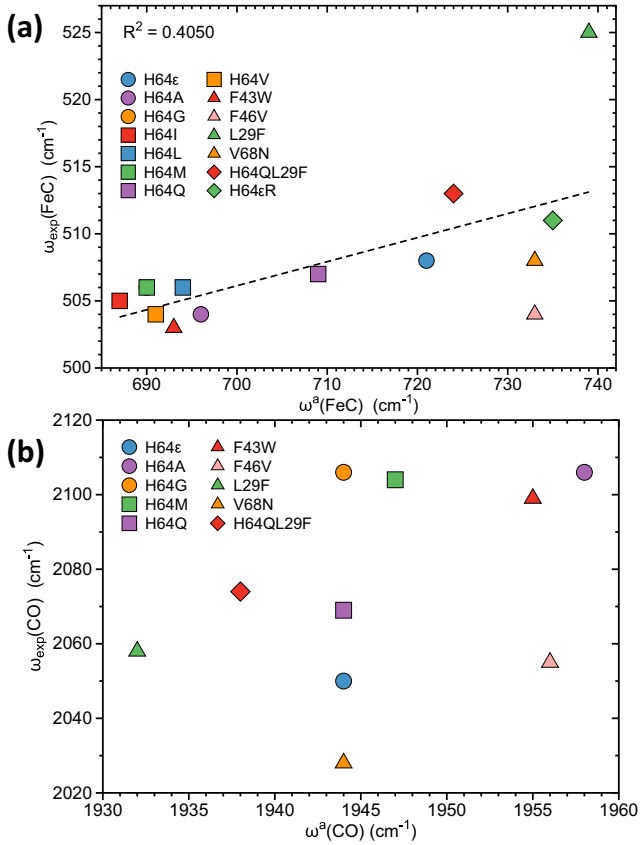

Fig. 10 **a** Calculated local mode frequencies ω^a(FeC) versus experimental normal mode frequencies ω_{exp}(FeC). **b** Calculated local mode frequencies ω^a(CO) versus experimental normal mode frequencies ω_{exp}(CO). Protein calculations at the PBE0/6-31G(d,p)/AMBER levels of theory

there is mode-mode coupling with other protein motions. It is interesting to note that the experimental resonance Raman spectrum of the **L29F** protein mutation shows only one maximum related to the FeC stretching vibration (526 cm^{-1}), which was originally assigned to the A_3 rather than A_1 conformation state [25]. This explains the large derivation from the other complexes shown in Fig. 10a. For the CO stretching mode, which is almost perfectly localized in the gas phase, there is no correlation between normal and local vibrational modes for the ten complexes shown in Fig. 10b. This suggests that the stronger exposure of the CO ligand to the electrostatic field created by the heme pocket leads to enhanced mode-mode coupling, clearly disqualifying the normal mode CO frequencies as bond strength descriptors in the protein environment. Only their local mode counterparts can serve for this purpose.

Conclusions

In this QM/MM computational study, we investigated the FeC and CO bond strength and potential hydrogen bonding in 20 MbCO models, involving a gas phase model, two models of the wild-type protein, and 17 protein mutations. As a key analysis tool, we used the local mode theory which provides via local mode force constants a quantitative measure of the strength, a chemical bond, and/or weak chemical interaction based on vibrational spectroscopy. The results of our investigation quantify for the first time the suggested inverse relationship between the FeC and CO bond strength in these systems, i.e., weaker FeC bonds are related to stronger CO bonds and vice versa. The FeC and CO bond strengths are determined by electronic structure reorganization of the Fe-CO moiety, which is initiated by the π back donation between Fe d_π and CO π^* orbitals. π back donation is modified by the electrostatic field of the distal side of the heme pocket, which polarizes the oxygen charge of carbon monoxide. In addition, CO bonding is weakened by potential hydrogen bonding between the CO oxygen and side chains of the surrounding amino acids. Such a complex scenario needs a qualified bond strength descriptor absorbing all these cumulative effects. We have shown that local mode force constants or local mode frequencies are the perfect tool for this purpose whereas normal mode frequencies are not qualified. Our comprehensive results provide new guidelines for the fine-tuning of existing and the design of MbCO models with specific FeC, CO, and CO···H bond strengths.

Acknowledgments The authors thank Southern Methodist University for providing excellent computational resources. We thank Niraj Verma for assistance with the CNM analysis of iron-porphyrin imidazole and Gernot Frenking, Vytor Oliveira, Dani Setiawan, and Daniel Sethio for fruitful discussions.

Funding This work was supported by the National Science Foundation (Grant CHE 1464906).

References

1. Gros G, Wittenberg BA, Jue T (2010). J Exp Biol 213(16):2713
2. Kendrew JC, Dickerson RE, Strandberg BE, Hart RG, Davies DR, Phillips DC, Shore VC (1960). Nature 185(4711):422
3. Springer BA, Sligar SG, Olson JS, Phillips GNJ (1994). Chem Rev 94(3):699
4. Ordway GA, Garry DJ (2004). J Exp Biol 207(Pt 20):3441
5. Tsai AL, Berka V, Martin E, Olson JS (2012). Biochemistry 51(1):172
6. Spiro TG, Ibrahim M, Wasbotten IH (Elsevier, Amsterdam, 2008), chap. 4, pp. 95–123
7. Buhrke D, Hildebrand P (2020). Chem Rev 120:3577
8. Schneider SH, Boxer SG (2016). J Phys Chem B 120:9672
9. Hush NS, Reimers JR (1995). J Phys Chem 99:15798
10. Fried SD, Boxer SG (2015). Acc Chem Res 48:998
11. Fried SD, Boxer SG (2017). Annu Rev Biochem 86:387

12. Ma J, Pazos IM, Zhang W, Culik RM, Gai F (2015). Annu Rev Phys Chem 66:357
13. Błasiak B, Londergan CH, Webb LJ, Cho M (2017). Acc Chem Res 50:968
14. Bishop DM (1993). J Chem Phys 98:3179
15. Tsubaki M, Srivastava RB, Yu NT (1982). Biochemistry 21(6): 1132
16. Ramsden J, Spiro TG (1989). Biochemistry 28(8):3125
17. Morikis D, Champion PM, Springer BA, Sligar SG (1989). Biochemistry 28(11):4791
18. Oldfield E, Guo K, Augspurger JD, Dykstra CE (1991). J Am Chem Soc 113(20):7537
19. Cameron AD, Smerdon SJ, Wilkinson AJ, Habash J, Helliwell JR, Li T, Olson JS (1993). Biochemistry 32(48):13061
20. Braunstein DP, Chu K, Egeberg KD, Frauenfelder H, Mourant JR, Nienhaus GU, Ormos P, Sligar SG, Springer BA, Young RD (1993). Biophys J 65(6):2447
21. Li T, Quillin ML, Phillips GN, Olson JS (1994). Biochemistry 33(6):1433
22. Ling J, Li T, Olson JS, Bocian DF (1994). Biochim Biophys Acta 1188(3):417
23. Ray GB, Li XY, Ibers JA, Sessler JL, Spiro TG (1994). J Am Chem Soc 116(1):162
24. Yang F, Phillips Jr GN (1996). J Mol Biol 256(4):762
25. Anderton CL, Hester RE, Moore JN (1997). Biochim Biophys Acta 1338(1):107
26. Nienhaus UG, Müller JD, McMahon BH, Frauenfelder H (1997). Physica D 107(2):297
27. Kachalova GS, Popov AN, Bartunik HD (1999). Science 284(5413):473
28. Vojtěchovský J, Chu K, Berendzen J, Sweet RM, Schlichting I (1999). Biophys J 77(4):2153
29. Phillips GN, Teodoro ML, Li T, Smith B, Olson JS (1999). J Phys Chem B 103(42):8817
30. Chu K, Vojtchovský J, McMahon BH, Sweet RM, Berendzen J, Schlichting I (2000). Nature 403(6772):921
31. Franzen S, Peterson ES, Brown D, Friedman JM, Thomas MR, Boxer SG (2002). Eur J Biochem 269(19):4879
32. Lamb DC, Nienhaus K, Arcovito A, Draghi F, Miele AE, Brunori M, Nienhaus GU (2002). J Biol Chem 277(14):11636
33. Nienhaus K, Deng P, Olson JS, Warren JJ, Nienhaus GU (2003). J Biol Chem 278(43):42532
34. Nienhaus K, Olson JS, Franzen S, Nienhaus GU (2005). J Am Chem Soc 127(1):40
35. Rwere F, Mak PJ, Kincaid JR (2008). Biochemistry 47(48):12869
36. Barends TRM, Foucar L, Ardevol A, Nass K, Aquila A, Botha S, Doak RB, Falahati K, Hartmann E, Hilpert M, Heinz M, Hoffmann MC, Köfinger J, Koglin JE, Kovacsova G, Liang M, Milathianaki D, Lemke HT, Reinstein J, Roome CM, Shoeman RL, Williams GJ, Burghardt I, Hummer G, Boutet S, Schlichting I (2015). Science 350(6259):445
37. Jewsbury P, Kitagawa T (1994). Biophys J 67(6):2236
38. Kushkuley B, Stavrov SS (1996). Biophys J 70(3):1214
39. Sigfridsson E, Ryde U (1999). J Biol Inorg Chem 4:99
40. Vogel KM, Kozlowski PM, Zgierski MZ, Spiro TG (2000). Inorg Chim Acta 297(1):11
41. Rabenstein B, Knapp EW (2001). Biophys J 80(3):1141
42. Rovira C, Schulze B, Eichinger M, Evanseck JD, Parrinello M (2001). Biophys J 81(1):435
43. Spiro TG, Zgierski MZ, Kozlowski PM (2001). Coord Chem Rev 219-221:923
44. Sigfridsson E, Ryde U (2002). J Inorg Biochem 91(1):101
45. Merchant KA, Noid WG, Akiyama R, Finkelstein IJ, Goun A, McClain BL, Loring RF, Fayer MD (2003). J Am Chem Soc 125(45):13804
46. Harvey JN (2004). Faraday Discuss 127(0):165
47. Spiro TG, Wasbotten IH (2005). J Inorg Biochem 99(1):34
48. Bikiel DE, Boechi L, Capece L, Crespo A, De Biase PM, Di Lella S, González Lebrero MC, Martí MA, Nadra AD, Perissinotti LL, Scherlis DA, Estrin DA (2006). Phys Chem Chem Phys 8(48): 5611
49. Strickland N, Mulholland AJ, Harvey JN (2006). Biophys J 90(4): L27
50. Alcantara RE, Xu C, Spiro TG, Guallar V (2007). PNAS 104(47): 18451
51. Devereux M, Meuwly M (2009). Biophys J 96(11):4363
52. Spiro TG, Soldatova AV, Balakrishnan G (2013). Coord Chem Rev 257(2):511
53. Shimada H, Caughey WS (1982). J Biol Chem 257(20):11893
54. Ansari A, Berendzen J, Braunstein D, Cowen BR, Frauenfelder H, Hong MK, Iben IET, Johnson JB, Ormos P, Sauke TB, Scholl R, Schulte A, Steinbach PJ, Vittitow J, Young RD (1987). Biophys Chem 26(2):337
55. Mourant JR, Braunstein DP, Chu K, Frauenfelder H, Nienhaus GU, Ormos P, Young RD (1993). Biophys J 65(4):1496
56. Johnson JB, Lamb DC, Frauenfelder H, Müller JD, McMahon B, Nienhaus GU, Young RD (1996). Biophys J 71(3):1563
57. Silvernail NJ, Roth A, Schulz CE, Noll BC, Scheidt WR (2005). J Am Chem Soc 127:14422
58. Spiro TG, Smulevich G, Su C (1990). Biochemistry 29(19):4497
59. Falahati K, Tamura H, Burghardt I, Huix-Rotllant M (2018). Nat Commun 9:4502
60. Konkoli Z, Cremer D (1998). Int J Quantum Chem 67:1
61. Konkoli Z, Larsson JA, Cremer D (1998). Int J Quantum Chem 67:11
62. Konkoli Z, Cremer D (1998). Int J Quantum Chem 67:29
63. Konkoli Z, Larsson JA, Cremer D (1998). Int J Quantum Chem 67:41
64. Cremer D, Larsson JA (1998) E. Kraka. In: Parkanyi C (ed) Theoretical and computational chemistry. Elsevier, Amsterdam, pp 259–327
65. Springer BA, Egeberg KD, Sligar SG, Rohlfs RJ, Mathews AJ, Olson JS (1989). J Biol Chem 264:3057
66. Olson JS, Phillips Jr GN (1997). J Biolog Inorg Chem 2(4):544
67. Phillips SEV, Schoenborn BP (1981). Nature 292(5818):81
68. Cheng X, Schoenborn BP (1991). J Mol Biol 220(2):381
69. Zhao X, Vyas K, Nguyen BD, Rajarathnam K, La Mar GN, Li T, Phillips GN, Eich RF, Olson JS, Ling J, Bocian DF (1995). J Biol Chem 270(35):20763
70. Wilson EB, Decius JC, Cross PC (1955) Molecular vibrations. McGraw-Hill, New York
71. Woodward LA (1972) Introduction to the theory of molecular vibrations and vibrational spectroscopy. Oxford University Press, Oxford
72. Califano S (1976) Vibrational states. Wiley, New York
73. Cremer D, Kraka E (2010). Curr Org Chem 14:1524
74. Kraka E, Larsson JA (2010) D. Cremer. In: Grunenberg J (ed) Computational spectroscopy. Wiley, New York, pp 105–149
75. Zou W, Filatov M, Cremer D (2011). Theor Chem Acc 130:633
76. Wilson Jr EB (1939). J Chem Phys 7:1047
77. Groner P (2006) Normal coordinate analysis. Wiley, New York
78. Kelley JD, Leventhal JJ, in Problems in Classical and Quantum Mechanics (Springer, 2017), pp. 95–117
79. Neto N (1984). Chem Phys 91:89
80. Stare J (2007). J Chem Inf Model 47(3):840
81. Zou W, Cremer D (2016). Chem Eur J 22:4087
82. Kraka E, Setiawan D, Cremer D (2015). J Comput Chem 37:130
83. Setiawan D, Kraka E, Cremer D (2015). J Phys Chem A 119:9541
84. Kalescky R, Kraka E, Cremer D (2013). J Phys Chem A 117:8981
85. Humason A, Zou W, Cremer D (2014). J Phys Chem A 119:1666
86. Kalescky R, Zou W, Kraka E, Cremer D (2014). J Phys Chem A 118:1948

87. Kraka E, Cremer D (2009). ChemPhysChem 10:686
88. Setiawan D, Sethio D, Cremer D, Kraka E (2018). Phys Chem Chem Phys 20:23913
89. Freindorf M, Kraka E, Cremer D (2012). Int J Quantum Chem 112:3174
90. Kalescky R, Zou W, Kraka E, Cremer D (2012). Chem Phys Lett 554:243
91. Kalescky R, Kraka E, Cremer D (2013). Mol Phys 111:1497
92. Kalescky R, Zou W, Kraka E, Cremer D (2014). Aust J Chem 67: 426
93. Tao Y, Zou W, Jia J, Li W, Cremer D (2017). J Chem Theory Comput 13:55
94. Tao Y, Zou W, Kraka E (2017). Chem Phys Lett 685:251
95. Freindorf M, Tao Y, Sethio D, Cremer D, Kraka E (2018). Mol Phys 117:1172
96. Makoś MZ, Freindorf M, Sethio D, Kraka E (2019). Theor Chem Acc 138:76
97. Lyu S, Beiranvand N, Freindorf M, Kraka E (2019). J Phys Chem A 123:7087
98. Oliveira V, Kraka E, Cremer D (2016). Phys Chem Chem Phys 18:33031
99. Oliveira V, Kraka E, Cremer D (2016). Inorg Chem 56:488
100. Oliveira V, Cremer D (2017). Chem Phys Lett 681:56
101. Yannacone S, Oliveira V, Verma N, Kraka E (2019). Inorganics 7: 47
102. Oliveira VP, Kraka E, Machado FBC (2019). Inorg Chem 58: 14777
103. Oliveira VP, Marcial BL, Machado FBC, Kraka E (2020). Materials 13:55
104. Setiawan D, Kraka E, Cremer D (2014). Chem Phys Lett 614:136
105. Setiawan D, Kraka E, Cremer D (2014). J Phys Chem A 119:1642
106. Setiawan D, Cremer D (2016). Chem Phys Lett 662:182
107. Oliveira V, Cremer D, Kraka E (2017). J Phys Chem A 121:6845
108. Oliveira V, Kraka E (2017). J Phys Chem A 121:9544
109. Sethio D, Oliveira V, Kraka E (2018). Molecules 23:2763
110. Li Y, Oliveira V, Tang C, Cremer D, Liu C, Ma J (2017). Inorg Chem 56:5793
111. Kalescky R, Kraka E, Cremer D (2013). Inorg Chem 53:478
112. Setiawan D, Kalescky R, Kraka E, Cremer D (2016). Inorg Chem 55:2332
113. Cremer D, Kraka E (2017). Dalton Trans 46:8323
114. Li Y, Liu C, Oliveira V, Cremer D, Chen Z, Ma J (2018). Mol Phys 117:1442
115. Kalescky R, Kraka E, Cremer D (2013). J Phys Chem A 118:223
116. Setiawan D, Kraka E, Cremer D (2016). J Organomet Chem 81: 9669
117. Zou W, Kalescky R, Kraka E, Cremer D (2012). J Chem Phys 137:084114
118. Verma N, Tao Y, Zou W, Chen X, Chen X, Freindorf M, Kraka E (2020). Sensors 20:2358
119. Tao Y, Zou W, Sethio D, Verma N, Qiu Y, Tian C, Cremer D, Kraka E (2019). J Chem Theory Comput 15:1761
120. Tao Y, Qiu Y, Zou W, Nanayakkara S, Yannacone S, Kraka E (2020). Molecules 25:1589
121. Mayer I (1983). Chem Phys Lett 97:270
122. Mayer I (1986). Int J Quantum Chem 29:477
123. Mayer I (2007). J Comput Chem 28(1):204
124. Cremer D, Kraka E (1984). Angew Chem Int Ed 23:627
125. Cremer D, Kraka E (1984). Croat Chem Acta 57:1259
126. Bader R (1984) Atoms in Molecules: A Quantum Theory. Oxford University press
127. Bader R (1998). Chem Rev 1:64
128. Adamo C, Barone V (1999). J Chem Phys 110(13):6158
129. Ditchfield R, Hehre WJ, Pople JA (1971). J Chem Phys 54:724
130. Chung LW, Sameera WMC, Ramozzi R, Page AJ, Hatanaka M, Petrova GP, Harris TV, Li X, Ke Z, Liu F, Li HB, Ding L, Morokuma K (2015). Chem Rev 115:5678
131. Case DA, Ben-Shalom IY, Brozell SR, Cerutti DS, Cheatham TE, Cruzeiro VWD, Darden TA, Duke RE, Ghoreishi D, Gilson MK, Gohlke H, Goetz AW, Greene D, Harris R, Homeyer N, Izadi S, Kovalenko A, Kurtzman T, Lee TS, LeGrand S, Li P, Lin C, Liu J, Luchko T, Luo R, Mermelstein DJ, Merz KM, Miao Y, Monard G, Nguyen C, Nguyen H, Omelyan I, Onufriev A, Pan F, Qi R, Roe DR, Roitberg A, Sagui C, Schott-Verdugo S, Shen J, Simmerling CL, Smith J, Salomon-Ferrer R, Swails J, Walker RC, Wang J, Wei H, Wolf RM, Wu X, Xiao L, York DM, Kollman PA, AMBER (2018). University of California, San Francisco
132. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Scalmani G, Barone V, Petersson GA, Nakatsuji H, Li X, Caricato M, Marenich AV, Bloino J, Janesko BG, Gomperts R, Mennucci B, Hratchian HP, Ortiz JV, Izmaylov AF, Sonnenberg JL, Williams-Young D, Ding F, Lipparini F, Egidi F, Goings J, Peng B, Petrone A, Henderson T, Ranasinghe D, Zakrzewski VG, Gao J, Rega N, Zheng G, Liang W, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Vreven T, Throssell K, Montgomery Jr JA, Peralta JE, Ogliaro F, Bearpark MJ, Heyd JJ, Brothers EN, Kudin KN, Staroverov VN, Keith TA, Kobayashi R, Normand J, Raghavachari K, Rendell AP, Burant JC, Iyengar SS, Tomasi J, Cossi M, Millam JM, Klene M, Adamo C, Cammi R, Ochterski JW, Martin RL, Morokuma K, Farkas O, Foresman JB, Fox DJ (2016) Gaussian 16. Gaussian Inc. Wallingford
133. Kraka E, Zou W, Filatov M, Gräfenstein J, Gauss J, He Y, Wu A, Konkoli Z, He Z, Cremer D et al (2016) COLOGNE16
134. Reed AE, Curtiss LA, Weinhold F (1988). Chem Rev 88:899
135. Keith TA, AIMALL (2017) TK gristmill software, Overland Park KS
136. Frenking G, Loschen C, Krapp A, Fau S, Strauss SH (2007). J Comput Chem 28:117
137. Dewar MJS (1953). Bull Soc Chim 48:112
138. Chatt J, Duncanson L (1953). J Chem Soc:2939
139. Gordon CP, Andersen RA, Coperet C (2019). Helv Chim Acta 102:e1900151
140. Bistoni G, Rampino S, Scafuri N, Ciancaleoni G, Zuccaccia D, Belpassi L, Tarantelli F (2016). Chem Sci 7:1174
141. Koch D, Chen Y, Golub P, Manzhos S (2019). Phys Chem Chem Phys 21:20814
142. Li M, Oliver AG, Scheidt WR (2018). Inorg Chem 57:5648
143. Zhou M, Andrews L, Bauschlicher CW (2001). Chem Rev 101: 1931
144. Tolman CA (1972). Chem Soc Rev 1(3):337
145. Tolman CA (1977). Chem Rev 77(3):313
146. Badger RM (1934). J Chem Phys 2:128
147. Kraka E, Cremer D (2012). Rev Proc Quim:39–42
148. Ibrahim M, Xu C, Spiro TG (2006). J Am Chem Soc 128(51): 16834