

CO Bonding in Hexa- and Pentacoordinate Carboxy-Neuroglobin - A QM/MM and Local Vibrational Mode Study

Marek Freindorf, Alexis Antoinette Ann Delgado, Elfi Kraka*

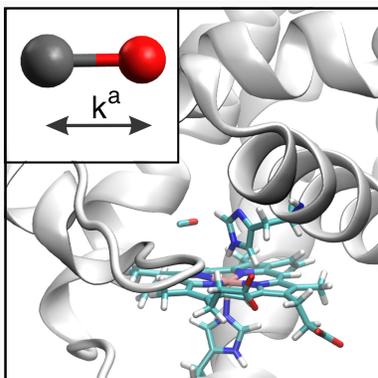
June 11, 2022

Abstract

We present a comprehensive investigation on the different role of CO in carboxy-neuroglobin i) as ligand of the heme group in the active site forming a bond with the heme iron and ii) dissociated from the heme group but still trapped inside the active site, focusing on two specific orientations, one with CO perpendicular to the plane defined by the distal histidine of the enzyme (form **A**) and one with CO located parallel to that plane (form **B**). Our study includes wild type carboxy-neuroglobin and nine known protein mutations. Considering that the distal histidine interacting with the heme group can adapt two different tautomeric forms and the two possible orientations of the dissociated CO, a total of 36 protein systems were analyzed in this study. Fully optimized geometries and vibrational frequencies were calculated at the QM/MM level, followed by the local mode analysis, to decode CO bond properties. The intrinsic bond strengths derived from the local mode analysis, complemented with NBO and QTAIM data, reveal that the strength of the CO bond, in the hexacoordinate (where CO is a ligand of the heme group) and pentacoordinate (where CO is dissociated from the heme group) scenarios, is dominated by through bond and through space charge transfer between CO and Fe, fine-tuned by electrostatic and dispersion interactions with the side chain amino acids in the distal heme pocket. Suggestions are made as to advise on how protein modifications can influence the molecular properties of the coordinated or dissociated CO, which could serve the fine-tuning of existing and the design of new neuroglobin models with specific FeC and CO bond strengths.

Keywords: Neuroglobin, CO, QM/MM, Local Mode Analysis, Intrinsic Bond Strength ■

*Department of Chemistry, Southern Methodist University, Dallas TX, USA



Using QM/MM and Local Mode Analysis we investigated molecular properties of the CO in the active site of carboxy-neuroglobin, with CO coordinated to the Fe of the heme group or dissociated from the heme group. In total 36 different neuroglobin systems were investigated.

INTRODUCTION

Neuroglobin (Ngb)¹⁻⁴ is a member of the globin superfamily of proteins sharing almost the same active site and similar protein properties with hemoglobin (Hb), myoglobin (Mb), and cytoglobin (Cygb).⁵⁻⁷ Ngb has been frequently studied during the past decade^{1,6} producing a vast amount of data and knowledge about this globin; however, a clear understanding of the biochemical role and function remains a challenge due to its engagement in various cell mechanisms including the potential involvement in cell survival under pathological conditions.¹ A multitude of experimental works suggest that Ngb is involved in a protective mechanism in response to brain hypoxia and ischemia,^{2,8-12} e.g., protecting the retina and optic nerve from degradation.^{13,14} Based on this ability to protect neurons from neuro-degenerative disorders and mitochondrial dysfunction, Ngb has been applied so far in therapies and treatments for strokes,^{8,9,15,16} Alzheimer's disease,^{17,18} and cancers.^{19,20}

Like Hb or Mb, Ngb can coordinate small signaling molecules L such as CO, NO, O₂, H₂S, CN⁻, NO[·] via the iron atom of the porphyrin ring. Their key role is to cause conformational changes of the protein resulting in diverse enzymatic activities,²¹⁻²³ where in higher doses, ligands such as CO or H₂S can lead to poisoning^{24,25} or to oxidative stress.²⁰ In the absence of an external ligand there are significant structural differences between Hb and Mb on the one hand and Ngb on the other hand. Hb and Mb are so-called pentacoordinate globins, in which the Fe of the heme group is bound by four nitrogen atoms of the porphyrin ring and the proximal histidine of the protein. Upon interaction with L the free sixth coordination site of the Fe is occupied. However, Ngb is a so-called hexacoordinate globin, i.e., without exogenous ligand the distal histidine (His64) binds to the sixth coordination position of the Fe.²⁶ Therefore, any external ligand has to compete with His64. Hexacoordination is a feature of various plant, bacterial, invertebrate and vertebrate globins, but its functional significance is not yet fully understood.^{27,28}

Ligand binding in Ngb occurs upon spontaneous rupture of His 64 and subsequent binding of the diatomic gas at the vacant sixth coordination site.^{29,30} Reported activation energies for human Ng-CO binding are generally about 50% higher than for Ng-His.^{31,32} The His64 side chain essentially maintains its position³³ and instead, the heme group slides deeper into the

protein pocket to enable ligand binding. This triggers a larger conformational reorganization of the protein, as confirmed for the crystal structure of carbon monoxide bound murine Ngb³⁴. Far-IR and THz domain spectroscopy,³⁵ resonance Raman,^{26,36,37} and time-resolved resonance Raman,³⁸ complemented with computational investigations^{39,40} have suggested that the biological function of Ngb and other globins is related in particular to the strength of the Fe-L bond, however these effects were not quantified.

One popular way to discuss metal–ligand (M-L) bonding is via the traditional Dewar-Chatt-Duncanson (DC) model is based on two charge transfer events: i) σ -donation from L towards M and ii) π -back donation from M (via d-orbitals) towards the anti-bonding π^* orbital of L.^{41,42} For L = CO the π -back donation from M \rightarrow CO is primarily responsible for weakening the CO bond and relates to experimentally observed red-shifts of the CO stretching frequency.^{43,44} Traditionally, spectroscopic shifts of the CO stretching frequency are associated with the so-called *Stark effect*.⁴⁵ which entered vibrational spectroscopy in 1995 with the seminal work of Chattopadhyay and Boxer on anisonitriles.⁴⁶ Since then, vibrational Stark effect (VSE) spectroscopy has become an important analytical method for assessing the influence of an electric field on a vibrational spectrum.^{47–52} The underlying assumption of VSE is that the normal stretching mode of the target bond such as CO^{53–55} or CN^{56–60} of a reporter molecule (termed vibrational Stark effect probe) is localized and free from coupling with other normal modes, so that its frequency shift directly reflects the influence of the electric field. For CO probe bonds it has been reported that the presence of d- π^* back bonding when CO is coordinated e.g., to heme results in highly increased sensitivity of the CO stretching frequency to the electric field.^{61,62} However, the validity of the underlying assumption of a localized Stark probe bond vibration has never been quantitatively assessed. Therefore, we recently performed a comprehensive analysis for a set of more than 100 currently used Stark effect probes including CO and CN probe bonds as well as potential candidates.⁶³ This led to a new protocol for the design of efficient VSE probes with optimally localized probe bonds.

As a secondary effect, the CO bond can further be weakened if the O atom gets involved in hydrogen bonding interactions, the opposite occurs if the C atom interacts with σ accep-

tors.⁶⁴ Comprehensive computational studies on transition metal carbonyl complexes, e.g., performed by Frenking and co-workers⁶⁵⁻⁷⁰ and other groups⁷¹⁻⁷⁷ have applied the DC model to explore the stabilizing role of CO ligands for both mono and poly-nuclear metal-carbonyl complexes. Some limitation of the DC and other Lewis type bonding model were recently pointed out.⁷⁸⁻⁸⁰ Two other popular strategies to describe M-L bond strengths are based on i) M-L bond dissociation energies (BDE)s^{76,81-84} and ii) M-L bond lengths. While all these attempts have certainly contributed to the chemical understanding of metal complexes, one has to realize that BDE values or bond lengths provide little insight into the intrinsic strength of the M-L bond. The BDE is a reaction parameter that includes all changes which take place during the dissociation process. It reflects the energy needed for bond breaking, but also contains energy contributions due to geometry relaxation and electron density reorganization in the dissociation fragments. Therefore, the BDE or related energy decomposition schemes⁸⁵ are not suitable measures of intrinsic chemical bond strength and their use has led in some cases to misinterpretation of bond strength.^{79,86-89} Also the M-L bond length is not a qualified bond strength descriptor. Numerous cases have been reported illustrating that a shorter bond is not always a stronger bond.⁹⁰⁻⁹⁴ Therefore, we have taken a different route by introduction an intrinsic bond strength measure based on the local vibrational mode analysis (LMA)⁹⁵ which has led to the metal electronic parameter (MLEP), for quantifying a M-L bonding,^{96,97} which was shown to be superior to the more qualitative Tolman electronic parameter (TEP).^{98,99} In two recent pilot studies applying a hybrid QM/MM (quantum mechanics/molecular mechanics) methodology,¹⁰⁰ followed by LMA, we could for the first time i) confirm and quantify the suggested inverse correlation of the CO and FeC bond strength^{37,101} and the special role of CO...H bonding in Mb¹⁰² and ii) could successfully clarify the binding mode of azanone (HNO) to the heme group of Mb¹⁰³, a long debated topic.^{23,104,105} Similar as in Mb, the binding of CO to Ngb is a reversible process. However, after dissociation CO stays in the docking site of Ngb, located about 3 Å above the porphyrin ring on the distal site,¹⁰⁶⁻¹¹¹ lacking so far an explanation as to why the ligand is obviously trapped in the docking site. Using the same computational protocol as in previous studies i.e., combining QM/MM with LMA, we investigated in this work the strength of the CO and FeC bonds in hexacoordinate NgbCO and that of the CO bond in pentacoordinate NgbCO,

in the native NgbCO and its mutations. We also addressed the experimental observation that after dissociation the vibrational CO frequency measured in the binding pocket of Ngb is redshifted compared to that of free CO, depending on the position in the active site and the protein mutation.^{106–111}

METHODOLOGY

Geometries, vibrational frequencies and molecular properties were calculated at the QM/MM level of theory i) for the hexacoordinate NgbCO complexes, i.e. complexes where the CO is bonded to the Fe of the heme group and ii) for two forms **A** and **B** where the heme Fe is pentacoordinated and CO dissociated from Fe as depicted in Figure 1. Calculations were performed for the wild protein based on three tautomers of distal histidine **H64 δ** , **H64 ϵ** and **H64 $\delta\epsilon$** (see 1), and nine protein mutations **F106A**, **F28L**, **F28W**, **H64A**, **H64Q**, **H64V**, **K67A**, **K67R**, and **K67T** which have been previously investigated spectroscopically.^{109,112–117} For these mutations both tautomeric forms of distal histidine **H64 δ** and **H64 ϵ** were considered, leading to **F106A δ** , **F106A ϵ** , **F28L δ** , **F28L ϵ** , **F28W δ** , **F28W ϵ** , **K67A δ** , **K67A ϵ** , **K67R δ** , **K67R ϵ** , **K67T δ** and **K67T ϵ** . Additionally we included in our set **H64A**, **H64Q**, and **H64V** protein mutations, where the distal histidine **H64** is replaced by another amino acid, which led to 18 protein systems in total. Results of all protein systems investigated in this study are reported in alphabetical order.

The experimental x-ray structure of carboxy murine Ngb¹¹⁸ (PDB entry: 1W92) served as a starting point for all calculations, from which the protein mutations investigated in this study were created by manual modification. Hydrogen atoms were placed according to the standard AMBER¹¹⁹ computational procedure, the protein system was then neutralized via the addition of two Na⁺ counter-ions. To mimic a water environment of the enzyme the heme center was surrounded by a TIP3P¹²⁰ water sphere of a 16 Å radius. The protein system was initially minimized with AMBER molecular mechanics. After minimization the protein system was divided into QM and MM parts, where the QM part included the heme group with CO, heme substituents, and both the distal (H64) and proximal (H93) histidine and the MM part included the rest of the protein system. Chemical bonds between the H64 and H93

histidine side chains and the protein backbone were cut and free valences were fulfilled via hydrogen atoms. The QM/MM calculations started from geometry optimization with mechanical embedding for the QM singlet electronic state, followed by geometry optimization with scaled electronic embedding using a PBE0/6–31G(d,p)/AMBER model chemistry,^{121,122} which was then followed by geometry optimization with scaled electronic embedding using the ω B97X–D/6–31G(d,p)/AMBER model chemistry^{122,123} applying ONIOM.¹⁰⁰ The initial calculations with a PBE0 functional provided better SCF convergence for subsequent calculations with the ω B97X–D functional. This functional was selected as it includes dispersion correction, which is important for the calculations of the dissociated CO forms **A** and **B** and because of its overall best performance in reproducing the molecular structure for a series of transition metals complexes derived from single–crystal x–ray diffraction experiments.¹²⁴ For all investigated protein systems geometry optimizations led to local minima on the potential energy surface, which was confirmed via harmonic vibrational frequency calculation reporting no imaginary frequencies. Anharmonicity effects were covered via the use of a scaling factor (0.9501) leading to an average decrease of the harmonic vibrational frequencies by about 5%.^{125,126} This applies also to the local mode frequencies in Table 4. The model Fe-porphyrin in the gas phase (**Gas**, see Figure 1) with CO bonded to the heme group was calculated with ω B97X–D/6–31G(d,p).

Calculations for the dissociated CO, i.e., NgbCO complexes with a pentacoordinated heme group, were performed using the same computational procedure as for the calculations of hexacoordinated NgbCO complexes(as described above). Initial CO positions in the active site of the wild type protein and the protein mutations were created manually in accordance with previous experimental and theoretical studies of NgbCO (see Figure 1).^{30,34,109,114} As suggested in the literature, the dissociated CO preferentially adapts two different orientations, where either CO is located perpendicular to the plane defined by the distal histidine (form **A**) or parallel to that plane (form **B**). Furthermore, in each of the two docking forms **A** and **B** CO can be placed initially in two opposite orientations before the geometry optimization, one where the C atom is closer to the heme center (denoted in our study as 1) and one where the O atom is closer to the heme center (denoted as 2) which led to 36 protein systems denoted in this work as: **F106A δ 1**, **F106A δ 2**, **F106A ϵ 1**, **F106A ϵ 2**,

F28L δ 1, **F28L δ 2**, **F28L ϵ 1**, **F28L ϵ 2**, **F28W δ 1**, **F28W δ 2**, **F28W ϵ 1**, **F28W ϵ 2**, **H64A1**, **H64A2**, **H64Q1**, **H64Q2**, **H64V1**, **H64V2**, **H64 δ 1**, **H64 δ 2**, **H64 ϵ 1**, **H64 ϵ 2**, **H64 $\delta\epsilon$ 1**, **H64 $\delta\epsilon$ 2**, **K67A δ 1**, **K67A δ 2**, **K67A ϵ 1**, **K67A ϵ 2**, **K67R δ 1**, **K67R δ 2**, **K67R ϵ 1**, **K67R ϵ 2**, **K67T δ 1**, **K67T δ 2**, **K67T ϵ 1**, and **K67T ϵ** , where **1** and **2** denote complexes obtained from different starting orientations. The actual geometries after geometry optimization with the corresponding protein mutation labels are presented in the Supplementary Material.

The QM/MM calculations for the dissociated CO (pentacoordinated heme group) were performed in our study for the quintet ground electronic state, which is consistent with previous experimental and theoretical studies.^{127–129} As a test we optimized the geometry of the **H64 ϵ 1** system for the triplet electronic state (including QM and MM parts) and found the total energy which is 15 kcal/mol higher than that of the quintet state which confirms that the ground electronic state is a quintet. Thus, all pentacoordinated heme calculations in this study were performed for the quintet state. Geometry optimization calculations of all protein systems with CO in the docking site were completed when a local minimum on the potential energy surface was reached, which was confirmed via vibrational frequency calculations, no imaginary frequencies were found. Notably, in our QM/MM geometry optimization confirmed the previously suggested preferred locations of the dissociated CO inside the enzyme pocket;¹⁰⁹ in 25 protein systems CO occupied the docking site **A**, and for the remaining 11 protein systems studied in this work CO occupied the docking site **B** (see Figure 1). Snapshots of all optimized QM parts investigated in this study can be found in the Supplementary Material. In addition, we calculated CO and other reference molecules such as Fe(CO)₅ in the gas phase, i.e., excluding the protein environment, using the ω B97X–D/6–31G(d,p) level of theory.

The idea of characterizing a chemical bond via the stretching force constant dates back to the 1920s and 1930s and includes the well-known *Badger rule*, an inverse power relationship between bond length and stretching force constant.¹³⁰ While this rule works fine for diatomic molecules, its extension to polyatomic molecules happens to be a major obstacle^{95,131} as normal vibrational modes tend to delocalize over the molecule rather than being localized in a specific bond.^{132,133} This disqualifies the related normal mode stretching force

constants as a suitable bond strength measure. LMA, originally suggested by Konkoli and Cremer,^{91–93,134,135} has solved this problem which has led to a new measure of the intrinsic strength of a chemical bond or weak chemical interaction and a generalized Badger rule¹³¹ being based on local mode force constants. A comprehensive discussion of the underlying theory of LMA, following two independent routes to derive local vibrational modes, is given in Ref. 95. Local mode force constants, contrary to normal mode force constants are independent of the choice of the coordinates used to describe the molecule in question.^{131,136,137} They are sensitive to differences in the electronic structure (e.g. caused by changing a substituent), and because they are, in contrast to frequencies, independent of the atomic masses, they capture pure electronic effects. In their landmark paper, Zou and Cremer¹³⁸ proved that the local stretching force constant $k_n^a(\text{AB})$ reflects the intrinsic strength of the bond/interaction between two atoms A and B being described by an internal coordinate q_n . In essence, LMA has advanced as a powerful analytical tool, extensively applied to a broad range of chemical systems from simple molecular systems to systems in solution^{139,140} to proteins^{102,141} accounting for both covalent bonds^{79,88,131,138,142–152} and non-covalent interactions^{89,150,153–166} including hydrogen bonds.^{167–177} Recently, a whole new scope of chemical systems were unlocked with the extension of LVM theory to periodic systems.¹⁷⁸

For the comparison of larger sets of k^a values, the use of a relative bond strength order (BSO) is more convenient. Both are connected via a power relationship of the form $BSO = A * (k^a)^B$ according to the generalized Badger rule derived by Cremer, Kraka and co-workers,^{131,147} where parameters A and B are obtained from two reference molecules with known BSO and k^a values and the request that for a zero force constant the BSO is also zero. In our study CH_3OH and CH_2O were used as the reference systems (CH_3OH : BSO=1, $k^a=5.286$ mDyn/Å, CH_2O : BSO=2, $k^a=14.195$ mDyn/Å) leading to A and B values of 0.3109 and 0.7017. For the transition metal-carbon bonds, Fe-C and Fe=C, we used Cu- CH_3 and Ni= CH_2 as a reference systems. The Mayer bond order^{179–181} of the single bond in Cu- CH_3 has a value of 0.992, while the Mayer bond order of the double bond in Ni= CH_2 has a value of 1.871, (corresponding k^a values are 3.296 and 5.455 mDyn/Å, respectively). Scaling the Mayer bond order of Cu- CH_3 to the value 1.0, the new bond order of Ni= CH_2 has a value of 1.886, which leads to the values $A = 0.2225$ and $B = 1.2598$ of the power relationship

between BSO and k^a for the FeC bond used in our study. For hydrogen bonding between CO and hydrogen atoms of the ϵ form of distal histidine, we used the HF molecule and the F_2H^- anion as a references (HF: BSO=1, $k^a= 9.456$ mDyn/Å, F_2H^- : BSO=0.5, $k^a= 1.203$ mDyn/Å) leading to parameters $A = 0.4699$ and $B = 0.3362$. All references were calculated at the ω B97X-D/6-31G(d,p) level of theory.

The quantum theory of atoms-in-molecules (QTAIM) developed by Bader¹⁸²⁻¹⁸⁵ presents a theoretical scheme for identifying, analyzing and characterizing chemical bonds and interactions via the topological features of the total electron density $\rho(\mathbf{r})$. In this work we used QTAIM as a complementary tool to the local mode analysis to determine the covalent character of the FeC and CO bonds via the Cremer-Kraka criterion¹⁸⁶⁻¹⁸⁸ of covalent bonding. The Cremer-Kraka criterion is composed of two conditions; necessary condition: (i) existence of a bond path and bond critical bond critical point ρ , i.e., (3,-1) saddle point of electron density $\rho(\mathbf{r})$ between the two atoms under consideration; (ii) sufficient condition: the energy density $H\rho$ at that point is smaller than zero. $H(\mathbf{r})$ is defined as:

$$H(\mathbf{r}) = G(\mathbf{r}) + V(\mathbf{r}) \tag{1}$$

where $G(\mathbf{r})$ is the kinetic energy density and $V(\mathbf{r})$ is the potential energy density. A negative $V(\mathbf{r})$ corresponds to a stabilizing accumulation of density whereas the positive $G(\mathbf{r})$ corresponds to depletion of electron density¹⁸⁷. As a result, the sign of H_c indicates which term is dominant.¹⁸⁸ If $H\rho < 0$, the interaction is considered covalent in nature, whereas $H\rho > 0$ is indicative of electrostatic interactions. In addition QTAIM we used the natural bond orbital (NBO) population analysis of Weinhold and co-workers^{189,190} in order to obtain atomic charges and the total population of antibonding orbitals (POP). All geometry and frequency calculations were performed with Gaussian09¹⁹¹, LMA properties were calculated with the LmodeA program,¹⁹² NBO atomic charges and antibonding populations were computed using the NBO program,¹⁹³ and the electron and energy density at the bond critical points¹⁹⁴ were obtained with AIMALL.¹⁹⁵

RESULTS

CO, FeC, and hydrogen bonding in hexacoordinate NgbCO complexes

CO bonding Table 1 presents CO bond properties of CO being coordinated to the heme Fe atom, the CO bond length R , the local mode force constant k^a , the bond strength order BSO, the local mode frequency ω^a , the electron density ρ and the energy density $H\rho$ taken at a bond critical point, the NBO atomic charges on C and O, and the total population POP of antibonding orbitals. In Figure 2 correlations between these properties are illustrated. Figure 2a shows the BSO values of the CO bond for in the hexacoordinate wild type NgbCO and its mutations, calculated from the power relation ship described above. The figure clearly reveals that the strength of the CO bond in the enzyme is generally smaller than the strength of this bond in the gas phase heme model **Gas** compared to that of an isolated CO molecule (BSO values are **Gas** = 2.420, isolated CO = 2.577, respectively, corresponding k^a values are 18.614 and 20.376 mDyn/Å, respectively). According to Figure 2a, the strongest CO bond in hexacoordinate NgbCO is observed for the **K67R δ** protein mutation ($k^a=18.620$ mDyn/Å), which almost matches the strength of the gas phase heme model **Gas**. In the **K67R δ** protein mutation, the distal histidine H64 is moved away from the heme group (see Figure S3 of the Supplementary Material) creating a relatively big distal heme pocket, which leads to the similar bond strength of CO in these two systems. Additionally, for the δ tautomer form of H64, the hydrogen atom of the distal histidine is on the opposite side of this residue side chain, which eliminates the interaction between the hydrogen atom and CO, responsible for a smaller strength of CO in the heme pocket. The weakest CO bond is observed for the **H64 $\delta\epsilon$** tautomer ($k^a=17.339$ mDyn/Å), which is the protonated form of this residue (see Figure S2 of the Supplementary Material). The **H64 $\delta\epsilon$** protein system is positively charged, in contrast to the other protein system investigated in this study, which makes the interaction between the H ϵ hydrogen atom of histidine with CO much stronger, leading to a weaker CO bond in the heme pocket. Among the neutral hexacoordinate NgbCO complexes, the weakest CO bond is found for **F28W ϵ** ($k^a=17.915$ mDyn/Å), where CO is oriented close to both tryptophan W28 and distal histidine H64 in its ϵ tautomer form (see Figure S1 of the Supplementary Material). The big side chain of tryptophan W28 occupies the distal heme

pocket making a relatively strong interaction with the CO oxygen atom, leading to a small strength of the CO bond.

Figure 2b shows a relatively good correlation ($R^2 = 0.9387$) between the CO force constant k^a and its bond length R . For the systems studied in this work we observe the general trend that a stronger CO bond has a shorter CO bond length (e.g., $R = 1.1453$ and 1.1441 Å for the **Gas** and **K67R δ** , protein system, respectively), and the weakest CO bond for the **H64 $\delta\epsilon$** tautomer is also the longest ($R = 1.1150$ Å). As shown in Figure 2c, there is also a relatively good correlation ($R^2 = 0.9568$) between the local force constant $k^a(\text{CO})$ and the energy density $H\rho$. According to a Cremer–Kraka criterion of a bond character,^{186,187} described above, a more negative value of $H\rho$ indicates a stronger covalent bond character. All CO bonds investigated are covalent in nature with the strongest covalent CO bonds found for both the **Gas** phase and **K67R δ** protein systems with -0.7710 and -0.7726 Hartree/Bohr³, respectively while the weakest CO bonds, of **H64 $\delta\epsilon$** and **F28W ϵ** , acquire much smaller negative energy density values of -0.7506 and -0.7592 Hartree/Bohr³, respectively. Therefore, Figure 2c indicates that the different strength of the CO bond in hexacoordinated NbgCO is dominated by electronic effects rather than electrostatic effects as revealed by $H\rho$.

Figure 2d compares the CO force constant k^a with the CO antibonding NBO orbital population. Although the correlation is not perfect ($R^2 = 0.7176$), it reflects the general trend that a stronger CO bond is related to a smaller antibonding NBO orbital population, which is inline with the DC model explaining M-L bond weakening by the π -back donation between CO and Fe of the heme group. The strongest CO bond in the **K67R δ** protein system has the smallest NBO antibonding population (0.3522 e), while the weakest CO bond in the **H64 $\delta\epsilon$** protein system, has the largest NBO antibonding population (0.4630 e). Figure 2e shows the relation between the CO force constant k^a and the NBO atomic charge on the O atom of of CO. According to Figure 2e, the strongest CO bond in the **K67R δ** protein system, has the less negative the NBO atomic charge of the CO oxygen atom (0.4573 e), while the weakest CO bond in **H64 $\delta\epsilon$** , has the most negative oxygen charge (-0.5222 e). These results indicate that the electrostatic field of the enzyme active site pocket, along with charge transfer and dispersion interactions between the CO molecule and side chains of amino acids in the heme pocket, modify the CO oxygen charge, influences the electron

density transfer between Fe and CO bond and in this way modulates the strength of the CO chemical bond. The same picture is revealed by Figure 2f which correlates CO force constant k^a and the sum of NBO charges on C and O. It is interesting to note that compared to the neutral CO molecule with (charge on O = -0.5145 e and on C = +0.5145 e), the CO ligand positively charged in all hexacoordinate NgbCO complexes, investigated in the work ranging from +0.0615 for e for **H64 δ** to +0.1296 e in **K67R δ** .

FeC bonding Table 2 summarizes FeC bond properties of the hexacoordinate NgbCO complex. The graphic representations of the relations between the FeC local mode force constant k^a and the other bond properties are presented in Figure 3. FeC bond orders range from 0.657 (**F28W δ**) to 0.991 (**H64Q**). As a comparison, the axial and equatorial FeC bond in Fe(CO)₅ have BSO values of 0.802 and 0.889, respectively which lies in the middle of the complexes investigated in this work. The strongest FeC bond is observed in our study for the **H64Q** protein mutation ($k^a=3.273$ mDyn/Å), where the distal histidine H64 is replaced by glutamine (see Figure S2 of the Supplementary Material), and the weakest FeC bond is found for **F28W δ** (2.363 mDyn/Å). Figure 3b shows the same general trend as observed for the CO bonds, the stronger FeC bond is related to a shorter FeC bond. **F28W δ** has the longest FeC bond of our set (R=1.7841 Å), **H64 δ** has the shorted (R=1.77464 Å). In comparison, the axial and equatorial FeC bonds in Fe(CO)₅ have bond lengths of R=1.8020 and 1.7891Å, respectively. As for the CO bonds, there is a general trend for the FeC bonds that the stronger FeC bond has a more covalent character as reflected by $H\rho$ values depicted in Figure 3c. Figure 3d correlates the FeC and the CO bond strengths. We find the same trend as in our previous study on MbCO complexes¹⁰² that a stronger the FeC bond is connected with a weaker CO bond and vice versa.

Hydrogen bonding The data presented in Table 3 and Figure 4 are devoted to hydrogen bonding interactions between the CO ligand and hydrogen atoms in the side chain of His64. These interactions are possible in **F106A ϵ** , **F28L ϵ** , **F28W ϵ** , **H64 ϵ** , **H64 δ** ϵ , **K67A ϵ** , **K67R ϵ** , and **K67T ϵ** . They involve O \cdots H-N hydrogen bonding between the O atom of CO and the HN hydrogen of His64, depicted in Fig 4a, non-classical hydrogen bonding¹⁹⁶ O \cdots H-C between the O atom of CO and the HC hydrogen of His64 depicted in Fig 4b and non-

classical hydrogen bonding $C \cdots H-N$ between the CO carbon and the HN hydrogen in His64 depicted in Figure 4c. According to Table 3, the length of the $O \cdots H-N$ hydrogen bond is in a range between 2.8 and 3.2 Å with local mode force constant values between 0.03 to 0.06 mDyn/Å, the smallest found for **K67T**ε ($k^a=0.031$ mDyn/Å) and the largest for **K67R**ε ($k^a=0.068$ mDyn/Å). Compared with the corresponding hydrogen bond in water imizadole dimer ($R = 1.922$ Å, $k^a = 0.205$ mDyn/Å, $BSO = 0.358$, $H\rho = 0.001$ Hartree/Bohr³ these $O \cdots H-N$ interactions are weak, mainly caused by the larger distance. No bond critical points were found.

Although non-classical in nature,¹⁹⁶ the $O \cdots H-C$ hydrogen bonds are stronger, which is reflected by force constant values in a range between 0.05 and 0.10 mDyn/Å, with the strongest interaction found for **H64**δε ($k^a=0.102$ mDyn/Å). Bond critical points were found in these cases with small positive $H\rho$ values in a range between 0.0009 and 0.0011 Hartree/Bohr³, denoting that these interactions are of weak electrostatic character. The non-classical $C \cdots H-N$ interactions are the strongest of the three types with the force constants in a range between 0.05 and 0.13 mDyn/Å as shown by the data in Table 3 and Figure 4. The smaller strength of the classical $O \cdots H-N$ hydrogen bonds is due to a specific orientation of the distal histidine relative to CO coordinated to the heme group, in which the N-H group of the histidine side chain is oriented closer to the CO carbon atom rather than the CO oxygen atom, making $C \cdots H-N$ dispersion interactions stronger than $O \cdots H-N$ hydrogen bonding. As pointed out in our previous study on hydrogen bonding in bases pairs,¹⁷⁷ non-classical hydrogen bonding plays a non-negligible role in biochemical systems, and depending on the topology of the system they can be stronger the classical hydrogen bonds.

CO vibrational frequencies data in hexa- and pentacoordinate NgbCO complexes

In the next section calculated local mode frequencies ω^a for FeC and CO bonds in hexa- and pentacoordinated NgbCO complexes will be discussed for which experimental normal mode frequencies ω_{exp} are available summarized in Table 4. The calculated local mode frequencies ω^a are scaled by a factor 0.9501 to account for anharmonicity effects, where this factor was selected according to the computational level of theory used in this study.¹²⁵ One caveat

is appropriate. For the hexacoordinate NgbCO where CO is bonded to Fe, the calculated CO local mode frequencies can not be directly compared with normal mode frequencies, because the latter couple with other vibrations, as outlined above. As a consequence, values of the normal mode frequencies tagged as CO stretching tend to be considerably smaller, as we showed e.g., for the water dimer¹⁶⁷ or the formaldehyde dimer.¹⁶⁸ Therefore, the local mode FeC and CO stretching frequencies presented in Table 4 are larger than their normal mode counterparts, however, general trends can be discussed. We have developed a unique way to decompose normal modes into local mode contribution.^{63,91,93,95,135} Work is in progress to extend the method for the analysis of large QM/MM systems.^{197,198} For the CO vibration of pentacoordinate NgbCO complexes, where the CO is dissociated from the ferrous heme group, there normal mode coupling can be neglected and therefore, the CO local mode vibrational frequency is closely related to the CO normal mode frequency, as reflected by the data in Table 4 which also confirms the quality of our chosen model chemistry.

According to Table 4, in the **H64A** and **H64V** protein systems, in which the distal histidine H64 is replaced by alanine and valine, only one FeC stretching vibration is observed experimentally (494 cm⁻¹),^{112,113} such is consistent with our results which predict only one FeC vibration. However, for the **H64A** protein mutation the local mode frequency of this vibration has a value of 633 cm⁻¹, while for the **H64V** the frequency of this local mode is 677 cm⁻¹. For the wild protein, where the distal histidine H64 can be in two tautomeric forms **H64 δ** and **H64 ϵ** , there are two (494 and 521 cm⁻¹),¹¹² or three (494, 505, and 521 cm⁻¹)¹¹³ vibrations observed experimentally, however the spectral band related to the 505 cm⁻¹ is not fully developed.¹¹³ Therefore, the two intense normal modes observed experimentally, can be assigned to two tautomer forms of the distal histidine H64, which, in our study, the local mode frequencies calculated for the **H64 δ** and **H64 ϵ** have values of 642 and 701 cm⁻¹, respectively. Similarly, there are two main spectral bands observed experimentally (494 and 521 cm⁻¹),¹¹³ for the K67A, K67R and K67T protein mutations, which correspond to the two protein systems **H64 δ** and **H64 ϵ** investigated in our study. Calculated local mode frequencies are reported in Table 4.

The experimental normal mode frequencies for CO bonded to Fe in the F28L protein mutation, have values 1937, 1957, and 1967 cm⁻¹.¹⁰⁹ According to our calculations, there

are two CO local mode vibrations with frequencies of a value of 2014 and 2018 cm^{-1} for the **F28L δ** and **F28L ϵ** protein systems. The similar relation between the three experimental normal modes and the two theoretical local modes of CO in the hexacoordinated active site, are observed in our study for the F28W, H64, K67A, K67R, and K67T protein systems, which indicates that the CO local mode can participate in two different normal modes in one of the H64 tautomers. For H64A protein mutation, two normal modes with the frequencies of a value of 1932 and 1972 cm^{-1} have been observed previously,¹¹³ it is noted that we observed similar frequency values for the H64V protein mutation. According to our calculations of the **H64A** and **H64V** protein systems, there is only one CO local mode with a frequency of a value of 2023 and 2027 cm^{-1} for H64A and H64V, respectively. Therefore, similarly as in the protein systems with the δ and ϵ forms of H64, the CO local mode can participate in two normal modes.

Comparing the CO local modes of the hexacoordinated Ngb and the pentacoordinated Ngb species we observe an interesting change in the local mode frequencies. For the **F28L δ** protein system, the frequency of the CO local mode in the hexacoordinated form has a value of 2003 cm^{-1} , while the frequency of this local mode in the pentacoordinate docking site has two values 2100 and 2118 cm^{-1} which depends on the orientation (1 or 2) of CO relative to the heme center. Therefore, there is a 97 cm^{-1} red-shift of the frequency for conformer **A** with the CO carbon atom close to Fe, and a 115 cm^{-1} red-shift for conformer **B** with the CO oxygen atom close to Fe (see Figure S9 of the Supplementary Material). These results are consistent with a weaker CO bond in **F28L δ 1** (19.875 mDyn/Å) than in **F28L δ 2** (20.305 mDyn/Å). Reported in Table 4 are similar red-shifts, which have been observed in the other protein systems.

There are six spectral bands observed experimentally for CO in the pentacoordinate docking site of the F28L protein mutation, 2107, 2120, 2130, 2136, 2142, and 2149 cm^{-1} where the bands, 2120 and 2130 cm^{-1} , have dominate intensities (39 and 24 %, respectively).¹⁰⁹ Previous work involving neutron crystallography confirmed that the ϵ tautomer form of the distal histidine is the dominant conformer of this residue in heme proteins.¹⁹⁹ Assuming that the ϵ tautomer of H64 dominates in Ngb, we can directly compare these two experimental frequencies with the calculated local mode frequencies of the **F28L ϵ 1** and **F28L ϵ 2** protein

system (2122 and 2125 cm^{-1}). They are in good agreement to experimental values within a few wave numbers. A similar agreement is observed for the F28W protein mutation (2128 and 2136 cm^{-1} from experiment,¹⁰⁹ and 2111 and 2130 cm^{-1} from our calculations), and for the wild type protein H64 (2129 and 2136 cm^{-1} from experiment,¹⁰⁹ and 2116 and 2127 cm^{-1} from our calculations). Overall, our results show good agreement between calculated and experimental frequencies of the CO molecular vibration in the pentacoordinate NgbCO complexes, confirming that the computational approach presented in this investigation, provides a deep and precise inside into the CO bond properties in this heme protein.

CO bonding in pentacoordinate NgbCO complexes

Form A Table 5, shows the results of our calculations for the CO molecule in the pentacoordinate docking site **A** of NgbCO which is located in our calculations on the opposite side of the heme pocket, relative to the propionic acid side chains, above the heme ring. As an example, for the **H64 ϵ 1** protein system, the carbon atom of CO in the docking site **A** is located 4.366 Å away from Fe, and 3.150 Å away from the closest carbon atom of the heme group. The CO oxygen atom is located 4.731 Å away from Fe, and 3.205 Å away from the closest carbon atom of the heme group. For the other protein systems investigated in this study, CO is located similarly in the docking site **A** (see Figures S4 – S8 of the Supplementary Material). CO is generally oriented more or less perpendicular to the ring site chain of the distal histidine H64. As expected and confirmed by the data in Table 5 and Figure 5a the dissociated CO bonds are considerably stronger as the CO bonds in the hexacoordinate NgbCO complexes. The strongest CO bond is found in the protein systems **F28W ϵ 2** and **F28W δ 2** ($k^a = 20.308$ and 20.305 mDyn/Å, respectively) where the CO bond strengths are in close proximity to that of an isolated CO molecule in the gas phase ($k^a = 20.376$ mDyn/Å).

Although the correlation between k^a and other molecular properties (see Figures 5b - 5f show general trends, they are weaker than for the CO bond of the hexacoordinate NgbCO complexes investigated in this study. Both the **F28W ϵ 2** and **F28W δ 2** protein systems show the smallest CO bond length values of 1.135 Å, the strongest covalent character with $H\rho = -0.768$ Hartree/Bohr³) and smallest negative charges for the O atom of CO (-0.4873

and -0.449 e) in this series. According to Table 5 the NBO antibonding populations of CO of the pentacoordinated docking site **A** are much smaller than those of the CO ligands in the hexacoordinate NgbCO complexes. In the hexacoordinated active site there is charge transfer between CO and Fe via a chemical bond, however for CO in the pentacoordinated docking site **A** the charge transfer is restricted to through space interactions between CO and the heme group and/or with the distal histidine H64. It is noted that through-space charge transfer events have been experimentally observed in heme protein systems,²⁰⁰ π -stacking systems,^{201,202} and transition metal complexes.²⁰³ Moreover, Ngb has been suggested to be an electron transfer species.²⁰⁴ According to our calculations, the **F28W δ 2** protein system has the smallest NBO antibonding population of CO in the docking site A (0.0047 e), while in **F28W ϵ 2** this population is slightly larger (0.007 e). In both the **F28W ϵ 2** and **F28W δ 2** protein systems, the big side chain of tryptophan W28, located between the distal histidine H64 and the docking site **A** (see Figure S6 in Supplementary Material), isolates CO from the interaction with distal histidine H64 and diminishes the through-space charge transfer from distal histidine resulting in the CO bond being relatively strong. For the protein systems **F28W ϵ 1** and **F28W δ 1**, with an opposite orientation of CO in the docking site (see Figure S4 of the Supplementary Material), we observe weak CO bonds (19.950 and 19.875 mDyn/Å) with long bond lengths (1.1374 Å and 1.1369 Å, respectively) smaller covalent character ($H\rho$ values of -0.7607 and -0.7612 Hartree/Bohr³), increased negative CO oxygen charges (-0.5422 and -0.5323 e), and larger antibonding populations (0.0080 and 0.0102 e), respectively, in comparison with the CO bond properties of the **F28W ϵ 2** and **F28W δ 2** protein systems.

In the **F28W ϵ 1** and **F28W δ 1** protein systems, the CO oxygen atom is in close contact with the hydrogen atom of the N-H group of the side chain in tryptophan W28 (2.148 and 2.381 Å, respectively), and the interaction between CO and the side chain of W28 makes the CO chemical bond much weaker. The weakest CO bond in this work, relevant to this series, was observed for the **K67A δ 1** protein mutation (19.829 mDyn/Å), however the small strength of this chemical bond is not well correlated with its bond length (1.1361 Å), energy density (-0.763 Hartree/Bohr³), the CO oxygen charge (-0.517 e) or its antibonding population (0.010 e) as depicted in Figure 4 where the **K67A δ 1** protein system appears as an outlier. For the corresponding protein mutation **K67A δ 2**, where CO is located in an

opposite orientation (see Figure S7 of the Supporting Material), the CO bond has a medium strength and medium bond properties, which indicates that the smallest CO strength in **K67A δ 1** is related to specific dispersion interactions between CO and the heme group as well side chains of the heme pocket in the docking site **A**.

Form B The results of our calculations for CO bond properties in the pentacoordinated docking site **B** of Ngb, are presented in Table 6. The CO molecule is located in the docking site of **B** in different manner compared to site **A**. For the **F106A ϵ 1** protein system the distance between the O atom of the carbonyl (i.e. CO) and Fe is 3.630 Å, the distance between the CO carbon atom and Fe is 4.149 Å, the distance between the O atom of CO and the proximal nitrogen atom of the heme group is 3.057 Å, and the distance between the CO carbon atom and the nearest carbon atom of the heme group is 3.274 Å. Figures S9 – S10 of the Supplementary Material show snapshots of the selected residues in the docking site **B** of pentacoordinate NgbCO complexes, suggesting that CO is generally oriented more or less parallel to the ring site chain of the distal histidine H64.

The relationships between the CO local mode force constant and the other CO bond properties in the docking site **B** are presented graphically in Figure 6. According to Figure 6a, the **H64A1** protein mutation acquires the strongest CO bond (20.270 mDyn/Å) which has a relatively short bond length (1.135 Å), an energy density $h\rho$ in a middle range (-0.768 Hartree/Bohr³), a small antibonding population (0.006 e), and a charge, on the oxygen atom, that is in the middle range (-0.509 e). For the **H64A1** protein mutation the distal histidine H64 is replaced by alanine (see Figure S9 of the Supplementary Material), making a relatively big distal heme pocket in which the interactions between CO and the side chains of the distal pocket are small. The CO bond in the corresponding protein system **H64A2**, where CO occupies the docking site **B** in an opposite orientation (see Figure S10 of the Supplementary Material), has a smaller strength (20.135 mDyn/Å), a longer bond length (1.1359 Å), a slightly more negative energy density (-0.769 Hartree/Bohr³), a larger antibonding population (0.017 e), and a less negative oxygen charge (-0.500 e), indicates on a stronger charge transfer via space between Fe the CO carbon atom compared to the charge transfer between Fe and the CO oxygen atom in the **H64A1** protein mutation. The weakest CO bond observed in our work is that of the **F28L δ 1** protein system (19.743 mDyn/Å),

which has a big bond length (1.137 Å), a relatively more negative energy density (-0.770 Hartree/Bohr³), a relative big antibonding population (0.027 e), and the oxygen charge in a middle range (-0.502 e). The **F28Lδ1** protein mutation has phenylalanine F28 replaced by leucine, orienting the CO carbon atom almost on the top of Fe (see Figure S9 of the Supplementary Material), which indicates on a strong charge transfer via space between Fe and the CO carbon atom, and a weak character of the CO bond. For the corresponding **F28Lδ2** protein mutation (see Figure S10 of the Supplementary Material), the CO bond has a bigger strength (20.078 mdyn/Å), a smaller bond length (1.135 Å), a less negative energy density (-0.768 Hartree/Å), a smaller antibonding population (0.012 e) and a more negative oxygen charge (-0.515 e), indicating a smaller through-space charge transfer between the CO oxygen and Fe atoms in contrast to that of the **F28Lδ1** protein mutation.

Comparison with MbCO

Myoglobin (Mb) has been used as a reference system to understand protein function since early structural work.²⁰⁵ Particularly important was to understand the interaction between diatomic ligands and the surrounding Mb active site pocket based on experimental X-ray structures for bonded ligands^{206,207} as well for dissociated ligands.²⁰⁸

For bonded ligands in the wild type sperm whale MbCO, four different substates related to CO stretching vibrations denoted as A₀, A₁, A₂ and A₃ (with the frequencies of a value of 1965, 1947, 1942, and 1932 cm⁻¹, respectively) were observed.²⁰⁹⁻²¹² The protein substates were assigned to His64–open and closed conformational states as well as local effects of the amino acids, different sizes of the cavities, or certain interactions between His64 and FeCO.²¹³⁻²¹⁷ It has been suggested that the differences observed amongst the various CO stretching conformers stem from changes in the electrostatic interaction between the heme pocket and the CO that is coordinated to the Fe atom of the heme group.²¹⁸ Moreover, because H64 can take on two tautomer forms, the interaction with the CO ligand coordinated to the Fe varies, and which tautomer form is the most stable often remains unclear.^{219,220} From exchange kinetics between the A₁ and A₃ states it has been concluded that these

states vary due to differences in the conformation of H64 in the distal pocket in which A_1 represented a weaker interaction between H64 and CO in contrast to A_3 .^{221,222} For MbCO it is widely recognized that substates A_0 and A_1 correspond to the His64–open and closed conformational states while the origin of substate A_3 has been a highly debated subject; some investigations having proposed the following to explain its origin: i) the local field–effect enabled by amino acids, ii) differing size of the cavities, iii) certain interactions between His64 and FeCO.^{213–216} However, more recent work has revealed that the origin of A_3 , as well as A_0 and A_1 , is solely attributed to the H–bond interaction between FeCO and His64 that, in turn, result in the His64–open and closed conformations; the free energy difference between the two conformations was found to be between 4 and 5 kJ/mol.²¹⁷ It also has been speculated that different protein conformations of varying substates, various mutations, and the presence of solvent near prosthetic groups may govern the time it takes for the excited state of the CO stretching mode in MbCO to reach the lowest excited vibrational state.^{223,224}

Other studies have suggested that strengthening of the FeC bonding in MbCO and its mutations leads to a weakening of the CO bonds where this inverse relationship is said to originate from the π –back donation effect from the occupied d_π orbitals of Fe into the vacant π^* orbitals of CO.^{225,226} In one of our most recent works we have investigated the bond strength for the FeC and CO bonds in the wild type MbCO and its 17 protein mutations and quantified for the first time the inverse relationship between FeC and CO bond strengths based on a comprehensive local vibrational mode analysis.¹⁰² The strength of the FeC bond is governed by π –back donation between Fe and the CO moiety, where this donation effect can be modified by the electrostatic field of the distal side of the heme pocket polarizing the CO oxygen atom. In addition, according to our investigation the CO bond is weakened by hydrogen bonding occurring between the O atom of CO and the side chains or surrounding amino acids.¹⁰² For CO coordinated to the active site in the wild type NgbCO our calculations reveal that there are two different CO local mode frequencies with scaled values of 2021 and 2013 cm^{-1} , for **H64 ϵ** and **H64 δ** , respectively. It is interesting to note that the strengthening of the FeC bonding in MbCO and its mutations, which leads to a weakening of the CO bonds,^{102,225,226} is also observed in our calculations of NgbCO.

In MbCO, after dissociation of CO from the heme group the ligand occupies a docking site, originally denoted as site **B**^{227,228} from which the ligand can rebound to the heme group or move to more isolated docking sites in the protein.²²⁹ The experimental vibrational spectrum of the wild type MbCO^{230,231} shows two maxima for CO in the docking site **B**, denoted as **B**₁ with a frequency value of about 2130 cm⁻¹ and **B**₂ with a frequency near 2120 cm⁻¹, i.e., a separation of 10 cm⁻¹. The maxima were originally assigned to two different orientations of the CO ligand relative to Fe, where the Fe···CO orientation corresponds to the **B**₂ peak and the Fe···OC orientation corresponds to the **B**₁ peak.²³² Gas phase calculations for constrained geometries of an active site heme model showed a peak separation of 4 cm⁻¹; however for the N_ε tautomer of distal histidine, there was a reverse ordering of these frequencies.²³³ Later, QM/MM calculations based on molecular dynamics assigned the **B**₁ peak to a Fe···CO ligand orientation and the **B**₂ peak to a Fe···OC orientation with a separation of the normal mode vibrational frequencies of about 10 cm⁻¹,²³⁴ for the N_ε tautomer of distal histidine, which was also confirmed by experimental data.^{107,222} In this work on NgbCO we found two CO docking sites, denoted as site **A** for a location further away from the heme center and as site **B** for a location closer to Fe. For each site we found two local minima, one in which the C atom of CO is closer to Fe and one where the O atom of CO is closer to Fe. In the wild type NgbCO of the N_ε tautomer of distal histidine, for the **H64ε1** conformer in docking site **A** and the Fe···CO orientation, the scaled CO local mode frequency has a value of 2227 cm⁻¹ and the Fe···C distance has a value of 4.366 Å. For the **H64ε2** conformer and the Fe···OC orientation, the scaled CO local mode frequency has a value of 2238 cm⁻¹ and the Fe···O distance has a value of 4.485 Å, i.e there is a small frequency difference of 11 cm⁻¹. The Fe···OC orientation corresponds to the larger local mode CO frequency in agreement with the original assignment of the measured **B**₁ and **B**₂ peaks in MbCO.²³² However, a caveat is appropriate. Local mode frequencies cannot directly be compared with normal mode frequencies because normal modes are generally delocalized and as such involve the movements of all atoms in a molecule, whereas local modes involve only the movements of the individual atoms of the local vibration under consideration. Consequently, one local mode can contribute to several normal modes, as shown in our previous studies (see e.g., Ref. 63) and a normal mode being composed of a

local stretching and bending vibration has therefore a lower normal mode frequency than the corresponding local stretching mode as found e.g., for the hydrogen bond in the water dimer.¹⁶⁷

The smaller local CO frequency in the Fe···CO orientation of the CO ligand can be seen as a result of the smaller Fe···C distance, which leads to a larger charge transfer between these atoms and increased occupation of the CO antibonding orbital (0.0096 and 0.0071 e for the **H64 ϵ 1** and **H64 ϵ 2** conformer, respectively) weakening the CO bond as reflected by the smaller local mode vibration.

We did not find for the wild type of NgbCO a stable structure with CO in the docking site **B** which could correspond to the position of the CO ligand responsible for the B₁ and B₂ peaks observed and calculated in wild type of MbCO discussed above. However, for the F106A ϵ protein mutation of NgbCO in the docking site **B**, which is similar to the wild type of this protein (see snapshots in the Supporting Material), the difference between the calculated local mode CO frequencies of the **F106A ϵ 1** and **F106A ϵ 2** conformers is 10 cm⁻¹, with the Fe···OC orientation having a larger frequency. For the Fe···OC orientation of the **F106A ϵ 1** conformer, the scaled local mode frequency has a value of 2124 cm⁻¹ and the Fe···O distance has a value of 3.630 Å, while for the Fe···CO orientation of the **F106A ϵ 2** conformer, the scaled local mode frequency has a value of 2114 cm⁻¹ and the Fe···C distance has a value of 2.916 Å. Similarly as in the docking site **A** the smaller frequency of the local CO mode in the Fe···CO orientation of the CO ligand is mostly because of the smaller Fe···C distance leading to charge transfer between these atoms and as such to a larger occupation of the CO antibonding orbitals (0.0335 and 0.0055 e, for the **F106A ϵ 2** and **F106A ϵ 1** conformer, respectively) and CO bond weakening. It is important to notice that although Mb and Ngb share similar active sites, the primary amino acid sequence of Ngb is only about 22% identical to that of Mb.² Another difference between these heme proteins is that Ngb has a large cavity (about 290 Å³) which connects the distal and proximal heme sides with the external solvent,^{235,236} and which could be responsible for different geometries and frequencies of CO in the docking sites of Mb and Ngb.

CONCLUSIONS

In this study we have calculated the CO bond properties for the active site of hexacoordinate NgbCO complexes, where CO is bonded to Fe, and in the two pentacoordinated docking sites, where CO is dissociated from the heme group but still trapped in the active site pocket of the protein. The investigations are based on QM/MM calculations of optimal geometries and molecular frequencies for the wild type NgbCO and several protein mutations, using the QM/MM method, the Local Mode Analysis, the QTAIM method, and the NBO analysis.

Our results confirm that the strength of the CO chemical bond in the hexacoordinate NgbCO complexes, depends on the amount of charge transfer between the FeC and CO bonds along the lines discussed in the DC model, which is modulated by the electrostatic and dispersion interactions with side chains of the distal heme pocket. For the CO in pentacoordinated docking sites, where CO is dissociated from the heme group, the strength of the CO bond depends on through-space charge transfer, between the heme group and the CO bond, which is modulated by the electrostatic and dispersion interactions with side chains of the distal heme pocket, similar to hexacoordinate NgbCO complexes, however effects are much smaller. As shown in Figure 7 the overall weakest CO bond is found for **H64 $\delta\epsilon$** which suggests that protein modifications with a positive charge should be envisioned if a weaker CO bond (stronger FeC bond) is desired. The wild type systems **H64 δ** and **H64 ϵ** are in mid-range, and the strongest CO bond in this series is obtained for **K67A δ** . Dissociated CO bonds are substantially stronger and inline with smaller POPs, the weakest CO bond for the dissociated CO forms is found for **F28L δ 1** and the overall strongest CO bond, close to that in isolated CO, is found for **F28W ϵ 2**. In **F28L δ 1** the phenylalanine is replaced with leucine and in **F28W ϵ 2** by a bulky tryptophan. This shows how changes in the active site of the protein influence the strength and nature of CO bonding. In essence, local mode CO force constants and related BSO values can be used to clarify whether a protein modification goes towards the desired direction.

Our work lays out a precise computational approach that involves the derivation of adequate bond strength descriptions, those of which account for cumulative electronic effects, in order to provide details about the CO bond in NgbCO complexes as to gain insight into

the effect of the protein environment on the strength of CO bonds and, in turn, the catalytic activity of heme enzymes. Based on the results of our study, a stronger CO bond is related to a weaker FeC bond, making the CO dissociation process easier. These findings rationalize and reinforce the use of CO as a probe to describe the catalytic activity of heme proteins in general as has been done in many theoretical and experimental studies, also to model the chemical activity of other important signaling molecules interacting with the heme group such as O₂, NO, H₂S, or HNO. Understanding the electronic effects which are responsible for easier CO dissociation processes and using the local mode CO force constants as a convenient probe to monitor changes of the electronic effects during protein modification, will increase our insight into the chemical activity of signaling molecules, and advance our understanding about the biological activity of heme proteins in general.

Future investigations will focus on similar bonding situations of Ngb complexed with other signal molecules, protein modification of Ngb involving the replacement of His64 and His96, as to provide a deeper insight about the role of these two histidines on protein function, and to apply our computational protocol to support the design of modified Ngbs with specific functions, such as the recent suggestion to convert human Ngb into a multifunctional peroxidase²³⁷ or the use of a Ngb derived enzyme to catalyze the synthesis of indigo.²³⁸

ACKNOWLEDGMENTS

This work was financially supported by the National Science Foundation (Grant CHE 2102461). We thank SMU for providing generous computational resources.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article including i) snapshots of the QM part in hexacoordinated neuroglobin, ii) snapshots of the QM part in pentacoordinated neuroglobin in site **A**; and iii) snapshots of the QM part in pentacoordinated neuroglobin in site **B**.

References

1. C. Exertier, L. C. Montemiglio, I. Fred, E. Gugole, G. Parisi, C. Savino, and B. Vallone, *Mol. Aspects Med.* pp. 1–11 (2021).
2. T. Burmester, B. Weich, S. Reinhardt, and T. Hankeln, *Nature* **407**, 520 (2000).
3. T. Burmester and T. Hankeln, *J. Exper. Biol.* **212**, 1423 (2009).
4. E. Luyckx, Z. P. V. Acker, P. Ponsaerts, and S. Dewilde, *Oxid. Med. Cell. Longev.* pp. 5728129.1–17 (2019).
5. A. Pesce, S. Dewilde, M. Nardini, L. Moens, P. Ascenzi, T. Hankeln, T. Burmester, and M. Bolognesi, *Structure* **11**, 1087 (2003).
6. P. Ascenzi, A. di Masi, L. Leboffe, M. Fiocchetti, M. T. Nuzzo, M. Brunori, and M. Marino, *Mol. Asp. Med.* **52**, 1 (2016).
7. M. Fiocchetti, M. Cipolletti, V. Brandi, F. Polticelli, and P. Ascenzi, *J. Mol. Recog.* **30**, e2654 (2017).
8. Y. Sun, K. Jin, X. O. Mao, Y. Zhu, and D. A. Greenberg, *PNAS* **98**, 15306 (2001).
9. Y. Sun, K. Jin, A. Peel, X. O. Mao, L. Xie, and D. A. Greenberg, *PNAS* **100**, 3497 (2003).
10. E. Fordel, E. Geuens, S. Dewilde, W. De Coen, and L. Moens, *IUBMB Life* **56**, 681 (2004).
11. T. T. H. Duong, P. K. Witting, S. T. Antao, S. N. Parry, M. Kennerson, B. Lai, S. Vogt, P. A. Lay, and H. H. Harris, *J. Neurochem.* **108**, 1143 (2009).
12. S. Watanabe, N. Takahashi, H. Uchida, and K. Wakasugi, *J. Biolog. Chem.* **287**, 30128 (2012).
13. G. De Simone, D. Sbardella, F. Oddone, A. Pesce, M. Coletta, and P. Ascenzi, *Cells* **10**, 557 (2021).

14. M. Schmidt, A. Giessler, T. Laufs, T. Hankeln, U. Wolfrum, and T. Burmester, *J. Biol. Chem.* **278**, 1932 (2003).
15. K. Jin, Y. Mao, X. Mao, L. Xie, and D. A. Greenberg, *Stroke* **41**, 557 (2010).
16. K. Jin, X. Mao, L. Xie, and D. A. Greenberg, *Neuroglobin Expression in Human Arteriovenous Malformation and Intracerebral Hemorrhage* (Springer, Ed. J. Zhang and A. Coloha, Vienna, 2011), pp. 315–319.
17. E. André, V. Derrien, P. Sebban, N. Assrir, E. Lescop, and S. Bernad, *JBIC* **24**, 39 (2019).
18. S. de Vidania, I. Palomares-Perez, A. Frank-Garcia, T. Saito, T. C. Saido, J. Draffin, M. Szaruga, L. Chavez-Gutierrez, M. Calero, M. Medina, et al., *Front. Neurosci.* **14**, 1001 (2020).
19. M. Fiocchetti, M. Cipolletti, S. Leone, A. Naldini, F. Carraro, D. Giordano, C. Verde, P. Ascenzi, and M. Marino, *PLOS ONE* **11**, e0154959 (2016).
20. M. Fiocchetti, V. S. Fernandez, E. Montalesi, and M. Marino, *Oxidat. Med. Cell. Long.* **2019**, 6315034 (2019).
21. J. S. Olson, *Antioxid. Redox Signal.* **32**, 228 (2020).
22. A. Keppner, D. Maric, M. Correia, T. W. Koay, I. M. C. Orlando, S. N. Vinogradov, and D. Hoogewijs, *Redox Bio.* **37**, 101687.1 (2020).
23. H.-J. Sun, W.-T. Lee, B. Leng, Z.-Y. Wu, Y. Yang, and J.-S. Bian, *Antioxid. Redox Sign.* **32**, 331 (2019).
24. J. J. Rose, K. A. Bocian, Q. Xu, L. Wang, A. W. DeMartino, X. Chen, C. G. Corey, D. A. Guimaraes, I. Azarov, X. N. Huang, et al., *J. Biol. Chem.* **295**, 6635 (2020).
25. J. M. Fukuto, V. S. Vega, C. Works, and J. Lin, *Curr. Opin. Chem. Biol.* **55**, 52 (2020).
26. C. Ferrante, G. Batignani, E. Pontecorvo, L. C. Montemiglio, M. H. Vos, and T. Scopigno, *J. Am. Chem. Soc.* **142**, 228 (2020).

27. J. Trent and M. Hargrove, *J. Biol. Chem.* **277**, 19538 (2002).
28. A. Bentmann, M. Schmidt, S. Reuss, U. Wolfrum, Hankeln, T., and T. Burmester, *J. Biol. Chem.* **280**, 20660 (2005).
29. D. Hamdane, L. Kiger, S. Dewilde, J. Uzan, T. Burmester, T. Hankeln, L. Moens, and M. C. Marden, *FEBS Journal* **272**, 2076 (2005).
30. X. Ming and W.-H. Fang, *J. Phys. Chem. B.* **112**, 990 (2008).
31. J. Uzan, S. Dewilde, T. Burmester, T. Hankeln, L. Moens, D. Hamdane, M. C. Marden, and L. Kiger, *Biophys. J.* **87**, 1196 (2004).
32. L. Kiger, J. Uzan, S. Dewilde, T. Burmester, T. Hankeln, L. Moens, D. Hamdane, V. Baudin-Creuzat, and M. C. Marden, *Life* **56**, 709 (2004).
33. L. Capece, M. A. Marti, A. Bidon-Chanal, A. Nadra, F. J. Luque, and D. A. Estrin, *Proteins: Structure, Function, and Bioinformatics* **75**, 885 (2009).
34. C. Ardiccioni, A. Arcovito, S. Della Longa, P. van der Linden, D. Bourgeois, M. Weik, L. C. Montemiglio, C. Savino, G. Avella, C. Exertier, et al., *IUCrJ* **6**, 832 (2019).
35. S. Bakels, M.-P. Gageot, and A. M. Rijs, *Chem. Rev.* **120**, 3233 (2020).
36. E. M. Jones, E. Monza, G. Balakrishnan, G. C. Blouin, P. J. Mak, Q. Zhu, J. R. Kincaid, V. Guallar, and T. G. Spiro, *J. Am. Chem. Soc.* **136**, 10325 (2014).
37. T. G. Spiro, A. V. Soldatova, and G. Balakrishnan, *Coord. Chem. Rev.* **257**, 511 (2013).
38. D. Buhrke and P. Hildebrandt, *Chem. Rev.* **120**, 3577 (2020).
39. A. V. Soldatova, M. Ibrahim, J. S. Olson, R. S. Czernuszewicz, and T. G. Spiro, *J. Am. Chem. Soc.* **132**, 4614 (2010).
40. P. Vidossich and A. Magistrato, *Biomolecules* **4**, 616 (2014).
41. J. Chatt and L. A. Duncanson, *J. Chem. Soc.* p. 2939 (1953).

42. J. Dewar, *Bull. Soc. Chim. Fr.* **18**, C71 (1951).
43. X. Wu, L. Zhao, J. Jin, S. Pan, W. Li, X. Jin, G. Wang, M. Zhou, and G. Frenking, *Science* **361**, 912 (2018).
44. G. Bistoni, S. Rampino, N. Scafuri, G. Ciancaleoni, D. Zuccaccia, L. Belpassi, and F. Tarantelli, *Chem. Sci.* **7**, 1174 (2016).
45. J. Stark, *Nature* **92**, 401 (1913).
46. A. Chattopadhyay and S. G. Boxer, *J. Am. Chem. Soc.* **117**, 1449 (1995).
47. S. D. Fried and S. G. Boxer, *Acc. Chem. Res.* **48**, 998 (2015).
48. J. Ma, I. M. Pazos, W. Zhang, R. M. Culik, and F. Gai, *Annu. Rev. Phys. Chem.* **66**, 357 (2015).
49. S. D. Fried and S. G. Boxer, *Annu. Rev. Biochem.* **86**, 387 (2017).
50. B. Błasiak, C. H. Londergan, L. J. Webb, and M. Cho, *Acc. Chem. Res.* **50**, 968 (2017).
51. V. V. Welborn and T. Head-Gordon, *Chem. Rev.* **119**, 6613 (2018).
52. V. V. Welborn, L. Ruiz Pestana, and T. Head-Gordon, *Nat. Catal.* **1**, 649 (2018).
53. S. H. Schneider and S. G. Boxer, *J. Phys. Chem. B* **120**, 9672 (2016).
54. Y. Wu and S. G. Boxer, *J. Am. Chem. Soc.* **138**, 11890 (2016).
55. S. D. Fried, S. Bagchi, and S. G. Boxer, *Science* **346**, 1510 (2014).
56. S. S. Andrews and S. G. Boxer, *J. Phys. Chem. A* **104**, 11853 (2000).
57. B. A. Lindquist, K. E. Furse, and S. A. Corcelli, *Phys. Chem. Chem. Phys.* **11**, 8119 (2009).
58. P. Deb, T. Haldar, S. M. Kashid, S. Banerjee, S. Chakrabarty, and S. Bagchi, *J. Phys. Chem. B* **120**, 4034 (2016).

59. J. D. Slocum and L. J. Webb, *Annu. Rev. Phys. Chem.* **69**, 253 (2018).
60. J. T. First, E. T. Novelli, and L. J. Webb, *J. Phys. Chem. B* **124**, 3387 (2020).
61. E. S. Park and S. G. Boxer, *J. Phys. Chem. B* **106**, 5800 (2002).
62. S. D. Fried, S. Bagchi, and S. G. Boxer, *J. Am. Chem. Soc.* **135**, 11181 (2013).
63. N. Verma, Y. Tao, W. Zou, X. Chen, X. Chen, M. Freindorf, and E. Kraka, *Sensors* **20**, 2358 (2020).
64. A. Krapp and G. Frenking, *J. Am. Chem. Soc.* **130**, 16646 (2008).
65. G. Frenking, I. Fernández, N. Holzmann, S. Pan, I. Krossing, and M. Zhou, *JACS Au* **1**, 623 (2021).
66. G. Frenking, C. Loschen, A. Krapp, S. Fau, and S. H. Strauss, *J. Comput. Chem.* **28**, 117 (2007).
67. X. Dong, C. Ding, Q. Zhang, M. Chen, L. Zhao, M. Zhou, and G. Frenking, *JACS* **143**, 14300 (2021).
68. A. O. Ortolan, G. F. Caramori, R. L. T. Parreira, R. P. Orenha, A. Muñoz-Castro, and G. Frenking, *Dalton Trans.* **49**, 16762 (2020).
69. S. Pan, L. Zhao, H. V. R. Dias, and G. Frenking, *Inorg. Chem.* **57**, 7780 (2018).
70. M. S. Nechaev, V. M. Rayón, and G. Frenking, *J. Phys. Chem. A* **108**, 3134 (2004).
71. M. Li, A. G. Oliver, and W. R. Scheidt, *Inorg. Chem.* **57**, 5648 (2018).
72. D. Koch, Y. Chen, P. Golub, and S. Manzhos, *Phys. Chem. Chem. Phys.* **21**, 20814 (2019).
73. S. A. Macgregor and O. Eisenstein, eds., *Structure and Bonding 167: Computational Studies in Organometallic Chemistry* (Springer, Heidelberg, 2016).
74. D. J. Durand and N. Fey, *Chem. Rev.* **119**, 6561 (2019).

75. S. Ahn, M. Hong, M. Sundararajan, D. Ess, and M. Baik, *Chem. Rev.* **119**, 6509 (2019).
76. K. A. Moltved and K. P. Kepp, *J. Chem. Theory and Comput.* **14**, 3479 (2018).
77. D. Wu, C. Dong, H. Zhan, and X.-W. Du, *J. Phys. Chem. Letters* **9**, 3387 (2018).
78. C. P. Gordon, R. A. Andersen, and C. Coperet, *Helv. Chim. Acta* **102**, e1900151 (2019).
79. L. Zhao, M. Zhi, and G. Frenking, *Int. J. Quantum Chem.* (2021).
80. S. Pan and G. Frenking, *Molecules* **26**, 4695 (2021).
81. Y.-R. Luo, *Comprehensive Handbook of Chemical Bond Energies* (Taylor and Francis, Boca Raton, FL, 2007).
82. N. Kosar, K. Ayub, M. A. Gilani, and T. Mahmood, *J Mol Model* **25** (2019).
83. M. D. Morse, *Acc. Chem. Res.* **52**, 119 (2018).
84. Z. Fang, M. Vasiliu, K. A. Peterson, and D. A. Dixon, *J. Chem. Theory Comput.* **13**, 1057 (2017).
85. J. Andrés, P. W. Ayers, R. A. Boto, R. Carbó-Dorca, H. Chermette, J. Cioslowski, J. Contreras-García, D. L. Cooper, G. Frenking, C. Gatti, et al., *J. Comput. Chem.* p. doi.org/10.1002/jcc.26003 (2019).
86. A. Krapp, F. M. Bickelhaupt, and G. Frenking, *Chem. Eur. J.* **12**, 9196 (2006).
87. L. Zhao, M. Hermann, W. H. E. Schwarz, and G. Frenking, *Nat. Rev. Chem.* **3**, 48 (2019).
88. D. Setiawan, D. Sethio, D. Cremer, and E. Kraka, *Phys. Chem. Chem. Phys.* **20**, 23913 (2018).
89. D. Sethio, V. Oliveira, and E. Kraka, *Molecules* **23**, 2763 (2018).
90. E. Kraka and D. Cremer, *ChemPhysChem* **10**, 686 (2009).

91. D. Cremer, J. A. Larsson, and E. Kraka, in *Theoretical and Computational Chemistry*, edited by C. Parkanyi (Elsevier, Amsterdam, 1998), pp. 259–327.
92. Z. Konkoli, J. A. Larsson, and D. Cremer, *Int. J. Quantum Chem.* **67**, 11 (1998).
93. Z. Konkoli and D. Cremer, *Int. J. Quantum Chem.* **67**, 29 (1998).
94. M. Kaupp, D. Danovich, and S. Shaik, *Coord. Chem. Rev.* **344**, 355 (2017).
95. E. Kraka, W. Zou, and Y. Tao, *WIREs: Comput. Mol. Sci.* **10**, 1480 (2020).
96. E. Kraka and M. Freindorf, in *Topics in Organometallic Chemistry - New Directions in the Modeling of Organometallic Reactions*, edited by A. Lledós and G. Ujaque (Springer, Berlin, Heidelberg, 2020), vol. 67, pp. 1–43.
97. D. Cremer and E. Kraka, *Dalton Trans.* **46**, 8323 (2017).
98. C. A. Tolman, *Chem. Soc. Rev.* **1**, 337 (1972).
99. C. A. Tolman, *Chem. Rev.* **77**, 313 (1977).
100. L. W. Chung, W. M. C. Sameera, R. Ramozzi, A. J. Page, M. Hatanaka, G. P. Petrova, T. V. Harris, X. Li, Z. Ke, F. Liu, et al., *Chem. Rev.* **115**, 5678 (2015).
101. N. J. Silvernail, A. Roth, C. E. Schulz, B. C. Noll, and W. R. Scheidt, *J. Am. Chem. Soc.* **127**, 14422 (2005).
102. M. Freindorf and E. Kraka, *J. Mol. Model.* **26**, 281 (2020).
103. W. Zou, Y. Tao, M. Freindorf, D. Cremer, and E. Kraka, *Chem. Phys. Lett.* **478**, 137337 (2020).
104. R. L. Khade, Y. Yang, Y. Shi, and Y. Zhang, *Angew. Chem. Int. Ed.* **55**, 15058 (2016).
105. Y. Zhang, *J. Inorg. Biochem.* **118**, 191 (2013).
106. K. Nienhaus, J. M. Kriegl, and G. U. Nienhaus, *J. Biolog. Chem.* **279**, 22944 (2004).
107. K. Nienhaus and G. U. Nienhaus, *J. Biolog. Phys.* **31**, 417 (2005).

108. K. Nienhaus and G. U. Nienhaus, *J. Biol. Phys.* **33**, 357 (2007).
109. K. Nienhaus, S. Lutz, M. Meuwly, and G. U. Nienhaus, *ChemPhysChem* **11**, 119 (2010).
110. M. Anselmi, A. Di Nola, and A. Amadei, *J. Phys. Chem. B* **115**, 2436 (2011).
111. S. Lutz and M. Meuwly, *Faraday Disc.* **150**, 375 (2011).
112. T. Uno, D. Ryu, H. Tsutsumi, Y. Tomisugi, Y. Ishikawa, A. J. Wilkinson, H. Sato, and T. Hayashi, *J. Biol. Chem.* **279**, 5886 (2004).
113. H. Sawai, M. Makino, Y. Mizutani, T. Ohta, H. Sugimoto, T. Uno, N. Kawada, K. Yoshizato, T. Kitagawa, and Y. Shiro, *Biochemistry* **44**, 13257 (2005).
114. S. Lutz, K. Nienhaus, G. U. Nienhaus, and M. Meuwly, *J. Phys. Chem. B* **113**, 15334 (2009).
115. A. Ivan, W. Ling, R. J. J., X. Qinzi, H. X. N., B. Andrea, W. Ying, G. Lanping, L. Chen, U. K. B., et al., *Sci. Trans. Med.* **8**, 368ra173 (2016).
116. C. Exertier, L. Milazzo, I. Freda, L. C. Montemiglio, A. Scaglione, G. Cerutti, G. Parisi, M. Anselmi, G. Smulevich, C. Savino, et al., *Sci. Rep.* **9**, 5326 (2019).
117. L. Milazzo, C. Exertier, M. Becucci, I. Freda, L. C. Montemiglio, C. Savino, B. Vallone, and G. Smulevich, *FEBS* **287**, 4082 (2020).
118. B. Vallone, K. Nienhaus, A. Matthes, M. Brunori, and G. U. Nienhaus, *PNAS* **101**, 17351 (2004).
119. D. A. Case, I. Y. Ben-Shalom, S. R. Brozell, D. S. Cerutti, T. E. Cheatham, V. W. D. Cruzeiro, T. A. Darden, R. E. Duke, D. Ghoreishi, M. K. Gilson, et al., *AMBER*, University of California, San Francisco (2018).
120. W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, *J. Chem. Phys.* **79**, 926 (1983).
121. C. Adamo and V. Barone, *J. Chem. Phys.* **110**, 6158 (1999).

122. R. Ditchfield, W. J. Hehre, and J. A. Pople, *J. Chem. Phys.* **54**, 724 (1971).
123. J.-D. Chai and M. Head-Gordon, *J. Chem. Phys.* **128**, 084106 (2008).
124. Y. Minenkov, Å. Singstad, G. Occhipinti, and V. R. Jensen, *Dalton Trans.* **41**, 5526 (2012).
125. M. K. Kesharwani, B. Brauer, and J. M. L. Martin, *J. Phys. Chem. A* **119**, 1701 (2015).
126. J. C. Z. Trujillo and L. K. McKemmish, *WIREs Comput Mol Sci.* pp. e1584–1–e1584–19 (2021).
127. B. A. Springer, S. G. Sligar, J. S. Olson, and G. N. J. Phillips, *Chem. Rev.* **94**, 699 (1994).
128. J. N. Harvey, *J. Am. Chem. Soc.* **122**, 12401 (2000).
129. K. P. Jensen and U. Ryde, *J. Biolog. Chem.* **279**, 14561 (2004).
130. R. M. Badger, *J. Chem. Phys.* **2**, 128 (1934).
131. E. Kraka, J. A. Larsson, and D. Cremer, in *Computational Spectroscopy*, edited by J. Grunenberg (Wiley, New York, 2010), pp. 105–149.
132. E. B. Wilson, J. C. Decius, and P. C. M. Cross, *Molecular Vibrations. The Theory of Infrared and Raman Vibrational Spectra* (McGraw-Hill, New York, 1955).
133. E. B. Wilson, J. C. Decius, P. C. Cross, and B. R. Sundheim, *J. Electrochem. Soc.* **102**, 235C (1955).
134. Z. Konkoli and D. Cremer, *Int. J. Quantum Chem.* **67**, 1 (1998).
135. Z. Konkoli, J. A. Larsson, and D. Cremer, *Int. J. Quantum Chem.* **67**, 41 (1998).
136. W. Zou, R. Kalescky, E. Kraka, and D. Cremer, *J. Chem. Phys.* **137**, 084114 (2012).
137. W. Zou and D. Cremer, *Theor. Chem. Acc.* **133**, 1451 (2014).
138. W. Zou and D. Cremer, *Chem. Eur. J.* **22**, 4087 (2016).

139. A. A. A. Delgado, D. Sethio, I. Munar, V. Aviyente, and E. Kraka, *J. Chem. Phys.* **153**, 224303 (2020).
140. R. Srinivas, N. Verma, E. Kraka, and E. C. Larson, *J. Chem. Inf. Model.* **61**, 2159 (2021).
141. N. Verma, Y. Tao, and E. Kraka, *J. Phys. Chem. B* **125**, 2551 (2021).
142. R. Kalescky, E. Kraka, and D. Cremer, *J. Phys. Chem. A* **117**, 8981 (2013).
143. A. Humason, W. Zou, and D. Cremer, *J Phys Chem A.* **119**, 1666 (2014).
144. R. Kalescky, E. Kraka, and D. Cremer, *Int. J. Quantum Chem.* **114**, 1060 (2014).
145. R. Kalescky, W. Zou, E. Kraka, and D. Cremer, *J. Phys. Chem. A* **118**, 1948 (2014).
146. J. A. Larsson and D. Cremer, *J. Mol. Struct.* **485-486**, 385 (1999).
147. D. Cremer and E. Kraka, *Curr. Org. Chem.* **14**, 1524 (2010).
148. D. Cremer, A. Wu, J. A. Larsson, and E. Kraka, *J. Mol. Model.* **6**, 396 (2000).
149. E. Kraka and D. Cremer, *Rev. Proc. Quim.* pp. 39–42 (2012).
150. D. Setiawan, E. Kraka, and D. Cremer, *J. Phys. Chem. A* **119**, 9541 (2015).
151. E. Kraka, D. Setiawan, and D. Cremer, *J. Comp. Chem.* **37**, 130 (2015).
152. A. A. A. Delgado, A. Humason, R. Kalescky, M. Freindorf, and E. Kraka, *Molecules* **26**, 950 (2021).
153. V. Oliveira, E. Kraka, and D. Cremer, *Phys. Chem. Chem. Phys.* **18**, 33031 (2016).
154. V. Oliveira, E. Kraka, and D. Cremer, *Inorg. Chem.* **56**, 488 (2016).
155. V. Oliveira and D. Cremer, *Chem. Phys. Lett.* **681**, 56 (2017).
156. S. Yannacone, V. Oliveira, N. Verma, and E. Kraka, *Inorganics* **7**, 47 (2019).
157. V. P. Oliveira, E. Kraka, and F. B. C. Machado, *Inorg. Chem.* **58**, 14777 (2019).

158. V. P. Oliveira, B. L. Marcial, F. B. C. Machado, and E. Kraka, *Materials* **13**, 55 (2020).
159. D. Setiawan, E. Kraka, and D. Cremer, *Chem. Phys. Lett.* **614**, 136 (2014).
160. D. Setiawan, E. Kraka, and D. Cremer, *J. Phys. Chem. A* **119**, 1642 (2014).
161. D. Setiawan and D. Cremer, *Chem. Phys. Lett.* **662**, 182 (2016).
162. V. Oliveira, D. Cremer, and E. Kraka, *J. Phys. Chem. A* **121**, 6845 (2017).
163. V. Oliveira and E. Kraka, *J. Phys. Chem. A* **121**, 9544 (2017).
164. X. Zhang, H. Dai, H. Yan, W. Zou, and D. Cremer, *J. Am. Chem. Soc.* **138**, 4334 (2016).
165. W. Zou, X. Zhang, H. Dai, H. Yan, D. Cremer, and E. Kraka, *J. Organometal. Chem.* **856**, 114 (2018).
166. Y. Tao, Y. Qiu, W. Zou, S. Nanayakkara, S. Yannacone, and E. Kraka, *Molecules* **25**, 1589 (2020).
167. R. Kalescky, W. Zou, E. Kraka, and D. Cremer, *Chem. Phys. Lett.* **554**, 243 (2012).
168. R. Kalescky, E. Kraka, and D. Cremer, *Mol. Phys.* **111**, 1497 (2013).
169. M. Freindorf, E. Kraka, and D. Cremer, *Int. J. Quantum Chem.* **112**, 3174 (2012).
170. R. Kalescky, W. Zou, E. Kraka, and D. Cremer, *Aust. J. Chem.* **67**, 426 (2014).
171. Y. Tao, W. Zou, J. Jia, W. Li, and D. Cremer, *J. Chem. Theory Comput.* **13**, 55 (2017).
172. Y. Tao, W. Zou, and E. Kraka, *Chem. Phys. Lett.* **685**, 251 (2017).
173. M. Freindorf, Y. Tao, D. Sethio, D. Cremer, and E. Kraka, *Mol. Phys.* **117**, 1172 (2018).
174. M. Z. Makoś, M. Freindorf, D. Sethio, and E. Kraka, *Theor. Chem. Acc.* **138**, 76 (2019).
175. S. Lyu, N. Beiranvand, M. Freindorf, and E. Kraka, *J. Phys. Chem. A* **123**, 7087 (2019).
176. S. Yannacone, D. Sethio, and E. Kraka, *Theor. Chem. Acc.* **139**, 125 (2020).

177. N. Beiranvand, M. Freindorf, and E. Kraka, *Molecules* **26**, 2268 (2021).
178. Y. Tao, W. Zou, D. Sethio, N. Verma, Y. Qiu, C. Tian, D. Cremer, and E. Kraka, *J. Chem. Theory Comput.* **15**, 1761 (2019).
179. I. Mayer, *Chem. Phys. Lett.* **97**, 270 (1983).
180. I. Mayer, *Int. J. Quantum Chem.* **29**, 477 (1986).
181. I. Mayer, *J. Comput. Chem.* **28**, 204 (2007).
182. R. F. W. Bader, *Acc. Chem. Res.* **18**, 9 (1985).
183. R. Bader, *Atoms in Molecules: A Quantum Theory*, International series of monographs on chemistry (Oxford:Clarendon Press, 1990), ISBN 9780198551683.
184. R. Bader, *Atoms in Molecules: A Quantum Theory* (Clarendon Press, Oxford, 1995).
185. P. Popelier, *Atoms in Molecules: An Introduction* (Prentice-Hall, Harlow, England, 2000).
186. D. Cremer and E. Kraka, *Angew. Chem. Int. Ed.* **23**, 627 (1984).
187. D. Cremer and E. Kraka, *Croatica Chem. Acta* **57**, 1259 (1984).
188. E. Kraka and D. Cremer, in *Theoretical Models of Chemical Bonding. The Concept of the Chemical Bond* (Z.B. Maksic, ed., Springer Verlag, Heidelberg, 1990), vol. 2, pp. 453–542.
189. A. Reed, L. Curtiss, and F. Weinhold, *Chem. Rev.* **88**, 899 (1988).
190. F. Weinhold and C. R. Landis, *Valency and Bonding: A Natural Bond Orbital Donor-Acceptor Perspective* (Cambridge University Press, 2003).
191. M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, et al., *Gaussian, Inc.*, Gaussian Inc., Wallingford CT (2009).

192. W. Zou, Y. Tao, M. Freindorf, M. Z. Makoś, N. Verma, D. Cremer, and E. Kraka, *Local vibrational mode analysis (lmodea)*, Computational and Theoretical Chemistry Group (CATCO), Southern Methodist University: Dallas, TX, USA (2021).
193. E. D. Glendening, J. K. Badenhoop, A. E. Reed, J. E. Carpenter, J. A. Bohmann, C. M. Morales, C. R. Landis, and F. Weinhold, *NBO6* (2013), theoretical Chemistry Institute, University of Wisconsin, Madison.
194. R. F. W. Bader, *Chem. Rev.* **91**, 893 (1991).
195. T. A. Keith, *AIMALL*, TK Gristmill Software, Overland Park KS (2017).
196. O. O. Brovarets', Y. P. Yurenko, and D. M. Hovorun, *J. Biomol. Struct. Dyn.* **32**, 993 (2014).
197. Y. Tao, C. Tian, N. Verma, W. Zou, C. Wang, D. Cremer, and E. Kraka, *J. Chem. Theory Comput.* **14**, 2558 (2018).
198. Y. Tao, W. Zou, S. Nanayakkara, M. Freindorf, and E. Kraka, *Theor. Chem. Acc.* **140**, 31 (2021).
199. S. Dajnowicz, S. Seaver, B. L. Hanson, S. Z. Fisher, P. Langan, A. Y. Kovalevsky, and T. C. Mueser, *Acta Cryst. D* **72**, 892 (2016).
200. J. H. van Wonderen, K. Adamczyk, X. Wu, X. Jiang, S. E. H. Piper, C. R. Hall, M. J. Edwards, T. A. Clarke, H. Zhang, L. J. C. Jeuken, et al., *PNAS* **118**, e2107939118 (2021).
201. A. Batra, G. Kladnik, H. Vázquez, J. S. Meisner, L. Floreano, C. Nuckolls, D. Cvetko, A. Morgante, and L. Venkataraman, *Nature Comm.* **3**, 1086 (2012).
202. X. Wang, S. Wang, J. Lv, S. Shao, L. Wang, X. Jing, and F. Wang, *Chem. Sci.* **10**, 2915 (2019).
203. S. A. Serron, W. S. Aldridge, C. N. Fleming, R. M. Danell, M.-H. Baik, M. Sykora, D. M. Dattelbaum, and T. J. Meyer, *J. Am. Chem. Soc.* **126**, 14506 (2004).

204. L. Paltrinieri, G. Di Rocco, G. Battistuzzi, M. Borsari, M. Sola, A. Ranieri, L. Zanetti-Polzi, I. Daidone, and C. A. Bortolotti, *JBIC* **22**, 615 (2017).
205. J. C. Kendrew, G. Bodo, H. M. Dintzis, R. G. Parrish, H. Wyckoff, and D. C. Phillips, *Nature* **181**, 662 (1958).
206. R. H. Austin, K. W. Beeson, L. Eisenstein, H. Frauenfelder, and I. C. Gunsalus, *Biochemistry* **14**, 5355 (1975).
207. I. Schlichting, J. Berendzen, G. N. Phillips, and R. M. Sweet, *Nature* **371**, 808 (1994).
208. R. F. Tilton, I. D. Kuntz, and G. A. Petsko, *Biochemistry* **23**, 2849 (1984).
209. J. Vojtěchovský, K. Chu, J. Berendzen, R. M. Sweet, and I. Schlichting, *Biophys. J.* **77**, 2153 (1999).
210. H. Shimada and W. S. Caughey, *J. Biol. Chem.* **257**, 11893 (1982).
211. A. Ansari, J. Berendzen, D. Braunstein, B. R. Cowen, H. Frauenfelder, M. K. Hong, I. E. Iben, J. Johnson, P. Ormos, T. B. Sauke, et al., *Biophys. Chem.* **26**, 337 (1987).
212. D. Braunstein, K. Chu, K. Egeberg, H. Frauenfelder, J. Mourant, G. Nienhaus, P. Ormos, S. Sligar, B. Springer, and R. Young, *Biophys. J.* **65**, 2447 (1993).
213. S. Bagchi, B. T. Nebgen, R. F. Loring, and M. D. Fayer, *J. Am. Chem. Soc.* **132**, 18367 (2010).
214. J.-H. Choi, K.-W. Kwak, and M. Cho, *J. Phys. Chem. B* **117**, 15462 (2013).
215. M. M. Teeter, *Protein Sci.* **13**, 313 (2004).
216. B. G. Schulze and J. D. Evanseck, *J. Am. Chem. Soc.* **121**, 6444 (1999).
217. A. Amadei and M. Aschi, *J. Phys. Chem. B* **125**, 13624 (2021).
218. G. N. Phillips, M. L. Teodoro, T. Li, B. Smith, and J. S. Olson, *J. Phys. Chem. B.* **103**, 8817 (1999).

219. C. Rovira, B. Schulze, M. Eichinger, J. D. Evanseck, and M. Parrinello, *Biophys. J.* **81**, 435 (2001).
220. M. Devereux and M. Meuwly, *Biophys. J.* **96**, 4363 (2009).
221. D. C. Lamb, K. Nienhaus, A. Arcovito, F. Draghi, A. E. Miele, M. Brunori, and G. U. Nienhaus, *J. Biol. Chem.* **277**, 11636 (2002).
222. K. Nienhaus, P. Deng, J. S. Olson, J. J. Warren, and G. Nienhaus, *J. Biol. Chem.* **278**, 42532 (2003).
223. M. C. Thielges, J. Y. Axup, D. Wong, H. S. Lee, J. K. Chung, P. G. Schultz, and M. D. Fayer, *J. Phys. Chem. B* **115**, 11294 (2011).
224. D. Kossowska, K. Kwak, and M. Cho, *Molecules* **23**, 3189 (2018).
225. T. G. Spiro, A. V. Soldatova, and G. Balakrishnan, *Coord. Chem. Rev.* **257**, 511 (2013).
226. K. Falahati, H. Tamura, I. Burghardt, and M. Huix-Rotllant, *Nat. Commun.* **9** (2018).
227. P. J. Steinbach, A. Ansari, J. Berendzen, D. Braunstein, K. Chu, B. R. Cowen, D. Ehrenstein, H. Frauenfelder, and J. B. Johnson, *Biochemistry* **30**, 3988 (1991).
228. A. Ansari, C. M. Jones, E. R. Henry, J. Hofrichter, and W. A. Eaton, *Biochemistry* **33**, 5128 (1994).
229. G. U. Nienhaus, K. Chu, and K. Jesse, *Biochemistry* **37**, 6819 (1998).
230. J. O. Alben, D. Beece, S. F. Bowne, W. Doster, L. Eisenstein, H. Frauenfelder, D. Good, J. D. McDonald, M. C. Marden, P. P. Moh, et al., *PNAS* **79**, 3744 (1982).
231. M. Lim, T. A. Jackson, and P. A. Anfinrud, *J. Chem. Phys.* **102**, 4355 (1995).
232. M. Lim, T. A. Jackson, and P. A. Anfinrud, *Nat. Struct. Biol.* **4**, 209 (1997).
233. K. Nienhaus, J. S. Olson, S. Franzen, and G. U. Nienhaus, *J. Am. Chem. Soc.* **127**, 40 (2005).

234. M. Meuwly, *ChemPhysChem* **7**, 2061 (2006).
235. B. Vallone, K. Nienhaus, A. Matthes, M. Brunori, and G. U. Nienhaus, *PNAS* **101**, 17351 (2004).
236. B. Vallone, K. Nienhaus, M. Brunori, and G. U. Nienhaus, *Proteins* **56**, 85 (2004).
237. S.-F. Chen, X.-C. Liu, J.-K. Xu, L. Li, J.-J. Lang, G.-B. Wen, and Y.-W. Lin, *Inorg. Chem.* **60**, 2839 (2021).
238. L. Chen, J.-K. Xu, L. Li, S.-Q. Gao, G.-B. Wen, and Y.-W. Lin, *Molecular Systems Design & Engineering* (2022).

Figure 1: Upper: The side view of sketches of the molecular systems presenting the QM model in the enzyme for **H64 ϵ** , **H64 δ** , and **H64 $\delta\epsilon$** tautomers of distal histidine for CO coordinated to Fe in hexacoordinate NgbCO; Lower: The side view sketch the QM heme model in the gas phase (**Gas**), and the side view sketches presenting CO above the heme group in the docking sites **A** and **B** of the pentacoordinate NgbCO.

Figure 2: a) The BSO value as a function of the CO local mode force constant k^a ; b) the CO local mode force constant k^a vs the CO bond length R ; c) the CO local mode force constant k^a vs the CO energy density H_ρ ; d) the CO local mode force constant k^a vs the CO antibonding population POP; e) the CO local mode force constant k^a vs the NBO charge on O; f) the CO local mode force constant k^a vs the sum of C and O charges of CO for CO bonded to the heme Fe in hexacoordinate NgbCO complexes. ω B97X-D/6-31G(d,p)/AMBER level of theory in the protein, and ω B97X-D/6-31G(d,p) in the gas phase. For descriptions of molecular labels, see the text.

Figure 3: a) The BSO value as a function of the local mode force constant k^a ; b) the FeC local mode force constant k^a vs the FeC bond length R ; c) the FeC local mode force constant k^a vs the FeC energy density H_ρ ; d) the CO local mode force constant k^a vs the FeC local mode force constant k^a for CO coordinated to the heme Fe in hexacoordinate NgbCO complexes. ω B97X-D/6-31G(d,p)/AMBER level of theory in the protein, and ω B97X-D/6-31G(d,p) in the gas phase. For descriptions of molecular labels, see the text.

Figure 4: BSO as a function of the local mode force constant of the weak interaction between CO and hydrogen atoms of H64 for hexacoordinate NgbCO complexes; a) O \cdots H-N interaction; b) O \cdots H-C interaction; c) C \cdots H-N interaction. ω B97X-D/6-31G(d,p)/AMBER level of theory. For descriptions of molecular labels, see the text.

Figure 5: a) The BSO value as a function of the CO local mode force constant k^a ; b) the CO local mode force constant k^a vs the CO bond length R ; c) the CO local mode force constant k^a vs the CO energy density H_ρ ; d) the CO local mode force constant k^a vs the CO antibonding population POP; e) the CO local mode force constant k^a vs the NBO charge on O; f) the CO local mode force constant k^a vs the sum of C and O charges of CO for dissociated CO in docking form **A** of pentacoordinate NgbCO. ω B97X-D/6-31G(d,p)/AMBER level of theory in the protein, and ω B97X-D/6-31G(d,p) in the gas phase. For descriptions of molecular labels, see the text.

Figure 6: a) The BSO value as a function of the CO local mode force constant k^a ; b) the CO local mode force constant k^a vs the CO bond length R ; c) the CO local mode force constant k^a vs the CO energy density H_ρ ; d) the CO local mode force constant k^a vs the CO antibonding population POP; e) the CO local mode force constant k^a vs the NBO charge on O; f) the CO local mode force constant k^a vs the sum of C and O charges of CO for dissociated CO in docking form **B** of pentacoordinate NgbCO. ω B97X-D/6-31G(d,p)/AMBER level of theory in the protein, and ω B97X-D/6-31G(d,p) in the gas phase. For descriptions of molecular labels, see the text.

Figure 7: BSO as a power function of the local mode force constant for CO; Black dots - hexacoordinate NgBCO complexes; Green dots - pentacoordinate form *A*; Blue dots - pentacoordinate form *B*. ω B97X-D/6-31G(d,p)/AMBER level of theory in the protein, and ω B97X-D/6-31G(d,p) in the gas phase.

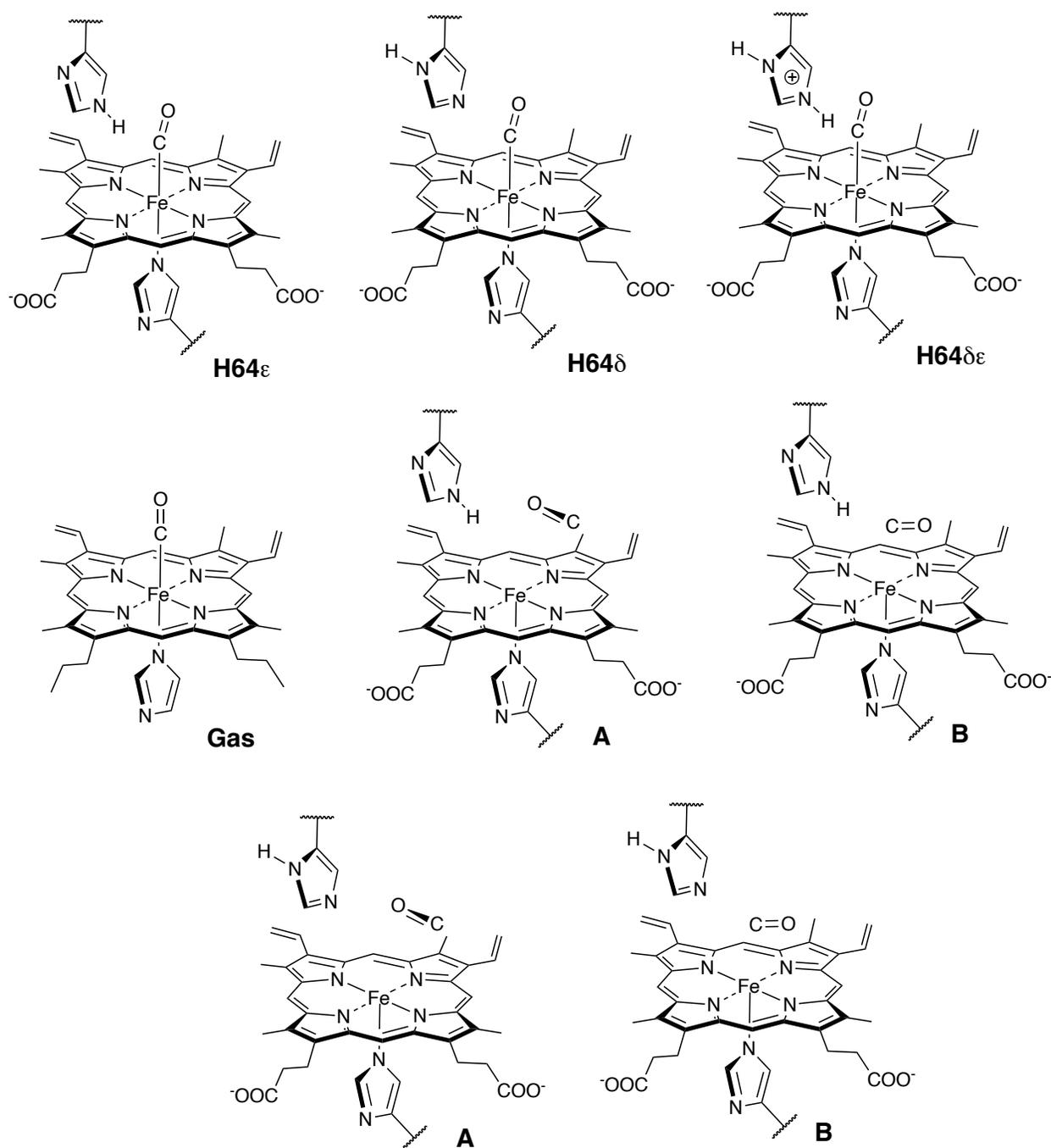
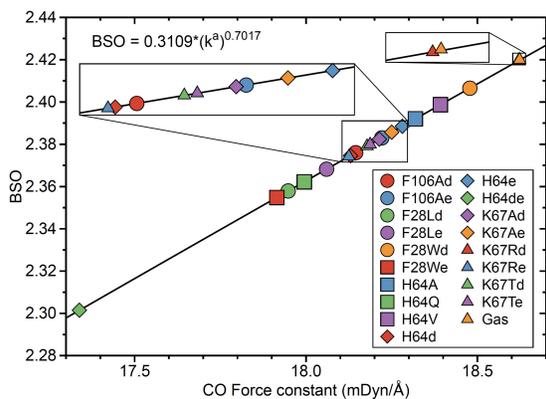
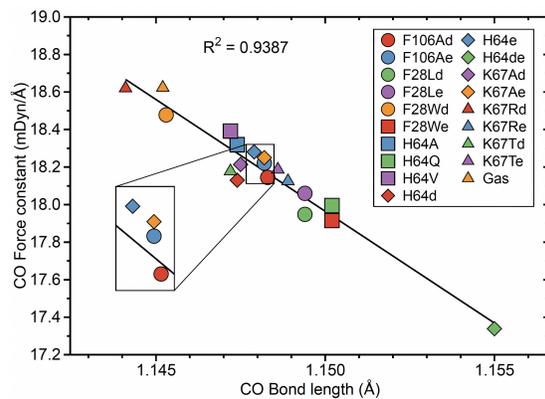


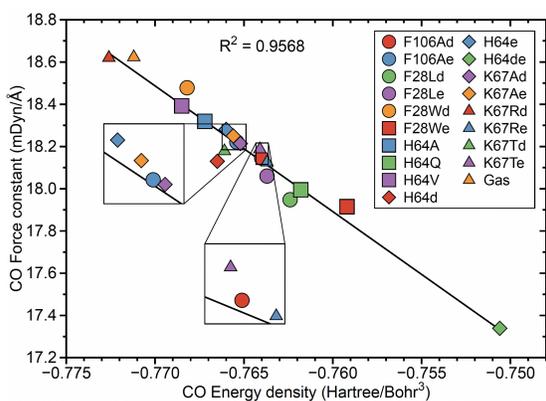
Figure 1
 Marek Freindorf, Alexis Delgado, Elfi Kraka
 J. Comput. Chem.



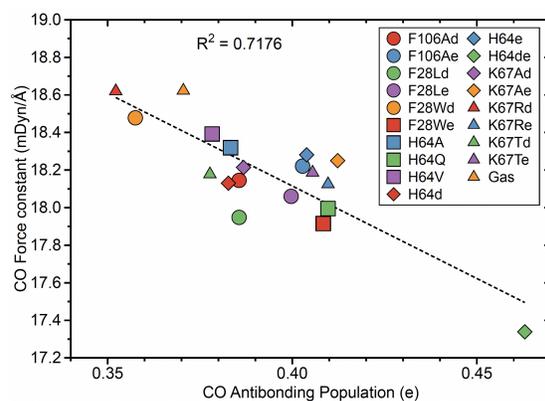
a)



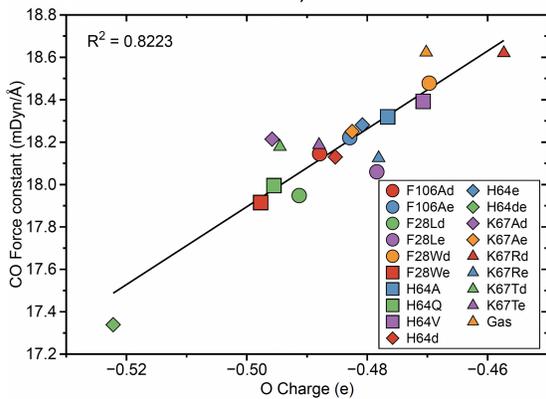
b)



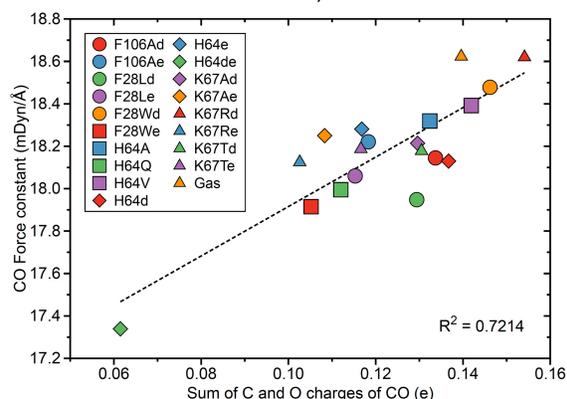
c)



d)



e)



f)

Figure 2
 Marek Freindorf, Alexis Delgado, Elfi Kraka
 J. Comput. Chem.

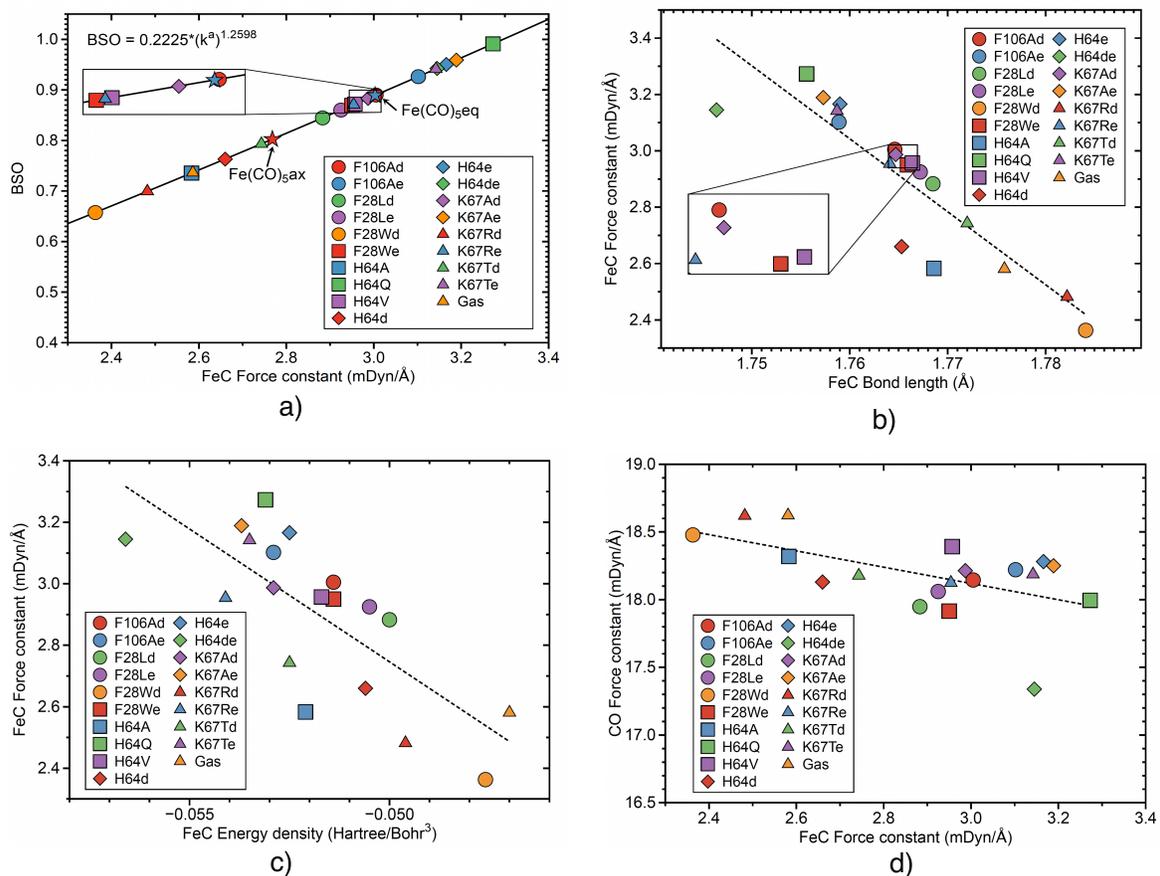


Figure 3
 Marek Freindorf, Alexis Delgado, Elfi Kraka
 J. Comput. Chem.

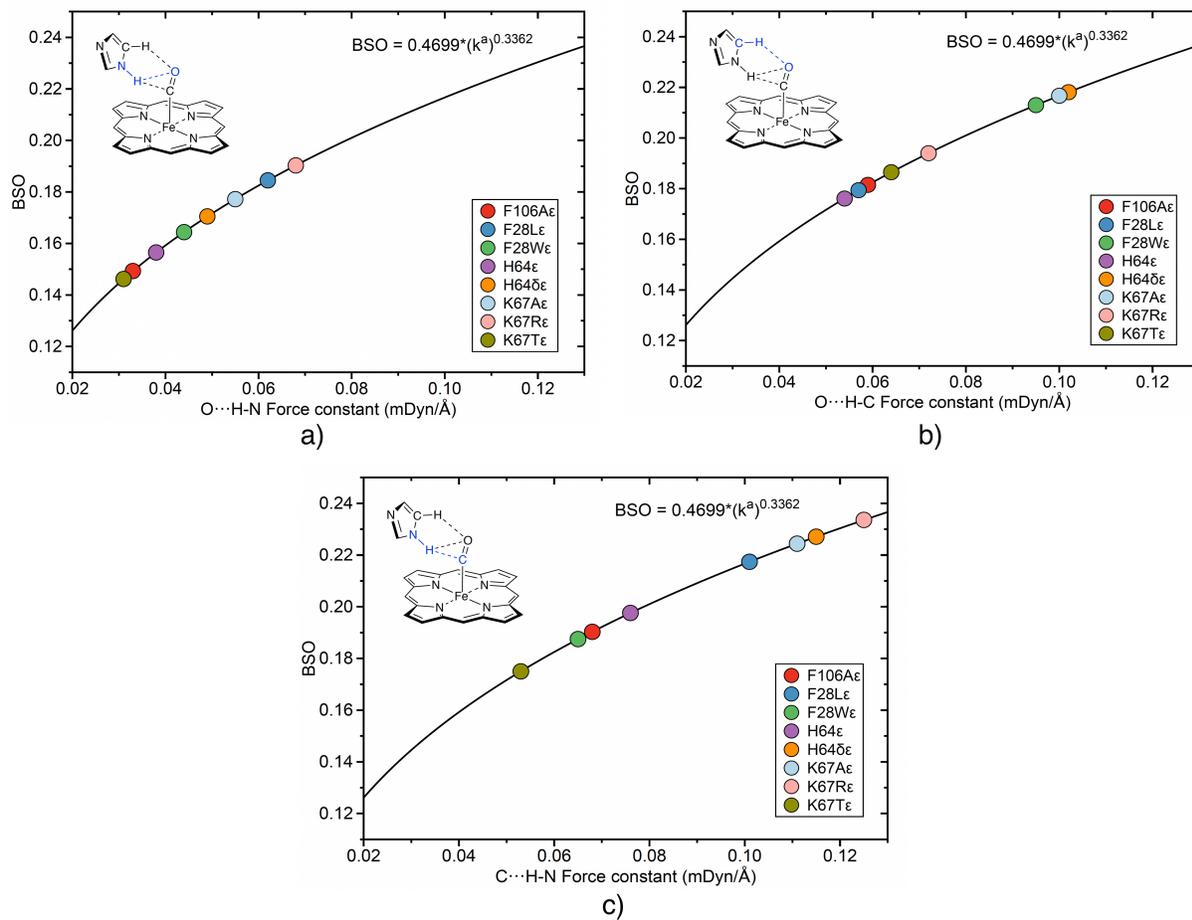


Figure 4
 Marek Freindorf, Alexis Delgado, Elfi Kraka
 J. Comput. Chem.

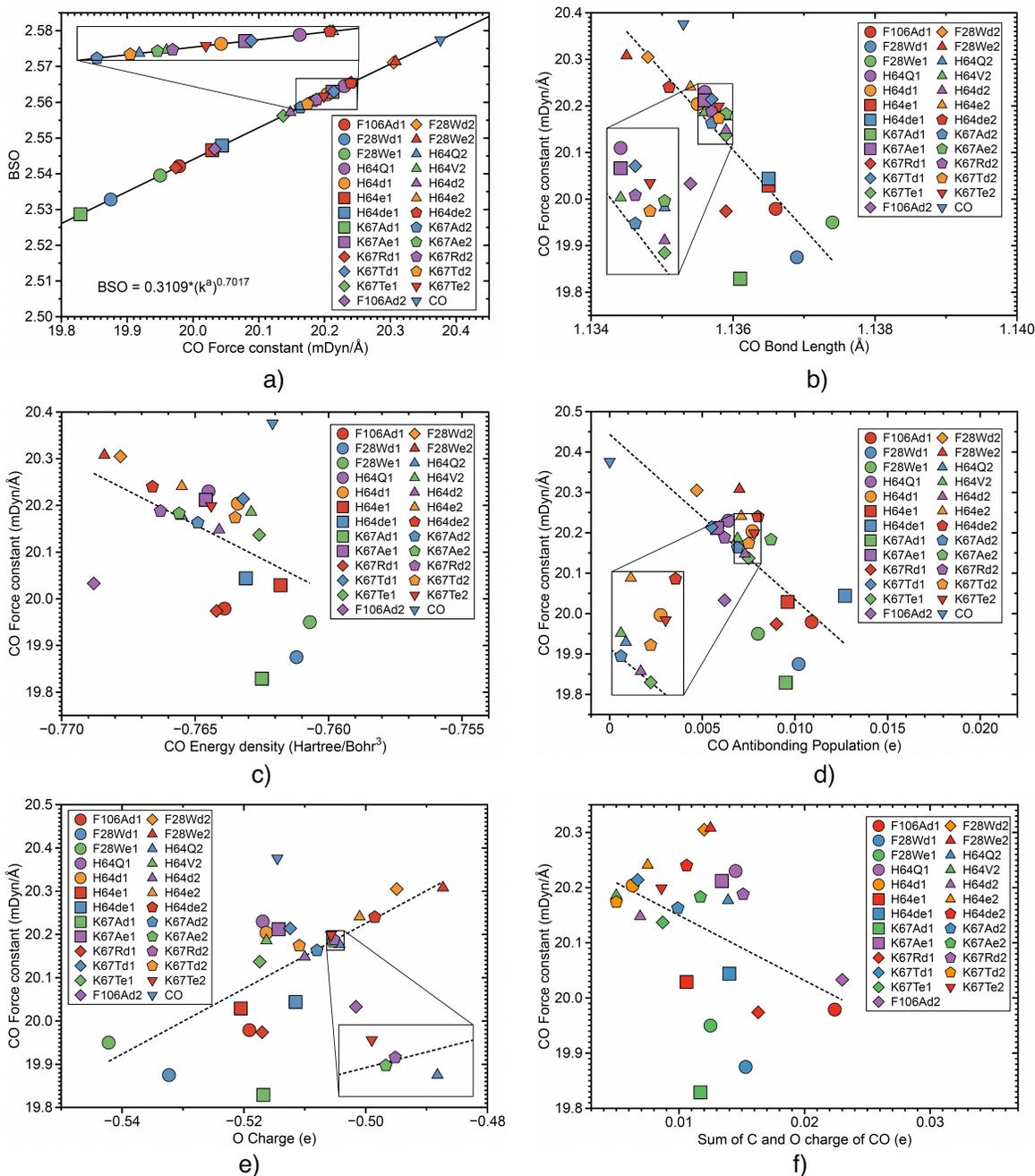


Figure 5
Marek Freindorf, Alexis Delgado, Elfi Kraka
J. Comput. Chem.

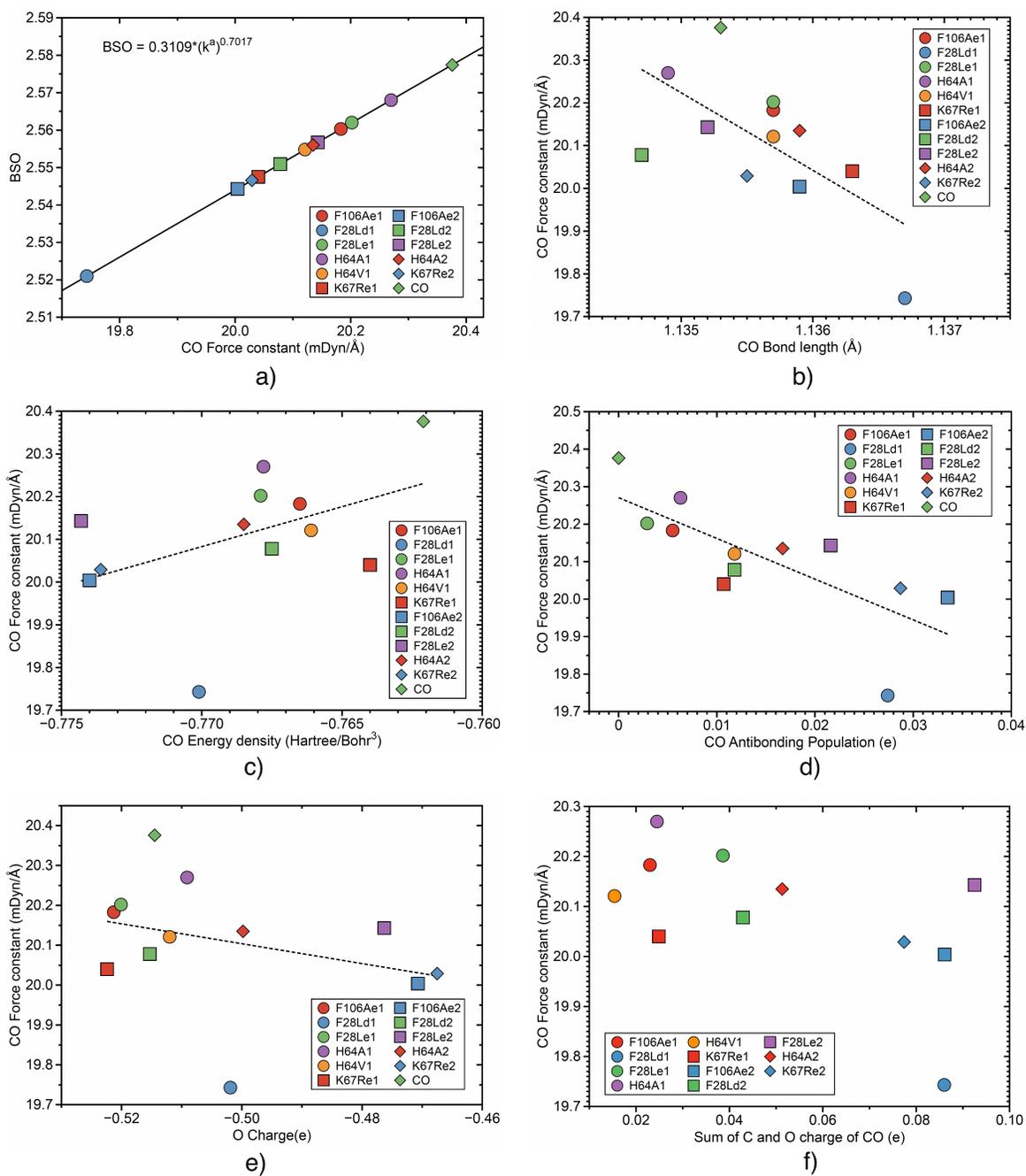


Figure 6
 Marek Freindorf, Alexis Delgado, Elfi Kraka
 J. Comput. Chem.

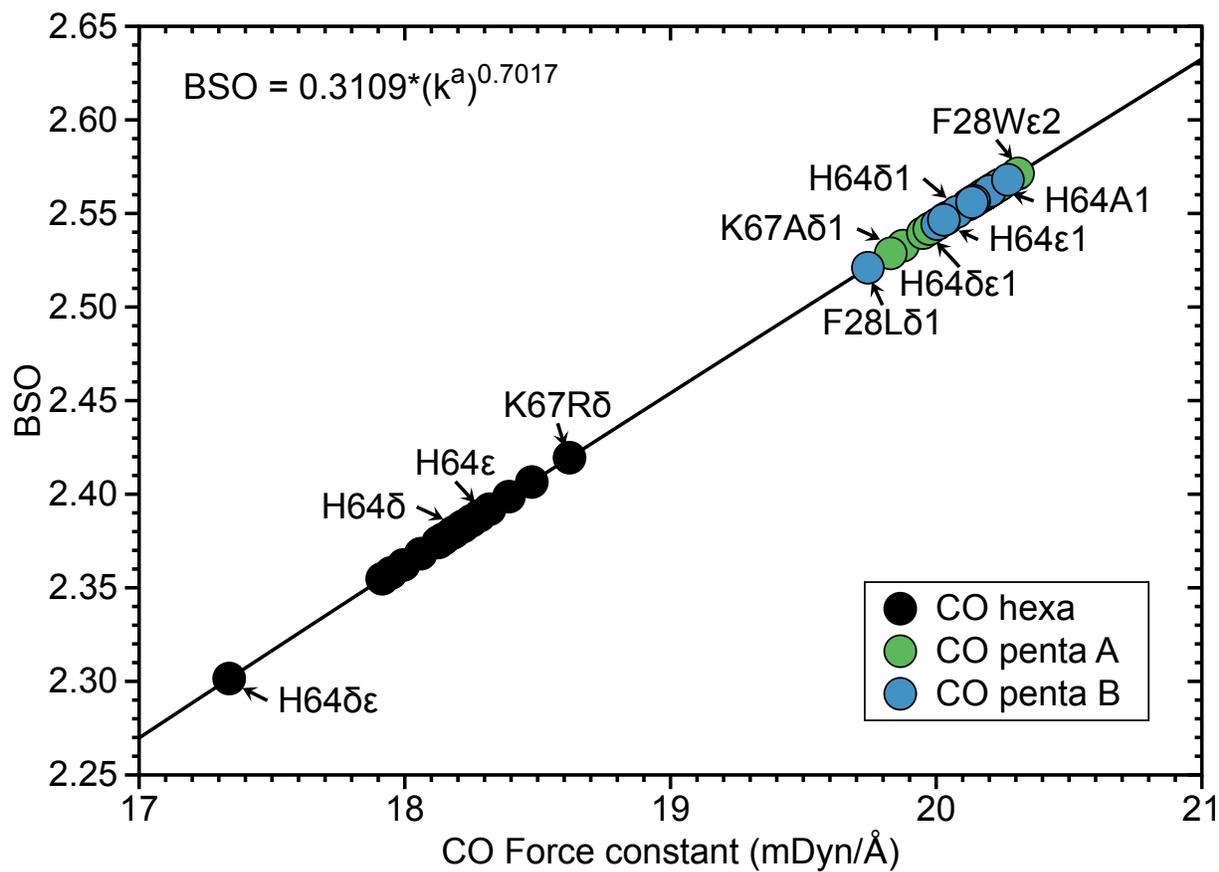


Figure 7
 Marek Freindorf, Alexis Delgado, Elfi Kraka
 J. Comput. Chem.

CO Bond	R	k^a	BSO	ω^a	ρ	H_ρ	C charge	O charge	POP
	Å	mDyn/Å		cm^{-1}	e/Bohr ³	Hr/Bohr ³	e	e	e
F106Aδ	1.1483	18.145	2.376	2119.4	0.4595	-0.7640	0.6216	-0.4879	0.3856
F106Aϵ	1.1482	18.221	2.383	2123.8	0.4599	-0.7654	0.6012	-0.4829	0.4028
F28Lδ	1.1494	17.948	2.358	2107.8	0.4585	-0.7624	0.6207	-0.4913	0.3856
F28Lϵ	1.1494	18.060	2.368	2114.4	0.4589	-0.7637	0.5937	-0.4784	0.3997
F28Wδ	1.1453	18.478	2.407	2138.8	0.4624	-0.7682	0.6158	-0.4697	0.3575
F28Wϵ	1.1502	17.915	2.355	2105.9	0.4572	-0.7592	0.6030	-0.4977	0.4084
H64A	1.1474	18.319	2.392	2129.6	0.4610	-0.7672	0.6090	-0.4766	0.3833
H64Q	1.1502	17.995	2.362	2110.6	0.4580	-0.7618	0.6076	-0.4955	0.4097
H64V	1.1472	18.392	2.399	2133.7	0.4614	-0.7685	0.6126	-0.4707	0.3783
H64δ	1.1474	18.130	2.375	2118.5	0.4607	-0.7665	0.6219	-0.4853	0.3827
H64ϵ	1.1479	18.281	2.389	2127.3	0.4603	-0.7660	0.5976	-0.4808	0.4039
H64$\delta\epsilon$	1.1550	17.339	2.301	2071.8	0.4523	-0.7506	0.5837	-0.5222	0.4630
K67Aδ	1.1475	18.214	2.382	2123.4	0.4602	-0.7652	0.6254	-0.4958	0.3868
K67Aϵ	1.1482	18.250	2.386	2125.5	0.4600	-0.7656	0.5908	-0.4825	0.4123
K67Rδ	1.1441	18.620	2.419	2146.9	0.4642	-0.7726	0.6114	-0.4573	0.3522
K67Rϵ	1.1489	18.125	2.374	2118.2	0.4591	-0.7637	0.5807	-0.4781	0.4097
K67Tδ	1.1472	18.178	2.379	2121.3	0.4606	-0.7661	0.6250	-0.4945	0.3777
K67Tϵ	1.1486	18.187	2.380	2121.9	0.4594	-0.7641	0.6046	-0.4880	0.4055
Gas	1.1453	18.614	2.420	2146.6	0.4631	-0.7710	0.6104	-0.4708	0.3711

Table 1: The bond length R, the local mode force constant k^a , the bond strength order BSO, the local mode frequency ω^a , the electron density at a bond critical point ρ , the energy density at a bond critical point H_ρ , the NBO atomic charge on C, the NBO atomic charge on O, and the total population of antibonding orbitals POP of CO bonded to Fe in hexacoordinate NgbCO complex. ω B97X-D/6-31G(d,p)/AMBER level of theory in the protein, and ω B97X-D/6-31G(d,p) in the gas phase. For descriptions of molecular labels, see the text.

FeC Bond	R	k^a	BCO	ω^a	ρ	H_ρ
	Å	mDyn/Å		cm^{-1}	e/Bohr ³	Hr/Bohr ³
F106A δ	1.7646	3.005	0.890	718.4	0.1489	-0.0514
F106A ϵ	1.7589	3.102	0.926	729.9	0.1508	-0.0529
F28L δ	1.7685	2.883	0.845	703.7	0.1472	-0.0500
F28L ϵ	1.7672	2.925	0.860	708.9	0.1474	-0.0505
F28W δ	1.7841	2.363	0.657	637.1	0.1424	-0.0476
F28W ϵ	1.7659	2.950	0.869	711.9	0.1481	-0.0514
H64A	1.7686	2.583	0.735	666.1	0.1481	-0.0521
H64Q	1.7556	3.273	0.991	749.9	0.1517	-0.0531
H64V	1.7664	2.957	0.872	712.8	0.1486	-0.0517
H64 δ	1.7653	2.660	0.763	676.0	0.1483	-0.0506
H64 ϵ	1.7590	3.166	0.950	737.5	0.1505	-0.0525
H64 $\delta\epsilon$	1.7464	3.145	0.942	735.0	0.1551	-0.0566
K67A δ	1.7647	2.987	0.883	716.4	0.1497	-0.0529
K67A ϵ	1.7573	3.189	0.959	740.2	0.1513	-0.0537
K67R δ	1.7822	2.482	0.699	653.0	0.1439	-0.0496
K67R ϵ	1.7641	2.954	0.871	712.3	0.1501	-0.0541
K67T δ	1.7720	2.743	0.793	686.4	0.1480	-0.0525
K67T ϵ	1.7587	3.142	0.941	734.7	0.1511	-0.0535
Gas	1.7756	2.586	0.736	666.5	0.1441	-0.0471
Fe(CO)_{5ax}	1.8020	2.768	0.802	689.5	0.1380	-0.0440
Fe(CO)_{5eq}	1.7891	3.003	0.889	718.3	0.1489	-0.0552

Table 2: The bond length R, the local mode force constant k^a , BSO, the local mode frequency ω^a , the electron density at a bond critical point ρ , the energy density at a bond critical point H_ρ of FeC in hexacoordinate NgbCO complexes. ω B97X-D/6-31G(d,p)/AMBER level of theory in the protein, and ω B97X-D/6-31G(d,p) in the gas phase. For descriptions of molecular labels, see the text.

	O...H-N				O...H-C				C...H-N			
	R	k ^a	BSO	H ρ	R	k ^a	BSO	H ρ	R	k ^a	BSO	H ρ
	Å	mDyn/Å		Hr/Bohr ³	Å	mDyn/Å		Hr/Bohr ³	Å	mDyn/Å		Hr/Bohr ³
F106A ϵ	3.1165	0.033	0.149	-	2.5863	0.059	0.181	0.0010	2.9114	0.068	0.190	0.0010
F28L ϵ	3.1338	0.062	0.185	-	2.6691	0.057	0.179	0.0010	2.8041	0.101	0.217	-
F28W ϵ	3.2489	0.044	0.164	-	2.4557	0.095	0.213	0.0009	2.9859	0.065	0.187	-
H64 ϵ	3.0394	0.038	0.157	-	2.6268	0.054	0.176	0.0010	2.7946	0.076	0.198	0.0012
H64 $\delta\epsilon$	2.9339	0.049	0.170	-	2.3582	0.102	0.218	0.0013	2.6406	0.115	0.227	-
K67A ϵ	2.8466	0.055	0.177	-	2.4391	0.100	0.217	0.0011	2.5895	0.111	0.224	0.0015
K67R ϵ	2.9880	0.068	0.190	-	2.5676	0.072	0.194	0.0011	2.5422	0.125	0.234	0.0015
K67T ϵ	3.1089	0.031	0.146	-	2.5324	0.064	0.186	0.0010	2.9400	0.053	0.175	-

Table 3: Local mode parameters for hydrogen bonding interactions between CO and hydrogens atom of His64, for hexacoordinate NgbCO complexes. Some gas phase references: O...H-O bond, water dimer: R = 1.926 Å, k^a = 0.225 mDyn/Å, BSO = 0.285, H ρ = -0.0014 Hartree/Bohr³; formaldehyde dimer: R = 2.457 Å, k^a = 0.044 mDyn/Å, BSO = 0.241, H ρ = 0.0004 Hartree; O...H-N bond, water imidazole dimer: R = 1.922 Å, k^a = 0.205 mDyn/Å, BSO = 0.358, H ρ = 0.0 Hartree/Bohr³. ω B97X-D/6-31G(d,p)/AMBER level of theory in the protein, and ω B97X-D/6-31G(d,p) in the gas phase.

	FeC(hexacoordinated)			CO(hexacoordinated)			CO(pentacoordinated)		
	ω^a	ω_{exp}	Ref.	ω^a	ω_{exp}	Ref.	ω^a	ω_{exp}	Ref.
	cm^{-1}	cm^{-1}		cm^{-1}	cm^{-1}		cm^{-1}	cm^{-1}	
F28L				2003 δ , 2009 ϵ	1937, 1957	¹⁰⁹	2100 δ 1, 2118 δ 2	2107, 2120	¹⁰⁹
					1967	¹⁰⁹	2122 ϵ 2, 2125 ϵ 1	2130, 2136	¹⁰⁹
								2142, 2149	¹⁰⁹
F28W				2001 ϵ , 2032 δ	1939, 1965	¹⁰⁹	2107 δ 1, 2111 ϵ 1	2117, 2128	¹⁰⁹
					1978	¹⁰⁹	2130 δ 2, 2130 ϵ 2	2136	¹⁰⁹
H64A	633	494	¹¹³	2023	1932, 1972	¹¹³			
H64V	677	494	¹¹²	2027	1932, 1972	¹¹³			
		494	¹¹³						
H64	642 δ , 701 ϵ	494, 521	¹¹²	2013 δ , 2021 ϵ	1932, 1956	¹¹³	2116 ϵ 1, 2122 δ 2	2105, 2114	¹⁰⁹
		494, 505	¹¹³		1972	¹¹³	2125 δ 1, 2127 ϵ 2	2123, 2129	¹⁰⁹
		521	¹¹³		1935, 1953	¹⁰⁹		2131, 2136	¹⁰⁹
					1966	¹⁰⁹			
K67A	681 δ , 703 ϵ	494, 521	¹¹³	2017 δ , 2019 ϵ	1934, 1956	¹¹³			
					1966	¹¹³			
K67R	620 δ , 677 ϵ	494, 505	¹¹³	2012 ϵ , 2040 δ	1934, 1962	¹¹³			
		521	¹¹³		1972	¹¹³			
K67T	652 δ , 698 ϵ	494, 521	¹¹³	2015 δ , 2016 ϵ	1934, 1962	¹¹³			
					1972	¹¹³			

Table 4: The calculated local mode frequencies ω^a and experimental normal mode frequencies ω_{exp} for selected NgbCO complexes investigated in this study. The labels δ and ϵ in the values of the local mode frequencies ω^a indicate the corresponding tautomer form of distal histidine H64. The labels 1 and 2 in the values of the local mode frequencies ω^a of pentacoordinate NgbCO complexes indicate the orientation of CO relative to the heme center, see text. Snapshots of selected residues of the active sites of hexacoordinate and pentacoordinate NgbCO complexes, are shown in Figures S1 – S10 of the Supplementary Material. The local mode frequencies ω^a are scaled by a factor 0.9501.¹²⁵ $\omega_{B97X-D/6-31G(d,p)/AMBER}$ level of theory.

CO Bond	R	k^a	BSO	ω^a	ρ	H_ρ	C charge	O charge	POP
Site A	Å	mDyn/Å		cm^{-1}	e/Bohr ³	Hr/Bohr ³	e	e	e
F106Aδ1	1.1366	19.979	2.542	2223.9	0.4700	-0.7639	0.5414	-0.5191	0.0109
F28Wδ1	1.1369	19.875	2.533	2218.2	0.4692	-0.7612	0.5476	-0.5323	0.0102
F28Wϵ1	1.1374	19.950	2.540	2222.3	0.4688	-0.7607	0.5547	-0.5422	0.0080
H64Q1	1.1356	20.230	2.564	2237.8	0.4708	-0.7645	0.5315	-0.5169	0.0064
H64δ1	1.1355	20.204	2.562	2236.4	0.4707	-0.7634	0.5226	-0.5163	0.0077
H64ϵ1	1.1365	20.029	2.547	2226.7	0.4696	-0.7618	0.5311	-0.5205	0.0096
H64δe1	1.1365	20.044	2.548	2227.5	0.4700	-0.7631	0.5255	-0.5115	0.0127
K67Aδ1	1.1361	19.829	2.529	2215.6	0.4701	-0.7625	0.5284	-0.5168	0.0095
K67Aϵ1	1.1356	20.212	2.563	2236.8	0.4708	-0.7646	0.5277	-0.5143	0.0058
K67Rδ1	1.1359	19.974	2.542	2223.7	0.4705	-0.7642	0.5333	-0.5170	0.0090
K67Tδ1	1.1357	20.214	2.563	2237.0	0.4707	-0.7632	0.5191	-0.5124	0.0055
K67Tϵ1	1.1359	20.137	2.556	2232.7	0.4702	-0.7626	0.5261	-0.5174	0.0075
F106Aδ2	1.1354	20.033	2.547	2226.9	0.4717	-0.7688	0.5246	-0.5016	0.0062
F28Wδ2	1.1348	20.305	2.571	2242.0	0.4720	-0.7678	0.5069	-0.4949	0.0047
F28Wϵ2	1.1345	20.308	2.571	2242.2	0.4724	-0.7684	0.4998	-0.4873	0.0070
H64Q2	1.1359	20.177	2.560	2234.9	0.4709	-0.7655	0.5182	-0.5043	0.0070
H64V2	1.1356	20.186	2.561	2235.4	0.4706	-0.7629	0.5214	-0.5163	0.0069
H64δ2	1.1359	20.148	2.557	2233.3	0.4706	-0.7641	0.5168	-0.5100	0.0073
H64ϵ2	1.1354	20.241	2.565	2238.5	0.4712	-0.7655	0.5086	-0.5010	0.0071
H64δe2	1.1351	20.240	2.565	2238.4	0.4717	-0.7666	0.5091	-0.4985	0.0080
K67Aδ2	1.1357	20.163	2.559	2234.1	0.4709	-0.7649	0.5179	-0.5080	0.0069
K67Aϵ2	1.1359	20.183	2.560	2235.3	0.4709	-0.7656	0.5171	-0.5054	0.0087
K67Rδ2	1.1357	20.188	2.561	2235.5	0.4711	-0.7663	0.5203	-0.5052	0.0062
K67Tδ2	1.1358	20.174	2.559	2234.8	0.4707	-0.7635	0.5158	-0.5109	0.0075
K67Tϵ2	1.1358	20.199	2.562	2236.1	0.4708	-0.7644	0.5143	-0.5057	0.0078
CO	1.1353	20.376	2.577	2245.9	0.4708	-0.7621	0.5145	-0.5145	0.0000

Table 5: Bond length R, local mode force constant k^a , bond strength order BSO, local mode frequency ω^a , electron density at bond critical point ρ , energy density at bond critical point H_ρ , NBO atomic charge on C, NBO atomic charge on O, and total population of antibonding orbitals POP of CO in pentacoordinate NgbCO complexes for docking form **A**. ω B97X-D/6-31G(d,p)/AMBER level of theory in the protein, and ω B97X-D/6-31G(d,p) in the gas phase. For descriptions of molecular labels, see the text.

CO Bond	R	k^a	BSO	ω^a	ρ	H_ρ	C charge	O charge	POP
Site B	Å	mDyn/Å		cm^{-1}	e/Bohr ³	Hr/Bohr ³	e	e	e
F106Aϵ1	1.1357	20.183	2.560	2235.3	0.4708	0.7665	0.5443	-0.5213	0.0055
F28Lδ1	1.1367	19.743	2.521	2210.8	0.4707	-0.7701	0.5880	-0.5019	0.0274
F28Lϵ1	1.1357	20.202	2.562	2236.3	0.4709	-0.7679	0.5586	-0.5201	0.0029
H64A1	1.1349	20.270	2.568	2240.1	0.4715	-0.7678	0.5336	-0.5091	0.0063
H64V1	1.1357	20.121	2.555	2231.8	0.4707	-0.7661	0.5273	-0.5120	0.0118
K67Rϵ1	1.1363	20.040	2.548	2227.3	0.4696	-0.7640	0.5473	-0.5224	0.0107
F106Aϵ2	1.1359	20.004	2.544	2225.3	0.4721	-0.7740	0.5568	-0.4707	0.0335
F28Lδ2	1.1347	20.078	2.551	2229.4	0.4707	-0.7675	0.5583	-0.5153	0.0118
F28Lϵ2	1.1352	20.143	2.557	2233.0	0.4727	-0.7743	0.5688	-0.4763	0.0216
H64A2	1.1359	20.135	2.556	2232.6	0.4712	-0.7685	0.5510	-0.4998	0.0167
K67Rϵ2	1.1355	20.029	2.547	2226.7	0.4724	-0.7736	0.5448	-0.4675	0.0287
CO	1.1353	20.376	2.577	2245.9	0.4708	-0.7621	0.5145	-0.5145	0.0000

Table 6: Bond length R, local mode force constant k^a , bond strength order BSO, local mode frequency ω^a , electron density at bond critical point ρ , energy density at bond critical point H_ρ , NBO atomic charge on C, NBO atomic charge on O, and total population of antibonding orbitals POP of CO in pentacoordinate NgbCO complexes for docking form **B**. ω B97X-D/6-31G(d,p)/AMBER level of theory in the protein, and ω B97X-D/6-31G(d,p) in the gas phase. For descriptions of molecular labels, see the text.