RSC Advances

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/advances

Journal Name

ARTICLE

Entry of water into the distal heme pocket of soluble Guanylate Cyclase β1 H-NOX domain alters the ligated CO structure: A resonance Raman and *in silico* simulation study

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

www.rsc.org/

Haoran Xu,[a] Yuebin Zhang,[ac] Lei Chen,[ad] Yan Li,[c] Chen Li,[b] Li Liu,[a] Takashi Ogura,[b] Teizo Kitagawa [b*] and Zhengqiang Li [a*]

Resonance Raman spectra (rRS) of the CO-*Ns* H-NOX (*heme nitric oxide and oxygen binding*) complex, a bacterial model for the mammalian CO-sGC (soluble Guanylyl Cyclase) complex, are reported. The CO-*Ns* H-NOX complex exhibited two Fe-CO stretching modes ($\nu_{Fe\text{-}CO}$) with an intense band at 470 cm^{-1} and a weaker band at 494 cm^{-1}, which are very similar to those observed for a CO-ligated full-length lung sGC. However, the rRS of truncated sGC β1 $_{(1\text{-}194)}$ H-NOX exhibited comparable bands, but with a reversed intensity ratio, the $\nu_{Fe\text{-}CO}$ at 492 cm^{-1} was dominant. This difference is similar to that observed for six-coordinate CO-sGC's in the absence and presence of YC-1, an effector. Molecular dynamic (MD) simulations of truncated sGC β1 $_{(1\text{-}194)}$ H-NOX demonstrate that water may enter the distal heme pocket to form a hydrogen bond with ligated CO; whereas it hardly happens to *Ns* H-NOX. By employing density functional theory (DFT) analysis, we calculated the C-O and Fe-CO stretching frequencies of the isolated imidazole (ImH)- and porphyrin (P)-containing complex, (ImH)FeP(CO), in the presence or absence of water molecules. The calculated rRS exhibited bands at a higher $\nu_{Fe\text{-}CO}$ frequency in the presence of water, in agreement with observations for truncated sGC, further supporting the hypothesis that water access to the distal heme pocket influences the ligation of CO, although the proximal effects remain to be elucidated.

Introduction

Soluble guanylyl cyclase (sGC) is an enzyme which catalyses the conversion of GTP to cyclic GMP, which subsequently regulates various physiological functions such as smooth muscle contraction, platelet aggregation and neuro-transmission.[1] Mammalian sGC is generally a heterodimeric hemeprotein composed of a larger α1 subunit (~80 kDa) and a smaller β1 subunit (~70 kDa).[2, 3] A protoheme resides only in the β1 subunit and remains in a reduced form when the protein is isolated from tissue and its heme iron is coordinated by His105 residue, although O_2 is not bound to it. Upon binding of NO to the heme cofactor of sGC (NO-sGC), the sGC catalytic activity increases more than 200-fold.[4, 5] CO can also ligate to the heme cofactor of sGC (CO-sGC) and activates the enzyme, but only 2 to 6-fold over basal enzymatic activity.[6] Notably, in the presence of an organic effector molecule such as YC-1, CO activates sGC to the same magnitude as does NO.[7-10] The binding of NO to sGC generates a five-coordinate nitrosyl complex of heme by disrupting the Fe-His bond. This structural change of the heme pocket in the N-terminal region is then transmitted to the catalytic domain located in the C-terminal region via conformational changes that alter the entire protein structure.[11, 12]

Unlike NO, CO forms a six-coordinate carbonyl complex (CO-heme),[13] which is distinguishable from the original six-coordinate CO-heme with regard to Fe-CO stretching frequencies.[9] In addition, from rRS observations it has been demonstrated that in the presence of effector YC-1, a five-coordinate carbonyl complex distinct from the six-coordinate CO-heme complex could also form.[9] We have examined this phenomenon in several ways using model systems of the heme binding domain of sGC. The smallest fragment of the β1 subunit that conserves the ligand binding characteristics of full-length sGC was mapped to the first 1-194 amino acids at the N-terminus.[8] Genome sequence alignments have determined that the heme binding region (1-194) of the β1 subunit is homologous to that of the H-NOX domain commonly observed in the sGC family.[14] Several prokaryotic homologs of sGC β1 H-NOX have been identified which exhibited high similarity to sGC β1 H-NOX in both sequence and ligand binding properties. Among those prokaryotic homologs, crystal structures of H-NOX domains have been determined with two

[a.] *Key Laboratory for Molecular Enzymology & Engineering, the Ministry of Education, School of Life Sciences, Jilin University, Changchun 130012, P.R. China. Email: lzq@jlu.edu.cn; Tel: +86-431-85155201.*
[b.] *Picobiology Institute, Graduate School of Life Science, University of Hyogo, RSC-UH Leading Program Center, 1-1-1 Koto, Sayo-cho, Sayo-gun, Hyogo 679-5148, Japan. Email: teizo@sci.u-hyogo.ac.jp.*
[c.] *State Key Laboratory of Molecular Reaction Dynamics, Dalian Institute of Chemical Physics, Chinese Academy of Sciences, Dalian 116023, P.R. China.*
[d.] *Institute of Atomic and Molecular Physics, Jilin University, Changchun, 130012, P.R. China.*
Electronic Supplementary Information (ESI) available: See
DOI: 10.1039/x0xx00000x

species; one from *Thermoanaerobacter tengcongensis* (*Tt* H-NOX)[15] and the other from *Nostoc spheroids* (*Ns* H-NOX).[11]

The structural information available from these prokaryotic proteins should provide deeper understanding of the mechanisms of sGC activation.[11, 16-18] *Ns* H-NOX shares higher sequence identities with sGC than *Tt* H-NOX (33% vs 18%) and this higher similarity also extends to the heme pocket region, in which 17 of 27 residues are identical between *Ns* H-NOX and sGC β1 H-NOX.[19] Moreover, *Ns* H-NOX exhibits the same ligand binding properties as sGC in that both NO and CO could bind to the heme cofactor, whereas the binding of O_2 is prohibited.[11, 20] In contrast, the *Tt* H-NOX can bind O_2 akin to a function expected for an oxygen sensor. Specifically, the Y140 residue in the distal heme pocket of *Tt* H-NOX forms a hydrogen bond to bound O_2, stabilizing the oxygen-bound form. On the contrary, the heme distal environments of both *Ns* H-NOX and sGC are predominantly hydrophobic.[21-23] Thus, *Ns* H-NOX appears to serve as a superior model for sGC than *Tt* H-NOX.

Accordingly, in this work we examined the rRS of the CO-bound form of *Ns* H-NOX (CO-*Ns* H-NOX). The rRS of CO-*Ns* H-NOX exhibited close resemblance to that of the full-length heterodimeric sGC purified from bovine lung tissue.[9] Two forms of CO-Ns H-NOX are generated in the same way as full-length sGC, giving rise to two $\nu_{Fe\text{-}CO}$ bands with different intensity. However, the rRS for the CO complex of truncated sGC β1 $_{(1\text{-}194)}$ H-NOX exhibited a reserved intensity which are similar to the bands observed for full-length sGC in the presence of YC-1 but not for native sGC and CO-*Ns* H-NOX. Therefore, protein truncation or addition of an effector both appear to alter the heme pocket structure and thus influence the maximal amount of heme-ligated CO. To gain a better understanding of the structural differences between *Ns* H-NOX and sGC β1 H-NOX, molecular dynamic (MD) simulations were employed to investigate the dynamic behaviours of CO-bound forms of these H-NOX domains. Water was found to be the key reason why there are two different vibration bands of CO complexes. That is water entering the distal pocket affects the heme-ligated CO. An appreciable proportion of sGC β1 H-NOX proteins undergo this water behaviour, whereas for *Ns* H-NOX there is no long-term retention of water molecules in the distal heme pocket. Density functional theory (DFT) was utilized to calculate the ν_{CO} and $\nu_{Fe\text{-}CO}$ frequencies for an isolated CO-heme model compound (ImH)FeP(CO) in the presence or absence of water molecules. The calculated vibrational frequencies are in agreement with the results observed for rRS, which further confirmed our hypothesis.

Nowadays, scientists have realized that water not only serves as a dispersion medium or solvent but plays a significant role in bio-signal transition system.[24] Combination of the Raman spectroscopy, MD simulation and DFT calculation technologies would discover more ultra-micro scale information of signal transfer and downstream functional mechanisms, from vibration information of chemical bond to conformation and coordinate information of the whole system, thus invite a new interpretation about the effects of protein truncation or the presence of an effector on the resulting molecular interactions within the heme pocket.

Experimental

Materials

The *hnox* gene coding for *Ns* H-NOX fused to an anaerobic promoter was constructed by Sangon Biotech (Shanghai, China). Plasmids pET-20b, pUC57 and pUC19 were purchased from Takara (Japan). *Pfu* polymerase was provided by Promega (USA) and *Bbs*I, *Pst*I and *Hind*III restriction enzymes were obtained from NEB (USA). *E.coli* DH 5α and BL21 (DE3) were purchased from Millipore (formerly Novagen, USA). The plasmid containing rat lung sGC cDNA was kindly provided by Michael A. Marletta, Ph.D.[8]

DEAE sepharose and G-75 sephadex were purchased from GE Healthcare Life Sciences (USA). The 10DG desalting column was purchased from Bio-Rad (USA). CO gas was purchased from Taiyo Nippon Sanso Corp. (Japan) and CO isotope was sourced from Cambridge Isotope Laboratories, Inc. (USA). All aqueous solutions were prepared with Milli-Q purified water. All other reagents were of analytical grade, and were commercially available.

Expression and purification of sGC β1 $_{(1\text{-}194)}$ and *Ns* H-NOX

The sGC β1 $_{(1\text{-}194)}$ subunit was expressed as previously described.[25] Briefly, a DNA fragment containing the *Ns* H-NOX gene fused to an upstream anaerobic promoter was amplified using a polymerase chain reaction (PCR) performed with *Pfu* polymerase. The PCR product was cloned into the pUC57 vector and transformed into *E. coli* DH 5α, after culture the plasmid was purified then digested using restriction enzymes *Hind*III and *Pst*I. The promoter-*Ns* H-NOX fragment was excised from plasmid DNA using restriction enzymes *Pst*I and *Bbs*I. Finally this promoter-*Ns* H-NOX fragment was cloned into the pUC19 plasmid and transformed into *E. coli* BL21 (DE3) cells for protein expression.

The purification procedure of the *Ns* H-NOX domain was performed as previously described for *Vitreoscilla* hemoglobin.[26] Protein expression was induced in an anaerobic environment at 25°C with shaking at 120 rpm for approximately 24 h. The supernatant was purified using a two-step ammonium sulfate precipitation procedure (40% and 60%). The resulting supernatant was loaded stepwise onto a DEAE ion exchange column followed by purification using a G-75 filtration column. Purification was completed using the AKTA Protein Purification System (GE Healthcare Life Sciences, USA). The sample was dissolved in 20 mM PBS containing 500 mM NaCl (pH 7.5).

A UV-2700 spectrophotometer (Shimadzu, Japan) was used to determine the Reinheitszahl (Rz) ratio, a measure of heme-protein purity. Only samples with Rz >1.5 and with Rz >2.5 were collected for further studies of sGC β1 $_{(1\text{-}194)}$ and *Ns* H-NOX, respectively. Protein concentrations were determined using both a BCA Assay Kit and the Lambert-Beer law using the molar extinction coefficient ε = 170 $mM^{-1}cm^{-1}$ at

430 nm.[27] SDS-PAGE further confirmed the purity and the molecular weight of purified proteins.

Resonance Raman spectroscopy

Preparation of CO-bound complex

Purified sGC β1 $_{(1-194)}$ and *Ns* H-NOX proteins were diluted to 50 µM in 20 mM PBS buffer and 63 µL of each solution separately analysed using an airtight spinning Raman cell. After several degassing steps, 7 µL 100 mM sodium dithionite was added to the cell via a syringe. Finally, CO was incorporated into the upper space of the Raman cell by injection through a rubber septum using an airtight syringe.

Resonance Raman measurement

Raman scattering was performed using excitation with the 413.1 nm line of a krypton laser (Spectra Physics, model BeamLok 2060). The scattered light at right angles was collected and dispersed using a single polychromator (Ritsu Oyo Kogaku, Japan), and detected using a liquid nitrogen-cooled charge-coupled device (CCD, Princeton Instruments, USA). Raman shifts were calibrated using indene. The wavenumber accuracy was ±1 cm^{-1} for well-defined peaks. The laser power was between 1 mW and 0.1 mW at the sample. The accumulation times were typically 0.5 h. UV-visible absorption spectra were measured using a U-3310 spectrophotometer (Hitachi, Japan) before and after each Raman measurement to confirm the integrity of each sample (Table S1, See supporting Information Table S1).

Molecular dynamic (MD) simulations of CO-bound *Ns* H-NOX and sGC β1 H-NOX

Since the structural information for the sGC H-NOX domain is not available now, we first built the homology model of the sGC β1 H-NOX domain (1-183) using the crystal structure of *Ns* H-NOX in the public database of the Research Collaboratory for Structural Bioinformatics (RCSB PDB ID: 2O09[11]) as a template in order to compare the differences in dynamic behaviours between *Ns* H-NOX and sGC β1 H-NOX. The sequence alignment of the two proteins is provided in the Supporting information and the building of the homology model was performed using the MODELLER package.[28] A total of 50 models were generated and the conformation with the lowest DOPE (Discrete Optimized Protein Energy) score was selected for MD simulation.

The MD simulations of both CO-ligated *Ns* H-NOX and sGC β1 H-NOX were conducted using Gromacs 5.0.4[29] with Charmm36 force field parameters.[30] First, the initial structures of *Ns* H-NOX (PDB ID: 2O09[11]) and sGC β1 H-NOX were centred into a 80 x 80 x 80 $Å^3$ cubic box. Next the systems were solvated using TIP3P water models.[31] Water molecules within the distance of 7 Å from the ligand CO were removed to set the initial hydrophobic environment at the distal heme pocket of both H-NOXs. 0.1M NaCl was added to balance the charges of the whole system. The steepest descent algorithm was employed to minimize the energy of the whole system before it was gradually heated to 300K with an increment of 20K in a total of 50,000 steps under an NVT ensemble with restraints set for the motion of the protein backbone. Next, we switched the system to an NPT ensemble using the Berendsen algorithm[32] for pressure coupling at 1 bar and the V-rescale algorithm[33] for temperature coupling at 300K. The NPT equilibration was maintained for 2 ns before performing the production run, in which the Parrinello-Rahman barostat[34] was used to control the pressure of the system at 1 bar and the temperature was coupled using the Nose-Hoover algorithm.[35] The particle mesh Ewald method[36] was used to compute the electrostatic interactions with a real-space cut off distance of 1.2 nm. The same cut off value was chosen when determining the van der Waals interactions. The SETTLE algorithm[37] was used to constrain water molecules and all non-water bonds were constrained using the LINE algorithm,[38] which ensured the stability of the system using an integration time-step of 2 fs.

Density functional theory (DFT) calculations of vibrational frequencies

All calculations were performed using DFT methods implemented in the Gaussian09[39] software package. The geometric optimization and vibrational analysis of the CO-Fe^{II} porphyrin complex in the presence or absence of water molecules were carried out using the B3LYP[40, 41] functional. The 6-31+g (d, p) standard basis set was used for N, C, O, and H, but only the Fe atom was analysed using the LanL2DZ pseudo-potential basis.[42] Vibrational frequency analysis was performed on the basis of the optimized geometries and showed that all of the electronic ground states were stable.

Results and Discussion

Electronic absorption spectra of Fe^{II}-unligated and CO-ligated *Ns* H-NOX and sGC β1 $_{(1-194)}$ H-NOX

The *Ns* H-NOX was isolated as described previously[11] in the Fe^{II}-unligated state. In contrast, sGC β1 $_{(1-194)}$ H-NOX was found to oxidize readily with a fast autoxidation rate of 0.073/min,[8] which differs from very slow rates observed for both full-length sGC and *Ns* H-NOX.[8, 19] To maintain sGC β1 $_{(1-194)}$ H-NOX in the reduced form, dithionite was needed. Fig. 1

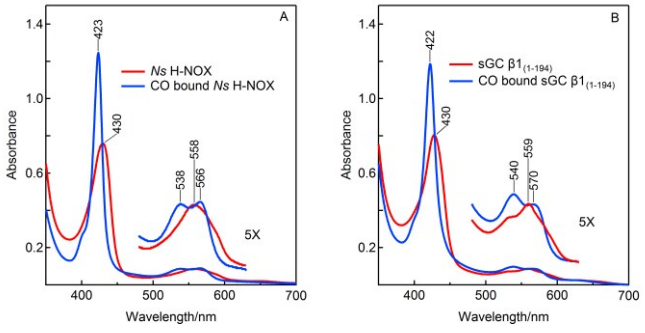


Fig. 1 Electronic absorption spectra for 5 µM *Ns* H-NOX (A) and sGC β1 $_{(1-194)}$ (B); inserts for these respective complexes depict 5-fold magnification of the 480-630 nm wavelength regions of the spectra; blue lines depict the reduced unligated form, red lines depict the CO-bound form.

Table 1. The peak positions (nm) of UV and visible spectra

Protein	Soret	β	α	Ref.
Fe^{II} unligated complex				
sGC	431	555		Ref.[43]
sGC β1 $_{(1-194)}$	430	559		This work
Ns H-NOX	430	558		This work
Fe^{II}-CO complex				
sGC	423	541	567	Ref.[43]
sGC β1 $_{(1-194)}$	423	541	571	Ref.[8]
sGC β1 $_{(1-194)}$	422	540	570	This work
Ns H-NOX	423	538	566	This work

depicts the electronic absorption spectra of both proteins in the unligated (red) and CO-ligated (blue) forms. The SDS-PAGE of these proteins is shown in Fig. S1 and demonstrates the purity of the samples. The peak wavelengths of spectra are shown in Fig. 1 and peak positions of proteins studied here are summarized in Table 1.

The *Ns* H-NOX electronic absorption spectrum is very similar to those of sGC β1 $_{(1-194)}$ H-NOX and full-length sGC.[8, 44] The position of the Soret band at 430 nm and a single α/β band at 558 nm suggest that the reduced H-NOX form adopts a five-coordinate, high-spin state. By adding CO into the systems under anaerobic conditions, the Soret band is shifted to 423 nm and the α/β band is split into two bands at 538 nm and 566 nm. These observations support the formation of a six-coordinated ferrous low-spin state, CO bound *Ns* H-NOX, whereas for sGC β1 $_{(1-194)}$ H-NOX, the Soret band is shifted to 422 nm and the α/β band split into 540 nm and 570 nm upon CO binding.

Resonance Raman spectroscopy of CO-bound *Ns* H-NOX and sGC β1 $_{(1-194)}$ H-NOX

The rRS of the Fe^{II}-CO complexes of *Ns* H-NOX and sGC β1 $_{(1-194)}$ are shown in Fig. 2 and Fig. S3. The assignments of major bands from these spectra are listed in Table 2, where their frequencies are compared to those of myoglobin, sGC full-length protein, sGC β1 $_{(1-385)}$ subunit and *Tt* H-NOX.

The π-electron density marker band frequency, ν_4, reflects the oxidation state of heme. The ν_{10}, ν_2, and ν_3 are marker band frequencies for the spin and coordination states, which are basically sensitive to the core size of the heme macrocycle.[45] The ν_2, ν_3, ν_4 and ν_{10} bands of CO-complexed *Ns* H-NOX appear at 1583 cm^{-1}, 1496 cm^{-1}, 1370 cm^{-1}, and 1627 cm^{-1}, respectively (Fig. 2A), indicating a six-coordinate, ferrous low-spin heme state. The sharp ν_4 band at 1370 cm^{-1} lacking a low-frequency shoulder indicates that no photolysis occurred during the rRS measurements. In the reduced form of *Ns* H-NOX, ν_2, ν_3, ν_4 and ν_{10} bands are observed at 1559 cm^{-1}, 1472 cm^{-1}, 1357 cm^{-1} and 1626 cm^{-1}, indicating a five-coordinate,

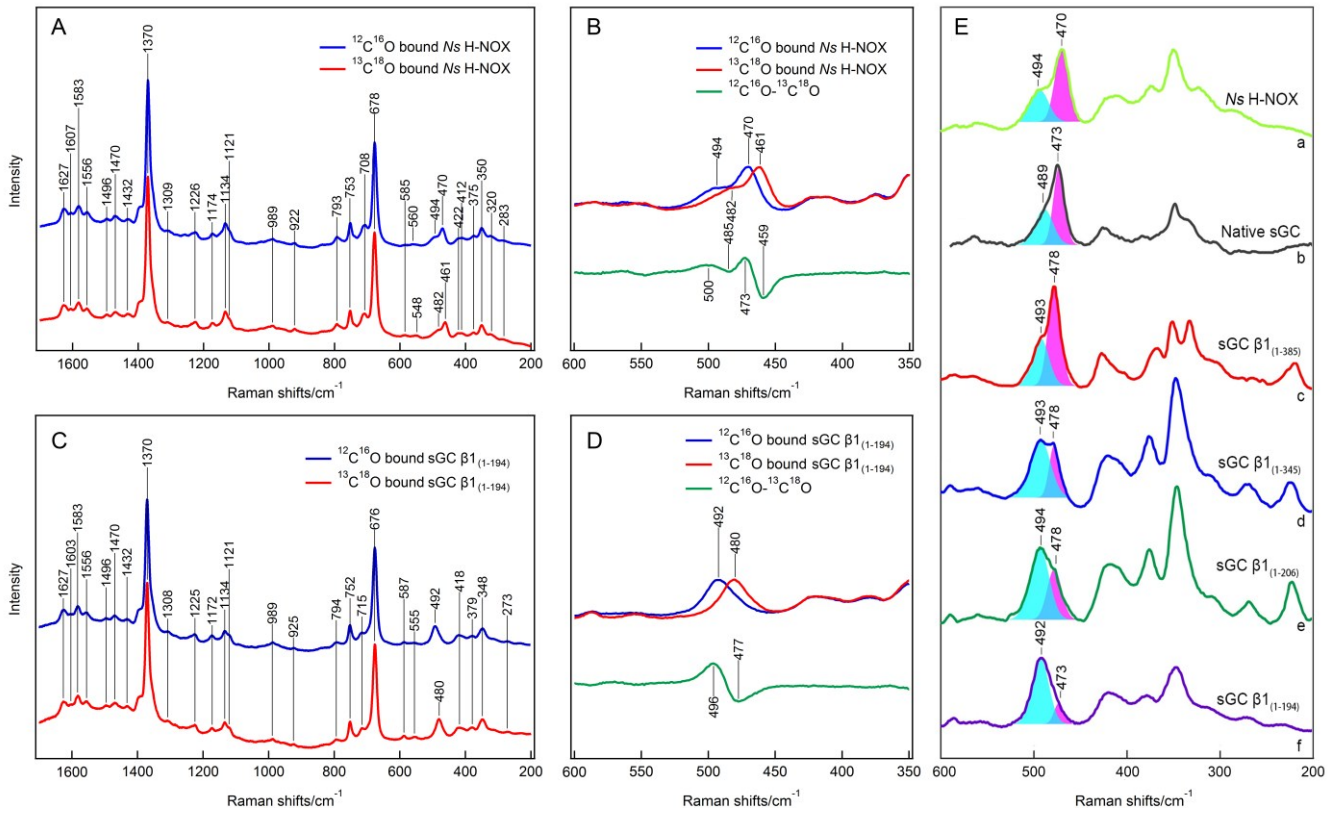


Fig. 2 (A, C) Resonance Raman spectra of H-NOX domains of 50 μM CO-*Ns* H-NOX and CO-sGC β1 $_{(1-194)}$ in PBS buffer; (B, D) magnification of their superimposed spectra in the Fe-CO stretching frequency region. $^{12}C^{16}O$ and $^{13}C^{18}O$ complexes are represented by blue and red lines, respectively, and their 1:1 difference spectrum ($^{12}C^{16}O$-$^{13}C^{18}O$) by green lines, the excitation wavelength is 413.1 nm, the laser power is 170 μW. (E) The resonance Raman spectra in the low frequency region (600-200 cm^{-1}) of proteins studied herein are also shown for comparison purposes (a part of the data is adapted from the literature (b),[9] (c)-(e)[13]: (a) *Ns* H-NOX, (b) sGC, (c) sGC β1 $_{(1-385)}$, (d) sGC β1 $_{(1-345)}$, (e) sGC β1 $_{(1-206)}$, (f) sGC β1 $_{(1-194)}$. Peak-differentiation is shown by blue and pink shading.

Table 2. Frequencies of several marker Raman bands of heme and CO-associated modes for myoglobin and different isoforms of Fe^{II}-sGC protein complexes.

Proteins	ν_{10}	ν_2	ν_3	ν_4	$\nu_{Fe\text{-}CO}$	$\delta_{Fe\text{-}C\text{-}O}$	ν_{CO}	Ref.
Mb	1637	1587	1498	1372	512	577	1944	Ref.[46]
sGC	1631	1583	1500	1372	472/488	565	1985/1969	Ref.[9]
sGC β1 $_{(1\text{-}385)}$	1629	1582	1496	1373	478/494	564	1985	Ref.[13, 47]
sGC β1 $_{(1\text{-}194)}$	1626	1583	1498	1371	493		1969	Ref.[47]
sGC β1 $_{(1\text{-}194)}$	1627	1583	1496	1370	492	555	1970	This work
Tt H-NOX		1581	1494	1369	490	569	1989	Ref.[10, 48]
Ns H-NOX	1627	1583	1496	1370	470/494	560	1986	This work

high-spin ferrous state of heme (Fig. S2). The low-frequency spectra of $^{12}C^{16}O$- or $^{13}C^{18}O$-bound *Ns* H-NOX and sGC β1 $_{(1\text{-}194)}$ are shown in Fig. 2B and Fig. 2D, respectively, together with their $^{12}C^{16}O$-$^{13}C^{18}O$ difference spectra (green). Two isotope sensitive bands are identified at 470 cm^{-1} and 494 cm^{-1} for $^{12}C^{16}O$-*Ns* H-NOX, which are assigned to the Fe-CO stretching vibration.[13, 49] A more intense band at 470 cm^{-1} and a weaker band at 494 cm^{-1} for $^{12}C^{16}O$-*Ns* H-NOX suggest that two distinct conformations of the Fe-CO complex are present whose relative amounts are proportional to their intensities. However, in the case of CO-sGC β1 $_{(1\text{-}194)}$, it seems that there is only a single isotope sensitive band at 492 cm^{-1} (Fig. 2D). The high-frequency spectra are shown in Fig. S3, C-O stretching vibration band at 1986cm^{-1} and 1970cm^{-1}, shift to 1897cm^{-1} and 1884cm^{-1}, with $^{13}C^{18}O$ isotope instead of $^{12}C^{16}O$, for *Ns* H-NOX and sGC β1 $_{(1\text{-}194)}$, respectively.

By further analysing the low-frequency spectrum using a Gaussian fitting method, we can identify a small shoulder at 473 cm^{-1} in the spectra of the CO-sGC β1 $_{(1\text{-}194)}$ complex (Fig. 2E). Moreover, in Fig. 2E, the low-frequency spectrum of CO-bound *Ns* H-NOX is compared to spectra of native sGC and recombinant truncated β subunits of various lengths.[9, 13, 50]. As can be seen from Fig. 3E, the two Fe-CO stretching modes of *Ns* H-NOX are very similar to that of full-length sGC. However, as the length of truncation increases, the intensities of these two bands become reversed. The sGC β1 $_{(1\text{-}345)}$ demonstrates nearly equal intensities at 493 cm^{-1} and 478 cm^{-1}, whereas in the case of sGC β1 $_{(1\text{-}194)}$, the $\nu_{Fe\text{-}CO}$ band at 492 cm^{-1} becomes dominant. The intensity ratios of the two bands are listed in Table 3. It is noted that in our previous rRS experiments[7, 9] the intensity ratio of the two $\nu_{Fe\text{-}CO}$ modes of full-length sGC was altered by addition of YC-1, with a similar trend seen with increasing truncation in this work.

Table 3. Frequencies and relative intensities of the Fe-CO stretching Raman bands ($\nu_{Fe\text{-}CO}$) for different sGC truncated heme proteins.

proteins	$\nu_{Fe\text{-}CO}$ conformation		Intensity	Ref.
	I	II	Ratio I/II	
Ns H-NOX	494	**470**	0.44	this work
sGC full-length	489	**473**	0.45	Ref.[9]
sGC β1 $_{(1\text{-}385)}$	493	**478**	0.49	Ref.[13]
sGC β1 $_{(1\text{-}345)}$	**493**	478	1.10	Ref.[13]
sGC β1 $_{(1\text{-}206)}$	**494**	478	1.42	Ref.[13]
sGC β1 $_{(1\text{-}194)}$	**492**	473	3.11	this work

Molecular dynamic (MD) simulations of CO-bound *Ns* H-NOX and sGC β1 H-NOX

To address the question of why the two H-NOX domains (*Ns* vs sGC β1) exhibit distinct $\nu_{Fe\text{-}CO}$ modes, we first employed MD simulations to investigate how the protein matrix modulates the heme environment within H-NOX domains.

Since the structural information of sGC H-NOX domain is not yet available, we used the crystal structure of *Ns* H-NOX as a template to build the homology model of the sGC β1 H-NOX domain (1-183). Next, we performed 400 ns MD simulations for both CO-*Ns* H-NOX and CO-sGC β1 H-NOX, respectively. In the first 100 ns MD simulation, both proteins experienced large fluctuations before reaching equilibrium after 150 ns (Fig. S4). By analysing our MD trajectories, we found that the equilibrated structures of CO-sGC β1 H-NOX demonstrated a large shifting motion of the heme group compared with that of *Ns* H-NOX. This shift is primarily modulated by a conformational bending of the C-terminal αF segment (Fig. 3B), where the residue H105 is located. Upon further inspection of the structure, we found the bending of αF is caused by a stable formation of a hydrogen bond between Tyr112 (Y112) at the end of αF and the conserved Tyr83 (Y83) within the hinge region between αD and αE (Fig. S5). However, in the case of *Ns* H-NOX, the residue at position 112 is Phe (F), and therefore, no hydrogen bond is formed between the two residues in *Ns* H-NOX. Moreover, the shifting motion of heme in sGC β1 $_{(1\text{-}194)}$ H-NOX also leads to a structural change in the protein which opens a portal and allows water access to the distal heme pocket. As a result, the water molecule can form a hydrogen bond with bound CO in the case of sGC β1 $_{(1\text{-}194)}$ H-NOX (Fig. 3C). We evaluated the minimum distance between the CO oxygen atom and all of the water hydrogen atoms over the two MD trajectories and the results are illustrated in Fig. 3A and Fig. 3C. It is apparent from these figures that water molecules can hardly approach the distal heme pocket to form a hydrogen bond with CO. However, in the case of *Ns* H-NOX, although there are a small number of water molecules which occasionally enter the distal pocket, water would be expelled rapidly due to the short path of the Y-shaped tunnel in the *Ns* H-NOX.[21, 22] Conversely, in the case of sGC β1 $_{(1\text{-}194)}$ H-NOX, interactions between CO oxygen and water hydrogen are observed for a large population after 150 ns of our MD simulation. Thus, we hypothesize that a portion of the water molecules in the distal heme pocket could form hydrogen bond with CO to change the conformation of the heme pocket, resulting in an additional Fe-CO stretching band at a distinct

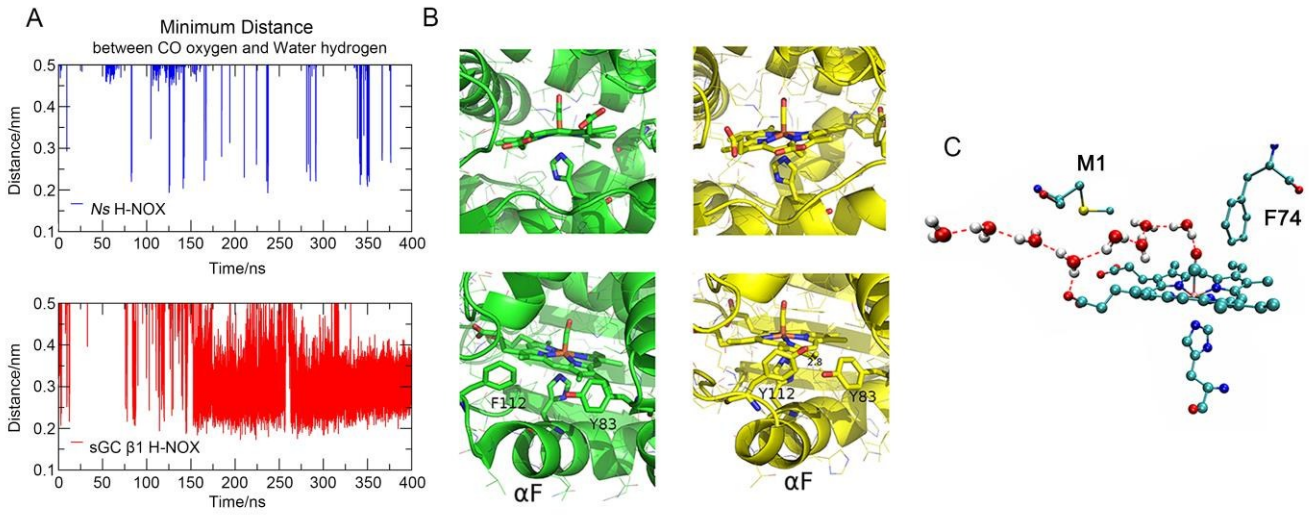


Fig. 3 (A) The minimum distance between the oxygen atom of the ligated CO and the hydrogen atoms of water molecules during 400 ns MD simulations; blue line depicts *Ns* H-NOX (top) and red line depicts sGC H-NOX (bottom). Only the minimum distances below 0.5 nm are recorded. (B) Depiction of protein conformations near the heme pocket during the simulation illustrating the heme positional shift in the equilibrated conformation of sGC H-NOX (yellow drawing) compared to that of *Ns* H-NOX (green drawing). The figures in the lower level of (B) represent the structures after rotating the conformations shown in the upper level by 90 degrees to the left. It is apparent that the bending of the αF segment in sGC H-NOX (yellow) is stabilized by the formation of hydrogen bond between Y112 and Y83, whereas no corresponding hydrogen bond is formed in *Ns* H-NOX (green). (C) Depicts the heme group, ligated CO and Y112 in sGC H-NOX, F112 in *Ns* H-NOX and depicts the conserved Y83 by using stick figures. The hydrogen bond networks of water (red circles for oxygen atoms) indicate the water molecules enter the distal heme pocket of sGC H-NOX via the portal between the propionate group of heme and the residue Met1 (M1). Formation of a hydrogen bond is apparent between one of the water molecules and the ligated CO.

frequency. In addition, the greater exposure of the distal heme environment of sGC β1 $_{(1\text{-}194)}$ H-NOX to the solvent also explains the fast oxidation rate of the truncated sGC β1 subunit compared to Ns H-NOX.[19] Accordingly, we propose that the dominant $\nu_{Fe\text{-}CO}$ band at 470 cm^{-1} in CO-*Ns* H- NOX is shifted to 492 cm^{-1} in sGC β1 $_{(1\text{-}194)}$ H-NOX and that this shift is caused by the formation of a hydrogen bond between water and ligated CO.

The calculations of vibrational frequencies using DFT

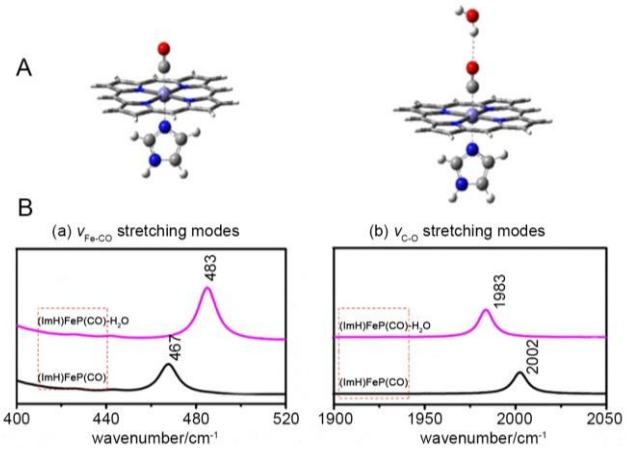


Fig. 4 (A) The optimized structures of the isolated (ImH)FeP(CO) (ImH: imidazole, P:porphyrin) generated by DFT calculations in the absence (left) and presence (right) of a water molecule. In the presence of water, a Fe^{II}-CO···H_2O structure is postulated. (B) Calculated rRS of the CO adduct of heme in the absence (left) or presence (right) of a water molecule for the low-frequency (a) and high-frequency spectral regions (b).

To further test our hypothesis that water entering the distal pocket affects the conformation of the ligated CO, we utilized the DFT method to calculate the vibrational frequencies of the isolated (ImH)FeP(CO) in the presence or absence of water molecules. The optimized structures and the calculated rRS results are shown in Fig. 4 and the optimized geometrical parameters are listed in Table S2. As can be seen from our calculated Raman spectra, both the Fe-CO ($\nu_{Fe\text{-}CO}$) and C-O vibrational frequencies (ν_{CO}) are indeed influenced by the presence of water. In the absence of water, the calculated vibrational modes for $\nu_{Fe\text{-}CO}$ and ν_{CO} are 467 cm^{-1} and 2002 cm^{-1}, respectively. However, in the presence of water, the calculated spectrum of the complex exhibits a higher $\nu_{Fe\text{-}CO}$ frequency at 483 cm^{-1} and a lower ν_{CO} frequency at 1983 cm^{-1}, which is qualitatively in agreement with our resonance Raman observations. This strongly supports our hypothesis that water molecules entering the distal heme pocket of sGC β1 $_{(1\text{-}194)}$ H-NOX indeed affect the vibrational mode of ligated CO, resulting in an increase of $\nu_{Fe\text{-}CO}$ and a decrease of ν_{CO}.

Conclusions

In summary, two $\nu_{Fe\text{-}CO}$ Raman bands were observed for the CO-*Ns* H-NOX complex, similar to the full-length sGC. The observation of two $\nu_{Fe\text{-}CO}$ modes indicate the presence of two conformations of ligated CO in CO-*Ns* H-NOX. However, the relative intensities of the two bands observed are reversed in sGC β1 $_{(1\text{-}194)}$ H-NOX. From the MD trajectories, a large shift in position of the heme group occurs upon CO-binding to sGC β1 H-NOX, this shift was modulated by the bending of the C-

terminal αF segment which is stably maintained by the formation of a hydrogen bond between Y112 at the end of αF and the conserved Y83 within the hinge region between αD and αE. Moreover, the shifting motion of heme in sGC β1 (1-194) H-NOX also leads to a protein conformational change resulting in the opening of a portal formed by the propionate group of heme and the protein matrix. The opening of this portal subsequently allows water to access the distal heme pocket of sGC β1 H-NOX. Thus, we propose that the water molecule in the distal heme pocket that contributes to the hydrogen bond interaction with CO is likely responsible for the formation of the two CO-ligated form of H-NOX. DFT calculations on the vibrational frequencies of the isolated (ImH)FeP(CO) in the presence or absence of water further confirmed this hypothesis. Which is of great significant in recognising the signal transfer of sGC and its activation mechanism. However, deeply studies are needed to elucidate how large changes in CO-associated vibrations can be induced from a certain distortion of the proximal Im group that is part of the imidazole ring of H105.

Acknowledgements

We appreciate the help from Dr. Sachiko Yanagisawa and Dr. Miyuki Sakaguchi for the technical support of measuring Raman Spectroscopy.

T. O. acknowledges the support of "Strategic Young Researcher Overseas Visits Program for Accelerating Brain Circulation" and Grant-in-Aid for Scientific Research (No. 26104532) both by JSPS. T. O. is a visiting scientist at RIKEN.

This work was supported in part by the Global COE Program, "Picobiology: Life Science at the Atomic Level" at the Graduate School of Life Science, University of Hyogo, from MEXT, Japan (T.K.), Grants-in-Aid for Scientific Research from MEXT, Japan (T.K.) (No. 24350086).

This work was also supported in part by the National Natural Science Foundation of China (Z.L.) (No. 30570404 and 30670458).

Notes and references

1 Y. Zhao, C. Hoganson, G. T. Babcock and M. A. Marletta, *Biochemistry*, 1998, **37**, 12458-12464.
2 H. H. Schmidt, S. M. Lohmann and U. Walter, *Biochim. Biophys. Acta*, 1993, **1178**, 153-175.
3 B. Wedel, P. Humbert, C. Harteneck, J. Foerster, J. Malkewitz, E. Bohme, G. Schultz and D. Koesling, *Proc. Natl. Acad. Sci. U.S.A.*, 1994, **91**, 2592-2596.
4 M. Ibrahim, C. Xu and T. G. Spiro, *J. Am. Chem. Soc.*, 2006, **128**, 16834-16845.
5 T. L. Poulos, *Curr. Opin. Struct. Biol.*, 2006, **16**, 736-743.
6 G. Deinum, J. R. Stone, G. T. Babcock and M. A. Marletta, *Biochemistry*, 1996, **35**, 1540-1547.
7 B. Pal, Z. Li, T. Ohta, S. Takenaka, S. Tsuyama and T. Kitagawa, *J. Inorg. Biochem.*, 2004, **98**, 824-832.
8 D. S. Karow, D. Pan, J. H. Davis, S. Behrends, R. A. Mathies and M. A. Marletta, *Biochemistry*, 2005, **44**, 16266-16274.
9 Z. Li, B. Pal, S. Takenaka, S. Tsuyama and T. Kitagawa, *Biochemistry*, 2005, **44**, 939-946.
10 D. S. Karow, D. H. Pan, R. Tran, P. Pellicena, A. Presley, R. A. Mathies and M. A. Marletta, *Biochemistry*, 2004, **43**, 10203-10211.
11 X. L. Ma, N. Sayed, A. Beuve and F. van den Akker, *EMBO J.*, 2007, **26**, 578-588.
12 J. A. Winger, E. R. Derbyshire and M. A. Marletta, *J. Biol. Chem.*, 2007, **282**, 897-907.
13 J. P. Schelvis, Y. Zhao, M. A. Marletta and G. T. Babcock, *Biochemistry*, 1998, **37**, 16289-16297.
14 L. M. Iyer, V. Anantharaman and L. Aravind, *BMC Genomics*, 2003, **4**, 5.
15 P. Pellicena, D. S. Karow, E. M. Boon, M. A. Marletta and J. Kuriyan, *Proc. Natl. Acad. Sci. U.S.A.*, 2004, **101**, 12854-12859.
16 W. K. Erbil, M. S. Price, D. E. Wemmer and M. A. Marletta, *Proc. Natl. Acad. Sci. U.S.A.*, 2009, **106**, 19753-19760.
17 J. Pan, F. Zhong, H. Wang, Z. Huang and X. Tan, *Scientia. Sinica. Chimica.*, 2014, **44**, 572-585.
18 E. S. Underbakke, A. T. Iavarone, M. J. Chalmers, B. D. Pascal, S. Novick, P. R. Griffin and M. A. Marletta, *Structure*, 2014, **22**, 602-611.
19 A. L. Tsai, V. Berka, F. Martin, X. Ma, F. van den Akker, M. Fabian and J. S. Olson, *Biochemistry*, 2010, **49**, 6587-6599.
20 P. Nioche, V. Berka, J. Vipond, N. Minton, A. L. Tsai and C. S. Raman, *Science*, 2004, **306**, 1550-1553.
21 Y. Zhang, M. Lu, Y. Cheng and Z. Li, *J. Mol. Graphics Modell.*, 2010, **28**, 814-819.
22 M. B. Winter, M. A. Herzik, Jr., J. Kuriyan and M. A. Marletta, *Proc. Natl. Acad. Sci. U.S.A.*, 2011, **108**, E881-889.
23 Y. Zhang, L. Liu, L. Wu, S. Li, F. Li and Z. Li, *Proteins*, 2013, **81**, 1363-1376.
24 J. Ostmeyer, S. Chakrapani, A. C. Pan, E. Perozo and B. Roux, *Nature*, 2013, **501**, 121-124.
25 L. Liu, D. Wang, H. Xu, M. Mi and Z. Li, *J. Mol. Struct.*, 2015, **1089**, 102-106.
26 W. Li, Y. Zhang, H. Xu, L. Wu, Y. Cao, H. Zhao and Z. Li, *Biochim. Biophys. Acta*, 2013, **1834**, 2124-2132.
27 J. W. Denninger and M. A. Marletta, *Biochim. Biophys. Acta*, 1999, **1411**, 334-350.
28 A. Fiser and A. Sali, *Methods Enzymol.*, 2003, **374**, 461-491.
29 D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark and H. J. Berendsen, *J. Comput. Chem.*, 2005, **26**, 1701-1718.
30 A. D. MacKerell, D. Bashford, M. Bellott, R. L. Dunbrack, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorkiewicz-Kuczera, D. Yin and M. Karplus, *J. Phys. Chem. B*, 1998, **102**, 3586-3616.
31 W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J. Chem. Phys.*, 1983, **79**, 926-935.
32 H. J. C. Berendsen, J. P. M. Postma, W. F. V. Gunsteren, A. Dinola and J. R. Haak, *J. Chem. Phys.*, 1984, **81**, 3684-3690.
33 G. Bussi, D. Donadio and M. Parrinello, *J. Chem. Phys.*, 2007, **126**, 014101.
34 M. Parrinello and A. Rahman, *J. Appl. Phys.*, 1981, **52**, 7182-7190.
35 S. Nose, *Mol. Phys.*, 1984, **52**, 255-268.
36 U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee and L. G. Pedersen, *J. Chem. Phys.*, 1995, **103**, 8577-8593.
37 S. Miyamoto and P. A. Kollman, *J. Comput. Chem.*, 1992, **13**, 952–962.
38 B. Hess, H. Bekker, H. J. C. Berendsen and J. G. E. M. Fraaije, *J. Comput. Chem.*, 1997, **18**, 1463–1472.
39 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, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J.

Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, O. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, and D. J. Fox, Gaussian'09, (Revision D.01), Gaussian, Inc., Wallingford CT, 2013.

40 A. D. Becke, *J. Chem. Phys.*, 1993, **98**, 5648-5653.

41 C. Lee, W. Yang and R. G. Parr, *Phys. Rev. B*, 1988, **37**, 785-789.

42 L. M. Blomberg, M. R. Blomberg and P. E. Siegbahn, *J. Inorg. Biochem.*, 2005, **99**, 949-958.

43 J. R. Stone and M. A. Marletta, *Chem. Biol.*, 1998, **5**, 255-261.

44 J. R. Stone and M. A. Marletta, *Biochemistry*, 1994, **33**, 5636-5640.

45 T. G. Spiro, *Analytica. Chimica. Acta.*, 1989, **218**, 356.

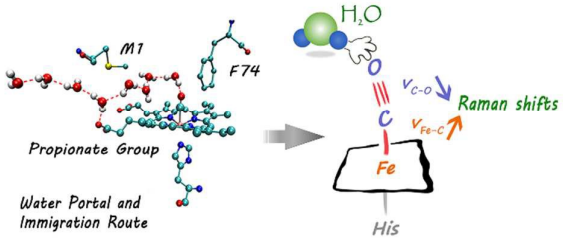
46 M. Tsubaki, R. B. Srivastava and N. T. Yu, *Biochemistry*, 1982, **21**, 1132-1140.

47 E. R. Derbyshire, M. B. Winter, M. Ibrahim, S. Deng, T. G. Spiro and M. A. Marletta, *Biochemistry*, 2011, **50**, 4281-4290.

48 R. Tran, E. E. Weinert, E. M. Boon, R. A. Mathies and M. A. Marletta, *Biochemistry*, 2011, **50**, 6519-6530.

49 K. M. Vogel, S. Hu, T. G. Spiro, E. A. Dierks, A. E. Yu and J. N. Burstyn, *J. Biol. Inorg. Chem.*, 1999, **4**, 804-813.

50 E. Martin, K. Czarnecki, V. Jayaraman, F. Murad and J. Kincaid, *J. Am. Chem. Soc.*, 2005, **127**, 4625-4631.



Water accessing into the heme pocket and alters the structures of CO-sGC(heme), exhibiting two different $\nu_{Fe\text{-}CO}$ stretching modes.