Supporting Information

1. Details of the QM/MM approach adopted in the present work.

All QM/MM optimizations were carried out with the COMQUM program suite.[i,ii] In the current version, it uses Turbomole 5.7[iii] for the QM part and AMBER 8[iv] (with the Amber 1999 force field[v]) for the MM part. In such a hybrid approach, the protein and solvent are divided into three subsystems: System 1 is treated at QM level, and contains the H-cluster atoms and relevant surrounding atoms (see below). System 2 consists of all residues with any atom within 12 Å of any atom in system 1 and it is optimized by a full MM minimization in each step of the QM/MM energy minimization. Finally, the remaining portion of the protein, together with the water molecules surrounding it are included in system 3, which is kept fixed at the crystallographic coordinates. Apart from system 1, which is represented by a wave function during the QM/MM geometry optimizations, each atom is represented by a partial point charge, taken from the Amber libraries.[v] All such MM charges are included in the Hamiltonian of the QM calculations, and thus the quantum chemical system is polarized by the atoms of system 2 and 3 in a self-consistent way. When the quantum and classical regions are connected by a chemical bond, the hydrogen link-atom approach is applied,[vi] i.e. the QM system is truncated with hydrogen atoms, the positions of which are linearly related to the corresponding carbon atom in the protein.

The total, QM/MM energy is calculated as

$$E_{QM/MM} = E_{QM} + E_{MM123} - E_{MM1} \quad (1)$$

Here, E_{QM} is the QM energy of the quantum system truncated by hydrogen atoms, including the interaction between system 1 and the surrounding point charges. E_{MM1} is the MM energy of the quantum system, still truncated by hydrogen atoms, but without any electrostatic interactions. Finally, E_{MM123} is the classical energy of all the atoms with carbon atoms at the junctions, and with the charges of the QM region zeroed, in order to avoid double counting of the electrostatic interactions. Such an approach, which is similar to the one used in the Oniom method,[vii] should lead to the cancellation of errors caused by the truncation of the quantum system.

Energy differences between $E_{QM/MM}$ and E_{QM} values are reported in Table S1 (see below) for all the model complexes considered in the present paper, together with energy differences obtained from single-point calculations carried out in a vacuum on the isolated quantum system of QM/MM optimized enzyme models. The latter energy differences, which will be termed ΔE_{QM}^{vacuum} in Table S1, are the only QM/MM energy values that are discussed in the main text of the present work.

Geometry optimizations were carried out in three steps. First, systems 2 and 3 were frozen and only the quantum system was optimized (geometries obtained by this optimization will be referred to as Protein_Fixed). Second, both systems 1 and 2 were allowed to relax. In the MM optimization of system 2, the charges of the quantum atoms were updated in each iteration of the QM/MM optimization.[ii] This optimization was performed with the looser convergence criterion of 10^{-4} a.u. for the change in total QM/MM energy and 10^{-2} a.u. for the maximum norm of the Cartesian gradient (0.26 kJ mol^{-1} and 50 kJ $(mol\ Å)^{-1}$). Then, system 2 was frozen again, and the geometry optimization was continued with default convergence criteria (10^{-6} and 10^{-3} a.u.). If not otherwise stated, the discussion is based on the results obtained with system 2 free to relax, because these structures are expected to be more realistic.

Deleted: s
Deleted: s
Deleted: s
Deleted: s
Deleted: v
Deleted: 24
Deleted: applied, i
Deleted: ·
Deleted: s
Deleted: s
Deleted: s
Deleted: s
Deleted: ii
Deleted: 21
Deleted: s
Deleted: s

1.1. The protein. All QM/MM calculations were based on the 1.6-Å resolution structure of the [FeFe]-hydrogenase from *Desulfovibrio desulfuricans* (PDB code 1HFE).[viii] This is a hetero-dimer composed of a large subunit that harbours the H-cluster and the two accessory Fe_4S_4 assemblies, and a small subunit. This crystal structure was selected because it has the highest resolution among the published [FeFe]-hydrogenases structures. Hydrogen atoms were added to the crystal structure, and the protein was solvated in a sphere of water molecules with a radius of 48 Å with the Amber routine leap. In order to optimize the positions of hydrogen atoms and solvent water molecules, a 90 ps simulated-annealing molecular dynamics calculation was carried out, followed by 10000 steps of conjugate gradient energy minimization. The protonation state of histidine side chains was chosen considering solvent exposure and the hydrogen-bond network around the residues; this means that, for each histidine side chain, all possible hydrogen bond donors and acceptors in close proximity of the imidazole nitrogen atoms were identified, and a congruent disposition of the proton(s) on the ring was established. As a result of such a procedure, we assigned protonation of $N^{\delta 1}$ for residues S89 (the letter S indicates the small subunit; residue numbers without a S refer to the large subunit) and 75; protonation of $N^{\varepsilon 2}$ for residues 351 and 371; and protonation on both the nitrogen atoms of the imidazole ring for residues S82, S85, S91, 14, 26, 58, 62, 141 and 196. All lysine and arginine residues were considered in their positively charged state, while aspartate and glutamate side chains were always included in the anionic form. Finally, the iron-bound cysteine residues (i.e. amino acids 36, 38, 41, 45, 66, 69, 72, 76, 179, 234, 378 and 382) were assumed to be deprotonated. All the ligands found in the PDB file were included in the QM/MM model, except a water molecule bridging Fe_d and Fe_p, which was replaced by a carbonyl group, following a more recent correction[ix] of the crystal structure.[viii]

The [FeFe]-hydrogenase from *D. desulfuricans* contains two accessory iron-sulphur clusters in addition to the H-cluster. For these Fe_4S_4 sites, we used Merz–Kollman electrostatic potential (ESP) charges,[x] taken from QM calculations at the B3LYP/6-31G* level for truncated models of each site. One of the iron-sulphur clusters belongs to system 2 (i.e., the MM region that is free to relax), and thus a set of MM parameters had to be also defined in this case. To this end, a QM frequency calculation was run, and force constants of all relevant bonds, angles, and dihedrals were extracted from the Hessian matrix, using the approach suggested by Seminario.[xi] As for the charge of the accessory Fe_4S_4 clusters, we adopted the $[Fe_4S_4]^{2+}$ state.

1.2. Quantum chemical calculations. All QM calculations described in the present paper were carried out within the density functional theory (DFT) framework, using the BP86 functional[xii] and an all-electron split-valence basis with polarization functions on all atoms (SVP).[xiii] Moreover, we applied the resolution-of-identity (RI) technique,[xiv,xv] which sped up the calculations by a factor of ~10. However, after all the QM/MM geometry optimizations carried out at BP86/SVP-RI level, a single point calculation at B3LYP/TZVP level was performed. In fact, the use of the hybrid functional B3LYP in conjunction with the large triple-ζ basis TZVP is expected to give more reliable QM energy values.

A theoretical investigation on the entire H-cluster is a challenge because the Fe_4S_4 cluster is composed of two Fe_2S_2 layers of high-spin Fe ions, coupled antiferromagnetically to give an overall low-spin ground state. The ground-state wave function of such spin-coupled systems corresponds to linear combinations of multiple determinants that cannot be treated within the single-determinant DFT approach. However, in the framework of the unrestricted formalism, these interactions can be modelled with the broken symmetry (BS) approach introduced by Noodleman et al.[xvi] The BS approach consists in the localisation of opposite spins of the mono-determinant wave function in different parts of the molecule.

As noted by Brunold,[xvii] the Fe atoms of Fe_4S_4 are not equivalent in the H-cluster, and therefore six different BS configurations can be generated for each species. However, we only considered one of

Deleted: Fe-only hydrogenase

Deleted: Fe-only hydrogenase

Deleted: s

Deleted: s

Deleted: Fe-only hydrogenase

Deleted: s

Deleted: the

the six possible BS wave function variants with the lowest energy. In fact, various QM/MM and QM calculations on the reduced form of the enzyme (data not shown) have indicated that the structural differences among H-cluster models characterized by different BS configurations are very small (the differences between corresponding bond lengths in the Fe_2S_2 cluster were always below 0.02 Å), and the QM/MM energies never differed by more than 10 kJ mol^{-1}. We also validated the use of the BP86/SVP level of theory by carrying out QM/MM geometry optimizations on models of the active-ready form of the enzyme, in the context of BP86/TZVP level of theory**.** The use of the TZVP basis (a triple-ζ polarized basis set) gave rise to structures very similar to those obtained at BP86/SVP level,[xviii] thus supporting the computational approach used in the rest of our work.

1.3. Composition of the model systems. System 1 - i.e. the QM system of the various [FeFe]-hydrogenase models here considered – always included the iron and sulphide ions of the Fe_6S_6 H-cluster, a PDT or a DTMA ligand bridging Fe_d and Fe_p, three CO groups, two CN^- ligands, and four CH_3S^- groups that represent the cysteine residues connecting the Fe_6S_6 cluster to the rest of the enzyme large subunit (Cys-179, Cys-234, Cys-378, Cys-382). In addition, the sidechain of Cys-178 was also included in the QM region, in the form of a CH_3SH group. This residue is close to the bidentate ligand of the binuclear subcluster, and it might act as the terminal element of the proton channel that supplies protons to the H-cluster during the dihydrogen-evolving route.

The QM systems of the various QM/MM models considered in this paper differ in terms of ligands nature and/or disposition in Fe_d coordination sphere (see main text). In particular, the coordination site *trans* to μ-CO on Fe_d can be either vacant or occupied by a CN^- group or by an exogenous CO ligand. The Fe_4S_4 subcluster was always modelled in the +2 state (in the antiferromagnetically coupled low-spin state). This gives a total of 57-59 atoms in system 1. Both active-ready and CO-inhibited enzyme models are neutral, apart from the case of the active-ready enzyme models that include a protonated DTMA residue (in this case, the overall charge of the enzyme is +1). System 2 consisted of 217 amino acids and 41 water molecules, whereas system 3 included the rest of protein and water molecules (267 residues and 7044 water molecules).

2. Details on the QTCP approach

All the free energy perturbation (FEP) calculations presented in this work were carried out in the context of the QM/MM thermodynamic cycle perturbation (QTCP) approach.[xix] In particular, the QTCP-U implementation of the QTCP method was applied,[xix] which means that the thermodynamic cycle reported in Figure S1 was employed to calculate free energy changes.

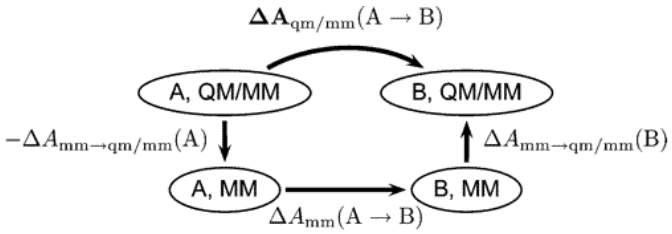


Figure S1. The thermodynamic cycle adopted in the QTCP calculations

The thermodynamic cycle depicted in Figure S1 makes it possible to compute high-level QM/MM free energy changes between two states A and B based on classical sampling. In this approach, the free energy change between A and B described by QM/MM is calculated as the sum of three terms: the negative free energy change between A described by MM and by QM/MM ($-\Delta A_{mm\rightarrow qm/mm}(A)$), the free energy change between A and B with both described by the MM potential ($\Delta A_{mm}(A \rightarrow B)$), and the free energy change between B described by the MM potential and by QM/MM ($\Delta A_{mm\rightarrow qm/mm}(B)$).

Therefore

$$\Delta A_{qm/mm}(A \rightarrow B) = -\Delta A_{mm\rightarrow qm/mm}(A) + \Delta A_{mm}(A \rightarrow B) + \Delta A_{mm\rightarrow qm/mm}(B) \qquad (2)$$

In our case, A and B are always two different forms of the active–ready or CO-inhibited [FeFe]-hydrogenase enzyme. In particular, the $\Delta A_{qm/mm}$ values have been evaluated for the following structural rearrangements: $\mathbf{I_{PDT}}^{\mathbf{CNtrans}} \leftrightarrows \mathbf{I_{PDT}}^{\mathbf{COtrans}}$; $\mathbf{I_{DTMA}}^{\mathbf{CNtrans}} \leftrightarrows \mathbf{I_{DTMA}}^{\mathbf{COtrans}}$; $\mathbf{A1_{PDT}} \leftrightarrows \mathbf{A2_{PDT}}$; $\mathbf{A1_{DTMA}} \leftrightarrows \mathbf{A2_{DTMA}}$ and $\mathbf{A1_{DTMAH+}} \leftrightarrows \mathbf{A2_{DTMAH+}}$ (the nomenclature here used for the various enzyme forms is the same as in the main text of the present work).

Each of the three terms in Eqn. 2 can be calculated by means of free energy perturbation (FEP).[xx] In the FEP approach, a free energy change, $\Delta A(0\rightarrow 1)$ is calculated as

$$\exp(-\Delta A(0\rightarrow 1)/k_B T) = \langle \exp(-(E_1 - E_0)/k_B T) \rangle_0 \qquad (3)$$

Using FEP, the middle term on the right side of Eqn. 2 can be written as

$$\Delta A_{mm}(A \rightarrow B) = -k_b T \ln \langle \exp(-(E_{MM123}(B) - E_{MM123}(A))/k_B T) \rangle_{MM123,A} \qquad (4)$$

and the first and last terms can be written as

$$\Delta A_{mm\rightarrow qm/mm}(X) = -k_b T \ln \langle \exp(-(E_{QM/MM}(X) - E_{MM123}(X))/k_B T) \rangle_{MM123,X} \qquad (5)$$

The $E_{QM/MM}$ and E_{MM123} terms in Eqns. 4 and 5 have been defined above (Eqn. 1).

Thus, the QTCP approach demands an MD simulation per state involved; a complete definition of the various energy terms reported in Eqns. 2-5 has been described in previous works.[xix] Here, it is only important to underline that the QTCP approach allows for the computation of QM/MM free energy differences by sampling only the phase space at the MM level, and that the calculation of $\Delta A_{qm/mm}(A \rightarrow B)$ implies a series of single-point DFT computations. Such single-point calculations are carried out on the QM system kept fixed at the geometry obtained by means of a QM/MM Protein_fixed optimization, and including in the QM representation the point charges of the MM region; the positions of the MM charges in the QM representation correspond to the atomic coordinates obtained from selected snapshots of the MD simulations (see section 2.2).

Deleted: Fe-only hydrogenase

Deleted: [xix]

Deleted: [xviii]

2.1. Molecular dynamics simulations in the context of QTCP calculations

For each of the enzyme forms of interest, a molecular dynamics (MD) calculation was run with the program AMBER 8.[xxi] In these simulations, the QM system was represented by using Merz–Kollman ESP point charges. The QM region was kept fixed in space during the simulations. Atoms of the cofactor and protein that are not part of the QM region were described by the Amber99 force field,[xxii] and the solvent was described explicitly using the TIP3P model.[xxiii] Periodic boundary conditions were employed using an octahedral unit cell and the particle-mesh Ewald (PME) method.[xxiv] To this end, the sphere of water molecules added prior to the Protein_fixed QM/MM calculation was removed from each model, while crystallization waters were left in the system. After that, solvent molecules reaching at least 6 Å outside the system composed by the protein and the crystallization waters were added.[xxv] Then, 100 steps of steepest decent MM minimization with all heavy atoms restrained to the original QM/MM structure were performed, followed by 20 ps of MD simulation at constant volume and 20 ps MD simulation at constant pressure (1 atm), still with heavy atoms restrained. Subsequently, a 50 ps MD equilibration was carried out with only the quantum system restrained to the original QM/MM structure, at constant pressure (1 atm), followed by a 1000 ps MD equilibration at constant volume with the quantum system fixed.
All the above reported MD calculations preceded the data-collection MD step, which was always carried out at constant volume for 400 ps.

Deleted: s

2.2. Single-point QM calculations in the context of the QTCP approach

For each of the data-collection MD simulation carried out in the present work, 80 configurations separated by 5 ps were stored. The atomic coordinates from each of these MD snapshots were then used for representing the protein in 80 distinct single-point QM calculations; the QM energy values thus obtained were then used to calculate the MM→QM/MM free energy change for the isomerization reactions of interest. These calculations were run in the same way as the QM/MM geometry optimizations (BP86/SVP).

Deleted: s

3. Optimized geometries of the various QM/MM models of the H-cluster considered in the main text.

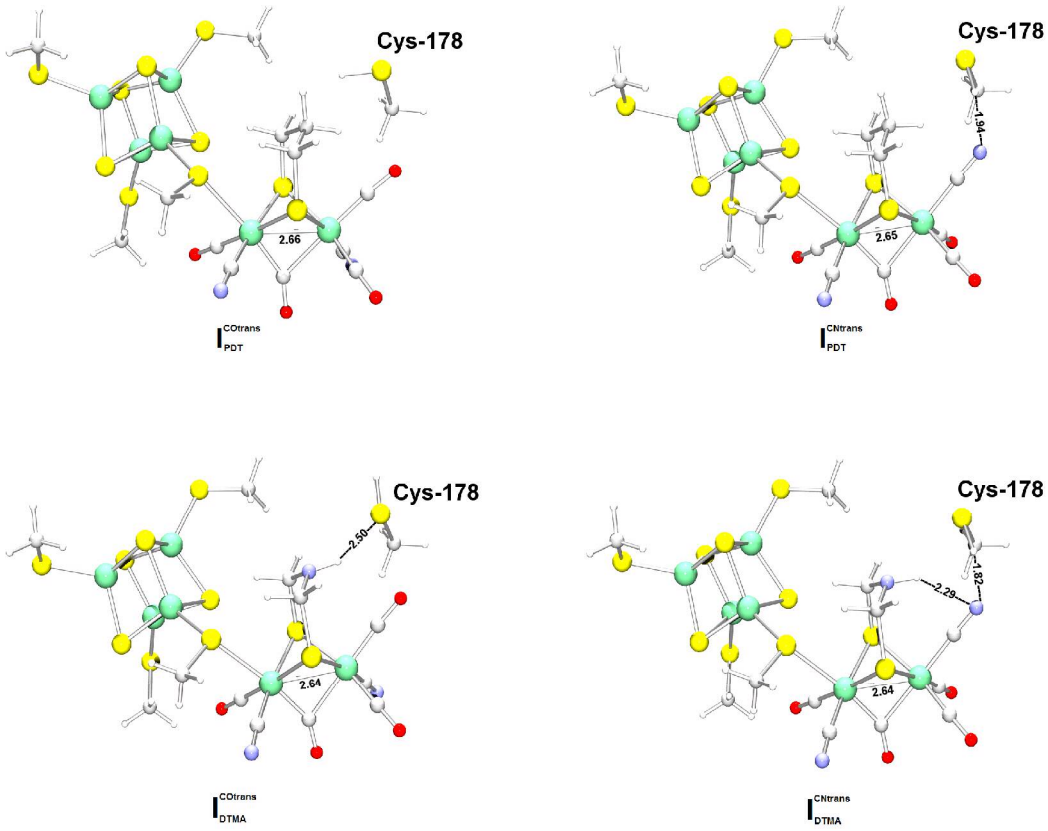


Figure S2. QM/MM optimized structures of the various CO-inhibited forms of the H-cluster investigated in the present paper ($I_{PDT}^{COtrans}$, $I_{PDT}^{CNtrans}$, $I_{DTMA}^{COtrans}$, $I_{DTMA}^{CNtrans}$). Selected distances are given in Å.

Deleted: s

Deleted:

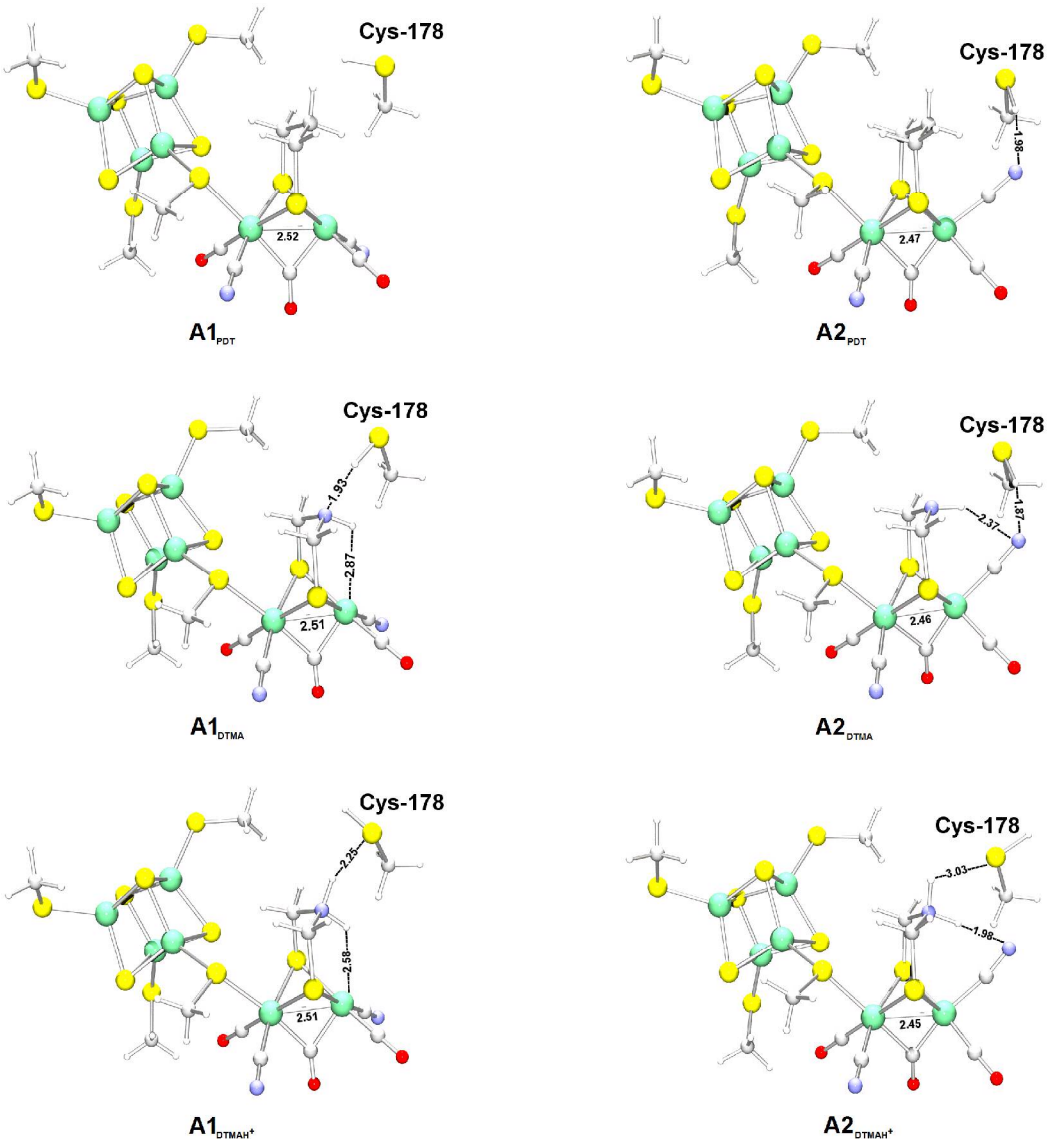


Figure S3. QM/MM optimized structures of the various active-ready forms of the H-cluster investigated in the present paper (**$A1_{PDT}$**, **$A2_{PDT}$**, **$A1_{DTMA}$**, **$A2_{DTMA}$**, **$A1_{DTMAH+}$**, **$A2_{DTMAH+}$**). Selected distances are given in Å.

4. QM/MM and QM energy values for the various QM/MM models considered in the main text

In table S1, the energy differences between the various models considered in the present work are reported, in terms of $E_{QM/MM}$, E_{QM} and E_{QM}^{vacuum} (see Eqn. 1 and section 1 of the Supporting Materials), and in terms of QTCP energy ($\Delta E_{QTCP} = \Delta A_{qm/mm}$, see Eqn. 2).

Table S1. Energy differences between the various models at QM, QM/MM and QTCP level. All values are given in kJ mol^{-1}.[a]

	$\Delta E_{QM/MM}$	ΔE_{QM}	ΔE_{QM}^{vacuum}	ΔE_{QTCP}
$I_{PDT}^{CNtrans} \rightarrow I_{PDT}^{COtrans}$	-53	-52	30	-14 ± 1.5
$I_{DTMA}^{CNtrans} \rightarrow I_{DTMA}^{COtrans}$	-72	-55	43	-6 ± 0.5
$A1_{PDT} \rightarrow A2_{PDT}$	107	100	-2	117 ± 12
$A1_{DTMA} \rightarrow A2_{DTMA}$	97	98	16	73 ± 6.5
$A1_{DTMAH+} \rightarrow A2_{DTMAH+}$	65	65	-20	76 ± 13

[a] QM/MM calculations in which Lys-237 was part of the QM system were also carried out, in order to assess how the results presented in Table S1 depend on the way the QM and the MM regions interface. Notably the calculated QM/MM energy differences remained substantially unchanged upon such an extension of the QM region, meaning that variations of no more than 10 kJ mol^{-1} were observed.

The hysteresis values associated with QTCP energy differences were obtained by calculating ΔE_{QTCP} for the QTCP thermodynamic cycle (Figure S1) in both directions (clockwise and anticlockwise). The ΔE_{QTCP} values reported in Table S1 are the average of the clockwise and anticlockwise QTCP energy differences, plus an uncertainty term that depends on the difference between the energy values used for average calculations. The accuracy of QTCP energy differences depends on hysteresis, which varies from time to time, and on the error associated with DFT calculations, which is in the order of 12 kJ mol^{-1}. However, it is relevant to underline that the uncertainty values reported in Table S1 and in the main text of the present work refer only to the hysteresis associated with molecular dynamics calculations.

5. Sequence alignment among several [FeFe]-hydrogenase homologues

A homology search was carried out using DdH sequence as a query for a BLAST[xxvi] run. Then, the first 20 hits were collected in order to obtain a multiple sequence alignment with ClustalW,[xxvii] using default parameters. The ClustalW results are below reported, using the one-letter aminoacid code; only the regions of sequence homologous to the DdH sequence segment 212-253 are shown, and the Lys237 residue is marked in red.

```
DdH (residues 212-253)              KTYGAERMKYDPKQVYTVSIMPCIAKKYEGLRPELK--------SSGMRD
sp|P07598|PHFL_DESVH                KTYGAERMKYDPKQVYTVSIMPCIAKKYEGLRPELK--------SSGMRD
tr|A1VD88|A1VD88_DESVV              KTYGAERMKYDPKQVYTVSIMPCIAKKYEGLRPELK--------SSGMRD
sp|P13629|PHFL_DESVO                KTYGADRMKYDRAKVYTVSIMPCTAKKYEGMRPQLW--------DSGHKD
tr|Q317L4|Q317L4_DESDG              KTYGPDVMKYDRSKVYTVSIMPCTAKKYEGMRADLW--------SSGYKD
tr|Q9AM36|Q9AM36_DESDE              KTYGPDVMKYDRSKVYTVSIMPCTAKKYEGMRADLW--------SSGYKD
tr|O08311|O08311_DESFR              KTYGAKELGYEPKQIYTVSIMPCTAKKFEGMRPEMD--------ASGFRD
tr|A0LGJ3|A0LGJ3_SYNFM              KTYGAQETKTPRDRMYTVSIMPCIAKKFEGLRPEMA--------ASGCRD
tr|Q30Z18|Q30Z18_DESDG              KTYGAQEAGVPAKKMYTVSIMPCIAKKFEGMRPEMN--------ASGYRD
tr|Q0AVN2|Q0AVN2_SYNWW              KTYLAEKLNVDPANMFSVSVMPCTAKKYECERPEFIAS--------GHQD
tr|Q2LSB7|Q2LSB7_SYNAS              KTYYAQVSGIDPANIFSVSIMPCTAKKYEAQRPEMK--------SSGYQD
tr|Q67J76|Q67J76_SYMTH              KTYGAKVDNVDPAKVYSVSVMPCTCKDFESGRPEMK--------ASGYRD
tr|Q24N91|Q24N91_DESHY              KTYFAEKNHVEPQKIFSVAIMPCTAKKFECQRPEMISAQTYWQDEQVSPD
tr|Q18T66|Q18T66_DESHD              KTYFAEKNHVEPQKIFSVAIMPCTAKKFECQRPEMISAQTYWQDEQVSPD
tr|Q1EVZ7|Q1EVZ7_9CLOT              KSYFAEKMAMDPKKIVTVSIMPCTAKKFESARGELS--------NDGLQD
tr|Q3CJE2|Q3CJE2_THEET              KSYYAEKKGLNPEDIYIVSIMPCTAKKLEIERPEMQ--------HNGIKD
tr|Q3ZA52|Q3ZA52_DEHE1              KTYYAEKSGIDPKDIINVSVMPCTAKKFECQRPEMN--------DSGFKD
tr|Q2RHS0|Q2RHS0_MOOTA              KTYYARKAGVDPARMVVVSIMPCTAKKFECQRPEMR--------DSGYQD
tr|A1HP37|A1HP37_9FIRM              KTYYAERVGIDPKDIVSVSIMPCTAKKAEAVRPEMR--------ASGYQD
tr|Q3ZWM9|Q3ZWM9_DEHSC              KTYYAEKAGIDPKTIVNVSVMPCTAKKFECQRPEMN--------DSGFKD
tr|Q2DWB9|Q2DWB9_9CHLR              KTYYAEKAGIDPKTIVNVSVMPCTAKKFECQRPEMN--------DSGFKD
                                    *:* .         :  *::*** .*. *  * ::              *
```

6. DFT analysis of models of the CO-inhibited binuclear subsite; computation of IR bands.

In a recent work, Zilberman et al. have developed an accurate approach for the prediction of IR bands of [FeFe]-hydrogenases synthetic models, using DFT with the PBE functional, and plane waves.[xxviii] Using the above reported computational scheme, those authors proposed a hypothesis for resolving the CO/CN$^-$ ligand arrangement in CO-inhibited [FeFe]-hydrogenases.[xxix] In fact, by computing IR bands of $\mathbf{I}^{\mathbf{CNtrans}}$ and $\mathbf{I}^{\mathbf{COtrans}}$ (see Figure S4; $\mathbf{I}^{\mathbf{COtrans}}$ has CO/CN$^-$ ligands disposed in the way that had been widely accepted as definitive before Zilberman's papers), and IR bands shifting consequent to inactivation with ^{13}CO instead of ^{12}CO, they found that the CO/CN$^-$ ligand assignment corresponding $\mathbf{I}^{\mathbf{COtrans}}$ is inconsistent with the IR data available for the CO-inhibited [FeFe]-hydrogenase from *Clostridium pasteurianum*. On the basis of their computational results, they propose that the actual disposition of CO/CN$^-$ ligands is the one corresponding to $\mathbf{I}^{\mathbf{CNtrans}}$ (see Figure S4).

Deleted: Fe-only hydrogenases

Deleted: To do

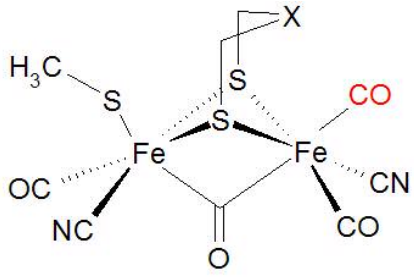
Deleted: Fe-hydrogenases

Deleted: Fe-hydrogenases

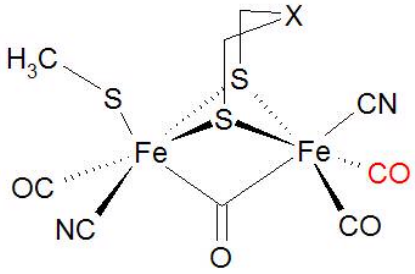
Deleted: 1

Deleted: Fe-hydrogenase

Deleted: 2



I^{COtrans}



I^{CNtrans}

Figure S4. Schematic structures of the models investigated by Zilberman et al.;[xxix] the CO group coloured in red was included either in the ^{12}CO or in the ^{13}CO form in the various models considered in the present work and in Zilberman et al.'s work.

In view of this, we have made a DFT analysis of the same CO-inhibited binuclear subsite models considered by Zilberman et al.,[xxix] but using both a pure functional and a hybrid one for the two species of interest. Full vacuum geometry optimizations and frequency calculations were carried out using Turbomole 5.7, at TZVP/BP86 and TZVP/B3LYP level. The results of such an analysis are summarized in Table S2; in this table, R^2 values from linear regression fit of the computed and experimental[xxx] IR bands are reported; both DTMA and PDT containing models were considered; models containing a CO or CN^- ligand *trans* to the bridging CO have the "**COtrans**" or "**CNtrans**" tag in their names, respectively. In Table S2, data regarding the all-^{12}CO models are reported in the first raw, while the second raw refers to models in which one of the CO ligands (i.e. the CO groups coloured in red in Figure S4) is a ^{13}CO.

Table S2. R^2 values from linear regression fitting of the computed and experimental IR bands for the various models of the CO-inibited binuclear subsite.[a]

Models:	**B3LYP-$I^{COtrans}$ - DTMA**	**B3LYP-$I^{CNtrans}$ - DTMA**	**B3LYP-$I^{COtrans}$ - PDT**	**B3LYP-$I^{CNtrans}$ - PDT**	**BP86-$I^{COtrans}$ - DTMA**	**BP86-$I^{CNtrans}$ - DTMA**	**BP86-$I^{COtrans}$ - PDT**	**BP86-$I^{CNtrans}$ - PDT**
R^2 (All-^{12}CO)	0.861	0.860	0.847	0.788	0.759	0.884	0.752	0.937
R^2 (One ^{13}CO)	0.956	0.907	0.948	0.902	0.833	0.937	0.834	0.941

[a] The nomenclature here used is different from that used in the main text for the QM/MM models of the entire CO-inhibited enzyme, in order to underline that the data reported in this Table regards DFT models of the *isolated* binuclear subsite.

Computed CO and CN^- vibration intensities were not considered in the present study because they have shown to be less reliable than vibration energies in previous studies carried out in our laboratories on DFT models of the H-cluster binuclear subsite.

Taken as a whole, the R^2 values reported in Table S2 show that BP86/TZVP data tend to be more consistent with the attribution of $\mathbf{I}^{CNtrans}$ geometry to the CO-inhibited enzyme, while B3LYP/TZVP results support the assignment of $\mathbf{I}^{COtrans}$ geometry for the same enzyme form. Such a situation is peculiar, since BP86/TZVP and B3LYP/TZVP calculations usually give very similar results on models of [FeFe]-hydrogenase active site, not only in terms of geometries[xxxi] but also in terms of linear fitting between computational and experimental IR frequencies; for example, when the

synthetic model $[Fe_2(S_2C_3H_6)(CO)_5(HCO)]^-$ is considered (see Figure S5), R^2 values from linear regression fitting of the computed and experimental CO bands are larger then 0.98 both at BP86/TZVP and B3LYP/TZVP level.

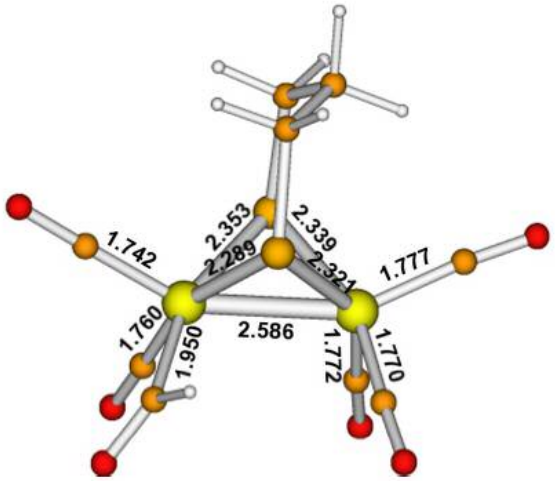


Figure S5. Ball and stick representation of complex $[Fe_2(S_2C_3H_6)(CO)_5(HCO)]^-$ optimized at BP86/TZVP level

We conclude that particular caution has to be adopted when DFT data on simple binuclear clusters are used to make structural assignments on the CO-inhibited [FeFe]-hydrogenase active site.

7. Results of the calculation of EPR parameters

Calculation of EPR parameters was carried on the various QM/MM optimized geometries of the H-cluster, using the program developed by Malkin and collaborators.[xxxii] In order to make the EPR calculation affordable in terms of computational cost, the cubane residue was always replaced by a hydrogen atom, meaning that a $SHCH_3$ group took the place of the Fe_d-bound $(SCH_3)_4[Fe_4S_4]$ moiety in all models. Then, a single-point B3LYP or B3PW91 calculation was carried out, using the TZVP basis set. The resulting wave functions were used for EPR parameters calculation; results of these computations are summarised in Table S3.

Table S3. Computed EPR parameters for DTMA- and PDT-containing models of the QM/MM optimized binuclear subsite.

	***g*1**	***g*2**	***g*3**
Experimental values[xxxiii]	2.010	2.010	2.060
$I_{PDT}^{CNtrans}$-B3LYP	2.009	2.018	2.033
$I_{PDT}^{COtrans}$-B3LYP	2.009	2.021	2.034
$I_{PDT}^{CNtrans}$-B3PW91	2.009	2.018	2.034
$I_{PDT}^{COtrans}$-B3PW91	2.008	2.022	2.035
$I_{DTMA}^{CNtrans}$-B3LYP	2.010	2.018	2.033
$I_{DTMA}^{COtrans}$-B3LYP	2.009	2.021	2.033
$I_{DTMA}^{CNtrans}$-B3PW91	2.010	2.018	2.034
$I_{DTMA}^{COtrans}$-B3PW91	2.009	2.022	2.035

The agreement between computational and experimental EPR data is always good for the *g* values ***g*1** and ***g*2**, while the 2.060 signal is reproduced with lower accuracy. However, in the context of the present work it is important to underline that the values of ***g*** calculated for the two possible geometries of **H_{ox}-CO** are very similar to each other both in PDT- and in DTMA-containing models; this suggests that resolving the CO/CN^- ligand disposition on Fe_d in **H_{ox}-CO** on the basis of EPR data is a non-trivial task.

References and notes:

[i] Ryde, U. *J. Comput.-Aided Mol. Des.* **1996**, *10*, 153.

[ii] Ryde, U.; Olsson, M. H. M. *Int. J. Quantum Chem.* **2001**, *81*, 335.

[iii] Ahlrichs, R.; Bar, M.; Haser, M.; Horn, H.; Kolmel, C. *Chem. Phys. Lett.* **1989**, *162*, 165.

[iv] Case, D.A. et al., *Amber 8*. **2004**, University of California: San Francisco, CA.

[v] Cornell, W. D.; Cieplak, P. I.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kolman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179.

[vi] Reuter, N.; Dejaegere, A.; Maigret, B.; Karplus, M. *J. Phys. Chem.* **2000**, *104*, 1720.

[vii] Svensson, M.; Humbel, S.; Froese, R. D. J.; Matsubara, T.; Sieber, S.; Morokuma, K. *J. Phys. Chem.* **1996**, *100*, 19357.

[viii] Nicolet, Y.; Piras, C.; Legrand, P.; Hatchikian, E. C.; Fontecilla-Camps, J. C. *Structure* **1999**, *7*, 13-23.

[ix] Nicolet, Y.; de Lacey, A. L.; Vernede, X.; Fernandez, V. M.; Hatchikian, E. C.; Fontecilla-Camps, J. C. *J. Am. Chem. Soc*. **2001**, *123*, 1596.

[x] Besler, B. H.; Merz, K. M.; Kollman, P. A. J. Cmput. Chem. **1990**, *11*, 431.

[xi] Seminario, J. M. *Intern. J. Quant. Chem.*, Quant. Chem. Symp. **1996**, *30*, 59-65. Nilsson, K.; Lecerof, D.; Sigfridsson, E.; Ryde, U. *Acta Crystallogr. D* **2003**, *59*, 274.

[xii] Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098-3100. Perdew, J. *Phys. Rev. B* **1986**, *33*, 8822.

[xiii] Schaefer, A.; Horn, H.; Ahlrichs, R. *J. Chem. Phys.* **1992**, *97*, 2571.

[xiv] Eichkorn, K.; Treutler, O.; Öhm, H.; Haser, M.; Ahlrichs, R. *Chem. Phys. Lett.* **1995**, *240*, 283.

[xv] Eichkorn, K.; Weigend, F.; Treutler, O.; Ahlrichs, R. *Theor. Chim. Acta* **1997**, *97*, 119.

[xvi] Noodleman, L.; Norman jr. J. G. *J. Chem. Phys.* **1979**, *70*, 4903-4906; Noodleman, L. *J. Chem. Phys.* **1981**, *74*, 5737-5743.

[xvii] Flieder, A. T.; Brunold, T. C. *Inorg. Chem.* **2005**, *44*, 9322.

[xviii] Greco C.; Bruschi M.; De Gioia L.; Ryde U. *Inorg. Chem.*, in press.

[xix] (a) Rod, T. H.; Ryde, U. *Phys. Rev. Lett.* **2005**, *94,* 138302. (b) Rod, T. H.; Ryde, U. *J. Chem. Theory Comput.* **2005,** *1,* 1240.

[xx] (a) Zwanzig, R. W. *J. Chem. Phys.* **1954**, *22,* 1420; (b) Beveridge, D. L.; DiCapua, F. M. *Annu. Rev. Biophys.* Biophys. Chem. **1989**, 18, 431.

[xxi] Case, D.A. et al., *Amber 8*. **2004**, University of California: San Francisco, CA.

[xxii] Cornell, W. D.; Cieplak, P. I.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kolman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179.

[xxiii] Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79,* 926.

[xxiv] Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. *J. Chem. Phys.* **1995**, *103,* 8577

[xxv] The use periodic boundary conditions and sizeable solvation shell reaching at least 9 Å outside the protein (9 Å considering both crystallization waters and the other water molecules added by means of the "leap" tool of Amber) is a standard procedure in MD simulations of proteins. However, the extension of the solvent box all around the protein could, in principle, influence the energetics of a QM/MM calculation in a significant way (a Protein_fixed QM/MM calculation always precedes a QTCP calculation, and the QTCP energy difference itself is always computed by means of series of single-point DFT computations that include the electrostatic field exerted by the protein on the QM region, see above). This point has been addressed in a recent study (Ryde et al., to be submitted): in this study, two diffrent approaches for carrying out QM/MM calculations on proteins are compared: in the first approach, the protein is truncated after 20-30 Å, and one or two layers of water molecules are added outside this radius; in the second approach, periodic boundary conditions and sizeable solvation shell of ~9 Å are used instead. The results obtained show that there is no difference between the two QM/MM approaches.

[xxvi] Altschul, S. F.; Madden, T. L.; Schäffer, A. A.; Zhang, J.; Zhang, Z.; Miller, W.; Lipman, D.J. *Nucleic Acids Res.* **1997**, *25*, 3389.

[xxvii] Higgins, D.; Thompson, J.; Gibson, T.; Thompson, J.D.; Higgins, D. G.; Gibson, T. J. *Nucleic Acids Res.* **1994**, *22*, 4673.

[xxviii] Zilberman, S.; Stiefel, E.; Cohen, M.; Car, R. *J. Phys. Chem. B* **2006**, *110*, 7049.

[xxix] Zilberman, S.; Stiefel, E.; Cohen, M.; Car, R. *Inorg. Chem.* **2006**, *45*, 5715.

[xxx] Chen, Z. J.; Lemon, B. J.; Huang, S.; Swartz, D. J.; Peters, J. W.; Bagley, K. A. *Biochemistry* **2002**, *41*, 2036

[xxxi] Greco, C.; Zampella, G.; Bertini, L.; Bruschi, M.; Fantucci, P.; De Gioia, L. *Inorg. Chem.* **2007**, *46*, 108; Zampella, G.; Bruschi, M.; Fantucci, .; Razavet, M.; Pickett C. J. and De Gioia, L. *Chemistry Eur. J.*, **2005**, *11*, 509.

Deleted: ,

Deleted: ,

Deleted: ,

Deleted: ,

Deleted: ,

Deleted: :

[xxxii] Malkin, V. G.; Malkina, O. L.; Reviakine, L.; Arbouznikov, A. V.; Kaupp, M.; Schimmelpfennig, B.; Malkin, I.; Helgaker, T.; Ruud, K. RESPECT program, Version 1.1

[xxxiii] Patil, D. S.; He, S. H.; DerVartanian, D. V.; Le Gall, J.; Huynh, B. H.; Peck, H. D. Jr. *FEBS Lett.* **1988**, *228*, 85.