Electronic Supplementary Material (ESI) for Physical Chemistry Chemical Physics

Supplementary Information

Insights into the structure of the active site of the O_2-tolerant membrane-bound [NiFe] hydrogenase of *R. eutropha* H16 by molecular modelling

Yvonne Rippers[†], Tillmann Utesch[†], Peter Hildebrandt, Ingo Zebger and Maria Andrea Mroginski*

Technische Universität Berlin Institut f. Chemie, Sekr. PC14,, Straße des 17. Juni 135, D-10623 Berlin (Germany). Fax: +49 30 31421122; Tel: +49 30 31426403; E-mail: andrea.mroginski@tu-berlin.de

Content

SI 1. Construction of the homology model

The homology model of the membrane bound hydrogenase (MBH) of R. eutropha has been constructed with the Modeller v9.5 [1]. We used the optimized standard hydrogenase structure of D. vulgaris Miyazaki F. as template, which shows a sequence identity of 44 % with respect to the MBH. The pre-optimization of the template with the Modeller v9.5 was performed by fitting the crystallographic structure (1WUK) [2] via a multiple structural alignment with similar standard [NiFe] hydrogenase structures of D. gigas (2FRV) [3], D. fructosovorans (1YQW) [4] and D. desulfuricans (1E3D) [5] (see figure S1). The outcome of this structural modification was used as input for the modelling of the MBH. For the modelling, the ca. 50 N-terminal and ca. 60 C-terminal amino acids of the small subunit of the MBH were discarded, because these regions are not present in [NiFe] hydrogenases with known 3D structure. The active site, the iron sulphur clusters and their linked atoms were kept fixed during the process. The catalytic site structure was obtained from the oxygen sensitive hydrogenase of D. gigas [3], with two CN^- and one CO ligated to the Fe atom, as validated by IR spectroscopy.[6] The other crystallographic structures exhibit variations in the active site constitution that are not consistent with available spectroscopic data. In the structural model for the MBH, the flexible loop regions originating form gaps in the template were energy-optimized in order to obtain a refined, energetically favourable structure according to the dope-score of the Modeller v9.5.[1] During the modelling, the active site and the iron sulphur clusters were treated as structural constraints and were not further optimized. The resultant model was validated by the analysis of the corresponding Ramachandran plot [7] showing a Z-value of -0.202 which describes a well-refined structure (Figure S2). Compared to standard hydrogenases, the modelled structure of the MBH is characterized by a conserved first coordination sphere of the active site. The only sequential difference is a serine to threonine mutation, which leads to minor structural differences. Additionally, we observed a very high match of secondary structure motifs between template and target (Figure S3; Table S1).

The *MBH-H_2O* model was constructed from the MBH structure by placing a manually a water molecule in the vicinity of the Ni-Fe center, as observed in the crystal structures of other hydrogenases. Whereas the structural model for the C81S variant of the MBH [8] was obtained from the WT model incorporating the mutation with the CHARMM V32b package [9] and a water molecule in the active site cavity in analogy to the *MBH-H_2O* model, followed by a short energy minimization of the structure. (Figure S4)

In a comparison of the secondary structure content of the *MBH-H_2O* model with those from the standard hydrogenase from *D. vulgaris* Miyazaki F, we could detect only a slight loss in secondary structure of 4% for α-helix and for β-sheets (Table S1).

SI 2. General system setup

For the MD simulations, partial charges for the active site of the modelled MBH were derived by electrostatic potential fits according to the Merz-Singh-Kollmann scheme. The quantum mechanical calculations were performed with GAUSSIAN 09 [10] using the BP86 functional [11; 12] (table S4). These computations were done with the QM/MM approach where the active site is embedded in the charge cloud of the protein matrix. The iron-sulfur clusters were treated with the partial charges obtained by Teixeira et al. [13] and were not adjusted to the protein environment, because they play only a minor role in the IR spectra calculation of the catalytic centre. The protonation of the protein side chains was set according to pH 7 and the protonation of the histidine residues was adjusted according to their specific environments. Subsequently, the protonated MBH was solvated in TIP3P water [14] using the VMD1.6 package and the system was neutralized by adding sodium and chloride ions.

In a first step, the systems were minimized, heated to 300 K and equilibrated in a 1 ns long MD simulation. The resultant structures and their partial charges were optimized in a second step by QM/MM calculations in order to obtain an optimal starting point for the IR spectra calculation.

The last part is the IR spectra calculation performed by a hybrid approach combining MD and QM/MM simulations. Out of a 2.5 ns long MD simulation, 25 snapshots in a 100 ps interval were extracted in order to calculate IR spectra on the QM/MM level.

SI 3. Molecular dynamics (MD) simulations

All MD simulations were performed with NAMD 2.7 [15] using the CHARMM 22 force field [16]. The system was handled under periodic boundary and NPT conditions facilitated by Langevin piston dynamics [17] with a cut-off of 12 Å for van der Waals interactions and short-range electrostatics. Long-range electrostatics interactions were calculated with the particle mesh Ewald summation [18]. The time step was set to 2 fs. The active site and all iron sulfur clusters were treated as nearly rigid bodies, as parameters for their internal motions were not available.

SI 4. QM/MM calculation

The QM part consists of 26 atoms containing the active site and side chains of the four coordinating cysteines, with nickel and iron in the oxidation state +III and +II, respectively. Therefore, the overall charge of the QM part is -2 e with a multiplicity of M=2. For calculations with the enlarged QM part (75 atoms), containing aminoacids that are involved in the hydrogen bond network (R530, T533 and H82), the overall charge is -1 e with the same multiplicity as for less QM atoms. Density functional theory (DFT) calculations were carried out at the BP86 [11; 12] level of theory using the 6-31g(d) basis set for all atoms excluding nickel and iron for which Ahlrichs triple-zeta polarization all-electron basis set (TZVP) was employed [19]. In a sphere with a radius of 20 Å around the iron of the active site, the protein is treated molecular mechanically (MM part).

IR spectra of the active site were calculated for all 25 snapshots of the three structural models. The geometry of each snapshot has been optimized at QM/MM level of theory using a limited memory quasi-Newton L-BFGS algorithm with the modular program package ChemShell [20] and choosing a maximum gradient component convergence criterion of 0.0008. Energies and gradients for the QM part were calculated with the TURBOMOLE 6.10 computational chemistry program. For the MM part, the empirical CHARMM22 MM force field [16] was employed. Covalent bonds were cut at the QM/MM border and saturated by hydrogen link atoms. The coupling between QM and MM was computed using the electrostatic embedding approach combined with a charge-shift scheme [21]. Frequency calculations were carried out with GAUSSIAN 09. Further degrees of freedom introduced by the additional hydrogen atoms were projected out of the Hessian matrix before diagonalization [22]. Final IR spectra (figure S5 and S6) were computed using the instantaneous normal mode analysis (INMA) approach [23].

Figure S1: *CLUSTAL 2.1 multiple sequence alignment of the target MBH protein and the templates employed for building the homology model.*

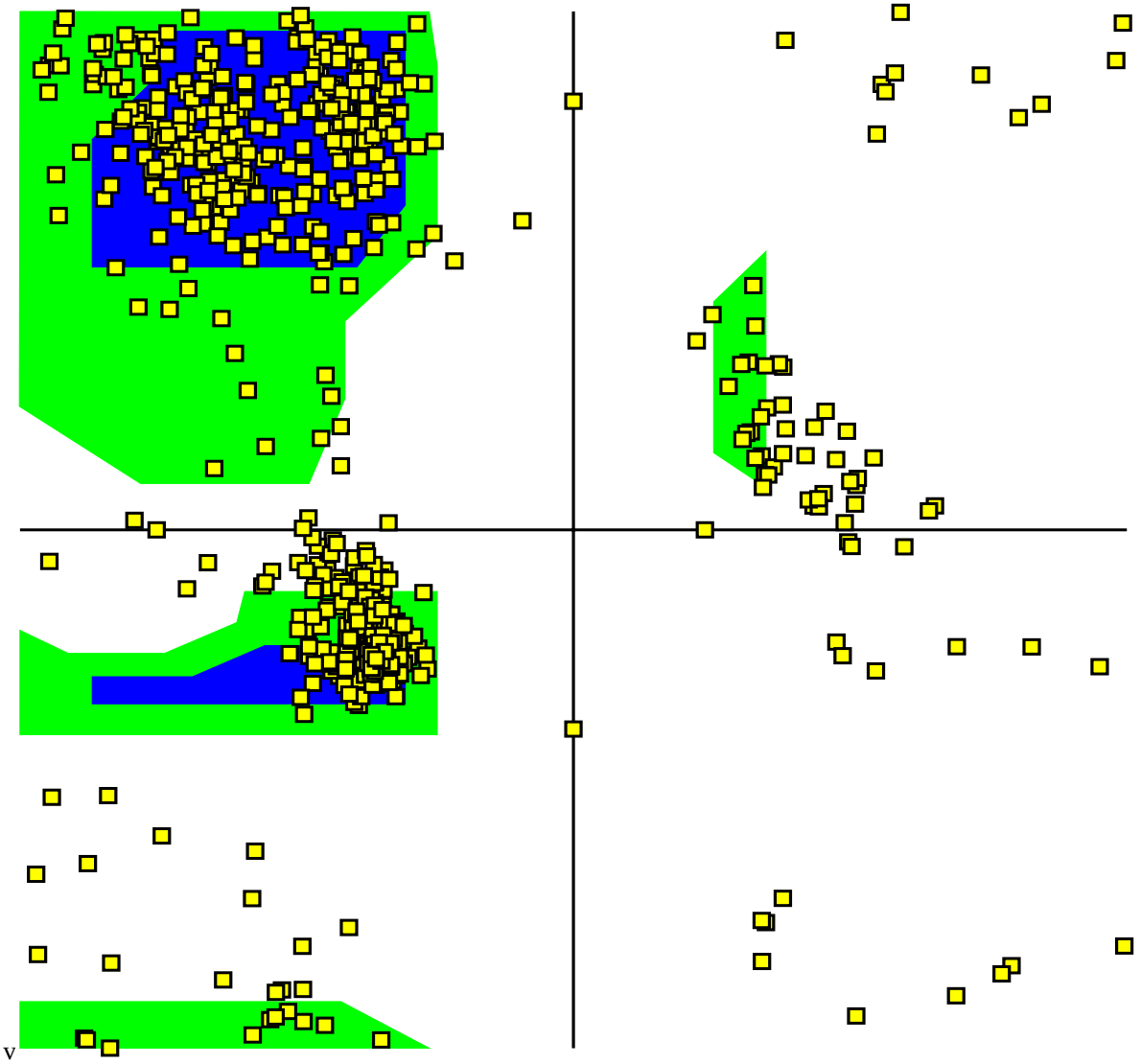

Figure S2. Ramachandran plot of the MBH model constructed with VMD1.8.7 [25].

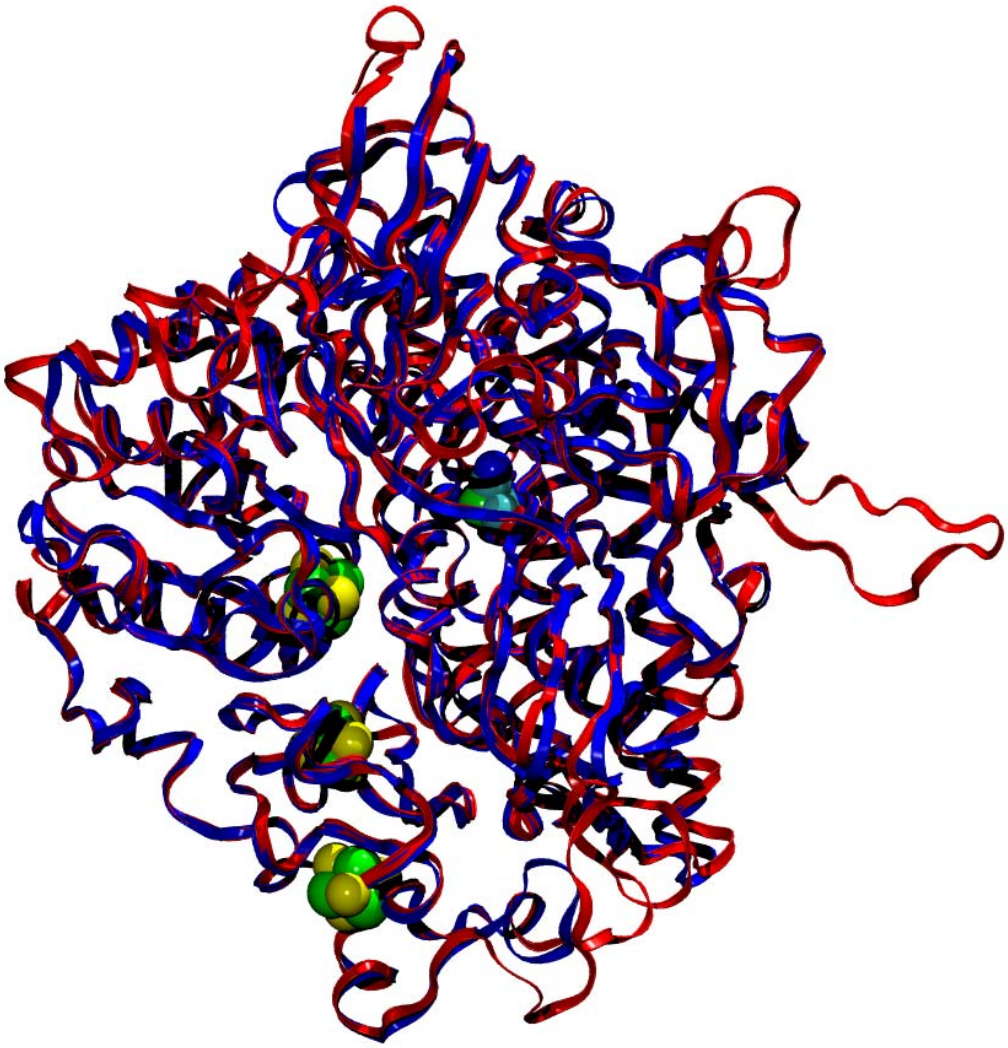

Figure S3: 3D alignment, performed with the combinatorial extension (CE) method [24], between the modeled MBH (red) and the [NiFe] hydrogenase of D. gigas (blue). The sequence identities between the two structures of the small (A) and large (B) subunits are 44.7 and 43.2 %, respectively. These similarities lead to root-mean-square deviations of 0.8 Å for both subunits.

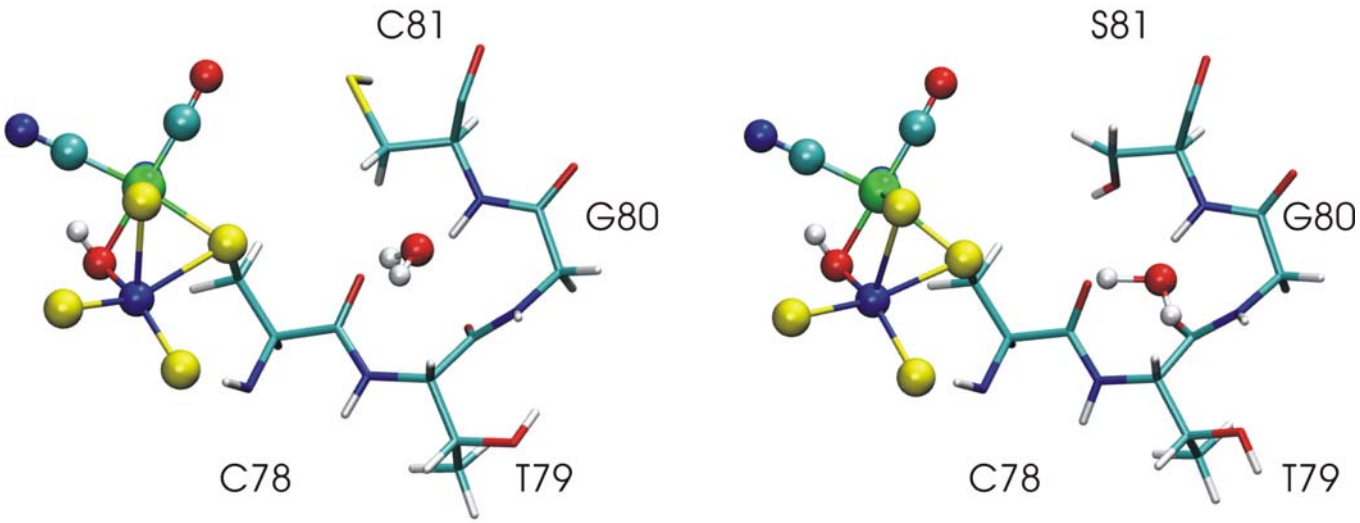

Figure S4: active site structure of wild type MBH (left) and the C81S mutant (right) including the water molecule responsible for the shift of the CO stretching frequency.

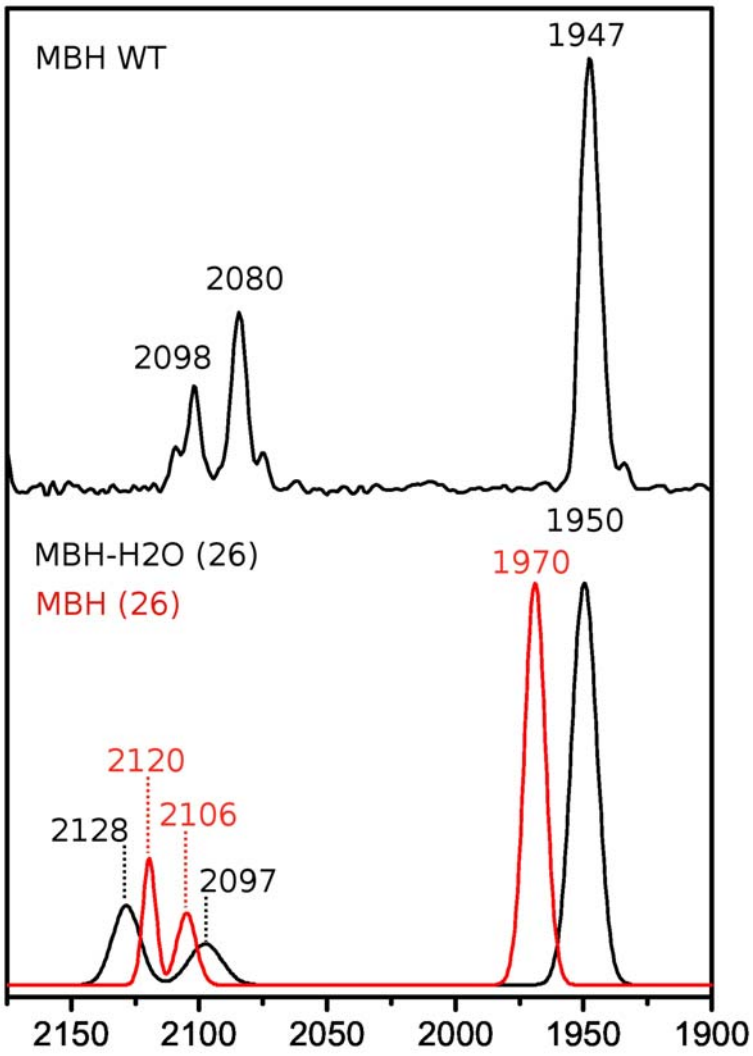

Figure S5. Experimental IR spectrum (top) of the Ni-B state of MBH [8] compared with the calculated IR spectra (bottom) of MBH-H_2O (black) and the MBH model (red, no internal water), both with 26 QM atoms.

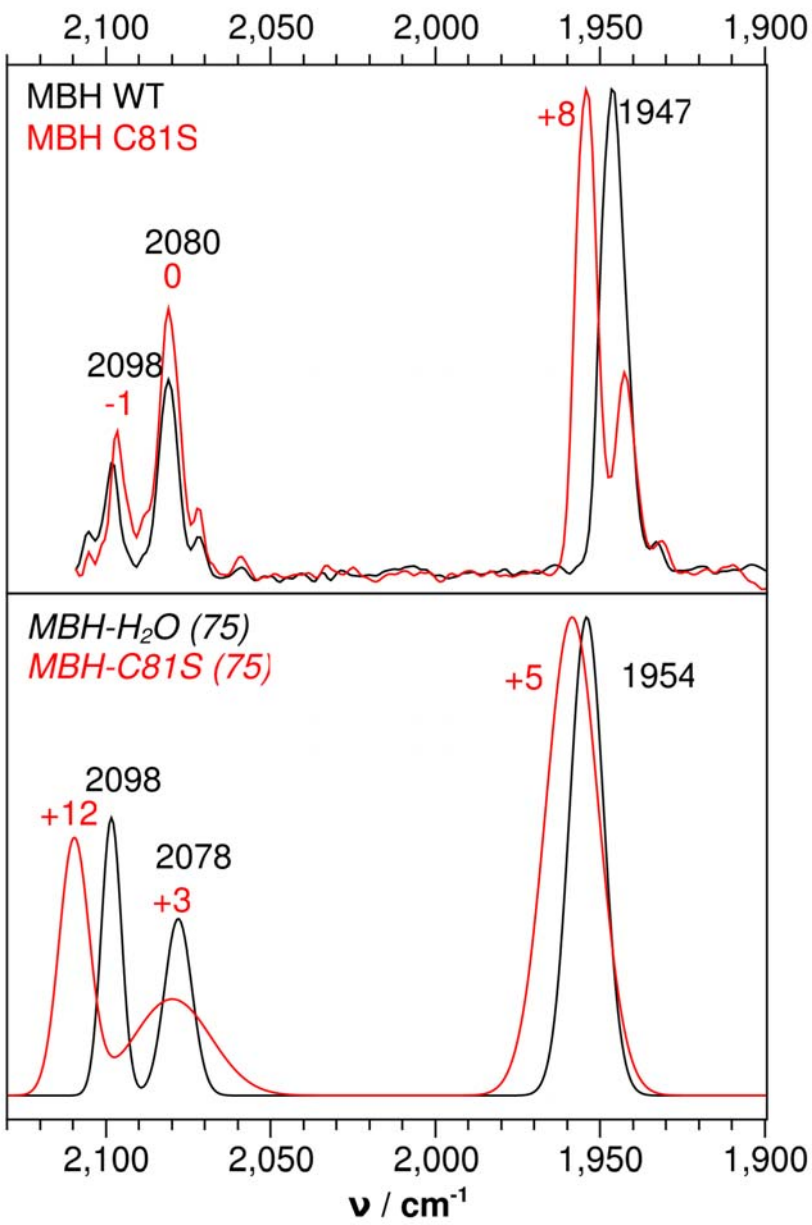

Figure S6. Experimental (top) [8] and calculated IR spectra (bottom) with 75 QM atoms of the MBH in Ni-B state. Black traces refer to the WT and red traces to the C81S mutant.

	1WUK	MBH-H_2O
α-helix	41	37
β-sheet	15	11
coils	45	52

Table S1. Secondary structure content (%)

	MBH	MBH-H_2O	MBH-C81S	D.gigas [3]
Ni - Fe / Å	2.85	2.91	2.91	2.90
Fe – CN1	1.85	1.84	1.87	1.87
Fe –CN2	1.89	1.91	1.88	1.67
Fe -CO	1.72	1.73	1.73	1.87
C-N1 / Å	1.19 ± 0.0008	1.19 ± 0.0007	1.19 ± 0.0018	1.20
C-N2 / Å	1.19 ± 0.0005	1.19 ± 0.0006	1.19 ± 0.0010	1.16
C-O / Å	1.19 ± 0.0009	1.18 ± 0.0011	1.18 ± 0.0014	1.17
Fe-OH / Å	2.05 ± 0.0052	2.01 ± 0.0100	2.03 ± 0.0100	2.14
Ni-OH / Å	1.92 ± 0.0048	1.96 ± 0.0063	1.96 ± 0.0095	1.73
O-H / Å	0.98 ± 0.0005	0.99 ± 0.0088	0.99 ± 0.0033	-
Fe-S(C78) / Å	2.33 ± 0.0076	2.39 ± 0.0133	2.36 ± 0.0309	2.23
Fe-S(C600) / Å	2.38 ± 0.0093	2.38 ± 0.0104	2.38 ± 0.0103	2.20
Ni-S(C75) / Å	2.21 ± 0.0115	2.21 ± 0.0073	2.22 ± 0.0160	2.16
Ni-S(C78) / Å	2.31 ± 0.0089	2.34 ± 0.0111	2.34 ± 0.0212	2.58
Ni-S(C597) / Å	2.24 ± 0.0056	2.23 ± 0.0170	2.25 ± 0.0230	2.27
Ni-S(C600) / Å	2.39 ± 0.0088	2.39 ± 0.0088	2.42 ± 0.0394	2.62
CO-SH(C81) / Å	6.43 ± 0.3993	3.51 ± 0.1240	-	-
CO-OH(S81) / Å	-	-	4.71 ± 0.7331	-
CO–CG2(V71) / Å	-	-	-	3.58
CO-water / Å	-	7.88 ± 0.3621	5.42 ± 0.5076	6.78
CN1-N(P552) / Å	3.43 ± 0.0539	3.29 ± 0.0515	3.44 ± 0.0730	4.07
CN1-N(T553) / Å	2.88 ± 0.0283	2.84 ± 0.0312	2.97 ± 0.0753	-
CN1-N(S553) / Å	-	-	-	2.99
CN1-N(T554) / Å	4.66 ± 0.1048	4.76 ± 0.1189	5.21 ± 0.0756	5.25
CN2-N(P529) / Å	3.46 ± 0.0749	3.49 ± 0.0750	3.42 ± 0.0722	3.86
CN2-N(R530) / Å	3.06 ± 0.0754	3.05 ± 0.0509	2.89 ± 0.0428	3.27
CN2-NE(R530) / Å	5.48 ± 0.0685	3.87 ± 0.1616	5.29 ± 0.1706	4.31
CN2-NH1(R530) / Å	5.55 ± 0.1341	3.63 ± 0.1949	4.03 ± 0.1524	3.20
OH-NH1(R530) / Å	5.44 ± 0.1541	2.74 ± 0.0248	2.95 ± 0.0626	3.34
S(C78)-NE2(H82) / Å	3.96 ± 0.3040	3.70 ± 0.2014	4.60 ± 0.1975	4.178
S(C600)-NE2(H82) / Å	3.53 ± 0.1524	3.90 ± 0.1344	3.54 ± 0.1844	3.484

Table S2. Selected structural parameters of the active site and its environment. The residue numbering refers to the MBH.

atom	(a) initial	(b) MBH	(c) MBH-H_2O	(d) MBH-C81S
NI	0.822712	0.493845	0.429857	0.508526
FE	0.506073	-0.714954	-0.097461	-0.512543
C1	0.073806	0.371088	0.224581	0.327768
N1	-0.577114	-0.748065	-0.650401	-0.761602
C2	0.073806	0.371088	0.272715	0.366494
N2	-0.577114	-0.748065	-0.675122	-0.712609
C3	0.157395	0.451380	0.165203	0.395880
O3	-0.323338	-0.332060	-0.283890	-0.319774
O4	-0.590084	-0.482670	-0.766420	-0.615794
H4	0.267808	0.391942	0.387199	0.378661
CB (75)	0.085107	0.017872	0.003677	0.047494
HB1 (75)	0.013506	0.023115	0.050628	0.038888
HB2 (75)	0.013506	0.023115	0.050628	0.038888
SG (75)	-0.602661	-0.548562	-0.565503	-0.523080
CB (78)	-0.042270	-0.238491	-0.144771	-0.165282
HB1 (78)	0.059652	0.171108	0.133469	0.145268
HB2 (78)	0.059652	0.171108	0.133469	0.145268
SG (78)	-0.503469	-0.203247	-0.266621	-0.337135
CB (597)	0.085107	-0.254554	-0.204408	-0.179922
HB1 (597)	0.013506	0.139560	0.166630	0.113707
HB2 (597)	0.013506	0.139560	0.166630	0.113707
SG (597)	-0.602661	-0.354212	-0.410744	-0.243713
CB (600)	-0.042270	-0.406052	-0.286409	-0.308113
HB1 (600)	0.059652	0.202056	0.157087	0.174468
HB2 (600)	0.059652	0.202056	0.157087	0.174468
SG (600)	-0.503469	-0.137962	-0.147110	-0.289918

Table S3. Partial charges for the MD simulation derived for the active site of the MBH for the initial configuration (a) and the optimized structures for the wild type without additional water molecule (b), the wild type with water (c) and the C81S mutant with water (d).

	D. gigas [3]	D. vulgaris [2]	our work	Jayapal [26]	Stein [27]
Fe – Cys600	2.20	2.33	2.38	2.49	2.47
Fe – Cys78	2.23	2.28	2.33 – 2.39	2.43	2.46
Ni – Cys600	2.62	2.54	2.39 – 2.42	2.38	2.51
Ni – Cys78	2.58	2.28	2.31 – 2.34	2.29	2.36
Ni – Cys597	2.27	2.14	2.23 – 2.25	2.27	2.31
Ni – Cys75	2.16	2.22	2.21– 2.22	2.25	2.29
Fe – OH^-	2.14	1.98	2.01 – 2.05	2.00	2.09
Ni – OH^-	1.74	1.63	1.92 – 1.96	1.89	1.98
Ni – Fe	2.90	2.67	2.85 – 2.91	2.89	3.05

Table S4. Comparison of selected bond lengths (in Å) of the active site with crystallographic structures and previous calculated data. The values obtained in the present work refer to the different models. The numbering refers to the MBH sequence.

References

[1] A. Sali, T. L. Blundell, *J. Mol. Biol.* **1993**, 234, 779-815
[2] H. Ogata, S. Hirota, A. Nakahara, H. Komori, N. Shibata, T. Kato, K. Kano, Y. Higuchi, *Structure* **2005**, 13, 1635-1642
[3] A. Volbeda, E. Garcin, C. Piras, A. L. de Lacey, V. M. Fernandez, E. C. Hatchikian, M. Frey, J. C. Fontecilla-Camps, *J. Am. Chem. Soc*. **1996,** 118, 12989-12996
[4] A. Volbeda, L. Martin, C. Cavazza, M. Matho, B. W. Faber, W. Roseboom, S. P. Albracht, E. Garcin, M. Rousset, J. C. Fontecilla-Camps, *J. Biol. Inorg. Chem.* **2005**, 10, 239– 249
[5] P. M. Matias, C. M. Soares, L. M. Saraiva, R. Coelho, J. Morais, J. Le Gall, M. A. Carrondo, *J. Biol. Inorg. Chem.* **2001**, 6, 63-81
[6] A. J. Pierik, W. Roseboom, R. P. Happe, K. A. Bagley, S. P. J. Albracht, *J. Biol. Chem.*, **1999**, 274, 3331-3337
[7] R. W. W. Hooft, G. Vriend, C. Sander, E. E. Abola, *Nature* **1996**, 381, 272-272
[8] M. Saggu, M. Ludwig, B. Friedrich, P. Hildebrandt, R. Bittl, F. Lendzian, O. Lenz, I. Zebger, *ChemPhysChem* **2010**, 11, 1215-1224
[9] B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, M. Karplus, *J. Comput. Chem.* **1983**, 4, 187-217
[10] 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. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, 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, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, D. J. Fox, *Gaussian, Inc., Wallingford CT* **2009**
[11] A. D. Becke, *Phys. Rev. A* **1998**, 38, 3098 - 3100
[12] J. P. Perdew, *Phys. Rev. B* **1986**, 33, 8822 - 8824
[13] V. H. Teixeira, A. M. Baptista, C. M. Soares, *Biophys. J.* **2006**, 91, 2035-2045
[14] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, *J. Chem. Phys.* **1983**, 79, 926-935
[15] J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kale, K. J. Schulten, *Comp. Chem.* **2005**, 26, 1781-1802.
[16] A. D. MacKerell et al, *J.Phys.Chem.B* **1998**, 102, 3586-3616
[17] S. E. Feller, Y. H. Zhang, R. W. Pastor, B. R. Brooks, *J. Chem. Phys.* **1995**, 103, 4613-4621
[18] T. Darden, D. York, L. Pedersen, *J. Chem. Phys.* **1993**, 98, 10089-10092
[19] F. Weigend, R. Ahlrichs, *Phys. Chem. Chem. Phys* **2005**, 7, 3297-3305
[20] P. Sherwood, A. H. de Vries, M. F. Guest, G. Schreckenbach, C. R. A. Catlow, S. A. French, A. A. Sokol, S. T. Bromley, W. Thiel, A. J. Turner, S. Billeter, F. Terstegen, S. Thiel, J. Kendrick, S. C. Rogers, J. Casci, M. Watson, F. King, E. Karlsen, M. Sjovoll, A. Fahmi, A. Schafer, C. Lennartz, *J. Mol. Struc. - Theochem* **2003**, 632 , 1-28
[21] D. Bakowies, W. Thiel, *J. Phys. Chem.* **1996**, 100, 10580-10594
[22] Q. Cui, M. Karplus, *J. Chem. Phys.* **2000**, 112, 1133-1149
[23] M. Nonella, G. Mathias, P. Tavan, *J. Phys. Chem. A* **2003**, 107, 8638-8647
[24] I. N. Shindyalov, P. E. Bourne, *Protein Eng.* **1998**, 11, 739-747
[25] W. Humphrey, A. Dalke, K. Schulten, *J. Mol. Graphics* **1996**, 14, 33-38
[26] P. Jayapal, M. Sundararajan, I. H. Hillier, N. A. Burton, *Phys. Chem. Chem. Phys.* **2008**, 10, 4249-4257
[27] M. Stein, W. Lubitz, *Phys. Chem. Chem. Phys.* **2001**, 3, 2668-2675