Supporting Information

A strong H-bond between a cysteine and the catalytic center of a [NiFe]-hydrogenase

Chara Karafoulidi-Retsou;[a] Sagie Katz;[a] Stefan Frielingsdorf;[a] Oliver Lenz;[a] Ingo Zebger[a]* and Giorgio Caserta[a]*

[a]Institut für Chemie, Technische Universität Berlin, Straße des 17. Juni 135, 10623 Berlin, Germany

E-mails: ingo.zebger@tu-berlin.de, giorgio.caserta@tu-berlin.de

Table of Contents

Experimental procedures

Protein purification p.1

Hydrogenase activity measurements p.1

IR Spectroscopy p.1

Figure S1: IR absorbance spectrum of the as-isolated (oxidized) native *Cn*MBH at 298 K. p.2

Figure S2: IR absorbance spectrum of the H_2-reduced native *Cn*MBH at 298 K. . p.3

Figure S3: IR difference spectrum of H_2-reduced regulatory [NiFe]-hydrogenase from *C. necator* (*Cn*RH) at 95 K. p.4

Figure S4: Sequence alignment of the large subunits from *Cn*MBH, *Ec*Hyd1, *Cn*RH, *Pf*SH1 and *Dv*MF hydrogenase. p.6

Figure S5: Sequence logo for the large subunit region around Cys81 of *Cn*MBH derived from the multiple sequence alignment of [NiFe]-hydrogenases belonging to subgroup *1d*. p.7

Figure S6: IR spectroscopic characterization of the *Cn*MBH^{Cys81Ser} variant. p.8

Figure S7: Comparison of the IR difference spectra of native *Cn*MBH and the *Cn*MBH^{Cys81Ser} variant. p.9

Supplementary discussion p.10

Supplementary references p.10

Experimental Procedures

Protein purification

Recombinant *C. necator* (*Cn*) strains HF649 [1] and HF677[2], carrying plasmids for overproduction of native *Cn*MBH and the *Cn*MBH^{Cys81Ser} variant, were cultivated as previously described in a basic mineral medium containing fructose and glycerol as carbon and energy sources.[3] When the bacterial cultures reached an optical density at 436 nm of 11-13, the cells were harvested by centrifugation (11,500 x *g*, 4 °C, 15 min), and the cell pellet was flash frozen in liquid nitrogen and stored at –80 °C until further use. Both *Cn*MBH versions were purified by affinity chromatography as described previously.[3] The purified enzymes were flash-frozen and stored in liquid nitrogen. The protein concentration was determined using a Pierce BCA Protein Assay kit (Thermo Scientific) using bovine serum albumin (BSA) as standard.

Hydrogenase activity measurements

H_2-oxidation activity was measured spectrophotometrically using a Cary50 UV−vis spectrophotometer (Varian, Agilent, Santa Clara, California) as described previously.[3]

IR Spectroscopy

For IR measurements, the MBH variants at a concentration of approx. 1 mM resided in 50 mM K_iPO_4 buffer, pH 5.5, containing 150 mM NaCl and 50 % glycerol. The addition of glycerol ensured the formation of a transparent glass at temperatures below 220 K. The pH of the buffer was adjusted after the addition of glycerol. The samples were reduced by exposure to 100 % humidified H_2 gas for at least 30 min at 10 °C. After reduction, the protein samples were transferred into an airtight sandwich cell for IR spectroscopy, consisting of two CaF_2 windows separated by a 20 µm Teflon (PTFE) spacer. Subsequently, the cell was transferred into a homemade liquid-nitrogen cooled bath cryostat. IR spectra were recorded with a resolution of 2 cm^{-1} using a Bruker Tensor 27 FT-IR spectrometer, equipped with a liquid-nitrogen cooled mercury cadmium telluride (MCT) detector. The sample compartment of the spectrometer was continuously purged with dry N_2 gas during the measurements. The Bruker OPUS software, version 7.8, was used for data analysis. IR single channel spectra were obtained by averaging 200 scans. Absorbance spectra were calculated from averaged single channel spectra (at least 7 reference spectra) of the sample using a buffer spectrum as reference. Illumination experiments were performed using a homemade LED ring system with an emission maximum at 460 nm (power density 1-10 mW/cm^2). Light-*minus*-dark spectra were calculated using the corresponding dark single spectra as reference.

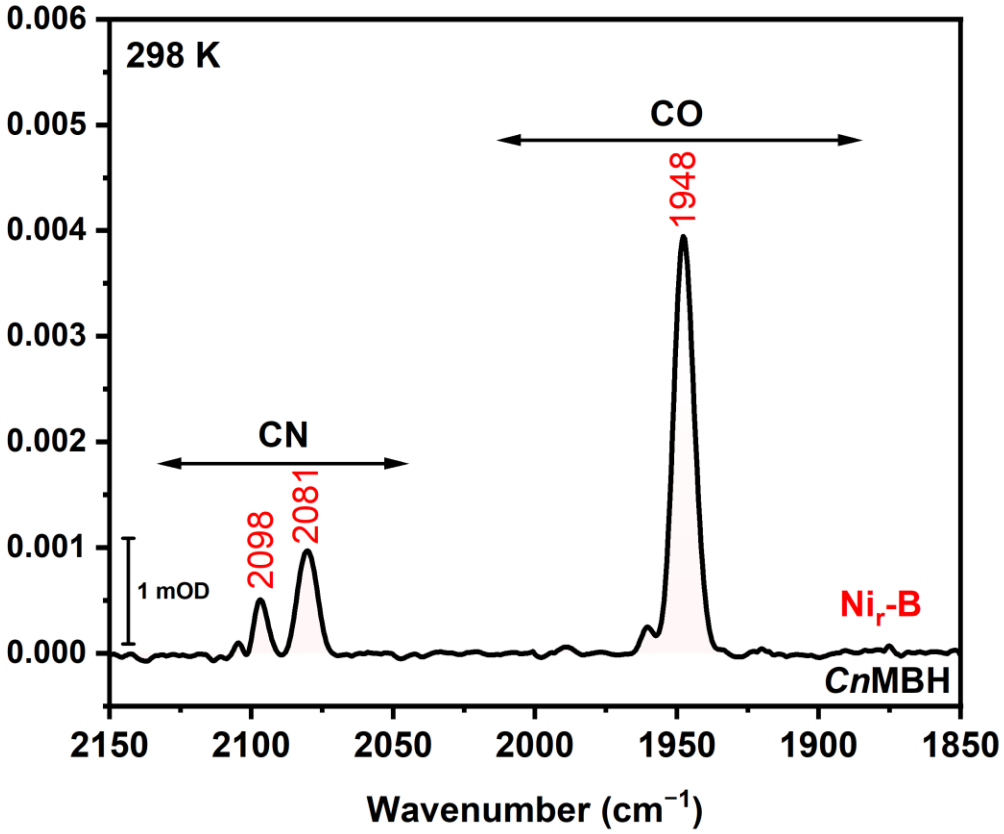


Figure S1. IR absorbance spectrum of the as-isolated (oxidized) native *Cn*MBH at 298 K. The spectrum of the as-isolated sample (ca. 1 mM MBH in 50 mM K_iPO_4, pH 5.5, 150 mM NaCl, 50 % glycerol) is dominated by absorptions attributed to the Ni_r-B state with a ν_{CO} stretching vibration at 1948 cm^{-1} and ν_{CN} absorptions at 2081 and 2098 cm^{-1}.

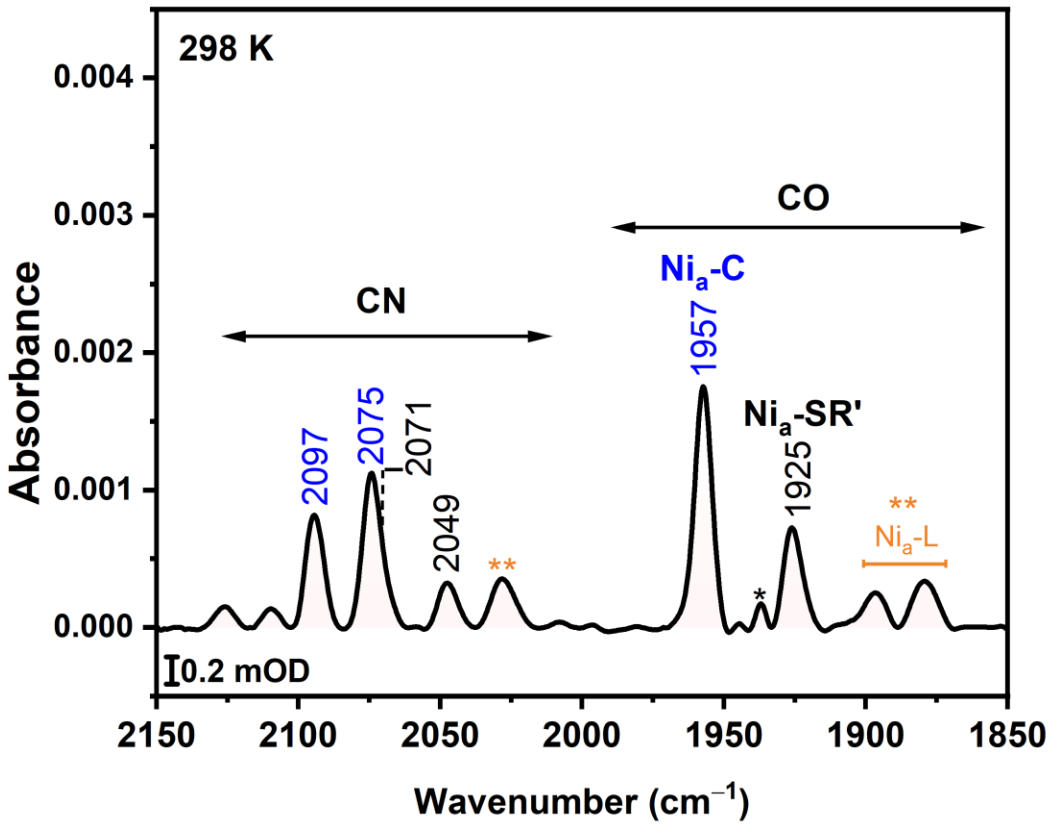


Figure S2. IR absorbance spectrum of the H_2-reduced native *Cn*MBH at 298 K. The solution IR spectrum (ca. 1 mM *Cn*MBH in 50 mM K_iPO_4, pH 5.5, 150 mM NaCl, 50 % glycerol) shows predominant signals of the Ni_a-C (ν_{CO} at 1957 cm^{-1} and ν_{CN} at 2075 and 2097 cm^{-1}) and the Ni_a-SR´ (ν_{CO} at 1925 cm^{-1} and ν_{CN} at 2049 and 2071 cm^{-1}) states of the catalytic center. Ni_a-SR' is reduced by one e^- more than Ni_a-C. The signal marked by a black asterisk is attributed to remnants of $Ni_{a/r}$-S species (1936 cm^{-1}). The signals marked with two orange asterisks are attributed to sub-stoichiometric Ni_a-L species.[5]

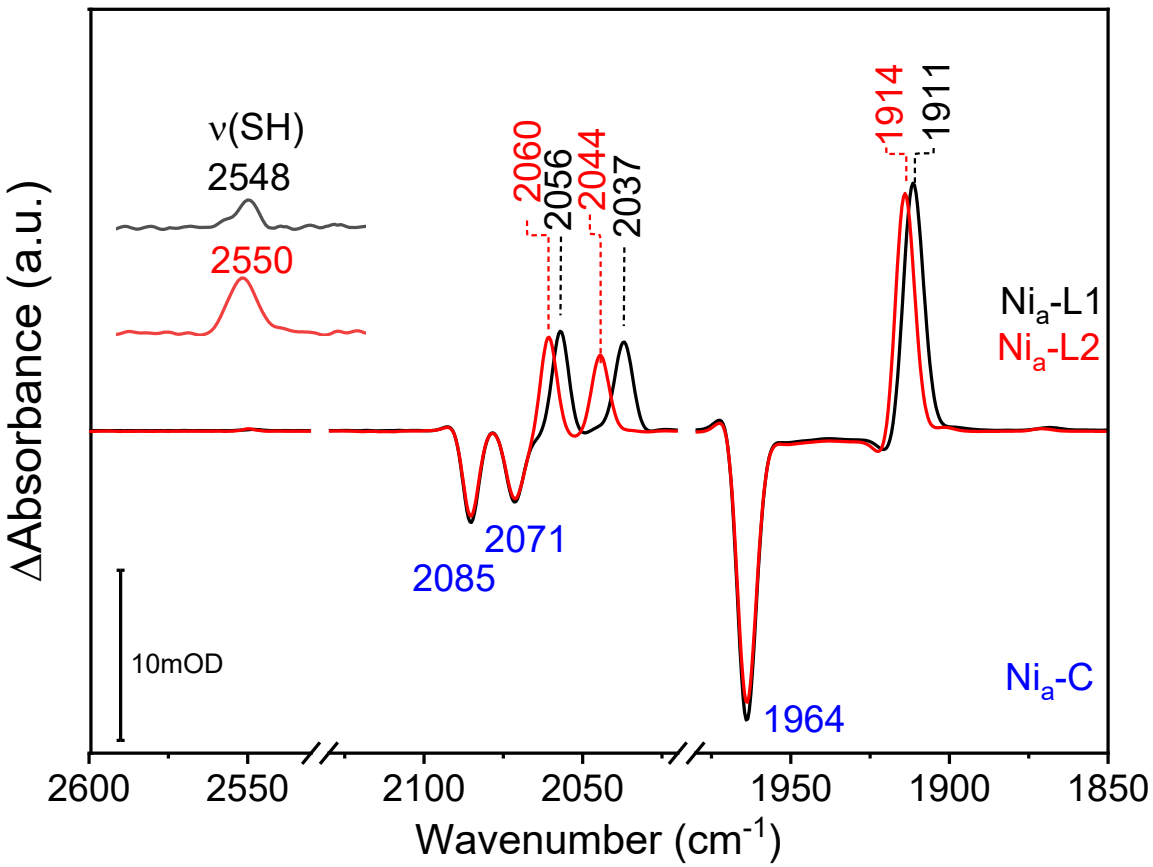
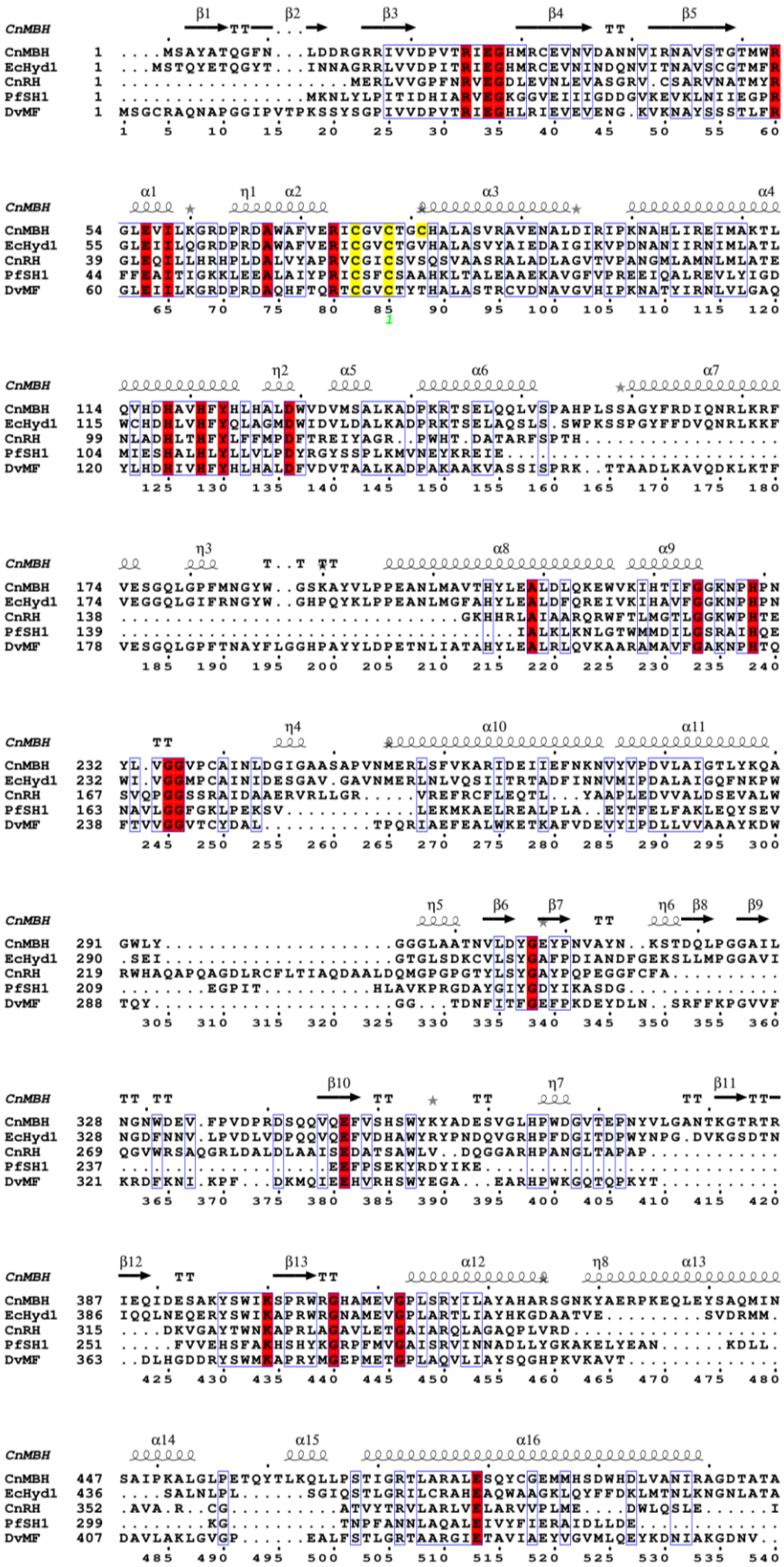


Figure S3. IR difference spectrum of H_2-reduced regulatory [NiFe]-hydrogenase from *C. necator* (*Cn*RH) at 95 K. Light-*minus*-dark IR difference spectra of H_2-reduced *Cn*RH (ca. 1.2 mM in 50 mM Tris, pH 8.0, 150 mM NaCl, 25% glycerol) displaying negative absorption bands for the Ni_a-C state and positive signals for Ni_a-L1 and Ni_a-L2 states. The spectral region comprises the frequencies associated with the ν_{CO} and ν_{CN} stretching vibrations of the diatomic ligands as well as those originating from the ν_{SH}. The latter are enlarged in the top left of the figure to make them clearly visible. Remarkably, the ν_{SH} stretchings of native *Cn*MBH (**Fig. 3a**) are visible without enlargement of the spectrum, suggesting that the detected protonated cysteine thiolate (Cys81, see main text) participates in a strong H–bond that polarizes the S–H bond. The figure was generated using data reported in https://doi.org/10.1021/jacs.3c01625.[6]


CnMBH
CnMBH 1 ....MSAYATQGFN...LDDRGRRIVVDPVTRIEGHMRCEVNVDANNVIRNAVSTGTMWR
EcHyd1 1 ...MSTQYETQGYT...INNAGRRLVVDPITRIEGHMRCEVNINDQNVITNAVSCGTMFR
CnRH 1 .....MERLVVGPFNRVEGDLEVNLEVASGRV.CSARVNATMYR
PfSH1 1 .....MKNLYLPITIDHIARVEGKGGVEIIIGDDGVKEVKLNIIEGPR
DvMF 1 MSGCRAQNAPGGIPVTPKSSYSGPIVVDPVTRIEGHLRIEVEVENG.KVKNAYSSSTLFR
CnMBH 54 GLEVILKGRDPRDAWAFVERICGVCTGCHALASVRAVENALDIRIPKNAHLIREIMAKTL
EcHyd1 55 GLEIILQGRDPRDAWAFVERICGVCTGVHALASVYAIEDAIGIKVPDNANIIRNIMLATL
CnRH 39 GLEQILLHRHPLDALVYAPRVCGICSVSQSVAASRALADLAGVTVPANGMLAMNLMLATE
PfSH1 44 FFEAITIGKKLEEALAIYPRICSFCSAAHKLTALEAAEKAVGFVPREEIQALREVLYIGD
DvMF 60 GLEILLKGRDPRDAQHFTQRTCGVCTYTHALASTRCVDNAVGVHIPKNATYIRNLVLGAQ
CnMBH 114 QVHDHAVHFYHLHALDWVDVMSALKADPKRTSELQQLVSPAHPLSSAGYFRDIQNRLKRF
EcHyd1 115 WCHDHLVHFYQLAGMDWIDVLDALKADPRKTSELAQSLS.SWPKSSPGYFFDVQNRLKKF
CnRH 99 NLADHLTHFYLFFMPDFTREIYAGR..PWHT.DATARFSPTH.....
PfSH1 104 MIESHALHLYLLVLPDYRGYSSPLKMVNEYKREIE.....
DvMF 120 YLHDHIVHFYHLHALDFVDVTAALKADPAKAAKVASSISPRK..TTAADLKAVQDKLKTF
CnMBH 174 VESGQLGPFMNGYW..GSKAYVLPPEANLMAVTHYLEALDLQKEWVKIHTIFGGKNPHPN
EcHyd1 174 VEGGQLGIFRNGYW..GHPQYKLPPEANLMGFAHYLEALDFQREIVKIHAVFGGKNPHPN
CnRH 138 .....GKHHRLAIAARQRWFTLMGTLGGKWPHTE
PfSH1 139 .....IALKLKNLGTWMMDILGSRAIHQE
DvMF 178 VESGQLGPFTNAYFLGGHPAYYLDPETNLIATAHYLEALRLQVKAARAMAVFGAKNPHTQ
CnMBH 232 YL.VGGVPCAINLDGIGAASAPVNMERLSFVKARIDEIIEFNKNVYVPDVLAIGTLYKQA
EcHyd1 232 WI.VGGMPCAINIDESGAV.GAVNMERLNLVQSIITRTADFINNVMIPDALAIGQFNKPW
CnRH 167 SVQPGGSSRAIDAAERVRLLGR.....VREFRCFLEQTL...YAAPLEDVVALDSEVALW
PfSH1 163 NAVLGGFGKLPEKSV.....LEKMKAELREALPLA..EYTFELFAKLEQYSEV
DvMF 238 FTVVGGVTCYDAL.....TPQRIAEFEALWKETKAFVDEVYIPDLLVVAAAYKDW
CnMBH 291 GWLY.....GGGLAATNVLDYGEYPNVAYN..KSTDQLPGGAIL
EcHyd1 290 .SEI.....GTGLSDKCVLSYGAFPDIANDFGEKSLLMPGGAVI
CnRH 219 RWHAQAPQAGDLRCFLTIAQDAALDQMGPGPGTYLSYGAYPQPEGGFCFA.....
PfSH1 209 .....EGPIT.....HLAVKPRGDAYGIYGDYIKASDG.....
DvMF 288 TQY.....GG...TDNFITFGEFPKDEYDLN..SRFFKPGVVF
CnMBH 328 NGNWDEV.FPVDPRDSQQVQEFVSHSWYKYADESVGLHPWDGVTEPNYVLGANTKGTRTR
EcHyd1 328 NGDFNNV.LPVDLVDPQQVQEFVDHAWYRYPNDQVGRHPFDGITDPWYNPG.DVKGSDTN
CnRH 269 QGVWRSAQGRLDALDLAAISEDATSAWLV..DQGGARHPANGLTAPAP.....
PfSH1 237 .....EEFPSEKYRDYIKE.....
DvMF 321 KRDFKNI.KPF...DKMQIEEHVRHSWYEGA...EARHPWKGQTQPKYT.....
CnMBH 387 IEQIDESAKYSWIKSPRWRGHAMEVGPLSRYILAYAHARSGNKYAERPKEQLEYSAQMIN
EcHyd1 386 IQQLNEQERYSWIKAPRWRGNAMEVGPLARTLIAYHKGDAATVE.....SVDRMM.
CnRH 315 ....DKVGAYTWNKAPRLAGAVLETGAIARQLAGAQPLVRD.....
PfSH1 251 .....FVVEHSFAKHSHYKGRPFMVGAISRVINNADLLYGKAKELYEAN.....KDLL.
DvMF 363 ..DLHGDDRYSWMKAPRYMGEPMETGPLAQVLIAYSQGHPKVKAVT.....
CnMBH 447 SAIPKALGLPETQYTLKQLLPSTIGRTLARALESQYCGEMMHSDWHDLVANIRAGDTATA
EcHyd1 436 ....SALNLPL.....SGIQSTLGRILCRAHEAQWAAGKLQYFFDKLMTNLKNGNLATA
CnRH 352 .AVA.R..CG.....ATVYTRVLARLVELARVVPLME....DWLQSLE.....I
PfSH1 299 .....KG.....TNPFANNLAQALEIVYFIERAIDLLDE.....
DvMF 407 DAVLAKLGVGP.....EALFSTLGRTAARGIETAVIAEYVGVMLQEYKDNIAKGDNV..

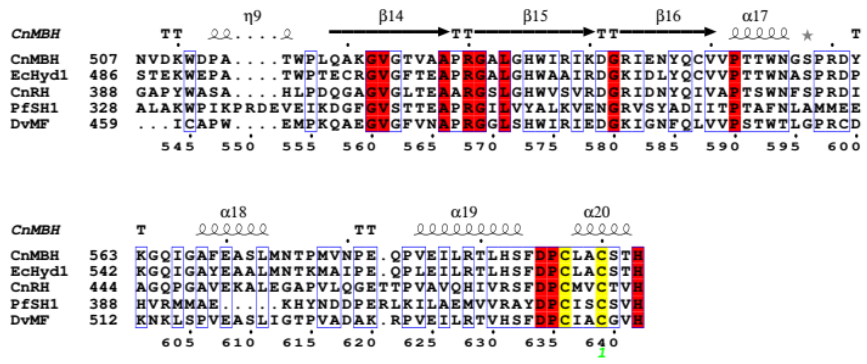


Figure S4. Sequence alignment of the large subunits from *Cn*MBH, *Ec*Hyd1, *Cn*RH, *Pf*SH1 and *Dv*MF hydrogenase. The four cysteine residues coordinating the NiFe(CN)$_2$CO cofactor (Cys 75, 78, 597 and 600 in *Cn*MBH) are highlighted in yellow, conserved residues are highlighted in red, similar residues are framed in blue boxes, secondary structural elements (derived from PDB:3RGW)[7] are shown on top of the alignment. The additional Cys81 of *Cn*MBH is also highlighted in yellow. The figure was generated using ESPript 3.0.[8]

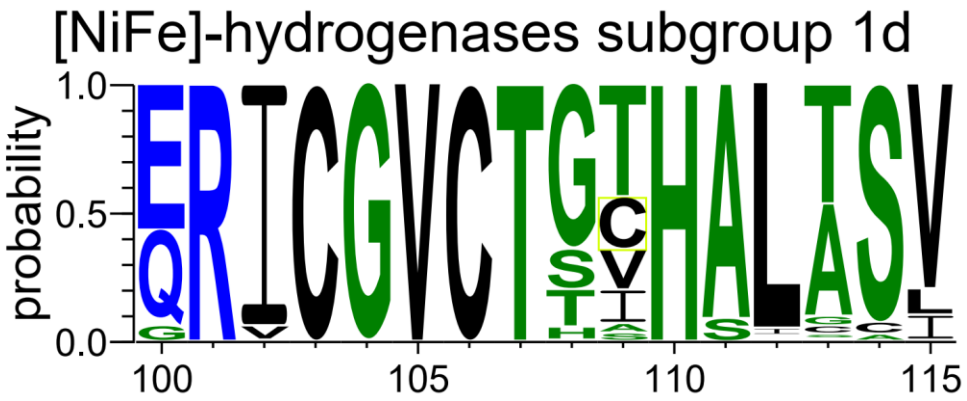


Figure S5. Sequence logo for the large subunit region around Cys81 of *Cn*MBH derived from the multiple sequence alignment of [NiFe]-hydrogenases belonging to subgroup *1d*. The multiple sequence alignment (provided as additional supporting material) was generated using Clustal Omega (https://www.ebi.ac.uk/jdispatcher/msa/clustalo)[9] and comprise 216 protein sequences retrieved using the web database HydDB for hydrogenase classification.[10] The figure was generated with the WebLogo3 online too (version 3.7.12, https://weblogo.threeplusone.com/create.cgi).[11] The size of letters is directly related to the abundance of the respective amino acid residue in all hydrogenase sequences from subgroup *1d*. Two of the four cysteines coordinating the NiFe site are part of the fully conserved CGVC motif (ruler positions 103/106). The yellow square indicates that cysteine (ruler position 109) is the 2nd more frequent amino acid found at position 81 of *Cn*MBH. Note that the residue numbering on x axis differs from that of *Cn*MBH. About 70% of all sequences of hydrogenases belonging to subgroup *1d* contain a polar residue (Cys, Thr and Ser), and all others contain small apolar amino acids whose function might be rescued by H_2O molecules.

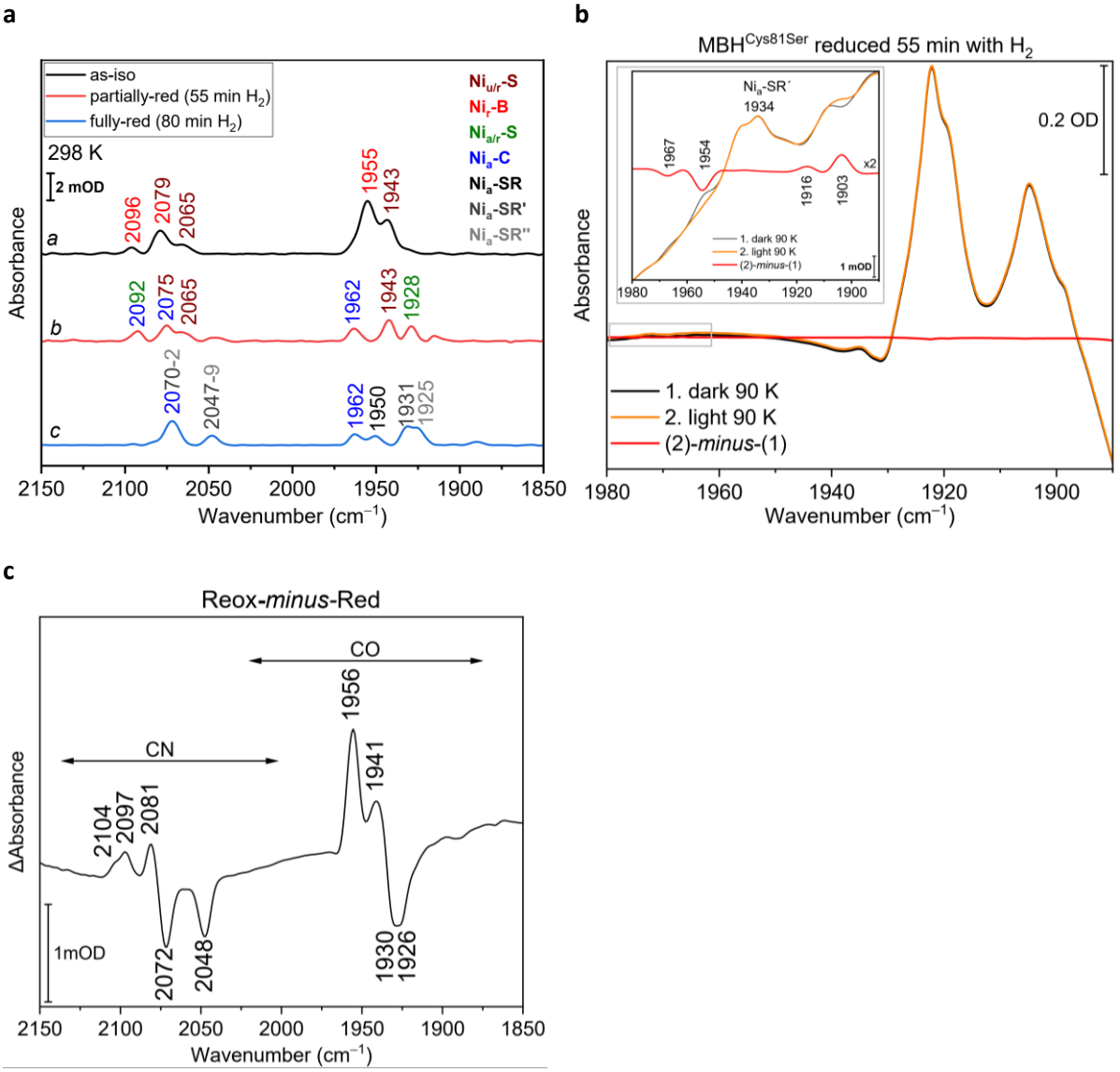


Figure S6: IR spectroscopic characterization of the *Cn*MBH^{Cys81Ser} variant. (**a**) IR spectra at 298 K of the as-isolated (oxidized) *Cn*MBH^{Cys81Ser} (ca. 1 mM in 50 mM K_iPO_4, pH 5.5, 150 mM NaCl, 50 % glycerol (trace *a*) and the corresponding reduced samples after incubation with 100 % H_2 either for 55 min (trace *b*) or 80 min (trace *c*). The as-isolated MBH^{Cys81Ser} resides in a mixture of Ni_r-B and Ni_u-S species. After 55 min of reduction, Ni_a-C and $Ni_{a/r}$-S states were detected together with residual Ni_u-S species. After prolonged incubation (80 min) with H_2 the IR spectra of *Cn*MBH^{Cys81Ser} show predominant Ni_a-SR' signals, similarly to native *Cn*MBH (**Fig. S2**). (**b**) IR spectra of H_2-reduced *Cn*MBH^{Cys81Ser} at 90 K before and after LED light irradiation at 460 nm show predominantly Ni_a-SR' (ν_{CO} band at 1934 cm^{-1}), accompanied by the formation of sub-stoichiometric amounts of Ni_a-C. For the Ni_a-C→Ni_a-L phototransformation, the partially reduced sample (trace *b* in **Fig. S6a**) was used because it contained the highest proportion of the Ni_a-C state. Illumination of the MBH^{Cys81Ser} sample with LED at 460 nm (90 K) resulted in the production of two distinct Ni_a-L species (ν_{CO} bands at 1916 and 1903 cm^{-1}). (**c**) IR difference spectrum (Reox-*minus*-Red) of the *Cn*MBH^{Cys81Ser} variant at 298 K, showing that the enrichment of reduced active site species is reversible as the reduced sample incubated with air returned to the original Ni_r-B (ν_{CO} at 1956 cm^{-1}) and Ni_u-S (ν_{CO} at 1941 cm^{-1}) states. Bands assignment follows those described in Saggu et *al*..[4]

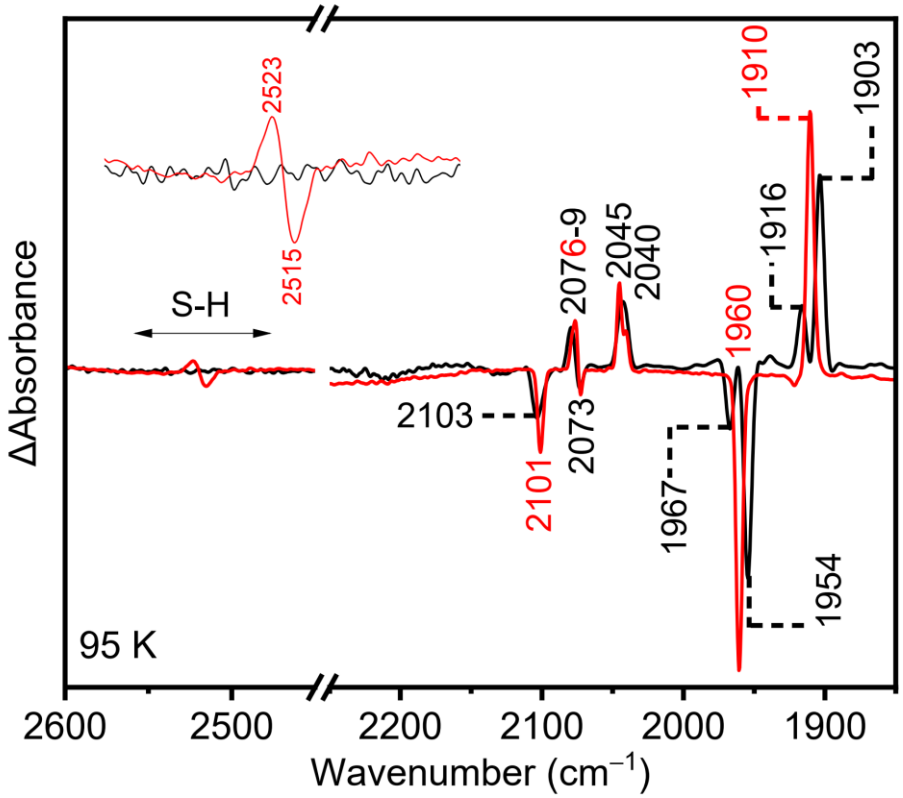


Fig. S7. **Comparison of the IR difference spectra of native *Cn*MBH and the *Cn*MBH^{Cys81Ser} variant.** Difference spectra (light-*minus*-dark, Ni_a-L-*minus*-Ni_a-C) of native *Cn*MBH (red trace) and the *Cn*MBH^{Cys81Ser} variant (black trace) (both at ca. 1 mM in 50 mM K_iPO_4, pH 5.5, 150 mM NaCl, 50 % glycerol) were recorded at 95 K, displaying the characteristic Ni_a-L and Ni_a-C absorptions as positive and negative bands, respectively. The inset in the upper left shows an enlargement of the SH spectral range. The IR data for *Cn*MBH^{Cys81Ser} show a lower signal-to-noise ratio compared to those of native *Cn*MBH, which is attributed to a lower Ni_a-C enrichment in the MBH variant (**Fig. S6a**, trace b vs **Fig. S2**). The exchange of Cys81 induces a splitting of the original ν_{CO} bands in native MBH at 1960 (Ni_a-C) and 1910 cm^{-1} (Ni_a-L), which now appear as two negative ν_{CO} absorptions at 1954 cm^{-1} and 1967 cm^{-1} (Ni_a-C) photoconverting to species with ν_{CO} at 1903 and 1916 cm^{-1} (see Supplementary Discussion). The exchange of Cys81 with Ser additionally results in the disappearance of the observed ν_{SH} bands, which aided their assignment to the protonated side chain of Cys81. Data of native *Cn*MBH were normalized to those of the *Cn*MBH^{Cys81Ser} variant (scaling factor 0.4) for better visualization.

Supplementary Discussion

The replacement of Cys81 by Ser resulted in the splitting of the ν_{CO} band of both the Ni_a-C and Ni_a-L states. In fact, the MBH variant exhibits two negative ν_{CO} absorptions at 1954 cm^{-1} and 1967 cm^{-1}, which are photoconverted to bands at 1903 and 1916 cm^{-1}. In contrast, the corresponding ν_{CN} bands remain essentially unaffected (**Fig. S7**). The O−H bond length (ca 1.0 Å) of serine is shorter than the S−H bond (ca 1.4 Å) of cysteine, which might prevent the formation of a H−bond between serine and the thiolate of Cys78 and/or the CO ligand. This, in turn, possibly led to the observed splitting of the CO band. We assigned the minor Ni_a-C and Ni_a-L species with ν_{CO} bands at higher energies (1967 and 1916 cm^{-1}) to active site arrangements in which serine adopts an identical conformation as Cys81 (Ni_a-C: 1960 cm^{-1}, Ni_a-L1/2: 1910 cm^{-1}, **Fig. 3c**) and thus forms an H−bond. Regarding the upshift of the ν_{CO} bands, this is related to the higher polarity of serine which weakens the back-bonding from the iron and strengthens the CO bond compared to that of native MBH. This assignment is in line with previous investigations on other active site species of the $MBH^{Cys81Ser}$ variant.[12] Conversely, we attributed the more abundant species with lower energy (ν_{CO} bands at 1954 and 1903 cm^{-1}) to Ni_a-C and Ni_a-L active site states without H−bonds, which are characterized by a higher electron density at the NiFe site that weakens the CO bonds. Importantly, temperature appears to influence the relative abundance of the two active site populations. Room temperature IR data predominantly show the high-energy species for the MBH variant (**Fig. S6**),[4] while data collected at lower temperatures (**Fig. S7**) favor the low-energy species.

Supplementary references

1 T. Schubert, O. Lenz, E. Krause, R. Volkmer and B. Friedrich, *Molecular Microbiology*, 2007, **66**, 453–467.
2 M. Ludwig, J. A. Cracknell, K. A. Vincent, F. A. Armstrong and O. Lenz, *J. Biol. Chem.*, 2009, **284**, 465–477.
3 O. Lenz, L. Lauterbach and S. Frielingsdorf, in *Methods in Enzymology*, Elsevier, 2018, vol. 613, pp. 117–151.
4 M. Saggu, M. Ludwig, B. Friedrich, P. Hildebrandt, R. Bittl, F. Lendzian, O. Lenz and I. Zebger, *ChemPhysChem*, 2010, **11**, 1215–1224.
5 M. Saggu, I. Zebger, M. Ludwig, O. Lenz, B. Friedrich, P. Hildebrandt and F. Lendzian, *J. Biol. Chem.*, 2009, **284**, 16264–16276.
6 A. F. T. Waffo, C. Lorent, S. Katz, J. Schoknecht, O. Lenz, I. Zebger and G. Caserta, *J. Am. Chem. Soc.*, 2023, **145**, 13674–13685.
7 J. Fritsch, P. Scheerer, S. Frielingsdorf, S. Kroschinsky, B. Friedrich, O. Lenz and C. M. T. Spahn, *Nature*, 2011, **479**, 249–252.
8 X. Robert and P. Gouet, *Nucleic Acids Research*, 2014, **42**, W320–W324.
9 F. Madeira, N. Madhusoodanan, J. Lee, A. Eusebi, A. Niewielska, A. R. N. Tivey, R. Lopez and S. Butcher, *Nucleic Acids Research*, 2024, **52**, W521–W525.
10 D. Søndergaard, C. N. S. Pedersen and C. Greening, *Sci Rep*, 2016, **6**, 34212.
11 T. D. Schneider and R. M. Stephens, *Nucl Acids Res*, 1990, **18**, 6097–6100.
12 Y. Rippers, T. Utesch, P. Hildebrandt, I. Zebger and M. A. Mroginski, *Phys. Chem. Chem. Phys.*, 2011, **13**, 16146.