An Unexpected P-Cluster like Intermediate *En Route* to the Nitrogenase FeMo-co.

Leon P. Jenner[a], Mickael V. Cherrier[a], Patricia Amara[a], Luis M. Rubio[b*] and Yvain Nicolet[a*]

[a]*Univ. Grenoble Alpes, CEA, CNRS, IBS, Metalloproteins Unit, F-38000 Grenoble, France.*
[b]*Centro de Biotecnología y Genómica de Plantas, Universidad Politécnica de Madrid, Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria, Pozuelo de Alarcón, 28223 Madrid, Spain.*

Supporting Information

Original structure analysis

The structure of *Mth*NifB from Kang et al. (PDB 7JMB)[1] features residues 13-276 in two chains A and B with R_{work} and R_{free} values of 0.262 and 0.314 respectively. The monomers of this crystallographic dimer were modeled nearly identically and had the same general ¾ β-barrel structure as *Mt*NifB (PDB 6Y1X) common to radical SAM enzymes.[2,3] The two chains were linked by a disulfide bond between C^{240} residues. An [Fe_4S_4] RS-cluster binds in the approximate center of the β-barrel *via* C^{38}, C^{42} and C^{45} (Fig. S1, right). Two more [Fe_4S_4] clusters, K1 and K2, bind to the 'bottom' of the β-barrel and to a loop formed from N-terminus, respectively. The C-terminus was ordered in a series of loops that wrapped both K-clusters, bringing them into close proximity. K1 was modeled with C^{18}, H^{31} and C^{115} as ligands and K2 with C^{260} and C^{263} as ligands (Fig. S1, left).
All clusters were modeled as perfect cubes forcing Fe coordination out of the expected tetrahedral geometry and giving rise to unlikely bond angles (shown in Fig. S1 as black dotted lines). Furthermore, H^{24} and H^{31} proposed as K-cluster ligands from X-ray spectroscopy and density functional theory analysis were modeled too far from the iron atoms of the cluster for them to be actually bonded (Fig. S1, green dotted lines).

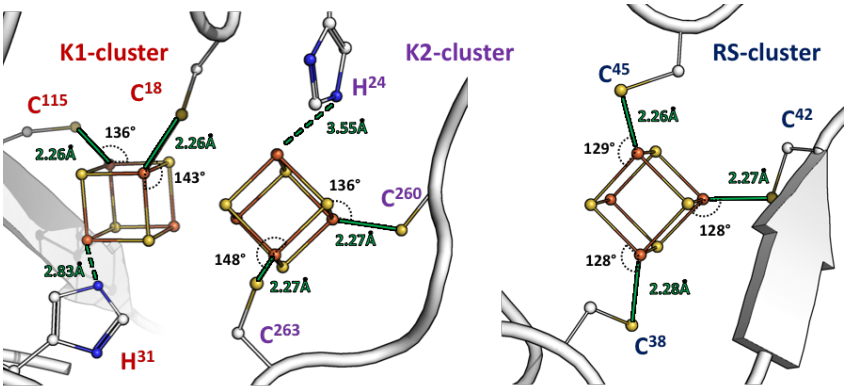

Figure S1 – Proposed MthNifB iron sulfur cluster ligation in Kang et al. *Clusters and ligands (as annotated) have a ball-and-stick representation with atoms colored by element (Fe brown, S yellow, N blue, C white). Interatomic distances shown as dashed green lines, solid where bonds are modeled and selected angles shown as dotted black lines. A typical Fe-Cys bond length is 2.30 Å, Fe-His is 2.25 Å and Fe coordination should approach tetrahedral symmetry with bond angles of 109.5°.*

Improved processing of the crystallographic data for *Mth*NifB

For the present study, we obtained X-ray diffraction data and the initial *Mth*NifB crystal structure initial model from the RSCB Protein Data Bank (PDB code: 7JMB).[1] The structure was then refined using BUSTER software[4] and manual modifications were performed using Coot.[5] During the first refinement cycles, target structure reference restraints were applied using the crystal structure of *Mt*NifB (PDB code: 6Y1X).[6] Non-Crystallographic Symmetry

(NCS) restraints were also applied during refinement. [Fe_4S_4] and P-cluster ligand geometry restraints (SF4 and CLF, respectively) were obtained from the RCSB Protein Data Bank. Figures were prepared using PYMOL.[7]

As a control, using the newly refined protein structure model (Table S1, R_{work} = 20.19 %; R_{free} = 23.95 %), we replaced the K-cluster with two independent [Fe_4S_4] clusters. After 5 cycles of such refinement and using the same parameters as before, the refinement statistics slightly increased (R_{work} = 20.22 %; R_{free} = 24.13 %). In addition to displaying bad ligand coordination geometry, the two [Fe_4S_4] clusters also exhibit lower correlation with the electron density when compared to that of the K-cluster. The original and corrected models for this region are presented in Figure 2.

Our improved structure refinement also revealed two additional unexpected features. The first one is the presence of an extra density for residue A^{276} in both chains A and B (Fig. S4). The most likely interpretation of this electron density peak, considering its distance to the A^{276} C_β and its geometry, is that residue 276 would have been mutated to serine. However, we do not have access to the actual protein sample used for crystallization and we cannot conclude it with certainty. The second feature is the presence of an extra electron density bound to the NifB RS-cluster (Fig. S5). Our best interpretation for this peak in the electron density is a nitrate molecule from the buffer, establishing three hydrogen bonds with residues A^{87}, G^{90} and S^{117}. Even though this rationalizes the electron density, the presence of a nitrate molecule at this position is not entirely satisfactory, because no direct interactions are made with the RS-cluster. As the resolution is rather low, another interpretation cannot be excluded.

Table S1. Refinement statistics

Resolution range	60.84 – 3.0 (3.11 – 3.0)
Reflections used in refinement	12029 (794)
Reflections used for R_{free}	1207 (76)
R_{work}	0.2019 (0.3471)
R_{free}	0.2395 (0.4316)
Number of non-hydrogen atoms	4124
macromolecules	4054
ligands	59
solvent	11
Protein residues	550
RMSD (bonds) (Å)	0.049
RMSD (angles) (°)	1.69
Ramachandran favored (%)	91.76
Ramachandran allowed (%)	6.23
Ramachandran outliers (%)	2.01
Rotamer outliers (%)	10.86
Clashscore	13.95
Average B-factor (Å^2)	77.22
Macromolecules (Å^2)	77.02
Ligands (Å^2)	96.42
Solvent (Å^2)	45.39

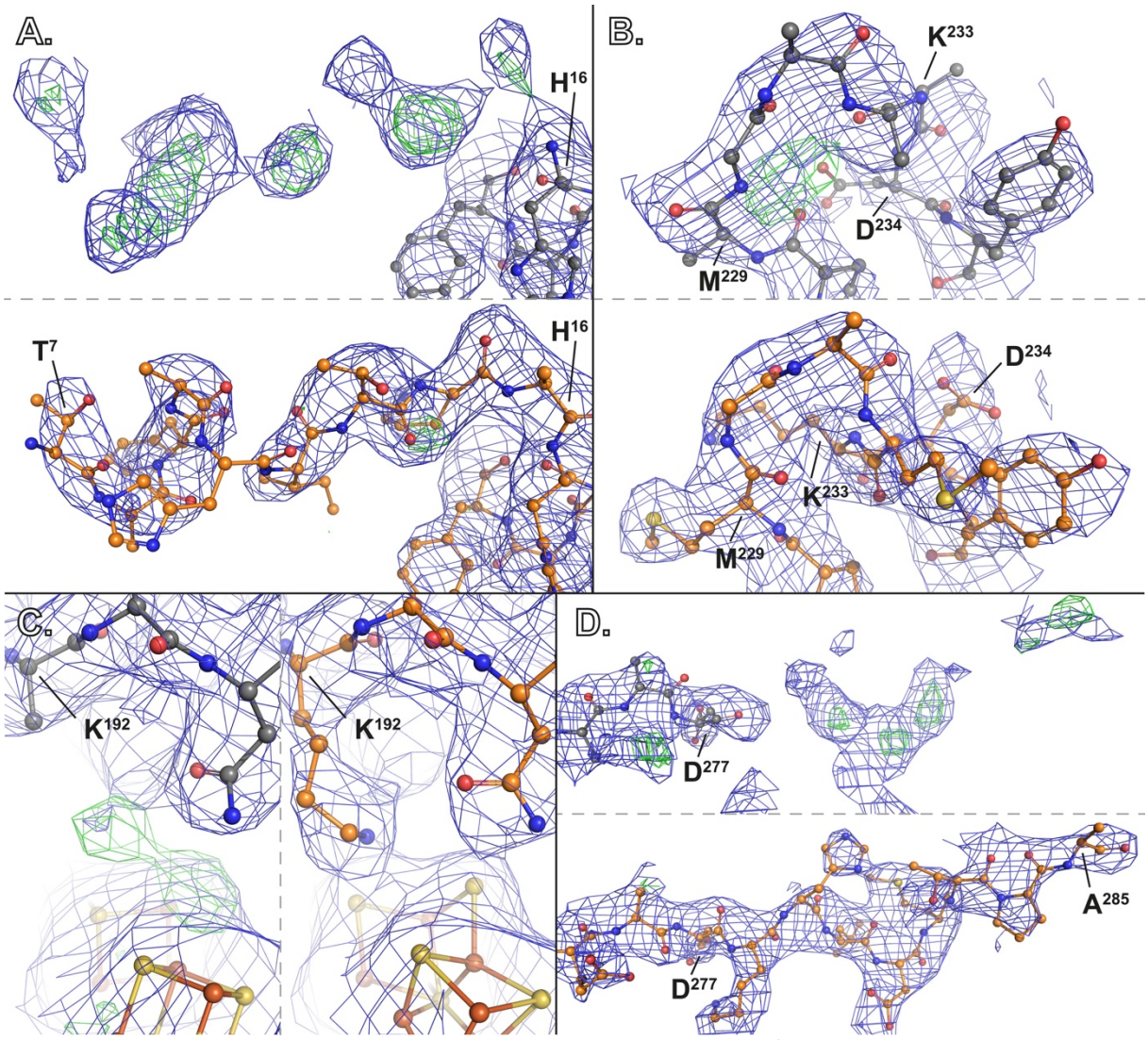

Figure S2. Revised MthNifB model using X-ray diffraction data from Kang et al.[1] ***A)*** *N-terminal stretch of chain A molecule.* ***B)*** *M^{229}-T^{234} loop in chain A.* ***C)*** *residue K^{192} and K-cluster from chain A.* ***D)*** *C-terminal stretch of chain A. The 2Fo-Fc Fourier and Fo-Fc difference Fourier electron density maps (blue and green meshes, respectively) were calculated using either PDB 7JMB (upper part in* ***A, B, D*** *and left part in* ***C****) or our revised model (lower part in* ***A, B,*** *and* ***D,*** *and right part in* ***C****) and were contoured at 1 and 3 σ, respectively. Here, we show only four representative examples of residual electron density that was either misinterpreted or not modeled in the 7JMB structure. Many others can be found throughout the structure.*

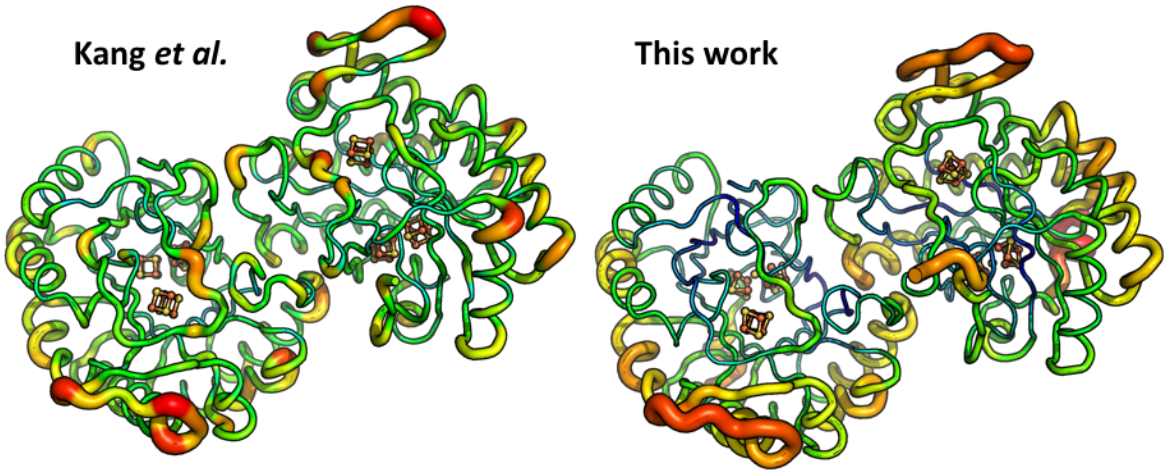

Figure S3. MthNifB models colored by residue B-factor. *Crystal structure from Kang et al. (left – 7JMB) and this work (right) with polypeptide chains shown as a cartoon tube with the size and color indicating the B-factor of the residues. Spatial spreading of the B-factor is a common criterion to assess the validity of a crystal structure, usually highlighting that the core structure is less agitated than the surface residues. Narrow, blue tubes represent low values, wide red tubes represent high values. The structure from Kang et al. (PDB 7JMB) displays B-factors ranging from 21.8 to 182.5 Å with an average B-factor = 70.8 Å. The reinvestigated structure (this work) displays B-factors ranging from 27.6 to 152.6 Å with an average B-factor = 74.9 Å. The B-factor determined from the Wilson-plot is 78.8 Å. Iron-sulfur clusters have a ball-and-stick representation. The structure from this work has low B-factors in the core and the surrounding of the K-cluster binding site, where the structure is expected to be rigid, progressing to higher factors in solvent exposed loops, which are expected to be flexible. B-factors in Kang et al. vary considerably between neighboring residues in a way that is hard to rationalize.*

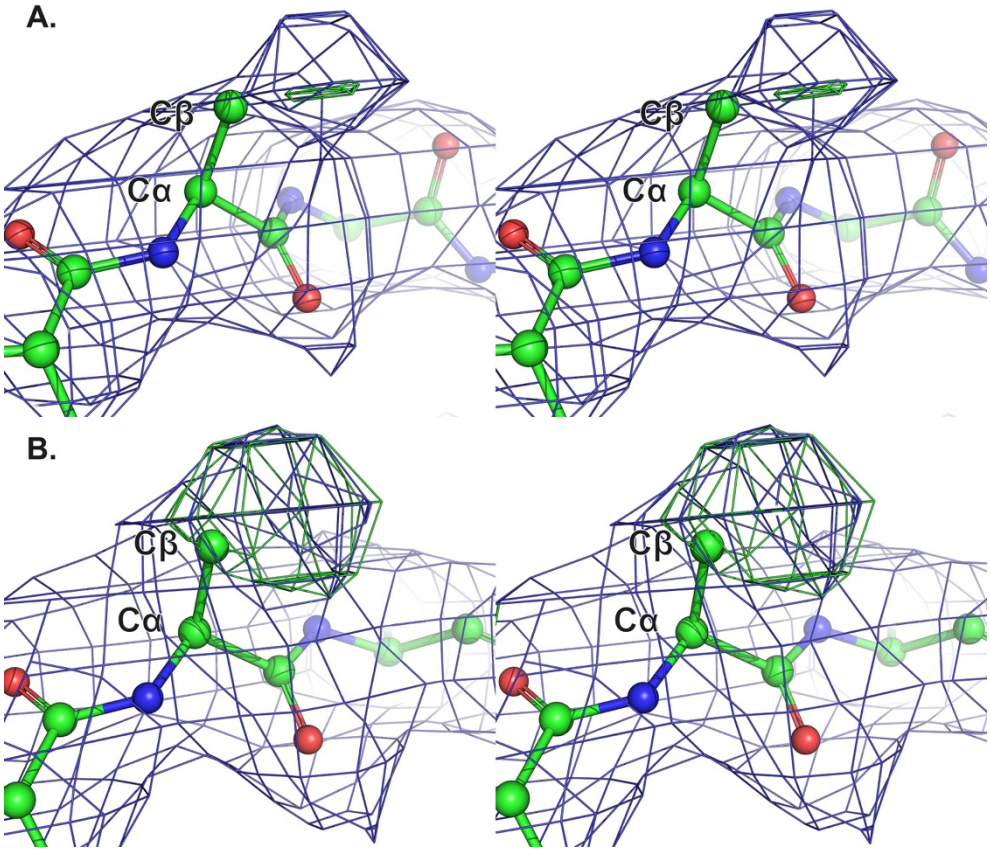

Figure S4. Electron density of the residue 276. *Stereo representation of the 2Fo-Fc (blue, 1 σ) and Fo-Fc (green, 3 σ) electron density map for residue A^{276} for (A) chain A and (B) chain B.*

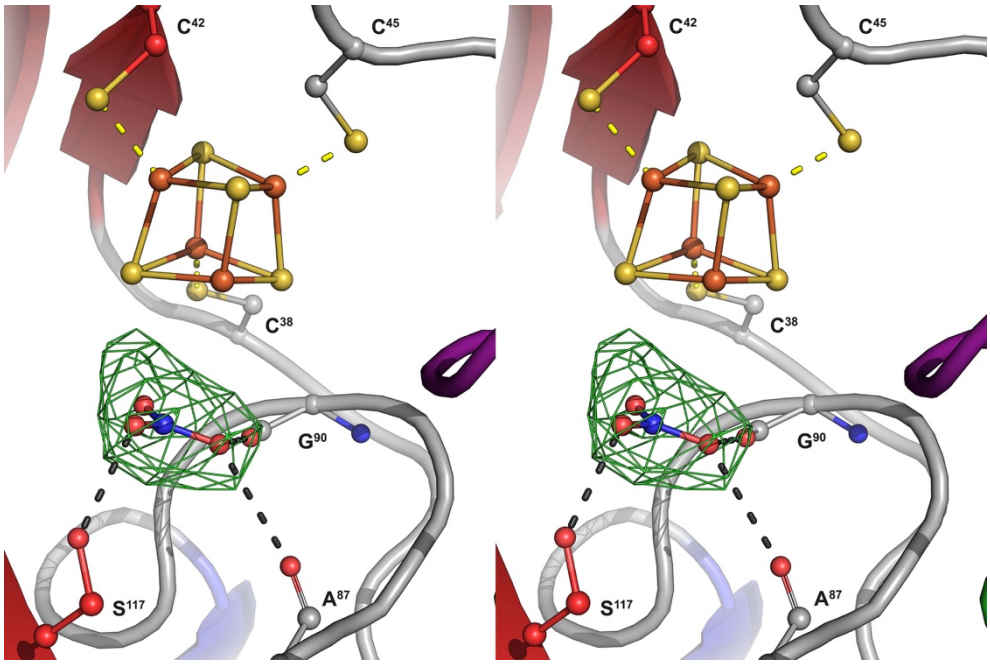

Figure S5. Stereoview of the residual electron density around the RS-cluster*. Fo-Fc difference Fourier electron density map (omit map) around the RS-cluster contoured at 3 σ. Hydrogen bonds are shown as dashed black lines. Only chain A is represented because no extra electron density peak is visible for chain B.*

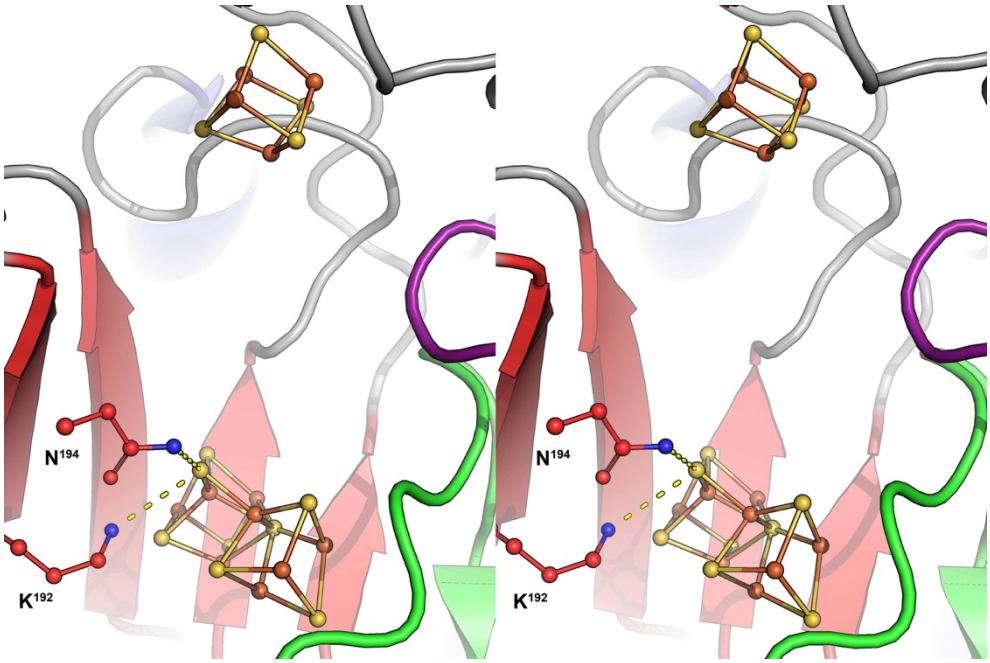

Figure S6. Stereoview of the belt-sulfide environment.

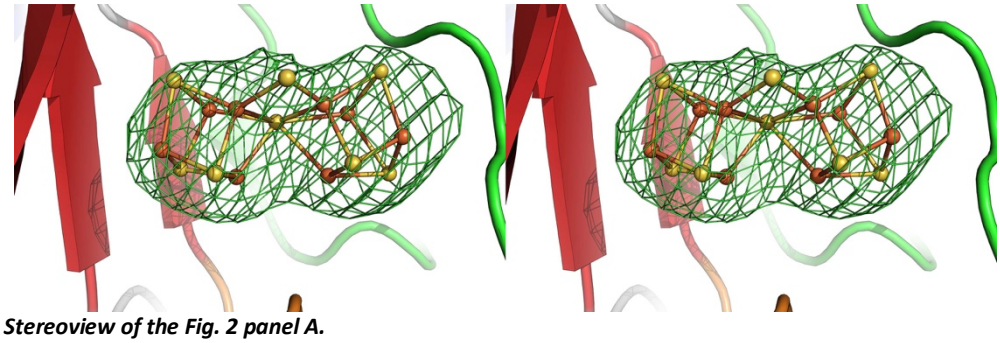

Figure S7. Stereoview of the Fig. 2 panel A.

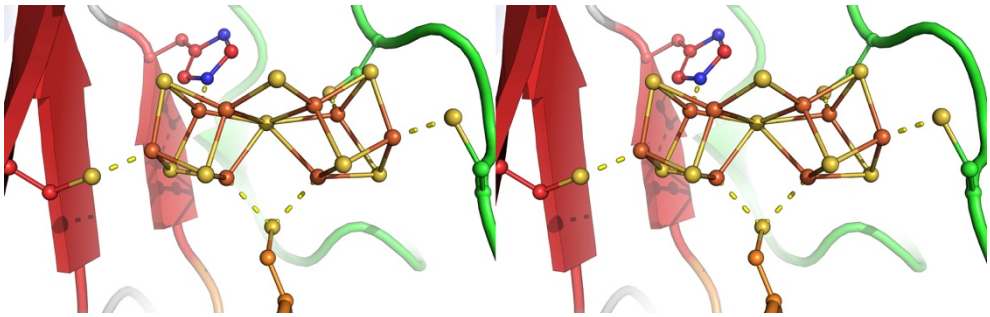

Figure S8. Stereoview of the Fig. 2 panel B.

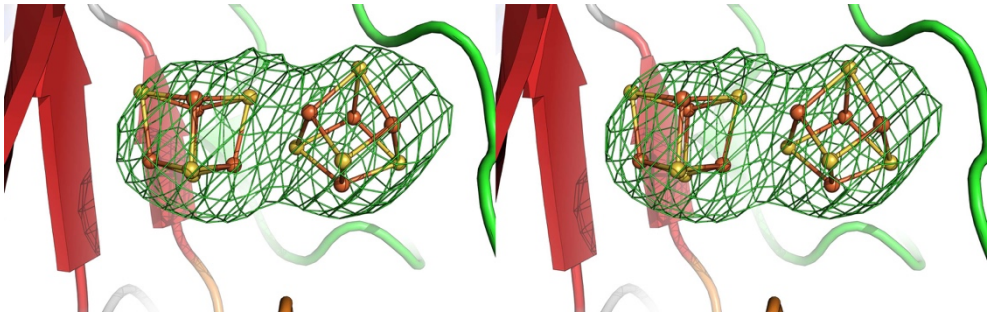

Figure S9. Stereoview of the Fig. 2 panel C.

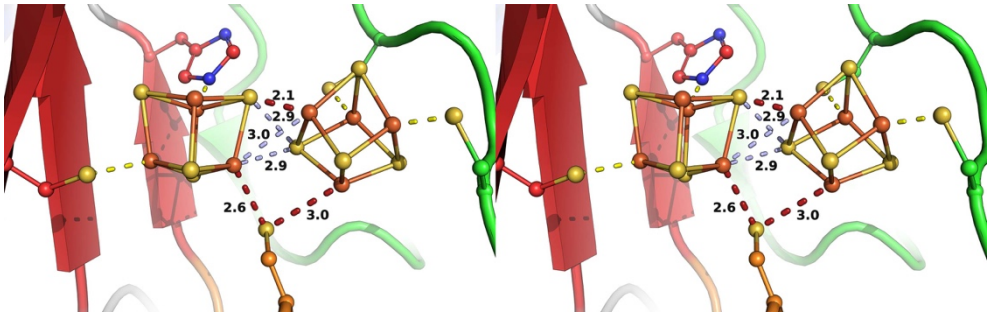

Figure S10. Stereoview of the Fig. 2 panel D.

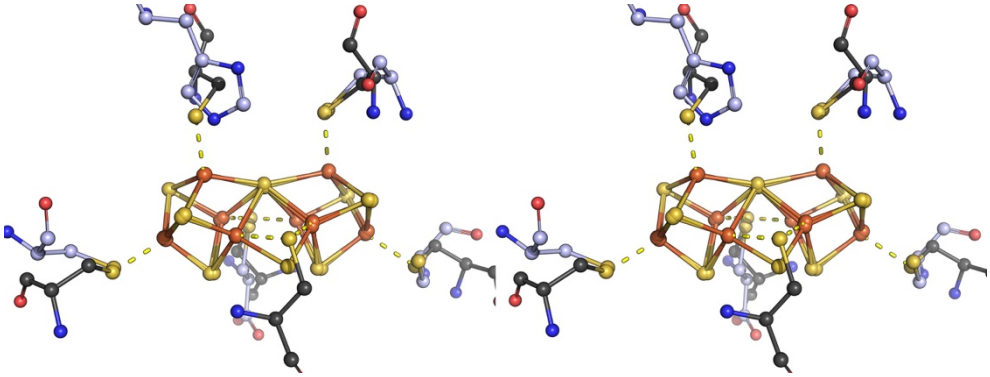

Figure S11. Stereoview of the Fig. 3.

Molecular docking

Method

The calculations to find alternative SAM-binding modes to NifB using molecular docking were performed with the Schrödinger suite[8] using the OPLS3e force field.[9] Starting from our revised crystallographic *Mth*NifB model, hydrogen atoms were built and protonation states were optimized with the *Protein preparation wizard*. Given the relatively low-resolution structure, only hydrogen atoms were geometry-optimized and formal charges were employed for the iron-sulfur cluster ions. The RS-cluster was given a total charge of +2 (a +1 reduced cluster gives essentially the same results, *not shown*) and the K-cluster was assumed to be fully reduced (total charge of 0 including the belt-sulfide ion). Water molecules were removed from the final model. We calculated a receptor grid centered at the centroid of residues F^{44}, T^{141}, T^{227} and the RS-cluster. The SAM ligand was prepared using *Ligprep*; the standard protonation state of the methionine fragment (NH_2, CO_2^-) observed in radical SAM enzymes (*i.e.* when bound to the RS-cluster for SAM cleavage) was chosen. To enhance sampling, 300 conformers were obtained with *Macromodel*. Then, all conformations were docked on the receptor grid using the SP and XP procedures in *Glide*.[10] Only poses where the methionine fragment was bound to the RS-cluster unique Fe ion were retained. These poses were sorted according to the XP docking score (from -16.0 to -10.9).

Results

The SAM-binding mode observed in radical SAM enzymes for SAM cleavage was not found in the docked poses; this could be due to the low structure resolution or, most probably, to a missing ion and structural water molecules present in the SAM-bound 22.3%-homologous radical SAM PFL-AE.[11]

Out of 6450 poses, 3100 had the methionine carboxylate group bound to the unique Fe ion, out of which only 80 poses had the methionine NH_2 group bound too. The majority of the poses had the adenine bound as radical SAM enzymes do (Fig. S12) and all of them had the hydrogen bond between the ribose O3' and the conserved T^{141} residue as observed with the conserved Ser126 in MoaA (see main text).[12] In contrast, the ribose O2' does not establish any hydrogen bonds with the receptor. This most probably explains the ribose moiety flexibility (see main text). In half of the poses, the belt-sulfide ion is located between 3.3 and 3.8 Å from the SAM methyl group (Fig. S12B).

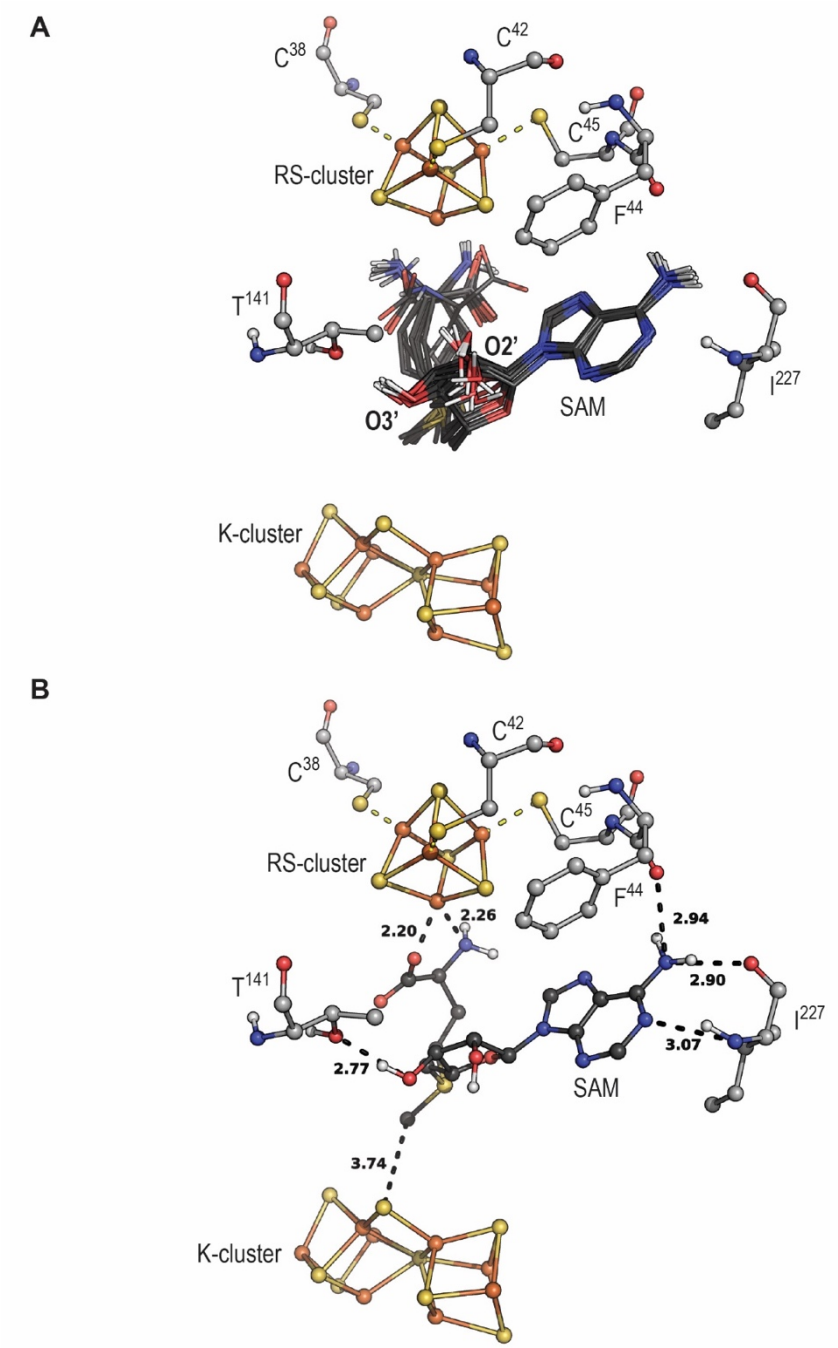

Figure S12. A) *The 20 poses of SAM docked in our MthNifB model with the highest score are represented as sticks. Key residues for SAM binding and both the K- and the RS- clusters are represented as ball-and-sticks. The hydroxyl-group oxygen atoms of the ribose moiety are labeled in bold.* ***B)*** *Representation of SAM-binding mode to MthNifB with hydrogen bonds shown as dashed lines. Heavy atom distances are given in Å. This pose corresponds to the highest docking score with the criteria we have chosen (see text). C, N, O, S, Fe and H atoms are colored in grey (dark grey for SAM), blue, red, yellow, orange and white, respectively.*

References

1 W. Kang, L. Rettberg, M. Stiebritz, A. Jasniewski, K. Tanifuji, C. C. Lee, M. W. Ribbe and Y. Hu, *Angew. Chem. Int. Ed Engl.*, DOI:10.1002/anie.202011367.

2 A. S. Fajardo, P. Legrand, L. A. Payá-Tormo, L. Martin, M. T. Pellicer Martı Nez, C. Echavarri-Erasun, X. Vernède, L. M. Rubio and Y. Nicolet, *J. Am. Chem. Soc.*, 2020, **142**, 11006–11012.

3 Y. Nicolet, *Nat. Catal.*, 2020, **3**, 337–350.

4 G. Bricogne, E. Blanc, M. Brandl, C. Flensburg, P. Keller, W. Paciorek, P. Roversi, A. Sharff, O. S. Smart, C. Vonrhein and T. O. Womack, BUSTER Version 2.10.3. Cambridge, United Kingdom: Global Phasing Ltd. 2017.

5 P. Emsley, B. Lohkamp, W. G. Scott and K. Cowtan, *Acta Crystallogr. Sect. -Biol. Crystallogr.*, 2010, **66**, 486–501.

6 A. S. Fajardo, P. Legrand, L. A. Payá-Tormo, L. Martin, M. T. Pellicer Martı Nez, C. Echavarri-Erasun, X. Vernède, L. M. Rubio and Y. Nicolet, *J. Am. Chem. Soc.*, 2020, **142**, 11006–11012.

7 The PyMOL Molecular Graphics System, Version 2.4.0 Schrödinger, LLC.

8 Schrödinger Release 2019-2: Maestro, Glide, Ligprep, Macromodel, Schrödinger, LLC, New York, NY, 2019.

9 E. Harder, W. Damm, J. Maple, C. Wu, M. Reboul, J. Y. Xiang, L. Wang, D. Lupyan, M. K. Dahlgren, J. L. Knight, J. W. Kaus, D. S. Cerutti, G. Krilov, W. L. Jorgensen, R. Abel and R. A. Friesner, *J. Chem. Theory Comput.*, 2016, **12**, 281–296.

10 R. A. Friesner, J. L. Banks, R. B. Murphy, T. A. Halgren, J. J. Klicic, D. T. Mainz, M. P. Repasky, E. H. Knoll, M. Shelley, J. K. Perry, D. E. Shaw, P. Francis and P. S. Shenkin, *J. Med. Chem.*, 2004, **47**, 1739–1749.

11 J. L. Vey, J. Yang, M. Li, W. E. Broderick, J. B. Broderick and C. L. Drennan, *Proc. Natl. Acad. Sci. U. S. A.*, 2008, **105**, 16137–16141.

12 P. Hänzelmann and H. Schindelin, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, **101**, 12870–12875.