Electronic Supplementary Material (ESI) for ChemComm.

Electronic Supplementary Information

Structural evidence for a reaction intermediate mimic in the active site of a sulfite dehydrogenase

Ahmed Djeghader,[a] Mélanie Rossotti,[b] Saleh Abdulkarim,[a,c] Frédéric Biaso,[b] Guillaume Gerbaud,[b] Wolfgang Nitschke,[b] Barbara Schoepp-Cothenet,[b*] Tewfik Soulimane,[a*] and Stéphane Grimaldi[b*]

---

[a] *Department of Chemical Sciences and Bernal Institute, University of Limerick, Ireland. Email: Tewfik.soulimane@ul.ie*
[b] *Aix Marseille Univ, CNRS, BIP UMR7281, Marseille, France. Email : schoepp@imm.cnrs.fr, stephane.grimaldi@univ-amu.fr*
[c] *Present address: Azzaytuna University, Faculty of Science, Biology Department, Tarhuna City, Libya.*

METHODS

Expression, purification and crystallization of *Tt*SDH

Expression of *Tt*SDH from *Thermus thermophilus* HB8 strain was conducted in *E. coli* TP1000 cells transformed with pQE-60-SO2 plasmid [1]. Cells were grown in ZY-5052 media (with ampicillin 100 µg/ml and kanamycin 25 µg/ml) at 37 °C. At OD=0.7, 1 mM of Na_2MoO_4 was added and the temperature was lowered to 17 °C. After overnight incubation, cells were harvested by centrifugation at 6000 x *g* for 20 min at 4 °C and the pellet was resuspended in lysis buffer containing 50 mM Tris, 500 mM NaCl, pH 8.0 supplemented with lysozyme (0.25 mg/ml) and Dnase I (50 µg/ml) and frozen at -80 °C for 2 hours. Subsequently, the cells were thawed, sonicated and the soluble fraction collected by centrifugation at 17000 x *g* for 30 min. The soluble fraction was filtered and then loaded onto a Ni^{2+} superflow immobilized affinity chromatography column pre-equilibrated with 50 mM Tris, 500 mM NaCl, 10 mM imidazole, pH 7.6. Elution was achieved by applying a step-wise gradient of imidazole from 0-500 mM. Fractions containing *Tt*SDH were pooled and dialyzed overnight against a buffer containing 50 mM Tris, 500 mM NaCl, pH 8.0 supplemented with 1 mM Na_2MoO_4. After dialysis, *Tt*SDH was concentrated and applied to a Hiload superdex 200 gel filtration column and eluted using a buffer containing 50 mM Tris, 150 mM NaCl, pH 8.0. Fractions containing *Tt*SDH were concentrated to 48 mg/ml, fast-frozen in liquid nitrogen and stored at -80 °C. For crystallization assays, 1 µl of protein (48 mg/ml) was mixed with 1 µl of crystallization solution on sitting drop vapor-diffusion plates. Diffraction quality crystals were obtained in a condition consisting of 1.8 M sodium phosphate monobasic monohydrate, potassium phosphate dibasic, pH 5.0. Crystals typically appeared after one to two days. To enable cryogenic data collection, 2 µl of the crystallization solution was mixed with 2 µl of 20 % glycerol, and 2 µl of this mixture was then added to the crystallization drops. Following a short equilibration period, single crystals were harvested using Molecular Dimensions MicroMesh loops. Loops containing the crystals were then plunged into liquid nitrogen.

Structure determination

Datasets were collected at Diamond I24 beamlines on a PILATUS 6M detector. Data were processed using XDS [2] and high resolution cut-off were selected according to Karplus & Diederichs [3]. Phasing was achieved through molecular replacement (phenix.MRage/Phaser)[4] using multiple search models. The best solution was obtained with the *S. novella* SorA (PDB ID: 2BLF)[5] in the space group *P*6122. The partial model was automatically built using Phenix.AutoBuild [6], improved manually with Coot [7] and refined with phenix.refine [8] and Refmac.[9]

Sample preparation for EPR measurements

For preparation of the sample at 1.8 M Phosphate, samples of purified *Tt*SDH prepared as described above and stored at -80 °C at approximately 10 mg/mL (i.e. approximately 200 μM) were diluted in 1.8 M sodium phosphate monobasic monohydrate, potassium phosphate dibasic/50 mM Glycine/50 mM 2-(N-morpholino)ethanesulfonic acid pH 5.7 and passed through a desalting PD-10 column equilibrated with the same buffer. The obtained sample was prepared at 30 μM (or 120 μM for Q-band EPR experiments) with the same buffer in presence of 2 mM EDTA. The following mediators at 50 μM each were used: 1,4 *p*-benzoquinone, 2,5-dimethyl-*p*-benzoquinone, 2-hydroxy 1,2-naphthoquinone, 1,4-naphthoquinone, duroquinone, 2,3-dimethyl-p-benzoquinone, 2,5-dihydroxy-p-benzoquinone, dihydroxy-1,4-naphthoquinone, anthraquinone-2-sulfonate. The sample was equilibrated under argon atmosphere at room temperature. Reduction was obtained by addition of titanium citrate while oxidation was obtained by addition of potassium ferricyanide. A 150 μL (or 50 μL for Q-band EPR experiments) sample with maximal Mo(V) signal for EPR was retrieved from the redox poised bulk solution. For preparation of the sample at 70 mM phosphate, the same procedure was followed by using a 70 mM sodium phosphate monobasic monohydrate, potassium phosphate dibasic/50 mM Glycine/50 mM 2-(N-morpholino)ethanesulfonic acid pH 5.7.

CW EPR and HYSCORE experiments

X-band CW EPR spectra were measured on a Bruker EleXsys E500 spectrometer equipped with an ER4102ST standard rectangular Bruker EPR cavity fitted to an Oxford

Instruments helium flow cryostat. Q-band EPR spectra were recorded using a standard Bruker resonator (ER5106QT) equipped with an Oxford Instruments CF 935 cryostat. Spectra were measured at 70 K with 0.1 mT field modulation amplitude at 100 kHz using 1 mW (X-band) or 0.1 mW (Q-band) applied microwave power. For *g*-value measurements, the microwave frequency was measured by using a Hewlett-Packard HP1552B frequency counter, and magnetic field values were corrected against a known *g* standard (weak pitch, $g = 2.0028 \pm 0.0001$). Estimated errors on Mo(V) g-values are ± 0.0004.

Two-dimensional (2D) HYSCORE spectra were measured at 50 K using a Bruker EleXsys E580 spectrometer equipped with an ER4118X-MD5 dielectric resonator and an Oxford Instruments CF 935 cryostat. In this four-pulse experiment ($\pi/2$-τ-$\pi/2$-t_1-π-t_2-$\pi/2$-τ-echo), the intensity of the echo after the fourth pulse was measured with varied t_2 and t_1 and constant τ. A τ value of 200 ns was used for all spectra shown herein. The length of a $\pi/2$ pulse was 12 ns, and the length of a π pulse 24 ns. HYSCORE data were collected in the form of 2D time-domain patterns containing 256 × 256 points with steps of 12 ns. Spectra were recorded at the magnetic field value corresponding to the maximum intensity of the Mo(V) signal measured in a two-pulse field sweep electron spin-echo sequence ($\pi/2$-τ-π-τ-echo), and the measurements were accumulated during 19 h.
HYSCORE spectra were processed using Bruker's Xepr software as previously detailed.[10] Relaxation decays were subtracted (fitting by polynomial functions) followed by zero-filling and tapering with a Hamming window, before 2D Fourier transformation, which finally gives the spectrum in frequency domain. Processed data were then imported into Matlab (The MathWorks Inc., Natick, MA) for plotting. HYSCORE spectra are shown in absolute value mode and the (-, +) quadrant is presented as contour plot.

Spectral simulations

Numerical simulations of EPR and HYSCORE spectra were performed with the EasySpin package (release 5.0.12) using Matlab (The MathWorks, Inc., US).[11,12] X- and Q-band EPR spectra were simulated using the same set of parameters, summing the contribution of both field-independent (unresolved hyperfine couplings, H-strain) and field-dependent (g-strain) linewidth models. For H-strain, the full width at half-maximum of Gaussian

lines along the three principal axes was adjusted to (12, 10, 11) MHz whereas the corresponding uncorrelated g-strain distributions were (0.0037, 0.0020, 0.0018). In addition to the ^{31}P hyperfine interaction, the contribution of the 25% I = 5/2 ^{95,97}Mo isotopes was taken into account using a hyperfine interaction tensor with principal values (A_1, A_2, A_3) = (167, 80, 66) MHz and Euler angles (0°, 27°, 0°) between the principal axes of the hyperfine and g tensors (in the zyz convention). Moreover, an axially symmetric ^{13}C hyperfine tensor with (A_{iso}, T) = ($\pm$ 10, $\mp$ 1.6) MHz was considered in the simulations of HYSCORE spectra to account for the additional cross peaks seen in the Mo(V) HYSCORE spectra correlating nuclear frequencies with maxima around $\pm$8.4 MHz and $\mp$1.3 MHz, separated by twice the ^{13}C Zeeman frequency ($\nu_C \approx 3.8$ MHz). They are tentatively assigned to the C_γ atom of the coordinated Cys102 residue.

Molecular modelling on *Tt*SDH

A large computational model was built from the phosphate-bound crystal structure of *Tt*SDH. All residues within 12 Å of the molybdenum ion were included except those far from the phosphate ligand (*i.e.* residues 25-28, 53, 115, 203, 204, 263, 365-369). The two crystal water molecules near the phosphate ligand were also included in the model. Molybdenum was considered as Mo(V) whereas the pyranopterin was assumed to be in the tetrahydro- form and its phosphate group was substituted by an OH group for computing time reasons. The PROPKA software [13] was used on the phosphate-bound crystal structure to predict the protonated states of all the residues at pH 5.7. Thus, all Arg, Lys, and Glu were assumed to be charged, as well as the Cys102 ligand. Three classes of structural models were tested according to the protonation state of the phosphate ligand: model 1-$[HPO_4]^{2-}$ for the monoprotonated form HPO_4^{2-}, model 2-$[H_2PO_4]^-$ for the diprotonated form $H_2PO_4^-$ and model 3-$[PO4]^{3-}$ for the deprotonated form PO_4^{3-}. As illustrated in Supplementary Figure S4 for bound HPO_4^{2-}, this led to models with a total of 672-674 atoms, which included residues E29, T30, Y48, I49, R50, N51, N52, L100, Q101, C102, S103, G104, N105, G106, R107, N117, P118, W119, G122, G123, V124, F165, V166, R167, R197, G205, V206, N207, N208, V209, K210, T224, A225, E226, R230, Y231, N262, Y367. Moreover, the three different possible proton locations over the phosphate molecule were explicitly considered for models 1-$[HPO_4]^{2-}$ and 2-$[H_2PO_4]^-$ by running correspondingly separate sets of calculations.

Geometry optimization was subsequently performed with the Gaussian16 program, using the two-layer ONIOM (QM:QM') method.[14] In this approach, each model is split into two subsystems: (i) the high-level region consists of the Mo cofactor, the axial oxo ligand, the phosphate ligand and the Cys102 side chain, (ii) the low-level region includes the entire model. The high layer was treated at the DFT level using B3LYP functional [15,16] and def2SVP basis set with effective core potential for Mo.[17,18] The semi-empirical PM6 method was used for the low layer. All atoms were left unconstrained except for the alpha-carbons and the remaining phosphate oxygen of the pyranopterin moiety which were kept frozen during geometry optimization.

No energy minimum could be achieved in the following cases: (i) for model 3-$[PO_4]^{3-}$, possibly due to the large electrostatic repulsions between fully deprotonated phosphate and other anionic ligands, (ii) upon protonation of the phosphate oxygen close to Trp119 in models 1-$[HPO_4]^{2-}$ and 2-$[H_2PO_4]^-$, most likely due to electrostatic and steric constraints with the side chain of Trp119 and Arg107, and to a neighboring water molecule. Consequently, magnetic resonance parameters were calculated on the three remaining possible models (referred to as models 1a–$[HPO_4]^{2-}$, 1b–$[HPO_4]^{2-}$ and 2–$[H_2PO_4]^-$). In all models, where present, the phosphate proton close to Arg50 and Arg230 makes a strong H-bond (1.8 – 1.9 Å) with a water molecule whereas the one that points towards Arg107 does not make any H-bond whatever the model. For computational cost reasons, a cluster model containing ~128 atoms was cut-off from each ONIOM-optimized structure to perform these calculations. The three corresponding models are shown in Supplementary Figure S5. They keep the molybdenum cofactor with the oxo ligand, the phosphate ligand, the cysteine ligand, the guanidinium group of Arg50, Arg107 and Arg230, the peptide part of Ser103, Gly 205 and Val206, the Tyr231 side chain and the two water molecules. EPR parameters were calculated at the DFT level of theory with ORCA 4.1 quantum chemistry package. The hybrid functional B3LYP was used in conjunction with the D3BJ dispersion correction[19,20] and the def2-TZVPP basis set with ZORA for relativistic treatments. The spin-orbit mean-field Hamiltonian (SOMF) was used to account for the spin-orbit coupling. The DFT-calculated Mo(V) g-values for each model as well as the eigenvalues of the ^{31}P and ^{95}Mo hyperfine coupling tensors are compared to the corresponding experimentally-determined ones in Table S3.

Data availability

The structure has been deposited at the Protein Data Bank under the code 6YoK.

Table S1. Data collection and refinement statistics.

***Tt*SDH-PO4 (6YoK)**	
Resolution range (Å)	42.75 - 1.7 (1.761 - 1.7)
Space group	P 61 2 2
Unit cell dimensions (Å) **Unit cell angles (°)**	130.6 130.6 114.59 90 90 120
Total reflections	866328 (87574)
Unique reflections	63544 (6247)
Multiplicity	13.6 (14.0)
Completeness (%)	99.94 (100.00)
Mean I/sigma(I)	18.68 (2.67)
Wilson B-factor	25.29
R-merge	0.08164 (1.063)
CC1/2	0.999 (0.826)
Reflections used in refinement	63527 (6247)
Reflections used for R-free	1999 (196)
R-work (%)	0.1707 (0.2410)
R-free (%)	0.1887 (0.2637)
RMS (bonds) (Å)	0.013
RMS (angles) (°)	1.71
Ramachandran favored (%)	97.60
Ramachandran allowed (%)	2.13
Ramachandran outliers (%)	0.27
Rotamer outliers (%)	0.31
Clashscore	2.13
Average B-factor	29.06
Macromolecules	27.71
Ligands	33.06
Solvent	38.05

Statistics for the highest-resolution shell are shown in parentheses.

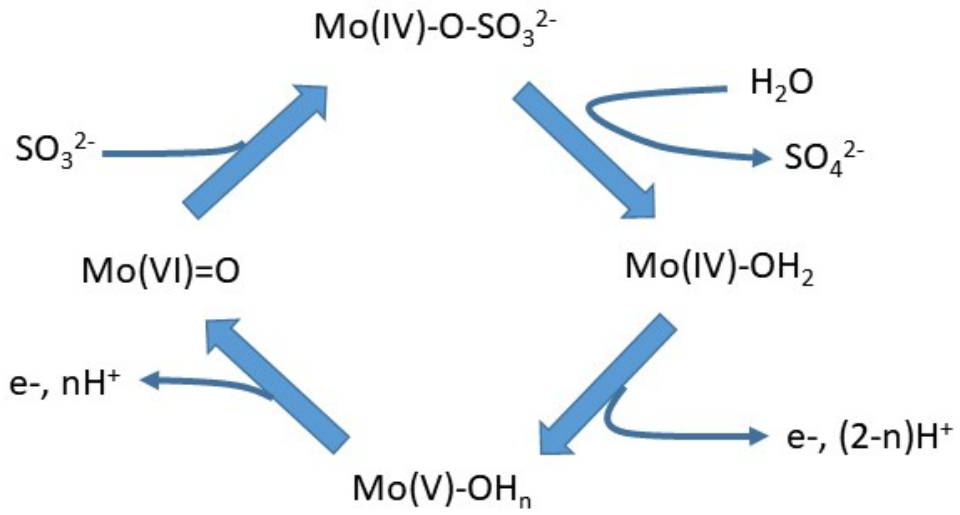

Fig. S1. General view of the Mo center oxidation state changes during SOEs' turnover. For the sake of simplification, only the Mo atom and the equatorial ligand are shown. n = 1 or 2. Adapted from [21].

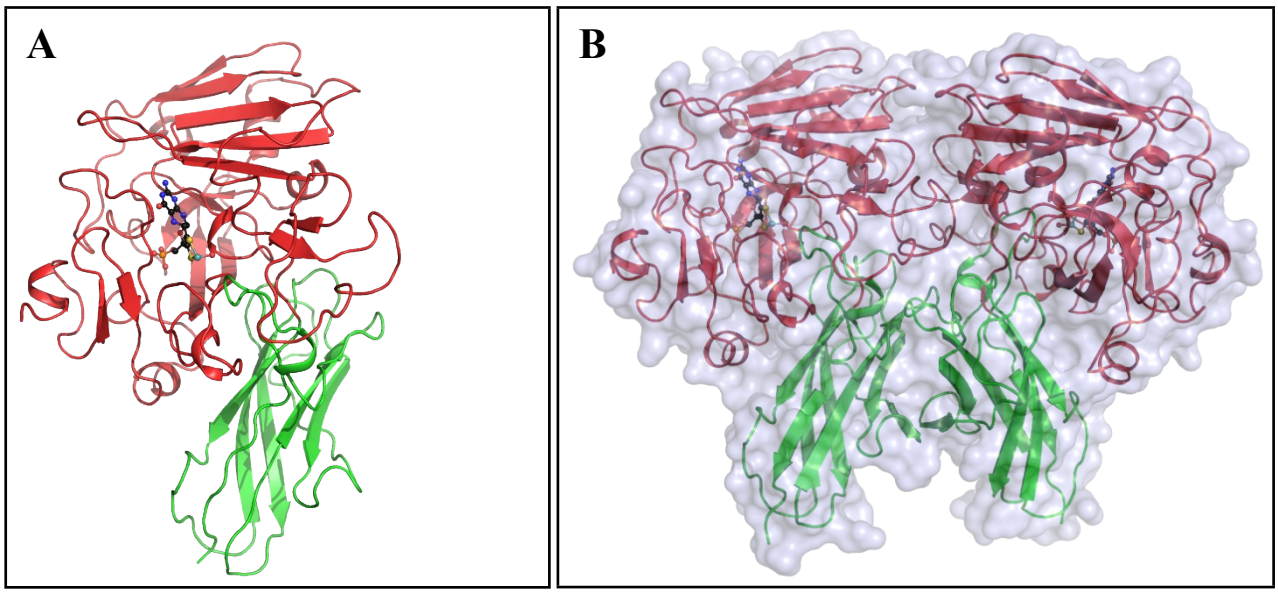

Fig. S2. Structure of *Tt*SDH. **A:** Cartoon representation of *Tt*SDH showing the classical organization of SOEs with a MoCo domain (red) and a dimerization domain (green). **B:** Dimeric organization of *Tt*SDH formed mainly by the interaction between the dimerization domains of each monomer. The Moco is shown in ball and stick representation.

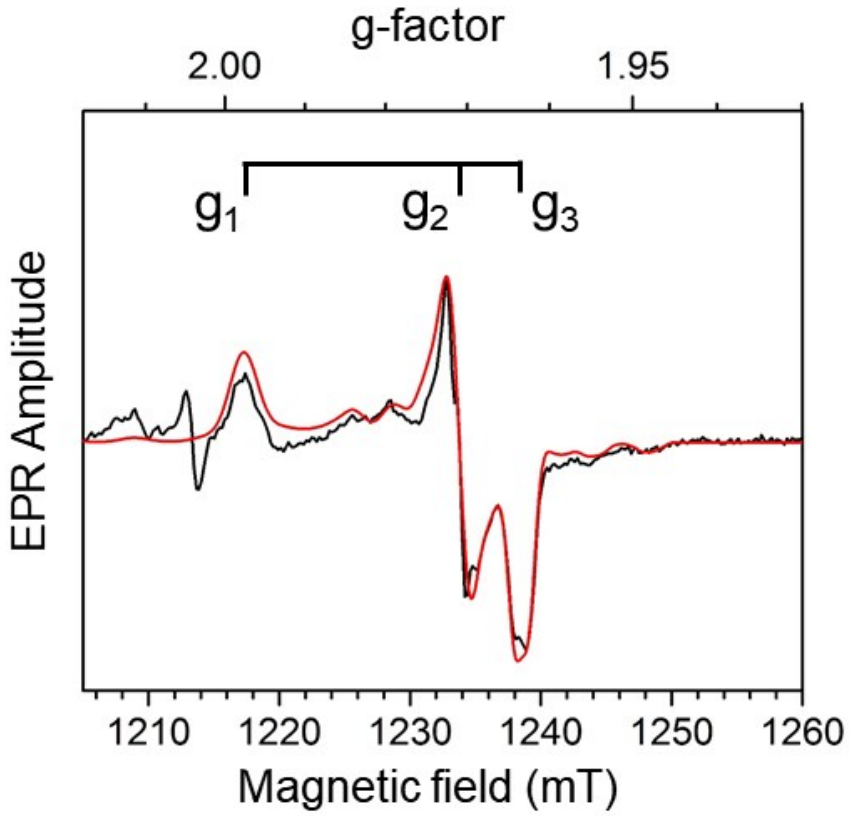

Fig. S3. Q-band (~34.0178 GHz) EPR spectrum of the Mo(V) species in *Tt*SDH incubated with 1.8 M sodium phosphate. The experimental spectrum (black trace) is shown together with its simulation (red trace) performed using (g_1, g_2, g_3) = (1.9967, 1.9700, 1.9624) whereas the other parameters including linewidths, ^{31}P and ^{95,97}Mo hyperfine tensors are provided in the ESI† text and Table S3.

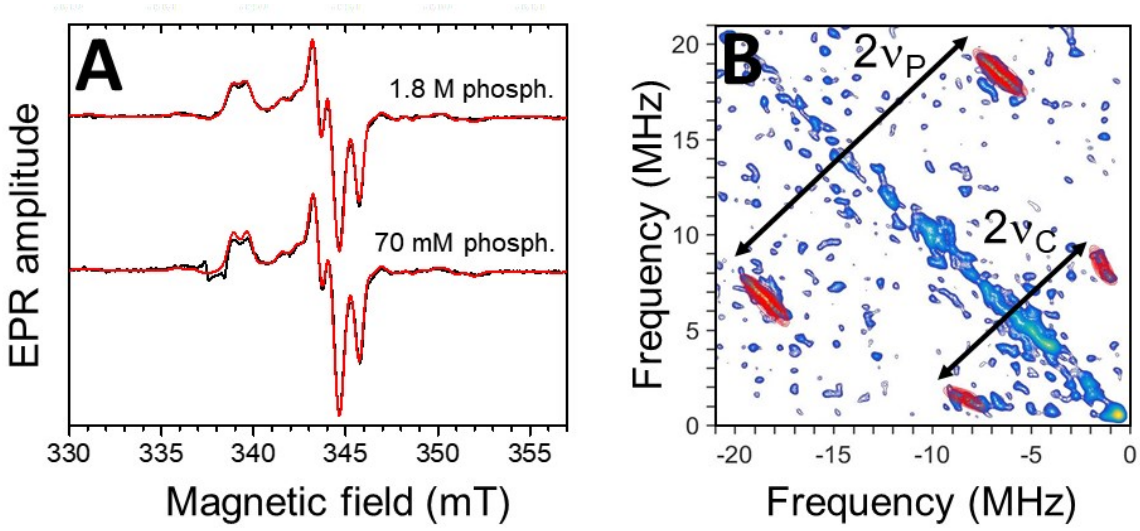

Fig. S4. X-band EPR (A) and HYSCORE (B) spectra of the Mo(V) species in *Tt*SDH prepared with 70 mM sodium phosphate at pH 5.7. Experimental (in black) and simulated (in red) spectra shown in the upper part of Figure S3A are identical to those depicted in Figure 2, while the bottom spectra correspond to the 70 mM phosphate-bound species, as indicated. Parameters used to simulate both spectra are identical (within experimental error) with the exception of the H-strain parameters that are slightly larger for the 70 mM phosphate Mo(V) than those used for the 1.8 M sample ((19, 15, 14) MHz vs. (12, 10, 11) MHz, respectively). Double arrows on Figure S3B schematically indicate double the ^{31}P and ^{13}C Larmor frequencies. Experimental parameters are identical to those given in the caption of figures 2 and 3 for A) and B), respectively. Exceptions are for B), magnetic field value, 352.2 mT, microwave frequency, 9.7135 GHz.

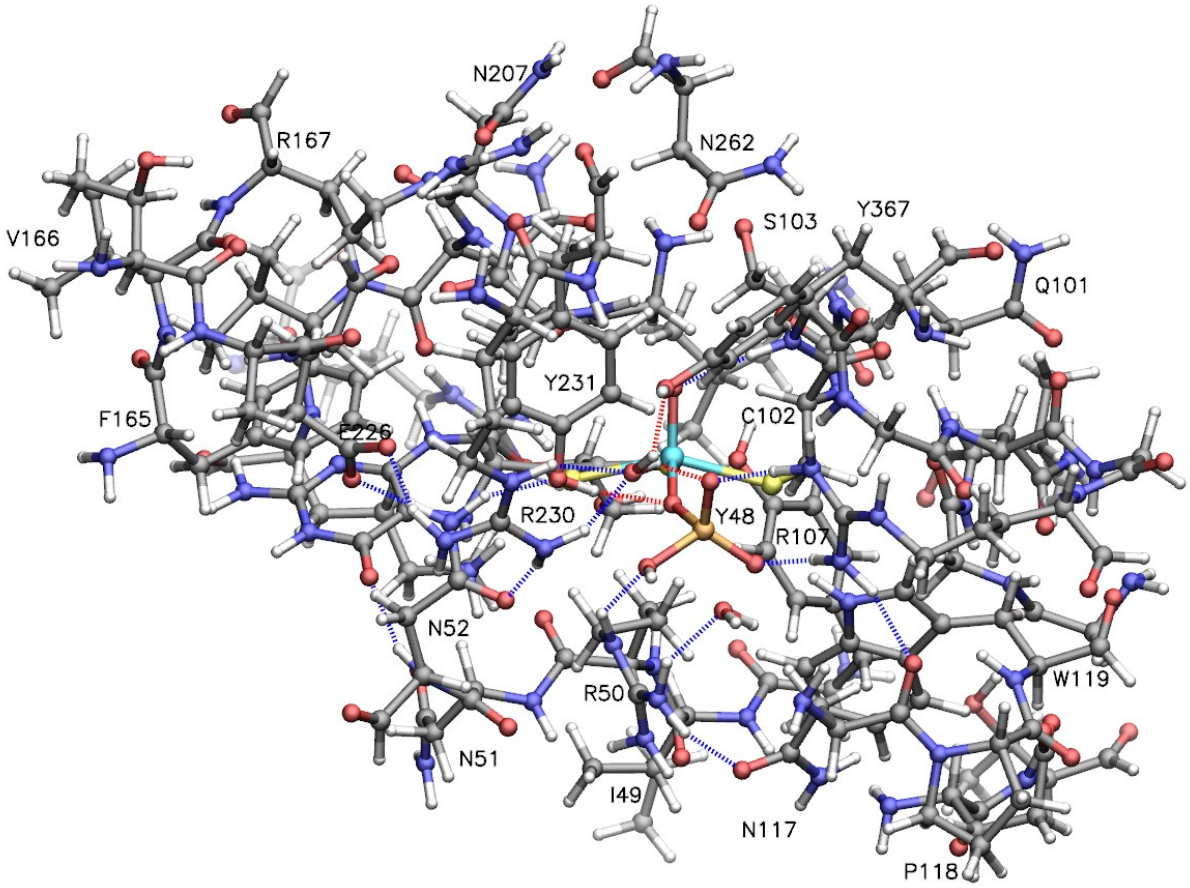

Fig. S5. Geometry-optimized structure of model 1a-$[HPO_4]^{2-}$. Molybdenum is shown as cyan sphere. Dashed lines indicate hydrogen bonds. The residue numbers are indicated for those that are in the foreground.

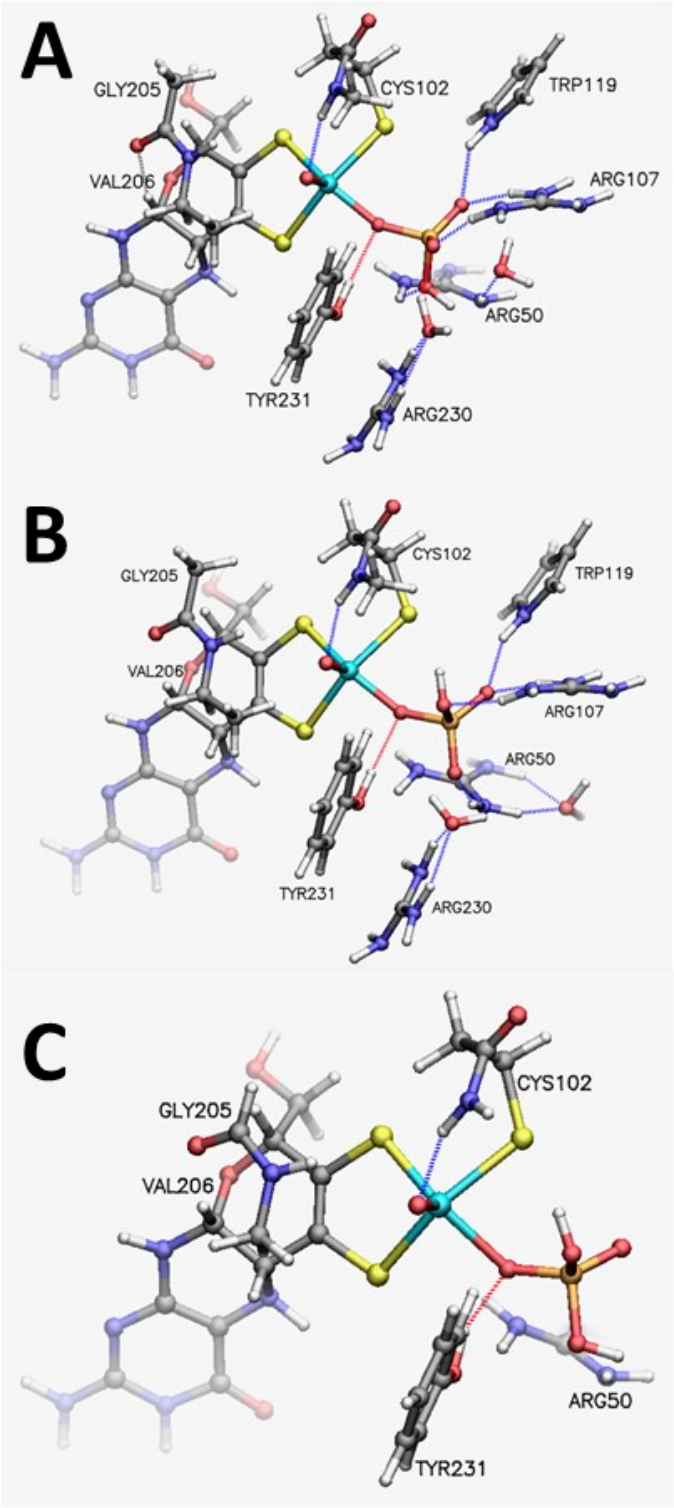

Fig. S6. Geometry-optimized model used for DFT calculations of EPR parameters for model 1a–$[HPO_4]^{2-}$ (A), model 1b–$[HPO_4]^{2-}$ (B) and model 2–$[H_2PO_4]^{-}$ (C). Dashed lines indicate hydrogen bonds.

Table S2. Values of selected structural parameters calculated from the geometry-optimized models and from the original *Tt*SDH X-ray structure with bound phosphate. Comparison between the geometry-optimized models shows that (i) the Mo-$O_{phosphate}$ bond is shorter for HPO_4^{2-} than for $H_2PO_4^-$, ii) the Mo-O-P angle value is significantly larger for HPO_4^{2-} than for $H_2PO_4^-$. Furthermore, the bound phosphate is closer to Arg107 for HPO_4^{2-} ligand (as observed in the crystal structure) and closer to Tyr231 for $H_2PO_4^-$.

	Mo-$O_{phosphate}$ bond length (Å)	Mo-O-P angle (°)	O_{ax}-Mo-O-P dihedral angle (°)	Distance Arg107 amine nitrogen /closest phosphate oxygen (Å)	Distance Tyr231 phenolic oxygen /closest phosphate oxygen (Å)
Model 1a-$[HPO_4]^{2-}$	1.96	152	-70	2.73	3.04
Model 1b-$[HPO_4]^{2-}$	1.99	146	-59	2.84	3.03
Model 2-$[H_2PO_4]^-$	2.24	131	-63	2.95	2.82
X-ray data (pdb code : 6YoK)	2.16	141	-62	2.84	3.04

Table S3. Calculated and experimental eigenvalues of the g-, ^{31}P and ^{95}Mo hyperfine tensors of the phosphate-bound Mo(V) species in *Tt*SDH. [1] determined from simulation of the X-band cw EPR spectrum.

Source	g-values	g_{av}	$g_1 - g_3$ ($\times 10^{-4}$)	$A_i(^{31}P)$ ($A_{iso}(^{31}P)$) [MHz]	$A_i(^{95}Mo)$ [MHz]
Model 1a-$[HPO_4]^{2-}$	$g_1 = 2.0003$ $g_2 = 1.9706$ $g_3 = 1.9620$	1.9776	383	46 35 35 (39)	55 50 147
Model 1b-$[HPO_4]^{2-}$	$g_1 = 2.0053$ $g_2 = 1.9732$ $g_3 = 1.9608$	1.9798	445	42 32 32 (35)	53 47 145
Model 2-$[H_2PO_4]^-$	$g_1 = 2.0178$ $g_2 = 1.9778$ $g_3 = 1.9604$	1.9853	574	19 13 13 (15)	52 48 145
EPR data[1]	$g_1 = 1.9965$ $g_2 = 1.9700$ $g_3 = 1.9620$	1.9762	345	29 22 21 (24)	66 80 167

REFERENCES

1 S. Robin, M. Arese, E. Forte, P. Sarti, A. Giuffre and T. Soulimane, *J Bacteriol*, 2011, **193**, 3988-3997.
2 W. Kabsch, *Acta Crystallogr D Biol Crystallogr*, 2010, **66**, 125-132.
3 P. A. Karplus and K. Diederichs, *Curr. Opin. Struct. Biol.*, 2015, **34**, 60-68.
4 A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni and R. J. Read, *J. Appl. Crystallogr.*, 2007, **40**, 658-674.
5 U. Kappler and S. Bailey, *J. Biol. Chem.*, 2005, **280**, 24999-25007.
6 T. C. Terwilliger, R. W. Grosse-Kunstleve, P. V. Afonine, N. W. Moriarty, P. H. Zwart, L.-W. Hung, R. J. Read and P. D. Adams, *Acta Crystallogr D Biol Crystallogr*, 2008, **64**, 61-69.
7 P. Emsley, B. Lohkamp, W. G. Scott and K. Cowtan, *Acta Crystallogr D Biol Crystallogr*, 2010, **66**, 486-501.

8 P. V. Afonine, R. W. Grosse-Kunstleve, N. Echols, J. J. Headd, N. W. Moriarty, M. Mustyakimov, T. C. Terwilliger, A. Urzhumtsev, P. H. Zwart and P. D. Adams, *Acta Crystallogr D Biol Crystallogr*, 2012, **68**, 352-367.
9 G. N. Murshudov, P. Skubak, A. A. Lebedev, N. S. Pannu, R. A. Steiner, R. A. Nicholls, M. D. Winn, F. Long and A. A. Vagin, *Acta Crystallogr D Biol Crystallogr*, 2011, **67**, 355-367.
10 J. Rendon, F. Biaso, P. Ceccaldi, R. Toci, F. Seduk, A. Magalon, B. Guigliarelli and S. Grimaldi, *Inorg. Chem.*, 2017, **56**, 4422-4434.
11 S. Stoll and R. D. Britt, *Phys. Chem. Chem. Phys.*, 2009, **11**, 6614-6625.
12 S. Stoll and A. Schweiger, *J. Magn. Reson.*, 2006, **178**, 42-55.
13 M. H. Olsson, C. R. Sondergaard, M. Rostkowski and J. H. Jensen, *J. Chem. Theory Comput.*, 2011, **7**, 525-537.
14 S. Dapprich, I. Komaromi, K. S. Byun, K. Morokuma and M. J. Frisch, *J. Mol. Struct.-Theochem*, 1999, **461**, 1-21.
15 A. D. Becke, *J. Chem. Phys.*, 1993, **98**, 5648-5652.
16 C. T. Lee, W. T. Yang and R. G. Parr, *Phys. Rev. B,* **37**, 785-789.
17 D. Andrae, U. Haussermann, M. Dolg, H. Stoll and H. Preuss, *Theor. Chim. Acta*, 1990, **77**, 123-141.
18 F. Weigend and R. Ahlrichs, *Phys. Chem. Chem. Phys.*, 2005, **7**, 3297-3305.
19 S. Grimme, S. Ehrlich and L. Goerigk, *J. Comput. Chem.*, 2011, **32**, 1456-1465.
20 S. Grimme, J. Antony, S. Ehrlich and H. Krieg, *J. Chem. Phys.*, 2010, **132**, 154104.
21 J. H. Enemark, *Dalton. Trans.*, 2017, **46**, 13202-13210.