ELECTRONIC SUPPLEMENTARY INFORMATION (ESI)

The binding modes of $V^{IV}O^{2+}$ ion in blood proteins and enzymes

Giuseppe Sciortino,*[a,b] and Eugenio Garribba*[b]

[a] *Departament de Química, Universitat Autònoma de Barcelona, 08193 Cerdanyola del Vallés, Barcelona, Spain. E-mail: giuseppe.sciortino@uab.cat*

[b] *Dipartimento di Chimica e Farmacia, Università di Sassari, Via Vienna 2, I-07100 Sassari, Italy. E-mail: garribba@uniss.it*

1. Molecular Modelling protocol: DFT and docking calculations

The X-ray structure of the proteins contained in the Protein Data Bank (PDB) (PDB codes: 2dn2 for Hb,[1] 3v83 for holo-hTf,[2] 1igt for IgG,[3] 1qi9 for VBrPO,[4] and 2f1d for IGPD[5]) were first prepared removing all the small molecules and crystallographic waters and, for each structure, the following protocol was applied: i) proteins were probed for zones in which a specific motif based on the electron paramagnetic resonance (EPR) evidences featuring α- and β-carbons of the coordinating amino acids within a range of distances from each queried grid point of 3.4-7.2 and 2.7-6.4 Å, respectively (Tables S3);[6,7] ii) the regions that satisfied these criteria were further evaluated with docking calculations (Table S1); iv) docking solutions were refined at full Density Functional Theory (DFT) level of theory (predicted distances in Table S2); and finally, v) on the optimized structures the hyperfine coupling (HFC) tensors $\boldsymbol{A}(^{51}V)$ or $\boldsymbol{A}(^{14}N)$ were DFT simulated and compared with the experimental ones (Table S1 for $A_z(^{51}V)$ constants).

$V^{IV}O^{2+}$ ion (modeled as $[V^{IV}O(H_2O)_4]^{2+}$) and all the adducts were optimized with Gaussian 09 (revision D.01)[8] at B3P86/6-311++g(d,p) level of theory using the SMD model[9] for water. Frequency calculations were computed to ensure that the structure was a minimum in the potential energy surface.

Docking calculations were carried out through GOLD 5.8 software[10] on the regions selected in the preliminary analysis. The PDB structure was cleaned removing all the small molecules and crystallographic waters, and hydrogen atoms were added with the UCSF Chimera program.[11] The proteins were docked with $V^{IV}O^{2+}$, $V^{IV}O(H_2O)^{2+}$ or $V^{IV}O(H_2O)_2^{2+}$ moieties, obtained from the optimization of the aquaion $[V^{IV}O(H_2O)_4]^{2+}$. The equatorial positions were activated replacing the equatorial water(s) with dummy hydrogen atom(s) according to what was recently established.[12] All dockings were computed considering both the protonation states at δ and ε nitrogens of His imidazole ring. The docking simulations were carried out constructing in the region of interest an evaluation sphere of 8 Å, taking into account side-chains flexibility using the GOLD implemented rotamers libraries.[13] Genetic algorithm (GA) parameters have been set to 50 GA runs and a minimum of 100000 operations. The other parameters of GA were set to default. The solutions were analyzed by means of GaudiView.[6] The scoring (*Fitness* of GoldScore) was evaluated applying the modified version of GoldScore scoring function, which was validated in previously published papers.[12] The best solutions (binding poses) were evaluated through i) the mean (F_{mean}) and the highest value (F_{max}) of the scoring (*Fitness* of GoldScore) associated with each pose and ii) the population of the cluster containing the best pose.

The refinement of the $V^{IV}O^{2+}$–protein adducts found by dockings was performed cutting out the region with the $V^{IV}O^{2+}$ ion and neighbor interacting amino acid side-chains within a radius of 5 Å. The extracted clusters were treated at full DFT theory level for all atoms, completing the valence and freezing the backbone atoms to simulate the protein constraints as reported by Siegbahn and Himo.[14] The geometry relaxation and $\Delta E_{binding}$ calculations were performed at B3P86/6-311++g(d,p) level of theory for all atoms within the framework of SMD model[9] for water. For coordination modes involving the deprotonation of amino acid side-chain donors (aaD), the binding energy was computed considering the proton diffusion to a cluster of 14 H_2O molecules:[15]

$$\text{aaD}\mathbf{H} + (H_2O)_{14} + [V^{IV}O(H_2O)_4]^{2+} \rightarrow \text{aaD}^-\text{–}VO(H_2O)_n + [\mathbf{H}(H_2O)_{14}]^+ + (4\text{–}n)H_2O \qquad \text{(eq. S1)}$$

where aaD**H** is a generic protonated amino acid containing the donor which binds to vanadium in its deprotonated form aaD^-.

2. Calculation of EPR and ESEEM parameters

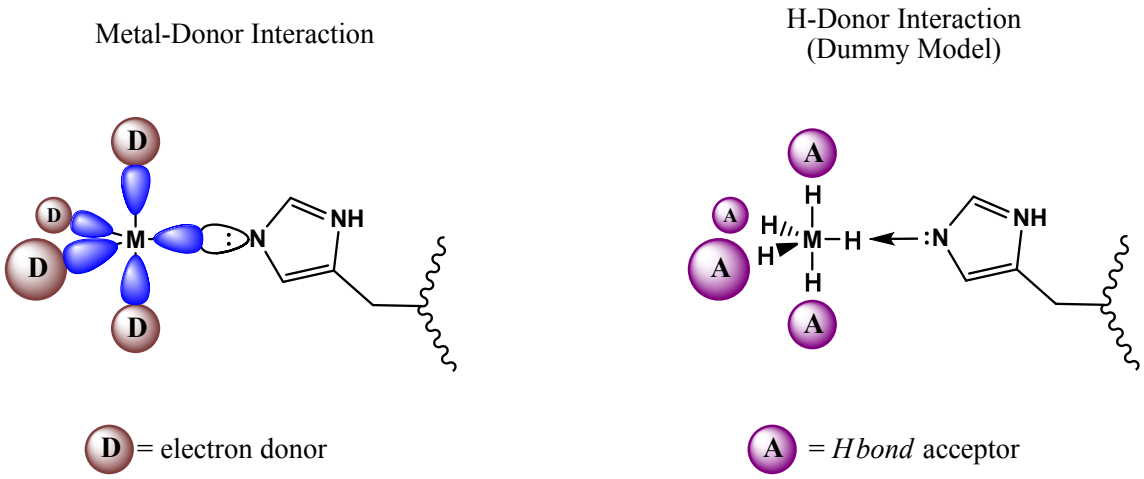
As suggested by the EPR theory, the contribution to the HFC constant at ^{51}V nucleus along the *z* axis [$A_z(^{51}V)$] of the equatorial donors has an approximately inverse relationship with the electron-donating capacity of the ligands, with the most donating donors contributing the least to the coupling constant. This effect is summarized in the "additivity relationship", which affirms that $A_z(^{51}V)$ for a specific coordination environment of $V^{IV}O^{2+}$ ion can be calculated from the sum of the contributions of each equatorial donor.[16] On the other hand, today DFT methods can be used to predict the HFC constants for paramagnetic metals, and vanadium in particular, with high accuracy.[17, 18] These two independent methods were employed to predict $A_z(^{51}V)$ and to confirm the data obtained with docking and QM methods, through a comparison of experimental [$A_z(^{51}V)^{exptl}$] and calculated [$A_z(^{51}V)^{calcd}$] values of A_z. Therefore, the coincidence of $A_z(^{51}V)^{exptl}$ and $A_z(^{51}V)^{calcd}$ was the first criterion to prove the correctness of our computational analysis. Analogous comments apply to the comparison between the experimental [$A_z(^{14}N)^{exptl}$] and calculated [$A_z(^{14}N)^{calcd}$] values of the ^{14}N superhyperfine coupling constants measurable from an electron spin echo envelope modulation (ESEEM) spectrum.[19]

The calculation of the ^{51}V HFC tensor [$\boldsymbol{A}(^{51}V)$] was performed Gaussian 09 (revision D.01)[8] on the optimized structures at BHandHLYP/6-311+g(d) level of theory, according to the methods previously published.[18] The interested readers can find the theory background in refs. 18, 20. The ^{14}N superhyperfine coupling tensor [$\boldsymbol{A}(^{14}N)$] was predicted with BHandH functional and 6-311g(d,p) basis set, according to the ref. 21.

3. Docking protocol details

Docking calculations were performed through the software GOLD 5.8[10] using GoldScore function[22] and the set of scoring parameters optimized in our recently published study.[12c]

Our strategy consists in defining the coordination bond using the hydrogen bond (*Hbond*) function included in GoldScore without source code modifications. The metal is virtually described as a hydrogen bond donor liable to interact with the *Hbond* acceptors of protein by means of dummy atoms positioned in the coordination vacancies. In terms of Lewis acid and base theory, the acid centre is transferred from the metal to the fictitious proton located along the virtual coordination bond axis (Scheme S1).



Scheme S1. Description of the dummy hydrogen approach to metal ion docking in proteins.

The S_{Hbond} Lennard-Jones (LJ) like functions (eqs. S2 and S3)[22] are able to describe the coordination bond lengths and angles by the multiplication of the potential for a weight block function $\left(wt = distance_{wt} + angle_{wt}\right)$, eq. S4:[22]

$$S_{i,j} = \sum_{i,j}\left(\frac{A}{r_{ij}^{8}} - \frac{B}{r_{ij}^{4}}\right) \qquad \text{(eq. S2)}$$

If the interatomic distance is equal to the sum of van der Waals (vdW) radius of the atom pairs, the potential utilized is a *Lennard-Jones 12-6*:

$$S_{i,j} = \sum_{i,j}\left(\frac{C}{r_{ij}^{12}} - \frac{D}{r_{ij}^{6}}\right) = \sum_{i,j}\left(\frac{C}{r_{ij}^{12}} - \frac{2}{3}\cdot 0.23 \cdot \frac{I_i I_j \alpha_i \alpha_j}{\left(I_i + I_j\right) r_{ij}^{6}}\right) \qquad \text{(eq. S3)}$$

where I_i, I_j, and α_i, α_j are, respectively, the ionization energy and the polarizability of the *i* and *j* atoms. The terms *A*, *B*, *C* and *D* are empirical coefficients present in Gold database.

$$x_{wt}\left(x, x_{ideal}, x_{max}\right) = \begin{cases} 1, & \text{if } x \le x_{ideal} \\ 1.0 - \dfrac{x - x_{ideal}}{x_{max} - x_{ideal}}, & \text{if } x_{ideal} \le x \le x_{max} \\ 0, & \text{if } x > x_{max} \end{cases} \qquad \text{(eq. S4)}$$

The strength of the coordination-like interactions are defined by a series of empirical parameters implemented in *GoldScore* that take into account the relative affinity of the donors for the different metals. The parameters have been validated for 15 metals (Mg, V, Cr, Mn, Fe, Co, Ni, Cu, Zn, Ru, Rh, Re, Os, Pt, Au) and all the amino acid acceptors.[12a-c] Particularly, in the case of $V^{IV}O^{2+}$ the atoms parameter libraries, the "coordination" energetic terms implemented in the Goldscore scoring function are reported in the next section.

2.1. Atoms parameter libraries for $V^{IV}O^{2+}$ atom type

a.1 vanadium

1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
M.V4	1.57	0.6	26.10	0.0	50.9	4.0	CO	6	Y	N

a.2 Water oxygen coordinated of the drug (i.e. maltol)

1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
O.H2O	1.52	0.116	24.39	5.4	16.0	0.0	TET	3	Y	Y

Each line in the atom table has 11 entries.

1. Atom type. Except for N.acid, N.plc, S.a, and S.min, these are equivalent to SYBYL atom types (see SYBYL theory manual).

2. van der Waals radius of the atom. The pairwise potentials used by GOLD assume that the interaction energy between two atoms is at a minimum at the sum of the vdW radii. Values of vdW radii used should reflect this. The values used are the TAFF vdW radii as used in SYBYL.

3. K value used by TAFF. This is used if the TAFF vdW interaction is used in preference to the dispersive potential.

4. Ionisation potential for atom. Used in the dispersive potential.

5. Polarizability of the atom. Used in the dispersive potential. Polarizabilities from ref. [23].

6. Atomic weight of atom. Not required by GOLD.

7. Formal charge of atom. Not required by GOLD.

8. Geometry of atom. One of NONE, TET (tetrahedral), TRI (trigonal), LIN (linear) or CO (coordinated metal ion). Used for adding lone pairs.

9. Number of neighbors of atom. Used for adding lone pairs.

10. Is the atom a donor? N.B. metals are modeled as donors.

11. Is the atom an acceptor?

2.2. Vanadium "coordination" energetic terms implemented in the GoldScore scoring function

Donor	Acceptor (GOLD abbreviation)	Strength
V	N2DA	-10
V	O2A	-5
V	OCO2A	-15
V	N1A	-10
V	N3A	-10
V	O3A	-10
V	N2A	-10
V	N3DA	-10
V	O3DA	0
V	NACIDA	-15
V	O2NA	-1
V	ONO2A	-10
V	SMIN	-20
V	SACC	-15
V	NARA	-20
V	O3MINUSA	-20
V	OMINUSA	-20

The optimal distance between the metal and the hydrogen that acts as a dummy was defined as 0.75 Å. Furthermore the *Hbond* mean distance (in the original scoring function 2.9 Å) was set to 1.9 Å. The methodology was applied to characterize the $V^{IV}O^{2+}$ binding site of carboxypeptidase.[12a]

4. Mol2 files including dummy H (in red)

$V^{IV}O^{2+}$ (4 dummies H)

```
@<TRIPOS>MOLECULE
Molecule Name
6 5 1 0 0
SMALL
NO_CHARGES

@<TRIPOS>ATOM
      1 V1          0.0092   -0.0097    0.1377 M.V4      1 VO1    0.0000
      2 O2         -0.0209    0.0889    1.7109 O.2       1 VO1    0.0000
      3 H1          0.4450    0.5659   -0.0653 H         1 VO1    0.0000
      4 H2          0.5858   -0.4511   -0.0501 H         1 VO1    0.0000
      5 H3         -0.4370   -0.5776   -0.0645 H         1 VO1    0.0000
      6 H4         -0.5847    0.3947   -0.0773 H         1 VO1    0.0000
@<TRIPOS>BOND
     1    1    2 1
     2    1    3 1
     3    1    4 1
     4    1    5 1
     5    1    6 1
@<TRIPOS>SUBSTRUCTURE
     1 UNK     1 RESIDUE           4 A     UNK     0 ROOT
```

$V^{IV}O(H_2O)^{2+}$ (3 dummies H)

```
@<TRIPOS>MOLECULE
Molecule Name
8 7 1 0 0
SMALL
NO_CHARGES

@<TRIPOS>ATOM
      1 V1          0.0092   -0.0097    0.1377 M.V4      1 VO2    0.0000
      2 O2         -0.0209    0.0889    1.7109 O.2       1 VO2    0.0000
      3 H1          0.4450    0.5659   -0.0653 H         1 VO2    0.0000
      4 H2          0.5858   -0.4511   -0.0501 H         1 VO2    0.0000
      5 H3         -0.4370   -0.5776   -0.0645 H         1 VO2    0.0000
      6 O3         -1.6089    1.0921   -0.4482 O.H2O     1 VO2    0.0000
      7 H4         -1.5344    1.6224   -1.2545 H         1 VO2    0.0000
      8 H5         -2.1542    1.5367    0.2155 H         1 VO2    0.0000
@<TRIPOS>BOND
     1    1    2 1
```

```
     2    6    7 1
     3    6    8 1
     4    1    6 1
     5    1    3 1
     6    1    4 1
     7    1    5 1
@<TRIPOS>SUBSTRUCTURE
     1 UNK     1 RESIDUE           4 A     UNK     0 ROOT
```

cis-$V^{IV}O(H_2O)^{2+}$ (2 dummies H)

```
@<TRIPOS>MOLECULE
VH8O5
10 9 1 0 0
SMALL
NO_CHARGES

@<TRIPOS>ATOM
      1 V1          0.0058    0.0145    0.1471 M.V4      1 VO3    0.0000
      2 O1         -0.0680    0.0110    1.7008 O.2       1 VO3    0.0000
      3 H1          0.0865   -0.7061   -0.0445 H         1 VO3    0.0000
      4 H2          0.7393    0.0706    0.0008 H         1 VO3    0.0000
      5 O3         -2.0034   -0.1783   -0.3749 O.H2O     1 VO3    0.0000
      6 H3         -2.1165   -0.6669   -1.2039 H         1 VO3    0.0000
      7 H4         -2.4936   -0.6764    0.2948 H         1 VO3    0.0000
      8 O4         -0.2390    1.9831   -0.4063 O.H2O     1 VO3    0.0000
      9 H5         -1.0553    2.1713   -0.8906 H         1 VO3    0.0000
     10 H6          0.4856    2.3909   -0.8996 H         1 VO3    0.0000
@<TRIPOS>BOND
     1    4    5 1
     2    4    6 1
     3    7    8 1
     4    7    9 1
     5    4    1 1
     6    3    1 1
     7    2    1 1
     8   10    1 1
     9    7    1 1
@<TRIPOS>SUBSTRUCTURE
     1 UNK     1 RESIDUE           4 A     UNK     0 ROOT
```

trans-$V^{IV}O(H_2O)^{2+}$ (2 dummies H)

```
@<TRIPOS>MOLECULE
```

```
VH8O5
10 9 1 0 0
SMALL
NO_CHARGES


@<TRIPOS>ATOM
      1 V1          0.0058    0.0145    0.1471 M.V4      1 VO4     0.0000
      2 O1         -0.0680    0.0110    1.7008 O.2       1 VO4     0.0000
      3 H1          0.0865   -0.7061   -0.0445 H         1 VO4     0.0000
      4 H2         -0.0834    0.7314   -0.0544 H         1 VO4     0.0000
      5 O2         -2.0034   -0.1783   -0.3749 O.H2O     1 VO4     0.0000
      6 H3         -2.1165   -0.6669   -1.2039 H         1 VO4     0.0000
      7 H4         -2.4936   -0.6764    0.2948 H         1 VO4     0.0000
      8 O3          2.0197    0.1685   -0.2545 O.H2O     1 VO4     0.0000
      9 H5          2.3283   -0.3229   -1.0291 H         1 VO4     0.0000
     10 H6          2.3805    1.0630   -0.3320 H         1 VO4     0.0000
@<TRIPOS>BOND
     1    4    5 1
     2    4    6 1
     3    8    9 1
     4    8   10 1
     5    7    1 1
     6    4    1 1
     7    8    1 1
     8    1    2 1
     9    1    3 1
@<TRIPOS>SUBSTRUCTURE
     1 UNK     1 RESIDUE           4 A     UNK     0 ROOT
```

5. Additional validation

The unique X-ray structures of the PDB[24] containing $V^{IV}O^{2+}$ ion bound to a protein were examined: the adducts with α-ketoglutarate-dependent taurine dioxygenase (TauD; PDB code: 6edh[25]) and with a variant of cyt cb_{562} ($V^{IV}O$-CH_3Y*; PDB: 6dyl[26]). The CH_3Y* variant of cyt cb_{562} was designed with three histidines in the positions 67, 71 and 97 (H67, H71 and H97), with the mutations T96C and G70Y and removing the heme group.[26] The two proteins (TauD and CH_3Y* variant of cyt cb_{562}) were cleaned and prepared for the docking as reported in the section 1. Subsequently, the $V^{IV}O^{2+}$ and coordinated H_2O were removed from the protein binding site and the relative coordinates saved in a new .mol2 file. The metal-containing ligands were pre-treated by virtually activating the vacant coordination sites through a dummy hydrogen atom according to the procedure recently established.[12, 27]

Docking calculations were performed with the software GOLD 5.8[10] using our modification of GoldScore function.[22] The metal complexes were blindly re-docked to the protein without any geometrical constraints or energy restraints building an evaluation sphere of 20 Å centered on the selected binding site. Genetic algorithm (GA) parameters were set at 50 runs and minimum of 100000 operations. All the other parameters – pressure, number of islands, niche size, crossover, mutation and migration – were set to default.

All the simulations were carried out considering full rotation of the V^{IV}–OH_2 bonds through the GOLD implemented algorithm. In a first simulation the protein residues were considered rigid as in the X-ray structure because the amino acid side-chains are already in the ideal conformation to bind the metal. To assess the general predictive capability of the method, in a second simulation the flexibility of the coordinating side-chains was taken into account using the rotamer libraries implemented in GOLD software.[13]

The results for the rigid docking between $V^{IV}O^{2+}(H_2O)_2$ and TauD show a unique cluster of 50 docking poses with the highest scoring (F_{max}) of 60.5 GoldScore *Fitness* units. The predicted solution indicates $V^{IV}O^{2+}$ binding with three different donors, two histidines (His99 and His255) and one aspartate residue (Asp101) corresponding to the crystallographic site. An RMSD of 0.091 Å is reported for the highest scoring solution (Fig. S1a). When considering side-chains flexibility the same binding site is found in the first ranked cluster containing the 51% of the solutions with the highest scoring of 62.4 GoldScore *Fitness* units and an RMSD respect to the experimental adduct of 0.161 Å. The slight differences in F_{max} and RMSD are attributed to the small changes in the side-chains able to rotate during docking (Fig. S1b). Finally, we push forward our method performing the prediction with the unbiased metal free structure of TauD (PDB: 3v15[28]) to ascertain the applicability of the technique starting from the apo structure. From this analysis results similar

to those with the metal-loaded structure were obtained (Fig. S1c), confirming that the side-chains full flexibility around the metal ion is generally enough to analyze the metal binding capability of the proteins and take into account the possible differences in the conformation between the apo and holo form. However, when proteins with high flexible regions or small peptides are examined, an extended conformational sampling is required.[29] In those cases preliminary Normal Mode Analysis (NMA) or Molecular Dynamics (MD) simulation of the apo structure should be performed, basing the subsequent metal binding prediction on the different sampled conformations.

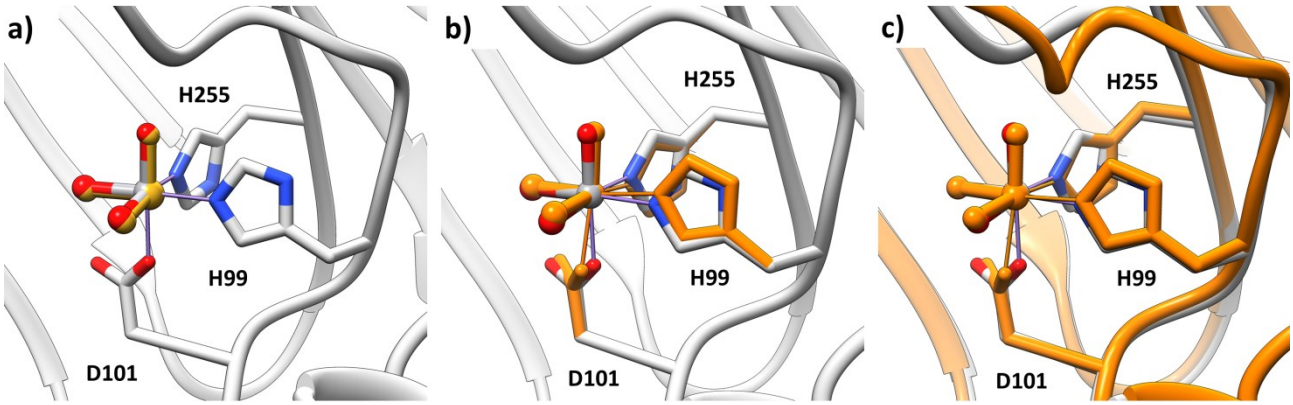


Fig. S1. Overlap of the X-ray structure (in gray) and the best docking pose (in yellow or orange) of the adduct between TauD and $V^{IV}O(H_2O)^{2+}$ moiety: a) rigid; b) flexible docking frameworks and c) docking using the unbiased metal free protein.

Similar results are found for dockings between $V^{IV}O^{2+}$ and CH_3Y* variant of cyt cb_{562}, for which using the rigid structure a fully populated cluster with F_{max} of 102.0 is found with five histidines bound to vanadium (His67.A, His71.A, His97.A, His67.B and His71.B); this structure corresponds to the crystallographic one (Fig. S2a). The best solution shows an RMSD of 0.039 Å. In the flexible framework the same binding site was identified in a fully populated cluster with F_{max} of 103.3 and RMSD 0.237 Å. As in the case of the adduct between $V^{IV}O^{2+}$ and Tau (Fig. 1), differences in RMSD and F_{max} are attributed to the small changes of the side-chain position due to the free rotation during docking (Fig. S2b).

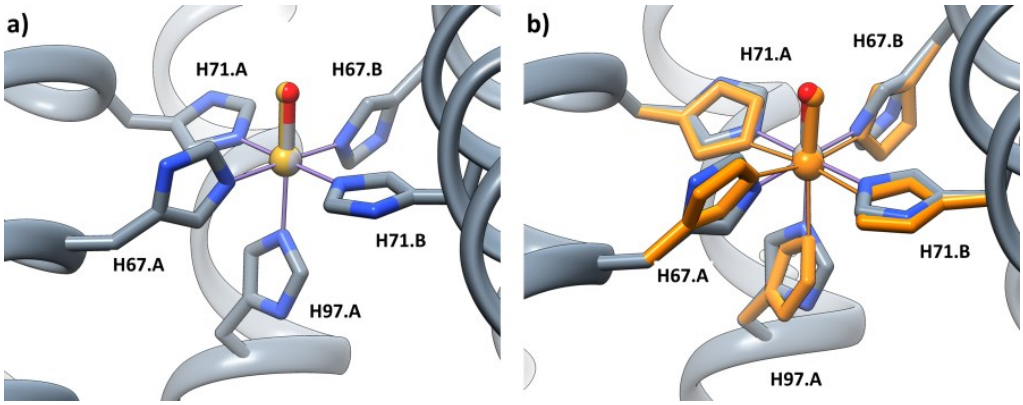


Fig. S2. Overlap of the X-ray structure (in gray) and the best docking pose (in yellow or orange) of the adduct between the CH_3Y* variant of cyt cb_{562} and $V^{IV}O^{2+}$ ion: a) rigid and b) flexible docking frameworks.

Table S1. Binding modes of serum proteins and enzymes for $V^{IV}O^{2+}$ ion identified by docking, binding energies and and $A_z(^{51}V)$ constants computed on their full QM optimized structures.

Protein	Binding Mode	Site	F_{max}[a]	F_{mean}[b]	Pop.[c]	$\Delta E_{binding}$[d]	$A_z(^{51}V)^{calcd}$[e]
Hb	N_{H116}, N_{H117}, COO^-_{E26}; H_2O	**γ¹**	68.2	65.8	9	-27.4	173.7 [f]
Hb	N_{H116}, N_{H117}, COO^-_{E26}; H_2O	**γ²**	61.3	55.8	22	-29.2	168.7 [f]
Hb	N_{H116}, N_{H117}; COO^-_{E26}; H_2O	**γ³**	58.4	53.9	15	-26.9	166.8 [f]
Hb	N_{H146}, COO^-_{E90}, COO^-_{D94}; H_2O	**β**	40.2	37.8	3	-30.9	171.2 [f]
Hb	N_{H112}, N_{H116}, HO_{Y21}; H_2O	**–**	38.0	37.9	3	-1.7	185.2 [f]
hTf	N_{H14}, N_{H289}, COO^-_{E13}; H_2O	**C¹**	66.0	59.0	27	-24.3	169.5 [g]
hTf	N_{H25}, N_{H273}; H_2O; OH^- [h]	**C²**	38.8	38.8	37	-32.8	167.9 [g]
hTf	N_{H473}, COO^-_{D478}, H_2O; H_2O [i]	**C³**	37.6	34.7	12	-26.4	168.6 [g]
hTf	N_{H349}, N_{H350}, COO^-_{E507}; H_2O; H_2O(ax)	**C⁴**	45.4	43.5	31	-13.1	172.3 [g]
hTf	COO^-_{D477}, COO^-_{D478}, H_2O; H_2O [i]	**C⁵**	41.4	35.1	9	-27.4	176.2 [g]
IgG	N_{H172}, COO^-_{D167}, O^-_{S174}; H_2O	**1**	75.0	69.3	47	-19.5	158.4 [j]
IgG	N_{H172}, COO^-_{D167}, COO^-_{D170}, NCO_{N138}	**2**	57.8	55.2	48	-31.8	165.8 [j]
IgG	N_{H460}, COO^-_{E461}, CO_{D399}; H_2O	**3**	42.4	37.7	45	-23.5	172.3 [j]
IgG	N_{H189}, COO^-_{E185}, CO_{E185}; H_2O	**3¹**	52.5	46.5	42	-26.2	171.0 [j]
VBrPO	N_{H418}, N_{H486}. O^-_{S416}; H_2O		66.3	65.4	50	-8.0	156.0 [k]
VBrPO	N_{H418}, N_{H486}, OH_{S416}; OH^-		56.3	55.4	50	-11.6	165.0 [k]
VBrPO	N_{H418}, N_{H486}, OH_{S416}; H_2O		56.3	55.4	50	-4.8	175.3 [k]
IGPD	N_{H74}, N_{H169}, N_{H47};COO^-_{E173}	**α¹**	54.9	52.2	16	-43.1	166.8 [l]
IGPD	N_{H74}, N_{H169}, N_{H47}(ax), COO^-_{E173}; H_2O	**α²**	54.8	52.2	10	-39.2	164.9 [l]
IGPD	N_{H73}, N_{H145}(ax), N_{H170}, COO^-_{E77}; H_2O	**β¹**	52.9	50.0	8	-47.4	164.3 [l]
IGPD	N_{H159}, N_{H113}, O^-_{S115}, H_2O	–	57.9	57.1	2	-15.3	156.0 [l]

[a] GoldScore *Fitness* value obtained for the more stable pose of each *cluster*. [b] Average value of GoldScore *Fitness* for each *cluster*. [c] Population of the *cluster*. [d] $\Delta E_{binding}$ in kcal/mol. [e] $A_z(^{51}V)$ reported in 10^{-4} cm^{-1} units as absolute values. [f] Absolute $A_z(^{51}V)$ value of 163.3 × 10^{-4} cm^{-1} for **site γ** and 166.8 × 10^{-4} cm^{-1} for **site β** (ref. 30). [g] Absolute $A_z(^{51}V)$ value of 165.4 × 10^{-4} cm^{-1} (ref. 31). [h] *cis* arrangement of H_2O and OH^- ligands. [i] *trans* arrangement of H_2O ligands. [j] Absolute $A_z(^{51}V)$ value of 158.8 × 10^{-4} cm^{-1} for **site 1**, 163.6 × 10^{-4} cm^{-1} for **site 2** and 167.4 × 10^{-4} cm^{-1} for site 3 (ref. 32). [k] Absolute $A_z(^{51}V)$ value of 160.1 × 10^{-4} cm^{-1} (ref. 33). [l] Absolute $A_z(^{51}V)$ value of 169.1 × 10^{-4} cm^{-1} for **site α**, 161.6 × 10^{-4} cm^{-1} for **site β** and 140.6 × 10^{-4} cm^{-1} for **site γ** (ref. 34).

Table S2. DFT optimized bond lengths for the $V^{IV}O^{2+}$ adducts.

Protein	Donors	V–N [a]	V–O [a]	V–OH_2 [a]
Hb	N_{H146}, COO^-_{E90}, COO^-_{D94}; H_2O (**site β**)	2.162	1.988, 1.952	2.094
Hb	N_{H116}, N_{H117}, COO^-_{E26}; H_2O (**site γ^1**)	2.096, 2.093	1.945	2.078
Hb	N_{H116}, N_{H117}, COO^-_{E26}; H_2O (**site γ^2**)	2.085, 2.105	1.935	2.091
Hb	N_{H116}, N_{H117}, COO^-_{E26}; H_2O (**site γ^3**)	2.061, 2.136	1.966	2.080
hTf	N_{H14}, N_{H289}, COO^-_{E13}; H_2O (**site C^1**)	2.122, 2.082	1.969	2.068
hTf	N_{H25}, N_{H273}; H_2O, OH^- (**site C^2**)[b]	2.091, 2.114	–	2.128, 1.902
hTf	N_{H473}, COO^-_{D478}; H_2O, H_2O (**site C^3**)[c]	2.072	1.990	2.067; 2.035
IgG	N_{H189}, COO^-_{E185}, CO_{E185}; H_2O (**site 3**)	2.083	1.959, 1.975 (CO)	2.070
IgG	N_{H460}, COO^-_{E461}, CO_{D399}; H_2O (**site 3**)	2.112	1.991,2.010 (CO)	2.091
IgG	N_{H172}, COO^-_{D167}, COO^-_{D170}, NCO_{N138} (**site 2**)	2.149	1.982, 1.994, 2.077 (Asn)	–
IgG	N_{H172}, COO^-_{D167}, O^-_{S174}; H_2O (**site 1**)	2.153	2.075, 1.874 (Ser)	2.095
VBrPO	N_{H418}, N_{H486}, O^-_{S416}; H_2O	2.051, 2.201	1.897	2.174
VBrPO	N_{H418}, N_{H486}, OH_{S416}; OH^-	2.073, 2.172	2.263	1.889
VBrPO	N_{H418}, N_{H486}, OH_{S416}; H_2O	2.054, 2.161	2.162	2.116
IGPD	N_{H74}, N_{H169}, N_{H47},COO^-_{E173} (**site α^1**)	2.115, 2.109, 2.085	1.962	–
IGPD	N_{H74}, N_{H169}, N_{H47}(ax), COO^-_{E173}; H_2O (**site α^2**)	2.119, 2.132, 2.399	2.013	2.076
IGPD	N_{H73}, N_{H145}(ax), N_{H170}, COO^-_{E77}; H_2O (**site β**)	2.116, 2.360, 2.124	2.049	2.012
IGPD	N_{H159}, N_{H113}, O^-_{S115}; H_2O	2.141; 2.148	1.867	2.105

[a] Distances in Å. [b] *cis* arrangement of H_2O and OH^- ligands. [c] *trans* arrangement of H_2O ligands.

Table S3. Coordinating motifs found during the first screening.[a]

Potential Donors	**Region**	**Number of Sites**
Hb (PDB code: 2dn2 [1])		
Queried Motifs: His; His; Asp/Glu and His; Asp/Glu; Asp/Glu		
His116.B/D; His117.B/D; Glu26.B/D	Chains B and D	2
His112.A/C; His20.A/C; Glu23/27.A/C	Chains A and C	2
His146.B/D; Glu90.B/D; Asp94.B/D	Chains B and D	2
holo-hTf (PDB code: 3v83 [2])		
Queried Motifs: His; His; Asp/Glu and His; Asp/Glu; Asp/Glu		
His25.A; His273.A; Glu281.A	N2-subdomain	1
Glu13.A; His14.A; His289.A	N2-subdomain	1
His349.A; His350.A; Glu507/372.A	C1/C2 subdomains interface	1
Glu410.A; His585.A; Asp634/628.A	C1/C2 Hinge (Fe-binding)	1
His473.A; **Asp477.A; Glu478.A**	C2 subdomain	1
IgG (PDB code: 1igt [3])		
Queried Motifs: His; Asp/Glu; Asp/Glu and His; Asp/Glu,Ser/Thr/Tyr		
His189.A/C; Asp151.A/C; Glu154.A/C	Fab C_L	2
His460.B/D; Asp399.B/D; Glu461.B/D	Fc C_H3	2
His172.B/D; Asp167.A/C; Asp170.A/C	Fab C_L/C_H1 Interface	2
His189.A/C; Asp184.A/C; Glu187.A/C	Fab C_L	2
His466.B/D; Asp399.B/D; Glu461.B/D	Fc C_H3	2
His189.A/C; Glu185.A/C; Tyr192.A/C	Fab C_L	2
His172.B/D; Asp167.A/C; Ser174.A/C	Fab C_L/C_H3 Interface	2
His467.B/D; Glu405.B/D; Tyr401/Ser457.B/D	Fc C_H3	2
His55.A/C; Asp101.A/C; Tyr49/Tyr102.A/C	Fab V_L	2
His302.B/D; Glu300.B/D; Ser289.B/D	Fc C_H2	2
His198.A/C; Asp110/143.A/C; Tyr173/Thr200.A/C	Fab C_L	2
VBrPO (PDB code: 1qi9 [4])		
Queried Motifs: His; His; Ser/Thr/Tyr		
His418.A/B His486.A/B; Ser416.A/B	Active Site	2
IGPD (PDB code: 2f1d [5])		
Queried Motifs: His; His; His; Asp/Glu and His; His; Ser/Thr/Tyr		
His169.C-D; His73.A-B; His74.A-B; Glu173.C-D/77.A-B/Asp71.A-B	Active Site	2
His169.C-D; His73.A; His47.C-D; Asp71.A-B/Glu173.C-D	Active Site	2/4
Hsi169.C-D; His73.A-B; His170.C-D Glu173.C-D/77.A-B	Active Site	2
His169.C-D; His73.A-B; His145.A-B; Glu77.A-B/Asp139.A-B	Active Site	2
His169.C-D; His145.A-B; His170.C-D Glu166.C-D/77.A-B/Asp108/109.C-D	Active Site	2
His169.C-D; His47.C-D; His74.A-B; Glu173.C-D/Asp71.A-B	Active Site	2
His169.C-D; His74.A-B; His170.C-D; **Glu173.C-D**	Active Site	2
His47.C-D; His74.A-B; His73.A-B; Asp67/71.A-B/Glu21.A-B/173.C-D	Active Site	2
His73.A-B; His170.C-D; His145.A-B; Glu77.A-B	Active Site	2
His73.A-B; His170.C-D; His74.A-B; Glu173.C-D	Active Site	2
His113.A-D; Ser115.A-D; His159.A-D	β6/β7 hinge	4

[a] In bold the suitable coordinating residues found by docking.

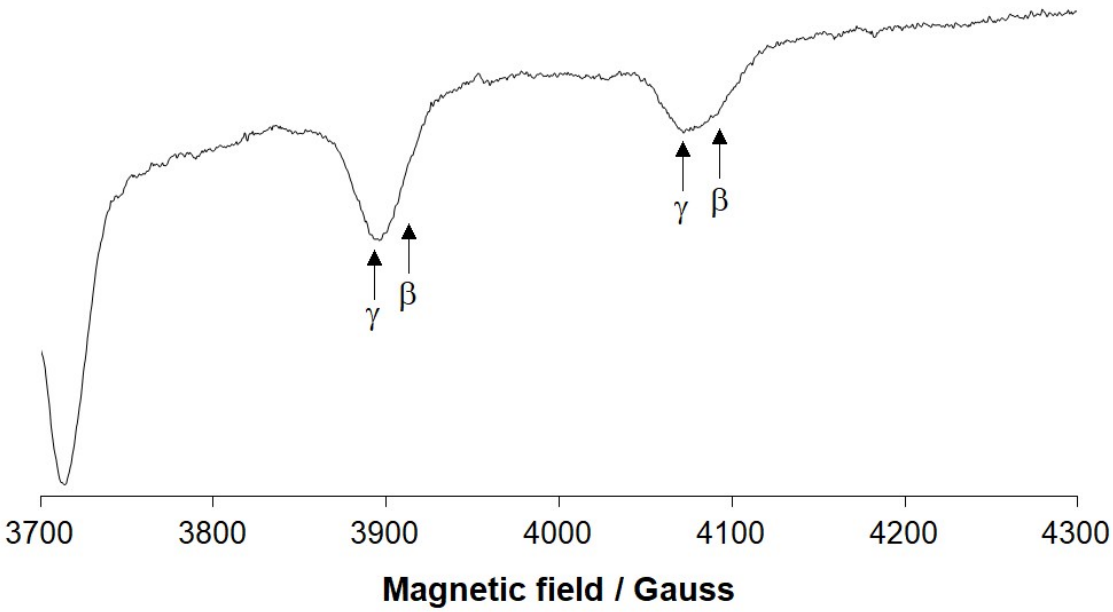


Fig. S3. High field region of the X-band anisotropic EPR spectrum recorded at pH 7.4 on a frozen solution (120 K) containing $V^{IV}O^{2+}$/Hb with molar ratio 2/1 ($V^{IV}O^{2+}$ concentration 6.2×10^{-4} M). The M_I = 5/2, 7/2 resonances of the **sites β** and **γ** are indicated. Adapted from ref. 30. The measured spin Hamiltonian parameters are: $g_z = 1.952$, $|A_z| = 166.8 \times 10^{-4}$ cm^{-1} (**site β**), and $g_z = 1.953$, $|A_z| = 163.3 \times 10^{-4}$ cm^{-1} (**site γ**).

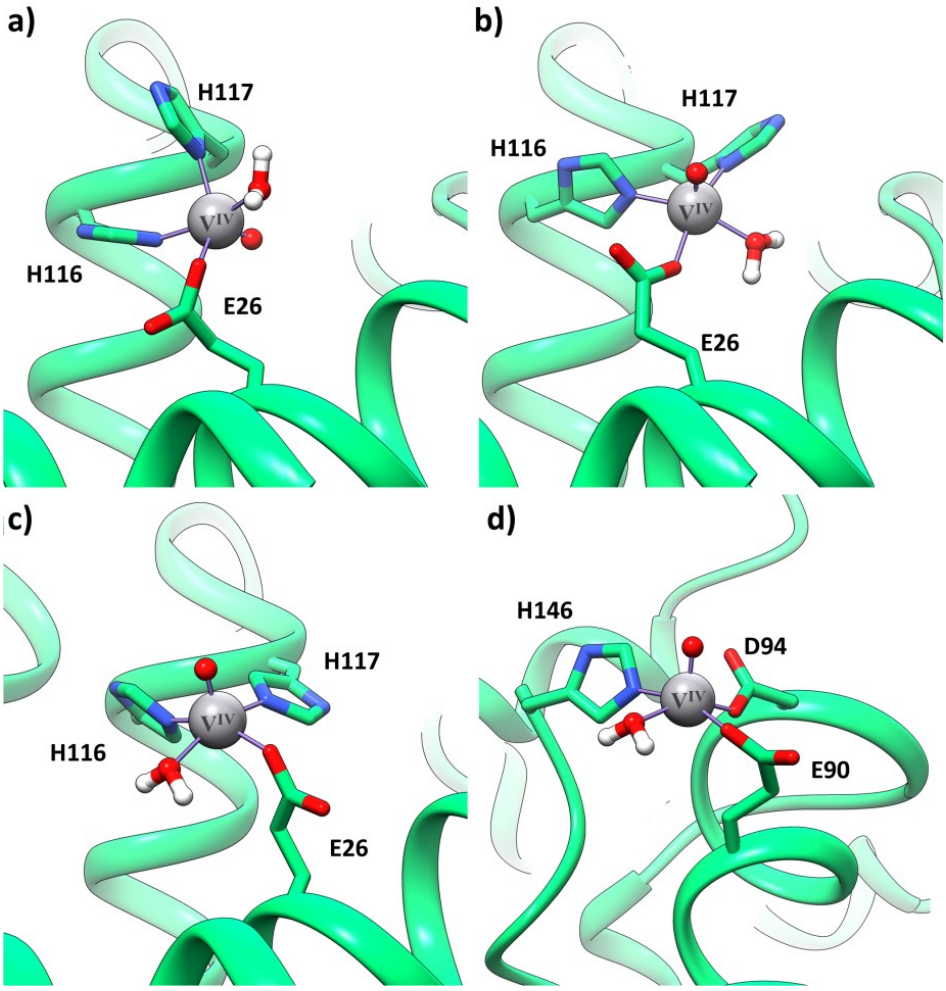


Fig. S4. Docking solutions for the binding of $V^{IV}O^{2+}$ ion to the **sites β** and **γ** of Hb: (a) mode γ^1, (b) mode γ^2, (c) mode γ^3 and (d) site β.

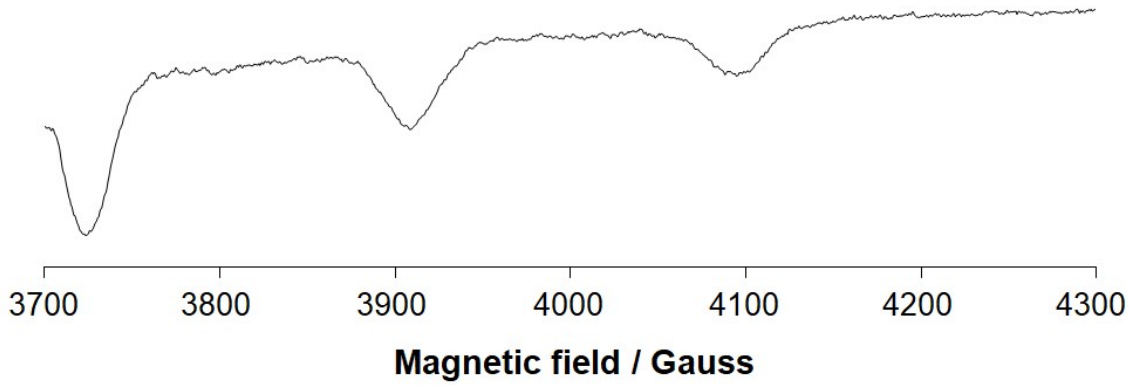


Fig. S5. High field region of the X-band anisotropic EPR spectrum recorded at pH 7.4 on a frozen solution (120 K) containing $V^{IV}O^{2+}$/holo-hTf with molar ratio 2/1 ($V^{IV}O^{2+}$ concentration 5.0×10^{-4} M). The resonances are those of the **site C**. Adapted from ref. 31. The measured spin Hamiltonian parameters are: $g_z = 1.944$, $|A_z| = 165.4 \times 10^{-4}$ cm^{-1}.

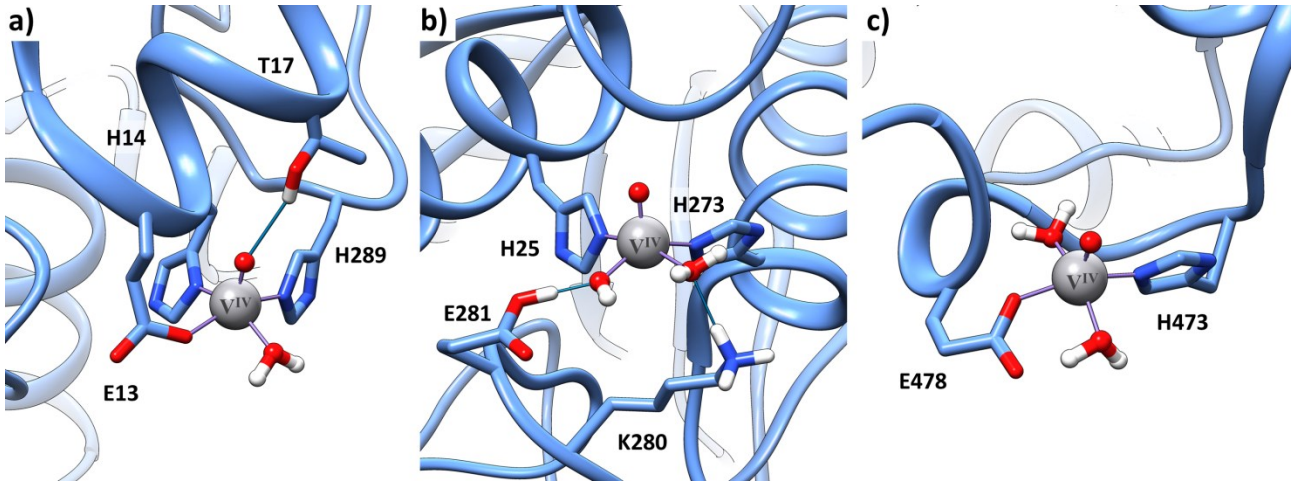


Fig. S6. Docking solutions for the binding of $V^{IV}O^{2+}$ ion to the **site C** of hTf: (a) mode $\mathbf{C^1}$, b) mode $\mathbf{C^2}$ and (c) mode $\mathbf{C^3}$.

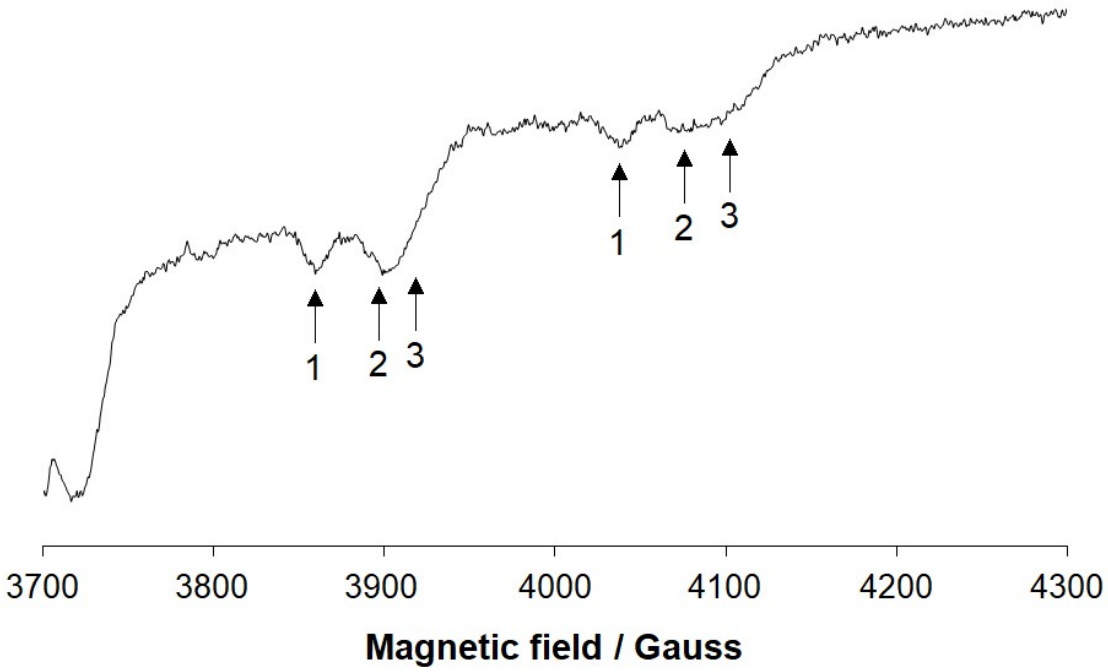


Fig. S7. High field region of the X-band anisotropic EPR spectrum recorded at pH 7.4 on a frozen solution (120 K) containing $V^{IV}O^{2+}$/IgG with molar ratio 1/1 ($V^{IV}O^{2+}$ concentration 3.0×10^{-4} M). The M_I = 5/2, 7/2 resonances of the **sites 1**, **2** and **3** are indicated. Adapted from ref. 32. The measured spin Hamiltonian parameters are: $g_z = 1.947$, $|A_z| = 167.4 \times 10^{-4}$ cm^{-1} (**site 3**), $g_z = 1.951$, $|A_z| = 163.6 \times 10^{-4}$ cm^{-1} (**site 2**), and $g_z = 1.960$, $|A_z| = 158.8 \times 10^{-4}$ cm^{-1} (**site 1**).

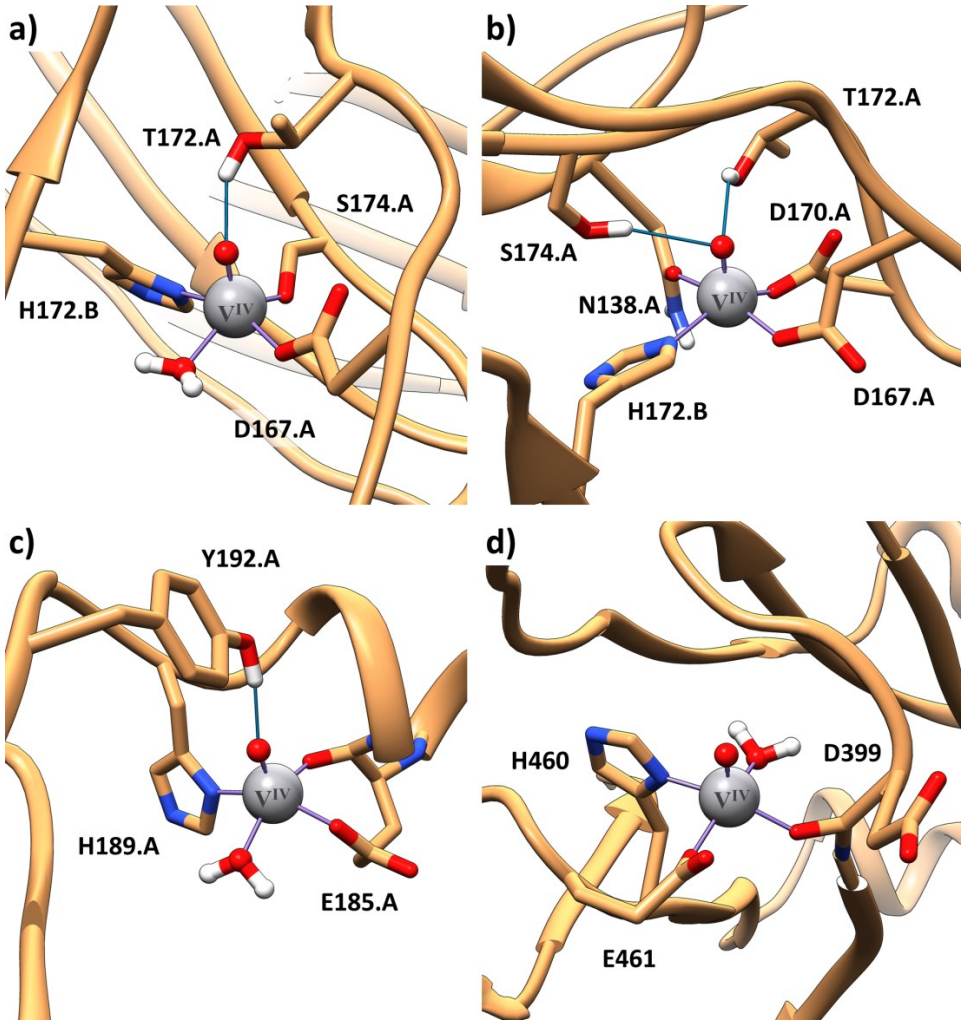


Fig. S8. Docking solutions for the binding of $V^{IV}O^{2+}$ ion to the **sites 1-3** of IgG: (a) **site 1**, (b) **site 2**, (c) **mode 3^1** and (d) **mode 3^2**.

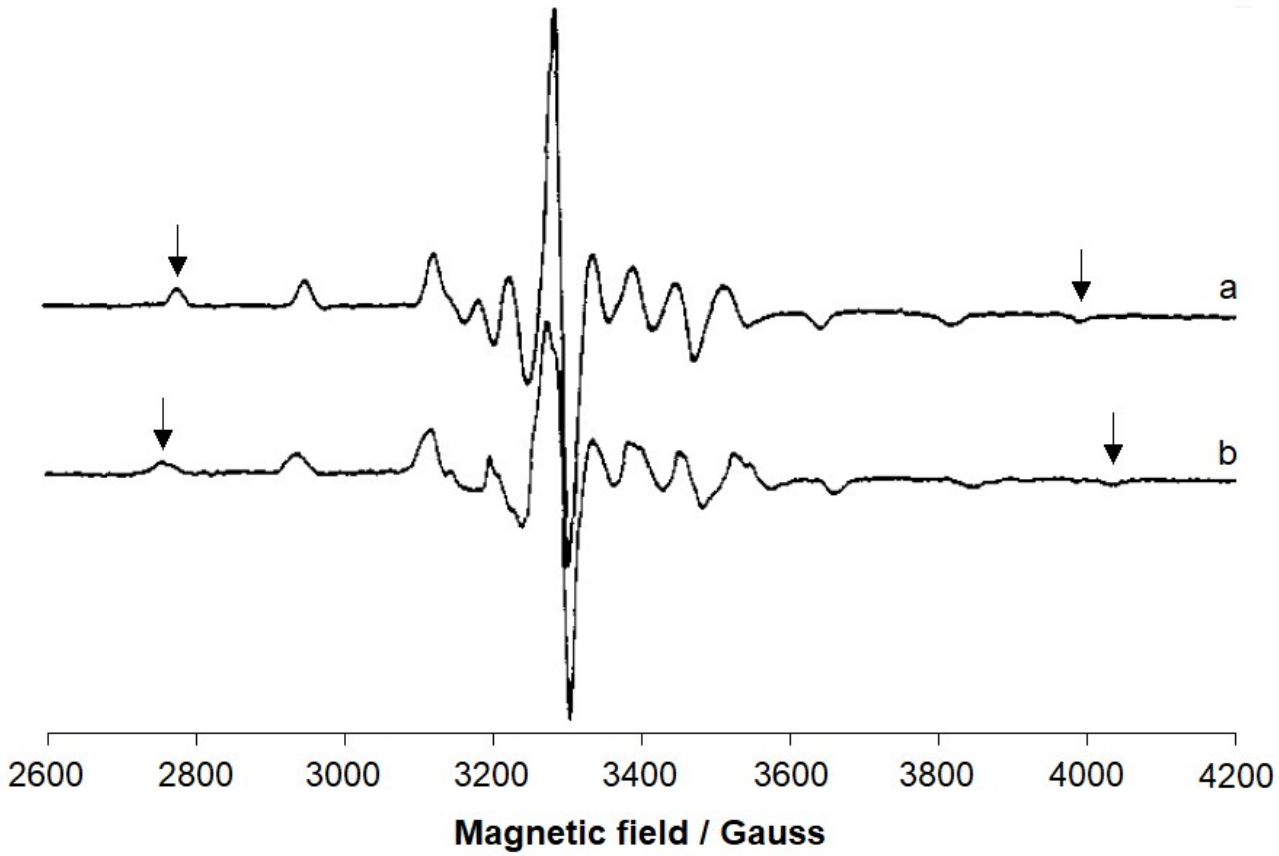


Fig. S9. X-band anisotropic EPR spectra recorded at 64 K on reduced VBrPO in H_2O: a) pH 8.4 (VBrPO concentration 140 μM); b) pH 4.2 (VBrPO concentration 226 μM). The M_I = -7/2, 7/2 resonances are indicated with the arrow. Adapted from ref. 33. The measured spin Hamiltonian parameters are: g_z = 1.948, $|A_z|$ = 160.1 × 10^{-4} cm^{-1} (pH 8.4), and g_z = 1.950, $|A_z|$ = 167.5 × 10^{-4} cm^{-1} (pH 4.2).

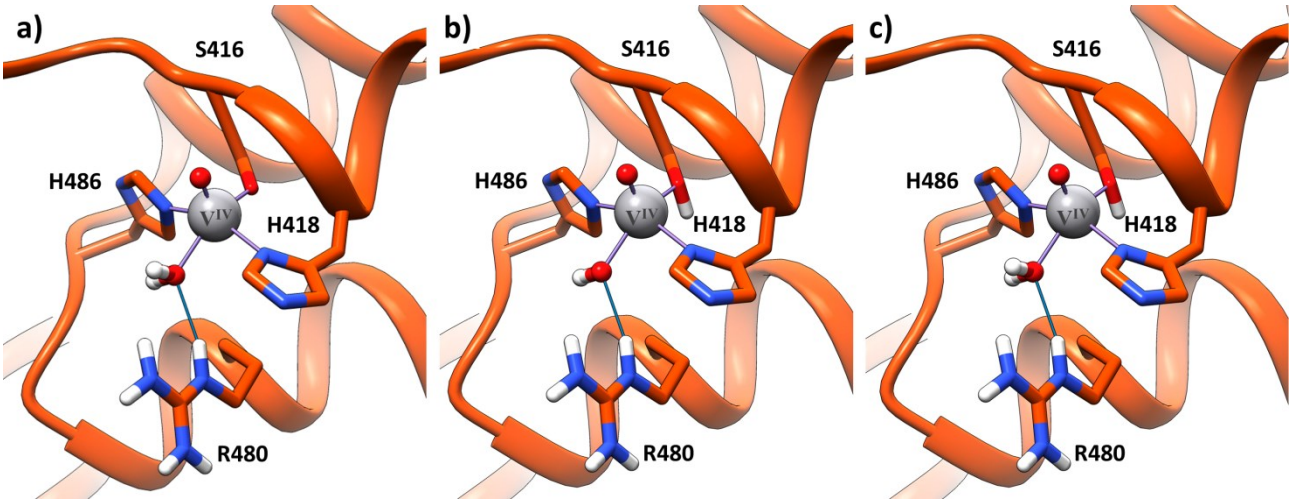


Fig. S10. Docking solutions for the binding of $V^{IV}O^{2+}$ ion to VBrPO: (a) **mode 1**, (b) **mode 2** and (c) **mode 3**.

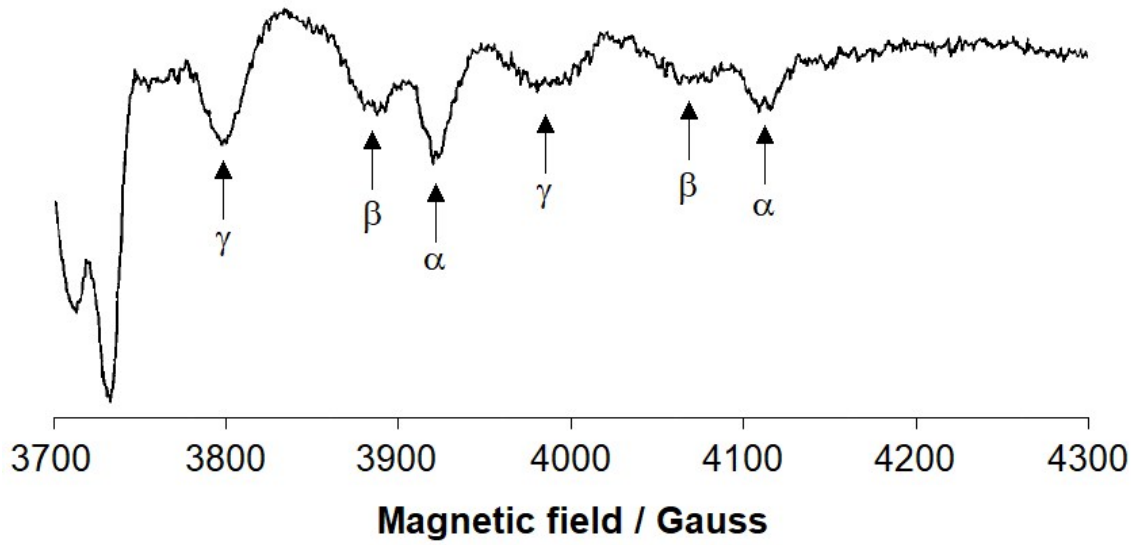


Fig. S11. High field region of the X-band anisotropic EPR spectrum recorded at pH 8.0 on frozen solutions (10 K) containing $V^{IV}O^{2+}$/IGPD with molar ratio 1/1. The M_I = 5/2, 7/2 resonances of the **sites α**, **β** and **γ** are indicated. Adapted from ref. 35. The measured spin Hamiltonian parameters are: g_z = 1.944, $|A_z|$ = 169.1 × 10^{-4} cm^{-1} (**site α**), g_z = 1.957, $|A_z|$ = 161.6 × 10^{-4} cm^{-1} (**site β**), and g_z = 1.955, $|A_z|$ = 140.6 × 10^{-4} cm^{-1} (**site γ**).

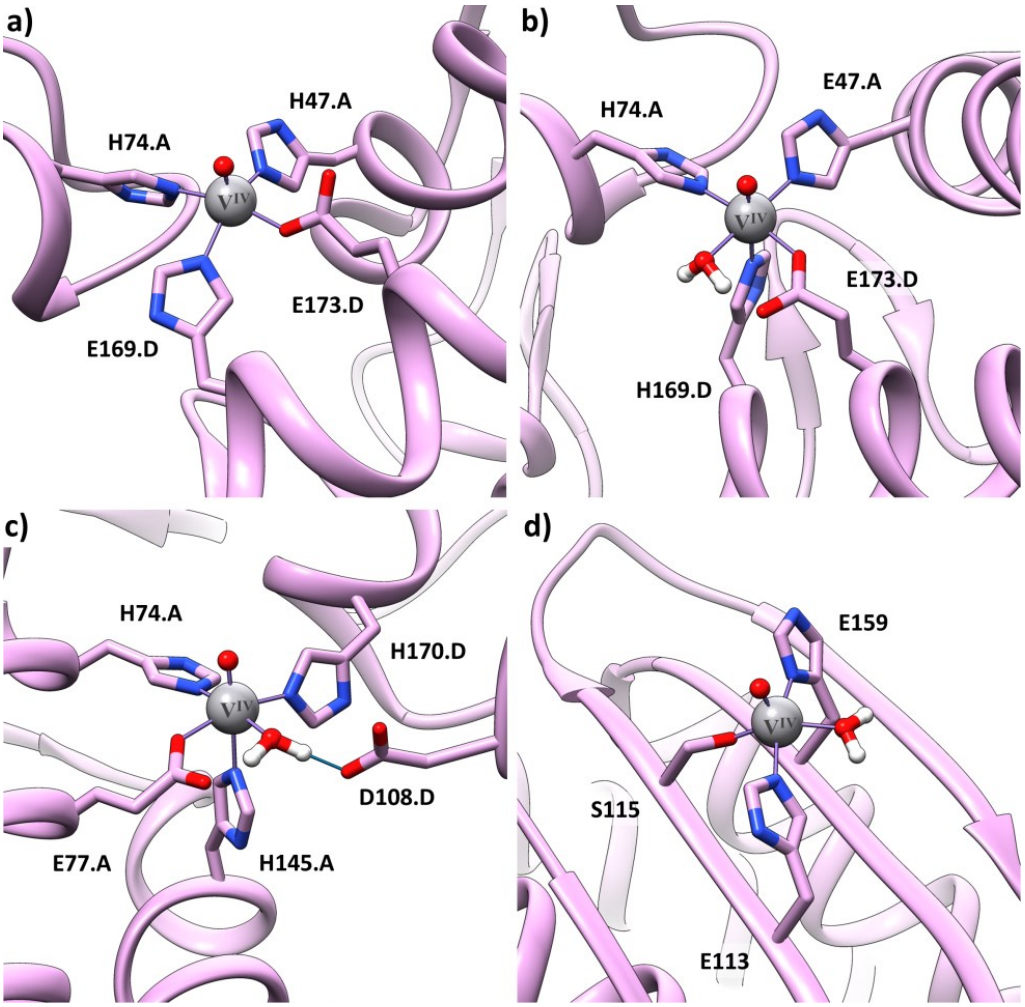


Fig. S12. Docking solutions for the binding of $V^{IV}O^{2+}$ ion to **sites α**, **β** and **γ** of $V^{IV}O^{2+}$-substituted IGPD: (a) **mode α^1**, (b) **mode α^2**, (c) **site β** and (d) N_{H159}, N_{H113}, O^-_{S115}; H_2O (unsuccessful attempt to characterize **site γ**).

6. References

1. S.-Y. Park, T. Yokoyama, N. Shibayama, Y. Shiro and J. R. H. Tame, *J. Mol. Biol.*, 2006, **360**, 690.
2. N. Noinaj, N. C. Easley, M. Oke, N. Mizuno, J. Gumbart, E. Boura, A. N. Steere, O. Zak, P. Aisen, E. Tajkhorshid, R. W. Evans, A. R. Gorringe, A. B. Mason, A. C. Steven and S. K. Buchanan, *Nature*, 2012, **483**, 53.
3. L. J. Harris, S. B. Larson, K. W. Hasel and A. McPherson, *Biochemistry*, 1997, **36**, 1581.
4. M. Weyand, H. J. Hecht, M. Kieß, M. F. Liaud, H. Vilter and D. Schomburg, *J. Mol. Biol.*, 1999, **293**, 595.
5. S. E. Glynn, P. J. Baker, S. E. Sedelnikova, C. L. Davies, T. C. Eadsforth, C. W. Levy, H. F. Rodgers, G. M. Blackburn, T. R. Hawkes, R. Viner and D. W. Rice, *Structure*, 2005, **13**, 1809.
6. (a) J. Rodríguez-Guerra, *Insilichem/gaudiview: Pre-alpha public releas, Zenodo*, 2017; (b) J. Rodríguez-Guerra Pedregal, G. Sciortino, J. Guasp, M. Municoy and J.-D. Maréchal, *J. Comput. Chem.*, 2017, **38**, 2118.
7. J.-E. Sánchez-Aparicio, L. Tiessler-Sala, V.-C. Lorea, L. Roldán-Martín, G. Sciortino and J.-D. Maréchal, *ChemRxiv*, 2020, DOI: 10.26434/chemrxiv.12668651.v1.
8. 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. L. Caricato, X., 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, T. Keith, 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. J. Adamo, J., 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 and D. J. Fox, *Gaussian 09, revision D.01*, Gaussian, Inc., Wallingford, CT, 2010.
9. A. V. Marenich, C. J. Cramer and D. G. Truhlar, *J. Phys. Chem. B*, 2009, **113**, 6378.
10. G. Jones, P. Willett, R. C. Glen, A. R. Leach and R. Taylor, *J. Mol. Biol.*, 1997, **267**, 727.
11. E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng and T. E. Ferrin, *J. Comput. Chem.*, 2004, **25**, 1605.

12. (a) G. Sciortino, E. Garribba and J.-D. Maréchal, *Inorg. Chem.*, 2019, **58**, 294; (b) G. Sciortino, D. Sanna, V. Ugone, G. Micera, A. Lledós, J.-D. Maréchal and E. Garribba, *Inorg. Chem.*, 2017, **56**, 12938; (c) G. Sciortino, J. Rodríguez-Guerra Pedregal, A. Lledós, E. Garribba and J.-D. Maréchal, *J. Comput. Chem.*, 2018, **39**, 42; (d) G. Sciortino, D. Sanna, V. Ugone, A. Lledós, J.-D. Maréchal and E. Garribba, *Inorg. Chem.*, 2018, **57**, 4456; (e) G. Sciortino, D. Sanna, V. Ugone, J. D. Marechal and E. Garribba, *Inorg. Chem. Front.*, 2019, **6**, 1561; (f) G. Sciortino, D. Sanna, V. Ugone, J.-D. Maréchal, M. Alemany-Chavarria and E. Garribba, *New J. Chem.*, 2019, **43**, 17647; (g) V. Ugone, D. Sanna, G. Sciortino, J. D. Marechal and E. Garribba, *Inorg. Chem.*, 2019, **58**, 8064.
13. S. C. Lovell, J. M. Word, J. S. Richardson and D. C. Richardson, *Proteins: Struct., Funct., Bioinf.*, 2000, **40**, 389.
14. P. E. M. Siegbahn and F. Himo, *Wiley Interdiscip. Rev. Comput. Mol. Sci.*, 2011, **1**, 323.
15. V. S. Bryantsev, M. S. Diallo and W. A. Goddard III, *J. Phys. Chem. B*, 2008, **112**, 9709.
16. (a) D. N. Chasteen, in *Biological Magnetic Resonance*, eds. L. J. J. Berliner and J. Reuben, Plenum Press, New York, 1981, vol. 3, pp. 53; (b) T. S. Smith II, R. LoBrutto and V. L. Pecoraro, *Coord. Chem. Rev.*, 2002, **228**, 1.
17. G. Micera and E. Garribba, *Dalton Trans.*, 2009, 1914.
18. G. Micera and E. Garribba, *J. Comput. Chem.*, 2011, **32**, 2822.
19. D. Rehder, *Bioinorganic Vanadium Chemistry*, John Wiley & Sons, Ltd, Chichester, 2008.
20. D. Sanna, G. Sciortino, V. Ugone, G. Micera and E. Garribba, *Inorg. Chem.*, 2016, **55**, 7373.
21. D. Sanna, V. L. Pecoraro, G. Micera and E. Garribba, *J. Biol. Inorg. Chem.*, 2012, **17**, 773.
22. G. Jones, P. Willett and R. C. Glen, *J. Mol. Biol.*, 1995, **245**, 43.
23. R. C. Glen, *J. Comput.-Aided Mol. Des.*, 1994, **8**, 457.
24. (a) S. K. Burley, H. M. Berman, C. Christie, J. M. Duarte, Z. Feng, J. Westbrook, J. Young and C. Zardecki, *Protein Sci.*, 2018, **27**, 316; (b) P. W. Rose, A. Prlić, A. Altunkaya, C. Bi, A. R. Bradley, C. H. Christie, L. D. Costanzo, J. M. Duarte, S. Dutta, Z. Feng, R. K. Green, D. S. Goodsell, B. Hudson, T. Kalro, R. Lowe, E. Peisach, C. Randle, A. S. Rose, C. Shao, Y.-P. Tao, Y. Valasatava, M. Voigt, J. D. Westbrook, J. Woo, H. Yang, J. Y. Young, C. Zardecki, H. M. Berman and S. K. Burley, *Nucleic Acids Res.*, 2017, **45**, D271; (c) H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov and P. E. Bourne, *Nucleic Acids Res.*, 2000, **28**, 235.
25. K. M. Davis, M. Altmyer, R. J. Martinie, I. Schaperdoth, C. Krebs, J. M. Bollinger and A. K. Boal, *Biochemistry*, 2019, **58**, 4218.
26. J. Rittle, M. J. Field, M. T. Green and F. A. Tezcan, *Nature Chem.*, 2019, **11**, 434.

27. D. Sanna, V. Ugone, G. Sciortino, P. Buglyo, Z. Bihari, P. L. Parajdi-Losonczi and E. Garribba, *Dalton Trans.*, 2018, **47**, 2164.
28. S. H. Knauer, O. Hartl-Spiegelhauer, S. Schwarzinger, P. Hänzelmann and H. Dobbek, *The FEBS Journal*, 2012, **279**, 816.
29. G. Sciortino, D. Sanna, G. Lubinu, J.-D. Maréchal and E. Garribba, *Chem.–Eur. J.*, 2020, DOI: 10.1002/chem.202001492.
30. D. Sanna, M. Serra, G. Micera and E. Garribba, *Inorg. Chem.*, 2014, **53**, 1449.
31. D. Sanna, G. Micera and E. Garribba, *Inorg. Chem.*, 2013, **52**, 11975.
32. D. Sanna, G. Micera and E. Garribba, *Inorg. Chem.*, 2011, **50**, 3717 and references therein.
33. E. De Boer, K. Boon and R. Wever, *Biochemistry*, 1988, **27**, 1629.
34. J. Petersen, T. R. Hawkes and D. J. Lowe, *J. Inorg. Biochem.*, 2000, **80**, 161 and references therein.
35. J. Petersen, T. R. Hawkes and D. J. Lowe, *JBIC, J. Biol. Inorg. Chem.*, 1997, **2**, 308.