RSC Advances

PAPER

View Article Online
View Journal | View Issue



Cite this: *RSC Adv.*, 2022, **12**, 5214

# Statistical analysis of $P^N$ clusters in Mo/VFe protein crystals using a bond valence method toward their electronic structures†

Chang Yuan, Wan-Ting Jin and Zhao-Hui Zhou *

Nowadays, large numbers of MoFe proteins have been reported and their crystal data obtained by X-ray crystallography and uploaded to the Protein Data Bank (PDB). By big data analysis using a bond valence method, we make conclusions based on 79 selected $P^N$ in all 119 P-clusters of 53 MoFe proteins and 10 P-clusters of 5 VFe proteins from all deposited crystallographic data of the PDB. In the condition of MoFe protein crystals, the resting state $P^N$ clusters are proposed to have the formal oxidation state of 2Fe(III)6Fe(II), hiding two oxidized electron holes with high electron delocalization. The calculations show that Fe1, Fe2, Fe5, Fe6 and Fe7 perform unequivocally as $Fe^{2+}$, and Fe3 is remarkably prone to Fe(III), while Fe4 and Fe8 have different degrees of mixed valences. For $P^N$ clusters in VFe protein crystals, Fe1, Fe2, Fe4, Fe5 and Fe6 tend to be $Fe^{2+}$, but the electron distributions rearrange with Fe7 and Fe8 being more oxidized mixed valences, and Fe3 presenting a little more reductive mixed valence than that in MoFe proteins. In terms of spatial location, Fe3 and Fe6 in P-clusters of MoFe proteins are calculated as the most oxidized and reduced irons, which have the shortest distances from homocitrate in the FeMo-cofactor and $[Fe_4S_4]$ cluster, respectively, and thus could function as potential electron transport sites. This work shows different electron distributions of $P^N$ clusters in Mo/VFe protein crystals, from those obtained from previous data from solution with excess reducing agent from which it was concluded that $P^N$ clusters are all ferrous according to Mössbauer and electron paramagnetic resonance spectra.

Received 21st November 2021
Accepted 26th January 2022

DOI: 10.1039/d1ra08507g

rsc.li/rsc-advances



## 1 Introduction

Nitrogenase is a biological enzyme that can activate the triple bond of $N_2$ to form ammonia at moderate temperature and pressure, as shown in eqn (1) below.

$$N_2 + 8e^- + 8H^+ + 16ATP \rightarrow 2NH_3 + H_2 + 16ADP + 16P_i \quad (1)$$

Given that over half of the fixed N inputs that sustain the earth's population are supplied biologically,[1] it is necessary to understand the mechanism of $N_2$ fixation. In nitrogenase, three metalloclusters participate in the catalytic process: $[Fe_4S_4]$ designated as the F-cluster, Mo*$Fe_7S_9C$(*R*-Hhomocit*) ($H_4$-homocit = homocitric acid, Hcys = cysteine, Hhis = histidine)[2–6] referred to as the FeMo-cofactor (FeMo-co) or M-cluster, and $[Fe_8S_7]$ named the P-cluster.[7,8] In an iron protein, $[Fe_4S_4]$ provides electrons along with the hydrolysis of adenosine triphosphate (ATP).[9,10] For an MoFe protein, numerous studies have confirmed that FeMo-co is the site where substrate reduction occurs,[11–14] and it has been proposed that $[Fe_8S_7]$ plays a pivotal role in transferring electrons between the iron protein and FeMo-co.[15–17] Therefore, it is essential to demonstrate the redox states of M/P-clusters so as to understand their catalytic mechanism.

Nowadays, several oxidation states of P-clusters have been found, including the resting states $P^N$, single-electron oxidized $P^{1+}$ [18] and double-electron oxidized $P^{2+}$ clusters.[19,20] The most stable structures of P-clusters observed in protein crystals are $P^N$ and $P^{2+}$. Their conformations can be transformed reversibly in the presence of reductant and oxidant.[21] With its unstable thermodynamics, $P^{1+}$ is a transient state whose crystal data was reported as PDB entry 6CDK with 60% completion.[22] However, it has the probability of being a mixture of $P^{1+}$ and $P^{2+}$ according to quantum refinement calculations.[23] Further oxidation states of $P^{3+}$ and the others have been observed, but only $P^N$, $P^{1+}$ and $P^{2+}$ were reported to be relevant to the catalytic cycle process.[20,24] In a previous report, the "Deficit-spending" model proposed that FeMo-co obtains one electron from $P^N$ at the moment an iron protein interacts with an MoFe protein. Meanwhile, $P^N$ is turned to $P^{1+}$ which is then rapidly refilled back to $P^N$ by electronic delivery from $[Fe_4S_4]$,[7] supposing no involvement of $P^{2+}$ in this mechanism. However, recent work has shown that a P-cluster performs as $P^{2+}$ while $N_2$ coordinates with FeMo-co, implying that $P^{2+}$ may play an important role in delivering

*State Key Laboratory for Physical Chemistry of Solid Surfaces, College of Chemistry and Chemical Engineering, Xiamen University, Xiamen 361005, P. R. China. E-mail: zhzhou@xmu.edu.cn*

† Electronic supplementary information (ESI) available: Details of bond distances and valence calculations of all relevant PDB entries in Tables S1–S116. See DOI: 10.1039/d1ra08507g

electrons.[25] Theoretical calculations have also proposed the catalytic involvement of $P^{2+}$ in the density of states.[26] Obviously, the roles of all oxidation states of the P-cluster are still uncertain, and the oxidation states of its irons are important for understanding the potential electron transfer sites and pathway.

The oxidation states of P-clusters have been proposed from electron paramagnetic resonance (EPR) and Mössbauer spectra,[18,20,27–29] which enumerated all kinds of signals and possible spin states of irons in $P^N$, $P^{1+}$ and $P^{2+}$ clusters. These early studies indicate that the resting state $P^N$ is all-ferrous.[29,30] $P^{1+}$ is the one-electron oxidized state of $P^N$, and correspondingly, $P^{2+}$ is commonly considered to come from double-electron oxidation.[18] Later, X-ray crystallography revealed three different conformations of P-clusters as $P^N$, $P^{2+}$ [8] and $P^{1+}$.[22] Magnetic circular dichroism (MCD) was also applied to suggest the capability of electron delivery by the P-cluster.[31] Nowadays, theoretical calculations with density functional theory (DFT)[23] have provided the viewpoint that Fe6 and Fe7 are the most oxidized irons, and many-electron quantum wavefunction simulations illustrate the possible spin states and electronic structure of the P-cluster in plenty of aspects based on crystal structures.[26] However, the detailed valence assignment of the P-cluster has not been analyzed specifically and agreed thus far. It will be helpful to extrapolate which irons play major roles in transferring electrons as the deficit-spending model describes. This inspired us to analyze the oxidation states of each iron in the P-cluster from the point of view of protein structures by the bond valence method.

The bond valence method was first used to analyze inorganic crystal structures[32–36] and was gradually applied in other fields.[37–41] It can be traced back historically to a proposal by Pauling.[42] It is a classic and valid approach for assessing the charge between a metal atom and its bound coordinated atoms, and has proved an effective method to evaluate the electron density in a delocalized system[43] and the oxidation states of metals in a metalloprotein.[44–46] Up to now, the crystallographic structures of MoFe proteins deposited in the PDB have supplied sufficient bond data for M- and P-clusters. We have used this method to evaluate the valences of molybdenum(III) and vanadium(III) in FeMo/V-cofactors and the corresponding oxidation states of seven irons.[46] In this work, we try to use the bond valence method to analyze the oxidation states of irons in P-clusters and to explore the function of $P^N$ in electron delivery between $[Fe_4S_4]$ and FeMo-cofactor from different oxidized iron sites.

## 2 Calculation method

Bond valence sums (BVSs) were calculated using eqn (2), as shown below:

$$S_i = \sum_j \exp\left[\left(R_0 - r_{ij}\right)/B\right] \quad (2)$$

$$S_t = \sum_i S_i \quad (3)$$

$S_i$ represents the calculated bond valence sum of each iron, and $S_t$ in eqn (3) refers to the calculated valence sum of all eight Fe1–Fe8 irons (abbreviated to 8Fe below) in the P-cluster. The term $r_{ij}$ is the bond distance between metal $i$ (Fe) and ligand $j$ (S/O/N/C), and $B$ is commonly related to the softness of the bond[47] and used as a constant equal to 0.37 Å.[32] $R_0$ is a constant for a specific bond and varies with the assumed metal valence ($Fe^{n+}$) and coordinated atom as shown in Table 1. The values of $R_0$ can be viewed on the web.[48–50]

For P-clusters in Mo/VFe proteins, $r_{ij}$ is measured from crystal structures of Mo/VFe proteins deposited in the RCSB Protein Data Bank (PDB), which presently contains data on 119 P-clusters in 53 MoFe proteins and 10 P-clusters in 5 VFe proteins. For 14 $P^N$-type model compounds, $r_{ij}$ is acquired from the Cambridge Crystallographic Data Centre (CCDC) and measured by Pymol. The valences of Fe atoms were calculated by using $R_0$ ($+n$) that corresponds to $Fe^{2+}$ and $Fe^{3+}$ coordinated with different ligands. All $r_{ij}$ values and their resulting $S_i$ values are estimated to the third decimal place. Detailed bond valence calculations of all PDB entries and model compounds are given in Tables S4–S116.†

As an evaluation index, the absolute deviation $|d|$ ($d = S_i - n$, $n$ is expected valence +2 or +3 for Fe atoms) represents the discrepancy between the calculated and expected valences, showing the fitting effects of $R_0$ (+2) and $R_0$ (+3). Due to the electron delocalization in P-clusters,[54] some valences calculated with $R_0$ (+2) and $R_0$ (+3) show similar values of $|d|$. In this situation, iron valences can properly be regarded as mixed valence rather than integral valence. When the differences in $|d|$ between $R_0$ (+2) and $R_0$ (+3) are distinct, the oxidation states of iron atoms should be assigned as the valence which has the smaller and more suitable value of $|d|$. The calculated valence sums of Fe1–Fe8 ($S_t$) also contrast with the assumed 8Fe all-ferrous valences which sum to "16" and all-ferric valences which sum to "24". The resulting values of $|d|$ shown in Fig. 2a and 4a imply the possible numbers of $Fe^{3+}$ covered in this electron delocalization system and the total electrons reserved in P-clusters.

As shown in Fig. 1b, in $P^{1+}$, after single-electron oxidation of $P^N$, Fe6 moves away from the central hexa-coordinated S1 atom and coordinates with the O atom in nearby amino acids such as *Ser*β188 in the MoFe protein of *Azotobacter vinelandii* (*Av*). With further one-electron oxidation, as shown in Fig. 1c, Fe5 in $P^{2+}$ leaves the central S1 and bonds with the backbone amide N atom of *Cys*α88 in *Av*. Thus, the bonds of Fe5–S1 and Fe6–S1 are disconnected and Fe5–N and Fe6–O are formed in $P^{1+}$ and $P^{2+}$ respectively. Due to there being only a small number of deposited PDB entries containing $P^{2+}$ or superposition of $P^{N/2+}$, where two oxidation states coexist in P-clusters, we focus on researching the abundant data on $P^N$ and pick out the part

Table 1 The values of $R_0$ corresponding to different types of bonds in P-clusters and their simulations

| M–L bonds | $R_0$ (Å) | M–L bonds | $R_0$ (Å) |
|---|---|---|---|
| $Fe^{2+}$–S | 2.120 [51] | $Fe^{3+}$–S | 2.149 [32] |
| $Fe^{2+}$–O | 1.715 [52] | $Fe^{3+}$–O | 1.749 [52] |
| $Fe^{2+}$–N | 1.769 [53] | $Fe^{3+}$–N | 1.815 [53] |

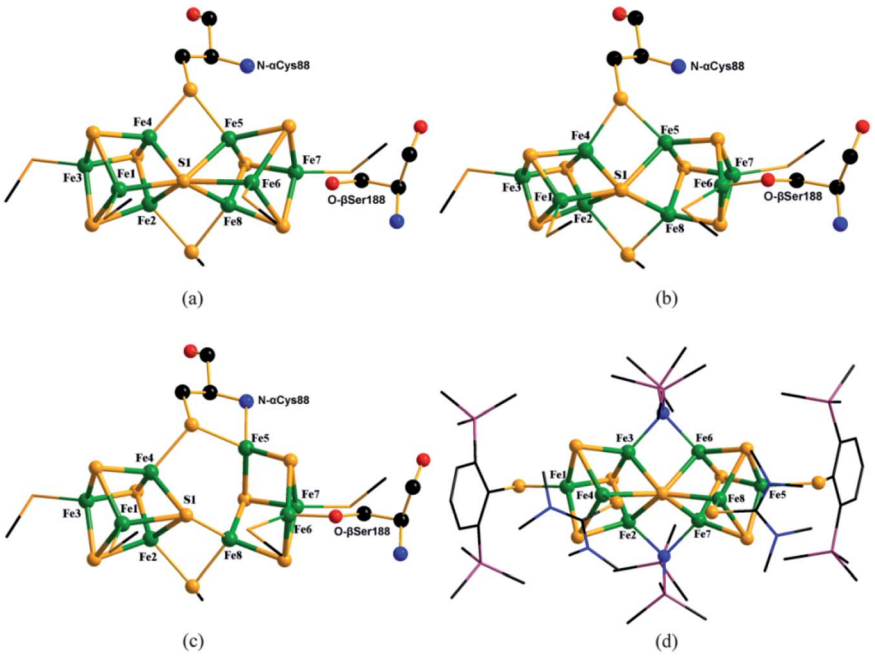


Fig. 1 Molecular structures of $P^N$ (a), $P^{1+}$ cluster (b) (PDB entry: 6CDK) and $P^{2+}$ (c) (PDB entry: 3U7Q) in MoFe proteins, and model compound of $P^N$ cluster (d) (CSD refcode: MUFQUA). Colors are Fe in green, S in yellow, O in red, N in blue, Si in plum and C in black.

relating to $P^N$ in the superposition structure, such as 3U7Q which is in the $P^{N/2+}$ mixed state. The only data assigned as $P^{1+}$ in MoFe protein (PDB entry: 6CDK) was also abandoned, because of the transient existence of $P^{1+}$, dubious Fe–Fe and Fe–S bond distances[23] and incompleteness of the crystal data.[22]

To achieve reasonable big data analysis of bond valences, on the one hand, we carefully picked out these valid data in terms of protein structures. The P-clusters (PDB entries: 1M34, 5CX1, 5VQ4) which were not clearly assigned as $P^{N/1+}$ in their conformational structures, were recognized as $P^N$ according to the reductive environments of protein purification, as Table S2† illustrates. Those data from unreasonable models (PDB entries: 1MIO, 3K1A) or structures containing deficient atoms (PDB entry: 6O7S) were abandoned, as shown in Table S3.† The protein data with unusually short Fe–S bonds (PDB entries: 1M1Y, 2AFI, 6BBL, 6OP1, 6OP2, 6OP4) which result in faulty $S_t$ of 8Fe by using $R_0$ (+2) above or approximating to 24 were not included in the analysis.

On the other hand, the bond valence method is empirical and has a great demand for high-precision data for bond distances. Thus, to deduce a more reasonable bond valence for each iron $S_i$ from all calculated P-clusters, it is crucial to select a suitable weighting formulation that includes resolutions of PDB data as weighting factors. Considering that a smaller value of the resolution $A_i$ should have a higher weight $w_i$, each $w_i(A_i)$ of PDB entries should be set as a function of the reciprocal resolution. It is appropriate to modify the inverse distance weighted (IDW) interpolation method to set a series of weights:

$$w_i = A_i^{-p} \Big/ \sum_i^N (A_i^{-p}) \quad (4)$$

$$\overline{S_w} = \sum_i^N S_i w_i \Big/ \sum_i^N w_i = \sum_i^N S_i w_i \quad (5)$$

Univariate IDW interpolation is used widely by earth scientists in geochemistry.[55–59] The character of eqn (4) is the same as the basic principle of IDW:[60] the calculated bond valences $S_i$ from smaller values of resolution $A_i$ such as 3U7Q have a greater influence on the weighted average valence than those from larger values of resolution, such as 1M34. $\overline{S_w}$ in eqn (5) is the weighted average of the calculated valences $S_i$ of different Fe atoms from all analyzed $P^N$ clusters. $N$ is the number of samples, including 69 $P^N$ clusters of MoFe proteins and 10 $P^N$ clusters of VFe proteins. $\sum_i^N w_i$ is actually equal to 1 in this equation. Parameter $p$ is an exponential parameter usually set to around 0.5–3.0 by the user.[55,61] By comparing different $w_i$ by using different $p$ values from a small amount of VFe protein data as shown in Table 2, we adopt $p = 1$,[62] which also

Table 2 The values of $w_i$ obtained by adopting different $p$ in P-clusters of VFe proteins along with different resolutions

| | | $w_i$ | | | |
|---|---|---|---|---|---|
| PDB entries | Res (Å) | $p = 0.5$ | $p = 1$ | $p = 2$ | $p = 3$ |
| 7ADR | 1.00 | 0.106 | **0.112** | 0.123 | 0.135 |
| 7ADY | 1.05 | 0.103 | **0.106** | 0.112 | 0.116 |
| 7AIZ | 1.05 | 0.103 | **0.106** | 0.112 | 0.116 |
| 6FEA | 1.20 | 0.097 | **0.093** | 0.086 | 0.078 |
| 5N6Y | 1.35 | 0.091 | **0.083** | 0.068 | 0.055 |

generate balanced $w_i$ for MoFe protein data of all various resolutions as in Fig. S1.†

Detailed supplemental illustrations of IDW, PDB classifications and bond valence calculations of all the above adopted and abandoned P-clusters can be seen in the ESI.†

# 3 Results and discussion

## 3.1. Criteria for valence assignment from 14 $P^N$ model compounds

Since BVS is an empirical method and considering the electron delocalization existing in an Fe–S cluster system, it is necessary to set a value of $D$ as a valence assignment criterion to identify different integral and mixed valences. We consider that the new criterion applied to P-clusters could refer to the BVS results of the $P^N$ model compounds by using the same $R_0$ parameter. 14 model compounds of $P^N$ have been selected from the Cambridge Crystallographic Data Centre (CCDC) and calculated by BVS. As shown in Fig. 1d, these model compounds have the same coordinated sites as natural $P^N$. From the viewpoint of electronic structures, the $P^N$ model compounds have the same electron delocalization as natural $P^N$, and are calculated by BVS as $2Fe^{3+}6Fe^{2+}$, which is consistent with previous reports.[63–65] After optimal selections of two calculated valences by using $R_0$ (+2) and $R_0$ (+3), two Fe(III) and irons with mixed valences commonly exist in model compounds, as shown in Table 3. In a previous article about model compounds, the terminal Fe1 and Fe5 of each doublet were proved to be $Fe^{3+}$,[64–67] which is completely consistent with the calculated values.

Although Thorp had stated that BVS values calculated from Brown's distance were reliable to ±0.25 units,[44,53] we think that the electron delocalization in P-clusters compared with inorganic crystals requires a larger error-tolerance interval. In Table 3, we presumed three acceptable $D$ values of 0.25, 0.3 and 0.35 as valence assignment criteria to compare the different valence assignments of iron. When the absolute deviation $|d|$ between the calculated valence and the assumed valence is less than the valence assignment criterion $D$, such as 0.25, the irons are assumed to have valence $Fe^{2+/3+}$; otherwise they ought to be assigned as having uncertainly mixed valence like $Fe^{2.5+}$.

In Table 3, Fe4 and Fe8 sometimes show the character of mixed valence, which may be due to their spatial locations being adjacent to ferric Fe1 and Fe5. When choosing $D$ as 0.25, several CSD entries like MUFREL and NIFWOQ show that Fe2/3/6/7 are assigned as mixed valences, which are obviously big deviations from the actual conclusion of $2Fe^{3+}6Fe^{2+}$. If we set our sights on $D = 0.35$, this high error tolerance leads to another completely unreasonable conclusion that three or four irons(III) exist in $P^N$ model compounds (CSD codes: MUFQOU and WUZDAW).

In contrast, adopting 0.3 as the valence assignment criterion $D$, we find that the calculated and observed values are in good agreement, except for one or two irons with mixed valence. As discussed above, it is more credible for P-clusters to take a valence assignment criterion $D = 0.3$, according to the deviation calibration with the 14 most structurally similar $P^N$ model compounds. Thus, the following discussions about valence distributions of $P^N$ in Mo/VFe nitrogenases are based on this adopted criterion $D$.

## 3.2. Valence analyses of $P^N$ clusters in MoFe proteins

Fig. 2 shows the absolute deviations $|d|$ of all 8 irons and each iron between calculated and assumed valences of $P^N$ in the resting state at a resolution of 2.3 Å. Detailed calculated results of each iron are shown in Table 4. In Fig. 2a, the discrepancies in the total valences of 8Fe between groups of $R_0$ (+2) (black) and $R_0$ (+3) (red) are obvious within a resolution of 1.6 Å. The weighted average value of $|d|$ in the group of $R_0$ (+2) is 2.15 if we assume $P^N$ is all-ferrous, and the corresponding $|d|$ in the group of $R_0$ (+3) is 4.47 when $P^N$ is assumed to be all-ferric. The smaller deviation calculated from all-ferrous parameters indicates that $P^N$ has a strong reductive property as previously reported[22] and could undertake the function of delivering electrons.Fig. 2(b–i)

Table 3 The optimal calculated bond valences of Fe atoms in 14 model compounds obtained by using $R_0$ (+2) or $R_0$ (+3). Numbers of irons assigned different valences by using different presumed $D$ are shown below

| CSD Refcodes | Fe(1) | Fe(2) | Fe(3) | Fe(4) | Fe(5) | Fe(6) | Fe(7) | Fe(8) | Sum | $n(+3) : n(+2) : n(\|d\| > 0.25)$ | $n(+3) : n(+2) : n(\|d\| > 0.3)$ | $n(+3) : n(+2) : n(\|d\| > 0.35)$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| DUGNEZ | 2.809 | 2.217 | 2.213 | 1.944 | 2.826 | 2.196 | 2.243 | 1.925 | 18.373 | 2 : 6 : 0 | 2 : 6 : 0 | 2 : 6 : 0 |
| MUFPOT | 2.814 | 2.234 | 2.145 | 2.608 | 2.842 | 2.249 | 2.140 | 2.606 | 19.639 | 2 : 4 : 2 | 2 : 4 : 2 | 2 : 4 : 2 |
| MUFPUZ | 2.804 | 2.217 | 2.146 | 2.322 | 2.804 | 2.217 | 2.146 | 2.322 | 18.978 | 2 : 4 : 2 | 2 : 4 : 2 | 2 : 6 : 0 |
| MUFQAG | 2.813 | 2.144 | 2.239 | 2.392 | 2.794 | 2.160 | 2.216 | 2.599 | 19.357 | 2 : 4 : 2 | 2 : 4 : 2 | 2 : 4 : 2 |
| MUFQEK | 2.809 | 2.232 | 2.169 | 2.628 | 2.886 | 2.277 | 2.153 | 2.607 | 19.760 | 2 : 3 : 3 | 2 : 4 : 2 | 2 : 4 : 2 |
| MUFQIO | 2.841 | 2.241 | 2.162 | 2.362 | 2.841 | 2.241 | 2.162 | 2.362 | 19.211 | 2 : 4 : 2 | 2 : 4 : 2 | 2 : 4 : 2 |
| MUFQOU | 2.898 | 2.212 | 2.201 | 2.660 | 2.764 | 2.228 | 2.185 | 2.347 | 19.495 | 2 : 4 : 2 | 2 : 4 : 2 | 3 : 5 : 0 |
| MUFQUA | 3.028 | 2.196 | 2.190 | 2.640 | 2.960 | 2.280 | 2.236 | 2.624 | 20.153 | 2 : 3 : 3 | 2 : 4 : 2 | 2 : 4 : 2 |
| MUFRAH | 2.983 | 2.206 | 2.198 | 2.267 | 2.932 | 2.226 | 2.236 | 2.361 | 19.409 | 2 : 4 : 2 | 2 : 5 : 1 | 2 : 5 : 1 |
| MUFREL | 3.000 | 2.268 | 2.188 | 2.335 | 3.004 | 2.290 | 2.350 | 2.166 | 19.601 | 2 : 2 : 4 | 2 : 4 : 2 | 2 : 6 : 0 |
| NIFWOQ | 2.855 | 2.139 | 2.267 | 2.350 | 2.855 | 2.139 | 2.267 | 2.350 | 19.223 | 2 : 2 : 4 | 2 : 4 : 2 | 2 : 6 : 0 |
| NIFWUW | 2.805 | 2.192 | 2.118 | 2.393 | 2.752 | 2.288 | 2.220 | 2.634 | 19.402 | 2 : 3 : 3 | 2 : 4 : 2 | 2 : 4 : 2 |
| NIFXAD | 2.840 | 2.230 | 2.196 | 2.620 | 2.840 | 2.230 | 2.196 | 2.620 | 19.772 | 2 : 4 : 2 | 2 : 4 : 2 | 2 : 4 : 2 |
| WUZDAW | 2.829 | 2.239 | 2.156 | 2.677 | 2.829 | 2.239 | 2.156 | 2.677 | 19.803 | 2 : 4 : 2 | 2 : 4 : 2 | 4 : 4 : 0 |
| Average | 2.866 | 2.212 | 2.185 | 2.443 | 2.852 | 2.233 | 2.208 | 2.443 | 19.441 | 2 : 4 : 2 | 2 : 4 : 2 | 2 : 4 : 2 |

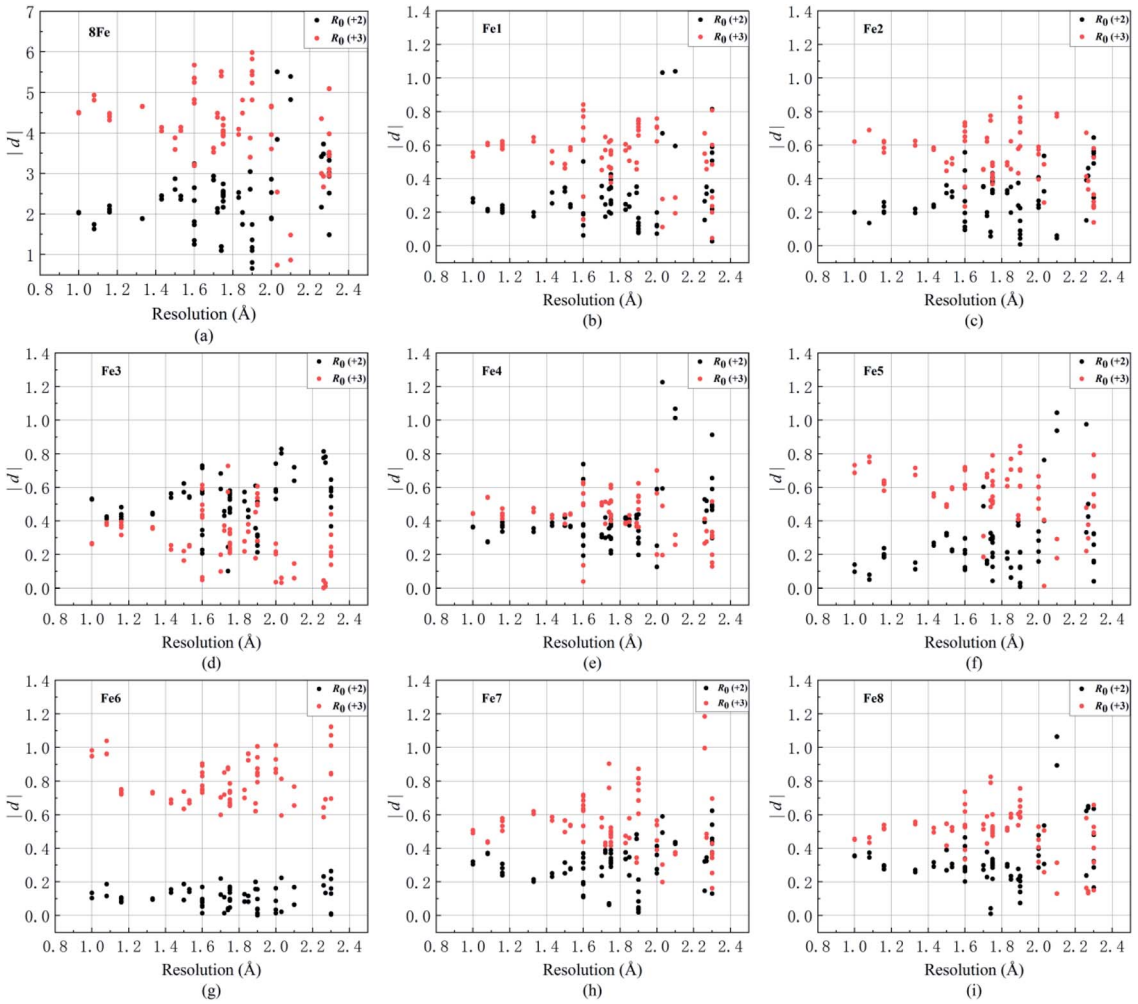


Fig. 2 The values of $|d|$ of (a) total eight irons and (b–i) Fe1 to Fe8 in $P^N$ of FeMo proteins in terms of $R_0$ (+2) (black) and $R_0$ (+3) (red) respectively. Resolution is on the horizontal axis and the value of $|d|$ is on the vertical axis. Some unusual $P^N$ data (PDB entries: 1M1Y, 2AFI, 6BBL, 6OP1, 6OP2, 6OP4) are excluded.

show the absolute deviations $|d|$ between calculated and presumed valences of each Fe, by using the parameters of $R_0$ (+2) and $R_0$ (+3) at different resolutions.

For Fe1, it is clear that most deviations $|d|$ of the two groups are separated by a value of 0.4 in Fig. 2b, with a weighted average value of 0.29 for group $R_0$ (+2) and 0.54 for $R_0$ (+3) within 2.3 Å resolution. As can be seen from Table 4, the difference $|d|$ between the BVS of $R_0$ (+2) and +2 valence even drops to 0.25 within a resolution of 1.6 Å, with both $|d|$ of $R_0$ (+2) below the adopted $D$, which implies that Fe1 tends to be ferrous and the assignment of Fe(II) might be appropriate. For Fe2, the weighted average value of $|d|$ is 0.27 for $R_0$ (+2) and 0.55 for $R_0$ (+3). Besides, Fe2 is prone to be iron(II) whose $|d|$ is 0.24, which is obviously below $D = 0.30$ in group $R_0$ (+2) within the resolution of 1.6 Å. For Fe3, it is the iron which is most prone to be Fe(III) compared with other irons, where its weighted average values of $|d|$ for $R_0$ (+2) and $R_0$ (+3) are 0.51 and 0.30, respectively. However, compared with Fe3 which possesses a smaller deviation $|d|$ of $R_0$ (+3) than $R_0$ (+2) in the overall data, the $|d|$ of the two groups $R_0$ (+2) and $R_0$ (+3) for Fe4 partly overlap and are tangled up over the whole range of resolutions, where its weighted averages $|d|$ are 0.40 and 0.41 for $R_0$ (+2) and $R_0$ (+3). This implies that Fe4 has a strong mixed valence character[29,68] with a more oxidized state than Fe(II), and the valence assignment of Fe4 could not be defined.

In the same way, Fe5, Fe6 and Fe7 are more inclined to be $Fe^{2+}$ rather than $Fe^{3+}$ in Fig. 2. The corresponding weighted average values of $|d|$ of $R_0$ (+2) and $R_0$ (+3) are 0.25 and 0.58 for Fe5, 0.04 and 0.80 for Fe6, and 0.28 and 0.54 for Fe7, respectively. The deviations 0.18 and 0.05 for group $R_0$ (+2) of Fe5 and Fe6 are even smaller than 0.2 within a resolution of 1.6 Å. However, unlike Fe5 or Fe6, as shown in Fig. 2f–g, the groups of $R_0$ (+2) and $R_0$ (+3) for Fe7 are separated by $D$ but are not so distinct, which implies that its electrons are not well localized. The $|d|$ values of Fe8 for $R_0$ (+2) groups are 0.31 and 0.32, respectively, no matter whether the data are for resolutions below 1.6 Å or 2.3 Å. Referring to the aforementioned valence assignment criterion $|D|$ of 0.3, Fe8 and Fe4 are alike and should be regarded as mixed valence, but Fe4 has a tendency to be more oxidative than Fe8. From Table 4, the weighted average

Table 4 The bond valence sums calculated with $R_0$ for $Fe^{2+}$ and $Fe^{3+}$ in valid P-cluster data, respectively, including 69 $P^N$ of 30 MoFe proteins and 10 $P^N$ of 5 VFe proteins. PDB entries with $P^{N/2+}$ superposition are focused on the part of $P^N$

| | | $R_0$ (+2) | | | | | | | | | $R_0$ (+3) | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| PDB codes | Res (Å) | Fe1 | Fe2 | Fe3 | Fe4 | Fe5 | Fe6 | Fe7 | Fe8 | $S_t$(8Fe) | Fe1 | Fe2 | Fe3 | Fe4 | Fe5 | Fe6 | Fe7 | Fe8 | $S_t$(8Fe) |
| **$P^N$ in MoFe proteins** | | | | | | | | | | | | | | | | | | | |
| **3U7Q**[4] | **1.00** | **2.259** | **2.200** | 2.527 | **2.365** | **2.097** | **1.897** | **2.320** | **2.356** | 18.022 | 2.444 | 2.379 | **2.734** | 2.557 | 2.268 | 2.052 | 2.510 | 2.548 | 19.492 |
| (Modeled as $P^N$) | | **2.282** | **2.199** | 2.532 | **2.362** | **2.139** | **1.866** | **2.305** | **2.352** | 18.037 | 2.468 | 2.378 | **2.738** | 2.555 | 2.313 | 2.018 | 2.493 | 2.544 | 19.508 |
| 4WES[72] | 1.08 | 2.218 | 2.135 | 2.426 | 2.273 | 2.050 | 1.814 | 2.367 | 2.345 | 17.627 | 2.399 | 2.309 | 2.623 | 2.458 | 2.217 | 1.962 | 2.560 | 2.536 | 19.064 |
| (Modeled as $P^N$) | | 2.207 | 2.135 | 2.413 | 2.276 | 2.079 | 1.884 | 2.372 | 2.373 | 17.741 | 2.387 | 2.309 | 2.610 | 2.462 | 2.249 | 2.038 | 2.565 | 2.567 | 19.187 |
| 1M1N[3] | 1.16 | 2.198 | 2.197 | 2.426 | 2.364 | 2.201 | 2.079 | 2.281 | 2.298 | 18.043 | 2.377 | 2.376 | 2.624 | 2.556 | 2.381 | 2.248 | 2.467 | 2.486 | 19.515 |
| | | 2.225 | 2.234 | 2.409 | 2.376 | 2.182 | 2.098 | 2.239 | 2.294 | 18.057 | 2.407 | 2.416 | 2.606 | 2.569 | 2.360 | 2.269 | 2.421 | 2.481 | 19.530 |
| | | 2.211 | 2.203 | 2.440 | 2.336 | 2.237 | 2.106 | 2.308 | 2.275 | 18.117 | 2.392 | 2.383 | 2.639 | 2.527 | 2.420 | 2.278 | 2.496 | 2.461 | 19.594 |
| | | 2.241 | 2.260 | 2.482 | 2.392 | 2.193 | 2.084 | 2.254 | 2.293 | 18.198 | 2.423 | 2.444 | 2.684 | 2.587 | 2.372 | 2.254 | 2.437 | 2.480 | 19.682 |
| 7JRF[73] | 1.33 | 2.199 | 2.220 | 2.447 | 2.334 | 2.112 | 2.094 | 2.215 | 2.258 | 17.879 | 2.378 | 2.401 | 2.647 | 2.524 | 2.285 | 2.265 | 2.395 | 2.442 | 19.338 |
| | | 2.175 | 2.196 | 2.439 | 2.355 | 2.151 | 2.100 | 2.201 | 2.271 | 17.888 | 2.352 | 2.375 | 2.638 | 2.547 | 2.326 | 2.272 | 2.380 | 2.456 | 19.347 |
| 4TKU[74] | 1.43 | 2.317 | 2.244 | 2.563 | 2.372 | 2.269 | 2.136 | 2.231 | 2.316 | 18.448 | 2.436 | 2.414 | 2.745 | 2.584 | 2.437 | 2.331 | 2.435 | 2.478 | 19.860 |
| | | 2.253 | 2.232 | 2.538 | 2.389 | 2.253 | 2.155 | 2.251 | 2.291 | 18.363 | 2.506 | 2.427 | 2.772 | 2.565 | 2.454 | 2.310 | 2.413 | 2.505 | 19.952 |
| 4TKV[74] | 1.50 | 2.347 | 2.314 | 2.571 | 2.420 | 2.326 | 2.187 | 2.315 | 2.389 | 18.869 | 2.538 | 2.503 | 2.781 | 2.617 | 2.516 | 2.365 | 2.504 | 2.584 | 20.408 |
| | | 2.323 | 2.361 | 2.623 | 2.371 | 2.314 | 2.092 | 2.251 | 2.269 | 18.603 | 2.513 | 2.554 | 2.837 | 2.564 | 2.502 | 2.262 | 2.434 | 2.454 | 20.120 |
| 5BVH[75] | 1.53 | 2.231 | 2.291 | 2.545 | 2.364 | 2.220 | 2.140 | 2.280 | 2.288 | 18.360 | 2.413 | 2.478 | 2.753 | 2.557 | 2.401 | 2.315 | 2.466 | 2.475 | 19.857 |
| | | 2.246 | 2.323 | 2.538 | 2.369 | 2.229 | 2.155 | 2.275 | 2.307 | 18.442 | 2.429 | 2.512 | 2.745 | 2.562 | 2.410 | 2.331 | 2.461 | 2.495 | 19.946 |
| 1QGU[76] | 1.60 | 2.193 | 2.144 | 2.206 | 2.308 | 2.123 | 2.059 | 2.110 | 2.202 | 17.344 | 2.372 | 2.319 | 2.386 | 2.497 | 2.296 | 2.227 | 2.282 | 2.381 | 18.759 |
| | | 2.061 | 2.095 | 2.227 | 2.193 | 2.112 | 2.085 | 2.187 | 2.288 | 17.247 | 2.229 | 2.265 | 2.408 | 2.372 | 2.284 | 2.255 | 2.365 | 2.475 | 18.653 |
| 1QH8 [76] | 1.60 | 2.122 | 2.112 | 2.316 | 2.253 | 2.108 | 1.987 | 2.369 | 2.465 | 17.731 | 2.295 | 2.285 | 2.504 | 2.437 | 2.280 | 2.148 | 2.563 | 2.666 | 19.177 |
| (Modeled as $P^N$) | | 2.187 | 2.196 | 2.345 | 2.317 | 2.123 | 1.948 | 2.281 | 2.413 | 17.810 | 2.366 | 2.375 | 2.536 | 2.506 | 2.297 | 2.107 | 2.467 | 2.610 | 19.262 |
| 5BVG[75] | 1.60 | 2.192 | 2.349 | 2.579 | 2.380 | 2.296 | 2.169 | 2.344 | 2.337 | 18.647 | 2.192 | 2.349 | 2.579 | 2.380 | 2.296 | 2.169 | 2.344 | 2.337 | 18.647 |
| | | 2.158 | 2.266 | 2.565 | 2.373 | 2.295 | 2.095 | 2.315 | 2.263 | 18.329 | 2.158 | 2.266 | 2.565 | 2.373 | 2.295 | 2.095 | 2.315 | 2.263 | 18.329 |
| 6OP3 [77] | 1.60 | 2.630 | 2.557 | 2.729 | 2.649 | 2.206 | 2.077 | 2.117 | 2.275 | 19.241 | 2.844 | 2.766 | 2.952 | 2.865 | 2.386 | 2.247 | 2.289 | 2.461 | 20.810 |
| | | 2.503 | 2.449 | 2.715 | 2.738 | 2.225 | 2.098 | 2.197 | 2.284 | 19.209 | 2.707 | 2.648 | 2.936 | 2.962 | 2.407 | 2.269 | 2.376 | 2.470 | 20.775 |
| **$\overline{S_w}$ (in 1.6 Å)** | | 2.247 | 2.239 | 2.482 | 2.368 | 2.183 | 2.048 | 2.271 | 2.314 | 18.152 | 2.418 | 2.409 | 2.669 | 2.547 | 2.348 | 2.203 | 2.443 | 2.490 | 19.526 |
| **Weighted $\|d\|$** | | 0.247 | 0.239 | 0.482 | 0.368 | 0.183 | 0.048 | 0.271 | 0.314 | 2.152 | 0.582 | 0.591 | 0.331 | 0.453 | 0.652 | 0.797 | 0.557 | 0.510 | 4.474 |
| 6O7P[78] | 1.70 | 2.356 | 2.354 | 2.590 | 2.303 | 2.603 | 2.124 | 2.236 | 2.272 | 18.839 | 2.475 | 2.543 | 2.901 | 2.507 | 2.691 | 2.401 | 2.474 | 2.487 | 20.477 |
| | | 2.288 | 2.351 | 2.683 | 2.318 | 2.488 | 2.220 | 2.287 | 2.299 | 18.934 | 2.548 | 2.546 | 2.801 | 2.491 | 2.815 | 2.297 | 2.419 | 2.457 | 20.375 |
| 5VQ3 [79] | 1.72 | 2.246 | 2.179 | 2.430 | 2.420 | 2.163 | 1.986 | 2.389 | 2.228 | 18.041 | 2.429 | 2.356 | 2.628 | 2.617 | 2.339 | 2.148 | 2.583 | 2.410 | 19.511 |
| | | 2.174 | 2.199 | 2.458 | 2.299 | 2.145 | 2.109 | 2.374 | 2.378 | 18.135 | 2.351 | 2.379 | 2.658 | 2.486 | 2.320 | 2.281 | 2.567 | 2.572 | 19.614 |
| 6VXT[25] | 1.74 | 2.201 | 2.056 | 2.244 | 2.355 | 2.291 | 1.968 | 1.938 | 2.042 | 17.095 | 2.381 | 2.224 | 2.427 | 2.547 | 2.478 | 2.128 | 2.096 | 2.209 | 18.490 |
| | | 2.338 | 2.083 | 2.101 | 2.307 | 2.327 | 1.959 | 2.071 | 2.010 | 17.196 | 2.529 | 2.253 | 2.272 | 2.495 | 2.517 | 2.119 | 2.240 | 2.174 | 18.599 |
| 5CX1 [80] | 1.75 | 2.348 | 2.333 | 2.526 | 2.222 | 2.293 | 2.089 | 2.329 | 2.296 | 18.436 | 2.540 | 2.523 | 2.732 | 2.403 | 2.480 | 2.259 | 2.519 | 2.484 | 19.940 |
| | | 2.393 | 2.316 | 2.468 | 2.292 | 2.209 | 2.170 | 2.370 | 2.315 | 18.535 | 2.588 | 2.505 | 2.669 | 2.479 | 2.389 | 2.347 | 2.564 | 2.504 | 20.046 |
| | | 2.249 | 2.330 | 2.450 | 2.206 | 2.270 | 2.168 | 2.319 | 2.322 | 18.315 | 2.433 | 2.520 | 2.650 | 2.386 | 2.455 | 2.345 | 2.508 | 2.511 | 19.808 |
| | | 2.193 | 2.389 | 2.555 | 2.381 | 2.042 | 2.097 | 2.290 | 2.217 | 18.165 | 2.371 | 2.584 | 2.764 | 2.575 | 2.209 | 2.268 | 2.477 | 2.398 | 19.647 |
| | | 2.270 | 2.315 | 2.581 | 2.222 | 2.302 | 2.135 | 2.389 | 2.284 | 18.498 | 2.455 | 2.504 | 2.792 | 2.403 | 2.489 | 2.309 | 2.584 | 2.470 | 20.006 |
| | | 2.426 | 2.433 | 2.567 | 2.405 | 2.186 | 2.047 | 2.341 | 2.336 | 18.741 | 2.624 | 2.632 | 2.776 | 2.601 | 2.364 | 2.214 | 2.532 | 2.526 | 20.269 |
| | | 2.256 | 2.407 | 2.476 | 2.370 | 2.308 | 2.153 | 2.309 | 2.285 | 18.563 | 2.440 | 2.603 | 2.678 | 2.563 | 2.496 | 2.328 | 2.497 | 2.471 | 20.076 |
| | | 2.344 | 2.377 | 2.538 | 2.299 | 2.126 | 2.164 | 2.312 | 2.306 | 18.466 | 2.535 | 2.571 | 2.745 | 2.486 | 2.300 | 2.340 | 2.500 | 2.494 | 19.972 |
| 5KOH[81] | 1.83 | 2.248 | 2.314 | 2.572 | 2.414 | 2.213 | 2.126 | 2.336 | 2.305 | 18.528 | 2.431 | 2.502 | 2.782 | 2.611 | 2.393 | 2.300 | 2.527 | 2.493 | 20.038 |
| | | 2.214 | 2.330 | 2.516 | 2.420 | 2.175 | 2.082 | 2.375 | 2.291 | 18.405 | 2.395 | 2.520 | 2.722 | 2.617 | 2.353 | 2.252 | 2.569 | 2.478 | 19.905 |

Table 4 (*Contd.*)

| PDB codes | Res (Å) | $R_0$ (+2) | | | | | | | | | $R_0$ (+3) | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Fe1 | Fe2 | Fe3 | Fe4 | Fe5 | Fe6 | Fe7 | Fe8 | $S_t$(8Fe) | Fe1 | Fe2 | Fe3 | Fe4 | Fe5 | Fe6 | Fe7 | Fe8 | $S_t$(8Fe) |
| 5VPW[79] | 1.85 | 2.233 | 2.197 | 2.423 | 2.374 | 2.122 | 1.919 | 2.239 | 2.234 | 17.740 | 2.493 | 2.543 | 2.665 | 2.606 | 2.230 | 2.037 | 2.539 | 2.395 | 19.508 |
| | | 2.305 | 2.351 | 2.464 | 2.409 | 2.062 | 1.884 | 2.347 | 2.215 | 18.037 | 2.415 | 2.376 | 2.620 | 2.567 | 2.295 | 2.075 | 2.421 | 2.416 | 19.186 |
| 6O7O[78] | 1.89 | 2.352 | 2.239 | 2.356 | 2.416 | 2.374 | 2.200 | 2.456 | 2.215 | 18.608 | 2.504 | 2.568 | 2.822 | 2.633 | 2.591 | 2.332 | 2.685 | 2.463 | 20.599 |
| | | 2.316 | 2.374 | 2.610 | 2.434 | 2.396 | 2.157 | 2.483 | 2.277 | 19.046 | 2.544 | 2.422 | 2.548 | 2.613 | 2.567 | 2.379 | 2.657 | 2.396 | 20.125 |
| 1H1L[82] | 1.90 | 2.077 | 2.149 | 2.317 | 2.341 | 1.991 | 2.039 | 2.213 | 2.230 | 17.358 | 2.246 | 2.324 | 2.506 | 2.532 | 2.154 | 2.206 | 2.393 | 2.412 | 18.773 |
| | | 2.101 | 2.225 | 2.514 | 2.438 | 2.120 | 1.964 | 2.141 | 2.235 | 17.738 | 2.272 | 2.406 | 2.719 | 2.637 | 2.293 | 2.124 | 2.315 | 2.417 | 19.184 |
| 4WZA[83] | 1.90 | 2.085 | 2.068 | 2.253 | 2.299 | 2.214 | 2.001 | 2.047 | 2.204 | 17.171 | 2.255 | 2.236 | 2.437 | 2.486 | 2.395 | 2.164 | 2.214 | 2.383 | 18.571 |
| | | 2.119 | 2.090 | 2.308 | 2.271 | 2.126 | 1.988 | 2.018 | 2.174 | 17.094 | 2.292 | 2.261 | 2.497 | 2.456 | 2.300 | 2.150 | 2.182 | 2.351 | 18.488 |
| 4WZB[84] | 1.90 | 2.137 | 2.008 | 2.213 | 2.266 | 2.029 | 1.903 | 1.966 | 2.139 | 16.660 | 2.311 | 2.171 | 2.394 | 2.451 | 2.195 | 2.058 | 2.126 | 2.314 | 18.019 |
| | | 2.165 | 1.956 | 2.278 | 2.197 | 2.211 | 1.844 | 2.083 | 2.074 | 16.809 | 2.341 | 2.115 | 2.464 | 2.376 | 2.391 | 1.994 | 2.253 | 2.243 | 18.179 |
| 4WNA[85] | 2.00 | 2.198 | 2.244 | 2.587 | 2.590 | 2.158 | 1.914 | 2.361 | 2.478 | 18.530 | 2.377 | 2.426 | 2.798 | 2.801 | 2.334 | 2.070 | 2.553 | 2.680 | 20.040 |
| (Modeled as $P^N$) | | 2.123 | 2.408 | 2.742 | 2.589 | 2.217 | 1.967 | 2.413 | 2.399 | 18.857 | 2.296 | 2.604 | 2.965 | 2.800 | 2.398 | 2.127 | 2.610 | 2.594 | 20.394 |
| 6O7Q[78] | 2.00 | 2.117 | 2.227 | 2.530 | 2.126 | 2.282 | 1.987 | 2.251 | 2.357 | 17.876 | 2.241 | 2.454 | 2.784 | 2.435 | 2.527 | 1.988 | 2.461 | 2.471 | 19.361 |
| | | 2.072 | 2.269 | 2.574 | 2.252 | 2.337 | 1.838 | 2.276 | 2.285 | 17.902 | 2.290 | 2.409 | 2.736 | 2.299 | 2.468 | 2.149 | 2.434 | 2.549 | 19.334 |
| 3MIN[8] | 2.03 | 3.031 | 2.536 | 2.829 | 3.226 | 2.763 | 2.223 | 2.590 | 2.306 | 21.505 | 3.279 | 2.743 | 3.060 | 3.489 | 2.988 | 2.405 | 2.801 | 2.494 | 23.259 |
| | | 2.671 | 2.324 | 2.803 | 2.593 | 2.400 | 2.021 | 2.494 | 2.536 | 19.841 | 2.889 | 2.514 | 3.031 | 2.804 | 2.595 | 2.186 | 2.697 | 2.742 | 21.459 |
| 2AFH[84] | 2.10 | 2.595 | 2.045 | 2.639 | 3.012 | 3.043 | 2.168 | 2.427 | 2.894 | 20.823 | 2.807 | 2.211 | 2.854 | 3.257 | 3.292 | 2.345 | 2.624 | 3.130 | 22.520 |
| | | 3.039 | 2.061 | 2.720 | 3.067 | 2.938 | 2.064 | 2.437 | 3.064 | 21.388 | 3.287 | 2.229 | 2.942 | 3.317 | 3.178 | 2.232 | 2.635 | 3.313 | 23.132 |
| 6O7L[78] | 2.26 | 2.266 | 2.391 | 2.775 | 2.528 | 2.976 | 2.179 | 1.679 | 2.623 | 19.417 | 2.450 | 2.586 | 3.002 | 2.734 | 3.219 | 2.357 | 1.815 | 2.836 | 21.000 |
| | | 2.153 | 2.151 | 2.815 | 2.394 | 2.331 | 2.232 | 1.853 | 2.238 | 18.167 | 2.329 | 2.326 | 3.044 | 2.589 | 2.521 | 2.414 | 2.005 | 2.420 | 19.648 |
| 6O7R[78] | 2.27 | 2.351 | 2.463 | 2.783 | 2.519 | 2.500 | 2.134 | 2.325 | 2.651 | 19.726 | 2.499 | 2.614 | 2.972 | 2.662 | 2.623 | 2.308 | 2.536 | 2.857 | 21.069 |
| | | 2.310 | 2.417 | 2.748 | 2.461 | 2.425 | 2.134 | 2.345 | 2.641 | 19.481 | 2.543 | 2.664 | 3.010 | 2.724 | 2.704 | 2.308 | 2.514 | 2.867 | 21.334 |
| 1L5H[86] | 2.30 | 2.815 | 2.552 | 2.367 | 2.297 | 2.258 | 1.783 | 2.541 | 2.402 | 19.015 | 3.045 | 2.760 | 2.560 | 2.485 | 2.442 | 1.928 | 2.748 | 2.598 | 20.565 |
| 1M34 [87] | 2.30 | 2.326 | 2.492 | 2.480 | 2.466 | 2.321 | 1.990 | 2.376 | 2.480 | 18.931 | 2.516 | 2.695 | 2.683 | 2.667 | 2.511 | 2.152 | 2.570 | 2.682 | 20.474 |
| | | 2.507 | 2.531 | 2.592 | 2.914 | 2.325 | 1.840 | 2.438 | 2.323 | 19.468 | 2.712 | 2.737 | 2.803 | 3.151 | 2.514 | 1.990 | 2.637 | 2.512 | 21.055 |
| | | 2.556 | 2.564 | 2.577 | 2.656 | 2.154 | 2.130 | 2.624 | 2.166 | 19.425 | 2.764 | 2.773 | 2.787 | 2.872 | 2.329 | 2.304 | 2.838 | 2.342 | 21.008 |
| | | 2.589 | 2.645 | 2.597 | 2.590 | 2.161 | 1.995 | 2.428 | 2.317 | 19.323 | 2.800 | 2.861 | 2.809 | 2.801 | 2.337 | 2.158 | 2.626 | 2.506 | 20.898 |
| 5VQ4 [79] | 2.30 | 2.218 | 2.287 | 2.646 | 2.493 | 2.040 | 1.737 | 2.457 | 2.635 | 18.513 | 2.399 | 2.474 | 2.862 | 2.697 | 2.206 | 1.878 | 2.657 | 2.850 | 20.023 |
| | | 2.026 | 2.236 | 2.548 | 2.483 | 2.039 | 1.736 | 2.130 | 2.286 | 17.484 | 2.191 | 2.419 | 2.755 | 2.686 | 2.206 | 1.877 | 2.304 | 2.472 | 18.910 |
| $\overline{S_w}$ (in 2.3 Å) | | 2.285 | 2.268 | 2.505 | 2.404 | 2.246 | 2.043 | 2.281 | 2.323 | 18.353 | 2.465 | 2.447 | 2.702 | 2.594 | 2.423 | 2.204 | 2.461 | 2.506 | 19.804 |
| **Weighted \|*d*\|** | | 0.285 | 0.268 | 0.505 | 0.404 | 0.246 | 0.043 | 0.281 | 0.323 | 2.353 | 0.535 | 0.553 | 0.298 | 0.406 | 0.577 | 0.796 | 0.539 | 0.494 | 4.196 |
| **$P^N$ in VFe proteins** | | | | | | | | | | | | | | | | | | | |
| 7ADR[88] | 1.00 | 2.154 | 2.085 | 2.324 | 2.290 | 2.183 | 2.050 | 2.397 | 2.365 | 17.846 | 2.329 | 2.255 | 2.513 | 2.476 | 2.361 | 2.217 | 2.592 | 2.557 | 19.301 |
| (Modeled as $P^N$) | | 2.137 | 2.081 | 2.315 | 2.265 | 2.204 | 2.084 | 2.405 | 2.372 | 17.862 | 2.312 | 2.250 | 2.504 | 2.450 | 2.384 | 2.254 | 2.601 | 2.565 | 19.319 |
| 7ADY[88] | 1.05 | 2.133 | 2.075 | 2.319 | 2.255 | 2.174 | 2.084 | 2.413 | 2.342 | 17.795 | 2.307 | 2.244 | 2.508 | 2.439 | 2.351 | 2.254 | 2.610 | 2.533 | 19.245 |
| (Modeled as $P^N$) | | 2.157 | 2.076 | 2.327 | 2.286 | 2.171 | 2.022 | 2.386 | 2.356 | 17.783 | 2.333 | 2.246 | 2.517 | 2.473 | 2.349 | 2.187 | 2.581 | 2.548 | 19.233 |
| 7AIZ[89] | 1.05 | 2.145 | 2.081 | 2.311 | 2.269 | 2.202 | 2.080 | 2.422 | 2.351 | 17.861 | 2.320 | 2.251 | 2.499 | 2.454 | 2.382 | 2.250 | 2.620 | 2.543 | 19.319 |
| (Modeled as $P^N$) | | 2.175 | 2.069 | 2.301 | 2.295 | 2.143 | 1.990 | 2.395 | 2.325 | 17.693 | 2.352 | 2.238 | 2.489 | 2.482 | 2.318 | 2.153 | 2.590 | 2.514 | 19.137 |
| 6FEA[90] | 1.20 | 2.155 | 2.103 | 2.344 | 2.320 | 2.213 | 2.046 | 2.438 | 2.411 | 18.030 | 2.331 | 2.274 | 2.535 | 2.509 | 2.393 | 2.213 | 2.637 | 2.608 | 19.500 |
| (Modeled as $P^N$) | | 2.181 | 2.108 | 2.383 | 2.295 | 2.253 | 2.072 | 2.457 | 2.400 | 18.148 | 2.359 | 2.279 | 2.578 | 2.482 | 2.437 | 2.240 | 2.657 | 2.595 | 19.627 |
| 5N6Y[91] | 1.35 | 2.249 | 2.145 | 2.432 | 2.324 | 2.300 | 2.176 | 2.511 | 2.461 | 18.599 | 2.433 | 2.320 | 2.631 | 2.513 | 2.488 | 2.353 | 2.716 | 2.662 | 20.115 |
| | | 2.192 | 2.165 | 2.336 | 2.269 | 2.261 | 2.115 | 2.449 | 2.430 | 18.216 | 2.371 | 2.341 | 2.526 | 2.453 | 2.445 | 2.288 | 2.648 | 2.628 | 19.701 |
| $\overline{S_w}$ | | 2.165 | 2.096 | 2.336 | 2.286 | 2.206 | 2.069 | 2.424 | 2.377 | 17.959 | 2.342 | 2.267 | 2.527 | 2.472 | 2.386 | 2.237 | 2.622 | 2.571 | 19.424 |
| **Weighted \|*d*\|** | | 0.165 | 0.096 | 0.336 | 0.286 | 0.206 | 0.069 | 0.424 | 0.377 | 1.959 | 0.658 | 0.733 | 0.473 | 0.528 | 0.614 | 0.763 | 0.378 | 0.429 | 4.576 |

valence sum of 8Fe is 18.353 by using $R_0$ (+2), and most valence sums of 8Fe are distributed in the range of 17 to 19, resembling $P^N$ model compound DUGNEZ.[66] The above discussions and the calculated results elucidate that resting state $P^N$ clusters in protein crystals contain one ferric and two iron atoms of mixed valence, and ought to have an oxidation state approximately equal to 6Fe(II)2Fe(III) with delocalized electrons. However, due to widespread electron delocalization, mixed valences existing in Fe3/4/8 influence their accurate valence assignments. Given the limitation of BVS applied in a mixed-valence system, we consider that Fe1/2/5/6/7 should be assigned to $Fe^{2+}$ and Fe4/8 are prone to being mixed valence as $Fe^{2.33+}$ or $Fe^{2.5+}$, which has been reported in $[MFe_3S_4]$ model compounds.[70,71] Fe3 has a larger possibility of possessing states of Fe(III) or $Fe^{2.5+}$, and more oxidative mixed valences compared with Fe4/8.

According to Fig. 2, we can see that Fe3 and Fe6 are obviously the most oxidized and reduced iron atoms in $P^N$. When changing our focus to the spatial locations of three clusters, as shown in Fig. 3, we found that Fe3 has the shortest 9.0 Å distance in the eight iron atoms to homocitrate in the M-cluster, which participates in catalysis as an electron-demander and takes charge of the reduction of the substrate. Fe4, which ranks as the second highest mixed valence in the eight irons, is also the iron second closest to the M-cluster. Besides, Fe6 with a distance of 15.0 Å is the nearest iron to the electron-donor $[Fe_4S_4]$ which is responsible for the "backfill" electron transfer to $P^{1+}$ in the deficit-spending mechanism. Similarly, Fe1/2, which side by side with Fe6 perform with an obviously reductive character, are the irons spatially second closest to the F-cluster.

From the perspective of $P^N$ structure, terminal Fe3 and Fe7 seem to be oxidized more easily, which is similar to the report that oxidation occurs preferentially at the peripheral iron sites Fe(1) and Fe(5) in $P^N$ model compounds.[64] Nearby Fe3 and Fe7, Fe4 and Fe8 are also calculated out mixed valence as Fe(4) and Fe(8). By maintaining $P^N$ in reductive solution with excess reducing agents, plenty of experiments give the conclusion that $P^N$ clusters are all-ferrous according to Mössbauer and EPR at an early stage.[20,28,29] But irons in model compounds of $P^N$ could have the character of 6Fe(II)2Fe(III), as reported. As Fig. 1d shows, structures of $P^N$ model compounds display the same coordinated sites as those amino acid residues bonding with natural $P^N$. Interestingly, the terminal Fe(III) sites Fe1 and Fe5 in $P^N$ model compounds correspond to terminal irons Fe3 and Fe7 in natural $P^N$.[64,65] Thus, partially oxidized $P^N$ model compounds could still maintain their original structure, which indicates that $P^N$ in the form of protein crystals has the possibility of being in the same oxidative state of 6Fe(II)2Fe(III) without structural change. On the other hand, the calculated results of the most accurate PDB entry 3U7Q indeed display consistent conclusions with the above discussion. Its Fe1, Fe2, Fe5 and Fe6 obviously approach $Fe^{2+}$ and Fe3 tends to $Fe^{3+}$ with an average value of 2.736 in the group of $R_0$ (+3), where different degrees of mixed valence are distributed in electron-delocalized Fe4, Fe7 and Fe8.

Thus, we could conclude that the oxidation states of the eight irons in $P^N$ of MoFe protein crystals are different and not completely all-ferrous. In protein crystals, Fe1, Fe2, Fe5 and Fe6 still perform with strong reductive character as in a reductive bio-environment, but Fe3 and Fe4/8/7 are more oxidized than other irons in consequence, even though $P^N$ is reduced in the initial stage.

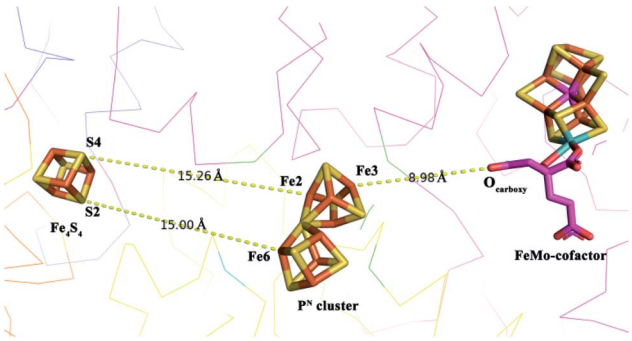


Fig. 3 The spatial position between the F-cluster, $P^N$ cluster and FeMo-cofactor from PDB entry 4WZB in 1.9 Å,[69] which simultaneously contains an MoFe protein and an Fe protein with the highest resolution in all deposited PDB data. Nearby amino acid residues are simplified.

### 3.3. Valence analyses of $P^N$ clusters in VFe proteins

From Fig. 4, P-clusters of five VFe proteins have similar degrees of oxidation states in each iron. The average valence sums $S_t$ 18.407 of 5N6Y and 18.089 of 6FEA by $R_0$ (+2) indicate that P-clusters of the latter are more reduced than those of the former as a whole, which is consistent with the discovery we also made before for FeMo-cofactors.[46]

For Fe1, Fe2, Fe4, Fe5 and Fe6 in Table 4 and Fig. 4a, their sideways $|d|$ of groups $R_0$ (+2) are below 0.3 and smaller than those of $R_0$ (+3), indicating that the assignment of iron(II) is appropriate, while Fe4 shows a small degree of mixed valence. Compared with the $P^N$ of MoFe proteins in Fig. 4b, P-clusters in VFe proteins have the similarity that Fe3, Fe7 and Fe8 perform with obvious characters of mixed valences while there are dramatic differences in Fe7 and Fe8 which are more oxidized than Fe3. In Table 4 it can be seen that Fe7 and Fe8 with strong electron delocalization have similar $|d|$ of 0.38 and 0.43 by using $R_0$ (+3), and 0.42 and 0.38 by using $R_0$ (+2), which manifest the tendency of Fe7 and Fe8 for being high mixed valence. Opposite to MoFe proteins, Fe3 in VFe proteins are more inclined to be iron(II) in mixed valence, with $|d|$ of 0.34 and 0.47 by using $R_0$ (+2) and $R_0$ (+3), respectively. From the perspective of the weighted average valence sum $S_t$ (17.96) of 8Fe by $R_0$ (+2), the $P^N$ in the VFe protein with a similar formal oxidation state of 6Fe(II) 2Fe(III), is more reductive than the MoFe protein to a small degree. Although the core structure $[Fe_8S_7]$ of the P-cluster is the same in MoFe and VFe proteins, it can be seen that their electron distributions of the P-cluster are apparently different, as the electron distribution of FeMo-co is also different from that of VFe-co.[46] The above difference could be attributed to different structures between MoFe and VFe proteins, which may result in different electron transfer channels to induce the formation of the relevant iron oxidation states.

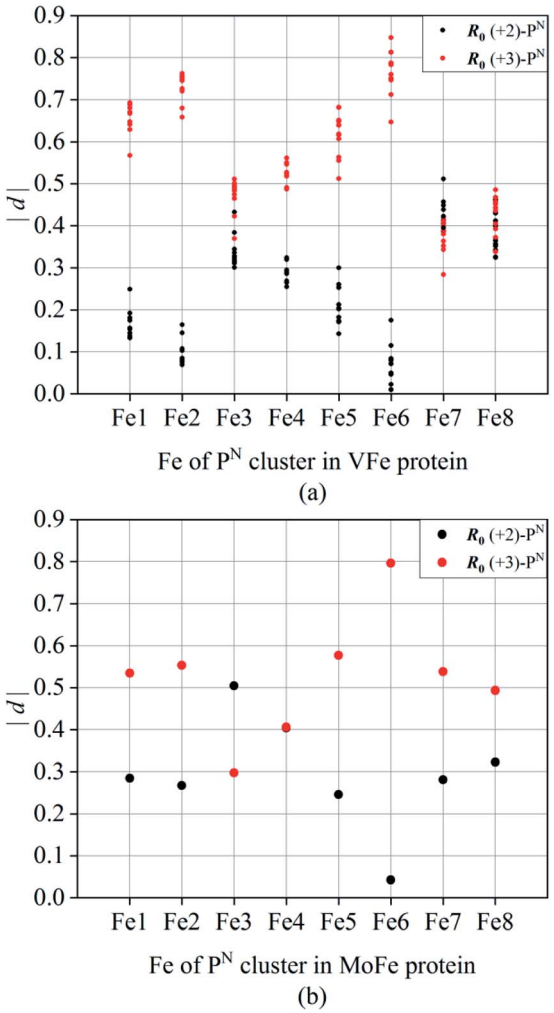


Fig. 4 (a) $|d|$ calculated for $R_0$ (+2) (black) and $R_0$ (+3) (red) of Fe1–Fe8 in P$^N$ from 5 VFe proteins; (b) the weighted average $|d|$ calculated for $R_0$ (+2) (black) and $R_0$ (+3) (red) of Fe1–Fe8 in P$^N$ from MoFe proteins.

## 4 Conclusions

We have studied all deposited P$^N$ in 119 P-clusters of 53 MoFe PDB entries and 10 P-clusters of 5 VFe PDB entries with the bond valence method. In Mo/VFe protein crystals, P$^N$ clusters are all supposed formally to be 6Fe(II)2Fe(III). All of their Fe1, Fe2, Fe5 and Fe6 ought to be assigned as $Fe^{2+}$, while the mixed valences $Fe^{2.33+/2.5+}$ in Fe3, Fe4, Fe7 and Fe8 are differently distributed, probably due to their different protein structures. These reflect which Fe atoms have a tendency to maintain oxidized or reduced states of Mo/VFe proteins in crystal form. The calculated results of P$^N$ in crystals seem not to be the same as in traditional ideas that P$^N$ clusters are "all-ferrous" from those analyses in reductive solutions with excess reducing agents. In view of the spatial position of the MoFe protein crystal, the most oxidized Fe3 and reduced Fe6 are simultaneously the nearest irons to FeMo-co and [$Fe_4S_4$], respectively. It seems that Fe3 and Fe6 could function as the most convenient electron transfer sites. This potential difference in P$^N$ clusters might be more suitable to correlate with the other two F- and M-clusters as an important electron transfer station.

This work first applied the bond valence method in P-clusters of Mo/VFe proteins statistically, which provided a new perspective in the general electron distributions of P-clusters in nitrogenase, and might be widely applied to other metalloenzyme systems with electron delocalization. Our work delivers a much more detailed evaluation of the oxidation states of the eight irons in the P-cluster, adding a special story to research into nitrogenase. More insightful pursuits still need further investigations. All of these studies were built on predecessors' work that supplied sufficient crystal data of Mo/VFe proteins in the PDB to help educe reasonable results.

## Conflicts of interest

The authors declare no competing financial interests.

## Acknowledgements

We thank the National Natural Science Foundation of China (22179110) for generous financial support.

## References

1 B. M. Hoffman, D. Lukoyanov, Z. Y. Yang, D. R. Dean and L. C. Seefeldt, *Chem. Rev.*, 2014, **114**, 4041–4062.
2 J. Kim, D. Woo and D. C. Rees, *Biochemistry*, 1993, **32**, 7104–7115.
3 O. Einsle, F. A. Tezcan, S. L. A. Andrade, B. Schmid, M. Yoshida, J. B. Howard and D. C. Rees, *Science*, 2002, **297**, 1696–1700.
4 T. Spatzal, M. Aksoyoglu, L. Zhang, S. L. A. Andrade, E. Schleicher, S. Weber, D. C. Rees and O. Einsle, *Science*, 2011, **334**, 940.
5 K. M. Lancaster, M. Roemelt, P. Ettenhuber, Y. Hu, M. W. Ribbe, F. Neese, U. Bergmann and S. DeBeer, *Science*, 2011, **334**, 974.
6 L. Deng, H. Wang, C. H. Dapper, W. E. Newton, S. Shilov, S. Wang, S. P. Cramer and Z. H. Zhou, *Commun. Chem.*, 2020, **3**, 145.
7 K. Danyal, D. R. Dean, B. M. Hoffman and L. C. Seefeldt, *Biochemistry*, 2011, **50**, 9255–9263.
8 J. W. Peters, M. H. B. Stowell, S. M. Soltis, M. G. Finnegan, M. K. Johnson and D. C. Rees, *Biochemistry*, 1997, **36**, 1181–1187.
9 R. V. Hageman and R. H. Burris, *Proc. Natl. Acad. Sci.*, 1978, **75**, 2699–2702.
10 D. Lukoyanov, S. A. Dikanov, Z. Y. Yang, B. M. Barney, R. I. Samoilova, K. V. Narasimhulu, D. R. Dean, L. C. Seefeldt and B. M. Hoffman, *J. Am. Chem. Soc.*, 2011, **133**, 11655–11664.
11 B. K. Burgess, *Chem. Rev.*, 1990, **90**, 1377–1406.
12 I. Čorić and P. L. Holland, *J. Am. Chem. Soc.*, 2016, **138**, 7200–7211.
13 D. F. Harris, D. A. Lukoyanov, H. Kallas, C. Trncik, Z. Y. Yang, P. Compton, N. Kelleher, O. Einsle, D. R. Dean,

B. M. Hoffman and L. C. Seefeldt, *Biochemistry*, 2019, **58**, 3293–3301.
14 J. Kim and D. C. Rees, *Science*, 1992, **257**, 1677–1682.
15 J. M. Chan, J. Christiansen, D. R. Dean and L. C. Seefeldt, *Biochemistry*, 1999, **38**, 5779–5785.
16 D. R. Dean, J. T. Bolin and L. Zheng, *J. Bacteriol.*, 1993, **175**, 6737–6744.
17 L. Ma, M. A. Brosius and B. K. Burgess, *J. Biol. Chem.*, 1996, **271**, 10528–10532.
18 R. C. Tittsworth and B. J. Hales, *J. Am. Chem. Soc.*, 1993, **115**, 9763–9767.
19 R. C. Tittsworth and B. J. Hales, *Biochemistry*, 1996, **35**, 479–487.
20 A. J. Pierik, H. Wassink, H. Haaker and W. R. Hagen, *Eur. J. Biochem.*, 1993, **212**, 51–61.
21 J. Christiansen, P. J. Goodwin, W. N. Lanzilotta, L. C. Seefeldt and D. R. Dean, *Biochemistry*, 1998, **37**, 12611–12623.
22 S. M. Keable, O. A. Zadvornyy, L. E. Johnson, B. Ginovska, A. J. Rasmussen, K. Danyal, B. J. Eilers, G. A. Prussia, A. X. LeVan, S. Raugei, L. C. Seefeldt and J. W. Peters, *J. Biol. Chem.*, 2018, **293**, 9629–9635.
23 L. Cao, M. C. Börner, J. Bergmann, O. Caldararu and U. Ryde, *Inorg. Chem.*, 2019, **58**, 9672–9690.
24 S. J. Yoo, H. C. Angove, V. Papaefthymiou, B. K. Burgess and E. Münck, *J. Am. Chem. Soc.*, 2000, **122**, 4926–4936.
25 W. Kang, C. C. Lee, A. J. Jasniewski, M. W. Ribbe and Y. Hu, *Science*, 2020, **368**, 1381–1385.
26 Z. Li, S. Guo, Q. Sun and G. K. Chan, *Nat. Chem.*, 2019, **11**, 1026–1033.
27 E. Münck, H. Rhodes, W. H. Orme-Johnson, L. C. Davis, W. J. Brill and V. K. Shah, *Biochim. Biophys. Acta, Protein Struct.*, 1975, **400**, 32–53.
28 P. A. McLean, V. Papaefthymiou, W. H. Orme-Johnson and E. Münck, *J. Biol. Chem.*, 1987, **262**, 12900–12903.
29 J. M. Mouesca, L. Noodleman and D. A. Case, *Inorg. Chem.*, 1994, **33**, 4819–4830.
30 P. A. Lindahl, V. Papaefthymiou, W. H. Orme-Johnson and E. Münck, *J. Biol. Chem.*, 1988, **263**, 19412–19418.
31 K. Rupnik, Y. Hu, C. C. Lee, J. A. Wiig, M. W. Ribbe and B. J. Hales, *J. Am. Chem. Soc.*, 2012, **134**, 13749–13754.
32 I. D. Brown and D. Altermatt, *Acta Crystallogr., Sect. B: Struct. Sci., Cryst. Eng. Mater.*, 1985, **41**, 244–247.
33 T. Seto and M. G. Brik, *RSC Adv.*, 2017, **7**, 40152–40157.
34 H. Ben Yahia, A. Alkhateeb and R. Essehli, *RSC Adv.*, 2020, **10**, 10420–10430.
35 R. K. Tiwari, I. Shruti and J. N. Behera, *RSC Adv.*, 2021, **11**, 10767–10776.
36 C. Zheng, M. Liang, H. Sun, J. Ma, X. Bi, Y. Zhao, W. Tan and H. Li, *RSC Adv.*, 2019, **9**, 39965–39969.
37 H. H. Thorp, *Inorg. Chem.*, 1998, **37**, 5690–5692.
38 I. D. Brown, *Chem. Rev.*, 2009, **109**, 6858–6919.
39 A. V. Morozov, A. A. Savina, A. O. Boev, E. V. Antipov and A. M. Abakumov, *RSC Adv.*, 2021, **11**, 28593–28601.
40 M. Dębowski, Z. Florjańczyk, A. Ostrowski, P. A. Guńka, J. Zachara, A. Krztoń-Maziopa, J. Chazarkiewicz, A. Iuliano and A. Plichta, *RSC Adv.*, 2021, **11**, 7873–7885.
41 X. Wang, X. Pan, X. Wang, Y. Li and G. Liu, *RSC Adv.*, 2020, **10**, 11046–11053.
42 L. Pauling, *J. Am. Chem. Soc.*, 1929, **51**, 1010–1026.
43 E. Levi, D. Aurbach and C. Gatti, *Phys. Chem. Chem. Phys.*, 2020, **22**, 13839–13849.
44 H. H. Thorp, *Inorg. Chem.*, 1992, **31**, 1585–1588.
45 P. Müller, S. Kopke and G. M. Sheldrick, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2003, **59**, 32–37.
46 W. T. Jin, M. Yang, S. S. Zhu and Z. H. Zhou, *Acta Crystallogr., Sect. D: Struct. Biol.*, 2020, **76**, 428–437.
47 H. Chen and S. Adams, *IUCrJ*, 2017, **4**, 614–625.
48 P. Paufler, *Acta Crystallogr., Sect. A: Found. Crystallogr.*, 2007, **63**, 374.
49 S. Adams., *SOFTBV. Version 0.96*, http://kristall.uni-mki.gwdg.de/softBV/, 2004.
50 I. Brown, , Accumulated table of bond-valence parameters, https://www.iucr.org/__data/assets/file/0011/150779/bvparm2020.cif, , 2020.
51 I. D. Brown, The Chemical Bond in Inorganic Chemistry, *The Bond Valence Model*, Oxford University Press, USA, 2006.
52 S. Z. Hu and Z. H. Zhou, *Z. Kristallogr. - Cryst. Mater.*, 2004, **219**, 614–620.
53 W. Liu and H. H. Thorp, *Inorg. Chem.*, 1993, **32**, 4102–4105.
54 L. Noodleman, C. Y. Peng, D. A. Case and J. M. Mouesca, *Coord. Chem. Rev.*, 1995, **144**, 199–244.
55 P. M. Bartier and C. P. Keller, *Comput. Geosci.*, 1996, **22**, 795–799.
56 D. Karunanidhi, P. Aravinthasamy, M. Deepali, T. Subramani and P. D. Roy, *RSC Adv.*, 2020, **10**, 4840–4859.
57 S. Karimian, S. Shekoohiyan and G. Moussavi, *RSC Adv.*, 2021, **11**, 8080–8095.
58 K. K. Yadav, N. Gupta, V. Kumar, P. Choudhary and S. A. Khan, *RSC Adv.*, 2018, **8**, 15876–15889.
59 X. Liu, Z. Bai, Q. Yu, Y. Cao and W. Zhou, *RSC Adv.*, 2017, **7**, 28029–28037.
60 Y. Shi, W. He, J. Zhao, A. Hu, J. Pan, H. Wang and H. Zhu, *J. Cleaner Prod.*, 2020, 253.
61 D. Shepard, presented in part at the *Proceedings of the 1968 23rd ACM national conference*, 1968.
62 L. H. Yin, N. Y. Ting, F. P. Shan, K. Shimizu and H. Lateh, presented in part at the *proceedings of the international conference on mathematical sciences and technology 2018 (MATHTECH2018): Innovative Technologies for Mathematics & Mathematics for Technological Innovation*, 2019.
63 S. Ohta, Y. Ohki, T. Hashimoto, R. E. Cramer and K. Tatsumi, *Inorg. Chem.*, 2012, **51**, 11217–11219.
64 Y. Ohki, M. Imada, A. Murata, Y. Sunada, S. Ohta, M. Honda, T. Sasamori, N. Tokitoh, M. Katada and K. Tatsumi, *J. Am. Chem. Soc.*, 2009, **131**, 13168–13178.
65 Y. Ohki, Y. Sunada, M. Honda, M. Katada and K. Tatsumi, *J. Am. Chem. Soc.*, 2003, **125**, 4052–4053.
66 Y. Ohki, A. Murata, M. Imada and K. Tatsumi, *Inorg. Chem.*, 2009, **48**, 4271–4273.
67 Y. Ohki, K. Tanifuji, N. Yamada, R. E. Cramer and K. Tatsumi, *Chem.–Asian J.*, 2012, **7**, 2222–2224.
68 Y. Sanakis, S. J. Yoo, F. Osterloh, R. H. Holm and E. Münck, *Inorg. Chem.*, 2002, **41**, 7081–7085.

69 F. A. Tezcan, J. T. Kaiser, J. B. Howard and D. C. Rees, *J. Am. Chem. Soc.*, 2015, **137**, 146–149.
70 J. L. Zuo, H. C. Zhou and R. H. Holm, *Inorg. Chem.*, 2003, **42**, 4624–4631.
71 C. Hauser, E. Bill and R. H. Holm, *Inorg. Chem.*, 2002, **41**, 1615–1624.
72 L. M. Zhang, C. N. Morrison, J. T. Kaiser and D. C. Rees, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2015, **71**, 274–282.
73 T. M. Buscagan, K. A. Perez, A. O. Maggiolo, D. C. Rees and T. Spatzal, *Angew. Chem., Int. Ed.*, 2021, **60**, 5704–5707.
74 T. Spatzal, K. Perez, O. Einsle, J. Howard and D. Rees, *Science*, 2014, **345**, 1620–1623.
75 T. Spatzal, K. A. Perez, J. B. Howard and D. C. Rees, *eLife*, 2015, **4**, e11620.
76 S. M. Mayer, D. M. Lawson, C. A. Gormal, S. M. Roe and B. E. Smith, *J. Mol. Biol.*, 1999, **292**, 871–891.
77 J. T. Henthorn, R. J. Arias, S. Koroidov, T. Kroll, D. Sokaras, U. Bergmann, D. C. Rees and S. DeBeer, *J. Am. Chem. Soc.*, 2019, **141**, 13676–13688.
78 H. L. Rutledge, J. Rittle, L. M. Williamson, W. A. Xu, D. M. Gagnon and F. A. Tezcan, *J. Am. Chem. Soc.*, 2019, **141**, 10091–10098.
79 C. N. Morrison, T. Spatzal and D. C. Rees, *J. Am. Chem. Soc.*, 2017, **139**, 10856–10862.
80 C. P. Owens, F. E. H. Katz, C. H. Carter, M. A. Luca and F. A. Tezcan, *J. Am. Chem. Soc.*, 2015, **137**, 12704–12712.
81 C. P. Owens, F. E. H. Katz, C. H. Carter, V. F. Oswald and F. A. Tezcan, *J. Am. Chem. Soc.*, 2016, **138**, 10124–10127.
82 S. M. Mayer, C. A. Gormal, B. E. Smith and D. M. Lawson, *J. Biol. Chem.*, 2002, **277**, 35263–35266.
83 F. A. Tezcan, J. T. Kaiser, J. B. Howard and D. C. Rees, *J. Am. Chem. Soc.*, 2015, **137**, 146–149.
84 F. Tezcan, J. Kaiser, D. Mustafi, M. Walton, J. Howard and D. Rees, *Science*, 2005, **309**, 1377–1380.
85 C. N. Morrison, J. A. Hoy, L. Zhang, O. Einsle and D. C. Rees, *Biochemistry*, 2015, **54**, 2052–2060.
86 B. Schmid, M. W. Ribbe, O. Einsle, M. Yoshida, L. M. Thomas, D. R. Dean, D. C. Rees and B. K. Burgess, *Science*, 2002, **296**, 352–356.
87 B. Schmid, O. Einsle, H. J. Chiu, A. Willing, M. Yoshida, J. B. Howard and D. C. Rees, *Biochemistry*, 2002, **41**, 15557–15565.
88 M. Rohde, K. Grunau and O. Einsle, *Angew. Chem., Int. Ed.*, 2020, **59**, 1–5.
89 M. Rohde, K. Laun, I. Zebger, S. T. Stripp and O. Einsle, *Sci. Adv.*, 2021, **7**, eabg4474.
90 D. Sippel, M. Rohde, J. Netzer, C. Trncik, J. Gies, K. Grunau, I. Djurdjevic, L. Decamps, S. L. A. Andrade and O. Einsle, *Science*, 2018, **359**, 1484–1489.
91 D. Sippel and O. Einsle, *Nat. Chem. Biol.*, 2017, **13**, 956–960.