ChemComm

Chemical Communications

Accepted Manuscript

ROYAL SOCIETY OF CHEMISTRY

rsc.li/chemcomm

COMMUNICATION

Redox- and metal-directed structural diversification in designed metalloprotein assemblies

Albert Kakkis[a], Eyal Golub[a], Tae Su Choi[a], and F. Akif Tezcan[a]

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

Abstract

Herein we describe a designed protein building block whose self-assembly behaviour is dually gated by the redox state of disulphide bonds and the identity of exogenous metal ions. This protein construct is shown–through extensive structural and biophysical characterization–to access five distinct oligomeric states, exemplifying how the complex interplay between hydrophobic, metal-ligand, and reversible covalent interactions could be harnessed to obtain multiple, responsive protein architectures from a single building block.

The propensity of a single protein sequence to form multiple conformations or assembly states has been crucial for the generation of structural and functional diversity during evolution.[1-3] For instance, protein folds such as the Rossman, four-helix bundle, and βαββββ motifs have been repeatedly used as modular building blocks for larger architectures or quaternary assemblies with a wide variety of functions.[4-6] Similarly, obtaining multiple structural outcomes from a single protein sequence also is a prerequisite for building switchable systems that transduce external stimuli into functionally relevant changes to their tertiary folds or quaternary assembly states.[7-10] Inspired by such natural examples, there has been great interest in designing proteins that can alter their conformations or alter their assembly states in response to different stimuli, such as ligand binding,[11, 12] metal coordination,[13, 14] phosphorylation,[15, 16] and cysteine oxidation/reduction.[17, 18] While there have indeed been several examples of such artificial multi-state systems,[3, 11-20] the ability to design proteins that respond to more than one type of stimulus or to obtain more than two structurally distinct states from a single protein sequence/structure has been limited (Figure 1a).[21] This is primarily due to the fact that most protein design strategies involve the implementation of extensive noncovalent interactions (in particular, hydrophobic packing) to obtain single, stable structures that correspond to deep free energy minima.[21-24] This strategy not only restricts the potential of structural diversification but lowers the potential for the resulting protein architecture to be stimuli-responsive and reconfigurable.

Due to their simultaneous strength and reversibility, metal-ligand and disulphide bonding interactions represent promising conduits for the design of protein constructs that can access multiple structural states in a stimuli-responsive manner..[18, 25-28] We have previously exploited metal coordination, disulphide bonding, and hydrophobic packing to construct cytochrome (cyt) cb_{562}-based assemblies with diverse conformations, oligomeric states and metal coordination environments.[13, 29-34] MBPC1, an early designed variant of cyt cb_{562}, was shown to form different assemblies with distinct oligomeric states and structures based on the coordination preferences of

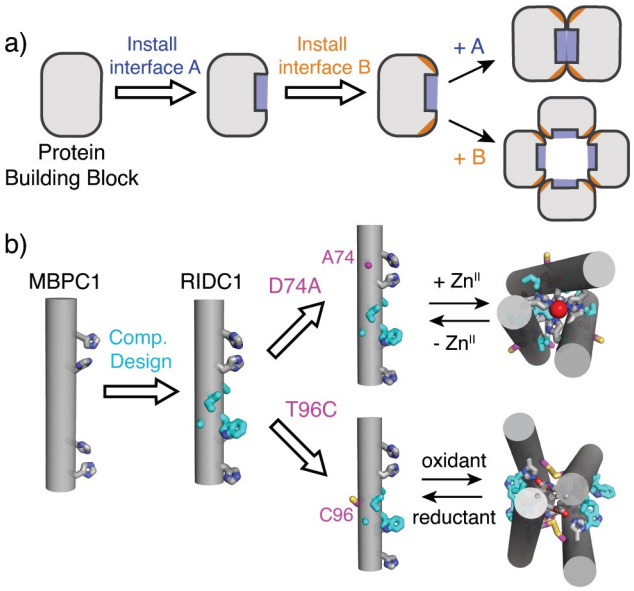

Fig. 1. (a) General workflow to design a multi-stimuli responsive protein construct. A and B represent different stimuli. (b) Cartoon schemes of previously designed cytochrome cb_{562} variants.

[a] Department of Chemistry and Biochemistry, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA, 92093 USA

Electronic Supplementary Information (ESI) available: [details of any supplementary information available should be included here]. See DOI: 10.1039/x0xx00000x

exogenously added metal ions.[13, 29] One of these assemblies, the Zn-directed tetramer Zn_4:MBPC1_4, served as a structural template for the computational design of a hydrophobic interface (highlighted in cyan in Figure 1b) on the surface of MBPC1.[31] Owing to the designed interactions between the hydrophobic surface residues, the resulting variant, RIDC1, formed a considerably more stable Zn-directed tetramer (Zn_4:RIDC1_4) with a nearly identical structure to that of Zn_4:MBPC1_4.[31] Importantly, Zn_4:RIDC1_4 served as a starting point for designing assemblies with functions that ranged from selective metal binding and metal-based allostery to *in vivo* enzymatic activity.[32, 35, 36]

In the course of our previous studies, we observed that single mutations of the RIDC1 construct alter the assembly outcomes.[37, 38] One RIDC1 variant, ^{C96}RIDC1, formed a redox-dependent but metal-independent tetramer (^{C96}RIDC1_4) stabilized by both hydrophobic and Cys96-Cys96 disulphide bonding (Figure 1b, S1a).[37] A second RIDC1 variant, ^{A74}RIDC1 (wherein a metal binding residue Asp74 was mutated to Ala) assembled into a Zn-dependent trimer (Zn_2:^{A74}RIDC1_3) stabilized by hydrophobic packing interactions and tetrahedral, Zn:His_4 coordination sites that had not been observed in other RIDC1 variants (Figure 1b, S1b).[38] Here, with the aim of developing a protein construct that can respond to redox and metal-based stimuli to access multiple structural states, we combined the A74 and C96 mutations to generate $^{A74/C96}$RIDC1 (Figure 2). We found that the interplay between hydrophobic, metal-ligand, and covalent interactions enabled this variant to form five discrete structural states in a redox- and metal-responsive fashion (Figure 2).

We surmised that in the oxidized state of $^{A74/C96}$RIDC1 ($^{A74/C96}$RIDC1^{ox}), the Cys96-Cys96 disulphide bonds would enforce tetramerization, as observed in the case of ^{C96}RIDC1.[37] Indeed, both in the absence and presence of metal ions (Co^{II}, Ni^{II}, Cu^{II}, Zn^{II}), $^{A74/C96}$RIDC1^{ox} formed a tetrameric species in solution in near quantitative yields as determined by sedimentation velocity-analytical ultracentrifugation (SV-AUC) measurements (Figure S2). The crystal structures of Co^{II}-, and Zn^{II}-bound [$^{A74/C96}$RIDC1^{ox}]$_4$ are nearly identical to one another, with a root-mean-square deviation (RMSD) of 1.24 Å between all α-C's (Figure S3-S5). On average, the buried surface area (BSA) of the metal-bound tetramers is about 40% smaller (1018 Å^2) than that of the apo structure (1388 Å^2) (Table S1). This indicates that the tetrameric assembly undergoes a significant structural change upon metal binding, with an average RMSD of 2.55 Å between apo and metal-bound structures (Figure S3). Both Co and Zn-bound [$^{A74/C96}$RIDC1^{ox}]$_4$ tetramers contain two C_2-symmetry-related coordination sites, with E81 and H77 residues from two different monomers serving as ligands. (Figure S4-7). The [$^{A74/C96}$RIDC1^{ox}]$_4$ structures illustrate that simultaneously exploiting the flexibility of and structural constraints imposed by disulphide bonds and hydrophobic packing interactions can engender a flexible protein assembly with well-defined metal coordination sites (Figure 2, S4-7).

We next turned to the reduced form of our construct, $^{A74/C96}$RIDC1^{red}, with the hypothesis that the lack of disulphide-mediated interfacial constraints could allow it to access

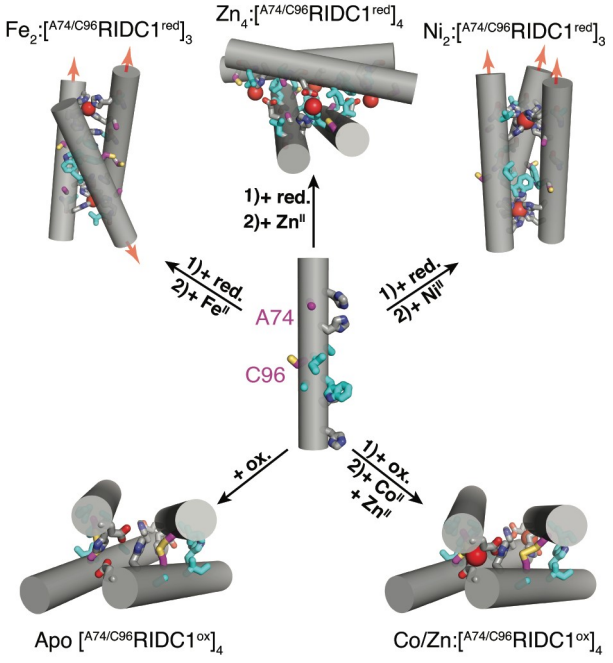

Fig. 2. Structural states of $^{A74/C96}$RIDC1 obtained through the addition of redox and/or metal-based stimuli. Hydrophobic mutations are highlighted in cyan. Ox. = oxidant, red. = reductant.

different oligomeric states upon metal coordination. $^{A74/C96}$RIDC1^{red} was obtained by adding 5-fold excess of the reductant tris(3-hydroxylpropyl) phosphine (THPP). At a protein concentration of ≥200 μM and up to 5-fold excess of metal ions, Fe^{II}, Ni^{II} and Cu^{II} addition to $^{A74/C96}$RIDC1^{red} led primarily to trimeric species in solution, whereas Zn^{II} and Co^{II} addition yielded tetrameric or higher-order assemblies (Figure S8, Table S2). All metal-directed $^{A74/C96}$RIDC1^{red} oligomers could be completely disassembled by the addition of a mixture of ethylenediaminetetraacetic acid (EDTA) and dipicolinic acid (DPA) (Figure S8).

The crystal structures of Fe^{II}-, Ni^{II}-, Cu^{II}-, and Zn^{II}-directed assemblies of $^{A74/C96}$RIDC1^{red} were determined at resolutions of 1.6 Å to 2.7 Å (Table S3). These structures revealed a correspondence between the oligomerization states observed in solution and crystals for the Fe^{II} (n=3), Ni^{II} (n=3) and Zn^{II} (n=4) complexes (Figure S5b, S9-10). By contrast, there was a deviation in the case of the Cu^{II}-directed $^{A74/C96}$RIDC1^{red} assembly (n=4 in crystals vs. n=3 in solution) (Figure S11). A closer look at the latter structure showed that all four Cu centres in the tetrameric assembly adopted a tetrahedral coordination geometry, strongly suggesting that they were in the +1 oxidation state and thus reduced by the excess THPP present in the crystallization solution (Figure S7a, S11, Table S4). In light of the complex redox equilibrium that exists between Cu ions, Cys-disulphide bonds, and THPP, we decided to focus our further analyses on Fe^{II}, Ni^{II} and Zn^{II} complexes of $^{A74/C96}$RIDC1^{red}.

Interestingly, the Fe^{II}- and Ni^{II}-directed assemblies, while both trimeric, adopt different structural conformations and metal coordination environments (Figure 3). The Fe_2:[$^{A74/C96}$RIDC1^{red}]$_3$ complex features two protein monomers with their C-termini projecting downward and one monomer

with its C-terminus projecting upward, resulting in an antiparallel, "up-up-down" arrangement similar to that observed for Zn_2:^{A74}RIDC1$_3$ (Figure 2, 3a, S9, S12).[38] The two Fe^{II} centres, termed Fe1 and Fe2, are distinct from one another, with Fe1 in a square pyramidal geometry formed by five His residues and Fe2 in a similar geometry but with three His and two aqua ligands (Figure 3a, S6a, S9).

By contrast, the Ni_2:[$^{A74/C96}$RIDC1^{red}]$_3$ assembly has C_3 symmetry with an all-parallel, "up-up-up" arrangement of protein monomers and two octahedral, hexa-His-coordinated

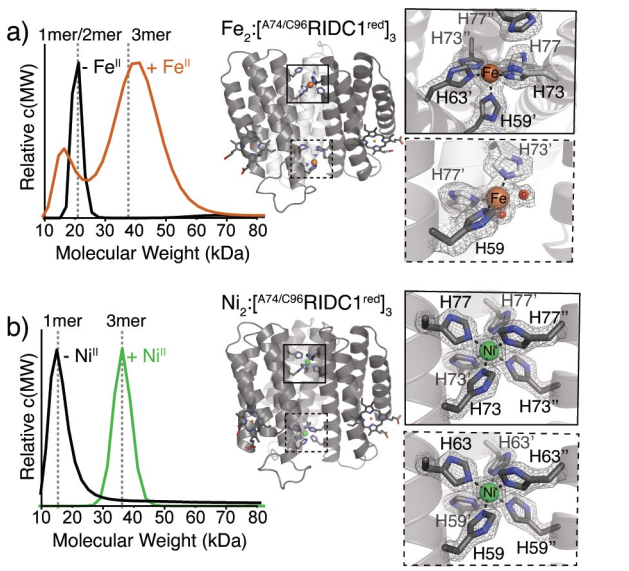

Fig. 3. Assembly properties of Fe- and Ni-directed $^{A74/C96}$RIDC1 trimers. (a) SV-AUC profile of 200 μM $^{A74/C96}$RIDC1 following the addition/removal of 5 equiv. Fe^{II}/monomer (left). Crystal structure of Fe_2:[$^{A74/C96}$RIDC1^{red}]$_3$ (PDB ID: 7RWY), highlighting His_3 and His_5 coordination sites (right). (b) SV-AUC profile of 200 μM $^{A74/C96}$RIDC1 following the addition/removal of 1 equiv. Ni^{II}/monomer (left). Crystal structure of Ni_2:[$^{A74/C96}$RIDC1^{red}]$_3$ (PDB ID: 7RWU), highlighting the His_6 coordination sites (right).

Ni^{II} centres (Figure 2, 3b, S6c, and S10). The structure of Ni_2:[$^{A74/C96}$RIDC1^{red}]$_3$ is nearly identical to that of the previously characterized Ni_2:MBPC1$_3$ assembly (RMSD = 0.66 Å, Figure S12).[13] While buried surface area and Rosetta interface calculations predict that the "up-up-up" trimer is less stable compared to the "up-up-down" configuration based purely on interfacial hydrophobic interactions (Table S1, S5), we propose based on DFT calculations of the metal coordination sites that the stability of the two octahedral Ni^{II}:His_6 coordination motifs is sufficiently high to favour the assembly of the "up-up-up" trimer (Figure S13-14, Table S6; also see Supplementary Discussion).

In contrast to the Fe^{II} and Ni^{II}-directed assemblies, the crystal structure of the Zn^{II}-directed assembly revealed a tetrameric, D_4-symmetric architecture (Zn_4:[$^{A74/C96}$RIDC1^{red}]$_4$) in which the free Cys96 residues coordinate the metal ion (Figure 4a, S5b). The assembly features four identical, tetrahedral $His_2GluCys$ coordination sites (Figure 4a, S7b). Each antiparallel dimer is oriented about 75° with respect to the other (Figure 4b). This canted arrangement contrasts with the nearly collinear arrangement (θ = 21°) of antiparallel dimers in Zn_4:[$^{A74/C96}$RIDC1^{ox}]$_4$ (Figure 4b). Taken together, the Zn_4:[$^{A74/C96}$RIDC1^{red}]$_4$ and Zn-bound [$^{A74/C96}$RIDC1^{ox}]$_4$ structures illustrate the dual functional role of cysteine as a metal coordinating ligand and covalent handle, which is critical to redox signalling in biological systems.[10, 39, 40] Our results demonstrate that $^{A74/C96}$RIDC1^{red} can adopt three unique architectures in the presence of different metal ions (Figure 2-4), exemplifying how metal coordination preferences and hydrophobic packing can collectively influence assembly outcomes.

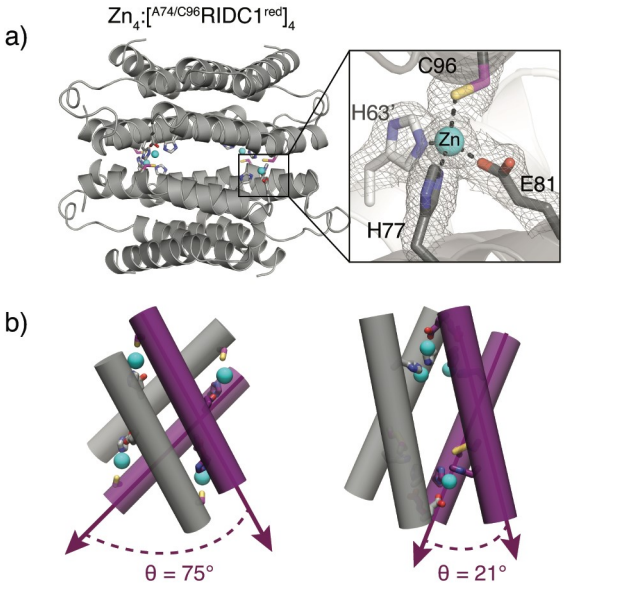

Fig. 4. (a) Crystal structure of Zn_4:[$^{A74/C96}$RIDC1^{red}]$_4$ (PDB ID: 7RWX), highlighting one of four identical tetrahedral Zn coordination sites featuring Cys96 in the primary sphere. (b) Cartoon depictions of Zn_4:[$^{A74/C96}$RIDC1^{red}]$_4$ (left) and Zn_4:[$^{A74/C96}$RIDC1^{ox}]$_4$ (right) illustrating significant conformational differences between the two assemblies.

The ability of a single protein construct to assemble into and interconvert between multiple structural states is crucial for generating functional diversity and the generation of switchable systems.[1, 2, 8, 9, 41] Herein, we have demonstrated that a designed protein ($^{A74/C96}$RIDC1) can be subjected to redox- and metal-based stimuli to obtain five structurally distinct assemblies (Figure 2). The large structural diversity of $^{A74/C96}$RIDC1 assemblies can be attributed to an intricate interplay between metal-ligand, disulphide bonding, and hydrophobic interactions. While potentially serving as starting points for engineering downstream functions, the dynamic $^{A74/C96}$RIDC1 assemblies also pave the path to the generation of multistate protein switches.

This work was funded by NIH (R01-GM138884 and T32-GM112584) and NASA (80NSSC18M0093; ENIGMA: Evolution of Nanomachines in Geospheres and Microbial Ancestors (NASA Astrobiology Institute Cycle 8)). E.G. acknowledges funding by EMBO (ALTF 1336-2015). Portions of this research were carried out at the Stanford Linear Accelerator Center (supported by the DOE, Office of Basic Energy Sciences contract DE-AC02-76SF00515 and NIH P30-GM133894) and the Advanced Light Source at the Lawrence Berkeley National Laboratory (supported by the DOE, Office of Basic Energy Sciences contract

DE-AC02-05CH11231and NIH P30-GM124169-01). Coordinate and structure factor files for the crystal structures have been deposited into the Protein Data Bank (www.rcsb.org) with the following accession codes: 7RWV (Apo [$^{A74/C96}$RIDC1^{ox}]$_4$), 7SU2 (Co_2:[$^{A74/C96}$RIDC1^{ox}]$_4$), 7RWW (Zn_4:[$^{A74/C96}$RIDC1^{ox}]$_4$), 7RWY (Fe_2:[$^{A74/C96}$RIDC1^{red}]$_3$), 7RWU (Fe_2:[$^{A74/C96}$RIDC1^{red}]$_3$), 7TEP (Cu_4:[$^{A74/C96}$RIDC1^{red}]$_4$), and 7RWX (Cu_4:[$^{A74/C96}$RIDC1^{red}]$_4$).

Conflicts of interest

There are no conflicts to declare.

Notes and references

1. L. C. James and D. S. Tawfik, *Trends Biochem. Sci.*, 2003, **28**, 361-368.
2. N. Tokuriki and D. S. Tawfik, *Science*, 2009, **324**, 203-207.
3. J. Zhu, N. Avakyan, A. Kakkis, A. M. Hoffnagle, K. Han, Y. Li, Z. Zhang, T. S. Choi, Y. Na, C. J. Yu and F. A. Tezcan, *Chem. Rev.*, 2021, **121**, 13701-13796.
4. B. D. Sanders, B. Jackson and R. Marmorstein, *Biochim. Biophys. Acta*, 2010, **1804**, 1604-1616.
5. N. D. Chasteen and P. M. Harrison, *J Struct Biol*, 1999, **126**, 182-194.
6. R. N. Armstrong, *Biochemistry*, 2000, **39**, 13625-13632.
7. C. Eicken, M. A. Pennella, X. Chen, K. M. Koshlap, M. L. VanZile, J. C. Sacchettini and D. P. Giedroc, *J. Mol. Biol.*, 2003, **333**, 683-695.
8. B. Groitl and U. Jakob, *Biochim. Biophys. Acta*, 2014, **1844**, 1335-1343.
9. D. M. Rosenbaum, S. G. Rasmussen and B. K. Kobilka, *Nature*, 2009, **459**, 356-363.
10. M. F. Fillat, *Arch. Biochem. Biophys.*, 2014, **546**, 41-52.
11. A. Ghanbarpour, C. Pinger, R. Esmatpour Salmani, Z. Assar, E. M. Santos, M. Nosrati, K. Pawlowski, D. Spence, C. Vasileiou, X. Jin, B. Borhan and J. H. Geiger, *J. Am. Chem. Soc.*, 2019, **141**, 17125-17132.
12. J. M. Karchin, J. H. Ha, K. E. Namitz, M. S. Cosgrove and S. N. Loh, *Sci. Rep.*, 2017, **7**, 44388.
13. E. N. Salgado, R. A. Lewis, S. Mossin, A. L. Rheingold and F. A. Tezcan, *Inorg. Chem.*, 2009, **48**, 2726-2728.
14. Y. Bai, Q. Luo, W. Zhang, L. Miao, J. Xu, H. Li and J. Liu, *J. Am. Chem. Soc.*, 2013, **135**, 10966-10969.
15. R. S. Signarvic and W. F. DeGrado, *J. Mol. Biol.*, 2003, **334**, 1-12.
16. L. Harrington, J. M. Fletcher, T. Heermann, D. N. Woolfson and P. Schwille, *Nat. Commun.*, 2021, **12**, 1472.
17. E. R. Ballister, A. H. Lai, R. N. Zuckermann, Y. Cheng and J. D. Mougous, *Proc. Natl. Acad. Sci. USA*, 2008, **105**, 3733-3738.
18. Y. Suzuki, G. Cardone, D. Restrepo, P. D. Zavattieri, T. S. Baker and F. A. Tezcan, *Nature*, 2016, **533**, 369-373.
19. K. Y. Wei, D. Moschidi, M. J. Bick, S. Nerli, A. C. McShan, L. P. Carter, P. S. Huang, D. A. Fletcher, N. G. Sgourakis, S. E. Boyken and D. Baker, *Proc. Natl. Acad. Sci. USA*, 2020, **117**, 7208-7215.
20. L. A. Campos, R. Sharma, S. Alvira, F. M. Ruiz, B. Ibarra-Molero, M. Sadqi, C. Alfonso, G. Rivas, J. M. Sanchez-Ruiz, A. Romero Garrido, J. M. Valpuesta and V. Munoz, *Nat. Commun.*, 2019, **10**, 5703.
21. R. G. Alberstein, A. B. Guo and T. Kortemme, *Curr. Opin. Struct. Biol.*, 2021, **72**, 71-78.
22. B. Kuhlman, J. W. O'Neill, D. E. Kim, K. Y. Zhang and D. Baker, *Proc. Natl. Acad. Sci. USA*, 2001, **98**, 10687-10691.
23. J. J. Gray, S. Moughon, C. Wang, O. Schueler-Furman, B. Kuhlman, C. A. Rohl and D. Baker, *J. Mol. Biol.*, 2003, **331**, 281-299.
24. A. Leaver-Fay, M. Tyka, S. M. Lewis, O. F. Lange, J. Thompson, R. Jacak, K. Kaufman, P. D. Renfrew, C. A. Smith, W. Sheffler, I. W. Davis, S. Cooper, A. Treuille, D. J. Mandell, F. Richter, Y. E. Ban, S. J. Fleishman, J. E. Corn, D. E. Kim, S. Lyskov, M. Berrondo, S. Mentzer, Z. Popovic, J. J. Havranek, J. Karanicolas, R. Das, J. Meiler, T. Kortemme, J. J. Gray, B. Kuhlman, D. Baker and P. Bradley, *Meth. Enzymol.*, 2011, **487**, 545-574.
25. J. D. Brodin, X. I. Ambroggio, C. Tang, K. N. Parent, T. S. Baker and F. A. Tezcan, *Nat. Chem.*, 2012, **4**, 375-382.
26. J. D. Brodin, S. J. Smith, J. R. Carr and F. A. Tezcan, *J. Am. Chem. Soc.*, 2015, **137**, 10468-10471.
27. M. Yang and W. J. Song, *Nat. Commun.*, 2019, **10**, 5545.
28. S. R. Rao, S. L. Schettler and W. S. Horne, *Chempluschem*, 2021, **86**, 137-145.
29. E. N. Salgado, J. Faraone-Mennella and F. A. Tezcan, *J. Am. Chem. Soc.*, 2007, **129**, 13374-13375.
30. R. J. Radford, P. C. Nguyen, T. B. Ditri, J. S. Figueroa and F. A. Tezcan, *Inorg. Chem.*, 2010, **49**, 4362-4369.
31. E. N. Salgado, X. I. Ambroggio, J. D. Brodin, R. A. Lewis, B. Kuhlman and F. A. Tezcan, *Proc. Natl. Acad. Sci. USA*, 2010, **107**, 1827-1832.
32. A. Medina-Morales, A. Perez, J. D. Brodin and F. A. Tezcan, *J. Am. Chem. Soc.*, 2013, **135**, 12013-12022.
33. J. Rittle, M. J. Field, M. T. Green and F. A. Tezcan, *Nat. Chem.*, 2019, **11**, 434-441.
34. A. Kakkis, D. Gagnon, J. Esselborn, R. D. Britt and F. A. Tezcan, *Angew. Chem. Int. Ed.*, 2020, **59**, 21940-21944.
35. L. A. Churchfield, A. Medina-Morales, J. D. Brodin, A. Perez and F. A. Tezcan, *J. Am. Chem. Soc.*, 2016, **138**, 13163-13166.
36. W. J. Song and F. A. Tezcan, *Science*, 2014, **346**, 1525-1528.
37. J. D. Brodin, A. Medina-Morales, T. Ni, E. N. Salgado, X. I. Ambroggio and F. A. Tezcan, *J. Am. Chem. Soc.*, 2010, **132**, 8610-8617.
38. T. W. Ni and F. A. Tezcan, *Angew. Chem. Int. Ed.* , 2010, **49**, 7014-7018.
39. W. Maret, *Biometals*, 2009, **22**, 149-157.
40. B. D'Autreaux, L. Pecqueur, A. Gonzalez de Peredo, R. E. Diederix, C. Caux-Thang, L. Tabet, B. Bersch, E. Forest and I. Michaud-Soret, *Biochemistry*, 2007, **46**, 1329-1342.
41. F. L. Takken, M. Albrecht and W. I. Tameling, *Curr. Opin. Plant Biol.*, 2006, **9**, 383-390.