



A De Novo Designed Metalloenzyme for the Hydration of CO₂**

Virginia M. Cangelosi, Aniruddha Deb, James E. Penner-Hahn, and Vincent L. Pecoraro*

Abstract: Protein design will ultimately allow for the creation of artificial enzymes with novel functions and unprecedented stability. To test our current mastery of nature's approach to catalysis, a Zn^{II} metalloenzyme was prepared using *de novo* design. α_3 DH₃ folds into a stable single-stranded three-helix bundle and binds Zn^{II} with high affinity using His₃O coordination. The resulting metalloenzyme catalyzes the hydration of CO₂ better than any small molecule model of carbonic anhydrase and with an efficiency within 1400-fold of the fastest carbonic anhydrase isoform, CAII, and 11-fold of CAIII.

Protein design is an increasingly popular approach for studying and modeling the structure–function relationships in proteins.^[1] There is a growing interest in the development of artificial enzymes that can perform with the efficiency of natural enzymes toward reactions not normally seen in nature. Specifically, artificial metalloenzymes are important design targets because over one-third of natural proteins use metal ions for structural, catalytic, and/or electron-transfer functions. There are two main metalloprotein design strategies: protein redesign and *de novo* design. The former approach involves the introduction of a metal-binding site into an existing, stable protein. The latter relies on first principles to design well-defined structures from amino acid sequences not found in nature. *De novo* design is challenging owing to its requirement for complete control over folding and function, but it can lead to significant insight into the nature of metal–enzyme interactions. Several recent examples showcase the power of *de novo* design in creating metalloenzymes with enzymatic activity for ester hydrolysis,^[2–5] nitrite reduction,^[6] oxidation,^[7,8] and *N*-hydroxylation.^[9]

One test of our current mastery of protein design is to model the activity of a natural enzyme. Carbonic anhydrase, which is ubiquitous in animals, plants, and bacteria, is essential for respiration, vision, regulation of acid–base equilibrium, and many other processes. One of the most

efficient enzymes in nature, human carbonic anhydrase II (CAII), catalyzes the reversible hydration of CO₂ with a catalytic efficiency approaching the diffusion limit in water. The importance of this metalloenzyme, in conjunction with its well-studied mechanism, solid-state structure, and inhibition, make it a highly appealing, yet challenging, target to assess our ability to replicate the activity of a natural metalloenzyme in a simplified system.

Previously, we presented a highly efficient CA-mimic by modeling the active site within a *de novo* designed three-stranded coiled-coil (3SCC). The bifunctional Hg^{II}₂Zn^{II}_N(TRIL9CL23H)₃ closely models the structure of the Zn^{II}His₃OH primary coordination sphere of CAII, yet places it within a very different fold (α -helices vs β -sheets in CAII), and contains an additional Hg^{II}Cys₃ site which provides structural stability. The metalloenzyme catalyzes CO₂ hydration with an efficiency comparable to some naturally occurring CAs and within 350-fold of the fastest isozyme, CAII. While this is the fastest CA-model to date, improvements to the system are limited by the inherent symmetry resulting from the self-assembly of three parallel α -helices. Not only are antiparallel helices more typical in nature,^[10] but CAII contains a network of H-bonds that cannot currently be modeled in this system. Therefore, a new approach is required.

To allow for asymmetry in the secondary sphere of the active site, we have designed a metalloenzyme starting from α_3 D, a *de novo* single-stranded antiparallel three-helix bundle. Designed by DeGrado and co-workers,^[13] this 73 amino acid protein folds with native protein-like stability, is tolerant of mutations within the hydrophobic core,^[14] and has been structurally characterized by NMR spectroscopy.^[12] Previously, our lab incorporated a Cys₃ metal binding site near the C-terminus of α_3 D and showed that the resulting protein, α_3 DIV, binds Hg^{II}, Pb^{II}, and Cd^{II} with high affinity in coordination geometries previously identified within the TRI family of 3SCCs.^[15] Herein, we report a new metalloenzyme, α_3 DH₃, which contains a His₃ site that, upon binding Zn^{II}, catalyzes the hydration of CO₂ (Figure 1).

α_3 DH₃ differs from α_3 D in that three leucine residues were replaced with histidine residues (L18H, L28H, L67H), a histidine residue was replaced with valine (H72V) to ensure no competition for Zn^{II} binding, and four extra residues were added to the end of the chain (GSGA) which improved expression yields (Table 1). Expressed from a synthetic gene in *Escherichia coli*, α_3 DH₃ was purified by high-performance liquid chromatography (HPLC) and characterized by electrospray ionization mass spectroscopy (EI-MS). The observed molecular weight of 8283.5 Da corresponds to the protein after deletion of the N-terminal methionine residue (MW_{calc} = 8283.1 Da). The protein is well-folded and stable according to circular dichroism spectroscopy (Supporting Information,

[*] Dr. V. M. Cangelosi, Dr. A. Deb, Prof. J. E. Penner-Hahn, Prof. V. L. Pecoraro
Department of Chemistry, University of Michigan
Ann Arbor MI, 48109 (USA)
E-mail: vlpec@umich.edu

[**] Research reported in this publication was supported by the National Institutes of Health under Award Numbers F32GM100543 (National Institute of General Medical Sciences) (V.M.C.) and R01ES012236 (National Institute of Environmental Health Sciences) (V.L.P.). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Supporting information for this article is available on the WWW under <http://dx.doi.org/10.1002/ange.201404925>.

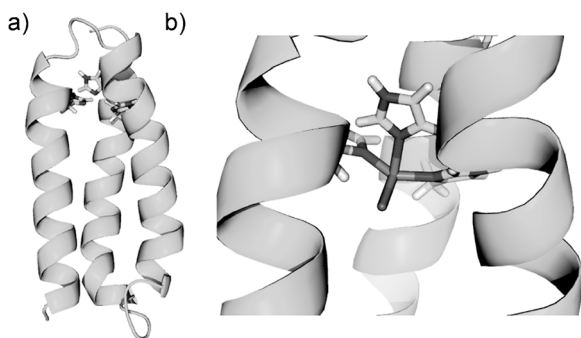


Figure 1. PyMol^[11] models of $\alpha_3\text{DH}_3$ showing a) the entire bundle (last four residues removed for simplicity) and b) the $\text{Zn}^{\text{II}}\text{His}_3\text{O}$ site incorporating EXAFS Zn-N/O distances. Models are based on the NMR solution structure of $\alpha_3\text{D}$ (PDB: 2A3D).^[12]

Table 1: Designed peptide sequences.^[a]

Peptide	Sequence
$\alpha_3\text{DH}_3$	MGSWAEFKQRLAAIKTR HQ ALGG SEAE H AFAFEKEIAAFESEIQAYKGKGNPE VEALRKEAAAIRDE HQ AYR V NGSGA
$\alpha_3\text{DH72V}$	MGSWAEFKQRLAAIKTRLQALGG SEAE L AFAFEKEIAAFESEIQAYKGKGNPE VEALRKEAAAIRDELQAYR V N

[a] Mutations with respect to $\alpha_3\text{D}$ are in bold and italics.

Figure S1). The observed double well at 208 and 222 nm is characteristic of α -helical proteins, and the molar ellipticities suggest that the protein is 82% folded at pH 9. Chemical denaturations with guanidine hydrochloride were fit to a two-state unfolding model, revealing a $\Delta G_u = 3.1 \text{ kcal mol}^{-1}$. Although the histidine substitutions destabilized the protein compared to native $\alpha_3\text{D}$ ($\Delta G_u = 5.1 \text{ kcal mol}^{-1}$), the construct is stable enough for use in kinetic studies. The presence of Zn^{II} caused no change in the folding or stability of $\alpha_3\text{DH}_3$ (Supporting Information, Figure S1).

Apparent Zn^{II} binding constants were measured by UV/Vis spectroscopy using Zincon as a colorimetric probe (Supporting Information, Figures S2–S4). $\alpha_3\text{DH}_3$ binds Zn^{II} with a $150 \pm 40 \text{ nM}$ affinity at pH 7.5, which strengthens to $59 \pm 9 \text{ nM}$ at pH 9.0. A control peptide, $\alpha_3\text{DH72V}$, which lacks the His_3 binding site, binds with 20-fold weaker affinity ($K_d = 1.2 \pm 0.4 \text{ }\mu\text{M}$ at pH 9.0), confirming that the histidine site is involved in specific Zn^{II} binding. These affinities are weaker than that of CAII measured by equilibrium dialysis ($K_d = 0.8 \pm 0.1 \text{ }\mu\text{M}$);^[16] however, recent isothermal titration calorimetry experiments suggest a three-orders of magnitude weaker affinity for CAII ($K_d = 0.45 \text{ nM}$).^[17] Based on this newer measurement, $\alpha_3\text{DH}_3$ has an affinity only two orders of magnitude weaker than CAII and stronger than those observed for our previously published $\text{Hg}^{\text{II}}_5\text{Zn}^{\text{II}}_N(\text{TRIL9CL23H})_3$ ($K_d = 0.8 \pm 0.1$ and $0.22 \pm 0.06 \text{ }\mu\text{M}$ at pH 7.5 and 9.0, respectively).^[18]

A sample of $\text{Zn}^{\text{II}}\alpha_3\text{DH}_3$ at pH 9.0 was analyzed by extended X-ray absorption fine structure (EXAFS) spectroscopy. The EXAFS data is dominated by nearest-neighbor scattering typical of oxygen or nitrogen ligation and can be

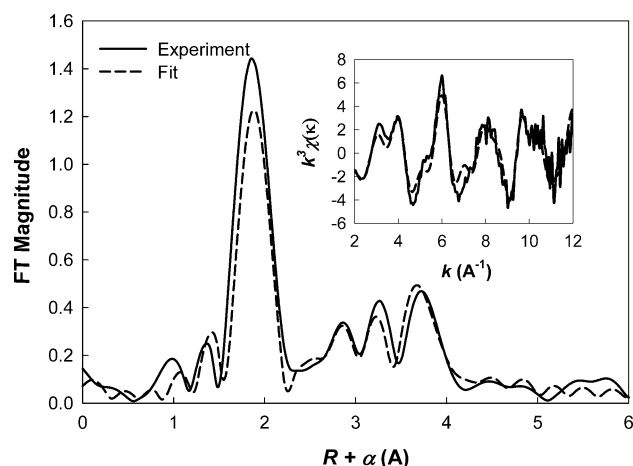


Figure 2. Fourier transform (FT) of the EXAFS spectra for $\text{Zn}^{\text{II}}\alpha_3\text{DH}_3$. Inset is k^3 -weighted EXAFS spectra used to calculate the FT. — Experimental data, ---- best-fit curve.

modeled with a single oxygen/nitrogen shell at about $1.98 \text{ }\text{\AA}$, a distance consistent with a four-coordinate Zn site. There is clear evidence of outer shell scattering typical of histidine ligated metals (that is, features at $R + \alpha > 2 \text{ }\text{\AA}$ in Figure 2). Given the relatively limited k range of these data, there are multiple models with similar fit quality, making it difficult to define the number of histidine ligands from EXAFS alone. However, the best fit, both in terms of mean-square deviation and in terms of the presence of physically reasonable fit parameters, uses one oxygen at $1.90 \text{ }\text{\AA}$ and 3 histidines at $1.99 \text{ }\text{\AA}$. These parameters are very similar to EXAFS distances measured for CAII (Zn-N/O of $1.98 \text{ }\text{\AA}$).^[17]

The true indication of a successfully designed metalloenzyme mimic is catalytic activity toward the physiological reaction of the natural enzyme, in this case, the hydration of CO_2 . Using Khalif's^[19] stopped-flow indicator technique, $\text{Zn}^{\text{II}}\alpha_3\text{DH}_3$ was found to be an efficient catalyst with activities that increase with pH (Table 2). Testing the activity at a pH higher than 9.5 is not feasible because the stability of the bundle is compromised when interhelical salt bridges are deprotonated. Fitting the catalytic efficiencies for pH 8–9.5, a maximal efficiency of $6.9 \times 10^4 \text{ L mol}^{-1} \text{ s}^{-1}$ and a kinetic $\text{p}K_a$ of 9.4 were derived (Figure 3). CAII has a $\text{p}K_a$ of 6.8^{20} for the deprotonation of Zn^{II} -bound water to give the active hydroxide complex. The $\text{p}K_a$ measured here might be for the same process (as was suggested for the observed $\text{p}K_a$ (8.8) for p -nitrophenylacetate hydrolysis by $\text{Hg}^{\text{II}}_5\text{Zn}^{\text{II}}_N(\text{TRIL9CL23H})_3$; however, it could also be for a lysine residue near the active site, which, upon deprotonation, would open up the bundle and allow improved substrate access. In either case, catalysis occurs at the $\text{Zn}^{\text{II}}\text{His}_3\text{O}$ site; even in the presence of Zn^{II} the control peptide $\alpha_3\text{DH72V}$ shows significantly lower activity (Supporting Information, Figure S5).

Numerous small molecule models for CAII have been reported, with varying degrees of structural similarity to the enzyme active site.^[21–28] Many of these show some catalytic activity toward the hydration of CO_2 , with the fastest catalysts listed in Table 2. Both $\text{Zn}^{\text{II}}(\text{tris}(4,5\text{-di-}n\text{-propyl-2-imidazolyl)})$ -

Table 2: Kinetic parameters for the catalysis of CO₂ hydration by enzymes and model complexes.

Enzyme/Model	pK _a	pH	k _{cat} [s ⁻¹]	K _M [mM]	k _{cat} /K _M [L mol ⁻¹ s ⁻¹]	k ₂ [L mol ⁻¹ s ⁻¹]	Reference
CAII	6.8 ^[a]	8.8	8.2 × 10 ⁵	8.9	9.2 × 10 ⁷		[29]
CAIII	8.5	9.0	8 × 10 ³	20	4 × 10 ⁵		[30]
Hg ^{II} ₅ Zn ^{II} _N (TRIL9CL23H) ₃ ⁿ⁺	8.82 ^[b]	9.5	(1.8 ± 0.4) × 10 ³	10.0 ± 2.4	(1.8 ± 0.3) × 10 ⁵		[2]
Zn ^{II} α ₃ DH ₃	9.4	9.5	134 ± 8	3.5 ± 0.6	(3.8 ± 0.5) × 10 ⁴		this work
		9.25	137 ± 8	4.1 ± 0.7	(3.2 ± 0.3) × 10 ⁴		
		9.0	103 ± 8	4.9 ± 1.0	(2.1 ± 0.3) × 10 ⁴		
		8.75	107 ± 12	7.3 ± 1.8	(1.5 ± 0.2) × 10 ⁴		
		8.5	82 ± 6	7.2 ± 1.1	(1.1 ± 0.1) × 10 ⁴		
		8.0	39 ± 4	6.2 ± 1.7	(0.63 ± 0.1) × 10 ⁴		
Zn ^{II} ([14]aneN4)	9.8	ind ^[c]				5040 ^[c]	[35]
Zn ^{II} ([12]aneN4)	8.1	9.1				3012 ± 193	[27]
Zn ^{II} (nitrilotris(2-benzimidazolymethyl-6-sulfonate))	8.3	9.5				3180 ^[d]	[28]
Zn ^{II} (tris(4,5-di- <i>n</i> -propyl-2-imidazolyl)phosphine)	8.0	6.55				2480 ^[e]	[25]

[a] Taken from Ref. [20]. [b] pK_a measured for the hydrolysis of *p*NPA. [c] pH-independent second-order rate constant (maximal rate) calculated using pK_a of 9.8 and measured rate of 690 M⁻¹ s⁻¹ at pH 9.0. [d] Second-order rate constant measured at 15 °C. [e] Second-order rate constant was measured for the dehydration of HCO₃⁻ + H⁺ in 80% EtOH/H₂O.

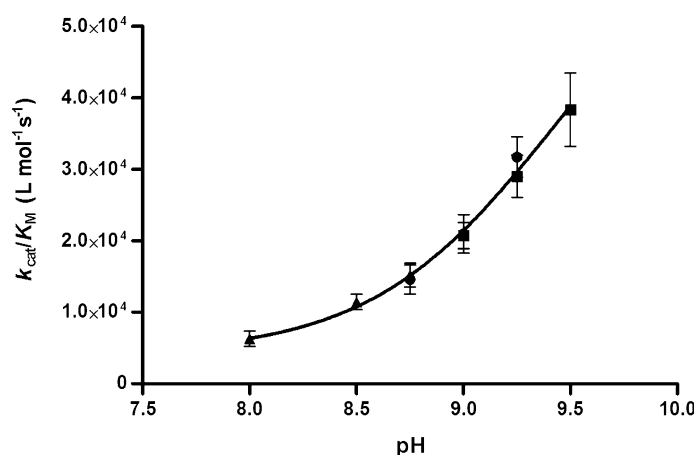


Figure 3. Plot of pH dependency of the catalytic efficiency of CO₂ hydration by Zn^{II}-α₃DH₃. Standard errors for the fit of the Michaelis–Menten equation to each plot of initial rate versus initial [CO₂]. The buffer/indicator pairs used: ▲ TAPS/*m*-cresol purple (pH 8–8.75), ● AMPSO/thymol blue (pH 8.75–9.25), ■ CHES/thymol blue (pH 9.25–9.5).

phosphine) and Zn^{II}(nitrilotris(2-benzimidazolymethyl-6-sulfonate)) have N₃O coordination of the Zn^{II} ion and show moderate activity.^[25,28] Zn^{II}α₃DH₃ outperforms each of these with a maximal catalytic efficiency that is at least 14-fold higher than their fastest second-order rate constants. Its efficiency is only 2.6-fold slower than that of Hg^{II}₅Zn^{II}_N(TRIL9CL23H)₃ (Table 2), the only other designed enzyme for CO₂ hydration.^[2] Compared to the native enzymes, Zn^{II}α₃DH₃ is 1400-fold slower than CAII and only 11-fold slower than CAIII.^[29,30]

A frequent problem with small-molecule models of enzymes is product inhibition.^[31–33] Using acetate as a more tractable mimic of the reaction product bicarbonate, product inhibition was tested (Supporting Information, Figure S5). Addition of 300 mM KOAc to the assay at pH 8.5 resulted in a decrease in k_{cat} from 82 ± 6 to 66 ± 4 s⁻¹. With an enzyme

concentration of 100 μM and HCO₃⁻ concentrations never exceeding 25 mM, no significant loss in catalytic activity is expected owing to product inhibition. This shows that the peptide is well-designed to avoid product inhibition, highlighting one benefit of de novo design over small molecule catalysts.

In conclusion, we successfully modified an existing de novo designed protein to contain a Zn^{II}His₃ binding site while retaining the desired fold. We report the first pH profile for CO₂ hydration by a de novo designed protein and show that Zn^{II}α₃DH₃ catalyzes this reaction better than any small molecule model and within 1–3 orders of magnitude of natural CA isozymes. The efficiency of our single-stranded construct is slightly lower than our previously reported 3SCC Hg^{II}₅Zn^{II}_N(TRIL9CL23H)₃ and the Zn^{II} affinity is stronger, likely a consequence of differences in the peptide scaffolds. The antiparallel bundle is expected to have a weaker dipole than the coiled-coil, imidazole rings in less-symmetric orientations, and different electrostatics around the binding site. Compared to Hg^{II}₅Zn^{II}_N(TRIL9CL23H)₃, Zn^{II}α₃DH₃ has greater potential for improving catalysis because it is far easier to make asymmetric protein modifications in this system. Furthermore, being an antiparallel construct, Zn^{II}α₃DH₃ more faithfully reproduces helical structures found in nature. In CAII, mutation of Thr199 to alanine lowers the activity 100-fold, showcasing the importance of second coordination sphere residues involved in hydrogen bonds.^[34] Future work will focus on site selective incorporation of acid–base catalytic residues or design of water channels that may tune the Lewis acidity of Zn^{II} to optimize activity.

Received: May 2, 2014

Published online: June 18, 2014

Keywords: carbonic anhydrase · de novo design · metalloenzymes · protein design · zinc enzymes

- [1] F. Yu, V. M. Cangelosi, M. L. Zastrow, M. Tegoni, J. S. Plegaria, A. G. Tebo, C. S. Mocny, L. Ruckthong, H. Qayyum, V. L. Pecoraro, *Chem. Rev.* **2014**, *114*, 3495–3578.
- [2] M. L. Zastrow, A. F. A. Peacock, J. A. Stuckey, V. L. Pecoraro, *Nat. Chem.* **2012**, *4*, 118–123.
- [3] B. S. Der, D. R. Edwards, B. Kuhlman, *Biochemistry* **2012**, *51*, 3933–3940.
- [4] S. D. Khare, Y. Kipnis, P. Greisen, R. Takeuchi, Y. Ashani, M. Goldsmith, Y. Song, J. L. Gallaher, I. Silman, H. Leader, J. L. Sussman, B. L. Stoddard, D. S. Tawfik, D. Baker, *Nat. Chem. Biol.* **2012**, *8*, 294–300.
- [5] C. M. Rufo, Y. S. Moroz, O. V. Moroz, J. Stöhr, T. A. Smith, X. Hu, W. F. DeGrado, I. V. Korendovych, *Nat. Chem.* **2014**, *6*, 303–309.
- [6] M. Tegoni, F. Yu, M. Bersellini, J. E. Penner-Hahn, V. L. Pecoraro, *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 21234–21239.
- [7] M. Faiella, C. Androzzi, R. T. M. de Rosales, V. Pavone, O. Maglio, F. Nastri, W. F. DeGrado, A. Lombardi, *Nat. Chem. Biol.* **2009**, *5*, 882–884.
- [8] J. Kaplan, W. F. DeGrado, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 11566–11570.
- [9] A. J. Reig, M. M. Pires, R. A. Snyder, Y. Wu, H. Jo, D. W. Kulp, S. E. Butch, J. R. Calhoun, T. Szyperski, E. I. Solomon, W. F. DeGrado, *Nat. Chem.* **2012**, *4*, 900–906.
- [10] J. S. Richardson, *Adv. Protein Chem.* **1981**, *34*, 167–339.
- [11] The PyMOL Molecular Graphics System, Version 1.5.0.4 Schrödinger, LLC.
- [12] S. T. R. Walsh, H. Cheng, J. W. Bryson, H. Roder, W. F. DeGrado, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 5486–5491.
- [13] J. W. Bryson, J. R. Desjarlais, T. M. Handel, W. F. DeGrado, *Protein Sci.* **1998**, *7*, 1404–1414.
- [14] S. T. Walsh, V. I. Sukharev, S. F. Betz, N. L. Vekshin, W. F. DeGrado, *J. Mol. Biol.* **2001**, *305*, 361–373.
- [15] S. Chakraborty, J. Y. Kravitz, P. W. Thulstrup, L. Hemmingsen, W. F. DeGrado, V. L. Pecoraro, *Angew. Chem.* **2011**, *123*, 2097–2101; *Angew. Chem. Int. Ed.* **2011**, *50*, 2049–2053.
- [16] J. A. Hunt, C. A. Fierke, *J. Biol. Chem.* **1997**, *272*, 20364–20372.
- [17] H. Song, D. L. Wilson, E. R. Farquhar, E. A. Lewis, J. P. Emerson, *Inorg. Chem.* **2012**, *51*, 11098–11105.
- [18] M. L. Zastrow, V. L. Pecoraro, *J. Am. Chem. Soc.* **2013**, *135*, 5895–5903.
- [19] R. G. Khalifah, *J. Biol. Chem.* **1971**, *246*, 2561–2573.
- [20] C. A. Fierke, T. L. Calderone, J. F. Krebs, *Biochemistry* **1991**, *30*, 11054–11063.
- [21] P. Woolley, *Nature* **1975**, *258*, 677–682.
- [22] J. Huguet, R. S. Brown, *J. Am. Chem. Soc.* **1980**, *102*, 7571–7572.
- [23] R. S. Brown, N. J. Curtis, J. Huguet, *J. Am. Chem. Soc.* **1981**, *103*, 6953–6959.
- [24] R. S. Brown, D. Salmon, N. J. Curtis, S. Kusuma, *J. Am. Chem. Soc.* **1982**, *104*, 3188–3194.
- [25] H. Slebocka-Tilk, J. L. Cocho, Z. Frakman, R. S. Brown, *J. Am. Chem. Soc.* **1984**, *106*, 2421–2431.
- [26] X. Zhang, R. van Eldik, T. Koike, E. Kimura, *Inorg. Chem.* **1993**, *32*, 5749–5755.
- [27] X. Zhang, R. van Eldik, *Inorg. Chem.* **1995**, *34*, 5606–5614.
- [28] K. Nakata, N. Shimomura, N. Shiina, M. Izumi, K. Ichikawa, M. Shiro, *J. Inorg. Biochem.* **2002**, *89*, 255–266.
- [29] J. E. Jackman, K. M. Merz, C. A. Fierke, *Biochemistry* **1996**, *35*, 16421–16428.
- [30] D. A. Jewell, C. Tu, S. R. Paranawithana, S. M. Tanhauser, P. V. LoGrasso, P. J. Laipis, D. N. Silverman, *Biochemistry* **1991**, *30*, 1484–1490.
- [31] A. Looney, R. Han, K. McNeill, G. Parkin, *J. Am. Chem. Soc.* **1993**, *115*, 4690–4697.
- [32] K. Nakata, M. K. Uddin, K. Ogawa, K. Ichikawa, *Chem. Lett.* **1997**, 991–992.
- [33] S. Yamaguchi, I. Tokairin, Y. Wakita, Y. Funahashi, K. Jitsukawa, H. Masuda, *Chem. Lett.* **2003**, *32*, 406–407.
- [34] J. F. Krebs, J. A. Ippolito, D. W. Christianson, C. A. Fierke, *J. Biol. Chem.* **1993**, *268*, 27458–27466.
- [35] L. Koziol, C. A. Valdez, S. E. Baker, E. Y. Lau, W. C. Floyd III, S. E. Wong, J. H. Satcher, Jr., F. C. Lightstone, R. D. Aines, *Inorg. Chem.* **2012**, *51*, 6803–6812.