Enzymatic C–S Coupling

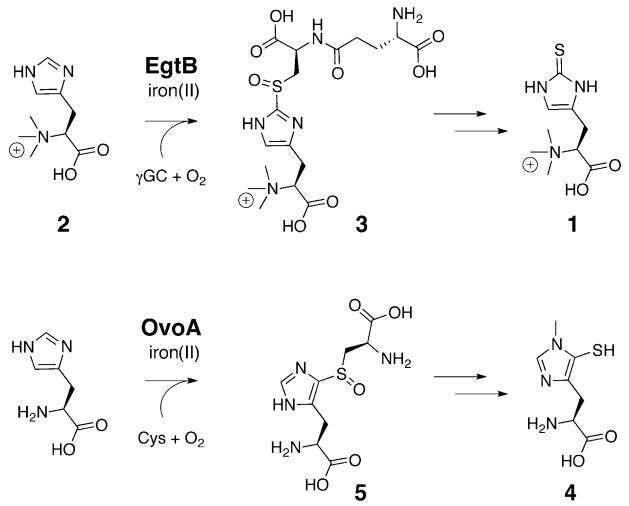
DOI: 10.1002/anie.201410045

Structure of the Sulfoxide Synthase EgtB from the Ergothioneine Biosynthetic Pathway**

Kristina V. Goncharenko, Allegra Vit, Wulf Blankenfeldt, and Florian P. Seebeck**

Abstract: *The non-heme iron enzyme EgtB catalyzes O_2-dependent C–S bond formation between γ-glutamyl cysteine and N-α-trimethyl histidine as the central step in ergothioneine biosynthesis. Both, the catalytic activity and the architecture of EgtB are distinct from known sulfur transferases or thiol dioxygenases. The crystal structure of EgtB from Mycobacterium thermoresistibile in complex with γ-glutamyl cysteine and N-α-trimethyl histidine reveals that the two substrates and three histidine residues serve as ligands in an octahedral iron binding site. This active site geometry is consistent with a catalytic mechanism in which C–S bond formation is initiated by an iron(III)-complexed thiyl radical attacking the imidazole ring of N-α-trimethyl histidine.*

Ergothioneine (**1**, Scheme 1) occurs in a broad range of prokaryotic and eukaryotic organisms, including humans and human pathogens such as *Mycobacterium tuberculosis*.[1] Higher eukaryotes absorb ergothioneine as a micronutrient, which stems from microbial production. The precise cellular function of ergothioneine is not known, but recent observations from animal, fungal, and bacterial cells suggest that this sulfur compound may be a protectant against oxidative stress.[2] *Mycobacteria* biosynthesize ergothioneine from glutamate, cysteine, and histidine.[1d,3] The central step in this pathway is catalyzed by the non-heme iron enzyme EgtB, which forms and sulfoxidizes a C–S bond between *N*-α-trimethyl histidine (**2**, TMH) and γ-glutamyl cysteine (γGC) with O_2 as the oxidant. Together with the ovothiol (**4**) biosynthesis enzyme OvoA (Scheme 1),[4] EgtB represents a distinct enzyme class (sulfoxide synthases) with no relation to sulfur oxidizing or C–S bond-forming iron enzymes such as cysteine dioxygenase or isopenicillin synthase.[5] Instead, sulfoxide synthases present a new entry to a rich collection of C–S bond-forming catalysts.[6]



Scheme 1. EgtB- and OvoA-catalyzed C–S bond formation and sulfoxidation between γ-glutamyl cysteine and N-α-trimethyl histidine (TMH) or between cysteine and histidine as the central steps in the syntheses of ergothioneine (**1**) and ovothiol (**4**), respectively.

To elucidate the structural basis for sulfoxide synthase activity, we determined the crystal structure of EgtB from *Mycobacterium thermoresistibile* ($EgtB_{thermo}$) in complex with iron and TMH, and as a quaternary complex with manganese, N-α-dimethyl histidine (DMH), and γGC. We produced this enzyme in *Escherichia coli* by following the protocols established for the EgtB homologue from *Mycobacterium smegmatis* ($EgtB_{smegmatis}$).[1d] The two enzymes share 81% sequence homology, both are monomeric (Figure S1 in the Supporting Information), and they display similar in vitro activities (Table 1). Both recombinant enzymes purified to a significant extent as iron-bound holoenzymes, as inferred from a ferrozine-based colorimetric assay ($EgtB_{thermo} > 95\%$, $EgtB_{smegmatis} > 50\%$; Table S1 in the Supporting Information) and from titration of EgtB activity with $FeSO_4$ (Figure S2). In vitro activity was assayed in HEPES-buffered solutions in the presence of TMH, γGC, $FeSO_4$ (4 μM), ascorbate (2 mM), NaCl (100 mM), and TCEP (2 mM) at 25 °C. Formation of the sulfoxide product (**3**; Scheme 1 and Figure S3) was monitored by cation-exchange HPLC.[1d] Under these conditions, $EgtB_{smegmatis}$ and $EgtB_{thermo}$ catalyzed up to one turnover per second and remained active for hundreds of turnovers (Table 1 and Figure S3). In the absence of ascorbate, the activity of $EgtB_{thermo}$ ceased after 120 ± 20 turnovers, but fully recovered after the addition of 2 mM ascorbate (Figure S4).

[*] K. V. Goncharenko,[+] F. P. Seebeck
Department for Chemistry, University of Basel
St. Johanns-Ring 19, 4056 Basel (Switzerland)
E-mail: florian.seebeck@unibas.ch
A. Vit,[+] W. Blankenfeldt
Structure and Function of Proteins
Helmholtz Centre for Infection Research
Inhoffenstr. 7, 38124 Braunschweig (Germany)
E-mail: wulf.blankenfeldt@helmholtz-hzi.de
[+] These authors contributed equally to this work.

[**] We would like to thank the Swiss Light Source (PSI, Villingen) and BESSY II Synchrotron (Berlin (Germany)) for beamline access and the European research council and the HZI Graduate School for Infection Research for financial support. F.P.S. is supported by the "Professur für Molekulare Bionik" and by an ERC starting grant.

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201410045.

Table 1: Kinetic parameters for the EgtB variants.[a]

Enzyme	Donor	Acceptor	$k_{cat,donor}$ [s^{-1}][b]	$k_{cat,donor}/K_M$ [$M^{-1}s^{-1}$]	$k_{cat,acceptor}$ [s^{-1}]	$k_{cat,acceptor}/K_M$ [$M^{-1}s^{-1}$]
$EgtB_{smegmatis}$	γGC	TMH	1.1	1.4×10^4	1.2	2.8×10^4
$EgtB_{thermo}$	γGC	TMH	8.6×10^{-1}	2.0×10^4	8.7×10^{-1}	2.2×10^4
$EgtB_{thermo}$	NGC	TMH	2.5×10^{-1}	2.3×10^2	n.a.	n.a.
$EgtB_{D416N}$	γGC	TMH	2.7×10^{-1}	1.3×10^2	1.1×10^{-1}	1.9×10^4
$EgtB_{D416N}$	NGC	TMH	1.0×10^{-1}	1.1×10^3	n.a.	n.a.

[a] γGC = gamma-glutamyl cysteine; TMH = N-α-trimethyl histidine; NGC = N-glutaryl cysteine; displayed values are averages from three independent measurements, errors correspond to less than 15% of the average value. [b] k_{cat} and k_{cat}/K_M were determined in the presence of cosubstrate at a concentration at least 3-fold higher than the corresponding K_M and in air-saturated buffers. Michaelis–Menten plots are displayed in the Supporting Information.

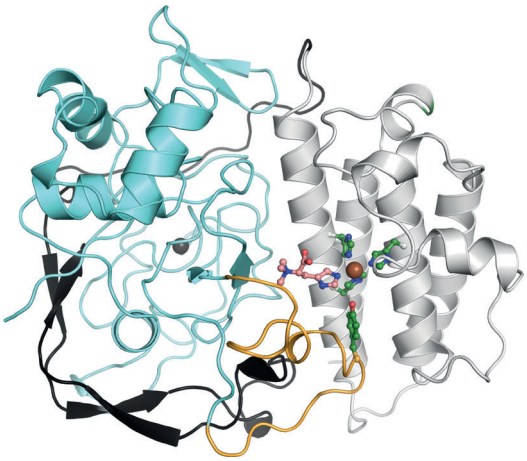


Figure 1. Cartoon diagram of $EgtB_{thermo}$ in complex with iron (brown) and TMH (pink). The protein consists of an N-terminal DinB domain (grey, residues 1–150), a two-stranded β-sheet region (black, residues 151–210), and a C-terminal C-type lectin domain (cyan/orange, residues 211–434). The active site, which contains a three-histidine facial triade (green; His51, His134, and His138) is formed between the DinB domain and residues 370–425 (orange)

Incubation with γGC alone also inactivated $EgtB_{thermo}$ (Figure S4) and induced an absorption band at 565 nm ($\varepsilon_{565} = 450\,M^{-1}cm^{-1}$; Figure S5), which also vanished after the addition of ascorbate. A similar absorption feature has been observed in cysteine-bound ferric cysteine dioxygenase, where it was assigned as a sulfur-to-iron(III) charge-transfer transition.[7] By contrast, no change in activity or absorbance was induced by incubation of $EgtB_{thermo}$ with TMH or buffer alone. We suspect that a γGC-dependent side reaction oxidizes EgtB to the inactive ferric state. The absorption at 565 nm of inactive $EgtB_{thermo}$ in complex with γGC provides a first indication that the sulfur atom of this ligand may directly interact with the catalytic iron center. The following structural analysis provides strong support for this notion.

We crystallized $EgtB_{thermo}$ as the apo protein (mthEgtB_apo), in complex with iron and TMH (mthEgtB_TMH), and in complex with manganese, DMH, and γGC (mthEgtB_DMH_γGC). The crystals belong to space group $P4_32_12$ with cell constants $a,b = 135$ and $c = 141$ Å. The apo structure was phased using the anomalous signal of the intrinsic sulfur atoms and the coordinated iron. The phases were later combined with native data to yield the structure of $EgtB_{thermo}$ at a resolution of 1.7 Å (apo). Ligand complexes of $EgtB_{thermo}$ were obtained to resolutions of 1.6 Å (TMH) and 1.98 Å (γGC and DMH). In all of the structures, the asymmetric unit includes two protein chains sharing only a relatively small and asymmetric interface.

The electron densities of the three $EgtB_{thermo}$ structures reveal a continuous polypeptide chain from Pro7 to Asp434 (Figure 1). Residues 7–150 fold to a DinB-like four-helix bundle with long linkers between helices 1 and 2 (18 residues), 2 and 3 (34 residues), and 3 and 4 (7 residues). The fourth helix is followed by an extended two-stranded β-sheet (residues 151–210) wrapped around the C-terminal domain, which adopts a C-type lectin fold (CLec).[8,9] This fold contains few secondary structure elements but instead is stabilized by a dense array of buried ionic interactions, such as salt bridges between Arg and Glu residues (Arg409:Glu196, Arg413:Glu296, Arg397:Glu300, and Arg428:Glu360). Furthermore, a calcium cation in the center of the C-terminal domain immobilizes six oxygen ligands from side chains and backbone amides (Met354 2.8 Å; Gly399 2.9 Å; Val358 2.6 Å; Gly356 2.7 Å; Gln353 3.6 Å; and Glu360 3.3 Å). This unusual loop-rich structure is conserved in at least two single-domain proteins with entirely different functions, with less than 30% sequence homology, and without dependence on transition metals.[8b,9] One of these structural homologues (r.m.s.d. of 0.741, Figure S6) is the human formylglycine-generating enzyme (FGE), which catalyzes the O_2-dependent posttranslational maturation of sulfatases.[8b] The second homologue is the diversity-generating retroelement variable protein TvpA from *Treponema denticola*.[9] Iron-dependent sulfoxide synthases such as EgtB and OvoA must have emerged from these FGE-like proteins through fusion with an N-terminal DinB domain. It is thus not surprising that the EgtB active site maps to the interface between the two domains.

The active site of $EgtB_{thermo}$ is housed in a 15 Å deep and 10 Å wide tunnel lined by residues 375 through 425 from the CLec domain and residues from helices 2 and 4 and the loops between helices 1, 2, and 3. At the bottom of this tunnel, three histidine residues from the DinB domain (residues 51, 134, and 138) coordinate the catalytic iron cation (Fe–N: 2.1 Å, 2.1 Å, and 2.3 Å). This structure revises our earlier prediction that Glu140 rather than His51 is the metal ligand.[1d,4a]

In hindsight, we note that this facial triad is fairly conserved among proteins from the DinB protein superfamily.[10] Several single-domain DinB proteins have been characterized as zinc-dependent thiol S-transferases or maleylpyruvate isomerases,[11] thus indicating that metal-catalyzed C–S bond formation may be a common activity among DinB-like proteins.[11a] However, the mechanisms of iron- and O_2-dependent oxidative C–S bond formation as catalyzed by EgtB and OvoA are most likely very different from that of zinc-promoted alkyl transfer.[6a,12] With this hypothesis in mind, we proceeded to study the substrate binding mode of $EgtB_{thermo}$.

In the ternary complex containing $EgtB_{thermo}$, TMH, and iron (ternary complex), the substrate imidazole ring (Fe–Nτ:

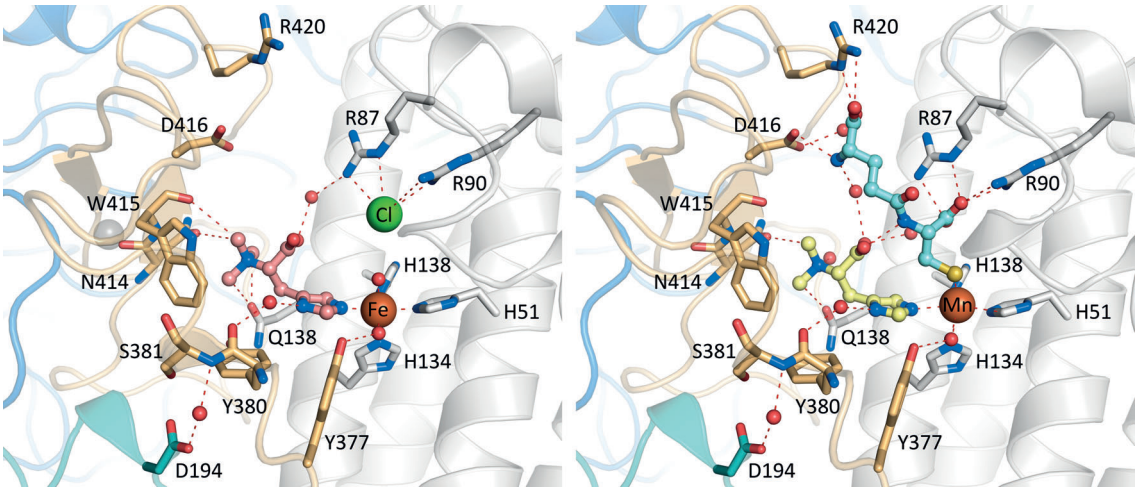


Figure 2. The active site of $EgtB_{thermo}$ in the ternary complex with TMH and iron (left), or with DMH, γGC, and manganese (right). Localized water molecules are shown in red. A chloride ion (green) in the ternary complex occupies the cationic docking site of γGC. Unbiased difference electron densities of bound ligands are shown in Figure S7.

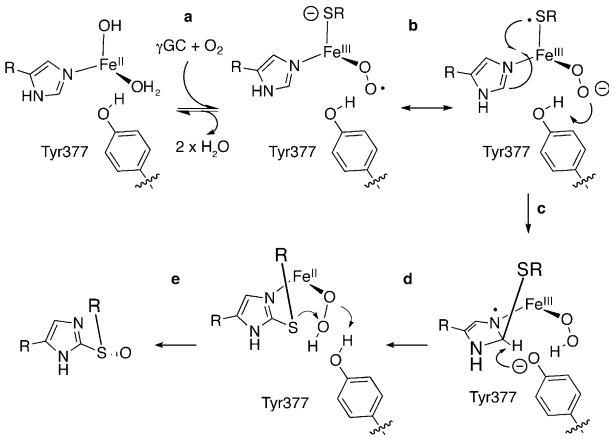
2.2 Å; Figure 2) and two water molecules (Fe–O: 2.1 and 2.2 Å) join the three-histidine facial triad in an octahedral coordination sphere around the iron center. The substrate imidazole ring also interacts with Tyr380 through a water-mediated hydrogen bond (5.4 Å), and the 1-carboxylate group is loosely connected to Arg87 (4.9 Å), again via a bridging water molecule. Two N-α-methyl groups of TMH pack against the indole side chain of Trp415. The third N-α-methyl group appears to make dipolar contacts to the amide side chains of Gln137 (3.2 Å) and Asn414 (3.5 Å).

To identify the binding mode of the second substrate γGC, we examined the structure of the quaternary complex between $EgtB_{thermo}$, DMH, γGC, and manganese(II). Initial attempts to soak the iron-containing binary complex with γGC resulted in disintegration of the crystals. As a solution to this problem, we used manganese-reconstituted $EgtB_{thermo}$. This enzyme complex is inactive[1d] but the corresponding crystals were tolerant to soaking with DMH and γGC and diffracted to a resolution of 1.98 Å (Figure 2 and Tables S2, S3). In this complex, the active-site residues and DMH adopt superimposable positions with respect to the ternary complex (r.m.s.d. = 0.041). γGC coordinates as the fifth ligand to the metal center (Mn–S: 2.6 Å). This direct metal–thiolate contact is consistent with the observed sulfur-to-iron charge-transfer transition absorption band in the γGC complex with ferric $EgtB_{thermo}$ (Figure S5). The α-amino group and the two carboxylates of γGC form salt bridges to Asp416 (2.8 Å), Arg420 (2.7 Å), Arg87 (3.0 Å), and Arg90 (2.6 Å). The amide function of γGC hydrogen bonds with the 1-carboxylate of DMH (2.8 Å).

To test whether the γGC-binding mode in the manganese complex properly reflects substrate binding by the iron-containing and therefore active form of $EgtB_{thermo}$, we probed γGC binding by site-directed mutagenesis. Specifically, we produced an $EgtB_{thermo}$ variant in which Asp416 is exchanged for Asn ($EgtB_{D416N}$). This mutation increases $K_{M,\gamma GC}$ by 200-fold but does not significantly change $K_{M,TMH}$ or k_{cat}. In a complementary experiment, we assayed wild-type $EgtB_{thermo}$ and $EgtB_{D416N}$ with a γGC derivative that lacks the α-amino function (N-glutaryl cysteine, NGC). NGC is a 100-fold less efficient sulfur donor for wild type $EgtB_{thermo}$ but a 10-fold better substrate for $EgtB_{D416N}$ than γGC. Evidently, the salt bridge between the α-amino group of γGC and Asp416 found in the manganese-containing quaternary complex of $EgtB_{thermo}$ is also important for substrate recognition during catalysis.

The sixth metal ligand in the quaternary complex is a water molecule (Mn–O: 2.5 Å), which also hydrogen bonds to the phenolic side chain of Tyr377 (2.8 Å). Additionally, this coordination site maps to the mouth of a narrow tunnel connecting the active site with the protein exterior (Figure S8). It is tempting to view this tunnel as the path through which O_2 accesses the active site.

Based on the observation that γGC, TMH, and possibly O_2 are direct ligands to the catalytic iron center, we propose the following mechanism for EgtB-catalyzed sulfoxide synthesis (Scheme 2): Binding of γGC and O_2 to the EgtB



Scheme 2. Proposed mechanism for EgtB-catalyzed C–S bond formation and sulfoxidation.

complex with TMH results in an iron(III) superoxo species (step a). This species may have partial iron(III) character and forms a complex with a peroxide anion (O_2^{-2}) and a thiyl radical (step b). The peroxide anion is stabilized by proton transfer from Tyr377 and the thiyl radical attacks the imidazole ring to form an iminyl radical (step c), which rearomatizes through deprotonation and ligand-to-metal electron transfer (step d). Sulfoxidation of the new thioether reduces the iron(II) peroxo species to return EgtB to its resting state (step e).

EgtB represents a novel type of non-heme iron enzyme, which catalyzes oxidative C–S bond formation and sulfoxidation. The crystal structure of EgtB provides the basis for testing mechanistic proposals by kinetic, spectroscopic, and computational methods[13] and also opens the door for protein engineering to generate tailor-made sulfur transferases. Our demonstration that mutation of a single residue (D416) suffices to change the sulfur donor specificity of $EgtB_{thermo}$ from γGC to NGC by almost three orders of magnitude bodes well for this objective. We anticipate that the structure of EgtB will also stimulate mechanistic investigations into the ovothiol biosynthetic sulfoxide synthase OvoA.[4]

Received: October 13, 2014
Published online: January 16, 2015

Keywords: enzyme catalysis · ergothioneine · non-heme iron enzymes · ovothiol · sulfur transfer

[1] a) D. B. Melville, W. H. Horner, C. C. Otken, M. L. Ludwig, *J. Biol. Chem.* **1955**, *213*, 61–68; b) D. B. Melville, D. S. Genghof, E. Inamine, V. Kovalenko, *J. Biol. Chem.* **1956**, *223*, 9–17; c) I. K. Cheah, B. Halliwell, *Biochim. Biophys. Acta* **2012**, *1822*, 784–793; d) F. P. Seebeck, *J. Am. Chem. Soc.* **2010**, *132*, 6632–6633.

[2] a) B. D. Paul, S. H. Snyder, *Cell Death Differ.* **2009**, *47*, S149–S149; b) T. Pluskal, M. Ueno, M. Yanagida, *PLoS One* **2014**, *9*, e97774; c) I. K. Cheah, R. L. Ong, J. Gruber, T. S. Yew, L. F. Ng, C. B. Chen, B. Halliwell, *Free Radical Res.* **2014**, *47*, 1036–1045; d) M. H. Bello, V. Barrera-Perez, D. Morin, L. Epstein, *Fungal Genet. Biol.* **2012**, *49*, 160–172; e) S. C. Emani, M. J. Williams, I. J. Wiid, N. F. Hiten, A. J. Viljoen, R. D. Pietersen, P. D. van Helden, B. Baker, *Antimicrob. Agents Chemother.* **2013**, *57*, 3202–3207.

[3] A. Vit, L. E. Misson, W. Blankenfeldt, F. P. Seebeck, *Acta Crystallogr. Sect. F* **2014**, *70*, 676–680.

[4] a) A. Braunshausen, F. P. Seebeck, *J. Am. Chem. Soc.* **2011**, *133*, 1757–1759; b) H. Song, A. S. Her, F. Raso, Z. Zheng, Y. Huo, P. Liu, *Org. Lett.* **2014**, *16*, 2122–2125; c) H. Song, M. Leninger, N. Lee, P. Liu, *Org. Lett.* **2013**, *15*, 4854–4857; d) G. T. Mashabela, F. P. Seebeck, *Chem. Commun.* **2013**, *49*, 7714–7716.

[5] a) C. A. Joseph, M. J. Maroney, *Chem. Commun.* **2007**, 3338–3349; b) P. L. Roach, I. J. Clifton, V. Fulop, K. Harlos, G. J. Barton, J. Hajdu, I. Andersson, C. J. Schofield, J. E. Baldwin, *Nature* **1995**, *375*, 700–704.

[6] a) B. Li, J. P. Yu, J. S. Brunzelle, G. N. Moll, W. A. van der Donk, S. K. Nair, *Science* **2006**, *311*, 1464–1467; b) E. Sasaki, X. Zhang, H. G. Sun, M. Y. Lu, M. Y. Liu, T. L. Liu, A. Ou, J. Y. Li, Y. H. Chen, S. E. Ealick, H. W. Liu, *Nature* **2014**, *510*, 427–431; c) A. Chatterjee, N. D. Abeydeera, S. Bale, P.-J. Pai, P. C. Dorrestein, D. H. Russell, S. E. Ealick, T. P. Begley, *Nature* **2011**, *478*, 542–U146; d) F. Berkovitch, Y. Nicolet, J. T. Wan, J. T. Jarrett, C. L. Drennan, *Science* **2004**, *303*, 76–79; e) P. L. Roach, I. J. Clifton, C. M. Hensgens, N. Shibata, C. J. Schofield, J. Hajdu, J. E. Baldwin, *Nature* **1997**, *387*, 827–830; f) D. H. Scharft, P. Chankhamjon, K. Scherlach, T. Heinekamp, K. Willing, A. A. Brakhage, C. Hertweck, *Angew. Chem. Int. Ed.* **2013**, *52*, 11092–11095; *Angew. Chem.* **2013**, *125*, 11298–11301.

[7] J. D. Gardner, B. S. Pierce, B. G. Fox, T. C. Brunold, *Biochemistry* **2010**, *49*, 6033–6041.

[8] a) S. A. McMahon, J. L. Miller, J. A. Lawton, D. E. Kerkow, A. Hodes, M. A. Marti-Renom, S. Doulatov, E. Narayanan, A. Sali, J. F. Miller, P. Ghosh, *Nat. Struct. Mol. Biol.* **2005**, *12*, 886–892; b) T. Dierks, A. Dickmanns, A. Preusser-Kunze, B. Schmidt, M. Mariappan, K. von Figura, R. Ficner, M. G. Rudolph, *Cell* **2005**, *121*, 541–552.

[9] J. Le Coq, P. Ghosh, *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 14649–14653.

[10] a) K. Nagata, J. Ohtsuka, M. Takahashi, A. Asano, H. IIino, A. Ebihara, M. Tanokura, *Proteins Struct. Funct. Bioinf.* **2008**, *70*, 1103–1107; b) S. S. Rajan, X. Yang, L. Shuzvalova, F. Collart, W. F. Anderson, *Biochemistry* **2004**, *43*, 15472–15479; c) D. R. Cooper, K. Grelewska, C.-Y. Kim, A. Joachimiak, Z. S. Derewenda, *Acta Crystallogr. Sect. F* **2010**, *66*, 219–224.

[11] a) G. L. Newton, S. S. Leuung, J. L. Wakabayashi, M. Rawat, R. C. Fahey, *Biochemistry* **2011**, *50*, 10751–10760; b) J. Feng, Y. Che, J. Milse, L. Yin, C. Rückert, X. H. Shen, S. W. Qi, J. Kalinowski, S. J. Liu, *J. Biol. Chem.* **2006**, *281*, 10778–10785; c) T. T. Liu, N. Y. Zhou, *J. Bacteriol.* **2012**, *194*, 3987–3994.

[12] J. Penner-Hahn, *Curr. Opin. Chem. Biol.* **2007**, *11*, 166–171.

[13] E. A. Bushnell, G. B. Fortowsky, J. W. Gauld, *Inorg. Chem.* **2012**, *51*, 13351–13356.

