Structural insights into the incorporation of the Mo cofactor into sulfite oxidase from site directed spin labeling

A. Hahn[1], C. Engelhard[2], S. Reschke[3], C. Teutloff[2,4], R. Bittl[2,4], S. Leimkühler[3], T. Risse[1,4,*]

[1]Institut für Chemie und Biochemie, Freie Universität Berlin, Takustr. 3, 14195 Berlin, Germany

[2]Fachbereich Physik, Freie Universität Berlin, Arnimallee 14,14195 Berlin, Germany

[3]Institut für Biochemie und Biologie, Universität Potsdam, Karl-Liebknecht-Str. 24-25, 14476 Golm, Germany

[4]Berlin Joint EPR Laboratory Freie Universität Berlin

Abstract:

Mononuclear *molybdoe*nzymes catalyze a broad range of redox reactions and are highly conserved in all kingdoms of life. This contribution addresses the question of how the Mo-cofactor (Moco) is incorporated into the apo-form of human sulfite oxidase (hSO) using site directed spin labeling to determine intramolecular distances in the nanometer range. Comparative measurements of the holo and apo-form of hSO allow for the first time to locate the corresponding structural changes, which are localized to a short loop (residues 263-273) of the Moco-containing domain. A flap-like movement of the loop provides access to the Moco binding-pocket in the apo-form of the protein and allows to understand the earlier studies on the *in vitro* reconstitution of apo-hSO with Moco. Remarkably, the loop motif can be found in a variety of structurally similar molybdoenzymes among various organisms suggesting a common mechanism of Moco incorporation.

Molybdoenzymes play essential roles in many aspects of cell metabolism[1]. These molybdoenzymes are broadly conserved throughout evolution and share a so-called molybdenum co-factor (Moco). Human SO (hSO), being localized in the intermembrane space of mitochondria, catalyzes the oxidation of sulfite to sulfate, the final step in the oxidative degradation of sulfur-containing amino acids, at its redox-active molybdenum cofactor (Moco). Sulfite as a strong nucleophile is able to cause severe damage to a multitude of cellular components including membrane components and sulfolipids[1a]. In the absence of SO sulfite accumulates in the body and especially causes severe damages to the brain. The corresponding diseases, isolated sulfite oxidase deficiency[2] or Moco deficiency[3], are characterized by rapidly progressing neurodegeneration[4] and death in early childhood.

In the past, molybdoenzymes were intensively studied to understand their enzymology including the biosynthesis[5], the cellular distribution[6] and the redox chemistry of the Moco[7]. However, elementary questions such as the molecular basis of the insertion mechanism of Moco into hSO remain open. Since hSO is localized in the intermembrane space of mitochondria, the mechanism of hSO translocation, assembly and cofactor insertion is a complex, multi-step maturation process, which has been proposed to be highly coordinated[8]. Due to its instability it is considered to be unlikely, that Moco exists in a free protein-unbound form in the cell. However, Moco-binding chaperones for enzymes of the sulfite oxidase family have not been identified so far. Thus, several unexplored hypotheses exist for the insertion of Moco into hSO[1a, 9]. Biochemical studies of SO and other molybdoenzymes suggest Moco to be the last cofactor inserted into the already assembled apo-enzyme[10], an observation that allows to address the structural details of the insertion step by comparing the apo- and the holo-forms of the protein.

This study aims at elucidating the structural basis of the Moco insertion mechanism into hSO using site directed spin labeling (SDSL). Human SO is a homodimer with a molecular mass of about 110 kDa and each monomer of the full-length protein consists of three domains referred to as the Moco, the heme-containing and the dimerization domain. The structure of hSO has not been determined yet, however, a crystal structure was determined for the highly homologous chicken liver SO (cSO)[11] (pdb: 1SOX, 67 % identical residues)[12]. Based on the structure of cSO suitable labeling sites to probe structural changes within the Moco domain were selected and distance distributions between pairs of site-specifically introduced spin labels of the respective holo- and apo-protein samples were determined using pulsed electron electron double resonance spectroscopy (PELDOR or DEER) to search for structural changes related to the presence or absence of Moco[13]. The results reveal a highly localized change involving a short loop (residues 263-273 in hSO) in the Moco domain. This loop provides access to the Moco binding pocket from the outer surface in the apo-form, while it blocks access in the holo-form. The conservation of the loop structure as well as its binding motif to the Moco in a variety of structurally related molybdoenzymes (e.g. s. Figure S12) points to a conserved Moco insertion mechanism.

SDSL is typically based on the labeling of cysteine residues, which requires a cys-less construct to avoid interference with labeling at native cysteine residues. The cys-less construct used within this study lacked the N-terminal heme domain of hSO and is referred to as hSOMD (hSO Moco domain) throughout this study. It has previously been shown that the activity of hSOMD is retained after the exchange of the four non-essential cysteine residues[14]. Additionally, 70 -90 % of wild type sulfite oxidation activity was obtained for all hSOMD variants containing the site-

specifically introduced cysteine residues used in this study (at positions 195, 235, 302, 217, 267, 341, and 353) as long as the essential cysteine residue 207 was retained. Mutation of C207 to serine results in an inactive enzyme [15] due to a loss of the cysteine-sulfur-molybdenum bond. Despite the lack of catalytic activity, all holo-form hSOMD variants lacking the essential residue C207 exhibited a molybdenum loading of about 80 %, which corresponds to the molybdenum incorporation level of the wild type protein as checked by inductively coupled plasma optical emission spectrometry (ICP-OES).

To verify the structural comparability of the cysteine-free hSOMD construct with the full-length wild type protein (hSO) three sites located on different secondary structure elements, namely N195[1] (α-helix), R235 (β-sheet), and R302 (loop) located around the Moco binding pocket were selected based on the cSO crystal structure (Figure 1). For the investigation of full-length hSO by SDSL, pairs of the non-natural amino acid p-acetylphenylalanine, namely N195-R235pAcPhe, N195-R302pAcPhe and R235-R302pAcPhe were introduced into the wild type enzyme. These variants were labeled using specific, catalytic ketoxime coupling chemistry[16]. The distance distributions obtained for hSO were compared to theoretical expectations based on the cSO crystal structure using the PyMOL script "MtsslWizard"[17], which calculates the conformational space of different spin labels considering a simple van-der-Waals model. For all measurements the resulting distance distributions as determined by using the software tool "DeerAnalysis 2013"[18] agreed well with expectations based on the crystal structure of cSO (Figure S2), a result that also holds for hSOMD variants containing the corresponding MTSSL labeled cysteine pairs

[1] The numbering scheme refers to hSO. Information on the sequence alignment as compared to cSO and its crystal structure may be found in the supplementary information.

only. Please note that the distance distribution reveals contributions from spin-spin interactions within one (intra) and between the two (inter) monomer units of the homodimer. The simulations allow to discriminate between them, and theoretical intra-monomer distances are colored in blue while the inter-monomer distances are shown in green. The inter monomer distances are usually too large to be probed by the DEER experiments, therefore the discussion of the distance distribution is restricted to the intra monomer distances. Hence, hSOMD lacking the essential cysteine residue 207 and the full-length protein are structurally indistinguishable from cSO in the examined region. Thus, the cysteine-free construct of hSOMD can be used as reference for the apo-protein. Please note that ketoxime coupling of spin labels to apo-protein

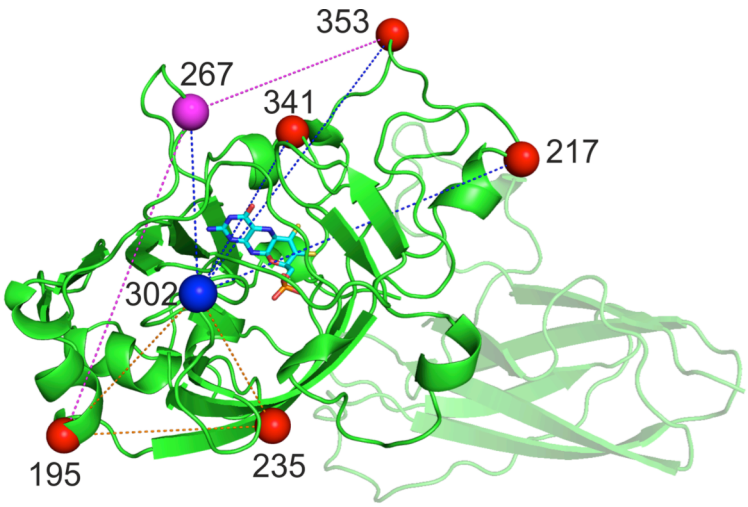

Figure 1: Residues selected for site directed spin labeling (Structure of cSO; pdb: 1SOX). Numbering corresponds to the respective positions in hSO. See SI for a sequence alignment.

variants containing unnatural amino acids was not possible, because of insufficient stability of the apo-protein at reaction conditions.

The measurements of the apo-forms of hSOMD variants N195C-R235C, N195-R302C and R235-R302C revealed no changes as expected based on the crystal structure of cSO (Figure S 3, the

distance distribution for all cysteine variants were calculated using both MtsslWizard as well as the Matlab toolbox "MMM2013.2" and both gave very similar results), hence, it is inferred that this region of the protein remains unperturbed in absence of the cofactor.

As a next step the remaining secondary structure elements of the Moco domain of hSOMD were

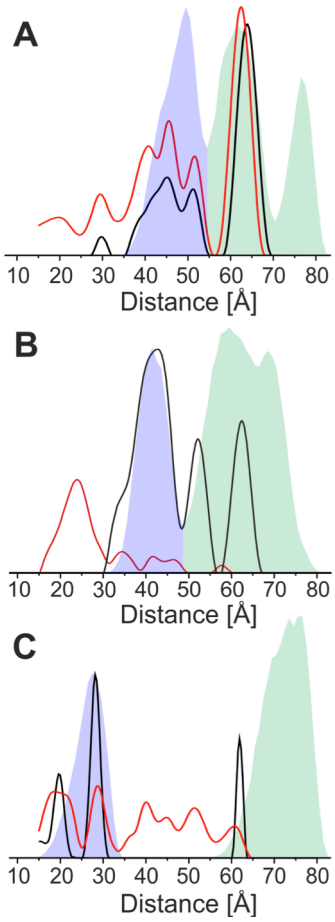

Figure 2: Distance distributions of holo- (black trace) and apo-protein (red trace) samples containing cysteine pairs (A) 302-353 and (B) 195-267 and (C) 267-302. Shaded areas are predictions based on the crystal structure of chicken SO (pdb: 1SOX) using the PyMOL script "MtsslWizard"; simulations using the MMM package can be found in Figure S8 of the SI. (blue: intra monomer distances; green: inter monomer contributions).

examined using a variety of MTSSL-labeled cysteine pairs containing residue 302, shown to be fixed, with sites 217, 267, 341 and 353 as well as two additional pairs of residue 267 with residues 195 and 353 (Figure 1), which exhibit predicted intra-monomer distances compatible

with DEER experiments (15- 60 Å[13]). For each pair, distance distributions of holo- and apo-protein were measured. For the traces with modulation depths approaching 0.5 (e.g. the holo form of variant 195-267 C, Figure S5) multiple spin interactions within the homodimer have to be considered. These may lead to artificial contributions ("ghost peaks") in the distance distribution. However, data analysis using the "power-scaling" approach[19] implemented in "DeerAnalysis 2013" gave no indication for "ghost-peaks". The measurements on pairs 302-217, 302-341 and 302-353 showed no significant differences between the holo and apo-protein samples as exemplified for the latter pair in Figure 2A (DEER traces s. Figure S4). Based on these results it can be concluded that the structure of the protein as probed by these residues is unchanged after insertion of the Moco. A different scenario is found for pairs containing residue 267. Comparing for example the distance distribution of apo- and holo-protein for the pair 195-267 shown in Figure 2B, a new peak at smaller distances together with some weight at distances between 30 and 50 Å is observed for the apo-form of the protein. The contribution at small distances is readily seen in the time traces (Figure S5), which show a considerably faster initial decay of the DEER trace for the apo-form, corresponding to shorter distances. While the interspin distance between 195-267 is reduced in the apo-form, increased distances are observed for 267-302 (Figure 2C). Together with the observation that the distance distribution is virtually unchanged for all other pairs investigated, it is concluded that the loop containing residue 267 undergoes a Moco dependent structural rearrangement. It should be noted that the distance distribution of pair 267-353 is broadened even for the holo-form of the protein as compared to expectations based on the crystal structure (Figure S7). This observation is in line with results from crystallography revealing a loss of electron density for the loop containing

residue 353 for a variant lacking the essential cysteine residue C207 pointing towards an increased mobility of this structural element (pdb: 3HBG)[20].

The structural constraints imposed by the measured distance distributions were used to describe the movement of the loop containing residue 267. DEER derived data were used to calculate the structural transition employing the "Elastic Network Model" (ENM) module of the Matlab toolbox "MMM2013.2"[21].

The distance distributions determined for the holo-protein fit well the expectation based on the cSO crystal structure (pdb: 1SOX). Hence, this structure was used as the reference point (without the heme domain) for the structural transition. The constraints (weighted mean value and width of the distribution, see Table 1) derived from the measurements of all cysteine pairs of the apo-protein mentioned above were used and resulted in a calculated structure of the apo-form shown in Figure 3A (orange) in comparison to the crystal structure of cSO (cyan). The r.m.s.d. value of the optimized structure is found to be 1.6 Å, which is a value considered to indicate an acceptable fit[21a].

The overlay of the two models reveals a large-scale movement of the loop-containing residue 267 (red spheres). A smaller movement is also observed for the loop including residue 353 (yellow spheres) while no significant differences are observed for the other residues (grey spheres) investigated here. The movement of the loop containing residue 353 is most likely due to structural differences originating from the C207S exchange as discussed above [20].

The analysis of the DEER experiments using the ENM approach on apo-protein samples demonstrate a flap-like movement of the short loop consisting of residues 263-273 in the absence of Moco, while the rest of the protein is virtually unperturbed. This is the first direct

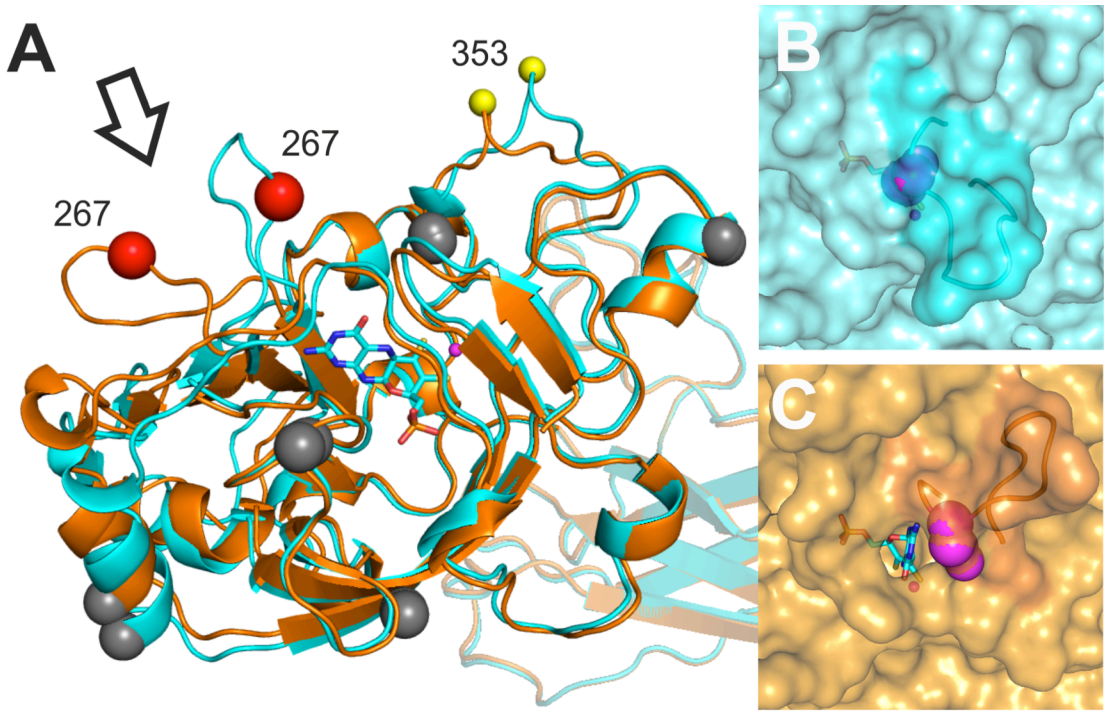

Figure 3: A: Structural model of the apo-form (orange) of hSOMD as calculated by the ENM approach (see text). Cyan: Holo-form of cSO (pdb: 1SOX) used as reference structure. The dimerization domain is shown transparent. B and C: Section of the Connolly surface containing the loop 263-273 for the hSO structures shown in A where the viewing direction is indicated by an arrow. B: Holo-protein structure; the surface of the loop region as well as the loop backbone are highlighted. Sidechain atoms of residue D265 are shown as blue/magenta spheres; Moco is shown as a stick model. C: Structure of the apo-protein; for better visualization of the access to the Moco binding pocket the Moco is shown as a stick model in the “apo”-form; the surface of the loop region as well as the loop backbone are highlighted. Sidechain of residue D265 are shown as orange/magenta spheres.

experimental evidence for a structural change between holo-hSO and apo-hSO and confirms a proposal by Hille et al. who suggested the movement of a secondary structure element as part of the Moco incorporation mechanism for SO from *Arabidopsis thaliana*[1b]. It may be noted in passing that this highly localized structural rearrangement is different from other well-studied protein-cofactor interactions, which typically result in extensive and long-range effects upon binding (or removal) of the cofactor [22]. Comparing the Connolly surface[23] of the holo- and the apo-form of the protein as shown in Figure 3b and c, respectively, shows that the clap-like movement in the apo-form provides direct access to the Moco binding pocket from the outer surface, while this access is blocked by the loop 263-273 in the holo-form. The conformation of the loop in the holo form is stabilized by two hydrogen bonds from the D265 side chain to the N1 ring-position and the 2-amino group of the Moco (Figure 3b)[11]. It is important to note that this analysis is not meant to provide information on the loop conformation itself. Its details depend on the specific input parameters of the simulation. However, all performed simulations show a movement of the loop such that the Moco binding pocket becomes accessible from the outer surface in the apo-protein (see SI for a more detailed discussion). This finding provides a possible structural explanation for the *in vitro* reconstitution of Moco into hSO in the absence of additional proteins [24]. However, the incorporation process *in vivo* may involve additional proteins to optimize the maturation process of the enzyme, similar to dedicated chaperones being identified for molybdoenzymes in prokaryotes [25].

Table 1: Distance constraints for the calculation of the apo protein structure by the ENM-approach derived from DEER distance distributions.

Labeled positions	Weighted mean distance [Å]	Width of distribution [Å]
195-235	25.7	17.2
195-302	26.9	17.1
235-302	25.8	22.2
217-302	40.9	15.0
302-341	36.3	19.7
267-302	34.0	41.3
302-353	39.0	37.0
195-267	27.0	34.3
267-353	28.8	31.2

As discussed above, residue D265 plays an important role in the Moco binding. It blocks the channel once the holo-protein is formed and keeps the Moco and the loop in a defined conformation. Comparison with other proteins from the sulfite oxidase family reveals remarkable similarities in this Moco binding motif: Analogous loop structures with similar coordination to the Moco mentioned above are e.g. found in sulfite oxidase from *Arabidopsis thaliana* (residues 160-171, pdb: 1OGP), sulfite dehydrogenase from *Starkeya novella* (residues 157-167, pdb: 2BPB) or *Escherichia coli* YedY (residues 43-52, pdb: 1XDQ) (structural models and additional examples can be found in the SI s. Figure S12). In light of the results found for hSO in this study, a similar mechanism of Moco incorporation involving the clap-like movement of the conserved loop structure for a range of vastly different organism can be expected.

In conclusion we could show that structural differences between the apo- and holo-forms of hSO are restricted to a small loop region of the protein containing residues 263-273. A clap-like motion of the loop provides a structural explanation for the observed ability to insert Moco into apo-sulfite oxidase *in vitro.* The conservation of this structural motif as well as the conserved interaction of analogous residues with the Moco throughout a range of mononuclear molybdoenzymes suggests that the structural changes observed for hSO and the proposed insertion mechanism apply to other homologouous proteins within this family as well.

Method section

Details of the methods used in this study can be found free of charge in the supplementary information.

Acknowledgement

The authors are grateful to the Center of Excellence 314 (UniCat) funded by the DFG for financial support. A.H. thanks the Studienstiftung des deutschen Volkes for a fellowship. We thank K.V. Rajagopalan (Duke University) for providing plasmids for the expression of hSO and hSOMD. The authors thank Ricarda Ebert for her help in the early stages of this project. We thank Gregor Hagelüken for adapting the PyMOL script "MtsslWizard" to the spin labeled version of the non-natural amino acid p-acetylphenylalanine.

[1] aG. Schwarz, R. R. Mendel, M. W. Ribbe, *Nature* **2009**, *460*, 839-847; bR. Hille, J. Hall, P. Basu, *Chem. Rev.* **2014**, *114*, 3963-4038; cG. Schwarz, R. R. Mendel, *Annu. Rev. Plant Biol.* **2006**, *57*, 623-647.

[2] S. H. Mudd, F. Irreverre, L. Laster, *Science* **1967**, *156*, 1599-1602.

[3] J. L. Johnson, W. R. Waud, K. V. Rajagopalan, M. Duran, F. A. Beemer, S. K. Wadman, *Proc Natl Acad Sci U S A* **1980**, *77*, 3715-3719.
[4] V. E. Shih, I. F. Abroms, J. L. Johnson, M. Carney, R. Mandell, R. M. Robb, J. P. Cloherty, K. V. Rajagopalan, *N. Engl. J. Med.* **1977**, *297*, 1022-1028.
[5] aJ. Kuper, A. Llamas, H. J. Hecht, R. R. Mendel, G. Schwarz, *Nature* **2004**, *430*, 803-806; bB. M. Hover, A. Loksztejn, A. A. Ribeiro, K. Yokoyama, *J. Am. Chem. Soc.* **2013**, *135*, 7019-7032.
[6] J. Teschner, N. Lachmann, J. Schulze, M. Geisler, K. Selbach, J. Santamaria-Araujo, J. Balk, R. R. Mendel, F. Bittner, *Plant Cell* **2010**, *22*, 468-480.
[7] aM. Kaupp, *Angew. Chem. Int. Ed. Engl.* **2004**, *43*, 546-549; bR. A. Rothery, B. Stein, M. Solomonson, M. L. Kirk, J. H. Weiner, *Proc Natl Acad Sci U S A* **2012**, *109*, 14773-14778.
[8] J. M. Klein, G. Schwarz, *J. Cell Sci.* **2012**, *125*, 4876-4885.
[9] T. Kruse, C. Gehl, M. Geisler, M. Lehrke, P. Ringel, S. Hallier, R. Hansch, R. R. Mendel, *J. Biol. Chem.* **2010**, *285*, 6623-6635.
[10] C. Iobbi-Nivol, S. Leimkühler, *Biochim. Biophys. Acta* **2013**, *1827*, 1086-1101.
[11] C. Kisker, H. Schindelin, A. Pacheco, W. A. Wehbi, R. M. Garrett, K. V. Rajagopalan, J. H. Enemark, D. C. Rees, *Cell* **1997**, *91*, 973-983.
[12] R. M. Garrett, D. B. Bellissimo, K. V. Rajagopalan, *Biochim. Biophys. Acta* **1995**, *1262*, 147-149.
[13] G. Jeschke, *Annu. Rev. Phys. Chem.* **2012**, *63*, 419-446.
[14] aC. A. Temple, T. N. Graf, K. V. Rajagopalan, *Arch. Biochem. Biophys.* **2000**, *383*, 281-287; bS. Reschke, D. Niks, H. Wilson, K. G. Sigfridsson, M. Haumann, K. V. Rajagopalan, R. Hille, S. Leimkuhler, *Biochemistry* **2013**, *52*, 8295-8303; cH. J. Cohen, S. Betcher-Lange, D. L. Kessler, K. V. Rajagopalan, *J. Biol. Chem.* **1972**, *247*, 7759-7766.
[15] aR. M. Garrett, K. V. Rajagopalan, *J. Biol. Chem.* **1996**, *271*, 7387-7391; bG. N. George, R. M. Garrett, R. C. Prince, K. V. Rajagopalan, *Inorg. Chem.* **2004**, *43*, 8456-8460.
[16] aM. R. Fleissner, E. M. Brustad, T. Kalai, C. Altenbach, D. Cascio, F. B. Peters, K. Hideg, S. Peuker, P. G. Schultz, W. L. Hubbell, *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 21637-21642; bA. Hahn, S. Reschke, S. Leimkuhler, T. Risse, *J. Phys. Chem. B* **2014**, *118*, 7077-7084; cC. C. Liu, P. G. Schultz, *Annu. Rev. Biochem.* **2010**, *79*, 413-444.
[17] G. Hagelueken, R. Ward, J. H. Naismith, O. Schiemann, *Appl Magn Reson* **2012**, *42*, 377-391.
[18] G. Jeschke, V. Chechik, P. Ionita, A. Godt, H. Zimmermann, J. Banham, C. R. Timmel, D. Hilger, H. Jung, *Applied Magnetic Resonance* **2006**, *30*, 473-498.
[19] T. von Hagens, Y. Polyhach, M. Sajid, A. Godt, G. Jeschke, *Phys. Chem. Chem. Phys.* **2013**, *15*, 5854-5866.
[20] J. A. Qiu, H. L. Wilson, M. J. Pushie, C. Kisker, G. N. George, K. V. Rajagopalan, *Biochemistry* **2010**, *49*, 3989-4000.
[21] aG. Jeschke, *Z. Phys. Chem.* **2012**, *226*, 1395-1414; bG. Jeschke, *J. Chem. Theory Comput.* **2012**, *8*, 3854-3863; cY. Polyhach, E. Bordignon, G. Jeschke, *Phys. Chem. Chem. Phys.* **2011**, *13*, 2356-2366; dY. Polyhach, G. Jeschke, *Spectroscopy* **2010**, *24*, 651-659.
[22] aM. Zhang, T. Tanaka, M. Ikura, *Nat. Struct. Biol.* **1995**, *2*, 758-767; bH. M. Berman, L. F. Ten Eyck, D. S. Goodsell, N. M. Haste, A. Kornev, S. S. Taylor, *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 45-50; cA. C. Newton, *Chem. Rev.* **2001**, *101*, 2353-2364.
[23] M. Connolly, *J. Appl. Crystallogr.* **1983**, *16*, 548-558.
[24] S. Leimkuhler, K. V. Rajagopalan, *J. Biol. Chem.* **2001**, *276*, 1837-1844.
[25] S. Leimkuhler, W. Klipp, *J. Bacteriol.* **1999**, *181*, 2745-2751.