

VIP Very Important Paper

Special  
Collection

# Lanthanide Tagging of Oligonucleotides to Nucleobase for Paramagnetic NMR

Sebastian Täubert<sup>+, [a]</sup> Yong-Hui Zhang<sup>+, [a]</sup> Mitchell Maestre Martinez,<sup>[a]</sup> Florian Siepel,<sup>[a]</sup> Edith Wöltjen,<sup>[a]</sup> Andrei Leonov,<sup>[a]</sup> and Christian Griesinger<sup>\*[a]</sup>

Although lanthanide tags, which have large anisotropic magnetic susceptibilities, have already been introduced to enrich NMR parameters by long-range pseudocontact shifts (PCSs) and residual dipolar couplings (RDCs) of proteins, their application to nucleotides has so far been limited to one previous report, due to the high affinities of lanthanides for the phosphodiester backbone of nucleotides and difficult organic synthesis. Herein, we report successful attachment of a lanthanide tag to a chemically synthesized oligonucleotide via a disulfide bond. NMR experiments reveal PCSs of up to 1 ppm and H–H RDCs of up to 8 Hz at 950 MHz. Although weaker magnetic alignment was achieved than with proteins, the paramagnetic data could be fitted to the known structure of the DNA, taking the mobility of the tag into account. While further rigidification of the tag is desirable, this tag could also be used to measure heteronuclear RDCs of <sup>13</sup>C, <sup>15</sup>N-labeled chemically synthesized DNA and RNA.

NMR spectroscopy has long been an indispensable tool for the study of the structure and dynamics of biomolecules with atomic resolution. In addition to the well-established NMR methods based on *J* coupling and nuclear Overhauser effect (NOE), paramagnetic NMR parameters such as paramagnetic relaxation enhancement (PRE), pseudocontact shift (PCS) or residual dipolar coupling (RDC) improve the accuracy of structures,<sup>[1]</sup> specifically structures of biomolecular complexes<sup>[2]</sup> and allow quantitative measurement of domain motion.<sup>[3–8]</sup> With the largest Curie spin and anisotropic magnetic susceptibilities, lanthanide ions are commonly used in paramagnetic NMR for proteins. Lanthanides can either be loaded into a metal-binding site of a metalloprotein,<sup>[9,10]</sup> or be site-specifically


attached to a biomolecule via a metal binding tag.<sup>[11–25]</sup> This has also been demonstrated for oligosaccharides,<sup>[26–32]</sup> but only in one case so far for oligonucleotides,<sup>[33]</sup> with attachment of the tag at a phosphorothioate group. Herein, we present a chemical approach developed for lanthanide tagging of oligonucleotides by using the modified nucleobase thymidine that allows to attach the Cys-Ph-TAHA tag<sup>[34]</sup> and the MesS-Ph-TAHA tag via a disulfide bond.


In a 24-mer hairpin oligomer (Figure 1) based on a STAT protein binding sequence, the thymine residue at position 3 was substituted by the modified nucleotide 1 which was shown to minimally perturb the DNA structure<sup>[35]</sup> (Table S8 in the Supporting Information). The presented synthesis, deprotection and purification procedure of the modified oligomer is conveniently similar to a conventional oligonucleotide protocol, yet it results in a thiol binding site for tagging. Here, we present a first synthesis strategy of MesS-Ph-TAHA tag based on the Cys-Ph-TAHA tag<sup>[34]</sup> (see the Supporting Information) with a total yield of 8%. Therein, the sulfur moiety needs to be introduced prior to the TAHA fragment which results in a 12-fold alkylation with a yield of only 20%. This crucial step was performed under mild conditions to avoid undesired side products that increased the reaction time up to 4 weeks. Due to the high affinity of lanthanide ions to the phosphate backbone, the Cys-Ph-TAHA/MesS-Ph-TAHA tags were preloaded with paramagnetic (Tm<sup>3+</sup>, Tb<sup>3+</sup>) or diamagnetic (Lu<sup>3+</sup>) lanthanide ions before performing the tagging reaction. In contrast to the protein-specific tagging reactions, the temperature was increased to 55 °C to ensure a complete turnover for tagging with Cys-Ph-TAHA, even this resulted in a dimerization of the oligonucleotide as a side product. The two products were, however, separable in the final HPLC purification step and the dimerized oligonucleotides were reactivated using the reducing agent TCEP. Tagging reactions using MesS-Ph-TAHA tag under the same conditions lead only to a partially tagged DNA strand. The more rigid and shorter MesS-Ph-TAHA tag gives rise to stronger electrostatic repulsion between the negatively charged TAHA fragment and the DNA backbone. To enable high quality NMR spectra, the tagged oligonucleotides had to be washed several times using a highly concentrated NaCl solution in order to remove traces of lanthanide ions from the backbone. In a final step, the tagged oligonucleotide was annealed to ensure the formation of the monomeric hairpin.


As is described in more detail in the Supporting Information, COSY (Figure 2) and NOESY spectra were acquired for chemical shift assignment (Tables S1–S7) and PCS measurement, HSQC (Figure S1) and E.COSY (Figure S2) spectra for RDC measure-

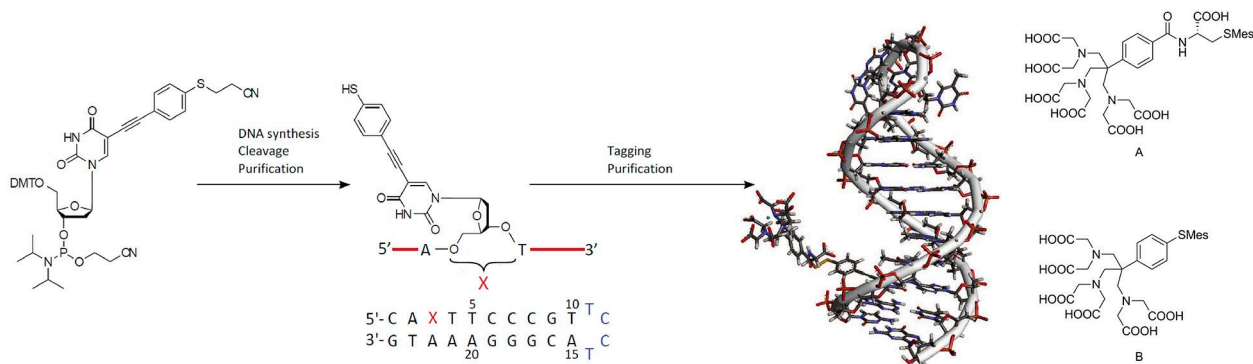
[a] S. Täubert,<sup>+</sup> Y.-H. Zhang,<sup>+</sup> M. M. Martinez, F. Siepel, E. Wöltjen, A. Leonov, C. Griesinger  
NMR Based Structural Biology  
Max Planck Institute for Biophysical Chemistry  
Am Fassberg 11, 37077 Göttingen (Germany)  
E-mail: cig@nmr.mpibpc.mpg.de

[<sup>+</sup>] This author is contributed equally to this work.

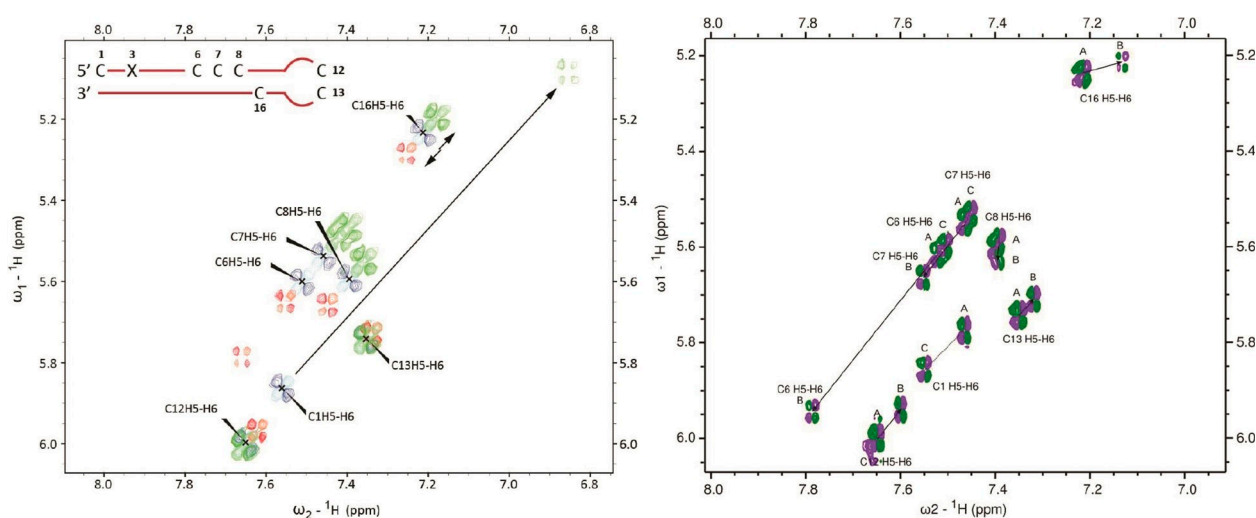
 Supporting information for this article is available on the WWW under <https://doi.org/10.1002/cbic.202000417>

 This article is part of a Special Collection on the occasion of Horst Kessler's 80th birthday. To view the complete collection, visit our homepage

 © 2020 The Authors. Published by Wiley-VCH GmbH. This is an open access article under the terms of the Creative Commons Attribution Non-Commercial NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.



**Figure 1.** Modified thymidine-phosphoramidite (X) is conveniently incorporated into the DNA hairpin, which is synthesized by standard methods to yield a hairpin structure with a modified nucleotide in position 3 (X). The DNA with the modified base (X) reacts with, for example, the mesylated Cys-Ph-TAHA tag (A, right) loaded with lanthanide ion to yield tagged and lanthanide-loaded DNA, which is purified from unreacted tag. Alternatively the MesS-Ph-TAHA tag (B) was used.



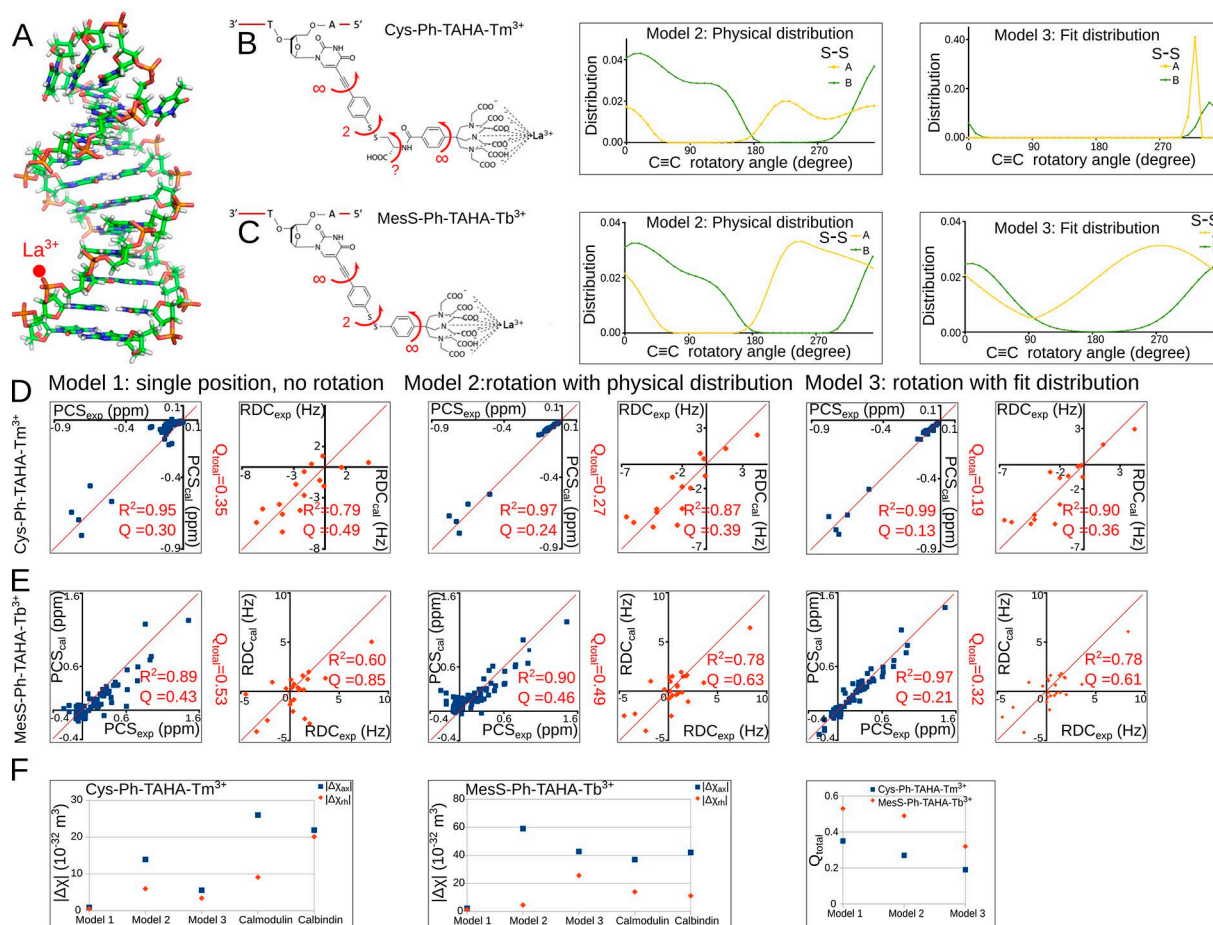
**Figure 2.** Left: overlay of the COSY of 24-mer DNA with Cys-Ph-TAHA tag, showing PCSs in opposite directions for  $\text{Tb}^{3+}$  (red) and  $\text{Tm}^{3+}$  (green), compared to  $\text{Lu}^{3+}$  (blue). Right: COSY of 24-mer DNA with MesS-Ph-TAHA tag and  $\text{Tb}^{3+}$  ions. Here three sets of peaks are observed; A and C are diamagnetic and B is paramagnetic.

ment. As a result, 67 PCSs and 15 H–C RDCs were obtained from 24-mer DNA with Cys-Ph-TAHA tag and  $\text{Tm}^{3+}$  ions (Tables S10 and S13); 121 PCSs and 25 H–H RDCs were obtained from 24-mer DNA with MesS-Ph-TAHA tag and  $\text{Tb}^{3+}$  ions (Tables S11, S12, S14, and S15). As introduced below, three models were explored for fitting these two data sets.

Both data sets were at first fit with a single position of the lanthanide and single magnetic susceptibility tensor, without any motion between the lanthanide and 24-mer DNA (model 1). Thus we had eight variables: three for the position of the lanthanide and five for the definition of the tensor. A B-DNA structure of 24-mer DNA was used (Table S16, Figure 3A, D and E), with the position and both the magnitude and orientation of magnetic susceptibility tensor of the lanthanide ion as fitting parameters. Due to the higher uncertainty of RDC data compared to the PCS, the quality factors of RDCs are larger than those of PCS (i.e.,  $Q_{\text{rdc}} > Q_{\text{pcs}}$ ). As is summarized in Figure 3F, the overall quality factors ( $Q_{\text{pcs+rdc}}$ ) are large and the

magnetic anisotropy tensors ( $\Delta\chi_{\text{ax}}$  and  $\Delta\chi_{\text{rh}}$ ) are far smaller than those in calcium binding proteins<sup>[34,35]</sup> as well as published lanthanide tags,<sup>[36,37]</sup> thus indicating that model 1 is unsatisfactory.

In order to improve the quality of data fitting and by that following work on other systems,<sup>[7,38]</sup> conformation ensembles taking structural flexibility into consideration (models 2 and 3) were constructed. In the chemical structures of Cys-Ph-TAHA tag (Figure 3B) and MesS-Ph-TAHA tag (Figure 3C), the carbon≡carbon triple bonds and phenyl groups allow for free rotation about the same axis, and disulfide bonds adopt two conformations with the S–S dihedral angle at  $+90^\circ$  and  $-90^\circ$  (see the Supporting movies). In addition, the peptide bond in the Cys-Ph-TAHA tag may introduce additional flexibility. In order to reduce the size of the conformation ensemble originating from structural flexibility, only the free rotation of the carbon≡carbon triple bond and the two conformations of the disulfide bond were considered in the data fitting. The free rotation of the



**Figure 3.** Fitting PCS and RDC data of 24-mer DNA with different lanthanide-tagging models. A) Single lanthanide ion located near the B-DNA structure of 24-mer DNA. B) and C) Left: chemical structures and flexibilities of Cys-Ph-TAHA and MesS-Ph-TAHA; free rotation (e.g., carbon≡carbon triple bond) is indicated as “∞”, two conformations (e.g., disulfide bond) as “2” and unknown flexibility as “?”. Center: the probability distribution of lanthanide-tagging conformations with different rotatory angles of the carbon≡carbon triple bond and double conformations of disulfide bond is calculated from electrostatic interaction. Right: probability distribution of the lanthanide-tagging conformation fit as two Gaussian functions from PCS and RDC data. D) and E) Comparison of experimental (*x*-axis) PCS and RDC data and those calculated (*y*-axis) from different models. F) Comparison of magnetic anisotropy tensors and quality factor. Center: the tensor determined from the MesS-Ph-TAHA-tagged DNA for model 3 matches very well the tensors for calmodulin<sup>[36]</sup> and calbindin,<sup>[37]</sup> as well as those for published tags.<sup>[39,40]</sup> Left: this could not be achieved for the more flexible Cys-Ph-TAHA tagged DNA. Reference to the literature is made as we could not determine the susceptibility tensor of the free tag. Right: Model 3 shows the lowest  $Q_{\text{total}} = Q_{\text{PCS}+\text{RDC}}$  factor for both tags.

phenyl group will not change the position of the lanthanide ions, and therefore we expect that its omission will not compromise the data fitting quality. For cross validation, we calculated the size and rhombicity of the tensor to compare with those of lanthanides in rigid binding pockets of the calcium binding proteins (Table S21). A motional model for the tag was acceptable, when the *Q* factors were small and the size and rhombicity of the tensor were well reproduced.

As the number of possible angles for a free rotation about the triple bond and the phenyl ring is large, we first made a reasonable guess of the conformational distribution, namely the Boltzmann distribution taking the electrostatic energy of charged atoms in each conformation into account while excluding steric clashes (model 2). With the calculated probability distribution of conformational ensembles (Figure 3B and C), the magnetic susceptibility tensor of the lanthanide ions were fit (Table S17, Figure 3D and E). As is summarized in

Figure 3F,  $Q_{\text{PCS}+\text{RDC}}$  decreased slightly and  $\Delta\chi_{\text{ax}}$  and  $\Delta\chi_{\text{th}}$  approach those of the calcium binding proteins, thus indicating that model 2 is slightly improved compared to model 1, although model 2 might be over-simplified, and the calculated probability distribution might not be accurate.

In model 3 we did not calculate electrostatic energies, rather, two Gaussian functions were used to fit probability distributions of lanthanide-tagging conformations (Table S18, Figure 3D and E), where each Gaussian function describes the probability distribution as a function of the rotation angle of the carbon-carbon triple bond for each of the two disulfide conformations (Figure 3B and C). As summarized in Figure 3F, the fitting quality was improved with  $Q_{\text{PCS}+\text{RDC}}=0.19$  for Cys-Ph-TAHA tag with  $\text{Tm}^{3+}$  ion and  $Q_{\text{PCS}+\text{RDC}}=0.33$  for MesS-Ph-TAHA tag with  $\text{Tb}^{3+}$  ion; this indicates that the flexibility of lanthanide-tagging is necessary for an accurate modeling of PCS and RDC data. In terms of the magnitude of anisotropy



tensors of the calcium binding proteins, the fitting result of Tb<sup>3+</sup>-loaded MesS-Ph-TAHA is similar, while the fitting result of Tm<sup>3+</sup>-loaded Cys-Ph-TAHA tag is smaller by a factor of more than 3, thus indicating that the model does not capture the mobility of the longer and far more flexible Cys-Ph-TAHA tag, whereas the shorter MesS-Ph-TAHA tag is well described.

The flexibility of lanthanide-tags is the main reason for scaling the magnetic alignment of proteins and oligonucleotides and thus their RDCs, which is put into perspective in the following. The largest H–N RDC at 21.14 Tesla was reduced from 35 Hz (Tb<sup>3+</sup>) and 30 Hz (Tm<sup>3+</sup>) in calmodulin to 15 Hz (Tb<sup>3+</sup>) and 12 Hz (Tm<sup>3+</sup>) in ubiquitin T12 C with a cysteine tag, and further down to 7 Hz (Tb<sup>3+</sup>) and 6 Hz (Tm<sup>3+</sup>) in ubiquitin S57 C with a cysteine tag (Table S19). For the nucleotides reported here, the largest H–C RDC for Cys-Ph-TAHA tag with Tm<sup>3+</sup> ion at 18.79 Tesla is 6 Hz and the largest H–H RDC for MesS-Ph-TAHA tag with Tb<sup>3+</sup> ion at 22.31 Tesla is 8 Hz, both equivalent to H–N RDC at 21.14 Tesla of 4 Hz, which are further reduced from RDCs of proteins. Comparing to RDCs in ubiquitin with cysteine tag, the reduction of RDCs in nucleotides is probably due to the flexibility of the carbon≡carbon triple bond where the lanthanide tag attaches to the nucleotide. As detailed in the Supporting Information, rigidification of the chemical bonds in the lanthanide tagging of 24-mer DNA would theoretically enlarge RDCs by 3–7 times (Figures S3 and S20). Yet, the PCSs and RDCs observed could be interpreted and, due to the availability of <sup>13</sup>C/<sup>15</sup>N-enriched DNA and RNA through solid-phase synthesis H–C and H–N RDCs will be available with this tagging approach.

In conclusion, a strategy for paramagnetic tagging of bases in oligonucleotides is described using a modified nucleobase in DNA with the Cys-Ph-TAHA and MesS-Ph-TAHA tags connected by a disulfide bond. The modified nucleobase can replace any nucleobase and, therefore, facilitate a tagging reaction at any desired position of a synthesized oligonucleotide. When analyzing complexes, of course the tag should not be at a binding site to avoid interference with complex formation. A convenient synthesis enables the preparation of the corresponding phosphoramidite on a 100-mg scale, followed by a fast and highly reproducible cleavage and purification protocol for the oligonucleotides. Lanthanide-induced PCSs and RDCs had sufficient sizes that they could be fitted to a motional model of the tag; this is so far unprecedented in the literature and could be extended to chemically synthesized <sup>13</sup>C/<sup>15</sup>N labeled DNA and RNA to measure H–C and H–N RDCs. With attachment to DNA that was not labeled with isotopes, only a few proton–proton and proton–carbon RDCs could be measured; this was insufficient for structure refinement. Further rigidification of the tag could lead in the future to even larger RDCs.

## Acknowledgement

The authors acknowledge support by the Max Planck Society and help with chromatography from Kerstin Overkamp. Open access funding enabled and organized by Projekt DEAL.

## Conflict of Interest

The authors declare no conflict of interest.

**Keywords:** lanthanide, NMR spectroscopy, oligonucleotide, paramagnetic NMR, tagging

- [1] J. C. Hus, D. Marion, M. Blackledge, *J. Mol. Biol.* **2000**, *298*, 927–936.
- [2] X. Xu, P. H. J. Keizers, W. Reinle, F. Hannemann, R. Bernhardt, M. Ubbink, *J. Biomol. NMR* **2009**, *43*, 247–254.
- [3] J. R. Tolman, J. M. Flanagan, M. A. Kennedy, J. H. Prestegard, *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 9279–9283.
- [4] X. C. Su, K. McAndrew, T. Huber, G. Otting, *J. Am. Chem. Soc.* **2008**, *130*, 1681–1687.
- [5] F. Rodriguez-Castañeda, P. Haberz, A. Leonov, C. Griesinger, *Magn. Reson. Chem.* **2006**, *44*, S10–S16.
- [6] X. C. Su, G. Otting, *J. Biomol. NMR* **2010**, *46*, 101–112.
- [7] L. Russo, M. Maestre-Martinez, S. Wolff, S. Becker, C. Griesinger, *J. Am. Chem. Soc.* **2013**, *135*, 17111–17120.
- [8] O. F. Lange, N.-A. Lakomek, C. Farès, G. F. Schröder, K. F. a Walter, S. Becker, J. Meiler, H. Grubmüller, C. Griesinger, B. L. de Groot, *Science* **2008**, *320*, 1471–1475.
- [9] M. Allegrozzi, I. Bertini, M. B. L. Janik, Y. M. Lee, G. Liu, C. Luchinat, *J. Am. Chem. Soc.* **2000**, *122*, 4154–4161.
- [10] I. Bertini, I. Gelis, N. Katsaros, C. Luchinat, A. Provenzani, *Biochemistry* **2003**, *42*, 8011–8021.
- [11] K. Barthelmes, A. M. Reynolds, E. Peisach, H. R. A. Jonker, N. J. Denunzio, K. N. Allen, B. Imperiali, H. Schwalbe, *J. Am. Chem. Soc.* **2011**, *133*, 808–819.
- [12] L. J. Martin, M. J. Hähnke, M. Nitz, J. Wöhnert, N. R. Silvaggi, K. N. Allen, H. Schwalbe, B. Imperiali, *J. Am. Chem. Soc.* **2007**, *129*, 7106–7113.
- [13] M. Nitz, M. Sherawat, K. J. Franz, E. Peisach, K. N. Allen, B. Imperiali, *Angew. Chem. Int. Ed.* **2004**, *43*, 3682–3685; *Angew. Chem.* **2004**, *116*, 3768–3771.
- [14] J. Wöhnert, K. J. Franz, M. Nitz, B. Imperiali, H. Schwalbe, *J. Am. Chem. Soc.* **2003**, *125*, 13338–13339.
- [15] W. M. Liu, P. H. J. Keizers, M. A. S. Hass, A. Blok, M. Timmer, A. J. C. Sarris, M. Overhand, M. Ubbink, *J. Am. Chem. Soc.* **2012**, *134*, 17306–17313.
- [16] Y. Yang, Q. F. Li, C. Cao, F. Huang, X. C. Su, *Chem. Eur. J.* **2013**, *19*, 1097–1103.
- [17] G. Sicoli, F. Wachowius, M. Bennati, C. Höbartner, *Angew. Chem. Int. Ed.* **2010**, *49*, 6443–6447; *Angew. Chem.* **2010**, *122*, 6588–6592.
- [18] P. Z. Qin, S. E. Butcher, J. Feigon, W. L. Hubbell, *Biochemistry* **2001**, *40*, 6929–6936.
- [19] P. Z. Qin, K. Hideg, J. Feigon, W. L. Hubbell, *Biochemistry* **2003**, *42*, 6772–6783.
- [20] P. Z. Qin, I. S. Haworth, Q. Cai, A. K. Kusnetzow, G. P. G. Grant, E. a Price, G. Z. Sowa, A. Popova, B. Herreros, H. He, *Nat. Protoc.* **2007**, *2*, 2354–2365.
- [21] N. Piton, Y. Mu, G. Stock, T. F. Prisner, O. Schiemann, J. W. Engels, *Nucleic Acids Res.* **2007**, *35*, 3128–3143.
- [22] S. A. Ingale, P. Leonard, H. Yang, F. Seela, *Org. Biomol. Chem.* **2014**, *12*, 8519–32.
- [23] C. H. Wunderlich, R. G. Huber, R. Spitzer, K. R. Liedl, K. Kloiber, C. Kreutz, *ACS Chem. Biol.* **2013**, *8*, 2697–2706.
- [24] J. M. Esquiaqui, E. M. Sherman, S. A. Ionescu, J. D. Ye, G. E. Fanucci, *Biochemistry* **2014**, *53*, 3526–3528.
- [25] Y. Takayama, G. M. Clore, *J. Biol. Chem.* **2012**, *287*, 14349–14363.
- [26] S. Yamamoto, T. Yamaguchi, M. Erdélyi, C. Griesinger, K. Kato, *Chem. Eur. J.* **2011**, *17*, 9280–9282.
- [27] A. Canales, I. Boos, L. Perkams, L. Karst, T. Luber, T. Karagiannis, G. Dominguez, F. J. Canada, J. Perez-Castells, D. Häussinger, C. Unverzagt, J. Jimenez-Barbero, *Angew. Chem. Int. Ed.* **2017**, *56*, 14987–14991; *Angew. Chem.* **2017**, *129*, 15183–15187.
- [28] B. Fernandez de Toro, W. J. Peng, A. J. Thompson, G. Dominguez, F. Javier Canada, J. Perez-Castells, J. C. Paulson, J. Jimenez-Barbero, J. A. Canales, Angeles, *Angew. Chem. Int. Ed.* **2018**, *57*, 15051–15055.
- [29] P. Valverde, J. I. Quintana, J. Santos, A. Arda, J. Jimenez-Barbero, *ACS Omega* **2019**, *4*, 13618–13630.
- [30] M. Erdélyi, E. D’Auvergne, A. Navarro-Vázquez, A. Leonov, C. Griesinger, *Chem. Eur. J.* **2011**, *17*, 9368–9376.

- [31] K. Kato, T. Yamaguchi, *Glycoconjugate J.* **2015**, *32*, 505–513.
- [32] P. M. Keizers, M. Ubbink, *Prog. Nucl. Magn. Reson. Spectrosc.* **2011**, *58*, 88–96.
- [33] Z. Wu, M. D. Lee, T. J. Carruthers, M. Szabo, M. L. Dennis, J. D. Swarbrick, B. Graham, G. Otting, *Bioconjugate Chem.* **2017**, *28*, 1741–1748.
- [34] F. Peters, M. Maestre-Martinez, A. Leonov, L. Kovačič, S. Becker, R. Boelens, C. Griesinger, *J. Biomol. NMR* **2011**, *51*, 329–337.
- [35] E. Wöltjen, *Ph.D. thesis*, University of Göttingen (Germany), **2009**.
- [36] I. Bertini, C. Del Bianco, I. Gelis, N. Katsaros, C. Luchinat, G. Parigi, M. Peana, A. Provenzani, M. A. Zoroddu, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 6841–6846.
- [37] I. Bertini, M. B. L. Janik, Y. M. Lee, C. Luchinat, A. Rosato, *J. Am. Chem. Soc.* **2001**, *123*, 4181–4188.
- [38] P. Trigo-Mourino, Pablo, T. Thestrup, O. Griesbeck, C. Griesinger, S. Becker, *Sci. Adv.* **2019**, *5*, eaaw4988.
- [39] P. H. J. Keizers, J. F. Desreux, M. Overhand, M. Ubbink, *J. Am. Chem. Soc.* **2008** *130*, 14802.
- [40] M. D. Lee, M. L. Dennis, B. Graham, J. D. Swarbrick, *Chem. Commun.* **2017**, *53*, 13205.

---

Manuscript received: June 29, 2020  
Revised manuscript received: July 19, 2020  
Accepted manuscript online: July 20, 2020  
Version of record online: August 17, 2020