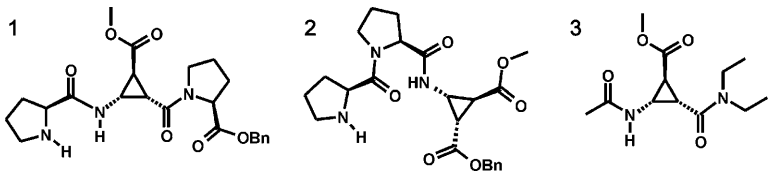
DOI: 10.1002/cbic.200800736

Residual Dipolar Couplings in Short Peptidic Foldamers: Combined Analyses of Backbone and Side-Chain Conformations and Evaluation of Structure Coordinates of Rigid Unnatural Amino Acids

Markus B. Schmid,[a] Matthias Fleischmann,[a] Valerio D'Elia,[a] Oliver Reiser,[a] Wolfram Gronwald,[b] and Ruth M. Gschwind*[a]



Scheme 1. H-(L)-Pro-▼-(L)-Pro-OBn **1**, H-(L)-Pro-(L)-Pro-▼-OBn **2** (for conformational analysis) and Ac-▼-NEt_2 **3** (for the parameterization of *cis*-β-ACC); ▼ = (−)-*cis*-β-ACC.

Although residual dipolar couplings (RDCs) have been established in the NMR structure determination of biomacromolecules in solution for several years,[1,2] their potential for the investigation of small molecules has not been fully explored due to difficulties in the evaluation of internal dynamics and conformational averaging. Thus, RDCs have been mostly employed for the determination of configurations in rigid or cyclic compounds[3–6] or for the structure refinement of cyclic peptides.[7,8] Only very recently could configurations in more flexible open-chain systems be elucidated.[9,10] To our knowledge, RDC applications addressing short linear peptides, however, have been restricted to the systematic elucidation of the conformational preferences of individual natural amino acids;[11,12] neither the conformations of short linear peptides nor the influence or the structure of cyclic unnatural amino acids have been analyzed in detail by RDCs. Since certain unnatural amino acids, for example, β-amino acids, are known to induce strong conformational preferences in peptides, foldamers containiing β-amino acids are promising targets for structural investigations supported by RDCs. In particular, the rigid *cis*-β-aminocyclopropanecarboxylic acid[13] (*cis*-β-ACC) stabilizes even short peptide sequences,[14] and high activities and selectivities have been found in medicinal chemistry[15,16] and asymmetric organocatalysis[17] for cis-β-ACC containing peptides. However, the structural analysis of these peptides by NMR and CD methods established for peptides comprised of natural amino acids has been problematic due to difficulties in assigning parameters for the unnatural building blocks and due to a missing body of reference compounds.

To fill this gap, NMR solution studies on tripeptides **1** and **2** (Scheme 1) are presented here. RDCs were applied as a quality check for the different *cis*-β-ACC parameterizations for molecular dynamics (MD) simulations and to detect preferences in proline side-chain conformations. Furthermore, the use of *cis*-β-ACC as an alignment probe allowed for the analysis of the peptide backbone.

In structural studies of small molecules and nonstandard amino acids, the appropriate parameterization of these compounds is essential for reliable MD simulations. The hydrogen positions, which are addressed by various NMR-derived restraints, especially influence the results of MD calculations. Since a crystal structure of the enantiomer of **3** (Scheme 1) was available,[18] **3** was used as a model compound for **1** and **2** in order to generate a parameterization of *cis*-β-ACC. For this purpose, two sets of coordinates for **3** were prepared, one with the Dundee PRODRG2 Server[19] and the second one from the crystal structure[18] by a subsequent DFT equilibrium geometry calculation resulting in significantly different hydrogen positions (see the structures in Figure 1). As RDCs provide in principle the potential to check the quality of coordinates in rigid molecules, both β-ACC structures were fitted to a set of six experimental RDCs of the rigid β-ACC moiety with the PALES software;[20,21] these RDCs had been determined from the well-resolved 1D ^1H and ^1H-^{13}C P.E.HSQC spectra[22] of samples of **2** in $CDCl_3$ and in a strained polydimethylsiloxane (PDMS)/$CDCl_3$ gel.[23] A comparison of the experimental and the back-calculated RDCs of the two coordinate sets is presented in Figure 1.

Besides demonstrating that the set of experimental data is accurate and the configuration of *cis*-β-ACC is correct, the significantly better Q factor of the X-ray-derived coordinates (Figure 1 B) indicates their congruence with the actual structure in solution. The coordinate file generated by PRODRG2, on the other hand, can be rejected; in addition to a poor Q factor, a significantly different alignment tensor was calculated, which would lead to the erroneous evaluation of further RDCs. However, the PRODRG2 structure can be improved by an energy minimization calculation, leading to results similar to the crystal structure (see the Supporting Information).

[a] *M. B. Schmid, M. Fleischmann, V. D'Elia, Prof. Dr. O. Reiser, Prof. Dr. R. M. Gschwind*
Institut für Organische Chemie, Universität Regensburg
Universitätsstrasse 31, 93053 Regensburg (Germany)
Fax: (+49) 941-943-4617
E-mail: ruth.gschwind@chemie.uni-regensburg.de

[b] *Dr. W. Gronwald*
Institut für Funktionelle Genomik, Universität Regensburg
Josef-Engert-Strasse 9, 93053 Regensburg (Germany)

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/cbic.200800736.

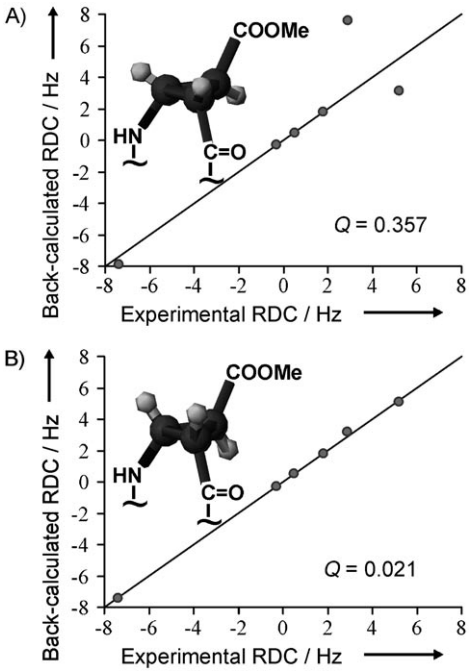
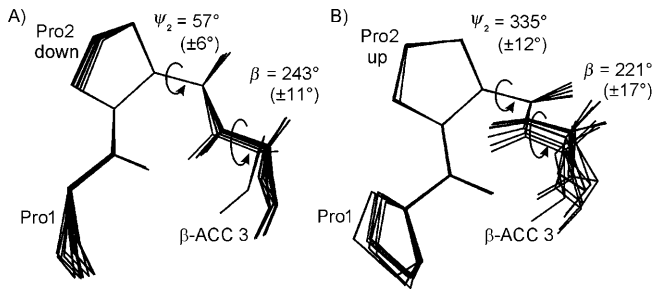


Figure 1. Experimental and back-calculated RDCs for *cis*-β-ACC with structure coordinates A) according to the Dundee PRODRG2 server[19] or B) derived from a crystal structure.[18] The Q factors allow for the identification of the correct amino acid structure.

The impact of such small differences in hydrogen positions on the conformational analyses of short peptides is illustrated by MD calculations on **2** (Scheme 2). The slightly changed



Scheme 2. Conformational changes of **2** caused by different β-ACC parameterizations derived from A) PRODRG2 or B) an X-ray structure of **3**. In both cases, the 50 structures with the lowest NOE energies are displayed, fitted on the N, C_α and CO of Pro 2. The C-terminal benzyl protecting group is omitted for the sake of clarity.

proton coordinates in the two different parameterizations not only lead to an erroneous peptide backbone but also to different proline side-chain conformations. Upon applying the correct parameterization, the most obvious changes occur in the backbone angles Ψ_2 and β (from 57° to 335° and from 243° to 221°, respectively) and in the flip of the side-chain conformation of Pro2 from down to up. The overall conformational alteration affects significantly the relative position of Pro1 and β-ACC, which is essential for the structural interpretation of the organocatalytic properties of unprotected **2**.[17] In the structural studies of small molecules such as **2**, it is essential to check whether the experimental NMR parameters originate from one prevalent conformation or are averaged over several coexisting or interconverting conformations. As criteria for the predominant adoption of a well-folded form of short linear peptides, it has been proposed that no NOE restraints are violated and all interproton distances below 3.5 Å that occur in more than half of the calculated structures give rise to observable NOEs.[14] In addition, the dominant population of a particular conformation can be evidenced experimentally by unusual large or small $^3J_{H\text{-}H}$ values.[24] Interestingly, in both structures shown in Scheme 2, the two NOE criteria are fulfilled; this indicates one conformation for **2**. But only with the correct parameterization, does the experimental β angle derived from $^3J_{HN\text{-}H\beta}$ agree with the calculated structure (Supporting Information). This example shows impressively that RDCs can be used as a tool to evaluate the correctness of the coordinates of rigid unnatural amino acids in solution, which is a crucial prerequisite for reliable restrained MD simulations, especially for short peptidic foldamers.

Inspired by the excellent performance of *cis*-β-ACC in the RDC back-calculation, we used *cis*-β-ACC as a probe for molecular alignment, and thus exploited the potential of RDCs as a source of conformational information for the structure refinement of short peptidic foldamers. In principle, the reported applications of β-ACC-containing foldamers in medicinal chemistry and organocatalysis suggest conformational investigations should be performed at room temperature. Since the degradation of **2** into the Pro-Pro-diketopiperazine is observed even at 273 K, **1** was selected for structural investigations with RDCs. At 300 K, the NOE interpretation of **1** was severely hampered by exchange processes involving the amide proton of β-ACC, which impeded NOE interpretation at this prominent position in the NOE network. In this case, RDC data are expected to be particularly valuable, because they allow for the determination of the orientation of vectors in the different amino acids relative to the alignment tensor.

Even in foldamers, linear tripeptides are very likely to show remaining internal dynamics or, even more seriously, several interconverting conformations; these situations would lead to a complex averaging of the experimental NMR parameters, including RDCs. In order to find out if parts of **2** were conformationally restricted enough to reduce the influence of conformational averaging to a minimum, and thus, to allow for RDC interpretation, the conformational stability of **1** was investigated at 240 K, and the resulting structure analyzed according to the NOE and *J* criteria discussed above. The NOE-restrained MD simulations indicated a surprisingly limited conformational flexibility as far as rotations around the backbone angles α and β are concerned (see Figure 2 A for nomenclature and Figure S3 for the calculated structure ensemble). The predominance of the conformation shown in Figure S3 at 240 K was revealed by the fulfillment of all the NOE restraints and by observable NOEs for all interproton distances below 3.5 Å. The unexpected stability of the β angle was experimentally confirmed by the unusual $^3J_{HN\text{-}H\beta}$[24] of 9.69 Hz (Supporting Information). At 300 K, small changes in the chemical shifts (<0.07 ppm, except for

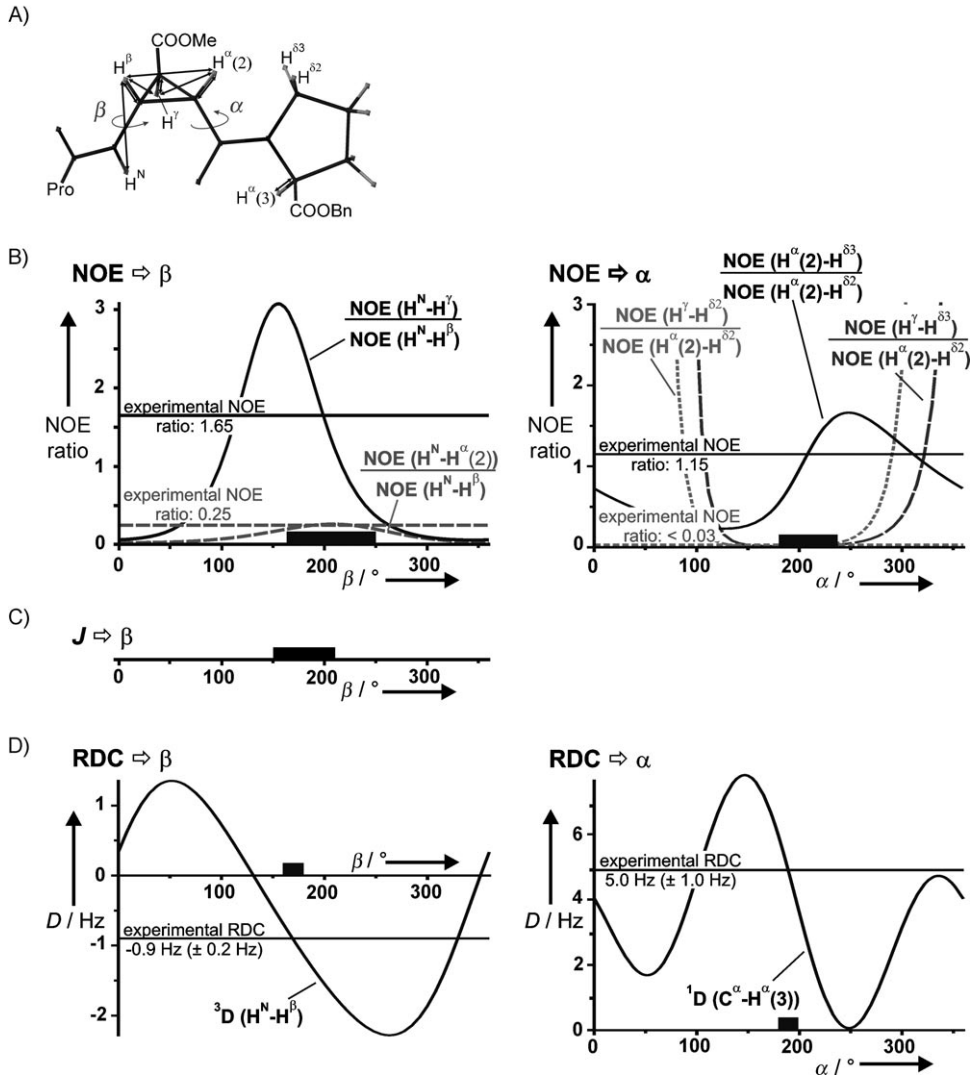


Figure 2. A) Conformational information on the angles α and β in **1** derived from B) NOEs, C) $^{3}J_{HN\text{-}H\beta}$ and D) RDCs. The black bars mark angle regions in agreement with the experimental parameters.

the amide proton), a qualitatively identical NOE intensity pattern and the slight decrease in the $^{3}J_{HN\text{-}H\beta}$ (from 9.69 Hz to 8.89 Hz) showed that this conformation was still preferred and only slightly loosened at higher temperatures. The formation of an intraresidual hydrogen bond, which is hinted at by IR spectroscopy at room temperature,[18] corroborates this conformation. On the other hand, the orientations of Pro1 and the benzyl protecting group are less defined. Therefore, only those RDCs solely affected by rotations around the angles α and β (with respect to the alignment probe *cis*-β-ACC) can be interpreted without significant conformational averaging, and thus, were used for a structural refinement of the backbone conformation of **1**.

Based on this structural model of **1** and *cis*-β-ACC as an alignment probe, RDC information was gathered on the angles α and β, which describe the orientation of the adjacent parts of the peptide relative to the probe (Figure 2A). For the angle β, $^{3}D_{HN\text{-}H\beta}$ was determined from 1D ^{1}H spectra, and $^{1}D_{C\alpha\text{-}H\alpha}$ within Pro3 (from PEHSQC spectra) was chosen as the sensor for the rotation around α, as it is the only α-relevant RDC that is practically independent of the proline side-chain conformation. The conformational space to be investigated by RDC analysis was mapped by calculating structures of **1** without any distance restraints. The alignment tensors for these structures were then determined by the six RDCs within the rigid *cis*-β-ACC residue, and the two RDCs relevant for α and β were back-calculated and checked against experimental values. Figure 2 displays the results of the relative NOE and RDC evaluation for α and β.

For α and β, the RDC evaluation matches angles of 175°–200° and 160°–180°, respectively, which are within the angle ranges derived from the relative NOE intensity pattern (Figure 2). In addition, the β angle range determined by RDCs is supported by an estimated Karplus curve for $^{3}J_{HN\text{-}H\beta}$ (8.89 Hz at 300 K) that indicates an angle of $180 \pm 30°$ (Figure 2C, see the Supporting Information for details). Order parameters, which take remaining internal dynamics into account, had not been included in the RDC interpretation so far. Therefore, the good agreement of the RDC data with other NMR parameters, which are each dynamically averaged in different ways, can be interpreted as further proof of a considerably stable core conformation of **1** in $CDCl_3$ at room temperature.

Focusing on the backbone rotations around α and β, we applied the RDC data in MD simulations with the software system CNS (Crystallography & NMR System).[25] Our set of eight RDCs (supplemented by the α-relevant $^{1}D_{C\delta\text{-}H\delta2}$, $^{1}D_{C\delta\text{-}H\delta3}$ and $^{2}D_{H\delta2\text{-}H\delta3}$) were used as the restraints on the basis of the alignment tensor determined for *cis*-β-ACC. For comparison, we quantified the α- and β-relevant NOE contacts, as far as this was possible, for the exchange-affected NOEs and applied them as distance restraints in separate calculations. The resulting structures (presented in Figure 3) show similar α and β values and a remarkable correspondence of the entire conformation for both parameter sets. The reduced conformational flexibility of the RDC-derived structures compared to that of the NOE-derived structures can be explained by the limited precision with which the NOEs could be obtained due to ex-

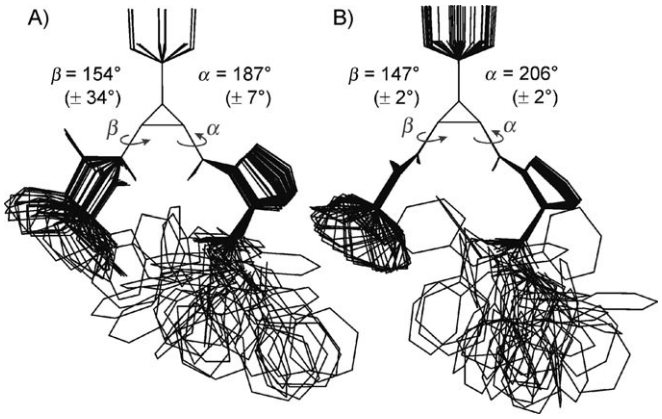


Figure 3. Comparison of A) NOE- and B) RDC-derived structures of **1** at room temperature. In both cases, the 40 structures with the lowest energies are presented, which are based on α/β-relevant NOEs or RDCs only.

change contributions; this required the use of relatively large error bounds for NOEs. In contrast, the RDCs could be determined very precisely. This example shows that, even in short linear peptides at natural abundance, conformational information can be derived from RDCs in the case of a conformationally stable compound and the presence of a reliable probe for molecular alignment.

Besides providing information on the peptide backbone, in principle, RDCs can also be used to gain insight into proline side-chain conformations. In order to cancel the influence of variable backbone conformations on the proline RDCs, separate alignment tensors were calculated for Pro1 and Pro3. As a result, the RDC analysis reflects the averaging of the proline side-chain conformations only and should allow for the detection of preferences. Based on a two-state model assumption for proline conformations ("up"⇄"down", Figure 4A),[26,27] structures of **1** were calculated in which *J* coupling restraints forced the proline residues into one of the two conformations. These structures were used to establish the conformation-dependent correlations of experimental and back-calculated proline RDCs (10 for Pro 1 and six for Pro 3) displayed in Figure 4B and C.

For Pro3, the excellent match of the down conformation and the experimental RDCs (Figure 4C) indicates a distinct preference of this conformation, which is supported by an analysis of scalar coupling constants.[28] For Pro1 on the other hand, the Q factors are poor for both the up and down conformations (0.686 and 0.747, respectively, Figure 4B). It can be optimized to 0.363 by assuming an equilibrium up/down population of 53:47, which is in agreement with the *J* analysis of Pro1.[28] The significantly larger Q factor of Pro1 compared to that of Pro3 can be explained by the presence of fast internal motion at single sites in Pro1 or the population of other conformations beside up and down.

In summary, we have presented an RDC-based approach to select appropriate structures for the force-field parameterization of rigid nonstandard amino acids. Conformational analyses of H-(L)-Pro-(L)-Pro-▼-OBn show that even slight alterations in the proton positions of unnatural amino acids can lead to significant deviations in backbone and side-chain conformations. In the presence of *cis*-β-ACC as a probe for molecular alignment, RDCs allowed us to obtain conformational information on the backbone of the short linear peptide H-(L)-Pro-▼-(L)-Pro-OBn, being especially valuable if the NOE analysis is affected by chemical exchange. In addition, RDCs were employed to detect the preferences of proline side-chain conformations.

This study shows that even in the case of short linear peptides with unnatural amino acids, RDCs at natural abundance can provide essential structural information. This example demonstrates that the RDC approach can be expanded to open-chain structures not only in the case of configuration determination, but also in the field of conformational analyses. In the context of β-ACC, RDC-supported conformational studies are expected to help establish the structure activity/selectivity relationships of neuropeptide Y (NPY) analogues, integrin ligands and organocatalysts that have been developed with this amino acid.

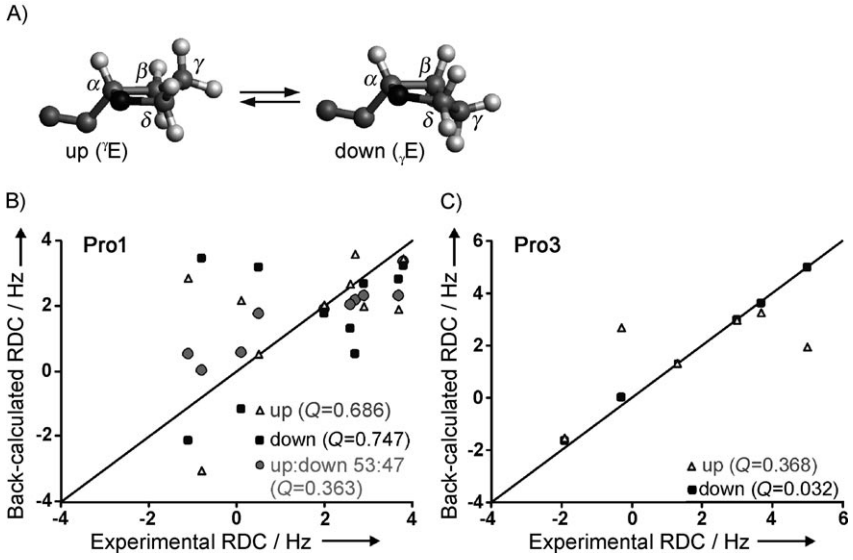


Figure 4. A) Up and down conformations of proline and experimental and back-calculated RDCs for B) Pro1 and C) Pro3.

Acknowledgements

This work was supported by the DFG (GRK 760 and SPP 1179) and the Bavarian Genomic Network (BayGene). Scholarships by the Cusanuswerk, the Studienstiftung des deutschen Volkes and the DAAD are gratefully acknowledged. We kindly thank B. Luy for providing PDMS-gels.

Keywords: conformation analysis · NMR spectroscopy · peptides · peptidic foldamers · residual dipolar couplings · unnatural amino acids

[1] J. H. Prestegard, C. M. Bougault, A. I. Kishore, *Chem. Rev.* **2004**, *104*, 3519.
[2] B. Simon, M. Sattler, *Angew. Chem.* **2002**, *114*, 453; *Angew. Chem. Int. Ed.* **2002**, *41*, 437.
[3] C. Farès, J. Hassfeld, D. Menche, T. Carlomagno, *Angew. Chem.* **2008**, *120*, 3782; *Angew. Chem. Int. Ed.* **2008**, *47*, 3722.
[4] C. M. Thiele, *Conc. Magn. Reson.* **2007**, *30A*, 65.
[5] C. M. Thiele, S. Berger, *Org. Lett.* **2003**, *5*, 705.
[6] C. M. Thiele, A. Marx, R. Berger, J. Fischer, M. Biel, A. Giannis, *Angew. Chem.* **2006**, *118*, 4566; *Angew. Chem. Int. Ed.* **2006**, *45*, 4455.
[7] J. Klages, C. Neubauer, M. Coles, H. Kessler, B. Luy, *ChemBioChem* **2005**, *6*, 1672.
[8] U. M. Reinscheid, J. Farjon, M. Radzom, P. Haberz, A. Zeeck, M. Blackledge, C. Griesinger, *ChemBioChem* **2006**, *7*, 287.
[9] A. Schuetz, T. Murakami, N. Takada, J. Junker, M. Hashimoto, C. Griesinger, *Angew. Chem.* **2008**, *120*, 2062; *Angew. Chem. Int. Ed.* **2008**, *47*, 2032.
[10] A. Schuetz, J. Junker, A. Leonov, O. F. Lange, T. F. Molinski, C. Griesinger, *J. Am. Chem. Soc.* **2007**, *129*, 15114.
[11] S. A. Dames, R. Aregger, N. Vajpai, P. Bernado, M. Blackledge, S. Grzesiek, *J. Am. Chem. Soc.* **2006**, *128*, 13508.
[12] S. Ohnishi, D. Shortle, *Proteins Struct. Funct. Genet.* **2003**, *50*, 546.
[13] R. Beumer, C. Bubert, C. Cabrele, O. Vielhauer, M. Pietzsch, O. Reiser, *J. Org. Chem.* **2000**, *65*, 8960.
[14] S. De Pol, C. Zorn, C. D. Klein, O. Zerbe, O. Reiser, *Angew. Chem.* **2004**, *116*, 517; *Angew. Chem. Int. Ed.* **2004**, *43*, 511.
[15] N. Koglin, C. Zorn, R. Beumer, C. Cabrele, C. Bubert, N. Sewald, O. Reiser, A. G. Beck-Sickinger, *Angew. Chem.* **2003**, *115*, 212; *Angew. Chem. Int. Ed.* **2003**, *42*, 202.
[16] S. Urman, K. Gaus, Y. Yang, U. Strijowski, N. Sewald, S. De Pol, O. Reiser, *Angew. Chem.* **2007**, *119*, 4050; *Angew. Chem. Int. Ed.* **2007**, *46*, 3976.
[17] V. D'Elia, H. Zwicknagl, O. Reiser, *J. Org. Chem.* **2008**, *73*, 3262.
[18] C. Zorn, Ph.D. Thesis, Universität Regensburg (Germany), **2001**.
[19] A. W. Schüttelkopf, D. M. F. van Aalten, *Acta Crystallogr. Sect. D: Biol. Crystallogr.* **2004**, *60*, 1355.
[20] M. Zweckstetter, *Nat. Protoc.* **2008**, *3*, 679.
[21] M. Zweckstetter, A. Bax, *J. Am. Chem. Soc.* **2000**, *122*, 3791.
[22] P. Tzvetkova, S. Simova, B. Luy, *J. Magn. Reson.* **2007**, *186*, 193.
[23] J. C. Freudenberger, P. Spiteller, R. Bauer, H. Kessler, B. Luy, *J. Am. Chem. Soc.* **2004**, *126*, 14690.
[24] H. J. Dyson, P. E. Wright, *Annu. Rev. Biophys. Biophys. Chem.* **1991**, *20*, 519.
[25] A. T. Brünger, P. D. Adams, G. M. Clore, W. L. DeLano, P. Gros, R. W. Grosse-Kunstleve, J.-S. Jiang, J. Kuszewski, M. Nilges, N. S. Pannu, R. J. Read, L. M. Rice, T. Simonson, G. L. Warren, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **1998**, *54*, 905.
[26] M. Cai, Y. Huang, J. Liu, R. Krishnamoorthi, *J. Biomol. NMR* **1995**, *6*, 123.
[27] Z. L. Madi, C. Griesinger, R. R. Ernst, *J. Am. Chem. Soc.* **1990**, *112*, 2908.
[28] M. Schmid, Diploma Thesis, Universität Regensburg (Germany), **2007**.

Received: November 11, 2008
Published online on January 20, 2009

