Polyketide Macrolides

DOI: 10.1002/anie.200800225

Simultaneous Determination of the Conformation and Relative Configuration of Archazolide A by Using Nuclear Overhauser Effects, *J* Couplings, and Residual Dipolar Couplings**

*Christophe Farès, Jorma Hassfeld, Dirk Menche, and Teresa Carlomagno**

Natural products with promising drug activity must undergo a long configuration and conformation characterization before their target-binding mode can be investigated and exploited. One prime candidate, the polyketide archazolide A (ArcA, Figure 1), is a unique 24-membered-macrolide molecule isolated from the myxobacterium *Archangium gephyra*.[1] This compound and its derivatives are of interest because they exhibit very potent cytotoxic activity by inhibiting the growth of numerous cell lines at concentrations down to the subnanomolar range.[2] Their activity is known to involve a selective interaction with the membrane-bound portion of the vacuolar-type adenosine triphosphatase (V-ATPase),[2] a proton-translocating enzyme that has been linked to various diseases such as osteoporosis, renal acidosis, and various cancers.[3,4] This strong interaction has stimulated the investigation of the binding mechanism of ArcA at a molecular level.[5–8]

For the determination of the receptor-bound structure of ligands, it is essential to elucidate the configuration at all stereogenic centers. For ArcA, this has been achieved by a demanding procedure based on combined chemical derivatization and NMR spectroscopy methods. The success of this analysis on ArcA has led to the design of a synthetic route, which has allowed the confirmation of the configuration, given access to more advantageous yields, and enabled strategic derivatization to modulate the drug activity.[6–9]

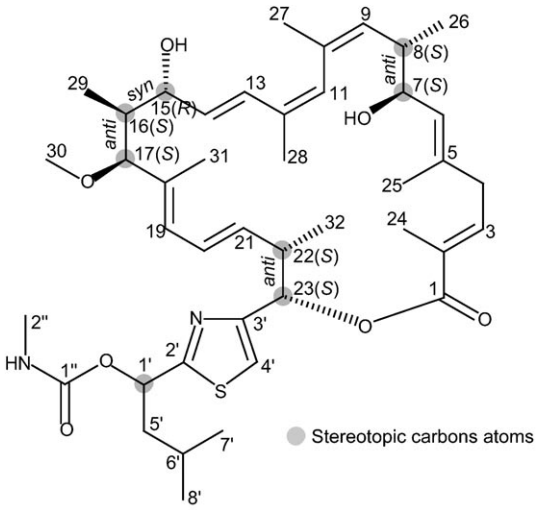

Figure 1. Structure of ArcA showing configuration, atom numbering, and nomenclature.

[*] Dr. C. Farès, Dr. T. Carlomagno
Max-Planck-Institute for Biophysical Chemistry
NMR-based Structural Biology
Am Fassberg 11, 37077 Göttingen (Germany)
Fax: (+49) 551-201-2202
E-mail: taco@nmr.mpibpc.mpg.de
Dr. D. Menche
Helmholtz-Zentrum für Infektionsforschung GmbH
Medzinische Chemie
Inhoffenstrasse 7, 38124 Braunschweig (Germany)
Dr. J. Hassfeld
Bayer Schering Pharma AG
Müllerstrasse 178, 13342 Berlin (Germany)

[**] This work was supported by the Max-Planck-Gesellschaft, the Volkswagenstiftung (Funding Initiative: "Interplay between Molecular Conformations and Biological Function"), and the Fonds der Chemischen Industrie ("Liebig-Stipendium" to D.M.). We thank Tatjana Arnold for technical support and Dr. Peter Harberz for his precious help in the preparation of the PH gel.

Supporting information for this article is available on the WWW under http://www.angewandte.org or from the author.

As an alternative to this approach, our group has explored a novel NMR-based route, centered on residual dipolar coupling (RDC) constraints, to solve the structure and stereochemistry problem of ArcA. RDCs are small spin–spin couplings obtained by imposing weak alignment on the molecule of interest, for instance, by dispersing it in liquid-crystalline media or in compressed gels. The dependence of RDCs on the orientation of the spin–spin vector in the magnetic field allows one to determine the relative angle between distant spin pairs and makes RDCs a convenient "long-range" complement to short-range NOEs. In recent years, the introduction of RDCs opened the way to the assignment of diastereotopic protons[10] (dihydropyridone,[11] strychnine,[12] and menthol[13]) and to the elucidation of the relative stereochemistry of nonconsecutive asymmetric centers in small molecules.[14] However, most of the reported elucidations have been for small and rigid organic molecules with relatively close stereocenters, such as spiroindene[15] and

a glutamate analogue,[16] while the problem of the relative configuration of neighboring stereocenters in more flexible molecules, such as saggitamide, has been addressed only recently.[17] Thus, the application of the RDC-based method to ArcA, with three well-separated groups of stereocenters and unknown conformation, constitutes an important methodological challenge. In this report, we show that a combination of NMR parameters (NOEs, *J* couplings, and RDCs) readily available in the 24-membered macrocyclic ring of ArcA (single sample in dimethylsulfoxide (DMSO)) accurately yielded the structure of this complex polyketide and permitted the correct relative configuration of the macrolactone ring to be singled out from 64 ($2^{(7-1)}$) pairs of enantiomers with 7 stereocenters.

The combination of two- and three-bond H–H and H–C scalar couplings and local NOEs may be used to determine the relative configuration and rotameric state of vicinal and proximal stereogenic centers in organic molecules, except in a few degenerate cases.[18] While this approach has been shown to be particularly valuable for conformationally rigid organic molecules, the usefulness of this method for flexible molecules remains more ambiguous.[19,20] Furthermore, the determination of the relationship between remote stereogenic centers is not possible through *J*-coupling-based configuration analysis.

A higher level of confidence in the determination of the correct structure is obtained by incorporating the following three strategies: 1) the inclusion of a large number of NOE restraints, both local and remote, to reach a self-consistent unique overall conformation; 2) the full-relaxation treatment of the ensemble of protons in the structure to account for possible spin-diffusion; and 3) the complementing of the short-range NOEs with long-range RDCs. In this work, all three of these techniques were implemented in a protocol that was simultaneously applied to all possible diastereoisomeric topologies of ArcA (see the Experimental Section).

We have previously shown that the combination of NOEs and *J* couplings (Table S1 in the Supporting Information) provides a straightforward answer for the relative configuration of all bonded chiral pairs with a well-defined conformation, whereas the analysis of *J* couplings alone fails.[5] Here, we submitted all 64 possible configurations to structure calculations by using the same set of restraints consisting of *J* couplings and NOEs (Table S2 in the Supporting Information) but no RDCs. In Figure 2, the lowest conformational energies out of the 50 calculated are shown for all of the configurations, grouped according to local topologies. In agreement with what was shown previously,[5] the results indicate that structures with the combination *anti–syn*, *anti–anti* (*a-sa-a*; see Figure 1 for the short-hand designation) are most compatible with the ensemble of restraints. Also, the next three combinations of configuration with lowest energy (*a-aa-a*, *a-ss-a*, *a-sa-s*),

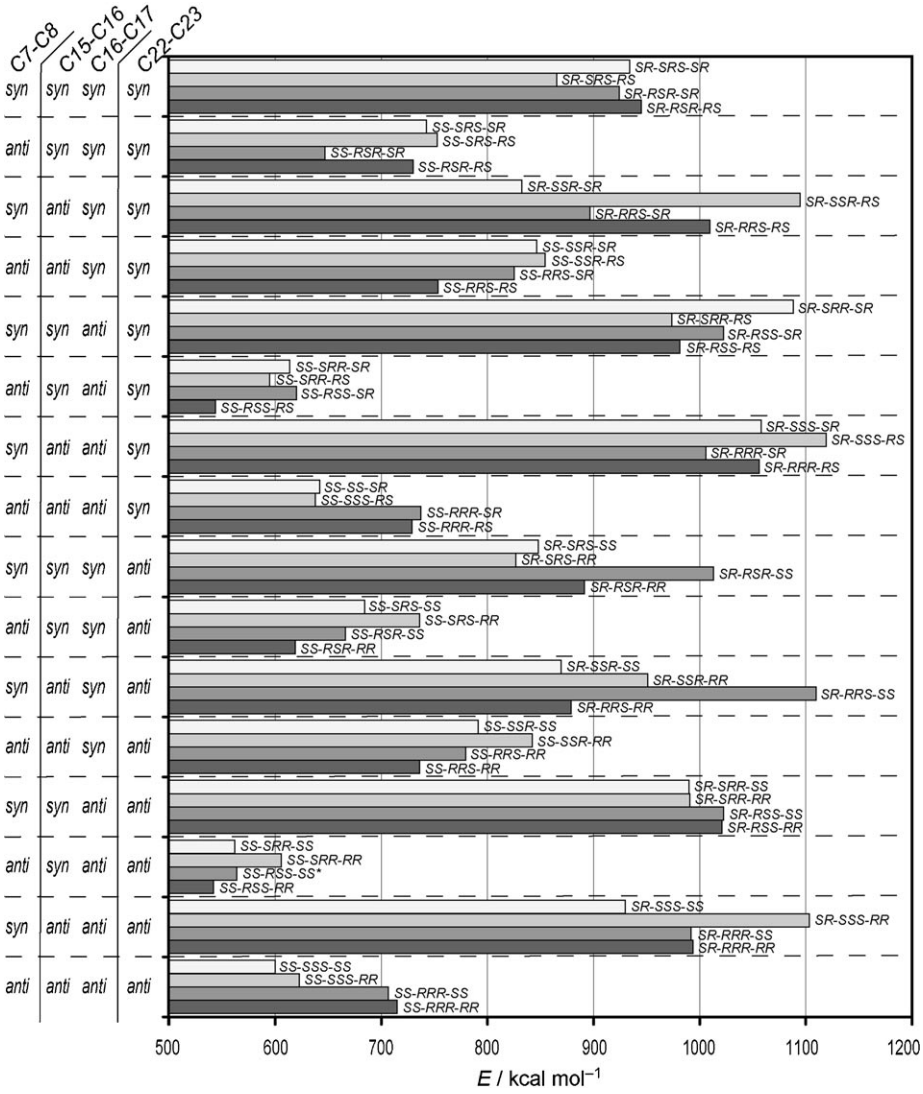

Figure 2. Histogram displaying the final lowest overall energies attained for 64 structures of ArcA with different relative configurations, computed from the XPLOR simulated annealing protocol in the presence of distance and torsion-angle restraints derived from NMR parameters (NOEs and *J* couplings). Different configurations are grouped into families with common vicinally related configurations for the bonded stereotopic pairs of C7—C8, C15—C16, C16—C17, and C22—C23. The family with the lowest overall energies corresponds to the *a-sa-a* family. Other low-energy families are closely related, with a single diastereomeric difference (*a-sa-s*, *a-aa-a*, and *a-ss-a*). These results emphasize the robustness of NOE and *J*-coupling restraints for deciphering the local relative configuration but not the distant relative configuration. Asterisk: correct ArcA configuration.

among which the *a-sa-s* group has only slightly higher overall energy than the preferred *a-sa-a* group, are all only one false vicinal diastereoisomeric configuration removed from the lowest energy *a-sa-a* combination. The residual ambiguity between the *a-sa-a* and the *a-sa-s* combinations can be explained by the lack of magnetically active atoms on one of the C23 substituents, which translates into a paucity of restraints for the [C22–C23] site.

These results for ArcA emphasize the fact that the ensemble of NOEs is quite effective at resolving ambiguous cases of vicinally related configurations (for example, *SS* versus *SR* for C7–C8) but does not allow one to distinguish between remotely related configurations (for example, *SS-RSS-SS* versus *SS-SRR-SS* for [C7–C8]-[C15–C16–C17]-[C22–C23]).

To lift the ambiguity left by the structure calculations made with *J*-coupling and NOE data, the structures for all possible diastereomers were tested for compatibility with the ensemble of one-bond non-methyl ^{1}H–^{13}C RDCs measured along the macrolactone ring of ArcA, obtained by dispersing the sample in a polyacrylamide gel (Table S3 in the Supporting Information).[21] The quality of the fit (Q) was evaluated as a normalized root mean square deviation (rmsd) against the measured RDCs.[22] The conformations obtained from the *J*-coupling and NOE restraints were not refined versus the RDCs prior to fitting. Figure 3 maps the lowest energy structures for each configuration according to their overall energy (calculated with the XPLOR program; see the Experimental Section) and their RDC-based Q value. Structures with low values for both of these parameters (bottom left) are in best agreement with the ensemble of measured NMR restraints. The complementary long-range nature of the RDCs is underlined by the lack of correlation between the two parameters. Here, the correct *SS-RSS-SS* configuration distinguishes itself from the competing candidates with its low overall energy, good RDC-based fit, and excellent convergence.

The RDC data discriminate very efficiently among the structures of the *a-sa-a* group with similar NOE and *J*-coupling energies: the *SS-RSS-SS* configuration (red circle, $Q = 0.37$) scores better than the *SS-RSS-RR* (red square, 0.43), *SS-SRR-SS* (red triangle, 0.49), and *SS-SRR-RR* (red diamond, 0.54) configurations. In other groups, the *SS-SRS-SS* (blue diamond, 0.38) and *SS-RSR-RR* (blue square, 0.40) configurations have at least one structure with a similar Q fit to the *SS-RSS-SS* configuration but display much higher energy. Other configurations, such as *SR-RSS-SS* (yellow circle), although closely related to the correct one, have very large Q values (0.77). However, for this configuration, the lowest energy structure has a backbone rmsd of 1.37 Å relative to the *SS-RSS-SS* configuration, thereby underlining the importance of the correct overall conformation for acceptable fitting of the RDCs. On the other hand, the very closely related *SS-RSS-RS* (black square, 0.41) and *SS-RSS-SR* (black circle, 0.41) topologies, with a single wrong stereotopic assignment near the low-proton-density flexible region of the C22 and C23 atoms, have energies and fit qualities close to the correct one. Their lowest energy structures have backbone rmsd values of only 0.28 Å relative to the *SS-RSS-SS* configuration; this implies that a change in chirality at the C22 or C23 atoms does not severely affect the lowest energy conformation. Despite the high similarity in the resulting conformations, the RDC fit slightly favors the *SS-RSS-SS* configuration, a result indicating that, in this case, the Q value is sensitive to a single change of stereotopic assignment. Furthermore, the convergence of the low-energy structures for both the *SS-RSS-RS* and *SS-RSS-SR* configurations is considerably worse than that for the *SS-RSS-SS* configuration; this result confirms that the *SS-RSS-SS* configuration is in best agreement with the complete set of NMR parameters, including the NOEs, *J* couplings, and RDCs.

It is interesting to note that the configuration validation hinges strongly on the RDCs measured near the chiral centers. Indeed, in a comparison of Q values obtained from different subsets of RDCs, the *SS-RSS-SS* topology ranks first when all 16 CH RDCs or only the 7 CH RDCs at the chiral sites are compared, but it only ranks seventh when the other 9

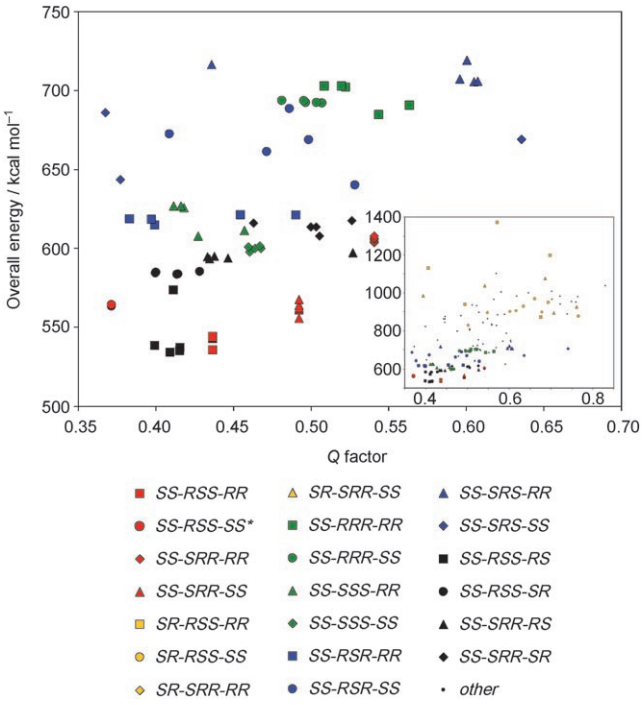

Figure 3. Plot representing the final energy (calculated with the XPLOR program) versus the RDC correlation parameter, Q, for the five lowest energy ArcA conformations (represented with a common symbol) for configurations with low conformational energies or in close stereotopic relation to the correct vicinal relative configuration, *a-sa-a*. In the inset plot, the same data is represented along with the average results for all other structures (single dots). Structures are color grouped by families with common vicinal relative configurations (red: *a-sa-a*; yellow: *s-sa-a*; green: *a-aa-a*; blue: *a-ss-a*; black: *a-sa-s*). The results show that a large number of configurations have poor performance for either the short-range (high *J*-coupling and NOE energies) or long-range (high Q values) structural parameters, or for both. The correct *SS-RSS-SS* configuration is distinguished by an energy value lower than 600 kcal mol^{-1} and a Q value lower than 0.4. This family of structures also displays excellent convergence. Asterisk: correct ArcA configuration.

CH RDCs are used (Figure S1 in the Supporting Information). Furthermore, when the Q values for all 25 RDCs (non-methyl and methyl) from the alignment tensor obtained from the SVD module of the PALES program[23] are compared, an even stronger demarcation of the correct *SS-RSS-SS* structure is observed (Figure S1 in the Supporting Information).

For ArcA, a clear result for the determination of the correct configuration is obtained, with the Q fit values reaching levels expected for good-quality biomolecular structures ($0.25 < Q < 0.40$), without any attempt to compensate for the possible dynamic fluctuations. This is possible only if these fluctuations do not move the structure far away from the average conformation determined from distance constraints. The relatively stiff backbone of ArcA, with its closed structure and a large number of double bonds, features a single flexible site characterized by a rotameric jump (*trans*/*gauche*) for the H22–C22–C23–H23 dihedral angle, according to averaged J couplings. Both rotameric states occur within the resulting ensemble, with the *gauche* rotamer being considerably more populated than the *trans* one, according to both NOE and J-coupling data. The less-populated *trans* rotamer can be accommodated in the cyclic structure of ArcA with little effect on the overall backbone conformation (backbone rmsd of 0.27 Å between the *gauche* and *trans* conformer). Due to the predominance of the *gauche* rotamer and to the very local effect of the conformational flexibility around the C22–C23 bond, the RDC data can be interpreted with a single conformation model. However, for more flexible molecules, fitting of the RDCs to an ensemble of conformations might be necessary.[17]

The correct *SS-RSS-SS* configuration is selected from the validation of the different structures against the measured RDCs prior to refinement. When RDC refinement is applied to the selected *SS-RSS-SS* configuration by increasing the RDC force constant in the XPLOR calculation, an outstanding compliance to the ensemble of RDCs is obtained. The correlation plot in Figure 4B highlights the improvement to a Q value of 0.15 upon refinement. More importantly, the resulting conformations (Figure 4A) are only slightly different from those used in the procedure of configuration validation described above. Indeed, the backbone carbon atoms of the main ring have an rmsd of only 0.18 Å when the structures obtained with and without RDC refinement are compared.

The final proposed structure (Figure 4A) features an elongated ring chain stabilized by a hydrophobic core composed of four of the eight methyl groups. Interestingly, the polar site of the 7-OH group, where a number of analogues have been shown to dramatically disrupt the binding efficiency of ArcA to its target V-ATPase, faces the opposite side of the structure compared to the 15-OH and 17-OMe groups, where derivatization has been shown to have no effect (Table S4 in the Supporting Information).

In conclusion, we have presented a solid and efficient approach to simultaneously determine the conformation and configuration at stereogenic centers of complex organic molecules by using NOEs, J couplings, and RDC data in solution. While RDCs have been used before to determine the configuration of small rigid molecules or the relative configuration of consecutive stereogenic centers in flexible molecules, the approach presented here is used to unravel the configuration and conformation of a large complex ligand with seven stereogenic centers located at three distant positions. This approach paves a new way for the determination of the configuration of organic compounds with no need for time-demanding chemical synthesis.

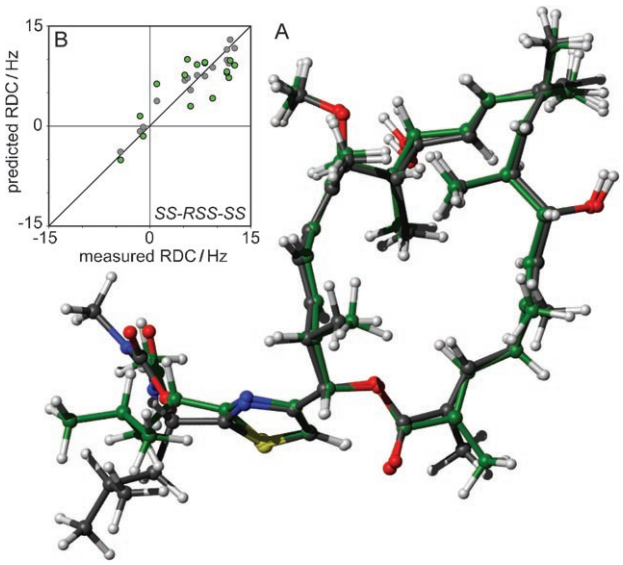

Figure 4. A) Overlay of the RDC-unrefined (green) and RDC-refined (gray) final structures of ArcA (rmsd = 0.18 Å) in the correct *SS-RSS-SS* configuration; red O, blue N, yellow S. B) Correlation plot of the 16 CH theoretical versus experimental RDCs for the structure of the *SS-RSS-SS* configuration of ArcA before (green) and after (gray) refinement.

Experimental Section

ArcA from *A. gephyra* was dissolved in [D_6]DMSO, to a concentration of 15 mM. Torsion-angle restraints were obtained from the interpretation of two- and three-bond homonuclear J_{HH} couplings, measured from 1D proton and DQF-COSY experiments, and of the two- and three-bond heteronuclear J_{CH} couplings, acquired by using the HECADE and HSQMBC sequences. Internuclear-distance restraints were inferred from the full-relaxation matrix treatment of proton–proton cross-relaxation rates (NOEs), with 6 mixing times of 100, 150, 200, 250, 300, and 350 ms. In all, 11 dihedral restraints and 97 complete 6-point NOE build-up restraints were used (Table S2 in the Supporting Information). In addition, 25 RDC orientation restraints were measured along the macrolactone ring (Table S3 in the Supporting Information). These were obtained by dispersing the sample in a polyacrylamide PH gel.[21] The $^1J_{CH}$ and $^1J_{CH} + {}^1D_{CH}$ splittings were measured for all CH bonds prior and subsequent to alignment in the gel, as described elsewhere.[21]

The structure calculations were set up for the 64 different relative diastereoisomers of ArcA. The NOE- and J-coupling-derived restraints were imposed by using the "RELAXATION"[24] and "DIHEDRAL"[25] modules of the XPLOR-NIH Ver. 2.9.7 program. Initially, all non-methyl CH RDCs were introduced into the calculations with a minimized force constant for the RDC target function (defined by the "TENSOR"[26] module of the XPLOR program), in order to calculate the fit of the measured RDCs to the back-calculated ones without refining the structure versus the RDC data. Later, RDC-refined structures were obtained by simply

augmenting the force constant to the same order of magnitude as the NOE and *J*-coupling restraints.

Received: January 16, 2008
Published online: April 10, 2008

Keywords: conformation analysis · natural products · NMR spectroscopy · polyketides · residual dipolar coupling

[1] F. Sasse, H. Steinmetz, G. Höfle, H. Reichenbach, *J. Antibiot.* **2003**, *56*, 520.
[2] M. Huss, F. Sasse, B. Kunze, R. Jansen, H. Steinmetz, G. Ingenhorst, A. Zeeck, H. Wieczorek, *BMC Biochem.* **2005**, *6*, 13.
[3] S. R. Sennoune, D. Luo, R. Martinez-Zaguilan, *Cell Biochem. Biophys.* **2004**, *40*, 185.
[4] K. W. Beyenbach, H. Wieczorek, *J. Exp. Biol.* **2006**, *209*, 577.
[5] J. Hassfeld, C. Fares, H. Steinmetz, T. Carlomagno, D. Menche, *Org. Lett.* **2006**, *8*, 4751.
[6] D. Menche, J. Hassfeld, J. Li, S. Rudolph, *J. Am. Chem. Soc.* **2007**, *129*, 6100.
[7] D. Menche, J. Hassfeld, F. Sasse, M. Huss, H. Wieczorek, *Bioorg. Med. Chem. Lett.* **2007**, *17*, 1732.
[8] D. Menche, J. Hassfeld, H. Steinmetz, M. Huss, H. Wieczorek, F. Sasse, *J. Antibiot.* **2007**, *60*, 328.
[9] P. A. Roethle, I. T. Chen, D. Trauner, *J. Am. Chem. Soc.* **2007**, *129*, 8960.
[10] A. Meddour, C. Canlet, L. Blanco, J. Courtieu, *Angew. Chem.* **1999**, *111*, 2558; *Angew. Chem. Int. Ed.* **1999**, *38*, 2391.
[11] C. Aroulanda, V. Boucard, F. Guibe, J. Courtieu, D. Merlet, *Chem. Eur. J.* **2003**, *9*, 4536.
[12] 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.
[13] L. Verdier, P. Sakhaii, M. Zweckstetter, C. Griesinger, *J. Magn. Reson.* **2003**, *163*, 353.
[14] J. Yan, F. Delaglio, A. Kaerner, A. D. Kline, H. Mo, M. J. Shapiro, T. A. Smitka, G. A. Stephenson, E. R. Zartler, *J. Am. Chem. Soc.* **2004**, *126*, 5008.
[15] J. C. Freudenberger, P. Spiteller, R. Bauer, H. Kessler, B. Luy, *J. Am. Chem. Soc.* **2004**, *126*, 14690.
[16] J. C. Freudenberger, S. Knor, K. Kobzar, D. Heckmann, T. Paululat, H. Kessler, B. Luy, *Angew. Chem.* **2005**, *117*, 427; *Angew. Chem. Int. Ed.* **2005**, *44*, 423.
[17] A. Schuetz, J. Junker, A. Leonov, O. F. Lange, T. F. Molinski, C. Griesinger, *J. Am. Chem. Soc.* **2007**, *129*, 15114.
[18] N. Matsumori, D. Kaneno, M. Murata, H. Nakamura, K. Tachibana, *J. Org. Chem.* **1999**, *64*, 866.
[19] G. Bifulco, P. Dambruoso, L. Gomez-Paloma, R. Riccio, *Chem. Rev.* **2007**, *107*, 3744.
[20] C. Bassarello, G. Bifulco, A. Zampella, M. V. D'Auria, R. Riccio, L. Gomez-Paloma, *Eur. J. Org. Chem.* **2001**, 39.
[21] P. Haberz, J. Farjon, C. Griesinger, *Angew. Chem.* **2005**, *117*, 431; *Angew. Chem. Int. Ed.* **2005**, *44*, 427.
[22] G. Cornilescu, J. L. Marquardt, M. Ottiger, A. Bax, *J. Am. Chem. Soc.* **1998**, *120*, 6836.
[23] M. Zweckstetter, A. Bax, *J. Am. Chem. Soc.* **2000**, *122*, 3791.
[24] P. Yip, D. A. Case, *J. Magn. Reson.* **1989**, *83*, 643.
[25] C. D. Schwieters, J. J. Kuszewski, G. M. Clore, *Prog. Nucl. Magn. Resonance Spectrosc.* **2006**, *48*, 47.
[26] H. J. Sass, G. Musco, S. J. Stahl, P. T. Wingfield, S. Grzesiek, *J. Biomol. NMR* **2001**, *21*, 275.

