JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

ARTICLE

pubs.acs.org/JACS

Challenge of Large-Scale Motion for Residual Dipole Coupling Based Analysis of Configuration: The Case of Fibrosterol Sulfate A

Han Sun,[†] Uwe M. Reinscheid,[†] Emily L. Whitson,[‡] Edward J. d'Auvergne,[†] Chris M. Ireland,[‡] Armando Navarro-Vázquez,[*,§] and Christian Griesinger[*,†]

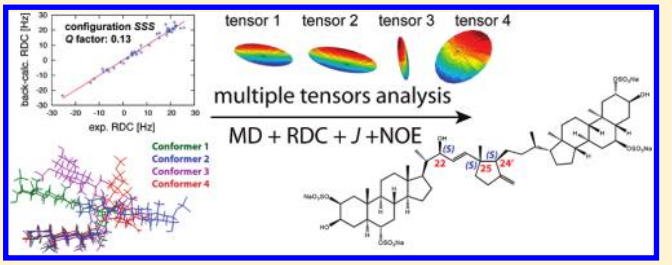
[†]Department of NMR Based Structural Biology, Max Planck Institute for Biophysical Chemistry, Am Fassberg 11, 37077 Göttingen, Germany

[‡]Department of Medicinal Chemistry, University of Utah, Salt Lake City, Utah 84112, United States

[§]Departmento de Química Orgánica, Universidade de Vigo, 36310, Vigo, Spain

Supporting Information

ABSTRACT: Fibrosterol sulfate A is a polysulfated bis-steroid with an atypical side chain. Due to the flexibility of the linker, large-scale motions that change dramatically the shape of the entire molecule are expected. Such motions pose major challenges to the structure elucidation and the correct determination of configuration. In this study, we will describe the determination of the relative configuration of fibrosterol sulfate A through a residual dipolar coupling based multiple alignment tensor analysis complemented by molecular dynamics. For completeness, we applied also the single tensor approach which is unreliable due to the large-scale motions and compare the results.



INTRODUCTION

Structure determination of complex natural products is in many cases a formidable task. Although high-resolution NMR provides a plethora of 2D correlation experiments to determine the constitution of compounds, the configuration is much harder to establish. NMR-based configuration analysis relies in most cases on a combination of proton−proton distance estimation from NOE-based experiments and determination of dihedral angles through Karplus analysis of ^{3}J scalar couplings. However, often ambiguities about the configuration remain, especially when conformationally heterogeneous molecules with multiple stereocenters are analyzed. This may require in many cases a synthesis of the proposed structure to compare the NMR spectra of synthesized and natural products,[1] a process that is time-consuming, cumbersome, and expensive. Thus, the use of NMR parameters that complement NOEs and *J*-couplings becomes desirable. Ab initio computations, mostly DFT based, of chemical shifts and scalar couplings have proven to be useful for the analysis of the constitution and configuration of medium-sized organic molecules.[2,3] Even more powerful is the use of NMR in alignment media, such as liquid crystal solutions or mechanically stretched cross-linked swollen polymers. These media impose certain degrees of orientational ordering on the studied solute molecules, and parameters hidden in isotropic solutions such as residual dipolar couplings (RDCs) or residual chemical shift anisotropies (RCSAs) become observable. RDC analysis is rooted in the seminal work of Saupe in the early 1960s,[4] and it is now a well-established technique for the structural and dynamic studies of biopolymers. The development of new weakly aligning media compatible with organic solvents has enabled the application of RDCs to configurational analysis of small molecules.[5−10] The dipolar coupling between two nuclei depends on their internuclear distance and on the angle between the internuclear vector and the external magnetic field. Dipolar couplings therefore provide long-range structural information which complements NOEs, *J*-couplings, and chemical shifts that are short-range in nature.

From a practical point of view, RDC analysis in small molecules makes use of one-bond proton carbon $^{1}D_{CH}$ couplings as they are easily measured in ^{13}C gated-decoupled or F1/F2-coupled HSQC experiments.[11,12] When the studied molecule is rigid or its conformational space is dominated by a single conformer, analysis is reduced to scoring the fitness of the different structural possibilities, allowing easy simultaneous determination of several stereocenters.[13−16] The same procedure can be applied also to the assignment of prochiral groups through scoring of the different assignment possibilities.[17−20]

When more than one conformation is present in solution, RDC analysis becomes problematic since orientational ordering is not independent of the conformation. Early approaches to the conformational distribution problem were the additive potential or maximum entropy methodologies.[21−24] However, many of the recent applications to flexible molecules make use of the rougher "single-tensor" approximation which assumes that the orientational ordering of a molecule is largely independent of its conformation. This assumption generally holds if the overall molecular shape is similar between different conformations as judged by structural

Received: February 20, 2011
Published: July 21, 2011

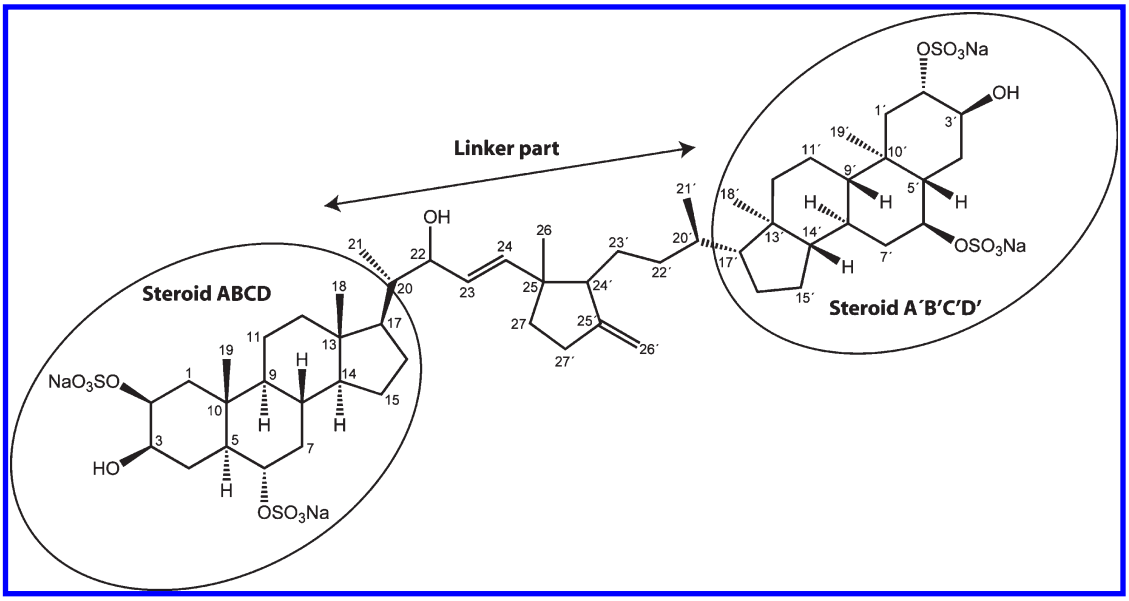


Figure 1. Structure of fibrosterol sulfate A with atom numbering. For convenience, the structure has been divided into three regions: steroid ABCD, linker part, and steroid A′B′C′D′. Determination of the configurations of positions 22, 25, and 24′ is the main focus of this study.

alignments minimizing the rmsd of atom positions. On the basis of the single-tensor approximation, RDC configurational analysis has been performed using $^1D_{CH}$ RDCs as additional restraints in NOE-*J* restrained molecular dynamics (r-MD) on flexible natural products such as sagittamide A,[25] archazolide A,[26] or neo-sucrolambertellin.[27] In another approach, determination of conformer structures is first performed by restrained molecular modeling, and the degree of conformational heterogeneity is determined by fitting the predicted RDCs to the observed ones as a function of the relative populations.[19,28−30] However, if the shapes of the various conformations are very different, the single-tensor approach may no longer be considered valid, and for each conformation, a different alignment tensor needs to be used. In this "multiple-tensor" approach, the number of required experimental anisotropic parameters increases in most cases beyond experimental measurement possibilities and has so far been used only in simple systems.[31] However, one can predict based on known alignment mechanisms the relative sizes of alignment tensors and their orientations for each conformation in the ensemble. Then, the experimental anisotropic parameters can be cross-validated against these predictions. Indeed, this approach has been successfully applied for the cross-validation of ensembles of intrinsically unfolded proteins against anisotropic parameters.[32−35]

In this manuscript, we used RDC-enhanced NMR to determine the configuration of fibrosterol sulfate A,[36] a polysulfated steroid isolated from the sponge *Lissodendoryx* (*Acanthodoryx*) *fibrosa*. It exhibited micromolar activity as an inhibitor of protein kinase C ζ (PKCζ), a protein involved in several types of cancer,[37−41] obesity,[42] and osteoarthritis.[38] The constitution of the compound was established on the basis of 2D NMR correlations as well as high-resolution electron spray ionization mass spectra (HRESIMS) analysis. The structure of this compound (Figure 1) consists of two steroid parts whose relative and absolute configuration was determined by comparison of the chemical shifts of steroids of known absolute configuration and biosynthetic considerations.[36] The linker between the two steroid parts contains five additional stereocenters, of which C20 and C20′ were assigned again by biosynthetic considerations. As described in the Supporting Information, the configuration of these two stereocenters is corroborated by our RDC analysis. Yet, three stereocenters in the linker, C22, C25, and C24′ (Figure 1), had not been assigned before. Long-range information is required to determine the configuration of these three unassigned stereocenters, which led us to use an RDC-based analysis. It is likely that conformational movement in the linker part can dramatically affect the global molecular shape. Thus, we anticipated the need for a multiple-tensor approach to determine the configuration of the linker stereocenters. In the following, we will describe the determination of the relative configuration of fibrosterol sulfate A through an RDC multiple alignment tensor analysis complemented by molecular dynamics computations. We compare these results to those obtained by the simpler single-tensor approach as described in the Supporting Information. Although the single-tensor approach relies on the incorrect assumption of the same alignment tensor for each conformation and is therefore unreliable, we found surprisingly that both methods give similar results despite the apparent large differences in molecular shape with different linker conformations.

METHODOLOGY

1. NMR Measurements. All experiments for the isotropic and anisotropic NMR measurements were performed on an 800 MHz Bruker spectrometer equipped with a triple resonance cryogenic probe. All spectra were recorded at 25 °C.

Isotropic NMR Measurements. Fibrosterol sulfate A (5 mg) was dissolved in 500 μL of DMSO-d_6 to a concentration of 6.6 mM. ^{1}H and ^{13}C resonances were assigned using a complete set of ^{1}H, COSY, NOESY, TOCSY, [^{1}H,^{13}C]-HSQC, and [^{1}H,^{13}C]-HMBC spectra. A [^{1}H,^{13}C]-CLIP-HSQC[11] spectrum was recorded in the absence of the alignment medium as a reference for the measurement of $^1D_{CH}$. A mixing time of 100 ms was used in the TOCSY experiment. For the detection of distance restraints, NOESY spectra were recorded with 10 different mixing times from 60 to 450 ms. We integrated the volume of the cross peaks in the

NOESY spectrum at the mixing time of 150 ms. The integrals were calibrated against the NOESY cross-peak integral for the geminal protons at C23′ with an assumed internuclear distance of 1.78 Å. The $^3J_{HH}$ and $^3J_{CH}$ couplings were extracted from P.E.COSY[43] and HSQMBC[44] spectra, respectively. In the HSQMBC sequence, the delay for the evolution of $^nJ_{CH}$ was set to 125 ms.

Anisotropic NMR Measurements. DMSO-d_6 (650 μL) was added to a 5 mm standard NMR tube together with a dried PAN polymer stick.[5] The swelling of the gel was monitored by observing the ^2D quadrupolar splitting of the solvent signal. After one month, the gel was equilibrated, giving a final splitting of 18.1 Hz. Fibrosterol sulfate A (5 mg) was dissolved in 50 μL of DMSO-d_6 and added on top of the gel which was further equilibrated during 3 weeks at room temperature. Forty-four $^1D_{CH}$ RDCs were extracted by measuring the splitting difference between isotropic and anisotropic signals from [^{1}H,^{13}C]-CLIP-HSQC[11] spectra recorded with a digital resolution of 23.6 Hz in the indirect dimension so that a clear separation of the carbon resonances of the two steroid parts was visible.

2. Computational Methodology. *Explicit Solvent MD Simulations.* The force field parameters for the system were taken from the 53A6 GROMOS united atom force field,[45] whereas the topology of fibrosterol sulfate A was parametrized manually. The starting structure of each configuration for MD was picked out from the simulated annealing conformations (described in the Supporting Information), always choosing a fully extended conformation. Negatively charged fibrosterol sulfate A with four sodium counterions was solvated in a DMSO solvent box using the standard GROMOS96 DMSO solvent model, placing the center of mass of the molecule in the center of the box. The energy of the whole system was first minimized using the conjugate gradient method. After 100 ps of equilibration at 298 K, simulations of 100 ns duration were performed at constant temperature (298 K) and pressure (1 atm) for the eight possible configurations. All MD trajectories were generated using the GROMOS96 software package.[46] The time-averaged NOE and *J*-coupling violations were calculated for the trajectories at snapshots taken at time intervals $\tau \in \{5, 10, 15, ..., 100\}$ ns. The NOE distance bounds were derived from the NOE intensities, and distances were r^{-6} averaged, i.e.

$$(\bar{r})^{-6} = \overline{r^{-6}} \tag{1}$$

which is a standard procedure for small molecules.[47] We should like to point out that NOEs are averaged in a much more complicated way than RDCs, and therefore the large estimated experimental error of 0.5 Å is justified.

Vicinal proton−proton $^3J_{HH}$ and proton−carbon $^3J_{CH}$ coupling constants were calculated from the simulations via the Karplus relationships[48,49]

$$^3J = a\cos^2\theta + b\cos\theta + c \tag{2}$$

where *a*, *b*, and *c* are equal to 9.5, −1.6, and 1.8 Hz, respectively, for $^3J_{HH}$; and for $^3J_{CH}$, *a*, *b*, and *c* are equal to 5.7, −0.6, and 0.5 Hz, respectively. Both $^3J_{HH}$ and $^3J_{CH}$ couplings were linearly averaged over the trajectory.

To map the structures sampled during the simulations onto a set of generic conformations, a clustering analysis was performed.[50] The MD trajectory (0−τ) for the conformational clustering was selected based on the smallest deviation of the back-calculated NOEs and ^3J couplings from the experimental ones. After choosing the best MD trajectory, the root-mean-square difference (rmsd) of the atomic coordinates between pairs of structures was calculated after superimposing all carbon atoms. Two conformations are considered to be similar if their rmsd is below 0.2 nm. The central member structure, defined as the structure with the highest number of neighbors in one cluster, is defined as the representative conformation of this cluster. Populations were determined from the number of structures contained in each cluster. The analysis above was performed for each possible configuration.

RDC Fitting Using a Multiple Alignment Tensor. In the multiple-tensor analysis, the alignment tensor for each conformer is determined individually by using eq 3

$$\langle D\rangle = -\sum_{i=1}^{N} p_i \frac{3\mu_0 h\gamma_A\gamma_B}{16\pi^3 R^3}\vec{r}_i^{\,T} A_i \vec{r}_i \tag{3}$$

in which γ_A and γ_B are the gyromagnetic ratios of each interacting nucleus; *R* is the internuclear distance; p_i is the population; $\vec{r}_i$ is the unit norm vector of the *i*th conformer; and A_i is the alignment of the *i*th conformer. Note that <*D*> depends on the product of p_i and A_i, and therefore the overall degree of alignment cannot be determined unless the populations are previously known;[31] alternatively, the populations can be determined if the norms of the tensors $|A_i|$ are known. The coordinates of the representative conformation for each cluster and its population obtained as described above were given as input for back-calculation of 44 $^1D_{CH}$. The calculation of the alignment tensor for each structure performed by the SVD method was implemented using the Python 2.6 programming language and the scipy[51] numeric library. Methyl $^1D_{CH}$ RDCs were treated by averaging of the corresponding SVD entries as previously described.[19] The corresponding linear equation for determining the alignment tensor for each conformer from the experimental RDCs using the SVD method is shown in eq 4

$$\begin{pmatrix} D_1^{red} \\ D_2^{red} \\ D_3^{red} \\ \cdot \\ \cdot \\ D_{n-1}^{red} \\ D_n^{red} \end{pmatrix} = \hat{M} \begin{pmatrix} A_{xy}^1 \\ A_{xz}^1 \\ A_{yy}^1 \\ A_{yz}^1 \\ A_{zz}^1 \\ \cdot \\ \cdot \\ A_{xy}^m \\ A_{xz}^m \\ A_{yy}^m \\ A_{yz}^m \\ A_{zz}^m \end{pmatrix} \tag{4}$$

where $D_i^{red} = D_i(R^3/\kappa)$ is the reduced RDC, where κ is defined as $\kappa = -(3/16\pi^3)\gamma_I\gamma_S\mu_0 h$; A^i is the alignment tensor of the *i*th conformer; and $\hat{M}$ is a matrix composed of the direction cosines of the internuclear vectors, defined in eq 5

$$\hat{M} = \begin{pmatrix} p_1a_{1,1} & p_1b_{1,1} & p_1c_{1,1} & p_1d_{1,1} & p_1e_{1,1} & ..... & p_na_{n,1} & p_nb_{n,1} & p_nc_{n,1} & p_nd_{n,1} & p_ne_{n,1} \\ p_1a_{1,2} & p_1b_{1,2} & p_1c_{1,2} & p_1d_{1,2} & p_1e_{1,2} & ..... & p_na_{n,2} & p_nb_{n,2} & p_nc_{n,2} & p_nd_{n,2} & p_ne_{n,2} \\ ... & ... & ... & ... & ... & ..... & ... & ... & ... & ... & ... \\ p_1a_{1,m} & p_1b_{1,m} & p_1c_{1,m} & p_1d_{1,m} & p_1e_{1,m} & ..... & p_na_{n,m} & p_nb_{n,m} & p_nc_{n,m} & p_nd_{n,m} & p_ne_{n,m} \end{pmatrix} \tag{5}$$

$$a = 2\cos\phi_x\cos\phi_y, b = 2\cos\phi_x\cos\phi_z, c = (\cos^2\phi_y - \cos^2\phi_x),$$
$$d = 2\cos\phi_y\cos\phi_z, e = (\cos^2\phi_z - \cos^2\phi_x)$$

where p_i is the population of the *i*th conformer, and ϕ is the internuclear vector relative to the molecular axis. The theoretical RDCs were back-calculated from the obtained tensor and the previously determined populations from the MD. We evaluated the

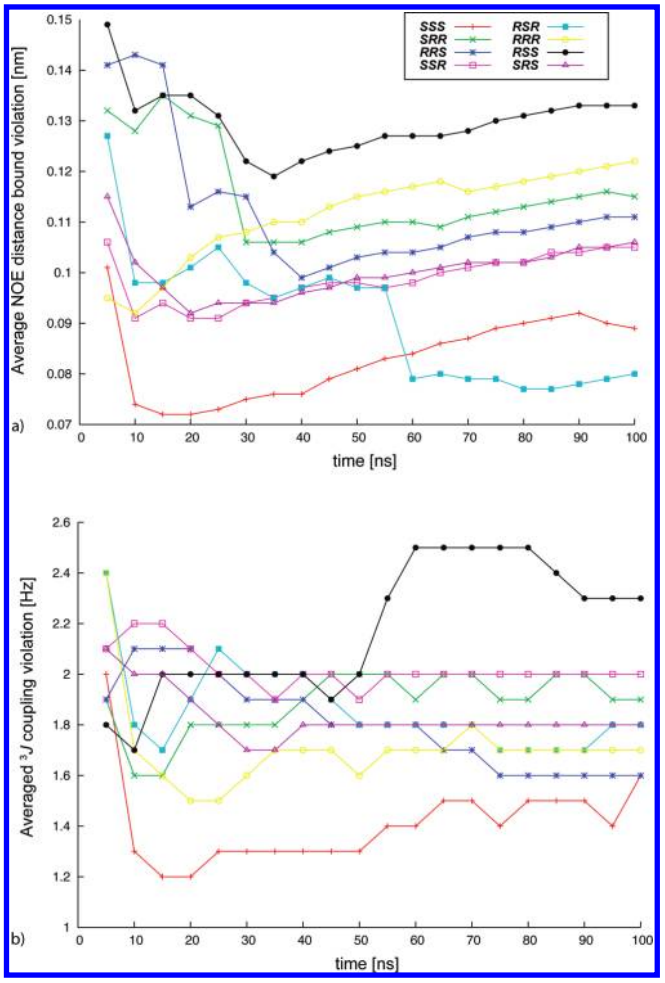


Figure 2. (a) Averaged (r^{-6} averaging) distance minus the NOE upper- or lower-bound distances calculated over all recorded structures after every 5 ns unrestrained MD simulation. The violations are the averaged deviations from 44 NOEs. (b) Averaged ^3J coupling violations calculated over all recorded structures after every 5 ns unrestrained MD simulation. The violations shown in the figure are the averaged deviations from three $^3J_{CH}$ and eight $^3J_{HH}$ couplings that are able to define the conformation in the flexible linker region. The unrestrained MD simulations were performed using the GROMOS 53A6 force field at 298 K.

fitting between theoretical and experimental RDCs according to Cornilescu's Q factor.[52]

RESULTS AND DISCUSSION

Fibrosterol sulfate A is a pseudosymmetric structure with two disulfated steroidal tetracyclic moieties (ABCD and A′B′C′D′, Figure 1) joined by a flexible linker containing a cyclopentyl unit. Complete chemical shift assignments for fibrosterol sulfate A in DMSO were determined using homo- and heteronuclear correlation experiments (Supporting Information, Table S1). Configurations for the rigid steroid ring moieties were previously established on the basis of extensive 1D and 2D NMR experiments.[36] In this study we confirmed the configuration of steroid rings ABCD and A′B′C′D′ by comparing the similarity of the $^1D_{CH}$ RDCs of parallel C−H vectors without calculation of the order tensor[16] since in the six-membered chairlike ring fragment of both steroid moieties [ABC/A′B′C′] all axial C−H bonds are parallel and therefore have nearly the same angle with respect to the alignment tensor. Thus, the $^1D_{CH}$ couplings coming from axial C−H bonds within each of the steroid ring systems should have the same size, which is indeed the case (see Supporting Information, Table S2). However, due to the flexibility of the linker and the relatively small number of coupling constants measured for this region, the configuration of the three stereocenters C22, C25, and C24′ in the linker part remained undetermined. There are eight possible configurations of fibrosterol sulfate A that we refer to using the configuration of C22, C25, and C24′; e.g., *SRS* is the diastereomer that has a configuration of *S* at C22, *R* at C25, and *S* at C24′.

To address this stereochemical question, NOE, $^3J_{HH}$, $^3J_{CH}$, and $^1D_{CH}$ RDCs of fibrosterol sulfate A were measured in DMSO. A set of 44 $^1D_{CH}$ RDCs were obtained by aligning the molecule in a PAN gel:[5] 14 for each of the steroid moieties ABCD and A′B′C′D′ and 16 for the linker moiety (Supporting Information, Table S2).

RDC-Based Multiple Alignment Tensor Analysis of Configuration and Conformation. Due to the flexibility of the linker, the shape of the entire fibrosterol sulfate A is expected to vary dramatically with linker conformation, and therefore the single alignment tensor approach, which has been extensively applied in previous studies, is expected to be problematic in this case.[25,28,30] Hence, we determined the alignment tensor for each main conformer individually according to eq 3 and eq 4, as mentioned in the Methodology section. Note that only the product of population and alignment tensor can be derived from the experimentally measured RDCs,[53] and the two unknown parameters cannot be separated unless additional information on either the population of each conformer in the ensemble or the relative sizes of the alignment tensors for each conformer is obtained from independent sources. In a previous study on α-methylene-γ-butyrolactone,[31] this problem was solved assuming that the magnitude of the alignment tensors was the same for the two conformers in the ensemble. This is not, however, a valid assumption for fibrosterol sulfate A since the degree of alignment induced by extended or collapsed structures is expected to be very different.

We solved this problem by determining the populations of the representative conformers of clusters from a MD simulation, so that in the fitting procedure only the tensor parameters had to be optimized. To obtain the populations with the best possible accuracy, we performed unrestrained MD simulations for each possible configuration of fibrosterol sulfate A in a DMSO solvent box using the GROMOS 53A6 force field, which is optimized for small molecules. 100 ns MD simulations at 298 K produced conformational ensembles that were used to interpret the NMR data. Figure 2a shows the ensemble averaged (r^{-6} averaging, see eq 1) distance violations with respect to the 44 experimental NOEs relevant for the analysis of the configuration of centers 22, 25, and 24′ (see Supporting Information, Table S3). Averaging was done over all recorded structures in steps of 5 ns. Few NOEs between the two steroid ring systems could be observed and integrated, due to strong chemical shift degeneracy of the pseudosymmetric bis-steroid. The only difference of the two steroid ring systems arises from the different configuration of C3 and C3′, which results in a significant deviation of the chemical shifts for the H2/H2′ and H3/H3′ pairs. Nevertheless, three long-range NOEs between the two steroid ring systems H2−H2′, H2−H4′e, and H3−H4′e could be unambiguously integrated and used in the calculation.

Table 1. $^3J_{HH}$ and $^3J_{CH}$ Couplings of the Linker of Fibrosterol Sulfate A in DMSO

atoms	$^3J_{HH}$ [Hz]	atoms	$^3J_{CH}$ [Hz]
H24′−H23′b	3.6	H24−C26	3.90
H24′−H23′a	11.9	H24−C27	4.31
H22′b−H23′a	11.8	H24−C24′	3.22
H22′a−H23′b	11.9		
H22′b−H23′b	4.8		
H22′a−H23′a	4.9		
H22′b−H20′	10.4		
H22′a−H20′	2.6		

They are especially important for defining the overall conformation of fibrosterol sulfate A since they validate the presence of the collapsed conformation; however, their weak intensities indicate that the collapsed conformation accounts for only part of the total population.

Averaged ^3J couplings were calculated over all recorded structures in steps of 5 ns shown in Figure 2b. The violations are the averaged deviations from three $^3J_{CH}$ and eight $^3J_{HH}$ couplings (Table 1) that are able to define the conformation in the flexible linker region. The time series of the averaged NOE and *J*-coupling violations were calculated for each possible configuration, which are depicted in different colors in Figure 2.

For all eight different configurations, we used an energy-minimized starting structure with a fully extended conformation. From Figure 2, one can see that the large NOE and ^3J coupling violations drop rapidly in the first 10 ns for most configurations because during this time collapsed conformations with contacts between the two steroid halves are populated, fulfilling the experimental NOEs and ^3J couplings better than the extended conformations. Some configurations like *RRS* and *SRR* need more time to populate such collapsed conformations. The NOE and ^3J coupling violations show a similar evolution for most configurations. Comparing the eight possible configurations, the lowest averaged NOE and ^3J coupling violations were obtained for configuration *SSS* after 15 ns simulation. To identify the populated conformers of fibrosterol sulfate A in solution, a conformational clustering analysis was performed as described in the Methodology section. Two structures are members in the same cluster if their mutual rmsd of all heavy atom positions is below 0.2 nm. For the conformational clustering, we chose only that part of the trajectory which has the lowest NOE and *J* coupling violations (Figure 2) so that the populations of the main clusters depend not only on the energy derived from the force field but also on the experimental NMR data. As an example, for configuration *SSS* the ensemble of the structures from the first 15 ns of the full trajectory, which has the lowest averaged NOE and ^3J coupling violations, was selected for the clustering. In the clustering analysis, we found that for each possible configuration only four clusters are significantly populated (Table 2).

Representative structures of the four most populated clusters for each configuration were manually selected and weighted with the determined populations, and their individual alignment tensors were determined by fitting against the 44 experimental $^1D_{CH}$ RDCs. This amounted to the determination of 4 (conformations) times 5 (components of the alignment tensor) = 20 unknowns fitted to the 44 knowns. The fit was accomplished by solving the linear equation system using the SVD method (eq 4). The correct configuration *SSS* was selected

Table 2. Multiple Alignment Tensor Fit: The Populations of the Main Conformations and *Q* Factor of Each Configuration Using the RDC Fitting to a Structure Ensemble[a]

	population				
configuration	conf. 1	conf. 2	conf. 3	conf. 4	*Q* factor
RSR	0.57	0.25	0.13	0.05	0.25
RRS	0.51	0.24	0.15	0.10	0.33
SSR	0.39	0.33	0.17	0.11	0.36
SRS	0.68	0.26	0.05	0.01	0.37
RRR	0.65	0.19	0.10	0.06	0.26
RSS	0.60	0.18	0.12	0.10	0.34
SRR	0.40	0.29	0.19	0.12	0.43
SSS	0.33	0.48	0.09	0.10	0.13

[a] The alignment tensors were determined individually from 44 $^1D_{CH}$ for each conformer using the SVD method. The populations were calculated using a conformational clustering analysis on a selected unrestrained MD trajectory based on NOEs and ^3J couplings.

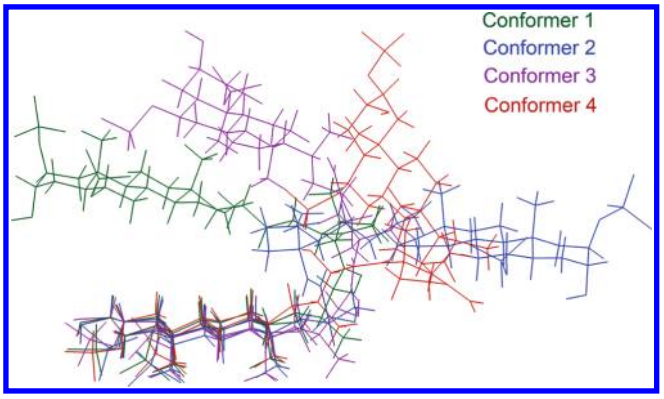


Figure 3. Multiple alignment tensor fit: the four populated conformations of the best-fitting configuration *SSS*, superimposed on steroid ABCD. The SVD fitting of the 44 $^1D_{CH}$ was performed against four conformers for *SSS* configuration. The four main conformers and the populations were obtained from a selected unrestrained MD simulation based on NOE and ^3J couplings. The alignment tensor was determined individually for each conformer.

based on the lowest *Q* factor (0.13) for experimental and back-calculated RDCs, which is summarized in Table 2 and Figure S2 in the Supporting Information. The other 7 configurations yielded significantly larger *Q* factors. The four representative cluster conformations of the configuration *SSS* (Figure 3) contain a collapsed conformation (conformer 1), an extended conformation (conformer 2), and two L-shaped conformations (conformers 3 and 4). The dihedrals of each rotatable bond in the linker reported for each conformation are listed in the Supporting Information Table S5.

Although in our multiple-tensor analysis we fitted 20 independent parameters, coming from four order tensors, to 44 experimental $^1D_{CH}$ values, some of them can be redundant due to nearly parallel CH vectors in the rigid steroid body. Therefore, the fitting process could be effectively underdetermined, making it necessary to perform an error analysis. Here, we performed a bootstrapping Monte Carlo analysis in which 250 sets of randomly generated RDC data were used as the input data for the fitting on the given structures. In each data set, a Gaussian

distributed noise with a standard deviation of 0.5 Hz was assumed. This standard deviation corresponds to a maximum RDC error of about 1.5 Hz for the 99% limits of confidence, which was estimated according to the signal-to-noise of the isotropic and anisotropic [^{1}H,^{13}C]-HSQC spectra. The error analysis was performed for each possible configuration, and each point in Figure 4 reveals the *Q* factor of one fitting procedure for one configuration. Since the distribution of *Q* values for configuration *SSS* over the expected experimental uncertainty ranges does not overlap with the *Q* distribution of other configurations,[15] the selection of the *SSS* configuration over the others is significant within the experimental error. Furthermore, to evaluate the errors on each derived tensor parameter and their correlations, we calculated the variance−covariance matrix of fitted parameters for the best fitting configuration *SSS* (Supporting Information, Table S6 and Table S7).

The orientations of the principle axes of the calculated alignment tensor for each conformer of the best-fitting configuration *SSS* are depicted in Figure 5, and the order parameters are shown in Table 3. These parameters give the information about the preferential orientation and the magnitude of the alignment of different conformers. The chosen alignment medium was a stretched poly(acrylonitrile) gel, which contains a nitrile group in

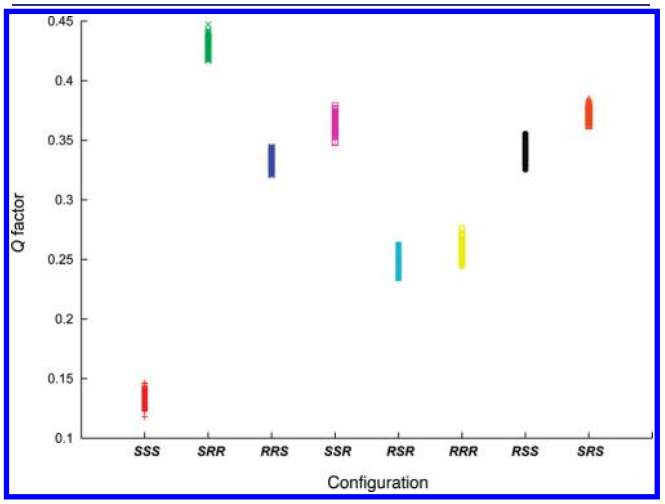


Figure 4. Multiple alignment tensor fit: the error bar of the *Q* factors of eight possible configurations obtained by bootstrapping error analysis using a standard deviation value of 0.5 Hz.

Table 3. Multiple Alignment Tensor Fit: Alignment Tensor Parameters of Each Conformer of the Best-Fitting Configuration *SSS*[a]

	conformer			
parameter	1	2	3	4
A_{xx}	-7.33×10^{-5}	4.58×10^{-5}	-4.44×10^{-4}	1.69×10^{-4}
A_{yy}	-2.88×10^{-4}	1.15×10^{-3}	-1.31×10^{-3}	1.62×10^{-3}
A_{zz}	3.61×10^{-4}	-1.19×10^{-3}	1.76×10^{-3}	-1.79×10^{-3}
A_a	5.42×10^{-4}	-1.79×10^{-3}	2.64×10^{-3}	-2.68×10^{-3}
A_r	2.15×10^{-4}	-1.10×10^{-3}	8.70×10^{-4}	-1.45×10^{-3}
α	24.8	−66.9	81.9	−10.2
β	47.5	5.1	−6.2	11.0
γ	−38.0	−65.0	−49.9	25.9

[a] A_{xx}, A_{yy}, and A_{zz} are the eigenvalues of the alignment tensor; A_a and A_r are axial and rhombic component of the alignment tensor, respectively; α, *β*, and *γ* are Euler angles in degrees for rotation of the alignment tensor into the principle axis frame. The reference frame is the *xyz* coordinate frame. The alignment tensor was determined individually from 44 $^{1}D_{CH}$ for each conformer using the SVD method.

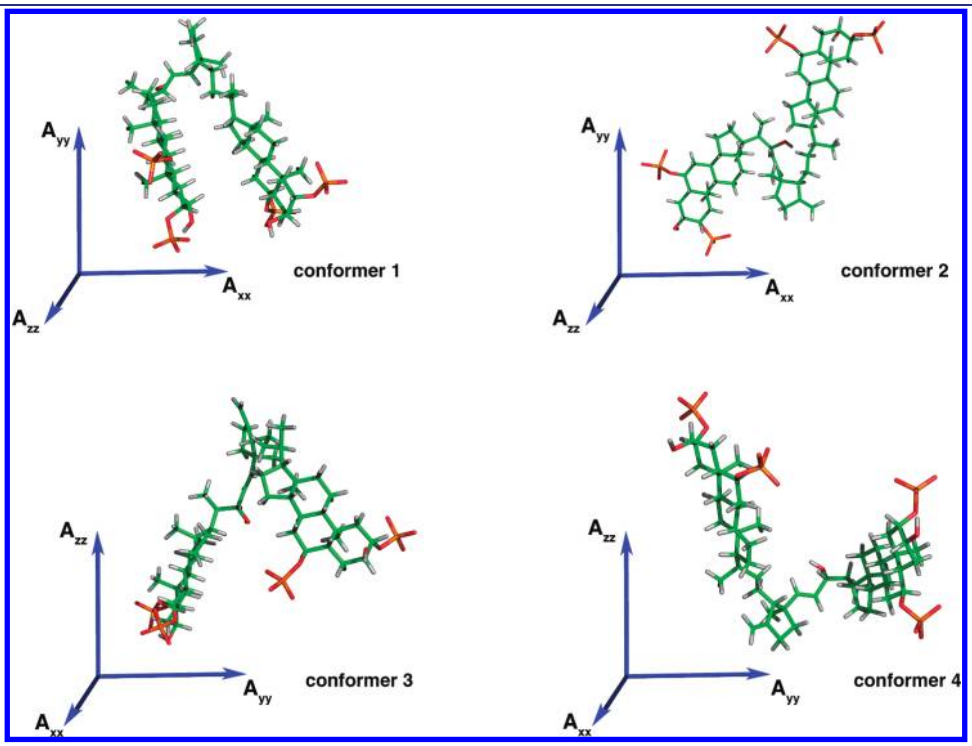


Figure 5. Multiple alignment tensor fit: orientations of four conformers of *SSS* configuration in the principal axis frame of the individual alignment tensors. The three blue arrows indicate the principal axes with components A_{xx}, A_{yy}, and A_{zz} of the molecular order tensors, respectively.

the main chain of the polymer. Since fibrosterol sulfate A contains four charged sodium sulfate groups, both steric and electrostatic interactions between the solute and the gel can contribute to the alignment mechanism. Therefore, prediction of the alignment tensor for each conformation only from the 3D structure[54−56] becomes particularly difficult.

To conclude this part of our work, we have presented a new approach using RDC data to establish the relative configuration and study the conformation of the molecule. We first determined the population from a MD trajectory evaluated for consistency with the experimental NOEs and ^3J values. A subsequent RDC-based analysis was then undertaken in which the alignment tensor was calculated individually for each populated conformer (schematic representation of the entire analysis procedure is shown in Supporting Information Figure S3).

Comparison of the RDC-Based Single-Tensor Analysis and Multiple-Tensor Analysis. The single alignment tensor approach has been used in many cases for solving stereochemical and conformational problems in small molecules, and it is often the only choice if a limited number of RDCs are available. In this study, we performed a single-tensor analysis as well, described in more detail in the Supporting Information, which led to the same configuration *SSS* of the three unknown stereocenters of fibrosterol sulfate A and similar conformer distribution. Although the single-tensor approach relies on an incorrect assumption and is therefore unreliable, a comparison between the two approaches is still worth to be conducted. We provide reasons why despite the inherent unreliability of the single-tensor analysis for flexible molecules it gives here still the correct answer despite the extensive motion of the fibrosterol sulfate A. A remarkable difference between the two approaches relies on the fact, that in the single-tensor approach, the populations of the conformations can be calculated directly using RDC data on a conformational ensemble obtained by a simple simulated annealing computation. By contrast, in the multiple-tensor fitting, since the magnitude of the order tensors and the populations are linearly dependent, information about the populations of the conformations or the relative sizes of the tensors are required. Here, we derived the populations from a MD simulation and could then determine the magnitudes of the tensors. Noteworthy, in the present case of fibrosterol sulfate A, although the populated conformers have a very different shape, the single-tensor fitting provides the same configuration of fibrosterol sulfate A as the multiple-tensor fitting approach. This can be attributed to the fact that in the single-tensor approach the major contribution to the dipolar coupling comes from the extended conformer (conformer 2: population: 48%) The extended conformer (conformer 2, Figure 3 and Figure S8, Supporting Information) in the single tensor and the multiple tensor approach looks very similar. According to the MD and multiple-tensor analysis, this conformer has the biggest population and a considerably larger magnitude of the tensor than the collapsed conformation.

CONCLUSION

In conclusion, in this work we have employed the multiple-tensor RDC analysis, combined with NOE and scalar coupling restraints, for studying the relative configuration and conformation of the complex and flexible natural product fibrosterol sulfate A. As this multiple-tensor approach requires previous knowledge of conformer populations, these were computed by means of explicit-solvent MD simulations. The three previously unknown stereocenters C22, C25, and C24′ were all determined to have *S* configuration, and the molecule exists in solution as a mixture of extended, collapsed, and L-shaped conformations. This finding demonstrates the power of RDCs in the determination of the configuration and conformation of flexible organic molecules, which in many cases cannot be achieved by using conventional isotropic NMR parameters in solution. Surprisingly, the single-tensor approximation yielded the same configuration, despite the large shape change between the different conformations, presumably by the fact that the major contribution to the RDCs originates from the extended conformation in both approaches. Nevertheless, fibrosterol sulfate A clearly belongs to the class of molecules which requires multiple tensor approach for reliable treatment.

ASSOCIATED CONTENT

Supporting Information. The assignments, experimental NOEs and RDCs data, validation of C20 and C20′ using RDC data, single-tensor analysis, and additional figures for the multiple-tensor analysis. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author
cigr@nmr.mpibpc.mpg.de; armando.navarro@uvigo.es

ACKNOWLEDGMENT

C.G. thanks the Max Planck Society, the Deutsche Forschungsgemeinschaft (GRK 782 and FOR 934), and the Fonds der Chemischen Industrie for financial support. CMI thanks NIH grant for financial support (CA 36633). A.N.V. thanks Xunta de Galicia-FEDER for financial support (Consellería de Educación 2009/071) and Ministerio de Ciencia e Innovación for a "Ramón y Cajal" contract and financial support (CTQ2007-65310). We thank Prof. Burkhard Luy for providing the PAN gel.

REFERENCES

(1) Nicolaou, K. C.; Snyder, S. A. *Angew. Chem., Int. Ed.* **2005**, *44*, 1012–1044.
(2) Chini, M. G.; Riccio, R.; Bifulco, G. *Magn. Reson. Chem.* **2008**, *46*, 962–968.
(3) Bifulco, G.; Dambruoso, P.; Gomez-Paloma, L.; Riccio, R. *Chem. Rev.* **2007**, *107*, 3744–3779.
(4) Saupe, A.; Englert, G. *Phys. Rev. Lett.* **1963**, *11*, 462.
(5) Kummerlowe, G.; Auernheimer, J.; Lendlein, A.; Luy, B. *J. Am. Chem. Soc.* **2007**, *129*, 6080–6081.
(6) Haberz, P.; Farjon, J.; Griesinger, C. *Angew. Chem., Int. Ed.* **2005**, *44*, 427–429.
(7) Freudenberger, J. C.; Spiteller, P.; Bauer, R.; Kessler, H.; Luy, B. *J. Am. Chem. Soc.* **2004**, *126*, 14690–14691.
(8) Thiele, C. M. *Concepts Magn. Reson., Part A* **2007**, *30A*, 65–80.
(9) Thiele, C. M. *Eur. J. Org. Chem.* **2008**, 5673–5685.
(10) Kummerlowe, G.; Luy, B. *TrAC* **2009**, *28*, 483–493.
(11) Enthart, A.; Freudenberger, J. C.; Furrer, J.; Kessler, H.; Luy, B. *J. Magn. Reson.* **2008**, *192*, 314–322.
(12) Feher, K.; Berger, S.; Kover, K. E. *J. Magn. Reson.* **2003**, *163*, 340–346.
(13) Yan, J. L.; Delaglio, F.; Kaerner, A.; Kline, A. D.; Mo, H. P.; Shapiro, M. J.; Smitka, T. A.; Stephenson, G. A.; Zartler, E. R. *J. Am. Chem. Soc.* **2004**, *126*, 5008–5017.

(14) Gil, R. R.; Gayathri, C.; Tsarevsky, N. V.; Matyjaszewski, K. *J. Org. Chem.* **2008**, *73*, 840–848.
(15) Garcia, M. E.; Pagola, S.; Navarro-Vazquez, A.; Phillips, D. D.; Gayathri, C.; Krakauer, H.; Stephens, P. W.; Nicotra, V. E.; Gil, R. R. *Angew Chem., Int. Ed.* **2009**, *48*, 5670–5674.
(16) Yan, J. L.; Kline, A. D.; Mo, H. P.; Shapiro, M. J.; Zartler, E. R. *J. Org. Chem.* **2003**, *68*, 1786–1795.
(17) Verdier, L.; Sakhaii, P.; Zweckstetter, M.; Griesinger, C. *J. Magn. Reson.* **2003**, *163*, 353–359.
(18) Thiele, C. M.; Berger, S. *Org. Lett.* **2003**, *5*, 705–708.
(19) Sanchez-Pedregal, V. M.; Santamaria-Fernandez, R.; Navarro-Vazquez, A. *Org. Lett.* **2009**, *11*, 1471–1474.
(20) Gayathri, C.; de la Fuente, M. C.; Luy, B.; Gil, R. R.; Navarro-Vazquez, A. *Chem. Commun.* **2010**, *46*, 5879–5881.
(21) Catalano, D.; Dibari, L.; Veracini, C. A.; Shilstone, G. N.; Zannoni, C. *J. Chem. Phys.* **1991**, *94*, 3928–3935.
(22) Emsley, J. W.; Luckhurst, G. R.; Stockley, C. P. *Proc. R. Soc. London A: Mater.* **1982**, *381*, 117–138.
(23) Stevensson, B.; Sandström, D.; Maliniak, A. *J. Chem. Phys.* **2003**, *119*, 2738–2746.
(24) Thaning, J.; Stevensson, B.; Östervall, J.; Naidoo, K. J.; Widmalm, G.; Maliniak, A. *J. Chem. Phys. B* **2008**, *112*, 8434–8436.
(25) Schuetz, A.; Junker, J.; Leonov, A.; Lange, O. F.; Molinski, T. F.; Griesinger, C. *J. Am. Chem. Soc.* **2007**, *129*, 15114–15115.
(26) Fares, C.; Hassfeld, J.; Menche, D.; Carlomagno, T. *Angew. Chem., Int. Ed.* **2008**, *47*, 3722–3726.
(27) Schuetz, A.; Murakami, T.; Takada, N.; Junker, J.; Hashimoto, M.; Griesinger, C. *Angew. Chem., Int. Ed.* **2008**, *47*, 2032–2034.
(28) Sun, H. d. A., E. J.; Reinscheid, U. M.; Carlos Dias, L.; Kleber Z. Andrade, C.; Oliveira Rocha, R.; Griesinger, C. *Eur. J. Chem.* **2011**, *17*, 1811–1817.
(29) Thiele, C. M.; Marx, A.; Berger, R.; Fischer, J.; Biel, M.; Giannis, A. *Angew. Chem., Int. Ed.* **2006**, *45*, 4455–4460.
(30) Trigo-Mouriño, P.; Santamaria-Fernandez, R.; Sanchez-Pedregal, V. M.; Navarro-Vazquez, A. *J. Org. Chem.* **2010**, *75*, 3101–3104.
(31) Thiele, C. M.; Schmidts, V.; Bottcher, B.; Louzao, I.; Berger, R.; Maliniak, A.; Stevensson, B. *Angew. Chem., Int. Ed.* **2009**, *48*, 6708–6712.
(32) Bernado, P.; Blanchard, L.; Timmins, P.; Marion, D.; Ruigrok, R. W. H.; Blackledge, M. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 17002–17007.
(33) Jha, A. K.; Colubri, A.; Freed, K. F.; Sosnick, T. R. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 13099–13104.
(34) Bernado, P.; Bertoncini, C. W.; Griesinger, C.; Zweckstetter, M.; Blackledge, M. *J. Am. Chem. Soc.* **2005**, *127*, 17968–17969.
(35) Mukrasch, M. D.; Markwick, P.; Biernat, J.; von Bergen, M.; Bernado, P.; Griesinger, C.; Mandelkow, E.; Zweckstetter, M.; Blackledge, M. *J. Am. Chem. Soc.* **2007**, *129*, 5235–5243.
(36) Whitson, E. L.; Bugni, T. S.; Chockalingam, P. S.; Concepcion, G. P.; Feng, X. D.; Jin, G. X.; Harper, M. K.; Mangalindan, G. C.; McDonald, L. A.; Ireland, C. M. *J. Org. Chem.* **2009**, *74*, 5902–5908.
(37) Guo, H.; Ma, Y. J.; Zhang, B. G.; Sun, B. C.; Niu, R. F.; Ying, G. G.; Zhang, N. *J. Leukocyte Biol.* **2009**, *85*, 911–918.
(38) Zhao, C. J.; Cai, M.; Zhang, Y.; Liu, Y.; Sun, R. H.; Zhang, N. *Anal. Biochem.* **2007**, *362*, 8–15.
(39) Cohen, E. E. W.; Lingen, M. W.; Zhu, B. M.; Zhu, H. Y.; Straza, M. W.; Pierce, C.; Martin, L. E.; Rosner, M. R. *Cancer Res.* **2006**, *66*, 6296–6303.
(40) Sun, R. H.; Gao, P.; Chen, L.; Ma, D. L.; Wang, J. M.; Oppenheim, J. J.; Zhang, N. *Cancer Res.* **2005**, *65*, 1433–1441.
(41) Mustafi, R.; Cerda, S.; Chumsangsri, A.; Fichera, A.; Bissonnette, M. *Mol. Cancer Res.* **2006**, *4*, 683–694.
(42) Sajan, M. P.; Standaert, M. L.; Nimal, S.; Varanasi, U.; Pastoor, T.; Mastorides, S.; Braun, U.; Leitges, M.; Farese, R. V. *J. Lipid Res.* **2009**, *50*, 1133–1145.
(43) Mueller, L. *J. Magn. Reson.* **1987**, *72*, 191–196.
(44) Williamson, R. T.; Marquez, B. L.; Gerwick, W. H.; Kover, K. E. *Magn. Reson. Chem.* **2000**, *38*, 265–273.
(45) Oostenbrink, C.; Villa, A.; Mark, A. E.; Van Gunsteren, W. F. *J. Comput. Chem.* **2004**, *25*, 1656–1676.
(46) Scott, W. R. P.; Hunenberger, P. H.; Tironi, I. G.; Mark, A. E.; Billeter, S. R.; Fennen, J.; Torda, A. E.; Huber, T.; Kruger, P.; van Gunsteren, W. F. *J. Phys. Chem. A* **1999**, *103*, 3596–3607.
(47) Tropp, J. *J. Chem. Phys.* **1980**, *72*, 6035–6043.
(48) Pardi, A.; Billeter, M.; Wuthrich, K. *J. Mol. Biol.* **1984**, *180*, 741–751.
(49) Tvaroska, I.; Hricovini, M.; Petrakova, E. *Carbohyd. Res.* **1989**, *189*, 359–362.
(50) Daura, X.; van Gunsteren, W. F.; Mark, A. E. *Proteins: Struct., Funct., Genet.* **1999**, *34*, 269–280.
(51) www.scipy.org.
(52) Cornilescu, G.; Marquardt, J. L.; Ottiger, M.; Bax, A. *J. Am. Chem. Soc.* **1998**, *120*, 6836–6837.
(53) Burnell, E. E.; Delange, C. A. *Chem. Phys. Lett.* **1980**, *76*, 268–272.
(54) Zweckstetter, M.; Bax, A. *J. Am. Chem. Soc.* **2000**, *122*, 3791–3792.
(55) Almond, A.; Axelsen, J. B. *J. Am. Chem. Soc.* **2002**, *124*, 9986–9987.
(56) Azurmendi, H. F.; Bush, C. A. *J. Am. Chem. Soc.* **2002**, *124*, 2426–2427.