Cross-Fitting of Residual Dipolar Couplings

Grit Kummerlöwe, Sebastian Schmitt and Burkhard Luy*

Technische Universität München, Department Chemie, Lehrstuhl II für Organische Chemie, Lichtenbergstraße 4, Garching, Germany

Abstract: Residual dipolar couplings (RDCs) are important NMR-parameters for the structure determination of organic molecules. In this article we describe how RDCs can be used to effectively transfer structural information by cross-fitting the anisotropic parameters of molecules with similar overall structure. Using the example of 5-α-cholestan-3-one and cholesterol, it is possible to distinguish diastereomers of the compounds by cross-fitting with transferred alignment tensors, even when strongly reduced subsets of RDCs are used. It is also demonstrated that RDCs can be used for direct cross-fitting even in flexible parts of the molecules that are sufficiently similar in structure and dynamic behavior. The cross-fitting approach as a general tool is discussed in details.

Keywords: NMR spectroscopy, residual dipolar couplings, alignment tensor, crossfitting, steroids, cholesterol, dynamics.

INTRODUCTION

Structure determination of organic molecules is one of the most challenging tasks for chemists and especially the determination of the stereo-configuration of natural or synthetic products is of high interest.

Among them steroids build a large class of molecules which are of special interest because of their variety of structures and biological activities [1]. Natural products with the steroid framework play an important role in medicinal chemistry and many strategies have been developed for their total synthesis or semisynthesis based on the modification of natural products [2]. In addition to the pharmaceutical and medicinal interests, steroid compounds are also quite suitable homochiral model compounds for chemo-, regio- and stereoselective investigations [3]. All this requires the reliable determination of constitution and configuration of the corresponding molecules.

While for X-ray-crystallography appropriate crystals are necessary, high resolution nuclear magnetic resonance spectroscopy (NMR) investigates molecules in solution. Based on classical NMR parameters like chemical shifts, *J*-couplings, and nuclear Overhauser enhancement (NOE) the structures of uncountable molecules have been solved. However, the structure determination by standard NMR parameters often fails because of a lack of information or because distant parts of the molecule can not be correlated.

In many cases a solution to the problem can be found by the use of residual dipolar couplings (RDCs) which contain information about internuclear distances and angles relative to an external reference. To measure anisotropic parameters like RDCs it is required to partially align the solute with the help of a so-called alignment medium. In recent years a variety of media for weak alignment of small to medium-sized organic molecules has been developed [4-6]. Especially mechanically stretched polymer gels are applicable to a wide range of solutes and solvents [7-18] and scaling of the alignment strength becomes very easy when combining polymer-based alignment media with an apparatus for arbitrary stretching [19-22].

Beside conformational studies of biologically active molecules [23-27] and e.g. the enantiomeric differentiation of small molecules in chiral alignment media, [15, 28-36] the central application for small organic molecules like natural or synthetic products is the determination of relative configurations of distant chiral and prochiral center [12, 17, 37-47]. Here, we describe the relative configurational analysis of molecules by cross-fitting of RDCs to RDCs from structurally similar molecules with known properties. After a brief theoretical introduction and a description of the RDC measurements performed on cholesterol and 5-α-cholestan-3-one as two test molecules with similar overall structure, cross-fitting based on transferring the alignment tensor or on direct comparison of RDCs will be demonstrated. In addition, the ability of distinguishing the diastereomeric 10-α-cholesterol from the measured cholesterol by cross-fitting with even strongly reduced RDC-datasets is studied in detail. As will be shown, the approach overcomes several limitations of classical RDC-analyses.

THEORY

In partially aligned samples residual dipolar couplings, D, add up to the corresponding heteronuclear scalar couplings, J and therefore spectra on oriented, anisotropic samples contain the sum J+2D. In corresponding spectra of isotropic samples, e.g. in solution, only the scalar coupling, J, is present and therefore the difference between couplings measured under anisotropic and isotropic conditions gives the desired RDCs.

*Address correspondence to this author at the Technische Universität München, Department Chemie, Lehrstuhl II für Organische Chemie, Lichtenbergstraße 4, Garching, Germany; Tel: ++49 (0)89/289-13275; Fax: ++49(0)89/289-13210; E-mail: Burkhard.Luy@ch.tum.de

1874-3838/10

As RDCs contain angular information relative to the external reference given by the magnetic field, their interpretation relies on the knowledge of the molecules orientation caused by the alignment medium. For rigid molecules this orientation is usually described by the so-called alignment tensor. Generally, the five independent components of the alignment tensor can be derived by mathematical methods like the singular value decomposition (SVD) as it is implemented in programs like PALES [48, 49] (-bestFit option). For this approach a set of 5 independent RDCs is required in which no two internuclear vectors for the RDCs are oriented parallel to each other and no more than three RDC vectors lie in a plane. Any further measured RDC directly contains valuable structural information given that the molecule can be considered as rigid.

As long as considerably more then 5 independent RDCs are given, this method can be used for the assignment of prochiral groups and/or for the configurational analysis of rigid molecules as various examples have shown (for reviews see [4, 6, 50]). Problems arise when the number of RDCs for fitting is not sufficient to unambiguously differentiate between different structural models. The situation gets significantly worse if flexible parts of a molecule must be taken into account [41-43, 51].

In steroids, for example, this lack of sufficient RDC-data might arise due to the small range in chemical shifts as most signals appear in the region between 15 ppm and 40 ppm in ^{13}C dimension and 0.5 ppm till 2.0 ppm in ^{1}H dimension respectively (see Fig. **1**). This does not only lead to signal overlap but might also cause immense strong coupling artifacts which prevent a reliable extraction of coupling constants [52]. In addition, a chair-like structure of six-membered rings in steroids reduces the number of independent RDCs as all axial CH-bonds are practically parallel. Therefore it might well happen that RDC-data measured on steroids are not sufficient for a desired configurational analysis.

Probably the most elegant and effective solution to the problem of a limited set of RDCs would be the prediction of alignment from first principles, which theoretically would allow to distinguish diastereomers by a single measured, decisive RDC. Many attempts have been made to predict the alignment of a molecule by considering steric [48] and electrostatic [53] interactions between the solute and the alignment medium, as it is for example done in PALES (-stPales option), [49] or by using inherent properties of the solute molecule like its tensor of gyration [54, 55]. The method works well for large biomacromolecules, but for small molecules, for which the effects of the fine-structure of the alignment medium and corresponding dynamics play a much more important role, it usually leads to wrong results with very few exceptions if the molecule of interest is void of functional groups [12].

A different approach, first used for the identification of absolute configuration *via* natural abundance deuterium NMR, [56] is the use of RDCs of a molecule of known, very similar structure for a cross-fit with the RDCs from the molecule of interest. As the alignment and therefore RDCs are strongly influenced by steric and charge interactions of the solute with the polymer of the alignment medium, the molecule used for cross-fitting should possess very similar overall shape and charge distribution and the dynamic behavior of the molecules should also match. If such a molecule is available, cross-fitting of RDCs should in principle allow the differentiation of diastereomers even in the case of otherwise insufficient RDC data.

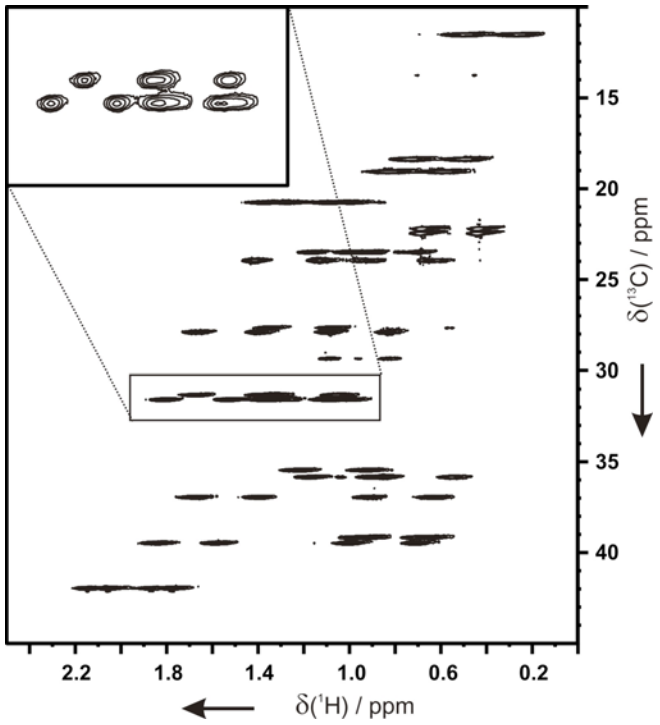

Fig. (1). Aliphatic region of the CLIP-HSQC spectrum [52] acquired on cholesterol in a stretched PDMS/$CDCl_3$ gel. Almost all steroid signals appear in a very narrow spectral range causing signal overlap and strong coupling artifacts. As an example of signal overlap the inset shows five doublets: C2-H2α, C2-H2β, C7-H7α, C7-H7β, and C8H8.

This approach can in principle be implemented in two ways: RDCs of the cross-fitting molecule can either be used to create an orientational model which is then applied to the solute of interest, or the couplings can be compared directly. If the molecule of interest contains large rigid parts and sufficient RDCs are available, simple singular value decomposition (SVD) [57] will allow to calculate an alignment tensor for the rigid parts which allows a detailed comparison of structural features in the solute. If the molecule is inherently flexible, more complex mean-field models should in principle be applicable like the AP-model, [58] the Chord-model, [59] the ME-model, [60] or combinations thereof [61-65].

If only the conservation of a specific part of the molecule shall be tested, a direct comparison of RDCs of this specific part should lead straightforwardly to the desired result without complicated fittings to a model.

In the following we test in how far the hypothesis of cross-fitting is applicable to real molecules. For this we apply the SVD-based alignment tensor model and the direct comparison of RDCs to two steroids with inherently high similarity. We also compare the result with steric RDC-prediction as implemented in PALES.

EXPERIMENTAL

Steroids have been selected as test molecules for our approach for several reasons. Firstly, steroids in general play a very important role in biology and medicine, where many drugs are based on the steroid scaffold. Secondly, they build a large group of molecules which differ in both configuration and substituents while they conserve the overall molecular shape determined by the four fused rings. The ring system contains also highly rigid parts as well as a relatively flexible five-membered ring, making steroids ideal systems to study the principle of cross-fitting.

We chose commercially available cholesterol and 5-α-cholestan-3-one as representative steroids. The structure of 5-α-cholestan-3-one is mainly determined by the keto-group and the fully saturated ring system (Fig. **2**). In contrast, cholesterol contains a double bond in ring B which bends the ring system and leads to a distinct variation in structure. Otherwise the two molecules have identical substituents and their overall shape can be considered to be quite similar.

The assignment of all NMR resonances (see Supporting Information) has been achieved by standard COSY, HSQC and ADEQUATE spectra performed on ~31mM samples of both steroids in $CDCl_3$. The isotropic samples were also used to measure scalar couplings. Anisotropic samples were prepared in a recently developed stretching apparatus: [21, 22] PDMS sticks (2.4 mm; 100 kGy) [12] were swollen inside the Kalrez® 8002 tube with a solution of 150 µl $CDCl_3$ and 16 mg of the corresponding steroid. The device was then stretched accordingly for obtaining equivalent alignment strength for both samples. The alignment strength was measured by the quadrupolar splitting of the deuterated solvent $CDCl_3$ and was adjusted to be 30.2 Hz for cholesterol and 30.8 Hz for 5-α-cholestan-3-one, respectively.

For the cross-fitting we measured heteronuclear $^1H,^{13}C$ one-bond couplings and homonuclear $^1H,^1H$ couplings between geminal protons in CH_2-goups, as they contain the most easily accessible RDCs.

For the measurement of $^1J_{CH}$ and $^1T_{CH} = {}^1J_{CH}+D_{CH}$ we used the so-called CLIP-HSQC and CLAP-HSQC pulse sequences as described by Enthart *et al*. [52] (D_{CH} in this case is defined as the dipolar splitting, which in the heteronuclear case corresponds to twice the dipolar coupling). The combination of this purely absorptive in-phase (CLIP) and anti-phase (CLAP) spectra allows the extraction of proton-carbon-couplings even in the case of overlapping CH_2-groups. Hard pulse versions of the CLIP/CLAP-HSQC were acquired on a Bruker DMX 500 MHz spectrometer with 4096 x 1536 data points in the ^1H and ^{13}C dimension and corresponding spectral widths of 10 and 70 ppm. For the solution samples 8 and for the stretched gel samples 20 scans per increment were recorded, leading to an overall experiment duration of 5 and 12.5 h for each experiment respectively. The $^1T_{CH}$ coupling constants were extracted by selecting a slice of the CLIP-HSQC at the appropriate carbon frequency and manually shifting a copy of the slice until the corresponding multiplet components were centered with respect to each other (see Fig. **3**). For overlapping signals of a CH_2-group the IPAP-approach was used as described in Ref [52].

$^2T_{HH}$ couplings were measured using the P.E.-HSQC pulse sequence as described by Tzvetkova *et al*. [66]. Resulting spectra allow the sign-sensitive extraction of geminal proton-proton-couplings. A hard pulse version of the experiment was used on a Bruker DMX 500 MHz spectrometer with 4096 x 3072 data points in the ^1H and ^{13}C dimension and corresponding spectral widths of 10 and 70 ppm. For the solution samples 12 and for the stretched gel samples 24 scans per increment were recorded, leading to an overall experiment duration of 16 and 32 h respectively. Fig. (**4**) shows the P.E.-HSQC region of two methylene groups of 5-α-cholestan-3-one for the isotropic and the anisotropic sample in comparison. The $^2T_{HH}$ coupling constants were extracted from the P.E.-HSQC spectra by selecting two slices at the appropriate carbon frequencies of one signal and manually shifting them until the corresponding multiplet components were centered with respect to each other as shown for the CLIP-HSQC in Fig. (**3**). The sign of the $^2T_{HH}$ is given by the tilt of the multiplet (see Fig. **4**).

It should be pointed out that special care was taken regarding the determination of experimental errors as line-broadening and strong coupling artifacts can have confound impact on the accuracy of determined coupling constants [52]. We therefore followed the procedure visualized in Fig. (**3**). After extraction of corresponding 1D-slices, α- and β-components of the multiplets were shifted relative to each other and the shift in Hz is taken as the coupling constant. For an error estimate, not the center of the signals was overlapped, but the left- and rightmost positions of overlap of the two multiplet components with special consideration of the flanks and feet of the signals of interest. We then took the average of the left- and rightmost shifts as the corresponding

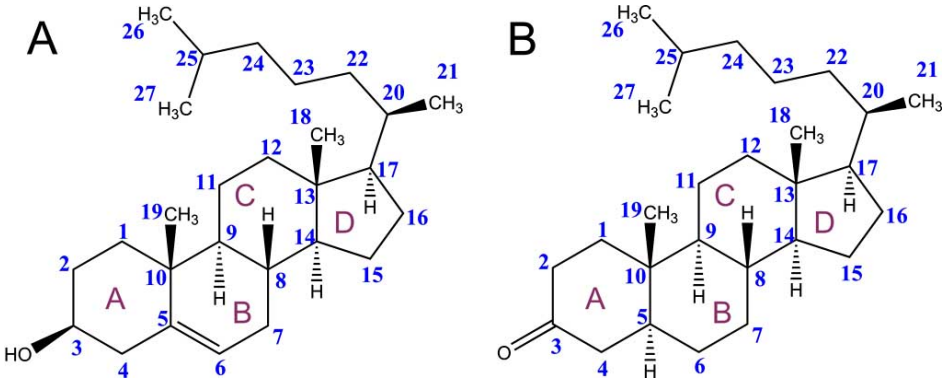

Fig. (2). Structure and numbering of cholesterol (**A**) and 5-α-cholestan-3-one (**B**).

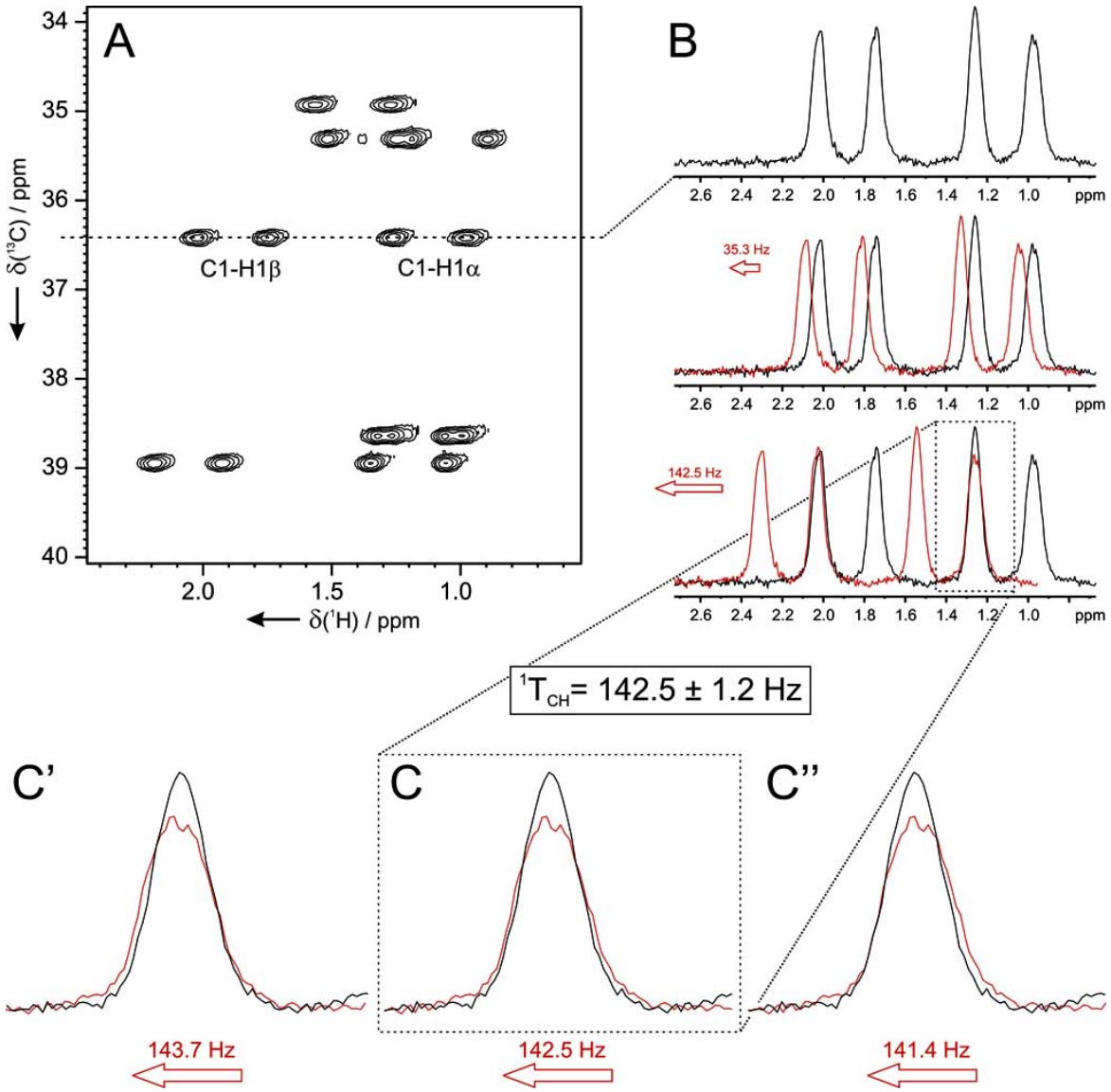

Fig. (3). Extraction of $^1T_{CH}$ couplings and the estimation of the individual maximum error of a specific coupling. Details of the CLIP-HSQC spectrum [52] of cholesterol in the stretched PDMS/$CDCl_3$ gel (**A**) with the two doublets of both protons attached to C1. A slice at the C1 carbon frequency was taken (**B**) and a copy of the slice (shown in red) is shifted relative to it to achieve maximum overlap of the two multiplet components as it is shown for the C1-H1α signal. While the overlap of the center of the multiplet components (**C**) reflects the coupling constant, the overlap at the right (**C'**) and left flank (**C''**) allow the estimation of the maximum individual error. For the example shown (C1-H1α) a coupling constant of $^1T_{CH}$ = 142.5±1.2 Hz was determined. (The coupling constant for C1-H1β was determined in analogy to be 137.4±1.0 Hz.).

coupling and half their difference as the individual maximum error estimate. The error estimates with this procedure are clearly larger than the conventional standard deviation, but they also ensure that theoretical couplings outside the specified limits will lead to a clear rejection of the model and overfitting of RDC data is avoided.

RDCs were calculated from the difference of couplings measured in the anisotropic sample and couplings measured in the isotropic sample (see Supporting Information for all couplings). In the case of cholesterol 16 $^1D_{CH}$ and 7 $^2D_{HH}$ couplings and for 5-α-cholestan-3-one 21 $^1D_{CH}$ and 8 $^2D_{HH}$ couplings could be extracted within the four rings. Additional 9 $^1D_{CH}$ couplings for the flexible side chain were extracted for each steroid but not used for SVD fitting as their orientational averaging most likely differs significantly from the more rigid part of the molecules. Missing couplings could not reliably be extracted due to signal overlap or strong coupling artifacts.

RESULTS AND DISCUSSION

Alignment Tensor Calculation

For the cross-fitting of RDCs between different molecules the assumption that both molecules possess the same alignment tensor must be fulfilled in a good approximation. Therefore we first calculated the alignment tensors of both steroids under investigation for comparison using the SVD approach. For this the -bestFit option of PALES [48, 49] was applied to fit measured RDCs to structural models of choles-

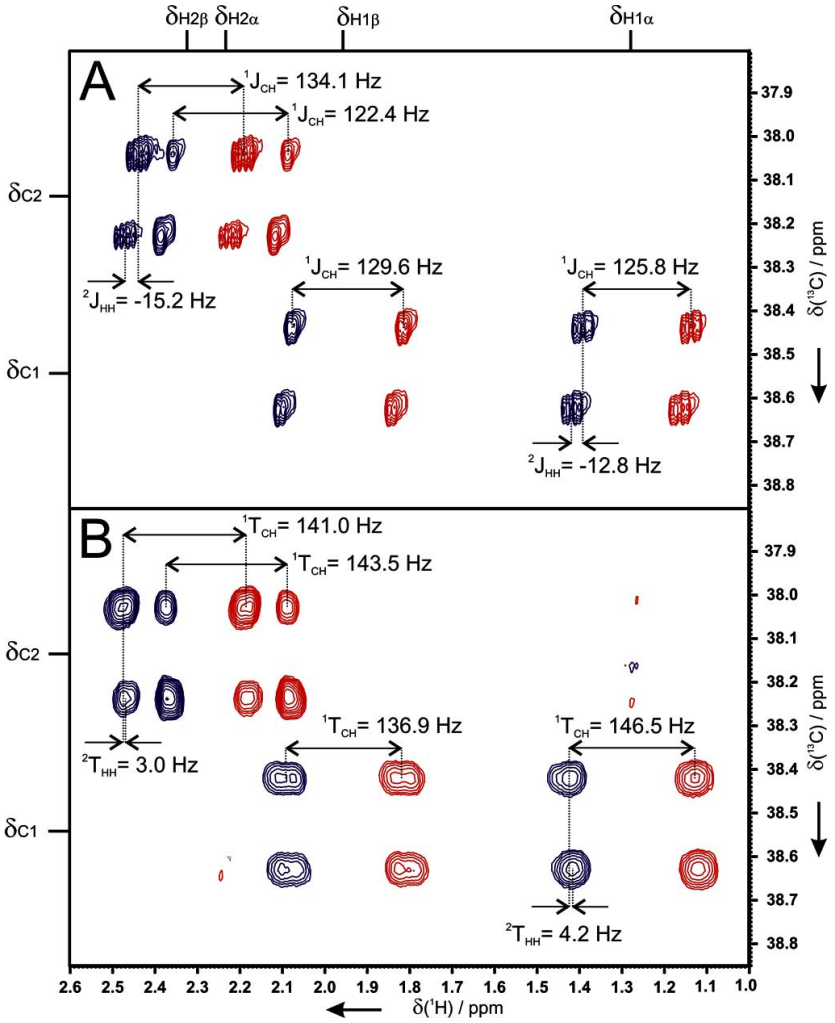

Fig. (4). Details of P.E.-HSQC spectra [66] acquired on 5-α-cholestan-3-one in a chloroform solution (**A**) and in a stretched PDMS/$CDCl_3$ gel (**B**), showing the multiplets of the methylene groups at C1 and C2. One-bond ^{13}C,^1H-couplings and geminal ^1H, ^1H-couplings are assigned in the spectra. Note that the sign-information of the homonuclear couplings is given by the tilt of the multiplet and that this tilt changes upon alignment in stretched PDMS gel as the negative $^2J_{HH}$-couplings (**A**) are compensated by the larger positive D_{HH}-couplings of both methylene groups in the aligned spectrum (**B**).

terol and 5-α-cholestan-3-one that were created and energy-minimized using the program SYBYL [67]. RDCs measured in the five-membered rings (D-ring) did not fit to back-calculated RDCs, which was expected as the flexible rings most likely experience a different orientational averaging and therefore a different average alignment compared to the rigid part of the steroids. As a result, the number of RDCs used for SVD calculations was reduced to 18 for cholesterol and 23 for 5-α-cholestan-3-one, respectively.

Since the prochiral assignment of methylene groups were not known *a priori*, we performed a permutation of all prochiral protons, resulting in 16 and 64 structural models for cholesterol and 5-α-cholestan-3-one, respectively. As was shown previously [68, 69] for small molecules, the best fit, as measured e.g. by the correlation factor R, the RMSD value, the Cornilescu Q, or n/χ^2, leads to the prochiral assignment for the entire molecule. Here, we consistently use the recently introduced n/χ^2 [47] which, from our point of view, gives the most reliable evaluation when identical experimental RDCs are compared to a variety of fits (see Supporting Information for details).

As expected for both molecules measured RDCs of the rigid parts fit well and corresponding alignment tensors could be calculated with high precision (see Supporting Information for couplings and alignment tensor parameters). A comparison of both calculated alignment tensors shows that they are very similar, as can also be seen in Fig. (**5**).

Cross-fitting of RDCs *via* the Alignment Tensor

As the alignment tensors of both steroids in the stretched PDMS/$CDCl_3$ gels are almost identical, a fit of measured RDCs for one of the molecules using the alignment tensor of the other molecule should also lead to a good agreement between measured and back-calculated RDC data. To test this we therefore cross-fitted both steroids to the alignment tensor derived for the other steroid using the PALES option for a user supplied order matrix (-saupe option) [48, 49].

As expected the quality of the corresponding fits drops (from n/χ^2 = 5.42 to 0.59 for cholesterol and from n/χ^2 = 1.13 to 0.26 for 5-α-cholestan-3-one) compared to the direct fit with the PALES -bestFit option, but measured and back-calculated RDCs are still in very good agreement and devia-

tions are small. For cholesterol fitted against the alignment tensor of 5-α-cholestan-3-one this can be seen in Fig. (**6C**) with the comparison of the corresponding direct fit in Fig. (**6A**). Detailed results on the fittings and analogue Figures for 5-α-cholestan-3-one fitted against the alignment tensor of cholesterol can be found in the Supporting Information.

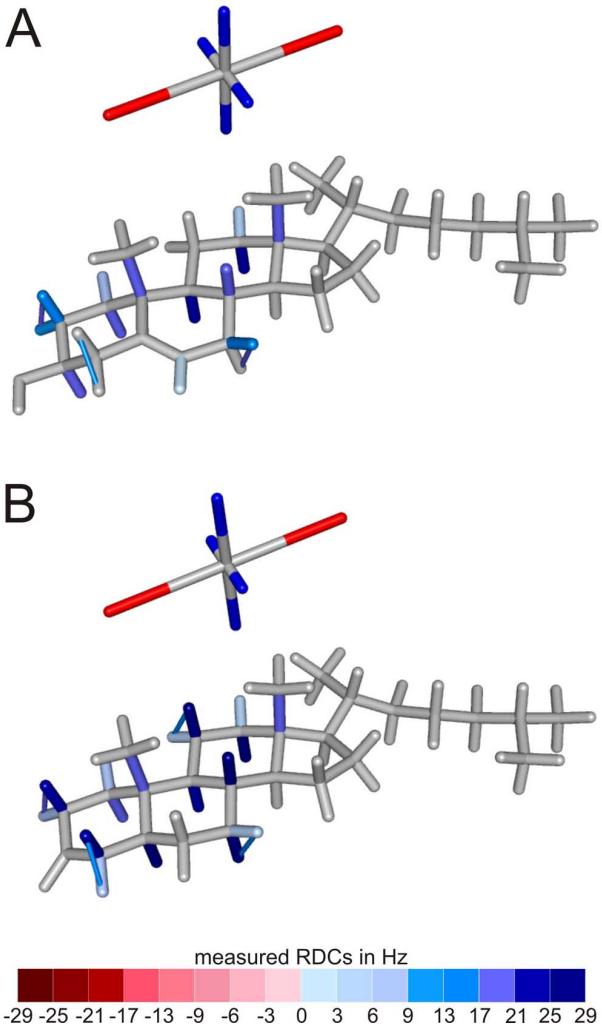

Fig. (5). Structures of cholesterol (**A**) and 5-α-cholestan-3-one (**B**) with color-coded bonds representing negative (red) and positive (blue) RDCs. The axes of the corresponding alignment tensors are drawn next to the structural models. Apparently the alignment tensors of both steroids are very similar but not fully identical.

Generally, this result can be seen as direct proof of principle for the cross-fitting approach based on the alignment tensor as the orientational model for rigid parts of molecules.

Differentiation of Potential Diastereomers

In a second step we wanted to test in how far the cross-fitting approach would be able to identify the correct relative configuration of an unknown molecule. We therefore cross-fitted measured RDCs against a potential diastereomer of one steroid using the alignment tensor of the other steroid to see if fitting results would be able to differentiate the correct diastereomer from a false one.

As a diastereomer of cholesterol we chose 10-α-cholesterol, which has an inverted chiral center at carbon atom C10. A structural model of 10-α-cholesterol has again been created and energy-minimized with the program SYBYL [67]. RDCs measured on cholesterol have been directly fitted against the structure of 10-α-cholesterol with the -bestFit option and cross-fitted with the fixed orientation of the alignment tensor derived for 5-α-cholestan-3-one (-saupe option) using PALES. The protons of all prochiral methylene groups were again permutated as their assignment was considered to be unknown and the best permutation was taken for comparison. Plots of RDCs measured on cholesterol and back-calculated for 10-α-cholesterol with the two methods can be seen in Fig. (**6B**/**D**) along with the deviation of the RDCs mapped to the structural model. Comparing the results with those for the cholesterol structure (Fig. **6A**/**C**), one can clearly differentiate the correct diastereomer cholesterol from the wrong 10-α-cholesterol.

In an analogue way we fitted RDCs measured on 5-α-cholestan-3-one against the structure of its potential diastereomer 5-β-cholestan-3-one with both methods and results for this fits can be found in the Supporting Information (Figs. **S4** and **S5**, Table **S5**). Also in this case the correct diastereomer could clearly be distinguished from the wrong one.

However, it should be mentioned that differentiation of diastereomers for the two steroids works with both cross-fitting and conventional RDC-based configurational analysis for which RDCs are directly fitted to the structural models of the diastereomers. It therefore can be concluded that if there is a sufficient number of RDCs for fitting available, a reliable differentiation between diastereomers is generally possible with either approach.

Differentiation of Diastereomers with Reduced RDC-Data

If the number of available RDCs is reduced, the situation can change significantly as the reliability of the alignment tensor of a direct SVD fit might be strongly decreased. For the simulation of the effect we therefore successively reduced the number of RDCs used for fitting. Out of the 18 measured RDCs of cholesterol, various subsets of 15, 12, 9, 8, 7 and 6 RDCs were generated by random selection of RDC-combinations. As with a decreasing number of RDCs within a subset the influence of the actual composition of the subset increases, we created the more subsets the less RDCs are contained within the subsets. In this way we were hoping to get an overall trend for what happens to the fitting quality with decreasing number of RDCs (See Supporting Information for actual combinations used).

In total we built 54 subsets of RDCs measured for cholesterol and fitted all of them to the structures of cholesterol and 10-α-cholesterol in order to differentiate the correct diastereomer. To compare the methods, we again used both the SVD-fit and the cross-fitting with the alignment tensor given by 5-α-cholestan-3-one for each subset and structure. Again all possible permutations for prochiral methylene groups were performed and the one with the best fitting result chosen.

A summary of the outcome is shown in Fig. (**7**) which compares n/χ^2 values for the direct SVD method (top) and the cross-fitting against the alignment tensor of 5-α-

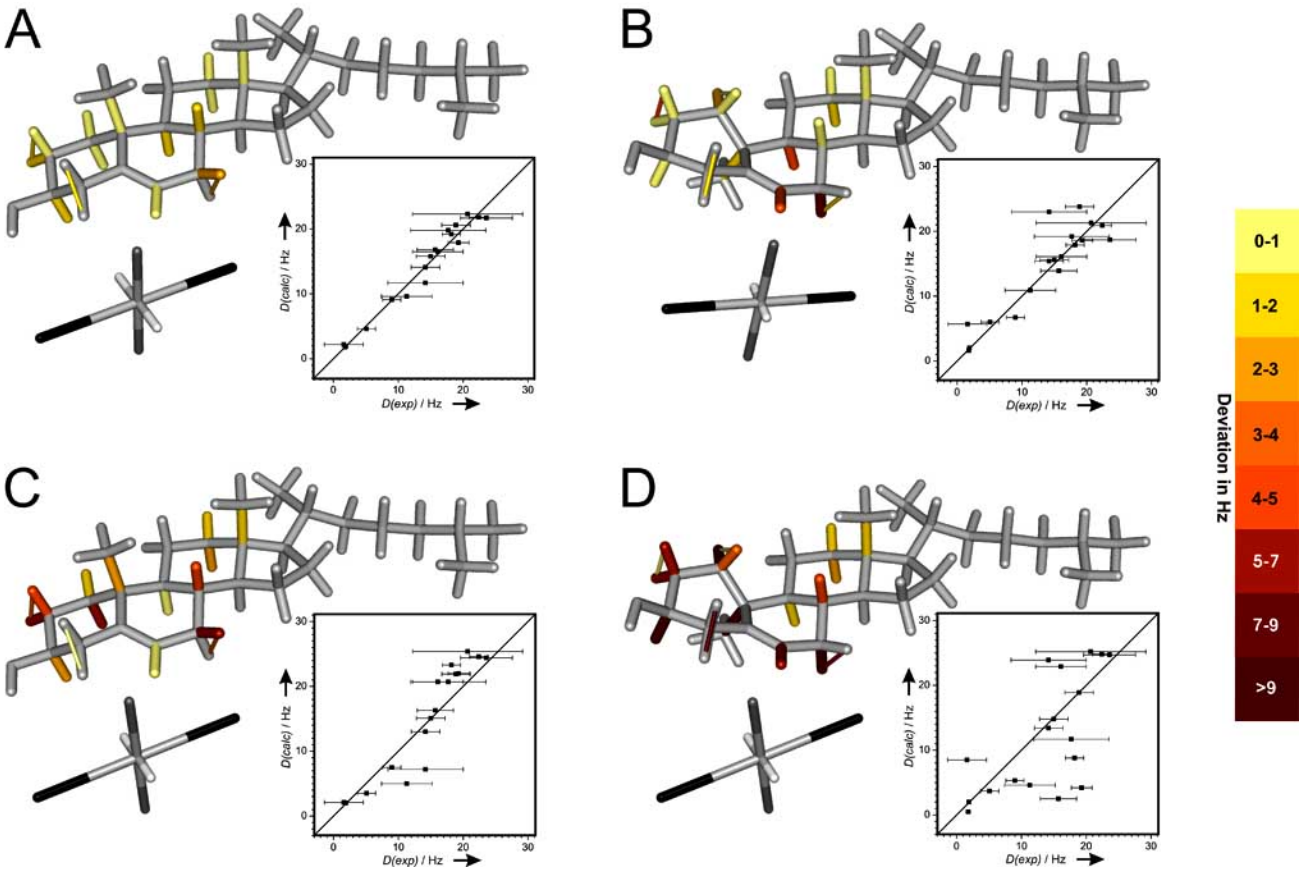

Fig. (6). (**A**, **B**) Plots of calculated vs. experimental RDCs and deviations of RDCs measured on cholesterol with directly back-calculated RDCs for the structures of cholesterol (**A**) and 10-α-cholesterol (**B**) using the -bestFit option of PALES (SVD-fit). (**C**,**D**) Identical plots with RDCs cross-fitted using the alignment tensor derived from 5-α-cholestan-3-one for cholesterol (**C**) and 10-α-cholesterol (**D**). Structures are shown with color-coded bonds denoting the deviation between measured and back-calculated RDCs for the different fits. The corresponding alignment tensors are visualized with their principal axis systems (black: A_{zz}; gray: A_{yy}; white: A_{xx}). For both methods the cholesterol structure (left) gives clearly the better fit.

cholestan-3-one (bottom) for both diastereomers and different RDC-subsets. One can clearly see the breakdown of the SVD approach with a decreasing number of RDCs: although n/χ^2 rises from an initial value on the order of 1 for all 18 RDCs to over 1000 if only 6 RDCs are used for the fitting routine (actually indicating a better fit to the structure), this rise is artificial since the lower number of restraints more easily allows a good fit to the five degrees of freedom of an alignment tensor. In fact less RDCs are more easily fitted to a structure but this does not mean that the fit reflects the real orientational behavior of the molecule. Considering that a fit with only 5 RDCs always gives a perfect match and therefore an infinitely high quality factor, it is no surprise that the quality factors increase with a decreasing number of RDCs, but it becomes also obvious that one can not trust a fit with only few RDCs. In this regard it is not surprising that also the differentiation of the two diastereomers fails with decreasing number of RDCs. For the randomly chosen subsets we used, the wrong diastereomer was favored for several combinations of 8, 7 and 6 RDCs. For one of the subsets with 9 RDCs and several combinations of 8, 7 and 6 RDCs a wrong permutation of the prochiral assignment for cholesterol was favored (marked with an * for each wrongly assigned prochiral center in Fig. **7**).

The cross-fitting approach, instead, leads to a completely different behavior for a decreasing number of RDCs. First of all, the quality of the fit does not depend on the number of RDCs and n/χ^2 values are generally on the same order close to 1 and vary only slightly due to the actual composition of the subset. While the average molecular orientation is allowed to vary freely in the SVD approach, the alignment tensor is fixed by the full set of 21 RDCs of 5-α-cholestan-3-one in the cross-fitting case and the reliability of the alignment tensor is independent of the available number of RDCs of the molecule of interest. As a consequence the full information content of the measured RDCs can be used for the differentiation of diastereomers and the correct diastereomer is favored for *all* randomly chosen subsets used for Fig. (**7**). In addition it should be noted, that for each subset the correct prochiral assignment of cholesterol is reflected in the fits.

The cross-fitting approach, in principle, will allow distinguishing the diastereomers even with less than 5 RDCs. However, its comparison with the SVD-fit makes only sense for more than 5 RDCs and it is also evident that with a decreasing number of RDCs the choice of RDCs becomes more and more important. As can be easily derived from Fig.

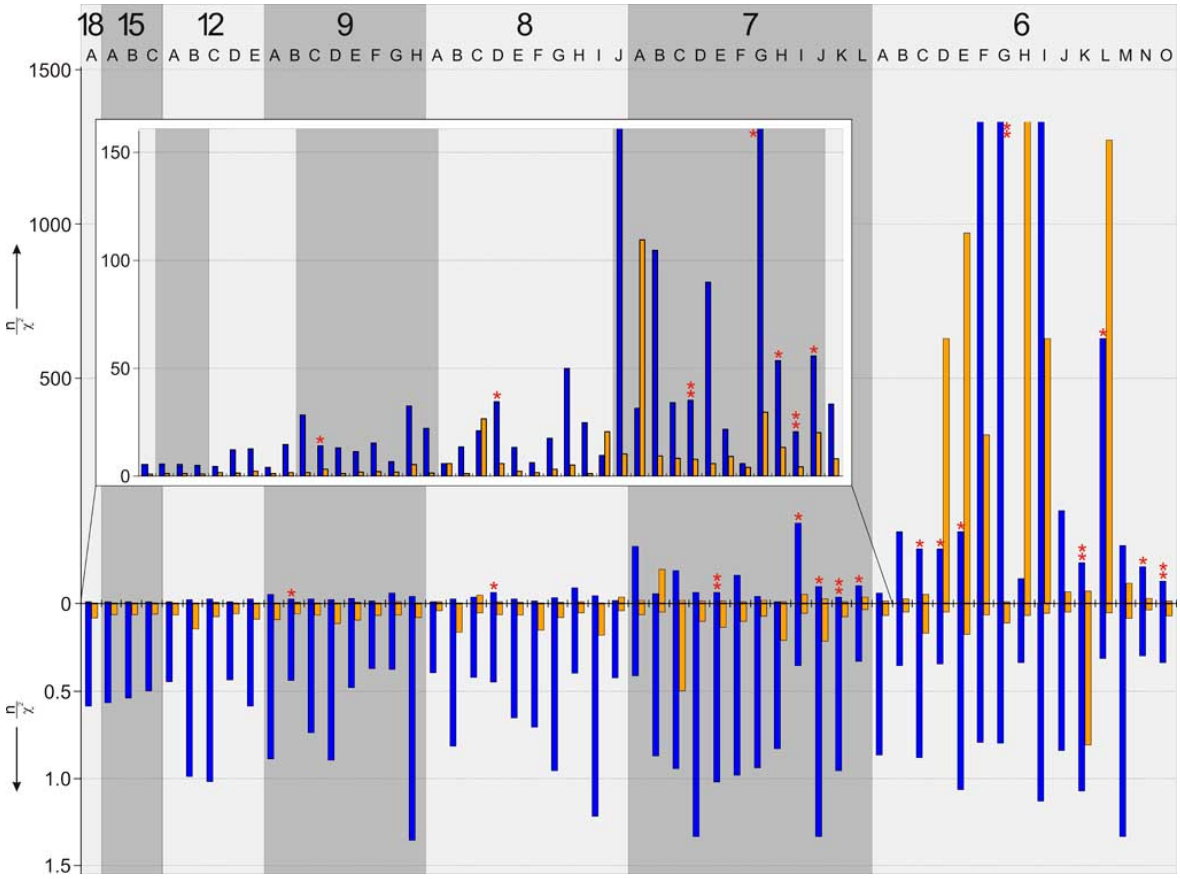

Fig. (7). Differentiation of diastereomers with reduced sets of RDCs. Compared are quality factors n/χ^2 obtained for the direct SVD-fitting method (top) and the cross-fitting approach (bottom) using subsets of RDCs measured on cholesterol in a stretched PDMS/$CDCl_3$ gel against the structures of cholesterol (blue) and 10-α-cholesterol (orange). RDC-data was reduced by building various subsets of the 18 measured RDCs with 15, 12, 9, 8, 7 and 6 RDCs. The individual subsets are marked with letters (15A-C, 12A-E, 9A-H, 8A-J, 7A-L, 6A-O) for which the corresponding RDC combinations can be found in Tables **S6** and **S7** of the Supporting Information. For each fit the prochiral assignment of CH_2-groups was permutated and only the fit with the best result is shown for each subset of RDCs. For each asterisk (*) one methylene group was obtained with incorrect prochiral assignment for the best permutation. Fewer RDCs for the SVD-fitting lead to highly unreliable results and strongly increased n/χ^2 values, while the cross-fitting approach maintains the correct differentiation of diastereomers and prochiral assignments even for sparse data.

(**6D**) by looking at the coupling deviations of individual RDCs, almost any RDC measured in the A-ring, for example, will be sufficient to differentiate the correct from the wrong diastereomer, while the whole set of RDCs of the C-ring will not lead to conclusive results.

Comparison with PALES RDC-prediction

As mentioned in the theory section, RDC prediction would be the most elegant way of deriving the alignment tensor of a given molecule. However, established methods only work for large proteins and sometimes small, rigid molecules without functional groups. In addition, the prediction of alignment is highly susceptible to flexible parts, whose influence on the induced orientation is generally overestimated.

The steroids investigated in our study both contain only a single functional group and their orientational order is most certainly dominated by the relatively rigid ring systems. They therefore represent almost ideal candidates for predicting the alignment tensor by PALES (with the -stPales option), but actual RDCs predicted by PALES for the whole steroid structure fit only reasonably to the measured RDCs.

We tried to improve the RDC prediction by shortening the flexible side chain attached to C17 (see Supporting Information). Indeed, predicted results became much better this way: for cholesterol best prediction results were achieved with the cholesterol fragment C1-C24 (cholesterol shortened by C25, C26 and C27 and adjacent protons) with a very good n/χ^2 value of 3.29 and for 5-α-cholestan-3-one prediction was best for the fragments C1-C23 (shortened by C24, C25, C26 and C27 and adjacent protons) and C1-C20 (shortened by the whole side chain) with n/χ^2 values of 0.92 and 0.87, respectively. However, the optimal length of the flexible side chain for the PALES prediction differs even in this case of two very similar molecules, showing how fast and unpredictable the method can get unreliable.

More problems with the PALES prediction arise when the interaction between solute and alignment medium is not purely steric but charges need to be considered. The steric

PALES prediction performed for sodium cholate, for example, does not at all reflect RDCs measured for sodium cholate in PAA/D_2O [39] (see Supporting Information). Considering how difficult and unreliable an accurate RDC-prediction with currently available methods is, cross-fitting appears to be the superior approach whenever applicable.

Cross-fitting by Direct Comparison of RDCs

The cross-fitting *via* an alignment tensor or another model describing the average orientational behavior allows the indirect comparison of RDCs measured even in differing parts of two similar molecules. However, quite often it is only necessary to transfer the assignment of e.g. prochiral groups or to identify signals in a flexible chain for which the alignment tensor approach is not applicable.

As has been shown previously for residual quadrupolar couplings (RQCs) by Ziani *et al*., [56] the anisotropic couplings in this case can directly be transferred between two molecules with very similar orientational behavior without the detour over an orientational model. A comparison between RDCs measured on cholesterol and 5-α-cholestan-3-one (Table **1**) leads to a unique assignment for practically all prochiral protons with the exception of C7-H7 close to the structural difference in ring B. Especially in the flexible side chain attached to C17 (atoms C20-C27) the match of RDCs measured for the two molecules is striking.

For more or less identical parts of two molecules that fulfill the conditions for cross-fitting, the direct comparison of anisotropic NMR parameters like RDCs therefore must be considered as an highly effective tool to obtain an assign-

Table 1. Comparison of $^1H,^{13}C$ one-bond RDCs Measured on Cholesterol and 5-α-Cholestan-3-one in Similarly Stretched PDMS/$CDCl_3$ gels. Individual Maximum Error Estimates are Propagated from the Individual Maximum Error Estimates of the Original Measurements of $^1J_{CH}$ and $^1J_{CH}+^1D_{CH}$ Couplings, Respectively (see also Tables S2 and S4 of the Supporting Information)

Group[a]	$^1D_{CH}$ (exp) cholesterol [Hz]	$^1D_{CH}$ (exp) 5-α-cholestan-3-one [Hz]
C19-H19	-6.6 ± 0.6	-7.2 ± 0.8
C18-H18	-7.0 ± 0.6	-7.2 ± 0.5
C16-H16a	0.3 ± 5.8	-1.9 ± 5.8
C16-H16b	5.1 ± 5.8	8.5 ± 5.4
C15-H15a	–[b]	11.6 ± 5.1
C15-H15b	14.0 ± 5.8	14.0 ± 5.1
C7-H7β	14.2 ± 5.8	4.1 ± 5.0
C7-H7α	–[b]	25.9 ± 1.5
C8-H8	20.7 ± 8.5	27.2 ± 1.0
C2-H2α	11.3 ± 3.9	6.9 ± 3.2
C2-H2β	16.1 ± 3.9	21.1 ± 3.2
C1-H1β	9.0 ± 1.4	7.3 ± 1.2
C1-H1α	18.2 ± 1.4	20.7 ± 1.7
C12-H12α	22.4 ± 1.4	27.4 ± 1.6
C12-H12β	5.1 ± 1.4	4.3 ± 1.1
C9-H9	23.6 ± 4.0	23.7 ± 1.1
C21-H21	-4.9 ± 0.6	-5.0 ± 0.7
C25-H25	10.8 ± 0.9	11.4 ± 0.8
C20-H20	23.3 ± 1.1	23.3 ± 1.0
C27-H27	-0.2 ± 0.4	-0.5 ± 0.4
C26-H26	0.2 ± 0.4	0.2 ± 0.4
C23-H23a	18.1 ± 5.1	20.5 ± 8.2
C23-H23b	5.2 ± 2.7	8.5 ± 8.5
C22-H22a	25.6 ± 6.4	26.7 ± 5.8
C22-H22b	8.7 ± 10.4	8.7 ± 5.1

[a] Methylene protons marked with a, b indicate that prochiral assignment is not available. Protons marked with α or β are assigned following standard steroid nomenclature.
[b] Couplings not extracted due to signal overlap or strong coupling artifacts.

ment or to independently verify a given assignment. Especially for molecules like the presented steroids with many signals within narrow chemical shift ranges this can be very useful as the potentially slightly better resolution for closely related compounds can be directly used by transferring corresponding RDCs.

Limitations and Potentials

In contrast to the prediction by PALES our approach does not need to consider structural details or charges of the alignment medium, as the orientational information is taken from experimental data. However, a number of conditions must be fulfilled: both the molecule of interest and the molecule used for cross-fitting RDCs must be rather similar in their overall shape and charge distribution, and RDCs must be measured for both molecules in the same alignment medium. Otherwise the cross-fitting of RDCs will fail.

The influence of the alignment medium and/or the influence of the molecules charge distribution might be seen when comparing experimental data measured on sodium cholate, another molecule with steroid scaffold, in PAA/D_2O [39] with our approach (see Supporting Information). The alignment of cholesterol and 5-α-cholestan-3-one in PDMS/$CDCl_3$ is fairly different compared to that of sodium cholate in PAA/D_2O and therefore the cross-fitting does not work. Notably the prediction of alignment with PALES in this case does also not work.

It should be mentioned that cross-fitting based on an alignment tensor requires rigid or at least partly rigid molecules, as only for those an alignment tensor is well-defined. The method, however, should be extendable to flexible molecules by the use of more general mean-field descriptions of orientational properties, which was not attempted here.

By directly comparing RDCs from a reference molecule with the molecule of interest, flexibility has no influence on the results as long as the molecules show similar dynamic behavior. The success of the method therefore is not limited by the cross-fitting approach in general, but by finding a reference molecule of known structure with sufficient similarity to the molecule of interest.

CONCLUSION

We could show that the alignment of the two steroids cholesterol and 5-α-cholestan-3-one in stretched PDMS/$CDCl_3$ gels is rather similar and that it is therefore possible to use the alignment tensor derived for one of the molecules to fit the RDCs of the other one. As has been shown in detail, cross-fitted RDCs then can be used to unambiguously distinguish diastereomers to the two measured compounds even in the case of massively reduced RDC datasets.

The approach of cross-fitting RDCs can also be applied without the detour of fitting an alignment tensor but by directly comparing RDCs of a known reference molecule with RDCs from the solute of interest. In this case, as long as the molecules are sufficiently similar, no restriction in terms of rigidity of the molecules apply and molecules can be compared even when no overall alignment tensor can be defined.

ACKNOWLEDGEMENTS

The authors thank the Center for Integrated Protein Science Munich (CIPSM), the Fonds der Chemischen Industrie (FCI), and the Deutsche Forschungsgemeinschaft Heisenberg grants LU835/2,3,4,7 and Forschergruppe FOR934) for financial support.

SUPPORTING INFORMATION

Supporting information is available on the publishers Web site along with the published article.

REFERENCES

[1] Hanson JR. Steroids: partial synthesis in medicinal chemistry. Nat Prod Rep 2007; 24: 1342-9.

[2] Nising CF, Bräse S. Highlights in steroid chemistry: total synthesis versus semisynthesis. Angew Chem Int Ed 2008; 47: 9389-91.

[3] Schönecker B, Lange C. Steroids as chiral model compounds for selective reactions with metals. J Organomet Chem 2006; 691: 2107-24.

[4] Thiele CM. Residual dipolar couplings (RDCs) in organic structure determination. Eur J Org Chem 2008; 5673-85.

[5] Kummerlöwe G, Luy B. Residual dipolar couplings as a tool in determining the structure of organic molecules. TrAC Trends Anal Chem 2009; 28: 483-93.

[6] Kummerlöwe G, Luy B. Residual dipolar couplings for the configurational and conformational analysis of organic molecules. Annu Rep NMR Spectrosc 2009; 68: 193-232.

[7] Tycko R, Blanco FJ, Ishii Y. Alignment of biopolymers in strained gels: A new way to create detectable dipole-dipole couplings in high-resolution biomolecular NMR. J Am Chem Soc 2000; 122: 9340-1.

[8] Sass HJ, Musco G, Stahl SJ, Wingfield PT, Grzesiek S. Solution NMR of proteins within polyacrylamide gels: Diffusional properties and residual alignment by mechanical stress or embedding of oriented purple membranes. J Biomol NMR 2000; 18: 303-9.

[9] Meier S, Häussinger D, Grzesiek S. Charged acrylamide copolymer gels as media for weak alignment. J Biomol NMR 2002; 24: 351-6.

[10] Cierpicki T, Bushweller JH. Charged gels as orienting media for measurement of residual dipolar couplings in soluble and integral membrane proteins. J Am Chem Soc 2004; 126: 16259-66.

[11] Luy B, Kobzar K, Kessler H. An easy and scalable method for the partial alignment of organic molecules for measuring residual dipolar couplings. Angew Chem Int Ed 2004; 43: 1092-4.

[12] Freudenberger JC, Spiteller P, Bauer R, Kessler H, Luy B. Stretched polydimethylsiloxane gels as NMR-alignment media for apolar and weakly polar organic solvents: ideal tool for measuring RDCs at low molecular concentrations. J Am Chem Soc 2004; 126: 14690-1.

[13] Freudenberger JC, Knör S, Kobzar K, *et al.* Stretched polyvinylacetate-gels as NMR-alignment media for the measurement of residual dipolar couplings in polar organic solvents. Angew Chem Int Ed 2005; 44: 423-6.

[14] Haberz P, Farjon J, Griesinger C. A DMSO-Compatible orienting medium: towards the investigation of the stereochemistry of natural products. Angew Chem Int Ed 2005; 44: 427-9.

[15] Kobzar K, Kessler H, Luy B. Streched gelatin gels as chiral alignment media for the discrimination of enantiomers by NMR spectroscopy. Angew Chem Int Ed 2005; 44: 3145-7.

[16] Kummerlöwe G, Auernheimer J, Lendlein A, Luy B. Stretched poly(acrylonitrile) as a scalable alignment medium for DMSO. J Am Chem Soc 2007; 129: 6080-1.

[17] Gil RR, Gayathri C, Tsarevsky NV, Matyjaszewski K. Stretched poly(methyl methacrylate) gel aligns small organic molecules in chloroform. stereochemical analysis and diastereotopic proton NMR assignment in ludartin using residual dipolar couplings and ^3J coupling constant analysis. J Org Chem 2008; 73: 840-8.

[18] Kummerlöwe G, Udaya Kiran M, Luy B. Covalently crosslinked gelatin allows chiral distinction at elevated temperatures and in DMSO. Chem Eur J 2009; 15: 12192-5.

[19] Kuchel PW, Chapman BE, Müller N, Bubb WA, Philp DJ, Torres AM. Apparatus for rapid adjustment of the degree of alignment of NMR samples in aqueous media: Verification with residual quadrupolar splittings in ^{23}Na and ^{133}Cs spectra. J Magn Reson 2006; 180: 256-65.

[20] Naumann C, Bubb WA, Chapman BE, Kuchel PW. Tunable-alignment chiral system based on gelatin for NMR spectroscopy. J Am Chem Soc 2007; 129: 5340-1.

[21] Kummerlöwe G, Halbach F, Laufer B, Luy B. Precise measurement of RDCs in water and DMSO based gels using a silicone rubber tube for tunable stretching. Open Spectrosc J 2008; 2: 29-33.

[22] Kummerlöwe G, McCord EF, Cheatham SF, Niss S, Schnell RW, Luy B. Tunable alignment for all polymer gel/solvent combinations for the measurement of anisotropic NMR parameters using a perfluorinated elastomer tube. 2010; (submitted).

[23] Bax A. Weak alignment offers new NMR opportunities to study protein structure and dynamics. Protein Sci 2003; 12: 1-16.

[24] Prestegard JH, Bougault CM, Kishore AI. Residual dipolar couplings in structure determination of biomolecules. Chem Rev 2004; 104: 3519- 40.

[25] Klages J, Neubauer C, Coles M, Kessler H, Luy B. Structure refinement of cyclosporin a in chloroform by using RDCs measured in a stretched PDMS-Gel. ChemBioChem 2005; 6: 1672-8.

[26] Reinscheid UM, Farjon J, Radzom M, *et al.* Effect of the solvent on the conformation of a depsipeptide: NMR-derived solution structure of hormaomycin in DMSO from residual dipolar couplings in a novel DMSO-compatible alignment medium. ChemBioChem 2006; 7: 287-96.

[27] Kiran MU, Sudhakar A, Klages J, Kummerlöwe G, Luy B, Jagadeesh B. RDC enhanced NMR spectroscopy in organic solvent media: the importance for the experimental determination of periodic hydrogen bonded secondary structures. J Am Chem Soc 2009; 131: 15590-1.

[28] Sarfati M, Lesot P, Merlet D, Courtieu J. Theoretical and experimental aspects of enantiomeric differentiation using natural abundance multinuclear NMR spectroscopy in chiral polypeptide liquid crystals. Chem Commun 2000; 2069-81.

[29] Farjon J, Merlet D, Lesot P, Courtieu J. Enantiomeric excess measurements in weakly oriented chiral liquid crystal solvents through 2D H-1 selective refocusing experiments. J Magn Reson 2002; 158: 169-72.

[30] Rivard M, Guillen F, Fiaud JC, Aroulanda C, Lesot P. Efficient enantiodiscrimination of chiral monophosphine oxides and boranes by phosphorus coupled C-13 NMR spectroscopy in the presence of chiral ordering agents. Tetrahedron: Asymmetry 2003; 14: 1141-52.

[31] Farjon J, Baltaze JP, Lesot P, Merlet D, Courtieu J. Heteronuclear selective refocusing 2D NMR experiments for the spectral analysis of enantiomers in chiral oriented solvents. Magn Reson Chem 2004; 42: 594-9.

[32] Lafon O, Lesot P, Rivard M, Chavarot M, Rose-Munch F, Rose E. Enantiomeric analysis of planar chiral (hexahapto-Arene) chromium tricarbonyl complexes using NMR in oriented solvents. Organometallics 2005; 24: 4021-8.

[33] Eliav U, Navon G. Collagen fibers as a chiral agent: a demonstration of stereochemistry effects. J Am Chem Soc 2006; 128: 15956-7.

[34] Beguin L, Courtieu J, Ziani L, Merlet D. Simplification of the ^{1}H NMR spectra of enantiomers dissolved in chiral liquid crystals, combining variable angle sample spinning and selective refocusing experiments. Magn Reson Chem 2006; 44: 1096-101.

[35] Manthati VL, Murthy ASK, Caijo F, *et al.* Enantioselective synthesis of methyl-5(R)-fluorohept-6-ynoate. Tetrahedron Asymmetry 2006; 17: 2306-10.

[36] Meddour A, Uziel J, Courtieu J, Jugé S. Enantiodifferentiation of acyclic phosphonium salts in chiral liquid crystalline solutions. Tetrahedron Asymmetry 2006; 17: 1424-9.

[37] Yan J, Kline AD, Mo H, Shapiro MJ, Zartler ER. A novel method for the determination of stereochemistry in six-membered chairlike rings using residual dipolar couplings. J Org Chem 2003; 68: 1786-95.

[38] Aroulanda C, Boucard V, Guibé F, Courtieu J, Merlet D. Weakly oriented liquid-crystal NMR solvents as a general tool to determine relative configurations. Chem Eur J 2003; 9: 4536-9.

[39] Mangoni A, Esposito V, Randazzo A. Configuration assignment in small organic molecules *via* residual dipolar couplings. Chem Commun 2003; 154-5.

[40] Yan JL, Delaglio F, Kaerner A, *et al.* Complete relative stereochemistry of multiple stereocenters using only residual dipolar couplings. J Am Chem Soc 2004; 126: 5008-17.

[41] Thiele CM, Marx A, Berger R, Fischer J, Biel M, Giannis A. Determination of the relative configuration of a five-membered lactone from residual dipolar couplings. Angew Chem Int Ed 2006; 45: 4455-60.

[42] Schuetz A, Junker J, Leonov A, Lange OF, Molinski TF, Griesinger C. Stereochemistry of sagittamide a from residual dipolar coupling enhanced NMR. J Am Chem Soc 2007; 129: 15114-5.

[43] Schuetz A, Murakami T, Takada N, Junker J, Hashimoto M, Griesinger C. RDC-Enhanced NMR spectroscopy in structure elucidation of sucro-neolambertellin. Angew Chem Int Ed 2008; 47: 2032-4.

[44] Farès C, Hassfeld J, Menche D, Carlomagno T. Simultaneous determination of the conformation and relative configuration of archazolide a by using nuclear overhauser effects, J couplings, and residual dipolar couplings. Angew Chem Int Ed 2008; 47: 3722-6.

[45] Kummerlöwe G, Knör S, Frank AO, Paululat T, Kessler H, Luy B. Deuterated polymer gels for measuring anisotropic NMR parameters with strongly reduced artefacts. Chem Commun 2008; 5722-4.

[46] Swarbrick JD, Ashton TD. NMR studies of dextromethorphan in both isotropic and anisotropic states. Chirality 2009; available online: DOI:10.1002/chir.20703.

[47] Intelmann D, Kummerlöwe G, Haseleu G, *et al.* Structures of storage-induced transformation products of the Beer's Bitter principle, revealed by sophisticated NMR and LC/MS techniques. Chem Eur J 2009; 15: 13047-58.

[48] Zweckstetter M, Bax A. Prediction of sterically induced alignment in a dilute liquid crystalline phase: Aid to protein structure determination by NMR. J Am Chem Soc 2000; 122: 3791-2.

[49] Zweckstetter M. NMR: prediction of molecular alignment from structure using the PALES software. Nat Protoc 2008; 3: 679-90.

[50] Kummerlöwe G, Luy B. Residual dipolar couplings as a tool in determining the structure of organic molecules. TrAC Trends Anal Chem 2009; 28: 483-93.

[51] Thiele CM, Schmidts V, Böttcher B, *et al.* On the treatment of conformational flexibility when using residual dipolar couplings for structure determination. Angew Chem Int Ed 2009; 48: 6708-12.

[52] Enthart A, Freudenberger JC, Furrer J, Kessler H, Luy B. The CLIP/CLAP-HSQC: pure absorptive spectra for the measurement of one-bond couplings. J Magn Reson 2008; 192: 314-22.

[53] Zweckstetter M, Bax A. Characterization of molecular alignment in aqueous suspensions of Pf1 bacteriophage. J Biomol NMR 2001; 20: 365-77.

[54] Azurmendi HF, Bush CA. Tracking alignment from the moment of inertia tensor (TRAMITE) of biomolecules in neutral dilute liquid crystal solutions. J Am Chem Soc 2002; 124: 2426-7.

[55] Wu B, Petersen M, Girard F, Tessari M, Wijmenga SS. Prediction of molecular alignment of nucleic acids in aligned media. J Biomol NMR 2006; 35: 103-15.

[56] Ziani L, Lesot P, Meddour A, Courtieu J. Empirical determination of the absolute configuration of small chiral molecules using natural abundance H-2 NMR in chiral liquid crystals. Chem Commun 2007; 4737-9.

[57] Losonczi JA, Andrec M, Fischer MWF, Prestegard JH. Order matrix analysis of residual dipolar couplings using singular value decomposition. J Magn Reson 1999; 138: 334-42.

[58] Emsley JW, Luckhurst GR, Stockley CP. A theory of orientational ordering in uniaxial liquid-crystals composed of molecules with alkyl chains. Proc R Soc London A 1982; 381: 117-38.

[59] Photinos DJ, Samulski ET, Toriumi H. Alkyl chains in a nematic field. 1. a treatment of conformer shape. J Phys Chem 1990; 94: 4688-94.

[60] Catalano D, Dibari L, Veracini CA, Shilstone GN, Zannoni C. A maximum-entropy analysis of the problem of the rotameric distribution for substituted biphenyls studied by H-1 nuclear-magnetic-resonance spectroscopy in nematic liquid-crystals. J Chem Phys 1991; 94: 3928-35.

[61] Stevensson B, Landersjö C, Widmalm G, Maliniak A. Conformational distribution function of a disaccharide in a liquid crystalline phase determined using NMR spectroscopy. J Am Chem Soc 2002; 124: 5946-7.

[62] Stevensson B, Sandström D, Maliniak A. Conformational distribution functions extracted from residual dipolar couplings: A hybrid model based on maximum entropy and molecular field theory. J Chem Phys 2003; 119: 2738-46.

[63] Thaning J, Stevensson B, Maliniak A. Molecular structure extracted from residual dipolar couplings: diphenylmethane dissolved in a nematic liquid crystal. J Chem Phys 2005; 123: 044-507.

[64] Landersjö C, Stevensson B, Eklund R, *et al.* Molecular conformations of a disaccharide investigated using NMR spectroscopy. J Biomol NMR 2006; 35: 89-101.

[65] Thaning J, Stevensson B, Ostervall J, Naidoo KJ, Widmalm G, Maliniak A. NMR studies of molecular conformations in alpha-cyclodextrin. J Phys Chem B 2008; 112: 8434-6.

[66] Tzvetkova P, Simova S, Luy B. P.E.HSQC: A simple experiment for simultaneous and sign-sensitive measurement of $^{1}J_{CH}$+D_{CH} and $^{2}J_{HH}$+D_{HH} couplings. J Magn Reson 2007; 186: 193-200.

[67] Sybyl®. Sybyl 7.2 program, Tripos, St. Louis, Missouri.

[68] Thiele CM, Berger S. Probing the diastereotopicity of methylene protons in strychnine using residual dipolar couplings. Org Lett 2003; 5: 705-8.

[69] Verdier L, Sakhaii P, Zweckstetter M, Griesinger C. Measurement of long range H,C couplings in natural products in orienting media: a tool for structure eluciation of natural products. J Magn Reson 2003; 163: 353-9.

Received: November 12, 2009 Revised: December 23, 2009 Accepted: December 25, 2009