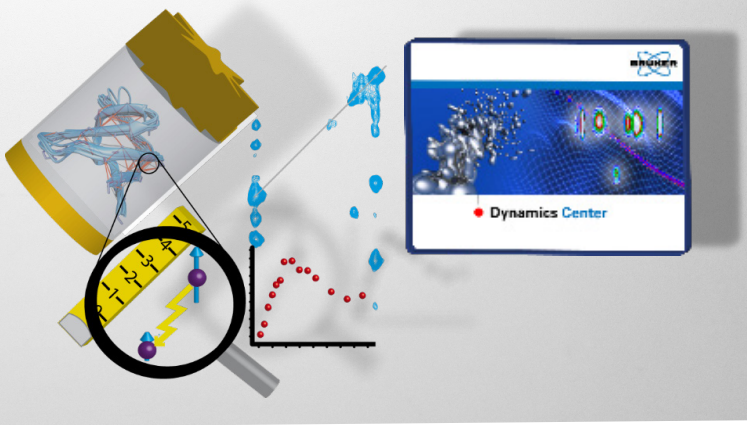

NMR

Exact Internuclear Distance Determination using the New Tool of the Dynamics Center “Exact Solid-State Distances”

Kristof Grohe and Peter Neidig

Innovation with Integrity

Determination of protein structure and dynamics by solid-state NMR plays a crucial role in structural biology, pharmacology, and biochemistry. To determine accurate protein structures, precise internuclear distance restraints are key. The use of very precise distance restraints is not only restricted to structure determination but also offers the analysis of directional motion. Therefore, we present a semi-automated, easy-to-use approach to determine those restraints by the Bruker Dynamics Center.

How ^{1}H-^{1}H distance determination is affected by different sources of error and how those are taken into account

In modern biological solid-state NMR, distance restraints are commonly derived from polarization-transfer ρ between protons. In solid-state NMR, the transfer efficiency is proportional to the internuclear distance r to the power of -3.

$$\rho \propto \frac{1}{r^3}$$

The transfer efficiency can be determined either by a single 3D (or 4D) spectrum using the cross-peak intensity or from a series of spectra with increasing ^{1}H-^{1}H mixing times, the buildup resembling a Bessel function. Polarization transfer can be achieved using transfer schemes like RFDR, BASS-SD or others. In **Figure 1 A,** the hhXH pulse sequence is shown, which will be available in the TopSpin library.

Determination of protein structure and dynamics by solid-state NMR plays a crucial role in structural biology, pharmacology, and biochemistry. To determine accurate protein structures, precise internuclear distance restraints are key. The use of very precise distance restraints is not only restricted to structure determination but also offers the analysis of directional motion. Therefore, we pr¬-esent a semi-automated, easy-to-use approach to determine those restraints by the Bruker Dynamics Center.

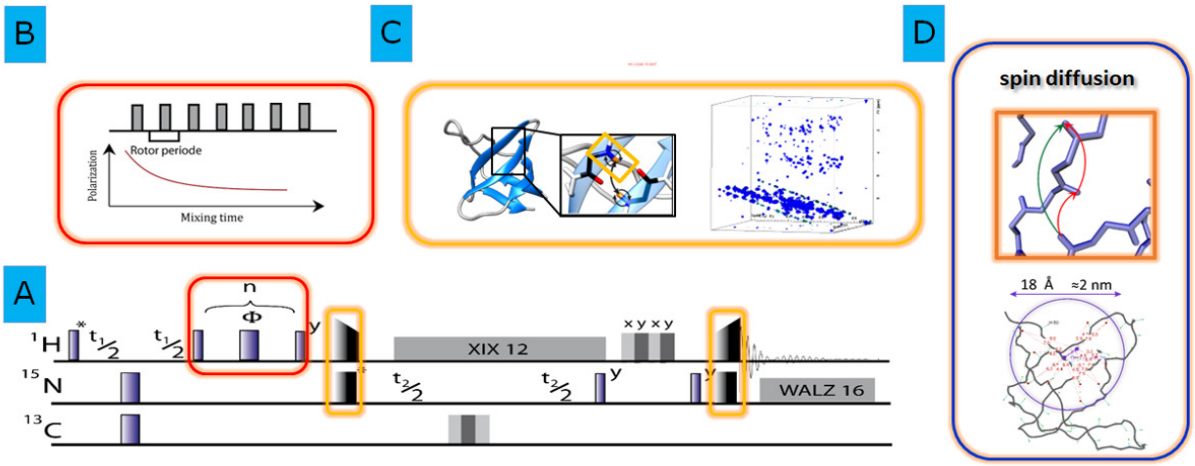

Figure 1: Depiction of errors influencing distance determination using an H-RFDR-hXH 3D spectrum. A) A simple version of the H-RFDR-hXH pulse sequence, where mixing can also be achieved using other schemes. B) Site-specific exponential polarization decay during mixing, originating from T_1 relaxation and pulse imperfection of the looped 180° pulse. C) Site-specific polarization transfer efficiencies between H and X. D) ^{1}H-^{1}H polarization transfer variation originating from spin diffusion and dipolar truncation.

While using a 3D H-hXH experiment with correlation to an X nucleus following an ^{1}H-^{1}H mixing block **(Figure 1 A)**, several sources of error are compromising the cross-peak intensity and therefore, the distance determination:

- During ^{1}H-^{1}H mixing, site-specific T_1 relaxation and – more importantly – 180° pulse-imperfections lead to an exponential decay as a function of mixing time **(Figure 1 B)**.
- Due to site-specific dynamics, the polarization transfer efficiency from ^{1}H to X differs for each pair of nuclei **(Figure 1 C)**.
- The polarization transfer between the protons is modulated by both: spin diffusion, increasing the polarization transfer rate via relay transfers involving other protons in the surrounding, and dipolar truncation, decreasing the transfer by polarization uptake from nearby protons.

All these sources of error are taken into account by the tool "exact Solid-State distances". The error originating from differential CP transfer efficiency is minimized by dividing the cross-peak intensity/volume by the diagonal peak intensity/volume. This leads to a cancellation of the error because polarization on cross and diagonal peak undergo the same polarization transfer steps while the polarization is at the same nucleus and hence experience an identical polarization transfer efficiency. This normalization simultaneously also reduces the error from site-specific polarization decay during mixing. This loss of intensity is partly derived from properties of the source proton and partly from those of the receiving proton. Therefore, the decay rate of the cross-peak is a mixture of the decay rates of both corresponding diagonal-peaks. Approximating the polarization to be at each nucleus for 50% of the time, an averaging of the buildup rates from the two complimentary cross-peaks normalized by its diagonal-peak takes the error into account. However, a decent error minimization is also achieved if only one pair of cross- and diagonal peak is unambiguously available. Normalization using the diagonal peak is automatically performed by the Dynamics Center.

The correction of the error originating from ^{1}H-^{1}H spin diffusion and dipolar truncation is more sophisticated. For the correction of spin-diffusion, all protons within a user-defined radius around both nuclei involved in the transfer are taken into account as relay transfer partners. (as determined from an initial structural model, which can be a preliminary structure from a standard structure calculation or an X-ray structure, dependent on the purpose of the exact solid-state distances, see below) The ratio between the rate for a theoretical direct polarization transfer and the cumulative rate derived from all theoretical polarization transfers, including the direct one and all transfers via spins within the described radii, is used as a correction factor. For more details see Söldner et al.[1]

$$F_{corr} = \frac{\rho_{IS}}{\rho_{IS} + \sum_K \min(\rho_{I\to K}, \rho_{K\to S})}$$

For this correction, an initial structural model is used by the Dynamics Center. This initial structure model can also be generated using restraints uncorrected for spin diffusion and dipolar truncation. As opposed for spin diffusion in terms of dipolar truncation, mostly those protons that are closest to the spin of polarization origin are relevant. The Dynamics Center determines only the closest ones, depending on the analyst's choice (e. g., the four nearest neighbors as in Söldner et al.[1]). For these protons, correction factors from a lookup table, which is included in the Dynamics Center, are chosen based on best geometric accordance (minimal Euclidean distance). The smallest one of the correction factors, i.e., the one for which the dipolar truncation is assumed to have the strongest effect[1], is then applied to the buildup rate determined by fitting with a Bessel function[4]. The lookup table represents theoretical values determined by SIMPSON simulations[2]. More details can be found in the publication of Söldner et al.[1]

First results obtained using the software

One of the first results, demonstrating the capabilities of the new toolset, is a structure determination of the SH3 domain of chicken α-spectrin, a 62-residue micro-crystalline protein, resulting in an accurate structure with root mean square deviation (RMSD) of 0.61 Å/1.45 Å (only backbone/all heavy atoms) and 0.45 Å/1.09 Å (only backbone/all heavy atoms, secondary-structural elements only) with respect to the X-ray structure **(Figure 2)**. For more details see Söldner et al.[1]

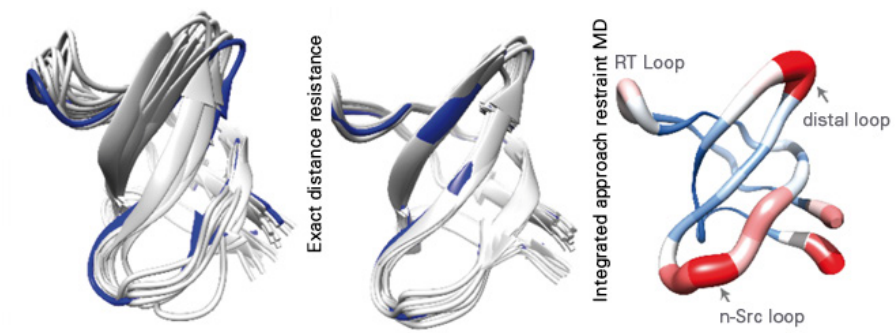

Figure 2: Excerpt from results investigated by Söldner et. al.[1] and Klein et. al. [3]. On the left side a structure ensemble (white) originating from restraints determined conventionally is shown. The structure in the middle shows a solid-state NMR structure using the exact distance restraints as provided by the presented software tool (PDB 8CF4). In blue, the X-ray structure 2NUZ is shown, which was used as an initial structural model. On the right, the root mean square fluctuation of an MD simulation that was restrained by the exact distance restraints is represented on the SH3 structure simultaneously as ribbon thickness and color coding. Thick ribbon and red color mean high fluctuation, whereas narrow ribbons and blue colors mean low fluctuation and high rigidity.

Maybe even more important, the exact distances were also used for determination of molecular motion in combination with MD simulation. In order to determine the dynamics, the over-proportional nature of the polarization transfer with respect to the internuclear distance is made use of I. e. The distance between 2 nuclei which are moving relatively to each other is usually determined being too short. This is because the buildup rises stronger when the protons get closer than it decreases when the protons move further away from each other **(Figure 3)**. This means the distance calculated from the determined buildup is shorter than the real average distance. During structure calculation those distance shortenings lead, in combination with chemical constraints, to violations. If now, the experimental error is smaller than the distance under-estimation, using multi-state structure determination this can be turned into dynamics information. Integrated approaches using MD simulations bury more complex background for dynamics determination.

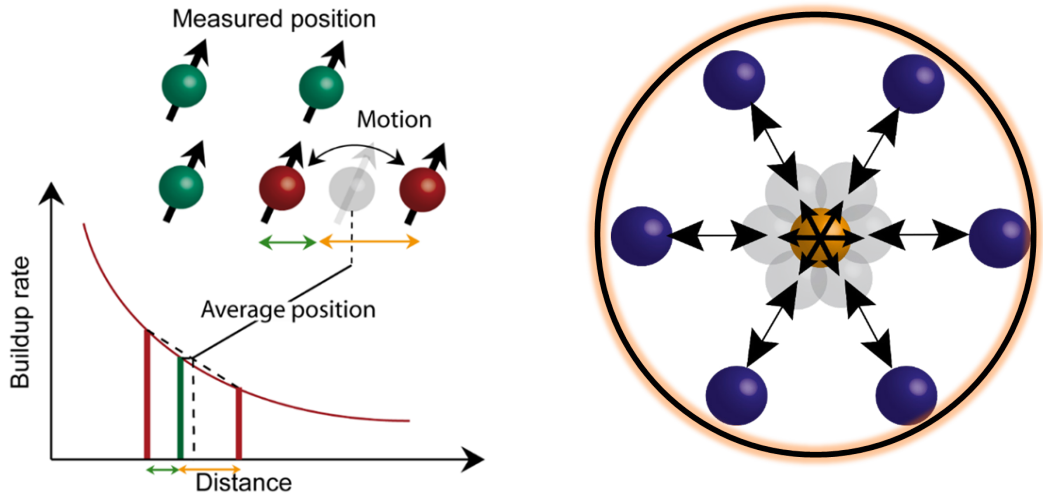

Figure 3: Depiction of underestimation of distances caused by motion. When motion decreases and increases the internuclear distance frequently (left side), the over-proportional relation between buildup and distance leads to an underestimation, if the distance is calculated from the buildup. During structure-calculation underestimated distances in combination with chemical restraints like bond length and angles leads to restrain-violations (right). Using ensemble averaged structure calculation or MD simulation this violation can be turned into dynamics information[4-7].

Workflow using "Exact SolidState Distances"

In the following, the major steps within the workflow of the new tool are briefly explained. For more detailed information read the manual accessible via the question mark in the top corner of the Dynamics Center main window **(Figure 4)**.

Figure 4: Top part of the Dynamics Center main window showing the button for accessing the manual in the red circle.

After clicking on "Sample", besides a description of sample properties, a PDB structure can be loaded, if it is available (see **Figure 5 A**). Since the analysis depends on the protonation state of the sample, the box "fully protonated" can be checked.

Clicking "Data" brings up a menu importing the NMR spectra either as a series of 2D or 3D. In order to create an assigned peak list, fully and semi-automated or manual peak-picking can be performed, with the opportunity of assigning it by mapping a shift list. The shift list can be imported in SPARKY format or directly as downloaded by the BMRB. It is also possible to import an assigned peak-list in SPARKY, XEASY, CCPNmr or TopSpin format. If a peak list exists in Topspin, this is directly used if the option "use peak list at spectrum" is chosen.

After loading the data, "Analysis" is performed as the next step. The buildups are fitted using a sum of six Bessel functions as also used for REDOR-fitting[4]. Further details are found in the Dynamics Center manual **(Figure 4)**.

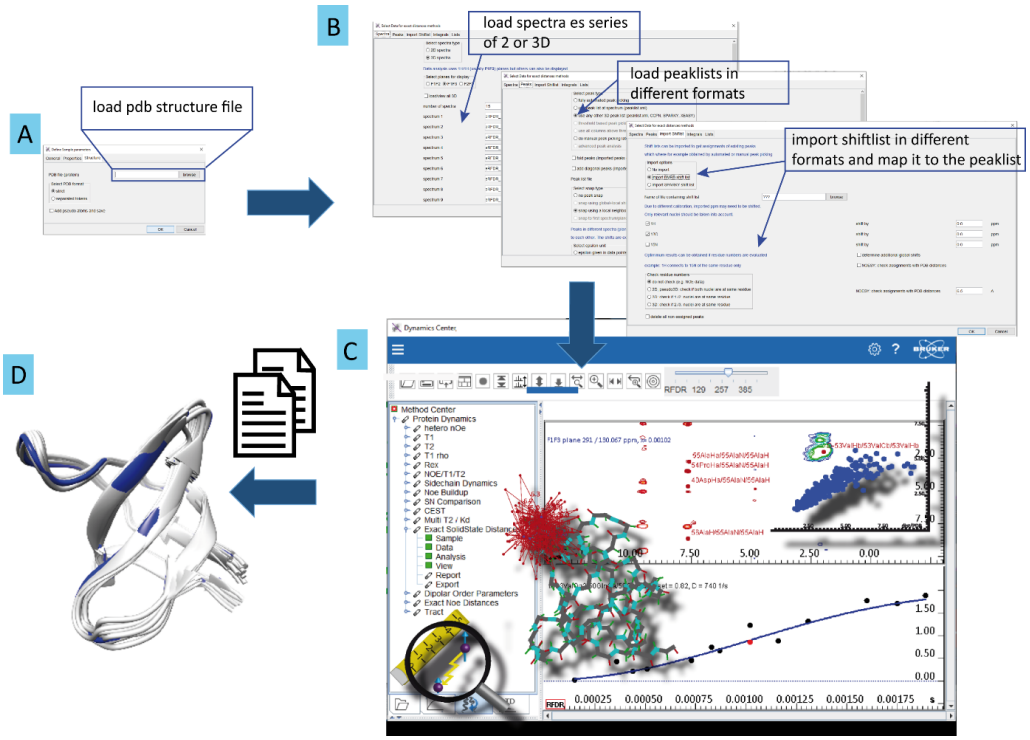

Figure 5: Workflow of the toolset "Exact SolidState Distances". In the step "Sample" a molecular structure file can be loaded if available **(A)**. The next step is "Data", where the spectra as well as a peak-list can be loaded **(B)**. Available peak list format types are SPARKY, CCPNmr, XEASY and Bruker. Instead of loading a peak list, automated peak-picking can be performed. This can be accompanied by mapping a shift-list in SPARKY or BMRB format onto the peak-list, including options of ^{1}H-^{1}H through space assignments (NOES-peak assignment) using the molecular structure. After the data is loaded, fitting is performed. The data can be exported as excel file or tab-separated text file. The export also includes a restraints file in CYANA, ARIA or CNS format which can be directly applied for structure calculation **(D)**. Prior to export, the data can be verified and calibrated interactively. Therefore, the graphical user interface includes a variety of functions like the molecular structure display, correlation plot and results table **(C)**.

After the fitting is done, generally the analysis is already complete. Now the results can be interactively investigated and if necessary, it can be corrected individually (peak by peak) **(Figure 5 C)**, with directly updated data analysis. Therefore, right clicking on the spectrum (not close to a peak) brings up a context menu including different functions. The most outstanding function is the "3D structure display". If a molecular structure in PDB format was loaded, correlations between nuclei with respect to individual cross-peaks are displayed on the structure. By right-click on the structure window, an additional context menu is accessible. Here, the function "show correlations plot" shows the determined distances as a function of the distances read out from the molecular input structure. This enables a judgement of data quality and interactive calibration on known distances. The function "results table" lists the determined distances together with peak assignments in an interactive table. Clicking at a row brings up the respective peak (jumping to the respective plain) in the display of the 3D spectrum. Simultaneously, a restraints file can be exported in CYANA, ARIA or CNS format. This file contains the exact distances and can directly be used for structure calculation or restrained MD simulation.

In the end all results can be exported as Excel file or tab separated text file. The detailed results file can also be used for refitting or further analysis using other programs, for instance based on MATLAB or Mathematica. A graphical export in pdf format is done after clicking report.

Conclusion

The tool "exact SolidState distances" of the Bruker Dynamics Center leads to reliable internuclear distance restraints in a fast and easy manner. The exported distance restraints can be directly used for structure determination by the program CYNA, ARIA and CNS or MD simulations restrained by NMR data. The results are precise protein structures and directional-dynamics information, which are valuable for structural biology.

Acknowledgement

We thank the whole research group of Prof. Rasmus Linser for the very fruitful collaboration resulting in a variety of research results and in the described software solution. Especially we thank Benedikt Söldner, Dr. Suresh Kumar Vasa, Dr. Alexander Klein and Prof. Rasmus Linser.

References

1. Söldner, B., et al., *Integrated Assessment of the Structure and Dynamics of Solid Proteins.* The Journal of Physical Chemistry Letters, 2023. **14**(7): p. 1725-1731.
2. Vosegaard, T., Z. Tošner, and N.C. Nielsen, *Numerical Simulations in Solid-State NMR with Simpson*, in *eMagRes*.
3. Klein, A., et al., *Unambiguous Side-Chain Assignments for Solid-State NMR Structure Elucidation of Nondeuterated Proteins via a Combined 5D/4D Side-Chain-to-Backbone Experiment.* The Journal of Physical Chemistry Letters, 2022. **13**(7): p. 1644-1651.
4. Mueller, K.T., et al., *The REDOR transform: direct calculation of internuclear couplings from dipolar-dephasing NMR* data. Chemical Physics Letters, 1995. **242**(6): p. 535-542.

Bruker BioSpin
info@bruker.com

bruker.com

Customer Support
https://www.bruker.com/en/services/support.html

Online information
bruker.com/