THE JOURNAL OF PHYSICAL CHEMISTRY A
A JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

pubs.acs.org/JPCA

Article

Minimizing Pervasive Artifacts in 4D Covariance Maps for Protein Side Chain NMR Assignments

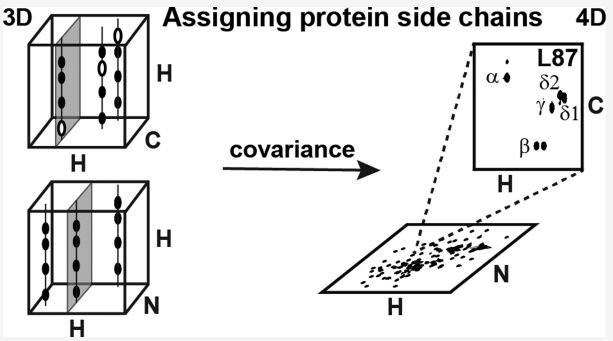
Aswani K. Kancherla, Kenneth A. Marincin, Subrata H. Mishra, and Dominique P. Frueh*

Cite This: *J. Phys. Chem. A* 2021, 125, 8313–8323

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Nuclear magnetic resonance (NMR) is a mainstay of biophysical studies that provides atomic level readouts to formulate molecular mechanisms. Side chains are particularly important to derive mechanisms involving proteins as they carry functional groups, but NMR studies of side chains are often limited by challenges in assigning their signals. Here, we designed a novel computational method that combines spectral derivatives and matrix square-rooting to produce reliable 4D covariance maps from routinely acquired 3D spectra and facilitates side chain resonance assignments. Thus, we generate two 4D maps from 3D-HcccoNH and 3D-HCcH-TOCSY spectra that each help overcome signal overlap or sensitivity losses. These 4D maps feature HC-HSQCs of individual side chains that can be paired to assigned backbone amide resonances of individual aliphatic signals, and both are obtained from a single modified covariance calculation. Further, we present 4D maps produced using conventional triple resonance experiments to easily assign asparagine side chain amide resonances. The 4D covariance maps encapsulate the lengthy manual pattern recognition used in traditional assignment methods and distill the information as correlations that can be easily visualized. We showcase the utility of the 4D covariance maps with a 10 kDa peptidyl carrier protein and a 52 kDa cyclization domain from a nonribosomal peptide synthetase.

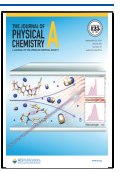


INTRODUCTION

Nuclear magnetic resonance (NMR) has emerged as a powerful biophysical tool to characterize structure,[1−4] dynamics,[5−8] kinetics,[9−12] and thermodynamics[13−16] of proteins at atomic resolution. Here, heteronuclear single quantum correlation spectroscopy (HSQC)[17] or transverse relaxation optimized spectroscopy (TROSY)[18] spectra provide the probes needed to study proteins at the molecular level, most often through correlations between backbone amide proton and nitrogen resonances or between methyl proton and carbon resonances. A fundamental prerequisite for obtaining such high-resolution information is to complete the resonance assignments of NMR spectra for the protein of interest. Thus, an incomplete assignment will bias the interpretation of data reporting on protein binding, allosteric responses, or structural changes, for example. Similarly, although amide and methyl moieties are useful probes to describe proteins owing to their global distribution, many biological processes rely on functional groups in side chains that are not represented by these moieties, and NMR experiments focusing specifically on these side chains must then be employed.[19−25] For example, side chains contain the functional groups involved in catalysis or ligand binding and their assignments become critical for a thorough investigation of enzyme mechanisms by NMR. Yet, many times, side chain resonance assignments are only partially completed due to experimental limitations and because significant time and effort is needed to complete the assignment. For example, side chain assignment completeness is not critical for determination of reliable structures, and hence, investigators often rightly employ a minimal effort strategy to obtain precise and accurate structural models without diminishing returns from unnecessary cumbersome assignment procedures. While completeness in side chain assignments may not be critical for structure determination, for functional studies, missing assignments may lead to erroneous molecular mechanisms. A current objective of our laboratory is to probe by NMR the molecular responses of specific side chains during complex enzymatic reactions involving multiple substrates, partner domains, and prosthetic groups, and the completeness of side chain resonance assignments will dictate the quality of our interpretation. To this aim, we seek to overcome spectral crowding in 3D side chain spectra without the need to collect time-consuming 4D experiments.

Here, we present a method to calculate reliable 4D covariance maps from typically acquired 3D spectra to facilitate protein side chain assignment. These 4D maps are obtained

Received: June 22, 2021
Revised: August 24, 2021
Published: September 12, 2021



https://doi.org/10.1021/acs.jpca.1c05507
J. Phys. Chem. A 2021, 125, 8313–8323

through a single computation from the covariance between 3D-HCcH-TOCSY[26] and itself and that between 3D-HCcH-TOCSY and 3D-HcccoNH.[27] Applying existing covariance methods[28−32] leads to unreliable maps contaminated with spurious correlations due to spectral overlap and frequency degeneracies in the crowded 3D spectra. We show that the combination of spectral derivatives[33,34] and matrix square-rooting[29] minimizes these otherwise pervasive artifacts to produce practical 4D maps. The added spectral dimension alleviates spectral crowding without additional experimental time and isolates the side chain signals of both aliphatic protons and carbons for a single residue within the protein. Further, we illustrate how covariance rescues the signal-to-noise of weak correlations when one of the spectra features stronger signals, a situation frequently encountered in larger proteins. Lastly, we show how conventional backbone spectra can be repurposed to generate 4D maps for obtaining asparagine amide side chain assignments.

The 4D covariance maps described here are mathematical translations of the cumbersome multistep visual pattern recognition process undertaken by a spectroscopist during assignment, and hence help overcome user bias or fatigue. These maps are generated using routine 3D experiments acquired for backbone and side chain assignments and do not need additional data acquisitions. We illustrate our method using spectra acquired on a protonated 10 kDa peptidyl carrier protein (PCP1), where spectral crowding is most severe, and an ILV methyl labeled[35] 52 kDa cyclization domain (Cy1), where sensitivity is most challenged, both from the yersiniabactin nonribosomal peptide synthetase.[36] By supplementing traditional approaches with our new method, we could assign 90% of PCP1 aliphatic proton and carbon resonances, 100% of PCP1 asparagine side chain amide resonances, and 98% of Cy1 methyl resonances.

METHODS

We have updated our covariance_4D.m MATLAB[50] script to incorporate the advances described in this publication. The new MATLAB script and supporting NMRPipe scripts can be downloaded at http://frueh.med.jhmi.edu/software-downloads/. The script was designed to be compatible with MATLAB 8.3+ (2014+) and the freeware GNU Octave 4.0+,[37] and we have tested it most recently with the 2017 edition of MATLAB.

PCP1 Sample Preparation. All NMR data were collected on a protonated, ^{15}N- and ^{13}C-labeled sample of the 10 kDa protein PCP1 harboring a phosphopantetheine group loaded with cysteine. The sample was prepared as described in ref 38. Briefly, apo PCP1 is expressed as a His_6-GB1 fusion protein with a Tobacco Etch Virus (TEV) cleavage site in minimal media containing 1 g/L $^{15}NH_4Cl$ and 2 g/L ^{13}C-glucose. Following cleavage, the protein is purified through reverse affinity and size-exclusion chromatography in phosphate buffer. The sample was then modified with an unlabeled, non-hydrolyzable mimic of cysteine-linked phosphopantetheine with a modified one-pot chemoenzymatic protocol[39,40] and further purified by size-exclusion chromatography.

PCP1 Data Collection. All side chain and backbone NMR spectra were collected at 25 °C on a 600 MHz Bruker Avance III spectrometer equipped with a QCI cryoprobe. The protein sample concentration was 314 μM in phosphate buffer containing 20 mM sodium phosphate (pH 6.59 at 22 °C), 150 mM NaCl, 1 mM ethylenediaminetetraacetic acid (EDTA), 2 mM tris(2-carboxyethyl)phosphine (TCEP), 10% D_2O and 200 μM sodium trimethylsilylpropanesulfonate (DSS) for internal referencing.

The 3D-HcccoNH spectrum was collected using uniform sampling with the following parameters: 32 scans and a recycling delay of 1 s; direct ^{1}H dimension (1024 complex points, 16.0192 ppm spectral width, and carrier at 4.698 ppm), indirect ^{15}N dimension (20 complex points, 31.25 ppm spectral width, and carrier at 117.0 ppm), and indirect ^{1}H dimension (50 complex points, 6.4 ppm spectral width, and carrier at 4.698 ppm); 23.022 ms total correlation spectroscopy (TOCSY) mixing time.

The 3D-HCcH-TOCSY spectrum was collected with uniform sampling using the following parameters: 8 scans and a recycling delay of 1 s; direct ^{1}H dimension (1024 complex points, 16.0192 ppm spectral width, and carrier at 4.698 ppm), indirect ^{13}C dimension (32 complex points, 43.0 ppm spectral width, and carrier at 24.2 ppm), and indirect ^{1}H dimension (50 complex points, 6.4 ppm spectral width, and carrier at 4.698 ppm); 43.464 ms TOCSY mixing time.

All four backbone 3D experiments, viz., 3D-HNCA,[41] 3D-HNcoCA,[42] 3D-HNcaCB,[43] and 3D-HNcocaCB[44] were acquired with nonuniform sampling and spectral widths of 16.0192 ppm (^{1}H) and 31.25 ppm (^{15}N), centered at 4.699 (^{1}H) and 117.0 ppm (^{15}N). 3D-HNCA and 3D-HNcoCA were acquired with spectral widths of 30.0 ppm and carriers at 52.5 ppm in the ^{13}C dimensions. 3D-HNcaCB and 3D-HNcocaCB were acquired with spectral widths of 60.0 ppm and a carrier frequency at 42.0 ppm in the ^{13}C dimensions. The 3D-HNCA and 3D-HNcoCA experiments were collected with 16 scans, a 1 s recycling delay, and nonuniform sampling using 10% sampling of 64(^{15}N) × 100(^{13}C) complex points. The 3D-HNcaCB and 3D-HNcocaCB experiments were collected with 16 scans, a 1 s recycling delay, and nonuniform sampling using a 10% sampling of 64(^{15}N) × 150(^{13}C) complex points. All NUS lists were generated using PoissonGap.[45]

Cy1 Sample Preparation. All NMR data for the 52 kDa Cy1 protein were acquired on ILV samples: (^{1}H,^{13}C) methyl labeled Leu, Val, and Ile (δ1 position only) side chains, in an otherwise uniform ^{2}H, ^{15}N, and ^{13}C labeling. Cloning of this construct,[46] and details of expression and purification of this sample[47] have been described previously.

Cy1 Data Collection. All Cy1 NMR experiments were acquired at 25 °C in phosphate buffer: 20 mM sodium phosphate pH 7.0 at 22 °C, 10 mM NaCl, 1 mM EDTA, 5 mM DTT, and 5% D_2O. The 3D-HCcH-TOCSY data set was recorded on a 640 μM Cy1 ILV sample on an 800 MHz Varian spectrometer equipped with a Chiliprobe, with the following parameters: 8 scans and a recycling delay of 1 s; direct ^{1}H dimension (910 complex points, 17.5 ppm spectral width, and carrier at 4.773 ppm); indirect ^{1}H methyl dimension (50 complex points; 3 ppm spectral width, and carrier at 0.5 ppm); indirect ^{13}C methyl dimension (65 complex points, 21 ppm spectral width, and carrier at 19 ppm); TOCSY mixing time of 11.7 ms. Water suppression was achieved using a WATERGATE[48] sequence before fid acquisition.

The TROSY-HcccoNH[49] data set was recorded on a Cy1 ILV sample (concentration < 640 μM, same sample as above after some protein losses due to precipitation in the NMR tube) using a 600 MHz Bruker AVANCE III spectrometer equipped with a QCI cryoprobe. The data was acquired nonuniformly with a schedule created to sample 374 out of 7500 indirect dimension points (~5% sampling) using

https://doi.org/10.1021/acs.jpca.1c05507
J. Phys. Chem. A 2021, 125, 8313−8323

PoissonGap. The data were acquired with the following parameters: 208 scans and a recycling delay of 1 s; direct ^{1}H dimension (1024 complex points, 16 ppm spectral width, and carrier at 4.695 ppm); indirect ^{15}N dimension (75 complex points, 33 ppm spectral width, and carrier at 119 ppm); indirect ^{1}H methyl dimension (100 complex points, 6.5 ppm spectral width, and carrier at 4.695 ppm); TOCSY mixing time of 32.9 ms. The ^{15}N encoding period was modified from the original pulse sequence to include a semiconstant time evolution period and concatenated to the reverse INEPT (insensitive nucleus enhancement by polarization transfer) before detection. The ^{13}C carrier was at 19 ppm for the first INEPT, at 42 ppm during the TOCSY transfer, and at 178.7 ppm during the ^{15}N encoding period.

Processing Considerations for Covariance NMR. Processing of the input spectra for the execution of the covariance MATLAB[50] script have been described in detail,[51] and we have recently provided a detailed protocol for calculating covariance maps[31] with examples that can be downloaded from our Web site (http://frueh.med.jhmi.edu/software-downloads/). We now provide an updated MATLAB script together with NMRPipe scripts that users can modify. When modifying the scripts, users should account for the following. The spectra need to be referenced consistently and feature a common digital resolution (points/Hz) along the covaried K dimensions, here the H-TOCSY dimension in the 3D-HCcH-TOCSY and 3D-HcccoNH spectra, and the ^{13}C dimension of the triple resonance experiments (HNCA, HNcoCA, HNcaCB, and HNcocaCB). For asparagine 4D maps, the spectra must feature a common digital resolution for the retained [I,J] and [L,M] dimensions (*vide infra*) across the different pairs of spectra used for element-wise (Hadamard) multiplication. Residual water signals in the detected H-TOCSY dimension of the 3D-HCcH-TOCSY water must be edited out of the spectrum (see next section) as they would otherwise introduce artificial noise into the covariance maps. We remind the reader that the file size of the 4D covariance map rapidly increases with resolution, and hence linear prediction and zero filling of the retained dimensions [I,J] and [L,M] has to be applied judiciously. In contrast, high resolution in the covaried dimension K does not increase the size of the final 4D as this dimension is subsumed.

Data was processed on an Intel PC workstation running Ubuntu Linux version 16.04.7. All nonuniformly sampled data were reconstructed using istHMS.[45] All processing steps other than reconstruction were carried out using NMRpipe.[52] All dimensions were apodized with a cosine-squared bell function.

PCP1 Data Processing. 3D-HCcH-TOCSY and 3D-HcccoNH were processed to have the H-TOCSY dimension along the *Z*-axis. The number of points in 3D-HCcH-TOCSY was doubled in each dimension by zero-filling, and the spectral region between 5.50 ppm and −2.00 ppm was extracted along the H-TOCSY dimension. Further, the NMRPipe[52] function SET was used to set the region between [4.68 ppm] and [4.90 ppm] around the water signal to zero. Circular-shift was used in the indirectly detected ^{1}H dimension to correct for the offset. The final data size of the processed 3D-HCcH-TOCSY spectrum was 100 (^{1}H) × 64 (^{13}C) × 959 (H_{toc}). For 3D-HcccoNH, the detected ^{1}H dimension was extracted between 6.20 and 9.50 ppm to contain all detected backbone amide signals. Zero-filling was avoided to keep the size of the spectrum small. The indirectly detected H-TOCSY dimension was linear predicted and zero-filled to a final size of 818 points to match the digital resolution of the TOCSY dimension in the 3D-HCcH-TOCSY spectrum. Circular-shift was used in the indirectly detected ^{1}H dimension to correct for frequency offsets between HCcH and HcccoNH TOCSY dimensions. The ^{15}N dimension was linear predicted and zero filled once. The final data size of the processed 3D-HcccoNH spectrum was 212 (^{1}H) × 80 (^{15}N) × 818 (H_{toc}).

3D-HNCA, 3D-HNcoCA, 3D-HNcaCB, and 3D-HNcocaCB, were processed to have the ^{13}C dimension along the *Z*-axis. The amide proton region between 9.5 and 6.5 ppm was extracted (195 real points) after Fourier transform. The nitrogen dimension was zero-filled to a final size of 128 complex points. Carbon dimensions were zero-filled to 400 (C^{α}) and 800 (C^{β}) complex points to achieve identical digital resolution in all the spectra. We applied a circular shift to the ^{13}C dimension of the 3D-HNcoCA spectrum to move aliased C^{β} signals correlated to side chain amides to their correct chemical shift position. The final data size of the processed spectra was 195 (^{1}H) × 128 (^{15}N) × 400 (C^{α}) for HNCA and HNcoCA, and 195 (^{1}H) × 128 (^{15}N) × 800 (C^{β}) for HNcaCB and HNcocaCB. To minimize artifacts, spectral regions specific to asparagine resonances were defined within the MATLAB script for the calculation of 4D-HN$H_{sc}N_{sc}$ covariance map: H [7.50−8.50 ppm], N [112.0−120.0 pm] for the dimensions [I,J] in 3D-HNCA and 3D-HNcaCB; H_{sc} [6.550−8.587 ppm] and N_{sc} [109.84−116.65 ppm] for dimensions [L,M] in 3D-HNcocaCB and 3D-HNcoCA. The spectral limits used in the covaried K dimension are [51.50−56.52 ppm] for C^{α} in 3D-HNCA and 3D-HNcocaCB and [37.50−40.50 ppm] for C^{β} in 3D-HNcaCB and 3D-HNcoCA.

Cy1 Data Processing. 3D-HCcH-TOCSY and 3D-HcccoNH were processed to have the TOCSY dimension along the *Z*-axis. The detected H-TOCSY dimension of 3D-HCcH-TOCSY was zero-filled to the nearest power of 2, and the region between −0.7 and 1.6 ppm was extracted. A linear prediction of indirect dimensions was carried out after Fourier transform of the remaining dimensions. Circular-shift was applied in the indirect ^{1}H dimension of 3D-HCcH-TOCSY. The final data size of the processed 3D-HCcH-TOCSY spectrum was 78 (^{1}H) × 130 (^{13}C) × 270 (H_{toc}). For 3D-HcccoNH, the detected ^{1}H dimension was extracted between 6.25 and 10.25 ppm. Zero-filling was avoided to keep the size of the spectrum small. The nitrogen dimension was zero-filled to a final size of 128 complex points. The indirect H-TOCSY dimension was zero-filled to a final size of 760 complex points to match the digital resolution of the H-TOCSY dimension in 3D-HCcH-TOCSY. Subsequently, the spectral region between −0.7 and 1.6 ppm that features methyl proton signals was extracted. The final data size of the processed 3D-HcccoNH spectrum was 240 (^{1}H) × 128 (^{15}N) × 270 (H_{toc}).

THEORY

Simultaneous Calculation of Symmetric and Non-symmetric Covariance Maps. Covariance NMR methods work on the principle that NMR spectra can be represented as numerical arrays[53,54] and hence are amenable to mathematical operations. Notably, Bruschweiler and co-workers introduced covariance as a means to establish correlations through processing,[28] and many related methods ensued.[55−60] (See refs 31 and 32 for recent reviews.) We highlight that since 2004 the spectra are subject to matrix multiplications rather than covariance,[29] although the term covariance NMR has been maintained to encapsulate all other efforts involved in

https://doi.org/10.1021/acs.jpca.1c05507
J. Phys. Chem. A 2021, 125, 8313−8323

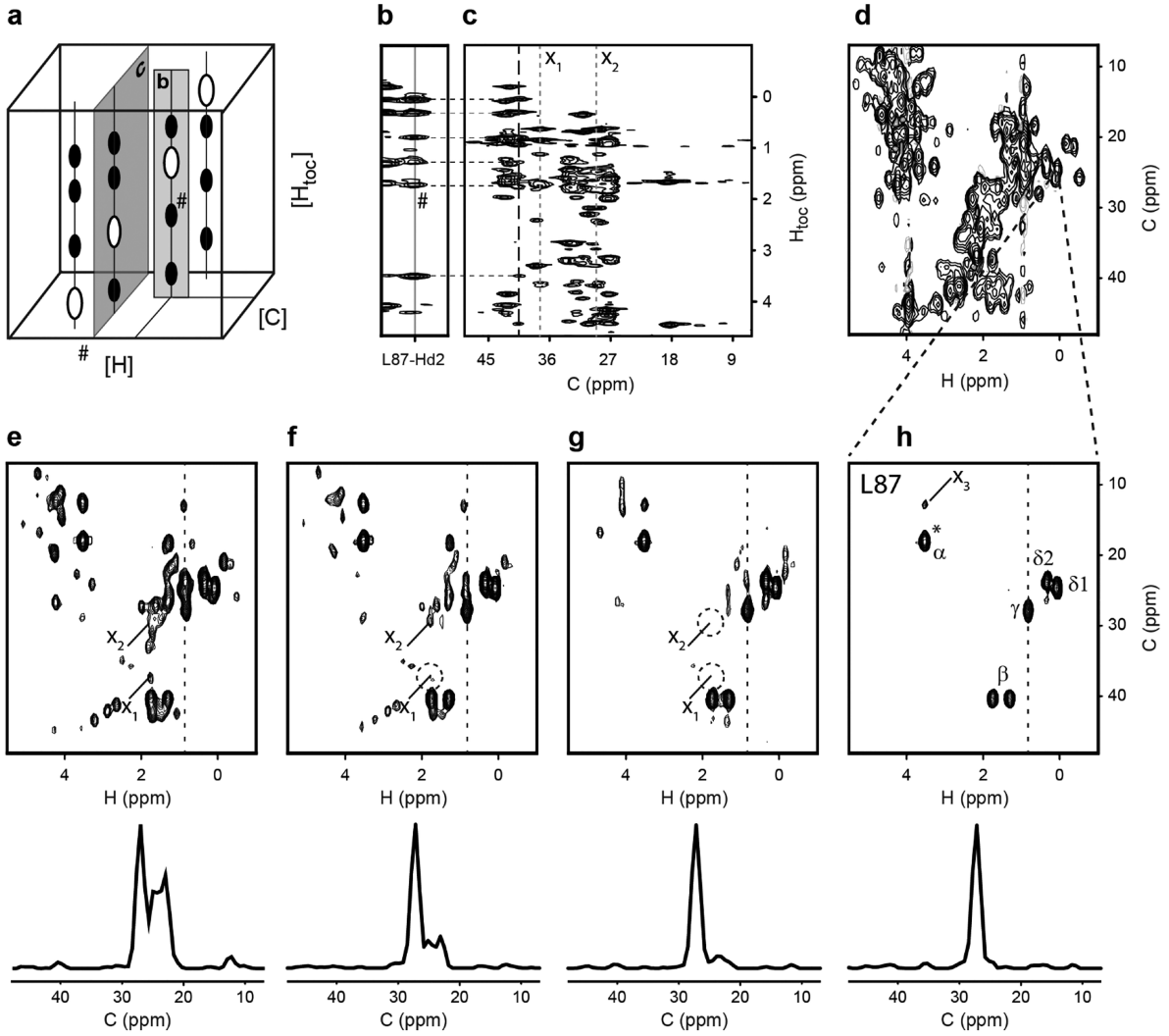


Figure 1. Assignment of side chain resonances using 3D-HCcH-TOCSY versus 4D-HCCH-TOCSY for the protein PCP1. (a) Schematic of the 3D-HCcH-TOCSY spectrum. Each (H,C) correlation features a TOCSY pattern along H_{toc}. The H_{toc}/C plane at position # (labeled c) is inspected to assign the ^{13}C frequency of the correlation labeled with # in the strip labeled with b. The labels b and c also refer to panels b and c. (b) TOCSY strip of Leu-87 of Cys-PCP1. (c) Orthogonal HC plane at the chemical shift of the β-protons marked # in panel b for identification of the chemical shift of the attached β-carbon. The black dashed line denotes the true ^{13}C frequency whereas the gray dotted lines at X_1 and X_2 highlight carbons where a subset of TOCSY correlations exhibit degeneracies or display a similar TOCSY pattern, respectively. (d) 2D-H/C projection of the 3D-HCcH-TOCSY spectrum highlighting the (H,C) correlation corresponding to the TOCSY strip in panel b. (e–h) H/C planes from 4D-HCCH-TOCSY covariance maps shown at contours 50% above the noise level. 1D slices through (H^{γ},C^{γ}) correlation taken along the carbon dimension (shown as dotted line) are displayed to illustrate reduction in artifacts.The artifacts arising from the TOCSY patterns at X_1 and X_2 in panel c are also indicated here. (e) 4D map without square-rooting or derivative, (f) without square-rooting but with derivative, (g) with square-rooting but without derivative, and (h) with square-rooting and with derivative. In panel h the artifacts reduce to the noise level and the side chain protons and carbons can be assigned at a single glance. The chemical shift of α-carbons are aliased as denoted by the asterisk (∗). See text for details.

improving the quality of the output spectra, such as those described below.

The current work introduces one new feature to facilitate computations and implements three procedures to minimize artifacts. The first procedure consists in taking the matrix square root of the covariance matrix to eliminate so-called pseudo-relayed artifacts.[28] These artifacts occur when two different spin-systems exhibit one or more accidental frequency degeneracies, that is, when two different nuclei have the same frequency leading to seemingly identical correlations. This situation frequently occurs between networks of scalar-coupled spin systems in TOCSY experiments and is observed in some of our applications. Matrix square-rooting is straightforward when the output of covariance is a square matrix, but implementation to unsymmetrical,[30] indirect covariance[61] necessitated the development of a generalized indirect covariance (GIC) formalism,[62] used in our applications. To describe the feature we introduce, we will represent 3D-spectra as a series of planes. Thus, a three-dimensional spectrum of size I × J × K is represented by a series of J matrices B_1, B_2, B_3, ..., and B_J, each with dimensions (I × K). This spectrum would correspond to a 3D-HCcH-TOCSY spectrum in our applications. The second spectrum of dimension L × M × K is similarly described by a series of M matrices A_1, A_2, A_3, ..., and A_M, each with matrix dimensions (L × K). This spectrum would correspond to that of 3D-HcccoNH in our applications. Note that the dimension of size K is common to both spectra and represents the dimension featuring TOCSY correlations in

https://doi.org/10.1021/acs.jpca.1c05507
J. Phys. Chem. A 2021, 125, 8313−8323

3D-HCcH-TOCSY and 3D-HcccoNH. Following Snyder et al.,[62] we build a GIC matrix carrying the combined multiplications of all M+J matrices as

$$C_{\mathrm{GIC(M+J)}} = S\cdot S^{\mathrm{T}} = \begin{bmatrix} A_1 \\ \vdots \\ A_{\mathrm{M}} \\ B_1 \\ \vdots \\ B_{\mathrm{J}} \end{bmatrix} \cdot [A_1^{\mathrm{T}} \quad \dots \quad A_{\mathrm{M}}^{\mathrm{T}} \quad B_1 \quad \dots \quad B_{\mathrm{J}}^{\mathrm{T}}] \tag{1}$$

or

$$C_{\mathrm{GIC}(m+n)} = \begin{bmatrix} A_1A_1^{\mathrm{T}} & \dots & A_1A_m^{\mathrm{T}} & A_1B_1^{\mathrm{T}} & \dots & A_1B_n^{\mathrm{T}} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ A_mA_1^{\mathrm{T}} & \dots & A_mA_m^{\mathrm{T}} & A_mB_1^{\mathrm{T}} & \dots & A_mB_n^{\mathrm{T}} \\ B_1A_1^{\mathrm{T}} & \dots & B_1A_m^{\mathrm{T}} & B_1B_1^{\mathrm{T}} & \dots & B_1B_n^{\mathrm{T}} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ B_nA_1^{\mathrm{T}} & \dots & B_nA_m^{\mathrm{T}} & B_nB_1^{\mathrm{T}} & \dots & B_nB_n^{\mathrm{T}} \end{bmatrix} \tag{2}$$

where the GIC matrix $C_{\mathrm{GIC(M+J)}}$ is a square matrix with dimensions $(N \times N)$, where $N = \mathrm{ML} + \mathrm{JI}$, which is amenable to matrix square-rooting performed through singular value decomposition. The spectrum of interest carrying the covariance between both spectra is obtained by extracting all $(\mathrm{M} \times \mathrm{J})$ submatrices in the upper right quadrant, that is, originated from $A_mB_j^{\mathrm{T}}$ matrices in eq 2, with m and j ranging from 1 to M and 1 to J, respectively. This spectrum corresponds to the 4D-HCccoNH map we describe below. However, in the current work, we also make use of correlations stemming from 3D-HCcH-TOCSY covaried with itself, and we modified our script to retrieve all $(\mathrm{J} \times \mathrm{J})$ planes corresponding to the lower right quadrant in eq 2, that is, with planes originating from $B_jB_j^{\mathrm{T}}$. This output corresponds to the 4D-HCCH-TOCSY described in the next section. In the end, the procedure described above permits simultaneous generation and application of matrix square-rooting to the 4D-HCCH-TOCSY and 4D-HCccoNH maps in a single calculation, which is a new feature of our updated covariance script.

The second procedure we implement consists of taking the spectral derivative along covaried dimensions before calculating the covariance matrices.[33,34] This procedure attenuates false correlations between systems featuring signals with similar but nonidentical frequencies along the covaried dimension, that is, signals that would display partial overlap if they were overlaid. These artifactual correlations arise because covariance generates correlations from matching signal amplitudes and not through their maxima. Taking the derivative converts maxima into inflection points, and mismatched frequencies lead to attenuated or suppressed correlations, whereas correlations from identical frequencies are preserved.[33,34] This procedure was employed in all our applications.

The third and final procedure relies on generating the same correlations from different sets of source spectra. When multiple sets of spectra are available to establish the same correlations, each set gives rise to different false positives reflecting different accidental frequency degeneracies. As long as the calculated maps feature the same dimension sizes, a simple, element-wise (Hadamard) multiplication will remove these false positives and enhance true correlations.[33,34] We implemented this procedure when calculating the 4D-$\mathrm{HNH_{sc}N_{sc}}$ map that was used for obtaining asparagine side chain assignments.

RESULTS AND DISCUSSION

4D-HCCH-TOCSY Map from a 3D Spectrum. Assigning resonances from 3D-HCcH-TOCSY[26] spectra is a well-established but cumbersome procedure and is limited by user interpretation. In this section, "assignment" refers to pairing signals that belong to the same side chain, without identifying the residue number. Figure 1a–c shows a schematic representation of the assignment process using a 3D-HCcH-TOCSY spectrum. The spectrum displays a TOCSY pattern for the side chain protons at every (H,C) correlation (Figure 1a). For a given side chain, every (H,C) correlation displays the same pattern, and ^{1}H and ^{13}C resonances can hence be assigned to the same side chain by identifying common patterns in the 3D spectrum. In practice, an H/H strip at a resolved correlation in HC-HSQC is first inspected to identify ^{1}H resonances within the side chain (Figure 1a, b). Next, orthogonal H/C planes (see definition in Figure 1a) are inspected at each of these proton frequencies, for example, that labeled # in Figure 1a, to identify common ^{1}H correlations at a given ^{13}C frequency, and hence pair this carbon frequency with that of the ^{1}H frequency defining the H/C plane (Figure 1c). This process is rather tedious and prone to human error such as missing partially overlapped or weaker correlations as illustrated in Figure 1c. The true ^{13}C frequency is shown by a black dashed line. A gray dashed line at X_1 highlights a ^{13}C frequency which exhibits a subset of correlations with accidental ^{1}H frequency degeneracy with the expected correlations. The gray dashed line at X_2 shows an example that exhibits a similar TOCSY pattern to that of the inspected signals. A closer inspection is needed to resolve these ambiguities, which reveals that signals missing in X_1 are not seen at the noise level and that the maxima of the signals seen in X_2 do not match those of the target system. How fast these ambiguities are resolved will depend on the alertness of the investigator. In contrast, we will show that each false candidate can be eliminated by a dedicated mathematical operation in the covaried spectrum.

Strip comparison within 3D-HCcH-TOCSY can be formulated as a covariance between the 3D-HCcH-TOCSY spectrum and itself. Here, we use a notation we previously introduced to describe spectra through their dimensions and through the coordinates of the data they contain.[31] Thus, the 3D spectrum of HCcH-TOCSY is described by $[\mathrm{H}][\mathrm{C}][\mathrm{H_{toc}}]$ (Figure 1a), and $[\mathrm{H}][\mathrm{C}][\mathrm{H_{toc}}](i,j,k)$ describes a specific data point at the indices (i,j,k) in the spectrum. Identifying common correlations along the $[\mathrm{H_{toc}}]$ dimension simply consists of performing covariance along the $[\mathrm{H_{toc}}]$ dimension:

$$[\mathrm{H}][\mathrm{C}][\mathrm{H}][\mathrm{C}](i,j,l,m) = \sum_{\mathrm{k}=1}^{\mathrm{K}} [\mathrm{H}][\mathrm{C}][\mathrm{H_{toc}}](i,j,k)\cdot[\mathrm{H}][\mathrm{C}][\mathrm{H_{toc}}](l,m,k) \tag{3}$$

The resulting 4D $[\mathrm{H}][\mathrm{C}][\mathrm{H}][\mathrm{C}]$ array can be visualized as an HC-HSQC of HC-HSQCs, where each (H,C) correlation has an associated 2D H/C plane that contains all aliphatic (H,C) correlations present in the corresponding side chain.

https://doi.org/10.1021/acs.jpca.1c05507
J. Phys. Chem. A 2021, 125, 8313−8323

Consequently, each (H,C) correlation immediately identifies an entire spin system, and hence amino acid type, through characteristic signals displayed in the corresponding H/C plane.

As discussed above, two challenges severely deteriorate the quality of the output 4D spectrum. First, degeneracies between ^{1}H frequencies of protons in different side chains, as seen in our example X_1, lead to so-called pseudo-relay artifacts.[62] Second, side chains with a set of signals with similar frequencies, that is, signals that would display partial overlap if they were overlaid as in example X_2, will lead to false positives. The combination of these effects renders the 4D maps unreliable and unusable as seen in Figure 1e. To remove pseudo-relay artifacts, we implemented matrix square-rooting[62] as shown in Figure 1f where the artifact stemming from X_1 disappears, among others. The artifacts due to partial spectral overlap are alleviated by applying a spectral derivative along the covaried dimension before calculating the covariance map[33,34] as evidenced by the disappearance of the artifact stemming from X_2 in Figure 1g. The combination of both methods shown in Figure 1h provides remarkably reliable correlation maps with minimal artifacts. The remaining artifacts in the 4D map reflect the limitations of the two methods in the presence of a large dynamic range, as seen for the correlation marked with X_3 in Figure 1h. We remind the reader that, when the 3D spectra are acquired in aqueous buffers, residual water signals must be removed before performing the covariance in fear of inducing noise artifacts (see Methods and Supporting Information, Figure S1).

The 4D map encapsulates all the steps from the side chain assignment procedure and presents them as an HC-HSQC for an individual side chain. Similarly, the operations we implemented in the current work in effect overcome the pitfalls described for assignments with 3D-HCcH-TOCSY. Thus, the artifacts removed by matrix square-rooting correspond to pairing strips that only display a subset of common frequencies such as those at X_1 in Figure 1c, and those removed by spectral derivative correspond to pairing strips with similar, yet slightly different patterns such as those at X_2 in Figure 1c. Just like every (H,C) coordinate corresponding to a side chain signal provides the same set of correlations in 3D-HCcH-TOCSY, the same coordinates now also provide H/C planes featuring all side chain (H,C) correlations. Clearly, the output map is subject to limitations of conventional multidimensional spectra. Thus, the H/C plane of a signal at the coordinate (H, C) may also display correlations of signals overlapping with the correlation of interest (so-called bleed-through correlations). The differences between the planes reflect different bleed-through correlations or variations in artifact suppression that stem from variations in correlation intensities in the original TOCSY spectrum. Thus, side chains are most easily assigned by first inspecting planes for isolated (H,C) correlations in the HC-HSQC where bleed-through correlations are minimal. If necessary, a subsequent comparison between all planes identifies true correlations seen in all planes and eliminates residual artifacts only seen in a given plane. In summary, covariance processing provides us with a 4D-HCCH-TOCSY for free from a 3D-HCcH-TOCSY spectrum and enables identification of all proton and carbon chemical shifts of a side chain from a single visual inspection.

4D-HCccoNH Map for Sequence-Specific Assignment of Side Chain Resonances. Traditionally, the information from 3D-HcccoNH[27] is compared with the information from 3D-HCcH-TOCSY[26] to complete sequence-specific assignments of side chain resonances. 3D-HcccoNH correlates the aliphatic resonances of a side chain with the backbone amide resonances of the successor residue in the sequence, enabling the transfer of sequence-specific assignments of backbone resonances to side chain resonances. As the experiment is rather insensitive, it is most often analyzed jointly with 3D-HCcH-TOCSY to palliate for weak or missing correlations. As described above, 3D-HCcH-TOCSY is also used to assign side chain ^{13}C resonances. The process involves comparisons between TOCSY patterns at (H,N) coordinates in the 3D-HcccoNH spectrum and TOCSY patterns at (H,C) coordinates in the 3D-HCcH-TOCSY spectrum. Again, this strip comparison can be formulated as a covariance along the common (compared) TOCSY dimension between 3D-HcccoNH and 3D-HCcH-TOCSY. We define $[H_s][N_s][H_{toc}]$ to describe the 3D-HcccoNH spectrum and $[H_s][N_s][H_{toc}]$-(l,m,k) to describe a specific data point in the spectrum, where the subscript "s" for the H and N dimensions highlights that these dimensions report on the successor residue. Using the notation defined above for 3D-HCcH-TOCSY, the calculation of the 4D map is described by

$$[H_s][N_s][H][C](l, m, i, j) = \sum_{k=1}^{K} [H_s][N_s][H_{toc}](l, m, k)\cdot[H][C][H_{toc}](i, j, k) \quad (4)$$

The resulting 4D $[H_s][N_s][H][C]$ map can be thought of as an HN-HSQC of HC-HSQCs, where each (H,N) coordinate in the HN-HSQC has an associated HC-HSQC depicting all (H,C) correlations of the side chain from the previous residue. The procedures described to remove artifacts in 4D-HCCH-TOCSY were implemented and revealed the same advantages (Supporting Figure S2). Figure 2 displays planes for a 4D-HCcccoNH correlation map where spectral derivatives and matrix square-rooting were performed. As described above, 4D-HCccoNH and 4D-HCCH-TOCSY spectra were obtained through a single calculation. Figure 2 provides various examples illustrating both the advantages of our method and how to use these spectra. The assigned backbone amide resonances of A88, E30, and K96 from the protein PCP1 are shown in the 2D-H^N/N projection of 3D-HcccoNH in Figure 2a. Figures 2 panels b−d display the H/C planes from the 4D-HCcccoNH map at the corresponding (H_s,N_s) correlations. In software packages such as CARA,[63] Sparky,[64] or CcpNmr Analysis,[65] the user simply clicks on one (H,N) coordinate and a synchronized 4D map reveals the corresponding H/C plane. The ^{1}H-TOCSY strips from the source 3D-HcccoNH spectrum at the same coordinates are shown as bottom panels. This layout mirrors that employed in our laboratory to rapidly and reliably benefit from our 4D covariance maps. Figure 2b shows the 4D H/C plane displaying the side chain resonances of L87, also used in Figure 1h. The assignment procedure is a simple visual inspection, in contrast to the lengthy protocols we described when using 3D spectra. The attentive reader will notice a few artifacts in Figure 2b absent in Figure 1h. They reflect the higher resolution in the covaried TOCSY dimension for 3D-HCcH-TOCSY, where TOCSY correlations appear in the detected dimension, compared to those in 3D-HcccoNH, where they appear in an indirect dimension of lower resolution. Thus, 4D-HCCH-TOCSY planes are occasionally inspected to identify such artifacts. Figure 2c highlights that in

https://doi.org/10.1021/acs.jpca.1c05507
J. Phys. Chem. A 2021, 125, 8313−8323

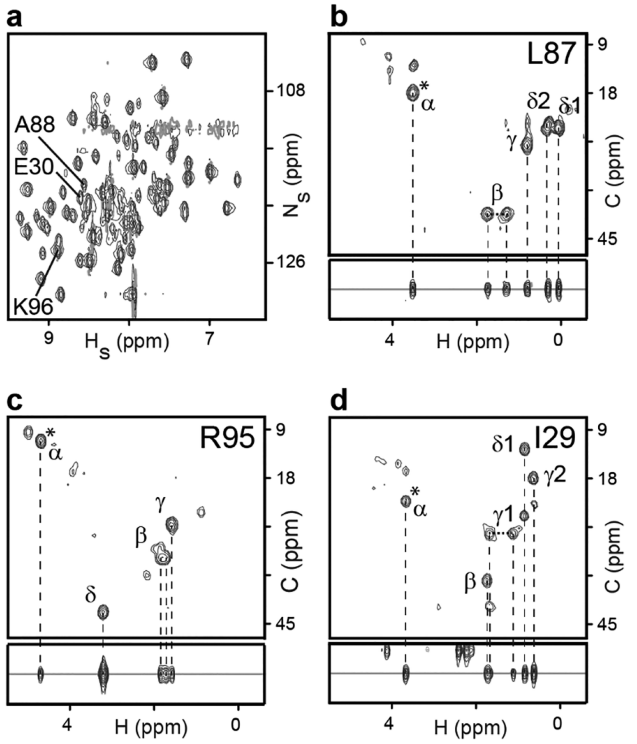


Figure 2. Using 4D-HCccoNH covariance map for transferring assignments from HN-HSQC to side chain resonances for the protein PCP1. (a) 2D-H/N projection of 3D-HcccoNH highlighting correlations for three examples. Panels b, c, and d show the HC planes from the 4D-HCccoNH map (top) and H-TOCSY strips from 3D-HcccoNH (bottom) at these correlations. The H-TOCSY strip from the 3D spectrum contains all the proton resonances from the sequential (i-1) residue which are further resolved in the 4D-map by the chemical shift of the attached carbon. The chemical shift of α-carbons are aliased as denoted by an asterisk (∗).

the 3D spectrum, three resonances could be paired with either β or γ moieties for R95 and only the 4D map reveals that two are paired to C^{β} and one to C^{γ}. Finally, Figure 2d shows how the 4D map seamlessly overcomes overlap between signals within the same side chain. Here, the ^{1}H signal of one γ1 proton overlaps with β, and both γ1 signals were initially erroneously thought to be equivalent when analyzing the 3D. The 4D map reveals resolved (H,C) correlations for the two γ protons. All examples also illustrate the importance of inspecting the original 3D spectra, as bleed-through or residual artifacts will easily be identified as such. Thus, in Figure 2c, four signals seen in the H/C plane belong to a different residue as seen from their absence in the strip from the 3D spectrum. Note that the 4D-HCccoNH map can also be navigated by viewing H/N planes associated with each (H,C) correlation. This application permits completion of the missing amide assignments from an assigned side chain signal, for example, when NOEs identify an aliphatic moiety for the predecessor of an unassigned residue.

Improving Signal-to-Noise within Side Chain Experiments of Large Proteins. The 3D-HcccoNH spectrum suffers from poor sensitivity due to transverse relaxation losses during the multiple magnetization transfers needed for the experiment. For larger proteins, traditionally, all nonexchangeable protons are replaced with deuterium save for those in the methyls of Ile(δ1), Leu, and Val.[35] Even so, 3D-HcccoNH spectra acquired on such a ^{1}H-methyl-ILV-^{2}H,^{13}C,^{15}N sample of the protein Cy1 (52 kDa) exhibited very low sensitivity. Notably, many times only one correlation is observed for Leu and Val when two signals are expected, and an investigator cannot distinguish between methyls displaying degenerate frequencies versus a single methyl detected with a second signal too weak to be observed. We have explored alternative assignment procedures in the past,[47,66,67] but we show here that covariance methods can be used to rescue this lack of sensitivity, a solution inspired by applications in studies of small molecules with NMR isotopes at natural abundance.[68]

In contrast to 3D-HcccoNH, 3D-HCcH-TOCSY is a sensitive experiment, and the procedure described by eq 4 can provide rescue in sensitivity. Indeed, Figure 3 reveals that

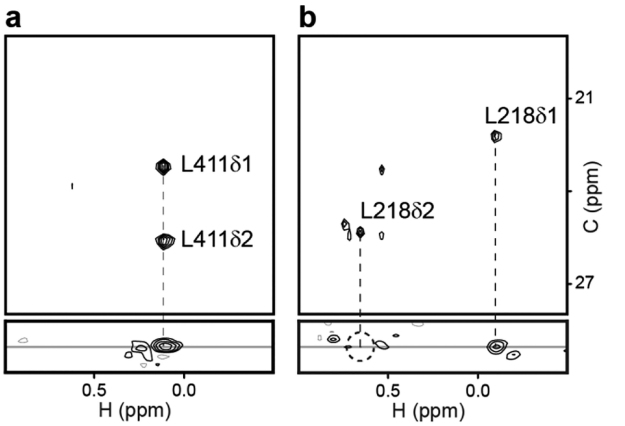


Figure 3. Resolving degeneracies and rescuing weak signals using 4D Covariance maps. (a) Methyl correlations of L411 overlap in the TOCSY strip from 3D-HcccoNH spectrum (bottom) but are resolved in the HC plane from 4D-HCccoNH map (top). (b) The signal for L218δ2 is not observed in the 3D-HcccoNH spectrum, as highlighted by a dashed circle (bottom), but rescued in the 4D-HCccoNH map (top).

the resulting 4D-HCcccoNH map could resolve the two scenarios described above using spectra of the 52 kDa Cy1 domain. Thus, the two degenerate signals of L411δ1 and L411δ2 become resolved through the carbon dimension in the H/C plane of the 4D-HCcccoNH shown in Figure 3a. In contrast, the signal for L218δ2 was too close to the noise level to be identified as a signal, but covariance with the sensitive 3D-HCcH-TOCSY provides the (H,C) correlation for L218δ2 shown in Figure 3b. Being able to faithfully distinguish these two scenarios is extremely valuable as considerable time may otherwise be wasted by lengthy inspections of 3D-HCcCH-TOCSY and NOESY (nuclear Overhauser effect spectroscopy) spectra, for example.

4D-HNH$_{sc}$N$_{sc}$ Covariance Map to Assign Asparagine Side Chain Amide Resonances. Assignments of side chain amide resonances of asparagine residues are often left out despite having the necessary information in conventional triple resonance experiments acquired for assigning backbone amide resonances. These assignments are invaluable when the side chains are involved in protein function or, for larger proteins, to supplement sparse distance constraints. Here, we exploit the incidental side chain signals that appear in triple resonance spectra to generate correlations with their assigned amide moieties through covariance.

https://doi.org/10.1021/acs.jpca.1c05507
J. Phys. Chem. A 2021, 125, 8313−8323

Figure 4c highlights backbone and side chain correlations observed for asparagine residues in triple resonance spectra

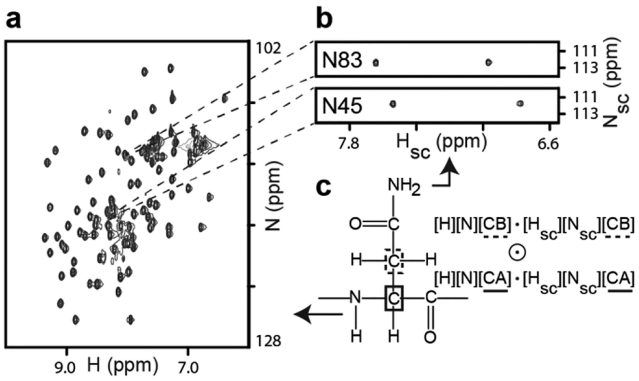


Figure 4. Assigning asparagine side chain amide resonances using the 4D-HN$H_{sc}N_{sc}$ covariance map. (a) HN-HSQC spectrum of PCP1 highlighting backbone correlations of two asparagine residues, N45 and N83. (b) H_{sc}/N_{sc} planes from 4D-HN$H_{sc}N_{sc}$ at these correlations reveal side chain amide resonances. (c) Schematic highlighting the asparagine nuclei and 3D spectra used for preparation of the 4D-HN$H_{sc}N_{sc}$ covariance map. ⊙ denotes element-wise product, and • denotes matrix multiplication.

and how they can be used to generate a 4D-HN$H_{sc}N_{sc}$ map correlating backbone and side chain amide resonances. Thus, the 3D-HNcocaCB spectrum will not only provide correlations between backbone amides and the β carbons of their predecessors but also correlations between side chain amides and their α carbons, denoted by (H_{sc},N_{sc},CA). That is, with respect to side chain amides, HNcocaCB provides a spectrum described as [H_{sc}][N_{sc}][CA] spanning coordinates [H_{sc}][N_{sc}]-[CA](l,m,k). Backbone amide resonances are correlated with the same C^{α} in HNCA spectra, denoted by [H][N][CA] and featuring [H][N][CA](i,j,k) coordinates. Hence, unassigned amide side chain resonances can be correlated with assigned backbone amide resonances through

$$[\mathrm{H}][\mathrm{N}][\mathrm{H}_{sc}][\mathrm{N}_{sc}](i,j,l,m) = \sum_{k=1}^{K} [\mathrm{H}][\mathrm{N}][\mathrm{CA}](i,j,k)\cdot[\mathrm{H}_{sc}][\mathrm{N}_{sc}][\mathrm{CA}](l,m,k) \quad (5)$$

which provides a 4D array carrying the desired (H,N,H_{sc},N_{sc}) correlations. We avoid undesired correlations by extracting a region of the common [CA] dimension that focuses on Asn C^{α} resonances before calculating the 4D map, as we previously suggested.[47] A second 4D-HN$H_{sc}N_{sc}$ map can be obtained by matching C^{β} chemical shifts using 3D-HNcaCB and 3D-HNcoCA, which provide the spectra [H][N][CB] and [H_{sc}][N_{sc}][CB] using the notation described above:

$$[\mathrm{H}][\mathrm{N}][\mathrm{H}_{sc}][\mathrm{N}_{sc}](i,j,l,m) = \sum_{k=1}^{K} [\mathrm{H}][\mathrm{N}][\mathrm{CB}](i,j,k)\cdot[\mathrm{H}_{sc}][\mathrm{N}_{sc}][\mathrm{CB}](l,m,k) \quad (6)$$

Note that depending on the spectral widths used in the carbon dimensions, correlations to the side chain amide resonances may be aliased in the spectrum, and circular shift may need to be applied so that the signals appear at the correct chemical shifts. The 4D maps obtained by eqs 5 and 6 each contain the same desired (H,N,H_{sc},N_{sc}) but likely different artifacts. Indeed, artifacts in the former are due to accidental C^{α} frequency degeneracies, while those in the latter reflect C^{β} chemical shift degeneracies. A simple element-wise product between the two 4D maps, denoted by ⊙ in Figure 4c, suppresses these artifacts and enhances the desired (H,N,H_{sc},N_{sc}) correlations. In our scripts, users only need to specify the four spectra involved in eqs 5 and 6, and all operations described in this section are performed during processing. Figure 4 panels a and b illustrate how asparagine side chain signals are then assigned from a simple visual inspection.

Note that here we have used residual signals of side chains in experiments designed to detect backbone resonances. Superior results are expected with experiments[69−71] designed to detect side chain resonances.

CONCLUSIONS

We have demonstrated that by incorporating matrix square-rooting and spectral derivatives reliable 4D covariance maps can be prepared from extremely crowded 3D spectra such as those of 3D-HcccoNH and 3D-HCcH-TOCSY. Revisiting the generalized indirect covariance formalism, we calculated a 4D-HCCH-TOCSY map conjointly with a 4D-HCccoNH map from 3D-HcccoNH and 3D-HCcH-TOCSY input spectra. Thus, a fourth dimension was obtained for each experiment, at no cost in experimental time. For larger proteins, these covariance maps can rescue weak side chain correlations when a spectrum of lower sensitivity is combined with a spectrum of higher sensitivity. Finally, we also demonstrated that with element-wise multiplications, reliable 4D maps can be calculated from a set of conventional triple resonance spectra to assign the side chain amide resonances of asparagine. In summary, 4D covariance maps prepared from existing 3D spectra can help improve the accuracy and completeness of side chain assignments by resolving signals along a fourth dimension and rescuing weak correlations without the need for lengthy acquisitions of 4D spectra. The covariance NMR methods presented here will be a valuable addition to the NMR spectral assignment toolkit and enable comprehensive protein signal assignments necessary for thorough investigations of protein function.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jpca.1c05507.

Plots from 4D-Covariance maps showing artificial noise due to residual water signal in 4D-HCCH-TOCSY maps and removal of artifacts in 4D-HCccoNH maps using spectral derivative and square-rooting (PDF)

AUTHOR INFORMATION

Corresponding Author

Dominique P. Frueh − *Department of Biophysics and Biophysical Chemistry, Johns Hopkins University School of Medicine, Baltimore, Maryland 21205, United States;* orcid.org/0000-0003-4605-3776; Email: dfrueh@jhmi.edu

https://doi.org/10.1021/acs.jpca.1c05507
J. Phys. Chem. A 2021, 125, 8313−8323

Authors

Aswani K. Kancherla − *Department of Biophysics and Biophysical Chemistry, Johns Hopkins University School of Medicine, Baltimore, Maryland 21205, United States;* orcid.org/0000-0002-8647-4666

Kenneth A. Marincin − *Department of Biophysics and Biophysical Chemistry, Johns Hopkins University School of Medicine, Baltimore, Maryland 21205, United States*

Subrata H. Mishra − *Department of Biophysics and Biophysical Chemistry, Johns Hopkins University School of Medicine, Baltimore, Maryland 21205, United States;* Present Address: Reference Standard Laboratory, United States Pharmacopeial Convention, 12601 Twinbrook Pkwy, Rockville, MD, 20852, USA

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jpca.1c05507

Author Contributions

A.K.K. calculated all maps, K.A.M. collected data for PCP1, and S.H.M. collected data for Cy1 and conceptualized applications of covariance for HCCH and $HNH_{sc}N_{sc}$ maps. D.P.F. designed and supervised the research. A.K.K. and D.P.F. wrote the manuscript; all authors have approved the final version of the manuscript.

Notes

The authors declare no competing financial interest.

NMRPipe and MATLAB (Octave) scripts available at http://frueh.med.jhmi.edu/software-downloads/.

■ ACKNOWLEDGMENTS

We thank Dr. Scott Nichols for acquiring the 3D-HcccoNH data of Cy1 and Dr. Bradley Harden for early MATLAB covariance scripts. We thank Drs. David Meyers and Yousang Hwang from the Chemical Core Facility at the Johns Hopkins Medical Institutions for synthesizing phosphopantetheine analogues. Research in the Frueh Lab is supported by the National Institutes of Health (Grant R01 GM 104257).

■ REFERENCES

(1) Xie, T.; Saleh, T.; Rossi, P.; Kalodimos, C. G. Conformational States Dynamically Populated by a Kinase Determine Its Function. *Science* **2020**, *370*, 6513.

(2) Sesterhenn, F.; Yang, C.; Bonet, J.; Cramer, J. T.; Wen, X.; Wang, Y.; Chiang, C. I.; Abriata, L. A.; Kucharska, I.; Castoro, G.; et al. De Novo Protein Design Enables the Precise Induction of RSV-Neutralizing Antibodies. *Science* **2020**, *368*, 6492.

(3) Majewski, D. D.; Okon, M.; Heinkel, F.; Robb, C. S.; Vuckovic, M.; McIntosh, L. P.; Strynadka, N. C. J. Characterization of the Pilotin-Secretin Complex from the Salmonella Enterica Type III Secretion System Using Hybrid Structural Methods. *Structure* **2021**, *29* (2), 125−138.e5.

(4) Dixit, K.; Megha Karanth, N.; Nair, S.; Kumari, K.; Chakrabarti, K. S.; Savithri, H. S.; Sarma, S. P. Aromatic Interactions Drive the Coupled Folding and Binding of the Intrinsically Disordered Sesbania Mosaic Virus VPg Protein. *Biochemistry* **2020**, *59* (49), 4663−4680.

(5) Fawzi, N. L.; Ying, J.; Ghirlando, R.; Torchia, D. A.; Clore, G. M. Atomic-Resolution Dynamics on the Surface of Amyloid-β Protofibrils Probed by Solution NMR. *Nature* **2011**, *480* (7376), 268−272.

(6) Whittier, S. K.; Hengge, A. C.; Loria, J. P. Conformational Motions Regulate Phosphoryl Transfer in Related Protein Tyrosine Phosphatases. *Science* **2013**, *341* (6148), 899−903.

(7) Chakrabarti, K. S.; Li, J.; Das, R.; Byrd, R. A. Conformational Dynamics and Allostery in E2:E3 Interactions Drive Ubiquitination: Gp78 and Ube2g2. *Structure* **2017**, *25* (5), 794−805.e5.

(8) Otten, R.; Liu, L.; Kenner, L. R.; Clarkson, M. W.; Mavor, D.; Tawfik, D. S.; Kern, D.; Fraser, J. S. Rescue of Conformational Dynamics in Enzyme Catalysis by Directed Evolution. *Nat. Commun.* **2018**, *9* (1), 1−11.

(9) Doucet, N.; Watt, E. D.; Loria, J. P. The Flexibility of a Distant Loop Modulates Active Site Motion and Product Release in Ribonuclease A. *Biochemistry* **2009**, *48* (30), 7160−7168.

(10) Meinhold, D. W.; Wright, P. E. Measurement of Protein Unfolding/Refolding Kinetics and Structural Characterization of Hidden Intermediates by NMR Relaxation Dispersion. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108* (22), 9078−9083.

(11) Fawzi, N. L.; Ying, J.; Torchia, D. A.; Clore, G. M. Probing Exchange Kinetics and Atomic Resolution Dynamics in High-Molecular-Weight Complexes Using Dark-State Exchange Saturation Transfer NMR Spectroscopy. *Nat. Protoc.* **2012**, *7* (8), 1523−1533.

(12) Charlier, C.; Alderson, T. R.; Courtney, J. M.; Ying, J.; Anfinrud, P.; Bax, A. Study of Protein Folding under Native Conditions by Rapidly Switching the Hydrostatic Pressure inside an NMR Sample Cell. *Proc. Natl. Acad. Sci. U. S. A.* **2018**, *115* (18), E4169−E4178.

(13) Schrank, T. P.; Bolen, D. W.; Hilser, V. J. Rational Modulation of Conformational Fluctuations in Adenylate Kinase Reveals a Local Unfolding Mechanism for Allostery and Functional Adaptation in Proteins. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106* (40), 16984−16989.

(14) Greenwood, A. I.; Rogals, M. J.; De, S.; Lu, K. P.; Kovrigin, E. L.; Nicholson, L. K. Complete Determination of the Pin1 Catalytic Domain Thermodynamic Cycle by NMR Lineshape Analysis. *J. Biomol. NMR* **2011**, *51* (1−2), 21−34.

(15) Rennella, E.; Morgan, G. J.; Yan, N.; Kelly, J. W.; Kay, L. E. The Role of Protein Thermodynamics and Primary Structure in Fibrillogenesis of Variable Domains from Immunoglobulin Light Chains. *J. Am. Chem. Soc.* **2019**, *141* (34), 13562−13571.

(16) Puglisi, R.; Brylski, O.; Alfano, C.; Martin, S. R.; Pastore, A.; Temussi, P. A. Quantifying the Thermodynamics of Protein Unfolding Using 2D NMR Spectroscopy. *Commun. Chem.* **2020**, *3* (1), 1−7.

(17) Bodenhausen, G.; Ruben, D. J. Natural Abundance Nitrogen-15 NMR by Enhanced Heteronuclear Spectroscopy. *Chem. Phys. Lett.* **1980**, *69* (1), 185−189.

(18) Pervushin, K.; Riek, R.; Wider, G.; Wuthrich, K. Attenuated T2 Relaxation by Mutual Cancellation of Dipole-Dipole Coupling and Chemical Shift Anisotropy Indicates an Avenue to NMR Structures of Very Large Biological Macromolecules in Solution. *Proc. Natl. Acad. Sci. U. S. A.* **1997**, *94*, 12366.

(19) LeMaster, D. M. NMR Relaxation Order Parameter Analysis of the Dynamics of Protein Side Chains. *J. Am. Chem. Soc.* **1999**, *121* (8), 1726−1742.

(20) Frederick, K. K.; Marlow, M. S.; Valentine, K. G.; Wand, A. J. Conformational Entropy in Molecular Recognition by Proteins. *Nat. 2007* **2007**, *448* (7151), 325−329, DOI: 10.1038/nature05959.

(21) Paquin, R.; Ferrage, F.; Mulder, F. A. A.; Akke, M.; Bodenhausen, G. Multiple-Timescale Dynamics of Side-Chain Carboxyl and Carbonyl Groups in Proteins by 13C Nuclear Spin Relaxation. *J. Am. Chem. Soc.* **2008**, *130* (47), 15805−15807.

(22) Stafford, K. A.; Ferrage, F.; Cho, J.-H.; Palmer, A. G. Side Chain Dynamics of Carboxyl and Carbonyl Groups in the Catalytic Function of Escherichia Coli Ribonuclease H. *J. Am. Chem. Soc.* **2013**, *135* (48), 18024−18027.

(23) Zeymer, C.; Werbeck, N. D.; Zimmermann, S.; Reinstein, J.; Hansen, D. F. Characterizing Active Site Conformational Heterogeneity along the Trajectory of an Enzymatic Phosphoryl Transfer Reaction. *Angew. Chem., Int. Ed.* **2016**, *55* (38), 11533−11537.

(24) Fenwick, R. B.; Vögeli, B. Detection of Correlated Protein Backbone and Side-Chain Angle Fluctuations. *ChemBioChem* **2017**, *18* (20), 2016−2021.

(25) Nguyen, D.; Chen, C.; Pettitt, B. M.; Iwahara, J. NMR Methods for Characterizing the Basic Side Chains of Proteins: Electrostatic Interactions, Hydrogen Bonds, and Conformational Dynamics. *Methods Enzymol.* **2019**, *615*, 285−332.

(26) Bax, A.; Clore, G. M.; Gronenborn, A. M. $^{1}H-^{1}H$ Correlation via Isotropic Mixing of ^{13}C Magnetization, a New Three-Dimensional Approach for Assigning ^{1}H and ^{13}C Spectra of ^{13}C-Enriched Proteins. *J. Magn. Reson.* **1990**, *88* (2), 425−431.
(27) Grzesiek, S.; Anglister, J.; Bax, A. A. Correlation of Backbone Amide and Aliphatic Side-Chain Resonances in 13C/15N-Enriched Proteins by Isotropic Mixing of 13C Magnetization. *J. Magn. Reson., Ser. B* **1993**, *101* (1), 114−119.
(28) Brüschweiler, R.; Zhang, F. Covariance Nuclear Magnetic Resonance Spectroscopy. *J. Chem. Phys.* **2004**, *120* (11), 5253−5260.
(29) Bruschweiler, R. Theory of Covariance Nuclear Magnetic Resonance Spectroscopy. *J. Chem. Phys.* **2004**, *121* (1), 409−414.
(30) Blinov, K. A.; Larin, N. I.; Williams, A. J.; Zell, M.; Martin, G. E. Long-Range Carbon-Carbon Connectivity via Unsymmetrical Indirect Covariance Processing of HSQC and HMBC NMR Data. *Magn. Reson. Chem.* **2006**, *44* (2), 107−109.
(31) Kancherla, A. K.; Frueh, D. P. Covariance Nuclear Magnetic Resonance Methods for Obtaining Protein Assignments and Novel Correlations. *Concepts Magn. Reson., Part A* **2017**, *46A* (2), e21437.
(32) Snyder, D. A. Covariance NMR: Theoretical Concerns, Practical Considerations, Contemporary Applications and Related Techniques. *Prog. Nucl. Magn. Reson. Spectrosc.* **2021**, *122*, 1−10, DOI: 10.1016/j.pnmrs.2020.09.001.
(33) Harden, B. J.; Nichols, S. R.; Frueh, D. P. Facilitated Assignment of Large Protein NMR Signals with Covariance Sequential Spectra Using Spectral Derivatives. *J. Am. Chem. Soc.* **2014**, *136* (38), 13106−13109.
(34) Harden, B. J.; Mishra, S. H.; Frueh, D. P. Effortless Assignment with 4D Covariance Sequential Correlation Maps. *J. Magn. Reson.* **2015**, *260*, 83−88.
(35) Goto, N. K.; Gardner, K. H.; Mueller, G. A.; Willis, R. C.; Kay, L. E. A Robust and Cost-Effective Method for the Production of Val, Leu, Ile (Delta 1) Methyl-Protonated N-15-, C-13-, H-2-Labeled Proteins. *J. Biomol. NMR* **1999**, *13* (4), 369−374.
(36) Keating, T. A.; Miller, D. A.; Walsh, C. T. Expression, Purification, and Characterization of HMWP2, a 229 KDa, Six Domain Protein Subunit of Yersiniabactin Synthetase. *Biochemistry* **2000**, *39* (16), 4729−4739.
(37) Eaton, J. W., Bateman, D., Hauberg, S., Wehbring, R. *GNU Octave Version 4.0.0 Manual: A High-Level Interactive Language for Numerical Computations.*; John W. Eaton: 2015.
(38) Marincin, K. A.; Pal, I.; Frueh, D. P. Using Delayed Decoupling to Attenuate Residual Signals in Editing Filters. *Magn. Reson. Discuss.* **2021**, *2*, 1−21.
(39) Worthington, A. S.; Burkart, M. D. One-Pot Chemo-Enzymatic Synthesis of Reporter-Modified Proteins. *Org. Biomol. Chem.* **2006**, *4* (1), 44−46.
(40) Kittilä, T.; Cryle, M. J. An Enhanced Chemoenzymatic Method for Loading Substrates onto Carrier Protein Domains. *Biochem. Cell Biol.* **2018**, *96* (3), 372−379.
(41) Ikura, M.; Kay, L. E.; Bax, A. A Novel Approach for Sequential Assignment of 1H, 13C, and 15N Spectra of Larger Proteins: Heteronuclear Triple-Resonance Three-Dimensional NMR Spectroscopy. Application to Calmodulin. *Biochemistry* **1990**, *29* (19), 4659−4667.
(42) Bax, A.; Ikura, M. An Efficient 3D NMR Technique for Correlating the Proton and 15N Backbone Amide Resonances with the α-Carbon of the Preceding Residue in Uniformly 15N/13C Enriched Proteins. *J. Biomol. NMR* **1991**, *1* (1), 99−104.
(43) Wittekind, M.; Mueller, L. HNCACB, a High-Sensitivity 3D NMR Experiment to Correlate Amide-Proton and Nitrogen Resonances with the Alpha-and Beta-Carbon Resonances in Proteins. *J. Magn. Reson., Ser. B* **1993**, *101* (2), 201−205.
(44) Yamazaki, T.; Lee, W.; Arrowsmith, C. H.; Muhandiram, D. R.; Kay, L. E. A Suite of Triple Resonance NMR Experiments for the Backbone Assignment of 15N, 13C, 2H Labeled Proteins with High Sensitivity. *J. Am. Chem. Soc.* **1994**, *116* (26), 11655−11666.
(45) Hyberts, S. G.; Milbradt, A. G.; Wagner, A. B.; Arthanari, H.; Wagner, G. Application of Iterative Soft Thresholding for Fast Reconstruction of NMR Data Non-Uniformly Sampled with Multidimensional Poisson Gap Scheduling. *J. Biomol. NMR* **2012**, *52* (4), 315−327.
(46) Mishra, S. H.; Harden, B. J.; Frueh, D. P. A 3D Time-Shared NOESY Experiment Designed to Provide Optimal Resolution for Accurate Assignment of NMR Distance Restraints in Large Proteins. *J. Biomol. NMR* **2014**, *60* (4), 265−274.
(47) Mishra, S. H.; Frueh, D. P. Assignment of Methyl NMR Resonances of a 52 KDa Protein with Residue-Specific 4D Correlation Maps. *J. Biomol. NMR* **2015**, *62* (3), 281−290.
(48) Piotto, M.; Saudek, V.; Sklenár, V. Gradient-Tailored Excitation for Single-Quantum NMR Spectroscopy of Aqueous Solutions. *J. Biomol. NMR* **1992**, *2* (6), 661−665.
(49) Hilty, C.; Fernández, C.; Wider, G.; Wüthrich, K. Side Chain NMR Assignments in the Membrane Protein OmpX Reconstituted in DHPC Micelles. *J. Biomol. NMR* **2002**, *23* (4), 289−301.
(50) Mathworks, I. *MATLAB: R2014a*; Mathworks Inc: Natick, 2014.
(51) Harden, B. J.; Frueh, D. P. Chapter 16 Covariance NMR Processing and Analysis for Protein. *Methods Mol. Biol.* **2018**, *1688*, 353−373.
(52) Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A. NMRPipe: A Multidimensional Spectral Processing System Based on UNIX Pipes. *J. Biomol. NMR* **1995**, *6* (3), 277−293.
(53) Lipkus, A. H.; Nieman, R. A.; Munk, M. E. A Manipulation of Two-Dimensional NMR Spectra Based on Graph Theory. *J. Magn. Reson., Ser. A* **1993**, *102* (1), 24−28.
(54) Bartels, C.; Wüthrich, K. A Spectral Correlation Function for Efficient Sequential NMR Assignments of Uniformly 15N-Labeled Proteins. *J. Biomol. NMR* **1994**, *4* (6), 775−785.
(55) Trbovic, N.; Smirnov, S.; Zhang, F.; Bruschweiler, R. *J. Magn. Reson.* **2004**, *171*, 277−283.
(56) Lescop, E.; Brutscher, B. Hyperdimensional Protein NMR Spectroscopy in Peptide-Sequence Space. *J. Am. Chem. Soc.* **2007**, *129* (39), 11916−11917.
(57) Kupče, E.; Freeman, R. Hyperdimensional NMR Spectroscopy. *J. Am. Chem. Soc.* **2006**, *128* (18), 6020−6021.
(58) Chen, K.; Delaglio, F.; Tjandra, N. A Practical Implementation of Cross-Spectrum in Protein Backbone Resonance Assignment. *J. Magn. Reson.* **2010**, *203* (2), 208−212.
(59) Benison, G.; Berkholz, D. S.; Barbar, E. Protein Assignments without Peak Lists Using Higher-Order Spectra. *J. Magn. Reson.* **2007**, *189* (2), 173−181.
(60) Wei, Q.; Chen, J.; Mi, J.; Zhang, J.; Ruan, K.; Wu, J. NMR Backbone Assignment of Large Proteins by Using13Cα-Only Triple-Resonance Experiments. *Chem. - Eur. J.* **2016**, *22* (28), 9556−9564.
(61) Zhang, F.; Bruschweiler, R. Indirect Covariance NMR Spectroscopy. *J. Am. Chem. Soc.* **2004**, *126* (41), 13180−13181.
(62) Snyder, D. A.; Bruschweiler, R. Generalized Indirect Covariance NMR Formalism for Establishment of Multidimensional Spin Correlations. *J. Phys. Chem. A* **2009**, *113*, 12898−12903.
(63) Keller, R. *The Computer Aided Resonance Assignment Tutorial*, 1st ed.; Cantina Verlag: Goldau, Switzerland, 2004.
(64) Goddard, T. D.; Kneller, D. G. *Sparky 3, NMR Assignment and Integration Software*; University of California: San Francisco. 2004.
(65) Vranken, W. F.; Boucher, W.; Stevens, T. J.; Fogh, R. H.; Pajon, A.; Llinas, M.; Ulrich, E. L.; Markley, J. L.; Ionides, J.; Laue, E. D. The CCPN Data Model for NMR Spectroscopy: Development of a Software Pipeline. *Proteins: Struct., Funct., Genet.* **2005**, *59* (4), 687−696.
(66) Tugarinov, V.; Kay, L. E. Ile, Leu, and Val Methyl Assignments of the 723-Residue Malate Synthase G Using a New Labeling Strategy and Novel NMR Methods. *J. Am. Chem. Soc.* **2003**, *125* (45), 13868−13878.
(67) Kerfah, R.; Hamelin, O.; Boisbouvier, J.; Marion, D. CH3-Specific NMR Assignment of Alanine, Isoleucine, Leucine and Valine Methyl Groups in High Molecular Weight Proteins Using a Single Sample. *J. Biomol. NMR* **2015**, *63* (4), 389−402.

(68) Martin, G. E.; Irish, P. A.; Hilton, B. D.; Blinov, K. A.; Williams, A. J. Utilizing Unsymmetrical Indirect Covariance Processing to Define 15N-13C Connectivity Networks. *Magn. Reson. Chem.* **2007**, *45* (8), 624−627.
(69) Farmer, B. T.; Venters, R. A. Assignment of Aliphatic Side-Chain 1HN/15N Resonances in Perdeuterated Proteins. *J. Biomol. NMR* **1996**, *7* (1), 59−71.
(70) McIntosh, L. P.; Brun, E.; Kay, L. E. Stereospecific Assignment of the NH2 Resonances from the Primary Amides of Asparagine and Glutamine Side Chains in Isotopically Labeled Proteins. *J. Biomol. NMR* **1997**, *9* (3), 306−312.
(71) Liu, A.; Yao, L.; Li, Y.; Yan, H. TROSY of Side-Chain Amides in Large Proteins. *J. Magn. Reson.* **2007**, *186* (2), 319−326.

https://doi.org/10.1021/acs.jpca.1c05507
J. Phys. Chem. A 2021, 125, 8313−8323