

# Proton Fingerprints Portray Molecular Structures: Enhanced Description of the $^1\text{H}$ NMR Spectra of Small Molecules

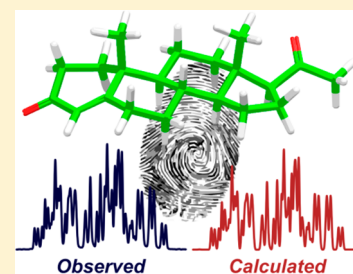
José G. Napolitano,<sup>†</sup> David C. Lankin,<sup>†</sup> James B. McAlpine,<sup>†</sup> Matthias Niemitz,<sup>‡</sup> Samuli-Petrus Korhonen,<sup>‡</sup> Shao-Nong Chen,<sup>†</sup> and Guido F. Pauli<sup>\*,†</sup>

<sup>†</sup>Department of Medicinal Chemistry & Pharmacognosy and Institute for Tuberculosis Research, College of Pharmacy, University of Illinois at Chicago, 833 South Wood Street, Chicago, Illinois 60612, United States

<sup>‡</sup>PERCH Solutions Ltd., Hyräräkatu 3A1, Kuopio 70500, Finland

## S Supporting Information

**ABSTRACT:** The characteristic signals observed in NMR spectra encode essential information on the structure of small molecules. However, extracting all of this information from complex signal patterns is not trivial. This report demonstrates how computer-aided spectral analysis enables the complete interpretation of 1D  $^1\text{H}$  NMR data. The effectiveness of this approach is illustrated with a set of organic molecules, for which replicas of their  $^1\text{H}$  NMR spectra were generated. The potential impact of this methodology on organic chemistry research is discussed.



NMR spectroscopy is acknowledged as a powerful, versatile tool to obtain structural information at the molecular level. For small organic molecules, solution-state NMR is the mainstay for structure elucidation and relative configuration, as well as conformational analysis. Moreover, NMR can simultaneously provide quantitation.

Although 2D/3D NMR techniques have facilitated structural analysis, the interpretation of complex resonance patterns observed in 1D  $^1\text{H}$  NMR spectra is still a challenge. Difficulties in analyzing these complex patterns originate from the intrinsically limited dispersion of the  $^1\text{H}$  dimension. Therefore, although these resonance patterns contain a wealth of structural information, a rigorous computational analysis is needed to mine it. Several computer programs for the analysis of high-resolution NMR spectra have been developed,<sup>1–3</sup> but have found limited application among non-NMR specialists. This may result from a widespread lack of awareness of their value or perhaps because the software packages are perceived as being user-unfriendly. Consequently, the development of platforms for semiautomated analysis of “multiplets” will provide scientists with new means to understand complex NMR spectra.

Inspired by the pioneering work of Raymond J. Abraham<sup>1</sup> and Ted Schaefer,<sup>4,5</sup> this study describes how modern computational tools for spectral prediction, simulation, and iteration can decode resonance patterns and allow the use of  $^1\text{H}$  NMR data to portray molecular structures. The application of  $^1\text{H}$  iterative full spin analysis (HiFSA)<sup>6</sup> using PERCH NMR software<sup>7,8</sup> and the automated consistency analysis (ACA)<sup>9,10</sup> module enables a thorough analysis of  $^1\text{H}$  NMR spectra, as demonstrated for a series of increasingly complex organic molecules. This approach produces replicas of  $^1\text{H}$  NMR spectra (i.e.,  $^1\text{H}$  fingerprints), provides complete assignments, and

extracts important spectral parameters, including chemical shifts ( $\delta$ ), coupling constants ( $J$ ), and effective linewidths ( $\Delta\nu_{1/2}$ ). These elements form a comprehensive *profile* that links the molecular structure and the NMR spectrum.

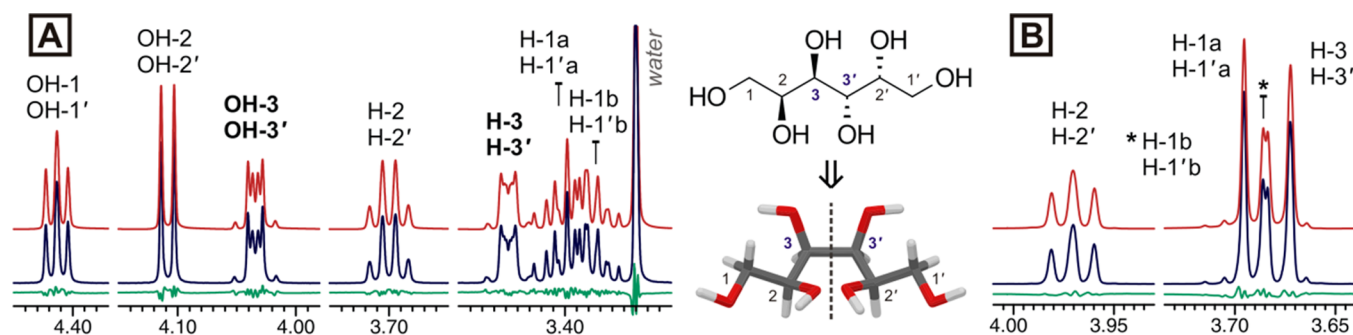
The development of characteristic HiFSA profiles relies on two essential components: first, the molecular structure, from which preliminary spectral parameters are predicted; and second, the 1D  $^1\text{H}$  NMR spectrum, which is used as reference during the assignment and iteration processes. The following paragraphs describe the preparation of these components and their role in the overall HiFSA workflow.

Molecular structures can be built from scratch using PERCH's molecular modeling software (MMS) or other 3D molecular editors.<sup>11,12</sup> Alternatively, X-ray structures deposited in the Cambridge Structural Database,<sup>13</sup> the Crystallography Open Database,<sup>14</sup> and the Protein Data Bank (PDB)<sup>15</sup> can be used as starting points. Particular attention must be paid to relative configuration and conformation, as both affect the outcome of the subsequent prediction step. This also provides an interface to the probing of alternative structures of the target molecule (e.g., stereoisomers). In addition, alternative minimum-energy structures must be considered, as well as multiple forms of the analyte in solution (e.g., anomers in reducing sugars).

The acquisition of high-quality NMR data is also critical. Careful sample preparation plus meticulous attention to acquisition parameters and postacquisition processing is necessary to achieve good line shape and high signal-to-noise ratio. Alternatively, NMR data can be obtained from web-based resources. The Human Metabolome Database (HMDB),<sup>16</sup> the

Received: June 13, 2013

Published: September 6, 2013



**Figure 1.**  $^1\text{H}$  fingerprints of galactitol in different solvents: (A)  $^1\text{H}$  NMR spectrum in  $\text{DMSO-}d_6$  (22 mM, 600 MHz); (B) 1D NOESY spectrum in water (12 mM, 10%  $\text{D}_2\text{O}$ , pH 7.0, 600 MHz). The calculated and observed NMR spectra are shown in red and blue, respectively. Residuals are shown in green.

Madison–Qingdao Metabolomics Consortium Database (MMCD),<sup>17</sup> the Biological Magnetic Resonance Data Bank,<sup>18</sup> and the Birmingham Metabolite Library<sup>19</sup> maintain repositories of raw NMR data for common metabolites. In addition, ChemSpider<sup>20</sup> and the Spectral Database for Instructors<sup>21</sup> contain growing collections of NMR spectra of small molecules.

ACA creates the nexus between the molecular structure and the NMR spectrum in a sequential manner. The structure is analyzed using molecular mechanics geometry optimization and Monte Carlo and molecular dynamics simulations to explore the conformational space. A subset of the conformers generated is used to define average chemical environments for the individual nuclei, which are utilized by PERCH's prediction engine to calculate  $\delta$  values as well as the magnitude and sign of  $J$  couplings. The prediction process is based on a semiempirical model that considers relative configuration, potential intramolecular interactions, solvent, pH, and ionic strength. Subsequently, coupling networks are generated and a trial spectrum is calculated using the predicted parameters and quantum-mechanical formalisms. This process includes the application of the X approximation<sup>1</sup> and splitting of large spin systems into manageable subsystems to reduce computation time.<sup>10</sup>

ACA evaluates potential sets of assignments (i.e., solutions) by comparing and refining the predicted spectral patterns against those in the experimental NMR spectrum. This process involves the use of quantum-mechanical total line shape (QMTLS) iterators.<sup>10</sup> Two main fitting modes are utilized: the integral transform mode for fast correction of  $\delta$  and  $J$  values and the total-line-shape mode for fine adjustment of the  $\delta$ ,  $J$ , and  $\Delta\nu_{1/2}$  parameters.<sup>7</sup> The analysis is repeated until the calculated spectrum resembles the experimental observations, and potential solutions are ranked according to the similarities between the predicted and optimized parameters.

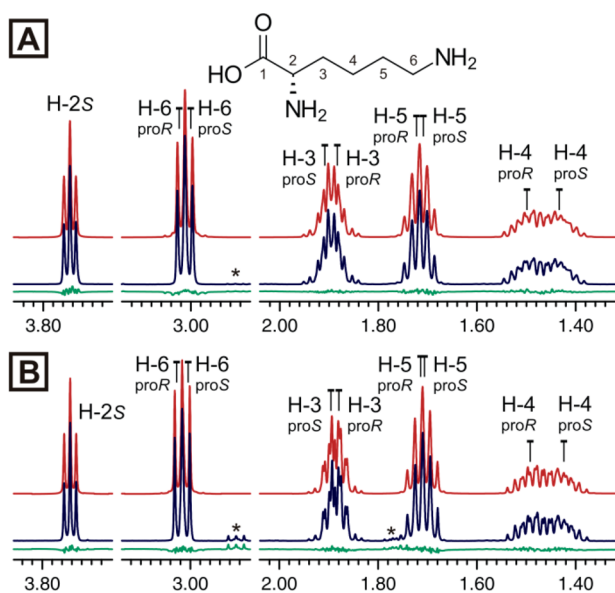
If the automated process is unable to find a consistent solution, the analysis can be repeated using 1D  $^{13}\text{C}$  NMR spectra and 2D  $^1\text{H}$ ,  $^{13}\text{C}$ -HSQC experiments as constraints. Moreover, an interactive spectral analysis can be performed through the ACA graphical user interface (ACA GUI), allowing the user to modify peak lists and integration areas, revise assignments, and manually adjust spectral parameters using information from additional NMR experiments (e.g., 1D selective TOCSY, 2D  $^1\text{H}$ ,  $^1\text{H}$ -COSY, and  $^1\text{H}$ ,  $^{13}\text{C}$ -HMBC). This procedure also can be performed using the spectral parameter editor and the parameter iterator built within the PERCH shell.

To exemplify the applicability of this approach, comprehensive HiFSA profiles of a set of organic compounds with molecular weights ranging from 150 to 1000 Da were generated. The examination of galactitol illustrates the analysis of chemical equivalence and higher-order effects. The  $^1\text{H}$  NMR spectrum in  $\text{DMSO-}d_6$  exhibits distorted signals at 3.448 and 4.034 ppm (Figure 1A). The former corresponds to the chemically equivalent (and spin-coupled) protons H-3 and H-3', whereas the latter belongs to the exchangeable protons OH-3 and OH-3', which are also affected by the strong coupling between H-3 and H-3'. In this case, HiFSA enabled a complete depiction of the unusual resonance patterns and the determination of difficult to measure coupling constants such as  $^3J_{\text{H-3,OH-3}}$  and  $^3J_{\text{H-3,H-3'}}$  (7.4 and 8.8 Hz, respectively).

Although HiFSA profiles are inherently solvent-dependent, they can be used as a foundation to build new profiles in different solvents. For instance, the HiFSA profile of galactitol in  $\text{DMSO-}d_6$  was used to explain the 1D NOESY spectrum in  $\text{D}_2\text{O}$  (Figure 1B). The spectral parameters from  $\text{DMSO-}d_6$  were fitted to the 1D NOESY spectrum and iterated, thereby creating an  $^1\text{H}$  fingerprint of galactitol in  $\text{D}_2\text{O}$  despite extensive signal overlap in the 3.6–3.8 ppm region (Figure 1B).

The analysis of L-lysine demonstrates that HiFSA can provide a detailed description of spin systems containing multiple coupled methylenes (Figure 2A). HiFSA enabled the relative assignment of prochiral protons and the determination of coupling constants in signals that otherwise would be described as multiplets. As the identification of spectral patterns is essential in NMR-based metabolomic studies,<sup>22,23</sup> the development of HiFSA profiles for endogenous metabolites will facilitate the NMR analysis of biofluids. Moreover, this approach can be used to refine an existing profile to consider variations in chemical shifts due to changes in analyte concentration, pH, and temperature (Figure 2B). The proposed assignments for L-lysine correspond to the best solution obtained by ACA, which not only yielded spectral parameters consistent with the molecular structure but also exhibited the lowest root-mean square deviation (rmsd) from the experimental results. Alternative solutions and their corresponding rmsd values are summarized in the Supporting Information.

This example also indicates that HiFSA profiles can reproduce NMR spectra recorded at different field strengths. This concept was introduced by Tiainen et al.,<sup>24</sup> who used adaptive spectral libraries to simulate the NMR spectra of amino acid isotopomers. We extended these results by using the 600 MHz HiFSA profile of ginkgolide A<sup>25</sup> to calculate the corresponding NMR spectrum at 60 MHz. Remarkably, the



**Figure 2.** Comparison of two  $^1\text{H}$  fingerprints of L-lysine: (A)  $^1\text{H}$  NMR spectrum (100 mM, 99.9%  $\text{D}_2\text{O}$ , pH 7.4, 400 MHz); (B) 1D NOESY spectrum (5.5 mM, 10%  $\text{D}_2\text{O}$ , pH 7.0, 500 MHz). The HiFSA profile in (A) was used to generate the profile in (B). The calculated spectra, observed spectra, and residuals are shown in red, blue, and green, respectively. (\*) denotes impurity signals.

calculated low-field spectrum was in excellent agreement with the experimental data (Figure 3). This correlation between high-field spectra and calculated low-frequency fingerprints should have considerable applications in reaction and process monitoring.<sup>26,27</sup>

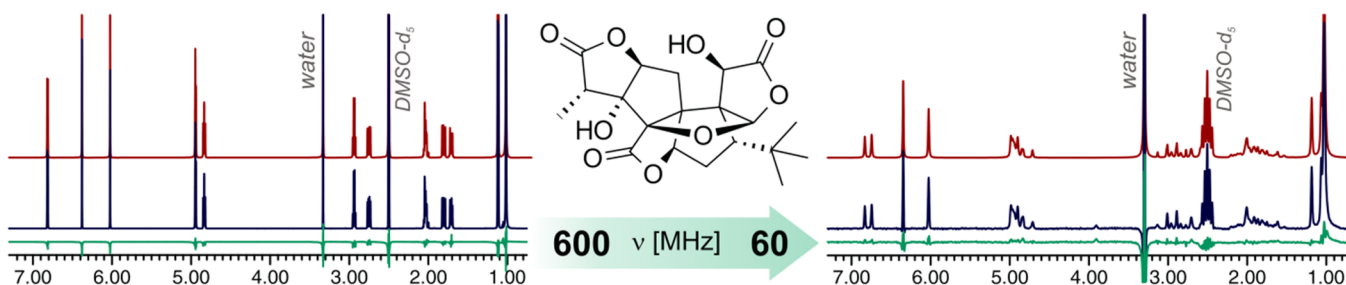
To test this approach on more complex molecules, thorough spectral analyses of progesterone and atosiban were performed. Progesterone contains a spin network that comprises 23 protons and more than 40 coupling constants, including ubiquitous geminal and vicinal couplings ( $^2,3J$ ) and several long-range couplings ( $^{\gamma}J$ ) that are essential in the interpretation of the experimental data.<sup>28,29</sup> The complex resonance patterns observed in the NMR spectrum, even at 900 MHz (Figure 4), were fully characterized using only the 1D  $^1\text{H}$  NMR data and a 2D  $^1\text{H}, ^1\text{H}$ -COSY experiment. This methodology enabled the assignment of all axial and equatorial protons in the ring system. Furthermore, HiFSA facilitated the analysis of distorted multiplets due to higher-order effects (Figure 4). For instance, although H-9 (1.032 ppm) is affected by the strongly coupled protons H-11 $\beta$  and H-12 $\alpha$  (1.519 and 1.504 ppm, respectively),

all of its  $\delta$  and  $J$  values were readily determined (see the Supporting Information).

The analysis of atosiban demonstrates the characterization of individual amino acids in a moderately sized peptide. Semiautomated spectral analysis was carried out with ACA using the 1D  $^1\text{H}$  NMR spectrum and a 2D  $^1\text{H}, ^{13}\text{C}$ -HSQC experiment. Additional 2D NMR experiments ( $^1\text{H}, ^1\text{H}$ -COSY and  $^1\text{H}, ^{13}\text{C}$ -HMBC) were used to verify all of the assignments, and the final optimization was carried out with the PERCH shell. As a result, HiFSA enabled the description of the NMR spectrum of atosiban, revealing the contribution of each constituent unit to the overall signal profile. HiFSA also facilitated the detection of impurities, including acetic acid and a structurally related peptide (Figure 5).

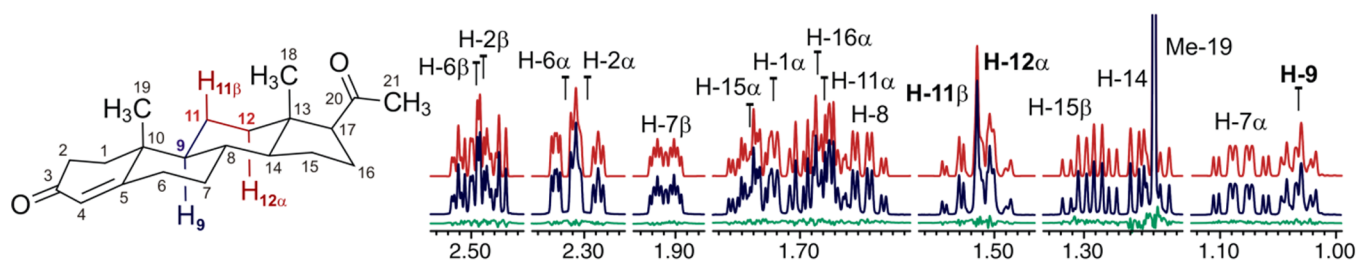
The examples described provide compelling evidence for the ability of HiFSA to resolve and fully assign individual resonances in complex NMR signal patterns. However, although the use of ACA is convenient, the automated analysis has limitations. Imperfections in the spectra (e.g., poor shimming or phasing) or sample impurities can affect the iteration process. Spectroscopic imperfections can be minimized by careful manual or gradient shimming and the use of proper postacquisition processing such as that proposed for qHNMR analysis.<sup>30</sup> While impurities may require specific attention, HiFSA is capable of addressing them both qualitatively and quantitatively, as recently demonstrated for mixtures of natural products.<sup>6,31</sup> Similarly, the success of the automated assignment protocol relies on the accuracy of the spectral prediction. Using a database of fully analyzed experimental data, ACA takes into account averaging on the NMR time scale by mapping the conformational space and correlating conformational parameters. Therefore, HiFSA does not require iterative optimization of NMR parameters for each single conformer. As is the case for any method that produces indirect structural evidence, ACA solutions must be reviewed to ensure that (i) an excellent fit is achieved, (ii) the assignments are correct, and (iii) the spectral parameters are consistent with the molecular structure.

A major strength of HiFSA fingerprinting is the ability to distinguish between closely related molecules, as recently shown for four silybin isomers with near-identical NMR spectra.<sup>31</sup> Moreover, HiFSA profiles are stored in manageable, easy-to-share text files (2–20 kB in size). These profiles are especially suitable for rapid identification of known structures, with a broad range of applications for monitoring synthetic reactions and, as databases of HiFSA profiles pass critical size thresholds,<sup>32–34</sup> for structural dereplication in natural product research. This approach also complements the information

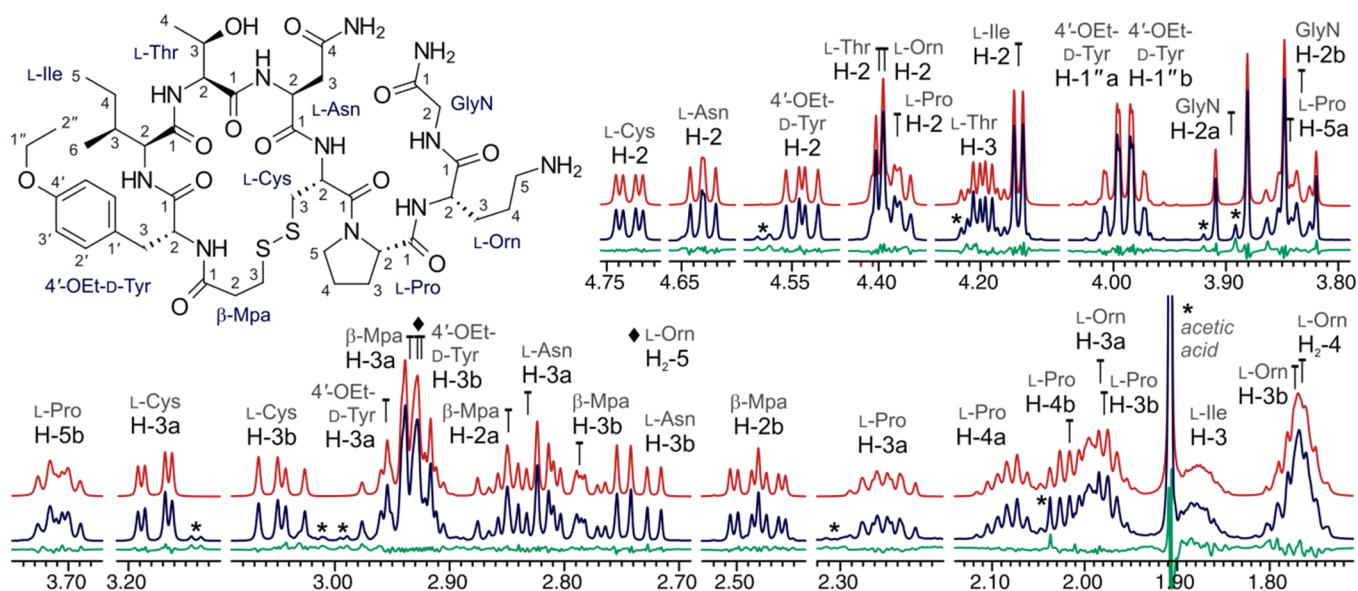


**Figure 3.** Simulation of NMR spectra of ginkgolide A in  $\text{DMSO}-d_6$  at different field strengths. Spectral parameters obtained at 600 MHz were used to calculate the NMR spectrum at 60 MHz. In both cases, excellent agreement between the calculated (red) and observed (blue) spectra was observed. Residuals are shown in green.





**Figure 4.** Sections of the calculated (red) and observed (blue) NMR spectra of progesterone (45 mM, methanol- $d_4$ , 900 MHz). Residuals are shown in green.



**Figure 5.** Sections of the calculated (red) and observed (blue) NMR spectra of atosiban (28 mM, methanol- $d_4$ , 600 MHz). Residuals are shown in green. (\*) denotes impurity signals.

obtained from 2D NMR experiments for characterization of new molecular entities, reducing the exploration of the configurational and conformational space to those structures compatible with the extracted spectral parameters. Moreover, as the recognition of characteristic  $^1\text{H}$  resonances becomes critical for mixture analysis in organic chemistry, forensic science, foodstuff quality assessment, and metabolomic research, the development of reproducible, digital NMR profiles will enable full exploitation of 1D  $^1\text{H}$  NMR for qualitative and quantitative analysis.

In summary, the computational approach described here enhances the description of small molecules and generates highly reproducible  $^1\text{H}$  fingerprints while yielding comprehensive NMR profiles, including all assignments and spectral parameters. Overall, these profiles are highly descriptive and provide insight into the relative configuration and conformation of small molecules in solution, and they ultimately represent a means of portraying molecular structures.

## EXPERIMENTAL SECTION

**Materials and Sample Preparation.** High-purity analytes, deuterated solvents, and NMR tubes were obtained from commercial suppliers. All organic compounds were used as received without further purification. Samples for NMR analysis were prepared by precisely weighing 1–10 mg of the analyte on an analytical balance ( $\pm 0.01$  mg) and then adding 600  $\mu\text{L}$  of deuterated solvent. The samples were placed in a vortex mixer for 2 min to ensure complete dissolution as judged by visual inspection and then transferred to 5

mm NMR tubes using gas-tight syringes. NMR spectra were recorded without sealing the NMR tubes or degassing the samples (see additional comments in the Supporting Information).

**Molecular Structures.** The 3D molecular structures of galactitol and L-lysine were built from scratch in PERCH's MMS module. The structure of progesterone was obtained from its crystal structure bound to the human progesterone receptor (PDB entry 1A28).<sup>35</sup> The 3D molecular model of atosiban was generated with Maestro software (v. 9.0.211) using the crystal structure of deamino-oxytocin (PDB entry 1XY2)<sup>36</sup> as a template.

**NMR Spectroscopy.** The 1D  $^1\text{H}$  NMR spectrum of L-lysine recorded in 99.9%  $\text{D}_2\text{O}$  was downloaded from MMCD (<http://mmcd.nmr.fam.wisc.edu>; entry cq\_00029; expnmr\_00082). The 1D NOESY spectra (ttnoesy pulse sequence) of galactitol and L-lysine, acquired in water with 10%  $\text{D}_2\text{O}$  and water presaturation, were downloaded from HMDB (<http://www.hmdb.ca>; entries HMDB00107 and HMDB00182, respectively).

NMR measurements were recorded at 600.13 and 899.94 MHz (14.1 and 21.1 T static magnetic fields, respectively) on NMR spectrometers equipped with 5 mm triple-resonance inverse detection TXI and TCI cryoprobes, respectively. NMR experiments were acquired under temperature-controlled conditions at 298 K (25  $^\circ\text{C}$ ), and the probes were frequency-tuned and impedance-matched prior to each sample run. The 60 MHz  $^1\text{H}$  NMR spectrum of ginkgolide A was recorded on a permanent-magnet NMR spectrometer operating at 60.01 MHz (1.4 T) with a dual-channel (C/H) probe maintained at 303 K (30  $^\circ\text{C}$ ). Chemical shifts ( $\delta$ ) are expressed in parts per million with reference to internal TMS or DSS (0.000 ppm). In cases where no internal reference was present, the residual solvent signal was used

as the reference. Scalar coupling constants ( $J$ ) and effective linewidths ( $\Delta\nu_{1/2}$ ) are given in hertz.

The 1D  $^1\text{H}$  NMR spectra were recorded under quantitative conditions using a  $90^\circ$  pulse experiment. The  $90^\circ$  pulse was calibrated by evaluating the null at  $360^\circ$  and back-calculating the corresponding pulse width. The following acquisition parameters were used: spectral width ( $sw$ ) of 30 ppm, center of the spectral window ( $o1p$ ) at 7.5 ppm, acquisition time ( $aq$ ) of 4.0 s, and a relaxation delay ( $d1$ ) of 60.0 s (more than 5 times the longest  $T_1$ ). For NMR experiments recorded at 900 MHz, eight transients were collected with 216 798 total data points and a receiver gain of 64. NMR experiments at 600 MHz were recorded with 64 transients, 143 882 total data points, and a receiver gain of 16.

NMR data were processed with TopSpin software (v. 3.2) or NUTS software (v. 201004) using a Lorentz-to-Gaussian window function for resolution enhancement (line broadening =  $-0.3$  Hz, Gaussian factor = 0.05). Zero filling was applied prior to Fourier transformation to increase the number of data points to 256k and 1024k in experiments recorded at 600 and 900 MHz, respectively. The resulting NMR spectra were subjected to manual phase adjustment and automatic baseline correction using polynomial functions.

The 2D NMR experiments were recorded at 600.13 MHz with 2k data points in  $F_2$  and 256 increments in  $F_1$ . Magnitude-mode  $^1\text{H}$ ,  $^1\text{H}$ -COSY experiments were recorded using the following parameters:  $sw = 12$  ppm in both  $F_1$  and  $F_2$ ,  $aq = 0.29$  s in  $F_2$ , and  $d1 = 1.0$  s. Phase-sensitive  $^1\text{H}$ ,  $^{13}\text{C}$ -HSQC and  $^1\text{H}$ ,  $^{13}\text{C}$ -HMBC experiments were recorded with  $sw = 12$  ppm in  $F_2$ ,  $aq = 0.29$  s in  $F_2$ , and  $d1 = 1.5$  s. In HSQC and HMBC experiments, the spectral widths in  $F_1$  were set to 170 and 220 ppm, respectively. In HMBC experiments, the interpulse delay for evolution of long-range heteronuclear coupling constants ( $d6$ ) was set to 0.065 s (i.e.,  $^2J_{\text{C,H}} = 7.7$  Hz). Subsequent 2D NMR processing was carried out with TopSpin software. Zero filling was applied to both dimensions to obtain matrices of  $4\text{k} \times 4\text{k}$  spectral data points. After Fourier transformation, the 2D NMR experiments were phase-adjusted (if necessary) and baseline-corrected using polynomial functions.

**Computer-Aided NMR Spectral Analysis.** Semiautomated NMR analysis was carried out with PERCH NMR tools (v. 2011.1). The experimental NMR data (i.e., the 1D  $^1\text{H}$  NMR spectrum and, optionally, a 1D  $^{13}\text{C}$  NMR spectrum and/or a 2D  $^1\text{H}$ ,  $^{13}\text{C}$ -HSQC experiment) were imported into PERCH's ACA module, together with the 3D molecular structure file (either in MDL Molfile format or prepared in PERCH's MMS module). ACA performed the complete spectral analysis largely in automation. This process includes postprocessing operations (peak picking and integration), conformational analysis, conformational sampling, and prediction of spectral parameters. Next, ACA evaluated potential solutions by fitting and honing the predicted spectral parameters of each solution against the observed  $^1\text{H}$  NMR spectrum using the QMTLS iterators. In cases where a consistent solution was not achieved in full automation (which is generally affected by spectral quality, molecular dynamics, purity, and overall complexity), the predicted spectral parameters were adjusted manually in the ACA GUI or PERCH shell. The iteration process was then repeated until convergence was reached and the rmsd was  $\leq 0.1\%$ .

## ■ ASSOCIATED CONTENT

### ■ Supporting Information

Spectral profiles and  $^1\text{H}$  NMR fingerprints. This material is available free of charge via the Internet at <http://pubs.acs.org>.

## ■ AUTHOR INFORMATION

### Corresponding Author

\*Tel: +1 (312) 355-1949. Fax: +1 (312) 355-2693. E-mail: [gfp@uic.edu](mailto:gfp@uic.edu).

### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

The authors acknowledge the vision and founding work of Dr. R. Laatikainen in establishing tools for QMTLS analysis. We also thank Dr. B. Ramirez for his assistance in the NMR facility at the UIC Center for Structural Biology (CSB) and Dr. A. Abe (Lake Forest College) for collecting the 60 MHz  $^1\text{H}$  NMR spectrum of ginkgolide A. This research work was funded by NIH through Grant RC2 AT005899 awarded to G.F.P. by NCCAM. During the preparation of this manuscript, J.G.N. was financially supported by the United States Pharmacopeial Convention through the 2012/2013 USP Global Fellowship Awards Program. The construction of the UIC CSB and the 900 MHz NMR spectrometer were funded by NIH through Grant P41 GM068944, awarded to Dr. P. G. W. Gettins by NIGMS.

## ■ REFERENCES

- (1) Abraham, R. J. *The Analysis of High Resolution NMR Spectra*; Elsevier: Amsterdam, 1971; pp 124–261.
- (2) Haigh, C. W. *Annu. Rep. Nucl. Magn. Reson. Spectrosc.* **1971**, *4*, 311–362.
- (3) Abraham, R. J.; Mobli, M. *Modelling  $^1\text{H}$  NMR Spectra of Organic Compounds: Theory, Applications and NMR Prediction Software*; Wiley: Chichester, U.K., 2008; Chapter 2, pp 23–65.
- (4) Parr, W. J. E.; Schaefer, T. *Acc. Chem. Res.* **1980**, *13*, 400–406.
- (5) Schaefer, T. In *Encyclopedia of Magnetic Resonance (eMagRes)*; Harris, R. K., Wasylishen, R. E., Eds.; Wiley: Chichester, U.K., 2012; pp 4726–4736.
- (6) Napolitano, J. G.; Gödecke, T.; Rodríguez-Brasco, M. F.; Jaki, B. U.; Chen, S.-N.; Lankin, D. C.; Pauli, G. F. *J. Nat. Prod.* **2012**, *75*, 238–248.
- (7) Laatikainen, R.; Niemitz, M.; Weber, U.; Sundelin, J.; Hassinen, T.; Vepsäläinen, J. *J. Magn. Reson., Ser. A* **1996**, *120*, 1–10.
- (8) Laatikainen, R.; Niemitz, M.; Malaisse, W. J.; Biesemans, M.; Willem, R. *Magn. Reson. Med.* **1996**, *36*, 359–365.
- (9) Thiele, H.; McLeod, G.; Niemitz, M.; Kühn, T. *Monatsh. Chem.* **2011**, *142*, 717–730.
- (10) Laatikainen, R.; Tiainen, M.; Korhonen, S.-P.; Niemitz, M. In *Encyclopedia of Magnetic Resonance (eMagRes)*; Harris, R. K., Wasylishen, R. E., Eds.; Wiley: Chichester, U.K., 2011; pp 1–12.
- (11) Boyd, D. B. *Rev. Comput. Chem.* **1995**, *6*, 317–354.
- (12) Hanwell, M. D.; Curtis, D. E.; Lonie, D. C.; Vandermeersch, T.; Zurek, E.; Hutchison, G. R. *J. Cheminf.* **2012**, *4*, 1–17.
- (13) Allen, F. H. *Acta Crystallogr., Sect. B* **2002**, *B58*, 380–388.
- (14) Gražulis, S.; Chateigner, D.; Downs, R. T.; Yokochi, A. F. T.; Quirós, M.; Lutterotti, L.; Manakova, E.; Butkus, J.; Moeck, P.; Le Bail, A. *J. Appl. Crystallogr.* **2009**, *42*, 726–729.
- (15) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. *Nucleic Acids Res.* **2000**, *28*, 235–242.
- (16) Wishart, D. S.; Tzur, D.; Knox, C.; Eisner, R.; Guo, A. C.; Young, N.; Cheng, D.; Jewell, K.; Arndt, D.; Sawhney, S.; Fung, C.; Nikolai, L.; Lewis, M.; Coutouly, M.-A.; Forsythe, I.; Tang, P.; Shrivastava, S.; Jeroncic, K.; Stothard, P.; Amegbey, G.; Block, D.; Hau, D. D.; Wagner, J.; Miniaci, J.; Clements, M.; Gebremedhin, M.; Guo, N.; Zhang, Y.; Duggan, G. E.; MacInnis, G. D.; Weljie, A. M.; Dowlatabadi, R.; Bamforth, F.; Clive, D.; Greiner, R.; Li, L.; Marrie, T.; Sykes, B. D.; Vogel, H. J.; Querengesser, L. *Nucleic Acids Res.* **2007**, *35*, D521–D526.
- (17) Cui, Q.; Lewis, I. A.; Hegeman, A. D.; Anderson, M. E.; Li, J.; Schulte, C. F.; Westler, W. M.; Eghbalnia, H. R.; Sussman, M. R.; Markley, J. L. *Nat. Biotechnol.* **2008**, *26*, 162–164.
- (18) Ulrich, E. L.; Akutsu, H.; Doreleijers, J. F.; Harano, Y.; Ioannidis, Y. E.; Lin, J.; Livny, M.; Mading, S.; Maziuk, D.; Miller, Z.; Nakatani, E.; Schulte, C. F.; Tolmie, D. E.; Kent Wenger, R.; Yao, H.; Markley, J. L. *Nucleic Acids Res.* **2008**, *36*, D402–D408.

- (19) Ludwig, C.; Easton, J. M.; Lodi, A.; Tiziani, S.; Manzoor, S. E.; Southam, A. D.; Byrne, J. J.; Bishop, L. M.; He, S.; Arvanitis, T. N.; Günther, U. L.; Viant, M. R. *Metabolomics* **2012**, *8*, 8–18.
- (20) Pence, H. E.; Williams, A. J. *Chem. Educ.* **2010**, *87*, 1123–1124.
- (21) Kalstabakken, K. A.; Harned, A. M. *J. Chem. Educ.* **2013**, *90*, 941–943.
- (22) Verpoorte, R.; Choi, Y. H.; Kim, H. K. *Phytochem. Rev.* **2007**, *6*, 3–14.
- (23) Lindon, J. C.; Nicholson, J. K. *Annu. Rev. Anal. Chem.* **2008**, *1*, 45–69.
- (24) Tiainen, M.; Maaheimo, H.; Niemitz, M.; Soininen, P.; Laatikainen, R. *Magn. Reson. Chem.* **2008**, *46*, 125–137.
- (25) Napolitano, J. G.; Lankin, D. C.; Chen, S.-N.; Pauli, G. F. *Magn. Reson. Chem.* **2012**, *50*, 569–575.
- (26) Edwards, J. C.; Giammatteo, P. J. In *Process Analytical Technology: Spectroscopic Tools and Implementation Strategies for the Chemical and Pharmaceutical Industries*, 2nd ed.; Bakeev, K. A., Ed.; Wiley: Chichester, U.K., 2010; pp 303–335.
- (27) Dalitz, F.; Cudaj, M.; Maiwald, M.; Guthausen, G. *Prog. Nucl. Magn. Reson. Spectrosc.* **2012**, *60*, 52–70.
- (28) Kolehmainen, E.; Laihia, K.; Laatikainen, R.; Vepsäläinen, J.; Niemitz, M.; Suontamo, R. *Magn. Reson. Chem.* **1997**, *35*, 463–467.
- (29) Niemitz, M.; Laatikainen, R.; Chen, S.-N.; Kleps, R.; Kozikowski, A. P.; Pauli, G. F. *Magn. Reson. Chem.* **2007**, *45*, 878–882.
- (30) Pauli, G. F.; Gödecke, T.; Jaki, B. U.; Lankin, D. C. *J. Nat. Prod.* **2012**, *75*, 834–851.
- (31) Napolitano, J. G.; Lankin, D. C.; Graf, T. N.; Friesen, J. B.; Chen, S.-N.; McAlpine, J. B.; Oberlies, N. H.; Pauli, G. F. *J. Org. Chem.* **2013**, *78*, 2827–2839.
- (32) Moco, S.; Tseng, L.-H.; Spraul, M.; Chen, Z.; Vervoort, J. *Chromatographia* **2006**, *64*, 503–508.
- (33) Roslund, M. U.; Tähtinen, P.; Niemitz, M.; Sjöholm, R. *Carbohydr. Res.* **2008**, *343*, 101–112.
- (34) Roslund, M. U.; Säwén, E.; Landström, J.; Rönnols, J.; Jonsson, K. H. M.; Lundborg, M.; Svensson, M. V.; Widmalm, G. *Carbohydr. Res.* **2011**, *346*, 1311–1319.
- (35) Williams, S. P.; Sigler, P. B. *Nature* **1998**, *393*, 392–396.
- (36) Wood, S. P.; Tickle, I. J.; Treharne, A. M.; Pitts, J. E.; Mascarenhas, Y.; Li, J. Y.; Husain, J.; Cooper, S.; Blundell, T. L.; Hruby, V. J.; Buku, A.; Fischman, A. J.; Wyssbrod, H. R. *Science* **1986**, *232*, 633–636.