J|A|C|S
ARTICLES
Published on Web 06/28/2006

Molecular Structure Determination in Powders by NMR Crystallography from Proton Spin Diffusion

Bénédicte Elena, Guido Pintacuda, Nicolas Mifsud, and Lyndon Emsley*

Contribution from the Laboratoire de Chimie, UMR 5182 CNRS/ENS, Laboratoire de Recherche Conventionné du CEA (DSV 23V / DSM 0432), Ecole Normale Supérieure de Lyon, 69364 Lyon, France

Received April 6, 2006; E-mail: Lyndon.Emsley@ens-lyon.fr

Abstract: The inability to determine molecular structures from powdered samples is a key barrier to progress in many areas of molecular and materials science. We report an approach to structure determination that combines molecular modeling with experimental spin diffusion data obtained from the high-resolution solid-state nuclear magnetic resonance of protons, and which allows the determination of the three-dimensional structure of an organic compound, in powder form and at natural isotopic abundance.

Introduction

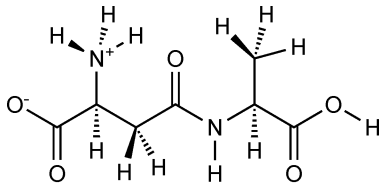
The ability to determine three-dimensional atomic or molecular structures by X-ray diffraction on single-crystal samples is the keystone on which much of our understanding of chemistry has developed over the last century. Today single-crystal diffraction methods (with either X-rays or neutrons) are capable of characterizing systems as diverse as membrane proteins,[1,2] whole virus particles,[3] complex inorganic materials,[4] supramolecular nanostructures,[5,6] or even transient time-resolved structures.[7,8] In contrast, if the sample is a powder, structural characterization represents an enormous challenge. Samples can be powders either because of their intrinsic nature (in the case of pharmaceutical preparations for example) or commonly because crystals large enough for diffraction cannot be formed. Such samples are becoming increasingly widespread, notably in the area of new materials, and the development of experimental methods to study the three-dimensional atomic structure of powdered solids is thus an area of great current interest.

Significant recent progress has been made in the application of X-ray and neutron diffraction methods,[9] and spectacular advances have been made in solid-state Nuclear Magnetic Resonance (NMR) methods for powdered solids.[10] This latter technique has recently led to the first determination of the three-dimensional structure of a powdered microcrystalline protein,[11] to structural models of amyloid fibrils,[12−14] and to models for extended inorganic networks.[15−18] Curiously, in some ways small molecular compounds are harder to handle by NMR than macromolecules. For example, macromolecules can often be treated without considering the crystalline environment, and sophisticated isotopic labeling schemes are available to facilitate spectral assignment and geometry measurements. Structural studies of small molecules, especially at natural isotopic abundance, remain challenging.

Here, we report the determination of the three-dimensional structure of an organic compound, in powder form and at natural isotopic abundance, obtained by an approach that combines molecular modeling with experimental spin diffusion data obtained from the high-resolution solid-state nuclear magnetic resonance of protons. We use a simple framework for a quantitative analysis of directly detected ^{1}H−^{1}H correlations due to spin diffusion, combined with molecular modeling. The approach is demonstrated with powdered (microcrystalline) β-L-aspartyl-L-alanine (**1**) (Scheme 1), where we determine the crystal structure to within an rmsd of 0.33 Å of the known coordinates.

(1) Edman, K.; Nollert, P.; Royant, A.; Belrhali, H.; Pebay-Peyroula, E.; Hajdu, J.; Neutze, R.; Landau, E. M. *Nature* **1999**, *401*, 822−826.
(2) Palczewski, K.; Kumasaka, T.; Hori, T.; Behnke, C. A.; Motoshima, H.; Fox, B. A.; Le Trong, I.; Teller, D. C.; Okada, T.; Stenkamp, R. E.; Yamamoto, M.; Miyano, M. *Science* **2000**, *289*, 739−745.
(3) Abrescia, N. G. A.; Cockburn, J. J. B.; Grimes, J. M.; Sutton, G. C.; Diprose, J. M.; Butcher, S. J.; Fuller, S. D.; Martin, C. S.; Burnett, R. M.; Stuart, D. I.; Bamford, D. H.; Bamford, J. K. H. *Nature* **2004**, *432*, 68−74.
(4) Trikalitis, P. N.; Rangan, K. K.; Bakas, T.; Kanatzidis, M. G. *Nature* **2001**, *410*, 671−675.
(5) Chichak, K. S.; Cantrill, S. J.; Pease, A. R.; Chiu, S. H.; Cave, G. W. V.; Atwood, J. L.; Stoddart, J. F. *Science* **2004**, *304*, 1308−1312.
(6) Bretonniere, Y.; Mazzanti, M.; Pecaut, J.; Olmstead, M. M. *J. Am. Chem. Soc.* **2002**, *124*, 9012−9013.
(7) Schotte, F.; Lim, M. H.; Jackson, T. A.; Smirnov, A. V.; Soman, J.; Olson, J. S.; Phillips, G. N.; Wulff, M.; Anfinrud, P. A. *Science* **2003**, *300*, 1944−1947.
(8) Ihee, H.; Lorenc, M.; Kim, T. K.; Kong, Q. Y.; Cammarata, M.; Lee, J. H.; Bratos, S.; Wulff, M. *Science* **2005**, *309*, 1223−1227.
(9) Harris, K. D. M.; Cheung, E. Y. *Chem. Soc. Rev.* **2004**, *33*, 526−538.
(10) Laws, D. D.; Bitter, H. M. L.; Jerschow, A. *Angew. Chem., Int. Ed.* **2002**, *41*, 3096−3129.
(11) Castellani, F.; van Rossum, B.; Diehl, A.; Schubert, M.; Rehbein, K.; Oschkinat, H. *Nature* **2002**, *420*, 98−102.
(12) Petkova, A. T.; Ishii, Y.; Balbach, J. J.; Antzutkin, O. N.; Leapman, R. D.; Delaglio, F.; Tycko, R. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 16742−16747.
(13) Tycko, R. *Curr. Opin. Struct. Biol.* **2004**, *14*, 96−103.
(14) Jaroniec, C. P.; MacPhee, C. E.; Bajaj, V. S.; McMahon, M. T.; Dobson, C. M.; Griffin, R. G. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 711−716.
(15) Brouwer, D. H.; Darton, R. J.; Morris, R. E.; Levitt, M. H. *J. Am. Chem. Soc.* **2005**, *127*, 10365−10370.
(16) Fyfe, C. A.; Diaz, A. C.; Grondey, H.; Lewis, A. R.; Forster, H. *J. Am. Chem. Soc.* **2005**, *127*, 7543−7558.
(17) Brouwer, D. H.; Kristiansen, P. E.; Fyfe, C. A.; Levitt, M. H. *J. Am. Chem. Soc.* **2005**, *127*, 542−543.
(18) Beitone, L.; Huguenard, C.; Gansmuller, A.; Henry, M.; Taulelle, F.; Loiseau, T.; Ferey, G. *J. Am. Chem. Soc.* **2003**, *125*, 9102−9110.

Scheme 1. β-L-Aspartyl-L-alanine



(**1**)

High-Resolution Proton NMR in Solids and Back Calculation of Spin Diffusion Curves

In the solid state, the presence of strong dipolar couplings between protons considerably broadens the spectral resonances, even under magic angle spinning (MAS). Despite this handicap the study of proton−proton contacts in the solid-state has nonetheless been shown to be particularly valuable, when possible, and such proton−proton NMR constraints are increasingly exploited for characterization of isotopically enriched molecular systems,[19−24] mostly microcrystalline proteins, using isotopic dilution and indirect detection schemes. Spiess and coworkers have shown that double-quantum (DQ) proton NMR spectra, obtained under fast MAS, can have very important applications in nonbiological systems.[25,26]

The use of combined rotation and multiple-pulse techniques (CRAMPS)[27] and considerable advances in the field of homonuclear dipolar decoupling have recently made the direct acquisition of highly resolved proton spectra possible.[28,29] Notably, the study of high-resolution ^{1}H−^{1}H transfers can be of major interest for the characterization of molecules at natural abundance, as illustrated for example by recent applications of high-resolution ^{1}H−^{1}H DQ MAS experiments.[30,31]

Proton spin diffusion (PSD) is a ubiquitous process in solids,[32] whereby magnetization is exchanged between protons according to a process driven by the internuclear distance dependent dipolar coupling. It has long been recognized that this provides in principle a probe of internuclear distances and therefore structures. However, the rate of spin diffusion also depends on the orientation of the internuclear vector in the sample, the details of the anisotropic chemical shifts of the two coupled nuclei, the coupling to other protons, and experimental factors such as the magic angle spinning rate.[32] In a powder under MAS it is not evident to calculate spin diffusion curves from trial structures to compare with experiment.

To circumvent this problem, we recently proposed to model spin diffusion with a phenomenological multispin kinetic rate matrix approach, summed over the structure.[33] In this model the rate of exchange between two types of spin i and j is given by

$$k_{ij} = A \sum_{\lambda} \left(\frac{\mu_0 \gamma_H{}^2 \hbar}{4\pi} \right)^2 \frac{1}{(r_{ij}{}^n)_\lambda} \qquad \text{for } i \neq j \tag{1}$$

where μ_0, γ, and $\hbar$ are the physical constants, where r_{ij} is the internuclear distance between atoms i and j, and where λ indicates the sum over exchange between sites i and j in different molecules in the crystalline lattice. A is a phenomenological scaling factor. The exponent n can in principle be a variable, but here it will be set to $n = 6$. The peak intensities, P_{ij}, observed in a two-dimensional exchange spectrum are then given by[33,34]

$$P_{ij}(\tau_{SD}) = \exp[-\mathbf{K}\,\tau_{SD}]_{ij}\, M^0_{zj} \tag{2}$$

where $\mathbf{K}$ is an $N \times N$ matrix of the rates k_{ij} of exchange between the N different resonances in the spectrum, where τ_{SD} is the spin diffusion mixing time, and where M^0_{zj} is the intensity of the jth peak at $\tau_{SD} = 0$.

This simplified model for spin diffusion, within its validity, provides a way of back calculating spectra from trial structures *and, therefore, a way of determining structures by evaluating the best fit between the data and trial structures*. We previously showed that if we assume that the molecular conformation is known, then this does provide a way to determine the crystalline unit cell parameters.[33] However, the more important question, which we address here, is if this model is sufficiently robust and sensitive to determine the molecular conformation inside a given unit cell.

Results and Discussion

Figure 1a shows experimental spin diffusion data for **1** obtained under magic angle spinning at 6.25 kHz and using eDUMBO-$1_{12.5}$ homonuclear decoupling[35] as described in the Materials and Methods section. The figure highlights the evolution as a function of the mixing time of 9 of the 49 peaks in the spectrum. The experimental data are compared to the predicted curves for two different trial geometries, which differ from each other by an rmsd of 0.12 Å. Clearly, we can evaluate that there are small but significant differences in the predicted spin diffusion behavior for these two structures and that the structure corresponding to the green curves provides a better fit to experiment. More precisely Figure 1b shows the variation in the least squares deviation of the curves predicted for the known crystal structure[36] from the experimental data as we introduce random variations into the data at the estimated level

(19) de Boer, I.; Bosman, L.; Raap, J.; Oschkinat, H.; de Groot, H. J. M. *J. Magn. Reson.* **2002**, *157*, 286−291.
(20) Reif, B.; van Rossum, B. J.; Castellani, F.; Rehbein, K.; Diehl, A.; Oschkinat, H. *J. Am. Chem. Soc.* **2003**, *125*, 1488−1489.
(21) Lange, A.; Seidel, K.; Verdier, L.; Luca, S.; Baldus, M. *J. Am. Chem. Soc.* **2003**, *125*, 12640−12648.
(22) Gehman, J. D.; Paulson, E. K.; Zilm, K. W. *J. Biomol. NMR* **2003**, *27*, 235−259.
(23) Zheng, L.; Fishbein, K. W.; Griffin, R. G.; Herzfeld, J. *J. Am. Chem. Soc.* **1993**, *115*, 6254−6261.
(24) Seidel, K.; Etzkorn, M.; Sonnenberg, L.; Griesinger, C.; Sebald, A.; Baldus, M. *J. Phys. Chem. A* **2005**, *109*, 2436−2442.
(25) Geen, H.; Titman, J. J.; Gottwald, J.; Spiess, H. W. *Chem. Phys. Lett.* **1994**, *227*, 79−86.
(26) Brown, S. P.; Spiess, H. W. *Chem. Rev.* **2001**, *101*, 4125−4155.
(27) Gerstein, B. C.; Pembleton, R. G.; Wilson, R. C.; Ryan, L. M. *J. Chem. Phys.* **1977**, *66*, 361−362.
(28) Vinogradov, E.; Madhu, P. K.; Vega, S. *Chem. Phys. Lett.* **2002**, *354*, 193−202.
(29) Lesage, A.; Sakellariou, D.; Hediger, S.; Elena, B.; Charmont, P.; Steuernagel, S.; Emsley, L. *J. Magn. Reson.* **2003**, *163*, 105−113.
(30) Madhu, P. K.; Vinogradov, E.; Vega, S. *Chem. Phys. Lett.* **2004**, *394*, 423−428.
(31) Brown, S. P.; Lesage, A.; Elena, B.; Emsley, L. *J. Am. Chem. Soc.* **2004**, *126*, 13230−13231.
(32) Meier, B. H. *Adv. Magn. Opt. Reson.* **1994**, *18*, 1−116.
(33) Elena, B.; Emsley, L. *J. Am. Chem. Soc.* **2005**, *127*, 9140−9146.
(34) Abel, E. W.; Coston, T. P. J.; Orrell, K. G.; Sik, V.; Stephenson, D. *J. Magn. Reson.* **1986**, *70*, 34−53.
(35) Elena, B.; De Paepe, G.; Emsley, L. *Chem. Phys. Lett.* **2004**, *398*, 532−538.
(36) Gorbitz, C. H. *Acta Chem. Scand. B* **1987**, *41*, 679−685.

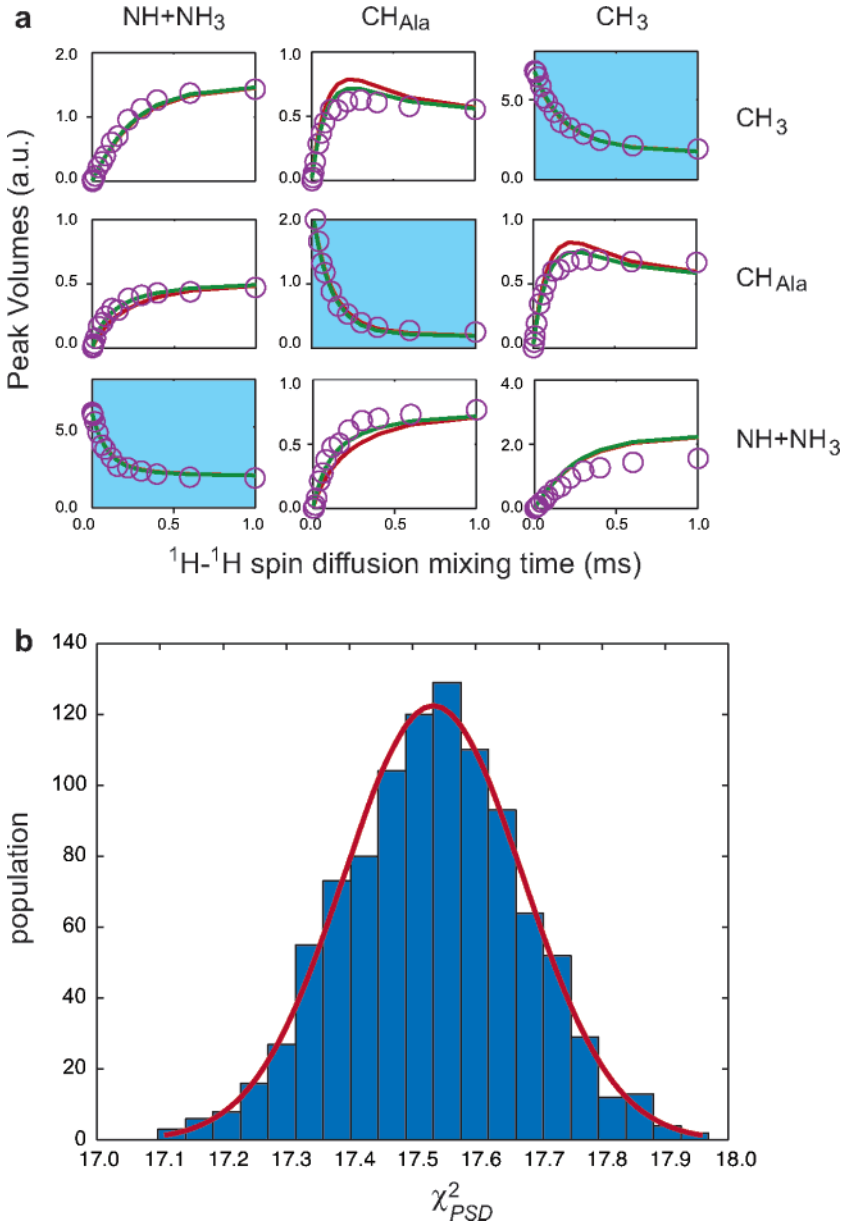


Figure 1. (a) Measured peak volumes as a function of spin diffusion mixing time for 9 of the 49 peaks observed in the 2D spin diffusion experiment for compound **1**. The measured values are compared to fits to the curves (in red and green) expected from the model of eq 2 for two trial structures that differ from each other by an rmsd of 0.12 Å. (b) Measured distribution of the values of χ^2_{PSD} for the comparison between the data and the known crystal structure, as random changes in the data are introduced. 1000 artificial datasets were generated by adding noise to the experimental data from a Gaussian distribution with a standard deviation of 1% of the volume of the OH peak at $\tau = 0$, which is the estimated value of the noise induced error in the measure of the cross-peak volumes. The red curve is a Gaussian function with full width at half-height of 0.35.

of the noise. The deviation is designated as

$$\chi_{PSD}{}^2 = \sum_i \frac{(calc_i - expt_i)^2}{\sigma_i{}^2} \qquad (3)$$

where σ_i is the estimated error on data point i. In this case we see that the data are of sufficient quality to detect changes in the χ^2_{PSD} of ~0.35. This means we can detect a significant difference between two structures that change the value of the χ^2_{PSD} by that amount. It is important to note that this level of structural sensitivity is obtained through full back calculation of the complete PSD buildup curves. Calculation of the intial rates is not sufficient to enable this analysis.

We have integrated the full back-calculation of the spectrum from a trial structure, and the evaluation of the difference with respect to the experimental data, as an external routine into the Xplor-NIH molecular modeling (MM) package.[37,38] Within this framework, we can optimize trial geometries by combining the internal MM energy term, which maintains reasonable covalent geometry, and a standard van der Waals potential, which maintains physically reasonable crystal structures, and the PSD pseudo energy (which is the measured χ^2 obtained by back calculation from the trial structure using the model of eqs 1 and 2).

(37) Schwieters, C. D.; Kuszewski, J. J.; Tjandra, N.; Clore, G. M. *J. Magn. Reson.* **2003**, *160*, 66−74.

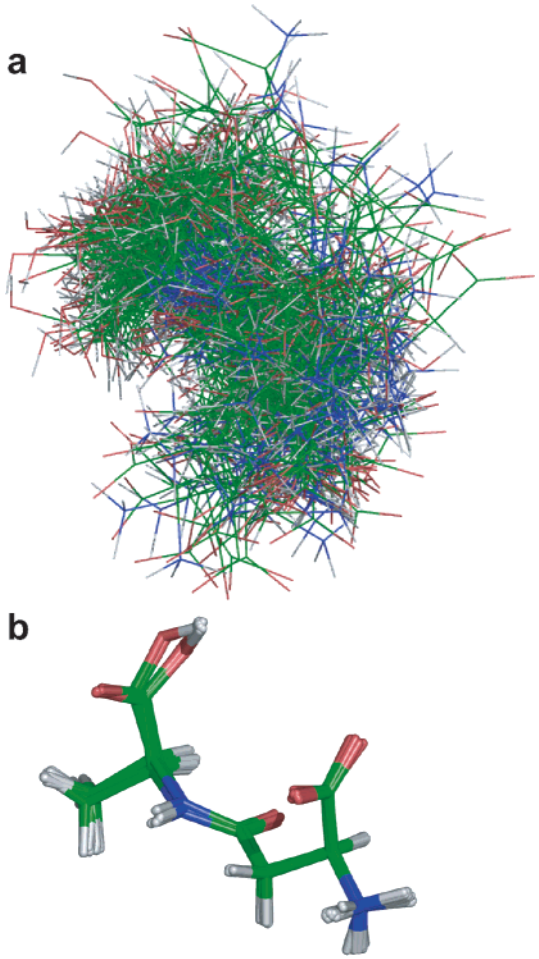


Figure 2. (a) A set of 150 structures from the ensemble of 3000 random structures used as the starting point for the structure refinement. (b) The 16 structures determined with the lowest E_{PSD} values after the optimization procedure described in the text.

Figure 2a shows a representative part of an ensemble of 3000 random starting structures roughly compatible with possible crystalline packing, generated as described in the Materials and Methods below. Note that these structures, which are randomly generated, do not occupy space in a random fashion but already limit the structure determination to more or less physically reasonable structures, thereby introducing a significant reduction in the structural space that needs to be explored using the experimental PSD constraints. Figure 2 shows the result of refining these structures using an energy given by

$$E_{tot} = E_{Xplor} + E_{PSD} \quad (4)$$

where E_{Xplor} is a standard MM force field including standard covalent terms, van der Waals, and electrostatic terms, and where $E_{PSD} = a_{PSD}\,\chi^2_{PSD}$ with a_{PSD} being the weight of the experimental constraints. E_{tot} is used as a cost function in a simplex minimization protocol of the Xplor-NIH package. The back-calculated PSD energy, E_{PSD}, was evaluated for all 3000 members of the ensemble, and the 200 structures which had the lowest values were then refined using the simplex procedure against E_{tot}. Figure 2b shows the 16 structures with the lowest

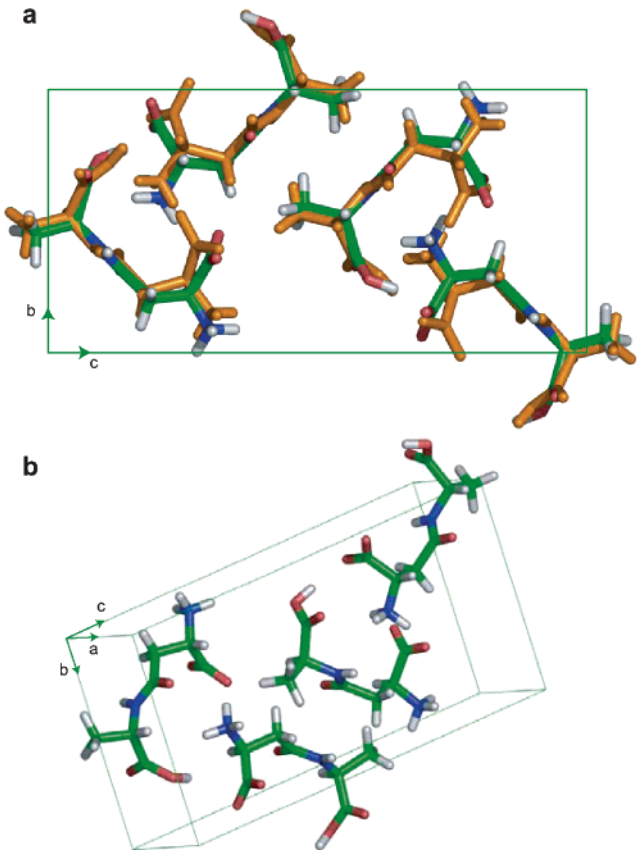


Figure 3. (a) Comparison between the known crystal structure of **1** in orange and the structure obtained from the average coordinates of the 16 structures determined here with the lowest E_{PSD} values as described in the text. The view is taken along the *a* axis of the unit cell. (b) A second view of the average structure determined here, illustrating the packing arrangement from another angle.

values of E_{PSD} after the refinement procedure and the exclusion of structures with unreasonably large intramolecular van der Waals energies. These structures deviate from each other by an ensemble rmsd for all atoms of only 0.09 Å (excluding the protons on the CH_3 and NH_3 groups) and deviate from the known crystal structure by an all atom rmsd of only 0.33 Å.

Figure 3 illustrates the crystal structure and packing determined here as represented by the average coordinates of the 16 lowest E_{PSD} structures. Figure 3a shows the deviation of the average structure found here from the known crystal structure. This clearly illustrates that, *at least in favorable conditions, and if the experimental data is of sufficient quality*, PSD data can be used to determine crystal structures in powdered samples at natural abundance. As such, this work constitutes a significant step further in the emerging domain of NMR crystallography.[15−17,33,39−41]

Furthermore, a particularly interesting factor is that this structure determination protocol uses proton positions as the basis for the determination. In that sense it is highly complementary to X-ray diffraction studies, which are guided primarily

(38) Brünger, A. T.; Adams, P. D.; Clore, G. M.; DeLano, W. L.; Gros, P.; Grosse-Kunsteleve, R. W.; Jiang, J.-S.; Kuszewski, J.; Nilges, M.; Pannu, N. S.; Read, R. J.; Rice, L. M.; Simonson, T.; Warren, G. L. *Acta Crystallogr., Sect. D* **1998**, *54*, 901−921.

(39) Interestingly, Crystallography is often assimilated today to X-ray studies on single crystals, due to the phenomenal success of this method. Crystallography is obviously a much wider discipline, defined (according to the Encyclopaedia Britannica) as "the branch of science that deals with discerning the arrangement and bonding of atoms in crystalline solids and with the geometric structure of crystal lattices." Since the powders we study here are microcrystalline, the term NMR Crystallography appears natural.

(40) Harris, R. K. *Solid State Sci.* **2004**, *6*, 1025−1037.

(41) Senker, J.; Seyfarth, L.; Voll, J. *Solid State Sci.* **2004**, *6*, 1039−1052.

by heavy atoms, and where proton positions are the hardest to determine. Notably, when we visualize the whole crystal lattice as determined here, in Figure 3b, we can see that the OH and NH protons are positioned in reasonable intermolecular hydrogen bonding configurations. This is determined directly in the procedure and does not rely on positioning the protons once the heavy atoms have been determined (rather the inverse). We also remark that diffraction and NMR data are further complementary in that where diffraction is by nature a global effect, and probes the whole lattice directly, into which one has to position the individual nuclei, NMR in contrast provides a local, atomic level probe, with interactions that "reach out" to determine the lattice.

The structure we determine here appears to have a systematic deviation from the X-ray determined structure. We can imagine several sources of this error, most of which could be improved in the future. The most obvious are that the spin diffusion data, which are the fruit of current state of the art NMR experiments, may contain artifacts in the intensity of certain peaks. Indeed, other data sets we have recorded, where we are conscious of problems in the spectra, lead to varying degrees of success in the convergence of the refinement, indicating that high-quality data are essential to obtain good structures. Structural deviations may also be induced because the model we use to fit the data is not correct. Apart from the validity of the single-exponential approximation, the geometrical model may itself be wrong, as we have approximated the motion of the protons of the CH_3 and NH_3 groups (which from the NMR spectrum are clearly undergoing rapid motion) by a simple fast three-site jump model (which is achieved by simply setting the exchange rate constants for magnetization transfer between these protons to be 2 orders of magnitude faster than the other, dipolar driven, exchange rates). This three-site jump model (slightly) improves the fit quality with respect to a single configuration. However, it is quite probable that these protons are undergoing continuous motion around the C_3 axis and that a model reproducing this type of dynamics would provide even better structures.

Finally, we remark that we have consciously made very little attempt to optimize the MM force field parameters, which could also lead to induction of unfavorable structures. This is particularly the case for the orientation of the two carboxyl groups present in this molecule. The orientation of the carboxyl groups depends only little on the proton positions and so is largely determined by the force field we use here. This highlights again how this NMR-based structure is defined by the proton positions, which are themselves found to be in remarkably good agreement with the known structure.

Conclusion

We have introduced a combined MM/NMR-PSD approach, using a model for the back calculation of the full PSD buildup curves, which enables us to determine the molecular structure of **1** *at natural abundance and in powdered form*, from a randomly generated starting ensemble to obtain a group of structures with a 0.09 Å rmsd, and which deviate on average from the known structure by only 0.33 Å. All the sources of error discussed above can be improved by further developing the model for PSD and the experimental methods used to acquire the spectra. We are thus confident that this method will improve and become more widespread in the future.

Materials and Methods

NMR Spectroscopy. The acquisition of the $^1H-^1H$ spin diffusion spectra analyzed here has already been reported in detail in ref 33. The data were acquired from high-resolution two-dimensional $^1H-^1H$ correlation spectra recorded at a Larmor frequency of 500 MHz using a single channel 2.5 mm MAS probe with about 10 mg of powdered sample, with homonuclear decoupling fields of $\nu_1 = 100$ kHz and the magic angle spinning frequency $\nu_r = 6.25$ kHz. The buildup curves used for the analysis here were obtained by integrating peak volumes in a series of 14 2D spectra recorded with spin diffusion mixing times τ_{SD} distributed from 2 μs to 1 ms.

β-L-Aspartyl-L-alanine was bought from Bachem (Bübendorf, Switzerland) and used without further purification or recrystallization.

Structure Calculations. Back calculation of the spectral intensities from trial coordinates, and comparison with the data, was achieved by a home written C++ routine. Comparison with the data to provide the value of χ^2_{PSD} used a fitting routine that only adjusted the value of A in eq 1 to provide the lowest possible value of χ^2_{PSD} for each trial structure. The C++ program was then interfaced to the Python scripting language using the SWIG interpreter and integrated into the Python scripting framework of Xplor-NIH as a new potential term, using the Xplor-NIH pyPot module.[37] In the structure calculations shown here a CHARMM19/22 empirical energy function was employed, with standard periodic van der Waals and electrostatic terms. The crystalline environment and the periodic interactions are enabled in Xplor through the "xrefin" statement. Calculations were carried out using the following unit cell parameters, $P2_12_12_1$, $a = 4.845$ Å, $b = 9.409$ Å, $c = 19.170$ Å, $\alpha = \beta = \gamma = 90°$.[36] There are four symmetry-related (NMR equivalent) molecules in the unit cell. The back calculation took account of all pairs of atoms within a radius of 10 Å of the atoms in the central molecule.

Each of the members of the starting ensemble of 3000 structures was generated by starting from a "template" coordinate set having an arbitrary extended conformation with an ideal geometry, obtained from randomly generated coordinates, on which dynamics and minimization were run to anneal the structure to a reasonable covalency. Then a standard molecular dynamics protocol was applied to "shuffle" the structure into an ensemble compatible with crystalline packing in three steps: (i) a high temperature (3000 K) loop with covalent restraints but no van der Waals repulsion term; (ii) a 100-step cycle with the inter- and intramolecular van der Waals radius and force constant being incrementally raised along with changes in the force constants for angles and improper dihedral angles; and (iii) a cooling loop, where all the weights were incrementally raised to their final values. 3000 structures out of a total of 45 000 were then selected on the basis of their van der Waals energy. The rmsd from the average for this group of 3000 random structures was 2.3 Å.

The structure refinement process was then carried out in five steps. Note that these steps constitute an empirical "recipe," constrained in most part by the absence of an analytical gradient function for the PSD energy term (see below). It is by no means intended as a definitive protocol and will certainly evolve in the future. First, an optimization of all 3000 structures was performed by systematically varying the CH_3, NH_3, and OH dihedral angles and retaining the structure with the lowest value of E_{PSD}. The 200 structures with the lowest E_{PSD} in the resulting ensemble were then optimized using E_{tot} in an all-atom simplex minimization of up to 3000 steps. The value of a_{PSD} was 35, and only intermolecular van der Waals terms were used (with no electrostatics). The ensemble of 200 structures was then regularized by a Powell type minimization using only E_{Xplor} with the same parameters. The ensemble was then reoptimized by a second simplex minimization using E_{tot} with van der Waals and electrostatics forces including both inter- and intramolecular terms, and finally the CH_3, NH_3, and OH dihedral angles were reoptimized against E_{PSD} alone. The 16 structures with the lowest E_{PSD} values were used in the ensemble shown in Figure 2b, and they vary in χ_{PSD}^2 from 15.9 to 16.2.

Simplex minimizations are, by far, not the most efficient structural optimization tools. However, we have currently not implemented a calculation of the gradient of the E_{PSD} function with respect to the atomic coordinates, and all other routines (Powell, Simulated Annealing, ...) require a gradient in order to function. Note that calculation of the gradient by finite differences is not practical, since it would greatly increase the number of calls required to the back calculation routine. As a guide, currently the optimization of 200 structures, as described above, takes about 6 h of CPU time on a 4 × 2.5 GHz Macintosh G5 computer running Xplor-NIH under MAC OSX, although it should be noted we have made little or no attempt so far to optimize our code. Evaluation of the analytical form of the gradient function is under way, and we are confident it will improve significantly the range of structures we can refine.

Acknowledgment. We are grateful to Dr. Charles Schwieters (NIH) for his help in getting to know Xplor-NIH.

Supporting Information Available: PDB files with the coordinates of the determined structures. This material is available free of charge via the Internet at http://pubs.acs.org.

JA062353P