

## Supramolecular Protein Structure Determination by Site-Specific Long-Range Intermolecular Solid State NMR Spectroscopy

Andrew J. Nieuwkoop and Chad M. Rienstra\*

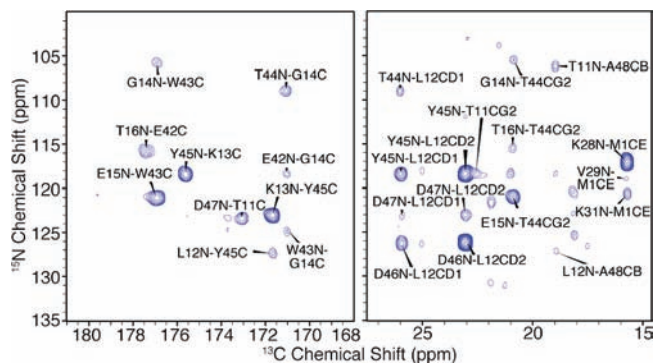
Department of Chemistry, University of Illinois at Urbana–Champaign, 600 South Mathews Avenue, Urbana, Illinois 61801

Received February 3, 2010; E-mail: rienstra@scs.uiuc.edu

Solid-state NMR (SSNMR) spectroscopy is a uniquely effective method for protein structure determination of insoluble protein fibrils or aggregates.<sup>1–4</sup> SSNMR spectra of such samples often contain intramolecular and intermolecular correlations with similar intensities, leading to ambiguities in interpretation of data sets obtained from undiluted, uniformly <sup>13</sup>C,<sup>15</sup>N-labeled samples. To faithfully reproduce tertiary and quaternary structural features, differential isotopic labeling strategies must be utilized. For example, intensities of intermolecular crosspeaks are attenuated by diluting the labeled sample in natural abundance material. Alternatively, these signals can be accentuated by utilizing regio-specific <sup>15</sup>N and <sup>13</sup>C labeling, as demonstrated in studies of reassembled thioredoxin,<sup>5</sup> or recrystallizing physical mixtures of <sup>15</sup>N and <sup>13</sup>C labeled proteins, such as Crh.<sup>6</sup> With such labeling patterns, heteronuclear distance techniques, such as REDOR,<sup>7</sup> TEDOR,<sup>8</sup> and NHHC methods,<sup>6</sup> can be applied to obtain site-specific distance restraints across molecular interfaces.

Here we extend this approach to the full, atomic-resolution structure determination of a quaternary protein assembly in the nanocrystalline state. We utilize 3D Z-filtered TEDOR,<sup>9</sup> as recently demonstrated to quantify intramolecular distances in proteins,<sup>10</sup> to detect site-resolved intermolecular correlations. The distance restraints are incorporated into simulated annealing calculations, resulting in a specific quaternary arrangement that is independent of the initial condition. For example, if the coordinates from the known trigonal X-ray crystal structure (PDB ID 2QMT)<sup>11</sup> are used as the starting point for the calculation, the NMR restraints refine the exact atom positions but do not change the quaternary arrangement; if instead the orthorhombic X-ray lattice coordinates are used (PDB ID 2GI9),<sup>12</sup> in the course of the calculation, the relative orientations of neighboring molecules change substantially, transforming into the trigonal form. Finally, calculations initialized with isolated monomers converge to the trigonal form. These results demonstrate that SSNMR methods can fully reproduce not only secondary and tertiary, but also quaternary, structural features at atomic-resolution detail.

Previously we have collected a range of structural data for GB1 in the nanocrystalline state, including homonuclear distance restraints and vector angles,<sup>13</sup> backbone chemical shift tensors,<sup>14</sup> and high-precision <sup>15</sup>N–<sup>13</sup>C distance restraints.<sup>10</sup> The 3D Z-filtered TEDOR pulse sequence produces hundreds of intramolecular <sup>15</sup>N–<sup>13</sup>C restraints, which are sufficient (with TALOS<sup>15</sup> restraints) to define an atomic resolution protein structure (PDB ID 2KQ4).<sup>10</sup> Extending this approach to physical mixtures (50:50) of <sup>13</sup>C- and <sup>15</sup>N-labeled molecules (see Supporting Information (SI) for sample preparation details) results in a rich set of intermolecular restraints. The spectra were collected with a 500 MHz Varian InfinityPlus spectrometer and 3.2-mm Balun <sup>1</sup>H–<sup>13</sup>C–<sup>15</sup>N probe, with ~100 kHz TPPM decoupling<sup>16</sup> during the REDOR periods (see SI for complete experimental details). Under these conditions, the <sup>15</sup>N and

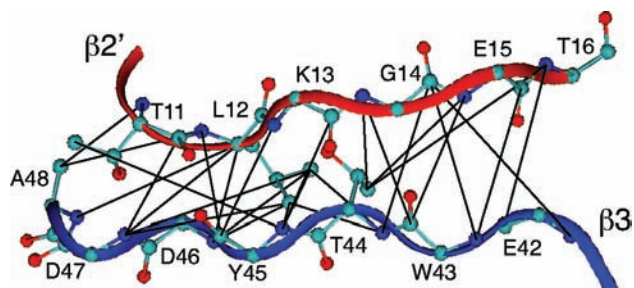


**Figure 1.** 2D <sup>15</sup>N–<sup>13</sup>C plane of 3D ZF-TEDOR spectrum, illustrating intermolecular correlations from a recrystallized physical mixture (50:50) of 1,3-<sup>13</sup>C-glycerol and <sup>15</sup>N-labeled GB1. Data were acquired at 500 MHz (<sup>1</sup>H) with 21.6 ms <sup>15</sup>N–<sup>13</sup>C mixing, 14.4 ms <sup>15</sup>N and 30.7 ms <sup>13</sup>C evolution. Total measurement time 26 h.

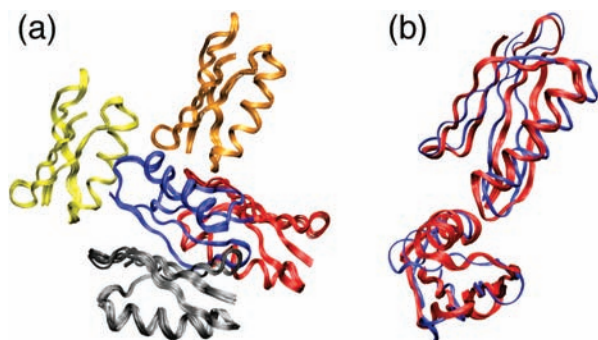
<sup>13</sup>C (especially methyl and carbonyl)  $T_2$  values are >25 ms, coherence lifetimes that in combination with the sparse labeling pattern (mitigating effects of scalar <sup>13</sup>C–<sup>13</sup>C couplings) enabled the detection of particularly strong TEDOR cross peaks. The peak intensities continue to increase up to mixing times of at least 20 ms (Figure 1, Figure S1). Most of these correlations are intermolecular correlations (natural abundance, intramolecular correlations are observed but are much weaker and can be readily identified with spectra at short mixing times; see SI). Reverse labeling of the <sup>15</sup>N sample with <sup>13</sup>C-depleted glucose would further suppress these undesirable peaks, and this strategy would be beneficial for proteins of higher molecular weight and/or greater spectral degeneracy.

The observed <sup>15</sup>N and <sup>13</sup>C line widths were ~0.5 and 0.2–0.3 ppm, respectively, enabling unique (unambiguous) assignment of several intermolecular correlations based on the known GB1 chemical shifts.<sup>17</sup> From these assignments, we identified two distinct molecular interfaces: (1) an antiparallel intermolecular beta-sheet ( $\beta 2'–\beta 3$ ), defined by long-range correlations with  $(i+j) = 58 \pm 2$  (Figure 2) and (2) contacts between helical residues 28–31 and residues 1 and 20–21. A subset of these correlations have been previously reported.<sup>10,13,18,19</sup> We obtained a sufficiently large number of unambiguous restraints that structure calculations generally converged well and clarified the assignment of ambiguous restraints (see SI for complete restraint lists).

We first performed XPLOR-NIH<sup>20</sup> calculations consisting of one subsection of the crystal lattice: a central GB1 molecule and all other molecules with a <sup>15</sup>N within 8 Å of a <sup>13</sup>C of the central GB1. The calculation was performed starting with each crystal form, applying identical restraints. Intermolecular restraints were assigned taking into account the inherent ambiguities; for each peak, all <sup>15</sup>N and <sup>13</sup>C resonance frequencies within half the peak line width were deemed possible assignments. The ambiguous restraints were then



**Figure 2.** Intermolecular  $^{15}\text{N}$ – $^{13}\text{C}$  distance restraints observed by TEDOR. The  $\beta 2'$ – $\beta 3$  intermolecular interface of GB1 nanocrystals with intermolecular restraints, with maximum distances of 5, 7, and 8 Å, determined by the TEDOR mixing time at which the peaks were first observed (see SI for details).



**Figure 3.** Crystal packing of GB1 determined by simulated annealing of five GB1 monomers with ambiguous intermolecular TEDOR restraints (pdb ID 2KWD). The bbRMSD of the 10 lowest energy structures (a) is  $0.42 \pm 0.08$  Å. Alignment of the region distinct to the trigonal lattice (b) yields a  $2.9 \pm 0.1$  Å bbRMSD of the ensemble (red) to the 2QMT crystal structure (blue).

formalized in XPLOR-NIH lists (provided in the SI). We then performed the calculation with several different initial conditions, to test the reproducibility of the resulting structure, which converged within 5 Å bbRMSD of the trigonal crystal structure (2QMT) but differed by  $>13$  Å from the orthorhombic structure (2GI9).

Finally, we performed a calculation starting with five GB1 monomers placed  $\sim 20$  Å apart (using VMD-XPLOR<sup>21</sup>) in several, randomly selected orientations. Intermolecular restraints were first applied with internal coordinates of the monomers held constant, allowing the intermolecular TEDOR restraints to dock the peripheral monomers to the central monomer. If docking was successful (as it was in half the cases), the structure was then refined. The resulting quaternary arrangements (Figure 3a) showed high precision (bbRMSD for all 5 monomers of  $<0.5$  Å) and accuracy, based on agreement with the trigonal form (Figure 3b). Hence, this is a *de novo* recreation of the crystal lattice of GB1.

In conclusion, we have demonstrated that 3D Z-filtered TEDOR experiments, when performed on mixtures of isotopically labeled protein samples, report on site-specific intermolecular distance

restraints. These data sets can be leveraged to perform rigorous structure calculations of the protein interface. In the example demonstrated here, we have determined the packing arrangement of our nanocrystalline GB1 preparation to be consistent with the trigonal form as determined by X-ray diffraction. Therefore this represents an important proof of principle, in a case where the results can be directly compared with other structural information. We envision the application of this approach to determining the registry and quaternary arrangement of protein fibrils, which most often cannot be determined by diffraction methods.

**Acknowledgment.** We thank the National Institutes of Health (R01-GM073770, an American Recovery and Reinvestment Act supplement, and the Molecular Biophysics Training Program at the University of Illinois) for financial support, Dr. Charles Schwieters for assistance in the application of XPLOR-NIH in multimeric systems, Lindsay Sperling for assistance with recreating crystal lattice coordinates, Damien Mathew for advice in performing VMD scripting, and the School of Chemical Sciences NMR Facility at the University of Illinois.

**Supporting Information Available:** Materials and Methods, additional spectra, graphical overlays of all structures, calculation statistics, restraint lists. This material is available free of charge via the Internet at <http://pubs.acs.org>.

## References

- (1) Petkova, A. T.; Leapman, R. D.; Guo, Z. H.; Yau, W. M.; Mattson, M. P.; Tycko, R. *Science* **2005**, *307*, 262.
- (2) Tycko, R. *Q. Rev. Biophys.* **2006**, *39*, 1.
- (3) Helmus, J. J.; Surewicz, K.; Nadaud, P. S.; Surewicz, W. K.; Jaroniec, C. P. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 6284.
- (4) Wasmer, C.; Lange, A.; Van Melckebeke, H.; Siemer, A. B.; Riek, R.; Meier, B. H. *Science* **2008**, *319*, 1523.
- (5) Yang, J.; Tasayco, M. L.; Polenova, T. *J. Am. Chem. Soc.* **2008**, *130*, 5798.
- (6) Etkorn, M.; Bockmann, A.; Lange, A.; Baldus, M. *J. Am. Chem. Soc.* **2004**, *126*, 14746.
- (7) Gullion, T.; Schaefer, J. *J. Magn. Reson.* **1989**, *81*, 196.
- (8) Hing, A.; Vega, S.; Schaefer, J. *J. Magn. Reson.* **1992**, *96*, 205.
- (9) Jaroniec, C. P.; Filip, C.; Griffin, R. G. *J. Am. Chem. Soc.* **2002**, *124*, 10728.
- (10) Nieuwkoop, A. J.; Wylie, B. J.; Franks, W. T.; Shah, G. J.; Rienstra, C. M. *J. Chem. Phys.* **2009**, *131*, 095101.
- (11) Frericks Schmidt, H. L.; Sperling, L. J.; Gao, Y. G.; Wylie, B. J.; Boettcher, J. M.; Wilson, S. R.; Rienstra, C. M. *J. Phys. Chem. B* **2007**, *111*, 14362.
- (12) Franks, W. T.; Wylie, B. J.; Stellfox, S. A.; Rienstra, C. M. *J. Am. Chem. Soc.* **2006**, *128*, 3154.
- (13) Franks, W. T.; Wylie, B. J.; Frericks Schmidt, H. L.; Nieuwkoop, A. J.; Mayrhofer, R. M.; Shah, G. J.; Graesser, D. T.; Rienstra, C. M. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 4621.
- (14) Wylie, B. J.; Schwieters, C. D.; Oldfield, E.; Rienstra, C. M. *J. Am. Chem. Soc.* **2009**, *131*, 985.
- (15) Cornilescu, G.; Delaglio, F.; Bax, A. *J. Biomol. NMR* **1999**, *13*, 289.
- (16) Bennett, A. E.; Rienstra, C. M.; Auger, M.; Lakshmi, K. V.; Griffin, R. G. *J. Chem. Phys.* **1995**, *103*, 6951.
- (17) Franks, W. T.; Zhou, D. H.; Wylie, B. J.; Money, B. G.; Graesser, D. T.; Frericks, H. L.; Sahota, G.; Rienstra, C. M. *J. Am. Chem. Soc.* **2005**, *127*, 12291.
- (18) Helmus, J. J.; Nadaud, P. S.; Hofer, N.; Jaroniec, C. P. *J. Chem. Phys.* **2008**, *128*, 052314.
- (19) Peng, X. H.; Libich, D.; Janik, R.; Harauz, G.; Ladizhansky, V. *J. Am. Chem. Soc.* **2008**, *130*, 359.
- (20) Schwieters, C. D.; Kuszewski, J. J.; Tjandra, N.; Clore, G. M. *J. Magn. Reson.* **2003**, *160*, 65.
- (21) Schwieters, C. D.; Clore, G. M. *J. Magn. Reson.* **2001**, *149*, 239.

JA100992Y