J|A|C|S
COMMUNICATIONS

Published on Web 03/21/2006

From Gene to HSQC in under Five Hours: High-Throughput NMR Proteomics

Shiranthi Keppetipola,† Wieslaw Kudlicki,† Bao D. Nguyen,‡ Xi Meng,‡ Kevin J. Donovan,‡ and A. J. Shaka*,‡

Invitrogen Corporation, 1610 Faraday Avenue, Carlsbad, California 92008, and Chemistry Department, University of California, Irvine, California 92697-2025

Received November 28, 2005; E-mail: ajshaka@uci.edu

Solution structure determination of recombinant, isotopically enriched proteins[1,2] by NMR[3,4] makes demands on protein purity, homogeneity, and absence of contaminants like proteases. With *Escherichia coli* or baculovirus-mediated insect cell expression, careful protein purification[5] is required for NMR spectroscopy. As the quality of the 2D ^{15}N−^{1}H heteronuclear single quantum coherence (HSQC) spectrum[6] itself reflects NMR suitability,[7] the time to obtain this spectrum, once the relevant gene is identified, measures potential throughput. A "quick and clean" way to do so would thus be a welcome advance. We show such a protocol using a rapid-fire, cell-free, *E. coli*-based in vitro expression system. Compatible with purification using histidine fusion tags and immobilized metal ion affinity chromatography,[8] it can also rapidly produce good NMR samples without time-consuming purification. Expressing SUMO-1[9] in this way, protein quantity was sufficient after a 4 h reaction to obtain 2D ^{15}N−^{1}H HSQC spectra with a conventional 5 mm HCN triple-resonance probe in under 1 h at 800 MHz. In contrast to previous experience with SUMO-1,[9,10] in which protein instability was noted, these samples remained stable in the NMR tube for at least 6 months. We used the filter diagonalization method (FDM)[11−14] to obtain high-resolution 2D NMR spectra. The in vitro expression system, optimized pulse sequence, and FDM combine to provide a good platform for high-throughput NMR proteomics.

The Expressway NMR cell-free *E. coli* expression system (catalog no. K9900-99, Invitrogen Corp., Carlsbad, CA) was used. It comprises high expression vectors with N-terminal (pEXP5-NT/TOPO) or C-terminal (pEXP5-CT/TOPO) histidine tags, high-yield *E. coli* extract, in vitro protein synthesis (IVPS), *E. coli* reaction buffer, 2.5x IVPS feed buffer, T7 enzyme mix, and individual amino acids used for precision labeling. In these experiments, uniformly ^{15}N-labeled SUMO-1 was produced as follows: 2 mL of high-yield *E. coli* extract, 2 mL of 2.5x IVPS reaction buffer, 0.1 mL of T7 enzyme mix, 0.5 mL of 100 mg/mL ^{15}N-labeled amino acid mix (catalog no. NLM-6695, CIL, Andover, MA), and 0.05 mg of pEXP5-CT/TOPO-SUMO-1 were mixed and brought to a final volume of 5 mL with nuclease-free water. The reaction, incubated at 30 °C with shaking at 300 rpm, ran just under 4 h. About 15 min after initiation, 2.5 mL of feed buffer with 10 mg/mL of ^{15}N-labeled amino acid mix was added; after 2 h, another identical aliquot was added to the reaction.

The solution was then centrifuged at 4000*g* for 5 min, loaded onto a pre-equilibrated (binding buffer, 20 mM Tris-HCl, 200 mM NaCl, pH 7.5) Ni−NTA column, and incubated in the column for about 5 min, and the flow-through was collected. Ten column volumes of wash buffer (20 mM Tris-HCl, pH 7.5, 200 mM NaCl), repeated thrice, washed the column. One column volume of elution buffer 1 (20 mM Tris-HCl, pH 7.5, 200 mM NaCl, 20 mM imidazole) was added to the column and eluted the nonspecific proteins bound to the column. One column volume of the elution buffer 2 (same as buffer 1, but 200 mM imidazole) was added and incubated for 5 min, and the protein was eluted. This was repeated twice with one column volume of elution buffer 2. Samples were run on a NuPAGE gel, and elutions containing protein were pooled and carefully loaded to the bed of a Zeba Desalt spin column (catalog no. 89893, Pierce, Rockford, IL). The labeled protein was collected by centrifugation at 1000*g* for 2 min. The column was pre-equilibrated with binding buffer; the flow-through was used for NMR analysis after concentration to 550 μL through a 5 kDa molecular weight cutoff polyethersulfone membrane (Vivaspin 20, Vivascience Inc., New York), pH adjustment to 6.1, and addition of 55 μL of D_2O for deuterium lock.

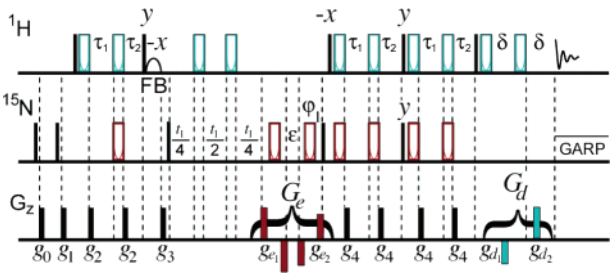

Figure 1. Sensitivity-enhanced ^{15}N−^{1}H HSQC pulse sequence. Pulses have phase *x*, unless otherwise noted, and have leading and trailing delays of 10 μs. Solid black icons on ^{15}N, ^{1}H staves are 90° pulses; on the G_z stave they are small "cleanup" pulsed field gradients (PFGs). Cyan open icons are 60 μs proton BIP 180° elements. A soft flip-back pulse (FB) places the strong H_2O magnetization along $+z$, improving sensitivity.[20] The phase $\varphi_1 = x, y, -x, -y$ with alternate addition and subtraction of scans selects ^{15}N transverse magnetization. The ^{15}N 90° and 180° ^{15}N BIP pulses preceding the pulse at the beginning of t_1 are incremented by 180° at every increment.[21] Decoupling of ^{15}N is accomplished using the GARP sequence.[22] See Supporting Information for further details.

The NMR spectra were obtained with a slightly modified HSQC pulse sequence[6] using pulsed field gradients for coherence-transfer pathway selection[15,16] and with sensitivity enhancement (SE),[17] as shown in Figure 1. To improve performance, conventional 180° pulses were replaced throughout with Broadband Inversion Pulses (BIPs),[18] and the encoding gradient was substituted with a CLUB sequence;[19] where magnetization phase was important, BIPs were substituted in pairs.

As FDM gives absorption mode, rather than "phase-twist"[23] peaks; only "N-type" data were needed, allowing twice as many ^{15}N increments. Using FDM, 2D HSQC SUMO-1 spectra were obtained in ∼40 min at 800 MHz, as shown in Figure 2. A 3D HNCO,[24] using 3D FDM,[13] is predicted to take ∼40 min. With protein purification, the time was ca. 6 h. Crude reaction extract, centrifuged at 4000*g* for 5 min, loaded onto a desalting column, and concentrated to 550 μL as previously described, but *not* purified,

† Invitrogen Corp.
‡ University of California, Irvine.

10.1021/ja0580791 CCC: $33.50

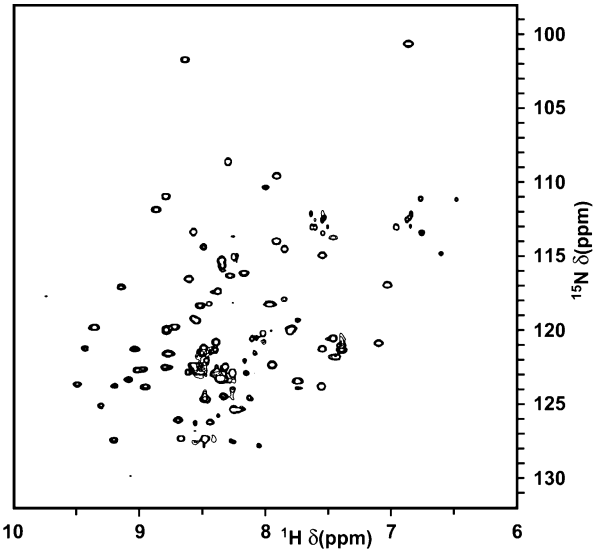

Figure 2. 800 MHz ^{15}N−^{1}H HSQC of SUMO-1 at 25 °C, obtained by multiwindow FDM implemented as a plug-in to the nmrPipe[24] package. A single set of N-type phase-modulated data with 96 increments was used as input. For experiment details see the Supporting Information.

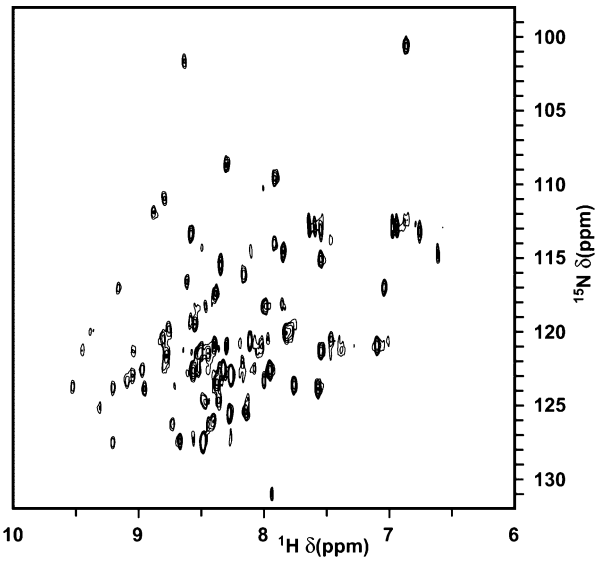

Figure 3. 800 MHz ^{15}N−^{1}H HSQC of crude SUMO-1 reaction mixture, 25 °C, by FT processing of N- and P-type data sets with 48 increments in the ^{15}N dimension. Time from gene to this spectrum is 5 h. A higher resolution FDM spectrum results in the same time, but this conventional FT spectrum shows that label incorporation is efficient and impurity peaks (e.g. at ∼7.95, 131 ppm) are not a concern. See Supporting Information for high-resolution FT data.

gave a viscous solution yielding the spectrum of Figure 3, showing structured protein, in under 5 h total.

The high protein yield shown here gives the ability to screen proteins for correct expression, folding, and solubility by NMR in a matter of hours. This method is thus attractive as a first step in more detailed study of solution structure, dynamics, or ligand binding, using selective labeling of specific amino acids, as has been demonstrated[26−30] using in vitro expression.[31] Cell-free protein expression for NMR is bound to see more use in the future.

Acknowledgment. The authors thank Prof. Melanie Cocco and Mr. E. J. Pone for help and encouragement. This research was made possible by support from the National Institutes of Health, GM-66763. K.J.D. was supported by the National Institutes of Health, National Research Service Award 5 T15 LM007443 from the National Library of Medicine. B.D.N. thanks the UCI School of Physical Sciences for partial support.

Supporting Information Available: Complete ref 7, C source code for the pulse sequence used, Varian format shape files for the ^{1}H and ^{15}N BIPs, high-resolution conventional 2D FT spectrum of SUMO-1 for comparison, and details of the experimental biochemistry of the in vitro system. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) McIntosh, L. P.; Dahlquist, F. W. *Q. Rev. Biophys.* **1990**, *23*, 1−38.
(2) Kay, L. E.; Clore, G. M.; Bax, A.; Gronenborn, A. M. *Science* **1990**, *249*, 411−414.
(3) Cavanagh, J.; Fairbrother, W. J.; Palmer, A. G., III; Skelton, N. J. *Protein NMR Spectroscopy: Principles and Practice*; Academic Press: San Diego, 1996.
(4) Wüthrich, K. *NMR of Proteins and Nucleic Acids (Baker Lecture Series)*; Wiley-Interscience: New York 1986.
(5) *Protein Purification: Principles, High-Resolution Methods, and Applications*; Janson, J.-C., Rydén, L., Eds.; Wiley-VCH: New York, 1998.
(6) Bodenhausen, G.; Ruben, D. J. *Chem. Phys. Lett.* **1980**, *69*, 185−189.
(7) Arrowsmith, C. H.; et al. *Nat. Struct. Biol.* **2000**, *7*, 903−909.
(8) Ueda, E. K. M.; Gout, P. W.; Morganti, L. *J. Chromatogr. A* **2003**, *988*, 1−23.
(9) Bayer, P.; Arndt, A.; Metzger, S.; Mahajan, R.; Melchior, F.; Jaenicke, R.; Becker, J. *J. Mol. Biol.* **1998**, *280*, 275−286.
(10) Jin, C. W.; Shiyanova, T.; Shen, Z. Y.; Liao, X. B. *Biol. Macromol.* **2001**, *28*, 227−234.
(11) Chen, J. H.; Mandelshtam, V. A.; Shaka, A. J. *J. Magn. Reson.* **2000**, *146*, 363−368.
(12) Chen, J. H.; De Angelis, A. A.; Mandelshtam, V. A.; Shaka, A. J. *J. Magn. Reson.* **2003**, *162*, 74−89.
(13) Chen, J. H.; Nietlispach, D.; Shaka, A. J.; Mandelshtam, V. A. *J. Magn. Reson.* **2004**, *169*, 215−224.
(14) Armstrong, G. S.; Cano, K. E.; Mandelshtam, V. A.; Shaka, A. J.; Bendiak, B. *J. Magn. Reson.* **2004**, *170*, 156−163.
(15) Hurd, R. E.; John, B. K. *J. Magn. Reson.* **1991**, *92*, 658−668.
(16) Davis, A. L.; Laue, E. D.; Keeler, J.; Moskau, D.; Lohman, J. *J. Magn. Reson.* **1991**, *94*, 637−644.
(17) Kay, L. E.; Keifer, P.; Saarinen, T. *J. Am. Chem. Soc.* **1992**, *114*, 10663−10665.
(18) Smith, M. A.; Hu, H.; Shaka, A. J. *J. Magn. Reson.* **2001**, *151*, 269−283.
(19) Hu, H.; Shaka, A. J. *J. Magn. Reson.* **1999**, *136*, 54−62.
(20) Grzesiek, S.; Bax, A. *J. Am. Chem. Soc.* **1993**, *115*, 12593−12594.
(21) Hammarström, A.; Otting, G. *J. Magn. Reson. Ser. A* **1994**, *109*, 246−249.
(22) Shaka, A. J.; Barker, P. B.; Freeman, R. F. *J. Magn. Reson.* **1987**, *64*, 145−150.
(23) Bodenhausen, G.; Freeman, R.; Niedermeyer, R.; Turner, D. L. *J. Magn. Reson.* **1977**, *26*, 133−164.
(24) Powers, P.; Gronenborn, A. M.; Clore, G. M.; Bax, A. *J. Magn. Reson.* **1991**, *162*, 209−213.
(25) Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A. *J. Biomol. NMR* **1995**, *6*, 277−293.
(26) Guignard, L.; Ozawa, K.; Pursglove, S. E.; Otting, G.; Dixon, N. E. *FEBS Lett.* **2002**, *524*, 159−162.
(27) Torizawa, T.; Shimuzu, M.; Taoka, M.; Miyano, H.; Masatsune, K. *J. Biomol. NMR* **2004**, *30*, 311−325.
(28) Shi, J.; Pelton, J. G.; Cho, H. S.; Wemmer, D. E. *J. Biomol. NMR* **2004**, *28*, 235−247.
(29) Ozawa, K.; Dixon, N. E.; Otting, G. *IUBMB Life* **2005**, *57*, 615−622.
(30) Trbovic, N.; Klammt, C.; Koglin, A.; Löhr, F.; Berhard, F.; Dötsch, V. *J. Am. Chem. Soc.* **2005**, *127*, 13504−13505.
(31) Katzen, F.; Chang, G.; Kudlicki, W. *Trend Biotechnol.* **2005**, *23*, 150−156.

JA0580791