J|A|C|S
ARTICLES
Published on Web 11/23/2004

Magic Angle Spinning Solid-State NMR Spectroscopy for Structural Studies of Protein Interfaces. Resonance Assignments of Differentially Enriched *Escherichia coli* Thioredoxin Reassembled by Fragment Complementation

Dabeiba Marulanda,[†] Maria Luisa Tasayco,*[,‡] Ann McDermott,[§] Marcela Cataldi,[‡] Vilma Arriaran,[‡] and Tatyana Polenova*[,†]

Contribution from the Department of Chemistry and Biochemistry, Brown Laboratories, University of Delaware, Newark, Delaware 19716, Department of Chemistry, Science Building, City College of the City University of New York, New York, New York 10031, and Department of Chemistry, Havemeyer Hall, Columbia University, New York, New York 10027

Received June 15, 2004; E-mail: tpolenov@chem.udel.edu

Abstract: De novo site-specific backbone and side-chain resonance assignments are presented for U-^{15}N-(1−73)/U-^{13}C,^{15}N(74−108) reassembly of *Escherichia coli* thioredoxin by fragment complementation, determined using solid-state magic angle spinning NMR spectroscopy at 17.6 T. Backbone dihedral angles and secondary structure predicted from the statistical analysis of ^{13}C and ^{15}N chemical shifts are in general agreement with solution values for the intact full-length thioredoxin, confirming that the secondary structure is retained in the reassembled complex prepared as a poly(ethylene glycol) precipitate. The differential labeling of complementary thioredoxin fragments introduced in this work is expected to be beneficial for high-resolution structural studies of protein interfaces formed by protein assemblies by solid-state NMR spectroscopy.

Introduction

The inherent capability of solid-state NMR for probing noncrystalline systems, such as protein aggregates and membrane proteins, intractable by other high-resolution structural techniques, determines the growing interest toward structural applications of solid-state NMR. A number of recent reports demonstrate the potential of solid-state NMR spectroscopy for analysis of complex biological systems, such as amyloid fibrils,[1,2] membrane proteins,[3,4] biomaterials,[5] and intact cells.[6] For the intrinsically soluble macromolecules, solid-state NMR very often offers additional information inaccessible from solution measurements. Internal protein and ligand dynamics on slow time scales,[7] electronic structure and geometry of paramagnetic centers[8] or quadrupolar metal sites,[9,10] and ionization state and hydrogen-bonding geometry[11,12] are directly inferred from the solid-state NMR spectra.

Structural and dynamics investigation of uniformly and extensively isotopically enriched proteins by solid-state NMR spectroscopy is a rapidly emerging field. Significant progress has been made recently by researchers in several laboratories in establishing the experimental protocols for resonance assignments and determination of structural constraints from multidimensional solid-state magic angle spinning spectra. Partial or nearly complete resonance assignments have been reported to date for four uniformly enriched proteins, including ubiquitin,[13−17] bovine pancreatic trypsin inhibitor (BPTI),[18] the α-spectrin SH_3 domain,[19,20] and the catabolite repression histidine-containing phosphocarrier protein Crh.[21]

* To whom correspondence should be addressed. Phone: (302) 831-1968. E-mail: tpolenov@chem.udel.edu.
[†] University of Delaware.
[‡] City College of the City University of New York.
[§] Columbia University.

(1) Tycko, R. *Prog. NMR Spectrosc.* **2003**, *42*, 53−68.
(2) 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.
(3) Park, S. H.; Mrse, A. A.; Nevzorov, A. A.; Mesleh, M. F.; Oblatt-Montal, M.; Montal, M.; Opella, S. J. *J. Mol. Biol.* **2003**, *333*, 409−424.
(4) Nishimura, K.; Kim, S. G.; Zhang, L.; Cross, T. A. *Biochemistry* **2002**, *41*, 13170−13177.
(5) Drobny, G. P.; Long, J. R.; Karlsson, T.; Shaw, W.; Popham, J.; Oyler, N.; Bower, P.; Stringer, J.; Gregory, D.; Mehta, M.; Stayton, P. S. *Annu. Rev. Phys. Chem.* **2003**, *54*, 531−571.
(6) Cegelski, L.; Kim, S. J.; Hing, A. W.; Studelska, D. R.; O'Connor, R. D.; Mehta, A. K.; Schaefer, J. *Biochemistry* **2002**, *41*, 13053−13058.
(7) Krushelnitsky, A. G.; Hempel, G.; Reichert, D. *Biochim. Biophys. Acta: Proteins Proteom.* **2003**, *1650*, 117−127.
(8) Lee, H.; de Montellano, P. R. O.; McDermott, A. E. *Biochemistry* **1999**, *38*, 10808−10813.
(9) Lipton, A. S.; Heck, R. W.; Ellis, P. D. *J. Am. Chem. Soc.* **2004**, *126*, 4735−4739.
(10) Lipton, A. S.; Buchko, G. W.; Sears, J. A.; Kennedy, M. A.; Ellis, P. D. *J. Am. Chem. Soc.* **2001**, *123*, 992−993.
(11) Gu, Z. T.; Drueckhammer, D. G.; Kurz, L.; Liu, K.; Martin, D. P.; McDermott, A. *Biochemistry* **1999**, *38*, 8022−8031.
(12) Emmler, T.; Gieschler, S.; Limbach, H. H.; Buntkowsky, G. *J. Mol. Struct.* **2004**, *700*, 29−38.
(13) Straus, S. K.; Bremi, T.; Ernst, R. R. *J. Biomol. NMR* **1998**, *12*, 39−50.
(14) Hong, M. *J. Biomol. NMR* **1999**, *15*, 1−14.
(15) Igumenova, T. I.; Wand, A. J.; McDermott, A. E. *J. Am. Chem. Soc.* **2004**, *126*, 5323−5331.
(16) Igumenova, T. I. Assignment of uniformly carbon-13-enriched proteins and optimization of their carbon lineshapes. Ph.D. thesis, Columbia University, New York, 2003.
(17) Igumenova, T. I.; McDermott, A. E.; Zilm, K. W.; Martin, R. W.; Paulson, E. K.; Wand, A. J. *J. Am. Chem. Soc.* **2004**, *126*, 6720−6727.
(18) McDermott, A. E.; Polenova, T.; Bockmann, A.; Zilm, K.; Martin, R.; Paulson, E.; Montellione, G. *J. Biomol. NMR* **2000**, *16*, 209−219.

10.1021/ja0464589 CCC: $27.50

In this work, we addressed a complex between complementary fragments U-^{15}N(1−73) and U-^{13}C,^{15}N(74−108) of *Escherichia coli* thioredoxin, by high-resolution solid-state magic angle spinning NMR spectroscopy at 17.6 T. *E. coli* thioredoxin is a 108-amino acid residue disulfide reductase and a member of a large superfamily of multifunctional protein modules. Thioredoxins are present in most organisms from Archea to humans, and are responsible for redox regulation of protein function and signaling via the thiol redox center.[22] Mammalian thioredoxins are critical in defense against oxidative stress, and in regulation of growth and apoptosis, immunomodulation, embryonic implantation, and developmental biology.[22] Thioredoxin is considered the ultimate "moonlighting protein",[22,23] with new functions being discovered. Due to its key role in many different metabolic pathways, the thioredoxin/thioredoxin reductase system has attracted recent attention as a target for a number of chemotherapeutic applications ranging from infectious diseases to cancer therapy.[24,25]

An interesting aspect of *E. coli* thioredoxin in solution is its ability to reassemble in vitro from the complementary fragments (generated by either proteolytic or chemical cleavage of the intact protein at a specific position), to yield a molecule with preserved tertiary structure.[26−28] This behavior is observed for a number of different cleavage sites in thioredoxin. Notably, in a wide variety of organisms thioredoxin fold is conserved, and the protein is very stable thermally and chemically.[29,30] Moreover, some of the thioredoxin reassemblies by fragment complementation also display high thermodynamic stability.[31−33] This is a remarkable achievement and perhaps is related to the fact that this protein has multifaceted roles in biology.

In this context, it is interesting to examine how the local structure and/or dynamic behavior of various thioredoxin reassemblies are altered with respect to the native protein and whether these changes can be correlated to the thermodynamic stability of these complexes. In a related system, a human−*E. coli* chimeric thioredoxin, the origin of the reduced thermodynamic stability was recently probed via structure and dynamics measurements by solution NMR.[34] While the overall structure and fast time scale dynamics were found to be very similar to those of the parent thioredoxin, subtle multiple structural changes were postulated to be responsible for the altered stability. Solid-state NMR methods will allow us to probe a broader range of time scales and assess whether slower motional modes are perturbed in different thioredoxin reassemblies.

In this study, we present the de novo resonance assignments and address the local secondary structure of the C-terminal portion of reassembled [U-^{15}N(1−73)/U-^{13}C,^{15}N(74−108)]-thioredoxin in the solid state, by a combination of homo- and heteronuclear two-dimensional (2D) correlation spectroscopy. The interface consists of two β-strands formed by residues 23−29 and 76−82, belonging to the complementary N- and C-terminal fragments resulting from the cleavage of the intact protein at Arg-73. In Figure 1, the primary amino acid sequence, the secondary structure generated using PDBsum,[35] and the tertiary fold of thioredoxin are shown.

We demonstrate that rapid precipitation of the thioredoxin complex with poly(ethylene glycol) (PEG) under controlled conditions yields excellent quality solid-state magic angle spinning spectra, from which an overwhelming majority of ^{13}C and ^{15}N resonances are readily assigned. This work contributes to the growing body of experimental evidence that a large number of globular proteins are amenable to high-resolution structural studies by multidimensional solid-state NMR spectroscopy as hydrated PEG precipitates.[19,21] We discuss the potential of employing differential labeling schemes for high-resolution structural studies of protein assemblies and protein interfaces by solid-state NMR spectroscopy.

Experimental Section

Materials. All chemicals were obtained from Sigma. [U-$^{13}C_6$]glucose and [^{15}N]ammonium chloride were purchased from Cambridge Isotope Laboratories. Biochemicals for bacterial cultures were purchased from Oxoid Inc. All reagents were used without further purification. The VDX crystallization plates and siliconized coverslips were from Hampton Research.

Expression and Purification of U-^{15}N(1−73)/U-^{13}C,^{15}N(74−108)-Enriched Reassembled Thioredoxin. Overexpression of oxidized wild-type thioredoxin was performed as described previously[36] in *E. coli* JF521 strain (gift of Dr. J. Fuchs and Dr. C. K. Woodward) containing the plasmid pTK100 with the Trx gene. This construct is identical to that used in the previous solution NMR studies by Dyson and co-workers.[37] Overexpression of uniformly labeled oxidized wild-type was performed in *E. coli* BL21 (DE3) strain from Novagen using published procedures.[38] Isotopic labels were introduced using M9 minimal media[39] containing either [^{15}N]ammonium chloride (1 g/L of minimal media) together with natural abundance glucose, or [U-$^{13}C_6$]-glucose (4 g/L of minimal media) together with [^{15}N]ammonium chloride (1 g/L of minimal media). Purification of wild-type thioredoxin with and without the NMR labels was done by molecular exclusion and ionic exchange chromatography using published procedures.[40] Characterization of wild-type thioredoxin was performed using polyacrylamide gel electrophoresis (PAGE) and electrospray mass spectroscopy. The concentration was determined spectrophotometrically using the average molecular mass and the extinction coefficient E280

(19) Pauli, J.; Baldus, M.; van Rossum, B.; de Groot, H.; Oschkinat, H. *ChemBioChem* **2001**, *2*, 272−281.
(20) Castellani, F.; van Rossum, B.; Diehl, A.; Schubert, M.; Rehbein, K.; Oschkinat, H. *Nature* **2002**, *420*, 98−102.
(21) Bockmann, A.; Lange, A.; Galinier, A.; Luca, S.; Giraud, N.; Juy, M.; Heise, H.; Montserret, R.; Penin, F.; Baldus, M. *J. Biomol. NMR* **2003**, *27*, 323−339.
(22) Arnér, E. S. J.; Holmgren, A. *Eur. J. Biochem.* **2000**, *267*, 6102−6109.
(23) Jeffery, C. *Trends Biochem. Sci.* **1999**, *24*, 8−11.
(24) Gromer, S.; Urig, S.; Becker, K. *Med. Res. Rev.* **2004**, *24*, 40−89.
(25) Hirt, R. P.; Muller, S.; Embley, T. M.; Coombs, G. H. *Trends Parasitol.* **2002**, *18*, 302−308.
(26) Yu, W.-F.; Tung, C.-C.; Wang, H.; Tasayco, M. L. *Protein Sci.* **2000**, *9*, 20−28.
(27) Tasayco, M. L.; Chao, K. *Proteins* **1995**, *22*, 41−44.
(28) Yang, X. M.; Yu, W. F.; Li, J. H.; Fuchs, J.; Rizo, J.; Tasayco, M. L. *J. Am. Chem. Soc.* **1998**, *120*, 7985−7986.
(29) Eklund, H.; Gleason, F. K.; Holmgren, A. *Proteins: Struct., Funct. Genet.* **1991**, *11*, 13−28.
(30) Martin, J. L. *Structure* **1995**, *3*, 245−250.
(31) Yang, X. M.; Georgescu, R. E.; Li, J. H.; Yu, W. F.; Haierhan; Tasayco, M. L. *Recognition between disordered polypeptide chains from cleavage of an alpha/beta domain: self-versus non-self-association*; Proceedings of the Pacific Symposium in Biocomputing, 1999; pp 590−600.
(32) Tasayco, M. L.; Fuchs, J.; Yang, X.-M.; Dyalram, D.; Georgescu, R. E. *Biochemistry* **2000**, *39*, 10613−10618.
(33) Georgescu, R. E.; García-Mira, M. D. M.; Tasayco, M. L.; Sánchez-Ruiz, J. M. *Eur. J. Biochem.* **2001**, *268*, 1477−1485.
(34) Dangi, B.; Dobrodumov, A. V.; Louis, J. M.; Gronenborn, A. M. *Biochemistry* **2002**, *41*, 9376−9388.
(35) Laskowski, R. A.; Hutchinson, E. G.; Michie, A. D.; Wallace, A. C.; Jones, M. L.; Thornton, J. M. *Trends Biochem. Sci.* **1997**, *22*, 488−490.
(36) Chaffotte, A. F.; Li, J. H.; Georgescu, R. E.; Goldberg, M. E.; Tasayco, M. L. *Biochemistry* **1997**, *36*, 16040−16048.
(37) Chandrasekhar, K.; Campbell, A. P.; Jeng, M.-F.; Holmgren, A.; Dyson, H. J. *J. Biomol. NMR* **1994**, *4*, 411−432.
(38) Marley, J.; Lu, M.; Bracken, C. *J. Biomol. NMR* **2001**, *20*, 71−75.
(39) Sambrook, J.; Fritsch, E. F.; Maniatis, T. *Molecular Cloning: A Laboratory Manual*; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, 1989; Vol. 3.
(40) Langsetmo, K.; Fuchs, J.; Woodward, C. *Biochemistry* **1989**, *28*, 3211−3220.

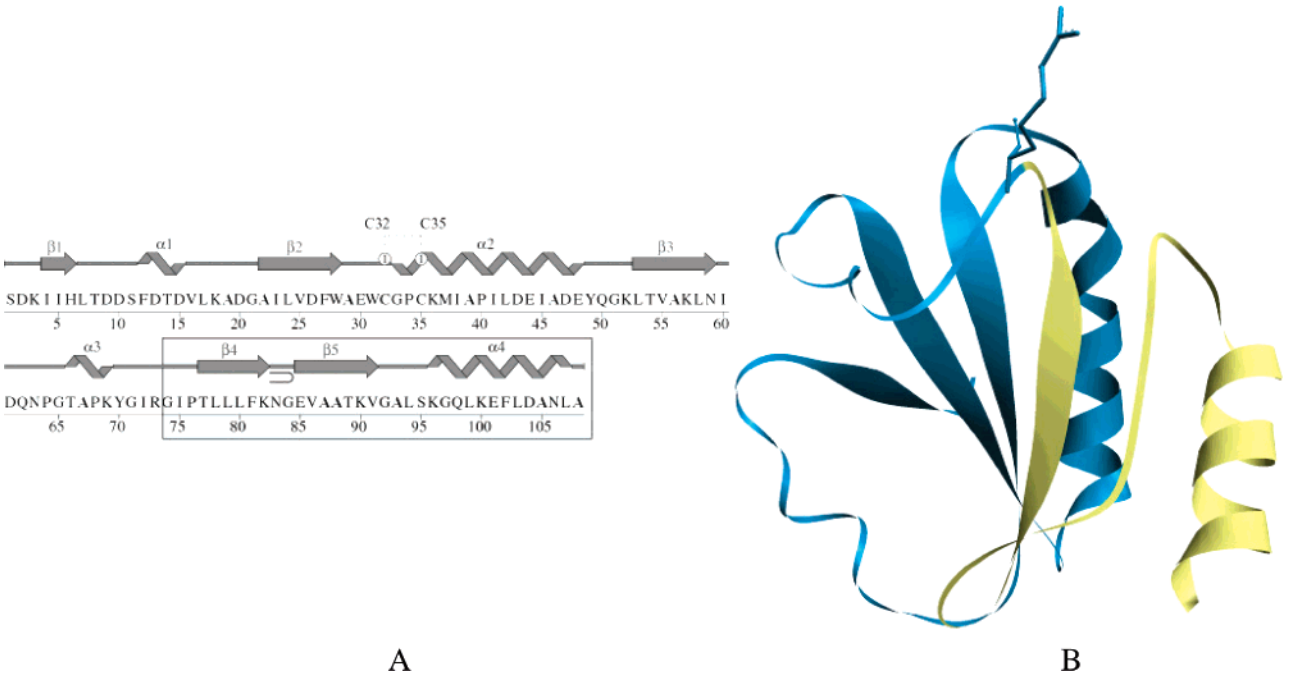

Figure 1. (A) Amino acid sequence and secondary structure of *E. coli* thioredoxin generated by PDBsum[35] using the PDB file 2trx.ent.[67] The C-terminal fragment encompassing amino acid residues 74−108 is highlighted in gray. (B) Tertiary structure. The N-terminal fragment (residues 1−73) is shown in blue; the C-terminus (residues 74−108) is in yellow. Arg-73 at the cleavage site is depicted using the stick representation. The interface consists of two β-strands, β2 and β4, formed by residues 23−29 and 76−82, belonging to the complementary N- and C-terminal fragments, respectively.

$= 14\,100\ M^{-1}\ cm^{-1}$.[41] In our hands, 40 mg of pure uniformly ^{13}C,^{15}N-enriched and 100 mg of pure uniformly ^{15}N-enriched thioredoxin per liter of minimal media were obtained. Thioredoxin was site-specifically cleaved at Arg-73 by Clostripain (Worthington Laboratories), according to the protocol established previously.[26] The fragments were isolated and purified by reverse-phase fast protein liquid chromatography. The concentration of the N and C fragments was determined using E280 $=$ $14\,100\ M^{-1}\ cm^{-1}$ and E215 $=$ $39\,700\ M^{-1}\ cm^{-1}$. The purity was evaluated by SDS−PAGE and electrospray mass spectroscopy. Solutions of the isolated fragments in 7 M Gu·HCl were mixed stoichiometrically at low concentration, dialyzed against potassium phosphate buffer, concentrated, and stored at −25 °C. Sample preparation for solid-state NMR measurements is described below.

Preparation of Microcrystalline and PEG-Precipitated Reassembled Thioredoxin. The crystallization conditions for thioredoxin were screened using a standard hanging drop procedure. The reservoir solutions contained 1 mL of 10 mM $NaCH_3COO$, 1 mM NaN_3, and the precipitant. A 0.5 μL aliquot of reassembled (1−73/74−108) thioredoxin solution was thoroughly mixed with 0.5 μL of the corresponding reservoir solution on a siliconized coverslip to form the hanging drop. The plates were left at 4 °C until crystals formed. In the initial screens, a number of conditions were varied, including the choice and the concentration of the precipitant, the concentration of reassembled thioredoxin, and the pH. In all the conditions, the overall buffer and salt concentration did not exceed 15 mM. Crystals were obtained at pH 3.5−3.8 from solutions containing final concentrations of 35 mg/mL of thioredoxin, 15−30% poly(ethylene glycol) 4000 (PEG-4000), 10 mM $NaCH_3COO$, and 1 mM NaN_3. These conditions were used to prepare solid-state NMR samples of mixed labeled thioredoxin complex by a controlled precipitation of the protein with PEG-4000. PEG-4000 was added to the solution containing 0.5 mL of 70 mg/mL of reassembled thioredoxin, 10 mM $NaCH_3COO$, and 1 mM NaN_3, pH 3.5, until no further protein precipitation was observed (ca. 2 h); the final concentration of PEG-4000 was estimated to be 30−35%. The PEG precipitate was centrifuged and transferred, after the supernatant was removed, into the 4 mm Bruker HRMAS rotor assembly. The samples were sealed using the upper spacer and the top spinner, according to the standard procedures.

Solid-State NMR Spectroscopy. All solid-state NMR spectra were acquired at 17.6 T on a wide-bore Bruker Avance spectrometer operating at Larmor frequencies of 750.22 MHz for ^1H, 188.64 MHz for ^{13}C, and 76.02 MHz for ^{15}N. A double-resonance 4 mm HX probe was employed for homonuclear experiments; the heteronuclear correlations were performed using a triple-resonance 4 mm narrow-bore HCN probe. Spectra were recorded using 8 mg of reassembled [U-^{15}N-(1−73)/U-^{13}C,^{15}N(74−108)]thioredoxin sample. The spinning speed was 7 and 10 kHz for hetero- and homonuclear experiments, respectively, and was controlled to within ±1 Hz by the Bruker MAS controller. The actual sample temperature was −15 °C, controlled to within ±0.5 °C throughout the experiments. ^{13}C chemical shifts were referenced with respect to adamantane used as an external standard, and assuming the chemical shift of its more deshielded carbon to be 38.6 ppm; 1.7 ppm was subsequently added to adjust the reference scale to 2,2-dimethylsilapentane-5-sulfonic acid DSS.[42] ^{15}N shifts were indirectly referenced with respect to solid NH_4Cl (39.2 ppm with respect to liquid ammonia at 25 °C).[43] This referencing procedure yields chemical shifts which are in close agreement (within the measurement error imposed by the ^{13}C and ^{15}N line widths) with the recently suggested protocol.[44] Table 1 summarizes the experimental solid-state chemical shifts in reassembled thioredoxin obtained by assigning the backbone and sidechain resonances as will be described in more detail below.

In Figure 2, the pulse sequences are shown for 2D experiments used for intraresidue and sequential assignments. Four experiments were conducted: NCO (sequential backbone correlations), NCA (intraresidue

(41) Holmgren, A.; Reichard, P. *Eur. J. Biochem.* **1967**, *2*, 187−196.

(42) Wishart, D. S.; Bigam, C. G.; Yao, J.; Abildgaard, F.; Dyson, H. J.; Oldfield, E.; Markley, J. L.; Sykes, B. D. *J. Biomol. NMR* **1995**, *6*, 135−140.

(43) McDermott, A.; Gu, Z. T. In *Encyclopedia of Nuclear Magnetic Resonance*; Grant, D. M., Harris, R. K., Eds.; Wiley: New York, 1996; Vol. 1, pp 1137−1147.

(44) Morcombe, C. R.; Zilm, K. W. *J. Magn. Reson.* **2003**, *162*, 479−486.

Table 1. ^{15}N and ^{13}C Solid-State Chemical Shifts of Reassembled *E. coli* Thioredoxin, Difference with Respect to Solution Shifts of Full-Length Thioredoxin, and Standard Deviations

	N Δ,soln SD	Cα Δ,soln SD	Cβ Δ,soln SD	Cγ Δ,soln SD	Cγ1 Δ,soln SD	Cγ1 & Cγ2 Δ,soln SD	Cγ2 Δ,soln SD	Cδ Δ,soln SD	Cδ1 Δ,soln SD	Cδ2 Δ,soln SD	Cε Δ,soln SD	Cε1 & Cε2 Δ,soln SD	Cξ Δ,soln SD	Co Δ,soln SD
G74	104.7	43.3												169.8
	−3.8	−1.2												n/a
	0.1	0.5												0.2
I75	114.4	59.4	38.6	n/a	24.5		20.4		16.6					173.7
	0.1	1.6	−2.8	n/a	−0.6		1.0		1.1					n/a
	0.1	0.5	0.6	n/a	0.5		0.1		0.1					0.3
P76	n/a	62.8	34.8	24.7				51.6						177.9
	n/a	−0.1	0.1	0.2				0.7						n/a
	n/a	0.1	0.2	0.2				0.3						0.3
T77	120.2	64.0	75.4				20.9							171.3
	1.5	−0.2	2.5				−1.1							n/a
	0.1	0.5	1.6				0.1							0.4
L78	126.1	52.6	44.1	26.9					n/a	n/a				175.3
	−1.1	−0.3	0.6	0.3					n/a	n/a				n/a
	0.1	0.2	0.2	0.1					n/a	n/a				0.1
L79	125.7	52.9	45.5	27.1					n/a	n/a				175.6
	3.5	−0.7	−0.62	−0.4					n/a	n/a				n/a
	0.1	0.1	0.2	0.1					n/a	n/a				0.1
L80	120	52.0	44.5	27.2					26.8	23.57				174.4
	−5.4	−1.3	−0.1	−0.3					0.1	−0.57				n/a
	0.1	2.4	0.3	0.1					0.1	0.3				0.1
F81	128.1	56.8	41.7	140.8					136.4	131.8		133.7	n/a	176.9
	−0.4	−0.6	−0.4	n/a					n/a	n/a		n/a	n/a	n/a
	0.1	0.4	0.3	0.1					0.1	0.1		n/a	n/a	0.6
K82	116.7	55.8	35.4	25.2				30.1			41.9			178.7
	−0.8	−0.4	−0.7	−0.2				0.2			n/a			n/a
	0.1	0.4	1.7	0.1				0.1			0.2			0.1
N83	123.9	54.3	37.5	n/a										176.1
	−0.4	−0.5	−0.1	n/a										n/a
	0.1	0.1	0.2	n/a										0.2
G84	104.2	46.3												173.9
	−0.3	0.01												n/a
	0.2	0.2												0.6
E85	117.3	55.3	35.5	38.1				183.9						177.2
	−1.5	1.3	2.3	2.5										n/a
	0.1	0.3	0.1	0.3				0.4						0.1
V86	124.3	64.4	29.1		−	22.2	−							175.5
	1.2	0.1	0.1		−	0.3	−							n/a
	0.4	0.1	0.1		−	0.2	−							0.1
A87	135.8	51.9	20.8											176.6
	2.5	−0.6	0.4											n/a
	0.3	0.3	0.4											0.1
A88	114.5	52.7	22.5											175.1
	−2.9	0.4	0.2											n/a
	0.2	0.1	0.3											0.4
T89	117.3	62.4	71.4				21.39							173.6
	1.9	0.74	−0.8				0.75							n/a
	0.1	0.2	0.3				0.2							0.2
K90	127.5	53.8	33.0	26.2				27.0			41.9			173.8
	1.2	−0.4	−1.9	2.2				−0.7			−0.2			n/a
	0.1	0.5	0.5	1.1				0.1			0.2			0.1
V91	132.1	62.2	33.5		−	21.2	−							177.3
	6.6	0.5	−0.5		−	0.5	−							n/a
	0.4	0.1	0.2		−	0.1	−							0.1
G92	114.9	44.0												171.7
	1.2	0.4												n/a
	0.3	0.1												0.1
A93	125.8	52.8	19.3											175.1
	7.3	0.1	0.3											n/a
	0.1	0.1	0.2											0.3
L94	124	54.7	43.6	26.7					26.9	24.0				n/a
	4.2	0.5	−0.6	0.1					0.7	−0.1				n/a
	0.2	0.2	0.7	0.2					0.1	0.1				n/a
S95	n/a	56.1	66.4											175.4
	n/a	−0.1	0.5											n/a
	n/a	0.1	0.1											0.1
K96	120.9	61.0	33.6	24.5				29.0			42.7			178.8
	−0.8	0.1	2.0	−1.2				−0.5			0.9			n/a
	0.2	0.1	1.1	2.04				0.8			0.2			0.1

Table 1 (Continued)

	N Δ,soln SD	Cα Δ,soln SD	Cβ Δ,soln SD	Cγ Δ,soln SD	Cγ1 Δ,soln SD	Cγ1 & Cγ2 Δ,soln SD	Cγ2 Δ,soln SD	Cδ Δ,soln SD	Cδ1 Δ,soln SD	Cδ2 Δ,soln SD	Cε Δ,soln SD	Cε1 & Cε2 Δ,soln SD	Cξ Δ,soln SD	Co Δ,soln SD
G97	106.1	46.9												177.0
	0.6	−0.2												n/a
	0.1	0.2												0.1
Q98	121.8	58.2	29.5	34.6				−						179.8
	1	−0.3	−0.1	−0.1				−						n/a
	0.2	0.2	0.1	0.1				−						0.2
L99	124	57.9	40.8	27.5					n/a	22.7				177.9
	0.5	−0.1	−0.2	0.0					n/a	−1.0				n/a
	0.1	0.1	0.1	0.1					n/a	0				0.7
K100	n/a	60.6	30.8	25.4				27.3			42.6			178.5
	n/a	−0.1	−0.4	−0.1				−1.5			1.3			n/a
	n/a	0.2	0.9	0.1				0.4			0.3			0.5
E101	118.7	59.9	28.5	34.5				181.8						179.5
	0.8	0.2	−1.3	−2.1				n/a						n/a
	0.2	0.2	0.4	0.5				0.9						0.6
F102	120.4	60.4	39.6	138.2					n/a	135.7		n/a	n/a	177.4
	0.1	0.4	0.5	n/a					n/a	n/a		n/a	n/a	n/a
	0.1	0.3	0.7	0.1					n/a	0.1		n/a	n/a	0.1
L103	n/a	57.8	40.9	26.2					23.2	27.5				177.4
	n/a	−0.1	−0.1	0.5					−2.7	4.8				n/a
	n/a	0.0	0.1	0.4					0.3	0.2				0.1
D104	120.8	57.5	39.2	180.7										179.2
	0.7	0.0	−0.2	n/a										n/a
	0.1	0.1	0.1	0.2										1.1
A105	n/a	53.8	19.2											179.1
	n/a	0.0	0.4											n/a
	n/a	0.3	0.1											0.1
N106	112.7	54.2	41.6	183.1										178.3
	−0.6	−0.3	−0.6	n/a										n/a
	0.3	0.2	0.1	0.1										0.5
L107	121	54.6	43.2	26.8					23.5	26.8				n/a
	−0.1	−0.3	0.7	2.3					0.1	0.2				n/a
	0.1	0.3	0.4	0.2					0.5	0.1				n/a
A108	n/a	54.4	19.3											178.0
	n/a	0.4	−0.4											n/a
	n/a	0.2	0.2											0.6

[a] Δ_{soln}, the difference between solid-state and solution chemical shifts for the reassembled thioredoxin and full-length intact thioredoxin. Solid-state chemical shifts were determined in this work; solution shifts were reported previously by Dyson and co-workers.[37] The ^{13}C solution shifts were re-referenced with respect to DSS. Solid-state ^{13}C chemical shifts were referenced using adamantane as external chemical shift reference (40.3 ppm for the deshielded carbon, which would correspond to 0 ppm with respect to DSS). Solid-state ^{15}N chemical shifts were referenced with respect to external NH_4Cl (39.2 ppm, corresponding to 0 ppm with respect to liquid NH_3 at 25 °C). The unassigned chemical shifts are denoted as n/a.

backbone correlations), CXCY (intraresidue and sequential side-chain correlations), and NN (sequential backbone correlations).

Two-dimensional $^{13}C-^{13}C$ chemical shift correlation spectra were acquired using a rotary-assisted proton spin diffusion experiment (RAD, DARR) at $N = 1$.[45] During the initial 1.5 ms cross polarization (CP) period, the ^1H amplitude was linearly ramped from 80 to 100%, with the center of the ramp corresponding to the first Hartmann−Hahn spinning sideband. During the ^{13}C chemical shift evolution period in the indirect dimension, continuous wave (CW) heteronuclear $^1H-^{13}C$ decoupling was used. Throughout RAD mixing and detection, the 70 kHz XiX decoupling scheme was applied.[46] Three sets of spectra were acquired using RAD mixing times of 10, 100, and 300 ms. The spectral widths for these three data sets were 50 kHz in both dimensions, and the ^{13}C frequency was centered at 95 ppm. A total of 32 scans were added for the final free induction decays (FIDs) in each t_1 transient; a total of 400 t_1 transients were acquired, and recycle delays of 2 s were employed. The total acquisition time was 7.1 h. The following field strengths were used on the ^1H channel: 125 kHz for hard pulses, 58 kHz (corresponding to the center of the ramp) for cross polarization, and 70 kHz during decoupling. Irradiation of 10 kHz was applied to ^1H during the mixing period. On the ^{13}C channel, field strength of 48 kHz was applied during cross polarization, and 53.5 kHz for $\pi/2$ and π pulses during the mixing period.

Two-dimensional $^{15}N-^{13}C$ correlation spectra were acquired using a SPECIFIC CP mixing sequence[47] to establish either the N−Cα or N−Co correlations. During the initial 3 ms cross polarization period, the ^1H amplitude was linearly ramped from 80 to 100%, with the center of the ramp corresponding to $\omega_1(^1H) = 35$ kHz, $\omega_1(^{15}N) = 28$ kHz. During the ^{15}N chemical shift evolution period in f_1, CW heteronuclear $^1H-^{13}N$ decoupling was used. The polarization was subsequently transferred from the amide nitrogen atoms to either Cα (NCA experiment) or Co (NCO experiment) carbons via selective double cross polarization, with the ^{15}N amplitude being linearly ramped from 80% to 100%. The cross polarization mixing time was 6 ms; $\omega_1(^{15}N) = 25$ kHz (center of the ramp), $\omega_1(^{13}C) = 18$ kHz. Due to the probe restrictions on the maximum amount of power, the heteronuclear decoupling field strength was kept at 50 kHz throughout the experiment, and the XiX decoupling scheme was employed. The ^{15}N frequency was centered at 119 ppm; the ^{13}C frequency was centered at 51 ppm (NCA experiment) or at 180 ppm (NCO experiment). A total of 208 (NCO) or 224 (NCA) scans were added for the final FIDs in each t_1

(45) Takegoshi, K.; Nakamura, S.; Terao, T. *Chem. Phys. Lett.* **2001**, *344*, 631−637.
(46) Detken, A.; Hardy, E. H.; Ernst, M.; Meier, B. H. *Chem. Phys. Lett.* **2002**, *356*, 298−304.
(47) Baldus, M.; Petkova, A. T.; Herzfeld, J.; Griffin, R. G. *Mol. Phys.* **1998**, *95*, 1197−1207.

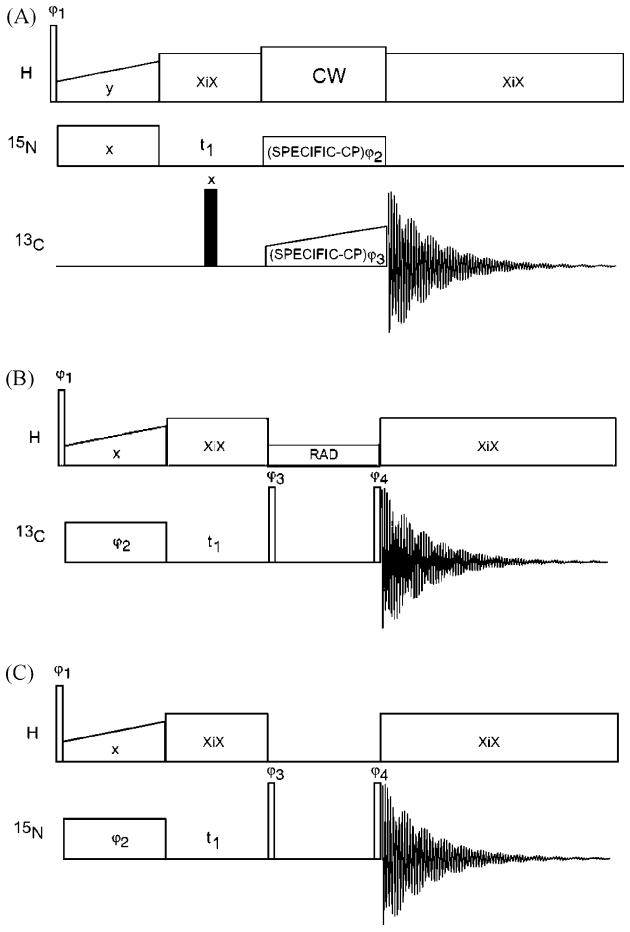

Figure 2. Pulse sequences for 2D NCA/NCO (A), CXCY (B), and NN (C) experiments used in the backbone and side-chain assignments. Open rectangles represent $\pi/2$ pulses; filled rectangles represent π pulses. SPECIFIC CP with 8 ms mixing time was utilized for the N−Cα(Co) transfers. The phase cycle was $\varphi_1 = xx\bar{x}\bar{x}$; $\varphi_2 = x\bar{x}x\bar{x}$; $\varphi_3 = (x)_4(y)_4$; $\varphi_{rec} = x\bar{x}\bar{x}x\bar{y}\bar{y}yy$. $^{13}C-^{13}C$ correlations were established via the RAD mixing scheme at $\omega_{rf} = \omega_r$. The phase cycle was $\varphi_1 = y\bar{y}$; $\varphi_2 = xx\bar{x}\bar{x}$; $\varphi_3 = (y)_4(\bar{y})_4$; $\varphi_4 = (x)_8(\bar{x})_8$; $\varphi_{rec} = x\bar{x}\bar{x}x(\bar{x}xx\bar{x})_2x\bar{x}\bar{x}x$. $^{15}N-^{15}N$ sequential correlations were achieved using the PDSD mixing sequence, $t_{mix} = 4$ s. The phase cycle was the same as in (B).

transient; a total of 185 t_1 transients were acquired, and recycle delays of 3 s were employed. The total experiment times were 32 and 34 h for the NCO and NCA experiments, respectively.

Two-dimensional $^{15}N-^{15}N$ correlation spectra were acquired using a proton-driven spin diffusion (PDSD) mixing sequence.[48] During the initial 1 ms cross polarization period, the ^1H amplitude was linearly ramped from 80 to 100%, with the center of the ramp corresponding to $\omega_1(^1H) = 59$ kHz, $\omega_1(^{15}N) = 49$ kHz. The PDSD mixing time was 4 s. XiX decoupling was applied during f_2. A total of 32 scans were added to form the FID in each t_1 increment; a total of 200 t_1 increments were acquired, and the recycle delay was 2 s. The total experiment time was 3.6 h.

Additionally, one-dimensional ^{13}C and ^{15}N CPMAS spectra were recorded at the spinning speed of 10 kHz. The ^1H amplitude was linearly ramped from 80 to 100%, with the center of the ramp corresponding to $\omega_1(^1H) = 59$ kHz, $\omega_1(^{15}N)$ (or $\omega_1(^{13}C)$) = 49 kHz. XiX decoupling at a field strength of 70 kHz was applied during the acquisition. A total of 128 or 512 scans were added to form the final FIDs in the ^{15}N or ^{13}C spectra, respectively; the acquisition time was 50 ms.

The data were processed in NMRPipe[49] under a Linux environment on a Dell Precision workstation. The NCA, NCO, and CC spectra were processed with cosine apodization in both dimensions, zero-filling up to 4096 points in f_2 and 512 points in f_1. Additionally, the CC spectrum acquired with 10 ms mixing time was processed with resolution enhancement, by applying a 60°-shifted sine apodization in f_2 and a 30°-shifted sine apodization in f_1. The NN spectra were processed with cosine apodization in both dimensions; the data were zero-filled up to 2048 points in f_2. Linear prediction up to 384 points was applied in f_1; zero-filling up to 1024 points was used. The spectral assignments were performed in Sparky.[50] The secondary shift statistical analysis and backbone dihedral angle prediction were performed in TALOS.[51]

(48) Bloembergen, N. *Physica* **1949**, *15*, 386-426.

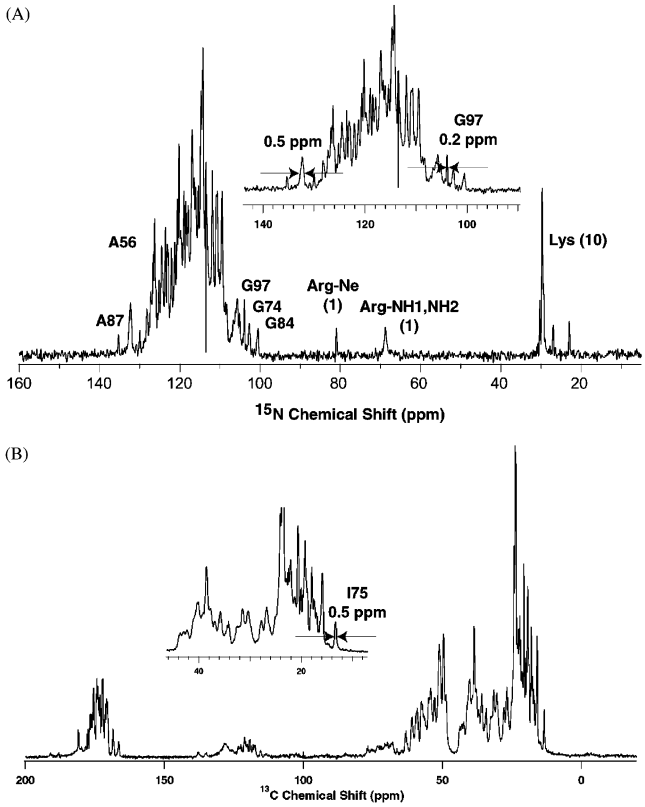

Figure 3. ^{15}N (A) and ^{13}C (B) CPMAS spectra of PEG-precipitated *E. coli* thioredoxin reassembled from complementary fragments U-^{15}N(1−73)/U-^{13}C,^{15}N(74−108). The line widths are 0.2−0.5 ppm (^{15}N) and 0.5−0.8 ppm (^{13}C). For the ^{15}N spectrum, 128 transients were added, and the recycle delay was 2 s. The ^{13}C spectrum was acquired with 512 transients. The spinning frequency was 10 kHz.

Results and Discussion

Sample Conditions for Solid-State NMR Experiments. In Figure 3, the ^{15}N and ^{13}C CPMAS spectra for the PEG-precipitated thioredoxin are shown. The line widths for the outliers are 15−38 Hz (0.2−0.5 ppm) and 35−94 Hz (0.5−0.8 ppm) for the ^{15}N and ^{13}C data sets, respectively, indicating a high degree of conformational homogeneity in the protein resulting from the above protocol.

As will be demonstrated below in the 2D data sets, the PEG-precipitated reassembled thioredoxin results in high-quality spectra suitable for resonance assignments. This finding, in conjunction with previous reports on solid-state NMR of the PEG-precipitated α-spectrin SH_3 domain[19,52] and the catabolite

(49) Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A. *J. Biomol. NMR* **1995**, *6*, 277−293.
(50) Goddard, T. D.; Kneller, D. G. *SPARKY 3*, version 3.110; University of California, San Francisco: San Francisco, CA, 2004.
(51) Cornilescu, G.; Delaglio, F.; Bax, A. *J. Biomol. NMR* **1999**, *13*, 289−302.
(52) Pauli, J.; van Rossum, B.; Forster, H.; de Groot, H. J. M.; Oschkinat, H. *J. Magn. Reson.* **2000**, *143*, 411−416.

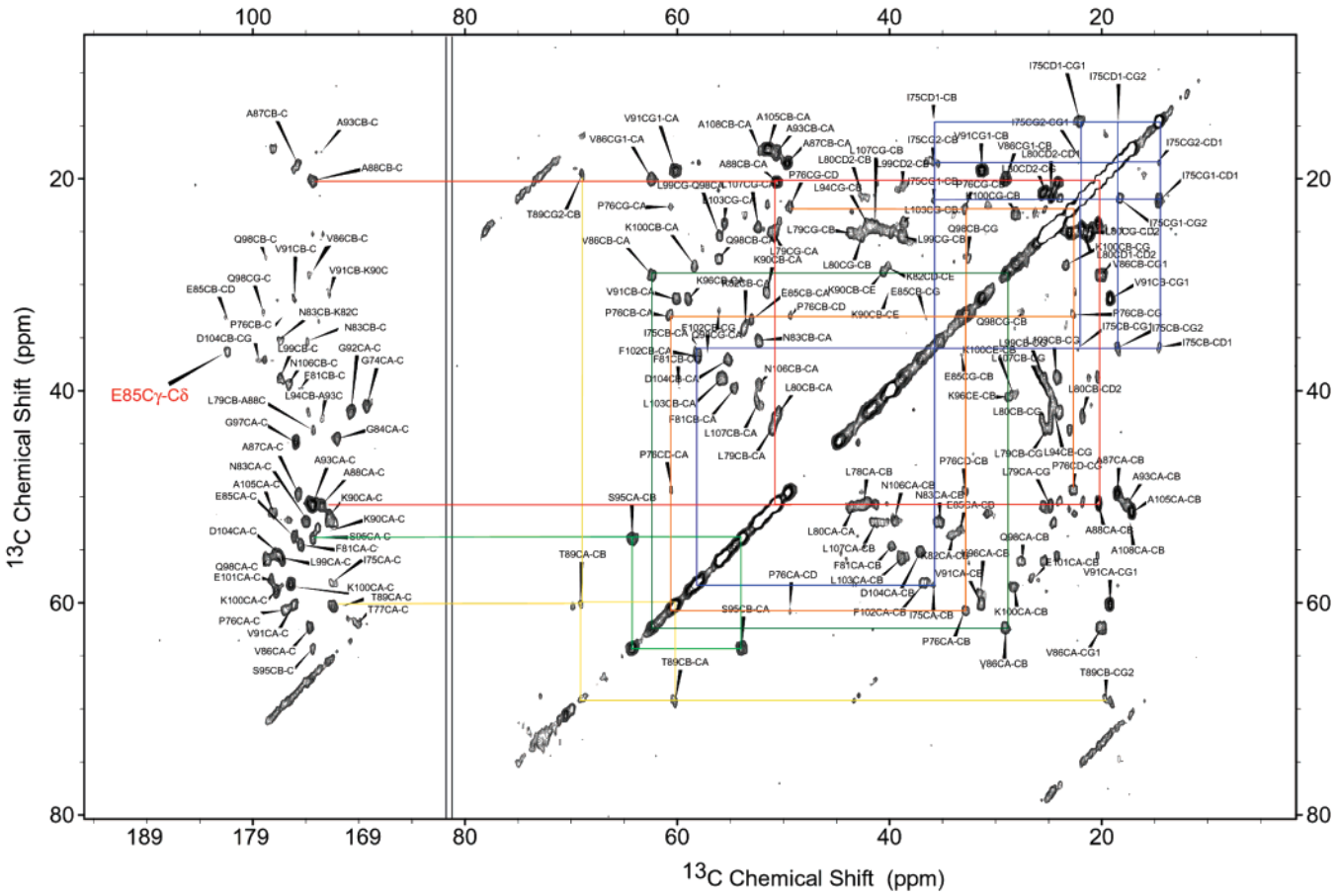

Figure 4. CXCY spectra of reassembled *E. coli* thioredoxin demonstrating intraresidue side-chain walks for identification of amino acid residue types. The CXCY spectrum was recorded with the RAD mixing element and a mixing time of 10 ms, resulting in predominantly one-bond correlations and a limited number of two-bond intraresidue correlations. The spectra were processed with cosine apodization in both dimensions, zero-filling up to 4096 points in the direct dimension and 512 points in the indirect dimension. Examples of amino acid chemical shift topologies are shown: Ala, red; Ile, blue; Pro, orange; Ser, light green; Thr, yellow; and Val, dark green. For Ala, strong $C_\beta-C_o$ cross-peaks were observed in addition to one-bond correlations.

repression histidine-containing phosphocarrier protein Crh,[21] as well as our studies of full-length thioredoxin, indicates that preparation of high-quality protein samples for solid-state NMR could be generally accomplished via a relatively facile generation of PEG precipitates. Our ongoing efforts on a number of other systems, including azurin, vanadium chloroperoxidase, and Max, suggest that this approach may work for a large number of globular proteins. Moreover, a number of previous studies demonstrate that PEG-precipitated proteins are functionally competent.[11, 53]

The observed ^{13}C and ^{15}N line widths were somewhat broader than the ^{13}C−^{15}N and ^{13}C−^{13}C scalar couplings (7−40 Hz),[54] indicating the contribution from the residual ^{1}H−^{13}C dipolar couplings, which is in agreement with the weak decoupling field strengths (50 kHz) applied throughout the measurements and imposed by the probe configuration. To further narrow the lines, it would be necessary to employ either higher decoupling field strengths or ultrafast magic angle spinning (spinning frequencies > 50 kHz).[55] Either of these two approaches may be crucial for larger proteins or intrinsically crowded spectra. However, for reassembled thioredoxin, moderate spinning frequencies and decoupling radio frequency fields resulted in adequate resolution in multidimensional spectra, to yield the majority of resonance assignments.

Experimental Protocols and Considerations for Resonance Assignments of Reassembled Thioredoxin at 17.6 T. For intraresidue and sequential backbone assignments, NCA, NCO, and NN experiments were conducted. In the NCA and NCO experiments, polarization is transferred from the backbone amide nitrogen atom to the α-carbon or the carbonyl carbon, thus establishing intraresidue or sequential backbone correlations. SPECIFIC CP mixing sequence[47] was employed with weak radio frequency fields centered at ^{15}N and ^{13}C (Cα or Co) frequencies and $\omega_1(^{15}N) = \omega_1(^{13}C) + \omega_r$, to accomplish selective transfers. The transfer efficiencies were 20−25%, which is probably the result of weak decoupling field strengths (vide supra). With these transfer efficiencies, high-quality spectra were acquired within 32−34 h. Transfer efficiencies were found to be hardware dependent, and with a different probe configuration those reached 45−50.

The homonuclear NN experiment employs a PDSD mixing period to transfer polarization between backbone amide nitrogens. Due to the small magnitude of the homonuclear ^{15}N dipole−dipole interaction (25 or 60 kHz, corresponding to the average distances of 2.7 or 3.5 Å in α-helices of β-sheets, respectively), only sequential N(i)−N(i+1) correlations are present in the spectra. At a mixing time of 4 s, cross-peak

(53) Williams, J. C.; McDermott, A. E. *Biochemistry* **1995**, *43*, 8309−8319.
(54) Clore, G. M.; Gronenborn, A. M. *Science* **1991**, *252*, 1390−1399.
(55) Ernst, M.; Meier, M. A.; Tuherm, T.; Samoson, A.; Meier, B. H. *J. Am. Chem. Soc.* **2004**, *126*, 4764−4765.

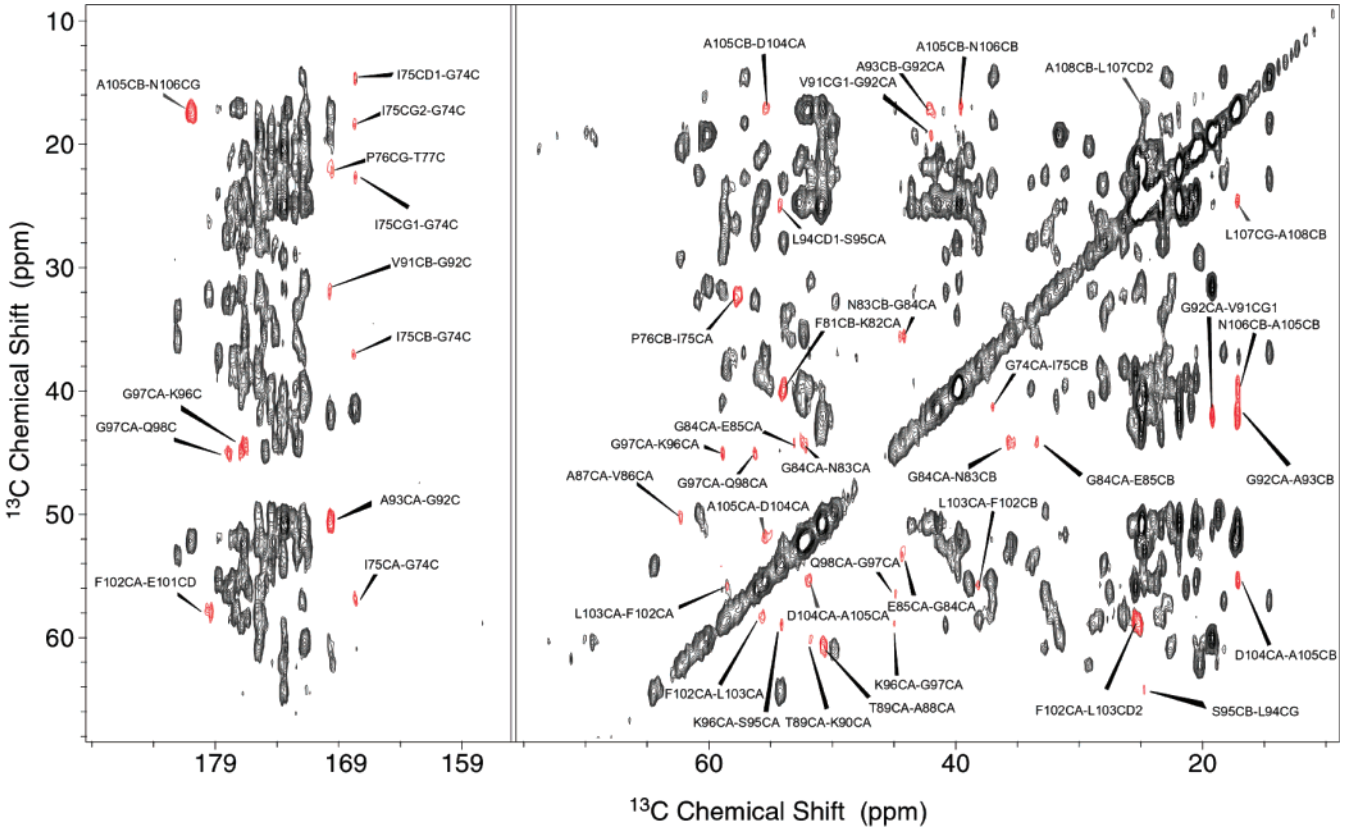

Figure 5. CXCY spectrum of reassembled *E. coli* thioredoxin demonstrating examples of sequential $^{13}C-^{13}C$ correlations. The experiment was recorded with RAD sequence and mixing time of 100 ms. The spectrum was processed using a cosine filter in both dimensions, zero-filling to 4096 points in the f_2 dimension and 1024 points in the f_1 dimension, and automatic baseline correction in the f_2 dimension.

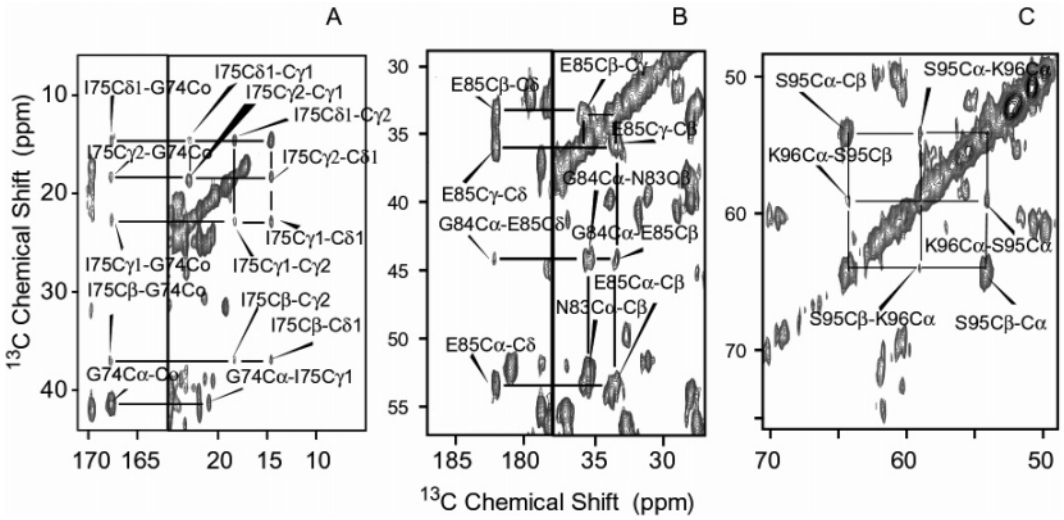

Figure 6. Expansions of the RAD spectrum acquired with a mixing time of 100 ms, demonstrating the sequential cross-peaks between (A) G74 and I75, (B) G84−E85, and (C) K96 and S95.

intensities were 10−30% of the diagonal, and were found to be independent of the secondary structure. Sequential nitrogen assignments based on these spectra alone are not feasible due to the large degree of spectral overlap. However, as described below, the NN experiment was useful in corroborating the nitrogen backbone assignments based on the NCO and NCA data.

For homonuclear CC recoupling, rotary-assisted spin diffusion sequence RAD[45] was utilized. This experiment is readily tuned for establishing dipolar interactions of the desired strength by the appropriate choice of mixing period. At short mixing times of 1−2 ms, predominantly one-bond correlations are established, which are utilized for intraresidue side-chain assignments. Upon increasing the mixing time to 10 ms, two-bond correlations emerge, and at mixing times of 100 ms or longer, a number of multibond correlations are present. These correspond to medium- and long-range $^{13}C-^{13}C$ distances and were useful for confirming sequential assignments as well as obtaining tertiary structure constraints. The recoupling efficiencies in RAD were found to be largely insensitive to the chemical shift offset, and uniform

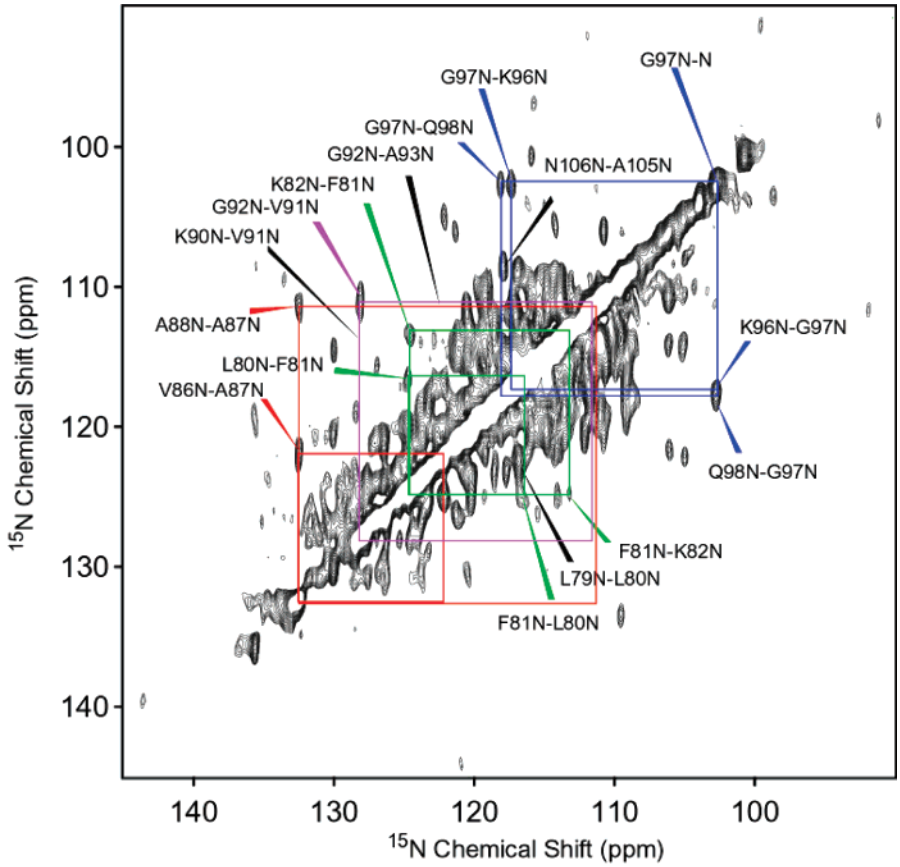

Figure 7. N−N correlation spectrum of reassembled *E. coli* thioredoxin demonstrating examples of sequential backbone walks. The $^{15}N-^{15}N$ correlations were established via a PDSD mixing sequence with 12 ms mixing time. The spectrum was processed with linear prediction up to 384 points in f_1, cosine apodization in both dimensions, zero-filling up to 2048 points in f_2 and 1024 points in f_1, and automatic baseline correction in f_2.

polarization transfer was achieved for the aliphatic side-chain carbons. Interestingly, the intensities of the cross-peaks corresponding to the two-bond correlations were found to be sensitive to the dihedral angles: the residues belonging to β-sheets exhibited stronger $C_\beta-C_o$ cross-peaks than those comprising turns and α-helices. These intensity differences might potentially serve as an independent qualitative indicator of the secondary structure context of a particular residue; however, more experimental evidence is needed.

Due to the small size of the U-^{13}C,^{15}N-enriched portion of reassembled thioredoxin and the high degree of chemical shift dispersion at 17.6 T, the above combination of 2D homo- and heteronuclear correlation experiments yielded nearly complete assignments of the backbone and side-chain ^{15}N and ^{13}C resonances, as detailed below for several residue types.

The approach presented above takes advantage of high magnetic fields and small size of reassembled thioredoxin. On the basis of our experience with thioredoxin and BPTI, and the work from other laboratories discussing alternative strategies,[13−21,52,56,57] several experimental considerations can be outlined for high-resolution MAS studies of uniformly enriched polypeptides. For moderate size proteins, backbone resonance assignments can be performed at lower field strengths (7.0 and 9.4 T);[13,14] however, three-dimensional experiments and a field strength of at least 9.4 T are required for nearly complete assignments.[15] Higher magnetic fields appear to be necessary for adequate resolution in homonuclear $^{13}C-^{13}C$ experiments.[18] Sensitivity is an additional important factor. At 750 MHz, 0.1−0.5 μmol of protein is sufficient for the measurements. Larger amounts of sample and longer experiment times will have to be utilized at lower fields.

Resonance Assignments. The majority of the amino acid types in the C-terminal fragment of reassembled thioredoxin were identified from intraresidue one- and two-bond CC correlations observed in the RAD experiment with a mixing time of 10 ms (Figure 4). Identification of the five alanines, two valines, two threonines, three out of four glycines, two out of four lysines, two phenylalanines, one serine, and one isoleucine was straightforward, based on the characteristic fingerprint region and their expected intraresidue cross-peaks, especially in the aliphatic region. In the carbonyl region, a number of $C_\beta-C_o$ and $C_\gamma-C_o$ intraresidue connectivities were observed as well. Several examples of site-specific assignment for individual residues are presented below.

Site-specific assignments of I75 and S95 were accomplished on the basis of the above experiment. The unique isoleucine in the sequence was easily identified by the multiple correlations among its aliphatic carbons, especially the methyl groups, which were the most shielded signals in the aliphatic region of the spectrum. These C−C connectivities are depicted in Figure 4. S95 was identified by its characteristically deshielded C_β resonance (64 ppm). The unequivocal assignment of I75 in

(56) Castellani, F.; van Rossum, B.-J.; Diehl, A.; Rehbein, K.; Oschkinat, H. *Biochemistry* **2003**, *42*, 11476−11483.
(57) Egorova-Zachernyuk, T. A.; Hollander, J.; Fraser, N.; Gast, P.; Hoff, A. J.; Cogdell, R.; de Groot, H. J. M.; Baldus, M. *J. Biomol. NMR* **2001**, *19*, 243−253.

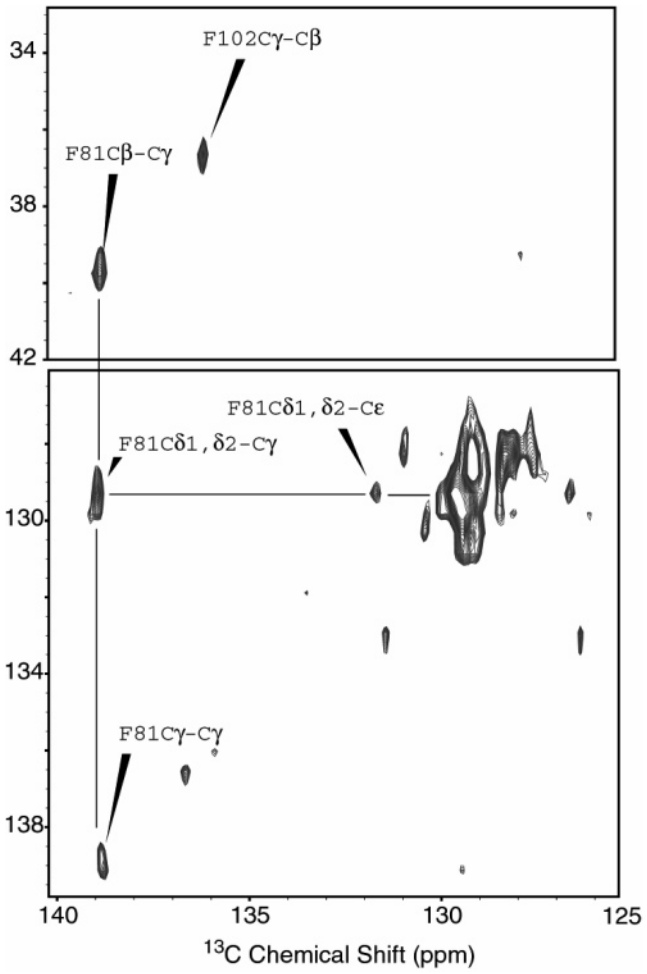

Figure 8. Expansion of the RAD spectrum acquired with a mixing time of 10 ms, demonstrating the aromatic side-chain assignments in F81 and F102.

addition revealed the identity of the unique proline, P76, via the multiple cross-peaks between its aliphatic carbons and the carbonyl carbon of I75 found in the carbonyl regions of the RAD experiment acquired with long mixing times of 100 and 300 ms. These three residues, I75, P76, and S95, served as the starting point in the assignments.

To establish sequential connectivities, a combination of four experiments was used: RAD with long mixing times (100 and 300 ms), NCO, NCA, and NN experiments (Figures 5−7, 9). The 100 and 300 ms RAD experiments were found particularly useful in corroborating the sequential heteronuclear assignments (Figures 5 and 6), as well as in resolving the ambiguities for the poorly dispersed leucine and lysine residues. In the following, the site-specific assignments for individual amino acid types are discussed.

Identification of glycines was relatively straightforward, due to their low abundance in the sequence (four) and their characteristic upfield C_α and C_o carbon chemical shifts. Three of the four glycines were found on the basis of their cross-peaks in the carbonyl regions of the RAD experiments at all mixing times, while the typical N−Cα correlations corresponding to all four glycines were present in the NCA experiment.

Sequential correlations between each of the four glycines and their neighboring residues were used to establish the site-specific assignments. For example, G84 was identified on the basis of its multiple CC sequential correlations with E85 and V86 (Figure 6). In the glutamate residues, the C_δ carbon is deshielded (ca. 180−186 ppm). Therefore, the characteristic $C_\gamma-C_\delta$ correlations of the two glutamate residues, E85 and E101, appeared in the isolated region of the spectrum and were used to report on these two residues. In Figure 4, the $C_\gamma-C_\delta$ correlation for E85 is highlighted in the carbonyl region of the spectrum. E85 was distinguished on the basis of the sequential correlations of its C_δ carbon to C_α of G84 and $C_{\gamma1}$,$C_{\gamma2}$ of V86 residues in the RAD spectra at long mixing times. Sequential cross-peaks between the aliphatic and the carbonyl carbons of G84, E85, and N83 were also observed under these conditions: $G84C_\alpha-E85C_\alpha$, $E85C_\beta-G84C_\alpha$, $G84C_\alpha-E85C_\delta$, $G84C_\alpha-N83C_\alpha$, $G84C_\alpha-N83C_\alpha$, and $G84C_\alpha-N83C_\beta$.

G74 was distinguished from the other glycines, G92 and G97, by several sequential cross-peaks with I75 in the aliphatic and carbonyl regions of the 100 and 300 ms RAD spectra (Figure 6): $G74C_o-I75C_\alpha$, $G74C_o-I75C_\beta$, $G74C_o-I75C_{\gamma1}$, $G74C_o-I75C_{\gamma2}$, $G74C_o-I75C_{\delta1}$, $G74C_\alpha-I75C_\alpha$, $G74C_\alpha-I75C_\beta$, and $G74C_\alpha-I75C_{\delta1}$. The absence of a cross-peak to the preceding residue in the NCO experiment confirmed this assignment (R73 belongs to the complementary thioredoxin fragment containing no ^{13}C labels).

Site-specific identification of the two remaining glycines, G92 and G97, was straightforward on the basis of the various sequential connectivities found in the NCO, NN, and 100 and 300 ms CC RAD experiments. The observed $G92C_\alpha-A93C_o$, $G92C_\alpha-A93C_\beta$, $G92C_\alpha-V91C_o$, $G92C_\alpha-V91C_\alpha$, $G92C_\alpha-V91C_{\gamma1}$, $G92C_\alpha-V91C_{\gamma2}$, G92N−C91C$_o$, and G92N−V91N correlations in conjunction with the intraresidue V91N−C_α cross-peaks yielded the ^{13}C and ^{15}N backbone and side-chain chemical shifts for G92. For G97, $G97C_\alpha-K96C_\alpha$, $G97C_\alpha-Q98C_o$, $G97C_\alpha-Q98C_\alpha$, $G97C_\alpha-K96C_\alpha$, G97N−K96N, and G97N−K96N correlations were employed to derive the ^{13}C and ^{15}N chemical shifts. Assignment procedures for the remaining residue types, including aromatic sidechains of the Phe-81 and Phe-82 (Figure 8), are contained in the Supporting Information.

NCO (A), NCA (B), and CXCY (C) spectra of reassembled *E. coli* thioredoxin, demonstrating examples of intraresidue and sequential backbone and side-chain assignment walks, are presented in Figure 9. In summary, we were able to unambiguously assign 93% of the ^{13}C and ^{15}N chemical shifts of the C-terminal of reassembled thioredoxin by using a combination of 2D homo- and heteronuclear correlation experiments. These chemical shifts are summarized in Table 1.

Secondary Chemical Shifts, Dihedral ϕ and ψ Angles, and Secondary Structure Predictions: Analysis and Comparison with Solution NMR. Secondary carbon and nitrogen chemical shifts are sensitive to polypeptide backbone conformation.[58] Dihedral ψ angle prediction based on the statistical analysis of secondary chemical shifts as implemented in TALOS[51] has been widely applied in solution NMR for deriving the secondary structure, and more recently, this approach has been demonstrated to yield accurate results for solid-state protein chemical shifts.[21,57] As solution chemical shifts for $^{13}C_\alpha$, $^{13}C_\beta$, and backbone amide ^{15}N atoms in reassembled 1−73/74−108 thioredoxin have not been reported, direct comparison between secondary shifts in the solid state and in solution is not possible. Therefore, we carried out a comparison of the dihedral ψ angles

(58) Wishart, D. S.; Sykes, B. D. *J. Biomol. NMR* **1994**, *4*, 171−180.

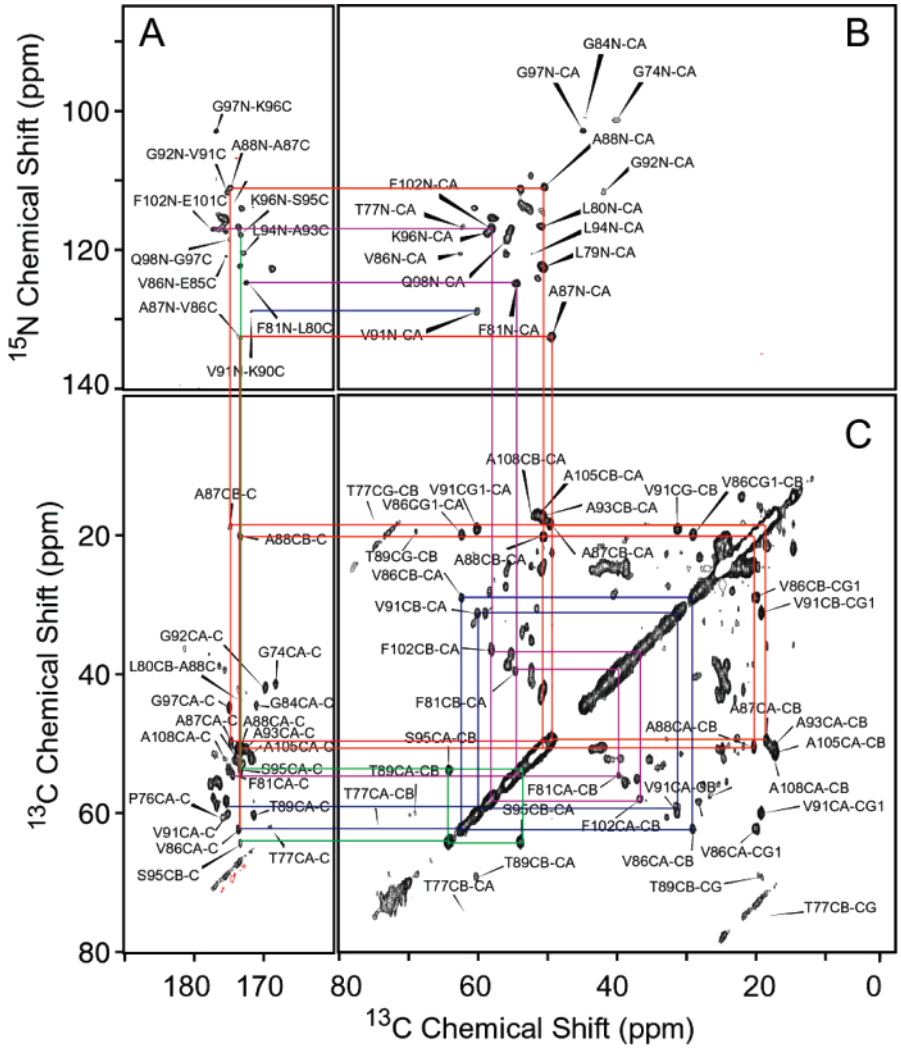

Figure 9. NCO (A), NCA (B), and CXCY (C) spectra of reassembled *E. coli* thioredoxin demonstrating examples of intraresidue and sequential backbone and side-chain assignment walks. The CXCY spectrum was recorded with the RAD mixing element and a mixing time of 10 ms, resulting in predominantly one-bond correlations and a limited number of two-bond intraresidue correlations. The spectra were processed with cosine apodization in both dimensions and zero-filling up to 4096 points in the direct dimension and 512 points in the indirect dimension.

predicted using the secondary shifts of the reassembled thioredoxin in the solid-state, and those of the full-length intact thioredoxin in solution. Such comparison is warranted since it was demonstrated previously that the overall tertiary structure in the reassembled thioredoxin in solution is preserved with respect to the intact protein.[26]

In Figure 10, the dihedral ϕ and ψ angles predicted from solid-state state shifts of reassembled thioredoxin and solution shifts for the full-length protein are plotted. The overwhelming majority (29 of 33) of ψ angles and 32 of 33 ϕ angles deviate by less than 20°, with the mean of 3.7 and 0.15° for ψ and ϕ, respectively, and correctly reflect the secondary structure of these 31 residues. The largest discrepancies are displayed by Asn-83, Ala-87, Lys-96, and Gln-98. Since Asn-83 and Lys-96 are located at the termini of a β-sheet and α-helix, respectively, the disagreement is not surprising. The ψ angle of −5° inferred from the solid-state shifts for Gln-98 accurately predicts that this residue is in an α-helical environment. The corresponding dihedral angle derived from the solution shifts is 137° and would indicate a β-sheet environment, in disagreement with the experimental X-ray and solution NMR structure. For Ala-87, solid-state shifts predict a ψ angle of −12°, indicative of an α-helical environment; on the contrary, the solution-predicted angle is 125°, in agreement with the expected location of this residue in a β-sheet. For Ala-87, the reason for the discrepancy between predictions based on solid-state NMR shifts and the expected angle is not clear at this point. The assignments for the ^{13}C and ^{15}N chemical shifts were based on redundant sets of experiments; therefore, it is unlikely that these were determined incorrectly and could be the source of error. An additional body of evidence for full-length thioredoxin and other proteins will be necessary to evaluate whether the observed discrepancy for Ala-87 is statistically meaningful. With the exception of Ala-87, the secondary structure motifs are correctly inferred from solid-state NMR data, providing additional evidence for the general applicability of the TALOS prediction algorithm to solid-state NMR chemical shifts.

Use of Differential Isotopic Labeling Strategies for High-Resolution Solid-State NMR Studies of Protein Interfaces and Protein Assemblies. Sparse and segmental isotopic labeling protocols have been widely employed for protein studies by both solution and solid-state NMR spectroscopy.[59−66] The current work demonstrates that introducing differential isotopic labels could be potentially beneficial for high-resolution struc-

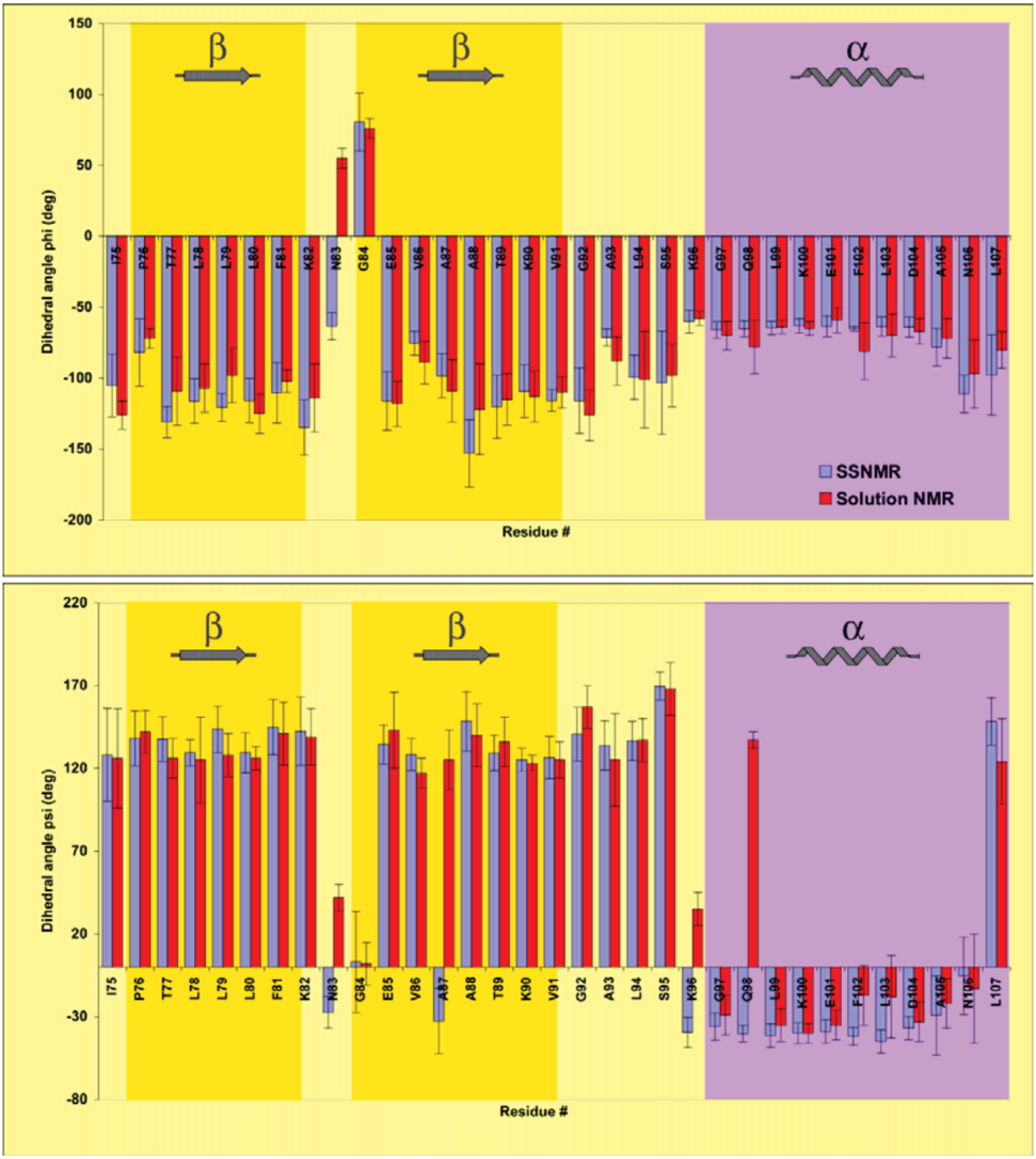

Figure 10. Comparison of the dihedral ψ angles predicted from the observed solid-state and solution chemical shifts of reassembled and intact thioredoxin, respectively, using TALOS.[51] The filled rectangles represent solid-state values, and the empty rectangles represent solution values. The α-helical and β-sheet regions in the secondary structure are highlighted in lighter and darker gray, respectively.

tural studies of protein interfaces and protein assemblies by solid-state NMR, especially in larger systems. One advantage of differential labeling is significantly simplified spectral assignments. For example, assignments of the C-terminal portion of the reassembled, differentially labeled thioredoxin complex reported in this work were straightforward due to the smaller number of carbon resonances compared with the full-length uniformly enriched protein. Another advantage is isotope editing of the polarization transfer, permitting for directing the transfer pathway. Specifically, the labeling scheme discussed in this work was developed for analysis of hydrogen-bonding patterns

(59) LeMaster, D. M.; Kushlan, D. M. *J. Am. Chem. Soc.* **1996**, *118*, 9255−9264.
(60) Hong, M. *J. Magn. Reson.* **1999**, *139*, 389−401.
(61) Gardner, K. H.; Kay, L. E. *Annu. Rev. Biophys. Biomol. Struct.* **1998**, *27*, 357−406.
(62) Muir, T. W.; Sondhi, D.; Cole, P. A. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 6705−6710.
(63) Waugh, D. S. *J. Biomol. NMR* **1996**, *8*, 184−192.
(64) Yabuki, T.; Kigawa, T.; Dohmae, N.; Takio, K.; Terada, T.; Ito, Y.; Laue, E. D.; Cooper, J. A.; Kainosho, M.; Yokoyama, S. *J. Biomol. NMR* **1998**, *11*, 295−306.
(65) Yamazaki, T.; Otomo, T.; Oda, N.; Kyogoku, Y.; Uegaki, K.; Ito, N.; Ishino, Y.; Nakamura, H. *J. Am. Chem. Soc.* **1998**, *120*, 5591−5592.
(66) Katti, S. K.; LeMaster, D. M.; Eklund, H. *J. Mol. Biol.* **1990**, *212*, 167−84.

at the reassembled thioredoxin interface via the $^{13}C-(^{1}H)-^{15}N$ experiments, and is applicable for assessing intermolecular interactions in protein interfaces formed by any arbitrary pair of proteins. Introduction of $^{13}C,^{15}N/^{15}N$ labels according to this protocol enables both the resonance assignments of one binding partner and the evaluation of interfacial contacts between the individual molecules. This approach is expected to be particularly beneficial for protein assemblies where only the interface structure is of interest and the complete assignment of the complementary partner may not be required or feasible. In this case, preparation of a $^{13}C/^{15}N$ mixed labeled complex would allow for directing the $^{15}N-^{13}C$ magnetization transfers via the interface and highlighting only the corresponding interface residues. These residues can be readily assigned via a combination of $^{15}N-^{13}C-^{13}C$ experiments establishing intraresidue and sequential ^{13}C correlations. A complementary labeling scheme would yield ^{13}C assignments for the interface residues of the binding partner. The assignment of ^{15}N resonances could be accomplished via a combination of $^{13}C-^{15}N$ and $^{15}N-^{15}N$ correlation experiments. Additional protocols employing a combination of uniform $^{13}C,^{15}N$ and sparse ^{13}C labels have been devised for analysis of long-range tertiary contacts at the reassembled thioredoxin interface. These studies will be reported in the forthcoming manuscripts.

Conclusions

Resonance assignments for the uniformly $^{13}C,^{15}N$-enriched C-terminal portion of *E. coli* thioredoxin reassembled from complementary fragments have been accomplished by two-dimensional solid-state magic angle spinning spectroscopy at 17.6 T. A sample preparation procedure developed for reassembled thioredoxin results in spectra of excellent resolution. Statistical analysis of ^{13}C and ^{15}N chemical shifts accurately predicts backbone dihedral angles and secondary structure for the overwhelming majority of amino acid residues in reassembled thioredoxin, and indicates no significant perturbations in the secondary structure of the complex with respect to the full-length native protein. The idea of utilizing differentially isotopically labeled fragments for preparation of reassembled thioredoxin, introduced in this report, is anticipated to be generally applicable for structural studies of protein interfaces and protein assemblies by solid-state NMR. Resonance assignments of reassembled thioredoxin presented here are the first step in our forthcoming efforts in this area.

This work additionally demonstrates that, at high magnetic fields, resonance assignments of moderate-size proteins and protein assemblies can be established in a straightforward and time-efficient way via a combination of 2D homo- and heteronuclear correlation experiments. This approach is anticipated to be especially advantageous for systems where increasing dimensionality of the experiments is undesirable (e.g., limited sample quantities are available, resulting in compromised sensitivity, or proteins undergo conformational exchange in the intermediate time scales). These experimental protocols are expected to be generally suited for structural studies of larger proteins and protein interfaces by high-resolution solid-state NMR spectroscopy.

Acknowledgment. We gratefully acknowledge instrument time on the 750 MHz solid-state NMR spectrometer at the New York Structural Biology Center. T.P. (as adjunct faculty of CUNY), M.L.T., and A.M. are members of the New York Structural Biology Center, supported in part by the National Institutes of Health (P41 GM066354) and its Member Institutions. T.P., M.L.T., and D.M. are supported by the National Institutes of Health (5S06GM060654-04; SCORE program individual subproject). T.P. acknowledges support by the National Science Foundation CAREER development award (CHE 0350385) and of the National Institutes of Health (P20-17716, COBRE individual subproject). M.L.T. acknowledges support by the National Science Foundation POWRE award (MB 0075115) and National Institutes of Health (5G12 RR03060 from the National Center for Research Resources).

Supporting Information Available: Resonance assignments for individual amino acid residues in reassembled [$^{15}N(1-73)/^{13}C,^{15}N(74-108)$]thioredoxin. This material is available free of charge via the Internet at http://pubs.acs.org.

JA0464589