Available online at www.sciencedirect.com
ScienceDirect
Journal of Magnetic Resonance 189 (2007) 157–165
JMR Journal of Magnetic Resonance
www.elsevier.com/locate/jmr

Assignment of congested NMR spectra: Carbonyl backbone enrichment via the Entner–Doudoroff pathway

Amir Goldbourt [a], Loren A. Day [b], Ann E. McDermott [a,*]

[a] *Department of Chemistry MC3113, Columbia University, New York, NY 10027, USA*
[b] *Public Health Research Institute, 225 Warren Street, Newark, NJ 07103, USA*

Received 22 April 2007; revised 10 July 2007
Available online 17 August 2007

Abstract

In NMR spectra of complex proteins, sparse isotope enrichment can be important, in that the removal of many ^{13}C–^{13}C homonuclear *J*-couplings can narrow the lines and thereby facilitate the process of spectral assignment and structure elucidation. We present a simple scheme for selective yet extensive isotopic enrichment applicable for production of proteins in organisms utilizing the Entner–Doudoroff (ED) metabolic pathway. An enrichment scheme so derived is demonstrated in the context of a magic-angle spinning solid-state NMR (MAS SSNMR) study of Pf1 bacteriophage, the host of which is *Pseudomonas aeruginosa*, strain K (PAK), an organism that uses the ED pathway for glucose catabolism. The intact and infectious Pf1 phage in this study was produced by infected PAK cells grown on a minimal medium containing 1-^{13}C D-glucose (^{13}C in position 1) as the sole carbon source, as well as $^{15}NH_4Cl$ as the only nitrogen source. The 37 MDa Pf1 phage consists of about 93% major coat protein, 1% minor coat proteins, and 6% single-stranded, circular DNA. As a consequence of this composition and the enrichment scheme, the resonances in the MAS SSNMR spectra of the Pf1 sample were almost exclusively due to carbonyl carbons in the major coat protein. Moreover, 3D heteronuclear NCOCX correlation experiments also show that the amino acids leucine, serine, glycine, and tyrosine were not isotopically enriched in their carbonyl positions (although most other amino acids were), which is as expected based upon considerations of the ED metabolic pathway. 3D NCOCX NMR data and 2D ^{15}N–^{15}N data provided strong verification of many previous assignments of ^{15}N amide and ^{13}C carbonyl shifts in this highly congested spectrum; both the semi-selective enrichment patterns and the narrowed linewidths allowed for greater certainty in the assignments as compared with use of uniformly enriched samples alone.

Keywords: Pf1 filamentous phage; Entner–Doudoroff pathway; Solid-state NMR; Magic-angle spinning; Resonance assignment; Isotope enrichment

1. Introduction

The assignment of NMR resonances in biological macromolecules is typically a prerequisite for obtaining site-specific information. In recent years, a growing number of assignments have been reported for globular and membrane proteins studied by magic-angle spinning (MAS[1]) solid-state NMR, yielding several structures [1–5]. One of the key factors for successful assignment is the spectral resolution. For proteins of large molecular weight or for proteins containing a dominant secondary structure element (i.e., all helical or all sheet), spectral congestion and peak overlap can hinder successful assignment. Intrinsic membrane proteins, which are both large and mostly helical, provide examples of this problem. Although successful assignments have been reported in some cases [6,7], and partial site-specific assignments and amino acids type assignments of relatively narrow line widths for isolated peaks (~0.5 ppm) have also been reported [8–10], spectral overlaps hampered large-scale site-specific assignments. Several methods have been used to overcome the problem of poorly resolved peaks. Larger signal dispersion has

* Corresponding author. Fax: +1 212 932 1289.
E-mail address: aem5@columbia.edu (A.E. McDermott).

[1] *Abbreviations used:* ATCC, the global bioresource center (http://www.atcc.org/Home.cfm); CSA, chemical shift anisotropy; DARR, dipolar assisted rotational resonance; ED, Entner–Doudoroff; EMP, Embden–Meyerhoff–Parnas; MAS SSNMR, magic-angle spinning solid-state NMR; NCOCX, 3D heteronuclear NMR experiment for sequential assignment; PAK, *Pseudomonas aeruginosa* strain K.

1090-7807/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmr.2007.07.011

been achieved by means of three-dimensional and double-quantum based experiments [3] or higher magnetic fields. Spectral editing techniques to select for specific functional groups [11,12] or specific spectral regions [13–15] have also been used. It is possible to obtain better site resolution by means of isotopic editing. This is achieved by producing protein samples from a medium containing partially enriched precursors. Methyl isotope enrichment [16], alternate carbon enrichment [17], segmental isotope enrichment [18], and deuteration [19,20] have all contributed to the study of the structure and dynamics of many macromolecules in solution and have facilitated especially the study of large (~100 kDa) molecular weight proteins [21]. For MAS solid-state NMR, the use of precursors such as 2-^{13}C glycerol, 1-^{13}C D-glucose, 1,3-^{13}C glycerol, and the additions of specifically enriched or non-enriched amino acids to the growth media have proved beneficial for assignment and for structure elucidation [4,17,22,23].

In this manuscript, we present a procedure for creating a network of isolated, isotopically enriched N–C′ spin-pairs throughout the backbone of Pf1 phage major capsid protein. Backbone carbonyl carbons have been shown to be an excellent probe for local mobility in proteins [24]. Their chemical shift anisotropy (CSA) values are directly correlated with their protonation states and hydrogen bonding patterns [25,26], and carbonyl CSA values are used as restraints for structure calculations [27]. Carbonyl enriched amino acids have been incorporated into large proteins and utilized to simplify data analysis and to probe protein–protein interactions [28,29]. With respect to Pf1, its major capsid protein is small, only 46 amino acids, but it is present in about 7300 copies in the 37 MDa filamentous virion, which also contains a few copies of minor coat proteins at the ends and a circular single-stranded DNA molecule of 7349 nucleotides [30–32]. Using the motionally averaged ^{31}P and ^{13}C CSA tensors of the glycine carbonyls and DNA phosphate groups in a 1-^{13}C glycine enriched, concentrated, Pf1 solution, rapid (10^4–10^6 Hz) reorientation of the entire virus particle along its long axis was detected at high pH (8.0) while no such motion was detected at low pH (4.0) [33]. In the scheme we present here, a set of 30 carbonyl carbons have been introduced into the 46 amino acid capsid subunit as a result of the manner in which 1-^{13}C D-glucose is broken down via the Entner–Doudoroff (ED) pathway to eventually produce a limited set of ^{13}C enriched amino acids. The ED pathway, discovered in a bacterium of the genus *Pseudomonas* [34], is the principal catabolic path for glucose for a wide variety of bacteria [35], including the Pf1 host *Pseudomonas aeruginosa*, K strain (PAK). Here we have used MAS NMR experimental data obtained for a uniformly ^{15}N and partially carbonyl enriched Pf1 sample to verify the previously assigned ^{15}N and ^{13}C backbone resonances in spectra of fully enriched intact Pf1 samples [36].

2. Results

Two cross-polarization (CP) MAS ^{13}C NMR spectra of the Pf1 filamentous phage are shown in Fig. 1, demonstrating qualitatively the carbonyl enrichment pattern. One is the spectrum of a fully {^{13}C, ^{15}N} enriched Pf1 virus sample (black, referred to as Pf1^{U}), which shows typical protein ^{13}C resonances, including signals from carbonyl carbons (170–185 ppm), aromatic carbons (110–160 ppm), Cα (44–70 ppm), and other aliphatic carbons (10–44 ppm). The second spectrum (red) is of a partially enriched Pf1 sample (referred to as Pf1^{NCO}) and is dominated by the carbonyl peaks. Pf1^{NCO} was prepared as previously described [36] from Pf1 infected PAK (both phage and host obtained originally from Prof. K. Amako [31]; ATCC #25102), grown in a minimal M9 medium containing $^{15}NH_4Cl$ as sole nitrogen source, but with 1-^{13}C D-glucose as a sole carbon source, rather than fully enriched ^{13}C D-glucose as a sole carbon source. The Cα carbons are almost completely suppressed, and some intensity is detected for several methyl carbons. The aromatic rings are also not enriched in this scheme.

It is well established that when proteins are expressed in *Escherichia coli*, the presence of 1-^{13}C glucose as the carbon source results in the selective isotopic enrichment of many aliphatic carbons and partial enrichment at the carbonyl position for about one-half of the natural amino acids [22]. These results are consistent with the expected fate of the C-1 of a D-glucose precursor over the course of glucose catabolism and subsequent biosynthesis of the amino acids

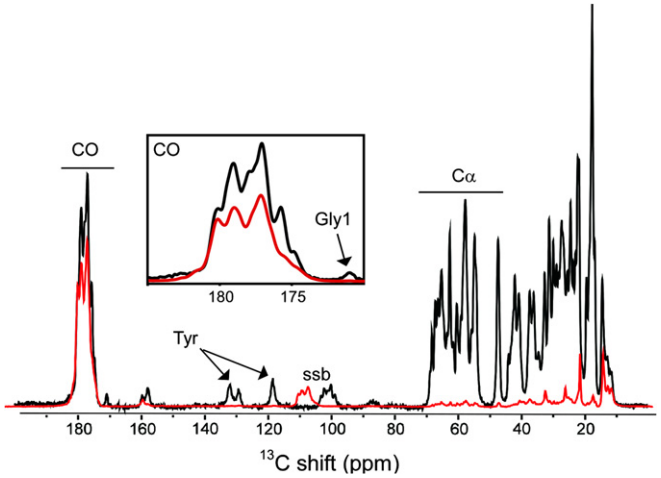

Fig. 1. 1D CP-MAS ^{13}C NMR spectra of Pf1 filamentous phage. In black is the spectrum of a fully enriched virus sample, Pf1^{U}. In red is the spectrum of Pf1^{NCO}. The relative intensities of the two spectra were set for clarity. The inset shows an expansion (170–185 ppm) of the carbonyl region. Arrows indicate the signals of the N-terminal glycine carbonyl and the tyrosine ring carbons as examples of carbon signals that are not enriched in Pf1^{NCO}. The backbone Cα and the ser and thr Cβ in the same spectral region are almost completely suppressed, and some methyl groups in the 10–40 ppm region (ile Cδ, thr Cγ, met Cγ, lys Cγ) have weak intensities. Spinning sidebands are noted as such (ssb). (For interpretation of the references in colour in this figure legend, the reader is referred to the web version of this article.)

via the Embden–Meyerhoff–Parnas pathway [37]. In contrast, *Ps. aeruginosa* utilizes the Entner–Doudoroff (ED) pathway in the first stages of glucose catabolism [34] and thus a different enrichment pattern of the amino acids results, which is dominated primarily by the carbonyl carbons.

In order to establish the enrichment pattern in Pf1^{NCO} and to verify our previously assigned carbonyl shifts we performed a 3D NCOCX heteronuclear experiment and a 2D ^{15}N–^{15}N free spin diffusion experiment. In the NCOCX experiment we observed many cross-peaks resulting from the magnetization transfer of the ^{15}N spins of a residue i (N_i) to the carbonyl carbon of the preceding residue $i-1$ (C'_{i-1}) and to additional carbons C'_j (cross-peak N_i–C'_{i-1}–C'_{i-1+j}). The main cross-peaks observed were for $j=-1$, 0, +1 and provided sequential contacts up to three residues ($i \rightarrow i-1 \rightarrow i-2$). The intensities of those cross-peaks varied between 12% and 25% of the corresponding diagonal carbonyl peaks. Some examples of those cross-peaks are shown in Fig. 2.

The reduction of the ^{13}Cα isotopes (and most other carbons) to the natural abundance level resulted in the elimination of the ^{13}C′–^{13}Cα one-bond J-coupling ($J_{C'-C\alpha}$), producing significantly narrower line widths, as shown in Fig. 3. Slices corresponding to the C′ peaks of Glu9 and Asp14 are shown for Pf1^{U} (a and c) and Pf1^{NCO} (b and d) samples. The line widths in Pf1^{NCO} are narrower by 50–70 Hz, corresponding approximately to the average value of $J_{C'-C\alpha}$, 55 Hz.

Another aspect of the ED pathway is that of the 14 different amino acids comprising the sequence of the Pf1 major coat protein, several are predicted not to be enriched when 1-^{13}C glucose is used as the sole carbon source. This prediction of non-enrichment, which is elaborated in Section 3, is confirmed on the basis of absent or very weak N_i–C'_{i-1}–C'_{i-1} cross-peaks for cases where C'_i corresponds to any of those non-enriched amino acids. In Fig. 4, 2D planes cut from the 3D NCOCX spectrum show an overlay for Pf1^{U} (blue) and the partially enriched sample Pf1^{NCO} (red). In the examples given below, the particular C′ auto-correlation peaks corresponding to site-specifically assigned Gly23 (a), Ser10 (b), Leu33 (c), and Leu38 (d) are missing from the spectrum. In some cases (not shown), peaks corresponding to one of the four amino acids tyr, ser, gly, and leu could not be resolved, however the consistent absence of C'_i–$C'_{i\pm1}$ peaks and the significantly reduced intensity of the auto-correlation peaks suggests that those signals are also suppressed.

Data from the 3D NCOCX experiment allowed us to verify many of our prior backbone nitrogen and carbonyl assignments. In order to verify further the ^{15}N assignments and support our results, we have also acquired a two-dimensional ^{15}N–^{15}N proton driven spin diffusion (PDSD) experiment [41]. Similar experiments by others detected cross-peaks that result mainly from sequential contacts along the backbone [42–44]. These contacts result from through-space magnetization transfer along typical ^{15}N–^{15}N distances of ~3.4 Å in a β-sheet and

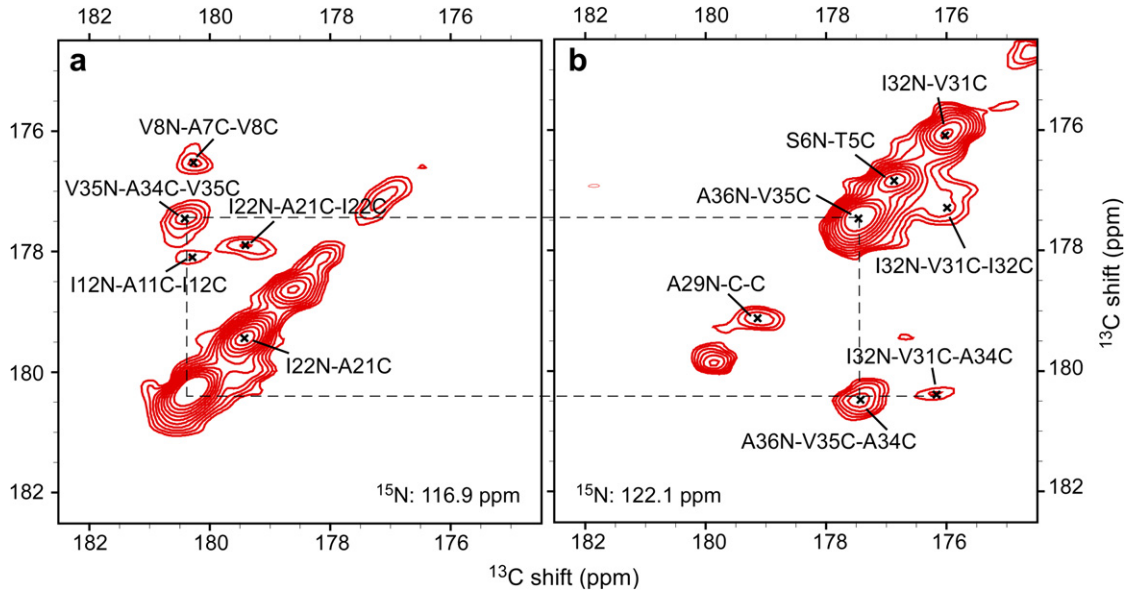

Fig. 2. Cross-peaks from a 3D NCOCX spectrum of the Pf1^{NCO} sample. Selective CP [13] was applied to obtain N–C′ transfer followed by carbon spin diffusion (500 ms) utilizing dipolar assisted rotational resonance [38,39]. Auto-correlation and cross-peaks were observed only for the carbonyl carbons. Cross-peaks to other aliphatic carbons were not observed even when if a 150 Hz line broadening apodization function was used in all dimensions. (a) A ^{13}C–^{13}C plane cut at a ^{15}N shift of 116.9 ppm showing primarily $N_i \rightarrow C'_{i-1} \rightarrow C'_i$ cross-peaks; the range of shifts for the actual maxima of the indicated peaks was 116.9–117.4 ppm. (b) A ^{13}C–^{13}C plane cut at a ^{15}N shift of 122.1 ppm showing auto-correlation peaks $N_i \rightarrow C'_{i-1} \rightarrow C'_{i-1}$, medium-range contacts $N_i \rightarrow C'_{i-1} \rightarrow C'_j$, and a signal from a single residue resulting from a weak two-bond N–C′ transfer (Ala29); the ^{15}N maxima range between 121.7 and 122.1 ppm for all indicated peaks. Ten contours are shown with multiplicative steps of 1.3. The minimum contour level is at a signal-to-noise of 10 in (a) and at a signal-to-noise level of 5.5 in (b). Noise levels were calculated using SPARKY [40] as the median of 1000 randomly sampled absolute value data points.

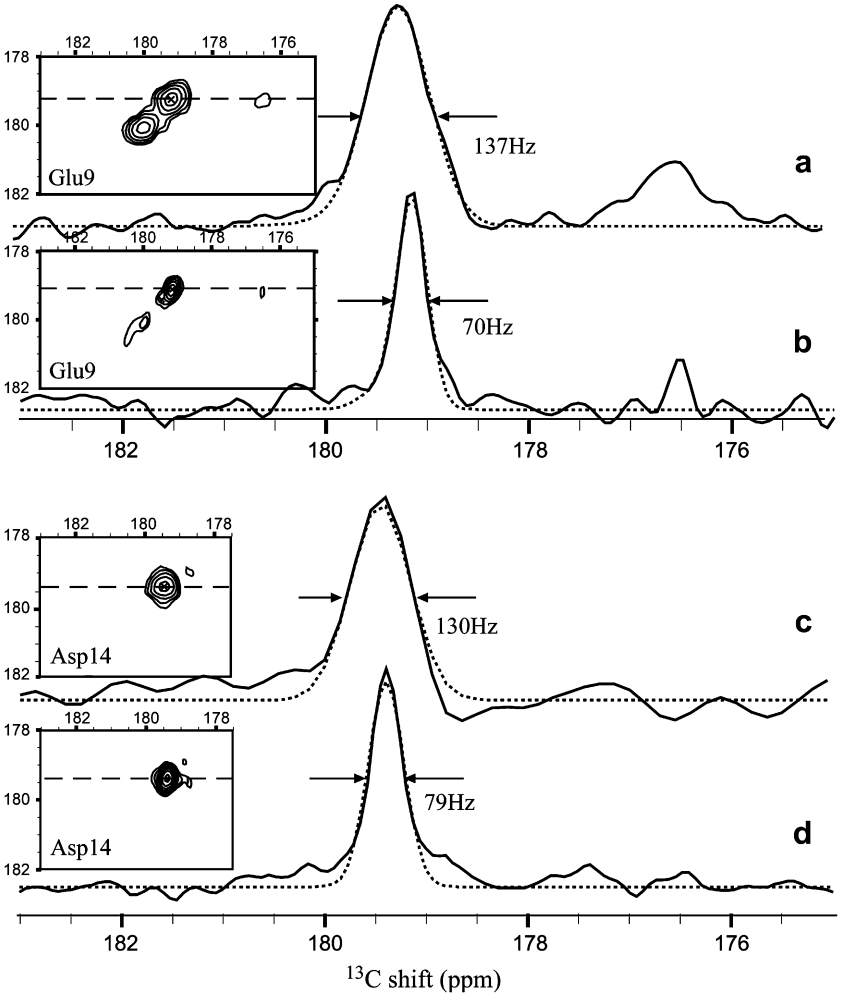

Fig. 3. Line narrowing of the carbonyl resonances from the NCOCX 3D spectrum. The line widths at half peak height (Δ_{FWHH}) were calculated using the SPARKY [40] line fitting routine. Both spectra of the uniformly enriched virus sample (Pf1^{U}) and the partially enriched virus sample (Pf1^{NCO}) were apodized using a cosine-bell function. Acquisition times were 25 and 20 ms for Pf1^{U} and Pf1^{NCO}, respectively. (a) A 1D slice extracted from the 3D N_i–C'_{i-1}–C'_{i-1} cross-peak of Glu9 of Pf1^{U}. $\Delta_{FWHH} = 137$ Hz. The 2D $^{13}C'$–$^{13}C'$ plane is shown in the inset at the ^{15}N shift of Ser10. (b) A slice corresponding to the Glu9 C′ peak taken from Pf1^{NCO}, $\Delta_{FWHH} = 70$ Hz. Inset shown for the ^{15}N shift of Ser10 from the same dataset. (c) Asp14 from Pf1^{U}, $\Delta_{FWHH} = 130$ Hz. (d) Asp14 of Pf1^{NCO}, $\Delta_{FWHH} = 79$ Hz. Insets in (c) and (d) are from a 2D $^{13}C'$–$^{13}C'$ plane extracted from the corresponding ^{15}N shift of Gly15. Experimental details are given in the caption of Fig. 2.

~2.8 Å in a right-handed α-helix. In three cases cross-peaks from longer distances have been reported. An inter-molecular contact of 4.75 Å (based on X-ray structure) has been observed for the Crh dimer [42] (a mixing time of $\tau = 4$ s was used). In SH3, mostly sequential contacts have been observed ($\tau = 3$ s) however several ambiguous long-range (3–6 Å) constraints have been used for the structure calculation [43], and in GB1 several two-bond transfers have been observed ($\tau = 4$ s) [44]. The ^{15}N–^{15}N spectrum of Pf1^{NCO} ($\tau = 4$ s) shows many sequential contacts and also several medium-range ($i \rightarrow i \pm 2$, 3) contacts. The region in the spectrum corresponding to the glycine cross-peaks was best resolved, and is shown in Fig. 5 with some key assignments. All Gly_i–$X_{i\pm1}$ cross-peaks are indicated (where X is a neighboring residue), with the exception of Gly23–Gly24, which was detected near the diagonal, outside the region shown here. We have observed several two-bond transfers (G24–I22, G24–I26, G17–M19, G37–V35), and some weaker three-bond transfers (G24 → A21 and G37 → Y40). A typical N_i–N_{i+2} distance in a right-handed α-helix is ~4.2 Å and a N_i–N_{i+3} distance is around 4.8 Å. For a 3_{10}-helix this distance would be ~5.9 Å, and very unlikely to be observed at all since the ^{15}N–^{15}N spin diffusion efficiency is ~6 times smaller than that of ^{13}C and for a mixing time of 4 s, distances above 5 Å have not been observed [42].

The chemical shifts of Pf1 derived from Pf1^{U} and Pf1^{NCO} have been deposited in the BioMagResBank under Accession Number 15138.

3. Discussion

Any isotopic dilution for NMR studies that results in line narrowing and spectral dilution will support and

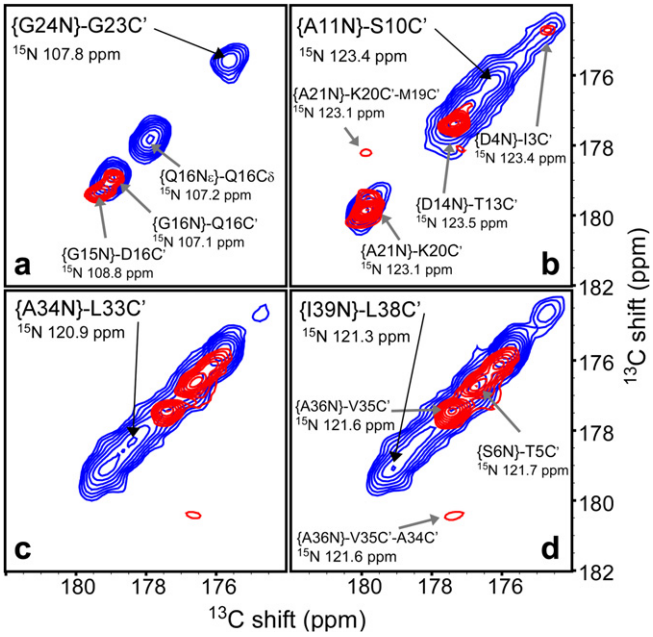

Fig. 4. ^{13}C–^{13}C planes extracted from 3D NCOCX spectra of Pf1 filamentous phage. Red and blue colors correspond to the partially ($Pf1^{NCO}$) and fully ($Pf1^{U}$) enriched samples, respectively. (a) The sequential correlation Gly24–Gly23 (175.7 ppm) is suppressed, confirming that glycine C′ is not enriched. (b) The sequential correlation Ala11–Ser10 (176.4 ppm) is absent, confirming the suppression of serine $^{13}C'$. (c and d) Lack of leucine $^{13}C'$ is demonstrated in two examples: Ala34–Leu33 (178.4 ppm) (c) and Ile39–Leu38 (179.1 ppm) (d). In all the examples above the auto-correlation C′–C′ peak is unique and based on our assignments, and no overlap is expected with other residues. Additional peaks observed in the spectra are indicated with gray arrows, and their ^{15}N shifts at the peak maximum are noted. Fifteen contours at multiplicative steps of 1.3 are shown in both spectra and are set to a minimal signal-to-noise of 10. Experimental details are similar to those given in Fig. 2. (For interpretation of the references in colour in this figure legend, the reader is referred to the web version of this article.)

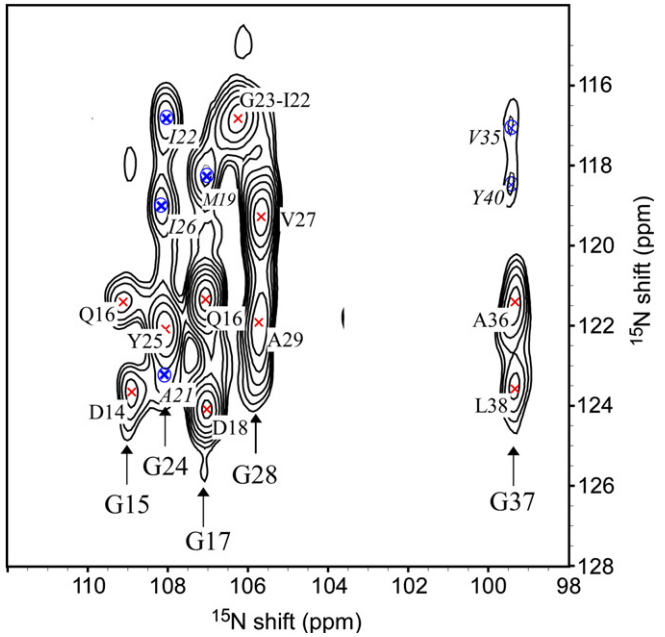

Fig. 5. A ^{15}N–^{15}N spin diffusion experiment from $Pf1^{NCO}$. The mixing time was 4 s and no proton irradiation was employed. The cross-peaks of the glycine residues with neighboring amino acids are shown. Peaks resulting from sequential contacts are indicated by a red mark, with the corresponding glycine neighbor noted by an arrow (G23–I22 is explicitly noted as such). Several peaks result from medium-range ($i \rightarrow i \pm 2, 3$) magnetization transfer and are indicated in blue and noted by an italic font. Eight contour levels are drawn starting from a signal-to-noise of 6 and with multiplicative steps of 1.3×. In cases of assignment ambiguities due to congested peaks, assignment was based upon the spin pair with the shortest distance. Additional experimental details are given in Section 4. (For interpretation of the references in colour in this figure legend, the reader is referred to the web version of this article.)

enable the assignment of large proteins with highly congested spectra. In the case of Pf1, both line width reduction and spin dilution have been achieved through the enrichment of a limited set of the backbone carbonyls resulting from our use of the glucose catabolic pathway of the host organism *Ps. aeruginosa*. A general scheme describing the pathways leading to the pattern we observed for the Pf1 coat protein is shown in Fig. 6. According to the ED pathway [34,45], a 1-^{13}C glucose precursor will be converted to a pyruvate enriched at the carboxylic carbon and to a non-enriched glyceraldehyde-3-phosphate (G3P) molecule. The non-enriched G3P then enters the lower Embden–Meyerhoff–Parnas (EMP) pathway to generate non-enriched serine and glycine amino acids, and an additional non-enriched pyruvate molecule. Tyrosine is a product of chorismate, which is generated from erythrose-4-phosphate (E4P), an intermediate of the pentose phosphate pathway (PPP), and phosphoenolpyruvate (PEP), an intermediate of the lower EMP pathway. The fourth amino acid that is not enriched in $Pf1^{NCO}$ is leucine. The carbonyl group in leucine originates from the acetyl group of acetyl coenzyme A when it condenses with the carbonyl-enriched α-ketoisovalerate as part of the valine biosynthesis pathway. The ^{13}C isotope of α-ketoisovalerate, originating from the pyruvate carboxylic carbon, is released as $^{13}CO_2$. As shown in Fig. 6 the residues serine, glycine, leucine and tyrosine remain non-enriched when glycolysis of 1-^{13}C glucose is by the ED pathway, and accordingly resonances from the 7 gly, 3 ser, 4 leu, and 2 tyr in Pf1 are indeed suppressed in our spectra (Fig. 4, and data not shown). However, several of them show very weak auto-correlation peak intensities on the order of 10% of other resolved auto-correlation peaks.

In addition to the enrichment of the backbone carbonyls, the CPMAS spectrum of $Pf1^{NCO}$ (Fig. 1, red) indicates that several methyl groups are also partially enriched. Based on the spectral assignment of Pf1 [36], those methyl peaks belong to isoleucine Cδ (11–15 ppm), threonine Cγ (22 ppm), lysine Cγ (K20, 26.6 ppm), and methionine Cγ (M19, 32.9 ppm). These observations are in agreement with the direct conversion of 1-^{13}C pyruvate and ^{13}C enriched carbonate to 1,4-^{13}C oxaloacetic acid, in which the carbon in position 4 originates from an enriched carbonate ion.

Prior spectral assignments of the resonances of 237 out of 257 ^{15}N and ^{13}C nuclei of $Pf1^{U}$ were achieved despite

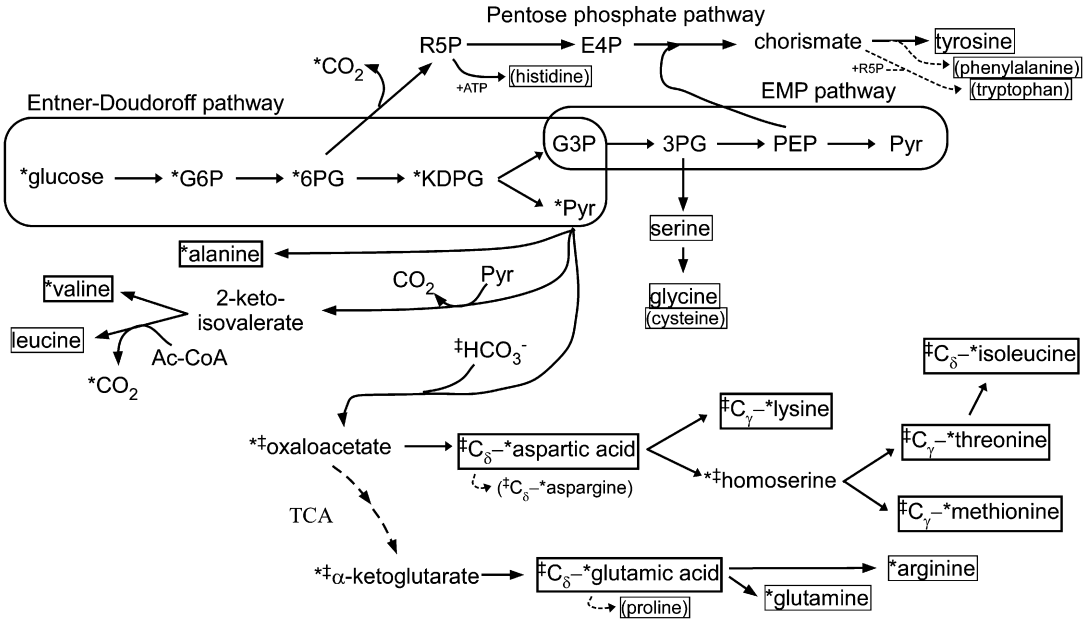

Fig. 6. The key biosynthetic pathways determining the enrichment pattern of $Pf1^{NCO}$ resulting from metabolism of a 1-^{13}C glucose precursor. A star (*) indicates amino acids that are enriched in the backbone carbonyl (e.g., *alanine) and pathway intermediates enriched at the carboxylic acid carbon. A double dagger (‡) indicates lower enrichment level stemming from $^{13}CO_2$. The enriched sidechain carbons of several amino acids are noted specifically (e.g., ‡Cγ-threonine). Oxaloacetate is enriched in positions 1 and 4, where carbon-1 corresponds to the pyruvate carboxylic group and carbon-4 comes from partially enriched carbonate ions in the growth medium. The dashed arrows indicate that α-ketoglutarate is a product of oxaloacetate in the TCA cycle, however the carbonyl enrichment in α-ketoglutarate (as observed in gln, glu, and arg) must come from another source since carbon-1 of oxaloacetate is released as CO_2. Amino acids not present in the Pf1 major coat protein are indicated in parenthesis. The following abbreviations have been used: G6P, glucose-6-phosphate; 6PG, 6-phosphogluconate; R5P, ribose-5-phosphate; E4P, erythrose-4-phosphate; KDPG, 2-keto-3-deoxy-6-phosphogluconate; G3P, glyceraldehyde-3-phosphate; 3PG, 3-phospho-glycerate; PEP, phosphoenolpyruvate; Pyr, pyruvate; Ac-CoA, acetyl coenzyme A; *‡oxaloacetate, *1,‡4-^{13}C oxaloacetate. *‡Homoserine, *1,‡4-^{13}C homoserine. Oxaloacetate and α-ketoglutarate are part of the TCA, tricarboxylic acid cycle.

the overlapping ^{15}N amide and ^{13}C carbonyl regions in this highly α-helical protein [36]. Verification of those assignments is crucial for successful structure determination. Assignments from the present 2D and 3D experiments on $Pf1^{NCO}$ are given in Fig. 7 together with our prior assignments. In Fig. 7a the ^{15}N shifts obtained from the 2D ^{15}N–^{15}N spin diffusion experiment (green triangles) and from the 3D NCOCX experiment (red squares) are com-

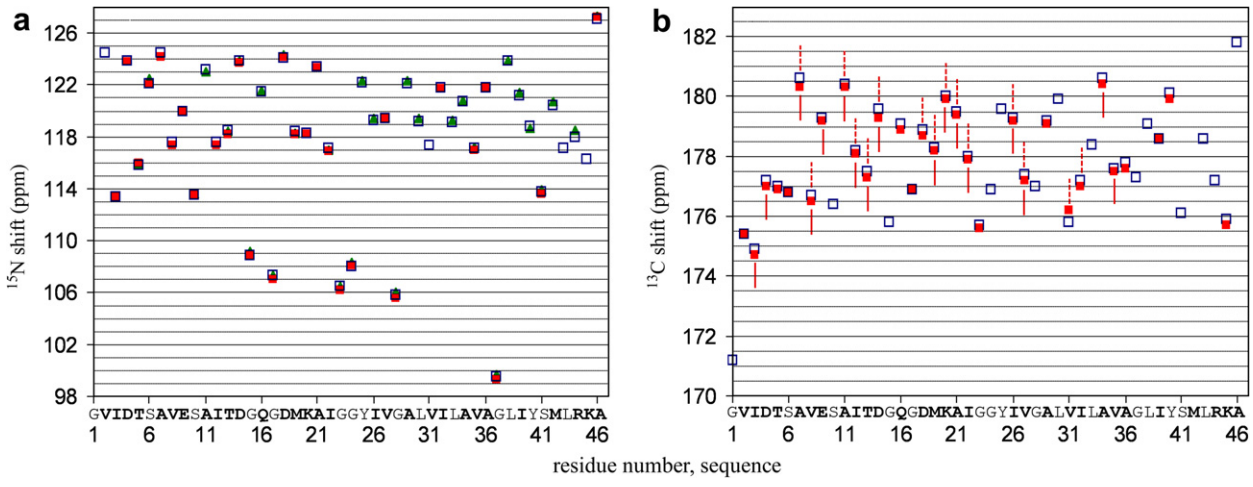

Fig. 7. (a) Comparison of the ^{15}N shifts from $Pf1^{U}$ (blue empty squares; size of 0.4 ppm) and the shifts extracted from the 2D ^{15}N–^{15}N (green triangles) and 3D NCOCX (red square) spectra of $Pf1^{NCO}$. (b) Comparison of the backbone carbonyl (^{13}C′) shifts from $Pf1^{U}$ (blue empty squares; size of 0.3 ppm) and the ^{13}C′ shifts from the NCOCX spectrum of $Pf1^{NCO}$ (red squares). The two types of cross-peaks observed in the NCOCX experiment are indicated as follows: red solid horizontal lines below the data points represent a two-residue-long $N_i \rightarrow C'_{i-1} \rightarrow C'_{i-2}$ cross-peaks. Red dashed horizontal lines above the data points represent $N_i \rightarrow C'_{i-1} \rightarrow C'_i$ cross-peaks. ^{13}C′ resonances that have no horizontal lines were identified only as auto-correlation $N_i \rightarrow C'_{i-1} \rightarrow C'_{i-1}$ peaks. The amino acid sequence of Pf1 is shown in the bottom with the residue number. Residues in bold are enriched in the C′ position. (For interpretation of the references in colour in this figure legend, the reader is referred to the web version of this article.)

pared to the previously assigned ^{15}N shifts (blue empty squares, size of 0.4 ppm). As can be seen in Fig. 7, the chemical shifts of 41 backbone ^{15}N nuclei observed by the two methods are in excellent agreement with our previous assignments.

In Fig. 7b the $^{13}C'$ shifts obtained from $Pf1^{NCO}$ (red squares) are compared to the previously assigned shifts (blue empty squares, size of 0.3 ppm). The backbone carbonyl shifts (red) have been inferred from several types of cross-peaks. Cross-peaks resulting from $N_i \rightarrow C'_{i-1} \rightarrow C'_{i-2}$ magnetization transfer are indicated by the solid lines below the data points. Cross-peaks resulting from $N_i \rightarrow C'_{i-1} \rightarrow C'_i$ magnetization transfer are indicated by the dashed lines above the data points. Auto-correlation peaks ($N_i \rightarrow C'_{i-1} \rightarrow C'_{i-1}$) account for the remaining observed C′ resonances, and have also been observed for many other residues. Those auto-correlation peaks have also proved useful for verifying many of our assignments and are comparable in signal-to-noise to those observed in $Pf1^{U}$ despite the much longer mixing time (500 ms in $Pf1^{NCO}$ vs. 10 ms in $Pf1^{U}$). In comparison, auto-correlation peaks of Cα carbons are significantly reduced in long mixing time spectra (500 ms, data not shown). Also, backbone carbonyl carbons have been shown to experience longer T_1 relaxation times then alpha-carbons or other aliphatic carbons [46], thus the use of carbonyl enriched proteins could prove useful for obtaining long-range distance constraints for use in a structure calculation.

Of the 16 amino acids in the coat protein sequence that were not enriched (noted with regular font in the abscissa label of Fig. 7), four of them (Ser6, Gly17, Gly23, Tyr40) have been observed on the diagonal as weak auto-correlation peaks, with signal intensities of around 10% when compared to other auto-correlation peaks. No cross-peaks in our spectra could be associated with those residues and their detection may be a result of some minor fractional scrambling of the ^{13}C isotopes. Overall the assignments of 31 carbonyl shifts, which include the four residues mentioned above, have been confirmed, and their shifts are in excellent agreement with our previous results.

The enrichment pattern we show here is based on the ED pathway that several bacterial species utilize as their main route for the break down of glucose, even when the enzymes of the EMP pathway are present [35]. We have directly observed high level of backbone carbonyl enrichment in Pf1 for ten amino acids (glu, gln, arg, asp, met, thr, lys, ile, ala, val). Two additional amino acids (pro and asn) that are not in the coat protein of Pf1 are likely to be enriched as well via the glucose catabolism of the ED pathway when 1-^{13}C glucose is the carbon source. The eight remaining amino acids that are not enriched in this manner are, in groups of similar biosynthetic pathways: gly, ser, cys; tyr, trp, phe, his; leu. For proteins expressed in *E. coli*, in which glucose is broken down via the EMP pathway, other schemes have been commonly employed that involve several different enrichment patterns of the backbone carbonyls. While growth on 1,3-^{13}C glycerol as a sole carbon source [4] (or with addition of $NaH^{12}CO_3$) results in enrichment of all C′ carbons except for that of leu, growth of *E. coli* on 2-^{13}C glycerol ($NaH^{13}CO_3$) results in eight amino acids not being enriched (gly, ser, cys; tyr, trp, phe; ala, val, with no information for his). The enrichment pattern reported for *E. coli* grown on 1-^{13}C glucose [22] via the EMP pathway has a low fractional enrichment (~25%) of several residues, and missing entirely are those amino acids we observed missing for the ED pathway (gly, ser, cys; tyr, trp, phe; leu), as well as three more (lys, ala, and val). The use of partially enriched succinate produced yet entirely different patterns [47]. Thus there is great potential for taking advantage of this natural ED pathway for achieving useful new enrichment patterns not obtainable by the other methods. That the ED pathway is also used by *E. coli* when it is grown on gluconate [48–50], expands the possible use of this scheme to many other proteins that are to be isotopically enriched for NMR studies.

In summary, the selective enrichment of backbone carbonyls in the Pf1 bacteriophage proved useful for verifying the backbone $^{13}C'$ and ^{15}N shifts. This scheme can be used as a building block for other new enrichment patterns that involve also the addition of enriched aliphatic carbons. Such patterns are expected to enable the acquisition by NMR of medium- and long-range order contacts as useful structural constraints. Also, the enrichment of spin-pairs along the backbone creates a network of isolated two-spin systems that gives an opportunity for the development of additional experiments, which can be easily interpreted theoretically since any effects of additional spins is almost negligible.

4. Materials and methods

4.1. Sample preparation

Uniformly enriched ^{13}C and ^{15}N Pf1 phage ($Pf1^{U}$) was purified and characterized as described before [36] by culturing infected *Pseudomonas aeruginosa* K strain (PAK) on a minimal M9 medium containing uniformly labeled ^{13}C D-glucose and $^{15}NH_4Cl$ as the sole carbon and nitrogen sources, respectively. The Pf1 sample enriched selectively in its carbonyls ($Pf1^{NCO}$) was prepared the same way, except that 1-^{13}C glucose was the sole carbon source. The virus samples were formulated for NMR with a cryo-protectant by making the 1 mg/ml solution 22% PEG 8000 and 30% ethylene glycol. Subsequent addition of 5 mM $MgCl_2$ resulted in quantitative precipitation. Approximately 40 μl of wet mass containing 10 mg virus were packed into a 4 mm Bruker rotor or a 4 mm thick-wall Varian rotor and used for the NMR experiments.

4.2. NMR experiments

1D ^{13}C CP-MAS (cross-polarization magic-angle spinning) [51] spectra of $Pf1^{NCO}$ and $Pf1^{U}$ were performed on

a Varian InfinityPlus spectrometer operating at a proton frequency of 600 MHz for protons, and on a 750 MHz Bruker DRX spectrometer, respectively. The following parameters have been used for Pf1^{NCO} and Pf1^{U}, respectively. CP-MAS was performed with the Hartman–Hahn (HH) condition ($\nu_{1H} - \nu_{1N} = n\nu r$) set to $n = 1$ (ν_{1X} the power level of channel X, νr the spinning frequency). CP-MAS contact times: 2.0 and 1.25 ms. Two-pulse phase modulation (TPPM) proton decoupling: $\nu_{1H} \sim 75$ kHz, phase 15° and pulse length 6.5 μs in both cases. Relaxation times 3 s. Acquisition lengths ~26 ms. Spinning frequencies 10.5 and 15 kHz. Processing was performed with zero-fill to 16K points and a cosine-bell apodization in both cases.

A 2D ^{15}N–^{15}N spin diffusion experiment [41] without proton irradiation was acquired on a 750 MHz Bruker spectrometer with the following parameters: proton–nitrogen CP contact time 2 ms with a 20% ramp on the proton channel and a CP-MAS at the $n = 1$ HH condition, ^{15}N–^{15}N spin diffusion mixing time 4 s, TPPM proton decoupling ~75 kHz/15°/6.4 μs, spinning rate 13.2 kHz, acquisition points 256 (t_1) × 1024 (t_2) employing time proportional phase increment (TPPI) [52] in t_1, spectral width 13.2 kHz (t_1) × 25 kHz (t_2). Several processing scripts were used. In t_1 (indirect dimension), apodization was done with a square cosine-bell or with exponential broadening, with and without forward linear prediction and with zero filling to 4096 points. In t_2 (acquisition dimension) a squared cosine-bell apodization function or a Lorentz-to-Gauss transformation were used, followed by zero filling to 4096 points and a polynomial baseline correction of the Fourier transformed spectrum.

A 3D NCOCX experiment was acquired on a Bruker 750 MHz spectrometer. In this experiment the magnetization of a given backbone ^{15}N nucleus is transferred to a carbonyl carbon (C′) of the preceding residue by means of a selective CP transfer [13,53]. C′ magnetization is then transferred to additional carbons by means of a homonuclear ^{13}C–^{13}C mixing enhanced by proton irradiation at the $n = 1$ rotational resonance condition, $\nu_{1H} = \nu r$ [38,39]. $^{15}N_i$, $^{13}C'_{i-1}$, and $^{13}C_j$ define the three dimensions t_1, t_2, and t_3. The following experimental parameters have been employed: proton–nitrogen CP contact time 2 ms (proton ramp 20%, HH condition $n = 1$), nitrogen-carbonyl ramped selective CP [13,53] of 4 ms (10% ramp on ^{13}C, $\nu_{1C} = 2.5 \times \nu r$, $\nu_{1N} = 1.5 \times \nu r$), TPPM proton decoupling ~75 kHz/15°/6.4 μs, DARR/RAD (dipolar assisted rotational resonance [38]/RF assisted spin diffusion [39]) mixing of 500 ms ($\nu_{1H} = \nu r = 13.2$ kHz), spectral widths 5.3 kHz (t_1); 3.3 kHz (t_2, corresponding to a dwell time equaling four times the rotor-spinning period of 13.2^{-1} μs); 100 kHz (t_3, acquisition dimension). Various processing schemes have been employed that optimize both line intensities (exponential broadening) and site resolution (Cosine-bell, Lorentz-to-Gauss transforms). Sample temperature (corrected for spinning) was set to 290 K, ensuring that Pf1 was in its high temperature form.

All data were processed using NMRPipe [54] and the multi-dimensional data were analyzed using SPARKY [40].

Acknowledgments

We thank Michael J. Harris for useful discussions regarding bacterial pathways. This work was supported by funds from PHRI (L.A.D.) and a grant from NSF, MCB 0316248 (A.E.M.). A.E.M. is a member of the New York Structural Biology Center (NYSBC). Support for the NYSBC has been provided by NIH/NIGMS through Grant P41 GM66354.

References

[1] S.G. Zech, A.J. Wand, A.E. McDermott, Protein structure determination by high-resolution solid-state NMR spectroscopy: application to microcrystalline ubiquitin, J. Am. Chem. Soc. 127 (2005) 8618–8626.
[2] F. Castellani, B.J. van Rossum, A. Diehl, K. Rehbein, H. Oschkinat, Determination of solid-state NMR structures of proteins by means of three-dimensional N-15–C-13–C-13 dipolar correlation spectroscopy and chemical shift analysis, Biochemistry 42 (2003) 11476–11483.
[3] C.E. Hughes, M. Baldus, G.A. Webb, Magic-Angle-Spinning Solid-State NMR Applied to Polypeptides and Proteins, Ann. Rep. NMR Spectrosc., Academic Press, 2005, pp. 121–158.
[4] F. Castellani, B. van Rossum, A. Diehl, M. Schubert, K. Rehbein, H. Oschkinat, Structure of a protein determined by solid-state magic-angle-spinning NMR spectroscopy, Nature 420 (2002) 98–102.
[5] A. Lange, S. Becker, K. Seidel, K. Giller, O. Pongs, M. Baldus, A concept for rapid protein-structure determination by solid-state NMR spectroscopy, Angew. Chem. Int. Ed. 44 (2005) 2089–2092.
[6] A.J.v. Gammeren, F.B. Hulsbergen, J.G. Hollander, H.J.M.d. Groot, Residual backbone and side-chain ^{13}C and ^{15}N resonance assignments of the intrinsic transmembrane light-harvesting 2 protein complex by solid-state magic angle spinning NMR spectroscopy, J. Biomol. NMR 31 (2005) 279–293.
[7] O.C. Andronesi, S. Becker, K. Seidel, H. Heise, H.S. Young, M. Baldus, Determination of membrane protein structure and dynamics by magic-angle-spinning solid-state NMR spectroscopy, J. Am. Chem. Soc. 127 (2005) 12965–12974.
[8] M. Hiller, L. Krabben, K.R. Vinothkumar, F. Castellani, B.J. van Rossum, W. Kuhlbrandt, H. Oschkinat, Solid-state magic-angle spinning NMR of outer-membrane protein G from *Escherichia coli*, Chembiochem 6 (2005) 1679–1684.
[9] H.L. Frericks, D.H. Zhou, L. Yap, R.B. Gennis, C.M. Rienstra, Magic-angle spinning solid-state NMR of a 144 kDa membrane protein complex: *E. coli* cytochrome bo3 oxidase, J. Biomol. NMR V36 (2006) 55–71.
[10] Ying Li, D.A. Berthold, H.L. Frericks, R.B. Gennis, C.M. Rienstra, Partial ^{13}C and ^{15}N chemical-shift assignments of the disulfide-bond-forming enzyme DsbB by 3D magic-angle spinning NMR spectroscopy, ChemBioChem 8 (2007) 434–442.
[11] S. Jehle, M. Hiller, K. Rehbein, A. Diehl, H. Oschkinat, B.J. van Rossum, Spectral editing: selection of methyl groups in multidimensional solid-state magic-angle spinning NMR, J. Biomol. NMR 36 (2006) 169–177.
[12] S. Jehle, K. Rehbein, A. Diehl, B.J. van Rossum, Amino-acid selective experiments on uniformly C-13 and N-15 labeled proteins by MAS NMR: filtering of lysines and arginines, J. Magn. Reson. 183 (2006) 324–328.
[13] M. Baldus, A.T. Petkova, J. Herzfeld, R.G. Griffin, Cross polarization in the tilted frame: assignment and spectral simplification in heteronuclear spin systems, Mol. Phys. 95 (1998) 1197–1207.

[14] D.H.H. Zhou, K.D. Kloepper, K.A. Winter, C.M. Rienstra, Band-selective C-13 homonuclear 3D spectroscopy for solid proteins at high field with rotor-synchronized soft pulses, J. Biomol. NMR 34 (2006) 245–257.
[15] R. Verel, M. Baldus, M. Ernst, B.H. Meier, A homonuclear spin-pair filter for solid-state NMR based on adiabatic-passage techniques, Chem. Phys. Lett. 287 (1998) 421–428.
[16] V. Tugarinov, V. Kanelis, L.E. Kay, Isotope labeling strategies for the study of high-molecular-weight proteins by solution NMR spectroscopy, Nat. Protocols 1 (2006) 749–754.
[17] D.M. LeMaster, D.M. Kushlan, Dynamical mapping of *E. coli* thioredoxin via ^{13}C NMR relaxation analysis, J. Am. Chem. Soc. 118 (1996) 9255–9264.
[18] G. Wider, K. Wuthrich, NMR spectroscopy of large molecules and multimolecular assemblies in solution, Curr. Opin. Struct. Biol. 9 (1999) 594–601.
[19] D.M. LeMaster, F.M. Richards, NMR sequential assignment of *Escherichia coli* thioredoxin utilizing random fractional deuteration, Biochemistry 27 (1988) 142–150.
[20] K.H. Gardner, L.E. Kay, The use of H-2, C-13, N-15 multidimensional NMR to study the structure and dynamics of proteins, Annu. Rev. Biophys. Biomol. Struct. 27 (1998) 357–406.
[21] N.K. Goto, L.E. Kay, New developments in isotope labeling strategies for protein solution NMR spectroscopy, Curr. Opin. Struct. Biol. 10 (2000) 585–592.
[22] M. Hong, Determination of multiple phi-torsion angles in proteins by selective and extensive C-13 labeling and two-dimensional solid-state NMR, J. Magn. Reson. 139 (1999) 389–401.
[23] L.-Y. Lian, D.A. Middleton, Labelling approaches for protein structural studies by solution-state and solid-state NMR, Prog. NMR Spectrosc. 39 (2001) 171–190.
[24] K.T. Dayie, G. Wagner, Carbonyl carbon probe of local mobility in ^{13}C, ^{15}N-enriched proteins using high-resolution nuclear magnetic resonance, J. Am. Chem. Soc. 119 (1997) 7797–7806.
[25] R.A. Burton, N. Tjandra, Residue-specific C-13′ CSA tensor principal components for ubiquitin: correlation between tensor components and hydrogen bonding, J. Am. Chem. Soc. 129 (2007) 1321–1326.
[26] Z.T. Gu, R. Zambrano, A. McDermott, Hydrogen-bonding of carboxyl groups in solid-state amino-acids and peptides—comparison of carbon chemical shielding, infrared frequencies, and structures, J. Am. Chem. Soc. 116 (1994) 6368–6372.
[27] R.S. Lipsitz, N. Tjandra, Carbonyl CSA restraints from solution NMR for protein structure refinement, J. Am. Chem. Soc. 123 (2001) 11065–11066.
[28] K. Kato, C. Sautes-Fridman, W. Yamada, K. Kobayashi, S. Uchiyama, H. Kim, J. Enokizono, A. Galinha, Y. Kobayashi, W.H. Fridman, Y. Arata, I. Shimada, Structural basis of the interaction between IgG and Fc gamma receptors, J. Mol. Biol. 295 (2000) 213–224.
[29] N. Nishida, F. Motojima, M. Idota, H. Fujikawa, M. Yoshida, I. Shimada, K. Kato, Probing dynamics and conformational change of the GroEL–GroES complex by C-13 NMR spectroscopy, J. Biochem. 140 (2006) 591–598.
[30] L.A. Day, C.J. Marzec, S.A. Reisberg, A. Casadevall, DNA packing in filamentous bacteriophages, Annu. Rev. Biophys. Biophys. Chem. 17 (1988) 509–539.
[31] K. Takeya, K. Amako, A rod-shaped Pseudomonas phage, Virology 28 (1966) 163–165.
[32] D.F. Hill, N.J. Short, R.N. Perham, G.B. Petersen, DNA sequence of the filamentous bacteriophage Pf1, J. Mol. Biol. 218 (1991) 349–364.
[33] P. Tsang, S.J. Opella, Pf1 virus particle dynamics, Biopolymers 25 (1986) 1859–1864.
[34] N. Entner, M. Doudoroff, Glucose and gluconic acid oxidation of Pseudomonas Saccharophila, J. Biol. Chem. 196 (1952) 853–862.
[35] T. Fuhrer, E. Fischer, U. Sauer, Experimental identification and quantification of glucose metabolism in seven bacterial species, J. Bacteriol. 187 (2005) 1581–1590.
[36] A. Goldbourt, B.J. Gross, L.A. Day, A.E. McDermott, Filamentous phage studied by magic-angle spinning NMR: resonance assignment and secondary structure of the coat protein in Pf1, J. Am. Chem. Soc. 129 (2007) 2338–2344.
[37] D. Voet, J.G. Voet, Biochemistry, New York, 1995.
[38] K. Takegoshi, S. Nakamura, T. Terao, ^{13}C–^{1}H dipolar-assisted rotational resonance in magic-angle spinning NMR, Chem. Phys. Lett. 344 (2001) 631–637.
[39] C.R. Morcombe, V. Gaponenko, R.A. Byrd, K.W. Zilm, Diluting abundant spins by isotope edited radio frequency field assisted diffusion, J. Am. Chem. Soc. 126 (2004) 7196–7197.
[40] T.D. Goddard, D.G. Kneller, SPARKY 3.1.1, University of California, San Francisco, CA.
[41] N.M. Szeverenyi, M.J. Sullivan, G.E. Maciel, Observation of spin exchange by two-dimensional fourier transform ^{13}C cross polarization-magic-angle spinning, J. Magn. Reson. 47 (1982) 462–475.
[42] N. Giraud, M. Blackledge, A. Bockmann, L. Emsley, The influence of nitrogen-15 proton-driven spin diffusion on the measurement of nitrogen-15 longitudinal relaxation times, J. Magn. Reson. 184 (2007) 51–61.
[43] B. Reif, B.J. van Rossum, F. Castellani, K. Rehbein, A. Diehl, H. Oschkinat, Characterization of ^{1}H–^{1}H distances in a uniformly ^{2}H, ^{15}N-labeled SH3 domain by MAS solid-state NMR spectroscopy, J. Am. Chem. Soc. 125 (2003) 1488–1489.
[44] W.T. Franks, B.J. Wylie, S.A. Stellfox, C.M. Rienstra, Backbone conformational constraints in a microcrystalline U-N-15-Labeled protein by 3D dipolar-shift solid-state NMR spectroscopy, J. Am. Chem. Soc. 128 (2006) 3154–3155.
[45] I.J. Stern, C.H. Wang, C.M. Gilmour, Comparative catabolism of carbohydrates in Pseudomonas species, J. Bacter. 79 (1960) 601–611.
[46] A. Perry, M.P. Stypa, B.K. Tenn, K.K. Kumashiro, Solid-state ^{13}C NMR reveals effects of temperature and hydration on elastin, Biophys. J. 82 (2002) 1086–1095.
[47] A.J.V. Gammeren, F.B. Hulsbergen, J.G. Hollander, H.J.M.d. Groot, Biosynthetic site-specific ^{13}C labeling of the light-harvesting 2 protein complex: a model for solid state NMR structure determination of transmembrane proteins, J. Biomol. NMR 30 (2004) 267–274.
[48] R. Zablotny, D.G. Fraenkel, Glucose and gluconate metabolism in a mutant of *Escherichia coli* lacking gluconate-6-phosphate dehydrase, J. Bacteriol. 93 (1967) 1579–1581.
[49] N. Peekhaus, T. Conway, What's for dinner?: Entner–Doudoroff metabolism in *Escherichia coli*, J. Bacteriol. 180 (1998) 3495–3502.
[50] R.C. Eisenberg, W.J. Dobrogosz, Gluconate metabolism in *Escherichia coli*, J. Bacteriol. 93 (1967) 941–949.
[51] E.O. Stejskal, J. Schaefer, J.S. Waugh, Magic-angle spinning and polarization transfer in proton-enhanced NMR, J. Magn. Reson. 28 (1977) 105–112.
[52] R.R. Ernst, G. Bodenhausen, A. Wokaun, Principles of Nuclear Magnetic Resonance in One and Two Dimensions, Clarendon Press, Oxford, 1987.
[53] J. Schaefer, R.A. McKay, E.O. Stejskal, Double-cross-polarization NMR of solids, J. Magn. Reson. 34 (1979) 443.
[54] F. Delaglio, S. Grzesiek, G.W. Vuister, G. Zhu, J. Pfeifer, A. Bax, Nmrpipe—a multidimensional spectral processing system based on unix pipes, J. Biomol. NMR 6 (1995) 277–293.