# 2D and 3D NMR Study of Phenylalanine Residues in Proteins by Reverse Isotopic Labeling

**Geerten W. Vuister,[§,†] Soon-Jong Kim,[‡] Carl Wu,[‡] and Ad Bax*[,§]**

*Contribution from the Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, and Laboratory of Biochemistry, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20892*

*Received June 16, 1994*[⊕]

**Abstract:** A protein isotopic labeling strategy is presented which offers improved NMR sensitivity and resolution relative to the commonly used uniform $^{13}C$ labeling approach. Incorporation of specific residues at natural abundance into an otherwise fully $^{13}C$-enriched protein yields $^{1}H$ line widths for the unlabeled residues which are not adversely affected by $^{13}C$ and makes it possible to selectively focus on interactions between the unlabeled residues and the remainder of the protein. Modifications of $^{13}C$ editing and $^{12}C$ filtering procedures are described which optimize their sensitivity and resolution. The experiments are used to obtain complete assignments for all 10 phenylalanine aromatic spin systems in the DNA-binding domain of *Drosophila* heat shock factor and to obtain a large number of structurally important long-range NOE constraints. A novel $J$ correlation experiment is also described which makes it possible to measure $H^{\alpha}$–$H^{\beta}$ $J$ couplings in larger proteins and yields quantitative values for all 10 phenylalanines in the DNA-binding domain of *Drosophila* heat shock factor.

In recent years, 3D and 4D NMR combined with uniform $^{13}C$ and $^{15}N$ enrichment has become an established procedure for determining the solution structures of proteins with spectra too complex for conventional homonuclear $^{1}H$ studies.[1-3] The resonance assignment part of this approach relies largely on determination of intraresidue $J$ connectivities via the large and well-resolved one-bond $^{1}H$–$^{13}C$, $^{1}H$–$^{15}N$, $^{13}C$–$^{13}C$, and $^{13}C$–$^{15}N$ $J$ couplings.[1-5] The $^{1}H$–$^{1}H$ NOE interactions, which are critical for structure determination, are typically severely overlapped in the conventional 2D NOESY spectrum but are dispersed in a 4D spectrum according to the frequencies of the $^{13}C$ or $^{15}N$ attached to each of the protons.

Although the approach outlined above has been applied successfully to a large range of proteins, there are inherent problems associated with the use of $^{13}C$ isotopic enrichment. Most importantly, the strong one-bond $^{1}H$–$^{13}C$ dipolar interaction is even larger than the dipolar interaction between geminal protons, causing a dramatic decrease in the $^{1}H$ $T_2$. Despite these short $T_2$ values, the high intrinsic sensitivity of many of the 3D $J$ correlation experiments allows their application to proteins larger than 30 kDa. However, the rapid transverse relaxation severely affects measurement of the inherently much weaker NOE interactions, particularly for $C^{\beta}$ methylene and other non-methyl-group side-chain protons in the rigid parts of the protein. Measurement of the structurally important $H^{\alpha}$–$H^{\beta}$ $J$ couplings is also adversely affected by the increase in $^{1}H$ line width caused by $^{13}C$, and very few accurate values have been reported for larger proteins.

Although the assignment problem has largely been solved by the uniform $^{13}C$ labeling approach, identification of the aromatic resonances of phenylalanine residues remains notoriously difficult. In addition to the $^{13}C$-induced $^{1}H$ line broadening, the aromatic Phe resonances also suffer from the non-first-order character of their $^{13}C$ spectrum, caused by the very poor $^{13}C$ chemical shift dispersion and large $^{13}C$–$^{13}C$ $J$ couplings (~55 Hz). Consequently, the assignment methods commonly used for other residues tend to give poor results for the phenylalanine ring system. For proteins with more than a few Phe residues, the resonance assignments therefore frequently do not extend much beyond the $C^{\delta}$–$H^{\delta}$ resonances. For example, in the study of the structure of a calmodulin/target peptide complex,[6] less than half of the aromatic protons and none of the $H^{\zeta}$ resonances of the nine calmodulin Phe residues had been assigned. $H^{\zeta}$ resonances are particularly critical to NMR structure determination, as they frequently are located in the hydrophobic core of the protein, where they exhibit disproportionately large numbers of long-range NOE constraints. Moreover, in contrast to the other Phe ring protons, the $H^{\zeta}$ position is independent of ring flip motions, and no pseudo-atom constraint-loosening approximations need to be made.

In practice, the small $^{13}C$ chemical shift dispersion of protonated aromatic ring carbons and the large width (110 Hz) of the $^{13}C$–$^{13}C$ triplet make the $^{13}C$ dispersion less useful than for other residues. Thus, the inherent $\sqrt{2}$ loss in sensitivity associated with the additional $^{13}C$ dimension is not offset by an increase in resolution. In fact, for a number of uniformly $^{13}C$-enriched proteins with long rotational correlation times (>10 ns), we have found that even the short $H^{\beta}$–$H^{\delta}$ NOE interactions may be difficult to observe in the 4D spectrum.

Here, we demonstrate the power of reverse labeling for alleviating the problems mentioned above. Incorporation of natural abundance residues into an otherwise uniformly $^{13}C$-enriched protein makes it possible to study the NOE interactions between the reverse labeled residues and the remainder of the protein with increased sensitivity. A novel experiment for quantitative measurement of $H^{\alpha}$–$H^{\beta}$ $J$ couplings in such residues is also demonstrated. The reverse labeling approach is demonstrated by incorporating natural abundance Phe into the otherwise uniformly $^{13}C/^{15}N$-enriched DNA-binding domain of *Drosophila* heat shock factor, dHSF(33–155), a 123-residue domain containing 10 Phe residues.

§ National Institute of Diabetes and Digestive and Kidney Diseases.
‡ National Cancer Institute.
† Present address: Bijvoet Center for Biomolecular Research, Padualaan 8, 3584CH Utrecht, The Netherlands.

⊕ Abstract published in *Advance ACS Abstracts*, September 1, 1994.

(1) Clore, G. M.; Gronenborn, A. M. *NMR of Proteins*; MacMillan: London, 1993.
(2) Wagner, G. *J. Biomol. NMR* **1993**, *3*, 375–385.
(3) Bax, A.; Grzesiek, S. *Acc. Chem. Res.* **1993**, *26*, 131–138.
(4) Palmer, A. G.; Fairbrother, W. J.; Cavanagh, J.; Wright, P. E.; Rance, M. *J. Biomol. NMR* **1992**, *2*, 103–108.
(5) Olejniczak, E. T.; Xu, R. X.; Fesik, S. W. *J. Biomol. NMR* **1992**, *2*, 655–659.
(6) Ikura, M.; Clore, G. M.; Gronenborn, A. M.; Zhu, G.; Klee, C. B.; Bax, A. *Science* **1992**, *256*, 632–638.

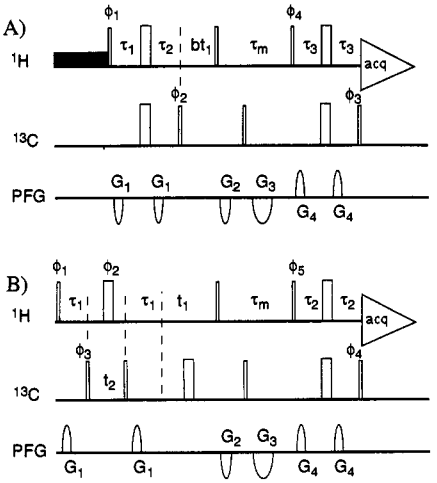


**Figure 1.** Pulse schemes of (A) the semi-constant-time doubly $^{12}C$-filtered 2D NOESY experiment and (B) the $^{13}C$-edited, $^{12}C$-filtered 3D NOESY experiment. Narrow pulses correspond to 90° flip angles, wide pulses to 180°. All pulses are applied along $x$, unless otherwise specified. All gradients are sine-bell shaped, 25 G/cm at center. The 180° $^{13}C$ pulses are of the composite type ($90_x220_y90_x$). Phase cycling for scheme A: $\phi_1 = x,-x$; $\phi_2 = 2(x), 2(-x)$, $\phi_3 = 8(x),8(-x)$; $\phi_4 = 2(x),2(y),2(-x),2(-y)$; acq = $x,-x,-y,y,-x,x,y,-y$; quadrature in $t_1$ using States-TPPI[36] on $\phi_1$. Delay durations for scheme A: $\tau_1 = 1/(4J_{CH}) - at_1$; $\tau_2 = 1/(4J_{CH}) + at_1$ with $a = [1/4J_{CH})]/t_{1max}$ and $b = 1-2a$; $\tau_3 = 1/(4J_{CH})$; $\tau_m$ = NOE mixing period; $G_{1,2,3,4}$ = 0.2, 1, 3.5, 0.6 ms. For scheme A, weak presaturation of the residual HDO resonance is used between scans. Phase cycling for scheme B: $\phi_1 = x,-x$; $\phi_2 = x$; $\phi_3 = 2(x),2(-x)$; $\phi_4 = 8(x), 8(-x)$; $\phi_5 = 4(x),4(y)$; acq = $x,2(-x),x,-y,2(y),-y$; quadrature in $t_1$ using States-TPPI on $\phi_1$ and $\phi_2$ simultaneously, quadrature in $t_2$ using States-TPPI on $\phi_3$. Delay durations for scheme B: $\tau_1 = 1/(2J_{CH})$; $\tau_2 = 1/(4J_{CH})$; $G_{1,2,3,4}$ = 0.2, 1, 3.5, 0.6 ms.

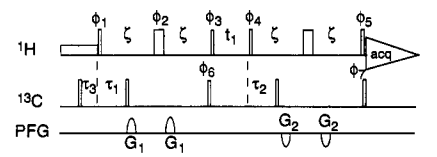


**Figure 2.** Pulse scheme of the $^{12}C$-filtered $H^{\alpha}$–$H^{\beta}$ quantitative $J$ correlation experiment. Narrow pulses have 90° flip angles, wide pulses 180°. All pulses are applied along $x$, unless otherwise specified. Gradients: sine-bell shaped, 25 G/cm at center; $G_{1,2}$ = 0.9, 0.5 ms. Phase cycling: $\phi_1 = x,-x$; $\phi_2 = x$; $\phi_3 = 4(y),4(-y)$; $\phi_4 = 2(y),2(-y)$; $\phi_5 = 8(y),8(-y)$; $\phi_6 = 4(x),4(-x)$; $\phi_7 = 16(x),16(-x)$; acq = $x,-x$; quadrature in $t_1$ using States-TPPI[36] on $\phi_1$, $\phi_2$, and $\phi_3$ simultaneously. Delays: $\tau_{1,2,3} = 3.6, 3.2, 20$ ms; $\zeta = 10.7$ ms. Weak presaturation of HDO is used during the delay between scans.

## Experimental Section

Sample Preparation. *Escherichia coli* (BL21/DE3) cells transformed with dHSF(33–155) plasmid were grown at 37 °C in M9 minimal media containing 0.3% [$^{13}C_6$]glucose/0.1% $^{15}NH_4Cl$ and 20 mg/L natural abundance L-Phe. The protein was purified as described previously,[7] and a 0.4 L culture yielded 5 mg of purified protein, sufficient for a 2 mM sample in a Shigemi microcell (Shigemi Inc., Allison Park, PA). The $^{12}C$ labeling of Phe residues was at least 90%, and no measurable increase in $^{12}C$ was observed for any of the other amino acids.

**NMR.** All NMR experiments were recorded in $D_2O$ solution, pH 6.3, 10 mM potassium phosphate, 50 mM KCl, at 27 °C, using a Bruker AMX600 spectrometer equipped with a Bruker triple-resonance probehead and a self-shielded $z$-gradient. Sine-bell shaped pulsed field gradients (25 G/cm at the center of the sine bell) were generated with an in-house developed pulse shaping unit and amplifier.

**$^{12}C$-Filtered HOHAHA.** The $^{12}C$-filtered HOHAHA spectrum of the aromatic region of dHSF(33-155) (Figure 3A) was recorded using a recently proposed $^{12}C$ filtering scheme that optimizes sensitivity by $^{12}C$ filtering during the $^1H$ isotropic mixing period,[8] avoiding the need for relaxation-sensitive filtering delays. Experimental parameters: $^1H$ carrier, 4.75 ppm; $^{13}C$ carrier, 63.9 ppm;[9] high-power $^1H$ radio frequency field strength, 26 kHz; high-power $^{13}C$ radio frequency field strength, 16.6 kHz); $^1H$ isotropic mixing and $^{13}C$ DIPSI irradiation, 9.1 kHz radio frequency field strength; matrix size, 300* × 1024* data matrix, where $n$* refers to $n$ complex data points; 32 transients acquired per complex $t_1$ increment; total measuring time, 4 h; digital filtering, 45°-shifted squared sine-bell filters in both $t_1$ and $t_2$; final digital resolution, 5.7 ($F_1$) and 3.4 Hz ($F_2$). The $^1H$ isotropic mixing period consisted of 11 DIPSI-2[10] cycles (3.1 ms each), of which the first and the last three were synchronized with DIPSI-2 $^{13}C$ irradiation.[8]

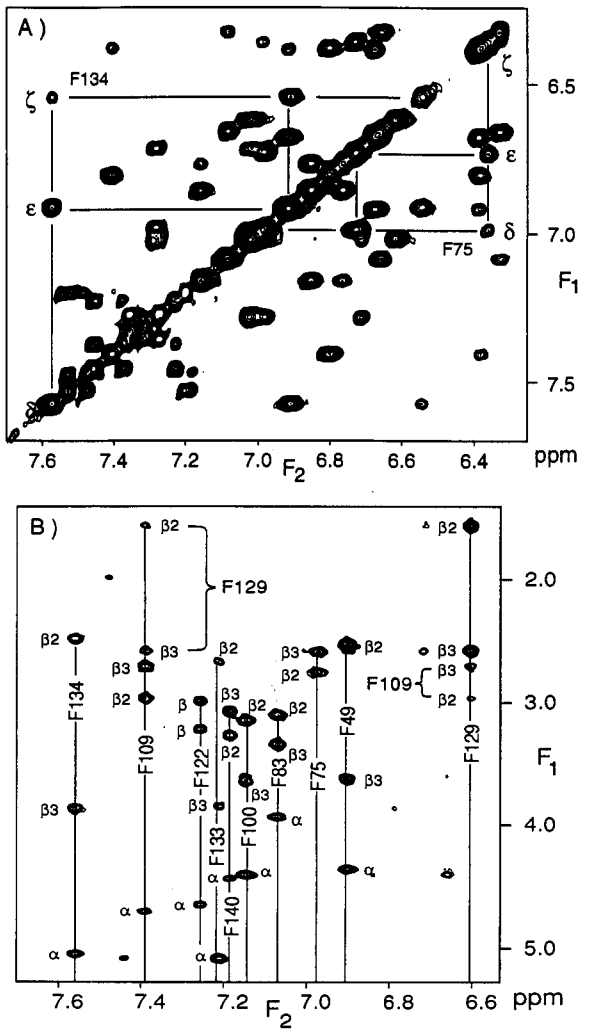


**Figure 3.** Small sections of (A) the $F_1/F_2$ $^{12}C$-filtered HOHAHA spectrum and (B) the $F_1/F_2$ $^{12}C$-filtered NOESY spectrum of {$^{12}C$-Phe,U-$^{13}C/^{15}N$} dHSF(33-155).

**$^{12}C$-Filtered NOESY.** The $^{12}C$-filtered spectrum of Figure 3B, which shows only interactions between $^{12}C$-attached protons, has been recorded with the pulse scheme of Figure 1A, which is a modified version of the previously proposed [$F1$-C,$F2$-$C$]NOESY filtering scheme.[11] Relaxation losses are minimized by using 2-fold shorter filtering delay durations ($\tau_1 + \tau_2 = 3.6$ ms; $2\tau_3 = 3.1$ ms) compared to the original scheme.[11] Moreover, the first evolution period is of the "mixed-constant-time" type,[12] where for longer $t_1$ durations an increasing fraction of the $^1H$ evolution takes place during the $\tau_1 + \tau_2$ filtering delay. This has the effect of reducing the decay rate of the magnetization in the $t_1$ dimension. Experimental

(7) Vuister, G. W.; Kim, S.-J.; Wu, C.; Bax, A. *Biochemistry* **1994**, *33*, 10–16.

(8) Bax, A.; Grzesiek, S.; Gronenborn, A. M.; Clore, G. M. *J. Magn. Reson., Ser. A* **1994**, *106*, 269–273.

(9) This $^{13}C$ carrier position was accidental. A better choice for filtering the aromatic region of the spectrum would have been to place the $^{13}C$ carrier in the aromatic region of the spectrum (~125 ppm) and to use shorter (two DIPSI-2 cycles, ~6.3 ms) durations for the synchronous $^{13}C$ irradiation.

(10) Shaka, A. J.; Lee, C. J.; Pines, A. *J. Magn. Reson.* **1988**, *77*, 274–293.

(11) Ikura, M.; Bax, A. *J. Am. Chem. Soc.* **1992**, *114*, 2433–2440.

(12) Grzesiek, S.; Bax, A. *J. Biomol. NMR* **1993**, *3*, 185–204.

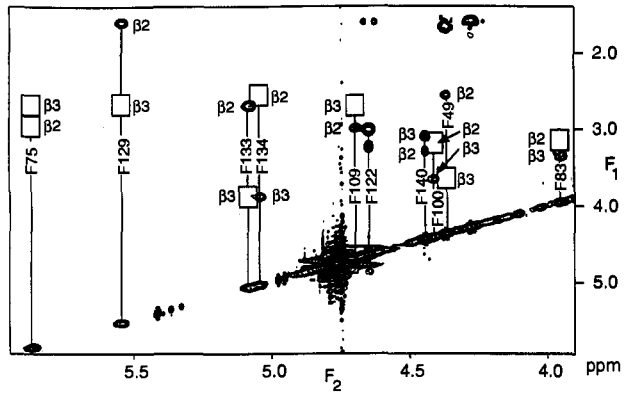


**Figure 4.** Quantitative $H^{\alpha}$–$H^{\beta}$ $J$ correlation spectrum of {$^{12}$C-Phe,U-$^{13}$C/$^{15}$N} dHSF(33-155), recorded with the pulse scheme of Figure 2. The spectrum displays $J$ connectivity for $^{12}$C-labeled residues (Phe) only, and the cross peak to diagonal peak intensity ratio is directly related to the magnitude of the $H^{\alpha}$–$H^{\beta}$ cross peak (see text).

parameters: $^{13}$C carrier, 64 ppm; $^{1}$H carrier, 4.75 ppm; $^{13}$C 90° pulse width, 11 μs; $^{1}$H 90° pulse width, 10 μs; NOE mixing period ($\tau_m$), 80 ms; 96 scans per complex $t_1$ increment; acquired data matrix size, 300* × 1024*; final digital resolution, 5.7 ($F_1$) and 3.4 Hz ($F_2$); digital filtering, 54°-shifted squared sine-bell ($t_1$) and 60°-shifted squared sine-bell ($t_2$) filters; total measuring time, 12 h.

**$^{13}$C-Separated, $^{12}$C-Filtered 3D NOESY.** The $^{13}$C-separated, $^{12}$C-filtered 3D NOESY experiment was recorded with the pulse scheme of Figure 1B. $^{13}$C frequency labeling in the $t_2$ dimension is obtained in the standard way by heteronuclear multiple quantum correlation (HMQC), using $J_{CH}$ de- and rephasing delays, $\tau_1$, of 3.5 ms. The scheme was optimized for suppressing detection of $^{13}$C-attached aromatic proton signals by using a $\tau_2$ value of 1.55 ms ($2\tau_2 = 1/(2J_{CHaromatic})$). Other experimental parameters: $^{13}$C carrier, 63.9 ppm; $^{1}$H carrier, 4.75 ppm; $^{13}$C 90° pulse width, 11 μs; $^{1}$H 90° pulse width, 9.6 μs; NOE mixing period ($\tau_m$), 80 ms; 16 scans per complex $t_1$ increment; acquired data matrix size, 96* × 32* × 512*; final digital resolution, 23 ($F_1$), 39 ($F_2$), and 7.1 Hz ($F_3$); digital filtering, 72°-shifted sine-bell ($t_1$), 72°-shifted squared-sine bell ($t_2$) 60°-shifted squared sine-bell ($t_3$); in the $t_1$ and $t_2$ dimensions, the end of the sine-bell window is truncated such that the last time domain data point is multiplied by 0.1; total measuring time, 38 h.

**Quantitative $H^{\alpha}$–$H^{\beta}$ $J$ Correlation.** The quantitative $H^{\alpha}$–$H^{\beta}$ $J$ correlation spectrum of Figure 4 has been recorded with the pulse scheme of Figure 2. A similar approach has been used previously for measurement of $^{13}$C–$^{13}$C, $^{13}$C–$^{1}$H, and $^{1}H^{N}$–$^{1}H^{\alpha}$ couplings in proteins.[13-15] A brief description of the pulse scheme of Figure 2 is given first. Very low power ($\gamma B_1 \sim 4$ Hz) saturation of the intense HDO resonance is used during the delay between scans. A 90° $^{13}$C pulse, followed by a $\tau_3$ delay of 20 ms, ensures that there is no net $^{13}$C magnetization when the first $^{1}$H 90°$_{\phi 1}$ is applied. Transverse magnetization generated for $^{13}$C-attached protons by the 90°$_{\phi 1}$ pulse is converted into multiple quantum coherence by the 90° $^{13}$C pulse, applied after a delay, $\tau_1$ ($\sim 1/(2J_{CH})$), and subsequently dephased by the two gradient pulses, $G_1$. The 90°$_{\phi 6}$ and subsequent $^{13}$C pulses serve to further suppress the signal from any $^{13}$C-attached protons. Although the use of four filtering steps to suppress $^{13}$C-attached proton signals may appear excessive, they do not lengthen the duration of the experiment, nor do they require any additional delays. Neglecting relaxation, the signal of an $H^{\alpha}$ proton, $A$, dephases during the period $2\zeta$ with respect to its $H^{\beta}$ protons, $M$ and $X$:

$$A_y \rightarrow \cos(2\pi J_{AM}\zeta)\cos(2\pi J_{AX}\zeta)A_y - 2\sin(2\pi J_{AM}\zeta)\cos(2\pi J_{AX}\zeta)A_xM_z - 2\cos(2\pi J_{AM}\zeta)\sin(2\pi J_{AX}\zeta)A_xX_z - 4\sin(2\pi J_{AM}\zeta)\sin(2\pi J_{AX}\zeta)A_yM_zX_z \quad (1)$$

The subsequent 90°$_{\phi 3}$ pulse ($\phi_3 = y$) converts this into the following:

$$\cos(2\pi J_{AM}\zeta)\cos(2\pi J_{AX}\zeta)A_y + 2\sin(2\pi J_{AM}\zeta)\cos(2\pi J_{AX}\zeta)A_zM_x + 2\cos(2\pi J_{AM}\zeta)\sin(2\pi J_{AX}\zeta)A_zX_x - 4\sin(2\pi J_{AM}\zeta)\sin(2\pi J_{AX}\zeta)A_yM_xX_x \quad (2)$$

During $t_1$, the first three terms oscillate with the chemical shift frequencies, $\delta_A$, $\delta_M$, and $\delta_X$ of spins $A$, $M$, and $X$. The last term represents three-spin coherence and is $t_1$-modulated by four frequencies, $\delta_A \pm \delta_M \pm \delta_X$. This three-spin coherence term may be safely ignored because under the experimental conditions used ($2\zeta \ll 1/J_{AM}, 1/J_{AX}$), it is very weak, and in practice, cross peaks resulting from it fall well below the signal-to-noise threshold. The effect of $J$ coupling during the relatively short $t_1$ evolution period ($t_{1max} \ll 1/J_{AM}, 1/J_{AX}$) may also be ignored as it only results in a small amount of line-broadening in the $t_1$ dimension. It can be shown that only those terms that are converted into $A$ spin transverse magnetization by the 90°$_{\phi 4}$ pulse, applied at the end of the $t_1$ evolution period, can give rise to observable in phase $A$ spin magnetization during $t_2$. Neglecting $J$ coupling evolution during the $t_1$ period, these terms are described by:

$$\cos(2\pi\delta_A t_1)\cos(2\pi J_{AM}\zeta)\cos(2\pi J_{AX}\zeta)A_y + 2\cos(2\pi\delta_M t_1)\sin(2\pi J_{AM}\zeta)\cos(2\pi J_{AX}\zeta)A_zM_x + 2\cos(2\pi\delta_X t_1)\cos(2\pi J_{AM}\zeta)\sin(2\pi J_{AX}\zeta)A_zX_x \quad (3)$$

The 90°$_{\phi 4}$ ($\phi_4 = y$) pulse converts these terms into the following:

$$\cos(2\pi\delta_A t_1)\cos(2\pi J_{AM}\zeta)\cos(2\pi J_{AX}\zeta)A_y - 2\cos(2\pi\delta_M t_1)\sin(2\pi J_{AM}\zeta)\cos(2\pi J_{AX}\zeta)A_xM_z - 2\cos(2\pi\delta_X t_1)\cos(2\pi J_{AM}\zeta)\sin(2\pi J_{AX}\zeta)A_xX_z \quad (4)$$

Subsequent refocusing during the final period, $2\zeta$, and a 90°$_{\phi 5}$ ($y$) purge yields the following terms for the in-phase $A$ spin magnetization:

$$\cos(2\pi\delta_A t_1)\cos^2(2\pi J_{AM}\zeta)\cos^2(2\pi J_{AX}\zeta)A_y - \cos(2\pi\delta_M t_1)\sin^2(2\pi J_{AM}\zeta)\cos^2(2\pi J_{AX}\zeta)A_y - \cos(2\pi\delta_X t_1)\cos^2(2\pi J_{AM}\zeta)\sin^2(2\pi J_{AX}\zeta)A_y \quad (5)$$

The first term gives rise to a diagonal $A$ spin resonance, whereas the second and third terms yield in-phase COSY cross peaks to spins $M$ and $X$, respectively. As is seen from these terms, the cross peak to a diagonal peak ratios for spins $M$ and $X$ depend directly on $\tan^2(2\pi J_{AM}\zeta)$ and $\tan^2(2\pi J_{AX}\zeta)$, respectively. As discussed previously for the measurement of heteronuclear $^{13}$C–$^{1}$H $J$ couplings, the strong geminal $M$–$X$ dipolar interaction can cause relatively rapid $M_z$–$X_z$ spin flips, which have the effect of increasing the weaker of the two $H^{\alpha}$–$H^{\beta}$ cross peaks and attenuating the more intense one.[14,15] Relaxation caused by dipolar interactions with other protons attenuates both cross peaks.[14,15] The intensity ratios of the cross and diagonal peaks are used directly to calculate the values of the $H^{\alpha}$–$H^{\beta}$ $J$ couplings, without accounting for these relaxation effects.

Experimental parameters: $^{13}$C carrier, 63.9 ppm; $^{1}$H carrier, 4.75 ppm; $^{13}$C 90° pulse width, 11 μs; $^{1}$H 90° pulse width, 9.6 μs; $\zeta$, 10.7 ms; $\tau_{1,2,3}$ = 3.6, 3.2, 20 ms; 128 scans per complex $t_1$ increment; acquired data matrix size, 128* × 2048*; $t_{1max}$, 21.7 ms; $t_{2max}$, 141 ms; final digital resolution, 23 ($F_1$) and 3.5 Hz ($F_2$); digital filtering, 72°-shifted sine-bell ($t_1$) and 60°-shifted squared sine-bell ($t_2$). Total measuring time, 6 h.

**4D $^{13}$C/$^{13}$C-Separated 4D NOESY.** The 4D NOESY spectrum was recorded on a 2 mM sample of uniformly $^{13}$C/$^{15}$N-labeled protein, at the same pH, temperature and ionic strength as the experiments described above. The HMQC-NOE-HSQC scheme of reference 24 was used, using 16 scans for each set of complex ($t_1$, $t_2$, $t_3$) values. Other experimental parameters; matrix size, 16* × 64* × 16* × 384*; spectral widths, 20.7 ($F_1$), 9.8 ($F_2$), 20.7 ($F_3$), and 12.1 ppm ($F_4$); final digital resolution, 97 Hz ($F_1$ and $F_3$), 23 ($F_2$), and 14 Hz ($F_4$); digital filtering, shifted (81° for $t_1$ and $t_3$, 72° for $t_2$ and $t_4$), squared sine-bell filtering, with the sine-bell truncated at the end so that the last data point is multiplied by 0.1. Total measuring time, 95 h.

## Results and Discussion

Resonance assignments in the 123-residue DNA-binding domain of *Drosophila* heat shock factor (dHSF(33-155)), used

(13) Bax, A.; Max, D.; Zax, D. *J. Am. Chem. Soc.* **1992**, *114*, 6924–6925.

(14) Vuister, G. W.; Yamazaki, T.; Torchia, D. A.; Bax, A. *J. Biomol. NMR* **1993**, *3*, 297–306.

(15) Vuister, G. W.; Bax, A. *J. Am. Chem. Soc.* **1993**, *115*, 7772–7777.

**Table 1.** $\chi_1$-Related $J$ Couplings and Stereospecific $H^\beta$ Assignments for the Phenylalanine Residues in the DNA-Binding Domain of *Drosophila* HSF

| residue | $J_{N-H\beta 2}$ | $J_{N-H\beta 3}$ | $J_{CO-H\beta 2}$ | $J_{CO-H\beta 3}$ | $J_{H\alpha-H\beta 2}$ | $J_{H\alpha-H\beta 3}$ | $H^{\beta 2}$ | $H^{\beta 3}$ | $\chi_1$ |
|---|---|---|---|---|---|---|---|---|---|
| $Phe^{49}$ | *a* | *a* | *a* | *a* | 2.9 | 4.6 | 2.57[f] | 3.65 | 60° |
| $Phe^{75}$ | large | n.o.[b] | n.o. | n.o. | 3.9 | 3.5 | 2.80 | 2.63 | 60° |
| $Phe^{83}$ | n.o. | n.o. | large | n.o. | small[d] | large[d] | 3.13 | 3.37 | 180° [e] |
| $Phe^{100}$ | n.o. | n.o. | large | n.o. | small[d] | large[d] | 3.14 | 3.63 | 180° |
| $Phe^{109}$ | small | large | n.o. | n.o. | large[d] | small[d] | 2.93 | 2.69 | –60° |
| $Phe^{122}$ | large | medium | small | small | 7.1[c] | 8.9[c] | 3.22[c] | 3.00[c] | averaging |
| $Phe^{129}$ | small | large | n.o. | n.o. | 11.2 | <5.9 | 1.61 | 2.61 | –60° |
| $Phe^{133}$ | small | large | n.o. | n.o. | 11.5 | <6.0[g] | 2.69 | 3.87 | –60° |
| $Phe^{134}$ | n.o. | n.o. | n.o. | n.o. | 9.5 | <5 | 2.52 | 3.88 | 180° [e] |
| $Phe^{140}$ | large | small | n.o. | medium | 4.1 | 5.3 | 3.28 | 3.09 | 60° [e] |

[a] Not observed due to short $T_2$ of $Phe^{49}$ and $Leu^{50}$ amide signals. [b] Not observed, *i.e.*, below the signal-to-noise threshold. [c] Not stereospecifically assigned. [d] Not measured quantitatively due to overlap of diagonal signal. [e] Based, in part, also on inspection of a set of 20 calculated structures. [f] $^1H$ chemical shifts are in ppm relative to TSP. [g] The intensity of the weak $Phe^{133}$ $H^\alpha$–$H^{\beta 3}$ correlation cannot be measured accurately because of the presence of intense nearby $Phe^{134}$ $H^\alpha$–$H^{\beta 3}$ cross peak.

in the present study, were found to be identical to those reported previously for a construct that included 11 additional unstructured carboxyterminal residues.[7] At the concentration and temperature used, $^{15}N$ relaxation indicates a rotational correlation time of ~10 ns, which is somewhat longer than expected for a monomeric 14 kDa protein. Sedimentation experiments indicate, however, that the domain is monomeric at lower concentration.[16] A detailed analysis of the NOE data also shows no evidence for intermolecular NOEs.

Due to the relatively large line widths in uniformly $^{13}C$-enriched dHSF(33-155), a detailed analysis of 3D HCCH[17,18] and (H$\beta$)C$\beta$-(C$\gamma$)C$\delta$H$\delta$[19] $J$ correlation (supplementary material) and 4D NOESY data previously had yielded tentative assignments for only half of the aromatic $^1H$ resonances of the 10 Phe residues. The difficulties in obtaining reliable assignments and NOE constraints for the Phe rings, most of which are located in the hydrophobic core, limited the precision at which the protein structure could be calculated. Therefore, incorporation of natural abundance Phe in dHSF(33-155) provided a good test case for the reverse labeling procedure.

Figure 3A shows the aromatic region of the $^{12}C$-filtered HOHAHA spectrum, displaying all 10 phenylalanine $^1H$–$^1H$ connectivity patterns, free of overlap from the Trp and Tyr residues. As can be seen, the unusually large number of aromatic residues in the hydrophobic core of the protein does not lead to an untractible overlap of resonances, presumably because the high density of aromatics also causes significant ring current effects throughout this region. Sequence-specific assignment of the ring systems can be obtained by correlating the $H^\delta$ protons with the previously assigned $H^\alpha$ and $H^\beta$ resonances in a $^{12}C$-filtered 2D NOESY spectrum (Figure 3B). Interactions between $Phe^{109}$ and $Phe^{129}$, anticipated on the basis of a low-resolution structure, are also seen in this figure.

The absence of $^{13}C$ results in relatively narrow Phe–$H^\alpha$ resonance ($T_2$ ~ 20–25 ms). Although the line widths of the $^{12}C$-attached protons, particularly those of the geminal $H^\beta$, remain too large for a detailed ECOSY[20] $J_{H\alpha H\beta}$ measurement, a good estimate for $J_{H\alpha H\beta}$ can be obtained from the quantitative $J$ correlation experiment of Figure 2. The spectrum of Figure 4 shows $H^\alpha$–$H^\beta$ $J$ cross peaks for 9 out of the 10 Phe residues; cross peaks to both $H^\beta$ protons of $Phe^{75}$ can be observed at contour levels lower than shown in the figure. As described in the Experimental Section, neglecting relaxation effects,[21] the cross peak to diagonal peak intensity ratio is $\tan^2(2\pi J\zeta)$, and $J$ couplings derived from this relationship are reported in Table 1. For $Phe^{129}$ and $Phe^{134}$, correlations to only one of the $H^\beta$ protons is observed, even when the spectrum is plotted at the noise level. The spectrum therefore provides an upper limit for $\tan^2(2\pi J\zeta)$ and thereby an upper limit for the corresponding $J_{H\alpha H\beta}$ couplings. Together with results from qualitative $^3J_{NH\beta}$ and $^3J_{COH\beta}$ measurements,[22,23] the $J_{H\alpha H\beta}$ data extracted from Figure 4 indicate $\chi_1$ averaging for $Phe^{122}$ ($J_{\alpha\beta}$ = 7.1/8.9 Hz), located in an extended flexible loop. All other Phe $\chi_1$ angles are well defined by the $J$ coupling data, permitting stereospecific assignments of the prochiral $H^\beta$ protons (Table 1).

Figure 5 illustrates the improvement in sensitivity that can be obtained by reverse labeling over conventional 4D $^{13}C/^{13}C$ separated NOESY. Figure 5A shows a small region of a $F_1/F_2$ cross section through the $^{13}C$-edited, $^{12}C$-filtered 3D NOESY spectrum of the reverse labeled sample. This cross section, taken at the $F_3$ frequency of $Phe^{83}$–$H^\zeta$, reveals NOE interactions between $Phe^{83}$–$H^\zeta$ and the methyl protons of seven different residues. Due to the lower $^1H$ resolution in the corresponding region of the 4-day 4D $^{13}C/^{13}C$-separated NOESY spectrum (Figure 5B), several of the NOE interactions to $Phe^{83}$–$H^\zeta$ (assigned using $^{12}C$ reverse labeling) cannot be distinguished from $Phe^{75}$–$H^\zeta$, and clearly the signal-to-noise ratio in the 4D suffers from the additional relaxation losses incurred during the final HSQC part of the 4D pulse scheme[24] and from the additional $^{13}C$-induced line broadening of the detected $H^\zeta$ resonance.

In recent years, the uniform $^{13}C$ labeling strategy has proven to be extremely powerful for structure determination of proteins up to about 20 kDa, and resonance assignments for proteins significantly larger than 20 kDa have also been reported.[25-27] In principle, the uniform labeling approach also is capable of yielding 3D structures for these larger proteins. However, due to the rapid drop in the signal-to-noise ratio of the NOESY spectrum with increasing correlation time, the average number of NOEs per residue, and thereby the resolution of the protein structure,[28] rapidly decreases for larger proteins.

Reverse labeling of otherwise uniformly $^{13}C/^{15}N$-enriched proteins presents an attractive approach for increasing the spectral information obtainable, particularly for spin systems such as

(16) Kim, S.-J.; Tsukiyama, T.; Lewis, M. S.; Wu, C. *Protein Sci.*, in press.
(17) Kay, L. E.; Ikura, M.; Bax, A. *J. Am. Chem. Soc.* **1990**, *112*, 888–889.
(18) Bax, A.; Clore, G. M.; Gronenborn, A. M. *J. Magn. Reson.* **1990**, *88*, 425–431.
(19) Yamazaki, T.; Forman-Kay, J. D.; Kay, L. E. *J. Am. Chem. Soc.* **1993**, *115*, 11054–11055.
(20) Griesinger, C.; Sørensen, O. W.; Ernst, R. R. *J. Chem. Phys.* **1986**, *85*, 6837–6852.
(21) Harbison, G. *J. Am. Chem. Soc.* **1993**, *115*, 3026–3027.
(22) Archer, S. J.; Ikura, M.; Torchia, D. A.; Bax, A. *J. Magn. Reson.* **1991**, *95*, 636–641.
(23) Grzesiek, S.; Ikura, M.; Clore, G. M.; Gronenborn, A. M.; Bax, A. *J. Magn. Reson.* **1992**, *96*, 215–221.
(24) Vuister, G. W.; Clore, G. M.; Gronenborn, A. M.; Powers, R.; Garrett, D. S.; Tschudin, R.; Bax, A. *J. Magn. Reson., Ser. B* **1993**, *101*, 210–213.
(25) Fogh, R. H.; Schipper, D.; Boelens, R.; Kaptein, R. *J. Biomol. NMR* **1994**, *4*, 123–128.
(26) Remerowski, M. L.; Domeke, T.; Groenewegen, A.; Pepermans, H. A. M.; Hilbers, C. W.; van de Ven, F. J. M. *J. Biomol. NMR* **1994**, *4*, 257–278.
(27) Grzesiek, S.; Döbeli, H.; Gentz, R.; Garotta, G.; Labhardt, A. M.; Bax, A. *Biochemistry* **1992**, *31*, 8180–8190.
(28) Clore, G. M.; Robien, M. A.; Gronenbron, A. M. *J. Mol. Biol.* **1993**, *231*, 82–102.

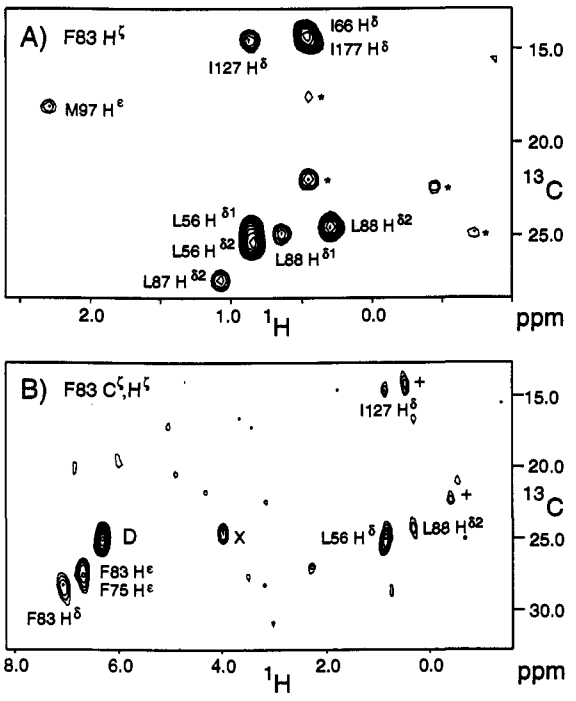


**Figure 5.** Comparison of NOE interactions observed to Phe[83]-H$^{\zeta}$ in (A) the $^{13}$C-edited, $^{12}$C-filtered 3D NOESY spectrum of reverse labeled dHSF-(33-155) and (B) the 4D $^{13}$C,$^{13}$C-edited NOESY spectrum. Spectrum A does not contain any intraresidue interactions, and only the cross peak-containing region of the spectrum is shown. The cross section through the 4D spectrum shows both intra- and interresidue NOE interactions. The intense resonance labeled "D" corresponds to the "diagonal" Phe[83]-C$^{\zeta}$,H$^{\zeta}$ resonance. Peaks marked with an asterisk in spectrum A are tails of cross peaks to Phe[75]H$^{\zeta}$; peaks marked + in spectrum B are cross peaks to Phe[83]-H$^{\zeta}$ but could not be assigned as such from the 4D spectrum because their $F_3/F_4$ coordinates are actually slightly closer to those of the nearby resonating Phe[75]-H$^{\zeta}$. The resonance marked × is a composite pulse modulation sideband. The $^{13}$C spectral width in the $F_1$ dimension is 20.7 ppm, and the aromatic resonances in panel B have been folded five times in this dimension, i.e., their $^{13}$C frequency is 103.5 ppm higher than that measured from the $F_1$ coordinate.

phenylalanine, where the aromatic $^{13}$C dispersion is poor. It has become well established that isotope filtering is ideally suited for studying the interaction between labeled and unlabeled residues.[29] The modified filtering procedures described above optimize sensitivity by reducing the duration of filtering delays. The longer $^1$H-{$^{12}$C} $T_2$ values, relative to $^1$H-{$^{13}$C}, improve both resolution and sensitivity compared to the study of uniformly $^{13}$C labeled material. This permits experiments such as the above-described quantitative H$^{\alpha}$-H$^{\beta}$ $J$ correlation, free from overlap of the multitude of $^{13}$C labeled residues, and it allows for additional experiments to complete the assignment process. For dHSF-(33-155), reverse labeling yielded 109 additional long-range NOEs, an increase of 39%. This resulted in a very significant reduction in the spread of the NMR structures, from 1.4 to 0.87 Å for the backbone plus ordered side-chain atoms.[30]

The labeling strategy described above is conceptually not new. In fact, before nonauxotrophic bacterial strains suitable for protein expression became widely available, essential amino acids at natural abundance were routinely added to the M9 growth medium. However, the lack of labeling of the essential amino acids was considered more of a nuisance than a useful avenue for extracting additional information. The reverse labeling strategy therefore is more comparable with the protein deuteration strategy, where a number of different protonated amino acids are incorporated into an otherwise perdeuterated protein. This deuteration strategy has proven to be quite powerful, both for obtaining resonance assignments and for obtaining detailed structural information.[31-35] In comparison to the $^{12}$C reverse labeling procedure described above, reverse deuteration requires a considerably larger number of selectively labeled samples to probe all possible interactions between protonated amino acids, and high levels of deuteration for metabolically active amino acids can be difficult to achieve.[33]

The $^{12}$C reverse labeling approach described in this paper requires a substantial amount of additional work in protein preparation and NMR data collection over the uniform $^{13}$C labeling approach. However, it is becoming increasingly clear that, for larger proteins, without this type of additional work it may frequently be impossible to gather a sufficiently high number of structural constraints for defining the solution structure with high accuracy.

**Acknowledgment.** We thank John Marquardt and Frank Delaglio for assistance. This work was supported by the AIDS Targeted Anti-Viral Program of the Office of the Director of the National Institutes of Health.

**Supplementary Material Available:** Figure showing strips through the 3D (H$\beta$)CB(C$\gamma$)C$\delta$H$\delta$ spectrum (1 page). This material is contained in many libraries on microfiche, immediately follows this article in the microfilm version of the journal, and can be ordered from the ACS; see any current masthead page for ordering information.

(29) Otting, G.; Wüthrich, K. *Q. Rev. Biophys.* **1990**, *23*, 39-96.

(30) Vuister, G. W.; Kim, S.-J.; Orosz, A.; Marquardt, J.; Wu, C.; Bax, A. *Nature. Struct. Biol.*, in press.

(31) Markley, J. L.; Putter, I.; Jardetzky, O. *Science* **1968**, *161*, 1249-1251.

(32) Brodin, P.; Drakenberg, T.; Thulin, E.; Forsen, S.; Grundstöm, T. *Prot. Eng.* **1989**, *2*, 353-358.

(33) Torchia, D. A.; Sparks, S. W.; Young, P. E.; Bax, A. *J. Am. Chem. Soc.* **1989**, *111*, 8315-8317.

(34) Arrowsmith, C. H.; Pachter, R.; Altman, R. B.; Iyer, S. B.; Jardetzky, O. *Biochemistry* **1990**, *29*, 6332-6341.

(35) Reisman, J. M.; Hsu, V. L.; Jariel-Encontre, I.; Lecou, C.; Sayre, M. H.; Kearns, D. R.; Parello, J. *Eur. J. Biochem.* **1993**, *213*, 865-873.

(36) Marion, D.; Ikura, M.; Tschudin, R.; Bax, A. *J. Magn. Reson.* **1989**, *85*, 393-399.