

Supporting Information

Monitoring ^{15}N Chemical Shifts During Protein Folding by Pressure-Jump NMR

Cyril Charlier, Joseph M. Courtney, T. Reid Alderson, Philip Anfinrud, and Ad Bax*

Laboratory of Chemical Physics, NIDDK, National Institutes of Health, Bethesda, MD, 20892-0520, USA

3-State Fitting. To extend the two-state model and include an on-pathway intermediate, we calculate the concentrations of the folding species by integration of the first order rate law

$$\frac{d}{dt} \begin{bmatrix} [\text{U}]_t \\ [\text{I}]_t \\ [\text{F}]_t \end{bmatrix} = \begin{bmatrix} -k_{u \rightarrow i} - k_{u \rightarrow f} & 0 & 0 \\ k_{u \rightarrow i} & -k_{i \rightarrow f} & 0 \\ k_{u \rightarrow f} & k_{i \rightarrow f} & 0 \end{bmatrix} \begin{bmatrix} [\text{U}]_t \\ [\text{I}]_t \\ [\text{F}]_t \end{bmatrix} \quad (\text{S1})$$

where $[\text{U}]_t$, $[\text{I}]_t$, and $[\text{F}]_t$ are the concentrations of the unfolded, intermediate, and folded species, respectively, at time t and $k_{u \rightarrow i}$, $k_{u \rightarrow f}$, and $k_{i \rightarrow f}$ are the rate constants for the transition of unfolded to intermediate, unfolded to folded, and intermediate to folded, respectively. Then, the amplitudes of the observed, folded protein magnetization for the two experiments are given by

$$M_x(\tau) \approx C \{([\text{U}]_\tau - [\text{U}]_T) \cos(\omega_U \tau) + ([\text{I}]_\tau - [\text{I}]_T) \cos(\omega_I \tau) + [\text{F}]_\tau \cos(\omega_F \tau)\} \quad (\text{S2a})$$

and

$$M_y(\tau) \approx C \{([\text{U}]_\tau - [\text{U}]_T) \sin(\omega_U \tau) + ([\text{I}]_\tau - [\text{I}]_T) \sin(\omega_I \tau) + [\text{F}]_\tau \sin(\omega_F \tau)\} \quad (\text{S2b})$$

where C is an instrumental constant. The multiplicative factor $([\text{X}]_\tau - [\text{X}]_T)$ accounts for the folding of the remaining unfolded and intermediate species after the chemical shift evolution period. As the chemical shift of the intermediate species is unknown, we fit it by non-linear least squares optimization of $M_x(\tau)$ and $M_y(\tau)$ to the observed peak intensities. The estimated uncertainty in ω_I is then

$$\sigma_{\omega_I} = \sqrt{\left(-\frac{d^2}{d\omega_I^2} [\ln L]\right)^{-1}} \quad (\text{S3})$$

where σ_{ω_I} is the estimated uncertainty in ω_I and $\ln L$ is the log likelihood of the fit model given the observed peak intensities given by,

$$\ln L = -\frac{n}{2} \ln(\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^n \left[(S_{x,i} - M_x(\tau_i))^2 + (S_{y,i} - M_y(\tau_i))^2 \right] \quad (\text{S4})$$

where n is the number of time points, σ , is the standard deviation of the noise and $S_{x,i}$ and $S_{y,i}$ are the measured signal intensities.

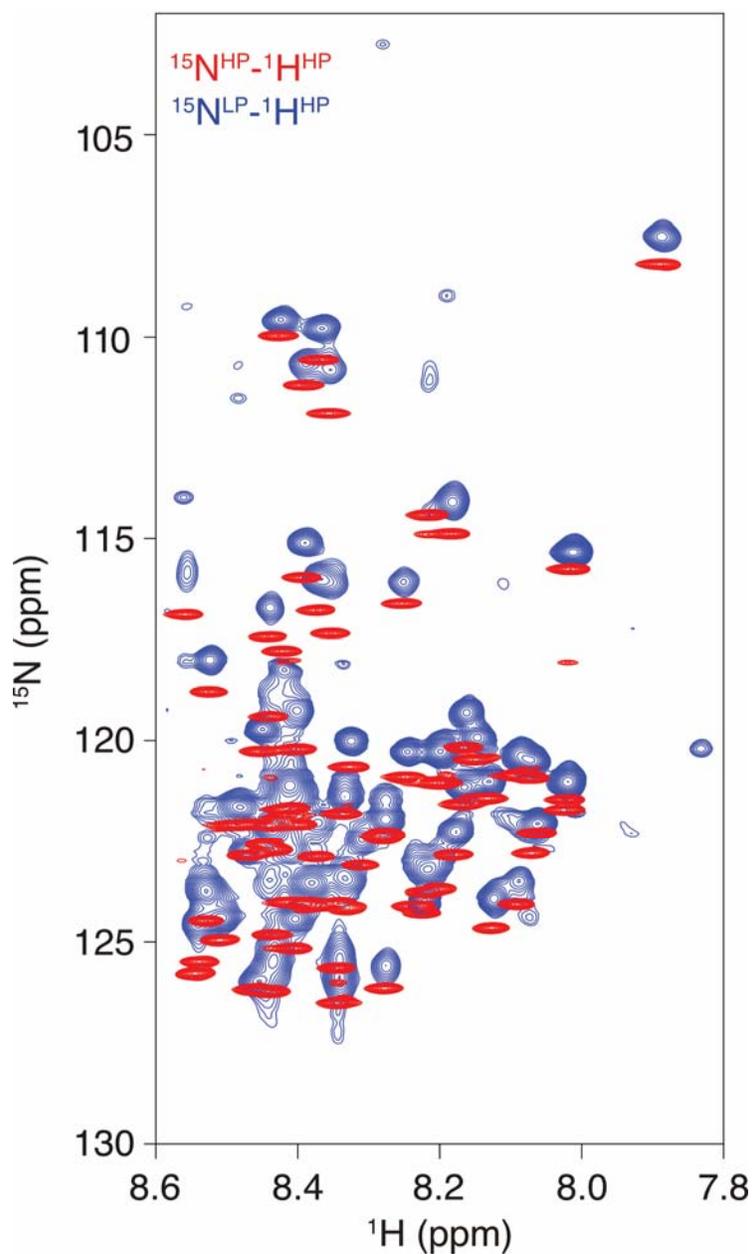


Figure S1. Superposition of the static pressure 600-MHz 2D HSQC measured at 2.5 kbar (red) and a pressure-jump 2D HSQC spectrum acquired with ^{15}N evolution at 1 bar and ^1H detection at 2.5 kbar (blue). Data were collected at 25 °C, 300 μM VA2-ubiquitin, 25 mM sodium phosphate buffer, pH 6.4 (at 1 bar). Strong ^{15}N line broadening during evolution at 1 bar (blue) compared to evolution at 2.5 kbar (red) is attributed to an exchange process associated with aborted folding attempts,¹ and is stronger at 25 °C than at 15 °C (for comparison, see Figure S6 of Charlier et al.¹)

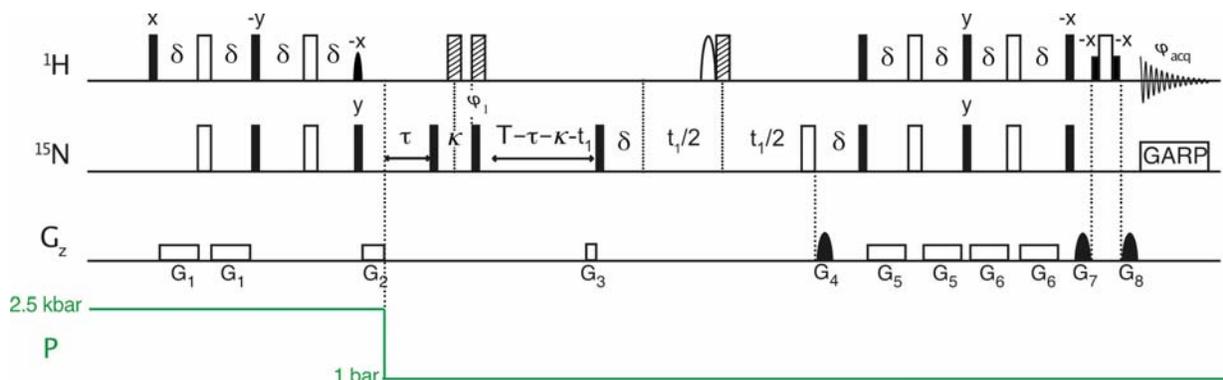


Figure S2. Pulse sequence for stroboscopic measurement of residue-specific, ensemble-averaged ^{15}N chemical shifts during protein folding. The pressure drop from 2.5 kbar to 1 bar immediately follows the first refocused INEPT magnetization transfer and initiates the protein folding process. This corresponds to the end of the preparation period (see Figure 1, main text), with a total duration of *ca* 10s, which serves as an equilibration period sufficiently long for the protein to fully unfold. At time τ after the pressure drop, the ensemble-averaged ^{15}N frequency is encoded by the stroboscopic element, $90_x\text{-}\kappa\text{-}90_{\phi_1}$, and stored as z magnetization until it is measured after protein folding has completed by a regular, gradient-enhanced HSQC element,² starting after gradient G_3 . For a given value of κ , the pulse scheme, is repeated for 16 values of τ and 100 complex t_1 increments, using a 2-step phase cycle for a total of 12,800 pressure cycles. Filled and open symbols on the ^1H and ^{15}N radiofrequency channels represent 90° and 180° pulses, respectively. Shaped pulses are selectively applied to the water resonance, as are the weak rectangular 90° pulses that are part of the standard WATERGATE element.³ Unless indicated, pulse phases are x . Hatched pulses are of the $90_x\text{-}210_y\text{-}90_x$ composite type.⁴ Composite-pulse decoupling during acquisition was performed on the ^{15}N channel with a GARP-scheme. The INEPT delays, δ , are 2.56 ms each. The total time from the pressure drop to the end of the ^{15}N evolution is kept constant at 300 ms. This is achieved by decrementing the delay between the $90^\circ\text{-}\kappa\text{-}90^\circ$ stroboscopic pulse pair and t_1 evolution, keeping the timing of the start of ^1H acquisition relative to the time point of the pressure drop constant. This avoids spectral distortion that otherwise is associated with the t_1 -dependent change in temperature caused by the adiabatic sample decompression following the pressure drop. Analogous effects on the ^{15}N evolution are much less severe, and therefore no such compensation for ^{15}N is needed. Note that the decreased total refolding time, prior to ^{15}N t_1 evolution, simply results in a slight ^{15}N line broadening. For each τ duration, two separate spectra are collected: The cos-modulated spectrum where $\phi_1 = \{x, -x\}$, $\phi_{\text{acq}} = \{x, -x\}$; and the sin-modulated spectrum where $\phi_1 = \{y, -y\}$, $\phi_{\text{acq}} = \{x, -x\}$. Pulsed field gradients are either sine-bell shaped ($G_{4,7,8}$) or weak, rectangular ($G_{1,2,3,5,6}$). Gradient durations and peak amplitudes are as follows: G_1 , 2.6 ms, 0.95 G/cm; G_2 , 2ms, 8.25 G/cm; G_3 , 1ms, 3.3 G/cm; G_4 , 2ms, 24.68 G/cm; G_5 , 2.6ms, 0.38 G/cm; G_6 , 2.6ms, 0.285 G/cm; G_7 , 0.5 ms, 15 G/cm; G_8 , 0.5 ms, 25 G/cm. Standard inversion of the sign of gradient G_4 together with inversion of the sign of the 90_y pulse after G_5 are used for generating the reverse precession signal in the gradient-enhanced detection scheme.²

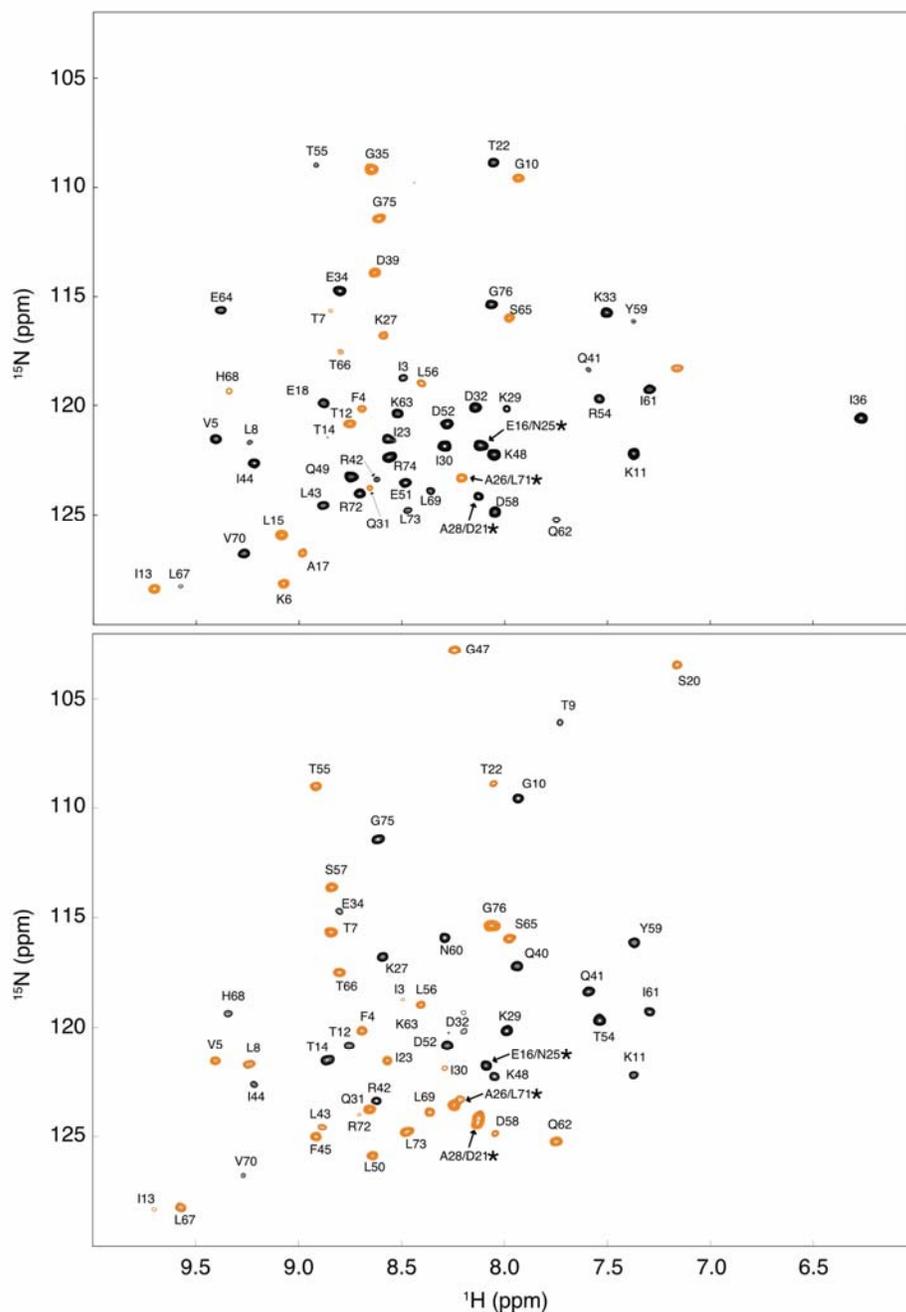
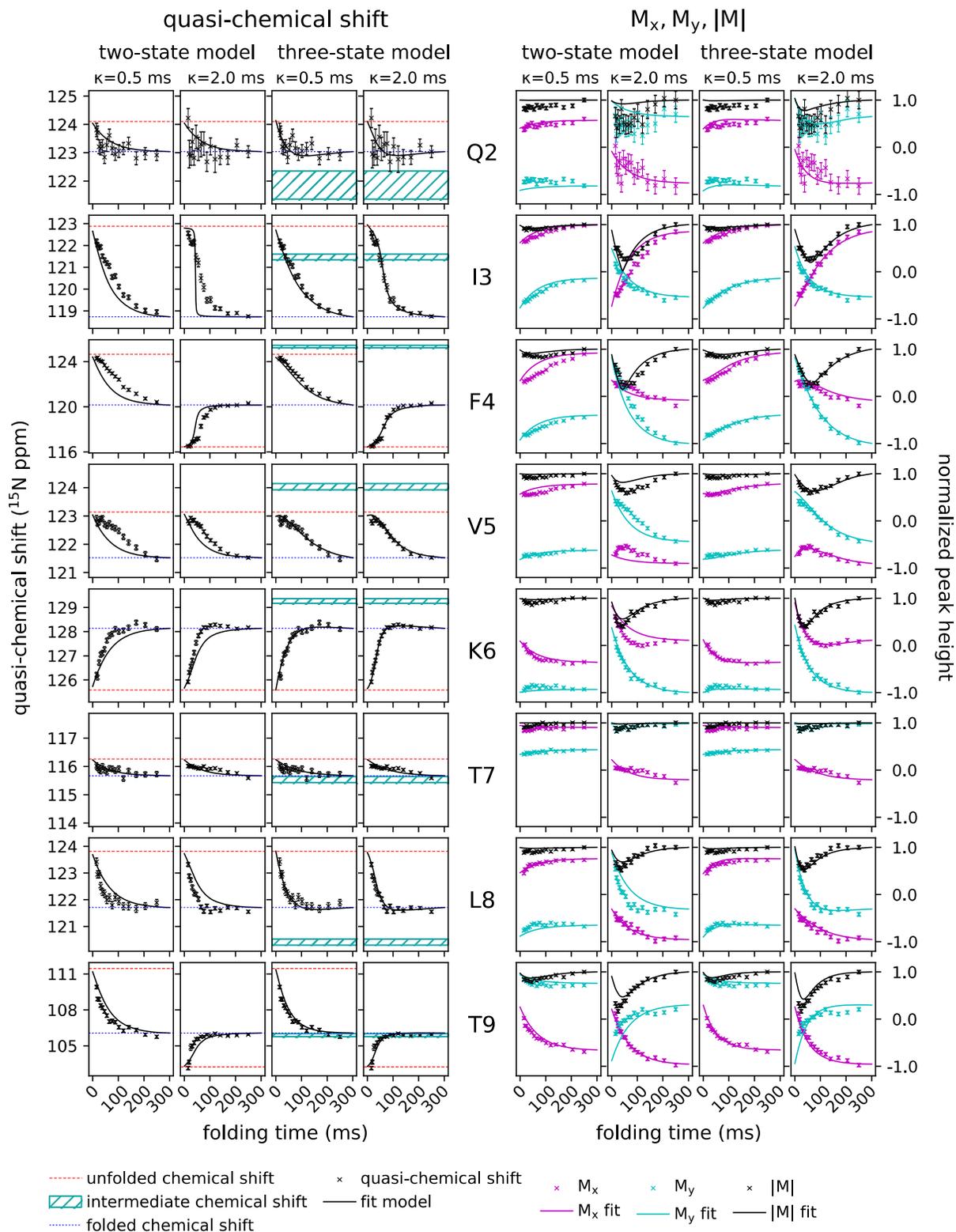
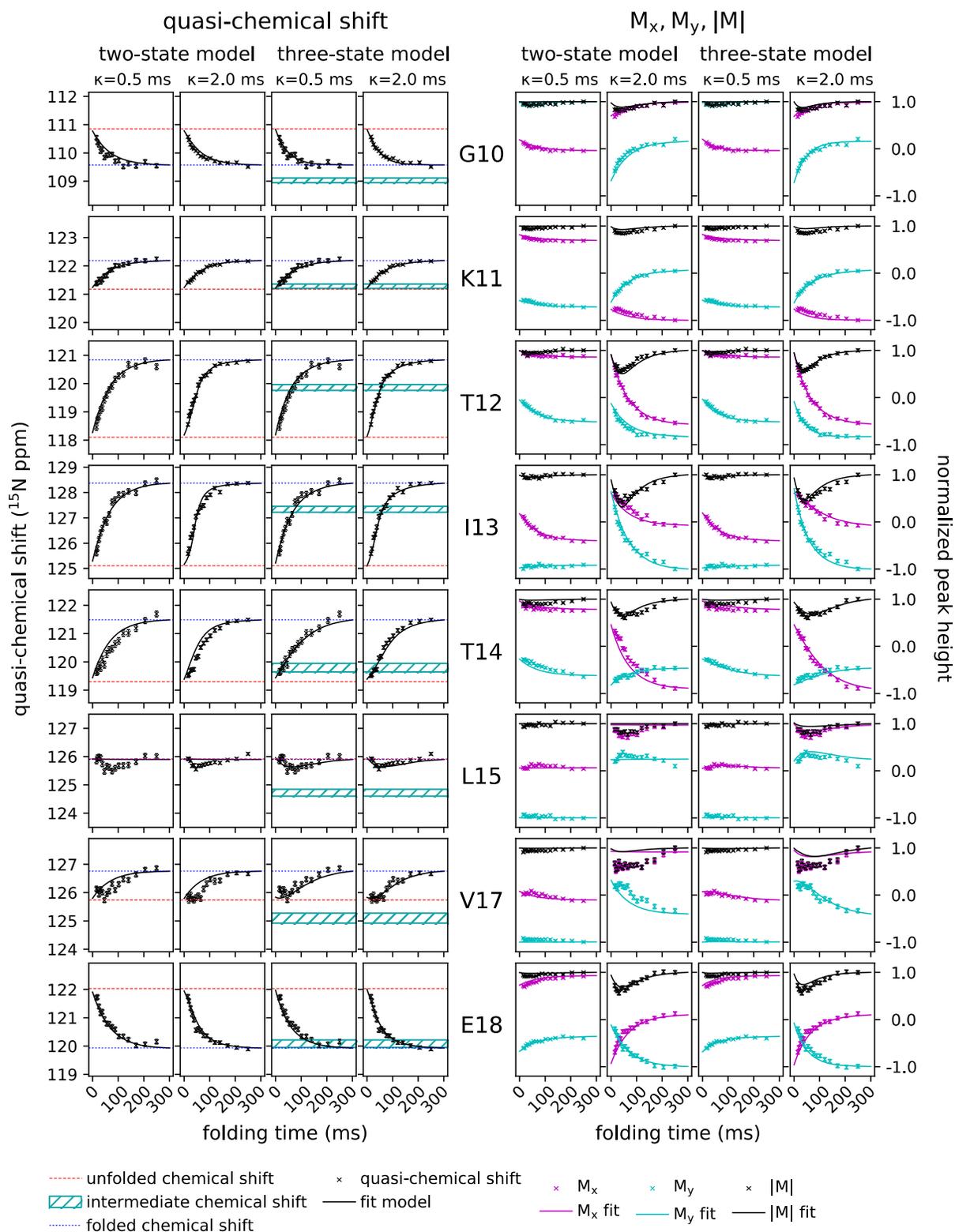
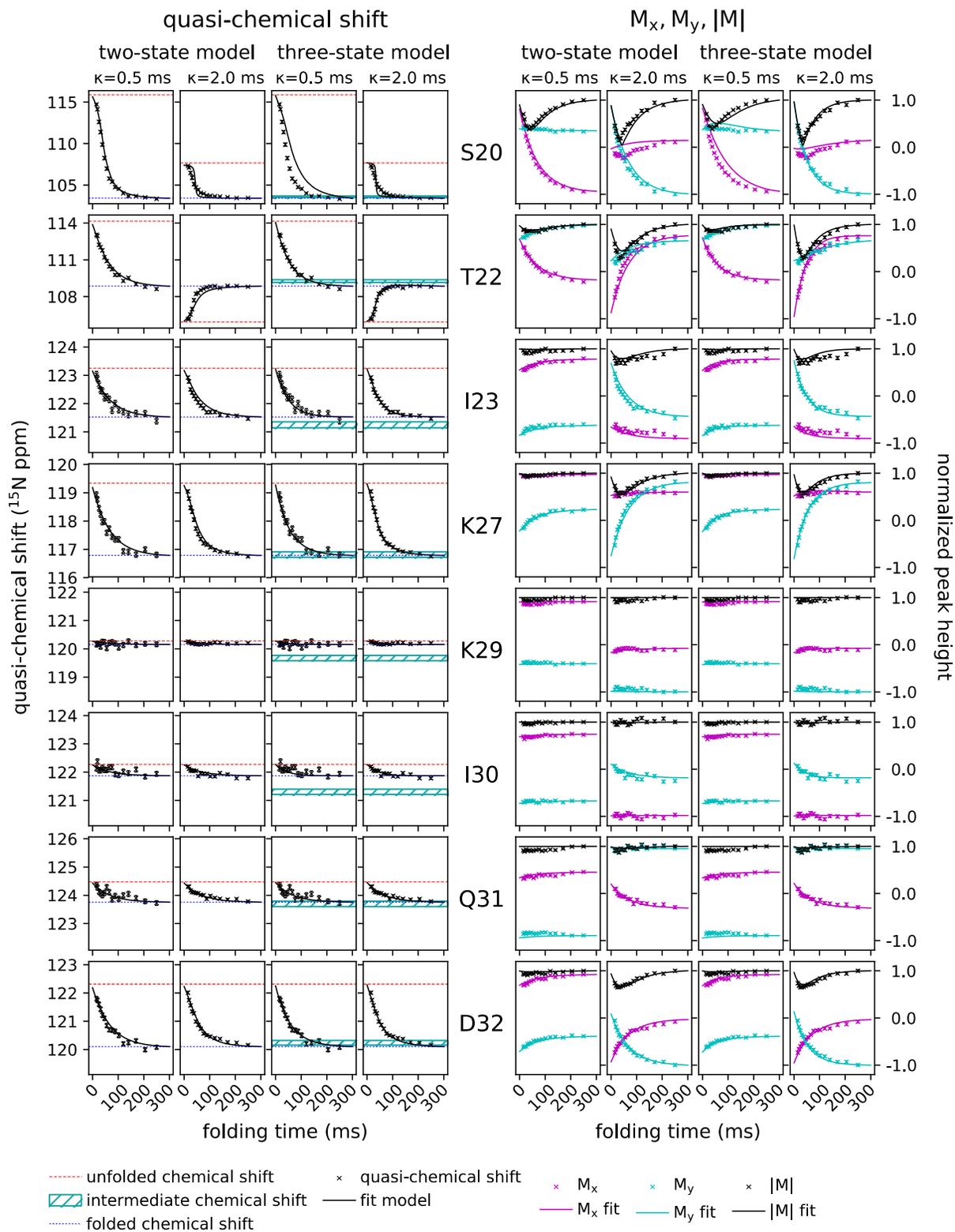
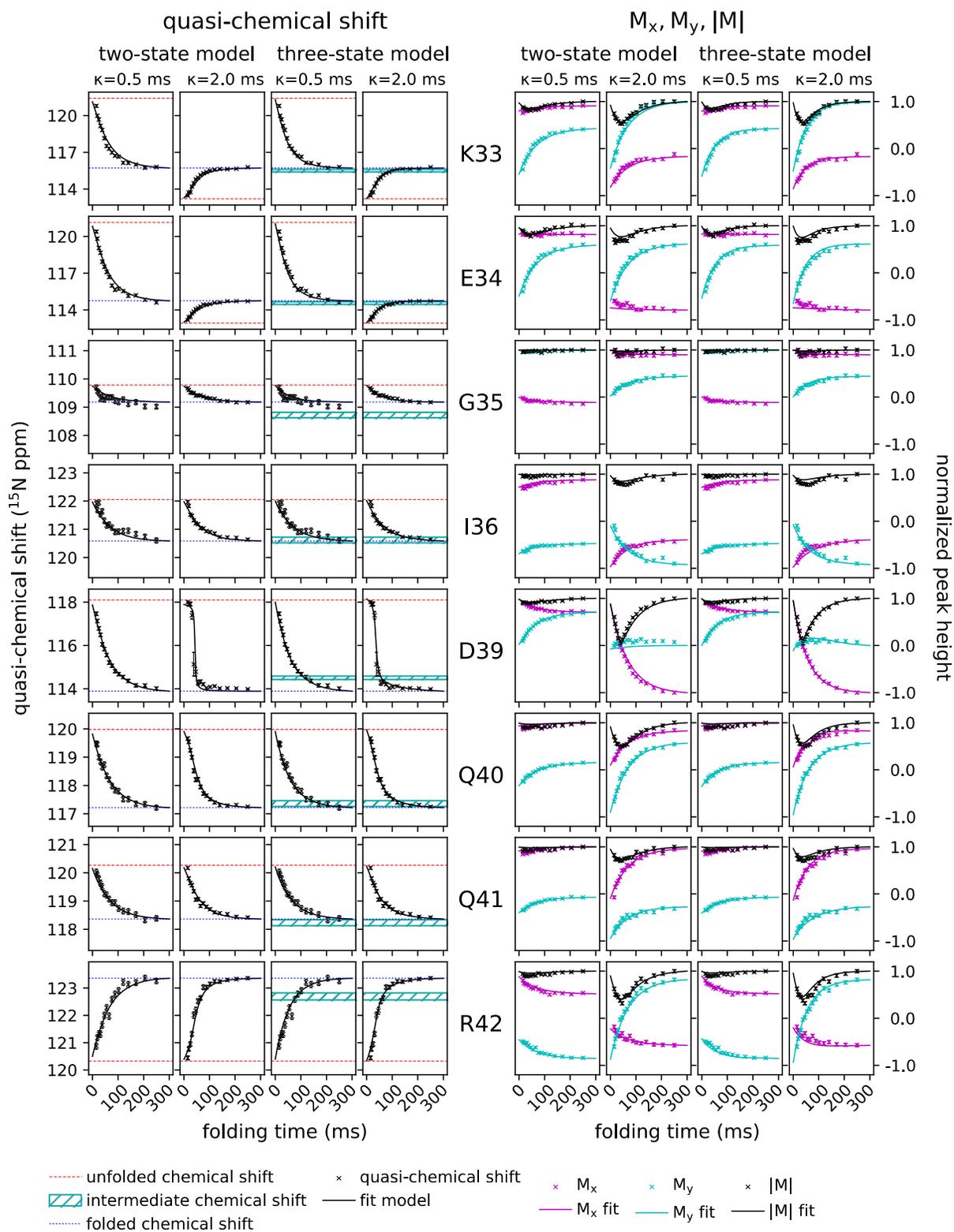


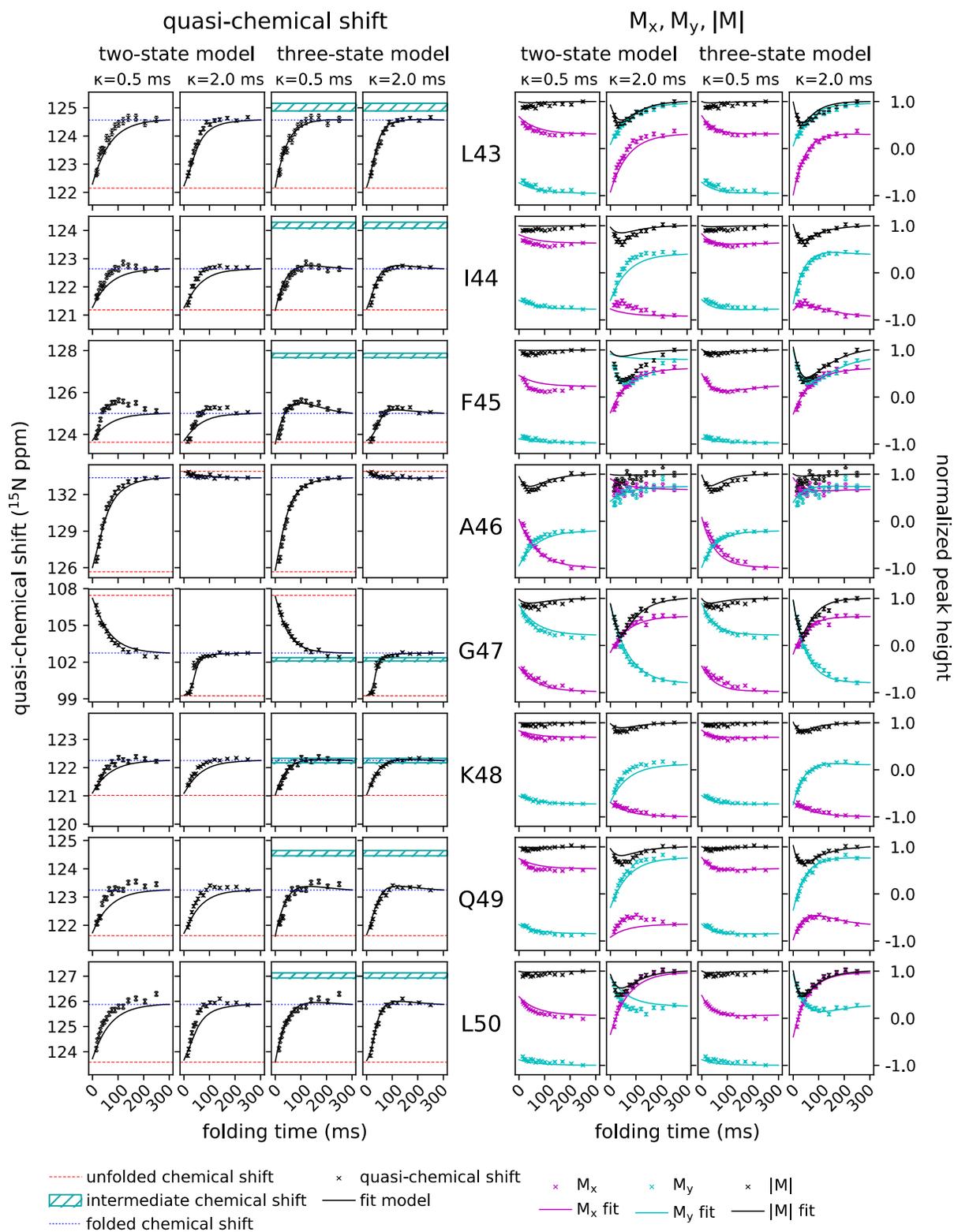
Figure S3: 600-MHz 2D HSQC spectra of VA2-ubiquitin (300 μ M), modulated by $\cos(\omega_N\kappa)$ (top panel) and $\sin(\omega_N\kappa)$ (bottom panel), obtained for $\kappa = 2$ ms and $\tau = 10$ ms. For analysis, peak positions are first assigned from a spectrum calculated from the magnitude spectrum, $M = (M_x^2 + M_y^2)^{1/2}$, where M_x and M_y refer to the cos- and sin-modulated absorptive mode spectra. Peak intensities of the M_x and M_y spectra were then extracted using the `autofit.tcl` routine of NMRPipe,⁵ with zero tolerance on the peak positions. This permits reliable intensities to be derived even if a peak falls below the noise threshold.⁶ Peaks marked * are overlapped and were removed from the analysis.

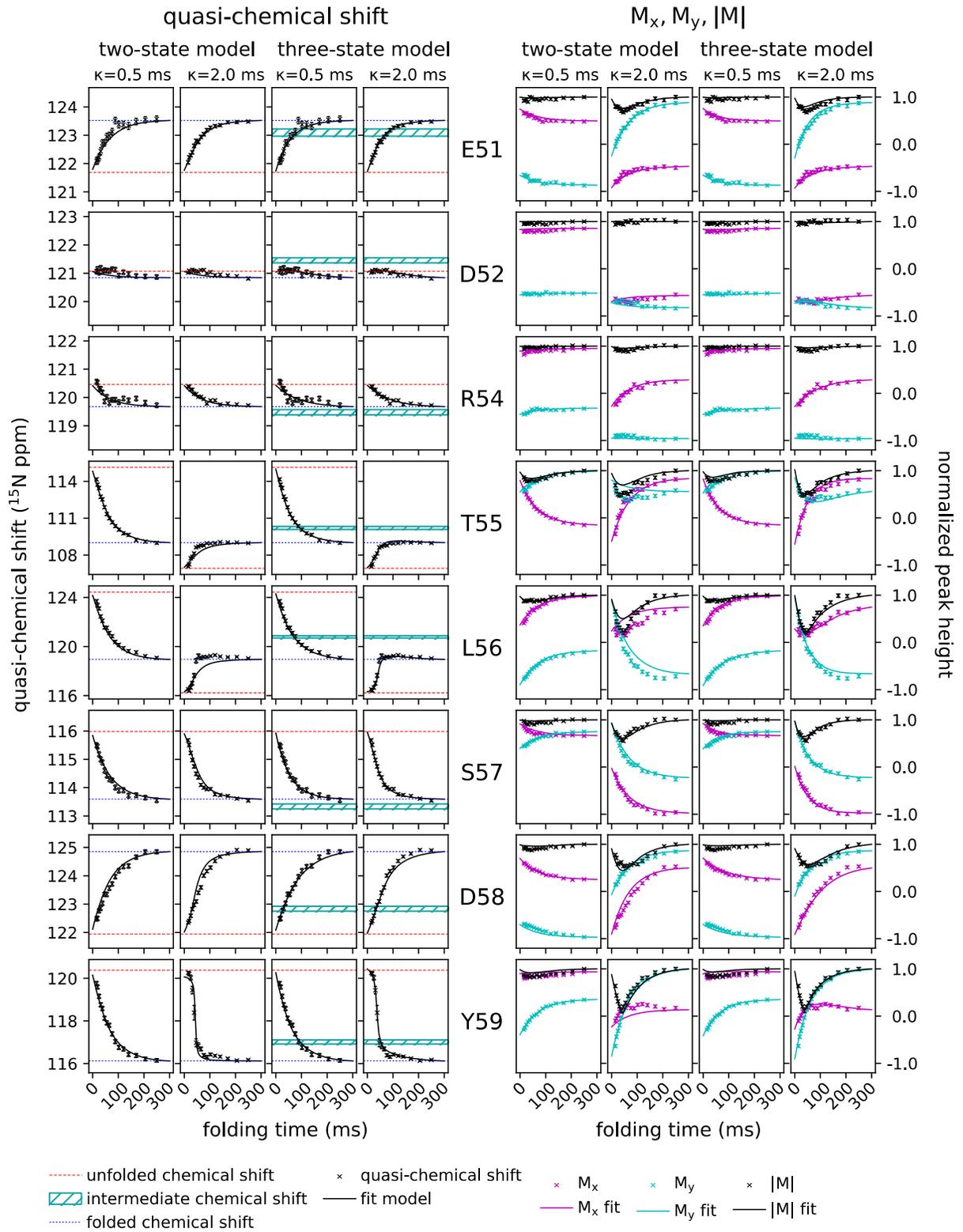


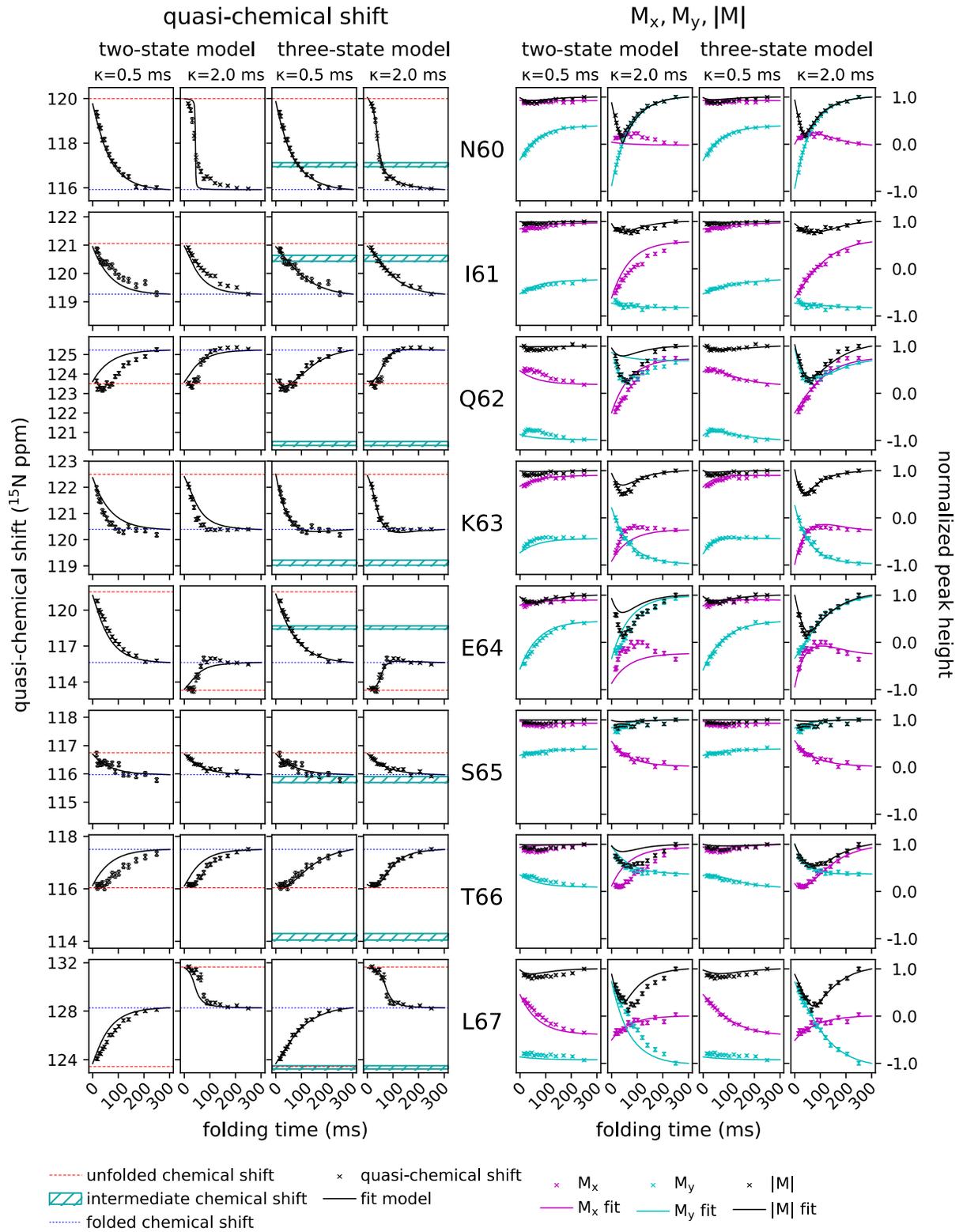


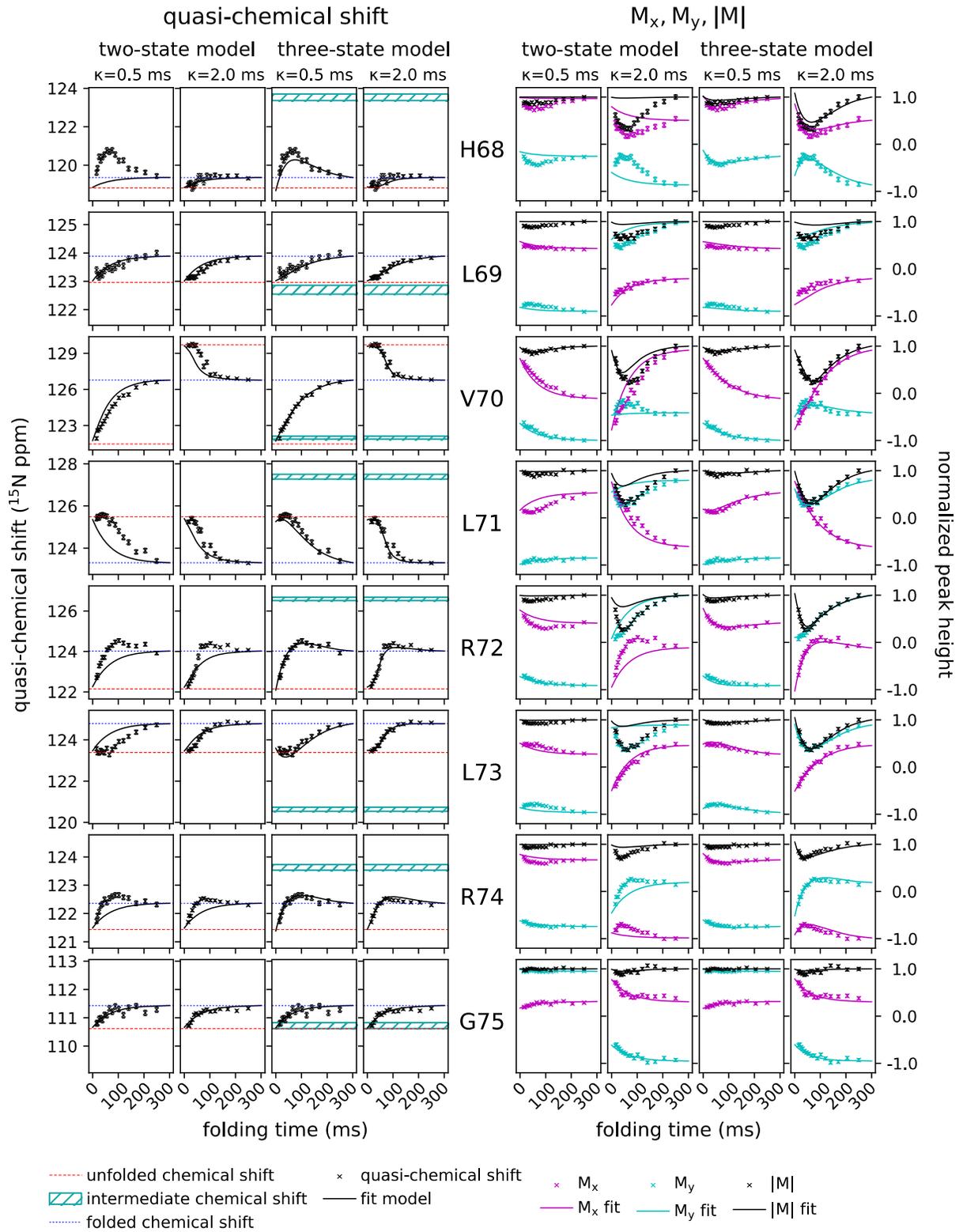












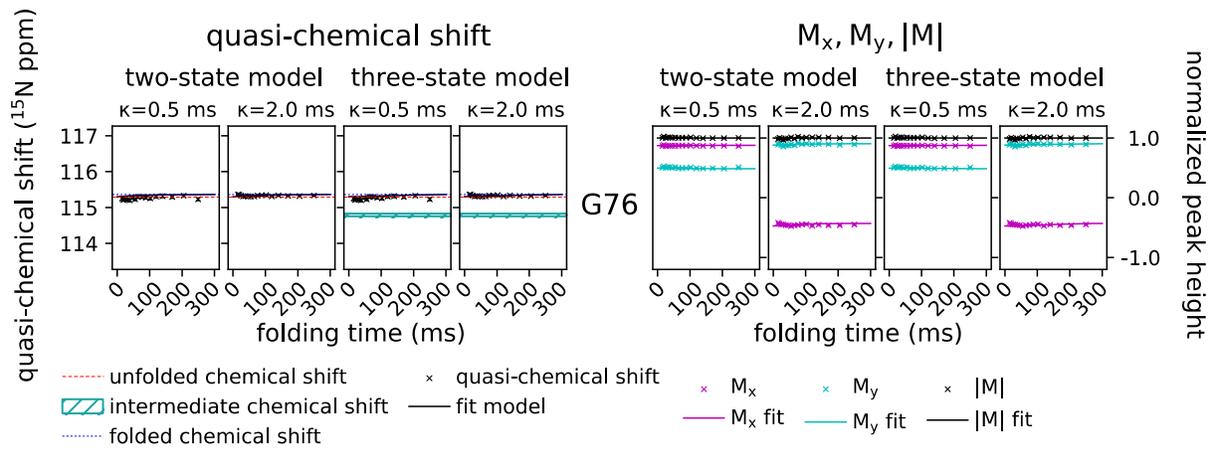


Table S1: F (Folded), U (Unfolded) and I (Intermediate) chemical shifts (in ppm) for the ^{15}N -labeled V17A/V26A ubiquitin labeled double mutant at 22 °C,^a pH 6.4 at 1 bar. Unfolded chemical shifts are from Charlier et al.¹ Folded chemical shifts are taken from the ^{15}N - ^1H HSQC spectrum recorded with the scheme of Figure S1 at 1 bar for $\tau = 248$ ms, after calculating $|M(\tau)| = \sqrt{[M_x(\tau)^2 + M_y(\tau)^2]}$. The intermediate chemical shifts were obtained as residue-specific parameters from joint fitting of the $\kappa=0.5$ ms and $\kappa=2$ ms data to the 3-state model, while treating $k_{u \rightarrow i}$, $k_{u \rightarrow f}$, and $k_{i \rightarrow f}$ as global variables.

Res	F	U	I		Res	F	U	I	
M1					D39	113.89	118.11	114.9	± 0.04
Q2	123.04	124.11	121.8	± 0.25	Q40	117.22	119.99	117.8	± 0.05
I3	118.73	122.89	121.6	± 0.06	Q41	118.36	120.26	118.9	± 0.05
F4	120.16	116.43	125.4	± 0.06	R42	123.36	120.33	123.0	± 0.07
V5	121.52	123.14	124.3	± 0.06	L43	124.57	122.15	125.0	± 0.07
K6	128.13	125.58	129.3	± 0.05	I44	122.64	121.19	123.8	± 0.06
T7	115.67	116.26	116.0	± 0.05	F45	125.01	123.64	128.0	± 0.06
L8	121.71	123.82	120.6	± 0.06	A46	133.36	133.90	133.0	± 0.11
T9	106.06	103.23	106.4	± 0.06	G47	102.75	99.24	102.7	± 0.07
G10	109.57	110.86	109.5	± 0.04	K48	122.25	121.03	122.8	± 0.04
K11	122.19	121.18	121.8	± 0.04	Q49	123.25	121.64	124.6	± 0.05
T12	120.83	118.11	120.1	± 0.05	L50	125.88	123.58	127.1	± 0.05
I13	128.37	125.11	127.2	± 0.06	E51	123.53	121.69	123.2	± 0.06
T14	121.48	119.31	119.4	± 0.07	D52	120.84	121.08	121.5	± 0.05
L15	125.90	125.92	124.9	± 0.06	G53				
E16					R54	119.68	120.47	120.2	± 0.05
V17	126.76	125.74	124.9	± 0.09	T55	109.00	106.91	110.5	± 0.06
E18	119.93	122.04	120.4	± 0.07	L56	118.95	116.22	121.0	± 0.05
P19					S57	113.60	115.99	113.8	± 0.05
S20	103.43	107.66	103.9	± 0.05	D58	124.85	121.94	123.0	± 0.05
D21					Y59	116.12	120.39	117.3	± 0.05
T22	108.85	105.93	109.7	± 0.05	N60	115.92	120.00	117.5	± 0.04
I23	121.53	123.26	121.5	± 0.06	I61	119.26	121.06	121.1	± 0.05
E24					Q62	125.22	123.51	120.6	± 0.05
N25					K63	120.39	122.50	119.2	± 0.05
A26					E64	115.63	113.31	118.7	± 0.07
K27	116.79	119.33	117.1	± 0.05	S65	115.97	116.75	116.3	± 0.05
A28					T66	117.50	116.04	114.2	± 0.06
K29	120.16	120.28	120.1	± 0.05	L67	128.28	131.65	123.8	± 0.07
I30	121.87	122.28	121.8	± 0.05	H68	119.36	118.82	123.5	± 0.08
Q31	123.76	124.47	124.0	± 0.05	L69	123.89	122.96	122.8	± 0.08
D32	120.10	122.32	120.6	± 0.04	V70	126.78	129.70	122.5	± 0.06
K33	115.72	113.19	115.7	± 0.05	L71	123.31	125.49	127.6	± 0.06
E34	114.75	112.92	114.9	± 0.05	R72	124.02	122.15	126.7	± 0.05
G35	109.18	109.79	109.3	± 0.05	L73	124.79	123.39	120.4	± 0.05
I36	120.59	122.06	121.2	± 0.05	R74	122.36	121.44	124.1	± 0.04
P37					G75	111.43	110.62	111.2	± 0.05
P38					G76	115.36	115.29	115.2	± 0.02

^a Note that due to the adiabatic cooling by ~ 3 °C associated with the pressure drop, chemical shift values are reported for 22 °C even though the sample was regulated at 25° during the long recycling delay (~ 10 s) at high pressure.

References

1. Charlier, C.; Alderson, T. R.; Courtney, J. M.; Ying, J.; Anfinrud, P.; Bax, A., Study of protein folding under native conditions by rapidly switching the hydrostatic pressure inside an NMR sample cell. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, E4169-E4178.
2. Kay, L. E.; Keifer, P.; Saarinen, T., Pure Absorption Gradient Enhanced Heteronuclear Single Quantum Correlation Spectroscopy with Improved Sensitivity. *J. Am. Chem. Soc.* **1992**, *114*, 10663-10665.
3. Piotto, M.; Saudek, V.; Sklenár, V., Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. *J. Biomol. NMR* **1992**, *2* (6), 661-665.
4. Levitt, M. H.; Freeman, R., Compensation for pulse imperfections in NMR spin-echo experiments. *J. Magn. Reson.* **1981**, *43* (1), 65-80.
5. Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A., NMRpipe - a multidimensional spectral processing system based on Unix pipes. *J. Biomol. NMR* **1995**, *6* (3), 277-293.
6. Viles, J. H.; Duggan, B. M.; Zaborowski, E.; Schwarzingler, S.; Huntley, J. J. A.; Kroon, G. J. A.; Dyson, H. J.; Wright, P. E., Potential bias in NMR relaxation data introduced by peak intensity analysis and curve fitting methods. *J. Biomol. NMR* **2001**, *21* (1), 1-9.