Supporting Information

Microsecond Time-Scale Conformational Exchange in Proteins: Using Long MD Trajectory to Simulate NMR Relaxation Dispersion Data.

Yi Xue, Joshua M. Ward, Tairan Yuwen, Ivan S. Podkorytov, Nikolai R. Skrynnikov*

Department of Chemistry, Purdue University, 560 Oval Drive, West Lafayette, Indiana 47907-2084

Appendix A. CPMG spin relaxation rate for arbitrary correlation function $g(\tau)$

We address a problem where single spin S moves stochastically between sites with different chemical shifts δ. In this case, the Zeeman Hamiltonian can be separated into time-independent and time-dependent parts:

$$H_Z = H_Z^0 + H_Z(t) = 2\pi\nu_0(1+\overline{\delta})S_z + 2\pi\nu_0(\delta(t)-\overline{\delta})S_z \qquad \text{(A1)}$$

The spin is also subject to the effect of CPMG train, $(\tau_{CP} - 180^\circ - \tau_{CP})_n$, with the Hamiltonian $H_{rf}^{CPMG}(t)$.

From the perspective of spin relaxation theory, it is convenient to separate the Hamiltonian into the spin part (S) and spin-lattice part (SL):

$$H_S = H_Z^0 + H_{rf}^{CPMG}(t) \qquad \text{(A2)}$$

$$H_{SL} = H_Z(t) \qquad \text{(A3)}$$

Assuming that time modulation of chemical shift $H_Z(t)$ is a source of spin relaxation, we can calculate the 'instantaneous' relaxation rate of magnetization S_x:

$$R(t) = \frac{(2\pi\nu_0)^2 \int_0^\infty g(\tau)\mathrm{Tr}\{S_x[S_z,[U_S(t-\tau;t)S_zU_S^\dagger(t-\tau;t),S_x]]\}d\tau}{\mathrm{Tr}\{S_xS_x\}} \qquad \text{(A4)}$$

This result is similar to the standard Redfield-theory expression for transverse relaxation rate [Canet *Prog. NMR Spectrosc.* **21**, 237–291 (1989)]. The difference is that Eq. (A4) takes into consideration the relatively complex character of spin evolution under the effect of CPMG sequence. This aspect has been discussed by Goldman [*J. Magn. Reson.* **149**, 160–187 (2001)]; in particular, Eq. (279) in the Goldman's paper is relevant for the problem at hand.

In Eq. (A4), the correlation function $g(\tau)$ has the same meaning as in the text, Eq. (1), and $U_S(t-\tau;t)S_zU_S^\dagger(t-\tau;t)$ describes the evolution of S_z under the effect of H_S beginning from the point of time $t-\tau$ and ending with t.* The character of $U_S(t-\tau;t)S_zU_S^\dagger(t-\tau;t)$ evolution is intuitively clear: operator S_z is inverted at regular intervals of time in response to the train of 180° pulses. The outcome is $S_z \cdot f$, where f is the square-wave function (see below).

* In $U(t-\tau;t)$ we reversed the order of arguments compared to the original definition used by Goldman.

Prior to evaluating $U_S(t-\tau;t)S_zU_S^{\dagger}(t-\tau;t)$ it is convenient to replace the time variable: $t-\tau=t'$. In practice, τ is restricted to a short interval from 0 to several times τ_{ex} (due to the presence of decaying correlation function, $g(\tau)$, in the integrand). In Redfield theory it is assumed that τ_{ex} is short on a time scale of spin relaxation; therefore, we do not distinguish between $R(t)=R(t'+\tau)$ and $R(t')$. Evaluating the spin commutators in Eq. (A4) we obtain:

$$R(t')=(2\pi\nu_0)^2\int_0^{\infty}g(\tau)f(t';t'+\tau)d\tau \qquad \text{(A5)}$$

The square-wave function $f(t';t'+\tau)$ is illustrated in Fig. A1. Note that $f(t';t'+\tau)$ always starts out positive (representing S_z operator prior to application of the first inversion pulse).

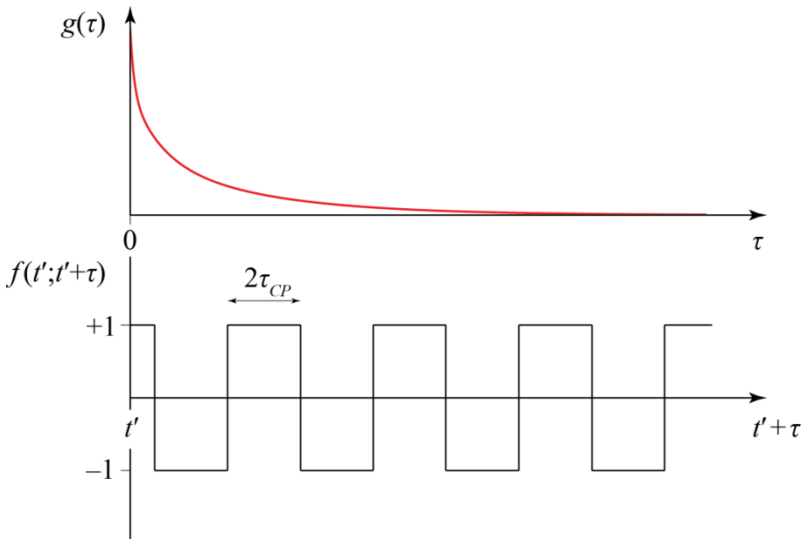


Fig. A1. Integrand functions in Eq. (A5): schematic representation

From Fig. A1 it becomes clear that the 'instantaneous' relaxation rate varies depending on the placement of t' relative to the pulses in CPMG train. To derive the effective rate, let us average the result Eq. (A5) over the interval between the two pulses, $2\tau_{CP}$:

$$R_{1\rho,CPMG}^{ex}=\frac{1}{2\tau_{CP}}\int_{t_0'}^{t_0'+2\tau_{CP}}(2\pi\nu_0)^2\int_0^{\infty}g(\tau)f(t';t'+\tau)d\tau dt' \qquad \text{(A6)}$$

Note that this kind of averaging implies that the relaxation curve can be linearized in the interval between the two consecutive pulses, $2\tau_{CP}$ (which is normally a safe assumption). Let us now change the order of integration in Eq. (A6) and start from computing the integral with respect to t'. The calculation is schematically illustrated in Fig. A2: the integral is evaluated by adding the

contributions from multiple traces such as traces B–D shown in the plot. The result of the integration, function $\Delta(\tau)$, is illustrated in panel E. The expression in Eq. (A6) is therefore reduced to:

$$R_{1\rho,CPMG}^{ex} = (2\pi\nu_0)^2 \int_0^\infty g(\tau)\Delta(\tau)d\tau \qquad \text{(A7)}$$

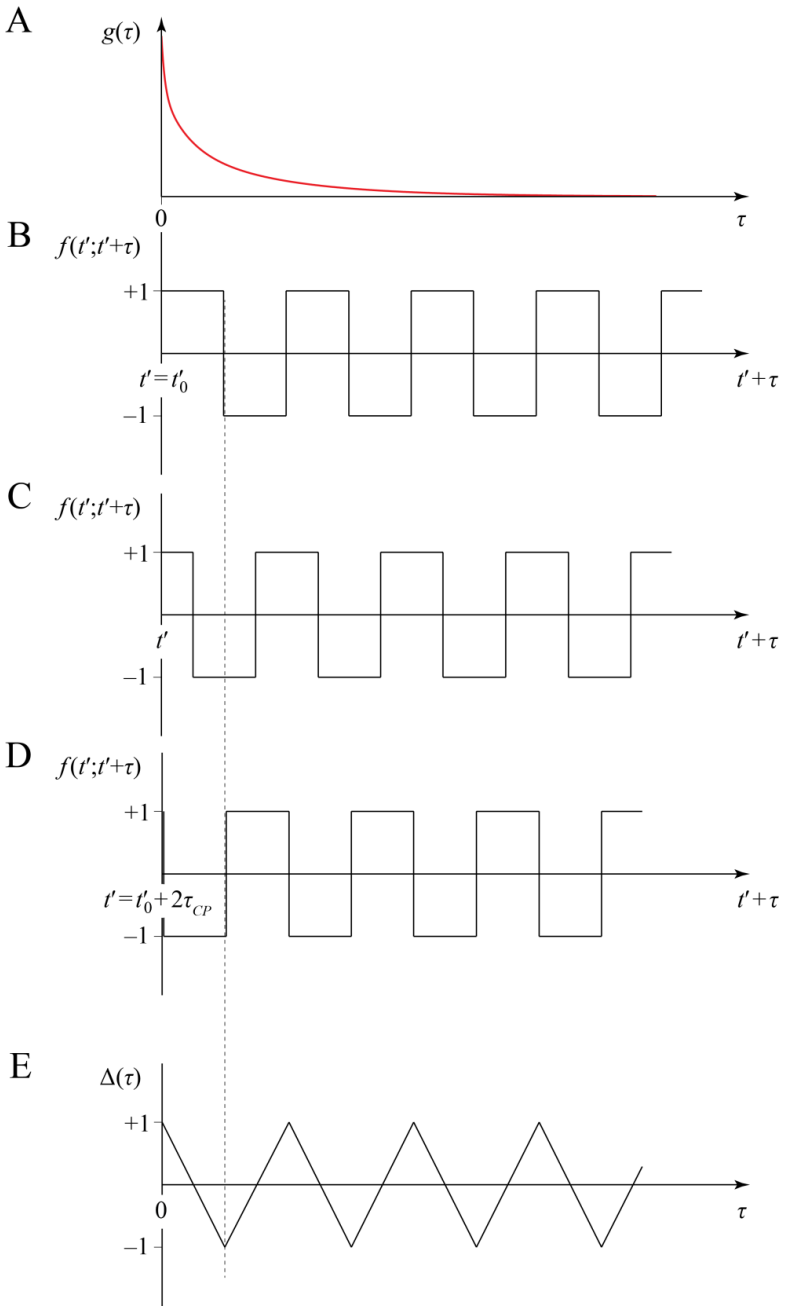


Fig. A2. Integrating Eq. (A6) with respect to t': graphical scheme.

To emphasize the similarity between Eq. (A7) and Eq. (3) in the text, we cast the result in a slightly different form:

$$R_{1\rho,CPMG}^{ex} = (2\pi\nu_0)^2 \int_0^\infty g(\tau)\,\mathrm{tri}(2\pi\nu_{CP}\tau)\,d\tau \qquad \text{(A8)}$$

where $\nu_{CP} = 1/4\tau_{CP}$ and $\mathrm{tri}(x)$ is a triangle-wave function shown in Fig. A(3), which closely resembles $\cos(x)$. The result Eq. (A8) reproduces Eq. (4) in the text.

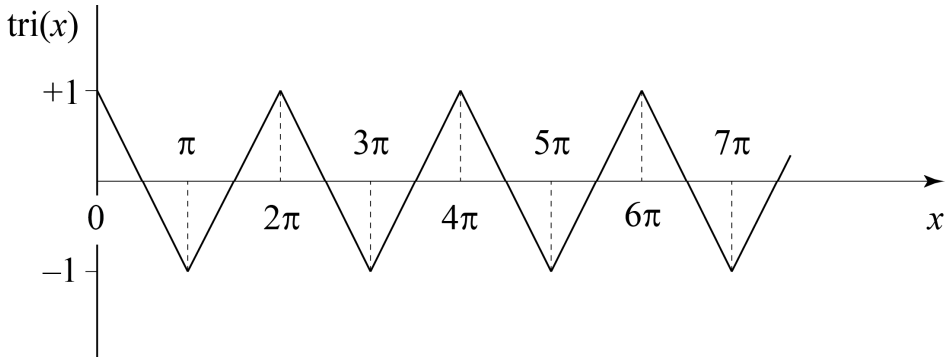


Fig. A3. Graph of triangle-wave function, as used in Eq. (A8)

For the particular case when $g(\tau)$ is a multiexponential function,

$$g(\tau) = \sum_{j=1}^{N} c_j \exp(-\tau / \tau_{ex}^j) \qquad \text{(A9)}$$

the expression Eq. (A8) can be readily integrated. Consider the contribution from an individual exponential component, $\exp(-\tau / \tau_{ex}^j)$. The domain of integration in Eq. (A8) can be divided into a series of intervals: $[0, 4\tau_{CP}]$, $[4\tau_{CP}, 8\tau_{CP}]$, $[8\tau_{CP}, 12\tau_{CP}]$, …, . The respective contributions to the integral, termed a_0, a_1, a_2, … , are related according to $a_{i+1} = a_i \cdot \kappa$, where $\kappa = \exp(-4\tau_{CP} / \tau_{ex}^j)$. Consequently, the sum of a_i can be computed as a sum of geometric series, $a_0 / (1-\kappa)$, ultimately leading to the desired result:

$$R_{1\rho,CPMG}^{ex}(\nu_{CP}) = (2\pi\nu_0)^2 \sum_{j=1}^{N} c_j \tau_{ex}^j \left\{ 1 - \frac{\tau_{ex}^j}{\tau_{CP}} \tanh \frac{\tau_{CP}}{\tau_{ex}^j} \right\} \qquad \text{(A10)}$$

This expression is equivalent to Eq. (5.2) in the text.

Figure S1. (A) Scatter plot of χ_1(C14) vs. χ_1(C38) derived from the 1-ms trajectory of BPTI. The data are colored according to χ_3 values; the clusters that contain conformational species M, m_{C14}, and m_{C38} are labeled in the plot. (B) Scatter plot of χ_2(C14) vs. χ_2(C38) for conformational species M, where M are defined as follows: $-120^\circ < \chi_1(\text{C14}) < 0^\circ$; $0^\circ < \chi_1(\text{C38}) < 120^\circ$; $0^\circ < \chi_3 < 180^\circ$. The cluster structure observed in this plot serves as a basis for separating the sub-states M1, M2, and M3. (C) Scatter plot of χ_2(C14) vs. χ_2(C38) for conformational species m_{C14}, where m_{C14} are defined as follows: $0^\circ < \chi_1(\text{C14}) < 120^\circ$; $0^\circ < \chi_1(\text{C38}) < 120^\circ$; $-180^\circ < \chi_3 < 0^\circ$. (D) Scatter plot of χ_2(C14) vs. χ_2(C38) for conformational species m_{C38}, where m_{C38} are defined as follows: $-120^\circ < \chi_1(\text{C14}) < 0^\circ$; $-120^\circ < \chi_1(\text{C38}) < 0^\circ$; $-180^\circ < \chi_3 < 0^\circ$. For the purpose of plotting, torsional angles χ_1 (χ_2) are defined over the interval from –120° to 240° (0° to 360°). In the main text, the standard range is used for both angles, –180° to 180°.

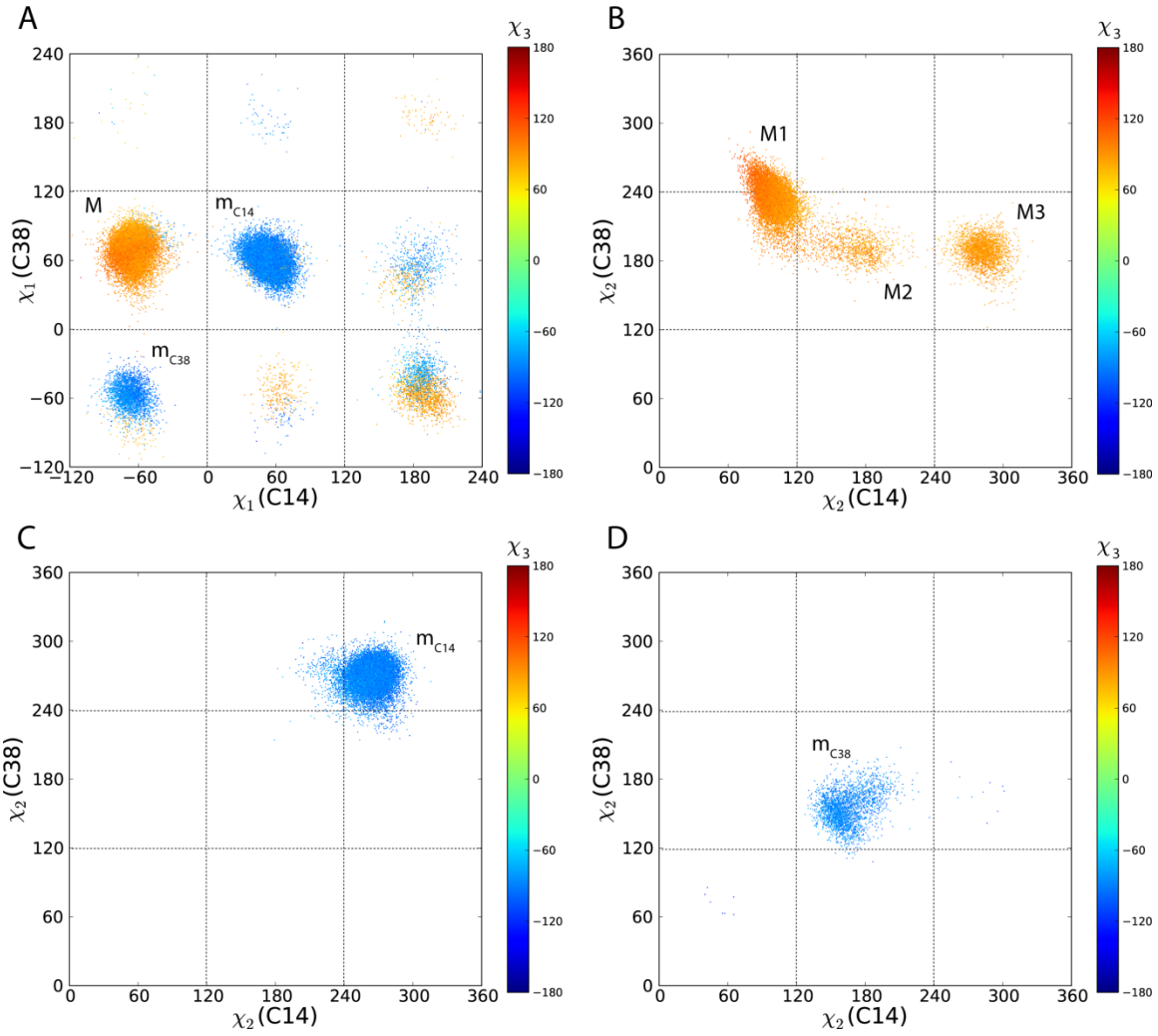


Figure S2. Time variation of (A,B) χ_1 torsional angles and (C,D) backbone $\delta(^{15}N)$ chemical shifts in disulfide-bonded residues C14, C38 as obtained from the 1-ms MD trajectory of BPTI. This figure is a copy of Fig. 1, color-coded according to the isomeric state of the C14–C38 disulfide bridge: M1 – blue, M2 – orange, M3 – magenta, m_{C14} – red, m_{C38} – green, other states – grey.

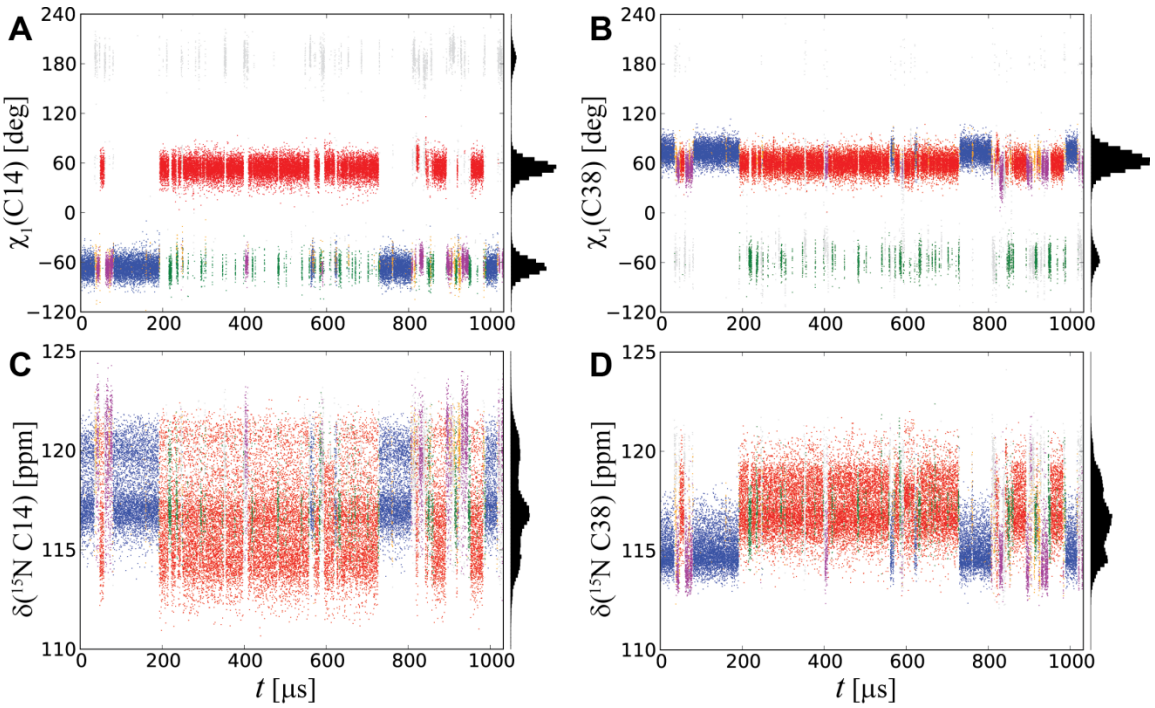


Table S1. ^{15}N chemical shift differences between the conformational states M1(M), m_{C14}, and m_{C38}. MD-based calculations have been performed using the customized version of the program SHIFTX+ which did not include BPTI as a part of the training set (courtesy of B. Han, University of Alberta).

Res.	$\Delta\delta_{M1,mC14}$ (ppm)			$\Delta\delta_{M1,mC38}$ (ppm)		
	Exptl.	Prev. calc.	Calc. MD	Exptl.	Prev. calc.	Calc. MD
C14	3.6	3.7	1.8	−0.4	0.1	0.4
K15	4.7	1.4	−1.3	−0.5	−0.4	−1.4
C38	|0.8|	0.6	−1.9	−1.7	−4.9	−2.0
R39	|1.2|	0.4	−0.3	−3.7	−2.3	−2.3

Table S2. ^{15}N chemical shift differences between the conformational states M1(M), m_{C14}, and m_{C38}. MD-based calculations have been performed using the program SPARTA+ [Shen & Bax *J.Biomol. NMR* **48**, 13–22 (2010)]. Similar to the standard version of SHIFTX+, this program includes BPTI as a part of the training set.

Res.	$\Delta\delta_{M1,mC14}$ (ppm)			$\Delta\delta_{M1,mC38}$ (ppm)		
	Exptl.	Prev. calc.	Calc. MD	Exptl.	Prev. calc.	Calc. MD
C14	3.6	3.7	1.7	−0.4	0.1	0.5
K15	4.7	1.4	−1.0	−0.5	−0.4	−1.7
C38	|0.8|	0.6	−0.3	−1.7	−4.9	−2.1
R39	|1.2|	0.4	−1.4	−3.7	−2.3	−1.8

Figure S3. Time variation of torsional angles χ_1 and chemical shifts $\delta(^{15}\mathrm{N})$ in the disulfide bridge C5-C55 (analogous to Fig. 1 in the text). Note rare occurrence of the distinctive long-lived species (lifetime ~10 μs). Note also the presence of two rapidly interconverting populations in panel (C). We have determined that these two populations are due to the presence of the nearby aromatic ring, residue F4. Their precise origin, however, remains unclear and may be related to the details of SHIFTX+ algorithm.

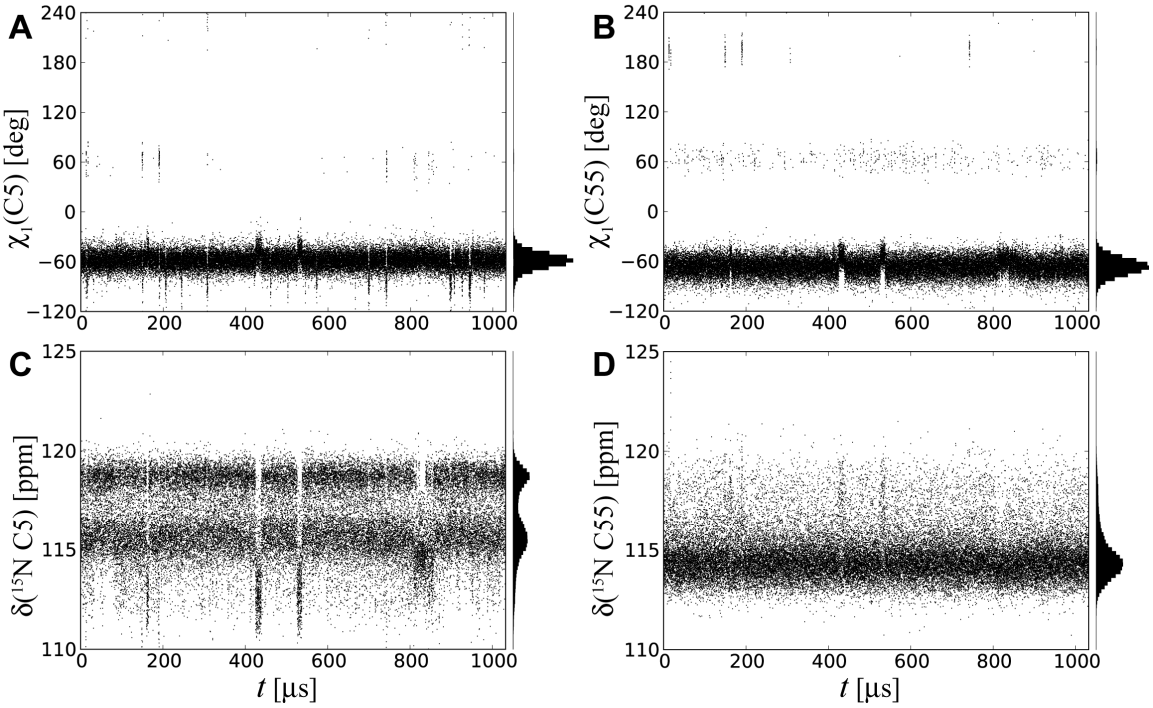
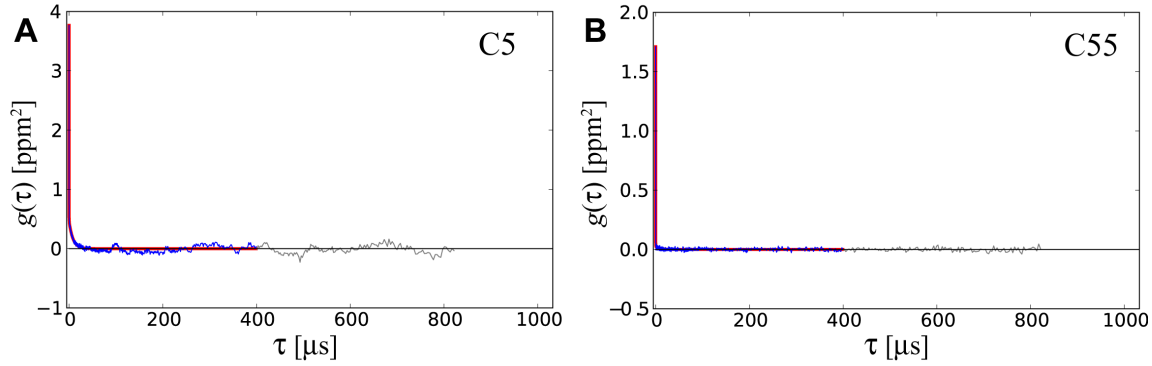


Figure S4. Chemical shift correlation functions for backbone $^{15}\mathrm{N}$ spins from residues C5, C55 (analogous to Fig. 2 in the text). Note the presence of μs component in the correlation function of residue C5. This component, however, is too small and decays too rapidly (~10 μs) to produce any detectable R^{ex} effect (cf. Fig. 3 in the text). In principle, conformational exchange on the time scale ~10 μs can be observed using proton-based relaxation dispersion experiments.



Appendix B. $^1H^N$,^{15}N multi-quantum relaxation rates ΔR_{MQ}^{ex} in BPTI.

The MD-based $^1H^N$,^{15}N multi-quantum relaxation rates can be calculated according to:

$$\Delta R_{MQ}^{ex} = R_{DQ}^{ex} - R_{ZQ}^{ex} \tag{B1}$$

$$R_{DQ/ZQ}^{ex} = \int_0^\infty g_{DQ/ZQ}(\tau) d\tau \tag{B2}$$

$$g_{DQ/ZQ}(\tau) = < (\omega_H(t) \pm \omega_N(t) - \overline{\omega_H \pm \omega_N})(\omega_H(t+\tau) \pm \omega_N(t+\tau) - \overline{\omega_H \pm \omega_N}) > \tag{B3}$$

where $\omega_H(t) = 2\pi\nu_0^H \delta^{1H}(t)$, $\omega_N(t) = 2\pi\nu_0^N \delta^{15N}(t)$ and Larmor frequencies ν_0^H and ν_0^N have the opposite signs. It is straightforward to show that Eqs. (B1−B3) can be reduced to:

$$\Delta R_{MQ}^{ex} = 2\int_0^\infty (g_{HN}(\tau) + g_{NH}(\tau)) d\tau \tag{B4}$$

$$g_{HN}(\tau) = < (\omega_H(t) - \overline{\omega_H})(\omega_N(t+\tau) - \overline{\omega_N}) > \tag{B5}$$

$$g_{NH}(\tau) = < (\omega_N(t) - \overline{\omega_N})(\omega_H(t+\tau) - \overline{\omega_H}) > \tag{B6}.$$

For infinitely large statistics the cross-correlation functions $g_{HN}(\tau)$ and $g_{NH}(\tau)$ are identical; in the case of limited statistics, however, they should be computed separately. Cross-correlation functions Eqs. (B5, B6) are arguably more interesting than auto-correlation functions Eq. (B3). Indeed, $g_{HN}(\tau), g_{NH}(\tau)$ highlight concerted response of $^1H^N$ and ^{15}N spins to one and the same μs motional modes. Given that chemical shifts of amide protons and nitrogens are sensitive to different structural factors, the partial correlation between the two can be revealing. On the other hand, $g_{DQ}(\tau)$, $g_{ZQ}(\tau)$ are more convenient from a practical standpoint – these functions are monotonous (decaying) and therefore well suited for multiexponential fitting, cf. Fig. 2. Hence, we choose Eqs. (B1–B3) to conduct the calculations.

Fig. B1 shows the comparison between the simulated ΔR_{MQ}^{ex} rates and the experimental data obtained by Wang and Palmer [*J.Biomol. NMR* **24**, 263-268 (2002)]. The set of residues exhibiting higher ΔR_{MQ}^{ex} rates is correctly reproduced in the simulations, cf. panel A vs. panel B in the plot. Aside from that, the quantitative agreement is obviously lacking. The reason for that is the exaggerated amount of $^1H^N$ chemical shift variation. The proton CS correlation functions $g(\tau)$ contain long-lived components with the amplitude $< (\delta(t) - \overline{\delta})^2 >^{1/2}$ of up to 0.6 ppm. The CS variation of this magnitude translates into excessive amount of line-broadening and leads to overestimation of ΔR_{MQ}^{ex}. At this point one should recall that the nominal error of $^1H^N$ chemical shift calculations by SHIFTX+ is 0.44 ppm [Han *et al*. *J.Biomol.NMR* **50**, 43-57 (2011)]. Thus, a

major portion of the apparent $\delta^{1H}(t)$ variation most likely stems from the error in chemical shift calculations.

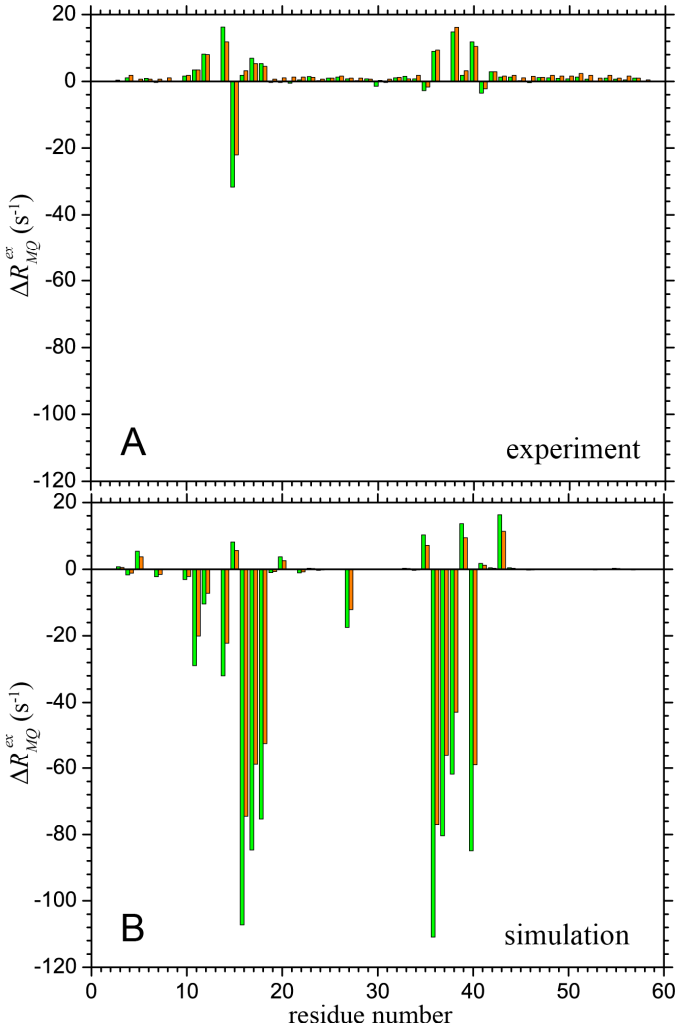


Fig. B1. Experimental (A) and simulated (B) values of $^{1}H^{N}$,^{15}N ΔR_{MQ}^{ex} rates in BPTI. Conditions: temperature 300 K, spectrometer frequency 500 and 600 MHz (orange and green bars, respectively). The experimental data are from the work of Wang and Palmer [*J.Biomol. NMR* **24**, 263-268 (2002)].

Of note, most of the large ΔR_{MQ}^{ex} values in our calculations are *negative*. In terms of two-state model, this means that δ_{ab}^{1H} and δ_{ab}^{15N}, as expressed in ppm, have the same sign. This is confirmed by the scatter plots $\delta^{1H}(t)$ vs. $\delta^{15N}(t)$, which can be displayed in HSQC-type format (not shown). In these plots the resonances tend to group along the main diagonal. In contrast, the majority of the experimental ΔR_{MQ}^{ex} values are *positive*. To determine whether one can expect any statistically significant trend with regard to the signs of ΔR_{MQ}^{ex}, we have examined a set of

randomly selected CS datasets from the BMRB depository. As it turns out, there is only a very weak correlation between $^{1}H^{N}$ and ^{15}N shifts in disulfide-bonded cystein residues, i.e. both positive and negative values of ΔR_{MQ}^{ex} can be found at these sites, depending on the details of the specific protein system.

Finally, note that the calculated ΔR_{MQ}^{ex} rates scale as a square of spectrometer frequency, as appropriate for the fast exchange regime (panel B in Fig. B1). The same is true for the experimental ΔR_{MQ}^{ex} rates in residues C14 and K15. This is no longer fulfilled, however, for residue C38, which is mostly affected by the slower form of exchange (panel A).

Figure S5 (pp. S13 – S21 below). Full complement of chemical shift correlation functions $g(\tau)$ for backbone ^{15}N spins in BPTI, cf. Fig. 2 in the text.

Fig. S5

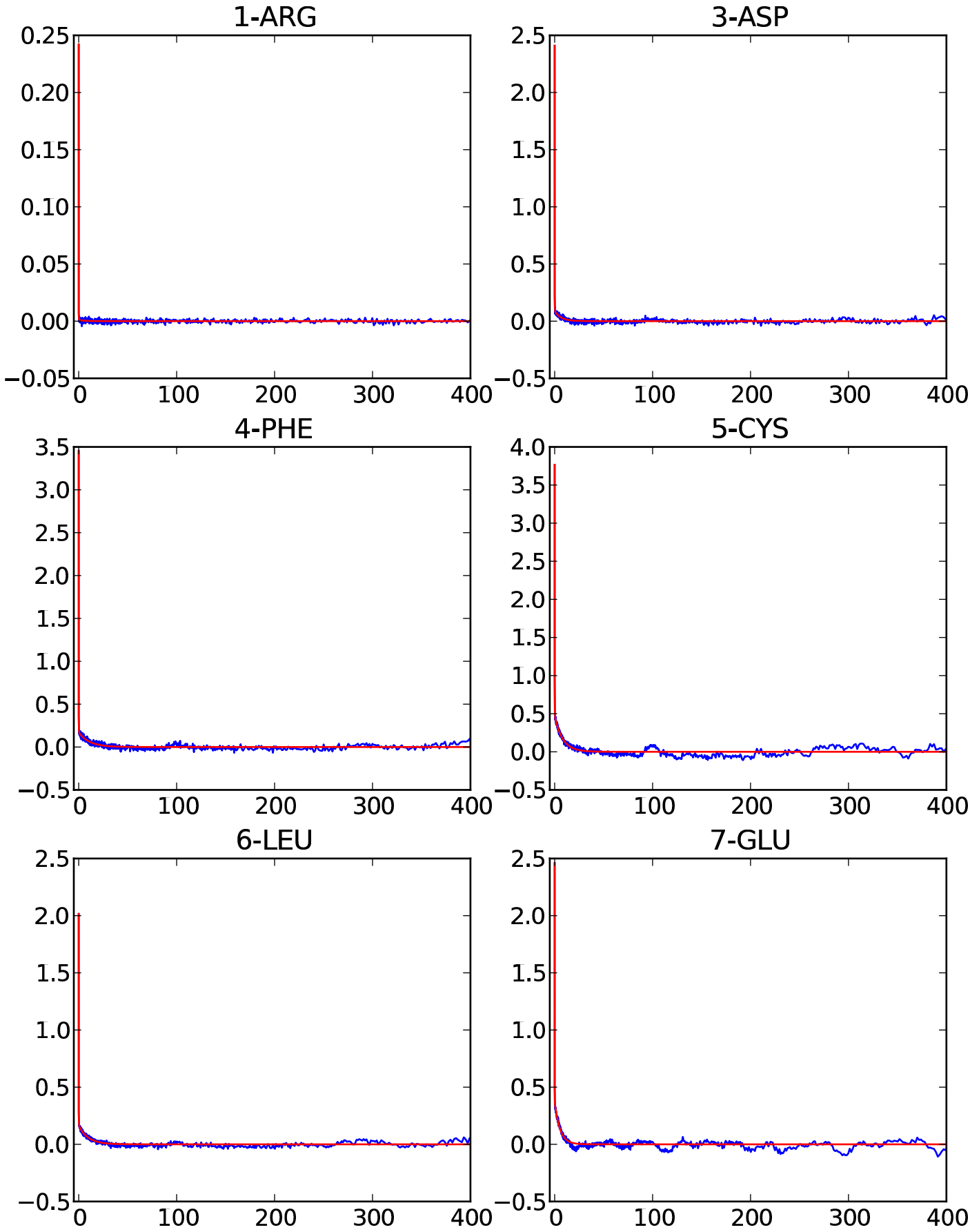


Fig. S5

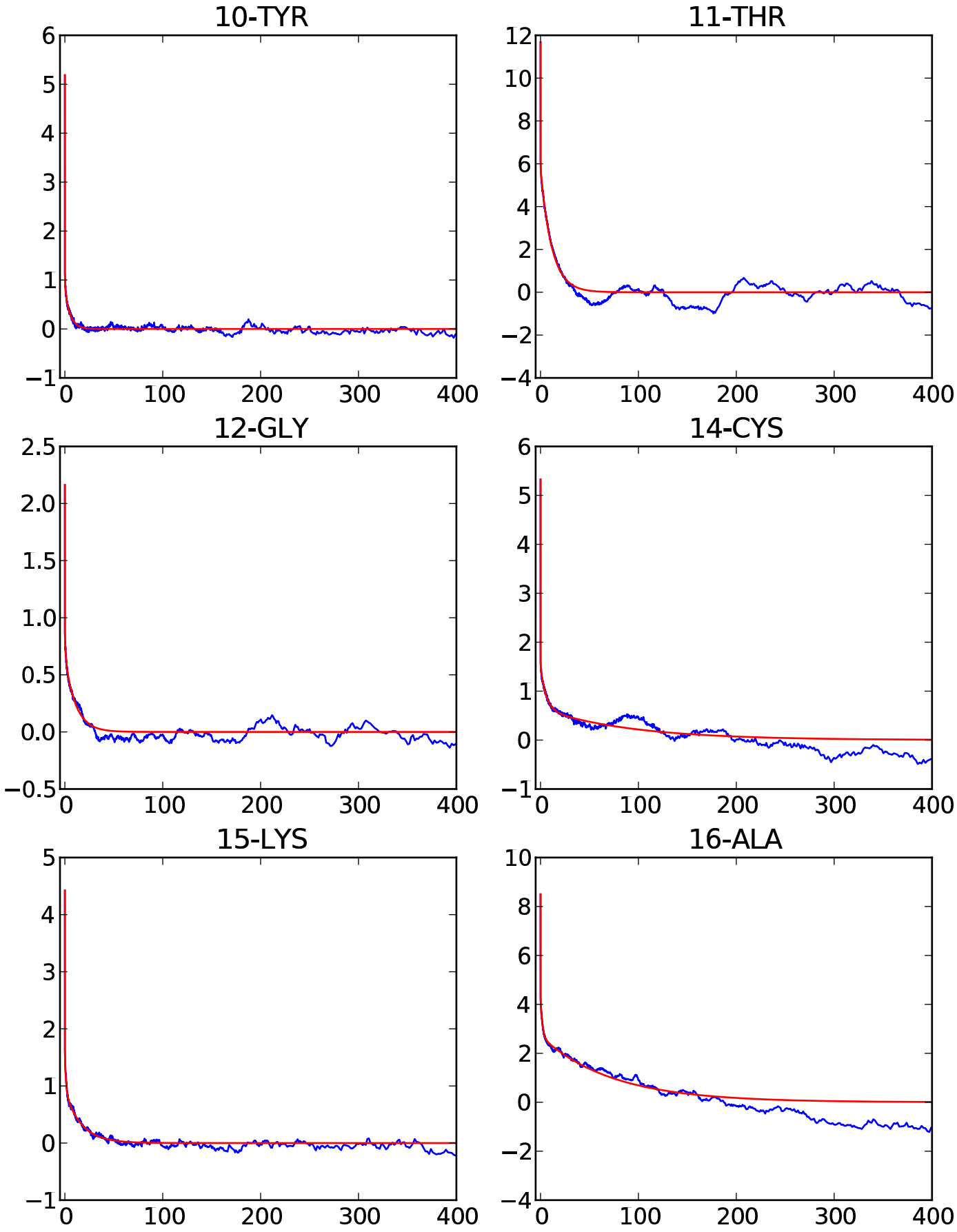


Fig. S5

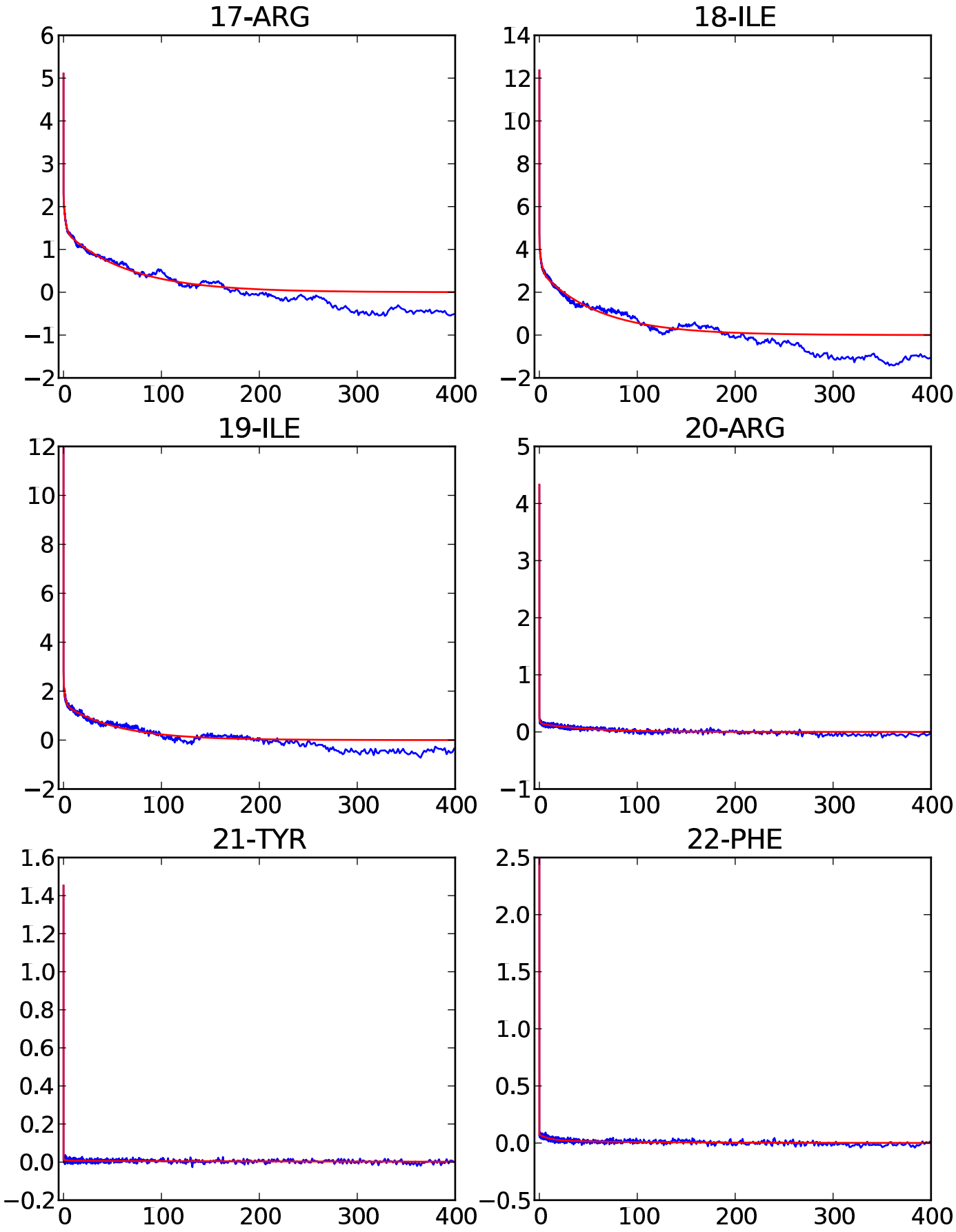


Fig. S5

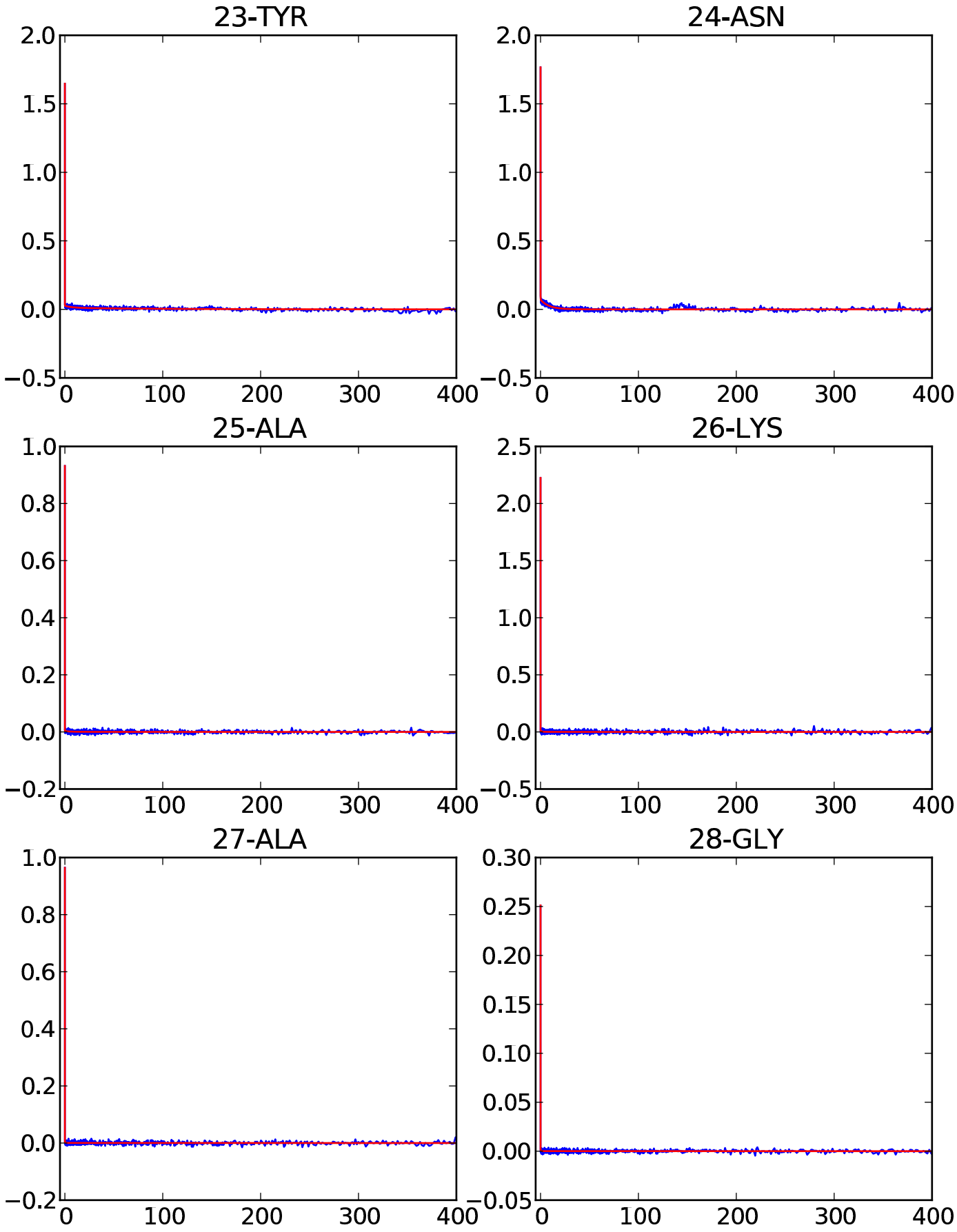


Fig. S5

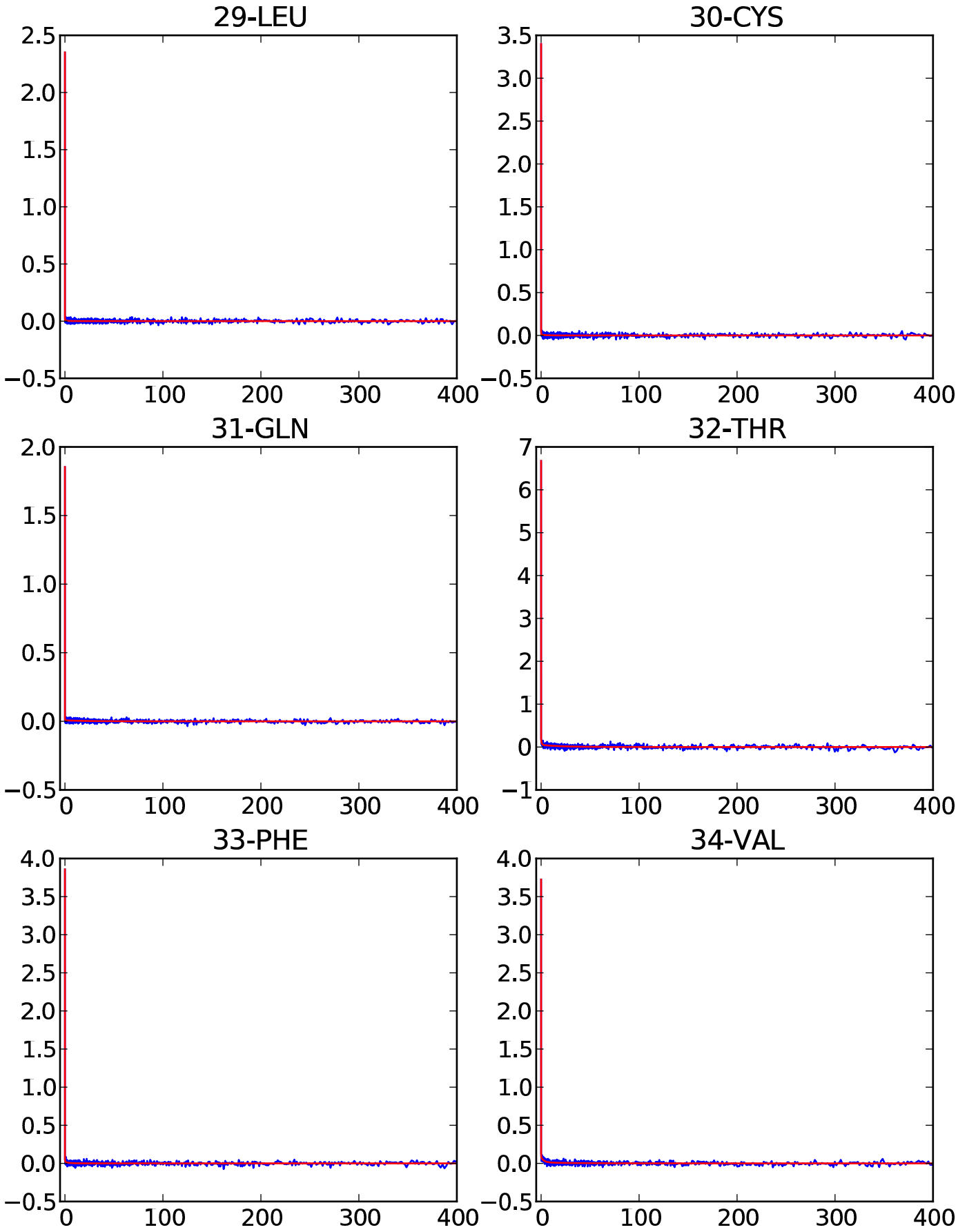


Fig. S5

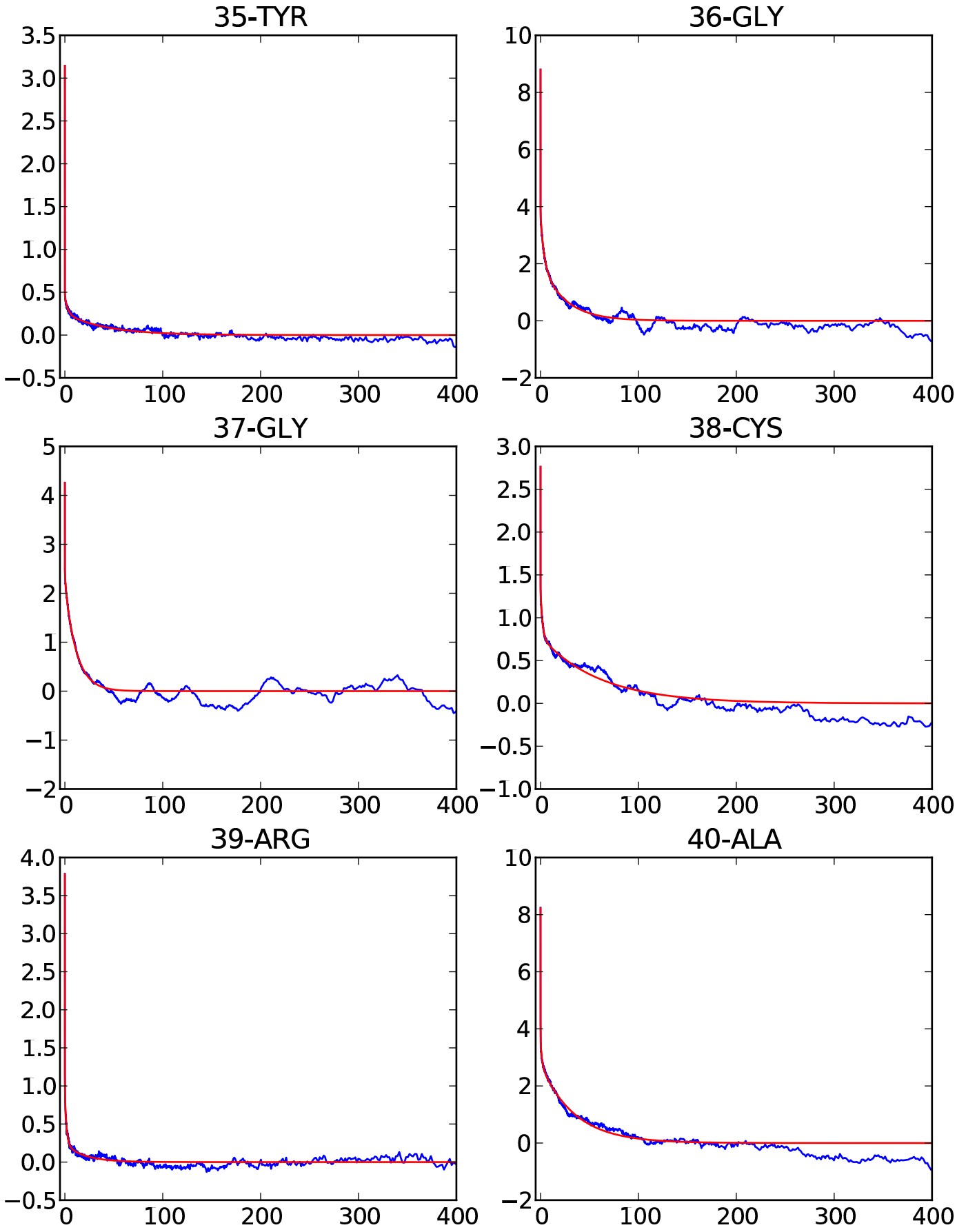


Fig. S5

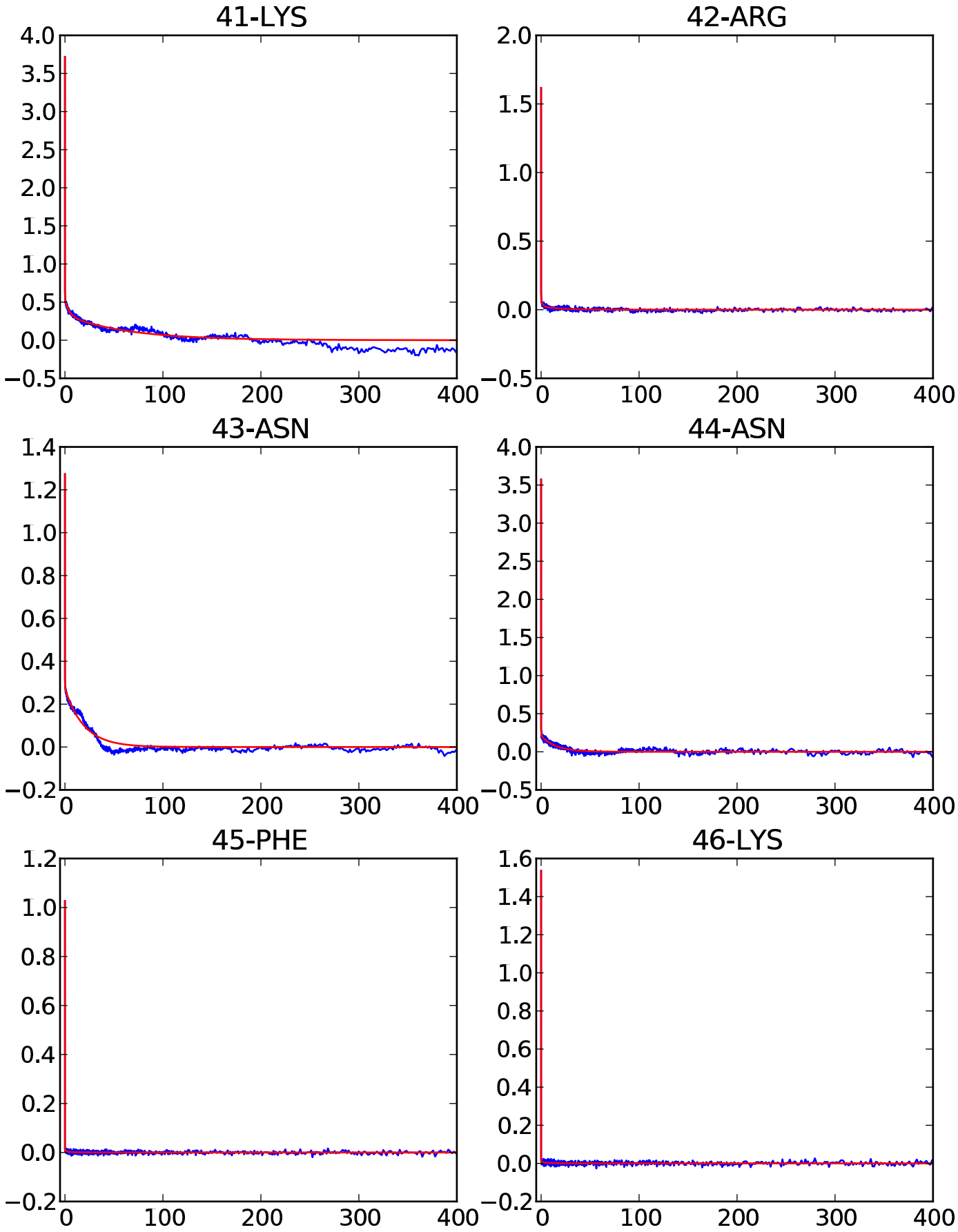


Fig. S5

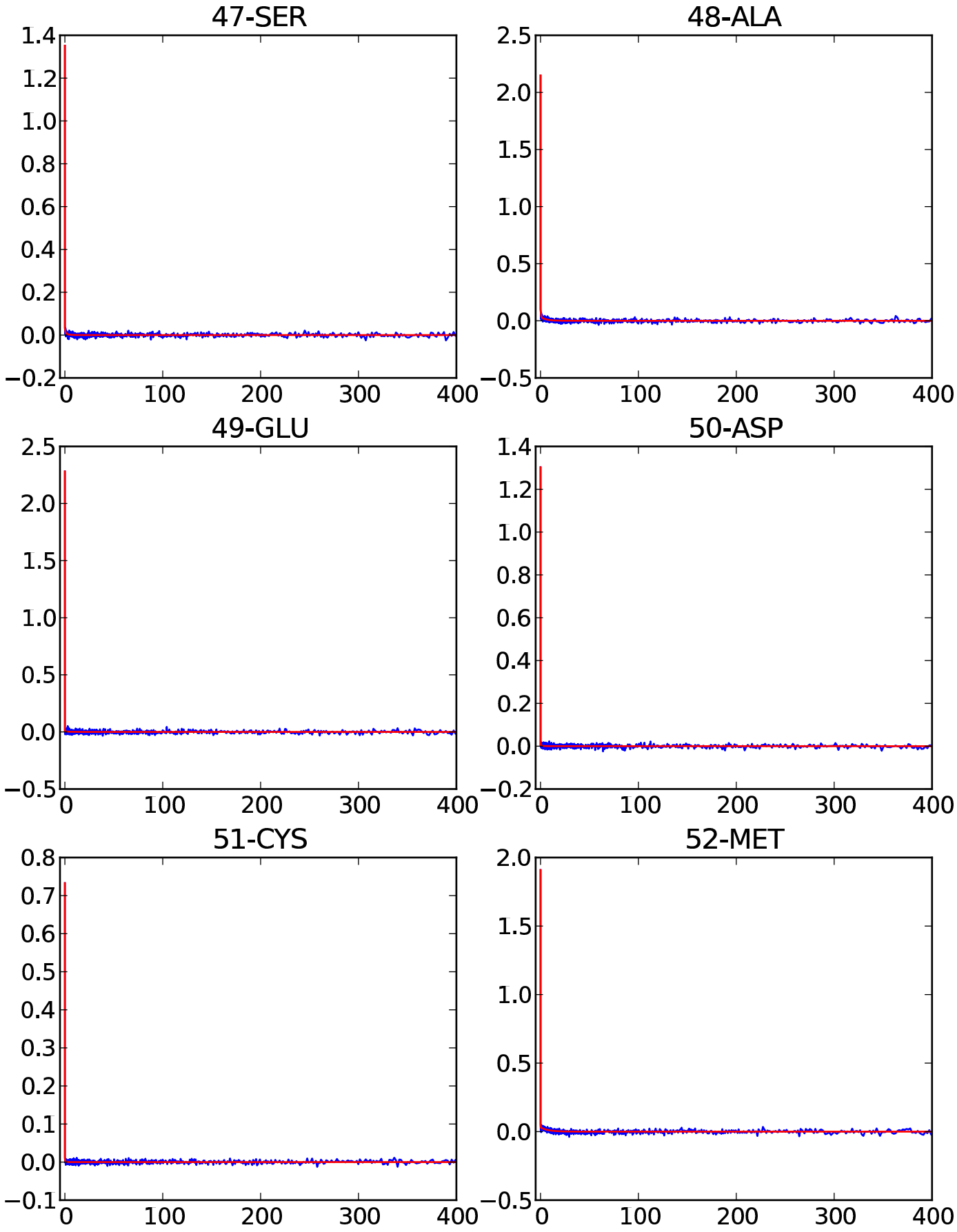
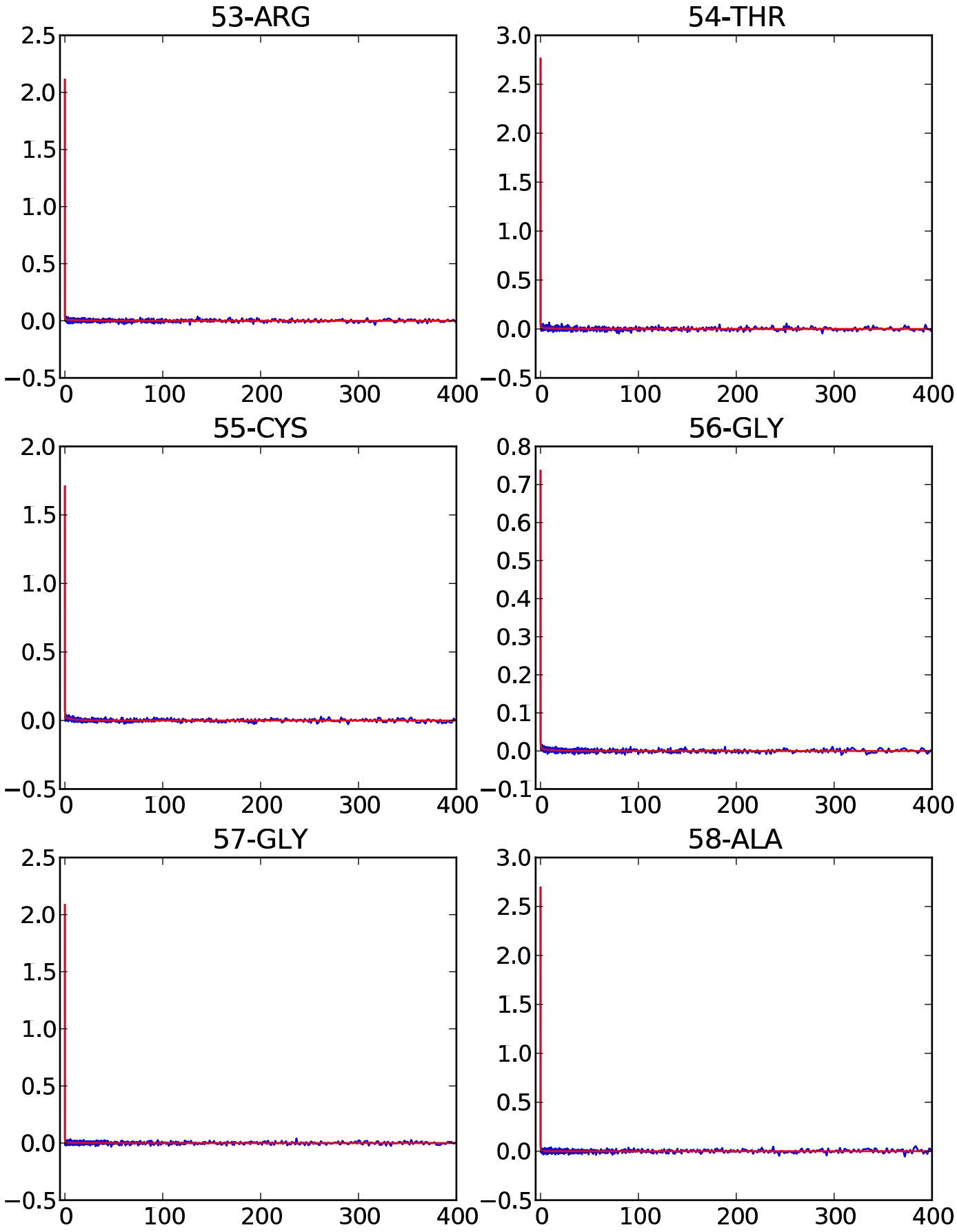


Fig. S5



Complete references:

14. Shen, Y.; Lange, O.; Delaglio, F.; Rossi, P.; Aramini, J. M.; Liu, G. H.; Eletsky, A.; Wu, Y. B.; Singarapu, K. K.; Lemak, A.; Ignatchenko, A.; Arrowsmith, C. H.; Szyperski, T.; Montelione, G. T.; Baker, D.; Bax, A. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 4685.