Measuring hydrogen exchange rates in invisible protein excited states

Dong Long[a], Guillaume Bouvignies[a,1], and Lewis E. Kay[a,b,2]

[a]Departments of Molecular Genetics, Biochemistry, and Chemistry, University of Toronto, Toronto, ON, Canada M5S 1A8; and [b]Program in Molecular Structure and Function, Hospital for Sick Children, Toronto, ON, Canada M5G 1X8

Edited by Adriaan Bax, National Institutes of Health, Bethesda, MD, and approved May 8, 2014 (received for review March 17, 2014)

Hydrogen exchange rates have become a valuable probe for studying the relationship between dynamics and structure and for dissecting the mechanism by which proteins fold to their native conformation. Typically measured rates correspond to averages over all protein states from which hydrogen exchange can occur. Here we describe a new NMR experiment based on chemical exchange saturation transfer that provides an avenue for obtaining uncontaminated, per-residue amide hydrogen exchange rates for interconverting native and invisible states so long as they can be separated on the basis of distinct ^{15}N chemical shifts. The approach is applied to the folding reaction of the Fyn SH3 domain that exchanges between a highly populated, NMR-visible native state and a conformationally excited, NMR-invisible state, corresponding to the unfolded ensemble. Excellent agreement between experimentally derived hydrogen exchange rates of the excited state at a pair of pHs is obtained, taking into account the expected dependence of exchange on pH. Extracted rates for the unfolded ensemble have been used to test hydrogen exchange predictions based on the primary protein sequence that are used in many analyses of solvent exchange rates, with a Pearson correlation coefficient of 0.84 obtained.

amide exchange with solvent | conformationally excited protein states | protein folding

The energy landscape of a protein is a multidimensional surface composed of many local minima in addition to the global minimum that is the native conformation (1, 2). An understanding of the relation between protein structure, dynamics, and function is predicated, therefore, on an analysis of the various conformational states that populate the minima on the landscape. This requires detailed structural and dynamics studies of each of the conformers and quantification of their relative energies as well as the kinetics of exchange between them. Biophysical techniques such as X-ray diffraction and NMR spectroscopy are available for obtaining detailed structural information on the molecules populating the lowest-energy regions of the landscape, providing insight into the structure–function paradigm for a great number of proteins. However, it is becoming increasingly well understood that studies of the ground states of proteins are not sufficient. Additional states that can be sparsely populated and transiently formed, referred to as excited states in what follows, are often important for processes that include molecular recognition, ligand binding, enzyme catalysis, and folding (3–9). Detailed studies of such excited states are, however, challenging because they are not "visible" to standard biophysical methods and as a consequence atomic resolution information is lacking.

Recent developments in solution NMR spectroscopy are changing this paradigm by providing an avenue for quantifying excited states at a level of detail that has typically only been possible in studies of highly populated native protein conformers (10). Backbone ^{1}H, ^{13}C, and ^{15}N chemical shifts for invisible protein states can now be measured so long as they are populated at levels of 0.5% or higher and exchange with an NMR visible state with rates on the order of approximately 100 to several thousand per second (10–12). These chemical shifts have, in turn, been used along with database computational approaches to generate atomic resolution models of excited states (7, 13, 14), providing detailed insights into a range of biochemical processes. It is now possible in some cases to measure scalar and residual dipolar couplings that, in turn, provide additional structural insights (15). Further advances have led to the measurement of side-chain ^{13}C and ^{1}H chemical shifts in invisible states (16–18) that can be sensitive to hydrophobic contacts in these rare conformers. Experiments for quantifying pico- to nanosecond time-scale side-chain dynamics have also emerged (19), showing in some cases large differences in motion between ground and excited states that directly relate to function (20).

Although the tool kit of NMR experiments for studying rare protein conformers is expanding, it remains significantly smaller than that for highly populated states, and there is a continuing need for the development of further methodologies. Notably absent from the tool kit is an approach for measuring hydrogen exchange rates in excited-state protein conformers. Ever since the pioneering work of Linderstrøm-Lang in the 1950s it has been recognized that the rates of exchange of amide hydrogens with solvent protons are important site-specific parameters of protein structure and dynamics (21). More recently, amide hydrogen exchange measurements have been used to investigate regions of local protein disorder, folding/unfolding processes, hydrogen bonding, allostery, and ligand binding (22–24). NMR spectroscopy has emerged as a primary tool for quantifying amide hydrogen exchange rates, $k_{H\text{-}EX}$, on a per-residue basis. However, because the overall transfer of hydrogens from solvent is measured,

Significance

Site-specific rates of exchange of amide hydrogens with solvent protons are important parameters for characterizing protein structure and dynamics, providing insight into a range of biomolecular processes. In typical hydrogen exchange experiments, measured rates reflect exchange from all accessible states of a protein. Here we describe an NMR approach for measuring exchange rates in sparsely populated, transiently formed conformational states that interconvert with a highly populated and long-lived ground state. An application to a protein folding reaction of the Fyn SH3 domain is provided to illustrate the methodology, establishing that robust measures of hydrogen exchange are obtained both for the ground, native state and the invisible, unfolded state that are uncontaminated by the folding exchange process.

Author contributions: D.L. and L.E.K. designed research; D.L. and L.E.K. performed research; G.B. contributed new reagents/analytic tools; D.L. analyzed data; and D.L. and L.E.K. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]Present address: Institut de Biologie Structurale, Université de Grenoble Alpes, CEA, CNRS, UMR5075, 38000 Grenoble, France.

[2]To whom correspondence should be addressed. E-mail: kay@pound.med.utoronto.ca.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1405011111/-/DCSupplemental.

the extracted rates often contain contributions from all of the exchange-accessible states of the protein (25–27). This is the case for H–D exchange experiments where time-dependent intensity changes of NMR signals are measured to extract $k_{H\text{-}EX}$ (22), as well as experiments that perturb the water signal and quantify the perturbation at exchangeable amide sites (26). Similar bulk measures of $k_{H\text{-}EX}$ are obtained using other methods as well (28), from which insight into a particular excited state can often only be inferred. Herein we describe an approach by which site-specific hydrogen exchange rates of amides in individual states along the energy landscape can be obtained for a protein system at equilibrium and under native conditions. An application to the folding reaction of the G48A Fyn SH3 domain is provided, where $k_{H\text{-}EX}$ values of the native conformation (ground state) and an excited state (unfolded ensemble) are measured. Because exchange rates are quantified by measurement of intensities of cross-peaks in NMR spectra of the ground state, the method is sensitive to $k_{H\text{-}EX}$ values as large as several hundred per second for the excited state. The measurement of "pure" $k_{H\text{-}EX}$ values provides a powerful new approach for understanding the dynamical properties of excited protein states.

Results

Description of the Method. The approach for obtaining "pure" $k_{H\text{-}EX}$ values can be understood by first considering an analogy with the well-known dependence of NMR lineshapes on the chemical exchange rate (11) (Fig. 1*A*). Consider a two-site exchanging system, $G \underset{k_{EG}}{\overset{k_{GE}}{\rightleftharpoons}} E$, where G and E denote the interconverting ground and excited states, respectively, $k_{ex} = k_{EG} + k_{GE}$ is the rate of exchange between the two, and the fractional population of each conformer is assumed equal, $p_G = p_E = 0.5$. In Fig. 1*A* the spectrum derived from a spin exchanging between two sites is highlighted, with a chemical shift difference (radians per second) for the spin in each state of $\Delta\omega_{EG} = \omega_E - \omega_G$. For $|k_{ex}/\Delta\omega_{EG}| << 1$ a pair of peaks is observed at the resonance frequencies of the spins in the absence of exchange. As the exchange rate increases the peaks move closer and for $|k_{ex}/\Delta\omega_{EG}| \sim 1$ they converge into a single broad resonance line that then becomes sharper as $|k_{ex}/\Delta\omega_{EG}|$ increases. A similar scenario holds when considering the relative position of ^{15}N doublet components in the case of a two-spin ^{15}N-^{1}HN spin system, as a function of the hydrogen exchange rate, that is germane for studies involving backbone amide groups considered here. In this case each component derives from a different ^{1}HN spin state (α or β) with a net separation of J_{HN} (hertz), where J_{HN} is the one-bond ^{15}N-^{1}HN scalar coupling constant. Analogous to the example of Fig. 1*A*, as $k_{H\text{-}EX}$ becomes on the order of $2\pi|J_{HN}|$ the multiplet components converge into a single broad peak that then sharpens as the exchange rate increases, resulting in a single decoupled line. Bodenhausen and coworkers have exploited this effect to quantify rapid hydrogen exchange rates in proteins (29). The situation is more complex in the case of the exchanging systems considered in this work because the populations of the interconverting states are highly skewed, $p_E << p_G$, so that it is not possible to directly observe and hence quantify the positions of multiplet components for state E that in principle could provide the desired exchange rates.

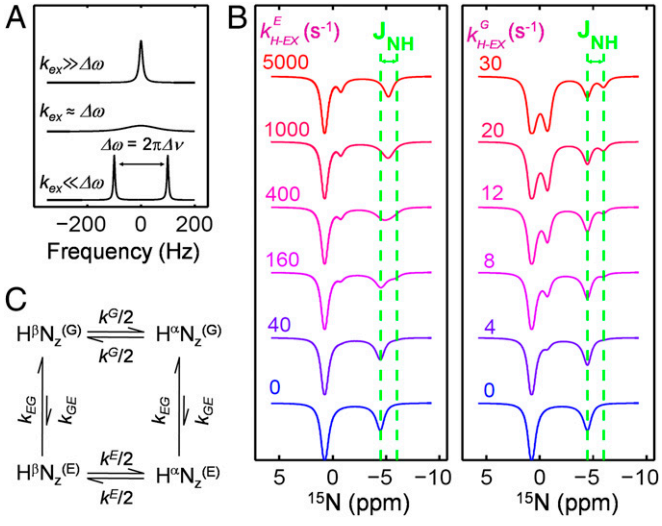

Fig. 1. Measurement of hydrogen exchange rates in invisible protein states. (*A*) Influence of $k_{ex}/\Delta\omega_{EG}$ on NMR lineshapes for a two-site chemical exchange process. Parameters used in the calculation: $p_G = p_E = 0.5$; $\Delta\omega_{EG}/(2\pi) = 200$ Hz; R_2 (transverse relaxation rate) = 7 s^{-1}; and $k_{ex} = 30$ (slow exchange), 1,200 (intermediate), and 10,000 (fast) s^{-1}. (*B*) Simulated TROSY CEST profiles, based on the pulse scheme of *SI Appendix*, Fig. S1, as a function of $k^E_{H\text{-}EX}$ from 0 to 5,000 s^{-1}, $\rho_{EXT} = k^G = 0$ s^{-1} (*Left*) or as a function of $k^G_{H\text{-}EX}$ from 0 to 30 s^{-1}, $\rho_{EXT} = k^E = 0$ s^{-1} (*Right*), where ρ_{EXT} is the contribution to relaxation from proximal protons (discussed in the text). Values of $k_{ex} = 120$ s^{-1}, $p_E = 5\%$, $T_{EX} = 300$ ms, $B_1 = 14.4$ Hz, $\Delta\omega_{EG} = -2{,}000$ rad/s, $B_o = 14.0$ T were used in the simulations, along with identical motional parameters for G and E states of $\tau_e = 10$ ps, $S^2 = 0.75$, $\tau_C = 5$ ns (see *SI Appendix* for details). (*C*) Magnetization pathway for TROSY and anti-TROSY ^{15}N components, including the effects of hydrogen exchange (k^G and k^E) and chemical exchange (k_{GE} and k_{EG}).

The spectrum of the excited state can, however, be recorded indirectly through the use of chemical exchange saturation transfer (CEST) (30, 31). Here a set of 2D ^{15}N-^{1}HN spectra are obtained whereby a weak ^{15}N rf field is applied during a delay when the magnetization of interest (^{15}N) is longitudinal. At each position of the rf field a separate spectrum is recorded and the intensity of each correlation of the ground state, I, plotted as a function of rf offset, normalized to the intensity in the absence of the field, I_b. If the field is applied at the resonance position of a spin in the ground state then I decreases to zero, owing to a well-known saturation effect. Similarly, when the rf field is proximal to the resonance position of a nucleus in state E the perturbation is transferred to the corresponding spin in G via chemical exchange, leading to a decrease in I again. Thus, a pair of dips is observed in a typical CEST profile, with the largest (smallest) at the resonance position of the spin in the ground (excited) state. Because the intensities of ground-state correlations are measured, the effect is to amplify peaks from the excited state so that they become observable, leading to a CEST profile that connects ground and excited state correlations.

SI Appendix, Fig. S1 illustrates the NMR experiment that has been used to measure amide hydrogen exchange rates in ground and invisible excited protein states. The details are discussed in *SI Appendix*. Briefly, the flow of magnetization can be summarized as follows:

$$H_Z \rightarrow H^\beta N_Z(m) \overset{CEST, T_{EX}}{\rightarrow} H^\beta N_Z(G) \rightarrow H^\beta N_{X,Y}(G)(t_1 acq) \rightarrow H_{X,Y} N^\beta(G)(t_2 acq), \quad [1]$$

where $H^\beta N_Z(m)$, $m \in (G,E)$, denotes ^{15}N longitudinal magnetization from either G or E that is associated with the amide proton β spin state and $H^\beta N_{X,Y}$ indicates transverse relaxation optimized (TROSY) ^{15}N magnetization. Thus, ^{1}HN polarization is transferred to $H^\beta N_Z$ (referred to as TROSY ^{15}N longitudinal magnetization in what follows) that is then allowed to evolve for a period T_{EX} during which a weak ^{15}N rf field is applied. At the conclusion of the CEST period ^{15}N and ^{1}HN chemical shifts of the ground (visible) state are recorded during t_1 and t_2 periods, respectively, to produce a standard 2D ^{15}N-^{1}HN TROSY correlation plot (32). As described above, a series of such spectra is obtained with the position of the rf field moved for each dataset and the intensities of individual correlations plotted as a function of rf position to generate CEST profiles.

Fig. 1*B* plots simulated ^{15}N CEST profiles obtained for a single two-spin ^{15}N-^{1}HN spin system that exchanges between G and

E with k_{ex} = 120 s^{-1}, p_E = 5%, as a function of $k_{H\text{-}EX}$ (k^G_{H-EX}, right; k^E_{H-EX}, left). The experimental scheme summarized above, Eq. **1**, (*SI Appendix*, Fig. S1) has been used. In the case where $k^G_{H-EX} = k^E_{H-EX} \sim 0$ a pair of peaks is recorded for each amide ^{15}N, centered at ϖ^{TR}_E and ϖ^{TR}_G (parts per million; bottom traces of Fig. 1*B*), that correspond to the positions of the TROSY peaks for the two chemically exchanging states. As k^E_{H-EX} increases so that $\frac{k^E_{H-EX}}{2\pi|J_{HN}|} \sim 0.2$, small dips are observed shifted by $|J_{HN}|$ Hz upfield (to the right in the diagram) of the main correlations, at frequencies of ϖ^{A-TR}_E and ϖ^{A-TR}_G, that are the anti-TROSY components (α ^{1}HN spin state). As k^E_{H-EX} increases even further such that $k^E_{H-EX} \sim 2\pi|J_{HN}|$ the TROSY and anti-TROSY lines from state E coalesce to form a broad dip that then narrows as k^E_{H-EX} increases further. The corresponding pattern observed as a function of k^G_{H-EX} when $k^E_{H-EX} = 0$ is illustrated on the right. Here much smaller exchange rates are considered because for large k^G_{H-EX} the ground-state signal that is detected in the experiment decreases significantly.

The profiles of Fig. 1*B* can be understood using the "exchange diagram" of Fig. 1*C* that depicts the flow of magnetization during the CEST interval, T_{EX}. Hydrogen exchange with solvent protons (k^m_{H-EX}) or dipolar relaxation with proton spins that are proximal to the amide proton in question (ρ_{EXT}) lead to the interconversion of the ^{1}HN spin state ($\alpha \leftrightarrow \beta$) according to (33)

$$\frac{dH^\beta N_Z(m)}{dt} \approx -\left(\rho_{TR} + \frac{k^m}{2}\right)H^\beta N_Z(m) + \frac{k^m}{2}H^\alpha N_Z(m)$$
$$\frac{dH^\alpha N_Z(m)}{dt} \approx -\left(\rho_{A-TR} + \frac{k^m}{2}\right)H^\alpha N_Z(m) + \frac{k^m}{2}H^\beta N_Z(m), \quad [2]$$

where ρ_{TR} and $\rho_{A\text{-}TR}$ are the self relaxation rates of the ^{15}N TROSY ($H^\beta N_Z$) and anti-TROSY ($H^\alpha N_Z$) longitudinal components, respectively, $k^m = k^m_{H-EX} + \rho^m_{EXT}$, and equilibrium magnetization is neglected for simplicity (*Materials and Methods*). It can be readily seen from Eq. **2** that $H^\beta N_Z$(m) and $H^\alpha N_Z(m)$ interconvert; thus, starting from the initial conditions of $H^\beta N_Z(m) \neq 0$ and $H^\alpha N_Z(m) = 0$ a buildup of the anti-TROSY component occurs that depends on the size of k^m, as observed in the simulations of Fig. 1*B*. Both the buildup of anti-TROSY magnetization and the resultant ^{15}N lineshapes are used to extract accurate k^m values in fits of CEST profiles (discussed below). In addition to the magnetization transfer pathway with rate k^m, denoted by the horizontal sides of the square in Fig. 1*C*, a second pathway (vertical sides) interconverts $H^j N_Z(G)$ and $H^j N_Z(E)$ owing to chemical exchange. Magnetization transfer pathways that give rise to each of the dips in Fig. 1*B* can be obtained from Fig. 1*C*, recalling that only magnetization of the form $H^\beta N_Z(G)$ is detected at the end of the CEST period (Eq. **1**). For example, a dip observed at the position of the anti-TROSY component of the spin in state E arises from both of the pathways, $H^\alpha N_Z^{(E)} \overset{k^E}{\leftrightarrow} H^\beta N_Z^{(E)} \overset{k_{EG}}{\leftrightarrow} H^\beta N_Z^{(G)}$ and $H^\alpha N_Z^{(E)} \overset{k_{EG}}{\leftrightarrow} H^\alpha N_Z^{(G)} \overset{k^G}{\leftrightarrow} H^\beta N_Z^{(G)}$.

The magnetization transfer diagram of Fig. 1*C* makes it clear that CEST profiles can depend on hydrogen exchange rates in both G and E states. We were interested, therefore, in ensuring that fitted values of k^G and k^E are not correlated. Extensive numerical simulations (*SI Appendix*, Fig. S2) and experimental studies (discussed below) establish that accurate rates can be obtained. By means of example, Fig. 2*A*, *Left* shows a synthetic ^{15}N-TROSY CEST profile calculated with k^E = 150 s^{-1}, k^G = 20 s^{-1}, and with random error added in the form of Gaussian noise with an SD of 0.02I_b, similar to what we have obtained experimentally. The resultant 2D surface, $y = \exp(-\chi^2/\chi^2_{min})$ (Fig. 2*A*, *Right*), clearly shows a maximum at the expected position (k^E = 156 s^{-1}, k^G = 19.7 s^{-1}) and very little correlation between the extracted rates. The simulations of Fig. 1*B*, *Left* show that ^{15}N

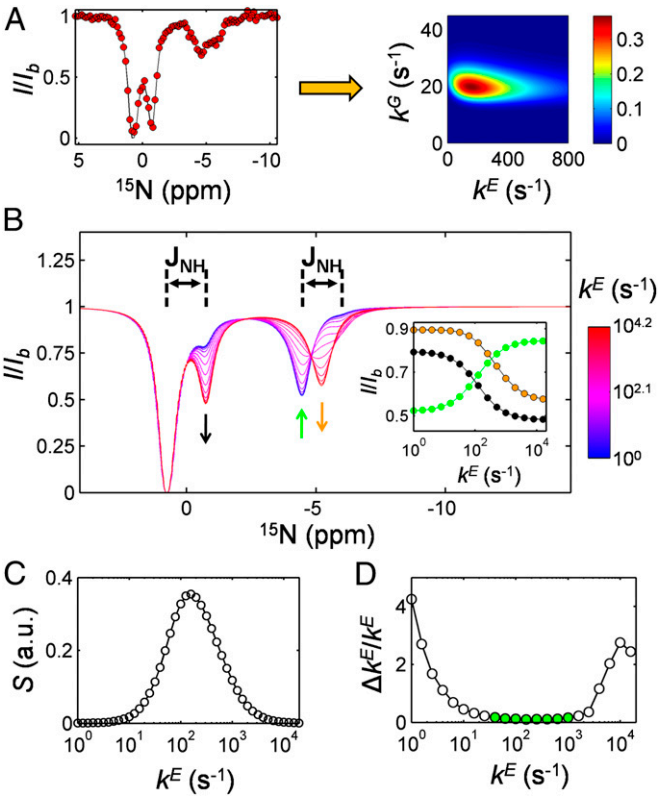

Fig. 2. Robustness of extracted k^E values. (*A*) (*Left*) Simulated ^{15}N TROSY CEST profile (red circles) using the same exchange parameters as in the legend to Fig. 1*B*, with the exception of k^E = 150 s^{-1} and k^G = 20 s^{-1}, along with added Gaussian noise with an SD of 0.02I_b. The profile is fit (black solid line) using the Bloch–McConnell equations (*SI Appendix*) to extract k^G and k^E. (*Right*) Two-dimensional k^G vs. k^E error surface, $y = \exp(-\chi^2/\chi^2_{min})$, with a maximum at (156 s^{-1}, 19.7 s^{-1}). (*B*) Superposition of ^{15}N TROSY CEST profiles simulated for values of k^E ranging from 1–$10^{4.2}$ s^{-1}, k^G = 4 s^{-1}, and other parameters as in Fig. 1*B*, along with the change in intensities of dips at $\varpi_E = \varpi^{TR}_E + \varpi^{A-TR}_E/2$ (orange), ϖ^{TR}_E (green), and ϖ^{A-TR}_G (black) as a function of increasing k^E (*Inset*). (*C*) Sensitivity of the CEST profile calculated using the parameters of Fig. 2*B* to k^E, defined in terms of the parameter S (Eq. **3**) as described in the text. (*D*) Relative uncertainties in k^E based on Monte Carlo simulations assuming experimental noise of 0.02I_b, with Δk^E defined as one SD of the extracted rates. Green circles correspond to those values of k^E for which fractional errors less than 20% are obtained. Other than k^E the parameters of Fig. 2*B* were used in the computations.

TROSY CEST profiles are relatively insensitive to k^E for values less than ~40 s^{-1} and greater than ~1,000 s^{-1}. By contrast, the profiles change significantly as a function of k^G over the range of values considered, 0–30 s^{-1} (Fig. 1*B*, *Right*). That such large values of k^E can be measured reflects the fact that the observed signal derives from ground-state correlations. By contrast, as k^G increases significantly the overall sensitivity of the experiment becomes adversely affected, as is the case in any amide-detection experiment as exchange rates with water become large.

A further set of simulations was performed to establish the range of k^E values that can be optimally detected with this technique, Fig. 2 *B–D*, assuming k_{ex} = 120 s^{-1}, p_E = 5% (see figure legend for additional details). Fig. 2*B* plots a superposition of ^{15}N TROSY CEST profiles simulated for values of k^E ranging from 1–$10^{4.2}$ s^{-1}. The directions of the arrows at the frequencies $\varpi_E = \frac{\varpi^{TR}_E + \varpi^{A-TR}_E}{2}$ (orange), ϖ^{TR}_E(green), and ϖ^{A-TR}_G(black) indicate how the intensities of the corresponding dips centered at these positions vary as a function of increasing k^E, with intensity vs. k^E plotted in the inset. It is clear that the methodology is maximally

sensitive to k^E values that are centered at ~100 s^{-1}. To quantify this further we have calculated

$$S = \sum_{\nu} \left[\frac{f_{CEST}(\nu, q^E + \varepsilon) - f_{CEST}(\nu, q^E - \varepsilon)}{2\varepsilon} \right]^2 \approx \sum_{\nu} \left[\frac{\partial f_{CEST}}{\partial q^E} \right]^2, \quad [3]$$

which provides a measure of the sensitivity of a CEST profile to the fractional change in k^E. In Eq. **3** the summation is over all ^{15}N frequencies, ν, in the CEST profile (i.e., each point of the CEST curve), $f_{CEST}(\nu)$ is the value of the profile at ν, $q^E = \log_{10}(k^E)$ and 2ε corresponds to a small increment in q^E. Fig. 2*C* plots S vs. k^E, with high S values (optimal sensitivity) over the range 40 $s^{-1} < k^E < 1{,}000$ s^{-1}. Finally, the fractional error in extracted k^E, $\Delta k^E/k^E$, has been calculated as a function of k^E, Fig. 2*D*. For k^E values between 40 and 1,000 s^{-1} (corresponding to green circles) fractional errors in extracted rates are under 20%. In absolute terms the errors remain relatively small for $k^E < 40$ s^{-1} so that the distinction between residues that are protected from exchange with solvent and those that are not remains clear when exchange rates are rapid, although quantitative comparison between small k^E rates is likely not feasible. It is worth noting that whereas relative and absolute measures of error will depend on experimental signal-to-noise as well as the chemical exchange parameters, further simulations show that the ^{15}N CEST approach remains most sensitive for values of k^E between several tens to 1,000 s^{-1}, over a range of chemical exchange parameters.

An Application to a Protein Folding Reaction. The Fyn SH3 domain is a small 60-residue module that plays an important role in protein recognition involved in a range of signaling events (34). The G48A mutant has been shown previously to interconvert on the millisecond timescale between a long-lived, highly populated native state and a transiently formed, sparsely populated unfolded ensemble, the latter invisible to traditional NMR experiments (35). Fig. 3 shows ^{15}N TROSY CEST traces of four residues recorded at pH 7.7 and 8.4, where small dips corresponding to the excited state are clearly observed. The expected approximately fivefold increase in solvent exchange rates in the excited state from pH 7.7 to 8.4 leads to a pH-dependent change in the dips derived from the sparsely populated conformer. For example, at pH 7.7 an anti-TROSY dip is not observed for Leu7 of the excited state; the dip is apparent, however, at pH 8.4, which reflects the increase in k^E and hence the more rapid flux through the pathway $H^{\alpha}N_Z^{(E)} \overset{k^E}{\leftrightarrow} H^{\beta}N_Z^{(E)} \overset{k_{EG}}{\leftrightarrow} H^{\beta}N_Z^{(G)}$. Asp9 and Glu11 have profiles similar to that of Leu7; however, the anti-TROSY dips of Asp9 are larger owing to the approximate factor of two increase in k^E for this residue. Notably, Ser52 shows an anti-TROSY dip already at pH 7.7, whereas at pH 8.4 a "decoupled" dip that results from rapid exchange is observed.

All TROSY CEST profiles have been fitted using the Bloch–McConnell equations (36) for a two-spin system exchanging between a pair of magnetically distinct sites (*SI Appendix*). Values of p_E, k_{ex}, and $\Delta\varpi_{EG}$ obtained from ^{1}HN-decoupled ^{15}N CEST profiles (31), which are not sensitive to solvent exchange, have been used as fixed parameters in the analysis. Examples of data fits are shown in Fig. 3 (solid lines) that are typical for the present study, and the excellent agreement with experimental data (circles) is clear.

As described above, values of $k^m = k^m_{H-EX} + \rho^m_{EXT}$, $m \in (G,E)$, are fitted from CEST profiles that include contributions from ^{1}H cross-relaxation with proximal protons. We have therefore measured site-specific values for ρ^G_{EXT} at pH 5.5, where exchange with solvent is minimal (*SI Appendix*). Values of ρ^G_{EXT} are, on average, 10- and 50-fold smaller than k^E rates at pH 7.7 and 8.4,

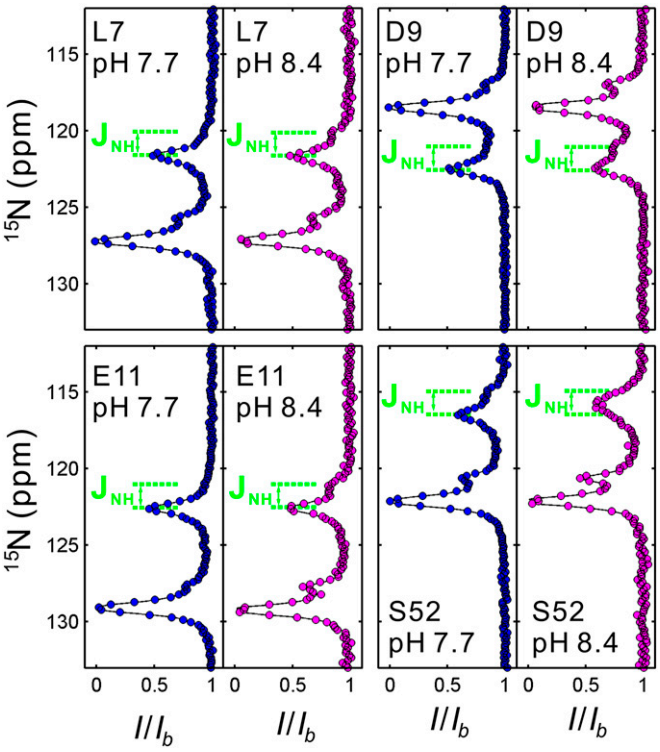

Fig. 3. ^{15}N TROSY CEST profiles of selected residues measured at pH 7.7 (blue) and pH 8.4 (magenta). Extracted k^E values (s^{-1}) for Leu7, Asp9, Glu11, and Ser52 at pH 7.7 and 8.4 are (20 ± 4, 75 ± 6), (35 ± 2, 146 ± 8), (14 ± 2, 62 ± 7), and (107 ± 6, 507 ± 41), respectively.

with ρ^G_{EXT} expected to be larger than ρ^E_{EXT}, so that to excellent approximation $k^E \approx k^E_{H-EX}$ (*SI Appendix*, Table S1). In cases where the effects of cross-relaxation are expected to be larger, they can be minimized by using highly deuterated proteins, which also has the benefit of improving the sensitivity of the experiment. Additionally, separation of exchange and dipolar relaxation contributions can be achieved from analysis of datasets recorded at two or more pH values, assuming that the structures of the interconverting states do not change with pH (27). Fig. 4*A* plots $k^{(E)}_{H-EX}$ values at pH 7.7 that have been scaled by a factor $\alpha = 10^{(8.4-7.7)}$ ($\alpha\, k^{E,pH=7.7}_{H-EX}$) vs. $k^{(E)}_{H-EX}$ values measured at pH 8.4. The value of α takes into account the pH dependence of exchange and is in excellent agreement with the best-fit value of $10^{0.67}$ that minimizes the differences in measured hydrogen exchange rates at the two pHs. In the analysis of the ^{15}N TROSY CEST profiles we have fit all residues, irrespective of the magnitude of the $\Delta\varpi_{EG}$ value. It is clear that for small $\Delta\varpi_{EG}$ values, where the overlap between ground and excited state dips is extensive, the accuracy of the extracted k^m values will be compromised because lineshape fitting becomes difficult. We were interested in establishing an approximate minimum $\Delta\varpi_{EG}$ value for which this occurs. In Fig. 4*A* $k^{(E)}_{H-EX}$ values are color-coded according to $\Delta\varpi_{EG}$ and the outliers correspond to shift differences of less than 1 ppm. Further, experiment (Fig. 4*A*, blue circles) and simulations (*SI Appendix*, Fig. S3) establish that for $\Delta\varpi_{EG} \geq 1$ ppm robust measures of exchange rates are obtained from analysis of CEST profiles recorded at a static magnetic field strength of 14 T. It is anticipated that this threshold (1 ppm) will decrease as experiments are recorded at higher magnetic fields. Fig. 4*B* plots $\alpha\, k^{E,pH=7.7}_{H-EX}$ (blue) and $k^{E,pH=8.4}_{H-EX}$ (red) as a function of residue for the excited state, including only residues for which $\Delta\varpi_{EG} \geq 1$ ppm. Similar plots for $k^G_{H-EX} (= k^G - \rho^G_{EXT})$ are illustrated in Fig. 4 *C* and *D* as well. Overall, the level of agreement between the two

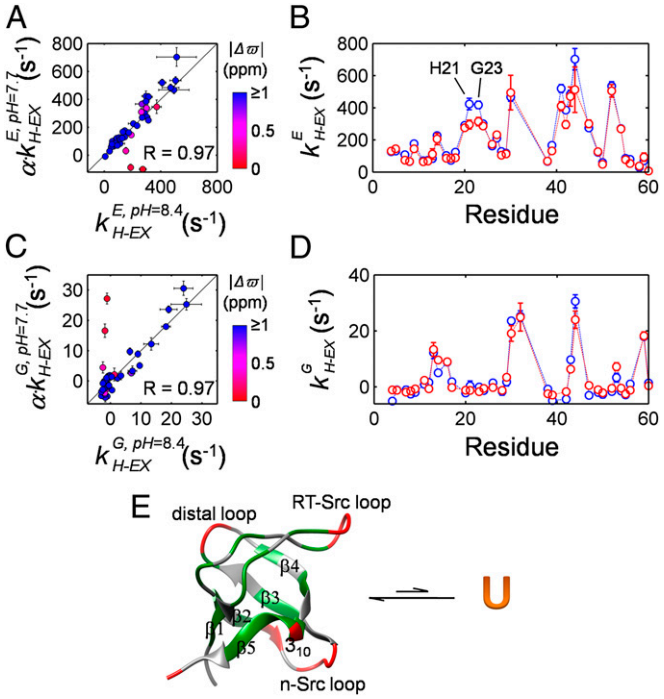

Fig. 4. Measuring separate hydrogen exchange rates for the native state and the unfolded ensemble of the Fyn SH3 domain. (*A* and *C*) Comparison of site-specific hydrogen exchange rates for residues in *E* (*A*) and *G* (*C*) at pH 7.7 and 8.4, with α a scaling factor based on the pH difference, $\alpha = 10^{0.7}$. Rates for residues with $|\Delta\varpi_{EG}| < 1$ ppm are denoted using spheres that are colored from red to blue as $|\Delta\varpi_{EG}|$ increases, as illustrated by the color bar. All data with $|\Delta\varpi_{EG}| \geq 1$ ppm are shown in blue. Pearson correlation coefficients are shown, calculated for residues with $|\Delta\varpi_{EG}| \geq 1$ ppm. (*B* and *D*) Comparison of $\alpha \cdot k_{H-EX}^{E(G),pH=7.7}$ (blue circles) and $k_{H-EX}^{E(G),pH=8.4}$(red circles) as a function of residue. Only data from residues with $|\Delta\varpi_{EG}| \geq 1$ are shown. (*E*) Mapping of residues with $k_{H-EX}^{G,pH=8.4} > 3s^{-1}$ (red) on the 3D structure of the Fyn SH3 domain (7) (PDB ID code 2LP5). Backbone positions of residues Glu15, Glu33, Gly34, and Asp35 that have large $k_{H-EX}^{G,pH=7.7}$ rates (>9 s^{-1}) but whose cross-peaks disappear at pH 8.4 are also colored in red. Colored in green are those residues for which $k_{H-EX}^{G,pH=8.4} < 3$ s^{-1}, and gray highlights residues whose rates could not be quantified or $|\Delta\varpi_{EG}| < 1$ ppm.

measurements is excellent, with the close similarity between calculated and predicted α values based on the pH difference between the two samples providing a further cross-validation. Notably, α $k_{H-EX}^{E,pH=7.7}$ and $k_{H-EX}^{E,pH=8.4}$ deviate somewhat for H21, K22, and G23 (Fig. 4*B* and *SI Appendix*, Table S1), which may reflect the fact that the titration of the sidechain of H21, the lone histidine residue in the protein, is not complete by pH 7.7, potentially affecting backbone hydrogen exchange rates in the vicinity of this residue. Fig. 4*E* shows a ribbon diagram of the structure of the Fyn SH3 domain highlighting regions that are protected from exchange based on extracted k_{H-EX}^{G} values at pH 8.4 (green) and those that are exposed (red). Interestingly, many residues in the long RT-Src loop are protected, consistent with a well-defined structure for the loop that is further confirmed by high backbone order parameters (37) that report on the amplitudes of motion.

Discussion

The hydrogen exchange method, originally conceived by Linderstrøm-Lang over 50 y ago, has become popular in protein science for addressing important questions relating structure with function and for probing the mechanism of protein folding (38). Traditional hydrogen exchange measurements often report bulk exchange rates with solvent that include contributions from all accessible states of a protein (25). Thus, in the case where the exchange between states, j, is fast compared with hydrogen exchange rates in j, k_{H-EX}^{j}, $k_{H-EX} \approx \sum_j f_j k_{H-EX}^{j}$ is measured for the exchange rate, as probed via some property of the native state (such as peak intensities in NMR or mass in mass spectrometry studies), where f_j is the fractional population of state j. Although under native conditions the population of the native state dominates, other more sparsely populated conformers can contribute significantly to k_{H-EX} because they are significantly less protected from hydrogen exchange. The relative contributions from each state j can be manipulated via changes to f_j through the addition of denaturants, as is typically done in native-state hydrogen exchange experiments (24, 38, 39). These measurements have been shown to be powerful for detecting and characterizing folding intermediates and have contributed greatly to our understanding of the mechanism of protein folding. However, averages of rates over all accessible states are still obtained for each condition examined, and it is often difficult to obtain parameters that report on the specific properties of each of the interconverting states.

Herein we describe a ^{15}N TROSY CEST-based NMR experiment that separates exchange rates from states that can be distinguished on the basis of differences in chemical shifts. The advantages of using this type of experiment over a non-TROSY CEST-based NMR scheme are discussed in *SI Appendix*, Fig. S4. By means of example, we consider a G48A Fyn SH3 domain that exchanges between a folded conformation and an unfolded ensemble, with the latter too sparsely populated to be detected by traditional NMR methods. However, because separate ^{15}N chemical shifts can be measured for the native and unfolded states in CEST experiments, k_{H-EX} values can be obtained on a site-specific basis for each. It is worth noting that in the case where a "state" corresponds to an average over multiple conformers that interconvert rapidly on the NMR chemical shift timescale the measured k_{H-EX} values would correspond to a weighted average over exchange rates in the ensemble.

Measured hydrogen exchange rates are often recast in terms of site-specific protection factors (PF), where $PF = k_{\rm int}/k_{H-EX}$ and k_{int} is the predicted intrinsic exchange rate that takes into account the primary amino acid sequence and the conditions of the experiment. A simple hydrogen exchange scheme originally pictured by Linderstrøm-Lang (21), $NH(closed) \underset{k_{cl}}{\overset{k_{op}}{\rightleftharpoons}} NH(open) \overset{k_{\rm int}}{\rightarrow} exchange$, is one occurring via a two-step reaction in which the amide site undergoes a closed to open transition, with exchange only in the open state. When the exchange reaction is carried out under conditions such that $k_{op} << k_{cl}$ and where $k_{int} << k_{cl}$ the ΔG for the opening reaction is given by $\Delta G = -RT\ln(1/PF)$, which depends on both measured k_{H-EX} rates and predicted k_{int} values (39). The measured site-specific values of k_{H-EX}^{E} from the unfolded ensemble of the G48A Fyn SH3 domain allow us to test predicted k_{int} values and hence provide an independent estimate of the accuracy of literature ΔG values from measured hydrogen exchange rates. Fig. 5 compares extracted k_{H-EX}^{E} values at pH 7.7 with predicted hydrogen exchange rates of the unfolded G48A Fyn SH3 domain computed using the program SPHERE (40, 41) (www.fccc.edu/research/labs/roder/sphere/). The predicted rates have been scaled by a single multiplicative factor (0.6) to best match experiment and the level of agreement is reasonably high ($R = 0.84$). Predicted values (without scaling) are approximately a factor of two higher than those measured, corresponding to an error in ΔG of ~0.4 kcal/mol. Although the difference could be explained by the fact that there is residual structure in the unfolded ensemble that provides a small level of protection from hydrogen exchange, the excellent agreement between measured chemical shifts and those predicted assuming an unfolded random coil-like structure (16) argues against this interpretation. It may be that some of the difference results from compaction of the unfolded Fyn SH3 domain, which leads to

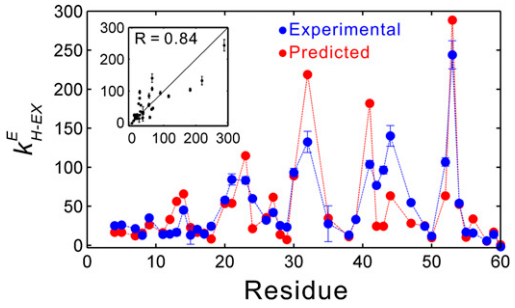

Fig. 5. Comparison of experimental (blue) and predicted (red) hydrogen exchange rates for residues of the unfolded ensemble of the G48A Fyn SH3 domain ($|\Delta\varpi_{EG}| \geq 1$ ppm). (*Inset*) Scatter plot of the experimental rates vs. predicted rates along with the Pearson correlation coefficient. Hydrogen exchange rates were predicted by the program SPHERE (www.fccc.edu/research/labs/roder/sphere/), which has an accuracy within about a factor of 2 for unfolded polypeptides (38, 40, 41). Predicted rates were uniformly scaled to best match the experimental k^E values.

a small attenuation of exchange rates over what is predicted from a random coil.

In summary, we have presented a CEST-based NMR approach for quantifying hydrogen exchange rates in invisible excited protein states, so long as they have ^{15}N chemical shifts that are distinct from the highly populated native state. The predicted exponential scaling of k_{H-EX} with pH has been observed both for the native and unfolded ensembles of the Fyn SH3 domain, and a very high level of agreement is obtained between hydrogen exchange rates at pH 7.7 and 8.4 when scaled for the pH difference. A series of simulations shows that the method is most sensitive to k^E_{H-EX} rates ranging from ~40 s^{-1} to 1,000 s^{-1} and as with other CEST experiments for k_{ex} between ~50–500 s^{-1} (16, 31). Notably, the approach provides an avenue for obtaining pure measures of hydrogen exchange rather than bulk rates that are typically recorded in most experiments. It is anticipated that the methodology will be particularly useful in studies of excursions from native-state conformations that often involve formation of transient, partially folded states (7, 14). As such it is a valuable addition to the toolbox of NMR experiments for studying sparsely populated protein conformers whether they be involved in protein folding or in other important biological processes.

Materials and Methods

^{15}N-labeled G48A *Gallus gallus* Fyn SH3 domain was expressed and purified as described before (16). The NMR sample comprised ~1.5 mM protein dissolved in 50 mM sodium phosphate, 0.2 mM EDTA, and 10% D_2O/90% H_2O (vol/vol). To measure pH-dependent changes of hydrogen exchange rates the pH of the NMR sample was adjusted to either 5.65, 7.65, or 8.38 by buffer exchange, with the same salt concentrations in each case. Experimental details and analysis of CEST profiles (including relevant equations) are provided in *SI Appendix*. The software used for data analysis is available from the authors upon request.

ACKNOWLEDGMENTS. D.L. is the recipient of a postdoctoral fellowship from the Canadian Institutes of Health Research (CIHR). This work was funded through a CIHR grant (to L.E.K.). L.E.K. holds a Canada Research Chair in Biochemistry.

1. Dill KA, Chan HS (1997) From Levinthal to pathways to funnels. *Nat Struct Biol* 4(1):10–19.
2. Frauenfelder H, Sligar SG, Wolynes PG (1991) The energy landscapes and motions of proteins. *Science* 254(5038):1598–1603.
3. Eisenmesser EZ, et al. (2005) Intrinsic dynamics of an enzyme underlies catalysis. *Nature* 438(7064):117–121.
4. Mulder FAA, Mittermaier A, Hon B, Dahlquist FW, Kay LE (2001) Studying excited states of proteins by NMR spectroscopy. *Nat Struct Biol* 8(11):932–935.
5. Tang C, Iwahara J, Clore GM (2006) Visualization of transient encounter complexes in protein-protein association. *Nature* 444(7117):383–386.
6. Eisenmesser EZ, Bosco DA, Akke M, Kern D (2002) Enzyme dynamics during catalysis. *Science* 295(5559):1520–1523.
7. Neudecker P, et al. (2012) Structure of an intermediate state in protein folding and aggregation. *Science* 336(6079):362–366.
8. Karplus M, Kuriyan J (2005) Molecular dynamics and protein function. *Proc Natl Acad Sci USA* 102(19):6679–6685.
9. Boehr DD, McElheny D, Dyson HJ, Wright PE (2006) The dynamic energy landscape of dihydrofolate reductase catalysis. *Science* 313(5793):1638–1642.
10. Sekhar A, Kay LE (2013) NMR paves the way for atomic level descriptions of sparsely populated, transiently formed biomolecular conformers. *Proc Natl Acad Sci USA* 110(32):12867–12874.
11. Palmer AG, 3rd, Kroenke CD, Loria JP (2001) Nuclear magnetic resonance methods for quantifying microsecond-to-millisecond motions in biological macromolecules. *Methods Enzymol* 339:204–238.
12. Hansen DF, Vallurupalli P, Kay LE (2008) Using relaxation dispersion NMR spectroscopy to determine structures of excited, invisible protein states. *J Biomol NMR* 41(3):113–120.
13. Bouvignies G, et al. (2011) Solution structure of a minor and transiently formed state of a T4 lysozyme mutant. *Nature* 477(7362):111–114.
14. Korzhnev DM, Religa TL, Banachewicz W, Fersht AR, Kay LE (2010) A transient and low-populated protein-folding intermediate at atomic resolution. *Science* 329(5997):1312–1316.
15. Vallurupalli P, Hansen DF, Stollar E, Meirovitch E, Kay LE (2007) Measurement of bond vector orientations in invisible excited states of proteins. *Proc Natl Acad Sci USA* 104(47):18473–18477.
16. Bouvignies G, Vallurupalli P, Kay LE (2014) Visualizing side chains of invisible protein conformers by solution NMR. *J Mol Biol* 426(3):763–774.
17. Hansen AL, Lundström P, Velyvis A, Kay LE (2012) Quantifying millisecond exchange dynamics in proteins by CPMG relaxation dispersion NMR using side-chain 1H probes. *J Am Chem Soc* 134(6):3178–3189.
18. Weininger U, Respondek M, Akke M (2012) Conformational exchange of aromatic side chains characterized by L-optimized TROSY-selected ^{13}C CPMG relaxation dispersion. *J Biomol NMR* 54(1):9–14.
19. Hansen DF, Vallurupalli P, Kay LE (2009) Measurement of methyl group motional parameters of invisible, excited protein states by NMR spectroscopy. *J Am Chem Soc* 131(35):12745–12754.
20. Baldwin AJ, et al. (2012) Probing dynamic conformations of the high-molecular-weight αB-crystallin heat shock protein ensemble by NMR spectroscopy. *J Am Chem Soc* 134(37):15343–15350.
21. Hvidt A, Nielsen SO (1966) Hydrogen exchange in proteins. *Adv Protein Chem* 21:287–386.
22. Wildes D, Marqusee S (2004) Hydrogen-exchange strategies applied to energetics of intermediate processes in protein folding. *Methods Enzymol* 380:328–349.
23. Englander SW, et al. (1992) Hydrogen exchange measurement of the free energy of structural and allosteric change in hemoglobin. *Science* 256(5064):1684–1687.
24. Maity H, Maity M, Krishna MMG, Mayne L, Englander SW (2005) Protein folding: The stepwise assembly of foldon units. *Proc Natl Acad Sci USA* 102(13):4741–4746.
25. Englander SW, Mayne L, Bai Y, Sosnick TR (1997) Hydrogen exchange: The modern legacy of Linderstrøm-Lang. *Protein Sci* 6(5):1101–1109.
26. Hwang T, Mori S, Shaka AJ, Van Zijl PCM (1997) Application of phase-modulated CLEAN chemical exchange spectroscopy (CLEANEX-PM) to detect water-protein proton exchange and intermolecular NOEs. *J Am Chem Soc* 119:6203–6204.
27. Fitzkee NC, Torchia DA, Bax A (2011) Measuring rapid hydrogen exchange in the homodimeric 36 kDa HIV-1 integrase catalytic core domain. *Protein Sci* 20(3):500–512.
28. Kaltashov IA, Bobst CE, Abzalimov RR (2013) Mass spectrometry-based methods to study protein architecture and dynamics. *Protein Sci* 22(5):530–544.
29. Kateb F, Pelupessy P, Bodenhausen G (2007) Measuring fast hydrogen exchange rates by NMR spectroscopy. *J Magn Reson* 184(1):108–113.
30. Fawzi NL, Ying J, Ghirlando R, Torchia DA, Clore GM (2011) Atomic-resolution dynamics on the surface of amyloid-β protofibrils probed by solution NMR. *Nature* 480(7376):268–272.
31. Vallurupalli P, Bouvignies G, Kay LE (2012) Studying "invisible" excited protein states in slow exchange with a major state conformation. *J Am Chem Soc* 134(19):8148–8161.
32. Pervushin K, Riek R, Wider G, Wüthrich K (1997) Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. *Proc Natl Acad Sci USA* 94(23):12366–12371.
33. Goldman M (1984) Interference effects in the relaxation of a pair of unlike spin-1/2 nuclei. *J Magn Reson* 60:437–452.
34. Pawson T (1995) Protein modules and signalling networks. *Nature* 373(6515):573–580.
35. Di Nardo AA, et al. (2004) Dramatic acceleration of protein folding by stabilization of a nonnative backbone conformation. *Proc Natl Acad Sci USA* 101(21):7954–7959.
36. McConnell HM (1958) Reaction rates by nuclear magnetic resonance. *J Chem Phys* 28:430–431.
37. Mittermaier A, Kay LE (2004) The response of internal dynamics to hydrophobic core mutations in the SH3 domain from the Fyn tyrosine kinase. *Protein Sci* 13(4):1088–1099.
38. Bai Y (2006) Protein folding pathways studied by pulsed- and native-state hydrogen exchange. *Chem Rev* 106(5):1757–1768.
39. Bai Y, Sosnick TR, Mayne L, Englander SW (1995) Protein folding intermediates: native-state hydrogen exchange. *Science* 269(5221):192–197.
40. Zhang YZ (1995) Protein and peptide structure and interactions studied by hydrogen exchange and NMR. PhD thesis (Univ of Pennsylvania, Philadelphia).
41. Bai Y, Milne JS, Mayne L, Englander SW (1993) Primary structure effects on peptide group hydrogen exchange. *Proteins* 17(1):75–86.