Optimizing frequency sampling in CEST experiments

Nicolas Bolik-Coulon[1-3*], D. Flemming Hansen[4*], Lewis E Kay[1-3,5*]

1) Department of Molecular Genetics, University of Toronto, Toronto, Canada, M5S 1A8

2) Department of Chemistry, University of Toronto, Toronto, ON, Canada, M5S 3H6.

3) Department of Biochemistry, University of Toronto, Toronto, ON, Canada, M5S 1A8.

4) Department of Structural and Molecular Biology, Division of Biosciences, University College London, London, United Kingdom, WC1E 6BT

5) Program in Molecular Medicine, Hospital for Sick Children Research Institute, Toronto, ON, Canada, M5G 0A4.

To whom correspondence should be addressed:

Nicolas Bolik-Coulon   E-mail: nicolas.bolikcoulon@utoronto.ca

D. Flemming Hansen   E-mail: d.hansen@ucl.ac.uk

Lewis E Kay   Email: kay@pound.med.utoronto.ca

Abstract

For the past decade Chemical Exchange Saturation Transfer (CEST) experiments have been successfully applied to study exchange processes in biomolecules involving sparsely populated, transiently formed conformers. Initial implementations focused on extensive sampling of the CEST frequency domain, requiring significant measurement times. Here we show that the lengthy sampling schemes often used are not optimal and that reduced frequency sampling schedules can be developed without *a priori* knowledge of the exchange parameters, that only depend on the chosen B_1 field, and, to a lesser extent, on the intrinsic transverse relaxation rates of ground state spins. The reduced sampling approach described here can be used synergistically with other methods for reducing measurement times such as those that excite multiple frequencies in the CEST dimension simultaneously, or make use of non-uniform sampling of indirectly detected time domains, to further decrease measurement times. The proposed approach is validated by analysis of simulated and experimental datasets.

Introduction

Biomolecules are inherently dynamic and the plethora of conformational states that they span can be critical for their function (Alderson and Kay, 2021; Boehr et al., 2006; Henzler-Wildman and Kern, 2007; Karplus and Kuriyan, 2005; Kimsey et al., 2018). Over the past several decades a variety of different NMR spin relaxation experiments have been developed that allow site-specific dynamics information to be obtained on biomolecules of different sizes and over a range of different timescales (Anthis and Clore, 2015; Mittermaier and Kay, 2006), informing on the energy landscapes of the molecules of interest. In particular, experiments for "observing" invisible states of molecules, so called conformationally excited states, provide kinetic and thermodynamic parameters of structural transitions from the ground, highly populated state to other, less populated states on the energy landscape (Korzhnev et al., 2004; Palmer et al., 2005), while in some cases also generating detailed structural information on these low populated conformations that is not available from other biophysical methods (Bouvignies et al., 2011; Korzhnev et al., 2010; Neudecker et al., 2012; Stiller et al., 2022; Vallurupalli et al., 2008).

Included in the list of NMR experiments for studies of sparsely formed, transiently populated conformers is Chemical Exchange Saturation Transfer (CEST), the physics of which was originally described by Forsén and Hoffman (Forsén and Hoffman, 1963), with subsequent important applications in the field of NMR imaging (van Zijl and Yadav, 2011; Ward et al., 2000). Details of these experiments for studies of invisible states can be found in a review by Vallurupalli and coworkers (Vallurupalli et al., 2017). Briefly, a weak magnetic field (generally between 5-50 Hz) is applied over a range of frequencies, typically one at a time, in "search" of the resonance positions of spins in the invisible state. When the frequency of the field coincides with the position of one such spin, the perturbation to its magnetization is transferred via conformational exchange to the corresponding spin in the ground state, decreasing its signal intensity. A plot of the intensity of the observable ground state signal as a function of the position of the weak magnetic field yields a CEST profile consisting of a pair of dips for each spin in a two-site exchanging system, corresponding to the resonance positions of the spin in the major and minor conformers. Fits of the CEST profiles to a model of chemical exchange then yields the kinetics and thermodynamics of the interconversion process, as well as the chemical shifts of spins in the excited state (Vallurupalli et al., 2012).

As described above, a weak B_1 field is necessary to sample the frequency range in which the resonance positions of spins in the excited state are expected to be found. Because

the sampling is, of necessity, discrete, the question of what the frequency spacing should be naturally emerges. This is of considerable practical utility as the measurement time scales linearly with the number of frequency values, much as the measurement time in a conventional multi-dimensional NMR experiment increases in proportion to the number of indirectly detected points. Typically, sampling steps are set to approximately the magnitude of the B_1 field in Hz (Vallurupalli et al., 2017, 2012), but this can lead to a large number of points, particularly for applications performed at high static magnetic fields. For instance, for a B_1 field of 15 Hz, a carbon-13 CEST experiment covering 5 kHz (*i.e.* 20 ppm on a 1 GHz NMR spectrometer) would require 330 points. The vast majority of these only contribute to the profile baseline, providing very little information about the exchange process. This issue has been recognized previously, and a number of approaches have been designed to minimize the number of sampled frequency points. These include multiple-frequency excitation techniques whereby several frequencies are excited simultaneously by using a B_1 field comprised of a set of cosine pulses (Leninger et al., 2018; Yuwen et al., 2018a), referred to as MF-CEST or cos-CEST or, alternatively, by applying a large number of very short, equally spaced DANTE-pulses (Bodenhausen et al., 1976) in the time domain (D-CEST) (Yuwen et al., 2018b). Both sets of related experiments lead to significant savings in measurement time because only the frequency region between successive excitation points generated via the cosine or DANTE-schemes must be sampled. Thus, for a spacing between DANTE pulses in the time domain of τ_D, a frequency range of $1/\tau_D$ (typically several hundreds of Hz) must be surveyed rather than potentially several kHz. Experimental measurement times can also be shortened by focusing on the "non-CEST" dimensions in multi-dimensional NMR experiments, where the use of Non-Uniform Sampling (NUS) schemes (Hyberts et al., 2007; Orekhov and Jaravine, 2011) have been shown to introduce minimal artifacts when applied to CEST experiments while leading to order of magnitude decreases in measurement times (Jameson et al., 2019; Long et al., 2015). A third approach in which the CEST profile is experimentally sampled by two-fold fewer points and then subsequent extended using a procedure involving inverse Fourier-transformation of the CEST frequency domain, linear prediction of the resulting time domain, followed by Fourier-transformation to regenerate a frequency profile has also been proposed (Carneiro et al., 2015).

Herein we examine what an "optimal" sampling profile in the CEST frequency domain might be to achieve significant decreases in measurement time without the introduction of artifacts. Such a sampling scheme could then be used in concert with other time-saving approaches such MF-/cos- or D-CEST, or with NUS schemes for sampling of

indirect detection dimensions, to further reduce experimental times. By calculating linewidths (Full-Width-at-Half-Height) of major and minor dips in CEST profiles over a broad range of exchange parameters we developed a simple relation between frequency sampling and the strength of the B_1 field ($\frac{\omega_1}{2\pi}$) that doesn't require *a priori* knowledge of the exchange parameters. The sampling schemes that emerge from our analysis have been validated using both simulated and experimental CEST and D-CEST data, and the extracted chemical shifts, kinetics, and relaxation parameters from analysis of ^{15}N datasets are in good agreement with those obtained from conventional sampling approaches. Significant savings in time are realized, typically a factor of two or more, that increases as the weak B_1 field strength decreases.

Materials and Methods

Protein production, purification, and sample preparation. The A39G FF domain from HYPA/FBP11 was overexpressed in *E.coli* BL21(DE3) cells transformed with the pET-29b plasmid containing the A39G FF gene. Cells were grown in M9 media containing 1 g/L ^{15}NH$_4$Cl as the sole nitrogen source and the protein was purified using a two-step protocol consisting of both cation exchange and size-exclusion chromatography, as described previously (Vallurupalli et al., 2009). The purified protein was concentrated to 2 mM in 50 mM sodium acetate (pH 5.7), 100 mM NaCl, 2 mM EDTA, 2 mM NaN$_3$ and 2.5% D$_2$O.

NMR spectroscopy. NMR experiments were performed using a 600 MHz Bruker Avance III HD spectrometer equipped with a triple resonance cryoprobe with x,y,z pulsed field gradients. CEST and D-CEST datasets were recorded (at 2°C) using previously published pulse sequences described by (Vallurupalli et al., 2012) and (Yuwen et al., 2018b), respectively. Spectral widths of 1650 Hz and {700 Hz, 800 Hz} were used in CEST and D-CEST experiments, respectively, along with weak B_1 fields of 15 Hz (D-CEST, CEST) or 30 Hz (CEST). In D-CEST experiments the DANTE pulses were applied with a 1.6 kHz field.

All spectra were acquired with 51 complex points in the ^{15}N-dimension and a spectral width of 24 ppm, for an acquisition time of approximately 35 ms in the indirect dimension. The number of frequency points in the CEST dimension of each experiment (*N*) was calculated as described in the main text for the reduced sampling approach; for standard sampling a spacing corresponding to the strength of the weak B_1 field was used. Values of *N*

are listed in Table 2. The ^{15}N weak B$_1$ field was calibrated using the nutation method of (Guenneugues et al., 1999). Although the B$_1$ field values used are listed as 15 Hz and 30 Hz for simplicity, the calibrated values were 16.6 Hz and 32.3 Hz (CEST) and 16.5 Hz (D-CEST).

Data analysis. NMR data were processed in *NMRPipe* (Delaglio et al., 1995) and visualized in *NMRFAM-SPARKY* (Lee et al., 2015) for peak picking. Peak intensities and chemical shifts for the ground state were obtained using *Peakipy* (https://github.com/j-brady/peakipy) and subsequently used for analysis in *Chemex* (https://github.com/gbouvignies/ChemEx) assuming a two-state exchange model. Analyses in *Chemex* were performed in two stages: an initial iteration where all residues showing a distinct minor dip were analyzed together to obtain values of the population of the minor state, p_E, and the exchange rate constant, k_{ex}, and a second iteration where all residues were included with p_E and k_{ex} held fixed to the value obtained in the first step. The initial iteration was performed with 24 residues for the CEST experiment with a B$_1$ field strength of 15 Hz (13, 22, 26, 28, 29, 33, 37, 38, 41, 42, 43, 45, 49, 50, 52, 53, 54, 55, 56, 59, 63, 66, 67, 68), 13 residues for the CEST experiment with a B$_1$ field strength of 30 Hz (26, 28, 29, 33, 37, 41, 42, 43, 50, 52, 55, 59, 67) and 24 residues for the D-CEST experiment (13, 22, 26, 28, 29, 33, 37, 38, 41, 42, 44, 45, 50, 52, 53, 54, 55, 56, 59, 60, 64, 66, 67, 69).

Simulation of CEST profiles. CEST profiles were computed by solving the set of coupled differential equations

$$\frac{d}{dt}\vec{M}(t) = -\hat{\mathcal{L}}\vec{M}(t) \tag{1}$$

where

$$\vec{M}(t) = (M_{x,G}(t) \quad M_{y,G}(t) \quad M_{z,G}(t) \quad M_{x,E}(t) \quad M_{y,E}(t) \quad M_{z,E}(t))^T \tag{2a}$$

$$\hat{\mathcal{L}} = \begin{pmatrix} R_{2,G}+k_{GE} & \omega_G-\omega_{rf} & 0 & -k_{EG} & 0 & 0 \\ -\omega_G+\omega_{rf} & R_{2,G}+k_{GE} & \omega_1 & 0 & -k_{EG} & 0 \\ 0 & -\omega_1 & R_{1,G}+k_{GE} & 0 & 0 & -k_{EG} \\ -k_{GE} & 0 & 0 & R_{2,E}+k_{EG} & \omega_E-\omega_{rf} & 0 \\ 0 & -k_{GE} & 0 & -\omega_E+\omega_{rf} & R_{2,E}+k_{EG} & \omega_1 \\ 0 & 0 & -k_{GE} & 0 & -\omega_1 & R_{1,E}+k_{EG} \end{pmatrix} \tag{2b}$$

and $\vec{M}(0) = (0 \quad 0 \quad p_G \quad 0 \quad 0 \quad p_E)^T$ (McConnell, 1958; Vallurupalli et al., 2017). In Eq. 2 T is the transpose operator, $R_{2,i}$ and $R_{1,i}$ are transverse and longitudinal relaxation rates of an exchanging spin in state i, k_{ij} is the exchange rate from state i to state j, $\omega_i - \omega_{rf}$ is the offset of the weak B$_1$ field from the resonance position of a spin in state i, and ω_1 is the B$_1$ field

strength. In order to evolve magnetization, the Liouvillian, $\hat{\mathcal{L}}$, was first diagonalized and any resulting eigenvalue with an imaginary part larger than 10^{-3} was multiplied by 10^9, effectively eliminating any frequency evolution of the spin system. This has the effect of 'smoothing' the resulting CEST profile, as would occur experimentally in the presence of radio frequency field inhomogeneity. CEST profiles were subsequently generated by evolving the spin system using the modified diagonal elements of $\hat{\mathcal{L}}$ to generate $\frac{M_{z,G}(T_{ex})}{M_{z,G}(0)}$ vs ω_{rf}, where T_{ex} is the exchange time in the CEST experiment. Linewidths of the dips in the CEST profiles were obtained by first inverting each profile so that the baseline is at zero (as in a "normal" spectrum), as shown in Figure 1a (left), and subsequently using the signal.peak_widths function from the Scipy Python library (SciPy 1.0 Contributors et al., 2020). First, peaks are defined using the signal.find_peaks function (SciPy 1.0 Contributors et al., 2020) which detects peak maxima based on values of adjacent points: a peak is defined at a particular point only if the two adjacent points have lower intensities. In simulations where noise is not included this procedure works well; in cases where noise is present it is possible to set a threshold below which peaks are not defined. Then, an array-search is performed until the closest half peak-intensity is found on both sides of the peak, thus defining the linewidth.

Generation of random profiles. We have evaluated the efficacy of the reduced sampling approach for recording CEST profiles via computation and experiment. In the case of computation, we considered a two-site exchanging system with exchange and relaxation parameter ranges as indicated in Table 1. One thousand sets, each containing two CEST profiles, were simulated, with one of the profiles of the set computed at a B_1 field strength randomly chosen between 5 and 20 Hz, and the second computed with a B_1 field strength between 21 and 50 Hz. Datasets were simulated with a static magnetic field chosen between 500 MHz and 1.2 GHz. Each CEST pair was analyzed only if both major and minor dips could be detected using the signal.find_peaks function from the Scipy Python library (SciPy 1.0 Contributors et al., 2020). In a second set of computations the difference in chemical shifts of ground and excited state spins, $\Delta\widetilde{\omega}$ (ppm), was decreased such that $|\Delta\widetilde{\omega}| < 0.5$ ppm and the resulting profiles analyzed even if only one peak could be detected with the signal.find_peaks function. In this way, the accuracy of extracted $\Delta\widetilde{\omega}$ values can be established for both large and small shift differences.

Results and Discussion

Typical CEST sampling schemes are non-optimal

CEST profiles have many features similar to 1D NMR spectra, most notably the fact that the positions of the dips correspond to those of peaks that would be obtained in conventional pulse-acquire schemes, at least in the limit where the minor state peaks could be observed in this way. In a conventional 1D experiment it is well established that acquisition times should not extend beyond where signal persists. Sampling beyond this limit adds noise but no new information. The inverse problem, that is what the optimal sampling in the frequency domain might be for CEST, is the question we address here. Insight can be obtained by examining CEST time domain profiles, obtained by inverse Fourier transformation of the frequency data; clearly, by analogy with conventionally recorded spectra, these should also not extend beyond where there is signal. Denoting the CEST profile as $I(v_{rf})/I_o$, where $I(v_{rf})$ and I_o are intensities of peaks in a CEST experiment recorded with ($I(v_{rf})$) and without (I_o) the CEST mixing period (duration T_{ex}), the corresponding time domain is generated by first inverting the CEST profile according to $\max(I(v_{rf})/I_o)$- $I(v_{rf})/I_o$ or cest_baseline - $I(v_{rf})/I_o$, where cest_baseline is the fitted baseline, followed by real Fourier transformation (Karunanithy et al., 2021). This is illustrated schematically in Figure 1a for a simulated profile generated for an exchanging system with parameters as indicated. Here we have chosen the frequency spacing between CEST points, Δv_{rf}, to be 40 Hz, significantly larger than 15 Hz that corresponds to the strength of the B_1 field, which we might have arbitrarily used in an experiment (a spacing set to the strength of the B_1 field is typically used in our laboratory). For this sampling interval the time domain signal extends until the end of the acquisition period, as might be the case had the data being optimally recorded in the time domain to start with. Note that a real discrete Fourier transform of a CEST profile generated with frequency sampling of $\Delta v_{rf} = sw/N$ between each CEST point, where sw and N are the spectral width of the region sampled and the total number of sampled frequency points, respectively, produces a time domain signal with approximately $N/2$ non-redundant real and imaginary points ($N/2+1$ and $(N+1)/2$ for even and odd N, respectively, Figure S1) (Karunanithy et al., 2021). In what follows, we display time domain signals of only the non-redundant points, extending to $N/(2sw)$ (even N) or $(N-1)/(2sw)$ (odd N) in the time-domain. The second half of the time domain can be regenerated from the first half, noting that $f_k = f_{N-k}^*$, where f_i is the ith complex point and * denotes complex conjugate (Figure S1).

Figures 1b,c show examples from experimental data recorded on an A39G mutant FF domain that has been established previously to undergo exchange between a number of different conformational states, including an unfolded ensemble that is responsible for the minor state dips shown here (Tiwari et al., 2021; Vallurupalli and Kay, 2013). Real Fourier transformation of the profiles in Figures 1a,b, recorded with a 15 Hz spacing between points and a B_1 field of 15 Hz, clearly show that the frequency domain data is sampled too densely (*i.e.*, N is too large), as 60 to 70% of the time domain points are essentially noise.

Influence of exchange parameters and B_1 field on the CEST linewidth

We have established that the rate of decay of the time domain CEST signal is an important parameter to consider in choosing appropriate frequency spacings for recording CEST profiles. This decay rate, in turn, relates to the linewidths of the dips in the frequency domain, and it is of interest, therefore, to establish how the linewidths of the major and minor dips in frequency domain profiles depend on exchange parameters and the B_1 field strength. To this end, we first derive approximate expressions for CEST dip linewidths for a two-site exchanging spin system. In the derivation that follows we assume that the populations of the ground (p_G) and excited (p_E) states are highly skewed ($p_G \gg p_E$), as is most often encountered in CEST experiments, and that the frequency difference $\Delta\omega = \omega_E - \omega_G$ between the spin in the interconverting states is sufficiently large such that well resolved dips are observed. The CEST profile can be approximated as (Palmer, 2014)

$$C(\omega_{rf}) = \frac{(\omega_G - \omega_{rf})^2}{(\omega_G - \omega_{rf})^2 + \omega_1^2} e^{-R_{1\rho}(\omega_{rf})T_{ex}}, \tag{3}$$

where ω_{rf} is the frequency position of the applied B_1 field (strength $\omega_1 = \gamma B_1$, with γ the gyromagnetic ratio for the nuclei of interest), T_{ex} is the length of the irradiation period during which exchange is monitored, and $R_{1\rho}$ is given by (Trott and Palmer, 2002)

$$R_{1\rho}(\omega_{rf}) = R_{1,G} \frac{(\omega_G - \omega_{rf})^2}{(\omega_G - \omega_{rf})^2 + \omega_1^2} + \frac{\omega_1^2}{(\omega_G - \omega_{rf})^2 + \omega_1^2} \left(R_{2,G} + \frac{p_G p_E \Delta\omega^2 k_{ex}}{(\omega_E - \omega_{rf})^2 + \omega_1^2 + k_{ex}^2} \right), \tag{4}$$

where $R_{1,G}$ and $R_{2,G}$ are the longitudinal and transverse intrinsic auto-relaxation rates for a spin in the ground state, respectively, and k_{ex} is the sum of the forward and reverse exchange rates between G and E. In order to calculate linewidths at half-height for each of the dips in the CEST profile, $\Delta\nu_{1/2}$, we invert the CEST profile, $C(\omega_{rf})$, as shown in Figure 1a

$$C'(\omega_{rf}) = e^{-R_{1,G}T_{ex}} - C(\omega_{rf}), \tag{5}$$

and solve for ω_{rf} using

$$\frac{c'(\omega_i)}{c'(\omega_{rf})} = 2, \quad i\epsilon\{G, E\}, \tag{6}$$

where it is understood that according to Eq. 6, ω_{rf} is the frequency position where the intensity of the CEST dip corresponding to state i is half maximal. Consider, first, $i=E$. Assuming that $\Delta\omega_{1/2} \ll \Delta\omega$, where $\Delta\omega_{1/2}$ is the dip linewidth (in rad/sec) at half height, gives $\omega_G - \omega_{rf} \approx -\Delta\omega$, and Eq. 6 can be expanded to

$$e^{-R_{1\rho}(\omega_{rf})T_{ex}} = \frac{1}{2}\left(\frac{\Delta\omega^2+\omega_1^2}{\Delta\omega^2}e^{-R_{1,G}T_{ex}} + e^{-R_{1\rho}(\omega_E)T_{ex}}\right). \tag{7}$$

from which it follows that

$$R_{1\rho}(\omega_{rf}) = -\frac{\mathcal{L}}{T_{ex}}, \tag{8}$$

where

$$\mathcal{L} = \ln\frac{1}{2}\left(\frac{\Delta\omega^2+\omega_1^2}{\Delta\omega^2}e^{-R_{1,G}T_{ex}} + e^{-R_{1\rho}(\omega_E)T_{ex}}\right). \tag{9}$$

Substituting Eq. 4 for $R_{1\rho}(\omega_{rf})$ in Eq. 8 yields

$$\left((\omega_E - \omega_{rf})^2 + \omega_1^2 + k_{ex}^2\right)R_{eff}^E = p_G p_E \Delta\omega^2 k_{ex}, \tag{10}$$

where

$$R_{eff}^E = -R_{2,G} - R_{1,G}\frac{\Delta\omega^2}{\omega_1^2} - \frac{\Delta\omega^2+\omega_1^2}{\omega_1^2}\frac{\mathcal{L}}{T_{ex}}. \tag{11}$$

Note that $\mathcal{L}$ is negative for standard values of T_{ex} (hundreds of milliseconds) and $\frac{\Delta\omega}{2\pi}$ is large, (recall that major and minor state dips are well resolved), leading to positive values for R_{eff}^E. Solving Eq. 10 for ω_{rf} gives the linewidth for the excited state, expressed in Hz, as

$$\Delta v_{1/2}^E = \frac{1}{\pi}\sqrt{\frac{p_G p_E \Delta\omega^2 k_{ex}}{R_{eff}^E} - \omega_1^2 - k_{ex}^2}. \tag{12}$$

A simpler equation can be derived for $\Delta v_{1/2}^E$ for very weak B$_1$ fields such that $\Delta\omega \gg \omega_1$. In this case, following from Eq. 4, we can write

$$R_{1\rho}(\omega_E) \approx R_{1,G} + \frac{p_G p_E \omega_1^2 k_{ex}}{\omega_1^2+k_{ex}^2}, \tag{13a}$$

$$R_{1\rho}(\omega_{rf}) \approx R_{1,G} + \frac{p_G p_E \omega_1^2 k_{ex}}{(\omega_E-\omega_{rf})^2+\omega_1^2+k_{ex}^2}. \tag{13b}$$

Substituting Eq. 13 into Eq. 7, we obtain,

$$e^{-R_{1\rho}(\omega_{rf})T_{ex}} = \frac{1}{2}\left(e^{-R_{1,G}T_{ex}} + e^{-R_{1\rho}(\omega_E)T_{ex}}\right), \tag{14}$$

from which it follows directly that,

$$e^{-\{R_{1\rho}(\omega_{rf})-R_{1,G}\}T_{ex}} \approx \frac{1}{2}\left(e^{\{R_{1\rho}(\omega_E)-R_{1,G}\}\frac{T_{ex}}{2}} + e^{-\{R_{1\rho}(\omega_E)-R_{1,G}\}\frac{T_{ex}}{2}}\right)e^{-\{R_{1\rho}(\omega_E)-R_{1,G}\}\frac{T_{ex}}{2}} \tag{15a}$$

$$\approx cosh\left(\{R_{1\rho}(\omega_E) - R_{1,G}\}\frac{T_{ex}}{2}\right)e^{\{R_{1\rho}(\omega_E)-R_{1,G}\}\frac{T_{ex}}{2}}. \qquad (15b)$$

Noting that $\ln(\cosh\frac{x}{2}) \approx \frac{x^2}{8}$ (so long as $x < \frac{\pi}{2}$; $(R_{1\rho}(\omega_E) - R_{1,G})T_{ex} < \frac{\pi}{2}$ in this case) and taking the log of both sides of Eq. 15, we obtain, after some algebra,

$$\Delta v_{1/2}^E \approx \frac{1}{\pi}\sqrt{(\omega_1^2 + k_{ex}^2)\frac{1+\alpha}{1-\alpha}}, \quad \text{where } \alpha = \frac{1}{4}\frac{\omega_1^2\, p_E\, p_G k_{ex}T_{ex}}{\omega_1^2 + k_{ex}^2}. \qquad (16)$$

A more accurate expression for $\Delta v_{1/2}^E$ can be obtained by considering a higher order expansion of $\ln(\cosh\frac{x}{2})$ but the resulting equation and Eq. 16 both simplify to

$$\Delta v_{1/2}^E \approx \frac{1}{\pi}\sqrt{\omega_1^2 + k_{ex}^2} \qquad (17)$$

in the limit where $\alpha \ll 1$, which is sufficient for a qualitative understanding of the curves in Figure 2a.

The linewidth of the major dip can be calculated by solving Eq. 6 for the case where $i=G$. We obtain

$$2\frac{(\omega_G-\omega_{rf})^2}{(\omega_G-\omega_{rf})^2+\omega_1^2}e^{-R_{1\rho}(\omega_{rf})T_{ex}} = e^{-R_{1,G}T_{ex}}. \qquad (18)$$

Noting that linewidths of the major state dips are, in general, considerably larger than the B_1 field strength (see Figure 2a), so that $(\omega_G - \omega_{rf})^2 > \omega_1^2$, Eq. 16 simplifies to

$$\ln(2) -\ R_{1\rho}(\omega_{rf})T_{ex} = -R_{1,G}\,T_{ex}. \qquad (19)$$

Substituting Eq. 4 into Eq. 17 with $\omega_E - \omega_{rf} \approx \Delta\omega$ produces

$$\ln 2 - \frac{\omega_1^2}{\left(\omega_G - \omega_{rf}\right)^2}\left(R_{2,G} + \frac{p_G p_E \Delta\omega^2 k_{ex}}{\omega_1^2 + \Delta\omega^2 + k_{ex}^2}\right)T_{ex} = 0 \qquad (20)$$

so that

$$\Delta v_{1/2}^G = \frac{\omega_1}{\pi}\sqrt{\frac{T_{ex}}{\ln 2}\left(R_{2,G} + \frac{p_G p_E \Delta\omega^2 k_{ex}}{\Delta\omega^2 + \omega_1^2 + k_{ex}^2}\right)}. \qquad (21)$$

Under the assumption that $\Delta\omega \gg \omega_1, k_{ex}$ and $p_G \approx 1$, Eq. 21 can be simplified to

$$\Delta v_{1/2}^G \approx \frac{\omega_1}{\pi}\sqrt{\frac{T_{ex}}{\ln 2}\left(R_{2,G} + p_E k_{ex}\right)}, \qquad (22)$$

and assuming that $R_{2,G} \gg p_E k_{ex}$, it follows that

$$\Delta v_{1/2}^G \approx \frac{\omega_1}{\pi}\sqrt{\frac{T_{ex}}{\ln 2}R_{2,G}}\left(1 + \frac{p_E k_{ex}}{2R_{2,G}}\right). \qquad (23)$$

Eqs. 17 and 23 are (very) approximate expressions for the linewidths of the minor and major states, respectively; these are simple enough so as to provide intuition as to how the widths vary as a function of exchange parameters and B₁ field strength. For example, $\Delta v_{1/2}^G$ varies linearly with $p_E k_{ex}$ and with ω_1. In contrast, $\Delta v_{1/2}^E$ varies in a more complex manner

with $p_E k_{ex}$. Consider Eq. 17 in the limit that $k_{ex} \gg \omega_1$ (in general $k_{ex} > \omega_1$ in CEST applications) so that $\Delta v_{1/2}^E \approx \frac{k_{ex}}{\pi}$: a plot of $\Delta v_{1/2}^E$ vs $p_E k_{ex}$ for constant p_E will be linear with a slope of $\frac{1}{\pi p_E}$. In contrast, a similar plot for constant k_{ex} will be flat (zero slope). Unlike the case for the major state dip where linewidths are always proportional to ω_1, when $k_{ex} \gg \omega_1$ there is essentially no ω_1 dependence for $\Delta v_{1/2}^E$.

Numerical evaluation of CEST linewidths

Figure 1 makes it clear that an optimal sampling of CEST profiles in the frequency domain requires knowledge of the decay rates of major and minor states (*i.e.*, their linewidths). Some insight is provided by the expressions derived above but, ultimately, it is necessary to resort to simulations to generate a complete, albeit, less intuitive picture. To this end we have simulated CEST profiles for a two-site exchanging system with a variety of exchange parameters and extracted linewidths of the major and minor dips as a function of p_E and k_{ex} (Figures 2a and S2, circles). Notably, we find that the major dip linewidth is only strictly dependent on the product $p_E k_{ex}$, while the minor dip shows strong variations with k_{ex} but limited changes with p_E. In addition, the broadening of the major dip increases faster than for the minor dip as B₁ gets stronger (Figures 2b and S2). The dependence of linewidths on exchange parameters is consistent with expectations based on the equations derived in the previous section. Indeed, Eqs. 12 and 21 (solid lines) reproduce the linewidths evaluated from the simulated CEST profiles reasonably well (Figures 2a and S2), although the simplified equations 17 and 23 are only valid in the limit of very weak B₁ fields or small p_E (1%) (Eq. 17) or for small $p_E k_{ex}$ values (Eq. 23) (see Figure S2, dashed lines).

A relation between B₁ field and minimum linewidths of CEST dips

The expressions for $\Delta v_{1/2}$ derived above, along with simulations of linewidths as a function of different exchange parameters (Figures 2 and S2), make it clear that it will be difficult to choose the optimum frequency sampling scheme without *a priori* knowledge of the exchange parameters, weak B₁ field used, and intrinsic relaxation properties of spins in each of the interconverting states. We sought, therefore, to obtain a simple expression for the minimum dip linewidth (considering both major and minor dips and a two-site exchanging system), $\Delta v_{1/2}^{min}$, as a function of B₁ over a wide spectrum of exchange parameters that encompass those typically observed experimentally. The minimum linewidth, *i.e.*, smallest

decay rate in the time domain, is a conservative choice as it ensures that a sufficient sampling of the frequency domain is taken, irrespective of the operative experimental and exchange parameters, so that robust estimates of the exchange kinetics and thermodynamics, as well as chemical shifts of the excited state, can be obtained from fits of the CEST profiles (see below).

Our goal here is to achieve a sampling of the CEST profile at a frequency rate of at least least one point per dip linewidth at half height (effectively two points per peak), corresponding to an acquisition time, after real Fourier transformation of the CEST profile, of $1.5\ T_2^{max}$ (see Supporting Information; $\Delta v_{1/2}^{min} = \frac{1}{\pi T_2^{max}}$). Simulations and analysis of experimental data show that accurate parameters are extracted from data sampled in this manner. In order to guarantee that at least two points are sampled for each dip in a CEST profile we have computed CEST curves, for given B_1, T_{ex}, and intrinsic transverse relaxation values, assumed to be the same in the ground and excited states. Values of exchange parameters were chosen at the low end of what is typically observed in CEST, p_E = 1% and k_{ex} = 50 s^{-1}, so that contributions from exchange to $\Delta v_{1/2}^{G/E}$ are relatively small (Figure 2a, see below). The minimum $\Delta v_{1/2} \in \{\Delta v_{1/2}^{G}, \Delta v_{1/2}^{E}\ \}$ is selected for each $\{B_1, T_{ex}, R_2\ \}$ set, generating the $\Delta v_{1/2}^{min}\ vs\ (\frac{\omega_1}{2\pi})$ relations shown in Figure 3. These are well approximated as linear and are relatively independent of T_{ex} (Figure 3a). As expected, there is some dependence on intrinsic relaxation rates (Figure 3b), but as we show with analysis of experimental data the extraction of accurate exchange parameters requires only very rough estimates of R_2 values. Using the profiles in Figure 3a,b relating $\Delta v_{1/2}^{min}$ and $(\frac{\omega_1}{2\pi})$, the number of CEST frequency points is given by

$$N = \frac{sw}{\Delta v_{1/2}^{min}}$$ (24)

and because $\Delta v_{1/2}$ values are computed directly from the CEST profiles, assumptions about the dip lineshape do not need to be made. Eq. 24 ensures that at least two points are recorded per dip (unless a point lies on the dip maximum) and because the values of p_E (1%) and k_{ex} (50 s^{-1}) are smaller than what is observed in most experimental applications $\Delta v_{1/2}^{min}$ is typically underestimated and N overestimated so that sufficient sampling of the CEST profiles is nearly always achieved. Exceptions would include cases where experimental exchange parameters are smaller than the values used to derive the relations in Figure 3b; however, it is worth noting that as p_E and k_{ex} values become increasingly small, minor dips will also

significantly decrease, challenging CEST applications. Figures 3c,d show plots of N vs $\left(\frac{\omega_1}{2\pi}\right)$ for standard and reduced sampling schemes as a function of R_2 rates. It is clear that large reductions in N are obtained via reduced sampling following Eq. 24, varying from factors of three or higher for small B_1 fields to greater than two for $\left(\frac{\omega_1}{2\pi}\right)$ = 50 Hz.

Establishing robustness of extracted exchange parameters based on simulations

Having established simple relationships to calculate N (Figure 3, Eq. 24), we then generated a set of simulated datasets that were subsequently fit to ensure that accurate estimates for the exchange ($\Delta\omega, p_E, k_{ex}$) and relaxation ($R_1$, $R_{2,G}$, $R_{2,E}$) parameters could be obtained. To this end a set of 2,000 CEST profiles was computed using 1,000 random two-site exchange and relaxation parameters (Table 1). For each set of parameters two profiles were simulated with different B_1 fields, between 5 and 20 Hz and between 21 and 50 Hz. Separate CEST profiles were computed using either the "standard" sampling approach (frequency increment of $\frac{\omega_1}{2\pi}$) or via a reduced sampling procedure based on the relation $\Delta v_{1/2}^{min}(Hz) =$ $2.15\frac{\omega_1}{2\pi} + 11.41$, corresponding to the case where R_2=20 s^{-1} (Figure 3b). Each pair of profiles, corresponding to different values of $\frac{\omega_1}{2\pi}$ but otherwise the same exchange and relaxation rates, was fit to a two-site exchange model using the program *Chemex* (https://github.com/gbouvignies/ChemEx). As the chosen values of $R_{2,G}$ rates for the simulated profiles range between 5 s^{-1} – 100 s^{-1} and $0.5R_{2,G} < R_{2,E} < 2R_{2,G}$, (Table 1), potentially quite different than the assumed $R_{2,G} = R_{2,E}$ = 20 s^{-1} (used to select the frequency spacing) a comparison of input *vs* fitted parameters is an excellent first step in testing the robustness of the reduced sampling approach. In the initial analysis, profiles were generated containing two distinct dips (see Materials and Methods) and the linear correlations between fitted and input parameters is excellent (Figure 4). We next repeated the computations (only 100 random sets of exchange and relaxation parameters) to generate CEST profiles with $|\Delta\widetilde{\omega}| < 0.5$ ppm. Each pair of CEST curves at two B_1 fields (randomly chosen within ranges as described above, Table 1) was analyzed together in *Chemex* with p_E and k_{ex} fixed to within 1% of their initial chosen values. (Note that in the analysis of CEST data one typically fits profiles with well-separated dips initially to obtain robust exchange parameters and then fixes these parameters in subsequent fits of all of the data). Correlations between input and fitted parameters obtained via analysis of CEST profiles generated from a single reduced sampling scheme are excellent (Figure S3), although in some cases fitted $R_{2,E}$ rates show

clear deviations from their accurate values. This is to be expected in the case where major and minor state dips overlap. These simulations, thus, provide a theoretical validation of our approach.

Experimental validation of the reduced sampling approach

The four-helix bundle FF domain from human HYPA/FBP11 (referred to in what follows as FF) has been used as a model system to understand protein folding (Jemth et al., 2005, 2004; Korzhnev et al., 2010, 2007; Tiwari et al., 2021). Studies of a variety of mutants of FF have provided a unified picture of the folding mechanism of this domain. Here we focus on the A39G mutant of FF that has recently been shown through extensive CEST data analysis to fold on a volcano energy landscape via at least two intermediate states (Tiwari et al., 2021). Notably, however, CEST data recorded on this system with one or two B_1 fields can be analyzed using a simple two-state model so long as $R_{2,G}$ and $R_{2,E}$ are fitted as separate parameters (Tiwari et al., 2021; Vallurupalli et al., 2012). In this case many of the obtained $R_{2,E}$ rates are large, reflecting additional exchange processes from the excited state.

Nitrogen-15 CEST profiles were measured on a sample of A39G FF using B_1 field strengths of 15 Hz and 30 Hz, with standard (112 and 57 points, respectively) and reduced (39 and 25 points, respectively) sampling approaches. In the latter case the relation $\Delta v_{1/2}^{min}(Hz) = 2.15\frac{\omega_1}{2\pi} + 11.41$ was used to select the frequency points. These datasets were analyzed individually (*i.e.*, one B_1 field at a time), using a two-state exchange model, by first fitting residues showing distinct minor dips to obtain values for p_E and k_{ex}, and then fixing these values in subsequent fits that included all profiles. In this way intrinsic relaxation rates and chemical shifts were generated for all ground and excited state spins. For both B_1 fields, the chemical shifts and relaxation rates agree well between the two sampling schemes (Figure 5), while extracted values for p_E and k_{ex} are similar (Table 2). Notably, even small $\Delta\varpi$ values, *i.e.* smaller than 0.5 ppm, can be obtained accurately, as shown in the insets to the $\Delta\varpi$ panels in Figure 5. Tyr49 stands out in the correlation plots of $R_{2,E}$ with significantly different values obtained for fits of datasets recorded with different sampling schemes but the same B_1 field strength, or between pairs of datasets both obtained using either reduced or standard sampling approaches but with different B_1 field strengths (*i.e.*, compare $R_{2,E}$ values for standard sampling with B_1 fields strengths of 15 and 30 Hz, for example). CEST profiles from this residue are less well-fit relative to others (compare Tyr49 and Leu59, that is well-fit, in Figure

S4), perhaps reflecting the fact that the FF exchange process is inherently more complex than two-state (Tiwari et al., 2021).

As described in the Introduction, a number of other methods have been developed for reducing the frequency sampling in the CEST dimension. These include MF-CEST (Leninger et al., 2018) or cos-CEST (Yuwen et al., 2018a), and D-CEST (Yuwen et al., 2018b). In D-CEST, the B$_1$ field is applied in the form of pulses at intervals of τ', with excitation frequencies $\omega_{rf} \pm 2\pi n/\tau'$, where n is an integer and ω_{rf} is the carrier frequency. This leads to a significant reduction in the CEST spectral width ($\frac{1}{\tau'}$), although it introduces ambiguities in the frequencies of the invisible state spins due to the periodic nature of the excitation profile (Yuwen et al., 2018b). This can be resolved by recording a pair of D-CEST experiments with different spectral widths, that is different τ' values, as described previously (Yuwen et al., 2018b). In practice, we choose the sampling lists for each of the two experiments in such a way as to ensure that the chosen frequencies for each spectral width do not overlap. For example, for sweep widths of 600 Hz and 560 Hz and a B$_1$ field strength of 20 Hz, we would choose offset frequencies of 300 - 40k Hz and 280 - 40k Hz, where k is a whole number, so that between ±280 Hz the CEST frequencies are sampled every 20 Hz when considering both profiles together, although each profile is comprised of points in frequency space sampled at twice this interval (40 Hz). Thus, although a pair of spectra must be recorded, the experimental time is essentially no longer than for a single dataset, with the number of frequency sampled points in each profile given by $\frac{N}{2}$, as calculated from Eq. 24 (Figure 3). The source code to generate the appropriate reduced sampling schemes for either CEST or D-CEST applications (ensuring that for each dataset the excitation frequencies lie exactly between each other for D-CEST), is given in Supporting Information.

In principle, it is possible to combine D-CEST with reduced sampling to achieve even greater savings in time than with either technique alone. Figure 6a compares D-CEST and reduced D-CEST profiles for a number of residues in A39G FF, recorded with spectral widths of 700 and 800 Hz and a B$_1$ field strength of 15 Hz. In total 52 and 19 frequency points were recorded for the standard and reduced D-CEST experiments respectively (this includes datasets at both spectral widths in each case). Analysis of the resulting profiles using *Chemex* shows that consistent values of $(\Delta\varpi, R_{2,G}, R_{2,E})$ (Figure 6b) and of the exchange parameters (Table 2) are obtained via the two sampling approaches, establishing that reduced sampling can be extended to D-CEST applications.

We were particularly interested in testing the limits of the reduced sampling approach by focusing on a residue where the dips for the ground and excited states are highly overlapped, such as for Gly39. Our experiments were performed with a concentrated sample (~ 2mM) and recorded on a spectrometer with a cryoprobe so that the sensitivity was very high (signal-to-noise ratios, SNR, of peaks recorded in the I_o spectrum are on the order of 1,000). In order to evaluate how decreased SNRs affect extracted $\Delta\varpi$ values when reduced frequency sampling is used we added increased amounts of noise to the experimental Gly39 CEST profile, shown with blue circles in Figure 7a, before fitting in *Chemex*. One hundred curves were generated by adding noise to each point in the experimental Gly39 CEST profile from a normal distribution of noise with standard deviation corresponding to the chosen noise level (0.1%, 1%, 3% and 5% of Gly39 I_o). CEST profiles were fit by fixing the exchange parameters p_E and k_{ex} to values reported in Table 2, determined from fits to all CEST profiles with well-separated dips. The extracted $\Delta\varpi$ values are plotted in the histograms shown in Figure 7b. The distributions obtained from reduced sampling are broader than those from the standard sampling approach for a given level of noise (noise levels are indicated adjacent to each histogram). However, the median peak position of the $\Delta\varpi$ distributions from fits of CEST data generated with standard and reduced sampling match to within 0.4 ppm across all error values, and to within 0.15 ppm of each other when errors are $\leq 3\%$. The distributions for the D-CEST data at large input errors (3% or 5%) are unacceptably wide, both for standard and reduced sampling, suggesting that D-CEST is best applied to high SNR samples.

Figure 7 shows that larger numbers of experimental CEST points lead to narrower distributions in fitted $\Delta\varpi$ values, and we wondered, therefore, if increasing the number of frequency points in a post-acquisition manner might improve the robustness of the extracted chemical shifts. This can be achieved, as illustrated in Figure 1a, where the frequency-domain CEST profile is treated as a 1D NMR spectrum. Thus, by inverse Fourier transformation, zero-filling and re-transformation, the number of points in the frequency dimension is increased. Figures S5-S11 show results based on an analysis of 13 experimental CEST profiles of A39G FF, measured by either standard ($\frac{\omega_1}{2\pi}$), or reduced (using the relation $\Delta v_{1/2}^{min}(Hz) = 2.15 \frac{\omega_1}{2\pi} + 11.41$, as described above) sampling approaches (see legend to Figure S5 for details). In the latter case, we have also generated datasets where the separation between CEST frequency points has been halved by one zero-fill in the time domain. As in the description of Figure 7, noise has been added to the data (0.1% - 5%), and the 100 generated profiles in this manner fit for each noise level. Mean values and standard deviations

of fitted values of $(\Delta\varpi, p_E, k_{ex})$ are plotted in either linear correlation or histogram formats. It is thus possible to evaluate (i) how doubling the number of CEST points in reduced sampled datasets (via zero-filling) influences extracted parameters, and, more generally, (ii) how noise affects the quality of extracted parameters from CEST datasets. These computations show that excellent correlations between $\Delta\varpi$ values are obtained via the different sampling schemes in all cases. When noise levels are low ($\leq 1\%$) robust exchange parameters are also obtained with all sampling approaches. For large noise values (5%) we find that zero-filling of the reduced sampling data leads to exchange parameters that are somewhat closer to those obtained from the standard sampling approach, although fits of profiles recorded with any of the sampling schemes, including the standard method, generate parameters that are considerably off, on average (Figures S6-S11). Notably, when datasets recorded at B_1 field strengths of 15 Hz and 30 Hz are fit together significantly, improved exchange parameters are obtained when high noise levels are present.

Concluding remarks

We have shown that the commonly used approach (at least by our laboratory) for frequency sampling of CEST profiles in which the step-size is set to the value of the B_1 field is not an optimal use of measurement time. Rather, in the absence of *a piori* knowledge of the exchange parameters, it is preferable to make use of simple $\Delta v_{1/2}^{min}(B_1)$ relations presented here that enable recording of CEST profiles with significantly reduced frequency sampling, typically by at least a factor of two. We have shown through simulation and experiment that accurate values of $(p_E, k_{ex}, \Delta\varpi, R_{2,G}, R_{2,E})$ are obtained in this manner. In principle, the time savings can be used to record a dataset with at least twice the number of scans, although it is not clear whether this is any better than using standard sampling (*i.e.*, defining CEST dips with fewer points of higher SNR, or increased points with less sensitivity). Rather the results of the simulations of Figures S5-S11 strongly suggest that recording a pair of reduced sampled experiments with different B_1 fields, especially when SNR is limiting, is the preferred method, as significantly more accurate exchange parameters can be obtained. We expect that reduced sampling will extend the applicability of CEST-type experiments to proteins that are not stable over long periods of time and especially to systems for which the exchange is best characterized using small B_1 fields where the number of sampled points is typically very high. Reduced sampling can be coupled with band excitation CEST methods

such as MF-, cos- or D-CEST and non-uniform sampling of the indirectly acquired chemical shift dimensions, leading to further decreases in measurement times.
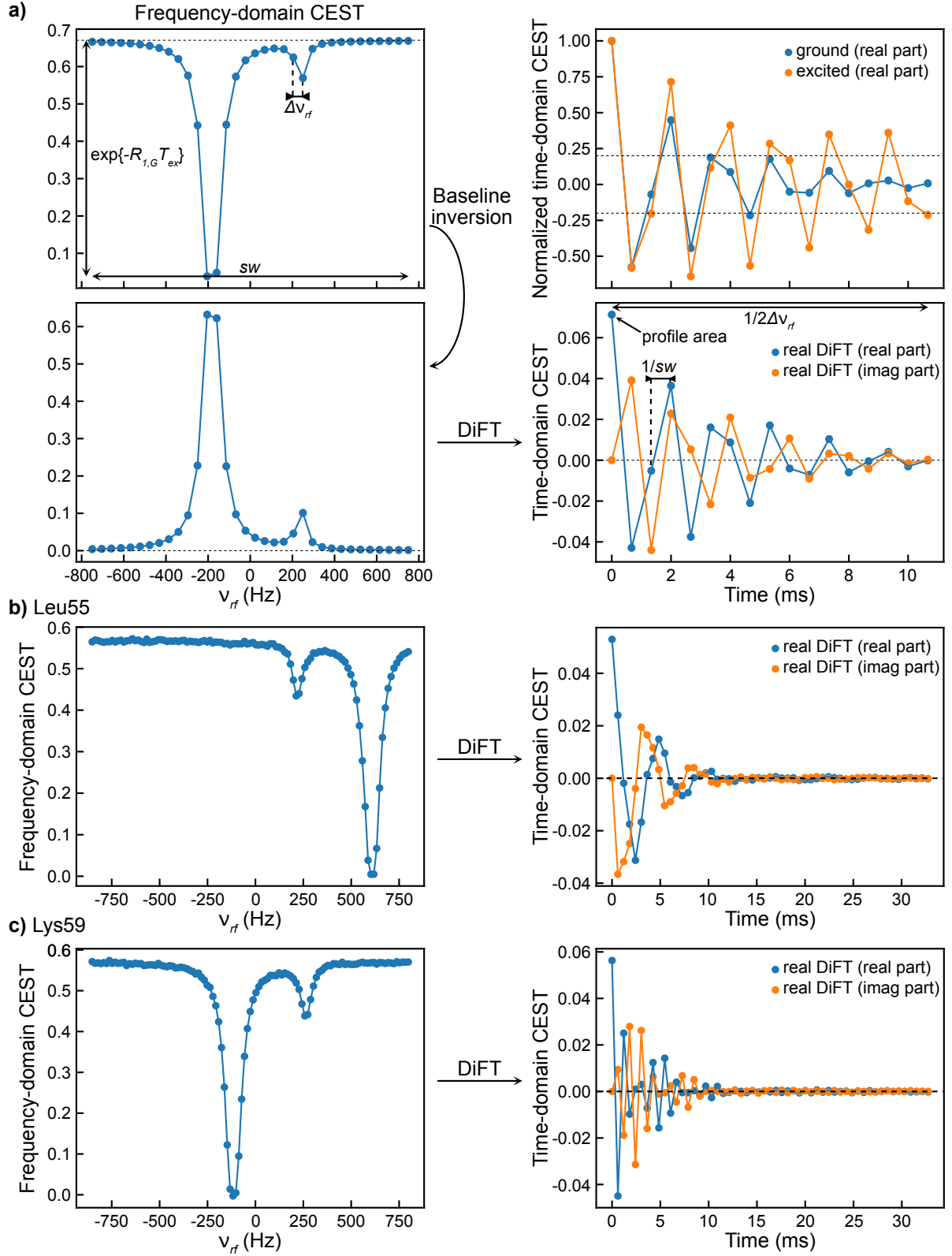
Acknowledgements

Figure 1. Standard CEST frequency domain sampling is typically non-optimal. (**a**) Transformation of frequency-domain CEST profiles into the time domain is obtained as shown schematically using simulated data. A CEST profile $\frac{I(v_{rf})}{I_o}$ is transformed to $\max(\frac{I(v_{rf})}{I_o}) - \frac{I(v_{rf})}{I_o}$, and subsequently a real discrete inverse Fourier transformation (DiFT) of the data produces the time domain response shown. The relations between the spectral width sampled in the frequency domain (sw = 1500 Hz), the acquisition time in the time domain ($\frac{1}{2x\Delta v_{RF}} = \frac{1}{2x45.46\ Hz} =$

12.3 ms), and the intervals between successive points in the frequency- (Δv_{RF}) and time- ($\frac{1}{sw}$) based profiles are shown. In the simulation a B_1 field of 15 Hz was used, T_{ex} = 400 ms., a two-site exchange process between ground (G) and excited (E) states is assumed, with population of the excited state p_E = 1%, and an exchange rate $k_{ex} = k_{GE} + k_{EG}$ = 100 s^{-1} where k_{ij} is the rate of transition from state i to state j. The intrinsic relaxation properties of both states are the same, with longitudinal relaxation and transverse relaxation rates equal to 1 s^{-1} and 20 s^{-1} respectively. The top-right panel shows the individual *normalized* time-domain CEST profiles for the ground and excited states, obtained under the assumption that both peaks can be modeled with a Lorentzian shape ($\frac{1}{2}\frac{\Delta v_{1/2}}{(v_{rf}-v_i)^2+(\Delta v_{1/2})^2/4}$, where $\Delta v_{1/2}$ is the dip linewidth obtained using the scipy signal.peak_width function and v_i is the frequency for spins in state i). The limit where 20% of the signal remains is shown with the horizontal dashed line. Experimental CEST profiles (left) of L55 (**b**) and K59 (**c**) of the A39G FF domain recorded at 600 MHz, 2°C, with a B_1 field strength of 15 Hz and a sampling frequency of 15 Hz were manipulated as shown schematically in (a). Points are connected by a line for visual clarity. It is clear that the acquisition in the time-domain is too long, corresponding to a smaller than optimal separation between successively sampled points in the frequency domain.
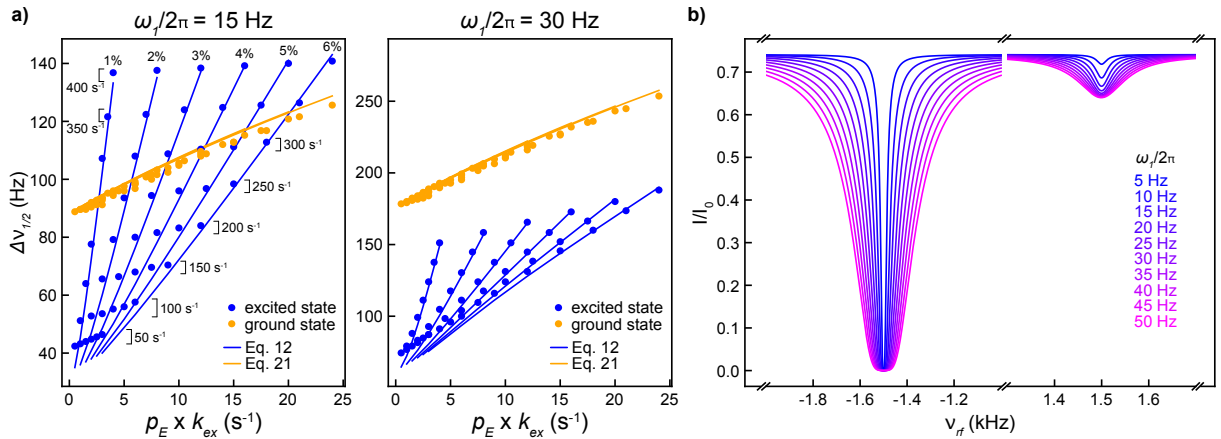

Figure 2. Linewidths at half height ($\Delta v_{1/2}$) for major and minor dips in CEST profiles. (**a**) Linewidths (circles) were obtained from simulated CEST profiles generated with longitudinal relaxation rates $R_{1,G} = R_{1,E}$ = 1 s^{-1}, transverse relaxation rates $R_{2,G} = R_{2,E}$ = 20 s^{-1}, T_{ex} = 300 ms, exchange parameters as indicated, and a (large) separation between major and minor dips of 3 kHz to ensure no overlap and, hence, accurate estimates of $\Delta v_{1/2}$. $\Delta v_{1/2}$ values are shown as a function of the product $p_E k_{ex}$ for two B_1 field strengths. The values of p_E and k_{ex} are indicated in the left panel. Linewidths predicted using Eq. 12 (minor state, blue) and Eq. 21 (major state, orange) are also shown as solid lines. (**b**) Simulated CEST profiles focusing on the major (left) and minor (right) dip regions for different B_1 field strengths, p_E = 1% and k_{ex} = 50 s^{-1}.
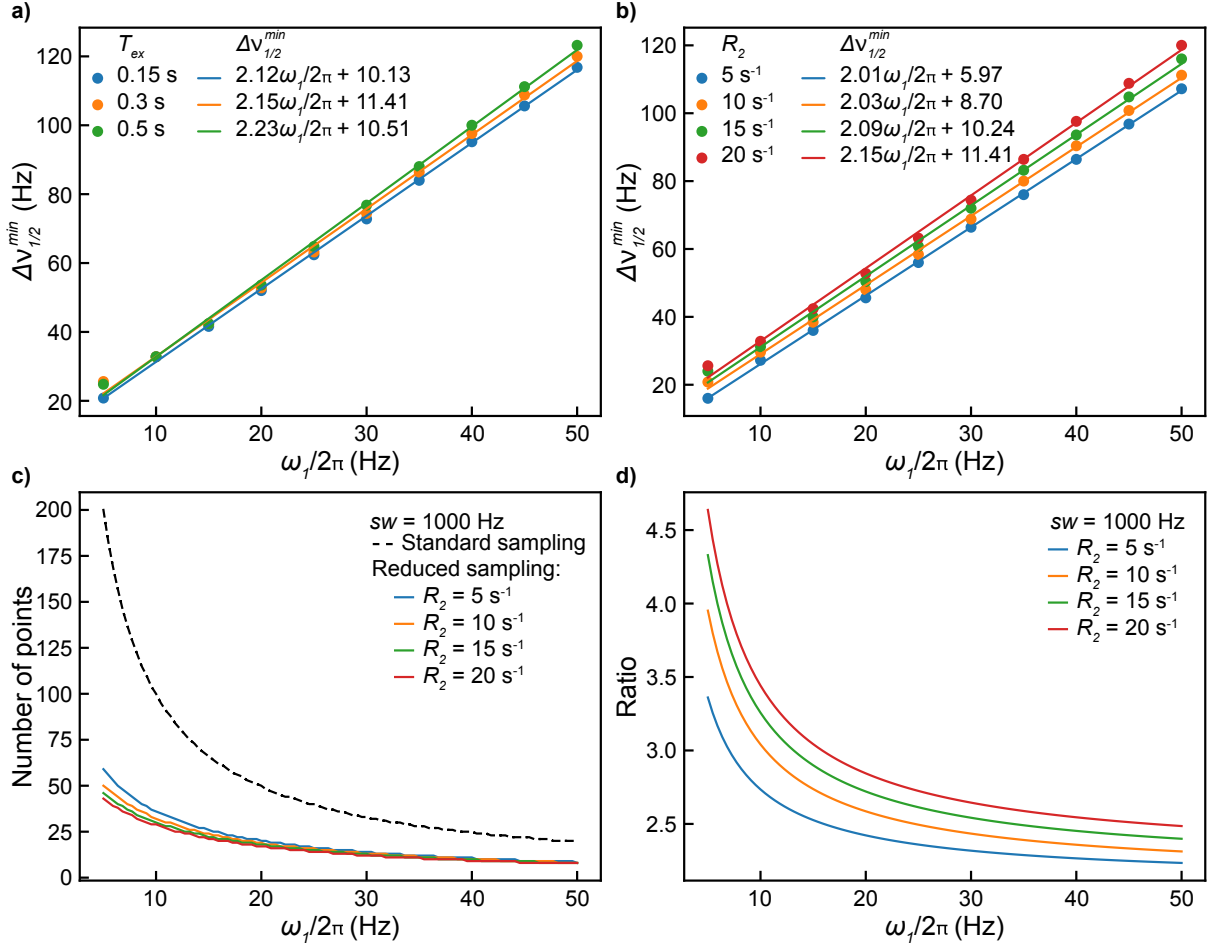
Figure 3. Optimization of the sampling frequency in CEST experiments. (**a**) Variation of $\min(\Delta v_{1/2}^G, \Delta v_{1/2}^E) = \Delta v_{1/2}^{min}$ as a function of the B_1 field strength $(\frac{\omega_1}{2\pi})$ measured from dips in simulated CEST profiles with $k_{ex} = 50$ s^{-1}, $p_E = 1\%$, $R_{2,G} = R_{2,E} = 20$ s^{-1}, and several typical T_{ex} values. As discussed in the text, the exchange parameters used are lower bounds for values typically fit via CEST and as such the obtained $\Delta v_{1/2}^{min}$ is a conservative estimate of the minimum dip linewidth over a wide range of exchange parameters, as dip widths will only increase with larger k_{ex}, p_E values. Fitted values of $\Delta v_{1/2}^{min}$ are shown with solid lines whose equations are given at the top of the panel. (**b**) Variation of $\Delta v_{1/2}^{min}$ (circles) as a function of the B_1 field strength from fits of CEST profiles simulated with $k_{ex} = 50$ s^{-1}, $p_E = 1\%$, $T_{ex} = 300$ ms, and $R_{2,G} = R_{2,E}$, as indicated. Linear fits of the linewidths are shown as solid lines; equations are given at the top of the panel for the transverse relaxation rates used. (**c**) Suggested number of CEST frequency points as a function of the B_1 field strength calculated using the expressions for $\Delta v_{1/2}^{min}$ listed in panel (b) along with Eq. 24 for a CEST spectral width of 1 kHz (solid lines). Also shown is the number of CEST points typically used where the frequency domain is sampled every $(\frac{\omega_1}{2\pi})$ Hz as a function of the B_1 field strength (dashed line). (**d**) Ratios of the number of points using the standard and reduced sampling approaches as a function of $(\frac{\omega_1}{2\pi})$ for $R_{2,G} = R_{2,E}$ values as shown (Ratio = standard/reduced), highlighting the considerable decrease in the numbers of frequency domain sampled points and hence decreased measurement times associated with the reduced sampling method.
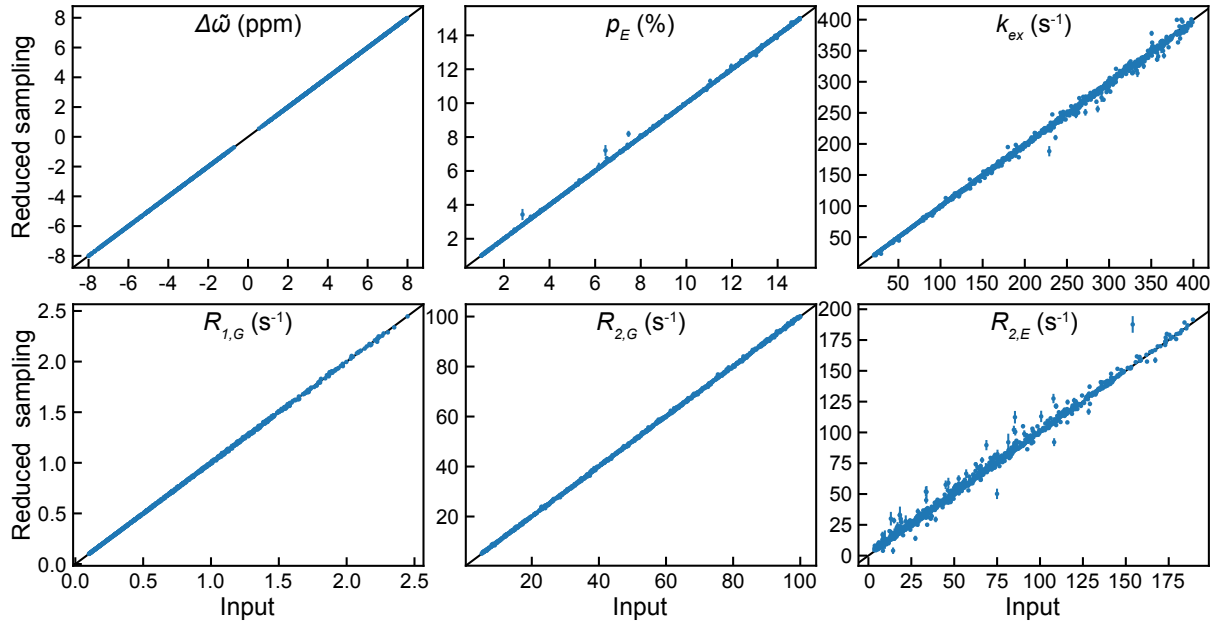
Figure 4. Linear correlation plots of exchange parameters fit from simulated CEST profiles generated as described in Materials and Methods and discussed in the text, using a reduced sampling approach *vs* input parameters. The relation, $\Delta v_{1/2}^{min}(Hz) = 2.15\frac{\omega_1}{2\pi} + 11.41$, was used along with Eq. 24 to obtain the number of CEST points and hence the frequency spacing. Each correlation panel is based on analysis of 1000 profiles (circles; a pair of CEST curves with different B_1 values for each profile were fit). The y=x line is shown in black. A two-site exchange model was used both for generating the CEST profiles and for fitting the data; parameters are listed in Table 1. Only profiles for which $|\Delta\varpi| > 0.5$ ppm were considered (see Figure S3 for the case where $|\Delta\varpi| \leq 0.5$ ppm).
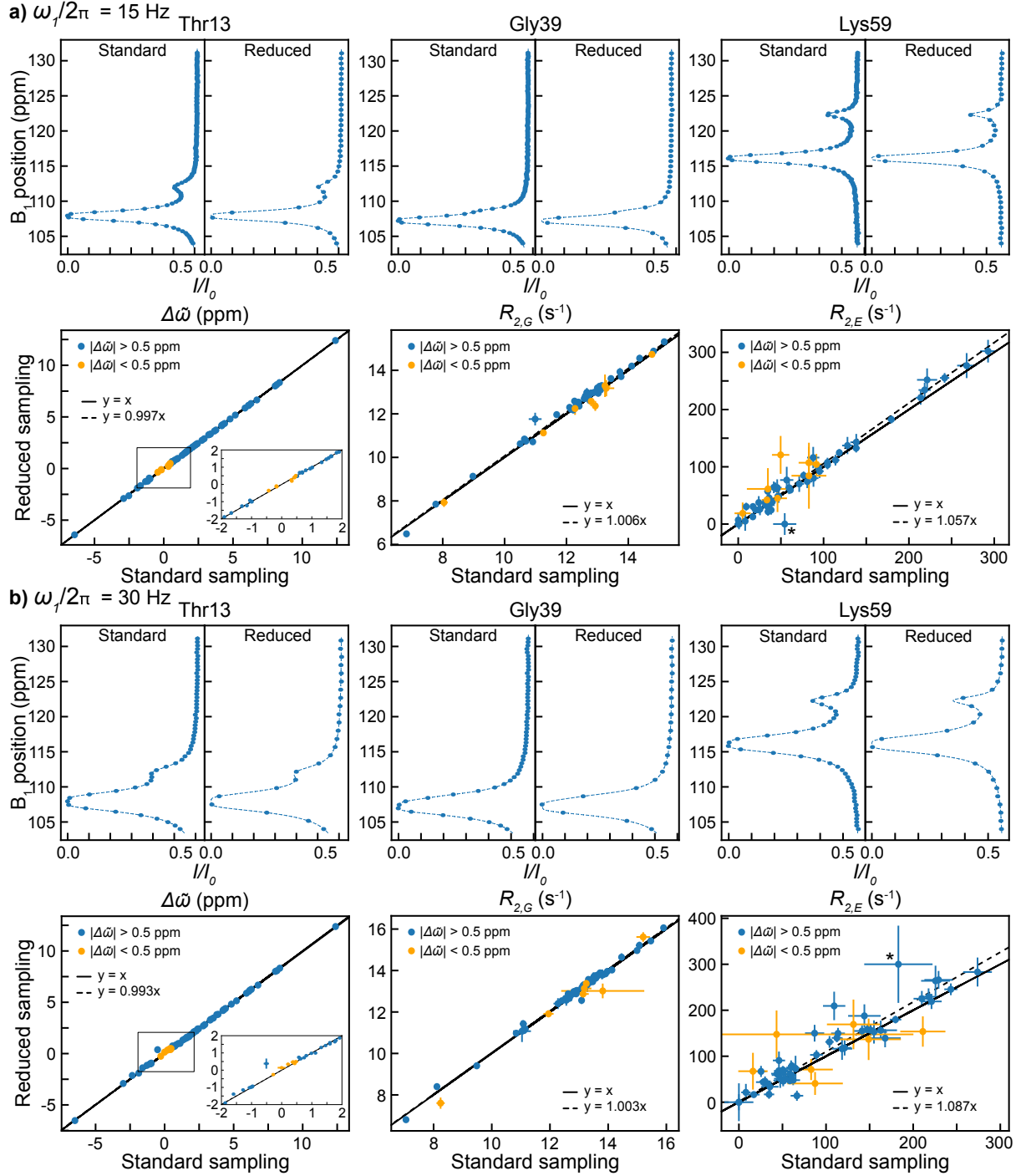
Figure 5. Experimental validation of the reduced sampling approach. Representative experimental CEST curves for A39G FF recorded with B_1 fields of 15 Hz (**a**) and 30 Hz (**b**), 600MHz, 2°C. Profiles for Thr13, Gly39 and Lys59 are shown in the top panels (circles) and best fits indicated (dashed lines) based on a global analysis of the data. The reduced datasets were sampled using the relation $\Delta v_{1/2}^{min}(Hz) = 2.15\frac{\omega_1}{2\pi} + 11.41$. Bottom panels show linear correlation plots for extracted $(\Delta\varpi, R_{2,G}, R_{2,E})$ values from profiles generated with reduced and standard sampling schemes. Residues for which $|\Delta\tilde{\omega}| < 0.5$ ppm are indicated as orange circles; insets in the $\Delta\varpi$ profiles show expanded correlations, highlighting that accurate chemical shift differences can be extracted even when they are small. The outlier in the $R_{2,E}$ correlation plots (*) derives from Tyr49 (see Figure S4). Note the increased uncertainties in $R_{2,E}$ values extracted from profiles recorded with either reduced or standard sampling approaches

with B_1 = 30 Hz (reduced χ^2 = 4.8 relative to 3.4 with B_1 = 15 Hz, $\chi^2 = \frac{1}{N}\sum_{n=1}^{N}\frac{(R_{2,E}^{red,n}-R_{2,E}^{std,n})^2}{\sigma_{red,n}\sigma_{std,n}}$ where the sum runs over all residues, $R_{2,E}^{red,n}$ and $R_{2,E}^{std,n}$ are the transverse relaxation rates of the excited state obtained from the analysis of the reduced and standard sampling datasets, respectively, with error $\sigma_{red,n}$ and $\sigma_{std,n}$ respectively, for residue n), as the dips are significantly more broadened when larger B_1 fields are used.
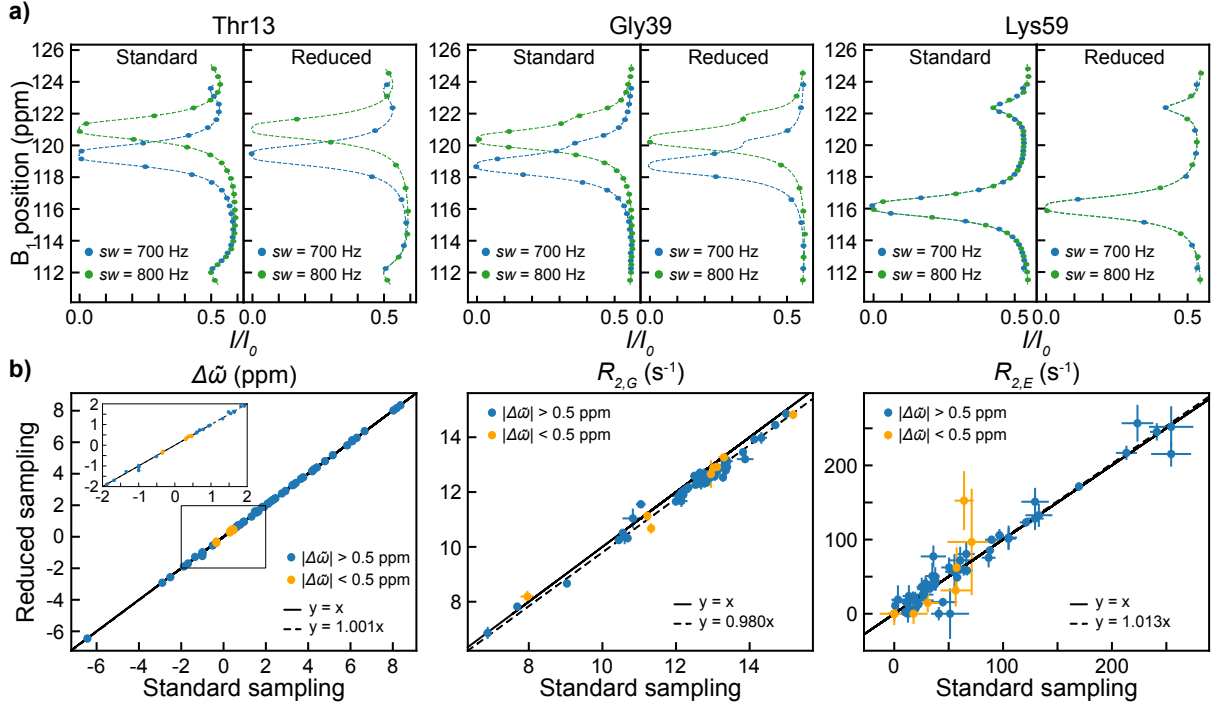

Figure 6. Experimental validation of the reduced sampling approach for D-CEST experiments. (**a**) Representative D-CEST profiles, highlighting Thr13, Gly39 and Lys59 from A39G FF, using spectral widths of 700 Hz (blue) and 800 Hz (green) and a B_1 field strength of 15 Hz (600 MHz, 2°C). Best fits of the experimental data (circles) are shown with the dash lines. The CEST data in each profile were sampled either (1) every $2x\frac{\omega_1}{2\pi}$ Hz (standard; recall that two datasets are recorded with sampled frequencies in each dataset located exactly between each other, so that the points in a "combined" profile are separated by $\frac{\omega_1}{2\pi}$ Hz; see text) or (2) using the relation $\Delta v_{1/2}^{min}(Hz) = 2.15\frac{\omega_1}{2\pi} + 11.41$ and Eq. 24, with the distance between sampled frequency points doubled, ensuring that the frequencies sampled in each of the pair of datasets are located between each other. Note that the profiles for Thr13 and Gly39 are aliased (*i.e.*, not within the 700 or 800 Hz windows used), while those for Lys59 are not. (**b**) Correlation plots for extracted ($\Delta\varpi, R_{2,G}, R_{2,E}$) as in Figure 5.
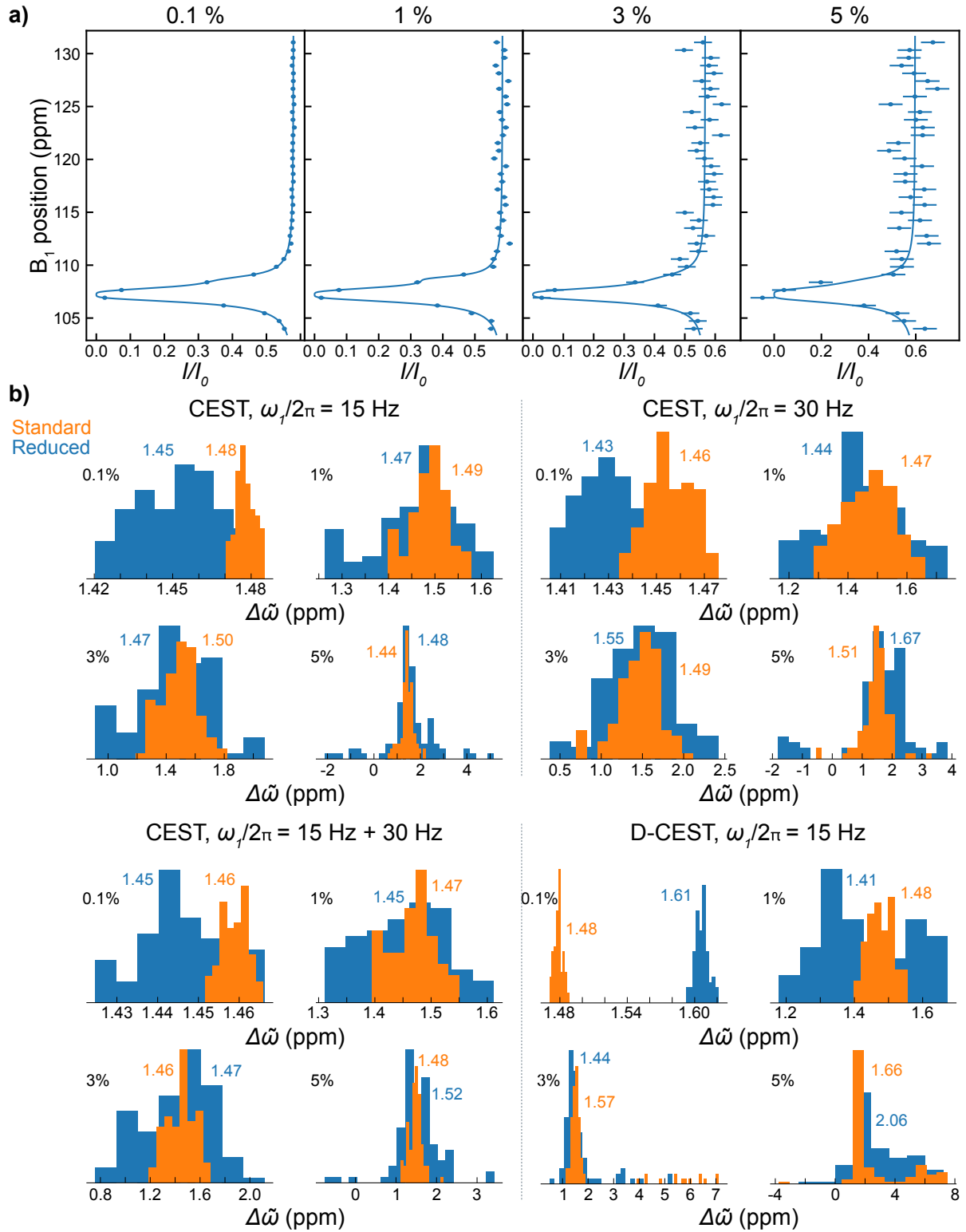
Figure 7. Analysis of CEST profiles for Gly39 of A39G FF (600 MHz, 2ºC) as a function of increasing noise levels. CEST and D-CEST profiles were re-analyzed in *Chemex* after increasing the experimental noise to 0.1, 1, 3 and 5% of the intensity of the Gly39 peak in the reference spectrum (I_o), as described in the text. For each noise level, 100 profiles were generated and analyzed. (**a**) Examples of CEST profiles (circles) at various noise levels, as indicated. The "original" profile was recorded with a B_1 field of 15 Hz and reduced sampling. Fits of

the data are indicated with solid lines. Shown in (**b**) are distributions of $\Delta\varpi$ values in histogram format for each error level and for datasets acquired using standard and reduced frequency spacings, as described in the text. Only $\Delta\varpi$ values within $\overline{\Delta\widetilde{\omega}} \pm 2\sigma$ are shown, along with $\overline{\Delta\widetilde{\omega}}$, the median difference in shifts.

Table 1. Parameter ranges used for simulations of CEST profiles. Distributions of parameters are uniform.

	Lower limit	Higher limit
B_0	500 MHz	1.2 GHz
p_E	1%	15%
k_{ex}	20 s^{-1}	400 s^{-1}
Spectral width (*sw*)	2,000 Hz	6,000 Hz
$\frac{\omega_G}{2\pi}$	-0.8 x *sw*	+0.8 x *sw*
$\Delta\varpi$	-8 ppm	8 ppm
R_{1G}	0.1 (s^{-1})	10 (s^{-1})
R_{1E}	0.9 x R_{1G}	1.1 x R_{1G}
R_{2G}	5 (s^{-1})	100 (s^{-1})
R_{2E}	0.5 x R_{2G}	2 x R_{2G}

Table 2. Values of N and extracted exchange parameters obtained from a global analysis of experimental CEST profiles. N does not include the reference plane (1 for CEST, 2 for D-CEST). Errors are obtained from *Chemex* fits.

	CEST, $\frac{\omega_1}{2\pi} = 15$ Hz		CEST, $\frac{\omega_1}{2\pi} = 30$ Hz		D-CEST, $\frac{\omega_1}{2\pi} = 15$ Hz	
	Standard	Reduced	Standard	Reduced	Standard	Reduced
N	111	38	56	24	52	19
p_E (%)	1.30 ± 0.01	1.33 ± 0.02	0.98 ± 0.01	1.04 ± 0.02	1.29 ± 0.02	1.30 ± 0.03
k_{ex} (s^{-1})	97 ± 2	90 ± 4	164 ± 6	144 ± 8	100 ± 3	100 ± 5

References

Alderson, T.R., Kay, L.E., 2021. NMR spectroscopy captures the essential role of dynamics in regulating biomolecular function. Cell 184, 577–595. https://doi.org/10.1016/j.cell.2020.12.034

Anthis, N.J., Clore, G.M., 2015. Visualizing transient dark states by NMR spectroscopy. Quart. Rev. Biophys. 48, 35–116. https://doi.org/10.1017/S0033583514000122

Bodenhausen, G., Freeman, R., Morris, G.A., 1976. A simple pulse sequence for selective excitation in Fourier transform NMR. Journal of Magnetic Resonance (1969) 23, 171–175. https://doi.org/10.1016/0022-2364(76)90150-5

Boehr, D.D., McElheny, D., Dyson, H.J., Wright, P.E., 2006. The Dynamic Energy Landscape of Dihydrofolate Reductase Catalysis. Science 313, 1638–1642. https://doi.org/10.1126/science.1130258

Bouvignies, G., Vallurupalli, P., Hansen, D.F., Correia, B.E., Lange, O., Bah, A., Vernon, R.M., Dahlquist, F.W., Baker, D., Kay, L.E., 2011. Solution structure of a minor and transiently formed state of a T4 lysozyme mutant. Nature 477, 111–114. https://doi.org/10.1038/nature10349

Carneiro, M.G., Reddy, J.G., Griesinger, C., Lee, D., 2015. Speeding-up exchange-mediated saturation transfer experiments by Fourier transform. J Biomol NMR 63, 237–244. https://doi.org/10.1007/s10858-015-9985-9

Delaglio, F., Grzesiek, S., Vuister, GeertenW., Zhu, G., Pfeifer, J., Bax, A., 1995. NMRPipe: A multidimensional spectral processing system based on UNIX pipes. Journal of Biomolecular NMR 6. https://doi.org/10.1007/BF00197809

Forsén, S., Hoffman, R.A., 1963. Study of Moderately Rapid Chemical Exchange Reactions by Means of Nuclear Magnetic Double Resonance. The Journal of Chemical Physics 39, 2892–2901. https://doi.org/10.1063/1.1734121

Guenneugues, M., Berthault, P., Desvaux, H., 1999. A Method for DeterminingB1Field Inhomogeneity. Are the Biases Assumed in Heteronuclear Relaxation Experiments Usually Underestimated? Journal of Magnetic Resonance 136, 118–126. https://doi.org/10.1006/jmre.1998.1590

Henzler-Wildman, K., Kern, D., 2007. Dynamic personalities of proteins. Nature 450, 964–972. https://doi.org/10.1038/nature06522

Hyberts, S.G., Heffron, G.J., Tarragona, N.G., Solanky, K., Edmonds, K.A., Luithardt, H., Fejzo, J., Chorev, M., Aktas, H., Colson, K., Falchuk, K.H., Halperin, J.A., Wagner, G., 2007. Ultrahigh-Resolution ^{1}H−^{13}C HSQC Spectra of Metabolite Mixtures Using Nonlinear Sampling and Forward Maximum Entropy Reconstruction. J. Am. Chem. Soc. 129, 5108–5116. https://doi.org/10.1021/ja068541x

Jameson, G., Hansen, A.L., Li, D., Bruschweiler-Li, L., Brüschweiler, R., 2019. Extreme Nonuniform Sampling for Protein NMR Dynamics Studies in Minimal Time. J. Am. Chem. Soc. 141, 16829–16838. https://doi.org/10.1021/jacs.9b08032

Jemth, P., Day, R., Gianni, S., Khan, F., Allen, M., Daggett, V., Fersht, A.R., 2005. The Structure of the Major Transition State for Folding of an FF Domain from Experiment and Simulation. Journal of Molecular Biology 350, 363–378. https://doi.org/10.1016/j.jmb.2005.04.067

Jemth, P., Gianni, S., Day, R., Li, B., Johnson, C.M., Daggett, V., Fersht, A.R., 2004. Demonstration of a low-energy on-pathway intermediate in a fast-folding protein by kinetics, protein engineering, and simulation. Proceedings of the National Academy of Sciences 101, 6450–6455. https://doi.org/10.1073/pnas.0401732101

Karplus, M., Kuriyan, J., 2005. Molecular dynamics and protein function. Proc. Natl. Acad. Sci. U.S.A. 102, 6679–6685. https://doi.org/10.1073/pnas.0408930102

Karunanithy, G., Yuwen, T., Kay, L.E., Hansen, D.F., 2021. Towards autonomous analysis of Chemical Exchange Saturation Transfer experiments using Deep Neural Networks (preprint). Chemistry. https://doi.org/10.26434/chemrxiv-2021-r1cmw

Kimsey, I.J., Szymanski, E.S., Zahurancik, W.J., Shakya, A., Xue, Y., Chu, C.-C., Sathyamoorthy, B., Suo, Z., Al-Hashimi, H.M., 2018. Dynamic basis for dG•dT misincorporation via tautomerization and ionization. Nature 554, 195–201. https://doi.org/10.1038/nature25487

Korzhnev, D.M., Religa, T.L., Banachewicz, W., Fersht, A.R., Kay, L.E., 2010. A Transient and Low-Populated Protein-Folding Intermediate at Atomic Resolution. Science 329, 1312–1316. https://doi.org/10.1126/science.1191723

Korzhnev, D.M., Religa, T.L., Lundström, P., Fersht, A.R., Kay, L.E., 2007. The Folding Pathway of an FF domain: Characterization of an On-pathway Intermediate State Under Folding Conditions by 15N, 13Cα and 13C-methyl Relaxation Dispersion and 1H/2H-exchange NMR Spectroscopy. Journal of Molecular Biology 372, 497–512. https://doi.org/10.1016/j.jmb.2007.06.012

Korzhnev, D.M., Salvatella, X., Vendruscolo, M., Di Nardo, A.A., Davidson, A.R., Dobson, C.M., Kay, L.E., 2004. Low-populated folding intermediates of Fyn SH3 characterized by relaxation dispersion NMR. Nature 430, 586–590. https://doi.org/10.1038/nature02655

Lee, W., Tonelli, M., Markley, J.L., 2015. NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy. Bioinformatics 31, 1325–1327. https://doi.org/10.1093/bioinformatics/btu830

Leninger, M., Marsiglia, W.M., Jerschow, A., Traaseth, N.J., 2018. Multiple frequency saturation pulses reduce CEST acquisition time for quantifying conformational exchange in biomolecules. J Biomol NMR 71, 19–30. https://doi.org/10.1007/s10858-018-0186-1

Long, D., Delaglio, F., Sekhar, A., Kay, L.E., 2015. Probing Invisible, Excited Protein States by Non-Uniformly Sampled Pseudo-4D CEST Spectroscopy. Angew. Chem. Int. Ed. 54, 10507–10511. https://doi.org/10.1002/anie.201504070

McConnell, H.M., 1958. Reaction Rates by Nuclear Magnetic Resonance. The Journal of Chemical Physics 28, 430–431. https://doi.org/10.1063/1.1744152

Mittermaier, A., Kay, L.E., 2006. New Tools Provide New Insights in NMR Studies of Protein Dynamics. Science 312, 224–228. https://doi.org/10.1126/science.1124964

Neudecker, P., Robustelli, P., Cavalli, A., Walsh, P., Lundström, P., Zarrine-Afsar, A., Sharpe, S., Vendruscolo, M., Kay, L.E., 2012. Structure of an Intermediate State in Protein Folding and Aggregation. Science 336, 362–366. https://doi.org/10.1126/science.1214203

Orekhov, V.Yu., Jaravine, V.A., 2011. Analysis of non-uniformly sampled spectra with multi-dimensional decomposition. Progress in Nuclear Magnetic Resonance Spectroscopy 59, 271–292. https://doi.org/10.1016/j.pnmrs.2011.02.002

Palmer, A.G., 2014. Chemical exchange in biomacromolecules: Past, present, and future. Journal of Magnetic Resonance 241, 3–17. https://doi.org/10.1016/j.jmr.2014.01.008

Palmer, A.G., Grey, M.J., Wang, C., 2005. Solution NMR Spin Relaxation Methods for Characterizing Chemical Exchange in High-Molecular-Weight Systems, in: Methods in Enzymology. Elsevier, pp. 430–465. https://doi.org/10.1016/S0076-6879(05)94018-4

SciPy 1.0 Contributors, Virtanen, P., Gommers, R., Oliphant, T.E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S.J., Brett, M., Wilson, J., Millman, K.J., Mayorov, N., Nelson, A.R.J., Jones, E., Kern, R., Larson, E., Carey, C.J., Polat, İ., Feng, Y., Moore, E.W., VanderPlas, J.,

Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E.A., Harris, C.R., Archibald, A.M., Ribeiro, A.H., Pedregosa, F., van Mulbregt, P., 2020. SciPy 1.0: fundamental algorithms for scientific computing in Python. Nature Methods 17, 261–272. https://doi.org/10.1038/s41592-019-0686-2

Stiller, J.B., Otten, R., Häussinger, D., Rieder, P.S., Theobald, D.L., Kern, D., 2022. Structure determination of high-energy states in a dynamic protein ensemble. Nature 603, 528–535. https://doi.org/10.1038/s41586-022-04468-9

Tiwari, V.P., Toyama, Y., De, D., Kay, L.E., Vallurupalli, P., 2021. The A39G FF domain folds on a volcano-shaped free energy surface via separate pathways. Proceedings of the National Academy of Sciences 118, e2115113118. https://doi.org/10.1073/pnas.2115113118

Trott, O., Palmer, A.G., 2002. R1ρ Relaxation outside of the Fast-Exchange Limit. Journal of Magnetic Resonance 154, 157–160. https://doi.org/10.1006/jmre.2001.2466

Vallurupalli, P., Bouvignies, G., Kay, L.E., 2012. Studying "Invisible" Excited Protein States in Slow Exchange with a Major State Conformation. Journal of the American Chemical Society 134, 8148–8161. https://doi.org/10.1021/ja3001419

Vallurupalli, P., Hansen, D.F., Kay, L.E., 2008. Structures of invisible, excited protein states by relaxation dispersion NMR spectroscopy. Proc. Natl. Acad. Sci. U.S.A. 105, 11766–11771. https://doi.org/10.1073/pnas.0804221105

Vallurupalli, P., Hansen, D.F., Lundström, P., Kay, L.E., 2009. CPMG relaxation dispersion NMR experiments measuring glycine 1Hα and 13Cα chemical shifts in the 'invisible' excited states of proteins. J Biomol NMR 45, 45–55. https://doi.org/10.1007/s10858-009-9310-6

Vallurupalli, P., Kay, L.E., 2013. Probing Slow Chemical Exchange at Carbonyl Sites in Proteins by Chemical Exchange Saturation Transfer NMR Spectroscopy. Angew. Chem. Int. Ed. 52, 4156–4159. https://doi.org/10.1002/anie.201209118

Vallurupalli, P., Sekhar, A., Yuwen, T., Kay, L.E., 2017. Probing conformational dynamics in biomolecules via chemical exchange saturation transfer: a primer. Journal of Biomolecular NMR 67, 243–271. https://doi.org/10.1007/s10858-017-0099-4

van Zijl, P.C.M., Yadav, N.N., 2011. Chemical exchange saturation transfer (CEST): What is in a name and what isn't?: CEST: What is in a Name and What Isn't? Magn. Reson. Med. 65, 927–948. https://doi.org/10.1002/mrm.22761

Ward, K.M., Aletras, A.H., Balaban, R.S., 2000. A New Class of Contrast Agents for MRI Based on Proton Chemical Exchange Dependent Saturation Transfer (CEST). Journal of Magnetic Resonance 143, 79–87. https://doi.org/10.1006/jmre.1999.1956

Yuwen, T., Bouvignies, G., Kay, L.E., 2018a. Exploring methods to expedite the recording of CEST datasets using selective pulse excitation. Journal of Magnetic Resonance 292, 1–7. https://doi.org/10.1016/j.jmr.2018.04.013

Yuwen, T., Kay, L.E., Bouvignies, G., 2018b. Dramatic Decrease in CEST Measurement Times Using Multi-Site Excitation. ChemPhysChem 19, 1707–1710. https://doi.org/10.1002/cphc.201800249