J Biomol NMR (2012) 53:125–138
DOI 10.1007/s10858-012-9630-9

ARTICLE

Increased precision for analysis of protein–ligand dissociation constants determined from chemical shift titrations

Craig J. Markin · Leo Spyracopoulos

Received: 27 January 2012 / Accepted: 11 April 2012 / Published online: 26 April 2012

Abstract NMR is ideally suited for the analysis of protein–protein and protein ligand interactions with dissociation constants ranging from $\sim$2 μM to $\sim$1 mM, and with kinetics in the fast exchange regime on the NMR timescale. For the determination of dissociation constants (K_D) of 1:1 protein–protein or protein–ligand interactions using NMR, the protein and ligand concentrations must necessarily be similar in magnitude to the K_D, and nonlinear least squares analysis of chemical shift changes as a function of ligand concentration is employed to determine estimates for the parameters K_D and the maximum chemical shift change ($\Delta\delta_{max}$). During a typical NMR titration, the initial protein concentration, [P_0], is held nearly constant. For this condition, to determine the most accurate parameters for K_D and $\Delta\delta_{max}$ from nonlinear least squares analyses requires initial protein concentrations that are $\sim 0.5 \times K_D$, and a maximum concentration for the ligand, or titrant, of $\sim 10 \times [P_0]$. From a practical standpoint, these requirements are often difficult to achieve. Using Monte Carlo simulations, we demonstrate that co-variation of the ligand and protein concentrations during a titration leads to an increase in the precision of the fitted K_D and $\Delta\delta_{max}$ values when $[P_0] > K_D$. Importantly, judicious choice of protein and ligand concentrations for a given NMR titration, combined with nonlinear least squares analyses using two independent variables (ligand and protein concentrations) and two parameters (K_D and $\Delta\delta_{max}$) is a straightforward approach to increasing the accuracy of measured dissociation constants for 1:1 protein–ligand interactions.

C. J. Markin · L. Spyracopoulos (✉)
Department of Biochemistry, University of Alberta, Edmonton, AB T6G 2H7, Canada
e-mail: leo.spyracopoulos@ualberta.ca

Keywords Chemical shift titration · Dissociation constant · Protein–protein interaction · Protein–ligand interaction

Introduction

Biological processes are driven by molecular recognition events involving protein–protein and protein–ligand interactions with dissociation constants (K_D) in the μM to mM range and kinetics that span the slow to fast NMR timescales (Zuiderweg 2002; Rintala-Dempsey et al. 2006; Fielding 2007; Haririnia et al. 2007; Marintchev et al. 2007; Baryshnikova et al. 2008; Markin et al. 2010a, b). For kinetics in the intermediate/fast exchange regimes, the observed NMR resonances for a protein being titrated with cognate binding partner represent weighted averages between the free and bound states of the protein. For 1:1 binding, the observed chemical shift changes follow a hyperbolic dependence on ligand concentration. In general, parameters for the maximum chemical shift change and K_D are estimated from nonlinear least squares fits of observed chemical shift changes for the protein as a function of ligand concentration during protein–ligand titrations. To obtain accurate values for K_D parameters derived from nonlinear least squares fits of 1:1 binding isotherms, statistical analyses of simulated experiments have established that the optimal protein concentration ([P_0]) should be held at $0.5 \times K_D$, and the ligand concentration should be varied between $\sim 0.4 \times [P_0]$ and $\sim 11 \times [P_0]$, with 15–20 repetitions of the titration (Granot 1983). From a practical standpoint, [P_0] is usually between 50 and 500 μM in order to achieve sufficient signal to noise ratio when employing standard ^{1}H–^{15}N or ^{1}H–^{13}C 2D NMR spectroscopic techniques. Thus, to conduct a titration for [P_0] $\sim$500 μM, and K_D $\sim$200 μM, the ligand

concentration must reach 2 mM during a titration in order to achieve reasonable accuracy, that is, the binding site on the protein must be saturated with ligand. Without prior knowledge of the actual K_D, it is difficult to choose the optimal value for $[P_0]$, and larger $[P_0]$ values are more desirable as they yield higher signal to noise ratios in NMR spectra. Another requirement to achieve accuracy is that the binding site should be saturated during the course of the titration; this requires high ligand concentrations, a condition which is often difficult to satisfy for larger K_D values. For example, if the ligand is another protein, it may be difficult to prepare in sufficient quantity, or it may not be soluble at higher concentrations. Furthermore, it is difficult to conduct a titration such that the concentration of the fixed protein component is not significantly altered. Ideally, this requires making separate NMR samples for each titration point from single stock solutions of protein and ligand. This is advantageous with respect to controlling buffer conditions, but problematic due to the large volumes of stock solutions required. Whilst there is no simple solution to address the difficulties associated with the experimental design of NMR-monitored titrations, we have devised two straightforward approaches for optimizing the experimental design of protein–ligand titrations such that significant increases in precision can be achieved. Conducting protein ligand titrations whereby the concentration of protein component is allowed to co-vary with the ligand concentration, and employing nonlinear least squares analyses with both ligand and protein concentrations as independent variables and K_D and $\Delta\delta_{\max}$ as variable parameters, can allow for a significant increase in the precision of K_D compared to the precision obtained at fixed $[P_0]$, particularly when $[P_0]$ exceeds K_D. We have analyzed two methods by which co-variation of protein and ligand concentration can be used to achieve increased precision for fitted K_D values. The first method involves addition of aliquots of ligand solution to a protein solution such that the protein concentration decreases by a constant factor as more ligand is added. The second method involves the serial dilution of a solution containing concentrated protein and concentrated ligand in a 1:2 ratio, respectively. The practicality of both methods was demonstrated by conducting 2D ^{1}H–^{15}N HSQC NMR chemical shift titration experiments for the interaction of human [*U*–^{15}N]-Mms2 (145 residues) with human ubiquitin (76 residues).

Theory and methods

1:1 Binding isotherms

At equilibrium, the relationship between the concentrations of protein, ligand, protein–ligand complex, and the equilibrium dissociation constant is given by:

$$K_D = \frac{[P][L]}{[PL]} \tag{1}$$

with the total protein and ligand concentrations:

$$P_T = [P] + [PL] \tag{2}$$

$$L_T = [L] + [PL] \tag{3}$$

Equations 1–3 can be solved to yield the following expression for the bound protein complex:

$$[PL] = \frac{1}{2}(K_D + L_T + P_T) - \sqrt{(-K_D - L_T - P_T)^2 - 4L_T P_T} \tag{4}$$

In the limit of fast exchange on the NMR timescale, the observed chemical shift changes for ^{13}C or ^{15}N labeled protein ($[P_T]$) being titrated with ligand are given by:

$$\Delta\delta_{obs} = f_b \Delta\delta_{\max} \tag{5}$$

where f_b is the fraction of bound protein, $[PL]/[P_T]$, and $\Delta\delta_{\max}$ is the maximum chemical shift change for a given residue given by the difference $\delta_{\text{free}} - \delta_{\text{bound}}$, where δ_{free} is the chemical shift for the free state of the protein being observed, and δ_{bound} is the chemical shift for protein being observed in the fully bound state. Thus Eq. 4 becomes:

$$\Delta\delta_{obs} = \frac{\Delta\delta_{\max}}{2[P_T]}\left[(K_D + L_T + P_T) - \sqrt{(-K_D - L_T - P_T)^2 - 4L_T P_T}\right] \tag{6}$$

Equation 6 is applicable in the fast exchange regime, defined by $k_{ex} \gg |\Delta\omega|$, and is roughly valid for exchange rates approaching the intermediate exchange regime, given by $k_{ex} \sim |\Delta\omega|$ (Cavanagh 2007). The precise limit for the applicability of Eq. 6 is case-specific, and depends on the magnitude of the line broadening.

Monte Carlo error analysis for simulated titrations

Protein chemical shift changes as a function of ligand concentration can be fit to Eq. 6 using nonlinear least squares methods with K_D and $\Delta\delta_{\max}$ as variable parameters. Typically, the ligand concentration is treated as an independent variable, whereas the initial protein concentration is treated as a constant. However, there is no a priori reason to make this distinction. For example, Eq. 6 describes a two-dimensional binding curve with $[P_T]$ and $[L_T]$ on the x and y-axes and $\Delta\delta$ on the z-axis. If this surface varies as a function of $[P_T]$, then it stands to reason that greater accuracy for a single titration can be achieved by sampling the surface more extensively with respect to $[P_T]$. To that end, we sought to determine the errors associated with fitting chemical shift data to Eq. 6 with K_D and $\Delta\delta_{\max}$

as adjustable parameters and $[P_T]$ and $[L_T]$ as independent variables. We assumed that the two main sources of error in a given titration arise from errors in $\Delta\delta_{obs}$ and the errors in the starting protein and ligand concentrations. Provided that titration points involve addition of a single stock solution of ligand, dilution, or a combination thereof, then the largest concentration error involves only the starting concentrations of protein and ligand, as subsequent errors due to pipetting/dilution are small.

Two specific types of titrations were simulated with the program *Mathematica* 8.0.4 (Wolfram 1999) using different concentrations of protein and ligand and seven different values for K_D, ranging from 2 to 2,000 μM. The first, Method 1, involves simulation of the standard method of conducting a titration, that is, addition of aliquots of ligand solution to a protein solution, with the exception that the protein concentration is allowed to decrease as more ligand is added (Table 1). The second type of simulated titration, Method 2, was one in which a solution of concentrated protein and concentrated ligand was serially diluted to produce a binding isotherm (Table 2). For Method 1, one thousand simulated data sets for titrations were generated for seven different values K_D (2, 20, 60, 200, 600, 1,000, and 2,000 μM), each with $\Delta\delta_{max} = 1$ ppm. The initial protein concentrations were chosen randomly from a normal distribution with mean 100 μM (Table 1, Case 1) or 500 μM (Table 1, Case 2) and a standard deviation of ± 5 %; likewise, the initial ligand concentrations were chosen randomly from a normal distribution with mean 0.05 mM and a standard deviation of ± 5 %, and increased by factors of 2, 3, 4, 5, 6, 7, 8, 12, and 16 to yield the ligand concentration for subsequent titration points (Table 1). The protein and ligand concentrations, as well as $\Delta\delta_{max}$ were substituted into Eq. 6 to yield an ensemble of Monte Carlo $\Delta\delta_{obs}$ values. In addition to the Monte Carlo ensembles corresponding to the seven values of K_D, two further sets of seven Monte Carlo ensembles were generated by choosing two constant factors $\Delta_{[P]} = 0.1$ and 0.25 (Table 1, Cases 3 and 4, respectively), such that the initial protein concentration $[P_0]$ decreased in the n subsequent titration points according to:

$$[P_n] = [P_{n-1}] - \Delta_{[P]} \times [P_{n-1}] \quad (7)$$

for titration points $n > 1$ and $\Delta_{[P]} \leq 1.0$.

For Method 2, the initial protein concentration $[P_0]$ was set to 500 μM, the initial ligand concentration ($[L_0]$) was set to 1 mM, seven values of K_D were chosen (2, 20, 60, 200, 600, 1,000, and 2,000 μM), with one $\Delta\delta_{max}$ (1 ppm) per K_D (Table 2). One thousand simulated data sets were generated by choosing the initial ligand and protein concentrations randomly from normal distributions with means 1 and 0.5 mM respectively, and standard deviations of ± 5 %. The protein and ligand concentrations were allowed to decrease by a constant increment of 50 μM for a total of ten titration points. These protein and ligand concentrations, along with $\Delta\delta_{max}$, were used to calculate $\Delta\delta_{obs}$ values for the Monte Carlo ensemble with Eq. 6.

Table 1 Protein and ligand concentrations for Method 1 simulations

Case	Method 1, K_D = 2, 20, 60, 200, 600, 1,000, and 2,000 μM										
1	$[P_T]$ (mM)	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
	$\Delta_{[P]} = 0$										
	$[L_T]$ (mM)	0.05	0.1	0.15	0.2	0.25	0.3	0.35	0.4	0.6	0.8
2	$[P_T]$ (mM)	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5
	$\Delta_{[P]} = 0$										
	$[L_T]$ (mM)	0.05	0.1	0.15	0.2	0.25	0.3	0.35	0.4	0.6	0.8
3	$[P_T]$ (mM)	0.5	0.45	0.405	0.364	0.328	0.295	0.266	0.239	0.215	0.194
	$\Delta_{[P]} = 0.1$										
	$[L_T]$ (mM)	0.05	0.1	0.15	0.2	0.25	0.3	0.35	0.4	0.6	0.8
4	$[P_T]$ (mM)	0.5	0.375	0.281	0.211	0.158	0.119	0.089	0.067	0.05	0.037
	$\Delta_{[P]} = 0.25$										
	$[L_T]$ (mM)	0.05	0.1	0.15	0.2	0.25	0.3	0.35	0.4	0.6	0.8

Table 2 Protein and ligand concentrations for Method 2 simulations

Method 2, K_D = 2, 20, 60, 200, 600, 1,000, and 2,000 μM										
$[P_T]$ (mM)	0.5	0.45	0.4	0.35	0.3	0.25	0.2	0.15	0.1	0.05
$[L_T]$ (mM)	1	0.9	0.8	0.7	0.6	0.5	0.4	0.3	0.2	0.1

The Monte Carlo ensembles of titration data were fit to Eq. 6 using the "NonlinearModelFit" nonlinear least squares regression package within the program *Mathematica* 8.04 (Wolfram 1999). The "NMinimize" and "DifferentialEvolution" options with the default parameters were chosen for the constrained nonlinear optimization algorithm. The resulting ensembles for the K_D and $\Delta\delta_{\max}$ parameters were either normal, i.e. Gaussian distributions, or gamma distributions. The normal distributions were fit to the probability density function:

$$P(x) = A_0 \frac{e^{-(x-\mu)^2/(2\sigma^2)}}{2\sigma^2} \tag{8}$$

where A_0 is a scaling parameter. The gamma distributions were fit to the following probability density function:

$$P(x) = A_0 \frac{e^{-x/\beta} x^{\alpha-1} \beta^{-\alpha}}{\Gamma(\alpha)} \tag{9}$$

where A_0 is a scaling parameter and

$$\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{-t} dt \tag{10}$$

For the normal and gamma distributions, the fitted probability distribution functions were used to calculate the probability of observing a given parameter to within a given error threshold. This approach is useful for gamma probability distribution functions that are typically asymmetric with respect to their median value. The Monte Carlo ensembles for K_D and $\Delta\delta_{\max}$ were binned by empirically adjusting the parameters for the *Mathematica* command "BinCounts" such that the output closely matched the default settings for the "Histogram" command given with the option "Probability".

^{15}N lineshape analyses for simulated titrations

Free induction decays (FIDs) for the various protein–ligand ratios of the simulated titrations were generated from the Bloch-McConnell equations for two-site chemical exchange (Palmer, Kroenke and Loria 2001), wherein the time course for transverse magnetization, or the FID, is given by:

$$\begin{aligned} M_A(t) &= M_A(0) a_{11}(t) \\ M_B(t) &= M_B(0) a_{22}(t) \end{aligned} \tag{11}$$

for two spins (A and B), with the coefficients for the autopeaks:

where R^0_{2A} and R^0_{2B} are the transverse relaxation rates for spins A and B, respectively, in the absence of chemical exchange, $\Delta\omega$ is the difference between the chemical shifts of spin A (Ω_A) and spin B (Ω_B) in rad s^{-1}, k_{ex} is the rate of chemical exchange, $k_{ex} = k_{on}[B] + k_{off}$, p_A and p_B are the populations of spin A and B, respectively, and

$$a_{11}(t) = \frac{1}{2}\left[\begin{array}{l} \left(1 - \dfrac{-i\Delta\omega + R^0_{2A} - R^0_{2B} + k_{ex}(p_B - p_A)}{\lambda_+ - \lambda_-}\right) \exp(-\lambda_- t) \\ + \left(1 + \dfrac{-i\Delta\omega + R^0_{2A} - R^0_{2B} + k_{ex}(p_B - p_A)}{\lambda_+ - \lambda_-}\right) \exp(-\lambda_+ t) \end{array}\right] \tag{12}$$

$$a_{22}(t) = \frac{1}{2}\left[\begin{array}{l} \left(1 + \dfrac{-i\Delta\omega + R^0_{2A} - R^0_{2B} + k_{ex}(p_B - p_A)}{\lambda_+ - \lambda_-}\right) \exp(-\lambda_- t) \\ + \left(1 - \dfrac{-i\Delta\omega + R^0_{2A} - R^0_{2B} + k_{ex}(p_B - p_A)}{\lambda_+ - \lambda_-}\right) \exp(-\lambda_+ t) \end{array}\right] \tag{13}$$

$$\lambda_\pm = \frac{1}{2}\left[\begin{array}{l} -i\Omega_A - i\Omega_B + R^0_{2A} + R^0_{2B} + k_{ex} \\ \pm\sqrt{\left(-i\Delta\omega + R^0_{2A} - R^0_{2B} + k_{ex}(p_B - p_A)\right)^2 + 4 p_A p_B k_{ex}^2} \end{array}\right] \tag{14}$$

For Method 1, ^{15}N NMR spectra were simulated using the program *Mathematica* 8.04 (Wolfram 1999) for a titration involving a 1:1 protein ligand interaction with $k_{on} = 1 \times 10^8$ M^{-1} s^{-1} and $k_{off} = 2 \times 10^4$ s^{-1} ($K_D = k_{off}/k_{on} = 200$ μM), ^{15}N $\Delta\delta_{\max} = 4.64$ ppm ($\Delta\delta_{\max} = {}^{15}$N $\Delta\delta_{\max}/5 = 0.93$ ppm or 278.56 Hz at 60 MHz), $[P_0] = 0.5 \times K_D = 100$ μM, and ligand concentrations of $[L_0]$, $2[L_0]$, $3[L_0]$, $4[L_0]$, $5[L_0]$, $6[L_0]$, $7[L_0]$, $8[L_0]$, $12[L_0]$, and $16[L_0]$, with $[L_0] = 50$ μM, with the final ligand concentration 16 $[L_0] = 8$ $[P_0]$. These conditions are similar to those in Method 1, Case 1 (Table 1), but with only one K_D and a slightly smaller $\Delta\delta_{\max}$. The resonance frequencies of the free and bound states were $\nu_A = 7{,}273.76$ and $\nu_B = 7{,}552.32$ Hz, respectively, and the intrinsic ^{15}N linewidths (R^0_{2A}/π and R^0_{2B}/π) of the free and bound states were taken to be 2 Hz, the expected value for an ~80 residue protein. FIDs were calculated over an

acquisition time of 94.4 ms. FIDs were multiplied by a cosine squared window function and the first point was halved prior to analytical Fourier transformation. ^{15}N NMR spectra for a second and third titration were simulated in a similar fashion with the exception that $[P_0] = 5 \times K_D = 500$ μM with $\Delta_{[P]} = 0$, and $[P_0] = 500$ μM with $\Delta_{[P]} = 0.25$, respectively, similar to Method 1, Cases 2 and 4, respectively (Table 2). For Method 2, ^{15}N NMR spectra were simulated in a similar fashion as Method 1, but with $k_{on} = 1 \times 10^8$ M^{-1} s^{-1}, $k_{off} = 2 \times 10^2$, 2×10^4, and 2×10^5 s^{-1} (K_D s of 2, 200, and 2,000 μM), with $[L_0] = 1$ mM, $[P_0] = 500$ μM, and a decrease in subsequent concentrations by a constant increment of 0.05 mM, for ten titration points, these conditions are similar to those in Method 2 (Table 2), but with only three K_Ds and a slightly smaller $\Delta\delta_{max}$.

Sample preparation and experimental NMR-monitored titrations

Human Mms2 (145 residues) was prepared as previously described (Spyracopoulos et al. 2005; Lewis et al. 2006), whereas human ubiquitin (76 residues) was purchased as a lyophilized powder from Boston Biochem (Cambridge, MA). NMR samples for Method 1 employed [U–^{15}N]-Mms2 at a starting concentration $[P_0]$ ~0.33 mM in 320 μL of 9:1 H_2O/D_2O (pH 7.3), 50 mM TRIS, 200 mM NaCl, 1 mM DTT, 1 mM DSS, 12 μL of 25 × stock protease inhibitor cocktail (prepared from Roche cOmplete protease inhibitor tablets, catalog #11697498001), in a Shigemi microcell NMR tube. A stock solution of 4.7 mM human ubiquitin was prepared using the identical buffer as that used to prepare [U–^{15}N]-Mms2. In addition to an initial titration point in the absence of ubiquitin, five titrations points were collected with Mms2 concentrations, $[P_T]$, of 0.28, 0.23, 0.19, 0.15, and 0.06 mM, with corresponding ubiquitin concentrations, $[L_T]$, 0.18, 0.39, 0.56, 0.71, and 1.06 mM (Table 3). Titrations were conducted by removing the sample from the NMR tube at each titration point using a glass pipette, and subsequently mixing a known amount of the recovered sample with the amount of stock ubiquitin solution and/or stock buffer solution necessary to prepare the next sample. To avoid concentration changes due to small volume losses upon sample recovery, a known amount of recovered sample is removed using a micropipette and buffer is added to this, allowing precise calculation of both concentrations in the new sample for the next titration point. The NMR tube was washed, rinsed, and dried between titration points to minimize unwanted sample dilution.

For Method 2, NMR samples employed [U–^{15}N]-Mms2 at a starting concentration $[P_0]$ ~0.59 mM and a starting concentration of 1.0 mM ubiquitin in 330 μL of 9:1 H_2O/D_2O (pH 7.3), 50 mM TRIS, 200 mM NaCl, 1 mM DTT, 1 mM DSS, 13.2 μL of 25 × stock protease inhibitor cocktail (prepared from Roche cOmplete protease inhibitor tablets, catalog #11697498001), in a Shigemi microcell NMR tube. In addition to the initial point, four titrations points were collected by diluting the initial sample with stock buffer to achieve Mms2 concentrations, $[P_T]$, of 0.49, 0.39, 0.29, and 0.20 mM, with corresponding Ub concentrations, $[L_T]$, 0.83, 0.67, 0.50, and 0.33 mM (Table 4). A separate NMR sample containing 0.44 mM [U–^{15}N]-Mms2 in the identical buffer employed in the titration was also prepared to collect free chemical shifts for Mms2. Titrations were conducted by removing the sample from the NMR tube after each titration point using a glass pipette, and subsequently mixing a known amount of the recovered sample with the amount of stock buffer solution necessary to prepare the next sample. Given the small volume loss upon sample recovery, a known amount of recovered sample is removed using a micropipette and buffer is added to this, allowing precise calculation of both concentrations in the new sample for the next titration point. Additionally, the NMR tube was washed, rinsed, and dried between titration points to prevent unwanted sample dilution. For both Methods 1 and 2, the concentrations of stock Mms2 and ubiquitin were determined by amino acid analysis.

NMR spectroscopy

For chemical shift titrations using Method 1 and 2, wherein [U–^{15}N]-Mms2 is monitored, 2D ^{1}H–^{15}N sensitivity

Table 4 Protein and ligand concentrations for experimental verification of Method 2

Method 2					
$[P_T]$ (mM)	0.59	0.49	0.39	0.29	0.20
$[L_T]$ (mM)	1.0	0.83	0.67	0.5	0.33
Transients collected	16	16	16	64	128

Table 3 Protein and ligand concentrations for experimental verification of Method 1

Method 1						
$[P_T]$ (mM)	0.33	0.28	0.23	0.19	0.15	0.06
$[L_T]$ (mM)	0.0	0.18	0.39	0.56	0.71	1.06
Transients collected	8	16	16	64	64	128

enhanced HSQC NMR spectra (Kay e al. 1992) were acquired using a Varian Unity INOVA 600 MHz spectrometer equipped with a room temperature 5-mm triple resonance probe and triple-axis pulsed field gradients. A total of 192 and 977 complex points were collected in the t_1 and t_2 domains, respectively. The number of titration points for Method 1 and Method 2, and the numbers of transients collected per titration point are given in Tables 3 and 4.

NMR data processing and analysis

All spectral processing was accomplished with the program NMRPipe (Delaglio et al. 1995). For 2D ^{1}H–^{15}N HSQC spectra, sorting and processing of the superposed orthogonal components for sensitivity enhancement were performed with the rance Y.M macro within the NMRPipe software. Post-acquisition processing of the t_2 interferograms for removal of residual water was employed for 2D ^{1}H–^{15}N HSQC NMR spectra. For 2D ^{1}H–^{15}N and NMR spectra acquired for Mms2, 85°-shifted sine and 90°-shifted sine-squared window functions were applied in t_2 and t_1, respectively. The t_2 and t_1 domains were extended to twice the number of points with zero filling. An automatic polynomial subtraction in the F_2 dimension was used for baseline correction, the region upfield of 6.0 ppm was discarded for 2D ^{1}H–^{15}N NMR spectra. Chemical shifts were assigned using those previously described (McKenna et al. 2003); spectra for the titrations were analyzed with the program Sparky (Goddard and Kneller 2008).

Nonlinear regression for estimation of K_D and $\Delta\delta_{max}$ from NMR-monitored titration data

For Methods 1 and 2, the per residue observed chemical shift changes for various Mms2 residues as a function of Mms2 and ubiquitin concentrations were fit to Eq. 6 to yield K_D and $\Delta\delta_{max}$ values using the default "NMinimize" constrained nonlinear least squares fitting algorithm implemented in *Mathematica* 8.0.4. The precision of K_D and $\Delta\delta_{max}$ were determined for each residue in both the ^{1}H and ^{15}N dimensions independently using Monte Carlo methods within *Mathematica*. For Methods 1 and 2, the initial protein concentrations by choosing randomly from normal distributions with means of 1 mM, and standard deviations of ± 10 %, and multiplying these values by the experimental protein concentrations; likewise, the initial ligand concentrations were chosen randomly in the same manner. For each Monte Carlo trial, the chemical shifts were also chosen randomly based on the per residue values of the chemical shift precision determined for a given cross-peak (vide infra). These data sets were then fit using the same nonlinear least squares fitting algorithm used to determine the experimental K_D and $\Delta\delta_{max}$. The standard deviations of the K_D and $\Delta\delta_{max}$ values from 1,000 Monte Carlo trials were taken as the errors in these parameters.

Analyses of chemical shift precision from 2D ^{1}H–^{15}N HSQC NMR spectra

The precision of various per residue ^{1}H and ^{15}N chemical shifts measured from 2D ^{1}H–^{15}N HSQC NMR spectra were determined using Monte Carlo methods within the program *Mathematica* 8.0.4. For a given 2D cross-peak, the maximum point and the first point on either side of the maximum in either the ^{1}H or ^{15}N dimension was assumed to follow a parabolic dependence on frequency, given by the system of polynomial equations (Goddard and Kneller 2008):

$$\begin{aligned} a_0 + a_1x_1 + a_2x_1^2 &= y_1 \\ a_0 + a_1x_2 + a_2x_2^2 &= y_2 \\ a_0 + a_1x_3 + a_2x_3^2 &= y_3 \end{aligned} \tag{15}$$

where y_2 is the height of the maximum, y_1 and y_3 are the heights of the points on either side, x_1, x_2, and x_3 are the chemical shifts of y_1, y_2, and y_3 respectively, and a_0, a_1, and a_3 are constants. This system of equations can be solved to yield the values of the constants, which are then substituted into the first derivative of the parabolic equation ($a_1 + 2a_2x$), which is set to zero and solved with respect to x to yield an estimate for the chemical shift at the interpolated peak maximum. Ten thousand Monte Carlo trials were conducted by adding random noise, estimated from the baseplane of the NMR spectrum, to every point of the region of the 2D spectrum surrounding the cross-peak of interest, then the y_1, y_2, and y_3 values for these data sets were chosen for either the ^{1}H or ^{15}N resonance for a given cross-peak, and used to interpolate the respective chemical shifts, as described above. The region around a given cross-peak was manually chosen such that the surrounding baseplane was well represented, and lacked other cross-peaks.

Results and discussion

Precision of K_D and $\Delta\delta_{max}$ parameters for simulated titrations conducted with fixed [P_0]

In a typical protein–ligand titration, the initial protein concentration (analyte or titrand), [P_0], is maintained nearly constant during a titration, and the resulting observed chemical shift changes are analyzed using nonlinear regression to estimate the parameters K_D and $\Delta\delta_{max}$. In the discussion that follows, we assume throughout that the protein–ligand interaction is described by 1:1 stoichiometry. It has previously been demonstrated that under

conditions of nearly fixed $[P_0]$, the most accurate K_D and $\Delta\delta_{max}$ parameters will be obtained for $[P_0] \sim 0.5 \times K_D$, a maximum concentration for the titrant (ligand) of ~ 10 $[P_0]$, and 15–20 repetitions of the titration (Granot 1983). This number of repetitions is typically not practical when employing isotopically labeled proteins and/or ligands that are difficult or costly to prepare. The optimal ligand concentration for achieving accuracy is also not often achievable in practice, particularly if the ligand is a peptide or protein that is difficult to produce, or not highly soluble. Furthermore, in the absence of prior knowledge of K_D, it is difficult to chose $[P_0] \sim 0.5 \times K_D$; moreover, it is often desirable to have millimolar concentrations of $[P_0]$ to maximize signal intensity in NMR spectra, whereas biologically relevant K_D values are typically in the μM range. Thus, increasing $[P_0]$ to mM concentrations for an initial titration, in the interest of sensitivity gains, will lead to deleterious effects on the accuracy of K_D and $\Delta\delta_{max}$ for biologically relevant K_Ds determined using a traditional titration. In this study, we developed two approaches whereby a titration can be conducted with millimolar starting concentrations for $[P_0]$, while achieving accuracy comparable to, or better than, maintaining $[P_0] \sim 0.5 \times K_D$.

Precision of K_D and $\Delta\delta_{max}$ parameters for simulations of Method 1: $[P_0]$ decreases by a constant factor as $[L_0]$ increases

The accuracy of a given titration where $[P_0]$ remains nearly constant can be assessed by considering a titration for a specific 1:1 protein ligand interaction (Method 1, Case 1, Fig. 1a) with $K_D = 200$ μM, $\Delta\delta_{max} = 1$ ppm, $[P_0] = 0.5 \times K_D = 100$ μM, ligand concentrations of $[L_0]$, $2[L_0]$, $3[L_0]$, $4[L_0]$, $5[L_0]$, $6[L_0]$, $7[L_0]$, $8[L_0]$, $12[L_0]$, and $16[L_0]$, with $[L_0] = 50$ μM, $16[L_0] = 8$ $[P_0]$, errors of ± 5 % for $[L_0]$ and $[P_0]$, and a chemical shift precision of 0.002 ppm. These conditions represent the traditional approach to conducting an NMR titration and can be considered as Method 1 with $\Delta_{[P]} = 0$. Nonlinear least squares regression for an ensemble of 1,000 simulated titrations for Method 1, Case 1, yields normally distributed parameters with median values $K_D = 200 \pm 17$ μM (Fig. 1b) and $\Delta\delta_{max} = 1.00 \pm 0.03$ ppm (Fig. 1c). For $[P_0] = 0.5$ mM, on the other hand (Case 2, Fig. 2a), the parameter ensembles appear as gamma distributions with median values $K_D = 200 \pm 143$ μM (Fig. 2b), $\Delta\delta_{max} = 1.0 \pm 0.2$ ppm (Fig. 2c); values substantially less precise than those obtained at the optimal concentration $[P_0] = 0.5 \times K_D$.

Given the previous titration (Case 1) for $[P_0] = 0.5 \times K_D$, or 100 μM, although good precision is theoretically possible for K_D (< ±10 %), the main problem is that conducting titrations with protein concentrations of

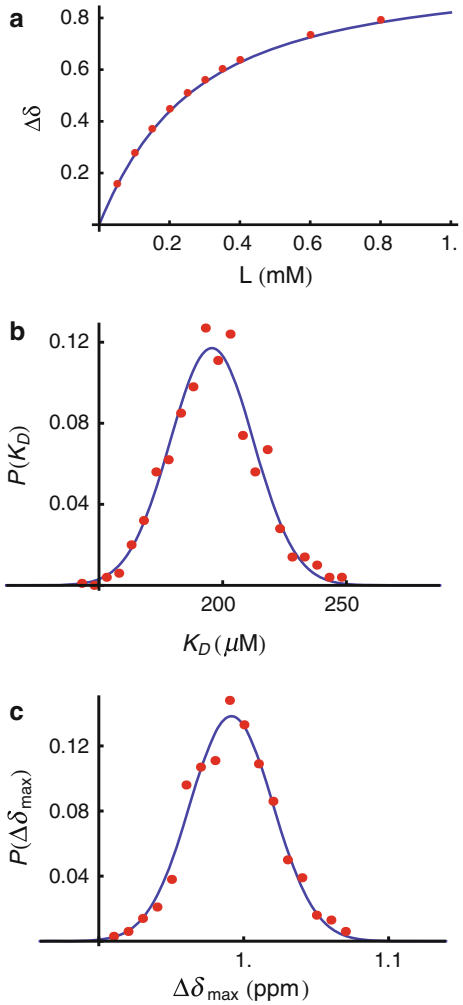

Fig. 1 **a** Binding isotherm for a 1:1 protein ligand interaction for Method 1, Case 1. **b** Histogram for 1,000 K_D values determined from Monte Carlo parameter estimation for Case 1 (*red dots*), and corresponding fit to a normal distribution (*blue line*, Eq. 8). **c** Histogram for 1,000 $\Delta\delta_{max}$ values determined from Monte Carlo parameter estimation for Case 1 (*red dots*), and corresponding fit to a normal distribution (*blue line*, Eq. 8)

100 μM or less becomes time consuming as more transients need to be collected to ensure good signal to noise ratio in standard 2D ^{1}H–^{15}N NMR spectra. The precision for Case 2 is poor (> ±50 %) when $[P_0]$ is 0.5 mM; however, if $[P_0]$ is allowed to decrease by a constant factor in subsequent titration points, precision comparable to Case 1 can be obtained ($[P_0] = 0.5 \times K_D$). For example, for $[P_0] = 0.5$ mM, using $\Delta_{[P]} = 0.25$ in Eq. 7 gives $[P_n] = 0.5$, 0.38, 0.28, 0.21, 0.16, 0.12, 0.09, 0.07, 0.05 and 0.04 mM for $n = 1,\ldots, 10$ (Case 3, Fig. 3a). Using these analyte concentrations yields normally distributed parameters with median values for $K_D = 200 \pm 23$ μM (Fig. 3b), and $\Delta\delta_{max} = 1.00 \pm 0.04$ ppm (Fig. 3c). Thus, the precision for both K_D and $\Delta\delta_{max}$ is similar to Case 1 where $[P_0] = 0.5$ K_D, or 100 μM, for which $\sigma(K_D)$ and

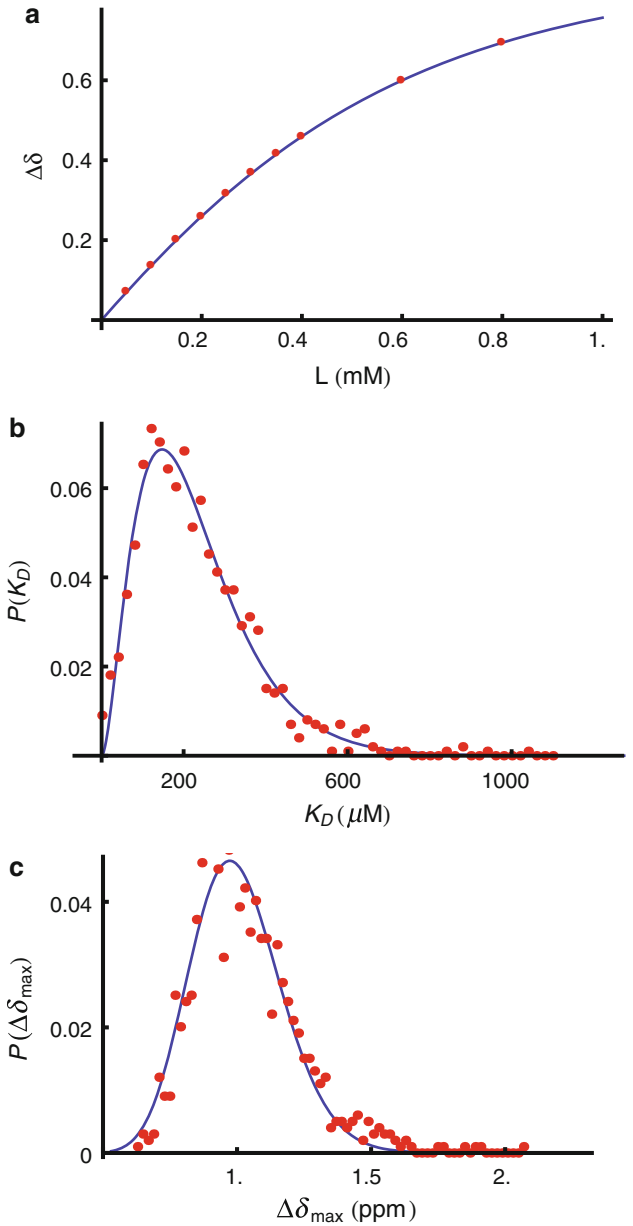

Fig. 2 **a** Binding isotherm for a 1:1 protein ligand interaction for Method 1, Case 2. **b** Histogram for 1,000 K_D values determined from Monte Carlo parameter estimation for Method 1, Case 2 (*red dots*), and corresponding fit to a gamma distribution (*blue line*, Eq. 8). **c** Histogram for 1,000 $\Delta\delta_{max}$ values determined from Monte Carlo parameter estimation for Method 1, Case 2 (*red dots*), and corresponding fit to a gamma distribution (*blue line*, Eq. 8)

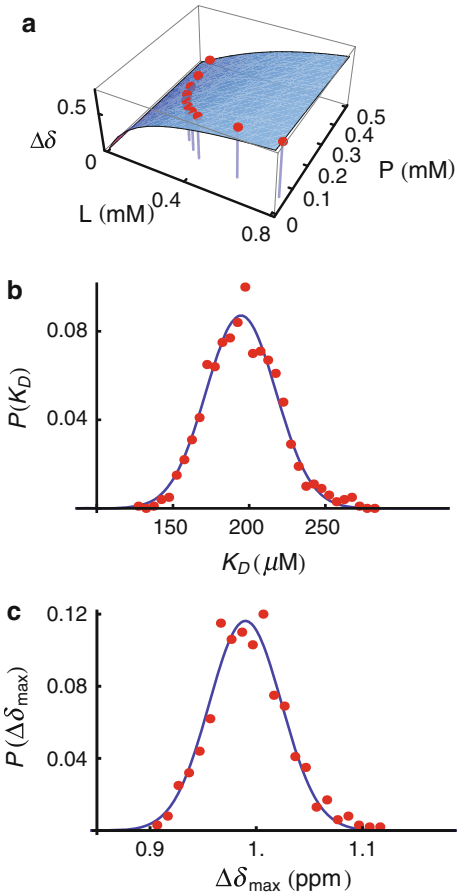

Fig. 3 **a** Binding isotherm for a 1:1 protein ligand interaction for Method 1, Case 3. **b** Histogram for 1,000 K_D values determined from Monte Carlo parameter estimation for Method 1, Case 3 (*red dots*), and corresponding fit to a normal distribution (*blue line*, Eq. 9). **c** Histogram for 1,000 $\Delta\delta_{max}$ values determined from Monte Carlo parameter estimation for Method 1, Case 3 (*red dots*), and corresponding fit to a normal distribution (*blue line*, Eq. 9)

$\sigma(\Delta\delta_{max})$ are ± 17 μM and ± 0.02 ppm, respectively. However, the main advantage for Case 3 is that six of the ten titration points contain protein concentrations exceeding 100 μM, and this is advantageous with respect to signal to noise ratio in NMR spectra in comparison to maintaining $[P_0] = 0.5\ K_D$. If we assume that NMR spectra are collected using identical parameters for the two different scenarios in Case 1 and Case 3, $[P_0] = 100$ μM and $[P_0] = 0.5$ mM with $\Delta_{[P]} = 0.25$, respectively, and linewidth differences between the two cases are small, then an overall increase in the average signal to noise of 1.9-fold can be expected for Case 3 with $[P_0] = 0.5$ mM and $\Delta_{[P]} = 0.25$ compared to Case 1 where $[P_0] = 100$ μM, simply on the basis of protein concentration (Fig. 4). Furthermore, in practice, it is not straightforward to conduct an NMR titration at a fixed protein concentration. For Case 1, where $[P_0] = 100$ μM, if subsequent titration points decreased in concentration by as little as 5 %, then the last five titration points would range between a maximum of 80 μM to a minimum of 60 μM. If NMR spectra are collected using identical parameters to Case 3, with minimal linewidth changes between the two cases, then an overall increase in the average signal to noise of 2.4-fold can be expected for Case 3 ($[P_0] = 500$ μM and $\Delta_{[P]} = 0.25$) compared to Case 1 ($[P_0] = 100$ μM).

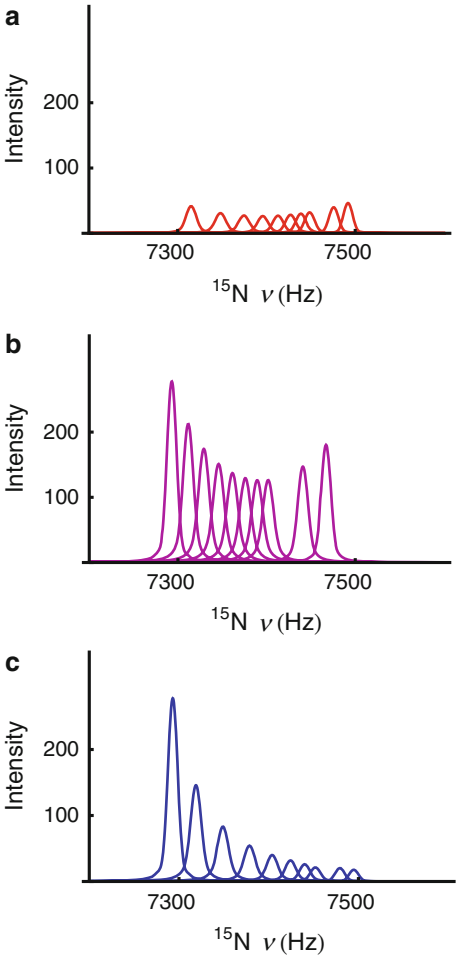

Fig. 4 **a** Simulated ^{15}N lineshapes for Method 1, Case 1, $[P_0] = 0.5 \times K_D = 100$ μM. **b** Simulated ^{15}N lineshapes for Method 1, Case 2, $[P_0] = 5 \times K_D = 500$ μM. **c** Simulated lineshapes for Method 1, Case 3, $[P_0] = 5 \times K_D = 500$ μM with $\Delta_{[P]} = 0.25$. For all spectra, the resonance frequencies for the free and bound protein are 7,273.76 and 7,552.32 Hz, respectively

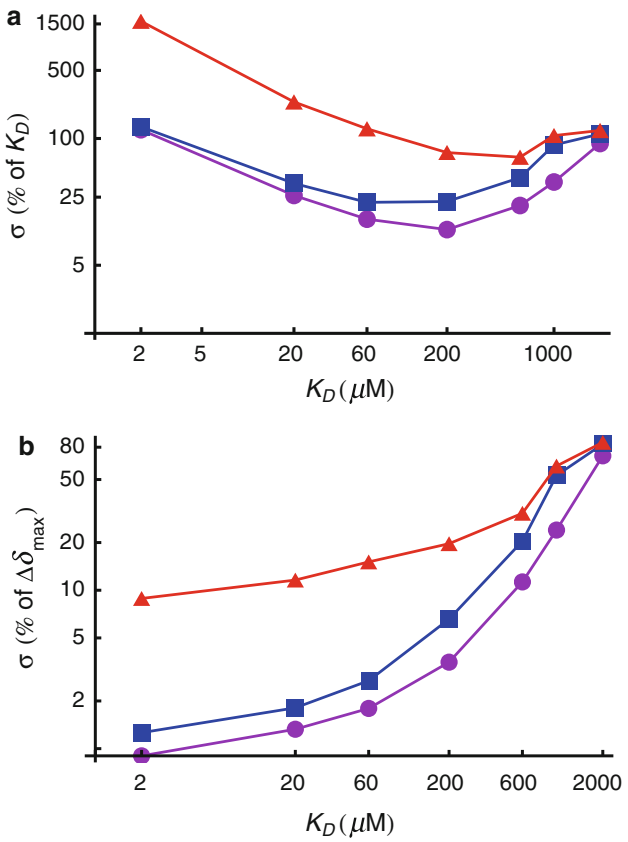

Fig. 5 **a** Standard deviations of 1,000 Monte Carlo trials for various K_D values for $[P_0] = 0.5 \times K_D = 100$ μM and $\Delta_{[P]} = 0$ (*red triangles*); $[P_0] = 500$ μM and $\Delta_{[P]} = 0.1$ (*blue*); $[P_0] = 500$ μM and $\Delta_{[P]} = 0.25$ (*magenta*). **b** Standard deviations of 1,000 Monte Carlo trials for $\Delta\delta_{max}$ corresponding to various values of K_D for $[P_0] = 0.5 \times K_D = 100$ μM and $\Delta_{[P]} = 0$ (*red*); $[P_0] = 500$ μM and $\Delta_{[P]} = 0.1$ (*blue squares*); $[P_0] = 500$ μM and $\Delta_{[P]} = 0.25$ (*magenta circles*)

General performance of Method 1 simulations over a range of K_D and with $\Delta_{[P]}$ values of 0.1 and 0.25

In general, the value of K_D for a given protein–ligand interaction is not known prior to planning an NMR titration. Thus, to determine the performance of Method 1 we conducted Monte Carlo simulations using $[P_0] = 500$ μM, for a series of K_D values (2, 20, 60, 200, 600, 1,000, and 2,000 μM), and $\Delta_{[P]}$ values (0.1 and 0.25, Eq. 7), and ligand concentrations as outlined in Table 1. Using these protein and ligand concentrations, a substantial increase in precision is obtained for K_Ds in the range 2–1,000 μM, as shown in Fig. 5a. However, reasonable accuracy is only obtained within the K_D range 20–600 μM, that is, $\sigma \leq \sim 50\%$ of the actual K_D. In comparison to the K_D parameter, reasonable accuracy for $\Delta\delta_{max}$ ($\sigma \leq \sim 50\%$ of the actual $\Delta\delta_{max}$) is obtained over a K_D range of 2–1,000 μM (Fig. 5b). Importantly, substantial increases in accuracy for K_D and $\Delta\delta_{max}$ are achieved over the range 20–600 μM. These results indicate that in the absence of prior knowledge for the actual value of K_D, Method 1 will produce K_D and $\Delta\delta_{max}$ values with reasonable accuracy for conditions where $[P_0] > (1–25 \times K_D)$.

Precision of K_D and $\Delta\delta_{max}$ parameters for simulations of Method 2: $[P_0]$ and $[L_0]$ decrease by a constant increment

For Method 1, co-variation of the ligand and protein concentrations involves successively diluting the initial protein concentration while simultaneously increasing the ligand concentration. Alternatively, for Method 2, a mixture of concentrated ligand and concentrated protein is successively diluted to conduct a titration (Fig. 6a). We performed Monte Carlo simulations using $[L_0] = 1$ mM and $[P_0] = 500$ μM, for a series of K_D values: 2, 20, 60, 200,

600, 1,000 and 2,000 μM, with a decrease in subsequent protein and ligand concentrations by a constant increment of 50 μM, for a total of ten titration points, as outlined in Table 2. The Monte Carlo ensembles for the simulation with $K_D = 200$ μM are shown in Fig. 6b and c. It is evident that the precision for $K_D = 200$ μM remains similar to that for Method 1. However, a substantial increase in precision is obtained for other K_D values in comparison to the traditional method of conducting a titration (Method 1 with $\Delta_{[P]} = 0$), as well as Method 1 with $\Delta_{[P]} = 0.25$ (Fig. 7). For example, for $K_D = 2$ μM, $\sigma(K_D)$ and $\sigma(\Delta\delta_{max})$ are ± 0.4 μM and ± 0.002 ppm, respectively, in comparison to Method 1, with $K_D = 2$ μM, a protein concentration of 100 μM, and $\Delta_{[P]} = 0$, for which $\sigma(K_D) = \pm$ 1.4 μM and $\sigma(\Delta\delta_{max}) = \pm$ 0.007 ppm. For $K_D = 2{,}000$ μM, $\sigma(K_D)$ and $\sigma(\Delta\delta_{max})$ are ± 653 μM and ± 0.2 ppm, respectively, in comparison to Method 1, with $K_D = 2{,}000$ μM, a protein concentration of 100 μM,

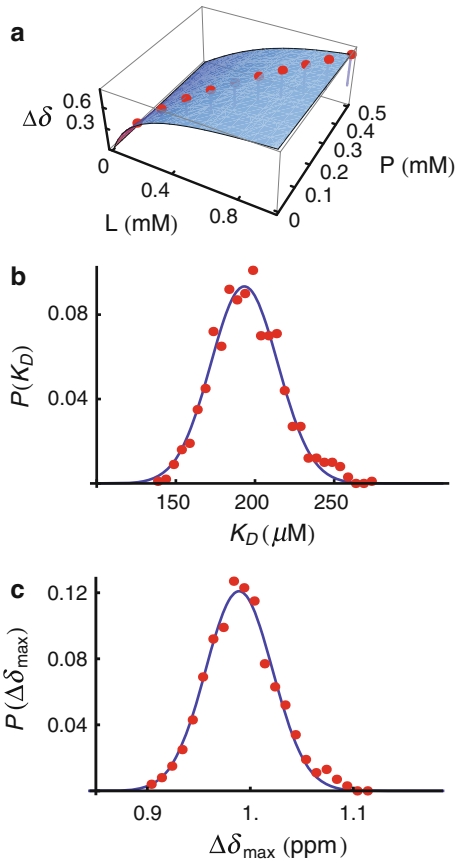

Fig. 6 **a** Binding isotherm for a 1:1 protein ligand interaction for Method 2, $K_D = 200$ μM. **b** Histogram for 1,000 K_D values determined from Monte Carlo parameter estimation for Method 2, $K_D = 200$ μM (*red dots*), and corresponding fit to a normal distribution (*blue line*, Eq. 9). **c** Histogram for 1,000 $\Delta\delta_{max}$ values determined from Monte Carlo parameter estimation for Method 2, $K_D = 200$ μM (*red dots*), and corresponding fit to a normal distribution (*blue line*, Eq. 9)

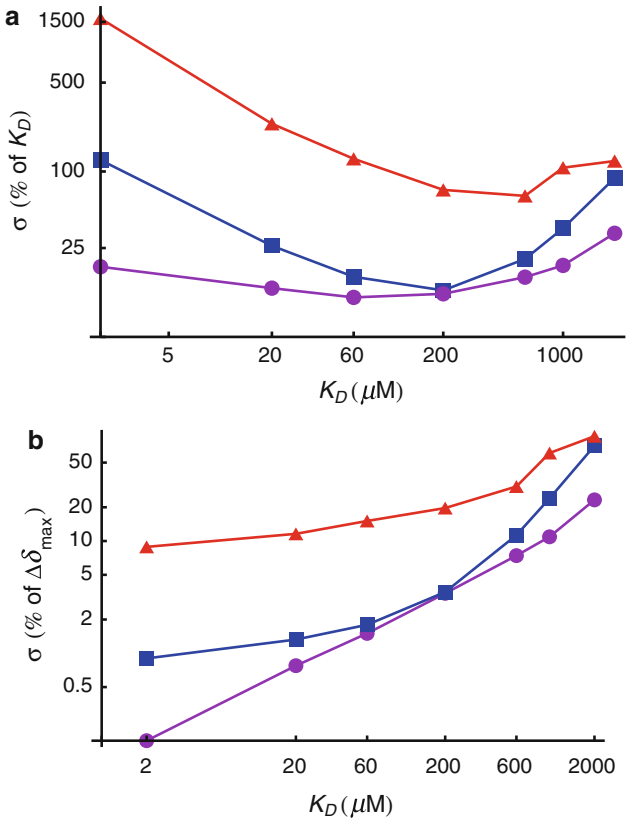

Fig. 7 **a** Standard deviations of 1,000 Monte Carlo trials for various K_D values for $[P_0] = 0.5 \times K_D = 100$ μM and $\Delta_{[P]} = 0$ (*red triangles*); $[P_0] = 500$ μM and $\Delta_{[P]} = 0.25$ (*blue squares*); $[P_0] = 500$ μM with $[L_0] = 1$ mM and with a decrease in subsequent protein and ligand concentrations by a constant increment of 0.05 mM for a total of ten titration points (*magenta circles*). **b** Standard deviations of 1,000 Monte Carlo trials for $\Delta\delta_{max}$ corresponding to various values of K_D for $[P_0] = 0.5 \times K_D = 100$ μM and $\Delta_{[P]} = 0$ (*red triangles*); $[P_0] = 500$ μM and $\Delta_{[P]} = 0.25$ (*blue squares*); $[P_0] = 500$ μM with $[L_0] = 1$ mM and with a decrease in subsequent protein and ligand concentrations by a constant increment of 0.1 mM for a total of ten titration points (*magenta circles*)

and $\Delta_{[P]} = 0$, for which $\sigma(K_D) = \pm$ 1,370 μM and $\sigma(\Delta\delta_{max}) = \pm$ 0.5 ppm. These results for Method 2 indicate that without knowledge of the actual value of K_D, Method 2 can produce K_D and $\Delta\delta_{max}$ values with excellent accuracy ($\sigma \leq \sim 25\%$) for conditions where $[P_0] > (1\text{–}25 \times K_D)$.

Simulated NMR spectra for titrations conducted for Method 2 with $K_D = 2$, 200, and 2,000 μM are shown in Fig. 8. These spectra highlight the fact that while Method 2 provides greater precision over a broad range of K_D values in comparison to Method 1, the observed chemical shift changes over the course of a titration are smaller, particularly at extremes of K_D (2 and 2,000 μM), and therefore, Method 2 has more stringent requirements for precise chemical shift measurements, and would be most accurate

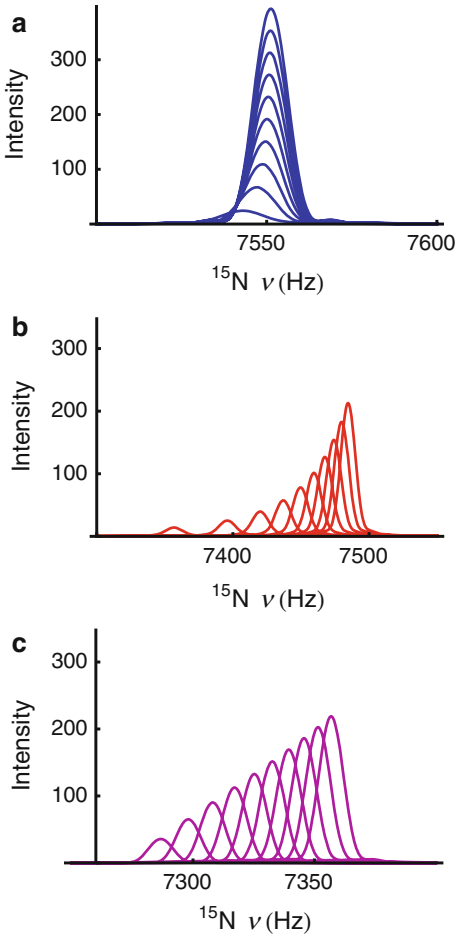

Fig. 8 **a** Simulated ^{15}N lineshapes for Method 2, $K_D = 2$ μM. **b** Simulated ^{15}N lineshapes for Method 2, $K_D = 200$ μM. **c** Simulated lineshapes for Method 2, $K_D = 2{,}000$ μM. For all spectra, the resonance frequencies for the free and bound protein are 7,273.76 and 7,552.32 Hz, respectively

for residues displaying the largest chemical shift changes in a titration ($\Delta\delta_{max} \sim 1$ ppm).

In addition to providing increased precision for K_D and $\Delta\delta_{max}$, the average signal to noise ratio for a titration conducted according to Method 2 is greater than that for a traditional NMR titration for which $[P_0]$ is held at a fixed concentration of $0.5 \times K_D$. Assuming that NMR spectra are collected using identical parameters for Method 2 with $[L_0] = 1$ mM and $[P_0] = 500$ μM, and a constant decrease in protein and ligand concentrations by 100 μM for a total of ten titration points, and Method 1 with a protein concentration of 100 μM, and $\Delta_{[P]} = 0$, the average signal to noise ratio for Method 2 is 2.75-fold greater. It should be noted that Method 2 is theoretically more accurate than Method 1 because the titration is designed to sample the portion of the two-dimensional binding isotherm that has significant curvature; this occurs at low concentrations of both protein and ligand (Fig. 6a). Thus, care must be taken to ensure sufficient transients are collected for the final points of the titration to ensure adequate signal to noise ratio and therefore, adequate chemical shift precision. In addition, Method 2 requires that δ_{free} for the resonances of the unbound protein are determined from a spectrum separate from the titration. Thus, $\Delta\delta_{max}$ can be expressed as a difference (Eq. 5), avoiding the use of δ_{free} as a third free parameter in fitting titration data to Eq. 6.

Experimental validation of Methods 1 and 2 by 2D ^{1}H–^{15}N HSQC NMR-monitored chemical shift titration

A caveat regarding the practical application of Methods 1 and 2 is the necessity of diluting the NMR-observed protein component. This may lead to long acquisition times for the last few NMR spectra to achieve sufficient signal to noise ratio, especially for NMR-monitored titrations for large protein–protein or protein–ligand complexes and/or processes in the intermediate exchange timescale. We conducted NMR-monitored titration experiments using Methods 1 and 2 for the interaction of human [*U*–^{15}N]-Mms2 (145 residues, 16 kDa) with human ubiquitin (76 residues, 8.5 kDa) (Tables 3 and 4, and Fig. 9). The interaction of Mms2 with ubiquitin has been previously characterized using NMR (McKenna et al. 2003; Lewis et al. 2006), and the binding occurs with 1:1 stoichiometery (Spyracopoulos et al. 2005). A total of five 2D ^{1}H–^{15}N HSQC NMR spectra were collected for both Methods 1 and 2; the associated fits to a 1:1 binding model to extract K_D and $\Delta\delta_{max}$ values for various Mms2 residues are shown in Fig. 10 and Tables 5 and 6. Overall, excellent chemical shift precision and signal to noise can be obtained in a reasonable amount of time using a room temperature triple resonance probe at a magnetic field strength of 600 MHz (14.1 Tesla). For example, for 2D ^{1}H–^{15}N NMR spectra collected at the lowest concentration of the observed protein component, [*U*–^{15}N]-Mms2, the chemical shift precision for the ^{1}H and ^{15}N dimensions is within the ppb range over a wide range of linewidths (Fig. 11). Another indication of the quality of the titration methods are the differences between experimental $\Delta\delta_{obs}$ values and those determined from the best fits of the data to a 1:1 binding isotherm, given by the expression $\sqrt{\left(\Delta\delta_{obs,calc} - \Delta\delta_{obs,exp}\right)^2}$. This average difference for T49 ^{15}N and ^{1}H^{N} chemical shifts for Method 1 is 0.3 and 0.2 ppb, respectively (Fig. 10a). For Method 2, the average difference for T49 ^{15}N and ^{1}H^{N} chemical shifts is 0.2 and 1.7 ppb, respectively (Fig. 10b). There is excellent agreement between the average K_Ds for Methods 1 and 2, 0.31 ± 0.02 and 0.35 ± 0.04 mM, respectively (Tables 5, 6, excluding V26 ^{15}N and T49 ^{1}H^{N}). For the respective experimental conditions, Method 2 is more precise than Method 1, with an average per residue K_D error of 0.04 compared to 0.07 mM (Tables 5, 6). This difference in

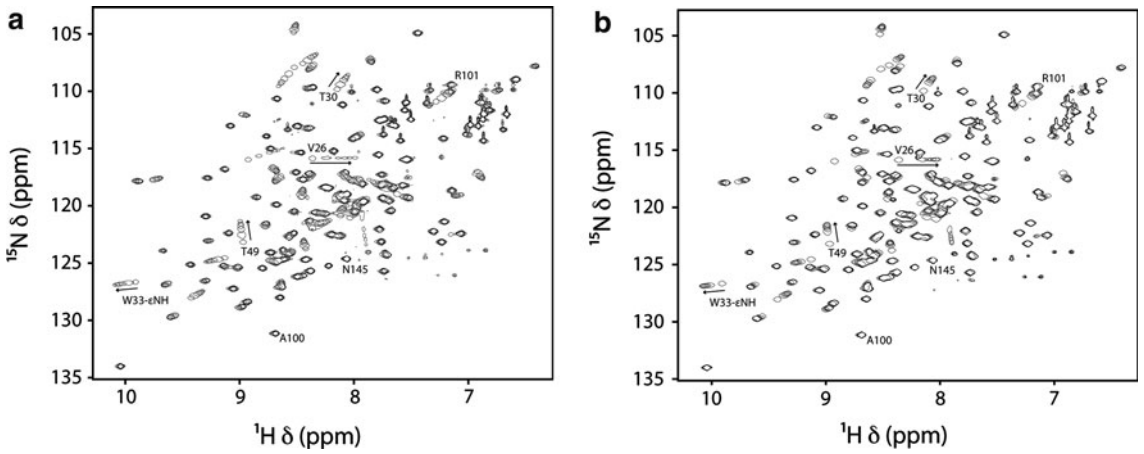

Fig. 9 2D ^{1}H–^{15}N HSQC 600 MHz NMR-monitored titrations for the interaction of [*U*–^{15}N]-Mms2 with ubiquitin for Method 1 (**a**), and Method 2 (**b**). A number of residues used in the various analyses are labeled. *Arrows* indicate the shift from the free state to the bound state

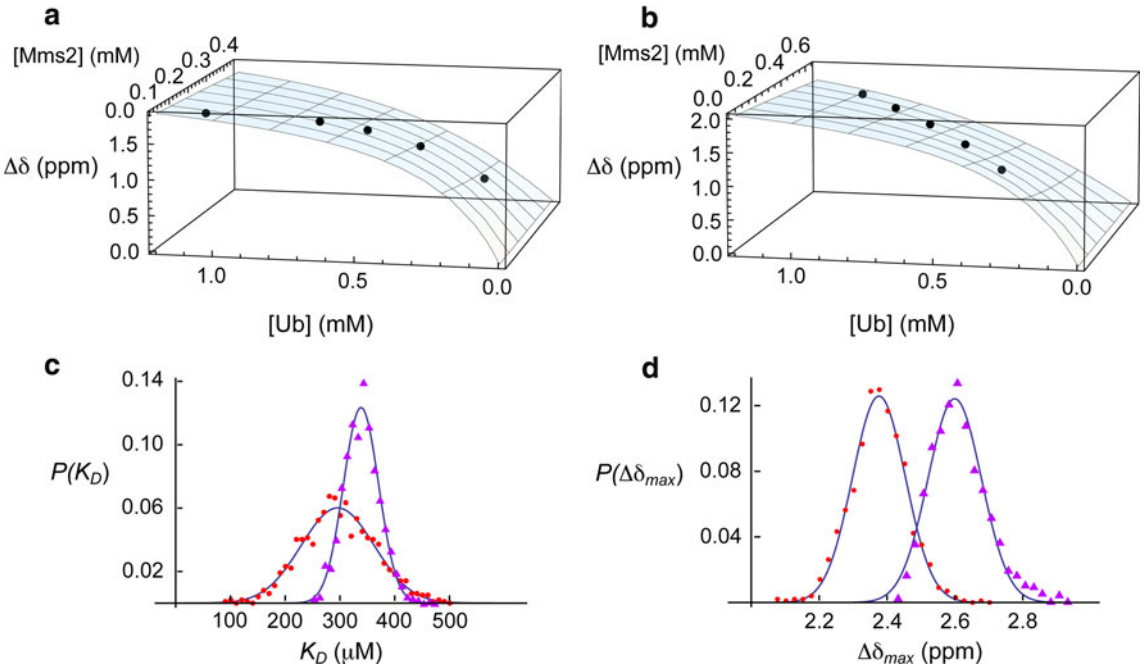

Fig. 10 **a** Experimental binding isotherms for T49 ^{15}N from NMR-monitored titrations using Method 1 for the interaction of [*U*–^{15}N]-Mms2 with ubiquitin. **b** Experimental binding isotherms for T49 ^{15}N from NMR-monitored titrations using Method 2 for the interaction of [*U*–^{15}N]-Mms2 with ubiquitin. **c** Error estimates for the K_D parameter from an ensemble of 1,000 Monte Carlo simulations for T49 ^{15}N, Method 1 (*red circles*) and Method 2 (*purple triangles*). **d** Error estimates for the $\Delta\delta_{max}$ parameter error for an ensemble of 1,000 Monte Carlo simulations for T49 ^{15}N, Method 1 (*red circles*), and Method 2 (*purple triangles*). For **c** and **d**, the lines through the points represent the best fit to a normal distribution (Eq. 8)

precision between Methods 1 and 2 is evident in the data in Fig. 10c, which shows that $\pm$ 1σ for the K_D value of T49 from Method 1 is about twice that of Method 2; consistent with the theoretical prediction that Method 2 should be slightly more precise for K_D $\sim$300 μM (Fig. 7a). Finally, both Methods 1 and 2 perform reasonably well with respect to the magnitude of $\Delta\delta_{max}$, the size of the chemical shift change upon binding for a given residue (Fig. 12). Generally, for larger values of $\Delta\delta_{max}$, between $\sim$0.2 and 0.5 ppm, with ^{15}N $\Delta\delta$/5 to facilitate comparison to ^{1}H^{N}, the average error is 18 % of the value of K_D. In contrast, values of $\Delta\delta_{max}$ that are smaller than $\sim$0.2 ppm have larger errors, $\sim$60 % of the value of K_D, on average. Thus, as a general rule, only residues with $\Delta\delta_{max}$ values exceeding $\sim$0.2 ppm should be used in determining the K_D for a given protein interaction.

Table 5 Fitted parameters from Method 1 for titration of [U–^{15}N]-Mms2 with ubiquitin

Residue	K_D (mM)	$\Delta\delta_{max}$ [a] (ppm)	χ^2 ($\times 10^{-7}$)
V26 ^{15}N	0.01 ± 1[b]	0.01 ± 0.01	4,675.38
V26 $^{1}H^{N}$	0.33 ± 0.08	0.49 ± 0.03	93.18
T49 ^{15}N	0.30 ± 0.07	0.48 ± 0.02	4,171.11
T49 $^{1}H^{N}$	1 ± 3[b]	0.06 ± 0.09	13.30
W33 $^{15}N_{\varepsilon 1}$	0.34 ± 0.1	0.068 ± 0.007	4.25
W33 $^{1}H_{\varepsilon 1}$	0.30 ± 0.07	0.24 ± 0.009	99.15
T30 ^{15}N	0.30 ± 0.06	0.33 ± 0.01	160.05
T30 $^{1}H^{N}$	0.27 ± 0.07	0.115 ± 0.006	4.88

[a] For ^{15}N chemical shifts, $\Delta\delta_{max}$ is scaled by a factor of 1/5

[b] Not fit

Table 6 Fitted parameters from Method 2 for titration of [U–^{15}N]-Mms2 with ubiquitin

Residue	K_D (mM)	$\Delta\delta_{max}$ [a] (ppm)	χ^2 ($\times 10^{-7}$)
V26 ^{15}N	10 ± 4[b]	0.10 ± 0.05	1,529.85
V26 $^{1}H^{N}$	0.42 ± 0.06	0.55 ± 0.03	801.40
T49 ^{15}N	0.34 ± 0.03	0.52 ± 0.02	2,489.66
T49 $^{1}H^{N}$	0.13 ± 0.08	0.028 ± 0.003	7.31
W33 $^{15}N_{\varepsilon 1}$	0.37 ± 0.05	0.071 ± 0.004	270.39
W33 $^{1}H_{\varepsilon 1}$	0.32 ± 0.03	0.253 ± 0.008	30.00
T30 ^{15}N	0.32 ± 0.03	0.348 ± 0.007	474.49
T30 $^{1}H^{N}$	0.36 ± 0.04	0.130 ± 0.005	8.71

[a] For ^{15}N chemical shifts, $\Delta\delta_{max}$ is scaled by a factor of 1/5

[b] Not fit

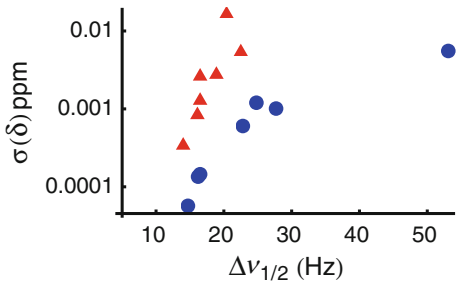

Fig. 11 Standard deviations of 10,000 Monte Carlo trials for chemical shifts determined from parabolic peak interpolation as a function of linewidth. Peak interpolation was conducted on a variety of cross-peaks from 2D ^{1}H–^{15}N HSQC 600 MHz NMR spectra corresponding to the lowest concentration of [U–^{15}N]-Mms2. Data for $^{1}H^{N}$ are shown as *blue circles*, whereas data for ^{15}N are shown as *red triangles*

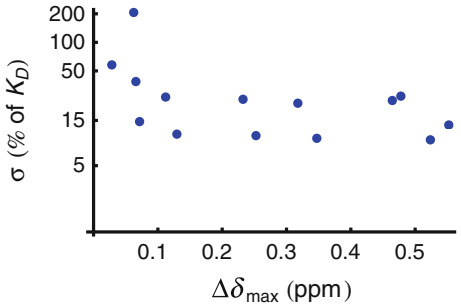

Fig. 12 Error in K_D as a function of the maximum chemical shift change upon binding ($\Delta\delta_{max}$). Monte Carlo parameter estimation was conducted on a variety of cross-peaks (labeled in Fig. 9) from NMR-monitored titration data for Methods 1 and 2. Data include both ^{15}N and $^{1}H^{N}$; ^{15}N chemical shifts are scaled by a factor of 1/5

Conclusions

For traditional NMR-monitored titrations, [P_0] remains fixed at a constant concentration, and without prior knowledge of the dissociation constant, it is difficult to experimentally measure an accurate K_D, given the requirement that the optimal value of [P_0] should be $\sim 0.5 \times K_D$. We have demonstrated that, in general, co-variation of the ligand and protein concentrations during the course of an NMR-monitored titration leads to increased precision in the parameters K_D and $\Delta\delta_{max}$ for 1:1 protein–protein and protein–ligand interactions. The general utility of co-variation of protein and ligand concentrations is two-fold: accurate K_Ds can be measured using higher starting concentrations for [P_0]; secondly, larger [P_0] values translate into greater average signal to noise ratios for the NMR spectra in a given titration. Of the two methods that were analyzed, Method 2, which involves serial dilution of a solution of concentrated protein and ligand, provides fairly robust precision ($\sigma \leq \sim 25$ %) over a broad range of K_D. Some caveats regarding both methods include the fact that observed chemical shift changes for Method 2 are smaller than Method 1, or traditional NMR titrations, and require that chemical shifts be measured with high precision. Secondly, both methods rely on dilution of the NMR-observed protein component, and may require longer acquisition times for the last few titration points to achieve sufficient signal to noise ratio. In this regard, we have demonstrated experimentally that the proposed methods are practical, and straightforward to apply to protein–protein interactions for complexes up to 24 kDa with binding kinetics in the fast exchange regime on the NMR timescale. NMR-monitored titrations involving larger protein–protein or protein–ligand complexes and/or processes in the intermediate exchange timescale become challenging to analyze due to extensive line broadening. However, approaches such as TROSY or CRINEPT NMR spectroscopy (Riek et al. 2000), coupled with deuteration of the protein components (Gardner and Kay 1998), can alleviate problems associated with linebroadening for NMR-monitored titrations (Markin et al. 2010b).

Acknowledgments This research was supported by the Canadian Institutes of Health Research (grant MOP 110964). L.S. is an Alberta Innovates—Health Solutions Senior Scholar, C.J.M. is the recipient of an Alberta Innovates—Health Solutions Studentship. We thank the Canadian National High Field NMR Centre (NANUC) for use of the facilities. Operation of NANUC is funded by the Natural Science and Engineering Research Council of Canada and the University of Alberta.

References

Baryshnikova OK, Robertson IM, Mercier P, Sykes BD (2008) The dilated cardiomyopathy G159D mutation in cardiac troponin C weakens the anchoring interaction with troponin I. Biochemistry 47:10950–10960

Cavanagh J (2007) Protein NMR spectroscopy: principles and practice. Academic Press, Boston

Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A (1995) NMRPipe—a multidimensional spectral processing system based on UNIX pipes. J Biomol NMR 6:277–293

Fielding L (2007) NMR methods for the determination of protein-ligand dissociation constants. Prog Nucl Mag Res Sp 51: 219–242

Gardner KH, Kay LE (1998) The use of ^{2}H, ^{13}C, ^{15}N multidimensional NMR to study the structure and dynamics of proteins. Ann Rev Bioph Biom 27:357–406

Goddard TD, Kneller DG (2008) SPARKY 3. Universityof California, San Fransico

Granot J (1983) Determination of dissociation constants of 1-1 complexes from NMR data—optimization of the experimental setup by statistical analysis of simulated experiments. J Magn Reson 55:216–224

Hariri nia A, D'Onofrio M, Fushman D (2007) Mapping the interactions between Lys48 and Lys63-linked di-ubiquitins and a ubiquitin-interacting motif of S5a. J Mol Biol 368:753–766

Kay LE, Keifer P, Saarinen T (1992) Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity. J Am Chem Soc 114:10663–10665

Lewis MJ, Saltibus LF, Hau DD, Xiao W, Spyracopoulos L (2006) Structural basis for non-covalent interaction between ubiquitin and the ubiquitin conjugating enzyme variant human MMS2. J Biomol NMR 34:89–100

Marintchev A, Frueh D, Wagner G (2007) NMR methods for studying protein–protein interactions involved in translation initiation. Method Enzymol 430:283–331

Markin CJ, Saltibus LF, Kean MJ, McKay RT, Xiao W, Spyracopoulos L (2010a) Catalytic proficiency of ubiquitin conjugation enzymes: balancing pK_a suppression, entropy, and electrostatics. J Am Chem Soc 132:17775–17786

Markin CJ, Wei XA, Spyracopoulos L (2010b) Mechanism for recognition of polyubiquitin chains: balancing affinity through interplay between multivalent binding and dynamics. J Am Chem Soc 132:11247–11258

McKenna S, Hu J, Moraes T, Xiao W, Ellison MJ, Spyracopoulos L (2003a) Energetics and specificity of interactions within Ub·Uev·Ubc13 human ubiquitin conjugation complexes. Biochemistry 42:7922–7930

McKenna S, Moraes T, Pastushok L, Ptak C, Xiao W, Spyracopoulos L, Ellison MJ (2003b) An NMR-based model of the ubiquitin-bound human ubiquitin conjugation complex Mms2·Ubc13—the structural basis for lysine 63 chain catalysis. J Biol Chem 278: 13151–13158

Palmer AG, Kroenke CD, Loria JP (2001) Nuclear magnetic resonance methods for quantifying microsecond-to-millisecond motions in biological macromolecules. Method Enzymol Pt B 339:204–238

Riek R, Pervushin K, Wuthrich K (2000) TROSY and CRINEPT: NMR with large molecular and supramolecular structures in solution. Trends Biochem Sci 25:462–468

Rintala-Dempsey AC, Santamaria-Kisiel L, Liao Y, Lajoie G, Shaw GS (2006) Insights into S100 target specificity examined by a new interaction between S100A11 and annexin A2. Biochemistry 45:14695–14705

Spyracopoulos L, Lewis MJ, Saltibus LF (2005) Main chain and side chain dynamics of the ubiquitin conjugating enzyme variant human Mms2 in the free and ubiquitin-bound states. Biochemistry 44:8770–8781

Wolfram S (1999) The mathematica book. Wolfram Media, Cambridge University Press, Champaign, IL, New York

Zuiderweg ERP (2002) Mapping protein–protein interactions in solution by NMR spectroscopy. Biochemistry 41:1–7