JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Article

pubs.acs.org/JACS

Quantitative Analysis of Multisite Protein–Ligand Interactions by NMR: Binding of Intrinsically Disordered p53 Transactivation Subdomains with the TAZ2 Domain of CBP

Munehito Arai,[†,‡,§,⊥] Josephine C. Ferreon,[†] and Peter E. Wright[†,*]

[†]Department of Molecular Biology and Skaggs Institute for Chemical Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037, United States

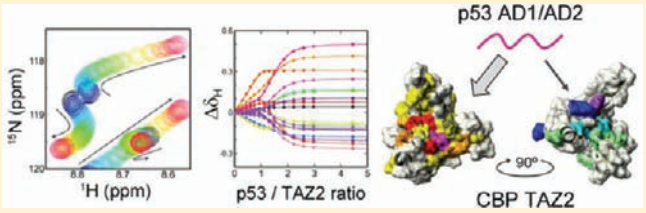
[‡]Institute for Biological Resources and Functions, AIST, 1-1-1 Higashi, Tsukuba, Ibaraki 305-8566, Japan

[§]Department of Life Sciences, Graduate School of Arts and Sciences, The University of Tokyo, 3-8-1 Komaba, Meguro, Tokyo 153-8902, Japan

[⊥]PRESTO, JST, 4-1-8 Honcho, Kawaguchi, Saitama 332-0012, Japan

Supporting Information

ABSTRACT: Determination of affinities and binding sites involved in protein–ligand interactions is essential for understanding molecular mechanisms in biological systems. Here we combine singular value decomposition and global analysis of NMR chemical shift perturbations caused by protein–protein interactions to determine the number and location of binding sites on the protein surface and to measure the binding affinities. Using this method we show that the isolated AD1 and AD2 binding motifs, derived from the intrinsically disordered N-terminal transactivation domain of the tumor suppressor p53, both interact with the TAZ2 domain of the transcriptional coactivator CBP at two binding sites. Simulations of titration curves and line shapes show that a primary dissociation constant as small as 1–10 nM can be accurately estimated by NMR titration methods, provided that the primary and secondary binding processes are coupled. Unexpectedly, the site of binding of AD2 on the hydrophobic surface of TAZ2 overlaps with the binding site for AD1, but AD2 binds TAZ2 more tightly. The results highlight the complexity of interactions between intrinsically disordered proteins and their targets. Furthermore, the association rate of AD2 to TAZ2 is estimated to be 1.7×10^{10} M^{-1} s^{-1}, approaching the diffusion-controlled limit and indicating that intrinsic disorder plus complementary electrostatics can significantly accelerate protein binding interactions.



INTRODUCTION

Proteins interact with a variety of ligands, including other proteins, DNA, and small molecules, at a single site or multiple sites on their surfaces. Determination of the affinities and binding sites involved in protein–ligand interactions is essential for understanding molecular mechanisms in biological systems. Heteronuclear NMR spectroscopy has been widely used to identify binding sites and to determine binding affinities by monitoring chemical shift perturbations that accompany complex formation.[1,2] However, the method has been applied mostly to one-site binding processes, and more complicated binding phenomena have rarely been studied quantitatively by NMR titrations. Since NMR is an extremely powerful technique for obtaining structural insights into protein–ligand interactions, development of a method to analyze quantitatively the chemical shift perturbations associated with multiple binding events will have a significant impact on molecular biology and biochemistry.

Two alternative approaches that are potentially applicable to the quantitative analysis of chemical shift perturbations are singular value decomposition (SVD) and global analysis. SVD is a multivariate analysis used to estimate the number of meaningful components, such as the number of binding sites, necessary for explaining the observed data set, and to remove the contribution of noise to experimental data.[3,4] However, it cannot provide information on affinities and binding sites. On the other hand, global analysis can provide reliable solutions for the affinities and locate binding sites by fitting multiple data sets simultaneously to a given physicochemical model.[4,5] Although global analysis fails if a fitting model is not reasonable, an appropriate physicochemical model can be inferred from the number of the binding sites provided by SVD. Therefore, the combination of SVD and global analysis is a powerful method for analyzing experimental biochemical and biophysical data.

The combined SVD/global analysis method has been applied to analyze structural transitions of proteins monitored by circular dichroism, fluorescence, absorption, and small-angle X-ray scattering.[4,6−8] However, relatively few studies have utilized SVD for heteronuclear NMR.[9−11] SVD has been applied to analyze chemical shift perturbations associated with the pH-dependent conformational transitions of a protein, but without an

Received: October 21, 2011
Published: January 24, 2012

   dx.doi.org/10.1021/ja209936u | J. Am. Chem. Soc. 2012, 134, 3792–3803

accompanying global analysis.[11] Here, we apply a combination of SVD and global analysis to study two-site protein−protein interactions measured by heteronuclear single-quantum correlation (HSQC) titrations.

As an example, we consider binding of the isolated AD1 (residues 13−37) and AD2 (residues 38−61) subdomains of the tumor suppressor p53, derived from the N-terminal transactivation domain (TAD), to the transcriptional adapter zinc finger 2 (TAZ2) domain of cyclic-AMP response element binding (CREB)-binding protein (CBP).[12] Direct interactions between TAZ2 and the p53 TAD are essential for activation of transcription from p53-responsive genes.[13,14] The AD1 and AD2 regions are intrinsically disordered but form stable helical structures upon binding to target proteins.[15,16] Combined SVD and global analysis of NMR titration data reveals that the AD1 and AD2 subdomains each bind at two distinct sites on the surface of TAZ2 and that AD2 has much higher affinities than AD1, with dissociation constants for primary and secondary binding, K_{d1} and K_{d2}, of 32 nM and 10.2 μM, respectively. To validate this small K_{d1} value, we performed extensive simulations by generating titration curves and line shapes with wide ranges of values for K_{d1} and K_{d2}. We find that K_{d1} in the 1−10 nM range can be determined accurately by NMR titrations, provided the secondary binding process is sufficiently tight and is coupled to the primary binding. We also find by line shape analysis that the association rate of AD2 to TAZ2 is 1.7×10^{10} M^{-1} s^{-1}, approaching the diffusion-controlled limit. This suggests that intrinsic disorder in the presence of favorable electrostatic interactions can significantly accelerate protein binding interactions.

RESULTS

Singular Value Decomposition (SVD). For independent multisite protein−protein interactions, the observed ^{1}H or ^{15}N chemical shift, δ_{obs}, can be represented by a linear combination of basis chemical shifts,

$$\delta_{obs} = p_F\delta_F + p_{B1}\delta_{B1} + p_{B2}\delta_{B2} + \cdots + p_{Bn-1}\delta_{Bn-1} \quad (1)$$

where δ_F and δ_{Bi} (i = 1, 2, ..., $n-1$) are the ^{1}H or ^{15}N chemical shifts for the free form and for the bound form in which only the i-th binding site is occupied, respectively, and p_F and p_{Bi} (i = 1, 2, ..., $n-1$) are the fraction of the free form and the occupancy of the i-th binding site, respectively. SVD analysis of a series of HSQC titrations can determine the number of the binding sites ($n-1$) involved in the system if the n species are spectrally distinguishable (see Supporting Information for details).[3]

We performed an HSQC titration in which the unlabeled p53 AD1 peptide was added incrementally to ^{15}N-labeled TAZ2 (Figure 1a and Figure S1a (Supporting Information)). The cross peaks in the spectrum of the TAZ2 domain shift in fast exchange and exhibit curvature, indicating the presence of at least two binding modes. To estimate the number of binding sites, we performed SVD of the raw chemical shift data set (see Materials and Methods for details). In SVD analysis, the number of non-noise components is estimated by taking into account (1) the significance of singular values, (2) the smooth shape (i.e., high autocorrelation) of the $\boldsymbol{v}_i$ vectors, which are related to the occupancies of the binding sites, and (3) a small root-mean-square deviation (rmsd) between the reconstructed and raw data sets.[6,8] For the ^{15}N-TAZ2 titration with AD1, the first three components have significantly large singular values (Figure 2a), and their corresponding $\boldsymbol{v}_i$ vectors have smooth shape and high autocorrelation (Figure 2a,b). The spectral data set reconstructed using only the three components is essentially the same as the raw data set with an rmsd of 0.007 ppm, and inclusion of additional components does not significantly improve the rmsd (Figure 2a). Thus, the SVD analysis indicates the presence of three non-noise components. According to eq 1, one of the non-noise components represents the spectrum of the free form of TAZ2. Thus, the presence of two additional non-noise components indicates that there are two AD1 binding sites on TAZ2.

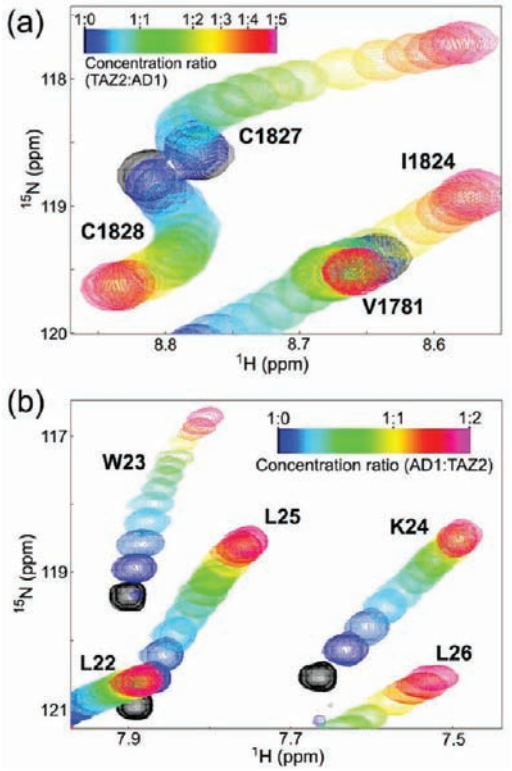


Figure 1. Portions of the ^{1}H−^{15}N HSQC spectra of (a) TAZ2 showing chemical shift changes upon titration with p53 AD1(13−37) and (b) p53 AD1 showing chemical shift changes upon titration of TAZ2. The cross-peak color changes gradually from black (free) to magenta (bound) according to the concentration ratio.

Local vs Global Analysis. Accordingly, we analyzed the HSQC titration curves (Figure S2) assuming a two-site binding model (Scheme 1), where F is the free form of TAZ2, B1 and B2 are the singly bound forms in which only the primary (high affinity) site or the secondary (low affinity) site is occupied by p53, respectively, B12 is the doubly bound form, and L is the free form of the p53 subdomain (see Materials and Methods for details). K_{d1}, k_{on1}, and k_{off1} are the equilibrium dissociation constant, the association rate, and the dissociation rate for primary binding, respectively, and K_{d2}, k_{on2}, and k_{off2} are those for secondary binding. Correction for protein concentration was implemented in the fitting, because accurate determination of protein concentration is crucial for obtaining reliable fits. Fitting each titration curve independently (i.e., local fitting) did not give reliable K_d values; only the curves with large chemical shift changes gave reasonable fits but with large scatter (Figure S3a,b and Table 1). When we used the noise-filtered data set that was reconstructed using only the three non-noise

dx.doi.org/10.1021/ja209936u | J. Am. Chem. Soc. 2012, 134, 3792−3803

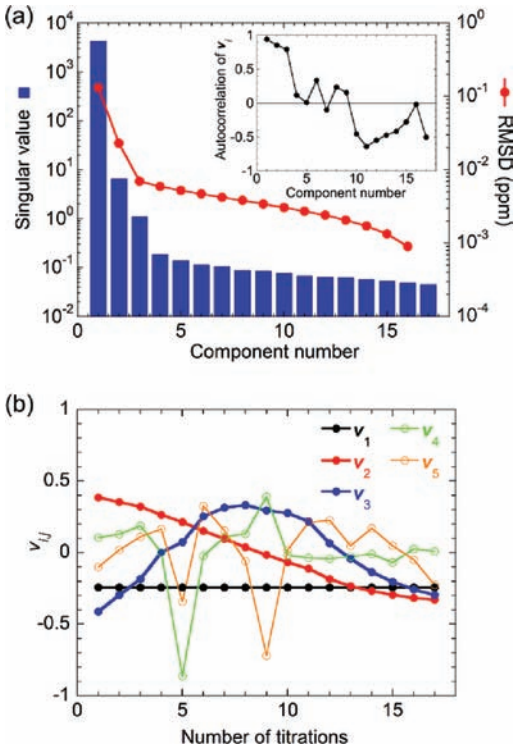
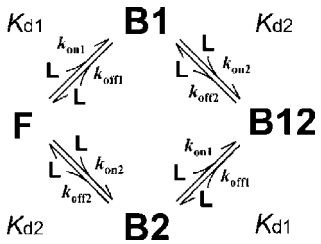


Figure 2. SVD analysis of the ^{15}N-TAZ2 titration with unlabeled p53 AD1. (a) Blue bars show the singular values sorted in decreasing order plotted in the logarithmic scale. Red circles show the rmsd (ppm) between the raw data set and the data set reconstructed using the first through *i*-th components (abscissa). Inclusion of all components (1−17) for the data reconstruction results in the rmsd of 0, which is not shown in the figure. Inset shows the autocorrelation of $\boldsymbol{v}_i$ vectors plotted against the component number. (b) The shape of $\boldsymbol{v}_i$ vectors for the first five components plotted against the number of titrations. The first three components have smooth shape (thick lines with filled circles), resulting in high autocorrelation.

Scheme 1. A Two-Site Binding Model



components, many curves fitted better to the two-site binding model, especially when a concentration correction was implemented (Figure S3c,d and Table 1). However, the K_d values were still scattered and had large standard deviations.

On the other hand, fitting all titration curves globally (i.e., global fitting) gave reliable estimates of K_{d1} and K_{d2} with smaller errors (Table 1). Here, dissociation constants and a concentration correction factor are common for all curves, but the chemical shift differences between the free and the bound forms differ for each curve. All ^{1}H and ^{15}N titration curves were fitted well to the common K_d values (Figure S2). Thus, global fitting is effective in accurately estimating dissociation constants. The use of the noise-filtered data set in the global fit with concentration correction further improved the model fitting and gave K_{d1} and K_{d2} of 24 and 164 μM, respectively, with smaller errors (Figures 3a and S4, and Table 1). Taken together, the global fit of titration curves after noise filtering is the most effective way of estimating accurate dissociation constants from HSQC titration experiments.

Table 1. Dissociation Constants K_{d1} and K_{d2} for Primary and Secondary Binding of Unlabeled p53 AD1 with ^{15}N-Labeled TAZ2

method	data set[a]	correction[b]	no. of curves[c]	K_{d1} (μM)	K_{d2} (μM)
local fit	raw	−	41	27 ± 27[d]	448 ± 788
local fit	raw	1.0 ± 0.3	30	24 ± 26	345 ± 638
local fit	SVD	−	57	33 ± 14	225 ± 280
local fit	SVD	1.1 ± 0.1	67	27 ± 14	126 ± 52
global fit[e]	raw	−	148	31 ± 2	180 ± 7
global fit	raw	1.029 ± 0.009	148	27 ± 3	177 ± 6
global fit	SVD	−	148	32 ± 2	169 ± 4
global fit	SVD	1.071 ± 0.005	148	24 ± 1	164 ± 3

[a]The data set used for fitting: the raw data set (raw) or the reconstructed data set after noise filtering (SVD). [b]With or without (−) correction for the TAZ2 concentration. The correction factor, c_P, is given when the correction was implemented in the fitting. [c]For local fits, only the good fits in which both K_{d1} and K_{d2} are within 10^{-6}−10^{-2} M and in which the fitting errors are smaller than the fitted values were used in calculating the average and standard deviation of the K_d values and the correction factor. The number of curves satisfying this criteria is shown. For global fits, all curves showing clear fast-exchange shifts were used for fitting. [d]Standard deviation. [e]For global fit, errors were estimated by the Monte Carlo error estimation method.

In addition to the dissociation constants, the global fit also gives the ^{1}H and ^{15}N chemical shift differences between the free and the bound forms for both primary and secondary binding (Figure S5). Effects of binding events on each residue are represented by an average chemical shift difference, $\Delta\delta_{av} = [(\Delta\delta_H)^2 + (\Delta\delta_N/5)^2]^{1/2}$ (Figure 3b).[17] For the primary binding, a large chemical shift difference was observed especially for Val1802, Leu1823, and Leu1826 ($\Delta\delta_{av}$ > 2 × standard deviation (SD) from the mean). On the other hand, the secondary binding largely affected Ala1825, Cys1827, Tyr1829, and Lys1832. Mapping these residues onto the NMR structure of TAZ2 shows that the residues affected by the primary and secondary AD1 binding are located at two opposite faces of TAZ2 (Figure 3c).

Complementary Titration. As an alternative way of measuring the same binding events, we measured the HSQC titrations of ^{15}N-labeled AD1 with unlabeled TAZ2 (Figures 1b and S1b). The AD1 peaks also show fast-exchange shifts and a slight curvature, again indicating the presence of at least two binding modes. The SVD analysis indicates the presence of three non-noise components, one of which is the spectrum for the free form of AD1 (Figure S6a,b). Thus, the data set reconstructed using only the three non-noise components was globally fitted with the two-site binding model to obtain K_{d1} and K_{d2} of 38 and 263 μM, respectively (Figure S6c and Table S1).

Although the K_d values are consistent with those obtained from the ^{15}N-TAZ2 titration with AD1, they have larger uncertainties. This is because the secondary binding event is less prominent for the ^{15}N-AD1 titration with TAZ2 compared with the ^{15}N-TAZ2 titration with AD1 (Figure S7). Thus, it is preferable to titrate ligand (that binds at multiple sites) into ^{15}N-labeled protein (that has multiple binding sites), in order to estimate accurate K_d values for multisite binding.

To estimate chemical shift changes for ^{15}N-AD1, the titration curves were fitted using the K_d values obtained by ^{15}N-TAZ2

dx.doi.org/10.1021/ja209936u | J. Am. Chem. Soc. 2012, 134, 3792−3803

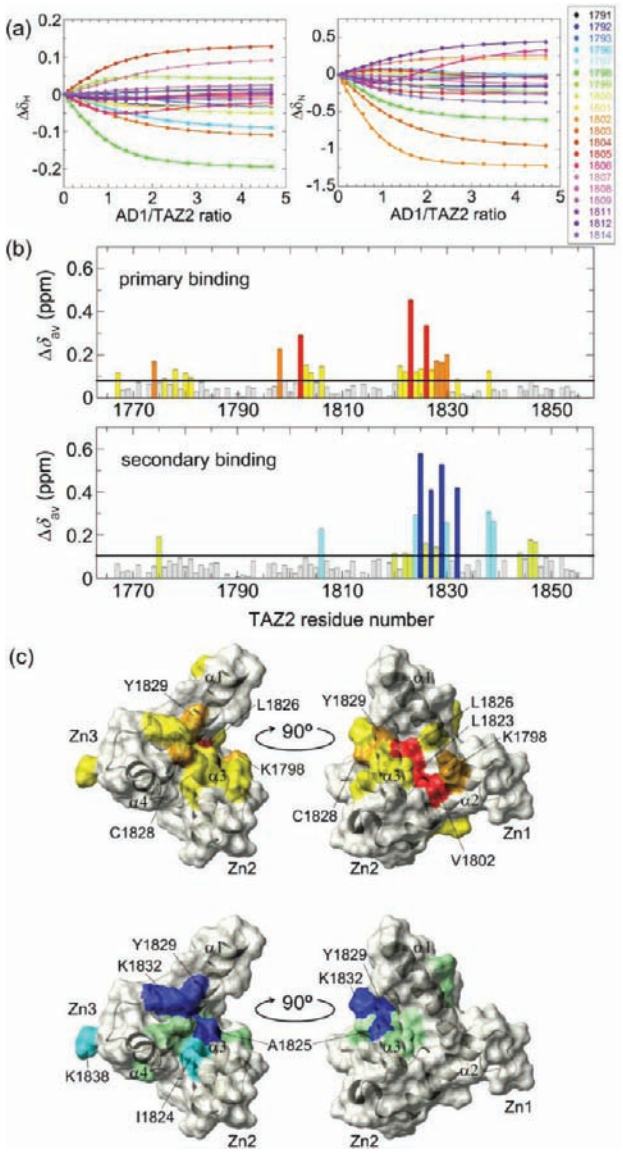


Figure 3. Global fitting of the noise-filtered titration curves for the ^{15}N-TAZ2 titration with unlabeled p53 AD1. (a) Selection of the data from the global fit of chemical shift changes of HSQC cross peaks of ^{15}N-TAZ2 as a function of the concentration ratio of AD1/TAZ2. Color codes are shown in the panel along with the residue number. (b) Histograms of the averaged chemical shift differences $\Delta\delta_{av}$ for the primary (upper) and secondary binding (lower). The black horizontal line shows the mean of all $\Delta\delta_{av}$ (0.081 and 0.104 ppm for primary and secondary binding, respectively). The residues are categorized into the following groups: (red or blue) greater than or equal to mean + 2SD; (orange or cyan) mean + 1SD to mean + 2SD; (yellow or pale green) mean to mean + 1SD; and (gray) less than mean. (c) Mapping the location of the residues having large $\Delta\delta_{av}$ onto the NMR structure for the primary (upper) and secondary binding (lower). The color codes are the same as in (b). Three zinc binding sites and four α-helices are shown. Helix α1 corresponds to residues 1765−1785, α2 to 1794−1806, α3 to 1818−1833, and α4 to 1842−1852.

titration with AD1 (Table 1). Upon binding of p53 AD1 to TAZ2, chemical shifts of residues 19−26 of p53 are affected (Figure 4), suggesting that this region interacts with TAZ2.

TAZ2:AD2 Interactions. Next we performed the HSQC titration of ^{15}N-labeled TAZ2 with the unlabeled AD2 peptide from p53 (Figure S8). The SVD analysis indicates the presence

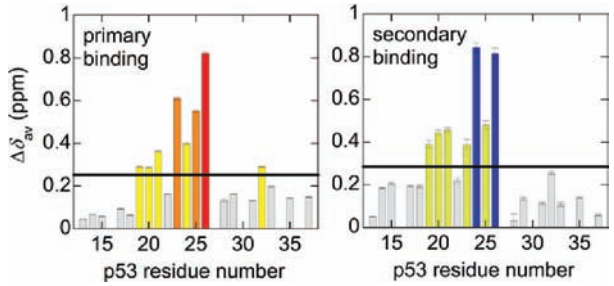


Figure 4. $\Delta\delta_{av}$ histograms of the primary (left) and secondary binding (right) obtained from the ^{15}N-p53 AD1 titration with unlabeled TAZ2 using the K_d values obtained by the ^{15}N-TAZ2 titration with AD1. The TAZ2 concentration correction factor was 1.08 ± 0.01. The black horizontal line shows the mean of all $\Delta\delta_{av}$ (0.251 and 0.286 ppm for primary and secondary binding, respectively). The residues are categorized as described in Figure 3b.

of three non-noise components, one of which is the spectrum for the free form of AD2 (Figure S9a,b; see Supporting Information for details). Thus, the data set reconstructed using only the three non-noise components was globally fitted to the two-site binding model to get K_{d1} and K_{d2} of 32 nM and 10.2 μM, respectively (Table S2, Figures 5a and S10). The K_{d1} value is close to the K_d for the binding of p53(13−61) to TAZ2 measured by fluorescence titrations (~10−20 nM), supporting the validity of this small value.[18] Large chemical shift differences were observed especially at Val1819, Ala1825, and Leu1826 for the primary binding and at Thr1806, Cys1827, Lys1832, and Lys1838 for the secondary binding (Figure S9c). Although the HSQC cross peaks for Val1802 and Leu1823 broadened at higher ratios and were not used for the global fit, their $\Delta\delta_{av}$ values at a 1:0.6 ratio were the seventh largest and the largest among all residues, respectively, indicating that both Val1802 and Leu1823 are greatly affected by the primary AD2 binding. The cross peak for Tyr1829 broadened above 1:2.5 ratios, but the chemical shift change between 1:1.0 and 1:2.5 ratios was the largest for Tyr1829 among all residues, indicating that Tyr1829 is strongly affected by the secondary binding. Mapping these residues onto the NMR structure of TAZ2 shows that the regions affected by the primary and secondary AD2 binding are very similar to those for the AD1 binding (Figure 5b).

Ranges of Accurate K_{d1} and K_{d2} Determined by NMR Titrations. To accurately determine dissociation constants for one-site binding processes from titration experiments, the protein concentration should be close to K_d.[2] Given the relatively high protein concentration typically used for NMR detection (ca. 100 μM or greater), it is difficult to determine K_d in the nanomolar range for single-site binding by NMR titrations, although the use of lower protein concentrations may expand the lower limit of K_d in favorable cases (see Supporting Information for details). In marked contrast, it is possible to accurately estimate a K_{d1} of less than 100 nM by NMR titrations for two-site binding, even at high protein concentrations, because secondary binding reduces the occupancy of the primary site.[12] However, accurate estimates of both K_{d1} and K_{d2} require that the primary and secondary binding events are coupled; if the secondary binding event is very weak, then the situation once again approximates to a one-site binding event. Our previous simulations show that K_{d1} of 10 nM−1 μM can be reliably obtained by global fit if K_{d2} is between 10^1- and 10^3-fold of K_{d1} and that, if the dissociation rate

dx.doi.org/10.1021/ja209936u | J. Am. Chem. Soc. 2012, 134, 3792−3803

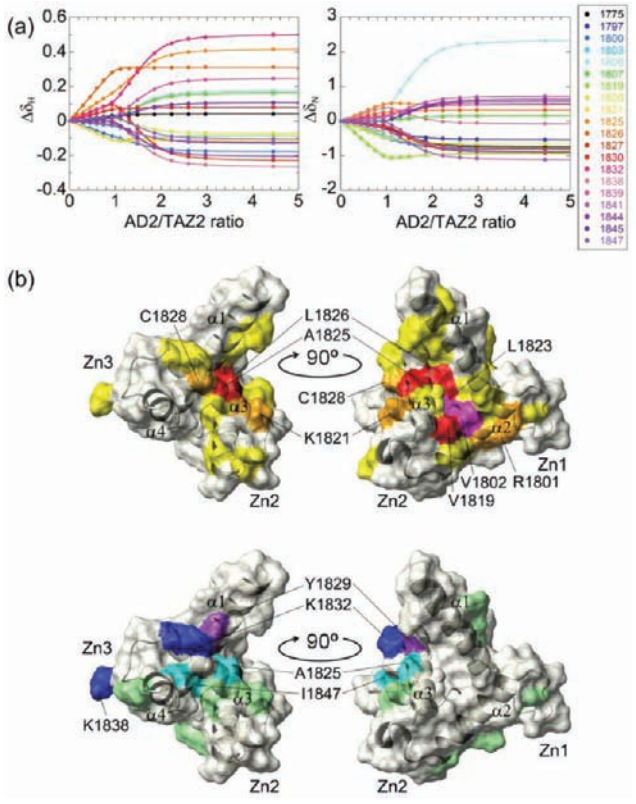


Figure 5. Global fitting of the noise-filtered titration curves for the ^{15}N-TAZ2 titration with unlabeled p53 AD2. (a) Selection of the data from the global fit of chemical shift changes of HSQC cross peaks of ^{15}N-TAZ2 as a function of the ratio of AD2/TAZ2. Color codes are shown in the panel along with the residue number. (b) Mapping the location of the residues having large $\Delta\delta_{av}$ onto the NMR structure for the primary (upper) and secondary binding (lower). The color codes are the same as in Figure 3b, except that Val1802 and Leu1823 are shown by magenta (upper) and Tyr1829 is shown by purple (lower).

is rapid enough such that resonances shift in fast exchange, even a K_{d1} of 1 nM can be measured if K_{d2} is in the range 0.1−10 μM.[12]

To examine whether this conclusion holds when experimental conditions are changed, such as the number of data points, the maximum titrant/protein ratio, the protein concentration, and the number of peaks used in the analysis, we simulated titration curves by changing these conditions and examined whether the K_{d1} and K_{d2} values used for simulation can be reproduced by global fitting (Figure 6). Here, K_{d1} and K_{d2} were set as multiples of 10 between 10^{-10} and 10^{-1} M ($K_{d1} < K_{d2}$), and chemical shift changes obtained for ^{15}N-TAZ2 titration with AD1 (Figure S5) were used for simulations (see Materials and Methods for details). First, we changed the number of data points by fixing the maximum titrant/protein ratio to 1:5. For example, the number of data points was 25 when the interval of titrations was 1:0.2. The fitting simulations showed that the K_{d1} of 10^{-9}−10^{-4} M was accurately reproduced by global fitting if K_{d2} was 10^{1}−10^{3}-fold of K_{d1} (Figures 6a and S11). On the other hand, when the interval of titrations was 1:1 and the number of data points was 5, only the K_{d1} of 10^{-7}−10^{-4} M was reproduced by global fitting (Figure S11). Thus, the increase in the number of data points, i.e., the decrease in the interval of titrations, improves fitting and widens the ranges of sets of K_{d1} and K_{d2} that are accurately reproduced by global fitting. Therefore, it is preferable to measure titration curves at an interval of 1:0.2 (Figure 6a). Interestingly, the titration curves with an interval of 1:0.5 also gave good estimates of K_{d1} and K_{d2} (Figure S11), probably because measurement at a 1:1 ratio is important for accurate estimates. In our experiments, we measured HSQC spectra at 1:0, 0.1, 0.2, 0.4, 0.6, 0.8, 1.0, 1.25, 1.5, 1.75, 2.0, 2.5, 3.0, 3.5, 4.0, 4.5, and 5.0 concentration ratios (17 data points), which are combinations of 1:0.2, 0.25, and 0.5 ratios. Thus, the present simulations show the validity of the use of these ratios for accurate estimates of K_d values. This set of concentration ratios was used in the following simulations.

Changing the maximum titrant/protein ratio in global fitting indicated that titration curves should be measured up to at least 1:2 ratio and, preferably, more than 1:3 ratio (Figures 6b and S12). Protein concentration dependence on titration experiments shows that lower K_ds are accurately estimated at lower protein concentrations and *vice versa* (Figures 6a,c and S13). Finally, changing the number of peaks used in global fitting indicated that the K_d values are more accurately estimated by using more resonance peaks (Figures 6d and S14); the use of a single peak (two curves from ^{1}H and ^{15}N chemical shifts) does not give an accurate K_{d1} in nM ranges, and the use of more than 20 peaks (40 curves) is preferable. With these points taken together, it is possible to accurately estimate a K_{d1} in nM ranges if the primary and secondary bindings are coupled and if the above criteria are satisfied. Our titration experiments fulfilled all these criteria, demonstrating the accuracy of the K_d values for the TAZ2:AD1/2 interactions.

Lower Limit of Accurate K_{d1}. An essential prerequisite for this analysis is that the chemical exchange process is in the fast-exchange regime. To estimate the K_{d1} range showing fast exchange, we simulated a series of one-dimensional ^{1}H line shapes for two-site binding at 1:0−1:5 concentration ratios, taking into account k_{on} and k_{off} rates explicitly. ^{1}H chemical shift changes for primary and secondary binding were assumed to be ∼0.1 ppm, because the mean $\Delta\delta_H$ values for the primary and secondary binding of p53 AD1/2 with TAZ2 were 0.05 and 0.07 ppm, respectively. k_{on} and k_{off} were set as multiples of 10 between 10^{4} and 10^{11} M^{-1} s^{-1} and between 10^{-2} and 10^{8} s^{-1}, respectively. Combinations of k_{on1}, k_{on2}, k_{off1}, and k_{off2} that satisfy K_{d1} and K_{d2} of 10^{-10}−10^{-1} M ($K_{d1} < K_{d2}$) were used to calculate line shapes (see Materials and Methods for details). For example, 336 combinations of rate constants satisfied a K_{d1} of 1 μM, and 132 combinations among them showed fast exchange throughout titrations from 1:0 to 1:5 ratios; in other words, the fraction of fast exchange was 39% (=132/336) at a K_{d1} of 1 μM (Figures 7a and S15a). The simulations showed that the fraction of fast exchange is 0% at a K_{d1} of 0.1 nM, increases with increasing K_{d1}, and finally reaches 100% at a K_{d1} of more than 10^{-2} M (Figure 7a). The results indicate that a K_{d1} of 0.1 nM cannot be determined by NMR titrations, and the lower limit of K_{d1} determined by the present method is 1 nM. Note that the lower limit of K_{d1} also depends on chemical shift changes for the primary and secondary binding. When the ^{1}H chemical shift differences were assumed to be ∼0.4 ppm in the line shape simulations, close to the maximum $\Delta\delta_H$ for ^{15}N-TAZ2 titrations with AD1 (Figure S5a), the lower limit of K_{d1} became 10 nM (Figure S15b). Therefore, the lower limit of K_{d1} that can be accurately determined by NMR titration ranges from 1 to 10 nM, depending on the particular system.

In Scheme 1, we neglected an interconversion between the B1 and B2 forms. However, it might be possible that the p53 AD1 or AD2 subdomain bound to one site of TAZ2 moves to another site without dissociation as shown in Scheme 2, where

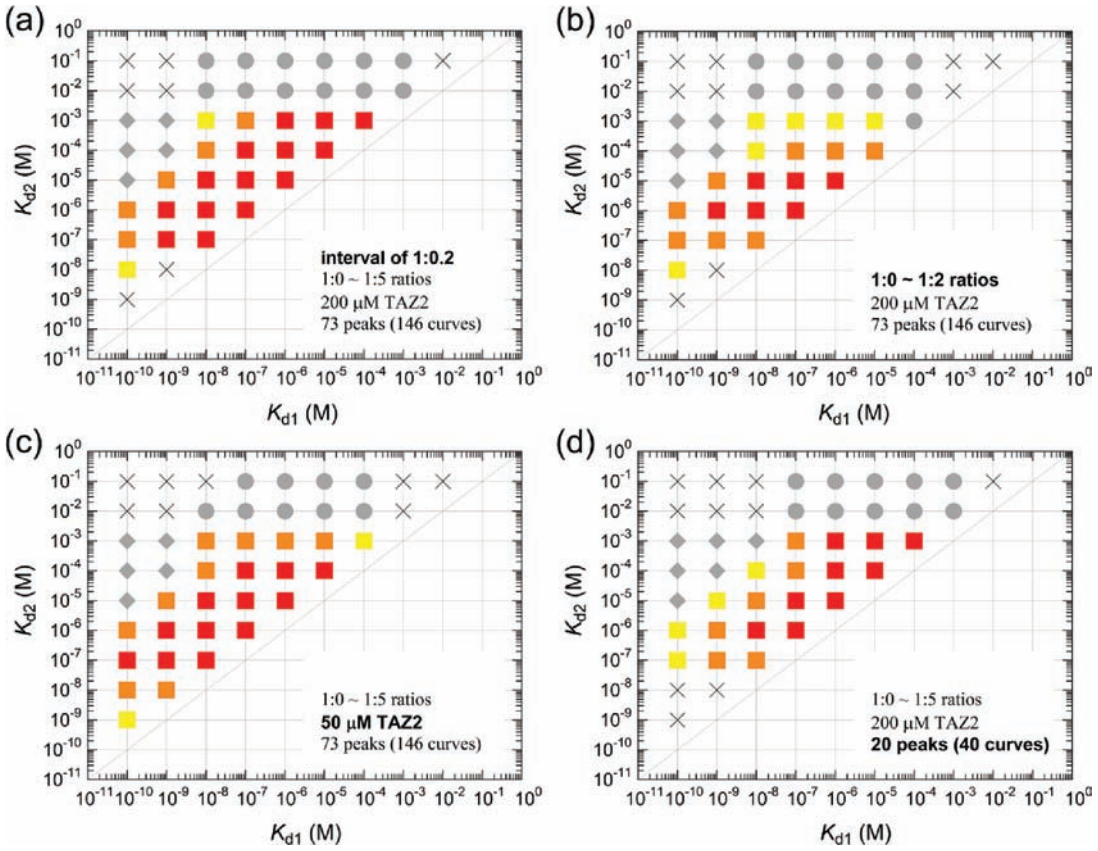


Figure 6. Dependence of accurate K_{d1} and K_{d2} ranges on (a) the number of data points (an interval of ratios), (b) the maximum titrant/protein ratio, (c) the protein concentration, and (d) the number of peaks used in the analysis. Representative results are shown (see Figures S11–S14 for details). Conditions of fitting simulations are described in each panel. The red, orange, and yellow squares and cross symbols show that both the fitting errors and the differences between the fitted K_d and the input K_d used for generating the curves are less than 5%, 5–10%, 10–20%, and more than 20% of the input K_d values, respectively. Gray circles and diamonds show that only K_{d1} or K_{d2} was accurately obtained by global fit, because the binding events approached one-site binding at higher K_{d2} or at lower K_{d1}, respectively. It should be noted that a K_{d1} of less than 1 nM cannot be estimated by NMR titrations, because such a tight binding does not show fast-exchange shifts (see Figure 7).

k_{12} and k_{21} are the rate constants for the interconversion from B1 to B2 and from B2 to B1, respectively. The presence of the interconversion may affect the chemical shift time scale. As expected, line shape simulations based on Scheme 2 showed that the presence of interconversion increases the fraction of fast exchange (Figure 7b); more than 40% increase in the fraction of fast exchange was observed at a K_{d1} of 10^{-8}–10^{-6} M (Figure 7c). Thus, fast exchange for a nanomolar dissociation constant is not unrealistic, and our NMR titration analysis is also applicable to the cases of tight binding phenomena accompanied by rapid association/dissociation equilibrium.

k_{on} and k_{off} Rates for TAZ2:AD2 Interactions. To estimate k_{on} and k_{off} that fulfill a small K_{d1} of 32 nM for the TAZ2 interactions with AD2, we performed fitting of line shapes of a series of titrations (see Materials and Methods for details). Here, one-dimensional slices of cross peaks in the ^{1}H dimension were used (Figures 8 and S16), and K_{d1} and K_{d2} were fixed to 32 nM and 10.2 μM, respectively. Global fitting of ^{1}H line shapes for many peaks gave a k_{on1} of $(1.7 \pm 0.2) \times 10^{10}$ M^{-1} s^{-1}, which is extremely fast but not beyond the diffusion-controlled limit (see below). k_{off1} was estimated to be 540 ± 80 s^{-1}, which is reasonably fast to show fast-exchange shifts. On the other hand, k_{on2} and k_{off2} for the secondary TAZ2:AD2 binding were $(7.1 \pm 0.8) \times 10^8$ M^{-1} s^{-1} and 7200 ± 900 s^{-1}, respectively. Thus, slower k_{on} and faster k_{off} compared with the primary binding resulted in larger K_{d2}.

It should be noted that the fast k_{on1} obtained by line shape analysis is not an artifact due to assumptions introduced in the fitting, because a k_{on1} of 10^{10} M^{-1} s^{-1} is readily predicted as follows. Fast-exchange shifts for TAZ2:p53 AD2 interaction imply that $k_{off1} \geq 500$ s^{-1}.[1] Here, K_{d1} is 32 nM, suggesting that $k_{on1} = k_{off1}/K_{d1} \approx 500/(32 \times 10^{-9}) \approx 1.6 \times 10^{10}$ M^{-1} s^{-1}. Our line shape analysis is fully consistent with this estimate.

DISCUSSION

p53 AD1 and AD2 Binding Sites on TAZ2. Primary binding of the p53 AD1 peptide largely affects the chemical shifts of the residues located on a hydrophobic surface at the interface between the α1, α2, and α3 helices. This region corresponds to the binding site previously identified by chemical shift mapping using the p53(14–28) peptide.[19] On the other hand, secondary binding of AD1 affects residues located on the opposite surface of TAZ2, composed mainly of residues in the α1, α3, and α4 helices. Because these two regions overlap with the binding sites of the STAT1-TAD[20] and E1A,[21] it is reasonable to consider that the large chemical shift changes are due to direct interactions with AD1. The chemical shifts of some residues that are distant from the binding surfaces, such as Thr1806 and Lys1838, are also slightly affected by the secondary binding process, probably due to propagation of perturbations from the binding surfaces or local conformational changes rather than direct binding (see

dx.doi.org/10.1021/ja209936u | J. Am. Chem. Soc. 2012, 134, 3792–3803

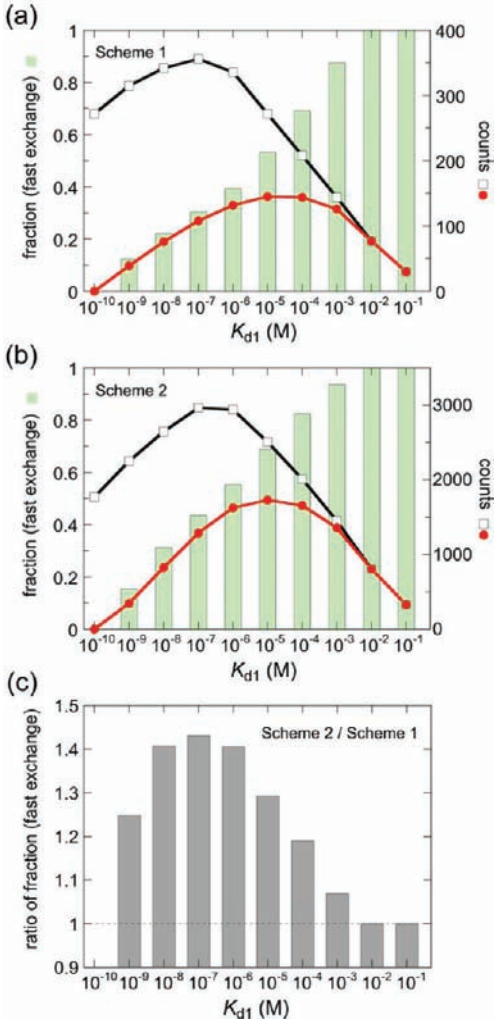
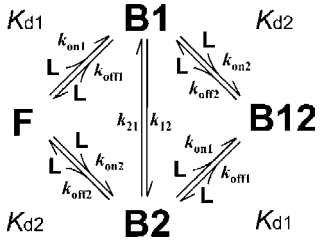


Figure 7. (a) K_{d1} dependence of the fraction of fast exchange for Scheme 1. Chemical shift changes were assumed to be ~0.1 ppm. Combinations of k_{on1}, k_{on2}, k_{off1}, and k_{off2} that satisfy K_{d1} and K_{d2} of 10^{-10}–10^{-1} M ($K_{d1} < K_{d2}$) were used to calculate line shapes. Black open squares show the counts of the combinations of rate constants that produce the indicated K_{d1} value. Red filled circles show the counts of the combinations of rate constants that show fast-exchange shifts throughout the titrations from 1:0–1:5 ratios. Green bars are the ratio of these values, corresponding to the fraction of fast exchange at the indicated K_{d1} value. (b) The same as (a), except that the interconversion between B1 and B2 is taken into account (Scheme 2). (c) Increase in the fraction of fast exchange by the presence of the B1–B2 interconversion. The gray bar shows the ratio of the fraction of fast exchange in the presence of the B1–B2 interconversion (Scheme 2) to that in the absence of the interconversion (Scheme 1).

Scheme 2. A Two-Site Binding Model with an Interconversion between B1 and B2



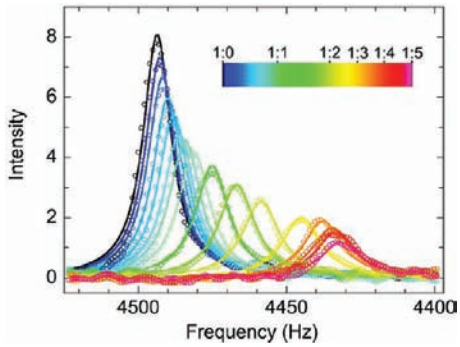


Figure 8. Series of ^{1}H line shapes for Val1841 observed in the ^{1}H–^{15}N HSQC titration of ^{15}N-TAZ2 with unlabeled p53 AD2 in which TAZ2:AD2 concentration ratios ranged from 1:0 to 1:5. The line shape color changes from black (free) to magenta (bound) according to the concentration ratio. Continuous lines are obtained from global fitting of line shapes.

Supporting Information for more discussion). Recently Feng et al. have determined the NMR structure of the complex between TAZ2 from p300 and the p53(2–39) peptide.[22] The structure is consistent with our chemical shift mapping results in that Val1802, Leu1823, and Leu1826 of TAZ2 have direct hydrophobic contacts with p53 and that residues 17–24 of the p53 AD1 subdomain form an α-helix upon binding.

Comparison between the primary and secondary binding sites indicates that the primary site is more hydrophobic than the secondary site. This suggests that hydrophobic interactions are important in stabilizing the TAZ2–p53 AD1 interactions.

Unexpectedly, the primary and secondary AD2 binding sites on TAZ2 overlap with those for AD1 binding. However, there is a slight difference between AD1 and AD2 binding. The primary binding site of AD2 overlaps with that of AD1 at Val1802, Leu1823, and Leu1826, but Val1819 and Ala1825 are also affected by AD2, indicating binding to a broader TAZ2 surface. The reason why AD2 binds more tightly than AD1 may be that highly negatively charged AD2 (−8.0) compared with AD1 (−2.0) is more favorable for the electrostatic interactions with the highly positively charged TAZ2 (+14.3).

Thus, AD1 and AD2 can bind to the same primary and secondary binding sites. However, in the full-length p53 TAD, AD1 and AD2 are connected by a linker, and it is unlikely that both of them can bind simultaneously to the primary site of a single TAZ2 molecule, considering the size of the binding site. Since the affinity of AD2 for the primary site is significantly higher than that of AD1 (K_{d1} 32 nM versus 24 μM), it is most likely that AD2 occupies the primary binding site and the AD1 the secondary site in the binary complex between TAZ2 and the full-length p53 TAD. In support of this, the titration of the AD2 peptide into the ^{15}N-AD1:TAZ2 (1:1) complex outcompetes the ^{15}N-AD1, but the titration of the AD1 peptide into the ^{15}N-AD2:TAZ2 (1:1) complex only inefficiently competes with the ^{15}N-AD2, indicating that both AD1 and AD2 bind the same high-affinity site on TAZ2 and that AD2 binds more tightly than AD1.

Extremely Fast Association Rate for TAZ2–p53 Binding. Previous studies have reported that a protein–protein association rate can be larger than 10^{9} M^{-1} s^{-1} (Table 2);[23–27] the hydroxylated HIF-1α/TAZ1 association rate was shown to be ~10^{9} M^{-1} s^{-1}.[28] However, a k_{on} larger than 10^{10} M^{-1} s^{-1} has not yet been reported. To our knowledge, a k_{on} of 1.7×10^{10} M^{-1} s^{-1}

dx.doi.org/10.1021/ja209936u | J. Am. Chem. Soc. 2012, 134, 3792–3803

Table 2. Fast Association Rates for Protein−Protein, Protein−Ligand, and Protein−Nucleic Acid Interactions

protein	target	k_{on} (M^{-1} s^{-1})	ref
	Protein−Protein		
TAZ2	p53 AD2	1.7×10^{10}	this study
barnase	barstar	$>5 \times 10^{9}$	26
E9 DNase	Im9	6×10^{9}	24
colicin E9	Im9	4×10^{9}	24
cytochrome peroxidase	cytochrome *c*	$>3 \times 10^{9}$	25
TAZ1	HIF−OH	1.3×10^{9}	28
acetylcholinesterase	fasciculin	1×10^{9}	27
thrombin	hirudin	$\sim 1 \times 10^{9}$	23
	Protein−Ligand		
acetylcholinesterase	TFK^{+}	$2-9 \times 10^{10}$	27
fumarase	fumarate	3×10^{10}	30
superoxide dismutase	superoxide	2×10^{10}	31
	Protein−Nucleic Acid		
lac repressor	operator DNA	2×10^{10}	44
restrictocin	rRNA	1.7×10^{10}	46
RNase A	RNA substrate	3×10^{9}	45

for TAZ2 binding with p53 AD2 is the fastest association rate for protein−protein interactions.

Although a diffusion-limited association rate was theoretically estimated to be $10^{9}-10^{10}$ M^{-1} s^{-1} for two uniformly reactive spheres with an equal size,[29] the association rate for binding of the p53 AD2 subdomain to TAZ2 is not exceptionally large if the size, geometry, and charges of the two interacting molecules are taken into account. First, an association rate can be faster if one molecule is small and diffuses rapidly while the other is large and provides a large target.[29] Indeed, fast association processes with a k_{on} of $>10^{10}$ M^{-1} s^{-1} have been reported for the interactions between a protein and a small ligand (Table 2).[27,30,31] Second, for a molecule with rod-like geometry (prolate ellipsoid), which can be adopted by nucleic acids, the effective capture radius approaches the length of the major semi-axis of the ellipsoid, leading to a higher association rate than for a spherical molecule.[29] Although intrinsically disordered proteins (IDPs) often have more compact structures than chemically denatured proteins, they are more expanded than molten globules.[32−34] In accord with this, Wolynes and co-workers have predicted on theoretical grounds that the association rates of IDPs will be enhanced substantially by the "fly casting" mechanism.[35,36] However, a recent study has suggested that a greater capture radius of an IDP also leads to slower translational diffusion, which opposes fast association.[37] Finally, particularly favorable, complementary electrostatic interactions can accelerate binding reactions dramatically, giving a k_{on} range of $10^{9}-10^{11}$ M^{-1} s^{-1}.[26,38−43] Consistent with this, association processes with a k_{on} of $\sim 10^{10}$ M^{-1} s^{-1} have been reported for protein−nucleic acid interactions (Table 2).[44−46]

The p53 AD2 motif fulfills these requirements: it is small, it is intrinsically unstructured in its free form, and it has negative charges that favor electrostatic interaction with the highly positively charged TAZ2. Thus, a k_{on} of 1.7×10^{10} M^{-1} s^{-1} is within the range expected for binding processes of IDPs with favorable electrostatic interactions (see Supporting Information for more discussion). Among these requirements for fast association, complementary electrostatics probably plays a major role in determining the very rapid binding kinetics of the AD2 motif. Accordingly, k_{on} for the secondary AD2 binding process is slower, because the second p53 AD2 molecule experiences a smaller electrostatic driving force for binding after the reduction of the net positive charge on TAZ2 due to binding of the first AD2 molecule.

In general, IDPs tend to be deficient in hydrophobic residues necessary for folding and rich in charged residues, resulting in a large net charge at neutral pH.[34] Therefore, it is plausible that a large net charge causes both intrinsic disorder and rapid binding, suggesting that fast association is an intrinsic nature of IDPs.

This fast association rate is comparable to the helix folding rate. A relaxation time constant for helix folding is typically 0.2−2 μs.[47] In our experimental conditions, the TAZ2 and p53 AD2 concentrations are on the order of 10^{-4} M. Here, the binding rate constant is 10^{6} s^{-1} ($=10^{10}$ M^{-1} s^{-1} $\times$ 10^{-4} M), which corresponds to 1 μs. This value is comparable to the helix folding rate. The protein concentrations *in vivo* are much lower than in our NMR studies, indicating that the binding time constant is longer than 1 μs, within which helix formation can occur. Therefore, even if the mechanism of TAZ2 recognition by p53 AD2 is conformational selection,[48] helix formation of AD2 can precede the TAZ2 binding reaction.

CONCLUSIONS

We have presented here a quantitative method to analyze two-site protein interactions monitored by NMR chemical shift perturbation. Our method extends the lower limit of K_d determined by NMR to the 1−10 nM range under favorable exchange conditions. The method is applicable not only for protein−protein interactions but also for protein−ligand interactions and for conformational changes in a single protein molecule. Application of the method to the binding of IDPs revealed the presence of two p53 AD1 binding sites on TAZ2. Moreover, these sites overlapped with the site of binding of AD2 on the hydrophobic surface of TAZ2, but AD2 binds to TAZ2 more tightly than AD1. These results highlight the complexity of the interactions between IDPs and their targets. Thus, careful investigation of binding interactions is necessary for understanding the molecular mechanisms involving IDPs and "hub" proteins such as transcriptional coactivators. The binding of AD2 to TAZ2 is extremely fast, approaching the diffusion-controlled limit and suggesting that intrinsic disorder plus complementary electrostatics can significantly accelerate protein binding interactions.

MATERIALS AND METHODS

Protein Expression and Purification. Unlabeled and ^{15}N-labeled TAZ2 (residues 1764−1855) domain of mouse CBP and unlabeled and ^{15}N-labeled p53 AD1 (13−37) and AD2 (38−61) subdomains were expressed and purified as described previously.[12]

NMR Spectroscopy. NMR spectra were recorded using Bruker 500, 600, and 800 MHz spectrometers and analyzed using NMRPipe[49] and NMRView.[50] Published backbone assignments for TAZ2 and p53 were used.[12] ^{1}H−^{15}N HSQC titrations were performed to characterize binding of the TAZ2 domain to the p53 AD1 and AD2 subdomains. Conditions were 20 mM Tris-HCl (pH 6.8), 50 mM NaCl, 1 mM DTT, and 10% D_2O at 35 °C. For ^{15}N-TAZ2 titration with p53 AD2, the initial concentration of ^{15}N-TAZ2 was 200 μM, and aliquots of concentrated p53 AD2 were titrated into the NMR tube. For ^{15}N-TAZ2 titration with p53 AD1 and for ^{15}N-p53 AD1 titration with TAZ2, the concentrations of ^{15}N-TAZ2 and ^{15}N-p53 AD1 were 200 μM throughout the titrations.

SVD Analysis. In the SVD analysis, the chemical shift data of each HSQC spectrum were represented as a one-dimensional column vector

that contains δ_H and δ_N values, $\boldsymbol{d} = \{\delta_{H1}, ..., \delta_{Hi}, ..., \delta_{Hm}, \delta_{N1}, ..., \delta_{Ni}, ..., \delta_{Nm}\}$, where m is the number of peaks (or residues) used in the analysis, and i (=1 − m) is the peak number. Then, the HSQC spectra obtained by a series of n titrations were represented as a $2m \times n$ data matrix $\boldsymbol{D}$, in which the j-th column $(\boldsymbol{D})_j$ is given by the chemical shift vector $\boldsymbol{d}$ for the j-th titration. The SVD algorithm expresses the data matrix in the following form:

$$\mathbf{D} = \mathbf{UWV}^{\mathrm{T}} \tag{2}$$

where $\boldsymbol{U} = (\boldsymbol{u}_1, ..., \boldsymbol{u}_n)$ is a $2m \times n$ matrix of orthogonal columns, which form a complete set of n basis chemical shift vectors $\boldsymbol{u}_i$; $\boldsymbol{W}$ is a diagonal $n \times n$ matrix of singular values sorted in decreasing order; and $\boldsymbol{V} = (\boldsymbol{v}_1, ..., \boldsymbol{v}_n)$ is an $n \times n$ matrix whose elements measure the proportion of each $\boldsymbol{u}_i$ basis vector attributable to each of the singular values.[6] Equation 2 indicates that the chemical shift vector at any concentration ratio can be expressed as a linear superposition of the orthogonal $\boldsymbol{u}_i$ bases. In the ideal case without noise, the rank of the matrix $\boldsymbol{D}$ corresponds to the number of mathematically independent states. However, for real data with noise, the number of non-noise components, n_c, must be estimated by taking into account (1) the singular values, (2) the shape and the autocorrelation of the $\boldsymbol{v}_i$ vectors, and (3) the rmsd of the SVD reconstruction.[6,8] The i-th singular value corresponds to the weight of the $\boldsymbol{u}_i$ basis vector in reconstituting the data matrix, and thus the large singular value indicates the significance of the corresponding base. Because each $\boldsymbol{v}_i$ vector is described as a linear combination of p_F and p_{Bi}, the $\boldsymbol{v}_i$ vector corresponding to a non-noise component should have a smooth shape and, consequently, a high autocorrelation. The autocorrelation and shape of each $\boldsymbol{u}_i$ basis vector were not used for this purpose, because a one-dimensional chemical shift vector, $\boldsymbol{d}$, usually does not have a smooth shape and its autocorrelation is close to 0 (Figure S17).

For independent multisite protein−protein interactions, the number of non-noise components thus obtained corresponds to the number of binding sites + 1 (for the free form) according to eq 1. To confirm that the spectrum for the free form is one of the non-noise components, we used a chemical shift difference referenced to the free form, $\Delta\delta = \delta_{obs} - \delta_F$, in constructing the one-dimensional chemical shift vector $\boldsymbol{d}$ and then the data matrix $\boldsymbol{D}$. The SVD results showed that the number of non-noise components is reduced by 1, demonstrating that the spectrum for the free form is indeed one of the non-noise components.

After the number of non-noise components n_c is determined, it is possible to reconstruct the data set using only the non-noise components and discarding the noise components. The component reduction (noise filtering) is done by constructing the reduced $\boldsymbol{U}$, $\boldsymbol{W}$, and $\boldsymbol{V}$ matrices using only the non-noise components, resulting in $2m \times n_c$, $n_c \times n_c$, and $n \times n_c$ matrices, respectively. Then, back calculation of the data matrix $\boldsymbol{D}$ using eq 2 provides the noise-filtered data matrix, the use of which will increase the reliability of the subsequent model fitting.

Binding Models. In the present study, interactions of the p53 AD1 and AD2 subdomains with the TAZ2 domain of CBP showed two-site binding, according to the SVD analysis. In the subsequent global analysis, we assumed an independent two-site binding model. In this model, the TAZ2 domain of CBP exists in the free form (F), the singly bound forms (B1 and B2) in which only the primary (high affinity) site or the secondary (low affinity) site is occupied, respectively, and the doubly bound form (B12). The p53 subdomain in the free form is denoted as L. For a two-site binding model with no interconversion between B1 and B2 (Scheme 1), the fractions of the F, B1, B2, and B12 forms, f_F, f_{B1}, f_{B2}, and f_{B12}, respectively, are described as

$$f_F = \frac{K_{d1}K_{d2}}{K_{d1}K_{d2} + (K_{d1} + K_{d2})[L] + [L]^2}$$

$$f_{B1} = \frac{K_{d2}[L]}{K_{d1}K_{d2} + (K_{d1} + K_{d2})[L] + [L]^2}$$

$$f_{B2} = \frac{K_{d1}[L]}{K_{d1}K_{d2} + (K_{d1} + K_{d2})[L] + [L]^2}$$

$$f_{B12} = \frac{[L]^2}{K_{d1}K_{d2} + (K_{d1} + K_{d2})[L] + [L]^2} \tag{3}$$

where K_{d1} and K_{d2} are the dissociation constants for the primary and secondary binding, respectively, and [L] is the concentration of the free form of AD1 or AD2, which is a solution of the following cubic equation:

$$[L]^3 + a[L]^2 + b[L] + c = 0$$

$$a = 2[P]_{tot} - [L]_{tot} + K_{d1} + K_{d2}$$

$$b = ([P]_{tot} - [L]_{tot})(K_{d1} + K_{d2}) + K_{d1}K_{d2}$$

$$c = -K_{d1}K_{d2}[L]_{tot} \tag{4}$$

where $[P]_{tot}$ and $[L]_{tot}$ are the total concentrations of TAZ2 and AD1 (AD2), respectively. The closed-form solution of eq 4 has been reported:[51]

$$[L] = -\frac{a}{3} + \frac{2}{3}\sqrt{(a^2 - 3b)}\cos\frac{\theta}{3} \tag{5}$$

where

$$\theta = \arccos\frac{-2a^3 + 9ab - 27c}{2\sqrt{(a^2 - 3b)^3}}$$

When TAZ2 is ^{15}N-labeled, the observed ^{1}H and ^{15}N chemical shift difference referenced to the free form, $\Delta\delta_{obs}$, is described as

$$\begin{aligned}\Delta\delta_{obs} &= \frac{[B1] + [B12]}{[P]_{tot}}\Delta\delta_{FB1} + \frac{[B2] + [B12]}{[P]_{tot}}\Delta\delta_{FB2} \\ &= (f_{B1} + f_{B12})\Delta\delta_{FB1} + (f_{B2} + f_{B12})\Delta\delta_{FB2} \\ &= \frac{[L]}{K_{d1} + [L]}\Delta\delta_{FB1} + \frac{[L]}{K_{d2} + [L]}\Delta\delta_{FB2}\end{aligned} \tag{6}$$

where [B1], [B2], and [B12] are the concentrations of the B1, B2, and B12 forms, respectively, $\Delta\delta_{FB1}$ and $\Delta\delta_{FB2}$ are the ^{1}H or ^{15}N chemical shift difference between the free form and the B1 or B2 form, respectively. On the other hand, when p53 AD1 is ^{15}N-labeled, $\Delta\delta_{obs}$ is described as

$$\begin{aligned}\Delta\delta_{obs} &= \frac{[B1] + [B12]}{[L]_{tot}}\Delta\delta_{FB1} + \frac{[B2] + [B12]}{[L]_{tot}}\Delta\delta_{FB2} \\ &= \frac{[P]_{tot}}{[L]_{tot}}\{(f_{B1} + f_{B12})\Delta\delta_{FB1} + (f_{B2} + f_{B12})\Delta\delta_{FB2}\} \\ &= \frac{[P]_{tot}}{[L]_{tot}}\left\{\frac{[L]}{K_{d1} + [L]}\Delta\delta_{FB1} + \frac{[L]}{K_{d2} + [L]}\Delta\delta_{FB2}\right\}\end{aligned} \tag{7}$$

For a two-site binding model in which the singly bound states B1 and B2 interconvert (Scheme 2), eqs 3−7 and the following equation hold:

$$\frac{k_{off1}k_{on2}k_{21}}{k_{on1}k_{off2}k_{12}} = \frac{K_{d1}k_{21}}{K_{d2}k_{12}} = 1 \tag{8}$$

dx.doi.org/10.1021/ja209936u | J. Am. Chem. Soc. 2012, 134, 3792−3803

Note that k_{12} and k_{21} cannot be determined by fitting the titration curves because the fitting function is the same for both Schemes 1 and 2. However, the ratio of k_{12} to k_{21} can be obtained from K_{d1} and K_{d2} using eq 8.

Global Analysis. The titration curves monitored using both ^{1}H and ^{15}N chemical shifts referenced to those of the free form, $\Delta\delta_H$ and $\Delta\delta_N$, respectively, were locally or globally fitted with the two-site binding model (Scheme 1). In performing global fits, the dissociation constants are common for all curves, but the chemical shift differences between the free and the bound forms, $\Delta\delta_{FB1}$ and $\Delta\delta_{FB2}$, differ for each curve. The fitting was done with an in-house fitting program *nmrKd* written in Fortran using algorithms taken from Numerical Recipes.[52] Accurate determination of the protein concentrations is crucial for obtaining reliable fits. Because TAZ2 concentration could not be determined accurately by UV absorbance due to the presence of DTT, concentration corrections were implemented in the fitting program. Thus, $[P]_{tot}$ was replaced by $c_p[P]_{tot}$ in the above equations, where c_p is a correction factor for the TAZ2 concentration and is set as a variable during the fitting procedure. The concentration correction led to improved fits. The global fit of $2m$ titration curves for m peaks includes $4m + 2$ adjustable parameters (or $4m + 3$ when concentration correction was performed).

Nonlinear least-squares fits with the Levenberg−Marquardt (LM) algorithm were used to minimize a global X^2:

$$X^2 = \sum_{i=1}^{2m}\sum_{j=1}^{n}\left[\frac{y_{obs}(i,j) - y_{fit}(i,j)}{\sigma(i,j)}\right]^2 \quad (9)$$

where $2m$ is the total number of titration curves, n is the total number of titration points, $y_{obs}(i,j)$ and $y_{fit}(i,j)$ are observed and fitted chemical shift differences at i-th curve and j-th titration point, respectively, and $\sigma(i,j)$ is a standard deviation of each data point, which works as a weight of the data point. Appropriate estimate of σ is necessary for global analysis, because only the curves with large $\Delta\delta_{obs}$ would be optimized if all data points had the same σ. When σ was not known, the σ for each titration curve was estimated by repeating the LM fit as follows: (1) an initial LM fit was performed assuming that all σ values are 1; (2) the rmsd was calculated for each titration curve; (3) the rmsd of the curve was used to determine σ for the data points involved in the curve, and the data were then refitted using the LM method; (4) steps (2) and (3) were repeated until the σ values for all curves converged. This procedure gave a reasonable estimate of σ for all curves.

The Monte Carlo error estimation method was used to estimate the fitted parameters and the corresponding errors.[52] Here, an initial global fit was performed using the actual data set. The rmsd between the actual and fitted curves was calculated for each curve. Then, ∼100 hypothetical data sets were synthesized by using the initially fitted parameters and adding to each data point a random error calculated as (a random number with a normal distribution) × (rmsd of the corresponding curve). Each of the synthetic data sets was globally fit to obtain its fitted parameter set. The mean and standard deviation of these ∼100 fitted parameter sets were used as the final fitted value and the corresponding error.

Peaks that are unassigned, noisy, or invisible because of intermediate exchange were not used in the combined SVD and global analysis. For ^{15}N-TAZ2 titration with AD1, a total of 74 out of 92 residues of TAZ2 were used (148 curves and 17 titration points; a total of 2516 data points). For ^{15}N-TAZ2 titration with AD2, a total of 68 residues of TAZ2 were used (136 curves and 16 titration points; 2176 data points). For ^{15}N-AD1 titration with TAZ2, a total of 21 peaks of AD1 were used (42 curves and 13 titration points; 546 data points).

Fitting Simulations. Fitting simulations were performed with various dissociation constants K_{d1} and K_{d2} by changing the number of data points (an interval of ratios), the maximum titrant/protein ratio, the protein concentration, and the number of peaks used for analysis. The titration curves were generated using the $\Delta\delta_{FB1}$ and $\Delta\delta_{FB2}$ values obtained by titrating AD1 into ^{15}N-labeled TAZ2 (Figure S5), assuming that the K_{d1} and K_{d2} values are a power of 10 between 10^{-10} and 10^{-1} M (see the grid points in Figure 6; $K_{d1} < K_{d2}$). A small amount of noise was added to each data point to simulate actual experimental data, as described above in the Monte Carlo error estimation method. Then, a global fit was performed for each K_d set, and the fitted K_d values were compared with those used in generating the titration curves. Fits were considered to be good if both the fitting errors and the differences between the fitted K_d and the input K_d used for generating the curves were within 10% of the input K_d values.

Line Shape Simulations. The quantum mechanical density matrix formalism was used for the line shape analysis.[53,54] Excited magnetization at each resonance frequency $M(\nu)$ is calculated by

$$M(\nu) = -\mathrm{Re}(\mathbf{I}^-\mathbf{M}_0^{-1}\boldsymbol{\sigma}) \quad (10)$$

where $\mathbf{I}^-$ is a row vector in which all elements are equal to 1. $\boldsymbol{\sigma}$ is a column vector represented by

$$\boldsymbol{\sigma} = C\{f_F, f_{B1}, f_{B2}, f_{B12}\} \quad (11)$$

where C is a constant and f_F, f_{B1}, f_{B2}, and f_{B12} are fractions of F, B1, B2, and B12 forms, respectively, which are calculated from eq 3. For Scheme 2 with the interconversion between B1 and B2, the complex matrix $\mathbf{M}_0$ is

$$\mathbf{M}_0 = \begin{pmatrix} a_1 & k_{off1} & k_{off2} & 0 \\ k_{on1}L & a_2 & k_{21} & k_{off2} \\ k_{on2}L & k_{12} & a_3 & k_{off1} \\ 0 & k_{on2}L & k_{on1}L & a_4 \end{pmatrix} \quad (12)$$

$$\begin{aligned} a_1 &= -k_{on1}L - k_{on2}L - R_{2F} + 2\pi i(\nu_F - \nu) \\ a_2 &= -k_{off1} - k_{on2}L - k_{12} - R_{2B1} + 2\pi i(\nu_{B1} - \nu) \\ a_3 &= -k_{on1}L - k_{off2} - k_{21} - R_{2B2} + 2\pi i(\nu_{B2} - \nu) \\ a_4 &= -k_{off1} - k_{off2} - R_{2B12} + 2\pi i(\nu_{B12} - \nu) \end{aligned} \quad (13)$$

where ν_i and R_{2i} (i = F, B1, B2, and B12) are the chemical shift and the transverse relaxation rate in the absence of chemical exchange, respectively. For Scheme 1 without the interconversion between B1 and B2, $k_{12} = k_{21} = 0$. The R_{2i} values were assumed to be 20 s^{-1} for all forms. ^{1}H chemical shift changes for primary and secondary binding were assumed to be ∼0.1 ppm. Consequently, the ν_i values were assumed to be −50, −10, 10, and 50 Hz for the F, B1, B2, and B12 forms, respectively. Calculations were also performed assuming ^{1}H chemical shift changes of ∼0.4 ppm (−200, −10, 10, and 200 Hz for the F, B1, B2, and B12 forms, respectively). Complex eigenvalues and eigenvectors of the matrix $\mathbf{M}_0$ were calculated using LAPACK. k_{on1}, k_{on2}, k_{off1}, and k_{off2} were set as multiples of 10 between 10^4 and 10^{11} M^{-1} s^{-1} and between 10^{-2} and 10^8 s^{-1}, respectively. k_{12} and k_{21} were assumed to be multiples of 10 between 10^{-2} and 10^8 s^{-1}. Combinations of k_{on1}, k_{on2}, k_{off1}, k_{off2}, k_{12}, and k_{21} that satisfy K_{d1} and K_{d2} of 10^{-10}−10^{-1} M ($K_{d1} < K_{d2}$) were used to calculate line shapes. Protein concentrations were the same as those used for the TAZ2 titration with AD1. A line shape was considered to be that of fast exchange if there is a single peak and its intensity is higher than one-tenth of the peak intensity for the free form at a 1:0 ratio.

Line Shape Fitting. To estimate k_{on} and k_{off}, the ^{1}H line shapes of a series of p53 AD2 titrations into ^{15}N-TAZ2, measured using the 500 MHz spectrometer, were globally fitted with the two-site binding model (Scheme 1). All spectra were processed without apodization but with 8192-point zero-filling to increase the number of data points in the ^{1}H dimension. One-dimensional slices in the ^{1}H dimension were obtained at the top of cross peaks from the NMRView-format files, regions including other peaks were removed from the slices, and a set of ^{1}H line shapes for a series of AD2 titrations into TAZ2 was obtained for each peak.

The fitting was done with an in-house fitting program *glove_lineshape* written in Fortran using algorithms taken from Numerical Recipes.[52] Similar to *nmrKd*, the program uses the LM fitting algorithm and can globally fit line shapes of many peaks. Line shapes

dx.doi.org/10.1021/ja209936u | J. Am. Chem. Soc. 2012, 134, 3792−3803

were calculated using the quantum mechanical density matrix formalism as described above.[53,54] In performing global fits, k_{on1}, k_{on2}, k_{off1}, and k_{off2} were common for all curves, but independent parameters were k_{on1} and k_{on2} only, because K_{d1} and K_{d2} were fixed to the values obtained from the titration curve analysis. Correction for TAZ2 concentration was implemented similar to *nmrKd*, and a correction factor c_p was fixed to the value obtained from the analysis of titration curves. The frequencies in hertz and the R_2 in s^{-1} for the F, B1, B2, and B12 forms differ for each peak. The initial estimates of frequencies were obtained from titration curve analysis. Because the population of B2 is very low, line shape fitting could not accurately estimate the frequency and R_2 for B2. Thus, the frequency of B2 was fixed to the value estimated from the titration curve analysis, and the R_2 of B2 was assumed to be the same as that of B1. In addition, an adjustable parameter for peak intensity was introduced for each line shape, to account for changes in peak intensities accompanied by titrations. The initial estimates of R_2 values and peak intensities were obtained by fitting the line shape of the free form. The global fit for m peaks and n titrations includes $m(n+7)+3$ adjustable parameters. Fitting errors were obtained from a covariance matrix.

Overlapping peaks were not used in the fitting. Moreover, line shapes were not fit well when they had low intensities, distorted shapes, and small peak shifts. Consequently, only the following 12 peaks gave a reasonably good global fit: Val1781, His1782, Cys1791, Gln1797, Arg1801, Val1803, Thr1806, Lys1807, Cys1820, Val1841, Lys1848, and His1849. A total of 192 line shapes and 18 009 data points were used for fitting (279 parameters). The fitted R_2 values for the F, B1 (B2), and B12 forms were, on average, 36, 41, and 49 s^{-1}, respectively.

■ ASSOCIATED CONTENT

Supporting Information

Details on fitting data, spectroscopic data, and simulations of fitting and line shapes. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

wright@scripps.edu

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Jane Dyson for valuable discussions. This work was supported by grant CA96865 from the National Institutes of Health and by the Skaggs Institute for Chemical Biology. M.A. was supported in part by Grants-in-Aid for Scientific Research from MEXT, by Kurata Grant from the Kurata Memorial Hitachi Science and Technology Foundation, by Grant for Basic Science Research Projects from The Sumitomo Foundation, and by Japan Science and Technology Agency, PRESTO. J.C.F. was a Leukemia and Lymphoma Society Special Fellow.

■ REFERENCES

(1) Cavanagh, J.; Fairbrother, W. J.; Palmer, A. G., III; Rance, M.; Skelton, N. J. *Protein NMR Spectroscopy: Principles and Practice*; Elsevier Academic Press: Burlington, MA, 2007.

(2) Fielding, L. *Prog. NMR Spectrosc.* **2007**, *51*, 219.

(3) Henry, E. R.; Hofrichter, J. *Methods Enzymol.* **1992**, *210*, 129.

(4) Jaumot, J.; Vives, M.; Gargallo, R. *Anal. Biochem.* **2004**, *327*, 1.

(5) Beechem, J. M. *Methods Enzymol.* **1992**, *210*, 37.

(6) Chen, L.; Hodgson, K. O.; Doniach, S. *J. Mol. Biol.* **1996**, *261*, 658.

(7) Ionescu, R. M.; Smith, V. F.; O'Neill, J. C. Jr.; Matthews, C. R. *Biochemistry* **2000**, *39*, 9540.

(8) Arai, M.; Iwakura, M. *J. Mol. Biol.* **2005**, *347*, 337.

(9) Jaumot, J.; Marchan, V.; Gargallo, R.; Grandas, A.; Tauler, R. *Anal. Chem.* **2004**, *76*, 7094.

(10) Matsuura, H.; Shimotakahara, S.; Sakuma, C.; Tashiro, M.; Shindo, H.; Mochizuki, K.; Yamagishi, A.; Kojima, M.; Takahashi, K. *Biol. Chem.* **2004**, *385*, 1157.

(11) Sakurai, K.; Goto, Y. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 15346.

(12) Ferreon, J. C.; Lee, C. W.; Arai, M.; Martinez-Yamout, M. A.; Dyson, H. J.; Wright, P. E. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 6591.

(13) Gu, W.; Shi, X. L.; Roeder, R. G. *Nature* **1997**, *387*, 819.

(14) Lill, N. L.; Grossman, S. R.; Ginsberg, D.; DeCaprio, J.; Livingston, D. M. *Nature* **1997**, *387*, 823.

(15) Bochkareva, E.; Kaustov, L.; Ayed, A.; Yi, G. S.; Lu, Y.; Pineda-Lucena, A.; Liao, J. C.; Okorokov, A. L.; Milner, J.; Arrowsmith, C. H.; Bochkarev, A. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 15412.

(16) Di Lello, P.; Jenkins, L. M.; Jones, T. N.; Nguyen, B. D.; Hara, T.; Yamaguchi, H.; Dikeakos, J. D.; Appella, E.; Legault, P.; Omichinski, J. G. *Mol. Cell* **2006**, *22*, 731.

(17) Grzesiek, S.; Bax, A.; Clore, G. M.; Gronenborn, A. M.; Hu, J. S.; Kaufman, J.; Palmer, I.; Stahl, S. J.; Wingfield, P. T. *Nat. Struct. Biol.* **1996**, *3*, 340.

(18) Lee, C. W.; Ferreon, J. C.; Ferreon, A. C.; Arai, M.; Wright, P. E. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 19290.

(19) De Guzman, R. N.; Liu, H. Y.; Martinez-Yamout, M.; Dyson, H. J.; Wright, P. E. *J. Mol. Biol.* **2000**, *303*, 243.

(20) Wojciak, J. M.; Martinez-Yamout, M. A.; Dyson, H. J.; Wright, P. E. *EMBO J.* **2009**, *28*, 948.

(21) Ferreon, J. C.; Martinez-Yamout, M. A.; Dyson, H. J.; Wright, P. E. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 13260.

(22) Feng, H.; Jenkins, L. M.; Durell, S. R.; Hayashi, R.; Mazur, S. J.; Cherry, S.; Tropea, J. E.; Miller, M.; Wlodawer, A.; Appella, E.; Bai, Y. *Structure* **2009**, *17*, 202.

(23) Stone, S. R.; Dennis, S.; Hofsteenge, J. *Biochemistry* **1989**, *28*, 6857.

(24) Wallis, R.; Moore, G. R.; James, R.; Kleanthous, C. *Biochemistry* **1995**, *34*, 13743.

(25) Mei, H.; Wang, K.; McKee, S.; Wang, X.; Waldner, J. L.; Pielak, G. J.; Durham, B.; Millett, F. *Biochemistry* **1996**, *35*, 15800.

(26) Schreiber, G.; Fersht, A. R. *Nat. Struct. Biol.* **1996**, *3*, 427.

(27) Radic, Z.; Kirchhoff, P. D.; Quinn, D. M.; McCammon, J. A.; Taylor, P. *J. Biol. Chem.* **1997**, *272*, 23265.

(28) Sugase, K.; Lansing, J. C.; Dyson, H. J.; Wright, P. E. *J. Am. Chem. Soc.* **2007**, *129*, 13406.

(29) Berg, O. G.; von Hippel, P. H. *Annu. Rev. Biophys. Biophys. Chem.* **1985**, *14*, 131.

(30) Alberty, R. A.; Hammes, G. G. *J. Phys. Chem.* **1958**, *62*, 154.

(31) Sharp, K.; Fine, R.; Honig, B. *Science* **1987**, *236*, 1460.

(32) Marsh, J. A.; Forman-Kay, J. D. *Biophys. J.* **2010**, *98*, 2383.

(33) Muller-Spath, S.; Soranno, A.; Hirschfeld, V.; Hofmann, H.; Ruegger, S.; Reymond, L.; Nettels, D.; Schuler, B. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 14609.

(34) Uversky, V. N. *Protein Sci.* **2002**, *11*, 739.

(35) Levy, Y.; Wolynes, P. G.; Onuchic, J. N. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 511.

(36) Shoemaker, B. A.; Portman, J. J.; Wolynes, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 8868.

(37) Huang, Y.; Liu, Z. *J. Mol. Biol.* **2009**, *393*, 1143.

(38) Eigen, M.; Hammes, G. G. *Adv. Enzymol.* **1963**, *25*, 1.

(39) Fersht, A. R. *Structure and mechanism in protein science: A guide to enzyme catalysis and protein folding*; W. H. Freeman and Co.: New York, NY, 1999.

(40) Sheinerman, F. B.; Norel, R.; Honig, B. *Curr. Opin. Struct. Biol.* **2000**, *10*, 153.

(41) Schreiber, G. *Curr. Opin. Struct. Biol.* **2002**, *12*, 41.

(42) Levy, Y.; Onuchic, J. N.; Wolynes, P. G. *J. Am. Chem. Soc.* **2007**, *129*, 738.

(43) Schreiber, G.; Haran, G.; Zhou, H. X. *Chem. Rev.* **2009**, *109*, 839.

(44) Winter, R. B.; Berg, O. G.; von Hippel, P. H. *Biochemistry* **1981**, *20*, 6961.
(45) Park, C.; Raines, R. T. *J. Am. Chem. Soc.* **2001**, *123*, 11472.
(46) Korennykh, A. V.; Piccirilli, J. A.; Correll, C. C. *Nat. Struct. Mol. Biol.* **2006**, *13*, 436.
(47) Mukherjee, S.; Chowdhury, P.; Bunagan, M. R.; Gai, F. *J. Phys. Chem. B* **2008**, *112*, 9146.
(48) Boehr, D. D.; Nussinov, R.; Wright, P. E. *Nat. Chem. Biol.* **2009**, *5*, 789.
(49) Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A. *J. Biomol. NMR* **1995**, *6*, 277.
(50) Johnson, B. A.; Blevins, R. A. *J. Biomol. NMR* **1994**, *4*, 603.
(51) Wang, Z. X.; Jiang, R. F. *FEBS Lett.* **1996**, *392*, 245.
(52) Press, W. H.; Flannery, B. P.; Teukolsky, S. A.; Vetterling, W. T. *Numerical recipes in Fortran: The art of scientific computing*, 2nd ed.; Cambridge University Press: Cambridge, UK, 1992.
(53) Binsch, G. *J. Am. Chem. Soc.* **1969**, *91*, 1304.
(54) Kern, D.; Kern, G.; Scherer, G.; Fischer, G.; Drakenberg, T. *Biochemistry* **1995**, *34*, 13594.