LETTERS

Mechanism of coupled folding and binding of an intrinsically disordered protein

Kenji Sugase[1,2], H. Jane Dyson[1] & Peter E. Wright[1]

Protein folding and binding are analogous processes, in which the protein 'searches' for favourable intramolecular or intermolecular interactions on a funnelled energy landscape[1,2]. Many eukaryotic proteins are disordered under physiological conditions, and fold into ordered structures only on binding to their cellular targets[3–6]. The mechanism by which folding is coupled to binding is poorly understood, but it has been hypothesized on theoretical grounds that the binding kinetics may be enhanced by a 'fly-casting' effect, where the disordered protein binds weakly and non-specifically to its target and folds as it approaches the cognate binding site[7]. Here we show, using NMR titrations and ^{15}N relaxation dispersion, that the phosphorylated kinase inducible activation domain (pKID) of the transcription factor CREB forms an ensemble of transient encounter complexes on binding to the KIX domain of the CREB binding protein. The encounter complexes are stabilized primarily by non-specific hydrophobic contacts, and evolve by way of an intermediate to the fully bound state without dissociation from KIX. The carboxy-terminal helix of pKID is only partially folded in the intermediate, and becomes stabilized by intermolecular interactions formed in the final bound state. Future applications of our method will provide new understanding of the molecular mechanisms by which intrinsically disordered proteins perform their diverse biological functions.

Following phosphorylation of Ser 133 in KID, CREB stimulates gene expression by association with the general transcriptional coactivator CBP (CREB binding protein) or its paralogue p300 (refs 8, 9). CBP and p300 bind directly to pKID via the KIX domain[10]. Free pKID is intrinsically disordered but folds to form two α-helices, designated α_A and α_B, on binding to KIX (ref. 11). Phosphorylation is required for high-affinity binding, as are hydrophobic contacts between α_B and a shallow groove on KIX (refs 12–14). To elucidate the mechanism by which pKID folds on binding to KIX, we used ^{15}N relaxation dispersion data[15,16] measured as a function of pKID:KIX concentration ratio and [1]H-^{15}N single quantum correlation (HSQC) titrations. These experiments identify weakly populated intermediate states, and provide insights into their structure, their population, and the kinetics of the coupled binding and folding processes.

The ^{1}H-^{15}N HSQC spectrum of free pKID is characteristic of an unstructured protein[11]. On titration of KIX into ^{15}N-labelled pKID (Fig. 1), the intensity of the free pKID cross-peaks diminishes, and new resonances appear and grow in intensity at the chemical shifts of the pKID:KIX complex. Exchange between the free and bound states is slow on the chemical shift timescale, and cross-peaks corresponding to both states are observed at intermediate pKID:KIX concentration ratios. However, several free pKID cross-peaks shift in fast exchange at substoichiometric concentrations of KIX (Fig. 1, top two insets). Many of t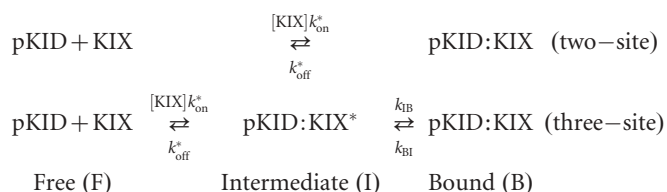he affected resonances do not move directly towards the bound cross-peak, which means that the shifts must arise from a secondary binding process to form a weak complex in fast exchange with free pKID. HSQC titrations using ^{15}N-labelled KIX confirm that pKID binds in slow exchange to the primary high-affinity site and in fast exchange to additional non-specific hydrophobic sites on the surface of KIX, including the binding site for the transactivation domain of the mixed myeloid leukaemia protein (MLL) (Supplementary Figs S2, S3).

Evidence that the fast-exchange shifts of pKID resonances result from transient interactions with KIX comes from competition and mutagenesis experiments. Titration with the MLL activation domain, which forms a ternary complex with pKID and KIX (refs 17, 18), reverses the fast-exchange shifts in the pKID spectrum by blocking non-specific interactions that overlap the MLL binding site (Fig. 1, middle two insets). Some interactions appear to be native-like: substitution of Tyr 658 of KIX by Phe, which disrupts a critical intermolecular hydrogen bond to the phosphoserine in the high-affinity complex[11], specifically decreases the shifts for pSer 133 and Tyr 134 while leaving other pKID cross-peaks unaffected (Fig. 1, bottom two insets). Thus, in addition to binding at the high-affinity site, pKID explores an ensemble of weak complexes at multiple sites on the KIX surface. This behaviour is suggestive of a diffusional encounter complex, a transient state in which the two molecules are held together by fluctuating short-range interactions as they search for the mutually favourable orientations and contacts that stabilize the fully bound state[19,20].

In contrast to encounter complexes formed between globular proteins[21,22], where association is driven by complementary electrostatic interactions, the transient complexes between pKID and KIX appear to be stabilized predominantly by hydrophobic contacts. The pKID residues that experience the largest chemical shift changes are, with the exception of pSer 133, bulky hydrophobic residues (Tyr 134, Ile 137 and Leu 138) that lie on one face of helix α_B and are essential for high-affinity binding[11,14] (Figs 1, 2). The magnitude and pattern of the shifts suggests that the α_B region is partly helical (up to 30%) in the low-affinity complex. Resonances from α_A are shifted only slightly by the fast exchange process.

To obtain further insights into the mechanism of coupled binding and folding, we performed [15]N R_2 dispersion experiments, where R_2 is the transverse relaxation rate. The effective R_2 relaxation rates (R_2^{eff}) depend on the KIX concentration, showing that the observed dispersion is associated with the binding process (Fig. 3a). Dispersion cannot arise from folding transitions in unbound pKID, to form a minor population of a binding-competent conformer, because no relaxation dispersion is observed for free pKID (Supplementary Fig. S4). The dispersion curves at two spectrometer frequencies and at four pKID:KIX concentration ratios were fitted simultaneously with two- and three-site exchange models:

[1]Department of Molecular Biology and Skaggs Institute for Chemical Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037, USA. [2]Suntory Institute for Bioorganic Research, 1-1-1 Wakayamadai, Shimamoto-cho, Mishima-gun, Osaka 618-8503, Japan.

$$\text{pKID} + \text{KIX} \underset{k^*_{\text{off}}}{\overset{[\text{KIX}]k^*_{\text{on}}}{\rightleftarrows}} \text{pKID:KIX} \quad (\text{two–site})$$

$$\text{pKID} + \text{KIX} \underset{k^*_{\text{off}}}{\overset{[\text{KIX}]k^*_{\text{on}}}{\rightleftarrows}} \text{pKID:KIX}^* \underset{k_{\text{BI}}}{\overset{k_{\text{IB}}}{\rightleftarrows}} \text{pKID:KIX} \quad (\text{three–site})$$

$$\text{Free (F)} \qquad\qquad \text{Intermediate (I)} \qquad \text{Bound (B)}$$

The two-site exchange model resulted in physically unrealistic parameters; the effective microscopic dissociation constant K^*_D derived from the fits (41 mM for pSer 133) is inconsistent with the macroscopic K_D (1.3 µM) determined by isothermal titration calorimetry[23]. In contrast, the three-site exchange model fits better (reduced $\chi^2 = 1.12$ versus 1.60) and gives a mean K^*_D (1.5 ± 1.0 µM) that is in excellent agreement with isothermal titration calorimetry (Table 1).

The dispersion curves for residues in the helical regions of bound pKID were analysed in clusters, with the folding and unfolding rate constants k_{IB} and k_{BI} fixed for all residues in each cluster. Three clusters were found (Table 1); residues 124–128 in α_A form a single cluster (I), while there are two clusters in α_B (clusters II and III, comprising residues 133–139 and 140–144, respectively) which appear to interact independently with KIX in both the encounter complex and the intermediate.

We simulated ^{15}N NMR spectra[24] using the parameters of Table 1. Even though the rate of exchange between bound and intermediate states ($k_{\text{ex,IB}}$) is fast ($k_{\text{ex,IB}} > \Delta\omega_{\text{IB}}$, where $\Delta\omega_{\text{IB}}$ is the chemical shift difference between the intermediate and bound states), the overall exchange process between free and fully bound states is slow (Fig. 3b),

as observed experimentally. Because exchange between I and B is fast, the 'bound' shift (given by $\Delta\omega^\star$) is a population weighted average of the chemical shifts in the I and B states. Comparison of the chemical shift difference between the free and intermediate states, $\Delta\omega_{\text{FI}}$, to $\Delta\omega^\star$ provides a measure of the extent of folding of each residue in the intermediate. For residues in cluster I, the average value of $\Delta\omega_{\text{FI}}/\Delta\omega^\star$ is greater than 0.9, suggesting that the α_A helix is nearly fully folded. Although α_A residues (and Arg 131) make nearly native inter-actions in the intermediate, they experience additional millisecond timescale fluctuations that contribute to relaxation dispersion. In contrast, the average $\Delta\omega_{\text{FI}}$ for residues 133–138 and 141 in the α_B helix is only about 70% of the fully bound shift, showing that this helix is incompletely folded (Fig. 3c). Thus, in the intermediate, the critical hydrophobic residues that anchor the pKID α_B helix to the KIX hydrophobic groove continue to search for the favourable inter-molecular contacts that stabilize the final bound state.

To elucidate the mechanism of folding of α_B, we looked for corre-lations between the chemical shifts determined from the dispersion data and equilibrium chemical shift changes measured from HSQC spectra, $\Delta\delta$ (Fig. 3d). A linear correlation is observed between $\Delta\omega^\star$ and $\Delta\delta$. Importantly, $\Delta\omega^\star$ correlates significantly better with the differences in chemical shift between the bound state and the ensemble of weak non-specific complexes, rather than with shift differences relative to free pKID. This strongly suggests that the tran-sient encounter complexes are essential for productive binding, and evolve to the high-affinity complex without dissociation of pKID from KIX.

As α_B is incompletely folded in the intermediate identified from the dispersion experiments, we conclude that this state is a pKID folding
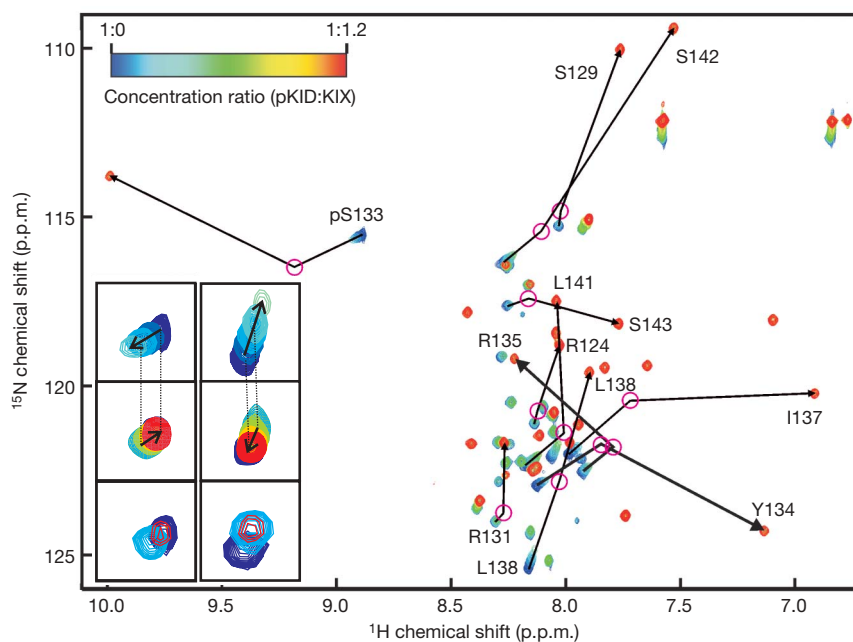

Figure 1 | Chemical shift changes on pKID binding to KIX. ^{1}H-^{15}N HSQC titration of [^{15}N]-pKID with unlabelled KIX in which pKID:KIX concentration ratios ranged from 1:0 to 1:1.2. The cross-peak colour changes gradually from blue (free) to red (bound) according to the concentration ratio. Exchange between the free and fully bound states of pKID is slow on the chemical shift timescale. At substoichiometric concentrations of KIX, several of the pKID cross-peaks shift away from their free positions in fast exchange. The open circles (magenta) represent the fully bound chemical shifts for the fast exchange process. These were estimated using the dissociation constant (680 µM) determined by fitting the shifts observed for KIX resonances in the presence of excess pKID (Supplementary Fig. S2). Selected resonances are labelled. An expanded view of the central part of this figure is shown in Supplementary Fig. S7. Insets show expanded views of the cross-peaks for pSer 133 and Leu 138 at substoichiometric pKID:KIX ratios, as follows. Top row, fast-exchange shift of the resonances of pSer 133 (left) and Leu 138 (right) at pKID:KIX mole ratios of 1:0, 1:0.1, 1:0.2, 1:0.3, 1:0.4. The contours are coloured according to the scale at the top of the figure. Middle row, reversal of fast-exchange shifts for pSer 133 (left) and Leu 138 (right) on titration with MLL. The spectra correspond to pKID:KIX:MLL mole ratios of 1:0:0 (dark blue contours), 1:0.3:0 (cyan), 1:0.3:0.1 (green), 1:0.3:0.2 (yellow) and 1:0.3:0.3 (red). Bottom row, HSQC cross-peaks of pSer 133 (left) and Leu 138 (right) in the presence of wild-type KIX (cyan contours) and the Tyr658Phe mutant (red) at a pKID:KIX mole ratio of 1:0.1. For reference, the cross-peaks for free pKID are shown in dark blue. The scale is slightly expanded relative to the other insets.

intermediate that differs significantly from the encounter complex. Strictly, the coupled folding and binding process should be modelled as four-site exchange (Free↔Encounter↔Intermediate↔Bound). However, exchange of pKID between the free state and the encounter complex is fast and does not contribute to ^{15}N R_2 relaxation; dispersion arises from exchange between the encounter complex, the folding intermediate, and the fully bound state. At steady state, the steps F↔E↔I (where E designates the encounter complex) can be replaced by a single-step process with effective association and dissociation rate constants $k_{on}^* = k_{on}k_{EI}/(k_{off} + k_{EI})$ and $k_{off}^* = k_{off}k_{IE}/(k_{off} + k_{EI})$ (ref. 25).

The values of k_{on}^* show that all residues of pKID except Arg 131 bind to KIX at similar rates (average $k_{on}^* = 6.3 \times 10^6\,M^{-1}\,s^{-1}$) to form the intermediate. The larger k_{on}^* for Arg 131 suggests a possible role for this residue in electrostatic steering[26]. In contrast, the effective rate of dissociation from the intermediate, k_{off}^*, differs significantly for individual residues. Residues in α_A (cluster I; average k_{off}^* is 12.7 s^{-1}) dissociate most slowly, whereas those in α_B have more than fivefold larger k_{off}^* values suggesting that they interact more weakly with KIX in the intermediate. The intermolecular contacts made by α_B residues are stabilized following evolution of the intermediate to the final state. For the fully bound state, the effective microscopic dissociation constants K_D^* and the apparent dissociation rates k_{off} are the same for the residues in each cluster. Thus, the critical interfacial hydrophobic residues in the α_A and α_B helices dissociate from the fully bound state at the same rates (average k_{off} is 6.5 s^{-1} for Leu 128, Tyr 134, Leu 138 and Leu 141). The increased apparent k_{off} rates for Asp 140 and Asp 144 cannot represent physical dissociation of pKID from KIX because neighbouring residues have much smaller k_{off}. Rather, the dispersion data probably contain contributions from additional exchange processes, reflecting either transient local distortions of this region of the α_B helix or fluctuating electrostatic interactions between the Asp side chains and His 602 and Lys 606 of KIX.

Our studies provide insights into the mechanism by which an intrinsically disordered protein folds on binding to its target. Application of our method to additional complexes will be needed to establish whether coupled folding and binding is ubiquitous, or if alternative mechanisms, such as conformational selection (Supplementary Fig. S1), also operate.
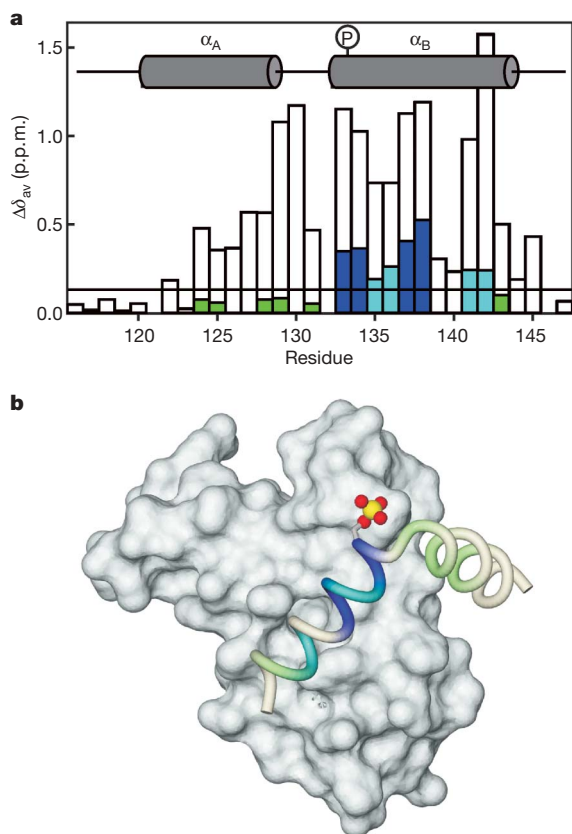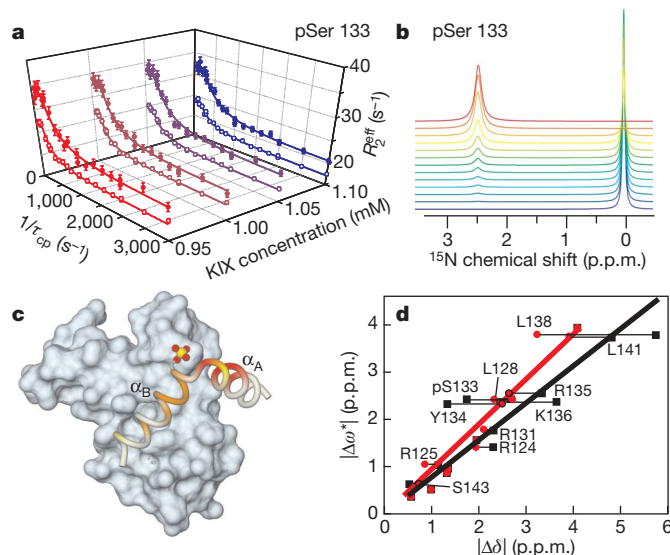

Figure 2 | Characteristics of the encounter complex. a, Chemical shift differences for pKID backbone amide (^1H and ^{15}N) on binding to KIX, calculated according to the formula $\Delta\delta_{av} = \sqrt{((\Delta\delta_{HN})^2 + (\Delta\delta_N/5)^2)}$, where $\Delta\delta_{HN}$ and $\Delta\delta_N$ are the amide proton and nitrogen chemical shift differences, respectively. The horizontal line corresponds to the average chemical shift difference (0.131 p.p.m.) between the free and the fast-exchanging states. Blue bars indicate residues that exhibit chemical shift deviations exceeding the average by more than 1 s.d. (0.289 p.p.m.), and cyan and pale green bars denote residues with intermediate and below average shift differences, respectively. The open bars indicate chemical shift differences between free pKID and fully bound pKID. The locations of the two helices α_A and α_B of pKID and the phosphorylation site are shown at the top of the figure. **b,** Structure of the high-affinity pKID:KIX complex[11], showing a backbone representation of pKID bound to the KIX surface. Chemical shift differences between the free and fast-exchanging states are mapped onto the pKID backbone using the same colour scheme as in **a**. The location of the pSer 133 side chain is shown. The figure was prepared using MOLMOL[29].

Figure 3 | Analysis of R_2 dispersion experiments. a, ^{15}N R_2 relaxation dispersion profile for pSer 133 of pKID recorded at 800 MHz (filled circles) and 500 MHz (open circles). Dispersion curves for 1 mM [^{15}N]-pKID in the presence of 0.95, 1.00, 1.05 and 1.10 mM KIX are shown. **b,** Simulated ^{15}N NMR spectra of pSer 133 at a ^{15}N resonance frequency of 81.1 MHz. The spectra were simulated using the parameters obtained by fitting the R_2 dispersion curves to the three-site exchange model. Chemical shifts in the free form are set to zero. The pKID:KIX concentration ratios are depicted using the same colour scheme as Fig. 1. **c,** Extent of folding in the intermediate state as indicated by $\Delta\omega_{FI}/\Delta\omega^*$ ratios, displayed on the structure of the high-affinity complex. Red, $\Delta\omega_{FI}/\Delta\omega^* > 0.9$; orange, $0.5 < \Delta\omega_{FI}/\Delta\omega^* < 0.9$; yellow, $\Delta\omega_{FI}/\Delta\omega^* < 0.5$. **d,** Correlation of ^{15}N chemical shift differences ($\Delta\omega^*$) determined from the R_2 dispersion measurements with equilibrium shift differences ($\Delta\delta$). Chemical shift differences between free pKID and the fully bound state ($\Delta\delta_{FB}$) are shown as black squares, and between the encounter complex and fully bound state ($\Delta\delta_{EB}$) are shown as red circles, with matching colours for the lines of best fit. The scatter plots show a better correlation between $\Delta\omega^*$ and $\Delta\delta_{EB}$ (slope = 0.95, $R^2 = 0.94$) than between $\Delta\omega^*$ and $\Delta\delta_{FB}$ (slope = 0.79, $R^2 = 0.79$). This difference was shown to be statistically significant using an F-test. The simulated and experimental chemical shift differences were obtained from the simulated ^{15}N NMR spectra (**b**) and from ^1H-^{15}N HSQC spectra (Fig. 1), respectively. $\Delta\delta_{EB}$ was calculated using the estimated fully bound chemical shifts of the encounter complex (magenta circles in Fig. 1). For resonances that undergo no fast-exchange shifts, the corresponding $\Delta\delta_{FB}$ values are used.

Table 1 | R_2 dispersion curve fitting with a three-site exchange model

Cluster	Residue	$\Delta\omega_{FI}$† (p.p.m.)	$\Delta\omega_{FB}$† (p.p.m.)	$k_{on}* \times 10^6$ ($M^{-1}s^{-1}$)	$k_{off}*$ (s^{-1})	Apparent k_{off}§ (s^{-1})	k_{IB} (s^{-1})	k_{BI} (s^{-1})	K_D*‖ (μM)
I	Arg 124	1.33 (0.02)‡	1.71 (0.02)‡	6.5 (0.1)‡	13.2 (0.6)‡	9.3	394 (8)‡	928 (10)‡	1.4
I	Arg 125	1.17 (0.03)	0.84 (0.03)	5.0 (0.2)	15 (1)	10.4			2.1
I	Ile 127	1.47 (0.03)	1.87 (0.03)	7.3 (0.2)	13 (1)	9.0			1.2
I	Leu 128	2.31 (0.03)	2.79 (0.03)	8.1 (0.2)	9.9 (0.7)	6.9			0.9
	Arg 130	2.12 (0.11)	4.01 (0.06)	5.4 (0.3)	155 (21)	5.2	2241 (25)	78 (6)	1.0
	Arg 131	1.77 (0.03)	3.25 (0.02)	15 (1)	4.5 (0.3)	4.4	34 (2)	957 (7)	0.3
II	pSer 133	1.76 (0.04)	2.53 (0.04)	5.5 (0.2)	54 (6)	6.2	1718 (10)	224 (10)	1.1
II	Tyr 134	1.72 (0.05)	2.44 (0.06)	6.4 (0.3)	56 (7)	6.5			1.0
II	Arg 135	1.89 (0.05)	2.66 (0.05)	8.9 (0.3)	65 (6)	7.5			0.9
II	Lys 136	1.66 (0.03)	2.48 (0.03)	7.1 (0.2)	55 (4)	6.4			0.9
II	Leu 138	3.12 (0.05)	3.90 (0.05)	7.3 (0.3)	48 (5)	5.6			0.8
II	Asn 139	1.22 (0.05)	0.89 (0.03)	4.2 (0.3)	137 (14)	15.8			3.8
III	Asp 140	0.6 (0.1)	0.59 (0.01)	5.5 (0.5)	601 (43)	61.6	2852 (16)	326 (18)	11.3
III	Leu 141	2.73 (0.05)	3.88 (0.06)	6.3 (0.3)	69 (6)	7.1			1.1
III	Ser 143	0.92 (0.06)	0.63 (0.03)	4.3 (0.2)	89 (12)	9.1			2.1
III	Asp 144	0.19 (0.06)	0.43 (0.02)	5.9 (0.4)	213 (33)	21.9			3.7
	Ala 145	1.11 (0.01)	0.81 (0.01)	7.62 (0.05)	16.9 (0.6)	10.1	340 (5)	511 (7)	1.3

See text for definitions of quantities in column headings.
† Chemical shift differences $\Delta\omega_{FI}$ and $\Delta\omega_{FB}$ (in rad s^{-1}) were used in curve fitting; values shown here, in units of p.p.m., are calculated according to the formula: $\Delta\omega$(p.p.m.) = $\Delta\omega$(rad s^{-1})/(2πB_O), where B_O is the ^{15}N resonance frequency.
‡ Values in parentheses are standard deviations.
§ Apparent k_{off} are calculated as described in Methods.
‖ Effective dissociation constants are calculated as described in Methods. Excluding the outlier Asp 140, the mean K_D* obtained by fitting the dispersion data with a three-site exchange model is 1.5 ± 1.0 μM.

METHODS SUMMARY

^{15}N R_2 relaxation rates were measured for [^{15}N]-pKID:KIX at 1:0.95, 1:1.0, 1:1.05 and 1:1.10 concentration ratios using relaxation-compensated constant-time Carr-Purcell-Meiboom-Gill pulse sequences[15,27]. Dispersion curves at all four pKID:KIX concentration ratios and at two spectrometer frequencies were fitted simultaneously for each residue using the program GLOVE[28]. Data were fitted to two-site exchange (F↔B) and three-site exchange (F↔I↔B) models. The fitting parameters are a population-average intrinsic relaxation rate R_2^0, effective association and dissociation rate constants $k_{on}*$ and $k_{off}*$, folding and unfolding rate constants k_{IB} and k_{BI}, chemical shift differences between each pair of states $\Delta\omega_{FI}$ and $\Delta\omega_{FB}$, and total protein concentrations [pKID]$_0$ and [KIX]$_0$. The parameters k_{IB}, k_{BI} and $\Delta\omega_{FI}$ apply only to three-site exchange. The rate constant for the binding process (F→B for two-site exchange and F→I for three-site exchange) depends on the concentration of free KIX, which is obtained from the effective dissociation constant K_D*, determined independently for each residue by fitting the dispersion data, and the total concentrations of pKID and KIX:

$$[KIX] = \frac{1}{2}\left\{ -K_D* + a[KIX]_0 - [pKID]_0 + \sqrt{\left(K_D* - a[KIX]_0 + [pKID]_0\right)^2 + 4a[KIX]_0K_D*} \right\},$$

where a is the pKID:KIX concentration ratio ($a = 0.95$, 1, 1.05 or 1.1). The dispersion data at the different concentration ratios are related to each other through this equation. For the two-site exchange model, $K_D* = k_{off}*/k_{on}*$, while for three-site exchange, $K_D* = k_{off}*/k_{on}* \times k_{BI}/(k_{BI} + k_{IB})$, in which $k_{off}*$ is modified by the rate constants k_{BI} and k_{IB} (apparent $k_{off} = k_{off}* \times k_{BI}/(k_{BI} + k_{IB})$). In the three-site exchange model, the intermediate is also a bound state and pKID can dissociate from KIX only from the intermediate state. Therefore, $k_{off}*$ is scaled by the population ratio: $p_I/(p_I + p_B) = k_{BI}/(k_{BI} + k_{IB})$.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

1. Tsai, C. J., Kumar, S., Ma, B. Y. & Nussinov, R. Folding funnels, binding funnels, and protein function. *Protein Sci.* **8**, 1181–1190 (1999).
2. Levy, Y., Cho, S. S., Onuchic, J. N. & Wolynes, P. G. A survey of flexible protein binding mechanisms and their transition states using native topology based energy landscapes. *J. Mol. Biol.* **346**, 1121–1145 (2005).
3. Wright, P. E. & Dyson, H. J. Intrinsically unstructured proteins: Re-assessing the protein structure-function paradigm. *J. Mol. Biol.* **293**, 321–331 (1999).
4. Dyson, H. J. & Wright, P. E. Coupling of folding and binding for unstructured proteins. *Curr. Opin. Struct. Biol.* **12**, 54–60 (2002).
5. Dyson, H. J. & Wright, P. E. Intrinsically unstructured proteins and their functions. *Nature Rev. Mol. Cell Biol.* **6**, 197–208 (2005).
6. Uversky, V. N. Natively unfolded proteins: A point where biology waits for physics. *Protein Sci.* **11**, 739–756 (2002).
7. Shoemaker, B. A., Portman, J. J. & Wolynes, P. G. Speeding molecular recognition by using the folding funnel: The fly-casting mechanism. *Proc. Natl Acad. Sci. USA* **97**, 8868–8873 (2000).
8. Chrivia, J. C. et al. Phosphorylated CREB binds specifically to nuclear protein CBP. *Nature* **365**, 855–859 (1993).
9. Kwok, R. P. S. et al. Nuclear protein CBP is a coactivator for the transcription factor CREB. *Nature* **370**, 223–226 (1994).
10. Parker, D. et al. Phosphorylation of CREB at Ser133 induces complex formation with CPB via a direct mechanism. *Mol. Cell. Biol.* **16**, 694–703 (1996).
11. Radhakrishnan, I. et al. Solution structure of the KIX domain of CBP bound to the transactivation domain of CREB: A model for activator:coactivator interactions. *Cell* **91**, 741–752 (1997).
12. Shih, H. M. et al. A positive genetic selection for disrupting protein-protein interactions: Identification of CREB mutations that prevent association with the coactivator CBP. *Proc. Natl Acad. Sci. USA* **93**, 13896–13901 (1996).
13. Shaywitz, A. J., Dove, S. L., Kornhauser, J. M., Hochschild, A. & Greenberg, M. E. Magnitude of the CREB-dependent transcriptional response is determined by the strength of the interaction between the kinase-inducible domain of CREB and the KIX domain of CREB-binding protein. *Mol. Cell. Biol.* **20**, 9409–9422 (2000).
14. Parker, D. et al. Analysis of an activator:coactivator complex reveals an essential role for secondary structure in transcriptional activation. *Mol. Cell* **2**, 353–359 (1998).
15. Loria, J. P., Rance, M. & Palmer, A. G. A relaxation-compensated Carr-Purcell-Meiboom-Gill sequence for characterizing chemical exchange by NMR spectroscopy. *J. Am. Chem. Soc.* **121**, 2331–2332 (1999).
16. Mulder, F. A., Mittermaier, A., Hon, B., Dahlquist, F. W. & Kay, L. E. Studying excited states of proteins by NMR spectroscopy. *Nature Struct. Biol.* **8**, 932–935 (2001).
17. Goto, N. K., Zor, T., Martinez-Yamout, M., Dyson, H. J. & Wright, P. E. Cooperativity in transcription factor binding to the coactivator CREB-binding protein (CBP). The mixed lineage leukemia protein (MLL) activation domain binds to an allosteric site on the KIX domain. *J. Biol. Chem.* **277**, 43168–43174 (2002).
18. De Guzman, R. N., Goto, N. K., Dyson, H. J. & Wright, P. E. Structural basis for cooperative transcription factor binding to the CBP coactivator. *J. Mol. Biol.* **355**, 1005–1013 (2006).
19. Gabdoulline, R. R. & Wade, R. C. Biomolecular diffusional association. *Curr. Opin. Struct. Biol.* **12**, 204–213 (2002).
20. Berg, O. G. & von Hippel, P. H. Diffusion-controlled macromolecular interactions. *Annu. Rev. Biophys. Biophys. Chem.* **14**, 131–160 (1985).
21. Tang, C., Iwahara, J. & Clore, G. M. Visualization of transient encounter complexes in protein-protein association. *Nature* **444**, 383–386 (2006).
22. Volkov, A. N., Worrall, J. A. R., Holtzmann, E. & Ubbink, M. Solution structure and dynamics of the complex between cytochrome c and cytochrome c peroxidase

determined by paramagnetic NMR. *Proc. Natl Acad. Sci. USA* **103,** 18945–18950 (2006).

23. Zor, T., De Guzman, R. N., Dyson, H. J. & Wright, P. E. Solution structure of the KIX domain of CBP bound to the transactivation domain of c-Myb. *J. Mol. Biol.* **337,** 521–534 (2004).

24. Johnson, C. S. Jr & Moreland, C. G. The calculation of NMR spectra for many-site exchange problems. *J. Chem. Ed.* **50,** 477–483 (1973).

25. Shoup, D. & Szabo, A. Role of diffusion in ligand binding to macromolecules and cell-bound receptors. *Biophys. J.* **40,** 33–39 (1982).

26. Selzer, T. & Schreiber, G. New insights into the mechanism of protein-protein association. *Proteins Struct. Funct. Genet.* **45,** 190–198 (2001).

27. Tollinger, M., Skrynnikov, N. R., Mulder, F. A., Forman-Kay, J. D. & Kay, L. E. Slow dynamics in folded and unfolded states of an SH3 domain. *J. Am. Chem. Soc.* **123,** 11341–11352 (2001).

28. McElheny, D., Schnell, J. R., Lansing, J. C., Dyson, H. J. & Wright, P. E. Defining the role of active-site loop fluctuations in dihydrofolate reductase catalysis. *Proc. Natl Acad. Sci. USA* **102,** 5032–5037 (2005).

29. Koradi, R., Billeter, M. & Wüthrich, K. MOLMOL: A program for display and analysis of macromolecular structures. *J. Mol. Graph.* **14,** 51–55 (1996).

METHODS

Sample preparation. Uniformly ^{15}N-labelled KID domain (residues 116–147) of rat CREB was expressed as a GB1 fusion protein[30] in BL21-DE3 cells in M9 minimal medium; Ser 133 phosphorylation was accomplished in *Escherichia coli* by coexpression with protein kinase A. The GB1 fusion protein was cleaved with Factor Xa and pKID was purified by reverse-phase HPLC. Unlabelled KIX domain (residues 586–672) of mouse CBP was prepared as described[11]. The proteins were dissolved separately in NMR buffer (90% H_2O/10% D_2O, 20 mM Tris-d$_{11}$-acetate-d$_4$ (pH 7.0 at 30 °C), 50 mM NaCl, 2 mM NaN$_3$) and concentrated. Samples of the [^{15}N]-pKID:KIX complex for the R_2 dispersion experiments were prepared from a single concentrated solution of each protein to make the concentration ratios accurate; the pKID concentration was kept at 1.0 mM while KIX concentration was 0.95, 1.0, 1.05 or 1.1 mM. The concentration was determined from the absorbance at 280 nm, using extinction coefficients of 12.95 mM^{-1} cm^{-1} for KIX and 1.49 mM^{-1} cm^{-1} for pKID.

NMR methods. ^{15}N R_2 relaxation rates were measured on Bruker 500 and 800 MHz spectrometers at a temperature of 30 °C using relaxation-compensated constant-time Carr–Purcell–Meiboom–Gill (CPMG) pulse sequences[15,27]. R_2 dispersion spectra were acquired as two-dimensional data sets with a constant relaxation delay of 40 or 60 ms. Four data points, including a reference spectrum acquired with the CPMG blocks omitted, were collected in duplicate and were used to estimate the absolute uncertainties and the signal-to-noise ratio of each spectrum. The K_D (680 µM) for the fast exchanging complex was estimated by fitting the shifts of KIX resonances induced by excess pKID.

Fitting of R_2 dispersion profiles. Dispersion curves were fitted to two-site exchange (F↔B) and three-site exchange (F↔I↔B) models using the in-house program GLOVE[28]. The fitting parameters are a population-average intrinsic relaxation rate R_2^0, effective association and dissociation rate constants k_{on}^* and k_{off}^*, folding and unfolding rate constants k_{IB} and k_{BI}, chemical shift differences between each pair of states $\Delta\omega_{FI}$ and $\Delta\omega_{FB}$, and total protein concentrations $[pKID]_0$ and $[KIX]_0$. The fitting parameters k_{IB}, k_{BI} and $\Delta\omega_{FI}$ apply only to the three-site exchange model. If the kinetic rate constants are much faster than the differences in intrinsic relaxation rate between the states, the intrinsic relaxation rate for each state cannot be derived[31]. The rate constant for the binding process (F→B for the two-site exchange model and F→I for the three-site exchange model) depends on the concentration of free KIX, which is obtained from the effective dissociation constant K_D^*, determined independently for each residue by fitting the dispersion data, and the total concentrations of pKID and KIX:

$$[KIX] = \frac{1}{2}\left\{-K_D^* + a[KIX]_0 - [pKID]_0 + \sqrt{\left(K_D^* - a[KIX]_0 + [pKID]_0\right)^2 + 4a[KIX]_0 K_D^*}\right\} \quad (1)$$

where a is the pKID:KIX concentration ratio ($a = 0.95, 1, 1.05$ or 1.1). For each residue, the dispersion data at the different concentration ratios are related to each other through this equation. For the two-site exchange model, $K_D^* = k_{off}^*/k_{on}^*$, while for three-site exchange, $K_D^* = k_{off}^*/k_{on}^* \times k_{BI}/(k_{BI} + k_{IB})$, in which k_{off}^* is modified by the rate constants k_{BI} and k_{IB} (apparent $k_{off} = k_{off}^* \times k_{BI}/(k_{BI} + k_{IB})$). In the three-site exchange model, the intermediate state is also a bound form and pKID can dissociate from KIX only from the intermediate state. Therefore, k_{off}^* must be scaled by the population ratio: $p_I/(p_I + p_B) = k_{BI}/(k_{BI} + k_{IB})$.

The R_2 dispersion curves at all four pKID:KIX concentration ratios and at the two spectrometer frequencies were fitted simultaneously for each residue. Fits were initially performed for each individual residue and the goodness of fit was assessed by the reduced χ^2 value (χ^2 divided by the degree of freedom, designated $\chi^2_{r,individual}$). However, it is reasonable to assume that the folding and unfolding processes for residues in a local element of secondary structure will be correlated; neighbouring residues within a given structural element are likely to fold and unfold at similar rates. The R_2 dispersion curves of pKID were therefore fitted in clusters; the folding and unfolding rate constants k_{IB} and k_{BI} were treated as global parameters for all residues in the cluster and the goodness of fit was assessed from $\chi^2_{r,cluster}$ values. The two pKID helices were treated separately in the cluster analysis. The initial cluster in each helix consisted of two neighbouring residues; additional residues were added to extend each cluster in subsequent rounds of data fitting. For each cluster, the goodness of fit was assessed by comparing the reduced χ^2 value to that obtained by fitting the constituent residues individually. The fitting quality does not change for a correctly identified cluster, but the reduced χ^2 value decreases with the global parameters because the degree of freedom increases. The cluster analysis was therefore performed by increasing the size of each cluster one residue at a time until the ratio of the reduced χ^2 values ($\chi^2_{r,cluster}/\chi^2_{r,individual}$) increased and became greater than 1. At this point, the added residue was omitted and new clusters were tested. Residues that lie outside the pKID helices were not included in the cluster analysis but their dispersion curves were fitted individually.

Exchange models. For the two-site exchange model, the following analytical equation[32] can be used:

$$R_2^{eff} = R_2^0 + \frac{1}{2}\left\{[KIX]k_{on}^* + k_{off}^* - \frac{1}{\tau_{cp}}\cosh^{-1}\left[D_+\cosh(\eta_+) - D_-\cos(\eta_-)\right]\right\}$$

where
$$D_\pm = \frac{1}{2}\left[\pm 1 + \frac{\Psi + 2\Delta\omega_{FB}^2}{\sqrt{\Psi^2 + \xi^2}}\right],$$

$$\eta_\pm = \tau_{cp}\sqrt{\frac{1}{2}\left(\pm\Psi + \sqrt{\Psi^2 + \xi^2}\right)},$$

$$\Psi = \left([KIX]k_{on}^* + k_{off}^*\right)^2 - \Delta\omega_{FB}^2 + 4[KIX]k_{on}^* k_{off}^*,$$

$$\xi = 2\Delta\omega_{FB}\left([KIX]k_{on}^* - k_{off}^*\right)$$

(2)

where τ_{cp} is the interval between 180° pulses in CPMG. In the case of a three-site exchange model, R_2^{eff} can be calculated according to[33]:

$$R_2^{eff} = R_2^0 - \frac{1}{\tau_{cp}}\ln(\lambda_1) \quad (3)$$

Under the experimentally accessible condition, R_2^{eff} is dominated by the largest eigenvalue λ_1 of the matrix

$$\begin{pmatrix} R\left[\exp\left(-\mathbf{A}\frac{\tau_{cp}}{2}\right)\exp\left(-\mathbf{A}^*\frac{\tau_{cp}}{2}\right)\right] & I\left[\exp\left(-\mathbf{A}\frac{\tau_{cp}}{2}\right)\exp\left(-\mathbf{A}^*\frac{\tau_{cp}}{2}\right)\right] \\ I\left[\exp\left(-\mathbf{A}\frac{\tau_{cp}}{2}\right)\exp\left(-\mathbf{A}^*\frac{\tau_{cp}}{2}\right)\right] & -R\left[\exp\left(-\mathbf{A}\frac{\tau_{cp}}{2}\right)\exp\left(-\mathbf{A}^*\frac{\tau_{cp}}{2}\right)\right] \end{pmatrix} \quad (4)$$

where $R[\ ]$ and $I[\ ]$ are functions to extract the real or imaginary elements, respectively, of the complex matrix. As the matrix $\mathbf{A}$ is a 3-by-3 evolution matrix as given below ($\mathbf{A}^*$ is its complex conjugate) for the three-site exchange model, the matrix size shown as (4) is 6-by-6. If the kinetic rate constants are much faster than differences in intrinsic relaxation rate between the sites,

$$\mathbf{A} = \begin{pmatrix} [KIX]k_{on}^* & -k_{off} & 0 \\ -[KIX]k_{on}^* & k_{off} + k_{IB} - i\Delta\omega_{FI} & -k_{BI} \\ 0 & -k_{IB} & k_{BI} - i\Delta\omega_{FB} \end{pmatrix} \quad (5)$$

Effective dissociation constant for the three-site exchange model. In our model, the intermediate state is one of the bound forms. Therefore, the effective dissociation constant K_D^* is given by

$$K_D^* = \frac{[pKID][KIX]}{[pKID:KIX]_I + [pKID:KIX]_B} \quad (6)$$

where $[pKID]$, $[KIX]$, $[pKID:KIX]_I$ and $[pKID:KIX]_B$ are the concentrations of free pKID, free KIX, complex in the intermediate state, and complex in the fully bound state, respectively. When both numerator and denominator are divided by the total pKID concentration, K_D^* can be expressed by the populations of the free, intermediate and bound states together with $[KIX]$,

$$K_D^* = \frac{p_F[KIX]}{p_I + p_B} \quad (7)$$

All populations can be derived from the kinetic rate constants using the condition of microscopic reversibility,

$$p_F = \frac{k_{off}^* k_{BI}}{k_{off}^* k_{BI} + k_{on}^*[KIX](k_{IB} + k_{BI})}$$

$$p_I = \frac{k_{on}^*[KIX]k_{BI}}{k_{off}^* k_{BI} + k_{on}^*[KIX](k_{IB} + k_{BI})} \quad (8)$$

$$p_B = \frac{k_{on}^*[KIX]k_{IB}}{k_{off}^* k_{BI} + k_{on}^*[KIX](k_{IB} + k_{BI})}$$

Substitution of equation (8) into (7) gives

$$K_D^* = \frac{k_{off}^*}{k_{on}^*} \times \frac{k_{BI}}{k_{BI} + k_{IB}} \quad (9)$$

Calculation of NMR spectra for multiple-site exchange. Excited magnetization at each resonance frequency $M(\omega)$ in the presence of multiple-site exchange can be calculated by[24]:

$$M(\omega) = -\omega M_0 R[\mathbf{P}\cdot(\mathbf{A}^t)^{-1}\cdot\mathbf{1}] \quad (10)$$

where M_0 is the equilibrium magnetization, which only scales the calculated NMR spectrum and therefore an arbitrary value can be used unless noise is taken into account. $R[\ ]$ is the function to extract the real part of the complex number.

$\mathbf{P}$ is a row vector whose elements are the population in each state. $\mathbf{A}$ is the Kubo-Sack matrix[24]. $\mathbf{1}$ is a column vector whose elements are all 1. In the case of our three site-exchange model,

$$\mathbf{P} = \begin{pmatrix} p_F & p_I & p_B \end{pmatrix}$$

$$\mathbf{A}(\omega) = \begin{pmatrix} -R_{2F}^0 - [\text{KIX}]k_{on}^* + i\omega & k_{off}^* & 0 \\ [\text{KIX}]k_{on}^* & -R_{2B}^0 - k_{off}^* - k_{IB} - i(\Delta\omega_{FI} - \omega) & k_{BI} \\ 0 & k_{IB} & -R_{2B}^0 - k_{BI} - i(\Delta\omega_{FB} - \omega) \end{pmatrix} \quad (11)$$

$$\mathbf{1} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

The free KIX concentration at each pKID:KIX concentration ratio used in the dispersion measurements was calculated by equation (1) and populations were then derived using equation (8). The intrinsic relaxation rates in the free form (R_{2F}^0) were experimentally determined by measuring ^{15}N R_2 dispersions with pKID in the free form. The R_2^0 rates obtained with the 1:1 concentration ratio sample were used as R_{2B}^0.

30. Huth, J. R. *et al.* Design of an expression system for detecting folded protein domains and mapping macromolecular interactions by NMR. *Protein Sci.* **6**, 2359–2364 (1997).
31. Grey, M. J., Wang, C. & Palmer, A. G. III. Disulfide bond isomerization in basic pancreatic trypsin inhibitor: Multisite chemical exchange quantified by CPMG relaxation dispersion and chemical shift modeling. *J. Am. Chem. Soc.* **125**, 14324–14335 (2003).
32. Carver, J. P. & Richards, R. E. General 2-site solution for chemical exchange produced dependence of T2 upon Carr-Purcell pulse separation. *J. Magn. Reson.* **6**, 89–105 (1972).
33. Allerhand, A. & Gutowsky, H. S. Spin-echo studies of chemical exchange. II. Closed formulas for two sites. *J. Chem. Phys.* **42**, 1587–1599 (1965).