Binding-Induced Folding of a Natively Unstructured Transcription Factor

Adrian Gustavo Turjanski[1], J. Silvio Gutkind[1], Robert B. Best[2¤], Gerhard Hummer[2]*

1 Oral and Pharyngeal Cancer Branch, National Institute of Dental and Craniofacial Research, National Institutes of Health, Bethesda, Maryland, United States of America, **2** Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland, United States of America

Abstract

Transcription factors are central components of the intracellular regulatory networks that control gene expression. An increasingly recognized phenomenon among human transcription factors is the formation of structure upon target binding. Here, we study the folding and binding of the pKID domain of CREB to the KIX domain of the co-activator CBP. Our simulations of a topology-based Gō-type model predict a coupled folding and binding mechanism, and the existence of partially bound intermediates. From transition-path and Φ-value analyses, we find that the binding transition state resembles the unstructured state in solution, implying that CREB becomes structured only after committing to binding. A change of structure following binding is reminiscent of an induced-fit mechanism and contrasts with models in which binding occurs to pre-structured conformations that exist in the unbound state at equilibrium. Interestingly, increasing the amount of structure in the unbound pKID reduces the rate of binding, suggesting a "fly-casting"-like process. We find that the inclusion of attractive non-native interactions results in the formation of non-specific encounter complexes that enhance the on-rate of binding, but do not significantly change the binding mechanism. Our study helps explain how being unstructured can confer an advantage in protein target recognition. The simulations are in general agreement with the results of a recently reported nuclear magnetic resonance study, and aid in the interpretation of the experimental binding kinetics.

Introduction

A central tenet of genomics is that sequence determines structure, and structure determines function. Even though this is true for many proteins, where structure and function are closely related, a rapidly increasing number of proteins have been found that are natively unfolded or that have long unstructured regions (>50 residues) [1–4]. For many of these proteins, and transcription factors in particular [2], this intrinsic disorder appears to be required for their function. Transcription factors are essential players in signal transduction by regulating gene expression in the nucleus in response to changes in the cellular environment. They are regulated themselves by post-transcriptional modifications, for example phosphorylation [5,6] or ubiquitination. To transmit the correct signal, transcription factors need to associate with many diverse proteins with high specificity and yet relatively low affinity, as they need to dissociate upon signal completion. Having substantial unstructured regions may help in binding to different targets with large surfaces of interaction, and with different conformations, to achieve the required specificity and relative affinity. Moreover, being unstructured may be relevant for the rapid degradation of transcription factors [7].

The transcription factor cAMP response-element binding protein (CREB) represents a well-characterized paradigm for an unstruc-

tured peptide that adopts a folded structure upon binding its co-activator CREB binding protein (CBP) [1,8]. CREB participates in diverse cellular processes, ranging from glucose homeostasis, survival and proliferation to learning and memory [9,10]. CREB is one of the best characterized transcription factors in terms of its structure and function [9–15]. Like many transcription factors, CREB consists of multiple domains (Figure 1A): (i) a C-terminal basic region/leucine zipper (bZIP) dimerization domain that binds DNA with high specificity for the conserved cAMP-responsive element (CRE) TGACGTCA [16,17]; (ii) a 60-amino-acid kinase-inducible domain (KID) (amino acids 100–160); and (iii) two hydrophobic glutamine-rich domains, designated Q1 and Q2, that function as constitutive transcriptional activators [10]. CREB has both basal and inducible transcriptional activity [14], the latter being triggered by phosphorylation of the KID domain at Ser[133] by the protein kinase A (PKA). Indeed, CREB was the first transcription factor whose activity was shown to be tightly regulated by phosphorylation/dephosphorylation [18]. The phosphorylated KID domain (pKID) promotes the recruitment of the transcriptional co-activator CBP (and the related molecule p300) by binding to its kinase-induced domain interacting domain (KIX) [12,19]. Upon binding to KIX, pKID undergoes a transition from unstructured to structured [12].

The nuclear magnetic resonance (NMR) structure of the pKID/KIX complex indicates that only residues 119 to 146 of pKID are

Author Summary

Protein-protein interactions are central to many physiological processes. Traditionally, the atomic structure of the isolated proteins, as obtained from X-ray crystallography or NMR spectroscopy, has been thought to determine the molecular recognition and binding process. However, this view has been challenged by the discovery of natively unstructured proteins, including many human transcription factors, which assume ordered molecular structures only upon binding to their targets. Understanding these transitions from a dissociated and unfolded state to a bound and folded state is key to unraveling how transcription factors, proteins involved in the system that controls the transcription of genetic information from DNA to RNA, perform their function. We conducted molecular simulations to study the binding of a well characterized transcription/co-transcription factor complex. We found that the transcription factor is unstructured when binding to its partner, with folding into an ordered structure occurring only after the initial binding event. The resulting coupled folding-and-binding mechanism is found to be in accord with state-of-the-art experimental data. Transcription-factor binding in an unstructured state may confer an advantage in protein target recognition by accelerating the rate of association without compromising the ability to bind to a diverse set of proteins with high specificity and yet relatively low affinity.

folded [12], forming a two-helix structure with helices α_A and α_B almost perpendicular to one another. The amphipathic α_B helix provides most of the intermolecular contact surface in the complex. Pro[132] is located in the hinge between the two helices.
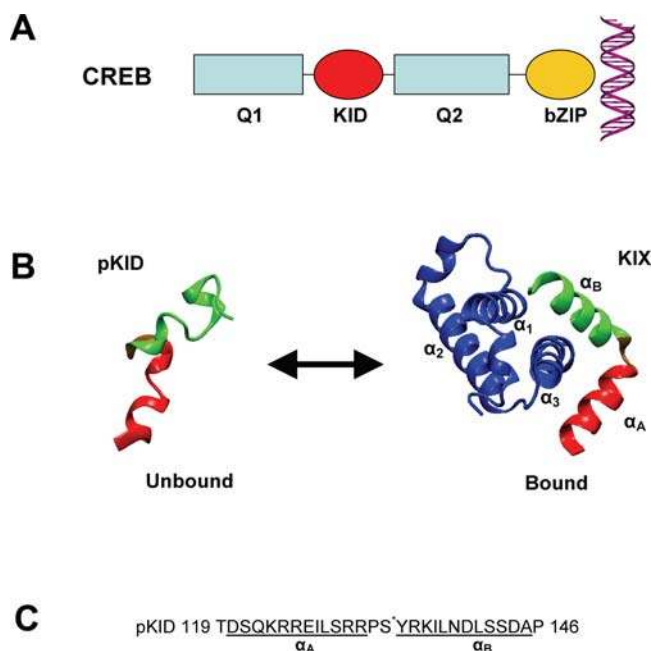
Figure 1. Structure of the pKID/KIX complex. (A) Domain structure of CREB, consisting of DNA-binding domain basic region/leucine zipper (bZIP), the cAMP activated domain KID, and the glutamine-rich domains Q1 and Q2. The sequence of the structured part of pKID is shown with residues in helices α_A and α_B underlined. (B) Folding and binding of pKID (red: helix α_A; green: helix α_B) to the KIX domain of CBP (blue). doi:10.1371/journal.pcbi.1000060.g001

Phospho-Ser[133], positioned at the beginning of the α_B helix, stabilizes the complex by interacting with Tyr[658] and Lys[662] of KIX [12]. NMR experiments have also shown that the unbound pKID domain is mostly unstructured except for the α_A helix which contains a significant fraction of helix (>50%), while the helical content in the α_B helix is very small (10–15%) [13]. The KIX domain, comprising residues 586–672 of CBP, folds into a three-helix structure with helices α_1 and α_3 arranged in parallel, forming a hydrophobic groove that forms the major binding interface for the amphipathic α_B helix of pKID.

Phosphorylation of Ser[133] in KID has been shown to increase the free energy of binding by ~1.5–3.0 kcal/mol [11,20,21], mostly by forming a hydrogen bond with Tyr[658] of KIX. Although a small increase in the pKID secondary structure has been detected upon phosphorylation, it does not seem to have an important role in the binding process [13]. However, it has been suggested that phosphorylation may promote the formation of a transient structural element similar to the bound conformation, which may thus play an extra role in the binding mechanism [22].

As we completed this work, a related NMR study was published on the mechanism of folding and binding of CREB to CBP [23] where it was suggested that pKID becomes structured only after committing to binding. The proposed mechanism involves the formation of transient encounter complexes that evolve to a single intermediate and then to the high-affinity complex [23]. Encounter complexes have been shown to play a key role during protein-protein association for globular proteins for quite some time [24,25] and a structural characterization by NMR has recently been reported [26]. Transient complexes in globular proteins are dominated by long-range electrostatic interactions [24,25,27], however in the pKID/KIX complex Sugase et. al [23] propose a predominant role for non-specific hydrophobic contacts.

Here, we study the folding and binding mechanism of pKID to KIX by simulation of a Gō-type topology-based model based on the native structure of the pKID:KIX complex. We analyze the binding mechanism quantitatively by identifying the key features of binding transition states (TS) using a Bayesian formula relating the equilibrium and transition-path ensembles [28,29]. Our simulations are in general agreement with recent NMR data [23] and also help refine the interpretation of the experimental results. We show that this process occurs via a coupled folding and binding mechanism that typically begins with the binding of an unstructured α_B region in pKID, followed by the folding of α_B into a helix, and finally the docking of the α_A helix. Our simulations show that the TS structures of this coupled binding-and-folding reaction are more closely related to the unstructured conformations of pKID than to the bound conformation. Surprisingly, simulations in which the helical propensity of the unbound pKID was increased resulted in a decrease in the rate of binding in comparison to the unstructured pKID. These findings are in contrast to models that invoke the binding of pre-folded conformations or a TS similar to the complex structure. By calculating Φ values from tertiary contacts we have identified a key residue for the coupled folding-and-binding mechanism, Leu[141], that interacts with the hydrophobic groove of KIX in the TS. Moreover, by adding non-native interactions in addition to those present in the Gō-like model we explored the role of non-specific interactions in the binding process. The simulations of the resulting non-Gō model show an enhanced formation of encounter complexes, in agreement with a recent NMR study [23]. We find that additional sequence-based ("non-native") attractive interactions do not significantly change the binding mechanism, as captured by the transition state ensemble and the pathways taken between unbound and bound states. However, strengthening the

non-native interactions and the resulting formation of non-specific encounter complexes increases the binding rate constant, reflecting the increased interaction cross section, consistent with results for binding of globular proteins and protein folding kinetics [27,30]. Overall, using a simple coarse-grained model we provide a detailed dynamic picture of the folding and binding process of an unstructured protein that is in accord with state-of-the-art experimental data.

Results/Discussion

Coupled Folding and Binding Mechanism

Figure 1A shows a schematic representation of the distinct domains within CREB, including Q1, Q2, BZIP and pKID as described above. Long equilibrium simulations of the complex were initiated from the NMR structure of the complex (PDB code 1KDX [12]) as depicted in Figure 1B. Only residues 119–146 of phosphorylated KID, here referred to as pKID, undergo folding and binding (Figure 1C). To monitor folding and binding of pKID to KIX separately, we calculated the fraction of intramolecular native amino-acid contacts of KIX (Q_{KIX}) and pKID (Q_{KID}), and the fraction of intermolecular native contacts (Q_C) that are formed over the course of the simulation (Figure 2). During most of the simulation time the conformation of KIX remained folded and close to the native structure, held together by an extensive network of inter-residue interactions (as compared to the relatively few interactions between pKID and KIX, which makes it energetically much easier to dissociate the complex than to unfold KIX). However, a minor population of a partially unfolded state was found for the KIX domain (data not shown). This observation is qualitatively consistent with the existence of a partially unfolded high-energy state of KIX that is sparsely populated under native conditions [31]. The conformation of unbound pKID retains part of the intramolecular native contacts, as measured by Q_{KID}, mainly because helix α_A remains partially folded, in agreement with previous experimental results [13]. The increased stability of helix α_A in comparison to helix α_B can be explained energetically, from α_A having more and stronger intramolecular contacts than α_B. The fraction of native intermolecular contacts between pKID and KIX, Q_C, allows us to monitor the exchange at equilibrium
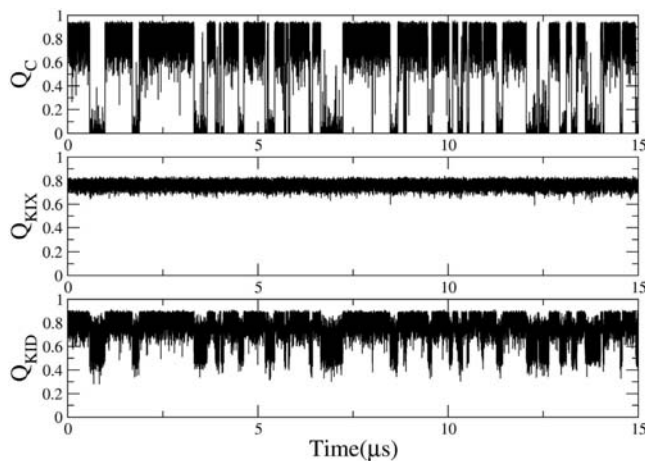
between the two most populated states of the system, the bound ($Q_C \sim 0.9$) and unbound conformations ($Q_C \sim 0.0$). The aggregate simulation data include a total of 889 binding and dissociation transitions, allowing us to characterize in detail the dynamic changes that occur during the folding/binding process.

To characterize the binding mechanism of CREB to CBP, we calculated the free energy profile along Q_C (Figure 3), which shows a free energy barrier to binding of ~ 4 kcal/mol. Notably, the free energy profile also suggests two major populated conformations in the bound state, one with the complex fully formed ($Q_C \sim 0.9$), and a second, "partially bound" intermediate conformation ($Q_C \sim 0.75$). We can gain further insight into the binding mechanism by separating the intermolecular contact fraction, Q_C, into contacts between KIX and helices α_A and α_B of pKID, Q_{CA} and Q_{CB}, respectively. The binding free energy surface as a function of these two coordinates (Figure 4; representative structures superimposed) shows a dominant L-shaped path from unbound ($Q_{CA} \sim Q_{CB} \sim 0$) to fully bound ($Q_{CA} \sim Q_{CB} \sim 0.9$), with the contacts to helix α_B formed before those to α_A. High values of Q_{CB} are observed at values of Q_{CA} as low as 0.0–0.2; in contrast, high values of Q_{CA} are not observed in the absence of high Q_{CB}. Two types of partially bound intermediates are evident in Figure 4: an intermediate I_A with $Q_{CA} < 0.1$ and $Q_{CB} > 0.4$, and an intermediate I_B with $Q_{CA} > 0.1$ and $Q_{CB} < 0.1$ (with a shallow free energy minimum at $Q_{CA} \sim 0.2$ and $Q_{CB} \sim 0$). In the high-population intermediate I_A, α_B is nearly completely bound while α_A mostly detaches from KIX. In the low-population intermediate I_B, α_A is partially bound, while α_B is detached. The local minimum at $Q_{CA} \sim 0.2$ and $Q_{CB} \sim 0.9$ corresponds to structures with transient native interactions formed between helix α_A and KIX, with helix α_B bound and folded.

It is not entirely surprising that binding is dominated by α_B, considering that only a small fraction of the intermolecular contacts are formed between α_A and KIX (5 out of 32, according
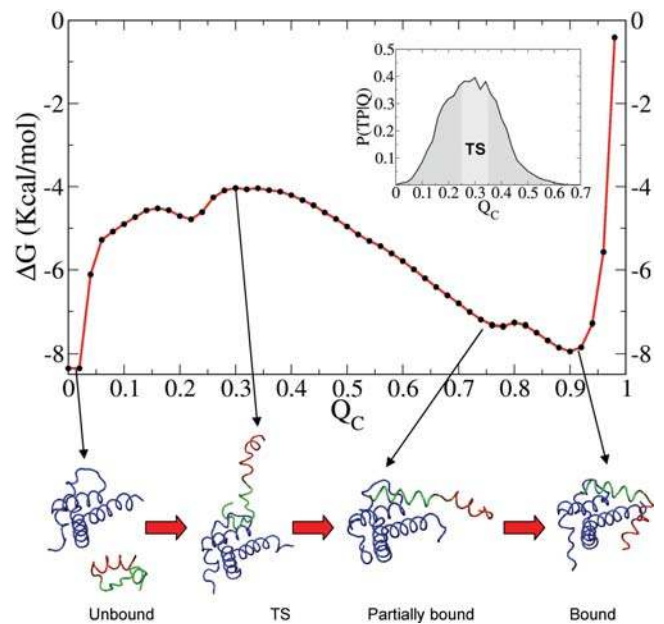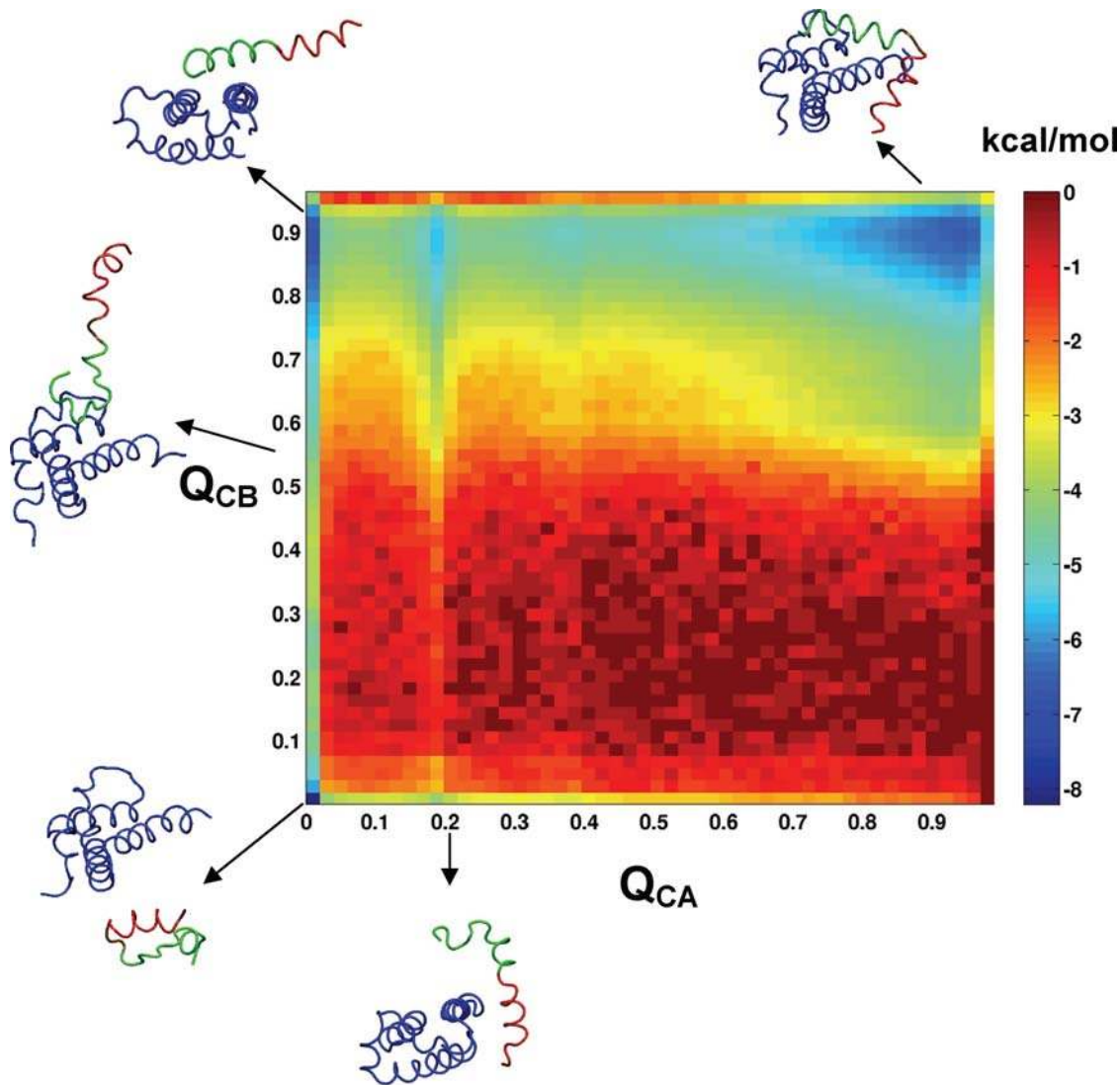

Figure 2. The fraction of native contacts for pKID, KIX, and the complex. Fraction of native amino-acid contacts as a function of time for the intermolecular complex (Q_C; top), KIX (Q_{KIX}; center), and pKID (Q_{KID}; bottom). Data are shown for one out of thirteen simulations of the same length.
doi:10.1371/journal.pcbi.1000060.g002

Figure 3. Free energy surface of the folding and binding process as function of Q_C. Free energy as a function of the fraction of native intermolecular residue-residue contacts Q_C. Representative conformations along the reaction coordinate are shown (blue: KIX, red: pKID-α_A, and green: pKID-α_B). Inset: The probability p(TP|Q_C) of being on a transition path for a given value of Q_C. The TS region is marked in light gray.
doi:10.1371/journal.pcbi.1000060.g003

Figure 4. 2D free energy surface for the binding-induced folding of pKID. Potential of mean force for binding as a function of the fraction of intermolecular native contacts between helix α_A (Q_{CA}) and helix α_B (Q_{CB}) of pKID and KIX. The black line depicts one representative transition path from unbound to bound. Representative structures are shown for important regions of the free energy landscape.
doi:10.1371/journal.pcbi.1000060.g004

to the contact definition in our topological model). The tighter binding between helix α_B and KIX seen here is consistent with previous mutational experiments aimed at destabilizing the interaction between helix α_B and pKID. When Arg[135] and Lys[136] in helix α_B were mutated to Gly the binding affinity was found to decrease [11]. However, no significant change in the binding affinity was found upon mutation of Arg[125] and Glu[126] to Gly in helix α_A, suggesting that formation of a helical structure in the α_A region contributes relatively little to the overall binding affinity [11].

To gain further insight into the binding mechanism we performed additional simulations using non-Gō potentials. In those simulations, we added sequence-based attractive interactions between non-native residue pairs of pKID and KIX to the otherwise unchanged native-topology based interactions. We explored a broad range of non-native interaction strengths, from 0 to 100% relative to the native interactions (whose strength was not varied). We found that for non-native interaction strengths >50%, the folding and binding process was very "glassy",

governed by escape from many deep local minima in the energy surface (data not shown), in contrast to the facile kinetics deduced from experiments [23]. For non-native interaction strengths between 0 to 50% we obtained similar binding free energy surfaces as with the Gō-type model, with only a small reweighting of the minima (e.g., an increase in complexes formed only by helix α_A; Figure S1). As the main effect of adding non-native interactions, we find increased flux on a secondary pathway of the folding and binding transition. This pathway involves the formation of intermediate I_B, from the unbound state to the minimum at $Q_{CA}{\sim}0.2$, $Q_{CB}{\sim}0$, proceeding to the minimum at $Q_{CA}{\sim}0.2$ $Q_{CB}{\sim}0.9$, and then to the bound state (Figure 4 and Figure S1). Whereas the relative contribution of this pathway to the overall reaction is negligible with only native attractions, it is found to have appreciable flux with non-native interactions added, but still much less than the main pathway involving the initial formation of intermediate I_A.

In summary, the following predominant binding mechanism emerges. Before binding helix α_A is mostly formed. Helix α_B, in

contrast, is largely unfolded. Initially, pKID binds to KIX with a largely unstructured α_B region that then folds over the surface of KIX. Once helix α_B is mostly folded and bound (I_A), binding of helix α_A locks the complex. A secondary pathway with much lower flux, involves the binding of helix α_A to the α_3 helix of KIX, with α_B unbound and unstructured (I_B), folding by the binding of the the initially unstructured α_B and its folding. Such independent dissociation of the two pKID helices in the two intermediates also seems reasonable in view of the bound structure (Figure 1B), which shows that the two bound helices lack tertiary interactions with each other, explaining their non-cooperative dissociation.

A recently published study used NMR relaxation-dispersion experiments [23] to infer the formation of a single binding intermediate. Interestingly, the rates of interconversion of the intermediate and the bound state were fitted separately to the relaxation-dispersion curves for certain "clusters" of residues, with marked differences between the rate coefficients for clusters belonging to α_A and those belonging to α_B (see Table 1 in [23]). The rate of conversion from the intermediate state to the fully-bound state was more than four times faster for clusters of residues belonging to α_B than for those belonging to α_A, whereas the rate for conversion from the fully-bound to the intermediate state was faster for residues in α_A than α_B, an unexpected result for a strictly two-state system. Our simulations suggest a simple mechanism which explains this finding: the intermediate state actually consists of two sub-states, one (I_A) with helix α_B folded and bound, and helix α_A unbound and only partially formed; the other (I_B) with α_A bound, and α_B unbound. One can associate the rates from helix α_A with the interconversion of intermediate I_A with the bound state, and those from helix α_B with the interconversion of I_B and the bound state. From the ratio of these experimentally measured rates (Table 1 [23]), it follows that in steady state the population of I_A is about 20 times larger than that of I_B. Indeed, we find in the simulations that the population of I_A is significantly larger than that of I_B (Figure 4 and Figure S1), with a population ratio of ~70, thus differing by a factor of about 3.5 from experiment (or about

$k_BT \ln 3.5 \sim 0.75$ kcal/mol in the free energy). Structurally, based on Figure 5A and 5B, we predict ~80% intramolecular helical contacts for residues on $á_A$, and ~70% for $á_B$. This is in reasonable agreement with the experimental estimates of ~90% folded for $á_A$ and ~70% folded for $á_B$, inferred from NMR chemical shift differences [23]. (Note that the measured chemical shift differences of the different residues report on the structure in the respective intermediates: I_A for residues in α_A, and I_B in α_B, not the population-weighted average.) Overall, we find nearly quantitative agreement with experiment, with respect to both the equilibrium populations of the two intermediates and their structural characteristics. Thus our work is a true prediction of binding mechanism from simulation.

Transition-State Ensemble

To quantify the structural properties of the binding TS, we use the fraction of intermolecular contacts, Q_C. To show that Q_C properly identifies the TS ensemble of the coupled folding-and-binding transition, we apply a Bayesian formalism [28,29] and compute the conditional probability $p(TP|Q_C)$ of being on a transition path (TP) given a particular fraction of intermolecular contacts, Q_C. Transition paths are defined as trajectory segments that connect unbound conformations ($Q_C<0.02$) with bound or partially bound conformations ($Q_C>0.7$), and vice versa, without re-crossings. The value of Q_C with the highest $p(TP|Q_C)$ is most indicative of being on a transition path, and is used to identify the most reactive states, or "transition states". The inset in Figure 3 shows that this largest value of $p(TP|Q_C)$ is obtained for $Q_C\sim0.3$ for which $p(TP| Q_C)\sim0.4$. With $p(TP|Q_C)$ being close to the theoretical maximum of 0.5 for a perfect reaction coordinate of a diffusive process, we conclude that Q_C is a relevant coordinate to describe the coupled folding-and-binding transition. We note, that for inertia-dominated processes, the theoretical maximum (in phase space) would be one; however, diffusion is expected to be relevant here for a process in condensed phase that involves the motions of large groups of atoms and solvent. We note further that, with Q_C being a good reaction coordinate, as deduced from the high value of $p(TP|Q_C)$ at the transition state, and with dynamic corrections being small, the maximum of the free energy along Q_C coincides nicely with the maximum of $p(TP|Q_C)$, leading to equivalent transition-state ensembles.

We selected all conformations from transition paths with Q_C values in the interval [0.25;0.35] as representatives of the TS ensemble. To characterize the TS structures and relate them to the bound and unbound states, we compared the fractions of native intramolecular contacts of helices α_A and α_B of pKID ($Q_{\alpha A}$, $Q_{\alpha B}$) in the bound ($Q_C>0.80$) and unbound conformations ($Q_C<0.02$), the partially-bound intermediates I_A and I_B, and the TS ensemble (Figure 5A and 5B). In the unbound state helix α_A has a smaller fraction of native contacts (maximum at $Q_{\alpha A}\sim0.65$) than in the bound state (maximum at $Q_{\alpha A}\sim0.95$), consistent with a partial unfolding of helix α_A in the I_A intermediate. Interestingly, our simulations show that at the TS the structures resemble the unbound conformation, with helix α_A partially folded, and helix α_B mostly unfolded. At the TS, the distributions of the intramolecular contact fractions $Q_{\alpha A}$ and $Q_{\alpha B}$ are nearly identical to the respective curves for the unbound state (Figure 5A and 5B), and substantially different from the bound ones.

We also calculated the fractions of intermolecular contacts made by each helix (Q_{CA}, Q_{CB}) in the TS and bound states (Figure 5C and 5D). The distributions of Q_{CA} and Q_{CB} indicate that at the TS helix α_B is partially bound to KIX, with α_A mostly dissociated. A representative conformation of the TS structure is shown in Figure 3, in which an unstructured helix α_B is bound to
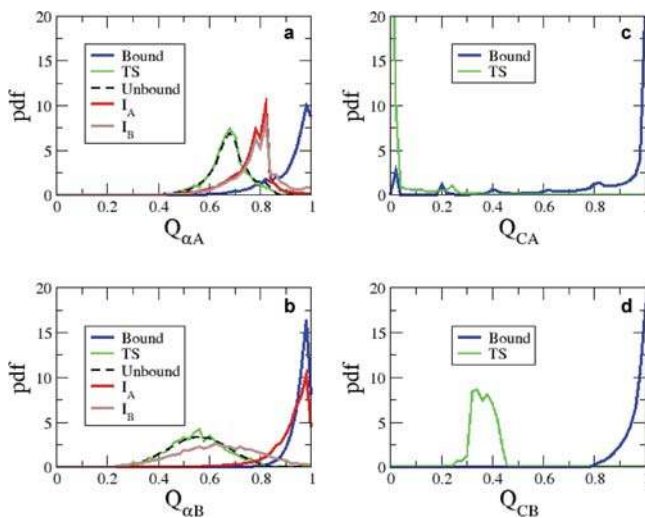
Figure 5. Structural characteristics of the transition state ensemble. (A) and (B) show the probability density function (pdf) of intramolecular native contacts of helix α_A ($Q_{\alpha A}$) and helix α_B ($Q_{\alpha B}$) of pKID for the unbound state, TS, intermediates I_A and I_B, and bound states. (C) and (D) show the probability density functions of native contact fractions Q_{CA} and Q_{CB} between KIX and helices α_A and α_B of pKID, respectively, in the bound and transition states.
doi:10.1371/journal.pcbi.1000060.g005

the hydrophobic groove formed between helices α_1 and α_3 on KIX. However, we emphasize that, because of the unstructured nature of pKID in the TS, the TS is best described by a broadly distributed ensemble of conformations in which pKID binds to KIX in different conformations.

Binding Kinetics of Structured and Unstructured pKID

Our results suggest that binding occurs with the α_B helix being largely unstructured. Based on this observation, one would expect little or no effect of α_B stability on the kinetics of association, as measured by the on-rate (k_{on}). We investigated the effect of increased helical structure in the unbound state by conducting simulations in which we strengthened the intramolecular contacts of pKID (Q_{KID}) to form pre-folded unbound peptides. By starting from different unbound configurations $(Q_C \sim 0.0, Q_{KID} \sim 1)$ randomly chosen from an equilibrium simulation with the original energy function, we performed 350 binding simulations with enhanced intra-helix interactions. The on-rate for the resulting pre-structured helix, $k_{on}(Q_{KID} \sim 1)$, estimated from mean first passage times, is in fact lower than the on rate for the original simulations by a factor ~ 1.6; remarkably, this value is in the same range as theoretical predictions for a "fly-casting" mechanism [32]. In previous experimental studies, the role of secondary structure formation in the folding/binding process was explored through extensive mutations in the α_B helix [11,33]. As stated above, mutations that destabilize the α_B helix of pKID decrease the binding affinity; however, mutations intended to increase the α_B helix propensity do not enhance binding towards KIX. In some cases, even a reduction in binding has been observed. These results can be rationalized based on our simulations. With the observed unstructured TS of the folding/binding reaction, destabilization of the α_B helix should increase the rate of dissociation, k_{off}, with not significant changes in k_{on}; in contrast, stabilization of α_B should produce at least partially compensating effects, diminishing k_{on}, by reducing the "fly-casting" effect, and k_{off}, by stabilizing the bound state, and therefore resulting in only small changes of the overall binding affinity, as observed experimentally [11,33]. To explore further the role of the α_B helix, measurements of the effects of mutations in pKID on the kinetics of association (k_{on}) and dissociation (k_{off}) might be insightful.

Φ Values for pKID-KIX Binding

In analogy to protein folding [34,35] we further characterized the TS by calculating Φ values for each amino acid residue of pKID that is involved in native contacts with KIX. We estimated Φ values by using a simple definition in which only native contacts are counted, $\Phi = Q_{xc}(TS)/Q_{xc}(bound)$ [34,36,37], where Q_{xc} is the number of native contacts in the respective state between protein 1 and residue x of protein 2. Φ values have been extensively used for determining key residues in the folding of proteins and remarkable agreement has been found between theoretical and experimental studies [34,38,39]. Φ-value analysis should quantify the amount of intermolecular structure formed at each site in the transition state. The calculated Φ values for the coupled folding-and-binding transition of pKID are shown in Figure 6. According to our analysis, only residues of helix α_B seem to be important for TS formation, with the highest Φ values being found for Leu[141] and the neighboring amino acids Asp[140] and Ser[142], implying that hydrophobic interactions involving amino acids at the center of the α_B region govern the association mechanism. A critical role for Leu[141], which is deeply bound into the hydrophobic groove of KIX, has been suggested based on the results of mutating Tyr[650] of CBP which is located in the wall of the hydrophobic groove [11]. Mutation to Ala resulted in a ~ 12-fold decrease in the
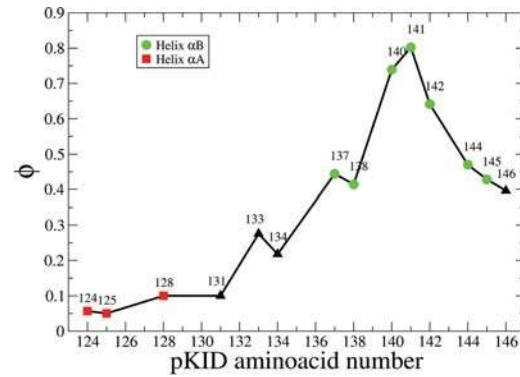
Figure 6. Φ value analysis for the transition state. Φ values for the folding/binding transition state for the residues of pKID forming native contacts with KIX. Red squares indicate amino acids belonging to helix α_A, green circles indicate amino acids belonging to helix α_B, and black triangles show Φ values for residues without secondary structure. doi:10.1371/journal.pcbi.1000060.g006

binding constant [11]. Our study suggests that corresponding kinetic measurements on this mutant should be insightful.

Effect of Non-Native Interactions

To clarify the role of non-native binding, as stated above, we performed additional simulations using non-Gō potentials. As discussed, the inclusion of non-native contacts at reasonable strengths does not substantially affect the binding mechanism (Figure S1). However, the additional non-native interactions result in the formation of non-specific encounter complexes. To quantify their population, we determined the binding free energy surface as functions of NNC and NC, the numbers of non-native and native contacts between pKID and KIX, respectively. The results are shown in Figure 7A and 7B for non-native interaction strengths of 0% and 40%, respectively, and in Figure S2 for other interaction strengths.

To explore the role of the non-native interactions in the binding mechanism we calculated the binding free energy surface as a function of NNC and NC selectively for the ensemble of transition paths that connect unbound to bound conformations. Figure 7C and 7D show that substantial non-native interactions are formed during the transition paths with 40% non-native interaction strength, resulting in a free energy minimum for NC~6 and NNC<25. However, the presence of a minimum does not in itself show whether the non-native interactions are essential elements of the transition state for binding.

We quantify the role of NNC in the binding transition state by calculating the probability p(TP|NC,NNC) of being on a transition path for given values of NC and NNC. Figure 7E and 7F show that the TS for binding, determined by the vertical ridge with p(TP|NC,NNC)~0.5 (which is the theoretical optimum), is at NC~10 independent of NNC. Therefore, even though the TS configurations typically have a substantial number of non-native interactions, the binding mechanism is not substantially changed upon the increase of non-specific binding. With the p(TP|NC,NNC)~0.5 ridge leaning slightly to the right, a larger number of NNC's at the TS appears to require a larger number of NC's as well.

To characterize the nature of the on-pathway non-native contacts, reflected in the local minimum in Figure 7D (NC~6, NNC<25), we calculated the fractions of intermolecular contacts made by each helix (Q_{CA}, Q_{CB}) and the fraction of intramolecular contacts for each helix $(Q_{\alpha A}, Q_{\alpha B})$ in this minimum (Figure S3).
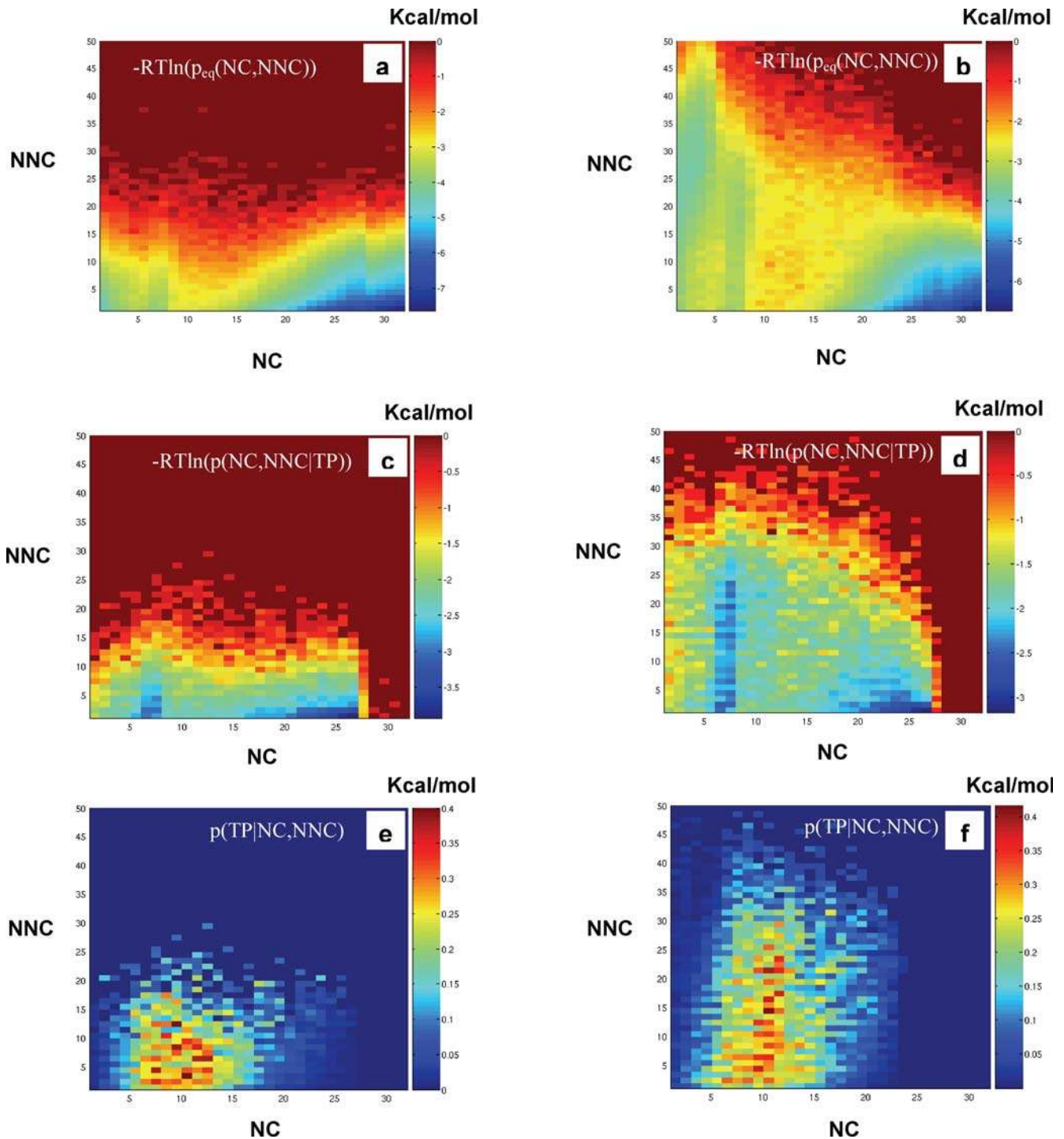
Figure 7. Non-specific encounter complexes. Potential of mean force as function of the number of native and non-native contacts between pKID and KIX. Free energy surfaces are shown as a function of all non-native contacts (NNC) and native contacts (NC) (A–B) at equilibrium and (C–D) in transition paths. (E–F) The probability p(TP|NC,NNC) of being in the transition path as a function of NC and NNC. (A), (C), and (E) correspond to a potential with only native attractive interactions and (B), (D), and (F) correspond to the same potential with added non-native interactions at a strength of 40% of the native contacts.
doi:10.1371/journal.pcbi.1000060.g007

The distributions of Q_{CA} and Q_{CB} indicate that in the *on-pathway* complexes α_B is partially bound to KIX, with α_A mostly dissociated. These conformations have an unstructured α_B region that is bound to the hydrophobic groove formed between helices α_1 and α_3 of KIX, resembling the binding TS ensemble (Figure 5), but with slightly fewer contacts between α_B and KIX (Q_{CB}).

We studied also the non-specific binding to the two major solvent exposed surfaces of KIX remote from the native interface (SF1 and SF2 as shown in Figure S4). SF2 includes most of the region previously reported to bind the transactivation domain of the mixed myeloid leukaemia protein (MLL) [40,41] and recently reported to be involved in transient encounter complexes with

pKID [23]. The free energy surfaces as function of the number of non-native contacts between SF1 or SF2 with pKID (NNC$_{SF1}$ and NC$_{SF2}$ respectively) for equilibrium simulations and transition paths are shown in Figure S4. Interestingly, for the MLL binding site SF2 we observe a considerable number of interactions with pKID, in agreement with the NMR results [23]; moreover these transient complexes also have a small number of native interactions, as has been deduced from experiment [23]. While being partially bound to the non-native solvent exposed surfaces of KIX, pKID is thus found to explore other surfaces, including the native interface. This observation is not entirely surprising because pKID is unstructured and KIX is a small protein. However, according to our simulations, SF1 and SF2 do not seem to be involved in the binding pathway towards the native conformation. In the respective free energy surfaces for the transition path ensemble (Figure S4B and S4D), there are no non-native interactions being formed on-pathway with either SF1 and SF2. How can this be reconciled with the NMR results reported by Sugase et al. [23]? The NMR evidence for the location of the encounter-complex binding comes from measurements at 300 μM pKID (MLL titration in Figure 1 of [23]) and ~1 mM pKID (pKID titration in Figure S2 of [23]). Those concentrations are much higher than the $K_d = 1.3$ μM [23] of the specific pKID:KIX complex, and it seems likely that under those conditions two pKID molecules could be bound to KIX. The second molecule may form a low affinity complex at the MLL binding site, which has been shown to bind another transcription factor and has high propensity to form non-native contacts with pKID (Figure S4C) [40].

Encounter complexes are thought to be important factors in the folding and binding kinetics [24,30]. To quantify this effect, we compared the on rate constants (k_{on}), estimated from mean first passage times, obtained from simulations with non-native interactions strengths of 0% and 40%. The observed ~4-fold increase in k_{on} could be explained in terms of an enlargement of the interaction surface, as has been previously shown for globular proteins [27]. Increasing the non-native interaction strength further will eventually produce a slowdown in the rate of specific binding, with non-specific encounter complexes acting as kinetic traps [30].

Induced-Fit Binding

Molecular complexes in which a "ligand" induces large changes in a "receptor" upon binding have conventionally been described by the "induced-fit" model [42]. Recently, a new view has challenged this simple picture, in which the ligand selectively binds and stabilizes a particular conformation of the receptor that is present at an appreciable population in the equilibrium unbound state [43,44]. In this "pre-equilibrium" picture, the bound conformation of the "receptor" is formed prior to binding, and "ligand binding" just causes a population shift. Here, it happens to be the "ligand" pKID that undergoes a structural change upon binding, but the above concepts are nevertheless relevant. Our analysis of the transition states shows that structure is formed only after binding, which makes pKID binding to KIX more consistent with an induced-fit interpretation, with the interesting difference that here it is the ligand that changes conformation.

It is easy to see why a population shift mechanism is disfavored in the pKID/KIX transcription factor binding. Stabilization of the folded and bound conformation of pKID comes primarily from the large binding interface with KIX, rather than the relatively weak intramolecular interactions, making the folded conformation, as seen in the bound state, extremely improbable in the unbound state. Although both the induced fit and population shift models originally considered transitions between different conformations of already-folded proteins, the unbound conformations of

folded proteins are frequently highly flexible and could be considered analogous to the unfolded peptide here. We stress that both mechanisms are possible in general: the distinction between them in a particular case would depend on the population of the bound conformation in the unbound state.

Conclusions

The phosphorylated KID domain (pKID) of the transcription factor CREB is known to be unstructured when it is not bound to a target, suggesting that folding and binding are concomitant [11–13,43,44]. The results of our simulations suggest that binding indeed occurs together with folding. However, we show in addition that folding into a two-helix structure occurs mostly after binding, with the transition state structure of pKID closely resembling the unfolded state.

This mechanism is different from models of protein complex formation in which structure is required for recognition prior to binding. Why would nature adopt such a model for molecular recognition? As mentioned above, there are many advantages to being unstructured, including the capacity to bind multiple targets and rapid termination of the signal *via* degradation of the peptide; since high affinity binding is not required, or even desirable for fast response, pre-existent structure may not be needed. This is also consistent with our finding that the bound complex is relatively loose and includes multiple bound structures.

Nonetheless, despite the acceleration of signal termination, one might imagine that the penalty for being unstructured would be a reduction in the rate of binding, and signal initiation. In fact, we find the opposite to be true: increasing the amount of structure in the unbound pKID slightly reduces the on rate. This slow-down can be understood in the context of the so called "fly-casting" mechanism [32], in which the increased "capture radius" of an unfolded polypeptide enhances the rate of target binding relative to that of compact folded proteins. While the acceleration that we observe, a factor of ~1.6, would probably not provide a significant advantage biologically, it demonstrates that being unstructured is certainly not a disadvantage from the point of view of binding rate.

A recently published experimental study broadly confirms our simulation results [23]. The main conclusion, of both the experimental study and our simulation study, is that pKID becomes structured only after committing to binding. By performing ^{15}N relaxation-dispersion experiments, Sugase et al. [23] inferred that the coupled folding and binding transition involves the formation of an intermediate. Remarkably, this result is also in agreement with our observation of partially bound intermediates states as shown in Figures 3 and 4; we are also able to rationalize the differences in fitted rate coefficients between the two helices based on our simulations.

Non specific transient encounter complexes have been proposed to play an important role in binding of globular proteins [24–26,45]. Kinetic data on a number of protein–protein associations have provided evidence for the initial formation of pre-equilibrium encounter complexes that subsequently relax to the final stereospecific complex [46]. Recently, Sugase et al. [23] found evidence for non-specific transient encounter complexes for the unstructured pKID upon binding to KIX, suggested to be on-pathway based on structural similarities inferred from chemical-shift correlations. Here, we found both on and off-pathway encounter complexes with non-specific interactions. However, we could show that the binding transition state, and thus the binding mechanism, was largely determined by the formation of specific native-like interactions. Nevertheless, non-specific interactions were found to enhance the on-rate for binding, consistent with the results for binding of globular proteins and protein folding [27,30,47].

In conclusion, our results support a model for binding of pKID to KIX in which the structure of pKID is induced by binding, conferring several advantages. The importance of unstructured or flexible regions in proteins is increasingly recognized on the basis of both experimental [1,4,48–50] and theoretical studies [3,32,51–57]. Since many proteins involved in cell signaling, including transcription factors other than CREB [15,58–63], also have unstructured regions that undergo a transition from unstructured to structured, the mechanism observed here may thus be of more general significance in their target recognition.

Methods

Model

The folding and binding of the pKID/KIX complex was studied by using a coarse-grained, native-topology based protein model. The Gō-type model was constructed from the structure of the complex, as determined by NMR (PDB code 1KDX [12]), following a standard procedure [64]. Only residues shown by the NMR data to be structured were used in building the models. The resulting energy surface consists of a transferable dihedral potential, a harmonic potential for angles with a minimum at the native angle, an amino-acid-type dependent (Miyazawa-Jernigan) attractive term [65] for interactions between pairs of residues which interact in the experimental structure (native contacts) and a repulsive potential for all other residue pairs. Each residue is represented by a single bead centered on its alpha-carbon (C_α) position. To account for the charge of the phosphoserine, amino-acid contact potentials for a glutamic acid were used instead. We scaled the nonbonded interactions in order to have similar populations of bound and unbound conformations. In simulations in the absence of KIX, the α_A helix of isolated pKID is only partially formed, and the α_B region is mostly unstructured, consistent with previous NMR observations [13].

To sample possible non-native complexes, we generated a model in which attractive potentials were added to all non-native residue pair interactions between pKID and KIX. We used a 12-6 potential with a minimum radius $\sigma = 5.5$ Å, and a well depth weighted by the same Miyazawa-Jernigan contact energies as in the Gō-type model, but scaled to 20%, 30%, 40% or 50% of the native value. The interactions of native residue pairs were left unchanged.

Simulations

A 1:1 complex of pKID/KIX was simulated in a 60 Å cubic box with periodic boundary conditions. Non-bonded interactions were smoothly cut off at 25 Å. The Langevin dynamics simulations were run in CHARMM [66–68] with a time step of 15 fs. All bonds were fixed at their native lengths using SHAKE [69]. The actual total simulation time is 195 μs (1.3×10^{10} steps). The chosen friction coefficient of 0.1 ps^{-1} is smaller than that used to mimic water (50–100 ps^{-1}) to increase the rate of transitions in the equilibrium simulations [70]. Although these simulations are not in the high friction limit, we can estimate, as a rough approximation, an upper bound for the times in water by scaling simulation times by a factor of 1000 [71], bringing the effective total simulation time to ~200 ms. We note further that to enhance the sampling of binding events, the binding constant is smaller than the experimental value; however, we find that increasing the affinity by an order of magnitude through enhanced intermolecular interactions does not affect the results. Simulations with the modified Gō model that includes non-native interactions were run in the same conditions as with the native model. In this case we ran 1.0×10^9 steps, equivalent to a simulation time of 15 μs, for each of the different potentials.

Supporting Information

Figure S1 2D free energy surface for different non-native interactions strength. Potential of mean force for binding as a function of the fraction of intermolecular native contacts Q_{CA} and Q_{CB} for potentials with non-native interaction strengths of 20%, 30%, 40%, and 50% of the native ones.
Found at: doi:10.1371/journal.pcbi.1000060.s001 (1.49 MB DOC)

Figure S2 Quantifying the role of non-native interactions. Potential of mean force for binding as a function of the number of intermolecular native contacts (NC) and non-native contacts (NNC) for potentials with non-native interaction strengths of 0%, 20%, 20%, and 50% of the native ones.
Found at: doi:10.1371/journal.pcbi.1000060.s002 (1.15 MB TIF)

Figure S3 On-pathway non-specific binding. (A) The free energy surface as a function of all non-native contacts (NNC) and native contacts (NC) between pKID and KIX in transition path for 40% non-native interaction strength relative to native interactions where the minimum for on-pathway transient complexes is marked. (B) The probability density function of intramolecular native contacts of helix αA ($Q_{\alpha A}$) and helix αB ($Q_{\alpha B}$) of pKID and of native contact fractions Q_{CA} and Q_{CB} between KIX and helices αA and αB of pKID for the selected minimum.
Found at: doi:10.1371/journal.pcbi.1000060.s003 (0.89 MB TIF)

Figure S4 Role of KIX surfaces not involved in the specific pKID/KIX complex. Free energy surfaces are shown as a function of: non-native contacts of pKID with surface 1 of KIX (NNC_{SF1}) during (A) equilibrium simulation and (B) transition paths respectively; non-native contacts of pKID with surface 2 of KIX (NNC_{SF2}) during (C) equilibrium simulation and (D) transition paths respectively. Regions SF1 and SF2 are illustrated on the structures to the right of (B) and (D), respectively. The potential has 40% non-native interaction strength relative to native interactions.
Found at: doi:10.1371/journal.pcbi.1000060.s004 (1.50 MB TIF)

Author Contributions

Conceived and designed the experiments: AT RB JG GH. Performed the experiments: AT. Analyzed the data: AT RB JG GH. Wrote the paper: AT RB JG GH.

References

1. Dyson HJ, Wright PE (2005) Intrinsically unstructured proteins and their functions. Nat Rev Mol Cell Biol 6: 197–208.
2. Minezaki Y, Homma K, Kinjo AR, Nishikawa K (2006) Human transcription factors contain a high fraction of intrinsically disordered regions essential for transcriptional regulation. J Mol Biol 359: 1137–1149.
3. Radivojac P, Obradovic Z, Smith DK, Zhu G, Vucetic S, et al. (2004) Protein flexibility and intrinsic disorder. Protein Sci 13: 71–80.
4. Tompa P (2002) Intrinsically unstructured proteins. Trends Biochem Sci 27: 527–533.
5. Iakoucheva LM, Radivojac P, Brown CJ, O'Connor TR, Sikes JG, et al. (2004) The importance of intrinsic disorder for protein phosphorylation. Nucleic Acids Res 32: 1037–1049.
6. Whitmarsh AJ, Davis RJ (2000) Regulation of transcription factor function by phosphorylation. Cell Mol Life Sci 57: 1172–1183.

7. Desterro JM, Rodriguez MS, Hay RT (2000) Regulation of transcription factors by protein degradation. Cell Mol Life Sci 57: 1207–1219.

8. Wright PE, Dyson HJ (1999) Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm. J Mol Biol 293: 321–331.

9. Conkright MD, Montminy M (2005) CREB: the unindicted cancer co-conspirator. Trends Cell Biol 15: 457–459.

10. Mayr B, Montminy M (2001) Transcriptional regulation by the phosphorylation-dependent factor CREB. Nat Rev Mol Cell Biol 2: 599–609.

11. Zor T, Mayr BM, Dyson HJ, Montminy MR, Wright PE (2002) Roles of phosphorylation and helix propensity in the binding of the KIX domain of CREB-binding protein by constitutive (c-Myb) and inducible (CREB) activators. J Biol Chem 277: 42241–42248.

12. Radhakrishnan I, Perez-Alvarado GC, Parker D, Dyson HJ, Montminy MR, et al. (1997) 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.

13. Radhakrishnan I, Perez-Alvarado GC, Dyson HJ, Wright PE (1998) Conformational preferences in the Ser133-phosphorylated and non-phosphorylated forms of the kinase inducible transactivation domain of CREB. FEBS Lett 430: 317–322.

14. Johannessen M, Delghandi MP, Moens U (2004) What turns CREB on? Cell Signal 16: 1211–1227.

15. Campbell KM, Lumb KJ (2002) Structurally distinct modes of recognition of the KIX domain of CBP by Jun and CREB. Biochemistry 41: 13956–13964.

16. Comb M, Birnberg NC, Seasholtz A, Herbert E, Goodman HM (1986) A cyclic AMP- and phorbol ester-inducible DNA element. Nature 323: 353–356.

17. Montminy MR, Sevarino KA, Wagner JA, Mandel G, Goodman RH (1986) Identification of a cyclic-AMP-responsive element within the rat somatostatin gene. Proc Natl Acad Sci U S A 83: 6682–6686.

18. Montminy MR, Bilezikjian LM (1987) Binding of a nuclear protein to the cyclic-AMP response element of the somatostatin gene. Nature 328: 175–178.

19. Chrivia JC, Kwok RP, Lamb N, Hagiwara M, Montminy MR, et al. (1993) Phosphorylated CREB binds specifically to the nuclear protein CBP. Nature 365: 855–859.

20. Dal PM, Alber F, Carloni P (2001) Ser133 phosphate-KIX interactions in the CREB-CBP complex: an ab initio molecular dynamics study. Eur Biophys J 30: 75–81.

21. Mestas SP, Lumb KJ (1999) Electrostatic contribution of phosphorylation to the stability of the CREB-CBP activator-coactivator complex. Nat Struct Biol 6: 613–614.

22. Solt I, Magyar C, Simon I, Tompa P, Fuxreiter M (2006) Phosphorylation-induced transient intrinsic structure in the kinase-inducible domain of CREB facilitates its recognition by the KIX domain of CBP. Proteins 64: 749–757.

23. Sugase K, Dyson HJ, Wright PE (2007) Mechanism of coupled folding and binding of an intrinsically disordered protein. Nature 447: 1021–1025.

24. Schreiber G, Fersht AR (1996) Rapid, electrostatically assisted association of proteins. Nat Struct Biol 3: 427–431.

25. Vijayakumar M, Wong KY, Schreiber G, Fersht AR, Szabo A, et al. (1998) Electrostatic enhancement of diffusion-controlled protein-protein association: comparison of theory and experiment on barnase and barstar. J Mol Biol 278: 1015–1024.

26. Tang C, Iwahara J, Clore GM (2006) Visualization of transient encounter complexes in protein-protein association. Nature 444: 383–386.

27. Schreiber G (2002) Kinetic studies of protein-protein interactions. Curr Opin Struct Biol 12: 41–47.

28. Best RB, Hummer G (2005) Reaction coordinates and rates from transition paths. Proc Natl Acad Sci U S A 102: 6732–6737.

29. Hummer G (2004) From transition paths to transition states and rate coefficients. J Chem Phys 120: 516–523.

30. Clementi C, Plotkin SS (2004) The effects of nonnative interactions on protein folding rates: theory and simulation. Protein Sci 13: 1750–1766.

31. Tollinger M, Kloiber K, Agoston B, Dorigoni C, Lichtenecker R, et al. (2006) An isolated helix persists in a sparsely populated form of KIX under native conditions. Biochemistry 45: 8885–8893.

32. Shoemaker BA, Portman JJ, Wolynes PG (2000) Speeding molecular recognition by using the folding funnel: the fly-casting mechanism. Proc Natl Acad Sci U S A 97: 8868–8873.

33. Parker D, Jhala US, Radhakrishnan I, Yaffe MB, Reyes C, et al. (1998) Analysis of an activator:coactivator complex reveals an essential role for secondary structure in transcriptional activation. Mol Cell 2: 353–359.

34. Daggett V, Li A, Itzhaki LS, Otzen DE, Fersht AR (1996) Structure of the transition state for folding of a protein derived from experiment and simulation. J Mol Biol 257: 430–440.

35. Garcia AE, Onuchic JN (2003) Folding a protein in a computer: an atomic description of the folding/unfolding of protein A. Proc Natl Acad Sci U S A 100: 13898–13903.

36. Li L, Shakhnovich EI (2001) Constructing, verifying, and dissecting the folding transition state of chymotrypsin inhibitor 2 with all-atom simulations. Proc Natl Acad Sci U S A 98: 13014–13018.

37. Vendruscolo M, Paci E, Dobson CM, Karplus M (2001) Three key residues form a critical contact network in a protein folding transition state. Nature 409: 641–645.

38. Shea JE, Onuchic JN, Brooks CL (2000) Energetic frustration and the nature of the transition state in protein folding. J Chem Phys 113: 7663–7671.

39. Clementi C, Garcia AE, Onuchic JN (2003) Interplay among tertiary contacts, secondary structure formation and side-chain packing in the protein folding mechanism: all-atom representation study of protein L. J Mol Biol 326: 933–954.

40. Goto NK, Zor T, Martinez-Yamout M, Dyson HJ, Wright PE (2002) 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.

41. De Guzman RN, Goto NK, Dyson HJ, Wright PE (2006) Structural basis for cooperative transcription factor binding to the CBP coactivator. J Mol Biol 355: 1005–1013.

42. Yankeelov JA Jr, Koshland DE Jr (1965) Evidence for conformation changes induced by substrates of phosphoglucomutase. J Biol Chem 240: 1593–1602.

43. Kumar S, Ma B, Tsai CJ, Sinha N, Nussinov R (2000) Folding and binding cascades: dynamic landscapes and population shifts. Protein Sci 9: 10–19.

44. Ma B, Shatsky M, Wolfson HJ, Nussinov R (2002) Multiple diverse ligands binding at a single protein site: a matter of pre-existing populations. Protein Sci 11: 184–197.

45. Volkov AN, Worrall JAR, Holtzmann E, Ubbink M (2006) Solution structure and dynamics of the complex between cytochrome c and cytochrome c peroxidase determined by paramagnetic NMR. Proc Natl Acad Sci U S A 103: 18945–18950.

46. Northrup SH, Boles JO, Reynolds JCL (1988) Brownian dynamics of cytochrome-c and cytochrome-c peroxidase association. Science 241: 67–70.

47. Zhang C, Chen J, DeLisi C (1999) Protein-protein recognition: exploring the energy funnels near the binding sites. Proteins 34: 255–267.

48. Brubaker K, Cowley SM, Huang K, Loo L, Yochum GS, et al. (2000) Solution structure of the interacting domains of the Mad-Sin3 complex: implications for recruitment of a chromatin-modifying complex. Cell 103: 655–665.

49. Swanson KA, Knoepfler PS, Huang K, Kang RS, Cowley SM, et al. (2004) HBP1 and Mad1 repressors bind the Sin3 corepressor PAH2 domain with opposite helical orientations. Nat Struct Mol Biol 11: 738–746.

50. Tompa P (2005) The interplay between structure and function in intrinsically unstructured proteins. FEBS Lett 579: 3346–3354.

51. Latzer J, Papoian GA, Prentiss MC, Komives EA, Wolynes PG (2007) Induced fit, folding, and recognition of the NF-kappaB-nuclear localization signals by IkappaBalpha and IkappaBbeta. J Mol Biol 367: 262–274.

52. Levy Y, Onuchic JN (2006) Mechanisms of protein assembly: lessons from minimalist models. Acc Chem Res 39: 135–142.

53. Levy Y, Wolynes PG, Onuchic JN (2004) Protein topology determines binding mechanism. Proc Natl Acad Sci U S A 101: 511–516.

54. Wang J, Lu Q, Lu HP (2006) Single-molecule dynamics reveals cooperative binding-folding in protein recognition. PLoS Comput Biol 2: e78. doi:10.1371/journal.pcbi.0020078.

55. Wang J, Zhang K, Lu H, Wang E (2006) Dominant kinetic paths on biomolecular binding-folding energy landscape. Phys Rev Lett 96: 168101.

56. Levy Y, Cho SS, Onuchic JN, Wolynes PG (2005) A survey of flexible protein binding mechanisms and their transition states using native topology based energy landscapes. J Mol Biol 346: 1121–1145.

57. Borg M, Mittag T, Pawson T, Tyers M, Forman-Kay JD, et al. (2007) Polyelectrostatic interactions of disordered ligands suggest a physical basis for ultrasensitivity. Proc Natl Acad Sci U S A 104: 9650–9655.

58. Campbell KM, Terrell AR, Laybourn PJ, Lumb KJ (2000) Intrinsic structural disorder of the C-terminal activation domain from the bZIP transcription factor Fos. Biochemistry 39: 2708–2713.

59. Dames SA, Martinez-Yamout M, De Guzman RN, Dyson HJ, Wright PE (2002) Structural basis for Hif-1 alpha/CBP recognition in the cellular hypoxic response. Proc Natl Acad Sci U S A 99: 5271–5276.

60. Freedman SJ, Sun ZY, Kung AL, France DS, Wagner G, et al. (2003) Structural basis for negative regulation of hypoxia-inducible factor-1alpha by CITED2. Nat Struct Biol 10: 504–512.

61. Freedman SJ, Sun ZY, Poy F, Kung AL, Livingston DM, et al. (2002) Structural basis for recruitment of CBP/p300 by hypoxia-inducible factor-1 alpha. Proc Natl Acad Sci U S A 99: 5367–5372.

62. Yoon MK, Shin J, Choi G, Choi BS (2006) Intrinsically unstructured N-terminal domain of bZIP transcription factor HY5. Proteins 65: 856–866.

63. Zor T, De Guzman RN, Dyson HJ, Wright PE (2004) Solution structure of the KIX domain of CBP bound to the transactivation domain of c-Myb. J Mol Biol 337: 521–534.

64. Karanicolas J, Brooks CL III (2003) Improved Go-like models demonstrate the robustness of protein folding mechanisms towards non-native interactions. J Mol Biol 334: 309–325.

65. Miyazawa S, Jernigan RL (1996) Residue-residue potentials with a favorable contact pair term and an unfavorable high packing density term, for simulation and threading. J Mol Biol 256: 623–644.

66. Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, et al. (1983) Charmm-A program for macromolecular energy, minimization, and dynamics calculations. J Comp Chem 4: 187–217.

67. Brunger A, Brooks CL, Karplus M (1984) Stochastic boundary-conditions for molecular-dynamics simulations of St2 water. Chem Phys Lett 105: 495–500.

68. Pastor RW, Brooks BR, Szabo A (1988) An analysis of the accuracy of Langevin and molecular-dynamics algorithms. Mol Phys 65: 1409–1419.
69. Ryckaert JP, Ciccotti G, Berendsen HJC (1977) Numerical-integration of Cartesian equations of motion of a system with constraints - Molecular-dynamics of N-alkanes. J Comput Phys 23: 327–341.
70. Best RB, Chen YG, Hummer G (2005) Slow protein conformational dynamics from multiple experimental structures: The helix/sheet transition of arc repressor. Structure 13: 1755–1763.
71. Best RB, Hummer G (2006) Diffusive model of protein folding dynamics with Kramers turnover in rate. Phys Rev Lett 96: 228104.