PROTEINS
STRUCTURE ■ FUNCTION ■ BIOINFORMATICS

Mechanism and rate constants of the Cdc42 GTPase binding with intrinsically disordered effectors

Xiaodong Pang and Huan-Xiang Zhou*

Department of Physics and Institute of Molecular Biophysics, Florida State University, Tallahassee, Florida, 32306

ABSTRACT

Intrinsically disordered proteins (IDPs) are often involved in signaling and regulatory functions, through binding to cellular targets. Many IDPs undergo disorder-to-order transitions upon binding. Both the binding mechanisms and the magnitudes of the binding rate constants can have functional importance. Previously we have found that the coupled binding and folding of any IDP generally follows a sequential mechanism that we term dock-and-coalesce, whereby one segment of the IDP first docks to its subsite on the target surface and the remaining segments subsequently coalesce around their respective subsites. Here we applied our TransComp method within the framework of the dock-and-coalesce mechanism to dissect the binding kinetics of two Rho-family GTPases, Cdc42 and TC10, with two intrinsically disordered effectors, WASP and Pak1. TransComp calculations identified the basic regions preceding the GTPase binding domains (GBDs) of the effectors as the docking segment. For Cdc42 binding with both WASP and Pak1, the calculated docking rate constants are close to the observed overall binding rate constants, suggesting that basic-region docking is the rate-limiting step and subsequent conformational coalescence of the GBDs on the Cdc42 surface is fast. The possibility that conformational coalescence of the WASP GBD on the TC10 surface is slow warrants further experimental investigation. The account for the differences in binding rate constants among the three GTPase-effector systems and mutational effects therein yields deep physical and mechanistic insight into the binding processes. Our approach may guide the selection of mutations that lead to redesigned binding pathways.

Proteins 2016; 84:674–685.

Key words: intrinsically disordered proteins; coupled binding and folding; dock-and-coalesce mechanism; transient-complex theory; binding kinetics.

INTRODUCTION

A significant portion of cellular proteins is now recognized as intrinsically disordered or as containing intrinsically disordered regions (IDPs/IDRs).[1,2] Many IDPs/IDRs play signaling or regulatory roles, through binding with target proteins (or other domains of the IDR-containing proteins).[3] Upon binding the IDPs often become ordered on the surfaces of the target proteins. The conformational ensembles of unbound IDPs,[4–7] structures of the bound IDPs,[8–14] and binding thermodynamics[15–17] have been intensely characterized. In comparison, the kinetics and mechanisms of the coupled folding and binding of IDPs have been less scrutinized, even though the binding pathways as well as the magnitudes of binding rate constants can be functionally important. Both experimental[18–24] and computational studies[25–32] have suggested that the structures of IDPs bound to their targets accrue sequentially on the latters' surfaces (hence cane be termed induced fit). More specifically, we have proposed dock-and-coalesce as a common mechanism for the coupled folding and binding, whereby one segment of the IDP first docks to its subsite on the target surface and the remaining segments subsequently coalesce around their respective subsites.[33] A number of questions then arise. What distinguish the "docking" segment from the "coalescing" segments? What physical properties dictate the magnitudes of the rate constants of the docking and coalescing steps, and therefore their

Additional Supporting Information may be found in the online version of this article.

Grant sponsor: National Institutes of Health Grant; Grant number: GM0585187.
*Correspondence to: Huan-Xiang Zhou, Department of Physics and Institute of Molecular Biophysics, Tallahassee, Florida 32306. E-mail: hzhou4@fsu.edu
Received 8 October 2015; Revised 25 January 2016; Accepted 3 February 2016
Published online 15 February 2016 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/prot.25018

relative contributions to the overall binding rate constant? Here we use computation to address these questions for the binding of the intrinsically disordered effectors of Cdc42 to this GTPase.

Cdc42, along with Rac1, belongs to the Rho family of p21 GTPases, which regulates actin cytoskeleton organization through interacting with various effectors. In particular, upon activation by GTP-loaded Cdc42, WASP, the protein associated with the immune disorder Wiskott–Aldrich syndrome, stimulates Arp2/3-mediated nucleation of actin polymerization.[34–40] The 502-residue WASP comprises a WASP-homology-1 domain (WH1), a basic region (BR; residues 230–237), a GTPase binding domain (GBD; residues 238–288) containing a Cdc42/Rac1 interactive binding motif (CRIB; residues 238–251), a proline-rich region, and the VCA domain (V: verprolin-homology region; C: cofilin-homology region; and A: acidic-rich region) [Fig. 1(a)]. Unbound WASP is autoinhibited, with the VCA domain trapped through intramolecular binding with the GBD.[34] In Cdc42-bound WASP, the BR and GBD are engaged in intermolecular interactions,[10] thereby releasing the VCA domain for interacting with Arp2/3 [Fig. 1(b)].

The WASP fragment containing the BR and GBD (residues 230–288; hereafter $WASP_{frag}$) undergoes disorder-to-order transition upon binding Cdc42. On the Cdc42 surface, the BR and CRIB adopt extended conformations while downstream residues of the GBD form a compact subdomain comprising a β hairpin and an α-helix [Fig. 1(c)].[10] The four basic residues in the BR [Fig. 1(d)] interact with acidic residues (Glu49 and Glu178) in the β2-β3 turn and α5 helix of Cdc42, respectively, while the CRIB is accommodated in the groove between α1/α5 and β2 of Cdc42. Experimental characterization of the $WASP_{frag}$-Cdc42 binding kinetics[20] showed very fast association, with an association rate constant (k_a) of $2.2\times10^7\ M^{-1}\ s^{-1}$ (at an ionic strength of 95 m*M*). Mutation of the $K_{230}KK_{232}$ motif in the BR to an oppositely charged EEE motif resulted in a ~250-fold decrease in k_a, while the charge reversal double mutation E49K/E178K on Cdc42 resulted in a ~35-fold decrease in k_a. These data suggest that electrostatic attraction between the BR and its subsite on Cdc42 makes a strong contribution to k_a. Significant salt effects also support the important role of electrostatic interactions in the fast $WASP_{frag}$-Cdc42 association. However, the ability of kinetic experiments alone to achieve further mechanistic dissection is limited.

Another Cdc42 effector involved in mediating actin cytoskeleton dynamics is Pak1, the founding member of p21-actived kinases. Like WASP, unbound Pak1 is autoinhibited, by dimerization (through residues 81–87) and interaction of the inhibitory switch domain (residues 88–136) and kinase inhibitory segment (residues 137–149) with the kinase domain (residues 249–545).[41] Binding to GTP-loaded Cdc42 activates Pak1, by dissociating the dimer and freeing the kinase domain from the inhibitory switch domain and kinase inhibitory segment. The freed kinase domain then initiates a signal transduction cascade, leading to decreased depolymerization of F-actin.[42]

Pak1 also contains a CRIB (residues 75–88), and the shortest fragment that binds Rac1 with high affinity extends to residue 118, with a modest increase in affinity provided by the N-terminal extension to residue 70.[43] Thus Cdc42/Rac1 activation of Pak1 can be explained by the partial overlap of the latter's GDB (residues 75–118) and the dimerization segment and inhibitory switch domain. On the N-terminal and C-terminal sides of the CRIB, Pak1 and WASP residues that potentially interact with Cdc42 show moderate to weak sequence identity [Fig. 1(d)]. The N-terminal BR of Pak1 contains four basic residues along with three acidic residues, so the net charge is less positive than in the Cdc42 BR. Correspondingly, less intense interaction between the Pak1 BR and Cdc42 can be anticipated, but no experimental structural information is available for this interaction. The structure of the Pak1 GBD bound to Cdc42 has been determined,[44] and reveals a similar interaction pattern as the WASP GBD, with the CRIB adopting an extended conformation while downstream residues of the GBD forming a compact subdomain comprising a β hairpin and an α-helix [Fig. 1(d) and Supporting Information Fig. S1a). The charge reversal double mutation E49K/E178K on Cdc42 resulted in a ~5-fold decrease in k_a for binding a Pak1 fragment containing the BR and GBD (hereafter $Pak1_{frag}$),[20] consistent with a modest effect on affinity by a 5-residue extension N-terminal to the CRIB[43] and supporting less intense electrostatic attraction of Cdc42 with the Pak1 BR than with the WASP BR.

The fast association between Cdc42 and $WASP_{frag}$ appears to be functionally essential, as another member of the Rho family of p21 GTPases, TC10, with 66.7% sequence identity [Fig. 1(e)] and a nearly identical dissociation rate constant but an apparent three orders of magnitude lower association rate constant, only minimally stimulated actin polymerization.[20] A yeast two-hybrid assay found that TC10 did not interact with WASP, suggesting that the latter is not an endogenous effector of TC10.[45] The sequence alignment between Cdc42 and TC10 [Fig. 1(e)] shows that the two acidic Cdc42 residues, Glu49 and Glu178, important for electrostatic interactions with the WASP BR, are replaced by oppositely charged and neutral TC10 residues (Lys63 and Thr192), respectively. However, mutating both Lys63 and Thr192 in TC10 to Glu (Supporting Information Fig. S1b) yielded only partial rescue of the rate constant for binding $WASP_{frag}$, reducing the k_a gap between Cdc42 and TC10 apparently from three to two orders of magnitude.[20] The remaining k_a gap could suggest different rate-limiting steps for the binding of these two GTPases with $WASP_{frag}$. Accounting for the differences in k_a among the three GTPase-effector systems and mutational effects therein, as done here, will yield deep physical and mechanistic insight into the binding processes.

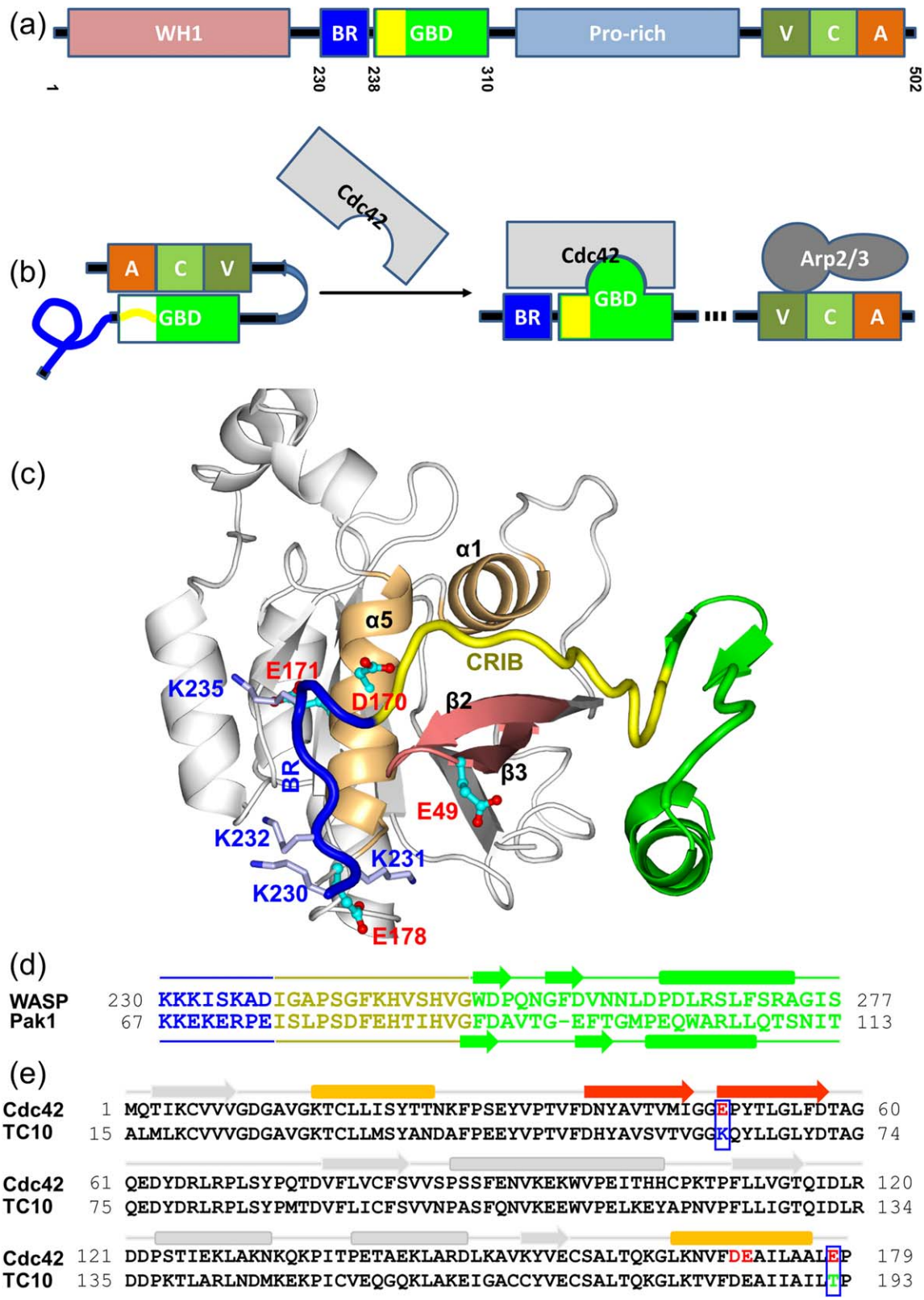

Figure 1

Structure and activation of WASP. (**a**) Domain organization of WASP. (**b**) Autoinhibition of WASP and activation by Cdc42. (**c**) Structure of Cdc42-bound $WASP_{frag}$ comprising the BR (residues 230–237; blue), the CRIB (residues 238–251; yellow), and the C-terminal β-hairpin and α-helix of the GBD (residues 253–277; green). Four basic residues of the WASP BR and four acidic residues on the Cdc42 surface are shown as stick and ball-and-stick, respectively. (**d**) Sequence alignment of $WASP_{frag}$ and $Pak1_{frag}$. The BRs, CRIBs, and C-terminal subdomains of the GBDs are shown by colors matching those in panel (c). (**d**) Sequence alignment of Cdc42 and TC10. Two key charge substitutions are highlighted by box; two other common acidic residues interacting with the WASP BR are highlighted by red letter.

MATERIALS AND METHODS

Rate constants were calculated by the TransComp method.[28,29] This method requires only the structure of the native complex as input, and the method of calculation is equivalent to producing 4000 trajectories in which the translational and rotational diffusion of the subunits is simulated. In such Brownian dynamics simulations, the steric interactions as well as long-range electrostatic interactions between the subunits are calculated, but the solvent is treated implicitly using the Poisson-Boltzmann model. The subunit molecules are treated as rigid; in essence, the internal conformational changes are assumed to be rapid so the molecules can quickly adapt to the native conformations when they reach the native complex by translational and rotational diffusion. The foregoing Brownian dynamics simulation based approach is well established and widely used.[46,47]

TransComp is based on the same general approach. Besides explicitly defining the concept of transient complex and prescribing its specification, a significant technical advance is the treatment of the long-range electrostatic interactions.[48] We found that, to a very high accuracy, the effect of the long-range electrostatic interactions during the Brownian dynamics simulations can be captured by the Boltzmann factor of the electrostatic interaction energy ΔG^*_{el} in the transient complex.[49] Therefore, in our Brownian dynamics simulations, the electrostatic interactions are turned off, yielding the basal rate constant k_{D0}, and the electrostatic contribution is recouped by calculating ΔG^*_{el}. The overall diffusion-limited binding rate constant is then

$$k_D = k_{D0} \exp(-\Delta G^*_{el}/k_B T) \quad (1)$$

Using TransComp for the binding kinetics of IDPs requires further adaption.[28,29] In particular, the TransComp calculation is done using the docking segment of an IDP as one subunit and the target protein as the other subunit. That is, we assume that the docking segment undergoes rapid conformational exchange, such that it adopts the native conformation when it approaches its subsite on the target surface. During this approach, the remaining segments can stay in non-native conformations but are assumed to minimally affect the interaction between the docking segment and the target protein (Fig. 2). Based on the latter assumption, the remaining segments are completely removed in the TransComp calculation.

Below some details on modeling the native complexes and the components of TransComp calculations are described.

Structural modeling of the TC10-WASP$_{frag}$ complex

The TC10-WASP$_{frag}$ complex was modeled after the Cdc42-WASP$_{frag}$ complex (Protein Data Bank (PDB) entry 1CEE[10]), with Cdc42 replaced by TC10 using their sequence alignment from the Dali server[50] [Fig. 1(e)]. Both Lys63 and Thr192 in TC10 were mutated to the corresponding Glu residues in Cdc42, so the modeled complex was for the TC10 K63E/T192E mutant bound with WASP$_{frag}$. After removal of steric clash by standard energy minimization in Amber, the WASP BR was relaxed manually to release the N-terminus from being trapped and refined again by energy minimization (Supporting Information Fig. S1b).

Structural modeling of the basic region of Cdc42-bound Pak1

Using Modeller[51] on the sequence alignment of Figure 1(d), an initial conformation of the Pak1 BR was modeled after the WASP BR, after superimposing the GBD Cα atoms in their respective complexes with Cdc42 (PDB entries 1E0A[44] and 1CEE, respectively). The Pak1 BR was then relaxed manually and refined by energy minimization (Supporting Information Fig. S1a).

k_D calculation by the TransComp method

TransComp and the associated parameters have undergone thorough validation,[28,29,47,48] and were used here without modification. Each TransComp calculation consists of three components: identification of the transient complex by mapping the inter-subunit interaction energy landscape around the native complex, determination of the basal rate constant k_{D0} by Brownian dynamics simulations, and computation of the electrostatic interaction energy ΔG^*_{el} in the transient complex by solving the Poisson-Boltzmann equation.

In identifying the transient complex, the interaction energy is simply modeled by the number of contacts (both native and non-native), N_c, between the apposing binding sites on the two subunits in any clash-free configuration. A total of 10^7 such configurations were generated by randomly sampling the six degrees of relative translational/rotational freedom in and around the bound-state energy well. As illustrated in Figure 3(a), the bound state is characterized by configurations with high N_c but a very restricted accessible range of the relative rotation angle χ, as indicated by small standard deviations, σ_χ, of χ calculated among configurations at a given N_c. When the two subunits move out of the bound-state energy well, there is a sharp increase in σ_χ. We fit the dependence of σ_χ on N_c to a function used for modeling two-state protein denaturation, and identify the midpoint, where N_c is designated N_c^*, of this fit as the transient complex. That is, configurations with $N_c = N_c^*$ constitute the transient-complex ensemble, and configurations with $N_c > N_c^*$ fall in the bound-state well.

The basal rate constant k_{D0} is calculated from Brownian dynamics simulations in which electrostatic interactions are turned off. Each Brownian dynamics trajectory starts from the bound-state well (that is, from a configuration

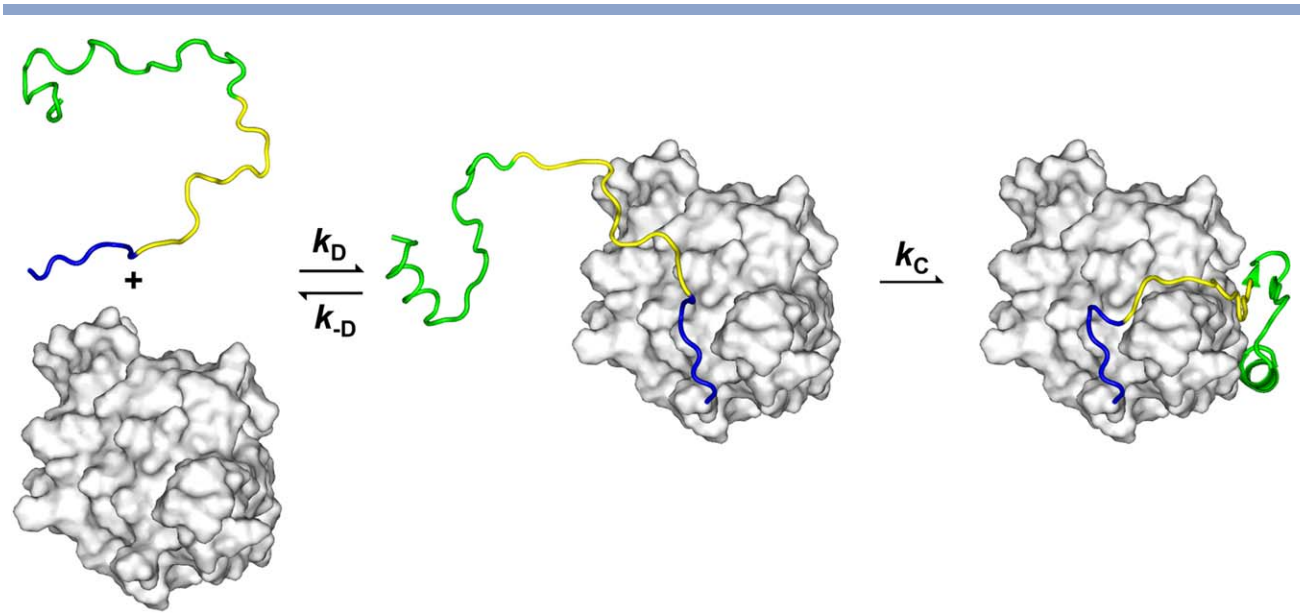

Figure 2

The dock-and-coalesce mechanism, illustrated on Cdc42-WASP$_{frag}$ binding. The WASP BR (blue) first docks to the Cdc42 surface (gray), with rate constant k_D. The resulting docked complex may subsequently either undock (with rate constant k_{-D}) or evolve to the native complex through conformational coalescence of the CRIB (yellow) and C-terminal subdomain (green) on the Cdc42 surface (with rate constant k_C). [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

with $N_c > N_c^*$) and is propagated in the translational/rotational space. At each time step where the protein pair is in the bound-state well, it is given a certain rate γ to form the native complex. If that happens, the trajectory is terminated. All trajectories are otherwise terminated at a cutoff time. The survival fraction of 4,000 trajectories as a function of time, upon extrapolating to infinite γ, allows k_{D0} to be calculated.[52]

The electrostatic interaction energy ΔG_{el}^* in the transient complex is calculated by numerically solving the Poisson-Boltzmann equation. We randomly choose 100 configurations from the transient-complex ensemble, calculate the electrostatic interaction energy for each at the desired ionic strength, and take the average to yield ΔG_{el}^*. Finally the diffusion-limited rate constant is calculated according to Eq. (1).

To obtain k_D for a mutant, the basal rate constant k_{D0} was assumed to be unaltered and only the electrostatic interaction energy ΔG_{el}^* was recalculated after molecular modeling of the mutation. Mutations were introduced by Pymol (www.pymol.org) and refined by side-chain energy minimization.

RESULTS

Framework for modeling the binding kinetics of IDPs

The extended conformation of WASP$_{frag}$ formed on the Cdc42 surface [Fig. 1(c)] typifies the bound structures of IDPs, with stability largely arising from intermolecular (rather than intramolecular) interactions distributed over elongated interfaces. In the binding process, different regions of an IDP are unlikely to form contact with the target protein all at once. Instead the binding may proceed via parallel pathways, each with a different region of the IDP acting as the initial docking segment, followed by structural consolidation on the target surface. Presumably, for a given IDP, a single such dock-and-coalesce pathway, or a very small number of them, yield much higher binding rate constants than alternative pathways, hence make the dominant contribution to the overall rate of the productive formation of the native complex.

Importantly, this dock-and-coalesce mechanism also provides a framework for quantitatively modeling the binding kinetics of IDPs[28,29,33] (Fig. 2). In the docking step, the docking segment approaches its subsite on the target protein by translational/rotational diffusion, and simultaneously undergoes rapid conformational exchange to reach a docked complex in which the docking segment is bound but the remaining segment(s) of the IDP are still loose. This step is likely to be rate-limited by the diffusional approach, which in turn can be sped up enormously by long-range electrostatic attraction of the target protein.[47] In the subsequent coalescing step, the remaining segments evolve toward their bound conformations, with energy barriers lowered by interactions with surface residues of the target protein.[27] If the rate constants are k_D and k_{-D}, respectively, for the docking step and the its reverse process, that is, the undocking step, and k_C for the coalescing step, then the overall association rate constant is

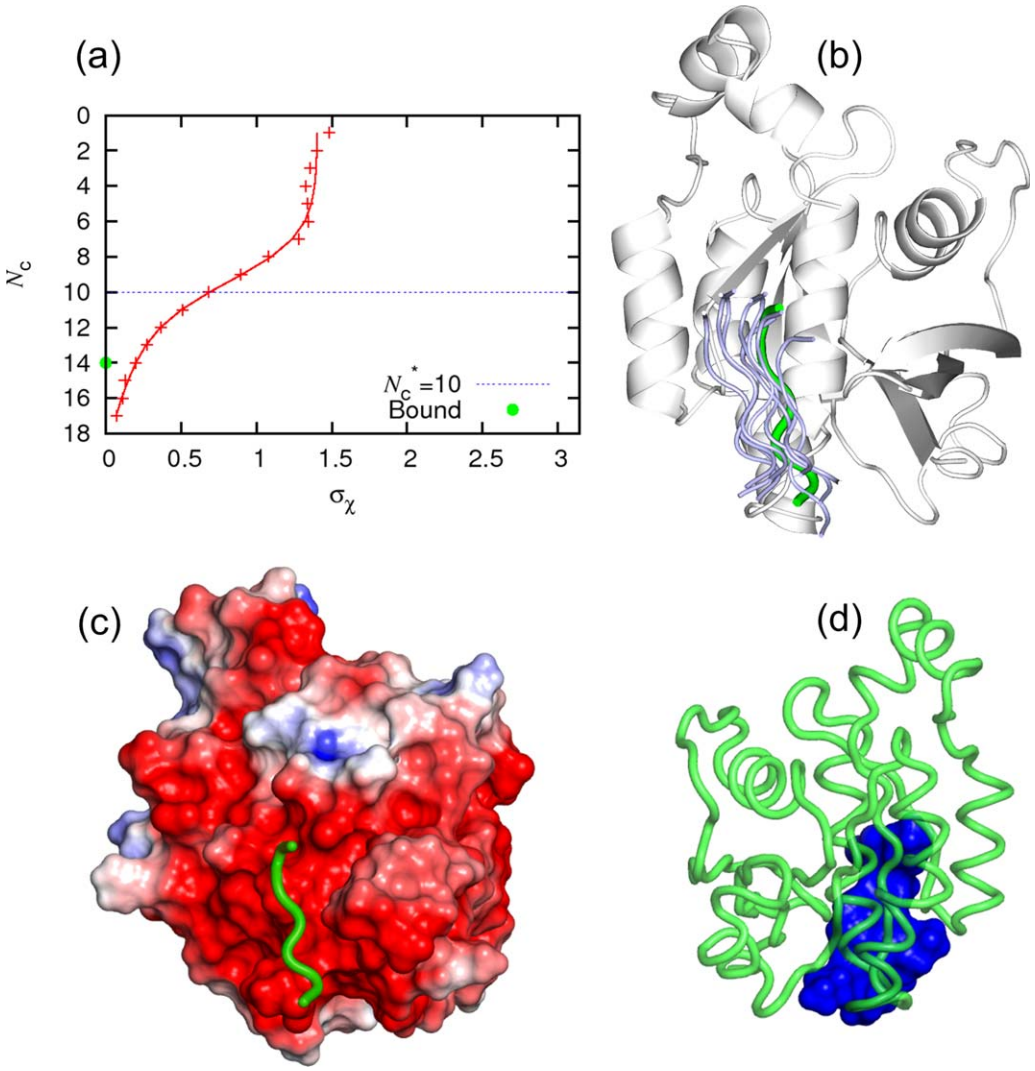

Figure 3

Calculation of the rate constant for the WASP BR (residues 230–236) docking to Cdc42. (**a**) Sharp transition from the unbound regime characterized by large relative translational/rotational freedom (i.e., high standard deviation, σ_χ, in the relative rotation angle χ) to the bound regime characterized by numerous intermolecular contacts (that is, high N_c). The native complex has $N_c = 14$ (green dot). The transient complex is located at the mid-point of the transition (blue line; $N_c^* = 10$). (**b**) Representative configurations of the BR (light blue) in the transient-complex ensemble, as illustrated by eight representative poses. The native-complex BR is shown as green for reference. (**c,d**) The electrostatic surfaces of Cdc42 and the WASP BR. Each surface is accompanied by a ribbon representation of the other subunit colored in green in the native complex, to indicate the binding site. Panel (c) shows the "front" view of panel (b); panel (d) shows the "back" view. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

$$k_a = \frac{k_D k_C}{k_{-D} + k_C} \tag{2}$$

For any diffusion-limited binding event, including the docking step of an IDP, we have developed a transient-complex theory for predicting the binding rate constant.[48] The transient complex refers to an on-pathway late intermediate in which the two binding partners have near-native separation and relative orientation, but have not yet formed most of the stereospecific short-range interactions. Note that the focus in this paragraph is the docking step, not the whole binding process of the IDP (Fig. 2). The binding partners are thus the docking segment and its subsite on the target surface. Correspondingly the transient complex here occurs intermediately before the successful docking of the docking segment only; the coalescing step is to occur subsequently. Once the docking segment and its subsite diffuse to the transient complex, located at the outer boundary of an energy well created by the native interactions between the two binding partners, they rapidly fall into it to generate the docked complex.

We have developed an automated implementation of the transient-complex theory, known as TransComp, and adapted it for IDP binding.[28] From our understanding

of the determinants for the wide variation in association rate constants,[47] we predict the dominant binding pathway to be the one where the docking step is rate-enhanced by electrostatic attraction and close to being rate-limiting for the overall binding process. By applying the TransComp method to different fragments of the IDP complexed with the target protein, we can identify the docking segment of the dominant pathway and predict the rate constant of the docking step, which serves as an upper bound for the rate constant of the overall binding process (see Equation (2)). This approach has yielded rate constants in good agreement with experimental results for a number of IDPs,[28,29,33] and is used below to account for the differences in k_a among the three GTPase-effector systems and mutational effects therein.

Table I
Comparison of Calculated (Calc) and Experimental (Expt) Results for the Association Rate Constants (in μM^{-1} s^{-1}) of Wild-Type and Mutant Cdc42-$WASP_{frag}$ Complexes

	Expt[a]	Calc
Wild-type	22	32.8
Cdc42 E49K	3.5	1.3
Cdc42 E49K/E178K	0.59	0.24
WASP K231E	0.59	0.33[b]
WASP $K_{230}KK_{232}$ to AAA	0.72	0.45[b]
WASP $K_{230}KK_{232}$ to EEE	0.09	0.17[b]

[a]Experimental results from Hemsath *et al.*[20]
[b]The rate constant for docking the CRIB, 1.7×10^5 M^{-1} s^{-1}, is added to the results calculated for the BR as the docking segment.

Cdc42-$WASP_{frag}$ binding

TransComp prediction of the diffusion-limited docking rate constant k_D starts with identification of the transient complex using the native complex as input.[28,29] As the two binding partners approach each other to form the docked complex, they abruptly lose relative translational/rotational freedom (as represented by the standard deviation, σ_χ, in the relative rotation angle χ) but gain intermolecular contacts (N_c) [Fig. 3(a)]. The transient complex is located at the mid-point of this transition [Fig. 3(a,b)]. In line with the assumption that rapid conformational exchange allows the docking segment of the IDP to achieve its bound conformation during the docking step, in the TransComp calculation the docking segment, like the target protein, takes the bound conformation. However, if the docking segment were extended to include coalescing segments, then binding would no longer be a single-step process. TransComp identifies this situation by a large gap in N_c, as occurs when the full $WASP_{frag}$ bound with Cdc42 (PDB entry 1CEE) was used for k_D prediction (Supporting Information Fig. S2a).

We broke $WASP_{frag}$ into three segments for separate TransComp runs in order to identify the likely docking segment. With the BR segment comprising residues 230–236 [Fig. 3(a,b)], the predicted docking rate constant k_D is 3.3×10^7 M^{-1} s^{-1} (at an ionic strength of 95 mM). This segment was selected for its highest k_D among all the segments starting at residue 230 and ending successively at residues 234 − 255 (Supporting Information Fig. S3a). In comparison, the docking rate constant would be two to three orders of magnitude lower, at 1.7×10^5 and 1.9×10^4 M^{-1} s^{-1}, respectively, for the CRIB segment (residues 237–251; Supporting Information Fig. S2b) and the C-terminal segment (residues 252–288; Supporting Information Fig. S2c). With this large separation in predicted docking rate constants, we can identify the BR with high confidence as the docking segment.

As critical support of the BR as docking segment, the predicted k_D, 3.3×10^7 M^{-1} s^{-1}, only slightly overestimates the experimental k_a, 2.2×10^7 M^{-1} s^{-1}, for the association of $WASP_{frag}$ with Cdc42.[20] The close agreement between the k_D and k_a suggests that the subsequent coalescing step is fast and has only a modest effect in slowing down the overall binding process. Hereafter we use the docking rate constant k_D as the prediction for k_a. The TransComp calculation further teases out this k_D, into a basal rate constant of 3.0×10^5 M^{-1} s^{-1}, and a two orders of magnitude enhancement by a very favorable ΔG^*_{el} of −2.8 kcal/mol. The latter reflects the strong electrostatic attraction across the binding interface [Figs. 1(c) and 3(c,d)].

Given the strong electrostatic rate enhancement, it is not surprising that a significant salt dependence of k_D is found. The calculated k_D values at six ionic strengths from 45 to 1045 mM agree closely with the experimental results (Supporting Information Table S1). The calculated k_D values, 1.3×10^6 and 2.4×10^5 M^{-1} s^{-1}, respectively, for the E49K single mutant and the E49K/E178K double mutant of Cdc42 modestly underestimate the measured k_a results (Table I). The discrepancy suggests, with the slowing down of the BR-initiated pathway by the charge reversal, contributions of alternative pathways may become important. For example, adding the contribution of the CRIB-initiated pathway (i.e., 1.7×10^5 M^{-1} s^{-1}) would nearly double the rate constant of the E49K/E178K mutant and improves the agreement with the experimental result (Table I).

The findings for charge neutralization and reversal mutations in the BR segment are similar (Table I). Notably, with the charge reversal of $K_{230}KK_{232}$, the BR-initiated pathway slows down to a k_D of 2×10^3 M^{-1} s^{-1}, and therefore the CRIB-initiated pathway now potentially becomes dominating and accounts largely for the observed association rate constant of 9×10^4 M^{-1} s^{-1}.

The foregoing TransComp results clearly indicates that Cdc42-$WASP_{frag}$ binding proceeds with the BR docking, followed by rapid conformational coalescence of the CRIB and C-terminal subdomain of the GBD on the Cdc42 surface (Fig. 2). To examine the generality of this binding mechanism, below we extend the study to the other two related GTPase-effector systems.

TC10-WASP$_{frag}$ binding

The structure of the TC10-WASP$_{frag}$ complex is not available in the PDB. Given the high sequence identity between TC10 and Cdc42, we used the structure of the Cdc42-WASP$_{frag}$ complex to build a model for the TC10-WASP$_{frag}$ complex. We then relaxed the BR of WASP$_{frag}$, and mutated Lys63 and Thr192 in TC10 to the corresponding Glu residues in Cdc42 (Supporting Information Fig. S1b). TransComp calculations were carried out on the latter structure to identify WASP residues 230–239 as constituting the docking segment, with a rate constant of 6.9×10^6 M^{-1} s^{-1} for docking to the TC10 K63E/T192E (EE) mutant [Supporting Information Figs. S3b and 4(a,b)]. This value is 30-fold higher than the reported k_a for WASP$_{frag}$ binding with the TC10 EE mutant.[20]

The high k_D can be attributed to strong electrostatic attraction between the docking segment and the TC10 EE mutant [Fig. 4(c,d)]. Indeed, the TC10 EE mutant appears to have a stronger negative electrostatic surface than Cdc42 [Figs. 3(c) and 4(c)], with the former's acidic residue Asp40 and neutral residue Cys167 corresponding to the latter's neutral residue Asn26 and basic residue Lys153, respectively [Fig. 1(e)]. With $\Delta G^*_{el} = -1.3$ kcal/mol, the electrostatic attraction between the docking segment and the TC10 EE mutant contributes a 8.5-fold enhancement to their docking rate constant. For the K63E single mutant and wild-type TC10, TransComp predicted 2.6- and 16-fold reductions, respectively, in k_D. These results somewhat overestimate the 1.2- and 7.9-fold reductions of the reported k_a values for these TC10 variants.

Several possibilities can be suggested for the 30-fold discrepancy between the calculated k_D and reported k_a for WASP$_{frag}$ binding with the TC10 EE mutant. First of all, the model for the complex built on the structure of the Cdc42-bound counterpart may be incorrect. However, given the high sequence identity, the structural model is unlikely to be grossly wrong and the docking of the BR segment is almost surely rate-enhanced by electrostatic attraction. Secondly, it is possible that, contrary to the case of binding with Cdc42, the coalescing step in binding with TC10 is very slow as to rate-limit the overall binding process. Hemsath *et al.*[20] noted that, among the small number of Cdc42 residues that contact WASP$_{frag}$ and differ from the corresponding TC10 residues in amino-acid identity, Ile21, Thr24, and Thr25 interact with the CRIB. When the TC10 residues, Met35, Ala38, and Asn39, were substituted into the Cdc42 counterparts, the association rate constant changed little. This observation does not support a significant difference in the coalescing rate constants of Cdc42 binding and TC10 binding.

The third possibility is that the reported k_a underestimates the rate constant for WASP$_{frag}$ binding with the TC10 EE mutant. As support for this contention, the TC10 EE mutant was almost as effective as Cdc42 and more effective than the Cdc42 E49K mutant in stimulating actin polymerization.[20] If the stimulation activity is a measure of the rate constant for binding WASP$_{frag}$, then the k_a of the TC10 EE mutant should be bracket by those of wild-type Cdc42 and the E49K mutant, that is, between 2.2×10^7 and 3.5×10^6 M^{-1} s^{-1}. Interestingly, our k_D for the TC10 EE mutant, at 6.9×10^6 M^{-1} s^{-1}, actually falls into this bracket.

Cdc42-Pak1$_{frag}$ binding

Relative to the WASP BR, the increase in acidic residues in the Pak1 BR [Fig. 1(d)] is expected to weaken its interaction with Cdc42. We modeled the Pak1 BR onto the structure of the Pak1 GBD (residues 75–118) bound with Cdc42 (PDB entry 1E0A), using the WASP BR in PDB entry 1CEE as a guide but loosened the interaction with Cdc42 (Supporting Information Fig. S1a). Based on TransComp calculations on various fragments containing the Pak1 BR (Supporting Information Fig. S3c), the docking segment was identified as comprising residues 66–77 [Fig. 5(a,b)]. The docking rate constant is 5.0×10^5 M^{-1} s^{-1}, with a modest 2.6-fold rate enhancement by electrostatic attraction [$\Delta G^*_{el} = -0.6$ kcal mol^{-1}; Fig. 5(c,d]. This calculated k_D is in accord with the experimental k_a of 6.1×10^5 M^{-1} s^{-1}.[20] Moreover, the reductions in k_D by 1.4- and 2.5-fold, respectively, for the E49K and E49KE178K mutants agree reasonably well with the 2.0- and 5.1-fold reductions in experimental k_a (Table II).

It appears that the binding of Pak1$_{frag}$, similar to the binding of WASP$_{frag}$, largely proceeds through initial docking by the BR, followed by rapid conformational coalescence of the GBD on the Cdc42 surface (Fig. 2). Due to the introduction of acidic residues in the BR, the binding rate constant in this pathway is lower. As a result the contribution of a CRIB-initiated pathway to the overall binding rate constant may not be negligible.

DISCUSSION

We have dissected the binding kinetics of the Cdc42-WASP and other two related GTPase-effector systems within the framework of the dock-and-coalesce mechanism. For Cdc42-WASP binding, our TransComp calculations have identified the BR of WASP as the docking segment, due to its much higher docking rate constant compared to alternatives. Moreover, the predicted docking rate constant is close to the observed association rate constant.[20] We therefore conclude that the coalescing step involving the GBD is sufficiently fast as to making only a modest contribution to the overall association rate constant. As support of this conclusion, TransComp predictions have rationalized the effects of salt and key mutations on the association rate constant.

The limitations of TransComp calculations, as with the whole Brownian dynamics simulation based approach,

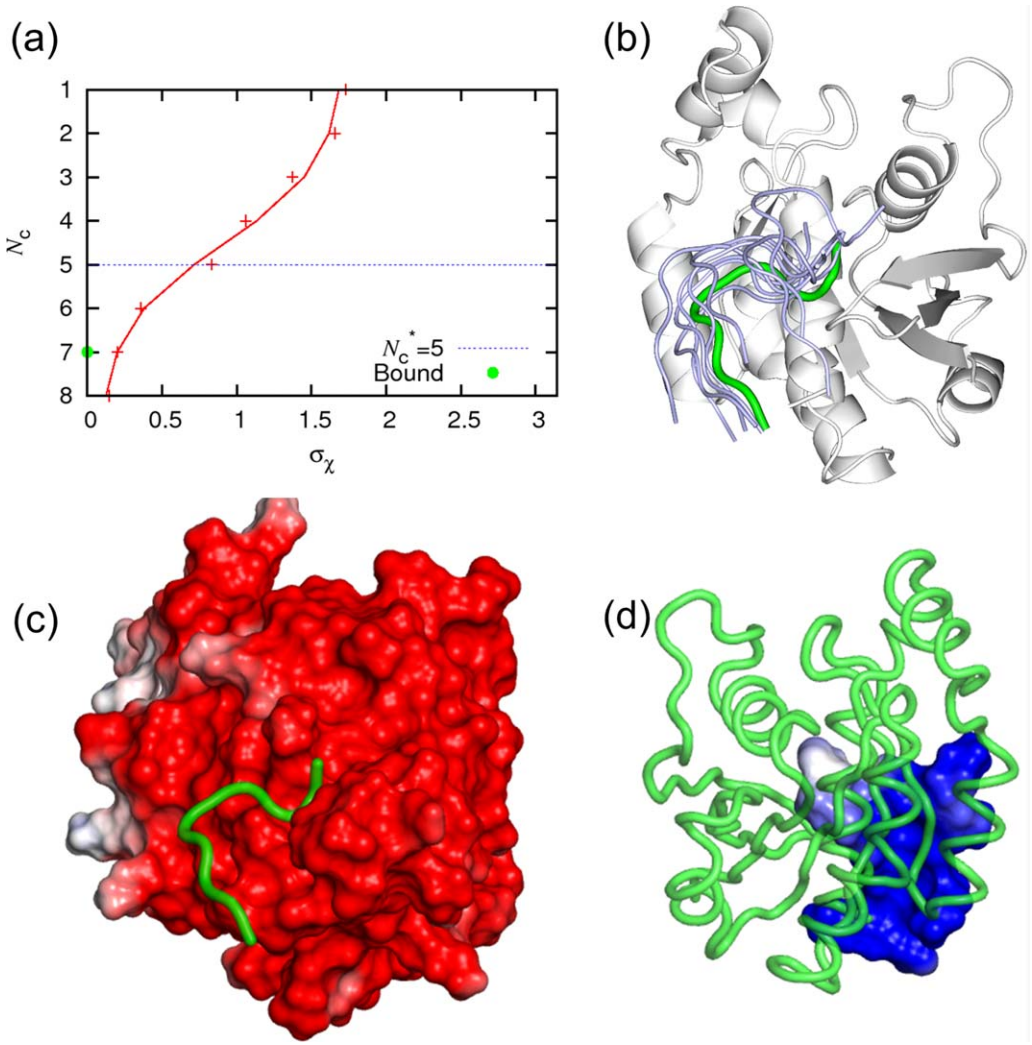

Figure 4

Calculation of the rate constant for the WASP BR (residues 230–239) docking to the TC10 K63E/T192E double mutant. (**a**–**d**) See Figure 3 legend. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

should be noted. First, the conformational exchange dynamics of the docking segment is assumed to be fast. In theory, slow conformational exchange can be treated in TransComp (by not extrapolating the exchange rate γ to infinity). However, the value of γ must be supplied. Given that the predicted diffusion-limited rate constant (that is, assuming infinite γ) is already close to the experimental value for overall binding, it seems likely that fast conformational exchange is not a bad assumption for Cdc42-WASP binding. Second, the effects of solvent and ions are treated not explicitly, but implicitly through the Poisson-Boltzmann model. Although a large body of literature supports the Poisson-Boltzmann model, its use in relation to an explicit solvent treatment for IDP binding kinetics warrants further investigation. There have been many simulations of IDPs in explicit solvent; however, calculating binding rate constants from these simulations remains a significant challenge.

The binding of WASP with TC10 and of Pak1 with Cdc42 also appears to start mostly with the docking of the effector BR. However, in these two cases, due to charge substitutions in either the GTPase or the effector, the docking rate constant of the BR is significantly reduced relative to the WASP case, and hence alternative pathways, including the CRIB-initiated pathway, may come into play. Additionally, after BR docking, conformational coalescence of the WASP GBD on the TC10 surface may possibly slow down the overall binding process, but this possibility warrants further experimental investigation. Overall, through TransComp calculations, account for the differences in k_a among the three GTPase-effector systems and mutational effects therein

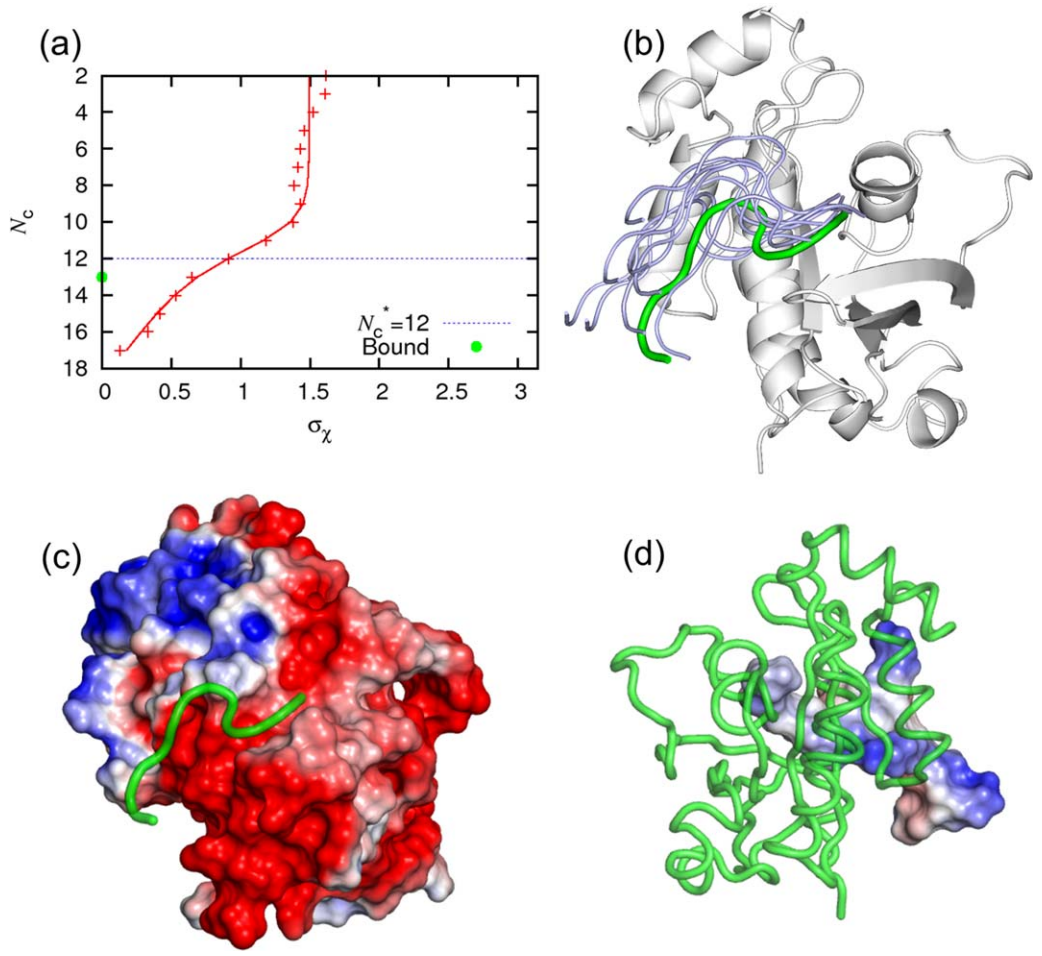

Figure 5

Calculation of the rate constant for the Pak1 BR (residues 67–77) docking to Cdc42. (a-d) See Figure 3 legend. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

have yielded deeper physical and mechanistic insight into the binding processes.

The prominent role of the effector BRs in initiating the docking step is identified here from studying the binding of the effector fragments that directly interact with the GTPases. A similar role can be suggested for the BRs in the activation of the full-length effectors by GTPases. For WASP, autoinhibition involves the interaction of the GBD with the VCA domain,[34] leaving the BR disordered and free from intramolecular interaction [Fig. 1(b)]. As first suggested by Hemsath *et al.*,[20] the free BR may initiate the docking with Cdc42. Once docked, Cdc42 can displace the VCA domain and form additional interaction with GBD. For Pak1, Lei *et al.*[41] proposed that activation by Cdc42 begins with attack of the CRIB. We suggest here that an alternative pathway is initiated by docking with the free BR. Once docked, Cdc42 can peel away the CRIB and the rest of the GBD and thereby dissociate the autoinhibited dimer. These examples illustrate that elucidating binding pathways of IDPs can lead to better understanding of their biological functions.

Table II

Comparison of Calculated (Calc) and Experimental (Expt) Results for the Association Rate Constants (in μM^{-1} s^{-1}) of Wild-Type and Mutant Cdc42-Pak1$_{frag}$ Complexes

	Expt[a]	Calc
Wild-type	0.61	0.50
Cdc42 E49K	0.31	0.36
Cdc42 E49K/E178K	0.12	0.20

[a]Experimental results from Hemsath *et al.*[20]

For both WASP and Pak1, the docking step involving the BRs appears to dictate the (high) association rate constants. For Cdc42-WASP binding, the high association rate constant comes about both because of the strong electrostatic attraction between the BR and its subsite on Cdc42 and because of fast conformational coalescence of the GBD on the Cdc42 surface. The rate enhancement of electrostatic attraction is now well

understood,[46–48] and, as illustrated here, can be predicted by TransComp calculations. By comparing the calculated docking rate constant with the observed overall association rate constant, we can isolate the coalescing rate constant and begin to characterize the physical determinants of its magnitude, which surely would include the interactions and internal dynamics of the molecular elements involved in the coalescing step. In this regard, valuable information is being generated by kinetic studies involving systematic mutations of interface residues and variations of solvent conditions (for example, temperature).[22,24]

It will be of both fundamental and practical interest to redesign binding pathways, for example, by switching docking and coalescing segments. The WASP mutant with charge reversal of $K_{230}KK_{232}$ may be an example with an altered binding pathway (Table I): the resulting slowdown of the BR-initiated pathway potentially allows the CRIB-initiated pathway to become dominant. In general, guided by TransComp calculations, it may be relatively straightforward to introduce charge mutations to change the docking rate constants of two competing pathways in opposite directions, slowing down one and enhancing the other. Using systematic mutations, it may also be possible to tune the coalescing rate constants of the competing pathways in desired directions. Combinations of these mutations may finally lead to a redesigned binding pathway.

REFERENCES

1. Uversky VN. (Introduction to intrinsically disordered proteins (IDPs). Chem Rev 2014;114:6557–6560.
2. Oldfield CJ, Dunker AK. Intrinsically disordered proteins and intrinsically disordered protein Regions. Annu Rev Biochem 2014; 83:553–584.
3. Wright PE, Dyson HJ. Intrinsically disordered proteins in cellular signalling and regulation. Nat Rev Mol Cell Bio 2015;16:18–29.
4. Fisher CK, Stultz CM. Constructing ensembles for intrinsically disordered proteins. Curr Opin Struct Biol 2011;21:426–431.
5. Ozenne V, Bauer F, Salmon L, Huang JR, Jensen MR, Segard S, Bernado P, Charavay C, Blackledge M. Flexible-meccano: a tool for the generation of explicit ensemble descriptions of intrinsically disordered proteins and their associated experimental observables. Bioinformatics 2012;28:1463–1470.
6. Knott M, Best RB. A preformed binding interface in the unbound ensemble of an intrinsically disordered protein: evidence from molecular simulations. PLoS Comp. Biol 2012;8:e1002605
7. Jensen MR, Ruigrok RWH, Blackledge M. Describing intrinsically disordered proteins at atomic resolution by NMR. Curr Opin Struct Biol 2013;23:426–435.
8. Russo AA, Jeffrey PD, Patten AK, Massague J, Pavletich NP. (Crystal structure of the p27(Kip1) cyclin-dependent-kinase inhibitor bound to the cyclin A Cdk2 complex. Nature 1996;382:325–331.
9. Radhakrishnan I, PerezAlvarado GC, Parker D, Dyson HJ, Montminy MR, Wright PE. Solution structure of the KIX domain of CBP bound to the transactivation domain of CREB: A model for activator:Coactivator interactions. Cell 1997;91:741–752.
10. Abdul-Manan N, Aghazadeh B, Liu GA, Majumdar A, Ouerfelli O, Siminovitch KA, Rosen MK. Structure of Cdc42 in complex with the GTPase-binding domain of the 'Wiskott-Aldrich syndrome' protein. Nature 1999;399:379–383.
11. Dames SA, Martinez-Yamout M, De Guzman RN, Dyson HJ, Wright PE. Structural basis for Hif-1 alpha/CBP recognition in the cellular hypoxic response. Proc Natl Acad Sci USA 2002;99:5271–5276.
12. De Guzman RN, Goto NK, Dyson HJ, Wright PE. Structural basis for cooperative transcription factor binding to the CBP coactivator. J Mol Biol 2006;355:1005–1013.
13. Wells M, Tidow H, Rutherford TJ, Markwick P, Jensen MR, Mylonas E, Svergun DI, Blackledge M, Fersht AR. Structure of tumor suppressor p53 and its intrinsically disordered N-terminal transactivation domain. Proc Natl Acad Sci USA 2008;105:5762–5767.
14. Lee CW, Martinez-Yamout MA, Dyson HJ, Wright PE. Structure of the p53 transactivation domain in complex with the nuclear receptor coactivator binding domain of CREB binding protein. Biochemistry 2010;49:9964–9971.
15. Mukherjee SP, Behar M, Birnbaum HA, Hoffmann A, Wright PE, Ghosh G. Analysis of the RelA:CBP/p300 interaction reveals its involvement in NF-kappaB-driven transcription. PLoS Biol 2013;11: e1001647
16. Papadakos G, Sharma A, Lancaster LE, Bowen R, Kaminska R, Leech AP, Walker D, Redfield C, Kleanthous C. Consequences of inducing intrinsic disorder in a high-affinity protein-protein interaction. J Am Chem Soc 2015;137:5252–5255.
17. Lee CW, Ferreon JC, Ferreon ACM, Arai M, Wright PE. Graded enhancement of p53 binding to CREB-binding protein (CBP) by multisite phosphorylation. Proc Natl Acad Sci USA 2010;107:19290–19295.
18. Stone SR, Hofsteenge J. Kinetics of the inhibition of thrombin by hirudin. Biochemistry 1986;25:4622–4628.
19. Lacy ER, Filippov I, Lewis WS, Otieno S, Xiao LM, Weiss S, Hengst L, Kriwacki RW. (p27 binds cyclin-CDK complexes through a sequential mechanism involving binding-induced protein folding. Nat Struct Mol Biol 2004;11:358–364.
20. Hemsath L, Dvorsky R, Fiegen D, Carlier MF, Ahmadian MR. An electrostatic steering mechanism of Cdc42 recognition by Wiskott-Aldrich syndrome proteins. Mol Cell 2005;20:313–324.
21. Sugase K, Dyson HJ, Wright PE. Mechanism of coupled folding and binding of an intrinsically disordered protein. Nature 2007;447: 1021U1011.
22. Bachmann A, Wildemann D, Praetorius F, Fischer G, Kiefhaber T. Mapping backbone and side-chain interactions in the transition state of a coupled protein folding and binding reaction. Proc Natl Acad Sci USA 2011;108:3952–3957.
23. Dogan J, Schmidt T, Mu X, Engstrom A, Jemth P. Fast association and slow transitions in the interaction between two intrinsically disordered protein domains. J Biol Chem 2012;287:34316–34324.
24. Rogers JM, Oleinikovas V, Shammas SL, Wong CT, De Sancho D, Baker CM, Clarke J. Interplay between partner and ligand facilitates the folding and binding of an intrinsically disordered protein. Proc Natl Acad Sci USA 2014;111:15420–15425.
25. Lu Q, Lu HP, Wang J. Exploring the mechanism of flexible biomolecular recognition with single molecule dynamics. Phys Rev Lett 2007;98:128105
26. Turjanski AG, Gutkind JS, Best RB, Hummer G. Binding-induced folding of a natively unstructured transcription factor. PLoS Comp Biol 2008;4:e1000060
27. Huang YQ, Liu ZR. Kinetic advantage of intrinsically disordered proteins in coupled folding-binding process: A critical assessment of the "fly-casting" mechanism. J Mol Biol 2009;393:1143–1159.
28. Qin S, Pang XD, Zhou HX. Automated prediction of protein association rate constants. Structure 2011;19:1744–1751.
29. Pang XD, Zhou KH, Qin SB, Zhou HX. Prediction and dissection of widely-varying association rate constants of actin-binding proteins. PLoS Comp Biol 2012;8:e1002696
30. Ganguly D, Zhang W, Chen J. Electrostatically accelerated encounter and folding for facile recognition of intrinsically disordered proteins. PLoS Comp Biol 2013;9:e1003363

31. Kurcinski M, Kolinski A, Kmiecik S. Mechanism of folding and binding of an intrinsically disordered protein as revealed by ab initio simulations. J Chem Theory Comput 2014;10:2224–2231.
32. Xue Y, Yuwen T, Zhu F, Skrynnikov NR. Role of electrostatic interactions in binding of peptides and intrinsically disordered proteins to their folded targets. I. NMR and MD characterization of the complex between the c-Crk N-SH3 domain and the peptide Sos. Biochemistry 2014;53:6473–6495.
33. Zhou HX, Pang X, Lu C. Rate constants and mechanisms of intrinsically disordered proteins binding to structured targets. Pccp 2012;14:10466–10476.
34. Kim AS, Kakalis LT, Abdul-Manan M, Liu GA, Rosen MK. Autoinhibition and activation mechanisms of the Wiskott-Aldrich syndrome protein. Nature 2000;404:151–158.
35. Higgs HN, Pollard TD. Activation by Cdc42 and PIP2 of Wiskott-Aldrich syndrome protein (WASp) stimulates actin nucleation by Arp2/3 complex. J Cell Biol 2000;150:1311–1320.
36. Machesky LM, Insall RH. Scar1 and the related Wiskott-Aldrich syndrome protein, WASP, regulate the actin cytoskeleton through the Arp2/3 complex. Curr Biol 1998;8:1347–1356.
37. Snapper SB, Rosen FS. The Wiskott-Aldrich syndrome protein (WASP): Roles in signaling and cytoskeletal organization. Annu Rev Immunol 1999;17:905–929.
38. Badour K, Zhang JY, Siminovitch KA. The Wiskott-Aldrich syndrome protein: Forging the link between actin and cell activation. Immunol Rev 2003;192:98–112.
39. Weisswange I, Newsome TP, Schleich S, Way M. The rate of N-WASP exchange limits the extent of ARP2/3-complex-dependent actin-based motility. Nature 2009;458:87–91.
40. Thrasher AJ, Burns SO. WASP: A key immunological multitasker. Nat Rev Immunol 2010;10:182–192.
41. Lei M, Lu WG, Meng WY, Parrini MC, Eck MJ, Mayer BJ, Harrison SC. Structure of PAK1 in an autoinhibited conformation reveals a multistage activation switch. Cell 2000;102:387–397.
42. Edwards DC, Sanders LC, Bokoch GM, Gill GN. Activation of LIM-kinase by Pak1 couples Rac/Cdc42 GTPase signalling to actin cytoskeletal dynamics. Nat Cell Biol 1999;1:253–259.
43. Thompson G, Owen D, Chalk PA, Lowe PN. Delineation of the Cdc42/Rac-binding domain of p21-activated kinase. Biochemistry 1998;37:7885–7891.
44. Morreale A, Venkatesan M, Mott HR, Owen D, Nietlispach D, Lowe PN, Laue ED. Structure of Cdc42 bound to the GTPase binding domain of PAK. Nat Struct Biol 2000;7:384–388.
45. Neudauer CL, Joberty G, Tatsis N, Macara IG. Distinct cellular effects and interactions of the Rho-family GTPase TC10. Curr Biol 1998;8:1151–1160.
46. Gabdoulline RR, Wade RC. Simulation of the diffusional association of barnase and barstar. Biophys J 1997;72:1917–1929.
47. Schreiber G, Haran G, Zhou HX. Fundamental aspects of protein-protein association kinetics. Chem Rev 2009;109:839–860.
48. Alsallaq R, Zhou HX. Electrostatic rate enhancement and transient complex of protein-protein association. Proteins 2008;71:320–335.
49. Zhou HX. Enhancement of protein-protein association rate by interaction potential: Accuracy of prediction based on local Boltzmann factor. Biophys J 1997;73:2441–2445.
50. Holm L, Rosenstrom P. Dali server: Conservation mapping in 3D. Nucleic Acids Res 2010;38:W545–W549.
51. Eswar N, Webb B, Marti-Renom MA, Madhusudhan MS, Eramian D, Shen MY, Pieper U, Sali A. Comparative protein structure modeling using MODELLER. Curr Protoc Protein Sci 2007. Chapter 2, Unit 2.9.
52. Zhou HX. Kinetics of diffusion-influenced reactions studied by Brownian dynamics. J Phys Chem 1990;94:8794–8800.