Published in final edited form as:

J Mol Biol. 2010 March 12; 396(5): 1508–1523. doi:10.1016/j.jmb.2010.01.022.

Molecular dynamics reveal the essential role of linker motions in the function of cullin-RING E3 ligases

Jin Liu¹ and Ruth Nussinov^{1,2,*}

¹Basic Science Program, SAIC-Frederick, Inc., Center for Cancer Research Nanobiology Program, NCI-Frederick, Frederick, MD 21702

²Sackler Institute of Molecular Medicine, Department of Human Genetics and Molecular Medicine, Sackler School of Medicine, Tel Aviv University, Tel Aviv 69978, Israel

Abstract

Tagging proteins by polyubiquitin is a key step in protein degradation. Cullin-RING E3 ubiquitin ligases (CRLs) facilitate ubiquitin transfer from the E2 conjugating enzyme to the substrate; yet, crystallography indicates a large distance between the E2 and the substrate, raising the question of how this distance is bridged in the ubiquitin transfer reaction. Here, we demonstrate that the linker motions in the substrate binding proteins can allosterically shorten this distance to facilitate this crucial ubiquitin transfer step, and increase this distance to allow polyubiquitination. We performed molecular dynamics simulations for five substrate binding proteins, Skp2, Fbw7, β -TrCP1, Cdc4, and pVHL, in two forms: bound to their substrates, and bound to both substrate and adaptor. The adaptor connects the substrate binding proteins to the cullin. In the bound-to-both forms of all cases, we observed rotations of the substrate binding domain, shortening the gap between the tip of the substrate peptide and the E2 active site by 7–12 Å compared to the crystal structures. Overall, together with our earlier simulations of the unbound and the bound-to-adaptor forms, the emerging picture is that the maximum 51–73 Å distance between the substrate binding domain and the E2 active site in the modeled unbound forms of these five proteins shrinks to a minimum of 39–49 Å in the bound-to-both forms. This large distance range, the result of allosterically-controlled linker motions, facilitates the ubiquitin transfer and polyubiquitination, and as such argues that the cullin-RING E3 ubiquitin ligase is under conformational control. We further observed that substrate binding proteins with multiple substrate acceptor lysines have larger distance range between the substrate and the E2 as compared to β -TrCP1, with only one acceptor lysine.

Keywords

linker; E3 ubiquitin ligase; polyubiquitin; conformational control; conformational selection

INTRODUCTION

The Ubiquitin-Proteasome System (UPS) plays a crucial role in regulating many cellular processes¹. Ubiquitin attaches to substrate proteins via a covalent bond between the C-terminus

© 2009 Elsevier Ltd. All rights reserved.

Corresponding Author: Ruth Nussinov, ruthnu@helix.nih.gov.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

of ubiquitin and a lysine residue of the substrate. The substrate protein labeled with a polyubiquitin chain is recognized by the 26S proteasome where it undergoes degradation. The transfer of ubiquitin to substrate proteins is achieved through the sequential action of the three enzymes, ubiquitin-activating enzyme E1, ubiquitin-conjugating enzyme E2 and ubiquitin ligase E3. First, ubiquitin is covalently attached to E1; next it is transferred to E2. E3 ligase binds both the substrate protein and the E2-ubiquitin, bringing them into proximity to facilitate the ubiquitin transfer from E2 to the substrate^{2–4}.

There are two main classes of E3 ubiquitin ligases based on the mechanism through which they mediate ubiquitin transfer: HECT E3s and RING/U-Box E3s². RING/U-Box E3s have two sub-categories: simple RING E3s, which have both an E2-binding domain and a substrate binding domain on one single polypeptide, and multi-module cullin-RING Ligases (or CRLs). CRLs, as shown in Figure 1, resemble two-arm machines: the substrate is bound to one arm via the substrate binding protein which connects to its adaptor protein; and the ubiquitin is covalently attached to the other arm, via the E2 enzyme which is bound to the RING-Box protein (RBX). These two arms are connected by the cullin scaffold, which comprises two domains: the N-terminal domain (NTD) connects to the adaptor and the C-terminal domain (CTD) connects to RBX. Substrate binding proteins have two domains. One domain binds the adaptor protein via a conserved three helices “box” motif, such as the F-box (Skp2⁵, Fbw76, β -TrCP17, Cdc48, Fbs19, and TIR110), VHL-box (e.g. pVHL11) and SOCS-box (e.g. SOCS212 and SOCS413). The other domain is the substrate binding domain, which selectively binds different substrates via different motifs, such as leucine-rich repeats (LRRs) (Skp2), WD-40 repeats (Fbw7, β -TrCP1, or Cdc4), sugar binding domain (Fbs1), β domain (pVHL), or SH2 domain (SOCS2 or SOCS4). These two domains are connected by linkers. The sequence and structure of the linkers are not conserved. pVHL has a short loop linker with only ~5 residues, but F-box proteins have long structured linkers with 40~70 residues. The linker of Skp2 adopts the structure of three non-canonical LRRs, whereas Fbw7, β -TrCP1 and Cdc4 have α -helical linker with 2, 3, 4 α -helices, respectively, and different loop lengths in-between. The crystal structures of Skp2 bound to both adaptor and substrate are shown in Figure 1; the structures of Fbw7, β -TrCP1, Cdc4 and pVHL are shown in SI Figure 1.

It is generally believed that CRL brings the substrate and ubiquitin into proximity to facilitate the ubiquitin transfer. However, according to models piecing together the crystal structures of the CRL components, there is a 50~60 Å gap between the ubiquitin-E2 binding site and the tip of the substrate binding protein, as shown in Figure 1. For example, Zheng et al¹⁴ superimposed the Rbx1-Cul1-Skp1-F box on the Skp1-Skp2 complex¹⁵, and docked the UbcH7 E2 onto the Rbx1 RING domain¹⁶, to build a model of the SCF^{Skp2} (Skp-Cullin-F box protein, where the F-box protein is Skp2)-Rbx1-E2 complex. In their model, the distance between the E2 active site cysteine on one arm and the tip of Skp2 on the other is ~50 Å^{14; 17}. When the p27 substrate complex is included in the model by Cardozo and Pagano¹⁷, the distance between the E2 active site and the substrate binding site is 59 Å. Similarly, a 59 Å separation was also measured for β -Trcp1¹⁷.

This 50–60 Å gap raises the question how the CRL brings the substrate acceptor lysine and the Ub-E2 active site into proximity. Accumulating evidence indicates that flexible linkers exist in the different CRL components and these can serve as hinges in this molecular machine. Duda et al. reported a flexible linker for Rbx1¹⁸ on one arm of this machine; On the other arm, based on molecular dynamics simulations, we observed that the flexibility of the linkers may be an intrinsic feature of substrate binding proteins¹⁹. We simulated the nine substrate binding proteins with available crystal structures: F-box proteins Skp2, Fbw7, β -TrCP1, Cdc4, Fbs1, TIR1; VHL-box protein pVHL; and SOCS-box proteins SOCS2, and SOCS4. All nine have two domains connected by a linker: one binds the substrate; the other binds the CRL adaptors. We simulated the free and the adaptor-bound forms of these proteins. We observed that in the

free form, the flexible linker rotates the substrate binding domain by up to ~80 degrees, and the rotation driven by hydrophobic or electrostatic interactions between the two domains. Of particular interest, the movements are allosterically correlated with the binding sites to the adaptor and to the substrate. On the other hand, when the substrate binding proteins are bound to their corresponding adaptor, the maximal rotation angles were much smaller, up to ~40 degrees, still with the movements correlated with the binding sites. Thus, in the free form, the substrate-binding proteins exist in ensembles of conformational states. Through conformational selection^{20–25}, binding-favored conformers presenting such rotational states bind to their corresponding adaptor. During binding, the substrate binding domain adjusts its orientation, further shrinking the gap between the substrate and ubiquitin. Experimentally^{26–32} and computationally^{24; 33–35}, allosteric effects propagating through linkers are well established.

In our previous studies¹⁹, we only considered the conformational states of the substrate binding protein when unbound and when bound to its adaptor protein, and compared these two states. However, the picture we obtained was incomplete, because 1) the substrate was not considered thus not reflecting the *in vivo* scenario; 2) our previous study showed that linker flexibility is an intrinsic common feature of substrate binding proteins, but the role of linker motions in the function of CRLs remained unclear. Thus the key question of whether the gap between the substrate binding protein and the E2 active site is further shortened following the binding of both substrate and adaptor was unanswered. Here, in the *in vivo* scenario with both substrate and adaptor bound, our simulations reveal the essential role of linker motions in the function of CRLs: allosteric shortening of the distance gap by additional 7~12Å compared to the corresponding crystal structures. Interestingly, our results further suggest that the linker motions may be tuned by evolution: for the Skp2^{23; 36}, Fbw7³⁷, Cdc4³⁸ and pVHL³⁹ substrates, which have *multiple* ubiquitin acceptor lysines, the correlation between the motions of the linker and the substrate binding site in the substrate-bound form is weak, suggesting that the fluctuations of the linker can increase the probability that any of the acceptor lysines would be brought into proximity with ubiquitin. However, for β -TrCP1, with only one principal ubiquitin acceptor lysine⁷, the correlation between the linker and the substrate is strong.

RESULTS

Conformational changes of the substrate binding proteins before and after substrate binding

Four F-box substrate binding proteins, Skp2, Fbw7, β -TrCP1 and Cdc4, and VHL-box substrate binding protein pVHL have crystal structures of substrate-bound complexes. The bound-to-substrate crystal structures for SOCS-box proteins, SOCS2 and SOCS4, are not available, thus SOCS-box proteins are not included in this study. F-box proteins Fbs1 and TIR1 are also not included. Instead of recognizing phosphorylated or hydroxylated protein substrates, Fbs1 specifically recognizes N-linked glycoproteins, whereas TIR1 binds the substrate in concert with small molecules. These two proteins will be investigated in our future studies. Here, we performed molecular dynamics simulations for five proteins, Skp2, Fbw7, β -TrCP1, Cdc4 and pVHL, in two forms: bound only to the substrate (bound-to-substrate form), and bound to both substrate and adaptor (bound-to-both form). These are compared with the simulations of the unbound form and of the bound-to-adaptor form. All five simulated proteins have two domains, the box domain and the substrate binding domain, joined by a linker. The structure of the box domain is conserved and serves as an anchor to the adaptor. In simulations of both bound-to-substrate and bound-to-both forms, when the box domain is superimposed, the substrate binding domains rotate significantly with respect to their corresponding box domains. The results are shown in Figure 2 and Figure 3 and Table 1. The results for the unbound and the bound-to-adaptor forms are also included in Figure 3 and Table 1 for comparison.

In all of these cases, the unbound forms have significant rotation angles. The maximum rotation angles for the F-box proteins are between 53 to 58 degrees, whereas pVHL has an 80 degree maximum rotation angle for the unbound form. However, when a substrate binding protein binds to either an adaptor or a substrate, the rotation angles decrease significantly for Skp2, Fbw7, Cdc4, and pVHL. The bound-to-substrate forms have larger rotation angles than the bound-to-adaptor forms: The maximum rotation angles for the bound-to-substrate form are 36.7, 37.0, 30.4, and 56.9 degrees for Skp2, Fbw7, and Cdc4, and pVHL, respectively, which are 3–16 degrees more than the maximum rotation angles of the bound-to-adaptor form. However, in β -TrCP1, which has the most significant rotation among the bound-to-substrate forms, with a maximum rotation angle of 62.7 degrees and a mean value of 30.5 degrees, the rotation is even larger than in the unbound form. For all cases, including β -TrCP1, the form that has the smallest rotation angles is the bound-to-both form. The mean values of rotation angles for the bound-to-both form are 6.3, 13.3, 20.1, 14.7 and 8.5, for Skp2, Fbw7, β -TrCP1, Cdc4, and pVHL, respectively. The maximum rotation angles for the bound-to-both form, between 14 to 35 degrees, are the smallest among the four forms for each protein.

In short, for Skp2, Fbw7, Cdc4, and pVHL, the unbound form has the largest rotation angles. When these substrate binding proteins are bound to either the adaptor or the substrate, the rotation angles decrease, with the bound-to-adaptor form decreasing more than the bound-to-substrate form. When bound to both adaptor and substrate, these proteins have the smallest rotation angles. The conformational flexibility of the β -TrCP1 is different: Even though the unbound form rotates significantly, when bound to either adaptor or substrate, the mean value of the rotation angles still increases. For the bound-to-substrate form, both the maximum and mean rotation angles are the largest among all four forms. The bound-to-both form has the smallest rotation angles among these four forms.

Conformational changes shorten the gap between E2 and substrate binding protein

It is widely accepted that in the E3 CRL, there is a 50~60 Å gap between the substrate binding protein and the Ub-E2 active site cysteine. We built models of the E3 CRL machine by superimposing each snapshot from the F-box substrate binding proteins simulations on the Cul1-Rbx1-Skp1-F box complex, and docking the UbcH7 E2 onto the Rbx1 RING domain. We then measured the shortest distance between the substrate binding proteins and the E2 ubiquitin active site cysteine during the simulations. The results are shown in Figure 4 and Table 2.

For Skp2, Fbw7 and Cdc4, the unbound forms have the largest distance range, from a maximum distance of 66~73 Å to a minimum distance of 42~48 Å. When these proteins bind to either the adaptor or the substrate, the maximum distances between the substrate binding protein and E2 decrease to 50~64 Å, whereas the minimum distances remain in the 42~46 Å range. The bound-to-both forms have the smallest distance ranges, which are 10~18 Å smaller comparing to the unbound form. The maximum distances of the bound-to-both are 53~59 Å, and the minimum distance stays in the 41~49 Å range. Unlike the other three F-box proteins, β -TrCP1 does not show a significant difference between the four forms. The ranges are less than 16 Å for all four forms of β -TrCP1. Overall, for all four forms, the maximum distances are 51~56 Å, and the minimum distances are 36~41 Å.

Conformational changes shorten the gap between E2 and substrate

All the simulated substrate-bound crystal structures have only short peptides available. We measured the shortest distance between the substrate peptide and the E2 cysteine during the simulations of the bound-to-both form. The results are shown in Figure 5 and Table 3. For Skp2, Fbw7, and Cdc4, the distance ranges are 20~30 Å, from the maximum distances 64~77 Å to minimum distances 43~51 Å. Comparing to distances between E2 and the tip of the

substrate in the crystal complexes, which are 55~58 Å, the minimum distances during the simulations are shorter by 7~12 Å. β -TrCP1 behaves again differently from the other three F-box proteins. The distance range is only 13.8 Å, from a maximum of 47.1 Å to a minimum of 33.3 Å. In the modeled crystal complex the distance between the substrate and E2 is 43.5 Å, comparing to 55~58 Å of the other three F-box proteins. However, the minimum distance during the simulation still shrinks by 10 Å compared to the crystal complex. In short, in all cases, the gap between E2 and the substrate could be shortened by 7~12 Å compared to the crystal complex.

Linker correlations with the substrate binding domain are different for each form

Covariance maps were generated for the five proteins from the simulations of four forms: unbound, bound-to-adaptor, bound-to-substrate, and bound-to-both. Figure 6 and SI Figure 2 show the covariance maps for four forms of Skp2, Fbw7, β -TrCP1, Cdc4 and pVHL. In all five cases, the correlations between the linker and the substrate binding domain are much stronger in the bound-to-adaptor form than in the unbound form, which suggests that after the substrate binding proteins bind to their adaptor, the movement of the substrate binding domain is more coupled allosterically to the linker movement. However, interestingly, the cases differ in the bound-to-substrate form. In the Fbw7 (Figure 6B) and Cdc4 (Figure 6D), the bound-to-substrate form has a stronger correlation than the bound-to-adaptor form. For pVHL (Figure 6E), the correlations for the bound-to-adaptor and bound-to-substrate forms are similar. For Skp2 (Figure 6A) and β -TrCP1 (Figure 6C), the bound-to-substrate correlation is weaker than the bound-to-adaptor form.

More intriguingly, while the bound-to-both form for Skp2, Fbw7, Cdc4 and pVHL show a much lower correlation than either the bound-to-adaptor or the bound-to-substrate forms, with the correlations similar to the unbound form, the bound-to-both form for β -TrCP1 shows the strongest correlation between the linker and the substrate binding domain among the four forms.

Linker flexibilities can change upon binding

Linker flexibilities are assessed by root mean square fluctuations (RMSF) during the simulations, as shown in Figure 7. To our surprise, even though the rotation angles are quite different for Skp2 and Cdc4, the linker flexibilities are quite similar in all four forms. For Fbw7 and β -TrCP1, the linker flexibilities are different for different forms. For all proteins, the residues with the biggest RMSF value are the loop region within the linker. To further understand the linker flexibility, we calculated the residue relative stability constants with COREX/BEST algorithm^{40; 41} for the unbound and the bound-to-substrate forms, as shown in SI figure 3. Consistent with the RMSF results, the linker stabilities for these two forms are quite similar for Skp2 and Cdc4, but very different for Fbw7 and TrCP1, suggesting that binding to the substrate may allosterically affect the linker stability. Our previous studies¹⁹ showed that for the unbound form, there are two rotation hinges within each linker, which is at the N-terminus of the linker, right next to the F-box, and the other hinge is near the C-terminus helix next to the substrate binding domain. Here, for the bound-to-substrate and bound-to-both forms, the rotation hinge for Skp2 is at the N-terminus of the linker, which includes two conserved residues, Leu and Trp, but for Fbw7, β -TrCP1 and Cdc4, the rotation hinge is at the C-terminus helix next to the substrate binding domain, which includes two conserved residues, Asn and Trp. The positions of Trp are marked in Figure 7 and SI figure 3. The relative stabilities of Asn and Trp observed by COREX in Fbw7, β -TrCP1 and Cdc4 decrease allosterically by substrate binding events, as shown in SI Figure 3.

Conformational entropy change upon binding

The order parameter, S^2 , is used in NMR to measure motion on the ps-ns timescale^{42; 43}. S^2 values range from 1, when the NH bond vector has no internal motion, to 0, when the bond vector quickly samples multiple orientations. Order parameters were estimated for all four forms and were converted to conformational entropy change, as shown in SI Table 1. For Skp2, Fbw7 and Cdc4, the mean order parameters, $\langle S^2 \rangle$, are smallest for the unbound form. The order parameters increase for the bound-to-adaptor or the bound-to-substrate form, and are largest for the bound-to-both form, suggesting a global rigidification in bound forms, with conformational entropy penalty $-T\Delta S$ ranging from 0.12 to 1.11 kcal/mol for all bound forms. β -TrCP1 is again different. The $\langle S^2 \rangle$ value is very high for the unbound form. Instead of conformational entropy penalty, the bound-to-substrate and bound-to-both forms have a favorable entropy change, with $-T\Delta S \sim -0.45$ and -0.09 kcal/mol, respectively, suggesting a favorable conformational entropy change upon binding to the substrate.

DISCUSSION

E3 CRL presents a mechanistic enigma: how is the large distance gap between E2 and the substrate binding protein bridged for the ubiquitin transfer from E2 to the substrate? When assessed based on the unbound form of the substrate binding proteins, the distances between the substrate and E2 fluctuate between 36~73 Å. However, when measured in the bound-to-the-adaptor form, the distance fluctuation shrinks to 41~61 Å. This large fluctuation in the distance derives from flexible linkers between the substrate binding domain and the box domain. In all available structures of the substrate binding proteins, F-box proteins Skp2, Fbw7, β -TrCP1, Cdc4, Fbs1, TIR1; VHL-box protein pVHL; and SOCS-box proteins SOCS2, and SOCS4, the flexible linkers rotate during the simulations by up to 30–80°. In particular, movements of the linker were observed to correlate allosterically with the binding sites to the adaptor and the substrate. Unfortunately, not all of these substrate binding proteins have available structures bound to their cognate substrate peptides. We performed molecular dynamics simulations for five substrate binding proteins, Skp2, Fbw7, β -TrCP1, Cdc4 and pVHL with available bound-to-substrate structures. The simulations were for two forms: the first is the bound-to-substrate form; the second, the bound-to-both form, with both the substrate and the adaptor bound to the different domains of the substrate binding protein. We compared these simulation results with our simulations for the unbound and bound-to-adaptor forms⁴⁴. Our goal is to elucidate the molecular mechanism of the ubiquitin transfer reaction, given the large distance gap between the E2-ubiquitin and the substrate binding domain. The large conformational change in the substrate binding protein (Figure 8), leads us to suspect that the E3 CRL is under conformational control.

Inspection of the distance gap in the bound-to-both form suggests that when the substrate binding protein binds both the adaptor protein and the substrate, the distance gap between the substrate binding protein and E2 can be shortened to 39~49 Å, compared to 50~60 Å in the crystal structure. The mean distances for the unbound state are significantly larger than those for the bound form, and the mean distances for all three bound forms are similar, suggesting that linker motions can shorten the distance in the bound state, regardless of whether the substrate binding protein is bound to the adaptor, substrate or both. To what extent can the distance be shortened? Even though all bound states have similar mean distances, their significantly different minimum, maximum and range of distances, as shown in Table 2, suggest differences in the distributions of the populations of the states, thus the likelihood of CRL reaching the shortest and longest distances.

The distances between the substrate peptide and E2 could be shortened from a maximum of 47~76 Å to a minimum of 33~51 Å (Figure 8), which is 7~12 Å shorter than the distance in the crystal structure. The substrates occupy certain space, easily filling this gap. For example,

during the simulations of the bound-to-both form of β -TrCP1, the distances between the tip of the substrate β -catenin and the E2 cysteine are in the range of 33~47 Å. Because only a fraction of β -catenin was crystallized, the acceptor lysine of β -catenin, which is not included in the crystal structure and 11 residues away from the tip of the crystallized β -catenin peptide, could reach up to 40 Å away from the tip of the crystallized β -catenin, which makes the contact of the substrate acceptor lysine and the E2 active site possible. The maximum distances, which are 14~29 Å larger than the minimum distances, may help the polyubiquitination process. Considering the ~20 Å distance between the ubiquitin Lys48 and Gly76, which connect consecutive ubiquitins in the polyubiquitin chain, the 14~29 Å difference between the minimum and maximum distances could accommodate ligated ubiquitins. The proposed mechanism is shown in Figure 8.

Unlike the substrates in the other four substrate binding proteins which have multiple acceptor lysines, β -catenin of β -TrCP1 has been reported to have only one primary acceptor lysine site. It was also reported that the β -catenin lysine-destruction motif spacing is a determinant of ubiquitination efficiency⁷. In addition, our previous study¹⁹ showed that when superimposing the box domain, the substrate binding sites of Skp2, Fbw7 and Cdc4 overlap each other, whereas β -TrCP1 substrate binding sites are more than 10 Å away from those of other proteins, implying a different mechanism for β -TrCP1. Nonetheless, even though β -TrCP1 has large rotation angles during the simulation and the distance range between the substrate and the E2 cysteine is significantly smaller than in other proteins, linker motions of β -TrCP1 still control the conformational change. This rotation could conceivably lead to an increase in the concentration of the substrate acceptor lysine near the E2 active site. This is consistent with the lysine spacing specificity. For Skp2, Fbw7 and Cdc4, the distance ranges between the substrate and E2 cysteine are much larger, which can provide sufficient space for the multiple acceptor lysines.

Our analysis of correlated movements supports linker motions as a key player in controlling the positioning of the substrate binding domains and thus in orienting the substrate acceptor lysines. Previously¹⁹ we observed a strong correlation between the movements of the linker and the binding sites on the two domains, particularly in the adaptor-bound form. Here, comparing the correlations between the movements of the linker and the substrate binding site in the bound-to-both forms versus other forms, β -TrCP1 has the strongest correlation among the five substrate binding proteins. This argues that when β -TrCP1 binds to both substrate and adaptor, the movements of the linker and the substrate binding domain are strongly correlated, leading to an accurate positioning of the substrate acceptor lysine for ubiquitin transfer. For Skp2, Fbw7, Cdc4 and pVHL, the movements of the linker and substrate binding sites are less correlated which suggest possible increase in the concentration of any of the potential acceptor lysines near the E2 active site.

For Skp2, Fbw7, Cdc4 and pVHL, the unbound form displays the largest linker motions. Linker motions in the bound-to-adaptor and bound-to-substrate forms decrease. However, for β -TrCP1, the form with the largest motions is not the unbound but the bound-to-substrate, suggesting that substrate binding, rather than decreasing the linker motions, allosterically enhances it²². We calculated the conformational entropy change upon binding. As expected, unlike Skp2, Fbw7 and Cdc4 which sustain conformational entropy penalties upon binding to the substrate, β -TrCP1 has a favorable conformational entropy change, suggesting that for β -TrCP1 substrate binding may be an entropy-driven process.

In the hinge regions, near the N- and C-termini of the linker, respectively, Leu-Trp and Asn-Trp are conserved in F-box proteins; Leu and Asn may constitute linker rotation positions and Trp may serve as an anchor, stabilizing the rotation in certain angles. In the unbound state, the N-terminus hinge is the major hinge; but when bound to the substrate, the C-terminus hinge

residues Asn-Trp are destabilized, serving as the major hinge. Based on these observations, we speculate that mutating Leu or Asn to Pro, or binding of small molecules to Trp, may affect the function of CRLs.

Additional linkers in other E3 components may also play key roles in facilitating the ubiquitin transfer reactions by bridging the distances between the arms and at the same time accommodating the space-demanding polyubiquitin chain. The E3 CRL has four components, the substrate binding protein, the adaptor protein, cullin, and the RING-box protein. Recently, based on two crystal structure conformers, Duda et al observed that the linker in the RING-box protein Rbx1¹⁸ is flexible, which suggests that this linker may also contribute to the E3 CRL conformational change during ubiquitination. We propose that CRL works like a two-arm machine. One arm is the Rbx1 whose flexibility could be further enhanced by Nedd8, and the other arm is the substrate binding protein, which is intrinsically flexible. Our simulations indicate that Nedd8 can allosterically affect the Rbx1 arm shortening the 50 Å gap by ~40 Å (data not shown). Here we show that the substrate binding protein arm can shorten the gap by ~10 Å. Thus these two arms may work cooperatively to reduce the distance gap between the substrate and the E2's cysteine.

We expect additional flexible links in other E3 CRL components, i.e. the adaptor and cullin. In crystal structures of the adaptor proteins Skp1, Elongin C and ASK1, there are disordered loops in the binding sites suggesting flexibility in the binding site of the adaptor to the substrate binding protein and the adaptor binding site to cullin. These flexible loops could act in concert with flexible linkers in substrate binding proteins and in the RING-box proteins. Furthermore, cullin is comprised of an N-terminal domain (NTD), which binds to the adaptor, and a C-terminal domain (CTD), which binds to the RING-box protein (Figure 1, Figure 8). While it has been reported that the introduction of an artificial linker to join the NTD and CTD of cullin can disrupt ubiquitination¹⁴, it does not exclude the possibility of flexible linkers in cullin; however, it does suggest a cooperative, allostery-driven E3 CRL. There are three repeats in NTD and these repeats are connected by linkers and the NTD and CTD are also connected by a linker. Cullin serves as a scaffold for the E3 CRL; as such, we suspect that even a small conformational change in a cullin linker may induce a large distance change between the substrate and the E2 ubiquitin binding sites.

In conclusion, our simulations of substrate binding proteins provide a mechanism for the E3 CRL ubiquitin transfer: the linker cooperatively rotates the substrate binding domain and substrate, shortening their distances to the E2 active site, or increasing their distance allowing polyubiquitination. As such, the linker could be a new therapeutic target for E3 CRL-related human diseases. Allostery is a key mechanism in protein function: it is a cooperative event, up-or down-regulating protein activities^{22; 45–48}. In our case here, the linker can be targeted directly or allosterically^{34; 49}. Point mutations or drug design targeting alteration of the linker motions may affect the overall function of CRLs.

METHODS

System setup

The starting structures of the bound-to-substrate and bound-to-both forms of five proteins, Skp2, Fbw7, β -TrCP1, Cdc4, pVHL were constructed from crystal structures (PDB codes: 2ast⁵, 2ovq6, 1p227, 1nex8, 1lm811). For the bound-to-substrate form, the adaptor, i.e. Skp1 or Elongin C, was removed. For the bound-to-both form, both the adaptor and the substrate were kept for simulations. For the pVHL adaptor Elongin C, since coordinates are not available between residues 49 and 58, and the solved Elongin C structure for residues 17 to 49 is far away from the pVHL-Elongin C binding site, only residues 58–112 were used in the simulations. For the F-box proteins binding protein adaptor Skp1, only residues 1084 to 1159

were used for a similar reason. The missing residues in β -TrCP1 and Cdc4 were added as random coils and were minimized for 1000 steps with the rest of the protein fixed. All models were solvated in TIP3P water box. A minimum distance from the edge of the box to any protein atom was 10 Å. The chloride or sodium ions were added to neutralize the systems.

Simulation protocol

Molecular dynamics simulations were performed with CHARMM 27⁵⁰ force field using the NAMD program⁵¹. To eliminate residual unfavorable interactions between the solvent and the protein, the solvated systems were first minimized for 3000 steps with the protein backbone fixed, followed by 3000 steps of minimization without any constraint. Then the systems were heated from 0K to 300K in 100 ps and equilibrated for 100 ps with backbone constrained, followed by 500 ps equilibration without any constraint. Production simulations were performed for 20 ns with the NPT ensemble at 300K and 1 bar. The temperature and pressure were controlled with Langevin thermostat and Nose-Hoover Langevin piston barostat, respectively. The short-range interactions employed a switch function with a cutoff of 12 Å and a switch distance of 10 Å. Particle mesh Ewald summation was used to treat the long-range electrostatic interactions. A SHAKE constraint on all bonds containing hydrogen atoms was used during the production simulations. The time step was 2 fs. Structural alignments and molecular graphics were produced with VMD⁵². The angle rotations analysis during the simulation were performed with Hingefind⁵³. The model for measurement of the distance change during the simulation was built based on the complex structure Rbx1-Cul1-Skp1-Skp2^{F box} (PDB code: 1ldk14). E2 UbcH7 (PDB code: 1fbv16) was docked into the Rbx1 RING subdomain, and snapshots of F-box of substrate binding proteins from the simulation trajectories were superimposed with Skp2^{F box}. The shortest distances were measured between E2 cysteine and the substrate binding protein (Table 2) or substrate (Table 3) for each snapshot. The distances for the crystal structure were obtained when crystal structures of substrate binding proteins, instead the snapshots from simulation trajectories, were included in the model.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This project has been funded in whole or in part with Federal funds from the National Cancer Institute, National Institutes of Health, under contract number HHSN261200800001E. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government. This research was supported (in part) by the Intramural Research Program of the NIH, National Cancer Institute, Center for Cancer Research. This study used the high-performance computational capabilities of the Biowulf Linux cluster at the National Institutes of Health, Bethesda, MD (<http://biowulf.nih.gov>).

REFERENCES

1. Hershko A, Ciechanover A. The ubiquitin system. *Annu Rev Biochem* 1998;67:425–479. [PubMed: 9759494]
2. Capili AD, Lima CD. Taking it step by step: mechanistic insights from structural studies of ubiquitin/ubiquitin-like protein modification pathways. *Curr Opin Struct Biol* 2007;17:726–735. [PubMed: 17919899]
3. Herrmann J, Lerman LO, Lerman A. Ubiquitin and ubiquitin-like proteins in protein regulation. *Circulation research* 2007;100:1276–1291. [PubMed: 17495234]
4. Nalepa G, Rolfe M, Harper JW. Drug discovery in the ubiquitin-proteasome system. *Nat Rev Drug Discov* 2006;5:596–613. [PubMed: 16816840]

5. Hao B, Zheng N, Schulman BA, Wu G, Miller JJ, Pagano M, Pavletich NP. Structural basis of the Cks1-dependent recognition of p27(Kip1) by the SCF(Skp2) ubiquitin ligase. *Mol Cell* 2005;20:9–19. [PubMed: 16209941]
6. Hao B, Oehlmann S, Sowa ME, Harper JW, Pavletich NP. Structure of a Fbw7-Skp1-cyclin E complex: multisite-phosphorylated substrate recognition by SCF ubiquitin ligases. *Mol Cell* 2007;26:131–143. [PubMed: 17434132]
7. Wu G, Xu G, Schulman BA, Jeffrey PD, Harper JW, Pavletich NP. Structure of a beta-TrCP1-Skp1-beta-catenin complex: destruction motif binding and lysine specificity of the SCF(beta-TrCP1) ubiquitin ligase. *Mol Cell* 2003;11:1445–1456. [PubMed: 12820959]
8. Orlicky S, Tang X, Willems A, Tyers M, Sicheri F. Structural basis for phosphodependent substrate selection and orientation by the SCFCdc4 ubiquitin ligase. *Cell* 2003;112:243–256. [PubMed: 12553912]
9. Mizushima T, Yoshida Y, Kumanomidou T, Hasegawa Y, Suzuki A, Yamane T, Tanaka K. Structural basis for the selection of glycosylated substrates by SCF(Fbs1) ubiquitin ligase. *Proc Natl Acad Sci U S A* 2007;104:5777–5781. [PubMed: 17389369]
10. Tan X, Calderon-Villalobos LI, Sharon M, Zheng C, Robinson CV, Estelle M, Zheng N. Mechanism of auxin perception by the TIR1 ubiquitin ligase. *Nature* 2007;446:640–645. [PubMed: 17410169]
11. Min JH, Yang H, Ivan M, Gertler F, Kaelin WG Jr, Pavletich NP. Structure of an HIF-1alpha-pVHL complex: hydroxyproline recognition in signaling. *Science* 2002;296:1886–1889. [PubMed: 12004076]
12. Bullock AN, Debreczeni JE, Edwards AM, Sundstrom M, Knapp S. Crystal structure of the SOCS2-elongin C-elongin B complex defines a prototypical SOCS box ubiquitin ligase. *Proc Natl Acad Sci U S A* 2006;103:7637–7642. [PubMed: 16675548]
13. Bullock AN, Rodriguez MC, Debreczeni JE, Songyang Z, Knapp S. Structure of the SOCS4-ElonginB/C complex reveals a distinct SOCS box interface and the molecular basis for SOCS-dependent EGFR degradation. *Structure* 2007;15:1493–1504. [PubMed: 17997974]
14. Zheng N, Schulman BA, Song L, Miller JJ, Jeffrey PD, Wang P, Chu C, Koepp DM, Elledge SJ, Pagano M, Conaway RC, Conaway JW, Harper JW, Pavletich NP. Structure of the Cul1-Rbx1-Skp1-F boxSkp2 SCF ubiquitin ligase complex. *Nature* 2002;416:703–709. [PubMed: 11961546]
15. Schulman BA, Carrano AC, Jeffrey PD, Bowen Z, Kinnucan ER, Finnin MS, Elledge SJ, Harper JW, Pagano M, Pavletich NP. Insights into SCF ubiquitin ligases from the structure of the Skp1-Skp2 complex. *Nature* 2000;408:381–386. [PubMed: 11099048]
16. Zheng N, Wang P, Jeffrey PD, Pavletich NP. Structure of a c-Cbl-UbcH7 complex: RING domain function in ubiquitin-protein ligases. *Cell* 2000;102:533–539. [PubMed: 10966114]
17. Cardozo T, Pagano M. The SCF ubiquitin ligase: insights into a molecular machine. *Nature reviews. Molecular cell biology* 2004;5:739–751.
18. Duda DM, Borg LA, Scott DC, Hunt HW, Hammel M, Schulman BA. Structural insights into NEDD8 activation of cullin-RING ligases: conformational control of conjugation. *Cell* 2008;134:995–1006. [PubMed: 18805092]
19. Liu J, Nussinov R. The mechanism of ubiquitination in the cullin-RING E3 ligase machinery: conformational control of substrate orientation. *PLoS Comput Biol* 2009;5 e1000527.
20. Kumar S, Ma B, Tsai CJ, Sinha N, Nussinov R. Folding and binding cascades: dynamic landscapes and population shifts. *Protein Sci* 2000;9:10–19. [PubMed: 10739242]
21. Boehr D, Nussinov R, Wright P. The role of dynamic conformational ensembles in biomolecular recognition. *Nat Chem Biol*.
22. del Sol A, Tsai CJ, Ma B, Nussinov R. The Origin of Allosteric Functional Modulation: Multiple Pre-existing Pathways. *Structure*. 2009
23. Shirane M, Harumiya Y, Ishida N, Hirai A, Miyamoto C, Hatakeyama S, Nakayama K, Kitagawa M. Down-regulation of p27(Kip1) by two mechanisms, ubiquitin-mediated degradation and proteolytic processing. *J Biol Chem* 1999;274:13886–13893. [PubMed: 10318797]
24. Swint-Kruse L, Larson C, Pettitt BM, Matthews KS. Fine-tuning function: correlation of hinge domain interactions with functional distinctions between LacI and PurR. *Protein Sci* 2002;11:778–794. [PubMed: 11910022]

25. Lange OF, Lakomek NA, Fares C, Schroder GF, Walter KF, Becker S, Meiler J, Grubmuller H, Griesinger C, de Groot BL. Recognition dynamics up to microseconds revealed from an RDC-derived ubiquitin ensemble in solution. *Science* 2008;320:1471–1475. [PubMed: 18556554]
26. Sondermann H, Nagar B, Bar-Sagi D, Kuriyan J. Computational docking and solution x-ray scattering predict a membrane-interacting role for the histone domain of the Ras activator son of sevenless. *Proc Natl Acad Sci U S A* 2005;102:16632–16637. [PubMed: 16267129]
27. Vogel M, Mayer MP, Bukau B. Allosteric regulation of Hsp70 chaperones involves a conserved interdomain linker. *J Biol Chem* 2006;281:38705–38711. [PubMed: 17052976]
28. Tribble RP, Emert-Sedlak L, Smithgall TE. HIV-1 Nef selectively activates Src family kinases Hck, Lyn, and c-Src through direct SH3 domain interaction. *J Biol Chem* 2006;281:27029–27038. [PubMed: 16849330]
29. Swain JF, Dinler G, Sivendran R, Montgomery DL, Stotz M, Gierasch LM. Hsp70 chaperone ligands control domain association via an allosteric mechanism mediated by the interdomain linker. *Mol Cell* 2007;26:27–39. [PubMed: 17434124]
30. McFadden K, Cocklin S, Gopi H, Baxter S, Ajith S, Mahmood N, Shattock R, Chaiken I. A recombinant allosteric lectin antagonist of HIV-1 envelope gp120 interactions. *Proteins* 2007;67:617–629. [PubMed: 17348010]
31. Zhan H, Taraban M, Trehwella J, Swint-Kruse L. Subdividing repressor function: DNA binding affinity, selectivity, and allostery can be altered by amino acid substitution of nonconserved residues in a LacI/GalR homologue. *Biochemistry* 2008;47:8058–8069. [PubMed: 18616293]
32. Tungtur S, Egan SM, Swint-Kruse L. Functional consequences of exchanging domains between LacI and PurR are mediated by the intervening linker sequence. *Proteins* 2007;68:375–388. [PubMed: 17436321]
33. Banavali NK, Roux B. Anatomy of a structural pathway for activation of the catalytic domain of Src kinase Hck. *Proteins* 2007;67:1096–1112. [PubMed: 17380483]
34. Liu J, Nussinov R. Allosteric effects in the marginally stable von Hippel-Lindau tumor suppressor protein and allostery-based rescue mutant design. *Proc Natl Acad Sci U S A* 2008;105:901–906. [PubMed: 18195360]
35. Sotomayor M, Schulten K. The allosteric role of the Ca²⁺ switch in adhesion and elasticity of C-cadherin. *Biophys J* 2008;94:4621–4633. [PubMed: 18326636]
36. Ryu KS, Choi YS, Ko J, Kim SO, Kim HJ, Cheong HK, Jeon YH, Choi BS, Cheong C. Direct characterization of E2-dependent target specificity and processivity using an artificial p27-linker-E2 ubiquitination system. *BMB reports* 2008;41:852–857. [PubMed: 19123975]
37. Welcker M, Clurman BE. FBW7 ubiquitin ligase: a tumour suppressor at the crossroads of cell division, growth and differentiation. *Nat Rev Cancer* 2008;8:83–93. [PubMed: 18094723]
38. Petroski MD, Deshaies RJ. Context of multiubiquitin chain attachment influences the rate of Sic1 degradation. *Mol Cell* 2003;11:1435–1444. [PubMed: 12820958]
39. Paltoglou S, Roberts BJ. HIF-1 α and EPAS ubiquitination mediated by the VHL tumour suppressor involves flexibility in the ubiquitination mechanism, similar to other RING E3 ligases. *Oncogene* 2007;26:604–609. [PubMed: 16862177]
40. Vertrees J, Barritt P, Whitten S, Hilser VJ. COREX/BEST server: a web browser-based program that calculates regional stability variations within protein structures. *Bioinformatics* 2005;21:3318–3319. [PubMed: 15923205]
41. Hilser VJ, Garcia-Moreno EB, Oas TG, Kapp G, Whitten ST. A statistical thermodynamic model of the protein ensemble. *Chem Rev* 2006;106:1545–1558. [PubMed: 16683744]
42. Forman-Kay JD. The 'dynamics' in the thermodynamics of binding. *Nat Struct Biol* 1999;6:1086–1087. [PubMed: 10581541]
43. Case DA. Molecular dynamics and NMR spin relaxation in proteins. *Acc Chem Res* 2002;35:325–331. [PubMed: 12069616]
44. Liu J, Nussinov R. The mechanism of ubiquitination in the cullin-RING E3 ligase machinery: Conformational control of substrate orientation. *PLoS Comp Bio*.
45. Horovitz A, Willison KR. Allosteric regulation of chaperonins. *Curr Opin Struct Biol* 2005;15:646–651. [PubMed: 16249079]

46. Kafri G, Horovitz A. Transient kinetic analysis of ATP-induced allosteric transitions in the eukaryotic chaperonin containing TCP-1. *J Mol Biol* 2003;326:981–987. [PubMed: 12589746]
47. Gunasekaran K, Ma B, Nussinov R. Is allostery an intrinsic property of all dynamic proteins? *Proteins* 2004;57:433–443. [PubMed: 15382234]
48. Leitner DM. Energy flow in proteins. *Annu Rev Phys Chem* 2008;59:233–259. [PubMed: 18393676]
49. Goodey NM, Benkovic SJ. Allosteric regulation and catalysis emerge via a common route. *Nature chemical biology* 2008;4:474–482.
50. MacKerell AD Jr, Bashford D, Bellott RL, Dunbrack RL Jr, Evanseck JD, Field MJ, Fischer S, Gao J, Guo H, Ha S, Joseph-McCarthy D, Kuchnir L, Kuczera K, Lau FTK, Mattos C, Michnick S, Ngo T, Nguyen DT, Prodhom B, Reiher WE III, Roux B, Schlenkrich M, Smith JC, Stote R, Straub J, Watanabe M, Wiorkiewicz-Kuczera J, Yin D, Karplus M. All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *J. Phys. Chem. B* 1998;102:3586–3616.
51. Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel RD, Kale L, Schulten K. Scalable molecular dynamics with NAMD. *J Comput Chem* 2005;26:1781–1802. [PubMed: 16222654]
52. Humphrey W, Dalke A, Schulten K. VMD: visual molecular dynamics. *J Mol Graph* 1996;14:33–38. 272–8. [PubMed: 8744570]
53. Wriggers W, Schulten K. Protein domain movements: detection of rigid domains and visualization of hinges in comparisons of atomic coordinates. *Proteins* 1997;29:1–14. [PubMed: 9294863]

Figure 1. Overview of the system. (A) A schematic illustration of the Ub-E2-CRL-substrate machinery. (B) Crystal structure of Skp2 bound to both adaptor (blue) and substrate (purple) (PDB code 2ast). Skp2 F-box domain is shown in red, linker in orange, and substrate binding domain in cyan. SBP stands for substrate binding protein.

Figure 2.

Comparison of snapshots from the simulations of (A) Skp2; (B) Fbw7; (C) β -TrCP1; (D) Cdc4; (E) pVHL. The bound-to-substrate form (left) and bound-to-both form (right) comparison is shown in each figure. Structural snapshots of the F-box or VHL-box domains are superimposed. The snapshots are taken at 0 ns (orange for substrate binding proteins, purple for substrate and blue for adaptor Skp1 or Elongin C) and at the time of maximum rotation (green). The rotation angles of the substrate binding domain are shown.

Figure 3.

Rotation angles of four forms for (A) Skp2; (B) Fbw7; (C) β -TrCP1; (D) Cdc4; (E) pVHL during the simulations. The four forms shown are unbound form (green line), bound-to-adaptor form (blue line), bound-to-substrate form (black line), and bound-to-both form (red line).

Figure 4.

The shortest distances between the substrate binding protein and the E2 active cysteine of four forms for (A) Skp2; (B) Fbw7; (C) β -TrCP1; (D) Cdc4; (E) pVHL during the simulations. The four forms shown are the unbound form (green line), bound-to-adaptor form (blue line), bound-to-substrate form (black line), and bound-to-both form (red line). adp stands for adaptor; sub for substrate.

Figure 5.

Distances between the tips of the substrates and the E2 active cysteine for the bound-to-both forms of substrate binding proteins. The model of E3 CRL complex with the substrate and E2 are shown on the left, and the distances during the simulations are shown on the right. E2-Rbx1-Cul1-Skp1 is modeled with (A) Skp2 and p27 (B) Fbw7 and Cyclin E (C) β -TrCP1 and β -catenin (D) Cdc4 and CPD. E2 (purple, PDB code 1fbv) is docked to Rbx1 (gray)-Cul1 (blue)-Skp1 (red)-Skp2 F-box (yellow) complex (PDB code 1ldk). Substrate binding proteins and substrates snapshots at longest (orange) and shortest distances (green) for the bound-to-both forms are superimposed on the crystal structure using the F-box domain as the pivot.

Figure 6.

Covariance maps of (i) unbound (ii) bound-to-adaptor (iii) bound-to-substrate (iv) bound-to-both form of (A) Skp2; (B) Fbw7; (C) β -TrCP1; (D) Cdc4; (E) pVHL. The more red, the stronger the positive correlation; the more blue the stronger the negative (anti-) correlation. The bar provides the scale.

Figure 7.

α RMSFs relative to the average structures of the trajectories for the linker. (A) Skp2; (B) Fbw7; (C) β -TrCP1; (D) Cdc4. The position for the rotation hinge Trp is marked. Secondary structures are shown at the top.

Figure 8.

A scheme showing that the rotation of the substrate binding domain (SBD) shortens the distance gap between the substrate acceptor lysine and the Ub-E2 active site. The minimum (left) and maximum distances (right) between the tip of the substrate peptide and Ub-E2 active site during the bound-to-both simulations are shown in the scheme. The unstructured part which includes the acceptor lysine(s) is shown with dashed lines.

Table 1

Rotation angles (degrees) for substrate binding proteins.

		Skp2	Fbw7	TrCP1	Cdc4	pVHL
Unbound	Max	53	56.2	53.3	58	80.2
	Mean	28.5	24.5	21.2	24.5	37.5
Bound-to-adaptor	Max	27.9	32.8	40.8	27	39.7
	Mean	10.2	15.8	25.1	11.8	22
Bound-to-substrate	Max	36.7	37	62.7	30.3	56.9
	Mean	16	16.3	30.5	14.9	29.3
Bound-to-both	Max	14.7	25	34.6	31.4	19.1
	Mean	6.3	13.3	20.1	14.7	8.5

Table 2

The shortest distances between substrate binding proteins and E2 cysteine.

	Skp2	Fbw7	TrCP1	Cdc4
Unbound	Max	73.1	51.3	65.7
	Min	41.9	35.9	44.4
	Mean	55.4	40.8	56.2
	Range	31.2	15.4	21.4
Bound-to-adaptor	Max	50.5	55.7	58.2
	Min	42.4	40.5	44.8
	Mean	46.3	48.7	51.3
	Range	8.1	15.2	13.3
Bound-to-substrate	Max	63.6	53.3	59
	Min	44.1	37.7	45.6
	Mean	49.7	44.7	52.1
	Range	19.6	15.6	13.5
Bound-to-both	Max	55.8	51.2	59.4
	Min	42.7	39.4	48.7
	Mean	46.9	44.7	53.8
	Range	13.4	11.8	10.7

Table 3

The shortest distance between tip of substrates and E2 cysteine for bound-to-both form.

	Skp2	Fbw7	TrCP1	Cdc4
Max	75.8	63.7	47.1	76.9
Min	46.8	43.7	33.3	50.7
crystal	57	55.2	43.5	57.6
Mean	60.5	52.3	38.9	62.9
Range	29	20	13.8	26.1