Human Proteome-scale Structural Modeling of E2–E3 Interactions Exploiting Interface Motifs

Gozde Kar,[†] Ozlem Keskin,^{*,†} Ruth Nussinov,^{‡,§} and Attila Gursoy^{*,†}[†]Center for Computational Biology and Bioinformatics and College of Engineering, Koc University, Rumelifeneri Yolu, 34450 Sariyer Istanbul, Turkey[‡]Center for Cancer Research Nanobiology Program, National Cancer Institute-Frederick, Frederick, Maryland 21702, United States[§]Sackler Institute of Molecular Medicine, Sackler School of Medicine, Tel Aviv University, Tel Aviv 69978, Israel

S Supporting Information

ABSTRACT: Ubiquitination is crucial for many cellular processes such as protein degradation, DNA repair, transcription regulation, and cell signaling. Ubiquitin attachment takes place via a sequential enzymatic cascade involving ubiquitin activation (by E1 enzymes), ubiquitin conjugation (by E2 enzymes), and ubiquitin substrate tagging (by E3 enzymes). E3 ligases mediate ubiquitin transfer from E2s to substrates and as such confer substrate specificity. Although E3s can interact and function with numerous E2s, it is still unclear how they choose which E2 to use. Identifying all E2 partners of an E3 is essential for inferring the principles guiding E2 selection by an E3. Here we model the interactions of E3 and E2 proteins in a large, proteome-scale strategy based on *interface structural motifs*, which allows elucidation of (1) which E3s interact with which E2s in the human ubiquitination pathway and (2) how they interact with each other. Interface analysis of E2–E3 complexes reveals that loop L1 of E2s is critical for binding; the residue in the sixth position in loop L1 is widely utilized as an interface hot spot and appears indispensable for E2 interactions. Other loop L1 residues also confer specificity on the E2–E3 interactions: HECT E3s are in contact with the residue in the second position in loop L1 of E2s, but this is not the case for the RING finger type E3s. Our modeled E2–E3 complexes illuminate how slight sequence variations in E2 residues may contribute to specificity in E3 binding. These findings may be important for discovering drug candidates targeting E3s, which have been implicated in many diseases.

KEYWORDS: ubiquitination, E2, E3, proteomics, protein–protein interactions, protein–protein interfaces, degradation, proteome-scale structural maps

■ INTRODUCTION

Protein modification by ubiquitin (ubiquitination) is a key mechanism for controlling many cellular processes such as protein degradation, DNA repair, signal transduction, transcription, immunity, endocytosis and cell death. Ubiquitination is achieved by a sequential enzymatic cascade of ubiquitin-activating (E1), ubiquitin-conjugating (E2), and ubiquitin-ligating (E3) enzymes.¹ In the human genome, there are only two known E1s,² which are conserved across different species while there are multiple E2s and many more E3s. Ubiquitin is activated by an ATP-dependent E1 and transferred to E2. E3s recognize specific target substrates and catalyze the ubiquitin transfer from the E2 to the substrate. Ubiquitin can be conjugated to the substrate as a monomer on one (monoubiquitination) or more substrate lysines (multiubiquitination) or as a polymer (polyubiquitination) by forming ubiquitin chains.³ There are three types of E3 ligases based on the protein families they belong to: Homologues of E6AP Carboxy Terminus (HECT), Really Interesting New Gene

(RING), and the UFD2 homology (U-box) family proteins. HECT is a domain of ~350 amino acids. It has a bilobal structure, in which the N-terminal lobe contains the E2-binding site, and the C-terminal lobe confers catalytic activity. The conserved Cys residue in the C-terminal lobe forms thioester bonds with ubiquitin.⁴ The RING finger domain is formed by a short motif rich in histidine and cysteine residues that coordinate zinc atoms in a cross-brace structure, characterized by a central α -helix and variable-length loops separated by small beta strands.⁵ The U-box domain constitutes a relatively small family of E3s and is similar to the structure of the RING domain with the exception that it lacks the conserved histidine and cysteine residues.⁶ The ubiquitination mechanisms for these three E3 classes differ. For HECT domain E3s, the ubiquitin is first transferred from E2 to the active site residue of HECT E3, with subsequent transfer from E3 to the substrate

Received: September 9, 2011

Published: December 12, 2011

protein. For RING and U-box type E3s, ubiquitin is directly transferred from E2 to the substrate without an E3 intermediate linkage. The E2 family is characterized by the presence of a highly conserved 150–200 amino acid catalytic core domain, which consists of four α -helices, a short 3_{10} helix and a four-stranded, antiparallel β -sheet^{7,8} (Figure 1A). The β sheet forms

Figure 1. Structural representation of E2 and E3 proteins. (A) Member of E2 family, UBE2D1 (PDB code: 2c4p:A) is visualized. The catalytic core domain consists of 150 residues. α -Helix 1 and highly flexible regions, loops L1 and L2, which are involved in E3 selection and binding, are displayed in yellow, green and orange color, respectively. Conserved catalytic cysteine residue, to which ubiquitin is attached, is shown as red spheres. (B) Interaction between an E3 ligase c-Cbl RING and UBE2L3 is shown. Loops L1 and L2 of UBE2L3 interact with α -helix and zinc-chelating loops of the c-Cbl RING domain, and α -helix 1 of UBE2L3 is in contact with the c-Cbl linker region.¹⁶

a central region bordered by helices, and there are two loop regions, L1 and L2, that are located at the C-terminal to β -strands 1 and 3.^{8–10} These loop regions show a high level of flexibility and are involved in E3 selection and binding.⁸ E2 families are distinguished by minor sequence differences in the core and amino and/or carboxyl terminal extension domains.³ On the basis of these, they are classified into four groups. Class I E2s consist only of the catalytic core domain. In addition to the core domain, Class II and Class III E2s contain amino terminal and carboxyl terminal extensions, respectively. Class IV E2s contain both amino and carboxyl terminal extensions.¹⁰

Currently, there is a limited number of known E2s (~40 in human); however, the number of known E3s is increasing rapidly (~500 or more have been proposed to exist in human),^{11–13} which suggests that one E2 can recognize several different E3s.¹⁴ E2 and E3 proteins are known to function combinatorially (Figure 2); however, the principles determining the E2–E3 selectivity are unclear and in many cases it is not known which E3s interact with which E2s.¹⁵ The first structural clues of E2–E3 interaction specificity were obtained from the crystal structure of the E3 ligase c-Cbl RING and UBE2L3¹⁶ (Figure 1B). Two loop regions, L1 and L2, of UBE2L3 are in contact with the α -helix and zinc-chelating loops of the RING domain. The central F63 residue in loop L1 and P97 and A98 in loop L2 of UBE2L3 mediate its binding to c-Cbl. In addition, the c-Cbl linker region interacts with α -helix 1 of UBE2L3.¹⁶ In another study, CNOT4 RING finger binding to the ubiquitin-conjugating enzyme UBE2D2 was found to be highly selective.¹⁷ Charge-alteration of residues E49 of CNOT4 and K63 of UBE2D2 results in altered specificity of a functional E2–E3 enzyme pair.¹⁷ In one experimental approach, two-hybrid experiments were performed to identify E2 partners of Brca1 ubiquitin ligase. Brca1 was found to interact with

multiple different E2s and to possess unique ubiquitin-transfer properties depending on the E2 used.^{18,19} In another study, *in vivo* cross-linking methods were used to identify HRD1 (HMG-CoA reductase degradation) ubiquitin ligase interactions.²⁰ Two large-scale studies have recently addressed this E2–E3 identification problem: Wijk et al.²¹ performed a global yeast-two hybrid screen and uncovered over 300 high quality interactions; Markson et al.²² combined yeast two-hybrid screens with homology modeling methods to generate a map of human E2–E3 RING interactions.

Although these studies identified new E2–E3 pairs, the nature and structural details of the interactions in the ubiquitin system are lacking. Here, we aim to model the human E2–E3 interactions on a large, proteome scale and to obtain an insight into their interaction specificity. To carry out this study, we have used Prism,^{23–25} which employs a highly efficient strategy to predict protein associations based on interface structural motifs. The Prism rationale argues that if any two proteins contain regions on their surfaces that are similar to complementary partners of a known interface, in principle these two proteins can interact with each other through these regions. This knowledge-based strategy, which utilizes structural and evolutionary similarity, is made more physical and biologically relevant by including flexibility and energetic assessment in the modeling. This is achieved by using FiberDock,²⁶ a flexible docking refinement server. Using Prism, we have constructed a human structural E2–E3 network consisting of 107 predicted interactions among 22 E2s and 16 E3s. Thirty-six percent of our predicted interactions were reported in earlier studies as interacting pairs; however, how they interact has been unclear. We first observed that E3 proteins could interact with multiple E2s and likewise E2 proteins could interact with multiple E3s, which is expected. However, analysis of the modeled interfaces of E2–E3 putative complexes revealed some structurally conserved residues that are present in almost all interfaces and as such are likely to be indispensable for E2 binding. Comparison of the E2-HECT domain E3 and E2-RING domain E3 interfaces suggests that the E2 loop L1 residues confer specificity in binding to different E3s. The structural E2–E3 network in this study, together with interface analysis, provides a resource for future studies of ubiquitination and E2–E3 selectivity.

MATERIALS AND METHODS

Template and Target Data Set

In this study, we predict and model complexes based on the known interfaces in a template data set. To construct the template data set, we extract all known E2–E3 complexes in the ubiquitination pathway from the PDB.²⁷ There are 9 available E2–E3 complex structures which are listed in Supporting Information, Table S3.

The target data set (Supporting Information, Table S4) contains the E2 and E3 protein structures among which we want to uncover possible interactions. The list of ubiquitin ligases (E3) and ubiquitin conjugating enzymes (E2) related to the human ubiquitination pathway are obtained from the KEGG database²⁸ and available 3D structures are extracted from the PDB. There are 24 E2 proteins, 20 RING finger type E3, 9 HECT type E3 and 3 U-box type E3 proteins with three-dimensional protein structures. Among these, Prism algorithm predicts interactions between 22 E2 and 16 E3 proteins (Supporting Information, Table S4).

Prediction Algorithm

The prediction algorithm is composed of four consecutive steps (Figure 3): extraction of the surface of target proteins, structural alignment, collision check and flexible refinement. In the first step, surface regions of the target proteins are extracted using Naccess²⁹ based on the relative accessible surface area of the residues. If the relative surface accessibility of a residue is more than 15%, then it is labeled a surface residue. In the second step, each interface in the template data set is split into its chain components. Using the MultiProt engine,³⁰ our algorithm searches whether the target surfaces are structurally similar to complementary partners of a template interface. At least 40% of the residues of template chains should be matched to the target surface residues to pass to the next step. If the template chains contain less than 50 residues, this threshold is 60%. In addition to the structural similarity, evolutionary similarity between the template and target surfaces is assessed; at least one hot spot in each template partner should match with the target surface. In the third step, each target protein is transformed onto the corresponding template interface to form the complex structure. After the transformation, the colliding residues are checked; if two partners have more than five spatially colliding residues, then the match is eliminated. At this step, side chain collisions are not considered and left to the last step for a more accurate treatment. The last step is the flexible refinement (backbone and side chain) of the predicted complexes using FiberDock²⁶ to resolve the steric clashes and rank the predicted complexes according to their calculated global energy. FiberDock uses both low and high frequency normal modes and models the backbone and side-chain movements according to the binding van der Waals forces between the receptor and ligand. We assign the larger protein chain to be the receptor and the smaller to be the ligand. In the side chain optimization, only clashing interface residues are considered as movable and 20% of clashes between the side-chain atoms are allowed. Following the refinement process, FiberDock calculates a global energy for each predicted complex. The putative complex structures with a global energy less than -10 kcal/mol are included in the human structural E2–E3 interaction network.

Interface Analysis

Interface regions of putative E2–E3 complexes are analyzed using the Hotpoint server,³¹ which uses an efficient method to determine computational hot spots based on conservation, solvent accessibility and statistical pairwise residue potentials of the interface residues. Hot spots are critical residues at the interface which account for the majority of the binding free energy.³² Structure-based multiple sequence alignment of human E2s are obtained from Christensen et al.¹⁹ and interface and computational hot spot residues are displayed. Computational mutagenesis analysis is performed using FoldX algorithm.³³ First, the putative complex structures are subjected to an optimization procedure using the repair function of FoldX. Next, quantitative estimations of the binding affinities of the wild-type and mutants are obtained by the PositionScan function.

RESULTS AND DISCUSSION

Interactions between ubiquitin-conjugating (E2) and ubiquitin-ligating (E3) enzymes are essential for ubiquitination (Figure 2). Here we computationally model E2–E3 interacting pairs in the human proteome on a broad scale to obtain the interaction

Figure 2. Ubiquitination mechanism. The ubiquitination process starts with the activation of ubiquitin by an E1 enzyme that then transfers it to the E2 protein. E2 together with ubiquitin bind to E3 protein ligase which interacts with the target substrate through another region. E3 ligase assists in ubiquitin transfer from E2 onto the substrate. The ubiquitin transfer mechanism depends on the E3 type; for HECT domain E3s, ubiquitin is first transferred from E2 to the active site residue of HECT E3, with subsequent transfer from E3 to the substrate protein whereas for RING and U-box type E3s, ubiquitin is directly transferred from E2 to the substrate without an E3 intermediate linkage.

network between all E2s and E3s whose structures have been determined, and in particular, to obtain structural insight into E2–E3 selectivity based on the large set. Modeling the structural network provides data not only relating to which E2s interact with which E3s but also to how they interact. Analysis of these interactions may yield patterns that distinguish between the E2 and E3 interaction classes.

In the following sections, first we explain how we constructed the human structural E2–E3 interaction network and present its topological characteristic. Next we provide an analysis of E2–E3 interfaces. We then describe the patterns that we observed and hypothesize that they may play a role in E2–E3 specificity.

Constructing a Human Structural E2–E3 Interaction Network

Incorporating protein structural information into networks is crucial for understanding the details of the interactions. Here, using the available structural proteome, we construct a human structural E2–E3 interaction network. First we compile the available structures for E2 and E3 proteins in the human ubiquitination pathway; there are 24 E2 proteins, 20 RING finger type E3, 9 HECT type E3 and 3 U-box type E3 proteins. To uncover possible interactions among these E2 and E3 proteins, we use our efficient prediction algorithm Prism.^{23–25} Prism has proven useful in our previous studies in the construction of pathways³⁴ and characterization of interactions³⁵ and networks.³⁶ On a docking benchmark,³⁷ Prism method is validated: it efficiently finds high-quality models for 87 out of 88 benchmark complexes and their binding regions (Tuncbag et al., in press). Prism exploits structural and evolutionary similarities, and a template interface data set which consists of known protein–protein interfaces. On the basis of these, it predicts potential interactions between target proteins. In this study, the known E2–E3 interfaces from the PDB constitute the template interfaces. Figure 3 presents a schematic representation of the Prism algorithm that is composed of four consecutive steps. In the first step, the surface regions of the target proteins are extracted. In the second step, each template

Figure 3. Schematic illustration of the Prism algorithm. In the first step, surface regions of target proteins are extracted. In the second step, each template interface is split into the complementary partner chains and the partners are structurally aligned with target protein surfaces by the MultiProt engine.³⁰ This concept is visualized here: Ube3a-UBE2L3 (PDB code: 1c4z:AD) is one of the complexes in our template interface set. When target proteins are structurally aligned with Ube3a and UBE2L3, an E3 ligase Huwe1 (PDB code: 3g1n:A) is found to have a structurally similar binding region to that of Ube3a. Likewise, one of the E2 proteins; UBE2D1 (PDB code: 2c4p:A) has a similar binding region to that of UBE2L3. Then, in principle, these two proteins Huwe1 and UBE2D1 can interact with each other through these regions. Predicted complexes are assessed in the third step: target proteins are transformed onto the template interface and if the residues of the target partners collide, these pairs are eliminated. In the last step, to obtain biologically more relevant interactions, flexible refinement of the rigid docking solutions of MultiProt is performed using FiberDock²⁶ and the predicted complexes are ranked according to the calculated global energy.

interface is split into the complementary partner chains and the partners are structurally aligned with the surfaces of the target proteins by the MultiProt engine.³⁰ Forty percent of the residues of template chains should match the target surfaces to pass to the next step. Besides the structural similarity, evolutionary similarity between target surface and template interface is assessed: at least one hot spot in each template partner should correctly match with the target surface. In the third step, target proteins are transformed onto the template interface and if the residues of the target partners collide (interpenetrate), then these pairs are eliminated. The last step is flexible refinement of the rigid docking solutions of MultiProt

to remove steric clashes, refine and rank the predicted complexes according to the global energy using FiberDock.²⁶ FiberDock considers both side chain and backbone flexibility.

The resulting network, which consists of Prism predictions, contains 107 interactions between 38 proteins and they are listed in Supporting Information, Table S1. There are 22 E2 proteins, 9 RING-finger domain E3, 5 HECT-domain E3 and 2 U-box domain E3 proteins. Among these 22 E2s and 16 E3s, the number of all possible E2–E3 pairs is 352 (i.e., 22×16), out of which Prism predicts 107 to be interacting and leaves the remaining 245 as unfavorable E2–E3 pairs. Out of 352 possibilities, 51 pairwise E2–E3 interactions were already

reported in earlier studies. Although information related to noninteracting E2–E3 pairs is limited, a detailed literature survey showed that 24 E2–E3 pairs were reported to be negative (listed in Supporting Information, Table S5). We recover 76% (39/51) of the known pairs through 107 predicted interactions verifying 36% (39/107) of the network (labeled in Supporting Information, Table S1). Of the 24 E2–E3 interactions reported as negative, 75% (18/24) are labeled as unfavorable by Prism. On the basis of known E2–E3 interaction data, the accuracy of our predictions is 76% $((39 + 18)/(51 + 24))$, which indicates that the predicted E2–E3 interactions are in agreement with validated functional E2–E3 pairs. The Prism method depends on the coverage of known interface architectures and at the time of this study, there are only 9 E2–E3 known complex structures in the PDB. Because the size and coverage of the PDB increases exponentially, we expect the prediction efficiency of Prism to increase.

Topological Characteristic of the Structural E2–E3 Interaction Network

In terms of network topology, since the network is structural and the structural database is incomplete, it is not a typical scale-free protein–protein interaction network and does not follow a power law distribution; however, it is a connected network with an average degree of 5.6. As we noted above, we observe that E3s interact with multiple E2s and likewise E2 proteins interact with multiple E3s. Some of the E3s and E2s have many interaction partners raising the possibility that they are “adaptable” and multifunctional. Among the E2s, UBE2D (UbcH5) family proteins show the highest number of interactions, consistent with previous findings that UbcH5 is active with most E3s.¹⁹ These E2s are involved in degradation of misfolded and short-lived proteins and their orthologs Ubc4 and Ubc4 in *Saccharomyces cerevisiae* have been shown to be functionally redundant.⁵⁸ Such redundancy among UBE2D E2s would be a physiological advantage because it protects the ubiquitination networks from genetic perturbation.²² Another highly interacting E2 in our network is UBE2N, which is involved in Lys63 ubiquitin chain assembly.³⁹ In contrast, E2s such as UBE2F, UBE2S, UBE2G1 and UBE2Q1/2 are involved in very few interactions. In comparison to previous E2–E3 RING interaction networks constructed by yeast two-hybrid screens,^{21,22} we encounter similar topological trends: E2s showing a high number of E3 interactions are reported to be UBE2N, UBE2D and UBE2E families,^{21,22} while UBE2F, UBE2S and UBE2G1 have only a few interactions.²² Thus, it appears that while some E2s such as the UBE2D family may function in the majority of the ubiquitination events in human cells,²² others such as UBE2F, which is known to conjugate ubiquitin-like protein NEDD8, are less important in maintaining the integrity of the ubiquitination network.²¹

Different patterns of substrate ubiquitination lead to different substrate fates. Monoubiquitination can regulate DNA repair, endocytosis and gene expression, whereas polyubiquitination through Lys48 ubiquitin chains generally results in proteasomal degradation, and Lys63-linked ubiquitin chains can function in endocytosis and signaling.⁴⁰ Although mechanisms that control lysine selection in substrates are not clearly identified, structural aspects of E2–E3 pairs and their binding to the substrate appear to be important.⁴¹ A RING E3 can utilize different E2s that have different linkage specificities,⁴² and its activity is more likely to depend on the nature of the E2s which are present.¹⁹ However, for HECT E3s, the chain linkages of the substrate are

usually determined by the E3 itself.⁴³ In our network, we observe that highly interacting RING and HECT E3s would mediate mono- and poly ubiquitination, indicating that they play roles in a broad range of cellular processes. Chain linkage properties of E2s and E3s (if available) are listed in Supporting Information, Table S4. We find that Ube3a, a HECT protein ligase, for which several substrates are identified including p53, Mcm7 and cell cycle regulator Cdkn1b, has 17 E2 partners. Ube3a specifically favors Lys48-linked polyubiquitination⁴³ and thus mediates degradation of cytoplasmic misfolded proteins.⁴⁴ As a multifunctional protein ligase, it is conceivable that Ube3a can interact with several ubiquitin-conjugating E2s, which are all able to catalyze Lys48-linked chains. We also predict Ube3a to be interacting with Sumo-conjugating enzyme UBE2I, which was reported earlier,⁴⁵ although the sumoylation function of Ube3a was not clearly identified. Ube3a mediates the ubiquitination of Pml (promyelocytic leukemia protein) tumor suppressor,⁴⁶ which is also known to undergo sumoylation required for nuclear body formation.⁴⁷ Thus, in addition to ubiquitination of Pml, Ube3a may function as an E3 mediating the sumoylation of Pml. Another highly interacting HECT E3 is Smurf2, which can induce degradation of receptor-regulated R-Smads and R-Smad bound partners,⁴⁸ mediating TGF β signaling via ubiquitination of the receptors or inhibitory Smad7 adaptor.⁴⁹ Smurf2 is widely expressed in tissues. In addition to its ubiquitin-conjugating function, we predict that it interacts with Sumo- and Nedd8-conjugating E2s, UBE2I and UBE2M, respectively. In the TGF β signaling pathway, Smad4 is known to be sumoylated in the presence of UBE2I, which results in the redistribution of Smad4 to subnuclear speckles.⁵⁰ Thus, Smurf2 may also interact with UBE2I in this pathway. For Huwe1, a recently identified HECT E3 implicated in the regulation of cell proliferation, apoptosis and DNA damage response,^{51–53} most of the E2 partners are not yet identified. We predict that UBE2C, -D1, -D2, -D3, -J2 and -L6 could be possible E2 partners. HECT E3s usually prefer to interact with E2s containing a phenylalanine at position 62 (by UBE2D1 numbering). We observe that all predicted E2 partners of Huwe1 contain a phenylalanine at this position except UBE2C, which has a tyrosine residue. Since both phenylalanine and tyrosine are hydrophobic residues,⁵⁴ and since UBE2C is similar to UBE2D1 in forming multiple different ubiquitin linkages,⁵⁵ we hypothesize that UBE2C would be a possible partner of HECT E3s. Considering the RING E3s, c-Cbl is a highly interacting ligase with 12 E2 partners in our network. c-Cbl, which targets numerous substrates, has both mono-ubiquitination (facilitating endocytosis,⁵⁶ lysosomal degradation⁵⁷) and polyubiquitination activities. In our network, in addition to UBE2D and UBE2L family interactions as described in the literature, and other polyubiquitin catalyzing E2s, c-Cbl is also found to interact with UBE2B, which may catalyze monoubiquitination of substrates targeted by c-Cbl. For another RING E3, Brca1 (Breast cancer type 1 susceptibility protein), although multiple different E2 partners, possessing unique ubiquitination activities, were identified in a yeast two-hybrid study by Christensen et al.,¹⁹ Prism predicts only three interactions (with UBE2B, -I and -N) and the latter two were reported before. Brca1 represents a special group of E3 enzymes which function as a heterodimeric complex with Bard1.¹⁸ Brca1 binds E2s and is an active E3 ligase only in association with Bard1. Christensen et al.¹⁹ tested the Brca1-Bard1 RING domains for E2 interactions and several E2 partners: UBE2D1, -D2, -D3, -E1, -E2, -I, -K, -L3 and -N were

Figure 4. Interactions of UBE2L3 with HECT E3s and RING-finger E3s are visualized. UBE2L3, HECT E3s and RING E3s are displayed in gray, pink and blue color, respectively. UBE2L3's binding regions are α -helix 1 (in yellow), loop L1 (in green) and loop L2 (in orange). On these regions, there are seven structurally conserved chemical contacts, which are identified using MAPPIS.⁷¹ Among these, loop L1 residues Pro62 and Phe63 and loop L2 residue Pro97 appear to be conserved through all interactions of UBE2L3, whereas loop L1 s position residue Ala59 (in red color) is observed only in HECT E3-UBE2L3 interfaces. Hydrophobic aliphatic residue contacts of Ala59 are visualized in the zoomed representation in the left panel. MAPPIS abbreviations of physicochemical properties of residue contacts are: Ali, aliphatic hydrophobic property; Pii, aromatic property; Dac, hydrogen bond donor and acceptor; and Acc, hydrogen bond acceptor.

detected. However, yeast two-hybrid screening with isolated Brcal was shown to yield different E2 interactions from screening with fused Brcal-Bard1 dimers.²¹ The reason we observe a limited set of E2s and many E2s as unfavorable for Brcal is that we only considered a single Brcal domain. A detailed analysis of the E2–E3 interactions is given in the next section.

Analysis of the E2–E3 Interaction Interfaces

The principles of selectivity between E2 and E3 proteins are still largely unclear. The factors that play a role in determining which E2 would bind E3 include (i) the substrate physiological/chemical state; that is, whether its concentration is too high, or it is damaged, etc.; (ii) the local concentration of the E2, which is affected by for example, its subcellular localization or distribution in tissues; and (iii) the E2 conformational state, which is determined by, for example, its post-translational modifications.^{58,59} All factors change dynamically with the fluctuating cellular conditions. Here we focus on preferred interactions from the structural standpoint. Below, we present an analysis of the interfaces of the E2–E3 complexes.

General Trends in E2–E3 Binding

Analysis of the interfaces of E2–E3 pairs is expected to reveal which residues are conserved among E2s interacting with a common E3 and which residues differ. Such analysis may provide clues to specificity. For each E2–E3 pair in our network, we analyzed the interacting interfaces and predicted hotspot residues at the interfaces using the Hotpoint server.³¹ The critical residues at the interface account for the majority of the binding energy are called hot spots.³² The Hotpoint server determines computational hot spot residues based on conservation, solvent accessibility and statistical pairwise residue potentials of the interface residues. The predicted hot

spots are observed to match with the experimental hot spots with an accuracy of 70%.^{31,60}

The E2 proteins which have been characterized so far are known to recognize E3s through the L1 and L2 loops and the N-terminal α -helix 1 on the E2 surface.⁶¹ In particular, loop L1 has been shown to be critical in E3 binding. Loop L1 residues phenylalanine at position 62 and lysine at position 63 of UBE2D2 have been observed to bind to RING E3s c-Cbl¹⁶ and CNOT4.¹⁴ In our network, E2 interface residues also mostly correspond to the α -helix 1, loop L1 and L2 regions. Considering all of the E2 interfaces, there seem to be some structurally conserved residues in E3 binding. We observe that the sixth residue of loop L1 (phenylalanine 62 by UBE2D2 numbering) is a critical residue in E3 binding which frequently acts as a hotspot. In addition, the fifth residue in α -helix 1 (arginine 5 by UBE2D2 numbering) and the Ser-Pro-Ala motif in loop L2 (UBE2D2 labeling) mostly participate in the E3 binding. Earlier work has also shown the importance of the Ser-Pro-Ala motif in E2s binding to U-box E3 CHIP⁶² and RING E3 Brcal.¹⁹ These residues are found in almost all E2 interfaces in our network, which suggests that they could be indispensable for E2 binding to an E3. To investigate whether these residues are structurally conserved among the E2 family, we used the ProBIS algorithm,⁶³ which detects structurally conserved regions of a protein by local structural alignments through ~23000 nonredundant PDB structures based on geometry and physicochemical properties. UBE2D1 (PDB code: 2c4p) is the query protein. 57 similar structures are found. The α -helix 1 residues Arg5 and Glu9, loop L1 residues Tyr60, Pro61, Pro64 and Pro65, loop L2 residues Trp 93 and Ser94 are observed to show high structural conservation. The Phe62 residue in loop L1, which we frequently observe in the E2–E3 interface in our network, is found to be moderately conserved. Thus, E2 family

E3 partner		UBE2L3 (UbcH7) interface									
		Helix-1			Loop L1			Loop L2			
HECT	NEDD4L	1	MAASRLMKEL	EEIRKC	:58	PAEYFFKPP	:93	EN	-----	WKPA	-----TK
	SMURF2	1	MAASRLMKEL	EEIRKC	:58	PAEYFFKPP	:93	EN	-----	WKPA	-----TK
	UBE3A	1	MAASRLMKEL	EEIRKC	:58	PAEYFFKPP	:93	EN	-----	WKPA	-----TK
RING	TRAF6	1	MAASRLMKEL	EEIRKC	:58	PAEYFFKPP	:93	EN	-----	WKPA	-----TK
	RBX2	1	MAASRLMKEL	EEIRKC	:58	PAEYFFKPP	:93	EN	-----	WKPA	-----TK
	C-CBL	1	MAASRLMKEL	EEIRKC	:58	PAEYFFKPP	:93	EN	-----	WKPA	-----TK

Figure 5. Interface and computational hotspot residues of UBE2L3 interacting with E3s. Interface residues and hotspot residues of UBE2L3 are shown in cyan and red color, respectively. Comparison of HECT type E3s (Nedd4l, Smurf2 and Ube3a) and RING finger type E3s (Traf6, Rbx2 and c-Cbl) reveals that loop L1 of UBE2L3 is utilized differently. UBE2L3 uses the second position residue in loop L1 frequently as hotspot while interacting with HECT E3s.

proteins can employ these conserved residues to interact with different E3s. Interface and computational hotspot residues on α -helix 1, loop L1 and L2 regions of E2s are provided for all E2–E3 interactions in the Supporting Information, Table S2.

For HECT E3s, which are composed of two subdomains connected by a flexible peptide linker, the C-terminal lobe contains the catalytic cysteine and the N-terminal lobe contains the E2 binding region.⁴ For Huwe1, in the E2 binding region (residues 4150–4200), 10 important residues for E2 binding were identified: F4153, G4156, L4157, Y4159, L4160, Y4170, L4172, V4178, Y4206 and C4211.⁵¹ Our Huwe1–E2 models indicates that Huwe1 utilizes most of these residues when interacting with its E2 partners. In particular, 6 of these residues (F4153, L4157, L4160, Y4170, L4172 and Y4206) are identified as hotspot by Hotpoint server. Similarly, for other HECT E3s—Nedd4l, Smurf2, Ube3a and Wwp1—our interface modeling and computational hotspots are in good agreement with experimental data obtained in previous works.^{51,64,65} From the RING E3 perspective, for c-Cbl, Trp408 and Ile383 residues were shown to have a critical role in E2 binding.¹⁶ The side chain of Trp408 is exposed to the solvent in the E2 binding cleft and hydrophobic residues are often found in an equivalent position in other RING E3s.⁴² In each of our c-Cbl–E2 models, both Trp408 and Ile383 are labeled as hotspots by the Hotpoint server, indicating their importance in binding. Another study showed that mutation of Trp408 to alanine reduces c-Cbl's affinity for the E2 and eliminates its ubiquitin-ligase activity in vitro.⁶⁶ For the Traf6–UBE2N complex structure, seven residues within the RING domain of Traf6, Glu69, Pro71, Ile72, Leu74, Met75, Ala101 and Pro106, were found to contribute significantly to E2 binding.⁶⁷ Among these, Ile72 and Leu74 were completely buried at the interface and contribute the most surface areas.⁶⁷ In our Traf6–E2 models, these residues reside at the interface although their contribution to binding differs, probably depending on the E2 partner present. Most of these residues are labeled as hotspots (especially Ile72 and Leu74). A comparison of HECT and RING E3s interacting with E2s is given in the next section.

HECT E3s Utilize Loop L1 of E2s Distinctively from the RING Finger-type E3s

As we noted above, ubiquitin transfer mechanisms differ among the E3 families. Here, we investigate whether this distinctive mechanistic feature, that is, of ubiquitin forming an intermediate with the HECT domain before being transferred to the substrate, is also observed in HECT E3–E2 interaction interfaces. The 16 E3s in the putative E2–E3 complexes that were predicted by Prism belong to different families: Rbx1,

Rbx2, Mdm2, Xiap, Trim37, c-Cbl, Brca1, Traf6 and Birc3 are RING finger type; Huwe1, Nedd4l, Wwp1, Smurf2 and Ube3a are HECT type; and Ube4a and Ube4b are of the U-box type. These E3s appear to share the same E2s; however, different from the RING finger and U-box type E3s, HECT-type E3s utilize the second position residue in loop L1 of E2s. The identity of this residue in loop L1 is not conserved among E2s, but its contacts with HECT E3s seem to be conserved structurally and chemically. One example shown in Figure 4 is UBE2L3, an E2 that is involved in degradation of many proteins. We find that HECT E3s Nedd4l, Smurf2, Ube3a and RING finger E3s Traf6, c-Cbl and Rbx2 interact with the same E2, UBE2L3, using a shared E2 binding site. For Ube3a and c-Cbl, the complex structures with UBE2L3 are already available in the PDB, whereas for Nedd4l, Smurf2 and Traf6 the structural complexes are not known, although these interactions were reported in earlier works.^{68–70} We modeled UBE2L3 interactions with Nedd4l and Smurf2 based on a known E2–E3 interface between UBE2D2 and Nedd4l (pdb code: 3jw0:AC) and the interaction with Traf6 is based on a known interface between the UBE2D1 and CHIP proteins (pdb code: 2oxq:AC). The interaction with Rbx2 is predicted by Prism. Figure 5 lists the interface and hotspot residues of UBE2L3. When we compare the chemical contacts of these interactions using MAPPIS,⁷¹ we observe seven structurally conserved contacts for HECT E3s among five UBE2L3 residues: Arg6, Ala59, Pro62, Phe63 and Pro97; and seven structurally conserved contacts for RING E3s among five UBE2L3 residues: Pro62, Phe63, Lys96, Pro97 and Ala98. Among these, loop L1 residues Pro62 and Phe63 and loop L2 residue Pro97 appear to be conserved through all interactions of UBE2L3. These residues mostly make hydrophobic aliphatic and aromatic (π) contacts with E3 residues according to the MAPPIS⁷¹ classification (see Figure 4) and thus they seem to be important in binding. Earlier work has supported this observation; Phe63 of UBE2L3 was shown to be recognized by both HECT and RING E3 ligases; c-Cbl and Ube3a.¹⁶ Comparison of the chemical contacts of HECT and RING E3s (left and right panel of Figure 4, respectively) reveals that the second position residue Ala59 in loop L1 of UBE2L3 appears to be utilized solely when interacting with HECT E3s. A previous study on E2 binding of Ube3a also reported this second position loop L1 residue as being important in HECT E3–E2 binding.⁷² We observe that this alanine residue makes hydrophobic aliphatic contacts with Thr765 of Nedd4l and Ile559 of Smurf2 and Pro668 of Ube3a. When MAPPIS⁷¹ is applied to known E2–HECT E3 (UBE2D2–Nedd4l, UBE2L3–Ube3a), E2–RING E3 (UBE2D2–Not4hp, UBE2L3–c-Cbl, and

Figure 6. Interface and computational hot spot residues of UBE2E family. (A) For the common interaction partners, Xiap and c-Cbl, interface residues on UBE2E1 and UBE2E2 are similar (first two rows). The major contribution to specificity might be inferred via the third position loop L1 residue of UBE2E1 and UBE2E2 (Glu105 and Asp59, respectively). (B) (Left) Xiap interactions with UBE2E1 and UBE2E2 are displayed (superimposed onto each other). Computational mutagenesis analysis indicate that mutation of Glu105 (blue color) and Asp59 (cyan color) of UBE2E1 and UBE2E2, respectively, does not affect binding affinities when interacting with Xiap ($\Delta\Delta G = 0.1$ kcal/mol). (Right) UBE2E2 specifically interacts with Ube3a utilizing loop L1 Asp59 residue as an interface residue and its mutation is significantly destabilizing ($\Delta\Delta G = 2.6$ kcal/mol). For all interactions of the UBE2E family, phenylalanine (red color) at sixth position in loop L1 appears to contribute significantly to binding.

UBE2N-Traf6) and E2-U-box E3 (UBE2D3-Ube4b, UBE2D1-STIP1) complexes in the PDB, similar trends are observed: loop L1 second position residue contacts are only employed when interacting with HECT E3s. Although the residue type is not conserved among other E2s, for example, UBE2D2 utilizes threonine and UBE2N uses glutamic acid at the second position in loop L1, the contacts of these residues through HECT E3s are structurally conserved. Interestingly, this observation holds among almost all E2-HECT E3 interactions which might be associated with the distinctive ubiquitin transfer mechanism of HECT E3s (see Supporting Information, Table S2 for E2 interaction interfaces). This unique residue contact might play a role in mediating the allosteric communication between E3 binding site and E3 active site, and thus make the ubiquitin transfer from E2 to HECT E3 possible.

Sequence Variations in E2 Residues May Contribute Specificity to E3 Binding: the UBE2E Example

The UBE2E family E2 proteins UBE2E1 and UBE2E2 are highly similar both in sequence and structure (97% sequence similarity and rmsd: 0.54 Å). The Prism results indicate that UBE2E1 and UBE2E2 share two E3 partners, Xiap and C-Cbl. Their remaining E3 partners are different: UBE2E1 specifically interacts with Mdm2 and Birc3 whereas UBE2E2 prefers to interact with Wwp1, Rbx2 and Ube3a. Earlier, van Wijk et al.²¹ reported that UBE2E2 has shown an E3 interaction pattern distinct from the profile of UBE2E1. Since UBE2E1 and UBE2E2 are quite similar, we investigate which residues would be responsible for this interaction specificity. Figure 6A displays the interface and computational hotspot residues on UBE2E1 and UBE2E2 which bind to different E3s. The residues in the binding regions of UBE2E1 and UBE2E2 are identical except for three: the residue in the 12th position in the α -helix 1 (Asp58 in UBE2E1 corresponding to Glu12 in UBE2E2), the residue in the first position in loop L1 (Thr103 in UBE2E1

corresponding to Ser57 in UBE2E2) and the residue in the third position in loop L1 (Glu105 in UBE2E1 corresponding to Asp59 in UBE2E2). Interestingly, these residues are not utilized at the common interface of E3 partners Xiap and C-Cbl. In this way, these E3s are likely to interact with both UBE2E1 and UBE2E2. In contrast, the residue at the third position in loop L1 seems to be crucial in ensuring the interaction specificity of the UBE2E family. For UBE2E2, this Asp59 favors specific binding to Wwp1, Rbx2 and Ube3a whereas for UBE2E1, the same position residue Glu105 plays a role in Mdm2 binding. Another specific binding of UBE2E1 is with Birc3. For this interaction, UBE2E1 utilizes an α -helix 1 residue, Leu44, which does not occur in the UBE2E2 sequence. A previous study²¹ reported the alteration of E3-interaction specificity of E2 UBE2N due to the mutation of its two helix 1 residues. This UBE2N mutant failed to interact with more than half of its E3 partners and gained some new interactions instead.²¹ Similarly, for the UBE2E family, the mutation of loop L1 residues Glu105 and Asp59 for UBE2E1 and UBE2E2, respectively, would probably alter the E3-interaction specificity. To investigate the importance of the contribution of these residues to the binding energy of the complexes, we performed a computational mutagenesis analysis using FoldX algorithm,³³ whose force-field is based on empirical energy terms correlated with experimental $\Delta\Delta G$ measurements.⁷³ Using FoldX, loop L1 residues Glu105 (UBE2E1) and Asp59 (UBE2E2) are mutated to alanine and quantitative estimations of the binding affinities of the wild-type and mutants are obtained. We observe that mutations of these residues in UBE2E1 and UBE2E2 that interact with the same E3s (Xiap and C-Cbl) do not affect the binding affinities significantly ($\Delta\Delta G = 0.3$ kcal/mol on average). In contrast, for the specific E3 interactions of UBE2E1 and UBE2E2 (with Mdm2, Birc3, Rbx2, Ube3a and Wwp1), mutations have a destabilizing effect ($\Delta\Delta G = 1.2$ kcal/mol on average). The most destabilizing is the mutation of loop L1 sixth position

Figure 7. Classification of the human structural E2–E3 interaction network according to the GO⁷⁵ cellular component terms. There are 72, 41, and 8 E2–E3 interactions in cytoplasm, nucleus and plasma membrane classification, respectively. Some interactions are likely to occur ubiquitously whereas others appear to be localization-specific. E2 proteins are displayed in yellow color and E3 proteins are in gray. E3s are visualized according to the domain types; RING-finger domain, HECT-domain and U-box domain E3s are in rectangular, diamond and triangular shape, respectively. Red color edges are the E2–E3 interactions for which the complex structures are available in the PDB.

residue phenylalanine, which is identified as a computational hotspot by the Hotpoint server ($\Delta\Delta G = 1.9$ kcal/mol on average). In Figure 6B left panel, the interactions of UBE2E1 and UBE2E2 with their common partner Xiap are illustrated. Xiap, which mediates the degradation of many proteins such as Caspase-3 and Smac, has been recently reported to interact with both UBE2E1⁷⁴ and UBE2E2.²² A specific interaction of UBE2E2 with Ube3a is displayed on the right panel of Figure 6B. We observe that Asp59 (shown in cyan color) is utilized as an interface residue and its mutation is significantly destabilizing ($\Delta\Delta G = 2.6$ kcal/mol). Consequently, while with mutation of this residue UBE2E1 and UBE2E2 may still interact with Xiap and C-Cbl, they will most probably fail to interact with other E3s. These findings illustrate how slight sequence variations in E2 residues may contribute to the specificity of E3 binding.

E2 Selection Based on Substrate and Subcellular Localization

E2 selection by an E3 may be affected or determined by the substrate that will be degraded since the substrate should be in close proximity to the E2 during the ubiquitin transfer. Subcellular localization of E2 and E3 enzymes may also affect the interaction preferences. To identify which E2–E3 interactions are more likely to occur in our network, we classified the E2 and E3 genes in our network according to their subcellular localization as defined by Gene Ontology⁷⁸ cellular component terms. We observe that many E2 and E3 enzymes are found both in the nucleus and in the cytoplasm and some are also found at the plasma membrane. The classification of E2–E3 interactions according to the cellular components, that is, the interactions in the cytoplasm, nucleus and plasma membrane, is illustrated in Figure 7. The interactions appear to be localization-specific; for example, although Mdm2 ligase is ubiquitous (existing everywhere in the cell), its interactions are dependent on the localization of the E2. One substrate of Mdm2 is p53 tumor suppressor protein, which localizes to

either the nucleus or cytoplasm and can be ubiquitinated and degraded by Mdm2.⁷⁶ In our network, to ubiquitinate p53, Mdm2 can select UBE2E1 or UBE2D1 in the nucleus, while in the cytoplasm it can additionally prefer UBE2D3, UBE2G2 and UBE2L6. Interactions with the UBE2D family and UBE2G2 were reported earlier.²¹ Mdm2 can also be found at the plasma membrane, where it was shown to bind to Beta-arrestin 2 (β arr2), which is a regulatory protein playing a central role in the endocytosis of most G-protein-coupled receptors,⁷⁷ to drive its ubiquitylation.⁷⁸ According to our network, at the plasma membrane, Mdm2 can only interact with UBE2D3 to ubiquitinate β arr2.

Other than Mdm2, Huwe1 and Ube3a (also called E6-AP) are known to ubiquitinate p53, targeting it for degradation as well.^{51,79} To perform this specific function, Mdm2, Huwe1 and Ube3a most likely select similar E2s. Consequently, common E2 partners for these E3s would reveal which E2s would be selected in a substrate-specific manner. In our network, common E2 partners for Mdm2, Huwe1 and Ube3a are UBE2D family E2s (UBE2D1, UBE2D2, UBE2D3) and UBE2L6. Interactions between Ube3a and these E2s are experimentally known^{68,72,80,81} without E2–E3 complex structures. Since Huwe1 is a recently identified E3 ligase, many of its interaction partners are not known. We observe that most of the interface residues on E2s are shared indicating that these E2s can interact with only one E3 at a time. When we compare the global energies of the putative E2–E3 complexes computed by FiberDock, we see that Mdm2 binds most favorably to UBE2D1 (global energy: -26.14 kcal/mol) in the cytoplasm and the nucleus whereas Ube3a prefers UBE2L6 (global energy: -45.32 kcal/mol) in the cytoplasm and UBE2D1 (global energy: -24.9 kcal/mol) in the nucleus. The most favorable E2 partner for Huwe1 is UBE2D2 (global energy: -23.38 kcal/mol), for which cellular component information is not available. The calculated Fiberdock energies for these E2–E3 complexes are listed in Table 1. For each E3 interacting with E2s, the binding energies seem to be close to

Table 1. Interactions and Binding Energies of E2–E3 Complexes Mediating Degradation of the Same Substrate, p53 Protein^a

E3 proteins	E2 proteins			
	UBE2D1	UBE2D2	UBE2D3	UBE2L6
MDM2	<u>-26.1</u>	-21.9	-15.7	-10.4
HUWE1	-10.6	<u>-23.4</u>	-11.5	-14.9
UBE3A	-24.9	-29.3	-32.3	<u>-45.3</u>

^aE3 proteins Mdm2, Huwe1 and Ube3a interact with E2 proteins UBE2D1, UBE2D2, UBE2D3 and UBE2L6 with different binding energies (kcal/mol) computed by Fiberdock. The most favorable partners for Mdm2, Huwe1 and Ube3a are UBE2D1, UBE2D2 and UBE2L6, respectively, for which the binding energies are underlined.

each other, which is consistent with earlier studies that reported similar binding affinities for a common E3 interacting with multiple E2s.^{19,45,82} Depending on which substrate is ubiquitinated and where it is ubiquitinated in the cell, E2 selection of E3s may vary. Our E2–E3 models provide a source for future investigation of selectivity.

CONCLUSIONS

Substrate ubiquitination is mediated by the interactions between E2 enzymes and E3 ligases. Although these E2 and E3 proteins function in a concerted manner, the principles of selectivity between them are still not entirely understood. Here we address E2–E3 interactions in the human proteome by taking a structural approach, and computationally model E2–E3 complexes based on interface structural motifs. By exploiting the available structural proteome and the powerful Prism algorithm,^{23–25} we construct a human structural E2–E3 interaction network. The results indicate which E2s are likely to interact with which E3s and how they could interact. Analysis of the modeled interfaces of E2–E3 pairs elucidates binding patterns: some residues are structurally conserved among E2 proteins and appear to be essential for all E2–E3 interactions, whereas others, particularly in loop L1, appear to play important roles in E3 selectivity. Further, classification of E2 and E3 enzymes according to their subcellular localization would reveal which E2–E3 pairs are more likely to occur in each cellular component and how the E2 preference may vary depending on the substrate.

Several E3s have been implicated in disease, including cancer and neurodegenerative diseases.⁸³ Recently various HECT-E3s have emerged as important regulators of cancer development.⁸⁴ RING-finger E3s are classified as either tumor suppressors or oncoproteins and are often overexpressed in cancer.⁸³ Therefore, in disease processes, to block the activity of E3s, inhibitory molecules can be developed which target E3–E2 interface or E3–ubiquitylation substrate interface. In this study, we focused on predicting and analyzing E2–E3 pairs in the human ubiquitination pathway. We believe that the structural E2–E3 network and the E2–E3 interface data in this study provide a resource for future studies of ubiquitination and E2–E3 selectivity, especially in discovering drug candidates targeting E3s.

ASSOCIATED CONTENT

Supporting Information

Table S1. Human structural E2–E3 interaction network constructed by Prism algorithm. Table S2. Interface and

hotspot residues on α -helix 1, loop L1 and L2 regions of E2s interacting with E3s. Table S3. List of protein complexes in the template data set. Table S4. Target proteins and their PDB codes. Table S5. Negative E2–E3 pairs and E2–E3 pairs that could not be predicted by Prism algorithm. This material is available free of charge via the Internet at <http://pubs.acs.org>.

AUTHOR INFORMATION

Corresponding Author

*Ozlem Keskin: okeskin@ku.edu.tr. Attila Gursoy: agursoy@ku.edu.tr.

ACKNOWLEDGMENTS

This work has been supported by TUBITAK (Research Grant Numbers: 109T343 and 109E207). O.K. acknowledges Turkish Academy of Sciences (TUBA). 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.

REFERENCES

- (1) Hershko, A.; Ciechanover, A. The ubiquitin system. *Annu. Rev. Biochem.* **1998**, *67*, 425–79.
- (2) Jin, J.; Li, X.; Gygi, S. P.; Harper, J. W. Dual E1 activation systems for ubiquitin differentially regulate E2 enzyme charging. *Nature* **2007**, *447* (7148), 1135–8.
- (3) Pickart, C. M. Mechanisms underlying ubiquitination. *Annu. Rev. Biochem.* **2001**, *70*, 503–33.
- (4) Rotin, D.; Kumar, S. Physiological functions of the HECT family of ubiquitin ligases. *Nat. Rev. Mol. Cell. Biol.* **2009**, *10* (6), 398–409.
- (5) Borden, K. L. RING domains: master builders of molecular scaffolds? *J. Mol. Biol.* **2000**, *295* (5), 1103–12.
- (6) Ohi, M. D.; Vander Kooi, C. W.; Rosenberg, J. A.; Chazin, W. J.; Gould, K. L. Structural insights into the U-box, a domain associated with multi-ubiquitination. *Nat. Struct. Biol.* **2003**, *10* (4), 250–5.
- (7) Hofmann, R. M.; Pickart, C. M. In vitro assembly and recognition of Lys-63 polyubiquitin chains. *J. Biol. Chem.* **2001**, *276* (30), 27936–43.
- (8) Burroughs, A. M.; Jaffee, M.; Iyer, L. M.; Aravind, L. Anatomy of the E2 ligase fold: implications for enzymology and evolution of ubiquitin/Ub-like protein conjugation. *J. Struct. Biol.* **2008**, *162* (2), 205–18.
- (9) Michelle, C.; Vourc'h, P.; Mignon, L.; Andres, C. R. What was the set of ubiquitin and ubiquitin-like conjugating enzymes in the eukaryote common ancestor? *J. Mol. Evol.* **2009**, *68* (6), 616–28.
- (10) Winn, P. J.; Religa, T. L.; Battey, J. N.; Banerjee, A.; Wade, R. C. Determinants of functionality in the ubiquitin conjugating enzyme family. *Structure* **2004**, *12* (9), 1563–74.
- (11) Semple, C. A. The comparative proteomics of ubiquitination in mouse. *Genome Res.* **2003**, *13* (6B), 1389–94.
- (12) von Arnim, A. G. A hitchhiker's guide to the proteasome. *Sci. STKE* **2001**, *2001* (97), e2.
- (13) Wong, B. R.; Parlati, F.; Qu, K.; Demo, S.; Pray, T.; Huang, J.; Payan, D. G.; Bennett, M. K. Drug discovery in the ubiquitin regulatory pathway. *Drug Discov. Today* **2003**, *8* (16), 746–54.
- (14) Dominguez, C.; Bonvin, A. M.; Winkler, G. S.; van Schaik, F. M.; Timmers, H. T.; Boelens, R. Structural model of the UbcH5B/CNOT4 complex revealed by combining NMR, mutagenesis, and docking approaches. *Structure* **2004**, *12* (4), 633–44.

- (15) Weissman, A. M. Themes and variations on ubiquitylation. *Nat. Rev. Mol. Cell. Biol.* **2001**, 2 (3), 169–78.
- (16) Zheng, N.; Wang, P.; Jeffrey, P. D.; Pavletich, N. P. Structure of a c-Cbl-UbcH7 complex: RING domain function in ubiquitin-protein ligases. *Cell* **2000**, 102 (4), 533–9.
- (17) Winkler, G. S.; Albert, T. K.; Dominguez, C.; Legtenberg, Y. I.; Boelens, R.; Timmers, H. T. An altered-specificity ubiquitin-conjugating enzyme/ubiquitin-protein ligase pair. *J. Mol. Biol.* **2004**, 337 (1), 157–65.
- (18) Christensen, D. E.; Brzovic, P. S.; Klevit, R. E. E2-BRCA1 RING interactions dictate synthesis of mono- or specific polyubiquitin chain linkages. *Nat. Struct. Mol. Biol.* **2007**, 14 (10), 941–8.
- (19) Christensen, D. E.; Klevit, R. E. Dynamic interactions of proteins in complex networks: identifying the complete set of interacting E2s for functional investigation of E3-dependent protein ubiquitination. *FEBS J.* **2009**, 276 (19), 5381–9.
- (20) Gardner, R. G.; Shearer, A. G.; Hampton, R. Y. In vivo action of the HRD ubiquitin ligase complex: mechanisms of endoplasmic reticulum quality control and sterol regulation. *Mol. Cell. Biol.* **2001**, 21 (13), 4276–91.
- (21) van Wijk, S. J.; de Vries, S. J.; Kemmeren, P.; Huang, A.; Boelens, R.; Bonvin, A. M.; Timmers, H. T. A comprehensive framework of E2-RING E3 interactions of the human ubiquitin-proteasome system. *Mol. Syst. Biol.* **2009**, 5, 295.
- (22) Markson, G.; Kiel, C.; Hyde, R.; Brown, S.; Charalabous, P.; Bremm, A.; Semple, J.; Woodsmith, J.; Duley, S.; Salehi-Ashtiani, K.; Vidal, M.; Komander, D.; Serrano, L.; Lehner, P.; Sanderson, C. M. Analysis of the human E2 ubiquitin conjugating enzyme protein interaction network. *Genome Res.* **2009**, 19 (10), 1905–11.
- (23) Aytuna, A. S.; Gursoy, A.; Keskin, O. Prediction of protein-protein interactions by combining structure and sequence conservation in protein interfaces. *Bioinformatics* **2005**, 21 (12), 2850–5.
- (24) Ogmen, U.; Keskin, O.; Aytuna, A. S.; Nussinov, R.; Gursoy, A. PRISM: protein interactions by structural matching. *Nucleic Acids Res.* **2005**, 33 (WebServer issue), W331–6.
- (25) Tuncbag, N.; Gursoy, A.; Nussinov, R.; Keskin, O. Predicting protein-protein interactions on a proteome scale by matching evolutionary and structural similarities at interfaces using PRISM. *Nat. Protoc.* **2011**, 6 (9), 1341–54.
- (26) Mashich, E.; Nussinov, R.; Wolfson, H. J. FiberDock: Flexible induced-fit backbone refinement in molecular docking. *Proteins* **2010**, 78 (6), 1503–19.
- (27) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, 28 (1), 235–42.
- (28) Kanehisa, M.; Goto, S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **2000**, 28 (1), 27–30.
- (29) Hubbard, S.; Thornton, J. NACCESS; Department of Biochemistry and Molecular Biology, University College London, 1993.
- (30) Shatsky, M.; Nussinov, R.; Wolfson, H. J. A method for simultaneous alignment of multiple protein structures. *Proteins* **2004**, 56 (1), 143–56.
- (31) Tuncbag, N.; Keskin, O.; Gursoy, A. HotPoint: hot spot prediction server for protein interfaces. *Nucleic Acids Res.* **2010**, 38 (Web Server issue), W402–6.
- (32) Bogan, A. A.; Thorn, K. S. Anatomy of hot spots in protein interfaces. *J. Mol. Biol.* **1998**, 280 (1), 1–9.
- (33) Schymkowitz, J.; Borg, J.; Stricher, F.; Nys, R.; Rousseau, F.; Serrano, L. The FoldX web server: an online force field. *Nucleic Acids Res.* **2005**, 33 (Web Server issue), W382–8.
- (34) Tuncbag, N.; Kar, G.; Gursoy, A.; Keskin, O.; Nussinov, R. Towards inferring time dimensionality in protein-protein interaction networks by integrating structures: the p53 example. *Mol. Biosyst.* **2009**, 5 (12), 1770–8.
- (35) Kar, G.; Keskin, O.; Gursoy, A.; Nussinov, R. Allosteric and population shift in drug discovery. *Curr. Opin. Pharmacol.* **2010**, 10 (6), 715–22.
- (36) Kar, G.; Gursoy, A.; Keskin, O. Human cancer protein-protein interaction network: a structural perspective. *PLoS Comput. Biol.* **2009**, 5 (12), e1000601.
- (37) Hwang, H.; Pierce, B.; Mintseris, J.; Janin, J.; Weng, Z. Protein-protein docking benchmark version 3.0. *Proteins* **2008**, 73 (3), 705–9.
- (38) Seufert, W.; Jentsch, S. Ubiquitin-conjugating enzymes UBC4 and UBC5 mediate selective degradation of short-lived and abnormal proteins. *EMBO J.* **1990**, 9 (2), 543–50.
- (39) Deng, L.; Wang, C.; Spencer, E.; Yang, L.; Braun, A.; You, J.; Slaughter, C.; Pickart, C.; Chen, Z. J. Activation of the I κ B kinase complex by TRAF6 requires a dimeric ubiquitin-conjugating enzyme complex and a unique polyubiquitin chain. *Cell* **2000**, 103 (2), 351–61.
- (40) Passmore, L. A.; Barford, D. Getting into position: the catalytic mechanisms of protein ubiquitylation. *Biochem. J.* **2004**, 379 (Pt 3), 513–25.
- (41) Sadowski, M.; Sarcevic, B. Mechanisms of mono- and poly-ubiquitination: Ubiquitination specificity depends on compatibility between the E2 catalytic core and amino acid residues proximal to the lysine. *Cell Div.* **2010**, 5, 19.
- (42) Deshaies, R. J.; Joazeiro, C. A. RING domain E3 ubiquitin ligases. *Annu. Rev. Biochem.* **2009**, 78, 399–434.
- (43) Wang, M.; Pickart, C. M. Different HECT domain ubiquitin ligases employ distinct mechanisms of polyubiquitin chain synthesis. *EMBO J.* **2005**, 24 (24), 4324–33.
- (44) Mishra, A.; Godavarthi, S. K.; Maheshwari, M.; Goswami, A.; Jana, N. R. The ubiquitin ligase E6-AP is induced and recruited to aggresomes in response to proteasome inhibition and may be involved in the ubiquitination of Hsp70-bound misfolded proteins. *J. Biol. Chem.* **2009**, 284 (16), 10537–45.
- (45) Eletr, Z. M.; Kuhlman, B. Sequence determinants of E2-E6AP binding affinity and specificity. *J. Mol. Biol.* **2007**, 369 (2), 419–28.
- (46) Louria-Hayon, I.; Alsheich-Bartok, O.; Levav-Cohen, Y.; Silberman, I.; Berger, M.; Grossman, T.; Matentzoglou, K.; Jiang, Y. H.; Muller, S.; Scheffner, M.; Haupt, S.; Haupt, Y. E6AP promotes the degradation of the PML tumor suppressor. *Cell Death Differ.* **2009**, 16 (8), 1156–66.
- (47) Muller, S.; Matunis, M. J.; Dejean, A. Conjugation with the ubiquitin-related modifier SUMO-1 regulates the partitioning of PML within the nucleus. *EMBO J.* **1998**, 17 (1), 61–70.
- (48) Izzi, L.; Attisano, L. Regulation of the TGF β signalling pathway by ubiquitin-mediated degradation. *Oncogene* **2004**, 23 (11), 2071–8.
- (49) Ebisawa, T.; Fukuchi, M.; Murakami, G.; Chiba, T.; Tanaka, K.; Imamura, T.; Miyazono, K. Smurf1 interacts with transforming growth factor- β type I receptor through Smad7 and induces receptor degradation. *J. Biol. Chem.* **2001**, 276 (16), 12477–80.
- (50) Lee, P. S.; Chang, C.; Liu, D.; Derynck, R. Sumoylation of Smad4, the common Smad mediator of transforming growth factor- β family signaling. *J. Biol. Chem.* **2003**, 278 (30), 27853–63.
- (51) Zhao, X.; Heng, J. I.; Guardavaccaro, D.; Jiang, R.; Pagano, M.; Guillemot, F.; Iavarone, A.; Lasorella, A. The HECT-domain ubiquitin ligase Huwe1 controls neural differentiation and proliferation by destabilizing the N-Myc oncoprotein. *Nat. Cell Biol.* **2008**, 10 (6), 643–53.
- (52) Hall, J. R.; Kow, E.; Nevis, K. R.; Lu, C. K.; Luce, K. S.; Zhong, Q.; Cook, J. G. Cdc6 stability is regulated by the Huwe1 ubiquitin ligase after DNA damage. *Mol. Biol. Cell* **2007**, 18 (9), 3340–50.
- (53) Adhikary, S.; Marinoni, F.; Hock, A.; Hulleman, E.; Popov, N.; Beier, R.; Bernard, S.; Quarto, M.; Capra, M.; Goettig, S.; Kogel, U.; Scheffner, M.; Helin, K.; Eilers, M. The ubiquitin ligase HectH9 regulates transcriptional activation by Myc and is essential for tumor cell proliferation. *Cell* **2005**, 123 (3), 409–21.
- (54) Hu, Z.; Ma, B.; Wolfson, H.; Nussinov, R. Conservation of polar residues as hot spots at protein interfaces. *Proteins* **2000**, 39 (4), 331–42.
- (55) Kirkpatrick, D. S.; Hathaway, N. A.; Hanna, J.; Elsasser, S.; Rush, J.; Finley, D.; King, R. W.; Gygi, S. P. Quantitative analysis of in vitro

ubiquitinated cyclin B1 reveals complex chain topology. *Nat. Cell Biol.* **2006**, *8* (7), 700–10.

(56) Jiang, X.; Sorkin, A. Epidermal growth factor receptor internalization through clathrin-coated pits requires Cbl RING finger and proline-rich domains but not receptor polyubiquitylation. *Traffic* **2003**, *4* (8), S29–43.

(57) Masson, K.; Heiss, E.; Band, H.; Ronnstrand, L. Direct binding of Cbl to Tyr568 and Tyr936 of the stem cell factor receptor/c-Kit is required for ligand-induced ubiquitination, internalization and degradation. *Biochem. J.* **2006**, *399* (1), S9–67.

(58) Pichler, A.; Knipscheer, P.; Oberhofer, E.; van Dijk, W. J.; Korner, R.; Olsen, J. V.; Jentsch, S.; Melchior, F.; Sixma, T. K. SUMO modification of the ubiquitin-conjugating enzyme E2–25K. *Nat. Struct. Mol. Biol.* **2005**, *12* (3), 264–9.

(59) Sadowski, M.; Mawson, A.; Baker, R.; Sarcevic, B. Cdc34 C-terminal tail phosphorylation regulates Skp1/cullin/F-box (SCF)-mediated ubiquitination and cell cycle progression. *Biochem. J.* **2007**, *405* (3), S69–81.

(60) Tuncbag, N.; Gursoy, A.; Keskin, O. Identification of computational hot spots in protein interfaces: combining solvent accessibility and inter-residue potentials improves the accuracy. *Bioinformatics* **2009**, *25* (12), 1513–20.

(61) Ye, Y.; Rape, M. Building ubiquitin chains: E2 enzymes at work. *Nat. Rev. Mol. Cell. Biol.* **2009**, *10* (11), 755–64.

(62) Xu, Z.; Kohli, E.; Devlin, K. I.; Bold, M.; Nix, J. C.; Misra, S. Interactions between the quality control ubiquitin ligase CHIP and ubiquitin conjugating enzymes. *BMC Struct. Biol.* **2008**, *8*, 26.

(63) Konc, J.; Janezic, D. ProBiS algorithm for detection of structurally similar protein binding sites by local structural alignment. *Bioinformatics* **2010**, *26* (9), 1160–8.

(64) Ogunjimi, A. A.; Briant, D. J.; Pece-Barbara, N.; Le Roy, C.; Di Guglielmo, G. M.; Kavsak, P.; Rasmussen, R. K.; Seet, B. T.; Sicheri, F.; Wrana, J. L. Regulation of Smurf2 ubiquitin ligase activity by anchoring the E2 to the HECT domain. *Mol. Cell* **2005**, *19* (3), 297–308.

(65) Verdecia, M. A.; Joazeiro, C. A.; Wells, N. J.; Ferrer, J. L.; Bowman, M. E.; Hunter, T.; Noel, J. P. Conformational flexibility underlies ubiquitin ligation mediated by the WWP1 HECT domain E3 ligase. *Mol. Cell* **2003**, *11* (1), 249–59.

(66) Joazeiro, C. A.; Wing, S. S.; Huang, H.; Leverson, J. D.; Hunter, T.; Liu, Y. C. The tyrosine kinase negative regulator c-Cbl as a RING-type, E2-dependent ubiquitin-protein ligase. *Science* **1999**, *286* (5438), 309–12.

(67) Yin, Q.; Lin, S. C.; Lamothe, B.; Lu, M.; Lo, Y. C.; Hura, G.; Zheng, L.; Rich, R. L.; Campos, A. D.; Myszk, D. G.; Lenardo, M. J.; Darnay, B. G.; Wu, H. E2 interaction and dimerization in the crystal structure of TRAF6. *Nat. Struct. Mol. Biol.* **2009**, *16* (6), 658–66.

(68) Nuber, U.; Schwarz, S.; Kaiser, P.; Schneider, R.; Scheffner, M. Cloning of human ubiquitin-conjugating enzymes UbcH6 and UbcH7 (E2-F1) and characterization of their interaction with E6-AP and RSP5. *J. Biol. Chem.* **1996**, *271* (5), 2795–800.

(69) Wiesner, S.; Ogunjimi, A. A.; Wang, H. R.; Rotin, D.; Sicheri, F.; Wrana, J. L.; Forman-Kay, J. D. Autoinhibition of the HECT-type ubiquitin ligase Smurf2 through its C2 domain. *Cell* **2007**, *130* (4), 651–62.

(70) Geetha, T.; Jiang, J.; Wooten, M. W. Lysine 63 polyubiquitination of the nerve growth factor receptor TrkA directs internalization and signaling. *Mol. Cell* **2005**, *20* (2), 301–12.

(71) Shulman-Peleg, A.; Shatsky, M.; Nussinov, R.; Wolfson, H. J. MultiBind and MAPPIS: web servers for multiple alignment of protein 3D-binding sites and their interactions. *Nucleic Acids Res.* **2008**, *36* (Web Server issue), W260–4.

(72) Huang, L.; Kinnucan, E.; Wang, G.; Beaudenon, S.; Howley, P. M.; Huibregtse, J. M.; Pavletich, N. P. Structure of an E6AP-UbcH7 complex: insights into ubiquitination by the E2–E3 enzyme cascade. *Science* **1999**, *286* (5443), 1321–6.

(73) Tokuriki, N.; Stricher, F.; Serrano, L.; Tawfik, D. S. How protein stability and new functions trade off. *PLoS Comput. Biol.* **2008**, *4* (2), e1000002.

(74) Dynek, J. N.; Goncharov, T.; Dueber, E. C.; Fedorova, A. V.; Izrael-Tomasevic, A.; Phu, L.; Helgason, E.; Fairbrother, W. J.; Deshayes, K.; Kirkpatrick, D. S.; Vucic, D. c-IAP1 and UbcH5 promote K11-linked polyubiquitination of RIP1 in TNF signalling. *EMBO J.* **2010**, *29* (24), 4198–209.

(75) Harris, M. A.; Clark, J.; Ireland, A.; Lomax, J.; Ashburner, M.; Foulger, R.; Eilbeck, K.; Lewis, S.; Marshall, B.; Mungall, C.; Richter, J.; Rubin, G. M.; Blake, J. A.; Bult, C.; Dolan, M.; Drabkin, H.; Eppig, J. T.; Hill, D. P.; Ni, L.; Ringwald, M.; Balakrishnan, R.; Cherry, J. M.; Christie, K. R.; Costanzo, M. C.; Dwight, S. S.; Engel, S.; Fisk, D. G.; Hirschman, J. E.; Hong, E. L.; Nash, R. S.; Sethuraman, A.; Theesfeld, C. L.; Botstein, D.; Dolinski, K.; Feierbach, B.; Berardini, T.; Mundodi, S.; Rhee, S. Y.; Apweiler, R.; Barrell, D.; Camon, E.; Dimmer, E.; Lee, V.; Chisholm, R.; Gaudet, P.; Kibbe, W.; Kishore, R.; Schwarz, E. M.; Sternberg, P.; Gwinn, M.; Hannick, L.; Wortman, J.; Berriman, M.; Wood, V.; de la Cruz, N.; Tonellato, P.; Jaiswal, P.; Seigfried, T.; White, R. The Gene Ontology (GO) database and informatics resource. *Nucleic Acids Res.* **2004**, *32* (Database issue), D258–61.

(76) Yu, Z. K.; Geyer, R. K.; Maki, C. G. MDM2-dependent ubiquitination of nuclear and cytoplasmic P53. *Oncogene* **2000**, *19* (51), S892–7.

(77) Pierce, K. L.; Lefkowitz, R. J. Classical and new roles of beta-arrestins in the regulation of G-protein-coupled receptors. *Nat. Rev. Neurosci.* **2001**, *2* (10), 727–33.

(78) Shenoy, S. K.; McDonald, P. H.; Kohout, T. A.; Lefkowitz, R. J. Regulation of receptor fate by ubiquitination of activated beta 2-adrenergic receptor and beta-arrestin. *Science* **2001**, *294* (5545), 1307–13.

(79) Yamamoto, Y.; Huibregtse, J. M.; Howley, P. M. The human E6-AP gene (UBE3A) encodes three potential protein isoforms generated by differential splicing. *Genomics* **1997**, *41* (2), 263–6.

(80) Anan, T.; Nagata, Y.; Koga, H.; Honda, Y.; Yabuki, N.; Miyamoto, C.; Kuwano, A.; Matsuda, I.; Endo, F.; Saya, H.; Nakao, M. Human ubiquitin-protein ligase Nedd4: expression, subcellular localization and selective interaction with ubiquitin-conjugating enzymes. *Genes Cells* **1998**, *3* (11), 751–63.

(81) Hatakeyama, S.; Jensen, J. P.; Weissman, A. M. Subcellular localization and ubiquitin-conjugating enzyme (E2) interactions of mammalian HECT family ubiquitin protein ligases. *J. Biol. Chem.* **1997**, *272* (24), 15085–92.

(82) Brzovic, P. S.; Keefe, J. R.; Nishikawa, H.; Miyamoto, K.; Fox, D. 3rd; Fukuda, M.; Ohta, T.; Klevit, R. Binding and recognition in the assembly of an active BRCA1/BARD1 ubiquitin-ligase complex. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100* (10), S646–51.

(83) Nalepa, G.; Rolfe, M.; Harper, J. W. Drug discovery in the ubiquitin-proteasome system. *Nat. Rev. Drug Discov.* **2006**, *5* (7), S96–613.

(84) Bernassola, F.; Karin, M.; Ciechanover, A.; Melino, G. The HECT family of E3 ubiquitin ligases: multiple players in cancer development. *Cancer Cell* **2008**, *14* (1), 10–21.