The reaction cycle of HECT domain ubiquitin ligases consists of three steps: 1) binding of an E2 protein, 2) transfer of ubiquitin from E2 to the HECT domain, and 3) transfer of ubiquitin to the substrate. We report the identification of a determinant that is specifically required for the last step of this cycle, a phenylalanine residue located four amino acids from the C terminus of most HECT domains, referred to here as the −4F. Alteration of this residue in human E6AP and Saccharomyces cerevisiae Rsp5p did not affect ubiquitin-thioester formation, but effectively blocked substrate ubiquitination. Alteration of the −4F to alanine with concomitant substitution of a nearby residue to phenylalanine only partially restored Rsp5p activity, indicating that precise spatial placement of this residue is important. C-terminal truncated E6AP and Rsp5p proteins were also defective for substrate ubiquitination, providing a likely biochemical understanding of a previously isolated Angelman syndrome-associated mutation of E6AP that alters the stop codon of an otherwise wild-type gene. We propose that the −4F may play a role in orienting ubiquitin when it is tethered to the HECT active site cysteine. This may be necessary to allow for approach of the incoming lysine ε-amino group of the substrate.

The best characterized function of protein ubiquitination is as a recognition signal for the 26 S proteasome (1). Ubiquitin is covalently linked to substrate proteins via isopeptide bond linkages formed between the terminal carboxyl group of ubiquitin and ε-amino groups of lysine side chains of the substrate, or in some cases to the N-terminal α-amino group. Additional ubiquitin molecules can be conjugated sequentially at specific lysine residues of ubiquitin to form multi-ubiquitinated proteins, which are the principal substrate of the 26 S proteasome.

Three groups of proteins cooperate in catalyzing ubiquitination: the E1 ubiquitin-activating enzyme, the E2 ubiquitin-conjugating proteins (Ubε1 proteins), and the E3 ubiquitin ligases (1). The E1 enzyme uses ATP to form a ubiquitin-adenylate intermediate at the terminal carboxyl group of ubiquitin. A ubiquitin-thioester is then formed at the active site cysteine of the E1 enzyme, with release of AMP. The E1 enzyme transfers ubiquitin to the active site cysteine of the highly conserved family of E2 proteins, maintaining a ubiquitin-thioester linkage. The E3 ubiquitin-ligases interact with both the E2 proteins as well as with substrate proteins. Whereas several classes of E3 enzymes are now recognized (SCF, APC, CBC, RING, CHIP, and HECT E3s), they fall into two groups with respect to overall mechanism. The first group, which includes all classes except for the HECT E3s, are thought to function essentially as docking proteins, orienting the E2 and substrate so that ε-amino groups of substrate lysines can attack the E2-ubiquitin thioester, forming a ubiquitin-isopeptide bond. HECT E3s participate more directly in catalysis and are defined by a conserved C-terminal domain of ~350 amino acids (the HECT domain, Ref. 2). Like the E1 and E2 proteins, HECT E3s also have an active site cysteine that forms a ubiquitin-thioester intermediate, accepting ubiquitin from an activated E2 enzyme (3). Substrates are then ubiquitinated by transfer from the E3 to the substrate, rather than from the E2 to the substrate. The HECT domain can be recharged with ubiquitin while the substrate remains bound, and therefore multiple ubiquitins can be ligated to the substrate.

Human E6AP was the first HECT E3 to be identified and characterized (4, 5). The human papillomavirus (HPV) E6 protein of the cancer-associated HPV types (e.g. types HPV16 and 18) binds to and redirects the activity of E6AP toward a set of cellular proteins that are not normally targeted by E6AP (5). The most notable of these substrates is the p53 tumor suppressor, although several cellular proteins have been reported to be ubiquitinated by the E6/E6AP complex (reviewed in Ref. 6). Whereas few of the normal physiologic (E6-independent) functions or substrates of E6AP have been identified (7–9), it is now clear that mutations in the gene or alteration in expression of human E6AP (gene designation UBE3A) is the cause of Angelman syndrome (AS), a severe neurological disorder (10, 11). In a remarkable case of tissue specific imprinting, only the maternal allele of E6AP/UBE3A is expressed in subregions of the brain, including the hippocampal and Purkinje neurons (12, 13). AS patients lack a functional maternal allele of E6AP, and therefore lack E6AP protein in these regions of the brain. The AS phenotype is therefore hypothesized to reflect the lack of E6AP-mediated ubiquitination of one or more proteins in the brain. The identification of these substrates and an understanding of the biochemistry of E6AP and HECT E3s, in general, are critical to understanding the molecular basis of this devastating disease.

HECT E3s are found in all eukaryotic organisms, and the human genome encodes ~50 HECT E3s (2). HECT E3 proteins range in size from 90 to over 500 kDa, with the ~350 amino acid HECT domain being the only defining element. The HECT domain is always located at the C terminus, with the active site cysteine located 32–34 amino acids from the end of the protein. The Saccharomyces cerevisiae genome encodes five HECT E3s.
The best characterized is Rsp5p, which contains three WW domains in the central portion of the protein. WW domains are protein-protein interaction motifs with affinity for proline-rich containing ligands (14), and in some cases the WW domains mediate direct substrate binding (15, 16). In other cases, as in the targeting of Spt23p by Rsp5p, enzyme-substrate recognition may be mediated by additional factors because an obvious WW domain recognition element does not exist in Spt23p. Yet the WW domains are required for Spt23p targeting (17). Perhaps related to the disordered nature of the C terminus, crystals of the WWP1 HECT domain could not be obtained until the last five amino acids were deleted (20). The last six amino acids of E6AP (residues 847–852; KGFGML) contain only two residues that are conserved among most HECT E3s, Gly and Phe (see Fig. 1A). For purposes of comparison between different HECT E3s, we often refer to conserved residues by their position relative to the C terminus (e.g. Phe404, the fourth from last amino acid of E6AP, will be −4F). Whereas the −4F residue is conserved among almost all HECT E3s, in some cases it is found at either the −2 or −5 position. The exceptions to this are a small number of HECT E3s that have an extended and divergent C-terminal sequence (e.g. human Herc1, Herc2; see Fig. 1A). This subgroup of HECT E3s is not a subject of the current study, primarily because functional assays for these proteins have not been established. Four of five S. cerevisiae HECT E3s (Rsp5p, Ufd4p, Hul5p, and Tom1p) possess the conserved −4F, while Hul4p has a tyrosine at the −4 position that, as suggested by results presented below, is likely to functionally substitute. We show here that −4F is the only determinant within the last six amino acids that is specifically required for the final step in catalysis of substrate ubiquitination by both human E6AP and yeast Rsp5p, that its precise positioning is critical to function, and that C-terminal extensions to the HECT domain result in a similar biochemical defect as alterations of the −4F. We present a model for possible function of the −4F in promoting substrate ubiquitination.

**EXPERIMENTAL PROCEDURES**

**Plasmids and Mutagenesis—**The WBP2 open reading frame was amplified by PCR from a human cDNA clone (gift of M. Sudol, Mount Sinai School of Medicine, New York) using a 5’-primer that also encoded the FLAG epitope (MDTKDDDDDK). The PCR product was cloned into vector pGEX-6p (Amersham Biosciences) for GST fusion protein expression in Escherichia coli and into pCDNA3 (Invitrogen) for in vitro translation using rabbit reticulocyte or wheat germ extract translation kits from Promega (Madison, WI). In both cases the insert was cloned using BamHI to NotI restriction sites. Mutagenesis of the RSP5 and E6AP genes was performed by PCR (Platinum Pfx DNA polymerase, Invitrogen) using specific primers encoding the indicated codon alterations, deletions, or extensions. The RSP5 constructs were cloned in pGEX-6p for bacterial expression, and E6AP mutants were cloned into pVL1393-GST for creation of recombinant baculovirus expressing GST-E6AP fusion proteins. Recombinant baculoviruses were created with the BaculoGold system (BD Pharmingen).

**Protein Expression and Purification—**GST-Rsp5p fusion proteins were expressed in E. coli by standard methods and affinity-purified on glutathione-Sepharose (Amersham Biosciences). Rsp5p proteins were cleaved from GST using PreScission protease (Amersham Biosciences) under the manufacturer’s recommended conditions. GST-E6AP proteins were expressed in High5 insect cells (Invitrogen) using recombinant baculoviruses. Cell lysates were made 36-h post-infection in Non-ident P-40 lysis buffer (100 mM Tris, pH 7.5, 100 mM NaCl, 1% Nonident P-40, 1 mM phenylmethylsulfonyl fluoride, 1 mM DTT), affinity-purified on glutathione-Sepharose, and cleaved from GST with PreScission protease. GST-ubiquitin was purified from E. coli lysate on glutathione-Sepharose and eluted with PBS containing 25 mM reduced glutathione. Human E1 ubiquitin-activating enzyme was expressed in insect cells infected with recombinant E1-expressing baculovirus. The protein was partially purified by DEAR batch chromatography, eluted with buffer containing 400 mM NaCl. Arabidopsis thaliana Ubc8 protein was prepared as described previously (5). HPV-33E6 protein was expressed as a GST fusion protein in the baculovirus/insect cells system, purified on glutathione-Sepharose, and cleaved from GST with PreScission protease. In vitro translation reactions were carried out in coupled transcription-translation reactions using either rabbit reticulocyte lysate or wheat germ systems (Promega) and 35S-labeled methionine.

**In Vitro Ubiquitin-Thioester and Ubiquitination Assays—**For thioester assays, E6AP or Rsp5p proteins (1 μg) were incubated in 50-μl reactions containing 25 mM Tris (pH 7.5), 50 mM NaCl, 0.1 mM DTT, 4
mM MgCl₂, 4 mM ATP. Reactions contained E1 and A. thaliana Ubc8 enzymes (250 ng each), with or without 2 μg of GST-ubiquitin. The reactions were incubated at room temperature for 10 min, then divided into two portions, and stopped with SDS-PAGE loading buffer, with or without 100 mM DTT. Samples were electrophoresed on 8% SDS-PAGE gels, transferred to polyvinylidene difluoride membrane, and immuno-blotted with either anti-RSP5 mouse antibody or anti-E6AP rabbit antibody. Secondary antibodies were HRP-conjugated anti-mouse or anti-rabbit IgG and visualized with the SuperSignal chemiluminescence detection system (Pierce). Ubiquitination assays are carried out under the same conditions as the ubiquitin-thioester assays but for 30 min, and in presence of unmodified bovine ubiquitin (Sigma). 500 ng of purified bacterially expressed and purified FLAG-WBP2 was used for Rsp5p ubiquitination assays, where indicated. The reaction products were analyzed by SDS-PAGE followed by autoradiography or immunoblotting using an anti-FLAG antibody (F3165 Sigma).

In Vivo Complementation Assays—As described previously (18), the rap5-1 strain, FW1808 (MATα, his4-912Δ5, lys2-128Δ, ura3-52, rap5-1), was used to assay for the ability of plasmid-expressed RSP5 variants to complement the growth arrest of this strain at 36 °C. All of the variants shown in Fig. 2 were subcloned into the pYES2 vector (Invitrogen) under control of the GAL1 promoter, and all contained an N-terminal hemagglutinin epitope. pYES2 plasmids were transformed into FW1808 or the closely related RSP5 strain FY56 (MATα, his4-912Δ, lys2-128Δ, ura3-52), with selection on uracil drop-out plates. Growth at 36 °C was scored on uracil drop-out plates with 2% galactose as the carbon source. Protein expression for all variants was assayed by preparing total cell extracts from cultures grown at 30 °C in 2% galactose containing uracil drop-out liquid medium, as described previously (18). Extracts were analyzed by SDS-PAGE and immuno blotting using anti-HA and anti-Rep5p antibodies.

RESULTS
Conservation of the −4F and Its Position in the Predicted E6AP-UbcH7-ubiquitin Structure—Fig. 1A shows the C-termi-

---

**FIG. 1. Structural context of the −4F.** A, alignment of C-terminal sequences of various human HECT domains and all five *S. cerevisiae* HECT domains. The active site cysteine is indicated in **red** and the conserved −4F is in **green**. The most highly conserved residues of this region are in **black** and **bold**. The region where Herc1 and Herc2 diverge from other HECT domains is **boxed**. The secondary structure over this region, as determined from the x-ray crystal structure of E6AP, is indicated above the E6AP sequence (**S**, β-sheet; **H**, α-helix). B, the structure of the E6AP HECT domain-UbcH7 complex (19), with ubiquitin modeled into the structure, based on the *S. cerevisiae* Ubc1-ubiquitin structure (22). The N lobe of the HECT domain is in **dark blue**, the C lobe is in **light blue**, UbcH7 is in **red**, and ubiquitin is in **white**. The 41 Å line connects the thiol groups of the active site cysteines of the UbcH7 (86C) and E6AP (820C). The distance from the HECT active site cysteine to the −4F side chain is also indicated.

---

mm MgCl₂, 4 mM ATP. Reactions contained E1 and A. thaliana Ubc8 enzymes (250 ng each), with or without 2 μg of GST-ubiquitin. The reactions were incubated at room temperature for 10 min, then divided into two portions, and stopped with SDS-PAGE loading buffer, with or without 100 mM DTT. Samples were electrophoresed on 8% SDS-PAGE gels, transferred to polyvinylidene difluoride membrane, and immuno-blotted with either anti-RSP5 mouse antibody or anti-E6AP rabbit antibody. Secondary antibodies were HRP-conjugated anti-mouse or anti-rabbit IgG and visualized with the SuperSignal chemiluminescence detection system (Pierce). Ubiquitination assays are carried out under the same conditions as the ubiquitin-thioester assays but for 30 min, and in presence of unmodified bovine ubiquitin (Sigma). 5 μl of *in vitro* translated WBP2 (rabbit reticulocyte lysate) or p53 (wheat germ extract) was used as source of substrates. 500 ng of purified bacterially expressed and purified FLAG-WBP2 was used for Rep5p ubiquitination assays, where indicated. The reaction products were analyzed by SDS-PAGE followed by autoradiography or immunoblotting using an anti-FLAG antibody (F3165 Sigma).
nal amino acid sequence of E6AP, beginning at Thr\(^{819}\), immediately preceding the active site cysteine (Cys\(^{820}\)), and continuing to the end of the protein at Leu\(^{825}\) (numbering according to protein isoform 1 of E6AP, Ref. 21). This is aligned with several other human HECT E3s and all five of the \(S.\ cerevisiae\) HECT E3s. Excluding Herc1 and Herc2, the most C-terminal of the highly conserved residues is a phenylalanine located either two, four, or five amino acids from the end of the protein. There are a small number of HECT E3s that have a tyrosine at the \(H11002\) position (KIAA1320 and Hul4p in Fig. 1A). The amino acid preceding the phenylalanine is generally a glycine, but 20–30\% of HECT E3s have various substitutions at this position (e.g. Ufd4p, Hul5p in Fig. 1A). Herc1 and Herc2 diverge from other family members over the last ten amino acids relative to E6AP, and continue with longer C-terminal extensions. Herc1 and Herc2 (the homolog of mouse runty-jerky-sterile; rjs) are closely related proteins, containing multiple RCC repeats, and to our knowledge, all HECT E3s that contain such divergent C-terminal sequences also contain RCC repeats. However, some RCC-HECT E3s have the more conventional C-terminal sequence containing the \(-4F\) (e.g. Herc3, cyclin E-binding protein, KIAA1593).

The 2.7 Å structure of the human E6AP HECT domain (amino acids 493–852) in complex with the UbcH7 E2 enzyme was reported previously (19). Ubiquitin was not present in this complex, so its position when bound to either the E2 or HECT E3 was unknown. A model of the \(S.\ cerevisiae\) Ubc1p E2 enzyme in complex with ubiquitin has been reported more recently (22). The structures of UbcH7 and Ubc1p are very similar and both belong to the subgroup of E2 enzymes known to function with HECT E3s. By aligning the E2 structures from the HECT-E2 and E2-ubiquitin structures it is possible to predict the position of ubiquitin within the HECT-UbcH7 structure when bound to UbcH7 (22). In this predicted ternary complex, ubiquitin lies to the side of the U-shaped structure formed by the HECT-E2 complex and does not cross the line-of-sight between the active site cysteines of the E2 and E3 (Fig. 1B). Fig. 1B also highlights the position of the \(-4F\) within the...
purification of bacterially expressed GST-Rsp5p, both wild-type and the protein relative to the wild-type control and no differences were determined to be optimal for obtaining biochemically active protein for the respective wild-type proteins. After purification, the proteins on glutathione-Sepharose the E3 proteins were cleaved from GST by site-specific proteolysis. The right panel (lanes 5–8) shows a similar purification of bacterially expressed GST-Rsp5p, both wild-type and the C-A mutant.

predicted HECT-E2-ubiquitin structure.

Analysis of C-terminal Mutations in E6AP and Rsp5p—To explore the function of the HECT C terminus in ubiquitination, a set of mutations was created in the context of both full-length human E6AP and yeast Rsp5p (Fig. 2). The set of mutants included single amino acid alterations within the last five amino acids, small deletions from the C terminus, or extensions to the C terminus. E6AP and Rsp5p proteins were expressed as GST fusion proteins, either in a baculovirus insect cell system for E6AP, or in bacteria for Rsp5p. These expression systems were determined to be optimal for obtaining biochemically active protein for the respective wild-type proteins. After purification of the proteins on glutathione-Sepharose the E3 proteins were cleaved from GST by site-specific proteolysis. Fig. 3 shows a Coomassie Blue-stained SDS-PAGE gel of purified E6AP and Rsp5p, both of the wild-type and the active site cysteine to alanine (C-A) mutated protein. All of the E6AP and Rsp5p proteins included in our analyses expressed similar amounts of protein relative to the wild-type control and no differences were observed in solubility of the altered proteins.

Purified E6AP proteins were assayed in vitro for ubiquitin thioester formation. The negative control in these experiments was the C820A mutant, which disrupts the active site cysteine and therefore eliminates the ability to form the ubiquitin-thioester intermediate (3). Fig. 4A shows ubiquitin-thioester formation for most of the E6AP mutants shown in Fig. 2, utilizing GST-ubiquitin as the source of ubiquitin to clearly distinguish the 100 kDa E6AP protein from the thioester-linked ubiquitinated intermediate. In control experiments not shown, we showed that the GST moiety did not affect efficiency of ubiquitin thioester formation. Assays were performed by incubating purified E6AP proteins with ATP, partially purified baculovirus-expressed human E1 protein, E2 protein expressed in bacteria (A. thaliana UBC8, Ref. 23), and GST-ubiquitin purified from E. coli. The reactions were stopped with SDS-PAGE loading buffer without or with DTT, analyzed by SDS-PAGE, and immunoblotted with anti-E6AP antibody. Approximately 20% of the input wild-type E6AP was incorporated into a thioester-linked complex with GST-ubiquitin under these assay conditions, with the adduct only being detected in the absence of reducing agent (DTT). No adduct was seen when GST-ubiquitin was omitted from the reaction, and as expected no E3-ubiquitin adduct was seen with the C820A active site mutant. All of the C-terminal variants of E6AP shown in Fig. 2 were still able to form the thioester intermediate to a similar degree as wild-type E6AP. These results were consistent with our earlier observation that E6AP deleted of the last six amino acids was still able to form the ubiquitin-thioester intermediate (2).

The same proteins were assayed for their ability to ubiquitinate p53 in the presence of the cancer-associated HPV33 E6 protein (closely related to HPV16 E6). HPV33 E6 was expressed as a GST protein using recombinant baculovirus and was cleaved from the purified GST by site-specific proteolysis. 35S-labeled p53 was synthesized in vitro in a wheat germ extract system, which lacks endogenous E6AP. Purified E6 and E6AP proteins were added to the complete p53 translation reaction, along with supplemental E1 and E2 enzymes. As shown in Fig. 4B, no ubiquitinated p53 was seen when E6 was omitted from the reaction, whereas over 70% of the input p53 was ubiquitinated in the presence of both E6 and wild-type E6AP. GST-ubiquitin functioned similarly in this assay (not shown). The active site cysteine mutant (C820A) did not multi-ubiquitinate p53, as expected, although a small amount of mono- or di-ubiquitinated species were detected. This was seen with all mutants that were negative for p53 multi-ubiquitination and may be a result of a small amount of ubiquitination being catalyzed by the E2 in the context of the [E6AP-E6-p53]-E2 complex. Deletion of one, two, or three amino acids from the C terminus of E6AP (851stop, 850stop, 849stop) did not block the ability of E6AP to multi-ubiquitinate p53, whereas deletion of six amino acids (not shown; see Ref. 2) effectively blocked ubiquitination. Alteration of the −4F to either tyrosine (F849Y) or alanine (F849A; not shown) blocked p53 ubiquitination as effectively as the active site cysteine mutant (C820A), whereas mutation of the −5G to alanine (G848A) had no effect on the ability to ubiquitinate p53. To determine whether the identity of the last one or two amino acids affected substrate ubiquitination, the terminal leucine residue was altered to either alanine or proline, or the last two amino acids (ML) were altered to QE. The latter alteration results in the last four amino acids of E6AP being identical to those of Rsp5p. These altered proteins ubiquitinated p53 as effectively as wild-type E6AP. Together, these results indicate that the −4F of E6AP is the single major determinant within the last six amino acids necessary for substrate ubiquitination by E6AP (see Fig. 2 for a tabulation of these results).

Most Angelman syndrome mutations that result in single amino acid changes within the E6AP open reading frames are found within the region encoding the HECT domain, and all mutations tested so far disrupt to some degree the catalytic activity of E6AP in vitro (Ref. 19).2 One Angelman syndrome mutation has been identified that affects the stop codon of E6AP, resulting in a terminally extended protein (24). To determine whether E6AP with a C-terminal extension would behave similarly to variants where the −4F is altered, the stop codon of E6AP was deleted within the E6AP cDNA sequence, resulting in a protein that is extended by eight amino acids (NKTQKNKT; E6AP-Ext8, see Fig. 2). As shown in Fig. 4, this protein, like the −4F to A mutant, could form a ubiquitin-thioester intermediate but could not catalyze p53 ubiquitination. To determine whether this was because of the actual sequence of the extension or if another extension would have a similar effect, the FLAG epitope (DTKDDDDK) was placed at the C terminus of E6AP. Again, substrate ubiquitination was blocked by this mutation without affecting thioester formation. These results suggest that C-terminal extensions interfere with the function of the −4F in ubiquitination. A similar set of C-terminal mutants was made in Rsp5p and assayed for thioester formation and substrate ubiquitination. In addition to confirming the generality of our results with a second HECT E3, the advantage of performing these experiments with Rsp5p was that a genetic system exists for easily

2 C. Salvat and J. M. Huibregtse, unpublished results.

FIG. 3. Expression and purification of E6AP and Rsp5p proteins. Left panel, GST-E6AP, both wild-type (WT) and the active site cysteine mutant (C-A), were purified from insect cell extract on glutathione-Sepharose. An aliquot of the bead-bound protein was loaded on an SDS-PAGE gel and stained with Coomassie Blue (lanes 1 and 3). Lanes 2 and 4 show an aliquot of the soluble E6AP protein released from the glutathione-Sepharose beads after site-specific proteolysis with PreScission protease. The right panel (lanes 5–8) shows a similar purification of bacterially expressed GST-Rsp5p, both wild-type and the C-A mutant.
assessing in vivo function. As reported previously, it is difficult to detect DTT-sensitive ubiquitin adducts in vitro with wild-type Rsp5p because Rsp5p very efficiently ubiquitinates itself at internal lysine residues (2). This is shown in Fig. 5, where purified Rsp5p proteins were incubated with E1, E2, ATP, and GST-ubiquitin and the products were analyzed by SDS-PAGE in the absence or presence of DTT, followed by immunoblotting with anti-Rsp5p antibody. A ladder of high molecular ubiquitinated Rsp5p products was seen with wild-type Rsp5p, and these conjugates remained after incubation with high concentrations of DTT, as expected for isopeptide-linked multi-ubiquitin conjugates. In contrast, the active site mutant (C777A) did not form any conjugates, either in the absence or presence of DTT. Whereas there is no evidence that the self-ubiquitination reaction is biologically meaningful, this in vitro assay simultaneously reflects the ability of the protein to form a thioester intermediate and to catalyze isopeptide-linked multi-ubiquitin conjugates. In contrast, the active site mutant (C777A) did not form any conjugates, either in the absence or presence of DTT. Therefore, we would expect to see only a DTT-sensitive mono-ubiquitinated conjugate in this assay, and this was indeed the case (Fig. 5), although a small amount of self-ubiquitination was seen for the F806A and F806P proteins. Whereas tyrosine could not substitute for phenylalanine in E6AP, the analogous F806Y mutation in Rsp5p resulted in a protein that could catalyze self-ubiquitination. This suggests that at least in some HECT E3s a tyrosine can functionally substitute for phenylalanine, consistent with the fact that a small number of HECT E3s do contain a tyrosine at the −4F position (Fig. 1A). Rsp5p deleted of the last three amino acids of Rsp5p (806stop) inactivated Rsp5p self-ubiquitination activity in vitro, in contrast to the equivalent mutation in E6AP, which did not affect p53 ubiquitination. The 806Astop mutation, which combines the −4F to alanine mutation with truncation of the last three amino acids, as well as further truncations (805stop and 803stop), were also negative for self-ubiquitination. Rsp5p with a FLAG epitope at the C terminus retained self-ubiquitination activity, although the level of activity was consistently decreased relative to wild type.

Because requirements for the self-ubiquitination reaction may not necessarily be identical to those for ubiquitination of exogenous substrates, we examined the ability of the Rsp5p mutants to ubiquitinate human WBP2. WBP2 was previously shown to interact with multiple WW domain proteins (25) and was isolated independently in a biochemical screen for substrates of Nedd4 and Rsp5p.2 In the experiment shown in Fig. 6A, WBP ubiquitination assay was assayed using in vitro translated WBP2 ([35S]methionine labeled in rabbit reticulo-
cyte lysate), adding purified Rsp5p proteins along with purified E1 and E2 proteins (A. thaliana Ubc8), ATP, and ubiquitin. WBP2 was very efficiently ubiquitinated in the presence of wild-type Rsp5p, whereas the C777A mutant (active site mutant) was completely inactive. All of the other altered proteins behaved similarly against WBP2 as they did in the self-ubiquitination assay, including the -4F to Ala mutant (F806A), which again showed a very small amount of activity against WBP2. The C-terminal FLAG tag blocked WBP2 ubiquitination significantly, consistent with the results of the self-ubiquitination assay. Overall, within this set of Rsp5p proteins the only mutant that disrupted the -4F that retained strong ubiquitination activity was the -4F to tyrosine (Tyr806) variant.

To address the importance of the positioning of the conserved -4F in Rsp5p, secondary mutations were introduced into the F806A mutant, placing a phenylalanine at either the -5 position (806A/805F), the -3 position (806A/807F), or as the last amino acid (806A/809F). In the experiment shown in Fig. 6B, the ubiquitination assay was performed using FLAG-WBP2 with detection by immunoblotting with anti-FLAG antibody. In this experiment we scored for disappearance of the input WBP2 protein, because high molecular weight ubiquitin conjugates of WBP2 were very inefficiently detected in this assay. This is the result of masking of the FLAG epitope of WBP2 following transthiolation, and isopeptide bond catalysis) the -4F residue.

The results presented here establish that the conserved -4F, near the C terminus of the HECT domain, plays an essential role in promoting substrate ubiquitination. In terms of the three step model for HECT E3 catalysis (binding of the E2, transthiolation, and isopeptide bond catalysis) the -4F residue...
is the first determinant identified that is required solely for the last step in the reaction cycle. In some HECT E3s a tyrosine is found at the −4 position, and indeed a tyrosine functionally substituted for the phenylalanine both in in vitro reactions and in vivo for Rsp5p. Together, our results suggest that the aromatic ring at this position is critical for the ubiquitination function. Whereas the crystal structure of E6AP indicates that the last few amino acids are disordered, the placement of the −4F was found to be crucial, as moving it to either the −5 or −3 position or terminal position in Rsp5p reduced ubiquitination in vitro and resulted in loss of in vivo complementation activity.

The only difference between the set of E6AP and Rsp5p mutants that is not readily explainable was the fact that Rsp5p was not functional when the protein terminated with the conserved phenylalanine (the 806stop mutant), unlike the case for E6AP (the 849stop mutant). We feel that it is unlikely that this indicates that there is another specific determinant within the last three amino acids of Rsp5p that is necessary for substrate ubiquitination given that these residues are highly divergent among HECT E3s.

What, specifically, might the −4F do to promote transfer of ubiquitin from the HECT active site cysteine to the substrate? The final step in the ubiquitination reaction is the nucleophilic attack on the ubiquitin-thioester carbonyl carbon by the ε-amino nitrogen of the lysine residue that is to be ubiquitinated. Given the separation of the −4F from the active site cysteine in the crystal structure (−18 Å; see Fig. 1B), and the nature of the phenylalanine side chain, it would seem unlikely that the −4F participates directly in catalysis. Assuming that the overall structure of the C lobe of the HECT domain is similar in the absence and presence of bound ubiquitin, we favor a model in which the −4F may contact ubiquitin whereas it is linked to the active site cysteine and play a role in orienting and positioning the tethered ubiquitin molecule. The active site cysteine of E6AP lies in a four-residue loop at the base of the junction formed by the N and C lobes of the HECT domain. When bound to the cysteine thiol, the orientation of the bulk of the ubiquitin molecule may be critical for allowing access of the incoming substrate lysine to the thioester bond. We propose that the −4F may contact ubiquitin and position it in an optimal orientation to allow for approach of the incoming lysine side chain. Thus, alteration of the −4F may prevent substrate ubiquitination by mis-orienting ubiquitin and thereby blocking access of the substrate lysine to the active site.

Whereas the above model is speculative, it predicts that specific amino acid substitutions on the surface of ubiquitin that mediate interaction with the −4F would result in a similar ubiquitination defect. Similarly, specific amino acid substitutions on the surface of ubiquitin might suppress the effect of alteration of the −4F. Given the nature of the phenylalanine side chain, hydrophobic patches on the surface of ubiquitin might be good candidate sites for −4F interaction. We have preliminarily tested alterations at hydrophobic surface residues of ubiquitin described by Beal et al. (26) at residues Leu6, Ile44, and Val70; purified proteins kindly provided by C. Pickart. The altered ubiquitin molecules could be conjugated to substrates by E6AP and Rsp5p only slightly less efficiently than wild-type ubiquitin, however the defects were proportional to their ability to form the ubiquitin-thioester intermediate (data not shown). Thus, these particular hydrophobic surface alterations did not mimic the effect of altering the −4F. Additional surface residues of ubiquitin are currently being investigated.

It has been suggested that the terminal C group of the HECT domain might participate in catalysis by acting as a general base to facilitate deprotonation of the incoming lysine residue side chain of the substrate (20). Thus, the terminal carboxylate might be a second determinant necessary specifically for substrate ubiquitination, and it is conceivable that this might also be linked to or cooperate in some way with −4F function. Rigorous testing of a potential role of the terminal carboxylate in ubiquitination will require chemical modification of the carboxyl group in an otherwise normal HECT E3. With respect to this hypothesis, however, it is perhaps interesting that eight amino acid C-terminal extensions to E6AP resulted in proteins with identical biochemical characteristics to the −4F to A mutants of these proteins: they could form the ubiquitin-thioester intermediate but could not transfer ubiquitin to the substrate. Therefore, if there is cooperation between the −4F and the terminal carboxylate, the separation between the phenylalanine side chain and the carboxylate might be expected to be critical. Alternatively, if the −4F contacts ubiquitin, the affect of the C-terminal extensions might be explained by a simple steric inhibition of the extended sequence on the −4F-ubiquitin interaction.

A single Angelman syndrome patient was identified in which the maternal E6AP allele contained a 15-bp deletion beginning with the last nucleotide before the stop codon and extending into the 3′-untranslated region (3142del15) (24). The investigators who identified this mutation suggested that the predicted protein would terminate in a string of lysines encoded by the poly(A) tail of mRNA (4); however, rigorous characterization of E6AP mRNAs (27) indicates the protein predicted from the 3142del15 alteration would be extended by 39 amino acids. Whereas this 39 amino acid-extended protein has not been analyzed biochemically, we predict, based on the results presented here, that if it is able to form the ubiquitin-thioester intermediate it would be unable to transfer ubiquitin to substrates. Indeed, Rsp5p with a large C-terminal extension (an −190 amino acid “TAP” tag; Ref. 28) is still able to form a ubiquitin-thioester in vitro, but cannot transfer that ubiquitin to itself in a self-ubiquitination assay (not shown). Together, this suggests that the protein encoded by the 3142del15 Angelman syndrome-associated mutation is likely to be specifically defective in catalysis of ubiquitin-isopeptide conjugates.

The HECT domain −4F is the first determinant to be identified that is required for the last step in transfer of ubiquitin to substrate, but is not required for upstream steps in the enzymatic cycle. Whereas other such determinants may yet be identified, the ability to separate the transattachment step from the ubiquitination step in vitro will permit more detailed biochemical investigations of each of both of these steps. In addition, in cases where there may be an advantage to therapeutically inactivating HECT E3s, such as in HPV E6-expressing cervical cancers and their associated metastases, the identification of the essential role of the −4F highlights a specific and exposed target for potential drug interaction.

Acknowledgments—We thank Melissa Kelley and Sylvie Beaudenot for critical reading of the manuscript.

REFERENCES
1. Pickart, C. M. (2001) Annu. Rev. Biochem. 70, 503–533
2. Huibregtse, J. M., Scheffner, M., and Howley, P. M. (1995) Proc. Natl. Acad. Sci. U. S. A. 92, 2563–2567
3. Scheffner, M., Nuher, U., and Huibregtse, J. M. (1995) Nature 373, 81–83
4. Huibregtse, J. M., Scheffner, M., and Howley, P. M. (1994) Cold Spring Harbor Symp. Quant. Biol. 59, 237–245
5. Scheffner, M., Huibregtse, J. M., Vierstra, R. D., and Howley, P. M. (1993) Cell 72, 495–505
6. Menger, K., and Howley, P. M. (2002) Virus Res. 89, 213–228
7. Kumar, S., Tais, A. L., and Howley, P. M. (1999) J. Biol. Chem. 274, 18785–18792
8. Chu, H., Kumar, S., and Howley, P. M. (1999) Proc. Natl. Acad. Sci. U. S. A. 96, 9557–9562
9. Kuhne, C., and Banks, L. (1998) J. Biol. Chem. 273, 34302–34309
10. Matsuzawa, T., Sottoliff, J. S., Fang, P., Galjaard, R. J., Jiang, Y. H., Benton, C. S., Rommens, J. M., and Beaudet, A. L. (1997) Nat. Genet. 15, 74–77
11. Kishino, T., Lalande, M., and Wagstaff, J. (1997) Nat. Genetics 15, 70–73
12. Yu, T. H., and Hoffman, A. R. (1997) Nat. Genet. 17, 12–13
13. Albrecht, U., Sutcliffe, J. S., Cattanach, B. M., Beechey, C. V., Armstrong, D.,
    Eichele, G., and Beaudet, A. L. (1997) Nat. Genetics 17, 75–78
14. Macias, M. J., Wiesner, S., and Sudol, M. (2002) FEBS Lett. 513, 30–37
15. Zhu, H., Kavsak, P., Abdollah, S., Wrana, J. L., and Thomsen, G. H. (1999)
    Nature 400, 687–693
16. Huibregtse, J. M., Yang, J. C., and Beaudenon, S. L. (1997) Proc. Natl. Acad.
    Sci. U. S. A. 94, 3656–3661
17. Hoppe, T., Matuschewski, K., Rake, M., Schlenker, S., Ulrich, H. D., and
    Jentsch, S. (2000) Cell 102, 577–586
18. Wang, G., Yang, J., and Huibregtse, J. M. (1999) Mol. Cell. Biol. 19, 342–352
19. Huang, L., Kinnucan, E., Wang, G., Beaudenon, S., Howley, P. M., Huibregtse,
    J. M., and Pavletich, N. P. (1999) Science 286, 1321–1326
20. Verdecia, M. A., Juarez, C. A., Wells, N. J., Ferrer, J. L., Bowman, M. E.,
    Hunter, T., and Noel, J. P. (2000) Mol. Cell 11, 249–259
21. Yamamoto, Y., Huibregtse, J. M., and Howley, P. M. (1997) Genomics 41,
    263–266
22. Hamilton, K. S., Ellison, M. J., Barber, K. R., Williams, R. S., Huzil, J. T.,
    McKenna, S., Ptak, C., Glover, M., and Shaw, G. S. (2001) Structure (Cam-
    bridge) 9, 897–904
23. Girod, P. A., Carpenter, T. B., van Noecker, S., Sullivan, M. L., and Vierstra,
    R. D. (1993) Plant J. 3, 545–552
24. Fang, P., Lev-Lehman, E., Tsai, T. F., Matsuura, T., Benton, C. S., Sutcliffe,
    J. S., Christian, S. L., Kubota, T., Halley, D. J., Meijers-Heijboer, H.,
    Langlois, S., Graham, J. M., Jr., Beuten, J., Willems, P. J., Ledbetter, D. H.,
    and Beaudet, A. L. (1999) Hum. Mol. Genet. 8, 129–135
25. Chen, H. I., Einbond, A., Kwaak, S. J., Linn, H., Koepf, E., Peterson, S., Kelly,
    J. W., and Sudol, M. (1997) J. Biol. Chem. 272, 17070–17077
26. Beal, R., Deveraux, Q., Xia, G., Rechsteiner, M., and Pickart, C. (1996) Proc.
    Natl. Acad. Sci. U. S. A. 93, 861–866
27. Kishino, T., and Wagstaff, J. (1998) Genomics 47, 101–107
28. Puig, O., Caspary, F., Rigaut, G., Rutz, B., Beauceret, E., Bragade-Nilsson, E.,
    Wilm, M., and Seraphin, B. (2001) Methods 24, 218–229
The -4 Phenylalanine Is Required for Substrate Ubiquitination Catalyzed by HECT Ubiquitin Ligases
Catherine Salvat, Guangli Wang, Anahita Dastur, Nancy Lyon and Jon M. Huibregtse

J. Biol. Chem. 2004, 279:18935-18943.
doi: 10.1074/jbc.M312201200 originally published online February 13, 2004

Access the most updated version of this article at doi: 10.1074/jbc.M312201200

Alerts:
• When this article is cited
• When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 28 references, 10 of which can be accessed free at http://www.jbc.org/content/279/18/18935.full.html#ref-list-1