Research Article

Redefining the catalytic HECT domain boundaries for the HECT E3 ubiquitin ligase family

Emma I. Kane, Steven A. Beasley, Johanna M. Schafer, Justine E. Bohl, Young Sun Lee, Kayla J. Rich, Elizabeth F. Bosia and Donald E. Spratt

Gustaf H. Carlson School of Chemistry and Biochemistry, Clark University, 950 Main Street, Worcester, MA 01610, U.S.A.

Correspondence: Donald E. Spratt (dspratt@clarku.edu)

There are 28 unique human members of the homologous to E6AP C-terminus (HECT) E3 ubiquitin ligase family. Each member of the HECT E3 ubiquitin ligase contains a conserved bilobal HECT domain of approximately 350 residues found near their C-termini that is responsible for their respective ubiquitylation activities. Recent studies have begun to elucidate specific roles that each HECT E3 ubiquitin ligase has in various cancers, age-induced neurodegeneration, and neurological disorders. New structural models have been recently released for some of the HECT E3 ubiquitin ligases, but many HECT domain structures have yet to be examined due to chronic insolubility and/or protein folding issues. Building on these recently published structural studies coupled with our in-house experiments discussed in the present study, we suggest that the addition of ∼50 conserved residues preceding the N-terminal to the current UniProt defined boundaries of the HECT domain are required for isolating soluble, stable, and active HECT domains. We show using in silico bioinformatic analyses coupled with secondary structural prediction software that this predicted N-terminal α-helix found in all 28 human HECT E3 ubiquitin ligases forms an obligate amphipathic α-helix that binds to a hydrophobic pocket found within the HECT N-terminal lobe. The present study brings forth the proposal to redefine the residue boundaries of the HECT domain to include this N-terminal extension that will likely be critical for future biochemical, structural, and therapeutic studies on the HECT E3 ubiquitin ligase family.

Introduction

In 1995, a novel class of enzymes containing an approximate 350 residue conserved domain at their respective C-terminus were found to be homologous to the C-terminal domain of E6-associated protein (E6AP) and had the capabilities of forming a thioester bond with ubiquitin [1]. The most unique aspect of this study revealed that a highly conserved catalytic cysteine residue within this C-terminal domain was responsible for polyubiquitin chain formation. This discovery led to the classification of the homologous to E6AP C-terminus (HECT) E3 ubiquitin ligase family, with currently 28 human members that differ in size and the presence of various N-terminal protein–protein interaction domains involved in substrate recruitment [2]. Further investigation through structural analysis revealed this 350 residue domain was bilobal with numerous conserved residues within the hinge region found between the N- and C-terminal lobes that provide flexibility to support HECT-dependent ubiquitylation activity [3]. This cleft region is suggested to also be critical to its HECT function as loss-of-function mutations in E6AP within this region have been linked to Angelman syndrome [4]. Both lobes were also identified to have unique characteristics involved in ubiquitin handling by the HECT E3 ubiquitin ligases – where the N-terminal lobe is responsible for E2 ubiquitin conjugating enzyme recruitment and binding, while the C-terminal lobe contains the highly conserved catalytic cysteine residue that facilitated ubiquitin transfer onto a specific substrate [5,6]. Conservation within the hinge between the N- and C-terminal lobes have been suggested to allow for the flexible rotation to bring the catalytic cysteine residue in the HECT C-terminal lobe into close proximity.
with the E2–ubiquitin complex bound to the HECT N-terminal lobe [6]. Studies have also suggested a competing proximal addition model where the HECT domain can form higher oligomeric structures to facilitate polyubiquitin chain formation [7–11].

Members of the HECT E3 ubiquitin ligase family have been broadly categorized into three subfamilies based upon their domain architecture in their variable N-terminal protein–protein interaction domains and their functional similarities. These subfamilies include the neuronal precursor cell–expressed developmentally down-regulated 4 subfamily (NEDD4, 9 members), the HECT and RLD domain subfamily (HERC, 6 members), and the unclassified ‘other’ subfamily (13 members) [12]. The NEDD4 E3 ubiquitin ligases possess a Ca2+-binding C2 domain for phospholipid recognition [13] and two-to-four WW domains required for binding specifically to proline-rich motif-containing substrates [14]. Members of the HERC subfamily are divided based upon size, with only the large HERCs (HERC1/HERC2) containing more than one RCC1-like domain (RLD) for multiple substrate recognition and binding opportunities [15]. The remaining 13 HECT E3 ubiquitin ligases have diverse protein–protein interacting domains within their N-terminus, with some homology existing such as ankyrin repeat and armadillo repeat domains [2]. Thorough phylogenetic analyses of the HECT family suggest that it may be more appropriate to have 16 separate HECT subfamilies, including the notable division of small and large members of the HERC subfamily [16,17].

HECT E3 ubiquitin ligases are involved in a myriad of biological processes and cellular pathways including protein degradation, DNA damage repair, apoptosis, subcellular localization, and immunological response [2,18–20]. Numerous studies have implicated the dysfunction of members from the HECT E3 ubiquitin ligase family in disease development such as various cancers, immunological disorders, neurodevelopmental disorders, and neurodegenerative disease [18,21–24]. Since the HECT domain possesses the most prominent functional regions of this protein family, it is logical that a research focus would be directed towards structural and functional characterization of this catalytic region of the HECT E3 ubiquitin ligases. New structural insights into the HECT domain have shown that the introduction of extending the N-terminal sequence of the HECT domain is important for folding. While many of these crystal structures exist for HECT domains with an additional α-helix [5,25–27], the HUWE1 structure contains up to three α-helices [27].

Here, we show through bioinformatic analysis and biochemical characterization that up to four N-terminal α-helices beyond the ∼350 residue UniProt defined boundary are predicted to be present within all 28 HECT E3 ubiquitin ligases and that including these N-terminal extensions for the HECT domain improves solubility and increases ubiquitylation activity. Our bioinformatic evaluation of the various members of the HECT E3 ubiquitin ligases through multiple sequence alignment (MSA) with Jalview [28] coupled with secondary structure prediction and evaluating available PDB structures using PyMol [29] suggests there exists at least one amphipathic α-helix – with some predicted to have two to four α-helices – prior to the N-terminal lobe of the HECT domain. Our assertion is further supported by the recent crystal structures of the HECT domains of apoptosis-resistant E3 ubiquitin protein ligase-1 (AREL1, PDB: 6JX5) and ubiquitin-protein ligase E3C (UBE3C; PDB: 6K2C) HECT domains that included an additional 50 amino acids preceding the N-terminal lobe to the isolated HECT domains [25,30]. Our studies coupled with recent structural reports on different HECT domains indicate that the currently annotated UniProt HECT domain residue boundaries need to be revisited and revised.

Results
In the years since the original E6AP HECT domain structure was published in 1999, every solved HECT structure has included at least one α-helical extension immediately upstream from the defined N-terminal lobe (Figure 1) [5,6,25–27,30–33]. The additional α-helix (often referred to in the literature as α1’) sits astride the HECT N-terminal lobe on the opposite face from the E2 enzyme binding site. The ClustalW and T-Coffee MSA of the 28 human HECT paralogs resulted in an alignment that includes a JPred predicted α-helix (Figure 2).

The predicted α-helix shows a periodicity for the highly conserved hydrophobic residues resulting in their collective orientation on the same face of the α-helix. Further analysis of the HECT domains with known structures reveals these residues form a hydrophobic core along with the hydrophobic residues from three distinct α-helices found in the HECT N-terminal lobe that are also conserved (Figure 3A). For example, in the AREL1 HECT domain structure, these comprise F439, V443, F446, L450, and V453 on the α-helical extension interacting with residues L563, Y564, L691, L692, I694, F695, and L703 on the HECT N-terminal lobe creating a hydrophobic pocket (Figure 3B,C). Further structural analysis revealed that removing this protective helix exposes a contiguous hydrophobic path for all the HECT proteins analyzed. For example, AREL1 has a hydrophobic path measured to be 627 Å and WWP1 has a hydrophobic patch of 698 Å (Figure 3C). This exposed pocket may have contributed to the insolubility of the shorter
Figure 1. Structural analysis of the solved HECT domains models of various HECT E3 ubiquitin ligases

Representative extended HECT domain models within the HECT E3 ubiquitin ligase family. With the amphipathic α-helix included, all models show a similar tertiary structure to its namesake E6AP (PDB: 1C4Z). The bilobal structure is shown with the C-lobe containing the catalytic cysteine (light gray), the N-lobe with the E2~ubiquitin binding interface and sites required for oligomerization (dark gray), and the N-terminal α-helical extension (magenta). These crystal models are also missing some atoms including loops and the C-terminal tail of the HECT domain.

Figure 1. Structural analysis of the solved HECT domains models of various HECT E3 ubiquitin ligases

Representative extended HECT domain models within the HECT E3 ubiquitin ligase family. With the amphipathic α-helix included, all models show a similar tertiary structure to its namesake E6AP (PDB: 1C4Z). The bilobal structure is shown with the C-lobe containing the catalytic cysteine (light gray), the N-lobe with the E2~ubiquitin binding interface and sites required for oligomerization (dark gray), and the N-terminal α-helical extension (magenta). These crystal models are also missing some atoms including loops and the C-terminal tail of the HECT domain.

constructs. It is interesting to note that the equivalent hydrophobic patch residues of the originally published E6AP structure missing the N-terminal α-helix (PDB: 1C4Z) are found in the core of the trimeric crystal interface.

The protection of this hydrophobic pocket probably provides structural integrity to the HECT domain, which is supported by the marked increase in solubility we observed for our extended HECT domain constructs (Figure 4A). Each extension of the HECT domain we incorporated into our extended HECT (extHECT) expression constructs was based off the location of predicted boundaries of α-helices, with the inclusion of each α-helix showing improved solubility and HECT-dependent activity. Taken together, our MSA results support and further demonstrate that identical or similar residues within these corresponding positions within the predicted α-helix extension are also found in all 28 human members of the HECT E3 ubiquitin ligase family.

The HECT E3 ubiquitin ligases we assessed in this study (HECW2, HERC4, G2E3, and HACE1) showed significantly increased solubility upon induction with IPTG and higher yields when comparing our extHECT constructs to the current UniProt HECT domain boundaries (Figure 4). Our results demonstrate that the susceptibility of the exposed hydrophobic cleft at the N-terminus of the HECT domain can be mitigated by the addition of these additional α-helices. Each hydrophobic residue within the extended α-helix/helices suggests that this conserved region is required to support normal HECT-dependent ubiquitylation. The solubility of each HECT E3 ubiquitin ligase appears to be dependent on the addition of at least one of these predicted α-helical extensions. Both paralog and ortholog secondary structure prediction indicate that the extended region would likely have α-helical content (Figures 2 and 3).

The overexpression and purification of heterologous proteins in Escherichia coli greatly depends on the overall stability and native fold, while mitigating solvent accessibility of hydrophobic regions of the expressed protein [34–36]. Despite our repeated attempts to modify buffers, reagents, and purification procedures, we were unable to improve stability, activity, and/or the yield of any of our original HECT constructs using the UniProt defined domain boundaries. Our experimental optimization included (i) varying temperatures for expression following IPTG induction, (ii) altering IPTG concentrations, (iii) scaling down production culture volumes, and (iv) subcloning the HECT domain construct to be expressed with other purification fusion tags. While we were able to successfully purify some of these UniProt defined HECT constructs, the overall yield we obtained for our new extHECT constructs resulted in approximately a 50-fold increase in protein yield (Figure 4).
Figure 2. Paralog and ortholog MSA of the HECT E3 ligase family with secondary structure prediction through Jalview and JPred

(A) Paralog MSA of the junction with the extended HECT α-helix (α1') and the initial β-sheet (β1) of the HECT domain. Sequences and residue numbers were obtained through UniProt for all human HECT E3 ubiquitin ligases. Structural elements from solved crystal models with the α-helical extension (magenta) and N-lobe β-sheet (green) are indicated. Residues facing towards the hydrophobic core are designated with an asterisk. The alignment was produced using ClustalW and T-Coffee and manually curated by Jalview.

(B) Ortholog results revealed high conservation for the HECT E3 ubiquitin ligase HACE1 where all four α-helices are predicted to fold preceding the N-lobe of the HECT domain. This finding compliments the paralog analysis shown in panel (A). Residues with absolute conservation are shown in gray.
Figure 3. Secondary structure prediction reveals a hydrophobic pocket protected by the N-terminal α-helical extension
(A) MSA analysis coupled with secondary structure prediction for all HECT E3 ubiquitin ligases indicates a conserved amphipathic α-helix (shown in magenta) with hydrophobic residues suggested to interact with residues found in the hydrophobic cleft of the HECT N-terminal lobe (shown in cyan). This alignment takes into consideration the conservation and secondary structure prediction with a focus on the 1′ α-helix preceding the HECT domain. (B) The AREL1 HECT (PDB: 6JX5) with conserved hydrophobic residues shown within the predicted α-helical extension (magenta) are predicted to interact with the conserved N-lobe residues (cyan). This alignment also confirms the role that extending the known HECT domain boundaries are important for domain stability and solubility.
(C) Surface maps of the extended (right) and UniProt defined (left) WWP1 and AREL1 HECT domains. The conserved sidechain HECT N-lobe residues highlighted in panel (B) reside within the hydrophobic pocket (cyan). The α-helical extension (magenta) reduces the surface area exposed from the hydrophobic patch, as observed through the decrease in color. The inset (right) shows a closer look at the conserved hydrophobic residues in the hydrophobic patch.
Figure 4. Solubility and purification of UniProt defined HECT domain against extended HECT domains

(A) The solubility of each HECT and extHECT construct were analyzed by SDS-PAGE. Solubility following induction showed that the extHECT constructs of the HECT E3 ubiquitin ligases evaluated in this paper (shown with yellow stars) were more soluble than the shorter HECT domains using the domain boundaries annotated on UniProt. (B) Example SDS-PAGE of the purification steps for HACE1 HECT (residues 574–909 based on UniProt defined domain boundaries) and extHECT (residues 523–909) constructs. The HECT construct was deemed to be insoluble (shown with red octagon) while the extHECT HACE1 was significantly more soluble and with a higher yield (shown with yellow star). (C) Analysis of protein purity for all proteins used in this study including the isolated extHECT constructs. All samples were loaded at 5 mg of protein, except for UBE2L3 and HACE1523-909 which were loaded at 10 mg.

We also assessed the ubiquitylation activity of our HECT and extHECT proteins using an in vitro fluorescent ubiquitylation assay that has been optimized to demonstrate E2~ubiquitin charging/discharging and improved extHECT~ubiquitin charging. We observed a marked increase in ubiquitylation activity by our extHECT constructs, whereas the HECT domains showed little to no activity (Figure 5). All extended HECT domains used for these assays demonstrated that the HECT domain could be charged with ubiquitin. To demonstrate optimization of UBE2L3~Ub charging, UBE2L3 discharging and HECT/extHECT~Ub charging, we performed an additional assessment to find the optimal concentration of UBE2L3 to ubiquitin to limit the excess free ubiquitin found at the bottom of the gel. Once determining the ideal concentration of ubiquitin (5 μM) and UBE2L3 (10 μM), we performed an optimization assessment of the E3 at a 1:1 ratio of UBE2L3 to reproducibly show E3~ubiquitin charging and UBE2L3 discharge of ubiquitin. We found that using a HECT or extHECT concentration at 10 μM at a 1:1 ratio showed the robust activity of our extended HECT domain constructs. With the difficulty in isolating the shorter HECT proteins, our data shows that the inclusion of additional N-terminal α-helices provide overall greater structural integrity coupled with increased solubility that aid in HECT-dependent ubiquitylation activity.

Discussion

Building on previous studies, the present study proposes and reinforces the need redefine of the HECT domain boundaries currently annotated on UniProt and the major databases. The HECT domain that was originally annotated using the BLAST search engine of the NCBI databases yielded only 14 hits [1]. The NCBI GenBank databases at the end of 1994 contained less than 238,000 total sequences as compared with nearly 250,000,000 sequences currently available.
Figure 5. Ubiquitin activity assays show that the extHECT domains are more active than the UniProt defined HECT domain boundary constructs.

(A) Assay optimization through varying concentrations of Alexa Fluor 647-ubiquitin:UBE2L3 to assess the ideal concentration to minimize unused free ubiquitin. The UBE2L3~ubiquitin complex (shown in cyan) is highlighted to demonstrate the ideal concentration of UBE2L3 used with minimal free ubiquitin present. (B) The UBE2L3:HECT/extHECT ratios were repeated as before to determine the ideal E3~ubiquitin charging ratio (yellow star). (C) Fluorescent ubiquitylation activity assays of each HECT E3 ubiquitin ligase evaluated in this study. Gel images were taken with an iBright FL1000 gel imager with fluorescence settings for Alexa Fluor 647. The images were converted to black/white for ideal visualization of each ubiquitin transfer step and product.
in the databases as of the writing of this study. The original boundaries for the HECT domain were defined from the sequences found in Pfam entry PF00643 and InterPro entry IPR000569. Subsequent HECT domain boundaries and analyses have used this definition as the basis for phylogenetic studies such as the identification of HECT proteins in soybeans [37].

Identifying the exact domain boundaries within a protein often requires access to solved structural models; however, bioinformatic approaches using MSA and secondary structure and disorder prediction have proved to be successful approaches to predict domain boundaries. For example, this bioinformatic MSA approach was used to successfully predict the boundaries for the novel Parkin RING0 domain, and the disordered regions found between domains such as the ubiquitin-like (UbLD) and RING–BRcat–Rcat (RBR) domains of Parkin [38] before 3D structures of full-length parkin were released in 2013 [39–42]. Although the POND results concerning the human HECT domains can show some differences, as it often predicts disorder in known structures at the N-terminal, the POND server reproducibly predicted that there was disorder before our new extended HECT domain boundaries. Previous bioinformatic analysis performed on the HECT family have identified conserved residues that regulate autoinhibition upon modification. As with NEDD4 and E6AP, there are lysine and threonine residues that reside within the first α-helix predicted to fold preceding the N-lobe that are subject to methylation and phosphorylation [43,44].

While it was found that this N-terminal extension also induces HECT-dependent oligomerization, the exact structural mechanisms used by many of the members in the HECT E3 ubiquitin ligase family that employ this modification remain only partially understood. Previous studies have also pointed out that the inclusion of the extended α-helices in-front of the HECT domain are necessary to perform structural analysis of the HECT domain that coincidentally also include highly conserved, essential residues that mediate HECT-dependent oligomerization [3,8,11,31,43,45]. Future studies are essential to examine the oligomerization and/or autoinhibition potential for the currently illusive members of the HECT E3 ubiquitin ligases.

Within the Human Gene Mutation Database (HGMD), numerous missense mutations have been annotated in the HECT E3 ubiquitin ligase family that are in proximity to the HECT domain (Table 1) [46]. Of these, some are relatively conserved and correlate to the onset of similar diseases and/or disorders. For example, an N-terminal missense mutation of an arginine to stop codon, irrespective of their subfamily classification, is linked to intellectual disability. Specifically assessing the HECW2 (R1330W) and HACE1 (R585W) mutants, this arginine residue is relatively conserved in the N-lobe and linked to the onset of neurological disorders [47,48]. Interestingly, the HECTD3 mutant (R478C), which has been linked to Tourette syndrome, lies i-3 from the conserved phenylalanine of the α1-helical extension that protects the hydrophobic pocket. Also, the HECW2 F1193V missense mutation identified in neurodevelopmental delay and hypotonia is the conserved α1 residue that interacts with the N-terminal α-helix to protect the hydrophobic pocket [49]. This demonstrates that some of these novel missense mutations within the HECT E3 ubiquitin ligase family likely are involved, whether directly or indirectly, with protection of this hydrophobic pocket to maintain its structure and ubiquitylation function.

HECT E3 ubiquitin ligases have been implicated in numerous diseases and disorders in relation to their altered expression and functional capabilities [2]. For example, the altered expression levels of HACE1 leads toward the development of Wilms’ tumor [77]. Interestingly, overexpression of HUWE1 has been correlated to breast, brain, and prostate cancer progression, but also is down-regulated in both colorectal and lung cancer [78]. The effect that these residue substitutions listed in Table 1 have on the structure and function of the HECT E3 ubiquitin ligase family members require further investigation to determine if these residue substitutions affect protein stability, autoregulation, or HECT domain concentration-dependent oligomerization.

While this manuscript was in preparation, the structure prediction AlphaFold database was publicly released [79]. The domain boundaries from our study and the predicted structures from AlphaFold are summarized in Table 2. The four largest human HECT E3 ubiquitin ligases are absent from the AlphaFold database as they exceed their 2700 amino acid length criterion (i.e. HERC1, HERC2, HECTD4, and UBR5). In general, there was consensus between the extended HECT and AlphaFold database domain boundary predictions with an obligate α-helix present to stabilize the HECT N-terminal lobe (Table 2).

The significant difference was the size of the loop between the canonical HECT and the helical extension, and this was most pronounced in the HERC subfamily (i.e. HERC1, HERC2, HERC3, HERC4, HERC5, and HERC6) as well as HECTD1, HECTD2, and HECTD3. It should also be noted that the extended α-helix is predicted to lie in the hydrophobic groove with its N-terminus proximal to the C-lobe, with the only exception being TRIP12 where the N-terminal helix is predicted to lie in the opposite orientation (Figure 6). Regardless, these new computational models further support our suggestion that the HECT domain boundaries need to be expanded to include the N-terminal helix.
Table 1 HECT domain mutations annotated in the HGMD database linked to disease

| HECT  | Residue substitution | Location in the HECT domain | Clinical manifestation |
|-------|----------------------|----------------------------|------------------------|
| NED4L | Y679C; Q974H | N-lobe                     | Periventricular neuronal heterotopia [50] |
|       | E893K; R897Q | C-lobe                     | Neurodevelopmental disorder [51] |
| SMURF2| T641A             | C-lobe                     | Neurodevelopmental delay and hypotonia [49] |
|       | R1191Q; F1193V   | ext-α2<sup>a</sup>    | Developmental delay, absent speech, epilepsy, encephalopathy, hypotonia, and macrocephaly [47] |
|       | R1330W           | N-lobe                     | Autism spectrum disorder [52] |
|       | D1442G           | E1445G                     | Neurodevelopmental delay and hypotonia [49] |
| HERC1 | L4154X<sup>b</sup> | ext-α3<sup>a</sup>    | Autism spectrum disorder [53] |
|       | G4520E           | N-lobe                     | Overgrowth, intellectual disability, and facial dysmorphism [54] |
| HERC2 | D4267E           | ext-α3<sup>a</sup>    | Neurological disease [55] |
| TRIP12| R1595Q           | S1840L                     | Intellectual disability [56] |
|       | Q1916X<sup>b</sup> | C-lobe                   | Intellectual disability [57] |
|       | R3267H           | R4013W                     | Ext-α4<sup>a</sup>    | Cranosynostosis [58] |
|       | R4023C; N4075K; Y4106C; L4157V; E4244D | N-lobe | Intellectual disability, microcephaly, and postnatal growth failure [61] |
|       | R4063Q           | N-lobe                     | Intellectual disability, microcephaly, and postnatal growth failure [61] |
| HACE1 | R585W           | Q618Fs; P674Fs; R748X<sup>b</sup> | Spastic paraplegia and psychomotor retardation with or without seizures [65,66] |
|       | L832X<sup>b</sup> | A961P                     | Hereditary spastic paraplegia [67] |
| HECTD3| R478C           | C-lobe                     | Tourette syndrome [68] |
| AREL1 | P779L           | C-lobe                     | Pulmonary inflammation [69] |
| UBE3B | Q700X; G779R; Q727P | N-lobe | Kaufman oculocerebrofacial syndrome [70] |
|       | R922C           | C-lobe                     | Blepharophimosis-ptosis-intellectual-disability syndrome [71] |
|       | R997P           | C-lobe                     | Autism spectrum disorder [72] |
| UBE3C | S845F           | C-lobe                     | Kaufman oculocerebrofacial syndrome [70] |
|       | F996C           | C-lobe                     | Kaufman oculocerebrofacial syndrome [73] |
| E6AP  | I827K; G870D   | C-lobe                     | Angelman syndrome [75,76] |

<sup>a</sup>ext-α = predicted α-helix N-terminal to current UniProt defined boundary
<sup>b</sup>X = stop codon
<sup>c</sup>fs = codon frameshift

In summary, our study demonstrates that our extHECT constructs have increased solubility, improved stability, higher yields, and robust ubiquitylation activity. While some HECT E3 ubiquitin ligases have solved HECT domain models [6,25–27,30,32,33], many others remain unknown. To date, most of these HECT domain structures have been solved for members of the NEDD4 subfamily, with only a few members in the ‘other’ subfamily having a HECT domain structure available on the Protein Data Bank. The present study clearly shows that extending the N-terminal
### Table 2: Uniprot PROSITE and predicted AlphaFold boundaries of the HECT E3 ligase family

| Protein ID | UniProt | MSA/Jalview | AlphaFold | Notes |
|------------|---------|-------------|-----------|-------|
| HERC Family |         |             |           |       |
| HERC1      | 4501-4848 | 4451-4861  | N/A       |       |
| HERC2      | 4457-4794 | 4400-4834  | N/A       |       |
| HERC3      | 951-1050  | 681-1050   | 669-1050  | Longer predicted loop and helix |
| HERC4      | 730-1057  | 689-1057   | 667-1057  | Longer predicted loop and helix |
| HERC5      | 702-1024  | 659-1024   | 638-1024  | Longer predicted loop and helix |
| HERC6      | 693-1017  | 650-1022   | 629-1022  | Longer predicted loop and helix |
| NEDD Family |      |            |           |       |
| NEDD4      | 984-1318  | 940-1319   | 941-1319  | PDB 2XBF; [33] |
| NED4L      | 640-974   | 594-975    | 595-975   | PDB 2ONI (X-ray: SGC) |
| ITCH       | 569-903   | 527-903    | 528-903   | PDB 3TUG (X-ray: SGC) |
| WWP1       | 588-922   | 547-922    | 547-922   | PDB 1ND7; X-ray [6] |
| WWP2       | 536-870   | 494-870    | 494-870   | PDB 4Y07; X-ray [32] |
| SMURF1     | 420-757   | 377-757    | 369-757   | Similar prediction |
| SMURF2     | 414-748   | 371-748    | 367-748   | PDB 12VD; X-ray [26] |
| HERCW1     | 1271-1606 | 1226-1606  | 1223-1606 | Similar prediction |
| HERC2      | 1237-1572 | 1192-1572  | 1189-1572 | Similar prediction |
| Other Family |    |            |           |       |
| TRIP12     | 1885-1992 | 1561-1992  | 1558-1992 | unique helix orientation PDB 3Q1N, 3H1D, 5LP8; |
| HUWE1      | 4038-4374 | 3993-4374  | N/A       | X-ray [27,31] PDB 7MWE; Cryo-EM [80] |
| HACE1      | 574-909   | 529-909    | 524-909   | Similar prediction |
| HECTD1     | 2151-2610 | 2192-2610  | 2091-2610 | Longer predicted loop and helix |
| HECTD2     | 437-776   | 395-776    | 377-776   | Longer predicted loop and helix |
| HECTD3     | 512-857   | 474-861    | 451-861   | Longer predicted loop and helix |
| HECTD4     | 3627-3996 | 3570-3996  | N/A       | - |
| UBR5       | 2462-2799 | 2395-2799  | N/A       | - |
| AREL1      | 483-823   | 439-823    | 439-823   | PDB 6JKS; X-ray [25] |
| Q2E3       | 371-698   | 341-706    | 345-706   | Similar prediction |
| UBE3B      | 702-1068  | 659-1068   | 650-1068  | Longer predicted loop and helix |
| UBE3C      | 744-1083  | 695-1083   | 695-1083  | PDB 6K2C; X-ray [30] |
| E6AP       | 778-875   | 500-875    | 486-875   | Longer predicted loop and helix |

**Figure 6. AlphaFold predicts an α-helix preceding the HECT domain**

Predicted models of the extended HECT domain of HERC2, HERC5, and TRIP12 from the AlphaFold database (https://alphafold.ebi.ac.uk/) showing different α-helical extension configurations. The HECT N-terminal lobe (dark gray), C-terminal lobe (light gray), α-helical extension that stabilizes the N-terminal lobe (magenta), and the catalytic cysteine (yellow) shown that AlphaFold models predicts that different orientations for the N-terminal extension in the extHECT construct may be possible.
lobe and redefining the HECT domain boundaries for all the human HECT E3 ubiquitin ligases is warranted. This knowledge will be an important consideration for future biochemical, structural, and therapeutic studies on the HECT E3 ubiquitin ligase family.

**Methods**

**Bioinformatic analysis of the HECT E3 ubiquitin ligases**

The sequences for the 28 human HECT paralogs were acquired from the UniProt database (www.uniprot.org) using the canonical HECT E3 ubiquitin ligase isoform and parsing out splice variants. The sequences were aligned using ClustalW [81] and T-Coffee [82] followed by manual curation in Jalview [28]. Secondary sequence prediction by JPred [83] was used on the aligned sequences within Jalview to assist in determining domain boundaries. Furthermore, the Protein Data Bank (PDB; www.rcsb.org) database was scoured to identify all the deposited PDB structures for isolated HECT domains for PyMol analysis [29,84] (Figures 1 and 2). Jalview was also used to assess the ortholog family of HACE1, showing both high conservation and α-helical prediction location immediately N-terminal to the annotated HECT domain (Figure 2B). Ortholog MSA analysis of both protein families were downloaded from Ensembl (ensembl.org) and analyzed in Jalview [28,85]. Sequences that were obvious fragments (lacking an intact HECT domain that included the catalytic cysteine), contained stretches of unknown residues, or isoforms from the same organism were removed from the analyses. The organisms were chosen to span disparate eukaryotic clades. PONDR VSL1 analysis was also used to search the HECT E3 ubiquitin ligases for predicted disordered regions to assist in domain boundary predictions (www.pondr.com). Expasy ProtParam was used to predict the physicochemical properties of the predicted extended HECT domain constructs (web.expasy.org/protparam) [86,87].

The proteins focused on in this study (HECW2, HERC4, G2E3, and HACE1) do not have any deposited PDB structures, and our initial studies with our constructs based upon UniProt HECT domain boundaries had proven difficult to express and purify for downstream biochemical and biophysical analysis. To determine the size of the hydrophobic patch, the PDB files for the known HECT structures that include the helical extension were analyzed using PyMol. The helical extension was removed from the original PDB file as the basis for the shorter HECT domain analysis. Only the A chain was considered, all waters and other heteroatoms removed, and hydrogens were added. Hydrophobic residues (A,G,V,I,L,F,M,Y,W) sidechain atoms only were selected and the PyMol command ‘get area’ was used to quantitatively determine the surface area coverage.

**HECT domain expression plasmid design and construction**

Expression plasmids containing the full-length open reading frame of the corresponding HECT E3 ubiquitin ligases were designed with *E. coli* codon optimization and synthesized by Synbio Technologies, Inc. (Monmouth Junction NJ, U.S.A.). The plasmids contained the essential components desired for protein overexpression in *E. coli* – a T7 promoter induced through the addition of β-D-1-thiogalactopyranoside (IPTG), a lac operon, origin of replication and terminator sequence, antibiotic resistant marker for ampicillin (amp) for selection, an N-terminal His6-SUMO tag for increased proteins solubility and downstream Ni2+-immobilized metal affinity chromatography (Ni2+-IMAC) purification, and flanking restriction sites for *Bam*HI (5′) and *Xho*I (3′) in the multiple cloning site.

Primers were designed to isolate both the UniProt defined HECT (HECT) and extended HECT (extHECT) domains for each protein from the full-length sequence plasmids by deleting residues from the full-length vector using single primers in parallel (SPRINP) [88] and quick-change mutagenesis (QCM) [89] strategies (Table 3). The boundaries for each extHECT construct were chosen based on bioinformatic predictions including the MSA and reference to the recent AREL1 and UBE3B HECT domain PDB structures [25,30]. These included the NEDD4 family member HECW2 (HECT: 1237-1572; extHECT: 1190-1572), HERC protein HERC4 (HECT: 708-1057; extHECT: 675-1057), and ‘other’ HECT subfamily members G2E3 (HECT: 369-706; extHECT: 340-706) and HACE1 (HECT: 574-909; extHECT: 523-909). All primer sequences were synthesized by Thermo-Fisher. SPRINP and QCM were performed using a T100 Thermal Cycler (Bio-Rad). The HECW2 and HERC4 extHECT expression constructs were made through PCR amplification of the desired ORF, restriction digestion with flanking *Bam*HI (5′) and *Xho*I (3′), and ligation with T4 DNA ligase into the His6-SUMO plasmid. Colonies were screened by colony PCR and restriction digestion, and positive transformants were verified by Sanger DNA sequencing (Psomagen, Cambridge MA, U.S.A.).

**Overexpression and purification of HECT and extHECT domains**

The His6SUMO-HECT and His6SUMO-extHECT expression plasmids for HECW2, HERC4, G2E3, and HACE1 were transformed into either BL21(DE3) or LOBSTR(DE3) *E. coli* cell lines and grown at 37°C at 200 rpm in Luria-Burtani media supplemented with 100 μg/ml ampicillin. When the cultures reached an OD600 of approximately 0.6–0.8,
Table 3 Oligonucleotide primers for recombinant HECT and extHECT plasmids

| Primers | Orientation | Nucleotide sequence |
|---------|-------------|---------------------|
| HACE1   | Sense       | 5′-GGAGGTGGATCCAAGGCTAATTGTGCTAAG-3′ |
|         | Antisense   | 3′-CTTAGACAAATTAGGTTGATCCACCTCCTG-5′ |
| 572-909 | Sense       | 5′-GGAGGTGGATCCAGCCGTTCAAGGATCGC-3′ |
|         | Antisense   | 3′-GCGATCCTTGAACGGCTGGGATCCACCTCC-5′ |
| G2E3    | Sense       | 5′-GATTGGAGGTGGATCCACGAAGCGTCTTTAC-3′ |
|         | Antisense   | 3′-CTAACCTCCACCTAGGTGCTTCGCAGAAATG-5′ |
| 369-706 | Sense       | 5′-GATTGGAGGTGGATCCTCAAAATTCCGGCG-3′ |
|         | Antisense   | 3′-CTAACCTCCACCTAGGAGTTTTAAGGCCGC-5′ |
| 523-909 | Sense       | 5′-GGAGGTGGATCCCAGCCGTTCAAGGATCGC-3′ |
|         | Antisense   | 3′-GCGATCCTTGAACGGCTGGGATCCACCTCC-5′ |
| 340-706 | Sense       | 5′-GATTGGAGGTGGATCCTCAAAATTCCGGCG-3′ |
|         | Antisense   | 3′-CTAACCTCCACCTAGGAGTTTTAAGGCCGC-5′ |
| HECW2   | Sense       | 5′-GGAGGTGGATCCAGCAGGAAGGAC-3′ |
|         | Antisense   | 3′-CCTCCACCTAGGTCGTCCTTCCTG-5′ |
| 1237-1572 | Sense     | 5′-GGAGGTGGATCCAAACGCGATTTCG-3′ |
|         | Antisense   | 3′-CCTCCACCTAGGTTTGCGCTAAAGC-5′ |
| 1190-1572 | Sense     | 5′-GGAGGTGGATCCAAACGCGATTTCG-3′ |
|         | Antisense   | 3′-CCTCCACCTAGGTTTGCGCTAAAGC-5′ |
| HERC4   | Sense       | 5′-CAGATTGGAGGTGGATCTGTCCTAATCTCGT-3′ |
|         | Antisense   | 3′-GTCTAACCTCCACCTAGGACGGATTAGAACCAA-5′ |
| 708-1057 | Sense       | 5′-CAGATTGGAGGTGGATCTGTCCTAATCTCGT-3′ |
|         | Antisense   | 3′-GTCTAACCTCCACCTAGGACGGATTAGAACCAA-5′ |
| 683-1057 | Sense       | 5′-CAGATTGGAGGTGGATCCCAAACTGAGATTGAC-3′ |
|         | Antisense   | 3′-GTCTAACCTCCACCTAGGACGGATTAGAACCAA-5′ |

In vitro fluorescent ubiquitylation assay

All soluble HECT and extHECT proteins were assessed for their ability to catalyze ubiquitin transfer between all cascading enzymes using a fluorescent auto-ubiquitylation assay. Specific attention was given to observing if the inclusion of the N-terminal extension impacted HECT-dependent ubiquitylation activity. Once HECT/extHECT−ubiquitin charging was confirmed through assay optimization, ubiquitylation assays were conducted with 5 μM Alexa Fluor 647 N-terminally tagged ubiquitin (0.1 μM E1 activating enzyme UBE1 (Uba1), 10 μM E2 conjugating enzyme UBE2L3 (UbcH7) or UBE2D3 (UbcH5c), and 10 μM HECT or extHECT, 0.5 mM DTT, 4 μM ATP, 10 μM MgCl₂ in 50 mM HEPES pH 7.5, 100 mM NaCl. Each reaction was incubated at 37 °C at 200 rpm for 1 to 2 h. To determine the presence of ubiquitin−thioester intermediates, appropriate samples were supplemented with 4 mM DTT. Reactions were terminated by adding SDS-PAGE loading dye and heating at 95 °C in a dry bath for one minute. The samples were...
then loaded on to a 12% Bis-Tris gel at pH 6.4 and run for 1 h at 130 V. The gels were removed from the apparatus and immediately visualized on an iBright FL1000 imaging system (Thermo-Fisher) using the setting for Alexa Fluor 647 fluorescence gel imaging.

Data Availability
All materials for this project could be made available upon request.

Competing Interests
The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Funding
This research was supported by the National Institutes of Health [grant number R15GM126432 (to D.E.S.)].

CRediT Author Contribution
Emma I. Kane: Conceptualization, Writing—original draft, Writing—review & editing. Steven A. Beasley: Conceptualization, Writing—original draft, Writing—review & editing. Johanna M. Schafer: Conceptualization. Justine E. Bohl: Conceptualization. Young Sun Lee: Conceptualization. Kayla J. Rich: Conceptualization. Elizabeth F. Bosia: Conceptualization. Donald E. Spratt: Conceptualization, Supervision, Writing—original draft, Writing—review & editing.

Abbreviations
AREL1, apoptosis-resistant E3 ubiquitin protein ligase-1; E6AP, E6-associated protein; HECT, homologous to E6AP C-terminus; HGMD, Human Gene Mutation Database; MSA, multiple sequence alignment; RLD, RCC1-like domain.

References
1 Huibregtse, J.M., Scheffner, M., Beaudenon, S. and Howley, P.M. (1995) A family of proteins structurally and functionally related to the E6-AP ubiquitin-protein ligase. Proc. Natl. Acad. Sci. U. S. A. 92, 2563–2567, https://doi.org/10.1073/pnas.92.7.2563
2 Wang, Y., Argüelles-Castillo, D., Kane, E.I., Zhou, A. and Spratt, D.E. (2020) HECT E3 ubiquitin ligases - emerging insights into their biological roles and disease relevance. J. Cell Sci. 133, jcs228072, https://doi.org/10.1242/jcs.228072
3 Huang, L., Kinnucan, E., Wang, G., Beaudenon, S., Howley, P.M., Huibregtse, J.M. et al. (1999) Structure of an E6AP-UbcH7 complex: insights into ubiquitination by the E2-E3 enzyme cascade. Science 286, 1321–1326, https://doi.org/10.1126/science.286.5443.1321
4 Kishino, T., Lalande, M. and Wagstaff, J. (1997) UBE3A/E6-AP mutations cause Angelman syndrome. Nat. Genet. 15, 70–73, https://doi.org/10.1038/ng0197-70
5 Kamadurai, H.B., Souphron, J., Scott, D.C., Duda, D.M., Miller, D.J., Stringer, D. et al. (2009) Insights into ubiquitin transfer cascades from a structure of a UbcH5B approximately ubiquitin-HECT(NEDD4L) complex. Mol. Cell 36, 1095–1102, https://doi.org/10.1016/j.molcel.2009.11.010
6 Verdecia, M.A., Joazeiro, C.A., Wells, N.J., Ferrer, J.L., Bowman, M.E., Hunter, T. et al. (2003) Conformational flexibility underlies ubiquitin ligation mediated by the WWP1 HECT domain E3 ligase. Mol. Cell. 11, 249–259, https://doi.org/10.1016/S1097-2765(02)00774-8
7 Ronchi, V.P., Kim, E.D., Summa, C.M., Klein, J.M. and Haas, A.L. (2017) In silico modeling of the cryptic E2 approximately ubiquitin-binding site of E6-associated protein (E6AP)/UBE3A reveals the mechanism of polyubiquitin chain assembly. J. Biol. Chem. 292, 18006–18023, https://doi.org/10.1074/jbc.M117.813477
8 Ronchi, V.P., Klein, J.M., Edwards, D.J. and Haas, A.L. (2014) The active form of E6-associated protein (E6AP)/UBE3A ubiquitin ligase is an oligomer. J. Biol. Chem. 289, 1033–1048, https://doi.org/10.1074/jbc.M113.517805
9 Ronchi, V.P., Klein, J.M. and Haas, A.L. (2013) E6AP/UBE3A ubiquitin ligase harbors two E2–ubiquitin binding sites. J. Biol. Chem. 288, 10349–10360, https://doi.org/10.1074/jbc.M113.458059
10 Todaro, D.R., Augustus-Wallace, A.C., Klein, J.M. and Haas, A.L. (2017) The mechanism of neural precursor cell expressed developmentally down-regulated 4-2 (Nedd4-2)/NEDD4L-catalyzed polyubiquitin chain assembly. J. Biol. Chem. 292, 19521–19536, https://doi.org/10.1074/jbc.M117.817882
11 Todaro, D.R., Augustus-Wallace, A.C., Klein, J.M. and Haas, A.L. (2018) Oligomerization of the HECT ubiquitin ligase NEDD4-2/NEDD4L is essential for polyubiquitin chain assembly. J. Biol. Chem. 293, 18192–18206, https://doi.org/10.1074/jbc.RA118.003716
12 Rotin, D. and Kumar, S. (2009) Physiological functions of the HECT family of ubiquitin ligases. Nat. Rev. Mol. Cell Biol. 10, 398–409, https://doi.org/10.1038/nrm2690
13 Plant, P.J., Yeger, H., Staub, O., Howard, P. and Rotin, D. (1997) The C2 domain of the ubiquitin protein ligase Nedd4 mediates Ca2+-dependent plasma membrane localization. J. Biol. Chem. 272, 32329–32336, https://doi.org/10.1074/jbc.272.51.32329
14 Staub, O., Dho, S., Henry, P., Correa, J., Ishikawa, T., McGlade, J. et al. (1996) WW domains of Nedd4 bind to the proline-rich PY motifs in the epithelial Na+ channel deleted in Liddle’s syndrome. EMBO J. 15, 2371–2380, https://doi.org/10.1002/j.1460-2075.1996.tb00593.x
15 Sanchez-Tena, S., Cubillos-Rojas, M., Schneider, T. and Rosa, J.L. (2016) Functional and pathological relevance of HERC family proteins: a decade later. Cell. Mol. Life Sci. 73, 1955–1968, https://doi.org/10.1007/s00018-016-2139-8

© 2022 The Author(s). This is an open access article published by Portland Press Limited on behalf of the Biochemical Society and distributed under the Creative Commons Attribution License 4.0 (CC BY).
45 Ronchi, V.P. and Haas, A.L. (2012) Measuring rates of ubiquitin chain formation as a functional readout of ligase activity. Methods Mol. Biol. 832, 197–218, https://doi.org/10.1007/978-1-61779-474-2_14
46 Stenson, P.D., Ball, E.V., Mort, M., Phillips, A.D., Shiel, J.A., Thomas, N.S. et al. (2003) Human Gene Mutation Database (HGMD): 2003 update. Hum. Mutat. 21, 577–581, https://doi.org/10.1002/humu.10212
47 Halvardson, J., Zhao, J.J., Zaghillo, A., Wentzel, C., Georgii-Hemming, P., Mansson, E. et al. (2016) Mutations in HECW2 are associated with intellectual disability and epilepsy. J. Med. Genet. 53, 697–704, https://doi.org/10.1136/jmedgenet-2016-103814
48 Iossifov, I., O’Roak, B.J., Sanders, S.J., Ronemus, M., Krumm, N., Levy, D. et al. (2014) The contribution of de novo coding mutations to autism spectrum disorder. Nature 515, 216–221, https://doi.org/10.1038/nature13906
49 Berko, E.R., Cho, M.T., Eng, C., Shao, Y., Sweetser, D.A., Waxler, J. et al. (2017) De novo missense variants in HECW2 are associated with neurodevelopmental delay and hypotonia. J. Med. Genet. 54, 84–86, https://doi.org/10.1136/jmedgenet-2016-103943
50 Broix, L., Jagline, H., Ivanova, E., Schmucker, S., Drouot, N., Clayton-Smith, J. et al. (2016) Mutations in the HECT domain of NEDD4L lead to AKT-mTOR pathway deregulation and cause periventricular nodular heterotopia. Nat. Genet. 48, 1349–1358, https://doi.org/10.1038/ng.3676
51 Reuter, M.S., Tawamie, H., Buchert, R., Hosny Gebril, O., Froukh, T., Thiel, C. et al. (2017) Diagnostic yield and novel candidate genes by exome sequencing in 152 consanguineous families with neurodevelopmental disorders. JAMA Psychiatry 74, 293–299, https://doi.org/10.1001/jamapsychiatry.2016.3798
52 Krumm, N., Turner, T.N., Baker, C., Vives, L., Mohajeri, K., Witherspoon, K. et al. (2015) Excess of rare, inherited truncating mutations in autism. Nat. Genet. 47, 582–588, https://doi.org/10.1038/ng.3303
53 Hashimoto, R., Nakazawa, T., Tsurusaki, Y., Yasuda, Y., Nagayasu, K., Matsumura, K. et al. (2016) Whole-exome sequencing and neurite outgrowth analysis in autism spectrum disorder. J. Hum. Genet. 61, 199–206, https://doi.org/10.1038/jhg.2015.141
54 Ortega-Recalejo, O., Beltran, D.I., Galvez, J.M., Palma-Montero, A., Restrepo, C.M., Mateus, H.E. et al. (2015) Biallelic HERC1 mutations in a syndromic form of overgrowth and intellectual disability. Clin. Genet. 88, e1–e3, https://doi.org/10.1111/cge.12634
55 Charr, W.L., Karaca, E., Coban Akdemir, Z., Gu, S. et al. (2016) Exome sequencing in mostly consanguineous Arab families with neurologic disease provides a high potential molecular diagnosis rate. BMC Med. Genomics 9, 42, https://doi.org/10.1186/s12920-016-0208-3
56 Bramswig, N.C., Ludecke, H.J., Pettersson, M., Albrecht, B., Bernier, R.A., Cremer, K. et al. (2017) Identification of new TRIP12 variants and detailed clinical study of 21 patients. Eur. J. Hum. Genet. 25, 197–218, https://doi.org/10.1007/978-1-61779-474-2_14
57 Leilieveld, S.H., Reijnders, M.R., Pfundt, R., Yntema, H.G., Kamsteeg, E.J., de Vries, P. et al. (2016) Meta-analysis of 2,104 trios provides support for 10 new genes for intellectual disability. Nat. Neurosci. 19, 1194–1196, https://doi.org/10.1038/nn.4352
58 Clarke, C.M., Fok, V.T., Timms, A.E., Frazar, C.D. et al. (2018) Single suture craniosynostosis: Identification of rare variants in genes associated with syndromic forms. Am. J. Med. Genet. A. 176, 290–300, https://doi.org/10.1002/ajmg.a.38540
59 Froyen, G., Corbett, M., Vandewalle, J., Vanhove, S. et al. (2008) Submicroscopic duplications of the hydroxysteroid dehydrogenase HSD17B10 and the E3 ubiquitin ligase HUWE1 are associated with mental retardation. Am. J. Hum. Genet. 82, 432–443, https://doi.org/10.1016/j.ajhg.2007.11.002
60 Moortgat, S., Berland, S., Aukrust, I., Pentland, R., Aukrust, I., Baker, L. et al. (2015) HUWE1 variants cause dominant X-linked intellectual disability: a clinical study of 21 patients. Eur. J. Hum. Genet. 23, 64–74, https://doi.org/10.1038/ejhg.2014.163
61 Friez, M.J., Brooks, S.S., Stevenson, R.E., Field, M., Basehore, M.J., Ades, L.C. et al. (2016) HUWE1 mutations in Joubert-Marsidi and Brooks syndromes: the results of an X-chromosome exome sequencing panel. BMJ Open 6, e009537, https://doi.org/10.1136/bmjopen-2015-009537
62 Strauss, K.A., Gonzaga-Jauregui, C., Briggli, K.W., Williams, K.B., King, A.K., Van Hout, C. et al. (2016) Genomic diagnostics within a medically underserved population: efficacy and implications. Genet. Med. 20, 31–41, https://doi.org/10.1038/jmg.2015.77.6
63 Niranjani, T.S., Skinner, C., May, M., Turner, T., Rose, R., Stevenson, R. et al. (2015) Affected kindred analysis of human X chromosome exomes to identify novel X-linked intellectual disability genes. PLoS ONE 10, e0116454, https://doi.org/10.1371/journal.pone.0116454
64 Rutterer, K., Juusola, J., Cho, M.T., Vlazka, P., Millan, F., Gibelendi, F. et al. (2016) Clinical application of whole-exome sequencing across clinical indications. Genet. Med. 18, 696–704, https://doi.org/10.1038/jmg.2015.148
65 Nagy, V., Hollstein, R., Pai, T.P., Herde, M.K., Buphamalai, P., Moeseneder, P. et al. (2019) HACE1 deficiency leads to structural and functional neurodevelopmental defects. Neurot. Genet. 5, e330, https://doi.org/10.1211/NXG.00000000000300
66 Deng, H.X. (2019) HACE1, RAC1, and what else in the pathogenesis of SPPRS? Neur. Genet. 5, e326, https://doi.org/10.1211/NXG.00000000000326
67 Sager, G., Turkylıma, A., Ates, E.A. and Kütluay, B. (2021) HACE1, GLRX5, and ELP2 gene variant cause spas tic paraplegies. Acta Neurol. Belg. 119, 391–399, https://doi.org/10.13767/590-201-01685-3
68 Wilks, B. and Kieloski, J.W. (2000) De novo coding variants are strongly associated with Tourette disorder. Nature 468, 486–489,499e498, https://doi.org/10.1016/j.neuron.2004.04.024
69 Leal, M.S., McKelvey, A.C., Evanko, J.W., Rajbhandari, S., Coon, T.A., Dunn, S.R. et al. (2019) KAA0317 regulates pulmonary inflammation through SOCS2 degradation. JCI Insight 4, e129110, https://doi.org/10.1172/jci.insight.129110
70 Biallelic UBE3B mutations. Hum. Genet. 133, 939–949, https://doi.org/10.1007/s00439-014-1436-2
71 Biallelic UBE3B mutations. Am. J. Hum. Genet. 91, 998–1010, https://doi.org/10.1016/j.ajhg.2012.10.011
72 Chahour, M.H., Yu, T.W., Lim, E.T., Ataman, B., Coultet, M.E., Hill, R.S. et al. (2012) Whole-exome sequencing and homozygosity analysis implicate depolarization-regulated neuronal genes in autism. PLoS Genet. 8, e1002635, https://doi.org/10.1371/journal.pgen.1002635
Yilmaz, R., Szakszon, K., Altmann, A., Altunoglu, U., Senturk, L., McGuire, M. et al. (2018) Kaufman oculocerebrofacial syndrome: Novel UBE3B mutations and clinical features in four unrelated patients. Am. J. Med. Genet. A. 176, 187–193, https://doi.org/10.1002/ajmg.a.38538

O’Roak, B.J., Vives, L., Girirajan, S., Karakoc, E., Krumm, N., Coe, B.P. et al. (2012) Sporadic autism exomes reveal a highly interconnected protein network of de novo mutations. Nature 485, 246–250, https://doi.org/10.1038/nature10989

Beasley, S.A., Kellum, C.E., Orlomoski, R.J., Idri, F. and Spratt, D.E. (2020) An Angelman syndrome substitution in the HECT E3 ubiquitin ligase C-terminal Lobe of E6AP affects protein stability and activity. PLoS ONE 15, e0235925, https://doi.org/10.1371/journal.pone.0235925

Martinez, F., Caro-Llopis, A., Rosello, M., Oltra, S., Mayo, S., Monfort, S. et al. (2017) High diagnostic yield of syndromic intellectual disability by targeted next-generation sequencing. J. Med. Genet. 54, 87–92, https://doi.org/10.1136/jmedgenet-2016-103964

Anglesio, M.S., Evdokimova, V., Melnyk, N., Zhang, L., Fernandez, C.V., Grundy, P.E. et al. (2004) Differential expression of a novel ankyrin containing E3 ubiquitin-protein ligase, Hace1, in sporadic Wilms’ tumor versus normal kidney. Hum. Mol. Genet. 13, 2061–2074, https://doi.org/10.1093/hmg/ddh215

Su, C., Wang, T., Zhao, J., Cheng, J. and Hou, J. (2019) Meta-analysis of gene expression alterations and clinical significance of the HECT domain-containing ubiquitin ligase HUWE1 in cancer. Oncol. Lett. 18, 2292–2303, https://doi.org/10.3892/ol.2019.10579

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O. et al. (2021) Highly accurate protein structure prediction with AlphaFold. Nature 596, 583–589, https://doi.org/10.1038/s41586-021-03819-2

Hunkeler, M., Jin, C.Y., Ma, M.W., Monda, J.K., Overwijn, D., Bennett, E.J. et al. (2021) Solenoid architecture of HUWE1 contributes to ligase activity and substrate recognition. Mol. Cell. 81, 3468e3467–3480e3467, https://doi.org/10.1016/j.molcel.2021.06.032

Li, K.B. (2003) ClustalW-MPI: ClustalW analysis using distributed and parallel computing. Bioinformatics 19, 1585–1586, https://doi.org/10.1093/bioinformatics/btg192

Procter, J.B., Carstairs, G.M., Soares, B., Mourao, K., Ofoegbu, T.C., Barton, D. et al. (2021) Alignment of Biological Sequences with Jalview. Methods Mol. Biol. 2231, 203–224, https://doi.org/10.1007/978-1-0716-1036-7_13

Cole, C., Barber, J.D. and Barton, G.J. (2008) The Jpred 3 secondary structure prediction server. Nucleic Acids Res. 36, W197–W201, https://doi.org/10.1093/nar/gkn238

Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H. et al. (2000) The Protein Data Bank. Nucleic Acids Res. 28, 235–242, https://doi.org/10.1093/nar/28.1.235

Howe, K.L., Achuthan, P., Allen, J., Allen, J., Alvarez-Jarreta, J., Amode, M.R. et al. (2021) Ensembl 2021. Nucleic Acids Res. 49, D884–D891, https://doi.org/10.1093/nar/gkaa942

Gaestelger, E., Gattiker, A., Hoogland, C., Ivanji, I., Appel, R.D. and Bairoch, A. (2003) ExPASy: The proteomics server for in-depth protein knowledge and analysis. Nucleic Acids Res. 31, 3784–3788, https://doi.org/10.1093/nar/gkg563

Yue, B., Dunbrack, R.L., Williams, R.W., Dunker, A.K. and Uversky, V.N. (2010) PONDR-FIT: a meta-predictor of intrinsically disordered amino acids. Biochim. Biophys. Acta 1804, 996–1010, https://doi.org/10.1016/j.bbapap.2010.01.011

Edelheit, O., Hanukoglu, A. and Hanukoglu, I. (2009) Simple and efficient site-directed mutagenesis using two single-primer reactions in parallel to generate mutants for protein structure-function studies. BMC Biotechnol. 9, 61, https://doi.org/10.1186/1472-6750-9-61

Liu, H. and Naismith, J.H. (2008) An efficient one-step site-directed deletion, insertion, single and multiple-site plasmid mutagenesis protocol. BMC Biotechnol. 8, 91, https://doi.org/10.1186/1472-6750-8-91