Mechanisms of HECT-mediated Ubiquitin Chain Elongation

DISSERTATION ZUR ERLANGUNG DES
DOKTORGRADES DER NATURWISSENSCHAFTEN (DR. RER. NAT.)
DER FAKULTÄT FÜR BIOLOGIE UND VORKLINISCHE MEDIZIN
DER UNIVERSITÄT REGENSBURG

vorgelegt von

Carsten Alexander Stollmaier

aus

Karlsruhe

im Jahr

2022

Das Promotionsgesuch wurde eingereicht am:

01.09.2022

Die Arbeit wurde angeleitet von:

Dr. Silke Wiesner

Unterschrift:

Table of Contents

A. Summary

Ubiquitination is the covalent attachment of the small 76 amino acid protein ubiquitin to target proteins. Since the discovery of the role of ubiquitination in targeting substrates for proteasomal degradation, a plethora of other roles in the regulation of cellular function have been identified. These roles depend on the way ubiquitin is attached: one or more substrate lysines can be conjugated to either mono ubiquitin or multiple ubiquitins, in which ubiquitins are attached to each other forming chains of different linkage and length. Ubiquitination is driven by a cascade of three enzymes, of which the last, the E3 ubiquitin ligase, is responsible for substrate selectivity.

In the HECT-domain (Homologous to the E6-AP Carboxyl Terminus) family of E3 ligases, which are the focus of this thesis, not only substrate selectivity but also linkage specificity is given by the E3. However, the mechanism by which linkage specificity is reached remains unclear. Previous studies showed that a part of the smaller, C-terminal lobe of the HECT domain is responsible for determining linkage specificity, though the molecular mechanism remains elusive. In this work we explored the role of the interface between HECT domain and ubiquitin, also located in the C-terminal lobe. We showed that this interface, and thus the orientation of ubiquitin with respect to the HECT domain, is conserved between HECT domains of different ubiquitin chain linkage specificity, and therefore unlikely to contribute to linkage specificity.

Since HECT domain E3 ligases have been identified as both oncogenes and tumour suppressors, their regulation is of obvious interest. In most members of the HECT domain family, sequences upstream of the catalytic domain, such as C2 or WW domains, are responsible for its inhibition. In this work, we describe a novel mechanism, in which the HECT domain of Huwe1 is inhibited due to a tight interface between its C- and N-terminal lobes, independently of other domains.

Furthermore, while usually HECT domains bind ubiquitin using their ubiquitin-binding surface, we did not detect any such binding for Huwe1. Weakening the inter-lobe interface with point mutations in key residues rendered the ubiquitin-binding surface accessible, allowing us to characterize it. The ubiquitin-binding surface had been postulated to have a positive effect on the activity of the HECT E3 ligases by keeping the ubiquitin chains near the HECT domain between rounds of elongation. Additionally, we showed that in Huwe1, the ubiquitin-binding surface exerts its effect on catalysis independently of chains, possibly by keeping the inter-lobe interface “open”.

Lastly, we sought to investigate factors controlling the chain length. To reduce the complexity of previous systems used in literature, which largely focused on complex kinetical models to fit elongation rates of chains of different lengths from observations of reaction mixtures, we devised a model in which elongation rates of chains of defined length could be observed separately. During set-up optimization, it became apparent that N-terminally epitope tagged ubiquitin, often used in place of wild-type ubiquitin for ease of detection, is unsuitable for this type of ubiquitination studies since it can be ubiquitinated by the HECT domains not only in its lysines but also in its N-terminal amine. This unspecific reaction can be performed by all HECT domains we tested. Therefore, in our experiments we used wild-type ubiquitin. Our results show that elongation rate decreases proportionally to chain length, consistent with a model where individual ubiquitins in a chain can all bind to the HECT domain with equal affinity. In summary, the results presented here provide novel insights into the function and regulation of HECT domain ubiquitin ligases.

B. Zusammenfassung

Ubiquitinierung ist die kovalente Bindung des kleinen 76 Aminosäure-Proteins Ubiquitin an Substratproteine. Seit der Entdeckung der Rolle der Ubiquitinierung bei der Markierung von Substraten zum proteasomalen Abbau wurde eine Fülle anderer regulatorischer Rollen identifiziert. Diese Rollen hängen davon ab, wie Ubiquitin gebunden ist: Ein oder mehrere Substrat-Lysine können entweder mit Mono-Ubiquitin oder verschieden verknüpften Ubiquitin-Ketten konjugiert werden. Die Ubiquitinierung wird durch eine Kaskade von drei Enzymen angetrieben, von denen das letzte, die E3-Ubiquitinligase, für die Substratselektivität verantwortlich ist.

In der HECT (Homolog zum E6-AP C-Terminus) Familie der E3-Ligasen, die im Fokus dieser Arbeit stehen, wird nicht nur die Substratselektivität, sondern auch der Verknüpfungstyp der gebildeten Ubiquitin-Ketten durch die E3-Ligase bestimmt. Der Mechanismus, der über die Verknüpfungstyp-spezifität der Ubiquitin-Ketten entscheidet, bleibt jedoch unklar. Frühere Studien zeigten, dass ein Teil der kleineren, C-terminalen Subdomäne der HECT-Domäne für die Bestimmung der Bindungsspezifität verantwortlich ist, wobei der genaue molekulare Mechanismus jedoch unbekannt ist. In dieser Arbeit untersuchten wir die Rolle des Interfaces zwischen HECT-Domäne und Ubiquitin, welches sich ebenfalls in der C-terminalen Subdomäne der HECT-Domäne befindet. Wir haben gezeigt, dass dieses Interface und somit die Orientierung von Ubiquitin in Bezug auf die HECT-Domäne gleich ist zwischen HECT-Domänen mit unterschiedlicher Ubiquitin-Ketten Verknüpfungstyp-Spezifität. Das Interface trägt daher wahrscheinlich nicht zur Spezifität des Verknüpfungstyps bei.

Da Ubiquitin-Ligasen der HECT-Familie sowohl als Onkogene als auch als Tumorsuppressoren fungieren können, ist ihre Regulation von offensichtlichem Interesse. Bei den meisten Mitgliedern der HECT-Domänenfamilie sind Sequenzabschnitte vor der C-terminalen HECT-Domäne, wie C2- oder WW-Domänen, für ihre Hemmung verantwortlich. In dieser Arbeit beschreiben wir einen neuartigen Mechanismus, bei dem die HECT-Domäne von Huwe1 unabhängig von anderen Domänen aufgrund einer engen Schnittstelle zwischen ihren C- und N-terminalen Subdomänen gehemmt wird.

Während HECT-Domänen Ubiquitin normalerweise in einer konservierten nicht-kovalenten Ubiquitin-Bindungsstelle binden, konnten wir für Huwe1 zuerst keine solche Bindung feststellen. Durch die gezielte Schwächung der Schnittstelle zwischen den Subdomänen der HECT-Domäne mit Punktmutationen, wurde die Ubiquitin-bindende Oberfläche zugänglich, sodass wir sie charakterisieren konnten. Während in der Literatur die Bindung der Ubiquitin ketten zwischen den Kettenverlängerungsschritten als Grund für die nicht-kovalenten Ubiquitin-Bindungsstelle angeführt wird, haben wir darüber hinaus gezeigt, dass in Huwe1 die Ubiquitin-bindende Oberfläche ihre Wirkung auf die Katalyse unabhängig von der Kettenbildung ausübt, möglicherweise indem sie das Interface zwischen den beiden Subdomänen „offen“ hält.

Im letzten Teil der Arbeit geht es darum Einblick in die Faktoren, die die Kettenlänge steuern, zu erlangen. Um die Komplexität gegenüber in der Literatur verwendeten Systemen zu reduzieren, welche sich weitgehend auf komplexe kinetische Modelle konzentrierten, welche die Verlängerungsraten von Ketten unterschiedlicher Länge aus Beobachtungen von Reaktionsmischungen bestimmten, entwickelten wir ein Modell, in dem die Verlängerungsraten von Ketten definierter Länge separat beobachtet werden konnten. Beim Aufsetzen des Systems wurde deutlich, dass Ubiquitin mit N-terminalen Epitop-Tags, welches häufig anstelle von Wildtyp-Ubiquitin aufgrund der leichteren Detektion verwendet wird, für diese Art von Ubiquitinierungsstudien ungeeignet ist, da es in den Lysinen im Epitop-Tag sowie im N-terminalen Amin ubiquitiniert werden kann. Diese unspezifische Reaktion wird von allen von uns getesteten HECT-Domänen katalysiert. Um dieses Problem zu vermeiden, verwendeten wir in unseren weiteren Experimenten Wildtyp-Ubiquitin ohne Epitop-Tag. Unsere Ergebnisse zeigen, dass die Elongationsrate proportional zur Kettenlänge abnimmt, was mit einem Modell übereinstimmt, bei dem einzelne Ubiquitine in einer Kette alle mit

gleicher Affinität an die HECT-Domäne binden können. Zusammenfassend liefern die hier vorgestellten Ergebnisse neue Einblicke in die Funktion und Regulation von Ubiquitin-Ligasen der HECT-Familie.

I. General Introduction

C. Ubiquitin and Ubiquitination

1. The beginnings of Ubiquitin research

The small, 76 amino acid protein ubiquitin was first identified in 1975 as one of the most abundant polypeptides between 4 and 12 kDa during the isolation of polypeptide hormones from calf thymus[1–3].

Originally called UBIP, *"ubiquitous immunopoietic polypeptide"*, it was judged to be likely universal to life as it had been found in all tested guinea pig tissues as well as all tested cell types of other organisms[2]. Soon afterwards its sequence was partially determined using Edman sequencing, and its name was shortened to its modern one, ubiquitin[4]. Several publications followed, revealing how ubiquitin is attached to histone 2A by an isopeptide linkage between the carboxyterminal glycine of ubiquitin and a lysine in histone 2A[5–7]. While by 1977 the initially postulated role of ubiquitin as a peptide hormone involved in the immune response had already been put into question[7,8], its role in protein degradation had not been determined.

In parallel, A. Ciechanover, A. Hershko, and I. Rose, who all received the Nobel prize for chemistry in 2004 for their discoveries, were investigating components of an ATP-dependent proteolysis system. The first factor they identified was named APF-1, *"ATP-dependent proteolytic factor 1"*[9], and was shown to be necessary and sufficient to induce ATP-dependent proteolysis of target proteins[9–11]. Subsequently, APF-1 was shown to actually be ubiquitin[12]. This discovery suggested that the bond between APF-1/ubiquitin and histone 2A was the same as the one between APF-1 and other proteolytic substrates[3,12].

In publications following the identification of ubiquitin, the future Nobel prize winners and colleagues succeeded in purifying the enzyme responsible for ubiquitin activation[13]. This was initially done using traditional protein purification methods, such as lysate fractionation and chromatography. Soon, however, they devised a protocol to purify the ubiquitin activating enzyme, now named E1[14], by using its enzymatic activity to couple it to column matrix-attached ubiquitin[15]. Using this method, they also succeeded in identifying further components of the ubiquitin-ligase system, which they named E2 and E3[16]. By exchanging E1 for E2 on their ubiquitin-bound column and investigating reductant sensitivities they could elegantly show that the activated ubiquitin was transferred from the E1 to the E2 as a thioester; further experiments showed that the third component (E3) was required to transfer ubiquitin to substrates[16].

Date(published)	Date(submitted)	Remarks	Reference
01.01.1975	30.09.1974	First isolation of ubiquitin, named UBIP. Wrongly claims it causes lymphocyte differentiation (due to an endotoxin contamination). Ubiquitous presence noted but wrongly extended to prokaryotes (sample contaminated with yeast extract).	Goldstein et al. 1975[2]
20.05.1975	30.12.1974	Edman sequencing of tryptic fragments of 74 amino acid ubiquitin.	Schlesinger et al. 1975[4]
04.08.1975	09.06.1975	Protein A24 is Histone 2A + unknown.	Goldknopf & Busch 1975[5]
24.01.1977	29.11.1976	Protein A24 is Histone 2A + Ubiquitin N-terminus + unknown.	Hunt and Dayhoff 1977[7]
01.03.1977	29.11.1976	(Ubiquitin) diglycine is isopeptide linked to Histone 2A lysine 119.	Goldknopf & Busch 1977[6]
28.04.1978	08.03.1978	Reticulocyte fraction I contains ~9kDa protein of unknown identity required for ATP-dependent proteolysis together with fraction II.	Ciechanover et al. 1978[9]
10.02.1979	02.06.1978	Endotoxin-free ubiquitin is not involved in immune response. Protein A24 - Ubiquitin linkage confirmed.	Low & Goldstein 1979[8]
01.07.1979	26.03.1979	APF-1 name used for ATP-dependent proteolysis factor 1; ~450 kDa APF-2 identified, probably the 26S proteasome[17].	Hershko et al. 1979[10]
01.03.1980	10.12.1979	Reversible covalent binding of APF-1 to reticulocyte components.	Ciechanover et al. 1980[18]
01.04.1980	10.12.1979	First mention of ubiquitin chains. Various substrates, including polylysine, investigated for discrete multiple APF-1 isopeptide conjugation. Free APF-1 is regenerated upon ATP removal.	Hershko et al. 1980[19]
25.08.1980	25.04.1980	APF-1 purified to show its specific activity.	Ciechanover et al. 1980[11]
25.08.1980	25.04.1980	Ubiquitin is APF-1.	Wilkinson et al. 1980[12]
01.02.1981	17.10.1980	E1 thioester by ATP activation of APF-1.	Ciechanover et al. 1981[13]
25.02.1981	15.10.1980	AMP activation of terminal glycine to form thioester and then isopeptide.	Hershko et al. 1981[20]
10.09.1981	05.03.1981	Previous reports of 74 aa ubiquitin lack the terminal diglycine due to tryptic like activity during tissue purification.	Wilkinson & Audhya 1981[21]
10.03.1982	07.10.1981	E1 purification via ubiquitin-sepharose column.	Ciechanover et al. 1982[15]
01.07.1982		Review, first use of the term “E1”.	Hershko and Ciechanover 1982[14]
10.07.1983	27.12.1982	E1 -> E2 -> E3 enzyme cascade for ubiquitin conjugation established.	Hershko et al. 1983[16]

Table I 1. Overview over the foundational publications on ubiquitin and ubiquitination. The table illustrates the different names used for ubiquitin, UBIP and APF-1 before and after their identity was established and includes in large part publications of winners of the 2004 Nobel prize for chemisty: A. Ciechanover, A. Hershko, and I. Rose.

2. *Basic paradigm of Ubiquitination*

The central paradigm of how ubiquitin is attached to substrates using a cascade of enzymes, established in the original publication by Hershko et al. in 1983[16], has stood the test of time and is still virtually identical in current publications on ubiquitination (see e.g. Wang et al. 2020[22], Deol et al. 2019[23]) (Figure I 1).

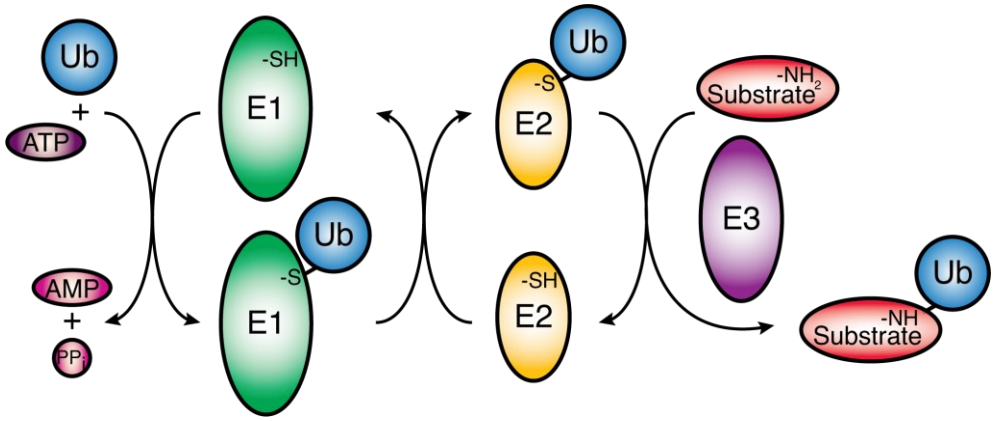


Figure I 1. The ubiquitination enzyme cascade: Ubiquitin is activated using ATP and successively transferred to a substrate through a series of enzymes named E1, E2 and E3.

In a first step, the carboxy terminus of ubiquitin is activated by a ubiquitin activating enzyme (E1) using ATP to form ubiquitin-AMP. The activated ubiquitin is then transferred to the E1 catalytic cysteine as a thioester, releasing AMP. Subsequently, the ubiquitin thioester is transferred to the catalytic cysteine of a ubiquitin conjugating enzyme (E2). In the final step, a ubiquitin ligase enzyme (E3) assists in the transfer of the activated ubiquitin to a substrate as an isopeptide.

3. *Models and mechanisms of chain formation*

The textbook model for ubiquitination (Figure I 1) allows for the amines of all 7 lysines of ubiquitin as well as its N-terminal amine to serve as further acceptor sites for ubiquitin conjugation. Already in 1980, before their elucidation of the E1-E2-E3 enzyme cascade, Hershko et al. recognized that, due to the preference for the elongation of an existing conjugated ubiquitin over fresh conjugation to a substrate, this process must be either processive or at least driven by a higher preference for proteins with ubiquitin already attached[19]. The source of this apparent processivity of chain assembly has been a matter of debate since [23,24], leading to the proposal of multiple models (Figure I 2).

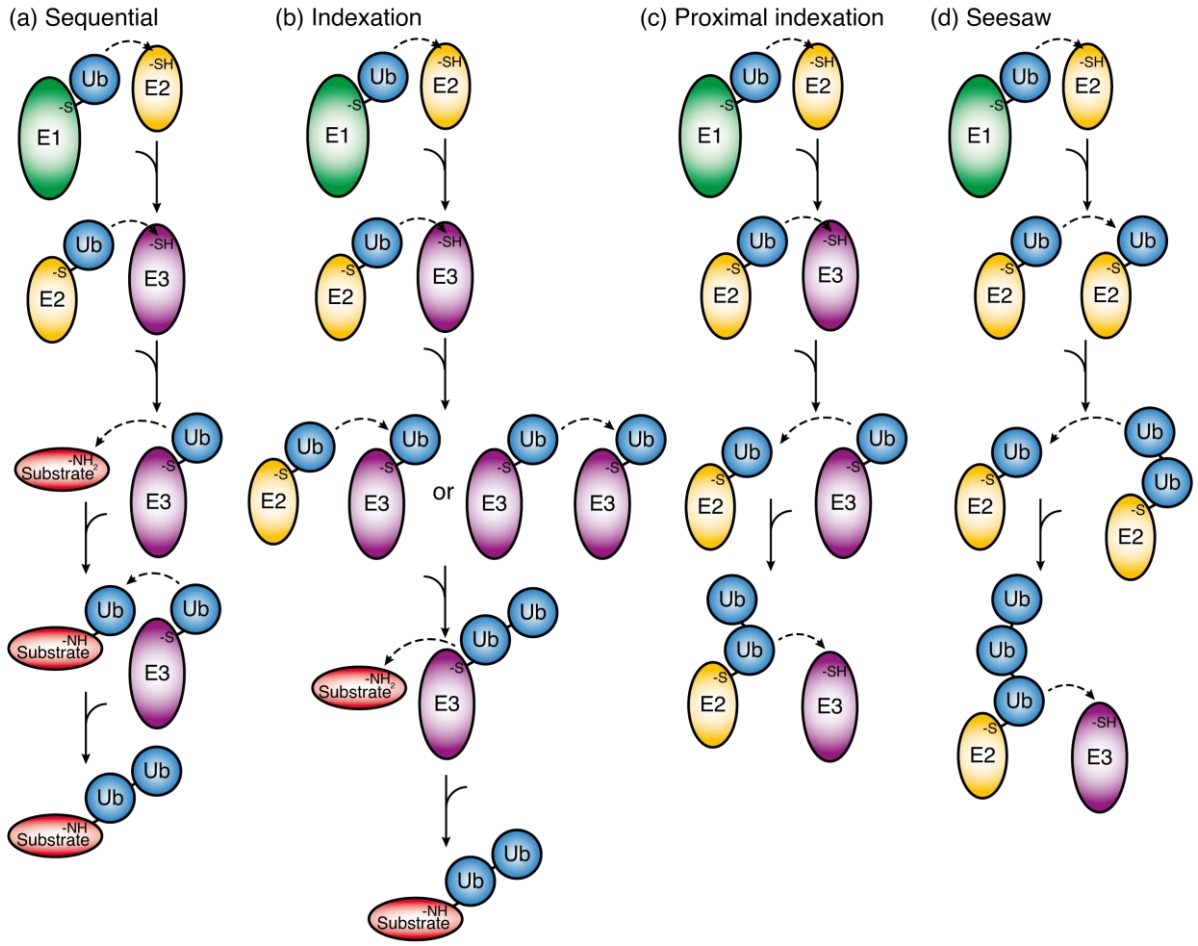


Figure I 2. Proposed models of chain formation. Sequential (a) and indexation (b) models as proposed in Wang and Pickart 2005[25] or Verdecia et. al. 2003[26]. Proximal indexation model (c) proposed by Ronchi et al. 2017[27]. Seesaw model (d) and figure concept from Hochstrasser 2006[24].

The standard model is the sequential addition model, in which a single ubiquitin is transferred through the enzyme cascade and added to the last ubiquitin of a growing chain on the substrate (Figure I 2(a)).

As an alternative to the sequential addition model, *en bloc* transfer models have been proposed to explain the apparent processivity of ubiquitin chain formation. In these models, the ubiquitin chain is thought to be pre-assembled and transferred to the substrate as a whole. Multiple potential *en bloc* models, differing in which component(s) of the E1-E2-E3 cascade carry the pre-assembled ubiquitin chain, have been reviewed in Hochstrasser, 2006[24].

The *en bloc* model hypothesis originated from the observation that the subdomains of HECT ("Homologous to the E6-AP Carboxyl Terminus") domain-containing E3 enzymes could assume different conformations, suggesting that these conformations might exist to accommodate a growing

ubiquitin chain assembled on the E3 via an *indexation* mechanism[26] (Figure I 2(b)). Indeed, in 2005, Wang and Pickart[25] found evidence that free chain formation catalysed by the K48-specific HECT-containing E3 ubiquitin ligase E6-AP ("E6-Associated Protein") could proceed by an indexation mechanism, while the HECT-containing E3 ligase KIAA10 (also called UBE3C), that assembles K48/K29-mixed chains, catalysed free ubiquitin chain formation via the sequential addition model. Since then, the pre-formation of chains on the cognate E2 enzymes of RING ("Really Interesting New Gene") domain-containing E3s has been described multiple times (for E2/E3 pairs UBC13/HLTF[28] and UBE2G2/GP78[29]) by direct visualisation of the E2~ubiquitin-chain thioesters.

An additional *en bloc* model, termed *proximal indexation*, has been proposed for HECT domain E3 ligases by the group of Arthur L. Haas. In this model, suggested by *in silico* evidence for a cryptic second E2~ubiquitin binding site and supported by kinetic data on ubiquitin chain assembly, the ubiquitin chain is transferred between E2 and E3, with a new ubiquitin subunit added at the E2/E3 proximal end of the chain at every step [27,30–33] (Figure I 2(c)).

Besides distal and proximal indexing mechanisms which could of course also be imagined with the chains being assembled on / by different members of the cascade, hybrid models as mixtures between both and models including the possibility that whole chains are passed between different members can be envisioned[24].

Evidence supporting *en bloc* transfer of ubiquitin chains from E2s~chain thioesters has been observed in multiple publications[34–36]. It has also been shown that the E1 enzyme is able to activate not only monomeric ubiquitin, but also free ubiquitin chains[37], suggesting that ubiquitin chains conjugated to the E2 might originate from naturally occurring free chains. Such ubiquitin chains can be synthesized *in vitro* by E3 ligases[25] and are *in vivo* the result of deubiquitinating enzyme (DUB)[38] activities, such as Rpn11 of the proteasome[39]. It is thought that these free chains are toxic *in vivo* and are therefore rapidly degraded by additional deubiquitinating enzymes[38], however their physiological role has not been studied in detail so far [38,40].

In sum, the question of which mechanism is used in chain assembly remains unsolved due to the scarce evidence unequivocally supporting any of the mechanisms. It is also conceivable that different mechanisms are being employed in the different enzymes of the cascade to different degrees, or that the ubiquitination mechanism could change between different enzymes of the same family.

4. Ubiquitin chain linkages

1. Physiological roles of different ubiquitin chain linkages

In ubiquitin chains, the carboxy group of the terminal glycine76 of one ubiquitin is linked to an amine in a preceding ubiquitin. This linkage can be established as an isopeptide bond with the ε-amino group of any of the seven lysines (K6, K11, K27, K29, K33, K48, K63)[41], or as a peptide bond to the free N-terminal amine of ubiquitin (M1)[41,42]. It is now known that ubiquitin conjugation can have different outcomes, besides the originally identified proteasomal degradation, and that the linkage type of ubiquitin chains modulates these different effects.

The chain linkage responsible for the first identified role of ubiquitination, targeting substrates for proteasomal degradation, was identified early on as K48[43]. A K48-linked ubiquitin chain of at least four ubiquitins was found to be required for efficient targeting[44], and structural investigations of the proteasome confirmed that the two proteasome subunits with the highest affinity for ubiquitin chains, Rpn10 and Rpn13[45] could indeed be bridged by a K48-linked chain of four ubiquitins[46–48]. More recently, it has however been shown that ubiquitination at multiple substrate lysines with K48 chains as short as two ubiquitins can, at least in some cases, provide a more efficient signal for proteasomal degradation[49]. K48-linked chains are the most abundant ubiquitin chains in cells[50] and, due to their involvement in the majority of intracellular protein turnover, are essential in all eukaryotes[41,51,52].

K63-linked ubiquitin chains are the second most abundant[50]. These modifications, unlike K48 linkages, are involved in several non-degradative processes, such as NF-κB signalling, DNA damage response, oxidative stress response as well as protein trafficking and sorting.

For example, the transcription factor NF-κB modulates the response to cytokines (Interleukin-1), viral antigens (via Toll-like receptors) and other stress stimuli. Its activation is mediated by a ubiquitin ligase, TRAF6 ("TNF receptor-associated factor 6"), which synthesizes K63-linked ubiquitin chains on itself and other substrates[53]. These K63 chains recruit and activate TAK1 ("TGF-beta-activated kinase 1"), which in turn phosphorylates the IKK ("IκB kinase") complex, including IKKα, IKKβ and NEMO[53]. IKK then phosphorylates IκB, leading to NF-κB translocation to the nucleus and the activation of its transcriptional activity[53].

In the DNA damage response, the nonhomologous end joining (NHEJ) pathway in double strand repair (DSR) is regulated by K63 ubiquitination of histones, mediated by the E2/E3 pairs UBC13/RNF8 as well as UBC13/RNF168[54]. The generated ubiquitin chains serve as a recruitment platform for NHEJ repair factors. Homologous recombination (HR), the other major DSR pathway together with NHEJ, is also regulated by ubiquitination: the ubiquitination and subsequent degradation of the Ku complex, which binds DNA double-strand break ends and is required for NHEJ[54], promotes HR.

In yeast, the E2/E3 ligases Rad6/Bre1 have been shown to generate K63-linked ubiquitin chains on several targets, including the ribosome, allowing for a rapid response to oxidative stress by down-regulating translation globally[55].

Besides targeting proteins for proteasomal degradation via K48-linked ubiquitin chains, ubiquitination is also involved in the other main cellular degradation pathways, lysosomal and vacuolar degradation[56–59] While mono- and multiubiquitylation are reported to be sufficient for the internalization of membrane proteins[60,61], K63-linked ubiquitin chains are essential for the sorting of membrane proteins into multivesicular bodies (MVBs) by the ESCRT ("Endosomal Sorting Complexes Required for Transport") apparatus, and their subsequent lysosomal / vacuolar degradation [56,57]. Interestingly, the HIV-1 Gag protein is able to hijack the ESCRT machinery for viral budding and release by interacting with NEDD4 ("Neuronal precursor cell-expressed developmentally downregulated 4") family ubiquitin ligases, which synthesize K63-linked ubiquitin chains[62].

The only M1 ubiquitin ligase known to date is the Linear Ubiquitin Assembly Complex (LUBAC). LUBAC is recruited in the above-mentioned NF-κB signalling pathway by interaction with K63-linked ubiquitin chains and in turn modifies them with M1-linked ubiquitin chains[63]. NEMO, a component of the IKK complex, has high affinity to M1-linked ubiquitin chains (compared to K63-linked chains) and is thus more efficiently recruited to the kinase, TAK1, responsible for its activation [53]. Interestingly, the pathway is also linked to canonical K48 ubiquitination, as, upon phosphorylation, IκB is ubiquitinated by $SCF^{\beta\text{-}TrCP1}$ with K48-linked chains [64].

Like K48-linked chains, K11-linked chains target proteins for degradation. The anaphase promoting complex/cyclosome (APC/C) uses the E2 UBE2S to K11 ubiquitinate cell cycle regulators[65] such as Nek2A and cyclin A[66]. However, recent results have shown that homotypic K11-linked chains alone are not an ideal substrate for the proteasome and that APC/C synthesizes instead K11/K48 branched chains, which are a better proteasome substrate[67,68].

Comparatively little is known about the other atypical linkages.

K6-linked ubiquitin chains, assumed to play non-degradative roles as their amount does not increase upon proteasome inhibition[69], have been shown to be generated by the BRCA1/BARD1 E3 ligase complex, which is involved in the regulation of the DNA damage response[54]. The E3 ubiquitin ligase Parkin positively regulates mitophagy via K6 ubiquitination of mitochondrial substrates, acting together with USP30, a deubiquitinase with a K6 linkage preference serving as a negative regulator[70,71].

K27-linked ubiquitin chains have been found to be synthesized by the above-mentioned E3 ligase RNF168, which is involved in DNA damage response. However, this observation is based on

overexpression of the ligase in cells together with FLAG-tagged ubiquitin[54,72]. The only other E3 reported to assemble K27-linked chains is HACE1, which functions in protein secretion and autophagic flux[65].

The E3 ubiquitin ligase KIAA10 (UBE3C) generates ubiquitin chains with mixed K29/K48 linkage. Interestingly, the Rpn13 subunit of the proteasome is reported to be a substrate of this modification, that is suggested to negatively regulate its ability to bind ubiquitin conjugates[73]. Trabid, a deubiquitinase for K29-linked ubiquitin chains, is involved in T-cell differentiation[74].

The only known K33 E3 ligase, AREL1 ("Apoptosis-resistant E3 ligase 1"), is sparsely studied. This modification is suggested to be involved in degradation and Golgi membrane protein trafficking[65].

2. Known mechanisms of ubiquitin chain linkage specificity

Since the type of ubiquitin chain linkage has been shown to regulate the physiological effect of the modification on the substrates, the mechanisms that ubiquitin ligases use to determine linkage specificity are of special interest.

Of the components of the ubiquitination cascade (E1, E2 and E3), only the last two (E2 and E3) can confer linkage specificity, as there are only two E1 enzymes in humans (Uba1(Ube1) and Uba6). Whether it is the E2 or the E3 that determines linkage specificity differs depending on the E3 family. The approximately 40 human E2 ligases do have varying degrees of linkage specificity[75] by themselves, but can only exert this when in combination with an E3 that does not impart its own linkage specificity.

Structural details of how linkage specificity is regulated by E2s have been shown for several enzymes. For the heterodimeric Mms2/Ubc13 E2, channels that could bring the K63 of the acceptor ubiquitin close to the catalytic cysteine were proposed based on the apo structure[76], experimentally validated[77] and finally shown in a structure with both donor and acceptor ubiquitin bound, occupying the postulated channel[78], structurally explaining the preference for the formation of K63-linked ubiquitin chains. Comparison of the acceptor ubiquitin binding surfaces of Ubc13 (K63 specific) and UbcH5A (K11 preference) led to the identification of a single amino acid largely responsible for linkage specificity[79].

In the E2 ligase UBE2S a substrate assisted catalytic mechanism was observed, in which a glutamate residue of the acceptor (substrate) ubiquitin can participate in catalysis only when the acceptor K11 is correctly placed[80].

Other investigations have confirmed ubiquitin-E2 interacting residues working to place the acceptor ubiquitin to reach linkage specificity for several E2s, such as Cdc34[81,82] and Ubc1[83].

While the RING family of E3 ubiquitin ligases is solely acting as a scaffold mediating the interaction between an E2 and the substrate, in the other two known families of E3 ubiquitin ligases, the HECT and RBR families, the E2 transfers the activated ubiquitin to a catalytic cysteine residue of the E3 itself[84], which in turn transfers it to the substrate. Comparatively less is known about the mechanisms of linkage specificity of these last two families.

D. E3 ligase families

1. RING

The RING ("Really Interesting New Gene") family of E3 ubiquitin ligases is the most numerous E3 family in the human genome, with over 600 members [85].

Its members are characterized by the RING finger motif, a zinc finger type domain containing the sequence Cys^{①}-X_2-Cys^{②}-$X_{(9\text{-}39)}$-Cys^{③}-$X_{(1\text{-}3)}$-His^{④}-$X_{(2\text{-}3)}$-Cys^{⑤}-X_2-Cys^{⑥}-$X_{(4\text{-}48)}$-Cys^{⑦}-X_2-Cys^{⑧} (X: any residue; subscript: number of residues; circled superscript: as in Figure I 3)[86], also known as a C_3HC_4 type zinc finger. This motif differs from other/classical zinc finger motifs in that the binding residues for the two bound zinc ions are interleaved in the sequence in a "cross-brace" manner[86], in which the

first two cysteines bind the first zinc ion together with the fifth and sixth, whereas the third, fourth, seventh and eight cysteine bind the second zinc ion.

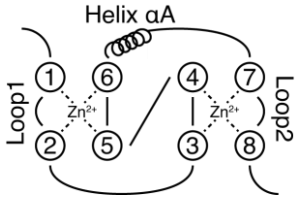


Figure I 3. Schema of the cross-brace fold of the RING domain.

In many RING domains the E2-binding interface is conserved and consists of a helix (Helix αA Figure I 3) downstream of the sixth zinc binding residue together with the first and last loop (Loop1, Loop2 Figure I 3) which connect the zinc binding residues[87]. However, there are also RING E3s which bind E2s via regions outside of the RING motif[87].

Substrate binding is typically not mediated by the RING domain itself, but by other domains either of the RING-containing protein itself or of associated proteins, as in the APC/C ("Anaphase-promoting complex/cyclosome") or the SCF ("Skp1-Cullin-F-box") E3 ligase complexes.

Mechanistically, RING E3 ubiquitin ligases act purely as a scaffold between the E2 carrying the activated ubiquitin and the target substrate. Unlike other E3 ligases, such as RBR and HECT family members, the activated ubiquitin is not transferred to the RING E3; instead, the E2~ubiquitin thioester is directly attacked by a substrate amine. Because of this, the linkage specificity of RING E3s typically depends on the E2 alone[88,89].

2. RBR

Originally the RBR ("RING1-in-Between-RINGs-RING2") family of E3 ligases was defined at the sequence level[90] by the presence of two RING domain (see section above) consensus motifs, with another conserved domain motif in between called IBR ("In-Between-RINGs") domain[91]. More recently, structural analyses of a subset of the approximately 15 RBR ubiquitin ligases in the human genome[92–94] showed that, while the RING1 domain is a typical RING domain, the RING2 and IBR adopt a very similar fold in which the zinc ions are coordinated not in a "cross-brace", but in a sequential manner[94].

Unlike in RING ligases, in RBR enzymes the RING1 domain recruits the E2 ligase carrying the activated ubiquitin. Subsequently, the activated ubiquitin is trans-thioesterified to a cysteine in the RING2 domain, from where it can be transferred to a substrate or a growing ubiquitin chain[93]. Thus, RBR ligases form a, albeit transient[93], thioester bond with ubiquitin and are themselves responsible for transfer specificity to a substrate or ubiquitin chain, thereby determining chain linkage type[42,95]. For these reasons, they were introduced as having a HECT-like mechanism[84].

The best researched RBR E3 ligase is HOIP, the catalytically active member of the LUBAC, which synthesizes specifically M1 chains[42]. This specificity however is not inherent to the RBR itself, but conferred by a downstream motif which positions the acceptor ubiquitin[96]. Determination of linkage specificity by the E3 has also been observed for another RBR ligase, RNF216, which produces K63-linked ubiquitin chains despite interacting with a K48-synthesizing E2[95].

However, other RBR family members do not possess inherent specificity and rely on the E2 enzyme to determine chain linkage specificity. The E3 ARIH2 ("Protein ariadne-2 homolog"), for example, produces K48-linked chains with the E2 Ube2L3, K63 with Ube2N/Ube2V2, and mixed linkages with the promiscuous Ube2D3. Similarly, Parkin synthesizes K63-linked chains with Ube2N/Ube2V2, but can also catalyse the formation of K48 chains [93].

5. *The HECT E3 ligase family*

The name of the HECT domain comes from "*homologous to the E6-AP carboxyl terminus*"[97], as this domain was first identified in E6-AP and subsequently found in 28 genes in the human genome[85,98,99]. Of these 28 HECT-containing proteins, 10 have been structurally characterized (Table I 2). Since the HECT domain in isolation is sufficient for catalysis and it is solely responsible for the linkage of the generated ubiquitin chains, studies often focus on the isolated HECT domain.

The HECT domain is approximately 45 kDa and, despite being called a domain, actually consists of two lobes: a 15 kDa C-terminal lobe (C-lobe) as well as an 30 kDa N-terminal lobe (N-lobe)[100], which usually have the ability to fold independently (i.e. Matta-Camacho et al. 2012[101]) (Figure I 4). The N-lobe can be further subdivided in a smaller and a larger subdomain[102,103].

The strictly conserved, catalytically active cysteine, that receives the activated ubiquitin via trans-thiolation from the E2, is located in the C-lobe[104–106] (Figure I 4).

Enzyme	Macromolecules	Shape	Comment	PDB IDs	Reference
E6AP	HECT + E2	L	E2 binding site identified	1C4Z, 1D5F	Huang et al. 1999[100]
WWP1	HECT	T	apo.	1ND7	Verdecia et al. 2003[26]
Smurf2	HECT	L	apo.	1ZVD	Ogunjimi et al. 2005[103]
NEDD4L	HECT	T	apo.	2ONI	SGC 2007 (unpublished)
NEDD4L	HECT + E2~Ub	T	E2 binding site	3JVZ, 3JW0	Kamadurai et al. 2009[107]
Huwe1	HECT	T	apo.	3G1N	SGC 2009 (unpublished)
Huwe1	HECT	T	apo.	3H1D	Pandya et al. 2010[108]
Rsp5	HECT + Ub(in UBS)	L(n.c.)	UBS (ubiquitin exosite) bound ub	3OLM	Kim et al. 2011[109]
ITCH	HECT	T	apo.	3TUG	SGC 2011 (unpublished)
Nedd4	HECT	L(n.c.)	apo.	2XBF	Maspero et al. 2011[102]
Nedd4	HECT	T	apo., A889F	4BBN, 4BE8	Maspero et al. 2013[105]
Rsp5	WW3-HECT~Subst.~Ub	L	substrate ubiq. transition state	4LCD	Kamadurai et al. 2013[104]
Nedd4	HECT + inhibitor(in UBS)	L(n.c.)	small molecule binding to UBS	5C91	Kathman et al. 2015[110]
WWP2	HECT	T	apo.	4Y07	Gong et al. 2015[111]
WWP1	HECT + UbV(in UBS) + E2	T	UbV binding to UBS	5HPT	Zhang et al. 2016[112]
Rsp5	HECT + UbV(in UBS)	T	UbV binding to UBS	5HPL	Zhang et al. 2016[112]
NEDD4L	HECT + UbV(in UBS)	T(n.c.)	UbV binding to UBS	5HPK	Zhang et al. 2016[112]
NEDD4	HECT + UbV(in UBS)	L(n.c.)	UbV binding to UBS	5C7J	Zhang et al. 2016[112]
WWP1	HECT + UbV(in UBS)	T	UbV binding to E2-BS	5HPS	Zhang et al. 2016[112]
ITCH	HECT + 2*UbV	T	UbVs binding to E2-BS and UBS(n.c.)	5C7M	Zhang et al. 2016[112]
Huwe1	HECT	T	dimer crystal	5LP8	Sander et al. 2017[113]
WWP2	HECT + WW2-L	T	autoinhibited by non-HECT elements	5TJ7,5TJQ,5TJ8	Chen et al. 2017[114]
ITCH	HECT + WW1-WW2-L	T	autoinhibited by non-HECT elements	5XMC	Zhu et al. 2017[115]
WWP1	HECT + WW2-L + WW4	T	autoinhibited by non-HECT elements	6J1X	Wang et al. 2019[116]
WWP1	HECT + L-WW3-WW4	T	autoinhibited by non-HECT elements	6J1Y	Wang et al. 2019[116]
WWP2	HECT + WW2-L	T	autoinhibited by non-HECT elements	6J1Z	Wang et al. 2019[116]
AREL1	HECT	T	apo.	6JX5	Singh et al. 2019[117]
UBE3C	HECT	T	apo.	6K2C	Singh and Sivaraman 2020[118]
Huwe1	full-length	T	first full-length HECT E3 structure (cryoEM)	7JQ9 (held)	Hunkeler et al. 2020 (preprint)

Table I 2. Overview of available HECT domain structures sorted by publication date. (n.c.: non-canonical)

Both lobes are joined by a linker, whose flexibility is thought to be required for activity[26,100,102,105,107,119,120]. This flexibility can be observed in the various structures of HECT domains determined so far (Table I 2), which differ in how N-lobe and C-lobe are oriented with respect to each other. Largely, the available structures can be divided in a set with a L-conformation and another with a ⊥ or inverted T-conformation. In the L-conformation, the C-lobe sits at one end of the more elongated N-lobe, while in the T-conformation the C-lobe is found in the middle of the N-lobe. However, even within the inverted T-conformations the C-lobe can be rotated with respect to the N-lobe to various degrees. Interestingly, all auto-inhibited forms of HECT domains show a nearly identical inverted T conformation in which the catalytic Cys is buried at the lobe interface. In contrast, catalytically active HECT domains possess a more open conformation where the catalytic Cys is exposed.

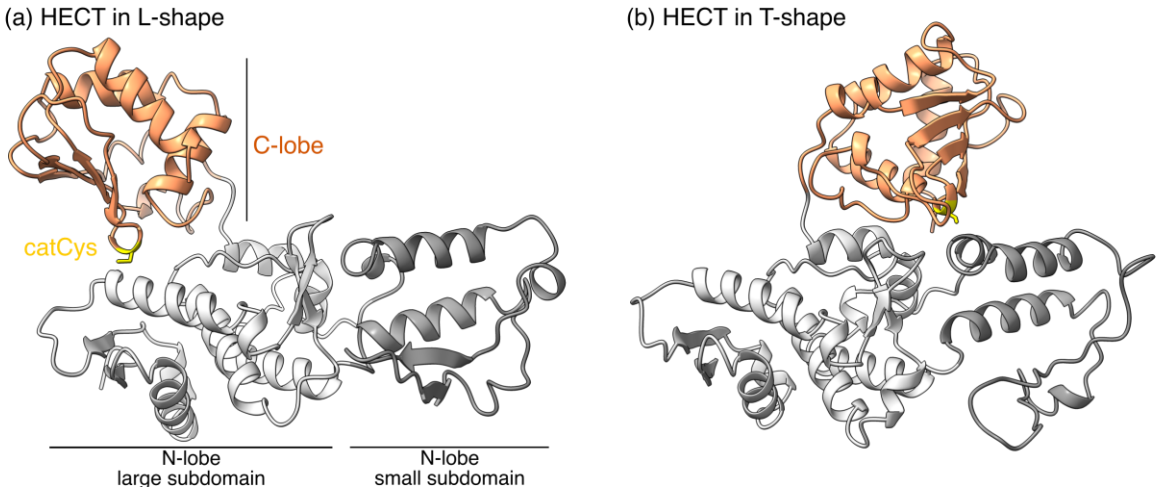


Figure I 4. Exemplary T- and L-shape structures from PDB IDs (a) 1C4Z (E6-AP)[100] and (b) 1ND7 (WWP1)[26].

E6-AP, the first HECT domain to be structurally characterized, was crystallized in complex with an E2, which binds to the smaller subdomain of the N-lobe[100]. This E2 binding site has since been found to be conserved and has been validated by subsequent structural studies[107,112].

While this initial E6-AP HECT-E2 complex structure crystallized in an L-conformation, the inverted T-conformation is likely to be closer to a E2-E3 trans-thiolation competent state[26], since the respective catalytically active cysteines are closer (~8 Å in Kamadurai et. al. 2009[107]) in the T-conformation than in the L-conformation (40 Å [100]).

Further insights into the role of the L-conformation has been gained from the structure of Rsp5, a Nedd4 family HECT E3, in a conformation analogous to the transition state of substrate ubiquitination[104]. In this structure, the substrate binding domain adjacent to the HECT domain is bound to a substrate peptide, whose acceptor residue is chemically linked to the catalytic cysteine of the HECT domain and ubiquitin. Thus, with the L-conformation required for ubiquitin to reach the lysine of the substrate bound to the substrate binding domain and the inverted T-state required for E2-E3 trans-thiolation, the necessity for the exchange between the two forms becomes apparent. However, while this structure gave insight into lysine selection of a small peptide, how ubiquitin lysine selection in polyubiquitination happens remains unexplained.

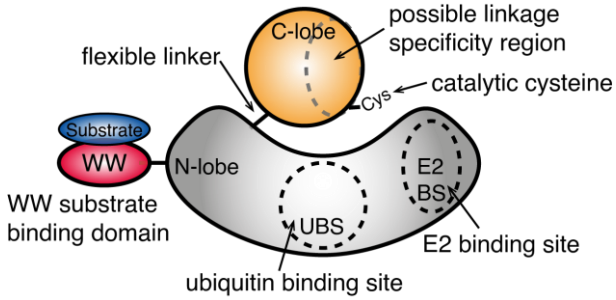


Figure I 5. Schema of a typical HECT domain.

Another important feature in the HECT domain is the absolute C-terminus, which is usually either unstructured, and therefore not visible in the electron density[105], or only visible due to (unphysiological) crystal contacts[104]. At position -4 (or -5 for Nedd4), there is a conserved phenylalanine, that, together with the absolute C-terminal (pos. -1) residue of the HECT domain, is essential for isopeptide formation, but not for E2 to E3 trans-thiolation[105–107,121]. The conserved phenylalanine is likely involved in positioning the C-terminus, and a corresponding pocket of residues that could rescue its mutation have been found in the N-lobe of Rsp5[104,122], a Nedd4 family HECT E3. The role for the remaining C-terminal residue(s) is less clear, but a catalytic triad, consisting in Nedd4

of Asp900 (-1 position), His865 and Cys867(catalytic cysteine), and conserved amongst Nedd4 family HECT E3s, has been proposed in analogy to a similar triad found in RBR E3s[94,98].

The HECT domain also contains a non-covalent ubiquitin binding site (UBS) on the N-lobe, which is sometimes also referred to as the ubiquitin exo-site[112]. This UBS was found to exist so far in all Nedd4 family HECT E3s, besides the little investigated HECW1 and HECW2[102,103,105,109,112,120,123,124]. Which non-Nedd4 family HECT domains contain a UBS is less clear. The binding affinity of the UBS for ubiquitin, however, varies greatly in Nedd4 family HECT domains. The affinity for monomeric ubiquitin is too weak to be detectable via pulldowns in ITCH, WWP1, WWP2 and E6AP[102] or via fluorescence anisotropy in WWP1[123]. However, evidence supporting the existence of a UBS like surface has been found in ITCH, WWP1 and WWP2 from crystal structures together with ubiquitin variants (UbVs)[112,114] which had been optimized by phage display for high affinity.

Among the crystal structures solved so far (see Figure I 6 and Table I 2), both the orientation of ubiquitin as well its binding site are conserved, and the location of the UBS has been confirmed in solution via NMR in multiple cases[120,125].

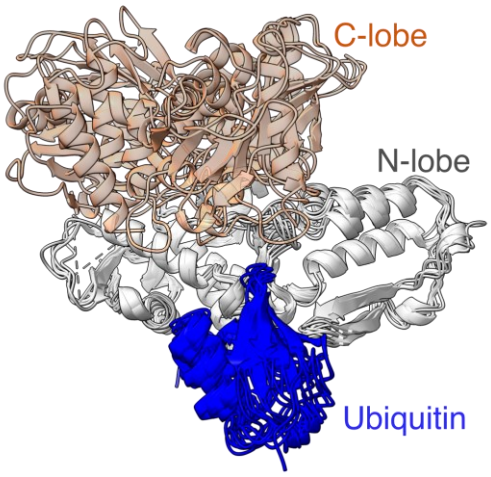


Figure I 6. Overlay over all published HECT domain - ubiquitin co-crystal structures with ubiquitin in the UBS (see Table I 2). Aligned on the highly conserved N-lobe with the C-lobe in a number of orientations between the T- and L- states. The orientation of ubiquitin with respect to the N-lobe is largely conserved.

Interestingly, ubiquitin binding stabilizes the conformation of a beta-hairpin in the smaller N-lobe subdomain and also the overall conformation of the N-lobe, which otherwise shows small deviations (<4°) in the angle between its small and large subdomains[26,100,102,107].

While it is tempting to assume that the UBS is the site of the acceptor ubiquitin, this is sterically unlikely, as the UBS-bound ubiquitin K63, the acceptor lysine in the Nedd4-family[105], is on the ubiquitin side opposite to the catalytic cysteine and thus out of reach even when taking the flexible linker attaching the HECT C-lobe into account. However, the UBS was found to be essential for efficient polyubiquitination[102,105,120]: mutation of critical residues in the UBS or of a hydrophobic patch on ubiquitin important for the interaction[102], as well as blocking the UBS with an inhibitor[110], abolish effective chain elongation. Various mechanisms for how the UBS could contribute to efficient polyubiquitination have been proposed: Ogunjimi et al. 2010 presented a model in which the interaction with monoubiquitinated substrates is stabilized[120], while Maspero et al. 2011 suggested a "slippage" mechanism, whereby the UBS would bind the terminal acceptor ubiquitin in the chain specifically, but loosely enough to allow for its elongation[102]. The same group further advanced the possibility of a mechanism, in which the second ubiquitin after the terminal acceptor ubiquitin in the ubiquitin chain could be positioned in the UBS to allow for elongation of the terminal acceptor ubiquitin[98].

Substrate binding is typically mediated by domains N-terminal to the HECT domain, such as the BH3 domain in Huwe1[126], the E6-binding region in E6-AP or the WW domains in the Nedd4 HECT family[127], which can either recruit substrates directly or via adaptor proteins[128].

Mechanistically, HECT ligases, unlike RING E3 ligases, receive the activated ubiquitin from the E2 before transferring it to a substrate or a growing ubiquitin chain. Consistently, members of the HECT family of E3s have intrinsic, E2 independent chain specificities[129]. The element sufficient and responsible for chain specificity had early been identified to be the catalytically active HECT domain itself[25,130]. Successively, in an elegant study employing chimeras between two HECT domains of differing linkage specificity, the part responsible for the linkage specificity was identified as the last 60 amino acids of the C-terminal lobe of the HECT domain[131].

The same study proposes a model whereby the positioning of the thioester-linked donor ubiquitin could control the positioning of the acceptor ubiquitin lysine for conjugation, in analogy to a linkage specificity mechanism previously identified in a E2 ligase[80]. Upon subsequent crystallization of a HECT domain together with a E2 carrying activated ubiquitin that showed a possible donor-ubiquitin - HECT C-lobe interface, this theory still remained a favourite[107]. Even if the way in which the acceptor ubiquitin is bound by the HECT domain remains elusive so far, defined residues on the side of the acceptor ubiquitin could be identified to be crucial for its utilization[132].

More recently, evidence has been found that the last three C-terminal residues of the HECT domain are involved in linkage selection in experiments where the C-terminus of the K48-specific E6AP enabled the K63-specific Nedd4 ligase to synthesize K48-linked chains[105].

E3	Linkage
E6AP	K48[25]
AREL1	K33/K11[133,134]
Huwe1	K6/K11/K48[135]
ITCH	K63[136]
Nedd4	K63[137]
Nedd4L	K63[105]
Rsp5	K63[138]
Smurf2	K63[105,131]
WWP1	K63>K48>K11[123]
WWP2	K63[105,129,131]
UBE3C	K48[118]

Table I 3. HECT ubiquitin ligases with their linkage specificities if known.

1. Regulation

As members of the HECT family of E3 ubiquitin ligases are crucial for the regulation of many cellular processes through the ubiquitination of various effector substrates, their activity must be tightly regulated both to control the rate of ubiquitination and to avoid off-target ubiquitination. In addition to mechanisms external to the function of the E3 proteins, like the regulation of their expression and substrate selectivity due to substrate binding domains recruiting different substrates to the E3 or the removal of ubiquitin from substrates by deubiquitinases, a number of auto-regulatory mechanisms exist for E3 proteins.

One such mechanism is autoubiquitination, whereby in the absence of substrate[104,139] and other inhibitory mechanisms (see below) the HECT E3 can ubiquitinate itself thereby triggering its degradation[140]. However, in addition to this inhibitory function, autoubiquitination has also been found to upregulate E3 function, for example by favouring the recruitment of substrates containing ubiquitin-binding domains[141]. Limiting autoubiquitination through other inhibitory mechanisms is

nevertheless beneficial to the cell, as it prevents futile cycles of autocatalytic ubiquitination leading to proteasomal degradation and thus requiring re-expression of the ligases[121].

One of the first reported autoinhibitory mechanisms of an HECT E3 itself was observed in the Nedd4 family protein Itch, whose WW domain region, N-terminal to the HECT domain, interacts with and inhibits the HECT domain if a proline rich linker region nearby is not phosphorylated by JNK1[142].

The first structural evidence for an autoinhibitory mechanism, however, was shown for the C2 domain, common to all Nedd4-family HECT E3s, which was found to exert an autoinhibitory effect by interacting with the HECT domain in the case of the Nedd4-family members Smurf2[121], Nedd4[121] and WWP2[121]. However, while present in all Nedd4-family HECT E3s, the C2 domain has been shown not to be responsible for autoinhibition in Rsp5[121], Itch[121,143] and Smurf1[144–147].

In the case of Itch, a more recent structural study identified a HECT-binding motif consisting of a WW domain and its following α-helical linker[115]. A similar WW domain - α-helical linker - WW domain motif has also been identified in other two Nedd4-family E3s, WWP1[116] and WWP2[114,116] and might also play a role in Nedd4 and Smurf2[116]. Additionally, our lab showed that the whole module comprising C2-domain – linker – WW-domain is responsible for binding the HECT domain, and thus mediating its autoinhibition, in Smurf2 but not Smurf1[147]. A concomitant publication from a different group showed that this motif is conserved in Nedd4[116].

C2 release can be mediated by calcium in Nedd4[148] and Nedd4L[148,149], and our lab demonstrated that this was due to the overlap of the binding surfaces for calcium and the HECT domain on the C2 domain of Nedd4[150].

Besides the aforementioned phosphorylation leading to activation in Itch, the C2 – HECT interaction in Nedd4[151] as well as the α-helical – HECT interaction in WWP2[114] have also been shown to be released from autoinhibition by phosphorylation.

In the case of Smurf2, an adaptor protein, Smad7, was shown to release autoinhibition by displacing the C2 domain from the HECT domain via binding to WW domains in Smurf2 via its proline rich (PPxY) binding motif[121]; this mechanism was also observed for WWP1[144].

Subsequently, many other examples of Nedd4-family HECT E3s being activated by displacing inhibitory domains from binding to the HECT domain have been found. For example, Nedd4L is activated by one of its substrates, ENaC[152]; WWP2 is activated by Dishevelled[146]; Itch can, in addition to phosphorylation, be activated by binding of the substrate adapter proteins NDFIP1/2[115,116,128,143,153] as can Nedd4, Nedd4L, Smurf2, WWP1 and WWP2[128].

In cases where binding surfaces of the inhibitory elements on the HECT domain are known to overlap with the UBS, this blocking explains the abrogation of effective polyubiquitination[120]. However, this cannot explain cases where the inhibited step has been determined to be the ubiquitin thioester transfer from the E2 to the E3, and not the E3 to substrate isopeptide transfer (see Table I 4).

During the course of this doctoral work, our lab and others[98,114–116,147] suggested that the common element of these autoinhibitory mechanisms are contacts between both lobes of the HECT domain to lock it[98,114–116,125] in an *inactive*[114] or *closed*[147] conformation where the catalytic cysteines of the E2 and E3 are possibly too far from each other to allow for trans-thiolation or where the catalytic cysteine is buried at the lobe interface.

In sum, while locking the HECT in an inactive/closed conformation may be a common pattern, not all Nedd4 ligases use the same mechanism of autoinhibition despite their common domain architecture. Also, little is known about the regulation of non Nedd4-family members.

Enzyme	inh. element	HECT surface	inhibited step	r.f.	Reference
Smurf2	n.d.	n.d.	n.d.	Smad7	Ogunjimi et al. 2005[103]
ITCH	L-WW1-4	n.d.	n.d.	phos.	Gallagher et al. 2006[142]
Smurf2	C2	includes N-lobe	thioester	Smad7	Wiesner et al. 2007[121]
Nedd4	C2	n.d.	n.d.	n.d.	Wiesner et al. 2007[121]
WWP2	C2	n.d.	n.d.	n.d.	Wiesner et al. 2007[121]
Nedd4L	WW1-4	n.d.	n.d.	ENaC	Bruce et al. 2008[152]
Smurf1	1,2-L	n.d.	n.d.	CKIP1	Lu et al. 2008[154]
ITCH	L-WW1-4	n.d.	n.d.	Ndfip1/2	Mund and Pelham 2009[128]
Nedd4	C2	n.d.	n.d.	Ndfip1/2	Mund and Pelham 2009[128]
Nedd4	C2	n.d.	n.d.	Ca^{2+}	Wang et al. 2010[148]
Nedd4L	C2	n.d.	n.d.	Ca^{2+}	Wang et al. 2010[148]
Smurf1	C2(trans dimer)	n.d.	n.d.	Cdh1	Wan et al. 2011[155]
E6AP	HERC2-binding$_{(150-200)}$	n.d.	n.d.	$HERC2^{RLD2}$	Kühnle et al. 2011[156]
Nedd4	C2	n.d.	n.d.	phos.	Persaud et al. 2014[151]
Nedd4L	C2	C-lobe not sufficient	n.d.	IP_3/Ca^{2+}	Escobedo et al. 2014[149]
Nedd4	C2	n.d.	thioester	n.d.	Mari et al. 2014[125]
Smurf2	C2	C-/N-lobe	thioester	n.d.	Mari et al. 2014[125]
ITCH	WW1-4	C-/N-lobe	thioester	Ndfip1	Riling et al. 2015[143]
E6AP	n.d.	n.d.	thioester	E6	Mortensen et al. 2015[157]
WWP1	C2 and/or WWs	n.d.	n.d.	Smad7	Courivaud et al. 2015[144]
Smurf1	*const. active.*	n.d.	n.d.	n.d.	Courivaud et al. 2015[144]
WWP2	C2-WW?	n.d.	n.d.	Dvl2, DEP	Mund et al. 2015[146]
Smurf1	WW?	n.d.	n.d.	Dvl2	Mund et al. 2015[146]
WWP2	WW2-L	C-/N-lobe	n.d.	phos.	Chen et al. 2017[114]
ITCH	(WW1-)WW2-L	C-/N-lobe	thioester	Ndfip1	Zhu et al. 2017[115]
dmITCH	WW12-L	n.d.	n.d.	dNdfip	Yao et al. 2018[153]
WWP1	WW2-L-WW3(4)	C-/N-lobe	thioester	n.d.	Wang et al. 2019[116]
WWP2	WW2-L-WW3(4)	C-/N-lobe	n.d.	n.d.	Wang et al. 2019[116]
ITCH	WW12-L-WW34	n.d.	n.d.	n.d.	Wang et al. 2019[116]
Nedd4	C2-WW1-L	n.d.	n.d.	n.d.	Wang et al. 2019[116]
Smurf2	C2-WW1-L	n.d.	n.d.	n.d.	Wang et al. 2019[116]
Smurf2	C2-WW1	C-/N-lobe	n.d.	n.d.	Ruetalo et al. 2019[147]
Smurf1	C2(only in trans)	includes N-lobe	n.d.	n.d.	Ruetalo et al. 2019[147]

Table I 4. Autoinhibitory mechanisms of HECT family E3 ligases. Given are the inhibitory element (always part of the HECT E3 itself), the surface on the HECT domain the inhibitory element has been shown to bind to, wether the catalytic step that was inhibited was determined as well as release factors (r.f.) that release the autoinhibition. (n.d. = not determined in respective publication)

E. Aims and significance of this Thesis

The overarching goal of this thesis was to gain a more comprehensive understanding of how a family of ubiquitin ligases, the HECT ligases, function, focusing especially on open questions in the field such as the regulation mechanisms of non-Nedd4-family HECT ligases, chain linkage selectivity as well as chain length control.

Ubiquitination is involved in the majority of cellular processes, not only in protein turnover as was initially believed, but also in other regulatory mechanisms. Thus, it is no surprise that several HECT E3 ubiquitin ligases have been found to be involved in diseases such as different types of cancer[22,158–160], neurodegenerative diseases[22], as well as viral infections, even bridging the gap between a viral infection and cancer in the case of the eponymous member E6-AP (E6 - associated protein) being hijacked by the human papilloma virus (HPV) E6 protein to degrade the p53 tumour suppressor protein.

Despite their important roles, little is known on the mechanistic details of HECT E3 ubiquitin ligases chain assembly on substrates. This is perhaps best exemplified by the scarce evidence on how members of the HECT family generate ubiquitin chains of different linkages (Table I 3) despite being structurally highly similar (Table I 2). A related open question is how the ubiquitin chain is bound by the ligase during its most likely processive elongation/assembly. Finally, and of utmost clinical relevance, is the need for a deeper understanding on how HECT domains are regulated – as for example Huwe1 has been described as both an oncogene as well as a tumour suppressor[161] – exemplifying the requirement for a tight regulation of these crucial enzymes *in vivo* and underlining the interest in drugs to correct their dysregulation in disease.

Previous investigations have clearly identified the C-lobe of the HECT domain as playing a role in thioester formation and chain specificity[25,80,102,105,107,109,130–132]. These topics are the focus of Chapter A, where we compare the interface between the C-lobe and the donor ubiquitin of two mechanistically different HECT domains, one with K48 (Huwe1) and one with K63 (Smurf2) ubiquitin chain-linkage specificity.

Members of the NEDD4 subfamily of the HECT domain family of E3 ligases are regulated via interactions between N-terminal domains and the catalytically active C-terminal HECT domain. This has been shown for various E3 ligases by our lab[121,125,147] as well as others (Table I 4). In contrast to this, the low baseline activity of Huwe1 has previously been attributed to a dimerization-dependent autoinhibitory mechanism[113]. However, since in our hands the Huwe1 HECT domain constructs did not dimerize in solution (as estimated from the τ_c influence on NMR spectra quality), we instead hypothesized an intra-domain regulatory mechanism. Indeed, we showed that the interface between the two lobes of the HECT domain in Huwe1, unlike in other HECT domains, is in a closed conformation in its "ground state". The results, presented in Chapter B, confirm the existence of this C-/N-lobe interface and its functional relevance.

Several members of the HECT E3 ubiquitin ligases have a ubiquitin binding exo-site (UBS) on their N-lobe. However, we had previously not observed ubiquitin binding to the Huwe1 HECT domain. The closed conformation found in Chapter B provided a possible explanation for this, raising the question of whether there is a functional UBS in Huwe1 that may not be accessible in its closed conformation. Indeed, the results of Chapter C demonstrate the existence of a UBS in Huwe1, which is only accessible for ubiquitin binding in its open conformation. Furthermore, we examined to which degree other HECT domains adopt a closed conformation in solution and whether the dynamic equilibrium between closed and open conformation correlates with ubiquitination activity and ubiquitin binding to the UBS.

During this work, I observed that it is crucial to use the native ubiquitin sequence to avoid ubiquitination at the unnatural ubiquitin N-terminus as a side reaction. Chapter D thus presents an evaluation of multiple HECT domains together with common N-terminal extensions (tags) of ubiquitin, which have been widely employed in several biochemical experiments. The data presented show that these side reactions are not solely due to the promiscuity of the specific HECT domain employed, but constitute a general principle and an explanation for the conserved ubiquitin N-terminus structure/sequence.

Finally, Chapter E details work towards establishing a model system whereby the kinetics of substrate ubiquitination, and also ubiquitin chain elongation, can be observed in a step-wise manner. This work was undertaken to gain insight into the processivity of HECT domain E3 ligases and how chain length is controlled. Additionally, these kinetic measurements answer the question of whether the UBS binds to the second to terminal ubiquitin in the elongating ubiquitin chain, which was one of the previously proposed roles of the UBS.

II. Results

A. β-sheet augmentation is a conserved mechanism of priming HECT E3 ligases for ubiquitin ligation

1. Disclaimer

This chapter has been published in the Journal of Molecular Biology (JMB) 2018 issue 430, pages 3218-3233 under the title "β-Sheet Augmentation Is a Conserved Mechanism of Priming HECT E3 Ligases for Ubiquitin Ligation" by the authors (in order) Magnus Jäckl, *Carsten Stollmaier*, Timo Strohäker, Karolina Hyz, Elena Maspero, Simona Polo and Silke Wiesner[162].

In this work, which compares aspects of Huwe1 with Smurf2, with the exception of the initial C-lobe crystallization/structure solution and the pull down in Figure 1d, I planned, cloned the constructs for and performed all experiments involving Smurf2 constructs (NMR spectra, *in vitro* assays).

2. Abstract

Ubiquitin (Ub) ligases (E3s) catalyze the attachment of Ub chains to target proteins and thereby regulate a wide array of signal transduction pathways in eukaryotes. In HECT-type E3s, Ub first forms a thioester intermediate with a strictly conserved Cys in the C-lobe of the HECT domain and is then ligated via an isopeptide bond to a Lys residue in the substrate or a preceding Ub in a poly-Ub chain. To date, many key aspects of HECT-mediated Ub transfer have remained elusive. Here, we provide structural and functional insights into the catalytic mechanism of the HECT-type ligase Huwe1 and compare it to the unrelated, K63-specific Smurf2 E3, a member of the Nedd4 family. We found that the Huwe1 HECT domain, in contrast to Nedd4-family E3s, prioritizes K6- and K48-poly-Ub chains and does not interact with Ub in a non-covalent manner. Despite these mechanistic differences, we demonstrate that the architecture of the C-lobe~Ub intermediate is conserved between Huwe1 and Smurf2 and involves a reorientation of the very C-terminal residues. Moreover, in Nedd4 E3s and Huwe1, the individual sequence composition of the Huwe1 C-terminal tail modulates ubiquitination activity, without affecting thioester formation. In sum, our data suggest that catalysis of HECT ligases hold common features, such as the β-sheet augmentation that primes the enzymes for ligation, and variable elements, such as the sequence of the HECT C-terminal tail, that fine-tune ubiquitination activity and may aid in determining Ub chain specificity by positioning the substrate or acceptor Ub.

3. Introduction

The post-translational modification of proteins with ubiquitin (Ub) plays a key role in cellular signaling[163]. The ubiquitination reaction is catalyzed by an enzymatic cascade consisting of an activating (E1), a conjugating (E2) and a ligating (E3) enzyme. Ub can be attached to substrates as a monomer at a single or multiple Lys residues (mono-/multi-ubiquitination) or as a poly-Ub chain (poly-ubiquitination) by the repeated action of the ubiquitination enzyme cassette. All seven lysines (K6, K11, K27, K29, K33, K48 and K63) and the N-terminal M1 residue can serve as linkage points in the distal (acceptor) Ub for isopeptide bond formation with the C-terminal carboxyl group of the donor Ub (Ub^D) molecule during chain elongation[164]. Notably, many E3 enzymes seem to prioritize a certain Lys residue in the acceptor Ub to form poly-Ub chains[105,137]. The information encoded in the linkage types dictates the cellular fate of the target protein. K11- and K48-linked poly-Ub chains direct substrates to

proteasomal degradation[43,165], while mono-ubiquitination and K63-linked poly-Ub chains have non-proteolytic functions in the assembly and trafficking of cell surface receptors and signalling complexes[166]. It is thus important to understand how E3 ligases catalyze Ub chain assembly and what determines Ub chain length and linkage.

Homologous to the E6-AP carboxyl terminus (HECT)-type E3 ligases use a two-step mechanism to ubiquitinate substrates. First, the carboxyl group at the Ub C-terminus is transferred from an E2 thioester intermediate to the catalytic Cys of the HECT domain in a transthiolation reaction[97,167]. Then the HECT~Ub thioester is attacked by the ε-amino group of a Lys in the substrate or the acceptor Ub to ultimately form an isopeptide bond[168]. For this last reaction step, HECT domains require a strictly conserved Phe or Tyr at the −4 position relative to the HECT C-terminus[106]. Although HECT-type ligases contain a diverse array of domains upstream of the catalytic domain[169], the HECT domain alone can perform both catalytic steps in the absence of N-terminal domains. Moreover, due to the two-step mechanism, linkage specificity is largely determined by the HECT-type ligase itself, in particular by the C-terminal 60 aa of the HECT domain[131]. We have recently shown that exchanging even only the three C-terminal residues in neural precursor cell-expressed developmentally downregulated gene 4 (Nedd4) to those of the K48-specific E6AP E3 is sufficient to alter Ub chain specificity from K63 to a mixture of K48- and K63-linked chains[105]. However, the exact mechanisms underlying HECT-mediated poly-Ub chain formation are still unclear since the catalytically important C-terminal residues could not be resolved in crystal structures of HECT domain intermediates.

Among HECT-type ligases, the Nedd4 family forms the largest group with nine members in humans. This enzyme family contains an N-terminal C2 domain, two to three central WW domains and the C-terminal HECT domain[99]. Nedd4-family E3s predominantly synthesizes K63-linked poly-Ub chains[105,133], and in contrast to other HECT ligases, the structures of their HECT domains have been characterized at multiple steps along their catalytic pathway[104,105,107]. HECT domains consist of two lobes that are tethered by a flexible hinge loop[100]. The larger N-terminal lobe (N-lobe) of the HECT domain engages the E2[100] and contains a non-covalent Ub binding surface that is important for Ub chain elongation[102,109,112,120,123]. The smaller C-lobe bears the catalytic Cys that forms the thioester intermediate with the Ub C-terminus. Interestingly, both the orientation of the donor Ub toward the C-lobe and the relative arrangement of the N- and C-lobes are preserved upon Ub transfer from the E2 to the E3[105,107]. The energy for pushing the transthiolation reaction forward is provided by the formation of an additional β-strand between the C-lobe and the C-terminus of the Ub^{D}. For isopeptide formation, the Ub-loaded C-lobe rotates as a whole with respect to the N-lobe to present the thioester to a substrate Lys for nucleophilic attack[104]. Once the isopeptide bond has been formed, the C-lobe most likely discharges the substrate-conjugated Ub after ligation and subsequently switches back into a conformation that allows it to accept another Ub from the E2-thioester. Once a substrate is mono-ubiquitinated, a non-covalent Ub binding surface on the N-lobe facilitates chain elongation[102,109,112,120,123]. However, the mechanistic details underlying this phenomenon are unclear. Although these studies have remarkably furthered our understanding of the catalytic mechanism underlying HECT-mediated Ub transfer, they fall short of providing insights as to how the C-terminal region enables isopeptide formation and contributes to Ub chain selectivity. Moreover, so far, all structural information on catalytic intermediates stems exclusively from K63-specific Nedd4 family members. Therefore, it is unclear whether these mechanisms apply to other if not all HECT domains.

In contrast to the Nedd4 family, structural and mechanistic information on reaction intermediates of other HECT-type ligases is sparse. Moreover, while Nedd4 ligases primarily synthesize K63-poly-Ub chains[105,123], HECT domains outside this family seem to prioritize other linkages, for example, K48-chains in the case of E6AP[169], K48- and K29-chains for UBE3C (a.k.a. KIAA10 or RAUL)[133,134,170], K33- and K11-chains for AREL1[133] and K6-chains for the HECT-like bacterial ligase NleL[171]. HECT, UBA and WWE domain containing 1 (Huwe1; a.k.a. MULE, ARF-BP1, Lasu1 or HECTH9) was shown to generate K6-,

K48- and/or K11-linked chains[133,135,172,173]. However, reports exist that Huwe1 also synthesizes K63-chains[174]. N-terminal to its HECT domain, Huwe1 contains two Armadillo-like repeat domains, a Ub-associated (UBA) domain, a UIM (Ub Interacting Motif), a WWE protein interaction domain and a BH3 motif for binding its bona fide substrate Mcl-1[126,175]. Huwe1 regulates important cellular processes including apoptosis and DNA damage repair and is closely linked to tumor development by targeting various key oncoproteins and tumor suppressors such as ARF, p53, Myc and Miz1, for proteasomal degradation[158,175–177].

To further our understanding of HECT domains, we characterized the catalytic mechanism of the Huwe1 HECT domain and compared it in detail to the Nedd4-family member SMAD-specific E3 ubiquitin protein ligase 2 (Smurf2). We show that in contrast to Nedd4-family E3s, the Huwe1 HECT domain preferentially synthesizes K6- and K48-linked Ub chains and does not interact with monomeric Ub despite being capable of generating poly-Ub chains. Regardless of these mechanistic differences, we find that the crystal structures of the Smurf2 and Huwe1 C-lobe~Ub^{D} complexes are highly similar. By NMR spectroscopy, we found that in both cases, the C-terminal residues of the C-lobe are affected by thioester formation. As for Nedd4[105,137], we show that the presence and the protein sequence of the C-terminal residues are not crucial for transthiolation, but strongly affect isopeptide formation. Given that the orientation of the Ub^{D} with respect to the C-lobe is preserved among Nedd4 family members and Huwe1, this suggests that the variable sequence of the C-terminal residues may create HECT~Ub^{D} binding interfaces that are important for catalytic efficiency and potentially specific for certain Lys residues in the substrate or acceptor Ub.

4. Results

1. The Huwe1 HECT domain forms preferentially K6- and K48-linked Ub chains

To gain insight into Huwe1 Ub chain specificity, we performed auto-ubiquitination assays with the isolated Huwe1 HECT domain using fluorescently labelled WT and single Lys Ub mutants (Figure 1a). We found that the Huwe1 HECT domain was most active with WT Ub and readily produced poly-Ub chains with both K48- and K6-only Ub. For a more detailed analysis, we examined the Ub chain type specificity of the Huwe1 HECT domain with absolute-quantification (AQUA) assays[105,178]. To this end, three independent auto-ubiquitination reactions (Figure 1b, left panel) were subjected to AQUA mass spectrometric analysis and the formed di-Gly peptides were quantified (Figure 1b, right panel). Consistent with the auto-ubiquitination assays with the single Lys Ub mutants, the Huwe1 HECT domain produced predominantly K48- and K6-linked oligomeric Ub chains. For longer Ub chains, we observed a large fraction of K48-linked chains and smaller fractions of K6- and K11-linked chains (Figure 1b, right panel). In sum, we find that the Huwe1 HECT domain prioritizes K6-, K11- and K48-linked poly-Ub chains and thus differs in chain specificity from Nedd4-family members such as Smurf2, Nedd4 and WWP1[105,123].

2. The Huwe1 HECT domain barely interacts with Ub in a non-covalent manner

Numerous members of the Nedd4 family have been shown to contain a non-covalent Ub binding surface that promotes Ub chain elongation[102,109,112,120,123]. To examine whether this also applies to the Huwe1 HECT domain, we performed NMR titration experiments with a Huwe1 HECT domain that was ^{15}N-labeled and partially deuterated. To this end, we added unlabelled Ub to the Huwe1 HECT domain at a 3-, 6- and 9-fold stoichiometric excess. Comparison of the respective ^{1}H,^{15}N transverse relaxation-enhanced spectroscopy (TROSY) spectra (Figure 1c and Figure S 1a) showed that only very few peaks exhibited significant chemical shift changes (larger than one peak width) even at a 9-fold excess of monomeric Ub. These peaks were already slightly shifted (less than a peak width) at a 3-fold stoichiometric excess of Ub and continued to shift from a 6- to a 9-fold stoichiometric excess (Figure S 1a). This suggests that the Huwe1 HECT domain interacts with monomeric Ub only very weakly and that the affinity for this interaction lies in the millimolar range. For comparison, we added unlabelled monomeric Ub at a 2- and 6-fold stoichiometric excess to the ^{15}N-labeled Smurf2 HECT domain. In contrast to the Huwe1 HECT domain, we observed extensive line broadening and chemical shift changes for a large number of peaks already at a 2-fold excess of Ub (Figure S 1b). This is well consistent with the high molecular weight of the complex that results in line broadening and the interaction being in the medium to lower micromolar affinity range. Taken together, these results suggest that, in contrast to Nedd4-family E3, the interaction of monomeric Ub with the Huwe1 HECT domain is very weak and may not lie in a physiologically relevant affinity range.

Since Huwe1 mainly generates K48-linked Ub chains and Nedd4-family HECT domains prioritize K63-linked Ub chains, we performed pull-down assays with commercial K63- and K48-linked polyUb chains and the GST-tagged HECT domains of Nedd4, Nedd4L, Smurf2 and Huwe1 to compare their poly-Ub binding capabilities (Figure 1d). In addition, we used GST and the Nedd4 F707A mutant that is defective in non-covalent Ub binding in negative control experiments[105]. As expected for Nedd4-family ligases, we found that the HECT domains of Nedd4, Nedd4L and Smurf2 readily interacted with K63-linked poly-Ub chains (Figure 1d, left panel) and to a lesser extent with K48-Ub chains (Figure 1d, right panel). On the contrary, the Huwe1 HECT domain and the negative control proteins GST and the Nedd4 F707A mutant were unable to pull down K63- or K48-linked poly-Ub chains. We thus conclude that in contrast to Nedd4-family HECT domains, the non-covalent interaction between the Huwe1 HECT domain and Ub is too weak to be detected even in a context where avidity caused by long poly-Ub chains is strongly favoured. Whether this interaction, if existing, is physiologically relevant remains questionable.

Together with the chain specificity that differs from Nedd4 ligases, this may point toward a catalytic mechanism for Huwe1 that is distinct from the Nedd4-family HECT ligases.

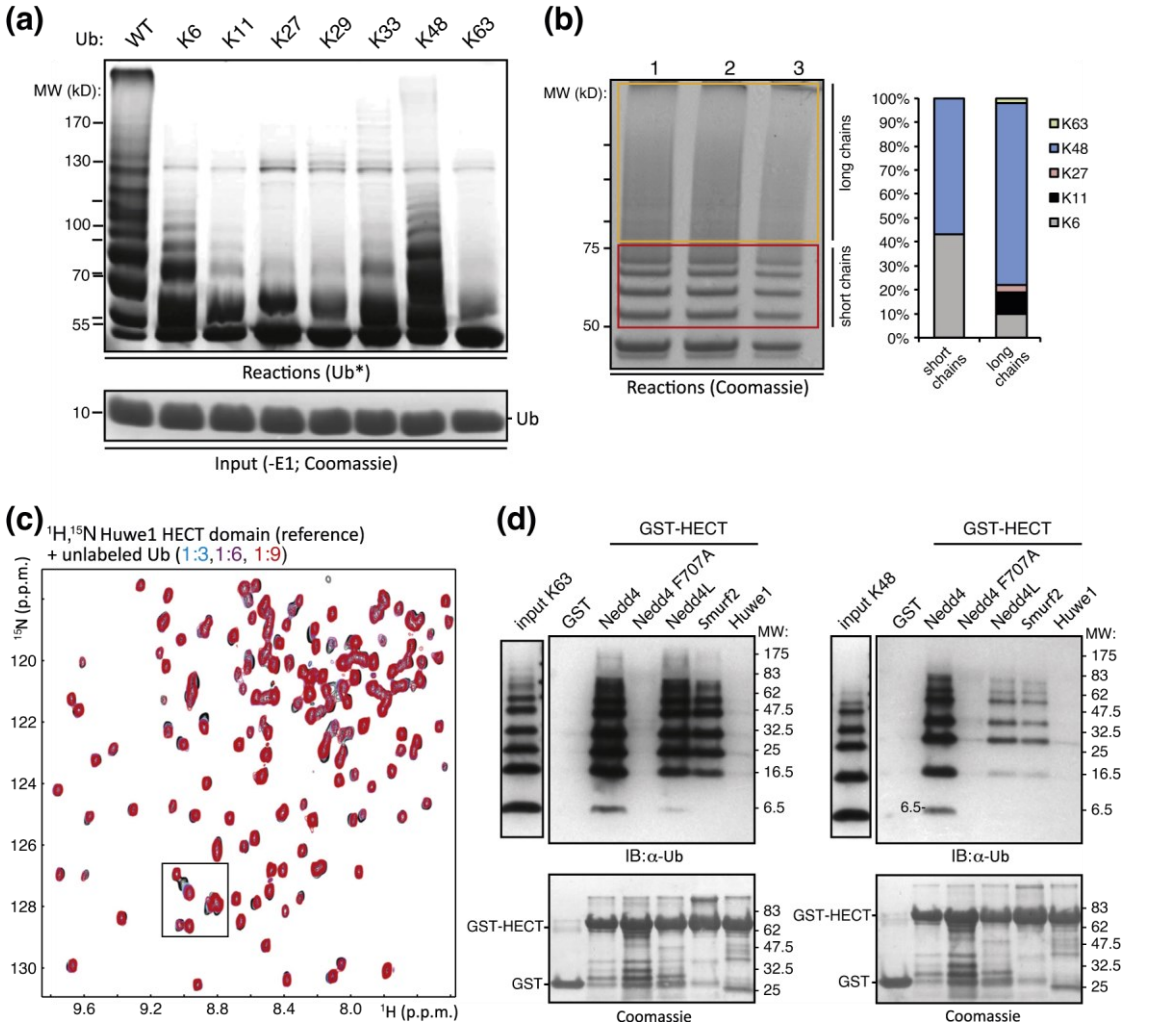


Figure 1. Ub chain type specificity and functional characterization of the Huwe1 HECT domain. (a) In vitro auto-ubiquitination of the Huwe1 HECT domain in the presence of E1 (Ube1) and E2 (UbcH7) with WT or single Lys Ub mutants as indicated. (b) AQUA analysis of poly-Ub chain linkages formed by the Huwe1 HECT domain. Left panel: SDS-PAGE gel of the poly-Ub chains formed by the Huwe1 HECT domain in the in vitro auto-ubiquitination reactions used for AQUA analysis. Right panel: AQUA proteomics quantification of the poly-Ub linkages showing the means and s.d. from measurements of the three reactions depicted in the gel on the left. (c) Overlay of a representative region of the $^{1}H,^{15}N$-TROSY NMR spectra of the Huwe1 HECT domain in the absence (black) and presence of increasing stoichiometric amounts of unlabeled monomeric Ub as indicated. The boxed area of the spectra is enlarged in Figure S 1a. (d) GST pull-down assays with commercial K63- (left panel) and K48-linked (right panel) poly-Ub chains and GST-tagged HECT domains or GST alone as indicated. Immunoblot (IB) and Coomassie as indicated.

3. The Huwe1 C-lobe~Ub^D complex is structurally highly similar to Nedd4 family E3s

To gain more detailed insight into the catalytic intermediates of the Huwe1 HECT domain and to compare it with HECT-type E3s of the Nedd4 class, we characterized the Huwe1 and Smurf2 thioesters on a structural and functional level. For structural analysis, we bypassed the instability of the thioesters by forming chemically similar, yet hydrolysis-resistant disulfides between a G76C mutant that served as donor Ub (Ub^D) and a Smurf2 or Huwe1 HECT domain that was Cys-free except for the catalytic Cys residue. Such disulfides have successfully been used to mimic E1~, E2~ and E3~Ub thioester intermediates for structural characterization[105,179–181].

Unfortunately, the Smurf2 HECT~Ub disulfide was prone to aggregation and did not crystallize. Since in the crystal structure of the Nedd4 HECT~Ub^D:Ub complex the donor Ub showed only limited contacts to the N-lobe[105], we generated a disulfide of the isolated Smurf2 and Huwe1 C-lobe. The Smurf2 C-lobe~Ub^D complex crystallized readily, and we solved its crystal structure at 2.5-Å resolution by molecular replacement (Figure 2a, Table S 1). The crystals contained two virtually identical complexes per asymmetric unit that superimpose with 0.49 Å r.m.s.d. of the backbone atoms (Figure S 2). The C-terminal tail of the Ub^D adopts an extended conformation with residues 73–75 forming an additional antiparallel β-strand (β') with the β9 strand of the Smurf2 C-lobe (Figure 2a, b). In addition, the side chain of R74 in the Ub^D forms hydrogen bonds with the backbone amides of Q673 and F674 in the α13 helix of the Smurf2 C-lobe (Figure 2b). Two additional non-covalent interaction surfaces complement β-sheet augmentation. Ub residues I36–Q40 are involved in an interaction network with Smurf2 residues N709 and L710 in the loop preceding the β9 strand and with residue T737 in the C-terminal α14 helix. Moreover, the Ub β1–2 loop (T7–T9) interacts with residues immediately C-terminal of the α14 helix of the Smurf2 C-lobe (E741–T742) (Figure 2b). Overall, the spatial arrangement of the Ub^D with respect to the Smurf2 C-lobe and the non-covalent interaction networks are highly similar to the Nedd4 HECT~Ub^D:Ub[105] and Rsp5 HECT~Ub^D:Sna3[104] complexes with an r.m.s.d. of 0.63 and 0.72 Å, respectively, for the backbone atoms (Figure 2c). We thus conclude that the overall architecture and interface composition of the C-lobe~Ub thioester complexes are highly conserved within the Nedd4 family of E3 ligases.

We then solved the Huwe1 C-lobe~Ub^D crystal structure at 2.8-Å resolution by molecular replacement (Figure 2d, Table S 2). The crystal contained one complex per asymmetric unit. Interestingly, many of the predominant intermolecular contacts are preserved between Huwe1 and the Nedd4 family, and, most importantly, the signature feature of the Nedd4 HECT~Ub thioester structures, the additional β-strand (β') between the Ub C-terminus and the C-lobe β9 strand (Figure 2d, e), is also present in Huwe1. While the overall architecture of the Huwe1 C-lobe~Ub^D complex is similar to Nedd4, Rsp5 and Smurf2, the spatial orientation of the Ub^D moiety is slightly rotated with respect to the C-lobe of the Nedd4-family structures (Figure 2f). This is reflected in a backbone r.m.s.d. of the Huwe1 C-lobe~Ub^D complex of 2.18 Å to Smurf2 for the entire C-lobe~Ub^D complex, while the C-lobes and Ub moieties alone align with a backbone r.m.s.d. of 0.80 and 0.47 Å, respectively. Of note, a symmetry-related Huwe1 C-lobe~Ub^D adduct intercalates into the Ub:C-lobe interface (Figure S 3a) in the crystal structure and thereby displaces the Ub moiety to some extent in comparison to the Nedd4-family Ub thioesters[104,105]. This affects most notably contacts of the Ub β1–β2 loop to the Huwe1 C-lobe (Figure S 3b) and restricts them to L8. Importantly, the extensive hydrogen bond network between the Ub^D and the C-lobe including the additional β' strand and the G35– Q40 loop is largely preserved between Huwe1 and Smurf2 (Figure 2b, e and Figure S 3c, d) and differs only in two contacts. The carbonyl oxygen of G35 in Ub forms a hydrogen bond with the ε-amino group of K734 in the Smurf2 C-lobe, while the side-chain carboxyl group of D39 forms a salt-bridge with the guanidinium group of R345 side chain in the Huwe1 C-lobe (Figure 2b, e and Figure S 3c, d).

In sum, we conclude that the overall structural arrangement of the Ub^D toward the HECT C-lobe is conserved among Nedd4 family members and Huwe1. This suggests that β-sheet augmentation is a

common conformation occurring during the HECT~Ub thioester formation irrespective of Ub chain specificity.

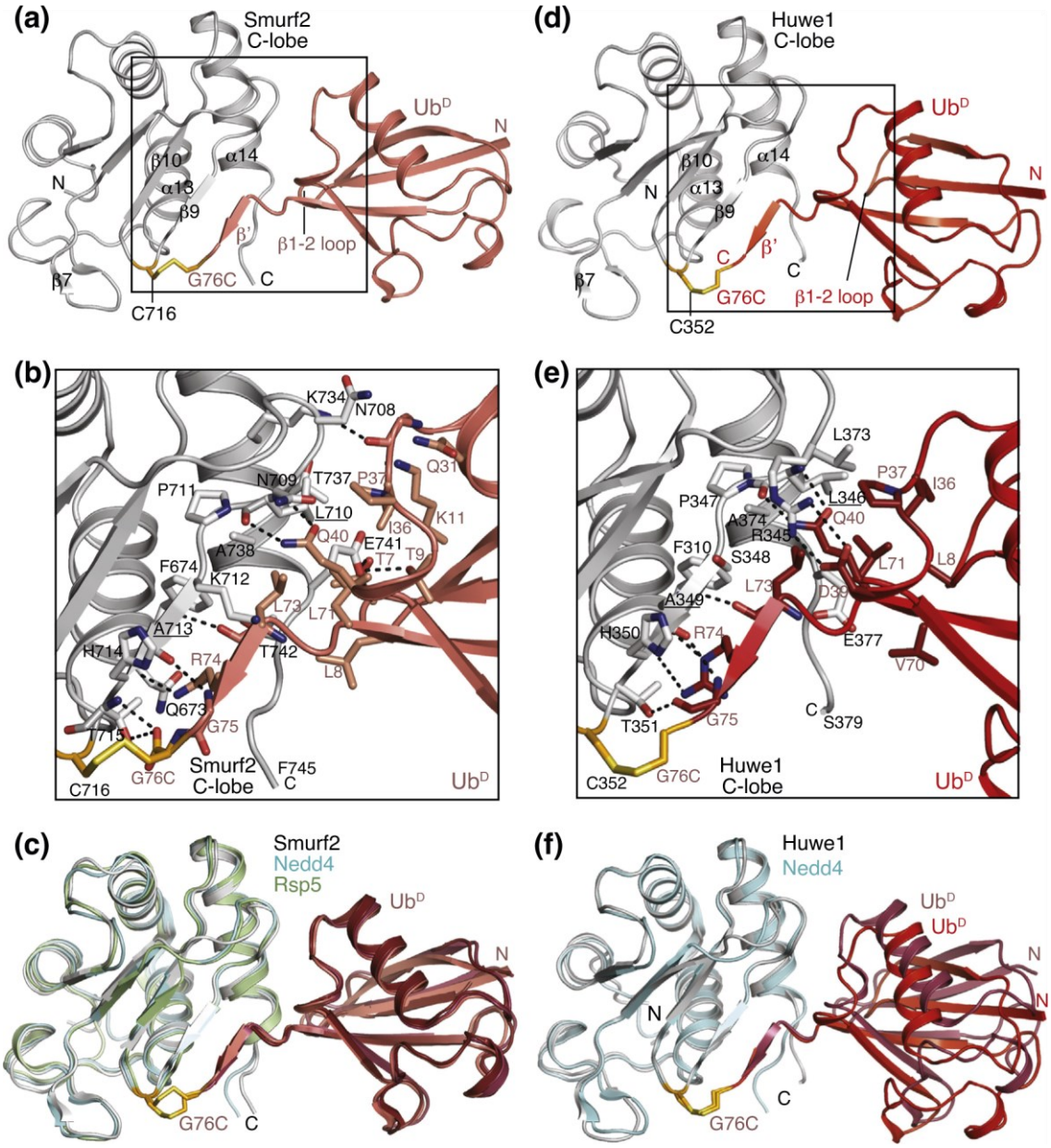


Figure 2. The structure of the Huwe1 C-lobe~Ub complex is highly similar to Nedd4 family HECT~Ub intermediates. (a) Ribbon representation of the x-ray structure of the Smurf2 C-lobe~Ub disulfide. Smurf2 is shown in gray and Ub G76C in salmon. The disulfide bridge (Smurf2 C716~Ub G76C) is depicted in stick representation. The region of the structure shown in panel b is indicated with a black box. (b) Zoom-in on covalent and non-covalent interactions that define the Smurf2 C-lobe~Ub^D complex. Hydrogen bonds are shown with dashed lines. (c) Overlay of the crystal structures of the Smurf2 (gray), Nedd4 (cyan) (PDB-ID: 4BBN[105]) and Rsp5 (green) (PDB-ID: 4LCD[104]) C-lobe thioester intermediates. (d) As panel a, but for the Huwe1 C-lobe~Ub disulfide. Huwe1 is shown in gray and Ub G76C in red. The disulfide bridge (Huwe1 C352~Ub G76C) is depicted in stick representation. For clarity, amino acid (aa) numbers for Huwe1 correspond to the HECT domain NMR construct (aa 1–385) that is equivalent to residues 3991–4375 in the Huwe1 sequence (NP_113584.3). (e) As panel b, but for the Huwe1 C-lobe~Ub^D complex. (f) Overlay of the crystal structures of the Huwe1 (gray) and Nedd4 (cyan) (PDB-ID: 4BBN[105]) C-lobe thioester intermediates.

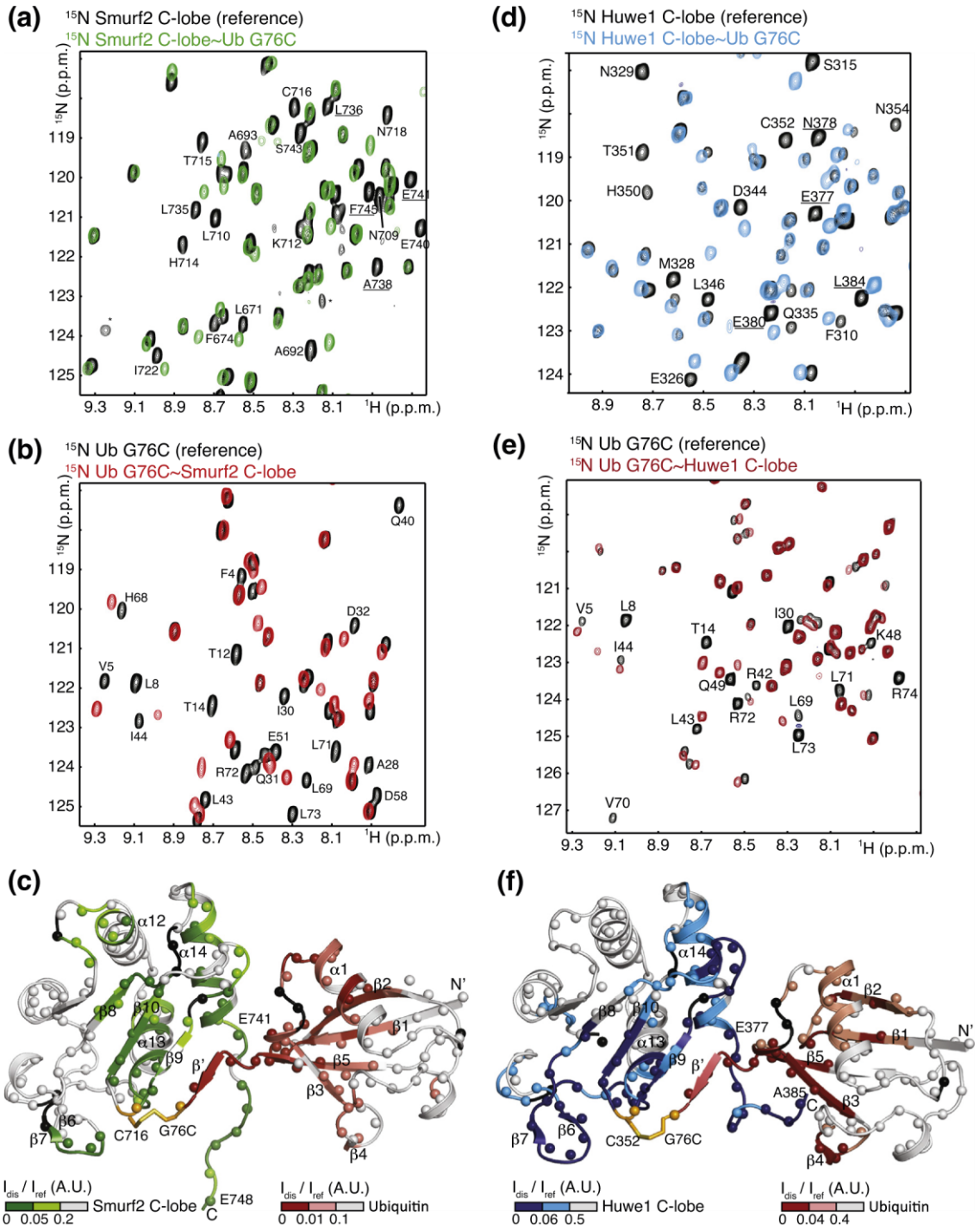


Figure 3. The C-terminal tail of the C-lobe is involved in thioester formation. (a) Overlay of a representative region of the $^{1}H,^{15}N$-TROSY spectra of the isolated ^{15}N-labeled C-lobe of the Smurf2 HECT domain (black) and the C-lobe in complex with unlabeled Ub G76C (green). Residues in the C-terminal tail of the C-lobe are underlined. (b) As panel a but for the isolated ^{15}N-labeled Ub G76C (black) or in complex with unlabeled Smurf2 C-lobe (red). (c) CSP mapping of the binding surface in the Smurf2 C-lobe and Ub G76C on the crystal structure of the Smurf2 C-lobe~Ub^{D} complex. CSPs were evaluated by comparing the signal intensities in the reference spectrum with those in the complex. Residues are color coded in green (C-lobe) or red (Ub) as indicated. Unassigned residues and prolines that are invisible in $^{1}H,^{15}N$-correlation spectra due to the lack of the amide proton are colored in black. The C-terminal tail of Smurf2 was modeled in MODELLER[182]. (d) As panel a, but for the isolated ^{15}N-labeled C-lobe of the Huwe1 HECT domain (black) and the C-lobe in complex with unlabeled Ub G76C (blue). (e) As panel d but for the isolated ^{15}N-labeled Ub G76C (black) or in complex with unlabeled Huwe1 C-lobe (dark red). (f) As panel c, but CSP mapping of the binding surface in the Huwe1 C-lobe and Ub G76C on the crystal structure of the Huwe1 C-lobe~Ub^{D} complex. Residues are color coded in blue (C-lobe) or red (Ub) as indicated.

4. Thioester formation affects the C-terminal tail of the Smurf2 and Huwe1

As for all HECT domain reaction intermediates[104,105,107], the very C-terminal residues of the Smurf2 (A746–E748) and Huwe1 (E380–A385) C-lobe could not be resolved in the crystal structures of the C-lobe~Ub^D complexes due to missing electron density. However, this region is well known to play a crucial role in isopeptide formation and Ub chain specificity[104–106,121].

To obtain structural insights into thioester formation in solution, we investigated the Smurf2 and Huwe1 C-lobe~Ub^D complexes by solution NMR spectroscopy. To this end, we assigned the back-bone chemical shift assignments of the isolated Smurf2 C-lobe and the Ub G76C mutant and compared the ^{1}H,^{15}N-TROSY spectra of the ^{15}N-labeled Smurf2 C-lobe in isolation and after disulfide linkage of the catalytic Cys to unlabelled Ub G76C (Figure 3a) and vice versa (Figure 3b). We observed that the peaks of numerous residues in the Smurf2 C-lobe and in Ub shifted by more than a peak width or were broadened upon Ub linkage (Figure 3a, b). To identify the residues affected by complex formation and to account for both peak shifts and peak broadening, we determined the relative signal intensities (Idis/Iref) of the free protein (Iref) and the disulfide complex (Idis) at the resonance frequencies of the free protein. For resonances not affected by ligand binding, the signal intensities were almost unchanged and thus Idis/Iref close to 1. In contrast, Idis/Iref was considerably smaller than 1 for residues that were affected by disulfide formation (Figure S 4a, b). We then mapped the obtained relative signal intensities for both binding partners on the crystal structure of the Smurf2 C-lobe ~ Ub^D complex and modeled the missing C-terminal residues of the C-lobe (A746–E748) for visualization (Figure 3c).

As expected from our crystal structure, the residues in the Smurf2 C-lobe affected by complex formation map to the immediate surroundings of the catalytic Cys (C716), that is, the β9-β10 loop, the β7 strand, the β9 strand involved in β-sheet augmentation, the neighboring β10 and β8 strands and the α13 helix, and the regions forming non-covalent contacts with the Ub^D such as the α14 helix (Figure 3c). Regions in the Smurf2 C-lobe displaying CSPs upon complex formation with Ub^D that are more remote from the interfaces observed in the crystal structure are residues in helix α12, but most notably included almost all residues C-terminal of the α14 helix and in particular the very C-terminal tail (A746–E748) that could not be resolved in the crystal structure (Figure 2b, Figure 3a, c and Figure S 4a).

As for the Smurf2 C-lobe, the CSPs in Ub were located to the region involved in β-sheet augmentation (aa 73–75, the β' strand) and non-covalent contacts (β1–β2, α1–β3 and β5–β' loops), but also regions in all β strands and the α1 helix that are not in direct contact with the C-lobe in the x-ray structure (Figure 3c and Figure S 4b). These CSPs may be caused by allosteric conformational changes in Ub upon thioester formation, by the C-lobe~Ub^D interface being more dynamic than the crystal structures would suggest, or by the C-terminal tail of the C-lobe interacting with the Ub^D in solution. In any case, the fact that all residues in the C-terminal tail of the Smurf2 C-lobe experience strong CSPs upon complex formation with the Ub^D demonstrates that the very C-terminal region of the Smurf2 HECT domain is affected by thioester formation.

We then extended these analyses to Huwe1 and acquired ^{1}H,^{15}N-TROSY spectra of a ^{15}N-labeled Huwe1 C-lobe in isolation or linked via a disulfide bond to unlabeled Ub G76C (Figure 3d) and vice versa (Figure 3e) and evaluated the observed CSPs as relative signal intensities in the same way as for Smurf2. As for Smurf2, we observed numerous CSPs in the Huwe1 C-lobe and in the Ub^D upon complex formation (Figure S 4c, d). Using backbone resonance assignments for the Huwe1 C-lobe, we mapped the observed CSPs on our crystal structure of the Huwe1 C-lobe~Ub^D complex and modeled the missing C-terminal residues (E380–A385) for visualization (Figure 3f). The CSPs mapped to the region close to the catalytic Cys (C352) and to regions that are largely consistent with the crystal structure of the Huwe1 C-lobe~Ub^D complex (Figure 3f and Figure S 4c, d). Similar to Smurf2, we found that residues in both the Huwe1 C-lobe and in the Ub^D that are not directly located at the interface exhibited substantial

CSPs. Most notably, also in the Huwe1 C-lobe, this included the entire C-terminal tail (E377– E385) that was partially invisible in the crystal structure (Figure 2e, Figure 3d, f and Figure S 4c).

Overall, our NMR studies of the Smurf2 and Huwe1 C-lobe~Ub complexes provide, to the best of our knowledge, the first direct structural evidence for a conserved conformational rearrangement of the C-terminal residues of HECT ligases occurring upon thioester formation.

5. Interfering with β-strand formation disrupts thioester formation

To examine the role of β-sheet augmentation in thioester and isopeptide formation, we mutated the central, conserved Ala in the β9 strand of the Smurf2 (A713) and Huwe1 (A349) HECT domains to proline (Figure 4a). We reasoned that this substitution should destabilize the hydrogen bond network between the C-lobe and the Ub^D C-terminus in the HECT~Ub thioester (Figure 4b), since prolines lack an amide proton. Indeed, while the WT HECT domains readily formed a Ub thioester that could be reduced with DTT, the Smurf2 A713P and Huwe1 A349P mutants were defective in thioester formation (Figure 4c, d) and consequently auto-ubiquitinantion (Figure S 5a, b). This shows that these residues in Smurf2 and Huwe1 are key to thioester formation and suggests that β-sheet augmentation indeed plays a crucial role in transthiolation as has been suggested for Nedd4[105].

To interfere with non-covalent interactions between the Ub^D and the C-lobe outside the region of β-sheet augmentation, we mutated a conserved Leu (Figure 4a) in the C-lobe of Smurf2 (L710) and Huwe1 (L346) to Pro. This Leu residue is located directly N-terminal of the β9 strand in the C-lobe and its backbone amide and carbonyl form hydrogen bonds to the side chain of Q40 in Ub in both the Huwe1 and the Smurf2 C-lobe~Ub^D crystal structures (Figure 2b, e). Similar to the mutations in the β9 strand, we found that these Pro substitution essentially abolished E2– E3 transthiolation (Figure 4c, d) and consequently poly-ubiquitination activity in both Smurf2 and Huwe1 (Figure S 5a, b). Of note, we verified by NMR spectroscopy that all Pro mutants were folded (Figure S 5c–f). Interestingly, previous biochemical analyses revealed that Q40 in the non-covalent interaction surface and two residues in the additional β-strand (L73 and R74) in the donor Ub are essential for ubiquitination activity in the K29- and K48-specific UBE3C HECT domain[132].

We thus conclude that irrespective of their distinct poly-Ub chain specificities, the crucial non-covalent contacts from the HECT C-lobe to the donor Ub are essential for thioester formation and are shared between Smurf2 and Huwe1 and potentially other HECT domains such as UBE3C. Given the distinct Ub chain specificities of these HECT domains, this strongly suggests that the orientation of the donor Ub alone with respect to the C-lobe is not a determinant of Ub chain specificity.

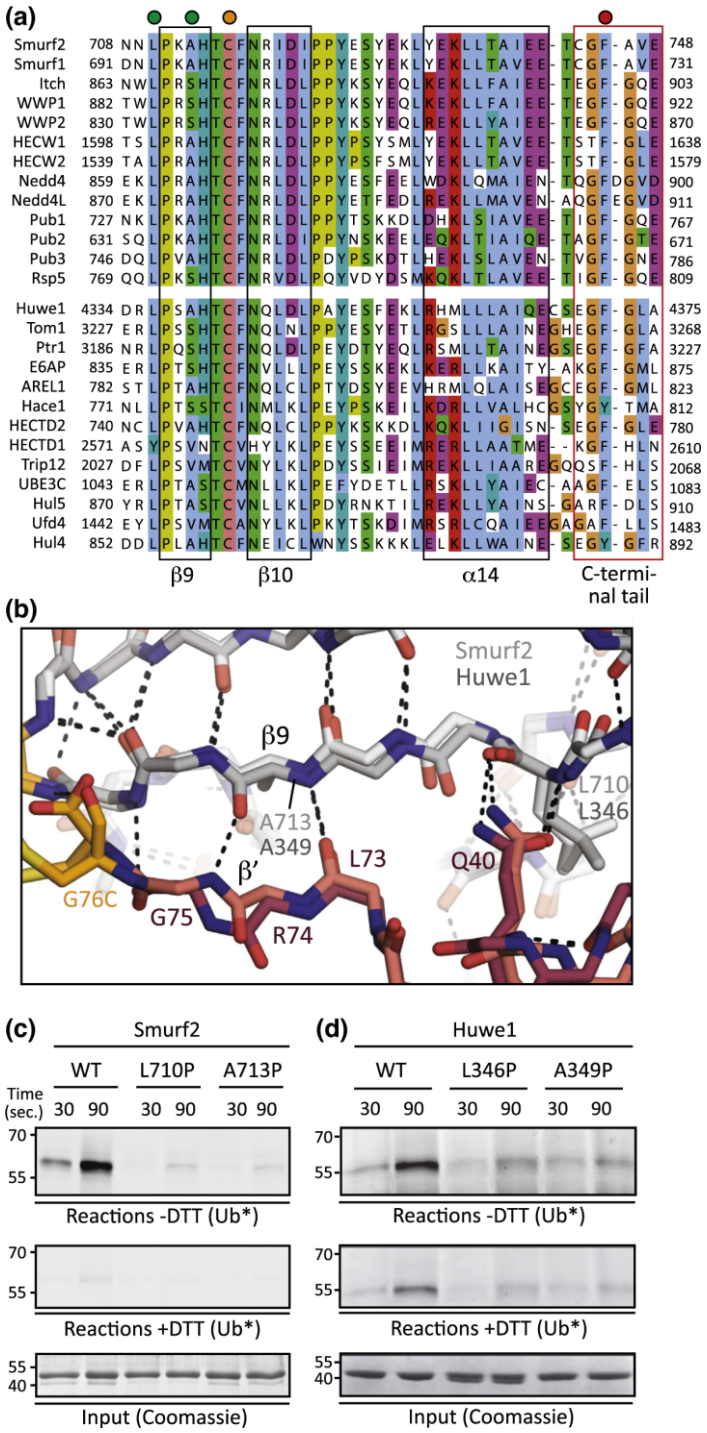


Figure 4. Mutations in the C-lobe~Ub^D interface interfere with thioester formation. (a) Multiple sequence alignment of the C-terminal residues of Nedd4 family (top) and other HECT-type (bottom) Ub ligases. Circles on top of the alignment highlight positions equivalent to L710 (Smurf2)/L346 (Huwe1) and A713 (Smurf2)/A349 (Huwe1) that form hydrogen bonds via their amide nitrogen atoms to the thioester Ub (green), the catalytic Cys (orange) and the –4F residue (red). The four or five (in Nedd4 E3s) C-terminal residues are boxed in red, while secondary structure elements are boxed in black and indicated at the bottom. Residue numbers are indicated to the left and right of the protein sequences. (b) Hydrogen bonding network at the C-lobe~Ub^D interface. The Huwe1 and Smurf2 C-lobe~Ub^D complexes are shown in stick representation with hydrogen bonds indicated by dashed lines. Residues in the C-lobes that were mutated to prolines in functional assays are labeled along with their hydrogen bonding partners in the Ub^D. (c, d) Thioester assays with Pro mutants in Smurf2 (c) and Huwe1 (d). Transthiolation reactions were performed with bacterially expressed WT and mutant HECT domains as indicated. HECT thioester formation was monitored by using fluorescently labeled Ub (Ub) after quenching the reaction at different time points by addition of Laemmli buffer without (top) or with (center) reducing agent. Equivalent levels of HECT domains were confirmed by incubation of the proteins in the absence of E1 and visualized by Coomassie staining (bottom).*

6. The sequence of the C-terminal tail modulates ligation activity

The C-terminal tail of HECT domains plays a crucial role in isopeptide formation and Ub chain specificity[104–106,121]. However, the underlying mechanistic details have remained elusive because these residues were missing in all HECT thioester structures determined to date (Figure 2a, d)[104,105,107]. Interestingly, previous studies revealed distinct roles of the C-terminal residues within the Nedd4 family and among the Nedd4 family and the K48-specific E6AP ligase[104–106].

We therefore examined the effect of the amino acid composition of the C-terminal tail in Huwe1 and Smurf2 on transthiolation and poly-ubiquitination activity. First, we generated HECT domain mutants that were successively truncated from the C-terminus, starting from the "Δ-4F" mutant (Δ745–748 and Δ382–385 mutants in Smurf2 and Huwe1) that abrogates HECT-mediated isopeptide bond formation, but not E2–E3 transthiolation[106]. While all deletion mutants were capable of forming DTT-sensitive HECT~Ub thioesters (Figure S 6a, b), they were defective in poly-ubiquitination (Figure 5a, b) with deletion of the very C-terminal residue (Δ748 in Smurf2 and Δ385 in Huwe1) already being sufficient to provoke this effect. Then, we generated mutants of Smurf2 (E748A and V747A/E748A) and Huwe1 (L384A and G383A/ L384A), where the C-terminal residues were successively replaced with alanines. Of note, the −1 position in the native Huwe1 sequence is already an Ala, while the same holds true for the −3 position relative to the C-terminus in Smurf2 (Figure 4a). Consistent with the C-terminal tail not being essential for thioester formation[104–106,121], mutation of the C-terminal residues to Ala resulted in thioester levels comparable to the WT for both the Smurf2 and Huwe1 HECT domains (Figure S 6c, d). Interestingly, substitution of the very C-terminal residue in Smurf2 (E748A) enhanced poly-ubiquitination activity (Figure 5c), which may explain the slightly weaker band for the E748A thioester in comparison to the WT and V747A/E748A HECT domains (Figure S 6c). In contrast, further Ala substitution of the Smurf2 C-terminus (V747A/E748A) led to reduced auto-ubiquitination compared to the WT HECT domain (Figure 5c). For Huwe1, we observed successively decreasing ubiquitination activity for the L384A and G383A/ L384A mutants (Figure 5d). Thus, in both Smurf2 and Huwe1, Ala substitutions of C-terminal residues did not affect thioester formation and were tolerated for poly-ubiquitination.

Taken together, we conclude that the sequence composition of the C-terminal tail of Smurf2 and Huwe1 has no role in the transthiolation reaction but influences Ub ligation independent of Ub chain specificity.

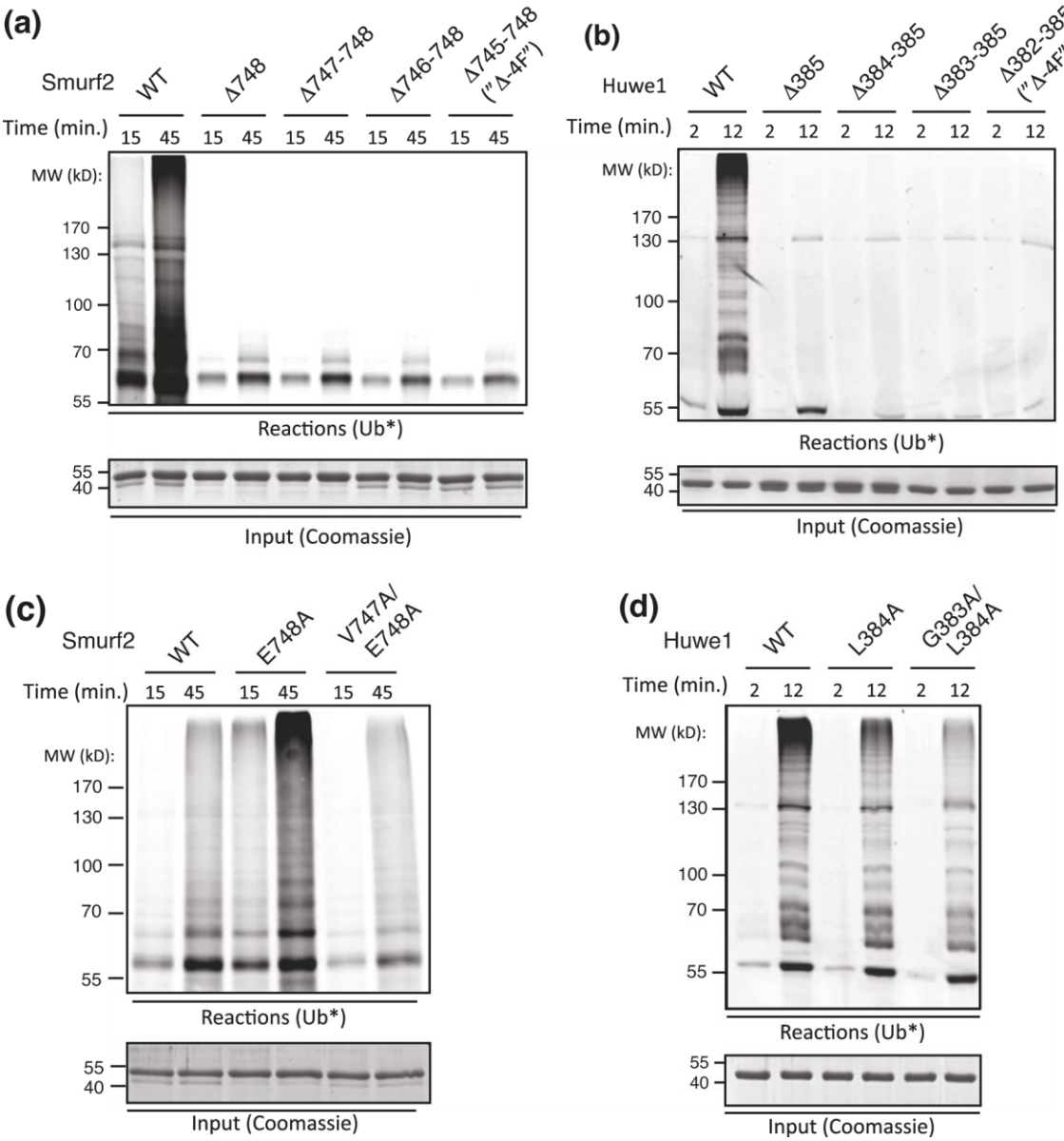


Figure 5. The length and composition of the HECT C-terminal tail is important for isopeptide formation. In vitro ubiquitination assay using the indicated bacterially expressed and purified Smurf2 (a) or Huwe1 (b) WT HECT domains or the indicated HECT truncation constructs. Reactions were stopped after the indicated time points and loaded on an SDS-PAGE gel to analyze HECT auto-ubiquitination activity with fluorescently labeled Ub (Ub) (top panels). Coomassie staining of the proteins incubated in the absence of E1 enzyme (bottom panel). (c, d) As panels a and b, but for Ala substitutions of C-terminal residues as indicated.*

5. Discussion

HECT-type ligases are key regulators of various signal transduction pathways that play a role in cellular homeostasis and embryonic development but adversely also in carcinogenesis[158]. Although significant progress has been made over the past decade in elucidating the mechanisms underlying HECT-mediated ubiquitination in Nedd4-family E3s, many general questions remain unresolved such as whether the mechanisms underlying Ub transfer are conserved within the entire HECT family, which residues form the catalytic center during isopeptide formation, and what determines Ub chain specificity.

Here, we find that the mechanisms of thioester formation are remarkably conserved between Smurf2 and Huwe1, although these enzymes belong to distinct families and possess different Ub chain specificities. Our functional and structural analyses suggest that the mechanisms underlying the first catalytic step in HECT-mediated Ub transfer, that is, E2–E3 transthiolation, is conserved among HECT domains and invariably involves β-sheet augmentation and a conserved network of non-covalent interactions between the C-lobe and the donor Ub that primes the enzymes for Ub ligation. Consistently, interfering with these C-lobe~Ub interactions abolished thioester formation in both Smurf2 and Huwe1. Most importantly, we identify a structural rearrangement of the C-terminal residues of the Smurf2 and Huwe1 C-lobes that occurs upon thioester formation. This conformational change in the C-terminal tail is not required for transthiolation, but rather participates in priming the enzyme for the subsequent catalytic step, that is, Ub ligation, as previously suggested[105].

Together with previous studies, our functional assays show that the impact of the individual residues in the C-terminal tail on ubiquitination activity seems to vary among HECT domains. While deletion of C-terminal residues abrogated Ub ligation in K63-specific Nedd4-family E3s and in Huwe1 (Figure 5a, b)[104–106], the K48-specific E6AP is still active with a truncated C-terminus[106]. Mutation of the C-terminal residues was tolerated in Huwe1 (Figure 5d), Rsp5 and E6AP[104,106] and even activated ubiquitination activity in Smurf2 (Figure 5c), but abolished isopeptide formation in Nedd4[104–106]. Of note, the C-terminal tails in Nedd4 and Nedd4L contain an additional residue in comparison to other HECT domains (Figure 4a). The requirement of an acidic residue at the C-terminal position may thus be more stringent in these enzymes[105] than in other HECT domains. In any case, these results suggest that the sequence composition of the C-terminal tails may play a rather individual role in HECT ubiquitination activity and ultimately Ub chain specificity.

In sum, we conclude that the mechanisms underlying thioester formation appear to be highly conserved in HECT domains irrespective of their Ub chain specificities. In contrast, Ub ligation and chain specificity seem to be determined at least partially by the individual C-terminal sequences of the HECT domains. It is tempting to speculate that the individual C-terminal sequence composition of the HECT domains fine-tunes ubiquitination activity and Ub chain specificity by creating particular binding surfaces with the preferred Lys in the substrate or in the acceptor Ub and/or the N-lobe. In support of this, mutations in the acceptor Ub differentially affected ubiquitination efficiency in E6AP and UBE3C, although both enzymes prioritize K48-Ub linkage[132]. However, unraveling the individual roles of the C-terminal tails in HECT domains on a structural level is highly challenging and future studies may hopefully shed light on this still unresolved, yet important aspect of HECT-mediated Ub ligation.

6. *Materials and Methods (Chapter A)*

1. Constructs and reagents

The nucleotide sequence of the human Huwe1 ligase (NP_113584.3) corresponding to aa 3993– 4375 (hereafter referred to as aa 3–385 for clarity) was amplified by PCR from a cDNA vector purchased from Imagenes and ligated into a pETM-11 vector (EMBL Heidelberg) to express a Huwe1 HECT domain with an N-terminal His6-tag followed by a TEV protease cleavage site. To generate a Huwe1 HECT domain where all Cys were replaced except for the catalytic Cys (hereafter referred to as "C1"), we performed the following substitutions: C107I, C134A, C192N, C219A and C375N. The Huwe1 C-lobe C1 construct was produced by amplification of the nucleotide sequence of aa 265–385 of the C1 HECT domain by PCR and ligation into a pETM-11 vector. The L346A and A349P mutations were introduced for NMR studies. For thioester and ubiquitination assays, the WT Huwe1 HECT domain DNA fragment was ligated into a modified pETM-30 vector that encodes an HAtag directly C-terminal of the TEV cleavage site. The following Huwe1 HECT domain variants were produced by using this vector as a template: Δ385, Δ384– 385, Δ383–385, Δ382–385 ("Δ-4F"), L384A, G383A/ L384A, L346P, A349P.

For Smurf2, a HECT domain where all Cys were replaced except for the catalytic Cys was generated using the human Smurf2 HECT domain sequence in a pProExHTb vector[121] as a template and included the following substitutions: C393I, C550S, C627G, C646A, C706N and C743S. The Smurf2 C-lobe C1 construct was produced by amplification of the nucleotide sequence of aa 610–748 of the C1 HECT domain by PCR and ligation into a pETM-11 vector. The L710A and A713P mutations were introduced for NMR studies. For thioester and ubiquitination assays, the WT Smurf2 HECT domain DNA fragment was inserted into a modified pETM-11 vector that encodes an HA-tag directly C-terminal of the TEV cleavage site via restriction-free cloning[183]. The following Smurf2 HECT domain variants were produced in this vector: Δ748, Δ747–748, Δ746–748, Δ745–748 ("Δ-4F"), E738A, V747A/ E748A, L710P, A713P.

For disulfide formation, the C-terminal Gly was replaced with a Cys using human Ub in a pETM-11 vector[121] as template. To produce fluorescein-labeled Ub, we first generated a modified pETM-10 vector by inserting the coding sequence of *Saccharomyces cerevisiae* Smt3 (a.k.a. Sumo) downstream of the His6-tag. Next, we cloned the nucleotide sequence of human Ub into the pETM-10-Sumo vector and introduced a Cys directly N-terminal of the M1 residue of the WT human Ub sequence (Ub C-1). For single Lys Ub variants, first all Lys residues were mutated to Arg (K0 Ub) in the Ub C-1 construct and single Lys reintroduced into the K0 clone at the sites of the native sequence (K6-, K11-, K27-, K29-, K33-, K48- and K63-only Ub). A vector for producing Sumo protease (Ulp1) was a gift from M. Gröttrup (University of Konstanz, Germany). WT Ub, UbcH7 (E2), Uba1 (E1) and GST-tagged HECT domains of Nedd4, Nedd4L, Smurf2 and Huwe1 were used as described[102,125,175]. All mutants were generated using the QuikChange mutagenesis protocol. All constructs were verified by sequencing.

2. Protein purification

All recombinant proteins were expressed in *Escherichia coli* BL21-CodonPlus (DE3) RIL cells (Stratagene). Cultures were grown at 37 °C in LB (unlabeled proteins) or in at least 90% D_2O (partial deuteration, ^{15}N-labeling)- or H_2O (^{15}N- or ^{13}C,^{15}N-labeling)-based M9 minimal medium. Cells were induced with 1 mM IPTG at 18 °C (unlabeled proteins) or 25 °C (labeled proteins) and harvested by centrifugation after ~16 h of induction. Cells were lysed by sonication on ice and cell debris removed by centrifugation. Proteins were purified by Ni-affinity chromatography. The affinity tag was cleaved with TEV protease where necessary or with Ulp1 in the case of Ub C-1 and proteins collected from the flow-through of a second Ni-affinity chromatography step. If necessary, size exclusion chromatography was performed to improve protein purity. Purified proteins were dialyzed into NMR buffer [20 mM sodium phosphate (pH 6.4), 150 mM NaCl, 1 mM DTT] or into reaction buffer [50 mM Tris (pH 7.5), 150 mM NaCl, 1 mM DTT]. Prior to disulfide formation, the DTT-free buffer was degassed and purged with argon.

3. Disulfide synthesis

To activate Ub for disulfide formation, His-tagged Ub G76C was immobilized on Ni-NTA beads (Qiagen). The beads were first washed with disulfide buffer A [20 mM sodium phosphate (pH 7.8), 150 mM NaCl, 1 mM DTT] to fully reduce all cysteines and then flushed with disulfide buffer B [20 mM sodium phosphate (pH 7.8), 150 mM NaCl, 50 mM Ellman's reagent (5,5'-dithiobis-(2-nitrobenzoicacid) or DTNB]. The resulting chemically activated thionitrobenzoic Ub was washed with disulfide buffer C [20 mM sodium phosphate (pH 7.8), 150 mM NaCl; degassed and argon purged] until all unreacted DTNB was removed.

Untagged HECT C-lobes where all Cys except for the catalytic Cys were replaced were buffer-exchanged to disulfide buffer C and passed over the Ni NTA beads containing the His-tagged thionitrobenzoic Ub to form C-lobe~Ub^D disulfides. The beads were first washed with disulfide buffer D [20 mM sodium phosphate (pH 6.4), 150 mM NaCl; degassed and argon purged] and then eluted with increasing imidazole concentrations in disulfide buffer D. The C-lobe~Ub^D disulfides were further purified by size exclusion chromatography.

4. NMR spectroscopy

NMR experiments were acquired on 600- and 800-MHz Bruker Avance III spectrometers operating with a room temperature probe head. Chemical shift perturbation studies for the C-lobes~Ub^D disulfides were performed by comparing the $^{1}H,^{15}N$-TROSY spectra of ^{15}N-labeled C1 C-lobes (reference spectra) with those in complex with unlabeled Ub G76C or vice versa. Non-covalent Ub binding to the Huwe1 HECT domain was examined by collecting 2D $^{1}H,^{15}N$-TROSY spectra of a 200 µM ^{15}N-labeled Huwe1 HECT domain sample in the absence or presence of increasing amounts of unlabeled Ub. All CSP spectra were recorded at 800 MHz. All NMR spectra were recorded in NMR buffer [20 mM sodium phosphate (pH 6.4), 150 mM NaCl in the absence (disulfides) or presence of 1 mM DTT (free proteins)] at 30 °C for Huwe1 and 25 °C for Smurf2. For backbone assignment of the Huwe1 and Smurf2 C1 C-lobes, 3D HNCACB, HN(CO)CACB and (H) CC(CO)NH-TOCSY spectra were recorded at 600 MHz on a 1 mM Huwe1 and a 2.1 mM Smurf2 sample. All NMR data were processed with the program suite NMRPipe/NMRDraw[184], analyzed with XEASY[185] for backbone assignment or Sparky[186] for CSP experiments and visualized with NMRView (OneMoonScientific).

Chemical shift perturbations were quantified as relative signal intensities (I_{dis}/I_{ref}) of the free protein (I_{ref}) and the disulfide complex (I_{dis}) at the resonance frequencies of the free protein to account for both peak shift and line broadening effects, where I_{dis} refers to the signal intensity of the disulfide at the resonance frequencies of the free protein and I_{ref} is the signal intensity of the free protein at the resonance frequencies of the free protein.

5. Crystallization, data collection and structure determination

The C-lobe~Ub^D complexes were crystallized in 0.8 M ZnSO4, 0.1 M sodium acetate (pH 4.0) for Huwe1 and in 1.36 M sodium citrate (pH 6.5), 15% (v/v) glycerol for Smurf2. Diffraction data were collected at 100 K using a wavelength of 1 Å and a PILATUS 6 M-F detector at the beam-line PXII of the Swiss Light Source (PSI, Villigen, Switzerland). Data were processed using XDS[187], and the structure solved by molecular replacement using Phaser[188], iterative manual modeling with Coot[189] and refinement with Phenix[190]. All structural representations were prepared with PyMOL (The PyMOL Molecular Graphics System, Version 1.7.6.6 Schroedinger, LLC).

6. AQUA assays

For Ub AQUA analysis, reaction mixtures (20 µL), containing purified enzymes (E1 50 nM, UbcH7 2.5 µM, Huwe1 HECT domain 3 µM) and 15 µM of Ub in ubiquitination buffer, were incubated at 37 °C for 2 h and stopped by addition of 4× Laemmli buffer with reducing agent (100 mM DTT).

Proteins were resolved by SDS-PAGE on a gradient gel (4%–12% TGX Precast Gel, Biorad) and stained with colloidal Coomassie. Gel bands corresponding to Ub-modified HECT domains were digested with trypsin. Briefly, samples were subjected to reduction in 10 mM DTT for 1 h at 56 °C. Digestion was

carried out overnight saturating the gel with 12.5 ng/µL sequencing grade modified trypsin (Promega) in 50 mM ammonium bicarbonate. Peptide mixtures were acidified with tri-fluoro acetic acid (final concentration 3%), extracted from gel slices with 30% acetonitrile/3% tri-fluoro acetic acid and concentrated to 100 µL in a vacuum concentrator. Peptides were loaded onto homemade C18-stage tips, dried and dissolved in 5% formic acid prior to analysis on an Agilent 1100 HPLC system (Agilent Technologies Inc.) coupled to Ultra LTQ-FT mass spectrometer (Thermo Fisher Scientific).

For AQUA analysis, samples were directly dissolved with a solution of reference peptides for Lys6, Lys11, Lys27, Lys29, Lys33, Lys48 and Lys63 poly-Ub branched chains (Cell Signaling Technology) in 5% formic acid. Samples were analyzed using a range of reference peptides concentrations spanning from 1 pmol to 100 fmol, injecting three technical replicates per concentration. All spectra were acquired in data-dependent mode. Ion chromatograms for reference and sample peptide pair precursor ions were manually extracted with Xcalibur v1.4 software (ThermoElectron). Chromatographic co-elution of reference and endogenous peptide pairs and accurate peak integration were manually confirmed.

7. Poly-Ub pull-down assays

For pull-down experiments, 2 µM of GST proteins immobilized onto glutathione beads was incubated with 250 ng of K63 or K48-linked Ub chains (Boston Biochem) for 2 h at 4 °C in YY buffer [50 mM Hepes (pH 7.5), 150 mM NaCl, 1 mM EDTA, 1 mM EGTA, 10% glycerol, 1% triton X-100]. After four washes with YY buffer, specifically bound proteins were resolved on Tricine-PAGE (11%) and detected by immunoblotting using the α-Ub P4D1 antibody (Santa Cruz Biotechnology). Coomassie staining of the membranes after transfer was used to show the input GST-fusion proteins.

8. Fluorescein labeling of Ub

Purified WT or single-Lys Ub C-1 was incubated with a 4-fold stoichiometric excess of 6-iodoacetamidofluorescein for 3 h at RT in 50 mM Tris (pH 7.8) and 150 mM NaCl buffer and subsequently applied to size exclusion chromatography to remove unreacted fluorescein and to exchange the buffer to 50 mM Tris (pH 7.5), 150 mM NaCl and 1 mM DTT.

9. Ubiquitination and thioester assays

In vitro auto-ubiquitination and thioester assays were performed in ubiquitination buffer [25 mM Tris (pH 7.5), 5 mM MgCl2, 100 mM NaCl 0.2 mM DTT, 2.5 mM ATP] at 30 °C for Huwe1 and at 25 °C for Smurf2. The reactions contained 0.5 (auto-ubiquitination) or 0.3 (thioester assay) µM E1 (Ube1), 10 µM E2 (UbcH7), 3 µM HECT domains and 60 µM fluorescein labeled Ub (Ub*). The reactions were stopped at the indicated time points with 3× Laemmli buffer containing 6 M urea with or without 0.5 M DTT (thioester detection). The samples were loaded on 8% and 16% acrylamide gels for SDS-PAGE and fluorescent Ub detected using an Amersham™ 600 or Typhoon 5 Imager (GE Healthcare Life Sciences).

10. Accession numbers

Coordinates of the crystal structures have been deposited in the PDB under accession codes 6FX4 and 6FYH for the Smurf2 and Huwe1 C-lobe~Ub G76C complexes, respectively. ^{1}H, ^{15}N, ^{13}C backbone resonance assignments of the Smurf2 and Huwe1 C-lobes and G76C Ub have been deposited in the BMRB under accession codes 27508, 27511 and 27512, respectively.

7. Acknowledgments

We are grateful to Marcus Hartmann, Ancilla Neu, and Iris Holdermann (MPI for Developmental Biology, Tübingen, Germany) for help with data collection and determination of the crystal structures and to Paolo Soffientini and the Proteomics/ MS Technological Development Unit at Cogentech/IFOM for the AQUA analysis. We also thank Samira Anders and Mira Schütz-Stoffregen (MPI for Developmental Biology, Tübingen, Germany) for cloning and purification of proteins used in this work. S.W. acknowledges support from the Max Planck Society and a Marie Curie Reintegration grant by the

Marie Skłodowska–Curie Actions Research Fellowship Programme of the European Commission (PIRG04-GA-2008-239418 UBLIGATION). S.P. acknowledges support from the Associazione Italiana per la Ricerca sul Cancro (IG11627).

8. Abbreviations

TROSY, transverse relaxation-enhanced spectroscopy;
HECT, homologous to the E6-AP carboxyl terminus;
Smurf2, SMAD-specific E3 ubiquitin protein ligase 2;
Huwe1, HECT, UBA And WWE domain containing 1;
Nedd4, neural precursor cell-expressed developmentally downregulated gene 4.

9. *Appendix A. Supplementary data*

1. Supplementary Figures

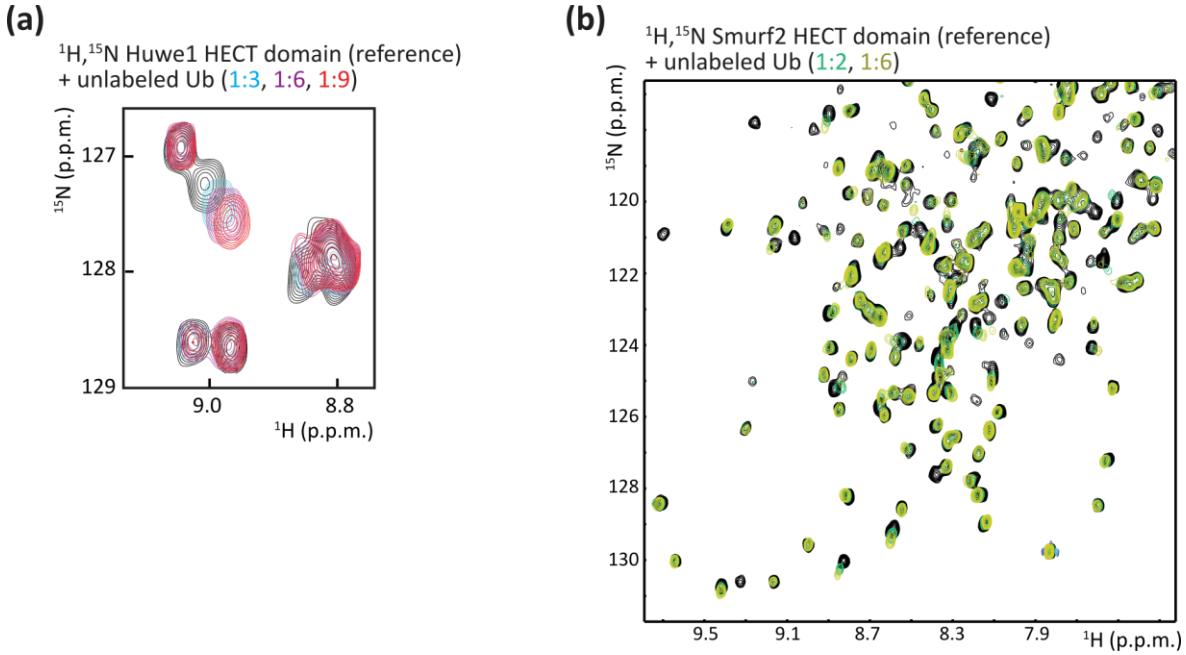


Figure S 1. The Huwe1 HECT domain barely interacts with Ub in a non-covalent manner. (a) Zoom of the area showing the largest chemical shift perturbations in the 1H,15N-TROSY NMR spectra of the Huwe1 HECT domain in the absence (black) and presence of increasing stoichiometric amounts of unlabeled monomeric Ub as indicated. (b) Overlay of a representative region of the ^{1}H,^{15}N-TROSY NMR spectra of the Smurf2 HECT domain in the absence (black) and presence of increasing stoichiometric amounts of unlabeled monomeric Ub as indicated. In contrast to the Huwe1 HECT domain (Figure 1c and Figure S 1a), numerous peaks are broadened beyond detection or show chemical shift changes already upon addition of a two-fold stoichiometric excess of monomeric Ub to the Smurf2 HECT domain.

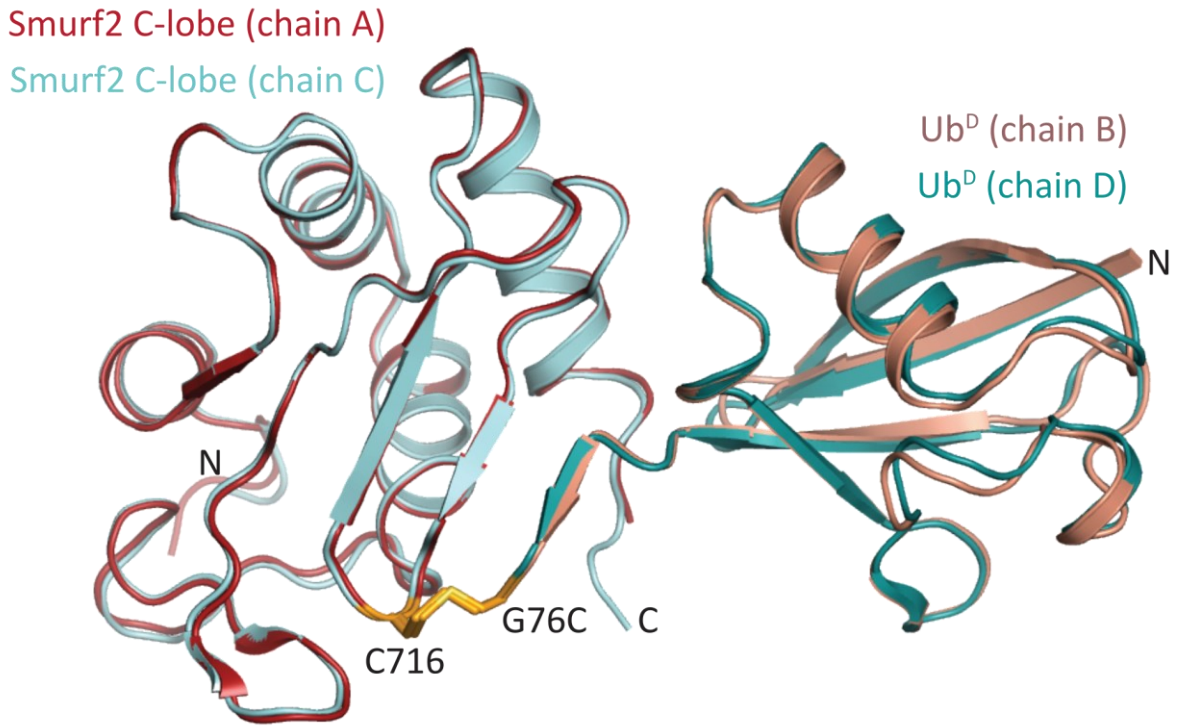


Figure S 2. Structure overlay of the two copies of Smurf2 C-lobe~Ub^{D} complex in the asymmetric unit. The two Smurf2 C-lobe~UbD complexes in the asymmetric unit are shown as ribbon representations. Color code of the protomer chains as indicated. The disulfide bridges (Smurf2 C716 ~ Ub G76C) are labeled and highlighted as stick representations and shown in orange.

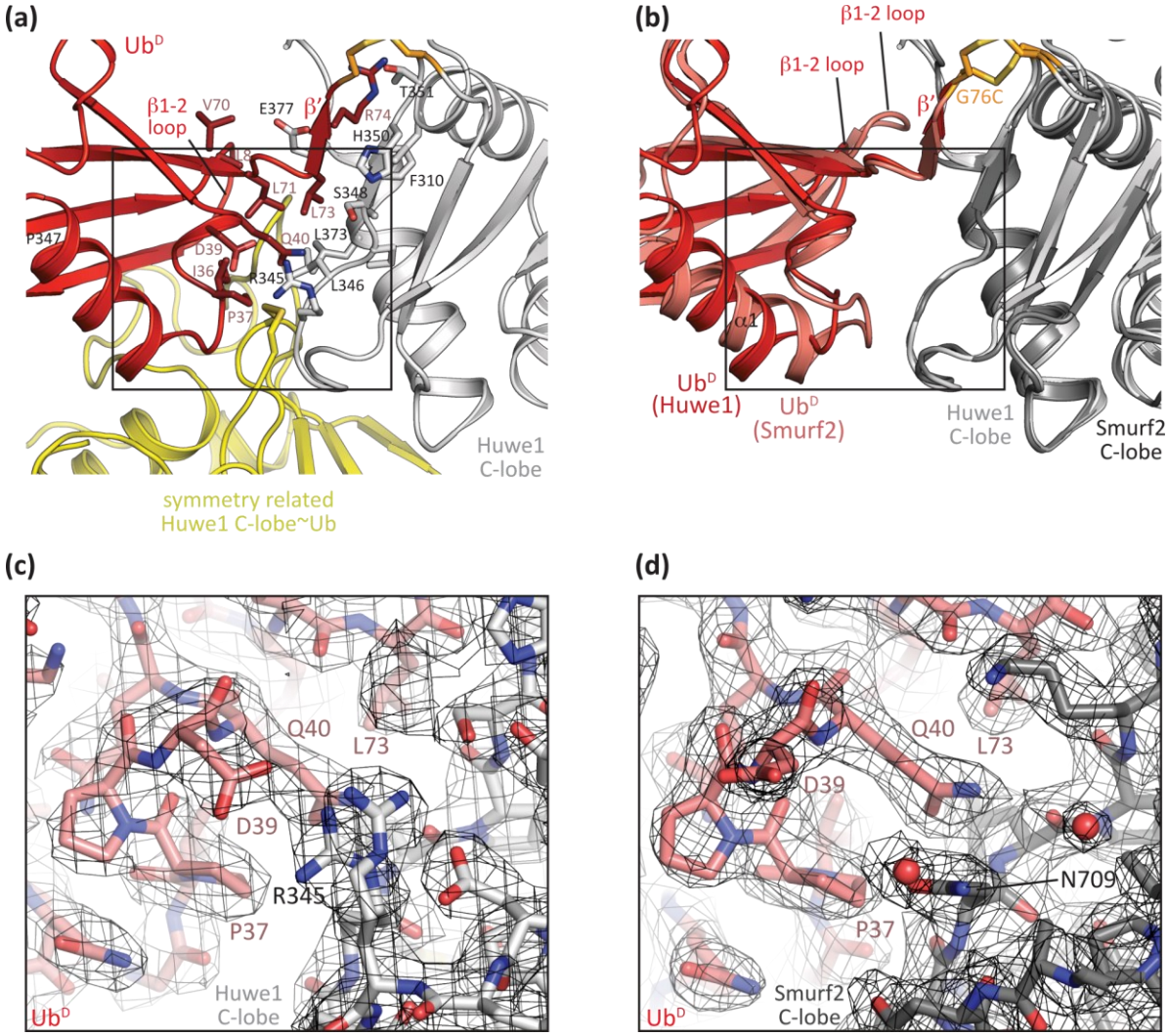


Figure S 3. Non-covalent contacts in the C-lobe:Ub interface are largely conserved in the crystal structures of the Huwe1 and Smurf2 C-lobe~Ub^D complexes. (a) Ribbon representation of the Huwe1 C-lobe~Ub disulfide shown in grey (C-lobe) and red (Ub) and the symmetry related complex (yellow) that inserts into the C-lobe:Ub^D interface. Side-chains at the C-lobe:Ub interface are labeled and shown in stick representation. The black box indicates of the area of the structure enlarged in (c). (b) Overlay of the Huwe1 (light grey) and Smurf2 (dark grey) C-lobe~Ub^D structures in the same orientation as in (a) highlighting differences in the orientation of the donor Ub in the two complexes. The black box indicates of the area of the structure enlarged in (c,d) Simulated-annealing 2Fo-Fc omit maps of side-chain contacts at the interface of the Huwe1 (c) and Smurf2 (d) C-lobe~Ub^D complexes contoured at a level of 1σ. Despite the proximity of a symmetry related complex, essential side-chain interactions in particular of P37 and Q40 (labeled and shown as sticks) in the C-lobe:Ub^D interface remain intact in the Huwe1 disulfide as compared to Smurf2 (d). Ub is shown in salmon, while the C-lobes are shown in light grey (Huwe1) or dark grey (Smurf2).

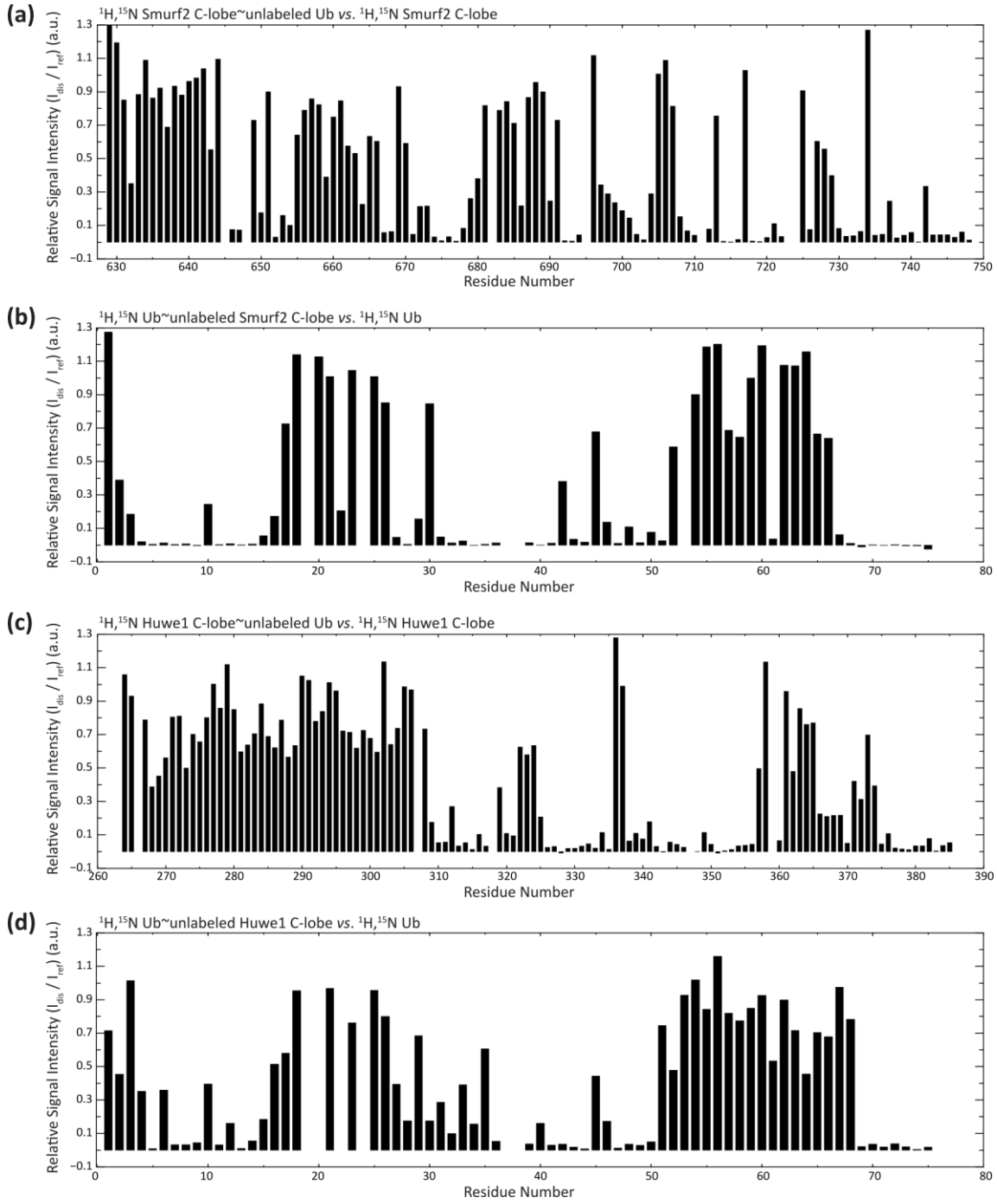


Figure S 4. Chemical shift perturbations upon disulfide formation. Chemical shift perturbations were evaluated as relative signal intensities, i.e. the ratio of signal intensity in the disulfide (I_{dis}) and the free protein (I_{ref}) both at the resonance frequencies of the free protein and plotted in arbitrary units (a.u.) versus the respective residue numbers in the Smurf2 C-lobe (a), Ub in the Smurf2 C-lobe~Ub disulfide (b), the Huwe1 C-lobe (c) and Ub in the Huwe1 C-lobe~Ub disulfide (d).

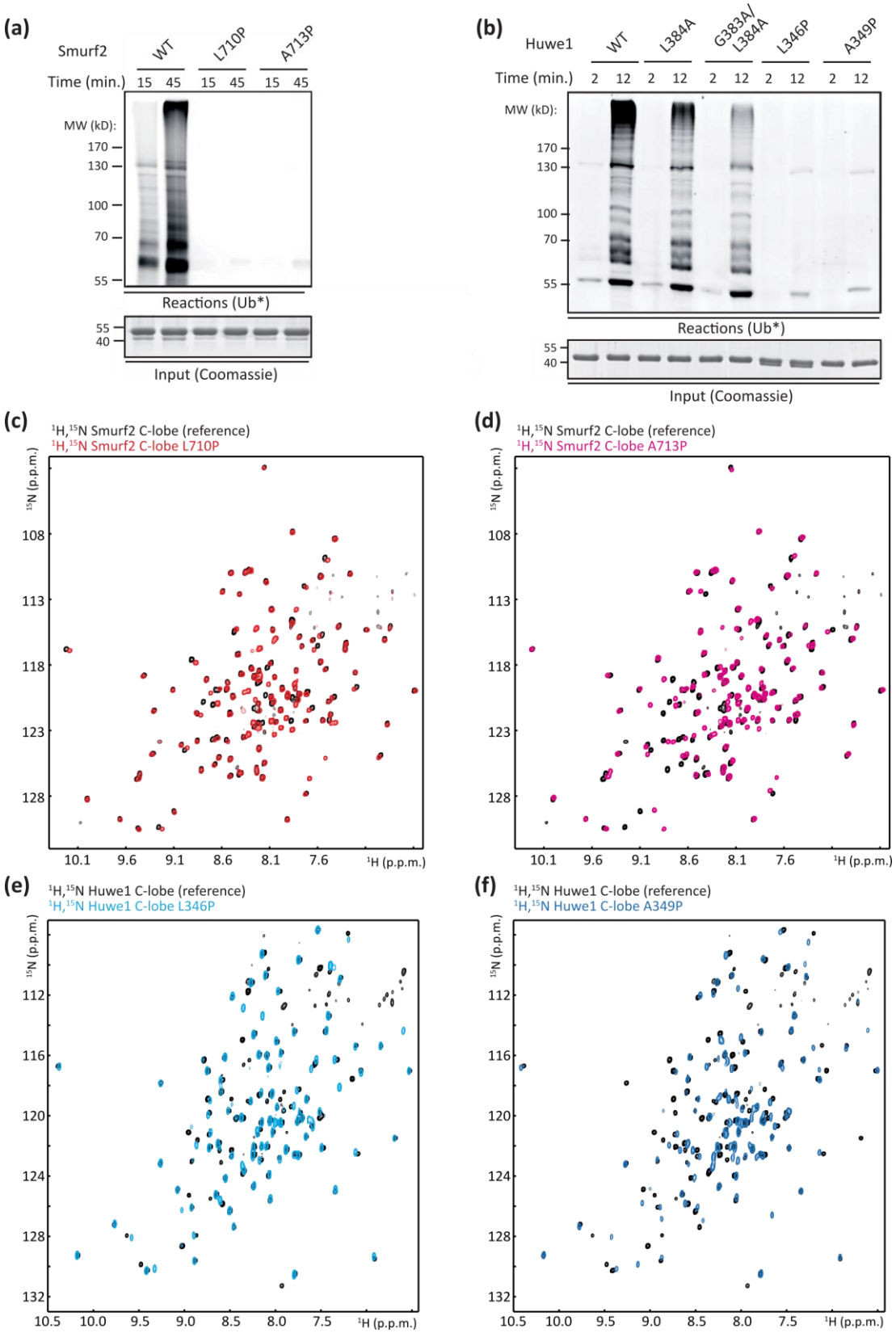


Figure S 5. Functional and structural characterization of the Smurf2 and Huwe1 Pro mutants. (a,b) In vitro ubiquitination assay using the indicated bacterially expressed and purified Smurf2 (a) or Huwe1 (b) WT or mutant HECT domains as indicated. Reactions were stopped after the indicated time points and loaded on an SDS-PAGE gel to analyze HECT auto-ubiquitination activity with fluorescently labeled Ub (Ub) (top panels). Coomassie staining of the reactions in the absence of E1 enzyme (bottom panel). (c-f) Overlay of the $^{1}H,^{15}N$-TROSY spectra of WT proteins and the Smurf2 L710P (c) and A713P (d) and Huwe1 L346P (e) and A349P (f) mutants. The well-dispersed spectra confirm the structural integrity of the mutants.*

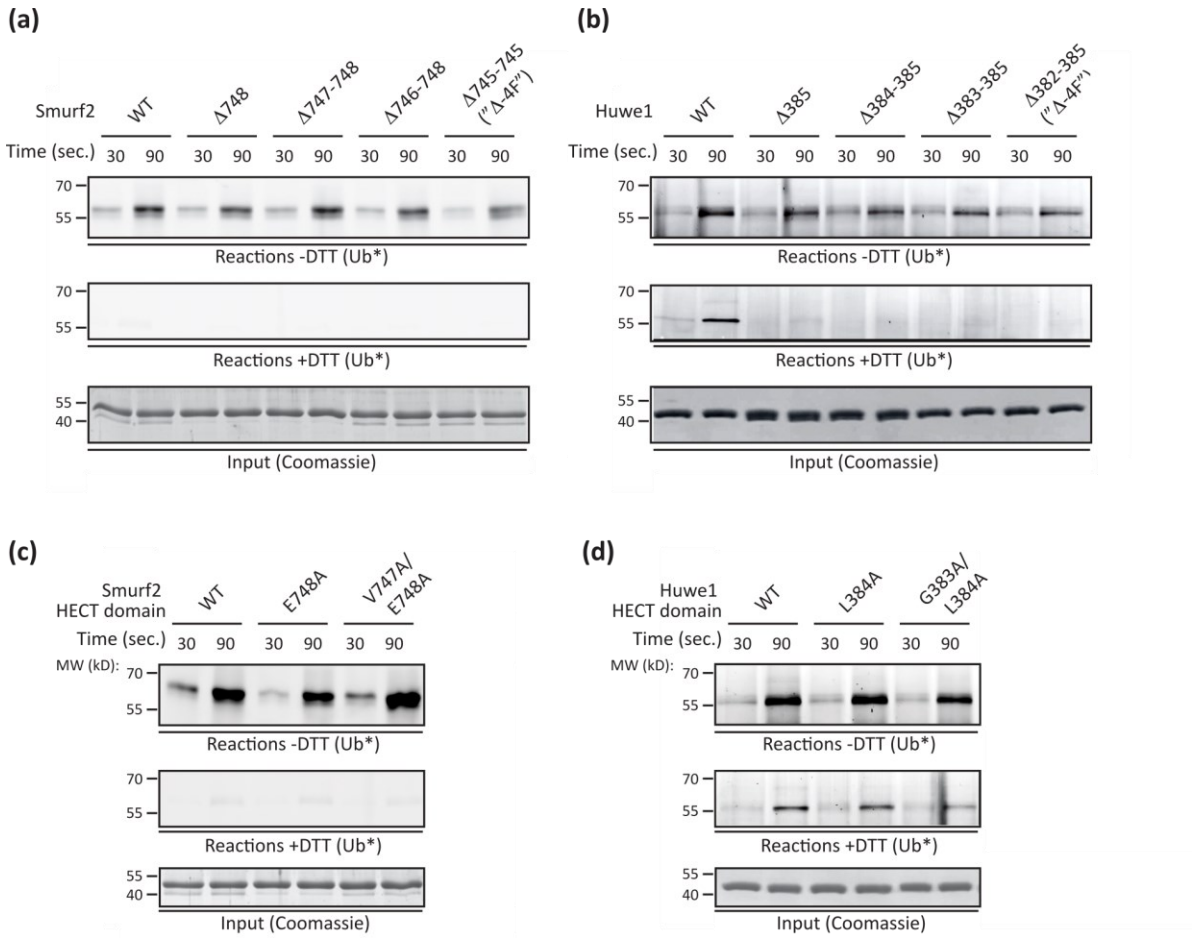


Figure S 6. Truncation or mutation of the HECT C-terminal tail does not impair thioester formation. Thioester assays with the indicated bacterially expressed and purified Smurf2 (a,c) or Huwe1 (b,d) WT or mutant HECT domains as indicated. Transthiolation reactions were performed with bacterially expressed WT and mutant HECT domains as indicated. HECT thioester formation was monitored by using fluorescently labeled Ub (Ub) after quenching the reaction at different time points by addition of Laemmli buffer without (top) or with (center) reducing agent. Equivalent levels of HECT domains were confirmed by incubation of the proteins in the absence of E1 and visualized by Coomassie staining (bottom).*

2. Supplementary Tables

Table S 1. Statistics of x-ray data collection and model refinement for the Smurf2 C-lobe~Ub G76C complex[1]

Data collection	
Resolution (Å)	19.57 – 2.5 (2.589 – 2.5)[2]
Completeness (%)	98.28 (97.76)
No. of unique reflections	14674 (1484)
Redundancy	3.5 (3.6)
CC1/2* (%)	0.98 (0.66)
I/σI	6.31 (1.91)
Wilson B-Factor (Å^2)	28.93
FreeR (% of reflections)	5
Crystal properties	
Space group	P 1
Unit cell dimensions:	
***a, b, c* (Å)**	38.63 45.95 70.46
***α, β, γ* (°)**	108.921 94.43 106.753
Solvent content (%)	42.92
Refinement	
Resolution (Å)	19.57 – 2.5 (2.589 – 2.5)[2]
R_{work} (%) / R_{free} (%)	18.73 / 23.89
R.m.s.d. bond lengths (Å)	0.002
R.m.s.d. bond angles (°)	0.57
B-Factor (Å^2) (Overall)	39.18
Protein / Ligands / Solvent	39.37 / 40.48 / 31.99
No. of atoms (Total)	3204
Protein / Ligands / Solvent	3108 / 12 / 84
Ramachandran (%) (Favored / allowed / outliers)	99.20 / 0.80 / 0

[1]Number of crystals equals one.

[2]Values in parentheses are for highest-resolution shell.

Table S 2. Statistics of x-ray data collection and model refinement for the Huwe1 HECT C-lobe~Ub G76C complex[1].

Data collection	
Resolution (Å)	41.12 - 2.906 (3.01 - 2.906)[2]
Completeness (%)	99.64 (97.28)
No. of unique reflections	114772 (10657)
Redundancy	16.7 (16.5)
CC1/2* (%)	0.999 (0.857)
I/σI	21.43 (3.09)
Wilson B-Factor (Å^2)	67.16
FreeR (% of reflections)	5
Crystal properties	
Space group	P 43 21 2
Unit cell dimensions:	
***a, b, c* (Å)**	82.234 82.234 85.852
***α, β, γ* (°)**	90 90 90
Solvent content (%)	59.06
Refinement	
Resolution (Å)	41.12 - 2.906 (3.01 - 2.906)[2]
R_{work} (%) / R_{free} (%)	22.96 / 26.61
R.m.s.d. bond lengths (Å)	0.002
R.m.s.d. bond angles (°)	0.46
B-Factor (Å^2) (Overall)	60.65
Protein / Ligands / Solvent	60.51 / 80.45 / 40.53
No. of atoms (Total)	1522
Protein / Ligands / solvent	1501 / 16 / 5
Ramachandran (%) (Favored / allowed / outliers)	94.12 / 4.81 / 1.07

[1]Number of crystals equals one.

[2]Values in parentheses are for highest-resolution shell.

B. Huwe1 activity is autoinhibited by a tight Interface between its C- and N- terminal lobes

1. Introduction

Most knowledge of HECT autoinhibition mechanisms comes from studies involving the Nedd4-family of HECT-domain E3s, while not much is known about other E3 ligases. Recently, insights into an alternative mechanism used by Huwe1, a HECT-domain E3 that is not part of the Nedd4-family, have been published. Sander et al. showed that a construct of the Huwe1 HECT domain including an additional α-helix ("pointer helix") immediately upstream of the HECT domain crystallized as a dimer and also formed a dimer in solution, albeit rather weak[113].

The authors also investigated the effect of HECT domain dimerization on activity and found that it indeed lowered autoubiquitination as well as substrate ubiquitination.

In this chapter, I will report a novel auto-inhibitory mechanism that has so far not been described for any other HECT domain, and which the Huwe1 HECT domain uses for autoinhibition.

2. Huwe1 HECT shows characteristic CSPs going towards a state where the C- and N-lobes are fully dissociated from each other.

The HECT domain consists of two lobes which, in many HECT domains, can successfully fold independently of each other (e.g. C-lobe structures in Chapter A). This is also the case for the lobes of the Huwe1 HECT domain, which allowed us to not only record NMR spectra of the full, bilobal HECT domain, but also spectra of its N- and C-lobes individually.

Since the naturally formed thioester bond between the C-terminal glycine (G76) of ubiquitin and the catalytic cysteine on the HECT domain is not stable in the conditions used for the measurements, we mimicked it with a disulfide link[105,180,191–194]. For this, we conjugated a ubiquitin modified with a C-terminal cysteine (ubiquitin^{G76C}) to a Huwe1 HECT domain that was cysteine free except for the catalytic cysteine (mutations outlined in Chapter A), to obtain a stable analogue of a HECT domain loaded with a donor-ubiquitin (Ubiquitin^{D}, Ub^{D}) on its catalytic cysteine. For comparison, NMR spectra of all HECT- and lobe-constructs were recorded in the same cysteine reduced background. To validate the observations gathered from the NMR experiments in a more physiological system, we used wild type HECT domain constructs when performing activity assays.

Of note, residue numbering of Huwe1 used in this work is relative to the start of the HECT domain: therefore, residue 1 of the HECT domain is residue 3990 of the *Homo Sapiens* Huwe1 full-length protein in the reference isoform 1 (UniProtKB: Q7Z6Z7-1).

Comparing spectra of the wild-type HECT domain, a catalytic cysteine mutant (C352K), a donor-ubiquitin loaded HECT domain, and the C-lobe in isolation (residues 266-385) we observed a set of residues which show chemical shift perturbations (CSPs) moving along a vector between the shifts corresponding to the wild-type HECT and the isolated C-lobe (Figure 6). A chemical shift perturbation in general indicates a change in the chemical environment around the residue. A chemical shift change along a vector between two endpoints is the result of a fast chemical exchange process between two states where the chemical shift represents the population-weighted average of the chemical shifts of the two states.

The amino acids displaying the highest, linearly progressing CSPs are isoleucines 270, 331 and 268. Interestingly, these residues are located in the C-lobe opposite of the C352K mutation and the donor-

ubiquitin binding surface and do not reside directly at the interface with the N-lobe, indicating a global rather than a local structural change in the Huwe1 HECT domain upon donor ubiquitin binding.

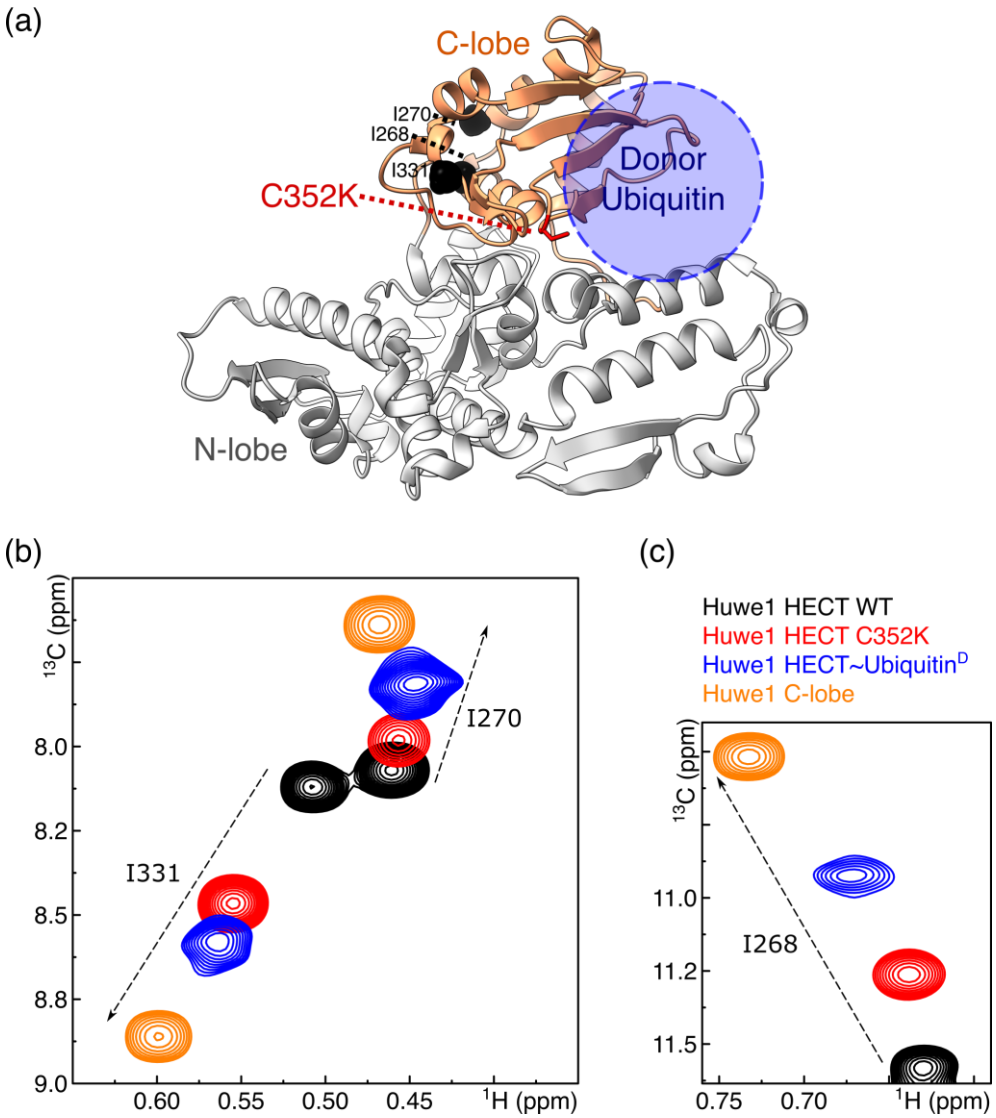


Figure 6. (a) Huwe1 HECT structure (PDB ID: 5LP8) with points of interest annotated. C-lobe (orange), catalytic cysteine (red), donor-ubiquitin position indicated as a blue circle (attached via thioester from its C-terminal glycine to the catalytic cysteine of Huwe1), sensor residues 13C labelled methyl groups (black). (b/c) SOFAST IM-13C-methyl labelled spectral changes (at 283K) of the residues depicted in (a), with Huwe1 HECT-domain wild-type used as a reference (black): the C352K mutant (red), the HECT~donor-ubiquitin thioester mimic (blue) and the isolated C-lobe (orange) spectra show CSPs moving in a linear fashion (full view of spectra overlay in Figure S 1).

In agreement with this notion, a superposition of the apo Huwe1 HECT domain[113] (PDB ID: 5LP8) with the Huwe1 C-lobe loaded with a donor-ubiquitin[162] (our lab, PDB ID: 6FYH) reveals extensive clashes between the donor-ubiquitin and the HECT N-lobe domain (Figure S 8a), suggesting that the C-lobe would have to move away from the N-lobe in order to accommodate the conjugated ubiquitin.

Moreover, the catalytic Cys is buried in the apo HECT domain structure and is thus not accessible for catalysis. Consistently, the CSPs observed for the Huwe1 HECT C325K mutant suggest that the interface between C- and N-lobe is very sensitive to even small additions near the catalytic cysteine. The clashes generated when replacing the catalytic cysteine with a lysine in the Huwe1 HECT structure (Figure S 8b), further confirm that the structure of the apo HECT domain represents a closed state with a significant lobe interface.

The effect of changes near the catalytic cysteine seems to be dependent on their size, as the CSPs observed for the C352K mutant are smaller than those detected upon addition of a whole ubiquitin (Figure 6). This size-dependency was further confirmed by a previous lab member, Dr. Magnus Jäckl, who performed a complementary experiment in which he attached a spin-label (4-maleimido-TEMPO(reduced)) to the catalytic cysteine, increasing its size, and observed CSPs of residues I268/I270/I331 closer to the changes (relative to the wild-type) induced by Huwe1 HECT~donor-ubiquitin disulfide, while smaller amino acids such as Ala and Asn induced smaller chemical shift changes.

The direct correlation between the extent of the CSPs and the size of the modification added to the catalytic cysteine suggests a conformational change where the conformational equilibrium is successively shifted towards the open state where the C-lobe is dissociated from the N-lobe.

The isolated C-lobe is the natural endpoint of this structural change, representing a completely open state of the HECT domain. On the contrary, we cannot determine whether the peak positions of the wild-type apo Huwe1 HECT domain represent the other end point of a completely closed state or this construct just samples the completely closed state the most out of the constructs we tested.

3. Confirming the existence of an N-/C-lobe interface by structure-based mutagenesis

The spectral changes progress towards a state where the two lobes of the HECT domain are fully separated from each other suggests the existence of an interface between the two lobes, holding them together in a closed conformation, in solution. To assess the biological relevance of this interface to Huwe1 function, and to determine its effect, we decided to probe the HECT domain with mutations chosen to disrupt this interface.

Since cysteine 325 is the catalytic cysteine in Huwe1, it is not possible to use the C352K mutant for activity assays. We instead focused on two residues in its vicinity (M328 and Y126), that possibly have a burying effect on it and could contribute to the lobe interface (via hydrophobic interactions to e.g., F353), and mutated these residues at the same time. To interfere with the lobe interface without secondary effects due to proximity to the catalytic cysteine, we mutated E259, which is on the N-lobe side remote from the catalytic cysteine. This residue is also identified as contributing to the C-/N-lobe interface in the HECT 5LP8 PDB structure using the PISA ("Proteins, Interfaces, Structures and Assemblies") server[195] (Table S 3), through a H-bond to S315 in the C-lobe. To minimize effects on the global fold, all residues were mutated to alanine.

We then recorded spectra of of these mutants (IM-^{13}C-Methyl labelled) (Figure 7). The mutants designed to interfere with the C-/N-lobe interface show spectral changes going in the same direction as the previously recorded C352K mutant. This supports the hypothesis that there is a C-/N-lobe interface whose binding can be interfered with in Huwe1. At the same time, since the mutations were designed based on the published 5LP8 structure of the Huwe1 HECT, this suggests that the same conformation, in which the C-/N-lobe interface is closed, is indeed the predominant one in solution.

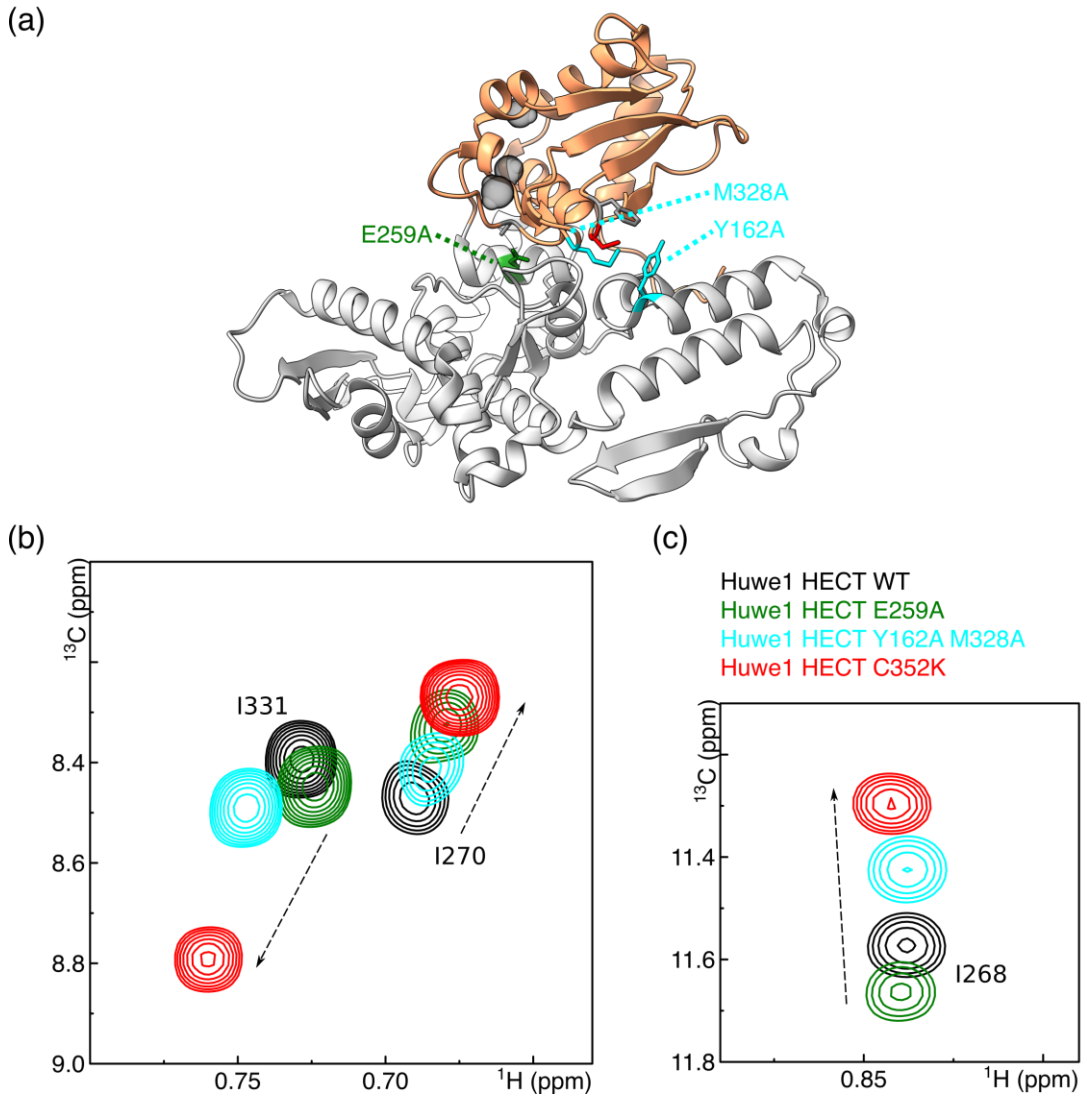


Figure 7. (a) Huwe1 HECT structure (PDB ID 5LP8) with mutants rationally chosen to disrupt the C-/N-lobe interface depicted. Putative interacting partners of mutated residues shown as sticks in dark grey. Van der Waals radii of methyl group reporters are shown in transparent black. (b/c) Changes in the SOFAST-HMQC spectra of IM-^{13}C-methyl labelled Huwe1 HECT domains (at 303 K) due to mutations chosen based on the PDB structure 5LP8 with the aim to destroy the C-/N-lobe interface (E259A, Y162A/M328A); Huwe1 HECT wild-type as a reference (black). The C352K mutant (See Figure 6) spectra (red) is also shown for comparison.

4. *The Y162A/E259A/M328A mutant shows CSPs similar to a donor ubiquitin-loaded Huwe1 HECT domain.*

The mutations described in section 3 at the interface between N- and C-lobe produced CSPs in the Huwe1 HECT domain spectra progressing towards a state in which the two lobes are dissociated (Section 3). However, the observed effects were not as big as those for the catalytic cysteine mutant (C352K), consistent with our observation that the extent of the conformational change scales with the size of the modification disrupting the interface. Since the E259A single mutant is located further away from the Y162A/M328A double mutant, yet the spectral changes it induces go in the same direction, we decided to combine these mutations to test whether they would have a synergistic effect on destabilizing the interaction interface.

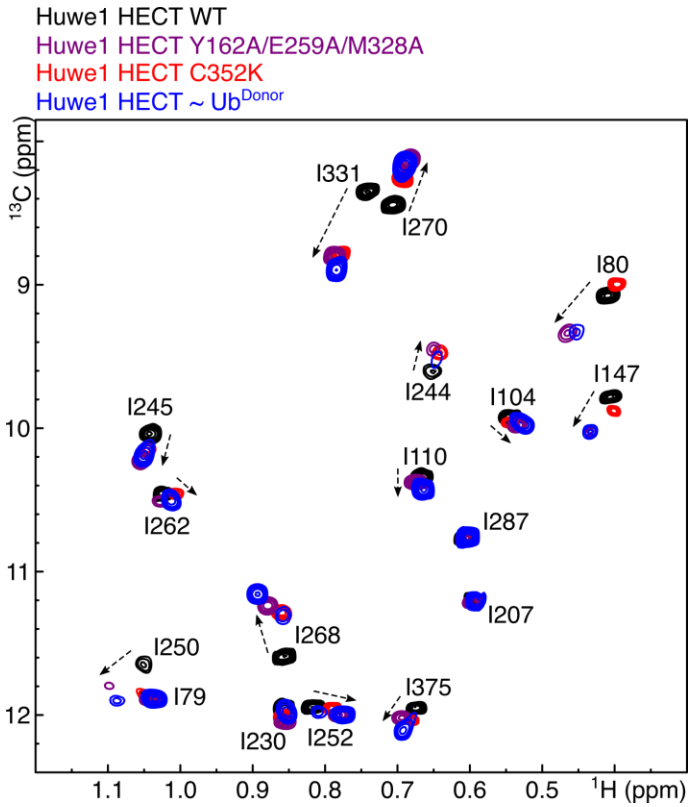


Figure 8. Changes in the spectra of IM-^{13}C-methyl labelled Huwe1 HECT domains (at 303 K) due to mutations chosen to impair the C-/N-lobe interface (E259A/Y162A/M328A); wild-type Huwe1 HECT spectra as a reference (black). Spectral changes due to this set of mutations are more similar to changes induced by having a donor-ubiquitin loaded onto the catalytic cysteine of the HECT domain (HECT~Ub^{D}: disulfide analogue) than changes induced by the single catalytic cysteine mutant (C352K).

These combined mutations (triple mutant: E259A/Y162A/M328A) indeed produced spectral effects closer to those produced by the catalytic cysteine mutant than the individual sets of mutants (E259A and Y162A/M328A) by themselves. In fact, the spectra of the triple mutant were closer to the spectra of the disulfide analogue of the HECT domain loaded with a donor-ubiquitin (Figure 8). This effect can be seen not only on the previously mentioned sensor residues, such as I270 in the C-lobe of the HECT domain (aa266-385), but also in N-lobe (aa1-263) residues, such as I80 and I147. The signals of these residues are not strongly affected by the catalytic cysteine mutant (Figure 8) but show very similar chemical shift perturbations in the triple mutant as in the donor-ubiquitin-loaded HECT domain (Figure 9).

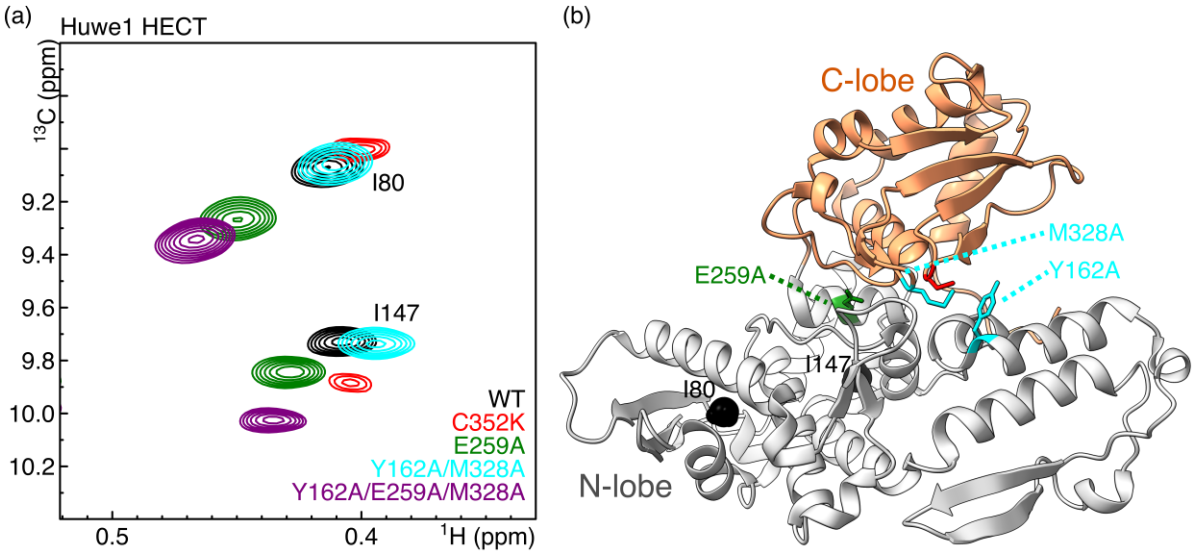


Figure 9. (a) Spectral changes (at 303 K) of the reporter residues I80 and I147 methyl groups in the Huwe1 HECT domain WT (black) or E259A (dark green), Y162A/M328A (blue), Y162A/E259/M328A (purple) sets of mutations. The Y162A/M328A mutations alone have little effect on N-lobe residues but enhance the effect of the E259A mutation when combined. (b) Huwe1 HECT structure (PDB ID 5LP8). Mutated residues are highlighted in green/blue, reporter methyl groups shown in black.

Since the Y162A/M328A double mutant targets residues in the vicinity of the catalytic cysteine, it is not surprising that these secondary effects on the N-lobe cannot be observed, similar to the C352K mutant spectra. The CSPs on N-lobe residues observed in the HECT domain linked to a donor-ubiquitin are however consistent with a shift in the conformational equilibrium towards the open state caused by steric hindrance; similar CSPs are also observed in the triple mutant, and to some extent in the single E259A mutant.

5. *The C-/N-lobe interface influences Huwe1 autoubiquitination and thioester formation activity*

Having found a set of mutants that disrupt the C-/N-lobe interface, we tested the effect of these mutants on the activity of the HECT domain. With the exclusion of the C352K mutant, which is inherently catalytically dead, autoubiquitination assays have been performed with all the mutants introduced in Section 3 (in a wild-type background while the spectra were in cysteine reduced background). This assay allows one to determine the activity of the HECT domain, using the domain itself as substrate. The wild-type Huwe1 HECT domain was used as reference for baseline autoubiquitination activity.

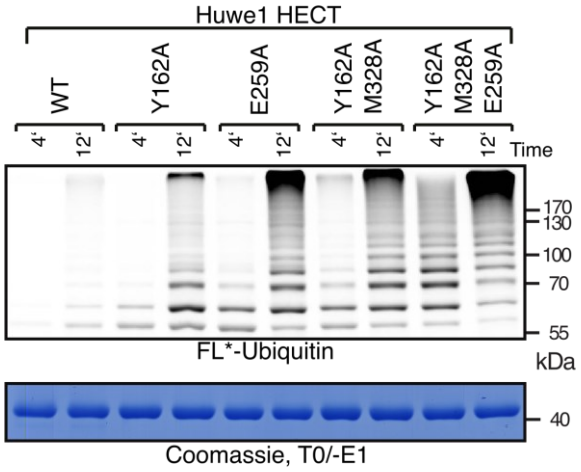


Figure 10. Auto-ubiquitination assay of C-/N-lobe interface mutants compared to the wild-type Huwe1 HECT domain (all His6-tagged). 60 µM fluorescently-labelled ubiquitin together with 10 µM UbcH7 (as E2) as well as 3 µM of the respective Huwe1 HECT mutants (as E3) were incubated together in ATP/Mg^{2+} containing assay-buffer at 30°C for indicated times after adding 0.5 µM Ube1 (as E1). Reactions were stopped by adding Lämmli sample buffer, separated via SDS-PAGE and fluorescence imaged (top panel). A sample was taken before starting the reaction by the addition of E1, and was loaded on a separate coomassie stained SDS-PAGE gel as a control for protein concentration and sample homogenity (bottom panel).

The autoubiquitination assay shows a clear increase in activity of the HECT domain in the interface opening mutants (Figure 10). Interestingly, this effect correlates very well with the degree of opening induced by the mutants as detected in the spectra: the single mutants are clearly more active than the wild-type, and this activation is further enhanced by their combination, indicating a synergistic effect, with the Y162A/M328A/E259A mutant being the most active.

In HECT E3 ligases, ubiquitin is transferred from a E2 ubiquitin-conjugating enzymes to the HECT domain as a reactive thioester moiety. The HECT domain then transfers ubiquitin to a substrate lysine via an isopeptide bond. Deletion of the last four residues in the HECT domain (HECT Δ-4) specifically impairs the latter step, without affecting trans-thiolation[105–107,121]. In this way, the HECT ~ ubiquitin thioester formation can be observed independently to determine at which reaction step the weakening of the lobe interface acts to increase autoubiquitination activity. When introduced in the HECT Δ-4 background, the interface opening mutants clearly show a positive effect on the rate of thioester formation (Figure 11) that, like the increase in autoubiquitination, correlates with the degree of opening induced by the respective mutant.

While these results do not exclude an effect on isopeptide bond creation, they clearly indicate that the stimulation of thioester bond formation is a major contributing factor to the increase in autoubiquitination (Figure 10) observed for the opening mutants. Unfortunately, as HECT domains are a substrate for themselves, it is not possible to observe isopeptide formation rates independently from thioester formation.

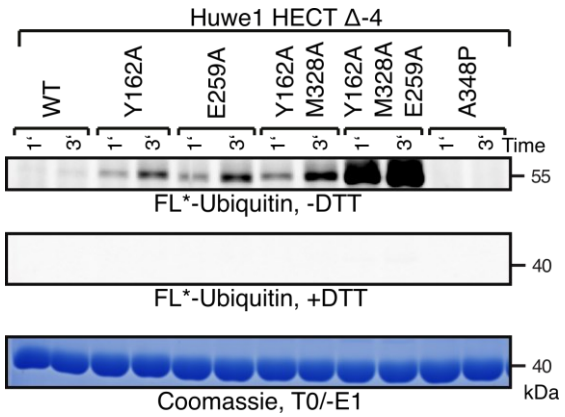


Figure 11. Thioester formation assay of C-/N-lobe interface mutants (as Δ-4), with Huwe1 HECT Δ-4 domain as reference. 15 µM fluorescently-labelled ubiquitin together with 1 µM Ube1 (as E1) as well as 30 µM UbcH7 (as E2) in ATP/Mg^{2+} was pre-incubated together for 5 minutes before ubiquitin activation was quenced with EDTA. Afterwards 15 µM of the respective Huwe1 HECT Δ-4 mutants (as E3) were added. The inactivating mutant A348P was used as negative control[162]. Reactions were stopped by the addition of Lämmli sample buffer after indicated times. Samples were split and half was treated with DTT before being separated via SDS-PAGE and fluorescence imaged (top panel -DTT, middle panel +DTT). An equivalent amount of the respective Huwe1 HECT Δ-4 domain was loaded on a separate SDS-PAGE gel (bottom panel) to show equivalent protein concentration and sample homogenity.

6. *CPMG relaxation dispersion experiments show that the Y162A/E259A/M328A mutant is introducing dynamics between an open and a closed state*

The C- and N-lobes in isolation naturally define the fully "open" state of the HECT domain. The gradual change in chemical shift induced by interface-disrupting mutations derives from a shift in a dynamic equilibrium between two states.

Many chemical shift changes proceed without substantial losses in intensity demonstrating that the exchange rate (k_{ex}) between the open and closed conformations lies in the fast chemical exchange regime, i.e., the exchange rate is significantly faster than the chemical shift difference ($\Delta\omega$; $k_{ex} >> |\Delta\omega|$). If the exchange rate were slower, one would either see a broad progressing signal (intermediate exchange; $k_{ex} \approx |\Delta\omega|$) or separate signals for each conformation, when $k_{ex} << |\Delta\omega|$ (slow exchange). The chemical shift difference between the full-length HECT domain ("closed" conformation) and the isolated C-lobe ("open" conformation) can be determined using the HMQC spectra of the respective states. For example, for one of the reporter residues, I331, $\Delta\omega$ (^{13}C) ≈ 0.756 ppm ≈ 151 Hz for the $^{1}H,^{13}C$ HMQC spectra recorded on an 800 MHz spectrometer. We can therefore assume that the exchange rate between the two conformations should be significantly higher than 151 Hz.

To characterize the dynamic equilibrium between the open and closed state of the Huwe1 HECT domain, we used Carr-Purcell Meiboom-Gill relaxation dispersion (CPMG RD) experiments, that measure the contribution of the chemical exchange (R_{ex}) to the transverse ("spin-spin") relaxation (R_2). Since in preliminary experiments it became evident that we were at the upper limit with respect to the observable exchange rate for CPMG RD experiments at 303 K (30°C), we recorded relaxation dispersion data at a lower temperature, 283 K (10°C), to decelerate the conformational dynamics.

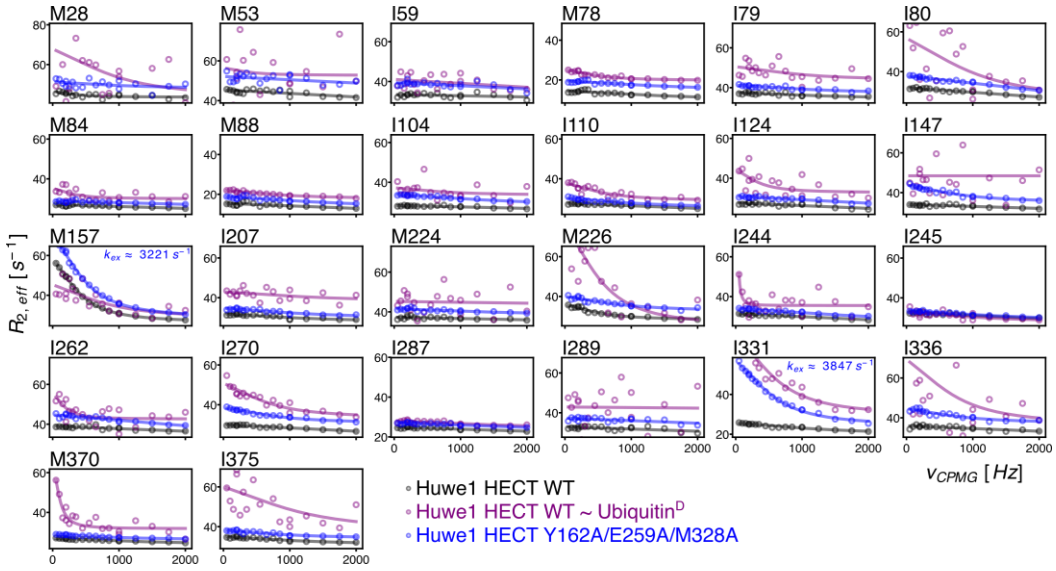


Figure 12. CPMG relaxation dispersion curves of the WT Huwe1 HECT domain (black), the Huwe1 HECT disulfide mimic linked to donor-ubiquitin (purple) as well as the Y162A/E259A/M328A triple mutant (blue), recorded at 283 K on an 800 MHz Bruker AVANCE spectrometer. Residue numbers are indicated on top of the individual panels. The curves show the individual fits against the Luz-Meiboom fast-equation[196–198]. The fitted K_{ex} values of the triple mutant are indicated for residues where the fit was visually distinct from linear.

We first recorded CPMG data on an isoleucine, methionine labelled WT Huwe1 HECT domain (catalytic-cysteine only background) (Figure 12, black). For most residues, there was a negligible effect on the effective relaxation rate ($R_{2,}$eff) due to an increase in CPMG frequency (ν_{CPMG}). This demonstrates that no dynamics can be observed on the CPMG time-scale for this HECT domain variant (note: dynamics on slower time-scales can be ruled out by the absence of very broad or split signals in HMQC spectra).

We then performed the same experiment with the disulfide mimic of a donor-ubiquitin-loaded Huwe1 HECT domain (Figure 12, purple). Here, we could clearly observe changes in $R_{2,eff}$ depending on the CPMG pulse frequency (ν_{CPMG}) for many of the reporter isoleucines and methionines, such as I331 and I270 (note: I268 was not included in this comparison as it shows spectral overlap with the signals of other residues at 283 K). This shows that the donor-ubiquitin-loaded Huwe1 HECT domain undergoes dynamics on the CPMG time-scale for multiple residues, including the previously mentioned "open" state reporters. However, signal-to-noise ratio of this construct is significantly lower than that of the reference, with a contributing factor being the increased size of the whole construct, due to the addition of ubiquitin, as well as additional motions at the ubiquitin/C-lobe interface, possibly obscuring the C-/N-lobe dynamics. Because of this, the interpretation of the CPMG relaxation dispersion effect is difficult for residues, such as M28, M53, I80, I124, I147, M226, I244, I289, I336 and I375, while most of the remaining residues signals show no effect. Residue M157 showed dynamics in the wild-type (catalytic cysteine only background) as well as in the donor-ubiquitin loaded construct and is thus unlikely to report on the C-/N-lobe interface dynamics. M370, which is located close to the flexible N-terminus of the C-lobe might instead be an interesting candidate for investigating the involvement of the HECT N-terminus in isopeptide formation.

To sidestep issues with low signal-to-noise ratio and the involvement of ubiquitin, we used the previously established Y162A/E259A/M328A triple mutant as a proxy for the exchange process of the lobe interface. As expected, we detected interface related dynamics on reporter residues I331 and I270. To perform a more detailed evaluation of the exchange process due to interface opening, we also recorded CPMG experiments where the side chain methyl groups of alanines, threonines, leucines and valines were labelled in addition to isoleucines and methionines. Using this extensive labelling strategy on both the Y162A/E259A/M328A triple mutant and the wild-type HECT domain, we found an additional set of residues (T140, L261, L296, V311, L325) that showed similar dynamics as the previously identified reporter residues, I331 and I270 (Figure 13). To improve the quality of the fit we recorded both samples at two different field strengths.

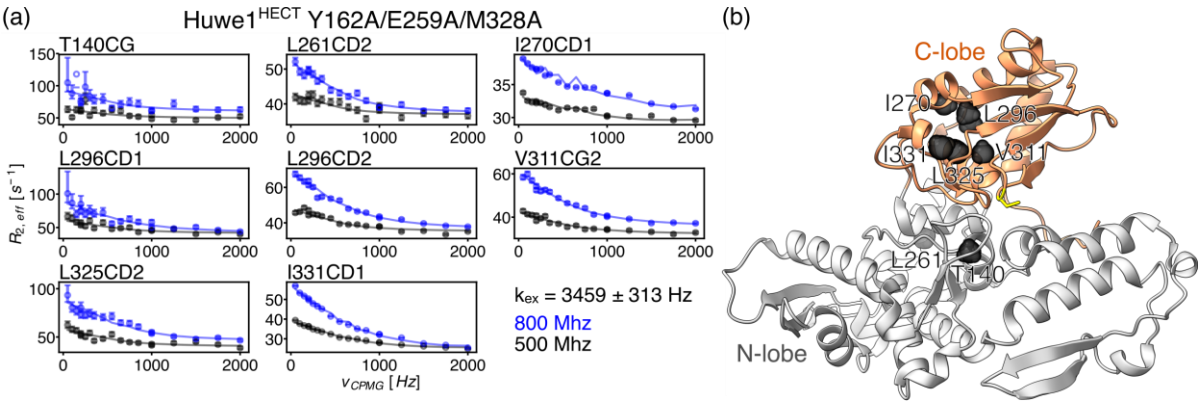


Figure 13. (a) CPMG relaxation dispersion curves of the Huwe1 HECT Y162A/E259A/M328A residues, recorded at 283 K on an 800 MHz spectrometer (blue), as well as a 500 MHz spectrometer (black). The curves are back calculated from the result of a global numeric fit of all shown using the program ChemEx assuming a 2-site exchange model. Residues were selected by their exchange behaviour. The fit results in an exchange rate of 3459 ± 313 s^{-1}. (b) Location of the fitted methyl groups in the Huwe1 HECT domain (PDB ID: 5LP8).

However, since we are in the fast-exchange limit, the experiment could still not determine a population estimate independently of the Δω between both states, since these values are correlated when fitting the data (Figure 14(a)).

Because of this, the fit does not allow the determination of population and Δω of the dynamic change. However, it confirms that the dynamic change observed is indeed the structural change between the wild-type conformation and the fully free lobes, as the determined R_{ex} values correlate well with the

(squared) chemical shift difference between the wild-type conformation and the free lobes (Figure 14(b))[199].

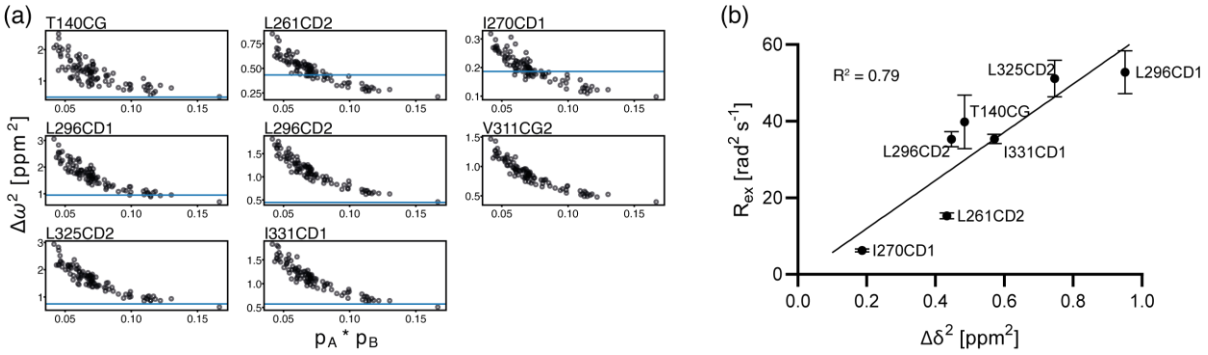


Figure 14. (a) Δω2(^{13}C) plotted against the population product estimated from 100 global numeric fits (bootstrap runs) of all indicated residues, using the program ChemEx on the CPMG relaxation dispersion data of Huwe1 HECT Y162A/E259A/M328A (data and fitted model as in Figure 13). Chemical shift difference between WT and individual lobes as thin blue line. (b) Per residue Rex at 800 MHz determined from bootstrap runs shown in (a) plotted against the square of the chemical shift differences between the wild-type and the individual lobes measured in the HMQC spectra. Error bars represent the standard deviation among the bootstrap runs.

Because the chemical shifts in the triple mutant are approximately halfway between the positions of the free lobes and the wild-type HECT domain (Figure S 9), one could hypothesize a population estimate of 50%. This is however unlikely, as fixing the population to 50% and the residue Δω's to half the difference between the wild-type and the free lobes does not yield a satisfactory fit result (Figure S 14).

One reason for this could be that the wild-type HECT domain does not represent the fully closed state. In fact, this is highly likely since the HECT domain would not be catalytically active if it adopted a fully closed conformation.

Of note, in the construct of the Huwe1 HECT domain including an additional α-helix ("pointer helix") N-terminal to the HECT domain, as introduced by Sander et al. 2017[113], our sensor residues (I80, I147, I270, I331) did not report a significantly more closed state (Figure S 12).

7. Discussion

Previous investigations of autoinhibitory mechanisms of HECT domain E3 ligases have focused on elements proceeding the HECT domain, such as the N-terminal C2 domain[121,125], the WW domains[143], or both[147]. Their autoinhibition activities have been predominantly ascribed to blocking access to the ubiquitin binding surface[125], the E2 binding surface[200] or the catalytic cysteine[121] by binding to overlapping sites on the HECT domain. In addition, a number of recent publications showed that various inter-domain linkers N-terminal of the HECT domain in various Nedd4 family members, such as Nedd4[201], WWP2[114,201] and Itch[115], lock the domain in a closed conformation by binding to and enforcing the interface between the N- and C-lobes. These more recent results point towards the restriction of inter-lobe flexibility and inaccessibility of the catalytic Cys as a general pattern of HECT domain regulation. This raises the question whether the inhibitory mechanisms that were ascribed to the simple blocking of various surfaces might also work by restricting inter-lobe conformational flexibility. Recent work from our lab shows that, at least in the case of the C2-WW1 module in Smurf2, such regulatory domains are indeed able to contact both the C- and the N-lobes at the same time[147].

The results described here reveal that a HECT domain, that of the Huwe1 E3, is inhibited independently of sequence elements N-terminal of the HECT domain, but solely due to an inter-lobe interface that is apparently more stable than in other HECT domains.

We identified here a set of residues that directly report on the interaction between the N- and C-lobe. This allowed us to follow the degree of interface opening in solution in NMR experiments. With this setup we detected significant opening of the N-/C-lobe interface in a HECT domain loaded with donor-ubiquitin, and we could mimic this opening by mutating residues important for stabilizing the interface. Since loading the HECT domain with ubiquitin must be accompanied by an open conformation for steric reasons, one can surmise that during the transfer of activated ubiquitin from an E2 to the HECT E3, the energy barrier of interface opening must be surpassed in the Huwe1 HECT domain (Figure 15 (2)). We could show that our set of mutations increases the rate of thioester formation (Figure 15 (2,3)) in a synergistic manner, which is consistent with an additive lowering of this energy barrier. Since these mutants also consistently increase Huwe1 HECT autoubiquitination activity overall (Figure 15 (4,5)), any negative effects, due to a disruption of another mechanistically important feature, such as impeding interface closing for example, are not likely to be significant.

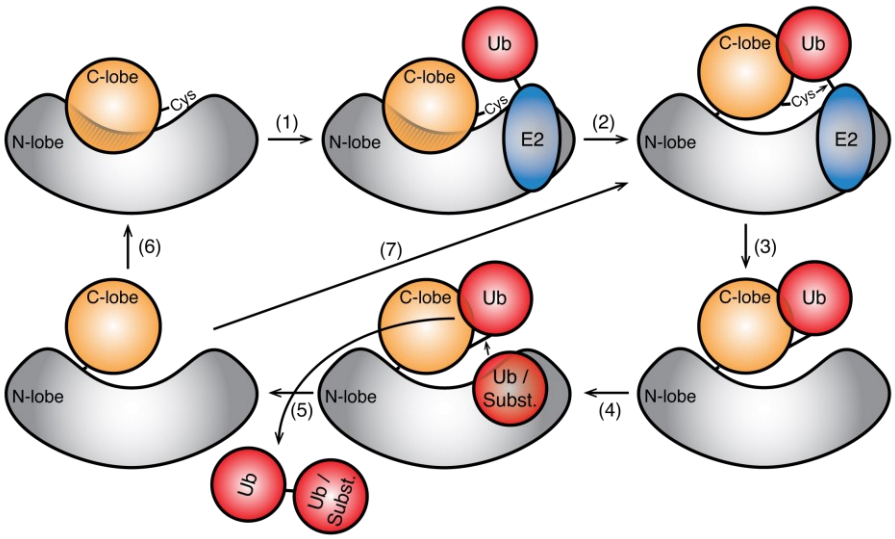


Figure 15. Proposed reaction cycle of the Huwe1 HECT domain. The lobe interface has to be open to allow ubiquitin thioester formation. Ubiquitin is successively transferred to substrates or a growing ubiquitin chain which are displaced after isopeptide formation from the catalytic centre by the subsequent acceptor ubiquitin.

While none of the Huwe1 mutants presented here has been characterized *in vivo*, the E798V mutation has been identified in a Nedd4 family HECT E3 member, WWP1[144], isolated from prostate cancer. This change was reported to have an activating effect in *in vivo* experiments[144], and is in a position structurally equivalent to E259 in Huwe1, whose mutation also stimulated ubiquitination in our assays. More recently, the same conserved glutamate has also been described in Arel1, another member of the HECT family, where the E701A mutation was found to increase protein activity in *in vitro* experiments.

E3 ubiquitin ligases have been called *"undruggable targets"*[202,203] before, and the main drug discovery approach to target the *"large and flat protein-protein interaction surfaces"*[203] with small molecule drugs has proven challenging. The relatively new paradigm of HECT E3 activity regulation via restriction of inter-lobe flexibility invites a re-consideration of the more traditional approach in which small molecule inhibitors fitting in small (active site) pockets were usually sought. Huwe1 has been identified in various studies as an oncogene as well as a tumour suppressor, depending on the cellular context[161] and on the type of cancer, and would be an interesting target for drugs designed to intercalate and open the interface, or bind both lobes and stabilize the closed conformation. Of note, NMR would have the advantage over *in vivo*- or even *in vitro*-based screening approaches, since it allows for an easy, reliable, and reproducible readout, and enables a measurement of the conformational change which

is direct and therefore independent from an upstream cascade of factors that need to be expressed and purified and could introduce confounding effects.

C. Huwe1 HECT non-covalent ubiquitin binding is enhanced by an open HECT domain conformation

1. Introduction

1. Ubiquitin binding abilities of different HECT domains

Despite substantial research done on the ubiquitin-binding exosite (ubiquitin-binding-site, UBS)[102,109,110,120,124,204], the exact role of this motif remains controversial. The earliest report on the UBS, by French et al.[124], suggested that it would restrict the length of polyubiquitin chains. However, since then, several publications[102,109,110,120,204] have proposed that it would instead promote polyubiquitination, or even represent a "processivity site"[110] that works by keeping ubiquitin chains bound during elongation. Even more recently, Zhang et al. showed that Ubiquitin-Variants (UbVs), selected by phage display to have high affinity for HECT domains, act not only at the level of chain elongation, but also have more direct, chain-elongation independent effects[112].

Interestingly, and in stark contrast to other Nedd4-subfamily members[112], a ubiquitin-binding exosite (ubiquitin-binding-site, UBS) on the N-lobe has so far not been reported for Huwe1[108,113,162].

In this chapter I aim to investigate the presence of a UBS on Huwe1 and compare its role with that of ubiquitin-binding exosites from Nedd4-subfamily members. To this end, we performed functional ubiquitin chain formation as well as E3 - ubiquitin thioester assays. A comparison between the activity of (putative) UBS mutants and wild-type Huwe1 confirmed the presence of a functional UBS. We further discovered that the UBS is occluded in the auto-inhibited state of Huwe1, explaining the lack of ubiquitin binding observed in previous studies. Using a Huwe1 mutant designed to obviate UBS occlusion in NMR experiments, we detected ubiquitin binding in solution, and mapped its interaction surface on the HECT domain.

2. Autoubiquitination activity depends on an intact non-covalent Ubiquitin Binding Surface

In contrast to Nedd4 family E3s, Huwe1 substrates are degraded by the proteasome in a K48-linked poly-ubiquitination-dependent manner[173]. In many Nedd4 E3s, the UBS is required for efficient poly-ubiquitination[120]. However, we failed to detect an interaction of the Huwe1 HECT domain with ubiquitin in NMR titration and GST pull down experiments[162]. Dimerization-dependent autoinhibition, reported by Sander et al. 2017[113], does not explain our results, as the construct for the Huwe1 HECT in use in our lab solely included the HECT domain (starting at residue 3992) and not the preceding pointer and thumb helices (3951-3991) reportedly required for dimerization. Additionally, the reported dimer interface does not overlap with the binding surface of a canonical UBS.

Since many residues in the UBS of Nedd4 enzymes are conserved in Huwe1, we decided to investigate whether Huwe1 possesses a UBS. For HECT domain E3 ligases, enzymatic activity can be tested using auto-ubiquitination assays, which is impaired in Nedd4 enzymes by mutations of the UBS[120]. Multiple candidate residues in structurally conserved positions on Huwe1's putative UBS (see Figure S 15) were mutated to alanines, to minimize the risk of disturbing the overall fold, and the corresponding constructs were tested for auto-ubiquitination.

Residue numbers given here are relative to the start of the HECT domain, thus residue 1 of the HECT domain is residue 3990 of the *Homo Sapiens* Huwe1 full-length (FL) protein in the canonical isoform 1 (UniProtKB: Q7Z6Z7-1).

F192 (F4181 in FL Huwe1) was chosen because the corresponding residue has been observed to interact with a hydrophobic patch around I44 on ubiquitin in multiple HECT-ubiquitin complex structures (Rsp5 F618[109,124], Nedd4 F707[125]). Q190 is in close proximity to F192 and could also contribute to the interaction. The residue at position C110 of the HECT domain (C4099 in FL Huwe1) corresponds to I537 in Rsp5, which is known to be part of the UBS and, when mutated to aspartate, impairs ubiquitin binding[109]. M88 was selected as a control, since it is located in the UBS but its mutation does not impair binding in Smurf2 (Y453[125]) or Nedd4 (Y604[125]).

To test the overall activity of these Huwe1 HECT domain variants, we then performed *in vitro* ubiquitination reactions using H6-tagged wild-type (WT) Huwe1 HECT-domain or the indicated mutants as E3 ligases, hUbcH7 as E2, and fluorescently labelled ubiquitin. The reaction was started by adding hUbe1 as the E1. In the absence of a substrate, ubiquitination was observed as HECT domain auto-ubiquitination.

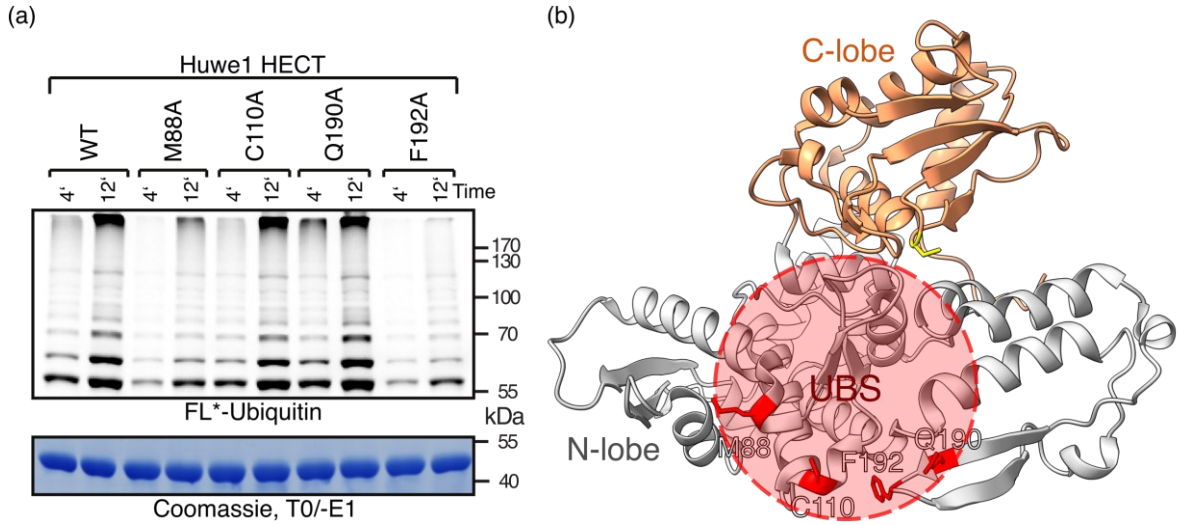


Figure 16. Mutations of the UBS interfere with Huwe1 HECT auto-ubiquitination. (a) H6-tagged WT Huwe1 HECT domain and indicated mutants are compared for auto-ubiquitination activity. Top panel: Fluorescently labelled ubiquitin is detected as it is attached to the Huwe1 HECT domain and forms chains. Long ubiquitin-chain products collect at the top of the gel but remain quantitatively measurable due to direct fluorescence detection. Bottom panel: Separately loaded coomassie-stained gel of the reaction mixture prior to the addition of E1, as a control for Huwe1 HECT WT/mutant protein levels (~50kDa band). (b) Structure of the Huwe1 HECT domain with the respective mutations shown as sticks and coloured in red (PDB ID: 5LP8).

Both the M88A and the F192A mutation considerably inhibit auto-ubiquitination activity compared to the wild-type Huwe1 HECT (Figure 16). In the case of F192A a strong effect is not surprising, since the residue and the structural element in which is it located (a short loop connecting the two strands of a beta hairpin) are completely conserved in Huwe1 and other HECT domains. In contrast, residues in structurally equivalent position to M88 had been described as “in” the UBS but not a “hot spot” for ubiquitin binding[125]. Since this residue is a conserved tyrosine in other HECT domains, but a methionine in Huwe1, it might contribute to the UBS-ubiquitin binding strength more in Huwe1 than in other HECT domains. Alternatively, it is also possible that our activity assays would be more sensitive to small variations in binding affinity than interaction assays. Neither C110A nor Q190A mutations showed a significant effect on autoubiquitination activity (Figure 16). C110A is only partially conserved, being an isoleucine in Rsp5 (I537) and Smurf1 (I455), a valine in Smurf2 (V475), and a cysteine in Nedd4 (C627). While the introduction of a charged residue at that position was shown to impair ubiquitin binding in Rsp5, it is possible that mutating it to alanine, while limiting the risk of disrupting the overall fold, may not be sufficient to disrupt binding. We therefore cannot exclude that these residues might still be part of the UBS.

Nevertheless, the impact of mutating residues predicted to contribute to the ubiquitin-binding surface suggests the presence of a functionally important UBS in Huwe1, despite the absence of ubiquitin binding in the wild-type HECT domain[162].

3. *The Y162A/E259/M328A triple mutant allows ubiquitin to bind to the Huwe1 HECT domain in a manner consistent with canonical UBS binding*

The functional evidence for the existence of a canonical UBS in Huwe1, combined with the lack of ubiquitin binding to the HECT domain of Huwe1 in binding assays[162], was puzzling. However, since we previously observed that Huwe1, unlike other HECT E3 ligases, is autoinhibited via a mechanism inherent to the HECT domain itself (Chapter B), it is conceivable that the same mechanism might also restrict access to the UBS despite the fact that the UBS is not blocked in the crystal structure of the inhibited Huwe1 HECT domain. To test whether the conformational state of the Huwe1 HECT domain determines its ubiquitin binding ability, we used the previously introduced Y162A/E259A/M328A triple mutant, which destabilizes the C-/N-lobe interface, for NMR binding studies.

To this end, we recorded $^{1}H,^{13}C$-SOFAST spectra of the isoleucine- and methionine- (IM-)^{13}C-Methyl labelled E259A/Y162A/M328A triple mutant in isolation and in the presence of successively increasing amounts of unlabelled ubiquitin (Figure 17a). As in NMR the chemical shift of labelled amino acids is highly sensitive to their chemical environment, the chemical shift perturbations that we observed upon addition of increasing amounts of ubiquitin clearly showed that the HECT domain triple mutant is able to bind ubiquitin, and allowed us to determine the dissociation constant (K_D) of binding (171 µM), fitted using line-shape analysis with the software package TITAN[205]. For the analysis, we used residues I270, I331, I244, I245, I147, I207, I336, M28, I375, I252, M224, M84, M88, I59, M53, I110, M78. In contrast, wild-type Huwe1 HECT did not show any CSPs even at a high molar excess of ubiquitin[162].

While almost all labelled residues showed chemical shift changes upon addition of monomeric ubiquitin, the signals of residues I124, I80, I104, I110, I230, M157, M226, M88, M84 and I59 broadened significantly quicker than others (Figure 17a). A quantitative evaluation of the exchange rate using TITAN was not possible for these signals, as spectra were recorded using the SOFAST pulse sequence, which is not appropriately filtered for fast-relaxing zero-quantum multiplet components of the signal[205–207]. However, a qualitative evaluation shows that these residues which show peak broadening upon ubiquitin titration, already at a sub-equimolar ratio, are located exclusively in the canonical UBS (Figure 17b/c), indicating an intermediate- to slow-exchange regime of ubiquitin binding and thus can be assume to have K_D values in the low micro- to nanomolar range. Residues showing intermediate- to fast-exchange behaviour (M53, I79, I147, I207, M224, I244, I245, I250, I252, I262, I268, I270, I289, I331, I336, I375) are instead located outside of the canonical UBS, such as on the back or side of the N-lobe, with respect to the UBS, or in the C-lobe (Figure 17b/c). The few residues which show no CSPs upon ubiquitin titration are located on top of the C-lobe or on the side of the N-lobe and tend to be the furthest away from the canonical UBS and the C-/N-lobe interface (I287, M370: C-lobe backside; M28, M78: N-lobe left side). These, changes could be indicative of a more global structural change of the HECT domain induced by ubiquitin binding, possibly of further opening.

As the HMQC spectrum of the Y162A/E259A/M328A triple mutant is similar spectra to the donor-ubiquitin loaded WT HECT domain, they seem to represent a similar conformational equilibrium. Consistent with an open conformation of the HECT domain being required for ubiquitin binding, a previous lab member, Dr. Magnus Jäckl, showed that the donor-ubiquitin-loaded Huwe1 HECT domain[194(5.3.7.)], and the N-lobe in isolation can interact with monomeric ubiquitin via the UBS[194(5.3.5.)]. Interestingly, chemical shift perturbations in the open-state reporter residues (Chapter B.2.) upon ubiquitin titration, when visible, go roughly into the direction of the open-state in most cases (I331, I147, I268), with the only exception of I270 (Figure 17a). This would be consistent with ubiquitin binding shifting the open/closed state equilibrium further towards the open state.

To achieve full occupancy of the ubiquitin-binding site Magnus Jäckl designed a mutation, N113C, which allowed us to link ubiquitin (G76C) to the HECT domain via a disulfide bond, in a position compatible with UBS binding. At the same time, we opened the C-/N-lobe interface by linking an additional ubiquitin (G76C) to the catalytic cysteine, to act as a donor-ubiquitin thioester mimic. The

signals in the spectra of this HECT domain with two linked ubiquitins are either at the final position of the free ubiquitin titration or, for residues that had not reached saturation, further in the direction in the perturbation observed during the titration (Figure S 16). Interestingly, when we titrated this double disulfide construct with further ubiquitin, in an attempt to reveal a potential binding surface for an acceptor-ubiquitin (substrate ubiquitin for chain formation/elongation), we did not observe chemical shift perturbations (Figure S 17). This suggests that an acceptor site, if present, could overlap with the ubiquitin-binding exosite, suggesting that polyubiquitination by HECT E3s could be a highly dynamic process requiring disassociation from the exosite before catalysis.

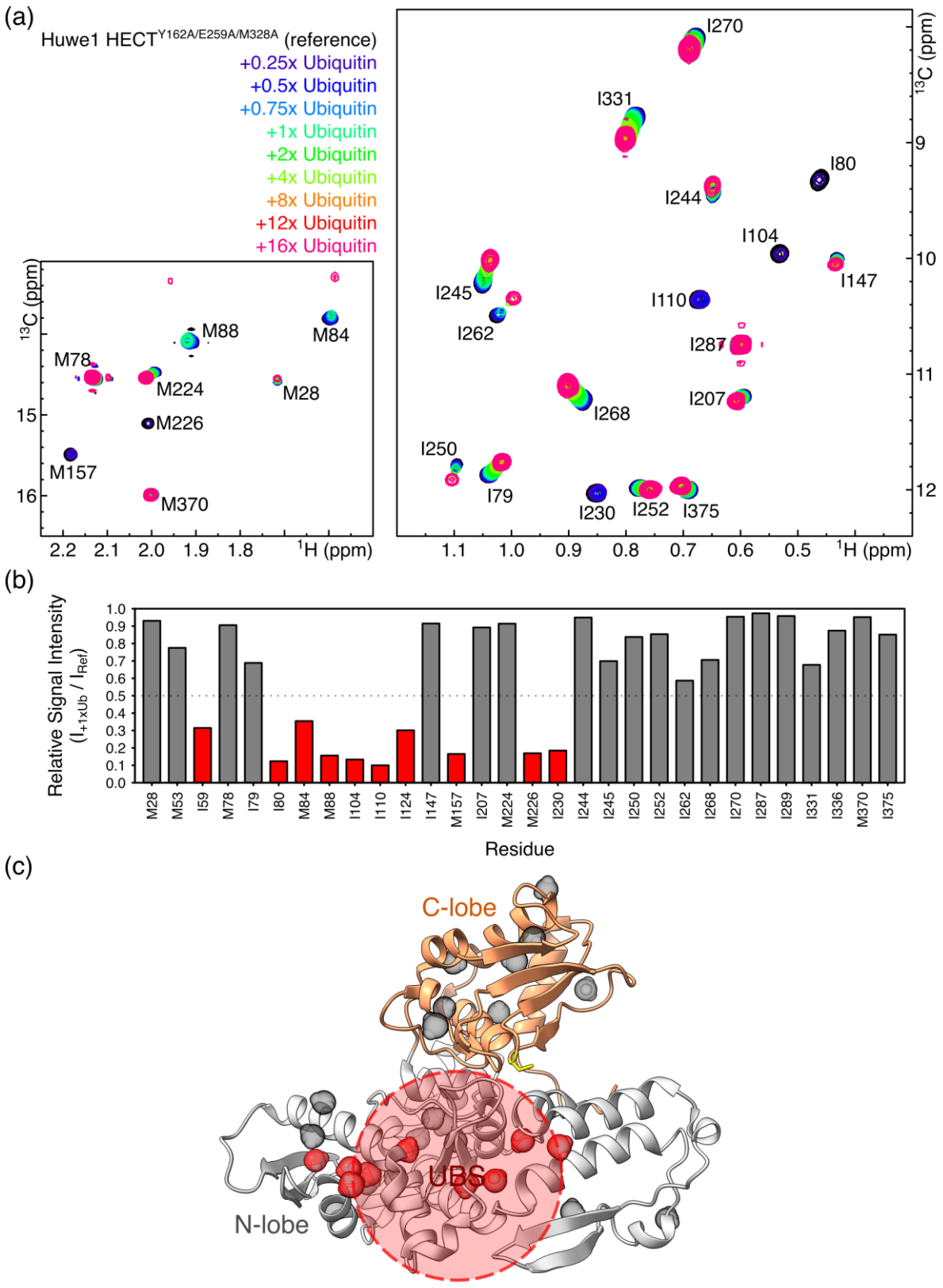


Figure 17. (a) Spectra of IM-^{13}C-Methyl labelled Huwe1 HECT Y126A/E259A/M328A triple mutant (at 303 K) without (reference) and with ubiquitin at stochiometric ratios of 0.25x to 16x added as indicated. (b) Signal intensity ratio between the reference spectrum of HECT Y126A/E259A/M328A and a spectrum of the same sample with 1x ubiquitin added, measured at the respective residues signal maxima in the respective spectra. Residues whose signals broaden below 50% of the original signal intensity are coloured red, others black, in the graph and on the (PDB ID: 5LP8) structure depiction (c).

4. *Thioester formation activity is indirectly influenced by the Ubiquitin Binding Surface*

Having established that the Huwe1 HECT domain possesses a UBS that is regulated via the tight C-/N-lobe interface, we sought to determine at which level the C-/N-lobe interface exerts its effect. HECT E3 ligases act via a two-step mechanism, in which ubiquitin is first transferred to the E3 catalytic cysteine, and successively linked to a substrate through an isopeptide bond to the ε-amine of a lysine or an N-terminal α-amine. Since the HECT~ubiquitin thioester formation precedes and is necessary for isopeptide formation, the latter activity depends on the former.

When the HECT domain thioester formation rate is quicker than its isopeptide formation rate, it is possible to use short reaction times to observe mainly thioester formation. Isopeptide formation can then by monitored by treating the reaction sample with a reducing agent, such as dithiothreitol (DTT) or beta-mercaptoethanol (β-Me), before analysing it on a gel. This ensures that the HECT-Ub band in the gel stems from the thioester and not from an isopeptide-linked ubiquitin on the HECT domain. This is the case for the HECT domains of e.g. Smurf2[162] and Nedd4[125].

When however, the rate of isopeptide formation is closer to the rate of thioester formation, the higher signal in the reductant-treated control can make quantification of thioester-specific signal difficult, since the difference will be very small and therefore hard to compare reliably. In addition, it may not be possible to shorten the reaction time sufficiently without dropping below the detection limit of the fluorescence readout. This is the case for Huwe1[162] (see Jäckl et al. Figure S6 for a comparison of Smurf2 with Huwe1). A way to circumvent this is to use a Huwe1 HECT construct lacking four C-terminal residues (HECT Δ-4), including a conserved Phenylalanine at position -4 which is essential for isopeptide formation, but does not influence thioester formation[106].

We therefore introduced the previously characterized UBS mutants (M88A, C110A, Q190A and F192A) in a HECT Δ-4 construct. The C-/N-lobe interface opening mutant E259A (Figure 11) was used as a positive control, while the inactivating mutant A348P, which disrupts the C-lobe – ubiquitin interface, served as a negative control[162].

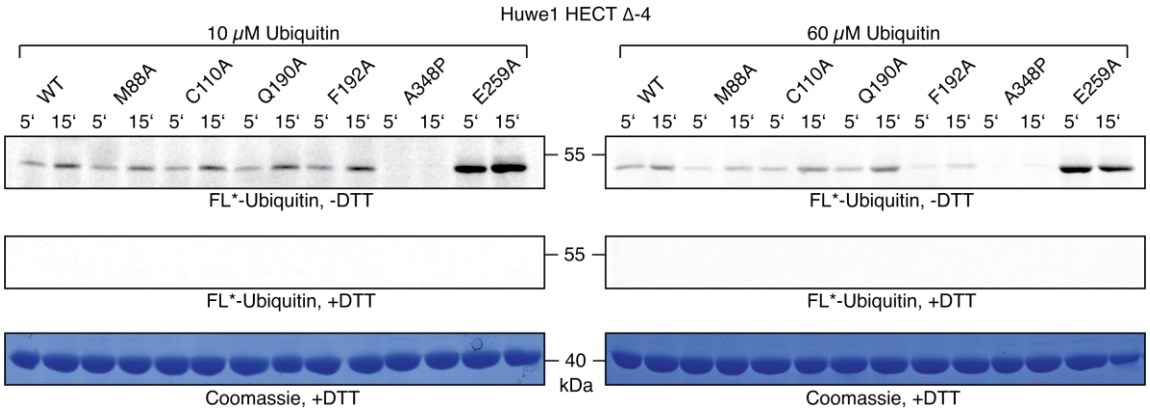


Figure 18. Thioester assay using fluorescently labelled (FL) ubiquitin and UBS mutants in the Huwe1 HECT Δ-4 construct that prevents isopeptide formation. Reactions were incubated at 30°C. Left panel: reactions at low (10 μM) ubiquitin concentration, stopped at 5 min and 15 min. Right panel: reactions at high (60 μM) ubiquitin concentration. All reactions contain 0.5 μM Ube1 (E1), 10 μM UbcH7 (E2), 5 μM indicated Huwe1 HECT Δ-4 WT or mutant. The thioester formation rate depends on the UBS at ubiquitin concentration in the upper range of physiological concentrations.

In the assay we used our standard ubiquitin concentration (10 μM) (close to single-turnover conditions as 5 μM HECT are used). Under these conditions, none of the UBS mutations affected thioester formation (Figure 18, left panel).

Given both the affinity range of HECT domain E3 UBS's for ubiquitin (~10 μM to ~400 μM[102,109,147]) and the wide range of ubiquitin cellular concentrations measured *in vivo* (up to ~85 μM in HEK293 cells[208]),

we repeated the assay with a higher ubiquitin : HECT ratio (Figure 18, right panel). Since we performed the experiment with the HECT Δ-4 mutation, in which isopeptide formation is prevented (Figure 18, +DTT fluorescence images), the use of close to single-turnover conditions, necessary for assays on the wild-type HECT domain, was not required. Under these conditions (60 µM ubiquitin) we could indeed observe a negative effect of the M88A and F192A mutations on thioester formation.

These results suggest that the ubiquitin : HECT ratio directly influences the thioester formation rate in a UBS-dependent manner. It is possible that the additional structural changes observed in the HECT domain outside of the UBS upon ubiquitin binding (see Figure 17, grey residues that still show CSP in (a) such as the C-lobe reporter residues I268, I270 and I331) are keeping the HECT domain in a conformation with an "open" N-/C-lobe interface, at least on a time-scale from one to the next transfer of ubiquitin from E2 to the HECT.

5. *Diminished UBS binding due to the stable C-/N-lobe interface seen in the Huwe1 HECT is not conserved amongst HECT-family E3's*

As described above, the Huwe1 HECT domain contains an interface between its C-/N-lobes which is important for ubiquitination activity regulation, via mechanisms both dependent and independent from UBS accessibility. Given that this regulatory mechanism had not been described before, we sought to investigate whether it might be conserved in other HECT family E3 ligases. To investigate the effects of a potential C-/N-lobe interface on UBS binding, we used the following approach: since the UBS is situated solely on the N-lobe, the affinity of ubiquitin for the full-length HECT domain should not differ from the affinity for the N-lobe in isolation, unless an interaction between N- and C-lobe interfered with UBS accessibility.

First, we tested the HECT domain of Rsp5, the sole Nedd4 family member in *Saccharomyces cerevisiae*. A construct of the HECT domain and one of the isolated N-lobe were expressed and titrated with ubiquitin. The K_Ds were determined using the program TITAN[205] to be ~35 µM for ubiquitin binding to the HECT domain (previously published to be 90 µM as measured by fluorescence anisotropy[109]) and ~33 µM for binding to the isolated N-lobe (Figure 19). Given the negligible difference between the two constructs, we concluded that, even if there would be a stable C-/N-lobe interface, its influence on ubiquitin - UBS binding is not relevant.

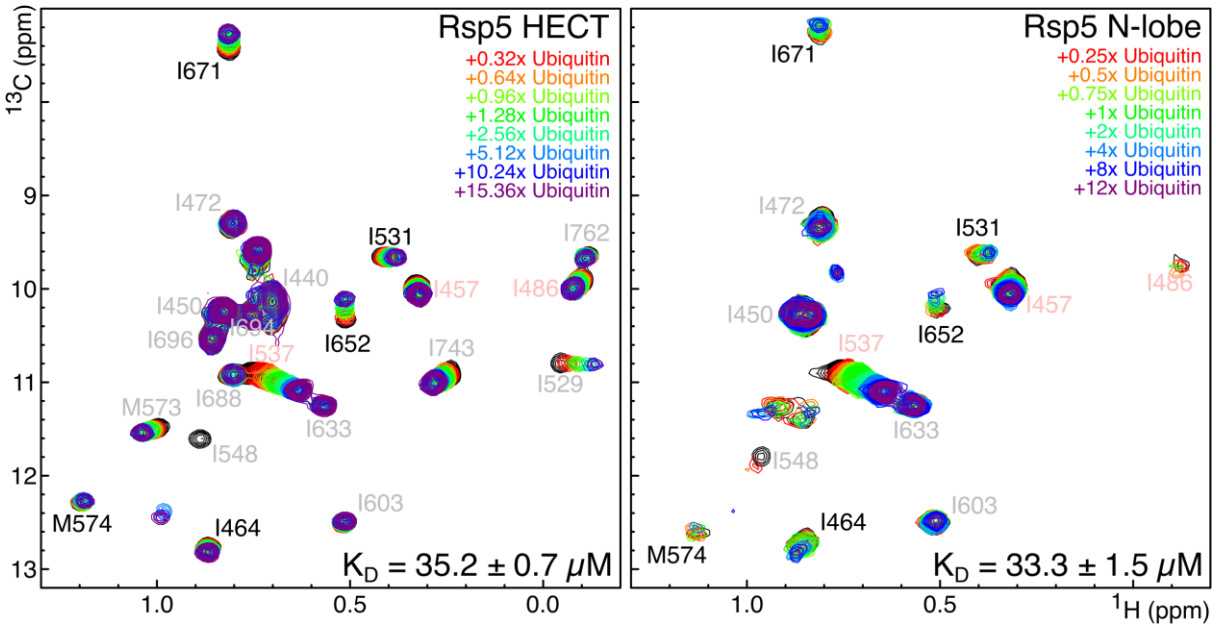


Figure 19. Spectra of Isoleucine and Methionine ^{13}C methyl labelled Rsp5 HECT (left) and N-lobe (right), both titrated with the indicated molar excess amounts of Ubiquitin. Black labelled residues exhibited the tightest binding and were used to fit the K_Ds of both titrations using the program TITAN[206]. (Note: the error given represents the uncertainty of the fit, the actual error in the determined values (due to typical concentration determination and pipetting imprecision) is plausibly larger by an order of magnitude.) Salmon labelled residues show binding with slightly slower saturating behaviour ($K_D \approx 55$ μM). Grey labelled residues were not included either due to the absence of CSP or much slower saturating behaviour. Neither salmon nor grey labelled residues were included in the fit as they might not report directly on ubiquitin binding, but on structural changes induced by it.

Interestingly, we observed that Rsp5 is able to form free ubiquitin chains, while Huwe1 does not (Figure 20). To investigate whether this behaviour might correlate with UBS accessibility, we further tested Smurf2, which, like Huwe1, also lacks the ability to produce ubiquitin chains not conjugated to a substrate.

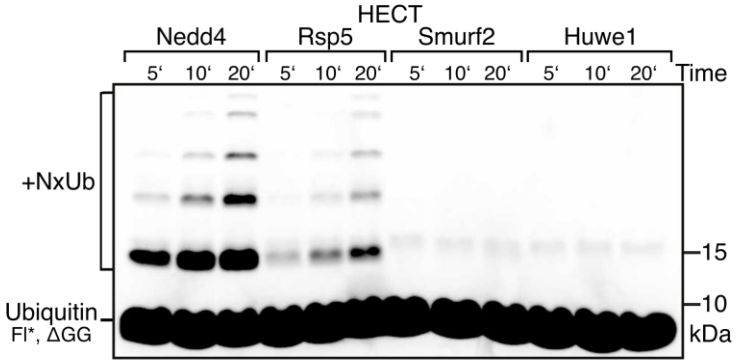


Figure 20. Panel assaying the ability of various HECT domains to synthesize free-standing Ubiquitin chains. 0.5μM of the respective HECT domain incubated with 12 μM Ubiquitin Fl,ΔGG (ΔGG: terminal diglycine removed (makes it unable to act as donor-Ubiquitin), FL*: Fluorescent Labelling), 120 μM Ubiquitin (native wild-type), 1 μM UbcH7 or Ubc4(for Rsp5) as E2 and 1 μM hUbe1 as E1.*

A previous member of the lab, Dr. Natalia Ruetalo, had already determined the K_D of ubiquitin binding to the HECT domain of Smurf2 to be 393 ± 21 μM[147]. To compare against that, an additional ubiquitin titration onto the Smurf2 N-lobe was performed, resulting in a K_D of 191 ± 2 μM (Figure 21). While this is only a twofold difference, it might still contribute to Smurf2 activity regulation and therefore be of biological relevance.

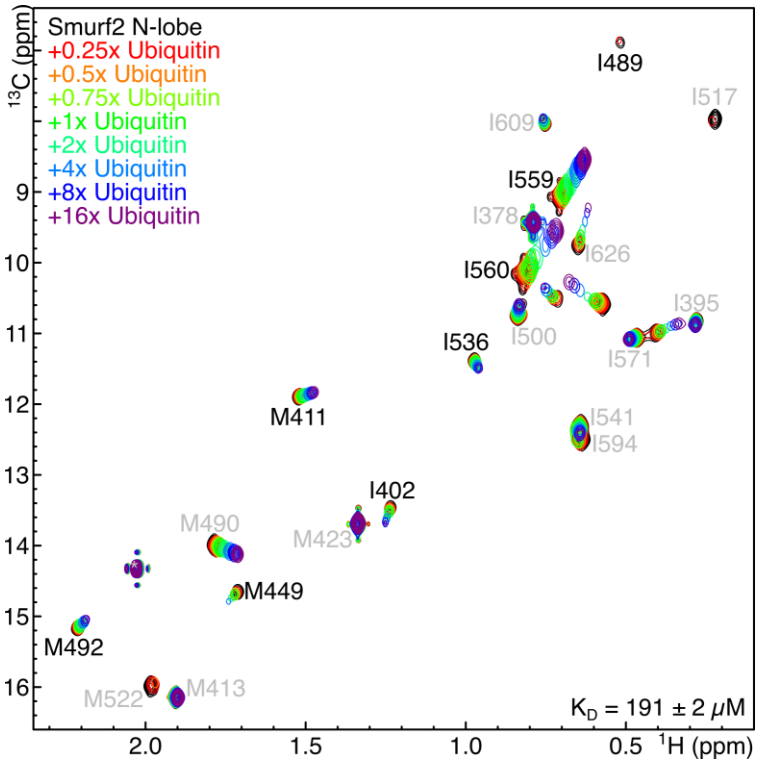


Figure 21. Spectra of Isoleucine and Methionine ^{13}C methyl labelled Smurf2 N-lobe titrated with the indicated molar excess amounts of ubiquitin. Black labelled residues correspond to residues originally used to determine the K_D for the Smurf2 HECT domain as published in Ruetalo et. al. 2019[147] and were therefore used to fit the K_D for the N-lobe construct using the program TITAN[205].

To test this, we generated a Smurf2 HECT construct mutated in residue E623, which is structurally equivalent to E259 in Huwe1, the residue whose mutation has the highest impact on enzymatic activity. Both the glutamate residue and its hydrogen-bonding partner (S315 in Huwe1, S678 in Smurf2) are conserved (Figure 22a). The effect of Smurf2 E623A mutant on the chemical shifts in the spectra was not nearly as clear as that of open state-inducing mutants in Huwe1 (see e.g., Figure 6). This could be due to the fact that the difference between the chemical shifts of the signals in the HECT domain versus the individual lobes is also significantly smaller in Smurf2 than in Huwe1 (Figure 22c).

We then compared the autoubiquitination activity of wild-type Smurf2 and of the E623A mutant, which showed a modest activation effect (Figure 22b), roughly corresponding to the K_D difference of ubiquitin binding to Smurf2 HECT and to the Smurf2 N-lobe.

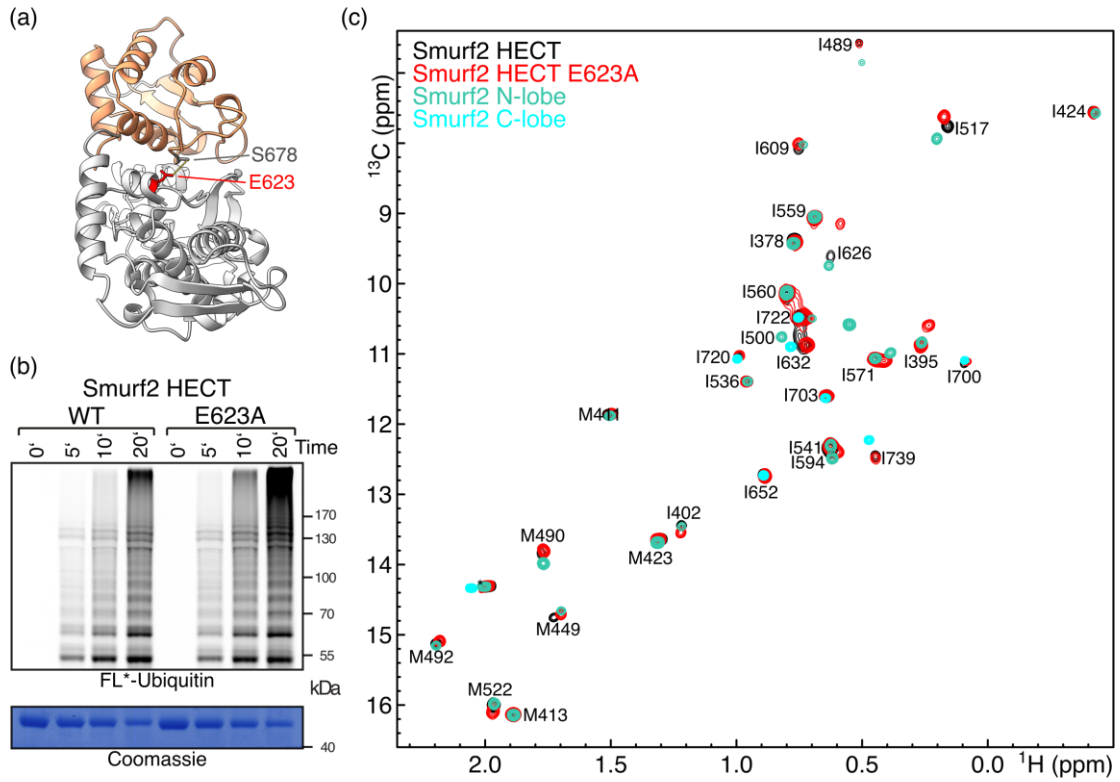


Figure 22. (a) Smurf2 C-lobe and N-lobes (PDB ID: 1ZVD) positioned based on structural alignment to the Huwe1 HECT structure (PDB: 5LP8). (c) ^{13}C-IM spectra of Smurf2 HECT WT and E623A, as well as individual N- and C-lobes. (b) Assay comparing autoubiquitination activities of the Smurf2 HECT wild-type and the E623A mutant. Assays were performed using 65 µM fluorescently-labelled ubiquitin, 1 µM Uba1 (E1), 10 µM UbcH7 (E2), 5 µM H6-Smurf2^{HECT} WT/E623A (E3).

6. Discussion

Here we show that the ubiquitin-binding exosite is indeed conserved widely amongst HECT domain E3 family members, and we confirm its presence in Huwe1, for which ubiquitin binding by the HECT domain had not been reported previously. We found that, in wild-type Huwe1, the catalytically important C-/N-lobe interface (Chapter B) impairs the UBS, preventing ubiquitin binding most likely through allosteric conformational changes. Using a set of mutants (Chapter B.4.) which weaken this interface and shift the equilibrium towards a more open conformation, we were indeed able to detect ubiquitin binding to the HECT domain, and mapped the location of the interaction interface to a region on the N-lobe which corresponds to the typical, conserved ubiquitin-binding-(exo)site of HECT family E3 ligases. Furthermore, mutating key residues in this interface reduces autoubiquitination activity, confirming the functional importance of the Huwe1 UBS.

The UBS was previously proposed to act as a “processivity site” that would, with micromolar binding affinity, stabilize the interaction of substrates with ubiquitin already conjugated, and therefore enhance their polyubiquitination. With the observed negative effect of UBS mutants on E3 thioester formation in Huwe1 depending on the concentration of free ubiquitin (Figure 18) and the fact that residues that report on the open/closed state equilibrium are clearly influenced by ubiquitin — UBS binding (Figure 17a), our results point towards a mechanism where ubiquitin binding to the UBS transiently stabilizes the open-state of the HECT domain (Figure 23 (6), right), obviating the need to open the interface again for the next cycle of ubiquitin thioester transfer from the E2 (Figure 23 (7)

instead of (6 → 8 → 1 → 2)). Of note, this effect could also be exerted by ubiquitin moieties in a growing ubiquitin chain that would keep the UBS occupied.

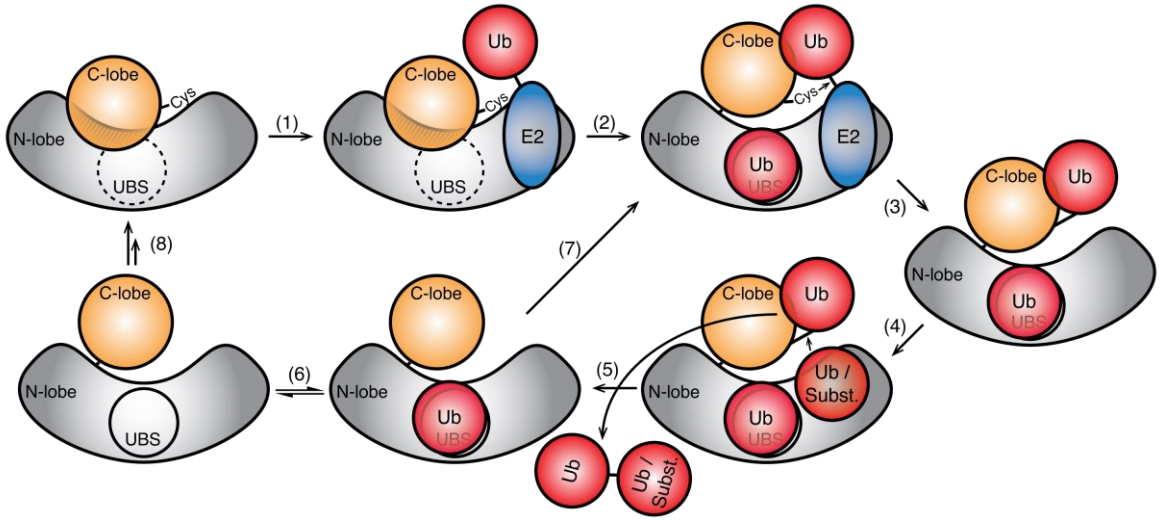


Figure 23. Proposed Huwe1 HECT reaction scheme. The C-/N-lobe interface opens for HECT ~ ubiquitin thioester formation and thereby the UBS becomes accessible (2). After conjugation of ubiquitin to a substrate, UBS occupancy either by monomeric ubiquitin or ubiquitin moieties in a growing ubiquitin chain has a stabilizing effect on the C-/N-lobe interface open-state (6). This could explain the positive effect on thioester formation observed for the Huwe1 HECT domain in the presence of higher amounts of Ubiquitin.

The influence of the UBS on thioester formation was observed at a ubiquitin concentration of 60 µM. It is entirely conceivable that this effect would be physiologically relevant as the cellular concentration of ubiquitin is within this range in different cell / tissue types[208] (e.g. a cellular concentration of ~85 µM in HEK293, ~30 µM in Human frontal cortex, ~20 µM in Mouse brain[208]), and is known to be affected by cellular stress conditions[209]. While the absolute level of ubiquitin in the cell will influence all ubiquitination enzymes, Huwe1 could nevertheless exhibit a stronger response to ubiquitin levels in comparison to other E3 enzymes. This hypothesis is consistent with our observations that in other HECT E3 ligases the C-/N-lobe interface, if present, does not inhibit ubiquitin — UBS binding to a degree comparable with Huwe1. It is conceivable that Huwe1 would not be the only HECT domain E3 with this mechanism. For example, in WWP1, the mutation of a conserved C-/N-lobe interface residue, E798V, has been described in prostate cancer[144]. Of course, ubiquitin – UBS binding could not be the only factor influencing the open-closed state equilibrium in Huwe1 and other HECT E3s.

Additional evidence supporting such a model comes from experiments using Ubiquitin Variant Probes (UbVs), ubiquitin variants specifically modified for improved HECT binding. It was observed that these UbVs can have both positive and negative regulatory functions depending on the HECT domain, despite binding in all cases to the UBS[112]. For example, UbV binding to WWP1 activated E3 ubiquitin thioester formation even in a construct of the HECT domain alone, whereas binding to Nedd4 caused a significant activating effect only in a full-length construct, where it presumably acts by displacing the autoinhibitory C2-WW1 module. This different behaviour could indicate an interplay between HECT inter-lobe interactions and UBS in WWP1, similar to what we observed in Huwe1, but may be absent in Nedd4 or the other Nedd4-family members investigated here that seem to inherently adopt an open HECT domain conformation (Chapter C).

D. The native N-terminus of ubiquitin prevents HECT-mediated synthesis of linear ubiquitin chains

1. Introduction

Post-translational modification of substrates by ubiquitination can take place in various forms, which only have in common the attachment of at least one ubiquitin molecule to the amino group of a lysine of a substrate or of another ubiquitin. The typical (i.e. first identified) way in which ubiquitin chains are formed is by the conjugation of one ubiquitin to lysine 48 or lysine 63 of another ubiquitin[89]. Since then, the formation of chains linked by any of the 7 lysines in ubiquitin have also been described[89] as *atypical ubiquitination*. This *atypical ubiquitination* also includes the branching of chains (switching linkage) and the conjugation to the (free) α-amino(N-) terminus of ubiquitin[210]. Of note, N-terminal ubiquitination is also known as a way by which the first ubiquitin subunit is attached to certain substrates[211].

The way in which ubiquitin is attached to target substrates and further forms chains of different linkages is now typically referred to as the ubiquitin code[41,212] in analogy to the histone code[213] (which includes ubiquitination amongst other post-translational modifications) or the genetic code. In this concept, ubiquitin ligases (E3s) are the writers of the code, which can then be edited and/or removed by deubiquitinating enzymes (DUBs). Both E3s as well as DUBs are often chain specific, i.e., they preferentially generate or remove certain chain types (see e.g., Chapter A, where we showed that Huwe1 generates mostly K48-linked chains). The ubiquitin code ultimately determines the effect of ubiquitination on the substrate. As ubiquitination is involved in the regulation of many cellular processes, it is of fundamental importance to determine the pattern / code of ubiquitination that a substrate protein receives under certain conditions, the kind of ubiquitination that a certain ubiquitin ligase imparts to its substrates, and also the way in which deubiquitinating enzymes modify this pattern.

Over the years, several approaches to gain information on how / when / where ubiquitination happens have been developed and extended to explore all the features of the ubiquitin code.

One of the classic techniques is immunoblotting (see Emmerich and Cohen, 2015[214] for an overview). Substrates can typically be isolated via immunoprecipitation or other affinity-based methods, after which the presence of conjugated ubiquitin can be detected by immunoblotting with an anti-ubiquitin antibody, or also an antibody for the substrate itself, as ubiquitin(chain) conjugation can be inferred from characteristic changes in size. Originally released generic poly-ubiquitin antibodies have varying specificities for the different types of chain linkages. In addition, since the idea of the ubiquitin code has gained prominence, antibodies against single chain linkage types of varying specificity have been made commercially available[214].

More recently, approaches based on the inherent chain linkages specificities of deubiquitinating enzymes, which also allow for better insight into branched chains, have been developed[215]. Furthermore, mass spectrometric methods have seen increased application in the discrimination of chain linkages (see e.g., Chapter A, Figure 1(b)). Tryptic digestion of ubiquitin results in a C-terminal di-glycine fragment, which remains isopeptide-bonded to lysine-containing fragments. These conjugates can be quantified, in a linkage-specific manner, using a set of peptide standards[178].

Another notable and easily accessible approach used especially *in vitro* but also *in vivo*, consists in creating ubiquitin variants where lysine residues are mutated to arginine. There are multiple ways in which this can be used to gain information on the chain linkage being generated. By replacing in turn each one of the 7 lysines with arginine, and comparing chain formation activity, one can infer which chains linkages are being generated by the decrease in activity resulting from mutating the respective

lysine. Conversely, one can generate ubiquitin where all lysines are mutated with arginine, except for one; poly-ubiquitination activity will only be observed for the construct(s) that retains the lysine forming the respective chain linkage(s) (see e.g., Chapter A, Figure 1(a)). Additionally, a completely lysine-free ubiquitin, where all lysines are mutated to arginine, can be used for example in *in vitro* experiments to enforce a single conjugation event to a substrate or ubiquitin-chain lysine. This approach has been used for instance to differentiate between multi-mono-ubiquitination and poly-ubiquitination of substrates. One limitation of this method is that it cannot directly be applied to study systems where an N-terminal linkage is involved. However, there is only a single ubiquitin ligase known to date that assembles linear chains, the LUBAC complex[63].

In both *in vitro* as well as *in vivo* experiments, it is often desirable to detect ubiquitin with higher sensitivity than with anti-ubiquitin antibodies; in many cases it is also necessary to differentiate between the different kinds of lysine to arginine mutants, or also between them and endogenous ubiquitin. Approaches where ubiquitin is tagged with well-known epitopes, recognized by commonly available and high affinity-binding antibodies have thus been developed[216]. Since the ubiquitin C-terminus is necessary for its conjugation, the N-terminus is used to attach such tags.

In this chapter, I will present an investigation into the use of N-terminally epitope-tagged ubiquitin for HECT domain E3 ubiquitin ligase ubiquitination assays and why their use should be avoided in any but the most basic experiments, especially if linkage specificity is of interest.

2. *All typical Antibody Epitope-Tags are Ubiquitinated*

For the experiments that will be described in Chapter E we required ubiquitin that cannot be further elongated by another ubiquitin. For this we intended to use a ubiquitin where all 7 native lysines are mutated to arginine (human ubiquitin K0, hUbK0). To detect its conjugation to our substrate of interest, or a growing chain on it, we had initially planned to use immunoblotting. Therefore, to distinguish it from native ubiquitin, we tagged it with an immuno-epitope tag, which could be easily detected by western blotting. To rule out any influence of the tag on the utilization of the ubiquitin K0 by our enzyme cascade, we tested different epitopes and compared them in an auto-ubiquitination assay in the presence of several K63 linkage-specific HECT E3 ligases (Table I 3). We expected to see multiple bands generated by multiple mono-ubiquitination events on the surface-exposed lysines of the HECT domain, our substrate. However, we observed a pattern of multiple bands that cannot be explained by mono-ubiquitination of the HECT domain, as the increase in molecular weight (as determined by SDS-PAGE) is higher than the one expected in the case of all lysines being mono-ubiquitinated (Figure 24(a)).

Besides autoubiquitination, we tested the reactivity of the tagged-ubiquitin moieties using an artificial model substrate for the WW domains of Rsp5 (GB1-Rim8ct, see Chapter E). This minimal substrate contains only one lysine and thus allows for only a single round of mono-ubiquitination. Initially designed based on a peptide recognized by the three WW domain module of Rsp5, here we could show that this model substrate is not only ubiquitinated by $Rsp5^{WW3\text{-}HECT}$, but also by $Smurf2^{WW3\text{-}HECT}$ as well as $Nedd4^{WW4\text{-}HECT}$. For detection, we labelled the model substrate at its single cysteine using the thiol reactive fluorescent dye Texas Red Maleimide. Using this substrate, we observed in each assay multiple ubiquitination events, although only one should be possible (Figure 24(b)).

In addition to autoubiquitination and substrate ubiquitination, a limited subset of the HECT domains is able to assemble free ubiquitin chains, unanchored to a substrate, starting from one ubiquitin initially serving as pseudo-substrate in the chain. To observe this, we used ubiquitin containing all 7 lysines, but missing the two glycines at its absolute C-terminus. This way, it can act as an acceptor, but not as a donor for chain elongation. To detect this ubiquitin, we labelled it using Iodoacetamido-Fluorescein at a cysteine introduced directly before its starting methionine. Also in this case, multiple ubiquitination events could be seen; however, these could also be explained as multiple mono-ubiquitination events at the different lysines of the ΔGG ubiquitin, as the used HECT domains are known to assemble chains linked by a different lysine to a lesser degree.

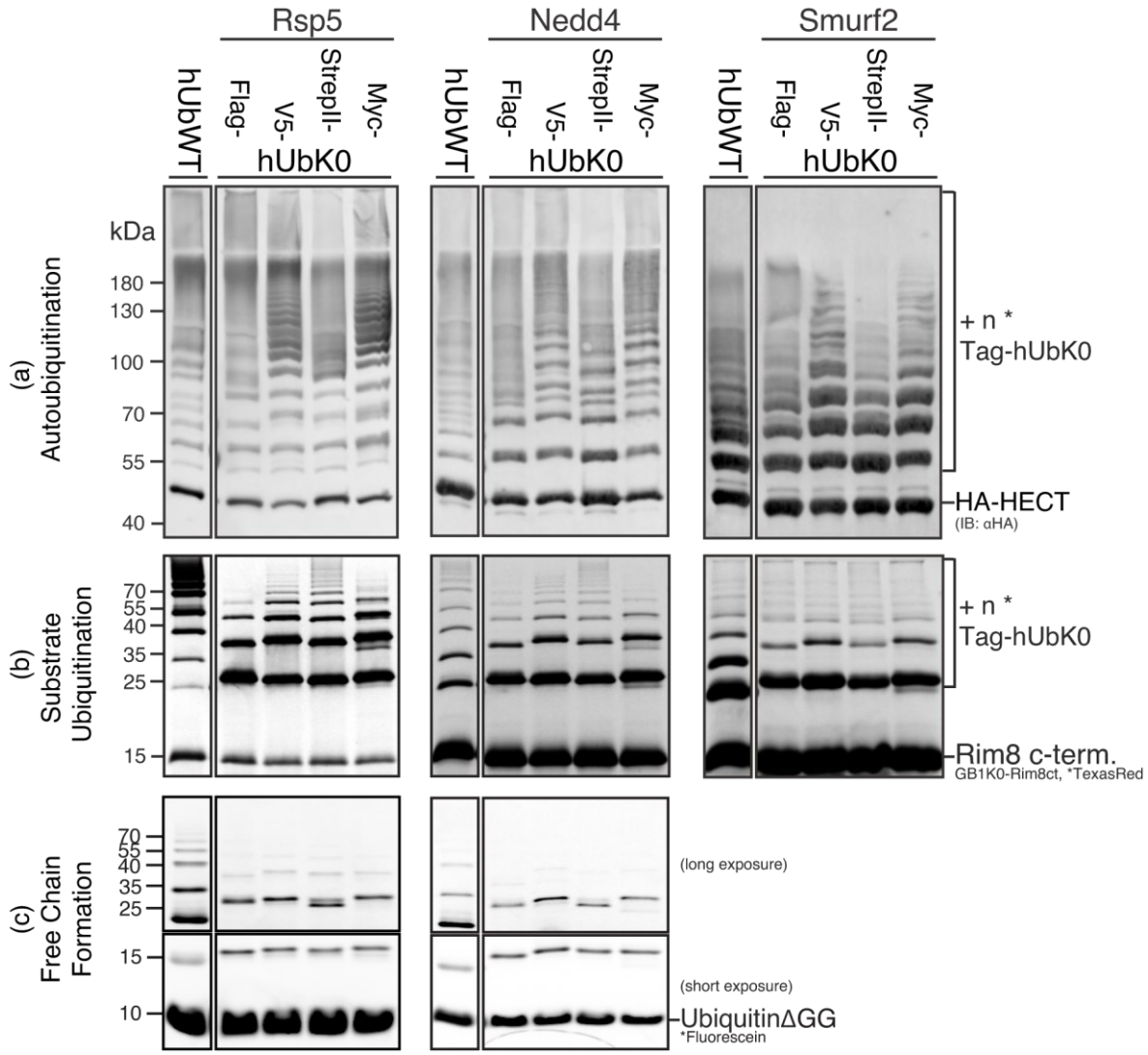


Figure 24. Ubiquitination reactions using wild-type (lane spatially separated on the SDS-PAGE, but using the same reaction mix and conditions) and epitope-tagged, lysine-free ubiquitins, and catalysed by (a, c) the HECT domains of Rsp5, Nedd4 and Smurf2 or (b) a construct consisting of the WW domain preceding the HECT domain and the HECT domain itself. Samples were resolved on SDS-PAGE and detected using (a) anti-HA antibody against the HA tagged HECT domains; (b) fluorescently labelled (TexasRed) substrate GB1-Rim8 C-terminus. (c) fluorescently labelled (Fluorescein) ubiquitin missing the C-terminal di-glycine motif (ΔGG) (untagged). Respective assay conditions can be found in Table S 5.

3. *Chain elongation of N-terminally extended Ubiquitins is extension length-dependent*

Having observed that tagged ubiquitin K0, despite having none of the 7 lysines available for conjugation, is able to form ubiquitin chains, we suspected that the lysines present in the epitope tags sequence may serve as ubiquitin acceptor sites. Therefore, we sought to express and purify an untagged ubiquitin K0.

For this, we cloned ubiquitin K0 into the multiple cloning sequence (MCS, between the NcoI and BamHI restriction sites) of a vector (pETM-11) bearing a His6 tag and a TEV protease cleavage site upstream of the MCS. After TEV-mediated cleavage of the expressed fusion protein, additional 4 amino acid residues (Gly-Ala-Met-Gly, GAMG) are left before the natural start methionine (M1) of ubiquitin K0. These additional residues still allowed Rsp5, Nedd4 and Smurf2 to assemble chains in autoubiquitination-, substrate ubiquitination- and unanchored chain formation-based assay. Since this

extension at the N-terminus of ubiquitin K0 contains no lysines, poly-ubiquitin conjugation must be mediated by the free N-terminal amine in this case.

To test if this poly-ubiquitination is influenced by the sequence and / or length of the extension, we cloned ubiquitin K0 constructs with one to three alanines preceding the native N-terminal methionine. For Rsp5, the ability to form ubiquitin chains was clearly dependent on the length of the N-terminal extension, with three alanines showing lower chain formation activity than the four amino acid artefact sequence, and the two and one alanines successively displaying even lower activity. The other tested HECT domains showed a higher decrease in activity from the GAMG sequence to the three alanines, suggesting some influence of the sequence of the N-terminal extension in spurious ubiquitination. While Nedd4 showed some activity in the presence of the ubiquitin K0 construct with three preceding alanines, no significant chain formation activity was detected for Smurf2 for any of the alanine constructs.

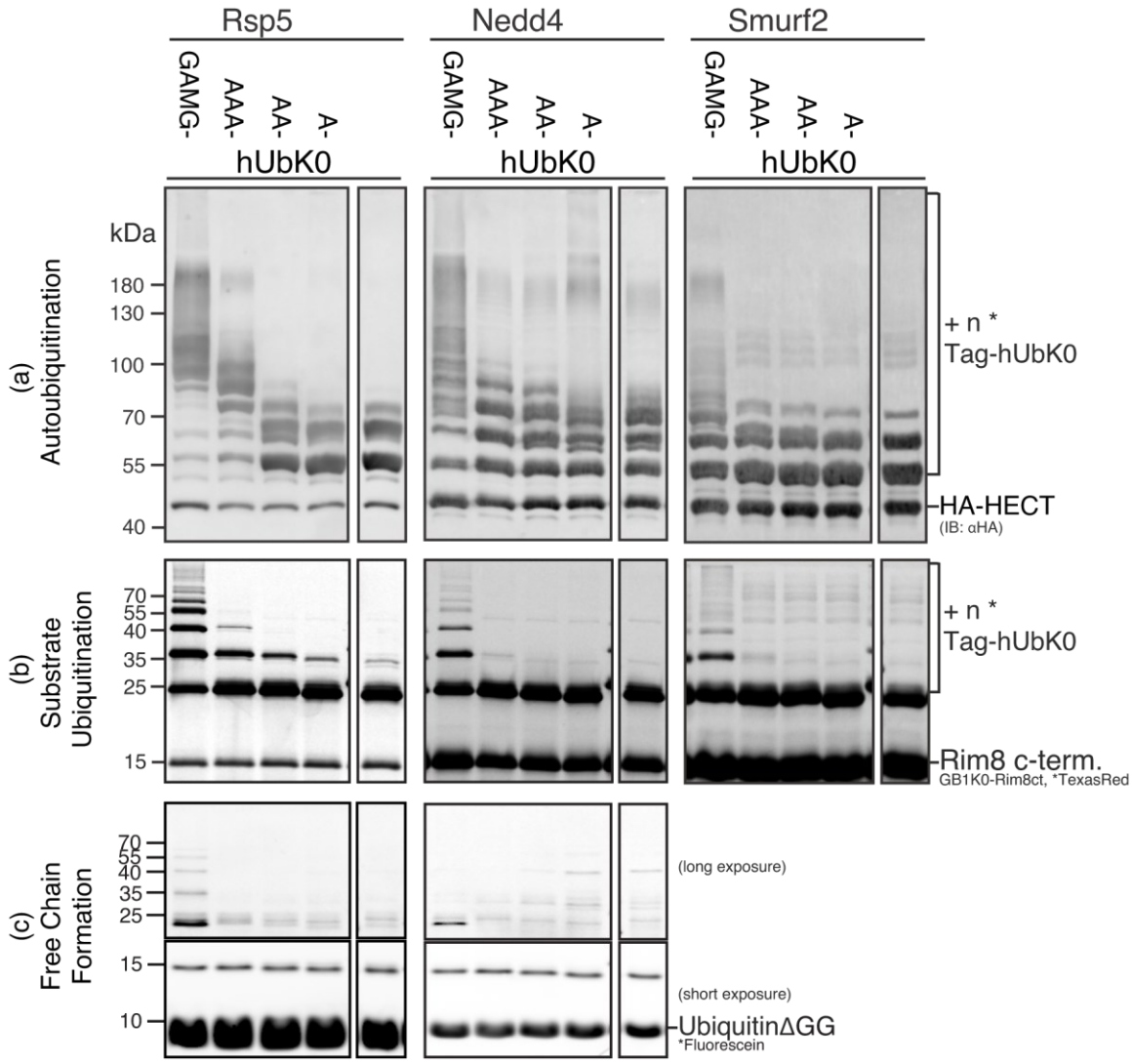


Figure 25. Ubiquitination reactions using lysine-free ubiquitin (lane spatially separated on the SDS-PAGE, but using the same reaction mix and conditions) as well as with N-terminal additions of the amino acids GAMG or the indicated number of alanines. Reactions were catalysed using (a, c) the HECT domains of Rsp5, Nedd4 and Smurf2 or (b) a construct consisting of the WW domain preceding the HECT domain and the HECT domain itself. Samples were resolved using SDS-PAGE and detected using (a) anti-HA antibody against the HA tagged HECT domains, (b) fluorescently labelled (TexasRed) substrate GB1K0-Rim8 C-terminus, (c) fluorescently labelled (fluorescein) ubiquitin missing the C-terminal di-glycine motif (ΔGG) (untagged). Respective assay conditions (identical to Figure 24) can be found in Table S 5.

A powerful method to investigate ubiquitin linkage uses mass spectrometry (MS) after tryptic digestion of the sample. Trypsin cleaves mainly after basic amino acids, thus producing a fragment of ubiquitin (which ends with the conserved sequence Arg-Gly-Gly) containing its C-terminal di-glycine. This C-terminal fragment will then be detected as a modification of lysines or N-terminal amines of ubiquitinated substrates. When we analysed ubiquitination reaction samples performed using epitope-tagged lysine-free ubiquitins (as in Figure 24), as well as lysine-free ubiquitin N-terminally extended with the GAMG sequence (as in Figure 25), we detected di-glycinated fragments in all of them, consistent with ubiquitin modification of the substrate.

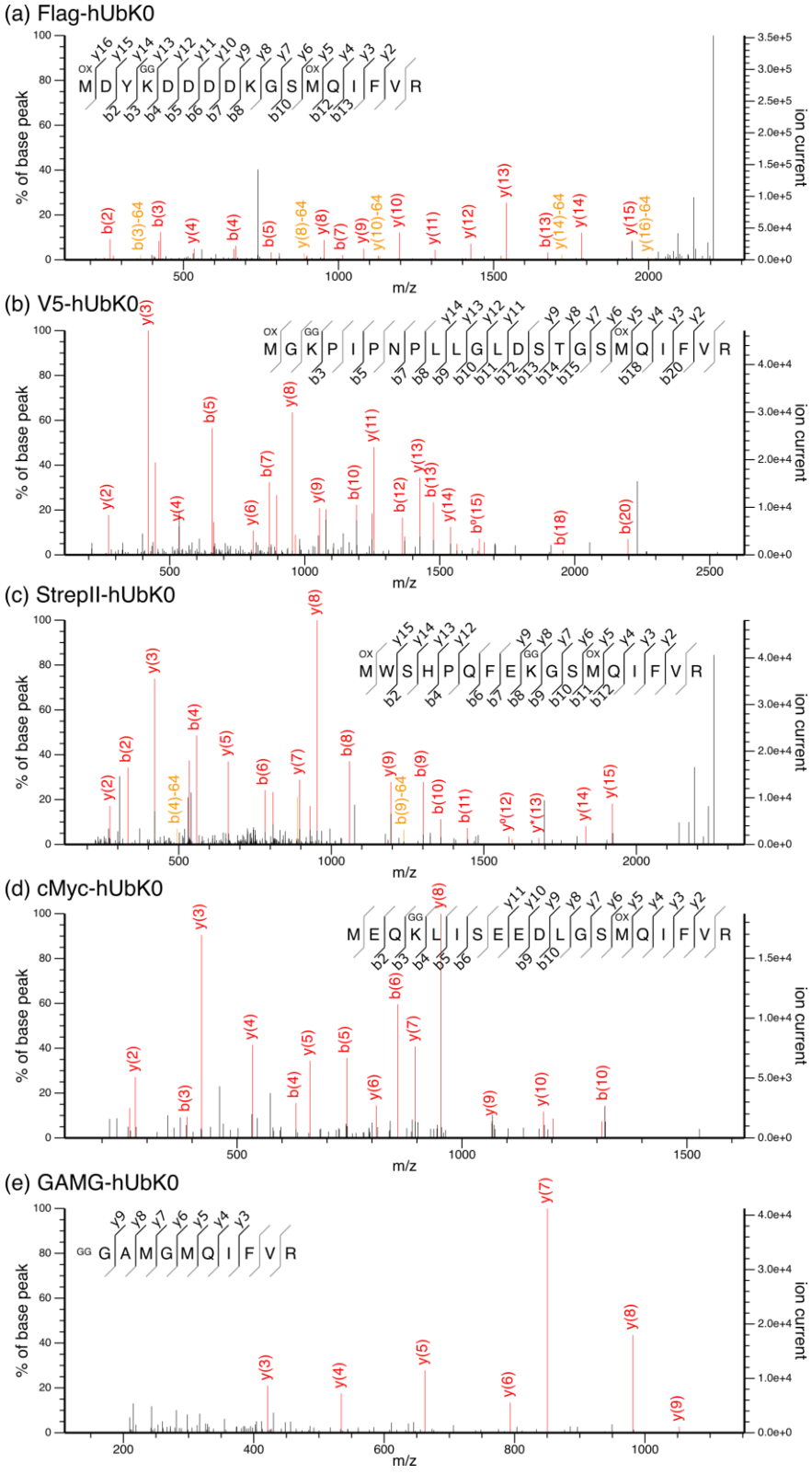


Figure 26. MS/MS fragmentation of trypsinized reaction mixtures as generated in Figure 1(b) and Figure 2(b). For the case of tagged lysine-free ubiquitins, fragments consistent with ubiquitin attachment to tag lysines were identified for all except the V5-tag, where the fragments found did not allow differentiation between the N-terminus and the tag-lysines. For the GAMG-extended lysine-free ubiquitin, fragments consistent with ubiquitin attachment to the N-terminus were found. (Trypsinization performed and spectra recorded by Dr. Astrid Bruckmann, head of the University of Regensburg MS facility.)

In the case of epitope-tagged ubiquitins, MS/MS fragmentation detected fragments consistent with ubiquitin attachment to the lysines present in the tag, whereas for the GAMG-extended ubiquitin these fragments were consistent with ubiquitin attachment to the N-terminus (Figure 26).

4. Discussion

In contrast to other ubiquitin-like molecules such as Sumo, the very N-terminus is already part of the ubiquitin secondary structure. In this chapter, we found that N-terminal extension of ubiquitin increases the chain formation reactivity. We show here that epitope tags containing lysine residues are used as additional acceptor sites in ubiquitin and thus enable ubiquitin chain formation, even if the ubiquitin itself does not contain a lysine (K0 mutant).

When the use of epitope-tagged ubiquitin was introduced as early as 1991 (by M. Hochstrasser[216]), it was immediately observed that, while ubiquitin was conjugated to a substrate protein and formed chains, it did not trigger the expected degradation of the target. At that time, the existence of different regulatory outcomes determined by the type of chain specificity in poly-ubiquitination had not yet been described. With the information presented in this chapter, it is now reasonable to propose that the lack of target degradation observed in the 1991 study was due to ubiquitin being linked to a lysine in the epitope tag (Myc-tag, also investigated here), instead of K48 (proteasomal degradation signal) or K63 (lysosomal degradation signal). This possible interference with the regulatory effect of ubiquitination makes it hard to justify the use of tagged-ubiquitins in *in vivo* experiments. In addition to the modification of ubiquitin itself, the N-terminal epitope tagging of ubiquitination substrates might be similarly problematic, and has been found in some cases to perturb their native-ubiquitination[211,217]. Having shown that the native N-terminus of ubiquitin is essential for its specific conjugation, we can conclude that special care also needs to be taken in studies where the chain type specificity is to be investigated. Since the typically flexible epitope tags could bridge the ~10-30 Å distance from the N-terminus to the lysines of ubiquitin, this could skew the population of chain linkage types produced in favour of lysines further away from the epitope tag.

Naturally, this also complicates the readout of assays which try to differentiate between polyubiquitination versus mono-ubiquitination at multiple sites using epitope-tagged lysine-free ubiquitin or even in-vivo assays where epitope-tagged lysine-free ubiquitin is used in an attempt to induce mono-ubiquitin to study mono-ubiquitin phenotypes.

The most remarkable observation in these assays and the subsequent mass spectrometry analyses (Figure 26) is, however, that HECT E3 ligases are capable of forming linear ubiquitin chains, albeit not with native ubiquitin. So far, the LUBAC complex was the only E3 known to form linear ubiquitin chains. We demonstrated here that a lysine-less N-terminal extension of K0 ubiquitin still results in chain formation when incubated with HECT E3 ligases. To some degree, this effect seems to depend on the sequence composition of the extension, since we detected a significant increase in activity when comparing a 3 alanine N-terminal extension to the 4 amino acid GAMG sequence (Figure 25). This difference could be explained by the distinct flexibility of the resulting chain. While both these short sequence stretches (AAA and GAMG) are expected to be flexible and not structured, they could still sample different preferred orientations. In particular, the increase in reactivity of the GAMG ubiquitin K0 compared to the AAA ubiquitin K0 (Figure 25), could for example be explained by the increased conformational flexibility provided by the glycine residues, which can sample a significantly larger area of the Ramachandran plot than any other amino acid and exhibit a propensity for the left -handed alpha-helix conformation.

In general, the observed chain formation seems to be independent of the presence of lysines in the epitope tag sequence and can in any case be mediated by the amino terminus (Figure 25). As TEV protease cleavage efficiency was significantly reduced when the cleavage site was directly upstream

of the ubiquitin N-terminus (data not shown), we used a different strategy to express the constructs bearing an N-terminal oligo-alanine extension. These ubiquitin constructs were instead fused to an N-terminal His6-tagged Sumo (a ubiquitin-like protein) that was cleaved using Ulp1, a Sumo-specific protease.

A structural model of TEV protease in complex with its recognition sequence directly fused with ubiquitin reveals extensive clashes (Figure S 18). This suggests that TEV binding to this substrate would be impaired by steric hindrance, and thus provides an explanation for the drop in cleavage efficiency that we observed for our untagged ubiquitin constructs (data not shown). Fusing a TEV recognition motif directly to the N-terminus of proteins is not typically reported to be an issue for TEV cleavage activity. This suggests that the N-terminus of ubiquitin might be generally less accessible. Indeed, when we compared the solvent accessible surface area of the N-terminal amine nitrogen of ubiquitin with the N-terminal amine in other structures of proteins of approximately the same size, we saw that the N-terminus of ubiquitin is in the lower range of the distribution (Figure 27(a)), demonstrating that the N-terminus of ubiquitin is less accessible than the N-terminus of a typical protein. Additionally, in available ubiquitin structures the N-terminal methionine forms multiple hydrogen bonds to neighbouring residues and is already part of the first beta-strand in ubiquitin. Consistent with this, the N-terminal Met in ubiquitin is significantly more rigid than the C-terminal Gly, as indicated by the lower B-factor (e.g., PDB ID: 1D3Z) and the absence of alternative conformation in NMR ensembles (e.g., PDB ID: 1D3Z in Figure 27(b), NMR structure ensemble).

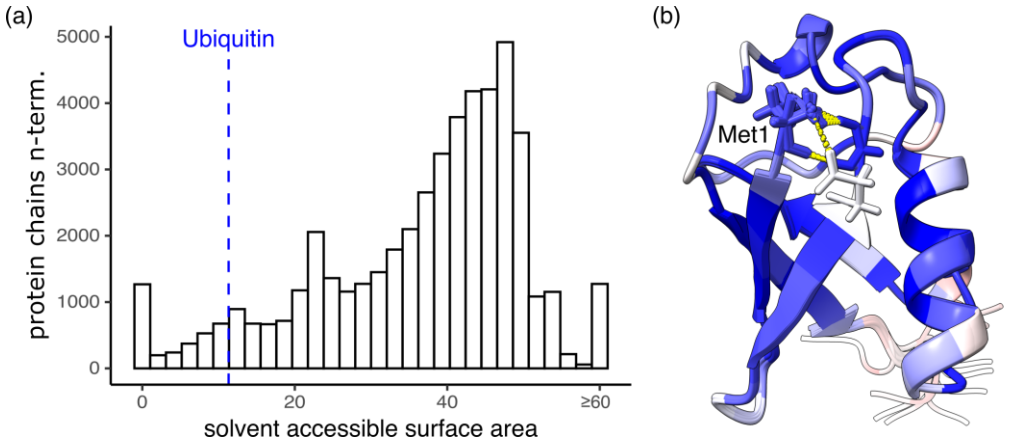


Figure 27. (a) Histogram depicting solvent accessible surface area of the N-terminal amine nitrogens of atomic resolution structure of proteins with similar size to ubiquitin (< 17.2 kDa = 2 x Ubiquitin), as determined using the get_area command in PyMOL[218]. The area of the ubiquitin N-terminus (from PDB ID: 1D3Z.2) is marked with a dotted blue line. (b) An NMR structure ensemble of ubiquitin (PDB ID:1D3Z) coloured by B-factor values in a gradient from blue (low B-factor) to red (high B-factor). Hydrogen bonds to methionine 1 are shown in yellow. Methionine 1 and its hydrogen bond partners are shown as sticks.

The native N-terminus of ubiquitin is thus important to regulate its conjugation with high specificity. Interestingly, while E3 enzymes often exhibit a preference for more than one lysine in ubiquitin (see e.g., Chapter A), LUBAC is the only one known to exclusively assemble M1-linked chains. In contrast to the promiscuity of HECT domain family E3 ligases, which in the assays shown in this chapter were able to assemble chains from N-terminally elongated ubiquitins, LUBAC exhibits an absolute specificity for the native N-terminus of ubiquitin, and the addition of an N-terminal His6-tag[210], or even a single amino acid to the N-terminus[219], have been reported to abolish its activity.

Groups working on LUBAC have suggested moving tags from the N-terminus to an internal loop in ubiquitin itself[68,220], to avoid interfering with M1 ubiquitination. However, care needs to be taken to not interfere with binding to the enzymes in the ubiquitination cascade. Additionally, while applicable

to the study of M1-linked chain formation, the introduction of an exogenous loop might create steric hindrance affecting lysine selection, and therefore be unsuitable for studies focusing on chain-linkage specificity.

In LUBAC strict positioning of both the acceptor as well as the donor ubiquitin and its importance for catalysis and specificity is well understood and could for example be demonstrated using single mutations in either the donor or acceptor ubiquitin interface which fully abolished ubiquitination activity [Stieglitz/Rana 2013. For HECT domains however, the picture is less clear. While we (Chapter A) managed to demonstrate a donor ubiquitin interface, this is conserved independently of the chain linkage preference of the HECT domain and can thus not explain it. Also, the placement of the acceptor ubiquitin in HECT domain ubiquitination has remained entirely elusive so far.

E. Rsp5 chain elongation: Establishing a model system to investigate ubiquitin chain elongation kinetics

1. Introduction

It is long known that the canonical signal for proteasomal recognition is a chain of at least four ubiquitins[43,44,221]. Early work on elucidating the mechanisms of ubiquitin ligases proposed multiple models for how these chains can be attached to substrates. Since ubiquitination requires a cascade of enzymes comprising the E1 ubiquitin activating enzyme, an E2 ubiquitin conjugating enzyme and finally a E3 ubiquitin ligase, multiple models, involving one or more components of the cascade were originally proposed[24].

From the observation that shortening the C-/N-lobe linker abolished activity, it was for example hypothesized that an E2-E3 heterodimer could transfer ubiquitin to a growing chain attached to the E3 catalytic cysteine(Figure I 2(b,1)), and that the linker between the two E3 lobes displaces the C-lobe to make space for the growing ubiquitin chain[26]. Other initially proposed mechanisms include a model where a homodimer of E2 pre-assembles a ubiquitin chain before transferring it to an E3 in a seesaw manner(Figure I 2(d))[24] or where an E3 homodimer preassembles a ubiquitin chain on the catalytic cysteine of one of the monomers(Figure I 2(b,2))[25]. In all these initial models, a pre-assembled ubiquitin chain is transferred to a substrate *en bloc*. This mechanism was proposed to explain the apparent processivity of ubiquitin ligases, whereby, in typical assays, chains significantly longer than the 4 subunits required for robust proteasome binding were attached to some of the substrates before every substrate protein was conjugated to at least a single ubiquitin.

While several HECT domains are able to produce free ubiquitin chains, seemingly supporting an *en bloc* mechanism, in fact the synthesis of these free chains occurs at a significantly lower rate than both auto- and substrate-ubiquitination, and generates mostly di-ubiquitin (see e.g. Chapter D, Figure 24(c) and other studies [123]). This, together with the evidence supporting formation of a thioester between the HECT domain and a single ubiquitin (see Chapter A, B and C figures for examples, as well as [120]), contrasts the notion that the *en bloc* transfer contributes significantly to the apparent processivity of ubiquitin chain formation by HECT E3 ubiquitin ligases.

The alternative to the *en bloc* mechanism is a model where ubiquitin is transferred sequentially (Figure I 2(a)), with a single ubiquitin being activated by the E1, transferred from the E1 to the E2, then from the E2 to the E3, and finally from the E3 to a substrate or a growing ubiquitin chain on the substrate. This sequential model is now regarded as the most likely mechanism for HECT E3 ligases for substrate ubiquitination.

Since a sequential mechanism does not involve the pre-assembly of ubiquitin chains, apparent processivity and control of chain length must have so far unknown underlying mechanisms. One possibility could be an active mechanism, where the HECT domain actively recognizes the acceptor ubiquitin among all ubiquitin moieties in the growing ubiquitin chain. This could be driven for example by a higher binding affinity of the terminal ubiquitin to the site of conjugation on the HECT domain or by dynamics in a ratchet-like mechanism (Figure 28(a)). Another model, first proposed by the group of S. Polo[98], states that chain elongation might be disfavoured for chains shorter than two ubiquitins, as the conserved ubiquitin binding site (UBS) that is present on the HECT domain N-lobe (see e.g. Chapter C) does not bind the terminal ubiquitin, but the second last ubiquitin in the chain (Figure 28(b)), therefore allowing the terminal ubiquitin in the chain to be more efficiently elongated.

Another possibility is that the kinetics of chain elongation by HECT E3 ligases could just follow a simple stochastic model (Figure 28(c), the “null hypothesis”), where all ubiquitins in the chain can bind the HECT domain equally well. In such a case elongation kinetics for short chains would follow roughly the

ratio of acceptor ubiquitin to total chain length. As there is only a single terminal ubiquitin in non-branched chains - which should be the predominant product, assuming specificity of the HECT domain for a single ubiquitin chain linkage type - the rate of elongation by one ubiquitin would be 1/N, with N being the chain length. The binding, and therefore the elongation could additionally be modulated by steric effects involving the intermediate ubiquitins in the chain attached to the bound substate and the terminal acceptor ubiquitin of the same chain.

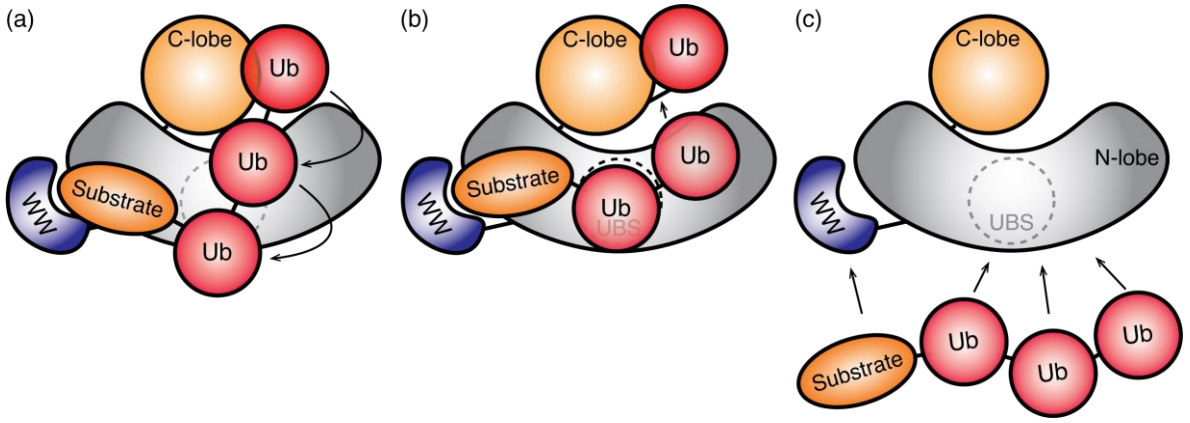


Figure 28. HECT domain features that may influence chain elongation kinetics: (a) The HECT domain could bind the acceptor ubiquitin in a chain with higher affinity than for other ubiquitins in a chain. (b) The ubiquitin binding site on the HECT domain N-lobe could bind the ubiquitin preceding the acceptor ubiquitin in the chain thereby placing the acceptor ubiquitin in spatial proximity of the catalytic cysteine. (c) Ubiquitin (and substrate) binding as a purely stochastic process dominating chain elongation kinetics.

To determine whether the HECT domain shows inherently different binding affinity for ubiquitins in different positions in a chain, a previous master student in the lab, Matthias Schwer, measured the individual affinity of each ubiquitin in a di-ubiquitin (the shortest ubiquitin "chain") to the Rsp5 HECT domain. While he found no significant difference in affinity for the two ubiquitins in this minimal setup, substrate binding, donor-ubiquitin presence, or longer ubiquitin chain length were not tested for.

Therefore, we sought to establish a more complete system. For this, we decided to measure individual reaction rates of for a stepwise ubiquitination process where a model system is extended by exactly one ubiquitin at a time by adding all necessary components for substrate ubiquitination.

2. *The components of the model system*

To measure the kinetics of chain elongation for each individual elongation step, we planned to start with chains of defined length and to measure their elongation by exactly one ubiquitin.

Having mostly observed ubiquitination activity through auto-ubiquitination - which requires only the HECT domain itself – the goal here was to observe substrate ubiquitination, since substrate binding to the E3 ligase might influence the kinetic of chain elongation. Since in Nedd4 family HECT E3 ubiquitin ligases the WW domains are the main substrate interaction platform, we extended the HECT domain construct of Rsp5, the *Saccharomyces cerevisiae* homologue of Nedd4, to include the WW3 domain N-terminal of the HECT domain.

As substrate, we chose the C-terminus of Rim8 (Rim8ct, residues 497-526), a known substrate of Rsp5[204]. Rim8ct represents a minimal substrate, containing not only a WW domain binding motif (PPxY or PxY), but also a lysine known to be ubiquitinated *in vivo*[222]. To avoid unspecific ubiquitination, the other lysines present in the construct, also subject to ubiquitination, albeit at lower levels[204], were

mutated to arginines (K507R, K513R). To allow for robust determination of protein production, concentration determination and visualization in a size range compatible with longer ubiquitin chains upon separation via SDS-PAGE, the Rim8ct sequence was expressed as a fusion to a variant of the B1 domain of the Streptococcal protein G (GB1) in which all lysines had been replaced by arginines (GB1 K0). Since this construct contained no cysteines, we introduced a single cysteine between the substrate peptide to allow for specific fluorescent labelling with iodoacetamide functionalized fluoresceine.

To avoid formation of branched chains, we used a ubiquitin variant with all but one lysine mutated to arginines for the preassembly of chains on the substrate. As Rsp5 is known to assemble predominantly lysine 63 (K63)-linked chains *in vivo*[223], K63-only ubiquitin was chosen. This choice also helps to not convolute the observation with the switch to a slower, more linkage promiscuous mode of chain elongation of very long chains that had been observed in HECT[123] as well as RING[224] family E3s. To be able to observe only a single elongation event of an existing chain, a ubiquitin that could not be further elongated was required. For this, a completely lysine-free variant of ubiquitin, where all lysines are mutated to arginines (K0), was used. All ubiquitin constructs were purified following a strategy without epitope and / or affinity tags to avoid the pitfalls of using ubiquitin variants with a non-native N-terminus (as discussed in Chapter D).

Lastly, ubiquitination requires an E1 ubiquitin activating enzyme as well as an E2 ubiquitin conjugating enzyme, in addition to the E3 ubiquitin ligase. As E1 enzyme, we used the human variant of Uba1 (Ube1), the main E1 enzyme that is highly conserved amongst eukaryotes[225]. As E2 we chose *Saccharomyces cerevisiae* Ubc4, an E2 enzyme already known to interact with Rsp5[226,227].

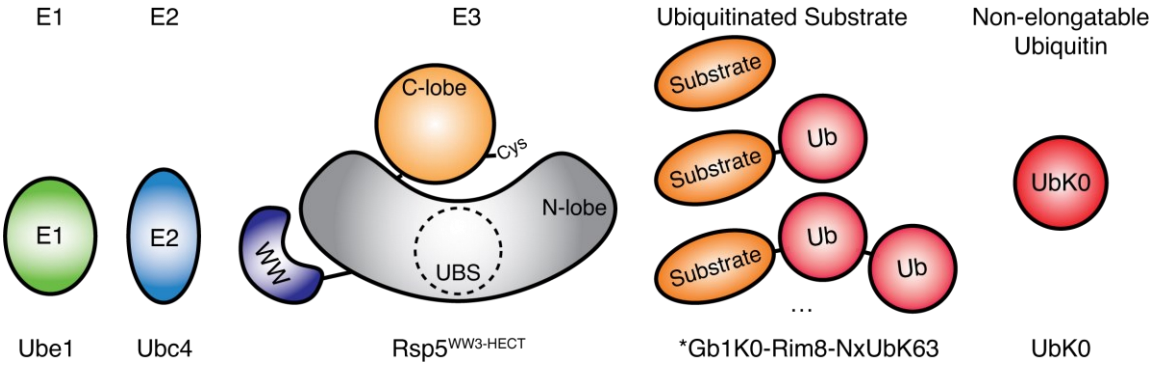


Figure 29. Schematic of components required for an in vitro chain elongation assay.

3. *The Rim8 C-terminus is a specific substrate of Rsp5 WW3-HECT*

We first tested the system for activity and specificity, using the components described in the previous paragraph (Figure 30). By comparing the Rsp5$^{WW3\text{-}HECT}$ construct against a construct encompassing the Rsp5 HECT domain alone, we confirmed that our Rim8 C-terminus-based construct is indeed a specific substrate of Rsp5 WW domain (Figure 30). Of note, we also tested a construct containing the GB1 domain alone, and we did not observe any ubiquitination in the presence of Rsp5$^{WW3\text{-}HECT}$ (Figure S 20).

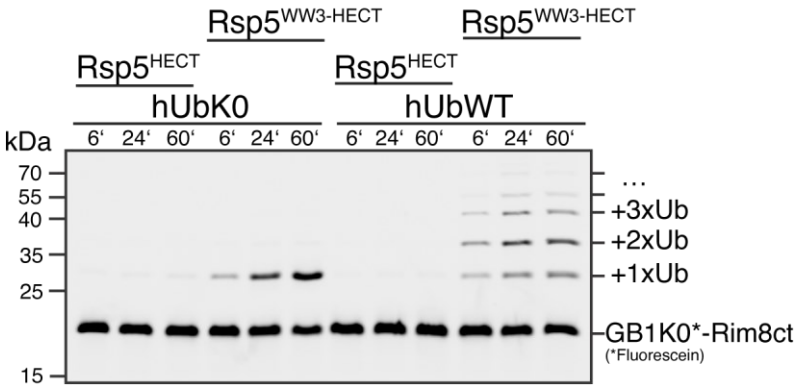


Figure 30. Proof of concept assay of the model system. The GB1K0-Rim8ct substrate was incubated with either $Rsp5^{HECT}$ or $Rsp5^{WW3\text{-}HECT}$, showing specific ubiquitination in the presence of the WW domain Single round ubiquitination as well as multi round ubiquitination was tested, using lysine-free or wild-type ubiquitin. Component concentrations given in Table S 6.*

At the same time, this assay confirmed that the substrate was indeed mono-ubiquitinated when using lysine-free ubiquitin (K0), despite the presence of the N-terminal amino group that could have acted as a second acceptor site for ubiquitination (see Chapter D). Thus, purification of tag-free ubiquitin, with a native N-terminus, can in fact prevent the occurrence of undesired chain linkages.

Finally, by using wild-type ubiquitin we could also confirm that the machinery of this model system is still able to catalyse the formation of longer ubiquitin chains on the substrate (Figure 30).

4. Production of polyubiquitin chains of defined length

Having set up this model system we went on to measure the kinetics of individual chain elongation steps. However, to do this, we needed not only the components of the test assay (Figure 30), but also ubiquitin chains of defined length, to serve as substrates for step-specific chain elongation.

The strategy we chose was to produce all conjugated ubiquitin chains in a single reaction where chain formation is initiated on the substrate, and chains are then elongated to yield a mixture of lengths. This mixture can subsequently be separated into chains of defined length, that could be used to measure the rate of elongation by exactly one ubiquitin.

For the chain separation step, we initially followed a protocol published by Pickart and Raasi[228] using cation exchange chromatography at pH 4.5 to separate free, unanchored ubiquitin chains. However, our model substrate (pI of 4.79[229–231]), in isolation or conjugated to ubiquitin chains, was not soluble under these conditions. While shifting the pH to 5.6 improved solubility and allowed substrate binding to a cation-exchange column, no separation of the ubiquitination products could be achieved (Figure 31(a)). Anion exchange chromatography, while being an improvement over cation exchange, provided sufficient resolution to separate longer ubiquitin chains (> 4), but not shorter chain lengths (<4, especially 1 and 2) (Figure 31(b)). Size-exclusion chromatography also failed to provide sufficient resolution for a preparative separation (Figure 31(c)).

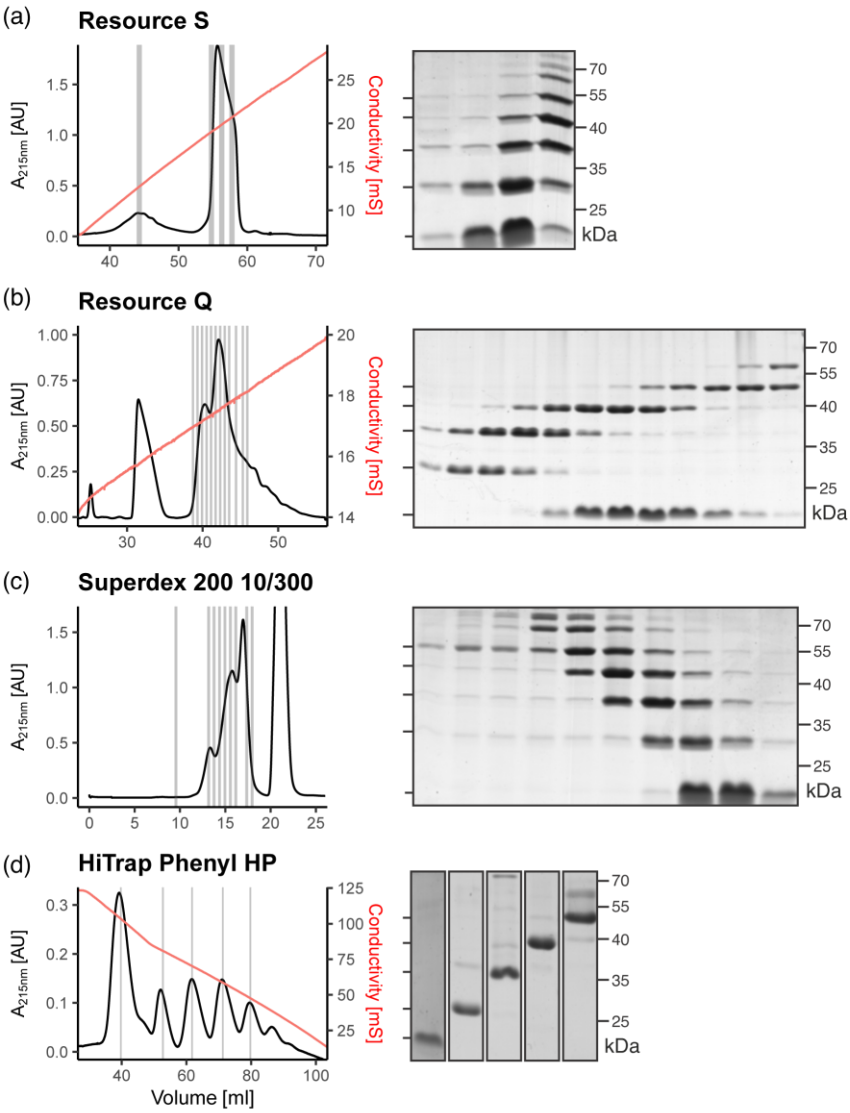


Figure 31. Screening for the preparative separation of ubiquitin chains (hUbK63-only) assembled on the GB1K0-Rim8ct model substrate. Shown are representative column elutions with selected fractions (gray bars) resolved by SDS-PAGE of: (a) Cation exchange using a GE Resource™ S column. (b) Anion exchange chromatography using a GE Resource™ Q column. (c) Size-exclusion chromatography using a GE Superdex™ 200 10/300. (d) Hydrophobic interaction chromatography using a GE HiTrap® Phenyl High Performance column.*

Finally, hydrophobic interaction chromatography (HIC) yielded suitable resolution to separate individual chain lengths in the required amounts (Figure 31(c)). Additionally, the high initial ionic strength required in HIC allowed efficient removal of the cascade of ubiquitination enzymes by precipitation. We confirmed the absence of substrate ubiquitination activity by testing the purified chains in an assay containing all components besides $Rsp5^{WW3\text{-}HECT}$ (data not shown).

5. *Kinetics of chain elongation by Rsp5 WW3-HECT*

Having successfully purified all required components, we could now measure the kinetics of chain elongation in a stepwise manner.

As we sought to quantify the rate of chain elongation by the E3 ligase ($Rsp5^{WW3\text{-}HECT}$), we limited the amounts of E3 to make ligation the rate limiting step in the reaction. The two enzymes acting prior to the E3 are the E1 (Uba1) and the E2 (Ubc4) enzyme. The E1 is typically used at sub-equimolar concentrations with respect to the E2 (see e.g., Zhang et. al. 2016[112] for an overview of typically used

E1/E2 concentrations). Initial testing yielded an increase in the reaction rate when either E1 or E2 concentrations were increased (data not shown). To simplify the optimization process, we therefore used a fixed ratio of 2:3 between E1 and E2. We determined the saturating amount of E1:E2 by varying their concentration (with a 2:3 fixed ratio) in a substrate ubiquitination assay where the concentration of E3 ($Rsp5^{WW3\text{-}HECT}$) was kept constant. The concentration of E3 was set to 20 nM, to maintain the substrate ubiquitination rate in a range that could be quantified with our setup, and saturating concentrations of E1 and E2 were determined to be 10 µM and 15 µM, respectively (Figure 32).

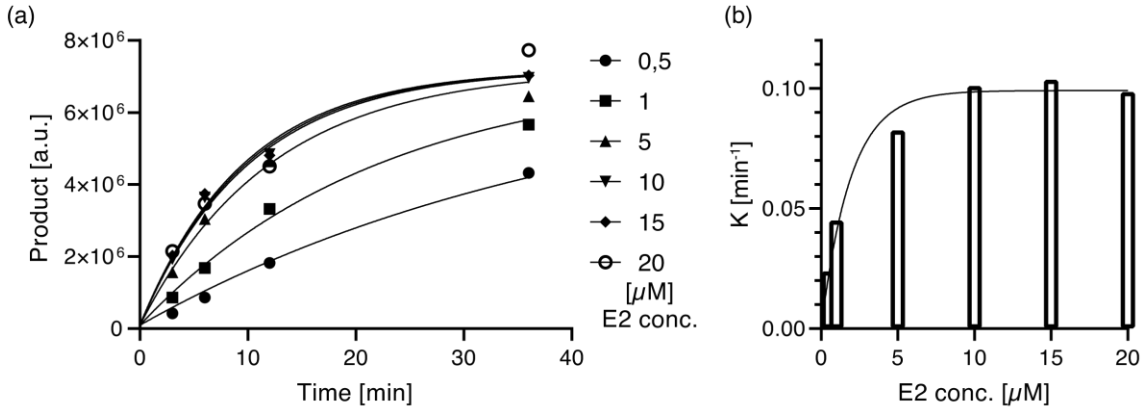


Figure 32. Test for saturating E1/E2 concentration. Different amounts of E2+E1 (at 3:2 ratio) were added to reaction mixes containing substrate and lysine free ubiquitin as well as $Rsp5^{WW3\text{-}HECT}$. (a) Amounts of product (ubiquitinated substrate) formed. Fitted to one phase decay (constrained to a common plateau ≙ all substrate ubiquitinated) with rates K depicted in (b). Raw data shown in Figure S 21.

To determine the kinetics of chain elongation in a stepwise manner, we used the fluorescently labelled GB1K0*-Rim8ct substrate alone or conjugated to K63-linked ubiquitin chains ranging from 1 to 4 ubiquitin moieties. For each substrate, we set up a ubiquitination reaction with the ubiquitination cascade enzymes concentrations determined above and lysine-free ubiquitin (K0), to ensure a single elongation step of the chain. These experiments showed that the chain elongation rate increases significantly once the substrate is mono-ubiquitinated, and progressively decreases the longer the ubiquitin chain is (Figure 33).

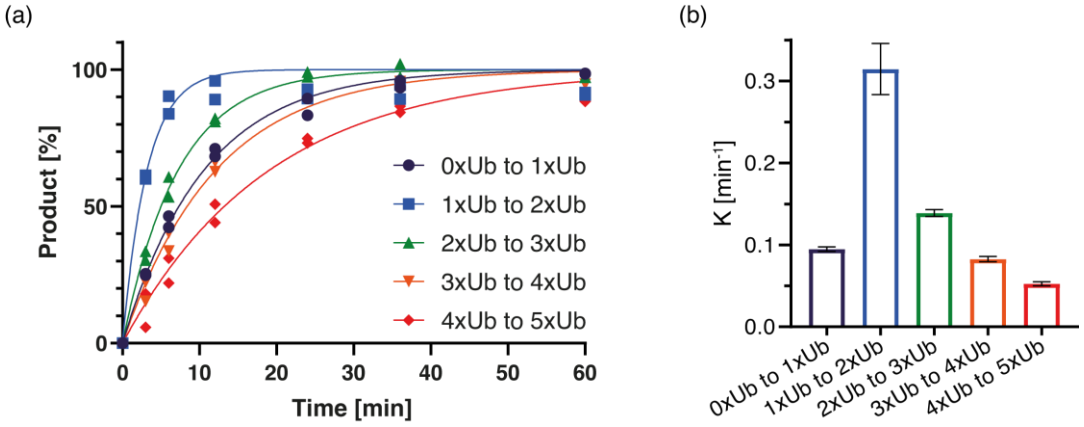


Figure 33. Rates of ubiquitin chain elongation in dependence of the chain length. Substrate in isolation or conjugated with a K63-only ubiquitin chain of specific length was elongated by one unit using K0 ubiquitin. (a) Product formation over time for chains of defined length was calculated from the inverse substrate depletion normalized to the initial educt signal and averaged with product formation normalized to the same range (raw data shown in Figure S 22). Data points from two experimental repeats (shown) were fitted to one phase decay equation (constrained to a common plateau ≙ all substrate ubiquitinated). (b) Calculated elongation rates (K) with the standard error of the mean of the fit.

6. Discussion

We observed that the ubiquitin chain elongation activity of the Rsp5 WW3-HECT ligase increases significantly between the substrate alone and the mono-ubiquitinated one and then decreases again progressively as the chains that are assembled on the substrate grow longer (Figure 33).

These data are not consistent with a model where the second ubiquitin in the chain is bound by the ubiquitin binding site[98], as in such a model one would expect a higher ubiquitination rate for the step from di-ubiquitinated to a tri-ubiquitinated substrate. Similarly, our results do not support a model where the terminal acceptor ubiquitin is actively searched and bound by the UBS, since the rate of elongation for oligo-ubiquitinated substrates decreases as fast as, or faster than, stochastic binding would dictate (i.e., the rate of elongation of a di-ubiquitin chain, $K = 0.14\ min^{-1}$, is almost half of that of a mono-ubiquitinated substrate, $K = 0.31\ min^{-1}$).

While the rate of the mono- to di-ubiquitinated substrate step determined here should be seen as a lower bound for the maximum possible rate, as the saturating E1:E2 concentration was determined on the substrate alone, it is interesting to observe how closely the rates of ubiquitination follow the ratio of acceptor ubiquitins in a chain (1) versus total number of ubiquitins (N).

The rates of ubiquitination for chain-conjugated substrates, $K = 0.08\ min^{-1}$ for a chain of 3 ubiquitin moieties ($< 0.10\ min^{-1}$ (1/3 * 1xUb)) and $K = 0.05\ min^{-1}$ for 4 ubiquitin moieties ($< 0.08\ min^{-1}$ (1/4 * 1xUb)), follow a similar trend, and are always slightly but consistently lower than what would be expected from a statistical binding of the terminal acceptor ubiquitin versus the other ubiquitin units in the chain to an acceptor site. This discrepancy could be explained by steric effects, for example a decrease in the degrees of freedom of the ubiquitins within a chain, caused by the higher number of ubiquitins that need to be accommodated between the substrate and the site of the acceptor ubiquitin. However, observed differences might also be below the accuracy of determination of the absolute rate values (in Figure 33), which is difficult to estimate precisely and could include systematic errors stemming from e.g., the determination of the concentration for the various components, imprecision of manual pipetting, or possible fluorescence quenching effects.

The observation that the kinetics of chain elongation follows the stochastic model upon the addition of a single ubiquitin, together with the ubiquitin binding site being the main source of affinity of HECT domains for ubiquitin, might suggest a role of the UBS in placing the acceptor ubiquitin for chain elongation. However, the orientation in which ubiquitin is bound by the UBS (conserved in a number of HECT domain crystal structures (Figure I 6)) does not allow the UBS-bound ubiquitin to act as an isopeptide bond acceptor, unless so far unobserved HECT domain rearrangements would occur.

The transition state in the ubiquitination mechanism of HECT E3 ubiquitin ligases remains thus uncharacterized. However, as the K_D of ubiquitin - UBS binding is typically in the lower micromolar range, one could still envision a mechanism where ubiquitin disassociates from the ubiquitin binding site just before isopeptide formation. A fine tuning of the avidity (per ubiquitin additive) of the whole chain for the HECT versus the requirement for a non-elongatable ubiquitin (any besides the last in the chain) to be able to dissociate (lower affinity) could be a reason for this intermediate binding strength.

In conclusion, our data would be consistent with a stochastic binding model for ubiquitin chains by the UBS (Figure I 4c), which does not directly explain HECT E3 ligases processivity. However, unfortunately this hypothesis was not experimentally addressed. Moreover, even if the acceptor ubiquitin is not actively recruited, micromolar affinity interactions between conjugated ubiquitins on the substrate and the UBS could help in keeping the already ubiquitinated substrate in proximity of the catalytic site, therefore stimulating a processive, rather than distributive, ubiquitination mechanism. For HECT domains exhibiting an auto-inhibitory mechanism, such as Huwe1 (Chapter B), this stochastic binding of chain ubiquitins to the UBS could also be important to maintain the C-/N-lobe interface open and therefore enhance catalytic activity.

F. Concluding discussion

The HECT domain, identified in 1995[97], has since been found as a common feature of numerous E3 ubiquitin ligases and has been shown to be sufficient for both catalysis and ubiquitin chain linkage determination. Multiple HECT domains, from different proteins, have been crystallized through the years, revealing remarkable flexibility in the arrangement of the C-lobe with respect to the N-lobe of the HECT domain. Early on this flexibility has been suggested to be essential for HECT domain activity[26]. Despite the abundance of structural information, it was so far not possible to define neither the determining factors for chain type specificity of the different HECT domains nor all steps involved in the catalytic cycle. Nevertheless, progress has been made and numerous features of HECT domains important for ubiquitination have been identified.

Using functional assays, previous investigations have shown that a region in the HECT C-terminus is likely the determining factor for chain linkage specificity[131]. However, the molecular details remained elusive. To provide structural insights into the role of this HECT domain region, in Chapter A we identify and characterize a conserved interface between the C-lobe of HECT domains from different E3 ligases and ubiquitin, when ubiquitin is conjugated to the catalytically active cysteine of the HECT domain. Interestingly, this interface is identical between a HECT domain E3 with primarily K48 chain linkage specificity as well as others with distinct K63 specificity – ruling it out as a determining factor for linkage specificity, further reducing the part of the HECT domains C-lobe that is responsible for the HECT domains linkage specificity.

Regulatory mechanisms of HECT domains E3 ligases are a major point of study in the field. The C2 and WW domains, which can be found N-terminally of the HECT domain in Nedd4 family E3s, have been known since the early 2000s to bind to the HECT domain, inhibiting it (Table I 4)[103,121,142]. During the course of this thesis, other laboratories as well as ours found evidence that this inhibitory mechanism is due to locking the HECT domain in a closed, catalytically inactive conformation, where both lobes interact with each other in the so-called T conformation. In Chapter B, we provide functional evidence for this mechanism, showing that Huwe1 activity strongly correlates with the strength of the interface between the N- and C- terminal lobes. Interestingly, we also found that Huwe1, unlike other HECT E3 ligases, does not require an inhibitory factor external to the HECT domain, suggesting that there may be external factors leading to its activation.

Mutations which weaken the Huwe1 N-/C-lobe interface cause a dramatically increased polyubiquitination activity. Such high activity levels were so far thought to require a functioning ubiquitin exo-site (UBS), which had not yet been identified in Huwe1. In Chapter C, we resolve this contradiction by showing that the open conformation of Huwe1 can indeed bind ubiquitin in the canonical location of the ubiquitin binding site and provide evidence that the ubiquitin binding site is also important for ubiquitination activity in Huwe1. Having confirmed the conservation of the UBS in Huwe1, we analysed other HECT domains to investigate whether the stability of the N-/C-lobe interface influences ubiquitin binding site accessibility and how that in turn affects polyubiquitination activity. For Smurf2, we found a modest difference in UBS accessibility leading to a modest activation in an interface-destroying mutant.

Finally, we examined the processivity of HECT domains from a kinetic perspective (Chapter E). Previous publications analysed polyubiquitination by applying complex kinetic models[232] to the change in levels of the individual chain lengths in polyubiquitination of a substrate where the individual product of N-times ubiquitinated substrate is available as the substrate right away. To reduce the complexity of the system, we set up a model in which we could observe a single ubiquitination event at a time, developing a singly ubiquitinable substrate. The use of this substrate alone or pre-conjugated to ubiquitin chains of defined length allowed us to measure the chain length-dependent rates of elongation. Thus, we observed that the rate of ubiquitination of Rsp5 dropped approximately linearly with the chain length, which is consistent with a stochastic binding model for ubiquitin chains.

To expand on an observation during my early experimental work, that a HECT domain could assemble ubiquitin chains from lysine-less tagged ubiquitin, I verified that several different HECT E3 ligases are able to assemble ubiquitin chains on N-terminally extended ubiquitins. These results warn against using N-terminally tagged ubiquitin in experimental settings, particularly when the chain type is a direct or indirect (e.g. via chain type leading to different functional outcomes) focus of the experiment.

III. Materials and Methods

A. Materials

1. Constructs

For each construct the internal database number and a brief description are indicated.

Huwe1 HECT:
#819 C1, WT
#1006 C0, C352K
#1441 C1, E259A
#1443 C1, Y162A/M328A
#1450 C1, Y162A/E259A/M328A
#342 WT, WT
#1415 WT, Y162A
#1437 WT, E259A
#1439 WT, Y162A/M328A
#1447 WT, Y162A/E259A/M328A
#1291 WT, A348P
#1475 WT, M88A
#1476 WT, C110A
#1477 WT, Q190A
#1478 WT, F192A
#1479 WT, Δ-4
#1480 WT, M88A, Δ-4
#1481 WT, C110A, Δ-4
#1482 WT, Q190A, Δ-4
#1483 WT, F192A, Δ-4
#1484 WT, Y162A, Δ-4
#1485 WT, E259A, Δ-4
#1486 WT, Y162A, M328A, Δ-4
#1487 WT, Y162A, E259A, M328A, Δ-4
#1488 WT, A348P, Δ-4

#587 Huwe1 C-lobe, C1
#586/1538(M0V) Huwe1 N-lobe, C0

#189 Rsp5 HECT, WT
#1536 Rsp5 N-lobe, WT

#104 Smurf2 C-lobe, WT
#102 Smurf2 N-lobe, WT
#506 Smurf2 HECT, WT
#1537 Smurf2 HECT, E623A

#1359 Rsp5 HA-HECT
#1358 Nedd4 HA-HECT
#1361 Smurf2 HA-HECT

#897 Rsp5 WW3-HECT
#1395 Nedd4 WW3-HECT
#96 Smurf2 WW3-HECT

#17 Uba1
#611 Ubc4
#1399 UbcH7

#892 GB1-Rim8ct, K1
#1430 GB1K0-C-Rim8ct, K1

#43 H6-TEV-GAMG-Ubiquitin
#1067 H6-Sumo-C-UbiquitinK0
#1058 H6-Sumo-A-UbiquitinK0
#1059 H6-Sumo-AA-UbiquitinK0
#1060 H6-Sumo-AAA-UbiquitinK0
#1380 H6-TEV-Flag-UbiquitinK0
#1381 H6-TEV-V5-UbiquitinK0
#1382 H6-TEV-StrepII-UbiquitinK0
#1383 H6-TEV-cMyc-UbiquitinK0
#1357 H6-Sumo-C-UbiquitinWT
#1398 H6-Sumo-C-UbiquitinWT ΔGG

#980 UbiquitinK0 (“native”)
#1199 UbiquitinWT
#594 UbiquitinG76C
#1062 UbiquitinK63only

#14/RS TEV
#1063 Ulp1

2. Cells

Escherichia coli DH5α
Escherichia coli BL21-CodonPlus(DE3)-RIL

3. Kits/Consumables

KAPA HiFi PCR Kit, Kapa Biosystems
NucleoSpin Gel and PCR Clean-up Mini kit, Macherey-Nagel
NucleoSpin Plasmid EasyPure Mini kit, Macherey-Nagel
pH-Fix 0–14, pH test strips, Macherey-Nagel
Ni-NTA Agarose, QIAGEN
Bio-Rad Protein Assay (Bradford 1976[233]), Bio-Rad
Regenerated Cellulose Membrane 3.5 kD, 12 - 14 kD, Spectra/Por®
Amicon™ Ultra Centrifugal Filter Units 3kDa, 10kDa, 30kDa, Merck Millipore
PD-10 Desalting Column, GE Healthcare
PD MiniTrap™ G-25 Desalting Column, GE Healthcare
premixed 10x Tris/glycine/SDS SDS-PAGE running buffer, Bio-Rad
PageRuler™ Prestained Protein Ladder, Thermo Scientific

Polypropylene tubes, conical, Falcon
Polypropylene tubes, self-standing, Corning
Safe seal tubes, Sarstedt

4. Chemicals

NaH_2PO_4, Roth
Na_2HPO_4, Roth
K_2HPO4, Roth
KH_2PO_4, Roth
Tris base, Roth
NaCl, Roth
DTT, Roth
Imidazole, Roth

Rotiphorese®40 (29:1), Roth
SDS, Roth

TEMED, Sigma-Aldrich
Coomassie Brilliant Blue G-250, Sigma-Aldrich
HCl 6M, Roth
Ethanol, Roth

6-Iodoacetamidofluorescein, Thermo Scientific

Texas Red™ C2 Maleimide, Thermo Scientific

Ampicilline (100 mg/mL in H_2O), Sigma-Aldrich
Kanamycin (50 mg/mL in H_2O), Sigma-Aldrich
Chlroamphenicol (34 mg/mL in EtOH), Sigma-Aldrich
IPTG (1M in H_2O), Roth

NH_4Cl (^{15}N), Sigma-Aldrich
L-Methionine (^{13}C), Sigma-Aldrich
2-Ketobutyric acid-4-^{13}C (L-Ile precursor), Sigma-Aldrich
alpha-Ketoisovaleric acid-$^{13}C_2$ (L-Leu, L-Val precursor), Cambridge Isotope Laboratories, Inc.
DLAM-$I^{\delta 1}T^{\gamma}$ (L-Ile, L-Thr ^{13}C labelling), NMR-Bio
L-Alanine (^{13}C), Cambridge Isotope Laboratories, Inc.

Triton X-100, Roth
TWEEN® 20, Sigma Aldrich
Powdered milk, blotting grade, low fat, Carl Roth
Pierce™ ECL Western Blotting Substrate, Thermo Scientific
anti-HA (HA-7, H3663) Mouse mAb, Sigma-Aldrich
anti-Mouse IgG Alexa-488, Thermo Scientific
NitroBind Nitrocellulose 0.22 µM, Maine Manufacturing

5. Equipment

Econo-Pac® Chromatography Column, Bio-Rad Laboratories, Inc.
HiLoad® 16/600 Superdex® 75 pg, GE Healthcare
HiLoad® 16/600 Superdex® 200 pg, GE Healthcare
HiTrap™ SP HP, GE Healthcare
HiTrap™ Q HP, GE Healthcare
Phenyl Sepharose™ 6 Fast Flow (high sub), GE Healthcare

NGC Chromatography System, Bio-Rad

6. Buffers/Solutions

1. Purification buffers

General lysis buffer:
50 mM Sodium Phosphate pH 7.5
150 mM NaCl
0.5 mM DTT
15 mM Imidazole
~0.1%$^{v/v}$ Triton X-100
~200 mg/L Lysozyme
~10 mg/L DNAse

IMAC wash buffer (NMR):
50 mM Sodium Phosphate pH 7.5
150 mM NaCl
0.5 mM DTT
15 mM Imidazole

IMAC wash buffer (Assay):
50 mM Tris/HCl pH 7.5
150 mM NaCl
0.5 mM DTT
15 mM Imidazole

IMAC elution buffer (NMR):
33 mM Sodium Phosphate pH 7.5
100 mM NaCl
0.33 mM DTT
343 mM Imidazole

IMAC elution buffer (Assay):
33 mM Tris/HCl pH 7.5^{RT}
100 mM NaCl
0.33 mM DTT
343 mM Imidazole

native ubiquitin lysis buffer:
25 mM Tris/HCl pH 7.5^{RT}
50 mM NaCl
~0.1%$^{v/v}$ Triton X-100
~200 mg/L Lysozyme
~10 mg/L DNAse

Cation exchange buffer (native, tag-less ubiquitin):
25 mM Sodium Acetate pH 4.5
A: 50 mM NaCl, B: 500 mM NaCl

Anion exchange buffer (fluorescently labelled ubiquitin & model substrate):
25 mM Tris/HCl pH 7.5^{RT}
0.5 mM DTT
A: 0 mM NaCl, B: 500 mM NaCl

HIC buffers:
Buffer A: assay enzyme storage buffer + 1 M $AmSO_4$
Buffer B: assay enzyme storage buffer

IMAC and IEC cleaning buffer:
50 mM Tris/HCl pH 7.5
1 M NaCl
5 M Urea

2. Final storage buffers

assay enzyme storage buffer:
20 mM Tris/HCl pH 7.5
150 mM NaCl
0.5 mM DTT

Nedd4 buffer changes:
pH 8.0, +5%$^{v/v}$ Glycerol

assay ubiquitin storage buffer:
25 mM Tris/HCl pH 7.5
50 mM NaCl
0.5 mM DTT

3. Controlled Disulfide formation buffers/solutions

Ellman's Buffer:
25 mM Sodium Phosphate/HEPES pH8.3
150 mM NaCl
5 mM DTNB*
0.69 M EtOH*
*100 mg of DTNB (Ellman's Reagent) (5,5-dithio-bis-(2-nitrobenzoic acid) is dissolved in 2 mL of Ethanol and added to 48 mL of buffer (0.1g / 50 ml of 396.35 g/mol = 5 mM)

Reaction Buffer:
25 mM Sodium Phosphate/HEPES pH8.3
150 mM NaCl

Elution/Stabilizing Buffer:
25 mM Sodium Phosphate/HEPES pH6.5
150 mM NaCl
20 mM to 250 mM Imidazole

4. Fluorescent labelling buffers

6-Iodoacetamidofluorescein labelling buffer:
25 mM Sodium Phosphate/Tris pH 7.5
150 mM NaCl

Texas Red™ C2 Maleimide labelling buffer:
50 mM Sodium Phosphate pH 7.2
150 mM NaCl

5. Assay Buffers

Ubiquitination buffer
25 mM Tris/HCl pH 7.5^{RT}
100 mM NaCl
0.2 mM DTT
2.5 mM ATP
5 mM $MgCl_2$

EDTA quenching buffer
25 mM Tris/HCl pH 7.5^{RT}
50 mM NaCl
50 mM EDTA

6. SDS-PAGE buffers/solutions

10x running buffer:
121 g Tris
576 g Glycine
40g SDS
∑4L dH2O

Polyacrylamide gel composition:

	8% (30 - 250 kD)	12% (14 - 150 kD)	16% (5 - 70 kD)	4% Stacking Gel
40% acrylamide (29:1)	14 mL	21 mL	28 mL	3.5 mL
H2O	37.5 mL	30.5 mL	23.5 mL	26.5 mL
1.5M Tris-HCl pH 8.8 (0.375 M)	17.5 mL			-
1.0M Tris-HCl pH 6.8 (0.125M)	-			4.35 mL
20% SDS (0.1%)	0.35 mL			0.175 mL
10% APS (0.1%)	0.35 mL			0.175 mL
TEMED (0.5 uL / mL)	35 uL			17.5 uL

Coomassie PAGE staining solution:
100 mL Ethanol
0.4 g Coomassie Brilliant Blue G-250
900 mL H_2O
5 mL 6M HCl

Sample buffer:
24 mM Tris-HCl pH 6.8
6% SDS
30% Glycerol
0.03 g Bromophenol blue
0.5M DTT
∑50 mL H_2O

7. Blotting Buffers

Blotting Buffer:
200 mL Bio-Rad 10x Tris/glycine/SDS SDS-PAGE running buffer
400 mL Ethanol
∑2L dH2O

TBS:
20 mM Tris-HCl pH 7.5
150 mM NaCl

TBS-T:
20 mM Tris-HCl pH 7.5
150 mM NaCl
0.05%$^{v/v}$ TWEEN 20

Blocking solution:
TBS-T with 5%$^{v/v}$ powdered skim milk

Primary antibody incubation solution:
TBS-T with 2%$^{w/v}$ powdered skim milk, antibody as indicated by manufacturer

Secondary antibody incubation solution:
TBS-T with 1%$^{w/v}$ powdered skim milk, antibody as indicated by manufacturer

8. Culture Media

LB

LB media, 1L (in H_2O)	
10 g	Bactotryptone
5 g	Yeast extract
5 g	NaCl

M9

M9 media, 1L (in H_2O or D_2O)	
0.5 g	$^{15}NH_4Cl$
4 g	Glucose
6 g	Na_2HPO_4
4 g	$KH_2PO_4 \cdot 3\ H_2O$
0.5 g	NaCl
1 ml	1M $MgSO_4$
1 ml	1mg/mL Biotin
1 ml	1mg/mL Thiamin
1 ml	1000× antibiotics
300 µl	1M $CaCl_2$
1 ml	1000 × trace elements
0.1 ml	10k × trace elements

A. Methods

1. Cloning

1. Preparation of $CaCl_2$ chemically competent *E. coli* and transformation

Both DH5α as well as BL21-CodonPlus(DE3)-RIL *Escherichia coli* cells are made transformation competent using the well-known method of permeabilizing the cell wall using calcium ions[234]. A culture grown to an OD_{600} of 0.5 is pelleted at 1500 RCF for 10 minutes at 4°C. After resuspension in ice cold 0.1M $CaCl_2$ and centrifugation, the pellet is resuspended in 0.1M $CaCl_2$/10% glycerol and flash frozen in 50 µl aliquots for long term storage at -80°C.

Cells were transformed by 4°C incubation with plasmid DNA for 5 to 20 minutes followed by a 42°C heat-shock for one minute. After recovery for antibiotic resistance expression in liquid LB media for 30 to 90 minutes they could be grown in the presence of the respective selection antibiotics on a LB agar plate or in LB liquid culture.

2. Polymerase chain reaction and linear DNA amplification

All polymerase chain reactions (PCR) as well as linear amplification reactions for restriction-free cloning were performed using the KAPA HiFi PCR Kit according to manufacturer protocol.

Volume	Component
variable	65 ng template DNA
0.5 µl	primer mix (10 µM each)
2.5 µl	KAPA HiFi Buffer (5x)
0.35 µl	KAPA dNTP Mix (10 mM)
0.25 µl	KAPA Hifi DNA Polymerase (1 U/µl)
∑ 12.5µl	H_2O to top up to

Table 1. Manufacturer protocol at used scale.

Step	Temp.	Duration	
Denaturation (initial)	95°C	3 min	
Denaturation	98°C	20 sec	} 25 cycles
Annealing	50-72°C	15-30 sec	
Extension	72°C	30 sec / kb	
Extension (final)	72°C	1 min / kb	
Storage	4-8°C	∞	

Table 2. Cycling protocol used.

3. Restriction-free cloning

Constructs used in this work, if not previously available in the lab[150,194], were made using restriction-free cloning techniques. Mutations were introduced based on the QuikChange™ methodology originally introduced in a kit by Stratagene. QuikChange and its further generalization to restriction-free cloning[235] works by replication of a whole vector using a high-fidelity polymerase. The polymerase synthesizes the target using a primer-pair that is (partially) self-complementary as well as partially complementary to the vector to be changed and contains the desired change, be it a single mutation or a longer insert or deletion. Of note, this reaction is a linear amplification as opposed to a polymerase chain reaction as the product cannot serve as template in later cycles due to the position of the nicks in the product (see Figure 34) and thus is uniquely dependent on the amount and quality of the used template vector DNA. After the reaction, the template vector can be removed by treatment using the restriction enzyme DpnI which cuts methylated or hemi-methylated *E. coli* DNA adenine methylase GATC sequences[236].

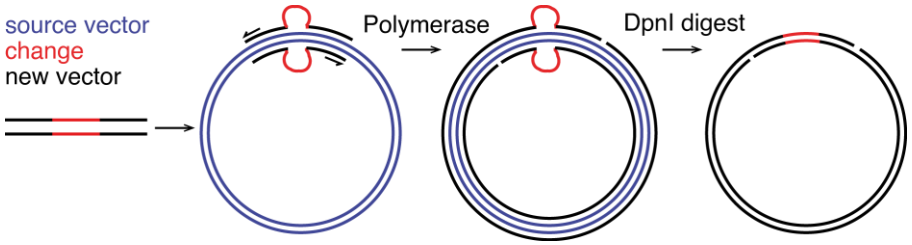


Figure 34. Principle of restriction-free cloning, adapted from Bond and Naus[183].

Small changes such as single point mutations, deletions or i.e., insertion of short tags (< 50 bp) were routinely performed using only standard DNA oligos ordered from Sigma-Aldrich or Eurofins as primers. Larger inserts were first PCR amplified from a source vector or commercially synthesized DNA (GeneArt Strings™) using primers complementary to both source as well as the destination vector. Primers were designed using online tools: PrimerX[237] for smaller inserts, rf-cloning.org[183] for larger inserts, or equivalent workflows aimed at optimizing annealing temperature equivalence between complementary regions upstream and downstream of the change as well as for GC rich primer ends.

Following DpnI treatment, the reaction product was transformed into DH5α which were plated on a LB agar plate containing the respective selection antibiotic for clone selection. After outgrowth of individual clones in antibiotic containing liquid LB media (37°C, 180 rpm, over-night), vector DNA was isolated using the NucleoSpin Plasmid EasyPure Mini kit by Macherey-Nagel and sent to be Sanger sequenced at Eurofins.

2. *Protein Purification*

1. Protein Production

Proteins were expressed in the BL21-CodonPlus(DE3)-RIL *E. coli* strain which carries a plasmid expressing tRNAs for rare Arg, Ile and Leu codons, which is selected for using chloramphenicol, and also expresses the λDE3 lysogen for T7 RNA polymerase-dependent expression.

Expression vectors used are the pETM set of vectors designed at the EMBL protein expression facility[238], using Kanamycin as selection antibiotic, as well as pProEX™ Htb (Invitrogen), using Ampicillin as selection antibiotic. Both sets of vectors use a T7 promoter under the control of a lac operator and express the lac repressor (lacI). Expression of the protein of interest is induced by relieving lac repressor binding using the allolactose analogue Isopropyl-β-D-thiogalactoside (IPTG) at 1 mM concentration.

After transformation, independently of final expression media, cells are first grown in LB media. They are then transferred into the expression medium by centrifugation at 3500 RCF for 10 minutes, removal of supernatant and resuspension at an initial optical density (OD_{600}) of 0.1 to 0.4. Choice of starting OD_{600} depends on whether the expression media is D_2O based and how carefully the presence of hydrogen is to be supressed.

In the expression media, growth at 37°C is allowed until an OD_{600} of 0.8, at which expression is induced using IPTG. Cells are allowed to express the protein at the temperature of 18 to 20°C overnight. In case of isotope labelling, appropriate metabolites are added during growth, one hour before induction, as detailed under the “Growth media” section above.

After over-night expression, cells are harvested by centrifugation at 4600 RCF for 15 minutes, with the pellet either flash frozen in liquid nitrogen and stored for medium term storage at -20°C or used directly.

2. Protein Purification

Lysis and Lysate clarification

Pelleted cells were resuspended in 35mL lysis buffer per 500mL original expression culture volume. From this point on, cell and/or protein containing solutions are stored at 4°C whenever possible in between purification steps/use. The lysis buffer contains lysozyme and Triton X-100 to aid in lysis and DNAse to reduce viscosity originating from cellular DNA. Cells are lysed using two 1min 30s cycles (0.5 s on, 0.5 s off) on a Bandelin Sonopuls sonicator with a KE76 probe at 40% amplitude. Cell debris is pelleted at 40000 RCF for 30 minutes and discarded, while the supernatant is additionally filtered with a 0.22 µm filter.

Immobilized Metal Ion Affinity Chromatography (IMAC)

As both in ubiquitin as well as in the HECT domain E3 ligases an unmodified C-terminus is important for catalytic activity, proteins here are mostly (see also Native/Tag-less ubiquitin Purification below) expressed with a N-terminal His6-tag. The His6-tag has intrinsic affinity to nickel ions. To purify protein, in a first step, the clarified supernatant is loaded onto 1-4 mL Ni-NTA bed volume per 1 L of expression volume of nickel-charged nitrilotriacetic acid agarose resin (Ni-NTA) pre-equilibrated with buffer. After the clarified supernatant has passed through the Ni-NTA resin, IMAC wash buffer is added until a Bradford assay of the liquid passing the Ni-NTA resin is indistinguishable (by eye) from the wash buffer Bradford assay control.

Then the Ni-NTA resin is eluted using Imidazole containing IMAC elution buffer until the Bradford assay of the eluate exiting the column is close to (90%+ of maximum Bradford intensity) an elution buffer only Bradford assay control.

As the high concentration of Imidazole in the elution buffer can have a detrimental effect on protein solubility/stability, it is lowered right after elution by dialysis. Dialysis is performed against a suitable volume of buffer (storage buffer) using a membrane of appropriate cut-off (protein size should be double the manufacturer specified cut-off). Depending on further purification steps, dialysis was performed once or multiple times against approximately 50x the sample volume of fresh dialysis buffer (storage buffer) for 4 h+, depending on the desired Imidazole end-concentration (see also Subtractive (reverse) IMAC).

Protease Cleavage

The His6-tag or the His6-tagged solubility tag here is typically separated from the native sequence of the protein of interest by a protease cleavage site. Protease sites used here include TEV, the tobacco etch virus protease, as well as Ulp1, the ubiquitin-like-specific protease 1 (“SUMO protease”) recognition sites. Between the His6-tag and the protease cleavage site, some constructs (see Constructs list) carry a solubility tag such as GST (Glutathione S-transferase) or the SMT3 SUMO (small ubiquitin-like modifier) homologue, serving also as the Ulp1 cleavage motif. Proteases were routinely produced in house, purified using a single step IMAC procedure as described in the last section and flash frozen in liquid nitrogen in 1 ml aliquots at 0.5 ml/mL with 25-50% glycerol as an additive. Per litre of initial expression volume, depending on overexpression yield, one to four protease aliquots were used.

Protease cleavage can conveniently be done concomitantly to the dialysis that is performed following the initial IMAC elution to remove the Imidazole used for elution.

Subtractive (reverse) IMAC

After protease cleavage and dialysis, the His6-tag (if it has not passed the dialysis membrane) or the His6-tagged purification tag can be removed from the mixture in a second IMAC step. In this step the tag binds to the column resin while the protein of interest is found in the flow-through. A potential pitfall that must be kept in mind is that proteins even without carrying a His6-tag can have intrinsic affinity to the Ni-NTA resin (for example by surface exposed histidines). Constructs of Smurf2 as well as Nedd4 need 5 mM Imidazole in the buffer used for the second IMAC to prevent them from interacting with the column resin, whereas for other proteins such as ubiquitin, Imidazole could be fully removed by repeated dialysis. By removing Imidazole fully, the impurities that bound to the column resin despite the 15 mM Imidazole used in the initial IMAC wash-buffer now bind even tighter to the column resin and can thus be removed. Imidazole, if remaining, was removed by buffer exchange either by further dialysis, by size-exclusion chromatography (SEC; see below), using a PD-10 Desalting Column or in a centrifugal concentrator

General Medium-Pressure Liquid Chromatography Workflow

Non gravity-flow based column purifications were performed using the Bio-Rad NGC chromatography system. Elution was monitored via the absorbance at 280 nm as well as at 215 nm for proteins with low amounts of aromatic amino acids (i.e., ubiquitin). Fractions are collected and analysed for purity using Coomassie-stained SDS-PAGE.

Size-Exclusion Chromatography (SEC)

As experiments routinely required proteins to be stable for multiple hours or longer, a final polishing size exclusion run was typically performed even when proteins were pure, as could be judged using Coomassie stained SDS-PAGE. For this the protein was concentrated to a volume (1-2 mL) that could be loaded on a HiLoad® 16/600 Superdex® and isocratically eluted using storage buffer.

Ion-exchange chromatography (IEC)

Constructs submitted for crystallization as well as constructs prone to partial degradation during expression (i.e., GB1-Substrates and E3 constructs including more domains besides the HECT domain) underwent an extra IEC purification step in addition to the aforementioned typical IMAC -> protease cleavage -> IMAC -> SEC workflow.

IEC was performed at a suitable pH at which the protein of interest is known to be soluble and at which the pH is at least one unit away from the predicted isoelectric point (pI)[229–231]. Depending on whether this pH was below or above the pI, cation exchange or anion exchange was selected.

HiTrap™ SP HP (cation exchange) as well as HiTrap™ Q HP (anion exchange) columns were used, and proteins were eluted using a salt gradient from 0-50 mM to 500 mM salt.

Hydrophobic interaction chromatography (HIC)

Ube1, displaying a prominent degradation/truncation product after expression, was purified using an HIC step instead of SEC. 2 mL Phenyl Sepharose™ 6 Fast Flow (high sub) per litre of expression culture were equilibrated with storage buffer in the presence of 1 M $AmSO_4$ and incubated with protein in the same buffer for 4-16 hours at 4°C. Using a stepwise elution from 500 mM to 0 mM $AmSO_4$ in the storage buffer, the full length Ube1 was eluted with the truncated isoform only eluting at the very end of the stepwise gradient. $AmSO_4$ was removed by dialysis and protein was flash frozen in liquid nitrogen in aliquots, at a concentration of 20 µM.

Native/Tag-less Ubiquitin Purification

For the experiments detailed in Chapter E it was necessary to produce ubiquitin without any non-native residues at either terminus. While the routinely used TEV and Ulp1 proteases are reported to be relatively promiscuous with respect to the P1' position of their recognition sites[239–241], constructs with the P1' being the starting methionine of ubiquitin were not a substrate of either protease and thus could not be used.

Purification of ubiquitin employing its resistance to acid precipitation has been described before by Pickart and Raasi 2005[228], who used perchloric acid followed by cation exchange in ammonium acetate. Here, the more hazardous perchloric acid was substituted with acetic acid.

After lysis in a lower salt lysis buffer, to facilitate loading on an ion exchanger, and removal of cell debris as described above, cellular protein was precipitated by dropwise addition of (glacial) acetic acid under continuous agitation until pH 4.5 was reached as monitored by pH indicator paper. After incubation at room-temperature (due to some cellular proteins being less soluble at RT than at 4°C) for 15 minutes, precipitated cellular protein was removed by 40000 RCF, 4°C centrifugation for 30 minutes.

Supernatant was filtered and dialysed against (or diluted 1:5 with) cation-exchange buffer A, loaded onto a HiTrap™ SP HP and eluted using a salt gradient (cation-exchange buffer A to buffer B).

Following a final SEC purification step (as described above), pure ubiquitin without terminal modifications was obtained.

Protein concentration

Proteins were concentrated in Merck Millipore Amicon™ Ultra Centrifugal Filter Units of appropriate cut-off size, according to manufacturer protocol. Of note, Centrifugal Filter Units were washed extensively prior to use to strip residual soluble substances present.

SDS-PAGE

Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) was performed according to well-known established principles[242,243], using the SE250 Mighty Small II Mini Vertical Electrophoresis Unit (Hoefer) and 1.0 mm thick SDS-PAGE gels prepared with a Hoefer Gel Caster. PageRuler™ Prestained Protein Ladder by Thermo Scientific was used to estimate the size of proteins of interest. Buffers and recipes used are given in detail above.

3. Protein Post-Processing

Disulfide Formation

For the controlled formation of asymmetric disulfides, one of the components was activated using Ellman's reagent (DTNB = 5,5-dithio-bis-(2-nitrobenzoic acid)). DTNB, as well as TNB (5-thio-2-nitrobenzoic acid), are colourless when attached to protein thiols, whereas the free thiol form of TNB in solution is yellow. Thus, both the attack of DTNB by protein thiol as well as the attack of the TNB-activated protein thiol by another protein's thiol can conveniently be followed optically due to the release of yellow TNB.

One of the components, usually ubiquitin^{G76C}, was purified without removal of the His6-tag. The His6-tagged protein was loaded onto Ni-NTA column material and activated using buffer containing Ellman's reagent in gravity-flow. In this step, it is not crucial to remove the thiol-containing DTT, present in purification and protein storage buffers, as the Ellman's reagent, with which it reacts, is used in significant excess.

After activation, remaining DTNB-containing buffer was washed off and afterwards the second component was added, so its free thiol could attack the TNB-activated thiol group of the column-fixed

component. The second component was used in excess and exchanged into DTT-free reaction buffer right before use (using a PD-10 column, according to manufacturer protocol). This step is crucial to avoid competition of the DTT thiol with the thiol of the activated component. The column was then washed to remove unreacted educt and the disulfide-conjugated complex was finally eluted using a stepwise imidazole gradient.

For increased disulfide stability, buffers at pH 6.5 were used to wash and elute the reacted disulfide. Additional clean-up of the resulting disulfide (side products, educts, and impurities) was performed using SEC, a finer IMAC gradient or HIC, as needed.

Fluorescent Labelling

Ubiquitin and ubiquitination model substrates were routinely fluorescently labelled on their cysteines.

Model substrates were designed with a single cysteine in a linker, whereas for ubiquitin a single cysteine was inserted between the protease cleavage site and the starting methionine of the native sequence.

Fluorescent labelling reagents used were 6-Iodoacetamidofluorescein (ubiquitin and model substrate) or Texas Red™ C2 Maleimide (model substrate) ($\varepsilon_{TexasRed}^{582nm}$ = 112 mM^{-1} cm^{-1}, $\varepsilon_{Fluorescein}^{490nm}$ = 80 mM^{-1} cm^{-1}). A 5x molar excess with respect to the number of cysteines (number of proteins times cysteines per protein) of the labelling reagent was dissolved in DMSO (typically in a 100 µl volume) and added to the proteins to be labelled. Proteins to be labelled were buffer-exchanged to be DTT-free immediately beforehand. Labelling was performed for 2 hours in the dark at room temperature, after which remaining unreacted labelling substance was quenched using an excess of DTT. Excess dye and labelled contaminants were removed using anion-exchange chromatography or SEC. Buffers used in the labelling reactions are given above.

3. Western-Blotting

Western blot transfer was performed using the TE 22 Mighty Small Transfer Tank, blot transfer system by Hoefer for 90 minutes at 370 mA at 4°C to a nitrocellulose membrane (0.22 µm). After transfer, the membrane was successively incubated with blocking solution (25 mL), primary antibody incubation solution (7 mL) and secondary antibody incubation solution (7 mL) with triplicate washes with 25 mL TBS-T after each step. Imaging of secondary antibody fluorescence was performed using the Amersham Imager 600 by GE Healthcare.

4. Ubiquitination Assays

Ubiquitination reactions were performed using E1, E2, and E3 enzyme cascade described. As the E1, *Homo sapiens* Ube1 was used. *Homo sapiens* Ube1 was used also in combination with E2/E3 enzymes from other organisms as it is well conserved.

As the E3, constructs consisting of the HECT domain alone, or of a module containing the HECT domain as well as the adjacent WW substrate-binding domain, were used. E3 enzymes used are the *Saccharomyces cerevisiae* Rsp5 as well as the *Homo sapiens* Smurf2, Nedd4 and Huwe1.

In the case of *Saccharomyces cerevisiae* Rsp5, the cognate E2 from *Saccharomyces cerevisiae,* Ubc4 was used, whereas for the *Homo sapiens* E3 ligases the *Homo sapiens* E2 UbcH7 was used.

In a typical reaction, 60 µM of ubiquitin (fluorescently labelled, wild-type or single/lysine-free mutant), 10 µM of an E2 and 3 µM of an E3 were employed. Temperatures ranged from 20°C to 37°C and were chosen depending on the stability of the given enzyme relative to the measurement duration. Exact amounts of enzymes, temperatures and reaction times were varied depending on the E3 used, the

observed component, the type of assay as well as other factors and are given directly in the respective figure captions.

Reactions were performed in a universal reaction buffer given in the buffer list above.

Reaction products (chains, mono-ubiquitinated or thioesterified components) were separated using SDS-PAGE and then detected employing the fluorescence of fluorescently-labelled ubiquitin, substrate, or secondary antibody, using an Amersham Imager 600 by GE Healthcare.

With this setup we could observe the two different reaction steps described below.

1. Isopeptide Assay

While ubiquitin is activated by the E1 and passed on to the E2 as a thioester, the E3 catalyses the actual ubiquitination reaction, whereby ubiquitin is attached to a substrate amine as an isopeptide.

To measure the difference in activity which distinct E3 mutants exert on this step, E1, E2, ubiquitin and substrate were kept constant in the reactions (typically by employing a master mix of common components) and all (ubiquitin) thioesters were cleaved by the inclusion of an excess of DTT in the Lämmli buffer used to stop the reaction.

Substrates used here were the HECT domain itself, as it can typically auto-ubiquitinate, a minimal substrate designed from a known WW domain-binding substrate peptide, or ubiquitin itself in the generation of free-standing chains as well as in chain elongation.

2. Thioester Assay

Besides the transfer of ubiquitin to the substrate, another step of interest was the transfer of activated ubiquitin from the E2 to the E3.

For E3 ligases where this step is significantly faster than the transfer of ubiquitin to substrate as an isopeptide, thioester formation could be trivially observed by stopping the reaction after a sufficiently short time and confirming the (relative) absence of formed isopeptide by treating half of the reaction with DTT (which removes the ubiquitin-thioester but leaves the isopeptide intact).

For E3 enzymes where the rate of thioester transfer was close to the isopeptide formation rate, this method was unsuitable as significant amounts of isopeptide were generated even with very short reaction times. Since the HECT domains C-terminus is important for isopeptide formation but is not required for thioester formation, a Δ-4 mutant (missing 4 C-terminal amino acids) was used here to observe thioester formation rates independently of isopeptide formation by blocking isopeptide formation entirely. Additionally, to limit the influence of E1 to E2 thioester transfer and only measure E2 to E3 thioester, these assays were performed after pre-incubation of all components, minus the E3, for a fixed time, to allow E2 loading with ubiquitin. Afterwards, the E1 activity was quenched by blocking its ATPase activity (using EDTA to sequester the required Mg^{2+} ions) and the pre-loaded E2 was incubated with the E3 for defined reaction times.

5. *Nuclear magnetic resonance*

NMR is a spectroscopic method, in which the energy difference between different states is probed by electromagnetic radiation. In NMR, the states of interest are the quantum states of the nuclear spin which emerges from the spins of protons and neutrons in a given nucleus. Nuclear spins are characterized by the angular momentum quantum number I, and their magnetic quantum number m. Depending on I, m can take a total number of $2I + 1$ values (Zeeman levels) from $-I$ to $+I$ in integer steps

$$m = -I, \dots, +I$$

The nuclei we employed for NMR here are ^{1}H, ^{15}N and ^{13}C, all nuclei with $I = 1/2$.

With the magnetic moment μ

$$\mu = \gamma \, \mathrm{I}$$

the projection of this angular momentum on an axis is given as

$$I_z = \hbar \, \mathrm{m}$$

with the magnetic moment μ_z along this axis being

$$\mu_z = \gamma \, \mathrm{I}_z = \gamma \, \hbar \, \mathrm{m}$$

The proportionality constant γ between the spin magnetic angular momentum I and the magnetic moment μ is called the gyromagnetic ratio and depends on the nucleus.

When put into a magnetic field, spins which are not perpendicular to the magnetic field "precess" around the field vector (B) at the larmor frequency (ω).

$$\omega = -\gamma \, B$$

or, corrected for the chemical environment (chemical shift σ) surrounding the nucleus

$$\omega = -\gamma \, (1 - \sigma) \, B_0$$

The energy of a spin angular momentum state in a magnetic field is given as

$$E = -\mu \, B$$

or, in the case of a B_0 field in direction of the z axis, as

$$E_m = -\gamma \, I_z \, B_0 = -m \, \hbar \, \gamma \, B_0$$

Thus, the energy of a transition between adjacent levels m and $m \pm 1$ is

$$\Delta E = \hbar \, \gamma \, B_0$$

$\Delta m = \pm 1$ is the quantum mechanical selection rule for allowed/directly observable transitions in NMR, which are called single quantum transitions. In the case of no external magnetic field (B), these levels are then degenerate.

The population difference of the states is given by the Boltzmann distribution, and even at high magnetic fields (B) energy differences (ΔE) are miniscule. As an example, measurement conditions typically used in this work yield a vanishingly small difference between two states of 0.999871:

$$\frac{N_{m+1}}{N_m} = e^{-\Delta E \, / \, (k_B \cdot T)} = e^{\frac{-\hbar \cdot 2.6752 \cdot 10^8 \frac{Hz}{T} \cdot 18.79 \, T}{k_B * 298 \, K}} = 0.999871$$

$$\Delta E = \hbar \, \gamma \, B_0$$

$$T = 298 \, K \, (24.85°C)$$

$$\gamma_{^1H} = 2.6752 \cdot 10^8 \, \frac{Hz}{T} \, (^1H)$$

$$B_0 = 18.79 \, T \, (800 \, MHz, ^1H)$$

This makes NMR a relatively insensitive technique even at high magnetic fields. However, as a finite population difference still persists, the bulk magnetic moment/nuclear polarization

$$M = \sum \mu$$

summed up over all spins in a sample does not vanish.

The behaviour of the bulk nuclear polarization in a magnetic field has been described by F. Bloch in 1946[244] using a set of equations now known as the Bloch equations, which give the phenomenological behaviour of the bulk magnetisation, including relaxation effects.

$$\frac{d\,M_{xy}(t)}{d\,t} = \gamma\,(M(t)\,\times\,B(t))_{xy} - \frac{M_{xy}(t)}{T_2}$$

$$\frac{d\,M_z(t)}{d\,t} = \gamma\,(M(t)\,\times\,B(t))_z - \frac{M_z(t)\,-\,M_0}{T_1}$$

The solution of these equations under the conditions of free precession in a static B_0 field with relaxation is:

$$M_z(t) = M_{z,eq}\cdot(1-2\,e^{\frac{-t}{T_1}})$$

$$M_{xy}(t) = M_{xy}(0)\,\cdot\,e^{-i\omega t}\,\cdot\,e^{\frac{-t}{T_2}}$$

In these equations, in addition to the forces of the B field (B) acting on the nuclear polarization (M), Bloch introduced two relaxation times, the "transversal" spin-spin relaxation time (T_2) as well as the "longitudinal" spin–lattice relaxation time (T_1). The spin-spin relaxation time (T_2) represents the relaxation generated by the interaction of the spin with other spins in the sample, whereas the spin–lattice relaxation time (T_1) represents the return to the equilibrium (with respect to the B_0 field in z-direction) magnetization.

In experiments here, depending on the signal to noise requirements, we tried to minimize transverse relaxation (maximize T_2 time) by using measurement buffer and growth media where deuterium substitutes hydrogen (replacing H_2O with D_2O, and adding deuterated glucose to a minimal growth medium), except for labelled methyl groups and/or amides. At the same time, the measurement temperature was chosen (within the physiological range) as high as protein stability allowed to increase the macromolecular tumbling rate to maximize T_2 time as well as potentially minimize T_1.

In NMR experiments performed, $M_{xy}(t)$ induces a signal in a measurement coil around the sample. While theoretically there could be two perpendicular coils x and y, in practice real and imaginary (x and y) are generated by mixing the signal with a carrier frequency as well as the carrier frequency 90° out of phase. This signal is sampled in discrete timesteps for a certain time t_1, and is referred to as the free induction decay (FID).

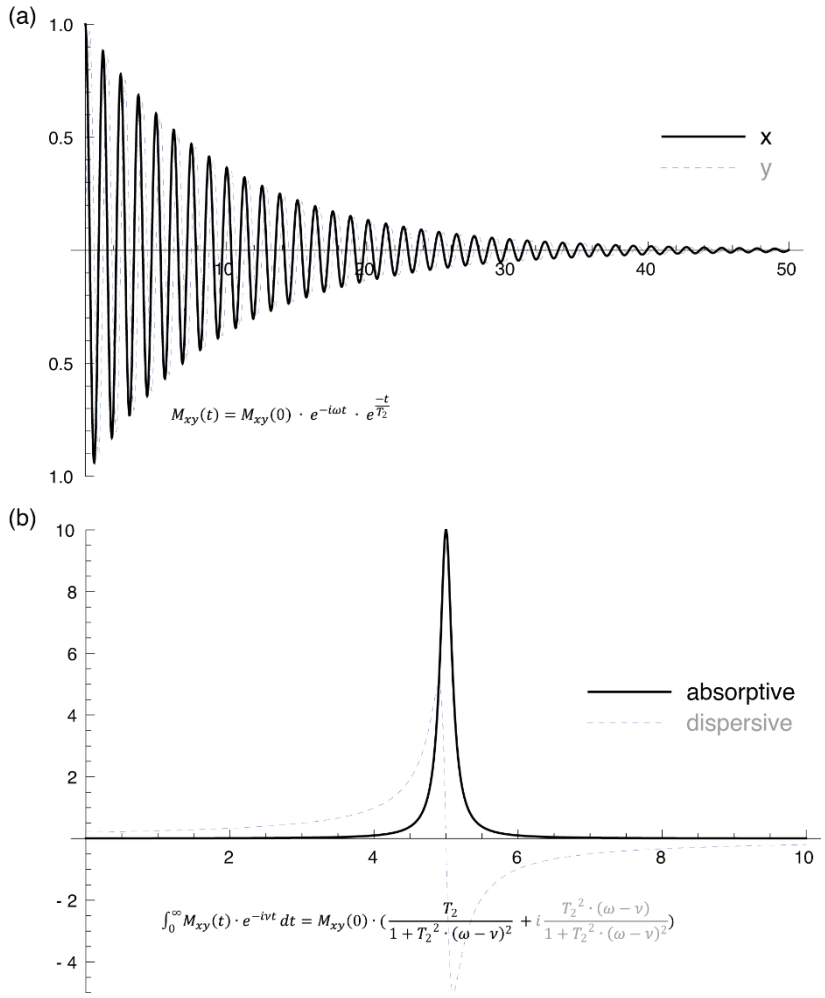


Figure 35. (a) Example of a FID: real and imaginary components (x,y in $M_{xy}(t)$ above) of a signal with (dimensionless) example values of $\omega = 5$, $M_{xy}(0) = 1$ and $T_2 = 10$. (b) Fourier transformed signal corresponding to (a).

The FID is further Fourier transformed from the time domain (t_1) into the frequency domain (f_1), giving the typical spectrum with its Lorentzian shaped peaks.

In the 2D experiments employed here (see following subsections), scalar ("J"-) coupling between different spins is exploited to transfer magnetization/coherence between different spins. In this way, a second, indirect time domain (t_2), during which such a coherence is allowed to evolve, is correlated to the first spin.

When recording such a 2D spectra, multiple FIDs are recorded in t_2 time increments, and the recorded data is additionally Fourier transformed into the frequency domain on the second axis (f_2).

1. Transverse relaxation-optimized spectroscopy (TROSY, ^{1}H ^{15}N)

For small (<25 kDa) proteins, a ^{15}N labelled TROSY 2D experiment was performed routinely. The naturally most abundant isotope of nitrogen is ^{14}N (99.6%), which has broad quadrupolar signals due to its integer spin of 1. For this reason, *E. coli* expressing the protein of interest were grown in minimal media with $^{15}NH_4Cl$ as the sole nitrogen source, allowing for high enrichment in the half integer spin isotope ^{15}N.

As in the classical HSQC (Heteronuclear Single-Quantum Correlation) experiment, the TROSY (Transverse Relaxation-Optimized Spectroscopy) experiment transfers magnetisation, initially generated using a 90° pulse on the hydrogens, to the less sensitive nitrogens with an INEPT (Insensitive Nuclei Enhancement by Polarization Transfer) block using J-coupling. Magnetisation then evolves with the nitrogen chemical shift, resulting in a second dimension that is also Fourier transformed. Since ^{1}H has higher sensitivity than ^{15}N, as in a standard (sensitivity enhanced) HSQC, the magnetisation is then transferred back to the more sensitive hydrogens for detection using two reverse INEPT elements[245–248].

In an HSQC the HN J-coupling is refocused by a 180° proton pulse in the middle of the nitrogen chemical shift evolution period (t_1), as well as by broadband decoupling during t_2. In contrast, in the TROSY experiment there is no decoupling (NH) during the chemical shift evolution on the nitrogen as well as during recording of the FID.

Whereas in HSQC slow- and fast-relaxing components of the magnetization are mixed, producing a signal of intermediate broadness, in TROSY they are kept separate and the signal resulting from the fast-relaxing components is removed using gradients and/or phase cycling. In the foundational paper in 1997, Pervushin et al. showed that the slow relaxing component is proportional to dipole-dipole (DD) minus the chemical shift anisotropy (CSA) and that this effect is maximized at fields between 900 to 1500 MHz[247,249,250].

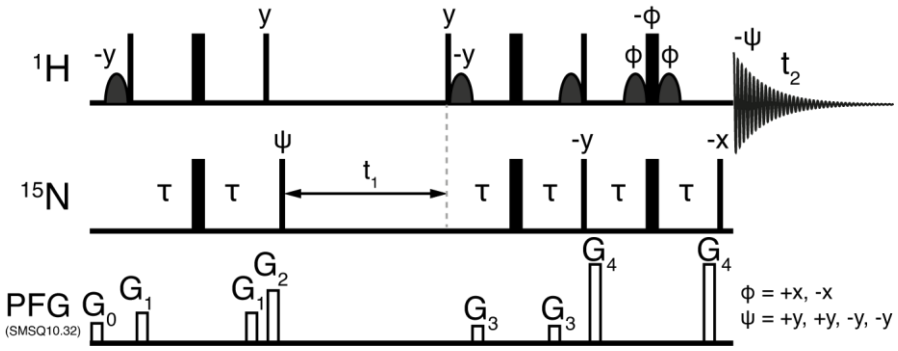


Figure 36. The ^{15}N-TROSY sequence used here. It originates from Lewis E. Kay (personal communication). As in the original ^{15}N-TROSY sequence[248] it contains a watergate sequence to supress signal from the water in the sample in addition to the elements described in the text above.

2. Band-Selective Optimized-Flip-Angle Short-Transient Heteronuclear Multiple-Quantum Correlation (SOFAST-HMQC, ^{1}H ^{13}C)

For larger proteins, ^{15}N-TROSY was unsuitable due to signal overlap resulting both from an increase in linewidth due to an imperfect TROSY effect and a greater number of signals from the greater number of amino acids. As an alternative to amide-TROSY, sidechain methyl groups were selectively labelled. Routinely, the ε-methyl groups of methionines as well as the δ1-methyl groups of isoleucines were $^{1}H^{13}C$ labelled by addition of labelled l-methionine-(methyl-^{13}C) and the isoleucine precursor 2-ketobutyrate-4-^{13}C to a minimal growth medium. To improve spectral quality, the CH_2 group protons in the isoleucine precursor were exchanged to deuterium by incubation of the precursor in D_2O at high pD before neutralization and use in the growth medium.

In methyl groups, the three equivalent (due to the usually fast rotation around the CH_2-CH_3 axis) protons provide an attractive sensitivity increase. In addition, while the cancellation between DD and CSA is negligible for methyl groups compared to the effect in ^{15}N TROSY, the HMQC (Heteronuclear Multiple-Quantum Correlation) experiment on methyl groups has nevertheless been shown to be optimal with respect to the TROSY principle of not mixing slow-relaxing (from the field-strength independent cancellation of intra-methyl dipolar interactions[249]) with fast-relaxing elements of the magnetisation[251,252].

The SOFAST (band-Selective Optimized Flip Angle Short Transient)-HMQC[253] employed here is an incremental improvement over FAST-HMQC (Flip-Angle Short-Transient HMQC)[254], which itself is an improvement over the classical HMQC pulse sequence[255–257]. FAST-HMQC introduces the use of a non-90° angle for the initial flip-angle called the Ernst angle, which is optimized for maximal signal over multiple repetitions of the pulse sequence via the observation that this depends on the recovered equilibrium magnetisation reached at the start of each individual repetition. It also introduces the watergate element, as well as a water selective flip back for water solvent signal suppression.

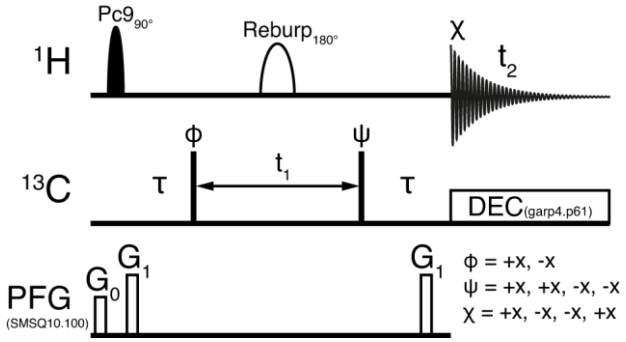


Figure 37. The SOFAST-HMQC pulse sequence used here. As indicated by the addition of "band-Selective Optimized" in SOFAST-HMQC, hard-pulses (on hydrogen) are replaced by (methyl-) band-selective pulses compared to a classical HMQC. The initial Ernst angle hard pulse and the following water flip back is replaced with a methyl-selective shaped pulse (Pc9), and the watergate element is simplified by using another methyl-selective pulse for the central 180° refocusing pulse (REBURP, Refocusing Band-Selective Pulse with Uniform Response and Phase).

3. CPMG RD (Carr-Purcell Meiboom-Gill Relaxation Dispersion)

Carr-Purcell Meiboom-Gill relaxation dispersion (CPMG RD) experiments can be used to measure the contribution of the chemical exchange (R_{ex}) to the transverse ("spin-spin") relaxation (R_2) ($R_{2,obs} = R_2 + R_{ex}$). In practice, the application of spin-echo elements, which refocus the effect of exchange broadening in the symmetric period before and after the central pair of 180° pulses, attempts to eliminate the chemical exchange contribution to the relaxation. By increasing the frequency of these spin-echo elements (ν_{CPMG}), relaxation due to exchange processes of higher and higher frequency can be refocused, eventually reaching a lower bound ($R_{2,obs} = R_2$)[198].

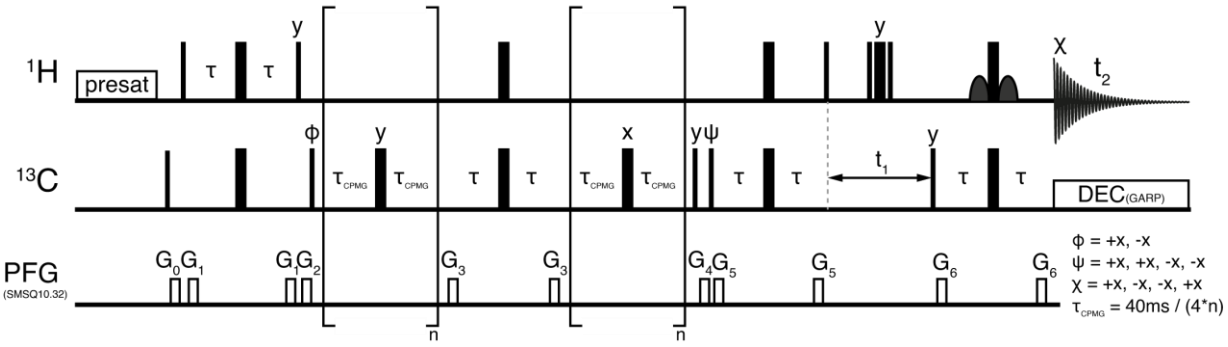


Figure 38. The single quantum Carr-Purcell Meiboom-Gill relaxation dispersion pulse sequence used here, developed by Lewis E. Kay et al. as "HtoC_CH3_exchange_*00_lek_ILV.c"[258], that has been rewritten in the pulse sequence format of Bruker spectrometers.

The sequence begins with a 90° pulse on ^{13}C to suppress remaining steady-state magnetisation[259]. After an INEPT block to generate single quantum magnetisation, the first of two CPMG blocks is applied. The CPMG blocks are centred around an element that exchanges the single quantum coherence from anti-phase to in-phase, which has the effect that differential relaxation resulting from interactions with non-exchanging, external spins (i.e. external solvent protons) is equalized[260–262]. The rest of the pulse sequence after the second CPMG element creates multi-quantum coherence, which is evolved for a time t_1 and ends with a watergate element before detection[258].

Practically, chemical exchange rates (k_{ex}) between 100 s^{-1} and 3000 s^{-1} can by determined via CPMG RD experiments (significant refocusing is reached at $\nu_{CPMG} = {k_{ex}}/{2}$)[198]. This timescale includes domain motions as well as folding events[198]. If the chemical exchange is not fast compared to the chemical shift difference between the states ($k_{ex} \gg \Delta\nu$), in addition to k_{ex} the populations of the exchanging states ($P_A;\ P_B = 1 - P_A$) can also be determined[198].

Here, spectra for ν_{CPMG} between 50 and 2000 Hz were typically recorded with a T_{CPMG} ($= 4 \cdot \tau_{CPMG}$) of 40 ms.

In analysing the data, we employed here the Luz-Meiboom fast-exchange equation[196–198], which is valid in the fast exchange regime. Additionally, for a more rigorous approach based on numerically solving the Bloch-McConnell equations[263,264], we employed the software ChemEx by Guillaume Bouvignies[265].

IV. Acknowledgements

I would like to express my thanks to my supervisor, Dr. Silke Wiesner, for an exciting research topic. She always found time for valuable support, encouragement, and advice throughout my PhD thesis.

I also would like to thank the members of the examination committee including the 2nd assessor Prof. Dr. Reinhard Sterner, the 3rd examiner Prof. Dr. Wolfgang Seufert and the Chairman, Prof. Dr. Frank Sprenger.

Special thanks go to Dr. Magnus Jäckl for the excellent collaboration on what makes up first chapter/project in this thesis and for always being available for input and encouraging discussions.

Thanks also go to the current and past members of the Wiesner group, especially Dr. Fabian Renschler, Dr. Natalia Ruetalo, Dr. Philip Rößler and Mira Schütz for being a joy to work with and for always being available to answer questions I had. Thanks go also to Jan Overbeck for proofreading the draft of the NMR Methods part. Besides for Jan, I would like to also thank the rest of the Sprangers group for their input during the lab meetings.

Most importantly, thanks also go to my parents who always supported me in anything I wanted to do allowing me to get where I am right now.

Finally, I would like to thank my love, Dr. Daniela Lazzaretti, who had great patience with me while I was sitting at home brooding over writing this thesis, not to mention her great help in proofreading it.

V. Appendix

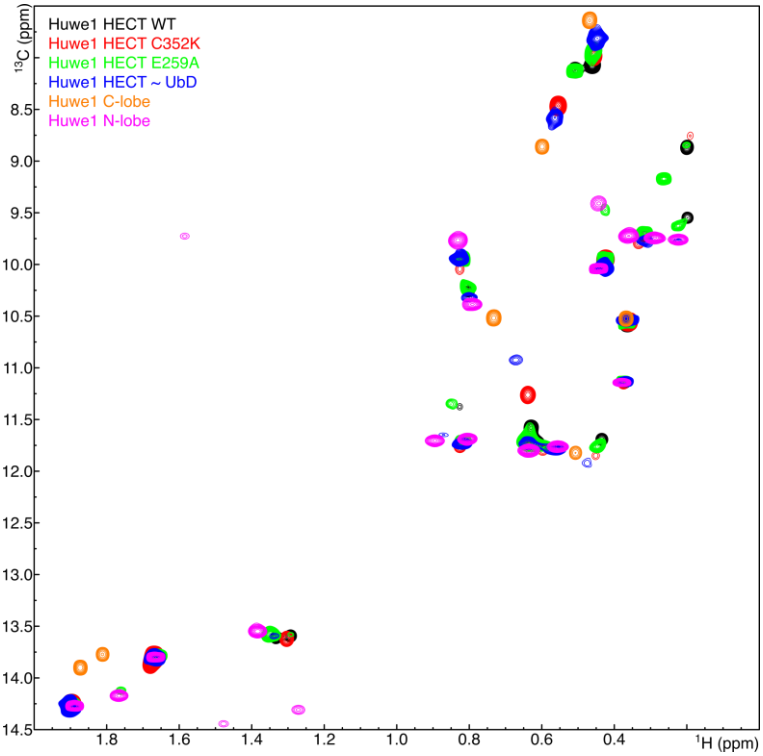


Figure S 7. SOFAST IM-^{13}C-Methyl labelled spectra (at 283 K) of the Huwe1 HECT-domain wild-type, the Huwe1 HECT C352K mutant, the E259A mutant, wild-type with donor-ubiquitin analogue disulfide linked to it as well as the individual C-lobe and the N-lobe (note, N-lobe also alanine labelled, extra with signals at ~14.5ppm ^{13}C). All in the cysteine reduced background construct.

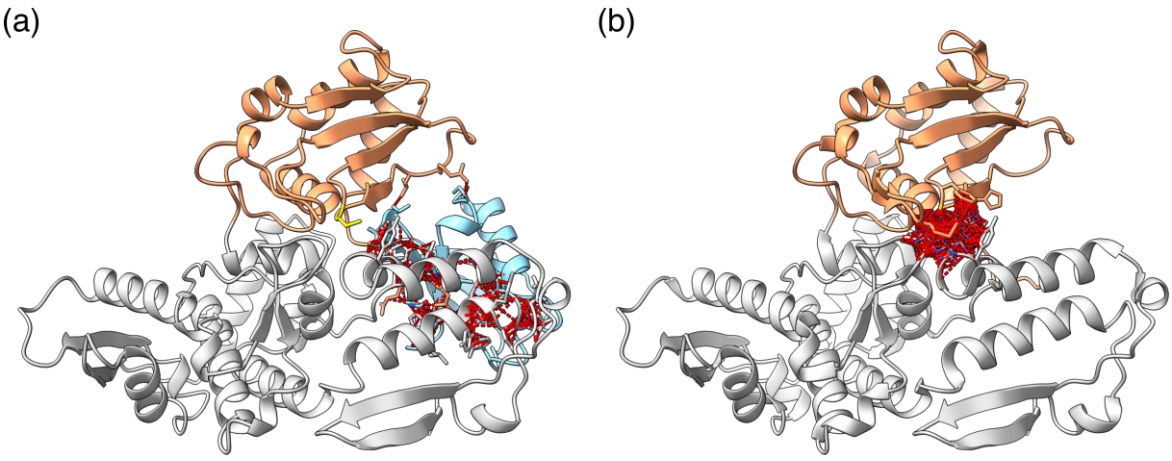


Figure S 8. (a) Structure of the Huwe1 HECT domain (PDB ID: 5LP8) overlaid with the structure of the Huwe1 C-lobe loaded with a donor ubiquitin (as disulfide mimic, PDB ID: 6FYH). N-lobe in gray, C-lobe in orange, Ubiquitin in blue, clashes between Ubiquitin and the HECT domain indicated in red. (b) Structure of the Huwe1 HECT domain (PDB ID: 5LP8) with the catalytic cysteine (C352) exchanged to rotamers of lysine (every entry of the Dunbrack2010[266] rotamer library) with clashes (VDW overlap > 0.6 Å) shown in red. Note: Every rotamer resulted in at least two clashes while the original cysteine did not clash.

EMBL-EBI PDBePISA[195] results:

N-lobe	Dist. [Å]	C-lobe
Hydrogen bonds		
ARG 141[NH2]	3.45	GLY 313[O]
ARG 99[NH1]	3.15	GLY 327[O]
ARG 99[NH1]	3.82	GLY 330[O]
SER 159[OG]	3.05	THR 351[OG1]
LYS 150[NZ]	2.73	GLU 377[OE1]
SER 151[N]	3.26	GLU 380[OE2]
LEU 265[O]	3.58	GLN 301[NE2]
GLU 257[OE2]	3.35	LYS 306[NZ]
GLN 256[O]	2.89	GLN 309[NE2]
GLU 259[OE2]	**3.16**	**THR 314[OG1]**
GLU 259[OE2]	**2.79**	**SER 315[N]**
GLU 259[OE2]	**2.60**	**SER 315[OG]**
GLY 97[O]	3.22	MET 328[N]
GLU 158[OE1]	2.97	HIS 350[ND1]
GLU 158[OE2]	3.00	HIS 350[ND1]
ASP 98[OD1]	2.75	THR 351[OG1]
Salt bridges		
LYS 150[NZ]	2.73	GLU 377[OE1]
LEU 265[O]	2.26	PRO 266[N]
GLU 257[OE2]	3.35	LYS 306[NZ]
GLU 158[OE1]	2.97	HIS 350[ND1]
GLU 158[OE2]	3.00	HIS 350[ND1]

Pymol script to split structure between C- and N-lobe

```
fetch 5lp8, async=0
remove /5lp8/A/B
remove solvent
remove hydrogens
remove /5lp8 and resi -3989
alter (all),resi=str(int(resi.replace("A",""))-3989)
alter (all),chain=str("B" if int(resi)>265 else "A")
alter (all),segi=""
save 5lp8_nlobe_clobe_chainsplit.pdb
```

Table S 3. (left) Results generated by an interface search on the 5LP8 (PDB ID) structure between its C- and N-lobes using the EMBL-EBI PDBePISA server. (right) The structures had to be split into separate chains for C- and N-lobe, a pymol script to do this is given.

Chapter B	Autoubiquitination	Thioester Formation
Ubiquitin*	60 μM	5 μM
Uba1	0.5 μM	0.33 μM
UbcH7	10 μM	10 μM
Huwe1^{HECT} (mutants)	3 μM	5 μM (Δ-4)

Chapter C	Autoubiquitination	Thioester Formation
Ubiquitin*	60 μM	10 μM / 60 μM
Uba1	0.5 μM	1.5 μM
UbcH7	12 μM	10 μM
Huwe1^{HECT} (mutants)	4 μM	15 μM (Δ-4)

Table S 4. Assay conditions for Autoubiquitination assays and Thioester Formation assays from chapters B and C. Assays were performed at 30°C. For thioester formation assays the E2 was pre-charged and quenched with EDTA by diluting 1:3 with quenching buffer (25 mM Tris pH 7.5, 50 mM NaCl, 50 mM EDTA). Given concentrations are the final concentrations under which E3 thioester formation was performed (after 1:3 dilution).

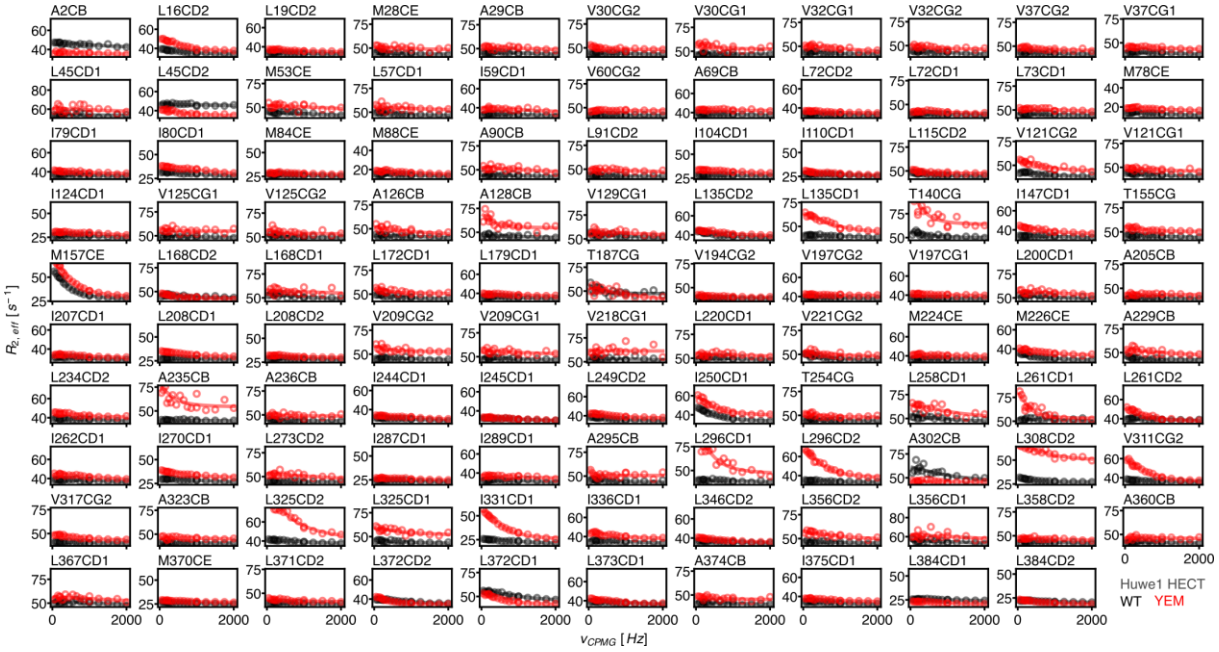


Figure S 9. CPMG relaxation dispersion profiles of all assigned residues for AILMTV labelled spectra of Huwe1 HECT WT (black) as well as the Y162A/E259A/M328A triple mutant (red) recorded on an 800 MHz spectrometer. Markers represent data (R2 effective at given ν_{CPMG}), while lines are residue and sample individual fits against the Luz-Meiboom fast-exchange equation.

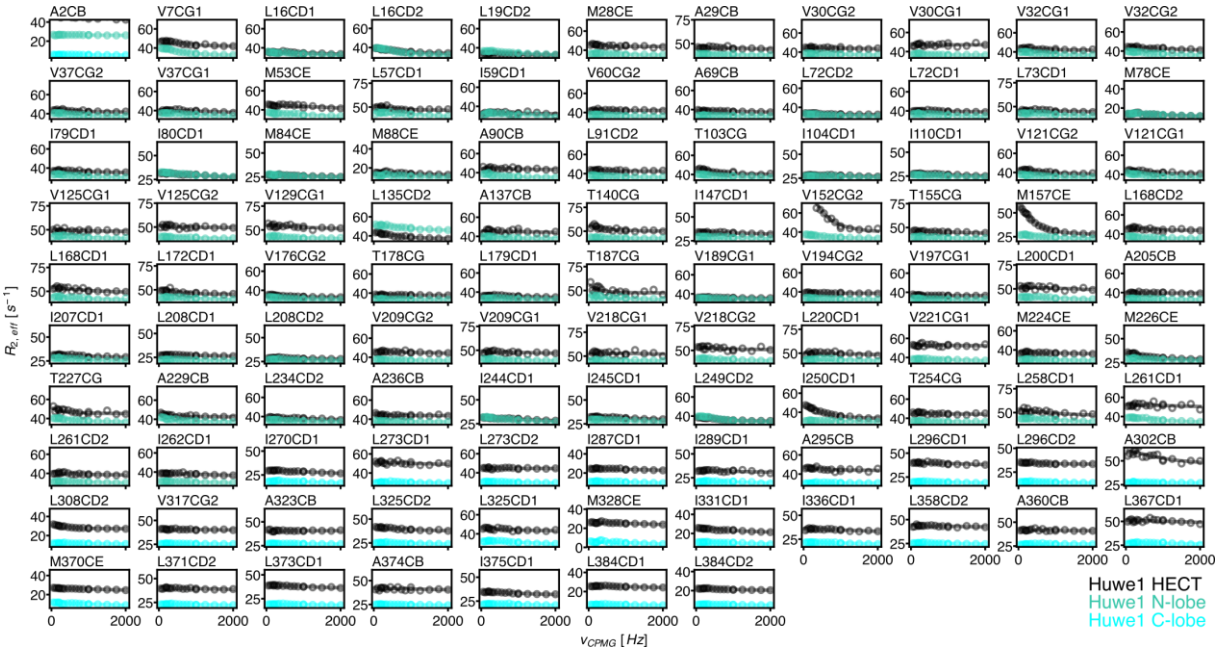


Figure S 10 CPMG relaxation dispersion profiles of all assigned residues for AILMTV labelled spectra of Huwe1 HECT WT (black) as well as the Huwe1 N-lobe (teal) and the Huwe1 C-lobe (turquoise) recorded on an 800 MHz spectrometer. Markers represent data (R2 effective at given vCPMG), while lines are residue and sample individual fits against the Luz-Meiboom fast-exchange equation.

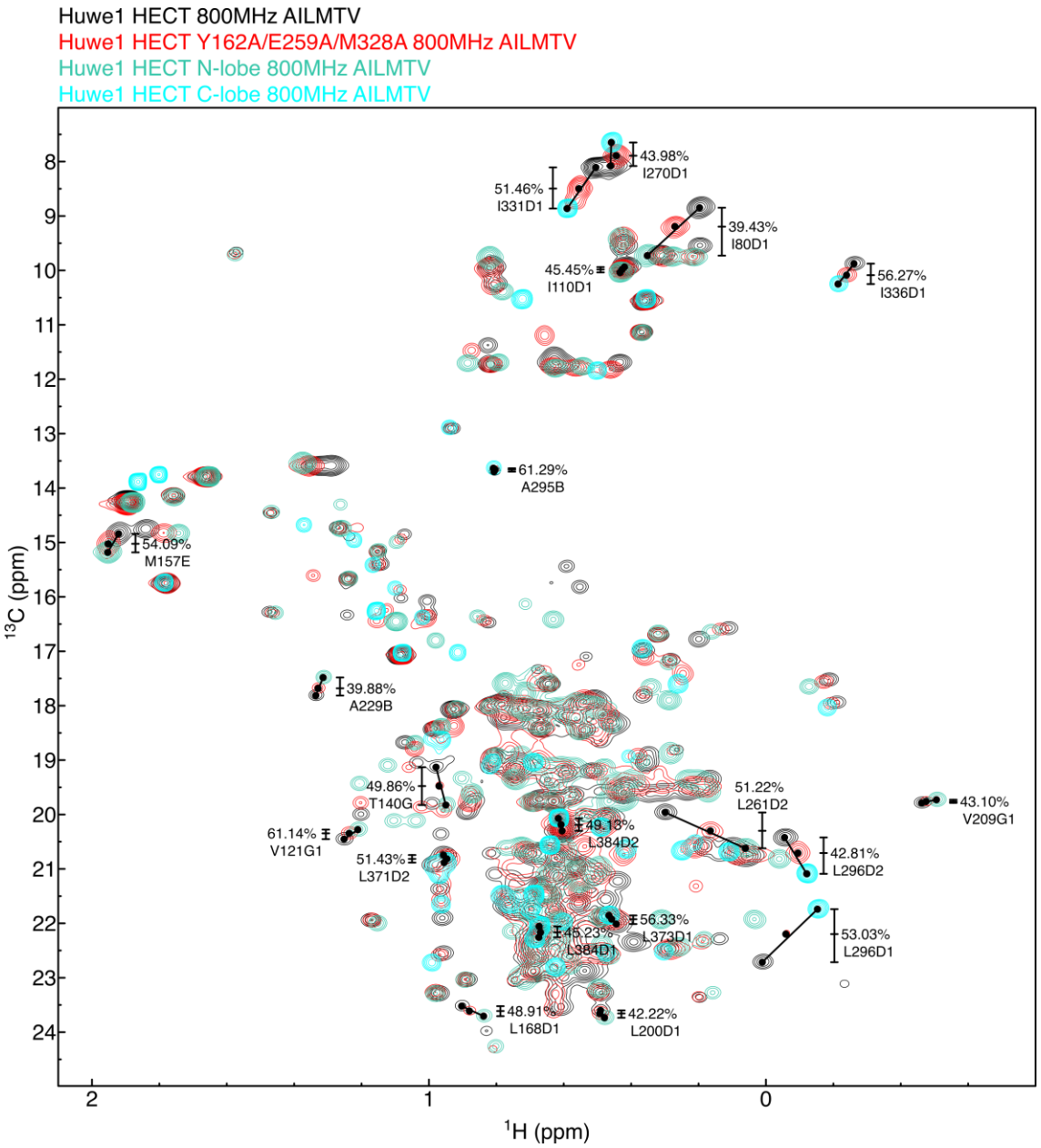


Figure S 11. Spectra of the AILMTV labelled Huwe1 HECT WT (black), Huwe1 HECT Y162A/E259A/M328A (red) and the Huwe1 N-lobe (teal) and C-lobe (turquoise) overlaid, all recorded on an 800Mhz spectrometer. Labels are methyl group signals whose ^{13}C shift of the Huwe1 HECT Y162A/E259A/M328A sample is in the middle (between 0.4 to 0.6) between the respective signals of the Huwe1 HECT WT and the N-/C-lobe spectra.

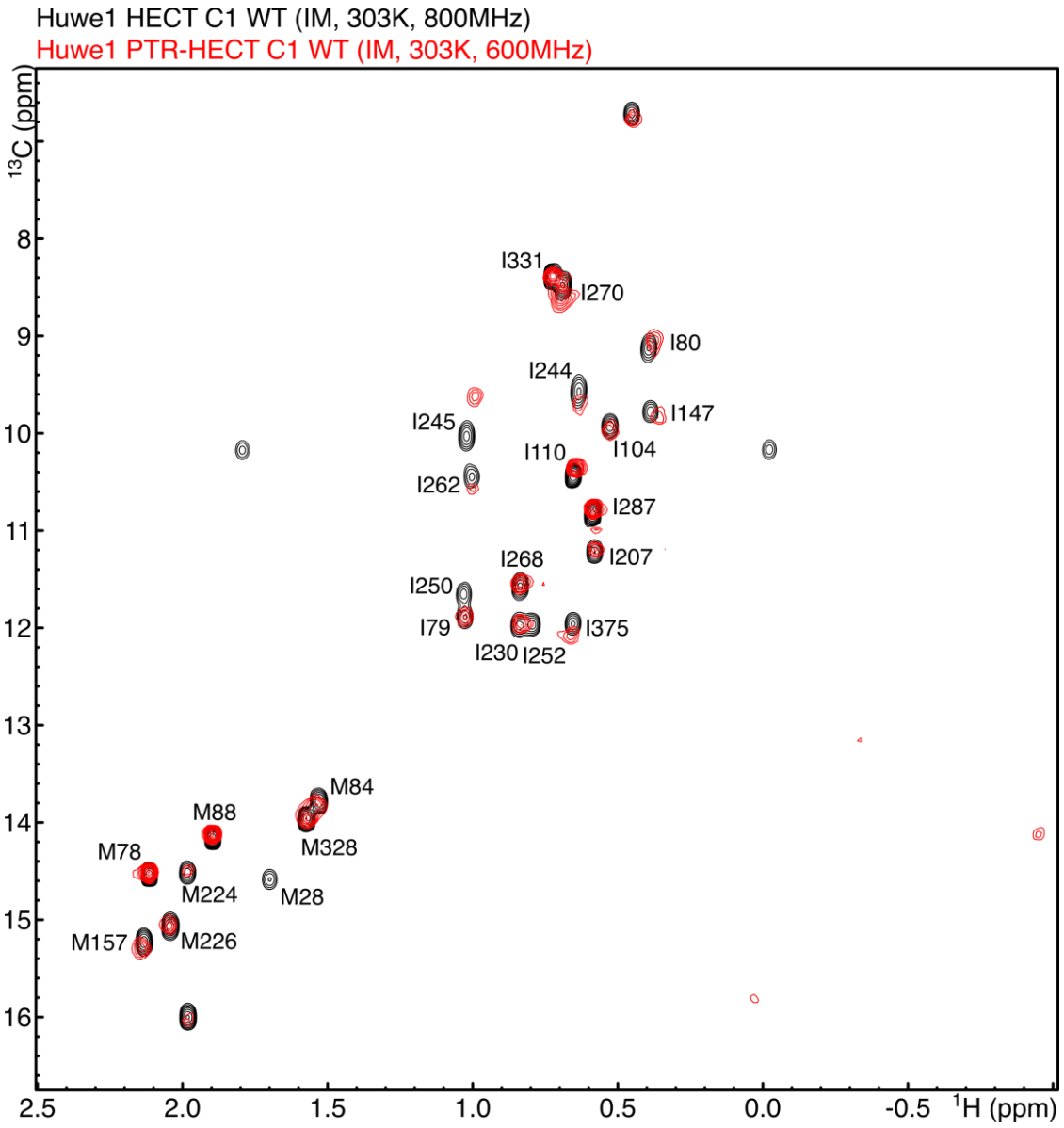


Figure S 12. SOFAST IM-^{13}C-Methyl labelled spectra (at 303 K) of the Huwe1 HECT-domain wild-type (3992-4374) and a Huwe1 HECT-domain wild-type preceded by the pointer and thumb helices (3951-4374) (construct as in Sander et al. 2017[113]). Both in the cysteine reduced background construct.

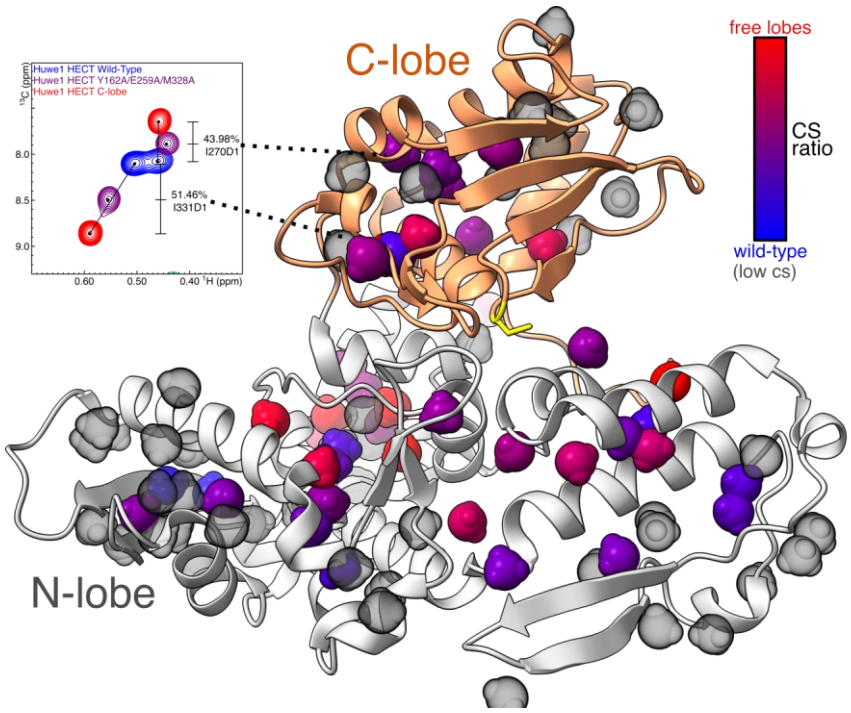


*Figure S 13. Huwe1 HECT (PDB-ID: 5LP8) structure with labelled methyl groups depicted in: Transparent/grey: methyl groups which show no significant change in chemical shift (^{13}C & ^{1}H * 4 < 0.2ppm) between wild-type and separated lobes. Blue to red: how close the signals in the open-state (Y162A/E259A/M328A) mutant are with respect to either the wild-type HECT (blue) or the individual lobe (red) spectra.*

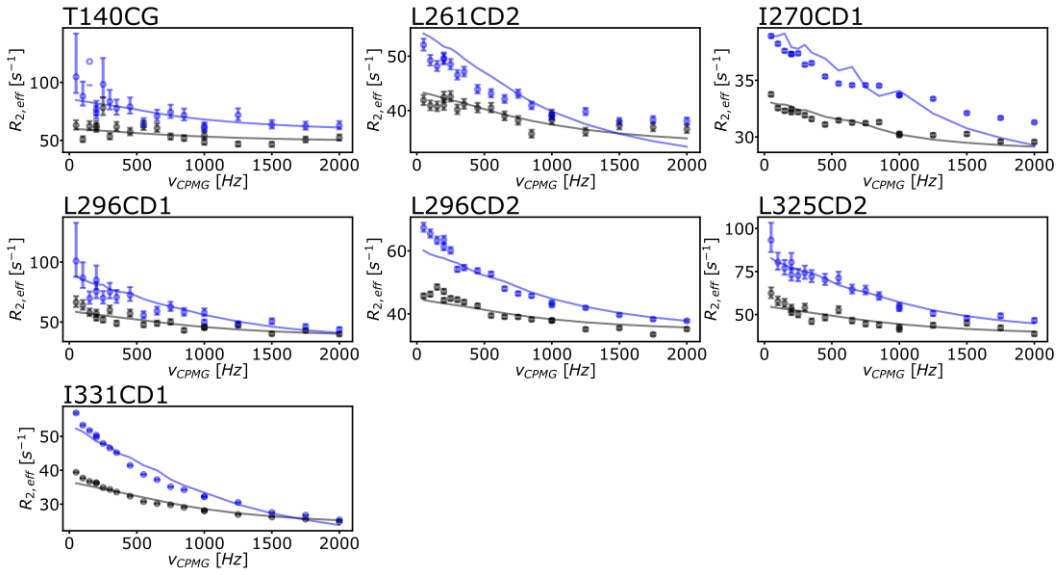


Figure S 14. CPMG relaxation dispersion curves of the Huwe1 HECT Y162A/E259A/M328A residues, recorded at 283 K on the 800 MHz spectrometer (blue), as well as a 500 MHz spectrometer (black). Markers represent data points of labeled residues methyl group signals (effective R2), lines are back calculated from the result of a global numeric fit of all shown residues using the program ChemEx assuming a 2-site exchange model with P_A (and P_B) set to 50%, and the Δω's to their differences between the wild-type and the free lobes.

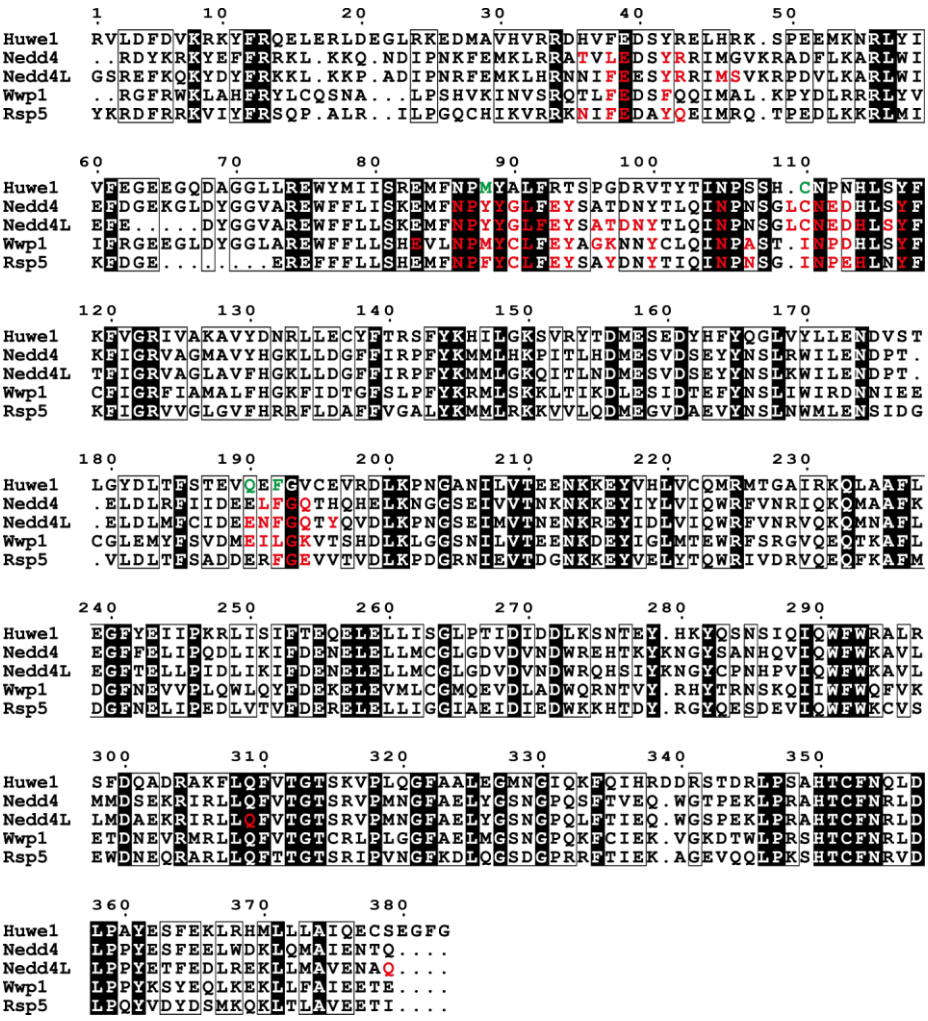


Figure S 15. Structure aided alignment using the tool PROMALS3D[267] between Huwe1 HECT (PDB ID: 5LP8) and HECT structures co-crystalized with ubiquitin in the ubiquitin-binding-site (Nedd4 HECT PDB ID: 2XBB; Nedd4L HECT PDB ID: 5HPK; Wwp1 HECT PDB ID: 5HPT; Rsp5 HECT PDB ID: 3OLM) rendered using the tool ESPript 3.0[268]. Residues marked in red are part of the HECT – ubiquitin interface in the respective co-crystal structures as determined using the PDBePISA server[195]. Huwe1 residues marked in green are the ubiquitin-binding-site mutants presented in Chapter C.

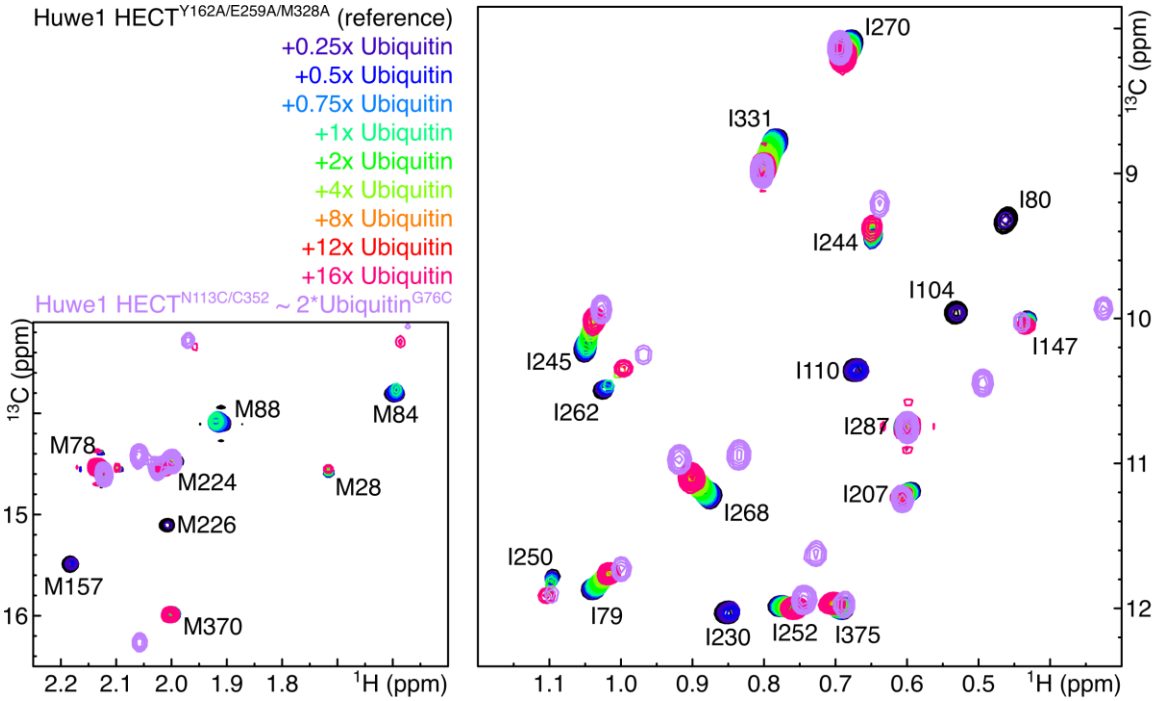


Figure S 16. SOFAST IM-^{13}C-Methyl labelled spectra (at 303 K) of Huwe1 HECT$^{Y126A/E259A/M328A}$ triple mutant without (reference) and with Ubiquitin at stochiometric ratios of 0.25x to 16x added. Additionally overlaid by a spectrum of Huwe1 HECT$^{N113C/C352}$ with both cysteine residues disulfide linked to Ubiquitin^{G76C}. The N113C mutant was chosen with the goal that it links ubiquitin in a way in which it can bind the ubiquitin-binding-site. Spectral similarity with an imaginary endpoint of the titration validates this double disulfide as a structural analogue of a ubiquitin-binding-site at full occupancy.

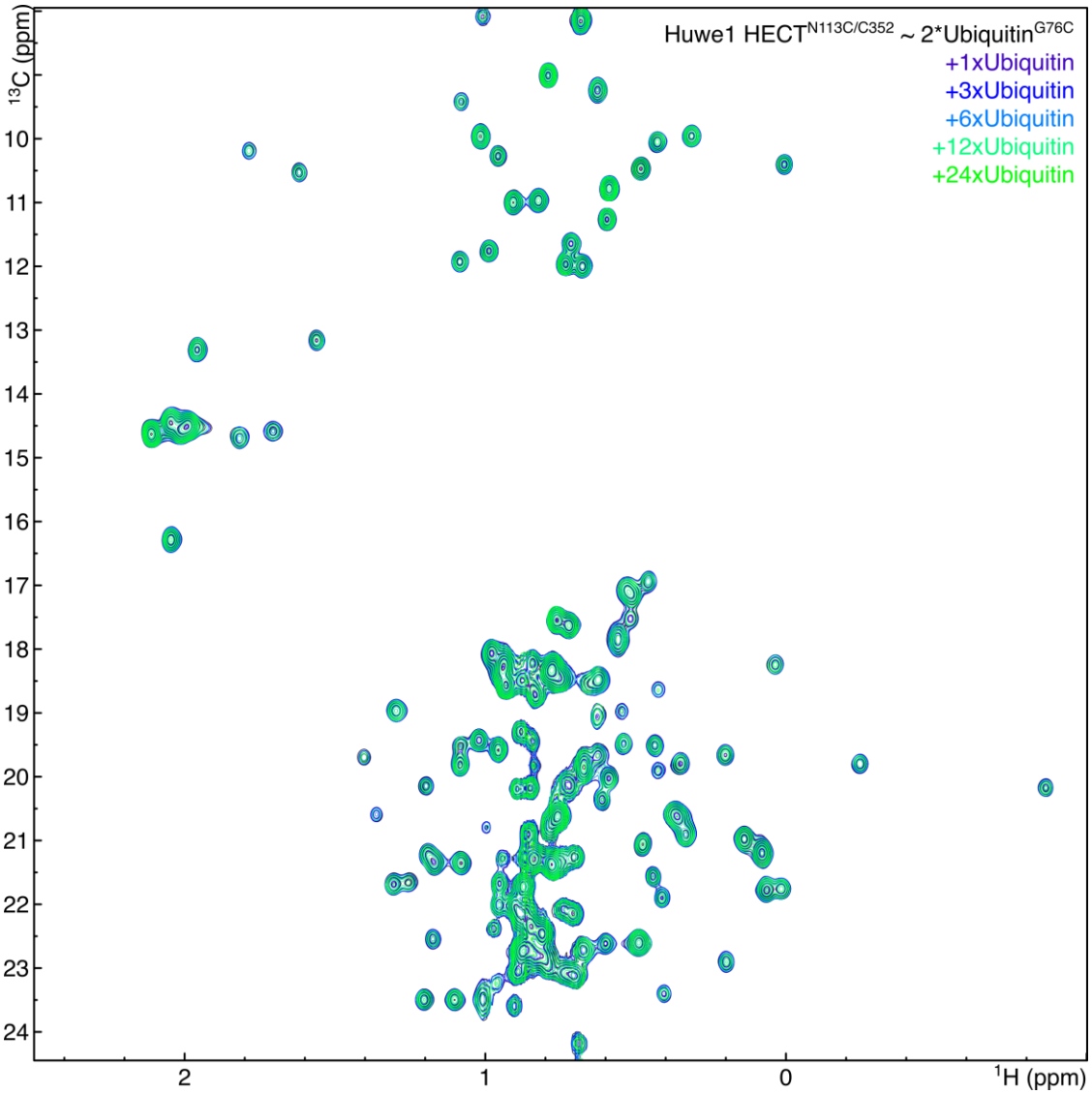


*Figure S 17. SOFAST IMLV-^{13}C-Methyl labelled spectra (at 303 K) of Huwe1 HECT$^{N113C/C352}$ ~ 2*Ubiquitin^{G76C} double disulfide titrated with (free) Ubiquitin at the indicated molecular ratios.*

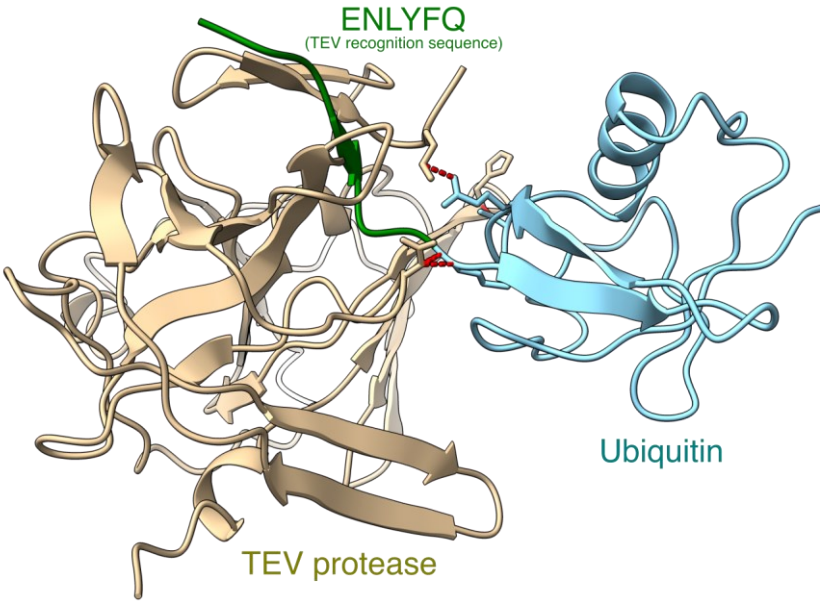


Figure S 18. The program CYANA was used calculate a minimum energy structure of the TEV protease (PDB: 1LVB) and ubiquitin (PDB: 1UBQ) fused to the TEV recognition peptide (ENLYFQS), as present in the TEV structure. For the modelling, the amino acids QS at the end of the recognition peptide adjacent to ubiquitin were kept flexible, while the rest of the TEV recognition peptide bound to TEV as well as ubiquitin were kept as two rigid bodies. Clashes could not be resolved by CYANA. The lowest energy structure is depicted with clashes, determined by ChimeraX, highlighted in red.

Tag	**Sequence**	**Ex. Length [~Å] (N*3.5)**	**[~Å] (source PDB ID)**
Flag	M*DY**K**DDDD**K***GS	28	12 (4WV9)
V5	*G**K**PIPNPLLGLDST*	49	14.7 (2B5L)
StrepII	M*WSHPQFE**K***GS	28	25 (4U4H)
cMyc	*MEQ**K**LISEEDLG*S	35	25.2 (2XA7)
GAMG	GAMG	14	
AAA	AAA	10.5	
AA	AA	7	
A	A	3.5	

Figure S 19. Epitope tags and other sequences used. Exact tag sequences used are given, with the minimal epitope sequence in italic and lysines shown in bold. Length was calculated for the maximally extended sequence (third column) and from the representative structures indicated (fourth column). Atom distance was calculated between the first amino acid N to last amino acid C.

		Rsp5	Nedd4	Smurf2
		120 μM Ubiquitin (various)		
Auto-ubiquitination	E1: Uba1	1.00 μM	1.00 μM	1.00 μM
	E2:	8.00 μM Ubc4	8.00 μM UbcH7	5.00 μM UbcH7
	Acceptor/E3: HA-HECT	4.00 μM	4.00 μM (H6-HA-tag)	10.00 μM
	Temperature:	25°C	25°C	25°C
	Time:	15 min	15 min	16 min
Substrate Ubiquitination	E1: Uba1	0.50 μM	0.50 μM	1.00 μM
	E2:	1.00 μM Ubc4	2.00 μM UbcH7	2.00 μM UbcH7
	E3: WW-HECT	0.25 μM	1.00 μM	1.00 μM
	Acceptor: GB1K0*-Rim8ct.	4 μM	4 μM	20 μM
	Temperature:	25°C	37°C	30°C
	Time:	30 min	13 min	90 min
Free Chain Formation	E1: Uba1	1.00 μM	1.00 μM	
	E2:	1.00 μM Ubc4	1.00 μM UbcH7	
	E3: HECT	0.50 μM	0.50 μM (H6-HA-tag)	
	Acceptor: FL-C-hUbWTdGG	12 μM	12 μM	
	Temperature:	32°C	37°C	
	Time:	15 min	15 min	

Table S 5. Assay conditions for non-native chain formation using ubiquitin variants.

	conc. [μM]
Uba1	10
Ubc4	15
$Rsp5^{HECT}$ / $Rsp5^{WW3\text{-}HECT}$	0,02
GB1K0*-Rim8ct.	10
hUbK0 / hUbWT	100

Table S 6. General assay conditions.

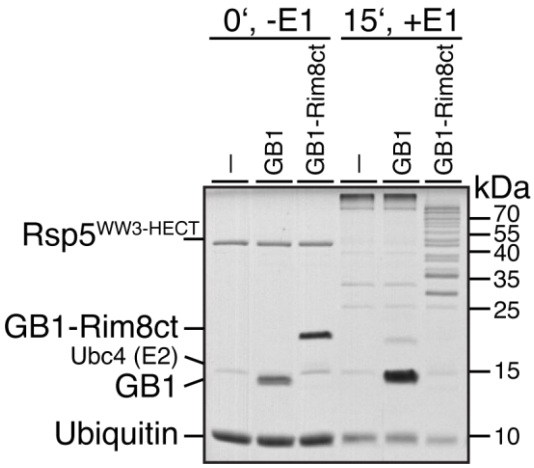
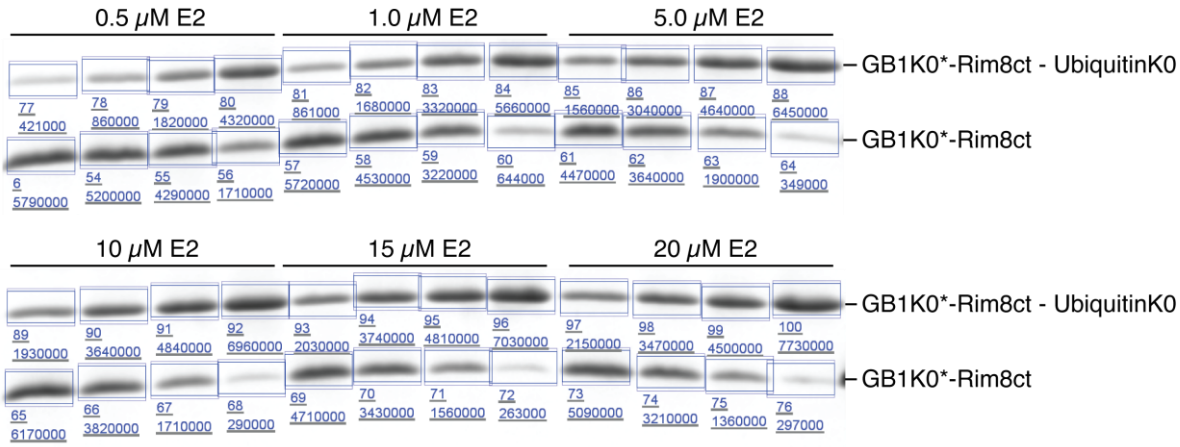
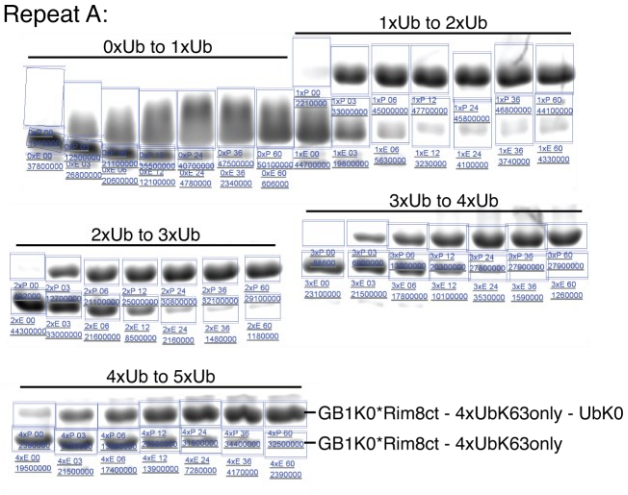


Figure S 20. Comparison of the utilization of GB1 alone and the GB1-Rim8ct substrate. Bands were detected using Coomassie staining. After 15 minutes, no non-ubiquitinated GB1-Rim8 remains, whereas GB1 is unaffected. The auto-ubiquitination of $Rsp5^{WW3\text{-}HECT}$ serves as an internal control showing ubiquitin ligase activity even when no substrate is available. Samples included 2.5 µM $Rsp5^{WW3\text{-}HECT}$, 2.5 µM Ubc4, 0.5 µM Uba1, 80 µM Ubiquitin (GAMG-hUbK0), 15 µM GB1-Rim8ct or 40 µM GB1.



⅔ E2	E1 (Uba1)
varying	E2 (Ubc4)
20 nM	E3 ($Rsp5^{WW3\text{-}HECT}$)
10 μM	Substrate (GB1K0*-Rim8ct)
100 μM	Ubiquitin (UbiquitinK0)

Figure S 21. Fluorescence readout, integrated regions for signal quantification, and enzyme/substrate concentrations used for the assays in Figure 32.



10 μM	E1 (Uba1)
15 μM	E2 (Ubc4)
20 nM	E3 ($Rsp5^{WW3\text{-}HECT}$)
10 μM	Substrate (GB1K0*-Rim8ct - NxUbiquitinK63only)
100 μM	Ubiquitin (UbiquitinK0)

Figure S 22. Fluorescence readout, integrated regions for signal quantification, and enzyme/substrate concentrations used for the assays in Figure 33.

VI. References

1. Goldstein, G. Isolation of Bovine Thymin: a Polypeptide Hormone of the Thymus. *Nature* **247**, 11–14 (1974).
2. Goldstein, G. *et al.* Isolation of a polypeptide that has lymphocyte-differentiating properties and is probably represented universally in living cells. *Proc Natl Acad Sci U S A* **72**, 11–15 (1975).
3. Ciechanover, A. Intracellular Protein Degradation: From a Vague Idea, through the Lysosome and the Ubiquitin–Proteasome System, and onto Human Diseases and Drug Targeting (Nobel Lecture). *Angewandte Chemie International Edition* **44**, 5944–5967 (2005).
4. Schlesinger, D. H., Goldstein, G. & Niall, H. D. Complete amino acid sequence of ubiquitin, an adenylate cyclase stimulating polypeptide probably universal in living cells. *Biochemistry* **14**, 2214–2218 (1975).
5. Goldknopf, I. L. & Busch, H. Remarkable similarities of peptide fingerprints of histone 2A and nonhistone chromosomal protein A24. *Biochemical and Biophysical Research Communications* **65**, 951–960 (1975).
6. Goldknopf, I. L. & Busch, H. Isopeptide linkage between nonhistone and histone 2A polypeptides of chromosomal conjugate-protein A24. *Proc Natl Acad Sci U S A* **74**, 864–868 (1977).
7. Hunt, L. T. & Dayhoff, M. O. Amino-terminal sequence identity of ubiquitin and the nonhistone component of nuclear protein A24. *Biochemical and Biophysical Research Communications* **74**, 650–655 (1977).
8. Low, T. L. & Goldstein, A. L. The chemistry and biology of thymosin. II. Amino acid sequence analysis of thymosin alpha1 and polypeptide beta1. *J. Biol. Chem.* **254**, 987–995 (1979).
9. Ciehanover, A., Hod, Y. & Hershko, A. A heat-stable polypeptide component of an ATP-dependent proteolytic system from reticulocytes. *Biochemical and Biophysical Research Communications* **81**, 1100–1105 (1978).

10. Hershko, A., Ciechanover, A. & Rose, I. A. Resolution of the ATP-dependent proteolytic system from reticulocytes: a component that interacts with ATP. *Proc Natl Acad Sci U S A* **76**, 3107–3110 (1979).
11. Ciechanover, A., Elias, S., Heller, H., Ferber, S. & Hershko, A. Characterization of the heat-stable polypeptide of the ATP-dependent proteolytic system from reticulocytes. *J. Biol. Chem.* **255**, 7525–7528 (1980).
12. Wilkinson, K. D., Urban, M. K. & Haas, A. L. Ubiquitin is the ATP-dependent proteolysis factor I of rabbit reticulocytes. *J. Biol. Chem.* **255**, 7529–7532 (1980).
13. Ciechanover, A., Heller, H., Katz-Etzion, R. & Hershko, A. Activation of the heat-stable polypeptide of the ATP-dependent proteolytic system. *Proceedings of the National Academy of Sciences of the United States of America* **78**, 761 (1981).
14. Hershko, A. & Ciechanover, A. Mechanisms of Intracellular Protein Breakdown. *Annu. Rev. Biochem.* **51**, 335–364 (1982).
15. Ciechanover, A., Elias, S., Heller, H. & Hershko, A. 'Covalent affinity' purification of ubiquitin-activating enzyme. *J. Biol. Chem.* **257**, 2537–2542 (1982).
16. Hershko, A., Heller, H., Elias, S. & Ciechanover, A. Components of ubiquitin-protein ligase system. Resolution, affinity purification, and role in protein breakdown. *J. Biol. Chem.* **258**, 8206–8214 (1983).
17. Wilkinson, K. D. The discovery of ubiquitin-dependent proteolysis. *Proc Natl Acad Sci U S A* **102**, 15280–15282 (2005).
18. Ciechanover, A., Heller, H., Elias, S., Haas, A. L. & Hershko, A. ATP-dependent conjugation of reticulocyte proteins with the polypeptide required for protein degradation. *Proceedings of the National Academy of Sciences of the United States of America* **77**, 1365 (1980).
19. Hershko, A., Ciechanover, A., Heller, H., Haas, A. L. & Rose, I. A. Proposed role of ATP in protein breakdown: conjugation of protein with multiple chains of the polypeptide of ATP-dependent

proteolysis. *Proceedings of the National Academy of Sciences of the United States of America* **77**, 1783 (1980).

20. Hershko, A., Ciechanover, A. & Rose, I. A. Identification of the active amino acid residue of the polypeptide of ATP-dependent protein breakdown. *J. Biol. Chem.* **256**, 1525–1528 (1981).
21. Wilkinson, K. D. & Audhya, T. K. Stimulation of ATP-dependent proteolysis requires ubiquitin with the COOH-terminal sequence Arg-Gly-Gly. *J. Biol. Chem.* **256**, 9235–9241 (1981).
22. Wang, Y., Argiles-Castillo, D., Kane, E. I., Zhou, A. & Spratt, D. E. HECT E3 ubiquitin ligases – emerging insights into their biological roles and disease relevance. *J Cell Sci* **133**, jcs228072 (2020).
23. Deol, K. K., Lorenz, S. & Strieter, E. R. Enzymatic Logic of Ubiquitin Chain Assembly. *Front Physiol* **10**, (2019).
24. Hochstrasser, M. Lingering Mysteries of Ubiquitin-Chain Assembly. *Cell* **124**, 27–34 (2006).
25. Wang, M. & Pickart, C. M. Different HECT domain ubiquitin ligases employ distinct mechanisms of polyubiquitin chain synthesis. *EMBO J.* **24**, 4324–4333 (2005).
26. Verdecia, M. A. *et al.* Conformational flexibility underlies ubiquitin ligation mediated by the WWP1 HECT domain E3 ligase. *Mol. Cell* **11**, 249–259 (2003).
27. Ronchi, V. P., Kim, E. D., Summa, C. M., Klein, J. M. & Haas, A. L. In silico modeling of the cryptic E2∼ubiquitin-binding site of E6-associated protein (E6AP)/UBE3A reveals the mechanism of polyubiquitin chain assembly. *J. Biol. Chem.* **292**, 18006–18023 (2017).
28. Masuda, Y. *et al.* En bloc transfer of polyubiquitin chains to PCNA in vitro is mediated by two different human E2–E3 pairs. *Nucleic Acids Res* **40**, 10394–10407 (2012).
29. Li, W., Tu, D., Brunger, A. T. & Ye, Y. A ubiquitin ligase transfers preformed polyubiquitin chains from a conjugating enzyme to a substrate. *Nature* **446**, 333–337 (2007).
30. Todaro, D. R., Augustus-Wallace, A. C., Klein, J. M. & Haas, A. L. 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 (2017).

31. Todaro, D. R., Augustus-Wallace, A. C., Klein, J. M. & Haas, A. L. Oligomerization of the HECT ubiquitin ligase NEDD4-2/NEDD4L is essential for polyubiquitin chain assembly. *J Biol Chem* **293**, 18192–18206 (2018).

32. Edwards, D. J., Streich, F. C., Ronchi, V. P., Todaro, D. R. & Haas, A. L. Convergent evolution in the assembly of polyubiquitin degradation signals by the Shigella flexneri IpaH9.8 ligase. *J Biol Chem* **289**, 34114–34128 (2014).

33. Ronchi, V. P., Klein, J. M. & Haas, A. L. E6AP/UBE3A Ubiquitin Ligase Harbors Two E2∼ubiquitin Binding Sites. *J Biol Chem* **288**, 10349–10360 (2013).

34. Haldeman, M. T., Xia, G., Kasperek, E. M. & Pickart, C. M. Structure and Function of Ubiquitin Conjugating Enzyme E2-25K: The Tail Is a Core-Dependent Activity Element. *Biochemistry* **36**, 10526–10537 (1997).

35. Ravid, T. & Hochstrasser, M. Autoregulation of an E2 enzyme by ubiquitin-chain assembly on its catalytic residue. *Nature Cell Biology* **9**, 422–427 (2007).

36. Bazirgan, O. A. & Hampton, R. Y. Cue1p is an activator of Ubc7p E2 activity in vitro and in vivo. *J Biol Chem* **283**, 12797–12810 (2008).

37. Singh, R. K., Kazansky, Y., Wathieu, D. & Fushman, D. Hydrophobic Patch of Ubiquitin is Important for its Optimal Activation by Ubiquitin Activating Enzyme E1. *Anal. Chem.* **89**, 7852–7860 (2017).

38. Blount, J. R., Libohova, K., Silva, G. M. & Todi, S. V. Isoleucine 44 Hydrophobic Patch Controls Toxicity of Unanchored, Linear Ubiquitin Chains through NF-κB Signaling. *Cells* **9**, (2020).

39. Yao, T. & Cohen, R. E. A cryptic protease couples deubiquitination and degradation by the proteasome. *Nature* **419**, 403–407 (2002).

40. Blount, J. R., Libohova, K., Marsh, G. B., Sutton, J. R. & Todi, S. V. Expression and Regulation of Deubiquitinase-Resistant, Unanchored Ubiquitin Chains in Drosophila. *Scientific Reports* **8**, 8513 (2018).

41. Komander, D. & Rape, M. The Ubiquitin Code. *Annual Review of Biochemistry* **81**, 203–229 (2012).

42. Kirisako, T. *et al.* A ubiquitin ligase complex assembles linear polyubiquitin chains. *EMBO J* **25**, 4877–4887 (2006).

43. Chau, V. *et al.* A multiubiquitin chain is confined to specific lysine in a targeted short-lived protein. *Science* **243**, 1576–1583 (1989).

44. Thrower, J. S., Hoffman, L., Rechsteiner, M. & Pickart, C. M. Recognition of the polyubiquitin proteolytic signal. *The EMBO Journal* **19**, 94–102 (2000).

45. Peth, A., Uchiki, T. & Goldberg, A. L. ATP-Dependent Steps in the Binding of Ubiquitin Conjugates to the 26S Proteasome that Commit to Degradation. *Molecular Cell* **40**, 671–681 (2010).

46. Lander, G. C. *et al.* Complete subunit architecture of the proteasome regulatory particle. *Nature* **482**, 186–191 (2012).

47. Matyskiela, M. E., Lander, G. C. & Martin, A. Conformational switching of the 26S proteasome enables substrate degradation. *Nat Struct Mol Biol* **20**, 781–788 (2013).

48. Sakata, E. *et al.* Localization of the proteasomal ubiquitin receptors Rpn10 and Rpn13 by electron cryomicroscopy. *PNAS* **109**, 1479–1484 (2012).

49. Lu, Y., Lee, B., King, R. W., Finley, D. & Kirschner, M. W. Substrate degradation by the proteasome: A single-molecule kinetic analysis. *Science* **348**, (2015).

50. Peng, J. *et al.* A proteomics approach to understanding protein ubiquitination. *Nature Biotechnology* **21**, 921–926 (2003).

51. Finley, D. *et al.* Inhibition of proteolysis and cell cycle progression in a multiubiquitination-deficient yeast mutant. *Mol Cell Biol* **14**, 5501–5509 (1994).

52. Wang, J. & Maldonado, M. A. The ubiquitin-proteasome system and its role in inflammatory and autoimmune diseases. *Cell Mol Immunol* **3**, 255–261 (2006).

53. Emmerich, C. H. *et al.* Activation of the canonical IKK complex by K63/M1-linked hybrid ubiquitin chains. *PNAS* **110**, 15247–15252 (2013).

54. Schwertman, P., Bekker-Jensen, S. & Mailand, N. Regulation of DNA double-strand break repair by ubiquitin and ubiquitin-like modifiers. *Nature Reviews Molecular Cell Biology* **17**, 379–394 (2016).

55. Silva, G. M., Finley, D. & Vogel, C. K63 polyubiquitination is a new modulator of the oxidative stress response. *Nat Struct Mol Biol* **22**, 116–123 (2015).

56. Erpapazoglou, Z., Walker, O. & Haguenauer-Tsapis, R. Versatile Roles of K63-Linked Ubiquitin Chains in Trafficking. *Cells* **3**, 1027–1088 (2014).

57. Huang, F. *et al.* Lysine 63-linked polyubiquitination is required for EGF receptor degradation. *PNAS* **110**, 15722–15727 (2013).

58. Kavsak, P. *et al.* Smad7 binds to Smurf2 to form an E3 ubiquitin ligase that targets the TGF beta receptor for degradation. *Mol Cell* **6**, 1365–1375 (2000).

59. Lauwers, E., Erpapazoglou, Z., Haguenauer-Tsapis, R. & André, B. The ubiquitin code of yeast permease trafficking. *Trends Cell Biol* **20**, 196–204 (2010).

60. Haglund, K. *et al.* Multiple monoubiquitination of RTKs is sufficient for their endocytosis and degradation. *Nat Cell Biol* **5**, 461–466 (2003).

61. Lauwers, E., Jacob, C. & André, B. K63-linked ubiquitin chains as a specific signal for protein sorting into the multivesicular body pathway. *J Cell Biol* **185**, 493–502 (2009).

62. Weiss, E. R. *et al.* Rescue of HIV-1 Release by Targeting Widely Divergent NEDD4-Type Ubiquitin Ligases and Isolated Catalytic HECT Domains to Gag. *PLOS Pathogens* **6**, e1001107 (2010).

63. Spit, M., Rieser, E. & Walczak, H. Linear ubiquitination at a glance. *J Cell Sci* **132**, (2019).

64. Oeckinghaus, A. & Ghosh, S. The NF-κB Family of Transcription Factors and Its Regulation. *Cold Spring Harb Perspect Biol* **1**, (2009).

65. Swatek, K. N. & Komander, D. Ubiquitin modifications. *Cell Res.* **26**, 399–422 (2016).

66. Grice, G. L. & Nathan, J. A. The recognition of ubiquitinated proteins by the proteasome. *Cell. Mol. Life Sci.* **73**, 3497–3506 (2016).

67. Grice, G. L. *et al.* The Proteasome Distinguishes between Heterotypic and Homotypic Lysine-11-Linked Polyubiquitin Chains. *Cell Rep* **12**, 545–553 (2015).

68. Meyer, H.-J. & Rape, M. Enhanced protein degradation by branched ubiquitin chains. *Cell* **157**, 910–921 (2014).

69. Kim, W. *et al.* Systematic and quantitative assessment of the ubiquitin modified proteome. *Mol Cell* **44**, 325–340 (2011).

70. Bingol, B. *et al.* The mitochondrial deubiquitinase USP30 opposes parkin-mediated mitophagy. *Nature* **510**, 370–375 (2014).

71. Cunningham, C. N. *et al.* USP30 and parkin homeostatically regulate atypical ubiquitin chains on mitochondria. *Nat Cell Biol* **17**, 160–169 (2015).

72. Gatti, M. *et al.* RNF168 Promotes Noncanonical K27 Ubiquitination to Signal DNA Damage. *Cell Reports* **10**, 226–238 (2015).

73. Besche, H. C. *et al.* Autoubiquitination of the 26S Proteasome on Rpn13 Regulates Breakdown of Ubiquitin Conjugates. *EMBO J* **33**, 1159–1176 (2014).

74. Jin, J. *et al.* Epigenetic regulation of the expression of Il12 and Il23 and autoimmune inflammation by the deubiquitinase Trabid. *Nat Immunol* **17**, 259–268 (2016).

75. David, Y., Ziv, T., Admon, A. & Navon, A. The E2 ubiquitin-conjugating enzymes direct polyubiquitination to preferred lysines. *J Biol Chem* **285**, 8595–8604 (2010).

76. Moraes, T. F. *et al.* Crystal structure of the human ubiquitin conjugating enzyme complex, hMms2-hUbc13. *Nat Struct Biol* **8**, 669–673 (2001).

77. VanDemark, A. P., Hofmann, R. M., Tsui, C., Pickart, C. M. & Wolberger, C. Molecular Insights into Polyubiquitin Chain Assembly: Crystal Structure of the Mms2/Ubc13 Heterodimer. *Cell* **105**, 711–720 (2001).

78. Eddins, M. J., Carlile, C. M., Gomez, K. M., Pickart, C. M. & Wolberger, C. Mms2–Ubc13 covalently bound to ubiquitin reveals the structural basis of linkage-specific polyubiquitin chain formation. *Nature Structural & Molecular Biology* **13**, 915–920 (2006).

79. Bosanac, I. *et al.* Modulation of K11-linkage formation by variable loop residues within UbcH5A. *J Mol Biol* **408**, 420–431 (2011).

80. Wickliffe, K. E., Lorenz, S., Wemmer, D. E., Kuriyan, J. & Rape, M. The Mechanism of Linkage-Specific Ubiquitin Chain Elongation by a Single-Subunit E2. *Cell* **144**, 769–781 (2011).

81. Sadowski, M., Suryadinata, R., Lai, X., Heierhorst, J. & Sarcevic, B. Molecular Basis for Lysine Specificity in the Yeast Ubiquitin-Conjugating Enzyme Cdc34. *Mol Cell Biol* **30**, 2316–2329 (2010).

82. Suryadinata, R. *et al.* Molecular and structural insight into lysine selection on substrate and ubiquitin lysine 48 by the ubiquitin-conjugating enzyme Cdc34. *Cell Cycle* **12**, 1732–1744 (2013).

83. Rodrigo-Brenni, M. C., Foster, S. A. & Morgan, D. O. Catalysis of lysine 48-specific ubiquitin chain assembly by residues in E2 and ubiquitin. *Mol Cell* **39**, 548–559 (2010).

84. Wenzel, D. M., Lissounov, A., Brzovic, P. S. & Klevit, R. E. UBCH7 reactivity profile reveals parkin and HHARI to be RING/HECT hybrids. *Nature* **474**, 105–108 (2011).

85. Li, W. *et al.* Genome-wide and functional annotation of human E3 ubiquitin ligases identifies MULAN, a mitochondrial E3 that regulates the organelle's dynamics and signaling. *PLoS One* **3**, e1487 (2008).

86. Borden, K. L. & Freemont, P. S. The RING finger domain: a recent example of a sequence—structure family. *Current Opinion in Structural Biology* **6**, 395–401 (1996).

87. Metzger, M. B., Pruneda, J. N., Klevit, R. E. & Weissman, A. M. RING-type E3 ligases: master manipulators of E2 ubiquitin-conjugating enzymes and ubiquitination. *Biochim Biophys Acta* **1843**, 47–60 (2014).

88. Deshaies, R. J. & Joazeiro, C. A. P. RING Domain E3 Ubiquitin Ligases. *Annual Review of Biochemistry* **78**, 399–434 (2009).

89. Kulathu, Y. & Komander, D. Atypical ubiquitylation — the unexplored world of polyubiquitin beyond Lys48 and Lys63 linkages. *Nature Reviews Molecular Cell Biology* **13**, 508–523 (2012).

90. Aguilera, M., Oliveros, M., Martínez-Padrón, M., Barbas, J. A. & Ferrús, A. Ariadne-1: a vital Drosophila gene is required in development and defines a new conserved family of ring-finger proteins. *Genetics* **155**, 1231–1244 (2000).

91. Morett, E. & Bork, P. A novel transactivation domain in parkin. *Trends Biochem Sci* **24**, 229–231 (1999).

92. Eisenhaber, B., Chumak, N., Eisenhaber, F. & Hauser, M.-T. The ring between ring fingers (RBR) protein family. *Genome Biol* **8**, 209 (2007).

93. Smit, J. J. & Sixma, T. K. RBR E3-ligases at work. *EMBO reports* **15**, 142–154 (2014).

94. Spratt, D. E., Walden, H. & Shaw, G. S. RBR E3 ubiquitin ligases: new structures, new insights, new questions. *Biochem. J.* **458**, 421–437 (2014).

95. Schwintzer, L., Aguado Roca, E. & Broemer, M. TRIAD3/RNF216 E3 ligase specifically synthesises K63-linked ubiquitin chains and is inactivated by mutations associated with Gordon Holmes syndrome. *Cell Death Discovery* **5**, 1–11 (2019).

96. Stieglitz, B. *et al.* Structural basis for ligase-specific conjugation of linear ubiquitin chains by HOIP. *Nature* **503**, 422–426 (2013).

97. Huibregtse, J. M., Scheffner, M., Beaudenon, S. & Howley, P. M. 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 (1995).

98. Fajner, V., Maspero, E. & Polo, S. Targeting HECT-type E3 ligases - insights from catalysis, regulation and inhibitors. *FEBS Lett* **591**, 2636–2647 (2017).

99. Rotin, D. & Kumar, S. Physiological functions of the HECT family of ubiquitin ligases. *Nat Rev Mol Cell Biol* **10**, 398–409 (2009).

100. Huang, L. *et al.* Structure of an E6AP-UbcH7 complex: insights into ubiquitination by the E2-E3 enzyme cascade. *Science* **286**, 1321–1326 (1999).

101. Matta-Camacho, E., Kozlov, G., Menade, M. & Gehring, K. Structure of the HECT C-lobe of the UBR5 E3 ubiquitin ligase. *Acta Cryst F* **68**, 1158–1163 (2012).

102. Maspero, E. *et al.* Structure of the HECT:ubiquitin complex and its role in ubiquitin chain elongation. *EMBO Rep.* **12**, 342–349 (2011).

103. Ogunjimi, A. A. *et al.* Regulation of Smurf2 ubiquitin ligase activity by anchoring the E2 to the HECT domain. *Mol. Cell* **19**, 297–308 (2005).

104. Kamadurai, H. B. *et al.* Mechanism of ubiquitin ligation and lysine prioritization by a HECT E3. *eLife* **2**, e00828 (2013).

105. Maspero, E. *et al.* Structure of a ubiquitin-loaded HECT ligase reveals the molecular basis for catalytic priming. *Nature Structural & Molecular Biology* **20**, 696–701 (2013).

106. Salvat, C., Wang, G., Dastur, A., Lyon, N. & Huibregtse, J. M. The -4 Phenylalanine Is Required for Substrate Ubiquitination Catalyzed by HECT Ubiquitin Ligases. *J. Biol. Chem.* **279**, 18935–18943 (2004).

107. Kamadurai, H. B. *et al.* Insights into ubiquitin transfer cascades from a structure of a UbcH5B approximately ubiquitin-HECT(NEDD4L) complex. *Mol. Cell* **36**, 1095–1102 (2009).

108. Pandya, R. K., Partridge, J. R., Love, K. R., Schwartz, T. U. & Ploegh, H. L. A Structural Element within the HUWE1 HECT Domain Modulates Self-ubiquitination and Substrate Ubiquitination Activities. *J. Biol. Chem.* **285**, 5664–5673 (2010).

109. Kim, H. C., Steffen, A. M., Oldham, M. L., Chen, J. & Huibregtse, J. M. Structure and function of a HECT domain ubiquitin-binding site. *EMBO reports* **12**, 334–341 (2011).

110. Kathman, S. G. *et al.* A Small Molecule That Switches a Ubiquitin Ligase From a Processive to a Distributive Enzymatic Mechanism. *J. Am. Chem. Soc.* **137**, 12442–12445 (2015).

111. Gong, W., Zhang, X., Zhang, W., Li, J. & Li, Z. Structure of the HECT domain of human WWP2. *Acta Cryst F* **71**, 1251–1257 (2015).

112. Zhang, W. *et al.* System-Wide Modulation of HECT E3 Ligases with Selective Ubiquitin Variant Probes. *Mol. Cell* **62**, 121–136 (2016).

113. Sander, B., Xu, W., Eilers, M., Popov, N. & Lorenz, S. A conformational switch regulates the ubiquitin ligase HUWE1. *eLife* **6**, e21036 (2017).

114. Chen, Z. *et al.* A Tunable Brake for HECT Ubiquitin Ligases. *Mol. Cell* **66**, 345-357.e6 (2017).

115. Zhu, K. *et al.* Allosteric auto-inhibition and activation of the Nedd4 family E3 ligase Itch. *EMBO Rep.* **18**, 1618–1630 (2017).

116. Wang, Z. *et al.* A multi-lock inhibitory mechanism for fine-tuning enzyme activities of the HECT family E3 ligases. *Nat Commun* **10**, 3162 (2019).

117. Singh, S., Ng, J., Nayak, D. & Sivaraman, J. Structural insights into a HECT-type E3 ligase AREL1 and its ubiquitination activities in vitro. *J. Biol. Chem.* jbc.RA119.010327 (2019) doi:10.1074/jbc.RA119.010327.

118. Singh, S. & Sivaraman, J. Crystal structure of HECT domain of UBE3C E3 ligase and its ubiquitination activity. *Biochemical Journal* **477**, 905–923 (2020).

119. Lorenz, S., Cantor, A. J., Rape, M. & Kuriyan, J. Macromolecular juggling by ubiquitylation enzymes. *BMC Biology* **11**, 65 (2013).

120. Ogunjimi, A. A. *et al.* The ubiquitin binding region of the Smurf HECT domain facilitates polyubiquitylation and binding of ubiquitylated substrates. *J. Biol. Chem.* **285**, 6308–6315 (2010).

121. Wiesner, S. *et al.* Autoinhibition of the HECT-Type Ubiquitin Ligase Smurf2 through Its C2 Domain. *Cell* **130**, 651–662 (2007).

122. Buetow, L. & Huang, D. T. Structural insights into the catalysis and regulation of E3 ubiquitin ligases. *Nature Reviews Molecular Cell Biology* **17**, 626–642 (2016).

123. French, M. E. *et al.* Mechanism of ubiquitin chain synthesis employed by a HECT domain ubiquitin ligase. *J. Biol. Chem.* **292**, 10398–10413 (2017).

124. French, M. E., Kretzmann, B. R. & Hicke, L. Regulation of the RSP5 Ubiquitin Ligase by an Intrinsic Ubiquitin-binding Site. *J. Biol. Chem.* **284**, 12071–12079 (2009).

125. Mari, S. *et al.* Structural and Functional Framework for the Autoinhibition of Nedd4-Family Ubiquitin Ligases. *Structure* **22**, 1639–1649 (2014).

126. Zhong, Q., Gao, W., Du, F. & Wang, X. Mule/ARF-BP1, a BH3-Only E3 Ubiquitin Ligase, Catalyzes the Polyubiquitination of Mcl-1 and Regulates Apoptosis. *Cell* **121**, 1085–1095 (2005).

127. Dunn, R. & Hicke, L. Domains of the Rsp5 ubiquitin-protein ligase required for receptor-mediated and fluid-phase endocytosis. *Mol. Biol. Cell* **12**, 421–435 (2001).

128. Mund, T. & Pelham, H. R. B. Control of the activity of WW-HECT domain E3 ubiquitin ligases by NDFIP proteins. *EMBO Rep* **10**, 501–507 (2009).

129. Sheng, Y. *et al.* A human ubiquitin conjugating enzyme (E2)-HECT E3 ligase structure-function screen. *Mol Cell Proteomics* **11**, 329–341 (2012).

130. You, J. & Pickart, C. M. A HECT domain E3 enzyme assembles novel polyubiquitin chains. *J Biol Chem* **276**, 19871–19878 (2001).

131. Kim, H. C. & Huibregtse, J. M. Polyubiquitination by HECT E3s and the Determinants of Chain Type Specificity. *Molecular and Cellular Biology* **29**, 3307–3318 (2009).

132. Wang, M., Cheng, D., Peng, J. & Pickart, C. M. Molecular determinants of polyubiquitin linkage selection by an HECT ubiquitin ligase. *EMBO J.* **25**, 1710–1719 (2006).

133. Kristariyanto, Y. A. *et al.* Assembly and structure of Lys33-linked polyubiquitin reveals distinct conformations. *Biochem J* **467**, 345–352 (2015).

134. Michel, M. A. *et al.* Assembly and Specific Recognition of K29- and K33-Linked Polyubiquitin. *Molecular Cell* **58**, 95–109 (2015).

135. Michel, M. A., Swatek, K. N., Hospenthal, M. K. & Komander, D. Ubiquitin Linkage-Specific Affimers Reveal Insights into K6-Linked Ubiquitin Signaling. *Mol Cell* **68**, 233-246.e5 (2017).

136. Scialpi, F. *et al.* Itch self-polyubiquitylation occurs through lysine-63 linkages. *Biochemical Pharmacology* **76**, 1515–1521 (2008).

137. Kim, H. T. *et al.* Certain pairs of ubiquitin-conjugating enzymes (E2s) and ubiquitin-protein ligases (E3s) synthesize nondegradable forked ubiquitin chains containing all possible isopeptide linkages. *J. Biol. Chem.* **282**, 17375–17386 (2007).

138. Galan, J. M. & Haguenauer-Tsapis, R. Ubiquitin lys63 is involved in ubiquitination of a yeast plasma membrane protein. *EMBO J* **16**, 5847–5854 (1997).

139. Nuber, U., Schwarz, S. E. & Scheffner, M. The ubiquitin-protein ligase E6-associated protein (E6-AP) serves as its own substrate. *Eur. J. Biochem.* **254**, 643–649 (1998).

140. Fang, S. & Weissman, A. M. A field guide to ubiquitylation. *Cell. Mol. Life Sci.* **61**, 1546–1561 (2004).

141. Woelk, T. *et al.* Molecular mechanisms of coupled monoubiquitination. *Nat. Cell Biol.* **8**, 1246–1254 (2006).

142. Gallagher, E., Gao, M., Liu, Y.-C. & Karin, M. Activation of the E3 ubiquitin ligase Itch through a phosphorylation-induced conformational change. *Proc Natl Acad Sci U S A* **103**, 1717–1722 (2006).

143. Riling, C. *et al.* Itch WW Domains Inhibit Its E3 Ubiquitin Ligase Activity by Blocking E2-E3 Ligase Trans-thiolation. *J. Biol. Chem.* **290**, 23875–23887 (2015).

144. Courivaud, T. *et al.* Functional Characterization of a WWP1/Tiul1 Tumor-derived Mutant Reveals a Paradigm of Its Constitutive Activation in Human Cancer. *J. Biol. Chem.* **290**, 21007–21018 (2015).

145. Lu, K. *et al.* Pivotal role of the C2 domain of the Smurf1 ubiquitin ligase in substrate selection. *J Biol Chem* **286**, 16861–16870 (2011).

146. Mund, T. *et al.* Disinhibition of the HECT E3 ubiquitin ligase WWP2 by polymerized Dishevelled. *Open Biol* **5**, 150185 (2015).

147. Ruetalo, N. *et al.* The WW1 Domain Enhances Autoinhibition in Smurf Ubiquitin Ligases. *Journal of Molecular Biology* **431**, 4834–4847 (2019).

148. Wang, J. *et al.* Calcium Activates Nedd4 E3 Ubiquitin Ligases by Releasing the C2 Domain-mediated Auto-inhibition. *J. Biol. Chem.* **285**, 12279–12288 (2010).

149. Escobedo, A. *et al.* Structural basis of the activation and degradation mechanisms of the E3 ubiquitin ligase Nedd4L. *Structure* **22**, 1446–1457 (2014).

150. Ruetalo Buschinger, N. Mechanisms underlying the regulation of Nedd4-family E3 Ubiquitin ligases. (Universität Tübingen, 2020). doi:http://dx.doi.org/10.15496/publikation-21031.

151. Persaud, A. *et al.* Tyrosine phosphorylation of NEDD4 activates its ubiquitin ligase activity. *Sci Signal* **7**, ra95 (2014).

152. Bruce, M. C. *et al.* Regulation of Nedd4-2 self-ubiquitination and stability by a PY motif located within its HECT-domain. *Biochemical Journal* **415**, 155–163 (2008).

153. Yao, W. *et al.* WW domain–mediated regulation and activation of E3 ubiquitin ligase Suppressor of Deltex. *J Biol Chem* **293**, 16697–16708 (2018).

154. Lu, K. *et al.* Targeting WW domains linker of HECT-type ubiquitin ligase Smurf1 for activation by CKIP-1. *Nat Cell Biol* **10**, 994–1002 (2008).

155. Wan, L. *et al.* Cdh1 Regulates Osteoblast Function Through an APC/C-Independent Modulation of Smurf1. *Mol Cell* **44**, 721–733 (2011).

156. Kühnle, S. *et al.* Physical and Functional Interaction of the HECT Ubiquitin-protein Ligases E6AP and HERC2. *J Biol Chem* **286**, 19410–19416 (2011).

157. Mortensen, F. *et al.* Role of ubiquitin and the HPV E6 oncoprotein in E6AP-mediated ubiquitination. *Proc Natl Acad Sci U S A* **112**, 9872–9877 (2015).

158. Bernassola, F., Karin, M., Ciechanover, A. & Melino, G. The HECT Family of E3 Ubiquitin Ligases: Multiple Players in Cancer Development. *Cancer Cell* **14**, 10–21 (2008).

159. Gong, X. *et al.* The structure and regulation of the E3 ubiquitin ligase HUWE1 and its biological functions in cancer. *Invest New Drugs* **38**, 515–524 (2020).

160. Zou, X., Levy-Cohen, G. & Blank, M. Molecular functions of NEDD4 E3 ubiquitin ligases in cancer. *Biochim. Biophys. Acta* **1856**, 91–106 (2015).

161. Crawford, L. J. *et al.* The E3 ligase HUWE1 inhibition as a therapeutic strategy to target MYC in multiple myeloma. *Oncogene* **39**, 5001–5014 (2020).

162. Jäckl, M. *et al.* β-Sheet Augmentation Is a Conserved Mechanism of Priming HECT E3 Ligases for Ubiquitin Ligation. *J. Mol. Biol.* **430**, 3218–3233 (2018).

163. Haglund, K. & Dikic, I. Ubiquitylation and cell signaling. *The EMBO Journal* **24**, 3353–3359 (2005).

164. Ikeda, F. & Dikic, I. Atypical ubiquitin chains: new molecular signals. *EMBO reports* **9**, 536–542 (2008).

165. Jin, L., Williamson, A., Banerjee, S., Philipp, I. & Rape, M. Mechanism of Ubiquitin-Chain Formation by the human Anaphase-Promoting Complex. *Cell* **133**, 653–665 (2008).

166. Tanno, H. & Komada, M. The ubiquitin code and its decoding machinery in the endocytic pathway. *The Journal of Biochemistry* **153**, 497–504 (2013).

167. Scheffner, M., Nuber, U. & Huibregtse, J. M. Protein ubiquitination involving an E1–E2–E3 enzyme ubiquitin thioester cascade. *Nature* **373**, 81–83 (1995).

168. Berndsen, C. E. & Wolberger, C. New insights into ubiquitin E3 ligase mechanism. *Nat Struct Mol Biol* **21**, 301–307 (2014).

169. Scheffner, M. & Kumar, S. Mammalian HECT ubiquitin-protein ligases: Biological and pathophysiological aspects. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* **1843**, 61–74 (2014).

170. You, J., Wang, M., Aoki, T., Tamura, T. & Pickart, C. M. Proteolytic Targeting of Transcriptional Regulator TIP120B by a HECT Domain E3 Ligase *. *Journal of Biological Chemistry* **278**, 23369–23375 (2003).

171. Hospenthal, M. K., Freund, S. M. V. & Komander, D. Assembly, analysis and architecture of atypical ubiquitin chains. *Nature Structural & Molecular Biology* **20**, 555–565 (2013).

172. Hao, Z. *et al.* K48-linked KLF4 ubiquitination by E3 ligase Mule controls T-cell proliferation and cell cycle progression. *Nat Commun* **8**, 14003 (2017).

173. Zhao, X. *et al.* The HECT-domain ubiquitin ligase Huwe1 controls neural differentiation and proliferation by destabilizing the N-Myc oncoprotein. *Nat. Cell Biol.* **10**, 643–653 (2008).

174. Groot, R. E. A. de *et al.* Huwe1-Mediated Ubiquitylation of Dishevelled Defines a Negative Feedback Loop in the Wnt Signaling Pathway. *Sci. Signal.* **7**, ra26–ra26 (2014).

175. Myant, K. B. *et al.* HUWE1 is a critical colonic tumour suppressor gene that prevents MYC signalling, DNA damage accumulation and tumour initiation. *EMBO Molecular Medicine* **9**, 181–197 (2017).

176. Chen, D. *et al.* ARF-BP1/Mule Is a Critical Mediator of the ARF Tumor Suppressor. *Cell* **121**, 1071–1083 (2005).

177. Peter, S. *et al.* Tumor cell-specific inhibition of MYC function using small molecule inhibitors of the HUWE1 ubiquitin ligase. *EMBO Mol Med* **6**, 1525–1541 (2014).

178. Kirkpatrick, D. S., Gerber, S. A. & Gygi, S. P. The absolute quantification strategy: a general procedure for the quantification of proteins and post-translational modifications. *Methods* **35**, 265–273 (2005).

179. Lv, Z. *et al.* S. pombe Uba1-Ubc15 Structure Reveals a Novel Regulatory Mechanism of Ubiquitin E2 Activity. *Molecular Cell* **65**, 699-714.e6 (2017).

180. Merkley, N., Barber, K. R. & Shaw, G. S. Ubiquitin Manipulation by an E2 Conjugating Enzyme Using a Novel Covalent Intermediate. *J. Biol. Chem.* **280**, 31732–31738 (2005).

181. Olsen, S. K. & Lima, C. D. Structure of a ubiquitin E1-E2 complex: insights to E1-E2 thioester transfer. *Mol. Cell* **49**, 884–896 (2013).

182. Webb, B. & Sali, A. Protein Structure Modeling with MODELLER. in *Protein Structure Prediction* (ed. Kihara, D.) 1–15 (Springer, 2014). doi:10.1007/978-1-4939-0366-5_1.

183. Bond, S. R. & Naus, C. C. RF-Cloning.org: an online tool for the design of restriction-free cloning projects. *Nucleic Acids Res* **40**, W209–W213 (2012).

184. Delaglio, F. *et al.* NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* **6**, 277–293 (1995).

185. Bartels, C., Xia, T., Billeter, M., Güntert, P. & Wüthrich, K. The program XEASY for computer-supported NMR spectral analysis of biological macromolecules. *J Biomol NMR* **6**, 1–10 (1995).

186. Lee, W., Tonelli, M. & Markley, J. L. NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy. *Bioinformatics* **31**, 1325–1327 (2015).

187. Kabsch, W. XDS. *Acta Crystallogr D Biol Crystallogr* **66**, 125–132 (2010).

188. McCoy, A. J. *et al.* Phaser crystallographic software. *J Appl Cryst* **40**, 658–674 (2007).

189. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 (2010).

190. Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Cryst D* **66**, 213–221 (2010).

191. Ellman, G. L. A colorimetric method for determining low concentrations of mercaptans. *Archives of Biochemistry and Biophysics* **74**, 443–450 (1958).

192. Habeeb, A. F. S. A. [37] Reaction of protein sulfhydryl groups with Ellman's reagent. in *Methods in Enzymology* vol. 25 457–464 (Academic Press, 1972).

193. Serniwka, S. A. & Shaw, G. S. The Structure of the UbcH8−Ubiquitin Complex Shows a Unique Ubiquitin Interaction Site. *Biochemistry* **48**, 12169–12179 (2009).

194. Jäckl, M. Structural characterization of reaction intermediates of HECT-mediated ubiquitin transfer. (Universität Tübingen, 2020). doi:http://dx.doi.org/10.15496/publikation-26330.

195. Krissinel, E. & Henrick, K. Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **372**, 774–797 (2007).

196. Luz, Z. & Meiboom, S. Nuclear Magnetic Resonance Study of the Protolysis of Trimethylammonium Ion in Aqueous Solution—Order of the Reaction with Respect to Solvent. *J. Chem. Phys.* **39**, 366–370 (1963).

197. Millet, O., Loria, J. P., Kroenke, C. D., Pons, M. & Palmer, A. G. The Static Magnetic Field Dependence of Chemical Exchange Linebroadening Defines the NMR Chemical Shift Time Scale. *J. Am. Chem. Soc.* **122**, 2867–2877 (2000).

198. Kleckner, I. R. & Foster, M. P. An introduction to NMR-based approaches for measuring protein dynamics. *Biochim. Biophys. Acta* **1814**, 942–968 (2011).

199. Lipchock, J. & Loria, J. P. Millisecond dynamics in the allosteric enzyme imidazole glycerol phosphate synthase (IGPS) from Thermotoga maritima. *J. Biomol. NMR* **45**, 73–84 (2009).

200. Malakhova, O. A. & Zhang, D.-E. ISG15 inhibits Nedd4 ubiquitin E3 activity and enhances the innate antiviral response. *J. Biol. Chem.* **283**, 8783–8787 (2008).

201. Jiang, H., Thomas, S. N., Chen, Z., Chiang, C. Y. & Cole, P. A. Comparative analysis of the catalytic regulation of NEDD4-1 and WWP2 ubiquitin ligases. *J. Biol. Chem.* **294**, 17421–17436 (2019).

202. Huang, X. & Dixit, V. M. Drugging the undruggables: exploring the ubiquitin system for drug development. *Cell Res.* **26**, 484–498 (2016).

203. Galdeano, C. Drugging the undruggable: targeting challenging E3 ligases for personalized medicine. *Future Medicinal Chemistry* **9**, 347–350 (2017).

204. Herrador, A., Léon, S., Haguenauer-Tsapis, R. & Vincent, O. A mechanism for protein monoubiquitination dependent on a trans-acting ubiquitin-binding domain. *J. Biol. Chem.* **288**, 16206–16211 (2013).

205. Waudby, C. A., Ramos, A., Cabrita, L. D. & Christodoulou, J. Two-Dimensional NMR Lineshape Analysis. *Sci Rep* **6**, 24826 (2016).

206. Waudby, C. A., Ouvry, M., Davis, B. & Christodoulou, J. Two-dimensional NMR lineshape analysis of single, multiple, zero and double quantum correlation experiments. *J Biomol NMR* **74**, 95–109 (2020).

207. Gill, M. L. & Palmer, A. G. Multiplet-filtered and gradient-selected zero-quantum TROSY experiments for 13C1H3 methyl groups in proteins. *J. Biomol. NMR* **51**, 245–251 (2011).

208. Kaiser, S. E. *et al.* Protein standard absolute quantification (PSAQ) method for the measurement of cellular ubiquitin pools. *Nat. Methods* **8**, 691–696 (2011).

209. Park, C.-W. & Ryu, K.-Y. Cellular ubiquitin pool dynamics and homeostasis. *BMB Rep* **47**, 475–482 (2014).

210. Stieglitz, B., Morris-Davies, A. C., Koliopoulos, M. G., Christodoulou, E. & Rittinger, K. LUBAC synthesizes linear ubiquitin chains via a thioester intermediate. *EMBO reports* **13**, 840–846 (2012).

211. Coulombe, P., Rodier, G., Bonneil, E., Thibault, P. & Meloche, S. N-Terminal Ubiquitination of Extracellular Signal-Regulated Kinase 3 and p21 Directs Their Degradation by the Proteasome. *Mol Cell Biol* **24**, 6140–6150 (2004).

212. Woelk, T., Sigismund, S., Penengo, L. & Polo, S. The ubiquitination code: a signalling problem. *Cell Div* **2**, 11 (2007).

213. Jenuwein, T. & Allis, C. D. Translating the Histone Code. *Science* **293**, 1074–1080 (2001).

214. Emmerich, C. H. & Cohen, P. Optimising methods for the preservation, capture and identification of ubiquitin chains and ubiquitylated proteins by immunoblotting. *Biochemical and Biophysical Research Communications* **466**, 1–14 (2015).

215. Hospenthal, M. K., Mevissen, T. E. T. & Komander, D. Deubiquitinase-based analysis of ubiquitin chain architecture using Ubiquitin Chain Restriction (UbiCRest). *Nat Protoc* **10**, 349–361 (2015).

216. Ellison, M. J. & Hochstrasser, M. Epitope-tagged ubiquitin. A new probe for analyzing ubiquitin function. *J. Biol. Chem.* **266**, 21150–21157 (1991).

217. McDowell, G. S. & Philpott, A. Non-canonical ubiquitylation: Mechanisms and consequences. *The International Journal of Biochemistry & Cell Biology* **45**, 1833–1842 (2013).

218. Schrödinger, LLC. The PyMOL Molecular Graphics System, Version 1.8. (2015).

219. Smit, J. J. *et al.* Target Specificity of the E3 Ligase LUBAC for Ubiquitin and NEMO Relies on Different Minimal Requirements. *J Biol Chem* **288**, 31728–31737 (2013).

220. Kliza, K. *et al.* Internally tagged ubiquitin: a tool to identify linear polyubiquitin-modified proteins by mass spectrometry. *Nature Methods* **14**, 504–512 (2017).

221. Martinez-Fonts, K. *et al.* The proteasome 19S cap and its ubiquitin receptors provide a versatile recognition platform for substrates. *Nature Communications* **11**, 477 (2020).

222. Herrador, A., Herranz, S., Lara, D. & Vincent, O. Recruitment of the ESCRT machinery to a putative seven-transmembrane-domain receptor is mediated by an arrestin-related protein. *Mol. Cell. Biol.* **30**, 897–907 (2010).

223. Kee, Y., Lyon, N. & Huibregtse, J. M. The Rsp5 ubiquitin ligase is coupled to and antagonized by the Ubp2 deubiquitinating enzyme. *EMBO J* **24**, 2414–2424 (2005).

224. Petroski, M. D. & Deshaies, R. J. Mechanism of Lysine 48-Linked Ubiquitin-Chain Synthesis by the Cullin-RING Ubiquitin-Ligase Complex SCF-Cdc34. *Cell* **123**, 1107–1120 (2005).

225. Schäfer, A., Kuhn, M. & Schindelin, H. Structure of the ubiquitin-activating enzyme loaded with two ubiquitin molecules. *Acta Cryst D* **70**, 1311–1320 (2014).

226. Hiraishi, H., Okada, M., Ohtsu, I. & Takagi, H. A functional analysis of the yeast ubiquitin ligase Rsp5: the involvement of the ubiquitin-conjugating enzyme Ubc4 and poly-ubiquitination in ethanol-induced down-regulation of targeted proteins. *Biosci. Biotechnol. Biochem.* **73**, 2268–2273 (2009).

227. Gwizdek, C. *et al.* The mRNA Nuclear Export Factor Hpr1 Is Regulated by Rsp5-mediated Ubiquitylation. *J. Biol. Chem.* **280**, 13401–13405 (2005).

228. Pickart, C. M. & Raasi, S. Controlled Synthesis of Polyubiquitin Chains. in *Methods in Enzymology* vol. 399 21–36 (Academic Press, 2005).

229. Bjellqvist, B. *et al.* The focusing positions of polypeptides in immobilized pH gradients can be predicted from their amino acid sequences. *ELECTROPHORESIS* **14**, 1023–1031 (1993).

230. Bjellqvist, B., Basse, B., Olsen, E. & Celis, J. E. Reference points for comparisons of two-dimensional maps of proteins from different human cell types defined in a pH scale where isoelectric points correlate with polypeptide compositions. *ELECTROPHORESIS* **15**, 529–539 (1994).

231. Gasteiger, E. *et al.* Protein Identification and Analysis Tools on the ExPASy Server. in *The Proteomics Protocols Handbook* (ed. Walker, J. M.) 571–607 (Humana Press, 2005). doi:10.1385/1-59259-890-0:571.

232. Pierce, N. W., Kleiger, G., Shan, S. & Deshaies, R. J. Detection of sequential polyubiquitylation on a millisecond timescale. *Nature* **462**, 615–619 (2009).

233. Bradford, M. M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry* **72**, 248–254 (1976).

234. Mandel, M. & Higa, A. Calcium-dependent bacteriophage DNA infection. *Journal of Molecular Biology* **53**, 159–162 (1970).

235. van den Ent, F. & Löwe, J. RF cloning: A restriction-free method for inserting target genes into plasmids. *Journal of Biochemical and Biophysical Methods* **67**, 67–74 (2006).

236. Geier, G. E. & Modrich, P. Recognition sequence of the dam methylase of Escherichia coli K12 and mode of cleavage of Dpn I endonuclease. *J Biol Chem* **254**, 1408–1413 (1979).

237. Lapid, C. PrimerX - Automated design of mutagenic primers for site-directed mutagenesis. http://www.bioinformatics.org/primerx/documentation.html (2003).

238. Dümmler, A., Lawrence, A.-M. & de Marco, A. Simplified screening for the detection of soluble fusion constructs expressed in E. coli using a modular set of vectors. *Microb Cell Fact* **4**, 34 (2005).

239. Kapust, R. B., Tözsér, J., Copeland, T. D. & Waugh, D. S. The P1′ specificity of tobacco etch virus protease. *Biochemical and Biophysical Research Communications* **294**, 949–955 (2002).

240. Lau, Y.-T. K. *et al.* Discovery and engineering of enhanced SUMO protease enzymes. *Journal of Biological Chemistry* **293**, 13224–13233 (2018).

241. Owerbach, D., McKay, E. M., Yeh, E. T. H., Gabbay, K. H. & Bohren, K. M. A proline-90 residue unique to SUMO-4 prevents maturation and sumoylation. *Biochemical and Biophysical Research Communications* **337**, 517–520 (2005).

242. Laemmli, U. K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680–685 (1970).

243. Maizel, J. V. 5 - Polyacrylamide Gel Electrophoresis of Viral Proteins. in *Methods in Virology* (eds. Maramorosch, K. & Koprowski, H.) vol. 5 179–246 (Elsevier, 1971).

244. Bloch, F. Nuclear Induction. *Phys. Rev.* **70**, 460–474 (1946).

245. Cavanagh, J. & Rance, M. Sensitivity-Enhanced NMR Techniques for the Study of Biomolecules. in *Annual Reports on NMR Spectroscopy* (ed. Webb, G. A.) vol. 27 1–58 (Academic Press, 1993).

246. Palmer, A. G., Cavanagh, J., Wright, P. E. & Rance, M. Sensitivity improvement in proton-detected two-dimensional heteronuclear correlation NMR spectroscopy. *Journal of Magnetic Resonance (1969)* **93**, 151–170 (1991).

247. Pervushin, K., Riek, R., Wider, G. & Wüthrich, K. Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 12366–12371 (1997).

248. Pervushin, K. V., Wider, G. & Wüthrich, K. Single Transition-to-single Transition Polarization Transfer (ST2-PT) in [15N,1H]-TROSY. *J Biomol NMR* **12**, 345–348 (1998).

249. Schütz, S. & Sprangers, R. Methyl TROSY spectroscopy: A versatile NMR approach to study challenging biological systems. *Progress in Nuclear Magnetic Resonance Spectroscopy* **116**, 56–84 (2020).

250. Takeuchi, K., Arthanari, H. & Wagner, G. Perspective: revisiting the field dependence of TROSY sensitivity. *J Biomol NMR* **66**, 221–225 (2016).

251. Ollerenshaw, J. E., Tugarinov, V. & Kay, L. E. Methyl TROSY: explanation and experimental verification. *Magnetic Resonance in Chemistry* **41**, 843–852 (2003).

252. Tugarinov, V., Hwang, P. M., Ollerenshaw, J. E. & Kay, L. E. Cross-Correlated Relaxation Enhanced 1H−13C NMR Spectroscopy of Methyl Groups in Very High Molecular Weight Proteins and Protein Complexes. *J. Am. Chem. Soc.* **125**, 10420–10428 (2003).

253. Schanda, P., Kupče, Ē. & Brutscher, B. SOFAST-HMQC Experiments for Recording Two-dimensional Deteronuclear Correlation Spectra of Proteins within a Few Seconds. *J Biomol NMR* **33**, 199–211 (2005).

254. Ross, A., Salzmann, M. & Senn, H. Fast-HMQC using Ernst angle pulses: An efficient tool for screening of ligand binding to target proteins. *J Biomol NMR* **10**, 389–396 (1997).

255. Bax, A., Griffey, R. H. & Hawkins, B. L. Correlation of proton and nitrogen-15 chemical shifts by multiple quantum NMR. *Journal of Magnetic Resonance (1969)* **55**, 301–315 (1983).

256. Bodenhausen, G. & Ruben, D. J. Natural abundance nitrogen-15 NMR by enhanced heteronuclear spectroscopy. *Chemical Physics Letters* **69**, 185–189 (1980).

257. Mueller, L. Sensitivity enhanced detection of weak nuclei using heteronuclear multiple quantum coherence. *J. Am. Chem. Soc.* **101**, 4481–4484 (1979).

258. Lundström, P., Vallurupalli, P., Religa, T. L., Dahlquist, F. W. & Kay, L. E. A single-quantum methyl 13C-relaxation dispersion experiment with improved sensitivity. *J Biomol NMR* **38**, 79–88 (2007).

259. Riek, R., Wider, G., Pervushin, K. & Wüthrich, K. Polarization transfer by cross-correlated relaxation in solution NMR with very large molecules. *PNAS* **96**, 4918–4923 (1999).

260. Loria, J. P., Rance, M. & Palmer, A. G. A Relaxation-Compensated Carr−Purcell−Meiboom−Gill Sequence for Characterizing Chemical Exchange by NMR Spectroscopy. *J. Am. Chem. Soc.* **121**, 2331–2332 (1999).

261. Mulder, F. A. A., Skrynnikov, N. R., Hon, B., Dahlquist, F. W. & Kay, L. E. Measurement of Slow (µs−ms) Time Scale Dynamics in Protein Side Chains by 15N Relaxation Dispersion NMR Spectroscopy: Application to Asn and Gln Residues in a Cavity Mutant of T4 Lysozyme. *J. Am. Chem. Soc.* **123**, 967–975 (2001).

262. Skrynnikov, N. R., Mulder, F. A. A., Hon, B., Dahlquist, F. W. & Kay, L. E. Probing Slow Time Scale Dynamics at Methyl-Containing Side Chains in Proteins by Relaxation Dispersion NMR Measurements: Application to Methionine Residues in a Cavity Mutant of T4 Lysozyme. *J. Am. Chem. Soc.* **123**, 4556–4566 (2001).

263. Long, D., Bouvignies, G. & Kay, L. E. Measuring hydrogen exchange rates in invisible protein excited states. *PNAS* **111**, 8820–8825 (2014).

264. McConnell, H. M. Reaction Rates by Nuclear Magnetic Resonance. *J. Chem. Phys.* **28**, 430–431 (1958).

265. Bouvignies, G. gbouvignies/ChemEx. (2021).

266. Shapovalov, M. V. & Dunbrack, R. L. A Smoothed Backbone-Dependent Rotamer Library for Proteins Derived from Adaptive Kernel Density Estimates and Regressions. *Structure* **19**, 844–858 (2011).

267. Pei, J., Kim, B.-H. & Grishin, N. V. PROMALS3D: a tool for multiple protein sequence and structure alignments. *Nucleic Acids Res* **36**, 2295–2300 (2008).

268. Robert, X. & Gouet, P. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res* **42**, W320-324 (2014).