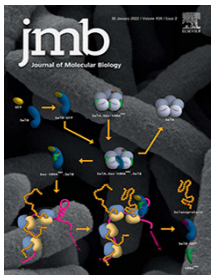
Research Article



Multifaceted N-Degron Recognition and Ubiquitylation by GID/CTLH E3 Ligases

Jakub Chrustowicz [1†], Dawafuti Sherpa [1†], Joan Teyra [2], Mun Siong Loke [1], Grzegorz M. Popowicz [4,5], Jerome Basquin [6], Michael Sattler [4,5], J. Rajan Prabu [1], Sachdev S. Sidhu [2,3] and Brenda A. Schulman [1*]

*1 - **Department of Molecular Machines and Signaling,** Max Planck Institute of Biochemistry, 82152 Martinsried, Germany*
*2 - **The Donnelly Centre,** University of Toronto, Toronto, ON M5S 3E1, Canada*
*3 - **Department of Molecular Genetics,** University of Toronto, Toronto, ON M5S 1A8, Canada*
*4 - **Helmholtz Zentrum München,** Ingolstädter Landstrasse 1, 85764 Neuherberg, Germany*
*5 - **Bavarian NMR Center,** Department of Chemistry, Technical University of Munich, Germany*
*6 - **Department of Structural Cell Biology,** Max Planck Institute of Biochemistry, 82152 Martinsried, Germany*

***Correspondence to Brenda A. Schulman:** schulman@biochem.mpg.de (B.A. Schulman)*
@dawafutisherpa (D. Sherpa), @chrustowicz_j (J. Chrustowicz), @rajanprabu (J.R. Prabu)
https://doi.org/10.1016/j.jmb.2021.167347
Edited by Igor Stagljar

Abstract

N-degron E3 ubiquitin ligases recognize specific residues at the N-termini of substrates. Although molecular details of N-degron recognition are known for several E3 ligases, the range of N-terminal motifs that can bind a given E3 substrate binding domain remains unclear. Here, we discovered capacity of Gid4 and Gid10 substrate receptor subunits of yeast "GID"/human "CTLH" multiprotein E3 ligases to tightly bind a wide range of N-terminal residues whose recognition is determined in part by the downstream sequence context. Screening of phage displaying peptide libraries with exposed N-termini identified novel consensus motifs with non-Pro N-terminal residues binding Gid4 or Gid10 with high affinity. Structural data reveal that conformations of flexible loops in Gid4 and Gid10 complement sequences and folds of interacting peptides. Together with analysis of endogenous substrate degrons, the data show that degron identity, substrate domains harboring targeted lysines, and varying E3 ligase higher-order assemblies combinatorially determine efficiency of ubiquitylation and degradation.

Introduction

Specificity of ubiquitylation depends on E3 ligases recognizing motifs, termed "degrons", in substrates to be modified. The first such motif to be identified was the N-terminal sequence - now called N-degron[1] - in substrates of the yeast E3 ligase Ubr1.[2,3] Subsequently, several E3 ligases in different families were discovered to recognize protein N-termini as degrons. Higher eukaryotes have one HECT-type and several RING-family E3s with "Ubr" domains homologous to those in yeast Ubr1 that either have been shown to or are presumed to recognize distinct N-terminal sequences.[4,5] Other N-degron-recognizing ubiquitin ligases were identified either through characterizing substrate sequences mediating E3-binding,[6,7] or through systematic genetic screens matching human protein N-terminal sequences with E3 ligases.[8] Some of the best-studied pathways recognize sequences with an N-terminal Arg,[9] Pro[6,10] or Gly[8,11] (termed

Arg/N-degron, Pro/N-degron or Gly/N-degron, respectively), or acetylated N-terminus.[12–15]

An N-degron-recognizing E3 of emerging importance is a suite of related multiprotein complexes termed "GID" in budding yeast (named due to mutations causing *g*lucose-*i*nduced degradation *d*eficiency of fructose-1,6-bisphosphatase, Fbp1)[7,16–20] or "CTLH" in higher eukaryotes (named due to preponderance of subunits containing CTLH motifs).[21] The yeast GID E3 mediates degradation of gluconeogenic enzymes Fbp1, Mdh2 and Icl1 during recovery from carbon starvation.[7] The GID E3 recognizes the N-terminal Pro in these substrates generated by cleavage of the initiator methionine.[6,7] In higher eukaryotes, corresponding CTLH complexes are involved in diverse biological processes including erythropoiesis, organ development, embryogenesis, and cell division.[22–32] However, the mechanistic roles of CTLH-mediated ubiquitylation in these pathways remain largely mysterious.

Recent genetic, biochemical and structural studies have revealed that the GID E3 is not a singular complex. Rather a core GID^{Ant} complex (comprising Gid1, Gid5, Gid8, Gid2, Gid9 subunits) essentially anticipates shifts in environmental conditions that stimulate expression of interchangeable and mutually exclusive substrate-binding receptors – Gid4 (termed "yGid4" for yeast Gid4 hereafter),[17,33,34] Gid10 (yGid10 hereafter)[34–36] and Gid11 (yGid11 hereafter).[37] Whereas yGid4 is expressed after glucose has been restored to carbon-starved yeast, yGid10 and yGid11 are upregulated upon other environmental perturbations including heat shock, osmotic stress as well as carbon, nitrogen and amino acid starvation. The resultant E3 complexes, GID^{SR4}, GID^{SR10}, and GID^{SR11} (where SR# refers to Gid substrate receptor), recognize distinct N-terminal sequences of their substrates.[6,7,34,35,37] In addition, another subunit, Gid7, can drive supramolecular assembly of two GID^{SR4} units into a complex named Chelator-GID^{SR4} to reflect its resemblance to an organometallic chelate capturing a smaller ligand through multiple contacts.[38] The cryo EM structure of a Chelator-GID^{SR4} complex with Fbp1 showed two opposing Gid4 molecules avidly binding N-degrons from different Fbp1 protomers. As such, Fbp1 is encapsulated within the center of the oval-shaped Chelator-GID^{SR4}. This assembly positions functionally-relevant target lysines from multiple Fbp1 protomers adjacent to two Chelator-GID^{SR4} catalytic centers.

The molecular details of GID/CTLH recognition of Pro/N-degrons were initially revealed from crystal structures of human Gid4 (referred to as hGid4 hereafter) bound to peptides with N-terminal prolines.[10] Although Pro/N-degron substrates of the CTLH E3 remain unknown, hGid4 is suitably well-behaved for biophysical and structural characterization, whereas yGid4 has limited solubility on its own.[10] Previously, the sequence PGLWKS was identified as binding hGid4 with highest affinity amongst all sequences tested, with a K_D in the low micromolar range.[10] The crystallized peptide-binding region of hGid4, which superimposes with the substrate-binding domains of yGid4 and yGid10 in GID^{SR4} and GID^{SR10}, adopts an 8-stranded β-barrel with a central tunnel that binds the N-terminus of a peptide, or of the intrinsically-disordered N-terminal degron sequence of a substrate.[10,34,36,38,39] Loops between β-strands at the edge of the barrel bind residues downstream of the peptide's N-terminus. Interestingly, although GID^{SR4} was originally thought to exclusively bind peptides with an N-terminal Pro, hGid4 can also bind peptides with non-Pro hydrophobic N-termini such as Ile or Leu, albeit with at best ≈8-fold lower affinity.[39] Furthermore, yGid11 is thought to use a distinct structure to recognize substrate Thr/N-degrons.[37] Collectively, these findings suggested that the landscape of GID/CTLH E3 substrates can extend beyond Pro/N-degron motifs.

Here, phage display screening identified peptides with various non-Pro N-termini that not only bind hGid4, yGid4 and yGid10, but do so with comparable or higher affinity than the previously identified Pro-initiating sequences including Pro/N-degrons of ubiquitylation substrates. Structural data reveal that loops in GID/CTLH substrate-binding domains adopt conformations complementary to partner peptide sequences downstream of the N-terminus. Thus, sequence context is a determinant of N-terminal recognition by GID/CTLH substrate-binding domains. In the context of natural substrates recognized by yGid4, not only the degron but also the associated domain harboring targeted lysine contribute to ubiquitylation by the core GID^{SR4} and its superassembly.

Results

hGid4 can bind peptides with a range of N-terminal sequences

We took advantage of the amenability of hGid4 to biophysical characterization to further characterize features of the PGLWKS sequence mediating interactions. To assess the importance of peptide length beyond the N-terminus, we examined chemical shift perturbations (CSPs) in 2D ^{1}H, ^{15}N-HSQC NMR spectra of [^{15}N]-labeled hGid4 mixed with the amino acid Pro, a Pro-Gly dipeptide, or the PGLWKS peptide (Figure 1(A)). Although prior studies emphasized the importance of an N-terminal Pro,[10,39] Pro alone only minimally influenced the spectrum. The Pro-Gly dipeptide elicited stronger CSPs, presumably due to the peptide bond directly interacting with hGid4, and suppressing repulsion by burying the negatively charged carboxylate of a single Pro in a hydrophobic environment (Figure S1(A)). The PGLWKS peptide showed the greatest CSPs and binding kinetics in the slow exchange regime at the NMR chemical shift

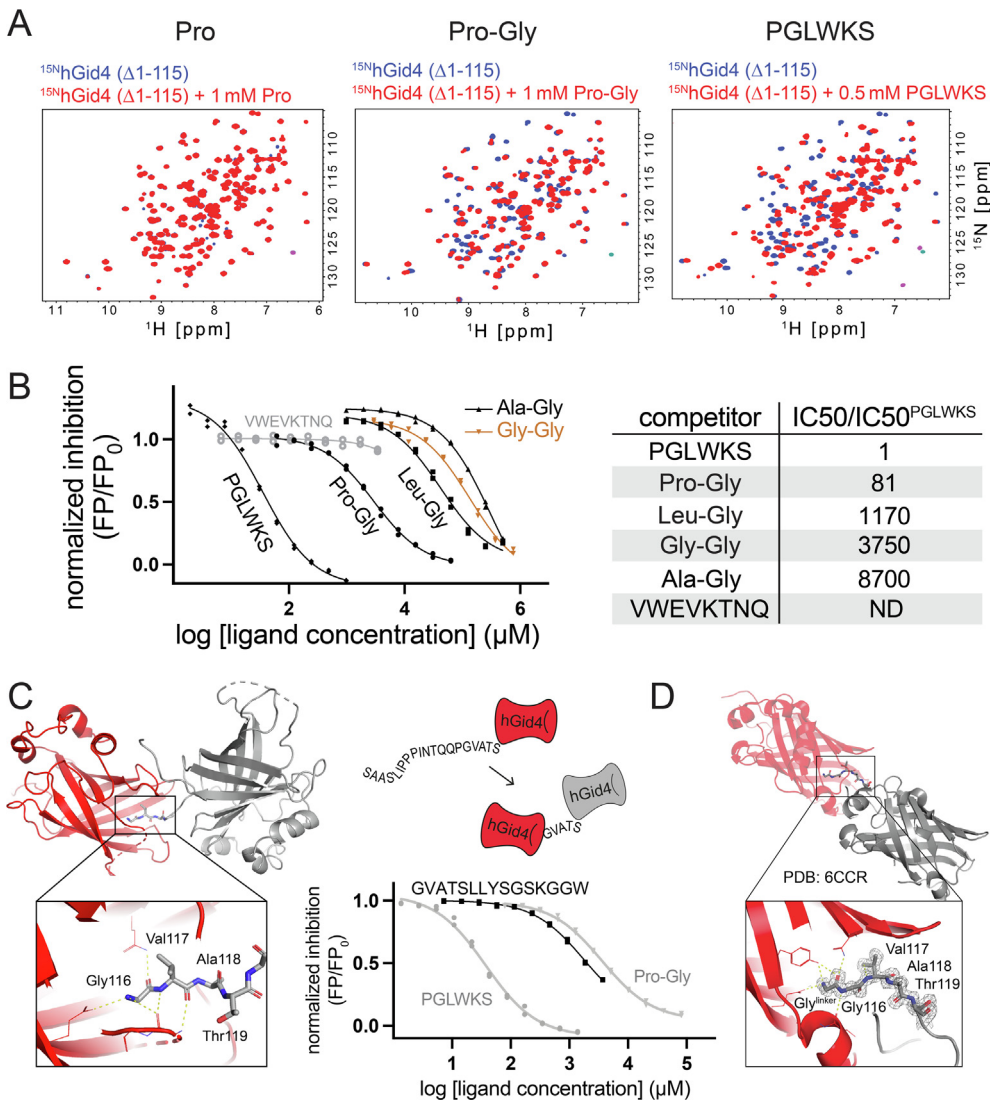


Figure 1. hGid4 recognizes various peptide N-termini and several downstream residues. A. Overlaid ^{1}H, ^{15}N-HSQC NMR spectra of 0.1 mM [^{15}N]-labeled 6xHis-hGid4 (Δ1-115) alone (blue) and upon addition of 1 mM Pro, 1 mM Pro-Gly or 0.5 mM PGLWKS peptide (red). B. Competitive fluorescence polarization (FP) experiments comparing different unlabeled ligands for inhibiting hGid4 (Δ1-115) binding to C-terminally fluorescein-labeled PGLWKS peptide. Ratios of FP signals at varying concentrations of unlabeled ligands to that in the absence of a competitor (FP/FP_0) were plotted as a function of log[ligand concentration] (left). Half-maximal inhibitory concentrations (IC50) for each ligand were determined by fitting to log[inhibitor] vs. response model and presented relative to IC50 of the unlabeled PGLWKS peptide (right). The peptide VWEVKTNQ corresponding to the N-terminus of Hbp1 (2–9) that is not an hGid4 substrate[88] was included as a negative control. C. Crystal structure of one hGid4 (red) accommodating serendipitously generated Gly116-initiating N-terminus of an adjacent hGid4 molecule (grey) in the crystal lattice. The binding strength of the newly generated N-terminal sequence (116–127) to hGid4 was compared to that of PGLWKS and Pro-Gly with competitive FP (right bottom). D. Previously published hGid4 crystal structure (PDB ID: **6CCR**) revealing one hGid4 binding the N-terminus bearing an additional Gly upstream Gly116 derived from cloning of an adjacent hGid4 molecule (grey) in the lattice of a distinct crystal form.

time scale, indicating tight binding, and, therefore, importance of downstream residues.

Given the ability of a Pro-Gly dipeptide to bind hGid4, we examined importance of the N-terminal residue by testing commercially-available variants (Leu-Gly, Ala-Gly, and Gly-Gly along with Pro-Gly) for competing with a fluorescently-labeled PGLWKS peptide whose binding to hGid4 can be measured by fluorescence polarization (FP) (Figure S1(B)). Although each of the dipeptides yielded sigmoidal curves, those with N-terminal Pro or Leu were superior (Figure 1(B)). Pro-Gly

showed a 15-fold lower IC_{50} than Leu-Gly, consistent with prior studies emphasizing the importance of an N-terminal Pro.[39]

To examine roles of individual positions in the 6-residue PGLWKS sequence, we employed peptide spot arrays testing all natural amino acids in position 1, positions 2 and 3 together, position 4 or position 5 (Figure S1(C)). Binding was detected after incubating the membranes with the substrate binding domain of hGid4, and immunoblotting with anti-hGid4 antibodies. Overall, the data confirm the previous findings that out of the peptides tested PGLWKS is an optimal binder, and that N-terminal non-Pro hydrophobic residues are tolerated in the context of the downstream GLWKS sequence albeit with lower binding.[10,39]

The peptide array data also highlighted the importance of context. Amongst the 400 possible combinations of residues 2 and 3, Gly, and to a lesser extent Ser, Val and Ala are preferred at position 2 and Ile or Leu at position 3, mirroring the previously defined sequence preferences.[10] The dynamic range of our assay suggested that downstream residues also contribute to specificity, by unveiling pronounced amino acid preference for bulky hydrophobics and some non-hydrophobic residues also at position 4. In agreement with the structural data,[10] the 5th position following the PGLW sequence tolerates many amino acids.

Despite this seemingly strong preference for an N-terminal Pro, we serendipitously visualized hGid4 recognizing a supposedly non-cognate sequence when we set out to visualize its structure in the absence of a peptide ligand by X-ray crystallography. Unexpectedly, the electron density from data at 3 Å resolution showed the first visible N-terminal residue of one molecule of hGid4 inserted into the substrate binding tunnel of an adjacent hGid4 molecule in the crystal lattice (Figure 1(C); Table S1). Perplexingly, this was not the first residue of the input hGid4 construct but Gly116 located 16 positions downstream. It appears that hGid4 underwent processing during crystallization, although it remains unknown if this neo-N-terminus was generated through enzymatic cleavage by a contaminating bacterial protease or chemical processing. Nonetheless, the potential for hGid4 to recognize a non-cognate N-terminal Gly was supported by re-examination of the published "apo" hGid4 crystal. In the previous structure of hGid4 (PDB ID: **6CCR)**, distinct crystal packing is also mediated by a peptide-like sequence (initiating with a Gly from the Tobacco Etch Virus (TEV) protease cleavage site, followed by hGid4 Gly116) inserting into the substrate binding tunnel of the neighboring molecule in the lattice (Figure 1(D)). The positions and interactions of the two N-terminal Gly are similar but not identical, as hGid4's Tyr258 does not hydrogen bond to the N-terminal amine of Gly116 in our structure.

To test binding of our fortuitously identified hGid4-binding sequence in solution, we examined competition with the fluorescently-labeled PGLWKS peptide (Figure 1(C)). Limited solubility of the GVATSLLYSGSKGGW peptide (hGid4 residues 116–127, with C-terminal Trp appended with a Gly-Gly linker to accurately measure peptide concentration) precluded accurate measurement of IC_{50} using our competitive FP assay. Nonetheless, the data qualitatively indicated that the GVATSLLYSGSKGGW peptide binds to hGid4 with significantly lower affinity than PGLWKS, but more tightly than the Pro-Gly dipeptide. Therefore, we speculate that these structurally-observed interactions were favored by the high concentration of protein during crystallization.

Taken together with published work, the data confirmed hGid4′s preference for binding to the previously-defined sequence PGLWKS, but they also highlighted capacity for hGid4 to recognize alternative N-termini. Moreover, given that specific combinations of residues downstream of the Pro-Gly substantially impact the interaction, we considered the possibility that hGid4 recognition of N-terminal sequences could be influenced by context.

Identification of superior hGid4-binding motifs not initiated by Pro

To discover alternative hGid4-binding sequences that do not initiate with Pro, we constructed a highly diverse N-terminal peptide phage-displayed library of 3.5×10^9 random octapeptides. The library was constructed after the signal peptide using 8 consecutive NNK degenerate codons encoding for all 20 natural amino acids and fused to the N-terminus of the phage coat protein. It is expected that Arg or Pro located next to the cleavage site (position + 1) will be inexistent or strongly underrepresented because they are known to either inhibit the secretion of phages[40,41] or the signal peptidase cleavage,[42,43] respectively.

The library was cycled through five rounds of selections following an established protocol[44] to enrich for phages displaying peptides that preferentially bound hGid4 (Figure 2(a)). Phages from individual clones that bound to GST-hGid4 (Δ1-99) but not a control GST based on phage ELISA were subjected to DNA sequence analysis.

The screen yielded 41 unique sequences, none of which were overtly similar to the previously defined hGid4-binding consensus motif PGLWKS (Figure 2(B); Table S2). A new consensus emerged with the following preferences: (1) hydrophobic residues at position 1, with Phe predominating; (2) Asp at position 2; (3) hydrophobic residues at positions 3 and 6, and to a lesser extent at position 5; and (4) small and polar residues at positions 4 and 7. Unlike the PGLWKS sequence wherein the striking selectivity

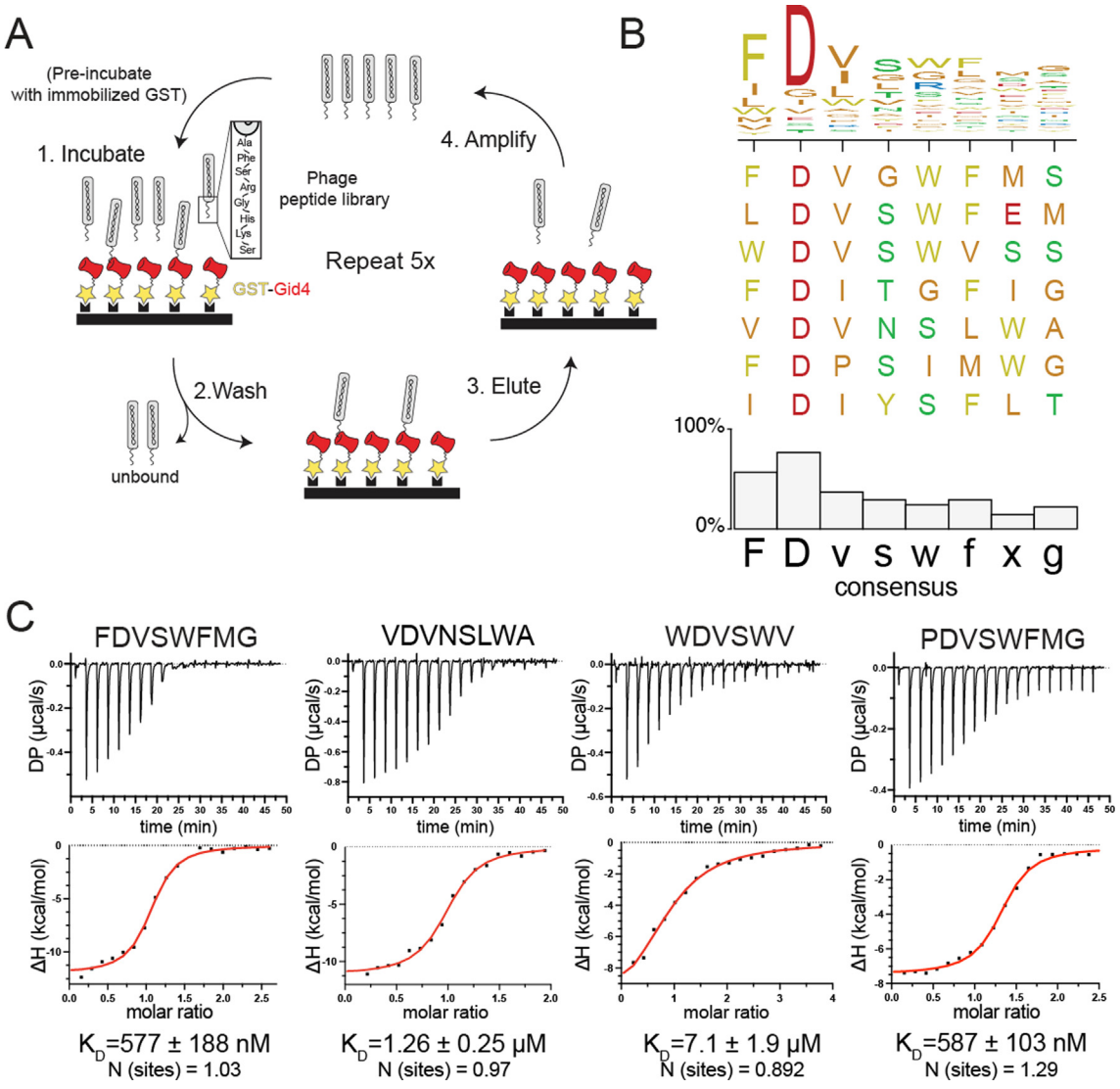


Figure 2. Identification of high-affinity hGid4-binding motifs initiating with non-Pro hydrophobic residues. A. Schematic of phage-display peptide library screen identifying peptides binding GST-tagged hGid4 (Δ1-99). B. Consensus motif obtained from multiple sequence alignment of 41 unique hGid4-binding peptide sequences listed in Table S2 (out of which a representative set of 7 sequences is shown). The height of the bars reflects the frequency of a given residue at different positions of the consensus. C. Isothermal titration calorimetry (ITC) to quantify binding of newly determined sequences to hGid4 (Δ1-115). The amount of heat released (ΔH) upon peptide injection was calculated from integrated raw ITC data (top) and plotted as a function of peptide:protein molar ratio (bottom). Dissociation constant (K_D) and the stoichiometry of the binding event (N) were determined by fitting to the One-Set-of-Sites binding model.

is predominantly for the first four residues, this new consensus extends through the seventh residue.

Although peptides with non-Pro hydrophobic N-termini were previously shown to bind hGid4, the tested sequences bound with one to two orders-of-magnitude lower affinity (K_D for IGLWKS 16 μM, VGLWKS 36 μM) than to PGLWKS (K_D = 1.9 μM) (Figure S2(A)).[39] To determine how the newly identified sequences compare, we quantified interactions by isothermal titration calorimetry (ITC). Notably, the peptides of sequences FDVSWFMG and VDVNSLWA showed superior binding (K_D = 0.6 and 1.3 μM, respectively) to the best binder with an N-terminal Pro (Figure 2(C) and S1(D)). Recognition of N-terminal Pro is also substantiated by the new consensus as substitution of Phe in FDVSWFMG with a Pro resulted in significantly tighter binding (K_D = 0.6 μM) than that of the PGLWKS motif (Figures 2(C) and S1(D)). Moreover, the affinity for a sequence starting with a Trp (K_D = 7.1 μM for WDVSWV) was superior to the previously identified best binder initiating with a non-Pro

hydrophobic residue (Figures 2(C) and S1(D)). Thus, hGid4 is able to accommodate even the bulkiest hydrophobic sidechain at the N-terminus of an interacting peptide. Taken together, the data show hGid4 binds a wide range of peptide sequences, with affinity strongly influenced by residues downstream of the N-terminus.

hGid4 structural pliability enables recognition of various N-terminal sequences

To understand how hGid4 recognizes diverse sequences, we determined its crystal structure bound to the FDVSWFMG peptide (Figure 3(A), Table S1; all peptide residues except C-terminal Gly visible in density). Overlaying this structure with published coordinates for other hGid4 complexes revealed diverse N-termini protruding into a common central substrate-binding tunnel (Figure S2(B), Phe (our study), or Pro, Leu, Val, or newly recognized Gly.[10,39] The N-terminal residues are anchored through contacts of their amine groups with hGid4 Glu237 and Tyr258 at the tip of the substrate binding tunnel, and common hydrogen bonds of the peptide backbone carbonyl to hGid4 Gln132.

The structures suggest that the varying peptide sequences are accommodated by complementary conformations of four hairpin loops (L1-L4) at the edge of the hGid4 substrate-binding tunnel (Figure 3(B)). The L2, L3, and L4 loops are fully or partially invisible, and are presumably mobile, in the structure of apo-hGid4 assembled in a subcomplex with its interacting subunits from the CTLH E3.[38] However, they are ordered and adopt different conformations when bound to the different peptides.

As compared to the structure with PGLWKS, the interactions with FDVSWFMG are more extensive and relatively more dominated by hydrophobic contacts rather than hydrogen bonding, which rationalizes improved binding of the new motif (Figure 3(C)). The L2 and L3 loops are relatively further from the central axis of the hGid4 β-barrel to interact with more residues in the peptide sequence. The different position of the L2 loop is also required to accommodate the hydrophobic Phe in the context of the new sequence (Figure 3 (D)). Meanwhile, repositioning of the L4 loop places hGid4 Gln282 to form a hydrogen bond with Asp2 in the peptide (Figure 3(C)). Moreover, upon binding to hGid4, FDVSWFMG itself adopts a structured conformation owing to multiple intrapeptide backbone hydrogen bonds as well as interaction of Asp2 sidechain with the sidechain and backbone amide of Ser4 (Figure 3(E)). Therefore, a strong bias towards Asp at position 2 of all identified sequences may stem from its importance for maintaining the complementary folds of both the peptide and the substrate binding pocket. Overall, the structures reveal pliability of the hGid4 substrate-binding tunnel enabling interactions with a range of N-terminal sequences, which themselves may also contribute interactions by conformational complementarity.

Yeast GID substrate receptors recognize natural degrons with suboptimal affinity

To extend our findings to the yeast GID system, we screened the phage peptide library for binders to the yGid4 and yGid10 substrate receptors. The selected consensus sequence binding yGid4 paralleled that for hGid4 (Figure 4(A); Table S2), in agreement with their being true orthologs. Remarkably, despite high similarity to the Gid4s, and its only known endogenous substrate likewise initiating with a Pro,[36] the selections with yGid10 identified 12 unique sequences, some with bulky hydrophobic residues and others with Gly prevalent at position 1, each followed by a distinct downstream pattern (Figure 4(B); Table S3). By solving an X-ray structure of yGid10 bound to FWLPANLW peptide and superimposing it on its prior structure with N-terminus of its *bona fide* substrate Art2,[36] we confirmed that the novel sequence is accommodated by the previously characterized binding pocket of yGid10 (Figures 4(C) and S3(A); Table S1). Moreover, conformations of the yGid10 loops varied in complexes with different peptides,[36,45] suggesting like hGid4, yGid10 structural pliability allows recognition of various N-terminal sequences (Figure S3(B)).

Then, we sought to quantitatively compare binding of the new sequences to respective substrate receptors. Affinities of yGid10 for Phe and Gly-initiating sequences, measured by ITC, were comparable to and 2-fold greater than for a peptide corresponding to the N-degron of a natural substrate Art2[36] (Figures 4(D) and S3(C)). Notably, the endogenous degron, and selected sequences, bind yGid10 10- to 20-fold more tightly than the Pro-initiating sequence previously identified by a yeast two-hybrid screen.[35] Although yGid4 is not amenable to biophysical characterization, we could rank-order peptides by inhibition of ubiquitylation of a natural GID^{SR4} substrate Mdh2 (Figure 4(E)). Comparing IC_{50} values for the different peptides led to two major conclusions: (1) the phage display-selected sequences are better competitors than N-terminal sequences of endogenous gluconeogenic substrates, and (2) natural substrate N-terminal sequences themselves exhibit varying suppressive effects, with degron of Mdh2 being the most potent, followed by those of Fbp1 and Icl1.

To test if the novel sequences can mediate binding of substrates for ubiquitylation, we performed two experiments. First, we connected a yGid4- and a yGid10-binding sequence to a lysine via a flexible linker designed based on prior structural modelling.[38] The peptides also had a C-terminal fluorescein for detection. Incubating the peptides with either GID^{SR4} or GID^{SR10} and ubiquitylation assay mixes revealed that each serves as a substrate only for its cognate E3, with low activity

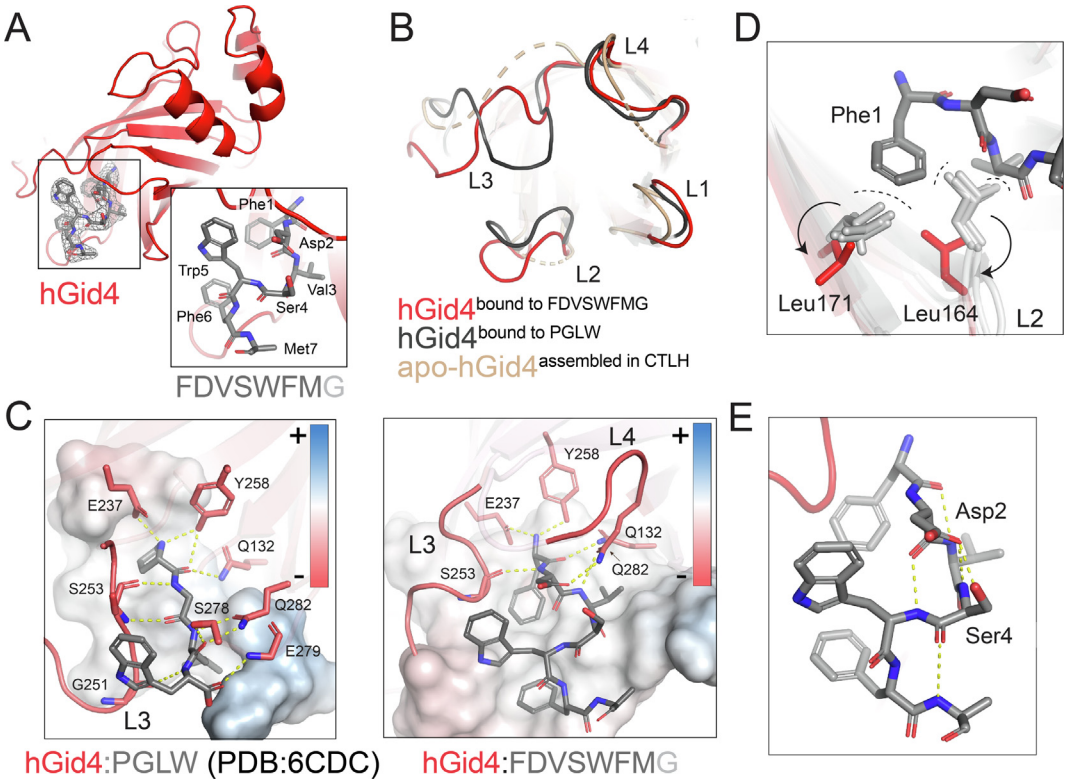
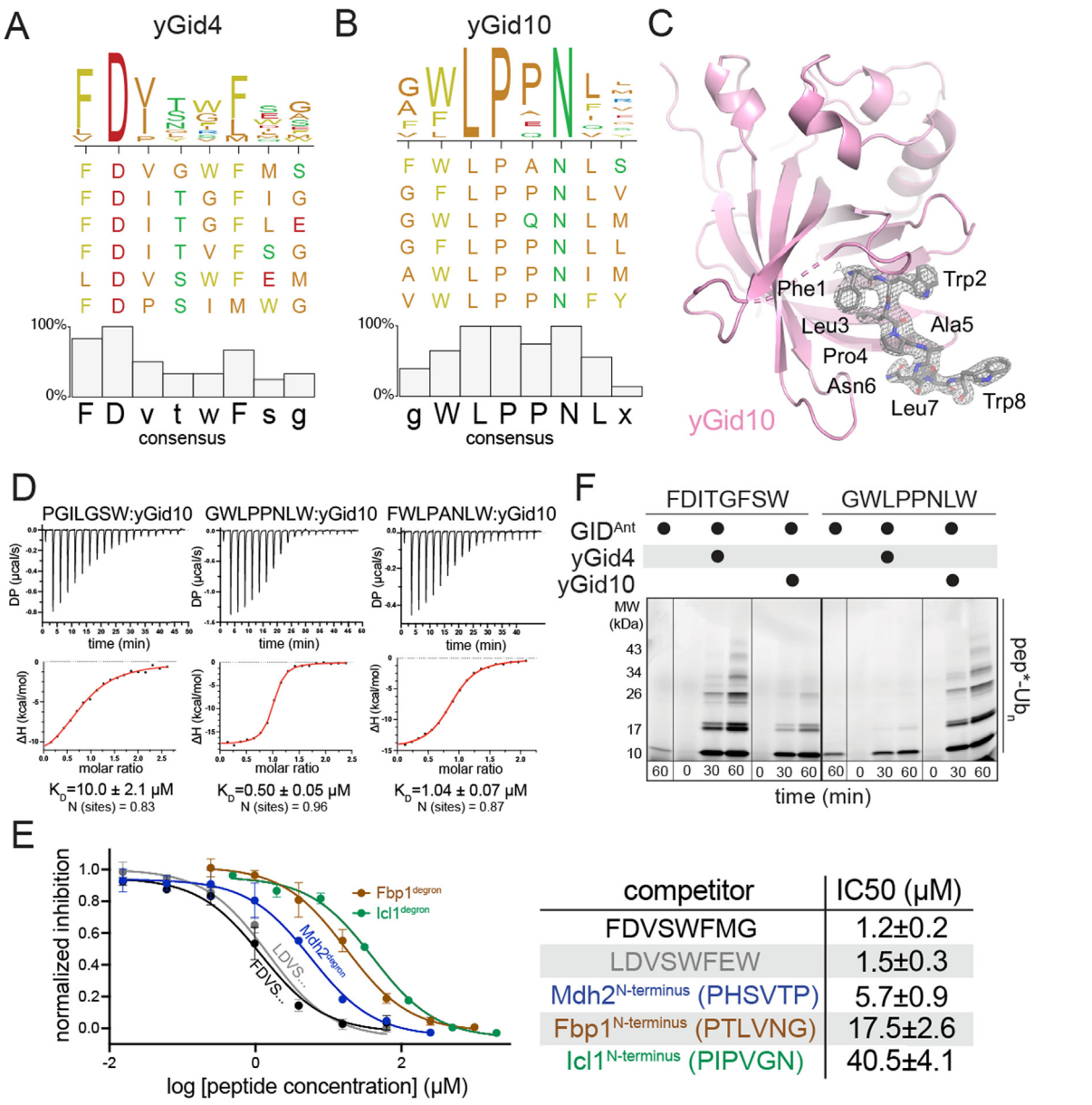


Figure 3. Molecular details of high-affinity peptide binding by hGid4. A. Crystal structure of hGid4 (Δ1-120, Δ294-300) bound to the FDVSWFMG peptide. Clear electron density ($2F_O$-F_C, contoured at 1.5 σ; grey mesh) was visible for all peptide residues besides the C-terminal Gly and the sidechain of Met7, presumably reflecting their mobility. B. Conformations of binding tunnel hairpin loops in apo-hGid4 assembled in CTLH^{SR4} (PDB ID: **7NSC**, light brown) as well as PGLW- (PDB ID: **6CDC**, dark grey) and FDVSWFMG-bound (red) hGid4. C. Comparison of PGLW (left) and FDVSWFMG (right) binding modes to hGid4. Hydrogen bonds between hGid4 residues (red sticks) and peptides (dark grey sticks) are depicted as yellow dashes, whereas the predominantly hydrophobic character of the binding tunnel is visualized as electrostatic potential surface (plotted at ± 7 kT/e; surface colored according to the potential: red – negative (-), blue – positive (+), white – uncharged). D. Overlay of hGid4 bound to PGLW (PDB ID: **6CDC**, light grey), IGLWKS (PDB ID: **6WZX**, light grey), VGLWKS (PDB ID: **6WZZ**, light grey) and FDVSWFMG (red) revealing conformational changes of L2 loop, which prevents steric clash (black dashes) between hGid4 Leu164 and Leu171 and N-terminal Phe of the FDVSWFMG peptide. E. Intrapeptide hydrogen bonding pattern (yellow dashes) within FDVSWFMG upon binding to hGid4.

of GID^{SR10} towards the yGid4-binding sequence (Figure 4(F)). Second, we replaced a native N-terminus of Fbp1 with the novel yGid4-binding consensus and performed *in vitro* ubiquitylation assay with two known forms of the GID E3 – the monomeric GID^{SR4} and the oligomeric Chelator-GID^{SR4} (Figure S4(A)). In both cases, the phage display-determined motif potentiated ubiquitylation of Fbp1 as compared to the WT control, indicating that it can mediate ubiquitylation of a full-length substrate.

Although ubiquitylation is typically a prerequisite, a multitude of processes control ubiquitin-mediated proteolysis in cells. Thus, we examined if the novel non-Pro initiating motifs would be sufficient to target Fbp1 for cellular degradation. We used the promoter reference technique, which was pioneered for examining degradation of GID E3 ligase substrates by normalizing for translation of an exogenously expressed substrate (here, C-terminally 3xFLAG-tagged versions of Fbp1) relative to a simultaneously expressed control (here, DHFR).[6,46] Since varying N-terminal sequences are differentially processed by Met-aminopeptidases,[47] or subjected to co-translational N-terminal acetylation[48] that would block binding to yGid4, we employed the previously-described technique of expressing assorted versions of Fbp1 as linear N-terminal fusions to ubiquitin. The N-terminal ubiquitin is cleaved off by deubiquitylating enzymes, revealing the residue following the ubiquitin sequence as a neo N-terminus.[39,49] As shown previously, Fbp1 harboring the native degron, or that replaced by the sequence IGLW that binds yGid4 with 8-fold lower affinity promoted timely degradation in this assay.[39] However, neither of the novel tight binders, initiating with either Phe or Leu, con-



competitor	IC50 (µM)
FDVSWFMG	1.2±0.2
LDVSWFEW	1.5±0.3
Mdh2$^{N\text{-terminus}}$ (PHSVTP)	5.7±0.9
Fbp1$^{N\text{-terminus}}$ (PTLVNG)	17.5±2.6
Icl1$^{N\text{-terminus}}$ (PIPVGN)	40.5±4.1

Figure 4. Identification of novel yGid4 and yGid10-binding sequence motifs superior to natural degrons. A. Consensus motif obtained by multiple sequence alignment of 12 unique yGid4-binding peptide sequences listed in Table S2 (out of which a representative set of 6 sequences is shown). B. Consensus motif obtained by multiple sequence alignment of 12 unique yGid10-binding peptide sequences listed in Table S3 (out of which a representative set of 6 sequences is shown). C. Crystal structure of yGid10 (Δ1-64, Δ285-292) (pink) bound to FWLPANLW (grey sticks). The $2F_O$-F_C electron density map corresponding to the peptide is shown as grey mesh contoured at 2σ. D. ITC binding assays as in Figure 2(C) but quantifying binding of several peptides to yGid10 (Δ1-56). E. Competitive *in vitro* ubiquitylation assays probing binding of two novel Phe- and Leu-initiating sequences to yGid4 (Δ1-115) as compared to N-termini of natural GID substrates (Mdh2, Fbp1 and Icl1). Unlabeled peptides were titrated to compete off binding of fluorescent Mdh2 (labeled with C-terminal fluorescein) to GID^{SR4}, thus attenuating its ubiquitylation. Normalized inhibition (fraction of ubiquitylated Mdh2 at varying concentration of unlabeled peptides divided by that in the absence of an inhibitor) was plotted against peptide concentration. Fitting to log[inhibitor] vs. response model yielded IC50 values and its standard error based on 2 independent measurements. F. Fluorescent scans of SDS-PAGE gels after *in vitro* ubiquitylation of fluorescent model peptides harboring either a yGid4 or yGid10-binding sequence by GID^{Ant} (comprising 2 copies each of Gid1 and Gid8, and one copy each of Gid5, Gid2 and Gid9) mixed with either yGid4 (Δ1-115) or yGid10 (Δ1-56) (forming GID^{SR4} or GID^{SR10}, respectively). The model peptides contained a corresponding phage display-determined consensus at the N-terminus connected to C-terminal fluorescein (indicated by an asterisk) with a flexible linker.

ferred instability (Figure S4(B)). At this point, future studies will be required to determine the molecular basis for defective proteasomal targeting. However, given that these sequences increased ubiquitylation *in vitro*, it is possible that accelerated ubiquitylation could impede degradation for example through mis-recruitment of deubiquitylating enzymes, mis-processing by Cdc48,[50] or more trivially, they may be subject to unknown modifications that inhibit binding to or ubiquitylation by the GID E3.

GID E3 supramolecular assembly differentially impacts catalytic efficiency toward different substrates

We were surprised by the differences in IC_{50} values for the naturally occurring degrons from the best-characterized GID E3 substrates, Fbp1 and Mdh2. We thus sought to compare ubiquitylation of the two substrates, which not only display different degrons but also distinct catalytic domains with unique constellations of lysines. Previous studies showed that ubiquitylation of both substrates depends on coordination of degron binding by yGid4 with placement of specific lysines in the ubiquitylation active site.[34,38] However, while GID^{SR4} is competent for Mdh2 degradation *in vivo*, a distinct E3 assembly – wherein the Gid7 subunit drives two GID^{SR4} complexes into an oval arrangement (Chelator-GID^{SR4}) is specifically required for optimal ubiquitylation and degradation of Fbp1.[38] Two yGid4 subunits in Chelator-GID^{SR4} simultaneously bind degrons from the oligomeric Fbp1, for simultaneous ubiquitylation of specific lysines on multiple Fbp1 protomers.

Much like for Fbp1, addition of Gid7 to GID^{SR4} was shown to affect Mdh2 ubiquitylation *in vitro*, albeit in a more nuanced way.[38] As a qualitative test for avid binding to two degrons from Mdh2 (whose dimeric state was confirmed by SEC-MALS (Figure S5(A)) and homology modeling (Figure S5(B))) we performed competition assays with monovalent (GID^{SR4} alone or with addition of a truncated version of Gid7 that does not support supramolecular assembly) and bivalent (GID^{SR4} with Gid7 to form Chelator-GID^{SR4}) versions of the E3, and lysineless monodentate (Mdh2 degron peptide) and bidentate (Mdh2 dimer) inhibitors (Figure S5(C)). While the two inhibitors attenuated ubiquitylation of Mdh2 to a similar extent in reactions with the monovalent E3s, only the full-length Mdh2 complex substantially inhibited the bivalent Chelator-GID^{SR4}. This suggested that Chelator-GID^{SR4} is capable of avidly binding to Mdh2.

Thus, we quantified roles of the Fbp1 and Mdh2 degrons by measuring kinetic parameters upon titrating the two different GID E3 assemblies. In reactions with monovalent GID^{SR4}, the K_m for Mdh2 was roughly 3-fold lower than for Fbp1, in accordance with differences in degron binding (Figure 5(A) and (B)). Although the higher-order Chelator-GID^{SR4} assembly improved the K_m values for Fbp1 and for Mdh2, the extents differ such that the values are similar for both substrates. Formation of the higher-order Chelator-GID^{SR4} assembly also dramatically increased the reaction turnover number (k_{cat}) for Fbp1, with a marginal increase for Mdh2 (8- vs. 1.4- times higher k_{cat}, respectively), which was already relatively high in the reaction with monomeric GID^{SR4} (Figures 5(C) and S4(C)). Combined with its effects on K_m, formation of the Chelator-GID^{SR4} assembly increased catalytic efficiency (k_{cat}/K_m) more than 100-times for Fbp1 and only 6-fold for Mdh2, which may rationalize Gid7-dependency of Fbp1 degradation.

Beyond avid substrate binding, the multipronged targeting of Fbp1 by Chelator-GID^{SR4} involves proper orientation of the substrate so that specific lysines in metabolic regulatory regions are simultaneously ubiquitylated.[38] To explain the lesser effect of Chelator-GID^{SR4} on catalytic efficiency toward Mdh2, we examined structural models. Briefly, after docking two substrate degrons into opposing yGid4 protomers, we rotated the tethered substrate to place the targeted lysines in the ubiquitylation active sites (Figure S5(E) and (F)). As shown previously, docking either Fbp1 targeted lysine cluster (K32/K35 and K280/K281) places the other in the opposing active site (Figures 5(D) and S5(E)). For Mdh2, upon mutating the individual clusters of preferentially targeted lysines determined previously,[34] we found that K330 is the major ubiquitylation target for Chelator-GID^{SR4} (Figure S5(D)). However, the structural locations of the two K330 residues within the Mdh2 dimer precludes their simultaneously engaging both Chelator-GID^{SR4} active sites (Figures 5(D) and S5(F)). Thus, the distinct constellations of targeted lysines may also contribute to differences in ubiquitylation efficiency.

Degron identity determines K_m for ubiquitylation but differentially impacts glucose-induced degradation of Mdh2 and Fbp1

To assess the roles of differential degron binding in the distinct contexts provided by the Fbp1 and Mdh2 experiments, we examined the effects of swapping their degrons. We first performed qualitative ubiquitylation assays using the simpler GID^{SR4} E3 ligase. Comparing ubiquitylation of fluorescently-labeled Fbp1 and Mdh2 side-by-side showed more Mdh2 is ubiquitylated with more ubiquitins during the time-course of reactions.[38] These properties are reversed when the N-terminal sequence of Mdh2 is substituted for the Fbp1 degron and vice-versa (Figure 6(A)).

Quantifying the K_m values showed that the values for degron-swapped substrates roughly scaled with degron identity (Figures 6(B) and S4(D); for Mdh2 $K_m \approx 1.3$ μM, for degron-swapped $Fbp1^{Mdh2\ degron}$ $K_m \approx 0.8$ μM, for Fbp1 $K_m \approx 3.6$ μM, for degron-swapped $Mdh2^{Fbp1\ degron} \approx 3.5$ μM). Furthermore,

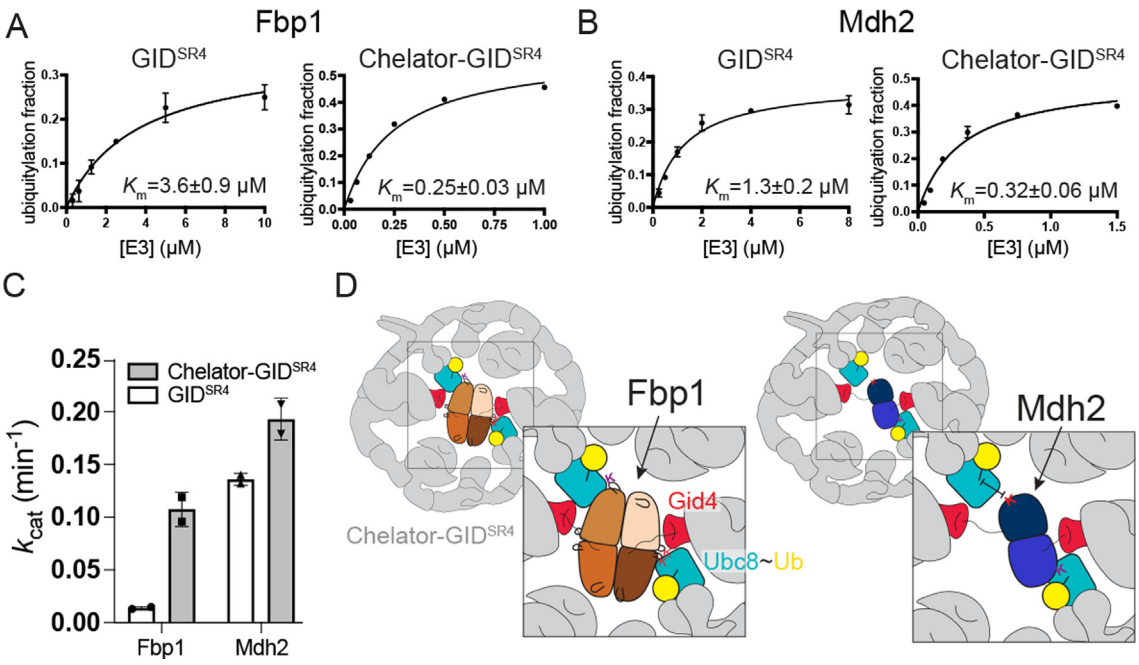


Figure 5. Differential targeting of Mdh2 and Fbp1 by GID E3. A. Plots showing fraction of *in vitro*-ubiquitylated Fbp1 as a function of varying concentration of GID E3 in either its monomeric GID^{SR4} or higher-order Chelator-GID^{SR4} form (co-expressed GID^{SR4} + Gid7). Fitting to Michaelis-Menten equation yielded K_m values. Error bars represent standard deviation ($n = 2$). B. Plots as in (A) but analyzing Mdh2 ubiquitylation. C. Comparison of k_{cat} values for Fbp1 and Mdh2 ubiquitin targeting by GID^{SR4} and Chelator-GID^{SR4} based on a time-course of substrate ubiquitylation (Figure S4(C)). D. Cartoons representing ubiquitylation of Fbp1 and Mdh2 by Chelator-GID^{SR4} based on structural modeling (Figure S5(E) and (F)).

as expected, the K_m values for all substrates improved in reactions with Chelator-GID^{SR4}. However, the relative impact seemed to scale with the way in which they are presented from the folded domain of a substrate rather than the degrons themselves (roughly 14-fold for Fbp1 and 11-fold for Fbp1$^{Mdh2\ degron}$ versus 4-fold for Mdh2 and 6-fold for Mdh2$^{Fbp1\ degron}$).

Effects *in vivo* were examined by monitoring glucose-induced degradation of the wild-type and mutant substrates with the promoter-reference technique.[6,46] As shown previously, Mdh2 was rapidly degraded in the wild-type yeast and the ΔGid7 strain (Figure 6(C)).[38] However, turnover of the mutant version bearing the weaker Fbp1 degron was significantly slower in both genetic backgrounds. Thus, the Mdh2 degron is tailored to the Mdh2 substrate. In striking contrast, although the Mdh2 degron did subtly impact degradation of Fbp1, it was not sufficient to overcome dependency on Gid7 (Figure 6(D)). Thus, substrate ubiquitylation, and turnover, depend not only on degron identity, but also on their associated targeted domains.

Discussion

Overall, our study leads to several conclusions. First, GID/CTLH E3 substrate receptors recognize a diverse range of N-terminal sequences, dictated not only by the N-terminal residue, but also the pattern of downstream amino acids (Figures 1 and S1). Second, such diverse N-terminal sequence recognition is achieved by the combination of (1) a deep substrate-binding tunnel culminating in conserved Glu and Tyr side-chains recognizing the N-terminal amine, (2) pliable loops at the entrance to the substrate binding tunnel that conform to a range of downstream sequences, and (3) the binders themselves forming distinct extended conformations that likewise complement the receptor structures (Figure 3). Remarkably, the hGid4 loops and the bound peptide reciprocally affect each other – peptide binding induces folding of the flexible loops whereas the arrangement of the loops dictates affinity for the bound peptide. This correlation rationalizes strong dependence of Gid4 specificity on the peptide sequence context. Third, the range of interactions result in a range of affinities (Figures 2, 4 and S2(A)). Notably, our randomized phage-display peptide library screen identified far tighter binders to yGid4 than known natural degrons. This approach also generated yGid10-binding sequences with affinities similar to or greater than the only known natural degron, and with significantly higher affinity than a sequence identified by yeast two-hybrid screening. Phage-display peptide library screening may thus prove to be a generally useful method for identifying E3

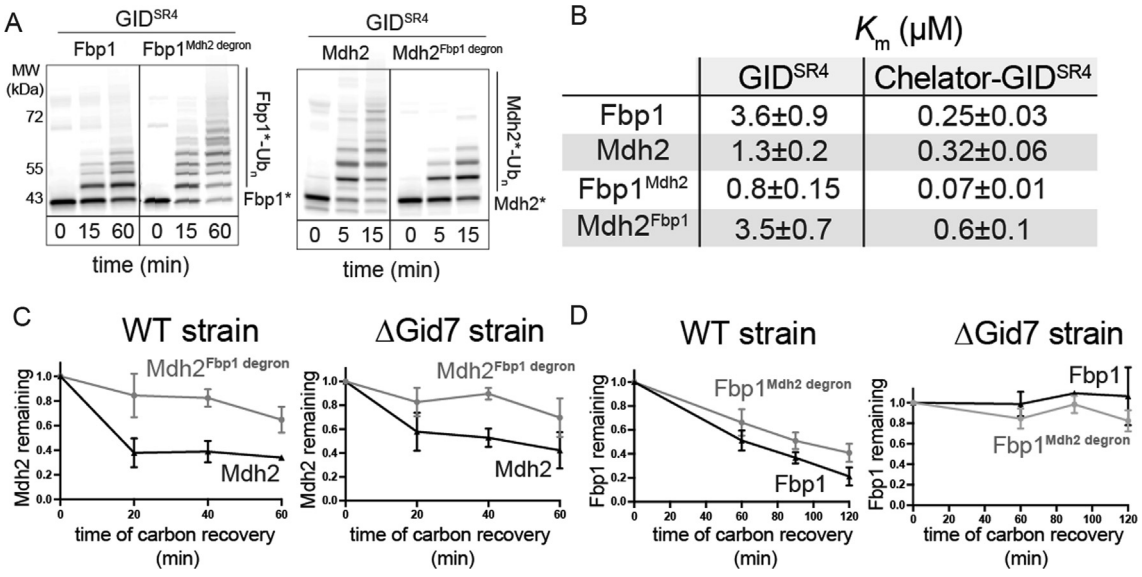


Figure 6. Combinatorial nature of substrate recognition by GID. A. Qualitative *in vitro* ubiquitylation assay probing effect of degron exchange between Fbp1 and Mdh2. Both WT and degron-swapped versions of Fbp1 and Mdh2 were C-terminally labelled with fluorescein (indicated by an asterisk) and ubiquitylated by GID^{SR4}. B. Table summarizing values of K_m for ubiquitylation of WT and degron-swapped substrates by the two versions of GID based on the plots in Figure S4(D). C. *In vivo* glucose-induced degradation of exogenously expressed and C-terminally 3xFlag-tagged Mdh2 as well as its degron-swapped versions quantified with a promoter-reference technique. Levels of the substrates (relative to the level of DHFR) at different timepoints after switch from gluconeogenic to glycolytic conditions were divided by their levels before the switch (timepoint 0). For each substrate, the experiment was performed in WT and ΔGid7 yeast strains. Error bars represent standard deviation ($n = 3$), whereas points represent the mean. D. *In vivo* assay as in (C) but with WT and degron-swapped Fbp1.

ligase binders. Fourth, degron binding is only part of substrate recognition by GID E3s (Figures 5 and 6). Rather, ubiquitylation and degradation depend on both the pairing of a degron with a substrate domain that presents lysines in a particular constellation, and configuration of the GID E3 in either a simplistic monovalent format or in a multivalent chelator assembly specialized for targeting some but not all oligomeric substrates.

Some features of the high-affinity peptide binding by Gid4s and yGid10 parallel other end-degron E3s. Although ubiquitin ligases in the Ubr family employ UBR-box 1 and UBR-box 2 domains with a shallower modes of N-degron recognition,[51–57] C-degron recognition by several cullin-RING ligase substrate receptors involves terminal peptide binding within deep clefts or tunnels[58–61] much like the high-affinity binder interactions with Gid4s and yGid10. Furthermore, end-degron E3 ligases use different strategies to recognize diverse degron sequences. For example, a single Ubr-family E3 can bind different N-terminal sequences through distinct N-degron-binding domains.[62–64] However, much like Gid4s and yGid10 recognize diverse N-terminal sequences, the substrate-binding site of a single cullin-RING ligase was recently shown to bind interchangeably to a C-degron or to a different substrate's internal sequence.[60,61,65]

To-date, few GID E3 substrates have unambiguously been identified. Thus, our findings may have implications for identifying new substrates. Most of the currently characterized substrates depend on co-translational generation of an N-terminal Pro. However, sequences initiating with bulky hydrophobic residues may be refractory to N-terminal processing enzymes such as Met aminopeptidases.[47] Nonetheless, post-translational processing could generate such N-termini. Several paradigms for post-translational generation of N-degrons have been established by studies of Ubr1 substrates. First, endoproteolytic cleavage – by caspases, calpains, separases, cathepsins and mitochondrial proteases[37,66–71] – is responsible for the generation of myriad Arg/N-degron pathway substrates recognized by some Ubr-family E3s.[9] Similarly, N-terminal trimming by aminopeptidases has recently been reported to expose Pro/N-degrons of two yGid4 substrates.[72] Notably, 15 hGid4 interactors reported in the BioGRID database[73] have a solvent-exposed internal [FIL]-D-[VIL] sequence (Figure S6), raising the possibility that the newly identified Gid4-and yGid10-binding motifs likewise could be exposed upon post-translational proteolytic cleavage. Second, some N-degrons are created by aminoacyl-tRNA protein transferases-catalyzed appendage of an

additional amino acid at the protein's N-terminus.[64,74] The bacterial N-degron pathway involves conjugation of hydrophobic residues such as Phe and Leu,[75,76] hence it is tempting to speculate that hydrophobic N-degrons in eukaryotes could likewise involve such N-terminal amino acid addition. Finally, yeast Ubr1 is modulated in an intricate manner: after HtrA-type protease cleavage, a portion of the protein Roq1 binds Ubr1 and alters its substrate specificity.[77] Notably, proteomic studies showed that the human CTLH complex itself associates with the HtrA-type protease HTRA2,[22,24,78–80] known to be involved in mitochondrial quality control.[81,82] This raises the tantalizing possibility that the CTLH E3 might form a multienzyme targeting complex that integrates a regulatory cascade to generate its own substrates or regulatory partners.

The identified sequences might also play various non-degradative functions. Some tight binders to other E3 ligases are pseudosubstrates that modulate subcellular localization,[83,84] or inhibit activity.[85–87] Irrespective of whether such sequences target endogenous proteins to GID/CTLH-family E3 ligases, the identification of nanomolar hGid4 binders and the structural insight into the substrate receptor plasticity may be useful for development of small molecules targeting these E3s.

Finally, our examination of degron-swapped GID E3 substrates Fbp1 and Mdh2 showed that N-terminal sequence is only part of the equation determining ubiquitylation and subsequent degradation. Mdh2 required its own degron and its ubiquitylation and degradation were impaired when substituted with the weaker degron from Fbp1, irrespective of capacity for GID^{SR4} to undergo Gid7-mediated superassembly. However, while either degron could support Fbp1 targeting, this requires Gid7-dependent formation of the chelate-like E3 configuration. Taken together, our data reveal that structural malleability of both the substrate receptor and the E3 supramolecular assembly endows GID E3 complexes – and presumably CTLH E3s as well – capacity to conform to diverse substrates, with varying degrons and associated targeted domains. Such structural malleability raises potential for regulation through modifications or interactions impacting the potential conformations of both the substrate binding domains and higher-order assemblies, and portends future studies will reveal how these features underlie biological functions of GID/CTLH E3s across eukaryotes. Moreover, our results highlight that turnover depends on structural complementarity between E3 and both the substrate degron and ubiquitylated domains, a principle of emerging importance for therapeutic development of targeted protein degradation.

Methods

Reagent table

Reagent/Resource	Reference or Source	Identifier or Catalog Number
Experimental models, cell lines and strains		
High Five Insect cells	Thermo Fisher	Cat#B85502
Saccharomyces cerevisiae: Strain S288C: BY4741; MATa his3Δ1leu2Δ0 met15Δ0 ura3Δ0	Euroscarf	Cat#Y00000
CRLY12; BY4741, Gid4::KANMX	38	N/A
CRLY14; BY4741, Gid7::KANMX	38	N/A
Recombinant DNA		
pCSJ95	6	N/A
pCSJ125	6	N/A
DSJC3; pRS313-pGPD-MPHSVTP-Fbp1(Δ1-7)-3xFLAG-CYC-pGPD-DHFR-HA-CYC	This study	N/A
DSJC4; pRS313-pGPD-MPTLVNG-Mdh2(Δ1-7)-3xFLAG-CYC-pGPD-DHFR-HA-CYC	This study	N/A
DSJC5; pRS313-pGPD-Ub-FDITGFSW-Fbp1(Δ1-9)-3xFLAG-CYC-pGPD-DHFR-HA-CYC	This study	N/A
DSJC6; pRS313-pGPD-Ub-LDVSWFEW-Fbp1(Δ1-9)-3xFLAG-CYC-pGPD-DHFR-HA-CYC	This study	N/A
DSJC7; pRS313-pGPD-Ub-IGLW-Fbp1(Δ1-5)-3xFLAG-CYC-pGPD-DHFR-HA-CYC	39	N/A
pLIB Gid4	34	N/A
pLIB Gid7	38	N/A
pBIG2 Gid1:Gid8-TEV-2xS:Gid5:Gid2:Gid9:Gid7	38	N/A
pBIG2 Gid1:Gid8-TEV-2xS:Gid5:Gid4:Gid2:Gid9:	38	N/A

(*continued*)

Reagent/Resource	Reference or Source	Identifier or Catalog Number
Gid7		
pBIG2 Gid1:Gid8-TEV-2xS:Gid5:Gid4:Gid2:Gid9	34	N/A
pBIG2 Gid1:Gid8-TEV-2xS:Gid5:Gid2:Gid9	34	N/A
pGEX GST-TEV-hGid4 (Δ1-115)	This study	N/A
pGEX GST-TEV-Gid7	38	N/A
pGEX GST-TEV-Gid7 (Δ1-285)	38	N/A
pGEX GST-TEV-hGid4 (Δ1-99)	38	N/A
pGEX GST-TEV-hGid4 (Δ1-120, Δ294-300)	This study	N/A
pGEX GST-TEV-yGid4 (Δ1-115)	34	N/A
pGEX GST-TEV-yGid10 (Δ1-57)	34	N/A
pGEX GST-TEV-yGid10 (Δ1-64, Δ285-292)	This study	N/A
pRSF Fbp1-GGGGS-sortag-6xHis	38	N/A
pRSF Mdh2-GGGGS-LPETGG-6xHis	34	N/A
pRSF MPHSVTP-Fbp1 (Δ1-7)-GGGGS-LPETGG-6xHis	This study	N/A
pRSF MPTLVNG-Mdh2 (Δ1-7)-GGGGS-sortag-6xHis	This study	N/A
pRSF GST-TEV-SUMO- FDITGFSW-Fbp1(Δ1-9)-GGGGS-sortag-6xHis	This study	N/A
pRSF Ubc8-6xHis	34	N/A
pRSF 6xHis-hGid4 (Δ1-115)	This study	N/A
pET3b Ub (ubiquitin)	34	N/A
pET29 sortase A	89	N/A
pRSF Mdh2-6xHis	34	N/A
pRSF Mdh2-6xHis K254R/K256R/K259R	This study	N/A
pRSF Mdh2-6xHis K330R	This study	N/A
pRSF Mdh2-6xHis K360R/K361R	This study	N/A
pRSF Mdh2-6xHis K254R/K256R/K259R; K330R; K360R/K361R	34	N/A
Antibodies		
Mouse anti-His	Cell Signaling Technology	Cat#9991
Sheep polyclonal anti-hGid4	38	N/A
Monoclonal ANTI-FLAG M2	Sigma Aldrich	Cat#F1804
Rabbit anti-HA	Sigma Aldrich	Cat#H6908
Goat anti-rabbit IgG Dylight488 conjugated	Invitrogen	Cat#35552
Goat anti-mouse IgG Dylight633 conjugated	Invitrogen	Cat#35512
Goat anti-mouse IgG HRP conjugated	Sigma Aldrich	Cat#A4416; PRID
Chemicals, Enzymes and peptides		
complete EDTA free	Roche	Cat#05056489001
Aprotinin from bovine lung	Sigma	A1153-10MG
Leupeptin	Sigma	L2884-250MG
Benzamidine	Sigma	B6506-25G
GGGGGFYVK-FAM	MPIB	N/A
PGLWKS	MPIB	N/A
IGLWKS	MPIB	N/A
Leu-Gly	Sigma	CAS# 686–50-0
Pro-Gly	Sigma	CAS# 2578–57-6
Ala-Gly	MPIB	CAS# 687–69-4
Gly-Gly	MPIB	CAS# 556–50-3
GVATSLLW	MPIB	N/A
FDVSWFMG	MPIB	N/A
PDVSWFMG	MPIB	N/A
LDVSWFMG	MPIB	N/A
VDVNSLWA	MPIB	N/A
WDVSWV	MPIB	N/A
FDITGFS	MPIB	N/A

(*continued on next page*)

(*continued*)

Reagent/Resource	Reference or Source	Identifier or Catalog Number
GWLPPNLW	MPIB	N/A
PGILGSW	MPIB	N/A
FWLPANLW	MPIB	N/A
PHSVTPWSI	MPIB	N/A
PTLVNGWPR	MPIB	N/A
PIPVGNWTK	MPIB	N/A
VWEVKTNQ	MPIB	N/A
PHSVTPSIEQDSLK	MPIB	N/A
PGLWKS-FAM	MPIB	N/A
GGGGRHDS(P)GLDS(P)MKDEE-FAM	MPIB	N/A
FDITGFSWRDSTEGFTGRGWSGRGWSKGGK-FAM	MPIB	N/A
GWLPPNLWRDSTEGFTGRGWSGRGWSKGGK-FAM	MPIB	N/A
Software		
Phyre[2]	90	http://www.sbg.bio.ic.ac.uk/~phyre2/html/page.cgi?id=index
UCSF Chimera	91	https://www.cgl.ucsf.edu/chimera/
UCSF ChimeraX	92	https://www.rbvi.ucsf.edu/chimerax/
PyMOL v2.1	Schrödinger	https://pymol.org/2/
Coot	93,94	https://www2.mrc-lmb.cam.ac.uk/personal/pemsley/coot/
Phenix	95–97	https://www.phenix-online.org/
Image Studio	LI-COR Biosciences	https://www.licor.com/bio/image-studio/
Fiji/ImageJ	98	https://imagej.net/
GraphPad Prism version 8.0	GraphPad Software	www.graphpad.com
ImageQuant TL Toolbox version 8.2	GE Healthcare	
MicroCal PEAQ-ITC Analysis Software	Malvern Panalytical	

Plasmid preparation and mutagenesis

All the genes encoding yeast GID subunits including the substrate receptors yGid4 and yGid10, as well as Fbp1 and Mdh2 substrates were amplified from *S. cerevisiae* BY4741 genomic DNA. The gene encoding hGid4 was codon-optimized for bacterial expression system and synthesized by GeneArt (Thermo Fisher Scientific).

All the recombinant constructs used for protein expression were generated by Gibson assembly method[99] and verified by DNA sequencing. The GID subunits were combined using the biGBac method[100] into a single baculoviral expression vector. All the plasmids used in this study are listed in the Reagent table.

Bacterial protein expression and purification

All bacterial expressions were carried out in *E. coli* BL21 (DE3) RIL cells in a Terrific Broth medium[101] overnight at 18 °C. All versions of yGid4, yGid10 and hGid4 (except for that used for NMR) were expressed as GST-TEV fusions. The harvested cell pellets were resuspended in the lysis buffer (50 mM HEPES pH 7.5, 200 mM NaCl, 5 mM DTT and 1 mM PMSF), disintegrated by sonication and subjected to glutathione affinity chromatography, followed by overnight cleavage of the eluted proteins at 4 °C with tobacco etch virus[87] protease to release the GST tag. Final purification was performed with size exclusion chromatography (SEC) in the final buffer containing 50 mM HEPES pH 7.5, 200 mM NaCl and 1 mM or 5 mM DTT (for assays and crystal trials, respectively), or 0.5 mM TCEP (for ITC binding assay). Additionally, pass-back over glutathione affinity resin was performed in order to get rid of the remaining uncleaved GST-fusion protein and free GST.

All versions of Ubc8, Fbp1 (except for FDITGFSW-Fbp1) and Mdh2 were expressed with a C-terminal 6xHis tag. The harvested cell pellets were resuspended in the lysis buffer (50 mM HEPES pH 7.5, 200 mM NaCl, 5 mM β-mercaptoethanol, 10 mM imidazole and 1 mM PMSF) and sonicated. Proteins were purified by nickel affinity chromatography, followed by anion exchange and SEC in the final buffer containing 50 mM HEPES pH 7.5, 200 mM NaCl and 1 mM DTT.

To purify FDITGFSW-Fbp1 for fluorescein labeling, it was expressed as N-terminal GST-SUMO fusion. After glutathione affinity chromatography, the GST-SUMO tag was cleaved

off with a SUMO-specific protease SENP2 generating a desired N-terminus. After cleavage, FDITGFSW-Fbp1 was further purified by SEC in the final buffer containing 50 mM HEPES pH 7.5, 200 mM NaCl and 1 mM DTT. The uncleaved GST-SUMO fusion and free GST-SUMO was removed by pass-back over the GST resin.

Untagged WT ubiquitin used for *in vitro* assays was purified via glacial acetic acid method,[102] followed by gravity S column ion exchange chromatography and SEC.

Insect cell protein expression and purification

All yeast GID complexes used in this study were expressed in insect cells. For protein expression, Hi5 insect cells were transfected with recombinant baculovirus variants and grown for 60–72 h in EX-CELL 420 Serum-Free Medium at 27 °C**.** The insect cells were harvested by centrifugation at 450xg for 15 min and pellets were resuspended in a lysis buffer (50 mM HEPES pH 7.5, 200 mM NaCl, 5 mM DTT, 10 μg/ml leupeptin, 20 μg/ml aprotinin, 2 mM benzamidine, EDTA-free complete protease inhibitor tablet (Roche, 1 tablet per 50 ml of buffer) and 1 mM PMSF). All the complexes were purified from insect cell lysates by StrepTactin affinity chromatography by pulling on a twin-Strep tag fused to the Gid8 C-terminus. Further purification was performed by anion exchange chromatography and SEC in the final buffer containing 25 mM HEPES pH 7.5, 200 mM NaCl and 1 mM DTT.

Preparation of fluorescent substrates for in vitro activity assays

C-terminal labelling of Fbp1, Mdh2 and their degron-swapped versions with fluorescein was performed through a sortase A-mediated reaction. The reaction mix contained 50 μM substrate (C-terminally tagged with a sortag (LPETGG) followed by a 6xHis tag), 250 μM fluorescent peptide (GGGGGFYVK-FAM), 50 μM sortase A[89] and a reaction buffer (50 mM Tris pH 8, 150 mM NaCl and 10 mM $CaCl_2$). The reaction was carried out at room temperature for 30 min. After the reaction, a pass-back over Ni-NTA resin was done to get rid of unreacted substrates. Further purification was done with SEC in the final buffer containing 50 mM HEPES pH 7.5, 200 mM NaCl and 1 mM DTT.

^{15}N labelling of hGid4

For NMR experiments, ^{15}N-labeling of 6xHis-hGid4 (Δ1-115) was carried out. Firstly, 50 ml of the preculture was spun at 3000 rpm for 20 mins. The supernatant was then removed and resuspended with 1x M9 cell growth medium (2 g glucose, 5 mg/ml thiamine chloride, 1 M $MgSO_4$, 1 M $CaCl_2$ and 1g $^{15}NH_4Cl$ per liter of 1x M9 medium) containing all essential ions and antibiotics. The cultures were then grown at 37 °C and 200 rpm until it reached the OD_{600} of 0.5–0.8. Subsequently, the temperature was reduced to 23 °C and kept for an hour before inducing with 0.6 M IPTG. The cultures were then kept growing overnight at 23 °C, 200 rpm, harvested and used for protein purification as described in the section "Protein expression and purification" but in the final SEC buffer containing 25 mM phosphate buffer pH 7.8, 150 mM NaCl and 1 mM DTT.

NMR (Nuclear Magnetic Resonance) spectroscopy

NMR experiments were recorded at 298 K on Bruker Avance III 600 MHz spectrometer (at ^{1}H Larmor frequency of 600 MHz) equipped with a 5 mm TCI cryoprobe. Samples at 0.1 mM ^{15}N-labeled hGid4 were prepared in NMR buffer (50 mM HEPES, 100 mM NaCl, pH 7.0) supplemented with 10% D_2O. ^{1}H,^{15}N HSQC (heteronuclear single quantum coherence) correlation spectra were acquired with 2048 × 256 complex points and a recycle delay of 1.2 s, with 24 scans. DMSO references were acquired at the beginning and end of the assay. No differences were observed between them. Spectra in the presence of ligands where measured at 1 mM Pro or Pro-Gly and 0.5 mM PGLWKS peptide.

Phage-displayed N-terminal peptide library construction and selections

A diverse octapeptide N-terminal phage-displayed library was generated for the identification of peptides binding to hGid4 (Δ1-99), yGid4 (Δ1-115) and yGid10 (Δ1-56). An IPTG-inducible P_{tac} promoter was utilized to drive the expression of open-reading frames encoding the fusion proteins in the following form: the stII secretion signal sequence, followed by a random octapeptide peptide, a GGGSGGG linker and the M13 bacteriophage gene-8 major coat protein (P8). The libraries were constructed by using oligonucleotide-directed mutagenesis with the phagemid pRSTOP4 as the template, as described.[103] The mutagenic oligonucleotides used for library construction were synthesized using with NNK degenerate codons (where N = A/C/G/T & K = G/T) that encode all 20 genetically encoded amino acids. The diversity of the library was 3.5×10^9 unique peptides.

The N-terminal peptide library was cycled through five rounds of binding selections against immobilized GST-tagged hGid4, yGid4, and yGid10, as described.[44] Pre-incubation of the phage pools against immobilized GST was performed before each round of selections to deplete non-specific binding peptides. For rounds four and five, 48 individual clones were isolated and tested for binding to the corresponding targets by phage ELISA,[104] and clones with a strong and specific pos-

itive ELISA signal were Sanger sequenced. A total of 41, 12, and 12 unique peptide sequences were identified binding to hGid4, yGid4, and yGid10, respectively, and their sequences were aligned to identify common specificity motifs.

Oligonucleotide used for the Kunkel reaction to construct the library:

GCTACAAATGCCTATGCANNKNNKNNKNNKN NKNNKNNKNNKGGTGGAGGATCCGGAGGA.

Fluorescence polarization (FP) assays

To determine conditions for a competitive FP assay, we first performed the experiment in a non-competitive format. A 2-fold dilution series of hGid4 (Δ1-115) was prepared in the FP buffer containing 25 mM Hepes pH 7.5, 150 mM NaCl, 0.5 mM DTT and 20 nM fluorescent PGLWKS-FAM and a non-binding GGGGRHDS(P)GLDS(P) MKDEE-FAM as a control peptide. The mixed samples were equilibrated at room temperature for 5 min before transferring to Greiner 384-well flat bottom black plates. Then, the polarization values were measured at the excitation and emission wavelengths of 482 nm and 530 nm, respectively using CLARIOstar microplate reader (BMG LABTECH). For each run, the gain was recorded with FP buffer-only control. The data were fit to one site-binding model in GraphPad Prism to determine K_D value.

To compare binding of several unlabeled ligands to hGid4, we performed the FP measurements in a competitive format. Based on the FP plot from hGid4 titration experiment, we identified hGid4 concentration, which resulted in ~60% saturation of the FP signal. Next, 2-fold dilution series of unlabeled competitors was prepared in the FP buffer mixed with 6.8 μM hGid4. After 5 min incubation, the measurement was performed as described above. The data were plotted relative to the FP signal in the absence of an inhibitor as a function of log(ligand concentration) and analyzed with log(inhibitor) vs. response model to determine IC50 values. To determine relative inhibitory strength of the ligands, the determined IC50 values were divided by that of PGLWKS.

Screening of PGLWKS sequence for hGid4 binding using peptide spot array

The array of peptides derived from the PGLWKS sequence with all 20 amino acid substituted at positions 1, 2 and 3 together, 4 and 5 were synthesized on a membrane in the MPIB biochemistry core facility . The membrane blot was first blocked with 3% milk in TBST buffer (20 mM Tris, 150 mM NaCl and 0.1% Tween 20) for 1 h at room temperature. hGid4 (Δ1-99) was diluted to 10 μg/ml in the buffer containing 150 mM NaCl, 25 mM HEPES pH 7.5, 0.5 mM EDTA pH 8.0, 10% glycerol, 0.1% Tween 20, 2% milk and 1 mM DTT and incubated with the blocked membrane overnight at 4 °C with gentle shaking. The membrane was then washed with TBST buffer 3 times, incubated with primary anti-hGid4 sheep polyclonal antibody (1:500) for 3 h with gentle shaking, followed by multiple washing steps with TBST and 1 h incubation with secondary HRP-conjugated anti-sheep (1:5000) antibody. The membranes were again washed multiple times with TBST and hGid4 binding was visualized by chemiluminescence in Amersham Imager 800 (Cytiva).

Isothermal titration calorimetry (ITC) binding assays

To quantify binding of peptides to hGid4 (Δ1-115) and yGid10 (Δ1-56), we employed ITC. All peptides were dissolved in the SEC buffer used for purification of substrate receptors containing 25 mM HEPES pH 7.5, 150 mM NaCl and 0.5 mM TCEP and their concentration was measured by absorbance at 280 nm (if not present in the original sequence, a single tryptophan residue was appended at peptides' C-termini to facilitate determination of peptide concentration). Binding experiments were carried out in the MicroCal PEAQ-ITC instrument (Malvern Pananalytica) at 25 °C by titrating peptides to either hGid4 or yGid10. Peptides were added to individual substrate receptors using 19 × 2 μl injections, with 4 s injection time and 150 s equilibration time between the injections. The reference power was set to 10 μcal/s. The concentration of the peptides and substrate receptors were customized according the estimated K_D values. Raw ITC data were analyzed using One-Set-of-Sites binding model (Malvern Pananalytica) to determine K_D and stoichiometry of the binding events (N). All plots presented in figures were prepared in GraphPad Prism.

Size exclusion chromatography with multiangle light scattering (SEC-MALS)

To determine the oligomeric state of Mdh2, we performed SEC-MALS (conducted in the MPIB Biochemistry Core Facility). For each run, 100 μl Mdh2 at 1 mg/mL were injected onto Superdex 200 column equilibrated with a buffer containing 25 mM HEPES pH 7.5, 150 mM NaCl and 5 mM DTT.

In vitro activity assays

All ubiquitylation reactions were performed in a multi-turnover format in the buffer containing 25 mM HEPES pH 7.5, 150 mM NaCl, 5 mM ATP and 10 mM $MgCl_2$. To quench the reactions at indicated timepoints, an aliquot of the reaction mix was mixed with SDS-PAGE loading buffer. Ubiquitylation of fluorescein-labelled substrates was visualized with a fluorescent scan of an SDS-

PAGE gel with a Typhoon imager (GE Healthcare) and quantified with ImageQuant (GE Healthcare; version 8.2).

To verify whether FDITGFSW and GWLPPNLW can be recognized by, respectively, yGid4 and yGid10 during ubiquitylation reaction (Figure 4(D)), we performed an *in vitro* activity assay with model peptides, consisting of the respective N-terminal sequences connected to a single acceptor lysine with a 23-residue linker and C-terminal fluorescein (the length of the linker was optimized based on the GID^{SR4} structure[38]). To start the reaction, 0.2 μM E1 Uba1, 1 μM E2 Ubc8-6xHis, 0.5 μM E3 GID^{Ant}, 20 μM ubiquitin, 1 μM yGid4 (Δ1-115) or yGid10 (Δ1-56) and 1 μM peptide substrate were mixed and incubated at room temperature.

In order to probe avid binding of Mdh2 to Chelator-GID^{SR4}, we employed a competition ubiquitylation assay (Figure S4(C)). The reactions were initiated by mixing 0.2 μM Uba1, 1 μM Ubc8-6xHis, 0.5 μM E3 GID^{SR4}, 0 or 2 μM Gid7 (WT or its N terminal deletion mutant, Δ1-284), 0.5 μM Mdh2-FAM, 20 μM unlabeled competitor (dimeric Mdh2-6xHis or a peptide comprising Mdh2 N-terminal sequence PHSVTPSIEQDSLK) and 20 μM ubiquitin. GID^{SR4} was incubated with Gid7 for 5 min on ice before the start of the reaction.

To test which of the preferred ubiquitylation sites within Mdh2 determined previously for GID^{SR4}[34] are major ubiquitylation targets of Chelator-GID^{SR4}, we performed an activity assay with WT and mutant Mdh2, in which putative target lysine clusters (K254/K256/K259; K330; K360/K361) were mutated to arginines individually and all together (Figure S5(A)). To start the reaction, 0.2 μM Uba1, 1 μM Ubc8-6xHis, 0.1 μM Chelator-GID^{SR4}, 1 μM WT or mutant Mdh2-6xHis and 20 μM ubiquitin were mixed. After quenching, Mdh2-6xHis and its ubiquitylated versions were visualized by immunoblotting with anti-6xHis primary antibody and HRP-conjugated anti-mouse secondary antibody.

To quantitatively compare recognition of phage display-identified sequences and degrons of natural GID substrates by yGid4, we employed competitive ubiquitylation assays (Figure 4(F)). Unlabeled peptide inhibitors comprising the analyzed sequences were titrated to compete off binding of Mdh2-FAM to GID^{SR4}, thus attenuating its ubiquitylation. Reactions were started by addition of 20 μM ubiquitin to the mixture of 0.2 μM E1 Uba1, 1 μM E2 Ubc8-6xHis, 0.5 μM E3 GID^{Ant}, 1 μM yGid4 (Δ1-115), 0.25 μM Mdh2-FAM and various concentrations of peptide competitors. After 3 min, the reactions were quenched so that their velocities were still in the linear range. The fractions of ubiquitylated Mdh2 in the presence of an inhibitor were divided by that for Mdh2 alone and plotted against peptide concentration. Fitting of the data to [inhibitor] vs. response model in GraphPad Prism yielded IC50 values.

To qualitatively compare degrons of Fbp1 and Mdh2 in the context of full-length substrates (Figure 6(A)), we performed activity assay with WT and degron-swapped versions ($Fbp1^{Mdh2\ degron}$ and $Mdh2^{Fbp1\ degron}$) of the substrates by mixing 0.2 μM E1 Uba1, 1 μM E2 Ubc8-6xHis, 1 μM E3 GID^{Ant}, 2 μM yGid4 (Δ1-115), 0.5 μM WT or mutant version of Fbp1-FAM or Mdh2-FAM and 20 μM ubiquitin. Similarly, we tested if the N-terminal FDITGFSW motif can promote *in vitro* ubiquitylation of Fbp1. The reactions contained 0.2 μM E1 Uba1, 1 μM E2 Ubc8-6xHis, 0.1 μM GID^{SR4} or Chelator-GID^{SR4}, 1 μM of WT or mutant Fbp1 and 20 μM ubiquitin.

Kinetic parameters for ubiquitylation of WT and degron-swapped versions of Fbp1 and Mdh2 were determined as described previously.[38] Briefly, to determine Michaelis-Menten constant (K_m), we titrated E3 (GID^{SR4} or Chelator-GID^{SR4}) at constant substrate concentration kept below K_m (0.5 and 0.1 μM for reactions with GID^{SR4} and Chelator-GID^{SR4}, respectively; Figures 5(A), (B) and S4(E)). The reaction time was optimized so that the velocity of all reactions was in the linear range. The fraction of ubiquitylated substrate was calculated and plotted as a function of E3 concentration in GraphPad Prism and fit to Michaelis-Menten equation to determine K_m. To calculate k_{cat}, time course assays were performed with the ratios of [E3]:K_m and [substrate]:K_m kept the same for all substrates and E3 versions (2.7 and 0.4, respectively; Figure 5(C)). The rates of the reactions were calculated by linear regression in GraphPad Prism from plots of fraction of ubiquitylated substrates vs. reaction time (Figure S4(D)) and converted into initial velocity using the following equation: $V_0 = rate \cdot [substrate]$.

Then, V_{max} was estimated using a modified form of the Michaelis-Menten equation:

$V_{max} = \frac{V_0 \cdot (K_m + [substrate])}{[substrate]}$. To obtain k_{cat} values, V_{max} was divided by the E3 concentration: $k_{cat} = \frac{V_{max}}{[E3]}$.

Yeast strain construction and growth conditions

The yeast strains used in this study are specified in the Reagents table. All the yeast strains were constructed as derivatives of BY4741 using standard genetic techniques and were verified using PCR, DNA sequencing and immunoblotting to confirm protein expression.

In vivo yeast substrate degradation assays

In order to test the effect of degron identity on glucose-induced degradation of GID substrates, we monitored turnover of WT and degron-exchanged versions of Mdh2 and Fbp1, using the promoter reference technique.[46] Initially, WT and ΔGid7 yeast strains were transformed with a plasmid harboring the open reading frame of either Fbp1-3xFLAG, Mdh2-3xFLAG or their mutant versions ($Fbp1^{Mdh2\ degron}$-3xFLAG and $Mdh2^{Fbp1\ degron}$-3xFLAG) and the control protein DHFR-3xHA, both

expressed from identical promoters. Cells were then grown in SD-glucose medium to OD^{600} of 1.0 followed by carbon starvation in SE medium (0.17% yeast nitrogen base, 0.5% ammonium sulfate, 2% ethanol, amino acid mix) for 19 h. Next, yeasts at the equivalent of 1 OD_{600} were transferred to SD-glucose medium containing 0.5 mM tetracycline resulting in translation inhibition induced by its binding to specific RNA-aptamers within ORFs of the examined and control proteins. At the indicated time points, 1 mL of cells were harvested and pellets were flash frozen in liquid nitrogen. Cell lysis was performed by thawing and resuspending the pellets in 800 μL 0.2 M NaOH, followed by 20 min incubation on ice and subsequent centrifugation at 11,200xg for 1 min at 4 °C. The supernatant was removed and pellets were resuspended in 50 μL HU buffer (8 M Urea, 5% SDS, 1 mM EDTA, 100 mM DTT, 200 mM Tris pH 6.8, protease inhibitor, bromophenol blue), heated at 70 °C for 10 min and then centrifuged again for 5 min at 11,200xg and at 4 °C. The substrates and the control protein DHFR were visualized by immunoblotting with, respectively, anti-FLAG or anti-HA primary and DyLight fluorophore conjugated secondary antibodies, and imaged using a Typhoon scanner (GE Healthcare). Quantification was done using the ImageStudioLite software (LI-COR). For the final graphs, the substrate signal was first normalized relative to the DHFR signal and then to the time point zero (before glucose replenishment). Three biological replicates were performed for all the assays.

A similar experiment was done to test is the novel high-affinity yGid4-binding sequences can confer glucose-induced instability onto Fbp1. To enable N-terminal exposure of sequences with N-terminal bulky hydrophobic residues, all Fbp1 versions (FDITGFSW-Fbp1(Δ1-9)-3xFLAG, LDVSWFEW-Fbp1(Δ1-9)-3xFLAG, a positive control IGLW-Fbp1 (Δ1-5)-3xFLAG[39] and Fbp1-3xFLAG) were expressed as N-terminal fusions to ubiquitin as described previously.[39,49] The cleavage of the ubiquitin fusion was confirmed by immunoblotting with anti-FLAG antibodies. The experiment was performed as described above.

X-ray crystallography

All crystallization trials were carried out in the MPIB Crystallization facility. All crystals were obtained by vapor diffusion experiment in sitting drops at room temperature. The diffraction datasets were recorded at X10SA beam line, Swiss Light Source (SLS) in Villingen, Switzerland.

Crystals of hGid4 (Δ1-99) (without a peptide) were obtained at a concentration of 10 mg/ml using 18% PEG 3350, 0.2 M ammonium nitrate and 0.1 M Bis-Tris buffer at pH 7. Crystals were cryoprotected in 20% ethylene glycol and flash-frozen in liquid nitrogen for data collection.

For hGid4 (Δ1-120, Δ294-300) crystals containing FDVSWFM peptide, 9.2 mg/mL of hGid4 was mixed with 600 μM FDVSWFMG peptide and incubated for 1 h on ice before setting up trays. Crystals were obtained using 1.1 M Sodium malonate, 0.3% Jeffamine ED-2001 pH 7 and 0.1 M HEPES pH 7 and cryoprotected using mix of 20% glycerol and 20% ethylene glycol.

Similarly, for yGid10 (Δ1-64, Δ285-292) crystals with the peptide FWLPANLW, the protein was concentrated to 10 mg/mL and mixed with the peptide to obtain final protein and peptide concentrations of 262 μM and 760 μM, respectively (~3-fold molar excess of the peptide). Crystals were obtained using 0.1 M MES pH 6.9 and cryoprotected using 20% ethylene glycol.

All the diffraction data were indexed, integrated, and scaled using XDS package. Phasing was performed through molecular replacement using the previous structure of hGid4 (PDB ID: **6CDC**, in the case of hGid4 with and without a peptide) or cryo EM structure of yGid4 (extracted from PDB ID: **7NS3**, in the case of peptide-bound yGid10) using PHASER module integrated into PHENIX software suite.[95–97] Model building was done using Coot,[93,94] and further refinements were carried out with phenix.refine. Details of X-ray diffraction data collection and refinement statistics are listed in Table S1.

Data availability

The accession codes for the PDB models are available in RCSB as follows: human Gid4 bound to a Gly/N-peptide, PDB ID: **7Q4Y**; human Gid4 bound to a Phe/N-peptide, PDB ID: **7Q50**; yeast Gid10 bound to a Phe/N-peptide, PDB ID: **7Q51**.

All the unprocessed image data have been deposited to Mendeley Data: http://dx.https://doi.org/10.17632/nz5mch8k2w.1.

CRediT authorship contribution statement

Jakub Chrustowicz: Conceptualization, Methodology, Investigation, Resources. **Dawafuti Sherpa:** Conceptualization, Methodology, Investigation, Resources. **Joan Teyra:** Methodology, Investigation, Resources. **Mun Siong Loke:** Investigation, Resources. **Grzegorz M. Popowicz:** Methodology, Investigation. **Jerome Basquin:** Methodology, Investigation. **Michael Sattler:** Methodology, Supervision, Funding acquisition. **J. Rajan Prabu:** Methodology. **Sachdev S. Sidhu:** Conceptualization, Methodology, Investigation, Supervision, Funding acquisition. **Brenda A. Schulman:** Conceptualization, Methodology, Supervision, Funding acquisition.

Acknowledgments

We thank G. Chen and A. Pavlenco for construction of N-terminal peptide phage-displayed

library; A. Varshavsky for promoter reference plasmids; S. Uebel and S. Pettera for peptide synthesis; K. Valer-Saldana and S. Pleyer for assistance with protein crystallization; J. Rech for the preparation of peptide spot arrays; Paul Scherrer Institut, Villigen, Switzerland for provision of synchrotron radiation beamtime at beamlines PXII and X10SA of the SLS; G. Kleiger for guidance regarding kinetics; I. Paron for technical assistance with mass spectrometry; O. Karayel and T. Vu for bioinformatic analysis of hGid4 interactors; S.v. Gronau for maintenance of insect cells and the Schulman lab for advice and support.

The project was funded by the Deutsche Forschungsgemeinschaft (DFG) SFB1035 (B.A.S. and M.S.), and Leibniz Prize SCHU 3196/1 (B.A. S.). Work in the Schulman lab is supported by the Max Planck Society.

Declaration of Competing Interest

B.A.S. is an honorary professor at Technical University of Munich, Germany and adjunct faculty at St. Jude Children's Research Hospital, Memphis, TN, USA, is on the scientific advisory boards of Interline Therapeutics and BioTheryX, and is co-inventor of intellectual property related to DCN1 small molecule inhibitors licensed to Cinsano.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jmb.2021.167347.

Received 31 August 2021;
Accepted 2 November 2021;
Available online 9 November 2021

Keywords:
N-degron pathway;
Phage display;
Ubiquitin;
Protein–protein interaction;
Structural biology

† These authors contributed equally.

References

1. Varshavsky, A., (2019). N-degron and C-degron pathways of protein degradation. *Proc. Natl. Acad. Sci. USA* **116**, 358–366.
2. Bachmair, A., Varshavsky, A., (1989). The degradation signal in a short-lived protein. *Cell* **56**, 1019–1032.
3. Bachmair, A., Finley, D., Varshavsky, A., (1986). In vivo half-life of a protein is a function of its amino-terminal residue. *Science* **234**, 179–186.
4. Tasaki, T., Mulder, L.C., Iwamatsu, A., Lee, M.J., Davydov, I.V., Varshavsky, A., et al., (2005). A family of mammalian E3 ubiquitin ligases that contain the UBR box motif and recognize N-degrons. *Mol. Cell Biol.* **25**, 7120–7136.
5. Tasaki, T., Kwon, Y.T., (2007). The mammalian N-end rule pathway: new insights into its components and physiological roles. *Trends Biochem Sci.* **32**, 520–528.
6. Chen, S.J., Wu, X., Wadas, B., Oh, J.H., Varshavsky, A., (2017). An N-end rule pathway that recognizes proline and destroys gluconeogenic enzymes. *Science* **355**
7. Hammerle, M., Bauer, J., Rose, M., Szallies, A., Thumm, M., Dusterhus, S., et al., (1998). Proteins of newly isolated mutants and the amino-terminal proline are essential for ubiquitin-proteasome-catalyzed catabolite degradation of fructose-1,6-bisphosphatase of Saccharomyces cerevisiae. *J. Biol. Chem.* **273**, 25000–25005.
8. Timms, R.T., Zhang, Z., Rhee, D.Y., Harper, J.W., Koren, I., Elledge, S.J., (2019). A glycine-specific N-degron pathway mediates the quality control of protein N-myristoylation. *Science* **365**
9. Varshavsky, A., (2011). The N-end rule pathway and regulation by proteolysis. *Protein Sci.* **20**, 1298–1345.
10. Dong, C., Zhang, H., Li, L., Tempel, W., Loppnau, P., Min, J., (2018). Molecular basis of GID4-mediated recognition of degrons for the Pro/N-end rule pathway. *Nature Chem. Biol.* **14**, 466–473.
11. Yan, X., Li, Y., Wang, G., Zhou, Z., Song, G., Feng, Q., et al., (2021). Molecular basis for recognition of Gly/N-degrons by CRL2(ZYG11B) and CRL2(ZER1). *Mol. Cell.*
12. Driessen, H.P., de Jong, W.W., Tesser, G.I., Bloemendal, H., (1985). The mechanism of N-terminal acetylation of proteins. *CRC Crit. Rev. Biochem.* **18**, 281–325.
13. Hwang, C.S., Shemorry, A., Varshavsky, A., (2010). N-terminal acetylation of cellular proteins creates specific degradation signals. *Science* **327**, 973–977.
14. Polevoda, B., Arnesen, T., Sherman, F., (2009). A synopsis of eukaryotic Nalpha-terminal acetyltransferases: nomenclature, subunits and substrates. *BMC Proc.* **3** (Suppl 6), S2.
15. Shemorry, A., Hwang, C.S., Varshavsky, A., (2013). Control of protein quality and stoichiometries by N-terminal acetylation and the N-end rule pathway. *Mol. Cell* **50**, 540–551.
16. Wolf, D.H., (2004). From lysosome to proteasome: the power of yeast in the dissection of proteinase function in cellular regulation and waste disposal. *Cell Mol. Life Sci.* **61**, 1601–1614.
17. Santt, O., Pfirrmann, T., Braun, B., Juretschke, J., Kimmig, P., Scheel, H., et al., (2008). The yeast GID complex, a novel ubiquitin ligase (E3) involved in the regulation of carbohydrate metabolism. *Mol. Biol. Cell.* **19**, 3323–3333.
18. Braun, B., Pfirrmann, T., Menssen, R., Hofmann, K., Scheel, H., Wolf, D.H., (2011). Gid9, a second RING finger protein contributes to the ubiquitin ligase activity of the Gid complex required for catabolite degradation. *FEBS Letters* **585**, 3856–3861.
19. Menssen, R., Schweiggert, J., Schreiner, J., Kusevic, D., Reuther, J., Braun, B., et al., (2012). Exploring the topology of the Gid complex, the E3 ubiquitin ligase involved in catabolite-induced degradation of gluconeogenic enzymes. *J. Biol. Chem.* **287**, 25602–25614.
20. Regelmann, J., Schule, T., Josupeit, F.S., Horak, J., Rose, M., Entian, K.D., et al., (2003). Catabolite degradation of fructose-1,6-bisphosphatase in the yeast Saccharomyces cerevisiae: a genome-wide screen identifies eight novel GID genes and indicates the existence of two degradation pathways. *Mol. Biol. Cell.* **14**, 1652–1663.

21. Kobayashi, N., Yang, J., Ueda, A., Suzuki, T., Tomaru, K., Takeno, M., et al., (2007). RanBPM, Muskelin, p48EMLP, p44CTLH, and the armadillo-repeat proteins ARMC8alpha and ARMC8beta are components of the CTLH complex. *Gene* **396**, 236–247.
22. Boldt, K., van Reeuwijk, J., Lu, Q., Koutroumpas, K., Nguyen, T.M., Texier, Y., et al., (2016). An organelle-specific protein landscape identifies novel diseases and molecular mechanisms. *Nature Commun.* **7**, 11491.
23. Cao, W.X., Kabelitz, S., Gupta, M., Yeung, E., Lin, S., Rammelt, C., et al., (2020). Precise Temporal Regulation of Post-transcriptional Repressors Is Required for an Orderly Drosophila Maternal-to-Zygotic Transition. *Cell Rep.* **31**, 107783
24. Lampert, F., Stafa, D., Goga, A., Soste, M.V., Gilberto, S., Olieric, N., et al., (2018). The multi-subunit GID/CTLH E3 ubiquitin ligase promotes cell proliferation and targets the transcription factor Hbp1 for degradation. *Elife* **7**
25. Liu, H., Ding, J., Kohnlein, K., Urban, N., Ori, A., Villavicencio-Lorini, P., et al., (2020). The GID ubiquitin ligase complex is a regulator of AMPK activity and organismal lifespan. *Autophagy* **16**, 1618–1634.
26. Liu, H., Pfirrmann, T., (2019). The Gid-complex: an emerging player in the ubiquitin ligase league. *Biol. Chem.* **400**, 1429–1441.
27. Zavortink, M., Rutt, L.N., Dzitoyeva, S., Henriksen, J.C., Barrington, C., Bilodeau, D.Y., et al., (2020). The E2 Marie Kondo and the CTLH E3 ligase clear deposited RNA binding proteins during the maternal-to-zygotic transition. *Elife* **9**
28. Soni, S., Bala, S., Gwynn, B., Sahr, K.E., Peters, L.L., Hanspal, M., (2006). Absence of erythroblast macrophage protein (Emp) leads to failure of erythroblast nuclear extrusion. *J. Biol. Chem.* **281**, 20181–20189.
29. Zhen, R., Moo, C., Zhao, Z., Chen, M., Feng, H., Zheng, X., et al., (2020). Wdr26 regulates nuclear condensation in developing erythroblasts. *Blood* **135**, 208–219.
30. Pfirrmann, T., Villavicencio-Lorini, P., Subudhi, A.K., Menssen, R., Wolf, D.H., Hollemann, T., (2015). RMND5 from Xenopus laevis is an E3 ubiquitin-ligase and functions in early embryonic forebrain development. *PLoS One* **10**, e0120342
31. Maitland, M.E.R., Kuljanin, M., Wang, X., Lajoie, G.A., Schild-Poulter, C., (2021). Proteomic analysis of ubiquitination substrates reveals a CTLH E3 ligase complex-dependent regulation of glycolysis. *FASEB J.* **35**, e21825
32. Yuan, S., Zhu, H., Gou, H., Fu, W., Liu, L., Chen, T., et al., (2012). A ubiquitin ligase of symbiosis receptor kinase involved in nodule organogenesis. *Plant Physiol.* **160**, 106–117.
33. Menssen, R., Bui, K., Wolf, D.H., (2018). Regulation of the Gid ubiquitin ligase recognition subunit Gid4. *FEBS Letters* **592**, 3286–3294.
34. Qiao, S., Langlois, C.R., Chrustowicz, J., Sherpa, D., Karayel, O., Hansen, F.M., et al., (2020). Interconversion between Anticipatory and Active GID E3 Ubiquitin Ligase Conformations via Metabolically Driven Substrate Receptor Assembly. *Mol. Cell* **77** 150–63 e9.
35. Melnykov, A., Chen, S.J., Varshavsky, A., (2019). Gid10 as an alternative N-recognin of the Pro/N-degron pathway. *Proc. Natl. Acad. Sci. USA* **116**, 15914–15923.
36. Langlois CR, Beier V, Karayel O, Chrustowicz J, Sherpa D, Mann M, et al. A GID E3 ligase assembly ubiquitinates an Rsp5 E3 adaptor and regulates plasma membrane transporters. bioRxiv [Preprint]..
37. Kong, K.E., Fischer, B., Meurer, M., Kats, I., Li, Z., Ruhle, F., et al., (2021). Timer-based proteomic profiling of the ubiquitin-proteasome system reveals a substrate receptor of the GID ubiquitin ligase. *Mol. Cell* **81** 2460–76 e11.
38. Sherpa, D., Chrustowicz, J., Qiao, S., Langlois, C.R., Hehl, L.A., Gottemukkala, K.V., et al., (2021). GID E3 ligase supramolecular chelate assembly configures multipronged ubiquitin targeting of an oligomeric metabolic enzyme. *Mol. Cell* **81** 2445–59 e13.
39. Dong, C., Chen, S.J., Melnykov, A., Weirich, S., Sun, K., Jeltsch, A., et al., (2020). Recognition of nonproline N-terminal residues by the Pro/N-degron pathway. *Proc. Natl. Acad. Sci. USA* **117**, 14158–14167.
40. Peters, E.A., Schatz, P.J., Johnson, S.S., Dower, W.J., (1994). Membrane insertion defects caused by positive charges in the early mature region of protein pIII of filamentous phage fd can be corrected by prlA suppressors. *J. Bacteriol.* **176**, 4296–4305.
41. Rodi, D.J., Soares, A.S., Makowski, L., (2002). Quantitative assessment of peptide sequence diversity in M13 combinatorial peptide phage display libraries. *J. Mol. Biol.* **322**, 1039–1052.
42. Nilsson, I., von Heijne, G., (1992). A signal peptide with a proline next to the cleavage site inhibits leader peptidase when present in a sec-independent protein. *FEBS Letters* **299**, 243–246.
43. Pluckthun, A., Knowles, J.R., (1987). The consequences of stepwise deletions from the signal-processing site of beta-lactamase. *J. Biol. Chem.* **262**, 3951–3957.
44. Tonikian, R., Zhang, Y., Boone, C., Sidhu, S.S., (2007). Identifying specificity profiles for peptide recognition modules from phage-displayed peptide libraries. *Nature Protoc.* **2**, 1368–1386.
45. Shin, J.S., Park, S.H., Kim, L., Heo, J., Song, H.K., (2021). Crystal structure of yeast Gid10 in complex with Pro/N-degron. *Biochem. Biophys. Res. Commun.* **582**, 86–92.
46. Oh, J.H., Chen, S.J., Varshavsky, A., (2017). A reference-based protein degradation assay without global translation inhibitors. *J. Biol. Chem.* **292**, 21457–21465.
47. Sherman, F., Stewart, J.W., Tsunasawa, S., (1985). Methionine or not methionine at the beginning of a protein. *Bioessays* **3**, 27–31.
48. Aksnes, H., Drazic, A., Marie, M., Arnesen, T., (2016). First Things First: Vital Protein Marks by N-Terminal Acetyltransferases. *Trends Biochem. Sci.* **41**, 746–760.
49. Varshavsky, A., (2005). Ubiquitin fusion technique and related methods. *Methods Enzymol.* **399**, 777–799.
50. Barbin, L., Eisele, F., Santt, O., Wolf, D.H., (2010). The Cdc48-Ufd1-Npl4 complex is central in ubiquitin-proteasome triggered catabolite degradation of fructose-1,6-bisphosphatase. *Biochem. Biophys. Res. Commun.* **394**, 335–341.
51. Choi, W.S., Jeong, B.C., Joo, Y.J., Lee, M.R., Kim, J., Eck, M.J., et al., (2010). Structural basis for the recognition of N-end rule substrates by the UBR box of ubiquitin ligases. *Nature Struct. Mol. Biol.* **17**, 1175–1181.
52. Kim, L., Heo, J., Kwon, D.H., Shin, J.S., Jang, S.H., Park, Z.Y., et al., (2021). Structural basis for the N-degron specificity of ClpS1 from Arabidopsis thaliana. *Protein Sci.* **30**, 700–708.
53. Matta-Camacho, E., Kozlov, G., Li, F.F., Gehring, K., (2010). Structural basis of substrate recognition and

specificity in the N-end rule pathway. *Nature Struct. Mol. Biol.* **17**, 1182–1187.
54. Pan, M., Zheng. Q., Wang. T., Liang, L., Mao, J., Zuo, C., et al. (2021). Structural Insights Into the Initiation and Elongation of Ubiquitination by Ubr1. BioRxiv [Preprint]..
55. Roman-Hernandez, G., Grant, R.A., Sauer, R.T., Baker, T. A., (2009). Molecular basis of substrate selection by the N-end rule adaptor protein ClpS. *Proc. Natl. Acad. Sci. USA* **106**, 8888–8893.
56. Schuenemann, V.J., Kralik, S.M., Albrecht, R., Spall, S.K., Truscott, K.N., Dougan, D.A., et al., (2009). Structural basis of N-end rule substrate recognition in Escherichia coli by the ClpAP adaptor protein ClpS. *EMBO Rep.* **10**, 508–514.
57. Wang, K.H., Roman-Hernandez, G., Grant, R.A., Sauer, R.T., Baker, T.A., (2008). The molecular basis of N-end rule recognition. *Mol. Cell* **32**, 406–414.
58. Li, Y., Jin, K., Bunker, E., Zhang, X., Luo, X., Liu, X., et al., (2018). Structural basis of the phosphorylation-independent recognition of cyclin D1 by the SCF (FBXO31) ubiquitin ligase. *Proc. Natl. Acad. Sci. USA* **115**, 319–324.
59. Rusnac, D.V., Lin, H.C., Canzani, D., Tien, K.X., Hinds, T. R., Tsue, A.F., et al., (2018). Recognition of the Diglycine C-End Degron by CRL2(KLHDC2) Ubiquitin Ligase. *Mol. Cell.* **72** 813–22 e4.
60. Chen, X., Liao, S., Makaros, Y., Guo, Q., Zhu, Z., Krizelman, R., et al., (2021). Molecular basis for arginine C-terminal degron recognition by Cul2(FEM1) E3 ligase. *Nature Chem. Biol.* **17**, 254–262.
61. Zhao, S., Ru, W., Chen, X., Liao, S., Zhu, Z., Zhang, J., et al., (2021). Structural insights into SMCR8 C-degron recognition by FEM1B. *Biochem. Biophys. Res. Commun.* **557**, 236–239.
62. Sriram, S.M., Kwon, Y.T., (2010). The molecular principles of N-end rule recognition. *Nature Struct. Mol. Biol.* **17**, 1164–1165.
63. Tasaki, T., Zakrzewska, A., Dudgeon, D.D., Jiang, Y., Lazo, J.S., Kwon, Y.T., (2009). The substrate recognition domains of the N-end rule pathway. *J. Biol. Chem.* **284**, 1884–1895.
64. Tasaki, T., Sriram, S.M., Park, K.S., Kwon, Y.T., (2012). The N-end rule pathway. *Annu. Rev. Biochem.* **81**, 261–289.
65. Manford, A.G., Rodriguez-Perez, F., Shih, K.Y., Shi, Z., Berdan, C.A., Choe, M., et al., (2020). A cellular mechanism to detect and alleviate reductive stress. *Cell* **183** 46–61 e21.
66. Piatkov, K.I., Colnaghi, L., Bekes, M., Varshavsky, A., Huang, T.T., (2012). The auto-generated fragment of the Usp1 deubiquitylase is a physiological substrate of the N-end rule pathway. *Mol. Cell* **48**, 926–933.
67. Rao, H., Uhlmann, F., Nasmyth, K., Varshavsky, A., (2001). Degradation of a cohesin subunit by the N-end rule pathway is essential for chromosome stability. *Nature* **410**, 955–959.
68. Ditzel, M., Wilson, R., Tenev, T., Zachariou, A., Paul, A., Deas, E., et al., (2003). Degradation of DIAP1 by the N-end rule pathway is essential for regulating apoptosis. *Nature Cell Biol.* **5**, 467–473.
69. Piatkov, K.I., Oh, J.H., Liu, Y., Varshavsky, A., (2014). Calpain-generated natural protein fragments as short-lived substrates of the N-end rule pathway. *Proc. Natl. Acad. Sci. USA* **111**, E817–E826.
70. Jin, S.M., Lazarou, M., Wang, C., Kane, L.A., Narendra, D. P., Youle, R.J., (2010). Mitochondrial membrane potential regulates PINK1 import and proteolytic destabilization by PARL. *J. Cell Biol.* **191**, 933–942.
71. Yamano, K., Youle, R.J., (2013). PINK1 is degraded through the N-end rule pathway. *Autophagy* **9**, 1758–1769.
72. Chen, S.J., Kim, L., Song, H.K., Varshavsky, A., (2021). Aminopeptidases trim Xaa-Pro proteins, initiating their degradation by the Pro/N-degron pathway. *Proc. Natl. Acad. Sci. USA* **118**
73. Oughtred, R., Rust, J., Chang, C., Breitkreutz, B.J., Stark, C., Willems, A., et al., (2021). The BioGRID database: A comprehensive biomedical resource of curated protein, genetic, and chemical interactions. *Protein Sci.* **30**, 187–200.
74. Gonda, D.K., Bachmair, A., Wunning, I., Tobias, J.W., Lane, W.S., Varshavsky, A., (1989). Universality and structure of the N-end rule. *J. Biol. Chem.* **264**, 16700–16712.
75. Tobias, J.W., Shrader, T.E., Rocap, G., Varshavsky, A., (1991). The N-end rule in bacteria. *Science* **254**, 1374–1377.
76. Dougan, D.A., Truscott, K.N., Zeth, K., (2010). The bacterial N-end rule pathway: expect the unexpected. *Mol. Microbiol.* **76**, 545–558.
77. Szoradi, T., Schaeff, K., Garcia-Rivera, E.M., Itzhak, D.N., Schmidt, R.M., Bircham, P.W., et al., (2018). SHRED Is a Regulatory Cascade that Reprograms Ubr1 Substrate Specificity for Enhanced Protein Quality Control during Stress. *Mol. Cell.* **70** 1025–37 e5.
78. Botham, A., Coyaud, E., Nirmalanandhan, V.S., Gronda, M., Hurren, R., Maclean, N., et al., (2019). Global Interactome Mapping of Mitochondrial Intermembrane Space Proteases Identifies a Novel Function for HTRA2. *Proteomics* **19**, e1900139
79. Huttlin, E.L., Ting, L., Bruckner, R.J., Gebreab, F., Gygi, M.P., Szpyt, J., et al., (2015). The BioPlex Network: A Systematic Exploration of the Human Interactome. *Cell* **162**, 425–440.
80. Huttlin, E.L., Bruckner, R.J., Paulo, J.A., Cannon, J.R., Ting, L., Baltier, K., et al., (2017). Architecture of the human interactome defines protein communities and disease networks. *Nature* **545**, 505–509.
81. Clausen, T., Kaiser, M., Huber, R., Ehrmann, M., (2011). HTRA proteases: regulated proteolysis in protein quality control. *Nature Rev. Mol. Cell Biol.* **12**, 152–162.
82. Vande Walle, L., Lamkanfi, M., Vandenabeele, P., (2008). The mitochondrial serine protease HtrA2/Omi: an overview. *Cell Death Differ.* **15**, 453–460.
83. Davis, M., Hatzubai, A., Andersen, J.S., Ben-Shushan, E., Fisher, G.Z., Yaron, A., et al., (2002). Pseudosubstrate regulation of the SCF(beta-TrCP) ubiquitin ligase by hnRNP-U. *Genes Dev.* **16**, 439–451.
84. Welcker, M., Larimore, E.A., Frappier, L., Clurman, B.E., (2011). Nucleolar targeting of the fbw7 ubiquitin ligase by a pseudosubstrate and glycogen synthase kinase 3. *Mol. Cell Biol.* **31**, 1214–1224.
85. Cappell, S.D., Mark, K.G., Garbett, D., Pack, L.R., Rape, M., Meyer, T., (2018). EMI1 switches from being a substrate to an inhibitor of APC/C(CDH1) to start the cell cycle. *Nature* **558**, 313–317.
86. Manford, A.G., Mena, E.L., Shih, K.Y., Gee, C.L., McMinimy, R., Martinez-Gonzalez, B., et al., (2021).

Structural basis and regulation of the reductive stress response. *Cell* **184** 5375–90 e16.
87. Miller, J.J., Summers, M.K., Hansen, D.V., Nachury, M.V., Lehman, N.L., Loktev, A., et al., (2006). Emi1 stably binds and inhibits the anaphase-promoting complex/cyclosome as a pseudosubstrate inhibitor. *Genes Dev.* **20**, 2410–2420.
88. Mohamed, W.I., Park, S.L., Rabl, J., Leitner, A., Boehringer, D., Peter, M., (2021). The human GID complex engages two independent modules for substrate recruitment. *EMBO Rep.*.
89. Chen, I., Dorr, B.M., Liu, D.R., (2011). A general strategy for the evolution of bond-forming enzymes using yeast display. *Proc. Natl. Acad. Sci. USA* **108**, 11399–11404.
90. Kelley, L.A., Mezulis, S., Yates, C.M., Wass, M.N., Sternberg, M.J., (2015). The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc.* **10**, 845–858.
91. Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., et al., (2004). UCSF Chimera–a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612.
92. Goddard, T.D., Huang, C.C., Meng, E.C., Pettersen, E.F., Couch, G.S., Morris, J.H., et al., (2018). UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Sci.* **27**, 14–25.
93. Emsley, P., Cowtan, K., (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132.
94. Emsley, P., Lohkamp, B., Scott, W.G., Cowtan, K., (2010). Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501.
95. Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Davis, I.W., Echols, N., et al., (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221.
96. Afonine, P.V., Klaholz, B.P., Moriarty, N.W., Poon, B.K., Sobolev, O.V., Terwilliger, T.C., et al., (2018). New tools for the analysis and validation of cryo-EM maps and atomic models. *Acta Crystallogr. D Struct. Biol.* **74**, 814–840.
97. DiMaio, F., Echols, N., Headd, J.J., Terwilliger, T.C., Adams, P.D., Baker, D., (2013). Improved low-resolution crystallographic refinement with Phenix and Rosetta. *Nature Methods* **10**, 1102–1104.
98. Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., et al., (2012). Fiji: an open-source platform for biological-image analysis. *Nature Methods* **9**, 676–682.
99. Gibson, D.G., Young, L., Chuang, R.Y., Venter, J.C., Hutchison 3rd, C.A., Smith, H.O., (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature Methods* **6**, 343–345.
100. Weissmann, F., Petzold, G., VanderLinden, R., Huis In't Veld, P.J., Brown, N.G., Lampert, F., et al., (2016). biGBac enables rapid gene assembly for the expression of large multisubunit protein complexes. *Proc. Natl. Acad. Sci. USA* **113**, E2564–9.
101. Lobanov, M.Y., Furletova, E.I., Bogatyreva, N.S., Roytberg, M.A., Galzitskaya, O.V., (2010). Library of disordered patterns in 3D protein structures. *PLoS Comput. Biol.* **6**, e1000958
102. Kaiser, S.E., Riley, B.E., Shaler, T.A., Trevino, R.S., Becker, C.H., Schulman, H., et al., (2011). Protein standard absolute quantification (PSAQ) method for the measurement of cellular ubiquitin pools. *Nature Methods* **8**, 691.
103. Chen, G., Gorelik, L., Simon, K.J., Pavlenco, A., Cheung, A., Brickelmaier, M., et al., (2015). Synthetic antibodies and peptides recognizing progressive multifocal leukoencephalopathy-specific point mutations in polyomavirus JC capsid viral protein 1. *MAbs* **7**, 681–692.
104. Deshayes, K., Schaffer, M.L., Skelton, N.J., Nakamura, G. R., Kadkhodayan, S., Sidhu, S.S., (2002). Rapid identification of small binding motifs with high-throughput phage display: discovery of peptidic antagonists of IGF-1 function. *Chem. Biol.* **9**, 495–505.