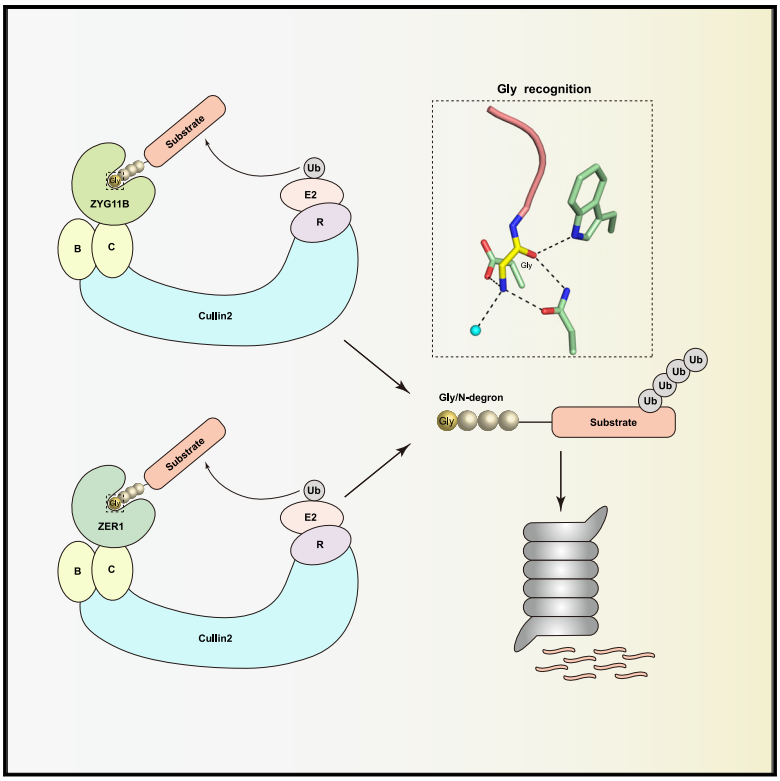
Molecular Cell

Article

Molecular basis for recognition of Gly/N-degrons by $CRL2^{ZYG11B}$ and $CRL2^{ZER1}$

Graphical abstract



Highlights

- Crystal structures of ZYG11B and ZER1 bound to various Gly/N-degrons are solved
- ZYG11B and ZER1 use a deep and narrow cavity to engage Gly/N-degron
- The first four residues of substrates are the key recognition elements
- ZYG11B and ZER1 prefer bulky aromatic residues following the first glycine

Authors

Xiaojie Yan, Yao Li, Guobin Wang, ..., Wenyi Mi, Zhenyi Ma, Cheng Dong

Correspondence

dongcheng@tmu.edu.cn

In brief

$CRL2^{ZYG11B}$ and $CRL2^{ZER1}$ are responsible for degradation of substrates bearing an N-terminal glycine through the proteasome-dependent Gly/N-degron pathway. Yan et al. provide structural insights into recognition of the Gly/N-degron by ZYG11B and ZER1.

Yan et al., 2021, Molecular Cell *81*, 1–13
August 19, 2021
https://doi.org/10.1016/j.molcel.2021.06.010

CellPress

Please cite this article in press as: Yan et al., Molecular basis for recognition of Gly/N-degrons by CRL2^{ZYG11B} and CRL2^{ZER1}, Molecular Cell (2021), https://doi.org/10.1016/j.molcel.2021.06.010

Molecular Cell

CellPress

Article

Molecular basis for recognition of Gly/N-degrons by CRL2^{ZYG11B} and CRL2^{ZER1}

Xiaojie Yan,[1,3] Yao Li,[1,3] Guobin Wang,[1] Zhili Zhou,[2] Guangyong Song,[1] Qiqi Feng,[1] Yueling Zhao,[2] Wenyi Mi,[2] Zhenyi Ma,[1] and Cheng Dong[1,4,*]

[1]Department of Biochemistry and Molecular Biology, The Province and Ministry Co-sponsored Collaborative Innovation Center for Medical Epigenetics, Key Laboratory of Immune Microenvironment and Disease (Ministry of Education), School of Basic Medical Sciences, Tianjin Medical University, Tianjin 300070, China

[2]Department of Immunology, The Province and Ministry Co-sponsored Collaborative Innovation Center for Medical Epigenetics, Key Laboratory of Immune Microenvironment and Disease (Ministry of Education), Tianjin Key Laboratory of Cellular and Molecular Immunology, School of Basic Medical Sciences, Tianjin Medical University, Tianjin 300070, China

[3]These authors contributed equally

[4]Lead contact

*Correspondence: dongcheng@tmu.edu.cn

https://doi.org/10.1016/j.molcel.2021.06.010

SUMMARY

N-degron pathways are a set of proteolytic systems that target the N-terminal destabilizing residues of substrates for proteasomal degradation. Recently, the Gly/N-degron pathway has been identified as a new branch of the N-degron pathway. The N-terminal glycine degron (Gly/N-degron) is recognized by ZYG11B and ZER1, the substrate receptors of the Cullin 2-RING E3 ubiquitin ligase (CRL2). Here we present the crystal structures of ZYG11B and ZER1 bound to various Gly/N-degrons. The structures reveal that ZYG11B and ZER1 utilize their armadillo (ARM) repeats forming a deep and narrow cavity to engage mainly the first four residues of Gly/N-degrons. The α-amino group of the Gly/N-degron is accommodated in an acidic pocket by five conserved hydrogen bonds. These structures, together with biochemical studies, decipher the molecular basis for the specific recognition of the Gly/N-degron by ZYG11B and ZER1, providing key information for future structure-based chemical probe design.

INTRODUCTION

Protein degradation plays an essential role in protein quality control and protein homeostasis maintenance (Balchin et al., 2016; Sontag et al., 2017). The ubiquitin (Ub)-proteasome system (UPS) is the major and selective protein degradation pathway by which cells eliminate dysfunctional proteins to ensure protein homeostasis (Kwon and Ciechanover, 2017). A substrate of the UPS is ubiquitinated through sequential actions of three enzymes—E1 (Ub-activating enzyme), E2 (Ub-conjugating enzyme), and E3 (Ub-protein ligase)—and subsequently degraded by the 26S proteasome (Finley, 2009; Pickart, 2001). Substrate selectivity of ubiquitination depends on the E3 ligase, which specifically recognizes the degradation signal (also known as degron) of a substrate (Bachmair and Varshavsky, 1989; Varshavsky, 2011).

The first degrons discovered, called N-degrons (Bachmair et al., 1986), are the destabilizing N-terminal residues of short-lived proteins that are targeted for proteasomal degradation by N-degron pathways (formerly known as the N-end rule pathways) (Varshavsky, 2019). Other degrons, including internal degrons and C-degrons, were subsequently identified (Chatr-Aryamontri et al., 2018; Koren et al., 2018; Lin et al., 2018; Lucas and Ciulli, 2017; Timms and Koren, 2020). The N-degron pathway is present in many species ranging from mammals, plants, and yeasts (Gonda et al., 1989; Graciet et al., 2009; Holdsworth et al., 2020; Tasaki et al., 2012) to bacteria, which lack the UPS but, instead, utilize the proteasome-like ATP-dependent protease (ClpAP) for substrate degradation (Tobias et al., 1991). In mammals, the N-degron pathways takes part in a broad range of biological processes, such as chromosome segregation, cardiovascular development, neurogenesis, and DNA repair (An et al., 2006; Hwang et al., 2009; Melby et al., 2020). Dysregulation of the N-degron pathway causes various pathological disorders, including neurodegeneration and cancer (Lee et al., 2016; Liu et al., 2016; Varshavsky, 2011).

In eukaryotes, there are several branches of N-degron pathways recognizing distinct N-terminal residues. The Arg/N-degron pathway targets N-terminal primary degrons (Arg, Lys, His, Leu, Trp, Phe, Ile, Tyr, or Met followed by a bulky hydrophobic residue), secondary degrons (Asp or Glu), and tertiary degrons (Cys, Asn, or Gln) (Varshavsky, 2011). The Pro/N-degron pathway recognizes proteins containing an N-terminal proline for degradation (Chen et al., 2017). The Ac/N-degron pathway

Please cite this article in press as: Yan et al., Molecular basis for recognition of Gly/N-degrons by CRL2^{ZYG11B} and CRL2^{ZER1}, Molecular Cell (2021), https://doi.org/10.1016/j.molcel.2021.06.010

recognizes N-alpha-terminal acetylation (Ac-Met, Ac-Ala, Ac-Ser, Ac-Thr, Ac-Val, or Ac-Cys) (Hwang et al., 2010). The formyl-methionine (fMet)/N-degron pathway recognizes N-terminal fMet (Kim et al., 2018). Among all of these pathways, the substrate recognition mechanisms of the Arg/N-degron pathway and Pro/N-degron pathway have been elucidated by structural studies (Choi et al., 2010; Dong et al., 2018, 2020;Matta-Camacho et al., 2010).

Recently, a new branch of the N-degron pathway, the Gly/N-degron pathway, was identified (Timms et al., 2019). An N-terminal glycine residue of proteins is exposed after initiator methionine excision and undergoes N-myristoylation for proper membrane localization (Wright et al., 2010). The Gly/N-degron pathway degrades the proteins that fail to be N-myristoylated and, thus, governs quality control of protein N-myristoylation (Eldeeb et al., 2019; Timms et al., 2019). An N-terminal glycine degron can also be generated by endoproteases such as caspase, suggesting that the Gly/N-degron pathway mediates degradation of caspase-cleaved products during apoptosis (Eldeeb et al., 2019; Timms et al., 2019). Indeed, the Gly/N-degron pathway has been shown to play an important role in activation of the human NLRP1 inflammasome through degradation of the autoinhibitory NLRP1 N-terminal fragment after enteroviral 3C protease cleavage (Robinson et al., 2020).

In the Gly/N-degron pathway, ZYG11B and ZER1, as the substrate receptors of Cullin 2-RING E3 Ub ligase (CRL2), are responsible for recognition of the N-terminal glycine degrons (Timms et al., 2019). However, the molecular mechanism underlying recognition of the unique Gly/N-degron by CRL2^{ZYG11B} and CRL2^{ZER1} remains unknown. Here we report the crystal structures of ZYG11B and ZER1 bound to various Gly/N-degrons. Our structural analysis, together with biochemical and cell-based experiments, unravels the essential elements required for specific recognition of the Gly/N-degron by CRL2^{ZYG11B} and CRL2^{ZER1}. Our study provides the molecular basis for future development of small-molecule inhibitors of CRL2^{ZYG11B} and CRL2^{ZER1}.

RESULTS

ARM repeats of ZYG11B and ZER1 mediate recognition of the Gly/N-degron

ZYG11B and ZER1 contain an N-terminal VHL (von Hippel-Lindau) motif, three LRRs (leucine-rich repeats) in the middle, and eight ARM (armadillo) repeats at the C terminus (Figure 1A). The CRL2 E3 ligase is known to utilize a variety of repeat-containing receptors to recognize C-degrons (Koren et al., 2018; Lin et al., 2018). For instance, CRL2^{KLHDC2} binds the Gly/C-degron through the Kelch repeats of KLHDC2, and CRL2^{FEM1C} engages the Arg/C-degron by the Ankyrin repeats of FEM1C (Chen et al., 2021; Rusnac et al., 2018; Yan et al., 2021). Therefore, we hypothesized that CRL2^{ZYG11B} and CRL2^{ZER1} bind the Gly/N-degrons through the ARM repeats of ZYG11B and ZER1. We generated multiple truncated fragments of ZYG11B (named ZYG11B-F1, -F2, -F3, -F4, and -F5) that contain different numbers of the ARM repeats (Figure 1A), and we also generated Gly/N-degron-glutathione S-transferase (GST) fusion proteins that contain a known Gly/N-degron heptapeptide derived from GIMAP5 or ZNF701 (Timms et al., 2019) to the N terminus of GST (hereafter referred to as GST-GIMAP5 degron and GST-ZNF701 degron, respectively) (Figure 1B). The N-terminal Gly/N-degron of GIMAP5 or ZNF701 can be exposed by TEV (tobacco etch virus) protease cleavage of an engineered cleavage site. Next we performed GST pull-down assays using GST-GIMAP5 degron or GST-ZNF701 degron incubated with ZYG11B fragments (Figure 1C). The results showed that ZYG11B-F1, -F2, and -F3 containing ARM3–ARM8 were able to bind both Gly/N-degrons. In contrast, ZYG11B-F4 and ZYG11B-F5 failed to pull down any Gly/N-degrons. These results suggest that the ARM3–ARM8 repeats of ZYG11B are essential for recognition of the Gly/N-degron. To further substantiate the conclusion, we synthesized two heptapeptides corresponding to the N-terminal sequences of GIMAP5 (GGFQRGK) and ZNF701 (GFLHVGQ) and carried out isothermal titration calorimetry (ITC) assays (Figures 1D–1G). ZYG11B-F3 exhibited robust interactions with the GIMAP5 peptide (K_D of 3.3 μM) and the ZNF701 peptide (K_D of 2.5 μM). Similarly, the ARM3–ARM8 repeats of ZER1 also showed strong binding to the GIMAP5 and ZNF701 peptides (K_D of 4.8 and 3.8 μM, respectively). These results suggest that the ARM repeats of ZYG11B and ZER1 participate in recognition of Gly/N-degrons and that ARM3–ARM8 are sufficient for binding.

Structures of ZYG11B and ZER1 bound to Gly/N-degrons

To determine the molecular basis of Gly/N-degron recognition by ZYG11B and ZER1, we sought to solve the crystal structures of ZYG11B and ZER1 in complex with a Gly/N-degron. To facilitate complex formation, we adopted a similar LC3B fusion technique (Kim et al., 2020) to fuse a polypeptide to the N terminus of ZYG11B or ZER1 protein, in which an N-terminal glycine degron can be exposed by TEV protease cleavage (Figure 2A). Using this strategy, we successfully solved the crystal structure of ZYG11B (ARM2–ARM8) bound to a GIMAP5-like peptide (GGFNR$_{1–5}$) (Table 1). The overall structure reveals that ARM2–ARM8 of ZYG11B stack together to form an arch-shaped conformation with the Gly/N-degron placed in the center (Figure 2B). Each ARM repeat is composed of three α helices (H1, H2, and H3) that are arranged in a right triangle shape. However, in the tandem ARM6 and ARM7, the regions (residues W638–F651 in ZYG11B) involved in formation of ARM6-H3 helix and ARM7-H1 helix form an elongated loop (e-loop) instead of canonical helices. Moreover, this e-loop bends toward to the center of the main body rather than packing against ARM6-H2 with an equivalent curvature, leading to a 90° twisted arrangement of ARM7-ARM8 (Figure 2C). As a result, the inner H3 helices of ARM3-ARM4 together with the e-loop form a deep cavity for recognition of the Gly/N-degron (Figure 2D).

We also determined the crystal structure of ZER1 (ARM3–ARM8) bound to the ZNF701 peptide (GFLHVG$_{1–6}$) at 2.1-Å resolution (Table 2). The overall structures of ZYG11B and ZER1 are very similar, with a root-mean-square deviation (RMSD) of 1.2 Å for the Cα atoms of ARM3–ARM8 (Figure 2E). Intriguingly, the upstream portions of the e-loops are structurally disordered with a sequence variant, but the downstream regions of the e-loops that contribute to the interactions with the Gly/N-degron can

Please cite this article in press as: Yan et al., Molecular basis for recognition of Gly/N-degrons by CRL2^{ZYG11B} and CRL2^{ZER1}, Molecular Cell (2021), https://doi.org/10.1016/j.molcel.2021.06.010

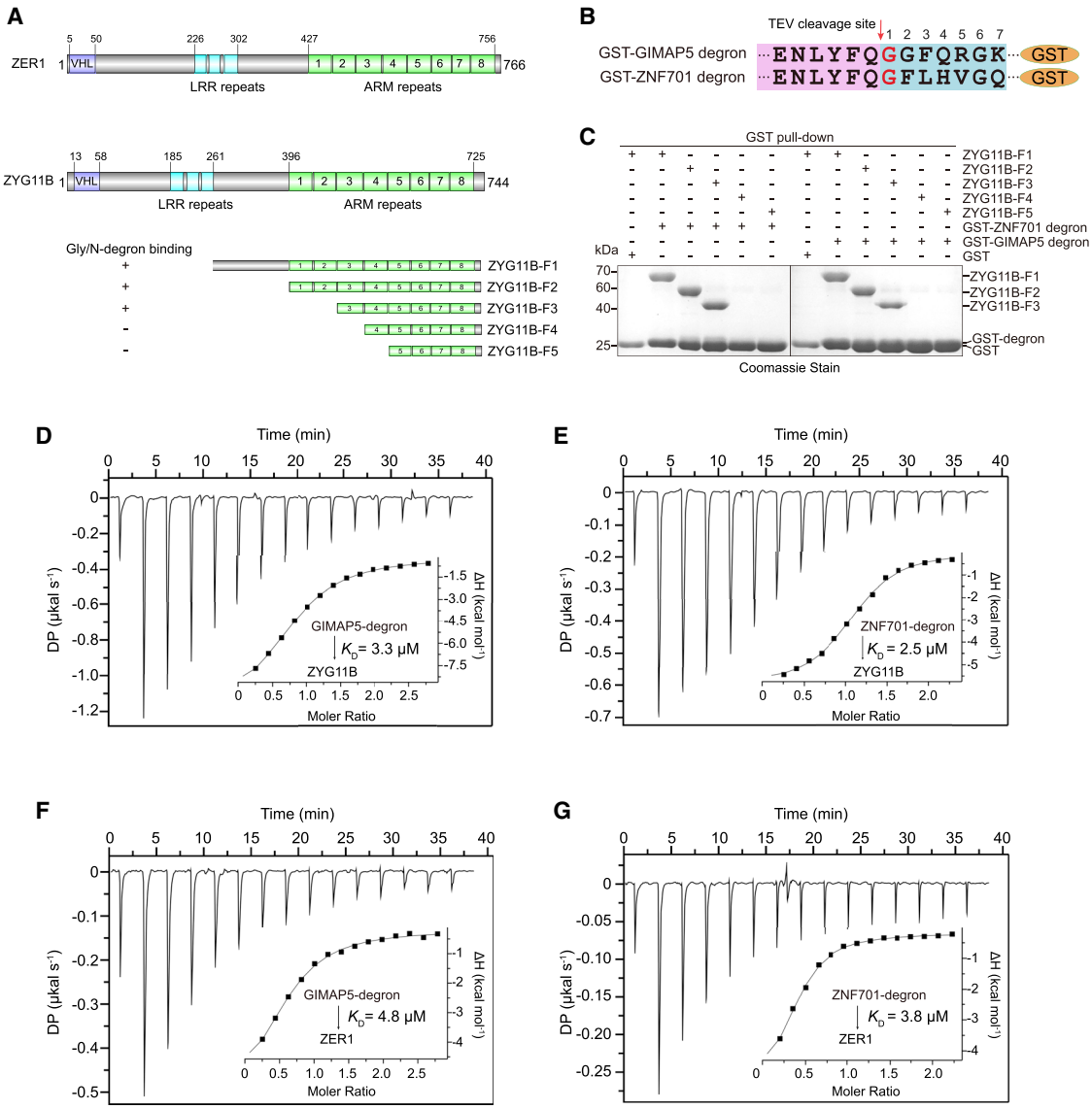


Figure 1. ARM3–ARM8 of ZYG11B and ZER1 are required for Gly/N-degron binding

(A) Top: domain architectures of human ZYG11B and ZER1. VHL, von Hippel-Lindau motif; LRR, leucine-rich repeat; ARM, Armadillo repeat. Bottom: fragments of ZYG11B used in GST pull-down experiments.

(B) Schematic of GST-GIMAP5 degron and GST-ZNF701 degron. The TEV protease cleavage site was introduced to expose the N-terminal glycine degron during the purification.

(C) GST-GIMAP5 degron or GST-ZNF701 degron was used to pull down purified SUMO (small ubiquitin-like modifier)-tagged ZYG11B fragments.

(D–G) ITC titration and fitting curves of ZYG11B (residues 490–728) and ZER1 (residues 469–766) with different Gly/N-degrons; the corresponding degrons and binding affinities are indicated.

be well superimposed. Furthermore, the residues composing the degron-binding pockets are conserved (Figure S1), suggesting that ZYG11B and ZER1 share an evolutionarily conserved substrate recognition mechanism in the Gly/N-degron pathway.

The recognition mode of Gly/N-degrons

To gain better insights into the molecular mechanism of ZYG11B- and ZER1-mediated recognition of Gly/N-degrons, we determined the crystal structures of ZYG11B bound to various Gly/N-degrons, including GGFN, GFLH, and GSTE degron peptides. The crystallography data and model refinement statistics are summarized in Table 1. In all complex structures, recognition of Gly/N-degrons is achieved in a similar fashion, and the electron density maps of the Gly/N-degron peptide are clearly traced (Figures S2A–S2C). In particular, the N-terminal glycine (G1) is deeply buried at the bottom of the cavity together with multiple water molecules that fill up the pocket (Figure 3A). This G1-binding pocket is enriched by acidic residues with a limited space, best designed for accommodating the free and small glycine. The second residue (such as G2, F2, or S2) is

Please cite this article in press as: Yan et al., Molecular basis for recognition of Gly/N-degrons by CRL2^{ZYG11B} and CRL2^{ZER1}, Molecular Cell (2021), https://doi.org/10.1016/j.molcel.2021.06.010

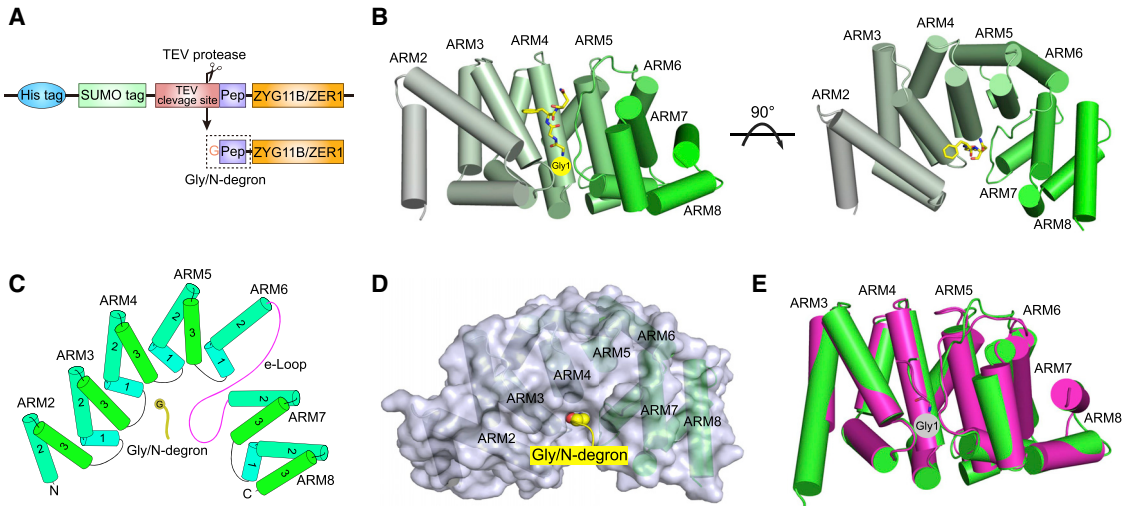


Figure 2. Crystal structures of ZYG11B and ZER1 bound to Gly/N-degrons

(A) Schematic of the recombinant constructs used for the crystallization assay. The Gly/N-degron peptide was fused to the N terminus of ZYG11B or ZER1 and was exposed by TEV protease cleavage during purification.

(B) Left panel: overall structure of ZYG11B (residues 445–728) bound to the GGFN peptide. The Gly/N-degron peptide (GGFN) is shown as yellow sticks, and ARM2–ARM8 are shown as green cylinders. Right panel: view of a 90° rotation around the horizontal axis.

(C) Topology diagram of the ARM repeat arrangement of ZYG11B. The e-loop is colored magenta, and Gly/N-degron is shown in yellow.

(D) The degron-binding pocket formed by ARM repeats of ZYG11B. The glycine degron is shown as yellow spheres, and ARM2–ARM8 are shown as cylinders and a gray surface.

(E) Superposition of ZYG11B (green) and ZER1 (magenta) bound to Gly/N-degrons. ZYG11B and ZER1 adopt conserved structural architectures.

See also Figure S1.

embedded into an inverted L-shaped pocket in which the backbone is inserted in the binding cavity and its side chain is extended into an open groove on the surface of ZYG11B. This open groove is large enough to allow binding of bulky hydrophobic phenylalanine (Figures 3B and 3C). Superimposition of the bound peptides in ZYG11B structures revealed that all backbones of the first two residues are rigid and structurally conserved (Figure 3G). However, the third and fourth residues are located adjacent to the entrance of the cavity, and their side-chain topologies are variable, depending on the specific intermolecular interactions (Figure 3F). The fifth and the succeeding residues are expected to be outside of the binding cavity (Figure S2G) and make minor contributions to recognition of Gly/N-degrons.

In parallel, we determined crystal structures of ZER1 bound to the GAGN, GFLH, or GKLH peptide (Figures S2D–S2F and S2H; Table 2). Overall, these tetrapeptides bind to ZER1 in a similar manner as to ZYG11B (Figures 3D, 3E, and 3H). This is reminiscent of the mode of GID4 bound to Pro/N-degrons, in which GID4 mainly recognizes the first four residues of Pro/N-degrons through a deep and narrow cavity (Dong et al., 2018; see below). Indeed, our ITC binding assays demonstrated that the 4-residue peptide retained strong interaction with ZYG11B, but longer peptides carrying five to seven amino acids displayed slightly increased binding (Figure 3I). In contrast, the 3-residue peptide exhibited a dramatic decrease in binding affinity, and the two-residue peptide showed no detectable binding to ZYG11B. Thus, we conclude that the first four residues of Gly/N-degron substrates suffice to confer a high-affinity interaction with ZYG11B.

Key determinants of the N-terminal glycine

The Gly/N-degron is stabilized in the cavity of ZYG11B and ZER1 through multiple intermolecular interactions (Figures 4A and 4B; Figures S3A–S3E). Structural alignment of GFLH-bound ZYG11B and ZER1 showed that the N-terminal glycine residues are constrained at the same site with subtle conformational changes upon the following residues of Gly/N-degrons, demonstrating a highly conserved G1 recognition mechanism shared by ZYG11B and ZER1 (Figure S3F). Specifically, the protonated α-amino group (NH_3^+) of G1 is oriented by three hydrogen bonds, two from the side chains of Asp526 (Asp556 in ZER1) and Asn567 (Asn597 in ZER1) and one water-mediated hydrogen bond with the carboxyl group of Glu570 (Glu600 in ZER1) (Figures 4A and 4B). Besides this conserved water molecule, there are many other water molecules buried in the bottom of this cavity. These water molecules fill up the G1 pocket and establish considerable sets of hydrogen bond networks with the surrounding residues of ZYG11B or ZER1, further stabilizing the α-amino group of G1 (Figures S3C and S3D). Post-translational modifications on the α-amino group would disrupt these contacts to ZYG11B and ZER1. Indeed, N-terminal acetylation of G1 completely abolished Gly/N-degron peptide binding to ZYG11B or ZER1 in our ITC assays (Figures 4C and 4D). Furthermore, N-terminal myristoylation of G1 has been shown previously to prevent ZYG11B- and ZER1-mediated recognition and degradation (Timms et al., 2019).

The carbonyl oxygen atom of G1 forms two hydrogen bonds with the side chains of ZYG11B Asn567 (Asn597 in ZER1) and Trp522 (Trp552 in ZER1). As a result, the G1 residue fits snugly at the bottom of the acidic cavity of ZYG11B/ZER1 through

Please cite this article in press as: Yan et al., Molecular basis for recognition of Gly/N-degrons by CRL2^{ZYG11B} and CRL2^{ZER1}, Molecular Cell (2021), https://doi.org/10.1016/j.molcel.2021.06.010

Table 1. Data collection and refinement statistics of ZYG11B

	ZYG11B-GSTE	ZYG11B-GFLH	ZYG11B-GGFN
PDB accession number	7EP0	7EP1	7EP2
Data collection			
Space group	P 4_3 2_1 2	P 2_1 2_1 2_1	P 2_1 2_1 2_1
Cell dimensions			
a, *b*, *c* (Å)	131.50, 131.50, 87.03	53.09, 97.58, 121.87	91.36, 102.32, 164.56
α, β, γ (°)	90.00, 90.00, 90.00	90.00, 90.00, 90.00	90.00, 90.00, 90.00
Resolution (Å)	92.98–2.16 (2.27–2.16)[a]	76.17–1.85 (1.90–1.85)	102.32–2.38 (2.44–2.38)
R_{sym} or R_{merge}	0.079 (0.812)	0.059 (0.493)	0.110 (1.536)
$I/\sigma I$	26.70(2.60)	26.00 (4.50)	15.80 (2.30)
Completeness (%)	94.0 (70.0)	98.5 (92.3)	99.9 (99.9)
Redundancy	21.9 (12.7)	12.1 (8.9)	13.0 (13.5)
Refinement			
Resolution (Å)	65.75–2.16 (2.232–2.155)	25.84–1.85 (1.918–1.852)	82.28–2.38 (2.465–2.380)
No. of reflections	39,083 (2,730)	53,738 (4,990)	62,511 (6,148)
R_{work}/R_{free}	0.1833/0.2177	0.1837/0.2266	0.1969/0.2542
No. of atoms			
Protein	4,026	3,962	9,119
Water	126	217	64
B-factors			
Protein	51.8	32.4	64.5
Water	51.3	45.8	58.3
RMSDs			
Bond lengths (Å)	0.008	0.007	0.008
Bond angles (°)	0.98	0.75	0.90

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

five conserved hydrogen-bonding contacts. This G1 binding pattern is similar to that of GID4 bound to N-terminal proline. However, GID4 is unable to recognize the N-terminal glycine (Dong et al., 2018), and our ITC measurements showed that neither ZYG11B nor ZER1 is capable of binding to the N-terminal proline-starting peptides (Figures 4C and 4D). These results suggest that the unique identity of the N-terminal residue is the principal recognition determinant and that diverse N-degrons are specifically recognized by cognate E3 ligases for selective degradation through the N-degron pathways.

Preference of the residues following the N-terminal glycine

In the complex structure, the backbone of the peptide position 2 residue is also inserted into the cavity of ZYG11B, running parallel to the indole moiety of Trp522 (Trp552 in ZER1), forming a backbone-π stacking interaction (Figures 4A and 4B). In addition, the backbone carbonyl and amide of position 2 form two strong hydrogen bonds with the backbone amide and carbonyl of Ala647 (Asn679 in ZER1). Unlike G1-interacting residues, this position 2-binding residue is not strictly conserved between ZYG11B and ZER1, but they interact with the second-position residue of Gly/N-degron by virtue of their backbone atoms independent of the side chains, emphasizing the role of the position 2 backbone-mediated hydrogen bond with ZYG11B or ZER1. To explore the specificity of degron recognition by ZYG11B and ZER1, we synthesized position 2-substituted peptides and performed ITC experiments (Figures 4E and 4H). The results showed that both ZYG11B and ZER1 bind preferentially to bulky aromatic residues at position 2 (such as F2 or Y2) but disfavored P2, E2, or I2 following the glycine. Based on the structural properties, even though the backbone of F2 is anchored in the channel, the side chain of F2 would run outside of the channel and extend to a spacious open groove (Figures 3E and 3F), unlike the Pro/N-degron binding pattern, which causes a steric hindrance in the presence of large residues in its narrow position 2-binding pocket (Dong et al., 2018). Consequently, the benzyl group of F2 forms an additional T-shaped π-π interaction with the indole group of Trp522 in ZYG11B (Trp552 in ZER1) (Figure 4B; Figures S3C and S3D), which helps to explain why ZYG11B and ZER1 prefer an aromatic residue rather than aliphatic residue (I2) at position 2. Notably, proline cannot be tolerated at position 2, probably because of the incompatibility between the specific pyrrolidine ring and the indole moiety of Trp522 (Trp552 in ZER1) in the narrow channel. On the other hand, this position 2-binding pocket is enriched with negative charges (Figures 3A–3E), precluding binding of a negatively charged residue (E2) but not positively charged K2 and H2, as revealed by ITC data (Figures 4E and 4H).

Please cite this article in press as: Yan et al., Molecular basis for recognition of Gly/N-degrons by CRL2^{ZYG11B} and CRL2^{ZER1}, Molecular Cell (2021), https://doi.org/10.1016/j.molcel.2021.06.010

Table 2. Data collection and refinement statistics of ZER1

	ZER1-GAGN	ZER1-GFLH	ZER1-GKLH
PDB accession number	7EP3	7EP4	7EP5
Data collection			
Space group	C 2 2 2_1	P $2_12_12_1$	P 2_12_12
Cell dimensions			
a, *b*, *c* (Å)	67.97, 120.42, 68.0	59.94, 70.15, 148.51	69.40, 150.67, 56.38
α, β, γ (°)	90.00, 90.00, 90.00	90.00, 90.00, 90.00	90.00, 90.00, 90.00
Resolution (Å)	59.19–1.51 (1.55–1.51)[a]	63.43–2.07 (2.12–2.07)	52.80–2.02 (2.07–2.02)
R_{sym} or R_{merge}	0.038 (0.440)	0.258 (0.627)	0.086 (0.735)
$I/\sigma I$	30.60 (2.80)	19.50 (8.50)	18.40 (2.30)
Completeness (%)	98.8 (89.5)	99.3 (94.6)	99.4 (95.3)
Redundancy	11.0 (4.8)	10.3 (7.0)	11.3 (5.8)
Refinement			
Resolution (Å)	22.32–1.51 (1.567–1.513)	63.43–2.07 (2.140–2.066)	52.80–2.02 (2.092–2.020)
No. of reflections	43,337 (3,919)	38,679 (3,628)	39,346 (3,720)
R_{work}/R_{free}	0.2189/0.2555	0.1786/0.2185	0.1854/0.2325
No. atoms			
Protein	1964	3955	4006
Water	95	334	193
B-factors			
Protein	28.3	13.3	32.6
Water	32.0	19.2	34.8
RMSDs			
Bond lengths (Å)	0.006	0.007	0.007
Bond angles (°)	0.75	0.84	0.83

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

The backbone amide of peptide position 3 is coordinated by a hydrogen bond with the side-chain carboxamide of Asn523 in ZYG11B (Figure 4A). ITC binding assays showed that ZYG11B can tolerate most amino acids with strong binding of hydrophobic (F3 and P3) and negatively charged residues (E3) at position 3 (Figure 4F). Consistent with these observations, F3 is further stabilized through hydrophobic interaction with Trp522 of ZYG11B (Figure 4A). Additionally, the side chain of the position 3 residue is flexible and can be positioned toward a basic path created by Lys515 and Lys560 (Figure 3F). In contrast to ZYG11B, our ITC results revealed that ZER1 disfavored charged residues E3, D3, and R3 (Figure 4I), probably because of its hydrophobic surface of position 3 (Figure 3E). The fourth-position residue is exposed to solvent and, thus, not involved in specific interactions with ZYG11B/ZER1 (Figures 4A and 4B; Figures S2G and S2H). Therefore, ZYG11B and ZER1 can tolerate diverse types of residues with a preferred hydrophobic tyrosine (Y4) at position 4, as supported by our ITC studies (Figures 4G and 4J).

Collectively, the backbone groups of the first three residues in Gly/N-degron predominantly mediate the interactions with ZYG11B and ZER1 through direct hydrogen bonds. Although ZYG11B and ZER1 share a similar recognition profile of the second position residues, they display distinct preferences for the third-position residues. The first two residues of Gly/N-degron are the key recognition motif for ZYG11B, whereas the recognition motif of ZER1 is highly restricted by the first three residues of Gly/N-degron. Nevertheless, ZYG11B and ZER1 exhibit strong preference for the bulky aromatic residues following the first glycine. These preferences of Gly/N-degron are in agreement with a previous study (Timms et al., 2019).

Key residues of ZYG11B and ZER1 in Gly/N-degron binding

To determine the roles of key residues of ZYG11B/ZER1 in mediating Gly/N-degron binding, we generated a series of single point mutants of ZYG11B and ZER1 and examined their binding affinities toward the GIMAP5 peptide by ITC (Figures 5A and 5B). In line with our structural analysis, substitution of Asp526, Asn567, or Glu570 with alanine, which would disrupt the hydrogen bonds with the α-amino group of G1, resulted in complete loss of degron binding. Accordingly, the alanine mutant of Trp522 that ablates one hydrogen bond with the carbonyl oxygen of G1 significantly decreased the binding affinity by more than 100-fold. These results demonstrate the critical role of hydrogen-bonding-mediated G1 recognition. Similarly, replacement of Asn523 with alanine, which disrupts the hydrogen bond with the backbone amide of the third position, displayed negligible interaction with Gly/N-degron. These results are consistent with GST pull-down assays (Figure 1C), in which the GIMAP5 or

Please cite this article in press as: Yan et al., Molecular basis for recognition of Gly/N-degrons by CRL2^{ZYG11B} and CRL2^{ZER1}, Molecular Cell (2021), https://doi.org/10.1016/j.molcel.2021.06.010

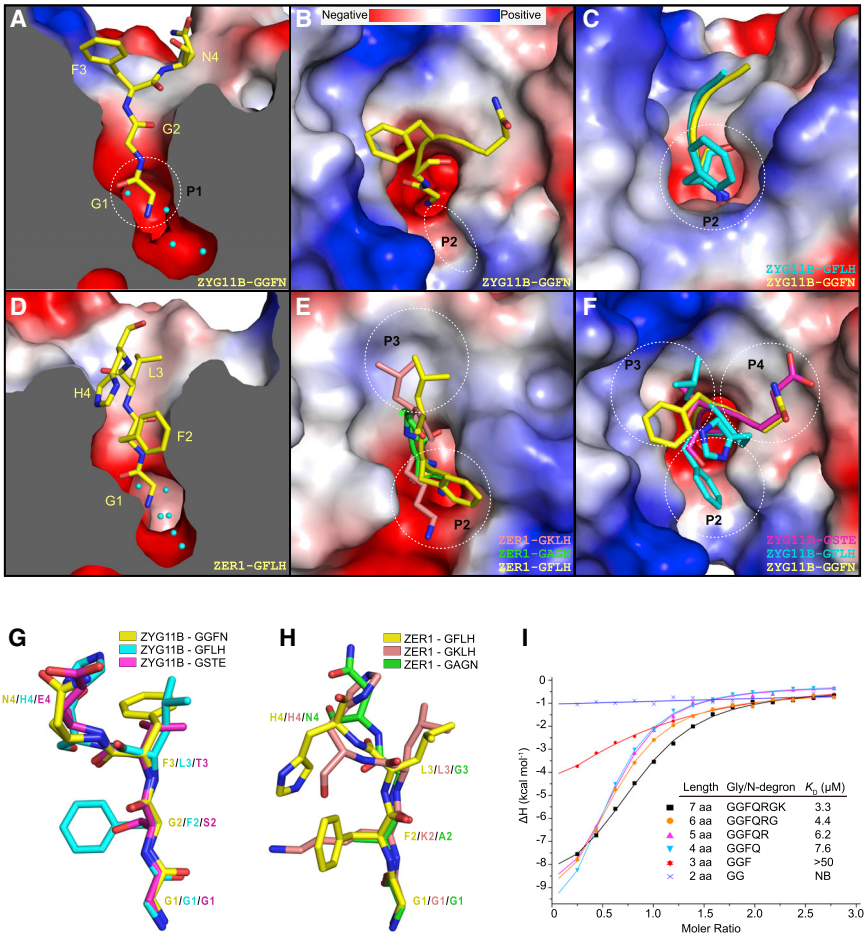


Figure 3. Characterization of Gly/N-degron recognition by ZYG11B and ZER1

(A) Cross-section view of the GGFN-binding pocket of ZYG11B. ZYG11B is shown as an electrostatic potential surface (red, negative; blue, positive). The degron is shown as yellow sticks. Water molecules are shown as cyan spheres.

(B) Top view of the GGFN-binding pocket of ZYG11B, in which ZYG11B forms an inverted L-shaped pocket to bind the second-position (P2) residue of the degron.

(C) Front view of the second-position binding pocket in the structure of ZYG11B with the GGFN or GFLH peptide.

(D) Cross-section view of the GFLH-binding pocket of ZER1. The GFLH peptide is shown as yellow sticks. Water molecules are shown as cyan spheres.

(E) Top view of the degron-binding pocket of ZER1 with different degrons (GKLH, GFLH, and GAGN).

(F) Top view of various Gly/N-degrons bound in the cavity of ZYG11B.

(G) Superposition analysis of different Gly/N-degrons. The entire ZYG11B-peptide complex structures were used to perform the superposition. For clarity, the ZYG11B structure is omitted.

(H) Superposition analysis of different Gly/N-degrons bound with ZER1.

(I) ITC fitting curves of ZYG11B (residues 490–728) titrated with variable lengths of the GIMAP5 degron peptide; the corresponding peptide lengths, sequences, and binding affinities are indicated. NB, no apparent binding under our experimental conditions.

See also Figure S2.

ZNF701 degron failed to bind the ZYG11B-F4 (ARM4–ARM8) or ZYG11B-F5 (ARM5–ARM8) fragment because of loss of the key degron-binding residues (such as Trp522, Asn523, and Asp526 in ARM3 and Asn567 and Glu570 in ARM4). Our ITC studies demonstrated that mutations of the degron-interacting residues in ZER1 (including D556A, N597A, W552A, and N553A) abolished binding with Gly/N-degron (Figure 5B), underscoring the vital role of these conserved residues mediating degron recognition by ZYG11B and ZER1.

To determine the roles of Gly/N-degron recognition by ZYG11B and ZER1 *in vivo*, we evaluated their effects on substrate stabilization of the DBNDD2 and SNX11 proteins using the global protein stability (GPS) system (Figure 5C). In brief, we constructed an Ub-(N-degron)-GFP- porcine teschovirus-1 2A self-cleaving peptide (P2A)-red fluorescent protein (RFP) cascade vector (Timms et al., 2019) in which the Ub-(N-degron)-GFP fusion protein and a RFP that is used as an internal control are produced from the same transcript and separated by P2A self-cleaving. The N-terminal glycine degron of 19-mer DBNDD2 and SNX11 peptides can be exposed through endogenous deubiquitinating enzymes (DUBs). Using this system, we assessed the protein stability of DBNDD2 and SNX11, quantified by GFP/RFP ratio, in HEK293T cells expressing wild-type (WT) ZYG11B or N-degron binding-deficient mutants. We found that overexpression of WT ZYG11B substantially reduced the GFP/RFP ratio by accelerated degradation of the DBNDD2-GFP fusion protein. In contrast, the G1 binding-defective mutants of ZYG11B, including W522A, D526A, and N567A, were severely retarded in DBNDD2 protein degradation, with cells retaining a high GFP/RFP ratio compared with cells expressing WT ZYG11B (Figures 5D and 5E). Moreover, the ZYG11B mutants failed to promote degradation of the SNX11-GFP fusion protein in HEK293T cells (Figures 5F and 5G). Similar to ZYG11B, the ZER1 mutants (W552A, D556A, and N597A) abrogated its interaction with the glycine degron and strongly impaired degradation of the SNX11-GFP fusion protein relative to WT ZER1 in the GPS assay (Figures 5H and 5I).These results corroborate our findings from structural and biochemical studies demonstrating the critical role of the key residues of ZYG11B- and ZER1-mediated Gly/N-degron recognition in substrate degradation *in vivo*.

DISCUSSION

The N-terminal amino acid of a protein acts as a key element to control the protein's half-life through the UPS-dependent N-

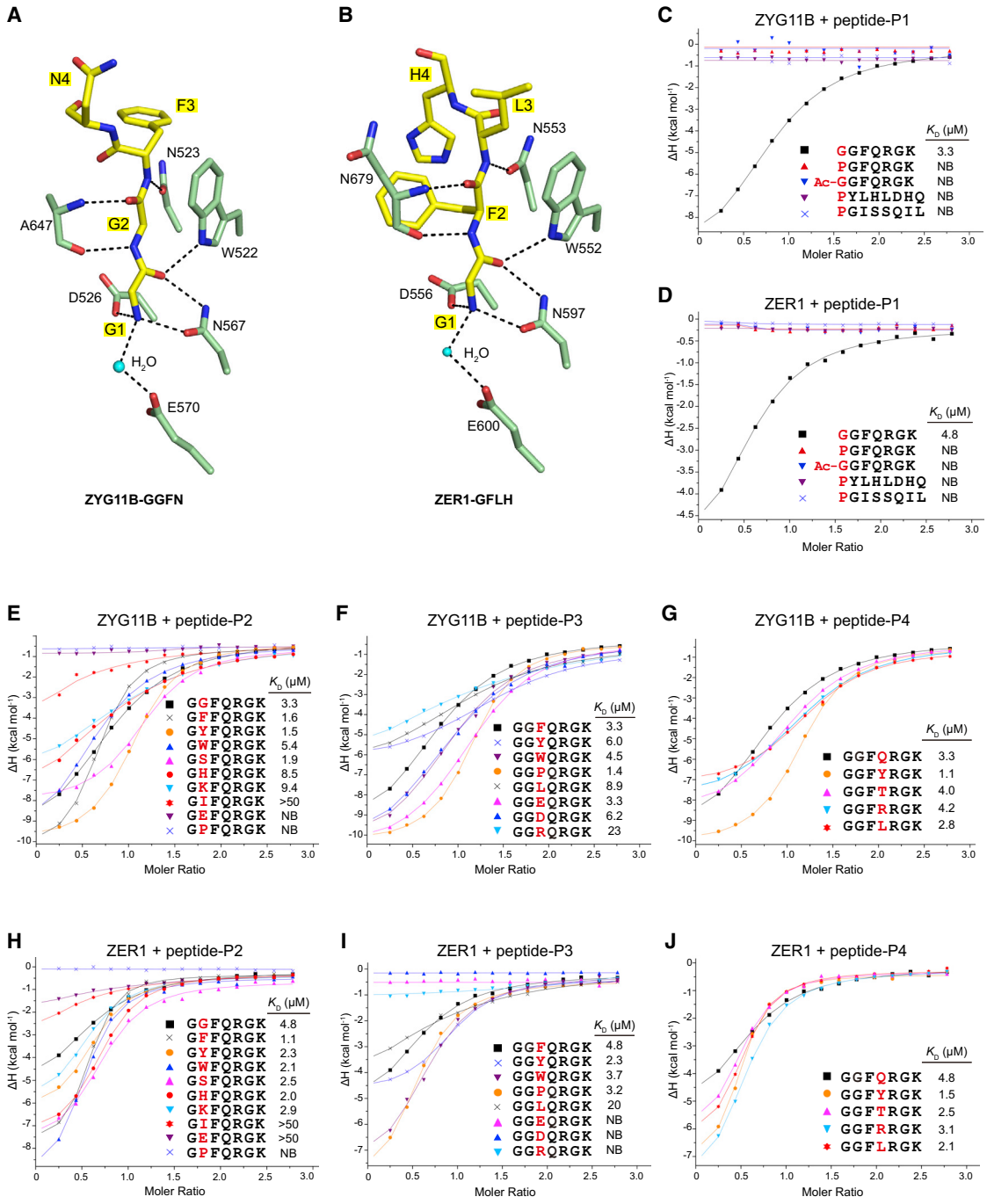


Figure 4. Substrate specificity of ZYG11B- and ZER1-mediated Gly/N-degron recognition

(A) Interactions of ZYG11B with the GGFN peptide. The peptide is shown as yellow sticks, the interacting residues in ZYG11B are shown as pale green sticks, the water molecule is shown as a cyan sphere, and the hydrogen bonds are shown as black dashed lines.

(B) Interactions of ZER1 with the GFLH peptide.

(C–J) ITC fitting curves of ZYG11B (residues 490–728) and ZER1 (residues 469–766) titrated by a series of peptides with position 1–4 substitutions. The corresponding peptide sequences and binding affinities are indicated.

See also Figure S3.

degron pathways. All 20 natural amino acids are thought to act as destabilizing N-terminal residues and compose N-degrons via direct exposure or post-translational modifications. Accordingly, a number of E3 ligases, also known as N-recognins, specifically recognize the N-degrons and promote substrate degradation through diverse N-degron pathways, including the Arg/N-

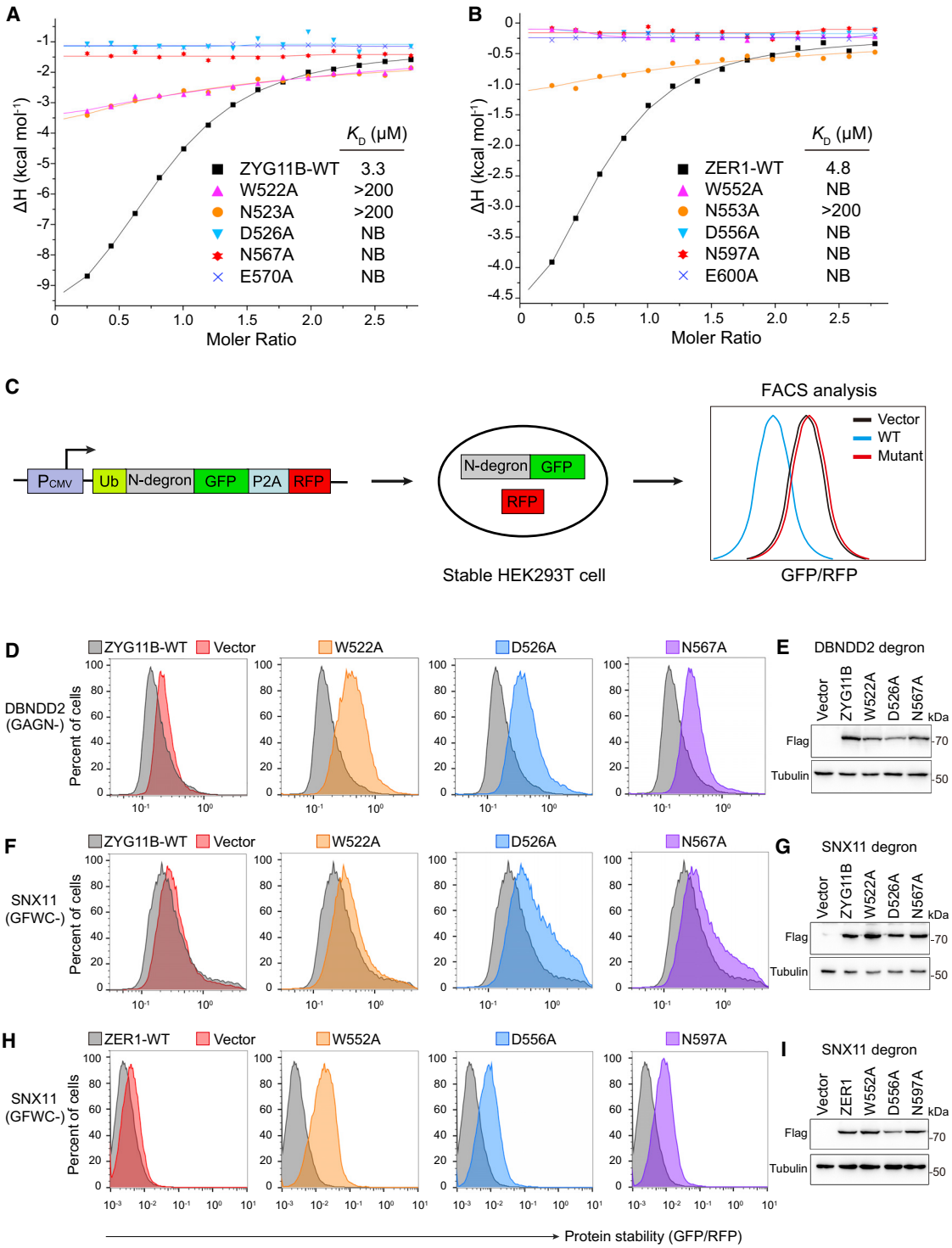


Figure 5. Mutational analysis of degron-binding residues of ZYG11B and ZER1

(A) ITC fitting curves of wild-type (WT) and mutant ZYG11B (residues 490–728) titrated with the GIMAP5 peptide (GGFQRGK); the corresponding mutant proteins and binding affinities are indicated.

(B) ITC fitting curves of WT and mutant ZER1 (residues 469–766) titrated with the GIMAP5 peptide.

(C) Schematic of the GPS assay. P_{CMV}, cytomegalovirus promoter; Ub, ubiquitin; P2A, porcine teschovirus-1 2A self-cleaving peptide; GFP, green fluorescent protein; RFP, red fluorescent protein.

(legend continued on next page)

Please cite this article in press as: Yan et al., Molecular basis for recognition of Gly/N-degrons by CRL2^{ZYG11B} and CRL2^{ZER1}, Molecular Cell (2021), https://doi.org/10.1016/j.molcel.2021.06.010

degron pathway, Ac/N-degron pathway, Pro/N-degron pathway, and Gly/N-degron pathway (Varshavsky, 2019). However, the molecular mechanisms of substrate recognition in the N-degron pathways remain largely unclear, especially for the recently identified Gly/N-degron pathway. Our study uncovers the molecular mechanism of recognition of N-terminal glycine degron by CRL2^{ZYG11B} and CRL2^{ZER1}. Although ZYG11B and ZER1 only have ~26% sequence identity, they share conserved structural architectures. Both proteins utilize a narrow cavity of their C-terminal ARM repeats to engage substrates through a network of hydrogen bonding with the backbone atoms of the Gly/N-degron.

Structural basis for recognition of N-degrons and C-degrons

Since the discovery of N-degron over 30 years ago, a large number of N-degron pathways as well as C-degron pathways, which are topologically analogous to N-degrons, have been identified (Chen et al., 2017; Hwang et al., 2010; Koren et al., 2018; Lin et al., 2018; Timms et al., 2019; Varshavsky, 2011). Structural studies have shed light on the substrate recognition mechanisms of the Arg/N-degron pathway (Choi et al., 2010), Pro/N-degron pathway (Dong et al., 2018), Gly/N-degron pathway (here), as well as the Arg/C-degron pathway (Yan et al., 2021) and Gly/C-degron pathway (Rusnac et al., 2018). Several common concepts have emerged from these terminal-degron recognition modes. (1) The residues in recognins that contribute to degron binding are evolutionarily conserved across different species (Figures 6A, 6D, 6G, 6J, and 6M), although they may adopt different conformations. For example, ZYG11B, GID4, and KLHDC2 use their respective ARM repeats, antiparallel β-barrel, and Kelch repeats to form a deep and narrow cavity bound to substrates (Figures 6B, 6E, and 6N). (2) The terminal residue of a substrate is the principal determinant in degron recognition. However, the adjacent sequence contexts in degrons, especially at the second position, also play an important role in substrate recognition (Figures 6C, 6F, 6I, 6L, and 6O). For instance, UBR1 favors a hydrophobic residue following the N-terminal arginine in the Arg/N-degron pathway (Choi et al., 2010), GID4 prefers a small side-chain amino acid at the second position of Pro/N-degron (Dong et al., 2018), ZYG11B favors an aromatic residue adjacent to the N-terminal glycine degron, KLHDC2 specifically recognizes a glycine at the penultimate position of Gly/C-degron (Rusnac et al., 2018), and FEM1C displays a strong preference for a positively charged residue at the antepenultimate position of Arg/C-degron (Yan et al., 2021). (3) N-recognins form an acidic pocket to capture the positively charged α-amino group of N-degrons (Figures 6B, 6E, and 6H). Additionally, the α-amino group is stabilized by three conserved hydrogen bonds (Figures 6C, 6F, and 6I). In contrast, the C-recognins (for example, KLHDC2 and FEM1C) form a basic pocket to engage the negatively charged carboxyl group of the C-terminal glycine (Figures 6K and 6N). In addition, FEM1C uses another acidic semi-open binding pocket to anchor the side chain of the C-terminal arginine (Figure 6K).

Recognition of the terminal arginine and glycine in the N- and C-degron pathways

Interestingly, UBR1 and FEM1C capture the N-terminal arginine and C-terminal arginine degron, respectively. The Arg/N-degron resides in a shallow binding groove of UBR1 through a set of hydrogen bonding interactions (Figures 6H and 6I), whereas the Arg/C-degron is installed in the semi-open binding pocket of FEM1C via a network of hydrogen bond and cation-π interactions (Figures 6K and 6L). Likewise, the glycine can be a degron at the N terminus or C terminus of proteins. The resulting Gly/N-degron can be recognized by CRL2^{ZYG11B}, whereas the Gly/C-degron can be targeted by CRL2^{KLHDC2}. Thus, the CRL2 E3 ligase contributes to protein degradation through N-degron pathway and C-degron pathway by employing different substrate receptors. Despite opposite electrostatic potential surfaces, ZYG11B (Figure 6B) and KLHDC2 (Figure 6N) use a narrow pocket to capture the terminal glycine degrons. The polypeptide backbone of the Gly/N-degron, particularly the first two residues, adopt a straight conformation, inserting into the binding pocket of ZYG11B, whereas the Gly/C-degron is embedded in the binding pocket of KLHDC2 with a highly coiled conformation. Nonetheless, they still have similar features in recognition of glycine degrons. For example, they both donate five conserved hydrogen bonds, including a water-mediated bond, to bind the free glycine (G1 or G-1), plus two hydrogen bonds interacting with the following glycine (G2 or G-2) (Figures 6C and 6O). Collectively, to achieve exquisite regulation of protein turnover by the proteasome, the E3 ligases take advantage of unique structural features to recognize respective substrates in a spatial and temporal manner for delivery into the terminal degron pathways.

Ub E3 ligases determine substrate specificity and selectivity in the UPS, making them attractive drug targets. Some E3 ligases have also been utilized for a promising therapeutic strategy referred to as PROTAC (proteolysis-targeting chimera) technology, inducing protein degradation via bifunctional degraders (Scudellari, 2019). The human genome encodes more than 600 E3 Ub ligases that target different degrons, but only a fraction of these enzymes have well-defined target degrons, limiting development of PROTAC degraders. The crystal structures of ZYG11B and ZER1 bound to Gly/N-degrons reported here not only provide the structural framework for an in-depth understanding of the molecular mechanism underlying Gly/N-degron recognition but also expand the spectrum of targeted E3 for

(D) Stability analysis of DBNDD2 degron-fused GFP with exogenous expression of WT and mutant ZYG11B proteins. The ratio of GFP/RFP was analyzed by flow cytometry.
(E) Western blot analysis of WT and mutant ZYG11B expression in DBNDD2 degron GPS reporter cell lines.
(F) Stability analysis of the SNX11 degron-fused GFP with exogenous expression of WT and mutant ZYG11B proteins.
(G) Western blot analysis of WT and mutant ZYG11B expression in SNX11degron GPS reporter cell lines.
(H) Stability analysis of SNX11 degron-fused GFP with exogenous expression of WT and mutant ZER1 proteins.
(I) Western blot analysis of WT and mutant ZER1 expression in SNX11degron GPS reporter cell lines.

Please cite this article in press as: Yan et al., Molecular basis for recognition of Gly/N-degrons by CRL2^{ZYG11B} and CRL2^{ZER1}, Molecular Cell (2021), https://doi.org/10.1016/j.molcel.2021.06.010

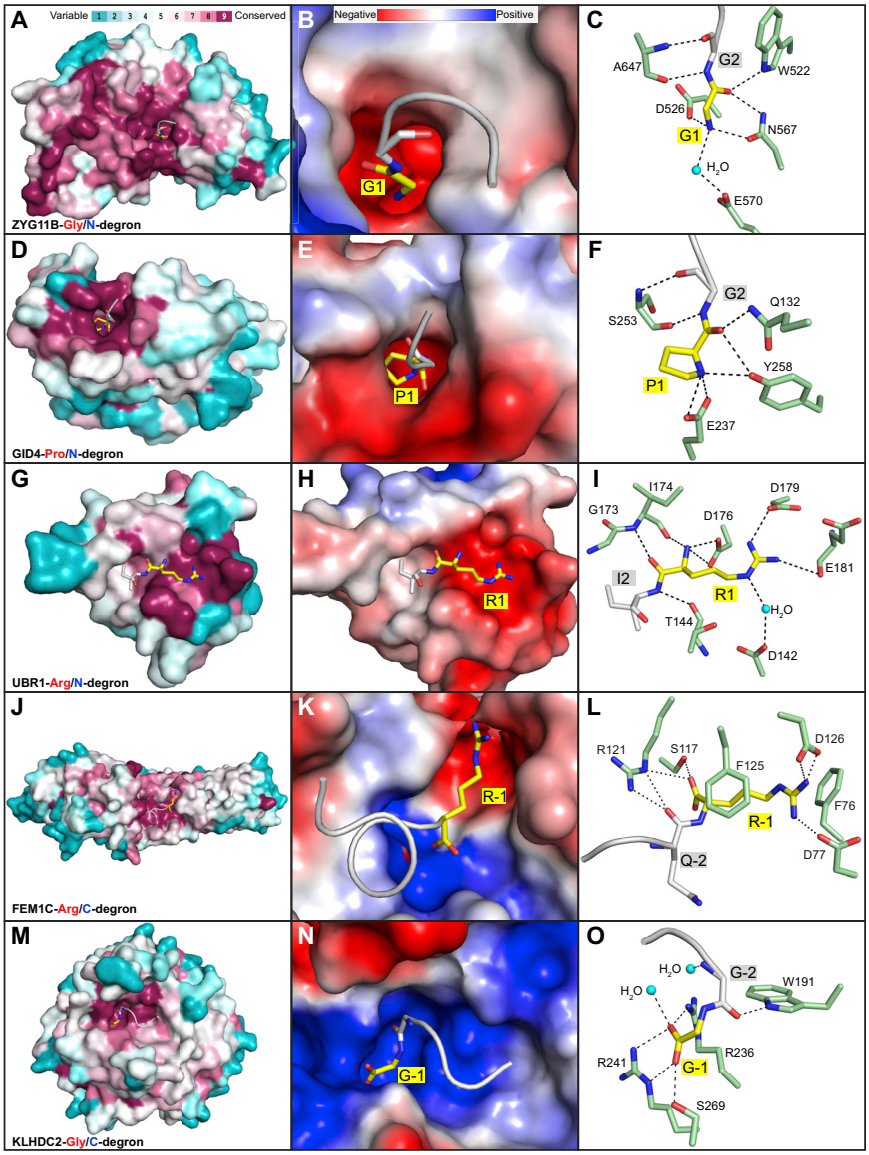


Figure 6. Structural comparison of different types of terminal-degron recognition

(A) Conservation analysis (Ashkenazy et al., 2010) of the Gly/N-degron (GGFN) binding pocket in ZYG11B. The residues of ZYG11B are colored based on conservation grades: conserved residues are indicated in maroon, and variable residues are indicated in turquoise.

(B) The electrostatic potential surface of ZYG11B bound to the Gly/N-degron (red, negative; blue, positive). The N-terminal glycine and the second glycine are shown as yellow and gray sticks, respectively.

(C) Interaction network of ZYG11B with the Gly/N-degron peptide. The N-terminal glycine is shown in yellow, the interacting residues in ZYG11B are shown as pale green sticks, and the water molecule is shown as a cyan sphere.

(D) Conservation analysis of the Pro/N-degron binding pocket in GID4 (PDB: 6CDC).

(E) The electrostatic potential surface of GID4 bound to the Pro/N-degron.

(F) Interactions of GID4 with the Pro/N-degron.

(G) Conservation analysis of the Arg/N-degron binding pocket in UBR1 (PDB: 3NIH).

(H) The electrostatic potential surface of UBR1 bound to the Arg/N-degron.

(I) Interactions of UBR1 with the Arg/N-degron.

(J) Conservation analysis of the FEM1C-mediated Arg/C-degron binding pocket (PDB: 7JYA).

(K) The electrostatic potential surface of FEM1C bound to the Arg/C-degron.

(L) Interactions of FEM1C with the Arg/C-degron.

(M) Conservation analysis of the KLHDC2-mediated Gly/C-degron binding pocket (PDB: 6DO5).

(N) The electrostatic potential surface of KLHDC2 bound to the Gly/C-degron.

(O) Interactions of KLHDC2 with the Gly/C-degron.

therapeutic applications, paving the way for developing potent small-molecule ligands of CRL2$^{ZYG11B/ZER1}$ in the future.

Limitations of the study

This study demonstrates that the ARM repeats of ZYG11B and ZER1 are required for binding of Gly/N-degrons. However, the contributions of LRR repeats in the Gly/N-degron pathway remain uncertain. A whole CRL2$^{ZYG11B/ZER1}$ complex structure is needed to answer this question. Although we did not detect substantial binding of ZER1 ARM3–ARM8 repeats to the N-terminal proline-starting peptides by ITC assay, ZER1 has been reported to mediate degradation of protein substrates (such as KCNT2) that contain N-terminal proline in a cellular context (Timms et al., 2019). Therefore, it is likely that other parts of the ZER1 protein (such as the LRR domain) or other proteins are required for N-proline substrate recognition of ZER1, an interesting topic warranting further investigation.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Bacteria
  - Human Cells
- METHOD DETAILS
  - Protein expression and purification
  - Protein crystallization
  - Data collection and structure determination

- Isothermal titration calorimetry
- GST pull-down assay
- Cell culture and viral transduction
- Global protein stability assay

- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.molcel.2021.06.010.

ACKNOWLEDGMENTS

We thank the staff at beamlines BL17U and BL19U of the Shanghai Synchrotron Radiation Facility and the Dr. Song Xiang and Dr. Yuhui Dong groups for assistance with X-ray data collection. We thank Dr. Xiaobing Shi for critically reading the manuscript. This work was supported by National Natural Science Foundation of China grants 31900865 (to C.D.), 32071193 (to C.D.), 81872350 (to Z.M.), 81974431 (to W.M.), and 81874039 (to X.Y.).

AUTHOR CONTRIBUTIONS

C.D. conceptualized the project. Z.M. and W.M. provided critical feedback and analyzed the data. X.Y. and Y.L. cloned all constructs and performed protein expression, purification, and crystallization with help from G.S. and Q.F. C.D. and X.Y. determined the crystal structures. X.Y. and Y.L. conducted the ITC assay. X.Y., Y.L., G.W., Z.Z., and Y.Z. performed the GPS assay. C.D. wrote the manuscript with critical input from all authors.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: February 9, 2021
Revised: May 8, 2021
Accepted: June 9, 2021
Published: July 1, 2021

REFERENCES

An, J.Y., Seo, J.W., Tasaki, T., Lee, M.J., Varshavsky, A., and Kwon, Y.T. (2006). Impaired neurogenesis and cardiovascular development in mice lacking the E3 ubiquitin ligases UBR1 and UBR2 of the N-end rule pathway. Proc. Natl. Acad. Sci. USA *103*, 6212–6217.

Ashkenazy, H., Erez, E., Martz, E., Pupko, T., and Ben-Tal, N. (2010). ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. Nucleic Acids Res. *38*, W529-33.

Bachmair, A., and Varshavsky, A. (1989). The degradation signal in a short-lived protein. Cell *56*, 1019–1032.

Bachmair, A., Finley, D., and Varshavsky, A. (1986). In vivo half-life of a protein is a function of its amino-terminal residue. Science *234*, 179–186.

Balchin, D., Hayer-Hartl, M., and Hartl, F.U. (2016). In vivo aspects of protein folding and quality control. Science *353*, aac4354.

Chatr-Aryamontri, A., van der Sloot, A., and Tyers, M. (2018). At Long Last, a C-Terminal Bookend for the Ubiquitin Code. Mol. Cell *70*, 568–571.

Chen, S.J., Wu, X., Wadas, B., Oh, J.H., and Varshavsky, A. (2017). An N-end rule pathway that recognizes proline and destroys gluconeogenic enzymes. Science *355*, eaal3655.

Chen, X., Liao, S., Makaros, Y., Guo, Q., Zhu, Z., Krizelman, R., Dahan, K., Tu, X., Yao, X., Koren, I., et al. (2021). Molecular basis for arginine C-terminal degron recognition by Cul2(FEM1) E3 ligase. Nat. Chem. Biol. *17*, 254–262.

Choi, W.S., Jeong, B.C., Joo, Y.J., Lee, M.R., Kim, J., Eck, M.J., and Song, H.K. (2010). Structural basis for the recognition of N-end rule substrates by the UBR box of ubiquitin ligases. Nat. Struct. Mol. Biol. *17*, 1175–1181.

Dong, C., Zhang, H., Li, L., Tempel, W., Loppnau, P., and Min, J. (2018). Molecular basis of GID4-mediated recognition of degrons for the Pro/N-end rule pathway. Nat. Chem. Biol. *14*, 466–473.

Dong, C., Chen, S.J., Melnykov, A., Weirich, S., Sun, K., Jeltsch, A., Varshavsky, A., and Min, J. (2020). Recognition of nonproline N-terminal residues by the Pro/N-degron pathway. Proc. Natl. Acad. Sci. USA *117*, 14158–14167.

Eldeeb, M., Esmaili, M., and Fahlman, R. (2019). Degradation of proteins with N-terminal glycine. Nat. Struct. Mol. Biol. *26*, 761–763.

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

Finley, D. (2009). Recognition and processing of ubiquitin-protein conjugates by the proteasome. Annu. Rev. Biochem. *78*, 477–513.

Gonda, D.K., Bachmair, A., Wünning, I., Tobias, J.W., Lane, W.S., and Varshavsky, A. (1989). Universality and structure of the N-end rule. J. Biol. Chem. *264*, 16700–16712.

Graciet, E., Walter, F., Ó'Maoiléidigh, D.S., Pollmann, S., Meyerowitz, E.M., Varshavsky, A., and Wellmer, F. (2009). The N-end rule pathway controls multiple functions during Arabidopsis shoot and leaf development. Proc. Natl. Acad. Sci. USA *106*, 13618–13623.

Holdsworth, M.J., Vicente, J., Sharma, G., Abbas, M., and Zubrycka, A. (2020). The plant N-degron pathways of ubiquitin-mediated proteolysis. J. Integr. Plant Biol. *62*, 70–89.

Hwang, C.S., Shemorry, A., and Varshavsky, A. (2009). Two proteolytic pathways regulate DNA repair by cotargeting the Mgt1 alkylguanine transferase. Proc. Natl. Acad. Sci. USA *106*, 2142–2147.

Hwang, C.S., Shemorry, A., and Varshavsky, A. (2010). N-terminal acetylation of cellular proteins creates specific degradation signals. Science *327*, 973–977.

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

Kim, J.M., Seok, O.H., Ju, S., Heo, J.E., Yeom, J., Kim, D.S., Yoo, J.Y., Varshavsky, A., Lee, C., and Hwang, C.S. (2018). Formyl-methionine as an N-degron of a eukaryotic N-end rule pathway. Science *362*, eaat0174.

Kim, L., Kwon, D.H., Heo, J., Park, M.R., and Song, H.K. (2020). Use of the LC3B-fusion technique for biochemical and structural studies of proteins involved in the N-degron pathway. J. Biol. Chem. *295*, 2590–2600.

Koren, I., Timms, R.T., Kula, T., Xu, Q., Li, M.Z., and Elledge, S.J. (2018). The Eukaryotic Proteome Is Shaped by E3 Ubiquitin Ligases Targeting C-Terminal Degrons. Cell *173*, 1622–1635.e14.

Kwon, Y.T., and Ciechanover, A. (2017). The Ubiquitin Code in the Ubiquitin-Proteasome System and Autophagy. Trends Biochem. Sci. *42*, 873–886.

Lee, K.E., Heo, J.E., Kim, J.M., and Hwang, C.S. (2016). N-Terminal Acetylation-Targeted N-End Rule Proteolytic System: The Ac/N-End Rule Pathway. Mol. Cells *39*, 169–178.

Liebschner, D., Afonine, P.V., Baker, M.L., Bunkóczi, G., Chen, V.B., Croll, T.I., Hintze, B., Hung, L.W., Jain, S., McCoy, A.J., et al. (2019). Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. Acta Crystallogr. D Struct. Biol. *75*, 861–877.

Lin, H.C., Yeh, C.W., Chen, Y.F., Lee, T.T., Hsieh, P.Y., Rusnac, D.V., Lin, S.Y., Elledge, S.J., Zheng, N., and Yen, H.S. (2018). C-Terminal End-Directed Protein Elimination by CRL2 Ubiquitin Ligases. Mol. Cell *70*, 602–613.e3.

Liu, Y., Liu, C., Dong, W., and Li, W. (2016). Physiological functions and clinical implications of the N-end rule pathway. Front. Med. *10*, 258–270.

Lucas, X., and Ciulli, A. (2017). Recognition of substrate degrons by E3 ubiquitin ligases and modulation by small-molecule mimicry strategies. Curr. Opin. Struct. Biol. *44*, 101–110.

Matta-Camacho, E., Kozlov, G., Li, F.F., and Gehring, K. (2010). Structural basis of substrate recognition and specificity in the N-end rule pathway. Nat. Struct. Mol. Biol. *17*, 1182–1187.

Melby, K., Fasmer, O.B., Henriksen, T.E., Gråwe, R.W., Aamo, T.O., and Spigset, O. (2020). Actigraphy assessment of motor activity and sleep in

patients with alcohol withdrawal syndrome and the effects of intranasal oxytocin. PLoS ONE *15*, e0228700.

Pickart, C.M. (2001). Mechanisms underlying ubiquitination. Annu. Rev. Biochem. *70*, 503–533.

Robinson, K.S., Teo, D.E.T., Tan, K.S., Toh, G.A., Ong, H.H., Lim, C.K., Lay, K., Au, B.V., Lew, T.S., Chu, J.J.H., et al. (2020). Enteroviral 3C protease activates the human NLRP1 inflammasome in airway epithelia. Science *370*, eaay2002.

Rusnac, D.V., Lin, H.C., Canzani, D., Tien, K.X., Hinds, T.R., Tsue, A.F., Bush, M.F., Yen, H.S., and Zheng, N. (2018). Recognition of the Diglycine C-End Degron by CRL2^{KLHDC2} Ubiquitin Ligase. Mol. Cell *72*, 813–822.e4.

Scudellari, M. (2019). Protein-slaying drugs could be the next blockbuster therapies. Nature *567*, 298–300.

Sontag, E.M., Samant, R.S., and Frydman, J. (2017). Mechanisms and Functions of Spatial Protein Quality Control. Annu. Rev. Biochem. *86*, 97–122.

Tasaki, T., Sriram, S.M., Park, K.S., and Kwon, Y.T. (2012). The N-end rule pathway. Annu. Rev. Biochem. *81*, 261–289.

Timms, R.T., and Koren, I. (2020). Tying up loose ends: the N-degron and C-degron pathways of protein degradation. Biochem. Soc. Trans. *48*, 1557–1567.

Timms, R.T., Zhang, Z., Rhee, D.Y., Harper, J.W., Koren, I., and Elledge, S.J. (2019). A glycine-specific N-degron pathway mediates the quality control of protein *N*-myristoylation. Science *365*, eaaw4912.

Tobias, J.W., Shrader, T.E., Rocap, G., and Varshavsky, A. (1991). The N-end rule in bacteria. Science *254*, 1374–1377.

Varshavsky, A. (2011). The N-end rule pathway and regulation by proteolysis. Protein Sci. *20*, 1298–1345.

Varshavsky, A. (2019). N-degron and C-degron pathways of protein degradation. Proc. Natl. Acad. Sci. USA *116*, 358–366.

Wright, M.H., Heal, W.P., Mann, D.J., and Tate, E.W. (2010). Protein myristoylation in health and disease. J. Chem. Biol. *3*, 19–35.

Yan, X., Wang, X., Li, Y., Zhou, M., Li, Y., Song, L., Mi, W., Min, J., and Dong, C. (2021). Molecular basis for ubiquitin ligase CRL2(FEM1C)-mediated recognition of C-degron. Nat. Chem. Biol. *17*, 254–262.

Please cite this article in press as: Yan et al., Molecular basis for recognition of Gly/N-degrons by CRL2^{ZYG11B} and CRL2^{ZER1}, Molecular Cell (2021), https://doi.org/10.1016/j.molcel.2021.06.010

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
HRP-conjugated Anti-Flag Antibody	Proteintech	Cat# 20543-1-AP; RRID:AB_11232216
HRP-conjugated Anti-Tubulin Antibody	Beyotime	AF0001
Bacterial and virus strains		
E.coli BL21 (DE3)	Transgene	CD601-02
E.coli T1	Transgene	CD501-02
Chemicals, peptides, and recombinant proteins		
pEASY-Basic Seamless Cloning and Assembly Kit	Transgene	CU201-02
DMEM	BioInd	06-1055-57-1ACS
Fetal bovine serum (FBS)	BioInd	04-001-1ACS
Penicillin/Streptomycin	Hyclone	SV30010
HRP Substrate	Millipore	WBKLS0100
GIMAP5, ZNF701 peptides	Apeptide	N/A
Chelating Sepharose Fast Flow	GE Healthcare	17-0575-02
Glutathione Sepharose 4B	GE Healthcare	17-0756-05
HiTrap Q HP, 5 mL	GE Healthcare	17-1154-01
Superdex 200 Increase 10/300 GL	GE Healthcare	28-9909-44
TEV protease	In house	N/A
Deposited data		
ZYG11B-GSTE	This paper	PDB: 7EP0
ZYG11B-GFLH	This paper	PDB: 7EP1
ZYG11B-GGFN	This paper	PDB: 7EP2
ZER1-GAGN	This paper	PDB: 7EP3
ZER1-GFLH	This paper	PDB: 7EP4
ZER1-GKLH	This paper	PDB: 7EP5
Experimental models: Cell lines		
HEK293T cell line	ATCC	CRL-3216
Recombinant DNA		
pET28-MKH8SUMO- ZYG11B	This paper	N/A
pET28-MKH8SUMO-ZER1	This paper	N/A
pET28-MHL-peptide-GST	This paper	N/A
pCDH-Ub-peptide-GFP-P2A-RFP	This paper	N/A
pNIC-CH-ZER1	This paper	N/A
Software and algorithms		
XDS	Kabsch, 2010	https://xds.mr.mpg.de/html_doc/xds_prepare.html
Phenix	Liebschner et al., 2019	https://phenix-online.org/
Pymol	Pymol	https://www.pymol.org/2/
COOT	Emsley et al., 2010	https://www2.mrc-lmb.cam.ac.uk/personal/pemsley/coot/
FlowJo	FLOWJO	N/A
MicroCal PEAQ-ITC Analysis Software	Malvern Panalytical	https://www.malvernpanalytical.com/en
Other		
ACEA NovoCyte	ACEA Biosciences, Inc.	N/A

Please cite this article in press as: Yan et al., Molecular basis for recognition of Gly/N-degrons by CRL2^{ZYG11B} and CRL2^{ZER1}, Molecular Cell (2021), https://doi.org/10.1016/j.molcel.2021.06.010

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Cheng Dong (dongcheng@tmu.edu.cn).

Materials availability

All unique/stable reagents generated in this study are available from the Lead Contact without restriction. Plasmids and strains are available from the authors upon request.

Data and code availability

The atomic coordinates included in this study have been deposited in the Protein Data Bank with accession codes 7EP0, 7EP1, 7EP2, 7EP3, 7EP4 and 7EP5.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Bacteria

E. coli T1 was used for DNA amplification and plasmid construction and *E. coli* BL21 (DE3) was used for wild-type and mutant protein expression.

Human Cells

Human HEK293T cells (ATCC CRL-3216) were maintained in Dulbecco's Modified Eagle's Medium (DMEM) (BioInd) supplemented with 10% fetal bovine serum (FBS, BioInd) and penicillin/streptomycin (HyClone) at 37°C, 5% CO_2.

METHOD DETAILS

Protein expression and purification

The genes encoding human ZYG11B (residues 445-728), ZYG11B (residues 483-728) and ZYG11B (residues 485-728) were amplified from cDNA with coding sequences of Gly-Phe-Asn (GFN), Ser-Thr-Glu (STE) and Phe-Leu-His-Val-Gly (FLHVG) appended to the 5′ ends, respectively. The fusion fragments were cloned into the pET28-MKH8SUMO vector (Addgene Plasmid #79526) using seamless cloning. The recombinant plasmid contains an 8 x His-SUMO tag followed by a TEV (Tobacco Etch Virus) cleavage site (LYFQ↓G, leaving an N-terminal glycine as the exposed N-degron). The human ZER1 fragments covering residues 520-766 were cloned into pET28-MKH8SUMO vector in the same way with an additional GAGN, GFLHVG or GKLHVG sequence at the N-termini. These recombinant plasmids were transformed into *E. coli* BL21 (DE3) cells with Kanamycin selection. The cells were cultured in Luria-Bertani (LB) medium at 37°C until the value of $OD_{600\ nm}$ reached 0.6 to 0.8, and then the temperature was reduced to 18°C. The fusion protein expression was induced by the addition of 0.2 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). After 16 hours, the bacteria were harvested by centrifugation and the pellets were resuspended in lysis buffer composed of 20 mM Tris-HCl pH 7.5, 400 mM NaCl and 2 mM beta-mercaptoethanol. The cells were lysed by sonication on ice and the cellular debris was removed by centrifugation at 14000 rpm for 40 minutes. The supernatant was then incubated with Ni-NTA beads for 1 hour at 4°C. Nonspecifically bound proteins were washed out by washing buffer (20 mM Tris-HCl pH 7.5, 400 mM NaCl and 25 mM imidazole pH 7.5), and the fusion proteins were eluted by elution buffer (20 mM Tris-HCl pH 7.5, 400 mM NaCl and 500 mM imidazole pH 7.5). The SUMO tag was cleaved by home-made TEV protease with a molar ration of 1:30 in dialysis buffer (20 mM Tris-HCl pH 7.5 and 300 mM NaCl) at 4°C overnight. The samples were reloaded onto the Ni-NTA column to remove the digested SUMO tag and TEV protease. Further purification was performed by anion exchange chromatography (HiTrap Q HP column, GE healthcare) followed by size exclusion chromatography (Superdex 200 Increase 10/300 GL, GE healthcare) pre-equilibrated with gel-filtration buffer (20mM Tris-HCl pH 7.5, 150mM NaCl and 1mM DTT). The purified proteins were concentrated and stored at −80°C for later use. The Se-methionine-substituted ZYG11B (residues 483-728) protein was expressed in M9 medium according to the manufacturer's instructions, and purified using the same protocol as the wild-type protein.

Protein crystallization

Crystallization trials were performed using the sitting-drop vapor diffusion method at 18°C by mixing 1 μL of protein and 1 μL of reservoir solution. The crystallization conditions are as follows: 0.2 M Ammonium Sulfate, 0.1 M HEPES pH 7.5 and 25% (wt/vol) Polyethylene glycol 3,350 for GGFN-fused ZYG11B (residues 445-728); 0.1 M MES pH 6.5, 14.4% (wt/vol) Polyethylene glycol 20,000, 0.033% (wt/vol) Anthrone, 0.033% (wt/vol) Congo Red, 0.033% (wt/vol) N-(2-Acetamido)-2-aminoethanesulfonic acid and 2 mM HEPES sodium pH 6.8 for Se-Met GSTE-fused ZYG11B (residues 483-728); 0.1 M MES pH 6.5 and 14.4% (wt/vol) Polyethylene glycol 20,000 for GFLHVG-fused ZYG11B (residues 485-728); 0.2 M Ammonium formate and 20% (wt/vol) Polyethylene glycol 3,350 for GFLHVG-fused ZER1 (residues 520-766); 0.22 M Ammonium formate and 18% (wt/vol) Polyethylene glycol 3,350 for GKLHVG-fused ZER1 (residues 520-766); 0.2 M Sodium formate, 23% (wt/vol) Polyethylene glycol 3,350, 0.033% (wt/vol) Anthrone, 0.033% (wt/vol)

Please cite this article in press as: Yan et al., Molecular basis for recognition of Gly/N-degrons by CRL2^{ZYG11B} and CRL2^{ZER1}, Molecular Cell (2021), https://doi.org/10.1016/j.molcel.2021.06.010

Congo Red, 0.033% (wt/vol) N-(2-Acetamido)-2-aminoethanesulfonic acid and 2 mM HEPES sodium pH 6.8 for GAGN-fused ZER1 (residues 520-766). The crystals were cryoprotected with the respective reservoir solution supplemented with 20% (vol/vol) glycerol and flash-frozen in liquid nitrogen.

Data collection and structure determination

Diffraction data were collected on the beamline BL17U or BL19U at Shanghai Synchrotron Radiation Facility (SSRF) and processed with XDS (Kabsch, 2010). The structure of selenomethionine derivative ZYG11B (residues 483-728) was solved by the single wavelength anomalous diffraction (SAD) method. Initial phase and model were obtained using AutoSol and AutoBuild implemented in the PHENIX package (Liebschner et al., 2019). The resulting model was improved by manual building in Coot (Emsley et al., 2010). The other structures were solved by molecular replacement with Phaser using the ZYG11B (residues 483-728) structure as a search template. Structural figures were prepared with PyMOL (https://www.pymol.org/2/).

Isothermal titration calorimetry

Isothermal titration calorimetry (ITC) measurements were performed at 16°C using a MicroCal PEAQ-ITC instrument (Malvern Panalytical). The sequence of ZYG11B (residues 490-728) was cloned into the pET28-MKH8SUMO vector carrying a serine instead of glycine at the TEV cleavage site (LYFQ↓S). As a result, the purified ZYG11B (residues 490-728) bearing an N-terminal SELF sequence. On the other hand, the fragment of ZER1 (residues 469-766) was cloned into the pNIC-CH (Addgene Plasmid #26117) vector with an N-terminal MEEL sequence was used to ITC assays. All proteins and peptides were prepared in an ITC buffer consisting of 20 mM Tris-HCl pH 7.5 and 150 mM NaCl. The concentration of proteins and peptides ranged from 50-70 μM and 0.5-1.4 mM, respectively. The peptide was titrated into the protein for each experiment with 15 injections of 1.5 μL each with the exception of the first injection of 0.5 μL spaced by 150 s with a reference power of 10 μcal sec^{-1}. The titration data was processed using MicroCal PEAQ-ITC Analysis Software version 1.30. ITC assays were repeated at least twice independently with similar results.

GST pull-down assay

The coding sequences of GFQRGK or FLHVGQ were appended to the N terminus of GST, and constructed into pET28-MHL (Addgene Plasmid #26096) vector containing 6 x His tag followed by a TEV cleavage site. The fusion proteins were purified by Ni-NTA agarose column and were digested by TEV protease. The resulting proteins bear an N-terminal GGFQRGK or GFLHVGQ polypeptide overhang of GST. The SUMO-tagged ZYG11B covering residues 572-728, 530-728, 490-728, 399-728 and 266-744 were purified as the target proteins. Approximately 150 μg of GST-fused Gly/N-degrons were immobilized on 10-20 μL GST beads followed by incubation with equal amounts of the target proteins for one hour at 4°C. After washing three times, the pulled-down samples were eluted with 50 mM glutathione (pH 7.5). The products were analyzed by SDS-PAGE followed by Coomassie Blue staining.

Cell culture and viral transduction

Human HEK293T cells (ATCC CRL-3216) were maintained in Dulbecco's Modified Eagle's Medium (DMEM) (BioInd) supplemented with 10% fetal bovine serum (BioInd) and penicillin/streptomycin (HyClone). For lentiviral production, HEK293T cells were co-transfected with either pCDH-puro or pCDH-Ub-GPS (containing an Ub-peptide-GFP-P2A-RFP cascade) transfer vector together with the third-generation packaging plasmids pMDLg/pRRE, pMD2.VSVG, and pRSV-REV. After 48 h, the lentivirus was filtered by a 0.45 μm filter, and applied to HEK293T cells in the presence of 8 μg mL^{-1} polybrene. The cells were washed and recovered 24 h prior to further procedure.

Global protein stability assay

The oligonucleotide encoding the 19-mer peptide from DBNDD2 or SNX11 degron was cloned into the pCDH-Ub-GPS vector using the Seamless Cloning method. The reconstructed vector encoded an ubiquitin-fused peptide followed by a short linker (ATSALGT) and GFP. The Ub-peptide-GFP fusion protein and RFP protein were expressed from the same transcript and separated by P2A self-cleaving. The N-terminal glycine degron of the peptides would be exposed through endogenous deubiquitinating enzymes (DUB). The stability of peptide-GFP was determined by measuring the cellular GFP/RFP ratio through flow cytometry using RFP as an internal control. Ub-peptide-GPS vectors were packaged into lentivirus that were used to transduce HEK293T cells to obtain GPS-reporter cells by blasticidine (5 μg mL^{-1}) selection. Human ZYG11B or ZER1 cDNA was cloned into pCDH-puro with an N-terminal Flag tag. Wild-type and mutant ZYG11B/ZER1 were overexpressed in GPS-reporter cells by lentiviral transduction after selection with puromycine (1 μg mL^{-1}) for 3 days. Cells were analyzed by flow cytometry using an ACEA NovoCyte instrument (ACEA Biosciences, Inc) and the data was analyzed using FlowJo. Expression of wild-type and mutant ZYG11B/ZER1 in GPS reporter cells were examined by western blot with anti-Flag (1:4,000; Proteintech, 20543-1-AP) and anti-alpha tubulin (1: 3,000; Beyotime, AF001) was used as an internal control.

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistics of X-ray diffraction data collection and structure refinement are summarized in Table 1 and Table 2. Softwares for crystal structure determination and analysis are presented in the Key resources table. Data from ITC experiments (Figures 1, 3, 4, and 5) were analyzed with MicroCal PEAQ-ITC Analysis Software using the one-site fitting model.