HHS Public Access

Author manuscript

Cell Rep. Author manuscript; available in PMC 2023 October 23.

Published in final edited form as:

Cell Rep. 2023 September 26; 42(9): 113100. doi:10.1016/j.celrep.2023.113100.

Mechanism and evolutionary origins of alanine-tail C-degron recognition by E3 ligases Pirh2 and CRL2-KLHDC10

Pratik Rajendra Patil[1], **A. Maxwell Burroughs**[2], **Mohit Misra**[3,4], **Federico Cerullo**[1], **Carlos Costas-Insua**[1], **Hao-Chih Hung**[1], **Ivan Dikic**[3,4], **L. Aravind**[2], **Claudio A.P. Joazeiro**[1,5,6,*]

[1]Center for Molecular Biology of Heidelberg University (ZMBH), DKFZ-ZMBH-Alliance, 69120 Heidelberg, Germany

[2]National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, MD 20894, USA

[3]Institute of Biochemistry II, Goethe University Faculty of Medicine, Theodor-Stern-Kai 7, 60590 Frankfurt am Main, Germany

[4]Buchmann Institute for Molecular Life Sciences, Goethe University Frankfurt, Max-von-Laue-Strasse 15, 60438 Frankfurt am Main, Germany

[5]Department of Molecular Medicine, UF Scripps Biomedical Research, Jupiter, FL 33458, USA

[6]Lead contact

SUMMARY

In ribosome-associated quality control (RQC), nascent polypeptides produced by interrupted translation are modified with C-terminal polyalanine tails ("Ala-tails") that function outside ribosomes to induce ubiquitylation by E3 ligases Pirh2 (p53-induced RING-H2 domain-containing) or CRL2 (Cullin-2 RING ligase2)-KLHDC10. Here, we investigate the molecular basis of Ala-tail function using biochemical and *in silico* approaches. We show that Pirh2 and KLHDC10 directly bind to Ala-tails and that structural predictions identify candidate Ala-tail-binding sites, which we experimentally validate. The degron-binding pockets and specific pocket residues implicated in Ala-tail recognition are conserved among Pirh2 and KLHDC10 homologs, suggesting that an important function of these ligases across eukaryotes is in targeting Ala-tailed substrates. Moreover, we establish that the two Ala-tail-binding pockets have convergently evolved, either from an ancient module of bacterial provenance (Pirh2) or via tinkering of a widespread C-degron-recognition element (KLHDC10). These results shed light on the recognition of a simple degron sequence and the evolution of Ala-tail proteolytic signaling.

*Correspondence: c.joazeiro@zmbh.uni-heidelberg.de.

AUTHOR CONTRIBUTIONS

Investigation, P.R.P., A.M.B., F.C., C.C.-I., and H.-C.H.; formal analysis, P.R.P., A.M.B., M.M., and L.A.; methodology, P.R.P. and M.M.; writing – original draft, P.R.P. and C.A.P.J.; writing – review & editing, P.R.P., A.M.B., M.M., F.C., C.C.-I., H.-C.H., I.D., L.A., and C.A.P.J.; supervision, I.D., L.A., and C.A.P.J.; conceptualization, C.A.P.J.; resource, C.A.P.J.

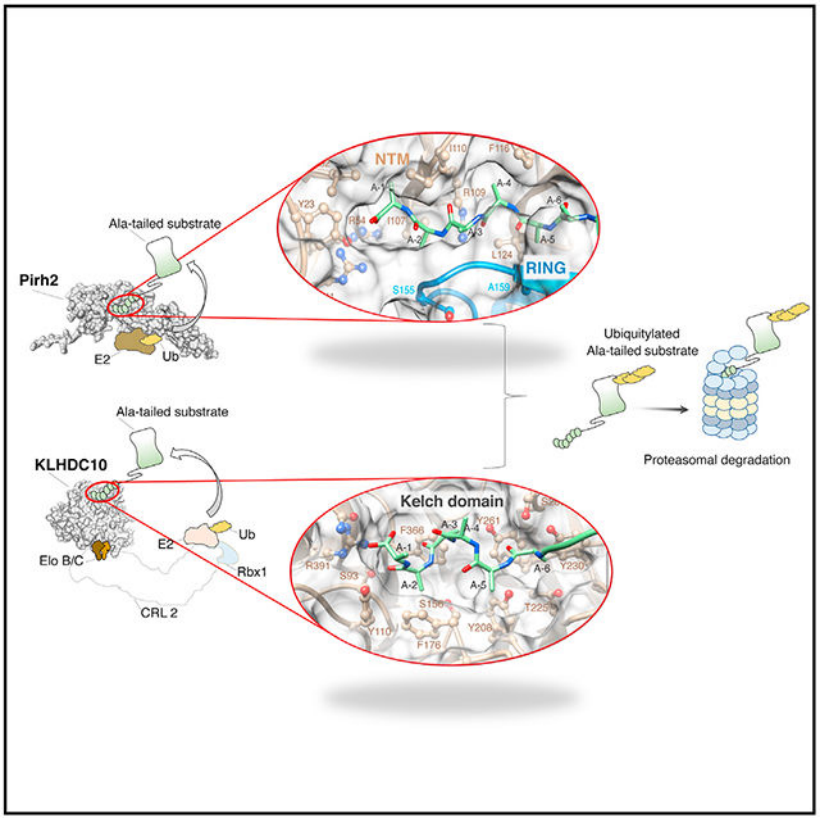
SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.celrep.2023.113100.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Graphical abstract



In brief

Ribosome-associated quality control (RQC) modifies incomplete polypeptides produced by ribosomal stalling with C-terminal alanine-tail degrons. In humans, alanine-tailed (Ala-tailed) proteins are ubiquitylated by E3 ligases Pirh2 and CRL2-KLHDC10, but how Ala-tails are sensed remained unclear. Patil et al. provide a structural basis for Ala-tail recognition and insights into its evolution.

INTRODUCTION

Protein quality control is essential for maintaining proteome homeostasis and cell viability.[1] This includes the proteasomal degradation of aberrant proteins mediated by E3 ubiquitin ligases.[2,3] E3 ligases–numbering over 650 in humans[4]–confer specificity in ubiquitylation, often by targeting substrates through recognition of linear motifs, or degrons, present internally, in the N terminus, or in the C terminus (C-end).[5] C-end degrons are a newly discovered class whose sequence diversity and cognate E3 ligases are beginning to be elucidated. They can be naturally present in some proteins, be exposed by proteolytic cleavage, be generated by damage, or be appended as a modification, as in ribosome-associated quality control (RQC).[6–20]

RQC eliminates incomplete polypeptides produced when ribosomes stall during translation.[21] Dedicated factors sense and dissociate stalled ribosomes, generating large (60S) ribosomal subunits that are still blocked by nascent polypeptide-tRNA conjugates. In mammals, NEMF (nuclear export mediator factor) recognizes this obstruction and recruits Ala-charged tRNA to extend the nascent polypeptides with a polyalanine tract ("Ala-tail").[8,22,23] Ala-tailing mediates protein degradation through at least two routes. In canonical RQC (RQC-L), C-terminal elongation exposes Lys residues buried in the ribosomal exit tunnel, facilitating on-ribosome ubiquitylation by the E3 ligase listerin/Ltn1[8,24]; in the RQC-C pathway, Ala-tails themselves serve as C-end degrons, mediating interaction outside the ribosome with either Pirh2, an E3 ligase, or KLHDC10, the substrate-recognition subunit of a Cullin-2 RING ligase (CRL2).[8]

Pirh2 (p53-induced RING-H2 domain-containing E3 ligase; also known as Rchy1) targets p53 for proteolysis,[25,26] along with tumor-suppressor proteins p63 and p73 and the oncoprotein c-Myc.[27–29] How Pirh2 recognizes these substrates remains unclear. Similarly, the biological roles of KLHDC10 remain poorly understood.[30] Mechanistically, KLHDC10 also targets C-end degrons ending in TrpGly, ProGly, or AlaGly.[10] Its paralog, KLHDC2, recognizes degrons ending in di-Gly, and KLHDC3 recognizes degrons ending in ArgGly, LysGly, or GlnGly.[10] Crystal structures of KLHDC2-bound C-end degron peptides provided first insights into recognition mechanisms.[12]

Here, we study the Ala-tail degron mechanism. Through biochemical and *in silico* structural analyses, we identify and validate unique Ala-tail-binding sites in Pirh2 and KLHDC10. The degron-binding sites and specific recognition residues are highly conserved among their homologs, widely distributed across eukaryotes. Pirh2 likely existed in the last eukaryotic common ancestor (LECA), suggesting the presence of an E3 ligase targeting Ala-tail-containing substrates since early eukaryotic origins. Whereas KLHDC10 evolved its substrate-binding pocket via subtle modifications of a C-degron-recognition site shared with paralogs KLHDC2/3, the Pirh2 pocket has a more dramatic history. It originated from an ancient bacterial module, subsequently reused via modification and augmentation to create a unique binding interface.

RESULTS

Reconstitution of Pirh2 and KLHDC10 binding to the Ala-tail degron *in vitro*

Pirh2 consists of an N-terminal module (NTM; residues 1–137), the RING domain (138–189), and a C-terminal domain (CTD; 190–261) (Figures 1A and S1A). Their individual structures have been solved (PDB: 2K2C, 2JRJ, and 2K2D, respectively).[26] Our analysis shows that the NTM is a composite, with three types of Zn-chelating domains (Figures S1B–S1E): (1) the N-terminal-most binuclear Zn-binding CHY Treble-clef domain, distantly related to the ZZ and RING domains.[31] (2) In the center, a classic Zn ribbon (ZnR) domain[32] coordinates a single Zn ion. (3) The C-terminal part comprises three repeats of a Zn-binding domain, evolved from a standalone C-terminal fragment of a Treble-clef domain.[31] The entire NTM binds 6 Zn ions.[26] The CTD has a single ZnR domain related to the NTM central domain and chelates one Zn ion. The NTM and CTD together recognize p53,[26] whereas the E3-catalytic RING domain binds ubiquitin-carrying E2 conjugases.[4]

To identify Pirh2's Ala-tail-recognition site, we expressed and purified its GST-tagged fragments from *E. coli* and evaluated binding to recombinant GFP-Ala_6, a fusion protein degraded *in vivo* in a Pirh2- and KLHDC10-dependent manner.[8] The Pirh2 NTM ($Pirh2_{1–137}$) sufficed to pull down GFP-Ala_6 (Figure 1B), although a construct containing both NTM and RING domains ($Pirh2_{1–195}$) was more efficient than the NTM alone. In contrast, a construct containing the RING domain and the CTD ($Pirh2_{120–261}$) did not have detectable GFP-Ala_6 binding activity.

To validate and quantify interactions between the Ala-tail and Pirh2, we developed an amplified luminescence proximity homogeneous assay (AlphaScreen). This assay relies on the proximity-dependent transfer of singlet oxygen from streptavidin-conjugated donor beads to anti-GST antibody-conjugated acceptor beads, generating fluorescence. For validation, we used GST-$Pirh2_{1–195}$ to bind anti-GST beads and designed a biotinylated peptide for streptavidin bead binding. This peptide, MDELYKAAAAAA ("Ala-tail peptide"), comprises six Ala residues preceded by the last six amino acids of GFP, whose polar and charged side chains enhance solubility. In the presence of both interacting partners, fluorescence emissions could be observed, which decreased upon titration of a non-biotinylated Ala-tail peptide competitor (Figure 1C). The dose-response curves provided a half-maximal inhibitory concentration (IC_{50}) value of ~90–120 nM for the interaction (Figure 1C). Finally, and consistent with the pull-down experiments, deleting the RING domain reduced Ala-tail binding ~10-fold.

We also examined Ala-tail binding to KLHDC10. This protein has a β-propeller domain with six Kelch repeats (residues 86–401; Figures 1D and S2), flanked by an N-terminal, likely unstructured 85-residue region, and a C-terminal trihelical domain comprised of a BC-box and a Cul2-box (residues 402–414 and 415–442, respectively).[33] In KLHDC2, the Kelch domain recognizes degrons,[12] whereas the BC-box binds to the Cullin-2 adaptors elongins B/C, and the Cul2-box provides additional Cullin-2 interaction, which in turn recruits the RING domain protein Rbx1.[34,35] Similar interactions likely apply for KLHDC10,[33] although we could not formally confirm that the Kelch domain suffices for C-degron binding due to the instability of proteins lacking the BC-box and/or the Cul2-box expressed in *E. coli*. Nonetheless, we confirmed that recombinant KLHDC10, unlike KLHDC2, specifically pulled down GFP-Ala_6 (Figure 1E), and we measured an IC_{50} of ~5–10 nM for the Ala_6 peptide-KLHDC10 interaction (Figure 1F).

Evaluation of features of the Ala-tail degron implicated in E3-ligase binding

The above results strongly indicate that Pirh2 and KLHDC10 bind directly and tightly to Ala-tails. Next, to identify degron features conferring specificity toward the E3 ligases, we performed titrations of Ala-tail-variant, unlabeled peptides and determined their IC_{50} values for GST-Pirh2 and -KLHDC10.

Deletion of KLHDC10's unstructured N terminus, except for residues 40–51, aided protein expression and solubility ($KLHDC10_{β1\text{-}88–442}$; Figure S2). Each Kelch repeat folds as a β-propeller blade, comprised of four β-strands denoted as A–D from inner to outer (Figure S2); residues 40–51 were retained in $KLHDC10_{β1\text{-}88–442}$ because they were predicted

by AlphaFold to contribute β-strand D of Kelch repeat 6. KLHDC10 full length and KLHDC10$_{β\text{-}188–442}$ had similar IC_{50} values for Ala-tail peptide (Figure 1F).

In one set of experiments, we explored our previous observation that polyAla sequences must be positioned at a protein's very C terminus to function as degrons.[8] For example, extending an Ala5 sequence with a single Thr caused complete loss of destabilizing activity *in vivo*, suggesting that both the degron's C-terminal Ala and its alpha-carboxyl group are sensed.[8] We thus assessed the impact of replacing the Ala-tail peptide's terminal alpha-carboxyl with an amide group, using AlphaScreen. Strikingly, this substitution alone caused a ~2,000-fold reduction in peptide affinity to both Pirh2 (Figure 2A) and KLHDC10 (Figure 2B), implying that these E3 ligases make critical contacts with the degron's terminal alpha-carboxyl group.

In another set, we investigated the requirement of at least four Ala residues for a homopolymeric Ala-tail to act as a degron *in vivo*[8] by determining the number of Ala in a peptide required for binding to each E3 ligase. A peptide containing five C-terminal Ala had an IC_{50} of ~50 nM for Pirh2, and extending it to six Ala did not improve binding further (Figure S3A). However, reducing the tail length to four Ala decreased affinity to ~2 μM, and further shortening rendered binding undetectable (Figure 2C).

Similar observations were made for KLHDC10. Peptides with four C-terminal Ala bound strongly (IC_{50} 56 nM), and extending to five Ala further improved binding (IC_{50} 5 nM). However, further elongating to six Ala had no additional effect (Figure S3B). In contrast, a peptide with three Ala showed a 400-fold binding reduction compared with that with four Ala (IC_{50} 19 μM; Figure 2D). The diminished binding of Pirh2 or KLHDC10 to peptides with three or fewer Ala was not due to overall peptide length — Ala-tail peptide N-terminal deletions affected binding to Pirh2 (Figure 2E) and KLHDC10 (Figure 2F) to a lesser extent.

In the above experiments, the tail length requirement could have been influenced by specific residues negatively affecting binding when positioned at a fixed distance from the C terminus—for instance, having Lys at position −5 might reduce binding of an Ala_4-containing peptide. To explore this, we tested peptides with fixed length while progressively decreasing the number of C-terminal Ala residues by Gly replacement, confirming the dependence of peptide binding on Ala-tail length for both E3 ligases, with Ala_5 having the highest affinity (Figures S3C and S3D).

Next, we investigated the requirement of the C-terminal-most residue (defined as −1), which is critical for recognition by C-end E3 ligases. For example, KLHDC2 binds to C-terminal Gly degrons, TRIM7 favors Gln-end degrons, and FEM1 binds Argend degrons.[12,14,17,18] The effect of replacing Ala −1 in the Ala-tail peptide with Gly, Leu, or Asp was examined. Gly has a smaller side chain than Ala and lacks hydrophobicity. Leu, in contrast, has a larger hydrophobic side chain. Lastly, Asp has a small but negatively charged side chain. While the Ala-tail peptide strongly binds to Pirh2 (IC_{50} 50–120 nM), the IC_{50} for an Ala-tail peptide ending in Gly was ~2 μM, whereas Asp at −1 position further decreased the IC_{50} to ~26 μM, and introducing Leu at −1 abolished binding (Figure 2G). These results suggest a strict requirement for Ala at −1 for Ala-tail recognition by Pirh2.

Unlike Pirh2, KLHDC10 can bind to substrates ending in -TrpGly, -ProGly, and -AlaGly[9,10,12] in addition to homopolymeric Ala-tails.[8] Consistent with this, KLHDC10 strongly bound to an Ala-tail peptide ending in Gly −1 (IC_{50} ~2 nM; Figure 2H), comparably to the Ala-ending "parental" peptide. However, similarly to Pirh2, Ala-tail peptides ending in Asp or Leu had reduced affinity for KLHDC10 (~500-fold lower or undetectable, respectively).

Although Pirh2 and KLHDC10 sense homopolymeric Ala-tails within the context of RQC, these E3 ligases might also recognize endogenous Ala-rich C-degrons in native proteins. To gain preliminary understanding of potential C-end sequences that Pirh2 and KLHDC10 can accommodate, we investigated the impact of substituting Ala with Leu or Gly at various positions (−6 to −2) within competing peptides. For Pirh2, replacing Ala −6 with Leu or Gly had no adverse effect on binding (IC_{50} ~60 nM), and similar changes at positions −3 or −5 reduced binding by only ~3- to 6-fold (Figures 2I and S3E). However, changes at −2 caused a 60- to 70-fold decrease, while substitutions at −4 caused either complete loss of binding (Leu) or 100-fold reduction (Gly). Similar analyses for KLHDC10 revealed high-affinity binding to peptides with Leu or Gly at −6; both substitutions at −5 decreased binding by 3- to 6-fold; at −3, a Leu-containing peptide bound with high affinity, but Gly −3 reduced binding ~11-fold; and finally, the most pronounced effects were for substitutions at −2 and −4, with Gly −4 or Leu −2 causing a 300-fold decrease, while Leu −4 or Gly −2 had relatively smaller effects (~70- and ~3-fold, respectively) (Figures 2J and S3F).

In summary, polyAla must reside at a protein's extreme C terminus to function as a degron; at least four C-terminal Ala and the terminal carboxyl group are essential for E3-ligase binding; both E3 ligases can tolerate diverse non-Ala residues at −3 and −5 positions; and, consistent with previous reports,[9,10,12] KLHDC10 can also accept peptides with Gly at −1.

AlphaFold2 predicts Ala-tail degron-binding sites in Pirh2 and KLHDC10

We next sought to understand how Pirh2 and KLHDC10 recognize Ala-tails structurally. *In silico* analyses utilizing AlphaFold2[36,37] proved insightful in predicting structures of Pirh2 and KLHDC10 complexes with the Ala-tail peptide and in delineating molecular principles underlying the interactions (Figures 3 and 4).

In one analysis, $Pirh2_{1–195}$ and Ala-tail peptide (MDELYKAAAAAA) were input as separate chains. In another, a flexible Gly_{50} linker connected an Ala-tail to the C terminus of $Pirh2_{1–195}$. Both analyses predicted almost identical Ala-tail peptide binding modes (Figures S4A and S4B). The highest-scoring models from the analysis treating Pirh2 and peptide as separate chains are discussed further.

Comparing nuclear magnetic resonance (NMR) structures of Pirh2's isolated NTM and RING domains[26] with the AlphaFold2 $Pirh2_{1–195}$ model showed a root-mean-square deviation (RMSD) of ~1.15 Å for the pruned atom pairs, validating the model's accuracy (Figure S4C). Closer inspection of Zn-binding Cys and His residues from NTM and RING domains indicated consistent orientation between the AlphaFold2 and NMR structures (Figure S4C).

In the AlphaFold2-predicted full-length Pirh2 structure, the NTM and RING domains establish close contacts (Figure S1A). The interaction is mediated by a hydrophobic interface between NTM residues Leu124, Leu126, Leu130, and Ile137 and RING domains Val140, Val161, and Leu167. Moreover, several residues contribute polar interactions—Arg109 and Lys135 from the NTM hydrogen bond with main-chain carbonyls of His153 and Cys164 of the RING domain, respectively; Ser141, in the linker between the NTM and RING domains, hydrogen bonds with the NTM Asn123 main-chain carbonyl and the RING domain His153; and the NTM His22 main-chain carbonyl potentially interacts with the RING domain Arg156 side chain.

Strikingly, in the predicted structure of the Pirh2-peptide complex, the Ala-tail peptide binding site lies in a pocket at the NTM-RING domain interface (Figure 3A). This model is consistent with our biochemical findings uncovering the NTM domain as both required and sufficient for binding, with the RING domain making a stimulatory contribution (Figures 1B and 1C).

The predicted Ala-tail-binding site in Pirh2 is a narrow channel where the Ala-tail peptide fits precisely, in an extended, flexible loop-like conformation. The peptide's C terminus reaches deep into the binding pocket, while its N terminus lies on the shallow, open end. This pocket has features suitable for Ala-tail binding — it is amphipathic, with dispersed hydrophilic and hydrophobic residues, and has a basic patch that complements the peptide's C-terminal alpha-carboxyl and main-chain carbonyl groups (Figures 3B and 3C). The Ala-tail peptide's C terminus lies near Pirh2's polar Tyr23 and positively charged Arg41 and Arg54 (Figure 3D), whose side chains are positioned to establish hydrogen bonds and/or electrostatic interactions with the degron's obligatory, negatively charged terminal alpha-carboxyl group. The peptide's Ala −1 side chain fits into a shallow hydrophobic cleft formed by the NTM Leu92 and Ile110. The cleft's spatial constraint and hydrophobicity support Pirh2's preference for the small, hydrophobic side chain of Ala at position −1. The side chains of Ala −2, −4, and −5 are likewise all enclosed into small clefts distributed throughout the binding pocket, lined by Arg109 and hydrophobic residues Ile107, Phe116, and Leu124. Ala −3 is positioned with its side chain pointing away from the binding pocket toward relatively open space, suggesting that this degron position might be more amenable for substitution. This aligns with two observations: (1) substituting Ala-3 for Leu in the Ala_6 peptide used for the AlphaScreen assay only caused a 3-fold binding reduction (Figure 2I), and (2) Pirh2 could recognize the *B. subtilis* SsrA tag (which ends in ALAA) *in vitro*,[38] although with moderate affinity (IC_{50} ~470 nM; Figure S5). Lastly, Ala −6 also orients its side chain away from the pocket. Thus, Pirh2's binding pocket perfectly accommodates the last five Ala of the degron peptide, consistent with the strongest binding observed for the Ala-tail peptide containing five Ala (Figures 2C and S3A).

The structure suggests mechanisms underlying the stimulation of Ala-tail binding by Pirh2's RING domain. First, the RING domain contributes a lateral surface to the Ala-tail-binding pocket, promoting a more hydrophobic environment by excluding water. Moreover, the Ala-tail peptide backbone runs parallel to and establishes contacts with a β-sheet of the RING domain, involving several main-chain interactions, such as the hydrogen bonding between

the Ser155 main-chain carbonyl and Ala-2 amide and between the Ala159 backbone carbonyl and amide groups and the Ala-5 main chain.

For KLHDC10$_{\beta 1\text{-}88\text{–}442}$, AlphaFold2 predicted Ala-tail binding on the Kelch domain's "top" surface (loops linking β-strands A to D and B to C; Sprague et al.[39]). However, initial analyses yielded non-uniform peptide orientations, with the highest ranked model lacking KLHDC10 contacts with the degron's alpha-carboxyl group. In subsequent analyses using an Ala-tail peptide with shorter GFP-derived sequence to avoid "distraction" (Figure S6), the peptide consistently bound on KLHDC10's top surface with the C-terminal Ala penetrating deep in the central pocket (Figures 4A, S6A, and S6B). An equivalent mode was found for the di-Gly C-end degron binding to the KLHDC2 Kelch domain by co-crystallization.[12] Notably, in the model of apo-full-length KLHDC10, its N terminus, which has Ala at positions 3 and 4, folds back and binds to the same pocket as the Ala-tail peptide (Figure S2). This interaction is reminiscent of the C-degron mimicry by the KLHDC2 C terminus, which causes autoinhibition.[20] Whether the KLHDC10 interaction is real and Ala dependent, and whether it regulates substrate binding, remains unexplored.

Although the revised analyses showed a consistent mode of Ala-tail binding to KLHDC10, specific features of the models were more heterogeneous than for Pirh2. In the top prediction, KLHDC10 recognized the degron in a coiled conformation, stabilized through intrachain hydrogen bonds between backbone carbonyl and amide groups (Figure 4A, left). This aligns with the polyAla helical nature,[40,41] but helical structure has also been observed for C-end degrons with unrelated sequences bound to their cognate E3 ligases.[12,14,15,18] This alpha-helical conformation positions Ala side chains spirally against the pocket wall, with those of Ala-1 and Ala-2 against the narrowed hydrophobic surface. In contrast, in the top prediction of the fused KLHDC10-Gly$_{50}$ linker-Ala-tail, the degron bound in a disordered conformation (Figure 4A, right).

In either case, the KLHDC10 central pocket appears favorable for Ala-tail recognition, with a basic patch at its deep bottom and hydrophobic residues lining its wall (Figures 4B and 4C). The degron's C-terminal carboxyl group sits near KLHDC10 Ser93, Ser156, and Arg391 (Figure 4D), forming hydrogen bonds and/or electrostatic interactions; in two out of five predictions, Tyr110 is near the C-terminal carboxyl group, Tyr208 hydrogen bonds with the backbone amide of Ala-5, and Tyr261 hydrogen bonds with the degron's main-chain carbonyl of Ala-6. The binding pocket at this end is narrowed by bulky, hydrophobic side chains of KLHDC10 Phe176 and Phe366, which flank the degron's Ala −1 and −2, respectively; steric hindrance caused by such bulky residues potentially explains why KLHDC10 recognizes both Ala- and Gly-ending degrons. The pocket is wider at the top such that there are minimal contacts beyond the peptide's Ala −4, flanked by KLHDC10 Phe176. In contrast, in a model where the Ala-tail binds non-helically, Ala-2 is flanked by KLHDC10 Phe176, Ala-3 by Phe366, Ala-5 by Tyr208, and Ala-6 by Thr225 and Tyr230 (Figure 4E), while the terminal carboxyl is stabilized by Arg391, Asn389, and Tyr110. Additionally, the main-chain Ala-5 amide and Ala-6 carbonyl hydrogen bond with Tyr208 and Ser280, respectively.

Therefore, AlphaFold2 identifies Ala-tail-binding pockets in both Pirh2 and KLHDC10 that are consistent with biochemical features implicated in degron binding from both the E3 ligase and degron perspectives.

Validation of the Ala-tail degron-binding sites in Pirh2 and KLHDC10 by analysis of rationally designed mutations

We next tested the AlphaFold2-predicted Ala-tail-binding sites by introducing single amino acid mutations in Pirh2 and KLHDC10 and examining their effect on GFP-Ala_6 binding *in vitro* and *in vivo* and on reporter ubiquitylation and steady-state levels *in vivo*.

In vitro, Pirh2 mutations affecting polar, charged, or hydrophobic residues on the peptide binding pocket either abolished (Y23A, R41A, L92D, R109A, F116D) or decreased (R54A, I110D, L124D) GFP-Ala_6 pull-down (Figure 5A). The specificity of the results is further highlighted by the several mutants that had no obvious binding defect, including point mutations intended to disturb the NTM-RING domain interface (I137D, K135A, V140D, and data not shown).

To examine Pirh2 mutants *in vivo*, GFP or GFP-Ala_6 was co-expressed with wild-type (WT) or mutant full-length Pirh2 in HeLa cells. All Pirh2 mutants tested were less efficient, to varying degrees, in decreasing GFP-Ala_6 levels (Figure 5B). Accordingly, all mutants seemed less active in ubiquitylation, based both on the loss of the high-molecular-weight ladder above GFP-Ala_6[8] and on the accumulation of unmodified GFP-Ala_6, which was clearer in anti-GFP immunoblot following GFP immunoprecipitation (IP). Interaction of Pirh2 with GFP-Ala_6 was also tested. Like the *in vitro* data, all Pirh2 mutants tested (Y23A, R41A, R54A, L92D, I110D, F116D, and L124D) failed to coIP with GFP-Ala_6 (Figure 5B), correlating with their decreased stimulation of GFP-Ala6 ubiquitylation. Unlike the *in vitro* data, Val140 mutation adversely affected GFP-Ala_6 coIP under *in vivo* conditions.

Similarly, KLHDC10's ability to coIP with GFP-Ala_6 *in vitro* was affected by the mutation of several residues in the predicted Ala-tail-binding pocket (Figure 5C): Ser93, Tyr110, Ser156, and Arg391, which interact with the degron's C-terminal carboxyl group; Phe176, which flanks Ala-1 or Ala-2; Phe366, which interacts with either Ala-2 or Ala-3; Tyr261, which interacts with the main-chain carbonyl of Ala-6; and Tyr208, predicted to interact with the main-chain amide of Ala-5.

We next examined KLHDC10 mutation effects *in vivo*. KLHDC10 WT specifically coIPed with GFP-Ala_6 (Figure 5D; Thrun et al.[8]). Endorsing the critical role of several KLHDC10 residues in Ala-tail binding, S93A, F176D, Y208A, Y261D, and R391A mutants did not coIP with GFP-Ala_6, while Y110A interacted partially (Figure 5D). GFP-Ala_6 ubiquitylation by KLHDC10 was less clear than for Pirh2. Upon KLHDC10 WT expression, a prominent high-molecular-weight band appeared on the anti-ubiquitin blot of the GFP IP rather than a ubiquitin ladder or smear (Figure 5D). This band, also observed with weaker intensity in the presence of KLHDC10 Y110A, may correspond to heavily polyubiquitylated GFP-Ala_6. In contrast, for the other mutants tested, this band did not appear above background; instead, a weak band of size consistent with monoubiquitylated GFP-Ala_6 was observed in the anti-GFP blot (S93A, F176D, R391A; less so for Y208A and Y261D). The results suggest

that the KLHDC10 mutants have largely impaired ubiquitylation activity, correlating with defective GFP-Ala_6 interaction.

In conclusion, rational mutagenesis data support the computationally predicted Ala-tail-binding sites and their critical residues in Pirh2 and KLHDC10.

Conservation of the C-degron-binding site in Pirh2 and KLHDC10 homologs

We next examined whether the identified C-degron-binding sites are shared among Pirh2 and KLHDC10 homologs.

A multiple sequence alignment (MSA) highlights conservation of several Pirh2 residues predicted to create the degron-binding pocket and to bind the Ala-tail across all major eukaryotic lineages (Figure 6A). The positions corresponding to Tyr23, Arg41, and Arg54 map to the NTM's N-terminal-most CHY Treble-clef domain (Figure S1). The first of these residues is almost always aromatic (Tyr or Trp) and is structurally adjacent to Arg41 and Arg54, which are both retained as positively charged positions in most homologs. This is consistent with the notion that the aromatic position engages in cation-π interactions with the two positively charged residues, stabilizing the end of the binding pocket and providing a chemical "counterpart" for the Ala-tail's terminal carboxyl group. Mapping the surface residue conservation on human Pirh2 revealed clear clustering of conservation at the binding pocket's deep end and at the RING domain surface that binds to E2 conjugases (Figure 6B). Finally, the Ala-tail-binding pocket was structurally conserved at the interface of the NTM and the RING domain from fission yeast to humans (Figure 6C). Together, these observations suggest that Pirh2 homologs in other organisms are equipped for functioning in Ala-tail binding. Substantiating this notion, *S. pombe* Pirh2 — the most distant and divergent sequence in Figure 6A—selectively interacted with Ala-tails (Figure 6D).

Evolutionary analysis of KLHDC10 through a phylogenetic tree clearly distinguished it from KLHDC2 and KLHDC3, the defining members of paralogous sister clades (Figures 7A, S7, and S8). KLHDC10 has a more restricted distribution than Pirh2, being present across animals, protist sister groups (choanoflagellates, filastereans, ichthyosporeans), and other lineages (stramenopiles, alveolates) but lost in fungi. KLHDC10 orthologs maintain sequence conservation (Figure 7B) and the predicted Ala-tail-binding pocket (Figure 7C), including residues for degron binding (Ser93, Ser156, Phe176, Tyr208, Thr225, Tyr261, Phe366, Arg391).

Residues predicted to contact the Ala-tail are selectively retained in the KLHDC10 clade, suggesting that Ala-tail binding is likely to be conserved (Figures S7–S9). The positions corresponding to Ser93, Tyr110, and Phe366 (as an aromatic) are also conserved in KLHDC2, which has distinct degron-binding specificity. Therefore, these residues might play a more general role in recognizing C-degrons across clades, whereas positions conserved only in KLHDC10 likely mediate specificity toward Ala-tails (Figures 7 and S8).

The above results strongly support the AlphaFold2-predicted Ala-tail degron-binding sites in Pirh2 and KLHDC10. Moreover, they implicate Pirh2 and KLHDC10 homologs in having a conserved Ala-tail-sensing function.

Evolutionary affinities and origins of the Ala-tail-recognition modules of Pirh2 and KLHDC10

Pirh2 and KLHDC10 are unrelated, indicating dual emergence of Ala-tail binding in eukaryotic evolution. Our analyses show Pirh2 orthologs widely distributed across major eukaryotic lineages. While there is some uncertainty about the eukaryotic tree, Pirh2's presence in "greater plants" (Archaeplastida) and lineages considered basal branches of the tree, like Heterolobosea (e.g., *Naegleria*) and Metamonada (*Giardia, Spironucleus*), suggests its presence in the LECA (Figure S10A). This ancestral form can be reconstructed as already having the tripartite NTM, central RING, and C-terminal ZnR organization, along with most key Ala-tail-recognition residues (Figures S10C and S10D).

In contrast, KLHDC10 has a patchier distribution, primarily found in animals, their immediate sister groups, and some Stramenopile-Alveolate-Rhizarian (SAR) lineages (Figure S8), suggesting non-LECA inheritance. The KLHDC2 clade is distributed primarily in animals and amoebozoans. However, our analyses indicate that KLHDC3 is more widespread, spanning animals and their closest sister groups, amoebozoans, SAR, and plants. KLHDC3 thus likely represents an older Kelch-repeat C-degron-binding clade (but cannot be confidently inferred as being in the LECA), with KLHDC10 and KLHDC2 evolving later for distinct specificities. Hence, the more widespread Pirh2 is probably the original Ala-tail-specific E3 ligase from the LECA.

However, the presence of two analogous E3 ligases with overlapping specificities might have allowed the loss of either version in certain lineages. For example, Pirh2 was likely secondarily lost in certain animals such as *C. elegans* and kinetoplastids. In *C. elegans*, a KLHDC10 homolog allows the use of Ala-tail C-degrons (Figure 7). In contrast, *S. cerevisiae*, an Ascomycete fungus, lacks both Pirh2 and KLHDC10, potentially due to the presence of Ala-Thr-tails instead of pure Ala-tails.[23] Basidiomycete fungi have similarly lost both Pirh2 and KLHDC10.

Moreover, several eukaryotic clades display unusual modifications of the ancestral Pirh2 system (Figure S10A). (1) In amoebozoans, the NTM combines with an F-box instead of a RING domain, suggesting that the same Ala-tail-recognition module operates via an E3 ligase with an F-box subunit. (2) Across "greater plants," Pirh2 fuses with 4 heme-binding hemerythrin domains, suggesting potential interface with light or redox sensing. (3) In cnidarians, Pirh2 fuses to N-terminal death-effector domains (DEDs) that mediate homotypic interactions specific to apoptosis in animals.

These observations raise the question of the origin of Pirh2's tripartite NTM in the first place. Sequence profile searches recover significant similarity to various proteins from eukaryotes, bacteria, and archaea (Figure S10). Their common denominator is the CHY Treble-clef and central ZnR domain (together defined as the CHY-finger module) and possibly including a rudimentary version of a single Treble-clef fragment repeat

(Figure S10). This minimal version occurs as a standalone protein in several bacteria and Haloarchaea and was secondarily transferred to fungi (e.g., *S. cerevisiae* Hot13). There is no evidence linking this core NTM version to Ala-tailing/RQC or prokaryotic ubiquitin-like-conjugating systems, and, accordingly, *S. aureus* AND44041.1 did not coIP with GFP-Ala_6 (Figure S11). Rather, conserved operonic linkages and direct fusions of the NTM precursor with biotin transporter (BioY) and ligase (BirA) suggest an adaptor function in aiding biotinylation or regulating biotin transport/conjugation through binding BioY's hydrophobic tail. In contrast, yeast Hot13 specifically binds Mia40, the receptor for mitochondrial intermembrane protein import, while also sequestering Zn ions from it.[42]

Our analysis showed that the minimal NTM version was acquired from bacteria and incorporated into two distinct ubiquitin-related functions in the LECA (Figure S10). In one case, it was expanded and fused to a RING domain, resulting in Pirh2. In the other, it merged with the RWD domain and nucleic-acid-binding CCCH domains (and sometimes additionally DNA-binding MYB domains) (Figure S10A). The RWD domain in the latter proteins might interact with E2 domains or nucleoprotein complexes, including ribosomes, while the NTM likely participates in peptide recognition. Comparing the NTM core between Pirh2, the Hot13-like clade, and the RWD-fused clade reveals conservation of the aromatic position equivalent to Pirh2 Tyr23 but not the basic positions equivalent to Arg41 and Arg54. The latter are aromatic residues in Hot13, forming an aromatic cage, and are replaced by an acidic residue in the RWD-fused clade. Thus, the Ala-tail-binding site appears to have emerged specifically in Pirh2, from a distinct ancestral binding site inherited from the bacterial versions.

In Pirh2, the peptide-binding site of the NTM was augmented by triplication of the C-terminal Treble-clef fragment (Figures S1B–S1E), with our results suggesting an important role for this augmentation in Ala-tail binding. The triplicated Treble-clef fragment additionally acquired an independent existence in eukaryotes, coupled either to transmembrane (TM) segments, creating the DHHC palmitoyltransferase catalytic domain,[43] or to the C-terminal RNA-binding domain of the 28S rRNA methyltransferase ZCCHC4 (Figures S10B and S1E). The former binds to palmitoyl CoA and the linker peptide that connects it to the TM-segment comparably to Pirh2-Ala-tail binding (Figure S10).

DISCUSSION

Our study provides insights into the selective recognition of a simple, homopolymeric polyAla degron sequence and sheds light onto the evolution of Ala-tails as a proteolytic signal.

AlphaFold as a prediction tool for protein-peptide complexes

AlphaFold successfully predicts protein structures based solely on sequences,[36] with recent applications expanding toward protein-peptide complex co-structures.[44,45] Indeed, AlphaFold2 generated reliable models for Pirh2 and KLHDC10 complexes with Ala-tails: first, these models explain experimentally determined features of both E3 ligases and the degron implicated in the interactions (Figures 1 and 2). Second, site-specific mutagenesis confirmed the contribution of residues predicted to mediate degron binding (Figures 3–

5). Third, the degron-binding pockets in both E3 ligases, pertinent chemical features on the pocket surfaces, and specific residues predicted to mediate Ala-tail binding are evolutionarily conserved (Figures 6 and 7).

The predicted binding modes explain experimentally determined features of both E3 ligases and the degrons that characterize the interactions

Pirh2 and KLHDC10 strongly bind to the Ala-tail, comparably to some other E3 ligase-degron pairs.[12] High-affinity binding likely evolved to degrade low-abundance substrates, eliminate substrates quickly, and/or keep levels of potential toxic substrates low,[12] aligning with Pirh2 and KLHDC10's protein quality control function.[8] Consistent with experimentally determined binding affinities, AlphaFold2 models for Pirh2- and KLHDC10-Ala-tail complexes feature extensive pocket interfaces. The Ala-tail C terminus is fully concealed in KLHDC10's deep binding pocket, while it is not entirely buried in the Pirh2 pocket, possibly explaining KLHDC10's stronger binding.

The structural predictions also explain Pirh2 and KLHDC10's Ala-tail selectivity based on the degron's sequence-, length-, and C terminus-dependence for function. The binding pockets accommodate 5 degron residues engaged in continuous contacts, consistent with high-affinity binding to both E3 ligases requiring at least four Ala and with the further increased binding to a degron with five Ala. For Pirh2, the side chains of Ala −1, −2, −4, and −5 are enclosed into small clefts. Notably, the choice of Ala −1 by both Pirh2 and KLHDC10 is specified by a small hydrophobic cleft that cannot accommodate larger side chains. Analogous selectivity mechanisms accommodating terminal degron residues have been reported for other E3-ligases–for example, FEM1 stabilizes the terminal Arg in its cognate C-degron by an acidic cleft.[15] The structures also suggest that the Ala-tail's C-terminal backbone carboxyl group mediates binding to both Pirh2 and KLHDC10, consistent with the requirement for a polyAla tract at the substrate's very C terminus to act as degron (Thrun et al.[8]; Figures 2A and 2B). In both E3 ligases, polar and charged residues forming a highly electropositive surface surround this C-terminal carboxyl group (Figures 3B and 4B). Similar polar residue distribution is observed in crystal structures of other E3 ligases complexed with C-end degron peptides.[12,14,17,18]

Finally, the structural model of the Pirh2 complex predicts close intramolecular contacts between the Pirh2 NTM and RING domains and that the Ala-tail peptide binding pocket lies in the NTM at that interface (Figure 3A). The model is consistent with the observed stimulatory effect of Pirh2's RING domain in Ala-tail binding to the NTM (Figures 1B and 1C).

Recognition of non-RQC substrates by Pirh2 and KLHDC10

Apart from Ala-tailed proteins generated by NEMF activity in RQC, Pirh2 can artificially bind to the bacterial SsrA degron,[38] which ends in ALAA. Predicting native full-length targets of Pirh2 or KLHDC10 is difficult due to absence of human proteins whose cytosol-accessible C termini have more than 3 consecutive Ala. Nevertheless, fusion to the C-terminal 23 residues of TRAPPC11, which ends with only 3 Ala (-MDDTSIAAA), caused GFP instability *in vivo* (contrasting with GFP-Ala3, which was more stable), and fusion

protein expression was restored upon Pirh2 and KLHDC10 depletion.[8] These results and the tolerance for residues other than Ala at certain terminal positions in the degron peptide (Figures 2I, 2J, S3E, and S3F) suggest that, beyond homopolymeric Ala-tails, these E3 ligases may use a comparable mechanism to target native proteins harboring C-end degrons rich in Ala. Additionally, Ala-tail-like degrons might conceivably arise from proteolytic cleavage.

Various substrates are reported for Pirh2[25–29] and KLHDC10,[33] none having C-terminal sequences resembling an Ala-tail. In p53 binding, Pirh2's CTD is central, while the NTM contributes transiently.[26] Likewise, Pirh2's NTM and CTD cooperate in binding to c-Myc.[26,29] Despite c-Myc's Ala-terminal residue, the corresponding peptide (KHKLEQLRNSCA) did not bind Pirh2 or KLHDC10 *in vitro* (Figure S12).

Evolutionary implications of Ala-tail signaling via Pirh2 and KLHDC10

Ala-tailing has ancient evolutionary history. NEMF/RqcH homologs, which mediate C-terminal tailing, are tRNA- and ribosomal 60S subunit-binding proteins represented across all domains of life.[46] The bacterial RqcH-mediated Ala-tailing mechanism[46] is related to yeast Rqc2-mediated Ala-/Thr-tailing and mammalian NEMF-mediated Ala-tailing.[8,47,48] Thus, Ala-tailing mediated by a NEMF/RqcH/Rqc2 ancestor likely already existed in the last universal common ancestor (LUCA). However, how Ala-tailing triggers proteolysis widely differs across the Tree of Life.

In the bacterial common ancestor, an additional mechanism arose to tag products of stalled ribosomes for degradation-transfer-messenger RNA (tmRNA)/SsrA. Although mechanistically unrelated, SsrA and RQC modify incomplete nascent chains with a C-terminal degron: polyAla in bacterial RQC[46] and an 8–35 residue-long SsrA peptide ending in ALAA in most bacteria.[49] These mechanisms have convergently evolved, with both tags being directly recognized by the ClpXP protease.[46,49]

In eukaryotes, with the ubiquitin system expansion, further convergent ubiquitin-dependent mechanisms arose targeting C-terminal tail-extended stalled polypeptides for degradation. In one such mechanism, C-terminal tailing facilitates Ltn1-mediated ubiquitylation on the 60S subunit (RQC-L; Thrun et al.[8]). Given Ltn1's prevalence across most eukaryotic lineages, this activity probably already existed in the LECA.

Here, we show that at least one E3 ligase from the RQC-C branch of eukaryotic RQC, viz., Pirh2, likely also existed in the LECA. In RQC-C, Ala-tailing has an extra-ribosomal degron function, analogous to bacterial RQC.[8] This in turn predicts that Ala-tail synthesis by eukaryotic NEMF orthologs runs in parallel with the presence of E3 ligases specifically recognizing homopolymeric Ala-tails. Keeping with this, the loss of both Ala-tail-binding E3 ligases in *S. cerevisiae* likely relates to the incorporation of both Ala and Thr in the Rqc2-dependent reaction.[23] In addition to supporting Ltn1 in RQC-L,[23,24,50] C-terminal tails containing both Ala and Thr tend to form amyloid aggregates,[22] suggesting that *S. cerevisiae* may have evolved a variant RQC activity mediating protein aggregation to either sequester aberrant translation products or enable their elimination via aggrephagy.[51]

Our analysis unravels the evolutionary trajectories of two distinct Ala-tail-recognition modules in eukaryotic E3 ligases. KLHDC10 evolved a binding pocket by subtly tinkering with a preexisting C-degron-recognition pocket shared with KLHDC2 and KLHDC3. In contrast, Pirh2 represents a more dramatic case, wherein a module comprised of three domains of bacterial provenance, predating the LECA, was repurposed for Ala-tail recognition. This process involved subtly modifying an ancient binding pocket and its augmentation via triplication of the Treble-clef fragment.

The results provide further evidence for Ala-tails' widely conserved role as a proteolysis tag and guide future analyses across diverse organisms regarding these protein modification and sensing mechanisms.

Limitations of the study

Our primary objective was to determine how Pirh2 and KLHDC10 recognize Ala-tails structurally and mechanistically. While AlphaFold2-based modeling provided valuable structural insights, the analyses could not distinguish between two KLHDC10-degron complexes with distinct degron conformations. Additionally, the predictions for both E3 ligases may be incomplete; for example, water molecules absent in AlphaFold models might contribute to Ala-tail binding. Therefore, experimentally determined structural models remain necessary to independently confirm our conclusions and verify interaction specifics.

STAR★METHODS

RESOURCE AVAILABILITY

Lead contact—Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Claudio A. Joazeiro (c.joazeiro@zmbh.uni-heidelberg.de).

Materials availability—All unique reagents generated in this study are available from the lead contact with completed material transfer agreement if there is potential commercial application.

Data and code availability

- All data reported in this paper will be shared by the lead contact upon request.
- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Bacterial strain for recombinant protein expression—E. coli BL21 (DE3) cells were used to express recombinant proteins. Cells were transformed with respective plasmids and grown in LB media with required antibiotics as explained below.

Cell lines—HeLa cells were cultured in DMEM (GIBCO) with 10% v/v FBS (GIBCO) and 2mM L-Glutamine (GIBCO) at 37°C and 5% CO_2.

METHOD DETAILS

Plasmid constructs—As described previously for Pirh2 and KLHDC10,[8] KLHDC2 cDNA was subcloned into a modified pET15b vector that carries an N-terminal GST tag sequence. Truncation constructs of Pirh2 and KLHDC10 were generated by amplifying each fragment by PCR and subcloned into pET15b vector with an N-terminal GST tag sequence. The gene sequence coding for the full-length *S. pombe* Pirh2 homolog was synthesized (IDT) and subcloned into the pET15b vector with an N-terminal GST tag. KLHDC10$_{\beta 1\text{-}88\text{-}442}$ was generated by deleting the flexible region at the N terminus (residues 1–39) and the internal loop (residues 52–87) of KLHDC10. KLHDC10$_{\beta 1\text{-}88\text{-}442}$ corresponds to the first β strand (residues 40–51) connected by a linker (GSGSG) to the rest of the protein (residues 88–442). The codon-optimized sequence with flanking NdeI and XhoI restriction sites was generated by IDT and subcloned in the pET15b vector with an N-terminal GST tag. The codon-optimized, bicistronic insert encoding Elongin B (residues 1–104) and Elongin C (residues 1–112) cloned in pET28b was described previously.[8] For Pirh2 and KLHDC10 mutants, single point mutations were introduced into GST-Pirh2$_{1\text{-}195}$ and GST-KLHDC10$_{\beta 1\text{-}88\text{-}44}$ in pET15b by site-directed mutagenesis.

The generation of 3X FLAG-tagged full length Pirh2 and KLHDC10 cloned into pcDNA5 for expression in mammalian cells was reported previously.[8] Mutants indicated in the study were generated by introducing single point mutations by site-directed mutagenesis. All constructs were verified by sequencing.

Recombinant protein purification—For expression of Pirh2 constructs, *E. coli* BL21 (DE3) cells were grown in LB media in presence of ampicillin (100 μg/mL), 10 mM $MgCl_2$ and 10 μM $ZnCl_2$. For expression of KLHDC10 and KLHDC2 constructs, *E. coli* BL21 (DE3) cells were co-transformed with pET15b vector carrying KLHDC10 and KLHDC2 and pET28b vector carrying a bicistronic insert for the expression of Elongin B (1–104) and Elongin C (1–112) as described before.[8] Cells were grown in LB media with ampicillin (100 μg/mL) and kanamycin (50 μg/mL). For expression of all recombinant proteins, cells were grown until OD 0.8 at 37°C, followed by overnight incubation at 18°C in presence of 0.75 mM IPTG. On the following day, cells were harvested and stored at −80°C until further processing.

GST-tagged Pirh2 proteins were purified by cell lysis using sonication in ice-cold lysis buffer (50 mM Tris pH 7.5 at 4°C, 150 mM NaCl, 1 mM DTT, 10 μM $ZnCl_2$, 1% Triton X-100 and EDTA-free protease inhibitor (Roche)). Lysates were clarified by centrifugation and supernatants were incubated with glutathione agarose bead slurry (Machery-Nagel) prewashed in wash buffer (50 mM Tris pH 7.5 at 4°C, 150 mM NaCl, 1 mM DTT, 10 μM $ZnCl_2$). After 2 h incubation at 4°C, beads were washed three times using wash buffer. Bound GST-Pirh2 was eluted from beads by glutathione-containing elution buffer (50 mM Tris pH 7.5 at 4°C, 150 mM NaCl, 1 mM DTT, 10 μM $ZnCl_2$, 10 mM reduced glutathione). To remove glutathione, protein fractions were pooled and buffer-exchanged with wash buffer

using Amicon Ultra-4, 10 kDa filters. Concentrated protein was aliquoted and stored at −80°C.

For purification of GST-tagged KLHDC10 and GST-KLHDC2 proteins, cells were lysed by sonication in ice-cold lysis buffer (50 mM Tris pH 7.5 at 4°C, 150 mM NaCl, 1 mM DTT, 0.5% Triton X-100 and EDTA-free protease inhibitors (Roche)). Proteins were purified using glutathione agarose bead as described above except that buffers did not contain $ZnCl_2$. Buffer-exchanged, concentrated protein fractions were aliquoted and stored at −80°C.

BL21 cells co-transformed with $KLHDC10_{\beta 1\text{-}88\text{-}442}$ and Elongin B/C constructs were grown in TB media at 37°C until OD 1, when the culture was cooled down to 18°C and protein expression induced with 0.4 mM IPTG. After 18 h, cells were harvested and re-suspended into 50 mM Tris-HCl pH-7.5, 300 mM NaCl, 10% glycerol, 1 mM PMSF, 0.1% lysozyme, 5 mM DTT and Roche protease inhibitor tablets. Cells were lysed by passing twice through a microfluidizer under 1,500 bar, at 4°C. The lysate was centrifuged for 1 h at 4°C at 15,000 × *g*. The supernatant was passed through a 0.8 μm filter and incubated with glutathione resin pre-equilibrated with re-suspension buffer. Beads were washed with the same buffer except for using a higher salt concentration (500 mM NaCl). The protein was eluted with 50 mM Tris-HCl pH 7.5, 500 mM NaCl, 15 mM glutathione and 5 mM DTT. Fractions containing the proteins of interest were pooled and diluted to 50 mM NaCl concentration using 50 mM Tris-HCl pH 7.5. This sample was subjected to anion exchange chromatography (Hiprep Q column) using a gradient between buffer A (50 mM Tris-HCl pH 7.5, 50 mM NaCl, 5 mM DTT) and buffer B (50 mM Tris-HCl pH 7.5, 1 M NaCl, 5 mM DTT). The peak fractions containing the proteins of interest were pooled and concentrated to 1 mL using Centricon filter (10 kDa cut-off). This sample was then used for size exclusion chromatography using Superdex 16/600 75pg. The complex of GST-KLHDC10 and Elongins B-C was concentrated to 1 mg/mL and flash frozen at −80°C for AlphaScreen assays used to evaluate Ala-tail peptides.

Following a previously reported protocol, GFP, GFP-Ala_6, GFP-Thr_6 and GFP-$(Ala\text{-}Thr)_3$ were purified from *B. subtilis* Δ*clpP* strains grown in LB media at 37°C until OD 2.0.[46] Cells were harvested and lysed as above in 50 mM Tris pH 8.0 at 4°C, 150 mM NaCl, 0.5% NP40, 1 mM DTT. Clarified lysates were mixed with GFP-Trap agarose slurry by rotating for 1 h at 4°C. Beads were washed in buffer without DTT. Elution was performed using 0.2 M glycine pH 2.5 at 4°C, and the pH was neutralized by adding 1M Tris base pH 10.4. Aliquots of purified protein were stored at −80°C with 10% (v/v) glycerol.

***In vitro* binding assay—**Binding assays were performed using recombinantly purified proteins GST-Pirh2 or GST-KLHDC10 and GFP-Ala_6.[8] Briefly, glutathione agarose beads (Machery-Nagel) were washed and blocked with 1% BSA in assay buffer (20 mM Tris pH 7.5, 100 mM NaCl, 0.1% Tween 20) for 2 h at 4°C. Glutathione beads were conjugated with 0.5–10 μg GST-Pirh2 or GST-KLHDC10 by incubating 10 μL BSA-blocked beads together with the respective protein in 100 μL assay buffer at 4°C. Following 2 h of incubation, beads were washed twice in assay buffer to remove excess unbound protein. Protein-conjugated beads were then incubated with 0.25 μg of either GFP or GFP-Ala_6 in 100 μL assay buffer supplemented with 0.1% BSA for 1 h at 4°C. Beads were then washed thoroughly to remove

unbound protein and were resuspended in sample buffer. Binding of GFP-Ala_6 was detected by anti-GFP and anti-GST blots.

CRISPR-mediated Pirh2 knockout (KO)—CRISPR-mediated genome editing was performed as described previously.[8] HeLa cells were cultured in 10 cm dishes until reaching 70% confluency, then transfected with a PX458 plasmid containing gRNA targeting RCHY1 (Pirh2) using Lipofectamine 3000 (Invitrogen), following the manufacturer's protocol. After 24 h incubation, cells were trypsinized and resuspended to create a single-cell suspension in a solution consisting of 1X DPBS with 0.5% FBS. Subsequently, cells were sorted using a BD FACSAria III instrument and BD FACSDiva software.

The top 5% of GFP-positive cells were isolated as single cells and sorted into 96-well plates, with 7–12 day recovery by incubation in High Survival Media (HSM, 20% log-phase conditioned media comprised 20% FBS, mixed in DMEM) at 37°C. Monoclonal cell lines were expanded, and knockout (KO) cells were identified by Western blot against Pirh2. Validated KO clones were further verified by extracting gDNA using DirectPCR (Viagen), followed by amplification of the relevant exon using PCR, TOPO TA Cloning (Invitrogen), and Sanger sequencing (Eurofins GATC) to confirm the presence of the intended mutation.

The HeLa Pirh2 knockout cell line (PKO) has a deletion of A6146/G6147 in compound heterozygosity with a 4-bp deletion from A6144 to G6147, in exon 2 of the Pirh2 gene (ENSG00000163743, GRCH38.p13). These deletions result in the addition of 0 and 9 amino acids after Pirh2 F55, respectively, and are followed by an in-frame stop codon.

Immunoprecipitation—One million HeLa cells (WT or PKO) were plated per 6-cm dish. The next day, cell transfection (5 μg of total DNA with a 1:1 mass ratio of each plasmid) was conducted with Lipofectamine 3000 reagent. 24 h post-transfection, cells were treated with 10 μM MG132 for 4 h, washed in ice-cold PBS, and lysed in ice-cold lysis buffer (20 mM Tris-HCl pH 7.5, 100 mM KOAc, 5 mM $MgCl_2$, 1% β-Octyl Gluco-side, 1 mM DTT, supplemented with EDTA-free protease inhibitor tablet (Roche)) for 30 min at 4°C. After a clarification step (12,000 × g for 5 min at 4°C), an aliquot was collected (Whole cell lysate, WCL fraction) and 0.5 mg of total protein were subsequently incubated with 10 μL GFP-Trap magnetic agarose beads (ChromoTek) for 3 h at 4°C. Beads were washed 3× in lysis buffer without DTT and proteins were eluted by boiling for 5 min at 95°C in 30 μL of 2× Laemmli buffer.

Immunoblotting—Samples from *in vitro* binding assays, cell lysates and GFP-IP were separated by SDS PAGE on 4–20% Bis-Tris gel (GenScript) and transferred to PVDF membrane (0.45 μm) for 1h 20 min at 35 V using Invitrogen mini blot module or to PVDF membrane (0.2 μm) using TransBlot turbo Transfer system (Bio-Rad). For protein detection, membranes were incubated with primary mouse monoclonal anti-GFP (Roche), goat polyclonal anti-GST (GE Healthcare), mouse monoclonal anti-FLAG M2 (Sigma Aldrich) or rabbit polyclonal GRP94 (Cell Signaling Technology), followed by HRP-conjugated secondary antibodies (Dianova) and treatment with enhanced chemiluminescence substrate kit (ECL, GE Healthcare). Images were acquired with LAS4000 (GE ImageQuant) or ImageQuant 600 (Cytiva).

AlphaScreen proximity assay—Ala-peptide binding to Pirh2 or KLHDC10 was evaluated using the AlphaScreen GST detection Kit (PerkinElmer) proximity-based assay. The two assay components, streptavidin-coated donor beads that associate with biotin-labeled peptide and anti-GST acceptor beads that associate with GST-fused protein, are brought in proximity depending on peptide-protein interaction. Upon excitation of donor beads, energy transfer to acceptor beads takes place that generates a luminescence signal. Experiments were performed in ProxiPlate 384-well microplates (PerkinElmer).

Competition experiments to determine IC_{50} were performed by titrating unlabeled peptides. All required dilutions of reaction components were made in assay buffer (20 mM Tris pH 7.5 at room temperature, 100 mM NaCl, 0.1% Tween 20, 0.05% BSA). 10 nM biotin-labeled Ala-tail peptide (MDELYKAAAAAA) (Biosyntan) and 1 nM GST-Pirh2 or GST-KLHDC10 were pre-incubated with varying concentrations of competing unlabeled Ala-tail variant peptide (GenScript) for 40 min at RT. Following this incubation, pre-mixed donor and acceptor beads (10 μg/mL each) were added to the reaction mixture, in a total 20 μL, and further incubated for 40 min at RT. AlphaScreen signal was measured using PHERAstar microplate reader (BMG Labtech) at RT. Each reaction was performed in triplicates.

Structure predictions using AlphaFold2—Peptide bound complexes of Pirh2 and KLHDC10 were predicted using AlphaFold2.[36,37] The experiments were run through ColabFold[52] assessed through UCSF ChimeraX 1.4 and 1.5.[53] For one experiment, sequences of $Pirh2_{1\text{-}195}$ or $KLHDC10_{\beta1\text{-}88\text{-}442}$ and Ala-tail (MDELYKAAAAAA) were input separately. In a second experiment, the Ala-tail peptide sequence was fused to the C terminus of the protein following a Gly_{50} linker. For KLHDC10, complexes were also predicted using a shorter Ala-tail peptide (LYKAAAAAA). Default settings without any templates from PDB database were used to run the predictions. For every run, the top 5 predictions were obtained as output from the server. Models were analyzed using UCSF Chimera.[64]

Sequence analysis—The sequences of the mouse Pirh2 and human KLHDC10 were obtained from the National Center for Biotechnology Information (NCBI) GenBank database[65] and dissected into constituent domains using known structures and reverse searches against sequence profiles. Sequence similarity searches were performed using the PSI-BLAST program[55] against the NCBI non-redundant (nr) database or the same database clustered down to 50% sequence identity using the MMseqs program[66] with a profile-inclusion threshold set at an e-value of 0.01. Profile-profile searches were performed with the HHpred program.[60] Multiple sequence alignments (MSAs) were constructed using the FAMSA[56] and MAFFT programs.[57] Multiple sequence alignments (MSA) of Pirh2 and KLHDC10 homologs were generated using Clustal Omega[54] and displayed using ESPript 3.0.[58] Conserved residues were mapped onto the AlphaFold predicted structures of full-length human Pirh2 and KLHDC10 using ConSurf webserver (https://consurf.tau.ac.il/consurf_index.php).[59]

Structure analysis—The JPred program was used to predict secondary structures using MSAs (see above). PDB coordinates of structures were retrieved from the Protein DataBank and rendered, compared, and superimposed using the Mol* program.[67] Structural models

were generated using the RoseTTAfold[61] and Alphafold2 programs.[36] Multiple alignments of related sequences (>30% similarity) were used to initiate HHpred searches for the step of identifying templates to be used by the neural networks deployed by these programs.

Comparative genomics and phylogenetic analysis—Phylogenetic analysis was performed using the maximum-likelihood method with the WAG or LG models with the IQTree[62] and FastTree program,[63] with 8–20 rate categories and 1 invariant category for the sites. The FigTree program (http://tree.bio.ed.ac.uk/software/figtree/) was used to render phylogenetic trees. Gene neighborhoods were extracted through custom PERL scripts from genomes retrieved from the NCBI Genome database. These were then clustered using MMSEQS adjusting the length of aligned regions and bit-score density threshold empirically and filtered using neighborhood distance cutoffs and phyletic patterns to identify conserved gene neighborhoods.

QUANTIFICATION AND STATISTICAL ANALYSIS

AlphaScreen assay data were plotted using GraphPad Prism 8. Mean and standard deviation were calculated for each data point (n = 3) and error bars represent standard deviation as indicated in the figure legends. Half-maximal inhibitory concertation (IC_{50}) was determined by non-linear curve fitting of the dose-response curve in GraphPad Prism 8.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

We thank S. Pfeffer, R. Wade, M. Mayer, S. Filbeck, S. Richter, and A.R. Carvajal for helpful discussions. We thank X. Zhao for assistance with KLHDC10 purification; A. Thrun, T. Dallinger, and M. Bobeldijk for constructs; and C.S. Umbaugh for the Pirh2-knockout HeLa cells. This work was supported by R01 grant NS102414 from the NINDS of the NIH (to C.A.P.J.; the content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH) and EU/EFPIA/OICR/McGill/KTH/Diamond Innovative Medicines Initiative 2 Joint Undertaking (EUbOPEN grant no 875510) (to I D.). L.A. and A.M.B. are supported by intramural funds of the NIH's National Library of Medicine.

REFERENCES

1. Pilla E, Schneider K, and Bertolotti A (2017). Coping with Protein Quality Control Failure. Annu. Rev. Cell Dev. Biol 33, 439–465. 10.1146/annurev-cellbio-111315-125334. [PubMed: 28992440]
2. Hershko A, Ciechanover A, and Varshavsky A (2000). The ubiquitin system. Nat. Med 6, 1073–1081. 10.1038/80384. [PubMed: 11017125]
3. Zheng N, and Shabek N (2017). Ubiquitin Ligases: Structure, Function, and Regulation. Annu. Rev. Biochem 86, 129–157. 10.1146/annurev-biochem-060815-014922. [PubMed: 28375744]
4. Deshaies RJ, and Joazeiro CAP (2009). RING domain E3 ubiquitin ligases. Annu. Rev. Biochem 78, 399–434. 10.1146/annurev.biochem.78.101807.093809. [PubMed: 19489725]
5. Varshavsky A (2019). N-degron and C-degron pathways of protein degradation. Proc. Natl. Acad. Sci. USA 116, 358–366. 10.1073/pnas.1816596116. [PubMed: 30622213]
6. Timms RT, and Koren I (2020). Tying up loose ends: the N-degron and C-degron pathways of protein degradation. Biochem. Soc. Trans 48, 1557–1567. 10.1042/BST20191094. [PubMed: 32627813]

7. Sherpa D, Chrustowicz J, and Schulman BA (2022). How the ends signal the end: Regulation by E3 ubiquitin ligases recognizing protein termini. Mol. Cell 82, 1424–1438. 10.1016/j.molcel.2022.02.004. [PubMed: 35247307]

8. Thrun A, Garzia A, Kigoshi-Tansho Y, Patil PR, Umbaugh CS, Dallinger T, Liu J, Kreger S, Patrizi A, Cox GA, et al. (2021). Convergence of mammalian RQC and C-end rule proteolytic pathways via alanine tailing. Mol. Cell 81, 2112–2122.e7. 10.1016/j.molcel.2021.03.004. [PubMed: 33909987]

9. Lin HC, Yeh CW, Chen YF, Lee TT, Hsieh PY, Rusnac DV, Lin SY, Elledge SJ, Zheng N, and Yen HCS (2018). C-Terminal End-Directed Protein Elimination by CRL2 Ubiquitin Ligases. Mol. Cell 70, 602–613.e3. 10.1016/j.molcel.2018.04.006. [PubMed: 29775578]

10. Koren I, Timms RT, Kula T, Xu Q, Li MZ, and Elledge SJ (2018). The Eukaryotic Proteome Is Shaped by E3 Ubiquitin Ligases Targeting C-Terminal Degrons. Cell 173, 1622–1635.e14. 10.1016/j.cell.2018.04.028. [PubMed: 29779948]

11. Yeh CW, Huang WC, Hsu PH, Yeh KH, Wang LC, Hsu PWC, Lin HC, Chen YN, Chen SC, Yeang CH, and Yen HCS (2021). The C-degron pathway eliminates mislocalized proteins and products of deubiquitinating enzymes. EMBO J. 40, e105846. 10.15252/embj.2020105846. [PubMed: 33469951]

12. Rusnac DV, Lin HC, Canzani D, Tien KX, Hinds TR, Tsue AF, Bush MF, Yen HCS, and Zheng N (2018). Recognition of the Diglycine C-End Degron by CRL2(KLHDC2) Ubiquitin Ligase. Mol. Cell 72, 813–822.e4. 10.1016/j.molcel.2018.10.021. [PubMed: 30526872]

13. Okumura F, Fujiki Y, Oki N, Osaki K, Nishikimi A, Fukui Y, Nakatsukasa K, and Kamura T (2020). Cul5-type Ubiquitin Ligase KLHDC1 Contributes to the Elimination of Truncated SELENOS Produced by Failed UGA/Sec Decoding. iScience 23, 100970. 10.1016/j.isci.2020.100970. [PubMed: 32200094]

14. Chen X, Liao S, Makaros Y, Guo Q, Zhu Z, Krizelman R, Dahan K, Tu X, Yao X, Koren I, and Xu C (2021). Molecular basis for arginine C-terminal degron recognition by Cul2(FEM1) E3 ligase. Nat. Chem. Biol 17, 254–262. 10.1038/s41589-020-00704-3. [PubMed: 33398168]

15. 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, 263–271. 10.1038/s41589-020-00703-4. [PubMed: 33398170]

16. Ravalin M, Theofilas P, Basu K, Opoku-Nsiah KA, Assimon VA, Medina-Cleghorn D, Chen YF, Bohn MF, Arkin M, Grinberg LT, et al. (2019). Specificity for latent C termini links the E3 ubiquitin ligase CHIP to caspases. Nat. Chem. Biol 15, 786–794. 10.1038/s41589-019-0322-6. [PubMed: 31320752]

17. Ru Y, Yan X, Zhang B, Song L, Feng Q, Ye C, Zhou Z, Yang Z, Li Y, Zhang Z, et al. (2022). C-terminal glutamine acts as a C-degron targeted by E3 ubiquitin ligase TRIM7. P Natl Acad Sci USA 119, e2203218119. 10.1073/pnas.2203218119.

18. Liang X, Xiao J, Li X, Liu Y, Lu Y, Wen Y, Li Z, Che X, Ma Y, Zhang X, et al. (2022). A C-terminal glutamine recognition mechanism revealed by E3 ligase TRIM7 structures. Nat. Chem. Biol 18, 1214–1223. 10.1038/s41589-022-01128-x. [PubMed: 35982226]

19. Ichikawa S, Flaxman HA, Xu W, Vallavoju N, Lloyd HC, Wang B, Shen D, Pratt MR, and Woo CM (2022). The E3 ligase adapter cereblon targets the C-terminal cyclic imide degron. Nature 610, 775–782. 10.1038/s41586-022-05333-5. [PubMed: 36261529]

20. Scott DC, King MT, Baek K, Gee CT, Kalathur R, Li J, Purser N, Nourse A, Chai SC, Vaithiyalingam S, et al. (2023). E3 ligase autoinhibition by C-degron mimicry maintains C-degron substrate fidelity. Mol. Cell 83, 770–786.e9. 10.1016/j.molcel.2023.01.019. [PubMed: 36805027]

21. Joazeiro CAP (2019). Mechanisms and functions of ribosome-associated protein quality control. Nat. Rev. Mol. Cell Biol 20, 368–383. 10.1038/s41580-019-0118-2. [PubMed: 30940912]

22. Udagawa T, Seki M, Okuyama T, Adachi S, Natsume T, Noguchi T, Matsuzawa A, and Inada T (2021). Failure to Degrade CAT-Tailed Proteins Disrupts Neuronal Morphogenesis and Cell Survival. Cell Rep. 34, 108599. 10.1016/j.celrep.2020.108599. [PubMed: 33406423]

23. Shen PS, Park J, Qin Y, Li X, Parsawar K, Larson MH, Cox J, Cheng Y, Lambowitz AM, Weissman JS, et al. (2015). Protein synthesis. Rqc2p and 60S ribosomal subunits mediate

mRNA-independent elongation of nascent chains. Science 347, 75–78. 10.1126/science.1259724. [PubMed: 25554787]

24. Kostova KK, Hickey KL, Osuna BA, Hussmann JA, Frost A, Weinberg DE, and Weissman JS (2017). CAT-tailing as a fail-safe mechanism for efficient degradation of stalled nascent polypeptides. Science 357, 414–417. 10.1126/science.aam7787. [PubMed: 28751611]
25. Leng RP, Lin Y, Ma W, Wu H, Lemmers B, Chung S, Parant JM, Lozano G, Hakem R, and Benchimol S (2003). Pirh2, a p53-induced ubiquitin-protein ligase, promotes p53 degradation. Cell 112, 779–791. 10.1016/s0092-8674(03)00193-4. [PubMed: 12654245]
26. Sheng Y, Laister RC, Lemak A, Wu B, Tai E, Duan S, Lukin J, Sunnerhagen M, Srisailam S, Karra M, et al. (2008). Molecular basis of Pirh2-mediated p53 ubiquitylation. Nat. Struct. Mol. Biol 15, 1334–1342. 10.1038/nsmb.1521. [PubMed: 19043414]
27. Jung YS, Qian Y, Yan W, and Chen X (2013). Pirh2 E3 Ubiquitin Ligase Modulates Keratinocyte Differentiation through p63. J. Invest. Dermatol 133, 1178–1187. 10.1038/jid.2012.466. [PubMed: 23235527]
28. Wu H, Zeinab RA, Flores ER, and Leng RP (2011). Pirh2, a Ubiquitin E3 Ligase, Inhibits p73 Transcriptional Activity by Promoting Its Ubiquitination. Mol. Cancer Res 9, 1780–1790. 10.1158/1541-7786.Mcr-11-0157. [PubMed: 21994467]
29. Hakem A, Bohgaki M, Lemmers B, Tai E, Salmena L, Matysiak-Zablocki E, Jung YS, Karaskova J, Kaustov L, Duan S, et al. (2011). Role of Pirh2 in mediating the regulation of p53 and c-Myc. PLoS Genet. 7, e1002360. 10.1371/journal.pgen.1002360. [PubMed: 22125490]
30. Yamaguchi N, Sekine S, Naguro I, Sekine Y, and Ichijo H (2016). KLHDC10 Deficiency Protects Mice against TNFalpha-Induced Systemic Inflammation. PLoS One 11, e0163118. 10.1371/journal.pone.0163118. [PubMed: 27631783]
31. Burroughs AM, Iyer LM, and Aravind L (2011). Functional diversification of the RING finger and other binuclear treble clef domains in prokaryotes and the early evolution of the ubiquitin system. Mol. Biosyst 7, 2261–2277. 10.1039/c1mb05061c. [PubMed: 21547297]
32. Kaur G, Iyer LM, Subramanian S, and Aravind L (2018). Evolutionary convergence and divergence in archaeal chromosomal proteins and Chromo-like domains from bacteria and eukaryotes. Sci. Rep 8, 6196. 10.1038/s41598-018-24467-z. [PubMed: 29670199]
33. Sekine Y, Hatanaka R, Watanabe T, Sono N, Iemura S.i., Natsume T, Kuranaga E, Miura M, Takeda K, and Ichijo H (2012). The Kelch Repeat Protein KLHDC10 Regulates Oxidative Stress-Induced ASK1 Activation by Suppressing PP5. Mol. Cell 48, 692–704. 10.1016/j.molcel.2012.09.018. [PubMed: 23102700]
34. Mahrour N, Redwine WB, Florens L, Swanson SK, Martin-Brown S, Bradford WD, Staehling-Hampton K, Washburn MP, Conaway RC, and Conaway JW (2008). Characterization of Cullin-box sequences that direct recruitment of Cul2-Rbx1 and Cul5-Rbx2 modules to elongin BC-based ubiquitin ligases. J. Biol. Chem 283, 8005–8013. 10.1074/jbc.M706987200. [PubMed: 18187417]
35. Bosu DR, and Kipreos ET (2008). Cullin-RING ubiquitin ligases: global regulation and activation cycles. Cell Div. 3, 7. 10.1186/1747-1028-3-7. [PubMed: 18282298]
36. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Žídek A, Potapenko A, et al. (2021). Highly accurate protein structure prediction with AlphaFold. Nature 596, 583–589. 10.1038/s41586-021-03819-2. [PubMed: 34265844]
37. Evans R, O'Neill M, Pritzel A, Antropova N, Senior A, Green T, Žídek A, Bates R, Blackwell S, Yim J, et al. (2022). Protein complex prediction with AlphaFold-Multimer. Preprint at bioRxiv. 10.1101/2021.10.04.463034.
38. Cerullo F, Filbeck S, Patil PR, Hung HC, Xu H, Vornberger J, Hofer FW, Schmitt J, Kramer G, Bukau B, et al. (2022). Bacterial ribosome collision sensing by a MutS DNA repair ATPase paralogue. Nature 603, 509–514. 10.1038/s41586-022-04487-6. [PubMed: 35264791]
39. Sprague ER, Redd MJ, Johnson AD, and Wolberger C (2000). Structure of the C-terminal domain of Tup1, a corepressor of transcription in yeast. EMBO J. 19, 3016–3027. 10.1093/emboj/19.12.3016. [PubMed: 10856245]
40. Gratzer WB, and Doty P (1963).A Conformation Examination of Poly-L-Alanine and Poly-D,L-Alanine in Aqueous Solution. J. Am. Chem. Soc 85, 1193–1197. 10.1021/ja00891a035.

41. Hong J-Y, Wang D-D, Xue W, Yue H-W, Yang H, Jiang L-L, Wang W-N, and Hu H-Y (2019). Structural and dynamic studies reveal that the Ala-rich region of ataxin-7 initiates α-helix formation of the polyQ tract but suppresses its aggregation. Sci. Rep 9, 7481. 10.1038/s41598-019-43926-9. [PubMed: 31097749]

42. Mesecke N, Bihlmaier K, Grumbt B, Longen S, Terziyska N, Hell K, and Herrmann JM (2008). The zinc-binding protein Hot13 promotes oxidation of the mitochondrial import receptor Mia40. EMBO Rep. 9, 1107–1113. 10.1038/embor.2008.173. [PubMed: 18787558]

43. Tabaczar S, Czogalla A, Podkalicka J, Biernatowska A, and Sikorski AF (2017). Protein palmitoylation: Palmitoyltransferases and their specificity. Exp. Biol. Med 242, 1150–1157. 10.1177/1535370217707732.

44. Chang L, and Perez A (2022). AlphaFold encodes the principles to identify high affinity peptide binders. Preprint at bioRxiv. 10.1101/2022.03.18.484931.

45. Tsaban T, Varga JK, Avraham O, Ben-Aharon Z, Khramushin A, and Schueler-Furman O (2022). Harnessing protein folding neural networks for peptide-protein docking. Nat. Commun 13, 176. 10.1038/s41467-021-27838-9. [PubMed: 35013344]

46. Lytvynenko I, Paternoga H, Thrun A, Balke A, Müller TA, Chiang CH, Nagler K, Tsaprailis G, Anders S, Bischofs I, et al. (2019). Alanine Tails Signal Proteolysis in Bacterial Ribosome-Associated Quality Control. Cell 178, 76–90.e22. 10.1016/j.cell.2019.05.002. [PubMed: 31155236]

47. Bengtson MH, and Joazeiro CAP (2010). Role of a ribosome-associated E3 ubiquitin ligase in protein quality control. Nature 467, 470–473. 10.1038/nature09371. [PubMed: 20835226]

48. Brandman O, Stewart-Ornstein J, Wong D, Larson A, Williams CC, Li GW, Zhou S, King D, Shen PS, Weibezahn J, et al. (2012). A ribosome-bound quality control complex triggers degradation of nascent peptides and signals translation stress. Cell 151, 1042–1054. 10.1016/j.cell.2012.10.044. [PubMed: 23178123]

49. Moore SD, and Sauer RT (2007). The tmRNA system for translational surveillance and ribosome rescue. Annu. Rev. Biochem 76, 101–124. 10.1146/annurev.biochem.75.103004.142733. [PubMed: 17291191]

50. Osuna BA, Howard CJ, Kc S, Frost A, and Weinberg DE (2017). In vitro analysis of RQC activities provides insights into the mechanism and function of CAT tailing. Elife 6, e27949. 10.7554/eLife.27949. [PubMed: 28718767]

51. Yonashiro R, Tahara EB, Bengtson MH, Khokhrina M, Lorenz H, Chen KC, Kigoshi-Tansho Y, Savas JN, Yates JR, Kay SA, et al. (2016). The Rqc2/Tae2 subunit of the ribosome-associated quality control (RQC) complex marks ribosome-stalled nascent polypeptide chains for aggregation. Elife 5, e11794. 10.7554/eLife.11794. [PubMed: 26943317]

52. Mirdita M, Schütze K, Moriwaki Y, Heo L, Ovchinnikov S, and Steinegger M(2022). ColabFold: making protein folding accessible to all. Nat. Methods 19, 679–682. 10.1038/s41592-022-01488-1. [PubMed: 35637307]

53. Pettersen EF, Goddard TD, Huang CC, Meng EC, Couch GS, Croll TI, Morris JH, and Ferrin TE (2021). UCSF ChimeraX: Structure visualization for researchers, educators, and developers. Protein Sci. 30, 70–82. 10.1002/pro.3943. [PubMed: 32881101]

54. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Söding J, et al. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol. Syst. Biol 7, 539. 10.1038/msb.2011.75. [PubMed: 21988835]

55. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, and Lipman DJ (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25, 3389–3402. 10.1093/nar/25.17.3389. [PubMed: 9254694]

56. Deorowicz S, Debudaj-Grabysz A, and Gudyś A (2016). FAMSA: Fast and accurate multiple sequence alignment of huge protein families. Sci. Rep 6, 33964. 10.1038/srep33964. [PubMed: 27670777]

57. Katoh K, Rozewicki J, and Yamada KD (2019). MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. Briefings Bioinf. 20, 1160–1166. 10.1093/bib/bbx108.

58. Robert X, and Gouet P (2014). Deciphering key features in protein structures with the new ENDscript server. Nucleic Acids Res. 42, W320–W324. 10.1093/nar/gku316. [PubMed: 24753421]

59. Ashkenazy H, Abadi S, Martz E, Chay O, Mayrose I, Pupko T, and Ben-Tal N (2016). ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. Nucleic Acids Res. 44, W344–W350. 10.1093/nar/gkw408. [PubMed: 27166375]

60. Söding J, Biegert A, and Lupas AN (2005). The HHpred interactive server for protein homology detection and structure prediction. Nucleic Acids Res. 33, W244–W248. 10.1093/nar/gki408. [PubMed: 15980461]

61. Humphreys IR, Pei J, Baek M, Krishnakumar A, Anishchenko I, Ovchinnikov S, Zhang J, Ness TJ, Banjade S, Bagde SR, et al. (2021). Computed structures of core eukaryotic protein complexes. Science 374, eabm4805. 10.1126/science.abm4805. [PubMed: 34762488]

62. Nguyen LT, Schmidt HA, von Haeseler A, and Minh BQ (2015). IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol. Biol. Evol 32, 268–274. 10.1093/molbev/msu300. [PubMed: 25371430]

63. Price MN, Dehal PS, and Arkin AP (2010). FastTree 2--approximately maximum-likelihood trees for large alignments. PLoS One 5, e9490. 10.1371/journal.pone.0009490. [PubMed: 20224823]

64. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, and Ferrin TE (2004). UCSF Chimera–a visualization system for exploratory research and analysis. J. Comput. Chem 25, 1605–1612. 10.1002/jcc.20084. [PubMed: 15264254]

65. Altschul SF, and Koonin EV (1998). Iterated profile searches with PSI-BLAST–a tool for discovery in protein databases. Trends Biochem. Sci 23, 444–447. 10.1016/s0968-0004(98)01298-5. [PubMed: 9852764]

66. Mirdita M, Steinegger M, and Söding J (2019). MMseqs2 desktop and local web server app for fast, interactive sequence searches. Bioinformatics 35, 2856–2858. 10.1093/bioinformatics/bty1057. [PubMed: 30615063]

67. Sehnal D, Bittrich S, Deshpande M, Svobodová R, Berka K, Bazgier V, Velankar S, Burley SK, Koča J, and Rose AS (2021). Mol* Viewer: modern web app for 3D visualization and analysis of large biomolecular structures. Nucleic Acids Res. 49, W431–W437. 10.1093/nar/gkab314. [PubMed: 33956157]

Highlights

- Basis for Ala-tail C-degron recognition by Pirh2 and KLHDC10 E3 ligases
- The Ala-tail-binding site in Pirh2 has bacterial origin
- The Ala-tail-specific E3, Pirh2, was inherited from the last eukaryotic common ancestor

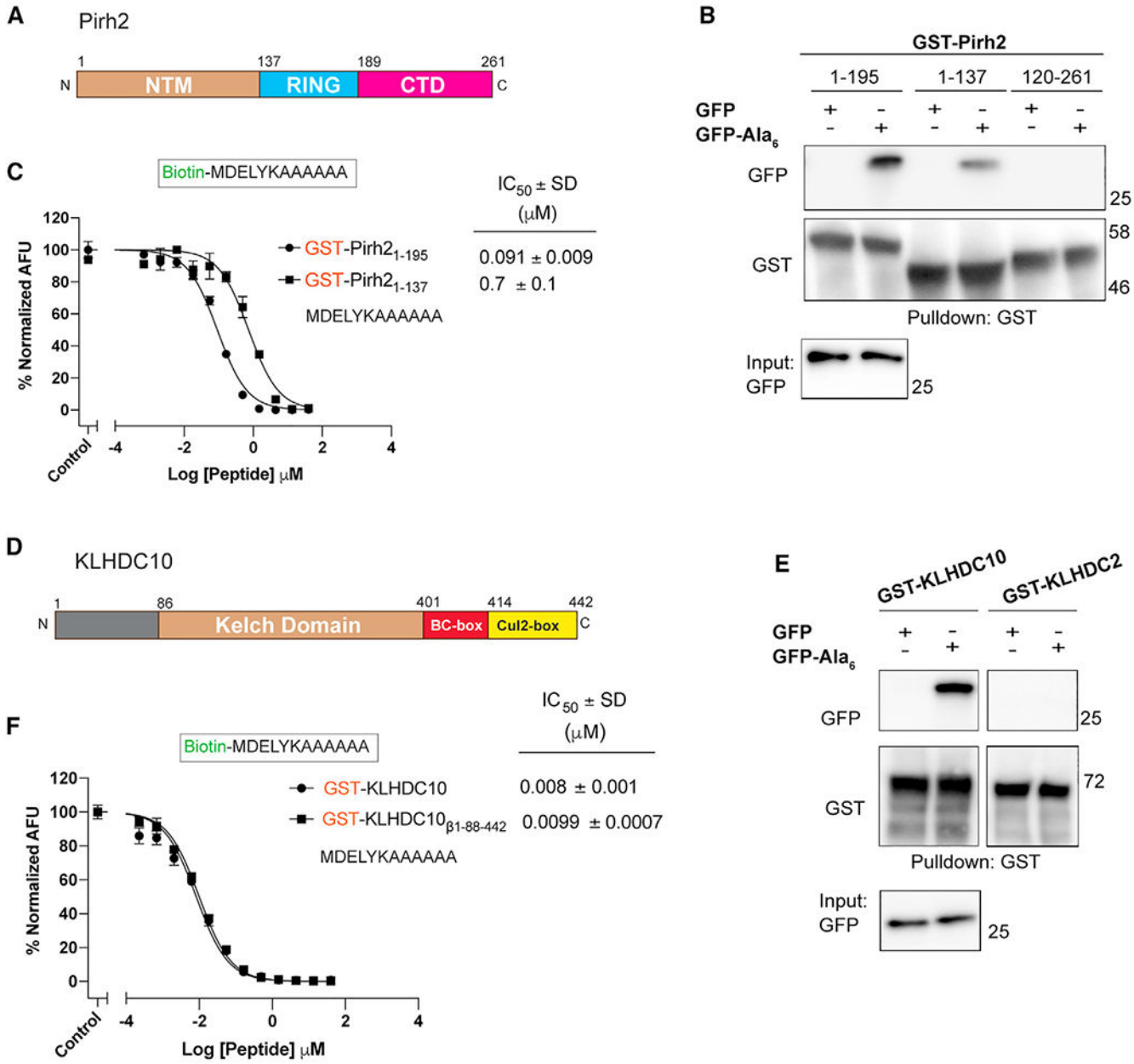


Figure 1. Pirh2 and KLHDC10 directly bind to Ala-tails

(A) Domain arrangement of human Pirh2: NTM, N-terminal module; RING, really interesting new gene domain; CTD, C-terminal domain.

(B) *In vitro* GST pull-down assay using the specified GST-Pirh2 truncation constructs and recombinantly purified GFP or GFP-Ala_6. Anti-GST and anti-GFP immunoblots are indicated.

(C) AlphaScreen assay to assess Pirh2-Ala-tail binding. Reactants were GST fusions with Pirh2 fragments containing both the NTM and RING domains ($Pirh2_{1–195}$) or the NTM alone ($Pirh2_{1–137}$) and a biotinylated Ala-tail peptide (MDELYKAAAAAA). Proximity-induced fluorescence signal (see STAR Methods) was monitored in the presence of

increasing amounts of a competing, non-biotinylated Ala-tail peptide to determine IC_{50} values from the dose-response curve. Each data point is a technical triplicate, and error bars represent ± SD. The y axis presents the normalized AlphaScreen signal as arbitrary fluorescence units (AFUs).

(D) Domain arrangement of human KLHDC10.

(E) *In vitro* GST pull-down assay using the GST-KLHDC10 and GST-KLHDC2 constructs and recombinantly purified GFP or GFP-Ala_6. Anti-GST and anti-GFP immunoblots are indicated.

(F) AlphaScreen assay to assess KLHDC10-Ala-tail binding. As in (C) but using GST-KLHDC10 and GST-$KLHDC10_{\beta 1\text{-}88–442}$ instead.

See also Figures S1, S2, S10, and S11.

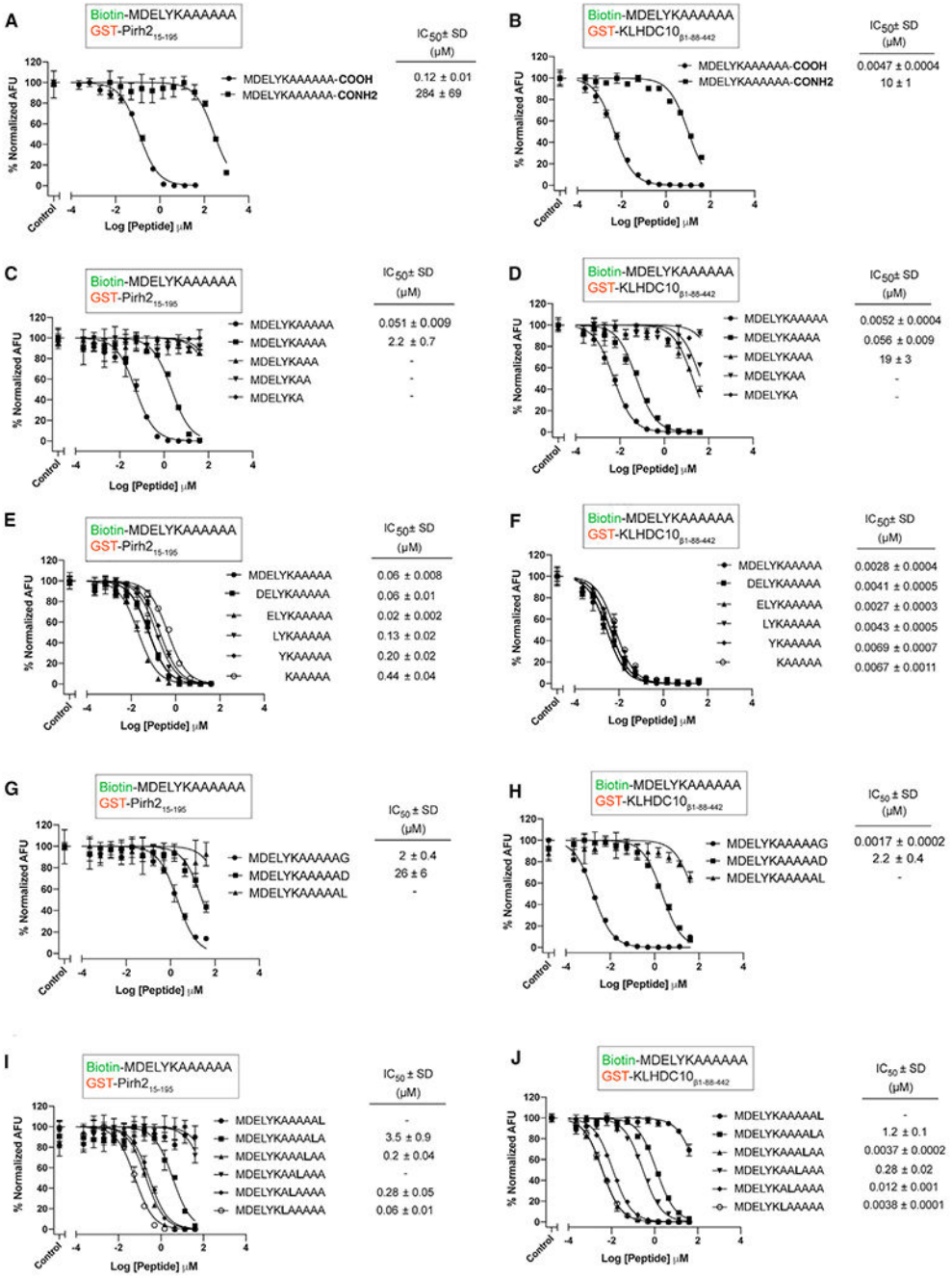


Figure 2. Four Ala residues at the C terminus and the terminal carboxyl group are essential for Ala-tail degron binding by Pirh2 and KLHDC10

(A–J) As in Figure 1C, but using GST-Pirh2$_{15\text{-}195}$ (A, C, E, G, and I), GST-KLHDC10$_{\beta 1\text{-}88\text{–}442}$ (B, D, F, H, and J), and competing peptides, as indicated, to measure the effects of C-terminal amide substitution (A and B); 1–5 Ala residues at the C terminus (C and D); shortened N-terminal sequences in Ala$_5$-containing peptides (E and F); 5 Ala followed by a C-terminal Gly, Asp, or Leu (G and H); and Leu substitution at each of the six terminal positions (I and J).

See also Figure S3. Each data point is a technical triplicate, and error bars represent ± SD.

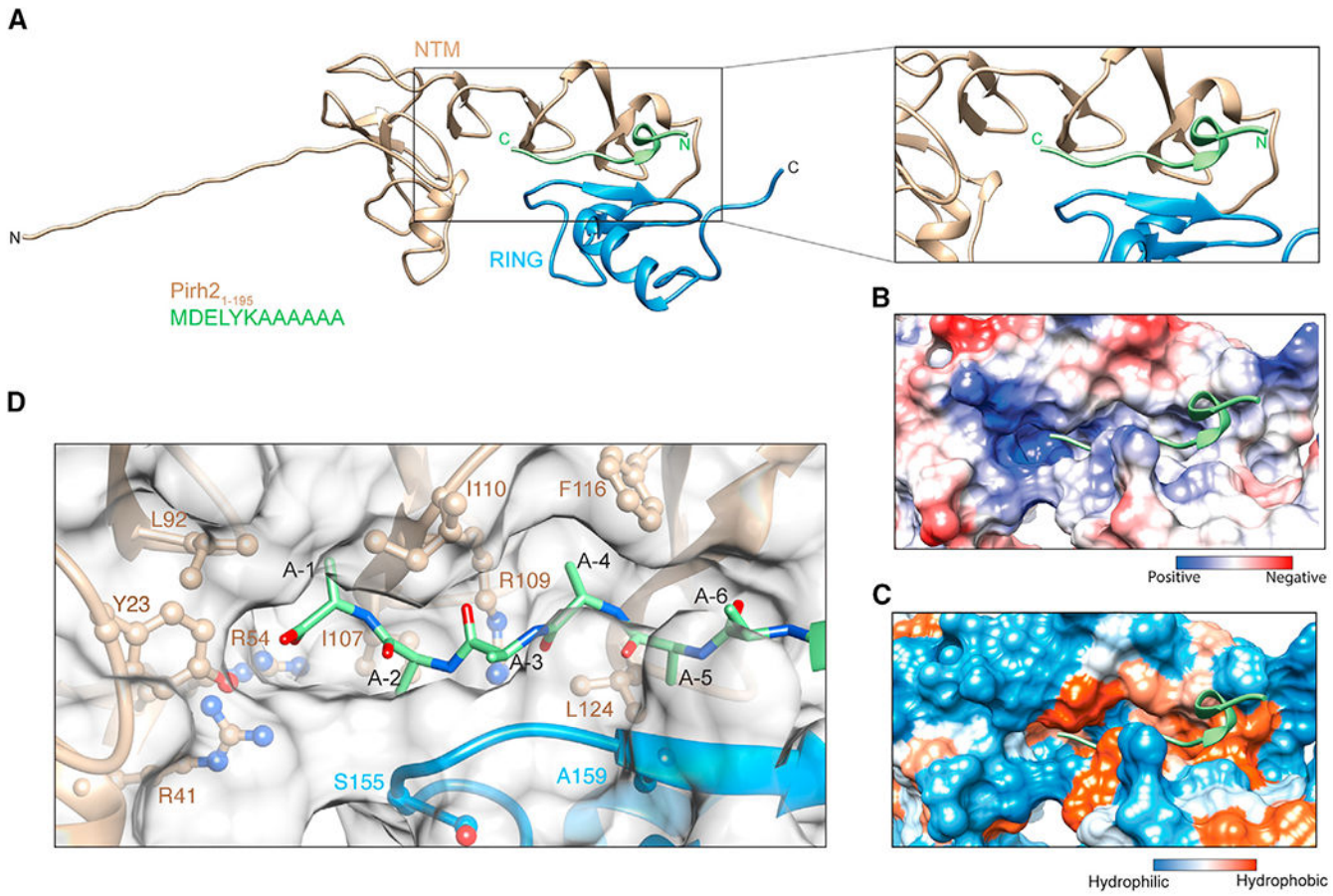


Figure 3. Identification of a candidate Ala-tail C-degron-binding site in Pirh2

(A) AlphaFold2 prediction of Pirh2$_{1–195}$ in complex with an Ala-tail peptide. The highest-ranking among predicted models is shown. Ribbon representation of Pirh2 NTM and RING domains colored according to Figure 1. The Ala-tail peptide is shown in green. The inset shows the Ala-tail degron-binding site in close-up view.

(B and C) Electrostatic surface potential (B) and surface hydrophobicity (C) represented for the focused region shown in (A); Ala-tail peptide in green.

(D) A close-up view of the Ala-tail peptide in complex with Pirh2$_{1–195}$, with the C-terminal Ala residues displayed in stick representation and colored in green. Pirh2$_{1–195}$ is represented as a transparent surface, with residues in the vicinity of the peptide labeled and sides chains in ball-stick representation.

See also Figures S4, S5, and S12.

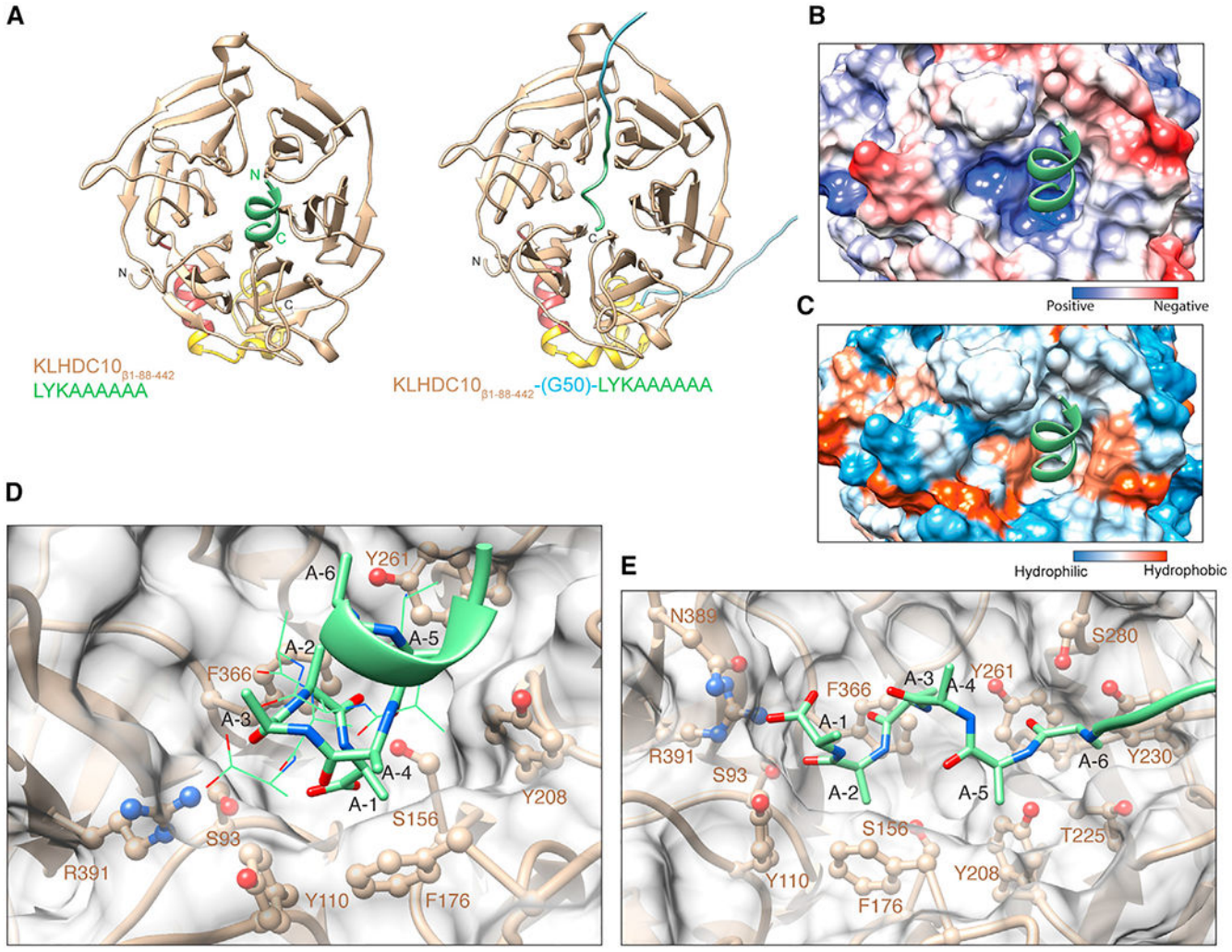


Figure 4. Identification of a candidate Ala-tail C-degron-binding site in KLHDC10

(A) As in Figure 3, but for KLHDC10$_{\beta 1\text{-}88\text{–}442}$. KLHDC10 is colored as in Figure 1D. Top predictions of KLHDC10$_{\beta 1\text{-}88\text{–}442}$ in complex with an LYKAAAAAA peptide (left) or fused to the LYKAAAAAA peptide through a Gly50 linker (right).

(B and C) Electrostatic surface potential (B) and surface hydrophobicity (C) represented for the focused region shown in (A), left; Ala-tail peptide in green.

(D) Close-up view of the left panel shown in (A). Two orientations of Ala-tail peptide (green) are presented, in stick and wire representations.

(E) Close-up of the right panel shown in (A).

See also Figures S6 and S12.

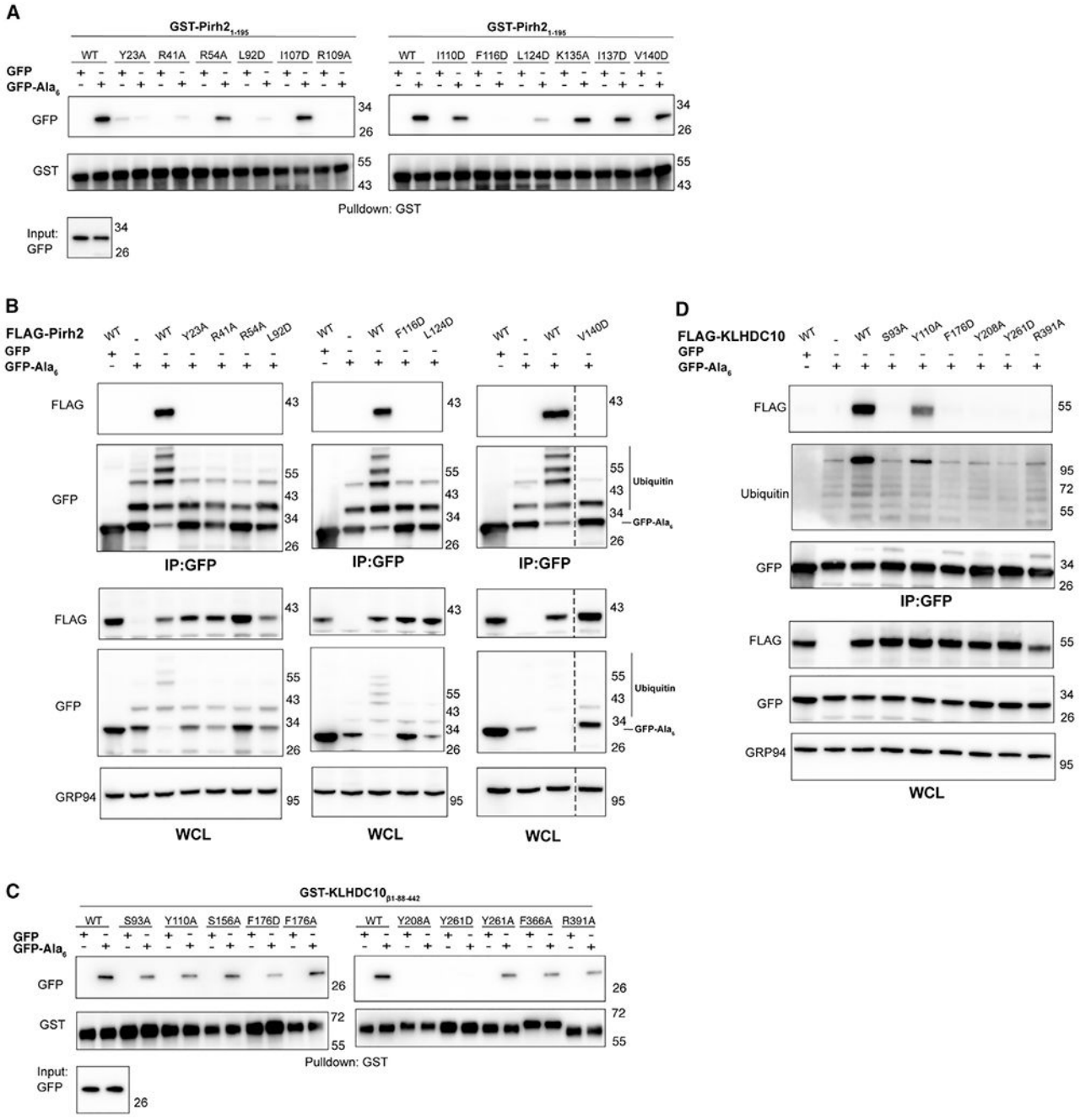


Figure 5. Validation of residues implicated in Ala-tail degron recognition by Pirh2 and KLHDC10

(A) *In vitro* GST pull-down assay using GST-Pirh2$_{1–195}$ WT and mutant constructs as indicated and recombinantly purified GFP or GFP-Ala$_6$. Anti-GST and anti-GFP immunoblots are shown.

(B) GFP or GFP-Ala$_6$ and WT or mutant full-length FLAG-Pirh2 were expressed in HeLa cells. Interaction between GFP-Ala$_6$ with Pirh2 and GFP-Ala$_6$ ubiquitylation was analyzed by GFP IP followed by anti-FLAG and anti-GFP immunoblots. The dashed line indicates that the blot image has been spliced.

(C) As in (A), but for GST-KLHDC10$_{\beta 1\text{-}88–442}$ WT and indicated mutant constructs.
(D) As in (B), but for WT and mutant full-length KLHDC10 constructs in HeLa Pirh2-knockout cells and including anti-ubiquitin blot.

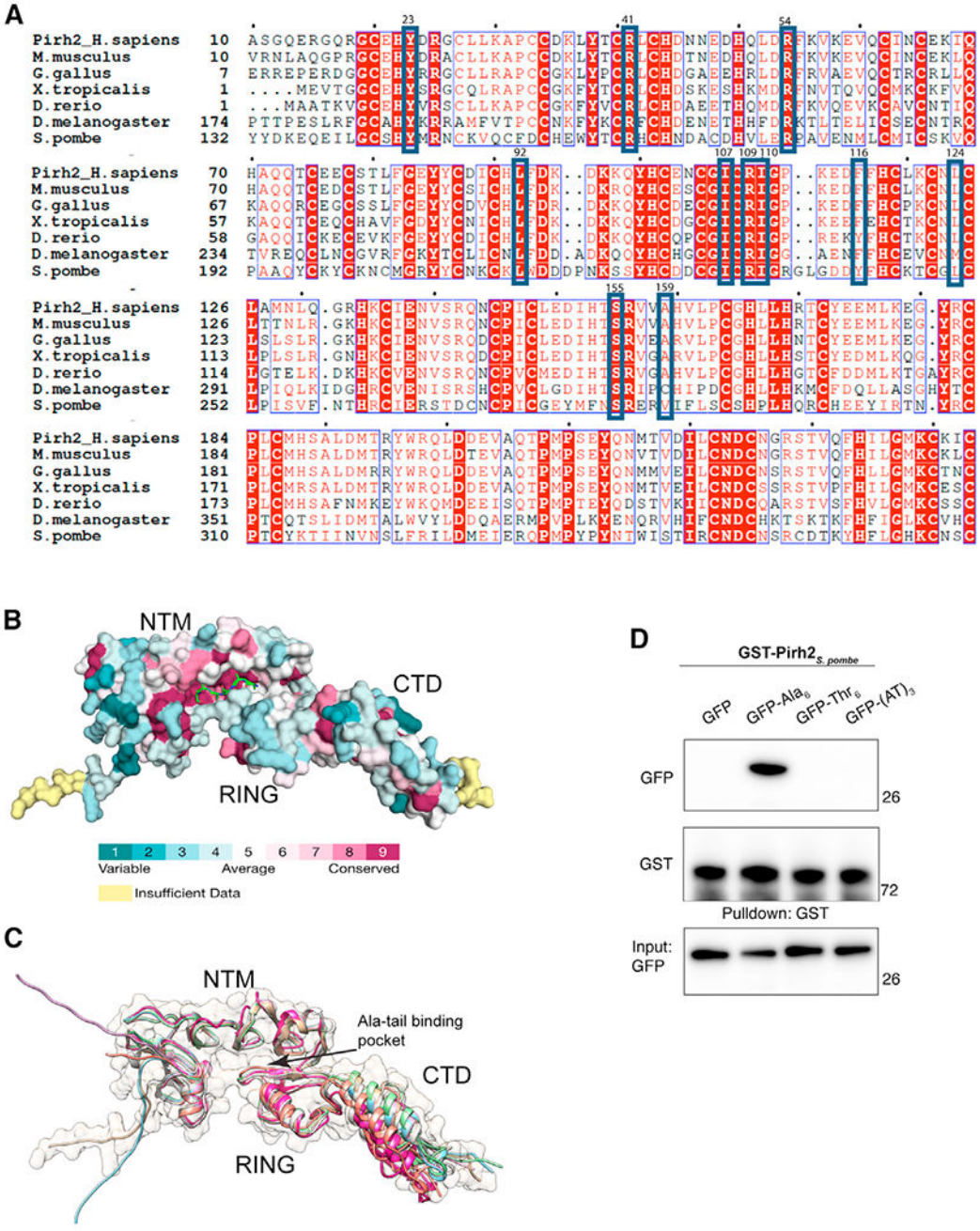


Figure 6. Conservation of residues implicated in Ala-tail binding among Pirh2 homologs

(A) Multiple sequence alignment for Pirh2 homologs. The blue boxes highlight residues predicted to stabilize the Ala-tail peptide at the binding pocket in Pirh2 that are shown in Figure 3D and corresponding residues in the homologs.

(B) Residues conserved in Pirh2 homologs mapped onto the AlphaFold structure of human Pirh2 (Uni-Prot: Q96PM5) using ConSurf webserver. The color key for variable to conserved residues is shown below. The Ala_6 peptide is shown in stick representation at the Ala-tail-binding site.

(C) AlphaFold models for Pirh2 homologs in (A) are superposed. Backbone of each protein is in ribbon representation in different colors, and human Pirh2 is represented with transparent surface.

(D) *In vitro* GST pull-down assay using GST-tagged full-length Pirh2 homolog from *S. pombe* and recombinantly purified GFP, GFP-Ala_6, GFP-Thr_6, or GFP-$(Ala\text{-}Thr)_3$. Anti-GST and anti-GFP immunoblots are indicated.

See also Figures S10 and S11.

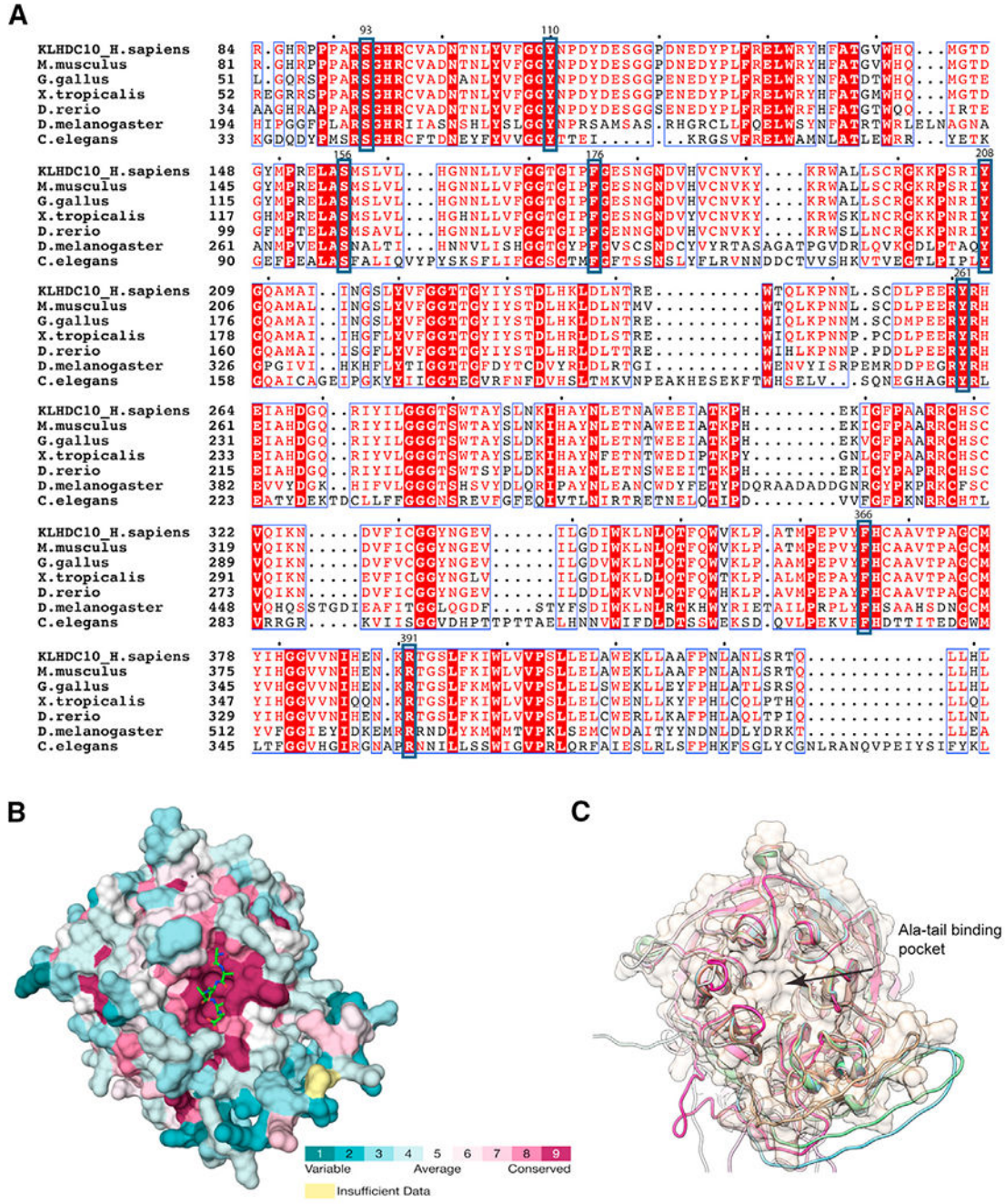


Figure 7. Conservation of residues implicated in Ala-tail binding among KLHDC10 homologs

(A–C) As in Figure 6, but for the KLHDC10 Kelch domain. In (A), blue boxes highlight residues implicated in Ala-tail binding as shown in Figure 4D.

See also Figures S7–S9.

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Mouse monoclonal anti-GFP	Roche	Cat# 11814460001; RRID: AB_390913
Goat polyclonal anti-GST	GE healthcare	Cat# 27–4577; RRID: AB_771432
Mouse monoclonal anti-FLAG M2	Sigma-Aldrich	Cat# F1804; RRID: AB_262044
Rabbit polyclonal anti-GRP94	Cell Signaling Technology	Cat# 2104: RRID: AB_823506
Rabbit polyclonal anti-Pirh2/Rchy1	Abcam	Cat# ab189247
Goat polyclonal anti-mouse IgG (H + L)-HRP	Dianova	Cat# 115-035-146; RRID: AB_2307392
Goat polyclonal anti-rabbit IgG (H + L)-HRP D	Dianova	Cat# 11-035-144; RRID: AB_2307391
Rabbit polyclonal anti-goat IgG (H + L)-HRP	Dianova	Cat# 305-035-045; RRID: AB_2339403
Bacterial and virus strains		
E. coli BL21 (DE3)	E. Schiebel lab	N/A
E. coli DH5aplha	E. Schiebel lab	N/A
Chemicals, peptides, and recombinant proteins		
Protino Glutathione Agarose 4B	Macherey-Nagel	Cat# 745500.10
EDTA free protease inhibitor mixture	Roche	Cat# 04693159001
Glutathione, reduced	Sigma-Aldrich	Cat# Y0000517
Lipofectamine 3000	Invitrogen	Cat# L3000015
GFP-Trap® magnetic agarose beads	ChromoTek	Cat# gtma-20
MG132	Sigma-Aldrich	Cat# C2211-5MG
Biotin-MDELYKAAAAAA	Biosyntan	N/A
MDELYKAAAAAA	Genscript	N/A
MDELYKAAAAAA-NH2	Genscript	N/A
MDELYKAAAAA	Genscript	N/A
MDELYKAAAA	Genscript	N/A
MDELYKAAA	Genscript	N/A
MDELYKAA	Genscript	N/A
MDELYKA	Genscript	N/A
DELYKAAAAA	Genscript	N/A
ELYKAAAAA	Genscript	N/A
LYKAAAAA	Genscript	N/A
YKAAAAA	Genscript	N/A
KAAAAA	Genscript	N/A
MDELYKAAAAAG	Genscript	N/A
MDELYKAAAAAD	Genscript	N/A
MDELYKAAAAAL	Genscript	N/A

REAGENT or RESOURCE	SOURCE	IDENTIFIER
MDELYKLAAAAA	Genscript	N/A
MDELYKALAAAA	Genscript	N/A
MDELYKAALAAA	Genscript	N/A
MDELYKAAALAA	Genscript	N/A
MDELYKAAAALA	Genscript	N/A
MDELAAAAA	Genscript	N/A
MDELGAAAA	Genscript	N/A
MDELGGAAA	Genscript	N/A
MDELGGGAA	Genscript	N/A
MDELGGGGA	Genscript	N/A
MDELYKGAAAAA	Genscript	N/A
MDELYKAGAAAA	Genscript	N/A
MDELYKAAGAAA	Genscript	N/A
MDELYKAAAGAA	Genscript	N/A
MDELYKAAAAGA	Genscript	N/A
MDELYKTAAAAA	Genscript	N/A
MDELYKATAAAA	Genscript	N/A
MDELYKAATAAA	Genscript	N/A
MDELYKAAATAA	Genscript	N/A
MDELYKAAAATA	Genscript	N/A
MDELYKAAAAAT	Genscript	N/A
Critical commercial assays		
AlphaScreen GST detection kit	PerkinElmer	Cat# 6760603C
Direct PCR Lysis Reagent	Viagen	Cat# 301-C
TOPO™ TA Cloning	Invitrogen	Cat# 450030
ExpressPlus™ PAGE Gel, 4–20%	GenScript	Cat# M42010 Cat# M42012 Cat# M42015
NuPAGE™ 4–12%, Bis-Tris	Thermo Fisher	Cat# NP0321BOX
Color Prestained Protein Standard, Broad Range (10–250 kDa)	New England Biolabs	Cat# P7719
Experimental models: Cell lines		
HeLa	Joazeiro Lab	N/A
HeLa PKO	This paper	N/A
Oligonucleotides		
Primers for PCR, see Table S1	This paper	N/A
sgRNA sequence targeting Pirh2 (Exon 2) Fw 5‘ CACCGACTAGATCGCTTTAAAGTGA 3’	This paper	N/A
sgRNA sequence targeting Pirh2 (Exon 2) Rv 3‘ CTGATCTAGCGAAATTTCACTCAAA 5’	This paper	N/A
Recombinant DNA		
pET15b GST-Pirh2	Thrun et al.[8]	N/A

REAGENT or RESOURCE	SOURCE	IDENTIFIER
pET15b GST-KLHDC10	Thrun et al.[8]	N/A
pET15b GST-KLHDC2	This paper	N/A
PET15b GST-KLHDC10$_{\beta 1\text{-}88\text{-}442}$	This paper	N/A
pETB28b Elongin B (a.a. 1–104)/Elongin C (a.a. 1–112)	Thrun et al.[8]	N/A
pET15b GST-Pirh2$_{15\text{-}195}$	This paper	N/A
pET15b GST-Pirh2$_{1\text{-}195}$ Y23A	This paper	N/A
pET15b GST-Pirh2$_{1\text{-}195}$ R41A	This paper	N/A
pET15b GST-Pirh2$_{1\text{-}195}$ R54A	This paper	N/A
pET15b GST-Pirh2$_{1\text{-}195}$ L92D	This paper	N/A
pET15b GST-Pirh2$_{1\text{-}195}$ I107D	This paper	N/A
pET15b GST-Pirh2$_{1\text{-}195}$ R109A	This paper	N/A
pET15b GST-Pirh2$_{1\text{-}195}$ I110D	This paper	N/A
pET15b GST-Pirh2$_{1\text{-}195}$ F116D	This paper	N/A
pET15b GST-Pirh2$_{1\text{-}195}$ L124D	This paper	N/A
pET15b GST-Pirh2$_{1\text{-}195}$ K135A	This paper	N/A
pET15b GST-Pirh2$_{1\text{-}195}$ I137D	This paper	N/A
pET15b GST-Pirh2$_{1\text{-}195}$ V140D	This paper	N/A
PET15b GST-KLHDC10$_{\beta 1\text{-}88\text{-}442}$ S93A	This paper	N/A
PET15b GST-KLHDC10$_{\beta 1\text{-}88\text{-}442}$ Y110A	This paper	N/A
PET15b GST-KLHDC10$_{\beta 1\text{-}88\text{-}442}$ S156A	This paper	N/A
PET15b GST-KLHDC10$_{\beta 1\text{-}88\text{-}442}$ F176D	This paper	N/A
PET15b GST-KLHDC10$_{\beta 1\text{-}88\text{-}442}$ F176A	This paper	N/A
PET15b GST-KLHDC10$_{\beta 1\text{-}88\text{-}442}$ Y208A	This paper	N/A
PET15b GST-KLHDC10$_{\beta 1\text{-}88\text{-}442}$ Y261D	This paper	N/A
PET15b GST-KLHDC10$_{\beta 1\text{-}88\text{-}442}$ Y261A	This paper	N/A
PET15b GST-KLHDC10$_{\beta 1\text{-}88\text{-}442}$ F366A	This paper	N/A
PET15b GST-KLHDC10$_{\beta 1\text{-}88\text{-}442}$ R391A	This paper	N/A
pCDNA5 3XFLAG-Pirh2	Thrun et al.[8]	N/A
pCDNA5 3XFLAG-Pirh2 Y23A	This paper	N/A
pCDNA5 3XFLAG-Pirh2 R41A	This paper	N/A
pCDNA5 3XFLAG-Pirh2 R54A	This paper	N/A
pCDNA5 3XFLAG-Pirh2 L92D	This paper	N/A
pCDNA5 3XFLAG-Pirh2 F116D	This paper	N/A
pCDNA5 3XFLAG-Pirh2 L124D	This paper	N/A
pCDNA5 3XFLAG-Pirh2 I137D	This paper	N/A
pCDNA5 3XFLAG-Pirh2 V140D	This paper	N/A
pCDNA5 3XFLAG-KLHDC10	Thrun et al.[8]	N/A

REAGENT or RESOURCE	SOURCE	IDENTIFIER
pCDNA5 3XFLAG-KLHDC10 S93A	This paper	N/A
pCDNA5 3XFLAG-KLHDC10 Y110A	This paper	N/A
pCDNA5 3XFLAG-KLHDC10 F176D	This paper	N/A
pCDNA5 3XFLAG-KLHDC10 Y208A	This paper	N/A
pCDNA5 3XFLAG-KLHDC10 Y261D	This paper	N/A
pCDNA5 3XFLAG-KLHDC10 R391A	This paper	N/A
pEGFP	Joazeiro lab	N/A
pEGFP-Ala_6	Thrun et al.[8]	N/A
Software and algorithms		
Prism8 for Mac OS X	Graphpad	https://www.graphpad.com
ImageJ	N/A	https://ImageJ.nih.gov/ij
ColabFold AlphaFold 2	Mirdita et al.[52]	https://colab.research.google.com/github/sokrypton/ColabFold/blob/main/AlphaFold2.ipynb
UCSF ChimeraX 1.4 and 1.5	Pettersen et al.[53]	RRID: SCR_015872
Clustal omega	Sievers et al.[54]	RRID: SCR_001591
PSI-BLAST	Altschul et al.[55]	RRID: SCR_001010
FAMSA	Deorowicz et al.[56]	RRID: SCR_021804
MAFFT	Katoh et al.[57]	RRID: SCR_011811
ESPript 3.0	Robert et al.[58]	https://espript.ibcp.fr/ESPript/ESPript/
ConSurf Webserver	Ashkenazy et al.[59]	https://consurf.tau.ac.il/consurf_index.php
HHpred	Soding et al.[60]	RRID: SCR_010276
RoseTTa Fold	Humphreys et al.[61]	N/A
IQ-TREE	Nguyen et al.[62]	RRID: SCR_017254
FastTree	Price et al.[63]	RRID: SCR_015501
FigTree	tree.bio.ed.ac.uk	RRID: SCR_008515
Adobe Ilustrator	Adobe	RRID: SCR_002798