

Degron motif definition: F{D,E}D{R,K}
Possible matches: FDDR, FDDK, FEDR, FEDK
Motif in the query: FDDR

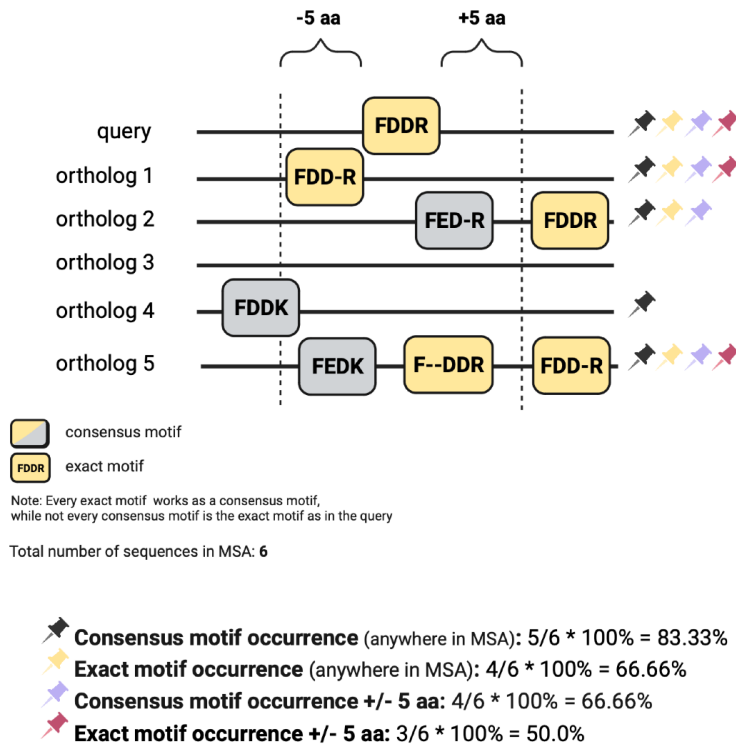


Figure S1. Four different conservation scores calculated by DEGRONOPEDIA. The server calculates the fraction of occurrences of the exact degron motif present in the query anywhere in the aligned sequences, as well as +/- 5 aa adjacent to the degron motif and the consensus motif (degrons are usually reported as a combination of consecutive amino acids), anywhere in the aligned sequences, and +/- 5 aa adjacent to the degron motif. The indicated threshold of +/- 5 aa in the query can be adjusted (see Customizable parameters in Materials and Methods).

Table S1. Pre-calculated MSAs of predicted orthologs at various taxonomic levels obtained from the eggNOG5 database available for selected model organisms. NCBI taxonomy identifiers are given in parentheses.

Species	Orthologous Group 1	Orthologous Group 2	Orthologous Group 3	Orthologous Group 4	Orthologous Group 5	Orthologous Group 6
<i>H. sapiens</i>	Hominidae (9604)	Euarchontoglires (314146)	Mammalia (40674)	Vertebrata (7742)	Opisthokonta (33154)	Eukaryota (2759)
<i>M. musculus</i>	Rodentia (9989)	Euarchontoglires (314146)	Mammalia (40674)	Vertebrata (7742)	Opisthokonta (33154)	Eukaryota (2759)
<i>D. rerio</i>	Actinopterygii (7898)			Vertebrata (7742)	Opisthokonta (33154)	Eukaryota (2759)
<i>D. melanogaster</i>	Drosophilidae (7214)	Diptera (7147)	Insecta (50557)		Opisthokonta (33154)	Eukaryota (2759)
<i>C. elegans</i>	Rhabditida (6236)				Opisthokonta (33154)	Eukaryota (2759)
<i>S. cerevisiae</i>	Saccharomycetaceae (4893)	Saccharomycetes (4891)	Ascomycota (4890)	Fungi (4751)	Opisthokonta (33154)	Eukaryota (2759)
<i>A. thaliana</i>	Brassicales (3699)		Viridiplantae (33090)			Eukaryota (2759)
<i>S. pombe</i>	Taphrinomycotina (451866)	Saccharomycetes (4891)	Ascomycota (4890)	Fungi (4751)	Opisthokonta (33154)	Eukaryota (2759)
<i>O. sativa</i>	Poales (38820)		Viridiplantae (33090)			Eukaryota (2759)
<i>Z. mays</i>	Poales (38820)		Viridiplantae (33090)			Eukaryota (2759)
<i>R. norvegicus</i>	Rodentia (9989)	Euarchontoglires (314146)	Mammalia (40674)	Vertebrata (7742)	Opisthokonta (33154)	Eukaryota (2759)

Table S2. Description of the user-customizable parameters available in DEGRONOPEDIA.

Parameter	Unit	Default value	Allowed values	Description
Primary degron-related				
1. Degron flanking region in sequence	aa	20	5-40	The maximum sequence distance to regions upstream and downstream of the degron motif to be considered as flanking
2. Degron flanking region in the structure	Å	20	5-40	The maximum structural distance to residues around the degron motif to be considered as flanking (such residues are not necessarily close in sequence to the degron motif)
3. Region length to calculate degron disorder	aa	10	1-20	The maximum sequence distance to regions upstream and downstream of the degron motif to be included in the degron mean disorder score
Secondary degron-related				
4. Region length to calculate secondary degron (K/C/T/S) ¹ disorder	aa	3	1-15	The maximum sequence distance to regions upstream and downstream of the secondary degron (K/C/T/S) to be included in the secondary degron mean disorder score
Tertiary degron-related				
5. Minimum IDR distance from the secondary degron (K/C/T/S)	aa	10	5-40	The minimum sequence distance from the secondary degron (K/C/T/S) to the continuous IDR of defined length (see parameter 6) to consider it as a tertiary degron
Structure-related				
6. Minimum continuous IDR length	aa	10	5-40	The minimum number of subsequent (in sequence) disordered residues to be considered as IDR

7. pLDDT/LDDT disorder threshold	%	70	40-90	The threshold below which the residue is considered as disordered based on its confidence pLDDT/LDDT score; the structure for which its disorder is to be predicted based on the pLDDT/LDDT scores must be either an AlphaFold2 (default when querying by UniProt ID) or RoseTTAFold model
8. IUPred3 disorder threshold	%	50	40-90	The threshold above which the residue is considered disordered based on predictions obtained from the IUPred3 software (sequence-based predictions)
9. Buried residue threshold	%	20	5-60	The threshold below which the residue is considered as buried based on its RSA
Multiple sequence alignment (MSA)-related				
10. Maximum distance from the degron motif in the query to consider the same degron motif from the aligned sequences as conserved	aa	5	5-40	The maximum distance to regions upstream and downstream of the degron motif (its ends) in the query to its aligned sequences in the MSA to consider it as evolutionarily conserved

¹ lysine/cysteine/threonine/serine residues

Table S3. Comparison of computation times and different types of queries with varying input protein lengths. All calculations included machine learning predictions of N-/C-terminus stability and disorder predictions based on IUPred3 software; all other settings remained default and no optional parameters were provided. Queries by structure were run on structures corresponding to the indicated UniProt identifiers (in parentheses) obtained from the AlphaFold Protein Structure Database v.4.

Protein	Query by UniProt ID		Query by sequence	Query by structure
	Degron conservation ✗	Degron conservation ¹ ✓		
p53 – 393 aa (P04637)	~ 25 s	~ 88 s	~ 25 s	~ 27 s
mTOR – 2549 aa (P42345)	~ 42 s	~ 180 s	~ 42 s	~ 45 s
AHNAK – 5890 aa (Q09666) ²	~ 90 s	~ 145 s	~ 78 s	~ 90 s

¹ When considering degnon conservation, calculation time depends on the number of degnon motifs found in the query, for which conservation is calculated based on the available Multiple Sequence Alignments (MSAs; also with varying number of member sequences). The conservation scores for p53 were calculated based on six different MSAs, while those for mTOR and AHNAK were calculated based on four.

² Since this protein sequence exceeded 2700 amino acids, an in-house script was used to merge the AlphaFold2 model fragments into a single PDB file.

Table S4. Comparison of functionalities of DEGRONOPEDIA and other online methods for degnon analysis.

	APC/C degnon repository	ELM prediction tool	DegPred	DegronMD	DEGRONOPEDIA
Detected/predicted degnons	ABBA, D box and KEN box degnons only	32 degnon motifs (N-/C- or internal motifs)	Predicts/maps degnons based on 303 degnon instances	Predicts/maps degnons based on 23 degnon motifs (133 degnon instances)	Maps N-/C- and internal degnons based on over 1800 degnons and predicts N-/C-terminal degnons
Screen by custom degnon motifs	Yes	No	No	No	Yes
N-/C-terminal stability	No	No	No	No	Yes
Analysis of degnons emergence following proteolysis	No	No	No	No	Yes
Degron tripartite model	No	No	No	No	Yes
Query by protein name/UniProt ID	Yes	Yes	Yes	Yes	Yes
Query by protein sequence	No	Yes	No	Yes	Yes

Query by protein structure	No	No	No	No	Yes
Query by species other than human	Yes	Yes	No	No	Yes
Adjust parameters for degron screen	No*	No	No	No	Yes
MSA of homologs	Yes	Yes	No	Yes	Yes
Analyze own MSA	No	No	No	No	Yes
Degron conservation scores	No	Yes	No	Yes	Yes
Post-translational modifications	Yes	Limited to phosphorylation	No	Yes	Yes
Mutations related to human disease	Yes	No	No	Yes	Yes
Structural data	Yes	Yes	Yes	Yes	Yes
Interacting E3 ubiquitin ligases	No	No	Yes	Yes	Yes

* Cut-offs for disorder score and flank length may be adjusted when screening a proteome