

Introduction to UniProtKB/Swiss-Prot

UniProt (Universal Protein Resource; <http://www.uniprot.org>) provides a central resource on protein sequences and functional annotation. The UniProt Consortium is comprised of the European Bioinformatics Institute (EBI), the Swiss Institute of Bioinformatics (SIB) and the Protein Information Resource (PIR). The UniProt Knowledgebase (UniProtKB) contains the manually annotated UniProtKB/Swiss-Prot section and the automatically annotated UniProtKB/TrEMBL section.

UniProtKB/TrEMBL
Computer-annotated protein sequences
Sequences with 100% identity from the same organism are merged

UniProtKB/Swiss-Prot
Manually annotated protein sequences
Non-redundant and high level of accuracy

The UniProt Consortium liaises with many of the research communities that concentrate on model organisms to ensure that we are supplying the best service.

Useful Features of UniProtKB/Swiss-Prot

- Manual referenced annotation
- Extensive cross-linking to other databases
- Feature table showing domains/sites of interest with illustrations
- Wide range of bioinformatics tools available for analysis

The screenshot shows the UniProtKB/Swiss-Prot entry for P30429 (CED-4_CAEEL). Key sections highlighted include:

- Search and Navigation:** Search bar, filters (Search, Blast, Align, Retrieve, ID Mapping).
- Entry Information:** Entry name (CED4_CAEEL), Accession (P30429), Entry history, and Entry status.
- Names and origin:** Protein names (Cell death protein-4), Gene names (ced-4), and Organism (Caenorhabditis elegans).
- Function:** Isoform a plays a major role in programmed cell death (PCD, apoptosis). Egr-1 binds to and directly initiates the activity of ced-3, releasing the cell death activator ced-4 from a cell-specific 4 containing protein complex and allowing ced-3 to activate the cell-killing caspase ced-3. Isoform b prevents PCD.
- Subunit structure:** Interacts with ced-3.
- Subcellular location:** Mitochondrion. Note: Isoform b is not cell death inducible. Ced-3 is required for mitochondrial localization. Perinuclear in cell death induced cells.
- Developmental stage:** Most abundant during embryogenesis and is also detected at later stages during periods of extensive programmed cell death.
- Miscellaneous:** Mutants without a stop codon almost all programmed cell deaths that normally occur during development.
- Sequence similarities:** Contains 1 CARD domain. Contains 1 NB-ARC domain.
- Ontologies:** Biological process (Apoptosis, Mitochondrion), Cellular component (Alternative splicing, ATR-binding, Nucleosome-binding), Ligand, Technical term (3D-structure, Complete proteome).
- Gene Ontology (GO):** Biological process (Positive regulation of apoptosis, Regulation of caspase activity), Cellular component (Cytosol, Membrane fraction, Perinuclear region, Penicillium region), Molecular function (BHL domain binding, BHL domain binding).
- Binary interactions:** Table showing interactions with BCL2L1, CASP8, ced-3, and ced-9.
- Alternative products:** This entry describes 2 isoforms produced by alternative splicing (A[sp]).
- References:** List of 6 references related to the protein's function and structure.
- Sequence annotation (Features):** Feature key (Chain), Position (1-571), Length (571), Description (Cell death protein-4). Includes a domain diagram showing CARD and NB-ARC domains.
- Sequences:** Table showing sequence details for Isoform a (Long) and Isoform a (Short).
- Relevant documents:** List of documents related to the protein.

Anatomy of a UniProtKB/Swiss-Prot entry

The detailed view shows the following sections:

- Cross-references:** Links to other databases like ENEM, PIR, UniGene, PDB, ModBase, and various protein interaction databases (DIP, IntAct, WormPep, HOENOM).
- Sequence annotation (Features):** Detailed view of the Chain feature (1-571) and domain diagrams for CARD and NB-ARC domains.
- Sequences:** Detailed view of the Isoform a (Long) sequence, showing the amino acid sequence and its length (571) and mass (65,336).
- Relevant documents:** List of documents related to the protein.

Acknowledgements

UniProt is mainly supported by the National Institutes of Health (NIH) grant 2 U01 HG027112-04. Additional support for the EBI's involvement in UniProt comes from the European Commission contract FELICS (021902) and from the NIH grant 5 P41 HG02273-06. UniProtKB/Swiss-Prot activities at the SIB are supported by the Swiss Federal Government through the Federal Office of Education and Science. PIR activities are also supported by the NIH grants for NIAID proteomic resource (HHSN26620040061C) and grid enablement (NCI-caBIG-ICR), and National Science Foundation grants for protein ontology (ITR-0205470) and BioTagger (IIS-0430743).