Enzymes in UniProt

Webinar
1st May 2019
www.uniprot.org

Universal Protein Resource which contains protein sequences for > 800’000 species together with functional annotation data.

The mission of UniProt is to provide the scientific community with a comprehensive, high-quality and freely accessible resource of protein sequence and functional information.
The two datasets in UniProtKB

**UniProtKB**
- Swiss-Prot (559,634)
  - Manually annotated and reviewed.
- TrEMBL (147,413,762)
  - Automatically annotated and not reviewed.

**UniProt release 2019_03**

- curated by experts
- data from scientific papers
- annotation of sequence features
- collates isoforms in one entry

- annotation from rule systems (incl. expert-curated rules)
- mapped experimental sequence features (3D structures)
- isoforms are kept separate
UniProtKB - Q9NUN7 (ACER3_HUMAN)

Stable and unique accession identifier

Field menu

Function


Catalytic activity

- H₂O + N-acyl-(4R)-hydroxysphinganine = 4-hydroxysphinganine + a fatty acid  ID: 2 Publications
  This reaction proceeds in the forward direction.
  Source: Rhea, Show

- H₂O + N-(5Z,8Z,11Z,14Z)-eicosatetraenoyle-4-sphing-4-ene = (5Z,8Z,11Z,14Z)-eicosatetraenoate + sphing-4-ene  ID: 1 Publication
  This reaction proceeds in the forward direction.
  Source: Rhea, Show

  This reaction proceeds in the forward direction.
  Source: Rhea, Show
- Catalytic activity: reaction description
- Cofactor(s)
- Activity regulation mechanisms
- Kinetic parameters
- Pathways
- Active site(s) position
- Cofactor and substrates binding sites

- Protein family
- Catalytic domain position
Catalytic activity

• Description of the reaction(s)
  • Multiples substrates (best substrates are selected for “catalytic activity”, others added in “function”)

• Physiological reactions

• Exceptions:
  • Drug metabolism reactions
  • reaction of biotechnological importance (for industrial/pharmaceutical purposes…)
  • Example: \textit{B. megaterium} cytochrome P450 (D5E3H2)
Catalytic activity

Alkaline ceramidase (Q9NUN7)

\[ \text{an } N\text{-acylsphing-4-enine} + \text{H}_2\text{O} = \text{a fatty acid} + \text{sphing-4-enine} \]

EC:3.5.1.23

This reaction proceeds in the forward direction.

Source: Rhea. « Hide
Catalytic activity-EC number

- 4 digit code => Classification and description of enzymatic reactions

<table>
<thead>
<tr>
<th>ID</th>
<th>3.5.3.12</th>
</tr>
</thead>
<tbody>
<tr>
<td>DE</td>
<td>Agmatine deiminase.</td>
</tr>
<tr>
<td>CA</td>
<td>Agmatine + H(2)O = N-carbamoylputrescine + NH(3).</td>
</tr>
</tbody>
</table>

- Not all reactions have associated EC number (→ partial EC ex.3.2.-.-) |
- No information on the directionality of the reaction

1.-.-.- Oxidoreductases |
2.-.-.- Transferases |
3.-.-.- Hydrolases |
4.-.-.- Lyases |
5.-.-.- Isomerases |
6.-.-.- Ligases |
7.-.-.- Translocases
Rhea results for: CHEBI:29748
Number of hits: 11

Information:
The query resolved to the following ChEBI compounds (by name): none.

<table>
<thead>
<tr>
<th>Accession</th>
<th>Equation / Hits details</th>
<th>Other directions</th>
</tr>
</thead>
<tbody>
<tr>
<td>RHEA:40051</td>
<td>chorismate = 3-[(1-carboxyvinyl)-oxy]benzoate + H₂O</td>
<td>RHEA:40052</td>
</tr>
<tr>
<td>RHEA:25512</td>
<td>(2S)-2-amino-4-deoxychorismate + L-glutamate = chorismate + L-glutamine</td>
<td>RHEA:25513</td>
</tr>
<tr>
<td>RHEA:25514</td>
<td></td>
<td>RHEA:25515</td>
</tr>
<tr>
<td>RHEA:13899</td>
<td></td>
<td>RHEA:13890</td>
</tr>
<tr>
<td>RHEA:13890</td>
<td></td>
<td>RHEA:13899</td>
</tr>
<tr>
<td>RHEA:33095</td>
<td>chorismate + inosine + phosphoenolpyruvate = futasoline</td>
<td>RHEA:33096</td>
</tr>
<tr>
<td>RHEA:33096</td>
<td></td>
<td>RHEA:33098</td>
</tr>
<tr>
<td>RHEA:33098</td>
<td></td>
<td>RHEA:33097</td>
</tr>
</tbody>
</table>

Stable identifier

ChEBI

RHEA:40051 (APPROVED)
chorismate + 3-[(1-carboxyvinyl)-oxy]benzoate + H₂O

EMBL-EBI
Catalytic activity

Alkaline ceramidase (Q9NUN7)

\[ \text{an } N\text{-acylsphing-4-enine} + \text{H}_2\text{O} = \text{a fatty acid} + \text{sphing-4-enine} \]

EC:3.5.1.23

This reaction proceeds in the forward direction.

Source: Rhea. « Hide
Catalytic activity-text format

- text reactions linked to complete EC number
- Examples: Proteases, DNA topoisomerase

**Cathepsin B (P07858)**

* Catalytic activity: 1
  - Hydrolysis of proteins with broad specificity for peptide bonds. Preferentially cleaves -Arg-Arg-|Xaa bonds in small molecule substrates (thus differing from cathepsin L). In addition to being an endopeptidase, shows peptidyl-dipeptidase activity, liberating C-terminal dipeptides. 2 Publications EC: 3.4.22.1
Cofactor

- non-protein substance required by an enzyme to be catalytically active
- Inorganic: metal atoms zinc, iron, etc.
- Organic: vitamins
- mapped to ChEBI identifiers
### Sites

<table>
<thead>
<tr>
<th>Feature key</th>
<th>Position(s)</th>
<th>Description</th>
<th>Actions</th>
<th>Graphical view</th>
</tr>
</thead>
<tbody>
<tr>
<td>Metal binding</td>
<td>153</td>
<td>Magnesium</td>
<td><img src="153_Magnesium" alt="Image" /></td>
<td><img src="153_Magnesium" alt="Graph" /></td>
</tr>
<tr>
<td>Active site</td>
<td>191</td>
<td>Proton acceptor</td>
<td><img src="191_Proton_Acceptor" alt="Image" /></td>
<td><img src="191_Proton_Acceptor" alt="Graph" /></td>
</tr>
<tr>
<td>Binding site</td>
<td>228</td>
<td>Substrate</td>
<td><img src="228_Substrate" alt="Image" /></td>
<td><img src="228_Substrate" alt="Graph" /></td>
</tr>
<tr>
<td>Binding site</td>
<td>347</td>
<td>Substrate</td>
<td><img src="347_Substrate" alt="Image" /></td>
<td><img src="347_Substrate" alt="Graph" /></td>
</tr>
</tbody>
</table>

**Catalytic site(s)**

**Cofactor binding site(s)**

**Substrate binding site(s)**
Biophysicochemical properties

• Information about enzyme biophysical and chemical properties
  • kinetic data
    • Michaelis-Menten constant (KM) and maximal velocity (Vmax).
  • Physiological substrates
  • optimum pH and optimal temperature for enzyme activity

---

**Kinetics**

- The $V_{\text{max}}$ with threo-D-(S)-isocitrate as substrate is 5-fold higher than with threo-D-(S)-2-methylisocitrate as substrate.  
  2 Publications

- $K_M$=1.4 mM for threo-D-(S)-isocitrate (at pH 7.0).  
  2 Publications

- $K_M$=0.6 mM for threo-D-(S)-2-methylisocitrate (at pH 7.0 and 30 degrees Celsius).  
  2 Publications

**pH dependence**

- Optimum pH is 7.  
  2 Publications

- Temperature dependence
  - Thermostable for 60 minutes up to 50 degrees Celsius.  
    2 Publications
Activity regulation

Describe components which regulate the enzyme activity:
• natural and synthetic products
• allosteric
• post-translational modification (ex. phosphorylation)

Activity regulation
Allosterically regulated and controlled by phosphorylation. 5-phosphoribose 1-diphosphate (PRPP) is an activator while UMP and UTP are inhibitors of the CPSase reaction.

1 Publication
Additional information

- Function: more substrates, context

```markdown
**Function**

Serine/threonine kinase which responds to activation by environmental stress and pro-inflammatory cytokines by phosphorylating downstream targets. Phosphorylates snk-1 which probably regulates skn-1 nuclear translocation in response to oxidative stress (PubMed:16166371). Probably by activating skn-1, involved in the up-regulation of gcs-1 and glutathione-S-transferase gst-4 expression upon bacteria infection
```

\[
\text{ATP} + \text{L-seryl-[protein]} = \text{ADP} + \text{H}^+ + \text{O-phospho-L-seryl-[protein]}
\]

EC:2.7.11.24

- Post translational modification (phosphorylation, etc.)

<table>
<thead>
<tr>
<th>Amino acid modifications</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Feature key</strong></td>
</tr>
<tr>
<td>Modified residue</td>
</tr>
<tr>
<td>Modified residue</td>
</tr>
<tr>
<td>Modified residue</td>
</tr>
</tbody>
</table>
Pathway

- describes the metabolic pathway(s) associated with an enzyme
Enzyme-related keywords and GO terms

Alkaline ceramidase (Q9D099)

### Keywords

<table>
<thead>
<tr>
<th>Molecular function</th>
<th>Hydrolase</th>
</tr>
</thead>
<tbody>
<tr>
<td>Biological process</td>
<td>Lipid metabolism, Sphingolipid metabolism</td>
</tr>
<tr>
<td>Ligand</td>
<td>Calcium, Metal-binding, Zinc</td>
</tr>
</tbody>
</table>

### GO - Molecular function

- ceramidase activity  [Source: UniProtKB-EC]
- dihydroceramidase activity  [Source: UniProtKB]
- N-acylsphingosine amidohydrolase activity  [Source: UniProtKB]
- phytoceramidase activity  [Source: UniProtKB]

View the complete GO annotation on QuickGO...
Feature viewer: summary of sequence features
How to search for enzyme-related information?

- Advanced search
- Search from the protein entry page
Search “catalytic activity” for a reaction participant

Search with:

- name
- CHEBI ID or RHEA ID
- InChIKey

Examples:

- succinate
- CHEBI:26806
- KDYFGRWQOYBRFD-UHFFFAOYSA-L
Example

Which enzymes catalyse reactions involving succinate?
<table>
<thead>
<tr>
<th>Entry</th>
<th>Entry name</th>
<th>Protein names</th>
<th>Gene names</th>
<th>Organism</th>
<th>Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>P28240</td>
<td>ACEA_YEAST</td>
<td>Isocitrate lyase</td>
<td>ICL1 YER065C</td>
<td>Saccharomyces cerevisiae (strain ATCC 204508 / S288c) (Baker's yeast)</td>
<td>557</td>
</tr>
<tr>
<td>P05050</td>
<td>ALKB_ECOLI</td>
<td>Alpha-ketoglutarate-dependent dioxygenase</td>
<td>alkB aidD, b2212, JW2200</td>
<td>Escherichia coli (strain K12)</td>
<td>216</td>
</tr>
<tr>
<td>Q3TSG4</td>
<td>ALKB5_MOUSE</td>
<td>RNA demethylase ALKB5</td>
<td>Alkbh5 Abh5, Ofoxid</td>
<td>Mus musculus (Mouse)</td>
<td>395</td>
</tr>
<tr>
<td>Q8K1E6</td>
<td>ALKB3_MOUSE</td>
<td>Alpha-ketoglutarate-dependent dioxygenase</td>
<td>Alkbh3 Abh3</td>
<td>Mus musculus (Mouse)</td>
<td>286</td>
</tr>
<tr>
<td>Q96Q83</td>
<td>ALKB3_HUMAN</td>
<td>Alpha-ketoglutarate-dependent dioxygenase</td>
<td>ALKBH3 ABH3, DEPC1</td>
<td>Homo sapiens (Human)</td>
<td>286</td>
</tr>
<tr>
<td>Q13688</td>
<td>ALKB1_HUMAN</td>
<td>Nucleic acid dioxygenase ALKB1</td>
<td>ALKBH1 ABH, ABH1, ALKBH</td>
<td>Homo sapiens (Human)</td>
<td>389</td>
</tr>
<tr>
<td>P0CB42</td>
<td>ALKB1_MOUSE</td>
<td>Nucleic acid dioxygenase ALKB1</td>
<td>Alkbh1 Abh, Alkbh</td>
<td>Mus musculus (Mouse)</td>
<td>389</td>
</tr>
<tr>
<td>P42449</td>
<td>ACEA_CORGL</td>
<td>Isocitrate lyase</td>
<td>aceA Cgl2331, cg2560</td>
<td>Corynebacterium glutamicum (strain ATCC 13032 / DSM 20300 / JCM 1318 / LMG 3730 / NCIMB 10025)</td>
<td>432</td>
</tr>
<tr>
<td>Q6PJ4</td>
<td>ALKB2_MOUSE</td>
<td>DNA oxidative demethylase ALKB2</td>
<td>Alkbh2 Abh2</td>
<td>Mus musculus (Mouse)</td>
<td>239</td>
</tr>
<tr>
<td>Entry</td>
<td>Entry name</td>
<td>Protein names</td>
<td>Gene names</td>
<td>Organism</td>
<td>Length</td>
</tr>
<tr>
<td>--------</td>
<td>-------------</td>
<td>----------------------------------------------------</td>
<td>--------------</td>
<td>---------------------------------</td>
<td>--------</td>
</tr>
<tr>
<td>P21912</td>
<td>SDHB_HUMAN</td>
<td>Succinate dehydrogenase [ubiquinone...</td>
<td>SDHB, SDH, SDH1</td>
<td>Homo sapiens (Human)</td>
<td>280</td>
</tr>
<tr>
<td>P31040</td>
<td>SDHA_HUMAN</td>
<td>Succinate dehydrogenase [ubiquinone...</td>
<td>SDHA, SDH2, SDHF</td>
<td>Homo sapiens (Human)</td>
<td>664</td>
</tr>
<tr>
<td>P53597</td>
<td>SUCA_HUMAN</td>
<td>Succinate--CoA ligase [ADP/GDP-forming]...</td>
<td>SUCLG1</td>
<td>Homo sapiens (Human)</td>
<td>346</td>
</tr>
<tr>
<td>Q96I99</td>
<td>SUCB2_HUMAN</td>
<td>Succinate--CoA ligase [GDP-forming]...</td>
<td>SUCLG2</td>
<td>Homo sapiens (Human)</td>
<td>432</td>
</tr>
<tr>
<td>Q9P2R7</td>
<td>SUCB1_HUMAN</td>
<td>Succinate--CoA ligase [ADP-forming]...</td>
<td>SUCLA2</td>
<td>Homo sapiens (Human)</td>
<td>463</td>
</tr>
</tbody>
</table>
How to search for enzyme-related information?

- Advanced search
- Search from the protein entry page
Search from the protein entry page

Which enzymes catalyse reactions involving glyoxylate?

**Catalytic activity**

- isocitrate = glyoxylate

EC: 4.1.3.1

Source: Rhea. « Hide

Search proteins in UniProtKB for this molecule.
Search chemical reactions in Rhea for this molecule.
See the description of this molecule in ChEBI.
Search proteins in UniProtKB for this molecule

Search chemical reactions in Rhea for this molecule

See the description of this molecule in CHEBI
Search from the protein entry page
Which enzymes catalyse a similar reaction?

Catalytic activity

- isocitrate $\rightarrow$
  EC: 4.1.3.1
  Source: Rhizobium

Search proteins in UniProtKB for this EC number. See the description of this EC number in ENZYME.
Search from the protein entry page

Which enzymes catalyse a similar reaction?

Catalytic activity

- isocitrate = glyoxylate + succinate

EC: 4.1.3.1

Source: Rhea.

Search for this reaction in UniProtKB.
See the description of this reaction in Rhea.
How to search for enzyme-related information?

- Advanced search

- Search from the protein entry page

- Programmatic access via restAPI
  http://uniprot.org/help/programmatic_access
Cross-references to external resources

- information related to enzymes and found in external databases

<table>
<thead>
<tr>
<th>Enzyme and pathway databases</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>UniPathway</td>
<td>UPA00655;UER00711</td>
</tr>
<tr>
<td>BioCyc</td>
<td>MetaCyc:HS05598-MONOMER</td>
</tr>
<tr>
<td>BRENDA</td>
<td>6.4.1.2 2681</td>
</tr>
</tbody>
</table>
| Reactome                     | R-HSA-163765 ChREBP activates metabolic gene expression  
                              | R-HSA-196780 Biotin transport and metabolism  
                              | R-HSA-200425 Import of palmitoyl-CoA into the mitochondrial matrix  
                              | R-HSA-2426168 Activation of gene expression by SREBF (SREBP)  
                              | R-HSA-3371599 Defective HLCS causes multiple carboxylase deficiency  
                              | R-HSA-75105 Fatty acyl-CoA biosynthesis |
| SABIO-RK                     | Q13085 |
| SIGNOR                       | Q13085 |
Cross-references to external resources

Enzymatic reaction:
- IntEnz
- BRENDA
- SABIO-RK
- swissprot
- KEGG

Enzyme families:
- ESTHER Database
- CAZY
- MEROPS
- RedOxibase

Pathways:
- Biocyc
- SignaLink 2.0
- UniPathway
- Signor 2.0
EBI-Enzyme Portal
Finding help

- http://www.uniprot.org/help/advanced_search
- https://www.uniprot.org/help/chemical_data_search
- http://www.uniprot.org/help/catalytic_activity
- http://www.uniprot.org/help/act_site
- http://www.uniprot.org/help/biophysicochemical_properties
- http://www.uniprot.org/help/cofactor

- Online training: http://www.ebi.ac.uk/training/online/course/uniprot-exploring-protein-sequence-and-functional

- Visit our YouTube Channel: https://www.youtube.com/user/uniprotvideos
- Contact us: help@uniprot.org
UniProt Team

**Pls:** Alex Bateman, Alan Bridge, Cathy Wu

**Key staff:** Cecilia Arighi (Curation), Lionel Breuza (Curation), Elisabeth Coudert (Curation), Hongzhan Huang (Development), Damien Lieberherr (Curation), Michele Magrane (Curation), Maria Martin (Development), Peter McGarvey (Content), Darren Natale (Content), Sandra Orchard (Content), Ivo Pedruuzzi (Curation), Sylvain Poux (Curation), Manuela Pruess (Coordination), Shriya Raj (Coordination), Nicole Redaschi (Development)

**Content / Curation:** Lucila Aimo, Ghislaine Argoud-Puy, Andrea Auchincloss, Kristian Axelsen, Emmanuel Boutet, Emily Bowler, Ramona Britto, Hema Bye-A-Jee, Cristina Casals-Casas, Anne Estreicher, Livia Famiglietti, Marc Feuermann, John S. Garavelli, Penelope Garmiri, George Georgiou, Arnaud Gos, Nadine Gruaz, Emma Hatton-Ellis, Ursula Hinz, Chantal Hulo, Nevila Hyka- Nouspikel, Florence Jungo, Kati Laiho, Philippe Lemercier, Yvonne Lussi, Alistair MacDougall, Patrick Masson, Anne Morgat, Sandrine Pilbout, Catherine Rivoire, Karen Ross, Christian Sigrist, Elena Speretta, Shyamala Sundaram, Nidhi Tyagi, C. R. Vinayaka, Qinghua Wang, Kate Warner, Lai-Su Yeh, Rossana Zaru


---

**EMBL-EBI**

European Bioinformatics Institute (EMBL-EBI), Hinxton, Cambridge, UK

**PIR**

Protein Information Resource (PIR), Washington DC and Delaware, USA

**SIB**

SIB Swiss Institute of Bioinformatics (SIB), Geneva, Switzerland
UniProt funding

UniProt is funded by the National Eye Institute (NEI), National Human Genome Research Institute (NHGRI), National Heart, Lung, and Blood Institute (NHLBI), National Institute of Allergy and Infectious Diseases (NIAID), National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK), National Institute of General Medical Sciences (NIGMS), and National Institute of Mental Health (NIMH) of the National Institutes of Health.