

Swiss Institute of
Bioinformatics

How to mine enzyme data ?

Practicals (and corrections)

Anne Morgat, Swiss-Prot group, SIB Swiss Institute of Bioinformatics



Lausanne, March 3rd 2020

www.sib.swiss

Q1: Retrieve human enzymes (annotated with EC numbers)

- How many human enzymes in UniProt ?
- How many expert-curated human enzymes?

Answer: Retrieve human enzymes (annotated with EC numbers)

- How many human enzymes in UniProt ? **14672**
ec:* AND organism:"Homo sapiens (Human) [9606]"
- How many expert-curated human enzymes (UniProtKB/Swiss-Prot)? **4345**
ec:* AND reviewed:yes AND organism:"Homo sapiens (Human) [9606]"

If you have questions related to the HUMAN proteome, please read
https://www.uniprot.org/help/human_proteome

Q2: Retrieve human enzymes with Catalytic activity annotations

- How many human enzymes in UniProtKB/Swiss-Prot have Catalytic activity annotations ?
- How many human enzymes in UniProtKB/Swiss-Prot have Catalytic activity annotations based on Rhea reactions?
- How do you explain the difference?

Answer: Retrieve human enzymes with Catalytic activity annotations (1)

- How many human enzymes in UniProtKB/Swiss-Prot have Catalytic activity annotations ? **3,617**

Query: annotation:(type:"catalytic activity") AND reviewed:yes AND organism:"Homo sapiens (Human) [9606]"

- How many human enzymes in UniProtKB/Swiss-Prot have Catalytic activity annotations based on Rhea reactions? **2,867**

Query: annotation:(type:"catalytic activity" rhea:*) AND reviewed:yes AND organism:"Homo sapiens (Human) [9606]"

- How do you explain the difference?

There are 750 HUMAN UniProtKB/Swiss-Prot that have a Catalytic activity annotation without Rhea reactions

Query: annotation:(type:"catalytic activity") NOT annotation:(type:"catalytic activity" rhea:*) AND organism:"Homo sapiens (Human) [9606]" AND reviewed:yes

Answer: Retrieve human enzymes with Catalytic activity annotations (2)

UniProtKB

BLAST Align Retrieve/ID mapping Peptide search Help Contact

UniProtKB results

About UniProtKB Basket

Filter by:

1 to 250 of 750 Show 250

Entry	Entry name	Protein names	Catalytic activity	ChEBI (Catalytic activity)	ChEBI (Cofactor)
<input type="checkbox"/> P42575	CASP2_HUMAN	Caspase-2	<ul style="list-style-type: none"> Strict requirement for an Asp residue at P1, with 316-asn being essential for proteolytic activity and has a preferred cleavage sequence of Val-Asp-Val-Ala-Asp- -.. EC:3.4.22.55 		
<input type="checkbox"/> Q8NB49	AT11C_HUMAN	Phospholipid-transporting ATPase IG	<ul style="list-style-type: none"> ATP + H(2)O + phospholipid(Side 1) = ADP + phosphate + phospholipid(Side 2). EC:7.6.2.1 		
<input type="checkbox"/> Q8TF62	AT8B4_HUMAN	Probable phospholipid-transporting ...	<ul style="list-style-type: none"> ATP + H(2)O + phospholipid(Side 1) = ADP + phosphate + phospholipid(Side 2). EC:7.6.2.1 		
<input type="checkbox"/> Q9UNE7	CHIP_HUMAN	E3 ubiquitin-protein ligase CHIP	<ul style="list-style-type: none"> S-ubiquitinyl-[E2 ubiquitin-conjugating enzyme]-L-cysteine + [acceptor protein]-L-lysine = [E2 ubiquitin-conjugating enzyme]-L-cysteine + N(6)-ubiquitinyl-[acceptor protein]-L-lysine. <input type="button" value="2 Publications"/> EC:2.3.2.27 		

Your results in sequence clusters with identity of:

Q3: Search UniProt by chemical name: 'cholesterol'

Simple search

- How many UniProt entries?
- How many Swiss-Prot entries?
- How many TrEMBL entries?

Advanced search (Catalytic Activity)

- How many Swiss-Prot entries?
- How many Swiss-Prot entries with any experimental evidence?
- How many HUMAN entries with any experimental evidence?

Answer: Search UniProt by chemical name: 'cholesterol'

Simple search

- How many UniProt entries? **156,727**
- How many Swiss-Prot entries? **2354**
- How many TrEMBL entries? **154373**

Advanced search (Catalytic Activity)

- How many Swiss-Prot entries? **131**
- How many Swiss-Prot entries with any experimental evidence? **47**
- How many HUMAN entries with any experimental evidence? **20**

Q4: Search UniProtKB/Swiss-Prot by chemical name: 'NAD+'

Using UniProt Advanced search for small molecule

- What is the ChEBI Id of NAD+?
- How many Swiss-Prot entries are annotated with NAD+?
- What is the role of NAD+?
Tips: Customize Result table to display:
Entry name, Protein names, Catalytic activity, ChEBI (Catalytic activity), ChEBI (Cofactor)
- Can NAD+ be both a cofactor and a reaction participant in the same enzyme?

Answer: Search UniProt by chemical name: 'NAD+' (1)

Advanced search (small molecule)

- What is the ChEBI ID? **CHEBI:57540**
- How many Swiss-Prot entries? **18,408**
chebi:"NAD(+)" [57540] AND reviewed:yes

- Customize Result table to display:

Entry name
Protein names
Catalytic activity
ChEBI (Catalytic activity)
ChEBI (Cofactor)

UniProt search results for: `chebi:"NAD(+)" [57540] AND reviewed:yes`

Customize results table

Columns to be displayed:

Drag and drop to re-order:

Entry name Protein names Catalytic activity ChEBI (Catalytic activity) ChEBI (Cofactor)

Add more columns:

Search:

Names & Taxonomy

- Entry name
- Gene names
 - Gene names (ordered locus)
 - Gene names (ORF)
 - Gene names (primary)
 - Gene names (synonym)
- Organism
- Organism ID
- Protein names
- Proteomes
- Taxonomic lineage
- Virus hosts

Sequences

- Alternative products (isoforms)
- Alternative sequence
- Erroneous gene model prediction
- Fragment
- Gene encoded by
- Length
- Mass
- Mass spectrometry
- Natural variant
- Non-identical residues
- Non-standard residue
- Non-terminal residue
- Polymerism
- RNA editing
- Sequence
- Sequence caution
- Sequence conflict
- Sequence uncertainty
- Sequence version

Function

- Absorption
- Active site
- Active regulation
- Binding site
- Calcium binding
- Catalytic activity
- Cofactor
 - DNA binding
 - EC number
 - Function [CC]
 - Kinetics
 - Metal binding
 - Nucleic acid binding
 - Pathway
 - pH dependence
 - Redox potential
 - Rhes IDs
 - Site
 - Temperature dependence

Miscellaneous

- Annotation
- Caution
- Features
- Keyword ID
- Keywords
- Matched text
- Miscellaneous [CC]
- Protein evidence
- Tools
- UniProt

Interaction

Expression

Gene Ontology (GO)

Chemical entities (CHEBI)

- ChEBI
- ChEBI (Catalytic activity)
- ChEBI (Cofactor)
- ChEBI IDs

Answer: Search UniProt by chemical name: 'NAD+' (2)

- What do you observe? **NAD(+)** [CHEBI:57540] is either a reaction participant or a cofactor

The screenshot shows the UniProt search results page for the query 'chebi:NAD(+) [57540] AND reviewed:yes'. The search results are displayed in a table with columns for Entry, Entry name, Protein names, Catalytic activity, ChEBI (Catalytic activity), and ChEBI (Cofactor). The table lists four entries: Q9TOA7, Q9C7W7, Q79AF6, and Q88FF8. The first two entries are UDP-glucose 4-epimerase 2 and 4, and the last two are Acetaldehyde dehydrogenase 4 and Quinone reductase. The table shows that NAD+ is involved in the catalytic activity of the first two entries and acts as a cofactor for the last two entries.

Entry	Entry name	Protein names	Catalytic activity	ChEBI (Catalytic activity)	ChEBI (Cofactor)
<input type="checkbox"/> Q9TOA7	UGE2_ARATH	UDP-glucose 4-epimerase 2	<ul style="list-style-type: none">• UDP-D-glucose = UDP-D-glucose EC:5.1.3.2 Source: Rhea.	UDP-D-glucose [CHEBI:66914] UDP-D-glucose [CHEBI:58885]	NAD ⁺ [CHEBI:57540]
<input type="checkbox"/> Q9C7W7	UGE4_ARATH	UDP-glucose 4-epimerase 4	<ul style="list-style-type: none">• UDP-D-glucose = UDP-D-glucose EC:5.1.3.2 Source: Rhea.	UDP-D-glucose [CHEBI:66914] UDP-D-glucose [CHEBI:58885]	NAD ⁺ [CHEBI:57540]
<input type="checkbox"/> Q79AF6	ACDH4_PARXL	Acetaldehyde dehydrogenase 4	<ul style="list-style-type: none">• acetaldehyde + CoA + NAD⁺ = acetyl-CoA + H⁺ + NADH EC:1.2.1.10 Source: Rhea.• 1-propanal + CoA + NAD⁺ = H⁺ + NADH + propanoyl-CoA EC:1.2.1.87 Source: Rhea.	propanoyl-CoA [CHEBI:57392] 1-propanal [CHEBI:17153] H ⁺ [CHEBI:15378] NAD ⁺ [CHEBI:57540] CoA [CHEBI:57287] acetyl-CoA [CHEBI:57288] NADH [CHEBI:57945] acetaldehyde [CHEBI:15343]	
<input type="checkbox"/> Q88FF8	CHRR_PSEPK	Quinone reductase	<ul style="list-style-type: none">• a quinone + H⁺ + NADH = a quinol + NAD⁺ 1 Publication• a quinone + H⁺ + NADPH = a quinol + NADP⁺ 1 Publication• Cr⁶⁺ + 2 NADH + O₂ = Cr³⁺ + 2 H⁺ + 2 NAD⁺ + superoxide 1 Publication• Cr⁶⁺ + 2 NADPH + O₂ = Cr³⁺ + 2 H⁺ + 2 NADP⁺ + superoxide 1 Publication	H ⁺ [CHEBI:15378] O ₂ [CHEBI:15379] NAD ⁺ [CHEBI:57540] superoxide [CHEBI:18421] a quinol [CHEBI:24646] NADPH [CHEBI:57783] Cr ³⁺ [CHEBI:49544] NADH [CHEBI:57945] NADP ⁺ [CHEBI:58349] a quinone [CHEBI:132124] Cr ⁶⁺ [CHEBI:33007]	FMN [CHEBI:58210]

Answer: Search UniProt by chemical name: 'NAD+' (3)

- Can NAD+ be both a cofactor and a reaction participant in the same enzyme? **no**

Searching in UniProtKB [Help](#)

Term: "NAD(+)" [57540]
Evidence: Any assertion method

AND Function > Cofactors > Cofactor

Term: CHEBI:"NAD(+)" [57540]
Evidence: Any assertion method

AND Reviewed > Reviewed

Search

UniProtKB - cofactor:(chebi:"NAD(+)" [57540]) annotation:(type:"catalytic activity" CHEBI:"NAD(+)" [57540]) reviewed:yes [Advanced](#) [Search](#)

[BLAST](#) [Align](#) [Retrieve/ID mapping](#) [Peptide search](#) [Help](#) [Contact](#)

UniProtKB results

[Filter byⁱ](#)
[View by](#)
[Demo](#)
[Help video](#)

[About UniProtKB](#) [Basket](#)

Sorry, no results found for your search term.

Can't find what you are looking for? Please [contact us](#)

Q5: Search UniProtKB/Swiss-Prot entries for RHEA:19253

Advanced search (Catalytic activity)

- How many UniProtKB/Swiss-Prot entries?
- What are the corresponding reactions in KEGG, MetaCyc and Reactome?
- What are parent/children reactions of RHEA:19253?

Answer: Search UniProtKB/Swiss-Prot entries for RHEA:19253

Advanced search (Catalytic activity)

- How many UniProtKB/Swiss-Prot entries? **34**
- What are the corresponding reactions in KEGG, MetaCyc and Reactome?

Go to <https://www.rhea-db.org/reaction?id=19253>

- What are parent/children reactions of RHEA:19253?

Parent reaction:

- RHEA:45300 an N-(acyl)-sphingosylphosphocholine + H₂O = H(+) + N-acyl-sphingoid base + phosphocholine

Children reactions

- RHEA:45324 H₂O + N-(tetracosanoyl)-sphing-4-enine-1-phosphocholine = H(+) + N-(tetracosanoyl)-sphing-4-enine + phosphocholine
- RHEA:54284 H₂O + N-(octadecanoyl)-sphing-4-enine-1-phosphocholine = H(+) + N-octadecanoylsphing-4-enine + phosphocholine

Q6: Search by InChIKey

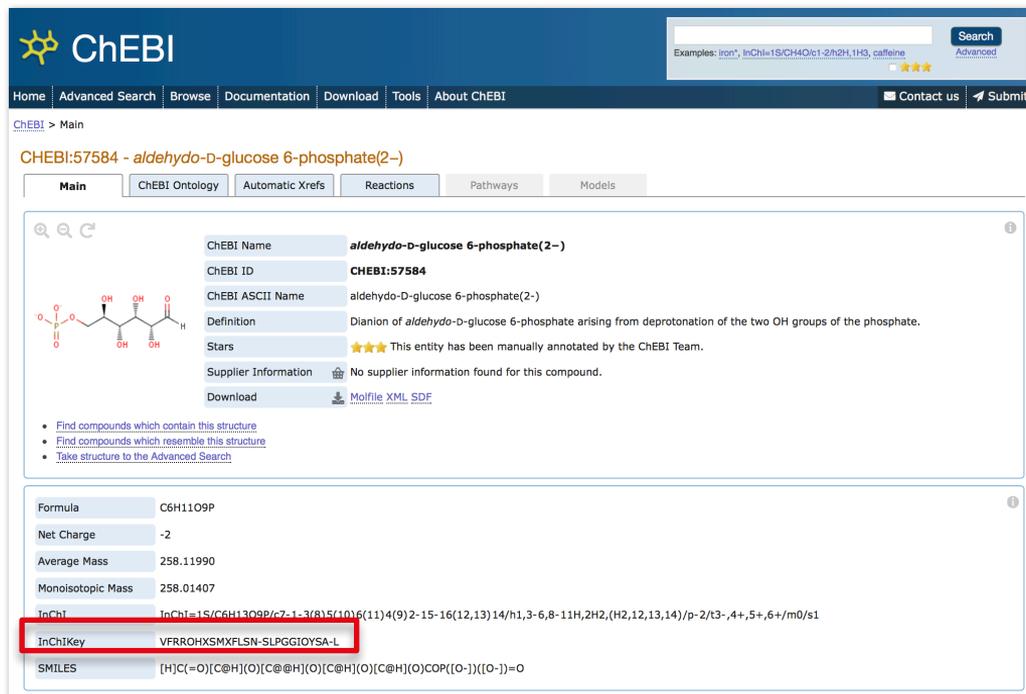
- Search ChEBI:57584 in ChEBI resource and copy its InChiKey
- How many UniProt entries are annotated with the full InChiKey?
- Remove the last two part, and see how many UniProt entries are found?

Answer: Search by InChIKey (1)

- Search ChEBI:57584 in ChEBI resource and copy its InChiKey

<https://www.ebi.ac.uk/chebi/searchId.do?chebid=CHEBI:57584>

VFRROHXSMXFLSN-SLPGGIOYSA-L



ChEBI

Home | Advanced Search | Browse | Documentation | Download | Tools | About ChEBI | Contact us | Submit

ChEBI:57584 - *aldehyde-D-glucose 6-phosphate(2-)*

Main | ChEBI Ontology | Automatic Xrefs | Reactions | Pathways | Models

ChEBI Name *aldehyde-D-glucose 6-phosphate(2-)*

ChEBI ID **CHEBI:57584**

ChEBI ASCII Name aldehyde-D-glucose 6-phosphate(2-)

Definition Dianion of *aldehyde-D-glucose 6-phosphate* arising from deprotonation of the two OH groups of the phosphate.

Stars ★★ This entity has been manually annotated by the ChEBI Team.

Supplier Information No supplier information found for this compound.

Download [Molfile](#) [XML](#) [SDF](#)

- Find compounds which contain this structure
- Find compounds which resemble this structure
- Take structure to the Advanced Search

Formula C6H11O9P

Net Charge -2

Average Mass 258.11990

Monoisotopic Mass 258.01407

InChI InChI=1S/C6H13O9P/c7-1-3(8)5(10)6(11)4(9)2-15-16(12,13)14/h1,3-6,8-11H,2H2,(H2,12,13,14)/p-2/t3-,4+,5+,6+/m0/s1

InChIKey VFRROHXSMXFLSN-SLPGGIOYSA-L

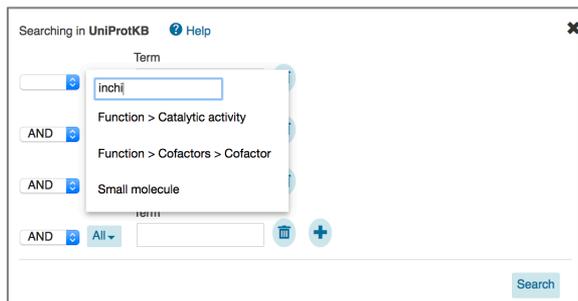
SMILES [H]C(=O)[C@H](O)[C@@H](O)[C@H](O)[C@H](O)COP([O-])([O-])=O

Answer: Search by InChIKey (2)

- How many UniProt entries are annotated with the full InChIKey? **38,647**

Advanced search

Term “inchi” – Select Function > Catalytic activity



Searching in UniProtKB [Help](#)

Term

inchi

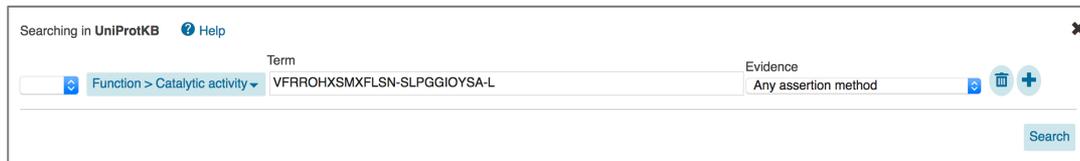
Function > Catalytic activity

Function > Cofactors > Cofactor

Small molecule

AND All

Search



Searching in UniProtKB [Help](#)

Term

Function > Catalytic activity

VFRROHXSXMFLSN-SLPGGIOYSA-L

Evidence

Any assertion method

Search

Query: annotation:(type:"catalytic activity" inchikey:VFRROHXSXMFLSN-SLPGGIOYSA-L)
640 reviewed (UniProtKB/Swiss-Prot)
38,007 unreviewed (UniProtKB/TrEMBL)

Answer: Search by InChIKey (3)

- Remove the last two part, and see how many UniProt entries are found? **40,458**
Query: `annotation:(type:"catalytic activity" inchikey:VFRROHXSMXFLSN)`
754 reviewed (UniProtKB/Swiss-Prot)
39,704 unreviewed (UniProtKB/TrEMBL)

Q6: Retrieve cholesterol and derivatives using structural search

- Search cholesterol in ChEBI and copy its SMILES
- Perform a substructural search in the Rhea web site
- How many cholesterol or cholesterol derivatives are use in Rhea?
- Retrieve the Rhea reaction(s) involving protodioscin
- Retrieve their catalyst(s) in UniProtKB/Swiss-Prot

Answer: Retrieve cholesterol and derivatives using structural search(1)

- Search cholesterol in ChEBI resource and copy its SMILES
C1C2(C3(CCC4(C(C3(CC=C2CC(C1)O)))(CCC4(C(C)CCCC(C)C)))C)C
- Rhea web site - Advanced search

The screenshot shows the Rhea website's advanced search page. The browser address bar displays <https://www.rhea-db.org/advancedsearch>. The page title is "Rhea - Annotated reactions database - Mozilla Firefox". The search bar contains the text "advanced / browse". Below the search bar, there are several search filters: "Wildcard characters: ? (one character), * (one or more characters)", "Excluding words: All of these words", "in field:", "Transport:", and "With cross-references to". The "Reaction participants search using ChEBI" section is visible at the bottom, with a note "This search is powered by ChEBI".

The screenshot shows the "Open File" dialog box for the Chemical Structure Search tool. The "Input" tab is selected, and the search criteria are defined by the SMILES string: C1C2(C3(CCC4(C(C3(CC=C2CC(C1)O)))(CCC4(C(C)CCCC(C)C)))C)C. The dialog box includes a "Load as a fragment and copy to the Clipboard" option and "Cancel" and "OK" buttons. On the right side, there are search options: "Find this entity", "Find compounds which contain this structure", and "Find compounds which resemble this structure". The "Results per page" is set to 15. The search criteria are "Search for" with "All in ChEBI" selected. The "Search" and "Reset" buttons are visible.

The screenshot shows the "Open File" dialog box for the Chemical Structure Search tool, displaying the chemical structure of cholesterol. The structure is a complex polycyclic molecule with a hydroxyl group, a double bond, and several methyl groups. The search criteria are the same as in the previous screenshot: C1C2(C3(CCC4(C(C3(CC=C2CC(C1)O)))(CCC4(C(C)CCCC(C)C)))C)C. The "Strictly Stereo" option is set to "Yes". The "Results per page" is set to 15. The search criteria are "Search for" with "All in ChEBI" selected. The "Search" and "Reset" buttons are visible.

Answer: Retrieve cholesterol and derivatives using structural search(2)

- How many cholesterol or cholesterol derivatives are use in Rhea? **68**
- Retrieve the Rhea reaction(s) involving protodioscin

Reaction participants search using ChEBI

This search is powered by ChEBI

Search Results for All in ChEBI

substructure AND Rhea in Cross_reference_sources



Edit Search

Download your results

68 entries found, displaying 1 to 15.

protodioscin CHEBI:8588 Stars: ★★★★★ Formula: C51H84O22 Mass: 1049.19950 Charge: 0 Rhea reactions	26-desglucoavenacoside B CHEBI:75931 Stars: ★★★★★ Formula: C51H82O23 Mass: 1063.18300 Charge: 0 Rhea reactions	avenacoside A CHEBI:2937 Stars: ★★★★★ Formula: C51H82O23 Mass: 1063.18300 Charge: 0 Rhea reactions
(25S)-3β-hydroxy-5-cholesten-26-oyl-CoA(4-)  CHEBI:83783 Stars: ★★★★★ Formula: C48H74N7O18P3S Mass: 1162.12600 Charge: -4 Rhea reactions	avenacoside B CHEBI:2938 Stars: ★★★★★ Formula: C57H92O28 Mass: 1225.32360 Charge: 0 Rhea reactions	cholesterol CHEBI:16113 Stars: ★★★★★ Formula: C27H46O Mass: 386.655 Charge: 0 Rhea reactions
24-methylenecholesterol CHEBI:19812 Stars: ★★★★★ Formula: C28H46O Mass: 398.66424 Rhea reactions	(25R)-3β-hydroxycholest-5-en-26-ol CHEBI:86096 Stars: ★★★★★ Formula: C27H44O2 Rhea reactions	3β-hydroxycholest-5-en-26-ol CHEBI:84145 Stars: ★★★★★ Formula: C27H44O2 Rhea reactions

Rhea

Search advanced / browse

Home SPARQL Web service Submit Download Statistics FAQ Documentation Changes Feedback

Rhea results for: CHEBI:8588

Number of hits: 1

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

Accession	Equation / Hits details	Other directions
RHEA:37895	H ₂ O + protodioscin = 26-deglucoprotodioscin + D-glucose	RHEA:37896 RHEA:37898 RHEA:37899

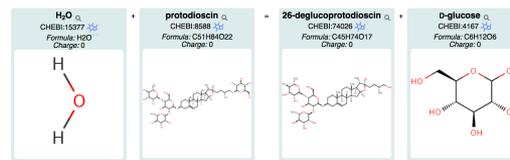
Rhea - Annotated reactions database - Mozilla Firefox

https://www.rhea-db.org/reaction?id=37895

RHEA:37895 (APPROVED)

H₂O + protodioscin = 26-deglucoprotodioscin + D-glucose

Last modified: 2019-12-09. Chemically balanced: yes.



Same participants, different directions

RHEA:37896 ⇌ H₂O + protodioscin ⇌ 26-deglucoprotodioscin + D-glucose
RHEA:37897 ⇌ 26-deglucoprotodioscin + D-glucose ⇌ H₂O + protodioscin
RHEA:37898 ⇌ H₂O + protodioscin ⇌ 26-deglucoprotodioscin + D-glucose

Links to other resources

Enzymes
Enzyme classification: EC 3.1.1.122 [ENZYME] [IntEnz]
UniProtKB: [reviewed protein] [all proteins(s)]
Reactions
KEGG reaction: R10530
MetaCyc: RXN-14633

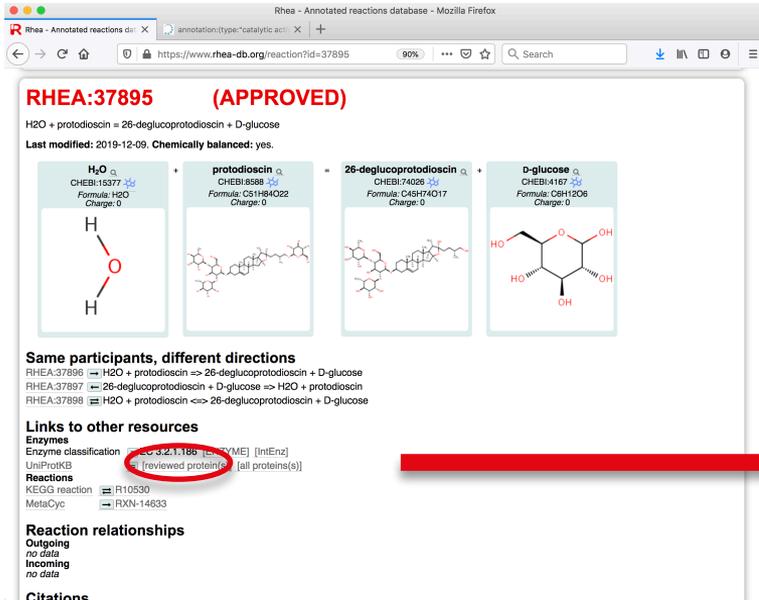
Reaction relationships

Outgoing: no data
Incoming: no data

Citations

Answer: Retrieve cholesterol and derivatives using structural search(3)

- Retrieve their catalyst(s) in UniProtKB/Swiss-Prot



RHEA:37895 (APPROVED)
H₂O + protodioscin = 26-deglucoprotodioscin + D-glucose
Last modified: 2019-12-09. Chemically balanced: yes.

Chemical structures shown:
- H₂O (CHEBI:15377)
- protodioscin (CHEBI:8568)
- 26-deglucoprotodioscin (CHEBI:74028)
- D-glucose (CHEBI:14167)

Same participants, different directions
RHEA:37896 ⇌ H₂O + protodioscin ⇒ 26-deglucoprotodioscin + D-glucose
RHEA:37897 ⇌ 26-deglucoprotodioscin + D-glucose ⇒ H₂O + protodioscin
RHEA:37898 ⇌ H₂O + protodioscin ⇌ 26-deglucoprotodioscin + D-glucose

Links to other resources
Enzymes
Enzyme classification: [EC:3.2.1.186](#) [Enzyme] [UniEnz]
UniProtKB: [reviewed protein\(s\)](#) [all protein(s)]

Reactions
KEGG reaction: [R10530](#)
MetaCyc: [RXN-14633](#)

Reaction relationships
Outgoing: no data
Incoming: no data

Citations



UniProtKB results for `annotation:(type:"catalytic activity" rhea:37895) AND reviewed:yes`

Filter by: Reviewed (2) Swiss-Prot

Entry	Entry name	Protein names	Catalytic activity	EC number
Q42707	F26G_CHESP	Furostanol glycoside 26-O-beta-glucosidase, CsF26G, EC 3.2.1.186 (Protodioscin 26-O-beta-D-glucosidase)	<ul style="list-style-type: none">H₂O + protodioscin = 26-deglucoprotodioscin + D-glucose 2 Publications EC:3.2.1.186 2 Publications Source: Rhea.	3.2.1.186
P0DKH4	F26G_SOLTO	Furostanol glycoside 26-O-beta-glucosidase, EC 3.2.1.186 (Torvosidase)	<ul style="list-style-type: none">H₂O + protodioscin = 26-deglucoprotodioscin + D-glucose 1 Publication EC:3.2.1.186 1 Publication Source: Rhea.	3.2.1.186

A case study (1)

« Biologists analyzing a metabolomics dataset from a cardiovascular disease cohort finds alterations in metabolites whose masses are consistent with cholesterol and related molecules. They also have proteomics data that show significant differences in some proteins. »

Questions:

- What are the enzymes involved in human cholesterol metabolism?
- Are the enzymes known to be linked to cardiovascular diseases (CVD) or other diseases?

SPARQL query: Retrieve the Rhea biochemical reactions that involve cholesterol or cholesterol derivatives (using sachem:substructureSearch)

Cholesterol SMILES = C1C2(C3(CCC4(C(C3(CC=C2CC(C1)O)))(CCC4(C(C)CCCC(C)C))C))C

endpoint: <https://sparql.rhea-db.org/sparql>

PREFIX ch:<http://purl.obolibrary.org/obo/>

PREFIX taxon:<http://purl.uniprot.org/taxonomy/>

PREFIX keywords:<http://purl.uniprot.org/keywords/>

SELECT distinct ?CHEBI

?CHEBI_UNIPROT_NAME

?RHEA

?REACTION_EQUATION

WHERE {

idsm:chebi service

SERVICE idsm:chebi {

?CHEBI sachem:substructureSearch [sachem:query

'C1C2(C3(CCC4(C(C3(CC=C2CC(C1)O)))(CCC4(C(C)CCCC(C)C))C))C] .

}

?RHEA_REACTION rdfs:subClassOf rh:Reaction .

?RHEA_REACTION rh:status rh:Approved .

?RHEA_REACTION rh:side / rh:contains / rh:compound / rh:chebi ?CHEBI .

?CHEBI up:name ?CHEBI_UNIPROT_NAME .

}

Rhea SPARQL endpoint (beta)

Your SPARQL query
Paste an example query from the list on the right, or write your query here.

Results Format: HTML

Run Query

About
This SPARQL endpoint contains Rhea reactions data (see statistics), and a snapshot of ChEBI data matching the Rhea release. It is free to access and supports the SPARQL 1.1 standard. The query timeout is 45 minutes. All triples are available in the default graph. There are two named graphs: <http://sparql.rhea-db.org/rhea> and <http://sparql.rhea-db.org/chebi>.

Documentation
See Rhea RDF diagrams.

Downloads
See Rhea FTP, containing rhea.rdf and chebi.owl. The chebi.owl file is a snapshot of chebi matching the Rhea (and Uniprot) release cycle.

Contact
See feedback form.

How to cite
Thierry Lambrard, Anne Morgat, Kristian B Avelsen, Lucilla Aimo, Nevila Hyka-Nouspikel, Anne Niknejad, Alex Ignatchenko, Ioannis Xenarios, Elisabeth Coudot, Nicole Redaschi, Alan Bridg
Updates in Rhea: SPARQLing biochemical reaction data
Nucleic Acids Research, Volume 47, Issue D1, 8 January 2019, Pages D586–D600; doi: 10.1093/nar/gky876

Examples

1. Select all Rhea reactions (unspecified direction), with "Approved" status. [Show](#)
2. Select all approved reactions using CHEBI:29985 (L-glutamate) as small molecule participant. [Show](#)
3. Select all approved reactions using L-glutamate as small molecule participant, by exact InChIKey. [Show](#)
4. Select all approved reactions using L-glutamate as small molecule participant, by partial InChIKey, allowing any charge. [Show](#)
5. Select the number of approved reactions involving lipids (subclass of CHEBI:18059) based on the ChEBI ontology. [Show](#)
6. Select all approved reactions involving lipids (subclass of CHEBI:18059) based on the ChEBI ontology. [Show](#)
7. Select all approved reactions using L-glutamate (CHEBI:29985) AND L-glutamine (CHEBI:56359) in different reaction sides. [Show](#)
8. Select all approved reactions annotated with a given Pubmed ID (2460082). [Show](#)
9. Select all approved transport reactions. [Show](#)
10. Select all cross-references to KEGG, MetaCyc, MACIE, ... for a given reaction (RHEA:11932). [Show](#)
11. Select all approved reactions linked to a given EC number (EC 1.1.1.353). [Show](#)
12. Distribution of reactions according to the first level of enzyme classification (federated query). [Show](#)
13. Select all compounds and count their occurrence in Rhea reactions. [Show](#)
14. Select children of CHEBI:35179 (a 2-oxo carboxylate) in the ChEBI hierarchy (using rdfs:subClassOf), used in Rhea reactions or not. [Show](#)
15. Select all the descendants of CHEBI:35179 (a 2-oxo carboxylate) in the ChEBI hierarchy (using rdfs:subClassOf), used in Rhea reactions or not. [Show](#)
16. Select children of CHEBI:35179 (a 2-oxo carboxylate) in the ChEBI hierarchy (using rdfs:subClassOf) used in Rhea reactions, and show the reaction(s). [Show](#)
17. Select all the descendants of CHEBI:35179 (a 2-oxo carboxylate) in the ChEBI hierarchy (using rdfs:subClassOf) used in Rhea reactions, and show the reaction(s). [Show](#)
18. Select the child reaction(s) of a given reaction (RHEA:11628) in the Rhea hierarchy (using rdfs:subClassOf). [Show](#)

Copy/paste the query in
<https://sparql.rhea-db.org/sparql>

A case study (2)

Query: Retrieve the number of UniProtKB/Swiss-Prot human enzymes that metabolize cholesterol or cholesterol derivatives (using sachem:substructureSearch)

See query example 22

Query: Retrieve the list of UniProtKB/Swiss-Prot human enzymes that metabolize cholesterol or cholesterol derivatives (using sachem:substructureSearch)

See query example 23

Query: Retrieve the number of UniProtKB/Swiss-Prot human enzymes that metabolize cholesterol or cholesterol derivatives and that are involved in diseases (federated query using Rhea, IDSM/Sachem and UniProt)

See query example 24

Query: Retrieve the list of diseases involving human enzymes that metabolize cholesterol or cholesterol derivatives and the number of proteins involved (federated query using Rhea, IDSM/Sachem and UniProt)

See query example 25