

Introduction to Biomedical Ontologies

Cecilia Arighi, PhD
Research Associate Professor
arighi@udel.edu

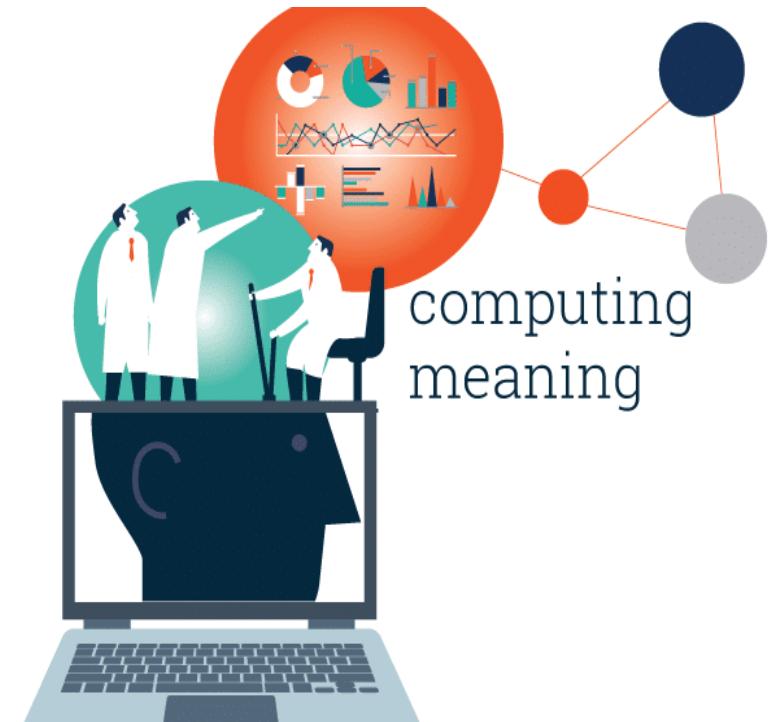


-of being or existence
Ontology:
-study

“The science of what is, of the kinds and structures of objects, properties, events, processes and relations in every area of reality”

Dr. Barry Smith, U Buffalo

Today's ontologies conceptualize the world by defining classes and relationships.



Source: <https://www.ontotext.com/blog/meaning-ontology/#>

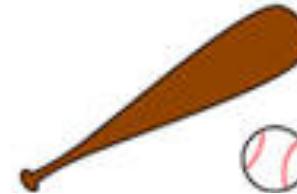
What are Ontologies?

What : Define domain and precise definition of classes.

BAT:



OR



?

How: Describe basic categories of and relationships between classes (`is_a`, `has_part`, `derives_from`, etc).

Mammal

is a



Bat

is a



Fruit Bat

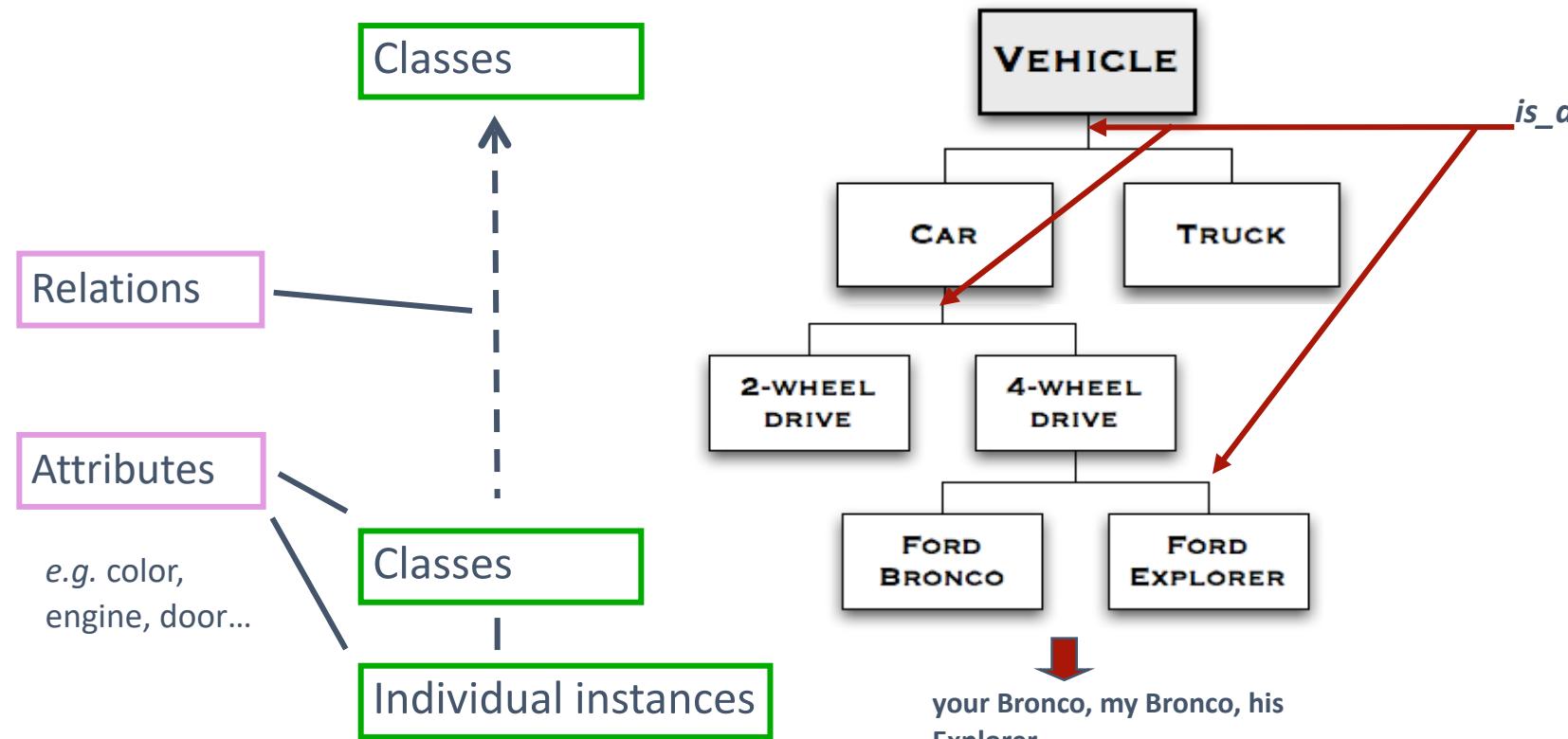


Vampire Bat

Why: Foster interoperability, communication, and analysis.



A small example...



Courtesy of Dr. Zhangzhi Hu



OBO Foundry

<https://obofoundry.org/>

The Open Biological and Biomedical Ontology (OBO) Foundry

Community development of interoperable ontologies for the biological sciences

Learn about OBO best practices and community resources

- More about the OBO Foundry
- OBO Foundry principles
- OBO tutorial
- Ontology browsers, tutorials, and tools

Participate

- Join the OBO mailing list
- OBO Foundry Operations and Working Groups
- Submit bug reports or suggestions for improvement via GitHub
- Submit your ontology to be considered for inclusion in the OBO Foundry

OBO Library: find, use, and contribute to community ontologies

The table below lists current OBO ontologies (in alphabetical order, but with the ontologies that have been manually reviewed by the OBO Foundry listed first, and obsolete ontologies listed last).

Download table as: [[YAML](#) | [JSON-LD](#) | [RDF/Turtle](#)]

Abbreviation	Name	Description	Action Buttons	Stars
bfo	Basic Formal Ontology	The upper level ontology upon which OBO Foundry ontologies are built. Detail		181
chebi	Chemical Entities of Biological Interest	A structured classification of molecular entities of biological interest focusing on 'small' chemical compounds. Detail		22
doid	Human Disease Ontology	An ontology for describing the classification of human diseases organized by etiology. Detail		264
go	Gene Ontology	An ontology for describing the function of genes and gene products Detail		148
obi	Ontology for Biomedical Investigations	An integrated ontology for the description of life-science and clinical investigations Detail		51

GOAL: create a suite of orthogonal interoperable reference ontologies in the biology and biomedical domains

- open use
- collaborative development
- non-overlapping
- common syntax and relations

Gene Ontology=> properties of gene products

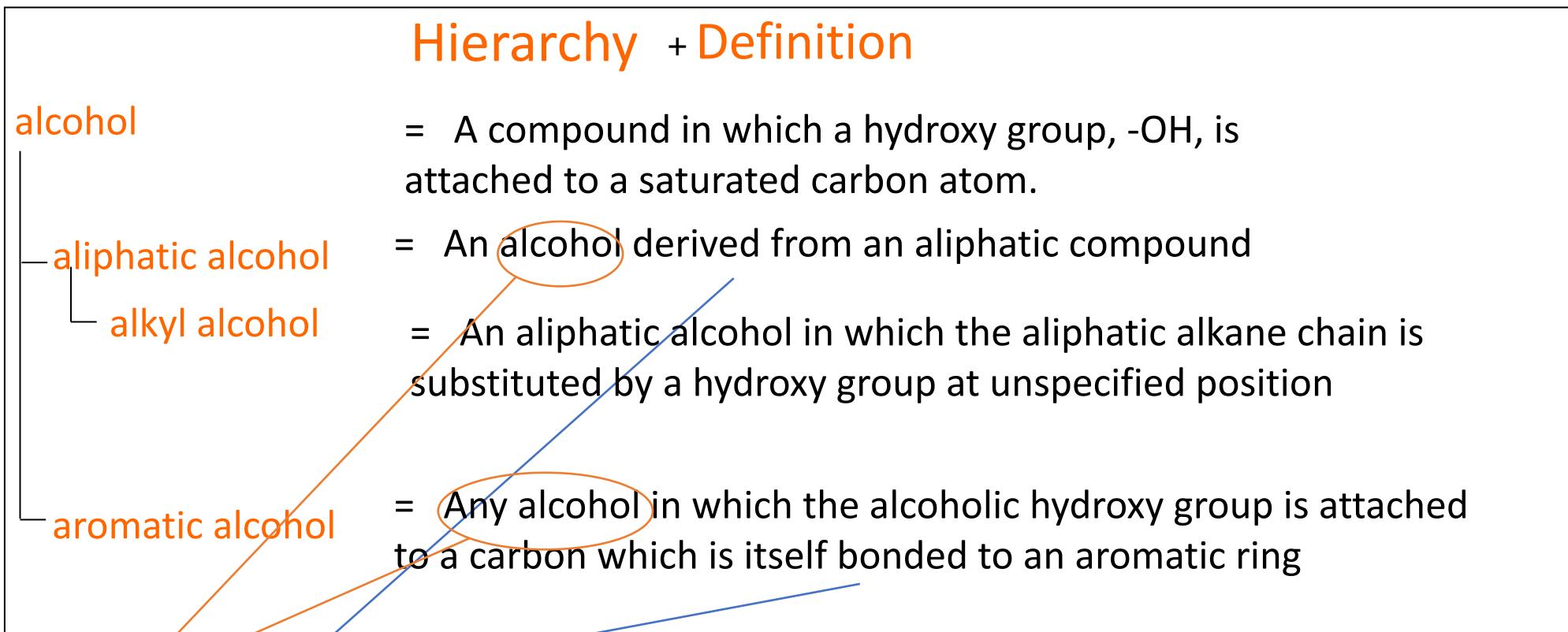
ChEBI => 'small' chemical compounds

DO => disease ontology

PRO => protein related entities



Elements of an ontology



It is the definition that allows you to unambiguously identify the entity in question

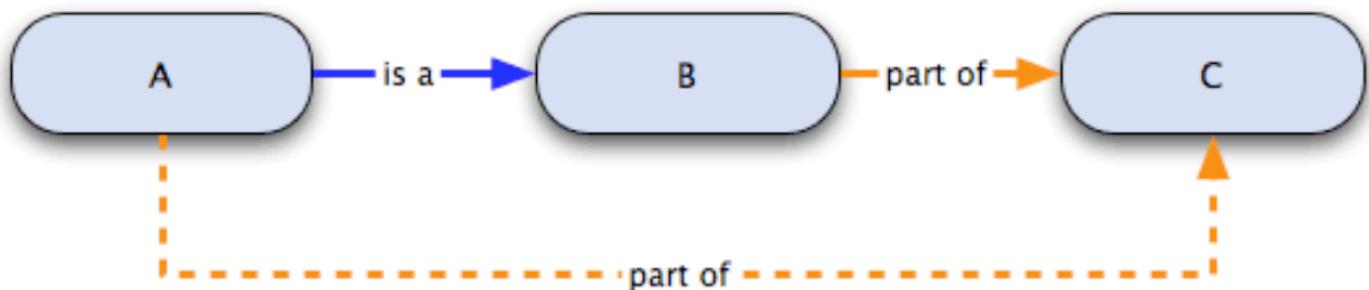
Alcohol: http://purl.obolibrary.org/obo/CHEBI_30879



Ontology structure

Terms (nodes) + Relations (edges)

This diagram would be interpreted as follows:



- A *is a* B
- B *is part of* C
- we can infer that A *is part of* C



Source: <http://geneontology.org/docs/ontology-relations/>



Relation	Description	Use for grouping annotations	Reasoning rules
<i>is a</i>	If we say A <i>is a</i> B, we mean that node A is a subtype of node B.	It is safe to use <i>is a</i> to group annotations.	
<i>part of</i>	The <i>part of</i> relation is used to represent part-whole relationships.	It is safe to use <i>part of</i> to group annotations.	
<i>regulates</i>	One process directly affects the manifestation of another process or quality, i.e., the former <i>regulates</i> the latter	<p><i>regulates</i> changes the relationship between the GO term and the gene product over the <i>is a</i> and <i>part of</i> relations. Some tools use <i>regulates</i> relations to group annotations. This can be useful for gene-set enrichment. The resulting gene sets include genes that are involved in processes that are causally related to the grouping term.</p>	

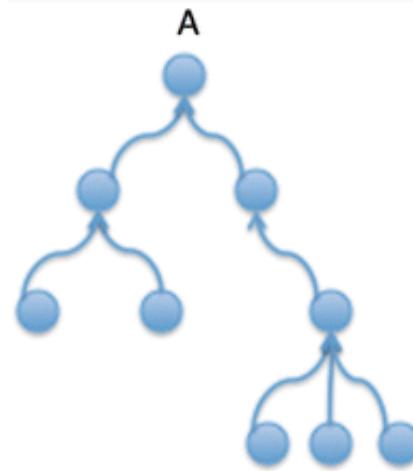
Source:<http://geneontology.org/docs/ontology-relations/>



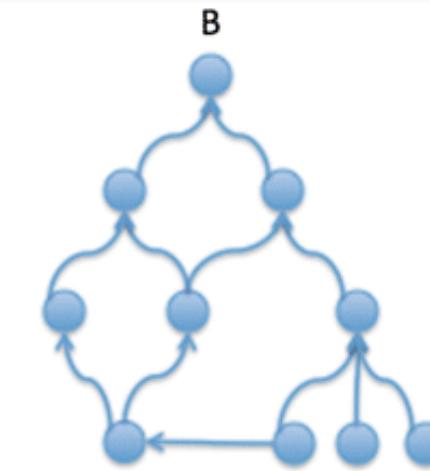
Hierarchical organization

Overall structure of the ontology should not contain cycles

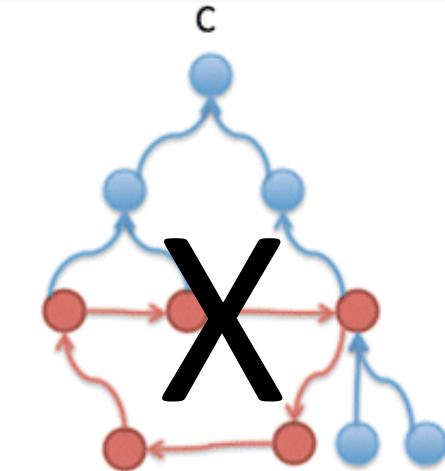
Simple tree



Directed, acyclic graph (DAG)



Graph with cycle



Source:<http://geneontology.org/docs/ontology-relations/>



Ontology types

Reference ontologies	Application ontologies	Aggregation ontologies
Aims at making the best possible description of a domain in reality.	Used to meet specific use cases	Merges multiple resources/ontologies
Intended to be re-used in multiple application contexts	Import all or parts of reference ontologies that are required to support the application use cases	Harmonize definitions
	Integrate ontologies across a common axis	
Disease Ontology (DO) Chemical ontology (ChEBI)	Experimental Factor Ontology (EFO) a systematic description of many experimental variables	MONDO disease ontology





<http://geneontology.org>

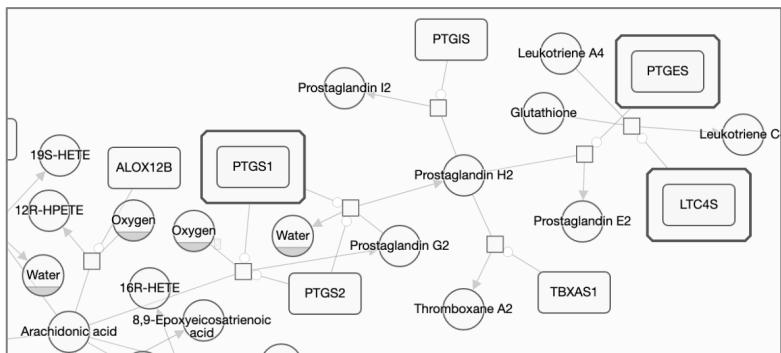
- The Gene Ontology is the most successful biomedical ontology
- GO was originally constructed in 1998 by a consortium of researchers studying the genome of three model organisms:
 - *Drosophila melanogaster* (fruit fly) (FlyBase)
 - *Mus musculus* (mouse) (MGD)
 - *Saccharomyces cerevisiae* (yeast) (SGD)
- Many databases have joined the GO Consortium (model organism and others)
- GO Consortium aims at standardizing the representation of gene and gene product attributes across species and databases.

What is GO?

GO is an ontology with concepts used to describe (annotate) the normal properties of gene products

Which process does it participate in?

GO Biological process

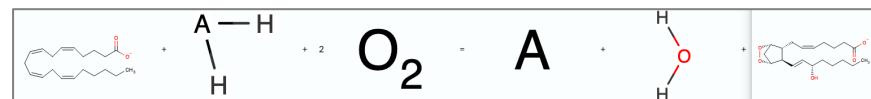


prostaglandin biosynthetic process

GO:0001516

What does it do?

GO Molecular function

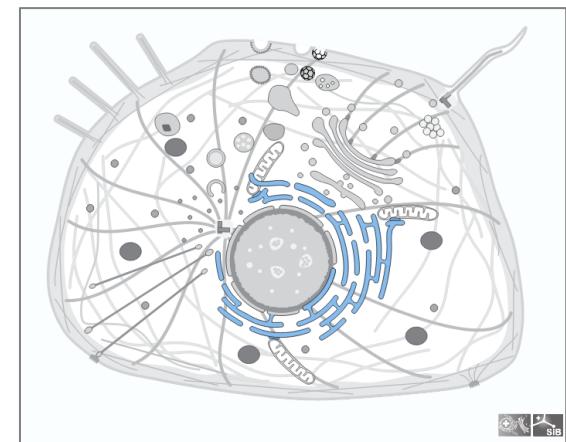


prostaglandin-endoperoxide synthase activity

GO:0004666

Where is it?

GO Cellular component



endoplasmic reticulum membrane

GO:0005789

PTGS1: Prostaglandin G/H synthase 1

X-ray of a GO term

Subontology

Links annotated genes

Term Information

Accession GO:0001516

Unique ID

Name prostaglandin biosynthetic process

Unique Name

Data

Ontology biological_process

Synonyms prostaglandin anabolism, prostaglandin biosynthesis, prostaglandin formation, prostaglandin synthesis

Alternate IDs None

Precise definition

Definition The chemical reactions and pathways resulting in the formation of prostaglandins, any of a group of biologically active metabolites which contain a cyclopentane ring. Source: GOC:ai

Comment None

History See term [history for GO:0001516](#) at QuickGO

Subset None

Related [Link](#) to all **genes and gene products** annotated to prostaglandin biosynthetic process (**excluding "regulates"**).

[Link](#) to all direct and indirect **annotations** to prostaglandin biosynthetic process (**excluding "regulates"**).

[Link](#) to all direct and indirect **annotations download** (limited to first 10,000) for prostaglandin biosynthetic process (**excluding "regulates"**).

[Include "regulates"](#)

For more information, please see the [ontology relation documentation](#).

Inferred tree view (partial snapshot)

- I GO:0033559 unsaturated fatty acid metabolic process
- I GO:0046456 icosanoid biosynthetic process
- I GO:0006692 prostanoid metabolic process
- I GO:0006636 unsaturated fatty acid biosynthetic process
- I GO:0006693 prostaglandin metabolic process
- I GO:0046457 prostanoid biosynthetic process
- ▼ GO:0001516 prostaglandin biosynthetic process
 - P GO:0019371 cyclooxygenase pathway
 - R GO:0031393 negative regulation of prostaglandin biosynthetic process
 - G GO:0031394 positive regulation of prostaglandin biosynthetic process
 - I GO:0100008 regulation of fever generation by prostaglandin biosynthetic process
 - R GO:0031392 regulation of prostaglandin biosynthetic process



What is an annotation?

Association between GO terms and a gene or gene products

An annotation is a statement that a gene product ...

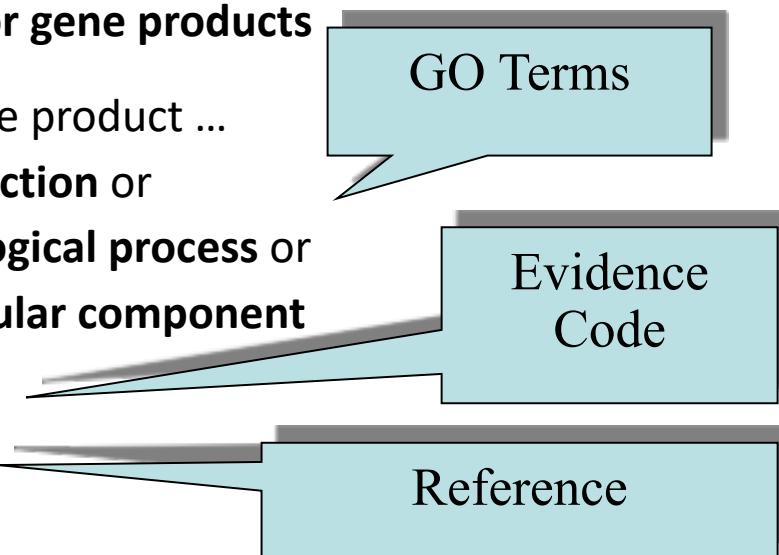
...has a particular **molecular function** or

...is involved in a particular **biological process** or

...is located within a certain **cellular component**

...as determined by a particular method

...as described in a particular source.



	Gene/product	Gene/product name	Annotation qualifier	GO class (direct)	Annotation extension	Contributor	Organism	Evidence	Evidence with	PANTHER family	Type	Isoform	Reference
<input type="checkbox"/>	PTGS1	Prostaglandin G/H synthase 1		prostaglandin biosynthetic process		UniProt	Homo sapiens	ISS	UniProtKB:P22437	prostaglandin g/h synthase pthr11903	protein		GO_REF:0000024
<input type="checkbox"/>	PTGS1	Prostaglandin G/H synthase 1		cyclooxygenase pathway		BHF-UCL	Homo sapiens	IDA		prostaglandin g/h synthase pthr11903	protein		PMID:1380156

Introduction

Experimental Evidence Codes

EXP: Inferred from Experiment

IDA: Inferred from Direct Assay

IPI: Inferred from Physical Interaction

IMP: Inferred from Mutant Phenotype

IGI: Inferred from Genetic Interaction

IEP: Inferred from Expression Pattern

Computational Analysis Evidence Codes

ISS: Inferred from Sequence or Structural Similarity

ISO: Inferred from Sequence Orthology

ISA: Inferred from Sequence Alignment

ISM: Inferred from Sequence Model

IGC: Inferred from Genomic Context

RCA: inferred from Reviewed Computational Analysis

Author Statement Evidence Codes

TAS: Traceable Author Statement

NAS: Non-traceable Author Statement

Curator Statement Evidence Codes

IC: Inferred by Curator

ND: No biological Data available

Automatically-assigned Evidence Codes

IEA: Inferred from Electronic Annotation

Uses of GO

- GO terms are used for functional annotations
 - Genome annotation
 - Protein annotation
- Interpretation of large-scale molecular biology experiments
 - -omics experiment measures levels of thousands of molecules, in different conditions (for example between a cancer cell and a normal cell)
 - GO enrichment analysis identifies relevant groups of genes that function together.



ChEBI

Describes chemical entities.
Collects nomenclature, structures,
synonyms and related chemical
information from a number of
freely accessible sources

The screenshot shows the ChEBI homepage with a dark blue header containing the ChEBI logo and navigation links for Home, Advanced Search, Browse, Documentation, Download, Tools, and About ChEBI. Below the header is a light blue search interface with a search bar, a 'Search' button, and options to search for 'All in ChEBI' or 'Search for ★★★ only'. A note below the search bar states: 'Chemical Entities of Biological Interest (ChEBI) is a freely available dictionary of molecular entities focused on 'small' chemical compounds.' At the bottom of the search interface are links for 'Advanced Search' and 'About ChEBI'.

CHEBI:5855

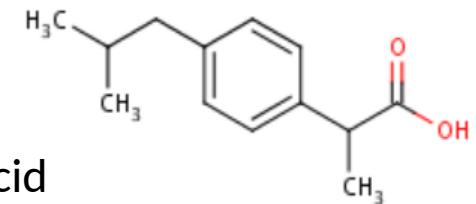
ibuprofen

2-[4-(2-methylpropyl)phenyl]propanoic acid
(RS)-ibuprofen

Motrin

Advil

4-isobutylhydratropic acid



Search with formula

The screenshot shows the ChEBI Advanced Search page. At the top, there is a search bar with examples like "iron*", "InChI=1S/CH4O/c1-2/h2H,1H3", and "caffeine". Below the search bar are links for "Contact us" and "Submit". The main area features a chemical structure editor with a toolbar at the top containing various icons for drawing and manipulating structures. To the right of the editor, there is a vertical column of colored boxes representing elements: H (white), C (light gray), N (light blue), O (light red), S (light green), P (light orange), F (light purple), Cl (light cyan), Br (light pink), I (light yellow), and Generic Groups (light gray). On the left side of the editor, there is a panel with various chemical symbols and structures, including A+, A-, →, [], R1, Q, and Arom. On the right side of the editor, there are search options: "Chemical Structure Search?", "Results per page: 15", "Search for", and "Editor powered by Ketcher. Structure Search powered by OrChem.".



Search

ChEBI > Advanced Search

Chemical Structure Search?

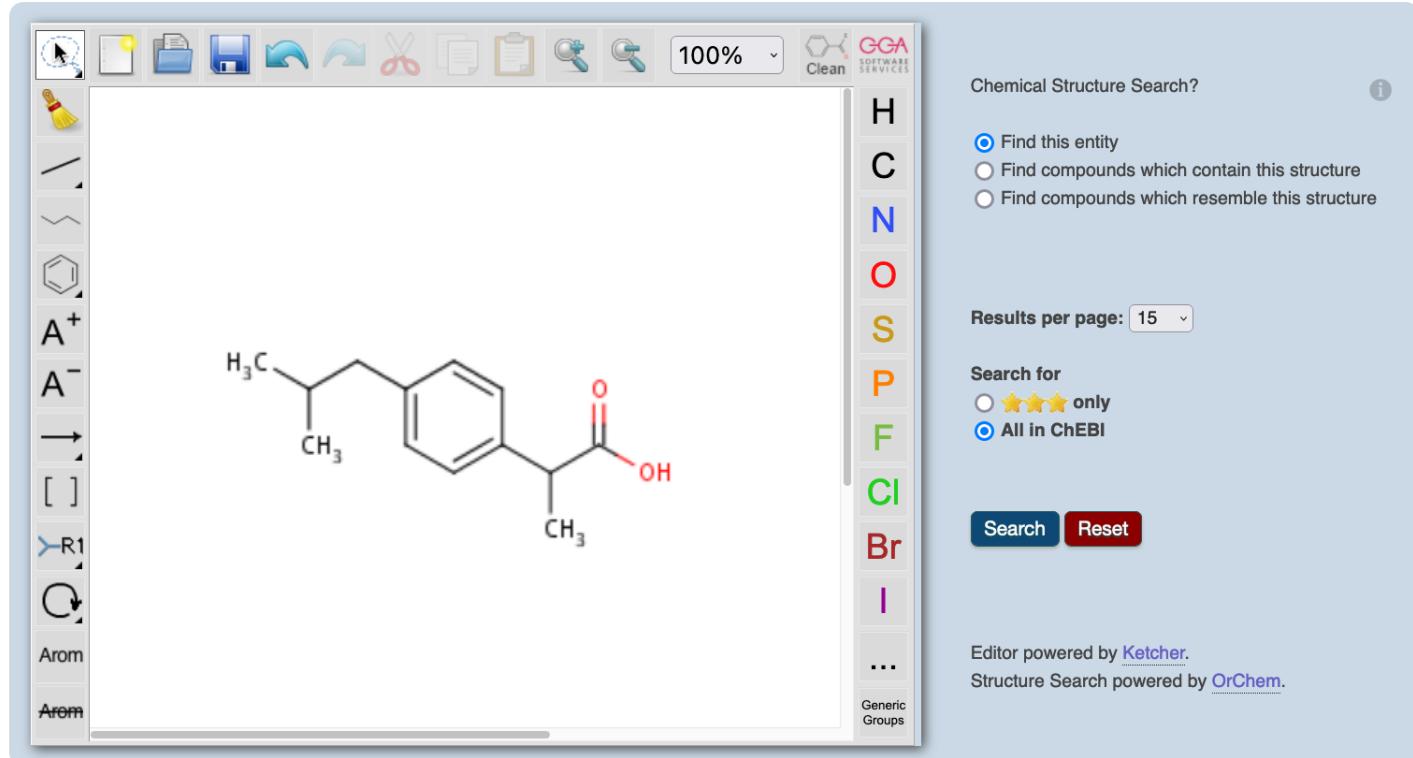
Find this entity
 Find compounds which contain this structure
 Find compounds which resemble this structure

Results per page: 15

Search for:
 ★★★ only
 All in ChEBI

Search **Reset**

Editor powered by [Ketcher](#).
Structure Search powered by [OrChem](#).



The interface includes a toolbar with various icons for file operations, zoom, and search. A vertical element periodic table shows atoms H, C, N, O, S, P, F, Cl, Br, I, and Generic Groups. On the left, there's a drawing tool with a palette of atoms and groups, and a scrollable list of aromatic and generic structures.

Chemical structure input: CC(C)Cc1ccc(CC(C)C(=O)O)cc1

Text Queries: (Example: water)
AND in Category All [+/-](#)

Filter by Ontology Term: (Example "is a CHEBI:15377" or "is a water"; Type the name, and choose one option from the dropdown menu)
AND Select relationship [+/-](#)

Formula: (Example: NaHCO3) *Case sensitive.
AND Formula: [+/-](#)

Average Mass range: (Example: 0 to 30.5)
AND range from to [+/-](#)

Monoisotopic Mass range: (Example: 0 to 30.5)
AND range from to [+/-](#)

Charge range: (Example: -1 to 1)
AND range from to [+/-](#)

Filter by Database:
AND contains a database cross-reference in All databases [+/-](#)

Filter by Chemical Structure:
Results only with chemical structures? Yes: No:

Search **Reset**

CHEBI:5855

<https://www.ebi.ac.uk/chebi/chebiOntology.do?chebivid=CHEBI:5855>

🔍 🔍 🌐

ChEBI Name ibuprofen

ChEBI ID CHEBI:5855

Definition A monocarboxylic acid that is propionic acid in which one of the hydrogens at position 2 is substituted by a 4-(2-methylpropyl)phenyl group.

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

Supplier Information ChemicalBook:CB3699656, ChemicalBook:CB4336930, ChemicalBook:CB92675917, eMolecules:539665, MolPort-001-791-802, Selleckchem:Ibuprofen(Advil)

Download Molfile XML SDF

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

△ CHEBI:36963 organooxygen compound

△ CHEBI:35605 carbon oxoacid

△ CHEBI:33575 carboxylic acid

△ CHEBI:25384 monocarboxylic acid

△ CHEBI:35366 fatty acid

△ CHEBI:26666 short-chain fatty acid

△ CHEBI:30768 propionic acid

△ CHEBI:5855 ibuprofen

← [b] CHEBI:132922 ibuprofen(1-)

→ [b] CHEBI:132922 ibuprofen(1-)

[f] CHEBI:76160 ibuproxam

[f] CHEBI:133199 carboxyibuprofen

△ CHEBI:43415 dexibuprofen

△ CHEBI:47835 levibuprofen

Formula information

Formula	C13H18O2
Net Charge	0
Average Mass	206.28082
Monoisotopic Mass	206.13068
InChI	InChI=1S/C13H18O2/c1-9(2)8-11-4-6-12(7-5-11)10(3)13(14)15/h4-7,9-10H,8H2,1-3H3,(H,14,15)
InChIKey	HEFNNWSXXWATRW-UHFFFAOYSA-N
SMILES	CC(C)Cc1ccc(cc1)C(C)C(O)=O

Molecule roles

Roles Classification ⓘ

radical scavenger

A role played by a substance that can react readily with, and thereby eliminate, radicals.

environmental contaminant

Any minor or unwanted substance introduced into the environment that can have undesired effects.

Bronsted acid

A molecular entity capable of donating a hydron to an acceptor (Bronsted base).

(via [oxoacid](#))

cyclooxygenase 2 inhibitor

A cyclooxygenase inhibitor that interferes with the action of cyclooxygenase 2.

cyclooxygenase 1 inhibitor

A cyclooxygenase inhibitor that interferes with the action of cyclooxygenase 1.

drug allergen

Any drug which causes the onset of an allergic reaction.

xenobiotic

A xenobiotic (Greek, *xenos* "foreign"; *bios* "life") is a compound that is foreign to a living organism. Principal xenobiotics include: drugs, carcinogens and various compounds that have been introduced into the environment by artificial means.

non-narcotic analgesic

A drug that has principally analgesic, antipyretic and anti-inflammatory actions. Non-narcotic analgesics do not bind to opioid receptors.

antipyretic

A drug that prevents or reduces fever by lowering the body temperature from a raised state. An antipyretic will not affect the normal body temperature if one does not have fever. Antipyretics cause the hypothalamus to override an interleukin-induced increase in temperature. The body will then work to lower the temperature and the result is a reduction in fever.

drug allergen

Any drug which causes the onset of an allergic reaction.

geroprotector

Any compound that supports healthy aging, slows the biological aging process, or extends lifespan.

non-steroidal anti-inflammatory drug

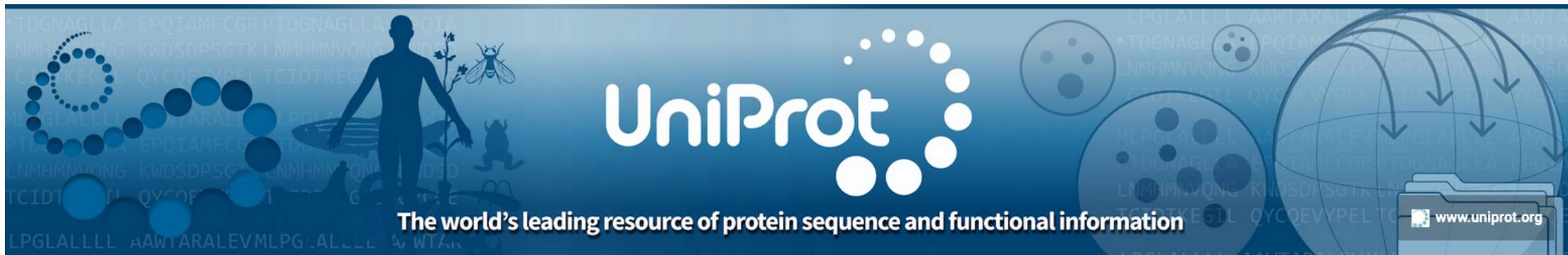
An anti-inflammatory drug that is not a steroid. In addition to anti-inflammatory actions, non-steroidal anti-inflammatory drugs have analgesic, antipyretic, and platelet-inhibitory actions. They act by blocking the synthesis of prostaglandins by inhibiting cyclooxygenase, which converts arachidonic acid to cyclic endoperoxides, precursors of prostaglandins.

non-narcotic analgesic

A drug that has principally analgesic, antipyretic and anti-inflammatory actions. Non-narcotic analgesics do not bind to opioid receptors.



Ontologies in action: case 1 UniProt



A screenshot of the UniProt homepage. At the top, there are links for UniProt (beta), BLAST, Align, Peptide search, ID mapping, and SPARQL. The main heading is "Find your protein". A search bar contains the text "UniprotKB" and a placeholder "Search". Below the search bar, it says "Examples: Insulin, APP, Human, P05067, organism_id:9606". On the right side of the search bar are buttons for "Advanced", "List", and "Search". A note at the bottom left says "UniProt is the world's leading high-quality, comprehensive and freely accessible resource of protein sequence and functional information. Cite UniProt." A feedback link is also present. At the bottom, there are four cards: "Proteins UniProt Knowledgebase" (red background), "Species Proteomes" (pink background), "Protein Clusters UniRef" (orange background), and "Sequence Archive UnPac" (blue background). Below these cards are sections for "Supporting Data" (Diseases, Keywords, Literature Citations, Cross-referenced databases) and "UniRule automatic annotation" (ARBA automatic annotation).

A landing page for the UniProt Knowledgebase. The title is "Proteins UniProt Knowledgebase". It compares two types of data: "Non-redundant Reviewed by curators" (represented by a star icon) and "Redundant Automatically generated" (represented by a document icon). Below this, there are two large icons: "Reviewed Swiss-Prot" (star icon) and "Unreviewed TrEMBL" (document icon).



UniProt protein entry

Prostaglandin G/H synthase 1

<https://beta.uniprot.org/uniprotkb/P23219/entry>

Gene namesⁱ

Name PTGS1 Imported

Synonyms COX1 1 Publication

P23219 · PGH1_HUMAN

Prostaglandin G/H synthase 1 · Homo sapiens (Human) · EC:1.14.99.1 · Gene: PTGS1 (COX1) · 599 amino acids · Evidence at protein level · Annotation score: 5/5

Entry Feature viewer Publications External links History

BLAST Align Download Add Add a publication Entry feedback

Functionⁱ

Dual cyclooxygenase and peroxidase in the biosynthesis pathway of prostanoids, a class of C20 oxylipins mainly derived from arachidonate, with a particular role in the inflammatory response. The cyclooxygenase activity oxygenates arachidonate (AA, C20:4(n-6)) to the hydroperoxy endoperoxide prostaglandin G2 (PGG2), and the peroxidase activity reduces PGG2 to the hydroxy endoperoxide PGH2, the precursor of all 2-series prostaglandins and thromboxanes. This complex transformation is initiated by abstraction of hydrogen at carbon 13 (with S-stereochemistry), followed by insertion of molecular O2 to form the endoperoxide bridge between carbon 9 and 11 that defines prostaglandins. The insertion of a second molecule of O2 (bis-oxygenase activity) yields a hydroperoxy group in PGG2 that is then reduced to PGH2 by two electrons (PubMed:7947975). Involved in the constitutive production of prostanoids in particular in the stomach and platelets. In gastric epithelial cells, it is a key step in the generation of prostaglandins, such as prostaglandin E2 (PGE2), which plays an important role in cytoprotection. In platelets, it is involved in the generation of thromboxane A2 (TXA2), which promotes platelet activation and aggregation, vasoconstriction and proliferation of vascular smooth muscle cells (Probable) 1 Publication 2 Publications

Miscellaneous

The conversion of arachidonate to prostaglandin H2 is a 2 step reaction: a cyclooxygenase (COX) reaction which converts arachidonate to prostaglandin G2 (PGG2) and a peroxidase reaction in which PGG2 is reduced to prostaglandin H2 (PGH2). The cyclooxygenase reaction occurs in a hydrophobic channel in the core of the enzyme. The peroxidase reaction occurs at a heme-containing active site located near the protein surface. The nonsteroidal anti-inflammatory drugs (NSAIDs) binding site corresponds to the cyclooxygenase active site.

Conversion of arachidonate to prostaglandin H2 is mediated by 2 different isozymes: the constitutive PTGS1 and the inducible PTGS2. PTGS1 is expressed constitutively and generally produces prostanoids acutely in response to hormonal stimuli to fine-tune physiological processes requiring instantaneous, continuous regulation (e.g. hemostasis). PTGS2 is inducible and typically produces prostanoids that mediate responses to physiological stresses such as infection and inflammation.

PTGS1 and PTGS2 are the targets of nonsteroidal anti-inflammatory drugs (NSAIDs) including aspirin and ibuprofen. Aspirin is able to produce an irreversible inactivation of the enzyme through a serine acetylation. Inhibition of the PGHs with NSAIDs acutely reduces inflammation, pain, and fever, and long-term use of these drugs reduces fatal thrombotic events, as well as the development of colon cancer and Alzheimer's disease. PTGS2 is the principal isozyme responsible for production of inflammatory prostaglandins. New generation PTGSs inhibitors strive to be selective for PTGS2, to avoid side effects such as gastrointestinal complications and ulceration.

Catalytic Activity

(5Z,8Z,11Z,14Z)-eicosatetraenoate + AH2 + 2 O2 = A + H2O + prostaglandin H2 1 Publication

This reaction proceeds in the forward direction. 1 Publication

EC:1.14.99.1 (UniProtKB | ENZYME | Rhea)

Source: Rhea 23728

Function annotation

Catalytic activity uses ChEBI for reaction description

Feedback

Help

Chemical Reaction Diagram:

(5Z,8Z,11Z,14Z)-eicosatetraenoate (CHEBI:32395) + AH2 (CHEBI:17499) + 2 O2 (CHEBI:15379) = A (CHEBI:13193) + H2O (CHEBI:15377) + prostaglandin H2 (CHEBI:57405)

zoom zoom zoom zoom zoom zoom

UniProt protein entry

Prostaglandin G/H synthase 1

<https://beta.uniprot.org/uniprotkb/P23219/entry>

Function

Names & Taxonomy

Subcellular Location

Disease & Drugs

PTM/Processing

Expression

Interaction

Structure

Family & Domains

Sequence & Isoforms

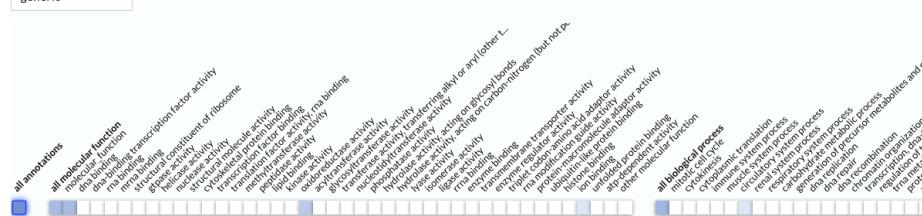
Similar Proteins

Gene Ontology Annotation

GO Annotations¹

Slimming set:

generic ▾



ASPECT	TERM	TERM SOURCE	ASSERTION SOURCE
Cellular Component	proreceptor outer segment	IEA:Ensembl	
Molecular Function	heme binding	IEA:InterPro	
Molecular Function	metal ion binding	IEA:UniProtKB-KW	
Molecular Function	oxidoreductase activity, acting on single donors with incorporation of molecular oxygen, incorporation of two atoms of oxygen		Manual Assertion Based On Experiment IBA:GO_Central
Molecular Function	peroxidase activity	IEA:UniProtKB-KW	
Molecular Function	prostaglandin-endoperoxide synthase activity		Manual Assertion Based On Experiment IDA:BHF-UCL
Biological Process	cyclooxygenase pathway		Manual Assertion Based On Experiment IDA:BHF-UCL
Biological Process	inflammatory response	IEA:InterPro	
Biological Process	prostaglandin biosynthetic process	ISS:UniProtKB	
Biological Process	regulation of blood pressure	ISS:UniProtKB	
Biological Process	regulation of cell population proliferation	IEA:Ensembl	
Biological Process	response to oxidative stress	IEA:InterPro	

Cross-references to external databases

Keywords¹

Molecular function	#Dioxygenase #Oxidoreductase #Peroxidase
Biological process	#Fatty acid biosynthesis #Fatty acid metabolism #Lipid biosynthesis #Lipid metabolism #Prostaglandin biosynthesis #Prostaglandin metabolism
Ligand	#Heme #Iron #Metal-binding

Enzyme and pathway databases

BRENDA	1.14.99.1 ↗ 2681
PathwayCommons	P23219 ↗
Reactome	R-HSA-140180 ↗ COX reactions R-HSA-2162123 ↗ Synthesis of Prostaglandins (PG) and Thromboxanes (TX)

SABIO-RK	P23219 ↗
SIGNOR	P23219 ↗
SignalLink	P23219 ↗
UniPathway	UPA00662
ENZYME	Search... ↗

Protein family/group databases

MoonDB	P23219 ↗ Curated
--------	------------------

PeroxiBase	3320 ↗ HsPGHS01
------------	-----------------

Chemistry

SwissLipids	SLP:000001103 ↗
-------------	-----------------



Enzymatic reactions described using ChEBI

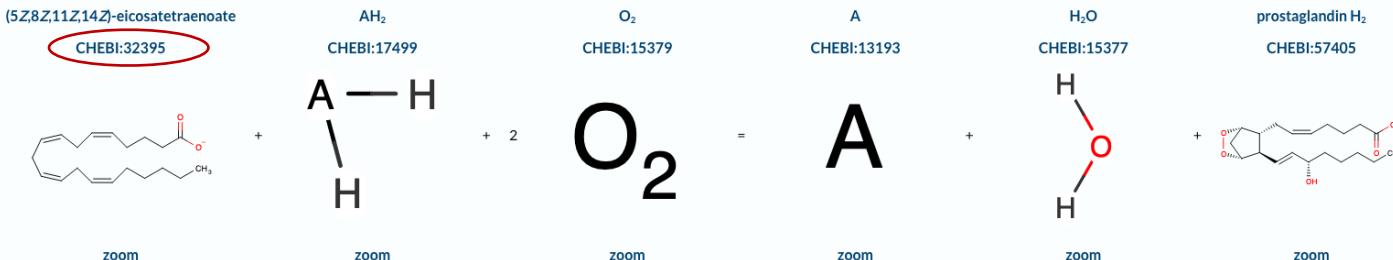
Catalytic Activity

(5Z,8Z,11Z,14Z)-eicosatetraenoate + AH₂ + 2 O₂ = A + H₂O + prostaglandin H₂ [1 Publication]

This reaction proceeds in the forward direction. [1 Publication]

EC:1.14.99.1 (UniProtKB | ENZYME | Rhea)

Source: Rhea 23728



Prostaglandin G/H synthase 1

<https://beta.uniprot.org/uniprotkb/P23219/entry>

<https://www.ebi.ac.uk/chebi/chebiOntology.do?chebid=CHEBI:32395>

ChEBI Name	arachidonate
ChEBI ID	CHEBI:32395
Chemical Structure	A long-chain fatty acid anion resulting from the removal of a proton from the carboxy group of arachidonic acid.
Stars	★★★ This entity has been manually annotated by the ChEBI Team.
Secondary ChEBI IDs	CHEBI:13852, CHEBI:22607
Supplier Information	No supplier information found for this compound.
Download	Molfile XML SDF



Using the power of ontologies to find enzymes with reactions involving arachidonate

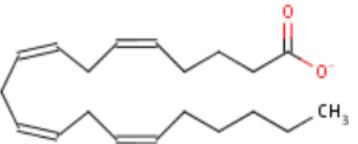
🔍 🔍 ⚡

ChEBI Name **arachidonate**

ChEBI ID **CHEBI:32395**

IUPAC Name ⓘ (5Z,8Z,11Z,14Z)-icos-5,8,11,14-tetraenoate

Synonyms ⓘ (20:4n6) (5Z,8Z,11Z,14Z)-eicosatetraenoate

Chemical Structure: 

Formula C₂₀H₃₁O₂

Net Charge -1

Average Mass 303.45894

Monoisotopic Mass 303.23295

InChI IC=IS/C20H32O2/c1-2-3-4-5-6-7-8-9-10-11-12-13-14-15-16-17-18-19-20(21)22/h6-7,9-10,12-13,15-16H,2-5,8,11,14,17-19H2,1H3,(H,21,22)/p-1/b7-6-,10-9-,13-12-,16-15-

InChIKey YZXBA PSDXZZRGB-DOFZRALJSA-M

SMILES CCCCC\C=C/C\C=C/C\C=C/C\C=C/CCCC([O-])=O

Advanced Search

Searching in
UniProtKB

Evidence
Catalytic Activity ▾ CHEBI:32395 ✕ Any assertion method ▾ Remove

Add Field

Cancel Search

Evidence
Catalytic Activity ▾ arachidonate ✕ Any assertion method ▾ Remove

Add Field

Cancel Search

Evidence
Catalytic Activity ▾ OFZRALJSA-M ✕ Any assertion method ▾ Remove

Add Field

Cancel Search



Proteins in UniProt with arachidonate in enzymatic reaction (human)

UniProtKB 75 results

UniProtKB (cc_catalytic_activity:"CHEBI:32395") Advanced | List Search

BLAST Align Peptide search ID mapping SPARQL UniProtKB (cc_catalytic_activity:"CHEBI:32395") Advanced | List Search

Status: Reviewed (Swiss-Prot) (75) Model organisms: Human (75) Taxonomy: Filter by taxonomy

Proteins with: 3D structure (30), Active site (40), Activity regulation (40), Alternative products (isoforms) (43), Alternative splicing (43). More items

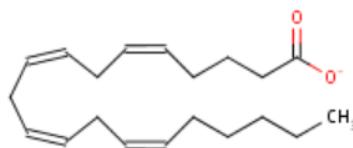
Protein existence: Protein level (75)

Annotation score: 5 / 75

Entry	Entry Name	Protein Names	Gene Names	Organism	Gene Ontology - Biological Process	Gene Ontology - Molecular Function
Q68DD2	PA24F_HUMAN	Cytosolic phospholipase A2 zeta[...]	PLA2G4F	Homo sapiens (Human)	arachidonic acid secretion, cellular response to antibiotic, cellular response to organic cyclic compound, glycerophospholipid catabolic process, phosphatidylglycerol acyl-chain remodeling, prostaglandin biosynthetic process	calcium ion binding, calcium-dependent phospholipase A2 activity, calcium-dependent phospholipid binding, lysophospholipase activity, phospholipase A1 activity, phospholipase A2 activity
P23219	PGH1_HUMAN	Prostaglandin G/H synthase 1[...]	PTGS1[...]	Homo sapiens (Human)	cyclooxygenase pathway, inflammatory response, prostaglandin biosynthetic process, regulation of blood pressure, regulation of cell population proliferation, response to oxidative stress	heme binding, metal ion binding, oxidoreductase activity, acting on single donors with incorporation of molecular oxygen, incorporation, peroxidase activity, prostaglandin-endoperoxide synthase activity

UNIVERSITY OF DELAWARE CENTER FOR BIOINFORMATICS & COMPUTATIONAL BIOLOGY

Using the power of ontologies to find enzymes with reactions involving polyunsaturated fatty acid anions



ChEBI Name **arachidonate**

ChEBI ID **CHEBI:32395**

- △ CHEBI:28868 fatty acid anion
- △ CHEBI:2580 unsaturated fatty acid anion
- △ **CHEBI:76567 polyunsaturated fatty acid anion**
- △ CHEBI:134019 hydroperoxy polyunsaturated fatty acid anion
- △ CHEBI:59720 HPETE anion
- △ CHEBI:146291 11-HPETE(1-)

CHEBI:32395 arachidonate

Search UniProt for ChEBI term CHEBI:76567 in catalytic activity

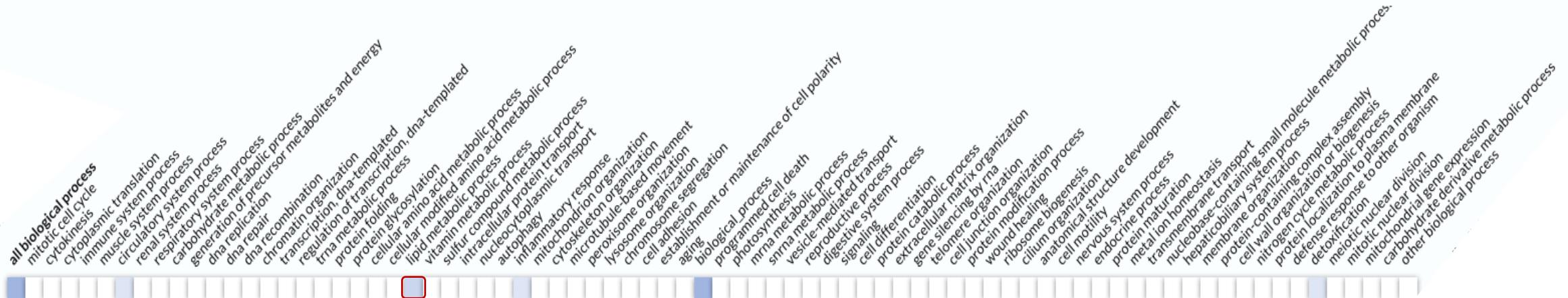
Filter results to get: human and reviewed entries

Can you spot PTGS1 in your results?

GO slims offer a way to group terms

Biological Process	cyclooxygenase pathway	Manual Assertion Based On Experiment	IDA:BHF-UCL
Biological Process	inflammatory response	IEA:InterPro	
Biological Process	prostaglandin biosynthetic process	ISS:UniProtKB	
Biological Process	regulation of blood pressure	ISS:UniProtKB	
Biological Process	regulation of cell population proliferation	IEA:Ensembl	
Biological Process	response to oxidative stress	IEA:InterPro	

GO slim



Enrichment analysis

- GO slims are useful for enrichment analysis where there is usually a large number of genes with a number GO annotations at varying level of granularity

Example: gene list from UniProt search with CHEBI:76567

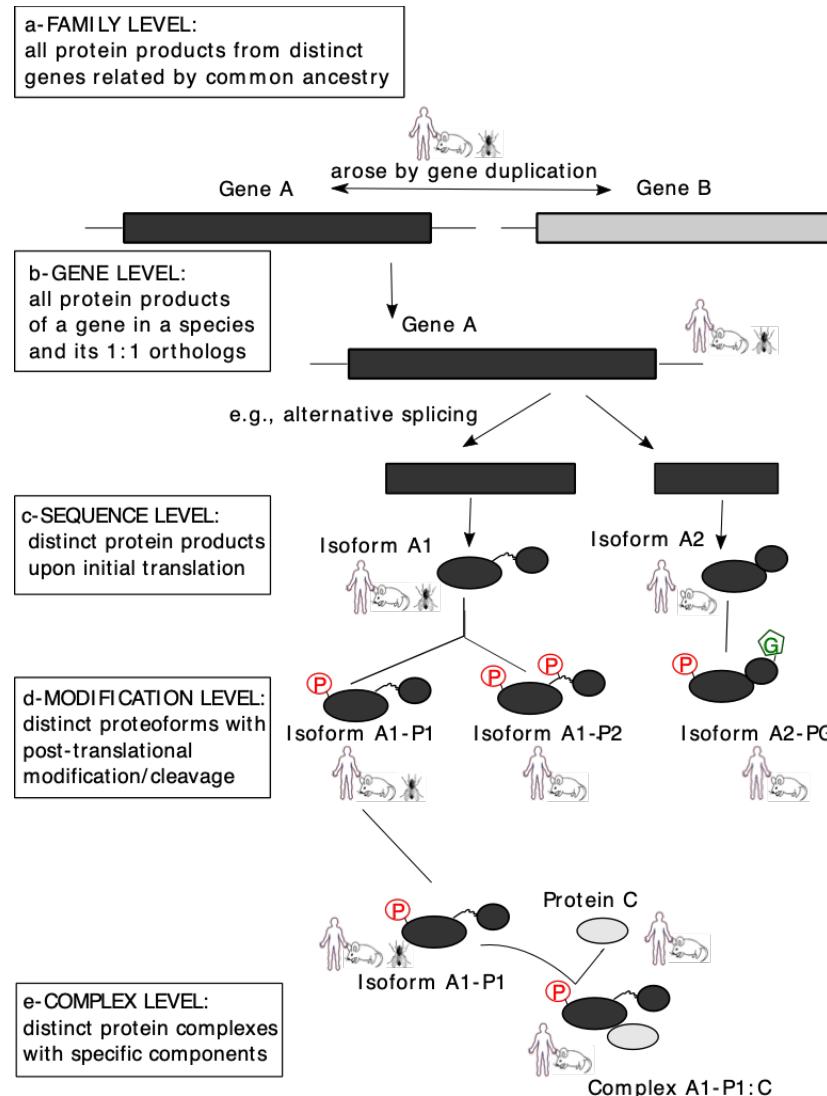
Complete GO BP

displaying only results for FDR P < 0.05, click here to display all results						
Homo sapiens (REF)	#	# expected	Fold Enrichment	+/-	raw P value	FDR
GO biological process complete	3	3	.02	> 100	+	2.97E-06
pressure natriuresis	15	3	.08	37.10	+	1.15E-04
↳ renal system process involved in regulation of blood volume	375	9	2.02	4.45	+	2.27E-04
↳ regulation of body fluid levels	3730	50	20.11	2.49	+	7.30E-11
↳ regulation of biological quality	24	3	.13	23.19	+	4.00E-04
↳ renal system process involved in regulation of systemic arterial blood pressure	192	6	1.04	5.80	+	6.99E-04
↳ regulation of blood pressure	405	10	2.18	4.58	+	8.01E-05
↳ blood circulation	495	14	2.67	5.25	+	5.86E-07
↳ circulatory system process	2	2	.01	> 100	+	1.70E-04
lipoxin B4 metabolic process	103	51	.56	91.84	+	2.59E-81
↳ unsaturated fatty acid metabolic process	327	85	1.76	48.22	+	2.79E-124
↳ fatty acid metabolic process	923	104	4.98	20.90	+	3.18E-128
↳ cellular lipid metabolic process	1215	109	6.55	16.64	+	5.20E-129
↳ lipid metabolic process	7091	110	38.23	2.88	+	1.52E-49
↳ primary metabolic process	8019	110	43.23	2.54	+	9.66E-44
↳ metabolic process	7577	110	40.85	2.69	+	2.04E-46
↳ organic substance metabolic process	7138	109	38.48	2.83	+	3.21E-47
↳ cellular metabolic process	15176	111	81.82	1.36	+	3.63E-15
↳ cellular process	499	85	2.69	31.60	+	3.07E-110
↳ monocarboxylic acid metabolic process	816	90	4.40	20.46	+	6.36E-103
↳ carboxylic acid metabolic process	842	90	4.54	19.83	+	9.01E-102
↳ oxoacid metabolic process	864	90	4.66	19.32	+	7.99E-101
↳ organic acid metabolic process	1616	100	8.71	11.48	+	9.33E-96
↳ small molecule metabolic process	6	6	.03	> 100	+	1.87E-11
↳ lipoxin metabolic process	73	18	.39	45.74	+	1.09E-23
↳ fatty acid derivative metabolic process	485	32	2.61	12.24	+	2.45E-25
↳ organic hydroxy compound metabolic process						

Panther slim GO BP

displaying only results for FDR P < 0.05, click here to display all results						
Homo sapiens (REF)	#	# expected	Fold Enrichment	+/-	raw P value	FDR
PANTHER GO-Slim Biological Process	6	4	.03	> 100	+	1.60E-07
↳ unsaturated fatty acid biosynthetic process	21	13	.11	> 100	+	1.24E-21
↳ unsaturated fatty acid metabolic process	107	26	.58	45.07	+	7.61E-34
↳ fatty acid metabolic process	152	26	.82	31.73	+	2.49E-30
↳ monocarboxylic acid metabolic process	314	28	1.69	16.54	+	1.79E-25
↳ carboxylic acid metabolic process	318	28	1.71	16.33	+	2.47E-25
↳ oxoacid metabolic process	333	35	1.80	19.50	+	2.94E-34
↳ organic acid metabolic process	5685	67	30.65	2.19	+	8.51E-13
↳ organic substance metabolic process	5907	68	31.85	2.14	+	1.74E-12
↳ metabolic process	5332	67	28.75	2.33	+	2.83E-14
↳ cellular metabolic process	648	38	3.49	10.88	+	1.69E-28
↳ small molecule metabolic process	367	50	1.98	25.27	+	3.13E-55
↳ cellular lipid metabolic process	419	51	2.26	22.58	+	3.78E-54
↳ lipid metabolic process	5251	52	28.31	1.84	+	1.65E-06
↳ primary metabolic process	41	6	.22	27.14	+	1.86E-07
↳ fatty acid biosynthetic process	46	6	.25	24.19	+	3.46E-07
↳ monocarboxylic acid biosynthetic process	88	8	.47	16.86	+	4.75E-08
↳ carboxylic acid biosynthetic process	88	8	.47	16.86	+	4.75E-08
↳ organic acid biosynthetic process	154	8	.83	9.64	+	2.64E-06
↳ small molecule biosynthetic process	212	10	1.14	8.75	+	3.30E-07
↳ lipid biosynthetic process	6	4	.03	> 100	+	1.60E-07
↳ arachidonic acid metabolic process	19	9	.10	87.86	+	1.67E-14
↳ icosanoid metabolic process	19	14	.10	> 100	+	5.23E-24
↳ long-chain fatty acid metabolic process						

Ontologies in action: case 2 PRotein Ontology



proconsortium.org

Ontology to formally represent proteoforms and protein complexes

Proteoforms may have distinct function, and localization.
Annotation at the gene level is not sufficient

PRO provides level of granularity needed for more specific annotations

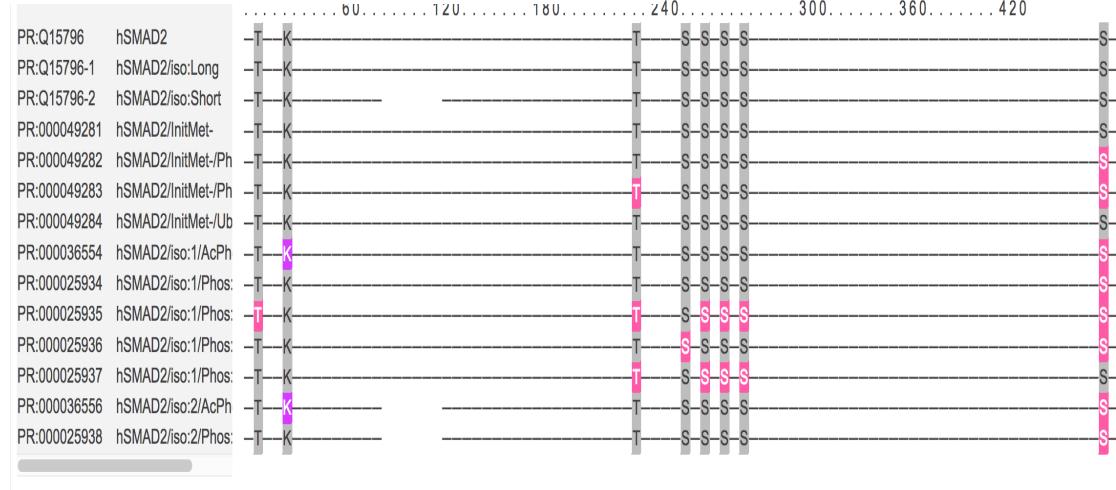
PRO connects evolutionary related proteoforms



Example human SMAD2

http://purl.obolibrary.org/obo/PR_Q15796

Proteoforms from human SMAD2



PRO Hierarchy

PR:000000027 smad protein
PR:000000066 TGF-beta receptor-regulated smad protein
PR:000000364 smad2
PR:Q15796-1 mothers against decapentaplegic homolog 2 (human)
PR:Q15796-1 mothers against decapentaplegic homolog 2 isoform Long (human)
PR:000036554 smad2 isoform 1 acetylated and phosphorylated 1 (human)
PR:000025934 smad2 isoform 1 phosphorylated 1 (human)
PR:000025935 smad2 isoform 1 phosphorylated 2 (human)
PR:000025936 smad2 isoform 1 phosphorylated 3 (human)
PR:000025937 smad2 isoform 1 phosphorylated 4 (human)
PR:000045371 smad2 isoform 1 unphosphorylated 1 (human)
PR:Q15796-2 mothers against decapentaplegic homolog 2 isoform Short (human)
PR:000036556 smad2 isoform 2 acetylated and phosphorylated 1 (human)
PR:000025938 smad2 isoform 2 phosphorylated 1 (human)
PR:000045494 smad2 isoform 2 unphosphorylated 1 (human)
PR:000049281 mothers against decapentaplegic homolog 2, initiator methionine removed form (human)
PR:000049282 mothers against decapentaplegic homolog 2, initiator methionine removed phosphorylated 1 (human)
PR:000049283 mothers against decapentaplegic homolog 2, initiator methionine removed phosphorylated 5 (human)
PR:000049284 mothers against decapentaplegic homolog 2, initiator methionine removed ubiquitinated form (human)
PR:000049284 mothers against decapentaplegic homolog 2, initiator methionine removed ubiquitinated form (human)
PR:000025960 smad2 isoform 1 unphosphorylated 1 (human)
PR:000045371 smad2 isoform 1 unphosphorylated 1 (human)
PR:000045494 smad2 isoform 2 unphosphorylated 1 (human)

Annotations for specific proteoforms

PRO Term	GO Annotation	Evidence
PR:000045371 hSMAD2/iso:1/UnPhos:1 Ser-465/Ser-467, PR:000026291	has_function GO:0005515 protein binding	PMID:8980228
	located_in GO:0005737 cytoplasm	PMID:15280432, PMID:17074053
	located_in GO:0005769 early endosome	PMID:12356868
	participates_in GO:0007179 transforming growth factor beta receptor signaling pathway	PMID:8752209
PR:000025946 smad2-smad4 protein complex 3 (human)	located_in GO:0005634 nucleus	PMID:12193595
	participates_in GO:0030511 positive regulation of transforming growth factor beta receptor signaling pathway	PMID:12193595
PR:000025947 smad2-smad4 protein complex 4 (human)	located_in GO:0005737 cytoplasm	PMID:11027280
	participates_in GO:0030512 negative regulation of transforming growth factor beta receptor signaling pathway	PMID:11027280
PR:000025952 smad2-smad4-ski protein complex 2 (human)	located_in GO:0005737 cytoplasm	PMID:12732634
	participates_in GO:0030512 negative regulation of transforming growth factor beta receptor signaling pathway	PMID:12732634
PR:000025951 smad2-smad4-ski protein complex 1 (human)	located_in GO:0005634 nucleus	PMID:10485843
	participates_in GO:0030512 negative regulation of transforming growth factor beta receptor signaling pathway	PMID:10485843
PR:000025938 hSMAD2/iso:2/Phos:1 Ser-435/Ser-437, MOD:00046	has_function GO:0003700 transcription factor activity, sequence-specific DNA binding	PMID:9873005
	has_function GO:0003672 DNA binding	PMID:14701940, PMID:9873005
	has_function GO:0030618 transforming growth factor beta receptor, pathway-specific cytosolic mediator activity	PMID:14701940, PMID:9873005
	participates_in GO:0007179 transforming growth factor beta receptor signaling pathway	PMID:9873005

Similar to GO, each annotation has an evidence code and an evidence source



Interested in phosphorylated proteoforms in human



NIH grant: 5R01GM080646-09

PRO provides an ontological representation of protein-related entities by explicitly defining them and showing the relationships between them. Each PRO term represents a distinct class of entities (including specific modified forms, orthologous isoforms, and protein complexes) ranging from the taxon-neutral to the taxon-specific (e.g. the entity representing all protein products of the human SMAD2 gene is described in PR:Q15796; one particular human SMAD2 protein form, phosphorylated on the last two serines of a conserved C-terminal SSxS motif is defined by [PR:000025934](#)). Current release: 54.0.

- [Consortium](#)
- [Dissemination](#)
- [PRO Wiki](#)
- [Documentation](#)
- [Downloads](#)
- [PRO tutorial](#)
- [PRO Publications](#)
- [PRO Statistics](#)

[Browse PRO](#)

-- Quick Browse

Example: methylated ([sample output](#))

Annotation: [RACE-PRO](#) [PRO tracker](#)

[Retrieve a PRO entry \(enter a PRO ID\):](#)

Example: PR:000025934 ([sample output](#))

[Search PRO \(enter text or ID\):](#)

Example: smad ([sample output](#))



[PRO SPARQL](#)



Interested in phosphorylated proteoforms in human



Phosphorylated forms Quick link to phosphorylated forms

[search] Any field AND Any field
homo sapiens

Narrow search to Homo sapiens

Display Options

1588 entries | 32 pages | 50 / page | 1 | 2 | 3 | 4 | 5 |

Save Result As:

click to show: selected [Hierarchy](#) | selected [OBO / PAF](#) OR related [OBO/PAF/Cytoscape View](#)

<input type="checkbox"/> PRO ID	PRO Name	PRO Term Definition	Category	Parent
<input type="checkbox"/> PR:000049554	tuberin phosphorylated 2 (human)	A tuberin (human) that has been post-translationally modified to include phosphoserine at position 1387 of the amino acid sequence represented by UniProtKB: P49815 . UniProtKB: P49815 , Ser-1387, MOD:00046. [PRO:DNx, Reactome:R-HSA-3132769]	organism-modification	PR:000037070 ; PR:P49815
<input type="checkbox"/> PR:000048628	nucleophosmin phosphorylated 2 (human)	A nucleophosmin (human) that has been post-translationally modified to include phosphoserine at position 4 of the amino acid sequence represented by UniProtKB: P06748 . UniProtKB: P06748 , Ser-4, MOD:00046. [PRO:DNx, Reactome:R-HSA-6801679]	organism-modification	PR:000037070 ; PR:P06748
<input type="checkbox"/> PR:000048090	kalirin phosphorylated form (human)	A kalirin (human) that has been post-translationally modified to include at least one phosphorylated residue. UniProtKB: Q60229 , MOD:00696. [PRO:DNx, Reactome:R-HSA-5692652]	organism-modification	PR:000037070 ; PR:Q60229
<input type="checkbox"/> PR:000045959	fascin phosphorylated 1 (human)	A fascin (human) that has been phosphorylated at Ser-39. UniProtKB: Q16658 , Ser-39, MOD:00046. [iPTMnet:Q16658, PMID: 8999969]	organism-modification	PR:000037070 ; PR:Q16658



Phosphorylated forms

(click here for Batch Retrieval page)

Any field

AND

Ontology term

homo sapiens

Nucleus

31 entries | 1 page | 50 / page |

Save

Now adding with annotation nucleus

click to show: selected [Hierarchy](#) | selected [OBO / PAF](#) OR related [OBO / PAF / Cytoscape View](#)

PRO ID	PRO Name	PRO Term Definition	Category	Parent	Annotation							
					Modifier	Relation	Ontology ID	Ontology Term	Relative To	Interaction With	Evidence Source	Evidence Code
<input type="checkbox"/> PR:000035486	breast cancer type 2 susceptibility protein phosphorylated 1 (human)	A breast cancer type 2 susceptibility protein phosphorylated 1 in human. [PMID: 10749118 , PRO:KER]	organism-modification	PR:000035485 ; PR:P51587		located_in	GO:0005634	nucleus			PMID:10749118	EXP
<input type="checkbox"/> PR:000035770	catenin beta-1 isoform 1 phosphorylated 2 (human)	A catenin beta-1 isoform 1 phosphorylated 2 in human. Tyr-64 is the major phosphorylation site for PTK6. UniProtKB: P35222-1 , Tyr-64, MOD:00048 Tyr-142, MOD:00048 (Tyr-331, MOD:00048 or Tyr-333, MOD:00048). [PMID: 20026641 , PRO:IC]	organism-modification	PR:000035769 ; PR:P35222-1		located_in	GO:0005634	nucleus			PMID:20026641	EXP
<input type="checkbox"/> PR:000026480	transcription factor PML isoform 1 phosphorylated 1 (human)	A transcription factor PML isoform 1 phosphorylated 1 in human. UniProtKB: P29590-1 , Ser-403, MOD:00046 Thr-409	organism-modification	PR:000026476 ; PR:P29590-1		located_in	GO:0005634	nucleus			PMID:20832753	EXP

Ontology browser and repositories

BioPortal

<https://bioportal.bioontology.org/>

The screenshot shows the BioPortal homepage with the following sections:

- Search for a class:** A text input field with placeholder "Enter a class, e.g. Melanoma" and a blue search icon.
- Find an ontology:** A text input field with placeholder "Start typing ontology name, then choose from" and a blue search icon. Below it is a teal button labeled "Browse Ontologies".
- Ontology Visits (March 2022):** A horizontal bar chart showing the number of visits for various ontologies. The x-axis ranges from 0 to 35,000. The data is as follows:

Ontology	Visits (March 2022)
MEDDRA	~34,000
SNOMEDCT	~13,000
RXNORM	~8,000
NDDF	~4,000
DTO	~1,000
- BioPortal Statistics:** A table showing the count of various entities.

Category	Count
Ontologies	978
Classes	13,918,657
Properties	36,286
Mappings	79,636,946

Search terms across ontologies/
vocabularies

Browse ontologies, provides usage
information

Annotate text with ontology terms



Term search across ontologies

glioblastoma - Human Disease Ontology (DOID)

http://purl.obolibrary.org/obo/DOID_3068

A malignant astrocytoma characterized by the presence of small areas of necrotizing tissue that is surrounded by anaplastic cells as well as the presence of hyperplastic blood vessels, and that ...

[details](#) [visualize](#)

Reuses in other ontologies

glioblastoma - BioAssay Ontology (BAO)

http://purl.obolibrary.org/obo/DOID_3068

An astrocytoma characterized by the presence of small areas of necrotizing tissue that is surrounded by anaplastic cells as well as the presence of hyperplastic blood vessels, and that has_material_basis_in ...

[details](#) [visualize](#)

glioblastoma - Cell Line Ontology (CLO)

http://purl.obolibrary.org/obo/DOID_3068

A malignant astrocytoma characterized by the presence of small areas of necrotizing tissue that is surrounded by anaplastic cells as well as the presence of hyperplastic blood vessels, and that ...

[details](#) [visualize](#)

glioblastoma - Human Health Exposure Analysis Resource (HHEAR)

http://purl.obolibrary.org/obo/DOID_3068

A malignant astrocytoma characterized by the presence of small areas of necrotizing tissue that is surrounded by anaplastic cells as well as the presence of hyperplastic blood vessels, and that ...

[details](#) [visualize](#) [8 more from this ontology](#)

glioblastoma - Natural Products Ontology (NATPRO)

http://www.owl-ontologies.com/NPontology.owl#DOID_3068

[details](#) [visualize](#)

glioblastoma multiforme - Drug Target Ontology (DTO)

http://purl.obolibrary.org/obo/DOID_3068

An astrocytoma characterized by the presence of small areas of necrotizing tissue that is surrounded by anaplastic cells as well as the presence of hyperplastic blood vessels, and that has_material_basis_in ...

[details](#) [visualize](#) [2 more from this ontology](#)

glioblastoma multiforme - EpilepsyOntology (EPIO)

http://purl.obolibrary.org/obo/DOID_3068

An astrocytoma characterized by the presence of small areas of necrotizing tissue that is surrounded by anaplastic cells as well as the presence of hyperplastic blood vessels, and that has_material_basis_in ...

[details](#) [visualize](#)

Annotate text with ontologies

Annotator

Get annotations for biomedical text with classes from the ontologies [?](#)

Glioblastoma or glioma is the most common malignant brain tumor.

Patients have a prognosis of approximately 15 months, despite the current aggressive treatment. Neurokinin-1 receptor (NK-1R) occurs naturally in human glioma, and it is necessary for the tumor development. OBJECTIVE: The purpose of the study was to increase the knowledge about the involvement of the substance P (SP)/NK-1R system in human glioma. METHODS: Cellular localization of NK-1R and SP was studied in GAMG and U-87 MG glioma cell lines by immunofluorescence. The contribution of both SP and NK-1R to the viability of these cells was also assessed after applying the tachykinin 1 receptor (TAC1R) or the tachykinin 1 (TAC1) small interfering RNA gene silencing method, respectively. RESULTS: Both SP and the NK-1R (full-length and truncated isoforms) were localized in the nucleus and cytoplasm of GAMG and U-87 MG glioma cells. The presence of full-length NK-1R isoform was mainly observed in the nucleus, while the level of truncated isoform was higher in the cytoplasm. Cell proliferation was decreased when glioma cells were transfected with TAC1R siRNA, but not with TAC1. U-87 MG cells were more sensitive to the effect of the TAC1R inhibition than GAMG cells.

[insert sample text](#)

Match longest only Match partial words Include mappings Exclude numbers Exclude synonyms

Select ontologies

Human Disease Ontology (DOID) [x](#)
[clear selection](#) [select from list](#)

Select UMLS semantic types [Start typing to select UMLS semantic types](#)

Include ancestors up to level

none

[Get Annotations](#)

Annotations

total results 10 (direct 10 / ancestor 0 / mapping 0

CLASS filter	ONTOLOGY filter	TYPE filter	CONTEXT filter	MATCHED CLASS filter	MATCHED ONTOLOGY filter
glioblastoma	Human Disease Ontology	direct	Glioblastoma or glioma is ...	glioblastoma	Human Disease Ontology
brain	Human Disease Ontology	direct	... common malignant brain tumor. ...	brain	Human Disease Ontology



Ontology browser and repositories

OLS



glioblastoma

Exact match Obsolete terms

Term type
Filter by type
class 549

Ontology
Filter by ontology
NCIT 284
BTO 57
CLO 53
EFO 40
MONDO 24
ORDO 16

Search results for glioblastoma
Showing 1 to 10 of 549 results

Glioblastoma NCIT:C3058
http://purl.obolibrary.org/obo/NCIT_C3058
The most malignant astrocytic tumor (WHO grade IV). It is composed of poorly differentiated i characterized by the presence of cellular polymorphism, nuclear atypia, brisk mitotic activity, proliferation and necrosis. It typically affects adults and is preferentially located in the cerebra diffuse astrocytoma WHO grade II or anaplastic astrocytoma (secondary glioblastoma, IDH-m

Ontology: NCI Thesaurus OBO Edition NCIT

glioblastoma MONDO:0018177
http://purl.obolibrary.org/obo/MONDO_0018177
Ontology: Mondo Disease Ontology MONDO

<https://www.ebi.ac.uk/ols/index>

Search terms across ontologies/vocabularies

Browse ontologies, provides usage information

OLS / Human Disease Ontology DOID / DOID:3068 Copy

glioblastoma

http://purl.obolibrary.org/obo/DOID_3068 Copy

A malignant astrocytoma characterized by the presence of small areas of necrotizing tissue that is surrounded by anaplastic cells as well as the presence of hyperplastic blood vessels, and that has_material_basis_in abnormally proliferating cells derives_from multiple cell types including astrocytes and oligodendrocytes. [url:https://www.ncbi.nlm.nih.gov/pubmed/23029035 url:http://en.wikipedia.org/wiki/Glioblastoma_multiforme url:https://www.ncbi.nlm.nih.gov/pubmed/20129251 http://purl.obolibrary.org/obo/ECO_0007638 http://purl.obolibrary.org/obo/ECO_0007645 url:http://cancergenome.nih.gov/cancersselected/glioblastomamultiforme]

Tree view **Term mappings**

- disease
 - disease of anatomical entity
 - musculoskeletal system disease
 - connective tissue disease
 - bone disease
 - bone cancer
 - high grade glioma
 - malignant astrocytoma
 - glioblastoma
 - IDH-mutant glioblastoma
 - IDH-wildtype glioblastoma
 - giant cell glioblastoma
 - glioblastoma classical subtype
 - glioblastoma mesenchymal subtype

Graph view
Reset tree
Show all siblings
Preferred root terms
All terms

Term information

database cross reference

 - MESH:D005909
 - NCI:C9094
 - GARD:2491
 - NCI:C39750
 - UMLS_CUI:C1514422
 - NCI:C3058
 - UMLS_CUI:C0017636
 - SNOMEDCT_US_2021_09_01:63634009
 - UMLS_CUI:C0278878
 - ICDO:9440/3

Ontologies and the semantic web

- The goal of the Semantic Web is to make internet data machine-readable.
- Linked Data refers to a set of best practices for publishing and connecting structured data on the Web
- Linked data enables data from heterogeneous sources to be shared, integrated and queried in a “Web of Data”
- Ontologies are a main vehicle for data integration, sharing, discovery, and reuse
- Possible via technologies such as Resource Description Framework (RDF) and Web Ontology Language (OWL)
- SPARQL is an RDF query language



SPARQL queries in UniProt

<https://sparql.uniprot.org/>

- Contain examples of queries that can be modified

Your SPARQL query

Add common prefixes

```
1 PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
2 PREFIX up: <http://purl.uniprot.org/core/>
3 SELECT
4   (CONCAT(SUBSTR(STR(?protein), 33)) AS ?uniprot)
5   (GROUP_CONCAT(?celtype; separator=";") AS ?celtypes)
6   (GROUP_CONCAT(?biotype; separator=";") AS ?biotypes)
7   (GROUP_CONCAT(?moltypes; separator=";") AS ?moltypes)
8 WHERE
9 {
10  VALUES (?ac) {("Q6GZX4") ("Q96375")}
11  BIND (IRI(CONCAT("http://purl.uniprot.org/uniprot/",?ac)) AS ?protein)
12  ?protein a up:Protein .
13  ?protein up:classifiedWith ?goTerm .
14  #Determine if the type is biological_process
15 OPTIONAL {
```

[Submit Query](#)

Examples

19. Find all Natural Variant Annotations if associated via an evidence tag to an article with a pubmed identifier [Use](#)
20. Find how often an article in pubmed was used in an evidence tag (ordered by most used to least) [Use](#)
21. Find where disease related proteins are known to be located in the cell [Use](#)
22. For two accessions find the GO term labels and group them into GO process,function and component [Use](#)
23. Number of reviewed entries (Swiss-Prot)

Run this example, once you get result, modify to see the GO terms for PTGS1 ([P23219](#)), PTGS2 ([P35354](#)) and PTGES ([O14684](#))



SPARQL queries in UniProt

<https://sparql.uniprot.org/>

Modify query for PTGS1 ([O14684](#)), PTGS2 ([P35354](#)) and PTGES ([O14684](#))

ORIGINAL

Add common prefixes

```
1 PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
2 PREFIX up: <http://purl.uniprot.org/core/>
3 SELECT
4   (CONCAT(SUBSTR(STR(?protein), 33)) AS ?uniprot)
5   (GROUP_CONCAT(?celtype; separator=";") AS ?celtypes)
6   (GROUP_CONCAT(?biotype; separator=";") AS ?biotypes)
7   (GROUP_CONCAT(?moltype; separator=";") AS ?moltypes)
8 WHERE
9  {
10    VALUES (?ac) {("Q6GZX4") ("Q96375")}
11    BIND (IRI(CONCAT("http://purl.uniprot.org/uniprot/",?ac)) AS ?protein)
12    ?protein a up:Protein .
13    ?protein up:classifiedWith ?goTerm .
14    #Determine if the type is biological_process
15    OPTIONAL {
```

Submit Query

MODIFIED

Add common prefixes

```
1 PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
2 PREFIX up: <http://purl.uniprot.org/core/>
3 SELECT
4   (CONCAT(SUBSTR(STR(?protein), 33)) AS ?uniprot)
5   (GROUP_CONCAT(?celtype; separator=";") AS ?celtypes)
6   (GROUP_CONCAT(?biotype; separator=";") AS ?biotypes)
7   (GROUP_CONCAT(?moltype; separator=";") AS ?moltypes)
8 WHERE
9  {
10    VALUES (?ac) {"P23219"} {"P35354"} {"014684"}
11    BIND (IRI(CONCAT("http://purl.uniprot.org/uniprot/",?ac)) AS ?protein)
12    ?protein a up:Protein .
13    ?protein up:classifiedWith ?goTerm .
14    #Determine if the type is biological_process
15    OPTIONAL {
```

Submit Query



CENTER FOR
BIOINFORMATICS &
COMPUTATIONAL
BIOLOGY

SPARQL queries in PRO

<https://lod.proconsortium.org/yasgui.html>

Federated Query

- FQ1: get UniProt AC and Mnemonic for PRO term
- FQ2: get all PRO human genes whose UniProt counterpart has variants with loss of function implicated in disease.
- FQ3: get all PRO human genes whose UniProt counterpart has variants with gain of function implicated in disease.
- FQ4: for proteins in ProKinO, count the number of homologs and modifications in PRO.
- FQ5: for proteins 'EGFR' 'PKCB' 'PKCT' in ProKinO, get their proteoforms that are part of complexes from PRO.
- FQ6: for proteins 'EGFR' 'PKCB' 'PKCT' in ProKinO, get their proteoforms with molecular functions from PRO.
- FQ7: for protein 'EGFR' in ProKinO, get its proteoforms from PRO that have a modification at site 768.
- FQ8: Find variants in UniProt or DisGeNET for AlzForum PRO terms.

Queries integrating data from different resources

General Query

- GQ1: get direct subclasses of TGF-beta superfamily receptor type-1.
- GQ2: get all subclasses of TGF-beta superfamily receptor type-1.
- GQ3: get all functional annotations for PR:000037190.
- GQ4: get PRO terms that pertain to products of a specified gene.
- GQ5: get information pertaining to the gene encoding a protein of interest.
- GQ6: get all PRO terms with functional annotation GO:0007179.

Queries within PRO



Concluding Remarks

- Ontologies offer a mean to represent the knowledge in a way that is understood both by humans and computers (allow automatic reasoning and interpretation)
- GO is the most widely used ontology, with terms to describe the properties of gene products. GO annotations are associated to genes and proteins
- Ontologies are a great tool for databases to organize and make data available and searchable
- Ontologies are the driving force for the semantic web





[This Photo](#) by Unknown Author is licensed under [CC BY-SA-NC](#)



CENTER FOR
BIOINFORMATICS &
COMPUTATIONAL
BIOLOGY