

Need Help?

ChEBI is hosted by the EMBL-European Bioinformatics Institute, a part of the European Molecular Biology Laboratory.

www.ebi.ac.uk/chebi

User manual: www.ebi.ac.uk/chebi/userManualForward.do

SourceForge forum: <http://sourceforge.net/projects/chebi/>

Submissions: www.ebi.ac.uk/chebi/submissions

Support: chebi-help@ebi.ac.uk

Phone: +44 (0) 1223 494444

ChEBI Team

EMBL-EBI

Wellcome Trust Genome Campus

Cambridge

CB10 1SD, UK

ChEBI: Chemical Entities of Biological Interest

www.ebi.ac.uk/chebi

As researchers begin to look beyond the genome with the aim of understanding all the processes of life, the need for a public database of biologically relevant 'small molecules' (molecules not encoded by the genome) grows stronger. Such a database should provide standardised descriptions of molecules that enable other databases to annotate their entries consistently. It should also bridge the gap between small molecules and the macromolecules that they interact with in living systems. ChEBI, EMBL-EBI's database of Chemical Entities of Biological Interest, has been designed with these aims in mind.

What is ChEBI?

ChEBI is a freely available, manually annotated database of small molecular entities. These could be any constitutionally or isotopically distinct atom, molecule, ion, ion pair, radical, radical ion, complex, conformer, or that is identifiable as a separately distinguishable entity, not directly encoded by the genome. The database focuses on chemical nomenclature and structures, and provides a wide range of related chemical data such as formulae, links to other databases and an ontology for the chemical space.

Download flat files, database dumps, and the ChEBI Ontology for local installation

View relationships in the ChEBI ontology

Link to other databases such as PDBeChem, KEGG COMPOUND and COME

View structure, nomenclature, formula, InChI, registry numbers and more

View mappings to other databases such as Reactome and UniProt

ChEBI:17992

CHBI Name sucrose

CHBI ID CHEBI:17992

Secondary CHBI IDs CHEBI:15128, CHEBI:45795, CHEBI:26812, CHEBI:519617, CHEBI:9314

InChI InChI=1S/C12H22O11/c13-1-4-6(16)/8(19)/11(21-4)23-12(3-15)10(20)/(17)5(2-14)22-12/h4-11,13-20H,1-3H2/4-5,6-7,8-9,10-11,12+1/m1

InChIKey InChIKey=CZMRCOWAGMRECN-UGDNZRGBSAN

SMILES OC[C@H]1O[C@H](O[C@@]2[C@@H](CO)O[C@H](O)[C@@H]2O)[C@@H](O)[C@@H](O)[C@@H]1O

Formula C12H22O11

ChEBI Ontology

Outgoing sucrose (CHEBI:17992) is a glycosyl glycoside (CHEBI:24407)

Incoming β -D-fructofuranosyl 6-O-octanoyl- α -D-glucopyranoside (CHEBI:39724) has functional parent sucrose (CHEBI:17992)
sucrose 6'-phosphate (CHEBI:16308) has functional parent sucrose (CHEBI:17992)

IUPAC Name β -D-fructofuranosyl α -D-glucopyranoside

Synonyms 1- α -D-Glucopyranosyl-2- β -D-fructofuranoside
 β -D-Fru(2-1)- α -D-Glcp
Cane sugar
saccharose
Saccharose
Saccharose
Sucrose
SUCROSE
table sugar

Database Links

Registry Numbers

Registry Number	Type
57-50-1	CAS Registry Number
57-50-1	CAS Registry Number
57-50-1	CAS Registry Number
90825	Beilstein Registry Number
9766	Mein Registry Number

Lipid Dictionary

PDBeChem: Molecules

SUC

Distinct chemical molecule that is composed of atoms and bonds

Code SUC

One-letter code SUCROSE

Molecular name C12H22O11

Systematic name β -D-fructofuranosyl 6-O-octanoyl- α -D-glucopyranoside

Formula C12H22O11

Polymer type Saccharide

Is a monomer of

Standard parent

Number of atoms 45

Number of atoms (except hydrogen) 23

Formal charge 0

Molecular weight 342.296

Defined at 1966-07-08

Last modified at 2006-10-14

SMILES OC[C@H]1O[C@H](O[C@@]2[C@@H](CO)O[C@H](O)[C@@H]2O)[C@@H](O)[C@@H](O)[C@@H]1O

SMILES OC[C@H]1O[C@H](O[C@@]2[C@@H](CO)O[C@H](O)[C@@H]2O)[C@@H](O)[C@@H](O)[C@@H]1O

What type of information does ChEBI contain?

The ChEBI database combines chemical nomenclature, structures, and related chemical information from a number of freely accessible sources. All data are manually annotated to a high standard before public release, using nomenclature, symbolism and terminology endorsed by the International Union of Pure and Applied Chemistry (IUPAC) and the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology (NC-IUBMB).

A major feature of ChEBI is that entries are interrelated by the ChEBI ontology, representing the meaning of the data in a structured manner as relationships between entities and their parents (less specialised) and/or children (more specialised terms). The ChEBI ontology is divided into sub-ontologies for molecular structure, role and sub-atomic particles.

ChEBI also allows you to search small molecule data from ChEMBL, a large database of drugs and drug-like molecules. To distinguish between fully annotated ChEBI entities and partially annotated ChEMBL data, a starring system has been devised in which fully annotated ChEBI entities are allocated three stars and partially annotated ChEMBL entities are allocated two. This is indicated in the search results on the ChEBI website: www.ebi.ac.uk/chebi

What can I do with ChEBI?

- Find the correct chemical terminology using name, formula or registry numbers, including CAS, Beilstein and Gmelin Registry Numbers.
- Visualise chemical structures, and use the chemical substructure and similarity search powered by the open source Oracle cartridge OrChem. The facility allows a user to draw or upload a chemical structure and then perform either an exact, substructure or similarity search.
- View the relationships between molecules using the ChEBI ontology, either from within a ChEBI entry or using the EBI's Ontology Lookup Service.
- Bridge the gap between small molecules and the macromolecules they interact with. Biological databases such as the UniProt Knowledgebase and Reactome allow users to view cross references to all entries featuring a particular chemical.
- Download chemical structures in MDL Molfile format and manipulate them using a Java applet.
- Request and discuss new entries using the ChEBI submission tool or ChEBI's SourceForge discussion forum <http://sourceforge.net/projects/chebi/>

Retrieving data from ChEBI

Quick search. Simply type in the term (e.g. cholesterol), formula (e.g. $C_6H_{12}O_6$), registry number (e.g. 64-17-5), IUPAC International Chemical Identifier (e.g. $InChI=1/H_2O/h1H2$) or ChEBI identifier (e.g. 30815). Wildcards can be used to search using a partial name; e.g. searching for *cholest** will find cholesterol, cholesteryl β -D-glucoside.

Advanced search. Search using several terms at once or specify the field to be searched (e.g. ChEBI name, synonym, formula). Structure-based searches can be performed by drawing or loading a chemical structure and the search can be further restricted by combining with a term-based search.

Data download. The entire ChEBI database can be downloaded from <ftp://ftp.ebi.ac.uk/pub/databases/chebi/> in several formats including SDF, Oracle and generic database dumps, flat files and the Open Biomedical Ontologies (OBO) format.

Web services. Programmatic access to ChEBI is available through the EBI's web services at www.ebi.ac.uk/Tools/webservices ●

Further reading

de Matos, P. *et al.* (2009) Chemical Entities of Biological Interest: an update. *Nucleic Acids Res.* (2009) Doi: 10.1093/nar/gkp886

Degtyarenko, K. (2003) Chemical vocabularies and ontologies for bioinformatics in *Proceedings of the 2003 International Chemical Information Conference, Nimes, France, 19-22 October 2003* (Collier, H., Ed.) Infonortics, Tetbury, pp. 144–162

About ChEBI

ChEBI is funded by the European Commission under SLING, grant agreement number 226073 (Integrating Activity) within Research Infrastructures of the FP7 Capacities Specific Programme, and by the BBSRC, grant agreement number BB/G022747/1 within the Bioinformatics and Biological Resources fund.