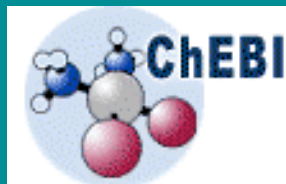


ChEBI, an Open-access Chemistry Resource for the Life Sciences: Facilities for On-line Submission and Curation

Marcus Ennis
European Bioinformatics Institute
Cambridge
UK



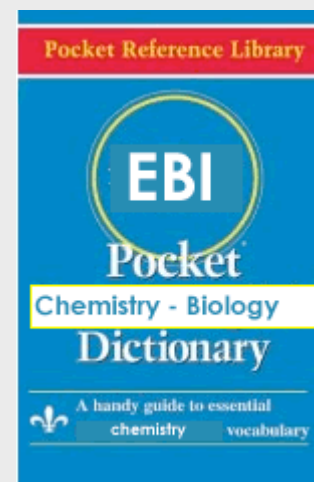
Biocuration 2010
12 October, 2010

EMBL-EBI 

EBI is an Outstation of the European Molecular Biology Laboratory.

What is ChEBI?

- **C**hemical **E**ntities of **B**iological **I**nterest
- Freely available
- Focused on 'small' chemical entities (no proteins or nucleic acids)
- Illustrated dictionary of **chemical nomenclature**
- Provides chemical **ontology**
- High quality, manual curation
- Public release: First Monday of every month
(latest release: 22,000 manually curated entries; 620,000 total entries).
- Data made available via
 - Public web application
 - Flat-file binary table dumps
 - Oracle table dumps
 - Generic SQL table dumps
 - MDL SD file export
- Query via Web services



Access ChEBI at <http://www.ebi.ac.uk/chebi/>

ChEBI entry view (i)

EBI > Databases > Small Molecules > ChEBI > Main

paracetamol (CHEBI:46195)

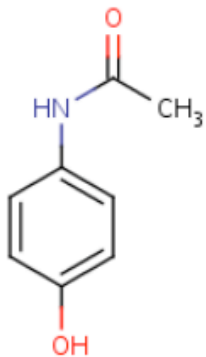
Quick search here!

GO

Search for **★★★★ only**

Main

Automatic Xrefs



 Molfile

ChEBI Name [?](#) **paracetamol**

ChEBI ID [?](#) **CHEBI:46195**

Definition [?](#) A derivative of phenol which has an acetamido substituent located *para* to the phenolic -OH group.

Last Modified [?](#) 29 June 2010


Stars [?](#) **★★★★** This entity has been manually annotated by the ChEBI Team.


Secondary ChEBI IDs [?](#) CHEBI:383982, CHEBI:116450, CHEBI:46191, CHEBI:2386

Image

Applet

[more structures >>](#)

InChI [?](#)  InChI=1/C8H9NO2/c1-6(10)9-7-2-4-8(11)5-3-7/h2-5,11H,1H3,(H,9,10)/f/h9H

InChIKey [?](#)  InChIKey=RZVAJINKPMORJF-BGGKNDAXCW

SMILES [?](#)  CC(=O)Nc1ccc(O)cc1

Formula [?](#)

C₈H₉NO₂

Source

KEGG COMPOUND

Charge [?](#) 0

Mass [?](#) 151.16260

ChEBI entry view (ii)

ChEBI Ontology

Outgoing and incoming view only

CHEBI:24431 molecular structure

- ↑ CHEBI:23367 molecular entity
- ↑ CHEBI:33579 main group molecular entity
- ↑ CHEBI:33675 p-block molecular entity
- ↑ CHEBI:33302 pnictogen molecular entity
- ↑ CHEBI:51143 nitrogen molecular entity
- ↑ CHEBI:32988 amide
- ↑ CHEBI:33256 primary amide
- ↑ CHEBI:37622 carboxamide
- ↑ CHEBI:29347 monocarboxylic acid amide
- ↑ CHEBI:22160 acetamides
- ↑ **CHEBI:46195 paracetamol**
- ↑ CHEBI:32639 acetaminophen glutathione conjugate
- ↑ CHEBI:32635 paracetamol sulfate

CHEBI:50906 role

- ↑ CHEBI:24432 biological role
- ↑ CHEBI:52206 biochemical role
- ↑ CHEBI:23924 enzyme inhibitor
- ↑ CHEBI:35544 cyclooxygenase inhibitor
- ↑ CHEBI:50629 cyclooxygenase 2 inhibitor
- ↑ **CHEBI:46195 paracetamol**
- ↑ CHEBI:32639 acetaminophen glutathione conjugate
- ↑ CHEBI:32635 paracetamol sulfate
- ↑ CHEBI:50630 cyclooxygenase 1 inhibitor
- ↑ **CHEBI:46195 paracetamol**
- ↑ CHEBI:32639 acetaminophen glutathione conjugate
- ↑ CHEBI:32635 paracetamol sulfate
- ↑ CHEBI:51086 chemical role
- ↑ CHEBI:35222 inhibitor
- ↑ CHEBI:23924 enzyme inhibitor
- ↑ CHEBI:35544 cyclooxygenase inhibitor

Relationship Types

- is a
- has part
- is conjugate base of
- is conjugate acid of
- is tautomer of
- is enantiomer of
- has functional parent
- has parent hydride
- is substituent group from
- has role

Status

- Checked
- Unchecked


ChEBI entry view (iii)

IUPAC Name

N-(4-hydroxyphenyl)acetamide

INNs

paracetamol

paracetamol 

paracétamol 

paracetamolium 

Sources

KEGG DRUG

ChEBI

ChEBI

ChemIDplus

Synonyms

4'-hydroxyacetanilide

4-(Acetylamino)phenol

4-acetamidophenol

Acenol

acetaminofén 

Acetaminofen

Acetaminophen

acétaminophène 

APAP

N-acetyl-*p*-aminophenol

p-acetamidophenol

p-acetaminophenol

p-Acetylamino phenol

p-hydroxyacetanilide

p-hydroxyphenolacetamide

Sources

ChemIDplus

ChemIDplus

NIST Chemistry WebBook

ChemIDplus

ChemIDplus 

ChemIDplus

KEGG COMPOUND

ChEBI

DrugBank

ChEBI

NIST Chemistry WebBook 

NIST Chemistry WebBook 

ChemIDplus

NIST Chemistry WebBook 

NIST Chemistry WebBook 

Brand Name

Tylenol

Source

KEGG DRUG

Database Links

[C06804](#)

[D00217](#)

[DB00316](#)

[TYL](#)

[116450](#)

Databases

KEGG COMPOUND

KEGG DRUG

DrugBank

PDBChem

ChEMBL

ChEBI entry view (iv)

Registry Numbers ?	Types	Sources
103-90-2	CAS Registry Number	KEGG COMPOUND
103-90-2	CAS Registry Number	NIST Chemistry WebBook
103-90-2	CAS Registry Number	ChemIDplus
2208089	Beilstein Registry Number	Beilstein

Citations [?](#)

Jorgensen WL, Duffy EM (2000)

Prediction of drug solubility from Monte Carlo simulations.

Bioorganic & medicinal chemistry letters **10**, 1155-8 (Source: *ChEMBL*) [[PubMed:10866370](#)]

[\[show Abstract\]](#)

Lombardo F, Shalaeva MY, Tupper KA, Gao F, Abraham MH (2000)

ElogP(Oct): a tool for lipophilicity determination in drug discovery.

Journal of medicinal chemistry **43**, 2922-8 (Source: *ChEMBL*) [[PubMed:10956200](#)]

[\[show Abstract\]](#)

Colmenarejo G, Alvarez-Pedraglio A, Lavandera JL (2001)

Cheminformatic models to predict binding affinities to human serum albumin.

Journal of medicinal chemistry **44**, 4370-8 (Source: *ChEMBL*) [[PubMed:11728183](#)]

[\[show Abstract\]](#)

Lombardo F, Obach RS, Shalaeva MY, Gao F (2002)

Prediction of volume of distribution values in humans for neutral and basic drugs using physicochemical measurements and plasma protein binding data.

Journal of medicinal chemistry **45**, 2867-76 (Source: *ChEMBL*) [[PubMed:12061889](#)]

[\[show Abstract\]](#)

Sharma V, Duffel MW (2002)

Comparative molecular field analysis of substrates for an aryl sulfotransferase based on catalytic mechanism and protein homology modeling.

Journal of medicinal chemistry **45**, 5514-22 (Source: *ChEMBL*) [[PubMed:12459019](#)]

[\[show Abstract\]](#)

Lombardo F, Shalaeva MY, Tupper KA, Gao F (2001)

ElogD(Oct): a tool for lipophilicity determination in drug discovery. 2. Basic and neutral compounds.

Journal of medicinal chemistry **44**, 2490-7 (Source: *ChEMBL*) [[PubMed:11448232](#)]

[\[show Abstract\]](#)

Lombardo F, Obach RS, Shalaeva MY, Gao F (2004)

Prediction of human volume of distribution values for neutral and basic drugs. 2. Extended data set and leave-class-out statistics.

The submission tool

- Web-based software utility to enable direct user submissions
- Needed to alleviate backlog of user requests
- An invitation to the community to participate in the growth and development of ChEBI
- Secure (submitter registration required)
- Simple stepwise procedure
- Instant generation of ChEBI ID and inclusion in next monthly public release

Access at <https://www.ebi.ac.uk/chebi/submissions/> or via ChEBI home page


Submission validations



Built-in integrity checker will check for:

- Uniqueness of name
- Uniqueness of chemical structure
- Correctness of ontology relationships
 - No non-allowed cycles
 - Relationships only between entities of relevant types
- Correctness of xref format
 - CAS Registry Number
 - DrugBank
 - KEGG LIGAND/DRUG/GLYCAN
 - Lipid Maps



Login page

EMBL-EBI  All Databases [Reset ?](#) [Give us feedback](#)

[Databases](#) | [Tools](#) | [EBI Groups](#) | [Training](#) | [Industry](#) | [About Us](#) | [Help](#) | [Site Index](#)  

[EBI](#) > [Databases](#) > [Small Molecules](#) > [ChEBI](#) > [Submissions](#)

Login to ChEBI Submissions Tool

Remember me?

[?](#)

Don't have a username? [Click here](#) to register.

Forgotten your password? [Click here](#) to reset.

[Terms of Use](#) | [EBI Funding](#) | [Contact EBI](#) | © European Bioinformatics Institute 2010. EBI is an Outstation of the European Molecular Biology Laboratory.

Submission Step 1 (i)

EBI > Databases > Small Molecules > ChEBI > Submissions

Edit ChEBI Submission

Submission Step 1

[Submission Step 2](#)

[Submission Step 3](#)

[Submission Step 4](#)

[Preview](#)



To complete this step in the submission process, you should enter a structure, name and optional definition.

[View Structure](#) [Edit Structure](#)

File Edit View Insert Atom Bond Structure Tools Help

Update structure ?

ChEBI name	<input type="text" value="shekanin"/>
Definition	<input type="text" value="The beta-D-glucoside of tectorigenin"/>
Formula	
InChI	
InChIKey	
SMILES	
Mass	
Charge	

Submission Step 1(ii)

EBI > Databases > Small Molecules > ChEBI > Submissions

Edit ChEBI Submission

Submission Step 1

Submission Step 2

Submission Step 3

Submission Step 4

Preview



To complete this step in the submission process, you should enter a structure, name and optional definition.

View Structure Edit Structure

File Edit View Insert Atom Bond Structure Tools Help

Update structure ?

ChEBI name	shekanin ?
Definition	The beta-D-glucoside of tectorigenin. ?
Formula	C ₂₂ H ₂₂ O ₁₀
InChI	InChI=1/C ₂₂ H ₂₂ O ₁₀ /c1-29-15-6-12-14(30-9-13(18(12)25)10-2-4-11(24)5-3-10) 7-16(15)31-22-21(28)20(27)19(26)17(8-23)32-22
InChIKey	InChIKey=OZBAVEKZGSOMOJ-MIUGBVL SBI
SMILES	<chem>COc1cc2c(cc1O[C@@H]1O[C@H](CO)[C@@H](O)[C@H](O)[C@H]1O)occ(-c1ccc(C</chem>
Mass	446.40410
Charge	0

Submission Step 2

EBI > Databases > Small Molecules > ChEBI > Submissions

Edit ChEBI Submission

[Submission Step 1](#)

[Submission Step 2](#)

[Submission Step 3](#)

[Submission Step 4](#)

[Preview](#)



You are editing shekanin. To complete this step in the submission process, you should capture any relevant database cross-references.

Database Cross-References

Cross-reference	Type	
<input type="text" value="611-42-7"/>	<input type="text" value="CAS Registry Number"/>	Delete cross-reference
<input type="text" value="68384"/>	<input type="text" value="Beilstein Registry Number"/>	Delete cross-reference

[Add cross-reference](#)

[Save changes](#)

[Cancel and return](#)

[Save and return](#)

[View Discussion...](#)

Submission Step 3

EBI > Databases > Small Molecules > ChEBI > Submissions

Edit ChEBI Submission

[Submission Step 1](#)

[Submission Step 2](#)

[Submission Step 3](#)

[Submission Step 4](#)

[Preview](#)



You are editing shekanin. To complete this step in the submission process, you should capture any relevant synonyms.

Synonyms

Synonym	Type	Language		
<input type="text" value="Tectoridin"/>	SYNONYM	English	?	Delete synonym
<input type="text" value="6-methoxy-4-oxo-3-phenyl-4H-ch"/>	IUPAC NAME	English	?	Delete synonym

[Add synonym](#)

Save changes

Cancel and return

Save and return

[View Discussion...](#)

Submission Step 4(i)

EBI > Databases > Small Molecules > ChEBI > Submissions

Edit ChEBI Submission

[Submission Step 1](#)

[Submission Step 2](#)

[Submission Step 3](#)

[Submission Step 4](#)

[Preview](#)



You are editing shekanin. To complete this step in the submission process, you should classify the submission within the ontology. The classification may be specified in the 'simple' or 'advanced' forms below.

Classification View

Simple

Select the classification which best matches your submission:

- organic molecular entity
- inorganic molecular entity
- group
- biological role
- application



Classification

shekanin *is_a* organic molecular entity (CHEBI:50860)



[Delete relationship](#)

Save changes

Cancel and return

Save and return

[View Discussion...](#)

Submission Step 4(ii)

EBI > Databases > Small Molecules > ChEBI > Submissions

Edit ChEBI Submission

[Submission Step 1](#)

[Submission Step 2](#)

[Submission Step 3](#)

[Submission Step 4](#)

[Preview](#)



You are editing shekanin. To complete this step in the submission process, you should classify the submission within the ontology. The classification may be specified in the 'simple' or 'advanced' forms below.

Classification View Advanced ▾

Direction: outgoing
 incoming

Relationship type: has_role ▾

Target ID:

Target Name:

[Add relationship](#)

[Select](#)

[Browse ChEBI Ontology](#)

Classification		
shekanin <i>is_a</i> flavones (CHEBI:24043)	?	Delete relationship
shekanin <i>is_a</i> β-D-glucoside (CHEBI:22798)	?	Delete relationship
shekanin <i>has_role</i> secondary metabolite (CHEBI:26619)	?	Delete relationship

[Save changes](#)

[Cancel and return](#)

[Save and return](#)

[View Discussion...](#)

Submission preview(i)

EBI > Databases > Small Molecules > ChEBI > Submissions

[Help](#) | [Logout](#)

Edit ChEBI Submission

[Submission Step 1](#)

[Submission Step 2](#)

[Submission Step 3](#)

[Submission Step 4](#)

Preview



Submission of shekanin.

Review your submission and, if there are no changes required, submit by clicking 'submit'. Once you have submitted you will not be able to alter this entry, ChEBI annotators will review this submission and if it is deemed appropriate it will be included in a ChEBI release.

Final Submission

Date submission to be active



Please enter the six characters you can see in the image on the left, in sequence.

[Can't read the image? Get a new code...](#)

Submit Entry to ChEBI

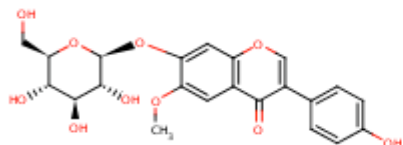
ChEBI name shekanin

Definition The beta-D-glucoside of tectorigenin.

Formula C₂₂H₂₂O₁₀

InChI InChI=1/C₂₂H₂₂O₁₀
/c1-29-15-6-12-14(30-9-13(18(12)25)10-2-4-11(24)5-3-10)
7-16(15)31-22-21(28)20(27)19(26)17(8-23)32-22

InChIKey InChIKey=OZBAVEKZGSOMOJ-MIUGBVL5BI



Submission preview(ii)

SMILES COc1cc2c(cc1O[C@@H]1O[C@H](CO)[C@@H](O)[C@H](O)[C@H]1O)occ(-c1ccc(O)cc1)c2=O
Mass 446.40410
Charge 0

Cross-reference	Type
611-42-7	CAS Registry Number
68384	Beilstein Registry Number

Synonym	Type	Language
Tectoridin	Type: SYNONYM Source: SUBMITTER	English
6-methoxy-4-oxo-3-phenyl-4H-chromen-7-yl beta-D-glucopyranoside	Type: IUPAC NAME Source: SUBMITTER	English

Classification
shekanin <i>is_a</i> flavones (CHEBI:24043)
shekanin <i>is_a</i> β -D-glucoside (CHEBI:22798)
shekanin <i>has_role</i> secondary metabolite (CHEBI:26619)

Save changes

Cancel and return

Save and return

[View Discussion...](#)

Submission confirmation

EBI > Databases > Small Molecules > ChEBI > Submissions

[Help](#) | [Logout](#)

Thank you!

You have successfully submitted *shekanin*. *shekanin* has been assigned ChEBI identifier CHEBI:60111. It will be available online either shortly or on your specified release date.

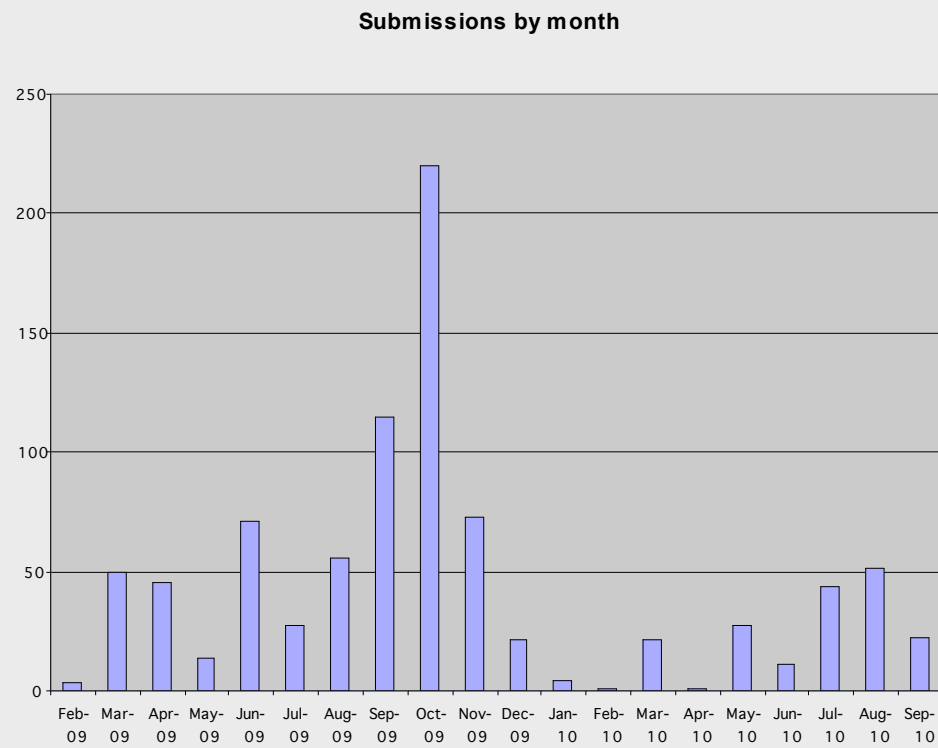
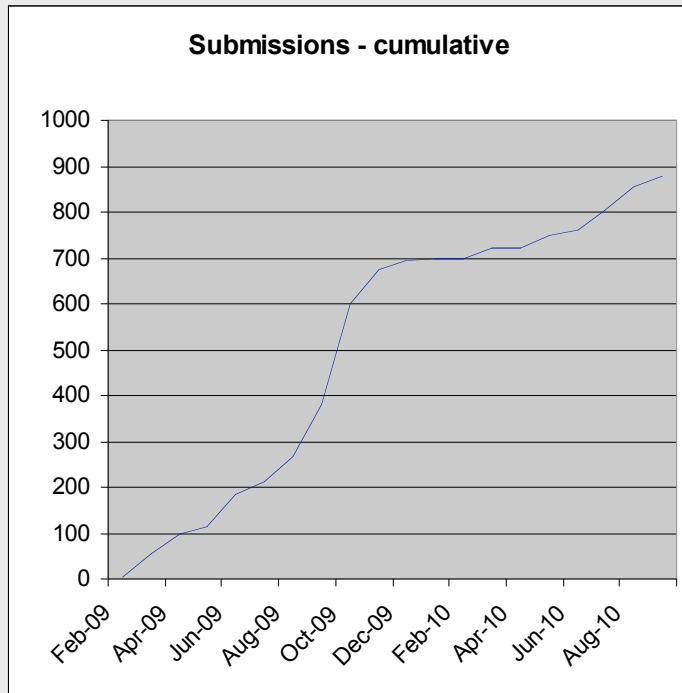


ChEBI annotators will review your submission; if it is deemed appropriate they will approve the entry and it will be made publicly visible in the following ChEBI release. However if it is inappropriate the entry will be deleted and will not be publicly visible as a normal entry but rather as a deleted entry in the next release. ChEBI annotators reserve the right to alter any of the data that is submitted.

You will receive an email notification of your submission.

[Return to start](#)

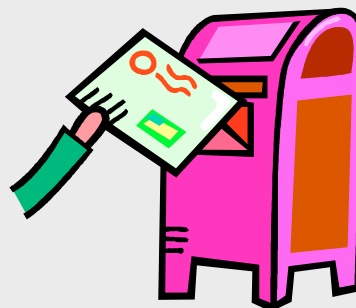
ChEBI submissions Feb 2009 – Sep 2010



(Data as at September 9th 2010)

ChEBI submitters

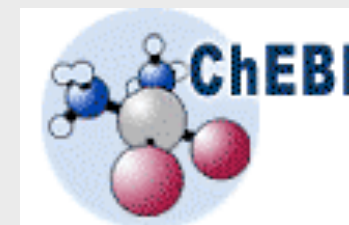
- Reactome
- Gene Ontology Consortium
- RSC Journals
- Rutgers Univ.
- Cornell Univ.
- Carleton Univ.
- Stanford Univ.
- SIB
- IntAct
- Rhea
- Array Express
- Manchester Centre for Integrative Systems Biology
- EML Research, Heidelberg
- Dept of Genetics, Cambridge Univ.



Submission tool - Future plans

- Bulk submissions (by invitation only) - imminent
- Modifications to existing ChEBI entries - ???

The ChEBI team



Curators

Marcus Ennis
Steve Turner
Zara Josephs
Nico Adams
Gareth Owen

Developers

Paula de Matos (team coordinator)
Janna Hastings
Adriano Dekker
Kenneth Haug

Group Leader

Christoph Steinbeck

www.ebi.ac.uk/chebi

ChEBI is funded by the European Commission under SLING, contract number 226073 within the Research Infrastructure Action of the FP7 "Capacities Specific" Programme.