

1

Workshop on CEFIC LRI Project EEM9.4
LRI AMBIT with IUCLID6 support and extended search capabilities

Hands on AMBIT2 functionality





Log in Info

- **Link:** <https://ambitlri.ideaconsult.net>
- Username: workshop
- Password: ambit2
- Please use the latest versions of browsers

Practical exercises

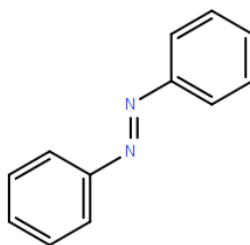
Exact Search

- Search structures using chemical names “Caffeine”, “caffei” (fragment search)
- Search structure using CAS No. 111-96-6 and display substances containing this structure as constituent/impurity
- Search structure in OpenFoodToxData database using SMILES “CCOC(=O)N”

Practical exercises

Similarity and substructure search

- Draw the following structure using the structure editor and search similar structures



Practical exercises

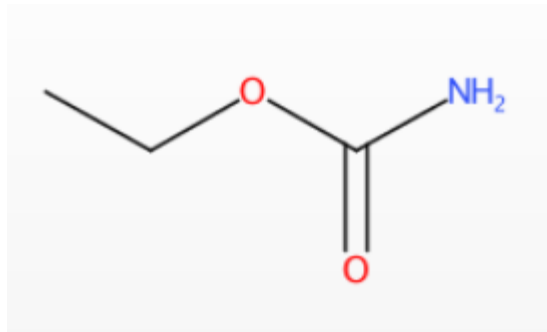
Substance search

- Search substance “Triglyme” and check the available composition and endpoint data

- Search substances which have LD50 values > 5000 (oral) and are not mutagenic (in-vitro)

Practical exercises

- Run VEGA Models and Toxtree prediction for the following chemical



- Create an assessment for Category amines, C12-18-alkyldimethyl (e.g. “N,N-dimethylhexadecan-1-amine”, CAS 112-69-6) with regard to genetic toxicity in-vivo and short-term toxicity to fish



Results

Results

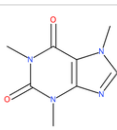
- Search structures using chemical names “Caffeine”, “cafei” (fragment search)

Search structures and associated data

Exact structure Similarity Substructure URL Enable fragment search Caffeine

Identifiers Datasets Export

Showing from 1 to 1 in pages of 20 entries Previous Next Filter...

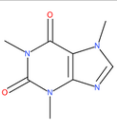
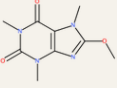
Diagram	CasRN	EC number	IUCLID 5 Re	Names	Trade Name	IUPAC name	SMILES	Std. InChI key	Std. InChI	RI
	58-08-2	200-362-1	ECHA-53...	1,3,7-trimethyl-3,7-dihydro-1H-purine-2,6-dione, caffeine 1H-Purine-2,6-dione, 3,7-dihydro-1,3,7-trimethyl- Caffeine	-	-	O=C1C2...	RYYVLZVUVIJVGH-UHFFFAOYSA-N	InChI=1S...	3C

Search structures and associated data

Exact structure Similarity Substructure URL Enable fragment search Caffe|

Identifiers Datasets Export

Showing from 1 to 9 in pages of 20 entries Previous Next Filter...

Diagram	CasRN	EC number	IUCLID 5 F	Names	Trade Name	IUPAC name	SMILES	Std. InChI key	Std. InChI
	58-08-2	200-362-1	ECHA-53...	1,3,7-trimethyl-3,7-dihydro-1H-purine-2,6-dione, caffeine 1H-Purine-2,6-dione, 3,7-dihydro-1,3,7-trimethyl- Caffeine	-	-	O=C1C2...	RYYVLZVUVIJVGH-UHFFFAOYSA-N	InChI=1S...
	569-34-6	209-312-3	-	8-methoxycaffeine	-	-	O=C1C2...	ATPSJRIIRXKPER-UHFFFAOYSA-N	InChI=1S...

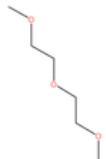












Results


- Search structure using CAS No. 111-96-6 and display substances containing this structure as constituent/impurity

Exact structure Similarity Substructure URL Enable fragment search

Identifiers Datasets

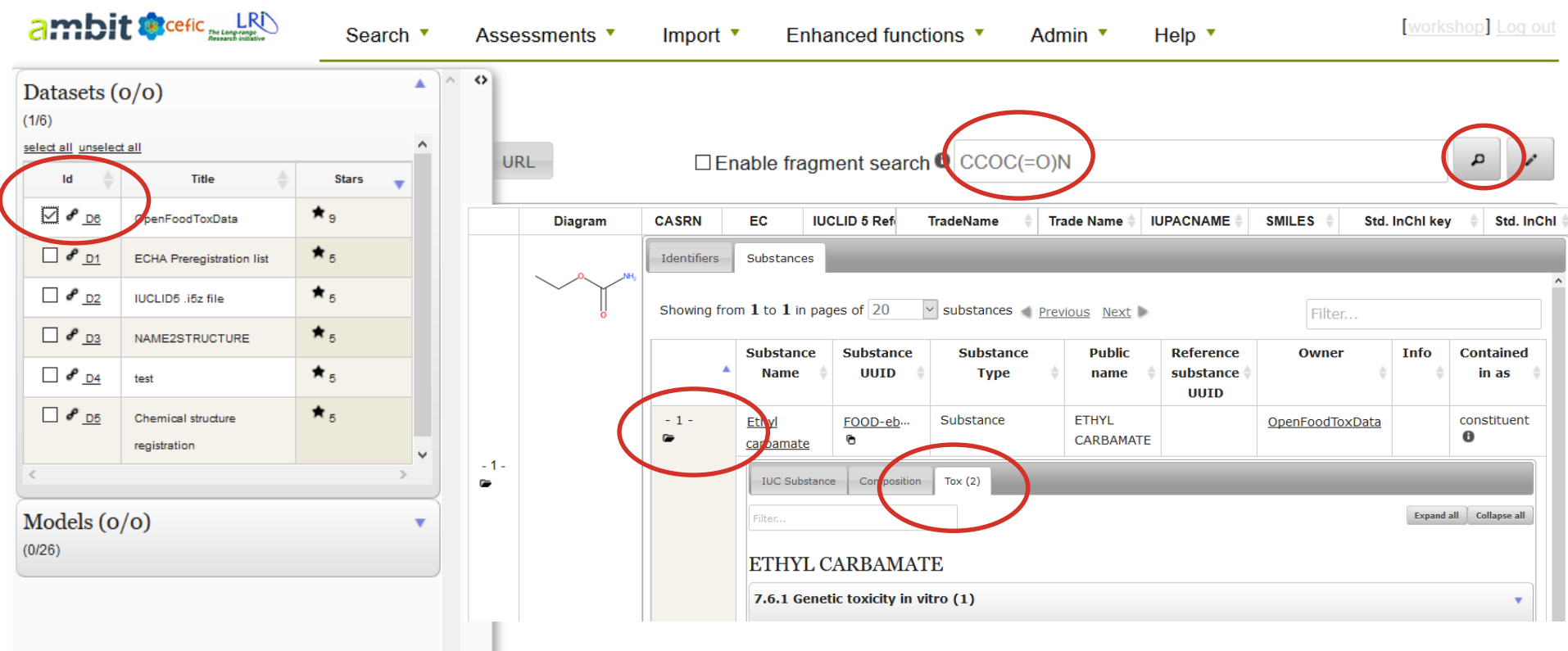
Showing from 1 to 1 in pages of 20 entries

Diagram	CasRN	EC number	IUCLID 5 R	Names	Trade Name	IUPAC name	SMILES	Std. InChI key	Std. InChI	REACH registration da
	- 2 - 	Glymes Diethylene glycol dimethyl ether (DEGDME, Diglyme) ICS MSe DE71	IUC5-a43... 	mono constituent substance	Diglyme			ECB5-19... 	Taunus / Germany Clariant Produkte (Deutschland) GmbH / Sulzbach am Taunus / Germany	constituent 
- 3 - 	bis(2-methoxyethyl)ether	IUC5-fb6... 	mono constituent substance			IUC5-4fd... 	European Chemicals Agency / Helsinki / Finland	constituent 		
- 4 - 	Glymes Ethylene glycol dimethyl ether (EGDME, Monoglyme) ICS MSe DE71	IUC5-eb8... 	mono constituent substance	Monoglyme	ECB5-73... 	Clariant Produkte (Deutschland) GmbH /	impurity 			



Results

- Search structure in OpenFoodToxData database using SMILES “CCOC(=O)N”



ambit ceficLRi The Long-Range Research Initiative

Search Assessments Import Enhanced functions Admin Help [workshop] Log out

URL Enable fragment search

Diagram CASRN EC IUCLID 5 Ref TradeName Trade Name IUPACNAME SMILES Std. InChI key Std. InChI

Identifiers Substances

Showing from 1 to 1 in pages of 20 substances

Substance Name	Substance UUID	Substance Type	Public name	Reference substance UUID	Owner	Info	Contained in as
- 1 - Ethyl carbamate	FOOD-eb...	Substance	ETHYL CARBAMATE		OpenFoodToxData		constituent

IUC Substance Composition Tox (2)

ETHYL CARBAMATE

7.6.1 Genetic toxicity in vitro (1)

Datasets (0/0) (1/6)
select all unselect all

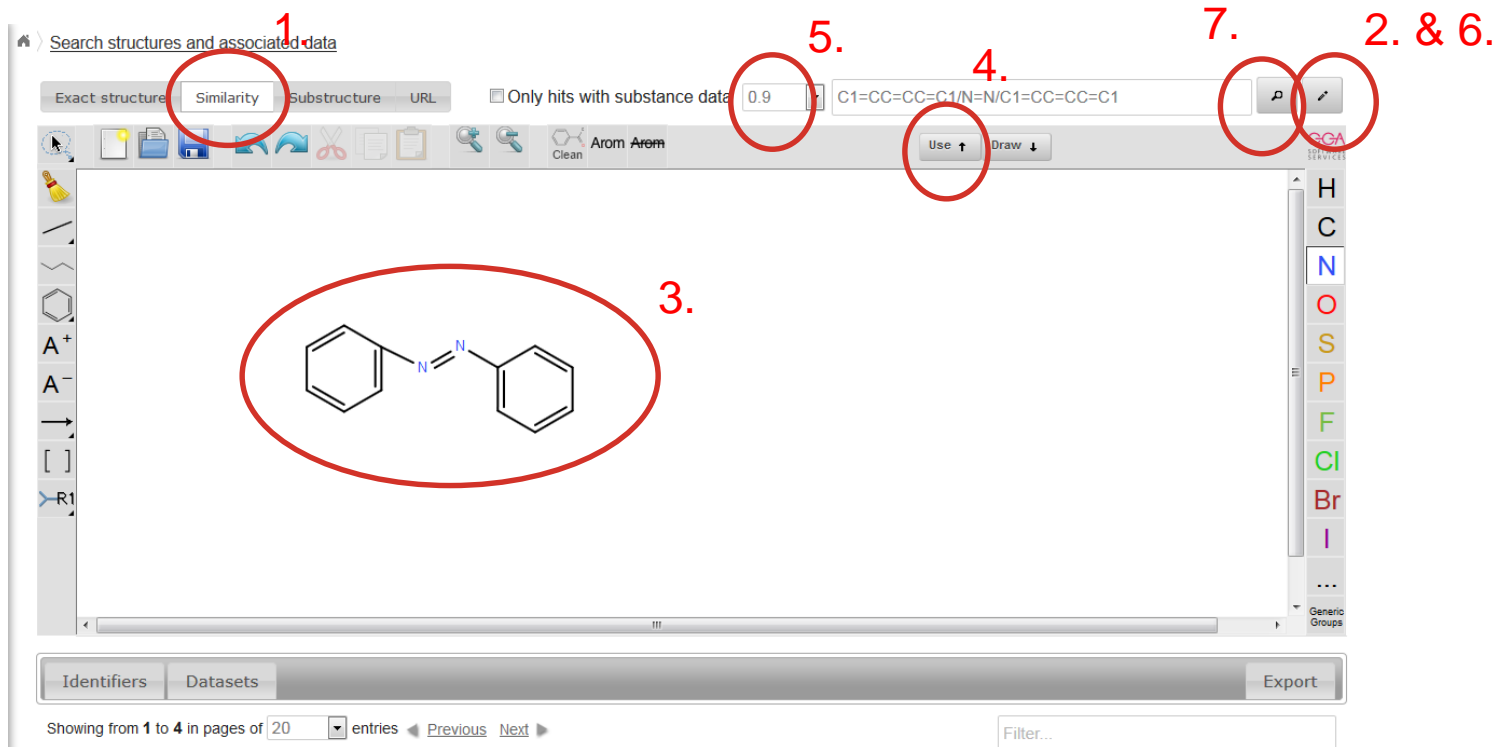
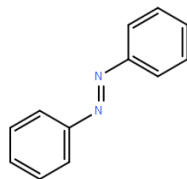
Id	Title	Stars
<input checked="" type="checkbox"/> D6	OpenFoodToxData	★ 9
<input type="checkbox"/> D1	ECHA Preregistration list	★ 5
<input type="checkbox"/> D2	IUCLID5 .ifz file	★ 5
<input type="checkbox"/> D3	NAME2STRUCTURE	★ 5
<input type="checkbox"/> D4	test	★ 5
<input type="checkbox"/> D5	Chemical structure registration	★ 5

Models (0/0) (0/26)

Results

Similarity and substructure search

- Draw the following structure using the structure editor and search similar structures



The screenshot displays a chemical structure editor interface. The main window shows the chemical structure of trans-stilbene, which is circled in red and labeled with the number 3. The search bar at the top contains the SMILES string C1=CC=CC=C1/N=C/C1=CC=CC=C1 and is labeled with the number 5. The search criteria are set to "Similarity" (labeled 1) and "Only hits with substance data" (labeled 2). The similarity threshold is set to 0.9 (labeled 4). The search results are displayed in a list below the main window, with the first four results shown (labeled 7). The interface includes a toolbar with various drawing tools and a vertical panel on the right with element selection buttons (H, C, N, O, S, P, F, Cl, Br, I). The status bar at the bottom indicates "Showing from 1 to 4 in pages of 20 entries".

Results

- Search substance “Triglyme” and check the available composition and endpoint data

Name UUID Name (regexp) External identifier

Showing from 1 to 2 in pages of 10 substances

	Substance Name	Substance UUID	Substance Type	Public name	Reference substance UUID	Owner	Info
<input type="checkbox"/> - 1 -	1,2-bis(2-methoxyethoxy)ethane	ECHA-91983fc6-...	mono constituent substance	Triglyme	ECHA-13bca83c-...		
<input type="checkbox"/> - 2 -	Glymes Triethylene glycol dimethyl ether (TEGDMF) (Triglyme) ICS_MSe_DZ71	IUC5-203e228e-...	mono constituent substance	Triglyme	ECB5-40dcb61-...	Clariant Produkte (Deutschland) GmbH / Sulzbach am Taunus / Germany	

How to search for substances ⓘ

Name UUID Name (regexp) External identifier

Triglyme
7.2.1 Acute toxicity - oral (2)

Results

- Search substances which have LD50 values > 5000 (oral) and are not mutagenic

Search substances by endpoint data | Hit list

Update result

7.2.1. Acute toxicity - oral (S) [26076]

Endpoint name Units

Value

>= 5000 <=

Interpretation of the results

7.2.2. Acute toxicity - inhalation (S) [9213]

7.2.3. Acute toxicity - dermal (S) [11852]

7.3.1. Skin irritation / Corrosion (S) [21694]

7.3.2. Eye irritation (S) [18430]

7.4.1. Skin sensitisation (S) [16404]

7.5.1. Repeated dose toxicity - oral (S) [20212]

7.5.2. Repeated dose toxicity - inhalation (S) [6657]

7.5.3. Repeated dose toxicity - dermal (S) [2139]

7.6.1. Genetic toxicity in vitro (S) [50366]

Genotoxicity

Interpretation of the result

Negative

Hit list | Download | Help

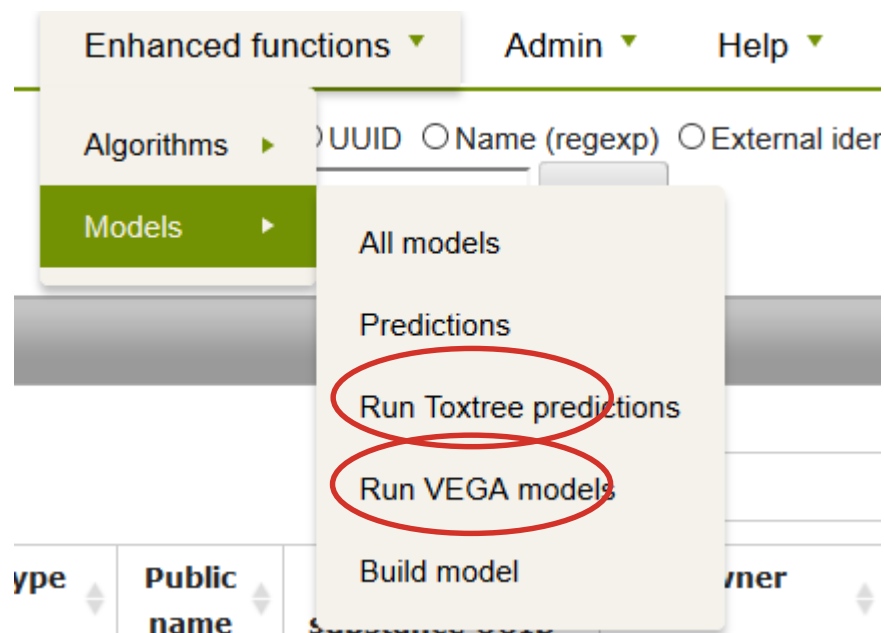
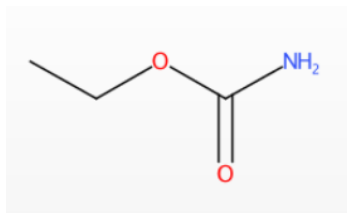
Showing from 1 to 10 in pages of 10 substances | Previous | Next

Filter...

	Substance Name	Substance UUID	Substance Type	Public name	Reference substance UUID	Owner	Info
<input type="checkbox"/> - 1 -	Glymes Ethylene glycol dimethyl ether (EGDME, Monoglyme) ICS_MSe_DE71	IUC5-eb8957ab...	mono constituent substance	Monoglyme	ECB5-730811f8...	Clariant Produkte (Deutschland) GmbH / Sulzbach am Taunus / Germany	
<input type="checkbox"/> - 2 -	Glymes Triethylene glycol dimethyl ether (TEGDME, Triglyme) ICS_MSe_DE71	IUC5-203e228e...	mono constituent substance	Triglyme	ECB5-40dcb61...	Clariant Produkte (Deutschland) GmbH / Sulzbach am Taunus / Germany	
<input type="checkbox"/> - 3 -	(2E, 13Z)-Octadecadien-1-yl Acetate	FOOD-3efbed96...	Substance	(2E, 13Z)-OCTADECADIEN-1-YL ACETATE		OpenFoodToxData	
<input type="checkbox"/> - 4 -	(3RS,4RS;3RS,4SR)-3-chloro-4-chloromethyl-1-(alpha, alpha, alpha-trifluoro-m-tolyl)-2-pyrrolidone	FOOD-e70cd311...	Substance	(3RS,4RS;3RS,4SR)-3-CHLORO-4-CHLOROMETHYL-1-(ALPHA,ALPHA,ALPHA-TRIFLUORO-M-TOLYL)-2-PYRROLIDONE		OpenFoodToxData	
<input type="checkbox"/> - 5 -	(7E, 9E)-Dodecadienyl Acetate	FOOD-340dd43...	Substance	(7E, 9E)-DODECADIENYL ACETATE		OpenFoodToxData	
<input type="checkbox"/> - 6 -	(7E, 9Z)-Dodecadien-1-yl Acetate	FOOD-1b545f71...	Substance	(7E, 9Z)-DODECADIEN-1-YL ACETATE		OpenFoodToxData	
<input type="checkbox"/> - 7 -	(7Z, 11E)-Hexadecadien-1-yl Acetate	FOOD-a9f66edb...	Substance	(7Z, 11E)-HEXADECADIEN-1-YL ACETATE		OpenFoodToxData	
<input type="checkbox"/> - 8 -	(7Z, 11Z)-Hexadecadien-1-yl Acetate	FOOD-9d6a2450...	Substance	(7Z, 11Z)-HEXADECADIEN-1-YL ACETATE		OpenFoodToxData	
<input type="checkbox"/> - 9 -	(9Z, 12E)-Tetradecadien-1-yl acetate	FOOD-db5e9ba2...	Substance	(9Z, 12E)-TETRADECADIEN-1-YL ACETATE		OpenFoodToxData	
<input type="checkbox"/> - 10 -	(E)-11-Tetradecenyl Acetate	FOOD-81f15ecd...	Substance	(E)-11-TETRADECENYL		OpenFoodToxData	

Results

- Run VEGA Models and Toxtree prediction for the following chemical



Enhanced functions ▾ Admin ▾ Help ▾

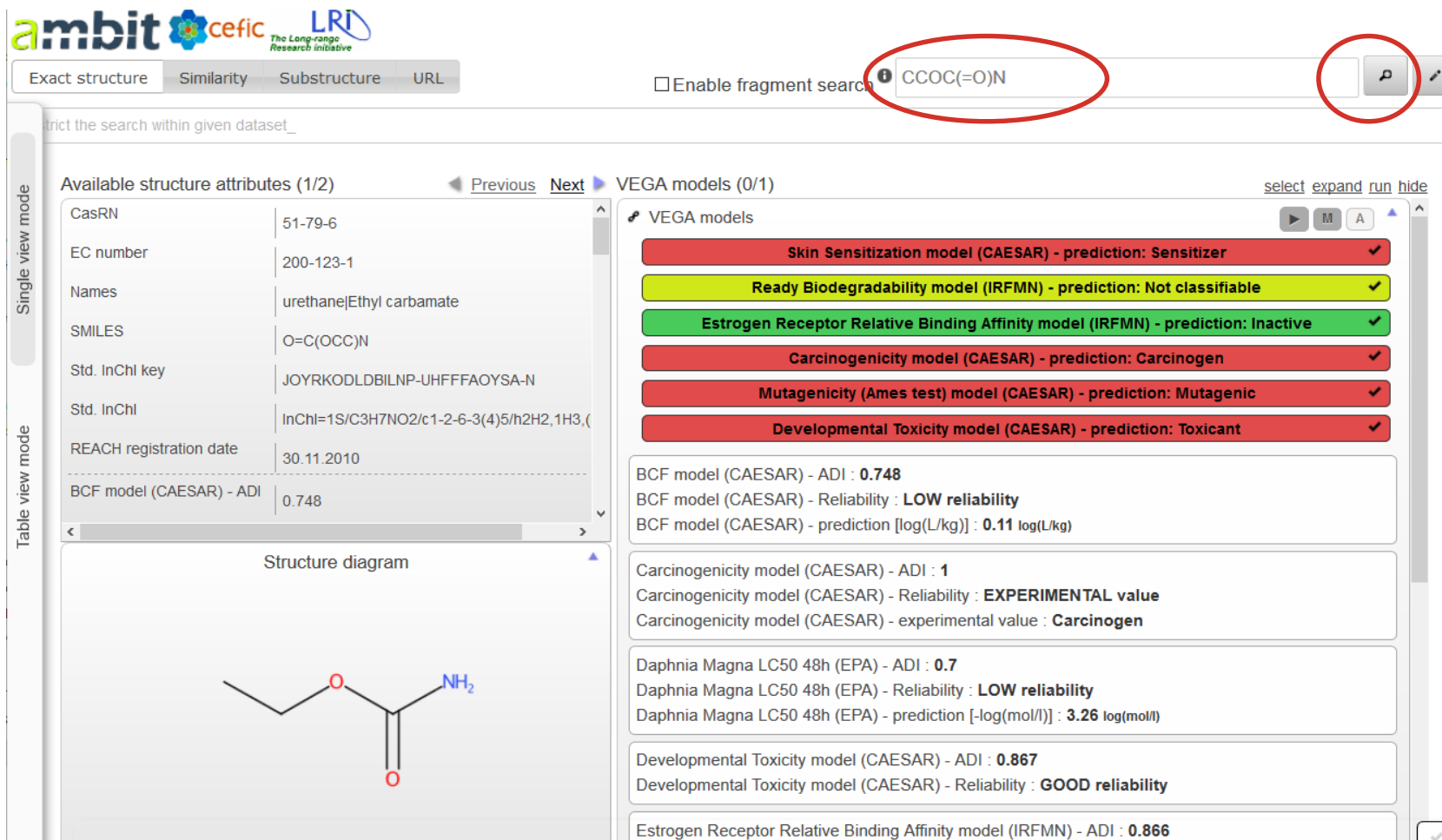
Algorithms ▶ UUID Name (regexp) External identifier

Models ▶

- All models
- Predictions
- Run Toxtree predictions
- Run VEGA models
- Build model

type ▶ Public name ▶ Substance name ▶

VEGA models Prediction



ambit cefic LRI The Long-range Research Initiative

Exact structure Similarity Substructure URL

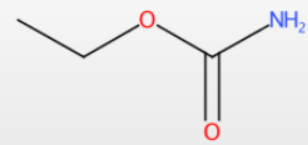
Enable fragment search **CCOC(=O)N**

restrict the search within given dataset

Available structure attributes (1/2) [Previous](#) [Next](#) VEGA models (0/1) [select](#) [expand](#) [run](#) [hide](#)

CasRN	51-79-6
EC number	200-123-1
Names	urethane Ethyl carbamate
SMILES	O=C(OCC)N
Std. InChI key	JOYRKODLDBILNP-UHFFFAOYSA-N
Std. InChI	InChI=1S/C3H7NO2/c1-2-6-3(4)5/h2H2,1H3,(
REACH registration date	30.11.2010
BCF model (CAESAR) - ADI	0.748

Structure diagram



VEGA models

- Skin Sensitization model (CAESAR) - prediction: Sensitizer ✓
- Ready Biodegradability model (IRFMN) - prediction: Not classifiable ✓
- Estrogen Receptor Relative Binding Affinity model (IRFMN) - prediction: Inactive ✓
- Carcinogenicity model (CAESAR) - prediction: Carcinogen ✓
- Mutagenicity (Ames test) model (CAESAR) - prediction: Mutagenic ✓
- Developmental Toxicity model (CAESAR) - prediction: Toxicant ✓

BCF model (CAESAR) - ADI : **0.748**
 BCF model (CAESAR) - Reliability : **LOW reliability**
 BCF model (CAESAR) - prediction [log(L/kg)] : **0.11 log(L/kg)**

Carcinogenicity model (CAESAR) - ADI : **1**
 Carcinogenicity model (CAESAR) - Reliability : **EXPERIMENTAL value**
 Carcinogenicity model (CAESAR) - experimental value : **Carcinogen**

Daphnia Magna LC50 48h (EPA) - ADI : **0.7**
 Daphnia Magna LC50 48h (EPA) - Reliability : **LOW reliability**
 Daphnia Magna LC50 48h (EPA) - prediction [-log(mol/l)] : **3.26 log(mol/l)**

Developmental Toxicity model (CAESAR) - ADI : **0.867**
 Developmental Toxicity model (CAESAR) - Reliability : **GOOD reliability**

Estrogen Receptor Relative Binding Affinity model (IRFMN) - ADI : **0.866**

Toxtree Prediction

Exact structure Similarity Substructure URL

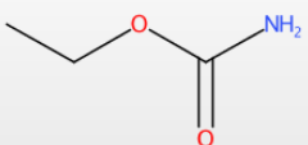
Enable fragment search ↻

Filter the search within given dataset_

Available structure attributes (1/2) ◀ Previous Next ▶

CasRN	51-79-6
EC number	200-123-1
Names	urethane Ethyl carbamate
SMILES	O=C(OCC)N
Std. InChI key	JOYRKODLDBILNP-UHFFFAOYSA-N
Std. InChI	InChI=1S/C3H7NO2/c1-2-6-3(4)5/h2H2,1H3,(
REACH registration date	30.11.2010
Alert for SNAr Identified.	NO

Structure diagram



Toxtree modules (14/14) unselect expand run hide

☞ Cramer rules ▶ M A

Intermediate (Class II)

Low (Class I) ✓

High (Class III)

Q1.Normal constituent of the body **No**

Q2.Contains functional groups associated with enhanced toxicity **No**

Q3.Contains elements other than C,H,O,N,divalent S **No**

Q5.Simplely branched aliphatic hydrocarbon or a common carbohydrate **No**

Q6.Benzene derivative with certain substituents **No**

Q7.Heterocyclic **No**

Q16.Common terpene **No**

Q17.Readily hydrolysed to a common terpene **No**

19.Open chain **Yes**

20.Aliphatic with some functional groups (see explanation) **Yes**

Q21.3 or more different functional groups **No**

Q18.One of the list (see explanation) **No** Class Low (Class I)

☞ Extended Cramer rules ▶ M A


Class 5 (Not possible to classify according to these rules)

Class 1 (narcosis or baseline toxicity)

☞ Verhaar scheme for predicting toxicity mode of action ▶ M A

Results

- Create an assessment for Category amines, C12-18-alkyldimethyl (e.g. “N,N-dimethylhexadecan-1-amine”, CAS 112-69-6) with regard to genetic toxicity in-vivo and short-term toxicity to fish
- Results: see the following assessment in the public server:

Title	Name	Code	Status	Owner
 Assessment demonstration No.2-DMA v2		Glymes20160121	draft	workshop