The Effectiveness of *In Silico* Methods in Predicting the Formation of Active Metabolites

Dr Steve Enoch

Reader in Computational Toxicology

Liverpool John Moores University, Liverpool, England

How good are *in silico* systems for predicting skin metabolism (pro-haptens)?

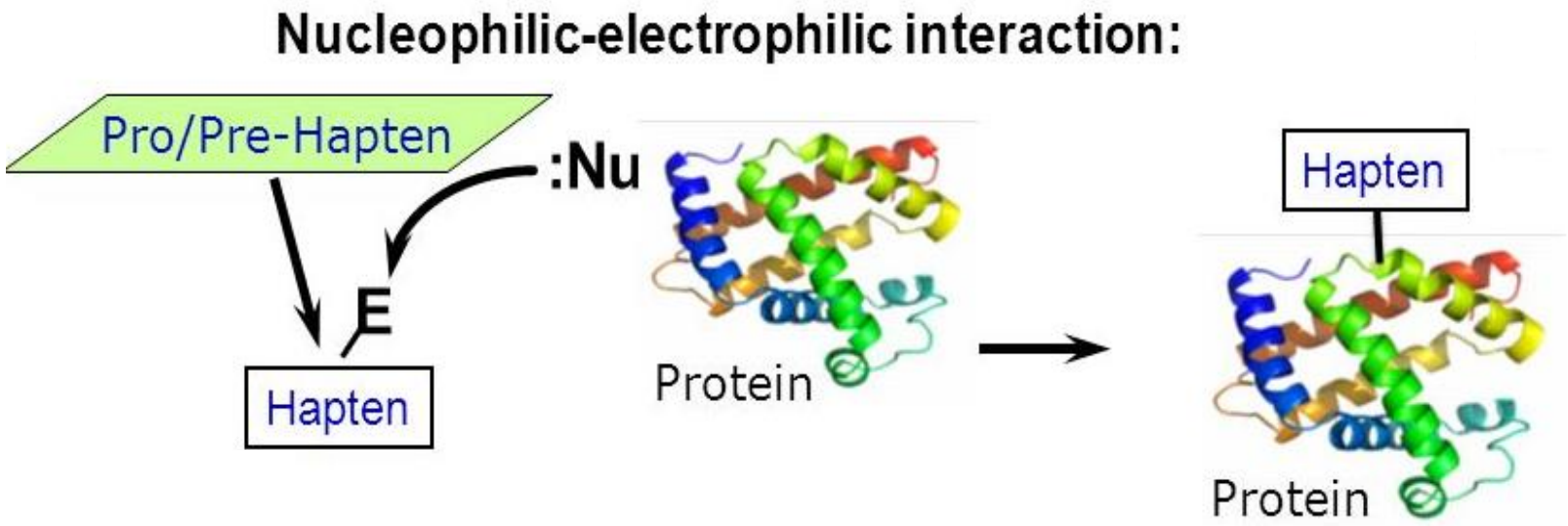
Why is this?

Do we need more experiments?

Skin Sensitisation: Electrophiles

- Pro-/pre-haptens are considered to make up about 25% of the chemical space related to skin sensitisation

- The chemistry of many pro-/pre-haptens is very closely related – indirect electrophiles is a more appropriate term



Patlewicz et al (2016) Regulatory Toxicology and Pharmacology, 82, p147

Skin Metabolism

- There are limited data around the exact expression levels of metabolising enzymes in the skin
- The most comprehensive review outlines (rat, mouse, pig, guinea pig and humans):
  - Low expression of the P450 system
  - Expression of oxidising enzymes (COX and FMO)
  - Expression of conjugating enzymes
- Essentially, the skin is capable of oxidation and conjugation

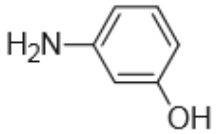
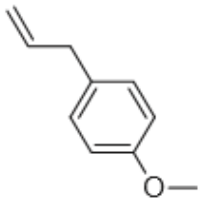
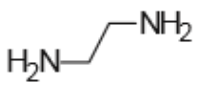
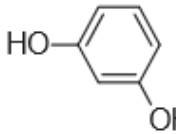
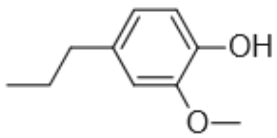
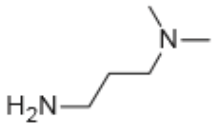
Oesch et al (2018) Archives of Toxicology, 92, p2411

Skin Sensitisation: Pre-/Pro-Haptens

- Numerous chemicals can behave as both pre- and pro-haptens (for example, hydroquinones)
- However, some chemicals are definite pro-electrophiles
- These chemicals are typically not identified in the DPRA
- Interestingly, some of these were identified by other assay from the ‘triple pack’ (Keratinosens and HCLat)

Patlewicz et al (2016) Regulatory Toxicology and Pharmacology, 82, p147

Table 2
Pro-haptens in the dataset.

Name	Structure	In vivo		Cell free in chemico	Cell based *in vitro*		Reaction mechanistic domain
		LLNA	Human data	DPRA	KeratinoSens™	h-CLAT	
3-Aminophenol		1	NA	0	0	1	pro-MA
4-Allylanisole		1	NA	1	0	1	pro-MA
Ethylenediamine (free base)		1	1	0	1	1	pro-SB
Resorcinol		1	1	0	0	1	pro-MA
Dihydroeugenol (2-methoxy-4-propyl-phenol)		1	NA	0	1	1	pro-MA
3-Dimethylamino propylamine		1	1	0	1	1	pro-SB[a]

Notes: NA – not available, 1 = positive, 0 = negative, SB = Schiff base former, MA = Michael acceptor.
[a] 3-Dimethylamino propylamine was originally assigned as a pre-SB in the original EC EURL-ECVAM report (Casati et al., 2016). It has been reconsidered as a pro-SB electrophile.

Table 2 taken from Patlewicz et al (2016) Regul Toxicol Pharmacol 82, p147

Pro-Hapten Identification for Skin Sensitisation: *In Silico* Tools

Derek nexus

Meteor nexus

Plus a number of other commercial software packages – for a review see reference below

Madden et al (2017) Computational Toxicology 3, p44

Toxtree - Toxic Hazard Estimation by decision tree approach

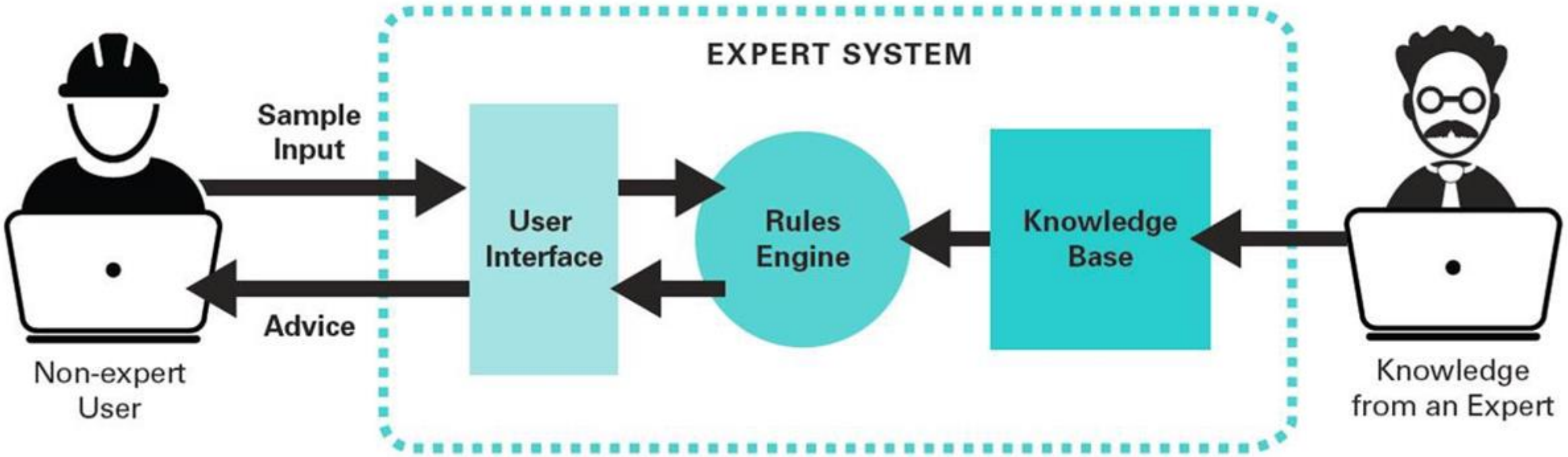
Toxtree is a full-featured and flexible user-friendly open source application, which is able to estimate toxic hazard by applying a decision tree approa

Toxtree could be applied to datasets from various compatible file types. User-defined molecular structures are also supported - they could be entered or by using the built-in 2D structure diagram editor.

The Toxtree application is suitable for a standalone PC and can run on any platform, supported by Java 1.6 or higher.

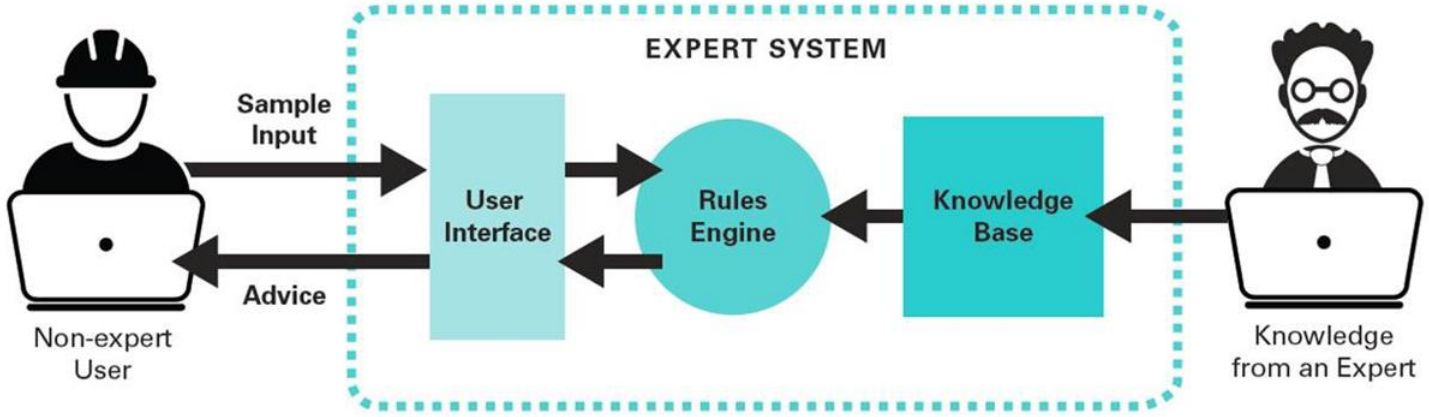
Toxtree has been designed with flexible capabilities for future extensions in mind (e.g. other classification schemes that could be developed at a futur decision trees with arbitrary rules can be built with the help of graphical user interface or by developing new plug-ins.

Expert Systems



Expert Systems – Knowledge Base

The knowledge base is constructed using structural alerts that relate chemistry to an effect (MIE, endpoint, metabolism etc.)

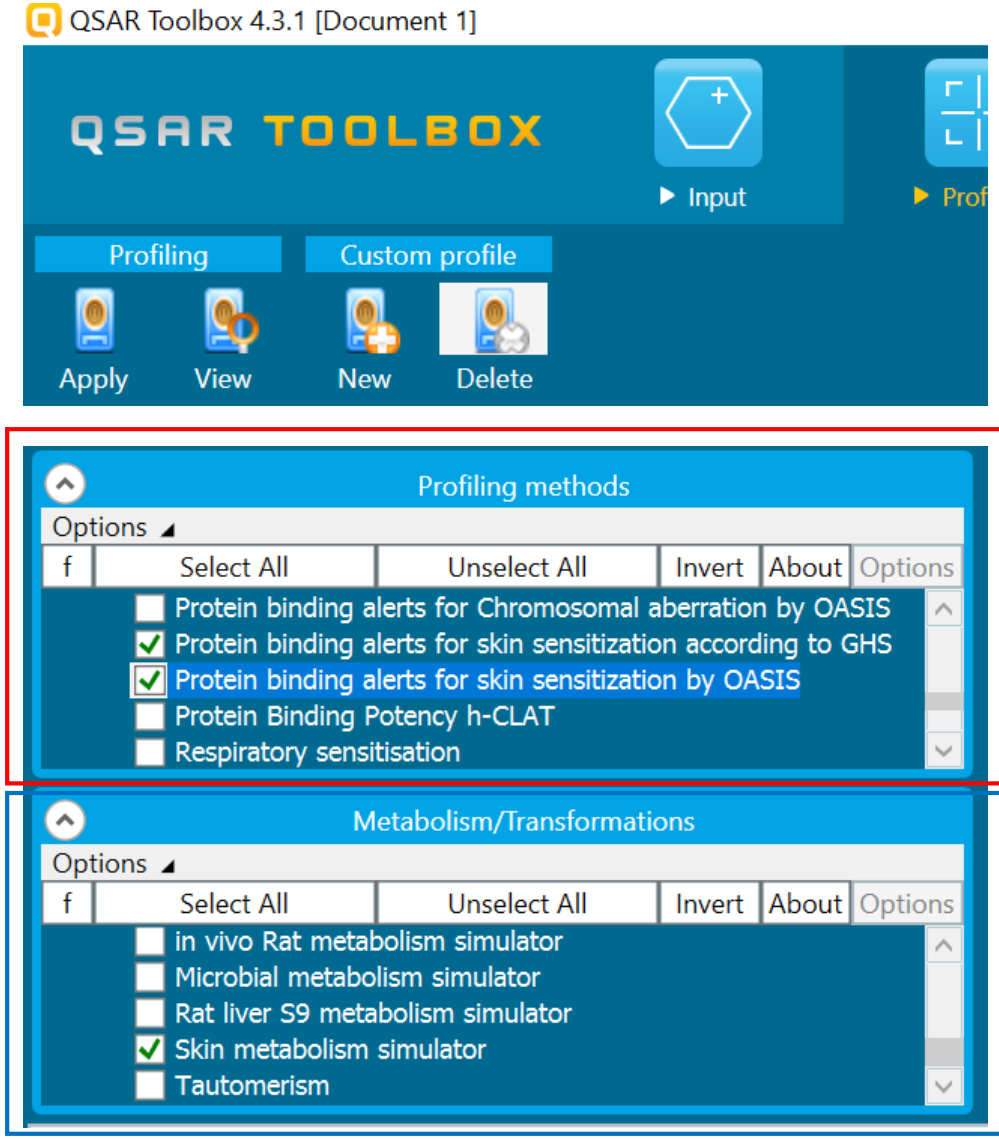


OECD QSAR Toolbox

- The OECD QSAR Toolbox is an *in silico* tool designed to facilitate read-across predictions (it is not a true expert system)
- It is chemistry-based, featuring a range of structural alert-based profilers
- These profilers are organised on an MIE and/or endpoint basis
- Contains the knowledge of other freely available tools such as ToxTree (certainly in terms of skin sensitisation)

OECD QSAR Toolbox is available from: qsartoolbox.org

OECD QSAR Toolbox – Relevant Profilers



Protein binding by OECD & OASIS
Protein binding alerts for skin sensitisation according to GHS & OASIS

Skin metabolism simulator
Hydrolysis simulator
Dissociation simulator

OECD QSAR Toolbox – Skin Metabolism Simulator

- Rules initially derived from an understanding of liver metabolism and skin sensitisation data derived from the LLNA, GPMT and other sources
- Subsequent analysis has improved the skin metabolism rule base with the availability of biotransformations for 151 chemicals
- Full implementation is found in the commercial TIMES software, with a simplified version being available in the OECD QSAR Toolbox
- Full version adds probability and reliability scores for each predicted biotransformation

Mekenyan et al (2012) SAR QSAR Environ Res, 23, p553; TIMES available from http://oasis-lmc.org/

Derek prediction

Derek Nexus (Lhasa Ltd.)

- A true knowledge base system for predictive toxicology
- Contains structural alerts for a number of endpoints, including skin sensitisation
- The Derek Nexus knowledge base for skin sensitisation contains alerts for direct and indirect electrophiles
- Contains a reasoning engine that offers a five levels of confidence in the predictions (based on the extent of the underlying knowledge)
- Levels of confidence being: impossible, improbable, doubted, equivocal, plausible, probable, certain

Meteor prediction

Meteor Nexus (Lhasa Ltd.)

- No specific skin metabolism model
- Instead, there's a range of potential biotransformations
- As with Derek Nexus, these predictions have an associated level of confidence associated with them: improbable, doubted, equivocal, plausible, probable, certain
- As outlined earlier it is known that skin and liver metabolism differ, with differing expression levels of the key enzymes
- Careful considerations of the relevant enzymes is required in order to prevent Meteor being over predictive of potential metabolites

Profiling Examples

- The next slides outline how we can identify potential haptens for:
- Cinnamyl formate (CAS No 104-65-4)
- Isoeugenol acetate (93-29-8)
- Cinnamic aldehyde dimethyl acetate (4364-06-1)
- Ethylenediamine (107-15-3)

Patlewicz et al (2016) Regulatory Toxicology and Pharmacology, 82, p147

Cinnamyl Formate (104-65-4)

QSAR Toolbox 4.3.1 [Document 1]

QSAR TOOLBOX

Input | Profiling | Data | Category definition | Data Gap Filling

Profiling | Custom profile

Apply | View | New | Delete

Documents

Document 1

CAS: 104654

Profiling methods

Options

f | Select All | Unselect All | Invert | About | Optior

- [ ] Protein binding alerts for Chromosomal al
- [x] Protein binding alerts for skin sensitizatio
- [x] Protein binding alerts for skin sensitizatio
- [ ] Protein Binding Potency h-CLAT

Metabolism/Transformations

Options

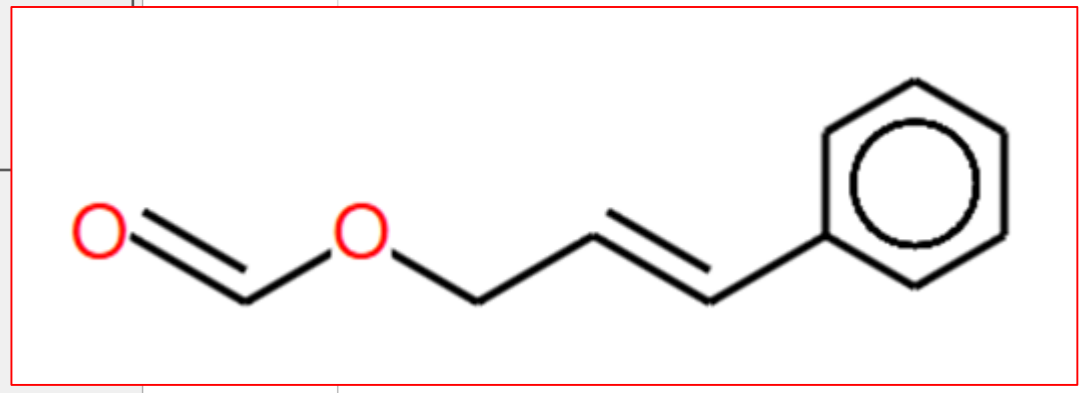
f | Select All | Unselect All | Invert | About | Optior

- [ ] Microbial metabolism simulator
- [ ] Rat liver S9 metabolism simulator
- [x] Skin metabolism simulator
- [ ] Tautomerism

Filter endpoint tree...

1 [target]

Structure



Parameters	
Physical Chemical Properties	
Environmental Fate and Transport	
Ecotoxicological Information	
Human Health Hazards	
Profile	
General Mechanistic	
Protein binding by OASIS	No alert found
Protein binding by OECD	No alert found
Endpoint Specific	
Protein binding alerts for skin sensitization according to GHS	No alert found
Protein binding alerts for skin sensitization by OASIS	No alert found
Metabolism/Transformations	
Skin metabolism simulator	8 metabolite(s)

Cinnamyl Formate (104-65-4)

QSAR Toolbox 4.3.1 [Document 1]

QSAR TOOLBOX

Input | Profiling | Data | Category definition | Data Gap Filling

Profiling | Custom profile

Apply | View | New | Delete

Documents

Document 1

CAS: 104654

Profiling methods

Options

f | Select All | Unselect All | Invert

Protein binding alerts for Chr

Protein binding alerts for skin

Protein binding alerts for skin

Protein Binding Potency h-CL

Metabolism/Transformatic

Options

f | Select All | Unselect All | Invert

Microbial metabolism simulato

Rat liver S9 metabolism simul

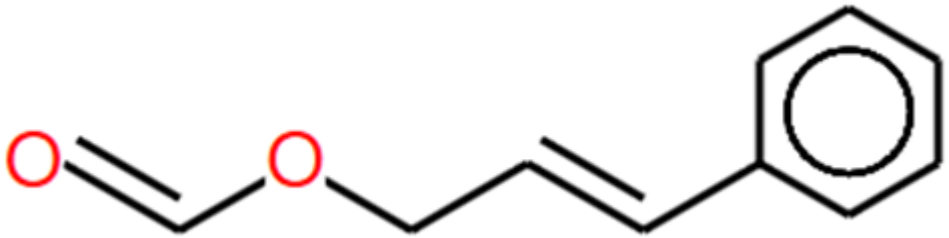
Skin metabolism simulator

Tautomerism

Filter endpoint tree...

1 [target]

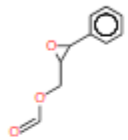
Structure



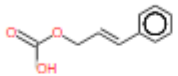
Parameters	
Physical Chemical Properties	
Environmental Fate and Transport	
Ecotoxicological Information	
Human Health Hazards	
Profile	
General Mechanistic	
Protein binding by OASIS	No alert found
Protein binding by OECD	No alert found
Endpoint Specific	
Protein binding alerts for skin sensitization according to GHS	No alert found
Protein binding alerts for skin sensitization by OASIS	No alert found
Metabolism/Transformations	
Skin metabolism simulator	8 metabolite(s)

Cinnamyl Formate (104-65-4)

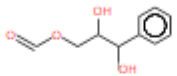
metabolite #1
No CAS number



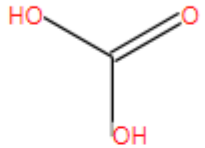
metabolite #2
No CAS number



metabolite #3
No CAS number

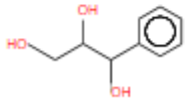


metabolite #4
No CAS number

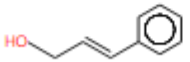


metabolite #5
No CAS number

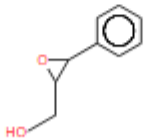
metabolite #6
No CAS number

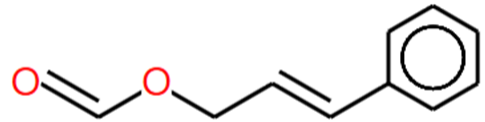


metabolite #7
No CAS number



metabolite #8
No CAS number





Metabolism/Transformations

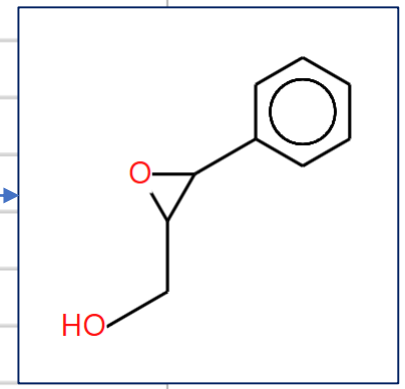
Skin metabolism simulator

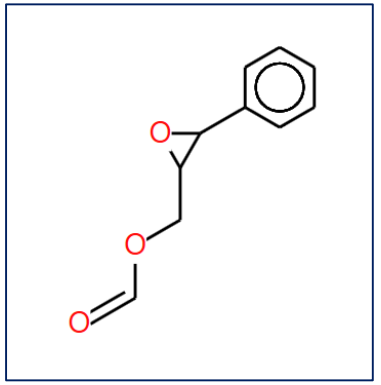
8 metabolite(s)

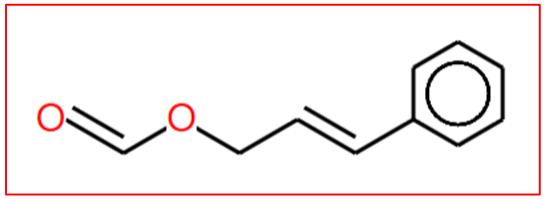
Cinnamyl Formate (104-65-4)

Profiling results

- 8 metabolite(s)
  - 2 x Skin sensitization Category 1B
    - 2 x Epoxides, Aziridines and Sulfuranes_low activity
      - O=COCC1OC1c1ccccc1
      - OCC1OC1c1ccccc1
  - 6 x No alert found
    - OC(=O)OCC=Cc1ccccc1
    - OC(COC=O)C(O)c1ccccc1
    - OC(O)=O
    - OC=O
    - OCC(O)C(O)c1ccccc1
    - OCC=Cc1ccccc1



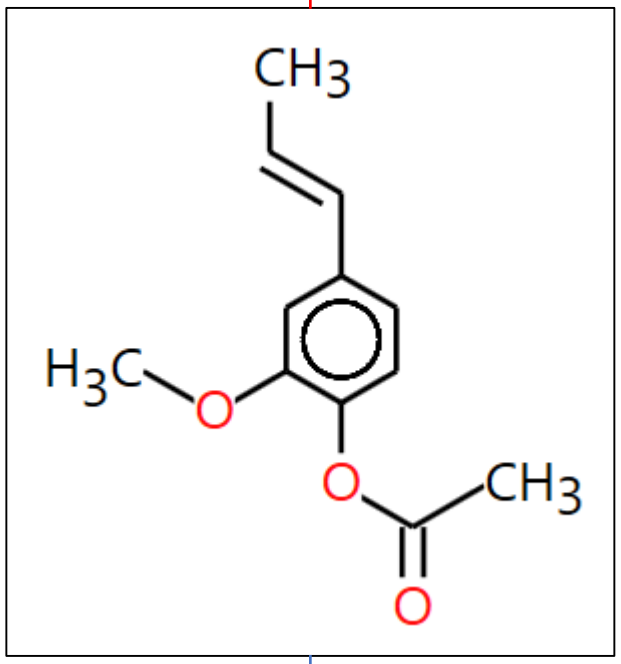


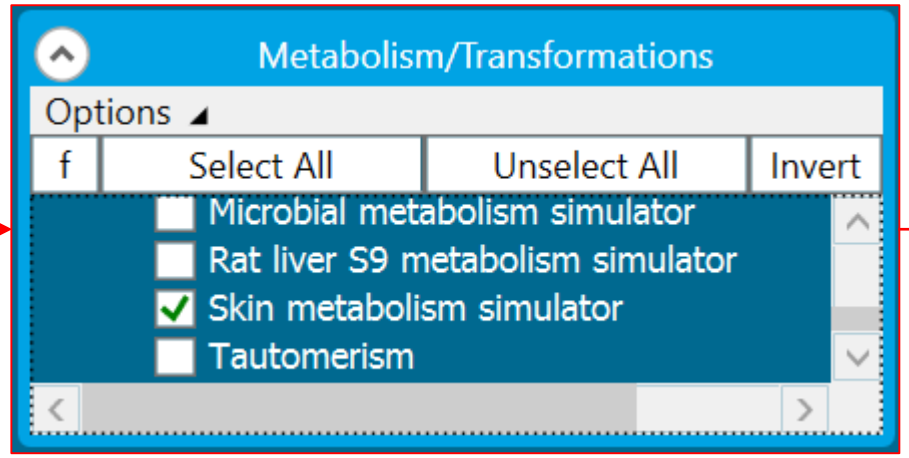


Protein binding by OECD	2 x SN2
Endpoint Specific	
Protein binding alerts for skin sensitization according to GHS	2 x Skin sensitization Category 1B
Protein binding alerts for skin sensitization by OASIS	2 x SN2

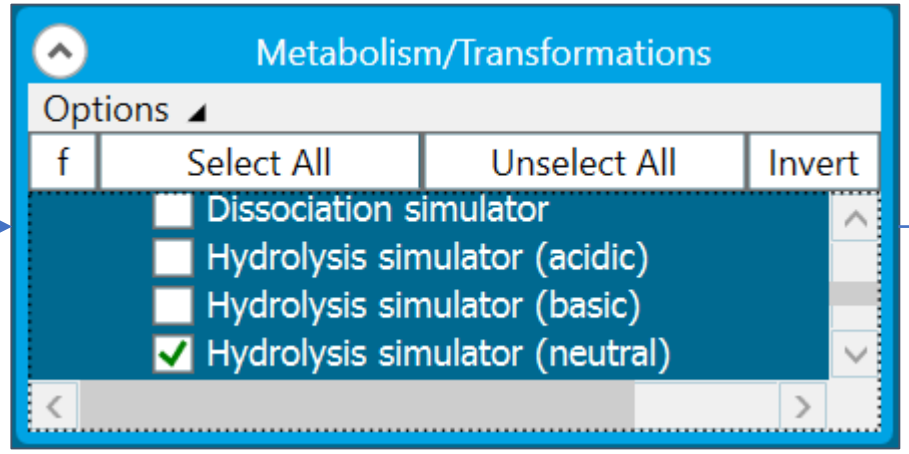
This is in contrast to the known liver metabolism in which cinnamic aldehyde is produced via the action of esterases followed by alcohol dehydrogenase
We might consider category 1B following metabolism to be a non-sensitiser

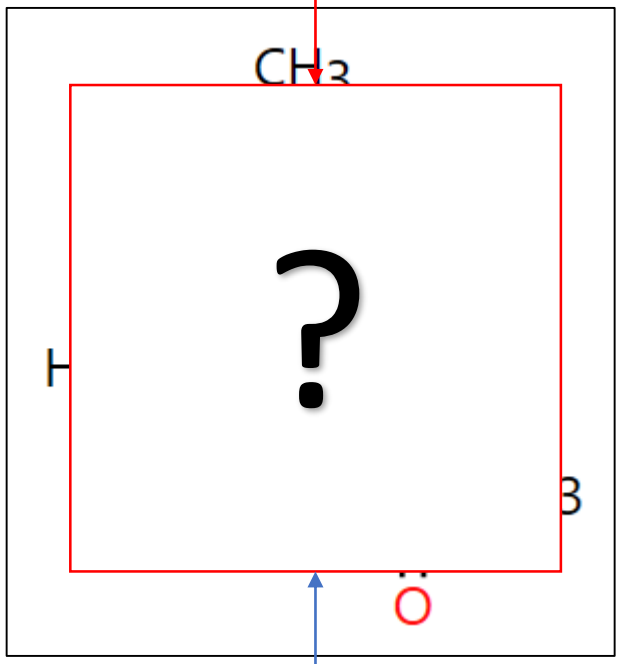
Isoeugenol Acetate (93-29-8)





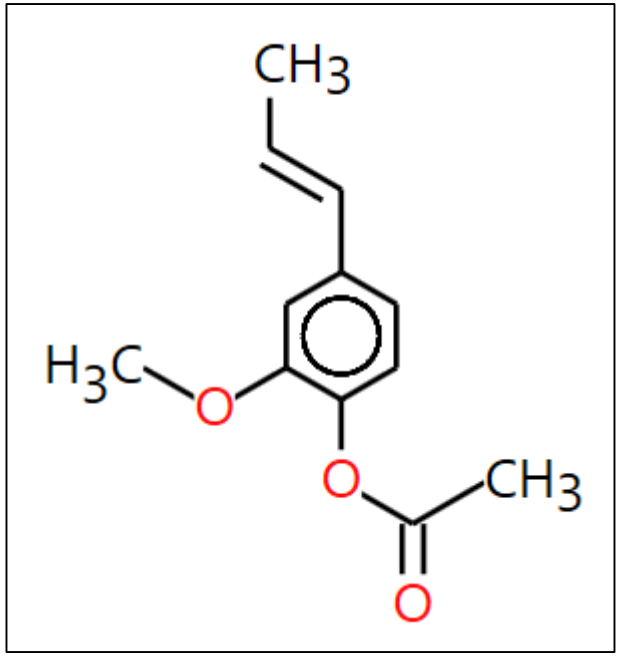
Metabolism versus hydrolysis



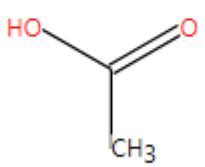


Isoeugenol Acetate (93-29-8)

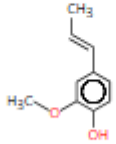
Profile	
General Mechanistic	
Protein binding by OASIS	No alert found
Protein binding by OECD	Acylation
Endpoint Specific	
Protein binding alerts for skin sensitiz...	No alert found
Protein binding alerts for skin sensitiz...	No alert found
Metabolism/Transformations	
Hydrolysis simulator (neutral)	2 metabolite(s)
Skin metabolism simulator	3 metabolite(s)



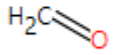
metabolite #1
No CAS number



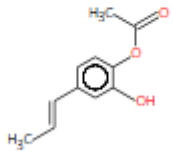
metabolite #2
No CAS number



metabolite #1
No CAS number



metabolite #2
No CAS number

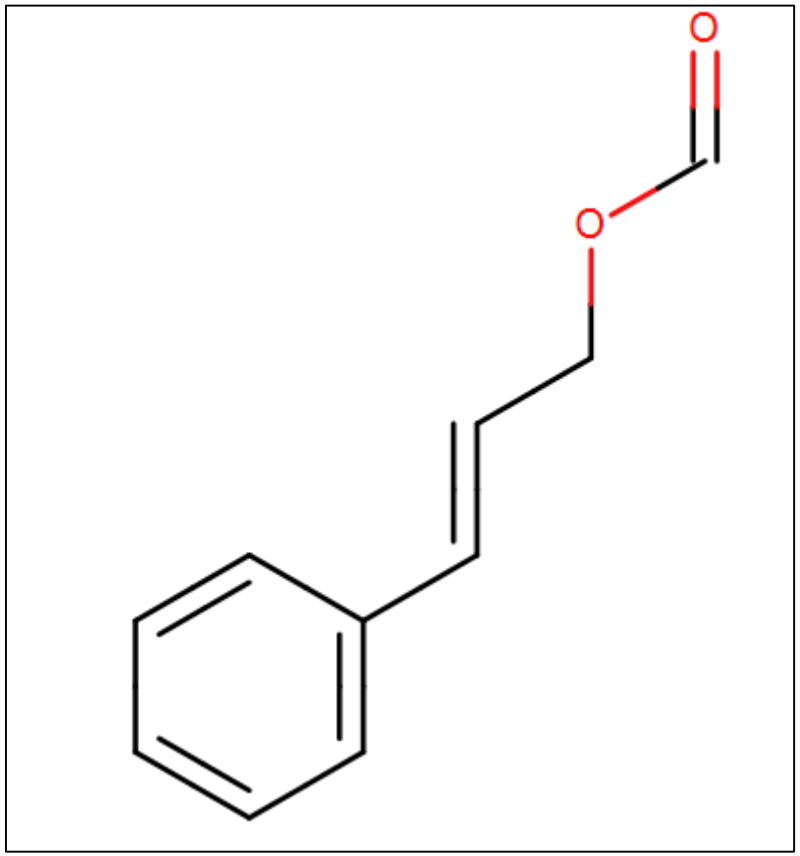


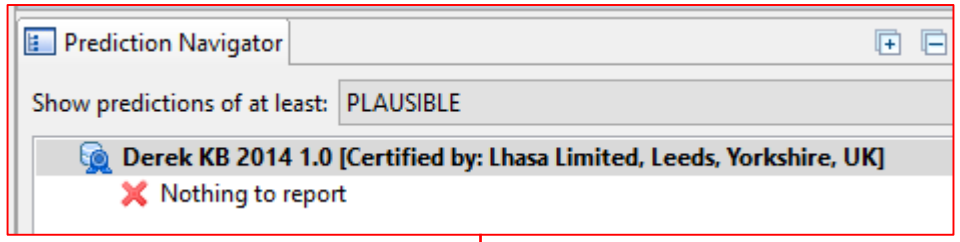
metabolite #3
No CAS number

Summary: OECD QSAR Toolbox

- The examples we have looked at showed that we can identify potential electrophiles due to skin metabolism and hydrolysis
- Profilers are available for the prediction of skin metabolism, hydrolysis and dissociation
- However, we cannot offer any indication of the rate of these reactions or the ratios in which the metabolites exist (probability/reliability scores are available in TIMES)
- Knowledge bases are only as good as the data on which they are built

Derek prediction





Derek Nexus returns ‘nothing to report’ if there is no information about the chemical of interest within the knowledge base

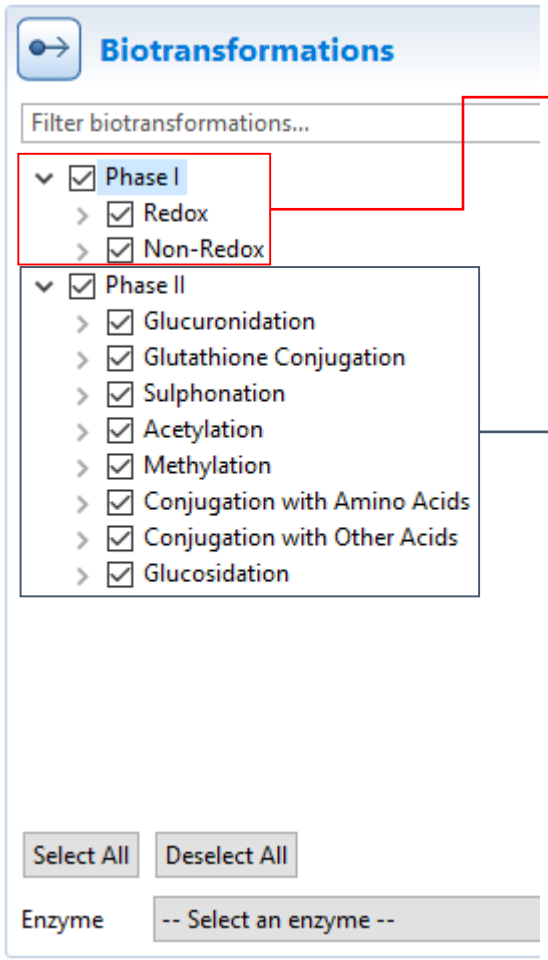
The next step is to investigate metabolism

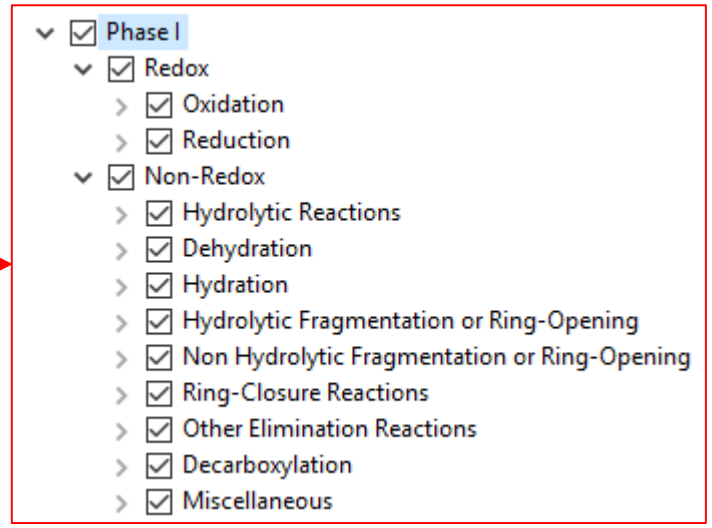
Meteor prediction

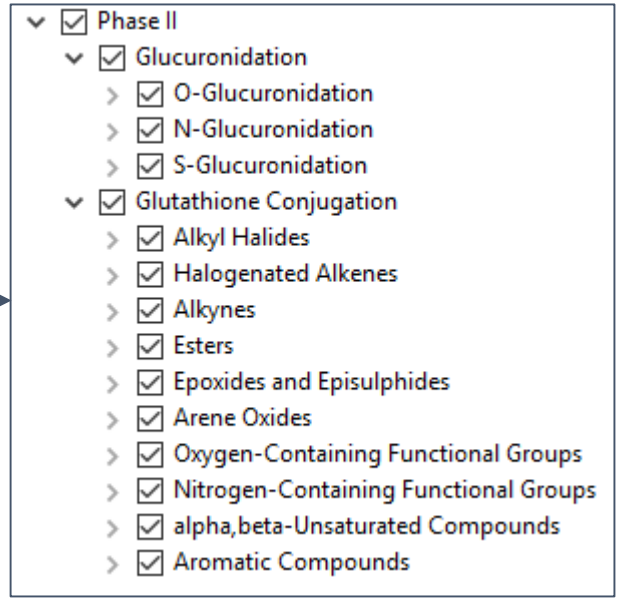
Lhasa Limited

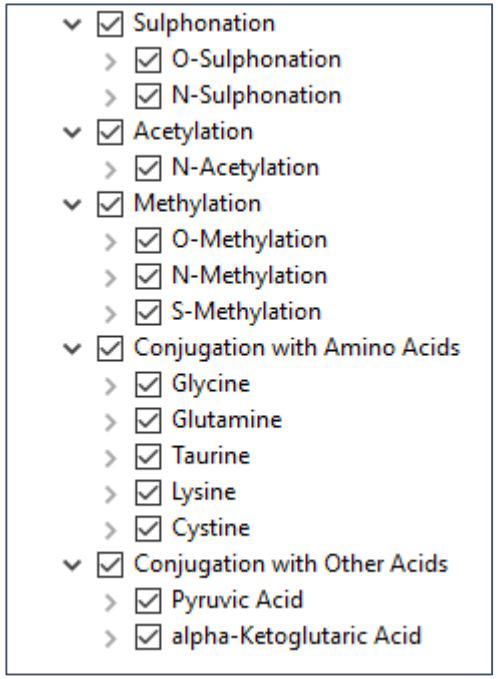
The benefit of this approach is that you can design your own profiling scheme for enzymes/biotransformations of interest

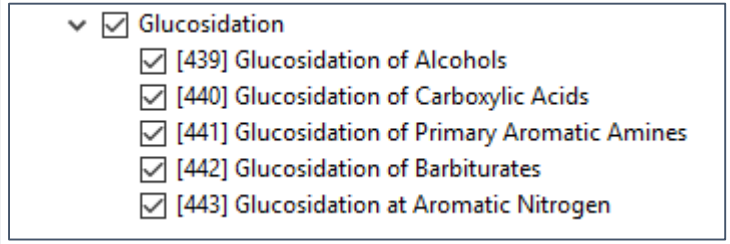
In this example we have selected oxidation and reduction phase 1 transformations only



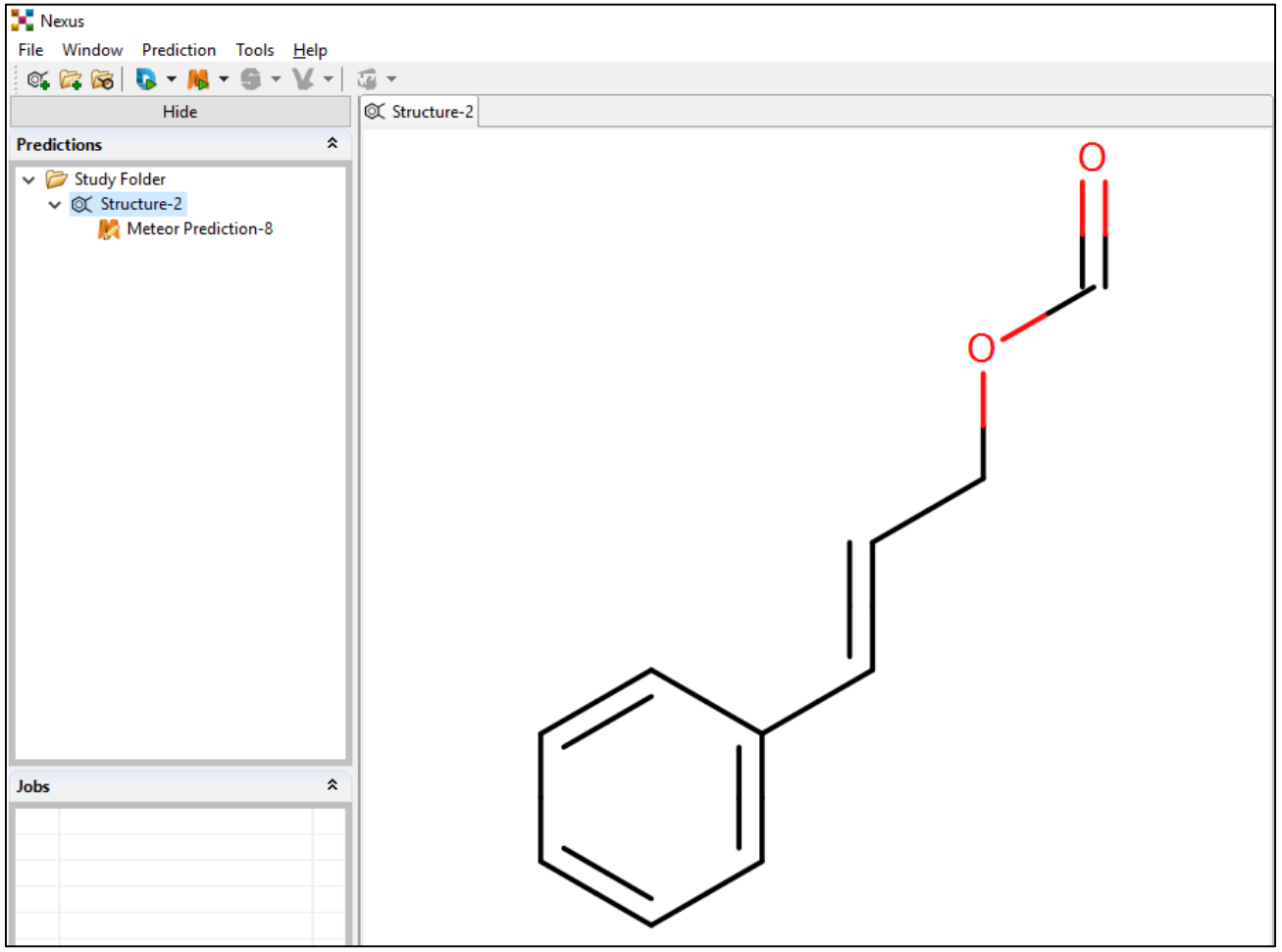


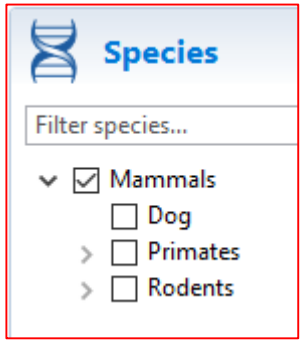


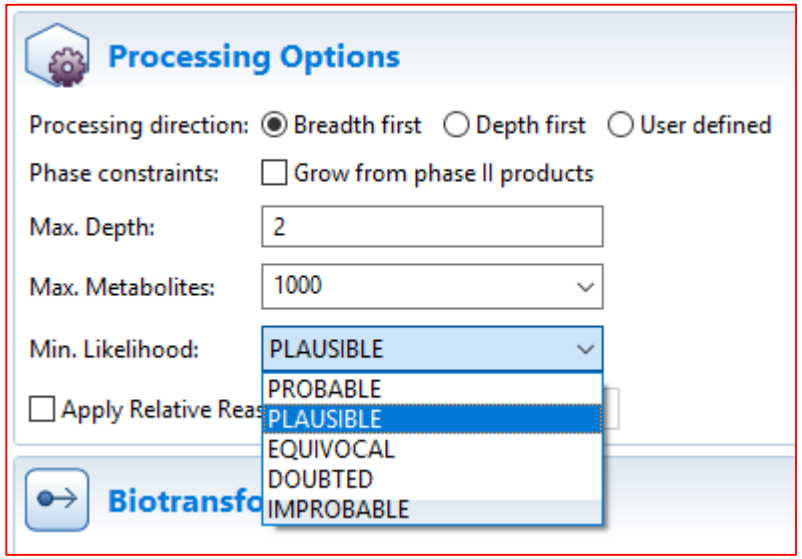




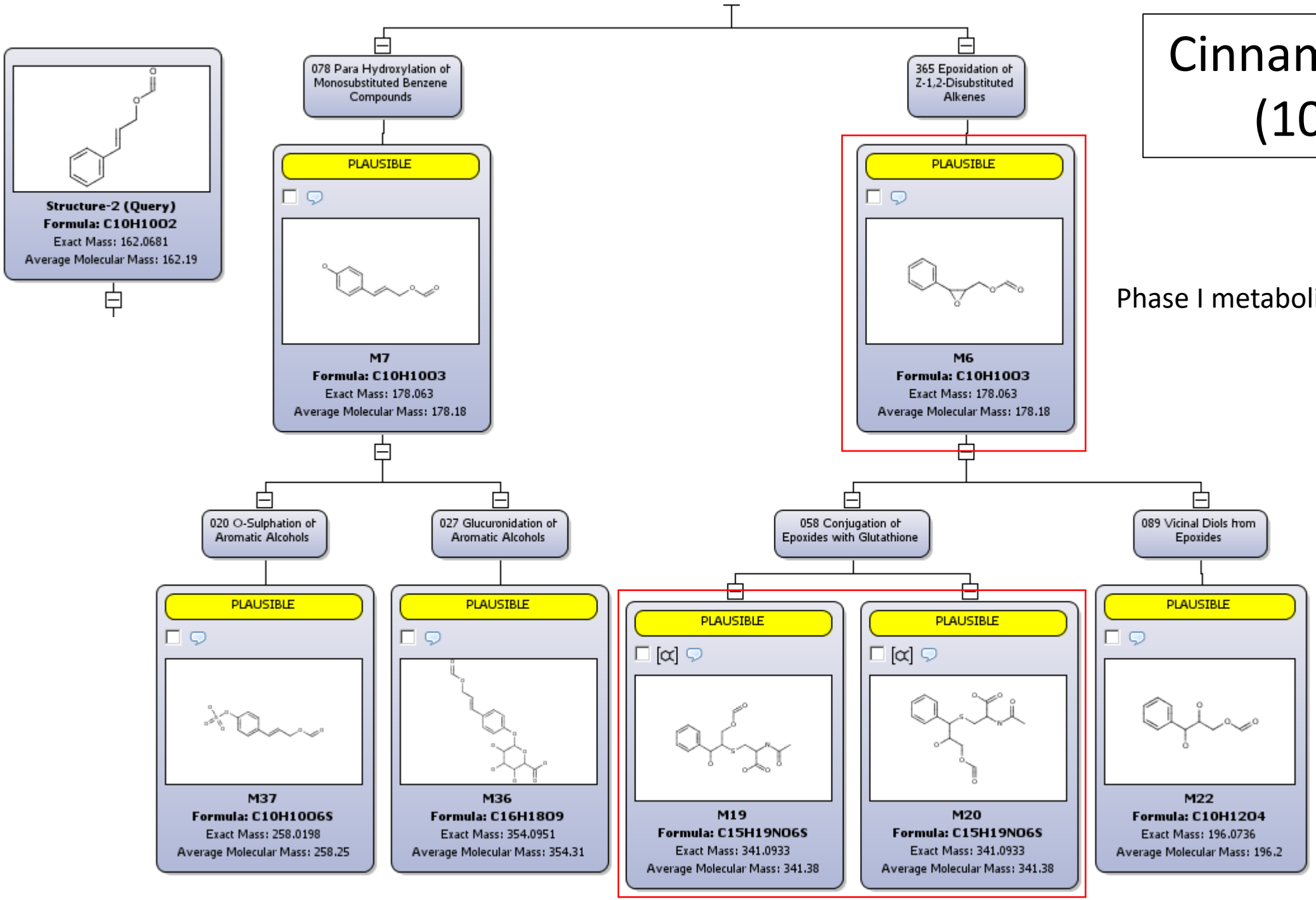
Cinnamyl Formate (104-65-4)







Cinnamyl Formate (104-65-4)

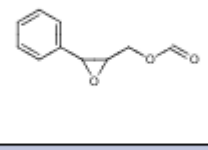


Phase I metabolism

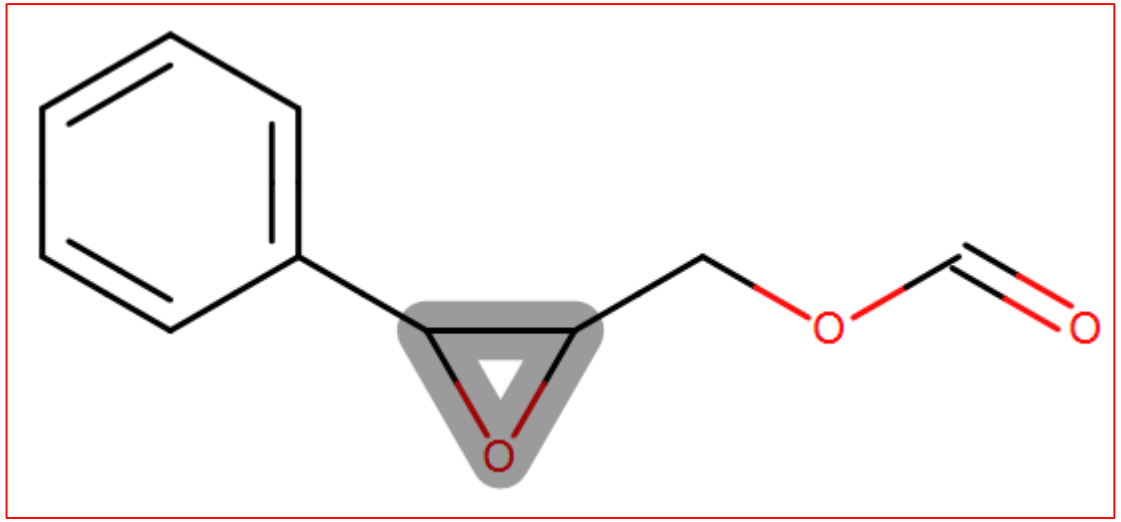
Phase II metabolism

Derek prediction

PLAUSIBLE



M6
Formula: C10H10O3
Exact Mass: 178.063
Average Molecular Mass: 178.18



Endpoints

Derek KB 2014 1.0

Filter endpoints...

- ☐ Carcinogenicity (ALL)
- ☐ Genotoxicity (ALL)
- ☐ Irritation (ALL)
- ☐ Miscellaneous endpoints (ALL)
- ☐ Neurotoxicity (ALL)
- ☐ Organ toxicity (ALL)
- ☐ Reproductive toxicity (ALL)
- ☐ Respiratory sensitisation (ALL)
- ■ Skin sensitisation (ALL)
  - ☐ Photoallergenicity
  - ☑ Skin sensitisation

Species

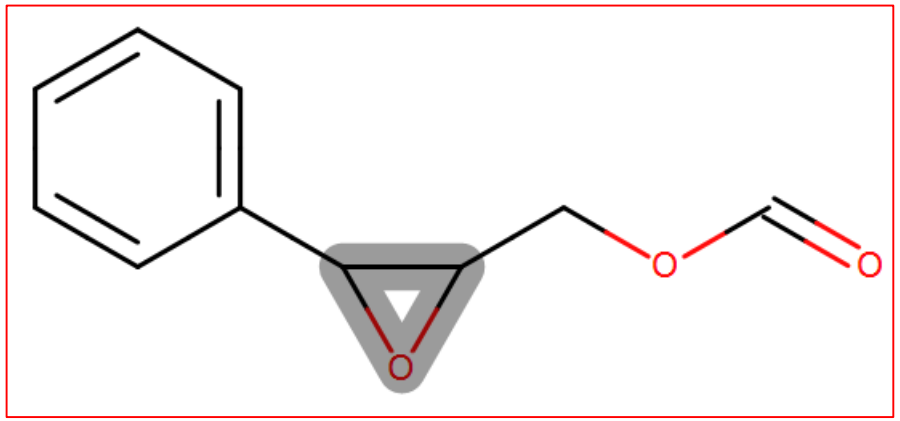
Filter species...

- ☑ bacterium
  - ☐ Escherichia coli
  - ☐ Salmonella typhimurium
- ☐ chicken
- ☑ mammal
  - ☐ dog
  - ☐ primate
    - ☐ human
    - ☐ monkey
  - ☐ rabbit
  - ☐ rodent
    - ☐ guinea pig
    - ☐ hamster
    - ☐ mouse
    - ☐ rat
- ☐ unspecified
- ☐ various

Prediction Navigator

Show predictions of at least: PLAUSIBLE

- **Derek KB 2014 1.0** [Certified by: Lhasa Limited, Leeds, Yorkshire, UK]
  - **Skin sensitisation**
    - **mammal** - PLAUSIBLE
      - **Alert** - 433: Epoxide



Prediction Navigator

Show predictions of at least: PLAUSIBLE

- **Derek KB 2014 1.0** [Certified by: Lhasa Limited, Leeds, Yorkshire, UK]
  - **Skin sensitisation**
    - **mammal** - PLAUSIBLE
      - **Alert** - 433: Epoxide

Comments

The activity of glycidyl ethers, amines, esters and amides is described in a separate alert.

The presence of a skin sensitisation structural alert within a molecule indicates the molecule has the potential to cause skin sensitisation. Whether or not the molecule will be a skin

Validation Comments

Skin sensitisation: guinea pig maximisation test, local lymph node assay

The alert has demonstrated the following predictive performance:

1) Cronin and Basketter data set: 2 compounds activate this alert of which 2 are reported positive (positive predictivity = 100%)

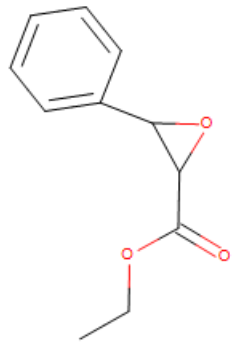
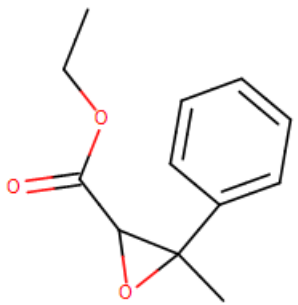
Endpoints

ID	Name	Parent
4189	Skin sensitisation	Skin sensitisation (ALL)

References

ID	Title	Author	Source	Year	Suppler
4853	Multivariate QSAR analysis of a skin sensitization databas	Cronin MTD and Basketter DA.	SAR and QSAR in Environmental Rese	1994	Note: d
4914	The use of nuclear magnetic resonance spectrometry (1H	Betso JE, Carreon RE and Miner VM.	Toxicology and Applied Pharmacolog	1991	

Examples

3-phenyl-oxirane-2-carboxylic acid	-methyl-3-phenyl-oxiranecarboxyli			
				

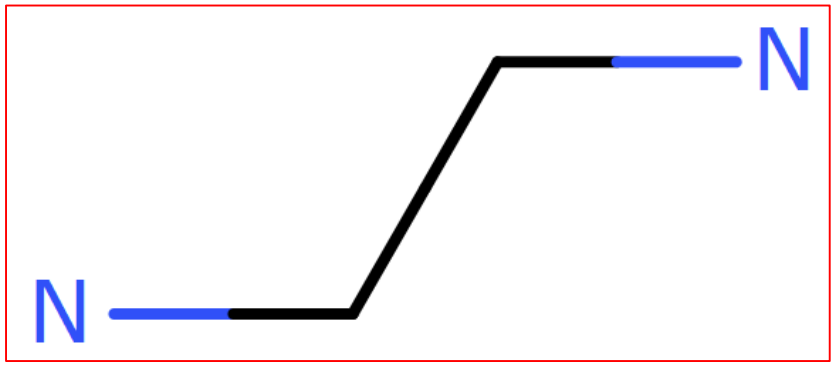
Summary: Meteor and Derek Nexus

- The example we have looked at showed that we can identify potential electrophiles due to skin metabolism
- The knowledge base in Meteor Nexus identifies potential metabolites, these can be screened for skin sensitisation potential using Derek Nexus
- Derek Nexus also contains structural alerts for some pro-electrophiles
- Predictions in both systems are given a level of confidence; however, potential biotransformation rates and ratios are not available
- Knowledge bases are only as good as the data on which they are built

OECD QSAR Toolbox & Nexus

- Occasionally these two leading tools will give differing profiling results
- This is usually around whether an alert has been encoded as a part of the ‘toxicity’ or ‘metabolism’ knowledge
- For example, consider profiling ethylenediamine (a pro-electrophile from the Patlewicz et al study)

Derek prediction



Derek Nexus contains direct and indirect alerts for skin sensitisation in its knowledge base

Ethylenediamine (107-15-3)

Prediction Navigator

Show predictions of at least: PLAUSIBLE

Derek KB 2014 1.0 [Certified by: Lhasa Limited, Leeds, Yorkshire, UK]

Skin sensitisation

mammal - PROBABLE

Alert - 435: Diamine

Example - 1,2-diaminoethane

Comments

The presence of a skin sensitisation structural alert within a molecule indicates the molecule has the potential to cause skin sens
sensitiser will also depend upon its percutaneous absorption. Generally, small lipophilic molecules are more readily absorbed int
sensitisation.

Validation Comments

Skin sensitisation: guinea pig maximisation test, local lymph node assay

The alert has demonstrated the following predictive performance:

1) Cronin and Basketter data set: 1 compound activates this alert of which 1 is reported positive (positive predictivity = 100%)

References

ID	Title	Author
4847	Chemikalien und Kontaktallergie: Eine Bewertende Zusa	Kayser D and Schlede E (editors).
4853	Multivariate QSAR analysis of a skin sensitization databa	Cronin MTD and Basketter DA.
4868	A comparison of three guinea-pig sensitization procedu	Goodwin BFJ, Crevel RWR and Johns
4910	Contact allergy to impurities in surfactants: amount, che	Angelini G, Rigano L, Foti C, Vena GA

Ethylenediamine (107-15-3)

QSAR Toolbox 4.3.1 [Document 1]

QSAR TOOLBOX

Input | Profiling | Data | Category definition | Data Gap Filling

Profiling | Custom profile

Apply | View | New | Delete

Documents

Document 1

Search chemical

Profiling methods

Options

f | Select All | Unselect All | Invert

- [ ] Oncologic Primary Classification
- [ ] Protein binding alerts for Chromosomal
- [x] Protein binding alerts for skin sensitizat
- [x] Protein binding alerts for skin sensitizat

Metabolism/Transformations

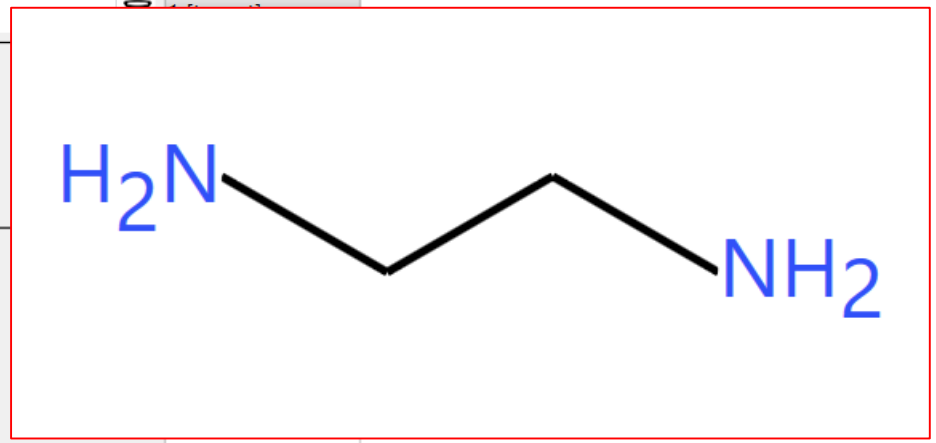
Options

f | Select All | Unselect All | Invert

- [ ] Microbial metabolism simulator
- [ ] Rat liver S9 metabolism simulator
- [x] Skin metabolism simulator
- [ ] Tautomerism

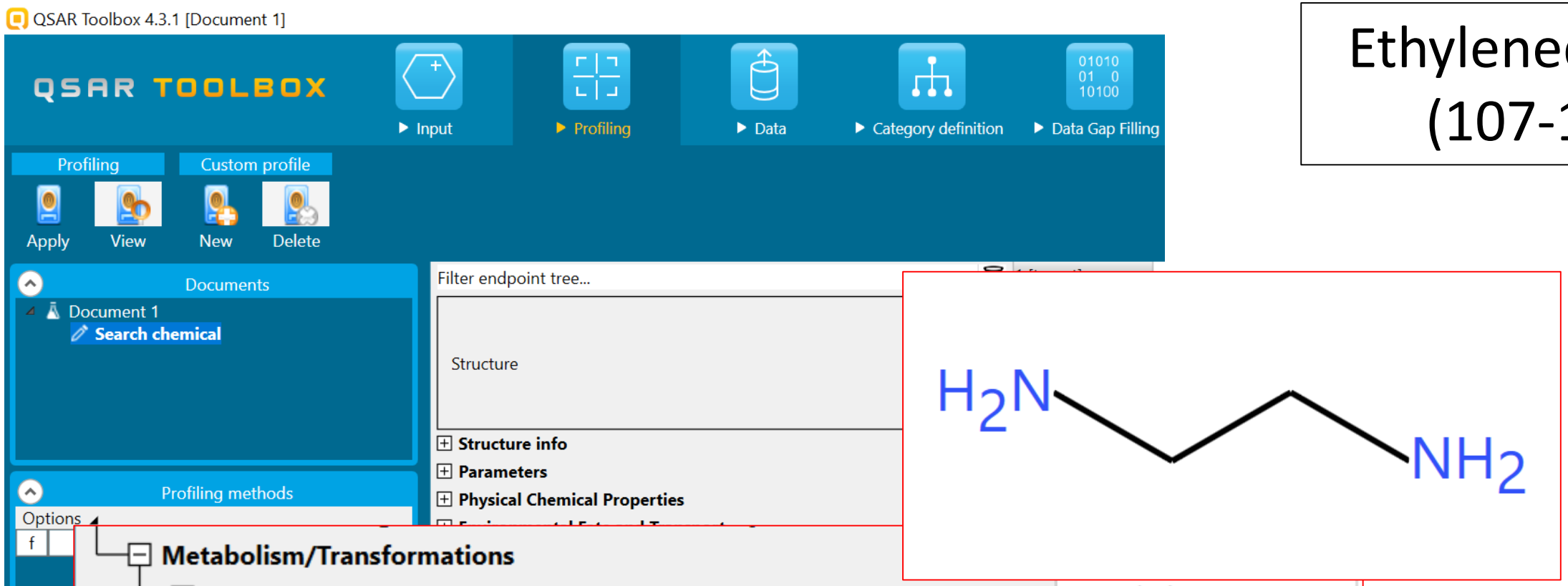
Filter endpoint tree...

Structure

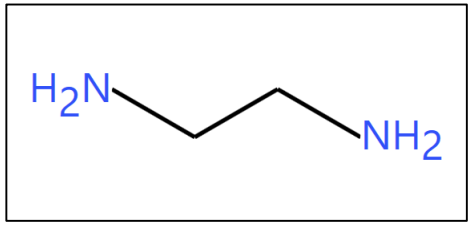


- Structure info
- Parameters
- Physical Chemical Properties
- Environmental Fate and Transport
- Ecotoxicological Information
- Human Health Hazards
- Profile
  - General Mechanistic
    - Protein binding by OASIS — No alert found
    - Protein binding by OECD — No alert found
  - Endpoint Specific
    - Protein binding alerts for skin sensitization according to GHS — No alert found
    - Protein binding alerts for skin sensitization by OASIS — No alert found
  - Metabolism/Transformations
    - Skin metabolism simulator — 5 metabolite(s)

Ethylenediamine (107-15-3)

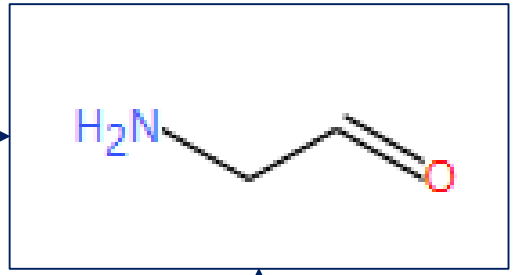


Ethylenediamine (107-15-3)



Profiling results

- 5 metabolite(s)
  - 2 x No alert found
    - N
    - NCC(O)=O
  - 2 x Skin sensitization Category 1B
    - 2 x Aldehydes
      - NCC=O
      - OC(=O)C=O
  - 1 x Skin sensitization Category 1A
    - 1 x Bis Aldehydes
      - O=CC=O



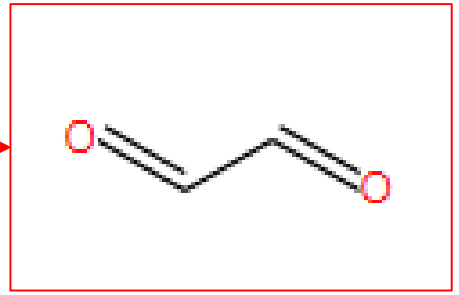
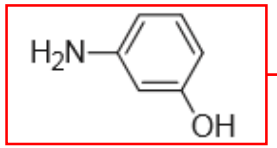
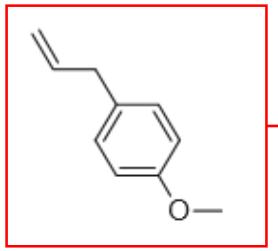
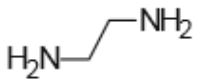
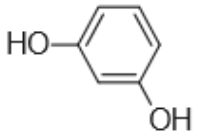
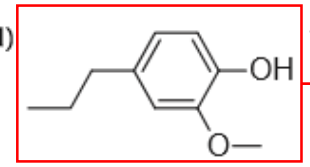
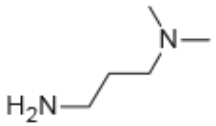
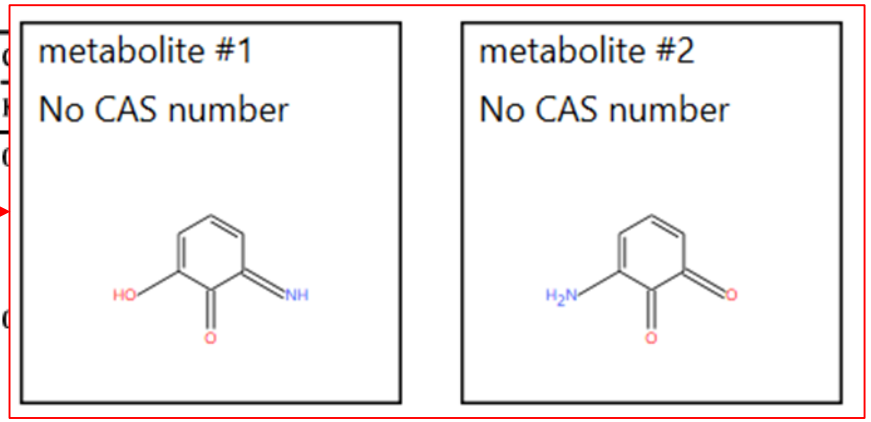
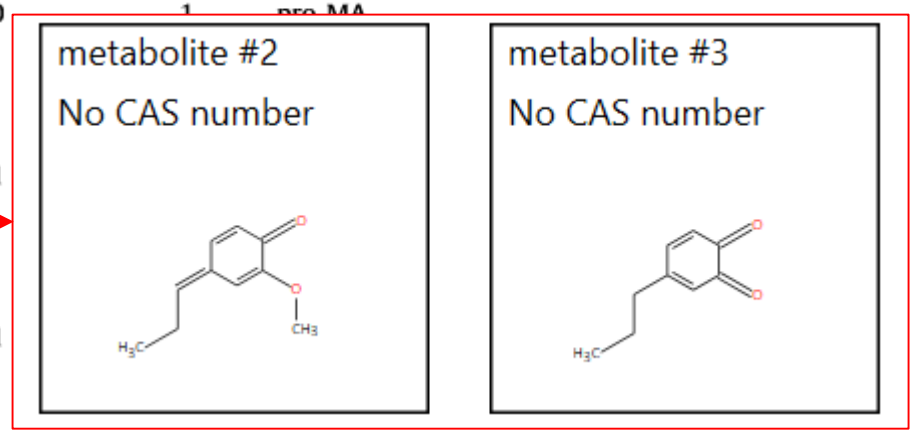


Table 2
Pro-haptens in the dataset.

Name	Structure	In vivo		Cell free in chemico			
		LLNA	Human data	DPRA			
3-Aminophenol		1	NA	0			
4-Allylanisole		1	NA	1			
Ethylenediamine (free base)		1	1	0	1	1	pro-SB
Resorcinol		1	1	0	0		
Dihydroeugenol (2-methoxy-4-propyl-phenol)		1	NA	0	1		
3-Dimethylamino propylamine		1	1	0	1		





Notes: NA – not available, 1 = positive, 0 = negative, SB = Schiff base former, MA = Michael acceptor.
[a] 3-Dimethylamino propylamine was originally assigned as a pre-SB in the original EC EURL-ECVAM report (Casati et al., 2016). It has been reconsidered as a pro-SB electrophile.

Table 2 taken from Patlewicz et al (2016) Regul Toxicol Pharmacol 82, p147

How good are *in silico* systems for predicting skin metabolism (pro-haptens)?

Why is this?

Do we need more experiments?

Conclusions

- There are number of ‘expert’ systems available that can predict the potential of a chemical to react covalently with proteins
- The Nexus programmes are true expert systems, whilst we have used the knowledge within the OECD QSAR Toolbox as it were
- The encoded knowledge in both systems is very good at identifying potential direct and indirect electrophiles capable of causing skin sensitisation
- However, neither system can predict the rates of metabolite formation or the ratios of the various metabolites
- Lhasa products are commercial, whilst the OECD QSAR Toolbox is free