AOP-Sens™

iS Tox Platform™
Machine-learning models for a new toxicology

Report

Client:
Altox

Username:
Tiago

Study Number:
AOPSens_Compound3_

Date:
2019/06/19 - 15:23:10

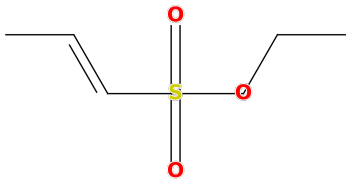
Program Version: 1.7

Molecular Query

Name:
Compound 3

CAS:
NA

SMILES:
C/C=C/S(=O)(=O)OCC



logK$_{ow}$:
0.89

logD:
1.61

Model Summary

AOP-Sens™ (Adverse Outcome Pathway-based model for Skin Sensitisation) is a computational tool developed by Altox using knowledge derived from a combination of mechanisms, key events, and outcomes data related to skin sensitization assessment.[1]

This model integrates *in chemico* , *in vitro* and *in vivo* data in a targeted and predictive framework that simulates an Integrated Approach to Testing and Assessment (IATA) for sensitization; the model uses a combination of structural alerts and artificial intelligence (AI) models to predict responses for each individual model: Direct Peptide Reactivity Assay (DPRA), KeratinoSens™ , human Cell Line Activation Test (h-CLAT), U-SENS ™ (generously provided by the L'oreal research and innovation team)[4-7], Local Lymph Node Assay (LLNA), and human sensitization. An AOP-based prediction is obtained by integration of individual models for the AOP components balanced by adjustments and weight of evidence assessment.

To assess both direct and indirect haptens, this tool predicts the potential for metabolic activation (pro-hapten formation) using known Phase I reactions. AOP-Sens integrates alerts, expert system, statistical and machine learning-models to support regulatory decision-making and integrated testing strategies.

This tool is powered by 197 Structural Alerts for skin protein reactivity, 128 for 1A and 69 for 1B GHS classes. The datasets were rigorously curated with results for DPRA assay, KeratinoSens™, h-CLAT, U-SENS™, LLNA, human and overall (Human and Animal) sensitization, containing 195, 190, 161,220, 997, 389 and 6971 structures, respectively.

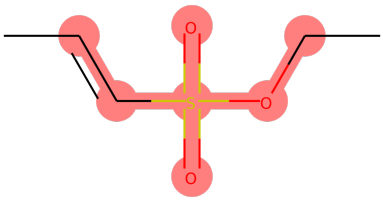
The final result is provided in a graphical plot visually summarising all the individual predictions, the applicability domain (AD), and relevant data for each endpoint.

Structural Alert Analysis

The Structural Alert Analysis module is based on fragments of the molecule assigned as reactive (i.e., *in vitro* protein binding alerts for skin sensitization according to GHS). These structural alerts can provide a mechanistic basis for the predicted skin sensitization, with the CLP/GHS potency sub-categories 1A or 1B, for interpretation of the mechanism based on theories and knowledge of toxicity mechanisms (OECD Principle 5).

Result: Positive (+)

Alerts were found in the molecule. The results are in the table below and a description is provided at the end of the report.

Category	Alert	Alert ID	References
Skin Sensitization Category 1A (Protein binding Alerts, EC3 (LLNA) ≤ 2%; NOEL (HRIPT)≤ 500 µg/cm2)		Polarised Alkenes - sulfonates	Roberts, D., Api, A.M., Safford, R., Lalko, J. Regulatory Toxicology and Pharmacology 72 (2015) 683–693. Aptula A.O. et al., (2006) Chemical Research in Toxicology, 19, 1097.

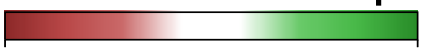
Machine Learning Models

The individual predictions below were obtained by Artificial Neural Networks (ANN) covering the AOP key events and outcomes:

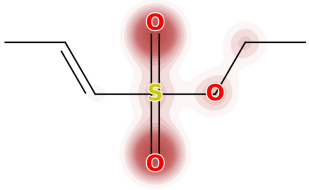
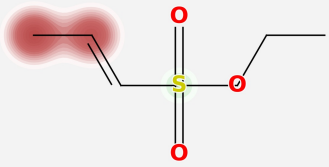
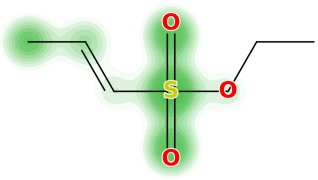
MIE (Molecular Initiating Event):	Reactivity and covalent interaction with cellular proteins (DPRA)
Key event 2 (KE2):	Activation of biochemical pathways by keratinocytes and release of pro-inflammatory mediators (KeratinoSens™)
Key event 3 (KE3):	Events in dendritic cells expression of co-stimulatory and adhesion molecules, by pathways-associated protein and gene expression (h-CLAT). Also, the change in the expression of a cell surface marker associated with the process of activation of monocytes and DC (U-Sens™).
Key event 4 (KE4) and Adverse Outcome:	T-cell proliferation and sensitization

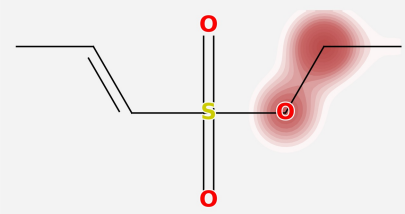
The ANN assign a category reactive (+) or non-reactive (-) for the DPRA, and a category active (+) or inactive (-) for the KeratinoSens™, h-CLAT and the U-Sens™. For the LLNA and human skin endpoints the ANN assign a category sensitizer (+) or nonsensitizer (-). To ensure transparency in the description of the deep learning model (an unambiguous algorithm - OECD Principle 2), more detailed information about each model is presented below. The STR (Structure-Toxicity Relationship) Probability Mapping indicates the fragments more related to the absence/decrease of toxicity (green) or presence/increase (red) as well, useful for hypotheses and mechanistic interpretations (OECD Principle 5).

STR Contribution Map



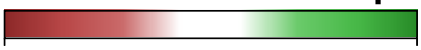
Reactive / Active (+) Non-Reactive / Inactive (-)

Assay/Event	Predicted class (Confidence)	STR Contribution Mapping
Direct Peptide Reactivity Assay (DPRA, OECD 442D) Deep Learning categorical model implemented with hybrid descriptors (ECFP6 fingerprint and physicochemical properties: MW, TPSA, $logK_{ow}$, logD)	Reactive (+) (75.8%)	
KeratinoSens™ (OECD 442D) Deep Learning categorical model implemented with hybrid descriptors (ECFP6 fingerprint and physicochemical properties: MW, TPSA, $logK_{ow}$, logD)	Active (+) (88.8%)	
Human Cell Line Activation Test (h-CLAT, OECD 442E) Deep Learning categorical model implemented with hybrid descriptors (ECFP6 fingerprint and physicochemical properties: MW, TPSA, $logK_{ow}$, logD)	Active (+) (68.0%)	

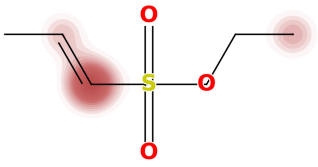
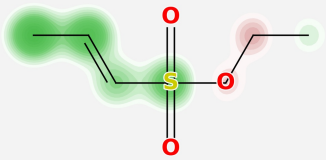
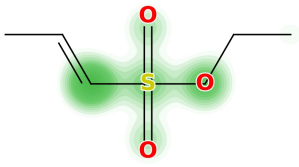
Assay/Event	Predicted class (Confidence)	STR Contribution Mapping
U-SENS™ (OECD 442D) Deep Learning categorical model implemented with hybrid descriptors (ECFP6 fingerprint and physicochemical properties: MW, TPSA, $logK_{ow}$, logD)	Active (+) (50.0%)	

The deep learning model has components with multiple hidden layers that can learn increasingly abstract representations of the chemical hybrid descriptors. The chemical hybrid descriptor is composed by ECFP6 fingerprint and physicochemical properties [molecular weight (MW), the topological polar surface area (TPSA), octanol-water partition coefficient for neutral compounds ($logK_{ow}$) or at different pH states - logD]. Then, every 4 subsequent layers learn more complex representations. Finally, the last layer can evaluate the skin sensitization of the given compound.

STR Contribution Map

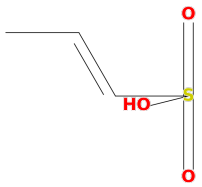
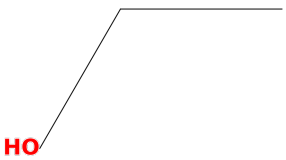
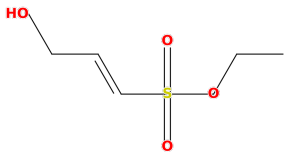
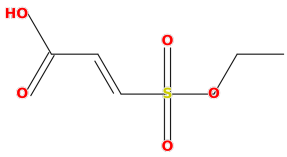


Sensitizer (+) Non-Sensitizer (-)

Predicted Outcome/Assay	Predicted class (Confidence)	STR Contribution Mapping
Local Lymph Node Assay (LLNA, OECD 429) Deep Learning categorical model implemented with hybrid descriptors (ECFP6 fingerprint and physicochemical properties: MW, TPSA, $logK_{ow}$, logD)	Sensitizer (+) (62.5%)	
Human Skin Deep Learning categorical model implemented with hybrid descriptors (ECFP6 fingerprint and physicochemical properties: MW, TPSA, $logK_{ow}$, logD)	Non-Sensitizer (-) (75.7%)	
Combined Model (Human and Animal) Deep Learning categorical model implemented with hybrid descriptors (ECFP6 fingerprint and physicochemical properties: MW, TPSA, $logK_{ow}$, logD)	Non-Sensitizer (-) (92.7%)	

Metabolism prediction and potential for haptenation

To assess both direct and indirect haptens, this module predicts the potential for metabolic activation (pro-hapten formation) by known Phase I reactions, i.e., it can be used to identify potential skin sensitizers which require some type of metabolism to an active metabolite (pro-haptens) before initiation of the key event 1 (KE1) in a skin sensitization AOP (OECD Principle 5).

Metabolite (predicted structure)				
SMILES	CC=CS(=O)(=O)O	CCO	CCOS(=O)(=O)C=CCO	CCOS(=O)(=O)C=CC(=O)O
Reaction Rule	Hydrolysis	Hydrolysis	Aliphatic Hydroxylation	Carboxylation
Metabolite Score	17.4 %	17.4 %	4.9 %	1.6 %
AOP-based Prediction	**Non-Sensitizer (-)**	**Non-Sensitizer (-)**	**Non-Sensitizer (-)**	**Non-Sensitizer (-)**

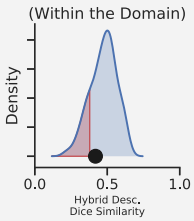
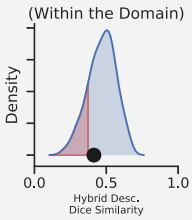
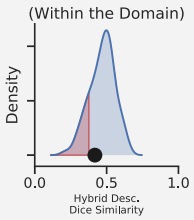
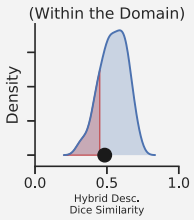
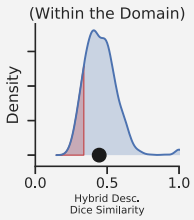
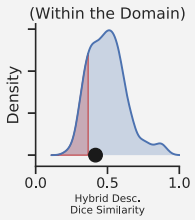
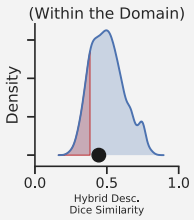
The Expert Knowledge algorithm is based on a broad range of metabolic phase 1 reactions reported in the human Metabolite Database[3]. The reactions found for the parent compound is presented in the column "reaction rule". The "metabolite score" is an empirical probability score assigned to each rule representing the fraction of correctly predicted metabolites in the training set to improve the predictivity. Evaluation of the rule-based predictions demonstrated a significant enrichment of true metabolites in the top list, e.g. the algorithm was able to identify 84 % of the cytochrome P450 specific metabolites and moreover, 66 % of the metabolites were in the top three predicted phase 1 metabolites.

Visual AD Inspection®

The applicability domain (AD) is defined by the chemical structure space and the toxicological response encoded by the developed model, to make new predictions with a given reliability (a defined domain of applicability - OECD Principle 3).[2] Our visual AS Inspection® is used to establish the scope and limitations of the models. Basically, new chemicals must be reasonably similar to training set compounds or a prediction cannot be accepted.

Our visual AD inspection is represented by a density plot of the average fingerprint-dice similarity for the k-nearest neighbors of each compound during the 5-Fold external validation. The chemical structure is represented by a hybrid descriptor composed by ECFP6 fingerprint and physicochemical measurements: molecular weight (MW), topological polar surface area (TPSA), octanol-water partition coefficient for neutral compounds ($logK_{ow}$) or at different pH states (logD). At the visual AD inspection, the black circle represents the evaluated compound, the highlighted red area represents the restricted similarity region, and the blue region is the allowed similarity of the chemical space to predict new compounds.

Even though a well-designed AD helps the user to assess the reliability of predictions made by the model, it should not automatically be assumed that all predictions within the defined AD are necessarily reliable.

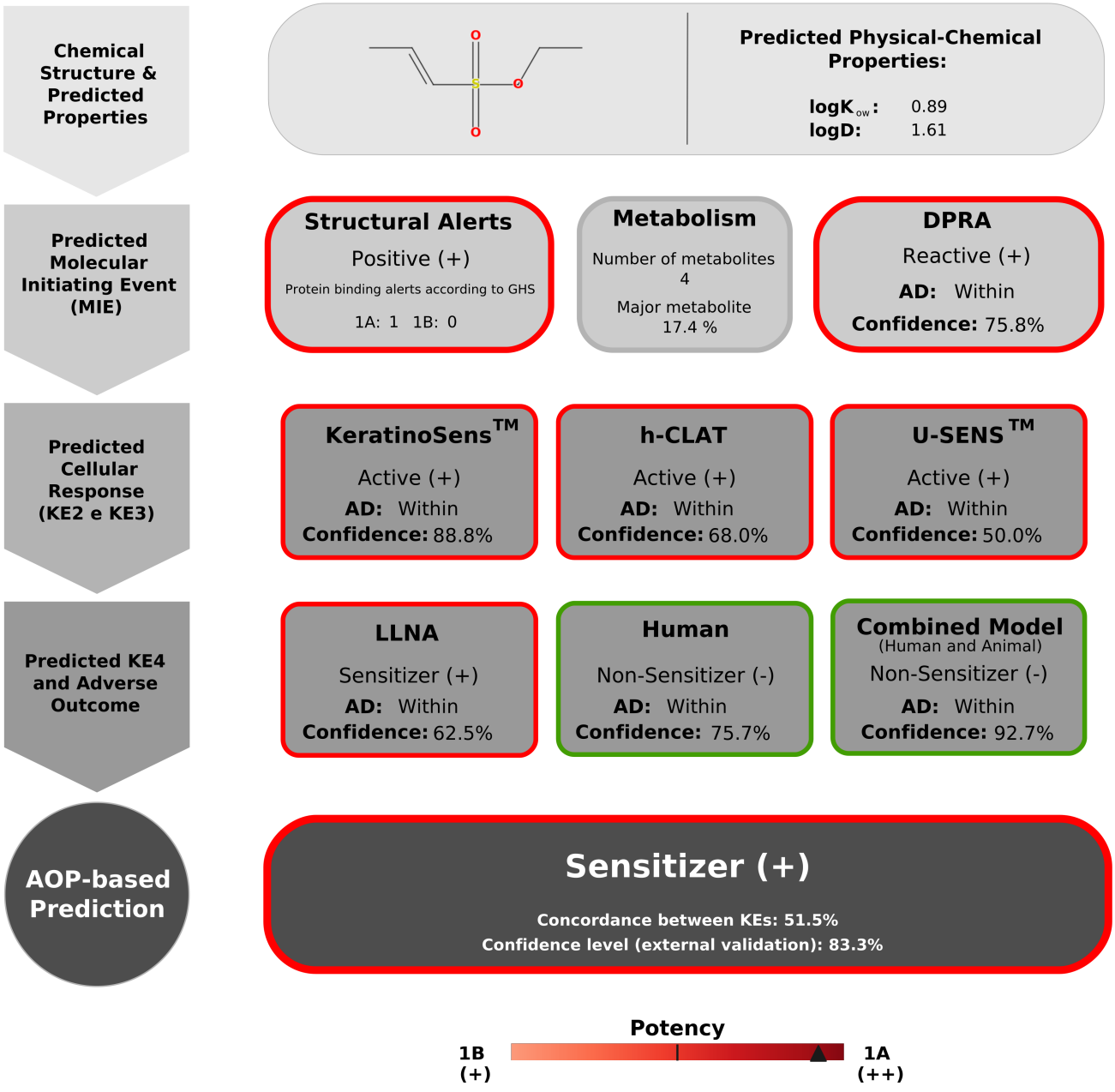
Direct Peptide Reactivity Assay (DPRA, OECD 442D)	Human Cell Line Activation Test (h-CLAT, OECD 442E)	KeratinoSens™ (OECD 442D)	U-SENS™ (OECD 442D)	Local Lymph Node Assay (LLNA, OECD 429)	Human Skin	Combined Model (Human and Animal)
						

Results

The results are presented below; positive (red) and/or negative (green) predictions are presented with applicability domain (AD) and confidence level (internal goodness-of-fit and robustness, OECD principle 4) for each AOP key event and outcome prediction:

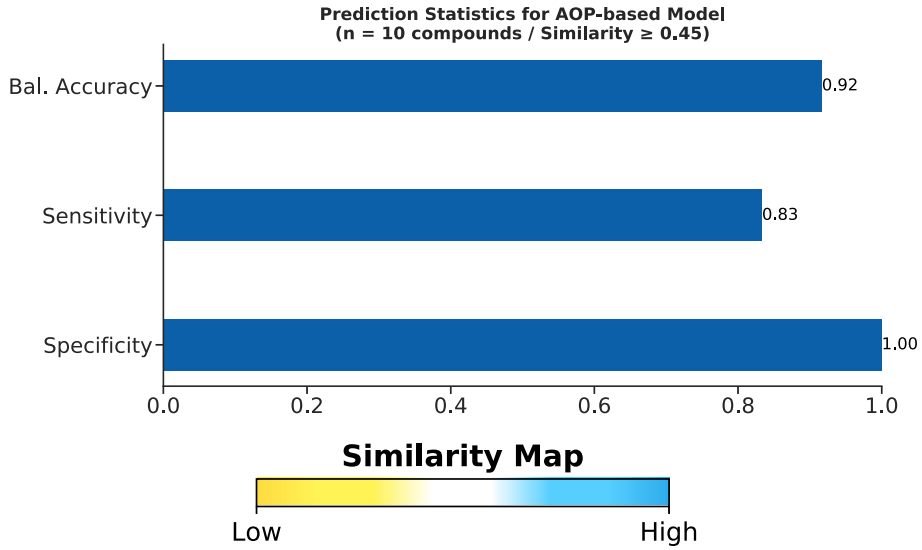
MIE (Molecular Initiating Event):	Reactivity and covalent interaction with cellular proteins (DPRA)
Key event 2 (KE2):	Activation of biochemical pathways by keratinocytes and release of pro-inflammatory mediators (KeratinoSens™)
Key event 3 (KE3):	Events in dendritic cells expression of co-stimulatory and adhesion molecules, by pathways-associated protein and gene expression (h-CLAT). Also, the change in the expression of a cell surface marker associated with the process of activation of monocytes and DC (U-Sens™).
Key event 4 (KE4) and Adverse Outcome:	T-cell proliferation and sensitization

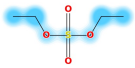
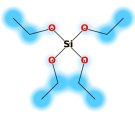
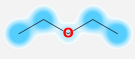
An AOP-based final prediction integrating all individual key events and outcomes is provided with external predictivity (OECD principle 4), balanced by key event relationships, adjustments, and weight-of-evidence assesstment. For each individual model is attributed a weight is assigned based in its confidence level.

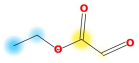
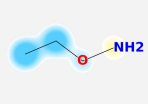
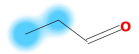
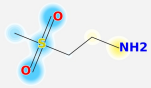
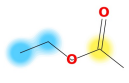


Additional Information

Hybrid descriptor Dice similarity was used to improve the deep learning confidence by interpolating the confidence equalized by the compound similarity criteria obtained from the dataset chemical space. This helps to improve the *in silico* toxicological model to reduce the false positive and negative error.



Molecule (Similarity)	Alerts	DPRA Prediction	KetatinoSens™ Prediction	h-CLAT Prediction	U-SENS™ Prediction	LLNA Prediction	Human Skin Prediction	Combined Model Prediction	Exp. Data/ Assay	AOP-based Prediction (Confidence)	Weight
 (0.62)	Positive (+)	Reactive (+)	Active (+)	Active (+)	Active (+)	Sensitizer (+)	Non-Sensitizer (-)	Non-Sensitizer (-)	LLNA (+), KeratinoSens (+), h-CLAT (+), U-SENS™ (+), DPRA (+)	**Sensitizer (+) (58.0%)**	+ + +
 (0.51)	Negative (-)	Reactive (+)	Active (+)	Active (+)	Active (+)	Sensitizer (+)	Non-Sensitizer (-)	Non-Sensitizer (-)	Not classified (ECHA) (-)	**Non-Sensitizer (-) (51.0%)**	+ + + +
 (0.47)	Negative (-)	Non-Reactive (-)	Active (+)	Active (+)	Active (+)	Non-Sensitizer (-)	Non-Sensitizer (-)	Non-Sensitizer (-)	Not classified (ECHA) (-)	**Non-Sensitizer (-) (72.7%)**	+ + + +
 (0.47)	Negative (-)	Reactive (+)	Active (+)	Active (+)	Active (+)	Non-Sensitizer (-)	Non-Sensitizer (-)	Non-Sensitizer (-)	Not classified (ECHA) (-), LLNA (-)	**Non-Sensitizer (-) (63.0%)**	+ + + +

Molecule (Similarity)	Alerts	DPRA Prediction	KetatinoSens Prediction	h-CLAT Prediction	U-SENS Prediction	LLNA Prediction	Human Skin Prediction	Combined Model Prediction	Exp. Data/ Assay	AOP-based Prediction (Confidence)	Weight
 (0.46)	Positive (+)	Reactive (+)	Active (+)	Active (+)	Active (+)	Sensitizer (+)	Sensitizer (+)	Sensitizer (+)	Category 1A (ECHA) (++)	**Sensitizer (+) (89.4%)**	+ + + +
 (0.46)	Negative (-)	Reactive (+)	Active (+)	Active (+)	Active (+)	Non-Sensitizer (-)	Non-Sensitizer (-)	Sensitizer (+)	Category 1A (ECHA) (++)	**Sensitizer (+) (62.9%)**	+ + + +
 (0.46)	Positive (+)	Reactive (+)	Active (+)	Active (+)	Active (+)	Non-Sensitizer (-)	Non-Sensitizer (-)	Non-Sensitizer (-)	LLNA (-)	**Non-Sensitizer (-) (54.4%)**	+ + +
 (0.45)	Negative (-)	Reactive (+)	Active (+)	Inactive (-)	Inactive (-)	Non-Sensitizer (-)	Sensitizer (+)	Sensitizer (+)	LLNA (-)	**Sensitizer (+) (61.9%)**	+ + +
 (0.45)	Positive (+)	Reactive (+)	Inactive (-)	Active (+)	Active (+)	Non-Sensitizer (-)	Non-Sensitizer (-)	Non-Sensitizer (-)	Not classified (ECHA) (-)	**Non-Sensitizer (-) (65.7%)**	+ + + +
 (0.45)	Positive (+)	Reactive (+)	Active (+)	Active (+)	Active (+)	Sensitizer (+)	Sensitizer (+)	Sensitizer (+)	Category 1B (ECHA) (+), Human Skin (+), LLNA (+), KeratinoSens (+), DPRA (+)	**Sensitizer (+) (90.9%)**	+ + + +

The reliability of the transformation is supported by Dr D. Roberts,

School of Pharmacy and Biomolecular Sciences, Liverpool John Moores University,

Liverpool, England L3 3AF

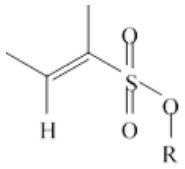
Mechanistic Domain: Michael addition

Mechanistic Alert: Michael type addition on Polarised Alkenes

Structural Alert: Polarised Alkenes – sulfonates

A chemical with this structural alert could interact with proteins via Michael type addition on conjugated alkenyl sulfonates. In TIMES SS (GHS) model such

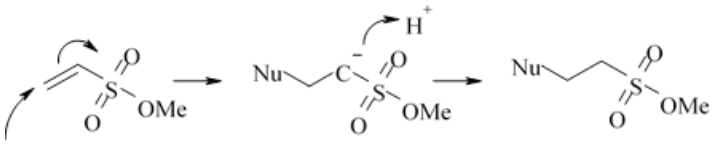
compounds are assessed as 1A. According to Roberts et al., 2015 Michael acceptors of general structure C=C-X, where X is a stronger activating group than carbonyl (e.g. NO, -NO2, SO2R, SOR, SO3R.....) should be treated as highly reactive skin sensitizers.



R = any carbon

Mechanism

The mechanism has been suggested to involve attack by a nucleophile at the β-carbon atom (Aptula et al., 2006, Hermens 1990, Roberts et al., 2007, Schultz et al., 2007, 2005, 2004, van der Ohe et al., 2005, Verhaar et al., 1992).



Nu

Figure 1: Michael addition for polarised alkene - sulfonate (Nu = biological nucleophile e.g.

cysteine or lysine)

References:

1. Roberts, D., Api, A.M., Safford, R., Lalko, J. Regulatory Toxicology and Pharmacology 72 (2015) 683–693.

2. Aptula A.O. et al., (2006) Chemical Research in Toxicology, 19, 1097.

4. Patlewicz G. (2003) QSAR and Combinatorial Science, 22, 196.

5. Roberts D.W. et al., (2007) Chemical Research in Toxicology, 20, 1019.

6. Roberts D.W. et al., (2006) Chemical Research in Toxicology, 19, 1228.

7. Schultz T.W. et al., (2007) Chemical Research in Toxicology, 20, 1359.

8. Schultz T.W. et al., (2005) SAR and QSAR in Environmental Research, 16, 313.

8. Schultz T.W. et al., (2004) SAR and QSAR in Environmental Research, 15, 139.

9. van der Ohe et al., (2005) Chemical Research in Toxicology, 18, 536.

10. Verhaar H.J.M. et al., (1992) Chemosphere, 25, 471.

References

1. The adverse outcome pathway for skin sensitisation initiated by covalent binding to proteins Part 1: Scientific Evidence, Series on Testing and Assessment No. 168 (2012). Online at http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=ENV/JM/MONO(2012)10/PART1&docLanguage=En

2. OECD; Guidance Document On The Validation Of (Quantitative)Structure-Activity (2007). (https://goo.gl/GgSwrj).

4. Piroird, C.; Ovigne, J. M.; Rousset, F.; Martinozzi-Teissier, S.; Gomes, C.; Cotovio, J.; Alépée, N. The Myeloid U937 Skin Sensitization Test (U-SENS) addresses the activation of dendritic cell event in the adverse outcome pathway for skin sensitization. Toxicol. Vitr. 2015, 29, 901–916.

5. Alépée, N.; Piroird, C.; Aujoulat, M.; Dreyfuss, S.; Hoffmann, S.; Hohenstein, A.; Meloni, M.; Nardelli, L.; Gerbeix, C.; Cotovio, J. Prospective multicentre study of the U-SENS test method for skin sensitization testing. Toxicol. Vitr. 2015, 30, 373–382.

6. Hoffmann, S.; Kleinstreuer, N.; Alépée, N.; Allen, D.; Api, A. M.; Ashikaga, T.; Clouet, E.; Cluzel, M.; Desprez, B.; Gellatly, N.; Goebel, C.; Kern, P. S.; Klaric, M.; Kühnl, J.; Lalko, J. F.; Martinozzi-Teissier, S.; Mewes, K.; Miyazawa, M.; Parakhia, R.; van Vliet, E.; Zang, Q.; Petersohn, D. Non-animal methods to predict skin sensitization (I): the Cosmetics Europe database*. Crit. Rev. Toxicol. 2018, 48, 344–358.

7. AU-SENS™ - Myeloid U937 Skin Sensitisation Test, EURL ECVAM - European Union, 2013. Online at https://tsar.jrc.ec.europa.eu/test-method/tm2013-02