VEGA QSAR MODELS

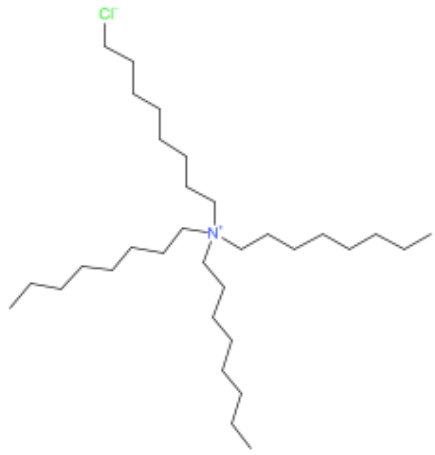
Report

Prediction and Applicability Domain analysis for models:

- Mutagenicity (Ames test) CONSENSUS model 1.0.4
- Mutagenicity (Ames test) model (CAESAR) 2.1.14
- Mutagenicity (Ames test) model (ISS) 1.0.3
- Mutagenicity (Ames test) model (SarPy-IRFMN) 1.0.8
- Mutagenicity (Ames test) model (KNN-Read-Across) 1.0.1
- Mutagenicity (Ames test) model for aromatic amines (CONCERT/IRFMN) 1.0.0
- Carcinogenicity model (CAESAR) 2.1.10
- Carcinogenicity model (ISS) 1.0.3
- Carcinogenicity model (IRFMN-ISSCAN-CGX) 1.0.2
- Carcinogenicity model (IRFMN-Antares) 1.0.2
- Carcinogenicity oral classification model (IRFMN) 1.0.1
- Carcinogenicity oral Slope Factor model (IRFMN) 1.0.1
- Carcinogenicity inhalation classification model (IRFMN) 1.0.1
- Carcinogenicity inhalation Slope Factor model (IRFMN) 1.0.1
- Carcinogenicity in male rat (CORAL) 1.0.0
- Carcinogenicity in female Rat (CORAL) 1.0.0
- Acute Toxicity (LD50) model (KNN) 1.0.0
- BCF model (CAESAR) 2.1.15
- BCF model (Meylan) 1.0.4
- BCF model (Arnot-Gobas) 1.0.1
- BCF model (KNN-Read-Across) 1.1.1

Core version: 1.3.18

Prediction for compound Molecule 0 -



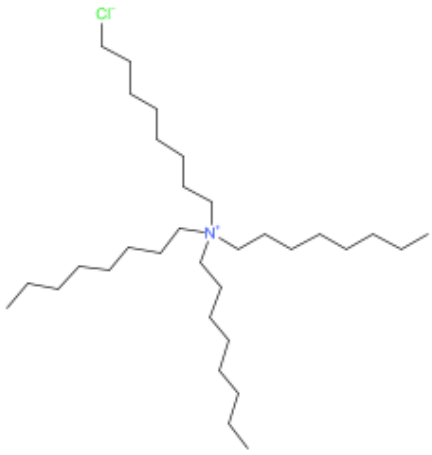
Prediction:

Prediction is Mutagenic with a consensus score of 0.25, based on 4 models.

Compound: Molecule 0
Compound SMILES: CCCCCCCC[N+](CCCCCCCC)(CCCCCCCC)CCCCCCCC[Cl-]
Used models: 4
Predicted Consensus Mutagen activity: Mutagenic
Mutagenic Score: 0.25
Non-Mutagenic Score: 0.15
Model Caesar assessment: Suspect Mutagenic (LOW reliability)
Model ISS assessment: Mutagenic (MODERATE reliability)
Model SarPy assessment: Mutagenic (LOW reliability)
Model KNN assessment: NON-Mutagenic (MODERATE reliability)
Remarks:
none

1. Prediction Summary

Prediction for compound Molecule 0 -



Prediction: Reliability:

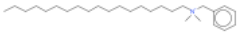
Prediction is Suspect Mutagenic, but the result may be not reliable. A check of the information given in the following section should be done, paying particular attention to the following issues:
- Only moderately similar compounds with known experimental value in the training set have been found
- similar molecules found in the training set have experimental values that disagree with the predicted value
The following relevant fragments have been found: SA8 Aliphatic halogens

Compound: Molecule 0
Compound SMILES: CCCCCCCC[N+](CCCCCCCC)(CCCCCCCC)CCCCCCCC[Cl-]
Experimental value: -
Predicted Mutagen activity: Suspect Mutagenic
Structural Alerts: SA8 Aliphatic halogens
Reliability: The predicted compound is outside the Applicability Domain of the model
Remarks:
none

3.1 Applicability Domain:
Similar Compounds, with Predicted and Experimental Values

Compound #1

CAS: 110-30-5
Dataset id:2884 (Training Set)
SMILES: O=C(NCCNC(=O)CCCCCCCCCCCCCCCCC)CCCCCCCCCCCCCCCCC
Similarity: 0.771
Experimental value : NON-Mutagenic
Predicted value : NON-Mutagenic

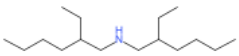


Compound #2

CAS: 37612-69-4
Dataset id:3982 (Training Set)
SMILES: c1ccc(cc1)C[N+](C)(C)CCCCCCCCCCCCCCCCCC
Similarity: 0.764
Experimental value : NON-Mutagenic
Predicted value : NON-Mutagenic

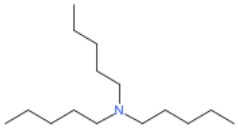
Compound #3

CAS: 124-30-1
Dataset id:2222 (Training Set)
SMILES: NCCCCCCCCCCCCCCCCCC
Similarity: 0.735
Experimental value : NON-Mutagenic
Predicted value : NON-Mutagenic



Compound #4

CAS: 106-20-7
Dataset id:357 (Training Set)
SMILES: N(CC(CC)CCCC)CC(CC)CCCC
Similarity: 0.731
Experimental value : NON-Mutagenic
Predicted value : NON-Mutagenic



Compound #5

CAS: 621-77-2
Dataset id:349 (Training Set)
SMILES: N(CCCCC)(CCCCC)CCCCC
Similarity: 0.716
Experimental value : NON-Mutagenic
Predicted value : NON-Mutagenic

Compound #6

CAS: 143-27-1
Dataset id:2208 (Training Set)
SMILES: NCCCCCCCCCCCCCCCCC
Similarity: 0.713
Experimental value : NON-Mutagenic
Predicted value : NON-Mutagenic

3.2 Applicability Domain:
Measured Applicability Domain Scores

✖	Global AD Index AD index = 0 Explanation: The predicted compound is outside the Applicability Domain of the model.
⚠	Similar molecules with known experimental value Similarity index = 0.756 Explanation: Only moderately similar compounds with known experimental value in the training set have been found..
✔	Accuracy of prediction for similar molecules Accuracy index = 1 Explanation: Accuracy of prediction for similar molecules found in the training set is good..
✖	Concordance for similar molecules Concordance index = 0 Explanation: similar molecules found in the training set have experimental values that disagree with the predicted value..
✔	Model's descriptors range check Descriptors range check = True Explanation: descriptors for this compound have values inside the descriptor range of the compounds of the training set..
✔	Atom Centered Fragments similarity check ACF index = 1 Explanation: all atom centered fragment of the compound have been found in the compounds of the training set..

Symbols explanation:

✔ The feature has a good assessment, model is reliable regarding this aspect.

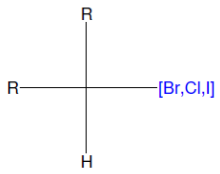
⚠ The feature has a non optimal assessment, this aspect should be reviewed by an expert.

✖ The feature has a bad assessment, model is not reliable regarding this aspect.

4.1 Reasoning:
Relevant Chemical Fragments and Moieties

(Molecule 0) Reasoning on fragments/structural alerts :.

Fragment found: SA8 Aliphatic halogens



R = any atom/group

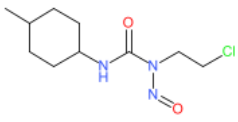
This alert contains non tertiary aliphatic halogens. Substances fired by Alerts SA2, SA4, SA5 and SA20 should be also excluded.

Following, the most similar compounds from the model's dataset having the same fragment.

CAS: 112-52-7
Dataset id:3322 (Test Set)
SMILES: CCCCCCCCCCCCCl
Similarity: 0.691

Experimental value : NON-Mutagenic
Predicted value : Suspect Mutagenic

Alerts (found also in the target): SA8 Aliphatic halogens

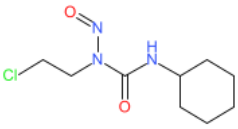


CAS: 13909-09-6
Dataset id:2393 (Test Set)
SMILES: O=NN(C(=O)NC1CCC(C)CC1)CCCl
Similarity: 0.564

Experimental value : Mutagenic
Predicted value : Mutagenic

Alerts (found also in the target): SA8 Aliphatic halogens

Alerts (not found also in the target): SA21 Alkyl and aryl N-nitroso groups



CAS: 13010-47-4
Dataset id:2238 (Training Set)
SMILES: O=NN(C(=O)NC1CCCCC1)CCCl
Similarity: 0.553

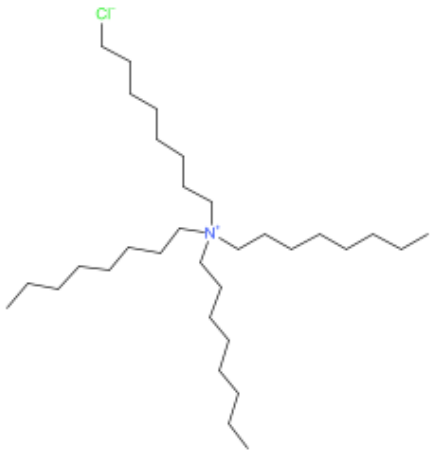
Experimental value : Mutagenic
Predicted value : Mutagenic

Alerts (found also in the target): SA8 Aliphatic halogens

Alerts (not found also in the target): SA21 Alkyl and aryl N-nitroso groups

1. Prediction Summary

Prediction for compound Molecule 0 -

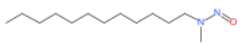


Prediction: Reliability:

Prediction is Mutagenic, but the result shows some critical aspects, which require to be checked:
- Only moderately similar compounds with known experimental value in the training set have been found
- some atom centered fragments of the compound have not been found in the compounds of the training set or are rare fragments (1 infrequent_fragments found)
The following alerts have been found: SA8 Aliphatic halogens

Compound: Molecule 0
Compound SMILES: CCCCCCCC[N+](CCCCCCCC)(CCCCCCCC)CCCCCCCC[Cl-]
Experimental value: -
Predicted Mutagen activity: Mutagenic
Structural Alerts: SA8 Aliphatic halogens
Reliability: The predicted compound could be out of the Applicability Domain of the model
Remarks:
none

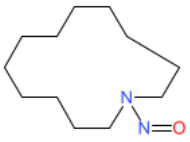
3.1 Applicability Domain:
Similar Compounds, with Predicted and Experimental Values



Compound #1

CAS: 55090-44-3
Dataset id:547 (Training Set)
SMILES: O=NN(C)CCCCCCCCCCCC
Similarity: 0.66
Experimental value : Mutagenic
Predicted value : Mutagenic

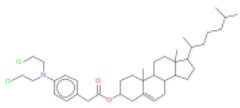
Alerts (not found also in the target): SA21 Alkyl and aryl N-nitroso groups



Compound #2

CAS: 40580-89-0
Dataset id:553 (Training Set)
SMILES: O=NN1CCCCCCCCCCC1
Similarity: 0.643
Experimental value : Mutagenic
Predicted value : Mutagenic

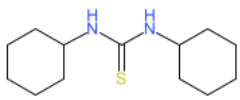
Alerts (not found also in the target): SA21 Alkyl and aryl N-nitroso groups



Compound #3

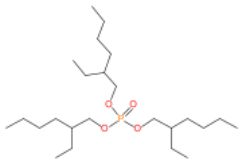
CAS: 3546-10-9
Dataset id:216 (Training Set)
SMILES:
O=C(OC4CC3=CCC1C(CCC2(C)(C(CCC12)C(C)CCCC(C)C))C3(C)CC4)Cc5ccc(cc5)N(CCCl)CCCl
Similarity: 0.639
Experimental value : NON-Mutagenic
Predicted value : Mutagenic

Alerts (not found also in the target): SA5 S or N mustard



Compound #4

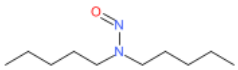
CAS: 1212-29-9
Dataset id:108 (Training Set)
SMILES: N(C(NC1CCCCC1)=S)C2CCCCC2
Similarity: 0.626
Experimental value : NON-Mutagenic
Predicted value : NON-Mutagenic



Compound #5

CAS: 78-42-2
Dataset id:69 (Training Set)
SMILES: O=P(OCC(CC)CCCC)(OCC(CC)CCCC)OCC(CC)CCCC
Similarity: 0.617
Experimental value : NON-Mutagenic
Predicted value : NON-Mutagenic

3.1 Applicability Domain:
Similar Compounds, with Predicted and Experimental Values



Compound #6

CAS: 13256-06-9
Dataset id:886 (Training Set)
SMILES: O=NN(CCCCC)CCCCC
Similarity: 0.615
Experimental value : Mutagenic
Predicted value : Mutagenic

Alerts (not found also in the target): SA21 Alkyl and aryl N-nitroso groups

3.2 Applicability Domain:
Measured Applicability Domain Scores

⚠	Global AD Index AD index = 0.686 Explanation: The predicted compound could be out of the Applicability Domain of the model.
⚠	Similar molecules with known experimental value Similarity index = 0.651 Explanation: Only moderately similar compounds with known experimental value in the training set have been found..
✓	Accuracy of prediction for similar molecules Accuracy index = 1 Explanation: Accuracy of prediction for similar molecules found in the training set is good..
✓	Concordance for similar molecules Concordance index = 1 Explanation: Similar molecules found in the training set have experimental values that agree with the predicted value..
⚠	Atom Centered Fragments similarity check ACF index = 0.85 Explanation: some atom centered fragments of the compound have not been found in the compounds of the training set or are rare fragments (1 infrequent_fragments found)..

Symbols explanation:

✓ The feature has a good assessment, model is reliable regarding this aspect.

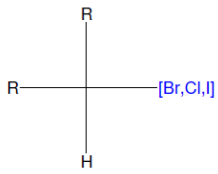
⚠ The feature has a non optimal assessment, this aspect should be reviewed by an expert.

✗ The feature has a bad assessment, model is not reliable regarding this aspect.

4.1 Reasoning:
Relevant Chemical Fragments and Moieties

(Molecule 0) Reasoning on fragments/structural alerts :.

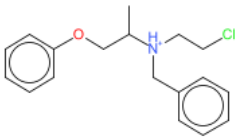
Fragment found: SA8 Aliphatic halogens



R = any atom/group

This alert contains non tertiary aliphatic halogens. Substances fired by Alerts SA2, SA4, SA5 and SA20 should be also excluded.

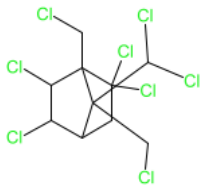
Following, the most similar compounds from the model's dataset having the same fragment.



CAS: 63-92-3
Dataset id:583 (Training Set)
SMILES: O(c1ccccc1)CC(C)[NH+](Cc2ccccc2)CCCl
Similarity: 0.53

Experimental value : Mutagenic
Predicted value : Mutagenic

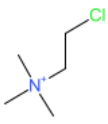
Alerts (found also in the target): SA8 Aliphatic halogens



CAS: 8001-35-2
Dataset id:727 (Training Set)
SMILES: C2C1C(C(C(CCl)(C1(CCl)C(Cl)Cl)C2(Cl)Cl)Cl)Cl
Similarity: 0.527

Experimental value : Mutagenic
Predicted value : Mutagenic

Alerts (found also in the target): SA8 Aliphatic halogens



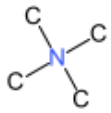
CAS: 999-81-5
Dataset id:100 (Training Set)
SMILES: C[N+](C)(C)CCCl
Similarity: 0.523

Experimental value : NON-Mutagenic
Predicted value : Mutagenic

Alerts (found also in the target): SA8 Aliphatic halogens

4.1 Reasoning:
Relevant Chemical Fragments and Moieties

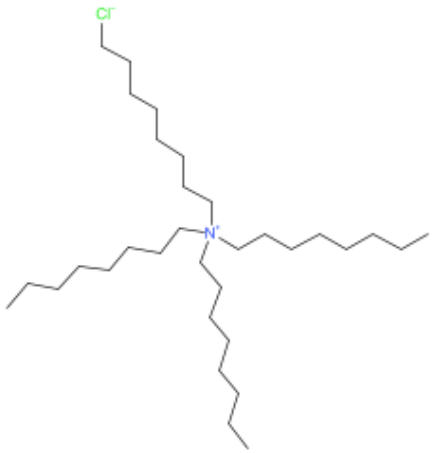
(Molecule 0) Reasoning on rare and missing Atom Centered Fragments .
The following Atom Centered Fragments have been found in the molecule, but they are not found or rarely found in the model's training set:



Fragment defined by the SMILES: CN(C)(C)C
The fragment has less than 3 occurrences in the model's training set

1. Prediction Summary

Prediction for compound Molecule 0 -

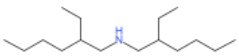


Prediction: Reliability:

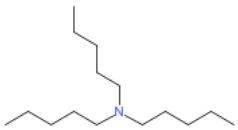
Prediction is Mutagenic, but the result may be not reliable. A check of the information given in the following section should be done, paying particular attention to the following issues:
- Only moderately similar compounds with known experimental value in the training set have been found
- similar molecules found in the training set have experimental values that disagree with the predicted value
The following relevant fragments have been found: SM106; SM142; SM163; SM175

Compound: Molecule 0
Compound SMILES: CCCCCCCC[N+](CCCCCCCC)(CCCCCCCC)CCCCCCCC[Cl-]
Experimental value: -
Predicted Mutagen activity: Mutagenic
No. alerts for mutagenicity: 1
No. alerts for non-mutagenicity: 3
Structural Alerts: SM106; SM142; SM163; SM175
Reliability: The predicted compound is outside the Applicability Domain of the model
Remarks:
none

3.1 Applicability Domain:
Similar Compounds, with Predicted and Experimental Values

	Compound #1 CAS: 110-30-5 Dataset id:2884 (Training Set) SMILES: O=C(NCCNC(=O)CCCCCCCCCCCCCCCCC)CCCCCCCCCCCCCCCCC Similarity: 0.771 Experimental value : NON-Mutagenic Predicted value : NON-Mutagenic Alerts (found also in the target): SM163 Alerts (not found also in the target): SM157; SM177
	Compound #2 CAS: 37612-69-4 Dataset id:3982 (Training Set) SMILES: c1ccc(cc1)C[N+](C)(C)CCCCCCCCCCCCCCCCCC Similarity: 0.764 Experimental value : NON-Mutagenic Predicted value : NON-Mutagenic Alerts (found also in the target): SM163; SM175 Alerts (not found also in the target): SM157
	Compound #3 CAS: 124-30-1 Dataset id:2222 (Training Set) SMILES: NCCCCCCCCCCCCCCCCCC Similarity: 0.735 Experimental value : NON-Mutagenic Predicted value : NON-Mutagenic Alerts (found also in the target): SM163 Alerts (not found also in the target): SM157
	Compound #4 CAS: 106-20-7 Dataset id:357 (Training Set) SMILES: N(CC(CC)CCCC)CC(CC)CCCC Similarity: 0.731 Experimental value : NON-Mutagenic Predicted value : NON-Mutagenic Alerts (found also in the target): SM163 Alerts (not found also in the target): SM169; SM182

3.1 Applicability Domain:
Similar Compounds, with Predicted and Experimental Values



Compound #5

CAS: 621-77-2
Dataset id:349 (Training Set)
SMILES: N(CCCCC)(CCCCC)CCCCC
Similarity: 0.716
Experimental value : NON-Mutagenic
Predicted value : NON-Mutagenic

Alerts (found also in the target): SM142

Compound #6

CAS: 143-27-1
Dataset id:2208 (Training Set)
SMILES: NCCCCCCCCCCCCCCCC
Similarity: 0.713
Experimental value : NON-Mutagenic
Predicted value : NON-Mutagenic

Alerts (found also in the target): SM163

Alerts (not found also in the target): SM157

3.2 Applicability Domain:
Measured Applicability Domain Scores

✖	Global AD Index AD index = 0 Explanation: The predicted compound is outside the Applicability Domain of the model.
⚠	Similar molecules with known experimental value Similarity index = 0.756 Explanation: Only moderately similar compounds with known experimental value in the training set have been found..
✔	Accuracy of prediction for similar molecules Accuracy index = 1 Explanation: Accuracy of prediction for similar molecules found in the training set is good..
✖	Concordance for similar molecules Concordance index = 0 Explanation: similar molecules found in the training set have experimental values that disagree with the predicted value..
✔	Atom Centered Fragments similarity check ACF index = 1 Explanation: all atom centered fragment of the compound have been found in the compounds of the training set..

Symbols explanation:

✔ The feature has a good assessment, model is reliable regarding this aspect.

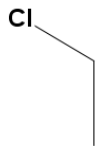
⚠ The feature has a non optimal assessment, this aspect should be reviewed by an expert.

✖ The feature has a bad assessment, model is not reliable regarding this aspect.

4.1 Reasoning:
Relevant Chemical Fragments and Moieties

(Molecule 0) Reasoning on fragments/structural alerts - 1 of 4:.

Fragment found: SM106



Sarpy alert n. 106 for Mutagenicity, defined by SMARTS: CCCl

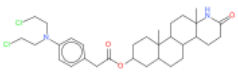
Following, the most similar compounds from the model's dataset having the same fragment.

CAS: 112-52-7
Dataset id:3322 (Test Set)
SMILES: CCCCCCCCCCCCCl
Similarity: 0.691

Experimental value : NON-Mutagenic
Predicted value : Mutagenic

Alerts (found also in the target): SM106; SM163

Alerts (not found also in the target): SM157

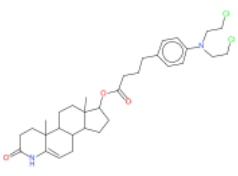


CAS: 43000-65-3
Dataset id:4260 (Test Set)
SMILES:
O=C(OC4CCC2(C)(C(CCC1C3CCC(=O)NC3(C)(CCC12))C4))Cc5ccc(cc5)N(CCCl)CCCl
Similarity: 0.655

Experimental value : Mutagenic
Predicted value : Mutagenic

Alerts (found also in the target): SM106; SM163

Alerts (not found also in the target): SM45; SM73; SM104; SM157; SM162; SM169; SM177; SM182; SM188; SM195



CAS: 107534-94-1
Dataset id:1563 (Test Set)
SMILES:
O=C(OC2CCC3C4CC=C1NC(=O)CCC1(C)C4(CCC23(C)))CCCc5ccc(cc5)N(CCCl)CCCl
Similarity: 0.641

Experimental value : Mutagenic
Predicted value : Mutagenic

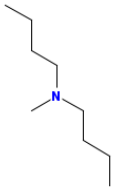
Alerts (found also in the target): SM106; SM163

Alerts (not found also in the target): SM45; SM73; SM104; SM123; SM157; SM162; SM169; SM177; SM178; SM182; SM188

4.1 Reasoning:
Relevant Chemical Fragments and Moieties

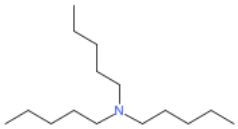
(Molecule 0) Reasoning on fragments/structural alerts - 2 of 4:.

Fragment found: SM142



Sarpy alert n. 142 for NON-Mutagenicity, defined by SMARTS: N(C)(CCCC)CCCC

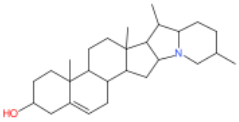
Following, the most similar compounds from the model's dataset having the same fragment.



CAS: 621-77-2
Dataset id:349 (Training Set)
SMILES: N(CCCCC)(CCCCC)CCCCC
Similarity: 0.716

Experimental value : NON-Mutagenic
Predicted value : NON-Mutagenic

Alerts (found also in the target): SM142

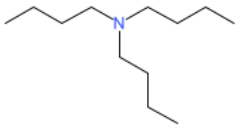


CAS: 80-78-4
Dataset id:4145 (Test Set)
SMILES: OC6CC5=CCC4C(CCC3(C)(C4(CC2N1CC(C)CCC1C(C)C23)))C5(C)CC6
Similarity: 0.707

Experimental value : NON-Mutagenic
Predicted value : NON-Mutagenic

Alerts (found also in the target): SM142; SM163

Alerts (not found also in the target): SM157; SM162; SM169; SM182



CAS: 102-82-9
Dataset id:2339 (Training Set)
SMILES: N(CCCC)(CCCC)CCCC
Similarity: 0.652

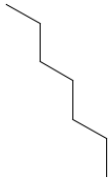
Experimental value : NON-Mutagenic
Predicted value : NON-Mutagenic

Alerts (found also in the target): SM142

4.1 Reasoning:
Relevant Chemical Fragments and Moieties

(Molecule 0) Reasoning on fragments/structural alerts - 3 of 4:.

Fragment found: SM163



Sarpy alert n. 163 for NON-Mutagenicity, defined by SMARTS: CCCCCCC

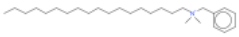
Following, the most similar compounds from the model's dataset having the same fragment.

CAS: 110-30-5
Dataset id:2884 (Training Set)
SMILES: O=C(NCCNC(=O)CCCCCCCCCCCCCCCCC)CCCCCCCCCCCCCCCCC
Similarity: 0.771

Experimental value : NON-Mutagenic
Predicted value : NON-Mutagenic

Alerts (found also in the target): SM163

Alerts (not found also in the target): SM157; SM177



CAS: 37612-69-4
Dataset id:3982 (Training Set)
SMILES: c1ccc(cc1)C[N+](C)(C)CCCCCCCCCCCCCCCCCC
Similarity: 0.764

Experimental value : NON-Mutagenic
Predicted value : NON-Mutagenic

Alerts (found also in the target): SM163; SM175

Alerts (not found also in the target): SM157

CAS: 124-30-1
Dataset id:2222 (Training Set)
SMILES: NCCCCCCCCCCCCCCCCCC
Similarity: 0.735

Experimental value : NON-Mutagenic
Predicted value : NON-Mutagenic

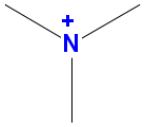
Alerts (found also in the target): SM163

Alerts (not found also in the target): SM157

4.1 Reasoning:
Relevant Chemical Fragments and Moieties

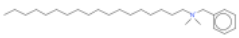
(Molecule 0) Reasoning on fragments/structural alerts - 4 of 4:.

Fragment found: SM175



Sarpy alert n. 175 for NON-Mutagenicity, defined by SMARTS: [N+](C)(C)C

Following, the most similar compounds from the model's dataset having the same fragment.

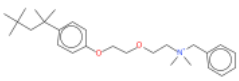


CAS: 37612-69-4
Dataset id:3982 (Training Set)
SMILES: c1ccc(cc1)C[N+](C)(C)CCCCCCCCCCCCCCCCCC
Similarity: 0.764

Experimental value : NON-Mutagenic
Predicted value : NON-Mutagenic

Alerts (found also in the target): SM163; SM175

Alerts (not found also in the target): SM157

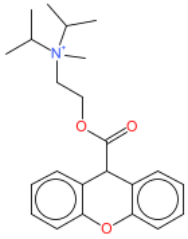


CAS: 121-54-0
Dataset id:3981 (Training Set)
SMILES: O(c1ccc(cc1)C(C)(C)CC(C)(C)C)CCOCC[N+](C)(C)Cc2ccccc2
Similarity: 0.561

Experimental value : NON-Mutagenic
Predicted value : NON-Mutagenic

Alerts (found also in the target): SM175

Alerts (not found also in the target): SM176; SM182



CAS: 298-50-0
Dataset id:787 (Training Set)
SMILES: O=C(OCC[N+](C)(C(C)C)C(C)C)C2c3ccccc3(Oc1ccccc12)
Similarity: 0.535

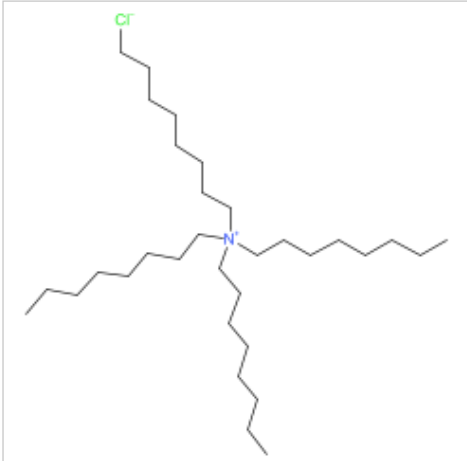
Experimental value : NON-Mutagenic
Predicted value : NON-Mutagenic

Alerts (found also in the target): SM175

Alerts (not found also in the target): SM176; SM195

1. Prediction Summary

Prediction for compound Molecule 0 -



Prediction: Reliability:

Prediction is NON-Mutagenic, but the result shows some critical aspects, which require to be checked:
- Only moderately similar compounds with known experimental value in the training set have been found

Compound: Molecule 0
Compound SMILES: CCCCCCCC[N+](CCCCCCCC)(CCCCCCCC)CCCCCCCC[Cl-]
Experimental value: -
Predicted Mutagen activity: NON-Mutagenic
Molecules used for prediction: 4
Reliability: The predicted compound could be out of the Applicability Domain of the model
Remarks:
none

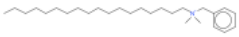
3.1 Applicability Domain:
Similar Compounds, with Predicted and Experimental Values

Compound #1

CAS: 10094-45-8
Dataset id:50 (Training Set)
SMILES: O=C(NCCCCCCCCCCCCCCCCCC)CCCCCCCCCCCC=CCCCCCCCC
Similarity: 0.804
Experimental value : NON-Mutagenic
Predicted value : NON-Mutagenic

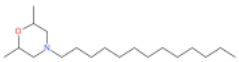
Compound #2

CAS: 110-30-5
Dataset id:486 (Training Set)
SMILES: O=C(NCCNC(=O)CCCCCCCCCCCCCCCCCC)CCCCCCCCCCCCCCCCCC
Similarity: 0.771
Experimental value : NON-Mutagenic
Predicted value : NON-Mutagenic



Compound #3

CAS: 122-19-0
Dataset id:903 (Training Set)
SMILES: c1ccc(cc1)C[N+](C)(C)CCCCCCCCCCCCCCCCCC
Similarity: 0.764
Experimental value : NON-Mutagenic
Predicted value : NON-Mutagenic

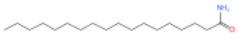


Compound #4

CAS: 24602-86-6
Dataset id:2380 (Training Set)
SMILES: O1C(C)CN(CCCCCCCCCCCCC)CC1C
Similarity: 0.755
Experimental value : NON-Mutagenic
Predicted value : NON-Mutagenic

Compound #5

CAS: 124-30-1
Dataset id:963 (Training Set)
SMILES: NCCCCCCCCCCCCCCCCCC
Similarity: 0.735
Experimental value : NON-Mutagenic
Predicted value : NON-Mutagenic



Compound #6

CAS: 124-26-5
Dataset id:962 (Training Set)
SMILES: O=C(N)CCCCCCCCCCCCCCCCCC
Similarity: 0.722
Experimental value : NON-Mutagenic
Predicted value : NON-Mutagenic

3.2 Applicability Domain:
Measured Applicability Domain Scores

⚠	Global AD Index AD index = 0.878 Explanation: The predicted compound could be out of the Applicability Domain of the model.
⚠	Similar molecules with known experimental value Similarity index = 0.772 Explanation: Only moderately similar compounds with known experimental value in the training set have been found..
✓	Accuracy of prediction for similar molecules Accuracy index = 1 Explanation: Accuracy of prediction for similar molecules found in the training set is good..
✓	Concordance for similar molecules Concordance index = 1 Explanation: Similar molecules found in the training set have experimental values that agree with the predicted value..
✓	Atom Centered Fragments similarity check ACF index = 1 Explanation: all atom centered fragment of the compound have been found in the compounds of the training set..

Symbols explanation:

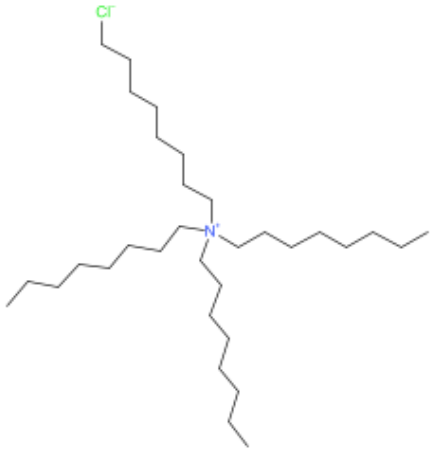
✓ The feature has a good assessment, model is reliable regarding this aspect.

⚠ The feature has a non optimal assessment, this aspect should be reviewed by an expert.

✗ The feature has a bad assessment, model is not reliable regarding this aspect.

1. Prediction Summary

Prediction for compound Molecule 0 -

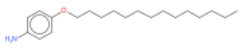


Prediction: Reliability:

Prediction is NON-Mutagenic, but the result may be not reliable. A check of the information given in the following section should be done, paying particular attention to the following issues:
- Only moderately similar compounds with known experimental value in the training set have been found
- Accuracy of prediction for similar molecules found in the training set is not adequate
- a prominent number of atom centered fragments of the compound have not been found in the compounds of the training set or are rare fragments (1 unknown fragments and 1 infrequent_fragments found)

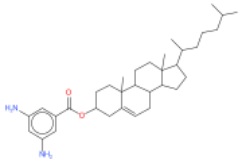
Compound: Molecule 0
Compound SMILES: CCCCCCCC[N+](CCCCCCCC)(CCCCCCCC)CCCCCCCC[Cl-]
Experimental value: -
Predicted Mutagen activity: NON-Mutagenic
Structural Alerts: -
Reliability: The predicted compound is outside the Applicability Domain of the model
Remarks:
none

3.1 Applicability Domain:
Similar Compounds, with Predicted and Experimental Values



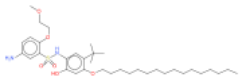
Compound #1

CAS: N.A.
Dataset id:6537 (Training Set)
SMILES: O(c1ccc(N)cc1)CCCCCCCCCCCCCC
Similarity: 0.617
Experimental value : NON-Mutagenic
Predicted value : Mutagenic



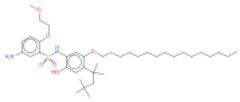
Compound #2

CAS: N.A.
Dataset id:7188 (Training Set)
SMILES:
O=C(OC4CC3=CCC1C(CCC2(C)(C(CCC12)C(C)CCCC(C)C))C3(C)CC4)c5cc(N)cc(N)c5
Similarity: 0.605
Experimental value : NON-Mutagenic
Predicted value : NA



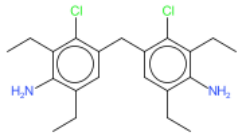
Compound #3

CAS: N.A.
Dataset id:4037 (Training Set)
SMILES:
O=S(=O)(Nc1cc(c(OCCCCCCCCCCCCCCCC)cc1(O))C(C)(C)C)c2cc(N)ccc2(OCCOC)
Similarity: 0.573
Experimental value : NON-Mutagenic
Predicted value : NON-Mutagenic



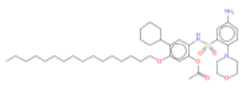
Compound #4

CAS: N.A.
Dataset id:6841 (Training Set)
SMILES:
O=S(=O)(Nc1cc(OCCCCCCCCCCCCCCCC)c(cc1(O))C(C)(C)CC(C)(C)C)c2cc(N)ccc2(OCCOC)
Similarity: 0.567
Experimental value : NON-Mutagenic
Predicted value : NON-Mutagenic



Compound #5

CAS: N.A.
Dataset id:6633 (Training Set)
SMILES: Nc1c(cc(c(c1CC)Cl)Cc2cc(c(N)c(c2Cl)CC)CC)CC
Similarity: 0.565
Experimental value : NON-Mutagenic
Predicted value : Mutagenic



Compound #6

CAS: N.A.
Dataset id:4359 (Training Set)
SMILES:
O=C(Oc3cc(OCCCCCCCCCCCCCCCC)c(cc3(NS(=O)(=O)c1cc(N)ccc1N2CCOCC2))C4CCCCC4)C
Similarity: 0.563
Experimental value : NON-Mutagenic
Predicted value : NON-Mutagenic

3.2 Applicability Domain:
Measured Applicability Domain Scores

✖	Global AD Index AD index = 0 Explanation: The predicted compound is outside the Applicability Domain of the model.
⚠	Similar molecules with known experimental value Similarity index = 0.611 Explanation: Only moderately similar compounds with known experimental value in the training set have been found..
✖	Accuracy of prediction for similar molecules Accuracy index = 0 Explanation: Accuracy of prediction for similar molecules found in the training set is not adequate..
✔	Concordance for similar molecules Concordance index = 1 Explanation: Similar molecules found in the training set have experimental values that agree with the predicted value..
✖	Atom Centered Fragments similarity check ACF index = 0.51 Explanation: a prominent number of atom centered fragments of the compound have not been found in the compounds of the training set or are rare fragments (1 unknown fragments and 1 infrequent_fragments found)..

Symbols explanation:

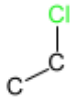
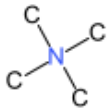
✔ The feature has a good assessment, model is reliable regarding this aspect.

⚠ The feature has a non optimal assessment, this aspect should be reviewed by an expert.

✖ The feature has a bad assessment, model is not reliable regarding this aspect.

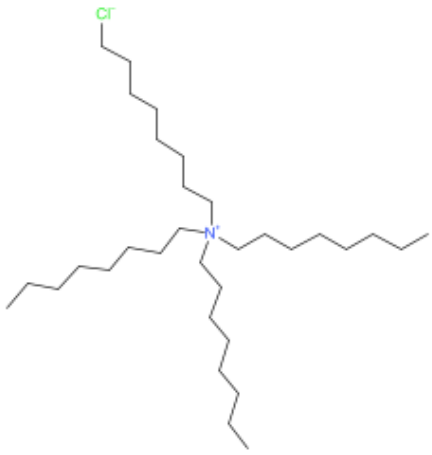
4.1 Reasoning:
Relevant Chemical Fragments and Moieties

(Molecule 0) Reasoning on rare and missing Atom Centered Fragments .
The following Atom Centered Fragments have been found in the molecule, but they are not found or rarely found in the model's training set:

	Fragment defined by the SMILES: CCCl The fragment has less than 3 occurrences in the model's training set
	Fragment defined by the SMILES: CN(C)(C)C The fragment has never been found in the model's training set

1. Prediction Summary

Prediction for compound Molecule 0 -

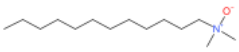


Prediction: Reliability:

Prediction is NON-Carcinogen, but the result may be not reliable. A check of the information given in the following section should be done, paying particular attention to the following issues:
- Only moderately similar compounds with known experimental value in the training set have been found
- Accuracy of prediction for similar molecules found in the training set is not adequate
- some similar molecules found in the training set have experimental values that disagree with the predicted value
- a prominent number of atom centered fragments of the compound have not been found in the compounds of the training set or are rare fragments (1 unknown fragments found)

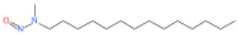
Compound: Molecule 0
Compound SMILES: CCCCCCCC[N+](CCCCCCCC)(CCCCCCCC)CCCCCCCC[Cl-]
Experimental value: -
Predicted Carcinogen activity: NON-Carcinogen
P(Carcinogen): 0.13
P(NON-Carcinogen): 0.87
Reliability: The predicted compound is outside the Applicability Domain of the model
Remarks:
none

3.1 Applicability Domain:
Similar Compounds, with Predicted and Experimental Values



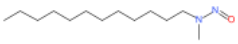
Compound #1

CAS: 1643-20-5
Dataset id:273 (Training Set)
SMILES: [O-][N+](C)(C)CCCCCCCCCCCC
Similarity: 0.718
Experimental value : NON-Carcinogen
Predicted value : Carcinogen



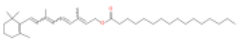
Compound #2

CAS: 75881-20-8
Dataset id:558 (Training Set)
SMILES: O=NN(C)CCCCCCCCCCCCCC
Similarity: 0.688
Experimental value : Carcinogen
Predicted value : Carcinogen



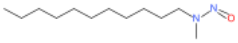
Compound #3

CAS: 55090-44-3
Dataset id:554 (Training Set)
SMILES: O=NN(C)CCCCCCCCCCCC
Similarity: 0.66
Experimental value : Carcinogen
Predicted value : Carcinogen



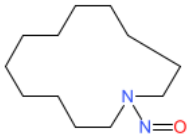
Compound #4

CAS: 79-81-2
Dataset id:693 (Training Set)
SMILES: O=C(OCC=C(C=CC=C(C=CC1=C(C)CCCC1(C)C)C)C)CCCCCCCCCCCCCCCC
Similarity: 0.649
Experimental value : NON-Carcinogen
Predicted value : NON-Carcinogen



Compound #5

CAS: 68107-26-6
Dataset id:603 (Training Set)
SMILES: O=NN(C)CCCCCCCCCCC
Similarity: 0.645
Experimental value : Carcinogen
Predicted value : Carcinogen



Compound #6

CAS: 40580-89-0
Dataset id:586 (Training Set)
SMILES: O=NN1CCCCCCCCCCC1
Similarity: 0.643
Experimental value : Carcinogen
Predicted value : Carcinogen

3.2 Applicability Domain:
Measured Applicability Domain Scores

✗	Global AD Index AD index = 0.355 Explanation: The predicted compound is outside the Applicability Domain of the model.
⚠	Similar molecules with known experimental value Similarity index = 0.702 Explanation: Only moderately similar compounds with known experimental value in the training set have been found..
✗	Accuracy of prediction for similar molecules Accuracy index = 0.484 Explanation: Accuracy of prediction for similar molecules found in the training set is not adequate..
⚠	Concordance for similar molecules Concordance index = 0.516 Explanation: some similar molecules found in the training set have experimental values that disagree with the predicted value..
✓	Model's descriptors range check Descriptors range check = True Explanation: descriptors for this compound have values inside the descriptor range of the compounds of the training set..
✗	Atom Centered Fragments similarity check ACF index = 0.6 Explanation: a prominent number of atom centered fragments of the compound have not been found in the compounds of the training set or are rare fragments (1 unknown fragments found)..
✓	Model class assignment reliability Pos/Non-Pos difference = 0.739 Explanation: model class assignment is well defined..
✓	Neural map neurons concordance Neurons concordance = 1 Explanation: predicted value agrees with experimental values of training set compounds laying in the same neuron..

Symbols explanation:

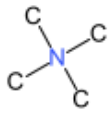
✓ The feature has a good assessment, model is reliable regarding this aspect.

⚠ The feature has a non optimal assessment, this aspect should be reviewed by an expert.

✗ The feature has a bad assessment, model is not reliable regarding this aspect.

4.1 Reasoning:
Relevant Chemical Fragments and Moieties

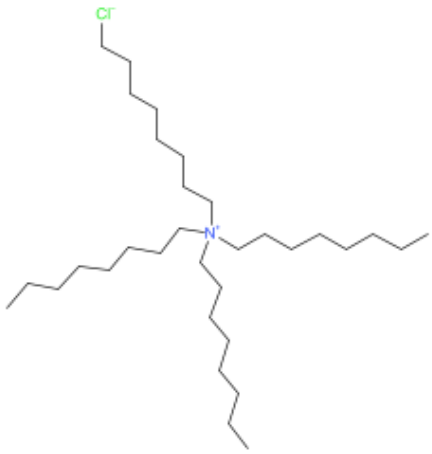
(Molecule 0) Reasoning on rare and missing Atom Centered Fragments .
The following Atom Centered Fragments have been found in the molecule, but they are not found or rarely found in the model's training set:



Fragment defined by the SMILES: CN(C)(C)C
The fragment has never been found in the model's training set

1. Prediction Summary

Prediction for compound Molecule 0 -

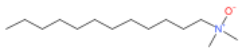


Prediction: Reliability:

Prediction is Carcinogen, but the result may be not reliable. A check of the information given in the following section should be done, paying particular attention to the following issues:
- Only moderately similar compounds with known experimental value in the training set have been found
- similar molecules found in the training set have experimental values that disagree with the predicted value
- some atom centered fragments of the compound have not been found in the compounds of the training set or are rare fragments (1 infrequent_fragments found)
The following alerts have been found: SA8 Aliphatic halogens

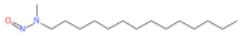
Compound: Molecule 0
Compound SMILES: CCCCCCCC[N+](CCCCCCCC)(CCCCCCCC)CCCCCCCC[Cl-]
Experimental value: -
Predicted Carcinogen activity: Carcinogen
Structural Alerts: SA8 Aliphatic halogens
Reliability: The predicted compound is outside the Applicability Domain of the model
Remarks:
none

3.1 Applicability Domain:
Similar Compounds, with Predicted and Experimental Values



Compound #1

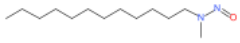
CAS: 1643-20-5
Dataset id:879 (Training Set)
SMILES: [O-][N+](C)(C)CCCCCCCCCCCC
Similarity: 0.718
Experimental value : NON-Carcinogen
Predicted value : NON-Carcinogen



Compound #2

CAS: 75881-20-8
Dataset id:579 (Training Set)
SMILES: O=NN(C)CCCCCCCCCCCCCC
Similarity: 0.688
Experimental value : Carcinogen
Predicted value : Carcinogen

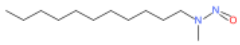
Alerts (not found also in the target): SA21 Alkyl and aryl N-nitroso groups



Compound #3

CAS: 55090-44-3
Dataset id:547 (Training Set)
SMILES: O=NN(C)CCCCCCCCCCCC
Similarity: 0.66
Experimental value : Carcinogen
Predicted value : Carcinogen

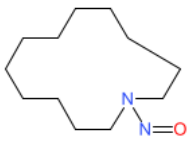
Alerts (not found also in the target): SA21 Alkyl and aryl N-nitroso groups



Compound #4

CAS: 68107-26-6
Dataset id:527 (Training Set)
SMILES: O=NN(C)CCCCCCCCCCC
Similarity: 0.645
Experimental value : Carcinogen
Predicted value : Carcinogen

Alerts (not found also in the target): SA21 Alkyl and aryl N-nitroso groups

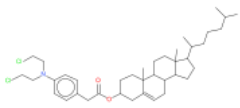


Compound #5

CAS: 40580-89-0
Dataset id:553 (Training Set)
SMILES: O=NN1CCCCCCCCCCC1
Similarity: 0.643
Experimental value : Carcinogen
Predicted value : Carcinogen

Alerts (not found also in the target): SA21 Alkyl and aryl N-nitroso groups

3.1 Applicability Domain:
Similar Compounds, with Predicted and Experimental Values



Compound #6

CAS: 3546-10-9
Dataset id:216 (Training Set)
SMILES:
O=C(OC4CC3=CCC1C(CCC2(C)(C(CCC12)C(C)CCCC(C)C))C3(C)CC4)Cc5ccc(cc5)N(CCCl)CCCl
Similarity: 0.639
Experimental value : Carcinogen
Predicted value : Carcinogen

Alerts (not found also in the target): SA5 S or N mustard

3.2 Applicability Domain:
Measured Applicability Domain Scores

✖	Global AD Index AD index = 0.594 Explanation: The predicted compound is outside the Applicability Domain of the model.
⚠	Similar molecules with known experimental value Similarity index = 0.702 Explanation: Only moderately similar compounds with known experimental value in the training set have been found..
✔	Accuracy of prediction for similar molecules Accuracy index = 1 Explanation: Accuracy of prediction for similar molecules found in the training set is good..
✖	Concordance for similar molecules Concordance index = 0.484 Explanation: similar molecules found in the training set have experimental values that disagree with the predicted value..
⚠	Atom Centered Fragments similarity check ACF index = 0.85 Explanation: some atom centered fragments of the compound have not been found in the compounds of the training set or are rare fragments (1 infrequent_fragments found)..

Symbols explanation:

✔ The feature has a good assessment, model is reliable regarding this aspect.

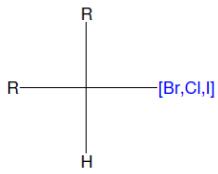
⚠ The feature has a non optimal assessment, this aspect should be reviewed by an expert.

✖ The feature has a bad assessment, model is not reliable regarding this aspect.

4.1 Reasoning:
Relevant Chemical Fragments and Moieties

(Molecule 0) Reasoning on fragments/structural alerts :.

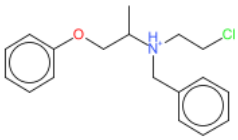
Fragment found: SA8 Aliphatic halogens



R = any atom/group

This alert contains non tertiary aliphatic halogens. Substances fired by Alerts SA2, SA4, SA5 and SA20 should be also excluded.

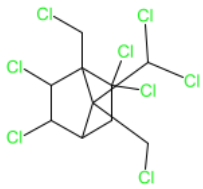
Following, the most similar compounds from the model's dataset having the same fragment.



CAS: 63-92-3
Dataset id:583 (Training Set)
SMILES: O(c1ccccc1)CC(C)[NH+](Cc2ccccc2)CCCl
Similarity: 0.53

Experimental value : Carcinogen
Predicted value : Carcinogen

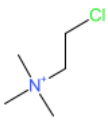
Alerts (found also in the target): SA8 Aliphatic halogens



CAS: 8001-35-2
Dataset id:727 (Training Set)
SMILES: C2C1C(C(C(CCl)(C1(CCl)C(Cl)Cl)C2(Cl)Cl)Cl)Cl
Similarity: 0.527

Experimental value : Carcinogen
Predicted value : Carcinogen

Alerts (found also in the target): SA8 Aliphatic halogens



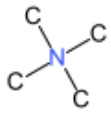
CAS: 999-81-5
Dataset id:100 (Training Set)
SMILES: C[N+](C)(C)CCCl
Similarity: 0.523

Experimental value : NON-Carcinogen
Predicted value : Carcinogen

Alerts (found also in the target): SA8 Aliphatic halogens

4.1 Reasoning:
Relevant Chemical Fragments and Moieties

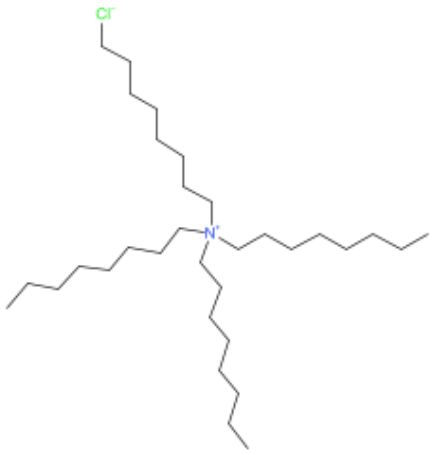
(Molecule 0) Reasoning on rare and missing Atom Centered Fragments .
The following Atom Centered Fragments have been found in the molecule, but they are not found or rarely found in the model's training set:



Fragment defined by the SMILES: CN(C)(C)C
The fragment has less than 3 occurrences in the model's training set

1. Prediction Summary

Prediction for compound Molecule 0 -

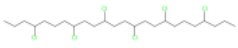


Prediction: Reliability:

Prediction is Carcinogen, but the result shows some critical aspects, which require to be checked:
- Only moderately similar compounds with known experimental value in the training set have been found
- some similar molecules found in the training set have experimental values that disagree with the predicted value
The following relevant fragments have been found: Carcinogenity alert no. 4

Compound: Molecule 0
Compound SMILES: CCCCCCCC[N+](CCCCCCCC)(CCCCCCCC)CCCCCCCC[Cl-]
Experimental value: -
Predicted Carcinogenic activity: Carcinogen
No. alerts for carcinogenicity: 1
Structural Alerts: Carcinogenity alert no. 4
Reliability: The predicted compound could be out of the Applicability Domain of the model
Remarks:
none

3.1 Applicability Domain:
Similar Compounds, with Predicted and Experimental Values



Compound #1

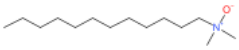
CAS: 63449-39-8
Dataset id:810 (Training Set)
SMILES: CCCC(CCCC(CCC(CCC(CCC(CCCC(CCC)Cl)Cl)Cl)Cl)Cl)Cl
Similarity: 0.743
Experimental value : Carcinogen
Predicted value : Carcinogen

Alerts (not found also in the target): Carcinogenity alert no. 18

Compound #2

CAS: 108171-27-3
Dataset id:675 (Training Set)
SMILES: CC(CCCC(CCCC(CCCC(CCCC(CCCC(CCCl)Cl)Cl)Cl)Cl)Cl)Cl
Similarity: 0.736
Experimental value : Carcinogen
Predicted value : Carcinogen

Alerts (not found also in the target): Carcinogenity alert no. 18



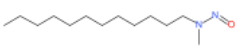
Compound #3

CAS: 1643-20-5
Dataset id:777 (Training Set)
SMILES: [O-][N+](C)(C)CCCCCCCCCCCC
Similarity: 0.718
Experimental value : NON-Carcinogen
Predicted value : Possible NON-Carcinogen

Compound #4

CAS: 75881-20-8
Dataset id:489 (Training Set)
SMILES: O=NN(C)CCCCCCCCCCCCC
Similarity: 0.688
Experimental value : Carcinogen
Predicted value : Carcinogen

Alerts (not found also in the target): Carcinogenity alert no. 1; Carcinogenity alert no. 14; Carcinogenity alert no. 27

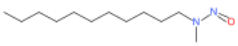


Compound #5

CAS: 55090-44-3
Dataset id:458 (Training Set)
SMILES: O=NN(C)CCCCCCCCCCC
Similarity: 0.66
Experimental value : Carcinogen
Predicted value : Carcinogen

Alerts (not found also in the target): Carcinogenity alert no. 1; Carcinogenity alert no. 14; Carcinogenity alert no. 27

3.1 Applicability Domain:
Similar Compounds, with Predicted and Experimental Values



Compound #6

CAS: 68107-26-6
Dataset id:439 (Training Set)
SMILES: O=NN(C)CCCCCCCCCCC
Similarity: 0.645
Experimental value : Carcinogen
Predicted value : Carcinogen

Alerts (not found also in the target): Carcinogenity alert no. 1; Carcinogenity alert no. 14; Carcinogenity alert no. 27

3.2 Applicability Domain:
Measured Applicability Domain Scores

⚠	Global AD Index AD index = 0.776 Explanation: The predicted compound could be out of the Applicability Domain of the model.
⚠	Similar molecules with known experimental value Similarity index = 0.732 Explanation: Only moderately similar compounds with known experimental value in the training set have been found..
✓	Accuracy of prediction for similar molecules Accuracy index = 1 Explanation: Accuracy of prediction for similar molecules found in the training set is good..
⚠	Concordance for similar molecules Concordance index = 0.676 Explanation: some similar molecules found in the training set have experimental values that disagree with the predicted value..
✓	Atom Centered Fragments similarity check ACF index = 1 Explanation: all atom centered fragment of the compound have been found in the compounds of the training set..

Symbols explanation:

✓ The feature has a good assessment, model is reliable regarding this aspect.

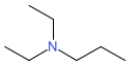
⚠ The feature has a non optimal assessment, this aspect should be reviewed by an expert.

✗ The feature has a bad assessment, model is not reliable regarding this aspect.

4.1 Reasoning:
Relevant Chemical Fragments and Moieties

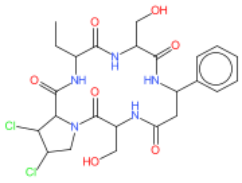
(Molecule 0) Reasoning on fragments/structural alerts :.

Fragment found: Carcinogenity alert no. 4



Structural alert for carcinogenity defined by the SMARTS:CCCN(CC)CC

Following, the most similar compounds from the model's dataset having the same fragment.

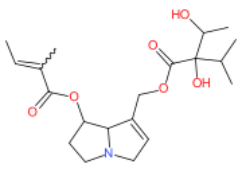


CAS: 12663-46-6
Dataset id:274 (Training Set)
SMILES:
O=C2NC(C(=O)N3CC(C(C3(C(=O)NC(C(=O)NC(C(=O)NC(c1ccccc1)C2)CO)CC))Cl)Cl)CO
Similarity: 0.544

Experimental value : Carcinogen
Predicted value : Carcinogen

Alerts (found also in the target): Carcinogenity alert no. 4

Alerts (not found also in the target): Carcinogenity alert no. 7; Carcinogenity alert no. 11; Carcinogenity alert no. 25; Carcinogenity alert no. 40

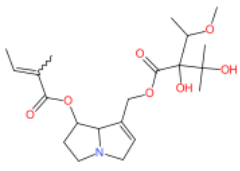


CAS: 22571-95-5
Dataset id:403 (Training Set)
SMILES: O=C(OC2CCN1CC=C(COC(=O)C(O)(C(O)C)C(C)C)C12)C(=CC)C
Similarity: 0.531

Experimental value : Carcinogen
Predicted value : Carcinogen

Alerts (found also in the target): Carcinogenity alert no. 4

Alerts (not found also in the target): Carcinogenity alert no. 20; Carcinogenity alert no. 29



CAS: 303-34-4
Dataset id:160 (Training Set)
SMILES: O=C(OC2CCN1CC=C(COC(=O)C(O)(C(OC)C)C(O)(C)C)C12)C(=CC)C
Similarity: 0.528

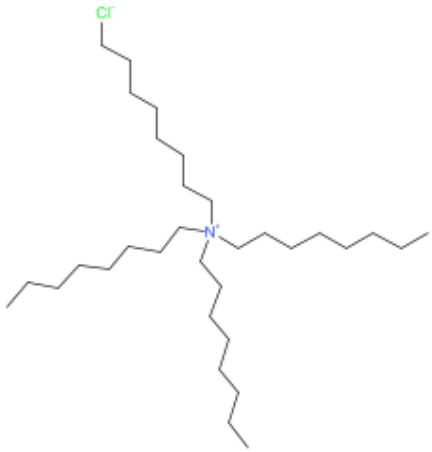
Experimental value : Carcinogen
Predicted value : Carcinogen

Alerts (found also in the target): Carcinogenity alert no. 4

Alerts (not found also in the target): Carcinogenity alert no. 20; Carcinogenity alert no. 29

1. Prediction Summary

Prediction for compound Molecule 0 -

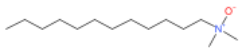


Prediction: Reliability:

Prediction is Carcinogen, but the result may be not reliable. A check of the information given in the following section should be done, paying particular attention to the following issues:
- Only moderately similar compounds with known experimental value in the training set have been found
- Accuracy of prediction for similar molecules found in the training set is not optimal
- some similar molecules found in the training set have experimental values that disagree with the predicted value
- a prominent number of atom centered fragments of the compound have not been found in the compounds of the training set or are rare fragments (1 unknown fragments found)
The following relevant fragments have been found: Carcinogenity alert no. 57

Compound: Molecule 0
Compound SMILES: CCCCCCCC[N+](CCCCCCCC)(CCCCCCCC)CCCCCCCC[Cl-]
Experimental value: -
Predicted Carcinogenic activity: Carcinogen
No. alerts for carcinogenicity: 1
Structural Alerts: Carcinogenity alert no. 57
Reliability: The predicted compound is outside the Applicability Domain of the model
Remarks:
none

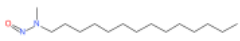
3.1 Applicability Domain:
Similar Compounds, with Predicted and Experimental Values



Compound #1

CAS: 1643-20-5
Dataset id:273 (Training Set)
SMILES: [O-][N+](C)(C)CCCCCCCCCCCC
Similarity: 0.718
Experimental value : NON-Carcinogen
Predicted value : Carcinogen

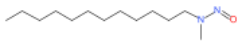
Alerts (not found also in the target): Carcinogenity alert no. 64



Compound #2

CAS: 75881-20-8
Dataset id:558 (Training Set)
SMILES: O=NN(C)CCCCCCCCCCCCCC
Similarity: 0.688
Experimental value : Carcinogen
Predicted value : Carcinogen

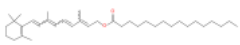
Alerts (not found also in the target): Carcinogenity alert no. 4; Carcinogenity alert no. 8; Carcinogenity alert no. 9; Carcinogenity alert no. 10; Carcinogenity alert no. 15; Carcinogenity alert no. 50; Carcinogenity alert no. 51; Carcinogenity alert no. 54; Carcinogenity alert no. 55; Carcinogenity alert no. 63



Compound #3

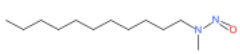
CAS: 55090-44-3
Dataset id:554 (Training Set)
SMILES: O=NN(C)CCCCCCCCCCCC
Similarity: 0.66
Experimental value : Carcinogen
Predicted value : Carcinogen

Alerts (not found also in the target): Carcinogenity alert no. 4; Carcinogenity alert no. 8; Carcinogenity alert no. 9; Carcinogenity alert no. 10; Carcinogenity alert no. 15; Carcinogenity alert no. 50; Carcinogenity alert no. 51; Carcinogenity alert no. 54; Carcinogenity alert no. 55; Carcinogenity alert no. 63



Compound #4

CAS: 79-81-2
Dataset id:693 (Training Set)
SMILES: O=C(OCC=C(C=CC=C(C=CC1=C(C)CCCC1(C)C)C)C)CCCCCCCCCCCCCCCC
Similarity: 0.649
Experimental value : NON-Carcinogen
Predicted value : Possible NON-Carcinogen

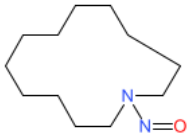


Compound #5

CAS: 68107-26-6
Dataset id:603 (Training Set)
SMILES: O=NN(C)CCCCCCCCCCC
Similarity: 0.645
Experimental value : Carcinogen
Predicted value : Carcinogen

Alerts (not found also in the target): Carcinogenity alert no. 4; Carcinogenity alert no. 8; Carcinogenity alert no. 9; Carcinogenity alert no. 10; Carcinogenity alert no. 15; Carcinogenity alert no. 50; Carcinogenity alert no. 51; Carcinogenity alert no. 54; Carcinogenity alert no. 55; Carcinogenity alert no. 63

3.1 Applicability Domain:
Similar Compounds, with Predicted and Experimental Values



Compound #6

CAS: 40580-89-0
Dataset id:586 (Training Set)
SMILES: O=NN1CCCCCCCCCCC1
Similarity: 0.643
Experimental value : Carcinogen
Predicted value : Carcinogen

Alerts (not found also in the target): Carcinogenity alert no. 4; Carcinogenity alert no. 5; Carcinogenity alert no. 8; Carcinogenity alert no. 9; Carcinogenity alert no. 10; Carcinogenity alert no. 15; Carcinogenity alert no. 50; Carcinogenity alert no. 51; Carcinogenity alert no. 53; Carcinogenity alert no. 54; Carcinogenity alert no. 55; Carcinogenity alert no. 63

3.2 Applicability Domain:
Measured Applicability Domain Scores

✖	Global AD Index AD index = 0.399 Explanation: The predicted compound is outside the Applicability Domain of the model.
⚠	Similar molecules with known experimental value Similarity index = 0.686 Explanation: Only moderately similar compounds with known experimental value in the training set have been found..
⚠	Accuracy of prediction for similar molecules Accuracy index = 0.644 Explanation: Accuracy of prediction for similar molecules found in the training set is not optimal..
⚠	Concordance for similar molecules Concordance index = 0.644 Explanation: some similar molecules found in the training set have experimental values that disagree with the predicted value..
✖	Atom Centered Fragments similarity check ACF index = 0.6 Explanation: a prominent number of atom centered fragments of the compound have not been found in the compounds of the training set or are rare fragments (1 unknown fragments found)..

Symbols explanation:

✔ The feature has a good assessment, model is reliable regarding this aspect.

⚠ The feature has a non optimal assessment, this aspect should be reviewed by an expert.

✖ The feature has a bad assessment, model is not reliable regarding this aspect.

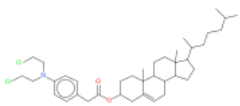
4.1 Reasoning:
Relevant Chemical Fragments and Moieties

(Molecule 0) Reasoning on fragments/structural alerts :.

Fragment found: Carcinogenity alert no. 57

Structural alert for carcinogenity defined by the SMARTS: CCl

Following, the most similar compounds from the model's dataset having the same fragment.

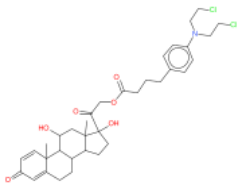


CAS: 3546-10-9
Dataset id:636 (Training Set)
SMILES:
O=C(OC4CC3=CCC1C(CCC2(C)(C(CCC12)C(C)CCCC(C)C))C3(C)CC4)Cc5ccc(cc5)N(CCCl)CCCl
Similarity: 0.639

Experimental value : Carcinogen
Predicted value : Carcinogen

Alerts (found also in the target): Carcinogenity alert no. 57

Alerts (not found also in the target): Carcinogenity alert no. 31; Carcinogenity alert no. 72; Carcinogenity alert no. 73; Carcinogenity alert no. 85

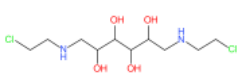


CAS: 29069-24-7
Dataset id:667 (Training Set)
SMILES:
O=C1C=CC4(C(=C1)CCC3C5CCC(O)(C(=O)COC(=O)CCCc2ccc(cc2)N(CCCl)CCCl)C5(C)(CC(O)C34))(C)
Similarity: 0.578

Experimental value : Carcinogen
Predicted value : Carcinogen

Alerts (found also in the target): Carcinogenity alert no. 57

Alerts (not found also in the target): Carcinogenity alert no. 31; Carcinogenity alert no. 72; Carcinogenity alert no. 73; Carcinogenity alert no. 86; Carcinogenity alert no. 106



CAS: 576-68-1
Dataset id:422 (Training Set)
SMILES: OC(CNCCCl)C(O)C(O)C(O)CNCCCl
Similarity: 0.563

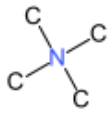
Experimental value : NON-Carcinogen
Predicted value : Carcinogen

Alerts (found also in the target): Carcinogenity alert no. 57

Alerts (not found also in the target): Carcinogenity alert no. 72

4.1 Reasoning:
Relevant Chemical Fragments and Moieties

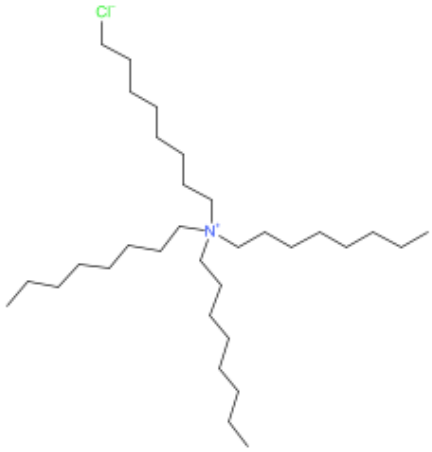
(Molecule 0) Reasoning on rare and missing Atom Centered Fragments .
The following Atom Centered Fragments have been found in the molecule, but they are not found or rarely found in the model's training set:



Fragment defined by the SMILES: CN(C)(C)C
The fragment has never been found in the model's training set

1. Prediction Summary

Prediction for compound Molecule 0 -

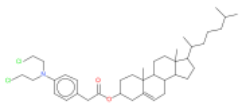


Prediction: Reliability:

Prediction is Carcinogen, but the result may be not reliable. A check of the information given in the following section should be done, paying particular attention to the following issues:
- Only moderately similar compounds with known experimental value in the training set have been found
- similar molecules found in the training set have experimental values that disagree with the predicted value
- some atom centered fragments of the compound have not been found in the compounds of the training set or are rare fragments (1 infrequent_fragments found)

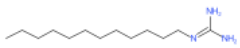
Compound: Molecule 0
Compound SMILES: CCCCCCCC[N+](CCCCCCCC)(CCCCCCCC)CCCCCCCC[Cl-]
Experimental value: -
Predicted Oral Carcinogenic class: Carcinogen
Reliability: The predicted compound is outside the Applicability Domain of the model
Remarks:
none

3.1 Applicability Domain:
Similar Compounds, with Predicted and Experimental Values



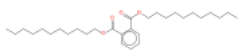
Compound #1

CAS: 3546-10-9
Dataset id:256 (Training Set)
SMILES:
O=C(OC4CC3=CCC1C(CCC2(C)(C(CCC12)C(C)CCCC(C)C))C3(C)CC4)Cc5ccc(cc5)N(CCCl)CCCl
Similarity: 0.639
Experimental value : Carcinogen
Predicted value : Carcinogen



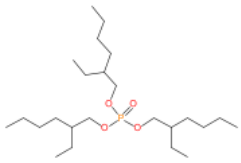
Compound #2

CAS: 2439-10-3
Dataset id:490 (Training Set)
SMILES: N(=C(N)N)CCCCCCCCCCCC
Similarity: 0.628
Experimental value : NON-Carcinogen
Predicted value : NON-Carcinogen



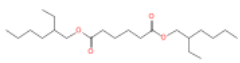
Compound #3

CAS: 3648-20-2
Dataset id:488 (Training Set)
SMILES: O=C(OCCCCCCCCCCC)c1ccccc1(C(=O)OCCCCCCCCCCC)
Similarity: 0.62
Experimental value : NON-Carcinogen
Predicted value : NON-Carcinogen



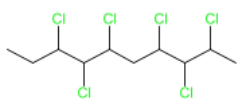
Compound #4

CAS: 78-42-2
Dataset id:313 (Training Set)
SMILES: O=P(OCC(CC)CCCC)(OCC(CC)CCCC)OCC(CC)CCCC
Similarity: 0.617
Experimental value : Carcinogen
Predicted value : Carcinogen



Compound #5

CAS: 103-23-1
Dataset id:94 (Training Set)
SMILES: O=C(OCC(CC)CCCC)CCCCC(=O)OCC(CC)CCCC
Similarity: 0.609
Experimental value : Carcinogen
Predicted value : NON-Carcinogen



Compound #6

CAS: 108171-26-2
Dataset id:65 (Training Set)
SMILES: CCC(C(C(CC(C(C(C)Cl)Cl)Cl)Cl)Cl)Cl
Similarity: 0.586
Experimental value : Carcinogen
Predicted value : Carcinogen

3.2 Applicability Domain:
Measured Applicability Domain Scores

✖	Global AD Index AD index = 0.571 Explanation: The predicted compound is outside the Applicability Domain of the model.
⚠	Similar molecules with known experimental value Similarity index = 0.633 Explanation: Only moderately similar compounds with known experimental value in the training set have been found..
✔	Accuracy of prediction for similar molecules Accuracy index = 1 Explanation: Accuracy of prediction for similar molecules found in the training set is good..
✖	Concordance for similar molecules Concordance index = 0.508 Explanation: similar molecules found in the training set have experimental values that disagree with the predicted value..
✔	Model's descriptors range check Descriptors range check = True Explanation: descriptors for this compound have values inside the descriptor range of the compounds of the training set..
⚠	Atom Centered Fragments similarity check ACF index = 0.85 Explanation: some atom centered fragments of the compound have not been found in the compounds of the training set or are rare fragments (1 infrequent_fragments found)..

Symbols explanation:

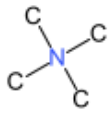
✔ The feature has a good assessment, model is reliable regarding this aspect.

⚠ The feature has a non optimal assessment, this aspect should be reviewed by an expert.

✖ The feature has a bad assessment, model is not reliable regarding this aspect.

4.1 Reasoning:
Relevant Chemical Fragments and Moieties

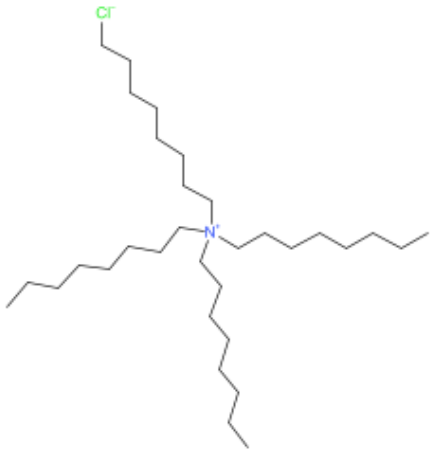
(Molecule 0) Reasoning on rare and missing Atom Centered Fragments .
The following Atom Centered Fragments have been found in the molecule, but they are not found or rarely found in the model's training set:



Fragment defined by the SMILES: CN(C)(C)C
The fragment has less than 3 occurrences in the model's training set

1. Prediction Summary

Prediction for compound Molecule 0 -

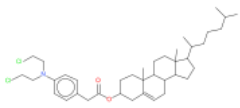


Prediction: Reliability:

Prediction is -1.74, but the result may be not reliable. A check of the information given in the following section should be done, paying particular attention to the following issues:
- No similar compounds with known experimental value in the training set have been found
- Accuracy of prediction for similar molecules found in the training set is not optimal
- similar molecules found in the training set have experimental values that disagree with the predicted value
- the maximum error in prediction of similar molecules found in the training set has a moderate value, considering the experimental variability
- a prominent number of atom centered fragments of the compound have not been found in the compounds of the training set or are rare fragments (1 unknown fragments found)

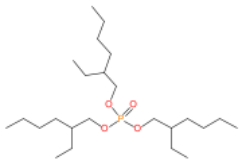
Compound: Molecule 0
Compound SMILES: CCCCCCCC[N+](CCCCCCCC)(CCCCCCCC)CCCCCCCC[Cl-]
Experimental value: -
Predicted Oral Carcinogenicity SF (log form) [log(1/(mg/kg-day))]: -1.74
Predicted Oral Carcinogenicity SF [1/(mg/kg-day)]: 0.0184
Experimental value [1/(mg/kg-day)]: -
Reliability: The predicted compound is outside the Applicability Domain of the model
Remarks:
none

3.1 Applicability Domain:
Similar Compounds, with Predicted and Experimental Values



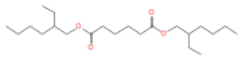
Compound #1

CAS: 3546-10-9
Dataset id:256 (Training Set)
SMILES:
O=C(OC4CC3=CCC1C(CCC2(C)(C(CCC12)C(C)CCCC(C)C))C3(C)CC4)Cc5ccc(cc5)N(CCCl)CCCl
Similarity: 0.639
Experimental value : 2.18
Predicted value : 0.78



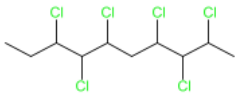
Compound #2

CAS: 78-42-2
Dataset id:313 (Training Set)
SMILES: O=P(OCC(CC)CCCC)(OCC(CC)CCCC)OCC(CC)CCCC
Similarity: 0.617
Experimental value : -2.49
Predicted value : -2.174



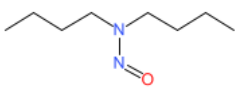
Compound #3

CAS: 103-23-1
Dataset id:94 (Test Set)
SMILES: O=C(OCC(CC)CCCC)CCCCC(=O)OCC(CC)CCCC
Similarity: 0.609
Experimental value : -2.92
Predicted value : -1.999



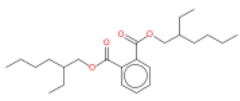
Compound #4

CAS: 108171-26-2
Dataset id:65 (Training Set)
SMILES: CCC(C(C(CC(C(C(C)Cl)Cl)Cl)Cl)Cl)Cl
Similarity: 0.586
Experimental value : -1.05
Predicted value : -1.089



Compound #5

CAS: 924-16-3
Dataset id:224 (Test Set)
SMILES: O=NN(CCCC)CCCC
Similarity: 0.562
Experimental value : 0.73
Predicted value : 0.473



Compound #6

CAS: 117-81-7
Dataset id:44 (Test Set)
SMILES: O=C(OCC(CC)CCCC)c1ccccc1(C(=O)OCC(CC)CCCC)
Similarity: 0.559
Experimental value : -1.85
Predicted value : -2.919

3.2 Applicability Domain:
Measured Applicability Domain Scores

✖	Global AD Index AD index = 0.377 Explanation: The predicted compound is outside the Applicability Domain of the model.
✖	Similar molecules with known experimental value Similarity index = 0.628 Explanation: No similar compounds with known experimental value in the training set have been found..
⚠	Accuracy of prediction for similar molecules Accuracy index = 0.858 Explanation: Accuracy of prediction for similar molecules found in the training set is not optimal..
✖	Concordance for similar molecules Concordance index = 2.335 Explanation: similar molecules found in the training set have experimental values that disagree with the predicted value..
⚠	Maximum error of prediction among similar molecules Max error index = 1.4 Explanation: the maximum error in prediction of similar molecules found in the training set has a moderate value, considering the experimental variability..
✔	Model's descriptors range check Descriptors range check = True Explanation: descriptors for this compound have values inside the descriptor range of the compounds of the training set..
✖	Atom Centered Fragments similarity check ACF index = 0.6 Explanation: a prominent number of atom centered fragments of the compound have not been found in the compounds of the training set or are rare fragments (1 unknown fragments found)..

Symbols explanation:

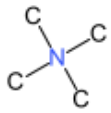
✔ The feature has a good assessment, model is reliable regarding this aspect.

⚠ The feature has a non optimal assessment, this aspect should be reviewed by an expert.

✖ The feature has a bad assessment, model is not reliable regarding this aspect.

4.1 Reasoning:
Relevant Chemical Fragments and Moieties

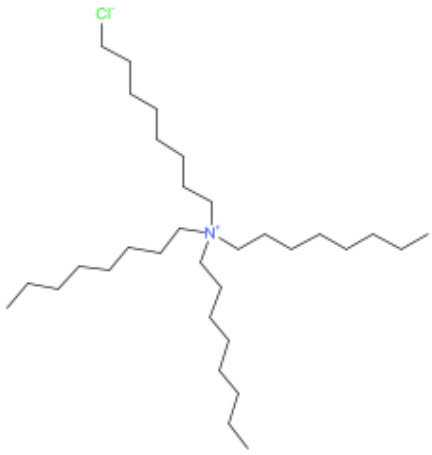
(Molecule 0) Reasoning on rare and missing Atom Centered Fragments .
The following Atom Centered Fragments have been found in the molecule, but they are not found or rarely found in the model's training set:



Fragment defined by the SMILES: CN(C)(C)C
The fragment has never been found in the model's training set

1. Prediction Summary

Prediction for compound Molecule 0 -

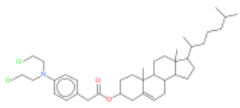


Prediction: Reliability:

Prediction is Carcinogen, but the result may be not reliable. A check of the information given in the following section should be done, paying particular attention to the following issues:
- Only moderately similar compounds with known experimental value in the training set have been found
- Accuracy of prediction for similar molecules found in the training set is not adequate
- similar molecules found in the training set have experimental values that disagree with the predicted value
- a prominent number of atom centered fragments of the compound have not been found in the compounds of the training set or are rare fragments (1 unknown fragments found)

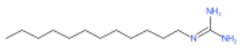
Compound: Molecule 0
Compound SMILES: CCCCCCCC[N+](CCCCCCCC)(CCCCCCCC)CCCCCCCC[Cl-]
Experimental value: -
Predicted Inhalation Carcinogenic class: Carcinogen
Reliability: The predicted compound is outside the Applicability Domain of the model
Remarks:
none

3.1 Applicability Domain:
Similar Compounds, with Predicted and Experimental Values



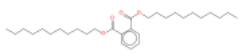
Compound #1

CAS: 3546-10-9
Dataset id:219 (Training Set)
SMILES:
O=C(OC4CC3=CCC1C(CCC2(C)(C(CCC12)C(C)CCCC(C)C))C3(C)CC4)Cc5ccc(cc5)N(CCCl)CCCl
Similarity: 0.639
Experimental value : Carcinogen
Predicted value : Carcinogen



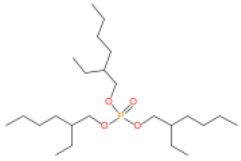
Compound #2

CAS: 2439-10-3
Dataset id:462 (Training Set)
SMILES: N(=C(N)N)CCCCCCCCCCCC
Similarity: 0.628
Experimental value : NON-Carcinogen
Predicted value : Carcinogen



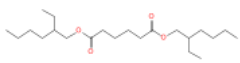
Compound #3

CAS: 3648-20-2
Dataset id:460 (Test Set)
SMILES: O=C(OCCCCCCCCCCC)c1ccccc1(C(=O)OCCCCCCCCCCC)
Similarity: 0.62
Experimental value : NON-Carcinogen
Predicted value : NON-Carcinogen



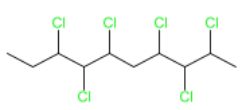
Compound #4

CAS: 78-42-2
Dataset id:741 (Training Set)
SMILES: O=P(OCC(CC)CCCC)(OCC(CC)CCCC)OCC(CC)CCCC
Similarity: 0.617
Experimental value : NON-Carcinogen
Predicted value : NON-Carcinogen



Compound #5

CAS: 103-23-1
Dataset id:391 (Training Set)
SMILES: O=C(OCC(CC)CCCC)CCCCC(=O)OCC(CC)CCCC
Similarity: 0.609
Experimental value : NON-Carcinogen
Predicted value : NON-Carcinogen



Compound #6

CAS: 108171-26-2
Dataset id:54 (Training Set)
SMILES: CCC(C(C(CC(C(C(C)Cl)Cl)Cl)Cl)Cl)Cl
Similarity: 0.586
Experimental value : Carcinogen
Predicted value : NON-Carcinogen

3.2 Applicability Domain:
Measured Applicability Domain Scores

✗	Global AD Index AD index = 0.34 Explanation: The predicted compound is outside the Applicability Domain of the model.
⚠	Similar molecules with known experimental value Similarity index = 0.633 Explanation: Only moderately similar compounds with known experimental value in the training set have been found..
✗	Accuracy of prediction for similar molecules Accuracy index = 0.508 Explanation: Accuracy of prediction for similar molecules found in the training set is not adequate..
✗	Concordance for similar molecules Concordance index = 0.508 Explanation: similar molecules found in the training set have experimental values that disagree with the predicted value..
✓	Model's descriptors range check Descriptors range check = True Explanation: descriptors for this compound have values inside the descriptor range of the compounds of the training set..
✗	Atom Centered Fragments similarity check ACF index = 0.6 Explanation: a prominent number of atom centered fragments of the compound have not been found in the compounds of the training set or are rare fragments (1 unknown fragments found)..

Symbols explanation:

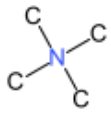
✓ The feature has a good assessment, model is reliable regarding this aspect.

⚠ The feature has a non optimal assessment, this aspect should be reviewed by an expert.

✗ The feature has a bad assessment, model is not reliable regarding this aspect.

4.1 Reasoning:
Relevant Chemical Fragments and Moieties

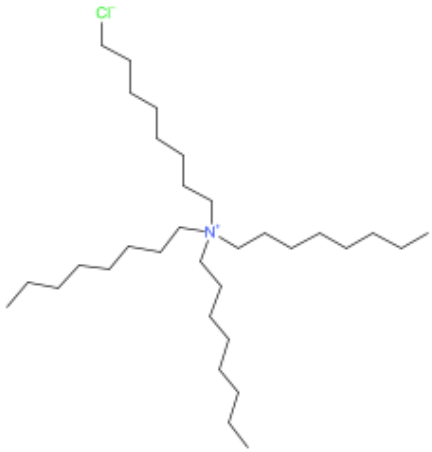
(Molecule 0) Reasoning on rare and missing Atom Centered Fragments .
The following Atom Centered Fragments have been found in the molecule, but they are not found or rarely found in the model's training set:



Fragment defined by the SMILES: CN(C)(C)C
The fragment has never been found in the model's training set

1. Prediction Summary

Prediction for compound Molecule 0 -

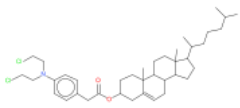


Prediction: Reliability:

Prediction is 0.3, but the result may be not reliable. A check of the information given in the following section should be done, paying particular attention to the following issues:
- No similar compounds with known experimental value in the training set have been found
- Accuracy of prediction for similar molecules found in the training set is not optimal
- similar molecules found in the training set have experimental values that disagree with the predicted value
- the maximum error in prediction of similar molecules found in the training set has a high value, considering the experimental variability
- a prominent number of atom centered fragments of the compound have not been found in the compounds of the training set or are rare fragments (1 unknown fragments found)

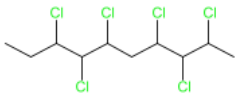
Compound: Molecule 0
Compound SMILES: CCCCCCCC[N+](CCCCCCCC)(CCCCCCCC)CCCCCCCC[Cl-]
Experimental value: -
Predicted Inhalation Carcinogenicity SF (log form) [log(1/(mg/kg-day))]: 0.3
Predicted Inhalation Carcinogenicity SF [1/(mg/kg-day)]: 1.97
Experimental value [1/(mg/kg-day)]: -
Reliability: The predicted compound is outside the Applicability Domain of the model
Remarks:
none

3.1 Applicability Domain:
Similar Compounds, with Predicted and Experimental Values



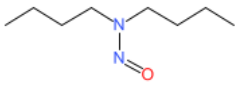
Compound #1

CAS: 3546-10-9
Dataset id:219 (Test Set)
SMILES:
O=C(OC4CC3=CCC1C(CCC2(C)(C(CCC12)C(C)CCCC(C)C))C3(C)CC4)Cc5ccc(cc5)N(CCCl)CCCl
Similarity: 0.639
Experimental value : 2.18
Predicted value : 1.825



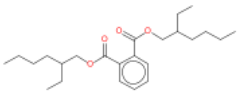
Compound #2

CAS: 108171-26-2
Dataset id:54 (Training Set)
SMILES: CCC(C(C(CC(C(C(C)Cl)Cl)Cl)Cl)Cl)Cl
Similarity: 0.586
Experimental value : -1.06
Predicted value : 0.58



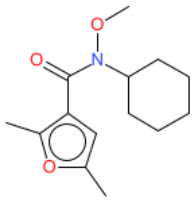
Compound #3

CAS: 924-16-3
Dataset id:192 (Training Set)
SMILES: O=NN(CCCC)CCCC
Similarity: 0.562
Experimental value : 0.75
Predicted value : -0.335



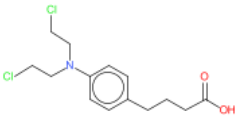
Compound #4

CAS: 117-81-7
Dataset id:38 (Training Set)
SMILES: O=C(OCC(CC)CCCC)c1ccccc1(C(=O)OCC(CC)CCCC)
Similarity: 0.559
Experimental value : -2.08
Predicted value : 0.163



Compound #5

CAS: 60568-05-0
Dataset id:123 (Training Set)
SMILES: O=C(c1cc(oc1C)C)N(OC)C2CCCCC2
Similarity: 0.54
Experimental value : -1.52
Predicted value : 0.268



Compound #6

CAS: 305-03-3
Dataset id:50 (Training Set)
SMILES: O=C(O)CCCc1ccc(cc1)N(CCCl)CCCl
Similarity: 0.529
Experimental value : 2.66
Predicted value : 1.001

3.2 Applicability Domain:
Measured Applicability Domain Scores

✗	Global AD Index AD index = 0.366 Explanation: The predicted compound is outside the Applicability Domain of the model.
✗	Similar molecules with known experimental value Similarity index = 0.611 Explanation: No similar compounds with known experimental value in the training set have been found..
⚠	Accuracy of prediction for similar molecules Accuracy index = 0.998 Explanation: Accuracy of prediction for similar molecules found in the training set is not optimal..
✗	Concordance for similar molecules Concordance index = 1.62 Explanation: similar molecules found in the training set have experimental values that disagree with the predicted value..
✗	Maximum error of prediction among similar molecules Max error index = 1.64 Explanation: the maximum error in prediction of similar molecules found in the training set has a high value, considering the experimental variability..
✓	Model's descriptors range check Descriptors range check = True Explanation: descriptors for this compound have values inside the descriptor range of the compounds of the training set..
✗	Atom Centered Fragments similarity check ACF index = 0.6 Explanation: a prominent number of atom centered fragments of the compound have not been found in the compounds of the training set or are rare fragments (1 unknown fragments found)..

Symbols explanation:

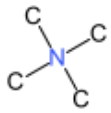
✓ The feature has a good assessment, model is reliable regarding this aspect.

⚠ The feature has a non optimal assessment, this aspect should be reviewed by an expert.

✗ The feature has a bad assessment, model is not reliable regarding this aspect.

4.1 Reasoning:
Relevant Chemical Fragments and Moieties

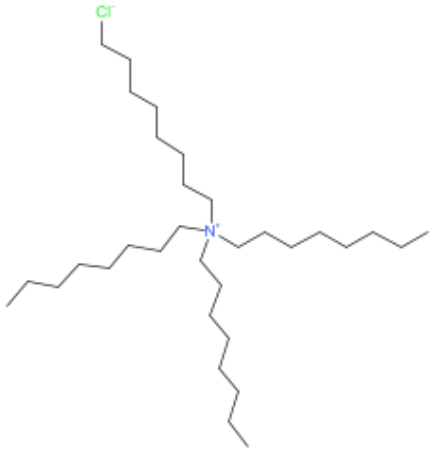
(Molecule 0) Reasoning on rare and missing Atom Centered Fragments .
The following Atom Centered Fragments have been found in the molecule, but they are not found or rarely found in the model's training set:



Fragment defined by the SMILES: CN(C)(C)C
The fragment has never been found in the model's training set

1. Prediction Summary

Prediction for compound Molecule 0 -

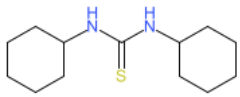


Prediction: Reliability:

Prediction is -2.311, but the result may be not reliable. A check of the information given in the following section should be done, paying particular attention to the following issues:
- No similar compounds with known experimental value in the training set have been found
- some similar molecules found in the training set have experimental values that disagree with the predicted value
- a prominent number of atom centered fragments of the compound have not been found in the compounds of the training set or are rare fragments (1 unknown fragments found)

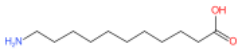
Compound: Molecule 0
Compound SMILES: CCCCCCCC[N+](CCCCCCCC)(CCCCCCCC)CCCCCCCC[Cl-]
Experimental value: -
Predicted TD50 [-log(mg/kg bw/day)]: -2.311
Predicted TD50 [mg/kg bw/day]: 204.66
Experimental TD50 [mg/kg bw/day]: -
Reliability: The predicted compound is outside the Applicability Domain of the model
Remarks:
none

3.1 Applicability Domain:
Similar Compounds, with Predicted and Experimental Values



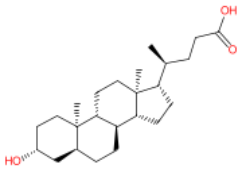
Compound #1

CAS: N.A.
Dataset id:176 (Test Set)
SMILES: C1C(CCCC1)NC(=S)NC1CCCCC1
Similarity: 0.626
Experimental value : -4.193
Predicted value : -3.794



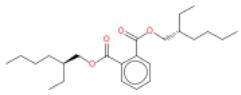
Compound #2

CAS: N.A.
Dataset id:34 (Training Set)
SMILES: C(CCCCCCCCCC(=O)O)N
Similarity: 0.601
Experimental value : -3.041
Predicted value : -2.357



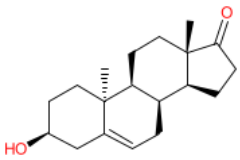
Compound #3

CAS: N.A.
Dataset id:129 (Training Set)
SMILES:
[C@@H]1(C[C@H]2[C@@](CC1)([C@H]1[C@H](CC2)[C@@H]2[C@@](CC1)([C@@H](CC2)[C@@H](C)CCC(=O)O)C)C)O
Similarity: 0.593
Experimental value : -3.352
Predicted value : -0.763



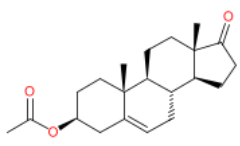
Compound #4

CAS: N.A.
Dataset id:122 (Training Set)
SMILES: c1ccc(c(c1)C(=O)OC[C@H](CCCC)CC)C(=O)OC[C@@H](CCCC)CC
Similarity: 0.559
Experimental value : -3.068
Predicted value : -2.992



Compound #5

CAS: N.A.
Dataset id:164 (Test Set)
SMILES:
C12=CC[C@H]3[C@@H]([C@@]2(CC[C@@H](C1)O)C)CC[C@]1([C@H]3CCC1=O)C
Similarity: 0.552
Experimental value : -1.833
Predicted value : -0.788



Compound #6

CAS: N.A.
Dataset id:81 (Training Set)
SMILES:
[C@H]1(CC2=CC[C@@H]3[C@@H]([C@]2(CC1)C)CC[C@]1([C@H]3CCC1=O)C)OC(=O)C
Similarity: 0.547
Experimental value : -1.559
Predicted value : -1.497

3.2 Applicability Domain:
Measured Applicability Domain Scores

✗	Global AD Index AD index = 0.368 Explanation: The predicted compound is outside the Applicability Domain of the model.
✗	Similar molecules with known experimental value Similarity index = 0.613 Explanation: No similar compounds with known experimental value in the training set have been found..
✓	Accuracy of prediction for similar molecules Accuracy index = 0.542 Explanation: Accuracy of prediction for similar molecules found in the training set is good..
⚠	Concordance for similar molecules Concordance index = 1.306 Explanation: some similar molecules found in the training set have experimental values that disagree with the predicted value..
✓	Maximum error of prediction among similar molecules Max error index = 0.684 Explanation: the maximum error in prediction of similar molecules found in the training set has a low value, considering the experimental variability..
✓	Model's descriptors range check Descriptors range check = True Explanation: descriptors for this compound have values inside the descriptor range of the compounds of the training set..
✗	Atom Centered Fragments similarity check ACF index = 0.6 Explanation: a prominent number of atom centered fragments of the compound have not been found in the compounds of the training set or are rare fragments (1 unknown fragments found)..

Symbols explanation:

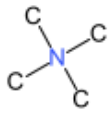
✓ The feature has a good assessment, model is reliable regarding this aspect.

⚠ The feature has a non optimal assessment, this aspect should be reviewed by an expert.

✗ The feature has a bad assessment, model is not reliable regarding this aspect.

4.1 Reasoning:
Relevant Chemical Fragments and Moieties

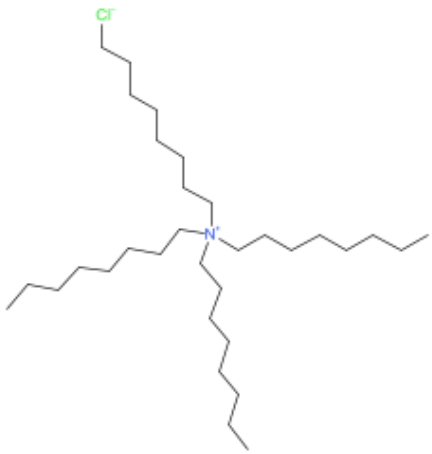
(Molecule 0) Reasoning on rare and missing Atom Centered Fragments .
The following Atom Centered Fragments have been found in the molecule, but they are not found or rarely found in the model's training set:



Fragment defined by the SMILES: CN(C)(C)C
The fragment has never been found in the model's training set

1. Prediction Summary

Prediction for compound Molecule 0 -

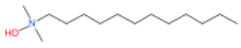


Prediction: Reliability:

Prediction is -2.1856, but the result may be not reliable. A check of the information given in the following section should be done, paying particular attention to the following issues:
- No similar compounds with known experimental value in the training set have been found
- Accuracy of prediction for similar molecules found in the training set is not optimal
- some similar molecules found in the training set have experimental values that disagree with the predicted value
- the maximum error in prediction of similar molecules found in the training set has a high value, considering the experimental variability
- a prominent number of atom centered fragments of the compound have not been found in the compounds of the training set or are rare fragments (1 unknown fragments found)

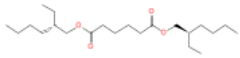
Compound: Molecule 0
Compound SMILES: CCCCCCCC[N+](CCCCCCCC)(CCCCCCCC)CCCCCCCC[Cl-]
Experimental value: -
Predicted TD50 [-log(mg/kg bw/day)]: -2.1856
Predicted TD50 [mg/kg bw/day]: 153.32
Experimental TD50 [mg/kg bw/day]: -
Reliability: The predicted compound is outside the Applicability Domain of the model
Remarks:
none

3.1 Applicability Domain:
Similar Compounds, with Predicted and Experimental Values



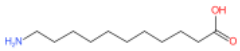
Compound #1

CAS: N.A.
Dataset id:93 (Training Set)
SMILES: CCCCCCCCCCCC[N](O)(C)C
Similarity: 0.709
Experimental value : -2.364
Predicted value : -0.843



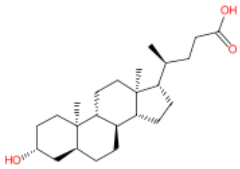
Compound #2

CAS: N.A.
Dataset id:113 (Training Set)
SMILES: CCCC[C@H](CC)COC(=O)CCCCC(=O)OC[C@H](CC)CCCC
Similarity: 0.609
Experimental value : -4.161
Predicted value : -2.71



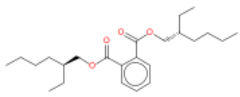
Compound #3

CAS: N.A.
Dataset id:117 (Training Set)
SMILES: C(CCCCCCCCCC(=O)O)N
Similarity: 0.601
Experimental value : -4.649
Predicted value : -4.782



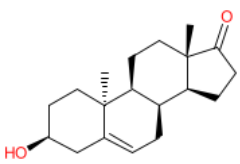
Compound #4

CAS: N.A.
Dataset id:47 (Training Set)
SMILES:
[C@@H]1(C[C@H]2[C@@](CC1)([C@H]1[C@H](CC2)[C@@H]2[C@@](CC1)([C@@H](CC2)[C@@H](C)CCC(=O)O)C)C)O
Similarity: 0.593
Experimental value : -3.276
Predicted value : -4.644



Compound #5

CAS: N.A.
Dataset id:42 (Training Set)
SMILES: c1ccc(c(c1)C(=O)OC[C@H](CCCC)CC)C(=O)OC[C@@H](CCCC)CC
Similarity: 0.559
Experimental value : -3.045
Predicted value : -2.938



Compound #6

CAS: N.A.
Dataset id:25 (Training Set)
SMILES:
C12=CC[C@H]3[C@@H]([C@@]2(CC[C@@H](C1)O)C)CC[C@]1([C@H]3CCC1=O)C
Similarity: 0.552
Experimental value : -1.921
Predicted value : -4.971

3.2 Applicability Domain:
Measured Applicability Domain Scores

✗	Global AD Index AD index = 0.391 Explanation: The predicted compound is outside the Applicability Domain of the model.
✗	Similar molecules with known experimental value Similarity index = 0.652 Explanation: No similar compounds with known experimental value in the training set have been found..
⚠	Accuracy of prediction for similar molecules Accuracy index = 1.486 Explanation: Accuracy of prediction for similar molecules found in the training set is not optimal..
⚠	Concordance for similar molecules Concordance index = 1.077 Explanation: some similar molecules found in the training set have experimental values that disagree with the predicted value..
✗	Maximum error of prediction among similar molecules Max error index = 1.521 Explanation: the maximum error in prediction of similar molecules found in the training set has a high value, considering the experimental variability..
✓	Model's descriptors range check Descriptors range check = True Explanation: descriptors for this compound have values inside the descriptor range of the compounds of the training set..
✗	Atom Centered Fragments similarity check ACF index = 0.6 Explanation: a prominent number of atom centered fragments of the compound have not been found in the compounds of the training set or are rare fragments (1 unknown fragments found)..

Symbols explanation:

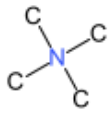
✓ The feature has a good assessment, model is reliable regarding this aspect.

⚠ The feature has a non optimal assessment, this aspect should be reviewed by an expert.

✗ The feature has a bad assessment, model is not reliable regarding this aspect.

4.1 Reasoning:
Relevant Chemical Fragments and Moieties

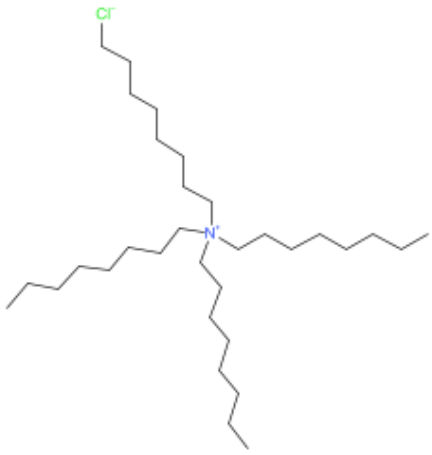
(Molecule 0) Reasoning on rare and missing Atom Centered Fragments .
The following Atom Centered Fragments have been found in the molecule, but they are not found or rarely found in the model's training set:



Fragment defined by the SMILES: CN(C)(C)C
The fragment has never been found in the model's training set

1. Prediction Summary

Prediction for compound Molecule 0 -



Prediction: Reliability:

Prediction is 3879.65 mg/kg, but the result may be not reliable. A check of the information given in the following section should be done, paying particular attention to the following issues:
- Accuracy of prediction for similar molecules found in the training set is not optimal
- the maximum error in prediction of similar molecules found in the training set has a high value, considering the experimental variability

Compound: Molecule 0
Compound SMILES: CCCCCCCC[N+](CCCCCCCC)(CCCCCCCC)CCCCCCCC[Cl-]
Experimental value: -
Predicted log LD50 [log(mmol/Kg)]: 0.889
Predicted log LD50 [mg/Kg]: 3879.65
Molecules used for prediction: 3
Experimental value [mg/Kg]: -
Reliability: The predicted compound is outside the Applicability Domain of the model
Remarks:
none

3.1 Applicability Domain:
Similar Compounds, with Predicted and Experimental Values

Compound #1

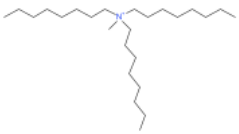
CAS: N.A.
Dataset id:3025 (Training Set)
SMILES: CCCCCCCCCCCCCC[N+](C)(C)CCCCCCCCCCCCCC
Similarity: 0.901
Experimental value : 0.29
Predicted value : 0.729

Compound #2

CAS: N.A.
Dataset id:3077 (Training Set)
SMILES: CCCCCCCCCCCCCCCCCC[N+](C)(C)CCCCCCCCCCCCCCCCCC
Similarity: 0.872
Experimental value : 1.29
Predicted value : 0.397

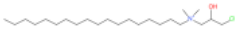
Compound #3

CAS: N.A.
Dataset id:2887 (Training Set)
SMILES: N(CCCCCCCCC)(CCCCCCCCC)CCCCCCCCC
Similarity: 0.869
Experimental value : 1.15
Predicted value : -0.19



Compound #4

CAS: N.A.
Dataset id:2818 (Training Set)
SMILES: CCCCCCCC[N+](C)(CCCCCCCC)CCCCCCCC
Similarity: 0.864
Experimental value : -0.26
Predicted value : 0.266



Compound #5

CAS: N.A.
Dataset id:5710 (Training Set)
SMILES: OC(C[N+](C)(C)CCCCCCCCCCCCCCCCCC)CCl
Similarity: 0.856
Experimental value : 0.78
Predicted value : 0.202

Compound #6

CAS: N.A.
Dataset id:2665 (Training Set)
SMILES: CCCCCCCCCC[N+](C)(C)CCCCCCCCCC
Similarity: 0.834
Experimental value : -0.63
Predicted value : 0.256

3.2 Applicability Domain:
Measured Applicability Domain Scores

✗	Global AD Index AD index = 0.7 Explanation: The predicted compound is outside the Applicability Domain of the model.
✓	Similar molecules with known experimental value Similarity index = 0.88 Explanation: Strongly similar compounds with known experimental value in the training set have been ..
⚠	Accuracy of prediction for similar molecules Accuracy index = 0.891 Explanation: Accuracy of prediction for similar molecules found in the training set is not optimal..
✓	Concordance for similar molecules Concordance index = 0.42 Explanation: Similar molecules found in the training set have experimental values that agree with the predicted value..
✗	Maximum error of prediction among similar molecules Max error index = 1.34 Explanation: the maximum error in prediction of similar molecules found in the training set has a high value, considering the experimental variability..
✓	Atom Centered Fragments similarity check ACF index = 1 Explanation: all atom centered fragment of the compound have been found in the compounds of the training set..

Symbols explanation:

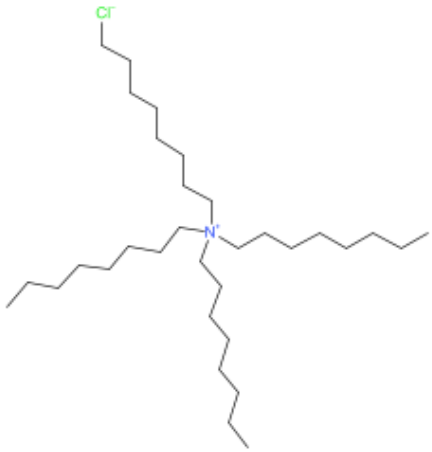
✓ The feature has a good assessment, model is reliable regarding this aspect.

⚠ The feature has a non optimal assessment, this aspect should be reviewed by an expert.

✗ The feature has a bad assessment, model is not reliable regarding this aspect.

1. Prediction Summary

Prediction for compound Molecule 0 -



Prediction: Reliability:

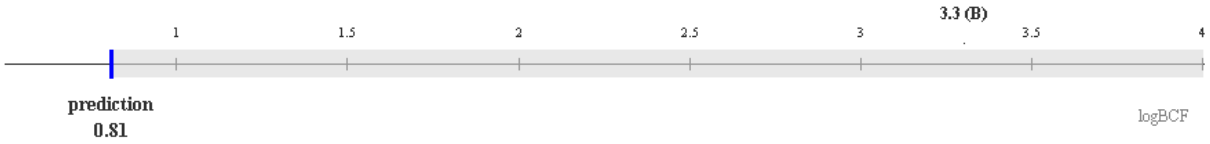
Prediction is 0.81 log(L/kg), but the result may be not reliable. A check of the information given in the following section should be done, paying particular attention to the following issues:
- No similar compounds with known experimental value in the training set have been found
- similar molecules found in the training set have experimental values that disagree with the predicted value
- the maximum error in prediction of similar molecules found in the training set has a moderate value, considering the experimental variability
- some atom centered fragments of the compound have not been found in the compounds of the training set or are rare fragments (1 infrequent_fragments found)
The following relevant fragments have been found: Tertiary amine (SR 05)

Compound: Molecule 0
Compound SMILES: CCCCCCCC[N+](CCCCCCCC)(CCCCCCCC)CCCCCCCC[Cl-]
Experimental value: -
Predicted BCF [log(L/kg)]: 0.81
Predicted BCF [L/kg]: 6
Predicted BCF from sub-model 1 (HM) [log(L/kg)]: 1.03
Predicted BCF from sub-model 2 (GA) [log(L/kg)]: 0.96
Predicted LogP (MLogP): 5.31
Structural Alerts: Tertiary amine (SR 05)
Reliability: The predicted compound is outside the Applicability Domain of the model
Remarks:
none

2. Possible Use and Uncertainty

Threshold 3.3 (bioaccumulative)

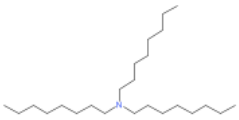
Following, a chart showing the predicted value together with its conservative confidence interval for safe classification.
For the threshold logBCF = 3.3, the current compound can not be associated (due to its Applicability Domain index value) to any conservative interval.
No safe classification can be done.



Threshold 3.7 (very bioaccumulative)

Following, a chart showing the predicted value together with its conservative confidence interval for safe classification.
For the threshold logBCF = 3.7, the current compound can not be associated (due to its Applicability Domain index value) to any conservative interval.
No safe classification can be done.

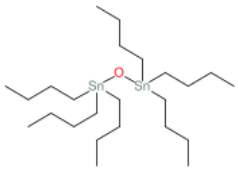
3.1 Applicability Domain:
Similar Compounds, with Predicted and Experimental Values



Compound #1

CAS: 1116-76-3
Dataset id:306 (Training Set)
SMILES: N(CCCCCCCC)(CCCCCCCC)CCCCCCCC
Similarity: 0.846
Experimental value : 1.92
Predicted value : 1.35

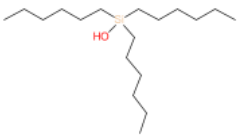
Alerts (found also in the target): Tertiary amine (SR 05)



Compound #2

CAS: 56-35-9
Dataset id:466 (Training Set)
SMILES: O([Sn](CCCC)(CCCC)CCCC)[Sn](CCCC)(CCCC)CCCC
Similarity: 0.652
Experimental value : 3.85
Predicted value : 3.686

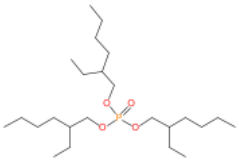
Alerts (not found also in the target): Sn atom in the molecule (SO 04)



Compound #3

CAS: 60782-58-3
Dataset id:465 (Training Set)
SMILES: O[Si](CCCCCC)(CCCCCC)CCCCCC
Similarity: 0.643
Experimental value : 1.48
Predicted value : 1.927

Alerts (not found also in the target): Si atom in the molecule (SO 03); OH group (PG 06)



Compound #4

CAS: 78-42-2
Dataset id:405 (Training Set)
SMILES: O=P(OCC(CC)CCCC)(OCC(CC)CCCC)OCC(CC)CCCC
Similarity: 0.617
Experimental value : 1.19
Predicted value : 1.31

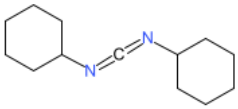
Alerts (not found also in the target): PO2 residue (SR 03)

Compound #5

CAS: 28299-29-8
Dataset id:290 (Training Set)
SMILES: O=C(O)CC(C(=O)O)CCCCCCCCC=CCCCCCCCC
Similarity: 0.61
Experimental value : 0.22
Predicted value : 0.993

Alerts (not found also in the target): Carbonyl residue (SR 02); COOH group (PG 01)

3.1 Applicability Domain:
Similar Compounds, with Predicted and Experimental Values



Compound #6

CAS: 538-75-0
Dataset id:373 (Training Set)
SMILES: C(=NC1CCCCC1)=NC2CCCCC2
Similarity: 0.606
Experimental value : 0.34
Predicted value : 1.39

3.2 Applicability Domain:
Measured Applicability Domain Scores

✗	Global AD Index AD index = 0.613 Explanation: The predicted compound is outside the Applicability Domain of the model.
✗	Similar molecules with known experimental value Similarity index = 0.721 Explanation: No similar compounds with known experimental value in the training set have been found..
✓	Accuracy of prediction for similar molecules Accuracy index = 0.367 Explanation: Accuracy of prediction for similar molecules found in the training set is good..
✗	Concordance for similar molecules Concordance index = 2.075 Explanation: similar molecules found in the training set have experimental values that disagree with the predicted value..
⚠	Maximum error of prediction among similar molecules Max error index = 0.57 Explanation: the maximum error in prediction of similar molecules found in the training set has a moderate value, considering the experimental variability..
✓	Model's descriptors range check Descriptors range check = True Explanation: descriptors for this compound have values inside the descriptor range of the compounds of the training set..
⚠	Atom Centered Fragments similarity check ACF index = 0.85 Explanation: some atom centered fragments of the compound have not been found in the compounds of the training set or are rare fragments (1 infrequent_fragments found)..

Symbols explanation:

✓ The feature has a good assessment, model is reliable regarding this aspect.

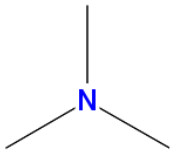
⚠ The feature has a non optimal assessment, this aspect should be reviewed by an expert.

✗ The feature has a bad assessment, model is not reliable regarding this aspect.

4.1 Reasoning:
Relevant Chemical Fragments and Moieties

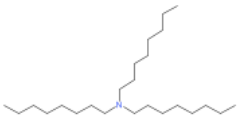
(Molecule 0) Reasoning on fragments/structural alerts :.

Fragment found: Tertiary amine (SR 05)



This chemical contains a tertiary amine. This residue has been found to be present in a large numberof non-bioaccumulative compounds (28), even when the logP value was higher than 3.

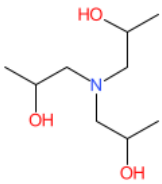
Following, the most similar compounds from the model's dataset having the same fragment.



CAS: 1116-76-3
Dataset id:306 (Training Set)
SMILES: N(CCCCCCCC)(CCCCCCCC)CCCCCCCC
Similarity: 0.846

Experimental value : 1.92
Predicted value : 1.35

Alerts (found also in the target): Tertiary amine (SR 05)

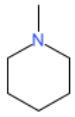


CAS: 122-20-3
Dataset id:311 (Training Set)
SMILES: OC(C)CN(CC(O)C)CC(O)C
Similarity: 0.532

Experimental value : -0.24
Predicted value : 0.004

Alerts (found also in the target): Tertiary amine (SR 05)

Alerts (not found also in the target): OH group (PG 06)



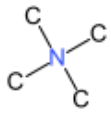
CAS: 626-67-5
Dataset id:441 (Training Set)
SMILES: N1(C)CCCCC1
Similarity: 0.495

Experimental value : 0.65
Predicted value : 0.459

Alerts (found also in the target): Tertiary amine (SR 05)

4.1 Reasoning:
Relevant Chemical Fragments and Moieties

(Molecule 0) Reasoning on rare and missing Atom Centered Fragments .
The following Atom Centered Fragments have been found in the molecule, but they are not found or rarely found in the model's training set:

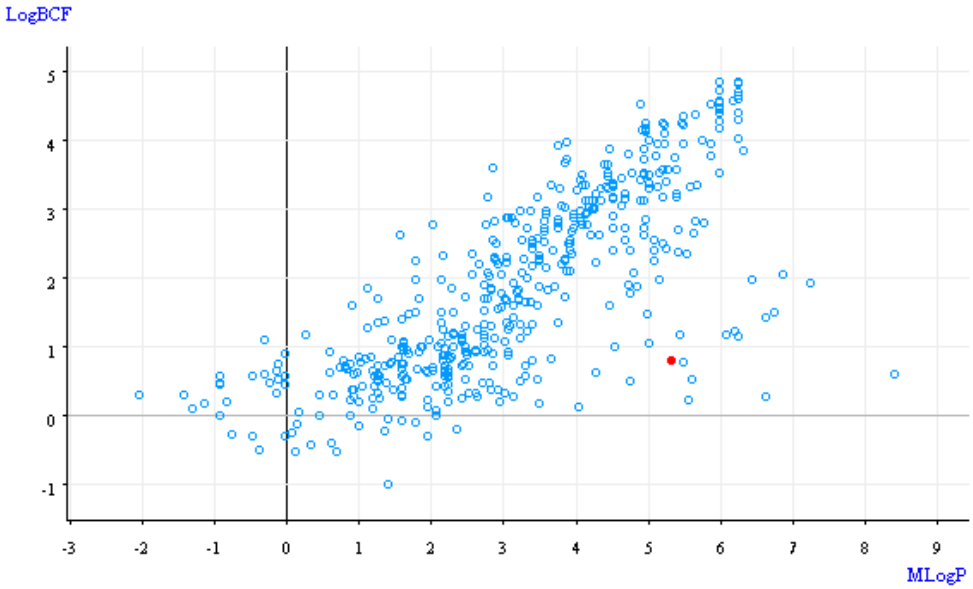


Fragment defined by the SMILES: CN(C)(C)C
The fragment has less than 3 occurrences in the model's training set

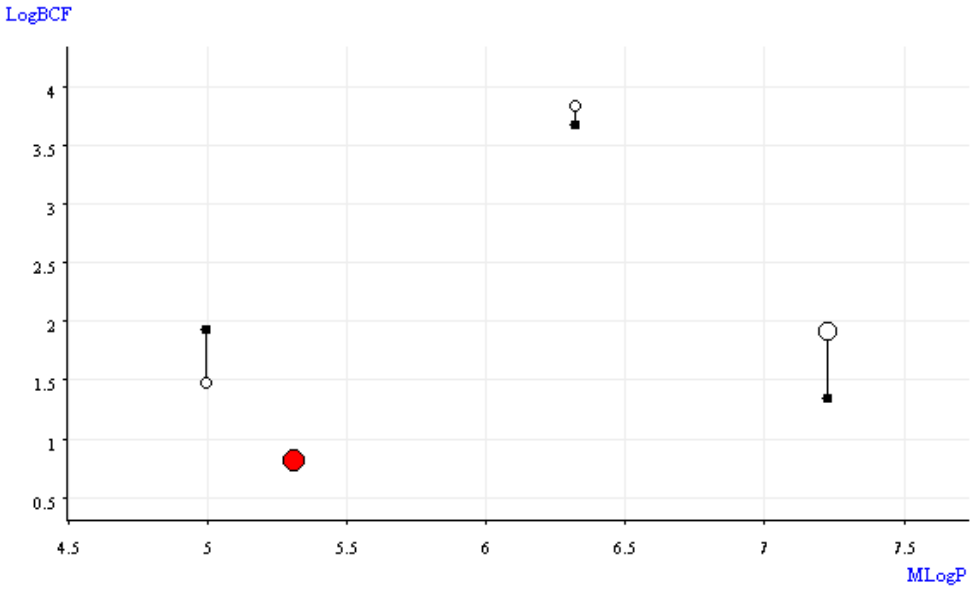
4.2 Reasoning:
Analysis of Molecular Descriptors

Descriptor name: MLogP
Descriptor name: LogP is directly correlated to the logBCF value.

Following, a scatterplot of MLogP against response values; experimental values are reported for the training set, predicted value for the studied compound. Light blue dots represent values of compounds from training set, red dot is the value of the studied compound.

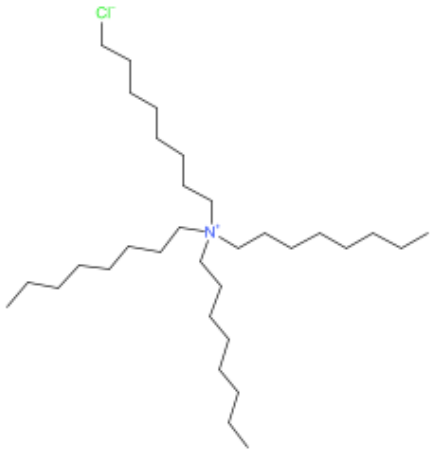


Following, a scatterplot of MLogP against response values only for 3 most similar compounds in the training set. Red dot is the value of the studied compound, black outlined circles represents experimental values of compounds from training set, black dots represents predicted value of the same compound; the size of the circle is proportional to the similarity to the studied compound.



1. Prediction Summary

Prediction for compound Molecule 0 -

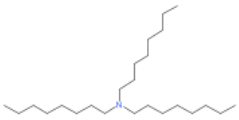


Prediction: Reliability:

Prediction is 0.5 log(L/kg), but the result may be not reliable. A check of the information given in the following section should be done, paying particular attention to the following issues:
- No similar compounds with known experimental value in the training set have been found
- Accuracy of prediction for similar molecules found in the training set is not optimal
- some similar molecules found in the training set have experimental values that disagree with the predicted value
- the maximum error in prediction of similar molecules found in the training set has a high value, considering the experimental variability
- reliability of logP value used by the model is not adequate
- a prominent number of atom centered fragments of the compound have not been found in the compounds of the training set or are rare fragments (1 unknown fragments found)

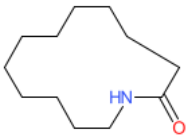
Compound: Molecule 0
Compound SMILES: CCCCCCCC[N+](CCCCCCCC)(CCCCCCCC)CCCCCCCC[Cl-]
Experimental value: -
Predicted BCF [log(L/kg)]: 0.5
Predicted BCF [L/kg]: 3
Predicted LogP (Meylan/Kowwin): 9.82
Predicted LogP reliability: Low
MW: 498.22
Ionic compound: yes
Reliability: The predicted compound is outside the Applicability Domain of the model
Remarks:
none

3.1 Applicability Domain:
Similar Compounds, with Predicted and Experimental Values



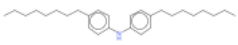
Compound #1

CAS: 1116-76-3
Dataset id:647 (Test Set)
SMILES: N(CCCCCCCC)(CCCCCCCC)CCCCCCCC
Similarity: 0.846
Experimental value : 1.93
Predicted value : 2.482



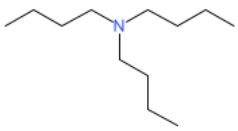
Compound #2

CAS: 947-04-6
Dataset id:219 (Training Set)
SMILES: O=C1NCCCCCCCCCCC1
Similarity: 0.653
Experimental value : 0.41
Predicted value : 1.594



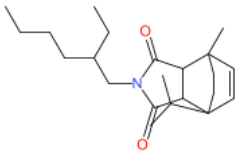
Compound #3

CAS: 26603-23-6
Dataset id:526 (Training Set)
SMILES: c1cc(ccc1Nc2ccc(cc2)CCCCCCCC)CCCCCCCC
Similarity: 0.652
Experimental value : 1.54
Predicted value : 2.036



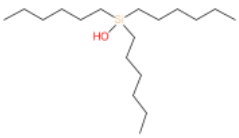
Compound #4

CAS: 102-82-9
Dataset id:614 (Test Set)
SMILES: N(CCCC)(CCCC)CCCC
Similarity: 0.652
Experimental value : 1.4
Predicted value : 1.234



Compound #5

CAS: 13358-11-7
Dataset id:450 (Training Set)
SMILES: O=C1N(C(=O)C3C1C2(C=CC3(CC2)C(C)C)(C))CC(CC)CCCC
Similarity: 0.646
Experimental value : 2.97
Predicted value : 3.599



Compound #6

CAS: 60782-58-3
Dataset id:518 (Training Set)
SMILES: O[Si](CCCCCC)(CCCCCC)CCCCCC
Similarity: 0.643
Experimental value : 2.7
Predicted value : 2.098

3.2 Applicability Domain:
Measured Applicability Domain Scores

✖	Global AD Index AD index = 0.433 Explanation: The predicted compound is outside the Applicability Domain of the model.
✖	Similar molecules with known experimental value Similarity index = 0.722 Explanation: No similar compounds with known experimental value in the training set have been found..
⚠	Accuracy of prediction for similar molecules Accuracy index = 0.868 Explanation: Accuracy of prediction for similar molecules found in the training set is not optimal..
⚠	Concordance for similar molecules Concordance index = 0.76 Explanation: some similar molecules found in the training set have experimental values that disagree with the predicted value..
✖	Maximum error of prediction among similar molecules Max error index = 1.184 Explanation: the maximum error in prediction of similar molecules found in the training set has a high value, considering the experimental variability..
✖	Reliability of logP prediction LogP reliability = 0 Explanation: reliability of logP value used by the model is not adequate..
✔	Model's descriptors range check Descriptors range check = True Explanation: descriptors for this compound have values inside the defined range..
✖	Atom Centered Fragments similarity check ACF index = 0.6 Explanation: a prominent number of atom centered fragments of the compound have not been found in the compounds of the training set or are rare fragments (1 unknown fragments found)..

Symbols explanation:

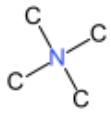
✔ The feature has a good assessment, model is reliable regarding this aspect.

⚠ The feature has a non optimal assessment, this aspect should be reviewed by an expert.

✖ The feature has a bad assessment, model is not reliable regarding this aspect.

4.1 Reasoning:
Relevant Chemical Fragments and Moieties

(Molecule 0) Reasoning on rare and missing Atom Centered Fragments .
The following Atom Centered Fragments have been found in the molecule, but they are not found or rarely found in the model's training set:

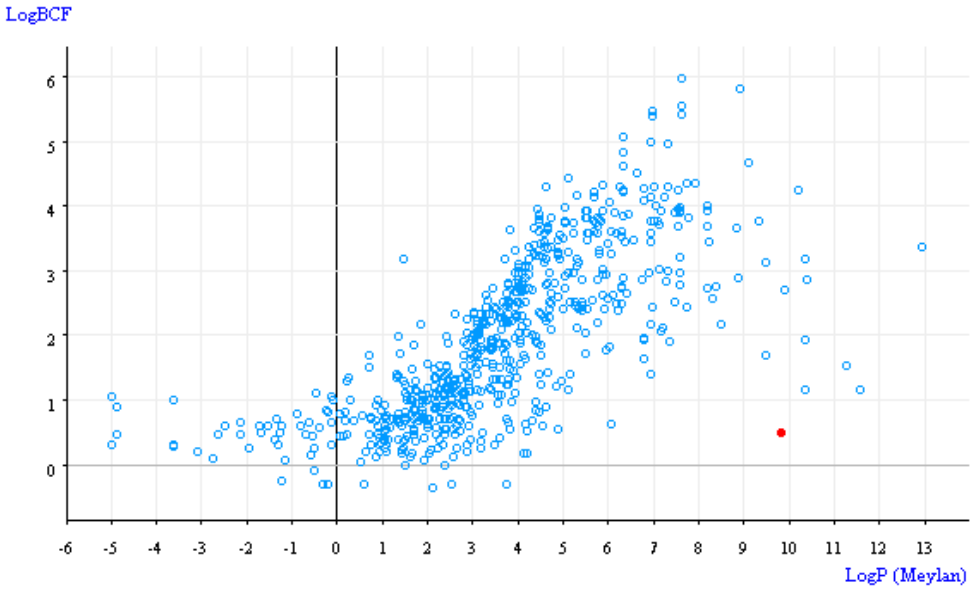


Fragment defined by the SMILES: CN(C)(C)C
The fragment has never been found in the model's training set

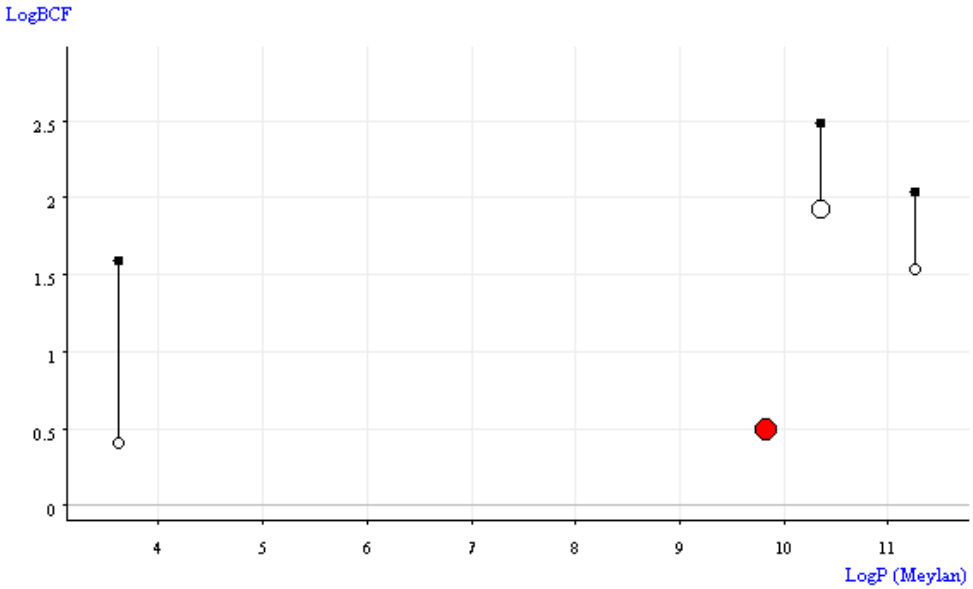
4.2 Reasoning:
Analysis of Molecular Descriptors

Descriptor name: LogP (Meylan)
Descriptor name: LogP is directly correlated to the logBCF value.

Following, a scatterplot of LogP (Meylan) against response values; experimental values are reported for the training set, predicted value for the studied compound. Light blue dots represent values of compounds from training set, red dot is the value of the studied compound.

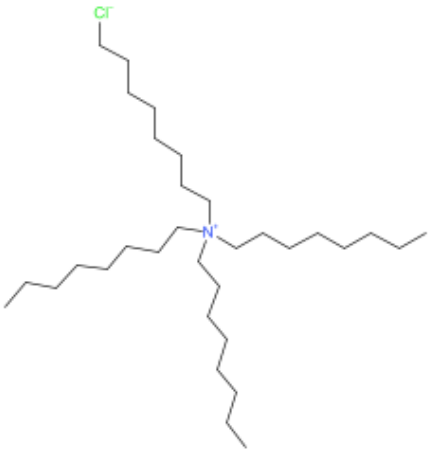


Following, a scatterplot of LogP (Meylan) against response values only for 3 most similar compounds in the training set. Red dot is the value of the studied compound, black outlined circles represents experimental values of compounds from training set, black dots represents predicted value of the same compound; the size of the circle is proportional to the similarity to the studied compound.



1. Prediction Summary

Prediction for compound Molecule 0 -

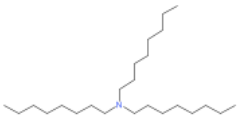


Prediction: Reliability:

Prediction is 1.47 log(L/kg), but the result may be not reliable. A check of the information given in the following section should be done, paying particular attention to the following issues:
- Only moderately similar compounds with known experimental value in the training set have been found
- Accuracy of prediction for similar molecules found in the training set is not adequate
- the maximum error in prediction of similar molecules found in the training set has a high value, considering the experimental variability
- reliability of logP value used by the model is not adequate
- some atom centered fragments of the compound have not been found in the compounds of the training set or are rare fragments (1 infrequent_fragments found)

Compound: Molecule 0
Compound SMILES: CCCCCCCC[N+](CCCCCCCC)(CCCCCCCC)CCCCCCCC[Cl-]
Experimental value: -
Predicted BCF (up) [log(L/kg)]: 1.47
Predicted BCF (up) [L/kg]: 30
Predicted BCF (low) [log(L/kg)]: 1.66
Predicted BCF (low) [L/kg]: 46
Predicted BCF (mid) [log(L/kg)]: 1.62
Predicted BCF (mid) [L/kg]: 41
Predicted LogP (Meylan/Kowwin): 9.82
Predicted LogP reliability: Low
Predicted kM (Meylan): 2.1
Predicted kM reliability: Low
Reliability: The predicted compound is outside the Applicability Domain of the model
Remarks:
none

3.1 Applicability Domain:
Similar Compounds, with Predicted and Experimental Values



Compound #1

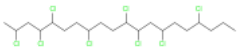
CAS: 1116-76-3
Dataset id:72 (Training Set)
SMILES: N(CCCCCCCC)(CCCCCCCC)CCCCCCCC
Similarity: 0.846
Experimental value : 1.86
Predicted value : 0.412

Compound #2

CAS: 4051-66-5
Dataset id:780 (Training Set)
SMILES: O=C(NCCCCCCCCCCCCCCCCCC)NCCCCCCCCCCCCCCCCCC
Similarity: 0.787
Experimental value : 1.235
Predicted value : -0.049

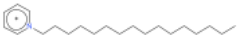
Compound #3

CAS: 110-30-5
Dataset id:315 (Training Set)
SMILES: O=C(NCCNC(=O)CCCCCCCCCCCCCCCCCC)CCCCCCCCCCCCCCCCCC
Similarity: 0.771
Experimental value : 0.32
Predicted value : -0.049



Compound #4

CAS: 61788-76-9
Dataset id:433 (Training Set)
SMILES: CCCC(CCC(CC(C(CCC(CCC(C(CC(C)Cl)Cl)Cl)Cl)Cl)Cl)Cl)Cl
Similarity: 0.685
Experimental value : 1.69
Predicted value : 0.304



Compound #5

CAS: 140-72-7
Dataset id:579 (Training Set)
SMILES: c1cc[n+](cc1)CCCCCCCCCCCCCCCC
Similarity: 0.683
Experimental value : 1.543
Predicted value : 1.614

Compound #6

CAS: 10496-18-1
Dataset id:99 (Training Set)
SMILES: CCCCCCCCCCSSCCCCCCCCCC
Similarity: 0.657
Experimental value : 1.15
Predicted value : 0.516

3.2 Applicability Domain:
Measured Applicability Domain Scores

✖	Global AD Index AD index = 0.692 Explanation: The predicted compound is outside the Applicability Domain of the model.
⚠	Similar molecules with known experimental value Similarity index = 0.814 Explanation: Only moderately similar compounds with known experimental value in the training set have been found..
✖	Accuracy of prediction for similar molecules Accuracy index = 1.366 Explanation: Accuracy of prediction for similar molecules found in the training set is not adequate..
✔	Concordance for similar molecules Concordance index = 0.312 Explanation: Similar molecules found in the training set have experimental values that agree with the predicted value..
✖	Maximum error of prediction among similar molecules Max error index = 1.448 Explanation: the maximum error in prediction of similar molecules found in the training set has a high value, considering the experimental variability..
✖	Reliability of logP prediction LogP reliability = 0 Explanation: reliability of logP value used by the model is not adequate..
⚠	Atom Centered Fragments similarity check ACF index = 0.85 Explanation: some atom centered fragments of the compound have not been found in the compounds of the training set or are rare fragments (1 infrequent_fragments found)..

Symbols explanation:

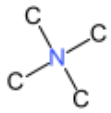
✔ The feature has a good assessment, model is reliable regarding this aspect.

⚠ The feature has a non optimal assessment, this aspect should be reviewed by an expert.

✖ The feature has a bad assessment, model is not reliable regarding this aspect.

4.1 Reasoning:
Relevant Chemical Fragments and Moieties

(Molecule 0) Reasoning on rare and missing Atom Centered Fragments .
The following Atom Centered Fragments have been found in the molecule, but they are not found or rarely found in the model's training set:

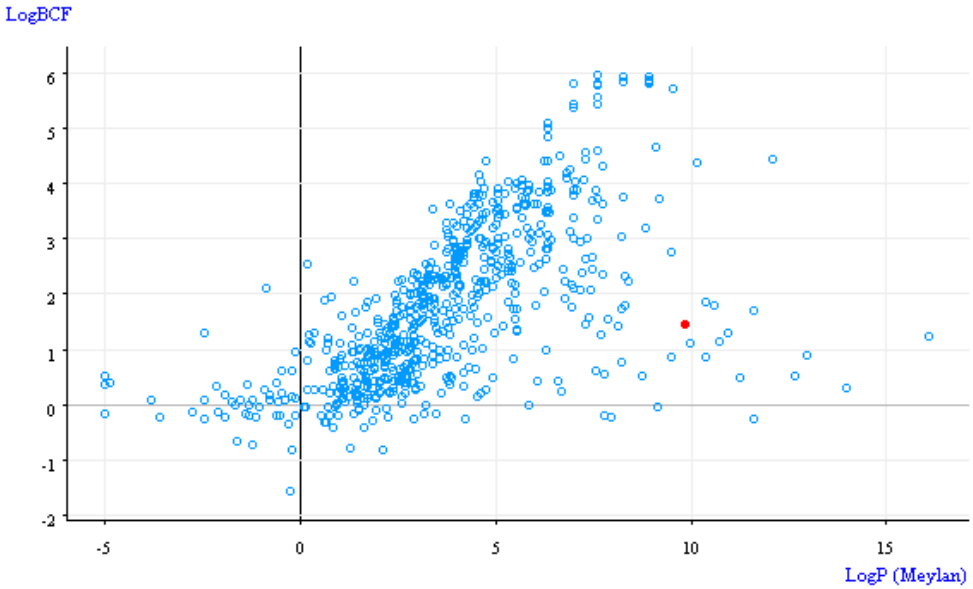


Fragment defined by the SMILES: CN(C)(C)C
The fragment has less than 3 occurrences in the model's training set

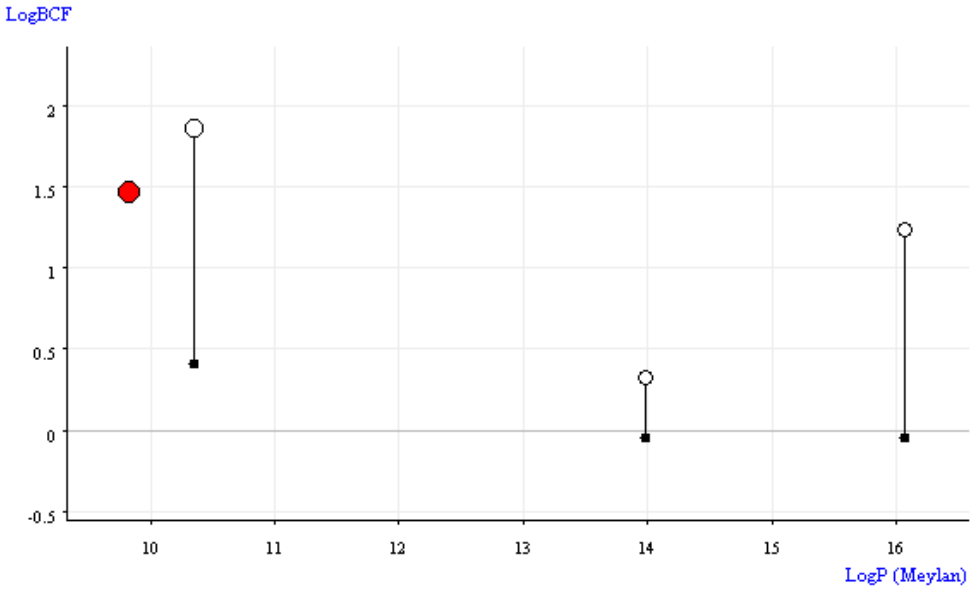
4.2 Reasoning:
Analysis of Molecular Descriptors

Descriptor name: LogP (Meylan)
Descriptor name: LogP is directly correlated to the logBCF value.

Following, a scatterplot of LogP (Meylan) against response values; experimental values are reported for the training set, predicted value for the studied compound. Light blue dots represent values of compounds from training set, red dot is the value of the studied compound.

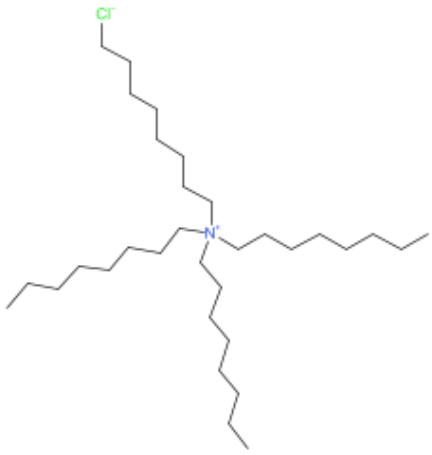


Following, a scatterplot of LogP (Meylan) against response values only for 3 most similar compounds in the training set. Red dot is the value of the studied compound, black outlined circles represents experimental values of compounds from training set, black dots represents predicted value of the same compound; the size of the circle is proportional to the similarity to the studied compound.



1. Prediction Summary

Prediction for compound Molecule 0 -

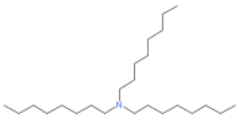


Prediction: Reliability:

Prediction is 1.33 log(L/kg), but the result may be not reliable. A check of the information given in the following section should be done, paying particular attention to the following issues:
- Accuracy of prediction for similar molecules found in the training set is not optimal
- some similar molecules found in the training set have experimental values that disagree with the predicted value
- the maximum error in prediction of similar molecules found in the training set has a high value, considering the experimental variability
- some atom centered fragments of the compound have not been found in the compounds of the training set or are rare fragments (1 infrequent_fragments found)

Compound: Molecule 0
Compound SMILES: CCCCCCCC[N+](CCCCCCCC)(CCCCCCCC)CCCCCCCC[Cl-]
Experimental value: -
Predicted BCF [log(L/kg)]: 1.33
Molecules used for prediction: 3
Reliability: The predicted compound is outside the Applicability Domain of the model
Remarks:
none

3.1 Applicability Domain:
Similar Compounds, with Predicted and Experimental Values



Compound #1

CAS: 1116-76-3
Dataset id:402 (Training Set)
SMILES: N(CCCCCCCC)(CCCCCCCC)CCCCCCCC
Similarity: 0.846
Experimental value : 1.969
Predicted value : 0.739

Compound #2

CAS: 4051-66-5
Dataset id:467 (Training Set)
SMILES: O=C(NCCCCCCCCCCCCCCCCCC)NCCCCCCCCCCCCCCCCCC
Similarity: 0.787
Experimental value : 1.477
Predicted value : 1.032

Compound #3

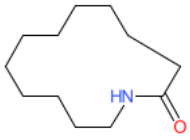
CAS: 110-30-5
Dataset id:213 (Training Set)
SMILES: O=C(NCCNC(=O)CCCCCCCCCCCCCCCCCC)CCCCCCCCCCCCCCCCCC
Similarity: 0.771
Experimental value : 0.319
Predicted value : 1.727

Compound #4

CAS: 26787-65-5
Dataset id:539 (Training Set)
SMILES: O=C(OCCO)CCCCCCCCCCCCCCCCCCCCCCCCCCCC
Similarity: 0.685
Experimental value : 1.271
Predicted value : 1.278

Compound #5

CAS: 10496-18-1
Dataset id:495 (Training Set)
SMILES: CCCCCCCCCCSSCCCCCCCCCC
Similarity: 0.657
Experimental value : 1.153
Predicted value : 0.52



Compound #6

CAS: 947-04-6
Dataset id:396 (Training Set)
SMILES: O=C1NCCCCCCCCCCC1
Similarity: 0.653
Experimental value : 0.164
Predicted value : 0.807

3.2 Applicability Domain:
Measured Applicability Domain Scores

✖	Global AD Index AD index = 0.677 Explanation: The predicted compound is outside the Applicability Domain of the model.
✔	Similar molecules with known experimental value Similarity index = 0.797 Explanation: Strongly similar compounds with known experimental value in the training set have been ..
⚠	Accuracy of prediction for similar molecules Accuracy index = 1.028 Explanation: Accuracy of prediction for similar molecules found in the training set is not optimal..
⚠	Concordance for similar molecules Concordance index = 0.6 Explanation: some similar molecules found in the training set have experimental values that disagree with the predicted value..
✖	Maximum error of prediction among similar molecules Max error index = 1.408 Explanation: the maximum error in prediction of similar molecules found in the training set has a high value, considering the experimental variability..
⚠	Atom Centered Fragments similarity check ACF index = 0.85 Explanation: some atom centered fragments of the compound have not been found in the compounds of the training set or are rare fragments (1 infrequent_fragments found)..

Symbols explanation:

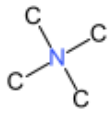
✔ The feature has a good assessment, model is reliable regarding this aspect.

⚠ The feature has a non optimal assessment, this aspect should be reviewed by an expert.

✖ The feature has a bad assessment, model is not reliable regarding this aspect.

4.1 Reasoning:
Relevant Chemical Fragments and Moieties

(Molecule 0) Reasoning on rare and missing Atom Centered Fragments .
The following Atom Centered Fragments have been found in the molecule, but they are not found or rarely found in the model's training set:



Fragment defined by the SMILES: CN(C)(C)C
The fragment has less than 3 occurrences in the model's training set

References and Documentation

You can find complete details on each model and on how to read results in the proper model's guide, available on-line at www.vega-qsar.eu or directly in the VegaNIC application.

Mutagenicity (Ames test) CONSENSUS model(version 1.0.4)

Mutagenicity (Ames test) Consensus model based on the predictions of VEGA mutagenicity models.

Mutagenicity (Ames test) model (CAESAR)(version 2.1.14)

QSAR classification model for Mutagenicity (from CAESAR project)

Mutagenicity (Ames test) model (ISS)(version 1.0.3)

Classification model for Mutagenicity (Ames test) based on Benigni-Bossa (Istituto Superiore di Sanità) rule set

Mutagenicity (Ames test) model (SarPy-IRFMN)(version 1.0.8)

QSAR classification model for Mutagenicity (SarPy/IRFMN)

References and Documentation

Mutagenicity (Ames test) model (KNN-Read-Across)(version 1.0.1)

KNN (Read-Across) model for Mutagenicity (Ames test)

Mutagenicity (Ames test) model for aromatic amines (CONCERT/IRFMN)(version 1.0.0)

A knowledge-based expert rule system for predicting mutagenicity (AMES test) of aromatic amines.

Carcinogenicity model (CAESAR)(version 2.1.10)

QSAR classification model for Carcinogenicity (from CAESAR project)

Carcinogenicity model (ISS)(version 1.0.3)

Classification model for Carcinogenicity based on Benigni-Bossa (Istituto Superiore di Sanità) rule set

References and Documentation

Carcinogenicity model (IRFMN-ISSCAN-CGX)(version 1.0.2)

QSAR classification model for Carcinogenicity (IRFMN/ISSCAN-CGX) based on the ISSCAN-CGX dataset

Carcinogenicity model (IRFMN-Antares)(version 1.0.2)

QSAR classification model for Carcinogenicity (IRFMN/Antares) based on the Antares dataset

Carcinogenicity oral classification model (IRFMN)(version 1.0.1)

Classification model for carcinogenicity (oral route).

Carcinogenicity oral Slope Factor model (IRFMN)(version 1.0.1)

Quantitative model for the carcinogenicity (oral route) Slope Factor.

References and Documentation

Carcinogenicity inhalation classification model (IRFMN)(version 1.0.1)

Classification model for carcinogenicity (inhalation route).

Carcinogenicity inhalation Slope Factor model (IRFMN)(version 1.0.1)

Quantitative model for the carcinogenicity inhalation route) Slope Factor.

Carcinogenicity in male rat (CORAL)(version 1.0.0)

Carcinogenicity in male rat quantitative (TD50) model model (CORAL)

Carcinogenicity in female Rat (CORAL)(version 1.0.0)

Carcinogenicity in female rat quantitative (TD50) model model (CORAL)

References and Documentation

Acute Toxicity (LD50) model (KNN)(version 1.0.0)

KNN model for acute toxicity (LD50)

BCF model (CAESAR)(version 2.1.15)

QSAR regression model for fish BCF (from CAESAR project)

BCF model (Meylan)(version 1.0.4)

QSAR regression model for fish BCF (based on Meylan model)

BCF model (Arnot-Gobas)(version 1.0.1)

QSAR regression model for fish BCF (based on Arnot-Gobas model)

References and Documentation

BCF model (KNN-Read-Across)(version 1.1.1)

KNN (Read-Across) model for fish BCF