ProtoICH

EXPERT RULE-BASED PREDICTION DOSSIER

ProtoQSAR Benigni-Bosa expert-rule model for mutagenicity

Prediction for CCCCC

www.protoqsar.com

Parque Tecnológico de Valencia
Carrer de Nicolau Copèrnic 6
46980 Paterna (Valencia, Spain)

protopred@protoqsar.com

+34 962 021 811

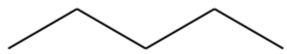
ProtoQSAR
Computational toxicology: fast, economical and ethical

Summary

Model: ProtoQSAR model for *in vitro* gene mutation study in bacteria (Ames test) (v1.0)

Human health effects: Mutagenicity. OECD 471: Bacterial reverse mutation test. Mutagenicity refers to the induction of permanent transmissible changes in the amount or structure of the genetic material of cells or organisms. The Bacterial reverse mutation test evaluates gene mutations. The test uses amino-acid requiring strains of bacteria to detect (reverse) gene mutations (point mutations and frameshifts) (see reference for "Guidance on information requirements and chemical safety assessment, Chapter R.7a" in section 9.2).

Molecule: CCCCC



Prediction: Potentially non-mutagenic, there are NO identified alerts for mutagenicity.

The presence of one or more structural alerts indicates potential for mutagenicity.

QPRF: ProtoQSAR expert-rule model for *in vitro* gene mutation study in bacteria (Ames test)

1. General information

1.1. Date of QPRF:

23-Jan-2026

1.2. QPRF author and contact details:

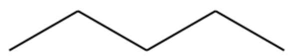
a. Authorship: ProtoQSAR

b. Address: Carrer de Nicolau Copèrnic 6

c. Phone:

d. Email: protopred@protoqsar.com

2. Substance



2.1. CAS number:

-

2.2. EC number:

-

2.3. Other regulatory numerical identifiers:

-

2.4. Chemical name:

-

2.5. Structural formula:

-

2.6. Structural and composition information:

a. SMILES: CCCCC

b. InChI: InChI=1S/C5H12/c1-3-5-4-2/h3-5H2,1-2H3

c. Other structural representation: A graphical representation above (not used in the prediction).

d. Stereochemical features: There is no stereochemical information codified in the SMILES, but the substance has stereocenters.

e. Composition information: The prediction corresponds to a single molecule, fully defined by its SMILES. It is unknown if this is the only component of the substance.

3. Model and software

3.1. Model:

a. Model or submodel name: ProtoQSAR expert-rule model for *in vitro* gene mutation study in bacteria (Ames test)

b. Model version: Version 1.0

c. Reference to QMRF: The corresponding QMRF, named ProtoQSAR model for *in vitro* gene mutation study in bacteria (Ames test) can be downloaded from https://protopred.protoqsar.com, more details can be requested to ProtoQSAR S.L., the owner of the model, by email to protopred@protoqsar.com

3.2. Software:

a. Software name: ProtoPRED® (ProtoQSAR proprietary software)

b. Software version: v1.0

c. Software reference: The software is proprietary and there is no reference

d. Software availability: The software to develop the models is proprietary and not publicly available. However, evaluators can contact protopred@protoqsar.com to get further details or to try the code.

4. Prediction

4.1. Property:

a. Predicted property: *In vitro* gene mutation in bacteria (Ames test)

b. Test guideline covered: Endpoint following the OECD: Test No. 471: Bacterial reverse mutation test.

c. Dependent variable: Original data was retrieved as a binary classification: positive (mutagenic) / negative (non-mutagenic).

4.2. Value:

a. Predicted value: Potentially non-mutagenic, there are no identified alerts for mutagenicity.

b. Predicted value (comments): The presence of one or more structural alerts indicates potential for mutagenicity.

c. Unit: N/A

ALERT	RESULT
SA1.-Acyl halides?	NO ALERT
SA2.-Alkyl (C<5) or benzyl ester of sulphonic or phosphonic acid?	NO ALERT
SA3.-N-methylol derivatives?	NO ALERT
SA4.-Monohaloalkene?	NO ALERT
SA5.-S or N mustard?	NO ALERT
SA6.-Propiolactones and propiosultones?	NO ALERT
SA7.-Epoxides and aziridines?	NO ALERT
SA8.-Aliphatic halogens?	NO ALERT

SA9.-Alkyl nitrite?	NO ALERT
SA10.-a,b unsaturated carbonyls?	NO ALERT
SA11.-Simple aldehyde?	NO ALERT
SA12.-Quinones?	NO ALERT
SA13.-Hydrazine?	NO ALERT
SA14.-Aliphatic azo and azoxy?	NO ALERT
SA15.-Isocyanate and isothiocyanate groups?	NO ALERT
SA16.-Alkyl carbamate and thiocarbamate?	NO ALERT
SA18.-Polycyclic Aromatic Hydrocarbons?	NO ALERT
SA19.-Heterocyclic Polycyclic Aromatic Hydrocarbons?	NO ALERT
SA21.-Alkyl and aryl N-nitroso groups?	NO ALERT
SA22.-Azide and triazene groups?	NO ALERT
SA23.-Aliphatic N-nitro?	NO ALERT
SA24.-a,b unsaturated alkoxy?	NO ALERT
SA25.-Aromatic nitroso group?	NO ALERT
SA26.-Aromatic ring N-oxide?	NO ALERT
SA27.-Nitro aromatic?	NO ALERT
SA28.-Primary aromatic amine, hydroxyl amine and its derived esters?	NO ALERT
SA28bis.-Aromatic mono- and dialkylamine?	NO ALERT
SA28ter.-Aromatic N-acyl amine?	NO ALERT
SA29.-Aromatic diazo with sulphonic groups?	NO ALERT
SA30.-Coumarins and Furocoumarins?	NO ALERT
SA37.-Pyrrolizidine Alkaloids?	NO ALERT
SA38.-Alkenylbenzenes?	NO ALERT
SA39.-Steroidal estrogens?	NO ALERT
SAaNNa.-Aromatic diazo?	NO ALERT
SAarNCH2.-Derived aromatic amines?	NO ALERT
QSA57_Ames.-DNA Intercalating Agents with a basic side chain?	NO ALERT
QSA58_Ames.-Haloalkene cysteine S-conjugates?	NO ALERT
QSA59_Ames.-Xanthones, Thioxanthones, Acridones?	NO ALERT
QSA60_Ames.-Flavonoids?	NO ALERT
QSA61_Ames.-Alkyl hydroperoxides?	NO ALERT
QSA62_Ames.-N-acyloxy-N -alkoxybenzamides?	NO ALERT
QSA63_Ames.-N-aryl-N-acetoxyacetamides?	NO ALERT
QSA64_Ames.-Hydroxamic acid derivatives?	NO ALERT
QSA65_Ames.-Halofuranones?	NO ALERT
QSA66_Ames.-Anthrones?	NO ALERT
QSA67_Ames.-Triphenylimidazole and related?	NO ALERT
QSA68_Ames.-9,10 - dihydrophenanthrenes?	NO ALERT
QSA69_Ames.-Fluorinated quinolines?	NO ALERT

5. Input

5.1. Structure:

a. Input structure: The prediction uses as input the SMILES of the molecule as shown in point 2.6a

b. Stereochemical features: There is no stereochemical information codified in the SMILES, so the substance is a non-stereochemical molecule or a racemic mixture.

c. Tautomerism: There is not automatic treatment of tautomers. The exact structure depicted in the SMILES is used.

5.2. Descriptors:

The model is based in a series of 48 structural alerts, represented by one or more SMARTS

a. Units: N/A

b. Experimental/calculated: Structures are based in expert judgement of experimental data.

c. Source: The structural rules have been adapted from the pubications of Benigni and Bossa.

[1] Benigni R, Bossa C, Richard A M, Yang C (2008) Ann. Ist. Super. Sanita, 44(1), 48-56

[2] Benigni R, Bossa C (2011) Chem Rev 111: 2507–2536

[3] Benigni R, Bossa C, Tcheremenskaia O (2013) Chem. Rev. 2013, 113, 5, 2940-2957

[4] Benigni R, Bossa C, Jeliazkova, N, Netzeva T, and Worth A. The Benigni / Bossa rulebase for mutagenicity and carcinogenicity - a module of Toxtree. European Commission report EUR 23241.

d. Differences with model development and validation: N/A

5.3. Model and/or software settings:

The model has been applied as described here and in the QMRF. ProtoPRED® does not have any customizable setting.

6. Applicability domain (AD) and limitations

6.1. Applicability domain (AD) and limitations:

a. AD assessment: N/A

b. AD assessment justification: N/A

c. Any other limitation: The model was built only for discrete organic chemicals. A prediction is considered to fall outside the AD if it does not match this.

7. Reliability assessment

7.1. Reproducibility:

The algorithm is unambiguous and perfectly defined by the equations encoded in the model. Additionally, the calculation of the descriptors also follows a reproducible methodology (including adequate and consistent seeds if there is any randomness). Furthermore, the model development methodology follow strict guidelines and it is properly reported and it should be also reproducible. Hence, the prediction can be reproduced using https://protopred.protoqsar.com

- Comments on reproducibility: Please, contact protopred@protoqsar.com if you need to assess a prediction with our software.

7.2. Overall performance of the model:

The model has been validated internally (by the expert rule authors) and externally (by ProtoQSAR S.L.). Accuracy values are shown in the tables below (extracted from the QMRF document).

Parameter	Training	Validation
Accuracy (ACC)	0.78	0.76
Sensitivity, recall or true positive rate (TPR)	0.85	0.84
Specificity or true negative rate (TNR)	0.72	0.66
Precision or positive predictive value (PPV)	-	0.74

7.3. Additional reliability aspects based on the training set:

a. Descriptor space: The expert-rule based models does not use descriptors.

b. Structural fragment space: The Benigni-Bossa expert-rule model only identifies a series of positives structural alerts. There is no a defined training set or list of negative structural alerts included in the model.

c. Response space: The model identifies a series of structural alerts of potential mutagenicity. Alternative responses are not possible.

d. Mechanism considerations: Mechanistic considerations have not been automatically included in the model. The user should consider if the target substance expected mechanism requires further investigation.

e. Metabolic considerations: Metabolic considerations have not been automatically included in the model. The model predicts the properties of the target substance exactly as reported above. Details on the metabolic stability of the substance and the effects of the potential metabolites should be considered separately.

7.4. Analogues:

No analogues have been identified for the expert-rule model.

7.5. Other reliable information on the property:

N/A

7.6. Conclusion on reliability:

The prediction is based in structural alerts based on Beningni Bossa expert rules. No quantitative assessment of the reliability of this particular predictions is provided.

8. Purpose of use (for regulatory applications)

8.1. Regulatory purpose:

This prediction has been performed with the purpose of following the guidelines for ICH M7 regulation.

8.2. Approach for regulatory interpretation of the prediction:

According to the ICH M7 regulation, if not experimental data on mutagenicity is available, an *in silico* approach combining two types of models: SAR models (based on expert rules) and QSAR models (statistical) can be used.

8.3. Regulatory interpretation of the result:

The Benigni-Bossa expert-rule model has been used for the prediction of mutagenicity. The compound has been predicted as Potentially non-mutagenic, there are NO identified alerts for mutagenicity..

8.4. Uncertainty:

The uncertainty of the expert-rule prediction could be estimated by reviewing the assessment elements outlined by the OECD in the prediction checklist.

8.5. Conclusion:

For ICH-M7 regulatory purposes, the negative prediction (non-mutagenic) reported above should be confirmed by a matching result from the expert rules based model. In case of contradiction, the result is considered uncertain.