

Évaluation de modèles (Q)SAR pour la prédiction de la génotoxicité d'ames : un exercice rétrospectif sur les substances chimiques enregistrées dans le cadre du règlement européen REACH

Enrico Mombelli

► **To cite this version:**

Enrico Mombelli. Évaluation de modèles (Q)SAR pour la prédiction de la génotoxicité d'ames : un exercice rétrospectif sur les substances chimiques enregistrées dans le cadre du règlement européen REACH. Rapport Scientifique INERIS, 2015, 2014-2015, pp.51-52. ineris-01869537

HAL Id: ineris-01869537

<https://hal-ineris.archives-ouvertes.fr/ineris-01869537>

Submitted on 6 Sep 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



ÉVALUATION DE MODÈLES (Q)SAR POUR LA PRÉDICTION DE LA GÉNOTOXICITÉ D'AMES: un exercice rétrospectif sur les substances chimiques enregistrées dans le cadre du règlement européen REACH

Certaines substances chimiques peuvent introduire des mutations dans le génome humain et provoquer des effets graves sur la santé tels que le cancer [1]. Pour ces raisons, la détection du potentiel mutagène des substances chimiques revêt une importance fondamentale dans le cadre du règlement européen REACH sur la toxicité des substances chimiques. Une des approches les plus utilisées pour l'identification de substances mutagènes est le test d'Ames [2]. Il permet d'évaluer les effets des substances chimiques sur l'expression des gènes pour la synthèse de l'histidine de plusieurs souches bactériennes de *S. typhimurium* et *E. coli*.

Le règlement REACH prévoit la possibilité d'utiliser les prédictions élaborées par des modèles (Q)SAR¹ afin de caractériser le potentiel de mutagenèse des substances chimiques. Cependant, et malgré la disponibilité de plusieurs outils (Q)SAR gratuits, les modèles (Q)SAR qui prédisent les effets de mutagenèse n'ont été utilisés que de façon exceptionnelle pour l'échéance du 30 novembre 2010 du règlement REACH [3].

L'objectif de ce travail était donc d'évaluer la performance prédictive des outils (Q)SAR

mis à disposition gratuitement sur internet, avec les substances qui ont été enregistrées pour cette échéance.

¹ Relations (quantitatives) entre structure moléculaire et activité biologique

Méthodes

Le [Tableau 1](#) présente les modèles (Q)SAR qui ont été analysés dans le cadre de ce travail. Le modèle « consensus-VEGA » (CNS-VEGA) intègre les prédictions des trois modèles de la plateforme VEGA selon le degré d'appartenance de la substance d'intérêt aux domaines d'applications de ces modèles. La performance prédictive des modèles a été estimée grâce à 4 indicateurs statistiques: concordance, sensibilité, spécificité (statistiques de Cooper) et coefficient de corrélation de Matthews (MCC). L'introduction du dernier indicateur a été nécessaire car les statistiques de Cooper se prêtent mal à la description de jeux de données déséquilibrés comme ceux analysés dans le cadre de ce travail (majorité de substances négatives). Le MCC varie entre -1 (désaccord complet entre toxicité expérimentale et prédite) et +1 (identité parfaite entre toxicité expérimentale et prédite). Tous ces indicateurs ont été calculés en fonction de substances

Notwithstanding the possibility of adopting (Q)SAR predictions for registration purposes, registrants providing information for the first REACH deadline (November 30, 2010) submitted (Q)SAR predictions only in an exceedingly small number of cases [3]. This observation prompted the interest of the project CALEIDOS that decided to provide a retrospective exercise on the reliability of freely available (Q)SAR models predicting Ames mutagenicity when applied to chemicals registered for this first deadline [8]. Our analysis showed that, with the only exception of one (Q)SAR tool (TEST), all the analyzed models were characterized by accuracies that were comparable to the experimental reliability of the Ames test. The best performance was displayed by the Benigni-Bossa rule-based system as implemented within the online VEGA platform (accuracy = 92 %, sensitivity = 83 %, specificity = 93 %, Matthews correlation coefficient = 0.68). The main conclusion is that our results support the accumulating evidence that the mechanistic relationship between electrophilicity and mutagenicity is properly described by existing (Q)SAR models.

Tableau 1
Modèles (Q)SAR analysés lors de l'exercice d'évaluation

Logiciel	Nom du modèle	Acronyme	Typologie
T.E.S.T. v4.1	Ames mutagenicity	T.E.S.T. MUTA	QSAR
TOXTREE v2.6.0	In vitro mutagenicity (Ames test) alerts by ISS	TTIDEA	SAR
VEGA v1.0.8	Mutagenicity model (CAESAR) v.2.1.12	CAESAR	QSAR
VEGA v1.0.8	Mutagenicity SarPy model v1.0.6-DEV	SARPY	QSAR
VEGA v1.0.8	Benigni-Bossa Mutagenicity (TOXTREE) v1.0.1-DEV	TTVEGA	SAR
VEGA v1.0.8	Consensus-VEGA	CNS-VEGA	QSAR/SAR





dont les données toxicologiques satisfaisaient plusieurs critères de qualité (ex. conformité avec la ligne directrice 471 de l'OCDE).

Lorsqu'on caractérise la performance prédictive d'un modèle (Q)SAR il est important de la comparer avec la reproductibilité des données expérimentales qui ont été utilisées pour le calibrer. La reproductibilité inter-laboratoire du test d'Ames est estimée à 85-90 % [4; 5] et ce pourcentage représente la limite supérieure de fiabilité qu'il est raisonnable de pouvoir obtenir en fonction d'une application systématique d'outils QSAR.

➤ Résultats

La [Figure 1](#) décrit une validation « externe » des modèles (Q)SAR retenus pour l'exercice. L'adjectif « externe » indique que les molécules dont la toxicité a été prédite par les modèles satisfont deux conditions: a) elles rentrent dans les domaines d'application des modèles et b) elles ne font pas partie des jeux d'apprentissage des modèles. Cette forme d'évaluation représente le critère de qualification le plus strict de la performance prédictive d'un modèle.

Nos résultats indiquent que le modèle individuel le plus performant est TTVEGA suivi par SARPY et CAESAR, caractérisés par des performances similaires. Un modèle « consensus », intégrant les prédictions de ces trois modèles, donne lieu à une performance prédictive améliorée aux dépens d'une réduction du domaine d'application ([Figure 1](#)).

La très bonne performance prédictive de TTVEGA peut être expliquée en fonction des 30 alertes structurelles sur lesquelles repose le modèle [6]. Ces alertes ont été sélectionnées après une analyse critique des données structure-génotoxicité disponibles dans la littérature.

➤ Conclusions

L'ensemble des résultats indique que les mécanismes de réactivité moléculaire qui déterminent la génotoxicité des substances chimiques sont bien pris en compte par plusieurs modèles (Q)SAR existants. Ces outils peuvent ainsi jouer un rôle important lors de prises de décisions sur la génotoxicité dans un contexte réglementaire. Il est intéressant d'observer que les statistiques de Cooper qui décrivent TTVEGA et CNS-VEGA sont comparables avec celles d'autres tests expérimentaux. Par exemple, un test réglementaire validé comme le test LLNA (essai de stimulation locale des ganglions lymphatiques) pour la sensibilisation cutanée est caractérisé par une sensibilité, une spécificité et une concordance égales à 72 %, 67 % et 72 % respectivement [7].

📖 Références

- [1] Alberts, B., et al. (2002) *The Preventable Causes of Cancer. Molecular Biology of the Cell*. Garland Science, New York.
- [2] Ames, B.N., McCann, J. and Yamasaki, E. (1975) *Methods for detecting carcinogens and mutagens with the Salmonella/mammalian-microsome*

mutagenicity test, Mutat Res, 31, 347-364.

[3] ECHA (2011) *The Use of Alternatives to Testing on Animals for the REACH Regulation*.

[4] Sushko, I., et al. (2010) *Applicability domains for classification problems: Benchmarking of distance to models for Ames mutagenicity set*, J Chem Inf Model, 50, 2094-2111.

[5] Sushko, I., et al. (2010) *Applicability domain for in silico models to achieve accuracy of experimental measurements*, J. Chemometrics 24, 202-208.

[6] Benigni, R., et al. (2008) *Benigni/Bossa rulebase for mutagenicity and carcinogenicity-A module of Toxtree*. JRC Scientific and Technical Reports.

[7] ICCVAM (1999) *The Murine Local Lymph Node Assay: A Test Method for Assessing the Allergic Contact Dermatitis Potential of Chemicals/Compounds. The Results of an Independent Peer Review Évaluation Coordinated by the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) and the National Toxicology Program Center for the Évaluation of Alternative Toxicological Methods*.

[8] Cassano, A., et al. (2014) *Évaluation of QSAR models for the prediction of ames genotoxicity: a retrospective exercise on the chemical substances registered under the EU REACH regulation, Journal of environmental science and health. Part C, Environmental carcinogenesis & ecotoxicology reviews*, 32, 273-298.

[Figure 1](#)

Performance prédictive des modèles analysés en fonction de substances chimiques qui rentrent dans leurs domaines d'application mais qui ne font pas partie des jeux d'apprentissage respectifs. Le nombre entre parenthèses indique le coefficient de corrélation de Matthews. Le nombre sur les histogrammes indique le nombre de substances évaluées.

