Les outils QSAR dans un contexte réglementaire : l'importance d'une validation scientifique externe

Enrico Mombelli

▶ **To cite this version:**

Enrico Mombelli. Les outils QSAR dans un contexte réglementaire : l'importance d'une validation scientifique externe. Rapport Scientifique INERIS, 2011, 2010-2011, pp.28-30. ineris-01869383

HAL Id: ineris-01869383
https://ineris.hal.science/ineris-01869383

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.

Contributeur

Enrico Mombelli

Les outils QSAR dans un contexte réglementaire

L'importance d'une validation scientifique externe

La mutagenèse et la carcinogenèse sont deux effets toxicologiques qui font partie des informations toxicologiques à fournir dans le cadre de plusieurs règlements sur les substances chimiques et notamment du règlement REACH [1].

Les effets mutagènes sont habituellement caractérisés de façon relativement économique et rapide par le biais du test d'Ames [2] qui quantifie le potentiel mutagène d'une substance grâce à une quantification du taux de révertants de souches de *S. typhimurium* auxotrophes pour l'histidine.

En revanche, la détermination expérimentale du potentiel cancérogène des produits chimiques est particulièrement onéreuse. Le coût d'une étude de cancérogenèse chez le rongeur peut aller jusqu'à un million de dollars par substance [3]. Ce coût élevé, ajouté à la préoccupation croissante pour le bien-être des animaux, encourage, en particulier, la mise au point de modèles QSAR pour la prédiction du potentiel cancérogène des produits chimiques.

Dans ce contexte, le système « Benigni/Bossa » pour l'analyse des propriétés carcinogènes et mutagènes [4] joue un rôle important en raison de son exhaustivité et de son intégration dans deux outils disponibles gratuitement : la boîte à outils QSAR de l'OCDE (OCDE, 2010) (figure 1) et Toxtree [5] (figure 2).

Le travail effectué en collaboration entre l'INERIS et le CTIS (Centre de traitement de l'information scientifique) se proposait d'analyser la validité du système « Benigni/Bossa » en tant que tel et de son implémentation dans ces deux outils (boîte à outils QSAR de l'OCDE versions 1.1.01 et 1.1.02 et Toxtree version 1.6).

Le système « Benigni/Bossa », pour l'analyse du potentiel mutagène et cancérigène d'une substance chimique, ne nécessite comme donnée d'entrée que la structure moléculaire (bidimensionnelle) de cette substance. Dans ce système, quatre cas de figures peuvent se présenter : (a) aucune alerte structurelle pour la carcinogénèse n'est reconnue ; (b) une ou plusieurs alertes structurelles pour la carcinogenèse sont reconnues ; (c) le recours à deux modèles QSAR pour la carcinogenèse et la mutagenèse est déclenché si la substance appartient à la classe des amines aromatiques homocycliques ; (d) le recours à un modèle QSAR pour la prédiction du potentiel mutagène est déclenché si le système détecte la présence d'un aldéhyde α, β insaturé à l'intérieur de la structure moléculaire. Les prédictions des trois modèles du système « Benigny/Bossa » (*voir encadré p. 30*) sont dichotomiques (carcinogène/non-carcinogène et mutagène/non mutagène).

Trente-cinq alertes structurelles (ex. azotures, coumarines, groupements nitroso…) composent le système « Benigni/Bossa » pour la reconnaissance de motifs structuraux qui peuvent induire des effets carcinogènes chez le rongeur, cinq d'entre elles décrivant des mécanismes non-génotoxiques [4]. Les auteurs du système ont basé leur sélection d'alertes sur une analyse approfondie

Références

[1] *UE, Regulation (EC) No 1907/2006 of the European Parliament and of the Council of 18 December 2006 concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals REACH), establishing a European Chemicals Agency, amending Directive 1999/45/EC and repealing Council Regulation (EEC) No 793/93 and Commission Regulation (EC) No 1488/94 as well as Council Directive 76/769/EEC and Commission Directives 91/155/EEC, 93/67/EEC, 93/105/EC and 2000/21/EC*, The European Parliament and the Council of the European Union, 2006 disponible à l'adresse http://eur-lex.europa.eu/LexUriServ/site/en/oj/2006/l_396/l_39620061230en00010849.pdf

[2] OCDE, *Bacterial Reverse Mutuation Test*, Ref. OECD TG 471, 1997, disponible l'adresse http://www.oecd.org/dataoecd/18/31/1948418.pdf

[3] *ECVAM, an overview – European Centre for the Validation of Alternative Methods* (2005), disponible à l'adresse http://ec.europa.eu/dgs/jrc/downloads/jrc_press_animal_ecvam_overview.pdf

[4] Benigni R., Bossa C., Jeliazkova N., Netzeva T., Worth, A. *The Benigni/Bossa rulebase for mutagenicity and carcinogenicity – a module of Toxtree*, disponible (2008) http://ecb.jrc.ec.europa.eu/DOCUMENTS/QSAR/EUR23241_EN.pdf

Figure 1
Interface graphique de la boîte à outils QSAR de l'OCDE (v1.1.01 et v1.1.02) après l'analyse de l'aniline par le système « Benigni/Bossa ».

Figure 2
Interface graphique de Toxtree v 1.6 après l'analyse de l'aniline par le système « Benigni/Bossa ».

LE COLLOQUE MODÉLISATION

Le 1er avril 2011, s'est tenu, à Paris, un colloque intitulé « Modélisation en (éco)toxicologie : vers l'âge de raison ? ». Organisé par l'INERIS, il a regroupé 80 personnes, principalement des chercheurs, des gestionnaires publics et des industriels. L'expansion des méthodes rend nécessaire un retour d'expérience et une réflexion prospective : comment répondre aux enjeux scientifiques tout en soutenant la prise de décision ? Faut-il envisager un référentiel commun pour accompagner le développement et l'utilisation des modèles ? Si oui, comment doit-on le structurer ?
Deux grandes conclusions se détachent :
• l'utilité des modèles en évaluation du risque est désormais largement reconnue, en termes d'extrapolation des résultats depuis l'expérimentation de laboratoire jusqu'aux cibles d'intérêt pour l'évaluation du risque (l'homme et les écosystèmes) et en termes de guide dans des situations complexes (regroupements ou hiérarchisation de substances, par exemple) ;
• il s'agit maintenant d'acquérir la confiance des utilisateurs. Or, tous les intervenants s'accordent à demander à la fois une plus grande transparence des outils, des critères de validité pour chaque modèle et la mise en place systématique d'une validation externe. Côté utilisateurs, il s'agit pour les agences et les industriels de renforcer à la fois la formation en modélisation mathématique et le recrutement de modélisateurs.

Références

[5] OCDE, *QSAR Application Toolbox*, 2010, disponible à l'adresse http://www.oecd.org/document/54/0,3343,en_2649_34379_42923638_1_1_1_1,00.html

[6] Ex-ECB, Toxtree, 2010, disponible à l'adresse http://ecb.jrc.it/qsar/qsar-tools/index.php?c=TOXTREE

[7] Devillers J., Mombelli E., *Evaluation of the OECD QSAR Application Toolbox and Toxtree for estimating the mutagenicity of chemicals*, Part 1, *Aromatic amines SAR and QSAR in Environmental Research, 21* (2010a), pp. 753-769.

[8] Devillers J., Mombelli E., *Evaluation of the OECD QSAR Application Toolbox and Toxtree for estimating the mutagenicity of chemicals*, Part 2. *a-b unsaturated aliphatic aldehydes SAR and QSAR in Environmental Research, 21* (2010b), pp. 771-783.

[9] Mombelli E., Devillers J., *Evaluation of the OECD QSAR Application Toolbox and Toxtree for predicting and profiling the carcinogenic potential of chemicals SAR and QSAR in Environmental Research*, 21 (2010), pp. 731-752.

[10] Kaiser K. L. E., Dearden J. C., Klein W., Schultz T. W., *A note of caution to users of ECOSAR*, Water Qual. Res. J. Can., 34 (1999), pp. 179-182.

[11] Reuschenbach P., Silvani M., Dammann M., Warnecke D., Knacker T., *ECOSAR model performance with a large test set of industrial chemicals*, CHEMOSPHERE, 71 (2008), pp. 1986-1995.

d'études scientifiques qui couvrent près de trois décennies de recherche en la matière.
Ainsi, le système « Benigni-Bossa » contient trois modèles QSAR [4] *(voir encadré p. 30)*.
Les principaux résultats de notre étude de validation du système « Benigni/Bossa », qui a généré trois publications scientifiques [7, 8, 9] sont reportés dans les paragraphes qui suivent.

S'agissant du système « Benigni/Bossa » en lui-même :

– le système de reconnaissance d'alertes structurelles pour la carcinogenèse a une bonne fiabilité globale mais peut présenter des limites lors de l'analyse de classes chimiques spécifiques (ex. hydrocarbures aromatiques polycycliques) ;

– les performances (sensibilité, spécificité et concordance) du modèle QSAR 6 pour la prédiction du potentiel mutagène des amines aromatiques sont en réalité inférieures à celles déclarées par le mode d'emploi des deux outils ;

– la sensibilité réelle du modèle QSAR 8 pour la prédiction du potentiel carcinogène des amines aromatiques est inférieure à celle qui est déclarée par le mode d'emploi des deux outils ;

– les paramètres du modèle QSAR 13 pour la prédiction du potentiel mutagène des aldéhydes α,β insaturés ont été estimés sur la base de vingt substances uniquement dont trois seulement sont non-mutagènes. Pour cette raison, la performance du modèle est très modeste (incapacité du modèle à prédire la toxicité de molécules qui ne font pas partie du jeu d'apprentissage) et son utilité pratique est, par conséquent, fortement limitée ;

– de façon surprenante, les prédictions pour les molécules qui composent les jeux d'apprentissage des trois modèles ne sont pas toujours identiques à celles rapportées par les publications originales.

S'agissant de l'implémentation du système « Benigni/Bossa » dans les logiciels :

– la reconnaissance des alertes structurelles est souvent perturbée par des problèmes relatifs à l'interface graphique de la boîte à outils QSAR de l'OCDE qui n'arrive pas toujours à bien représenter l'aromaticité des molécules et/ou à bien gérer la structure des groupements nitro lors de l'application des modèles QSAR. Par rapport à ces dysfonctionnements, l'outil de l'OCDE manque de transparence et seul un utilisateur averti peut se rendre compte de tous les problèmes ;

– l'outil Toxtree présente une transparence supérieure à l'outil OCDE ;

– les équations qui sont censées corriger les valeurs des descripteurs calculées par les outils afin de reproduire les valeurs numériques →

→ rapportées par les publications originales ne remplissent pas leur fonction.
Plusieurs idées pour une amélioration des outils évalués ont été proposées dans nos articles (nouvelles alertes putatives, jeux d'évaluation élargis, corrections d'erreurs d'implémentation conceptuelle et informatique).
Les résultats de l'étude de validation montrent donc que l'intégration informatique d'un système QSAR peut provoquer plusieurs dysfonctionnements imprévisibles et qu'une évaluation de la performance opérationnelle d'une telle implémentation doit être considérée comme une priorité avant d'entreprendre toute sorte d'évaluation toxicologique *in silico*. À titre de comparaison avec les outils que nous avons étudiés, les outils QSAR les plus fiables (ex. ECOSAR de l'EPA américaine) ont été évalués à plusieurs reprises par des experts qui ne faisaient pas partie du groupe de concepteurs du modèle [10, 11]. Cette validation externe est incontournable pour améliorer les outils existants et pour leur utilisation fiable de la part de toutes les parties prenantes d'un règlement sur la toxicité des substances chimiques.

ABSTRACT

Toxicological effects and chemical structure can be related quantitatively or qualitatively through QSAR modeling (Quantitative structure-activity relationship). During the past decade, the development of QSAR tools has continued to progress especially because of the requirements of new European legislations (REACH and 7th amendment to the EU Cosmetics Directive) that encourage and support the implementation of alternative methods to animal testing including QSAR models.
In order to facilitate the regulatory use of QSAR tools, the management of the environment directorate of the OECD and the former ECB (European Chemicals Bureau) have made available two free software tools (the "QSAR Toolbox" and "Toxtree"). Both tools allow the user to use several QSAR models and in particular the "Benigni/Bossa" rule-based system for predicting the mutagenic and carcinogenic potential of chemicals on the sole basis of chemical structure.
Given the regulatory role of the OECD QSAR Toolbox and Toxtree, the INERIS and the CTIS jointly carried out a scientific validation of the QSAR models that form the "Benigni-Bossa" rule-based system.
The main results of this study show that several conceptual and software related problems can, in some specific cases, give raise to wrong predictions.

MODÈLES QSAR

1_ Le modèle « QSAR 6 » qui discrimine amines aromatiques mutagènes et non mutagènes grâce à une fonction discriminante monodimensionnelle :

$w = -3{,}14\ HOMO + 1{,}76\ LUMO + 0{,}62\ MR_2 + 0{,}75\ MR_3 + 1{,}88\ MR_6 + 3{,}75\ Idist$

w(moyenne,Class1) = 28,42
w(moyenne,Class2) = 26,44
Seuil = 27,43

Ce modèle a été dérivé pour un jeu d'apprentissage composé de 111 amines aromatiques (64 positives au test d'Ames pour la mutagenèse et 47 négatives)

2_Le modèle « QSAR 8 » qui discrimine amines carcinogènes et non carcinogènes grâce à une fonction discriminante monodimensionnelle :

$w = -3{,}79\ L(R) + 3{,}52\ B_5(R) - 4{,}12\ HOMO + 4{,}41\ LUMO + 3{,}09\ MR_3 + 2{,}60\ MR_5 + 4{,}63\ MR_6 - 3{,}49\ I(An) + 1{,}80\ I(NO_2) - 1{,}78\ I(BiBr)$

w(moyenne,Class1) = 27,82
w(moyenne,Class2) = 30,34
Seuil = 29,08

Ce modèle a été dérivé pour un jeu d'apprentissage composé de 64 amines aromatiques (52 carcinogènes et 12 non carcinogènes.

3_Le modèle « QSAR 13 » qui discrimine aldéhydes α,β insaturés mutagènes et non mutagènes.

$w = 0.387\ MR - 3{,}12\ log\ P + 3{,}23\ LUMO$

w(moyenne,class1) = 9.69
w(moyenne,class2) = 6.37
Seuil = 8.03

Ces modèles se basent sur les descripteurs suivants : les énergies orbitalaires HOMO et LUMO, les réfractivités moléculaires MR (l'index indique la position du substituant), une variable indicatrice Idist qui indique la présence/absence de substituants (l'index indique la position du substituant), les paramètres stériques L(R) (longueur) et B5(R) (largeur maximale) pour les substituants des amines aromatiques, les variables indicatrices I(An), $I(NO_2)$ et I(BiBr) qui indiquent la présence ou absence d'anilines, groupements nitro et « ponts » entre deux phényles, le coefficient de partage log P entre l'n-octanol et l'eau.