HAL
open science

Prédiction des propriétés physico-chimiques dangereuses des substances chimiques : vers l'application des modèles QSPR dans un contexte réglementaire

Guillaume Fayet, Patricia Rotureau

To cite this version:

Guillaume Fayet, Patricia Rotureau. Prédiction des propriétés physico-chimiques dangereuses des substances chimiques : vers l'application des modèles QSPR dans un contexte réglementaire. Rapport Scientifique INERIS, 2012, 2011-2012, pp.32-34. ineris-01869418

HAL Id: ineris-01869418

https://ineris.hal.science/ineris-01869418v1

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.

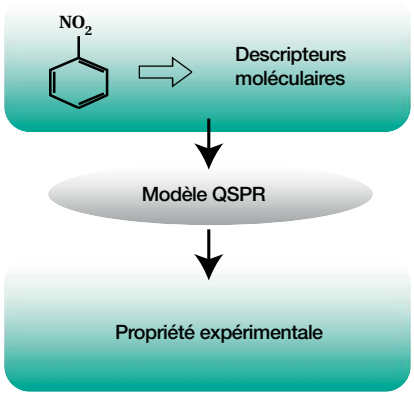

Figure 1
Principe d'une méthode QSPR.

Contributeurs

Guillaume Fayet

Patricia Rotureau

PRÉDICTION DES PROPRIÉTÉS PHYSICO-CHIMIQUES DANGEREUSES DES SUBSTANCES CHIMIQUES

Vers l'application des modèles QSPR dans un contexte réglementaire

L'INERIS est reconnu au plan international dans le domaine de l'évaluation de l'explosibilité et de l'inflammabilité des substances chimiques (risques d'explosion et d'incendie). Dans ce domaine, l'expertise repose essentiellement sur la réalisation d'essais selon des référentiels réglementaires.

Or, l'approche par modélisation moléculaire a connu ces dernières années un essor important lié à l'accès à des ordinateurs toujours plus puissants. Ainsi, elle peut dans certains cas se substituer à l'expérience ou aider à identifier les substances les plus à risque qu'il est indispensable de tester. Aussi, depuis plusieurs années, l'INERIS travaille au développement de méthodes prédictives telles que les méthodes QSAR/QSPR *(Quantitative Structure-Activity/Property Relationship)* pour l'évaluation des propriétés physico-chimiques dangereuses des substances chimiques [A]. Ces méthodes sont très développées et très utilisées dans la recherche pharmaceutique ou biologique ainsi que pour l'évaluation des risques toxicologiques et pour l'environnement (écotoxicité, etc.).

Dès 2007, le programme de recherche REPLACE (Recherche sur les propriétés et l'activité de composés explosifs, 2007-2010)

Références

Ⓐ Rotureau P., Fayet G., Marlair, G., Michot C., Joubert L., Adamo C. *Évaluation de l'explosibilité des substances chimiques. Des approches expérimentales classiques à la prédiction par la chimie quantique et les méthodes statistiques QSPR.* Actualité Chimique, 2010, 337, p. 51.

Ⓑ Fayet G., Rotureau P., Joubert L., Adamo C., *Development of a QSPR model for predicting thermal stabilities of nitroaromatic compounds taking into account their decomposition mechanism.* J. Mol Model, 2011, 17, p. 2443.

Ⓒ Fayet G., Del Rio A., Rotureau P., Joubert L., Adamo C., *Predicting the thermal stability of nitroaromatic compounds using chemoinformatic tools.* Mol Inf, 2011, 30, p. 623.

Ⓓ *Gaussian 09.* Gaussian Inc., Pittsburg, PA, USA, 2009.

Ⓔ *Guidance document on the validation of (quantitative) structure-activity relationships [(Q)SAR] models.* OECD, 2007.

Ⓕ Papa E., Kovarich S., Gramatica P. *Development, validation and inspection of the applicability domain of QSPR models for physicochemical properties of polybrominated diphenyl ethers.* QSAR Comb Sci, 2009, 28, p. 790.

Ⓖ Dearden J. C., Cronin M. T. D., Kaiser K. L. E. *How to not develop a quantitative structure-activity or structure-property relationship (QSAR/QSPR).* SAR QSAR Environ Res, 2009, 20, p. 241.

Ⓗ Fayet G., Rotureau P., Prana V., Adamo C. *Global and local QSPR models to predict the impact sensitivity of nitro compounds.* Process Saf Prog, accepté.

Ⓘ Kamlet M.J. *The relationship of impact sensitivity with structure of organic high explosives. 1. Polynitroaliphatic explosives.* 6th Symposium International on Detonation, Coronado, CA, 1976, p. 69.

Ⓙ Wang R., Jiang J., Pan, Y. Kao H., Cui Y. *Prediction of impact sensitivity of nitro energetic compounds by neural network based on electrotopological-state indices.* J Hazard Mater, 2009, 166, p. 155.

Ⓚ Prana V., Fayet G., Rotureau P., Adamo C. *Development of validated QSPR models for impact sensitivity of nitroaliphatic compounds.* J Hazard Mater, accepté.

Ⓛ QSAR Toolbox: www.qsartoolbox.org.

Ⓜ (Q)SAR *Model Reporting Format Inventory:* qsardb.jrc.it/qmrf/.

Ⓝ Projet PREDIMOL: www.ineris.fr/predimol

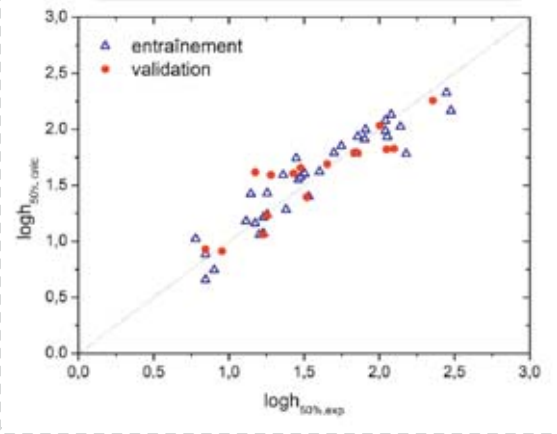

Figure 2
Corrélation entre les valeurs de sensibilité à l'impact expérimentales et calculées par le modèle (Eq. 1).

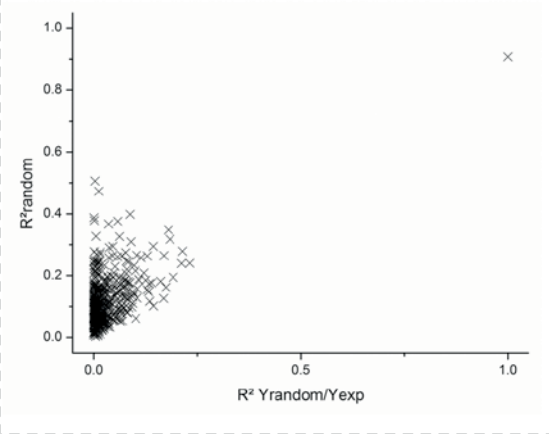

Figure 3
Résultats du test de Y-scrambling représentant la valeur des coefficients R²random, en fonction de la corrélation entre les valeurs Y expérimentales et redistribuées aléatoirement (R²Yrandom/Yexp).

a initié le développement de premiers modèles QSPR fiables pour la chaleur de décomposition des composés nitroaromatiques [B] [C]. Fort de ce succès, le développement de modèles QSPR est désormais poursuivi pour d'autres propriétés physico-chimiques dangereuses (notamment pour la réglementation REACh) et d'autres familles de composés chimiques selon une méthode robuste permettant d'envisager leur utilisation dans un contexte réglementaire.

Relations quantitatives structures-propriétés (QSPR)

Les méthodes QSPR consistent à relier de manière quantitative une propriété expérimentale à la structure moléculaire d'une substance, comme illustré en figure 1. Afin de décrire cette structure moléculaire, des descripteurs de plusieurs types peuvent être employés : constitutionnels (par exemple, nombre de groupements, atomes), topologiques (basés sur la connectivité des atomes dans la molécule), géométriques (caractérisant la géométrie 3D de la molécule) ou quantiques (regroupant les informations énergétiques, électroniques et de réactivité). Ces derniers peuvent être obtenus par des méthodes de chimie quantique telle que la Théorie de la fonctionnelle de la densité (DFT) à l'aide du logiciel Gaussian 09 [D].

Il s'agit de mettre en place une équation mathématique (ou un algorithme) entre la propriété macroscopique à prédire et une série de descripteurs à l'échelle moléculaire. Pour ce faire, différents outils de traitement de données peuvent être employés : algorithmes génétiques, réseaux de neurones ou régressions multilinéaires.

Pour utiliser les données prédites par ces modèles QSAR/QSPR à des fins réglementaires, l'OCDE a mis en place en 2004 une série de cinq principes de validation auxquels doivent répondre les modèles [E] :

- **une définition précise de la propriété prédite** par le modèle, incluant le protocole et les conditions expérimentales ;
- **une équation mathématique (ou un algorithme) sans équivoque (reproductible),** incluant la définition des différents paramètres employés ainsi que les méthodes de calcul éventuellement utilisées pour les obtenir ;
- **un domaine d'applicabilité défini,** permettant de déterminer pour quelles molécules les prédictions sont fiables ;
- **des mesures appropriées des performances** du modèle en termes de corrélation et de prédiction, incluant donc la mesure de son pouvoir prédictif pour un jeu de molécules de validation ;
- si possible, **une interprétation des mécanismes moléculaires** mis en jeu au travers des descripteurs employés et de la structure du modèle.

Les performances des modèles sont évaluées sur la base de différents critères statistiques [F] [G]. La qualité d'ajustement du modèle est vérifiée par la corrélation entre les valeurs calculées et expérimentales de son jeu d'entraînement. La validation interne du modèle est réalisée à l'aide d'une méthode de validation croisée *(cross-validation method)* qui permet d'évaluer sa robustesse. Une méthode dite de « Y-randomisation » est également réalisée afin de s'assurer que le modèle n'a pas été obtenu par chance. Enfin, l'évaluation de son pouvoir prédictif est réalisée sur un jeu de données externes (jeu de validation), non employées pour le développement du modèle.

Prédiction de la sensibilité à l'impact des nitroaliphatiques

Dans le cadre des travaux menés par l'INERIS, des modèles QSPR ont été développés pour prédire la sensibilité à l'impact de différentes substances explosives nitrées (nitroaromatiques, nitramines et nitroaliphatiques) avec de bonnes performances [H].

La sensibilité à l'impact ($h_{50\,\%}$) est une propriété importante car elle caractérise la tendance du matériau à engendrer une décomposition (potentiellement explosive) sous l'effet d'un impact. Pour cette propriété, des modèles QSPR ont déjà été développés par le passé, mais aucun ne répond à l'intégralité des principes OCDE (pas de validation externe dans les modèles de Kamlet [I], une mauvaise robustesse et un domaine d'applicabilité non défini pour ceux de Wang [J]).

Des modèles multilinéaires ont donc été développés à l'INERIS pour prédire la sensibilité à l'impact de nitroaliphatiques (à partir d'une base de données de 50 composés) en ➡

➡ répondant à tous les principes OCDE [K]. À titre d'exemple, en considérant 66 descripteurs constitutionnels, un modèle quantitatif représenté par une équation multilinéaire à trois descripteurs (Eq. 1) a été obtenu :

$$\log h_{50\%} = -2{,}53\, n_N/n_{atom} + 0{,}07\, n_{single} - 0{,}25\, n_{NO2} + 1{,}94$$

où n_N/n_{atom} est le nombre relatif d'atomes d'azote, n_{single} le nombre de liaisons simples, et n_{NO2} le nombre de groupements NO_2. D'un point de vue chimique, la présence du groupement NO_2 était attendue, puisqu'il est connu que le mécanisme principal impliqué dans la décomposition des composés nitroaliphatiques est la dissociation de la liaison C-NO_2.

Ce modèle est caractérisé par une bonne corrélation (R^2=0,88), comme le montre la figure 2. Il est robuste (Q^2_{LOO}=0,85) et présente un pouvoir prédictif important (R^2_{ext}=0,81), en particulier dans son domaine d'applicabilité (R^2_{in}=0,78). La méthode de Y-scrambling a également validé le modèle, puisque les coefficients R^2 obtenus après redistribution aléatoire des valeurs expérimentales sont faibles (avec une valeur moyenne R^2_{YS}=0,09) comme indiqué sur la figure 3.

Ce modèle permet donc de prédire la valeur de la sensibilité à l'impact de composés nitroaliphatiques par la seule connaissance de trois descripteurs constitutionnels, qui sont facilement calculés à partir de la structure 2D des molécules.

Ce modèle QSPR est, à notre connaissance, le premier modèle répondant aux exigences d'une utilisation dans un cadre réglementaire (cinq principes OCDE) pour prédire la sensibilité à l'impact de composés nitroaliphatiques. En effet, il a été validé avec des méthodes de validation interne (pour caractériser la robustesse) et externe (caractérisant la prédictivité) et son domaine d'applicabilité a été clairement défini.

Mise à disposition et acceptabilité des modèles QSPR

Une fois validés, les modèles peuvent être utilisés pour prédire la propriété étudiée pour des substances chimiques de structure moléculaire « proche ». Ils peuvent aider notamment à la prédiction de la propriété pour des substances chimiques encore non synthétisées (optimisation de procédés ou substitution de substances, par exemple). Dans certains cas, les modèles peuvent également permettre la compréhension des mécanismes mis en jeu dans la propriété étudiée.

L'utilisation des modèles QSPR pour prédire les propriétés des substances chimiques est même recommandée dans le cadre de la nouvelle réglementation REACh pour peu qu'ils répondent aux exigences des principes OCDE. Des modèles validés sont ainsi disponibles dans :

- la QSAR Toolbox [L] (de l'OCDE et de l'ECHA), une boîte-outil accessible gratuitement, connue et utilisée largement par les industriels et les instances réglementaires. Elle est principalement dédiée à la toxicité des substances, qui regroupe différentes méthodes alternatives (*read across* et modèles QSAR/QSPR). Néanmoins, la version 2.1 propose un module dédié à la prédiction des propriétés physico-chimiques (y compris celles d'explosibilité, comme montré en figure 4), même si aucun modèle lié aux propriétés physico-chimiques dangereuses n'y figure jusqu'à présent.
- une base de modèles QSAR/QSPR mise en place par le JRC (*Joint Research Center*) [M]. Les modèles intégrés y sont validés selon un certain nombre de critères définis dans un fichier QMRF (QSAR *Model Reporting Format*), fichier très complet assurant la transparence des modèles développés et leurs performances (nécessaire également pour l'intégration des modèles dans la QSAR Toolbox).

Les démarches nécessaires à l'acceptation des nouveaux modèles QSPR développés par l'INERIS dans la QSAR Toolbox ont été initiées et le modèle précédemment présenté (Eq. 1) a reçu un avis favorable pour une intégration prochaine au sein de l'outil.

Le succès de ces premiers travaux a mené à la mise en place du projet ANR PREDIMOL [N] (PREDIction des propriétés physico-chimiques des produits par modélisation MOLéculaire) piloté par l'INERIS pour la période 2010-2013 au sein duquel de nouveaux modèles pour la prédiction des propriétés physico-chimiques requises par REACh sont en développement pour les peroxydes organiques et les amines. ●

Collaboration

Chimie ParisTech – Équipe de modélisation des systèmes complexes.

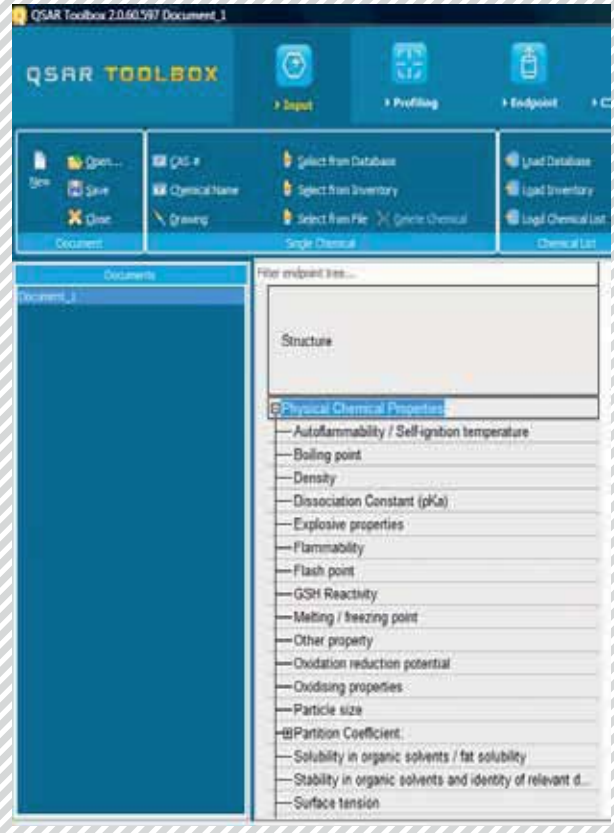

Figure 4
Aperçu de la QSAR Toolbox de l'OCDE et de l'ECHA.

Abstract

New theoretical methodologies were developed to access robust QSPR models to predict the properties of explosibility of chemicals. Several models dedicated to the prediction of explosibility properties (thermal stability and sensitivity to impact) of nitro compounds were derived following the requirements of OECD for acceptance within the regulatory framework of chemicals (e.g. REACh). To allow their use by industrials and by regulatory instances, they will be integrated into the OECD/ECHA QSAR Toolbox and the JRC QSAR database. Current researches, in the PREDIMOL project, are dedicated to the development of new models for other hazardous physico-chemical properties and other families of compounds (organic peroxides and amines).