

### REACH et les (Q)SAR DE QUOI S'AGIT IL ?

La réglementation REACH va nécessiter la détermination des caractéristiques toxicologiques et écotoxicologiques de plusieurs dizaines de milliers de substances chimiques. Sous ces termes il faut considérer des propriétés intrinsèques à chaque substance aussi variées que la mutagenicité, la neurotoxicité, la toxicité par ingestion (DL50), les caractéristiques irritantes, allergènes ou bien la CL50 pour la toxicité aquatique ... Au total, à l'échelle européenne, des millions de tests devraient être réalisés dans un futur proche. Il est absolument unimaginable de mettre en oeuvre ce chantier par l'expérimentation animale « in vivo » ou même par des méthodes « in vitro » qui seraient préférables aux premières mais pas forcément réalisables et qui de toutes façons submergeraient rapidement les laboratoires. La réglementation REACH fait référence à ces aspects et les (Q)SAR sont mentionnés dans le cas des méthodes dites alternatives :

*Extrait :*

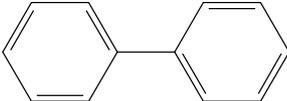
**Essais sur les animaux:** la limitation des essais sur les animaux a été l'un des principes directeurs dans l'élaboration de la proposition. Le comité scientifique de la toxicité, de l'écotoxicité et de l'environnement (CSTEE) a exprimé la crainte que les essais envisagés sur les animaux ne fourniraient pas des informations suffisantes pour éviter les risques et affirmé que davantage d'essais seraient nécessaires.

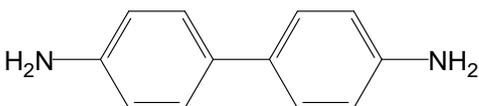
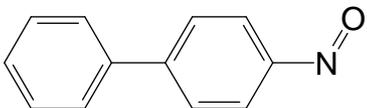
- En raison de la forte pression du public en faveur de la limitation des essais sur les animaux, le nombre des essais n'a pas été augmenté.
- Pour réduire encore le besoin d'essais sur les animaux sans mettre en péril la santé humaine et l'environnement, l'utilisation de modèles de relation qualitative ou quantitative structure-activité- (Q)SAR - est encouragée; le texte du règlement précise aussi que le partage des données sera obligatoire.

#### **Les QSAR (Quantitative Structure Activity Relationship ou RSA Relation Structure Activité)**

Les QSAR constituent une méthode de prédiction « in silico » des caractéristiques mesurables des molécules à partir des informations relatives à leur structure. Les molécules sont constituées de l'assemblage d'atomes liés entre eux par des liaisons chimiques et dont les possibilités d'agencement sont infinies. C'est cet agencement et les atomes qui les constituent qui vont conférer à la molécule ses propriétés intrinsèques.

*Ci-dessous exemple de molécules proches sur le plan structural mais dont les propriétés mutagènes pour deux d'entre elles sont conférées par la présence des groupes amino et nitro aromatiques.*

<b>Nom</b>	<b>Structure</b>	<b>Caractéristiques mutagènes</b>	<b>Cause</b>
Diphenyl		non	

Benzidine		oui	Aminoaromatique
Nitrosobiphenyl		oui	Nitroaromatique

*Prédictions obtenues avec la plateforme Bio-rad KnowItAll Adme /tox avec deux modèles distincts (1)*

Construire un modèle QSAR consiste à établir une relation mathématique entre une propriété mesurable et la structure chimique qui peut-être décrite par des variables chimiques telles que « l'hydrophobicité », la distribution électronique ou les propriétés stériques (volume, géométrie spatiale). Ces relations mathématiques sont des fonctions des « descripteurs » judicieusement choisis pour une propriété donnée. Ces relations sont construites à l'aide de logiciels utilisant des méthodes mathématiques (2) plus ou moins sophistiquées permettant l'analyse de données expérimentales d'un échantillon de molécules.

Exemple de QSAR :

Expression donnant la caractéristique anesthésiante d'un éther à partir des caractéristiques hydrophobes (LogP)

$$\text{Log}1/C = -0.22(\text{LogP})^2 + 1.04 \text{LogP} + 2.16 \quad (3)$$

L'utilisation d'un QSAR n'est applicable que pour une substance pure. Ce qui exclut leur utilisation dans le cas de mélanges de substances (cas complexe des huiles essentielles et soumis à la réglementation). Dans ce cas, on pourra, tout au plus, déterminer les propriétés des composants majeurs (purs) du mélange.

Par ailleurs, pour pouvoir « construire » un modèle QSAR pour une propriété donnée, il faut disposer d'une base de données contenant des structures et la mesure expérimentale de la propriété associée. C'est le seul moyen pour corrélérer une propriété aux descripteurs.

Sachant que la qualité de prédiction d'un QSAR pour une molécule donnée est dépendante de la base de données de structures à partir de laquelle on construit le modèle QSAR, il conviendra, pour un modèle solide et fiable, de construire un QSAR à partir de données provenant d'une famille de molécules « similaires » à la molécule étudiée. Disposer de ce type de données n'est pas facile et l'utilisateur sera souvent contraint d'utiliser des QSAR « prêts à l'emploi » construits à partir de bases de données conséquentes de molécules d'origines très variées. On parlera dans ce cas de modèles « globaux ». Leurs qualités prédictives sont très variables d'une molécule à l'autre. Les résultats obtenus sont parfois surprenants de justesse pour certaines molécules et totalement décevants pour d'autres, d'où l'importance du regard de l'expert dans l'interprétation finale d'une prédiction.

De nombreux modèles globaux pour diverses propriétés sont disponibles sur le marché. Ces modèles sont intégrés dans des logiciels permettant de dessiner ou d'importer des structures chimiques et d'obtenir les résultats par un simple « clic ». Certains de ces logiciels disposent d'outils complémentaires de gestion des données et de réalisation de rapports (4).

*Exemple d'interface où apparaissent des prédictions effectuées sur une base de données de structures traitées en batch, les résultats du traitement de la base sont triés en fonction de critères pré-établis*

The screenshot shows the MinelT Chemistry software interface. The main window displays a table with the following columns: logP So, Mol Weight, log P Predicted, and Blood-Brain Barrier Permeability. The table contains 46 rows of data, with values ranging from approximately -5.14 to 6.01 for logP So, and 395.166 to 683.803 for Mol Weight. The log P Predicted values range from 3.01 to 4.99, and Blood-Brain Barrier Permeability values range from -1.88 to 1.14. A chemical structure is shown on the right side of the interface, and a table of properties is visible at the bottom right.

Mol_ID	Name	Value	Unit
Mol_ID	1		
Formula	C29H23FN4O2S		
Mol Weight	510.585		g/mol
logP So	-6.01		
log P Comments Predicted	Log P predicted by ChemSilico Log P Predictor version 1.1.		
log P Predicted	3.01		
Blood-Brain Barrier Permeab:	1.14		

Interface de la plateforme Bio-rad knowItAll ADME/Tox

Les QSAR sont-ils fiables ? Il n'existe pas de QSAR universel et les résultats obtenus, comme nous l'avons vu, dépendent de la source de données, du choix des descripteurs et de la méthode utilisée. Les résultats doivent être interprétés avec la plus grande précaution. Seul l'expert en toxicologie ou le chimiste averti sera en mesure, non pas de garantir, mais de valider avec la mention « probable », le résultat obtenu. Avec un modèle « local », des prédictions exactes à 90% seraient remarquables, il en demeurerait non moins vrai que, même dans ce cas, 10% des molécules seraient nécessairement classées pour une caractéristique donnée « vrai négatif » ou « faux positif ».

En conclusion, les QSAR constituent un outil puissant d'information sur les caractéristiques probables d'une substance à condition d'être maniés et interprétés par un utilisateur averti. Ils peuvent permettre d'éviter des tests expérimentaux ou de ne les effectuer que lorsque l'interprétation des résultats laisse planer une suspicion. Ils peuvent permettre aussi de compléter l'information expérimentale. Les QSAR sont faciles à mettre en œuvre, ils fournissent un résultat immédiat et leurs coûts sont sans commune mesure avec ceux engendrés par la réalisation de tests expérimentaux.

Ils peuvent être réalisés par des laboratoires équipés des logiciels adaptés ou par le demandeur lui-même s'il dispose des outils informatiques et des personnes compétentes pour valider les résultats car on ne saurait trop insister sur les risques de conclusions hâtives.

Hugues Arnould

ITtekm- Agent distributeur de solutions informatiques pour la recherche et l'industrie  
fr

(1) régression, Analyse multivariées en composante principale, PLS, méthode « Naïve Bayes », réseaux de neurones, SVM...

(2) CSMutagenicity (ChemSilico) et PredictIt Toxicity (CompuDrug).

(3) Extrait de Chimie Pharmaceutique Graham L. PATRICK

(4) Quelques éditeurs liste non exhaustive (Bio-rad Informatics, Lhasa, Syracus Research Corporation-EPISuite, ChemSilico, MultiCase, Accelrys...)

Lien utile : European Chemicals Bureau <http://ecb.jrc.it/qsar/>