# Méthodes de corrélation Quantitative

## QSAR : quantitative structure-activité relationships

Objectif: Prédire l'activité biologique de composés non testés
à partir de « descripteurs »

Pr. Yves Blache

Laboratoire Matériaux Polymères Interfaces - Environnement Marin (MAPIEM-EA 4323)
Université du Sud Toulon-Var
BP 20132
83957 La Garde cedex
http://mapiem.univ-tln.fr
blache@univ-tln.fr

**Objectifs du cours**

- Comprendre et savoir utiliser les concepts et les outils *in silico* dans le contexte d'évaluation des risques : **les principes de la modélisation moléculaire**

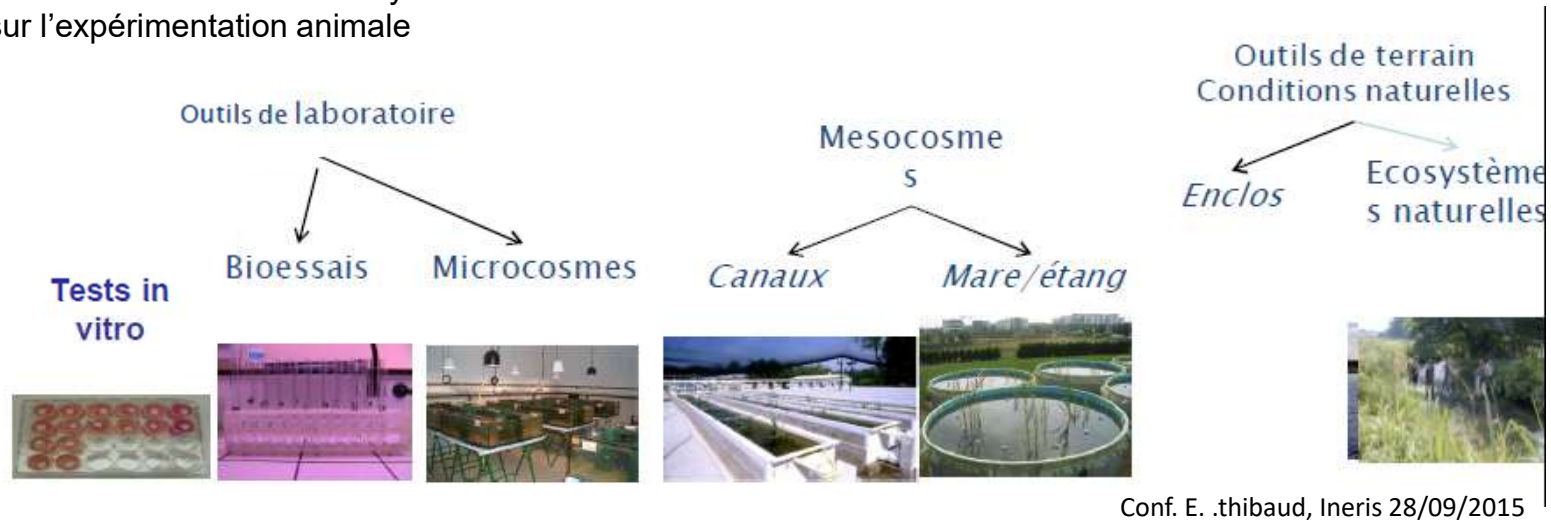
- Appréhender la prédiction de la (**éco**)**toxicité de molécules par la modélisation** des propriétés chimiques des substances à l'aide de **modèles QSAR**

## Contexte: REACH

*En2018, la conformité vis-à-vis de REACH est devenue obligatoire pour les substances chimiques importées ou fabriquées à plus de1t par an (TPA)*

**REACH est un règlement de l'Union européenne adopté pour mieux protéger la santé humaine et l'environnement contre les risques liés aux substances chimiques, tout en favorisant la compétitivité de l'industrie chimique de l'UE. Il encourage également des méthodes alternatives pour l'évaluation des dangers liés aux substances afin de réduire le nombre d'essais sur animaux.**

->25000 substances sur la période 2013-2018 (19 à 35 études/substance)
-70000 à >200000€ par dossier (coût d'enregistrement incluant les études)
-risque de saturation des laboratoires d'analyses
-impact éthique sur l'expérimentation animale




Conf. E. .thibaud, Ineris 28/09/2015

Cout élevé ➡ Utilisation des modèles comme substituts à l'expérimentation

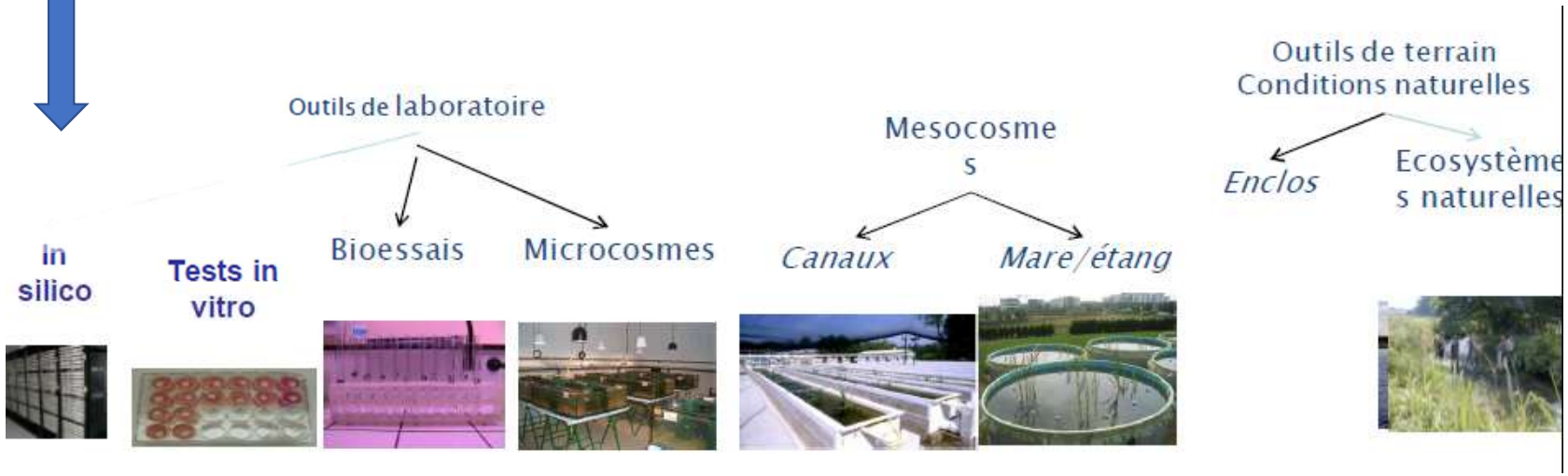
**Outils de modélisation moléculaire**

-1er filtre go/no-go en recherche
-constitution des dossiers pour l'AMM de produits

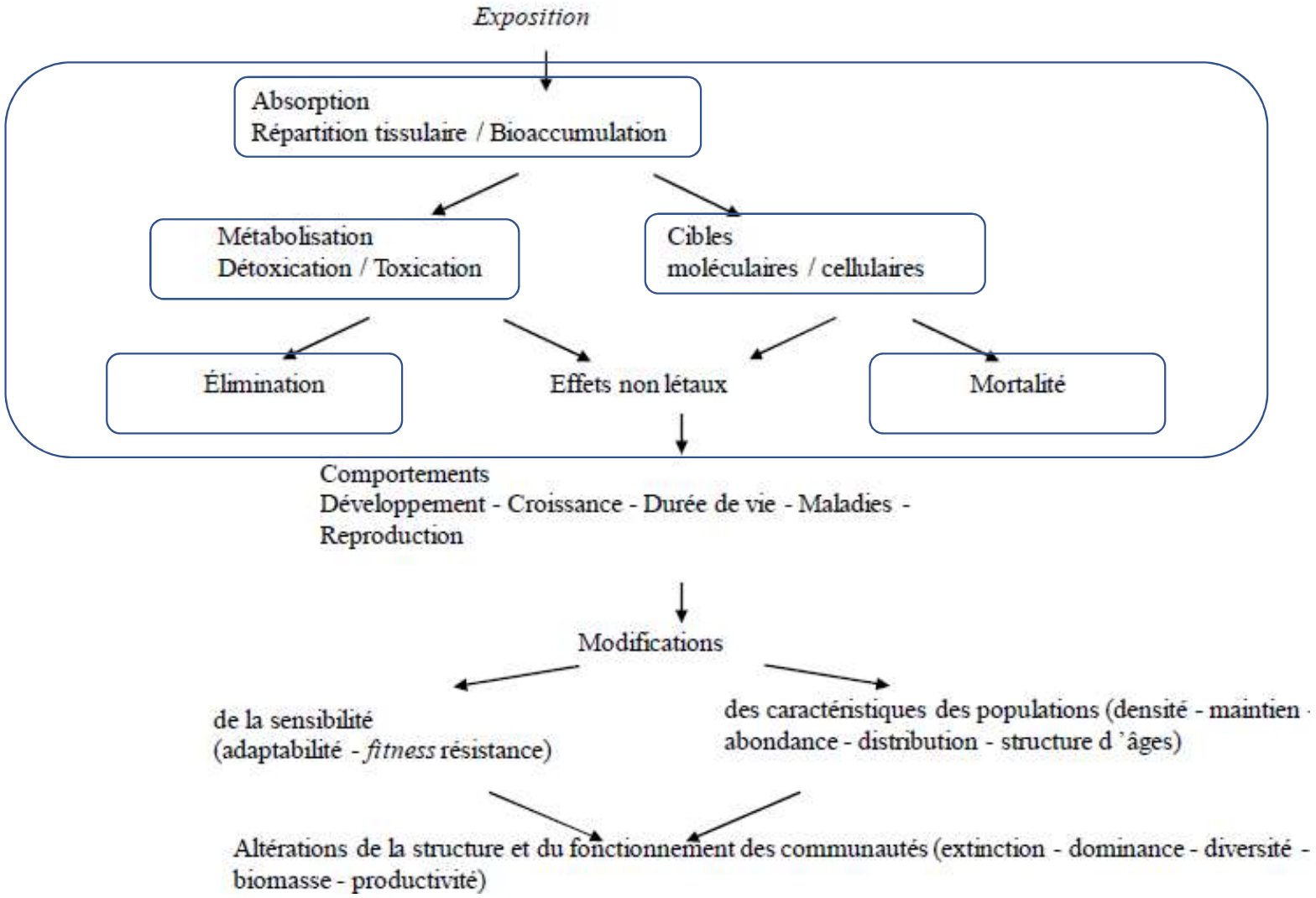
**3 avantages majeurs:**

1. Réduction du coût de constitution des dossiers REACH jusqu'à 40%
2. Réduction du délais de constitution des dossiers REACH
3. Limitations des expérimentations animales

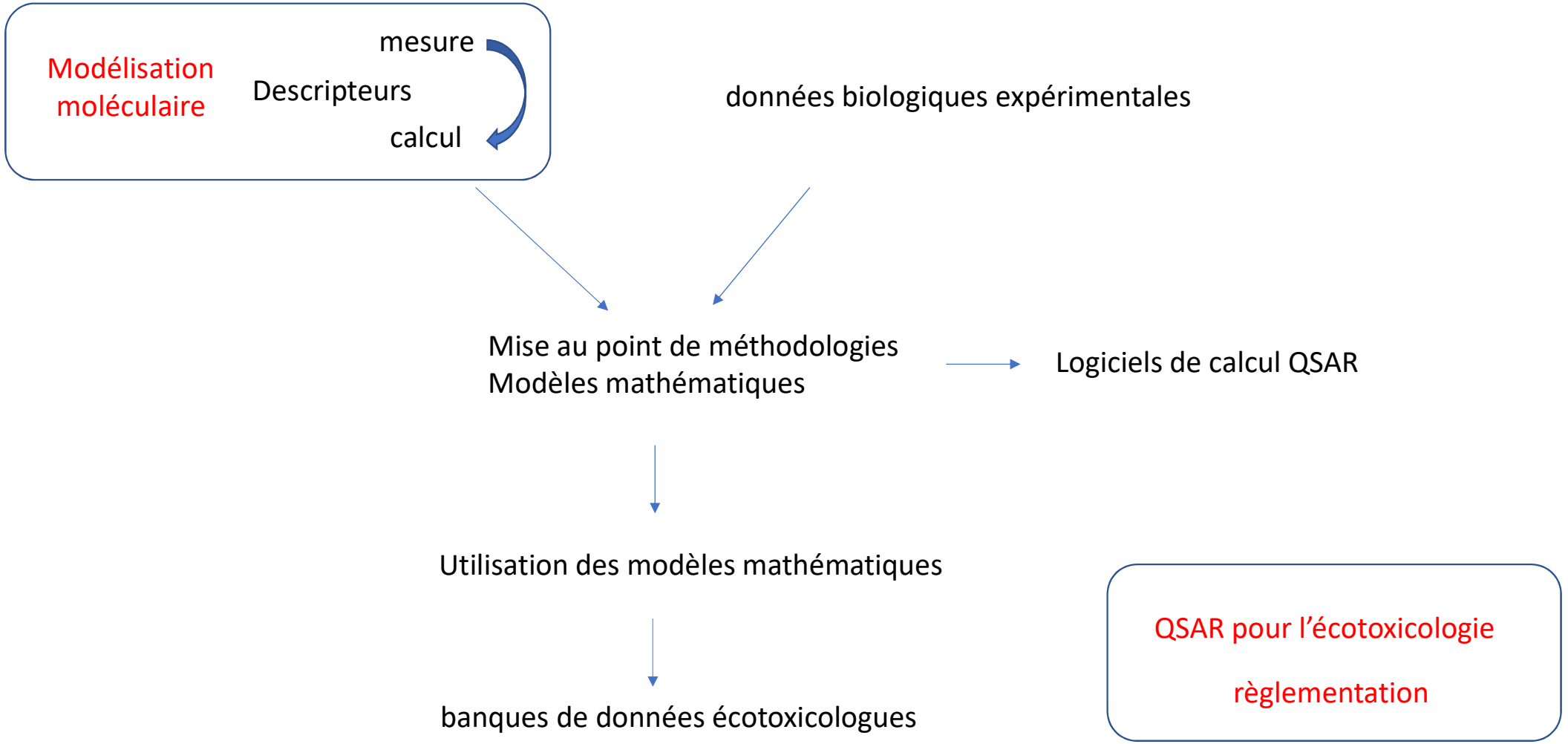
- aide à l'expertise des produits déjà sur le marché



Séquence chronologique des évènements biologiques affectant les organismes exposés aux polluants de l'environnement




Conf. E. .thibaud, Ineris 28/09/2015


Modélisation moléculaire
mesure
Descripteurs
calcul
données biologiques expérimentales
Mise au point de méthodologies
Modèles mathématiques
Logiciels de calcul QSAR
Utilisation des modèles mathématiques
banques de données écotoxicologues
QSAR pour l'écotoxicologie
règlementation

**1865**

Crum Brown and Fraser[1] expressed the idea that there was a relationship between activity and chemical structure.

**1893**

**Richet[2] correlated toxicities of simple organic molecules with their solubility in water.**

**1900**

Meyer[3] and Overton[4] found linear relationships between the toxicity of organic compounds and their lipophilicity.

**1969**

**Hansch[5] published a free-energy related model to correlate biological activities with physicochemical Properties.**

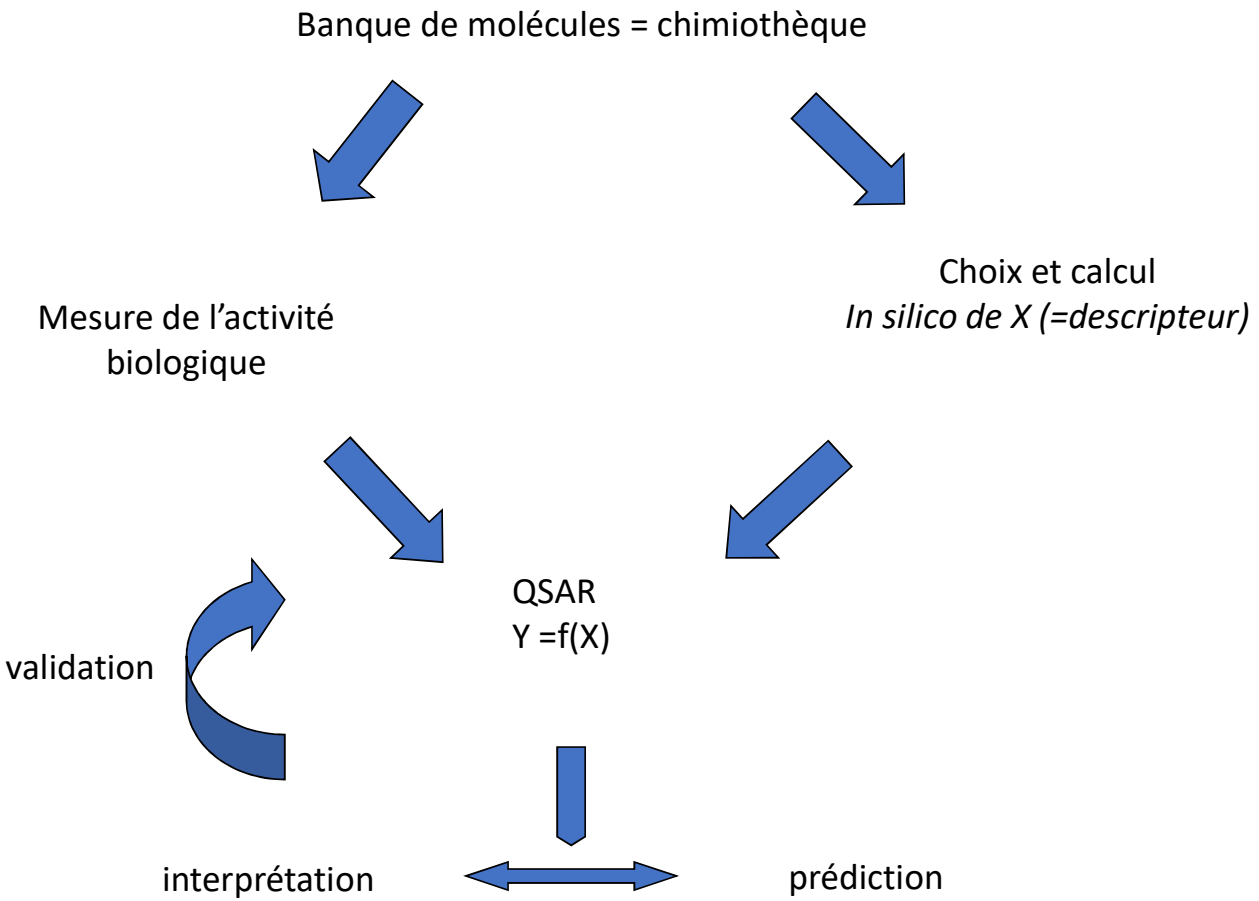
The father of the concept of quantitative structure-activity relationship (QSAR), the quantitative correlation of the physicochemical properties of molecules with their biological activities[5]

**Corwin Herman Hansch**

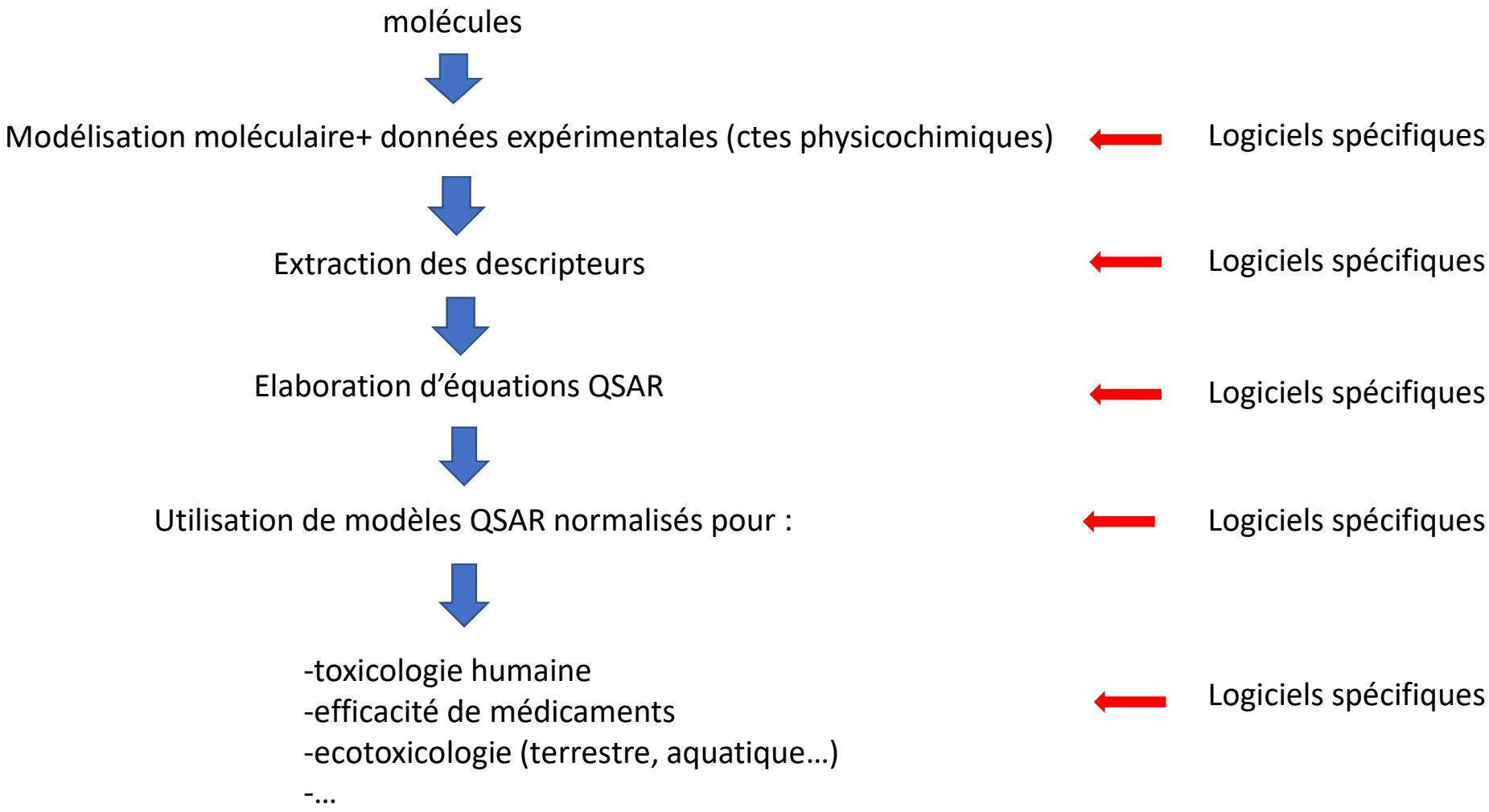
(1918 – 2011)

[1] A. Crum-Brown, T.R. Fraser, Trans. R. Soc. Edinb. 25 (1968–1969) 257. [2] M.C. Richet, Compt. Rend. Soc. Biol. 45 (1893) 775.
[3] H. Meyer, Arch. Exp. Pathol. Pharmakol. 42 (1899) 109. [4] E. Overton, Studien über die Narkose, Gustav Fischer, Jena, 1901
[5] Hansch, C. (1969) A quantitative approach to biochemical structure-activity relationships *Accounts of Chemical Research*, 2, 232-239.

# Elaboration d'un modèle qsar



## Elaboration d'un modèle qsar



## Principe : équation de corrélation

- Permet de déterminer les valeurs des paramètres qui correspondent à une activité maximale et ainsi de prédire l'activité des molécules qui n'ont pas encore été synthétisée

- Act. Bio= Cte + $C_1P1$ + $C_2P_2$ + $C_nP_n$

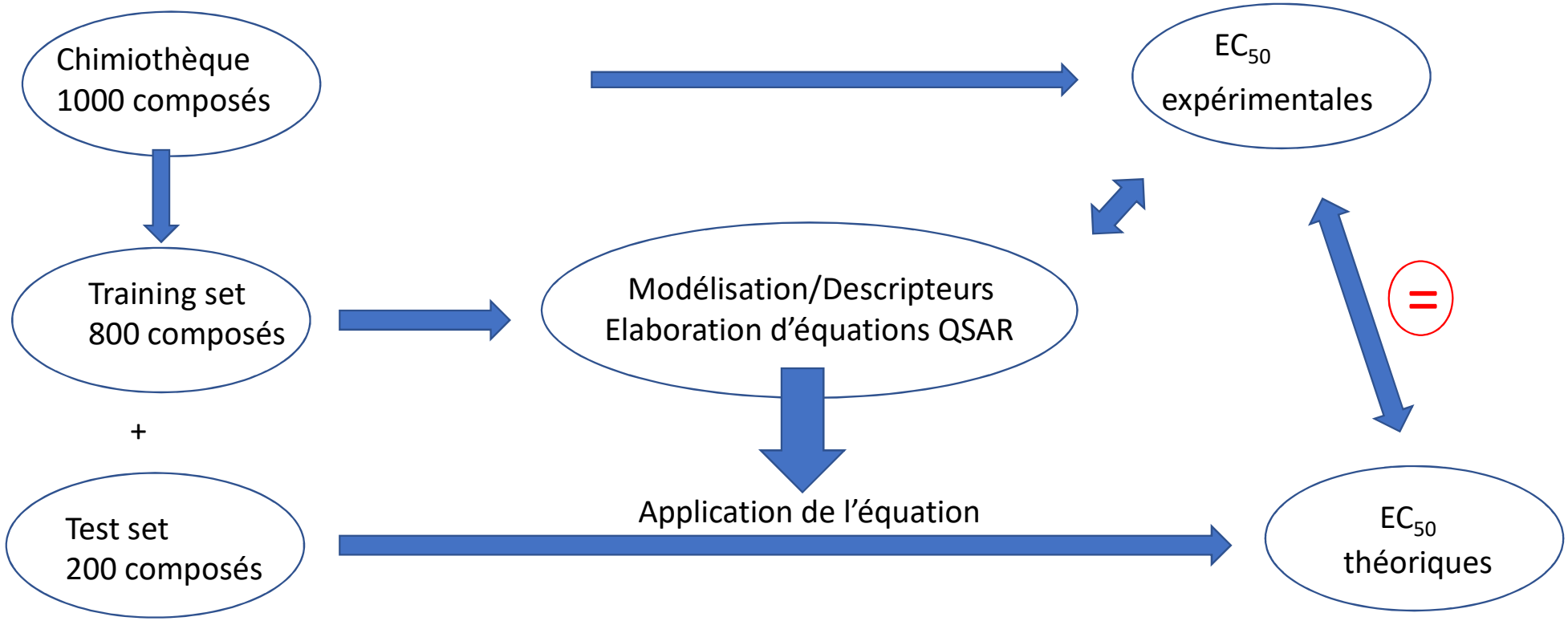
P : paramètres calculés = descripteurs

- La validité d'un modèle QSAR dépendra donc du choix que l'on aura fait sur les paramètres

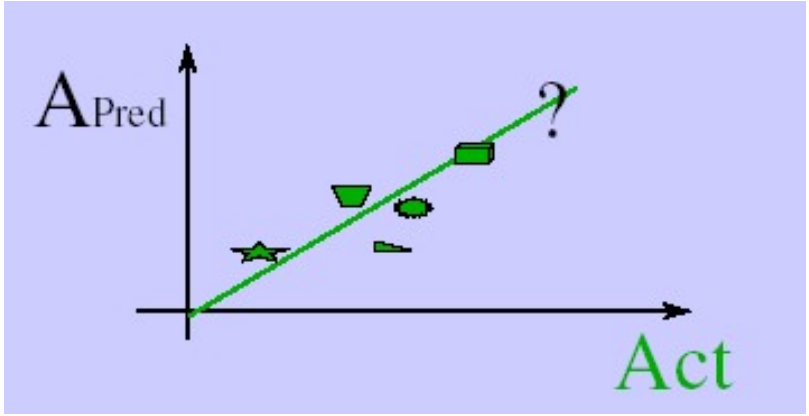
# Principe : méthodologie

L'élaboration d'un modèle nécessite en amont des données expérimentales importantes

Exemple



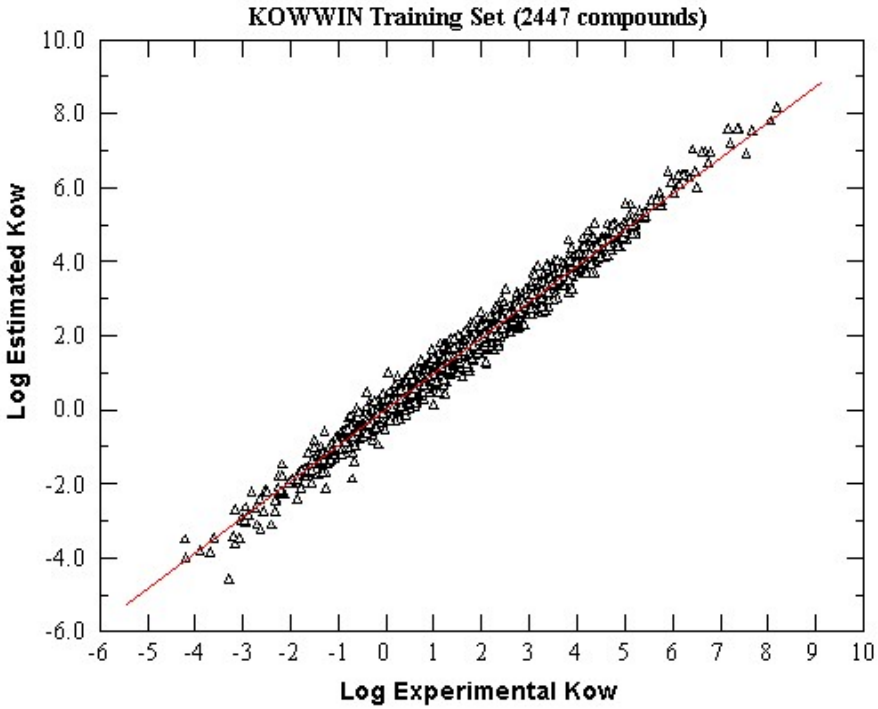
Un bon modèle QSAR!




Points expérimentaux proches de la droite théorique

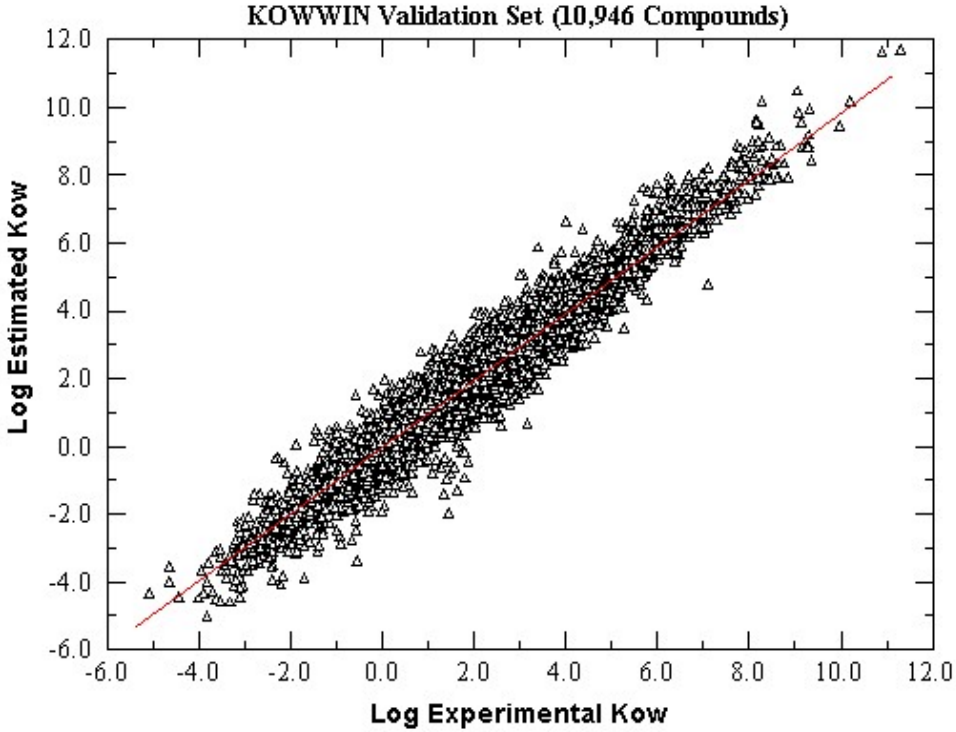
Bon caractère prédictif

séparation entre bons substrats / substrats intermédiaires / mauvais substrats (classement dans l'ordre)




**Total Training Set Statistics:**

| | |
|---|---|
| number in dataset | = 2447 |
| correlation coef ($r^2$) | = 0.982 |
| standard deviation | = 0.217 |
| absolute deviation | = 0.159 |
| avg Molecular Weight | = 199.98 |




**Total Validation Set Statistics:**

| | |
|---|---|
| number in dataset | = 10946 |
| correlation coef ($r^2$) | = 0.943 |
| standard deviation | = 0.479 |
| absolute deviation | = 0.356 |
| avg Molecular Weight | = 258.98 |

**Molecular Descriptors** are numerical values that characterize properties of molecules.

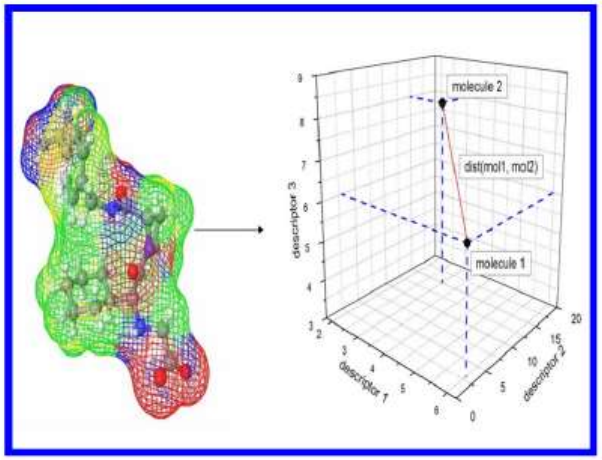
calculated by

**Experimental measurements**

**Theoretical Indexes**

Mathematical formulas

Computational algortihms




**Different Theories** (e.g. Quantum-chemistry, information theory, organic chemistry)

derived from

More than 5000[2]

**Molecular Descriptors**

processed by

statistics
chemometrics
chemoinformatics

applied in

**Scientific Fields** (e.g. QSAR, medicinal chemistry, drug design, toxicology, analytical chemistry, Environmetrics, virtual screening)

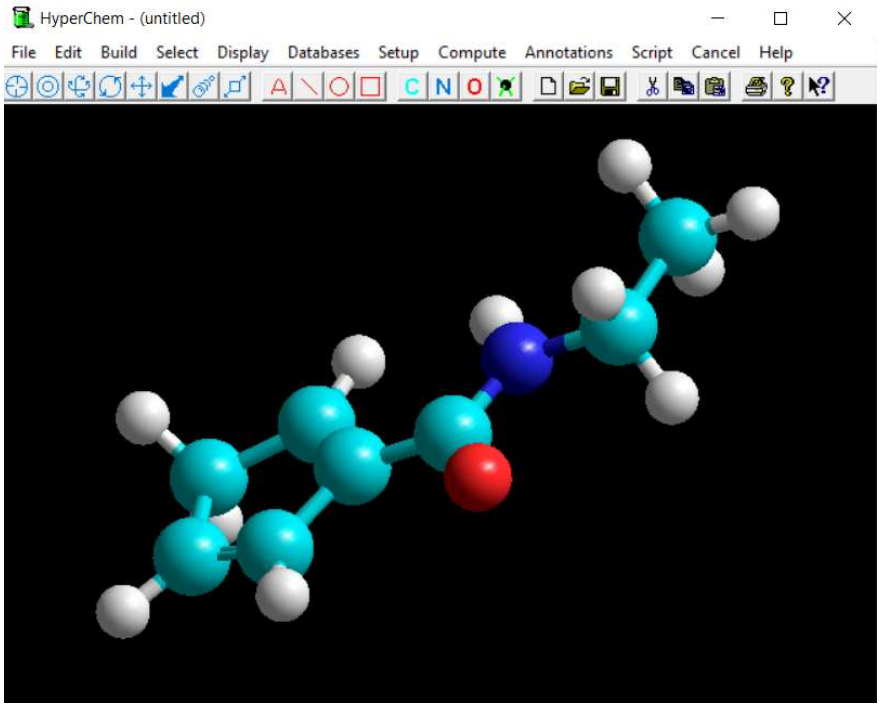
[1] *T. Puzyn et al. (eds.), Recent Advances in QSAR Studies, 103–125.* DOI 10.1007/978-1-4020-9783-6_4, C Springer Science+Business Media B.V. 2010
[2] Todeschini R, Consonni V (2000) Handbook of molecular descriptors. Wiley-VCH, Weinheim

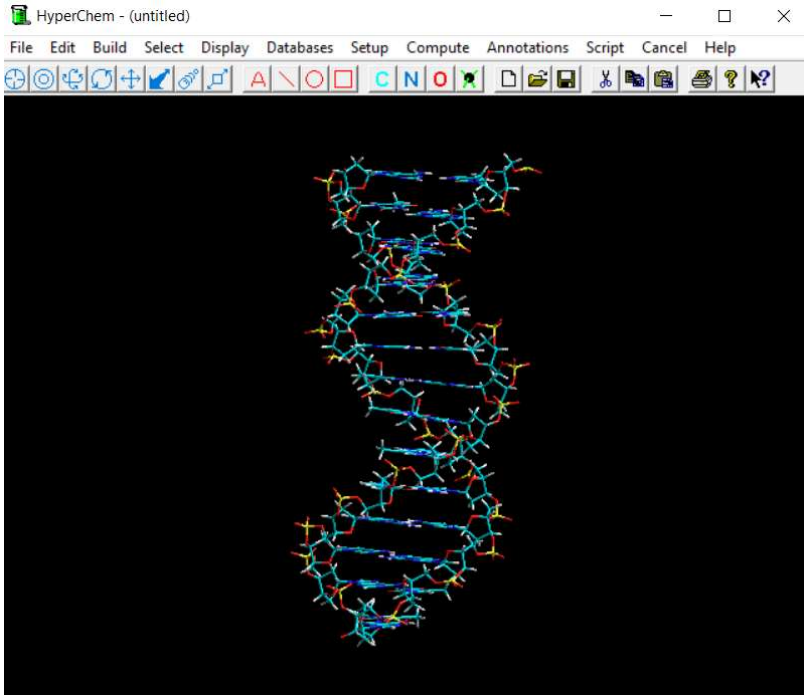


Ceyda Oksel, Xue Z. Wang, University of Leeds, Lecture « quality nano trainning school »27.03.2013

# Comment calculer les propriétés d'une molécule?

## La modélisation moléculaire

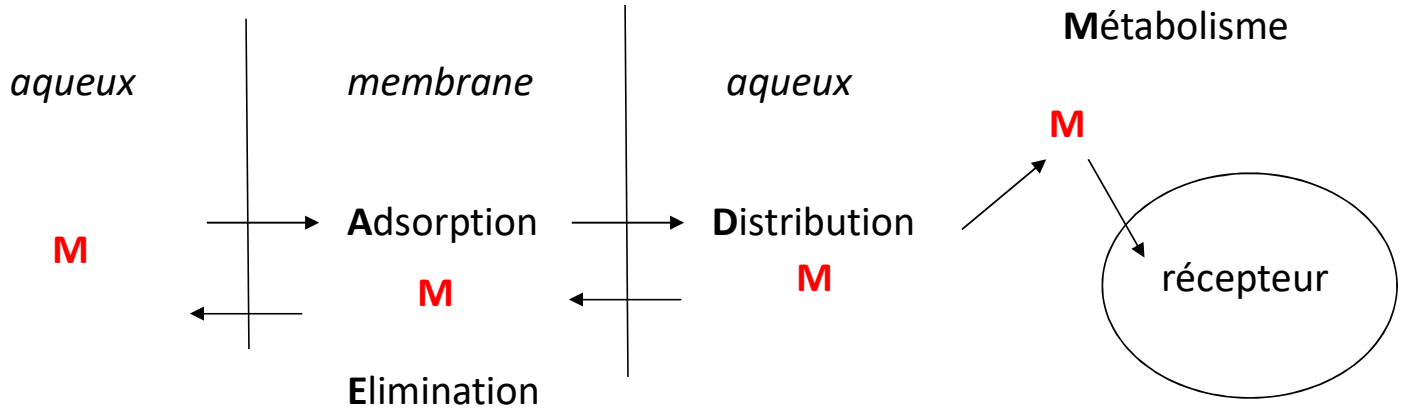


## Quels sont les paramètres (descripteurs) les plus pertinents?

### Phase ADME

**ADME** : -**A**dsorption (solubilité, ionisation, passage de membrane)
-**D**istribution (idem)
-**M**étabolisme (interaction avec les enzymes responsables de métabolisation)
-**E**limination (solubilité, ionisation)

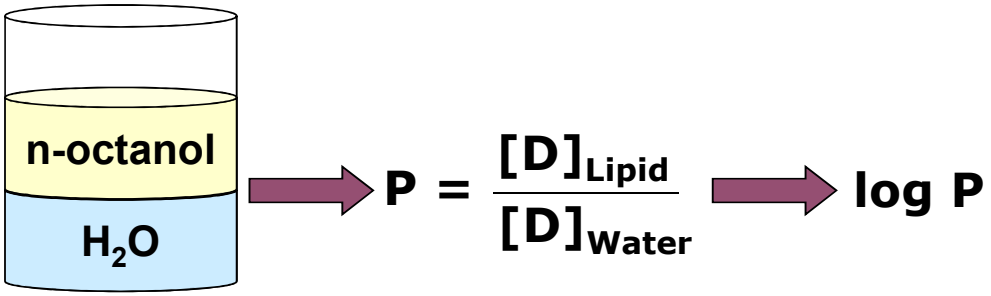



***Paramètres pertinents: solubilité, lipophilie, paramètres électroniques***

# Les paramètres de lipophilie

Coefficient de partage octanol/eau

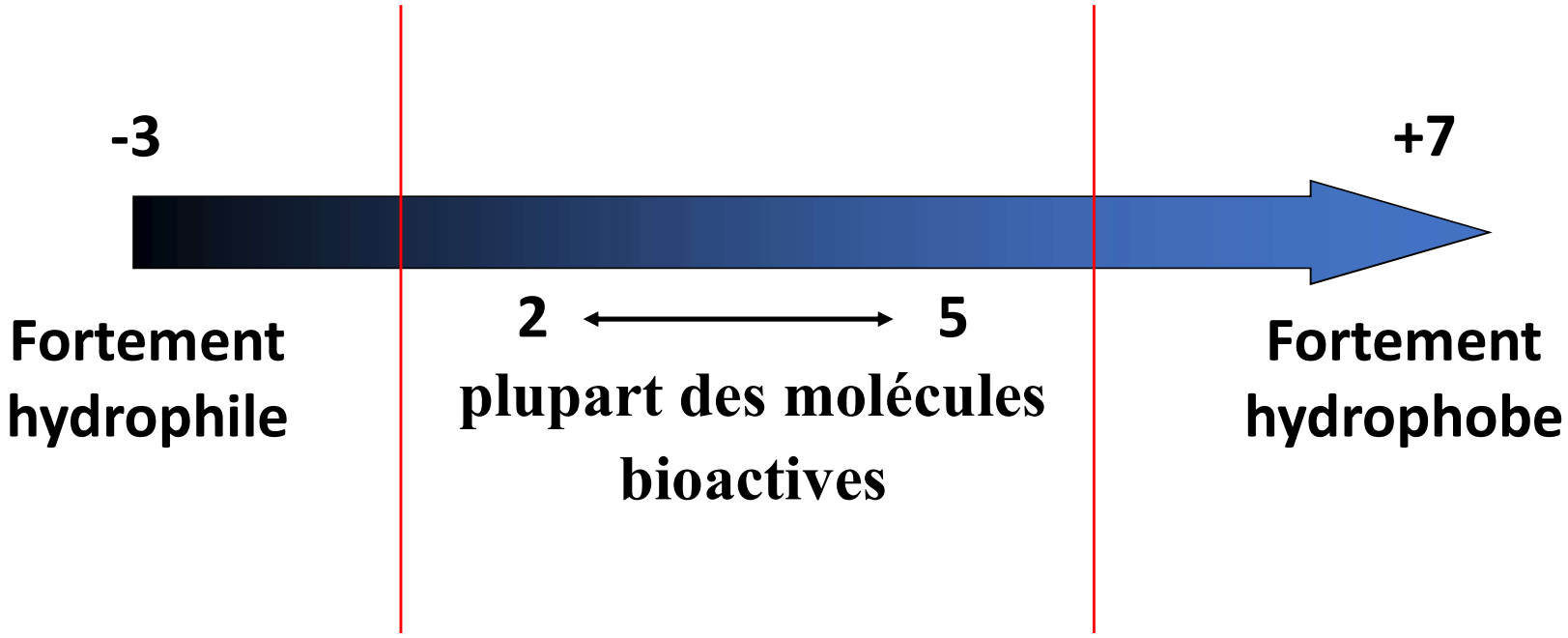
Peut se déterminer experimentalement
Peut se calculer *in silico*: plusieurs méthodes, certaines « en ligne »




Modèle qsar

$$Log(act) = log(P) + Cte$$

# logP


-3
+7
Fortement hydrophile
2
5
plupart des molécules bioactives
Fortement hydrophobe

inhibition de l'agrégation des plaquettes de sang de bovins par des composés alcooliques (ROH)

$$\log A = 1{,}00.\log P + 0{,}18$$

inhibition de la contraction du muscle sartorius de grenouille par différents composés :

$$\log A = 0{,}88.\log P + 0{,}63$$

Autres paramètres physico-chimiques

hydrophilie

logP en fonction du pH

Logiciels permettant de calculer uniquement les paramètres physico-chimiques sont nombreux

Certains sont utilisables « en ligne »

Ex: molinspiration, EPISUITE…

Calculation of Molecular Properties and Drug-likeness - Microsoft Internet Explorer

Fichier Edition Affichage Favoris Outils ?

Précédente Rechercher Favoris

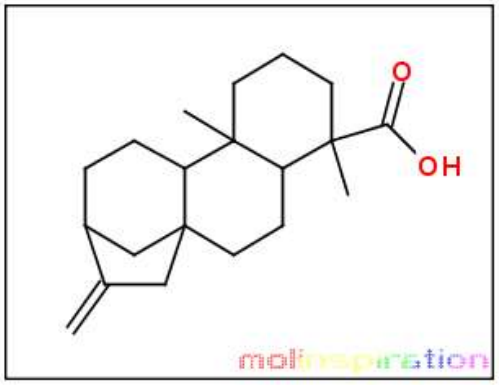
Adresse http://www.molinspiration.com/cgi-bin/p OK Liens SciFinder Scholar® Explore as Research Topic Explore

molinspiration Rechercher Traduire Ouvrir session Y! Mail Questions/Réponses Messenger

molinspiration

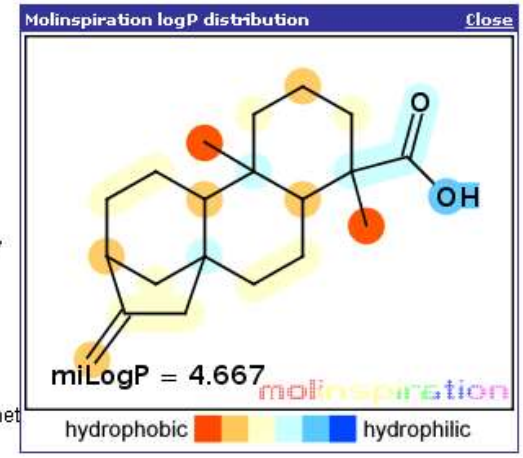
Calculation of Molecular Properties

SMILES C42(C)C(C(C)(C(=O)O)CCC4)CCC13C2CCC(C1)C(=C)C3
miSMILES CC12CCCC(C)(C(O)=O)C1CCC34CC(CCC23)C(=C)C4




| | | |
|---|---|---|
| miLogP | 4.667 | distribution NEW |
| TPSA | 37.299 | |
| natoms | 22 | |
| MW | 302.458 | |
| nON | 2 | |
| nOHNH | 1 | |
| nviolations | 0 | |
| nrotb | 1 | |
| volume | 308.332 | |

Get data as text (for copy /




This was request 1 out of 100 available this month for your site 193.52.233.18.
With technology from Molinspiration you can easily setup similar service also directly on your intranet
Comments or questions ? See FAQ and do not hesitate to contact us by email !

**Molecular polar surface area**

Rotable bonds

javascript:void(0); Internet

Démarrer cours Microsoft PowerPoin... Calculation of Mo... Pckocwin v1.66 8:51 AM

## Quels sont les paramètres les plus pertinents?

**Les paramètres électroniques**

**(interviennent sur le complexe récepteur/ligand et ADME)**

Certains paramètres (descripteurs) ⇢ Interaction récepteur/ligand Mécanisme d'action

Molécule + récepteur $\underset{}{\overset{K_a}{\rightleftharpoons}}$ complexe

$$\Delta G = RT.\ln(K_a) = \Delta H - T.\Delta S$$

## Les paramètres électroniques

## (interviennent sur le complexe récepteur/ligand et ADME)

Orbitales moléculaires: HOMO et LUMO

Calcul de l'énergie (non la forme et la localisation)

Charges atomiques partielles

Liaisons Hydrogène

……..

**Plus difficiles à mesurer expérimentalement**

**Faciles à calculer**

**Plusieurs logiciels de modèles QSAR existent**

EPISUITE/ECOSAR
T.E.S.T.
VEGA
DanishQSAR Database
SPARC
ChemSilico
ADMET Predictor
ACD/Labs
CATALOGIC *etc…*

-peuvent être imprécis (conçus pour du screening→valeur approximative)
-manque de fiabilité (basés sur des données expérimentales hétérogènes et/ou pauvres)
-source d'ambiguïté
-utilisation complexe (!)
-définition du domaine d'applicabilité
-forte variabilité dans le cas des méthodes par consensus

**Certains sont développésvalidés et utilisés par des agences internationales**

## Les agences internationales

https://www.epa.gov/

The mission of EPA is to protect human health and the environment.
•EPA works to ensure that: Americans have clean air, land and water;
•National efforts to reduce environmental risks are based on the best available scientific information;
•Federal laws protecting human health and the environment are administered and enforced fairly, effectively and as Congress intended;
•Environmental stewardship is integral to U.S. policies concerning natural resources, human health, economic growth, energy, transportation, agriculture, industry, and international trade, and these factors are similarly considered in establishing environmental policy;
•All parts of society--communities, individuals, businesses, and state, local and tribal governments--have access to accurate information sufficient to effectively participate in managing human health and environmental risks;
•Contaminated lands and toxic sites are cleaned up by potentially responsible parties and revitalized; and
•Chemicals in the marketplace are reviewed for safety

---

https://echa.europa.eu/fr/home

L'ECHA joue un rôle central au sein des autorités de réglementation pour la mise en œuvre de la nouvelle législation européenne sur les produits chimiques et entend promouvoir la protection de la santé humaine et de l'environnement ainsi que l'innovation et la compétitivité.
L'ECHA aide les entreprises à se conformer à la législation, favorise l'utilisation sûre des substances chimiques, fournit des informations sur les produits chimiques et examine les produits chimiques préoccupants.

ECHA aspire à devenir la première autorité de réglementation au monde en matière de sécurité des produits chimiques

***REACH : règlement de l'Union européenne adopté pour mieux protéger la santé humaine et l'environnement contre les risques liés aux substances chimiques, tout en favorisant la compétitivité de l'industrie chimique de l'UE. Il encourage également des méthodes alternatives pour l'évaluation des dangers liés aux substances afin de réduire le nombre d'essais sur animaux.***

Toues les nouvelles substances nécessitent une autorisation avant de pouvoir être mis sur le marché.

Les biocides sont particulièrement concernés qui servent à protéger des êtres humains, des animaux, des matériaux ou des objets contre des organismes nuisibles tels que des insectes nuisibles ou des bactéries (pesticides, fongicides, algicides, antifouling…) .

Le règlement sur les produits biocides (BPR, règlement (UE) 528/2012) concerne leur mise sur le marché et leur utilisation.

Le texte a été adopté le 22 mai 2012

Ce règlement vise à :

- améliorer le fonctionnement du marché des produits biocides dans l'UE
- garantir un niveau élevé de protection de l'homme et de l'environnement

les produits biocides sont classés en 4 grands groupes

https://echa.europa.eu/fr/home

ECHA
EUROPEAN CHEMICALS AGENCY

| Number | Product-type | Description |
|---|---|---|
| **Main group 1: Disinfectants**<br>These product types exclude cleaning products that are not intended to have a biocidal effect, including washing liquids, powders and similar products. | | |
| PT 1 | Human hygiene | Products in this group are biocidal products used for human hygiene purposes, applied on or in contact with human skin or scalps for the primary purpose of disinfecting the skin or scalp. |
| PT 2 | Disinfectants and algaecides not intended for direct application to humans or animals | Used for the disinfection of surfaces, materials, equipment and furniture which are not used for direct contact with food or feeding stuffs. Usage areas include, inter alia, swimming pools, aquariums, bathing and other waters; air conditioning systems; and walls and floors in private, public, and industrial areas and in other areas for professional activities.<br><br>Used for disinfection of air, water not used for human or animal consumption, chemical toilets, waste water, hospital waste and soil.<br><br>Used as algaecides for treatment of swimming pools, aquariums and other waters and for remedial treatment of construction materials.<br><br>Used to be incorporated in textiles, tissues, masks, paints and other articles or materials with the purpose of producing treated articles with disinfecting properties. |
| PT 3 | Veterinary hygiene | Used for veterinary hygiene purposes such as disinfectants, disinfecting soaps, oral or corporal hygiene products or with anti-microbial function.<br><br>Used to disinfect the materials and surfaces associated with the housing or transportation of animals. |
| PT 4 | Food and feed area | Used for the disinfection of equipment, containers, consumption utensils, surfaces or pipework associated with the production, transport, storage or consumption of food or feed (including drinking water) for humans and animals.<br><br>Used to impregnate materials which may enter into contact with food. |
| PT 5 | Drinking water | Used for the disinfection of drinking water for both humans and animals. |

https://echa.europa.eu/fr/home

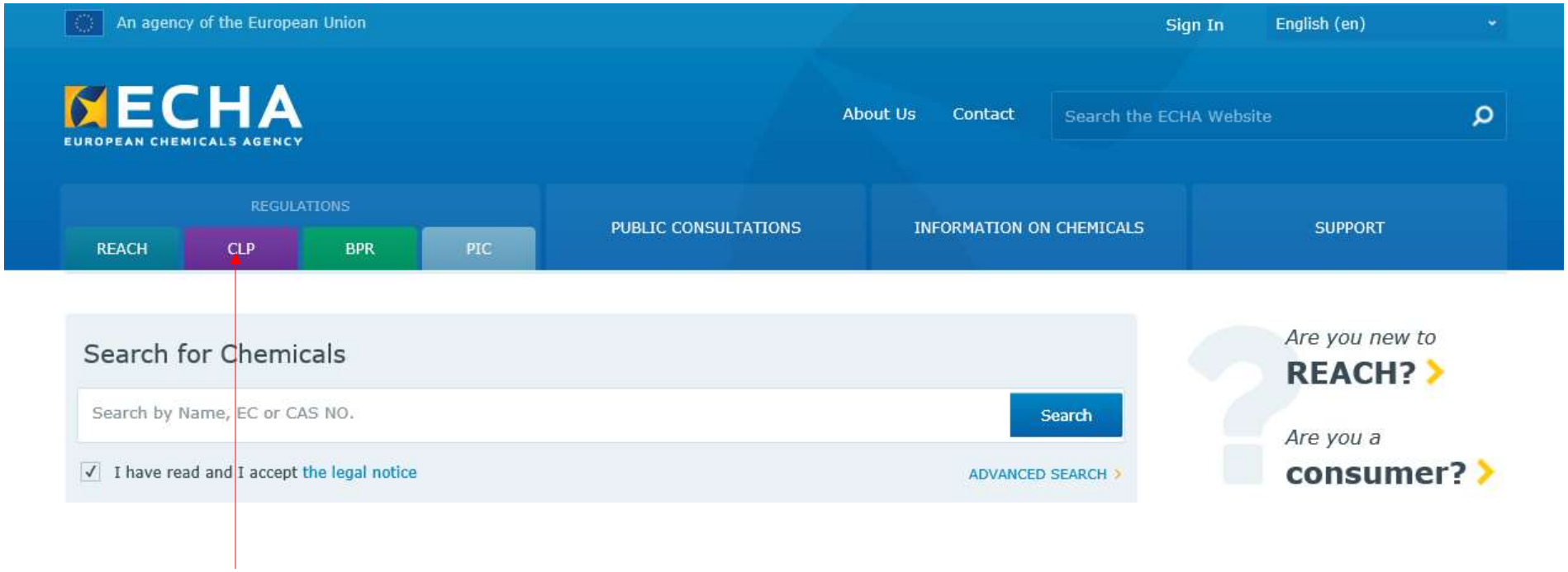
| Number | Product-type | Description |
|---|---|---|
| **Main group 2: Preservatives**<br>Unless otherwise stated these product-types include only products to prevent microbial and algal development. | | |
| PT 6 | Preservatives for products during storage | Used for the preservation of manufactured products, other than foodstuffs, feeding stuffs, cosmetics or medicinal products or medical devices by the control of microbial deterioration to ensure their shelf life.<br>Used as preservatives for the storage or use of rodenticide, insecticide or other baits. |
| PT 7 | Film preservatives | Used for the preservation of films or coatings by the control of microbial deterioration or algal growth in order to protect the initial properties of the surface of materials or objects such as paints, plastics, sealants, wall adhesives, binders, papers, art works. |
| PT 8 | Wood preservatives | Used for the preservation of wood, from and including the saw-mill stage, or wood products by the control of wood-destroying or wood-disfiguring organisms, including insects. This product type includes both preventive and curative products. |
| PT 9 | Fibre, leather, rubber and polymerised materials preservatives | Used for the preservation of fibrous or polymerised materials, such as leather, rubber or paper or textile products by the control of microbiological deterioration.<br>This product-type includes biocidal products which antagonise the settlement of micro-organisms on the surface of materials and therefore hamper or prevent the development of odour and/or offer other kinds of benefits. |
| PT 10 | Construction material preservatives | Used for the preservation of masonry, composite materials, or other construction materials other than wood by the control of microbiological and algal attack. |
| PT 11 | Preservatives for liquid-cooling and processing systems | Used for the preservation of water or other liquids used in cooling and processing systems by the control of harmful organisms such as microbes, algae and mussels.<br>Products used for the disinfection of drinking water or of water for swimming pools are not included in this product-type. |
| PT 12 | Slimicides | Used for the prevention or control of slime growth on materials, equipment and structures, used in industrial processes, e.g. on wood and paper pulp, porous sand strata in oil extraction. |
| PT 13 | Working or cutting fluid preservatives | Products to control microbial deterioration in fluids used for working or cutting metal, glass or other materials. |

https://echa.europa.eu/fr/home

| Number | Product-type | Description |
|---|---|---|
| **Main group 3: Pest control** | | |
| PT 14 | Rodenticides | Used for the control of mice, rats or other rodents, by means other than repulsion or attraction. |
| PT 15 | Avicides | Used for the control of birds, by means other than repulsion or attraction. |
| PT 16 | Molluscicides, vermicides and products to control other invertebrates | Used for the control of molluscs, worms and invertebrates not covered by other product types, by means other than repulsion or attraction. |
| PT 17 | Piscicides | Used for the control of fish, by means other than repulsion or attraction. |
| PT 18 | Insecticides, acaricides and products to control other arthropods | Used for the control of arthropods (e.g. insects, arachnids and crustaceans), by means other than repulsion or attraction. |
| PT 19 | Repellents and attractants | Used to control harmful organisms (invertebrates such as fleas, vertebrates such as birds, fish, rodents), by repelling or attracting, including those that are used for human or veterinary hygiene either directly on the skin or indirectly in the environment of humans or animals. |
| PT 20 | Control of other vertebrates | Used for the control of vertebrates other than those already covered by the other product types of this main group, by means other than repulsion or attraction. |

| Number | Product-type | Description |
|---|---|---|
| **Main group 4: Other biocidal products** | | |
| PT 21 | Antifouling products | Used to control the growth and settlement of fouling organisms (microbes and higher forms of plant or animal species) on vessels, aquaculture equipment or other structures used in water. |
| PT 22 | Embalming and taxidermist fluids | Used for the disinfection and preservation of human or animal corpses, or parts thereof. |

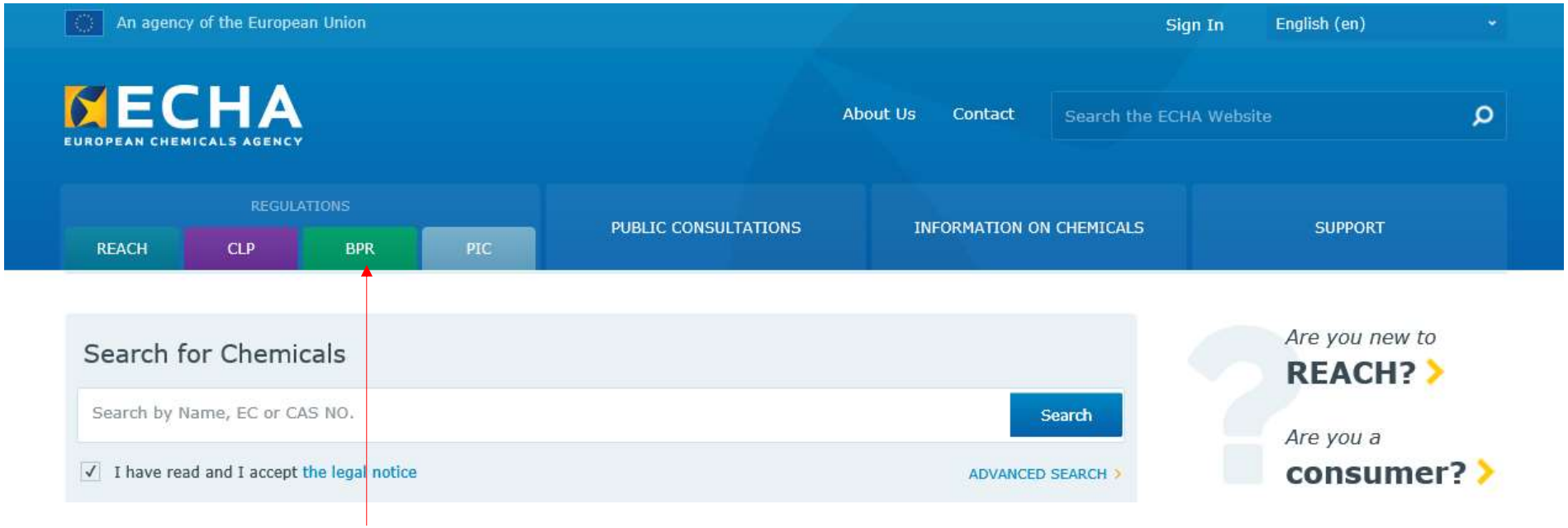
https://echa.europa.eu/fr/home




Le règlement (CE) n ° 1272/2008 sur la **classification, l'étiquetage et l'emballage (CLP)** est basé sur le Système général harmonisé (SGH) des Nations unies

Assurer un niveau élevé de protection de la santé et de l'environnement, ainsi que la libre circulation des substances, mélanges et objets.

Depuis le 1 er juin 2015, c’est la seule législation en vigueur dans l'UE pour la classification et l'étiquetage des substances et des mélanges.




Le **règlement sur les produits biocides** (BPR, règlement (UE) 528/2012) concerne la mise sur le marché et l'utilisation de produits biocides, qui servent à protéger des êtres humains, des animaux, des matériaux ou des objets contre des organismes nuisibles tels que des insectes nuisibles ou des bactéries. des substances actives contenues dans le produit biocide.

Tous nécessitent une autorisation avant de pouvoir être mis sur le marché ainsi que les substances actives contenues dans ce produit

Le texte a été adopté le 22 mai 2012

améliorer le fonctionnement du marché des produits biocides dans l'UE

garantir un niveau élevé de protection de l'homme et de l'environnement

les produits biocides sont classés en 4 grands groupes

Toutes ces molécules ont un impact potentiel sur l'environnement

Leur toxicité doit être étudiée sur différents type d'organismes :

QSAR et ecotoxicologie : définition des « endpoints »

-Dépend du champ d'applicabilité (santé, industrie agro, agriculture, aquaculture peintures…)

-Dépend de l'environnement (terrestre, eau douce, eau de mer…

# QSAR models related to REACH endpoints

## a. Physicochemical endpoints

| Endpoint | Software tool | Models/Modules | Free or Commercial |
|---|---|---|---|
| Melting/freezing point | EPI Suite (US EPA) | MPBPVP | Free |
| Boiling point | EPI Suite (US EPA) | MPBPVP | Free |
| | T.E.S.T. (US EPA) | Normal boiling point | Free |
| | ACD/Percepta (ACD/Labs) | Boiling Point/Vapor Pressure Module | Commercial |
| Relative density | T.E.S.T. (US EPA) | Density | Free |
| Vapour pressure | EPI Suite (US EPA) | MPBPVP | Free |
| | T.E.S.T. (US EPA) | Vapor pressure at 25°C | Free |
| | ACD/Percepta (ACD/Labs) | Boiling Point/Vapor Pressure Module | Commercial |
| Surface tension | T.E.S.T. (US EPA) | Surface tension at 25°C | Free |
| Water solubility | EPI Suite (US EPA) | WSKOW and WATERNT | Free |
| | T.E.S.T. (US EPA) | Water solubility at 25°C | Free |
| | ACD/Percepta (ACD/Labs) | Aqueous Solubility Module | Commercial |
| | ADMET Predictor (Simulations Plus) | Physicochemical and Biopharmaceutical Module | Commercial |
| | Discovery Studio (Accelrys) | ADMET Descriptors | Commercial |
| Partition coefficient (log Kow) | EPI Suite (US EPA) | KOWWIN | Free |
| | VEGA (IRFMN) | LogP Models | Free |
| | ACD/Percepta (ACD/Labs) | LogP Module | Commercial |
| | ADMET Predictor (Simulations Plus) | Physicochemical and Biopharmaceutical Module | Commercial |
| | JChem (ChemAxon) | LogP/logD predictor | Commercial |
| Flash point | T.E.S.T. (US EPA) | Flash point | Free |
| Dissociation constant | Danish QSAR Database (DTU) | pKa from ACD/Labs | Free |
| | ACD/Percepta (ACD/Labs) | pKa Module | Commercial |
| | ADMET Predictor (Simulations Plus) | Physicochemical and Biopharmaceutical Module | Commercial |
| | JChem (ChemAxon) | pKa predictor | Commercial |
| Viscosity | T.E.S.T. (US EPA) | Viscosity at 25°C | Free |

## b. Environmental fate and pathways endpoints

| Endpoint[11] | Software tool | Models/Modules | Free or Commercial |
|---|---|---|---|
| Hydrolysis | EPI Suite (US EPA) | HYDROWIN | Free |
| Ready biodegradability | Danish QSAR Database (DTU) | Not Ready Biodegradability model from DTU | Free |
| | EPI Suite (US EPA) | BIOWIN and BioHCwin | Free |
| | VEGA (IRFMN) | IRFMN model | Free |
| | CATALOGIC (LMC) | Several OECD 301 models | Commercial |
| | Discovery Studio (Accelrys) | Aerobic Biodegradability model | Commercial |
| | Meta-PC (MultiCASE) | Aerobic Microbial Biodegradation expert rules | Commercial |
| Bioaccumulation in aquatic species | EPI Suite (US EPA) | BCFBAF | Free |
| | T.E.S.T. (US EPA) | Bioaccumulation factor | Free |
| | VEGA (IRFMN) | CAESAR, Meylan and KNN/Read-Across models | Free |
| | CASE Ultra (MultiCASE) | EcoTox model bundle | Commercial |
| | CATALOGIC (LMC) | Two BCF base-line models | Commercial |
| Adsorption/desorption screening | EPI Suite (US EPA) | KOCWIN | Free |

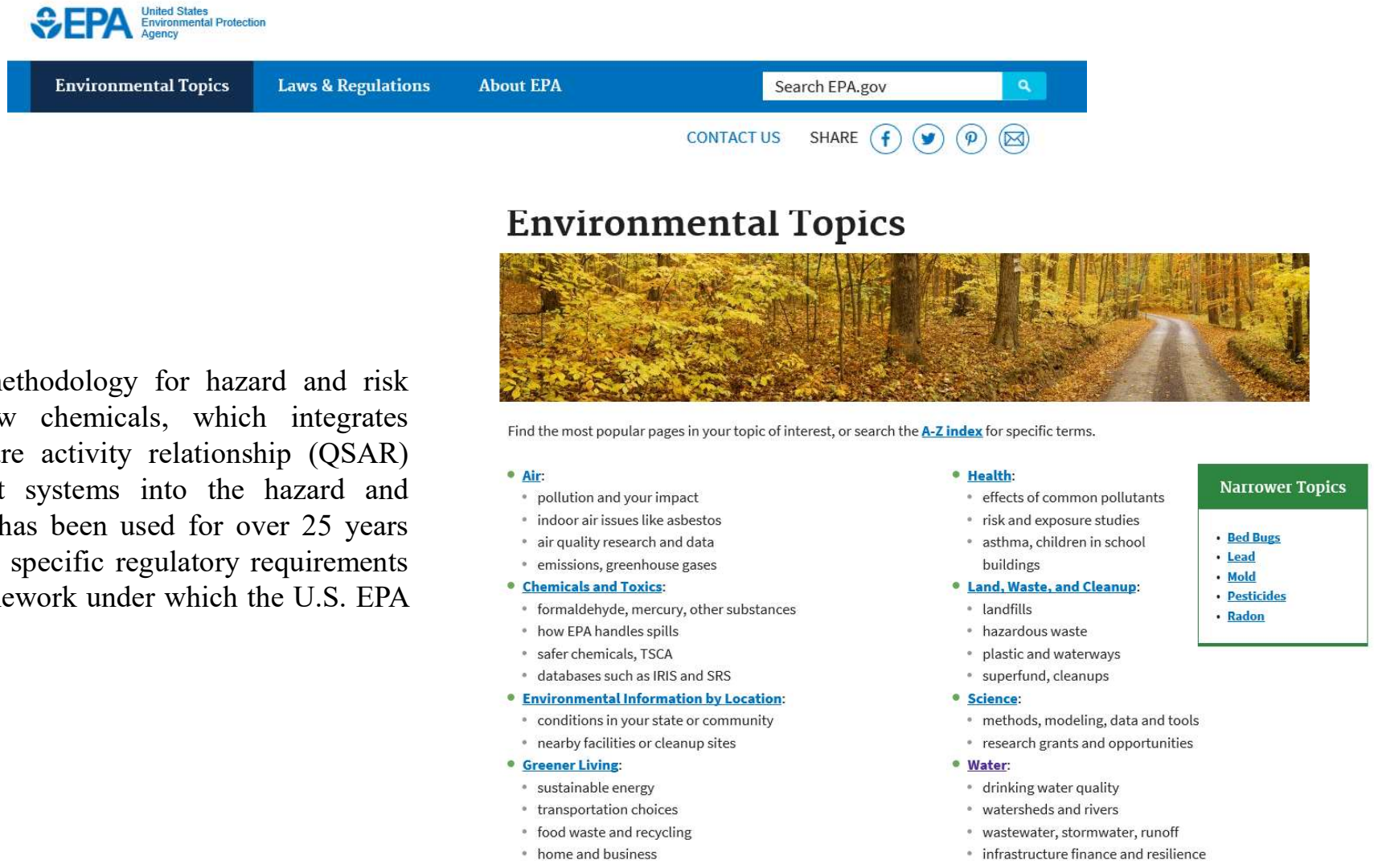
## c. Ecotoxicological endpoints

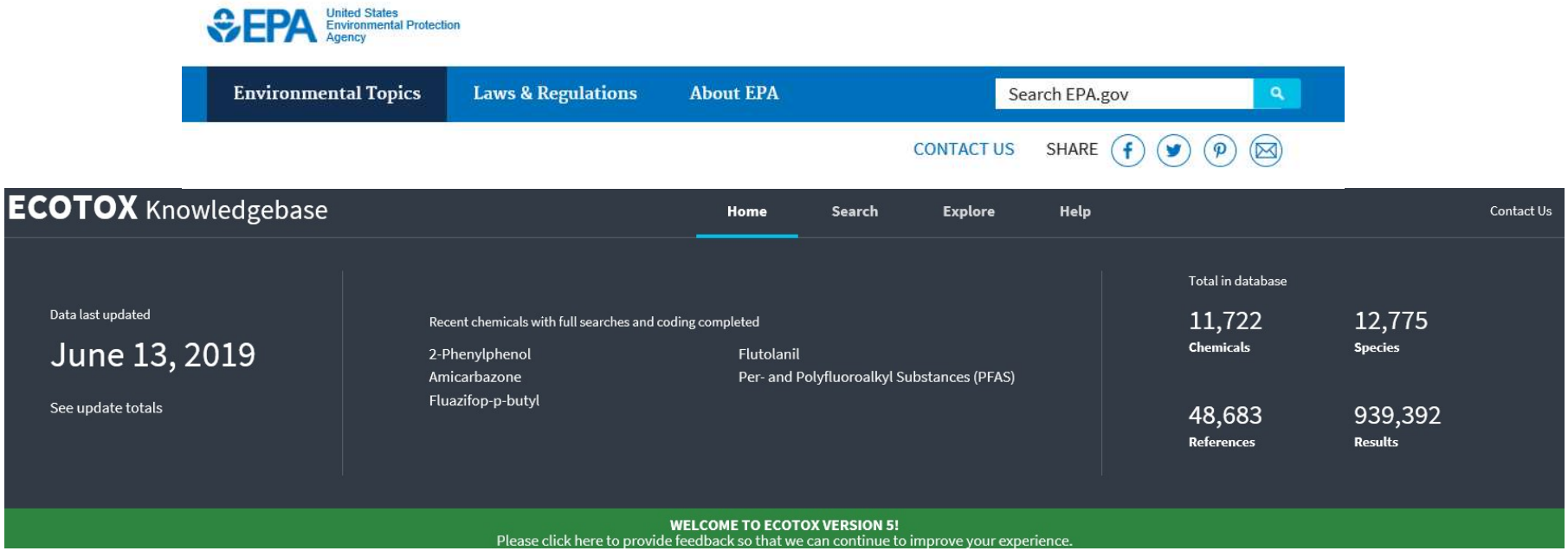
| Endpoint[12] | Software tool | Models/Modules | Free or Commercial |
|---|---|---|---|
| Short-term toxicity to fish | Danish QSAR Database (DTU) | Fathead minnow 96h LC50 from DTU | Free |
| | ECOSAR (US EPA) | Fish, 96-hr, LC50 | Free |
| | T.E.S.T. (US EPA) | Fathead minnow LC50 (96 hr) | Free |
| | VEGA (IRFMN) | SarPy/IRFMN classification and KNN/Read-Across model | Free |
| | ADMET Predictor (Simulations Plus) | Toxicity module | Commercial |
| | CASE Ultra (MultiCASE) | EcoTox model bundle | Commercial |
| | Discovery Studio (Accelrys) | Fathead Minnow LC50 | Commercial |
| Long-term toxicity to fish | ECOSAR (US EPA) | Fish, ChV[13] | Free |
| Short-term toxicity to aquatic invertebrates (daphnia) | Danish QSAR Database (DTU) | Daphnia magna 48h EC50 from DTU | Free |
| | ECOSAR (US EPA) | Daphnid, 48-hr, LC50 | Free |
| | T.E.S.T. (US EPA) | Daphnia magna LC50 (48 hr) | Free |
| | VEGA (DEMETRA) | Daphnia Magna LC50 (48 h) | Free |
| | ADMET Predictor (Simulations Plus) | Toxicity module | Commercial |
| | Discovery Studio (Accelrys) | Daphnia EC50 | Commercial |
| Long-term toxicity to aquatic invertebrates (daphnia) | ECOSAR (US EPA) | Daphnid, ChV[11] | Free |
| Toxicity to aquatic plants (algae) | Danish QSAR Database (DTU) | Pseudokirchneriella s. 72h EC50 from DTU | Free |
| | ECOSAR (US EPA) | Green Algae, 96-hr, EC50 | Free |
| Short-term toxicity to terrestrial invertebrates | ECOSAR (US EPA) | Earthworm, 14-day, LC50 | Free |

## d. Toxicological endpoints

| Endpoint | Software tool | Models/Modules | Free or Commercial |
|---|---|---|---|
| Acute toxicity | Danish QSAR Database (DTU) | Models for acute toxicity in rodents from ACD/Labs | Free |
| | T.E.S.T. (US EPA) | Oral rat LD50 | Free |
| | ACD/Percepta (ACD/Labs) | Acute Toxicity Module | Commercial |
| | ADMET Predictor (Simulations Plus) | Toxicity module | Commercial |
| | CASE Ultra (MultiCASE) | AcuteTox model bundle | Commercial |
| | Discovery Studio (Accelrys) | Rat oral LD50 and rat inhalation toxicity LC50 | Commercial |
| Skin irritation or skin corrosion | Danish QSAR Database (DTU) | Skin irritation model | Free |
| | OECD QSAR Toolbox | Skin irritation/corrosion Inclusion (and Exclusion) rules by BfR | Free |
| | ToxTree (JRC) | Skin irritation / skin corrosion | Free |
| | ACD/Percepta (ACD/Labs) | Irritation Module | Commercial |
| | CASE Ultra (MultiCASE) | SkinEye Toxicity model bundle | Commercial |
| | Derek (Lhasa) | Irritation (of the skin) alerts | Commercial |
| | Discovery Studio (Accelrys) | Skin irritancy | Commercial |
| Eye irritation | OECD QSAR Toolbox | Eye irritation/corrosion Inclusion (and Exclusion) rules by BfR | Free |
| | ToxTree (JRC) | Eye irritation and corrosion | Free |
| | ACD/Percepta (ACD/Labs) | Irritation Module | Commercial |
| | CASE Ultra (MultiCASE) | SkinEye Toxicity model bundle | Commercial |
| | Derek (Lhasa) | Irritation (of the eye) alerts | Commercial |
| | Discovery Studio (Accelrys) | Ocular irritancy | Commercial |
| Skin sensitisation | Danish QSAR Database (DTU) | Allergic Contact Dermatitis model | Free |
| | OECD QSAR Toolbox | Protein binding alerts for skin sensitisation by OASIS | Free |
| | ToxTree (JRC) | Skin sensitisation reactivity domains | Free |
| | VEGA (IRFMN) | CAESAR model | Free |
| | ACD/Percepta (ACD/Labs) | Irritation Module | Commercial |
| | CASE Ultra (MultiCASE) | SkinEye Toxicity model bundle | Commercial |
| | Derek (Lhasa) | Skin sensitisation | Commercial |
| | Discovery Studio (Accelrys) | Skin sensitization | Commercial |
| | TIMES (LMC) | Skin sensitization with autoxidation | Commercial |
| Repeated dose toxicity | ADMET Predictor (Simulations Plus) | Toxicity module | Commercial |
| | CASE Ultra (MultiCASE) | Several model bundles associated with repeated dose toxicity | Commercial |
| | Derek (Lhasa) | Several endpoints associated with repeated dose toxicity | Commercial |
| | Discovery Studio (Accelrys) | Rat Chronic (Oral) LOAEL | Commercial |
| | Leadscope | Several models associated with repeated dose toxicity | Commercial |
| *In vitro* gene mutation in bacteria (Ames test) | Danish QSAR Database (DTU) | Models for Ames test | Free |
| | OECD QSAR Toolbox | Several profilers (alerts) associated with this endpoint | Free |
| | T.E.S.T. (US EPA) | Mutagenicity | Free |
| | ToxTree (JRC) | *In vitro* mutagenicity (Ames test) alerts by ISS | Free |
| | VEGA (IRFMN) | CAESAR, SarPy/IRFMN, ISS and KNN/Read-Across models | Free |
| | ACD/Percepta (ACD/Labs) | Genotoxicity Module | Commercial |
| | CASE Ultra (MultiCASE) | Bacterial mutagenicity model bundle | Commercial |
| | Derek and Sarah (Lhasa) | Mutagenicity *in vitro* | Commercial |
| | Discovery Studio (Accelrys) | Ames Mutagenicity | Commercial |
| | Leadscope | Genetox Expert Alerts Suite and Non-human Genetic Toxicity Suite | Commercial |
| | TIMES (LMC) | Ames mutagenicity | Commercial |
| Mutagenicity (other endpoints than *in vitro* gene mutation in bacteria) | Danish QSAR Database (DTU) | Models for genotoxicity endpoints | Free |
| | OECD QSAR Toolbox | Several profilers (alerts) associated with mutagenicity | Free |
| | ToxTree (JRC) | Several decision trees associated with mutagenicity | Free |
| | CASE Ultra (MultiCASE) | EcoTox model bundle | Commercial |
| | Derek (Lhasa) | Chromosome damage *in vitro* | Commercial |
| | Leadscope | Non-human Genetic Toxicity Suite | Commercial |
| | TIMES (LMC) | Several models associated with mutagenicity | Commercial |
| Reproductive toxicity | Danish QSAR Database (DTU) | Models for Endocrine endpoints and model for Teratogenic Potential in Humans | Free |
| | VEGA (IRFMN) | CAESAR and PG models | Free |
| | ADMET Predictor (Simulations Plus) | Toxicity module | Commercial |
| | CASE Ultra (MultiCASE) | Several model bundles associated with reproductive and developmental toxicity | Commercial |
| | Derek (Lhasa) | Several endpoints associated with reproductive toxicity | Commercial |
| | Discovery Studio (Accelrys) | Developmental Toxicity Potential | Commercial |
| | Leadscope | Several models associated with reproductive and developmental toxicity | Commercial |
| | TIMES (LMC) | Androgen, AHR and Estrogen (receptor) binding affinity models | Commercial |

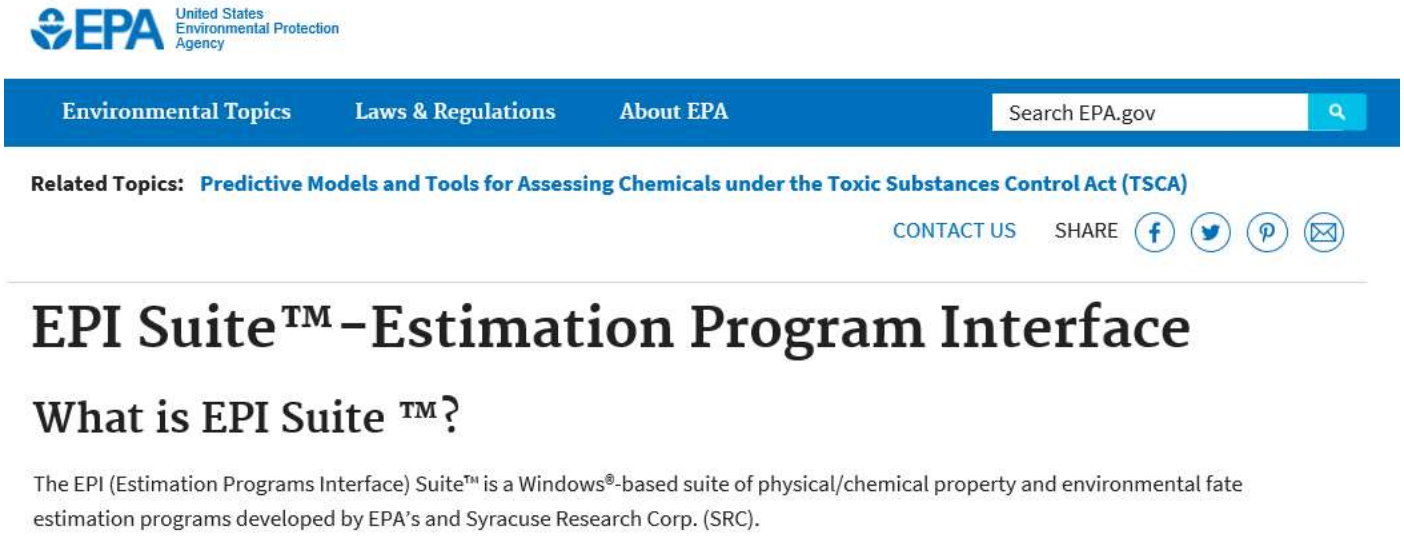
The U.S. EPA's methodology for hazard and risk assessment of new chemicals, which integrates quantitative structure activity relationship (QSAR) models and expert systems into the hazard and exposure analysis, has been used for over 25 years and reflects several specific regulatory requirements that define the framework under which the U.S. EPA must operate.



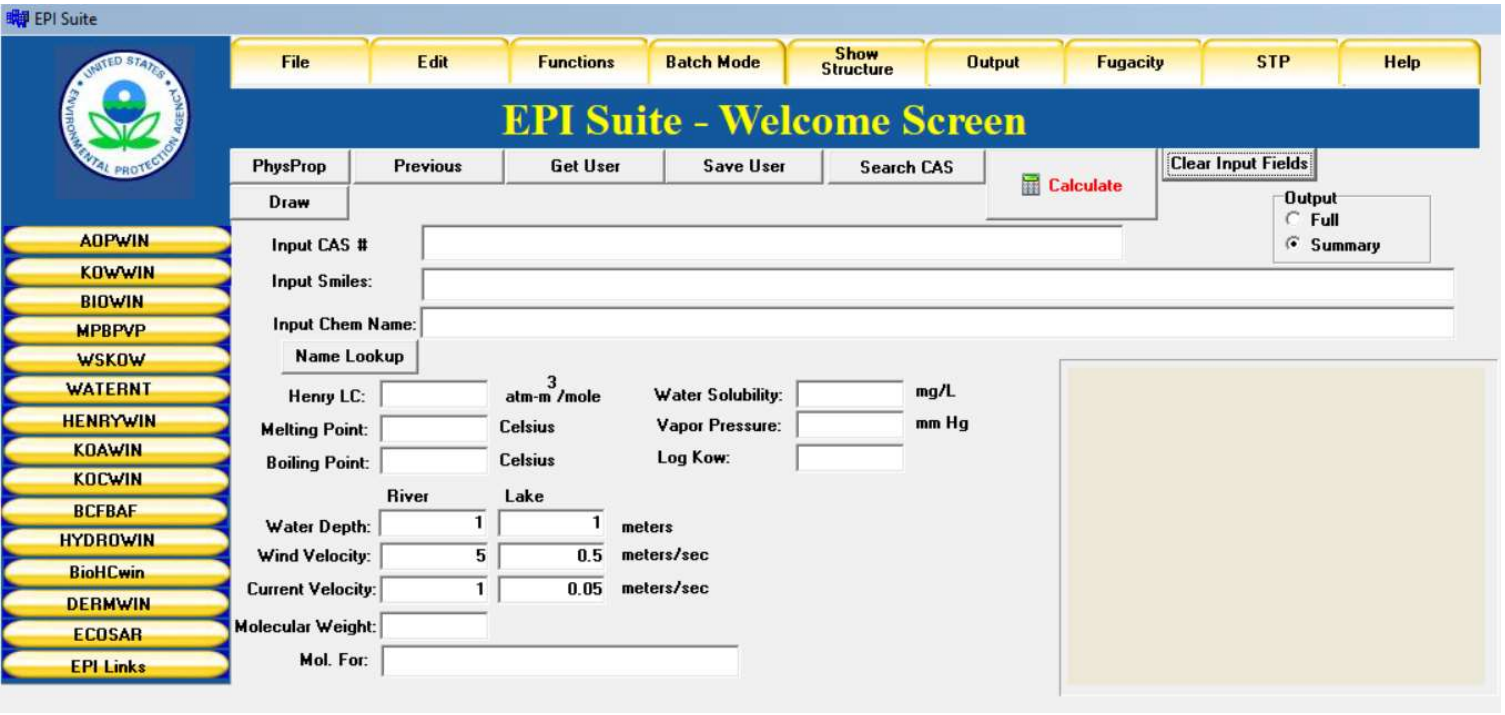



Recherche par molécule, organisme, écosystème (terrestre aqueux…), effets biologiques…

Les solutions QSAR




Utilisation relativement facile




**AOPWIN** - estimates atmospheric oxidation rates
**BCFBAF** - estimates bioconcentration factor (BCF) and biotransformation rate (kM)
**BioHCwin** - estimates biodegradation of hydrocarbons
**BIOWIN** - estimates biodegradation probability
**ECOSAR** - estimates aquatic toxicity (LD50, LC50)
**HENRYWIN** - estimates Henry's law constant
**HYDROWIN** - estimates aqueous hydrolysis rates (acid-, base-catalyzed)
**KOAWIN** - estimates octanol-air partition coefficient
**KOCWIN** - estimates soil sorption coefficient (Koc)
**KOWWIN** - estimates octanol-water partition coefficient
**MPBPVP** - estimates melting pt, boiling pt, and vapor pressure (also referred to as MPBPWIN)
**WSKOWWIN** - estimates water solubility (from log octanol-water partition coefficient)
**WATERNT** - estimates water solubility (using atom-fragment methodology)

The BCFBAF Program pertains to estimation of Bioconcentration Factor (BCF). The BCFBAF program estimates BCF of an organic compound using the compound's log octanol-water partition coefficient (Kow).

Methodology for Non-Ionic : separation of compounds into three divisions by Log Kow value as follows:

For Log Kow < 1.0 the derived QSAR estimation equation is: All compounds with a log Kow of less than 1.0 are assigned an estimated log BCF of 0.50

For Log Kow 1.0 to 7.0 the derived QSAR estimation equation is:

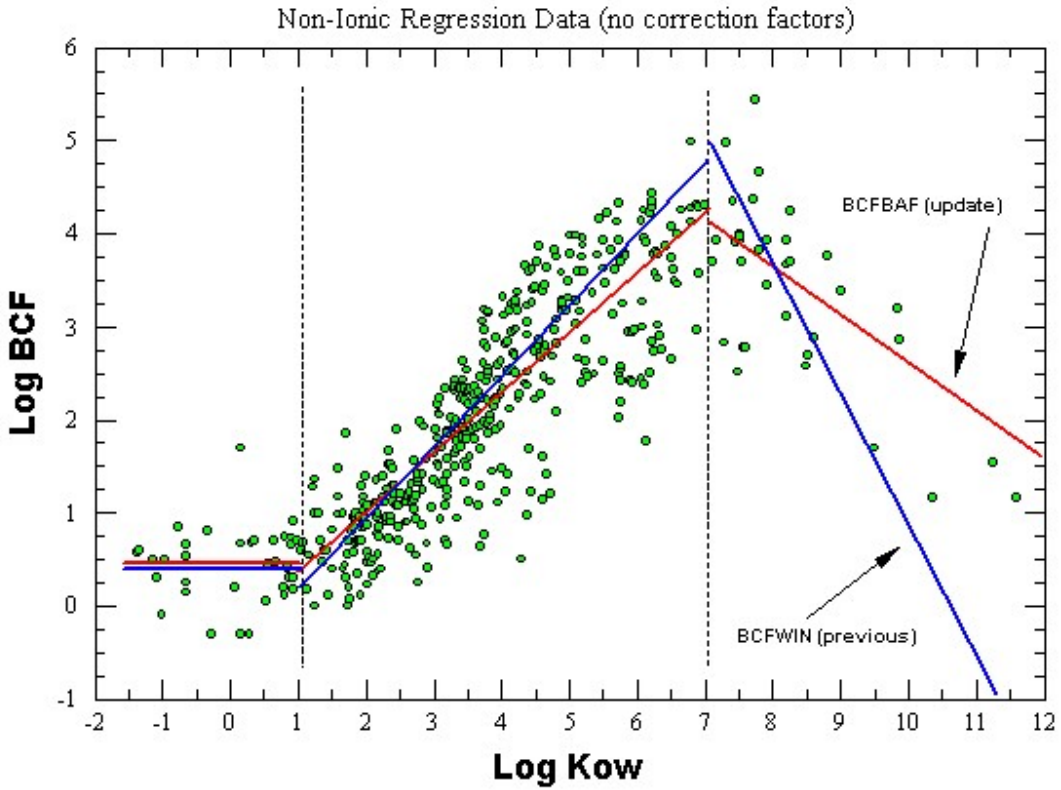
$$\text{Log BCF} = 0.6598 \text{ Log Kow} - 0.333 + \sum \text{correction factors}$$

(n = 396, $r^2$ = 0.792, $Q^2$ = 0.78, std dev = 0.511, avg dev = 0.395)

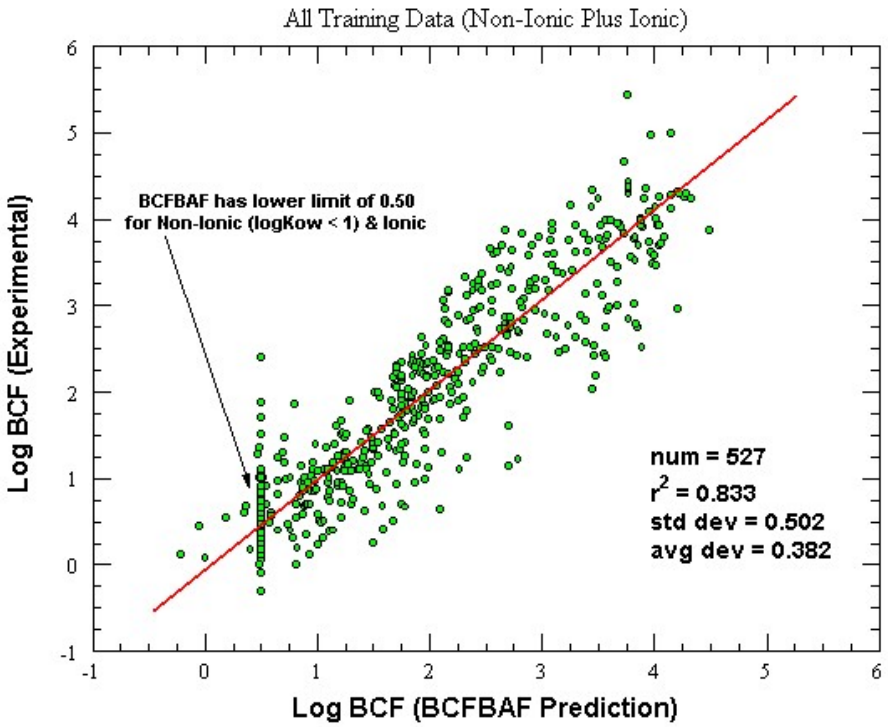
For Log Kow > 7.0 the derived QSAR estimation equation is:

$$\text{Log BCF} = -0.49 \text{ Log Kow} + 7.554 + \sum \text{correction factors}$$

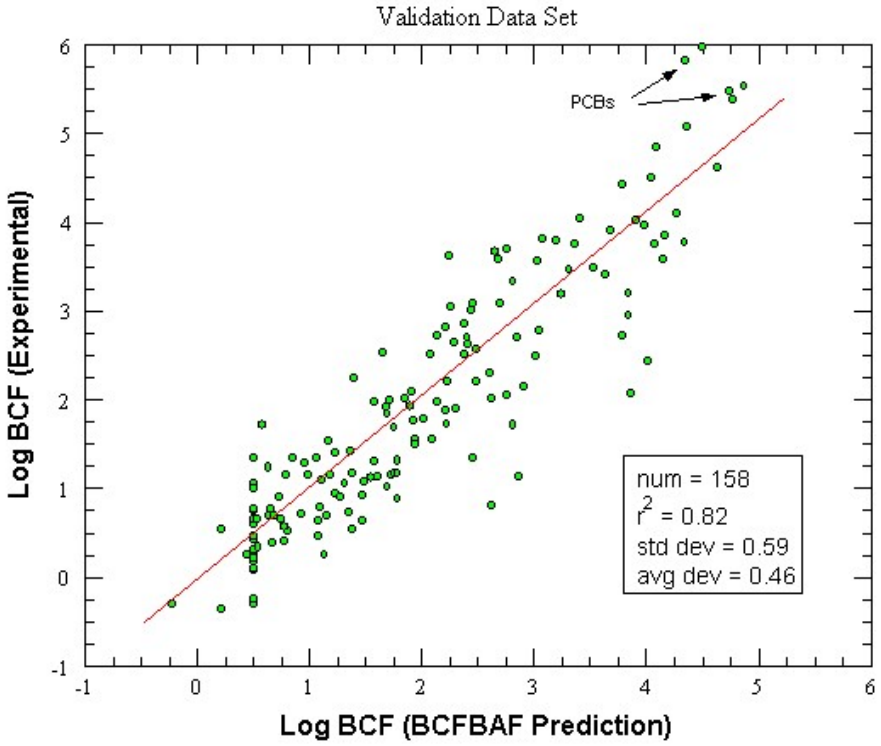
(n = 35, $r^2$ = 0.634, $Q^2$ = 0.57, std dev = 0.538, avg dev = 0.396)



Accuracy of the Training Set:




Accuracy of the Validation Set:




For freshwater fish data, species frequently include bluegill sunfish (*Lepomis macrochirus*), common carp (*Cyprinus carpio*), fathead minnow (*Pimephales promelas*), guppy (*Poecilia reticulate*), rainbow trout (*Oncorhynchus mykiss*), red killifish (*Oryzias latipes*), or zebrafish (*Brachydanio rerio*). For freshwater invertebrates, species frequently include *Daphnia magna* or *Daphnia pulex*. For freshwater algae, species frequently include *Desmodesmus subspicatus* or *Pseudokirchneriella subcapitata*. Therefore, the equations in ECOSAR are derived from surrogate species of fish, zooplankton, and phytoplankton. While these surrogate species can comprise several genera as well as families, the equations are not intended to assess toxicity to only those species, but rather to the general trophic levels they represent (fish, aquatic invertebrates, and aquatic plants)

| Endpoint | 2012 Neutral Organics Equations |
|---|---|
| Fish 96hr LC50 | Log Toxicity (mmol/L) = -0.8981(log Kow) + 1.7108 |
| Daphnid 48hr LC50 | Log Toxicity(mmol/L) = -0.8580(log Kow) + 1.3848 |
| Green Algae 96hr EC50 | Log Toxicity (mmol/L) = -0.6922(log Kow) + 0.9253 |
| Fish Chronic (ChV) | Log Toxicity (mmol/L) = -0.8508(log Kow) + 0.6063 |
| Daphnid Chronic (ChV) | Log Toxicity (mmol/L) = -0.7464(log Kow) + 0.1507 |
| Green Algae Chronic (ChV) | Log Toxicity (mmol/L) = -0.6029(log Kow) + 0.1648 |
| Fish 96 hr - Salt Water | Log Toxicity (mmol/L) = -0.8955(log Kow) + 1.8056 |
| Mysid Shrimp 96hr - Salt Water | Log Toxicity (mmol/L) = -1.1897(log Kow) + 2.2651 |
| Fish Chronic (ChV) -Salt Water | Log Toxicity (mmol/L) = -0.6303(log Kow) + 0.3108 |
| Mysid Shrimp Chronic (ChV) - SW | Log Toxicity (mmol/L) = -1.3184(log Kow) + 1.4592 |
| Earthworm 14d | Log Toxicity (mmol/L) = -0.1037(log Kow) + 0.4476 |

**Acute Effects:**
Fish 96 hr LC50
Daphnid 48 hr LC50
Algae 72 or 96 hr EC50

**Chronic Effects:**
Fish ChV
Daphnid ChV
Algae ChV

The ChV, or Chronic Value, is defined as the geometric mean of the no observed effect concentration (NOEC) and the lowest observed effect concentration (LOEC). This can be mathematically represented as: ChV = 10^([log (LOEC x NOEC)]/2)

For freshwater fish data, species frequently include bluegill sunfish (*Lepomis macrochirus*), common carp (*Cyprinus carpio*), fathead minnow (*Pimephales promelas*), guppy (*Poecilia reticulate*), rainbow trout (*Oncorhynchus mykiss*), red killifish (*Oryzias latipes*), or zebrafish (*Brachydanio rerio*). For freshwater invertebrates, species frequently include *Daphnia magna* or *Daphnia pulex*. For freshwater algae, species frequently include *Desmodesmus subspicatus* or *Pseudokirchneriella subcapitata*. Therefore, the equations in ECOSAR are derived from surrogate species of fish, zooplankton, and phytoplankton. While these surrogate species can comprise several genera as well as families, the **equations are not intended to assess toxicity to only those species, but rather to the general trophic levels they represent (fish, aquatic invertebrates, and aquatic plants)**

# Toxicity Estimation Software Tool (TEST)

The Toxicity Estimation Software Tool (TEST) was developed to allow users to easily estimate the toxicity of chemicals using Quantitative Structure Activity Relationships (QSARs) methodologies. QSARs are mathematical models used to predict measures of toxicity from the physical characteristics of the structure of chemicals (known as molecular descriptors). Simple QSAR models calculate the toxicity of chemicals using a simple linear function of molecular descriptors:

$$\text{Toxicity} = ax_1 + bx_2 + c$$

where $x_1$ and $x_2$ are the independent descriptor variables and a, b, and c are fitted parameters. The molecular weight and the octanol-water partition coefficient are examples of molecular descriptors. Additional examples are provided in our Molecular Descriptors

**Ask a Technical Expert**

Got a question about our research model? Want to give us feedback? Contact a technical expert about TEST.

- 96-hour fathead minnow 50 percent lethal concentration (LC50)
- 48-hour daphnia magna 50 percent lethal concentration (LC50)
- Tetrahymena pyriformis 50 percent growth inhibition concentration (IGC50) EXIT
- Oral rat 50 percent lethal dose (LD50) EXIT
- Bioconcentration Factor (BCF) The bioconcentration factor data set was compiled by researchers at the Mario Negri Istituto Di Ricerche Farmacologiche EXIT
- Developmental Toxicity (DevTox) EXIT
- Ames Mutagenicity (Mutagenicity) EXIT

https://www.epa.gov/chemical-research/toxicity-estimation-software-tool-test

ECHA
EUROPEAN CHEMICALS AGENCY

QSAR TOOLBOX

About Features Resources Support Developers Download

Databases Inventories Profilers Metabolisms QSARs

## DATABASES

**57 databases** containing **84 291 chemicals** with almost **2.5 million measured data points** are implemented in QSAR Toolbox 4.3.

The databases are grouped in four categories based on the type of information they are providing:

- **physical chemical properties:** 48,485 substances and 204,151 experimental values
- **environmental fate and transport:** 11,557 substances and 127,560 endpoint values
- **ecotoxicological information:** 19,842 substances and 1,029,830 endpoint values
- **human health hazards:** 39,130 substances and 1,038,214 endpoint values

The databases and the main type of metadata they contain are shown in the table below.

Databases Inventories Profilers Metabolisms QSARs

## (Q)SARS

**902 (Q)SAR models** for predicting different properties are available in QSAR Toolbox 4.3.

There are four general groups of (Q)SAR models based on the endpoint for predicting:

- **physical chemical properties:** 28 (Q)SAR models available
- **environmental fate and transport:** 41 (Q)SAR models available
- **ecotoxicological information:** 688 (Q)SAR models available
- **human health hazards:** 145 (Q)SAR models available

More information for the available (Q)SAR models could be seen following the links in the table below.

Très complet, utilisation plus complexe

## Physical-Chemical Properties

| Endpoint tree | Number of relevant (Q)SAR models |
|---|---|
| Boiling point | 1 |
| Dissociation contant (pKa) | 7 |
| Explosive properties | 1 |
| Melting / freezing point | 4 |
| Partition Coefficient | 8 |
| Vapour pressure | 5 |
| Water solubility | 2 |

## Human Health Hazards

| Endpoint tree | Number of relevant (Q)SAR models |
|---|---|
| Acute toxicity | 6 |
| Carcinogenicity | 18 |
| Developmental toxicity/Teratogenicity | 5 |
| Genetic toxicity | 68 |
| Irritation/ Corrosion | 6 |
| Photoinduced toxicity | 1 |
| Repeated dose toxicity | 4 |
| Sensitization | 8 |
| Toxicity to reproduction | 29 |

## Environmental Fate and Transport

| Endpoint tree | Number of relevant (Q)SAR models |
|---|---|
| Bioaccumulation: aquatic | 13 |
| Biodegradation | 15 |
| Photodegradationt | 4 |
| Stability in Water | 3 |
| Transport and Distribution between Environmental Compartments | 6 |

## Ecotoxicological Information

| Endpoint tree | Number of relevant (Q)SAR models |
|---|---|
| Aquatic toxicity | 683 |
| Terrestial Toxicity | 5 |