



Mise en place d'une base de données sur les propriétés des substances prioritaires

Rapport technique

Ministère de l'Ecologie et du Développement Durable
Direction de l'Eau
20, avenue de Ségur – 75302 PARIS 07 SP

Convention DE n° CV03000081 – Opération n°3

François LE GOFF

Responsable d'affaire : H. MAGAUD

*Unité « Evaluation des Risques Ecotoxicologiques »
Direction des Risques Chroniques*

15 JANVIER 2004

Mise en place d'une base de données sur les propriétés des substances prioritaires

Rapport technique

Ministère de l'Ecologie et du Développement Durable
 Direction de l'Eau
 20, avenue de Ségur – 75302 PARIS 07 SP

Convention DE n° CV03000081 – Opération n°3

15 JANVIER 2004

PERSONNES AYANT PARTICIPE A L'ETUDE :

M. CHOUGRANI - F. LE GOFF - H. MAGAUD - K. MOIREZ - O. SAINT-JEAN - V. VEG

Ce document comporte 24 pages (hors couverture et annexes).

	Rédaction	Vérification	Approbation
NOM	F. LE GOFF	H. MAGAUD	M. COQUERY
Qualité	Ingénieur à l'Unité Evaluation des Risques Ecotoxicologiques	Ingénieur à l'Unité Evaluation des Risques Ecotoxicologiques	Ingénieur à l'Unité Chimie Analytique et Environnementale
Visa			

Table des matières

Glossaire	4
Résumé	5
Introduction	6
1. PARCOURS DES BASES DE DONNÉES PRÉEXISTANTES	7
1.1. LES BASES DE DONNÉES EN LIGNE	7
1.1.1. AGRITOX	7
1.1.2. European Chemical Bureau : dossiers d'évaluation des substances existantes	8
1.1.3. OECD Integrated HPV Database	8
1.1.4. OSPAR	9
1.1.5. IPCS (OMS)	10
1.1.6. AQUIRE	10
1.2. LES BASES DE DONNÉES HORS LIGNE	10
1.2.1. IUCLID	11
1.2.2. Nordic Substance DataBase (NSDB)	11
1.3. BILAN	12
2. MISE EN PLACE D'UNE BASE DE DONNÉES ENVIRONNEMENTALES	13
2.1. RAPPEL DES OBJECTIFS	13
2.2. SUBSTANCES CONCERNÉES	13
2.2.1. Substances UE	13
2.2.2. Substances OCDE	13
2.2.3. Substances OSPAR	14
2.2.4. Substances INERIS	14
2.3. CARACTÉRISATION DES DONNÉES	14
2.4. POINTS TECHNIQUES CONCERNANT LA CRÉATION DE LA BASE	15
2.5. OUTIL DE SAISIE DES INFORMATIONS	15
2.5.1. Menu général	16
2.5.2. Formulaires secondaires	16
2.5.3. Formulaire principal : gestion des dossiers	17
2.6. FORMAT DES DONNÉES EN SORTIE	18
2.7. OUTIL DE RECHERCHE	18
2.7.1. Présentation générale	18
2.7.2. Utilisation du moteur de recherche : recherche par substance (affichage des « fiches »)	20
2.7.3. Utilisation du moteur de recherche : recherche par type de données (affichage des « catalogues »)	22
Conclusion	23
Annexes	24

Glossaire

AQUIRE : AQUatic toxicity Information REtrieval database

DRC : Direction des Risques Chroniques

ECB : European Chemical Bureau

EINECS : European Inventory of Existing Commercial Chemical Substances

EUSES : European Union System for the Evaluation of Substances

HPV : High Production Volume

ICCA : International Council of Chemical Associations

IPCS : International Programme on Chemical Safety

IUCLID : International Uniform Chemical Information Database

IUPAC : International Union of Pure and Applied Chemistry

NICNAS : National Industrial Chemicals Notification and Assessment Scheme (Australia)

NSDB : Nordic Substance DataBase

OECD : Organisation for Economic Cooperation and Development

OSPAR : Convention Oslo Paris

QSAR : Quantitative Structure – Activity Relationship

RAR : Risk Assessment Report

SIAM : SIDS Initial Assessment Meeting

SIAP : SIDS Initial Assessment Profile

SIAR : SIDS Initial Assessment Report

SIDS : Screening Information DataSet

TSCA : Toxic Substances Control Act (US EPA)

UNEP : United Nations Environment Programme

WHO : World Health Organization (Organisation Mondiale de la Santé : OMS)

Résumé

Depuis plusieurs années des démarches sont engagées, par différents pays ou organisations (Union Européenne, OCDE, France – Ineris etc.), pour évaluer les risques environnementaux posés par certaines substances chimiques. Le principe général de ces études consiste à se baser sur l'état des connaissances, au niveau de l'exposition et des effets des substances, pour caractériser le risque, après validation des données utilisées par des comités d'experts.

A l'heure actuelle, diverses bases de données consultables sur Internet proposent un accès à des documents d'évaluation des risques sur les substances. Cependant les données accessibles ne concernent qu'un nombre de substances limité, étudiées dans un cadre particulier. De plus, les informations sur les propriétés physico-chimiques et écotoxicologiques apparaissent soit dans des documents différents, sur des sites différents, soit, lorsqu'elles figurent dans un même document, ne sont pas clairement mises en évidence dans le texte.

La base de données mise au point par l'INERIS vise la valorisation des informations concernant les propriétés physico-chimiques et écotoxicologiques des substances chimiques. En effet, il est important de diffuser à un large public des données sur les substances validées par différentes instances internationales. Cette base propose ainsi :

- un « mélange » de toutes les sources de données disponibles sur les substances chimiques (dans le domaine de l'évaluation des risques). Il n'est ainsi plus nécessaire de naviguer sur différents sites pour extraire des informations des documents disponibles,
- la compilation de données, si possible validées, est présentée sous une forme condensée : création de fiches synthétiques, pour chaque substance,
- la possibilité d'obtenir des informations sur une liste de substances (e. g. $\log K_{ow}$ et PNEC pour les substances inscrites à l'annexe X de la directive cadre sur l'eau),
- La compilation, dans une base, de données physico-chimiques et écotoxicologiques validées permettra, via des analyses sur un grand nombre de substances, de mieux appréhender les liens possibles entre les paramètres (e. g. développement de QSAR¹, "validation" de la théorie du coefficient de partage pour les sédiments).

La base de données a été mise en ligne et rendue accessible au public début décembre 2003. Elle est consultable à l'adresse suivante :

http://chimie.ineris.fr/lien/basededonnees/environnementale/recherche/search1_1.php

Au cours de l'année 2003, 300 dossiers regroupant des informations sur les substances prioritaires ont été traités de manière à proposer, sur le site Internet, un nombre de "fiches substance" correspondant.

Cette première année, qui a vu la création et la mise en ligne de la base de données, a permis de mettre à disposition du public francophone un nombre conséquent d'informations sur les substances chimiques. Cependant, toutes les sources d'informations retenues n'ont pas encore pu être intégrées. Un travail continu d'alimentation de la base accompagné d'une activité de mise à jour des données entrées antérieurement, qui s'amplifiera avec la taille de la base, devra être mené dans les années à venir. Ceci afin de garantir, au cours du temps, pour les utilisateurs, un niveau de validité équivalent à celui proposé lors de la mise en ligne de la base.

¹ Quantitative Structure Activity Relationship

Introduction

Depuis plusieurs années des démarches sont engagées, par différents pays ou organisations (Union Européenne, OCDE, France – Ineris etc.), pour évaluer les risques environnementaux posés par certaines substances chimiques. Le principe général de ces études consiste à se baser sur l'état des connaissances, au niveau de l'exposition et des effets des substances, pour caractériser le risque, après validation des données utilisées par des comités d'experts.

Les dossiers réalisés dans ce cadre, qu'ils soient finalisés ou toujours en cours d'évaluation, sont, pour la plupart, disponibles en ligne :

- rapports d'évaluation des risques des substances existantes sur le site de l'ECB²
- rapports SIDS sur le site de l'OCDE³ ou de l'UNEP⁴

La base de données mise au point par l'INERIS⁵ vise la valorisation des informations concernant les propriétés physico-chimiques et écotoxicologiques des substances chimiques. Celles-ci sont contenues dans différents dossiers d'évaluation des risques auxquels l'Institut participe dans le cadre de l'évaluation des risques des substances existantes (Dir. CE/793/93), l'évaluation des dangers des substances étudiées par l'OCDE, etc. Il est important de diffuser à un large public des données sur les substances validées par différentes instances internationales. Cette valorisation passe par une gestion de ces données grâce à la création d'une base.

Dans un premier temps, une recherche des bases de données existantes en la matière a été réalisée. Le but étant d'éviter la mise au point d'un produit qui ferait "doublet" avec les outils déjà disponibles pour les évaluateurs des risques. Cette étape primordiale a permis de définir les besoins réels et donc de décider du contenu et de l'organisation de la base de données. Un aperçu des différentes bases de données existantes est présenté dans la première partie de ce rapport.

Une deuxième partie présente le contenu de la base de données, sa structure et enfin les outils mis en ligne via Internet pour que le grand public puisse extraire les données qu'il recherche concernant les substances chimiques.

² <http://ecb.jrc.it/existing-chemicals/> rubrique "Documents \ Public access \ Risk Assessment"

³ http://www.oecd.org/document/63/0,2340,en_2649_34379_1897983_1_1_1_1,00.html

⁴ http://www.chem.unep.ch/irptc/Publications/sidsindex/sidsindex_chem.htm

⁵ http://chimie.ineris.fr/lien/basededonnees/environnementale/recherche/search1_1.php

1. Parcours des bases de données préexistantes

La construction d'une base de données contenant des informations sur les substances chimiques en terme d'évaluation des risques environnementaux nécessite une bonne connaissance préalable des outils similaires disponibles et facilement accessibles. Dans un premier temps, les bases de données disponibles sur Internet, correspondant aux critères de recherches définis, ont été recensées et étudiées.

Les banques de données compulsées dans cette partie ont été sélectionnées de manière à être comparées à celle mise au point. En effet, il existe de nombreuses bases de données regroupant des informations sur les substances chimiques mais seulement quelques-unes fournissent les données nécessaires à l'évaluation des risques.

Cet état des lieux a été effectué au premier trimestre 2003.

1.1. Les bases de données en ligne

1.1.1. AGRITOX

Accès : accès à la base : <http://www.inra.fr/agritox/>. Cette base est gérée par l'INRA (Institut National de la Recherche Agronomique) et a été créée par le département de Phytopharmacie et d'Ecotoxicologie de l'INRA. L'accès vers la procédure d'élaboration des fiches (contenu et définitions des différentes rubriques) se fait à l'adresse suivante : <http://www.inra.fr/agritox/guides/guide-agritox.html>.

Contenu : sous forme de fiches, par substance

- Identité de la substance : synonymes, famille chimique, activité biologique, acteurs industriels
- Données physico-chimiques
- Toxicité : aiguë, à long terme, reproduction et développement
- Ecotoxicité sur la faune et la flore sauvage (aiguë et chronique)
- Devenir dans l'environnement
- Données réglementaires des substances actives phytopharmaceutiques autorisées en France : classement toxicologique, dose journalière acceptable, niveau acceptable d'exposition pour l'opérateur, limite maximale de résidus, dose de référence aiguë

Origine des données

- Dossiers toxicologiques de demande d'homologation déposés par les industriels au niveau français et européen (80 %)
- Source bibliographique (20 %)

Outil de recherche : la liste alphabétique des fiches disponibles. Les noms des substances proposées sont accompagnés des sources d'où proviennent les données.

Mise à jour : les mises à jour ont lieu après chaque séance de la commission française d'étude de la toxicité des produits antiparasitaires à usage agricole.

Format des données : publiées sur Internet, les fiches sont au format HTML.

Exemple : voir l'exemple de l'atrazine en annexe (annexe 1).

1.1.2. European Chemical Bureau : dossiers d'évaluation des substances existantes

Accès : accès au site général : <http://ecb.jrc.it/existing-chemicals/> (voir, sur le site de l'ECB, la section « Documents \ Public Access \ Risk Assessment »).

Contenu : il n'existe pas, en ligne, de véritable base de données sur les substances chimiques. En revanche, divers documents sont accessibles sur Internet : les évaluations des risques publiées au niveau européen ainsi que les versions en projet, le guide technique européen pour l'évaluation des substances chimiques (TGD : Technical Guidance Documents), différents textes réglementaires, etc.

1.1.3. OECD Integrated HPV Database

Accès : site web : <http://cs3-hq.oecd.org/scripts/hpv/>.

Contenu : le contenu proposé reprend, sous forme de fiches, les informations obtenues par la procédure SIDS (Screening Information Data Set). Ce programme établit le jeu de données minimum nécessaire à une première évaluation d'une substance. Les informations suivantes apparaissent pour chaque fiche :

- Identité du composé : nom, synonymes, n° CAS
- Informations référencées par l'état "sponsor" de la substance :
 - Catégories d'utilisation
 - Informations additionnelles sur les usages
 - Volume produit
 - Information sur le danger (substance reconnue ou non comme peu dangereuse)
 - Rapports existants (affichage d'une liste mais pas d'accès aux différentes publications)
 - Prise en charge ou non du dossier par l'ICCA⁶
 - Code permettant d'identifier le projet
 - Référence dans Exichem⁷ et lien vers cette base de données
 - Informations concernant l'avancement du dossier d'étude de la substance
 - Lien vers les documents SIAR⁸ et SIAP⁹

Origine des données : informations sur les substances chimiques produites en grandes quantités évaluées dans le cadre d'un programme de l'OCDE. Les informations sont fournies par les états membres qui « parrainent » la molécule en question. Pour les substances n'ayant pas encore été retenues pour une évaluation, une note spécifique est rédigée par le secrétariat de l'OCDE qui gère cette base.

⁶ International Council of Chemical Associations

⁷ base de données mise en place pour la gestion des dossiers étudiés dans le cadre de l'OCDE. Cette base fournit des informations sur les responsables du dossier ainsi que sur son état d'avancement.

⁸ SIDS Initial Assessment Report

⁹ SIDS Initial Assessment Profile

Outil de recherche : plusieurs modes de recherche sont proposés :

- Par nom
- Numéro CAS
- Pays sponsor
- Niveau actuel d'investigation
- Catégories de substances

Mise à jour : les informations sont communiquées par les états membres et centralisées au secrétariat de l'OCDE, gestionnaire de la base de données.

Format des données : la présentation générale est faite au format HTML auquel sont incorporés des scripts de recherche. Les produits de sortie sont au format PDF interactifs (liens web et internes au fichier) pour les documents SIDS.

1.1.4. OSPAR

Accès : l'accès à la base de données OSPAR (conventions Oslo Paris) se fait par l'adresse suivante : <http://www.ospar.org/fr/html/welcome.html>

Contenu : le site propose l'ensemble des fiches caractéristiques des substances étudiées par OSPAR (liste des substances potentiellement préoccupantes).

- Identification par nom, n° CAS, n° EINECS
- Raison de la sélection dans la liste (critères PBT, perturbateurs endocriniens, autres raisons – notion de « filet de sécurité »)
- Identification des travaux sous OSPAR
- Propriétés des substances (lorsque ces substances ont été validées au niveau européen – CE/93/793 – seules les données validées sont indiquées).

Dans les fiches, les points caractéristiques, ayant conduit à l'inscription de la substance à la liste des substances potentiellement préoccupantes, sont mis en évidence.

Les différentes informations apparaissant dans les fiches des substances sont classées et présentées selon le schéma suivant : Paramètre / Valeur / Source et référence / Remarques.

Origine des données : les propriétés indiquées proviennent pour l'essentiel des bases de données nationales des parties contractantes.

- NSDB : base de données des pays nordiques (Cf. point 1.2.2).
- Utilisation de divers modèles pour le calcul des QSAR¹⁰ (dont EPIWIN : <http://esc.syrres.com/>).
- Des résultats de QSAR ont été extraits des bases de données des Pays Bas et du Danemark (base de données sur les QSAR de l'agence danoise de protection de l'environnement – EPA).

Outil de recherche : la recherche peut se faire à l'aide des critères suivants :

- Par nom de substance
- Par n° CAS
- Par groupe d'utilisation ou fonction de la molécule

¹⁰ QSAR : Quantitative Structure Activity Relationship (données obtenues par modélisation). Pour le classement prioritaire des substances, une plus grande attention a été portée aux informations provenant de sources expérimentales.

Mise à jour : il s'agit d'une liste de substances dynamique. Elle est revue, après validation du groupe d'experts OSPAR, par le secrétariat de la commission (dernière mise à jour 28 juin 2002).

Format des données : le téléchargement des fiches se fait sous format compressé (.zip). Les fiches sont proposées au format Excel¹¹.

Exemple : voir l'exemple de l'antracène (annexe 2).

1.1.5. IPCS (OMS)

Accès : l'accès se fait par le site de l'Organisation Mondiale de la Santé : <http://www.who.int/fr/index.html> ou celui de l'IPCS (International Programme on Chemical Safety) : <http://www.who.int/pcs/index.htm>.

Contenu : IPCS propose, en libre accès, une documentation sur les évaluations des risques, essentiellement pour la santé humaine mais certains rapports incluent aussi des considérations pour l'environnement : http://www.who.int/pcs/ra_main.html.

Origine des données : monographies EHC (Environmental Health Criteria) : http://www.who.int/pcs/ra_site/ehc.html. La méthodologie d'évaluation des risques est développée conjointement avec l'OCDE. Les publications concernant ce point sont consultables en ligne : http://www.who.int/pcs/pubs/pub_meth.htm.

Outil de recherche : la recherche sur les EHC se fait selon différents moyens :

- Sélection parmi les EHC en cours d'élaboration
- Liste alphabétique des produits chimiques
- Choix par numéro de dossier

Format des données : la consultation des résumés ou des dossiers complets a lieu au format html pour la plupart des documents. Certains (les plus récents) sont au format Adobe Acrobat Reader¹² (.pdf).

1.1.6. AQUIRE

La base de données AQUIRE (AQUatic Information Retrieval) recense des informations, extraites de la littérature, sur les effets toxiques des substances chimiques sur les organismes. Il s'agit d'une compilation de résultats d'essais écotoxicologiques de laboratoire, sans validation particulière. Les informations contenues dans cette base doivent par conséquent être validées avant leur utilisation pour évaluer le risque représenté par une substance.

Cette base, mise au point par l'US EPA (US Environmental Protection Agency), est accessible à l'adresse suivante : <http://www.epa.gov/ecotox/>.

1.2. Les bases de données hors ligne

Dans un second temps, des bases de données proposées hors ligne (le plus souvent sur support cd-rom) ont été explorées. En effet, ce type de produit est quelquefois amené à évoluer vers une diffusion plus large (en ligne). C'est pourquoi, les bases de données utilisées lors de l'évaluation des risques environnementaux des substances, ou regroupant les résultats de telles évaluations, ont été parcourues.

¹¹ Microsoft Corp.

¹² <http://www.adobe.com/products/acrobat/readmain.html>

1.2.1. IUCLID

IUCLID (International Uniform Chemical Information Database) est l'outil de base nécessaire à la collecte d'informations sur les substances chimiques dans le cadre des évaluations des risques européennes (ce système est également utilisé, depuis 1999, par l'OCDE dans le cadre de son programme sur les substances chimiques existantes). Cette base est commune aux différents acteurs dans ce domaine : les états membres, l'industrie et le bureau européen des substances chimiques (ECB).

Les substances visées par cette collecte d'informations sont celles produites ou importées à plus de 1000 tonnes par an en Europe (HPVC : High Production Volume Chemicals – Annexe I de la directive CE/793/93). Dans la liste des substances existantes (liste EINECS), à l'exception des pesticides et médicaments, 1884 substances remplissent ce critère (sur 100195 substances existantes) pour lesquelles les producteurs et importateurs sont tenus de soumettre des informations.

Dans une seconde phase, les industriels ont indiqué les HPVC non listés dans EINECS. La nouvelle liste mise au point est connue sous le nom de EU-HPVC.

La troisième phase a élargi le champ des substances à renseigner en incluant celles produites ou importées à des tonnages annuels compris entre 10 et 1000 tonnes (tonnage mis sur le marché européen).

Au niveau européen, les industriels peuvent intégrer leurs informations à la base de données IUCLID via un programme appelé HEDSET (Harmonized Electronic Data SET). En ce qui concerne le programme de l'OCDE, les informations recueillies servent à générer des rapports appelés : SIDS.

A l'heure actuelle, la base IUCLID de l'ECB recueille environ 30000 dossiers sur près de 10500 substances.

IUCLID se décline en deux outils distincts :

- Une version complète de la base qui permet aux utilisateurs autorisés de mettre à jour des informations concernant les substances chimiques. Cette base est utilisée dans une configuration Client – Serveur.
- Une version, disponible sur cd-rom, qui ne permet qu'une simple consultation des informations répertoriées. Cette version est disponible sous deux formes qui requièrent des autorisations différentes : une version non-confidentielle (la version 2000 contient des informations sur 2604 substances : renseignements sur la substance fournis par les industriels ainsi que les éventuels rapports d'évaluation), diffusée à tous les demandeurs, et une version incluant des données confidentielles qui n'est accessible qu'aux personnes autorisées.

Le programme IUCLID fonctionne sous environnement Windows et permet un accès aux dossiers sous format pdf.

1.2.2. Nordic Substance DataBase (NSDB)

Accès : cette base est en cours de finalisation et d'actualisation. Elle a été diffusée sur cd-rom auprès des parties contractantes à OSPAR, en mars 2000. Un accès en ligne a été mis en place fin 2002. Cet accès permet le téléchargement de la dernière version de la base de données (décembre 2002) au format Microsoft Access (<http://www.norden.org/miljoe/sk/nsdp.asp>).

Contenu : cette base de données a été mise au point pour faciliter les premières études des risques environnementaux engendrés par une substance chimique dans le cadre de la priorisation des substances dangereuses organisée par OSPAR. Les critères de cette évaluation sont basés sur les effets et l'exposition de ces substances. Une banque de données réunissant toutes ces caractéristiques a donc été créée de manière à regrouper les informations de ce type déjà présentes dans d'autres bases de données. Après cette phase de récupération, plusieurs outils de sélection et de recherche ont été intégrés de manière à faciliter le travail des utilisateurs.

Les catégories d'informations répertoriées sont les suivantes :

- Toxicité aquatique
- Bioaccumulation
- Biodégradation
- Toxicité aquatique calculée (QSAR)
- Biodégradation calculée (QSAR)
- Exposition
- Empoisonnement secondaire

Parallèlement à ce stockage d'informations, la base de données permet d'appliquer des grilles d'interprétation des données de manière à obtenir un score de risque pour une substance sélectionnée. Il s'agit d'un outil d'aide à l'affectation de priorités aux substances chimiques.

Origine des données : cette base de données en regroupe plusieurs autres :

- IUCLID
- AQUIRE
- Sunset (Suède)
- Envichem (Finland), etc.

Y sont également listés différents résultats tirés de la bibliographie scientifique ainsi que d'autres informations venant d'ONG ou de l'industrie. Toutes les molécules apparaissant sur les listes de l'Union Européenne sont également incluses. Aucune de ces informations n'est validée.

Format des données : l'organisation de cette base s'est faite sous Access¹³ 97. Les produits de sortie susceptibles d'être créés sont donc ceux autorisés par ce logiciel.

1.3. Bilan

Les diverses **bases de données consultables sur Internet** étudiées ici proposent, pour la plupart, un accès à des documents d'évaluation des risques sur les substances. On peut d'ors et déjà noter les principaux manques de ces bases de données existantes :

- Les données proposées ne concernent qu'un nombre de substances limité. De plus, les sites internet visités ne proposent des informations que sur les substances étudiées dans un cadre particulier.
- Les informations sur les propriétés physico-chimiques et écotoxicologiques sont "dispersées", c'est à dire que :
 - soit elles apparaissent dans des documents différents, sur des sites différents,
 - soit elles figurent dans un même document mais ne sont pas clairement mises en évidence dans le texte.

¹³ Microsoft Corp.

En parallèle aux manques des bases de données existantes il est possible d'indiquer les avantages de la base de données faisant l'objet du projet présenté ici :

- La base propose un « mélange » de toutes les sources de données disponibles sur les substances chimiques (dans le domaine de l'évaluation des risques). Il n'est ainsi plus nécessaire de naviguer sur différents sites pour extraire des informations des documents disponibles.
- La compilation de données, si possible validées, est présentée sous une forme condensée : création de fiches synthétiques, pour chaque substance.
- La possibilité d'obtenir des informations sur une liste de substances (e. g. log K_{ow} et PNEC pour les substances inscrites à l'annexe X de la directive cadre sur l'eau).
- La compilation, dans une base, de données physico-chimiques et écotoxicologiques validées permettra, via des analyses sur un grand nombre de substances, de mieux appréhender les liens possibles entre les paramètres (e. g. développement de QSAR¹⁴, "validation" de la théorie du coefficient de partage pour les sédiments).

2. Mise en place d'une base de données environnementales

http://chimie.ineris.fr/lien/basededonnees/environnementale/recherche/search1_1.php

2.1. Rappel des objectifs

- Classement et référencement des données disponibles sur les dossiers européens, internationaux (OCDE, OSPAR), etc.
- Diffusion de ces informations à un large public (Internet). Les produits de sortie de la base de données ainsi que toute la documentation associée seront disponibles en français. La mise en place d'un accès en anglais devra être envisagé par la suite.

2.2. Substances concernées

2.2.1. Substances UE

Cette catégorie comprend toutes les substances qui bénéficient d'une procédure d'évaluation des risques environnementaux au niveau européen. Cela comprend les substances prioritaires pour l'évaluation des substances existantes. Elles apparaissent au sein de différentes listes prioritaires établies dans le cadre de la législation sur l'évaluation des substances existantes énoncée dans la directive européenne n°793/63 (principes d'évaluation des risques pour l'homme et l'environnement dans la directive CE/94/1488). Lien vers les listes sur le site de l'ECB : <http://ecb.jrc.it/existing-chemicals/>, section « Priority lists ».

2.2.2. Substances OCDE

Liste des substances « parrainées » par un pays membre de l'OCDE (substances SIDS). La liste des substances est disponible à : <http://cs3-hq.oecd.org/scripts/hpv/>.

¹⁴ Quantitative Structure Activity Relationship

2.2.3. Substances OSPAR

Liste des substances potentiellement préoccupantes établie grâce au processus DYNAMEC et la sélection de substances répondant aux critères PBT, pour la plupart. Consultation de la liste complète sur le site d'OSPAR : <http://www.ospar.org/fr/html/welcome.html>, section « Liste des substances potentiellement préoccupantes ».

2.2.4. Substances INERIS

Substances chimiques ayant fait l'objet d'une étude spécifique par l'INERIS, dans le cadre d'une évaluation des risques environnementaux.

2.3. Caractérisation des données

Les caractéristiques des données intégrées dans la base de données environnementales sont **détaillées dans un guide, en annexe 3**. Elles sont également récapitulées dans le Tableau 1.

Tableau 1 : Récapitulatif des informations contenues dans la base de données

Catégorie	Information
Identification de la substance	N° CAS
	N° EINECS
	Nom commun
	Classification environnement
	Classification PBT
	Perturbateur endocrinien
	Usages (catégories IC/UC, EUSES)
	Apparition dans la réglementation
Propriétés physico-chimiques	Poids moléculaire
	Point d'ébullition
	Densité
	Kow ou log
	Koc
	Pression de vapeur
	Solubilité
	Constante de Henry
	Biodégradation (facile, inhérente, ou rien)
	Persistance eau douce : temps ½ vie
	Persistance eau marine : temps ½ vie
	Persistance sol : temps ½ vie
	Persistance atmosphère (photolyse) : temps ½ vie
	Persistance atmosphère (oxydation radicaux OH) : temps ½ vie
	Bioaccumulation (BCF)
	Coefficient de partage eau / sol
Coefficient de partage eau / sédiment	
Coefficient de partage eau / matières en suspension	
Ecotoxicologie pour : Eau douce Eau marine Sol	Résultat algues (microorganismes du sol pour le sol) aigu validé et utilisé pour la détermination de la PNEC
	Résultat invertébrés aigu validé et utilisé pour la détermination de la PNEC
	Résultat poissons (plantes pour le sol) aigu validé et utilisé pour la détermination de la PNEC
	Résultat algues (microorganismes du sol pour le sol) chronique validé et utilisé pour la détermination de la PNEC chronique

Catégorie	Information
	Résultat invertébrés chronique validé et utilisé pour la détermination de la PNEC chronique
	Résultat poissons (plantes pour le sol) chronique validé et utilisé pour la détermination de la PNEC chronique
	Facteur d'extrapolation utilisé
	PNEC chronique
	Facteur d'extrapolation utilisé pour la PNEC statistique
	PNEC statistique
	Moyenne
	Ecart type
Ecotoxicologie sédiments	Résultat organismes benthiques aigu validé et utilisé pour la détermination de la PNEC
	Résultat organismes benthiques chronique validé et utilisé pour la détermination de la PNEC chronique
	Facteur d'extrapolation sédiment utilisé
	PNEC sédiment expérimentale
	PNEC sédiment coefficient de partage
	PNEC sédiment sélectionnée
Ecotoxicologie Micro organismes	Résultat d'essai utilisé
	PNEC station d'épuration
Source globale pour chaque fiche	Source des données : <ul style="list-style-type: none"> • Finalisé ou projet • Date • Lien vers le site ou les fiches
	Niveau de validation
	Date de modification des données

2.4. Points techniques concernant la création de la base

La base de données a été mise au point en utilisant le langage MySQL (version 3.23.49.). Sa création a été réalisée sous un environnement PHP avec l'aide de phpMyAdmin (version 2.2.6.). Le tout étant géré par un serveur Apache.

L'organigramme présenté en annexe 4 reprend toutes les tables servant à enregistrer les données ainsi que leurs liaisons logiques (arborescence de la base de données).

2.5. Outil de saisie des informations

La gestion de la base de données environnementales, créée au format MySQL, est réalisée grâce à une série de formulaires, réalisés sous Microsoft Access, qui permettent d'enregistrer les données dans la base. Des liens vers les tables de données sont effectués sous cet environnement (lien ODBC).

2.5.1. Menu général

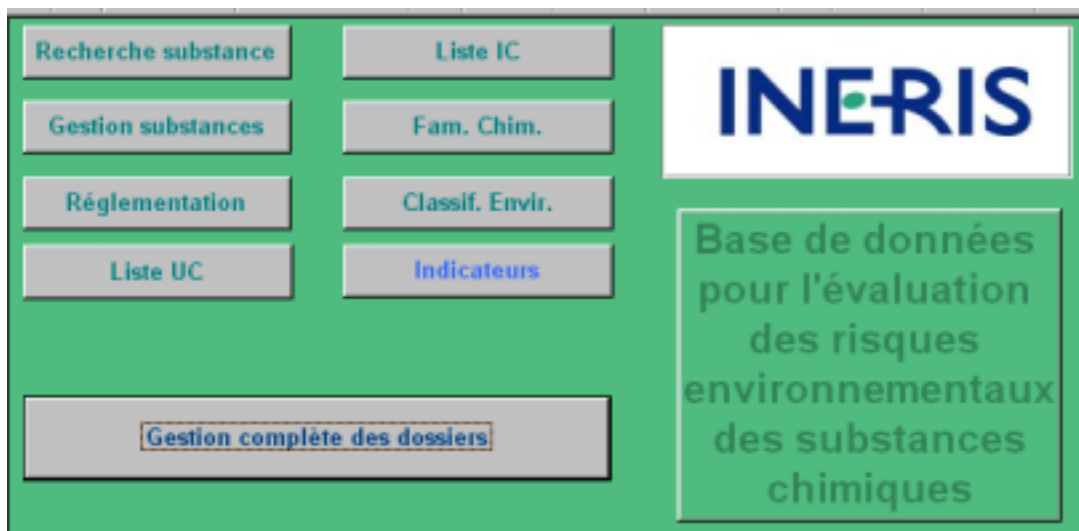


Figure 2-1 : Menu général des formulaires de saisie des données sous MS Access

2.5.2. Formulaires secondaires

◆ Gestion substances et recherche substances

Toutes les données entrées dans un dossier font référence à une substance chimique particulière. Il est nécessaire, avant d'ouvrir un dossier, de s'assurer de l'existence de la substance voulue (grâce au formulaire « Recherche substances ») ou de la créer grâce au formulaire présenté par la Figure 2-2. La recherche est possible par n° de CAS et par nom.

Figure 2-2 : Formulaire de gestion des substances de la base

- ◆ **Réglementation** : ce formulaire permet d'enregistrer, pour une substance chimique sélectionnée, les listes de substances prioritaires dans lesquelles elle apparaît.
- ◆ **Liste UC** : saisie de tous les codes et désignations des catégories d'utilisation et de fonction qui existent pour une molécule. Ces catégories¹⁵ sont celles indiquées dans EUSES (voir annexe 5).
- ◆ **Liste IC** : saisie de tous les codes et désignations des catégories industrielles qui existent pour une molécule. Ces catégories sont celles indiquées dans EUSES (annexe 5).
- ◆ **Classification environnementale** : code et phrases de risques environnementaux correspondantes.

2.5.3. Formulaire principal : gestion des dossiers

Figure 2-3 : Formulaire principal pour la gestion des données relatives à une substance

La saisie d'informations pour un dossier s'applique pour une substance précise, sélectionnée au préalable (onglet supérieur du formulaire) – l'identification de la substance en cours s'affiche sur la partie droite. Lorsque des dossiers ont déjà été enregistrés, il est possible d'y accéder en les sélectionnant dans le premier cadre supérieur du formulaire.

¹⁵ Ce champ reprend les différentes catégories définies dans le cadre de l'évaluation européenne des substances existantes. Il existe 16 catégories de domaines industriels (IC) dans lesquels les substances peuvent être utilisées. Certaines substances peuvent appartenir à plusieurs catégories. 55 catégories d'utilisations (UC) des substances sont également disponibles. Certaines d'entre-elles se divisent en sous-catégories. Il est bien entendu possible d'attribuer plusieurs fonctions ou usages à une même substance chimique. Ces catégories apparaissent souvent sous la forme de couples IC / UC (ex. : 2 / 48 : Chemical industry : basic chemicals / Solvents).

Une fois la substance et le dossier sélectionnés, les étapes de saisie d'informations peuvent commencer :

- ◆ Informations générales + assignation des catégories industrielles, d'utilisation et de la classification environnementale à l'aide des boutons de sélection en bas de formulaire.
- ◆ Indication des paramètres physico-chimiques (attention, plusieurs onglets possibles pour les paramètres pour lesquels plusieurs valeurs sont disponibles).
- ◆ Entrée des paramètres d'écotoxicologie :
 - Eau douce
 - Eau marine
 - Sol
 - Sédiments
 - Micro-organismes
- ◆ Et enfin, entrée des liens Internet vers les sites hébergeant les documents de référence.

2.6. Format des données en sortie

Le module de recherche associé à la base est accessible à partir du portail substances chimiques de l'INERIS¹⁶. Il permet l'extraction des divers éléments qui la composent sous deux formes principales :

- Des fiches synthétiques regroupant toutes les informations disponibles pour une substance donnée (classification environnementale, propriétés physico-chimiques, devenir dans l'environnement, paramètres écotoxicologiques, etc.),
- Des catalogues de données réunissant une sélection d'informations pour toutes les substances ou des groupes de substances sélectionnés.

Pour permettre d'extraire ces éléments de la base de données, un outil de recherche est mis à disposition sur le site Internet.

2.7. Outil de recherche

2.7.1. Présentation générale

Afin de répondre à l'objectif de mise à disposition de la base en libre accès sur Internet, un outil de recherche à été mis au point aux formats HTML et PHP (version 4.2.0.) - le descriptif technique de l'architecture du site dédiée à la base de données environnementale est disponible en annexe 6.

Ce dernier permettant d'établir une liaison directe entre la base de données (mySQL) et l'utilisateur sur le web.



Figure 2-4 : Liens entre les différents formats informatiques utilisés pour la création de la base de données

¹⁶ http://chimie.ineris.fr/lien/basededonnees/environnementale/recherche/search1_1.php

La base de données a été intégrée à un site Internet développé par l'INERIS dédié aux substances chimiques. Le "Portail Substances Chimiques" est accessible à l'adresse : <http://chimie.ineris.fr>

La base de données est l'objet de la partie "Environnement" du site. L'accès au moteur de recherche de la base se fait via le lien "Base de données environnementales" (voir Figure 2-5 et Figure 2-6). Une présentation de l'outil est également consultable.

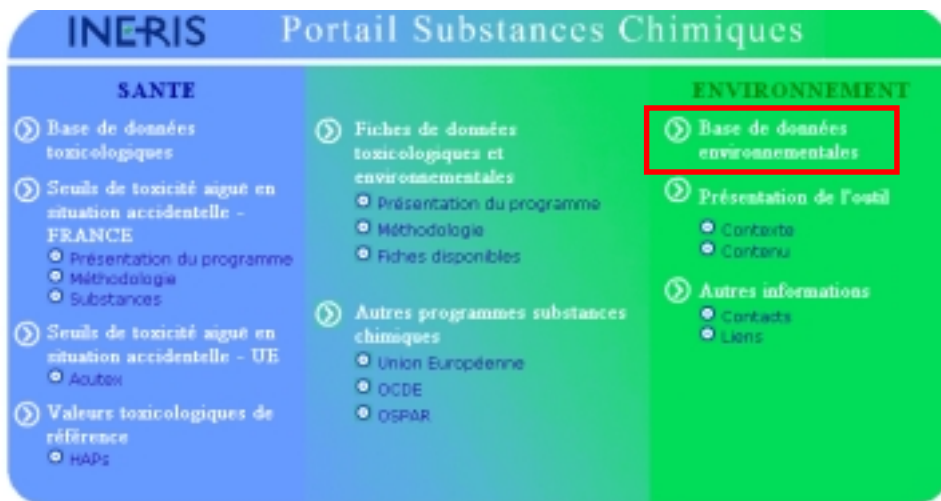


Figure 2-5 : Page d'accueil du Portail Substances Chimiques

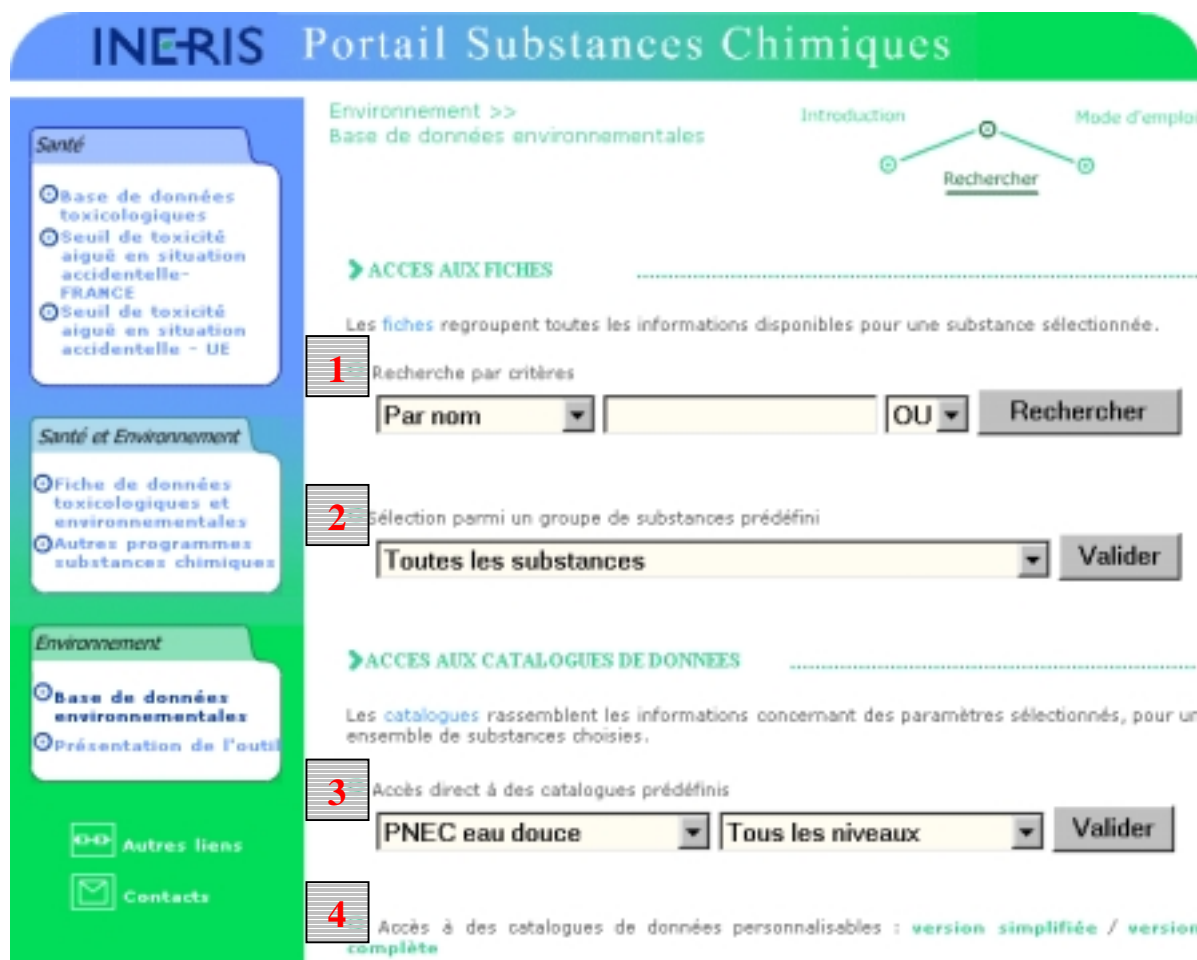
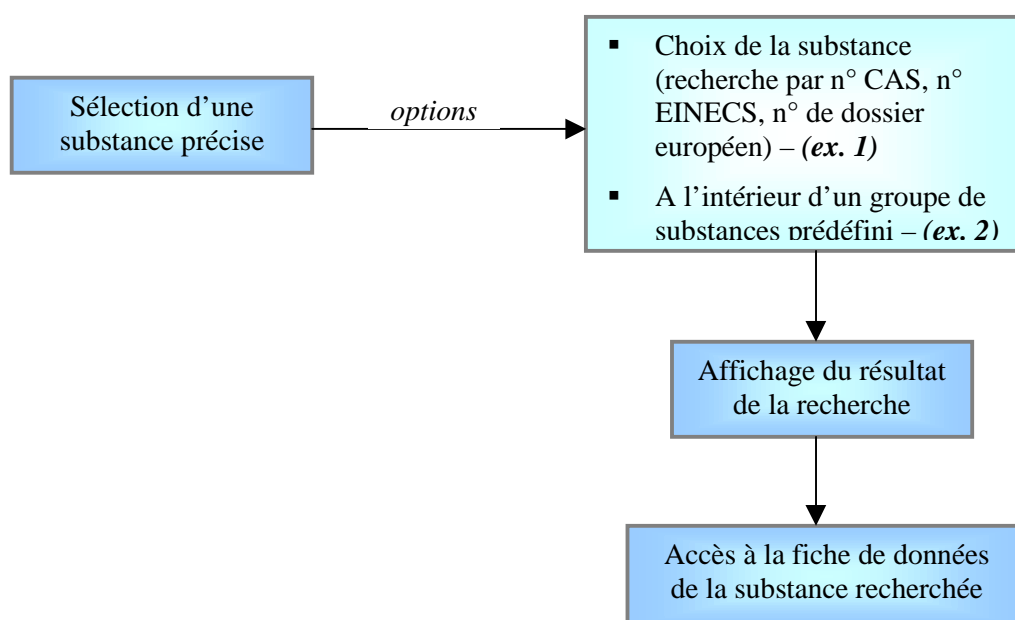


Figure 2-6 : Moteur de recherche de la base de données environnementales

2.7.2. Utilisation du moteur de recherche : recherche par substance (affichage des « fiches »)

Ce type de recherche permet d'extraire les données de la base sous la forme de fiches synthétiques récapitulant les propriétés physico-chimiques et écotoxicologiques d'une substance. Ce type de recherche peut se faire de deux manières différentes :

- recherche par mots-clés (voir exemple 1 et numéro 1 sur la Figure 2-6). La recherche est effectuée selon les critères spécifiés par l'utilisateur. L'accès à la fiche recherchée se fait ensuite après un choix dans la liste des substances correspondant au(x) critère(s) de recherche préalablement entré(s).
- recherche par groupes de substances (voir exemple 2 et numéro 2 sur la Figure 2-6). La recherche de la fiche d'une substance se fait par le choix parmi un groupe de substances prédéfini (e. g. listes réglementaires, catégories d'utilisation, etc.).



◆ Exemple 1 (ex. 1) : recherche d'une fiche de substance par mots-clés

- Recherche des substances comportant les termes « ethanol » ET « ethoxy » dans leur nom.

Moteur de recherche

1 : choix du critère de recherche

2 : saisie du ou des mot(s) clé(s)

- Après validation, on accède à l'affichage du résultat de la recherche :

11 substances trouvées

N° CAS	Nom	Fiches
112-34-5	2(2-butoxyethoxy)ethanol	✓
111-77-3	2-(2-methoxyethoxy)-ethanol	✓
110-80-5	2-Ethoxyethanol	-
109-86-4	2-Methoxyethanol	-
109-59-1	Ethanol, 2-(1-methylethoxy)-	-
112-35-6	Ethanol, 2-(2-(2-methoxyethoxy)ethoxy)-	-
124-17-4	Ethanol, 2-(2-butoxyethoxy)-, acetate	-
111-90-0	Ethanol, 2-(2-ethoxyethoxy)-	-
6881-94-3	Ethanol, 2-(2-propoxyethoxy)-	-
1704-62-7	Ethanol, 2-[2-(dimethylamino)ethoxy]-	-
112-59-4	Ethanol, 2-[2-(hexyloxy)ethoxy]-	-

- Puis on accède à la fiche souhaitée en cliquant sur la marque verte. Exemple de la fiche du 2(2-butoxyethoxy)ethanol : voir annexe 7.

◆ **Exemple 2 (ex. 2) : recherche d'une fiche de substance par groupes**

- Recherche d'une substance parmi un groupe prédéfini de substances, par exemple la première liste prioritaire pour l'évaluation des risques des substances existantes (dir. CE/793/93).

Sélection d'une substance dans des groupes prédéfinis

Choix d'un groupe

- Après validation, on accède à l'affichage du résultat de la recherche.

42 substances trouvées

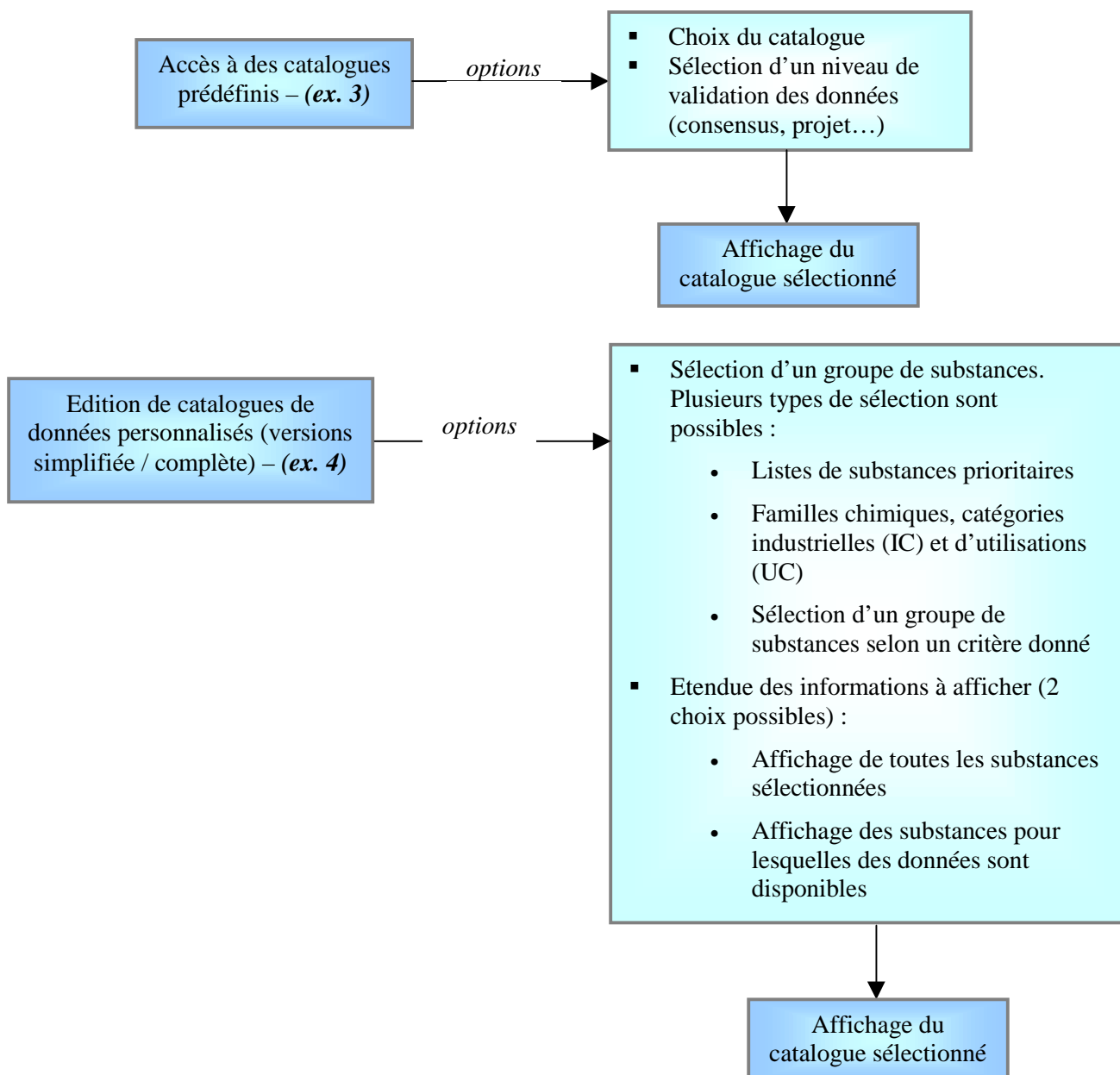
N° CAS	Nom	Fiches
106-99-0	1,3-Butadiene	✓
106-46-7	1,4-Dichlorobenzene	-
112-34-5	2(2-butoxyethoxy)ethanol	✓
95-80-7	2,4-Diaminotoluene	-
111-77-3	2-(2-methoxyethoxy)-ethanol	✓
103-11-7	2-ethylhexyl acrylate	-
110-49-6	2-methoxyethyl acetate	-
95-76-1	3,4-Dichloroaniline	-
101-77-9	4,4'-methylenedianiline	✓
1570-64-5	4-chloro-2-methylphenol (4-chloro-cresol)	-

- Puis accès à la fiche souhaitée en cliquant sur la marque verte. Exemple de la fiche du 2(2-butoxyethoxy)ethanol : voir annexe 7.

2.7.3. Utilisation du moteur de recherche : recherche par type de données (affichage des « catalogues »)

Ce type de recherche permet d'extraire les données de la base sous la forme de tableaux récapitulatifs (ou catalogues). A ce niveau, différents modes d'extraction des données sont proposés :

- Affichage d'un catalogue prédéfini (voir exemple 3 et numéro 3 sur la Figure 2-6). Ce sont les tableaux récapitulatifs des PNEC (concentrations prédites sans effet) pour différents milieux : eau douce, eau marine, sédiments, sol et stations d'épuration. Il est également possible de préciser le niveau de validation des données présentes dans ces catalogues : consensus international, projet international, proposition INERIS.
- Affichage d'un catalogue de données personnalisé (voir exemple 4 et numéro 4 sur la Figure 2-6). Ces catalogues ont la forme de tableaux reprenant, pour une liste de substances sélectionnée, les paramètres physico-chimiques ainsi que les propriétés écotoxicologiques choisis par l'utilisateur. Deux versions de ces catalogues sont disponibles : l'une qualifiée de "simplifiée" dont les propriétés écotoxicologiques sont regroupées, et l'autre, qualifiée de "complète" pour laquelle le choix des paramètres à afficher est laissé entièrement à l'utilisateur.



◆ **Exemple 3 (ex. 3) : affichage d'un catalogue prédéfini, voir numéro 3 sur la Figure 2-6.**

- Choix du catalogue prédéfini qui regroupe les PNEC pour les eaux douces faisant l'objet d'un consensus international (sélection du niveau de validation pour les données).

Accès direct à des catalogues prédéfinis

choix parmi les catalogues

option, sélection d'un niveau de validation pour les données

- Après validation, on accède à l'affichage du résultat de la recherche. Exemple du catalogue obtenu en annexe 8.

◆ **Exemple 4 (ex. 4) : affichage d'un catalogue personnalisé (version simplifiée), voir numéro 4 sur la Figure 2-6.**

- Sélection d'un groupe de substances (plusieurs critères de sélection possibles).

Edition de catalogues de données personnalisés (version simplifiée) - accès à la version complète ✓

1 : sélection d'un groupe de substances

a Réglementation

b Autres groupes (familles chimiques, IC, UC)

c / / Critère (paramètre / opérateur / valeur)

- Sélection des paramètres physico-chimiques et écotoxicologiques à afficher.

3 : sélection des paramètres à afficher dans le tableau (par défaut, le nom et le n° CAS identifie une substance)

Informations générales			
N° EINECS	<input type="checkbox"/>	Code SMILES	<input checked="" type="checkbox"/>
Formule brute	<input type="checkbox"/>	Perturbateur endocrinien	<input type="checkbox"/>
Critère PBT	<input type="checkbox"/>	Classification environnementale	<input checked="" type="checkbox"/>
Paramètres physico-chimiques			
Masse molaire	<input type="checkbox"/>	Point de fusion	<input type="checkbox"/>
Densité	<input type="checkbox"/>	log Kow	<input checked="" type="checkbox"/>
Biodégradabilité	<input type="checkbox"/>	Solubilité	<input checked="" type="checkbox"/>
Koc	<input type="checkbox"/>	Kséd-eau	<input type="checkbox"/>
Pression de vapeur	<input checked="" type="checkbox"/>	Persistance eau douce	<input type="checkbox"/>
Persistance eau marine	<input type="checkbox"/>	Photolyse	<input type="checkbox"/>
Persistance sol	<input type="checkbox"/>	Bioaccumulation (BCF)	<input type="checkbox"/>
Oxydation radicaux OH	<input type="checkbox"/>		
EAU DOUCE écotoxicologie			<input checked="" type="checkbox"/>
EAU MARINE écotoxicologie			<input type="checkbox"/>
SEDIMENTS écotoxicologie			<input type="checkbox"/>
SOL écotoxicologie			<input type="checkbox"/>
MICRO-ORGANISMES écotoxicologie			<input type="checkbox"/>

4 : valider votre sélection

- Après validation, on accède à l'affichage du résultat de la recherche. Exemple du catalogue obtenu en annexe 9.

Conclusion

La base de données regroupant des informations sur les propriétés physico-chimiques et écotoxicologiques des substances prioritaires a été mise en ligne et rendue accessible au public début décembre 2003. Elle est consultable sur le nouveau site INERIS "Portail Substances Chimiques" (<http://chimie.ineris.fr>), rubrique "base de données environnementales". Au cours de l'année 2003, 300 dossiers regroupant des informations sur les substances prioritaires ont été traités de manière à proposer, sur le site Internet, un nombre de "fiches substance" correspondant.

Cette première année, qui a vu la création et la mise en ligne de la base de données, a permis de mettre à disposition du public francophone un nombre conséquent d'informations sur les substances chimiques. Cependant, toutes les sources d'informations retenues n'ont pu être intégrées. Un travail continu d'alimentation de la base accompagné d'une activité de mise à jour des données entrées antérieurement, qui s'amplifiera avec la taille de la base, devra être mené dans les années à venir. Ceci afin de garantir, au cours du temps, pour les utilisateurs, un niveau de validité équivalent à celui proposé lors de la mise en ligne de la base.