ECHA
EUROPEAN CHEMICALS AGENCY

Preparation of an inventory of substances suspected to meet REACH Annex III criteria

Technical documentation

18 May 2016

General

ECHA prepared an inventory of substances (Annex III inventory) suspected to meet REACH Annex III criteria.

According to REACH, substances registering in the tonnage band of 1 – 10 tonnes/year meeting either one or both of Annex III criteria have to provide full Annex VII information and those not meeting either of the criteria only data on physicochemical properties.

REACH Annex III criteria are the following:

a. substances predicted (i.e. by the use of QSARs or other evidence) to likely meet criteria for CMR category 1A or 1B or Annex XIII criteria (i.e. PBT and vPvB);
b. substances with dispersive or diffuse use(s) AND predicted to likely meet criteria for any health or environmental hazard classes or differentiations under CLP Regulation.

If a substance (or any of its constituents, impurities or additives) is on this Annex III inventory, it means that there are indications that either one or both of the Annex III criteria are fulfilled. In this case full Annex VII information has to be submitted unless there are substantiated reasons to disregard the information provided in the inventory.

Substances included in the inventory

The Annex III inventory is intended to cover substances that have been pre-registered under REACH but not yet registered, and for which:

- a structure can be derived to make a prediction, or
- there is available information regarding hazard in one of the experimental databases used for the creation of the inventory.

This subset of the pre-registered substances is called the "starting pool". Substances in the starting pool for which ECHA has found any human health or environmental concerns have been flagged with an indication on the type and source of concern.

The concern for the substance in this starting pool has been derived with information collected by ECHA from:

1. Annex VI to CLP regulation (Annex VI substances)
2. QSAR models (predictions)
3. Experimental databases available in the QSAR Toolbox or on the internet (experimental data)

The appendices of this document provide the description and references to all the models and experimental sources used, allowing to check the information supporting each indication provided in the Annex III inventory.

1. Annex VI of CLP substances

Substances whose CAS or EC number are listed in Annex VI of CLP (i.e. that have a harmonised classification) for any human health or environmental concerns have been included in the inventory.

The harmonised classifications from Annex VI of CLP can be consulted in the ECHA website: http://echa.europa.eu/information-on-chemicals/annex-vi-to-clp

2. Predictions

Substances that are predicted by using in-silico techniques to be of human health or environmental concern have been included in the inventory. Where possible, predictions with low reliability have been excluded.

The list of software, models and endpoints used to compile the inventory is provided in Appendix 2, together with some considerations on the reliability of the models.

3. Experimental data

Substances that have data in external databases indicating any human health or environmental concern have been included in the inventory. The CAS number of the substances in the starting pool was queried in those databases, and those data points highlighting potential to meet the criteria for human health or environmental classification were included.

The list of databases considered is included in Appendix 3 of this document.

Format of the inventory

The Annex III inventory consists of four columns:

- **EC number**: EC number of the substance in the inventory.
- **CAS number**: CAS registry number of the substance in the list.
- **Indication**: endpoint for which there is a concern.
- **Source**: source of data indicating the concern.

A list of all the possible combination of indications and sources is given as a table in Appendix 4.

Appendices

Appendix 1: Structures used for the (Q)SAR predictions

The structures for the predictions have been derived by the Danish Technical University in the context of the Danish QSAR Database project. All models in the database require pre-processing of the molecular structures to ensure that predictions are only used for molecular structures that are structurally acceptable for QSAR processing.

In particular, the following pre-processing/filtering of molecular structures was carried out by the developers of the Danish QSAR Database:

- molecular structures that are not discrete organics are not subject to predictions;
- molecular structures that do not contain at least two carbon atoms are not subject to predictions;
- molecular structures with atoms other than H, Li, B, C, N, O, F, Na, Mg, Si, P, S, Cl, K, Ca, Br and I are not subject to predictions;
- molecular structures that correspond to mixtures or salts are not subject to predictions, unless if the salt is formed by an organic part and a small counter ion that is not expected to affect the observed toxicity, such as Na+ salts.

In the case of salts that are acceptable for processing, the small counter ion that is not expected to affect toxicity is stripped off prior to running the predictions.

Appendix 2: List of models and criteria for inclusion in the inventory

ECHA used an extensive list of QSAR software to generate predictions covering a variety of endpoints. All the models and structural alerts described below were used in the elaboration of the Annex III inventory:

Danish QSAR Database:

The following predictions included in the Danish QSAR Database were included in the Annex III inventory:

- Predictions developed by the Danish EPA using their consensus system for skin irritation.
- Predictions generated with ACD labs software for acute oral toxicity to rat.
- Predictions included in the Danish Database using the EpiSUITE models for ready biodegradability, short term toxicity to fish, daphnids and green algae, and bioconcentration factor from the Arnot-Gobas model (upper trophic level).

The Danish QSAR database, containing all the predictions used and the documentation for the models is available free of charge at: http://qsardb.food.dtu.dk/db/index.html

Predictions for acute aquatic toxicity

The predictions for acute aquatic toxicity to fish, daphnids and green algae generated with ECOSAR were included when the calculated EC50 or LC50 was equal to or below 100 mg/L, which is the threshold for aquatic toxicity classification[1].

ECOSAR produces a warning if the predicted effect level (EC50 or LC50) is higher than the water solubility of the substance, and therefore it is likely that the chemical will not be soluble enough to show acute toxicity. In addition, the Danish QSAR database warns when a substance has a log Kow higher than the maximum for a given ECSOAR model. Predictions flagged with any of these two warnings have been disregarded, since they are considered of lower reliability.

The ECOSAR software together with the documentation explaining the calculation methodology can be downloaded from the US EPA website http://www.epa.gov/tsca-screening-tools/ecological-structure-activity-relationships-ecosar-predictive-model.

Predictions generated with ACD labs software for acute oral toxicity to rat

Apart from the LD50 predictions, the Danish QSAR database includes the calculation of a reliability index (RI) for each prediction. This RI ranges between 0 and 1, so that higher values indicate more reliable predictions. The predictions for this endpoint were included in the database if the calculated LD50 value was equal to or below 2000 mg/L and the RI was over 0.75 (so that predictions with low reliability were excluded). The documentation for this model can be found in the Danish QSAR database website under the link http://qsardb.food.dtu.dk/download/qmrf/ACD.LD50_(Rat_OR).pdf.

[1] Note that other surrogate criteria also apply for environmental classification under GHS, depending on the value of the LC50/EC50.

Predictions generated with BIOWIN for ready biodegradability

The BIOWIN predictions contained in the Danish QSAR Database have been used to predict whether a chemical is potentially persistent. BIOWIN contains 7 models with different approaches. For the prediction of ready biodegradability, the approach recommended in the BIOWIN documentation is followed. The models number 3 and 5 of BIOWIN are used. If BIOWIN 3 yields a value below 2.75, or BIOWIN 5 yields a value below 0.5, the substance is considered as not readily biodegradable, and potentially persistent.

The documentation for the BIOWIN models can be found in the help files of the EPISUITE package, which can be downloaded free of charge from http://www.epa.gov/tsca-screening-tools/download-epi-suitetm-estimation-program-interface-v411

Predictions generated with BCFBAF for bioconcentration

The predictions generated with the Arnot-Gobas model (upper trophic level, including biotransformation) were used to give indication of bioaccumulation potential. If the predicted BCF with this model is >= 2000 L/kg (3.3 log L/kg), then an indication for bioaccumulation potential is given in the inventory.

The documentation for the Arnot-Gobas model can be found in the help files of the EPISUITE package, which can be downloaded free of charge from http://www.epa.gov/tsca-screening-tools/download-epi-suitetm-estimation-program-interface-v411

Predictions generated with the Danish Database battery approach for skin irritation

The model for "Severe skin irritation in rabbit, battery" developed by the Danish Technical University was used for generating an indication for skin irritation. This model uses a consensus approach from three other skin irritation models developed by the Danish Technical University using Case Ultra, SciQSAR and Leadscope platforms. If the prediction gives a positive result, and is considered within domain (which will be displayed in the database with the message "POS_IN"), an indication for skin irritation is given in the inventory.

The documentation for the three modelling approaches and the battery approach followed to generate the predictions can be found at the website of the Danish QSAR Database:

- Case Ultra model: http://qsardb.food.dtu.dk/download/qmrf/CU_ARA.pdf
- Leadscope model: http://qsardb.food.dtu.dk/download/qmrf/LS_ARA.pdf
- SciQSAR model: http://qsardb.food.dtu.dk/download/qmrf/SQ_ARA.pdf
- Battery approach (from page 14): http://qsardb.food.dtu.dk/Danish_QSAR_Database_Draft_User_manual.pdf

Endpoint	Model	Criteria for inclusion*
Acute Oral toxicity in rat	Acute toxicity (LD50) in rats, oral administration	<=2000 mg/kg and RI >=0.75
Bioaccumulation	Arnot Gobas model (upper trophic level)	>= 3.3 log (L/kg)
Daphnia acute toxicity	ECOSAR v1.11 (most toxic class)	<= 100 mg/L And the substance is not flagged as having a water solubility lower than the effect level or a log Kow higher than the maximum for the relevant ECOSAR model.

Green Algae	ECOSAR v1.11 (most toxic class)	<= 100 mg/L And the substance is not flagged as having a water solubility lower than the effect level or a log Kow higher than the maximum for the relevant ECOSAR model.
Fish acute toxicity	ECOSAR v1.11 (most toxic class)	<= 100 mg/L And the substance is not flagged as having a water solubility lower than the effect level or a log Kow higher than the maximum for the relevant ECOSAR model.
Ready biodegradability	BIOWIN 3 and BIOWIN 5	BIOWIN 3 <2.75 or BIOWIN 5 <0.5
Skin Irritation	Severe skin irritation in rabbit, Battery	POS_IN

VEGA:

VEGA models are described in details in the help file of the software freely downloadable on the website http://www.vega-qsar.eu/. Below we present a summarised description extracted from these help files.

Mutagenicity

- CAESAR: The model provides a qualitative prediction of mutagenicity on Salmonella typhimurium (Ames test). An integrated model was arranged cascading two models: Model A, a trained Support Vector Machine(SVM) classifier, and an additional Model B for false negatives (FNs) removal based on Structural Alerts (SAs) matching. Model B works with two sets of SAs: the first one is related to mutagenicity activity, if no matching is found the second set is checked, and if some matches are found, the prediction is "suspect mutagen".

- ISS: The model provides a qualitative prediction of mutagenicity on Salmonella typhimurium (Ames test). The model has been built as a set of rules, taken from the work of Benigni and Bossa (ISS) as implemented in the software ToxTree version 2.6 (http://toxtree.sourceforge.net). The model implements all the rules related to mutagenicity and does not implement the full decision tree used by ToxTree. If at least one mutagenicity rule is matching with the given compound, a "mutagen" prediction is given; otherwise, a "non-mutagen" prediction is given. The training set for the model has been extracted from ToxTree, and consists of 670 compounds.

- KNN: The model uses as prediction the value of the closest neighbours and provides a qualitative prediction of mutagenicity on Salmonella typhimurium (Ames test). The model performs the assessment on a dataset of 5770 chemicals. This dataset has been made by Istituto di Ricerche Farmacologiche Mario Negri, merging experimental data from a benchmark dataset compiled by Hansen and colleagues and from a collection of data made available by the Japan Health Ministry within their Ames (Q)SAR project.

- SarPy: The model provides a qualitative prediction of mutagenicity on Salmonella

typhimurium (Ames test). The model has been built as a set of rules, extracted with Sarpy software from the original training set from the Mutagenicity Caesar model. The original work has been extended, resulting in two sets of rules for mutagenicity (112 rules) and non-mutagenicity (93 rules). If at least one mutagenicity rule is matching with the given compound, a "mutagen" prediction is given; if only one or more non-mutagenicity rule is matching, a "non-mutagen" prediction is given; if no rules match with the given compound, a "possible non-mutagen" prediction is given.

Carcinogenicity

- CAESAR: The model has been built as a Counter Propagation Artificial Neural Network (CP ANN). The neural network output consists of two values labelled as Positive and Non-Positive, both in the range [0,1] and with sum equal to 1; they represent how much the neuron in which the predicted compound falls belongs to the class of carcinogenic or non-carcinogenic compounds. The higher between these two values leads to the prediction.

- ISS: The model provides a qualitative prediction of carcinogenic potency according to specific requirements of Chemical regulation. The model has been built as a set of rules, taken from the work of Benigni and Bossa (ISS) as implemented in the software ToxTree version 2.6. The model implement all the rules related to carcinogenicity and does not implement the full decision tree used by ToxTree. If at least one carcinogen rule is matching with the given compound, a "carcinogen" prediction is given; otherwise, a "non-carcinogen" prediction is given. The training set for the model has been extracted from ToxTree, and consists of 797 compounds.

Skin sensitisation

- CAESAR: The model provides a qualitative prediction of skin sensitisation on mouse (local lymph node assay model). The model consists in an Adaptive Fuzzy Partition (AFP) based on 8 descriptors. The AFP produces as output two values O(positive) and O(negative) that represent the belonging degree respectively to the sensitizer and non-sensitizer classes. The input compound is assigned to the class having this degree value higher than 0.5, unless the difference between the values of the two degrees is lower than the threshold 0.001; in this case, the belonging to one class or the other is not sure, thus no prediction is made.

Developmental toxicity

- CAESAR: The model provides a qualitative prediction of developmental toxicity, based on a binary classification of FDA criteria (FDA categories A and B are considered as non toxicant, categories C, D and X are considered toxicant). The model is a QSAR classification model based on a Random Forest method, implemented using WEKA open-source libraries.

- PG: The model provides a qualitative prediction of developmental and reproductive toxicity. The model implements a virtual library of toxicant compounds as described in a study from Procter & Gamble. In the work from Procter & Gamble, 25 categories of possible toxicant have been identified, and for each category an extended list of virtual compounds have been generated. The model implements these categories, and tries to find an exact match between the given compound and any of the virtual compounds in the library. If a match is found, a prediction of "Toxicant" is given, otherwise a "NONToxicant" prediction is provided.

  The original dataset used in the cited work, consisting of 641 compounds, has been implemented as the training set.

Fish acute toxicity

- Fathead EPA: The model provides a quantitative prediction for Daphnia Magna LC50 (48 hour), given in -log(mol/l) and its conversion in mg/L. The model is a linear regression made on 17 molecular descriptors. The regression coefficients have been calculated on the T.E.S.T. original dataset, that contains 337 compounds extracted from the ECOTOX aquatic toxicity database (http://cfpub.epa.gov/ecotox/), splitted in 269 compounds for the training set and 68 for the test set.

- IRFMN: The model provides a qualitative evaluation (four toxicity classes) of fish toxicity. The model has been built as three sets of rules, extracted with Sarpy software, related to different toxicity classes. Each of the three set contains a list of relevant fragment (expressed in SMARTS notation) related to the first three toxicity classes, defined on the basis of the classification for toxicity to fish provided by Directive 92/32/EEC of the EU for dangerous substances. If one or more rules are verified for the given compound, the model will assign the compound to the most toxic class available among the verified rules. If no rules apply to the given compound, the prediction will be class 4. The extraction of the rules has been performed on a training set consisting of 567 compounds.

- KNN: The model uses as prediction the value of the closest neighbours and provides a quantitative prediction of acute toxicity in fish, given in -log(mg/L). The model performs a prediction based on a dataset of 972 chemicals. This dataset has been made by Istituto di Ricerche Farmacologiche Mario Negri, merging experimental data from several reliable sources: the database compiled by the MED-Duluth group, the OECD Toolbox, the DEMETRA Project (Rainbow Trout toxicity model) and the work of Li et al.

Daphnia acute toxicity

- DEMETRA: The model provides a quantitative prediction for Daphnia Magna LC50 (48 hour), given in -log(mol/l) and its conversion in mg/L.

- EPA: The model is a QSAR model for acute toxicity for Water Flea (Daphnia Magna), as implemented in the Demetra project. The model is a hybrid model base on multiple linear regressions, using on 16 molecular descriptors.

Ready biodegradability

- IRFMN: The model is based on the OECD TG 301C - modified MITI -I test data and provides a qualitative evaluation (binary classification) of ready biodegradability properties. The model has been built as a set of rules, extracted from the training set with Sarpy software. The final set of fragments obtained come from a work that involved both a statistical part and an expert-based part. The overall model is conservative, and in case of the presence of conflicting fragments the prediction is for "not readily biodegradable". The logical scheme of the model comes directly from a chemical reasoning: a substance is always considered not biodegradable if at least one fragment related to "not biodegradable" is found, even if easily biodegradable fragments are found; this means that a part of the compound is anyway persistent.

Bioaccumulation

- CAESAR: The model provides a quantitative prediction of bioconcentration factor (BCF) in fish, given in log(L/kg). Two models, Model A and Model B, have been used to build hybrid

model, Model C. In the proposed approach, the outputs of the individual models (Model A and B) were used as inputs of the hybrid model. Model A was developed by Radial Basis Function Neural Networks (RBFNN) using a heuristic method to select the optimal descriptors; Model B was developed by RBFNN using genetic algorithm for the descriptors selection. RBFNN was used with a Matlab function for building the models. An in-house was used to combine Models A and B within the Model C.

- KNN: The model uses as prediction the value of the closest neighbours and provides a quantitative prediction of bioconcentration factor (BCF) in fish, given in log(L/kg). The model performs a read-across on a dataset of 860 chemicals. This dataset has been made by Istituto di Ricerche Farmacologiche Mario Negri, merging experimental data from several reliable sources, including the original dataset of the CAESAR BCF model (note that experimental values may differ from the ones in the CAESAR BCF dataset, as this new dataset has been built including more sources).

- Meylan: The model provides a quantitative prediction of bioconcentration factor (BCF) in fish, given in log(L/kg). The model is based on the method proposed by Meylan et al. implemented in the EPI Suite BCFBAF module (http://www.epa.gov/oppt/exposure/pubs/episuite.htm). The model provides a BCF prediction based on different regression equations or fixed values, selected on the basis of an initial classification between ionic and non-ionic compounds, and on the value of the predicted logP value.

The following models implemented in the VEGA platform have been used:

Endpoint	Model	Criteria for inclusion*
Mutagenicity	CAESAR ISS KNN SarPy	Mutagen, suspect mutagen Mutagen Mutagen Mutagen
Carcinogenicity	CAESAR ISS	Carcinogen Carcinogen
Skin sensitisation	CAESAR	Sensitizer
Developmental toxicity	CAESAR PG	Toxicant Toxicant
Fish acute toxicity	Fathead EPA IRFMN KNN	<= 100 mg/L Toxicant <= 100 mg/L
Daphnia acute toxicity	DEMETRA EPA	<= 100 mg/L <= 100 mg/L
Ready biodegradability	IRFMN	Possible NON Readily Biodegradable, NON Readily Biodegradable
Bioaccumulation	CAESAR KNN Meylan	>= 3.3 log (L/kg) >= 3.3 log (L/kg) >= 3.3 log (L/kg)

***** VEGA provides both prediction results and the so called "assessment", which reports the experimental value for the target, when known. For compiling the Annex III inventory, the "assessment" results have been used if available. VEGA also performs automatically some domain checks to highlight whether the prediction is considered reliable. These checks are summarised in an overall "applicability domain index" ranging between 0 and 1. Based on the "applicability domain index", VEGA assigns a reliability score to the prediction. For this exercise, predictions with "low reliability" were excluded.

QSAR Toolbox:

The following profilers from the QSAR Toolbox were used. Not all of the structural classes within a profiler were used; those with the lowest predictivity were excluded in order to avoid raising concerns that were likely irrelevant.

Endpoint	Model	Classes used
Mutagenicity	*In vitro* mutagenicity (Ames test) alerts by ISS	All classes except: - Alkyl (C<5) or benzyl ester of sulphonic or phosphonic acid - Alkyl carbamate and thiocarbamate - alpha,beta-unsaturated carbonyls - Aromatic mono-and dialkylamine - Coumarins and Furocoumarins - Hydrazine - Simple aldehyde
	DNA alerts for AMES, MN and CA by OASIS v.1.3	All classes except: - Acylation >> Direct acylation involving a leaving group >> Geminal Polyhaloalkanes - SN2|SN2 >> Epoxidation of Aliphatic Alkenes >> Polarized Haloalkene Derivatives - SN2|SN2 >> SN2 at sp3-carbon atom|SN2 >> SN2 at sp3-carbon atom >> Alkylphosphates, Alkylthiophosphates and Alkylphosphonates
Developmental toxicity	DART scheme v1.0	All classes except: - Inorganic chemical;Metal atoms were identified;Metals (1a);Not covered by current version of the decision tree - Inorganic chemical;Metal atoms were identified;Metals (1a);Not covered by current version of the decision tree;Not known precedent reproductive and developmental toxic potential - Inorganic chemical;Metal atoms were identified;Metals (1a);Not covered by current version of the decision tree;Organophosphorus compounds (1b) - Inorganic chemical;Metal atoms were identified;Not covered by current version of the decision tree - Inorganic chemical;Metal atoms were identified;Not covered by current version of the decision tree;Not known precedent reproductive and developmental toxic potential - Inorganic chemical;Metal atoms were identified;Not covered by current version of the decision tree;Organophosphorus compounds (1b) - Inorganic chemical;Not covered by current version of the decision tree - Inorganic chemical;Not covered by current version of the decision tree;Not known precedent reproductive and developmental toxic potential - Inorganic chemical;Not covered by current version of the decision tree;Not known precedent reproductive and developmental toxic potential;Organophosphorus compounds (1b) - Inorganic chemical;Not covered by current version of the

		decision tree;Organophosphorus compounds (1b) - Inorganic chemical|Metal atoms were identified|Metals (1a)|Not covered by current version of the decision tree - Inorganic chemical|Metal atoms were identified|Metals (1a)|Not covered by current version of the decision tree|Not known precedent reproductive and developmental toxic potential - Inorganic chemical|Metal atoms were identified|Metals (1a)|Not covered by current version of the decision tree|Organophosphorus compounds (1b) - Inorganic chemical|Metal atoms were identified|Not covered by current version of the decision tree - Inorganic chemical|Metal atoms were identified|Not covered by current version of the decision tree|Not known precedent reproductive and developmental toxic potential - Inorganic chemical|Metal atoms were identified|Not covered by current version of the decision tree|Not known precedent reproductive and developmental toxic potential|Organophosphorus compounds (1b) - Inorganic chemical|Metal atoms were identified|Not covered by current version of the decision tree|Organophosphorus compounds (1b) - Inorganic chemical|Not covered by current version of the decision tree - Inorganic chemical|Not covered by current version of the decision tree|Not known precedent reproductive and developmental toxic potential - Inorganic chemical|Not covered by current version of the decision tree|Not known precedent reproductive and developmental toxic potential|Organophosphorus compounds (1b) - Inorganic chemical|Not covered by current version of the decision tree|Organophosphorus compounds (1b) - Metal atoms were identified;Metals (1a);Not covered by current version of the decision tree - Metal atoms were identified;Not covered by current version of the decision tree - Metal atoms were identified;Not covered by current version of the decision tree;Not known precedent reproductive and developmental toxic potential - Metal atoms were identified|Metals (1a)|Not covered by current version of the decision tree - Metal atoms were identified|Metals (1a)|Not covered by current version of the decision tree|Not known precedent reproductive and developmental toxic potential - Metal atoms were identified|Not covered by current version of the decision tree - Metal atoms were identified|Not covered by current version of the decision tree|Not known precedent reproductive and developmental toxic potential - Metal atoms were identified|Not covered by current version of the decision tree|Organophosphorus compounds (1b) - Not covered by current version of the decision tree;Organophosphorus compounds (1b) - Not covered by current version of the decision tree|Not known precedent reproductive and developmental toxic potential|Organophosphorus compounds (1b) - Not covered by current version of the decision

		tree|Organophosphorus compounds (1b) - Not covered by current version of the decision tree|Organosiloxanes (1c-1) - Not known precedent reproductive and developmental toxic potential - Not known precedent reproductive and developmental toxic potential;Steroid derivatives - Not known precedent reproductive and developmental toxic potential|Steroid derivatives - Steroid derivatives
Respiratory sensitisation		All classes

Appendix 3: List of databases and criteria for inclusion in the inventory

Data was extracted from the following databases in the QSAR Toolbox, and compared with the starting pool of substances for the inventory using their CAS numbers as identifiers:

Endpoint	Database	Criteria for inclusion
Carcinogenicity	IARC monographies[2] IMAP assessment[3] ISSCAN[4] NTP RoC[5]	Group 1, 2A, 2B Recommendation for C classification
Mutagenicity	ISSSTY[6] ISSCTA[7] ISSMIC[8]	Positive or Equivocal result Positive or Positive in tumour promotion Positive or Equivocal result
Persistence	IMAP assessment Soil OASIS[9] Biodegradation NITE[10]	Assessed as P Soil biodegradation > 120 days Results from OECD TG 302C > 70% Results from OECD TG 301C-D > 60%
Respiratory sensitisation	IMAP assessment	Recommended for Resp sens classification
Skin sensitisation	IMAP assessment	Recommended for Skin sens classification
Toxicity	IMAP assessment	Assessed as T
Toxicity to reproduction	IMAP assessment DART database[11]	Recommended for R classification "Known developmental potential"
Specific target organ toxicity	IMAP assessment	Recommended for STOT RE classification

[2] http://monographs.iarc.fr/ENG/Classification/index.php
[3] http://www.nicnas.gov.au/chemical-information/imap-assessments/imap-assessments
[4] http://www.iss.it/meca/index.php?lang=2&id=199&tipo=25
[5] http://ntp.niehs.nih.gov/pubhealth/roc/roc13/index.html
[6] Same as footnote 3
[7] Same as footnote 3
[8] Same as footnote 3
[9] Extracted from the OECD QSAR Toolbox http://www.qsartoolbox.org/
[10] Same as footnote 8
[11] Same as footnote 8

Appendix 4: Possible combination of indication and sources in the inventory

Indication	Source
Harmonised classification for acute toxicity	The substance is listed in Annex VI of CLP as: Acute Tox. 1 The substance is listed in Annex VI of CLP as: Acute Tox. 2 The substance is listed in Annex VI of CLP as: Acute Tox. 3 The substance is listed in Annex VI of CLP as: Acute Tox. 4
Harmonised classification for aquatic toxicity	The substance is listed in Annex VI of CLP as: Aquatic Acute 1 The substance is listed in Annex VI of CLP as: Aquatic Chronic 1 The substance is listed in Annex VI of CLP as: Aquatic Chronic 2 The substance is listed in Annex VI of CLP as: Aquatic Chronic 3 The substance is listed in Annex VI of CLP as: Aquatic Chronic 4
Harmonised classification for aspiration toxicity	The substance is listed in Annex VI of CLP as: Asp. Tox. 1
Harmonised classification for carcinogenicity	The substance is listed in Annex VI of CLP as: Carc. 1A The substance is listed in Annex VI of CLP as: Carc. 1B The substance is listed in Annex VI of CLP as: Carc. 2
Harmonised classification for effects on or via lactation	The substance is listed in Annex VI of CLP as: Effects on or via lactation
Harmonised classification for eye damage	The substance is listed in Annex VI of CLP as: Eye Dam. 1
Harmonised classification for eye irritation	The substance is listed in Annex VI of CLP as: Eye Irrit. 2
Harmonised classification for mutagenicity	The substance is listed in Annex VI of CLP as: Muta. 1B The substance is listed in Annex VI of CLP as: Muta. 2
Harmonised classification for reprotoxicity	The substance is listed in Annex VI of CLP as: Repr. 2 The substance is listed in Annex VI of CLP as: Repr. 1B The substance is listed in Annex VI of CLP as: Repr. 1A
Harmonised classification for respiratory sensitisation	The substance is listed in Annex VI of CLP as: Resp. Sens. 1
Harmonised classification for skin corrosion	The substance is listed in Annex VI of CLP as: Skin Corr. 1A The substance is listed in Annex VI of CLP as: Skin Corr. 1B The substance is listed in Annex VI of CLP as: Skin Corr. 1C
Harmonised classification for skin sensitisation	The substance is listed in Annex VI of CLP as: Skin Sens. 1 The substance is listed in Annex VI of CLP as: Skin Sens. 1B The substance is listed in Annex VI of CLP as: Skin Sens. 1A
Harmonised classification for specific target organ toxicity	The substance is listed in Annex VI of CLP as: STOT RE 1 The substance is listed in Annex VI of CLP as: STOT RE 2 The substance is listed in Annex VI of CLP as: STOT SE 2 The substance is listed in Annex VI of CLP as: STOT SE 3 The substance is listed in Annex VI of CLP as: STOT SE 1
Suspected acutely toxic via the oral route	The Danish QSAR database contains information indicating that the substance is predicted as toxic via the oral route.
Suspected bioaccumulative	Bioaccumulation fish CEFIC LRI database in the Toolbox reports at least one value above the B threshold of 2,000 L/Kg (3.3 log units) Bioaccumulation Canada database in the Toolbox reports at least one value above the B threshold of 2,000 L/Kg (3.3 log units) CAESAR BCF model in VEGA (Q)SAR platform predicts a BCF of x

	(x≥3.3) log(L/kg) (moderate reliability) CAESAR BCF model in VEGA (Q)SAR platform predicts a BCF of (x≥3.3) log(L/kg) (good reliability) CAESAR BCF model in VEGA (Q)SAR platform predicts a BCF of (x≥3.3) log(L/kg) (EXPERIMENTAL value) KNN BCF model in VEGA (Q)SAR platform predicts a BCF of x (x≥3.3) log(L/kg) (moderate reliability) KNN BCF model in VEGA (Q)SAR platform predicts a BCF of (x≥3.3) log(L/kg) (good reliability) KNN BCF model in VEGA (Q)SAR platform predicts a BCF of (x≥3.3) log(L/kg) (EXPERIMENTAL value) Meylan BCF model in VEGA (Q)SAR platform predicts a BCF of x (x≥3.3) log(L/kg) (moderate reliability) Meylan BCF model in VEGA (Q)SAR platform predicts a BCF of x (x≥3.3) log(L/kg) (good reliability) Meylan BCF model in VEGA (Q)SAR platform predicts a BCF of x (x≥3.3) log(L/kg) (EXPERIMENTAL value) EpiSuite data included in the Toolbox contain at least one experimental log Kow value equal to or higher than 4.5
Suspected carcinogen	The Toolbox profiler 'Carcinogenicity (genotox and nongenotox) alerts by ISS' gives an alert for carcinogenicity CAESAR Carcinogenicity model in VEGA (Q)SAR platform predicts that the chemical is Carcinogen (moderate reliability) CAESAR Carcinogenicity model in VEGA (Q)SAR platform predicts that the chemical is Carcinogen (good reliability) CAESAR Carcinogenicity model in VEGA (Q)SAR platform predicts that the chemical is Carcinogen (EXPERIMENTAL value) ISS Carcinogenicity model in VEGA (Q)SAR platform predicts that the chemical is Carcinogen (moderate reliability) ISS Carcinogenicity model in VEGA (Q)SAR platform predicts that the chemical is Carcinogen (EXPERIMENTAL value) ISS Carcinogenicity model in VEGA (Q)SAR platform predicts that the chemical is Carcinogen (good reliability) IARC monographs classified the substance as carcinogenic or probably/possibly carcinogenic Recommended for C category 2 by IMAP Recommended for C category 1A or 1B by IMAP carcinogen according to ISSCAN equivocal carcinogenicity data according to ISSCAN known to be a human carcinogen according to NTP 13th RoC reasonably anticipated to be a human carcinogen according to NTP 13th RoC
Suspected hazardous to the aquatic environment	DEMETRA Daphnia Magna toxicity model in VEGA (Q)SAR platform predicts that the chemical has a 48h EC50 of x (x<100) mg/L (good reliability) DEMETRA Daphnia Magna toxicity model in VEGA (Q)SAR platform predicts that the chemical has a 48h EC50 of x (x<100) mg/L (moderate reliability) DEMETRA Daphnia Magna toxicity model in VEGA (Q)SAR platform predicts that the chemical has a 48h EC50 of x (x<100) mg/L (EXPERIMENTAL value) EPA Daphnia Magna toxicity model in VEGA (Q)SAR platform

	predicts that the chemical has a 48h EC50 of x (x<100) mg/L (good reliability) EPA Daphnia Magna toxicity model in VEGA (Q)SAR platform predicts that the chemical has a 48h EC50 of x (x<100) mg/L (moderate reliability) EPA Daphnia Magna toxicity model in VEGA (Q)SAR platform predicts that the chemical has a 48h EC50 of x (x<100) mg/L (EXPERIMENTAL value) Fathead Minnow toxicity model (EPA) in VEGA (Q)SAR platform predicts that the chemical has a 48h EC50 of x (x<100) mg/L (good reliability) Fathead Minnow toxicity model (EPA) in VEGA (Q)SAR platform predicts that the chemical has a 48h EC50 of x (x<100) mg/L (moderate reliability) Fathead Minnow toxicity model (EPA) in VEGA (Q)SAR platform predicts that the chemical has a 48h EC50 of x (x<100) mg/L (EXPERIMENTAL value) Fish Acute Toxicity model (KNN/Read-Across) in VEGA (Q)SAR platform predicts that the chemical has a 48h EC50 of x (x<100) mg/L (good reliability) Fish Acute Toxicity model (KNN/Read-Across) in VEGA (Q)SAR platform predicts that the chemical has a 48h EC50 of x (x<100) mg/L (moderate reliability) Fish Acute Toxicity model (KNN/Read-Across) in VEGA (Q)SAR platform predicts that the chemical has a 48h EC50 of x (x<100) mg/L (EXPERIMENTAL value) Fish toxicity classification (SarPy/IRFMN) model in VEGA (Q)SAR platform predicts that the chemical is 'Toxic-1 (less than 1 mg/l) (EXPERIMENTAL value)' Fish toxicity classification (SarPy/IRFMN) model in VEGA (Q)SAR platform predicts that the chemical is 'Toxic-1 (less than 1 mg/l) (good reliability)' Fish toxicity classification (SarPy/IRFMN) model in VEGA (Q)SAR platform predicts that the chemical is 'Toxic-1 (less than 1 mg/l) (moderate reliability)' Fish toxicity classification (SarPy/IRFMN) model in VEGA (Q)SAR platform predicts that the chemical is 'Toxic-2 (between 1 and 10 mg/l) (EXPERIMENTAL value)' Fish toxicity classification (SarPy/IRFMN) model in VEGA (Q)SAR platform predicts that the chemical is 'Toxic-2 (between 1 and 10 mg/l) (good reliability)' Fish toxicity classification (SarPy/IRFMN) model in VEGA (Q)SAR platform predicts that the chemical is 'Toxic-2 (between 1 and 10 mg/l) (moderate reliability)' Fish toxicity classification (SarPy/IRFMN) model in VEGA (Q)SAR platform predicts that the chemical is 'Toxic-3 (between 10 and 100 mg/l) (EXPERIMENTAL value)' Fish toxicity classification (SarPy/IRFMN) model in VEGA (Q)SAR platform predicts that the chemical is 'Toxic-3 (between 10 and 100 mg/l) (good reliability)' Fish toxicity classification (SarPy/IRFMN) model in VEGA (Q)SAR platform predicts that the chemical is 'Toxic-3 (between 10 and 100 mg/l) (moderate reliability)'

Suspected mutagen	The Toolbox profiler 'DNA alerts for AMES MN and CA by OASIS v.1.3' gives an alert for mutagenicity The Toolbox profiler 'in vitro mutagenicity (Ames test) alerts by ISS' gives an alert for mutagenicity The Toolbox profiler 'Protein binding alerts for Chromosomal aberration by OASIS v1.1' gives an alert for mutagenicity CAESAR Mutagenicity model in VEGA (Q)SAR platform predicts that the chemical is Mutagen (good reliability) CAESAR Mutagenicity model in VEGA (Q)SAR platform predicts that the chemical is Suspect Mutagen (moderate reliability) CAESAR Mutagenicity model in VEGA (Q)SAR platform predicts that the chemical is Mutagen (moderate reliability) CAESAR Mutagenicity model in VEGA (Q)SAR platform predicts that the chemical is Suspect Mutagen (good reliability) CAESAR Mutagenicity model in VEGA (Q)SAR platform predicts that the chemical is Mutagen (EXPERIMENTAL value) ISS Mutagenicity model in VEGA (Q)SAR platform predicts that the chemical is Mutagen (moderate reliability) ISS Mutagenicity model in VEGA (Q)SAR platform predicts that the chemical is Mutagen (EXPERIMENTAL value) ISS Mutagenicity model in VEGA (Q)SAR platform predicts that the chemical is Mutagen (good reliability) KNN Mutagenicity model in VEGA (Q)SAR platform predicts that the chemical is Mutagen (good reliability) KNN Mutagenicity model in VEGA (Q)SAR platform predicts that the chemical is Mutagen (EXPERIMENTAL value) KNN Mutagenicity model in VEGA (Q)SAR platform predicts that the chemical is Mutagen (moderate reliability) SARPY Mutagenicity model in VEGA (Q)SAR platform predicts that the chemical is NON-Mutagen (EXPERIMENTAL value) SARPY Mutagenicity model in VEGA (Q)SAR platform predicts that the chemical is Mutagen (good reliability) SARPY Mutagenicity model in VEGA (Q)SAR platform predicts that the chemical is Mutagen (moderate reliability) SARPY Mutagenicity model in VEGA (Q)SAR platform predicts that the chemical is Mutagen (EXPERIMENTAL value) Recommended for M category 2 by IMAP Recommended for M category 1A or 1B by IMAP The outcome in CTA assay is "positive" according to ISSCTA The outcome in CTA assay is "positive in tumor promotion" according to ISSCTA In vivo micronucleus test outcome "positive" according to ISSMIC In vivo micronucleus test outcome "equivocal" according to ISSMIC mutagen according to ISSSTY equivocal mutagenicity data according to ISSSTY
Suspected persistent in the environment	IMAP assessed the substance as P Biodegradation NITE database in the Toolbox contains at least one experimental data from a 14 days ready biodegradability test (OECD TG 302C) reporting a value lower than 70% Biodegradation NITE database in the Toolbox contains at least one experimental data from 28 days ready biodegradability test (OECD TG 301C or 301D) reporting a value lower than 60%

	Soil OASIS database in the Toolbox contains at least one experimental data from biodegradation in soil test reporting a value equal or greater than 120 days Ready biodegradability model (IRFMN) in VEGA (Q)SAR platform predicts that the chemical is Possible NON Readily Biodegradable (moderate reliability) Ready biodegradability model (IRFMN) in VEGA (Q)SAR platform predicts that the chemical is NON Readily Biodegradable (moderate reliability) Ready biodegradability model (IRFMN) in VEGA (Q)SAR platform predicts that the chemical is NON Readily Biodegradable (EXPERIMENTAL value) Ready biodegradability model (IRFMN) in VEGA (Q)SAR platform predicts that the chemical is NON Readily Biodegradable (good reliability) Ready biodegradability model (IRFMN) in VEGA (Q)SAR platform predicts that the chemical is Possible NON Readily Biodegradable (good reliability)
Suspected respiratory sensitiser	The Toolbox profiler 'Respiratory sensitisation' gives an alert for respiratory sensitisation Recommended for Resp. Sens 1 by IMAP
Suspected skin irritant	The Danish QSAR database contains information indicating that the substance is predicted as skin irritant
Suspected skin sensitiser	The Toolbox profiler 'Protein binding alerts for skin sensitization by OASIS v1.3' gives an alert for skin sensitisation CAESAR skin sensitisation model in VEGA (Q)SAR platform predicts that the chemical is Sensitizer (moderate reliability) CAESAR skin sensitisation model in VEGA (Q)SAR platform predicts that the chemical is Sensitizer (good reliability) CAESAR skin sensitisation model in VEGA (Q)SAR platform predicts that the chemical is Sensitizer (EXPERIMENTAL value) Recommended for Skin Sens 1 by IMAP
Suspected to meet STOT RE classification	Recommended for STOT RE 2 by IMAP Recommended for STOT RE 1 by IMAP
Suspected toxic	IMAP assessed the substance as Uncertain T IMAP assessed the substance as T
Suspected toxic to reproduction	The Toolbox profiler 'DART scheme v.1.0' gives an alert for toxicity to reproduction CAESAR developmental toxicity model in VEGA (Q)SAR platform predicts that the chemical is Toxicant (good reliability) CAESAR developmental toxicity model in VEGA (Q)SAR platform predicts that the chemical is Toxicant (moderate reliability) CAESAR developmental toxicity model in VEGA (Q)SAR platform predicts that the chemical is Toxicant (EXPERIMENTAL value) Developmental/Reproductive Toxicity library (PG) in VEGA (Q)SAR platform predicts that the chemical is Toxicant (good reliability) Developmental/Reproductive Toxicity library (PG) in VEGA (Q)SAR platform predicts that the chemical is Toxicant (EXPERIMENTAL value) Developmental/Reproductive Toxicity library (PG) in VEGA (Q)SAR platform predicts that the chemical is Toxicant (moderate reliability) DART database in the Toolbox reports that this substance as

	"Known developmental potential" Recommended for R category 1A or 1B by IMAP Recommended for R category 2 by IMAP
Suspected very bioaccumulative	CAESAR BCF model in VEGA (Q)SAR platform predicts a BCF of x (x≥3.7) log(L/kg) (moderate reliability) CAESAR BCF model in VEGA (Q)SAR platform predicts a BCF of (x≥3.7) log(L/kg) (good reliability) CAESAR BCF model in VEGA (Q)SAR platform predicts a BCF of (x≥3.7) log(L/kg) (EXPERIMENTAL value) KNN BCF model in VEGA (Q)SAR platform predicts a BCF of x (x≥3.7) log(L/kg) (moderate reliability) KNN BCF model in VEGA (Q)SAR platform predicts a BCF of (x≥3.7) log(L/kg) (good reliability) KNN BCF model in VEGA (Q)SAR platform predicts a BCF of (x≥3.7) log(L/kg) (EXPERIMENTAL value) Meylan BCF model in VEGA (Q)SAR platform predicts a BCF of x (x≥3.7) log(L/kg) (moderate reliability) Meylan BCF model in VEGA (Q)SAR platform predicts a BCF of x (x≥3.7) log(L/kg) (good reliability) Meylan BCF model in VEGA (Q)SAR platform predicts a BCF of x (x≥3.7) log(L/kg) (EXPERIMENTAL value)