Annexes

Annex 1: The detailed contribution of the author of present thesis

Annex 2: The final SOP used for the QC/QA of the COSMOS database chemical domain (chapter 2)

Annex 3: Use functions of cosmetics ingredients from the EC COSING inventory (chapter 3)

Annex 4: The final SOP used for skin permeability data harvesting (chapter 4)

Annex 5: Dataset used for skin permeability classification analysis (chapter 5)

Annex 6: The values of calculated descriptors (Corina Symphony, Molecular Networks GmbH, Nüremberg, Germany) and 3 Principal Component's scores (JMP, SAS Inc.) used for the skin permeability classification analysis (chapter 5)

Annex 7: Summary statistics calculated for particular properties within each category of skin permeability potential (chapter 5)

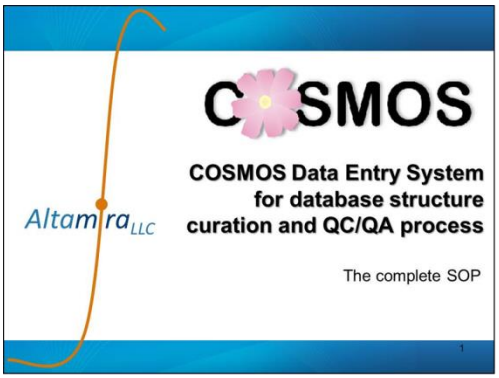
Annex 8: Abstracts of conference presentations related to the present PhD programme

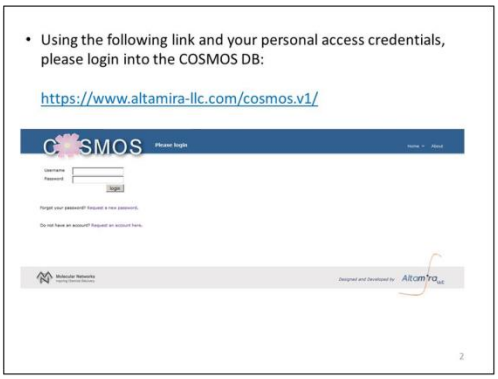
Annex 1: The detailed contribution of the author of present thesis

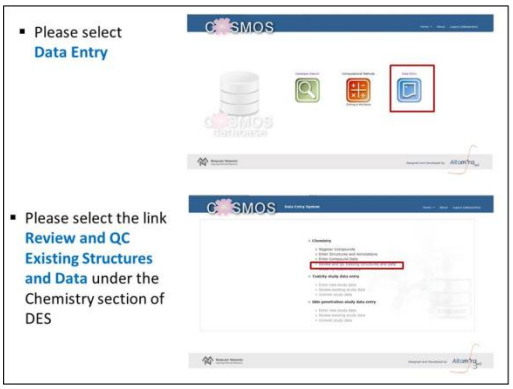
CHAPTER	CHAPTER TITLE	AUTHOR'S CONTRIBUTION
Chapter 2	**Quality Control of the COSMOS Database Chemical Domain**	○ Co-design of the sets of controlled vocabularies for chemical compounds and structures' annotations, with a specific goal to address the problematic issues related to the representation and identification of cosmetics related substances during the collation of chemical part of the COSMOS database and COSMOS Cosmetics Inventory ○ Curation of the chemical records from the U.S. EPA DSSTox inventory (approximately 12,000 records) for the purpose of populating them into the COSMOS database ○ Preparing the Standard Operating Procedure (SOP) for conducting the Quality Control/Quality Assurance (QC/QA) process of the COSMOS database chemical domain ○ Conducting a training session for the participating COSMOS partners: "COSMOS Data Entry System training for database structure curation" ○ Performing the QC/QA for 38 compounds
Chapter 3	**Chemical Space Analysis of the COSMOS Cosmetics Inventory**	○ Analysis performed by the author
Chapter 4	**The COSMOS Skin Permeability Database: Harvesting, Curating and Quality Control of the Data**	○ Curation and QC of the Kent database for the purpose of merging it with the EDETOX content ○ Preparation of the data entry tables for the new data harvesting and leading two cycles of pilot data harvesting ○ Conducting a training session for the participating COSMOS partners: "COSMOS Skin Permeability/Absorption Data Harvesting" ○ Preparing the SOP and final entry tables ("data harvesting package") for the data harvesting team ○ Harvesting 100 skin permeability/absorption studies (47 *in vitro* and 53 *in vivo*) for 25 compounds ○ Gathering the harvested data from all the harvesters, performing the format QC, integrating the results into one final file ready to be merged with EDETOX/Kent content ○ Preparing data entry tables for the COSMOS/ILSI Expert Group QC ○ Gathering QC comments from the Expert Group members and incorporating them into the database ○ Analysis of the final COSMOS Skin Permeability Database content
Chapter 5	**Classification of Skin Permeability Potential Following Dermal Exposure to Support the Prediction of Repeated Dose Toxicity of Cosmetics-Related Compounds**	○ Analysis performed by the author

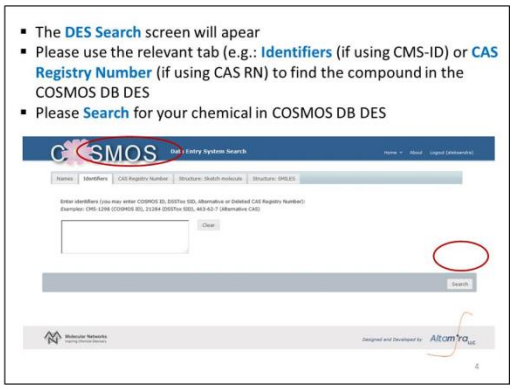
CHAPTER	CHAPTER TITLE	AUTHOR'S CONTRIBUTION
Chapter 6	**COSMOS Oral Repeated Dose Toxicity Database (oRepeatToxDB): Harvesting, Curating and Quality Control of the Data**	○ Harvesting oral repeated dose toxicity studies for 43 compounds ○ Performing the database normalisation QC/QA of 2722 records (approximately 2%) sampled from the COSMOS oRepeatToxDB ○ Analysis of the final COSMOS oRepeatToxDB content
Chapter 7	**Mechanistic, ontology-based liver toxicity data mining in the COSMOS oRepeatToxDB**	○ Participation in the validation of the liver toxicity ontology ○ Ontology-based mechanistic data mining (liver steatosis/steatohepatitis/fibrosis endpoints) and identification of 59 hepatotoxicants ○ Structural analysis (ToxPrint chemotypes) of identified hepatotoxicants and identification of potential PPAR γ agonists among them ○ Interpretation of the results of molecular modelling delivered by COSMOS partners from BAS

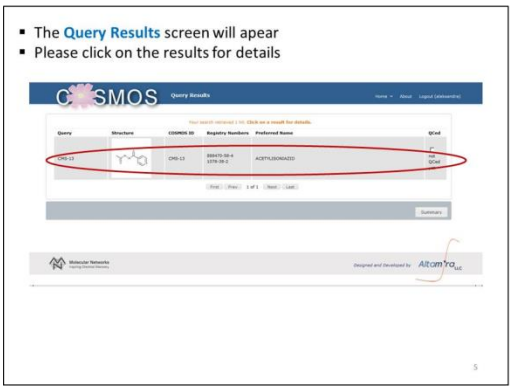
Annex 2: The final SOP used for the QC/QA of the COSMOS database chemical domain (chapter 2)

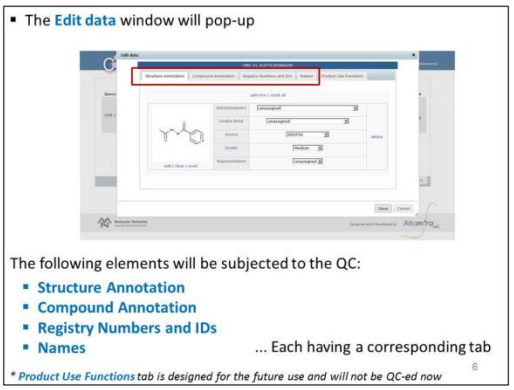


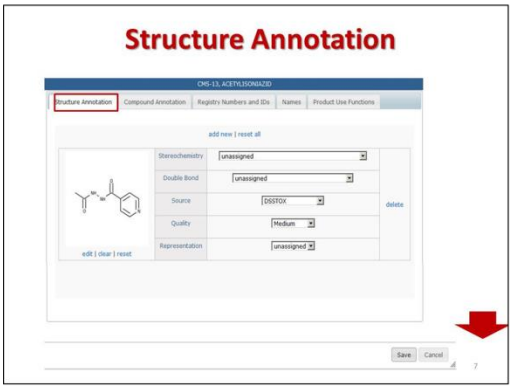


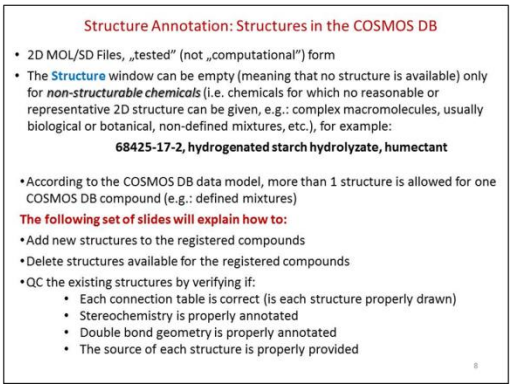












Structure Annotation: Adding and Deleting Structures

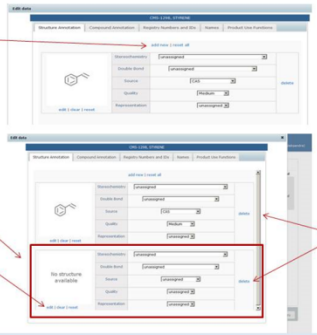
1) New structure can be added by clicking **add new**

2) The fileds for new structure will appear in separate row

3) Please click **edit** below the new structure window

This will activate embedded Molecular Editor, allowing to draw the new structure planned to add, or to upload (paste) it from the MOL/SDFile

4) Any structure can be removed by clicking **delete**

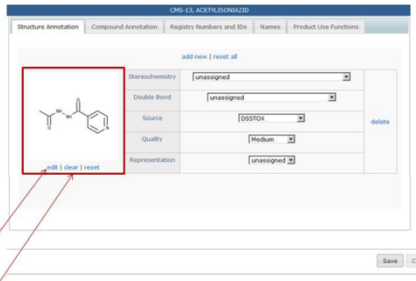


- Please add new structures only when the structure that could/should be associated with registered compound is missing („No structure available" message in the Structure window)!
- When the structure is available but it is wrong – please edit it (do not add the new, „correct" one)!

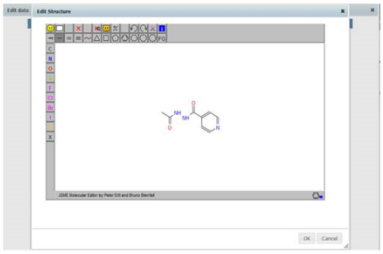
Structure Annotation: Structure QC

Please review (SciFinder) if the existing structure is properly drawn:

- If YES: Do not make any edits
- If NO: **edit** or **clear** and re-draw structure using the links below the structure image

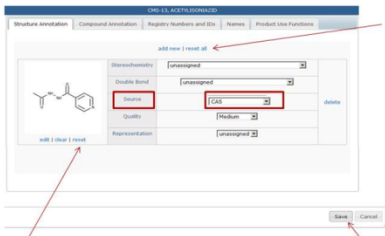


The **Edit Structure** window will pop-up with activated embedded JSME Molecular Editor (by P. Ertl and B. Bienfait, available at: http://peter-ertl.com/jsme/)



For the detailed instructions concerning the Molecular Editor applet, please refer to the authors' online help, available at: peter-ertl.com/jsme/2013_03/help.html

- Please update the structure **Source ONLY** if any changes in the structure were made. In such case, please select the source of corrected structurefrom the drop-down pick list – if the structure was checked against SciFinder – please select CAS as the source
- If no edits were made – please **DO NOT** change the structure source

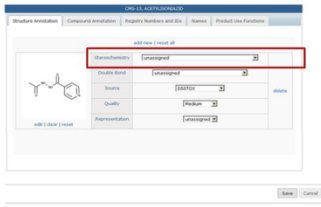


Clicking **reset all** will reset all changes made (including changes made in the connection table as well as in the annotations)

Any structure edits made can be discarded by clicking **reset** below the structure image (this will result in returning to the orignial structure)

The changes may be discarded by clicking reset/reset all only before using the Save button!

Structure Annotation: Stereochemistry



- Please review the stereochemistry information presented in the structure
- Please review/correct the annotation, selecting the relevant one from the available drop-down list

Please see the following slide for details

- Absolute stereochemistry
- Absolute stereochemistry, rotation (+)
- Absolute stereochemistry, rotation (-)
- Relative stereochemistry
- Relative stereochemistry, racemic mixture
- No stereochemistry
- Unassigned

Structure Annotation: Stereochemistry details (1)

The following Stereochemistry annotations are available:

- **Absolute stereochemistry:** The absolute configuration of the chiral center(s) is provided; A known, specific, single configuration is depicted
  - **Absolute stereochemistry, rotation (-)**
  - **Absolute stereochemistry, rotation (+)**
- **Relative stereochemistry:** The relative configuration of two or more chiral centers is provided (but not their absolute configuration); there is a known relationship to some/all of the other centers, but the absolute value in the relationship is unknown
- **Relative stereochemistry, racemic mixture:** The substance is a racemic mixture of the substance as drawn and its mirror image. The substance has an identical mirror image (that is, it is a "meso" compound)
- **No stereochemistry:** No stereochemistry is associated with the compound
- **Unassigned:** No decision on sterochemistry can be made

Structure Annotation: Stereochemistry details (2)

Update (October, 2013):

Detailed information on annotating racemic mixtures :

(Part of) structure name...

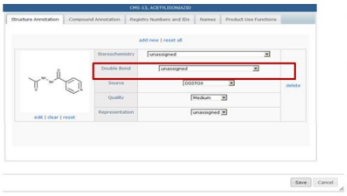
(±)
RS
DL

Following IUPAC nomenclature, this is indication of 1:1 (±) isomers ratio, **racemic mixture annotation**

(+/-)
R/S
D/L

Following IUPAC nomenclature, this is not an indication of 1:1 (+/-) isomers ratio, **not a racemic mixture**

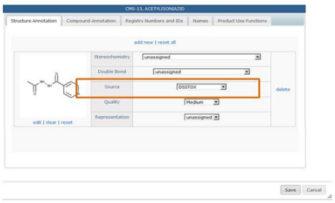
Structure Annotation: Double Bond Geometry



- Please review the double bond geometry in the structure
- Please review/correct the annotation, selecting the relevant one from the available drop-down list

- **Double bond geometry (E)**: Structure has one or more double bond(s), all with trans (E) geometry
- **Double bond geometry (Z)**: Structure has one or more double bond(s), all with cis (Z) geometry
- **Double bond geometry (E,Z)**: Structure has multiple double bond(s) with both, trans (E) and cis (Z) geometry
- **Double bond geometry**: Structure has double bonds but the geometry is not specified (can vary)
- **No double bond geometry:** Compound does not have any associated double bond geometry
- **Unassigned**: No decision on double bond geometry can be made

Structure annotation: Source

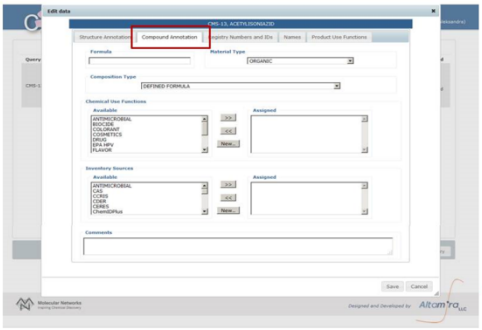


The source identifies the origin of the structure incorporated into the COSMOS DB

- The source should be changed accordingly **only** when any edits were done in the structure (please see #12)
- Please make sure that the structure source information was updated after structure edits/corrections

17

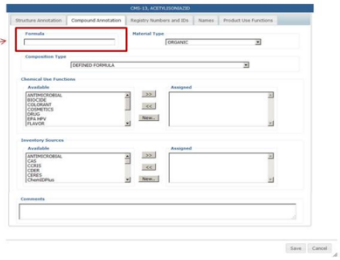
Compound Annotation



18

Compound Annotation: Formula

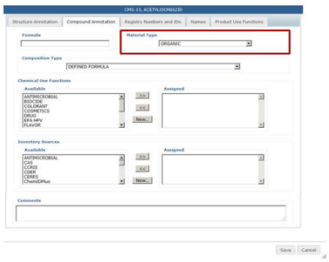
- If missing - Please fill-in the formula field using the information provided in SciFinder



- Field **Formula** will be blank if no structure is available (unstructurable/unspecified chemicals)

19

Compound Annotation: Material Type



- Entry Material Type refers to the nature of the chemical compound
- Please review if the Material Type is properly annotated
- If NO – please select the proper annotation from the available drop-down list:
  - Biologics
  - Botanicals
  - IOM – inorganic
  - IOM – metal complex
  - IOM – metalloid complex
  - IOM – organometal
  - IOM – organometalloid
  - IOM – polymer
  - Organic
  - Organic - polymer
  - Unspecified
  - Unassigned

For the definitions please see the following slide

20

Compound Annotation: Material Type definitions (1)

- **Biologics** – usually difficult to structure/ unstructurable macromolecules of biological importance (protein/ nucleic acid sequences, lipids, enzymes, etc.)
- **Botanicals** – usually complex, difficult to structure/ unstructurable mixtures of natural/plant origin; extracts; sometimes can be unspecified
- **Organics** - chemical structures containing organic carbon atom (i.e. not carbon monoxide, carbon dioxide, carbonates and cyanides), but not organometals/organometalloids
- **Polymers** – can be **IOM or Organic** – chemical structures constituted by multiple regularly or irregularly repeated units (mers)
- **Unspecified** – little or no information is available for chemical (e.g.: complex mixtures of botanical extract, exhaust gases, reaction products, etc.)
- **Unassigned** – no decision on the compound's material type can be made

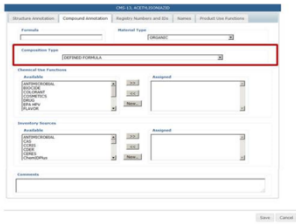
21

Compound Annotation: Material Type definitions (2)

- **IOMs** – category including:
  - **Inorganics** - chemical structures containing ***no organic carbon atom***; elements including ***metal atoms*** and ***metalloid atoms*** (boron (B), silicon (Si), germanium (Ge), arsenic (As), antimony (Sb), tellurium (Te)); ***ions*** (e.g. borate, chromate); ***minerals***
  - **Organometallics** - chemical structures containing organic carbon directly bonded to any metal atom other than alkali (I) or alkaline earth (II) metals (salts)
  - **Organometalloids** - chemical structures containing organic carbon directly bonded to any metalloid atom
  - **Metal complexes** - chemical structures containing central metallic atom covalently bonded to the total number of ligands either larger or smaller than indicated by the central metal atom's oxidation state
  - **Metalloid complexes** - chemical structures containing central metalloid atom covalently bonded to the total number of ligands either larger or smaller than indicated by the central metalloid atom's oxidation state

22

Compound Annotation: Composition Type



- Composition Type refers to the constitution of the chemical compound and elements listed in the formula field
- In majority of cases it can be inferred from the molecular formula (please see the following slide for details)
- Please review if the Composition Type is properly annotated
- If NO – please select the proper annotation from the available drop-down list:
  - defined formula
  - defined formula – varying isomers
  - formulation
  - ill-defined formula
  - ill-defined formula – polydispersed
  - ill-defined formula – polydispersed – varying isomers
  - ill-defined formula – varying composition
  - ill-defined formula – varying composition – varying isomers
  - ill-defined formula – varying isomers
  - unspecified
  - unassigned

For the definitions please see the following slide

23

Compound Annotation: Composition Type definitions (1)

Composition Type: In majority of cases can be inferred from the molecular formula solely. The information that cannot be derived from the molecular formula and must be inferred from the Name/Connection Table regards the **configuration** (stereochemistry, double bond geometry, position isomers). **Varying configuration** can be thus relevant for chemical compounds with both, ***ill-defined and defined***, molecular formulas

Ill-defined formula – refers to many cases when the chemical structure is only partially represented in the molecular formula. This can include the ***varying configuration*** but also ***varying composition*** and ***varying number of repeating units*** (polydispersion)

- **Varying composition**: Chemical structures in which one or more components, or the amount of one or more components, are unknown; This is reflected in the molecular formula as X
- **Polydispersion**: Varying number of repeating units; Relevant for polymers and usually reflected in the molecular formula as ()x or ()n, where ()x is more undefined than ()n. When ()n is used the repeating units are varying within a known range. Usually, polydispersion is also reflected in the name of the polymer (poly-)

24

Compound Annotation: Composition Type definitions (2)

Defined formula: Chemical compound with the chemical structure fully represented in the molecular formula (except configuration, please see above)

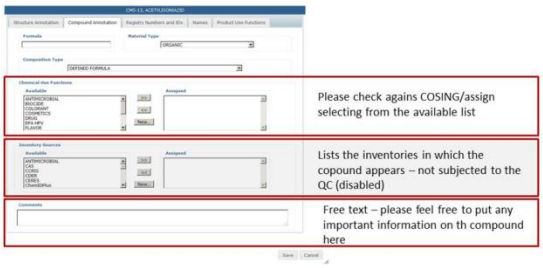
Formulation: Well-defined (usually commercial, may be proprietary) composition of two or more substances

Unspecified: Little or nothing is known about the composition of the chemical (e.g.: comlex botanical extracts, reaction products) – no molecular formula

Unassigned: No decision on compound's composition type can be made

25

Compound Annotation: Chemical Use Functions, Inventory Sources and Comments



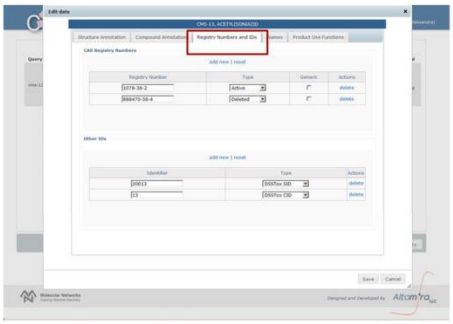
Please check agains COSING/assign selecting from the available list

Lists the inventories in which the copound appears – not subjected to the QC (disabled)

Free text – please feel free to put any important information on th compound here

26

Registry Numbers and IDs



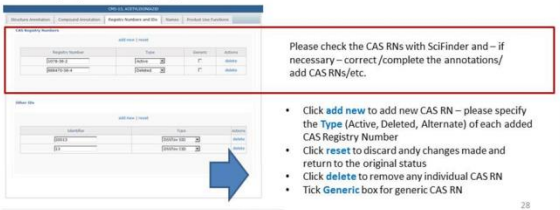
27

Registry Numbers and IDs: CAS Registry Numbers (CAS RNs)

One chemical compound can have more than one CAS Registry Number (active and several alternate or deleted CAS RNs can be also in use in different sources):

- **Active** – the current one, most recently assigned to the chemical compound
- **Alternate** – a second RN generated by CAS for a second structural representation of a substance; These records have a more preferred structure
- **Deleted** – RN assigned to a substance but later changed to another (active) RN

Generic CAS RN refers to the CAS RN that covers the whole category of chemicals, including multiple individual compounds (usually having their own, individual CAS RNs as well). Please tick the proper box, if appropriate

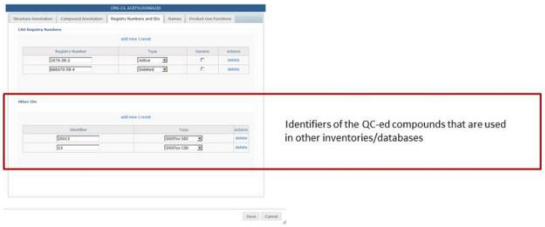


Please check the CAS RNs with SciFinder and – if necessary – correct /complete the annotations/ add CAS RNs/etc.

- Click **add new** to add new CAS RN – please specify the **Type** (Active, Deleted, Alternate) of each added CAS Registry Number
- Click **reset** to discard andy changes made and return to the original status
- Click **delete** to remove any individual CAS RN
- Tick **Generic** box for generic CAS RN

28

Registry Numbers and IDs: Other IDs

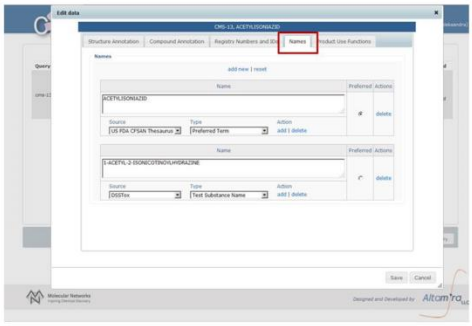


Identifiers of the QC-ed compounds that are used in other inventories/databases

Please feel free to check the IDs of the COSMOS DB compounds in other sources (inventories, databases) and add them, if necessary

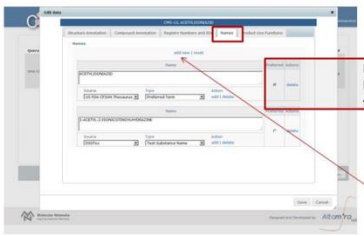
29

Names



30

Names



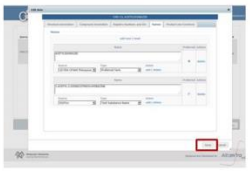
Preferred name should not be changed – disabled from the QC

- Please check other names and **delete** any (if any) outright wrong ones
- Please check the CAS REGISTRY NAME in SciFinder and correct it (if wrong) or add it (if missing). Please specify the source as CAS
- Please feel free to check for compound's names in other sources (inventories, databases) and add them (if necessary)
- Please remember about specifying the source for each added name!

31

Finalizing the QC process (1)

- After the QC please **Save** the results
- The **Save** button applies to all tabs: It can be used at any tab and all changes will be saved
- After using the **Save** button the changes made cannot be discarded/reset anymore and there is no possibility to come back to the original data status
- After using the Save button the updated DB content will be processed
- The comment confirming the successful DB update will appear
- Please click **OK** button



Finalizing the QC process (2)

- Saving the QC results (updating the DB content) will lead back to the **Query Results** window
- Now, as the QC of the chemical is finished, the relevant box at the right hand side of the screen can be ticked

... And we are ready to continue with the next compound from the list ☺

Annex 3: Use functions of cosmetics ingredients from the EC COSING inventory (chapter 3)

Name	Description
ABRASIVE	Removes materials from various body surfaces or aids mechanical tooth cleaning or improves gloss
ABSORBENT	Takes up water- and/or oil-soluble dissolved or finely dispersed substances
ANTICAKING	Allows free flow of solid particles and thus avoids agglomeration of powdered cosmetics into lumps or hard masses
ANTICORROSIVE	Prevents corrosion of the packaging
ANTIDANDRUFF	Helps control dandruff
ANTIFOAMING	Suppresses foam during manufacturing or reduces the tendency of finished products to generate foam
ANTIMICROBIAL	Helps control the growth of micro-organisms on the skin
ANTIOXIDANT	Inhibits reactions promoted by oxygen, thus avoiding oxidation and rancidity
ANTIPERSPIRANT	Reduces perspiration
ANTIPLAQUE	Helps protect against plaque
ANTISEBORRHOEIC	Helps control sebum production
ANTISTATIC	Reduces static electricity by neutralising electrical charge on a surface
ASTRINGENT	Contracts the skin
BINDING	Provides cohesion in cosmetics
BLEACHING	Lightens the shade of hair or skin
BUFFERING	Stabilises the pH of cosmetics
BULKING	Reduces bulk density of cosmetics
CHELATING	Reacts and forms complexes with metal ions which could affect the stability and/or appearance of cosmetics
CLEANSING	Helps to keep the body surface clean
COSMETIC COLORANT	Colours cosmetics and/or imparts colour to the skin and/or its appendages. All colours listed are substances on the positive list of colorants (Annex IV of the Cosmetics Directive)
DENATURANT	Renders cosmetics unpalatable. Mostly added to cosmetics containing ethyl alcohol
DEODORANT	Reduces or masks unpleasant body odours
DEPILATORY	Removes unwanted body hair
DETANGLING	Reduces or eliminates hair intertwining due to hair surface alteration or damage and, thus, helps combing
EMOLLIENT	Softens and smooths the skin
EMULSIFYING	Promotes the formation of intimate mixtures of non-miscible liquids by altering the interfacial tension
EMULSION STABILISING	Helps the process of emulsification and improves emulsion stability and shelf-life
FILM FORMING	Produces, upon application, a continuous film on skin, hair or nails
FLAVOURING	Gives flavour to the cosmetic product
FOAM BOOSTING	Improves the quality of the foam produced by a system by increasing one or more of the following properties: volume, texture and/or stability
FOAMING	Traps numerous small bubbles of air or other gas within a small volume of liquid by modifying the surface tension of the liquid
GEL FORMING	Gives the consistency of a gel (a semi-solid preparation with some elasticity) to a liquid preparation
HAIR CONDITIONING	Leaves the hair easy to comb, supple, soft and shiny and/or imparts volume, lightness, gloss, etc.

Name	Description
HAIR DYEING	Colours hair
HAIR FIXING	Permits physical control of hair style
HAIR WAVING OR STRAIGHTENING	Modifies the chemical structure of the hair, allowing it to be set in the style required
HUMECTANT	Holds and retains moisture
HYDROTROPE	Enhances the solubility of substance which is only slightly soluble in water
KERATOLYTIC	Helps eliminate the dead cells of the stratum corneum
MASKING	Reduces or inhibits the basic odour or taste of the product
MOISTURISING	Increases the water content of the skin and helps keep it soft and smooth
NAIL CONDITIONING	Improves the cosmetic characteristics of the nail
NOT REPORTED	NOT REPORTED
OPACIFYING	Reduces transparency or translucency of cosmetics
ORAL CARE	Provides cosmetic effects to the oral cavity, e.g. cleansing, deodorising, protecting
OXIDISING	Changes the chemical nature of another substance by adding oxygen or removing hydrogen
PEARLESCENT	Imparts a nacreous appearance to cosmetics
PERFUMING	Used for perfume and aromatic raw materials (Section II)
PLASTICISER	Softens and makes supple another substance that otherwise could not be easily deformed, spread or worked out
PRESERVATIVE	Inhibits primarily the development of micro-organisms in cosmetics. All preservatives listed are substances on the positive list of preservatives (Annex VI of the Cosmetics Directive)
PROPELLANT	Generates pressure in an aerosol pack, expelling contents when the valve is opened. Some liquefied propellants can act as solvents
REDUCING	Changes the chemical nature of another substance by adding hydrogen or removing oxygen
REFATTING	Replenishes the lipids of the hair or of the top layers of the skin
REFRESHING	Imparts a pleasant freshness to the skin
SKIN CONDITIONING	Maintains the skin in good condition
SKIN PROTECTING	Helps to avoid harmful effects to the skin from external factors
SMOOTHING	Seeks to achieve an even skin surface by decreasing roughness or irregularities
SOLVENT	Dissolves other substances
SOOTHING	Helps lightening discomfort of the skin or of the scalp
STABILISING	Improves ingredients or formulation stability and shelf-life
SURFACTANT	Lowers the surface tension of cosmetics as well as aids the even distribution of the product when used
TANNING	Darkens the skin with or without exposure to UV
TONIC	Produces a feeling of well-being on skin and hair
UV ABSORBER	Protects the cosmetic product from the effects of UV-light
UV FILTER	Filters certain UV rays in order to protect the skin or the hair from harmful effects of these rays. All UV filters listed are substances on the positive list of UV filters (Annex VII of the Cosmetics Directive)
VISCOSITY CONTROLLING	Increases or decreases the viscosity of cosmetics

Annex 4: The final SOP used for skin permeability data harvesting (chapter 4)

Dermal Absorption Data Harvesting Guidelines

I. Data sources:

a) ECHA resources available at:
http://echa.europa.eu/web/guest/information-on-chemicals/registered-substances

(Toxicological Information --> Toxicokinetics, metabolism and distribution --> Dermal absorption)

b) SCCS opinions available at:
http://ec.europa.eu/health/scientific_committees/consumer_safety/index_en.htm

If the original reference in ECHA or SCCS opinion is a report (unpublished) – then the information provided in ECHA or SCCS opinion should be harvested (as we usually do not have the access to original data).

If the original reference in ECHA or SCCS opinion is publication or book chapter – the original data source should be harvested. It is extremely important to harvest original papers – in ECHA many, many things are missing or not precisely reported.

II. Dermal absorption data entry tables:

Dermal absorption data harvesting will be carried out according to the 11 data entry tables, designed after pilot studies, and provided in the form of excel worksheets for each harvester.

Controlled vocabulary is provided in the form of drop-down pick lists (Table: **Lists** in excel file), which should be used during the data harvesting, and – in majority of cases – should not be modified by the harvesters without the prior consultation with ALTAMIRA (non-editable lists are marked in red). There are, however, several lists that may be extended *ad hoc* – these include, for e.g. units or strains, and are marked in green.

The detailed guidelines on how the data should be entered into particular tables are provided in the **ANNEX 1: Dermal absorption data entry tables**. Several general suggestions are listed below:

1. Please do not try to fill-in each „cell" in the worksheets – some of them can be left empty (while others will be always filled-in):
    - NULL (please refer to the ANNEX 1) attributes: can be left empty
    - NOT NULL (please refer to the ANNEX 1) attributes: cannot be left empty
2. Please do not calculate any means – just report the exact values from the investigateted studies (even if provided as an interval (range) – this is also a numerical type of data)
3. Please do not make any units conversions – add the units to the drop-down lists if necessary (lists with units are editable)

ANNEX 1: Dermal absorption data entry tables

Table 1: Chemistry (Attributes: 7) **All provided by ALTAMIRA**			
Attribute	**Data Type** **NULL/NOT NULL**	**Description**	**Examples**
Chemistry_Row#	Integer NOT NULL	Provided by ALTAMIRA	13
COSMOS ID	Formatted text NOT NULL	Provided by ALTAMIRA	CMS-5507
CAS	Number NOT NULL	Provided by ALTAMIRA	5064-31-3
INCI NAME	Text NULL	Provided by ALTAMIRA	TRISODIUM NTA
Other NAME	Text NOT NULL	Provided by ALTAMIRA	trisodium 2,2',2''-nitrilotriacetate
Data Entry Institution & Name	Formatted text NOT NULL	Provided by ALTAMIRA	CMBE-BAS, PA
Links/Comments	Free text NULL	Provided by ALTAMIRA (where relevant)	LINK TO ECHA

Table 2: StudyINFO (Attributes: 8)			
Attribute	**Data Type NULL/NOT NULL**	**Description**	**Examples**
StudyInfo_Row#	Integer NOT NULL	Please number the rows consecutively (separately for each Table)	1
StudyID	Formatted text NOT NULL	• Please assign using the format: YourInitials-Number (start with 1 and number consecutively) • There will be multiple StudyIDs if Study_Substance, Test_Animals or Test_Skin conditions vary (e.g. occluded/unoccluded; differrent species; different pre-treatment of animals; different membrane), even if the experiment is described in one original publication or retrieved in one ECHA hit • If only Test_DoseGroup conditions vary (e.g. dose level, vehicle, concentration of applied formulation, pH) and Study_Substance, Test_Animals, Test_Skin conditions are the same – one StudyID and multiple (2 or more) DoseGroupIDs should be assigned – please refer to Table 8	AMS-1
COSMOS ID	Formatted text NOT NULL	Please use COSMOS ID provided by ALTAMIRA	CMS-1371
StudyType	Formatted text NOT NULL	• Please pick from non-editable drop-down list • If the study type is different than in the list – please do not harvest	
StudyDescription	Text NULL	• Please report the description as provided in the harvested database • If not provided – please leave empty	Exp Key Dermal absorption.001
GLP_Compliance	Formatted text NULL	• Please pick from non-editable drop-down list • If not provided – please leave empty	
ECHA Klimish Score	Formatted text NULL	• Relevant only for ECHA resources • Please report as provided in ECHA using the format from example • Please leave empty for other harvested resources	2 (reliable with restrictions)
StudyInformation_Comment	Free text NULL	• Please insert any other important information • Please do not repeat information	

Table 3: StudyREF (Attributes: 13)			
Attribute	**Data Type NULL/NOT NULL**	**Description**	**Examples**
StudyReference_Row#	Integer NOT NULL	Please number the rows consecutively (separately for each Table)	1
StudyID	Formatted text NOT NULL	Please use assigned StudyID (Table 2)	AMS-1
HarvestedStudySource_Type	Formatted text NOT NULL	Please pick from non-editable drop-down list	
HarvestedStudySource_DatabaseName	Formatted text NULL	• Please pick from non-editable drop-down list • Please leave empty if not relevant (e.g. if HarvestedStudySource_Type is Primary literature publication)	
OriginalReference_Type	Formatted text NULL	• Please pick from non-editable drop-down list • Please leave empty if not relevant	
OriginalReference_ Title	Formatted text NULL	• Please use the formats from example • If not provided or not relevant : please leave empty • Please report 1st Author's Name or Organisation	The human stratum corneum layer: An effective barrier against dermal uptake of different forms of topically applied micronised titanium dioxide
OriginalReference_ JournalName			Skin Pharmacol Appl Skin Physiol
OriginalReference_ 1stAuthorName			Pflücker F WHO
OriginalReference_ Volume(Issue)	Number NULL	• Please use the formats from example • If not provided or not relevant: please leave empty	14(1)
OriginalReference_ Page-Page			92-97
OriginalReference_Year			2001
OriginalReference_OtherInformation	Text NULL	• Please insert all other details that might be important for tracking the original data source, e.g. Document number, Report date, etc.	Study Nr. 02073979/02073989 Report date: 2000-08-13
StudyReference_Comment	Free text NULL	• Please insert any other important information • Please do not repeat information	

Table 4: StudySubstance (Attributes: 8)			
Attribute	**Data Type NULL/NOT NULL**	**Description**	**Examples**
StudySubstance_Row#	Integer NOT NULL	Please number the rows consecutively (separately for each Table)	1
StudyID	Formatted text NOT NULL	Please use assigned StudyID (Table 2)	AMS-1
TestSubstance_Name	Text NOT NULL	Please repord as provided in the harvested source	Zinc Pyrithione
TestSubstance_TestedForm	Formatted text NOT NULL	• Please pick from non-editable drop-down list • Please consult ALTAMIRA if you need to add vocabulary to the list	Parent-Neutral-Radiolabelled
MetaboliteID	Formatted text NULL	• If necessary, please assign using the format: YourInitials-MB-Number • Please start with 1 and number consecutively • There might be multiple MetaboliteID(s) for one StudyID • If information on metabolites is not provided: please leave empty	AMS-MB-1
SpecificActivity_Value	Number NULL	• Relevant only when radiolabelling was used • Please leave empty if information is not provided or not relevant	19.19
SpecificActivity_Unit	Formatted text NULL	• Relevant only when radiolabelling was used • Please pick from the editable drop-down list • Please feel free to add new unit(s) to the list if necessary • Please leave empty if information is not provided or not relevant	mCi/mmol
StudySubstance_Comment	Free text NULL	• Please insert any other important information • Please do not repeat information	No radiolabelling

Table 5: Metabolites (Attributes: 14) **Please use this Table only if information on metabolites is provided. Otherwise, leave empty.**			
Attribute	**Data Type NULL/NOT NULL**	**Description**	**Example**
Metabolites_Row#	Integer NOT NULL	Please number the rows consecutively (separately for each Table)	1
StudyID	Formatted text NOT NULL	Please use assigned StudyID (Table 2)	AMS-6
MetaboliteID	Formatted text NOT NULL	Please use assigned MetaboliteID (Table 3)	AMS-MB-1
Metabolite_Name	Text NULL	Please leave empty if information is not provided	2,2'-pyridyldisulfide
Metabolite_SMILES			
Metabolic_PathwayStep_Precursor_Name			
Metabolic_PathwayStep_MetabolicPathway_Name			
Metabolic_PathwayStep_Biotransformation_Phase			
Metabolic_PathwayStep_ReactionType			
Metabolic_PathwayStep_Enzyme_Name			
MetabolicPathway_Description			
MetabolicPathway_Species			
MetabolicPathway_OrganTissue			
Metabolites_Comment	Free text NULL	• Please insert any other important information • Please do not repeat information	

Table 6: Test_Animal (Attributes: 10)			
Attribute	**Data Type NULL/NOT NULL**	**Description**	**Example**
TestAnimal_Row#	Integer NOT NULL	Please number the rows consecutively (separately for each Table)	1
StudyID	Formatted text NOT NULL	Please use assigned StudyID (Table 2)	AMS-1
Species	Formatted text NOT NULL	• Please pick from non-editable drop-down list • If the species is different than in the list – please do not harvest	
Strain	Formatted text NULL	• Please pick from the editable drop-down list • Please feel free to add new strain to the list if necessary • Please leave empty if information is not provided	
Sex	Formatted text NULL	• Please pick from non-editable drop-down list • Please leave empty if information is not provided	
Initial Age	Text NULL	• Please report using the format from examples • Please leave empty if information is not provided	12-14 weeks
Initial Weight			100-150 g
Supplier			NHS Lothian, St. John's Hospital, Livingston, UK
Number of animals			5 per group
TestAnimal_Comment	Free text NULL	• Please insert any other important information • Please do not repeat information	

Table 7: Test_Skin (Attributes: 12)			
Attribute	**Data Type NULL/NOT NULL**	**Description**	**Example**
TestSkin_Row#	Integer NOT NULL	Please number the rows consecutively (separately for each Table)	1
StudyID	Formatted text NOT NULL	Please use assigned StudyID (Table 2)	AMS-1
SkinMembrane_Type	Formatted text NULL	• Relevant for both, in vivo and in vitro studies • Please pick from non-editable drop-down list • Please consult ALTAMIRA if you need to add vocabulary to the list • Please leave empty if information is not provided	
SkinMembrane_Thickness_Value	Number NULL	• Relevant for in vitro studies (skin thickness in vivo can be deducted from the information on the site of application) • Please leave empty if information is not provided	0.2 0.2-0.4
SkinMembrane_Thickness_Unit	Formatted text NULL	• Relevant for in vitro studies • Please pick from the editable drop-down list • Please feel free to add new unit(s) to the list if necessary • Please leave empty if information is not provided	
SkinMembrane_DiskSize_Value	Number NULL	• Relevant for in vitro studies • Please leave empty if information is not provided	0.2 0.2-0.4
SkinMembrane_DiskSize_Unit	Formatted text NULL	• Relevant for in vitro studies • Please pick from the editable drop-down list • Please feel free to add new unit(s) to the list if necessary • Please leave empty if information is not provided	
SkinMembrane_Storage	Text NULL	• Relevant for in vitro studies and for excised skin from in vivo studies, if not analyzed immediately/fresh/frozen skin, etc. • Please leave empty if information is not provided	-20 degC

Table 7, continued

Table 7: Test_Skin (Attributes: 12), CONTINUED			
Attribute	**Data Type** **NULL/NOT NULL**	**Description**	**Example**
Skin_Site	Formatted text NULL	• Relevant for both, in vivo and in vitro studies • Please pick from non-editable drop-down list • Please consult ALTAMIRA if you need to add vocabulary to the list • Please leave empty if information is not provided	
Site_Area_Value	Number NULL	• Relevant for in vivo studies • Please leave empty if information is not provided	0.2 0.2-0.4
Site_Area_Unit	Formatted text NULL	• Relevant for in vivo studies • Please pick from the editable drop-down list • Please feel free to add new unit(s) to the list if necessary • Please leave empty if information is not provided	
TestSkin_Comment	Free text NULL	• Please insert any other important information • Please do not repeat information	

Table 8: Test_DoseGroup (Attributes: 19)			
Attribute	**Data Type NULL/NOT NULL**	**Description**	**Example**
TestDose_Row#	Integer NOT NULL	Please number the rows consecutively (separately for each Table)	1
StudyID	Formatted text NOT NULL	Please use assigned StudyID (Table 2)	AMS-1
DoseGroupID	Formatted text NOT NULL	• Please assign using the format: YourInitials-DG-Number • Please start with 1 and number consecutively • **Please note that for one StudyID multiple DoseGroupIDs might be assigned! (please refer to Table 2 description)**	AMS-DG-1
DoseDeliveryType	Formatted text NULL	• Please pick from non-editable drop-down list • Please leave empty if information is not provided	
SolventVehicle	Formatted text NULL	• Please pick from editable drop-down list • Please add vocabulary to the list if necessary • Please leave empty if information is not provided • **If there is no vehicle (unchanged): please use Neat from the list. Please remenmber that in this case the concentration is 100 % and the volume applied is equal to the test substance applied**	
SolubilityComments	Text NULL	Please leave empty if information is not provided	
ReceptorSolution	Text NULL	Please leave empty if information is not provided	PBS
AssayTechnique	Formatted text NULL	• Please pick from the editable drop-down list • Please feel free to add new assay technique(s) to the list if necessary • Please leave empty if information is not provided	

Table 8, continued

Table 8: Test_DoseGroup (Attributes: 19), CONTINUED			
Attribute	**Data Type NULL/NOT NULL**	**Description**	**Example**
DoseVolume	Text NULL	• Please report dose volume (**only in VOLUME UNITS**) – if provided • If applied dose is reported in other units (eg. volume/area) – please report it in the attributes DoseApplication_Solution(Formulation)_Value and DoseApplication_Solution(Formulation)_Unit • Please leave empty if information is not provided	6.4 micro-l
DoseConcentration_Value	Number NULL	• Please provide the concentration of applied formulation/solution – if provided • **For Neat (no vehicle) – the concentration is 100%** • Please leave empty if information is not provided	67
DoseConcentration_Unit	Formatted text NULL	• Please pick from the editable drop-down list • Please feel free to add new CONCENTRATION unit(s) to the list if necessary (**do not add mass or volume/area!**) • Please leave empty if information is not provided	%
DoseApplication_Solution(Formulation)_Value	Number NULL	• Please report the amount of applied **SOLUTION (FORMULATION)** reported in units other than VOLUME units • Here mass/volume per area values can be reported – if provided • Please leave empty if information is not provided	14
DoseApplication_Solution(Formulation)_Unit	Formatted text NULL	• Please pick from the editable drop-down list • Please feel free to add new unit(s) to the list if necessary (**mass or volume/area can be added**) • Please leave empty if information is not provided	mg/cm2
DoseApplication_TestSubstance_Value	Number NULL	• Please report the amount of applied **TEST SUBSTANCE** if provided • It can be reported as mass (volume)/area, mass (volume)/animal/day (for repeated dose), only as mass, etc. • Please leave empty if information is not provided	14

Table 8, continued

Table 8: Test_DoseGroup (Attributes: 19), CONTINUED			
Attribute	**Data Type NULL/NOT NULL**	**Description**	**Example**
DoseApplication_TestSubstance_Unit	Formatted text NULL	• Please pick from the editable drop-down list • Please feel free to add new unit(s) to the list if necessary (**mass or volume/area can be added**) • Please leave empty if information is not provided	mg/cm2
ScoringTechnique	Formatted text NULL	• Usually: Liquid Scintillation Counting (if radiolabelled) • Please pick from the editable drop-down list • Please feel free to add new scoring technique(s) to the list if necessary • Please leave empty if information is not provided	
Exposure_Duration	Formatted text NULL	• Report following the format from example • Please leave empty if information is not provided	2 hours 2-4 hours
Length_of_Study	Formatted text NULL	• Report following the format from example • Please leave empty if information is not provided	2 hours 2-4 hours
TestConditions_Comment	Free text NULL	• Please insert any other important information • Please do not repeat information	

Table 9: Test_Diffusion (Attributes: 15)			
Attribute	**Data Type NULL/NOT NULL**	**Description**	**Example**
TestDiff_Row#	Integer NOT NULL	Please number the rows consecutively (separately for each Table)	1
StudyID	Formatted text NOT NULL	Please use assigned StudyID (Table 2)	AMS-1
DoseGroupID	Formatted text NOT NULL	• Please use assigned DoseGroupID (Table 8)	AMS-DG-1
Diffusion_Cell_Type	Formatted text NULL	• Please pick from non-editable drop-down list • Please leave empty if information is not provided	
Diffusion_Cell_DosedArea_Value	Number NULL	• Relevant for in vitro studies (in vivo: Table: Test_Skin, Attribute: Site_Area_Unit) • Please leave empty if information is not provided	0.65
Diffusion_Cell_DosedArea_Unit	Formatted text NULL	• Relevant for in vitro studies (in vivo: Table: Test_Skin, Attribute: Site_Area_Unit) • Please pick from the editable drop-down list • Please feel free to add new unit(s) to the list if necessary • Please leave empty if information is not provided	
Diffusion_Cell_ApertureSize_Value	Number NULL	• Relevant for in vitro studies • Please leave empty if information is not provided	0.65
Diffusion_Cell_ApertureSize_Unit	Formatted text NULL	• Relevant for in vitro studies • Please pick from the editable drop-down list • Please feel free to add new unit(s) to the list if necessary • Please leave empty if information is not provided	

Table 9, continued

Table 9: Test_Diffusion (Attributes: 15), CONTINUED			
Attribute	**Data Type NULL/NOT NULL**	**Description**	**Example**
Equilibration_Bath_Temperature	Number NULL	Please leave empty if information is not provided	37.0 +/- 0.5
Equilibration_Skin_Temperature			32 +/- 1
Equilibration_TemperatureUnit	Formatted text NULL	• Please pick from the editable drop-down list • Please feel free to add new unit(s) to the list if necessary • Please leave empty if information is not provided	
Equilibration_FlowRate_Value	Number NULL	• Relevant mostly for flow-through cells (in vitro) and receptor fluid flow rate (in vivo) • Please leave empty if information is not provided	10
Equilibration_FlowRate_Unit	Formatted text NULL	• Please pick from the editable drop-down list • Please feel free to add new unit(s) to the list if necessary • Please leave empty if information is not provided	
Equilibration_Duration	Formatted text NULL	• Please report using the format from example • Please leave empty if information is not provided	2 hours
Diffusion_Comment	Free text NULL	• Please insert any other important information • Please do not repeat information	

Table 10: TestResults (Attributes: 23)			
Attribute	**Data Type NULL/NOT NULL**	**Description**	**Example**
TestRes_Row#	Integer NOT NULL	Please number the rows consecutively (separately for each Table)	1
StudyID	Formatted text NOT NULL	Please use assigned StudyID (Table 2)	AMS-1
DoseGroupID	Formatted text NOT NULL	• Please use assigned DoseGroupID (Table 8)	AMS-DG-1
LagTime_Value	Number NULL	Please leave empty if information is not provided	
LagTime_Unit	Formatted text NULL	• Please pick from the editable drop-down list • Please feel free to add new unit(s) to the list if necessary • Please leave empty if information is not provided	
Flux_Value	Number NULL	Please leave empty if information is not provided	
Flux_Unit	Formatted text NULL	• Please pick from the editable drop-down list • Please feel free to add new unit(s) to the list if necessary • Please leave empty if information is not provided	
Flux_Maximal_Value	Number NULL	Please leave empty if information is not provided	
Flux_Maximal_Unit	Formatted text NULL	• Please pick from the editable drop-down list • Please feel free to add new unit(s) to the list if necessary • Please leave empty if information is not provided	
Flux_Npoints	Number NULL	Please leave empty if information is not provided	
TotalPercentAbsorbed	Number NULL	Please leave empty if information is not provided	

Table 10, continued

Table 10: TestResults (Attributes: 23), CONTINUED			
Attribute	**Data Type NULL/NOT NULL**	**Description**	**Example**
TotalAmountAbsorbed_Value	Number NULL	Please leave empty if information is not provided	
TotalAmountAbsorbed_Unit	Formatted text NULL	• Please pick from the editable drop-down list • Please feel free to add new unit(s) to the list if necessary • Please leave empty if information is not provided	
AbsorptionPenetration_Score	Number NULL	Please leave empty if information is not provided	
AbsorbtionPenetration_Comments	Free text NULL	• Please insert any other important information • Please do not repeat information	
kp_Value	Number NULL	Please leave empty if information is not provided	
kp_Unit	Formatted text NULL	• Please pick from the editable drop-down list • Please feel free to add new unit(s) to the list if necessary • Please leave empty if information is not provided	
Histology_Status	Text NULL	• This attribute informs wheather histology studies were conducted • Please leave empty if information is not provided	
Microautoradiography_Status	Text NULL	Please leave empty if information is not provided	
TimePoints_Information	Formatted text NOT NULL	• This attribute flags the studies with timepoints data available • Please pick from non-editable drop-down list	provided
Distribution_OtherMedia_Information	Text NULL	• Please specify if the information on test substance distribution in other media or organs is provided • Please leave empty if information is not provided	Blood; Carcass; Liver

Table 10, continued

Table 10: TestResults (Attributes: 23), CONTINUED			
Attribute	**Data Type NULL/NOT NULL**	**Description**	**Example**
Elimination_Information	Text NULL	• Please specify if information on test substance elimination is provided • Please leave empty if information is not provided	Exhaled air; Urine; Faeces; Expired volatiles
Results_Comment	Free text NULL	• Please insert any other important information • Please do not repeat information	

Table 11: TestResults_Recovery (Attributes: 16)			
Attribute	**Data Type NULL/NOT NULL**	**Description**	**Example**
TestRes_Rec_Row#	Integer NOT NULL	Please number the rows consecutively (separately for each Table)	1
StudyID	Formatted text NOT NULL	Please use assigned StudyID (Table 2)	AMS-1
DoseGroup_ID	Formatted text NOT NULL	• Please use assigned DoseGroupID (Table 8)	AMS-DG-1
TotalPercentRecovery	Number NULL	Please leave empty if information is not provided	43
TotalRecovery_Value	Number NULL	Please leave empty if information is not provided	65.3
TotalRecovery_Unit	Formatted text NULL	• Please pick from the editable drop-down list • Please feel free to add new unit(s) to the list if necessary • Please leave empty if information is not provided	micro-g/cm2
Recovery_MeasureUnit	Formatted text NULL	• This unit refers to the following 8 numerical attributes (please report it here and do not repeat) • Please pick from the editable drop-down list • Please feel free to add new unit(s) to the list if necessary • Please leave empty if information is not provided	% of applied dose
Recovery_SurfaceWashUnabsorbed	Number NULL	• Please leave empty if information is not provided • Please report only value (these are attributes for numerical data) • Please report the unit in the Recovery_MeasureUnit attribute	11.05
Recovery_StratumCorneumTapeStripping			2.76
Recovery_Epidermis			
Recovery_UpperDermis			
Recovery_Epidermis&Dermis			
Recovery_TotalSkin			1.2
Recovery_ReceptorFluid			3.63
Recovery_Material-On-Cell			0.08
Recovery_Comment	Free text NULL	• Please insert any other important information • Please do not repeat information	

Annex 5: Dataset used for skin permeability classification analysis (chapter 5)

CMS ID	COSMOS DB PREFERRED NAME	# Studies with JMAX data	JMAX RANGE (micro-g/cm2/h)	MIN JMAX (micro-g/cm2/h)	MAX JMAX (micro-g/cm2/h)	MEAN JMAX (micro-g/cm2/h)	Log MEAN (micro-g/cm2/h)	SKIN PERMEABILITY POTENTIAL CATEGORY
CMS-2331	ALDOSTERONE	1	0	0.000000793	0.000000793	0.000000793	-6.10	LOW
CMS-2703	CORTISONE	1	0	0.0000035	0.0000035	0.0000035	-5.46	LOW
CMS-2367	CORTICOSTERONE	1	0	0.000036	0.000036	0.000036	-4.44	LOW
CMS-18896	CORTODOXONE	1	0	0.000259	0.000259	0.000259	-3.59	LOW
CMS-2292	PROGESTERONE	1	0	0.000943	0.000943	0.000943	-3.03	LOW
CMS-7461	PREGNENOLONE	1	0	0.001623	0.001623	0.001623	-2.79	LOW
CMS-7930	ARBUTIN	6	0.0016	0.0009	0.0025	0.0018	-2.74	LOW
CMS-8285	17ALPHA-HYDROXYPROGESTERONE	1	0	0.00198	0.00198	0.00198	-2.70	LOW
CMS-33483	CLIMBAZOLE	1	0	0.004	0.004	0.004	-2.40	LOW
CMS-3122	MORPHINE	1	0	0.006	0.006	0.006	-2.22	LOW
CMS-3386	TERBINAFINE	1	0	0.01	0.01	0.01	-2.00	LOW
CMS-691	GRISEOFULVIN	2	0.0051	0.0104	0.0155	0.01295	-1.89	LOW
CMS-143	BENZO(A)PYRENE	2	0	0.015	0.015	0.015	-1.82	LOW
CMS-1315	T-2 TOXIN	6	0.08182	0.00038	0.0822	0.017428333	-1.76	LOW
CMS-26935	2-NITRO-5-GLYCERYL METHANOLANILINE	1	0	0.0183	0.0183	0.0183	-1.74	LOW
CMS-6532	CLOTRIMAZOLE	1	0	0.02	0.02	0.02	-1.70	LOW
CMS-874	PARATHION-METHYL	2	0.0316	0.0105	0.0421	0.0263	-1.58	LOW
CMS-2940	HYDROMORPHONE	1	0	0.032	0.032	0.032	-1.49	LOW
CMS-4402	LINOLEIC ACID	1	0	0.036	0.036	0.036	-1.44	LOW
CMS-72028	TRANS-RETINYL ASCORBATE	2	0.0855	0.0125	0.098	0.05525	-1.26	LOW
CMS-3741	PROCHLORAZ	2	0.038563	0.041437	0.08	0.0607185	-1.22	LOW
CMS-345	CODEINE	1	0	0.09	0.09	0.09	-1.05	LOW
CMS-8923	ASCORBYL PALMITATE	2	0.0681	0.071	0.1391	0.10505	-0.98	LOW
CMS-729	HYDROCORTISONE	8	0.189998	0.000002	0.19	0.11287525	-0.95	LOW
CMS-7046	D&C BLUE NO. 4	6	0.124	0.068	0.192	0.142166667	-0.85	LOW
CMS-7741	AMMONIUM PERFLUOROOCTANOATE	1	0	0.19	0.19	0.19	-0.72	LOW
CMS-2865	FENTANYL	1	0	0.26	0.26	0.26	-0.59	LOW
CMS-1777	DIBUTYL PHTHALATE	2	0.52	0.07	0.59	0.33	-0.48	LOW
CMS-72019	3-O-ISOVALERYL NALTREXONE	1	0	0.36125	0.36125	0.36125	-0.44	LOW
CMS-1256	VITAMIN A PALMITATE	2	0.238	0.262	0.5	0.381	-0.42	LOW
CMS-3355	SUFENTANIL	1	0	0.4	0.4	0.4	-0.40	LOW
CMS-72003	DIBUTYL SQUARATE	1	0	0.4	0.4	0.4	-0.40	LOW
CMS-72025	3-O-ISOPROPYLOXYCARBONYL NALTREXONE	1	0	0.5124	0.5124	0.5124	-0.29	LOW
CMS-72026	3-O-PIVALYL NALTREXONE	1	0	0.544	0.544	0.544	-0.26	LOW
CMS-3049	MEPERIDINE	1	0	0.6	0.6	0.6	-0.22	LOW
CMS-72031	N,N-DIISOPROPYL NALTREXONE-3-O-CARBAMATE	1	0	0.70668	0.70668	0.70668	-0.15	LOW

CMS ID	COSMOS DB PREFERRED NAME	# Studies with JMAX data	JMAX RANGE (micro-g/cm2/h)	MIN JMAX (micro-g/cm2/h)	MAX JMAX (micro-g/cm2/h)	MEAN JMAX (micro-g/cm2/h)	Log MEAN (micro-g/cm2/h)	SKIN PERMEABILITY POTENTIAL CATEGORY
CMS-9533	OCTYL SALICYLATE	2	0.100132	0.700924	0.801056	0.75099	-0.12	MED
CMS-72020	3-O-(2-ETHYLBUTYRYL) NALTREXONE	1	0	0.86483	0.86483	0.86483	-0.06	MED
CMS-72021	3-O-ISOBUTYRYL NALTREXONE	1	0	0.92064	0.92064	0.92064	-0.04	MED
CMS-72752	N,N-DIETHYL NALTREXONE-3-O-CARBAMATE	1	0	0.9812	0.9812	0.9812	-0.01	MED
CMS-620	DI(2-ETHYLHEXYL) PHTHALATE	1	0	1.06	1.06	1.06	0.03	MED
CMS-922	MONOCHLOROACETIC ACID--PROHIBITED	1	0	1.1	1.1	1.1	0.04	MED
CMS-17099	NALTREXONE	2	0.21142	1.04346	1.25488	1.14917	0.06	MED
CMS-72032	N,N-DIMETHYL NALTREXONE-3-O-CARBAMATE	1	0	1.21128	1.21128	1.21128	0.08	MED
CMS-1776	DIETHYL PHTHALATE	1	0	1.27	1.27	1.27	0.10	MED
CMS-1517	BROMOACETIC ACID	1	0	1.4	1.4	1.4	0.15	MED
CMS-311	CHLOROFORM	1	0	1.6	1.6	1.6	0.20	MED
CMS-3937	BROMOCHLOROACETIC ACID	1	0	1.6	1.6	1.6	0.20	MED
CMS-72027	3-O-TERTIARYBUTYLOXYCARBONYL NALTREXONE	1	0	1.64493	1.64493	1.64493	0.22	MED
CMS-72018	PROPRANOLOL BENZOATE	2	0.1	1.6	1.7	1.65	0.22	MED
CMS-466	CHLORPYRIFOS	1	0	1.7	1.7	1.7	0.23	MED
CMS-263	CATECHOL	2	0.17	1.71	1.88	1.795	0.25	MED
CMS-205	BROMODICHLOROMETHANE	1	0	1.8	1.8	1.8	0.26	MED
CMS-72005	DIETHYL SQUARATE	1	0	1.8	1.8	1.8	0.26	MED
CMS-431	DICHLOROACETIC ACID	1	0	1.9	1.9	1.9	0.28	MED
CMS-1398	TCA	1	0	1.9	1.9	1.9	0.28	MED
CMS-304	DIBROMOCHLOROMETHANE	1	0	2	2	2	0.30	MED
CMS-1394	BROMOFORM	1	0	2.1	2.1	2.1	0.32	MED
CMS-72033	PENTYL NALTREXONE-3-O-CARBAMATE	1	0	2.33748	2.33748	2.33748	0.37	MED
CMS-3511	DIBROMOACETIC ACID	1	0	2.6	2.6	2.6	0.41	MED
CMS-72029	BUTYL NALTREXONE-3-O-CARBAMATE	1	0	3.5156	3.5156	3.5156	0.55	MED
CMS-2352	DIMETHYL PHTHALATE	1	0	3.95	3.95	3.95	0.60	MED
CMS-1046	N-NITROSODIETHANOLAMINE	1	0	4.1	4.1	4.1	0.61	MED
CMS-1091	BENZOYL PEROXIDE	1	0	5.1	5.1	5.1	0.71	MED
CMS-3147	NICARDIPINE	9	11.56	0.74	12.3	6.086666667	0.78	MED
CMS-72030	ETHYL NALTREXONE-3-O-CARBAMATE	1	0	6.39836	6.39836	6.39836	0.81	MED
CMS-777	ISOPHORONE	2	11.29	0.91	12.2	6.555	0.82	MED
CMS-1163	O-PHENYLPHENOL	10	11.69	1.11	12.8	6.955	0.84	MED
CMS-72017	PROPRANOLOL OLEATE	2	0.3	7	7.3	7.15	0.85	MED
CMS-72034	PROPYL NALTREXONE-3-O-CARBAMATE	1	0	7.412	7.412	7.412	0.87	MED
CMS-49099	HEXYL NICOTINATE	2	8.62	3.58	12.2	7.89	0.90	MED
CMS-4058	2,4-D, DIMETHYLAMINE SALT	2	1.3	7.9	9.2	8.55	0.93	MED
CMS-1706	1,1,1-TRICHLOROACETONE	1	0	9.6	9.6	9.6	0.98	MED
CMS-31782	DDT	1	0	10	10	10	1.00	MED

CMS ID	COSMOS DB PREFERRED NAME	# Studies with JMAX data	JMAX RANGE (micro-g/cm2/h)	MIN JMAX (micro-g/cm2/h)	MAX JMAX (micro-g/cm2/h)	MEAN JMAX (micro-g/cm2/h)	Log MEAN (micro-g/cm2/h)	SKIN PERMEABILITY POTENTIAL CATEGORY
CMS-3963	4-TERT-BUTYLCATECHOL	2	4.28	8.52	12.8	10.66	1.03	HIGH
CMS-61741	BENZYL NICOTINATE	2	6.5	13.9	20.4	17.15	1.23	HIGH
CMS-1597	1,1-DICHLOROPROPANONE	1	0	17.2	17.2	17.2	1.24	HIGH
CMS-5235	N,N-DIMETHYLETHYLAMINE	3	15	11	26	17.66666667	1.25	HIGH
CMS-435	P-DICHLOROBENZENE	2	21.6	15.3	36.9	26.1	1.42	HIGH
CMS-1541	DIETHYLENE GLYCOL MONOBUTYL ETHER	1	0	35	35	35	1.54	HIGH
CMS-4068	DIACETONE ALCOHOL	2	19.3	37.3	56.6	46.95	1.67	HIGH
CMS-3660	HEPTANE	3	91.2	22.1	113.3	66.2	1.82	HIGH
CMS-4603	PENTANE	3	155.5	13.5	169	69.03333333	1.84	HIGH
CMS-58536	BUTYL NICOTINATE	2	16.9	62.1	79	70.55	1.85	HIGH
CMS-934	NAPHTHALENE	2	117.6	25	142.6	83.8	1.92	HIGH
CMS-5106	DIETHYLENE GLYCOL MONOBUTYL ETHER ACETATE	2	103	59	162	110.5	2.04	HIGH
CMS-1908	DIETHYLENE GLYCOL MONOETHYL ETHER	1	0	125	125	125	2.10	HIGH
CMS-1000	2-NITROPROPANE	3	219.1	66.8	285.9	157.2333333	2.20	HIGH
CMS-167	BIPHENYL	2	199.2	59.1	258.3	158.7	2.20	HIGH
CMS-9	ACETONITRILE	3	309.6	66	375.6	194.8666667	2.29	HIGH
CMS-950	NICOTINE	1	0	206	206	206	2.31	HIGH
CMS-4141	DIETHYLENE GLYCOL MONOMETHYL ETHER	1	0	206	206	206	2.31	HIGH
CMS-10507	METHYL NICOTINATE	2	139	170	309	239.5	2.38	HIGH
CMS-11753	ETHYLENE GLYCOL ISOPROPYL ETHER	2	6	240	246	243	2.39	HIGH
CMS-845	4-METHOXYPHENOL	2	60	223	283	253	2.40	HIGH
CMS-158	BENZYL CHLORIDE	2	290.7	156.8	447.5	302.15	2.48	HIGH
CMS-2413	METHYL P-HYDROXYBENZOATE	16	950.93	76.51	1027.44	319.944375	2.51	HIGH
CMS-5434	ETHYLENE GLYCOL MONOPROPYL ETHER	2	171	394	565	479.5	2.68	HIGH
CMS-999	1-NITROPROPANE	3	1040.1	178.9	1219	525.9666667	2.72	HIGH
CMS-1765	METHYL ACETATE	3	977	250	1227	577.6	2.76	HIGH
CMS-455	1,2-DICHLOROPROPANE	3	177	501	678	614.9333333	2.79	HIGH
CMS-1847	1,4-XYLENE	2	1016.6	192.4	1209	700.7	2.85	HIGH
CMS-1895	ETHYLENE GLYCOL MONOETHYL ETHER ACETATE	1	0	800	800	800	2.90	HIGH
CMS-302	CHLOROBENZENE	3	387.4	614.6	1002	824.4666667	2.92	HIGH
CMS-4428	ETHYLENE GLYCOL MONOMETHYL ETHER ACETATE	2	71	831	902	866.5	2.94	HIGH
CMS-595	ETHYL ALCOHOL	4	1269	584	1853	1037.25	3.02	HIGH
CMS-1459	VINYLIDENE CHLORIDE	3	2028.3	144.7	2173	1039.466667	3.02	HIGH
CMS-443	ETHYLENE DICHLORIDE	3	1325.4	329.6	1655	1060.533333	3.03	HIGH
CMS-509	N,N-DIMETHYLACETAMIDE	3	1505	1069	2574	1914.333333	3.28	HIGH
CMS-3690	ETHYLENE GLYCOL MONOMETHYL ETHER	1	0	2820	2820	2820	3.45	HIGH
CMS-527	N,N-DIMETHYLFORMAMIDE	1	0	8400	8400	8400	3.92	HIGH
CMS-1346	TETRAHYDROFURAN	2	13900	6100	20000	13050	4.12	HIGH

Annex 6: The values of calculated descriptors (Corina Symphony, Molecular Networks GmbH, Nüremberg, Germany) and 3 Principal Component's scores (JMP, SAS Inc.) used for the skin permeability classification analysis (chapter 5)

CMS ID	NAME	Skin pereability category	PC 1	PC 2	PC 3	BondsRot	HAcc	HDon	Stereo	Weight	Complex	Complex Ring	McGowan	TPSA	Polariz	LogS	XlogP	Diameter	Rgyr	Span
CMS-2331	ALDOSTERONE	LOW	2.33	3.06	-0.69	3	5	2	7	360.44	681.89	1.35	275.46	91.67	37.55	-1.74	0.00	12.26	3.58	6.55
CMS-2703	CORTISONE	LOW	2.42	2.90	-0.53	2	5	2	6	360.44	723.91	1.35	275.46	91.67	37.55	-1.68	-0.09	13.69	3.82	7.37
CMS-2367	CORTICOSTERONE	LOW	2.16	2.40	-1.13	2	4	2	7	346.46	637.81	1.35	273.89	74.60	37.47	-2.47	1.21	12.32	3.61	6.60
CMS-18896	CORTODOXONE	LOW	2.27	1.98	-1.04	2	4	2	6	346.46	652.28	1.35	273.89	74.60	37.47	-2.96	1.93	12.45	3.72	6.86
CMS-2292	PROGESTERONE	LOW	1.36	-0.04	-2.54	1	2	0	6	314.46	588.65	1.35	262.15	34.14	36.19	-4.37	3.89	11.36	3.52	5.99
CMS-7461	PREGNENOLONE	LOW	1.63	0.55	-2.30	1	2	1	7	316.48	550.05	1.35	266.45	37.30	36.74	-4.11	3.93	11.93	3.51	6.27
CMS-7930	ARBUTIN	LOW	1.42	4.55	1.67	3	7	5	5	272.25	279.25	1.00	186.41	119.61	25.13	-0.89	-0.48	10.07	3.15	5.71
CMS-8285	17ALPHA-HYDROXYPROGESTERONE	LOW	1.71	1.15	-1.80	1	3	1	6	330.46	635.00	1.35	268.02	54.37	36.83	-3.30	2.54	12.13	3.57	6.22
CMS-33483	CLIMBAZOLE	LOW	0.70	-0.75	-0.48	5	4	0	1	292.76	335.26	1.00	218.63	44.12	31.05	-3.96	3.31	11.07	3.56	5.80
CMS-3122	MORPHINE	LOW	0.66	2.34	-1.35	0	4	2	5	285.34	494.43	1.61	206.48	52.93	30.59	-1.99	0.78	9.02	2.73	5.00
CMS-3386	TERBINAFINE	LOW	1.39	-2.71	-1.13	6	1	0	0	291.43	427.80	1.20	260.61	3.24	37.94	-5.40	5.70	13.95	4.48	7.91
CMS-691	GRISEOFULVIN	LOW	1.55	0.57	-0.65	3	6	0	2	352.77	575.43	1.21	239.47	71.06	33.53	-3.40	2.01	11.95	3.51	6.52
CMS-143	BENZO(A)PYRENE	LOW	0.29	-2.57	-2.51	0	0	0	0	252.31	372.24	1.50	195.36	0.00	36.04	-6.74	6.41	11.32	3.16	5.80
CMS-1315	T-2 TOXIN	LOW	4.47	2.84	-0.42	9	9	1	8	466.52	881.16	1.43	341.21	120.89	45.61	-2.74	1.51	13.54	3.88	7.90
CMS-26935	2-NITRO-5-GLYCERYL METHANOLANILINE	LOW	1.20	2.25	2.03	6	7	3	1	242.23	245.34	1.00	173.01	107.54	23.14	-1.63	0.54	11.73	3.72	7.10
CMS-6532	CLOTRIMAZOLE	LOW	0.99	-2.33	-1.67	4	2	0	0	344.84	396.25	1.00	262.30	17.82	40.47	-6.98	6.20	9.30	3.08	5.24
CMS-874	PARATHION-METHYL	LOW	0.44	-0.21	0.25	5	6	0	0	263.21	278.64	1.00	171.66	73.51	22.92	-3.53	3.06	10.66	3.20	5.63
CMS-2940	HYDROMORPHONE	LOW	0.42	1.70	-1.52	0	4	1	4	285.34	494.43	1.61	206.48	49.77	30.23	-2.05	0.90	8.85	2.81	5.08
CMS-4402	LINOLEIC ACID	LOW	2.53	-3.15	2.00	14	2	1	0	280.45	266.59	0.00	263.32	37.30	34.14	-4.93	6.46	19.57	6.28	10.36
CMS-72028	TRANS-RETINYL ASCORBATE	LOW	5.83	0.42	1.27	9	7	3	2	458.54	957.13	1.00	362.17	113.29	49.14	-5.71	5.00	20.30	6.55	10.44
CMS-3741	PROCHLORAZ	LOW	1.47	-0.99	-0.28	6	5	0	0	376.67	376.98	1.00	253.09	47.36	36.38	-4.68	3.44	11.68	3.92	6.28
CMS-345	CODEINE	LOW	0.82	1.61	-1.62	1	4	1	5	299.36	508.56	1.61	220.57	41.93	32.43	-2.14	1.09	10.29	2.89	5.81
CMS-8923	ASCORBYL PALMITATE	LOW	7.13	-1.33	3.22	18	7	3	2	414.53	515.29	1.00	338.17	113.29	43.54	-5.80	6.49	27.90	8.41	16.61
CMS-729	HYDROCORTISONE	LOW	2.59	3.42	-0.48	2	5	3	7	362.46	683.89	1.35	279.76	94.83	38.10	-1.93	0.29	12.62	3.76	6.64
CMS-7046	D&C BLUE NO. 4	LOW	9.37	1.06	1.51	12	11	3	0	749.89	1548.31	1.00	533.08	169.36	81.69	-8.48	4.00	21.24	6.48	11.83
CMS-7741	AMMONIUM PERFLUOROOCTANOATE	LOW	0.45	-1.41	0.70	7	2	1	0	414.07	530.14	0.00	157.42	37.30	14.81	-4.38	4.70	11.80	3.39	6.58
CMS-2865	FENTANYL	LOW	1.79	-1.87	-0.40	6	3	0	0	336.47	390.83	1.00	283.99	23.55	41.13	-4.66	3.94	15.12	4.54	8.71
CMS-1777	DIBUTYL PHTHALATE	LOW	1.49	-1.64	0.51	10	4	0	0	278.34	270.85	1.00	227.42	52.60	30.23	-4.09	4.33	14.62	4.12	8.08
CMS-72019	3-O-ISOVALERYL NALTREXONE	LOW	3.39	1.17	-0.69	6	6	1	4	425.52	779.87	1.52	315.78	76.07	45.04	-3.32	2.03	14.73	4.10	8.15
CMS-1256	VITAMIN A PALMITATE	LOW	8.57	-6.48	0.96	21	2	0	0	524.86	803.20	1.00	493.18	26.30	65.82	-9.58	11.38	32.62	9.74	19.64
CMS-3355	SUFENTANIL	LOW	1.99	-1.25	-0.17	8	4	0	0	386.55	459.23	1.00	310.51	32.78	44.66	-4.03	2.82	13.96	4.07	7.91
CMS-72003	DIBUTYL SQUARATE	LOW	0.18	-0.61	0.51	8	4	0	0	226.27	274.14	1.00	179.66	52.60	23.27	-2.39	2.38	10.77	3.55	6.52
CMS-72025	3-O-ISOPROPYLOXYCARBONYL NALTREXONE	LOW	3.49	1.41	-0.53	6	7	1	4	427.49	782.03	1.52	307.56	85.30	43.85	-3.37	1.99	14.70	4.10	7.89
CMS-72026	3-O-PIVALYL NALTREXONE	LOW	3.22	1.39	-0.79	5	6	1	4	425.52	802.22	1.52	315.78	76.07	45.04	-3.06	1.66	14.43	4.00	7.79
CMS-3049	MEPERIDINE	LOW	-0.28	-0.74	-0.48	4	3	0	0	247.33	276.22	1.00	205.01	29.54	28.25	-2.90	2.61	9.88	3.00	5.36
CMS-72031	N,N-DIISOPROPYL NALTREXONE-3-O-CARBAMATE	LOW	4.06	1.14	-0.77	6	7	1	4	468.59	849.96	1.52	353.94	79.31	50.19	-3.85	2.49	15.43	4.29	8.23

CMS ID	NAME	Skin pereability category	PC 1	PC 2	PC 3	BondsRot	HAcc	HDon	Stereo	Weight	Complex	Complex Ring	McGowan	TPSA	Polariz	LogS	XlogP	Diameter	Rgyr	Span
CMS-9533	OCTYL SALICYLATE	MED	1.89	-2.06	0.72	9	3	1	0	250.33	227.78	1.00	211.76	46.53	28.31	-4.84	5.84	16.75	4.87	9.79
CMS-72020	3-O-(2-ETHYLBUTYRYL) NALTREXONE	MED	3.75	0.95	-0.61	7	6	1	4	439.54	793.09	1.52	329.87	76.07	46.88	-3.56	2.34	16.31	4.26	8.34
CMS-72021	3-O-ISOBUTYRYL NALTREXONE	MED	3.07	1.44	-0.68	5	6	1	4	411.49	764.70	1.52	301.69	76.07	43.21	-2.89	1.45	15.20	3.96	7.70
CMS-72752	N,N-DIETHYL NALTREXONE-3-O-CARBAMATE	MED	3.54	1.48	-0.58	6	7	1	4	440.53	797.20	1.52	325.76	79.31	46.52	-3.13	1.55	14.66	4.13	7.96
CMS-620	DI(2-ETHYLHEXYL) PHTHALATE	MED	4.20	-3.08	0.26	16	4	0	2	390.56	394.32	1.00	340.14	52.60	44.91	-6.79	7.85	18.91	5.23	10.35
CMS-922	MONOCHLOROACETIC ACID--PROHIBITED	MED	-3.38	0.79	0.94	1	2	1	0	94.50	42.91	0.00	58.72	37.30	7.10	-0.12	0.16	4.68	1.66	2.97
CMS-17099	NALTREXONE	MED	1.74	2.21	-0.67	2	5	2	4	341.40	620.82	1.52	243.76	70.00	35.60	-2.18	0.71	11.00	3.32	6.89
CMS-72032	N,N-DIMETHYL NALTREXONE-3-O-CARBAMATE	MED	2.96	1.91	-0.61	4	7	1	4	412.48	768.69	1.52	297.58	79.31	42.85	-2.47	0.69	14.38	3.89	7.64
CMS-1776	DIETHYL PHTHALATE	MED	-0.10	-0.59	0.17	6	4	0	0	222.24	223.40	1.00	171.06	52.60	22.89	-2.86	2.76	10.55	3.08	5.90
CMS-1517	BROMOACETIC ACID	MED	-3.14	0.66	0.89	1	2	1	0	138.95	42.91	0.00	63.98	37.30	7.79	-0.37	0.51	4.82	1.65	3.59
CMS-311	CHLOROFORM	MED	-3.85	-0.82	-0.41	0	0	0	0	119.38	8.00	0.00	61.67	0.00	8.39	-2.06	2.07	2.94	1.62	1.70
CMS-3937	BROMOCHLOROACETIC ACID	MED	-2.78	0.67	0.59	1	2	1	1	173.39	64.57	0.00	76.22	37.30	9.72	-0.84	0.92	4.82	1.79	3.53
CMS-72018	PROPRANOLOL BENZOATE	MED	2.62	-1.41	-0.28	9	4	1	1	363.45	446.90	1.13	291.24	47.56	42.92	-5.81	5.52	13.46	4.28	7.07
CMS-72027	3-O-TERTIARYBUTYLOXYCARBONYL NALTREXONE	MED	3.72	1.28	-0.65	6	7	1	4	441.52	819.74	1.52	321.65	85.30	45.68	-3.70	2.41	14.68	4.15	7.97
CMS-466	CHLORPYRIFOS	MED	1.12	-1.75	-0.54	6	4	0	0	350.59	302.78	1.00	215.03	40.58	29.69	-5.54	5.44	11.32	3.38	6.65
CMS-263	CATECHOL	MED	-2.30	0.91	0.04	0	2	2	0	110.11	62.93	1.00	83.38	40.46	11.71	-1.48	1.63	5.67	1.84	3.09
CMS-205	BROMODICHLOROMETHANE	MED	-3.63	-0.94	-0.48	0	0	0	0	163.83	13.51	0.00	66.93	0.00	9.09	-2.30	2.42	3.08	1.64	2.02
CMS-72005	DIETHYL SQUARATE	MED	-1.30	0.37	0.23	4	4	0	0	170.16	225.78	1.00	123.30	52.60	15.93	-1.19	0.81	7.76	2.63	4.69
CMS-431	DICHLOROACETIC ACID	MED	-3.09	0.62	0.82	1	2	1	0	128.94	60.57	0.00	70.96	37.30	9.02	-0.59	0.57	4.68	1.79	3.09
CMS-1398	TCA	MED	-2.80	0.44	0.67	1	2	1	0	163.39	83.43	0.00	83.20	37.30	10.95	-1.15	1.03	4.68	1.86	3.18
CMS-304	DIBROMOCHLOROMETHANE	MED	-3.41	-1.07	-0.56	0	0	0	0	208.28	13.51	0.00	72.19	0.00	9.79	-2.55	2.77	3.21	1.73	2.12
CMS-1394	BROMOFORM	MED	-3.24	-1.17	-0.66	0	0	0	0	252.73	8.00	0.00	77.45	0.00	10.49	-2.80	3.13	3.21	1.81	1.85
CMS-72033	PENTYL NALTREXONE-3-O-CARBAMATE	MED	4.89	1.24	0.29	8	7	2	4	454.56	801.55	1.52	339.85	88.10	48.36	-3.75	2.21	20.04	4.93	11.51
CMS-3511	DIBROMOACETIC ACID	MED	-2.67	0.38	0.66	1	2	1	0	217.84	60.57	0.00	81.48	37.30	10.42	-1.09	1.27	4.82	1.85	3.55
CMS-72029	BUTYL NALTREXONE-3-O-CARBAMATE	MED	4.44	1.51	0.17	7	7	2	4	440.53	786.29	1.52	325.76	88.10	46.52	-3.48	1.88	18.79	4.62	10.57
CMS-2352	DIMETHYL PHTHALATE	MED	-0.89	-0.06	0.03	4	4	0	0	194.18	200.16	1.00	142.88	52.60	19.22	-2.18	1.91	8.54	2.64	4.83
CMS-1046	N-NITROSODIETHANOLAMINE	MED	-2.02	2.12	2.47	5	5	2	0	134.13	72.21	0.00	100.49	73.13	12.02	0.79	-1.69	7.54	2.19	3.85
CMS-1091	BENZOYL PEROXIDE	MED	0.54	-1.03	0.06	5	4	0	0	242.23	258.17	1.00	175.48	52.60	25.21	-4.00	3.43	13.45	3.81	6.72
CMS-3147	NICARDIPINE	MED	4.71	0.37	0.83	11	9	1	1	479.53	855.64	1.00	362.48	113.69	50.88	-5.48	3.54	15.02	4.49	8.30
CMS-72030	ETHYL NALTREXONE-3-O-CARBAMATE	MED	3.39	2.17	-0.12	5	7	2	4	412.48	755.88	1.52	297.58	88.10	42.85	-2.80	0.98	14.74	3.99	8.43
CMS-777	ISOPHORONE	MED	-2.28	-0.35	-0.95	0	1	0	0	138.21	186.91	1.00	124.08	17.07	16.41	-1.82	1.75	6.65	2.14	3.60
CMS-1163	O-PHENYLPHENOL	MED	-1.22	-0.79	-0.69	1	1	1	0	170.21	149.20	1.00	138.29	20.23	21.82	-3.34	3.55	9.18	2.60	4.67
CMS-72017	PROPRANOLOL OLEATE	MED	8.87	-5.35	1.56	24	4	1	1	537.82	619.01	1.20	479.78	47.56	66.10	-9.43	10.79	31.24	9.33	18.22
CMS-72034	PROPYL NALTREXONE-3-O-CARBAMATE	MED	3.83	1.91	0.00	6	7	2	4	426.51	771.07	1.52	311.67	88.10	44.69	-3.07	1.32	15.85	4.23	9.41
CMS-49099	HEXYL NICOTINATE	MED	0.30	-1.27	0.46	7	3	0	0	207.27	182.46	1.00	173.60	39.19	23.29	-2.79	2.92	14.21	4.05	8.08
CMS-4058	2,4-D, DIMETHYLAMINE SALT	MED	-0.52	-0.16	0.13	3	3	1	0	221.04	186.23	1.00	137.61	46.53	19.32	-2.74	2.56	9.97	3.16	5.79
CMS-1706	1,1,1-TRICHLOROACETONE	MED	-3.07	-0.42	0.08	1	1	0	0	161.41	82.67	0.00	91.42	17.07	12.15	-1.48	1.32	4.83	1.88	3.37
CMS-31782	DDT	MED	0.42	-2.96	-1.85	3	0	0	0	354.49	250.05	1.00	221.80	0.00	33.40	-6.99	6.65	9.81	3.56	5.52

CMS ID	NAME	Skin pereability category	PC 1	PC 2	PC 3	BondsRot	HAcc	HDon	Stereo	Weight	Complex	Complex Ring	McGowan	TPSA	Polariz	LogS	XlogP	Diameter	Rgyr	Span
CMS-3963	4-TERT-BUTYLCATECHOL	HIGH	-1.10	0.16	-0.12	1	2	2	0	166.22	148.41	1.00	139.74	40.46	19.05	-2.82	3.25	7.96	2.53	4.12
CMS-61741	BENZYL NICOTINATE	HIGH	-0.31	-0.76	-0.12	4	3	0	0	213.23	223.97	1.00	163.93	39.19	23.78	-2.88	2.43	11.61	3.46	5.96
CMS-1597	1,1-DICHLOROPROPANONE	HIGH	-3.37	-0.24	0.22	1	1	0	0	126.97	59.81	0.00	79.18	17.07	10.22	-0.92	0.85	4.78	1.81	3.21
CMS-5235	N,N-DIMETHYLETHYLAMINE	HIGH	-3.77	-0.23	0.25	1	1	0	0	73.14	17.61	0.00	77.20	3.24	9.47	-0.36	0.49	5.44	1.61	2.92
CMS-435	P-DICHLOROBENZENE	HIGH	-2.46	-1.28	-1.33	0	0	0	0	147.00	54.93	1.00	96.12	0.00	14.29	-3.24	3.26	6.24	2.41	3.12
CMS-1541	DIETHYLENE GLYCOL MONOBUTYL ETHER	HIGH	-1.23	-0.01	2.07	8	3	1	0	162.23	66.36	0.00	141.19	38.69	17.37	-0.16	0.44	11.05	3.42	6.39
CMS-4068	DIACETONE ALCOHOL	HIGH	-2.78	0.65	1.03	2	2	1	0	116.16	94.70	0.00	102.84	37.30	12.51	-0.11	0.21	6.61	1.97	3.48
CMS-3660	HEPTANE	HIGH	-2.37	-1.96	0.18	4	0	0	0	100.20	19.22	0.00	109.49	0.00	13.62	-2.86	3.81	9.28	2.64	4.64
CMS-4603	PENTANE	HIGH	-3.25	-1.36	-0.02	2	0	0	0	72.15	7.51	0.00	81.31	0.00	9.95	-2.18	2.92	6.78	1.94	3.40
CMS-58536	BUTYL NICOTINATE	HIGH	-0.61	-0.65	0.24	5	3	0	0	179.22	159.00	1.00	145.42	39.19	19.62	-2.09	2.03	11.77	3.30	6.63
CMS-934	NAPHTHALENE	HIGH	-2.28	-1.23	-1.52	0	0	0	0	128.17	80.61	1.20	108.54	0.00	17.70	-3.20	3.29	7.10	2.09	3.55
CMS-5106	DIETHYLENE GLYCOL MONOBUTYL ETHER ACETATE	HIGH	-0.42	-0.70	1.98	10	4	0	0	204.26	136.08	0.00	170.94	44.76	21.12	-0.98	1.18	12.27	3.90	7.26
CMS-1908	DIETHYLENE GLYCOL MONOETHYL ETHER	HIGH	-2.09	0.57	1.88	6	3	1	0	134.17	47.57	0.00	113.01	38.69	13.70	0.51	-0.45	8.88	2.77	5.03
CMS-167	BIPHENYL	HIGH	-1.64	-1.73	-1.28	1	0	0	0	154.21	100.00	1.00	132.42	0.00	21.18	-3.89	3.96	9.18	2.62	4.59
CMS-1000	2-NITROPROPANE	HIGH	-3.20	0.28	0.58	1	3	0	0	89.09	54.30	0.00	70.55	45.82	8.13	-1.07	1.11	4.28	1.60	2.74
CMS-9	ACETONITRILE	HIGH	-4.31	0.33	0.29	0	1	0	0	41.05	29.30	0.00	40.42	23.79	4.46	-0.13	0.04	3.14	1.17	1.95
CMS-950	NICOTINE	HIGH	-1.78	-0.09	-0.78	1	2	0	1	162.23	146.77	1.00	137.10	16.13	19.48	-1.67	1.11	8.15	2.48	4.38
CMS-4141	DIETHYLENE GLYCOL MONOMETHYL ETHER	HIGH	-2.50	0.85	1.80	5	3	1	0	120.15	38.66	0.00	98.92	38.69	11.86	0.85	-0.88	8.08	2.52	4.30
CMS-10507	METHYL NICOTINATE	HIGH	-1.97	0.25	-0.07	2	3	0	0	137.14	124.78	1.00	103.15	39.19	14.12	-1.04	0.71	8.06	2.33	4.27
CMS-11753	ETHYLENE GLYCOL ISOPROPYL ETHER	HIGH	-2.86	0.47	1.19	3	2	1	0	104.15	35.06	0.00	93.05	29.46	11.22	0.05	0.20	7.09	2.15	3.67
CMS-845	4-METHOXYPHENOL	HIGH	-2.09	0.25	-0.10	1	2	1	0	124.14	74.99	1.00	97.47	29.46	13.54	-1.34	1.51	7.74	2.20	4.20
CMS-158	BENZYL CHLORIDE	HIGH	-2.59	-1.03	-1.06	1	0	0	0	126.58	55.41	1.00	97.97	0.00	14.20	-2.40	2.49	6.19	2.16	3.78
CMS-2413	METHYL P-HYDROXYBENZOATE	HIGH	-1.40	0.39	0.21	2	3	1	0	152.15	136.27	1.00	113.13	46.53	15.46	-1.60	1.55	8.83	2.60	4.69
CMS-5434	ETHYLENE GLYCOL MONOPROPYL ETHER	HIGH	-2.65	0.37	1.42	4	2	1	0	104.15	29.26	0.00	93.05	29.46	11.22	0.15	0.07	8.00	2.40	4.44
CMS-999	1-NITROPROPANE	HIGH	-2.99	0.19	0.82	2	3	0	0	89.09	47.25	0.00	70.55	45.82	8.13	-0.93	0.93	5.56	1.83	3.56
CMS-1765	METHYL ACETATE	HIGH	-3.65	0.26	0.56	1	2	0	0	74.08	40.16	0.00	60.57	26.30	7.00	-0.17	0.23	5.35	1.52	2.83
CMS-455	1,2-DICHLOROPROPANE	HIGH	-3.41	-0.79	-0.32	1	0	0	1	112.99	20.85	0.00	77.61	0.00	10.14	-1.85	2.02	4.88	1.77	3.03
CMS-1847	1,4-XYLENE	HIGH	-2.71	-1.01	-1.17	0	0	0	0	106.17	48.44	1.00	99.82	0.00	14.10	-2.42	2.68	6.82	2.02	3.41
CMS-1895	ETHYLENE GLYCOL MONOETHYL ETHER ACETATE	HIGH	-2.17	-0.12	1.28	5	3	0	0	132.16	80.37	0.00	108.71	35.53	13.14	-0.40	0.47	8.97	2.67	5.00
CMS-302	CHLOROBENZENE	HIGH	-2.85	-0.98	-1.20	0	0	0	0	112.56	46.14	1.00	83.88	0.00	12.36	-2.52	2.64	5.58	1.91	3.42
CMS-4428	ETHYLENE GLYCOL MONOMETHYL ETHER ACETATE	HIGH	-2.59	0.16	1.18	4	3	0	0	118.13	70.07	0.00	94.62	35.53	11.31	-0.07	0.04	8.11	2.39	4.20
CMS-595	ETHYL ALCOHOL	HIGH	-4.14	0.74	0.66	0	1	1	0	46.07	2.75	0.00	44.91	20.23	5.08	0.30	-0.09	4.14	1.20	2.18
CMS-1459	VINYLIDENE CHLORIDE	HIGH	-3.83	-0.81	-0.32	0	0	0	0	96.94	27.02	0.00	59.22	0.00	8.11	-1.94	1.92	3.65	1.52	2.47
CMS-443	ETHYLENE DICHLORIDE	HIGH	-3.74	-0.82	-0.11	1	0	0	0	98.96	6.00	0.00	63.52	0.00	8.30	-1.59	1.71	4.36	1.91	2.18
CMS-509	N,N-DIMETHYLACETAMIDE	HIGH	-3.63	0.36	0.41	0	2	0	0	87.12	58.57	0.00	78.77	20.31	9.68	-0.01	-0.18	5.38	1.67	2.74
CMS-3690	ETHYLENE GLYCOL MONOMETHYL ETHER	HIGH	-3.51	0.92	1.20	2	2	1	0	76.09	14.36	0.00	64.87	29.46	7.55	0.75	-0.70	5.72	1.72	3.06
CMS-527	N,N-DIMETHYLFORMAMIDE	HIGH	-3.91	0.48	0.44	0	2	0	0	73.09	33.87	0.00	64.68	20.31	7.84	0.20	-0.40	4.28	1.53	2.51
CMS-1346	TETRAHYDROFURAN	HIGH	-3.64	0.16	-0.75	0	1	0	0	72.11	22.83	1.00	62.23	9.23	7.98	-0.51	0.50	4.14	1.38	2.23

Annex 7: Summary statistics calculated for particular properties within each category of skin permeability potential (chapter 5)

Skin permeability category	BondsRot (the number of rotational bonds in a molecule)					
	n	Minimum	Maximum	Range	Mean	Median
HIGH	38	0	10	10	2.13	1
LOW	36	0	21	21	5.61	5
MED	38	0	24	24	4.63	4

Skin permeability category	H-Acc (the number of hydrogen bond acceptors in a molecule)					
	n	Minimum	Maximum	Range	Mean	Median
HIGH	38	0	4	4	1.63	2
LOW	36	0	11	11	4.56	4
MED	38	0	9	9	3.63	4

Skin permeability category	H-Don (the number of hydrogen bond donors in a molecule)					
	n	Minimum	Maximum	Range	Mean	Median
HIGH	38	0	2	2	0.32	0
LOW	36	0	5	5	1.14	1
MED	38	0	2	2	0.82	1

Skin permeability category	Stereo (the number of tetrahedral stereo- centers in a molecule)					
	n	Minimum	Maximum	Range	Mean	Median
HIGH	38	0	1	1	0.05	0
LOW	36	0	8	8	2.86	2
MED	38	0	4	4	1.21	0

Skin permeability category	MW (molecular weight)					
	n	Minimum	Maximum	Range	Mean	Median
HIGH	38	41.05	213.23	172.18	116.68	114.57
LOW	36	226.27	749.89	523.62	356.31	345.65
MED	38	94.50	537.82	443.32	278.43	232.23

Skin permeability category	Complex (complexity of a molecule)					
	n	Minimum	Maximum	Range	Mean	Median
HIGH	38	2.75	223.97	221.22	64.72	51.37
LOW	36	245.34	1548.31	1302.97	564.59	522.72
MED	38	8.00	855.64	847.64	342.20	224.59

Skin permeability category	ComplexRing (ring complexity of a molecule)					
	n	Minimum	Maximum	Range	Mean	Median
HIGH	38	0.00	1.20	1.20	0.37	0.00
LOW	36	0.00	1.61	1.61	1.17	1.21
MED	38	0.00	1.52	1.52	0.83	1.00

Skin permeability category	McGowan (McGowan volume)					
	n	Minimum	Maximum	Range	Mean	Median
HIGH	38	40.42	170.94	130.52	95.79	95.37
LOW	36	157.42	533.08	375.66	271.04	264.89
MED	38	58.72	479.78	421.06	194.52	172.33

Skin permeability category	TPSA (topological polar surface area)					
	n	Minimum	Maximum	Range	Mean	Median
HIGH	38	0.00	46.53	46.53	22.10	25.05
LOW	36	0.00	169.36	169.36	64.10	53.65
MED	38	0.00	113.69	113.69	48.81	47.05

Skin permeability category	Polariz (mean molecular polarisability of a molecule)					
	n	Minimum	Maximum	Range	Mean	Median
HIGH	38	4.46	23.78	19.32	12.58	12.11
LOW	36	14.81	81.69	66.87	37.89	37.15
MED	38	7.10	66.10	59.00	27.16	23.09

Skin permeability category	Log S (solubility of a molecule in water)					
	n	Minimum	Maximum	Range	Mean	Median
HIGH	38	-3.89	0.85	4.74	-1.17	-1.01
LOW	36	-9.58	-0.89	8.69	-3.87	-3.47
MED	38	-9.43	0.79	10.22	-2.97	-2.79

Skin permeability category	Log P (octanl/water partition coefficient of a molecule)					
	n	Minimum	Maximum	Range	Mean	Median
HIGH	38	-0.88	3.96	4.84	1.25	1.02
LOW	36	-0.48	11.38	11.86	3.07	2.57
MED	38	-1.69	10.79	12.48	2.55	1.99

Skin permeability category	Diameter (molecular diameter)					
	n	Minimum	Maximum	Range	Mean	Median
HIGH	38	3.14	12.27	9.13	6.99	6.80
LOW	36	8.85	32.62	23.77	13.87	12.29
MED	38	2.94	31.24	28.30	10.83	10.26

Skin permeability category	Rgyr (molecular radius of gyration)					
	n	Minimum	Maximum	Range	Mean	Median
HIGH	38	1.17	3.90	2.73	2.19	2.12
LOW	36	2.73	9.74	7.01	4.16	3.72
MED	38	1.62	9.33	7.71	3.26	3.24

Skin permeability category	Span (molecular span)					
	n	Minimum	Maximum	Range	Mean	Median
HIGH	38	1.95	7.26	5.31	3.83	3.55
LOW	36	5.00	19.64	14.64	7.68	6.62
MED	38	1.70	18.22	16.52	6.17	5.85

Annex 8: Abstracts of conference presentations related to the present PhD programme

Profiling Data-Rich Areas of Cosmetics Inventories to Increase Confidence in Read-Across

M. T. D. Cronin, C. Yang, A. Bassan, E. Fioravanzo, J. Liu, J. C. Madden, A. S. Mostrag-Szlichtyng, J. F. Rathman, C. H. Schwab, A. Tarkhov

Poster Presentation at the Society of Toxicolgy (SOT) 56th Annual Meeting and ToxExpo, Baltimore, Maryland, USA, 13-16 March 2017

The Toxicologist: Late-Breaking Supplement, Abstract #3390

Chemoinformatics tools allow for the investigation and mining of chemical inventories linked to toxicological data and effects. This study has characterized inventories of cosmetics ingredients associated with repeat-dose toxicity data. The purpose was to identify overlaps and areas of unique chemical space between inventories to determine toxicologically data-rich areas to facilitate data mining and read-across. Three cosmetics inventories available through COSMOS DataShare Point were characterized, namely the US Cosmetics Ingredient Review (CIR), CosIng (European Union) and Korean Cosmetics Industry Institute (KCII). After removing botanicals and polymers, over 7,000 unique chemical structures were analyzed for chemical and biological activity space based on physico-chemical properties and ToxPrint chemotypes. Data-rich regulatory inventories of food-related chemicals from the European Food Safety Authority (EFSA), US FDA's Priority-based Assessment of Food Additives (PAFA) and the Registered Substance Database of European Chemicals Agency (ECHA), in addition to the toxicity data from COSMOS DB v2, were projected onto the chemical space of the three cosmetics inventories. Chemical space was analyzed with Principal Component Analysis (PCA) and 2-D clustering for grouping and visualization. Analyses identified areas of overlap between the cosmetics and the data-rich inventories. Although the cosmetics inventories showed significant overlap, only 10% of the structures appeared in all three inventories. Therefore, the geographical dependence of chemical space could be leveraged to expand the general data profile. Data-rich, with regard to repeat-dose toxicity data, chemical space increases confidence in techniques such as read-across, as common drivers to organ-level toxicity may be observed. There are clear advantages in bringing together inventories of cosmetics ingredients, especially when the underlying toxicity data are available, as they increase the number and quality of data points. The analysis also demonstrated the need to include information on bioavailability in a more comprehensive manner to support read-across predictions.

In Vivo Data Mining and In Silico Metabolic Profiling to Predict Diverse Hepatotoxic Phenotypes: Case study of Piperonyl Butoxide

V. Vitcheva, A. Mostrag-Szlichtyng, O. Sacher, B. Bienfait, C. H. Schwab, A. Richarz, I. Tsakovska, M. Al Sharif, I. Pajeva, C. Yang

Poster presentation at 51st Congress of the European Societies of Toxicology, EUROTOX 2015, Porto, Portugal, September 13-16, 2015

Toxicology Letters 238(2), Supplement, 2015, S173

http://dx.doi.org.proxy.lib.ohio-state.edu/10.1016/j.toxlet.2015.08.586

Piperonyl butoxide (PBO) is a synergist used in a wide variety of insecticides. Its toxicity was extensively investigated in animal studies and liver was identified as the main target organ. The dependence of severity and type of hepatotoxic effects on the duration of exposure to PBO was confirmed in the scientific literature: the short-term exposure leads to the mild changes (liver steatosis and enlargement associated with hepatocyte hypertrophy), whereas the long-term exposure (or higher dosage) yields more severe effects, including necrosis and liver cancer. The potential of PBO for binding to the peroxisome proliferator-activated receptor gamma (PPARgamma), involved in the liver steatosis adverse outcome pathway, was suggested in our previous research, involving mechanistic mining of the *in vivo* data from COSMOS oral repeated dose toxicity database (oRepeatToxDB) and molecular modelling methods (pharmacophore modeling and docking), and was confirmed by the extensive literature search. In the current study we investigate the role of different metabolic pathways in diverse hepatotoxic effects elicited by PBO. Two compounds were used as reference: safrole – weak hepatocarcinogen structurally similar to PBO, and ethyleneglycol – supposedly associated with liver steatosis. MetaboGen (Molecular Networks GmbH) software tool was used to predict the formation of PBO metabolites, and showed that PBO undergoes two major metabolic pathways: opening of the methylenedioxyphenol ring and oxidation on the glycol side chains. Hepatocarcinogenicity observed in long-term studies (but not steatogenic activity) associated with the conversion of the ring methylenedioxy group to a carbene forming ligand complexes with the haem moiety of cytochromes *P*-450 was proposed for PBO, due to its structural similarity to safrole, acting through this pathway. On the contrary, the glycol side chain of PBO is proposed to be responsible for the prosteatogenic mode of action upon short-term exposure. The present case study demonstrates how metabolic profiling can be applied for investigating chemically induced liver toxicity, underlying mechanisms and modes of action, as well as for providing rationales and basis for further discovery of chemotypes associated with the liver toxicity.

Supported by EU FP7 COSMOS Project (266835).

COSMOS DB as an International Share Point for Exchanging Regulatory and Toxicity Data of Cosmetics Ingredients and Related Substances

C. Yang, D. P. Hristozov, A. Tarkhov, T. Kleinoeder, I. Boyer, M. T. D. Cronin, E. Fioravanzo, H. J. Kim, B. Heldreth, A. Mostrag-Szlichtyng, J. F. Rathman, A. N. Richarz, C. H. Schwab, V. Vitcheva, A. P. Worth

Poster presentation at 51st Congress of the European Societies of Toxicology, EUROTOX 2015, Porto, Portugal, September 13-16, 2015

Toxicology Letters 238(2), Supplement, S382

http://dx.doi.org.proxy.lib.ohio-state.edu/10.1016/j.toxlet.2015.08.1090

Since the public release of the COSMOS database v1.0 in December 2013, there has been much interest in connecting the database with other external sites to incorporate regulatory content as well as to enhance the repeated dose toxicity data. The ultimate consortium goal of SEURAT is to develop methods for the eventual replacement of animal testing of cosmetic products for repeated-dose toxicity and biokinetics. To this end, the legacy data and opinions housed at the Scientific Committee of Consumer Safety (SCCS) are important resources. Currently in Europe, the regulatory opinions related to chemicals used as cosmetics ingredients or in formulations are only available from the Scientific Committee. The COSMOS team has remodeled the data model in order to accommodate regulatory data such that document-centered regulatory needs can be compatible with the chemical-centered COSMOS DB. During the assessment workflow, it is essential to easily identify the critical NOAEL values from key studies leading to risk assessment decisions, whilst intuitively linking to the underlying toxicity data that support the decision. This database also houses the critical point of departure data identified by the COSMOS TTC project in collaboration with ILSI Europe as well as other TTC datasets such as Munro (non-cancer) and CPDB (cancer). These TTC datasets can be exported from the database as relevant tables. In summary, this poster will demonstrate the power of the COSMOS DB as an international share point in a variety of regulatory use cases. This abstract does not reflect the policy of CIR, JRC, EC, or KCII (Korea).

Supported by the EU FP7 and Cosmetics Europe.

In Silico Approaches to Support Liver Toxicity Screening of Chemicals: Case Study on Molecular Modelling of Ligands - Nuclear Receptors Interactions to Predict Potential Steatogenic Effects

I. Tsakovska, M. Al Sharif, E. Fioravanzo, A. Bassan, S. Kovarich, V. Vitcheva, A. Mostrag-Szlichtyng, C. Yang, F. Steinmetz, M. Cronin

Poster presentation at 51st Congress of the European Societies of Toxicology, EUROTOX 2015, Porto, Portugal, September 13-16, 2015

Toxicology Letters 238(2): S173

http://dx.doi.org.proxy.lib.ohio-state.edu/10.1016/j.toxlet.2015.08.585

In the Mode of Action/Adverse Outcome Pathway (MoA/AOP) framework addressing repeated dose toxicity, liver steatosis has been recognised as one of the initial manifestations of liver toxicity. The interaction of exogenous chemicals with nuclear receptors (NRs) involved in lipid homeostasis is one of the molecular initiating events (MIEs) triggering the development of liver steatosis. Within the EU COSMOS project different in silico methodologies, including (Q)SAR and molecular modelling, have been employed and integrated for the evaluation of potential binding to NRs involved in the development of liver steatosis, namely LXR (liver X receptor), and PPARγ (peroxisome proliferator-activated receptor γ). The present study further tests and exploits the use of molecular modelling approaches in the AOP framework. It is based on: (i) theoretically described AOPs leading to liver steatosis whose molecular initiating event is a ligand interaction with LXR and PPARγ; (ii) the knowledge about PPARγ as positive transcriptional regulator of LXR expression. Exploring binding to both LXR and PPARγ is the main objective of the study since dual PPARγ/LXR binders could be of higher concern in relation to potential prosteatotic effects. Pharmacophore models were first built on the knowledge of interactions with NRs and validated by means of datasets including known LXR and PPARγ binders. A dataset of chemicals with liver phenotypic effects was then extracted from the COSMOS repeated dose toxicity database (http://cosmosdb.cosmostox.eu), and it was screened with the developed approach hitting some potential dual PPARγ/LXR ligands. This study confirms the utility of molecular modelling approaches to assist in the screening of chemicals to prioritise potential liver toxicants according to given MIEs.

Supported by the EU FP7 COSMOS Project.

Chemical and mechanistic similarity based assessment of the cosmetics space supporting the evaluation of cosmetics-related substances

A-N. Richarz, S. J. Enoch, E. Fioravanzo, S. Kovarich, J. C. Madden, C. Mellor, A. Mostrag-Szlichtyng, A. Palczewska, K. Przybylak, F. Steinmetz, I. Tsakovska, C. Yang, M. T. Cronin

Poster presentation at 51st Congress of the European Societies of Toxicology, EUROTOX 2015, Porto, Portugal, September 13-16, 2015

Toxicology Letters 238(2), Supplement, 2015, S170

http://dx.doi.org.proxy.lib.ohio-state.edu/10.1016/j.toxlet.2015.08.578

The evaluation of cosmetics-related substances by alternative methods is encouraged with the full cosmetics testing ban of the Cosmetics Regulation entering into force in March 2013. Furthermore, the support provided by computational approaches contributes to guide the assessment process from an early stage. In order to support this process, a practical workflow has been developed and implemented in a user-friendly online tool using the KNIME technology. The aim of the workflow is to help the user assess a new compound with regard to its position within "cosmetics space" relative to known cosmetics-related substances as well as similar chemicals in a user-defined sub-group. The cosmetics space was defined by the compilation of cosmetics-related substances in the COSMOS Cosmetics Inventory, which includes over 19000 unique substances, of which more than 5500 have defined structures. The user is able to choose sub-spaces based on cosmetics use classes according to CosIng or on functional groups. The most similar compounds to the target chemical within these sub-spaces are identified and evaluated in the chemical space taking into consideration physico-chemical properties, general molecular fragments of concern, specific structural features or in silico profilers flagging, e.g., potential binding to nuclear receptors, proteins or liver toxicity. The target chemical can be assessed compared to the categorised similar substances or within overall cosmetics space and thus support the evaluation in view of further safety assessment. The workflow also has the flexibility to be extended further, for example to include assessment related to metabolism and bioavailability or to take route of exposure into account.

The funding from the European Community's 7th Framework Program (FP7/2007–2013) COSMOS Project (Grant Agreement N° 266835) and Cosmetics Europe is gratefully acknowledged.

From PPARγ ligand dependent dysregulation to liver steatosis; MoA description and molecular modelling study

M. Al Sharif, I. Tsakovska, I. Pajeva, P. Alov, E. Fioravanzo, A. Bassan, S. Kovarich, A. Mostrag-Szlichtyng, V. Vitcheva, C. Yang

Poster presentation at 8th International Symposium on Computational Methods in Toxicology and Pharmacology Integrating Internet Resources (CMPTI), 2015, 21-25 June 2015, Chios, Greece

Mode of Action and Adverse Outcome Pathway (MoA/AOP) are key elements in the toxicological knowledge framework that are being built to support chemical risk assessment based on mechanistic reasoning. Peroxisome Proliferator-Activated Receptor gamma (PPARγ) is a nuclear receptor with wide tissue expression. In adipocytes it regulates insulin sensitivity and lipid synthesis and storage. PPARγ activation in hepatocytes has been recently proposed as one of the molecular initiating events (MIE) involved in liver steatosis/steatohepatitis [1]. This presentation summarises the application of different methodologies to investigate the involvement of PPARγ in the pathogenesis of fatty liver disease. As a first step MoA is proposed starting with MIE PPARγ ligand activation, passing through a number of intermediate events, and ending with liver steatosis [2]. Further a combination of different molecular modelling methodologies (docking, pharmacophore modelling, 3D QSAR) are applied in order to screen chemicals based on their potential to interact with and activate PPARγ. The results provide the basis for both prioritizing compounds potentially of major concern (for liver toxicity) and / or grouping chemicals potentially sharing the specific AOP [3, 4].

Acknowledgment: Funding from the European Community's 7th Framework Program COSMOS Project (grant n°266835) and from Cosmetics Europe, and from the Ministry of Education, Youth and Science, Bulgaria (grant n°D01-169/14.07.2014) is gratefully acknowledged.

References

[1] B. Landesmann, , M. Goumenou, S. Munn, M. Whelan, Reference Report by the Joint Research Centre of the European Commission, Institute for Health and Consumer Protection, 2012, EUR 25631 EN 40 Pages.

[2] M. Al Sharif, P. Alov, V. Vitcheva, I. Pajeva, I.Tsakovska, PPAR Res. 2014(2014) 1-13.

[3] I. Tsakovska, M. Al Sharif, P. Alov, A. Diukendjieva, E. Fioravanzo, MTD Cronin, I. Pajeva, Int. J. Mol. Sci. 15(2014) 7651-7666.

[4] A. S. Mostrag-Szlichtyng, V. Vitcheva, M.D. Nelms, P. Alov, I. Tsakovska, S.J. Enoch , A. P. Worth, M.T.D. Cronin, C. Yang, SOT 53rd Annual Meeting, 24–27 March 2014, Phoenix, Arizona, USA.

Classification of skin permeability potential following dermal exposure to chemicals to support safety assessment

A. Mostrag-Szlichtyng, J. Rathman, B. Hobocienski, F. Steinmetz, J. Madden, M. Cronin, C. Schwab, D. Hristozov, C. Yang

Poster Presentation at the Society of Toxicolgy (SOT) 54th Annual Meeting and ToxExpo, San Diego, California, USA, 23-26 March 2015

The Toxicologist: Late-Breaking Supplement 144(1), Abstract #2643

Degree of dermal absorption/permeation of chemical has impact on its bioavailability and potential toxicity after topical exposure. We present a set of rules to categorize a query molecule based on skin permeability potential (low/med/high). Skin Permeability Database (developed in the EU COSMOS Project) contains >450 chemicals with data rigorously curated from existing databases and by harvesting literature/ regulatory sources. Systematic quality control was used to minimize concerns about data accuracy and reliability. For the rules formulation and validation we used 280 compounds (split into training/test sets) with data on 2 parameters key to understanding skin permeability: in vitro steady-state flux, J and permeability coefficient, Kp. Computational methods for classifying compounds as low/med/high with respect to J and Kp were developed; the descriptors used were structural fragments encoded with electronic properties (ToxPrint chemotypes) and selected physicochemical properties. Principle component (PC) analysis was used to identify differentiating descriptors and compensate for descriptors intercorrelations. The chemotype frequencies and mean values and ranges of properties were determined and used to develop profile for each category. For instance, chemotype-based PC projection plots reveal the chemotypes useful for assigning the molecules to low J category (cyclic alkane/alkene ketones, cyclic alkanes, fused rings, alicyclic amines), while the physicochemical property-space plots indicated the usefulness of H-bond donors/acceptors number, polar surface area, McGowan volume, molecular weight, and logP for identification of high J category compounds. This research supports further modeling of dermal absorption/permeation and skin sensitization to assess safety of dermal exposure to chemicals.

Description of the MoA/AOP linked with PPAR gamma receptor dysregulation leading to liver fibrosis

V. Vitcheva, M. Al Sharif, I. Tsakovska, P. Alov, A. Mostrag-Szlichtyng, M. T. D. Cronin, C. Yang, I. Pajeva

Poster Presentation at 50th Congress of the European Societies of Toxicology, EUROTOX 2014, 7-10 September 2014, Edinburgh, Scotland, UK

Toxicology Letters, 229, Supplement, S49, P-1.32

http://dx.doi.org/10.1016/j.toxlet.2014.06.211

Modes of Action and Adverse Outcome Pathways (MoAs/AOPs) are key elements in the toxicological knowledge framework that are being built to support chemical risk assessment based on mechanistic reasoning. Peroxisome Proliferator-Activated Receptor gamma (PPARgamma) is a nuclear receptor that regulates adipocyte differentiation, insulin sensitivity and lipid synthesis and storage in adipocytes. PPARgamma activation in hepatocytes is regarded as one of the molecular initiating events (MIE) involved in liver steatosis/stetatohepatitis. However its inhibition in the hepatic stellate cells (HSCs) results in their activation that is essential for the pathogenesis of liver fibrosis. In the current study a systematic literature search has been performed and a MoA scheme based on the PPARgamma dysregulation in stellate cells and resulting in hepatic fibrosis is proposed. Literature data revealing the role of PPARgamma in HSCs are consistent and associate its depletion with HSC activation and fibrosis, whereas increasing PPARgamma expression results in HSC quiescence. A large body of literature confirms that PPARgamma agonists have anti-proliferative and anti-fibrotic effects on activated HSCs. Two applications have been defined from the AOP for fibrosis from PPARgamma dysfunction in stellate cells: (i) the description of possible MIEs triggering PPARgamma inhibition/downregulation that result in fibrosis and allowing for the development of structural alerts; (ii) the identification of key events downstream from PPARgamma dysregulation leading to fibrosis that would be suitable for assay development.

Acknowledgement: The funding from the European Community's 7th Framework Program (FP7/2007-2013) COSMOS Project under grant agreement no. 266835 and from Cosmetics Europe is gratefully acknowledged.

In silico ligand screening based on a pharmacophore model of PPARγ full agonists

I. Tsakovska, M. Al Sharif, P. Alov, V. Vitcheva, E. Fioravanzo, A. Mostrag-Szlichtyng, C. Yang, M. T. D. Cronin, I. Pajeva

Poster Presentation at 16th International Workshop on Quantitative Structure-Activity Relationships in Environmental and Health Sciences, 16-20 June 2014, Milan, Italy

Description of the toxicological modes of action (MoAs) from ligand dependent dysregulation of transcriptional regulators to liver toxicity is among the important concepts in the predictive toxicology. The activation of the hepatic peroxisome proliferator-activated receptor gamma (PPARγ) has been outlined as one of the probable molecular initiating events leading to liver steatosis [1, 2]. Thus, modelling of interactions between PPARγ and its full agonists could facilitate understanding of the molecular mechanisms that further trigger downstream events and promote development of liver toxicity. To this aim a pharmacophore model of the PPARγ full agonists has been recently developed based on X-ray complexes of the receptor in the Protein Data Bank (http://www.rcsb.org/) [3]. In this study the model is externally evaluated on a PPARγ ligand database. The database has been created by analysing and systemising literature data. The model is further applied for the *in silico* screening of toxicity databases, including COSMOS Database (http://cosmosdb.cosmostox.eu/). A pharmacophore search is performed for ligands with liver adverse effects (Figure 1). Potential PPARγ full agonists are outlined. The approach could be used for the *in silico* screening of agonists of hepatic PPARγ that can function as steatosis inducers facilitating in this way the process of MoA development.

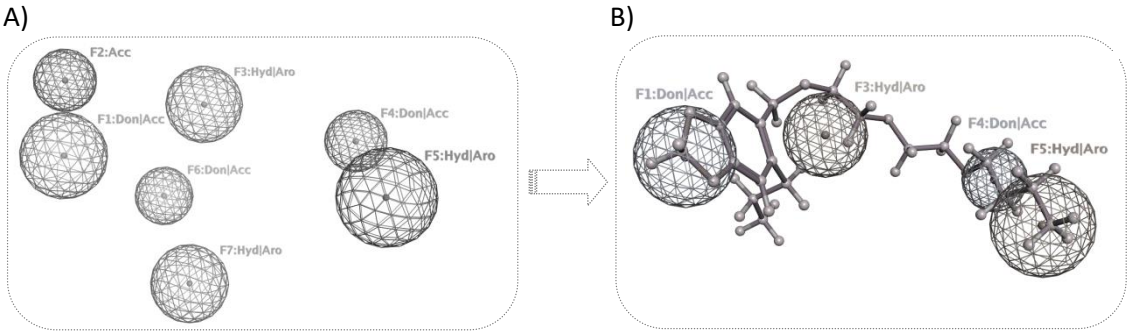


Figure 1: Screening of COSMOS DB: A) The pharmacophore model of PPAR full agonists; B) Pharmacophore search on piperonyl butoxide

The funding from the European Community's 7th Framework Program (FP7/2007-2013) COSMOS Project under grant agreement n°266835 and from Cosmetics Europe is gratefully acknowledged.

1. Landesmann, B.; Goumeou, M.; Munn, S.; Whelan, M. JRC *Scientific and Policy Reports, Office of the European Union*. **2012**.
2. Al Sharif, M.; Alov, P.; Vitcheva, V.; Pajeva, I.; Tsakovska, I. *PPAR Research*, *2014*, **2014**, Article ID 432647.
3. MOE v.2013 .08, Chemical Computing Group Inc., 1010 Sherbooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7, **2013**.

Data Mining Approach to Formulate Alerting Chemotypes for Liver Steatosis / Steatohepatitis / Fibrosis

A. Mostrag-Szlichtyng, V. Vitcheva, M. D. Nelms, P. Alov, I. Tsakovska, S. J. Enoch, A. P. Worth, M. T. D. Cronin, C. Yang

Poster Presentation at the Society of Toxicology (SOT) 53nd Annual Meeting and ToxExpo, Phoenix, Arizona, USA, 24-27 March 2014

The Toxicologist 138(1): Abstract #2254

COSMOS oRepeatToxDB, oral repeat-dose toxicity database, is designed with an ontology describing toxicological effects at each dose level using controlled vocabulary, thus enabling mechanistic data mining. Observations are also coupled to organism-level sites and more specific effects at lower levels are formulated within hierarchical framework: organs/systems -> segments/tissues -> cells/organelles. The majority of biological/chemical processes occur at the cell/organelle level, and so interactions between chemicals and proteins/genes are investigated in order to associate chemical structures with phenotypic effects resulting from related toxicity mechanisms. Furthermore, common structural fragments are extracted and refined into mechanistic chemotypes representing underlying molecular initiating events. We present a data mining case for liver steatosis, steatohepatitis and fibrosis. Over 20% of cosmetics-related chemicals in this database were associated with lipid deposition, fatty changes, cytoplasmic vacuolization, cellular infiltration and inflammation in various hepatocytes, ultimately leading to fibrosis. Combined phenotypic effects and morphological changes at various sites were mapped onto chemical compounds. Applying the ToxPrint chemotypes to these compounds, the set of alerting chemotypes for liver steatosis/steatohepatitis/fibrosis was identified. They include alcohols, diols, glycol ethers, aminophenols, tertiary amines, aromatic amines, polychlorinated short alkanes, halogenated amines, and Michael acceptors. Identification of these alerting chemotypes can be considered as the initial step in developing the categories used in safety/risk assessment. This approach also provides a way to investigate molecular pathways relevant to toxicological mechanisms.

Supported by EU FP7 COSMOS Project.

Alerting Chemotypes for Liver Steatosis, Steatohepatitis and Fibrosis Identified by Mining COSMOS DB

A. Mostrag-Szlichtyng, V. Vitcheva, M. D. Nelms, P. Alov, I. Tsakovska, S. J. Enoch, A. P. Worth, M. T. D. Cronin, C. Yang

Poster Presentation at the SEURAT-1 Fourth Annual Meeting, Barcelona, Spain, 5-6 February 2014

The COSMOS oral repeated dose toxicity database (oRepeatToxDB) includes an ontology for phenotypic effects at each dose level using controlled vocabulary. Toxicity effects observed at target organ sites have been organised hierarchically to relate organs to tissues to cells. The majority of biological/chemical processes occur at the cell/organelle level. Therefore interactions between chemicals and proteins/genes are investigated in order to associate chemical structures with phenotypic effects initiated by related toxicity mechanisms. Common structural fragments are extracted and refined into mechanistic chemotypes representing underlying molecular initiating events (MIE). Liver steatosis, steatohepatitis and fibrosis were chosen as a case sudy for data mining. Over 20% of cosmetics-related chemicals in oRepeatToxDB were associated with lipid deposition, fatty changes, cytoplasmic vacuolisation, cellular infiltration and inflammation in various hepatocytes, ultimately leading to fibrosis. Combinations of phenotypic effects and morphological changes at various sites were mapped onto chemical classes. A set of alerting chemotypes for liver steatosis, steatohepatitis, fibrosis was identified by application of the ToxPrint chemotypes and will be further used for developing chemical categories to be used in safety assessment. This approach also provides a way to elucidate the underlying molecular pathways and mechanisms for hepatotoxicity.

Molecular modelling studies of LXR and PPAR gamma receptors in relation to the MoA/AOP framework for liver steatosis

S. Kovarich, M. Al Sharif, P. Alov, A. Bassan, M.T.D. Cronin, E. Fioravanzo, A. Mostrag-Szlichtyng, I. Pajeva, I. Tsakovska, V. Vitcheva, A. Worth, C. Yang (2014)

Poster presentation at the SEURAT-1 Fourth Annual Meeting, Barcelona, Spain, 5-6 February 2014

The SEURAT-1 cluster adopted the Mode-of-Action/Adverse Outcome Pathway (MoA/AOP) framework to understand human adverse health effects caused by repeated exposure to chemicals, that initiate the sequence of events from the molecular (molecular initiating event, MIE) through higher levels (organelles/cells/tissues/organs) and lead to the perturbations observed at the whole organism level. Within the COSMOS Project innovative in silico approaches are being explored to study the MIEs involved in liver steatosis. This implies the investigation of applicability of molecular modelling (MM) methods to predict the binding of small molecules to two nuclear receptors involved in the liver steatosis MoA, namely the liver X receptor (LXR) and peroxisome proliferator-activated receptor gamma (PPARγ) and to study the ligand-dependent activation of them. The poster presents the MM results of the binding of selected ligands to LXR and PPARγ, including the characterisation of the ligand-binding pocket of the receptors, the identification of ligand-receptor interactions and essential structural features involved in LXR/PPARγ binding. The challenging objective of these studies is to lay the foundations for the application of MM in predictive toxicology as a part of an integrated strategy which combines multiple methods and approaches (e.g., in silico, in vitro, mechanistic information) to support toxicity prediction in the MoA/AOP framework.

Data mining toxicity effects through an ontology approach to investigate toxicity mode of action

V. Vitcheva, A. Mostrag-Szlichtyng, M. Nelms, P. Alov, S. Enoch, I. Tsakovka, J. Rathman, M. Cronin

Poster Presentation at 49th Congress of the European Societies of Toxicology, EUROTOX 2013, Interlaken, Switzerland, 1-4 September 2013

Toxicology Letters 221S (2013) S59-S256, P05-5

http://dx.doi.org/10.1016/j.toxlet.2013.05.081

Toxicological modeling and structure knowledge development begin by connecting biological effects and chemicals involved in pathways. A systematic data mining method has been established to link biological observations of cellular events to chemical reactivity. This method is based on an in vivo oral repeated dose toxicity database equipped with a controlled vocabulary for describing phenotypic effects at the cellular level. For example, hepatocytes, Kupffer cells, sinusoids, and stellate cells are associated with fatty/lipid storage (accumulation, deposits, etc.), Toxicity effects observed at target organ sites have been organized hierarchically to relate organs to tissues to cells, while also mapping biological processes to phenotypic effects. Data mining to elucidate site/effect combinations can suggest causal relationships in toxicity pathways, and plausible hypotheses can then be generated by mapping these combinations onto chemical classes relevant to the compounds responsible for the phenotypic effects. Groupings of chemicals with biologically similar functions can then be generated. As a case study, liver steatosis and fibrosis have been chosen and the relationship between these phenotypic effects and the underlying morphological changes caused by, for example, analogs of vitamin A/retinoids as well as aromatic amines are discussed from mechanistic perspective. This methodology provides a systematic approach for investigating chemically-induced toxicity and elucidating the underlying mechanism, and may further guide studies to determine the mode of action for hepatotoxicity.

Development of new COSMOS oRepeatDose and non-cancer Threshold of Toxicological Concern (TTC) databases to support alternative testing methods for cosmetics related chemicals

C. Yang, M. Ambrosio, K. Arvidson, S. Barlow, A. Boobis, M. Checheva, M. Cronin, S. Felter, E. Fioravanzo, H. Hollnagel, D. Hristozov, K. Jacobs, D. Keller, A. Mostrag-Szlichtyng, M. Nelms, J. Rathman, A. Richarz, I. Tsakovska, S. Vidry, V. Vitcheva, A. Worth A

Poster Presentation at 49th Congress of the European Societies of Toxicology, EUROTOX 2013, Interlaken, Switzerland, 1-4 September 2013

Toxicology Letters 221S (2013) S59–S256, P05-6

http://dx.doi.org/10.1016/j.toxlet.2013.05.082

The Seventh Amendment of the Cosmetics Directive requires replacement of animal testing of cosmetic products for repeated dose/reproductive toxicity and toxicokinetics. To this end, the COSMOS consortium within SEURAT, a cluster of research jointly funded by the European Commission and Cosmetics Europe, has been engaged in development of computational methods and tools. COSMOS has prepared a new Cosmetics Inventory based on the chemical records from the EU COSING database and the list from the US Personal Care Products Council. COSMOS has also developed a new toxicity database enriched with oral repeated dose studies for cosmetics-related chemicals. The sources for toxicity data include US Food and Drug Administration, US Environmental Protection Agency, EU Scientific Committee on Consumer Safety, European Chemical Agency, US National Toxicology Program, and literature publications. A new non-cancer TTC database for cosmetics-related chemicals has been compiled by augmenting the COSMOS database with substances from the Munro dataset found in the Cosmetics Inventory. The resulting TTC database contains over 580 chemical structures with no-observed-adverse-effect levels (NOAELs); the toxicity data for the chemicals in the lowest 10th percentile of the distribution of NOAELs have been further subjected to detailed quality control. The inclusion and selection criteria of the NOAEL decisions have been documented. The chemical space of the new TTC database has been compared with existing TTC databases to demonstrate that the coverage is suitable for the assessment of cosmetics products. The TTC database will be made public and serve as a resource for alternative methods.