

# Key substance identity concepts and dossier preparation

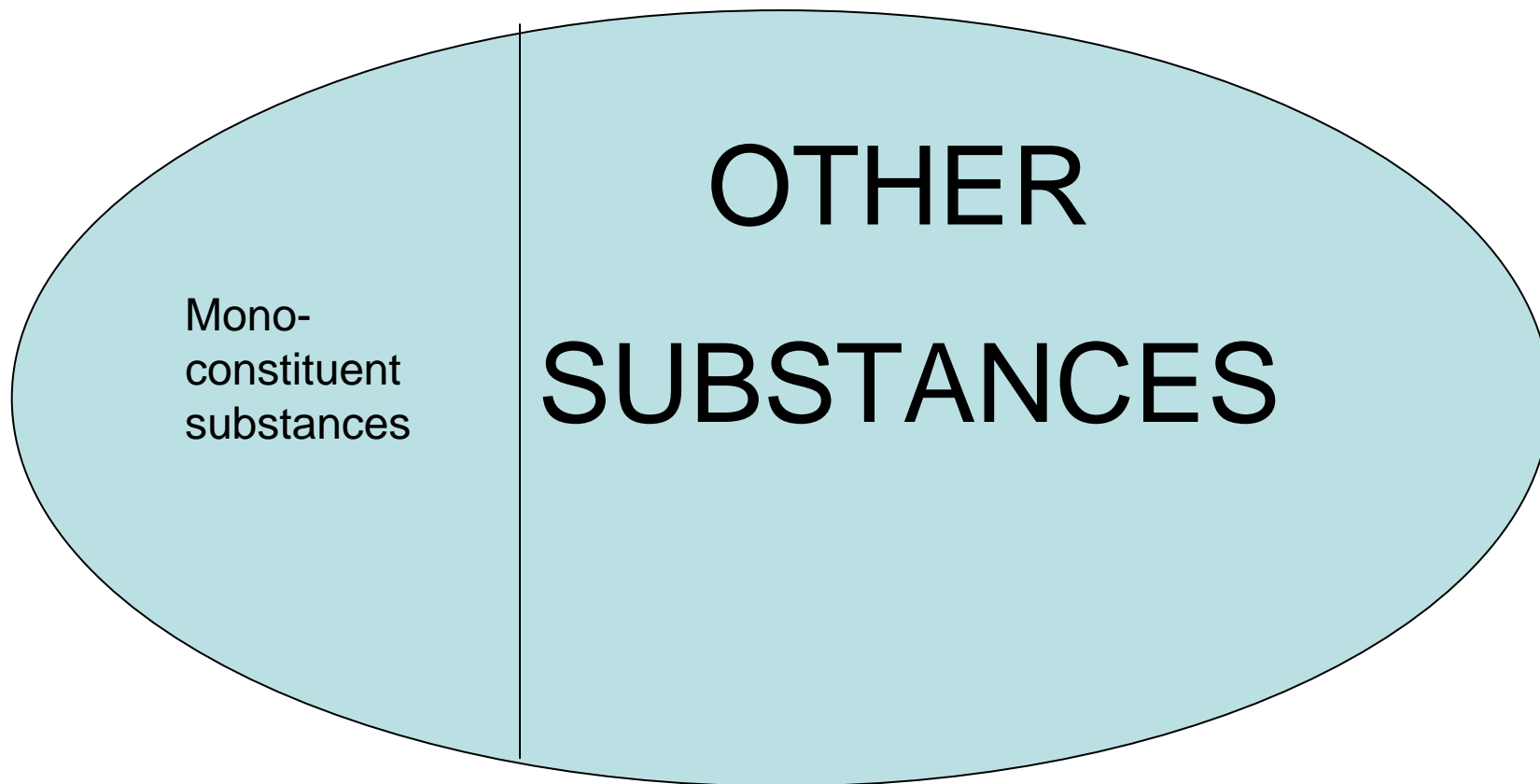
Andreas Uphoff  
Alberto Martin Aparicio

# Overview

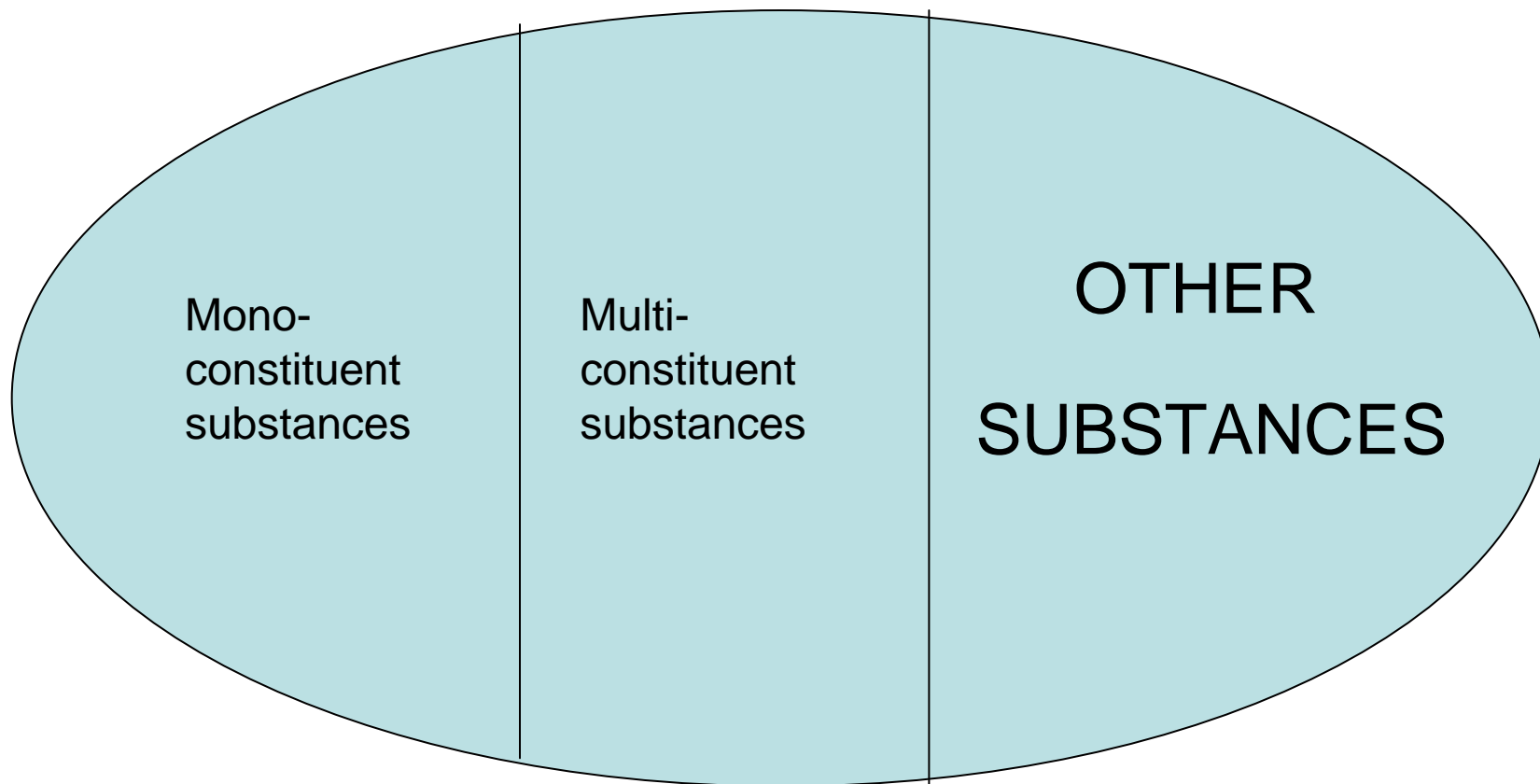
- Introduction
- Required data for substance identity purposes
- Dossier preparation (introduction)
- Substance identity
  - Mono-constituent substances
  - Multi-constituent substances
  - ... with additional identifiers
- Dossier preparation (mono & multi-constituent substances)
  
- BREAK
  
- Substance identity of UVCB substances
- Dossier preparation (UVCB substances)

# Overview

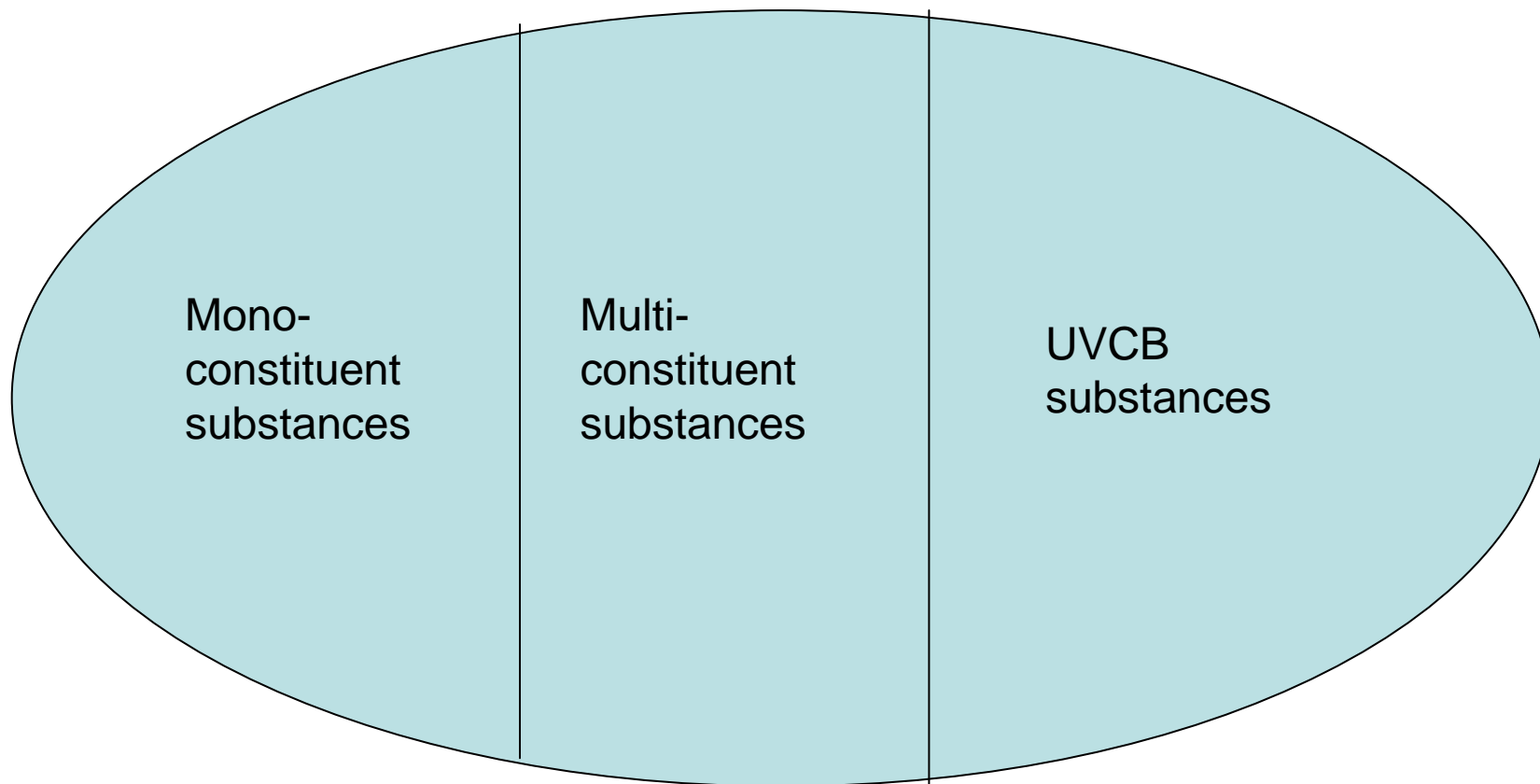
- **Introduction**
  - Required data for substance identity purposes
  - Dossier preparation (introduction)
  - Substance identity
    - Mono-constituent substances
    - Multi-constituent substances
    - ... with additional identifiers
  - Dossier preparation (mono & multi-constituent substances)
  - BREAK
  - Substance identity of UVCB substances
  - Dossier preparation (UVCB substances)
-



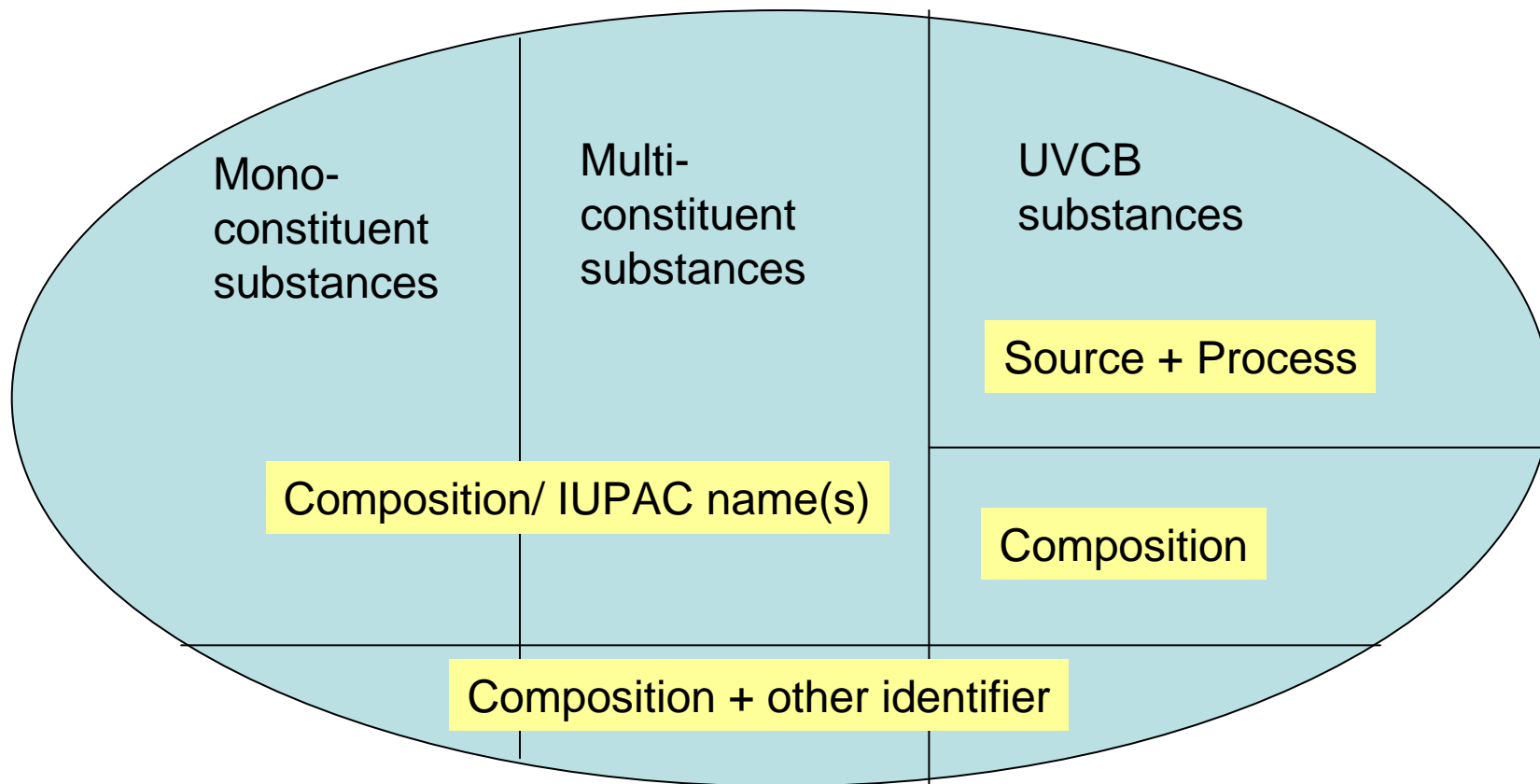
# Introduction



# Introduction



# Introduction



# Overview

- Introduction
- **Required data for substance identity purposes**
- Dossier preparation (introduction)
- Substance identity
  - Mono-constituent substances
  - Multi-constituent substances
  - ... with additional identifiers
- Dossier preparation (mono & multi-constituent substances)
  
- BREAK
  
- Substance identity of UVCB substances
- Dossier preparation (UVCB substances)



## Required data

1. Substance identifiers
2. Composition
3. Qualitative analytical methods and results (identification)
4. Quantitative analytical methods and results (composition)

# Required data

1. Substance identifiers
2. Composition
3. Qualitative analytical methods and results (identification)
4. Quantitative analytical methods and results (composition)

# Required data

## 1. Substance identifiers

- Name derived according to guidance on substance identity (often **IUPAC name**)
- Numerical identifiers if available (**EC** number, **CAS** number, C.I. number, ...)
- Structure if appropriate (human readable structural formula & SMILES code)
- If appropriate: molecular formula (Hill format), molecular weight range,

## Required data

1. Substance identifiers
2. **Composition**
3. Qualitative analytical methods and results (identification)
4. Quantitative analytical methods and results (composition)

# Required data

## 2. Composition

- Purity
- Concentration and identification of each constituent > 1%
  - The composition has to give the ratio of all isomers present
  - Solvents are not considered part of the substance
  - Additives in the sense of REACH need to have stabilising function

## Required data

1. Substance identifiers
2. Composition
3. Qualitative analytical methods and results (identification)
4. Quantitative analytical methods and results (composition)

## Required data

### 3. Qualitative analytical methods and results (identification)

- “shall be sufficient to enable each substance to be identified”
  - UV/Vis, IR,  $^1\text{H}$  NMR
    - If analysis technically impossible: justification needed
  - If necessary:  $^{13}\text{C}$  NMR, MS, XRD, ...
  - All (ionic) parts of the substance need to be determined

## Required data

1. Substance identifiers
2. Composition
3. Qualitative analytical methods and results (identification)
4. Quantitative analytical methods and results (composition)



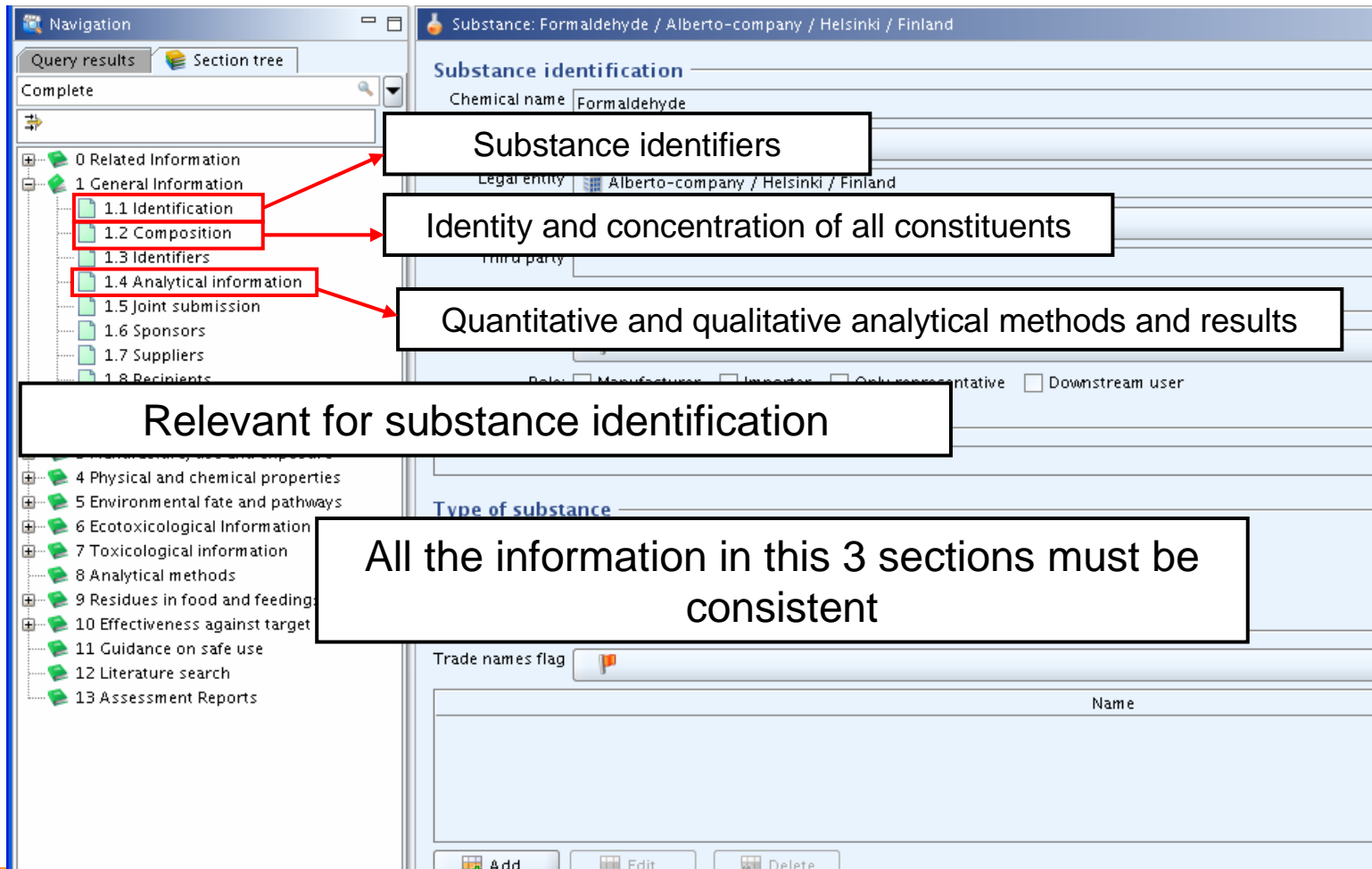
# Required data

## 4. Quantitative analytical methods and results (composition)

- Quantitation of all constituents > 1%
  - For ionic constituents each ion should be quantified
- Typical concentration and concentration ranges
- GC or HPLC
  - If necessary other methods: AAS, titration, ...
- Provide a table of results with assignment to constituents mentioned in the composition
- Explain method of quantitative evaluation (areas, standards, calibration curve)
- Number and concentration range of unknown impurities have to be given
  - Justification needed if significant amount of constituents unknown

# Overview

- Introduction
  - Required data for substance identity purposes
  - Dossier preparation (introduction)
  - Substance identity
    - Mono-constituent substances
    - Multi-constituent substances
    - ... with additional identifiers
  - Dossier preparation (mono & multi-constituent substances)
  - BREAK
  - Substance identity of UVCB substances
  - Dossier preparation (UVCB substances)
-



The screenshot shows the ECHA Substance Identification interface for Formaldehyde. The left sidebar contains a 'Section tree' with 14 categories. Three categories are highlighted with red boxes and red arrows pointing to callout boxes:

- 1.1 Identification** points to a box labeled "Substance identifiers".
- 1.2 Composition** points to a box labeled "Identity and concentration of all constituents".
- 1.4 Analytical information** points to a box labeled "Quantitative and qualitative analytical methods and results".

A larger box labeled "Relevant for substance identification" encompasses sections 1.1, 1.2, and 1.4. Another box at the bottom states "All the information in this 3 sections must be consistent". The main interface shows fields for "Chemical name" (Formaldehyde), "Legal entity" (Alberto-company / Helsinki / Finland), and "Type of substance".

Reference substance: formaldehyde / formaldehyde / Formaldehyde / 50-00-0

---

**General information**

Reference substance name

---

**EC inventory**

EC number  CAS number

EC name

Molecular formula

Description

---

**No EC information available**

Justification

---

**Reference substance information**

---

**CAS information**

CAS number

CAS name

---

**IUPAC name**

---

**Description**

The image shows a software interface for managing chemical substances. In the background, a substance profile for 'Sodium pyrazole' is visible, with sections for 'Substance identification', 'Role in the supply chain', 'Reference substance', 'Type of substance', and 'Trade names'. Two dialog boxes are overlaid on the interface:

- Query**: A dialog box titled 'Find the reference substance' with input fields for 'Reference substance name', 'CAS number (EC Inventory)', 'EC name', 'EC number' (containing '2\*'), and 'CAS number'.
- Reference substance assistant**: A dialog box titled 'Create a new reference substance' with a molecular structure icon. It contains a search field for 'Reference substance name' (filled with 'Sodium Pyrazole'), a checked checkbox for 'Active reference substance', and buttons for '< Back', 'Next >', 'Finish', and 'Cancel'.

Below the 'Reference substance assistant' dialog, a table of search results is displayed:

	<b>Peanut oil /</b>	IUC5-9ab79f04-9...	2009-06-20
	<b>EPBC-1 /</b>	IUC5-cfc04268-b...	2009-06-20
	<b>EPBC /</b>	IUC5-e950df38-2...	2009-06-20
	<b>ECHAsubstance:</b>	IUC5-f3bbb1fb-ce...	2009-07-23

Below the table, it says 'Number of results: 4/4'. At the bottom of the 'Reference substance assistant' dialog, the 'New' button is circled in red.

## Before you start

- Make sure that you have all the relevant identifiers for your substance and all its constituents (name according to guidance, EC#, CAS#, etc)
- If you are using IUCLID:
  - Create beforehand all the reference substances you are going to need
  - Fill them with all the necessary information (identifiers, structural formula, molecular weight, etc)
- Make sure that you have all the necessary electronic files (documents and/or image files) with the relevant analytical information.

# Overview

- Introduction
  - Required data for substance identity purposes
  - Dossier preparation (introduction)
  - **Substance identity**
    - Mono-constituent substances
    - Multi-constituent substances
    - ... with additional identifiers
  - Dossier preparation (mono & multi-constituent substances)
  - BREAK
  - Substance identity of UVCB substances
  - Dossier preparation (UVCB substances)
-

# Mono-constituent substance

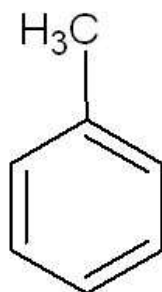
- Composition known
- Concentration of main constituent  $\geq 80\%$
- Main constituent completely identified by IUPAC name



# Mono-constituent substance

- **Borderline cases**
  - Typical concentration  $\geq 80\%$  but lower concentration limit  $< 80\%$ 
    - possibly mono-constituent, justification needed
    - possibly multi-constituent, justification needed
  - Considerable amounts of second constituent that is “important for the desired technical effect of the substance”
    - possibly multi-constituent substance (justification needed)

# Mono-constituent substance



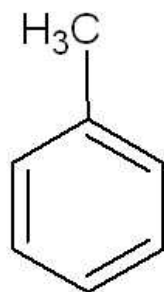
100%

toluene



0%

# Mono-constituent substance



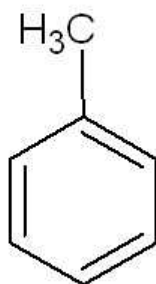
95 – 100%

toluene



0 - 5%

# Mono-constituent substance



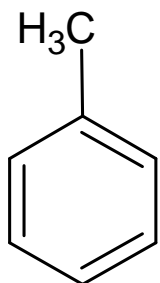
75 – 85 %

toluene



15 – 25 %

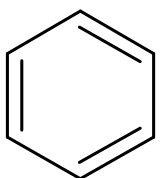
# Mono-constituent substance



75 – 85 %

toluene

reaction mass of toluene and benzene



15 – 25 %

?

# Overview

- Introduction
- Required data for substance identity purposes
- Dossier preparation (introduction)
- **Substance identity**
  - Mono-constituent substances
  - **Multi-constituent substances**
  - ... with additional identifiers
- Dossier preparation (mono & multi-constituent substances)
- BREAK
- Substance identity of UVCB substances
- Dossier preparation (UVCB substances)

## Multi-constituent substance

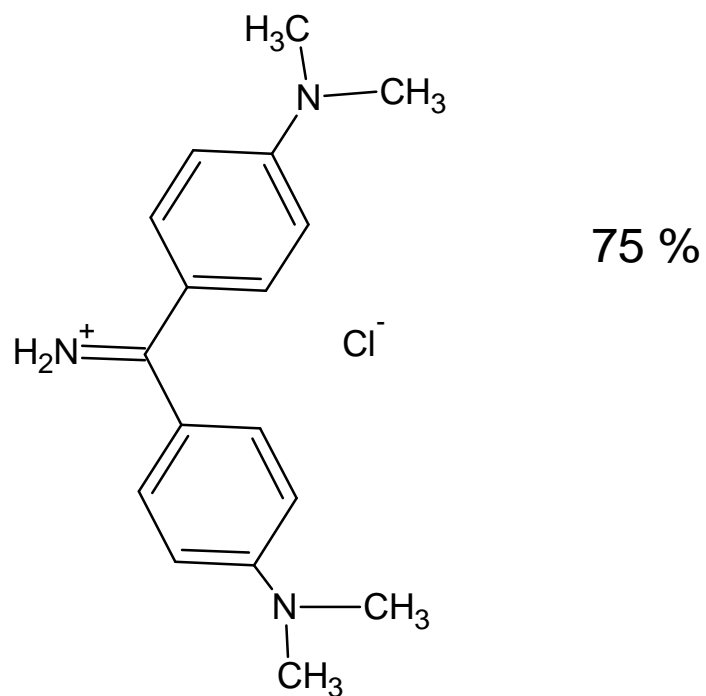
- Composition known and defined
- More than one constituent  $\geq 10\%$  and  $< 80\%$
- Each main constituents completely identified by IUPAC name
  - “Reaction mass of [IUPAC name] and [IUPAC name] and ...”

# Multi-constituent substance

- **Borderline cases**
  - Typical concentration < 80% but higher concentration limit  $\geq 80\%$ 
    - possibly multi-constituent, justification needed
    - possibly mono-constituent, justification needed
  - Constituent  $\geq 10\%$  but clearly “impurity”
    - possibly not included in name (justification needed)



# Multi-constituent substance?



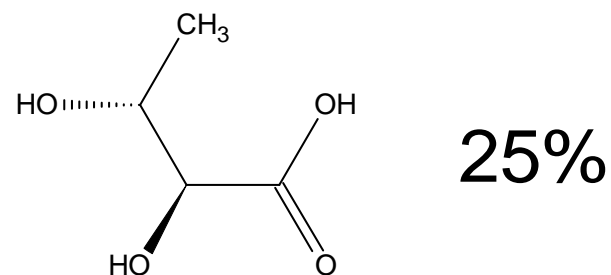
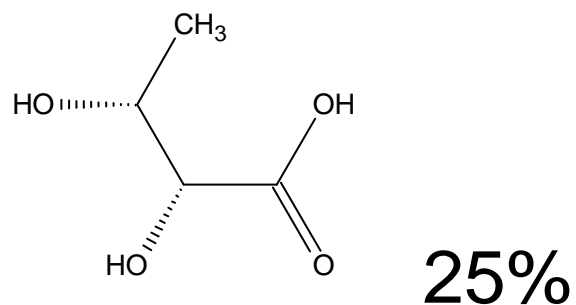
bis[4-(dimethylamino)phenyl]  
methaniminium chloride  
(C.I. Basic Yellow 2)

NaCl 25 %

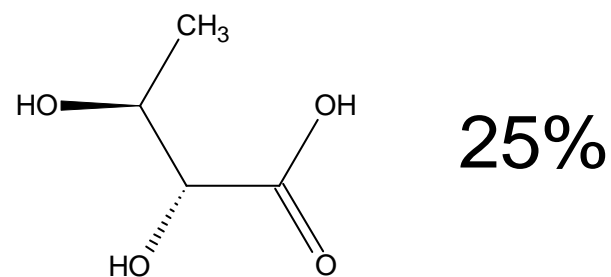
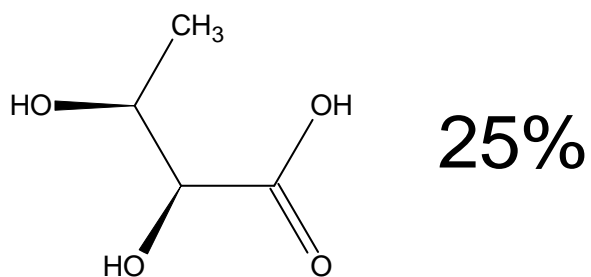
# Multi-constituent substance

- **Isomers**
  - racemic mixture of enantiomers
  - two enantiomers
  - multiple structural isomers
- **Unspecific IUPAC name acceptable**
  - if all isomers present
- **Specification of all isomers in composition**
  - except racemic mixture w/o optical activity
- **“Reaction mass of ...”**
  - if not all isomers present  $\geq 10\%$

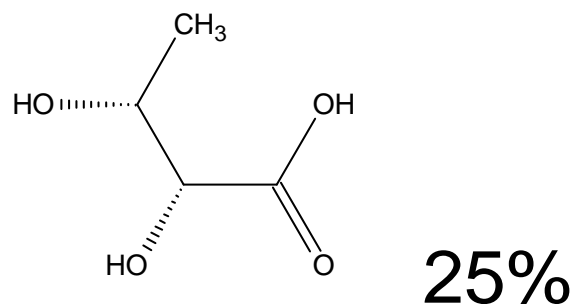
# Multi-constituent substance



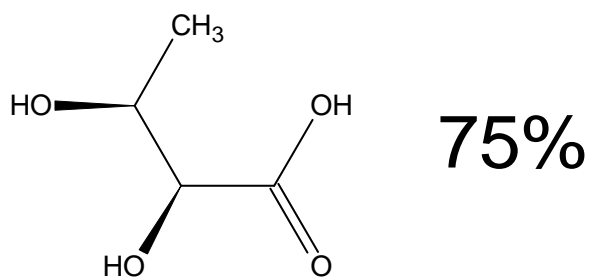
## 2,3-dihydroxybutanoic acid



# Multi-constituent substance



Reaction mass of  
(2R,3R)-2,3-dihydroxybutanoic acid and  
(2S,3S)-2,3-dihydroxybutanoic acid



# Overview

- Introduction
- Required data for substance identity purposes
- Dossier preparation (introduction)
- **Substance identity**
  - Mono-constituent substances
  - Multi-constituent substances
  - ... with additional identifiers
- Dossier preparation (mono & multi-constituent substances)
- BREAK
- Substance identity of UVCB substances
- Dossier preparation (UVCB substances)

# Additional identifiers

- Composition known
- Otherwise mono- or multi-constituent as before
- BUT: IUPAC name not enough to identify the substance
  - E.g. inorganic, crystalline material
- If possible structure as additional identifier
  - If not: hardness, swelling capacity, density, surface area,...

# Additional identifiers

- Examples:
  - Rutile (TiO<sub>2</sub>)  
(**EC#** : 215-282-2, **CAS#** : 1317-80-2)
  - Anatase (TiO<sub>2</sub>)  
(**EC#** : 215-280-1, **CAS#** : 1317-70-0)
  - Spinel (Mg(AlO<sub>2</sub>)<sub>2</sub>)  
(**EC#** : 215-105-9, **CAS#** : 1302-67-6)

# Additional identifiers

- Note:
    - Data requirements remain the same:  
“shall be sufficient to enable each substance to be identified”
    - Structural information is required even if the IUPAC name is sufficient for identification
  
    - Applies also to UVCB substances:
      - **EC#** : 310-193-6
      - **CAS#** : 1345-16-0
      - **Substance Name:** Cobalt aluminate blue spinel
      - **Description:** An inorganic pigment that is the reaction product of high temperature calcination in which cobalt (II) oxide and aluminum oxide in varying amounts are homogeneously and ionically interdiffused to form a crystalline matrix of spinel. Its composition may include any one or a combination of the modifiers MgO, ZnO, Li<sub>2</sub>O, or TiO<sub>2</sub>.
-



# Overview


- Introduction
  - Required data for substance identity purposes
  - Dossier preparation (introduction)
  - Substance identity of
    - Mono-constituent substances
    - Multi-constituent substances
    - ... with additional identifiers
  - Dossier preparation (mono & multi-constituent substances)
  - BREAK
  - Substance identity of UVCB substances
  - Dossier preparation (UVCB substances)
-

# IUCLID 5 Dossiers


Dossier preparation for mono-  
constituent substances

**Substance identification**

Chemical name: Sodium pyrazole


Legal entity flags: 

Legal entity: Alberto-company / Helsinki / Finland

Third party flags: 

Third party:

**Role in the supply chain**

Role flags: 

Role:  Manufacturer  Importer  Only representative  Downstream user

**Reference substance**

Sodium Pyrazole / sodium pyrazol-1-ide / 1H-pyrazole, sodium salt (1.1) / 40958-82-5

EC number	EC name
255-164-8	1H-pyrazole, sodium salt
CAS number	CAS name
40958-82-5	1H-pyrazole, sodium salt (1.1)
IUPAC name	
sodium pyrazol-1-ide	

**Type of substance**

Composition: mono constituent substance

Origin:

**Fill in the fields of your reference substance as much as possible**

**Select from the picklist**

# 1.1 Identity (reference substance)

Reference substance: Sodium Pyrazole

**CAS information**

CAS number: 40958-82-5

CAS name: 1H-pyrazole, sodium salt (1.1)

**IUPAC name**

sodium pyrazol-1-ide

**Description**

**Synonyms**

Name
Additional identifiers can be provided in this field (e.g. Colour Index number)

Add... Edit... Delete

If using a reference substance downloaded from IUCLID inventory: check pre-filled information for correctness

Additional identifiers can be provided in this field (e.g. Colour Index number)

# 1.1 Identity (reference substance)

**Molecular and structural information**

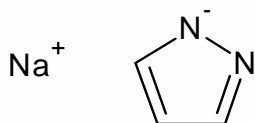
Molecular formula

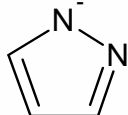
Molecular weight range

SMILES notation

InChI

Structural formula



Na<sup>+</sup> 

Remarks **Provide any necessary further information about the substance in this field**

# 1.2 Composition (I)

Substance composition

Sodium pyrazole

Name: Sodium pyrazole

Brief description:

Degree of purity

> 96 < 100 % (w/w)

Same as concentration range for mono-constituent substances

Constituents

98 % (w/w) Sodium Pyrazole / sodium pyrazol-1-ide / 1H-pyrazole, sodium salt (1.1) / 40958-82-5

Reference substance: Sodium Pyrazole / sodium pyrazol-1-ide / 1H-pyrazole, sodium salt (1.1) / 40958-82-5

EC number	EC name
255-164-8	1H-pyrazole, sodium salt
CAS number	CAS name
40958-82-5	1H-pyrazole, sodium salt (1.1)
IUPAC name	sodium pyrazol-1-ide

Typical concentration: 98 % (w/w)

Concentration range: > 96 < 100 % (w/w)

Provide concentration as a range

Remarks:

# 1.2 Composition (II)

Substance composition

Sodium pyrazole

Name: Sodium pyrazole

Brief description:

Degree of purity

> 96 < 100 % (w/w)

Constituents

98 % (w/w) Sodium Pyrazole / sodium pyrazol-1-ide / 1H-pyrazole, sodium salt (1.1) / 40958-82-5

Reference substance: Sodium Pyrazole / sodium pyrazol-1-ide / 1H-pyrazole, sodium salt (1.1) / 40958-82-5

EC number	EC name
255-164-8	1H-pyrazole, sodium salt
CAS number	CAS name
40958-82-5	1H-pyrazole, sodium salt (1.1)

Typical

Concentration range > 96 < 100 % (w/w)

Remarks

Use this remarks field to justify possible deviations from the guidance (80% rule)

# 1.2 Composition (III)

**Impurities**

ca. 2 % (w/w) Water / Water / Water / 7732-18-5

Reference substance: Water / Water / Water / 7732-18-5

EC number	EC name
231-791-2	water
CAS number	CAS name
7732-18-5	Water
IUPAC name	
Water	

Typical concentration: ca. 2 % (w/w)

Concentration range: > 0.5 < 3 % (w/w)

Remarks:

---

< 0.5 % (w/w) Unknown impurities / Unknown impurities

Reference substance: Unknown impurities / Unknown impurities

EC number	EC name
CAS number	CAS name
IUPAC name	
Unknown impurities	

Typical concentration: < 0.5 % (w/w)

Concentration range: > 0 < 1 % (w/w)

Remarks: 2 unknown impurities



# 1.2 Composition

**Additives**

50 ppm Buffer agent / 144-55-8

Only additives in the sense of REACH should be indicated here, i.e., substances with stabilising function

Function

Typical concentration  ppm

Concentration range   ppm

Remarks

# 1.4 Analytical information

**Analytical information**

Analytical methods and spectral data

The following analytical methods were used to characterize the substance:

- UV/Vis, IR, <sup>1</sup>H-NMR and <sup>13</sup>C-NMR Spectra
- HPLC
- Titration according to Karl Fischer (determination of water content)
- AAS (determination of Sodium content)

Methods of detection and analysis.pdf / 114.24 KB

Optical activity

**Results of analysis**

Ultraviolet /Visible spectrum

Analysis type: Ultraviolet /Visible spectrum

Tested substance: 1H-pyrazole, sodium salt

Method used: See attachment

UV spectrum.pdf / 114.24 KB

Remarks

***Provide any relevant comment here, or a justification as to why one of the analysis could not be performed if relevant***

# IUCLID 5 Dossiers

Dossier preparation for multi-  
constituent substances

# 1.1 Identity

**Substance identification**

Chemical name

Legal entity flags

Legal entity

Third party flags

Third party

**Role in the supply chain**

Role flags

Role:  Manufacturer  Importer  Only representative  Downstream user

**Reference substance**

**Type of substance**

Composition

Origin

**Trade names**

Trade names flag

Name
------

# 1.1 Identity (reference substance)

Reference substance: Lactic acid

---

**General information**

Reference substance name

---

**EC inventory**

EC number  CAS number

EC name

Molecular formula

Description

---

**No EC information available**

Justification

---

**Reference substance information**

---

**CAS information**

CAS number

CAS name

---

**IUPAC name**

# 1.1 Identity (reference substance)

Molecular and structural information

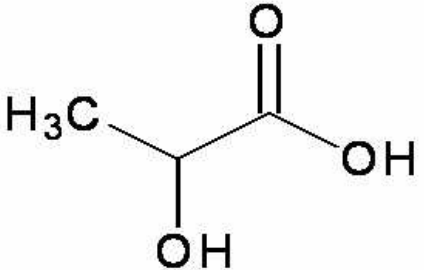
Molecular formula

Molecular weight range

SMILES notation

InChI

Structural formula



Remarks

# 1.1 Identity (reference substance)

Molecular and structural information

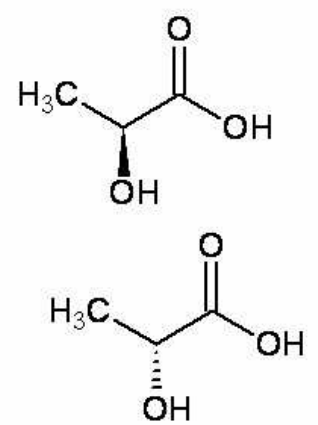
Molecular formula

Molecular weight range

SMILES notation

InChI

Structural formula



Remarks

# 1.1 Identity (reference substance)

Molecular and structural information

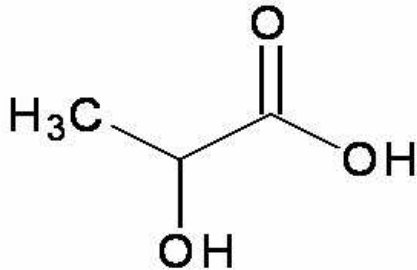
Molecular formula C<sub>3</sub>H<sub>6</sub>O<sub>3</sub>

Molecular weight range > 90 < 90.1

SMILES notation O=C(O)C(C)O

InChI

Structural formula



Remarks Generic structural formula for the 2 isomers

Load... Zoom... Delete



# 1.2 Composition







Lactic acid

Name




Brief description

Degree of purity




Constituents

66 % (w/w) S-Lactic acid / (2S)-2-hydroxypropanoic acid	  
33 % (w/w) R-lactic acid / (2R)-2-hydroxypropanoic acid	  

Impurities

< 1 % (w/w) Unknown impurities / Unknown impurities	  
---	---

Additives

	  
--	---

# 1.2 Composition

Constituents

66 % (w/w) S-Lactic acid / (2S)-2-hydroxypropanoic acid

Reference substance: S-Lactic acid / (2S)-2-hydroxypropanoic acid

EC number	EC name
<input type="text"/>	<input type="text"/>
CAS number	CAS name
<input type="text"/>	<input type="text"/>
IUPAC name	
(2S)-2-hydroxypropanoic acid	

Typical concentration: 66 % (w/w)

Concentration range: 60 - 68 % (w/w)

Remarks:

---

33 % (w/w) R-lactic acid / (2R)-2-hydroxypropanoic acid

Reference substance: R-lactic acid / (2R)-2-hydroxypropanoic acid

EC number	EC name
<input type="text"/>	<input type="text"/>
CAS number	CAS name
<input type="text"/>	<input type="text"/>
IUPAC name	
(2R)-2-hydroxypropanoic acid	

Typical concentration: 33 % (w/w)

Concentration range: 30 - 34 % (w/w)

Remarks:

# 1.2 Composition

Reference substance: Lactic acid

**General information**

Reference substance name

**EC inventory**

EC number  CAS number

EC name

Molecular formula

Description

**No EC information available**

Justification


**Reference substance information**

**CAS information**

CAS number

CAS name

**IUPAC name**



# 1.2 Composition

Reference substance: R-lactic acid / (2R)-2-hydroxypropanoic acid / 2-hydroxypropanoic acid, (2R) / 10326-41-7

---

**General information**

Reference substance name

---

**EC inventory**

EC number  CAS number

EC name

Molecular formula

Description

**Only include information for that specific constituent**

---

**No EC information available**

Justification

---

**Reference substance information**

---

**CAS information** Data protection and regulatory purposes flags

CAS number

CAS name

---

**IUPAC name**

# 1.2 Composition

**Molecular and structural information**

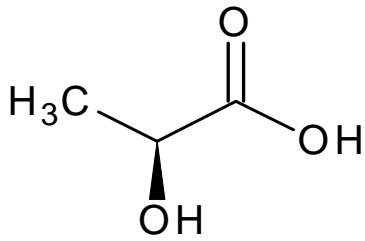
Molecular formula

Molecular weight range

SMILES notation

InChI

Structural formula



Load... Zoom... Delete

Remarks